

MAPA DA AULA

CURSO

Fluência em Dados para Iniciantes

Aula 00 - Conceituação de Dados

PROFESSOR

Thiago Ultra



RESUMOS

21



QUESTÕES

101



BANCAS

16



CARACTERES

176,3 MIL

ESTRUTURA

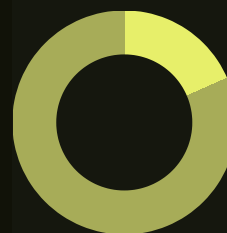
Seções	29	Destaques	28
Tabelas	7	Diagramas	6
Imagens	2	Code blocks	2
Exemplos	14		

QUESTÕES POR BANCA

Cebraspe	54
Inédita	14
FGV	6
Quadrix	6
Outras bancas	21

VOLUME

TEORIA X QUESTÕES



Teoria	18%
Questões	82%

TEORIA

4.917

palavras

QUESTÕES

21.882

palavras

Total de Palavras 26.799

Esquemas do resumo ficam fora da contagem de teoria; questões não numeradas e a lista final entram no volume de questões.

LINGUAGENS DE CÓDIGO

SQL 1

XML 1

Sumário

1 Bem-vindos!	2
1.1 Conteúdo das aulas	3
2 Fundamentos de Dados	4
2.1 Conceituação	4
2.1.1 Dado	5
2.1.2 Informação	6
2.1.3 Conhecimento	8
2.1.4 Inteligência	10
2.1.5 Pra sintetizar	11
3 Dados estruturados, semiestruturados e não estruturados	31
3.1 Dados estruturados	31
3.2 Dados semiestruturados	34
3.3 Dados não estruturados	37
3.4 Para resumir	39
3.5 Questões para exercitar	39
4 Observações, atributos e métricas	63
4.1 Observações e granularidade	63
4.2 Atributos e Variáveis	64
4.2.1 Atributos qualitativos	65
4.2.2 Atributos quantitativos	67
4.3 Dimensões, medidas e métricas	69
4.3.1 Dimensões	69
4.3.2 Medidas e Métricas	70
4.4 Questões para exercitar	70
5 Ciclo de vida dos dados e da informação	77
5.1 Ciclo de Vida dos Dados (CVD)	77
5.2 Ciclo de Vida da Informação (CVI)	78
5.3 Comparativo entre CVD x CVI	79
5.4 Questões para exercitar	79
6 Questões da aula	83
6.1 Gabarito	108
7 Resumos	110

1 Bem-vindos!

Olá Pessoal!

Meu nome é Thiago Ultra. Sou Auditor Fiscal da Receita Federal, formado em Ciência da Computação pela Universidade de Londres, e professor de TI para concursos públicos.

Tenho um história longa nos concursos públicos. Minha primeira aprovação veio em 2001 (Caixa Econômica Federal). Mais tarde fui aprovado no BNDES (2. lugar para Técnico Administrativo) e TRF 2 (Técnico Judiciário), ambos em 2008. E finalmente Auditor Fiscal da RFB em 2009.

De 2011 a 2013 fui professor de Contabilidade no Estratégia Concursos, mas parei de dar aulas por dois motivos: filhos pequenos e falta de prazer em ensinar Contabilidade.

Agora, os filhos cresceram. E depois de cursar Ciência da Computação, voltei a dar aulas, mas sobre um assunto que me apaixona.

O Curso de Fluência em Dados ainda não está completo. Estou escrevendo as aulas e comentando questões aos poucos, mas nos próximos meses ele estará pronto para publicação.

Por ora, vocês podem conhecer meu método com essa aula demonstrativa. Se quiserem, podem me acompanhar no [Instagram](#) ou se inscrever na minha **lista de emails** para o lançamento do curso, em <https://www.profthiagoultra.com.br>.

Bons estudos!



Questões com links para o TEC Concursos!

A maioria das questões deste curso possui link direto para a página da questão do TEC Concursos.

Então, durante os estudos, você pode seguir o pdf, **clicar no cabeçalho da questão**, abrir a página da questão no TEC Concursos, resolver a questão lá e retornar para ver os comentários no pdf.

Resolver a questão no TEC é opcional, claro. E este curso não possui qualquer relação com a plataforma de questões. É só uma comodidade para quem é assinante de lá também.

1.1 Conteúdo das aulas

Aula	Conteúdo
Aula 00	Hierarquia DIKW. Dados estruturados, semiestruturados e não estruturados. Observações, Atributos e Métricas. Ciclo de Vida dos dados e da informação.
Aula 01	Analytics e Análise Exploratória
Aula 02	Ciência de Dados, CRISP-DM, Data Mining
Aula 03	Aprendizado de Máquina - Conceitos
Aula 04	Aprendizado de Máquina - Métricas de Avaliação
Aula 05	Aprendizado de Máquina - Algoritmos
Aula 06	Aprendizado Profundo
Aula 07	Transformação de Dados.
Aula 08	IA Generativa, Processamento de Linguagem Natural e LLMs
Aula 09	Big Data
Aula 10	...
Aula 11	...
Aula 12	...

2 Fundamentos de Dados

Dados são um tema clássico de Sistemas de Informação, mas também são algo mais atual do que nunca. Grandes instituições públicas e privadas vivem inundadas por dados. Cadastros de contribuintes, notas fiscais, declarações, lançamentos, transações, medições, documentos eletrônicos, *logs* de sistemas: tudo isso é produzido em volume contínuo e crescente.

A capacidade de gerar registros, por si só, não gera compreensão, controle nem decisão de qualidade. Acumular registros não é o mesmo que entender o que eles significam.

É justamente por isso que fundamentos de dados precisam ser o ponto de partida dos estudos. Antes de perguntar como os dados serão armazenados em tabelas, consultados por SQL ou explorados em processos analíticos, precisamos responder a uma pergunta mais elementar: **o que é, afinal, um dado?** E de que modo o dado apoia a decisão?

Em prova, é muito comum a banca explorar a confusão entre dado e informação, ou embaralhar os diferentes tipos de dados para testar se o candidato realmente domina a base conceitual do tema. Por isso, começaremos pelo ponto mais fundamental: a conceituação de dado e a progressão conceitual pela qual ele ganha contexto, significado e utilidade, sumarizada na Hierarquia DIKW.

2.1 Conceituação

O uso de dados tem, entre seus propósitos mais importantes, o de orientar decisões. Mas entre o dado bruto e a tomada de decisão costuma-se representar uma progressão conceitual pela qual o **dado** ganha contexto, significado e utilidade. Esta progressão é representada pelos conceitos **dado, informação, conhecimento e inteligência** (ou **sabedoria** na formulação clássica).

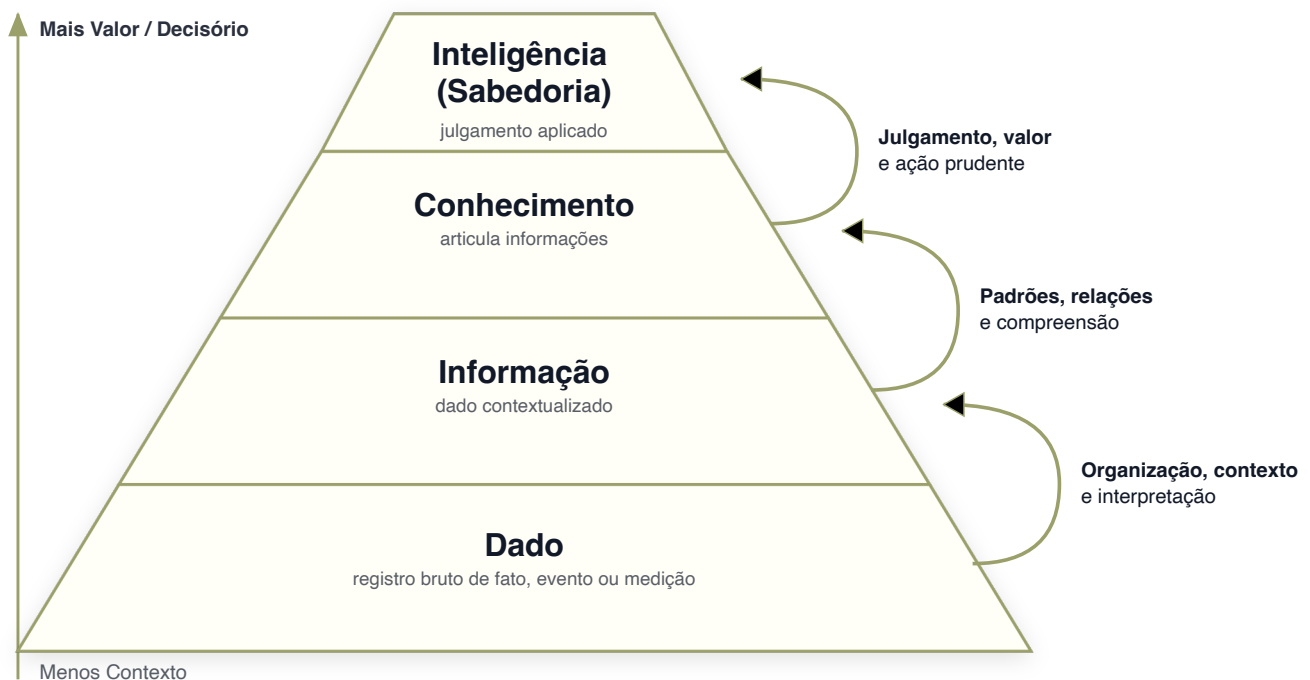
O **dado**, por si só, é apenas um **registro bruto de um fato, evento ou elemento observado**. Em outras palavras, o dado é produto da observação, mas ainda não possui valor decisório imediato. Para **ganhar valor**, ele precisa ser organizado, contextualizado e interpretado.

Quando isso ocorre, o dado se converte em **informação**. Quando diferentes informações passam a ser relacionadas, compreendidas e incorporadas a um raciocínio, chegamos ao plano do **conhecimento**.

Essa progressão se completa com a **inteligência (ou sabedoria)**, entendida como a capacidade de julgar e agir de maneira adequada com base no conhecimento acumulado.

Toda essa conceituação é consolidada na **Hierarquia DIKW (Data, Information, Knowledge, Wisdom)**, também conhecida com **pirâmide informacional**:

Hierarquia DIKW (Pirâmide informacional)



2.1.1 Dado

Dado é o registro bruto, o símbolo que representa um objeto, um evento, uma medição do ambiente. É produto da observação, costuma ser obtido por meio de máquinas e sensores, além de ser transferível e frequentemente quantificável. **Dado** representa a **matéria-prima da informação**.



Pra exemplificar...

Quando um termômetro captura a temperatura da pele, ou do ambiente, ele gera um dado: a temperatura medida.

- 39 °C

Quando um enfermeiro, ou um meteorologista, interpretam o dado, eles produzem informação: "criança tá com febre" ou "calor intenso". Vejam que o contexto é fundamental para a informação produzida.

O importante é que o dado bruto, isolado de contexto, tem utilidade decisória limitada até que seja interpretado: no exemplo acima, a temperatura, por si só, não significa muita coisa. O que agrega utilidade é a interpretação do dado: no caso da criança, é febre; no clima, é calor intenso. A interpretação do dado produz informação.

Dado

- Registro bruto
- Elemento ainda não interpretado, não organizado, nem contextualizado no contexto decisório
- Matéria-prima da informação
- Dados podem ser obtidos por sensores, sistemas, máquinas ou registros humanos.

Quadrix - 2025 - Ana Sistemas - CFBIO



No que diz respeito às ferramentas de automação de processos, aos conceitos de dados e aos índices, julgue o item seguinte.

Dados podem ser definidos como elementos brutos que, quando contextualizados, se tornam informação.

Comentários:

Correto, é a definição clássica de dado e de informação.

Gabarito: Certa

2.1.2 Informação

Informação é o **dado contextualizado**. É o dado que foi **organizado, interpretado ou associado a significado**, passando a ser útil para algum fim.

A informação costuma aparecer como resposta a perguntas do tipo *quem, o que, onde, quando e quanto*. Isso ajuda a perceber que a informação já não é mero registro bruto: ela **incorpora contexto e permite descrever uma situação de modo mais inteligível**.

O dado, isoladamente, apenas registra; a informação, por sua vez, já permite compreender algo sobre a realidade observada. No exemplo da medição de temperatura (mais acima), a informação foi gerada a partir "do quê" foi medido: a temperatura da criança, ou a temperatura do ambiente. A informação, portanto, pressupõe contexto, organização e interpretação, e não está contida no dado bruto de forma imediata.

Logo, temos que a **dados geram informação quanto contextualizados e organizados**, não o contrário.

Informação

- É o dado organizado, interpretado ou associado a significado
- Dado que passou a ser útil para algum fim
- Dados geram informação quando organizados e contextualizados, não o contrário

Cebraspe - 2025 - APC / Infra e Cibersegurança - FUNPRES-EXE



Julgue o seguinte item, relativo a banco de dados.

Considere que, em uma análise sobre a sustentabilidade de um sistema de previdência social, sejam cadastrados dados acerca da quantidade de contribuições mensais de servidores, a idade de aposentadoria e o valor das aposentadorias concedidas. Considere, ainda, que esses dados, por si só, sejam números isolados e sem significado direto. Considere, por fim, que quando esses dados são analisados e comparados com projeções de crescimento populacional e expectativa de vida, eles forneçam informações úteis para avaliar a viabilidade do sistema. Com base nessa situação hipotética, é correto afirmar que dados são fatos brutos e sem contexto, enquanto informação é o resultado da análise desses dados, tornando-os úteis para a tomada de decisões.

Comentários:

Questão corretíssima. Dado é registro/fato bruto sem contexto. Informação é o dado organizado, interpretado ou associado a significado.

No cenário descrito pelo enunciado, vejam que os dados foram *analisados e comparados com projeções*. Essa **organização em gráficos e projeções acrescenta contexto** e permite extrair sentido do conjunto. Por isso, neste caso, as projeções e gráficos devem ser compreendidos como informação.

Gabarito: Certa

Cebraspe - 2025 - Tec / Tecnologia da Informação - FUB



Julgue o item a seguir, relativo aos conceitos de administração de dados.

Informação é uma unidade básica de dado que pode ser medida, registrada ou armazenada sem que esteja associada a um significado ou conhecimento, ao passo que dado é uma informação dotada de significado que pode ser usada para a geração de conhecimento ou a tomada de decisões.

Comentários:

A questão inverte as definições de dado e de informação.

- **Dado** é a unidade básica, elemento bruto sem contexto/organização suficiente.
- A **informação** é dotada de significado que pode ser usada para a geração de conhecimento.

Alternativa errada.

Gabarito: Errada

2.1.3 Conhecimento

O **conhecimento** resulta da compreensão construída a partir das informações obtidas dos dados. Ele surge quando essas **informações** deixam de ser apenas descrições isoladas e passam a ser articuladas entre si, permitindo **reconhecer padrões, relações e implicações relevantes**. Isto é, o conhecimento não se limita ao acúmulo de informações: ele envolve **compreensão**.

O conhecimento costuma ser adquirido:

1. pela **transmissão**, quando alguém nos ensina o que sabe, ou
2. pela **experiência**, quando ele é adquirido a partir da prática e da vivência.

Em ambos os casos, o processo de aquisição de conhecimento recebe o nome de **aprendizado**.

Isso ajuda a perceber que o conhecimento já ocupa um patamar distinto em relação ao dado e à informação. O dado registra. A informação descreve e contextualiza. O **conhecimento**, por sua vez, **permite compreender, explicar e orientar a ação**. É justamente por isso que, no contexto das organizações, o conhecimento costuma ser tratado como um ativo de alto valor: ele não apenas informa sobre a realidade, mas **permite agir sobre ela com maior compreensão**.



Pra exemplificar...

Se o dado é 39 °C, e a informação extraída a partir dele *a criança está com febre*, o conhecimento aparece em um nível seguinte: *febre alta em criança porque pode indicar um quadro infeccioso ou inflamatório que demanda acompanhamento e possível intervenção*.

Veja que o conhecimento geralmente implica no reconhecimento de padrões: *febre alta costuma ser quadro infeccioso ou inflamatório*. E também orienta a tomada decisão.

No campo da gestão do conhecimento, também é comum distinguir o conhecimento **explícito** do conhecimento **tácito**:

- **explícito** é aquele que pode ser formalizado e registrado em documentos, normas, relatórios, manuais ou procedimentos.

- **tácito** é pessoal, advém da experiência e mais difícil de verbalizar, estando muitas vezes incorporado à prática de indivíduos e equipes. Em muitas abordagens de gestão do conhecimento, o conhecimento tácito representa uma parcela relevante do conhecimento organizacional.

Essa distinção é relevante porque mostra que nem todo conhecimento está imediatamente disponível em bancos de dados ou documentos: **uma parte importante dele permanece embutida na experiência (na cabeça) humana, sendo difícil de transferir.**

Conhecimento

- Construído a partir das informações obtidas dos dados
- Implica o reconhecimento de padrões, relações e implicações relevantes
- Permite compreender, explicar e orientar a ação (é acionável)
- Pode ser tácito (obtido pela experiência) ou explícito
- Pode ser difícil de estruturar e de transferir, especialmente quando é tácito

FGV - 2024 - Aud CE / Ciência de Dados - TCE PA



Com relação aos conceitos de dado, informação e conhecimento, avalie as afirmativas a seguir.

- I. Os dados são itens elementares, são cadeias de símbolos e não possuem significado.
- II. São exemplos de conhecimento: tendência de vendas de um produto A em uma região B; relação entre o aumento ou queda do preço de uma ação X, na bolsa de valores, e a variação do câmbio.
- III. As informações correspondem ao dado processado, com significado e um contexto indefinido.

Está correto o que se afirma em

- A) I, apenas.
- B) I e III, apenas.
- C) I e II, apenas.
- D) II, apenas.
- E) I, II e III.

Comentários:

I - Correta. Dados são elementos elementares, representam símbolos sem significado contextual.

II - Correta. O conhecimento implica reconhecimento de padrões, relações e implicações relevantes. A tendência de vendas de um Produto A em uma região B pode ser interpretado como um padrão; a relação entre o aumento ou queda do preço de uma ação X, na bolsa, e a variação do câmbio é um exemplo de uma "relação relevante".

III - Incorreta. O erro está na expressão 'contexto indefinido', pois a informação é resultado de *contextualização* do dado.

Logo, apenas I e II são corretas.

Gabarito: C

2.1.4 Inteligência

A **inteligência** (sabedoria, na formulação clássica da hierarquia **DIKW**) ocupa o nível mais alto da progressão conceitual, e orienta a **tomada de decisões concretas e mais eficazes**. Ela representa a **aplicação qualificada do conhecimento para decisão, inferência ou ação**.

Por isso, **inteligência** não se confunde com mero acúmulo de conhecimento: envolve **juízo, discernimento e capacidade de agir de maneira adequada diante de uma situação real**. Em geral, é a sabedoria que explica o fato de duas pessoas distintas atuarem de modo diferente para resolver um mesmo problema, vez que a sabedoria/inteligência implica **juízo**, e o juízo é marcadamente pessoal.



Pra exemplificar...

- **Dado:** 39 °C
- **Informação:** a criança está com febre (temperatura medida na pele)
- **Conhecimento:** febre alta em criança pode indicar quadro que exige avaliação cuidadosa
- **Sabedoria/Inteligência:** diante do contexto concreto, decidir monitorar, buscar atendimento ou seguir orientação profissional

Atualmente, em alguns contextos de computação e inteligência artificial, a camada superior da hierarquia é aproximada da ideia de **inteligência**, isto é, da aplicação do conhecimento à inferência, à decisão e à ação.

Inteligência (sabedoria)

- Orienta a tomada de decisões concretas e eficazes
- Representa a **aplicação qualificada do conhecimento** para decisão, inferência ou ação
- Envolve **juízo, discernimento e capacidade de agir de maneira adequada** diante de uma situação real
- Na formulação clássica, está associada ao julgamento humano

IBFC - 2024 - Téc Jud. / Apoio Esp. Desenv. Sistemas - TRF 5



Leia a frase abaixo.

O ____ é a assimilação e compreensão de informações adquiridas, enquanto a ____ envolve a capacidade de aplicar esse ____ de maneira adaptativa, analítica e inovadora para resolver problemas e tomar decisões eficazes em diferentes contextos.

Assinale a alternativa que preencha correta e respectivamente as lacunas.

- A) inteligência / conhecimento / artefato
- B) conhecimento / inteligência / conhecimento
- C) dado / informação / atributo
- D) conjunto / informação / artefato

Comentários:

De acordo com a Hierarquia DIKW, a assimilação e compreensão de **informações** adquiridas corresponde ao **conhecimento**. E a capacidade de aplicar esse conhecimento de maneira adaptativa representa justamente a **inteligência**.

Logo, o preenchimento das lacunas segue a ordem conhecimento / inteligência / conhecimento.

Gabarito: B

2.1.5 Pra sintetizar

Para prova, a síntese mais segura dos conceitos deste tópico é:



Decorar!

- **Dado:** registro bruto
- **Informação:** dado interpretado e contextualizado
- **Conhecimento:** compreensão (padrões, implicações) obtida a partir das informações
- **Inteligência/Sabedoria:** uso do conhecimento, com julgamento, para agir e decidir.

1. (Cebraspe - 2025 - Tec / Tecnologia da Informação - FUB)



Julgue o item a seguir, relativo aos conceitos de administração de dados.

Informação é uma unidade básica de dado que pode ser medida, registrada ou armazenada sem que esteja associada a um significado ou conhecimento, ao passo que dado é uma informação dotada de significado que pode ser usada para a geração de conhecimento ou a tomada de decisões.

Comentários:

A questão inverte as definições de dado e de informação.

- **Dado** é a unidade básica, elemento bruto sem contexto/organização suficiente.
- A **informação** é dotada de significado que pode ser usada para a geração de conhecimento.

Alternativa errada.

Gabarito: Errada

2. (Cebraspe - 2025 - APC / Infra e Cibersegurança - FUNPRESP-EXE)



Julgue o seguinte item, relativo a banco de dados.

Considere que, em uma análise sobre a sustentabilidade de um sistema de previdência social, sejam cadastrados dados acerca da quantidade de contribuições mensais de servidores, a idade de aposentadoria e o valor das aposentadorias concedidas. Considere, ainda, que esses dados, por si só, sejam números isolados e sem significado direto. Considere, por fim, que quando esses dados são analisados e comparados com projeções de crescimento populacional e expectativa de vida, eles forneçam informações úteis para avaliar a viabilidade do sistema. Com base nessa situação hipotética, é correto afirmar que dados são fatos brutos e sem contexto, enquanto informação é o resultado da análise desses dados, tornando-os úteis para a tomada de decisões.

Comentários:

Questão corretíssima. Dado é registro/fato bruto sem contexto. Informação é o dado organizado, interpretado ou associado a significado.

No cenário descrito pelo enunciado, vejam que os dados foram *analisados e comparados com projeções*. Essa **organização em gráficos e projeções acrescenta contexto** e permite extrair sentido do conjunto. Por isso, neste caso, as projeções e gráficos devem ser compreendidos como informação.

Gabarito: Certa

3. (Cebraspe - 2024 - ACE / Esp. Arquivologia - TC DF)



Julgue o próximo item, relativos a dado, informação, conhecimento e inteligência.

Informação é a capacidade de compreender, interpretar e aplicar dados de forma estratégica.

Comentários:

A definição da assertiva não é a de **informação**. A **informação** é obtida a partir de dados que foram contextualizados, interpretados e organizado. A capacidade de compreender, interpretar e aplicar dados de forma estratégica tende a se aplicar mais do conceito de **conhecimento**.

Gabarito: Errada

4. (Cebraspe - 2023 - Ana TI - FUB)



A respeito de dados e informações, julgue o item que se segue.

Quando ocorre o tratamento de um ou mais dados, é gerada a informação, a qual constitui conhecimento quando aplicada em determinado contexto.

Comentários:

Dado bruto tratado e organizado gera informação. O conhecimento da compreensão de padrões, implicações e relações identificados nas informações, especialmente quando elas são aplicadas num determinado contexto. Por isso, a afirmativa correta.

Gabarito: Certa

5. (Cebraspe - 2022 - ADP / Redes e Comunicação de Dados - DPE RO)



Em uma situação hipotética, o fato de pessoas atuarem de forma diferenciada para resolver um mesmo problema é explicado pelo conceito de

- A) informação.
- B) comportamento.
- C) dado.
- D) conhecimento.
- E) inteligência.

Comentários:

O que explica o fato de pessoas atuarem de forma diferenciada para resolver um mesmo problema é a **inteligência**, pois que a inteligência implica julgamento e discernimento, atributos marcadamente pessoais.

Gabarito: E

6. (Cebraspe - 2022 - Ag Inv - PC PB)



Na teoria da informação, os insumos para a tomada de decisão na organização são apresentados na fase de

- A) geração da informação.
- B) tomada de decisão.
- C) provisão de conhecimento.
- D) acompanhamento de resultados.
- E) aquisição de dados.

Comentários:

O que habilita a tomada de decisões é o **conhecimento**, ou, no caso do enunciado, a **provisão do conhecimento**. Podemos entender a provisão do conhecimento, neste sentido, como a compreensão das informações geradas e partir de dados brutos que foram interpretado e contextualizados. O conhecimento **permite orientar a ação e a tomada de decisão**. Assim, a alternativa correta é a letra C.

Gabarito: C

7. (Cebraspe - 2022 - PPNS / Ana. Sistemas / Eng. de Software - PETROBRAS)



Julgue o item a seguir, quanto a conceitos de dado, informação, inteligência e business intelligence (BI).

A inteligência é a informação que é expressa na forma de números, letras ou símbolos, os quais são os insumos dos gráficos.

Comentários:

Incorreta. O que é expresso na forma de número, letras ou símbolos é o **dado** bruto. A **informação** representa o dado enriquecido, contextualizado, organizado. A **inteligência**, em última instância, representa a aplicação de conhecimento ao caso de concreto de maneira eficaz, implicando julgamento e discernimento sobre as ações a serem tomadas.

Gabarito: Errada

8. (Cebraspe - 2022 - Aud Est / Tecnologia da Informação - SECONT ES)



Com relação a dado, informação e conhecimento, julgue o item subsecutivo.

O conhecimento é obtido a partir de um conjunto de informações inseridas em um mesmo e específico contexto.

Comentários:

A afirmativa está incorreta, pois o conhecimento não depende, necessariamente, de informações situadas em um mesmo e específico contexto.

Em termos conceituais, o conhecimento resulta da interpretação e da articulação de informações, com a identificação de padrões, relações e implicações relevantes, inclusive quando essas informações decorrem de **contextos distintos**, como relacionar dados de vendas, clima e sazonalidade para inferir padrão de demanda.

Gabarito: Errada

9. (Cebraspe - 2022 - Aud Est / TI - SECONT ES)



Com relação a dado, informação e conhecimento, julgue o item subsecutivo.

Os termos CÉU e AZUL, quando utilizados separadamente, representam dados, enquanto a expressão CÉU É AZUL representa uma informação.

Comentários:

Correta. CÉU e AZUL, isoladamente, não estão agrupados nem contextualizados adequadamente. Agora, em CÉU É AZUL, a simples adição do verbo de ligação "é" já representa, por si, só uma contextualização do dado, gerando a informação.

Gabarito: Certa

10. (Cebraspe - 2022 - ACE / TI - TCE RJ)



Julgue o item a seguir, referente à informação e a dados, estruturados e não estruturados.

A informação normalmente envolve a participação humana na sua estruturação e na busca do seu significado, enquanto os dados são mais facilmente obtidos e quantificados por máquinas.

Comentários:

A informação é o dado interpretado e contextualizado, e isso normalmente implica avaliação humana para agregação de contexto e interpretação. Já os dados são mais facilmente obtidos e quantificados por máquinas (sistemas, sensores, e demais aparelhos eletrônicos). Afirmativa correta.

Gabarito: Certa

11. (Cebraspe - 2022 - AFCE / Tecnologia da Informação - TCE SC)



Em relação a conceitos de informação e conhecimento e de banco de dados relacionais, julgue o item subsequente.

Informação pode ser definida como um conjunto de fatos diferentes relativos a uma transação e que é, geralmente, desprovida de significado; ela é considerada a matéria-prima dos sistemas de informação.

Comentários:

A definição expressa no enunciado se refere ao **dado**. O dado é registro bruto sem utilidade imediata para a tomada de decisão, sendo considerado a matéria-prima dos sistemas de informação. A **informação**, por sua vez, é extraída do dado interpretado e contextualizado.

Gabarito: Errada

12. (Cebraspe - 2022 - AFCE / Tecnologia da Informação - TCE SC)



Em relação a conceitos de informação e conhecimento e de banco de dados relacionais, julgue o item subsequente.

O conhecimento serve para dar suporte ao processamento decisório; ele representa a informação tratada, confiável e íntegra.

Comentários:

Afirmativa correta, pois o conhecimento dá, de fato, suporte ao processo decisório. Ademais, no contexto da afirmativa, a expressão "informação tratada, confiável e íntegra" deve ser entendida em sentido amplo, como informação já compreendida e validada para apoiar decisões. Logo, afirmativa correta.

Gabarito: Certa

13. (Cebraspe - 2021 - Ag Pol - PC DF)



Sobre a teoria da informação, julgue o item a seguir.

Embora não haja conceitos únicos para dados e informação, é possível afirmar que os dados são matéria-prima para a informação.

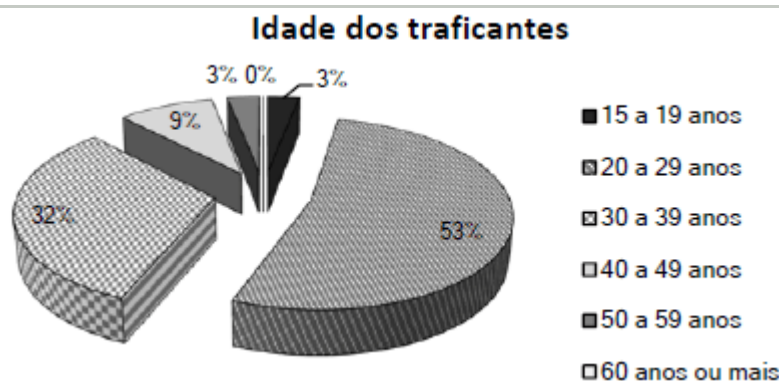
Comentários:

Correta. Dados são a **matéria-prima** para a informação. A informação é o que resulta do dado contextualizado, interpretado e organizado.

Gabarito: Certa

14. (Cebraspe - 2021 - APF - PF)





Considere que a Polícia Federal tenha registrado, em determinado período, a prisão de 1.789 traficantes de drogas pertencentes a facções criminosas, conforme faixas etárias mostradas no gráfico.

Com referência às informações e ao gráfico precedentes, julgue o item subsequente.

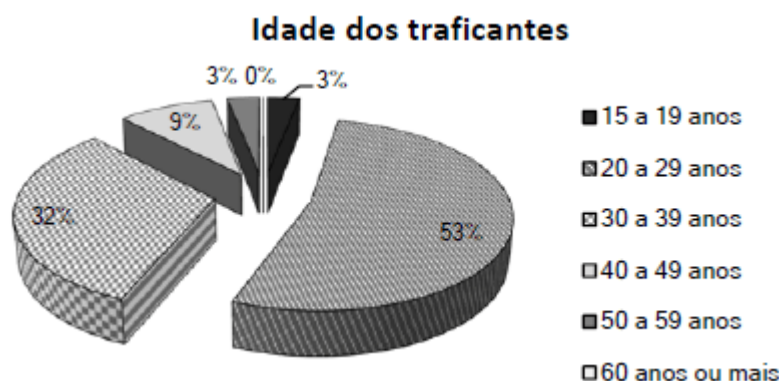
Considerando-se a classificação dados, informação, conhecimento e inteligência, é correto afirmar que o gráfico representa, por si só, a inteligência.

Comentários:

Incorreta. O gráfico, por si só, representa uma organização visual da informação (dados organizados com contexto). A inteligência, por sua vez, envolve **julgamento, discernimento e capacidade de agir de maneira adequada diante de uma situação real**.

Gabarito: Errada

15. (Cebraspe - 2021 - APF - PF)



Considere que a Polícia Federal tenha registrado, em determinado período, a prisão de 1.789 traficantes de drogas pertencentes a facções criminosas, conforme faixas etárias mostradas no gráfico.

Com referência às informações e ao gráfico precedentes, julgue o item subsequente.

O número 1.789 sozinho caracteriza uma informação, independentemente do contexto.

Comentários:

Incorreta. O número 1.789 é meramente um **dado bruto**, sem contexto nenhum. Para se tornar informação ele precisaria ser contextualizado e organizado.

Gabarito: Errada

16. (Cebraspe - 2021 - ACE / Controle Externo / Ciências Contábeis - TCE RJ)



Com relação a dado, informação, conhecimento e inteligência, julgue o item que se segue.

Um dado que está inserido em um contexto pode ser denominado de informação.

Comentários:

Certa. A informação é o dado contextualizado. Logo, dado inserido num contexto pode ser considerado uma informação.

Gabarito: Certa

17. (Cebraspe - 2021 - ACE / Controle Externo / Controle Externo - TCE RJ)



Com relação aos conceitos de análise de dados e informações, julgue o item a seguir.

A informação pode ser facilmente obtida por meio de máquinas, além de ser transferível e frequentemente quantificável.

Comentários:

Incorreta. O que pode ser facilmente obtido por meio de máquinas, além de ser transferível e frequentemente quantificável é o **dado**. A informação pressupõe interpretação e contextualização de dados brutos, que normalmente requer ação humana, mas também pode ser extraída a partir de sistemas.

Gabarito: Errada

18. (Cebraspe - 2021 - ACE / Controle Externo / TI - TCE RJ)



Com relação aos conceitos de análise de dados e informações, julgue o item a seguir.

O conceito de inteligência está relacionado à capacidade de julgamento, ou seja, o ser humano usa o seu conhecimento para tomar a melhor decisão em uma situação real.

Comentários:

Corretíssima, é exatamente essa a definição da inteligência na pirâmide informacional.

Gabarito: Certa

19. (Cebraspe - 2018 - APF - PF)



Julgue o item, a respeito da teoria da informação e de sistemas de informação.

O conceito de conhecimento é mais complexo que o de informação, pois conhecimento pressupõe um processo de compreensão e internalização das informações recebidas, possivelmente combinando-as.

Comentários:

Corretíssima. Afirmativa irretocável.

Gabarito: Certa

20. (Cebraspe - 2018 - EPF - PF)



Julgue o item a seguir, a respeito da teoria da informação e de metadados de arquivos.

O conhecimento é embasado na inteligência das informações que são coletadas e analisadas para uma organização.

Comentários:

Incorreta. Uma versão correta da afirmativa seria:

- O conhecimento é embasado na **compreensão** das informações **geradas a partir de dados** que são coletados e analisados para uma organização.

O problema principal da afirmativa é afirmar que o conhecimento é *embasado* inteligência, quando o correto seria afirmar que a *inteligência representa uma aplicação qualificada do conhecimento*.

Gabarito: Errada

21. (Cebraspe - 2018 - PCF / Área 14 - PF)



Julgue o item, a respeito de computação na nuvem, sistemas de informações e teoria da informação.

A informação se caracteriza pela compreensão e internalização do conteúdo recebido, por meio do seu uso em nossas ações; o dado, por sua vez, é um elemento bruto dotado apenas de significado e relevância que visem fornecer uma solução para determinada situação de decisão.

Comentários:

Afirmativa incorreta. Em partes:

- O **conhecimento** se caracteriza pela compreensão e internalização do conteúdo recebido, por meio do seu uso em nossas ações (conhecimento é compreensão e experiência).
- O **dado**, por sua vez, é um elemento bruto **ainda sem contexto ou utilidade decisória imediata**.

Gabarito: Errada

22. (Cebraspe - 2018 - PPF - PF)



Acerca da definição de dados e informação e sua representação em sistemas de informação, julgue o item que se segue.

Dados são fatos que descrevem os objetos de informação, por exemplo, eventos e entidades.

Comentários:

Correto. Dados apenas descrevem objetos, eventos, medições. São brutos, coletáveis por máquinas, transferíveis e compõem a matéria-prima da informação.

Gabarito: Certa

23. (Cebraspe - 2018 - PPF - PF)



Acerca da definição de dados e informação e sua representação em sistemas de informação, julgue o item que se segue.

Informação é constituída por um conjunto de dados com características específicas. O ponto de análise é que os dados devem ser irrelevantes para o sistema a que se destinam.

Comentários:

Incorreta. Dados são a **matéria-prima** dos sistemas de informação, logo, devem ser **relevantes** para o sistema a que se destinam.

Gabarito: Errada

24. (Cebraspe - 2018 - Aud Est / Controle Externo - TCM BA)



O diretor de uma montadora de veículos necessita tomar uma decisão acerca da continuidade ou não de um dos produtos vendidos no Brasil. Para tanto, solicitou um relatório sobre as vendas de carros da marca do último trimestre de 2018, por faixa de preço, região, modelo e cor. Nessa situação, no contexto de análise da informação, o relatório representa

- A) conhecimento.
- B) inteligência.
- C) dados.
- D) informação.
- E) sabedoria.

Comentários:

O relatório sobre as vendas de carros da marca por faixa de preço, região, modelo e cor representa **dado organizado e contextualizado**. Logo, por si só, o relatório representa **informação**. Letra D.

Gabarito: D

25. (Cebraspe - 2017 - ACE / Auditoria de Contas Públicas - TCE PE)



Acerca dos conceitos de dado, informação e conhecimento, julgue o item a seguir.

A informação caracteriza-se por ser frequentemente tácita, bem como por ser de estruturação e captura difíceis em máquinas.

Comentários:

Incorreta. O **conhecimento** caracteriza-se por ser frequentemente **tácito**, isto é, não explícito e embutido na experiência humana. Por isso mesmo, é difícil de estruturar e capturar em máquinas.

Mas não podemos esquecer que existe, também, conhecimento **explícito**, que pode ser verbalizado ou escrito e registrado em sistemas de informação.

Gabarito: Errada

26. (Cebraspe - 2016 - AFCE / Tecnologia da Informação - TCE SC)



Julgue o item a seguir, acerca de dado, informação, conhecimento e inteligência.

Define-se informação como significado, ou seja, como registros icônicos e simbólicos — fonéticos ou numéricos — e signos — linguísticos, lógicos ou matemáticos —, por meio dos quais se representam atos, conceitos ou instruções.

Comentários:

Incorreta. A informação traduz significado, mas *registros icônicos e simbólicos — fonéticos ou numéricos — e signos — linguísticos, lógicos ou matemáticos —, por meio dos quais se representam atos, conceitos ou instruções* define o **dado**.

Gabarito: Errada

27. (Cebraspe - 2016 - AFCE / Tecnologia da Informação - TCE SC)



Julgue o item a seguir, acerca de dado, informação, conhecimento e inteligência.

O atributo de inteligência depende mais da qualidade da informação disponível do que da sua quantidade, tendo, portanto, natureza qualitativa.

Comentários:

Correta. A inteligência envolve **julgamento, discernimento e capacidade de agir de maneira adequada** diante de uma situação real. Ela representa uma **aplicação do conhecimento a situações concretas**. Com isso, informações de maior qualidade tende a ser mais relevante do que simplesmente possuir *mais* informação, de modo que a inteligência possui, sim, natureza qualitativa.

Gabarito: Certa

28. (Cesgranrio - 2023 - Esc BB / Agente de Tecnologia - BB)



No processo de aprendizagem de máquina, busca-se identificar padrões compreensíveis em bases de dados.

Desse modo, supõe-se a existência de uma hierarquia de três níveis: a base, com os itens elementares, captados e armazenados por recursos de Tecnologia da Informação; o nível intermediário, no qual esses itens são processados, com significados e contextos bem definidos; e o nível mais alto, contendo os padrões ou os conjuntos cuja formulação pode envolver ou relacionar os níveis inferiores.

Os três níveis dessa hierarquia, listados, respectivamente, da base para o nível mais alto, são:

- A) conhecimento, informação e dado
- B) dado, conhecimento e informação
- C) dado, informação e conhecimento
- D) informação, conhecimento e dado
- E) informação, dado e conhecimento

Comentários:

Vejam os trechos da afirmativa:

- a base, com os itens elementares, captados e armazenados por recursos de Tecnologia da Informação: a base representa o **dado** bruto, capturado com recursos de TI (medições, logs, sistemas de registro, etc).
- o nível intermediário, no qual esses itens são processados, com significados e contextos bem definidos: a adição de contexto ao dado bruto produz **informação**.
- e o nível mais alto, contendo os padrões ou os conjuntos cuja formulação pode envolver ou relacionar os níveis inferiores: a compreensão de informações, como o reconhecimento de padrões, gera **conhecimento**.

Logo, a sequência correta é **dado, informação, conhecimento**.

Gabarito: C

29. (COGEPE FIOCRUZ - 2024 - TecnoSP / TI - FIOCRUZ)



Uma das mais importantes atividades da gestão conhecimento nas organizações é a distinção entre dado, informação e conhecimento. Sobre esta distinção é correto afirmar que:

- A) dado é um caso qualquer, desde que seja intangível; informação é a transformação do dado para possibilitar a sua utilização prática; e conhecimento é caracterizado pelo uso da informação para resolver problemas, sejam eles reais ou não.
- B) dado é um fato discreto e objetivo sobre um evento; informação é um dado dotado de relevância e propósito; e o conhecimento é uma mistura de experiências, valores, informações contextuais e insights.
- C) dado é a informação interpretada; conhecimento é a habilidade de processar dados eficientemente para resolver problemas.
- D) dado é qualquer entrada digital em um computador; informação é a exibição destes dados em formato de relatórios e gráficos; e conhecimento é a interpretação dos relatórios e gráficos, com base na compreensão profunda da informação.

E) dado consiste em uma representação visual; informação é uma interpretação que leva ao conhecimento; e conhecimento é a aplicação teórica da informação.

Comentários:

Vejamos as alternativas:

A) Incorreta. A definição de dado como "caso qualquer" está muito abrangente e pouco precisa. Ademais, o dado não precisa ser intangível (muito embora as medições, por natureza, o sejam). Também incorreto assumir que "qualquer" transformação de dado gera informação: o dado precisa ser interpretado, organizado e contextualizado.

B) Correta. O dado é um fato discreto e objetivo sobre um evento, como a medição de temperatura: 39 °C. Informação é, de fato, o dado dotado de relevância e propósito: 39 °C *de temperatura corporal é febre*. Já o conhecimento é uma mistura de experiências, valores, informações contextuais e insights: 39 °C *de temperatura corporal (febre) costuma indicar infecções ou inflamações e requer acompanhamento médico*.

C) Incorreta. Dado é registro bruto. Informação é o dado interpretado, e conhecimento resulta da compreensão as informações obtidas a partir dos dados.

D) Incorreta. Dado, hoje em dia, costuma ser digital, mas nem tudo que é digital é dado. Tanto informações quanto conhecimento podem ter entrada digital, sem, no entanto, serem dados. Informação é o dado com contexto, e **não precisa de suporte visual** (relatórios ou gráficos) para existir. Por fim, conhecimento não advém, necessariamente, da interpretação de relatórios e gráficos (há muito conhecimento tácito derivado da experiência humana).

E) Incorreta. Essa afirmativa tem alguns problemas, já que o dado é meramente um registro bruto, e registra um objeto, evento ou medição, de modo que a representação visual não é um requisito do dado. O trecho que trata da informação está correto, já que a interpretação (dos dados) pode levar ao conhecimento (quanto as informações são articuladas entre si). O conhecimento deve orientar a tomada de decisão, logo, não se trata meramente de aplicação teórica da informação, mas também prática.

Gabarito: B

30. (Consulplan - 2024 - Prof / Informação e Comunicação - SEED PR)



O dado é considerado um elemento fundamental para a constituição de sistemas organizados. Segundo o autor Silberschatz, os dados são a representação de fatos do mundo real, que podem ser numéricos, alfanuméricos, imagens, sons entre outros tipos. Eles possuem um papel importante em um sistema de informação, pois são coletados pelos usuários para serem processados dentro de uma lógica previamente estabelecida, gerando o resultado da sua interpretação e organização que é conhecido como:

- A) Solução.
- B) Produto.
- C) Algoritmo.
- D) Informação.

Comentários:

A organização e contextualização do dado dá origem à **informação**. Com isso, a resposta correta é a D.

Gabarito: D

31. (FCC - 2018 - APE / Administração Pública ou de Empresas - TCE RS)

Os conceitos de dados, informação e conhecimento são de grande importância no contexto de sistemas de informação. Sobre eles, é correto afirmar que

- A) não são necessários os dados para que se obtenha o conhecimento.
- B) a informação é obtida acrescentando-se significado aos dados.
- C) a informação é obtida a partir do conceito de conhecimento.
- D) o processo de tomada de decisão em um sistema de informação tem por base apenas os dados brutos.
- E) os dados consistem do conhecimento analisado sob diferentes pontos de vista.

Comentários:

Vejamos as alternativas:

- A) Incorreta. Dados são a matéria-prima dos sistemas de informação. A partir dos dados extrai-se a informação e o conhecimento (nessa ordem).
- B) Correta. A informação é obtida acrescentando-se significado aos dados (contextualização, interpretação e organização).
- C) Incorreta. A informação é obtida a partir do **dado**.
- D) Incorreta. O processo de tomada de decisão pode se apoiar em dados e informações, mas tendem a se apoiar principalmente no **conhecimento**: a compreensão contextualizada das informações.
- E) Incorreta. O dado é registro bruto não interpretado.

Gabarito: B

32. (FGV - 2024 - Ana TI / Análise de Negócios de TI - DATAPREV)

Sobre os conceitos de dados, informação, conhecimento e inteligência, é correto afirmar que

- A) dados são informações já interpretadas, enquanto informação é a simples coleta de fatos brutos sem organização.
- B) informação é a aplicação do conhecimento para a tomada de decisões eficazes e resolução de problemas.
- C) conhecimento é a compreensão e internalização da informação, permitindo sua aplicação em diferentes contextos.

- D) inteligência é a coleta de dados brutos de diversas fontes sem qualquer processamento ou análise.
- E) dados e inteligência são termos intercambiáveis no contexto de Business Intelligence (BI), representando o mesmo conceito.

Comentários:

Vejamos as alternativas:

- A) Incorreta. Informações são dados interpretados, e dado representa a simples coleta de fatos brutos sem organização.
- B) Incorreta. A aplicação do conhecimento para a tomada de decisões eficazes representa a **inteligência** da pirâmide informacional.
- C) Correta. É justamente a definição de conhecimento.
- D) Incorreta. A mera coleta de dados brutos de diversas fontes sem processamento ou análise representa meramente um processo e coleta de dados brutos.
- E) Incorreta. Dados e inteligência são conceitos distintos, **não são sinônimos**.

Gabarito: C

33. (FGV - 2024 - APE / Economia de Energia - EPE)



No contexto de ciência de dados, os dados podem ser interpretados como itens elementares, captados e armazenados por recursos de Tecnologia da Informação. Seu propósito é expressar fatos do mundo real de forma a serem tratados no contexto computacional, sendo a base de uma pirâmide hierárquica.

Um nível acima, as informações representam os dados processados, com significado e contextos bem definidos.

Assinale a opção que apresenta corretamente o conceito do último nível dessa pirâmide, correspondendo a um padrão ou conjunto de padrões cuja formulação pode envolver e relacionar fatos e informações.

- A) Aprendizado.
- B) Inteligência.
- C) Conhecimento.
- D) Durabilidade.
- E) Ciência.

Comentários:

Essa questão tem uma pegadinha porque ela não especifica, diretamente, a pirâmide informacional (hierarquia DIKW), cujo último nível é associado à inteligência ou sabedoria. O candidato desatento poderia, por isso, marcar a B.

Em vez disso, o enunciado delimita, ele mesmo, a pirâmide:

1. No primeiro nível estão os dados;
2. No segundo nível as informações;
3. No terceiro nível (o último da estrutura), o conceito correspondente *a um padrão ou conjunto de padrões cuja formulação pode envolver e relacionar fatos e informações*.

O conceito que corresponde à compreensão de padrões, relações e implicações a partir de informações é o **conhecimento**. E por esse motivo a resposta correta é a letra C.

Gabarito: C

34. (FGV - 2024 - Aud CE / Ciência de Dados - TCE PA)



Com relação aos conceitos de dado, informação e conhecimento, avalie as afirmativas a seguir.

- I. Os dados são itens elementares, são cadeias de símbolos e não possuem significado.
- II. São exemplos de conhecimento: tendência de vendas de um produto A em uma região B; relação entre o aumento ou queda do preço de uma ação X, na bolsa de valores, e a variação do câmbio.
- III. As informações correspondem ao dado processado, com significado e um contexto indefinido.

Está correto o que se afirma em

- A) I, apenas.
- B) I e III, apenas.
- C) I e II, apenas.
- D) II, apenas.
- E) I, II e III.

Comentários:

I - Correta. Dados são elementos elementares, representam símbolos sem significado contextual.

II - Correta. O conhecimento implica reconhecimento de padrões, relações e implicações relevantes. A tendência de vendas de um Produto A em uma região B pode ser interpretado como um padrão; a relação entre o aumento ou queda do preço de uma ação X, na bolsa, e a variação do câmbio é um exemplo de uma "relação relevante".

III - Incorreta. O erro está na expressão 'contexto indefinido', pois a informação é resultado de *contextualização* do dado.

Logo, apenas I e II são corretas.

Gabarito: C

35. (IBFC - 2024 - Téc Jud. / Apoio Esp. Desenv. Sistemas - TRF 5)



Leia a frase abaixo.

O ____ é a assimilação e compreensão de informações adquiridas, enquanto a ____ envolve a capacidade de aplicar esse ____ de maneira adaptativa, analítica e inovadora para resolver problemas e tomar decisões eficazes em diferentes contextos.

Assinale a alternativa que preencha correta e respectivamente as lacunas.

- A) inteligência / conhecimento / artefato
- B) conhecimento / inteligência / conhecimento
- C) dado / informação / atributo
- D) conjunto / informação / artefato

Comentários:

De acordo com a Hierarquia DIKW, a assimilação e compreensão de **informações** adquiridas corresponde ao **conhecimento**. E a capacidade de aplicar esse conhecimento de maneira adaptativa representa justamente a **inteligência**.

Logo, o preenchimento das lacunas segue a ordem conhecimento / inteligência / conhecimento.

Gabarito: B

36. (Instituto AOCP - 2017 - Sec Exec - UFBA)



Em tempos de concorrência acirrada, as informações, se utilizadas de forma estratégicas, contribuem para o sucesso das organizações. Em relação ao gerenciamento de informações, julgue, como VERDADEIRO ou FALSO, o item a seguir.

O Dado é um elemento isolado, que ainda não passou por contextualização. O cruzamento ou a organização de dados produz uma informação que pode ser utilizada por gestores e, a partir da aplicação da informação, pode ser produzido conhecimento.

Comentários:

A afirmativa está corretíssima:

- O Dado é um elemento isolado, que ainda não passou por contextualização.
- O cruzamento ou a organização de dados gera informação que pode ser utilizada por gestores.
- A partir da aplicação da informação (no sentido de compreensão e articulação), pode ser produzido conhecimento.

Gabarito: Certa

37. (NEC UFMA - 2016 - Tec / Tecnologia da Informação - UFMA)



Qual a denominação de um conjunto de dados devidamente ordenados e organizados de forma a terem significado?

- A) Informação
- B) Processamento
- C) Cluster
- D) Memória
- E) Dado

Comentários:

O conjunto de dados devidamente ordenados e organizados de forma a terem significado representa a **informação**.

Gabarito: A

38. (Quadrix - 2025 - Ana Sistemas - CFBIO)



No que diz respeito às ferramentas de automação de processos, aos conceitos de dados e aos índices, julgue o item seguinte.

Dados podem ser definidos como elementos brutos que, quando contextualizados, se tornam informação.

Comentários:

Correto, é a definição clássica de dado e de informação.

Gabarito: Certa

39. (Quadrix - 2024 - Ana TI - COREN PR)



Em relação ao conceito de dado, à Lei Geral de Proteção de Dados (LGPD) e à linguagem de programação Qlik, julgue o item a seguir.

Dado é uma coleção organizada de informações relevantes para uma determinada função ou tarefa.

Comentários:

Incorreta. A afirmativa inverte os conceitos. A forma correta seria:

- Informação é uma coleção organizada de dados relevantes para uma determinada função ou tarefa.

Gabarito: Errada

40. (Quadrix - 2023 - Ana Área / Tecnologia da Informação - CREA GO)



Quanto aos conceitos utilizados no tratamento de dados e na análise de dados e informações, julgue o item.

A capacidade de aprender, adaptar, raciocinar, resolver problemas e aplicar conhecimento de forma eficaz em contextos variados é definida como informação.

Comentários:

Incorreta. **Informação** é o dado contextualizado. A capacidade de aprender, adaptar, raciocinar, resolver problemas e aplicar conhecimento de forma eficaz em contextos variados é definida como **inteligência**.

Gabarito: Errada

41. (Quadrix - 2023 - Ana Área / Tecnologia da Informação - CREA GO)



Quanto aos conceitos utilizados no tratamento de dados e na análise de dados e informações, julgue o item.

Conhecimento pode ser definido como uma representação simbólica ou descritiva de fatos, conceitos ou instruções em vários formatos, como, por exemplo, números, palavras, imagens ou símbolos.

Comentários:

A definição da alternativa é a de **dado**, não de conhecimento:

- **Dado** pode ser definido como uma representação simbólica ou descritiva de fatos, conceitos ou instruções em vários formatos, como, por exemplo, números, palavras, imagens ou símbolos.

Logo, incorreta.

Gabarito: Errada

42. (Quadrix - 2023 - Ana Adm / Tecnologia da Informação - CREFITO 14)



É o resultado do processamento de dados, dando-lhes contexto e significado. É quando os dados são organizados, estruturados e relacionados a um determinado contexto, tornando-os compreensíveis.

O texto acima refere-se ao conceito de

- A) informação.
- B) conhecimento.
- C) inteligência.
- D) conhecimento tácito.
- E) conhecimento explícito.

Comentários:

O enunciado refere-se à **informação**:

-
- **Informação** é o resultado do processamento de dados, dando-lhes contexto e significado. É quando os dados são organizados, estruturados e relacionados a um determinado contexto, tornando-os compreensíveis.

Gabarito: A

3 Dados estruturados, semiestruturados e não estruturados

Agora que estudamos os fundamentos, precisamos conhecer um pouco mais sobre a estrutura dos dados, que também é matéria de prova.

Em sistemas de informação, alguns dados seguem um **formato rígido e previsível**, como tabelas de banco de dados. Outros apresentam apenas uma **organização parcial**, como arquivos JSON, XML e documentos com marcações (*markup*). E há também dados produzidos de **forma livre**, como textos, imagens, áudios e vídeos.

Em vista disso, costuma-se classificar os dados em **estruturados**, **semiestruturados** e **não estruturados**. Entender essa distinção é fundamental, porque o grau de organização dos dados influencia diretamente a forma pela qual eles são armazenados, processados, consultados e transformados em informação útil.



Pegadinha!

Cada tipo de dado precisa ser **avaliado individualmente**, de modo que não podemos **consolidar os dados de uma organização inteira** (que geralmente contém dados estruturados, semiestruturados e não estruturados) **com base no tipo predominante**. O que vale é a avaliação individualizada de cada dado.

Nos próximos tópicos, estudaremos cada um destes tipos.

3.1 Dados estruturados

Dados estruturados são dados que estão representados num **formato estrito**. O exemplo mais comum são **tabelas em bancos de dados relacionais**, em que a estrutura dos dados é cuidadosamente definida numa estrutura chamada **esquema** (*schema*) antes dos dados serem coletados. Outro exemplo típico são as planilhas eletrônicas, como as do **Excel**, embora nesse caso a estrutura costume ser mais flexível e menos rigorosamente controlada do que em bancos de dados relacionais.



O que é o esquema?

O **esquema** descreve a estrutura dos dados. E por estrutura devemos entender os tipos das **colunas** (ou, mais formalmente, **atributos**) a serem criadas, as relações entre as diversas tabelas existentes numa mesma base e também as restrições aplicadas a cada coluna da

tabela (por exemplo, uma restrição pode ser que **um valor nunca poderá se repetir** numa coluna).

Esse conceito de esquema será aprofundado quando estudarmos bancos de dados relacionais, mas, para não deixá-los no *vácuo*, uma esquema possível para uma tabela numa base relacional seria o seguinte:

SQL

```
CREATE TABLE aluno (  
    id_aluno INT PRIMARY KEY,  
    nome VARCHAR(100) NOT NULL,  
    cpf CHAR(11) UNIQUE NOT NULL,  
    data_nascimento DATE,  
    email VARCHAR(150) UNIQUE  
);
```

Nesse exemplo, o esquema (em linguagem SQL) define que a tabela se chama `aluno` e possui cinco atributos: `id_aluno`, `nome`, `cpf`, `data_nascimento`, `email`. Também define o tipo de dado de cada atributo: INT para número inteiro, VARCHAR para texto de tamanho variável, CHAR para texto de tamanho fixo e DATE para data.

Posteriormente, a tabela seria alimentada com **dados que necessariamente precisariam seguir o esquema definido**. Do ponto de vista visual, a tabela poderia ser assim representada:

id_aluno	nome	cpf	data_nascimento	email
1	Zé Luis	00000000000	20/03/2000	zeluis@email.com

Assim, dados estruturados representam **estruturas rígidas e previsíveis** que tendem a mudar com pouca frequência. E, quando os dados comecem a ser coletados, em geral teremos um sistema que irá **garantir** que os dados inseridos no banco observem o esquema previamente definido. Também dizemos que os dados estruturados são **tipados**.

A vantagem de dados estruturados é que eles oferecem **maior facilidade para armazenamento, validação e consulta**, especialmente em bancos de dados relacionais. Por essa razão, são amplamente utilizados em sistemas transacionais, aplicações corporativas e diversos serviços da internet.

Dados estruturados

- Estrutura **rígida e previsível**
- Dados são **tipados**
- Definidos num **esquema** (inclusive restrições) a parte.
- Exemplos principais são tabelas em bancos de dados relacionais e algumas estruturas tabulares como planilhas em Excel.
- Mais fáceis para consultar.

Cebraspe - 2024 - Ana Sistemas - CAGEPA



Uma empresa está coletando os dados de campanhas sazonais de publicidade e mantém organizada uma planilha com as informações de cada campanha, com os respectivos períodos de duração e valores investidos.

Nessa situação hipotética, a empresa possui dados do tipo

- A) metadado.
- B) bruto.
- C) não estruturado.
- D) semiestruturado.
- E) estruturado.

Comentários:

O armazenamento de informações em planilha (tabular) é usualmente interpretado como um dado estruturado, de modo que cada coluna da planilha conterá valores com tipos pré-determinados (o esquema é pré-definido). Logo, a alternativa correta é a letra E.

Gabarito: E

Cebraspe - 2023 - ACE - TC DF



A respeito de dados estruturados e dados não estruturados, julgue o item que se segue.

A organização de dados estruturados é totalmente aberta.

Comentários:

Perguntinha capciosa. Em tese, a forma pela qual os dados estruturados são modelados é aberta: cabe à organização decidir, com base no problema a ser resolvido, quais tabelas e relações precisará criar.

Agora, dizer que a organização dos dados é **totalmente aberta** é incorreto, porque dados estruturados possuem requisitos próprios que precisam ser observados, **como a própria definição de esquemas rígidos previamente à existência dos dados**. Logo, não podemos considerar que sua organização é totalmente aberta. Afirmativa incorreta.

Gabarito: Errada

3.2 Dados semiestruturados

Nos **dados semiestruturados**, a rigidez da separação entre o **esquema** e os **dados**, como ocorre nos dados estruturados, deixa de existir. Aqui, costuma-se dizer que o **esquema é informado juntamente com os dados**, sendo mais flexível. Por essa razão, dizemos que dados semiestruturados são **auto-descritivos**. E, regra geral, em se tratando de dados semiestruturados, o **dado precisa ser analisado para que o esquema seja inferido a partir dele**.

Note, portanto, que dados semiestruturados **têm representação heterogênea**, de modo que não são **nem rigidamente estruturados, nem possuem forma livre** (como os não estruturados).



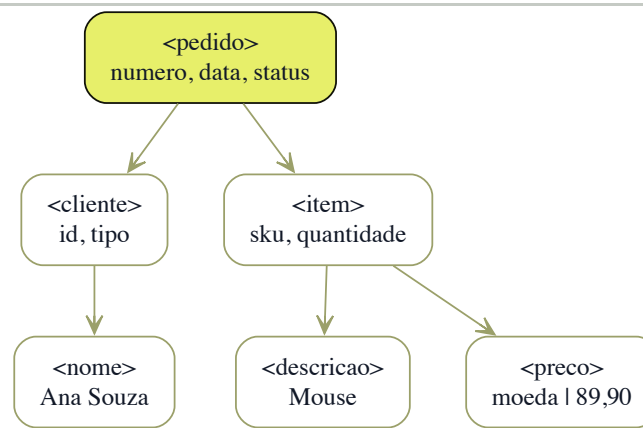
Pra exemplificar...

Documentos XML são exemplos típicos de **dados semiestruturados**. Neles, a informação da estrutura (o esquema) aparece junto aos dados. Veja o exemplo do pedido abaixo, em formato **XML**:

XML

```
<pedido numero="123" data="2026-04-21" status="pago">
  <cliente id="c45" tipo="pf">
    <nome>Ana Souza</nome>
  </cliente>
  <item sku="A10" quantidade="2">
    <descricao>Mouse</descricao>
    <preco moeda="BRL">89,90</preco>
  </item>
</pedido>
```

O elementos como `<pedido>`, `<cliente>` e `<item>` organizam os valores em uma hierarquia, enquanto atributos como numero, id, sku e moeda adicionam metadados ao próprio documento. Notem, portanto, que a estrutura aparece junto aos dados, e até mesmo tipos podem ser inferidos do XML (data, números e textos).



Arquivos **JSON**, similarmente, possuem estrutura hierárquica e também costumam ser associados a dados semiestruturados.

Outro exemplo são os arquivos **CSV**. CSVs **costumam** ser considerados semiestruturados, muito embora tenham um **formato tabular**. Isto porque, CSV pode conter colunas organizadas e separadores bem definidos, mas não contém regras específicas de tipos, chaves, relacionamentos e restrições de integridade. Por isso, ele não é totalmente livre como um texto comum, mas também não possui a rigidez típica de um dado estruturado.

A doutrina brasileira na área define dados **semiestruturados** da seguinte maneira.

- **Definição à posteriori:** esquemas para dados semi-estruturados são usualmente definidos após a existência dos dados, com base em uma investigação de suas estruturas particulares e da análise de similaridades e diferenças.
- **Estrutura irregular:** coleções extensas de dados semanticamente similares estão organizados de maneiras diferentes, podendo algumas ocorrências terem informações incompletas ou adicionais em relação a outras.
- **Estrutura implícita:** muitas vezes existe uma estrutura básica para os dados, porém, essa estrutura está implícita na forma como os dados são apresentados. É necessário realizar uma computação para obter essa estrutura;
- **Estrutura parcial:** o esquema para estes dados nem sempre é completo do ponto de vista semântico e nem sempre todas as informações esperadas estão presentes;
- **Estrutura extensa:** a ordem de magnitude de uma estrutura para estes dados é grande, uma vez que os mesmos são muito heterogêneos. Supondo diferentes formatos para um curriculum vitae, uma união de atributos significativos em cada formato pode produzir um esquema extenso;
- **Estrutura evolucionária:** a estrutura dos dados modifica-se tão freqüentemente quanto os seus valores. Dados Web apresentam este comportamento, uma vez que existe o interesse em manter dados sempre atualizados;
- **Estrutura descritiva e não prescritiva:** dada a natureza irregular e evolucionária dos dados semi-estruturados, as estruturas de representação implícitas ou explícitas normalmente se restringem a descrever o estado corrente de poucas ocorrências de dados similares.

- **Distinção entre estrutura e dados não é clara:** como a estrutura está embutida na descrição dos dados, muitas vezes a distinção lógica entre estrutura e valor não é clara.

Dados semiestruturados

- Para prova, guarde principalmente: dados semiestruturados têm estrutura flexível, irregular, parcialmente implícita e frequentemente embutida nos próprios dados.
- São considerados **auto-descritivos**.
- XML, JSON e CSV costumam ser associados a dados semiestruturados

Cebraspe - 2023 - Ana TI - DATAPREV



A respeito de dados estruturados e não estruturados, julgue o item a seguir.

Arquivos .xml, .csv e .json são exemplos de dados estruturados.

Comentários:

Incorreta. `.xml`, `.csv` e `.json`, genericamente falando, costumam ser associados a dados semiestruturados, por possuírem uma estrutura interna flexível, irregular (que pode variar de dado para dado) e parcialmente implícita nos dados (como a própria hierarquia dos dados, tipos implícitos e metadados). Isto é, não são dados não estruturados, pois possuem esquema flexível, embora possa ser incompleto semanticamente falando, e também não são considerados como dados estruturados, por não possuírem, em regra, um esquema rígido, pré-definido.

Gabarito: Errada

Cebraspe - 2023 - Ana TI - DATAPREV



A respeito de dados estruturados e não estruturados, julgue o item a seguir.

Se, em certa organização, grande parte dos dados for constituída de dados estruturados, mas também houver muitos dados de arquivos .pdf e .docx, então os dados, em geral, serão classificados como semiestruturados.

Comentários:

Incorreta. Quando avaliamos os dados de uma organização, precisamos avaliar cada tipo de dado individualmente. Ademais, **o tipo de dado predominante não define o tipo de dado "da organização"**. Se uma organização possui dados estruturados, não estruturados e semiestruturados, então, **a organização possui dados estruturados, não estruturados e**

semiestruturados. Não cabe aqui tentar encontrar uma "média geral dos dados" e convergir apenas para semiestruturado ou coisa assim.

Por fim, .pdf e .docx são exemplos de dados **não estruturados**.

Gabarito: Errada

3.3 Dados não estruturados

Existe um terceiro tipo de classificação de dados conhecido por **dados não estruturados**. Dados não estruturados são aqueles que não possuem **não possuem uma estrutura (esquema) definido**, de modo que são criados de forma livre.



Pra exemplificar...

Um exemplo típico de dado não estruturado é um arquivo de texto simples, pois ele não segue, necessariamente, um esquema rígido de organização em campos, tabelas ou atributos previamente definidos num esquema. O mesmo pode ser dito a respeito de arquivos de vídeo, de som e de imagens, que embora tenham um padrão técnico definido, não possuem um esquema especificado.

Grande parte dos **arquivos no nossos computador e das empresas são não estruturados: videos, áudios, imagens e arquivos em geral**, assim como **posts de redes sociais (normalmente aparece em provas)**. Dados não estruturados favorecem o armazenamento de qualquer tipo de dado, mas por outro lado dificultam consultas analíticas e pesquisas.



Atenção!

O **Cebraspe** costuma considerar que **dados não estruturados não podem ser armazenados em tabelas de bancos relacionais**, que são associadas a dados estruturados.

Atenção, porque **bancos de dados relacionais permitem, há muito tempo, o armazenamento de conteúdos não estruturados em colunas do tipo BLOB** (e outros tipos binários), incluindo imagens, videos e arquivos em geral.

No entanto, esse tipo de armazenamento **reduz as vantagens dos bancos de dados relacionais, como a facilidade de validação interna do conteúdo e consulta (já que não contém estrutura interna definida)**. Portanto, do ponto de vista prático, é possível armazenar dados não estruturados em bancos de dados relacionais, embora não seja algo desejável.

Dados não estruturados

- **Não possuem** um esquema/estrutura pré-definido
- Produzidos de forma livre
- Menos apropriados para pesquisas
- Maior parte dos arquivos no nosso computador: imagens, videos e arquivos em geral.

Cebraspe - 2024 - ATI - TCE AC



A respeito de Big Data, julgue o próximo item.

Assim como os dados estruturados, os dados não estruturados também podem ser armazenados em bancos de dados relacionais.

Comentários:

Pessoal, a afirmativa foi considerada incorreta pela banca, eu discordo.

Os bancos de dados relacionais permitem, sim, que dados não estruturados sejam armazenados. Para tanto, costumamos utilizar o tipo **BLOB** (binário) para as colunas respectivas (ou seja, uma coluna de uma tabela do banco pode armazenar dados de imagens, videos ou textos). Isto não costuma ser o padrão recomendado para sistemas em produção, mas que é possível, é.

Agora, será que existe alguma interpretação que torne a afirmativa verdadeira? Na minha opinião, só se fizermos uma análise conceitual muito restrita de que bancos de dados relacionais, por serem a representação mais *pura* de dado estruturado com esquema rígido e previsível, só admite dado também dado estruturado puramente tipado (conceitualmente falando), de modo que tanto dados semiestruturados quanto não estruturados não seriam *possíveis* de serem armazenados. Muito forçado, mas é a posição reiterada da banca.

Gabarito: Errada

Cebraspe - 2024 - ERSPT - ANATEL



Julgue o item a seguir, relativo a dados e bancos de dados.

É possível a análise de dados estruturados e não estruturados armazenados no banco de dados, apesar de o formato dos dados não estruturados ser mais diversificado e complexo que o dos dados estruturados.

Comentários:

Correta. É sim possível a análise de dados estruturados e não estruturados armazenados no banco de dados. Dados estruturados destacam-se pela facilidade de análise e pesquisa, principalmente com a linguagem SQL. Dados não estruturados, embora sejam diversificados e de forma livre, também podem ser analisados.

Gabarito: Certa

3.4 Para resumir

Tipo	Definição	Diferença-chave
Dados estruturados	Estrutura rígida e previsível definida em um esquema.	Fácil para realizar buscas e consultas analíticas
Dados semiestruturados	Dado possui estrutura flexível e identificável por tags e marcas (XML, JSON). Auto-descritivo.	Flexível, mais fácil de organizar e de processar do que dados não estruturados.
Dados não estruturados	Não possui esquema de dados pré-definido	Estrutura de forma livre, mas é difícil de processar e analisar

3.5 Questões para exercitar

43. (Cebraspe - 2025 - EPF - PF)



Julgue o item subsequente, relativo à teoria da informação e a banco de dados.

No contexto de banco de dados, dados estruturados são aqueles que não possuem um formato fixo e são armazenados em sistemas como bancos NoSQL, ao passo que dados não estruturados possuem um formato rígido e são armazenados em tabelas relacionais.

Comentários:

Incorreta. A definição está invertida.

No contexto de banco de dados, **dados não estruturados** são aqueles que não possuem um formato fixo e são armazenados em sistemas como bancos NoSQL, ao passo que dados **estruturados** possuem um formato rígido e normalmente armazenados em tabelas relacionais.

Gabarito: Errada

44. (Cebraspe - 2025 - AFT - SEFAZ SE)



Dados estruturados são tipos de dados que

- A) não podem ser indexados, mesmo com organização tabular, devendo ser percorridos integralmente em buscas sequenciais.
- B) dependem de metadados embutidos em arquivos binários para permitirem qualquer forma de busca ou filtragem.
- C) podem ser armazenados somente em sistemas de gerenciamento de banco de dados relacionais.
- D) não podem conter valores nulos, pois a não atribuição de valores descaracterizaria a consistência estrutural exigida.
- E) têm formato e organização que, definidos por esquemas rígidos que estabelecem tipos e restrições para cada campo, permitem consultas por meio de linguagens como SQL.

Comentários:

Essa questão avança um pouco sobre a matéria de Sistemas de Bancos e Dados, SQL e Dados relacionais, mas acho que vale a prática:

A) Incorreta. Dados estruturados em bancos de dados relacionais (tabulares) podem sim ser indexados, de modo que a pesquisa pode ser feita via índice, sem necessidade de percorrer todos os registros em buscas sequenciais. Um índice nada mais é que uma estrutura de dados muito otimizada para pesquisas (similar a um dicionário composto por pares de chave/valor)

B) Incorreta. Dados estruturados costumam ser representados por estruturas tabulares, logo, não há necessidade de embutir metadados em arquivos binários que seja possível realizar buscas ou filtrações.

C) Incorreta. Embora seja a forma mais típica de encontrá-los, dados estruturados podem ser armazenados em outros meios que não apenas em sistemas de gerenciamento de banco de dados relacionais, como o próprio sistema de arquivos.

D) Incorreta. Dados estruturados podem sim conter valores nulos. O próprio esquema que define a estrutura do dado pode prever a existência de valores nulos, ou não, para determinada coluna/campo do dado.

E) É a afirmativa correta.

Gabarito: E

45. (Cebraspe - 2024 - ERSPT - ANATEL)



Julgue o item a seguir, relativo a dados e bancos de dados.

É possível a análise de dados estruturados e não estruturados armazenados no banco de dados, apesar de o formato dos dados não estruturados ser mais diversificado e complexo que o dos dados estruturados.

Comentários:

Correta. É sim possível a análise de dados estruturados e não estruturados armazenados no banco de dados. Dados estruturados destacam-se pela facilidade de análise e pesquisa, principalmente com a linguagem SQL. Dados não estruturados, embora sejam diversificados e de forma livre, também podem ser analisados.

Gabarito: Certa

46. (Cebraspe - 2024 - Ana Sistemas - CAGEPA)



Uma empresa está coletando os dados de campanhas sazonais de publicidade e mantém organizada uma planilha com as informações de cada campanha, com os respectivos períodos de duração e valores investidos.

Nessa situação hipotética, a empresa possui dados do tipo

- A) metadado.
- B) bruto.
- C) não estruturado.
- D) semiestruturado.
- E) estruturado.

Comentários:

O armazenamento de informações em planilha (tabular) é usualmente interpretado como um dado estruturado, de modo que cada coluna da planilha conterá valores com tipos pré-determinados (o esquema é pré-definido). Logo, a alternativa correta é a letra E.

Gabarito: E

47. (Cebraspe - 2024 - Tecno P2 - CTI)



Considerando processos de análise e mineração de dados, julgue o item subsecutivo.

Os dados utilizados como fonte para análise e data mining são considerados semiestruturados quando possuem uma estrutura não homogênea, como arquivos XML e JSON.

Comentários:

Dados como XML e JSON são considerados semiestruturados por serem heterogêneos, com estrutura irregular entre instâncias diferentes de um mesmo tipo de dado. Logo, afirmativa correta.

Gabarito: Certa

48. (Cebraspe - 2024 - APO - MPO)



Julgue o item a seguir, a respeito de conceitos e fundamentos de dados.

Dados estruturados são aqueles que se encaixam perfeitamente nas tabelas de dados e incluem tipos de dados discretos, como números, textos curtos e datas.

Comentários:

A afirmativa está correta. Como a estrutura de dados é definida num esquema rígido e pré-existente, podemos afirmar que o **dado estruturado encaixa perfeitamente nas tabelas de dados**. Quando o dado não se encaixa perfeitamente à tabela (ou ao esquema), ele costuma ser rejeitado pelo sistema de gerenciamento de banco de dados, justamente por não se conformar à estrutura pré-definida.

Gabarito: Certa

49. (Cebraspe - 2024 - ACE - TC DF)

Julgue o próximo item, relativos a dado, informação, conhecimento e inteligência.

É possível que grandes dados numéricos ou textuais não sejam estruturados em razão da ineficiência em modelá-los como uma tabela.

Comentários:

Correta. Em certos casos, grandes volumes de dados textuais ou numéricos não se ajustam bem a um esquema tabular rígido, de linhas e colunas. Com isso, a definição de um esquema rígido (comum para tabelas) para os dados se torna pouco eficiente ou pouco prática.

Nessa hipótese, pode ser mais adequado armazená-los em formatos mais flexíveis, sem esquema pré-definido, caracterizando dados não estruturados.

Gabarito: Certa

50. (Cebraspe - 2024 - ATI - TCE AC)

A respeito de Big Data, julgue o próximo item.

Assim como os dados estruturados, os dados não estruturados também podem ser armazenados em bancos de dados relacionais.

Comentários:

Pessoal, a afirmativa foi considerada incorreta pela banca, eu discordo.

Os bancos de dados relacionais permitem, sim, que dados não estruturados sejam armazenados. Para tanto, costumamos utilizar o tipo **BLOB** (binário) para as colunas respectivas (ou seja, uma coluna de uma tabela do banco pode armazenar dados de imagens, vídeos ou textos). Isto não costuma ser o padrão recomendado para sistemas em produção, mas que é possível, é.

Agora, será que existe alguma interpretação que torne a afirmativa verdadeira? Na minha opinião, só se fizermos uma análise conceitual muito restrita de que bancos de dados relacionais, por serem a representação mais *pura* de dado estruturado com esquema rígido e previsível, só admite dado também dado estruturado puramente tipado (conceitualmente falando), de modo que tanto dados semiestruturados quanto não estruturados não seriam *possíveis* de serem armazenados. Muito forçado, mas é a posição reiterada da banca.

Gabarito: Errada

51. (Cebraspe - 2024 - TJ TSE - TSE)



Em relação a banco de dados em memória, soluções para Big Data e dados estruturados e não estruturados, julgue o item que se segue.

Documentos de uma empresa e postagens nas redes sociais são exemplos de dados estruturados.

Comentários:

Incorreto. Documentos de uma empresa (no sistema de arquivos, por exemplo) e postagens nas redes sociais (texto de forma livre, imagens ou videos) costumam ser exemplos de dados **não estruturados**.

Gabarito: Errada

52. (Cebraspe - 2023 - Ana TI - DATAPREV)



A respeito de dados estruturados e não estruturados, julgue o item a seguir.

Arquivos .xml, .csv e .json são exemplos de dados estruturados.

Comentários:

Incorreta. `.xml`, `.csv` e `.json`, genericamente falando, costumam ser associados a dados semiestruturados, por possuírem uma estrutura interna flexível, irregular (que pode variar de dado para dado) e parcialmente implícita nos dados (como a própria hierarquia dos dados, tipos implícitos e metadados). Isto é, não são dados não estruturados, pois possuem esquema flexível, embora possa ser incompleto semanticamente falando, e também não são considerados como dados estruturados, por não possuírem, em regra, um esquema rígido, pré-definido.

Gabarito: Errada

53. (Cebraspe - 2023 - Ana TI - DATAPREV)



A respeito de dados estruturados e não estruturados, julgue o item a seguir.

Se, em certa organização, grande parte dos dados for constituída de dados estruturados, mas também houver muitos dados de arquivos .pdf e .docx, então os dados, em geral, serão classificados como semiestruturados.

Comentários:

Incorreta. Quando avaliamos os dados de uma organização, precisamos avaliar cada tipo de dado individualmente. Ademais, **o tipo de dado predominante não define o tipo de dado "da organização"**. Se uma organização possui dados estruturados, não estruturados e semiestruturados, então, **a organização possui dados estruturados, não estruturados e semiestruturados**. Não cabe aqui tentar encontrar uma "média geral dos dados" e convergir apenas para semiestruturado ou coisa assim.

Por fim, .pdf e .docx são exemplos de dados **não estruturados**.

Gabarito: Errada

54. (Cebraspe - 2023 - APO - SEPLAN RR)



Com base nos conceitos de dados estruturados e não estruturados, de análise exploratória de dados e de dados abertos, julgue o item seguinte.

Dados de redes sociais são exemplos de dados não estruturados.

Comentários:

O contexto da afirmativa, na minha opinião, dá a entender que o foco é o **conteúdo** dos dados das redes sociais (ou seja, as postagens que nós vemos, como usuários).

Isto porque, internamente, dados das redes sociais podem ser tanto estruturados (em tabelas de bancos de dados relacionais) ou semi estruturados (caso armazenados como JSON, por exemplo, em bancos de dados não relacionais, como NoSQL). Mas como o enunciado não referencia especificamente a forma de armazenamento do dado no sistema da rede social, entendo que a interpretação mais segura é assumir que a proposta de análise se dá em relação ao conteúdo visto.

Assim, no sentido do conteúdo publicado, podemos assumir que dados de redes sociais (o conteúdo que nós vemos) **são exemplos de dados não estruturados**, pois, marcadamente, tais dados representam textos de forma livre, vídeos e imagens, todos exemplos típicos de dados não estruturados.

Gabarito: Certa

55. (Cebraspe - 2023 - ACE - TC DF)



A respeito de dados estruturados e dados não estruturados, julgue o item que se segue.

A organização de dados estruturados é totalmente aberta.

Comentários:

Perguntinha capciosa. Em tese, a forma pela qual os dados estruturados são modelados é aberta: cabe à organização decidir, com base no problema a ser resolvido, quais tabelas e relações precisará criar.

Agora, dizer que a organização dos dados é **totalmente aberta** é incorreto, porque dados estruturados possuem requisitos próprios que precisam ser observados, **como a própria definição de esquemas rígidos previamente à existência dos dados**. Logo, não podemos considerar que sua organização é totalmente aberta. Afirmativa incorreta.

Gabarito: Errada

56. (Cebraspe - 2023 - ACE - TC DF)



A respeito de dados estruturados e dados não estruturados, julgue o item que se segue.

Dados não estruturados são aqueles que, desde a elaboração da estrutura, são pensados estritamente com uma finalidade.

Comentários:

Incorreta. Precisamos interpretar o enunciado à luz do contexto. Como estamos lidando com dados estruturados e não estruturados, a melhor leitura de **estrutura** seria por **esquema**.

Desta forma, **não é correto** afirmar que dados não estruturados, **desde a elaboração da estrutura, são pensados estritamente com uma finalidade**, pois dados não estruturados não possuem estrutura (esquema) elaborado, sendo criados de forma livre.

A presença de um esquema rígido, pré-definido, típica de dados estruturados que implicaria na ideia de **pensados estritamente com uma finalidade**.

Gabarito: Errada

57. (Cebraspe - 2022 - ADP - DPE RO)



Assinale a opção que mostra um dado não estruturado.

- A) data de nascimento
- B) cadastro de pessoa física (CPF)
- C) nome
- D) biografia
- E) sexo

Comentários:

O segredo aqui é verificar se as alternativas representam dados que podem ser armazenados em campos específicos, com formato, tipo ou domínio de valores previsível, como ocorre em bancos de dados relacionais.

Data de nascimento pode ser representada em campo próprio de data. CPF pode ser representado em campo textual ou numérico, com formato esperado. Nome também pode ser armazenado em campo textual específico. Sexo, por sua vez, pode ser representado por valores previamente definidos ou por campo textual controlado.

Restou-nos, portanto, a **biografia**. A biografia é um texto livre, potencialmente longo e com organização interna variável. Logo, é tipicamente um dado não estruturado, pois, embora possa ser armazenada em um campo textual, seu conteúdo não segue um esquema rígido de campos e tipos.

Logo, a alternativa correta é a letra D.

Gabarito: D

58. (Cebraspe - 2022 - ADP - DPE RO)



As fontes de dados não estruturados incluem

- A) a planilha eletrônica.
- B) a rede social.
- C) a linguagem de programação.
- D) o arquivo XML.
- E) o banco de dados.

Comentários:

Vejamos as alternativas:

A) Incorreta. A planilha eletrônica é exemplo costumeiro de **dado estruturado** (colunas possuem tipos definidos).

B) Correta. Acho que o examinador poderia ser mais preciso, mas diante das demais alternativas, só essa pode ser a resposta correta. O caminho aqui é ter um olhar de usuário, de quem consome o **conteúdo** das redes sociais, que são marcadamente textuais e de mídia (áudio, vídeo, imagens). Neste contexto, o conteúdo representa **dados não estruturados**.

C) Incorreta. A **linguagem de programação** representa uma abstração que especifica **regras que regem como um código deve ser criado** e posteriormente traduzido em operações de máquina. Veja que estamos tratando da linguagem (regras definida num padrão), não dos arquivos de código fonte que contém as instruções a serem executadas (esses, sim, podem ser considerados dados não estruturados). A linguagem em si não costuma ser considerada uma fonte de dado.

D) Incorreta. Arquivo XML é exemplo típico de **dado semiestruturado**, com estrutura irregular, tipos que podem ser implícitos ou explícitos, e sem esquema de dados rígido pré-definido.

E) Incorreta. O banco de dados pode ser um exemplo de **dado estruturado**, quando tabular, sendo o exemplo comum os bancos de dados relacionais que se conformam ao padrão SQL.

Gabarito: B

59. (Cebraspe - 2022 - ATCG - MCom)



No que diz respeito aos dados estruturados e não estruturados, aos bancos de dados relacionais, ao modelo de referência CRISP-DM e à modelagem preditiva, julgue o item a seguir.

Dados estruturados são aqueles representados em um formato estrito como, por exemplo, a informação armazenada nos bancos de dados.

Comentários:

Correta. É definição clássica de dados estruturados: um esquema rígido (formato estrito) define a estrutura dos dados.

Gabarito: Certa

60. (Cebraspe - 2022 - ATCG - MCom)



No que diz respeito aos dados estruturados e não estruturados, aos bancos de dados relacionais, ao modelo de referência CRISP-DM e à modelagem preditiva, julgue o item a seguir.

A principal característica dos dados não estruturados é que eles não possuem metadados.

Comentários:

Incorreta. Dados não estruturados **podem possuir metadados**. Uma imagem pode ter metadados EXIF, um vídeo pode ter metadados como duração, codec e data de criação, e um PDF pode registrar autor, título ou data de modificação.

Gabarito: Errada

61. (Cebraspe - 2022 - Esc Pol - PC PB)



Em arquivos do tipo texto, os metadados que permitem armazenar informações por meio de estruturas como XML ou JSON são do tipo

- A) estruturados.
- B) semiestruturados.
- C) hierárquicos.
- D) categorizados.
- E) não estruturados.

Comentários:

XML e JSON são exemplos comuns de dados semiestruturados, por possuírem esquema flexível, evolução irregular (estrutura de dados muda com o tempo para dados de um mesmo domínio) e a estrutura interna do dado (o esquema) encontra-se informado junto aos dados em si (auto-descritivos).

Gabarito: B

62. (Cebraspe - 2021 - APF - PF)



No que se refere aos conceitos de dados estruturados e não estruturados, julgue o próximo item.

Os dados estruturados diferenciam-se dos dados não estruturados pela rigidez em seu formato e pelo fato de poderem ser armazenados em campos de tabelas de um banco de dados relacional.

Comentários:

Dados estruturado têm **formato (esquema) rígido e pré-definido**, enquanto que dados não estruturados **não possuem esquema definido**.

No entanto, não é tecnicamente verdade que **dados não estruturados não podem ser armazenados em campos de tabelas de um banco de dados relacional**. A maior parte dos bancos de dados relacional suporta armazenamento deste tipo de dado (com tipo BLOB ou outros tipos binários).

Porém, a **Cebraspe** adora essa pegadinha, e no caso específico da banca, vamos aceitar o seguinte: **dados não estruturados não podem ser armazenados em tabelas de bancos de dados relacional**, que são associados a dados estruturados.

Logo, afirmativa correta.

Gabarito: Certa

63. (Cebraspe - 2021 - Ana - SERPRO)



Julgue o item a seguir relativos a conceitos de ingestão de dados estruturados, semiestruturados, não estruturados e em streaming.

Os dados não estruturados constam de bancos de dados relacionais que são eminentemente pesquisáveis tanto por meio de consultas realizadas por humanos quanto por meio daquelas realizadas por algoritmos que usam tipos de dados e nomes de campos, como alfabéticos ou numéricos, moeda ou data.

Comentários:

Incorreta. Os dados que constam de bancos de dados relacionais são os **estruturados**. Estes dados, justamente por possuírem uma estrutura rígida e definida que especifica nomes e tipos de campos (colunas), são pesquisáveis por seres humanos ou por algoritmos que usam os respectivos tipos e nomes para recuperar informações de interesse.

Gabarito: Errada

64. (Cebraspe - 2020 - Profissional de TI - ME)



A respeito de dados, informação, conhecimento e inteligência, julgue o próximo item.

Embora com características particulares, dados não estruturados podem ser classificados em sua totalidade, assim como os dados estruturados.

Comentários:

Afirmativa está incorreta. A chave é entender o que **totalidade de classificação** significa.

É possível alcançar **totalidade de classificação de dados estruturados**? Sim, é possível. Precisamos lembrar que dados estruturados **seguem um esquema rígido** em que os nomes e tipos de campos são previamente definidos num esquema. Isto significa que, por exemplo, dada uma tabela de dados com 10 mil linhas, é possível alcançar a classificação dos dados em sua totalidade, porque temos acesso às informações específicas de esquema, de modo que a comparação entre um ou mais campos é plenamente possível.

Agora, vejamos o caso dos **dados não estruturados**. Supondo que tenhamos, também, 10 mil arquivos num pasta do nosso computador (no sistema de arquivos). Estes 10 mil arquivos representam dados não estruturados, logo, podem representar vídeos, imagens, textos e arquivos de som, entre outros arquivos.

Neste cenário, é possível **classificar os dados em sua totalidade**? A resposta é não. Nós podemos classificá-los parcialmente, principalmente por meio de metadados e outras técnicas, mas é extremamente difícil classificar os dados em sua totalidade, pois para isso precisaríamos olhar para o conteúdo dos arquivos e tentar classificá-los de modo equivalente, e este isto é muito difícil para dados não estruturados.

Portanto, afirmativa incorreta.

Gabarito: Errada

65. (Cebraspe - 2018 - EPF - PF)



- CPF
- NOME
- DATA DE NASCIMENTO
- NOME DO PAI
- NOME DA MAE
- TELEFONE
- CEP
- NUMERO

As informações anteriormente apresentadas correspondem aos campos de uma tabela de um banco de dados, a qual é acessada por mais de um sistema de informação e também por outras tabelas. Esses dados são utilizados para simples cadastros, desde a consulta até sua alteração, e também para prevenção à fraude, por meio de verificação dos dados da tabela e de outros dados em diferentes bases de dados ou outros meios de informação.

Considerando essas informações, julgue o item que se segue.

Os dados armazenados na referida tabela são considerados não estruturados.

Comentários:

Incorreta. Os dados armazenados na tabela são considerados **estruturados**, já que seguem um esquema rígido, previsível, que pode ser facilmente **tipado** (formato textual, número, etc). Ademais, estruturas tabulares são tipicamente associadas a dados estruturados.

Gabarito: Errada

66. (Cebraspe - 2018 - ACP - TCE PB)



Com relação a dados estruturados e não estruturados, assinale a opção correta.

- A)** Dados não estruturados podem ser caracterizados por arquivos de diversos tipos — textos, imagens, vídeos, entre outros —, cujas estruturas não são descritas implicitamente.
- B)** Por padrão, documentos do tipo XML (eXtensible Markup Language) são estruturados.

- C) Dados não estruturados de um mesmo grupo possuem as mesmas descrições e, consequentemente, os mesmos atributos.
- D) Por padrão, dados não estruturados são organizados em blocos semânticos.
- E) A alta heterogeneidade facilita as consultas aos dados não estruturados, desde que estes estejam ligados por ponteiros.

Comentários:

Vejamos as alternativas:

A) Correta. Arquivos de textos, imagens, vídeos, entre outros, são exemplos de dados não estruturados. sua estruturas não estão descritas implicitamente (típico de dados semiestruturados).

B) Incorreta. Por padrão, documentos do tipo XML são associados a dados semiestruturados.

C) Incorreta. Dados não estruturados, ainda que parte de um mesmo *grupo*, podem conter descrições e atributos diferentes, afinal, possuem forma livre.

D) Incorreta. Por padrão, dados não estruturados **não possuem** organização semântica entre os blocos que os compõem.

E) Incorreta. A alta heterogeneidade **dificulta** as consultas aos dados não estruturados, especialmente por não haver um esquema prévio que defina a estrutura dos dados (campos/tipos a serem buscados). É a existência de um esquema (quanto mais rígido, melhor) que facilita as consultas, por isso dados estruturados são mais fáceis de consultar/pesquisar.

Gabarito: A

67. (Cesgranrio - 2022 - PNS - ELETRONUCLEAR)



Há dados não estruturados que podem ser categorizados e preparados para uso em ambiente de BI.

Esses dados não estruturados podem ser, por exemplo:

- A) operações de débito e crédito realizadas em uma conta-corrente bancária
- B) postagens e respectivos comentários em uma rede social
- C) quantidades em estoque por produto em um site de comércio eletrônico, ao final do dia
- D) temperatura média de uma cidade, dia a dia
- E) valores de cotação de cada ação negociada em uma bolsa de valores, hora a hora

Comentários:

Vejamos as alternativas:

A) Incorreta. Operação de débito e crédito realizadas em uma conta-corrente são, possivelmente, representados de modo estruturado em bases de dados relacionais, ou, quem sabe, de modo semiestruturado em arquivos JSON, XML, entre outros. É bastante improvável que estejam armazenados de modo não estruturado (em arquivos textos, por exemplo).

B) Correta. Postagens e respectivos comentários em uma rede social são exemplo corriqueiros, que as bancas adoram, associados a dados não estruturados.

C) Incorreta. Quantidades em estoque por produto podem ser armazenados em estruturas tabulares, com esquema rígido e previsível, por isso são normalmente associados a dados estruturados.

D) Incorreta. A temperatura média de uma cidade, dia a dia, representa uma série temporal que pode ser coletada e armazenada em estruturas tabulares com campos de tipo específico, num esquema pré-definido. Portanto, seria normalmente associada a um dado estruturado.

E) Incorreta. Valores de cotação de cada ação negociada em uma bolsa de valores, hora a hora, também representam uma série temporal que pode ser coletada e armazenada em estruturas tabulares com campos de tipo específico, num esquema pré-definido (dado estruturado).

Gabarito: B

68. (FCC - 2026 - Ges Gov - SCGE PE)



Durante o processo de análise de grandes volumes de informações, um gestor precisa identificar corretamente o tipo de dado com o qual está trabalhando. Em um mesmo projeto, ele utiliza uma planilha contendo:

- I. arquivos JSON com informações de navegação.
- II. postagens em redes sociais contendo textos e imagens.
- III. registros tabulares de clientes.

Nessa situação, os dados são classificados como

- A) I. semiestruturados; II. estruturados; III. não estruturados.
- B) I. estruturados; II. não estruturados; III. semiestruturados.
- C) I. semiestruturados; II. não estruturados; III. estruturados.
- D) I. estruturados; II. semiestruturados; III. não estruturados.
- E) I. não estruturados; II. estruturados; III. semiestruturados.

Comentários:

Vamos avaliar as afirmativas:

I. arquivos JSON com informações de navegação são tipicamente **dados semiestruturados**, postagens em redes sociais contendo textos e imagens.

II. postagens em redes sociais contendo textos e imagens são exemplos de **dados não estruturados**, pois não possuem esquema prévio que determine a estrutura interna do dado, sendo produzidos de forma livre.

III. Registros tabulares de clientes são notadamente considerados **dados estruturados**, pois costumam possuir um esquema prévio que define a estrutura e o tipo dos dados (dados tipados).

Logo, a sequência correta é, **I. semiestruturado, II. não estruturado, III. estruturado.**

Gabarito: C

69. (FCC - 2026 - FTE - SEFAZ MT)



O Diretor de TI de uma Secretaria da Fazenda precisa modernizar a arquitetura de dados da fiscalização, que atualmente lida com cadastros de contribuintes armazenados em tabelas relacionais, Notas Fiscais Eletrônicas em formato XML e relatórios de fiscalização em PDF e texto livre (linguagem natural), em geral tratados como dados não estruturados para fins arquiteturais e analíticos, podendo demandar técnicas de extração e processamento para uso sistemático. Para garantir que a estratégia de armazenamento e análise seja eficiente e compatível com as características desses diferentes tipos de dados, o analista responsável deve classificá-los e tratá-los considerando que

- A)** a infraestrutura fazendária deve focar em dados estruturados, pois qualquer informação textual ou documental de fiscalização pode ser convertida automaticamente em colunas de tabelas sem perda de contexto ou significado para o fisco.
- B)** os dados estruturados possuem esquema rígido e organização tabular; os semiestruturados possuem metadados e organização hierárquica que permitem flexibilidade; e os não estruturados carecem de um modelo predefinido, exigindo ferramentas específicas de processamento e análise para coexistirem no ecossistema de fiscalização.
- C)** os dados estruturados, como os cadastros de contribuintes, possuem esquema flexível e ausência de organização fixa, sendo ideais para armazenar documentos XML sem necessidade de validação ou transformação para o modelo relacional.
- D)** os dados não estruturados possuem esquemas rígidos e tabelas normalizadas previamente definidas, o que permite sua inserção direta em bancos de dados SQL transacionais para consultas imediatas.
- E)** os dados semiestruturados, como o XML das Notas Fiscais, são caracterizados pela ausência de metadados e organização hierárquica, devendo ser tratados apenas como objetos binários para evitar a complexidade da extração de campos específicos.

Comentários:

Questão bem elaborada e muito boa para exercitarmos os conceitos:

A) Incorreta. O problema aqui é afirmar que dados não estruturados podem ser **automaticamente convertidos** em colunas de tabelas sem perda de contexto ou significado. Isto não é verdade, pois a conversão de dados não estruturados, como informação textual (um relatório fiscal extenso), para estrutura tabular é **difícil, requer processamento adicional**, e, regra geral, **resulta em perda de informação**.

B) Correta! A definição de dados não estruturados, semiestruturados e não estruturados está corretíssima.

C) Incorreta. O problema é afirmar que dado estruturado tem esquema flexível e ausência de organização fixa, o não é verdade. Dados estruturados possuem **esquema rígido e pré-definido**.

D) Incorreta. Dados não estruturados **não possuem esquema**, nem rígido nem flexível. Possuem forma livre, por isso são mais difíceis de pesquisar e analisar.

E) Incorreta. Dados semiestruturados **possuem** metadados e organização hierárquica, como documentos XML, em que o esquema (ainda que semanticamente incompleto) anda junto com os dados.

Gabarito: B

70. (FGV - 2026 - AJ - TJ RJ)



Diferentes sistemas produzem dados em formatos variados, que podem ser classificados em dados estruturados, semiestruturados ou não estruturados.

Nesse contexto, relacione os tipos de dados às suas respectivas descrições.

1. Dados estruturados
2. Dados semiestruturados
3. Dados não estruturados

() Gravações em áudio e vídeo de audiências públicas, armazenadas em arquivos MP4, acompanhadas apenas de nome do arquivo e data de criação.

() Registros de protocolo eletrônico armazenados em tabelas de banco de dados relacional, com campos bem definidos (número do processo, data, unidade, assunto) e chaves primárias/estrangeiras.

() Arquivos de log de acesso ao portal de serviços do governo, registrados em formato JSON, contendo campos como timestamp, user_id, endpoint, status_code, com alguns campos opcionais variando conforme o tipo de requisição.

A sequência correta é:

- A) 1, 2 e 3;
- B) 1, 3 e 2;
- C) 2, 1 e 3;
- D) 3, 1 e 2;
- E) 3, 2 e 1.

Comentários:

Vamos avaliar as afirmativas:

- Gravações em áudio e vídeo de audiências públicas, armazenadas em arquivos MP4, acompanhadas apenas de nome do arquivo e data de criação.

Arquivos MP4 (vídeo) são exemplos típicos de **dados não estruturados**. Embora possuam um formato técnico de arquivo, o conteúdo audiovisual não está organizado em campos tabulares ou em um esquema analítico rígido, como ocorre em registros estruturados de bancos relacionais.

- Registros de protocolo eletrônico armazenados em tabelas de banco de dados relacional, com campos bem definidos (número do processo, data, unidade, assunto) e chaves primárias/estrangeiras.

Tabelas de banco e dados relacional são tipicamente representativos de **dados estruturados** que possuem um **esquema rígido** que define a estrutura dos dados a serem armazenados.

- Arquivos de log de acesso ao portal de serviços do governo, registrados em formato JSON, contendo campos como timestamp, user_id, endpoint, status_code, com alguns campos opcionais variando conforme o tipo de requisição.

Arquivos JSON, tal qual arquivos XML, são marcadamente **dados semiestruturados**, já que o informações do esquema andam juntas dos dados (a hierarquia, tipos implícitos ou explícitos, etc), além de possuírem estrutura irregular (dois JSONs diferentes com dados de um mesmo domínio podem ter estrutura de campos diferentes).

Desta forma, a sequência correta é **3, 1, 2**.

Gabarito: D

71. (FGV - 2024 - Tecno PD - Pref Caraguatatuba)



Os dados podem ser classificados de diversas formas, uma das categorias de classificação diz respeito a forma da estruturação dos dados em repositórios de dados, analise os itens a seguir.

I. Dados estruturados são dados que foram formatados e transformados para se adaptar a um modelo de dados relacional bem definido. Os dados brutos são mapeados em campos predefinidos que podem ser extraídos e lidos por meio de declarações SQL.

II. Dados semiestruturados é categoria localizada entre dados estruturados e não estruturados. São um tipo de dado que possui algumas características consistentes e definidas e não se limitam a uma estrutura rígida como a necessária para bancos de dados relacionais.

III. Dados não estruturados podem ser qualquer tipo de dado não esteja em um formato específico. São suportados apenas pelos sistemas de arquivos e não pelos sistemas de bancos de dados, exemplos podem imagens, sons ou desenhos.

Está correto o que se afirma em

- A) I apenas.
- B) III apenas.
- C) I, II e III.
- D) I e II, apenas.
- E) II e III, apenas.

Comentários:

Vejamos as afirmativas:

- I. Correta. A definição de dado estruturado está condizente com o que prega a doutrina.
- II. Correta também. Dados semiestruturados (como JSON e XML), não possuem um esquema rígido, são auto-descritivos e ficam de fato localizados entre dados estruturados e não estruturados.
- III. Incorreta. O erro da afirmativa está em dizer que dados não estruturados são suportados **apenas** pelo sistema de arquivos, o que não é verdade. Mesmo **sistemas de bancos de dados podem armazenar dados não estruturados** (até mesmo como **binário**).

Logo, a alternativa correta é a letra D.

Gabarito: D

72. (FGV - 2024 - AFTRM - Pref Cuiabá)



Dados podem ser classificados em algumas definições, que podem variar a depender de sistemas, processos e aplicações a que se destinam. Com relação a sua estrutura, podemos classificá-los, dentre outras formas, entre estruturados, semiestruturados e não estruturados.

Os itens a seguir contém alguns exemplos de dados que podem se enquadrar dentro dessas três definições.

- I. Um banco de dados relacional.
- II. Grafos que representam relações de usuários em uma rede social.
- III. Um e-mail, que contém texto, imagens e um campo destinado a informar a data e a hora onde foi enviado.

O(s) item(s) que contém dados do tipo não estruturado é(são):

- A) II, apenas.
- B) I e II, apenas.
- C) II e III apenas.
- D) I, II e III.
- E) III, apenas.

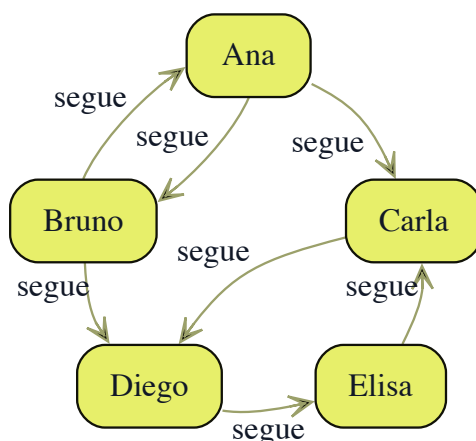
Comentários:

Questão difícil, por conta da assertiva 2, relacionada a grafos.

- I. Incorreta. Um banco de dados relacional é representação tradicional de **dado estruturado**.
- II. Correta. Grafos, por definição, representam uma estrutura de nós e arestas que poderiam, em tese, ser considerados semiestruturados: não possuem esquema rígido, pré-definido, mas possuem estrutura interna organizada em nós e arestas, de modo que esquema, ainda que incompleto semanticamente, estaria junto aos dados.

O problema é que, no contexto do enunciado, os grafos que representam relações de usuários em uma rede social conteriam **dados não estruturados em seus nós**, tais como curtidas, postagens, imagens, preferências de perfil e atributos de usuário, de modo que o grafo em si se tornaria não estruturado por conter elementos deste tipo. Este foi o argumento da banca no julgamento de recurso contra a questão.

Na minha opinião, seria muito difícil adotar tal entendimento unicamente com base no enunciado. Isto porque, a *relação entre usuários* poderia indicar tão-somente o relacionamento entre os usuários (seguidores), que faria o grafo cair em dado semiestruturado, algo simples como:



Interpretação possível de grafo de relações entre usuários

III. Correta. Um e-mail com texto, imagens e campos para informar data e hora representa dados **não estruturados**. A questão aqui é entender que o e-mail, puramente como colocado, possuirá um conjunto de texto, imagens e campos sem uma estrutura/esquema pré-definido, desta forma não há como considerá-lo um dado estruturado ou semiestruturado. Diferente seria se o email fosse uma "tabela de um banco de dados", com colunas para texto, imagens, data e hora (cada qual com tipo pré-definido num esquema).

Logo, alternativa correta é a letra C.

Gabarito: C

73. (FUNDATEC - 2023 - Ana Sistemas - BRDE)



Assinale a alternativa que contém um exemplo de dados não estruturados.

- A) Dados extraídos de uma rede social de um banco podem ser classificadas em polaridades como positivas ou negativas, ou agrupadas por similaridade de assunto.
- B) Um sistema bancário armazena dados sobre seus clientes, sobre suas respectivas contas bancárias e sobre as transações executadas sobre tais contas.
- C) Funcionários de um departamento de recursos humanos observam os candidatos a uma vaga de emprego e fazem anotações sobre as suas atitudes durante a realização de atividades de um processo de seleção.
- D) Uma empresa deseja lançar um novo produto e, para isso, realiza pesquisas sobre sua aceitação com um conjunto específico de pessoas que representa o público-alvo.

- E) Uma corrida de Fórmula 1 na qual se medem os tempos que cada piloto gasta para chegar em diferentes trechos de um circuito.

Comentários:

Vejamos as alternativas:

- A) Correta. Dados extraídos de rede social são associados a dados não estruturados (o conteúdo dos textos dos posts). A classificação em polaridades é irrelevante para o enunciado.
- B) Incorreta. O armazenamento de dados de clientes num sistema bancário muito provavelmente se dá em um banco de dados relacionado, de modo que se tratam de dados estruturados.
- C) Incorreta. As anotações podem ser feitas de modo estruturado, através de um formulário que posteriormente será salvo numa tabela de banco de dados.
- D) Incorreta. A realização de pesquisas sobre aceitação de produto com um conjunto de pessoas também pode ser considerado um dado estruturado, caso essa pesquisa seja tabulada em campos/colunas com tipos definidos, para posterior análise de resultados.
- E) Incorreta. Medições de tempo em intervalos podem ser associados a séries temporais, frequentemente organizadas em uma estrutura previsível, com campos rígidos. Por isso, podem ser associadas a dados estruturados.

Gabarito: A

74. (Fuvest - 2023 - Ana Sistemas - USP)



Suponha que você precisa analisar posts em uma rede social sobre opiniões de clientes da sua empresa.

O conjunto de textos desses posts são considerados dados

- A) categóricos.
- B) estruturados.
- C) intervalares.
- D) nominais.
- E) não estruturados.

Comentários:

Posts de rede social são marcadamente **não estruturados**, já que além da descrição textual, podem conter mídias associadas (vídeos e imagens), todos dados não estruturados sem esquema que os defina.

Logo, a única alternativa correta é a letra E.

Gabarito: E

75. (IBFC - 2024 - AC - ECT)



Os dados dentro do sistema de gerenciamento de banco de dados (SGBD) podem estar organizados de três formas: estruturados, não estruturados ou semiestruturados. Assinale a alternativa que apresenta qual elemento podemos encontrar em uma estrutura de dados não estruturados.

- A) Imagens
- B) Tags de um arquivo HTML
- C) Um atributo Nome do tipo Varchar(40)
- D) Arquivo JSON

Comentários:

Vejamos as alternativas:

A) Correta. Imagens são exemplo típicos de dado não estruturado.

B) Incorreta. Tags de um arquivo HTML são considerados dados semiestruturados, tal qual arquivos XML. Isto porque, as tags representam a estrutura implícita e flexível inerentes aos dados semiestruturados.

C) Incorreta. Um atributo Nome do tipo Varchar(40) é uma unidade de dado estruturada (um texto). Poderia corresponder a uma definição de campo tipado no esquema de uma tabela de banco de dados relacional.

D) Arquivo JSON, assim como o XML, corresponde a um dado semiestruturado. Arquivos JSON possuem estrutura hierárquica entre campos, tipos implícitos dos dados, nomes de campos, de modo que o esquema dos dados, ainda que eventualmente com semântica incompleta, está junto dos dados.

Gabarito: A

76. (Instituto Consulplan - 2025 - AJ - TJ RO)



No contexto de análises estatísticas, os dados podem estar disponíveis em diferentes tipos de organização. Dependendo da forma de definição e armazenamento, os conjuntos de dados são tipicamente caracterizados como estruturados, semiestruturados e não estruturados.

Diante do exposto, analise as afirmativas a seguir.

I. Enquanto dados não estruturados são normalmente armazenados em seu formato nativo (como vídeos, imagens ou documentos de texto), dados estruturados são armazenados em linhas e colunas e podem ser mapeados para campos predefinidos.

II. Ao contrário dos dados estruturados, que podem ser organizados e acessados por meio de bancos de dados relacionais, dados não estruturados não têm um modelo rígido de organização predefinido.

III. Conjuntos de dados semiestruturados usam tags e marcadores (ou seja, metadados) em vez de exigir um esquema predefinido em tabelas, linhas e colunas, como nos dados estruturados.

IV. Os data lakes podem ser projetados para armazenar dados semiestruturados juntamente com dados estruturados, permitindo a coexistência de dados de ambos os tipos.

Está correto o que se afirma em

- A) I, II, III e IV.
- B) I e III, apenas.
- C) II e IV, apenas.
- D) I, II e III, apenas.
- E) I, III e IV, apenas.

Comentários:

Todas as afirmativas estão corretas. O problema maior poderia advir da assertiva IV, que trata de data lakes (ainda não estudado nas aulas). No entanto, neste ponto, cabe ressaltar que data lakes podem ser projetados para armazenar qualquer tipo de dado, assim como o sistema de arquivos do nosso computador.

Desse modo, a alternativa correta é a A.

Gabarito: A

77. (Instituto Consulplan - 2023 - Ana MP - MPE MG)



Dado é a representação utilizada para gerar uma informação. Um dado pode ser estruturado ou não estruturado. Em relação às diferenças entre dados estruturados e dados não estruturados, marque V para as afirmativas verdadeiras e F para as falsas.

- () Em dados estruturados, dados de um mesmo registro possuem relação entre eles.
- () Textos, documentos, banco de dados, redes sociais, imagens, vídeos e áudios são exemplos de dados não estruturados.
- () Dados estruturados permitem que tipos de dados diferentes das estruturas preestabelecidas sejam carregados.
- () Em dados estruturados, atributos e campos são definidos por um esquema.

A sequência está correta em

- A) F, V, V, F.
- B) V, V, V, V.
- C) V, F, F, V.
- D) F, F, F, F.

Comentários:

Vejamos as assertivas:

I - Verdadeira. Em dados estruturados, dados de um mesmo registro possuem relação entre eles. A relação aqui pode ser o mero pertencimento a uma mesma tabela, e de complemento da informação.

II - Falsa. O erro está em incluir **banco de dados** como exemplo de dado não estruturado. Bancos de dados relacionais (espécie de banco de dados) são exemplos de dados estruturados.

III - Falsa. Dados estruturados **não permitem** que tipos de dados diferentes das estruturas preestabelecidas sejam carregados. Isto porque, dados estruturados seguem um esquema rígido pré-definido que precisa, necessariamente, ser observado quando um novo dado é carregado.

IV - Verdadeira. Em dados estruturados, atributos e campos são definidos num esquema rígido de dados, que define campos e tipos de dados existentes na estrutura (geralmente tabular).

Logo, a sequência correta é V, F, F, V, letra C.

Gabarito: C

78. (Instituto Consulplan - 2022 - Tec - MPE PA)



Dados são um conjunto de fatos em estado bruto que podem ser utilizados para tomadas de decisão. Os dados podem ser divididos entre estruturados e não-estruturados. Sobre os dados estruturados, assinale a afirmativa correta.

- A) Mensagens de e-mail são dados estruturados.
- B) Dados estruturados são organizados em linhas e colunas.
- C) Dados estruturados são armazenados em arquivos no formato ".TXT".
- D) Mecanismos de busca possuem dificuldade para pesquisar dados estruturados.

Comentários:

Vejamos as afirmativas:

A) Incorreta. Mensagens de e-mail são dados **não estruturados**, já que a natureza de uma mensagem de texto é possuir forma livre, sem um esquema predeterminado para estruturação do dado.

B) Correta. Dados estruturados são organizados de forma tabular, em linhas e colunas, como em bancos de dados relacionais.

C) Incorreta. Arquivos ".TXT" com texto livre são associados a dados não estruturados, logo, não são apropriados para o armazenamento de dados estruturados. Dados estruturados são frequentemente armazenados em sistemas de bancos de dados relacionais.

D) Incorreta. Mecanismos de busca encontram **facilidade** para pesquisar dados estruturados. O esquema pre-definido de campos e tipos facilita demais as pesquisas.

Gabarito: B

79. (Quadrix - 2024 - Tec ESI - CREF 9)



A principal vantagem dos dados estruturados em comparação aos dados não estruturados é que eles

- A) são mais difíceis de processar e analisar.
- B) requerem menos armazenamento de dados.
- C) podem ser facilmente organizados, pesquisados e analisados por meio de consultas e algoritmos de banco de dados.
- D) são mais flexíveis e adaptáveis a diferentes tipos de informações.
- E) incluem conteúdo multimídia, como imagens e vídeo.

Comentários:

Dados estruturados têm, como principal vantagem, a facilidade de organização, pesquisa, classificação e análise, porque sua estrutura é rígida e previsível, estabelecida num esquema.

Justamente por isto, são também difíceis de adaptar para diferentes tipos de informações e não costumam incluir conteúdo multimídia, como imagens e vídeos (que são associados a dados não estruturados).

Logo, a alternativa correta é a letra C.

Gabarito: C

80. (Vunesp - 2024 - Analista - SPCINE)



É possível estabelecer uma classificação dos dados de uma organização em dados estruturados e dados não estruturados, sobre os quais é correto afirmar que

- A) dados estruturados possuem comprimento e tipo definidos, tais como datas e nomes.
- B) dados de temperatura ao longo dos dias representam tipos de dados não estruturados.
- C) bases de dados orientadas a objetos armazenam exclusivamente dados não estruturados.
- D) dados não estruturados seguem exclusivamente formatos numéricos.
- E) dados de mídias sociais representam tipos de dados estruturados.

Comentários:

A) Correta. Dados estruturados possuem tipo e, muitas vezes, comprimento definido dentro de um esquema rígido, como ocorre com campos de nome, data ou valor numérico em uma tabela.

B) Incorreta. Dados de temperatura ao longo dos dias são valores numéricos associados a uma série temporal, normalmente com tipo definido e estrutura previsível. Por isso, são tipicamente classificados como dados estruturados, e não como dados não estruturados.

C) Incorreta. Bases de dados orientadas a objetos armazenam dados de **estrutura fixa** (um objeto representa uma estrutura fixa, tipada).

D) Incorreta. Dados não estruturados podem ter formato textual, com textos longos.

E) Incorreta. Dados de mídias sociais são exemplos recorrentes de tipos de dados **não estruturados**.

Gabarito: A

81. (Vunesp - 2023 - APPGG - Pref SP)



Um analista de políticas públicas trabalha com diversos formatos e tipos de dados. Identificar corretamente essas características no começo das atividades é essencial para realizar a escolha correta das ferramentas e técnicas que serão utilizadas.

Assinale a alternativa que traz apenas dados semiestruturados.

- A)** Arquivo CSV (Comma Separated Values) e arquivos de áudio.
- B)** Arquivo JSON (JavaScript Object Notation) e arquivos de imagens.
- C)** Arquivo JSON (JavaScript Object Notation) e arquivo XML (eXtensible Markup Language).
- D)** Arquivos PDF (Portable Document Format) e arquivos de vídeo.
- E)** Tabela de um banco de dados relacional e arquivo XML (eXtensible Markup Language).

Comentários:

Vamos analisar as alternativas:

- A) Incorreta. Arquivos de áudio representam dados não estruturados.
- B) Incorreta. Arquivos de imagens representam dados não estruturados.
- C) Correta. Tanto arquivos JSON quanto XML representam dados semiestruturados.
- D) Incorreta. Arquivos PDF e arquivos de vídeo são ambos exemplos de dados não estruturados.
- E) Incorreta. Tabela de um bancos de dados relacional é sempre associada a dado estruturado.

Gabarito: C

4 Observações, atributos e métricas

Antes de iniciarmos o aprofundamento do nosso estudo de Fluência em Dados, precisamos definir alguns conceitos que serão fundamentais para estruturar nosso entendimento para aulas futuras. Este módulo, portanto, se propõe a expor pequenos *tijolos* que serão de fundação para o que veremos adiante em Ciência de Dados e Mineração de Dados.

E, mais do que ser base para o futuro, costuma ser diretamente cobrado em prova.

4.1 Observações e granularidade

A melhor forma de entendermos dados, observações e granularidade é por meio de um exemplo. Vejamos a tabela abaixo, que ilustra o registro de produtos vendidos por uma empresa num banco de dados:

data	nr_nfe	cnpj	produto	quantidade	total	icms
20/03/26	1001	12345678912345	notebook	1	8.000	1.440
20/03/26	1001	12345678912345	mouse	2	200	38

Vejam que cada linha da tabela representa um item da nota fiscal. Em análise de dados, cada linha representa uma **observação** ou exemplo a ser analisado. **Observações** correspondem a eventos, objetos ou experimentos medidos no mundo real. Dados de um paciente, ou os registros de uma venda acima, podem ambos ser considerados observações.

O **nível de detalhe de cada observação** é o que determina sua **granularidade**. A **granularidade** é a definição precisa do que cada registro representa.

No exemplo acima, os itens da nota fiscal tem **alta granularidade**, já que os registros estão detalhados com o número da nota, o cnpj do vendedor, o nome do produto vendido, a quantidade, etc.

Por outro lado, **baixa granularidade** geralmente envolve visões agregadas por algum critério de agrupamento. Por exemplo, em vez de detalhar item a item como no exemplo acima, o tabela poderia trazer valores agregados, como "total vendido por mês".

Neste caso, nós continuaríamos a ter informações dos itens vendidos, mas numa granularidade muito mais baixa, já que os itens deixam de ser detalhados individualmente em favor da visão macro do total vendido pela empresa no mês.

Observações e Granularidade

- **Observações** correspondem a eventos, objetos ou experimentos medidos no mundo real, como um paciente ou registros de venda.
- **Granularidade** determina o nível de detalhe que cada observação possui: **granularidade fina** indica alto detalhe; **granularidade grossa** indica visões agregadas por mês, dia, empresa, ou qualquer outro critério de agrupamento.
- **Conjunto de dados** (ou apenas dados) representa uma coleção de observações com seus respectivos atributos.

Inédita - 2026 - Ultra TI

Julgue o próximo item, relativo a observações e granularidade.

A granularidade define precisamente o que cada registro representa; por isso, dois conjuntos formados a partir das mesmas vendas podem apresentar granularidades distintas se um registrar cada item vendido e o outro registrar apenas totais mensais.

Comentários:

A granularidade define o grau de detalhe do registro. O que é mais detalhado é mais granular. Agregações e totalizações, por outro lado, por terem menos detalhes individuais dos elementos, são menos granulares. Por isso, afirmativa correta.

Gabarito: Certa

4.2 Atributos e Variáveis

Um **atributo** representa uma **característica de uma entidade, evento ou observação**. Em uma tabela, os atributos são representados pelas colunas da tabela. Em uma imagem, os atributos são representados pelos valores que representam cada pixel.

Na tabela anterior de itens vendidos, os atributos são:

- **data**
- **nr_nfe**
- **CNPJ**
- **produto**
- **total**
- **quantidade**
- **ICMS**

Uma **coleção de atributos** é o que **descreve um objeto ou evento**.

As **variáveis**, por sua vez, podem assumir diferentes valores entre as observações que serão analisadas. Na prática, contudo, variáveis e atributos são utilizados como sinônimos. Mas, conceitualmente, **atributo corresponde à descrição formal dos dados, já a variável é esse atributo no contexto de uma análise de dados.**

A título de exemplo, ainda nos socorrendo da tabela anterior, o **produto** é um atributo da nota fiscal. Mas se precisarmos investigar a receita da empresa por produto vendido, ele passa a funcionar como uma variável analítica.



Atenção!

Em **Mineração de Dados**, um conjunto de atributos utilizados para identificar uma determinada observação é denominado como **vetor de atributos**, ou **vetor de características**. No exemplo acima, um vetor representativo do primeiro item seria:

[20/03/26, 1001, 12345678912345, notebook, 1, 8000, 1440]

Vejam que os atributos estão necessariamente ordenados conforme a representação das colunas.

Isto, contudo, não significa dizer que o vetor precisa seguir exatamente a ordem da disposição das colunas numa tabela. O importante é que **alguma ordem seja estabelecida entre os atributos e que tal ordem seja mantida em observada em diferentes vetores**, de modo que os elementos numa mesma posição em diferentes vetores representem exatamente o mesmo atributo.

Por fim, os atributos podem ser classificados em diferentes tipos, mas os dois grupos principais são os de **atributos qualitativos** (nominal, ordinal, binário) e **atributos quantitativos** (discreto, contínuo, intervalar e razão)

Atributos e Variáveis

- Um atributo representa uma **característica de uma entidade, evento ou observação.**
- Uma variável representa um atributo **no contexto de uma análise de dados.**

4.2.1 Atributos qualitativos

Atributos Nominais

Atributos nominais, de "nomes", referem-se a dados categóricos cujos valores representam diferentes categorias ou rótulos sem nenhuma ordem implícita. Atributos nominais são

normalmente utilizados para representar nomes, descrições ou rótulos de objetos, entidades ou conceitos.

Atributos Binários

Atributos binários são aqueles em que os valores possíveis para os dados são apenas dois distintos. Estes atributos são normalmente utilizados para representar sim ou não, presença ou ausência, verdadeiro ou falso, 0 ou 1.



Atenção!

Observem que atributos qualitativos não precisam ser, necessariamente, textuais. Eles podem ser também representados por números, desde que o valor numérico represente um valor categórico em vez de uma medição qualquer. Um código de cidade do IBGE, como 1302603 (Manaus) é um atributo nominal. Mas a renda média *per-capta* de Manaus, não.

O número do CNPJ também um atributo nominal, já que corresponde a um identificador.

Atributos Ordinais

Atributos ordinais são atributos qualitativos cujos valores apresentam uma ordem ou um ranking. Assim, é possível estabelecer que determinado valor ocupa uma posição anterior ou posterior a outro, embora a distância entre essas posições não seja necessariamente mensurável.

Eles podem ser representados por expressões textuais, como "primeiro", "segundo" e "terceiro", ou por números, como 1, 2 e 3. Nesse último caso, os números representam posições ou categorias ordenadas, e não necessariamente quantidades.

Atributo Qualitativos

- **Nominais** - valores representam categorias ou rótulos. Podem ser textuais ou numéricos.
- **Binários** - os valores possíveis são apenas dois distintos: 0 ou 1, Sim ou Não, etc.
- **Ordinais** - valor indica uma ordem ou *ranking* atribuída à observação.

Inédita - 2026 - Ultra TI

Acerca da classificação dos atributos, julgue o item subsequente.

Por ser composto exclusivamente de algarismos, o CNPJ deve ser classificado como atributo quantitativo, de modo que operações aritméticas efetuadas sobre seus valores possuem significado analítico.

Comentários:

Regra geral, atributos utilizados exclusivamente para identificar um evento, um objeto ou uma entidade — como o CNPJ, que identifica uma entidade cadastrada — são classificados como **qualitativos nominais**, ainda que sejam representados apenas por algarismos. Isso porque, do ponto de vista analítico, seus valores funcionam como identificadores: operações aritméticas entre identificadores não produzem resultados significativos. Afinal, somar diferentes números de CNPJ não costuma fornecer qualquer informação analítica útil.

Isto, contudo, não impede a contagem de CNPJs distintos numa base, para extrair a quantidade de empresas existentes. Notem que, neste caso, contamos apenas as ocorrências individuais do CNPJ, mas não realizamos qualquer operação matemática sobre seus algarismos.

Gabarito: Errada

4.2.2 Atributos quantitativos

Um atributo é quantitativo porque ele é mensurável, representado por um inteiro ou um número real.

Atributos Discretos

Atributos quantitativos discretos são geralmente representados por números inteiros associados à contagem. Quantidade de itens numa nota fiscal, por exemplo, é um atributo discreto.



Atenção!

Em mineração de dados, “discreto” pode vir a ser empregado em sentido mais amplo para **qualquer atributo com domínio finito ou enumerável**. Nesse uso, atributos categóricos também podem ser considerados discretos.

Atributos Contínuos

Um atributo é contínuo quando ele pode assumir qualquer valor num determinado intervalo, com a possibilidade, inclusive, de o intervalo ser infinito. Exemplos de atributos contínuos são peso, altura e temperatura, onde medimos também os decimais.

Atributos Intervalares

Um atributo intervalar é um valor numérico que possui **ordem clara e distâncias iguais e constantes** entre seus valores, mas que **não possuem ponto zero absoluto, isto é, zero não representa ausência total**. Por isso, ainda que a distância entre eles tenham significado, suas razões não tem.

O exemplo típico é o calendário. O ano zero não representa ausência total de tempo, mas apenas uma marcação histórica: a terra no ano 2.000 DC não tem o "dobro de existência" do que tinha no ano 1.000 DC.

Atributos Razão

Um atributo razão é um valor numérico contínuo ou discreto que possui **ponto zero absoluto**, onde o **zero representa a ausência total**, permitindo todas as operações matemáticas, como peso, altura e idade.

Vejamos alguns exemplos de atributos e suas classificações:

Atributo	Tipo
Quantidade de itens	Quantitativo, discreto e razão
Número da nota	Qualitativo e nominal
Temperatura em Celsius	Quantitativo, contínuo e intervalar
Valor da operação	Quantitativo, contínuo e razão
CNPJ	Qualitativo e nominal

Atributos Quantitativos

- **Discretos** - geralmente representados por números inteiros associados à contagem.
- **Contínuos** - podem assumir qualquer valor num determinado intervalo.
- **Intervalos** - valor numérico que possui ordem clara e distâncias iguais e constantes entre seus valores, mas que **não possui ponto zero absoluto**. Ex: temperatura em Celsius.
- **Razão** - valor numérico que possui **ponto zero absoluto**. Ex: peso ou altura.

Inédita - 2026 - Ultra TI

Julgue o item seguinte, relativo à classificação de atributos quantitativos.

Como os anos do calendário apresentam ordem e intervalos regulares, eles constituem uma escala de razão, sendo correto afirmar, por exemplo, que o ano 2000 representa o dobro do ano 1000.

Comentários:

O calendário conta a passagem do tempo, e por convenção o ano zero está fixado no ano de nascimento de Jesus Cristo. De fato, os anos do calendário possuem ordem e intervalos regulares, mas ele não possui ponto zero absoluto, por isso, não estão na escala razão, mas intervalar.

Faz sentido afirmar que entre os anos 1000 e 2000 existe um intervalo de mil anos. Contudo, não é correto dizer que o ano 2000 representa o dobro do ano 1000, pois comparações de proporção exigem um zero absoluto.

Gabarito: Errada

4.3 Dimensões, medidas e métricas

Depois de identificarmos os atributos existentes em um conjunto de dados, precisamos entender o papel que cada um deles exerce na análise. Vamos recuperar a tabela anterior:

data	nr_nfe	cnpj	produto	quantidade	total	icms
20/03/26	1001	12345678912345	notebook	1	8.000	1.440
20/03/26	1001	12345678912345	mouse	2	200	38

Na acima, temos atributos que ajudam a **contextualizar a operação realizada**, como data, CNPJ e produto. Também temos atributos que **quantificam essa operação**, como quantidade, valor total e ICMS.

4.3.1 Dimensões

Dimensões são atributos utilizados para **contextualizar, organizar, filtrar ou agrupar os dados durante uma análise**. Normalmente, respondem a perguntas como "quem", "quando", "qual" ou "onde".

No nosso exemplo anterior, as dimensões são:

- **data**: dimensão temporal;
- **cnpj**: dimensão relacionada à empresa/contribuinte;
- **produto**: dimensão relacionada ao item comercializado.

Por meio delas, poderíamos investigar questões como:

- quanto foi vendido em cada data?
- qual foi o ICMS associado a cada produto?
- como as vendas se distribuem entre os dias da semana?

Dimensões

Dimensões são atributos que fornecem o **contexto da análise**. Elas são utilizadas para filtrar, segmentar, organizar e agrupar as observações.

4.3.2 Medidas e Métricas

Uma **medida** é um valor quantitativo associado **diretamente a uma observação**.

Uma **métrica**, por sua vez, é uma **quantificação utilizada para analisar determinado fenômeno, frequentemente obtida por meio de agregações ou cálculos sobre as observações**.

Na tabela de itens vendidos, funcionam como **medidas** os atributos **quantidade**, **total** e **icms**. A partir da tabela, podemos derivar algumas **métricas**:

Métrica	Calculo	Resultado
Qtd total de itens	Soma dos itens vendidos	3
Valor total vendido	Soma do total	8200
ICMS total	Soma do ICMS	1478
Quantidade de notas	Contagem distinta de notas	1
Alíquota efetiva	ICMS total ÷ valor total	0,18 (18%)

Por fim, algumas referências e ferramentas utilizam **medida** e **métrica** como sinônimos.

Medidas e Métricas

A **medida** é um valor quantitativo associado à observação. A **métrica** é uma quantificação utilizada na análise, obtida diretamente de uma medida ou por meio de agregação e cálculo.

4.4 Questões para exercitar

82. (Inédita - 2026 - Ultra TI -)

Acerca dos conceitos de observação e conjunto de dados, julgue o item a seguir.

Em um conjunto de dados tabular, cada linha pode representar uma observação, isto é, um evento, um objeto ou um experimento medido no mundo real, enquanto as colunas representam atributos dessas observações.

Comentários:

É exatamente isto: uma observação ou *exemplo* representa um evento, objeto ou experimento medido no mundo real, e as colunas são os atributos dessas observações.

Numa tabela com registro de transações de vendas, cada observação corresponderá a uma transação de venda; numa tabela com informações com informação de veículos, cada observação corresponderá a um veículo e seus atributos, entre outros exemplos. Por isso, afirmativa correta.

Gabarito: Certa

83. (Inédita - 2026 - Ultra TI -)

No que se refere à granularidade dos dados, julgue o item subsequente.

Ao substituir os registros individualizados dos itens de notas fiscais pelo total vendido em cada mês, aumenta-se a granularidade do conjunto de dados, pois cada registro mensal passa a condensar maior quantidade de vendas.

Comentários:

Veja que a proposta da afirmativa é sair de um registro individualizado dos itens de notas fiscais, que contém informações detalhadas acerca do que foi vendido, para uma visão agregada do total vendido por mês. Esta transição **reduz** a granularidade do conjunto de dados, pois perde-se o detalhe dos itens em favor de uma totalização do que foi vendido. Logo, incorreta.

Gabarito: Errada

84. (Inédita - 2026 - Ultra TI -)

Julgue o próximo item, relativo a observações e granularidade.

A granularidade define precisamente o que cada registro representa; por isso, dois conjuntos formados a partir das mesmas vendas podem apresentar granularidades distintas se um registrar cada item vendido e o outro registrar apenas totais mensais.

Comentários:

A granularidade define o grau de detalhe do registro. O que é mais detalhado é mais granular. Agregações e totalizações, por outro lado, por terem menos detalhes individuais dos elementos, são menos granulares. Por isso, afirmativa correta.

Gabarito: Certa

85. (Inédita - 2026 - Ultra TI -)

Com relação aos conceitos de atributo e variável, julgue o item a seguir.

Embora atributo e variável sejam frequentemente tratados como sinônimos, pode-se considerar, conceitualmente, que o atributo descreve formalmente os dados, ao passo que a variável corresponde a esse atributo no contexto de uma análise.

Comentários:

Geralmente variáveis e atributos utilizados de forma intercambiável, como sinônimos. Mas, conceitualmente, eles significam coisas diferentes: o atributo é uma descrição do dado, a variável é o atributo utilizado no contexto de uma análise de dados.

Gabarito: Certa

86. (Inédita - 2026 - Ultra TI -)

Acerca da classificação dos atributos, julgue o item subsequente.

Por ser composto exclusivamente de algarismos, o CNPJ deve ser classificado como atributo quantitativo, de modo que operações aritméticas efetuadas sobre seus valores possuem significado analítico.

Comentários:

Regra geral, atributos utilizados exclusivamente para identificar um evento, um objeto ou uma entidade — como o CNPJ, que identifica uma entidade cadastrada — são classificados como **qualitativos nominais**, ainda que sejam representados apenas por algarismos. Isso porque, do ponto de vista analítico, seus valores funcionam como identificadores: operações aritméticas entre identificadores não produzem resultados significativos. Afinal, somar diferentes números de CNPJ não costuma fornecer qualquer informação analítica útil.

Isto, contudo, não impede a contagem de CNPJs distintos numa base, para extrair a quantidade de empresas existentes. Notem que, neste caso, contamos apenas as ocorrências individuais do CNPJ, mas não realizamos qualquer operação matemática sobre seus algarismos.

Gabarito: Errada

87. (Inédita - 2026 - Ultra TI -)

Julgue o item seguinte, a respeito dos atributos qualitativos.

Um atributo binário somente pode ser considerado qualitativo quando seus dois valores forem representados por textos, como “sim” e “não”; se forem codificados como 0 e 1, o atributo será necessariamente quantitativo.

Comentários:

Novamente, a representação numérica de um atributo (binário ou não) não é determinante para classificá-lo, por si só, como quantitativo. Há de se analisar o significado do dado.

No enunciado, se um atributo binário tiver o significado de "sim" e "não", então ele é um **dado qualitativo nominal**, qualquer que seja sua representação — numérica ou textual. Por isso, afirmativa incorreta.

Gabarito: Errada

88. (Inédita - 2026 - Ultra TI -)

No que se refere aos tipos de atributos, julgue o próximo item.

Uma classificação de atendimento formada pelas categorias "baixo", "médio" e "alto" constitui atributo qualitativo ordinal, pois as categorias apresentam uma ordem, ainda que não expressem uma mensuração quantitativa.

Comentários:

Atributos ordinais podem ser representados por categorias textuais ou numéricas, desde que expressem a ideia de ordem entre as categorias. As categorias "baixo", "médio" e "alto" claramente passam a ideia de alguma ordenação relativa, por isso podem ser considerados atributos qualitativos ordinais. Afirmativa correta.

Gabarito: Certa

89. (Inédita - 2026 - Ultra TI -)

Com relação aos atributos quantitativos, julgue o item a seguir.

A quantidade de itens de uma nota fiscal pode ser classificada simultaneamente como atributo quantitativo discreto e atributo de razão, uma vez que decorre de contagem e possui zero absoluto.

Comentários:

Quantidade de itens em uma nota fiscal são normalmente contados como número inteiros: 1, 2, 3... Por isso, são classificadores como discreto.

Atributos razão são aqueles que possuem **ponto zero absoluto**, em que o 0 representa a **ausência de itens**.

Logo, afirmativa correta.

Gabarito: Certa

90. (Inédita - 2026 - Ultra TI -)

Acerca das escalas intervalar e de razão, julgue o item subsequente.

A temperatura em graus Celsius é um exemplo de atributo intervalar: diferenças entre valores possuem significado, mas o zero da escala não representa ausência total de temperatura, razão

por que comparações de proporção não são adequadas.

Comentários:

Atributos intervalares são aqueles que **não possuem ponto zero absoluto**, ainda que a diferença entre dois valores seja significativo, como a diferença entre 10 °C e 20 °C, e exista ordem e intervalos iguais

Por outro lado, a ausência de zero absoluto implica que comparações de proporções não sejam, mesmo adequadas. Se uma medição acusa 100 °C, e outra acusa 50 °C, não podemos afirmar que a segunda tem a "metade" da temperatura da primeira.

Por ausência de ponto zero absoluto podemos entender que, na medição da temperatura em graus Celsius, 0 °C não representa ausência total: trata-se apenas de um ponto na escala.

Gabarito: Certa

91. (Inédita - 2026 - Ultra TI -)

Julgue o item seguinte, relativo à classificação de atributos quantitativos.

Como os anos do calendário apresentam ordem e intervalos regulares, eles constituem uma escala de razão, sendo correto afirmar, por exemplo, que o ano 2000 representa o dobro do ano 1000.

Comentários:

O calendário conta a passagem do tempo, e por convenção o ano zero está fixado no ano de nascimento de Jesus Cristo. De fato, os anos do calendário possuem ordem e intervalos regulares, mas ele não possui ponto zero absoluto, por isso, não estão na escala razão, mas intervalar.

Faz sentido afirmar que entre os anos 1000 e 2000 existe um intervalo de mil anos. Contudo, não é correto dizer que o ano 2000 representa o dobro do ano 1000, pois comparações de proporção exigem um zero absoluto.

Gabarito: Errada

92. (Inédita - 2026 - Ultra TI -)

No que se refere a dimensões, medidas e métricas, julgue o próximo item.

Em uma tabela na qual cada observação representa um item vendido, a quantidade e o valor daquele item são medidas; já a soma dos valores de diversos itens para apurar o total vendido constitui uma métrica.

Comentários:

Em uma tabela cuja granularidade corresponde ao item vendido, cada item representa uma observação descrita por múltiplos atributos. Alguns desses atributos podem ser quantitativos e registrar medidas associadas diretamente à observação, como quantidade e valor.

Uma métrica, por sua vez, é uma quantificação utilizada na análise. Ela pode ser obtida diretamente de uma medida ou derivada por meio de agregações e cálculos, como somas, médias, contagens e razões. O valor total vendido e a quantidade total de itens são exemplos de métricas.

Gabarito: Certa

93. (Inédita - 2026 - Ultra TI -)

Acerca do papel dos atributos em uma análise de dados, julgue o item a seguir.

Dimensões são necessariamente atributos textuais, pois atributos representados por números somente podem desempenhar o papel de medidas.

Comentários:

Lembrem-se de que, em análise de dados, dimensões são atributos utilizados para contextualizar, organizar, agrupar ou filtrar os dados. Frequentemente, são atributos qualitativos ou descritivos, que podem ser representados tanto por textos quanto por números, como ocorre com códigos e identificadores. Por isso, a afirmativa está errada ao restringir o papel de dimensão aos atributos textuais e considerar que todo atributo numérico deve funcionar como medida.

Gabarito: Errada

94. (Inédita - 2026 - Ultra TI -)

Julgue o item subsequente, relativo a vetores de atributos.

Para que diferentes observações sejam representadas de forma coerente por vetores de atributos, deve-se manter a mesma ordem dos atributos em todos os vetores, embora essa ordem não precise coincidir com a disposição original das colunas de uma tabela.

Comentários:

Isso mesmo. Cada observação é representada por um vetor, e cada posição desse vetor contém o valor de determinado atributo. É fundamental que a ordem dos atributos seja consistente entre os diferentes vetores, de modo que uma **mesma posição sempre corresponda ao mesmo atributo**. Podemos imaginar cada posição do vetor como uma coluna da tabela.

Essa ordem não precisa coincidir com a disposição original das colunas; basta que seja previamente estabelecida e mantida em todos os vetores.

Gabarito: Certa

95. (Inédita - 2026 - Ultra TI -)

Com relação aos conceitos de medida e métrica, julgue o próximo item.

Para ser considerada métrica, uma quantificação deve obrigatoriamente resultar da agregação de várias observações, não podendo ser obtida diretamente de uma medida associada a uma única

observação.

Comentários:

A agregação de várias observações é uma forma comum de obter métricas — por exemplo, valor total vendido ou média de itens por nota —, mas não é obrigatória.

Uma métrica também pode ser obtida diretamente de uma medida associada a uma única observação. Assim, o valor de R\$ 8.000 de um item vendido pode ser usado como uma métrica de análise, mesmo sem agregação. Portanto, **não é obrigatório** realizar agregações.

Gabarito: Errada

5 Ciclo de vida dos dados e da informação

5.1 Ciclo de Vida dos Dados (CVD)

O estudo do Ciclo de Vida dos Dados ganhou muita relevância nos últimos anos por conta, principalmente, do aumento exponencial dos dados disponíveis no mundo digital. Em um contexto marcado por dados massivos, diversos e produzidos em fluxos enormes e contínuos, o CVD propõe uma análise estruturada, delimitando o ciclo de vida em fases e ressaltando os diferentes desafios e particularidades de cada uma delas.

Para estudo do ciclo de Vida dos Dados, utilizaremos como base o artigo "Ciclo de vida dos dados: uma perspectiva a partir da Ciência da Informação" de Sant'Ana (2016), referência no Brasil.

A fase inicial do estudo do Ciclo de Vida dos Dados corresponde à **fase de coleta**. Esta fase representa a necessidade de **obtenção inicial de dados** para um propósito específico. Aqui, devem ser identificados quais dados são necessários, onde estão disponíveis e como obtê-los. Também se avalia se os dados precisam algum tratamento. Além disso, devem ser considerados aspectos legais e éticos, como privacidade e permissão de coleta, bem como elementos que permitam sua identificação para uso futuro.

A segunda fase corresponde à **fase de armazenamento** dos dados, que envolve a **persistência**. Esta fase preocupa-se em como os dados serão efetivamente guardados. Também se ocupa com a definição da estrutura física e lógica para persistência, de como devem ser preservados os metadados e demais informações sobre a coleta, além da avaliação de risco à privacidade, direito de armazenamento, condições de uso e fatores que permitam sua utilização ao longo do tempo.

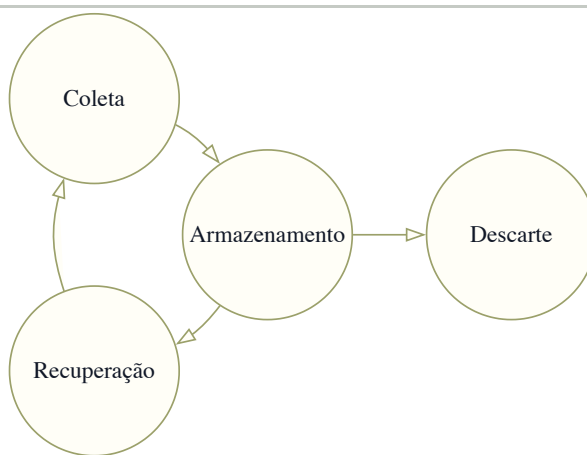
A terceira fase é a **fase de recuperação** dos dados. Nesta fase, o dado armazenado será **acessado para utilização**, por isso, a organização deve preocupar-se com a definição de quais dados armazenados serão oferecidos, para qual público-alvo; a forma de acesso aos dados; requisitos quanto à frequência de atualização, controle de acesso e privacidade; avaliação de direito ou autorização para disponibilização dos dados, bem como assegurar que os dados possam ser encontrados e interpretados.

Por fim, a quarta fase refere-se à **fase de descarte** do dado. Aqui, o foco é **eliminação dos dados** que perderam utilidade. Algumas atividades que orientam o descarte, são:

- O dados ainda são úteis?
- O dado pode ser excluído, ou existe alguma obrigação normativa de mantê-los?
- A remoção do dado pode prejudicar a qualidade do conjunto de dados como um todo?

Assim, podemos esquematizar as fases da seguinte maneira:

Ciclo de Vida dos Dados (CVD)



5.2 Ciclo de Vida da Informação (CVI)

A informação não permanece parada dentro de uma organização. Ela nasce em algum ponto, é utilizada por pessoas ou sistemas, pode ser armazenada para consultas futuras, circular entre áreas ou instituições e, em algum momento, deixar de ser necessária.

É essa trajetória que chamamos de **Ciclo de Vida da Informação (CVI)**, citado de modo expresseo em normativos públicos de segurança da informação, como a Portaria GM/MS nº 3.207, de 20 de outubro de 2010, do Ministério da Saúde.

Em termos simples, o Ciclo de Vida da Informação descreve as fases pelas quais a informação passa, desde sua produção até seu descarte. Esse enfoque é importante principalmente do ponto de vista da segurança. Nas organizações, a informação precisa ser protegida e administrada durante todo o seu ciclo de vida, e não apenas quando está armazenada em um banco de dados ou em um arquivo.



Pra exemplificar...

Imagine, por exemplo, um relatório de arrecadação tributária. Ele começa a existir quando informações são produzidas a partir de dados de sistemas fiscais. Depois, esse relatório é manuseado por servidores, analisado por gestores, armazenado em repositórios institucionais, compartilhado com áreas autorizadas e, ao final, pode ser arquivado ou descartado conforme as regras aplicáveis.

O ciclo de vida da informação olha justamente para essa trajetória e suas particularidades: do "nascimento" ao "fim" da informação.

É importante notar que a informação **se movimenta**. Ela possui as seguintes fases de vida:

Na fase de **produção**, a informação é **criada ou gerada**. Isso pode ocorrer a partir de dados coletados em sistemas ou documentos recebidos. É o momento em que a informação passa a existir para a organização.

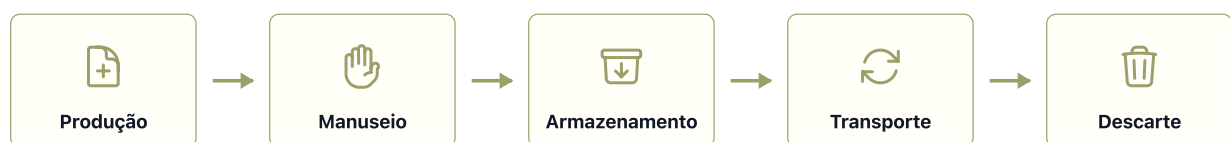
Na fase de **manuseio**, a informação é **utilizada, consultada, alterada, analisada ou manipulada** durante as atividades organizacionais.

Na fase de **armazenamento**, a informação é **guardada para uso posterior**. Esse armazenamento pode ocorrer em meio físico, como documentos impressos, ou em meio digital.

Na fase de **transporte**, a **informação circula**. Ela pode ser enviada por e-mail, transmitida entre sistemas, compartilhada em uma plataforma, encaminhada em processo administrativo, comunicada por telefone ou transportada em mídia física.

Por fim, na fase de **descarte**, a informação é **eliminada quando deixa de ser necessária ou quando se encerra seu prazo de guarda**. Mesmo esta fase representa riscos de segurança que não podem ser negligenciados.

Ciclo de Vida da Informação (CVI)



5.3 Comparativo entre CVD x CVI

CVD x CVI

- **Ciclo de Vida dos Dados** enfatiza o fluxo dos dados como registros em sistemas e repositórios
- **Ciclo de Vida da informação** ressalta a trajetória da informação dentro da organização

5.4 Questões para exercitar

96. (Cebraspe - 2024 - ERAC/Qualquer Área de Formação - ANAC)



A respeito de dado, informação, conhecimento e inteligência, julgue o item a seguir.

Na fase de armazenamento do ciclo de vida dos dados, o foco principal é prover meios que ampliem os níveis de utilização desses dados, seja por ampliação das possibilidades de acesso via cópia, por obtenção de conjuntos para análise, ou por meio da disponibilização de recursos de visualização desses dados.

Comentários:

Na fase de **armazenamento** de dados, o **foco principal é persistência dos dados**. Esta fase preocupa-se em como os dados serão efetivamente guardados. Também se ocupa com a definição da estrutura física e lógica para persistência, de como devem ser preservados os metadados e demais informações sobre a coleta, além da avaliação de risco à privacidade, direito de armazenamento, condições de uso e fatores que permitam sua utilização ao longo do tempo.

A definição da assertiva está mais relacionada à fase de **recuperação** dos dados, que se preocupa em disponibilizar os dados para uso.

Desta forma, afirmativa incorreta.

Gabarito: Errada

97. (Cebraspe - 2020 - ProTI/Ciência de Dados - ME)



A respeito de dados, informação, conhecimento e inteligência, julgue o próximo item.

O ciclo de vida da informação compreende as fases de produção, manuseio, armazenamento, transporte e descarte.

Comentários:

Conforme esquema visto em aula, o ciclo de vida da informação compõe:



Logo, afirmativa correta.

Gabarito: Certa

98. (Cebraspe - 2014 - AJ/Análise de Sistemas/Segurança da Informação - TJ SE)



Acerca de gestão, risco e conformidade, julgue o item a seguir.

O ciclo de vida da informação, inclui as etapas de manuseio, de armazenamento, de transporte e de uso, mas não a de descarte, uma vez que esta, mesmo no caso de ser descarte inadequado, não é fonte de vulnerabilidade.

Comentários:

Incorreta. A etapa de descarte **faz parte do ciclo de vida da informação**. Nesta etapa, a informação é eliminada quando deixa de ser necessária ou quando se encerra seu prazo de guarda. E mesmo no descarte devem ser adotadas medidas de segurança para que não resulte em descarte inadequado, expondo a organização a vulnerabilidades.

Gabarito: Errada

99. (FAIBC - 2012 - Ana/Gestão de Sistemas - CM Nova Iguaçu)



Levando em consideração o ciclo de vida da informação marque o item em que caracteriza-se o fim deste ciclo.

- A) Manuseio.
- B) Armazenamento.
- C) Transporte.
- D) Descarte.
- E) Criação.

Comentários:

O fim do ciclo de vida da informação é caracterizado pelo **descarte**.

Gabarito: D

100. (FCC - 2026 - GG POG - SEPLAG PE)



Uma equipe de Tecnologia da Informação inicia um projeto para aprimorar o processo de planejamento orçamentário. Nessa etapa, são reunidos documentos em PDF, mensagens de e-mail e imagens de notas fiscais digitalizadas, provenientes de diversas unidades administrativas, com o objetivo de consolidar essas informações em um sistema de apoio à decisão.

Essa situação caracteriza-se por uma atuação sobre dados

- A) estruturados na etapa de coleta.
- B) estruturados na etapa de tratamento.
- C) estruturados na etapa de armazenamento.
- D) não estruturados na etapa de coleta.
- E) não estruturados na etapa de recuperação.

Comentários:

Todos os documentos citados no enunciado referem-se a **dados não estruturados**, sem esquema definido, de modo que possuem forma livre: o PDF, as mensagens de e-mail e imagens de notas fiscais digitalizadas.

Ademais, a reunião de documentos para consolidação em um sistema de apoio à decisão se refere à **etapa de coleta** do ciclo de vida dos dados. Desta forma, a única alternativa correta é a letra D.

Gabarito: D

101. (FCC - 2018 - TJ/Adm/Agente Jud Policial - TRT 15)



O técnico de segurança deve conhecer o ciclo de vida da informação, que é composto por

- A) produção, proteção, contenção e arquivamento.
- B) criação, contenção, divulgação e arquivamento.
- C) manuseio, armazenamento, transporte e descarte.
- D) manuseio, armazenamento, confinamento e controle.
- E) criação, manuseio, confinamento e descarte.

Comentários:

A FCC considerou por correta a alternativa C. Ela não está estritamente correta, pois omite a primeira fase do ciclo de vida da informação: a **criação**. No entanto, diante das demais alternativas, a **mais correta** é realmente a C.

A) A proteção, contenção e o arquivamento não representam etapas do ciclo de vida da informação.

B) A contenção, divulgação e o arquivamento não representam etapas do ciclo de vida da informação.

C) Faltou referenciar a fase de **criação**. Mas aceita como correta pela banca.

D) Confinamento e controle não representam etapas do ciclo de vida da informação.

E) Confinamento não representa uma etapa do ciclo de vida da informação.

Gabarito: C

6 Questões da aula

1. (Cebraspe - 2025 - Tec / Tecnologia da Informação - FUB)



Julgue o item a seguir, relativo aos conceitos de administração de dados.

Informação é uma unidade básica de dado que pode ser medida, registrada ou armazenada sem que esteja associada a um significado ou conhecimento, ao passo que dado é uma informação dotada de significado que pode ser usada para a geração de conhecimento ou a tomada de decisões.

2. (Cebraspe - 2025 - APC / Infra e Cibersegurança - FUNPRESP-EXE)



Julgue o seguinte item, relativo a banco de dados.

Considere que, em uma análise sobre a sustentabilidade de um sistema de previdência social, sejam cadastrados dados acerca da quantidade de contribuições mensais de servidores, a idade de aposentadoria e o valor das aposentadorias concedidas. Considere, ainda, que esses dados, por si só, sejam números isolados e sem significado direto. Considere, por fim, que quando esses dados são analisados e comparados com projeções de crescimento populacional e expectativa de vida, eles forneçam informações úteis para avaliar a viabilidade do sistema. Com base nessa situação hipotética, é correto afirmar que dados são fatos brutos e sem contexto, enquanto informação é o resultado da análise desses dados, tornando-os úteis para a tomada de decisões.

3. (Cebraspe - 2024 - ACE / Esp. Arquivologia - TC DF)



Julgue o próximo item, relativos a dado, informação, conhecimento e inteligência.

Informação é a capacidade de compreender, interpretar e aplicar dados de forma estratégica.

4. (Cebraspe - 2023 - Ana TI - FUB)



A respeito de dados e informações, julgue o item que se segue.

Quando ocorre o tratamento de um ou mais dados, é gerada a informação, a qual constitui conhecimento quando aplicada em determinado contexto.

5. (Cebraspe - 2022 - ADP / Redes e Comunicação de Dados - DPE RO)



Em uma situação hipotética, o fato de pessoas atuarem de forma diferenciada para resolver um mesmo problema é explicado pelo conceito de

- A) informação.
- B) comportamento.

- C) dado.
- D) conhecimento.
- E) inteligência.

6. (Cebraspe - 2022 - Ag Inv - PC PB)



Na teoria da informação, os insumos para a tomada de decisão na organização são apresentados na fase de

- A) geração da informação.
- B) tomada de decisão.
- C) provisão de conhecimento.
- D) acompanhamento de resultados.
- E) aquisição de dados.

7. (Cebraspe - 2022 - PPNS / Ana. Sistemas / Eng. de Software - PETROBRAS)



Julgue o item a seguir, quanto a conceitos de dado, informação, inteligência e business intelligence (BI).

A inteligência é a informação que é expressa na forma de números, letras ou símbolos, os quais são os insumos dos gráficos.

8. (Cebraspe - 2022 - Aud Est / Tecnologia da Informação - SECONT ES)



Com relação a dado, informação e conhecimento, julgue o item subsecutivo.

O conhecimento é obtido a partir de um conjunto de informações inseridas em um mesmo e específico contexto.

9. (Cebraspe - 2022 - Aud Est / TI - SECONT ES)



Com relação a dado, informação e conhecimento, julgue o item subsecutivo.

Os termos CÉU e AZUL, quando utilizados separadamente, representam dados, enquanto a expressão CÉU É AZUL representa uma informação.

10. (Cebraspe - 2022 - ACE / TI - TCE RJ)



Julgue o item a seguir, referente à informação e a dados, estruturados e não estruturados.

A informação normalmente envolve a participação humana na sua estruturação e na busca do seu significado, enquanto os dados são mais facilmente obtidos e quantificados por máquinas.

11. (Cebraspe - 2022 - AFCE / Tecnologia da Informação - TCE SC)



Em relação a conceitos de informação e conhecimento e de banco de dados relacionais, julgue o item subsequente.

Informação pode ser definida como um conjunto de fatos diferentes relativos a uma transação e que é, geralmente, desprovida de significado; ela é considerada a matéria-prima dos sistemas de informação.

12. (Cebraspe - 2022 - AFCE / Tecnologia da Informação - TCE SC)



Em relação a conceitos de informação e conhecimento e de banco de dados relacionais, julgue o item subsequente.

O conhecimento serve para dar suporte ao processamento decisório; ele representa a informação tratada, confiável e íntegra.

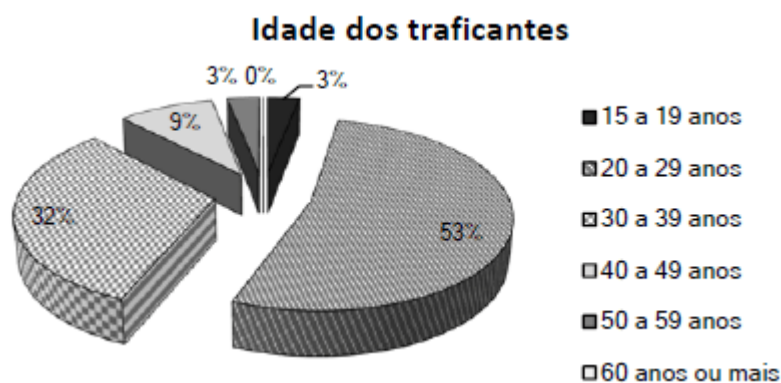
13. (Cebraspe - 2021 - Ag Pol - PC DF)



Sobre a teoria da informação, julgue o item a seguir.

Embora não haja conceitos únicos para dados e informação, é possível afirmar que os dados são matéria-prima para a informação.

14. (Cebraspe - 2021 - APF - PF)



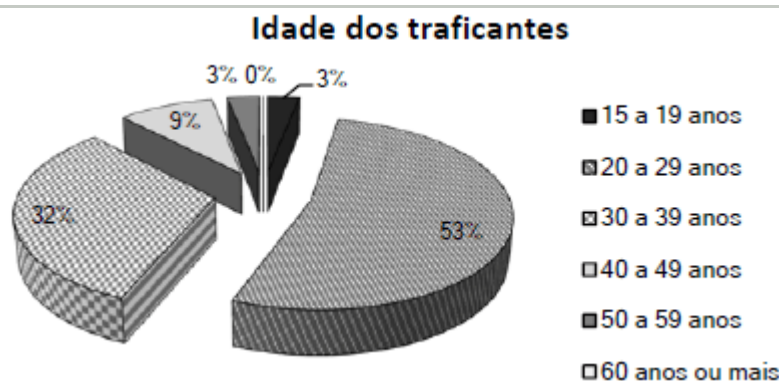
Considere que a Polícia Federal tenha registrado, em determinado período, a prisão de 1.789 traficantes de drogas pertencentes a facções criminosas, conforme faixas etárias mostradas no gráfico.

Com referência às informações e ao gráfico precedentes, julgue o item subsequente.

Considerando-se a classificação dados, informação, conhecimento e inteligência, é correto afirmar que o gráfico representa, por si só, a inteligência.

15. (Cebraspe - 2021 - APF - PF)





Considere que a Polícia Federal tenha registrado, em determinado período, a prisão de 1.789 traficantes de drogas pertencentes a facções criminosas, conforme faixas etárias mostradas no gráfico.

Com referência às informações e ao gráfico precedentes, julgue o item subsequente.

O número 1.789 sozinho caracteriza uma informação, independentemente do contexto.

16. (Cebraspe - 2021 - ACE / Controle Externo / Ciências Contábeis - TCE RJ)



Com relação a dado, informação, conhecimento e inteligência, julgue o item que se segue.

Um dado que está inserido em um contexto pode ser denominado de informação.

17. (Cebraspe - 2021 - ACE / Controle Externo / Controle Externo - TCE RJ)



Com relação aos conceitos de análise de dados e informações, julgue o item a seguir.

A informação pode ser facilmente obtida por meio de máquinas, além de ser transferível e frequentemente quantificável.

18. (Cebraspe - 2021 - ACE / Controle Externo / TI - TCE RJ)



Com relação aos conceitos de análise de dados e informações, julgue o item a seguir.

O conceito de inteligência está relacionado à capacidade de julgamento, ou seja, o ser humano usa o seu conhecimento para tomar a melhor decisão em uma situação real.

19. (Cebraspe - 2018 - APF - PF)



Julgue o item, a respeito da teoria da informação e de sistemas de informação.

O conceito de conhecimento é mais complexo que o de informação, pois conhecimento pressupõe um processo de compreensão e internalização das informações recebidas, possivelmente combinando-as.

20. (Cebraspe - 2018 - EPF - PF)



Julgue o item a seguir, a respeito da teoria da informação e de metadados de arquivos.

O conhecimento é embasado na inteligência das informações que são coletadas e analisadas para uma organização.

21. (Cebraspe - 2018 - PCF / Área 14 - PF)



Julgue o item, a respeito de computação na nuvem, sistemas de informações e teoria da informação.

A informação se caracteriza pela compreensão e internalização do conteúdo recebido, por meio do seu uso em nossas ações; o dado, por sua vez, é um elemento bruto dotado apenas de significado e relevância que visem fornecer uma solução para determinada situação de decisão.

22. (Cebraspe - 2018 - PPF - PF)



Acerca da definição de dados e informação e sua representação em sistemas de informação, julgue o item que se segue.

Dados são fatos que descrevem os objetos de informação, por exemplo, eventos e entidades.

23. (Cebraspe - 2018 - PPF - PF)



Acerca da definição de dados e informação e sua representação em sistemas de informação, julgue o item que se segue.

Informação é constituída por um conjunto de dados com características específicas. O ponto de análise é que os dados devem ser irrelevantes para o sistema a que se destinam.

24. (Cebraspe - 2018 - Aud Est / Controle Externo - TCM BA)



O diretor de uma montadora de veículos necessita tomar uma decisão acerca da continuidade ou não de um dos produtos vendidos no Brasil. Para tanto, solicitou um relatório sobre as vendas de carros da marca do último trimestre de 2018, por faixa de preço, região, modelo e cor. Nessa situação, no contexto de análise da informação, o relatório representa

- A) conhecimento.
- B) inteligência.
- C) dados.
- D) informação.
- E) sabedoria.

25. (Cebraspe - 2017 - ACE / Auditoria de Contas Públicas - TCE PE)



Acerca dos conceitos de dado, informação e conhecimento, julgue o item a seguir.

A informação caracteriza-se por ser frequentemente tácita, bem como por ser de estruturação e captura difíceis em máquinas.

26. (Cebraspe - 2016 - AFCE / Tecnologia da Informação - TCE SC)



Julgue o item a seguir, acerca de dado, informação, conhecimento e inteligência.

Define-se informação como significado, ou seja, como registros icônicos e simbólicos — fonéticos ou numéricos — e signos — linguísticos, lógicos ou matemáticos —, por meio dos quais se representam atos, conceitos ou instruções.

27. (Cebraspe - 2016 - AFCE / Tecnologia da Informação - TCE SC)



Julgue o item a seguir, acerca de dado, informação, conhecimento e inteligência.

O atributo de inteligência depende mais da qualidade da informação disponível do que da sua quantidade, tendo, portanto, natureza qualitativa.

28. (Cesgranrio - 2023 - Esc BB / Agente de Tecnologia - BB)



No processo de aprendizagem de máquina, busca-se identificar padrões compreensíveis em bases de dados.

Desse modo, supõe-se a existência de uma hierarquia de três níveis: a base, com os itens elementares, captados e armazenados por recursos de Tecnologia da Informação; o nível intermediário, no qual esses itens são processados, com significados e contextos bem definidos; e o nível mais alto, contendo os padrões ou os conjuntos cuja formulação pode envolver ou relacionar os níveis inferiores.

Os três níveis dessa hierarquia, listados, respectivamente, da base para o nível mais alto, são:

- A) conhecimento, informação e dado
- B) dado, conhecimento e informação
- C) dado, informação e conhecimento
- D) informação, conhecimento e dado
- E) informação, dado e conhecimento

29. (COGEPE FIOCRUZ - 2024 - TecnoSP / TI - FIOCRUZ)



Uma das mais importantes atividades da gestão conhecimento nas organizações é a distinção entre dado, informação e conhecimento. Sobre esta distinção é correto afirmar que:

- A) dado é um caso qualquer, desde que seja intangível; informação é a transformação do dado para possibilitar a sua utilização prática; e conhecimento é caracterizado pelo uso da

informação para resolver problemas, sejam eles reais ou não.

- B)** dado é um fato discreto e objetivo sobre um evento; informação é um dado dotado de relevância e propósito; e o conhecimento é uma mistura de experiências, valores, informações contextuais e insights.
- C)** dado é a informação interpretada; conhecimento é a habilidade de processar dados eficientemente para resolver problemas.
- D)** dado é qualquer entrada digital em um computador; informação é a exibição destes dados em formato de relatórios e gráficos; e conhecimento é a interpretação dos relatórios e gráficos, com base na compreensão profunda da informação.
- E)** dado consiste em uma representação visual; informação é uma interpretação que leva ao conhecimento; e conhecimento é a aplicação teórica da informação.

30. (Consulplan - 2024 - Prof / Informação e Comunicação - SEED PR)



O dado é considerado um elemento fundamental para a constituição de sistemas organizados. Segundo o autor Silberschatz, os dados são a representação de fatos do mundo real, que podem ser numéricos, alfanuméricos, imagens, sons entre outros tipos. Eles possuem um papel importante em um sistema de informação, pois são coletados pelos usuários para serem processados dentro de uma lógica previamente estabelecida, gerando o resultado da sua interpretação e organização que é conhecido como:

- A)** Solução.
- B)** Produto.
- C)** Algoritmo.
- D)** Informação.

31. (FCC - 2018 - APE / Administração Pública ou de Empresas - TCE RS)



Os conceitos de dados, informação e conhecimento são de grande importância no contexto de sistemas de informação. Sobre eles, é correto afirmar que

- A)** não são necessários os dados para que se obtenha o conhecimento.
- B)** a informação é obtida acrescentando-se significado aos dados.
- C)** a informação é obtida a partir do conceito de conhecimento.
- D)** o processo de tomada de decisão em um sistema de informação tem por base apenas os dados brutos.
- E)** os dados consistem do conhecimento analisado sob diferentes pontos de vista.

32. (FGV - 2024 - Ana TI / Análise de Negócios de TI - DATAPREV)



Sobre os conceitos de dados, informação, conhecimento e inteligência, é correto afirmar que

- A)** dados são informações já interpretadas, enquanto informação é a simples coleta de fatos brutos sem organização.

- B)** informação é a aplicação do conhecimento para a tomada de decisões eficazes e resolução de problemas.
- C)** conhecimento é a compreensão e internalização da informação, permitindo sua aplicação em diferentes contextos.
- D)** inteligência é a coleta de dados brutos de diversas fontes sem qualquer processamento ou análise.
- E)** dados e inteligência são termos intercambiáveis no contexto de Business Intelligence (BI), representando o mesmo conceito.

33. (FGV - 2024 - APE / Economia de Energia - EPE)



No contexto de ciência de dados, os dados podem ser interpretados como itens elementares, captados e armazenados por recursos de Tecnologia da Informação. Seu propósito é expressar fatos do mundo real de forma a serem tratados no contexto computacional, sendo a base de uma pirâmide hierárquica.

Um nível acima, as informações representam os dados processados, com significado e contextos bem definidos.

Assinale a opção que apresenta corretamente o conceito do último nível dessa pirâmide, correspondendo a um padrão ou conjunto de padrões cuja formulação pode envolver e relacionar fatos e informações.

- A)** Aprendizado.
- B)** Inteligência.
- C)** Conhecimento.
- D)** Durabilidade.
- E)** Ciência.

34. (FGV - 2024 - Aud CE / Ciência de Dados - TCE PA)



Com relação aos conceitos de dado, informação e conhecimento, avalie as afirmativas a seguir.

- I. Os dados são itens elementares, são cadeias de símbolos e não possuem significado.
- II. São exemplos de conhecimento: tendência de vendas de um produto A em uma região B; relação entre o aumento ou queda do preço de uma ação X, na bolsa de valores, e a variação do câmbio.
- III. As informações correspondem ao dado processado, com significado e um contexto indefinido.

Está correto o que se afirma em

- A)** I, apenas.
- B)** I e III, apenas.
- C)** I e II, apenas.

- D) II, apenas.
- E) I, II e III.

35. (IBFC - 2024 - Téc Jud. / Apoio Esp. Desenv. Sistemas - TRF 5)



Leia a frase abaixo.

O ____ é a assimilação e compreensão de informações adquiridas, enquanto a ____ envolve a capacidade de aplicar esse ____ de maneira adaptativa, analítica e inovadora para resolver problemas e tomar decisões eficazes em diferentes contextos.

Assinale a alternativa que preencha correta e respectivamente as lacunas.

- A) inteligência / conhecimento / artefato
- B) conhecimento / inteligência / conhecimento
- C) dado / informação / atributo
- D) conjunto / informação / artefato

36. (Instituto AOCP - 2017 - Sec Exec - UFBA)



Em tempos de concorrência acirrada, as informações, se utilizadas de forma estratégicas, contribuem para o sucesso das organizações. Em relação ao gerenciamento de informações, julgue, como VERDADEIRO ou FALSO, o item a seguir.

O Dado é um elemento isolado, que ainda não passou por contextualização. O cruzamento ou a organização de dados produz uma informação que pode ser utilizada por gestores e, a partir da aplicação da informação, pode ser produzido conhecimento.

37. (NEC UFMA - 2016 - Tec / Tecnologia da Informação - UFMA)



Qual a denominação de um conjunto de dados devidamente ordenados e organizados de forma a terem significado?

- A) Informação
- B) Processamento
- C) Cluster
- D) Memória
- E) Dado

38. (Quadrix - 2025 - Ana Sistemas - CFBIO)



No que diz respeito às ferramentas de automação de processos, aos conceitos de dados e aos índices, julgue o item seguinte.

Dados podem ser definidos como elementos brutos que, quando contextualizados, se tornam informação.

39. (Quadrix - 2024 - Ana TI - COREN PR)



Em relação ao conceito de dado, à Lei Geral de Proteção de Dados (LGPD) e à linguagem de programação Qlik, julgue o item a seguir.

Dado é uma coleção organizada de informações relevantes para uma determinada função ou tarefa.

40. (Quadrix - 2023 - Ana Área / Tecnologia da Informação - CREA GO)



Quanto aos conceitos utilizados no tratamento de dados e na análise de dados e informações, julgue o item.

A capacidade de aprender, adaptar, raciocinar, resolver problemas e aplicar conhecimento de forma eficaz em contextos variados é definida como informação.

41. (Quadrix - 2023 - Ana Área / Tecnologia da Informação - CREA GO)



Quanto aos conceitos utilizados no tratamento de dados e na análise de dados e informações, julgue o item.

Conhecimento pode ser definido como uma representação simbólica ou descritiva de fatos, conceitos ou instruções em vários formatos, como, por exemplo, números, palavras, imagens ou símbolos.

42. (Quadrix - 2023 - Ana Adm / Tecnologia da Informação - CREFITO 14)



É o resultado do processamento de dados, dando-lhes contexto e significado. É quando os dados são organizados, estruturados e relacionados a um determinado contexto, tornando-os compreensíveis.

O texto acima refere-se ao conceito de

- A) informação.
- B) conhecimento.
- C) inteligência.
- D) conhecimento tácito.
- E) conhecimento explícito.

43. (Cebraspe - 2025 - EPF - PF)



Julgue o item subsequente, relativo à teoria da informação e a banco de dados.

No contexto de banco de dados, dados estruturados são aqueles que não possuem um formato fixo e são armazenados em sistemas como bancos NoSQL, ao passo que dados não estruturados

possuem um formato rígido e são armazenados em tabelas relacionais.

44. (Cebraspe - 2025 - AFT - SEFAZ SE)



Dados estruturados são tipos de dados que

- A) não podem ser indexados, mesmo com organização tabular, devendo ser percorridos integralmente em buscas sequenciais.
- B) dependem de metadados embutidos em arquivos binários para permitirem qualquer forma de busca ou filtragem.
- C) podem ser armazenados somente em sistemas de gerenciamento de banco de dados relacionais.
- D) não podem conter valores nulos, pois a não atribuição de valores descaracterizaria a consistência estrutural exigida.
- E) têm formato e organização que, definidos por esquemas rígidos que estabelecem tipos e restrições para cada campo, permitem consultas por meio de linguagens como SQL.

45. (Cebraspe - 2024 - ERSPT - ANATEL)



Julgue o item a seguir, relativo a dados e bancos de dados.

É possível a análise de dados estruturados e não estruturados armazenados no banco de dados, apesar de o formato dos dados não estruturados ser mais diversificado e complexo que o dos dados estruturados.

46. (Cebraspe - 2024 - Ana Sistemas - CAGEPA)



Uma empresa está coletando os dados de campanhas sazonais de publicidade e mantém organizada uma planilha com as informações de cada campanha, com os respectivos períodos de duração e valores investidos.

Nessa situação hipotética, a empresa possui dados do tipo

- A) metadado.
- B) bruto.
- C) não estruturado.
- D) semiestruturado.
- E) estruturado.

47. (Cebraspe - 2024 - Tecno P2 - CTI)



Considerando processos de análise e mineração de dados, julgue o item subsequente.

Os dados utilizados como fonte para análise e data mining são considerados semiestruturados quando possuem uma estrutura não homogênea, como arquivos XML e JSON.

48. (Cebraspe - 2024 - APO - MPO)



Julgue o item a seguir, a respeito de conceitos e fundamentos de dados.

Dados estruturados são aqueles que se encaixam perfeitamente nas tabelas de dados e incluem tipos de dados discretos, como números, textos curtos e datas.

49. (Cebraspe - 2024 - ACE - TC DF)



Julgue o próximo item, relativos a dado, informação, conhecimento e inteligência.

É possível que grandes dados numéricos ou textuais não sejam estruturados em razão da ineficiência em modelá-los como uma tabela.

50. (Cebraspe - 2024 - ATI - TCE AC)



A respeito de Big Data, julgue o próximo item.

Assim como os dados estruturados, os dados não estruturados também podem ser armazenados em bancos de dados relacionais.

51. (Cebraspe - 2024 - TJ TSE - TSE)



Em relação a banco de dados em memória, soluções para Big Data e dados estruturados e não estruturados, julgue o item que se segue.

Documentos de uma empresa e postagens nas redes sociais são exemplos de dados estruturados.

52. (Cebraspe - 2023 - Ana TI - DATAPREV)



A respeito de dados estruturados e não estruturados, julgue o item a seguir.

Arquivos .xml, .csv e .json são exemplos de dados estruturados.

53. (Cebraspe - 2023 - Ana TI - DATAPREV)



A respeito de dados estruturados e não estruturados, julgue o item a seguir.

Se, em certa organização, grande parte dos dados for constituída de dados estruturados, mas também houver muitos dados de arquivos .pdf e .docx, então os dados, em geral, serão classificados como semiestruturados.

54. (Cebraspe - 2023 - APO - SEPLAN RR)



Com base nos conceitos de dados estruturados e não estruturados, de análise exploratória de dados e de dados abertos, julgue o item seguinte.

Dados de redes sociais são exemplos de dados não estruturados.

55. (Cebraspe - 2023 - ACE - TC DF)



A respeito de dados estruturados e dados não estruturados, julgue o item que se segue.

A organização de dados estruturados é totalmente aberta.

56. (Cebraspe - 2023 - ACE - TC DF)



A respeito de dados estruturados e dados não estruturados, julgue o item que se segue.

Dados não estruturados são aqueles que, desde a elaboração da estrutura, são pensados estritamente com uma finalidade.

57. (Cebraspe - 2022 - ADP - DPE RO)



Assinale a opção que mostra um dado não estruturado.

- A) data de nascimento
- B) cadastro de pessoa física (CPF)
- C) nome
- D) biografia
- E) sexo

58. (Cebraspe - 2022 - ADP - DPE RO)



As fontes de dados não estruturados incluem

- A) a planilha eletrônica.
- B) a rede social.
- C) a linguagem de programação.
- D) o arquivo XML.
- E) o banco de dados.

59. (Cebraspe - 2022 - ATCG - MCom)



No que diz respeito aos dados estruturados e não estruturados, aos bancos de dados relacionais, ao modelo de referência CRISP-DM e à modelagem preditiva, julgue o item a seguir.

Dados estruturados são aqueles representados em um formato estrito como, por exemplo, a informação armazenada nos bancos de dados.

60. (Cebraspe - 2022 - ATCG - MCom)



No que diz respeito aos dados estruturados e não estruturados, aos bancos de dados relacionais, ao modelo de referência CRISP-DM e à modelagem preditiva, julgue o item a seguir.

A principal característica dos dados não estruturados é que eles não possuem metadados.

61. (Cebraspe - 2022 - Esc Pol - PC PB)



Em arquivos do tipo texto, os metadados que permitem armazenar informações por meio de estruturas como XML ou JSON são do tipo

- A) estruturados.
- B) semiestruturados.
- C) hierárquicos.
- D) categorizados.
- E) não estruturados.

62. (Cebraspe - 2021 - APF - PF)



No que se refere aos conceitos de dados estruturados e não estruturados, julgue o próximo item.

Os dados estruturados diferenciam-se dos dados não estruturados pela rigidez em seu formato e pelo fato de poderem ser armazenados em campos de tabelas de um banco de dados relacional.

63. (Cebraspe - 2021 - Ana - SERPRO)



Julgue o item a seguir relativos a conceitos de ingestão de dados estruturados, semiestruturados, não estruturados e em streaming.

Os dados não estruturados constam de bancos de dados relacionais que são eminentemente pesquisáveis tanto por meio de consultas realizadas por humanos quanto por meio daquelas realizadas por algoritmos que usam tipos de dados e nomes de campos, como alfabéticos ou numéricos, moeda ou data.

64. (Cebraspe - 2020 - Profissional de TI - ME)



A respeito de dados, informação, conhecimento e inteligência, julgue o próximo item.

Embora com características particulares, dados não estruturados podem ser classificados em sua totalidade, assim como os dados estruturados.

65. (Cebraspe - 2018 - EPF - PF)



- CPF

- NOME
- DATA DE NASCIMENTO
- NOME DO PAI
- NOME DA MAE
- TELEFONE
- CEP
- NUMERO

As informações anteriormente apresentadas correspondem aos campos de uma tabela de um banco de dados, a qual é acessada por mais de um sistema de informação e também por outras tabelas. Esses dados são utilizados para simples cadastros, desde a consulta até sua alteração, e também para prevenção à fraude, por meio de verificação dos dados da tabela e de outros dados em diferentes bases de dados ou outros meios de informação.

Considerando essas informações, julgue o item que se segue.

Os dados armazenados na referida tabela são considerados não estruturados.

66. (Cebraspe - 2018 - ACP - TCE PB)



Com relação a dados estruturados e não estruturados, assinale a opção correta.

- A) Dados não estruturados podem ser caracterizados por arquivos de diversos tipos — textos, imagens, vídeos, entre outros —, cujas estruturas não são descritas implicitamente.
- B) Por padrão, documentos do tipo XML (eXtensible Markup Language) são estruturados.
- C) Dados não estruturados de um mesmo grupo possuem as mesmas descrições e, consequentemente, os mesmos atributos.
- D) Por padrão, dados não estruturados são organizados em blocos semânticos.
- E) A alta heterogeneidade facilita as consultas aos dados não estruturados, desde que estes estejam ligados por ponteiros.

67. (Cesgranrio - 2022 - PNS - ELETRONUCLEAR)



Há dados não estruturados que podem ser categorizados e preparados para uso em ambiente de BI.

Esses dados não estruturados podem ser, por exemplo:

- A) operações de débito e crédito realizadas em uma conta-corrente bancária
- B) postagens e respectivos comentários em uma rede social
- C) quantidades em estoque por produto em um site de comércio eletrônico, ao final do dia
- D) temperatura média de uma cidade, dia a dia
- E) valores de cotação de cada ação negociada em uma bolsa de valores, hora a hora

68. (FCC - 2026 - Ges Gov - SCGE PE)



Durante o processo de análise de grandes volumes de informações, um gestor precisa identificar corretamente o tipo de dado com o qual está trabalhando. Em um mesmo projeto, ele utiliza uma planilha contendo:

- I. arquivos JSON com informações de navegação.
- II. postagens em redes sociais contendo textos e imagens.
- III. registros tabulares de clientes.

Nessa situação, os dados são classificados como

- A) I. semiestruturados; II. estruturados; III. não estruturados.
- B) I. estruturados; II. não estruturados; III. semiestruturados.
- C) I. semiestruturados; II. não estruturados; III. estruturados.
- D) I. estruturados; II. semiestruturados; III. não estruturados.
- E) I. não estruturados; II. estruturados; III. semiestruturados.

69. (FCC - 2026 - FTE - SEFAZ MT)



O Diretor de TI de uma Secretaria da Fazenda precisa modernizar a arquitetura de dados da fiscalização, que atualmente lida com cadastros de contribuintes armazenados em tabelas relacionais, Notas Fiscais Eletrônicas em formato XML e relatórios de fiscalização em PDF e texto livre (linguagem natural), em geral tratados como dados não estruturados para fins arquiteturais e analíticos, podendo demandar técnicas de extração e processamento para uso sistemático. Para garantir que a estratégia de armazenamento e análise seja eficiente e compatível com as características desses diferentes tipos de dados, o analista responsável deve classificá-los e tratá-los considerando que

- A) a infraestrutura fazendária deve focar em dados estruturados, pois qualquer informação textual ou documental de fiscalização pode ser convertida automaticamente em colunas de tabelas sem perda de contexto ou significado para o fisco.
- B) os dados estruturados possuem esquema rígido e organização tabular; os semiestruturados possuem metadados e organização hierárquica que permitem flexibilidade; e os não estruturados carecem de um modelo predefinido, exigindo ferramentas específicas de processamento e análise para coexistirem no ecossistema de fiscalização.
- C) os dados estruturados, como os cadastros de contribuintes, possuem esquema flexível e ausência de organização fixa, sendo ideais para armazenar documentos XML sem necessidade de validação ou transformação para o modelo relacional.
- D) os dados não estruturados possuem esquemas rígidos e tabelas normalizadas previamente definidas, o que permite sua inserção direta em bancos de dados SQL transacionais para consultas imediatas.
- E) os dados semiestruturados, como o XML das Notas Fiscais, são caracterizados pela ausência de metadados e organização hierárquica, devendo ser tratados apenas como objetos binários para evitar a complexidade da extração de campos específicos.

70. (FGV - 2026 - AJ - TJ RJ)



Diferentes sistemas produzem dados em formatos variados, que podem ser classificados em dados estruturados, semiestruturados ou não estruturados.

Nesse contexto, relacione os tipos de dados às suas respectivas descrições.

1. Dados estruturados
2. Dados semiestruturados
3. Dados não estruturados

() Gravações em áudio e vídeo de audiências públicas, armazenadas em arquivos MP4, acompanhadas apenas de nome do arquivo e data de criação.

() Registros de protocolo eletrônico armazenados em tabelas de banco de dados relacional, com campos bem definidos (número do processo, data, unidade, assunto) e chaves primárias/estrangeiras.

() Arquivos de log de acesso ao portal de serviços do governo, registrados em formato JSON, contendo campos como timestamp, user_id, endpoint, status_code, com alguns campos opcionais variando conforme o tipo de requisição.

A sequência correta é:

- A) 1, 2 e 3;
- B) 1, 3 e 2;
- C) 2, 1 e 3;
- D) 3, 1 e 2;
- E) 3, 2 e 1.

71. (FGV - 2024 - Tecno PD - Pref Caragatatuba)



Os dados podem ser classificados de diversas formas, uma das categorias de classificação diz respeito a forma da estruturação dos dados em repositórios de dados, analise os itens a seguir.

I. Dados estruturados são dados que foram formatados e transformados para se adaptar a um modelo de dados relacional bem definido. Os dados brutos são mapeados em campos predefinidos que podem ser extraídos e lidos por meio de declarações SQL.

II. Dados semiestruturados é categoria localizada entre dados estruturados e não estruturados. São um tipo de dado que possui algumas características consistentes e definidas e não se limitam a uma estrutura rígida como a necessária para bancos de dados relacionais.

III. Dados não estruturados podem ser qualquer tipo de dado não esteja em um formato específico. São suportados apenas pelos sistemas de arquivos e não pelos sistemas de bancos de dados, exemplos podem imagens, sons ou desenhos.

Está correto o que se afirma em

- A) I apenas.
- B) III apenas.
- C) I, II e III.
- D) I e II, apenas.
- E) II e III, apenas.

72. (FGV - 2024 - AFTRM - Pref Cuiabá)



Dados podem ser classificados em algumas definições, que podem variar a depender de sistemas, processos e aplicações a que se destinam. Com relação a sua estrutura, podemos classificá-los, dentre outras formas, entre estruturados, semiestruturados e não estruturados.

Os itens a seguir contém alguns exemplos de dados que podem se enquadrar dentro dessas três definições.

I. Um banco de dados relacional.

II. Grafos que representam relações de usuários em uma rede social.

III. Um e-mail, que contém texto, imagens e um campo destinado a informar a data e a hora onde foi enviado.

O(s) item(s) que contém dados do tipo não estruturado é(são):

- A) II, apenas.
- B) I e II, apenas.
- C) II e III apenas.
- D) I, II e III.
- E) III, apenas.

73. (FUNDATEC - 2023 - Ana Sistemas - BRDE)



Assinale a alternativa que contém um exemplo de dados não estruturados.

- A) Dados extraídos de uma rede social de um banco podem ser classificadas em polaridades como positivas ou negativas, ou agrupadas por similaridade de assunto.
- B) Um sistema bancário armazena dados sobre seus clientes, sobre suas respectivas contas bancárias e sobre as transações executadas sobre tais contas.
- C) Funcionários de um departamento de recursos humanos observam os candidatos a uma vaga de emprego e fazem anotações sobre as suas atitudes durante a realização de atividades de um processo de seleção.
- D) Uma empresa deseja lançar um novo produto e, para isso, realiza pesquisas sobre sua aceitação com um conjunto específico de pessoas que representa o público-alvo.

- E) Uma corrida de Fórmula 1 na qual se medem os tempos que cada piloto gasta para chegar em diferentes trechos de um circuito.

74. (Fuvest - 2023 - Ana Sistemas - USP)



Suponha que você precisa analisar posts em uma rede social sobre opiniões de clientes da sua empresa.

O conjunto de textos desses posts são considerados dados

- A) categóricos.
- B) estruturados.
- C) intervalares.
- D) nominais.
- E) não estruturados.

75. (IBFC - 2024 - AC - ECT)



Os dados dentro do sistema de gerenciamento de banco de dados (SGBD) podem estar organizados de três formas: estruturados, não estruturados ou semiestruturados. Assinale a alternativa que apresenta qual elemento podemos encontrar em uma estrutura de dados não estruturados.

- A) Imagens
- B) Tags de um arquivo HTML
- C) Um atributo Nome do tipo Varchar(40)
- D) Arquivo JSON

76. (Instituto Consulplan - 2025 - AJ - TJ RO)



No contexto de análises estatísticas, os dados podem estar disponíveis em diferentes tipos de organização. Dependendo da forma de definição e armazenamento, os conjuntos de dados são tipicamente caracterizados como estruturados, semiestruturados e não estruturados.

Diante do exposto, analise as afirmativas a seguir.

I. Enquanto dados não estruturados são normalmente armazenados em seu formato nativo (como vídeos, imagens ou documentos de texto), dados estruturados são armazenados em linhas e colunas e podem ser mapeados para campos predefinidos.

II. Ao contrário dos dados estruturados, que podem ser organizados e acessados por meio de bancos de dados relacionais, dados não estruturados não têm um modelo rígido de organização predefinido.

III. Conjuntos de dados semiestruturados usam tags e marcadores (ou seja, metadados) em vez de exigir um esquema predefinido em tabelas, linhas e colunas, como nos dados estruturados.

IV. Os data lakes podem ser projetados para armazenar dados semiestruturados juntamente com dados estruturados, permitindo a coexistência de dados de ambos os tipos.

Está correto o que se afirma em

- A) I, II, III e IV.
- B) I e III, apenas.
- C) II e IV, apenas.
- D) I, II e III, apenas.
- E) I, III e IV, apenas.

77. (Instituto Consulplan - 2023 - Ana MP - MPE MG)



Dado é a representação utilizada para gerar uma informação. Um dado pode ser estruturado ou não estruturado. Em relação às diferenças entre dados estruturados e dados não estruturados, marque V para as afirmativas verdadeiras e F para as falsas.

- () Em dados estruturados, dados de um mesmo registro possuem relação entre eles.
- () Textos, documentos, banco de dados, redes sociais, imagens, vídeos e áudios são exemplos de dados não estruturados.
- () Dados estruturados permitem que tipos de dados diferentes das estruturas preestabelecidas sejam carregados.
- () Em dados estruturados, atributos e campos são definidos por um esquema.

A sequência está correta em

- A) F, V, V, F.
- B) V, V, V, V.
- C) V, F, F, V.
- D) F, F, F, F.

78. (Instituto Consulplan - 2022 - Tec - MPE PA)



Dados são um conjunto de fatos em estado bruto que podem ser utilizados para tomadas de decisão. Os dados podem ser divididos entre estruturados e não-estruturados. Sobre os dados estruturados, assinale a afirmativa correta.

- A) Mensagens de e-mail são dados estruturados.
- B) Dados estruturados são organizados em linhas e colunas.
- C) Dados estruturados são armazenados em arquivos no formato ".TXT".
- D) Mecanismos de busca possuem dificuldade para pesquisar dados estruturados.

79. (Quadrix - 2024 - Tec ESI - CREF 9)



A principal vantagem dos dados estruturados em comparação aos dados não estruturados é que eles

- A) são mais difíceis de processar e analisar.
- B) requerem menos armazenamento de dados.
- C) podem ser facilmente organizados, pesquisados e analisados por meio de consultas e algoritmos de banco de dados.
- D) são mais flexíveis e adaptáveis a diferentes tipos de informações.
- E) incluem conteúdo multimídia, como imagens e vídeo.

80. (Vunesp - 2024 - Analista - SPCINE)



É possível estabelecer uma classificação dos dados de uma organização em dados estruturados e dados não estruturados, sobre os quais é correto afirmar que

- A) dados estruturados possuem comprimento e tipo definidos, tais como datas e nomes.
- B) dados de temperatura ao longo dos dias representam tipos de dados não estruturados.
- C) bases de dados orientadas a objetos armazenam exclusivamente dados não estruturados.
- D) dados não estruturados seguem exclusivamente formatos numéricos.
- E) dados de mídias sociais representam tipos de dados estruturados.

81. (Vunesp - 2023 - APPGG - Pref SP)



Um analista de políticas públicas trabalha com diversos formatos e tipos de dados. Identificar corretamente essas características no começo das atividades é essencial para realizar a escolha correta das ferramentas e técnicas que serão utilizadas.

Assinale a alternativa que traz apenas dados semiestruturados.

- A) Arquivo CSV (Comma Separated Values) e arquivos de áudio.
- B) Arquivo JSON (JavaScript Object Notation) e arquivos de imagens.
- C) Arquivo JSON (JavaScript Object Notation) e arquivo XML (eXtensible Markup Language).
- D) Arquivos PDF (Portable Document Format) e arquivos de vídeo.
- E) Tabela de um banco de dados relacional e arquivo XML (eXtensible Markup Language).

82. (Inédita - 2026 - Ultra TI -)

Acerca dos conceitos de observação e conjunto de dados, julgue o item a seguir.

Em um conjunto de dados tabular, cada linha pode representar uma observação, isto é, um evento, um objeto ou um experimento medido no mundo real, enquanto as colunas representam atributos dessas observações.

83. (Inédita - 2026 - Ultra TI -)

No que se refere à granularidade dos dados, julgue o item subsequente.

Ao substituir os registros individualizados dos itens de notas fiscais pelo total vendido em cada mês, aumenta-se a granularidade do conjunto de dados, pois cada registro mensal passa a condensar maior quantidade de vendas.

84. (Inédita - 2026 - Ultra TI -)

Julgue o próximo item, relativo a observações e granularidade.

A granularidade define precisamente o que cada registro representa; por isso, dois conjuntos formados a partir das mesmas vendas podem apresentar granularidades distintas se um registrar cada item vendido e o outro registrar apenas totais mensais.

85. (Inédita - 2026 - Ultra TI -)

Com relação aos conceitos de atributo e variável, julgue o item a seguir.

Embora atributo e variável sejam frequentemente tratados como sinônimos, pode-se considerar, conceitualmente, que o atributo descreve formalmente os dados, ao passo que a variável corresponde a esse atributo no contexto de uma análise.

86. (Inédita - 2026 - Ultra TI -)

Acerca da classificação dos atributos, julgue o item subsequente.

Por ser composto exclusivamente de algarismos, o CNPJ deve ser classificado como atributo quantitativo, de modo que operações aritméticas efetuadas sobre seus valores possuem significado analítico.

87. (Inédita - 2026 - Ultra TI -)

Julgue o item seguinte, a respeito dos atributos qualitativos.

Um atributo binário somente pode ser considerado qualitativo quando seus dois valores forem representados por textos, como "sim" e "não"; se forem codificados como 0 e 1, o atributo será necessariamente quantitativo.

88. (Inédita - 2026 - Ultra TI -)

No que se refere aos tipos de atributos, julgue o próximo item.

Uma classificação de atendimento formada pelas categorias "baixo", "médio" e "alto" constitui atributo qualitativo ordinal, pois as categorias apresentam uma ordem, ainda que não expressem uma mensuração quantitativa.

89. (Inédita - 2026 - Ultra TI -)

Com relação aos atributos quantitativos, julgue o item a seguir.

A quantidade de itens de uma nota fiscal pode ser classificada simultaneamente como atributo quantitativo discreto e atributo de razão, uma vez que decorre de contagem e possui zero absoluto.

90. (Inédita - 2026 - Ultra TI -)

Acerca das escalas intervalar e de razão, julgue o item subsequente.

A temperatura em graus Celsius é um exemplo de atributo intervalar: diferenças entre valores possuem significado, mas o zero da escala não representa ausência total de temperatura, razão por que comparações de proporção não são adequadas.

91. (Inédita - 2026 - Ultra TI -)

Julgue o item seguinte, relativo à classificação de atributos quantitativos.

Como os anos do calendário apresentam ordem e intervalos regulares, eles constituem uma escala de razão, sendo correto afirmar, por exemplo, que o ano 2000 representa o dobro do ano 1000.

92. (Inédita - 2026 - Ultra TI -)

No que se refere a dimensões, medidas e métricas, julgue o próximo item.

Em uma tabela na qual cada observação representa um item vendido, a quantidade e o valor daquele item são medidas; já a soma dos valores de diversos itens para apurar o total vendido constitui uma métrica.

93. (Inédita - 2026 - Ultra TI -)

Acerca do papel dos atributos em uma análise de dados, julgue o item a seguir.

Dimensões são necessariamente atributos textuais, pois atributos representados por números somente podem desempenhar o papel de medidas.

94. (Inédita - 2026 - Ultra TI -)

Julgue o item subsequente, relativo a vetores de atributos.

Para que diferentes observações sejam representadas de forma coerente por vetores de atributos, deve-se manter a mesma ordem dos atributos em todos os vetores, embora essa ordem não precise coincidir com a disposição original das colunas de uma tabela.

95. (Inédita - 2026 - Ultra TI -)

Com relação aos conceitos de medida e métrica, julgue o próximo item.

Para ser considerada métrica, uma quantificação deve obrigatoriamente resultar da agregação de várias observações, não podendo ser obtida diretamente de uma medida associada a uma única observação.

96. (Cebraspe - 2024 - ERAC/Qualquer Área de Formação - ANAC)



A respeito de dado, informação, conhecimento e inteligência, julgue o item a seguir.

Na fase de armazenamento do ciclo de vida dos dados, o foco principal é prover meios que ampliem os níveis de utilização desses dados, seja por ampliação das possibilidades de acesso via cópia, por obtenção de conjuntos para análise, ou por meio da disponibilização de recursos de visualização desses dados.

97. (Cebraspe - 2020 - ProTI/Ciência de Dados - ME)



A respeito de dados, informação, conhecimento e inteligência, julgue o próximo item.

O ciclo de vida da informação compreende as fases de produção, manuseio, armazenamento, transporte e descarte.

98. (Cebraspe - 2014 - AJ/Análise de Sistemas/Segurança da Informação - TJ SE)



Acerca de gestão, risco e conformidade, julgue o item a seguir.

O ciclo de vida da informação, inclui as etapas de manuseio, de armazenamento, de transporte e de uso, mas não a de descarte, uma vez que esta, mesmo no caso de ser descarte inadequado, não é fonte de vulnerabilidade.

99. (FAIBC - 2012 - Ana/Gestão de Sistemas - CM Nova Iguaçu)



Levando em consideração o ciclo de vida da informação marque o item em que caracteriza-se o fim deste ciclo.

- A) Manuseio.
- B) Armazenamento.
- C) Transporte.
- D) Descarte.
- E) Criação.

100. (FCC - 2026 - GG POG - SEPLAG PE)



Uma equipe de Tecnologia da Informação inicia um projeto para aprimorar o processo de planejamento orçamentário. Nessa etapa, são reunidos documentos em PDF, mensagens de e-mail e imagens de notas fiscais digitalizadas, provenientes de diversas unidades administrativas, com o objetivo de consolidar essas informações em um sistema de apoio à decisão.

Essa situação caracteriza-se por uma atuação sobre dados

- A) estruturados na etapa de coleta.
- B) estruturados na etapa de tratamento.
- C) estruturados na etapa de armazenamento.
- D) não estruturados na etapa de coleta.
- E) não estruturados na etapa de recuperação.

101. (FCC - 2018 - TJ/Adm/Agente Jud Policial - TRT 15)



O técnico de segurança deve conhecer o ciclo de vida da informação, que é composto por

- A) produção, proteção, contenção e arquivamento.
- B) criação, contenção, divulgação e arquivamento.
- C) manuseio, armazenamento, transporte e descarte.
- D) manuseio, armazenamento, confinamento e controle.
- E) criação, manuseio, confinamento e descarte.

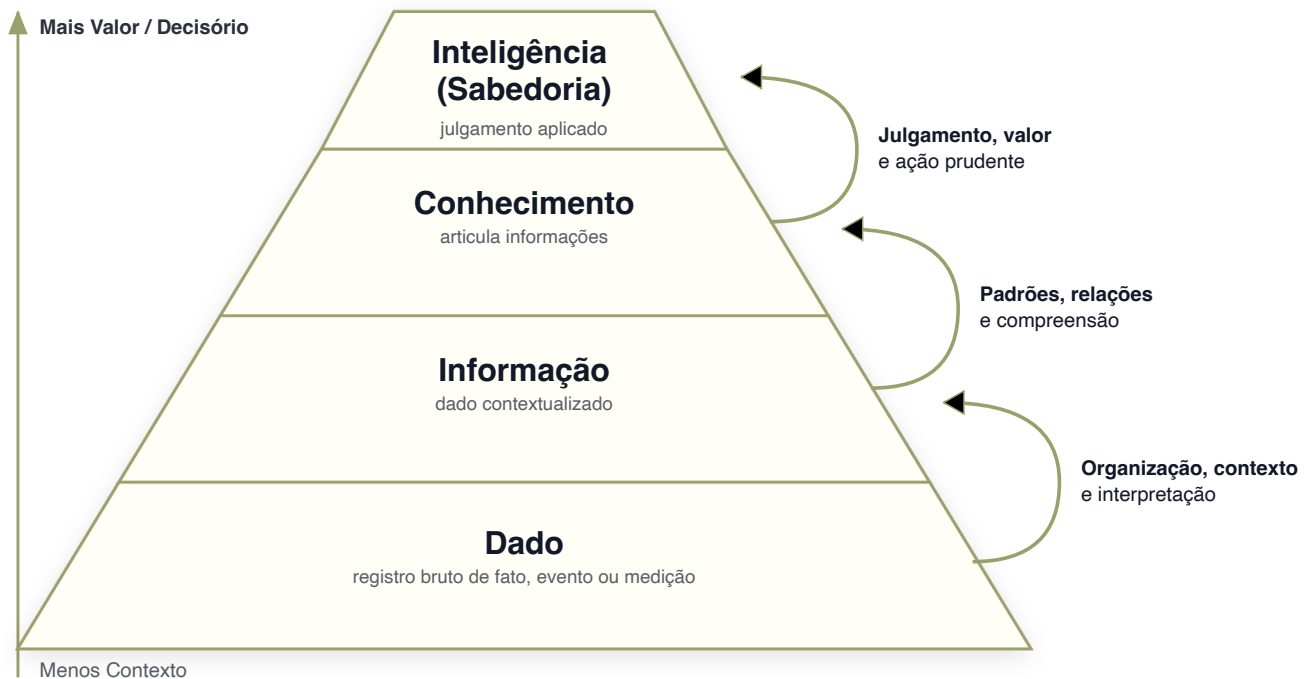
6.1 Gabarito

Nº	Resposta	Nº	Resposta	Nº	Resposta	Nº	Resposta
1	Errada	26	Errada	51	Errada	76	A
2	Certa	27	Certa	52	Errada	77	C
3	Errada	28	C	53	Errada	78	B
4	Certa	29	B	54	Certa	79	C
5	E	30	D	55	Errada	80	A
6	C	31	B	56	Errada	81	C
7	Errada	32	C	57	D	82	Certa
8	Errada	33	C	58	B	83	Errada
9	Certa	34	C	59	Certa	84	Certa
10	Certa	35	B	60	Errada	85	Certa
11	Errada	36	Certa	61	B	86	Errada
12	Certa	37	A	62	Certa	87	Errada
13	Certa	38	Certa	63	Errada	88	Certa
14	Errada	39	Errada	64	Errada	89	Certa
15	Errada	40	Errada	65	Errada	90	Certa
16	Certa	41	Errada	66	A	91	Errada
17	Errada	42	A	67	B	92	Certa
18	Certa	43	Errada	68	C	93	Errada
19	Certa	44	E	69	B	94	Certa
20	Errada	45	Certa	70	D	95	Errada
21	Errada	46	E	71	D	96	Errada
22	Certa	47	Certa	72	C	97	Certa
23	Errada	48	Certa	73	A	98	Errada
24	D	49	Certa	74	E	99	D
25	Errada	50	Errada	75	A	100	D

Nº	Resposta	Nº	Resposta	Nº	Resposta	Nº	Resposta
101	C						

7 Resumos

Hierarquia DIKW (Pirâmide informacional)



Dado

- Registro bruto
- Elemento ainda não interpretado, não organizado, nem contextualizado no contexto decisório
- Matéria-prima da informação
- Dados podem ser obtidos por sensores, sistemas, máquinas ou registros humanos.

Informação

- É o dado organizado, interpretado ou associado a significado
- Dado que passou a ser útil para algum fim
- Dados geram informação quando organizados e contextualizados, não o contrário

Conhecimento

- Construído a partir das informações obtidas dos dados
- Implica o reconhecimento de padrões, relações e implicações relevantes
- Permite compreender, explicar e orientar a ação (é acionável)
- Pode ser tácito (obtido pela experiência) ou explícito
- Pode ser difícil de estruturar e de transferir, especialmente quando é tácito

Inteligência (sabedoria)

- Orienta a tomada de decisões concretas e eficazes
- Representa a **aplicação qualificada do conhecimento** para decisão, inferência ou ação
- Envolve **juízo, discernimento e capacidade de agir de maneira adequada** diante de uma situação real
- Na formulação clássica, está associada ao juízo humano



Decorar!

- **Dado:** registro bruto
- **Informação:** dado interpretado e contextualizado
- **Conhecimento:** compreensão (padrões, implicações) obtida a partir das informações
- **Inteligência/Sabedoria:** uso do conhecimento, com juízo, para agir e decidir.

Dados estruturados

- Estrutura **rígida e previsível**
- Dados são **tipados**
- Definidos num **esquema** (inclusive restrições) a parte.
- Exemplos principais são tabelas em bancos de dados relacionais e algumas estruturas tabulares como planilhas em Excel.
- Mais fáceis para consultar.

Dados semiestruturados

- Para prova, guarde principalmente: dados semiestruturados têm estrutura flexível, irregular, parcialmente implícita e frequentemente embutida nos próprios dados.
- São considerados **auto-descritivos**.
- XML, JSON e CSV costumam ser associados a dados semiestruturados



Atenção!

O **Cebraspe** costuma considerar que **dados não estruturados não podem ser armazenados em tabelas de bancos relacionais**, que são associadas a dados estruturados.

Atenção, porque **bancos de dados relacionais permitem, há muito tempo, o armazenamento de conteúdos não estruturados em colunas do tipo BLOB** (e outros tipos binários), incluindo imagens, vídeos e arquivos em geral.

No entanto, esse tipo de armazenamento **reduz as vantagens dos bancos de dados relacionais, como a facilidade de validação interna do conteúdo e consulta (já que não contém estrutura interna definida)**. Portanto, do ponto de vista prático, é possível armazenar dados não estruturados em bancos de dados relacionais, embora não seja algo desejável.

Dados não estruturados

- **Não possuem** um esquema/estrutura pré-definido
- Produzidos de forma livre
- Menos apropriados para pesquisas
- Maior parte dos arquivos no nosso computador: imagens, vídeos e arquivos em geral.

Tipo	Definição	Diferença-chave
Dados estruturados	Estrutura rígida e previsível definida em um esquema.	Fácil para realizar buscas e consultas analíticas
Dados semiestruturados	Dado possui estrutura flexível e identificável por tags e marcas (XML, JSON). Auto-descritivo.	Flexível, mais fácil de organizar e de processar do que dados não estruturados.
Dados não estruturados	Não possui esquema de dados pré-definido	Estrutura de forma livre, mas é difícil de processar e analisar

Observações e Granularidade

- **Observações** correspondem a eventos, objetos ou experimentos medidos no mundo real, como um paciente ou registros de venda.
- **Granularidade** determina o nível de detalhe que cada observação possui: **granularidade fina** indica alto detalhe; **granularidade grossa** indica visões agregadas por mês, dia, empresa, ou qualquer outro critério de agrupamento.
- **Conjunto de dados** (ou apenas dados) representa uma coleção de observações com seus respectivos atributos.

Atributos e Variáveis

- Um atributo representa uma **característica de uma entidade, evento ou observação**.
- Uma variável representa um atributo **no contexto de uma análise de dados**.

Atributo Qualitativos

- **Nominais** - valores representam categorias ou rótulos. Podem ser textuais ou numéricos.
- **Binários** - os valores possíveis são apenas dois distintos: 0 ou 1, Sim ou Não, etc.
- **Ordinais** - valor indica uma ordem ou *ranking* atribuída à observação.

Atributos Quantitativos

- **Discretos** - geralmente representados por números inteiros associados à contagem.
- **Contínuos** - podem assumir qualquer valor num determinado intervalo.
- **Intervalos** - valor numérico que possui ordem clara e distâncias iguais e constantes entre seus valores, mas que **não possi ponto zero absoluto**. Ex: temperatura em Celsius.
- **Razão** - valor numérico que possui **ponto zero absoluto**. Ex: peso ou altura.

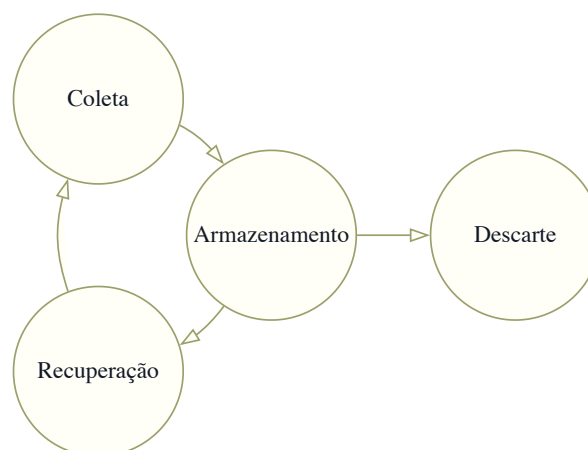
Dimensões

Dimensões são atributos que fornecem o **contexto da análise**. Elas são utilizadas para filtrar, segmentar, organizar e agrupar as observações.

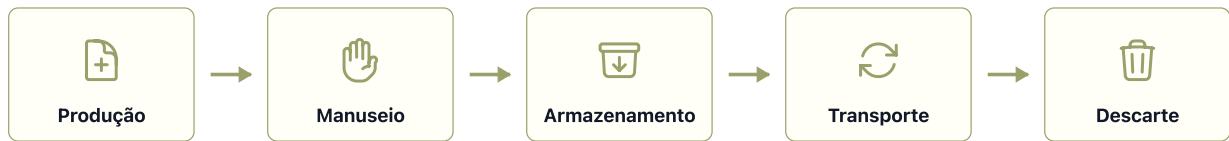
Medidas e Métricas

A **medida** é um valor quantitativo associado à observação. A **métrica** é uma quantificação utilizada na análise, obtida diretamente de uma medida ou por meio de agregação e cálculo.

Ciclo de Vida dos Dados (CVD)



Ciclo de Vida da Informação (CVI)



CVD x CVI

- **Ciclo de Vida dos Dados** enfatiza o fluxo dos dados como registros em sistemas e repositórios
- **Ciclo de Vida da informação** ressalta a trajetória da informação dentro da organização