



COMPANHIA DE INFORMÁTICA DE JUNDIAÍ – CIJUN

ESTADO DE SÃO PAULO

CONCURSO PÚBLICO

009. PROVA OBJETIVA

ANALISTA DE DADOS PLENO (ESTATÍSTICA) (PRESIDÊNCIA)

- ◆ Você recebeu sua folha de respostas e este caderno contendo 50 questões objetivas.
- ◆ Confira seus dados impressos na capa deste caderno e na folha de respostas.
- ◆ Quando for permitido abrir o caderno, verifique se está completo ou se apresenta imperfeições. Caso haja algum problema, informe ao fiscal da sala para a devida substituição desse caderno.
- ◆ Leia cuidadosamente todas as questões e escolha a resposta que você considera correta.
- ◆ Marque, na folha de respostas, com caneta de tinta preta, a letra correspondente à alternativa que você escolheu.
- ◆ A duração da prova é de 3 horas, já incluído o tempo para o preenchimento da folha de respostas.
- ◆ Só será permitida a saída definitiva da sala e do prédio após transcorridos 75% do tempo de duração da prova.
- ◆ Ao sair, você entregará ao fiscal a folha de respostas e este caderno.
- ◆ Até que você saia do prédio, todas as proibições e orientações continuam válidas.

AGUARDE A ORDEM DO FISCAL PARA ABRIR ESTE CADERNO.

Nome do candidato _____

RG _____

Inscrição _____

Prédio _____

Sala _____

Carteira _____

CONHECIMENTOS GERAIS

LÍNGUA PORTUGUESA

Leia o texto para responder às questões de números **01** a **03**.

Enquanto um mundo novo repleto de oportunidades se abre ao adolescente que vislumbra a fase adulta, é possível que aos seus responsáveis reste uma sensação oposta, de um mundo com menos brilho do que havia antes de os filhos crescerem. E não precisa nem que o tal ninho vazio se concretize — muitas vezes, a dor já começa antes.

A educadora e pesquisadora do comportamento adolescente Carolina Delboni diz que, para muitos pais, a sensação nessa fase da vida é de que perderam os filhos. “E, como os pais não sabem se relacionar com essa ‘perda’, muitas vezes eles acabam se relacionando pela raiva”, afirma. Com isso, a distância em casa só aumenta: adultos num canto, magoados, e os jovens no outro, se sentindo incompreendidos e buscando compreensão lá fora.

Também é papel dos pais, segundo Delboni, quebrar os estereótipos da adolescência — evitar cair na tentação, por exemplo, de acusar os filhos de serem lacônicos, de não saírem do quarto e do celular, ou de dizer que conversar com eles é impossível. “A gente tem o costume de rebater tudo o que vem do universo adolescente, de dizer que é coisa fútil, de chamar a música que eles gostam de ruim, de falar que o lugar que eles escolheram para ir passear não é o melhor etc. Sempre temos algo a dizer que mostra que eles deveriam estar fazendo diferente”, afirma a pesquisadora.

Para ela, este hábito “corta todos os vínculos” com o jovem. “O que isso vai acrescentar às relações? Nada. Não vai melhorar o vínculo, pelo contrário, e ainda não vai tirar ele dali do celular. Por outro lado, você já experimentou perguntar o que ele está vendo? Pode ser que ele não mostre, mas você fez uma pergunta, demonstrou interesse. Nesse cotidiano, estamos sempre querendo tirar o adolescente à força do seu lugar em vez de nos deslocarmos para o canto dele, sem reclamar”.

(Marcella Franco. *Estamos sempre querendo tirar o adolescente à força do seu lugar, diz educadora. Folha de S. Paulo*, 12.10.2023. Adaptado)

01. Com base no texto, é correto afirmar que a adolescência

- (A) é uma fase com menos brilho do que a infância, visto que os jovens lidam com dificuldade com a separação em relação aos pais.
- (B) deve ser mais bem compreendida pela família e pela escola, para que se orientem os jovens acerca dos tipos de relações e do estilo de vida ideais.
- (C) é um período que apresenta desafios para muitos pais, os quais têm dificuldade em lidar com a transformação dos filhos em relação à infância.
- (D) pode ser comparada à infância pela dificuldade de comunicação e pela necessidade constante de orientação dos jovens.
- (E) dá aos pais a sensação de ninho vazio, já que os jovens cortam todo e qualquer vínculo de afeto em relação a eles.

02. Em “Enquanto um mundo novo repleto de **oportunidades** se abre **ao adolescente...**” (1º parágrafo), as expressões em destaque exercem, respectivamente, as funções de

- (A) complemento nominal e objeto indireto.
- (B) adjunto adnominal e complemento nominal.
- (C) complemento nominal e agente da passiva.
- (D) adjunto adverbial e agente da passiva.
- (E) adjunto adnominal e objeto indireto.

03. Assinale a alternativa em que a exclusão da vírgula no trecho original mantém a norma-padrão de emprego da vírgula.

- (A) Enquanto um mundo novo repleto de oportunidades se abre ao adolescente que vislumbra a fase adulta é possível que aos seus responsáveis... (1º parágrafo)
- (B) E não precisa nem que o tal ninho vazio se concretize — muitas vezes a dor já começa antes. (1º parágrafo)
- (C) ... muitas vezes eles acabam se relacionando pela raiva” afirma. (2º parágrafo)
- (D) ... de dizer que é coisa fútil de chamar a música que eles gostam de ruim... (3º parágrafo)
- (E) Pode ser que ele não mostre mas você fez uma pergunta... (4º parágrafo)

04. Assinale a alternativa em que se observa emprego e colocação dos pronomes de acordo com a norma-padrão.

- (A) É necessário pensar antes de dizer algo aos filhos de modo que não lhes magoemos.
- (B) Interessar-se pelo que os adolescentes fazem os permitirá agir com mais confiança.
- (C) Há pais que relacionam-se com os filhos certos de que eles não têm vontade própria.
- (D) Pais queriam que os filhos se comportassem de novo como se deles dependessem.
- (E) No dia a dia, pais não sabem mais como tratar os filhos e têm criado-os incorretamente.

05. Pesquisas comprovam que adolescentes _____ pais fumam são mais propensos _____ fumarem. Além disso, nas famílias _____ os pais são fumantes verificou-se maior prejuízo _____ saúde dos filhos, mesmo que estes ainda não tenham adquirido esse hábito.

Assinale a alternativa que completa, correta e respectivamente, as lacunas do texto.

- (A) cujos ... a ... onde ... a
- (B) os quais ... à ... em que ... à
- (C) os quais ... a ... onde ... à
- (D) cujos ... a ... em que ... à
- (E) cujos os ... a ... onde ... a

06. Uma loja mandou fabricar 4 500 calendários que serão distribuídos entre suas 3 filiais A, B e C, de maneira diretamente proporcional ao número de cartões fidelidade emitidos por cada filial. Sabe-se que o número de cartões fidelidade emitidos pela filial B superou em 20 o número de cartões emitido pela filial A e que a filial C, cujo número de cartões fidelidade emitidos é igual a $\frac{7}{8}$ do número

de cartões emitidos pela filial B, recebeu 200 calendários a menos do que a filial B.

Sabendo que ao todo foram emitidos 900 cartões fidelidade, o número de calendários recebidos pela filial C foi

- (A) 1800.
- (B) 1600.
- (C) 1500.
- (D) 1400.
- (E) 1200.

07. A média aritmética dos valores de venda de 5 imóveis é R\$ 380.000,00. A média aritmética do valor de venda dos 2 imóveis mais caros supera em R\$ 100.000,00 a média aritmética do valor de venda dos outros 3 imóveis. Se a diferença entre o valor de venda dos 2 imóveis mais caros é R\$ 20.000,00, o valor do imóvel mais caro é igual a

- (A) R\$ 370.000,00.
- (B) R\$ 410.000,00.
- (C) R\$ 430.000,00.
- (D) R\$ 450.000,00.
- (E) R\$ 470.000,00.

08. No início de um congresso, a razão do número de palestrantes convidados para o número de inscrições recebidas era $\frac{1}{40}$. No decorrer desse congresso, 3 palestrantes

convidados não puderam comparecer e ocorreram 60 novas inscrições, de modo que ao final do congresso, a razão do número de palestrantes convidados que compareceram para o número de inscrições totais recebidas passou a ser $\frac{3}{140}$. Para que essa razão de $\frac{3}{140}$ tivesse

sido de $\frac{1}{50}$, o número de novas inscrições, teria que superar as 60 novas inscrições recebidas, em

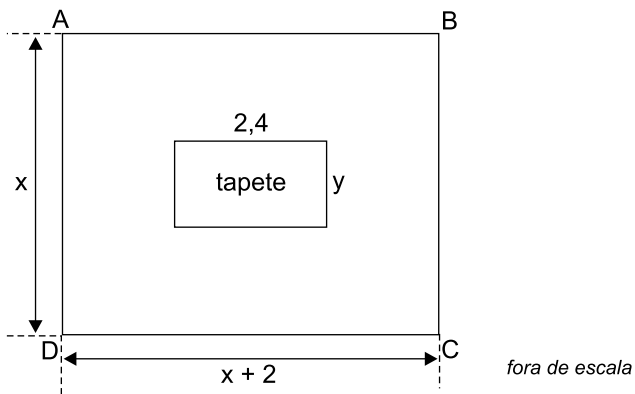
- (A) 30.
- (B) 50.
- (C) 90.
- (D) 120.
- (E) 150.

09. A tabela apresenta algumas informações sobre o número de unidades vendidas dos produtos A, B e C, o valor unitário de custo e o respectivo valor unitário de venda.

Produto	Nº de unidades vendidas	Valor unitário de custo	Valor unitário de venda
A	30	R\$ 5,00	4x
B	y	R\$ 4,00	R\$ 10,00
C	25	x	R\$ 8,00

Sabendo que o custo de todas essas unidades foi de R\$ 385,00 e que o valor obtido com a venda de todas elas foi de R\$ 960,00, então, se todas as unidades do produto B tivessem sido vendidas pelo valor unitário de venda do produto A, o valor obtido com a venda de todas as unidades do produto B teria sido de

- (A) R\$ 240,00.
 (B) R\$ 300,00.
 (C) R\$ 360,00.
 (D) R\$ 420,00.
 (E) R\$ 480,00.
10. Em uma sala retangular ABCD, a medida do comprimento excede a medida da largura em 2 m. Nessa sala foi colocado um tapete retangular com 2,4 m de comprimento e y metros de largura, conforme mostra a figura, onde as medidas indicadas estão em metros.



Sabendo que a medida do comprimento da sala é 4 vezes a medida da largura do tapete, e que a diferença entre os perímetros da sala e do tapete é 12,2 m, então, a área da sala ocupada pelo tapete é de

- (A) 15%.
 (B) 16%.
 (C) 18%.
 (D) 21%.
 (E) 24%.

11. A negação da afirmação “Se Alberto visitou todos os Estados, então ele sabe o nome de todas as capitais” é logicamente equivalente à

- (A) Alberto visitou todos os Estados e ele não sabe o nome de alguma capital.
- (B) Alberto visitou todos os Estados e ele sabe o nome de alguma capital.
- (C) Alberto visitou todos os Estados e ele não sabe o nome de capital alguma.
- (D) Alberto visitou pelo menos um Estado e ele sabe o nome de alguma capital.
- (E) Alberto visitou pelo menos um Estado e ele não sabe o nome de alguma capital.

12. Considere as proposições:

p: Todo pescador tem alicate.

q: Se João é bom pescador, então ele tem alicate.

r: João pesca de barco ou João não usa isca artificial.

Sabendo que o valor lógico de $\neg p \rightarrow (q \vee r)$ é falso, então tem valor lógico verdadeiro o que se afirma em

- (A) João não pesca de barco e tem alicate.
- (B) João pesca de barco e não tem alicate.
- (C) João usa isca artificial e não é bom pescador.
- (D) João é bom pescador e pesca de barco.
- (E) João usa isca artificial e não tem alicate.

13. José foi conhecer o novo restaurante de sua amiga Sueli. Ao chegar ao local, José viu que havia 3 restaurantes lado a lado, um de portas brancas, um de portas pretas e um de portas vermelhas e ele sabia que apenas um restaurante era o de Sueli e que cada atendente ou só mentia ou só dizia a verdade. José perguntou ao atendente de cada restaurante qual era o de Sueli, obtendo as seguintes respostas de cada atendente, identificado pela cor das portas do restaurante que trabalha:

(branca) O restaurante de portas brancas não é o de Sueli.

(preta) O restaurante de Sueli é o de portas vermelhas.

(vermelha) Eu sou o único atendente que fala a verdade.

Como José não foi capaz de identificar o restaurante com as respostas dadas, ele perguntou a cada atendente: “Você trabalha no restaurante de Sueli?”, e todos os atendentes responderam com “Não”.

A cor da porta do restaurante de Sueli e a cor da porta do restaurante que trabalha um atendente que disse a verdade são, respectivamente,

- (A) branca e branca.
- (B) branca e preta.
- (C) preta e branca.
- (D) vermelha e preta.
- (E) vermelha e vermelha.

14. Um relógio cuco está com um defeito e ele só badala a cada 37 minutos, uma única vez, não badalando, necessariamente, a cada hora cheia (0h, 1h, 2h, 3h, etc...). Se no dia 15 de outubro, às 8h, esse relógio badalou, a próxima vez em que ele badalou em uma hora cheia foi no dia 16 de outubro às

- (A) 0h.
- (B) 7h.
- (C) 13h.
- (D) 18h.
- (E) 21h.

15. André, Carlos, Elton e Gustavo são amigos e cada um é irmão de apenas um outro desses rapazes. André e seu irmão nasceram no mesmo Estado. Gustavo nasceu em Pernambuco e seu irmão nasceu em Rondônia. André nasceu no mesmo Estado que Elton. No dia que Carlos nasceu, o irmão de Gustavo completou 10 anos.

A partir das informações dadas, é necessariamente verdade que

- (A) André nasceu em Pernambuco.
- (B) Carlos é irmão de André.
- (C) Elton é irmão de André.
- (D) Gustavo é mais velho do que Carlos.
- (E) André é mais velho do que Elton.

16. Os 234 entrevistados de uma pesquisa responderam à pergunta “Qual ou quais serviços de assinatura de vídeos sob demanda você possui?”. As únicas respostas foram sobre os serviços A, B ou C, sendo que cada entrevistado assina pelo menos um desses serviços. Do total de entrevistados, um terço assina o serviço A e 5 pessoas assinam os três serviços. O número de pessoas que assinam apenas o serviço B é o quádruplo do número de pessoas que assinam apenas os serviços B e C. Ninguém assina apenas os serviços A e B e 92 pessoas assinam o serviço C. São 31 pessoas que assinam apenas os serviços A e C.

O número de entrevistados que assinam apenas o serviço A excede o número de entrevistados que assinam apenas o serviço C em

- (A) 7.
- (B) 9.
- (C) 11.
- (D) 13.
- (E) 15.

17. Considere a sequência 7, 13, 20, 11, 21, 32, 15, 29, 44, 19, ... construída a partir de um padrão lógico. O 500º termo dessa sequência é

- (A) 1277.
- (B) 1301.
- (C) 1341.
- (D) 1421.
- (E) 1453.

18. Considere a sequência 0, 1, 0, 2, 0, 0, 3, 0, 0, 0, 4, 0, 0, 0, 0, 5... Nessa sequência o termo igual a 0 ocorre a primeira vez na 1ª posição, a segunda vez na 3ª posição e a décima vez na 14ª posição.

Nessa sequência, a centésima ocorrência do termo igual a 0 é na

- (A) 114ª posição.
- (B) 119ª posição.
- (C) 123ª posição.
- (D) 126ª posição.
- (E) 129ª posição.

19. Uma equipe é formada por 33 ciclistas e 26 corredores, sendo que quem é corredor não é ciclista. Dessa equipe 28 pessoas usam relógio e, entre os corredores, 5 não usam relógio.

O número de ciclistas que não usam relógio é

- (A) 23.
- (B) 24.
- (C) 25.
- (D) 26.
- (E) 27.

20. Três avós, Alzira, Beatriz e Carlota têm, cada uma, uma neta, cujos nomes são, não necessariamente nessa ordem, Dolores, Edite e Florinda. Essas seis mulheres jogaram várias partidas de dominó entre si, cada partida entre duas mulheres e cada partida com uma vencedora. No total foram 73 partidas e Beatriz foi a grande vencedora, com 24 vitórias. Carlota venceu 3 partidas a mais do que Alzira e a avó de Florinda venceu um número de partidas igual a um terço do número de partidas que a avó de Edite venceu. Dolores e sua avó venceram um total de 18 partidas e a neta de Alzira foi a neta que mais venceu, com 14 vitórias. Uma das netas venceu a metade do número de partidas vencidas por outra neta.

O total de vitórias das netas de Beatriz e Carlota foi

- (A) 12.
- (B) 16.
- (C) 21.
- (D) 22.
- (E) 23.

Leia o texto para responder às questões de números 21 a 25.

Information Security (InfoSec) defined

Information security, often abbreviated InfoSec, is a set of security procedures and tools that broadly protect sensitive enterprise information from misuse, unauthorized access, disruption, or destruction. InfoSec encompasses physical and environmental security, access control, and cybersecurity. It often includes technologies like cloud access security brokers (CASB), deception tools, endpoint detection and response (EDR), and security testing for DevOps (DevSecOps), among others.

Key elements of information security

InfoSec comprises a range of security tools, solutions, and processes that keep enterprise information secure across devices and locations, helping to protect against cyberattacks or other disruptive events.

- Application security - Policies, procedures, tools, and best practices enacted to protect applications and their data.
- Cloud security - Policies, procedures, tools, and best practices enacted to protect all aspects of the cloud, including systems, data, applications, and infrastructure.
- Cryptography - An algorithm-based method of securing communication meant to ensure only intended recipients of a specific message can view and decipher it.
- Disaster recovery - A method to reestablish functional technological systems in the wake of an event like a natural disaster, cyberattack, or another disruptive event.
- Incident response - An organization's plan for responding to, remediating, and managing the aftermath of a cyberattack, data breach, or another disruptive event.
- Infrastructure security - Security that encompasses an organization's entire technological infrastructure, including both hardware and software systems.
- Vulnerability management - The process an organization takes to identify, assess, and remediate vulnerabilities in its endpoints, software, and systems.

([www.microsoft.com/en-ww/security/business/security-101/what-is-information-security-infosec#:~:text=Information%20security%2C%20often%20abbreviated%20\(InfoSec,access%2C%20disruption%2C%20or%20destruction.](https://www.microsoft.com/en-ww/security/business/security-101/what-is-information-security-infosec#:~:text=Information%20security%2C%20often%20abbreviated%20(InfoSec,access%2C%20disruption%2C%20or%20destruction.) Adaptado.)

21. According to the definition of Information Security provided in the first part of the text, the aim of InfoSec is to

- (A) extend access controls to most business partners.
- (B) search for the best technologies to improve information retrieval.
- (C) shield sensitive information against various kinds of threats.
- (D) protect information and data from overuse.
- (E) make sure that at least a stable physical security environment is rendered.

CONHECIMENTOS ESPECÍFICOS

22. No trecho da primeira parte do texto “It often includes technologies **like** cloud access security brokers (CASB), deception tools, endpoint detection and response (EDR), and security testing for DevOps (DevSecOps), among others”, o termo destacado em negrito indica

- (A) preferência.
- (B) exemplificação.
- (C) comparação.
- (D) probabilidade.
- (E) similaridade.

23. In the excerpt from the first part of the text “**It** often includes technologies”, the word in bold refers to

- (A) cybersecurity.
- (B) technologies.
- (C) information security.
- (D) cloud access security brokers.
- (E) security testing for DevOps.

Leia a parte do texto *Key elements of information security* para responder às questões de números 24 e 25.

24. No texto, os elementos chave de Segurança da Informação que mencionam explicitamente o uso de políticas são

- (A) application security e cloud security.
- (B) cloud security e cryptography.
- (C) cryptography e infrastructure security.
- (D) infrastructure security e vulnerability management.
- (E) vulnerability management e application security.

25. Os termos que evidenciam a diferença entre os objetivos de “Disaster recovery” e de “Incident response” são, respectivamente,

- (A) method e plan.
- (B) systems e cyberattack.
- (C) cyberattack e data breach.
- (D) reestablish e remediating.
- (E) recovery e managing.

26. Um analista de dados com diversos programas abertos em seu ambiente MS-Windows 10, em configuração padrão, está processando dados para um relatório e deseja visualizar sua Área de Trabalho, para acessar um arquivo específico.

O atalho por teclado que permite exibir a Área de Trabalho é

- (A) Ctrl + F
- (B) Ctrl + E
- (C) Tecla Windows + E
- (D) Tecla Windows + R
- (E) Tecla Windows + M

27. Um usuário preparou um grande documento por meio do MS-Word 2016, em sua configuração padrão, e para criar um _____, acionou o ícone de mesmo nome exibido a seguir, dentro da guia _____.



Assinale a alternativa que preenche, correta e respectivamente, as lacunas do enunciado.

- (A) Sumário ... Referências.
- (B) Índice ... Inserir.
- (C) Sumário ... Inserir.
- (D) Índice ... Referências.
- (E) Índice ... Design.

28. Um analista de dados, por meio do MS-Excel 2016, em sua configuração padrão, elaborou a planilha a seguir.

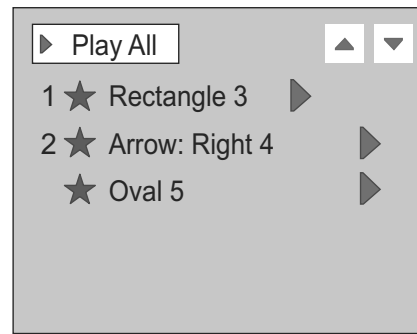
	A	B	C	D
2	1	Novo		
3	2	Alterado		
4	3	Excluído		
5				
6	Registro	Status	Data	Descrição
7	AAA	1	01/01/2023	Novo
8	BBB	1	02/02/2023	Novo
9	CCC	2	03/03/2023	Alterado
10	DDD	3	04/04/2023	Excluído
11	EEE	2	05/05/2023	Alterado
12				

Nas linhas de 2 a 4, foi preparada uma legenda representando os códigos e suas respectivas descrições para indicar o status dos registros. Em seguida, nas colunas de A a C, a partir da linha 6, foram digitados dados de registros e seus respectivos código do status e data do registro. Então, na coluna D, o analista inseriu uma fórmula na célula D7, para preencher automaticamente com a descrição definida nas linhas de 2 a 4, de acordo com o código de status. Após preencher a fórmula em D7, arrastou a alça de preenchimento para preencher as células D8 até D11.

Assinale a alternativa que apresenta a fórmula adicionada em D7 que está de acordo com o enunciado e a imagem exibida.

- (A) =PROCV(\$A\$2:\$B\$4;B7;2)
- (B) =PROCV(B7;\$A\$2:\$B\$4;2)
- (C) =PROCV(B7;\$A;\$2;\$B;\$4;2)
- (D) =PROCV(\$A\$2;B7;\$B\$4;2)
- (E) =PROCV(\$A\$2:\$B\$4;2;B7)

29. Um usuário preparou uma apresentação por meio do MS-PowerPoint 2016, em sua configuração padrão, contendo apenas 1 slide e 3 Formas. Definiu animações de entrada sobre as Formas, cujo Painel de Animação é visto a seguir.



De acordo com a imagem, quantos cliques serão necessários para que todas as Formas apareçam durante o Modo de Apresentação?

- (A) 5
- (B) 4
- (C) 3
- (D) 2
- (E) 1

30. Um recurso de internet bastante utilizado para quem manipula grandes quantidades de dados é a transferência de arquivos pela internet.

Assinale a alternativa que apresenta uma URL que indica o uso de protocolo de transferência de arquivos.

- (A) <http://site.com.br>
- (B) <https://site.com.br>
- (C) <ftp://site.com.br>
- (D) <mms://site.com.br>
- (E) <smtp://site.com.br>

31. A ferramenta Microsoft Excel é muito utilizada na coleta e transformação de dados. Para facilitar a manipulação dos dados, o Excel fornece a opção de formatação de dados como tabelas ou criação de tabelas. Cada tabela deve possuir um nome que segue determinados critérios de validade.

Assinale a alternativa correta sobre nomes válidos para tabelas no Excel.

- (A) Nomes de tabelas seguem os mesmos critérios de nomenclatura de planilhas.
- (B) Nomes de tabelas seguem os mesmos critérios de nomenclatura de arquivos do Sistema Operacional Windows.
- (C) Nomes de tabelas podem ter no máximo 256 caracteres.
- (D) Podem existir tabelas com o mesmo nome dentro de uma mesma pasta de trabalho ou arquivo do Excel.
- (E) Nome de tabelas devem iniciar somente com caracteres de barra invertida (\), sublinhado (_) ou letras.

32. Considere o trecho de código na linguagem R destacado.

```
df <-  
data.frame(  
  natureza_despesa = c(  
    '3.1.90.11.01',  
    '3.1.90.11.01',  
    '3.1.90.11.33',  
    '3.1.90.11.33'  
  ),  
  valor_empenhado = c(45291.59, 134963.54,  
    4538.38, 1200.64)  
)
```

função placeholder(df\$valor_empenhado,
df\$natureza_despesa, sum)

Um analista de dados da prefeitura deseja somar o valor total empenhado por natureza de despesa.

Qual função da família *apply* da linguagem R ele deve substituir no lugar de **função placeholder?**

- (A) apply
- (B) mapply
- (C) sapply
- (D) tapply
- (E) lapply

33. A linguagem Python oferece diversos pacotes nativos muito utilizados por cientistas de dados. Um dos mais utilizados é o pacote *collections*. Considere o código abaixo:

```
from collections import Counter  
counter1 = Counter({'Constitucionais e Royalties': 1302.58  
  , 'Legais, Voluntárias e Específicas': 12000})  
counter2 = Counter({'Constitucionais e Royalties': 93939.08  
  , 'Legais, Voluntárias e Específicas': 3893.57})  
Counter_max = _____
```

Para que os dois contadores sejam mesclados em um único que apresente somente os valores máximos para cada chave, qual código o analista de dados deve utilizar no lugar do sublinhado?

- (A) Counter(counter1, counter2)
- (B) counter1.update(counter2)
- (C) Counter([counter1, counter2])
- (D) counter1 | counter2
- (E) counter1 & counter2

34. Muitos fatores são considerados por uma analista de dados para selecionar a melhor configuração de suas ferramentas. Portanto, para obter um desempenho adequado em seu relatório do Microsoft Power BI é necessário conhecer os modos de armazenamento.

Sobre os modos de armazenamento no Power BI, assinale a alternativa correta.

- (A) No modo “Dupla” (*Dual* em inglês), as tabelas sempre são armazenadas em cache.
- (B) No modo *DirectQuery* as tabelas são armazenadas diretamente no cache.
- (C) Alterar o modo de armazenamento de uma tabela para o modo Importação é uma operação que pode ser revertida.
- (D) No modo Importação as consultas são realizadas diretamente na fonte de dados.
- (E) No modo *DirectQuery* as consultas DAX são atendidas por meio da execução de consultas sob demanda para a fonte de dados.

35. Acerca da ferramenta de Business Intelligence Tableau, assinale a alternativa correta.

- (A) Não é possível definir relações com base em campos geográficos.
- (B) Não é possível configurar outros tipos de conexão de dados além das conexões listadas diretamente dentre os conectores de dados nativos da ferramenta.
- (C) É possível configurar relações com campos que possuem tipos de dados diferentes.
- (D) Quando uma relação é definida, o Tableau mescla automaticamente as tabelas para melhorar o desempenho das consultas.
- (E) Por ser uma ferramenta paga, o Tableau não fornece interfaces para criação de Plug-ins.

36. Um analista de dados da prefeitura recebeu a tarefa de explorar os dados de remuneração dos servidores da prefeitura utilizando a linguagem Java. Considere o trecho de código incompleto gerado pelo analista abaixo:

```
1. import java.util.Arrays;
2. import java.util.List;
3. import java.util.DoubleSummaryStatistics;
4. public class Main {
5.     public static void main(String[] args) {
6.         List < Double > remuneracoes = Arrays.asList
            (4911.82, 2950.99, 19315.51, 3056.44);
7.         DoubleSummaryStatistics ed = remuneracoes.
            stream()
8.             .mapToDouble(remuneracao ->
                remuneracao)
9.             ._____();
10.        System.out.println(ed); } }
```

O código deve utilizar a biblioteca correta do Java para gerar estatísticas descritivas dos dados da lista de remunerações.

Assinale a alternativa que indica corretamente o método que deve ser chamado na linha 9 e quais são as estatísticas descritivas retornadas por esse método.

- (A) summary e as estatísticas: contagem, soma, mínimo, média e máximo.
- (B) summaryStatistics e as estatísticas: contagem, soma, mínimo, média e máximo.
- (C) summary e as estatísticas: mínimo, primeiro quartil, mediana, terceiro quartil e máximo.
- (D) summaryStatistics e as estatísticas: contagem, média, desvio padrão, mediana e máximo.
- (E) doubleSummaryStatistics e as estatísticas: contagem, média, desvio padrão, mediana e máximo.

37. Os Indicadores-Chave de Desempenho, ou KPIs (Key Performance Indicators), são ferramentas essenciais para construção de bons painéis de dados para a gestão. Sobre as características desses indicadores, assinale a alternativa correta.

- (A) Indicadores não precisam ter impacto significativo para o negócio.
- (B) Indicadores não podem ter relação com mais de uma perspectiva do balanced scorecard.
- (C) A responsabilidade sobre um indicador pode ser atribuída a uma equipe ou conjunto de equipes.
- (D) Indicadores devem ser complexos com linguagem focada somente na alta gestão.
- (E) Indicadores-Chave devem ser medidos em intervalos de tempos mensais, bimestrais ou semestrais.

38. Para definir corretamente Indicadores-Chave de Desempenho é necessário que os Fatores Críticos de Sucesso sejam bem delineados. Para essa etapa são utilizados os critérios de qualidade de métricas S.M.A.R.T (Specific - Específico, Measurable - Mensurável, Achievable - Atinável, Relevant – Relevante e Time Sensitive - Temporal).

Assinale a alternativa que apresenta o critério e sua definição correta associada.

- (A) Temporal: devem ser estabelecidos fatores com limites de tempo bem definidos.
- (B) Relevante: devem ser estabelecidos fatores alcançáveis.
- (C) Específico: devem ser estabelecidos fatores que possam fornecer o acompanhamento do progresso.
- (D) Atinável: devem ser estabelecidos fatores orientados pela missão, visão e objetivos da empresa.
- (E) Mensurável: devem ser estabelecidos fatores de forma clara e sem ambiguidade.

39. A análise exploratória de dados e a estatística descritiva são muito utilizadas pelos analistas de dados governamentais na pesquisa de preços para licitações públicas. Considere que após coletar preços para monitores de computador o analista chegou a seguinte tabela de distribuição de frequência:

Valor Unitário	Frequência
R\$ 420,00	2
R\$ 435,00	2
R\$ 457,00	5
R\$ 490,00	1
R\$ 510,00	1
R\$ 560,00	2
R\$ 577,00	5
R\$ 640,00	3
R\$ 835,00	2
R\$ 840,00	2

Para a tabela apresentada, os valores de média, de moda e da mediana são respectivamente:

- (A) 570,8 ; 517,00 ; 560,00
- (B) 570,8 ; 457,00 e 577,00; 562,00
- (C) 570,8 ; 457,00 e 577,00; 560,00
- (D) 576,4 ; 457,00 e 577,00; 560,00
- (E) 576,4; 517,00 ; 562,00

40. O coeficiente de variação de Pearson (CV) é uma medida de dispersão relativa que fornece a variação dos dados em relação à média. Considere a distribuição de frequência de valores de monitores em licitações públicas da tabela abaixo:

Valor Unitário	Frequência
R\$ 420,00	2
R\$ 435,00	2
R\$ 457,00	5
R\$ 490,00	1
R\$ 510,00	1
R\$ 560,00	2
R\$ 577,00	5
R\$ 640,00	3
R\$ 835,00	2
R\$ 840,00	2

Considere que o desvio padrão dos dados apresentados é igual 137,9. Qual o valor do CV em termos percentuais (calcule somente até a segunda casa decimal sem arredondamentos)?

- (A) 4,13%
- (B) 29,97%
- (C) 24,46%
- (D) 24,15%
- (E) 26,49%

41. Medidas de assimetria de uma distribuição são essenciais para verificar distorções nos dados na fase de análise exploratória dos dados. Considere a fórmula de assimetria baseada em quartis dada pela fórmula:

$$SK = \frac{(Q3 - Q2) - (Q2 - Q1)}{Q3 - Q1}$$

Q1, Q2 e Q3 representam o primeiro, segundo e terceiro quartis respectivamente. Considere a distribuição de frequência de valores de monitores em licitações da tabela abaixo:

Valor Unitário	Frequência
R\$ 420,00	2
R\$ 435,00	2
R\$ 457,00	5
R\$ 490,00	1
R\$ 510,00	1
R\$ 560,00	2
R\$ 577,00	5
R\$ 640,00	3
R\$ 835,00	2
R\$ 840,00	2

Qual o valor da assimetria utilizando a fórmula apresentada (calcule até a segunda casa decimal sem arredondamentos)?

- (A) - 0,60
- (B) - 0,12
- (C) 0,12
- (D) 0,60
- (E) 0,32

42. Assinale a alternativa que apresenta uma comparação correta entre as linguagens de programação Scala e Java.

- (A) Scala não compila o código para arquivos “.class” como a linguagem Java.
- (B) Scala é uma linguagem funcional e não utiliza orientação a objetos como a linguagem Java.
- (C) Scala é uma linguagem dinamicamente tipada, enquanto Java é estaticamente tipada.
- (D) Scala não fornece suporte para funções **lambda**, enquanto o Java fornece.
- (E) O código Scala idiomático não utiliza o valor nulo e, portanto, não sofre de NullPointerException como a linguagem Java.

43. A arquitetura do *Hadoop* possui um componente responsável pelo armazenamento distribuído dos dados. Assinale a alternativa que indica esse componente.

- (A) YARN
- (B) Oozie
- (C) Spark
- (D) HDFS
- (E) Ambari

44. Os bancos de dados NoSQL utilizam um modelo transacional mais flexível em relação ao modelo ACID (atomicidade, consistência, isolamento e durabilidade) tradicional. Esse modelo é conhecido como BASE (Basically Available, Soft State e Eventually Consistent).

Assinale a alternativa que traz corretamente uma característica do modelo BASE e sua respectiva definição.

- (A) Eventually Consistent: em algum momento do tempo, quando as atualizações pararem, todos os dados estarão consistentes.
- (B) Soft State: os dados são distribuídos em diversos nós para que o sistema nunca fique fora do ar.
- (C) Basically Available: a responsabilidade pela consistência dos dados é do desenvolvedor da aplicação.
- (D) Eventually Consistent: o banco deve continuar a funcionar apesar de qualquer falha de comunicação entre nós.
- (E) Basically Available: todos os clientes devem ver o mesmo dado, sem quebra de consistência.

45. Para construção de um Data Warehouse é necessário escolher a tecnologia mais adequada para as necessidades de negócio. Em relação as tecnologias ROLAP (*Relational Online Analytical Processing*) e MOLAP (*Multidimensional Online Analytical Processing*), assinale a alternativa correta.

- (A) MOLAP armazena os dados em memória com valores pré-agregados.
- (B) ROLAP fornece melhor tempo de resposta por utilizar diretamente bancos relacionais.
- (C) ROLAP é utilizado, geralmente, para um volume menor de dados.
- (D) MOLAP não utiliza cubo de dados.
- (E) MOLAP, geralmente, são mais simples e demandam menos esforço de manutenção e modelagem.

46. Na modelagem multidimensional de banco de dados para Data Warehouse, tabelas de dimensão armazenam os dados descritivos relacionados aos fatos.

Quando uma dimensão é denominada *Junk* (lixo),

- (A) está contida dentro da própria tabela fato por ter somente sua chave como atributo.
- (B) pode estar relacionada com a tabela fato em visões diferentes pela mesma chave substituta.
- (C) tem como vantagem diminuir o tamanho da tabela fato associada.
- (D) deve possuir conexão com outras dimensões.
- (E) está normalizada e subdividida em tabelas menores de domínio.

47. Um analista de dados da prefeitura deseja calcular a média da diferença entre o valor empenhado e o valor pago das despesas. Para isso, encontrou no banco de dados as duas tabelas:

tabela_empenhado

numero_empenho	valor
24288/2023	10309.50
31199/2023	6626.55
30607/2023	10444.17

tabela_pago

numero_empenho	valor
24288/2023	10309.50
31199/2023	5000.00
30607/2023	null

Para realizar o cálculo elaborou a seguinte consulta SQL:

```
SELECT AVG(diferenca)
FROM
(
    SELECT tabela_pago.numero_empenho
    , tabela_empenhado.valor - tabela_pago.valor AS
    diferenca
FROM tabela_pago
    INNER JOIN tabela_empenhado
    ON tabela_pago.NUMERO_EMPENHO =
    tabela_empenhado.NUMERO_EMPENHO
) AS subquery;
```

Ao executar a consulta, o resultado numérico obtido pelo analista foi (considerando até a terceira casa decimal):

- (A) null
- (B) 542.183
- (C) 813.275
- (D) 9 126.94
- (E) 7 654.75

48. Após a definição do problema a ser estudado, é necessário delimitar corretamente qual o tipo de tarefa de mineração de dados. Em relação as tarefas de estimação denominadas de classificação e de regressão, assinale a alternativa correta.

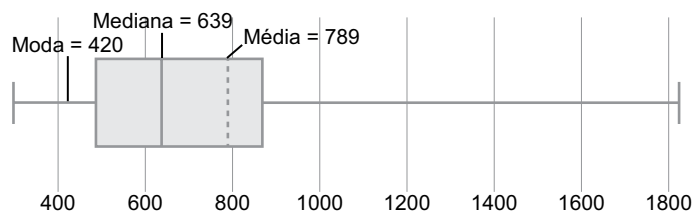
- (A) Não há diferença entre os dois tipos de tarefa.
- (B) Estimar a função orçamentária de uma despesa é um exemplo de tarefa de regressão.
- (C) Estimar o valor gasto com educação para o ano seguinte é um exemplo de tarefa de regressão.
- (D) Estimar o valor do crescimento do PIB é uma tarefa de classificação.
- (E) As tarefas se diferenciam por uma ser supervisionada e a outra semi-supervisionada.

49. Um analista de dados recebeu a tarefa de criar um classificador de imagens e fotos capturadas por câmeras da prefeitura de Jundiaí. Foi escolhida a arquitetura de rede neural profunda mais popular para a tarefa em questão. Assinale a alternativa que indica essa arquitetura.

- (A) RNN
- (B) CNN
- (C) LSTM
- (D) BiLSTM
- (E) LSTM-CRF

50. Um analista de dados da prefeitura deseja realizar uma análise visual da assimetria da distribuição de valores de determinado item em aquisições públicas. Para isso, construiu o seguinte diagrama de caixa (*boxplot*) com os dados coletados:

Valores de aquisição de peças de computadores



Ao analisar o gráfico, o analista concluiu corretamente que

- (A) a distribuição tem assimetria negativa ou à esquerda.
- (B) a distribuição tem assimetria positiva ou à esquerda.
- (C) a distribuição é simétrica.
- (D) a distribuição tem assimetria negativa ou à direita.
- (E) a distribuição tem assimetria positiva ou à direita.

