

Cargo: **ANALISTA DE TECNOLOGIA DA INFORMAÇÃO - CIÊNCIA DE DADOS**

PROVA OBJETIVA

ORIENTAÇÕES

- A Prova Objetiva possui 50 (cinquenta) questões, que deverão ser respondidas no período máximo de quatro horas.
- O tempo de duração das provas abrange a assinatura da Folha de Respostas e a transcrição das respostas do Caderno de Questões da Prova Objetiva para a Folha de Respostas.
- Não será permitido ao candidato ausentar-se em definitivo da sala de provas antes de decorrida 1 (uma) hora do início das provas.
- O candidato não poderá levar o seu Caderno de Questões da Prova Objetiva.
- Os três últimos candidatos deverão permanecer na sala até que todos os demais tenham terminado a prova. Apenas podendo retirar-se, concomitantemente, após a assinatura do relatório de aplicação de provas.
- Depois de identificado e instalado, o candidato somente poderá deixar a sala mediante consentimento prévio, acompanhado de um fiscal, ou sob a fiscalização da equipe de aplicação de provas.
- Será proibido, durante a realização das provas, fazer uso ou portar, mesmo que desligados, telefone celular, relógios, *paggers*, *beep*, agenda eletrônica, calculadora, *walkman*, *tablets*, *notebook*, *palmtop*, gravador, transmissor/receptor de mensagens de qualquer tipo ou qualquer outro equipamento eletrônico. A organização deste Concurso Público não se responsabilizará pela guarda destes e de outros equipamentos trazidos pelos candidatos.
- Durante o período de realização das provas, não será permitida qualquer espécie de consulta ou comunicação entre os candidatos ou entre estes e pessoas estranhas, oralmente ou por escrito, assim como não será permitido o uso de livros, códigos, manuais, impressos, anotações ou quaisquer outros meios.
- Durante o período de realização das provas, não será permitido também o uso de óculos escuros, boné, chapéu, gorro ou similares, sendo o candidato comunicado a respeito e solicitada a retirada do objeto.
- Findo o horário limite para a realização das provas, o candidato deverá entregar as folhas de resposta da prova, devidamente preenchidas e assinadas, ao Fiscal de Sala.
- O candidato não poderá amassar, molhar, dobrar, rasgar ou, de qualquer modo, danificar sua Folha de Respostas, sob pena de arcar com os prejuízos advindos da impossibilidade de sua correção. Não haverá substituição da Folha de Respostas por erro do candidato.
- Ao transferir as respostas para a Folha de Respostas, use apenas caneta esferográfica preta; preencha toda a área reservada à letra correspondente à resposta solicitada em cada questão (conforme exemplo a seguir); assinale somente uma alternativa em cada questão. Sua resposta NÃO será computada se houver marcação de mais de uma alternativa, questões não assinaladas ou questões rasuradas.

	A	B	C	D
01	(A)	(B)	(C)	●

1 De acordo com a sintaxe da norma padrão da Língua Portuguesa, no período composto as orações podem ser coordenadas ou subordinadas. No caso das orações subordinadas, elas são classificadas de acordo com as funções sintáticas que desempenham com relação à oração principal. Com base nessa informação, analise a tirinha abaixo:



FONTE DA IMAGEM: <https://www.facebook.com/tirinhasanesia/>

Assinale a alternativa que classifica corretamente a oração subordinada presente no primeiro quadrinho: **“Você disse que o filme deu a louca na rainha era um filme emocionante”**:

- (A) Oração subordinada substantiva subjetiva
- (B) Oração subordinada substantiva completiva nominal.
- (C) Oração subordinada substantiva predicativa.
- (D) Oração subordinada substantiva objetiva direta.

2 Considere o texto da coluna “nacional”, publicado pela CNN Brasil, em 24/02/2022:

Exclusão escolar no Brasil atinge mais as meninas, diz pesquisa do Fundo Malala

À CNN, a doutora em educação e integrante do Malala Fund no Brasil, Cleo Manhas, disse que meninas assumiram mais responsabilidades no período da pandemia

Um estudo do Fundo Malala apontou que a taxa de exclusão escolar no Brasil aumentou de forma drástica desde o início da pandemia de Covid-19.

Em entrevista à CNN, a doutora em educação e integrante da Rede de Ativistas pela Educação do Malala Fund no Brasil, Cleo Manhas, avaliou que a desigualdade se acentuou durante esse período.

“Somos onze ativistas de diferentes organizações que compõem o Malala Fund e constatamos que recaem maiores responsabilidades sobre as meninas”, afirmou.

Entre elas, por exemplo, estão as atividades domésticas e mesmo fora de casa. “Aliado a isso, problemas de acesso à internet fizeram com que elas perdessem muitas aulas e acabaram desistindo da escola”.

A especialista analisou que não há um plano de ação do Ministério da Educação “para busca ativa de quem não retornou às aulas”. “Temos uma realidade complicada e difícil de ser mapeada porque soubemos agora, nos últimos dias, que há falta de acesso aos dados do censo escolar”.

Cleo defende a adoção de um plano de ação nacional e mais investimento em recursos orçamentários e de infraestrutura para as escolas, com o objetivo de minimizar os efeitos da pandemia.

“A realidade das escolas públicas é de que são superlotadas e sem infraestrutura para receber alunos durante a pandemia e resolver questões aprofundadas”, disse. “O peso da responsabilidade recai de diferentes formas entre meninas e meninos, sobram funções de cuidado às meninas, como de irmãos menores, cuidar da casa, são elas as mais vulneráveis à gravidez na adolescência, recai com mais força sobre meninas e negras, nossas desigualdades se aprofundaram”, completou.

Garcia, A.; Campos, B. CNN Brasil – São Paulo. Disponível em: cnnbrasil.com.br/nacional. Acesso em: 24 fev. 2022.

Durante a leitura do texto, o leitor é levado a compreender que:

- I – a exclusão escolar atinge mais as meninas, pelas responsabilidades assumidas no período da pandemia.
- II – muitas meninas abandonaram a escola, sobretudo pela falta de acesso à internet.

III – a desigualdade se acentuou durante o período da pandemia; muitas meninas desistiram da escola porque a responsabilidade com atividades domésticas e problemas de acesso à internet determinaram a perda de muitas aulas. Esta exclusão escolar é inerente à falta de recursos nas escolas públicas.
IV – falta um plano do Ministério da Educação para minimizar os efeitos da pandemia e resolver os problemas das escolas públicas, desde o levantamento dos(as) estudantes que não retornaram às aulas, até investimentos em infraestrutura para as escolas.

São assertivas as alternativas:

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

3 Observe o trecho abaixo, retirado de um artigo científico:

Observando a luta por igualdade de gênero e seus efeitos sobre o processo eleitoral, três considerações são dignas de registro: a primeira é de que é impossível falar em tratamento igualitário entre homens e mulheres sem compreender as causas da profunda desigualdade que estrutura as opções dos eleitores antes mesmo da possibilidade de escolha dos candidatos. Se a baixa representatividade das mulheres na ALBA (Assembleia Legislativa do Estado da Bahia) permanece gerando efeitos negativos para as lutas do movimento feminista na Bahia, uma avaliação que contemple suas causas – sociais e econômicas, precisam de um diagnóstico mais contundente.

SILVA TEIXEIRA, Erica; ZAIDAN DE CARVALHO, Alexandre Douglas. Mulheres na assembleia legislativa da Bahia no pós-1988. Revista Eletrônica Direito e Política, Programa de Pós-Graduação Stricto Sensu em Ciência Jurídica da UNIVALI, Itajaí, v.16, n.1, 1º quadrimestre de 2021. Disponível em: www.univali.br/direitoepolitica - ISSN 1980-7791. Acesso em: 23 fev 2022.

De acordo com Fiorin e Savioli (2008), o discurso dissertativo de caráter científico deve ser ela-

borado de maneira a criar um efeito de sentido de objetividade, pois pretende dar destaque ao conteúdo das afirmações feitas ao enunciado e não à subjetividade do enunciador. Sendo assim, para neutralizar a presença do enunciador, recomenda-se evitar certos procedimentos linguísticos, tais como:

- (A) A exploração do sentido denotativo das palavras visando à precisão conceitual.
- (B) Usar a 1ª pessoa do plural ou sujeito indeterminado (sabemos, sabe-se) para indicar que o conteúdo do enunciado é produto de um saber coletivo (ciência).
- (C) Utilizar a língua padrão na sua expressão formal, pois não se ajusta ao texto científico o uso de gírias ou quaisquer usos linguísticos distanciados da modalidade culta e formal da língua.
- (D) Usar os verbos de dizer na 1ª pessoa do singular (digo, afirmo, penso) para certificar que o conteúdo do enunciado é opinião de quem o proferiu (enunciador).

4 Ao redigir um documento oficial, o órgão público, como comunicador da mensagem, deve atentar-se ao uso adequado da linguagem com base no padrão culto formal da Língua Portuguesa prezando sempre pela transparência e clareza da redação. Com base nessa informação, analise os períodos:

- I – “Conforme acordado na última reunião, encaminho anexo os documentos necessários para dar seguimento ao processo licitatório”.
- II – “Seguimos as orientações do documento oficial onde está explícita a necessidade de anexar todos os arquivos ao processo”.
- III – “Após a análise dos recursos interpostos, segue anexo o resultado final do processo seletivo.”
- IV – “Informamos que a decisão do órgão competente foi tomada com base na lei onde são fixadas as penas para casos como esse”.
- V – “Encaminho em anexo as atas para apreciação do Conselho”.

Considerando as orientações dispostas no Manual de Redação da Presidência da República, assinale a alternativa que apresenta os trechos com construções que **DEVEM SER EVITADAS** na redação oficial:

- (A) I, II e IV.
- (B) I, II, III e IV.
- (C) I, III e V.
- (D) I, II, e V.

5 Leia o texto a seguir para responder à questão que segue.

‘Miúdos’ que falam ‘brasileiro’ em Portugal – Sérgio Rodrigues.

A reportagem publicada nesta quarta (10) no Diário de Notícias de Lisboa acrescenta uma página de comédia – agridoce, mas engraçada – ao novelão dos encontros e desencontros entre as variedades de português faladas aqui e em Portugal.

“Há crianças portuguesas que só falam ‘brasileiro’”, anuncia o título com alarme. O subtítulo explica: “Dizem grama em vez de relva, autocarro é ônibus, rebuçado é bala, riscas são listras e leite está na geladeira em vez de no frigorífico”.

O curioso fenômeno que faz os miúdos soarem como pequenos brasileiros tem explicação mais simples do que aquele abordado pelo escritor português Valter Hugo Mãe no romance “A Máquina de Fazer Espanhóis”. Afinal, a língua no presente caso é a mesma – ou não?

A reportagem resume assim o problema do deslocamento linguístico: “Os educadores notam-no sobretudo depois do confinamento – à conta de muitas horas de exposição a conteúdos feitos por youtubers brasileiros”.

Sim, a raiz do fenômeno está naquilo que “a maioria das crianças portuguesas vê nos ecrãs de tablet, computador ou telemóvel”. Luccas Neto, 36 milhões de seguidores, é citado como o principal desencaminhador linguístico de tuguinhas. Como António, 4 anos, que “começou a dar sinais há já algum tempo. Ao princípio, a família até achava alguma piada à forma como ele falava (...). Mas à medida que o tempo foi passando, a educadora de infância começou a preocupar-se e foi dando sinais, porque o menino não conseguia dizer os r’s nem os l’s”.

Claro que há muito alarmismo e superficialidade nessa abordagem – como é habitual quando o jornalismo, espelhando o senso comum, trata de temas linguísticos –, para não falar do antibrasileirismo que parece andar em alta na cultura portuguesa. De todo modo, recomenda-se pôr a questão em perspectiva.

Será que pais brasileiros deixariam de se preocupar se, num hipotético – e impensável – caso reverso de imperialismo cultural via Youtube, seus herdeiros em idade pré-escolar comessem a engolir vogais e a chamar a camisa do Flamengo ou do Corinthians de camisola?

Não acho justo acusar a reportagem do DN de xenofobia, embora ela exale sem dúvida um nacionalismo rançoso. O texto até busca algum equilíbrio por trás do ar assustado.

Seu maior mérito nesse sentido é dar voz à linguista Catarina Menezes, que relativiza o problema. “Quando eu era menina havia o mesmo pânico social com os livros do tio Patinhas, que eram traduzidos em português do Brasil”, lembra.

A professora propõe ver no caso uma oportunidade de ampliar a consciência das novas gerações sobre a linguagem, debatendo as diferenças com naturalidade em sala de aula, “porque aí também existe uma interculturalidade que não existia”.

Bingo. Como a que há na Inglaterra diante da influência avassaladora do inglês americano. Português europeu e português brasileiro são duas vertentes da mesma língua – uma falada por 10 milhões de pessoas, a outra por uma multidão mais de 20 (vinte!) vezes maior. Um amigo meu diz, com alguma maldade, que “sem o Brasil o português seria uma espécie de búlgaro”. Estamos diante de um fato que não se pode abafar, com sua história complexa – o imperialismo sempre inclui a volta do cipó de aroeira – e suas oportunidades de inseminação cultural. Aos interessados no tema, recomendo a leitura do texto integral do DN, que está aberto a não assinantes.

Disponível em <https://www1.folha.uol.com.br/colunas/sergio-rodrigues/2021/11/miudos-que-falam-brasileiro-em-portugal.shtml> Acesso em: 25 fev. 2022.

Para construção do sentido global do texto, é correto afirmar que o autor

(A) recorre a elementos da variedade linguística falada em Portugal, a fim de ironizar os usos em relação à variedade brasileira. Com isso, aponta que há uma língua brasileira e outra portuguesa e que, embora haja semelhanças

entre ambas, a primeira tem mais prestígio sobre a segunda, em razão de sua maior veiculação aos meios digitais.

- (B) utiliza palavras e expressões da variedade portuguesa, somente quando cita trechos da reportagem, sobre a qual faz comentários ao longo do artigo. Com isso, defende o posicionamento do jornal português acerca da necessidade de permanência da variedade europeia, ainda que o acuse de sensacionalista.
- (C) apresenta uma seleção lexical da variedade portuguesa, por meio de citação da reportagem em discussão e de trechos de sua autoria. Conforme o texto progride, defende um posicionamento reconciliador, ainda que com ressalvas, em torno da polêmica.
- (D) faz uso de palavras e expressões portuguesas, a fim de mostrar a diferença entre as variedades linguísticas. A partir disso, sugere a supremacia da variedade brasileira, dado o expressivo número de falantes em relação à lusitana, bem como sua presença massiva em meios digitais.

6 A construção de sentido de um texto se dá com base nas relações estabelecidas com os elementos do próprio texto e com elementos exteriores como a situação comunicativa e o diálogo com outras obras. Considerando essas informações, leia o texto abaixo e analise as afirmações que aparecem na sequência:

“Com a fé de quem olha do banco a cena
Do gol que nós mais precisava na trave
A felicidade do branco é plena
A pé, trilha em brasa e barranco, que pena
Se até pra sonhar tem entrave
A felicidade do branco é plena
A felicidade do preto é quase

Olhei no espelho, Ícaro me encarou
Cuidado, não voa tão perto do Sol
Eles num guenta te ver livre, imagina te ver rei
O abutre quer te ver de algema pra dizer:
Ó, num falei?!

No fim das conta é tudo Ismália, Ismália
Ismália, Ismália
Ismália, Ismália

Quis tocar o céu, mas terminou no chão
Ela quis ser chamada de morena

Que isso camufla o abismo entre si e a humanidade plena
A raiva insufla, pensa nesse esquema
A ideia imunda, tudo inunda
A dor profunda é que todo mundo é meu tema”

(Ismália – Emicida, 2019) - Adaptado

*Ícaro: personagem da mitologia grega

*Ismália: poema de Alphonsus de Guimaraens

Fonte do texto: www.letras.mus.br. Acesso em: 22 fev. 2022.

I – Pela leitura do texto de Emicida e considerando o contexto de produção da canção, é possível inferir que Ismália representa a mulher negra que sofre com a exclusão social.

II – A compreensão da relação intratextual estabelecida com a figura mitológica de Ícaro e com a personagem Ismália do poema de Alphonsus de Guimarães é fundamental para a compreensão do texto de Emicida.

III - A polissemia presente na palavra Ismália confere um efeito de sonoridade ao texto.

IV – O verso “Que isso camufla o abismo entre si e a humanidade plena” estabelece uma relação semântica de alternância com relação ao verso anterior.

V – Pelos versos da segunda estrofe, é possível inferir que o autor faz alusão ao preconceito existente com relação à ascensão social da população negra no Brasil.

É possível afirmar que estão **INCORRETAS** as sentenças expostas em:

(A) I, II, III e IV.

(B) I, III e IV.

(C) II, III e IV.

(D) II e III e V.

7 O microconto a seguir foi publicado na antologia de textos selecionados em um concurso de textos de humor. Considere-o para responder à questão.

– Vê uma completa!

Quando o cardiologista lhe perguntou se comia frutas e verduras, Jorjão inflou o peito e respondeu de bate-pronto:

- Couve e laranja toda quarta.
- Algo mais, Seu Jorge?
- Claro, Doutor. Também aos sábados.

Lucas Oliveira Campos Verzola | São Paulo/SP

Disponível em: <http://biblioteca.piracicaba.sp.gov.br/site/wp-content/uploads/2021/11/Microcontos-de-Humor-2021.pdf>. Acesso em: 23 fev. 2022.

Para produzir o efeito humorístico pretendido no texto, o autor recorre, principalmente

- (A) à variação em grau do nome do personagem, de modo a sugerir certos hábitos alimentares indesejados pelo cardiologista.
- (B) a um conhecimento cultural implícito sobre alimentação brasileira, que o autor supõe ser subentendido pelo leitor.
- (C) ao uso da expressão adverbial ‘de bate-pronto’, de modo a reproduzir aspectos da oralidade na interação entre médico e paciente.
- (D) ao uso dos vocativos – ‘seu Jorge’ e ‘Doutor’ – de modo a marcar uma relação assimétrica entre os personagens.

8 Observe a letra da canção “Pais e Filhos” – Renato Russo

Estátuas e cofres, e paredes pintadas
 Ninguém sabe o que aconteceu
 Ela se jogou da janela do quinto andar
 Nada é fácil de entender
 Dorme agora
 É só o vento lá fora
 Quero colo! Vou fugir de casa
 Posso dormir aqui com vocês?
 Estou com medo, tive um pesadelo
 Só vou voltar depois das três
 Meu filho vai ter nome de santo
 Quero o nome mais bonito
 É preciso amar as pessoas
 Como se não houvesse amanhã
 Porque se você parar pra pensar
 Na verdade, não há
 Me diz, por que que o céu é azul?
 Explica a grande fúria do mundo
 São meus filhos
 Que tomam conta de mim
 Eu moro com a minha mãe
 Mas meu pai vem me visitar
 Eu moro na rua, não tenho ninguém

Eu moro em qualquer lugar
 Já morei em tanta casa
 Que nem me lembro mais
 Eu moro com meus pais
 É preciso amar as pessoas
 Como se não houvesse amanhã
 Porque se você parar pra pensar
 Na verdade, não há
 Sou uma gota d’água
 Sou um grão de areia
 Você me diz que seus pais não o entendem
 Mas você não entende seus pais
 Você culpa seus pais por tudo, isso é um absurdo
 São crianças como você
 O que você vai ser
 Quando você crescer?

RUSSO, R. Pais e filhos. Disponível em: <https://www.letras.mus.br/renato-russo/75858/>. Acesso em: 20 fev. 2022.

A letra acima inicia-se com uma narrativa, mas ao longo do texto o que predomina é:

- (A) Justaposição de discursos diretos.
- (B) Descrição subjetiva.
- (C) Sequência de discursos indiretos.
- (D) Argumentação por comparação.

9 Considere os seguintes títulos de notícias:

(1) DECISÃO: TRF1 _____ pedido de novo bloqueio de valores no sistema BacenJud até o limite do crédito em execução

Fonte: <https://portal.trf1.jus.br>

(2) TSE cassa _____ de deputado do Paraná por infidelidade partidária

Fonte: <https://agenciabrasil.ebc.com.br>

(3) Brasil _____ Protocolo de Nagoya, mas não está claro como vai cumprir

Fonte: <https://revistagloborural.globo.com>

(4) A rocha Moutonnée, monumento geológico localizado em Salto/SP, representa um importante testemunho da glaciação ocorrida _____ 300 milhões de anos no planeta terra

Fonte: <https://www.infraestruturameioambiente.sp.gov.br>

O contexto das notícias permite assinalar o conjunto de palavras e expressões que melhor preencham os espaços:

- (A) (1) defere – (2) mandato – (3) ratifica – (4) há cerca de.
- (B) (1) difere – (2) mandato – (3) retifica – (4) há cerca de.
- (C) (1) defere – (2) mandato – (3) ratifica – (4) a cerca de.
- (D) (1) defere – (2) mandado – (3) retifica – (4) acerca de.

10 Segundo o *Manual de redação da Presidência da República*, de 2018, um dos problemas da redação oficial são os chamados erros de paralelismo. Esses dizem respeito à inconsistência na apresentação de ideias gramaticais. Dentre os enunciados que seguem, o único que atende a essa recomendação, bem como está de acordo com a norma-padrão, é:

- (A) Pelo ofício circular, recomendaram-se aos Ministérios economizar energia e que elaborassem planos de redução de despesas.
- (B) Pelo ofício circular, recomendou-se aos Ministérios que economizassem energia e elaborassem planos de redução de despesas.
- (C) Pelo ofício circular, recomendaram-se aos Ministérios a economia de energia e que elaborassem planos de redução de despesas.
- (D) Pelo ofício circular, recomendou-se aos Ministérios que se economizasse energia e a elaboração de planos de redução de despesas.

11 A educação é um direito fundamental que está previsto na Carta Magna brasileira. Refletir sobre a educação enquanto dever do Estado é contribuir para promover a igualdade entre os cidadãos e cidadãs.

Consoante o art. 208, da Constituição Federal de 1988, o dever do Estado com a educação será efetivado mediante a garantia de:

- (A) educação básica obrigatória e gratuita dos 7 (sete) aos 17 (dezessete) anos de idade.
- (B) atendimento educacional especializado aos portadores de deficiência, exclusivamente na rede regular de ensino.
- (C) oferta de ensino noturno regular, adequado às condições do estudante.
- (D) acesso aos níveis mais elevados de ensino, da pesquisa e da criação artística, segundo a idade de cada um.

12 De acordo com a LDB (Lei 9394/96), a educação profissional técnica de nível médio será desenvolvida nas formas: articulada com o ensino médio e subsequente, em cursos para quem já tenha concluído o ensino médio. Com relação à educação profissional técnica de nível médio articulada, marque a alternativa que corresponde a uma forma em que poderá ser desenvolvida:

- (A) Integrada, oferecida somente a quem já tenha concluído o ensino médio, sendo o curso planejado de modo a conduzir o aluno à habilitação profissional técnica de nível médio, na mesma instituição de ensino, efetuando-se matrícula única para cada aluno.
- (B) Concomitante, oferecida a quem ingresse no ensino médio ou já o esteja cursando, efetuando-se uma única matrícula para cada aluno, e podendo ocorrer na mesma instituição de ensino, aproveitando as oportunidades de ensino disponíveis.
- (C) Concomitante, oferecida a quem ingresse no ensino médio ou já o esteja cursando, efetuando-se matrículas distintas para cada curso, não podendo ocorrer em instituições distintas de ensino, aproveitando as oportunidades de ensino disponíveis.
- (D) Concomitante, oferecida a quem ingresse no ensino médio ou já o esteja cursando, efetuando-se matrículas distintas para cada curso, e podendo ocorrer em instituições de ensino distintas, mediante convênios de intercomplementaridade, visando ao planejamento e ao desenvolvimento de projeto pedagógico unificado.

13 A lei nº 8.429, de 2 de junho de 1992, em um de seus artigos diz o seguinte sobre os atos de improbidade administrativa:

“receber, para si ou para outrem, dinheiro, bem móvel ou imóvel, ou qualquer outra vantagem econômica, direta ou indireta, a título de comissão, percentagem, gratificação ou presente de quem tenha interesse, direto ou indireto, que possa ser atingido ou amparado por ação ou omissão decorrente das atribuições do agente público.”

O agente público que se enquadre na situação descrita acima, independentemente do ressarcimento integral do dano patrimonial, se efetivo, e das sanções penais comuns e de responsabilidade, civis e administrativas previstas na legislação es-

pecífica, estará sujeito à seguinte cominação, que pode ser aplicada isolada ou cumulativamente, de acordo com a gravidade do fato:

- (A) perda dos bens ou valores acrescidos ilicitamente ao patrimônio, perda da função pública, suspensão dos direitos políticos até 12 (doze) anos, pagamento de multa civil equivalente ao valor do acréscimo patrimonial e proibição de contratar com o poder público ou de receber benefícios ou incentivos fiscais ou creditícios, direta ou indiretamente, ainda que por intermédio de pessoa jurídica da qual seja sócio majoritário, pelo prazo não superior a 12 (doze) anos.
- (B) perda dos bens ou valores acrescidos ilicitamente ao patrimônio, perda da função pública, suspensão dos direitos políticos até 12 (doze) anos, pagamento de multa civil equivalente ao valor do acréscimo patrimonial e proibição de contratar com o poder público ou de receber benefícios ou incentivos fiscais ou creditícios, direta ou indiretamente, ainda que por intermédio de pessoa jurídica da qual seja sócio majoritário, pelo prazo não superior a 14 (catorze) anos.
- (C) perda dos bens ou valores acrescidos ilicitamente ao patrimônio, perda da função pública, suspensão dos direitos políticos até 14 (catorze) anos, pagamento de multa civil equivalente ao valor do acréscimo patrimonial e proibição de contratar com o poder público ou de receber benefícios ou incentivos fiscais ou creditícios, direta ou indiretamente, ainda que por intermédio de pessoa jurídica da qual seja sócio majoritário, pelo prazo não superior a 12 (doze) anos.
- (D) perda dos bens ou valores acrescidos ilicitamente ao patrimônio, perda da função pública, suspensão dos direitos políticos até 14 (catorze) anos, pagamento de multa civil equivalente ao valor do acréscimo patrimonial e proibição de contratar com o poder público ou de receber benefícios ou incentivos fiscais ou creditícios, direta ou indiretamente, ainda que por intermédio de pessoa jurídica da qual seja sócio majoritário, pelo prazo não superior a 14 (catorze) anos.

14 O tratamento de dados pessoais de crianças e de adolescentes deverá ser realizado em seu melhor interesse, nos termos deste artigo e da legislação pertinente, para tal:

- (A) deverá ser realizado com o consentimento específico e em destaque dado por pelo menos um dos pais ou pelo responsável legal.
- (B) os controladores deverão manter em sigilo a informação sobre os tipos de dados coletados, a forma de sua utilização e os procedimentos.
- (C) poderão ser coletados dados pessoais de crianças sem o consentimento específico de pelo menos um dos pais ou responsável legal quando a coleta for necessária para contatar os pais ou o responsável legal, utilizados e guardados em bancos de dados próprios informando a quem manifestar interesse pelas informações.
- (D) o controlador deve realizar todos os máximos esforços para verificar que o consentimento a que se dado por pelo menos um dos pais foi dado pelo responsável pela criança, considerada a confiança em quem prestou a informação.

15 Com base na **Lei Nº 11.892, de 29 de dezembro de 2008**, marque a alternativa que NÃO indica uma das finalidades da Rede Federal de Educação Profissional, Científica e Tecnológica:

- (A) promover a produção, o desenvolvimento e a transferência de tecnologias sociais, notadamente as voltadas à preservação do meio ambiente.
- (B) constituir-se em centro de referência na produção de tecnológica, em geral, e, estimulando o desenvolvimento de espírito competitivo na formação profissional.
- (C) realizar e estimular a pesquisa aplicada, a produção cultural, o empreendedorismo, o cooperativismo e o desenvolvimento científico e tecnológico.
- (D) desenvolver programas de extensão e de divulgação científica e tecnológica.

16 Considere as seguintes afirmações:

- 1 – O produto de um número real por um número irracional é sempre um número irracional;
- 2 – O produto de um número irracional pelo seu inverso, é sempre um número natural;
- 3 – A soma de dois números irracionais é sempre um número irracional.

É correto afirmar que as alternativas verdadeiras são:

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

17 Recentemente, a particularidade de uma data do mês de fevereiro de 2022 teve grande repercussão nas mídias sociais.

As pesquisas sobre o significado de “22/02/2022” para a astrologia e numerologia dispararam nesta terça-feira (22), e centenas de casais aproveitaram para ir ao cartório e começar uma vida nova em um dia tão singular. Mas, afinal, o que a data significa? Nada, necessariamente. Mas o dia é especial para quem gosta de simetrias e curiosidades com números – não só para os supersticiosos [...]. Para datas simétricas, é preciso ignorar espaços, traços e barras (22/02/2022, então, fica como 2 2 0 2 2 0 2 2).

COSTA, L. Palíndromos e ambigramas: por que a data 22/02/2022 é tão especial? Superinteressante. Disponível em <https://super.abril.com.br/cultura/palindromos-e-ambigramas-por-que-a-data-22-02-2022-e-tao-especial/>. Acesso em: 28 fev. 2022.

A data em questão exemplifica um palíndromo: algo que independente da ordem da leitura (tanto da esquerda para a direita quanto o contrário) pode ser compreendido da mesma forma. É claro que, no caso de datas, tal como consta no trecho da matéria publicada no site da Revista Superinteressante indicada acima, é preciso ignorar as barras.

No quadro a seguir, estão reunidas as datas do período 2001-2070 que também são palíndromos.

Períodos	2000-2009	2010-2019	2020-2029	2030-2039	2040-2049	2050-2059	2060-2069
Datas	10/02/2001	01/02/2010	02/02/2020	03/02/2030	04/02/2040	05/02/2050	06/02/2060
	20/02/2002	11/02/2011	12/02/2021 22/02/2022	13/02/2031 23/02/2032	14/02/2041 24/02/2042	15/02/2051 25/02/2052	16/02/2061 26/02/2062

De acordo com as informações presentes no quadro, é correto afirmar que:

- (A) Nas duas primeiras décadas, a probabilidade de ocorrer um palíndromo nas datas é de 20%.
- (B) Nas quatro primeiras décadas, a probabilidade de ocorrer um palíndromo nas datas é menor que 25%.
- (C) Nas duas últimas décadas, a probabilidade de ocorrer um palíndromo nas datas é maior que 30%.
- (D) No período analisado (2000-2069) a probabilidade de ocorrer um palíndromo nas datas é 19%.

18 Na madrugada do dia 24 de fevereiro de 2022, após a decisão do presidente russo, Vladimir Putin, teve início uma invasão da Rússia ao território da Ucrânia. Na tentativa de buscar informações a respeito dessa guerra, um estudante acessou, via link encaminhado para o seu aplicativo de mensagens, um *site* de buscas que, até então, ele não conhecia.

No campo de buscas desse *site*, ele digitou UCRANIA, sem acento, e clicou em “pesquisar”. Como resultado, surgiu a seguinte mensagem:

Não encontramos resultados para a busca por “ABQEMOE”

Após algum tempo olhando a tela, o estudante entendeu o que aconteceu: por algum motivo desconhecido, as letras que ele digitou no campo de pesquisa do site foram modificadas. Entretanto, o que ocorreu com as vogais digitadas foi diferente do que ocorreu com as consoantes. Ele, então, atualizou a página e fez uma outra tentativa de busca, agora utilizando o nome do presidente ucraniano (que estava constantemente na mídia), digitando VOLODYMYR ZELENSKY. Ao clicar em “pesquisar”, o mesmo problema ocorreu.

Sabendo disso, é correto que o resultado visto pelo estudante tenha sido:

- (A) Não encontramos resultados para a busca por “TUKUCXLXQ YIKIMRJX”.
- (B) Não encontramos resultados para a busca por “TIKICXLXQ YAKAMRJX”.
- (C) Não encontramos resultados para a busca por “WUMUFZNZS BIMIPTLZ”.
- (D) Não encontramos resultados para a busca por “WIMIFZNZS BAMAPTLZ”.

19 Um grupo de 6 (seis) amigos está em viagem pelo interior e encontra a placa a seguir, com indicação dos caminhos:



Fonte: IFSP, 2022.

Para definir qual caminho cada um seguiria, decidiram que cada um jogaria um dado para o alto e seguiria o caminho correspondente ao número indicado pelo dado, e, eventualmente, estariam separados no restante da viagem. Considerando os números atribuídos a cada destino (1 a 6), conforme a imagem, qual a probabilidade de todos irem para o mesmo destino?

- (A) $\frac{1}{6}$
- (B) $\frac{1}{6^6}$
- (C) $\frac{1}{6^5}$
- (D) $\frac{1}{36}$

20 Com as letras da palavra BRASIL, podemos formar, com ou sem sentido em língua portuguesa, 720 palavras diferentes. Ao se fazer um ordenamento alfabético dessas possíveis formações de letras em palavras iremos encontrar, em primeiro lugar, a palavra ABILRS e, em último, a palavra SRLIBA. Seguindo essa ordenação, podemos afirmar que a palavra IABRSL estará na posição:

- (A) 241º
- (B) 242º
- (C) 243º
- (D) 244º

21 O Microsoft Word é um processador de texto desenvolvido pela Microsoft, o qual possui diversos recursos que permitem a criação de documen-

tos profissionais. Um desses recursos permite que o documento possa ser dividido em partes independentes. Assinale a alternativa que indica o recurso responsável por permitir a inserção de cabeçalhos diferentes dentro das páginas de um mesmo documento do Microsoft Word.

- (A) Quebra de página.
- (B) Quebra de seção.
- (C) Inserir página em branco.
- (D) Inserir cabeçalho.

22 Os ataques virtuais estão sempre evoluindo em sua forma e aplicação, conforme novas vulnerabilidades de sistemas são descobertas. Um tipo de ataque virtual que vêm se popularizando e que acontece quando um software malicioso obtém acesso a uma rede, se espalha e criptografa as informações a que tem acesso, e depois cobra um valor de resgate pela chave de decifração da informação roubada, é chamado de:

- (A) Backdoor.
- (B) Spyware.
- (C) Ransomware.
- (D) Adware.

23 O envio de e-mails facilita a comunicação e registro de ações dentro de uma instituição, mas um envio para destinatários de maneira errada pode expor a privacidade de endereços de destinatários para quem não deveria ter acesso a essas informações. Sobre o controle de privacidade de destinatários de um e-mail, considere as afirmações abaixo:

- I – O campo “Para” é onde se coloca os destinatários principais da mensagem;
- II – O campo “Cc” serve para enviar uma cópia da mensagem para um ou mais destinatários, mas oculta os destinatários que receberam as cópias da mensagem;
- III – O campo “Para” pode ter somente um destinatário;
- IV – O campo “Cco” envia cópia da mensagem para os destinatários e deixa público todos os endereços que receberam a mensagem;
- V – O campo “Cco” oculta os endereços dos destinatários que receberam a mensagem.

São CORRETAS as afirmações:

- (A) As afirmações I e V.
- (B) As afirmações I, II e IV.
- (C) As afirmações I e II.
- (D) As afirmações III e V.

24 Como forma de organizar o setor de licitações, o Coordenador de Licitações e Contratos do IFSP solicitou ao seu subordinado que criasse no Windows uma estrutura de diretórios conforme ilustrado na figura a seguir. O subordinado encontrou dificuldades para concluir a tarefa e não conseguiu criar uma das pastas solicitadas. Todas as demais foram criadas com sucesso.

- 1 Coordenadoria de Licitações e Contr@to\$
- 2 Cobranças =\$=
- 3 'Comunicados'
- 4 \\Concorrência//
- 5 Preg@o!!!

Fonte: IFSP, 2022.

Dessa forma, é possível afirmar que existem alguns caracteres especiais que não são permitidos para nomear e renomear diretórios e arquivos no Windows. Considerando o caso relatado, assinale a alternativa em que todos os caracteres são proibidos:

- (A) = \$!
- (B) \ / @
- (C) \ / =
- (D) \ / :

25 O setor de gestão de pessoas do IFSP deseja automatizar uma planilha de controle de servidores em estágio probatório. Conforme a imagem abaixo, indique qual função deve ser construída na célula D9 para retornar o total de servidores em estágio probatório.

	A	B	C	D
1				
2	SERVIDOR	DATA POSSE	DATA EXERCÍCIO	ESTÁGIO PROBATÓRIO
3	João Arantes	02/03/2018	17/03/2018	NÃO
4	Fabiano Luiz	15/10/2015	26/10/2015	NÃO
5	Ana Clara Lins	21/11/2020	07/11/2020	SIM
6	Paulo Teixeira	13/01/2013	15/01/2013	NÃO
7	Juliana Filgueiras	30/08/2021	09/09/2021	SIM
8				
9	SERVIDORES EM ESTÁGIO PROBATÓRIO			2

Fonte: IFSP, 2022.

- (A) =SOMA(SE((HOJE()-C3)<1095;"SIM";"NÃO"))
- (B) =CONT.SE(D3:D7;"SIM")
- (C) =SOMASE(D3:D7;"SIM";D3:D7)
- (D) =SOMA(D3:D7;"SIM")

CONHECIMENTOS ESPECÍFICOS

26 Trabalhar com álgebra linear está associado aos vetores e matrizes, com suas regras e cálculos. Para isso, é necessário conhecer tanto a matemática envolvida quanto os recursos da linguagem de programação. Assinale a alternativa correta que implementa o cálculo algébrico apresentado:

$$A = \begin{bmatrix} 10 & 20 & 30 \\ 40 & 50 & 60 \\ 70 & 80 & 90 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Sendo $D = (A+B) \cdot C$

(A)

```
import numpy as np
A = np.arange(10,100,10).reshape(3,3)
B = np.arange(1,9,1).reshape(3,3)
C = np.eye(3)
D = (A+B)*C
print(D)
```

(B)

```
import numpy as np
A = np.arange(10,100,10).reshape(3,3)
B = np.arange(1,9,1).reshape(3,3)
C = np.eye(3)
D = A+B*C
print(D)
```

(C)

```
import numpy as np
A = np.arange(10,100,10).reshape(3,3)
B = np.linspace(1,9,9).reshape(3,3)
C = np.eye(3)
D = (A+B)*C
print(D)
```

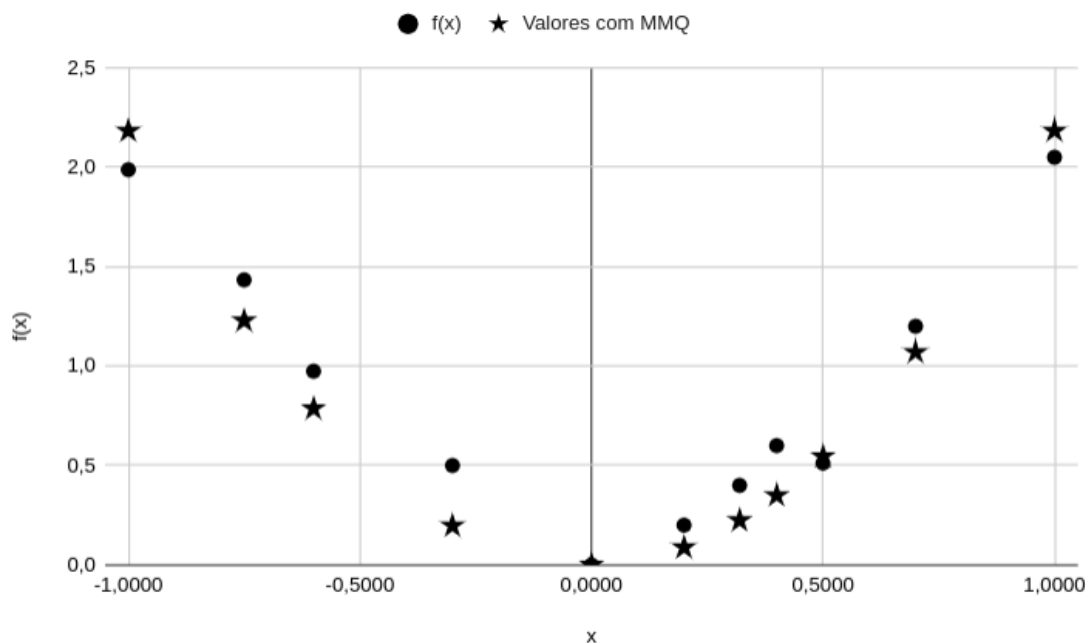
(D)

```
import numpy as np
A = np.arange(10,100,10).reshape(3,3)
B = np.linspace(1,9,9).reshape(3,3)
C = np.ones(3)
D = (A+B)*C
print(D)
```

27 Os métodos de regressão linear são utilizados para se buscar soluções por polinômios ou métodos iterativos que possibilitem estimar os valores de $f(x)$, considerando-se um certo x de entrada. Uma das soluções é pelo Método dos Mínimos quadrados, que consiste em escolher os α de tal forma que a soma dos quadrados dos desvios seja mínima, ou seja, de acordo com o critério dos quadrados mínimos, os coeficientes α_k , que fazem com que $\varphi(x)$ se aproxime ao máximo de $f(x)$ descrito por

$$\sum_{k=1}^m d_k^2 = \sum_{k=1}^m (f(x_k) - \varphi(x_k))^2$$

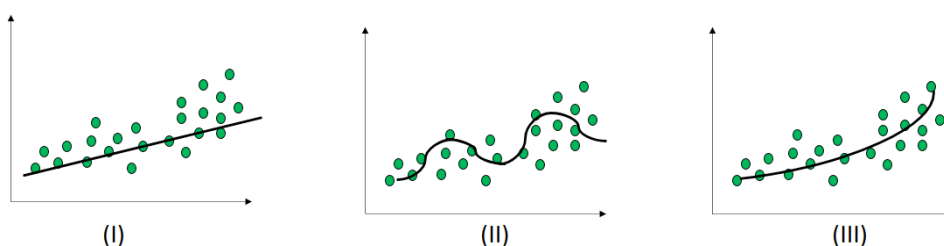
Assinale a alternativa em que se tem valor mais aproximado do cálculo pelo método dos mínimos quadrados, considerando $\alpha = 2,1831$, $x = 0,32$ e o $f(x) = x^2$. Repare que os dados dessa análise podem ser visualizados no gráfico a seguir.



Fonte: IFSP, 2022.

- (A) 0,2235
- (B) 0,1437
- (C) 0,4270
- (D) 0,6540

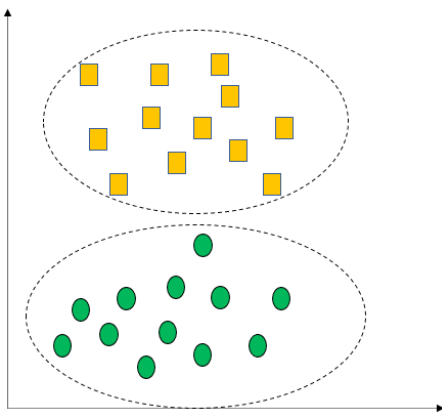
28 Nos modelos de aprendizado de máquina, existem dois problemas básicos que devem ser considerados. O primeiro problema é o sobreajuste (*overfitting*) que acontece quando se tem um modelo com bom desempenho com os dados treinados, mas que não trabalha bem com novos dados. Já o segundo problema, é o sub-ajuste (*underfitting*) que já sequer trabalhar com os dados de treino e, conseqüentemente, na aplicação em si. Para isso, é necessário implementar modelos que sejam equilibrados para atender as demandas. Considerando essas afirmações e a figura a seguir, selecione a alternativa que melhor associa o problema com os dados dispostos no espaço com a solução equilibrada, com sub-ajuste e sobreajuste de acordo com sua indicação (I, II e III).



Fonte: IFSP, 2022.

- (A) (I) equilibrado, (II) sobreajustado e (III) sub-ajustado
- (B) (I) equilibrado, (II) sub-ajustado e (III) sobreajustado
- (C) (I) sobreajustado, (II) equilibrado e (III) sub-ajustado
- (D) (I) sub-ajustado, (II) sobreajustado e (III) equilibrado

29 No aprendizado de máquina, é necessário conhecer os modelos e os recursos numéricos que melhor se adequam ao problema de interesse. Em problemas de agrupamento, o modelo deve possuir recursos para realizar medições intra-agrupamento e também entre os elementos de um mesmo grupo. Considerando a figura a seguir, descreva qual modelo matemático apresenta uma solução para retornar a distância média entre os objetos dos dois grupos.



Fonte: IFSP, 2022.

- (A)
$$d(C_1, C_2) = \frac{1}{n_1 \cdot n_2} \cdot \prod_{x_i \in C_1; x_j \in C_2} d(x_i, x_j)$$
- (B)
$$d(C_1, C_2) = \frac{1}{n_1 \cdot n_2} \cdot \sum_{x_i \in C_1; x_j \in C_2} d(x_i, x_j)$$
- (C)
$$d(C_1, C_2) = \prod_{x_i \in C_1; x_j \in C_2} d(x_i, x_j)$$
- (D)
$$d(C_1, C_2) = \sum_{x_i \in C_1; x_j \in C_2} d(x_i, x_j)$$

30 O algoritmo K-NN (*K Nearest Neighbor*) é um dos algoritmos mais utilizados em *Machine Learning*, além de ser simples se comparado aos demais. Esse algoritmo é muito utilizado em problemas de classificação, mas pode ser aplicado para regressão. Selecione a opção que NÃO representa uma prática eficiente para preparação dos dados com a aplicação desse algoritmo.

- (A) Normalização dos dados.
- (B) Remoção de dados ausentes (sem valor).
- (C) Divisão dos dados em 80% para treinamento e 20% para testes.
- (D) Trabalhar com poucas entradas (dimensionalidade).

31 A árvore de decisão é um tipo de algoritmo de aprendizagem de máquina supervisionado, que se baseia na ideia de divisão dos dados em grupos homogêneos, e é utilizado para classificação ou regressão. O objetivo da árvore de decisão é encontrar o atributo que gera a melhor divisão dos dados, isto é, para gerar subconjuntos com maior “pureza”. Selecione a alternativa que NÃO apresenta métrica empregada para a definição do melhor atributo para dividir os dados.

- (A) Coeficiente de Gini.
- (B) Entropia.
- (C) Redução de variância.
- (D) Poda de nós.

32

```
import datetime
d = datetime.datetime(2022,1,2,3,4,5)
print(type(d))
print(d.strftime("%j"))
```

Com base no algoritmo, considere as afirmativas:
I – A primeira linha realiza a importação do módulo ‘datetime’

II – A segunda linha cria um objeto

III – A terceira linha retorna <class ‘datetime.datetime’>

IV – A última linha imprime o valor 002

Assinale a alternativa correta:

- (A) Todas as afirmações estão corretas.
- (B) Somente III e IV estão corretas.
- (C) Somente I, II e III estão corretas.
- (D) Somente I e II estão corretas.

33 Uma tarefa muito comum durante a etapa de pré-processamento de dados é o tratamento de valores ausentes. A linguagem Python possui uma biblioteca muito utilizada pelos cientistas de dados, denominada Pandas, que permite realizar o processo de transformação dos dados de maneira bem prática. Considere os dados exibidos pela figura abaixo:

	Gravidez	Glicose	Pressão Sanguínea	Espessura da Pele	IMC	Idade	Resultado
0	6.00	148	72.00	35.00	33.60	50	1.00
1	1.00	85	66.00	29.00	26.60	31	NaN
2	8.00	183	64.00	NaN	23.30	32	1.00
3	1.00	89	66.00	23.00	28.10	21	NaN
4	NaN	137	40.00	35.00	43.10	33	1.00
5	5.00	116	74.00	NaN	25.60	30	NaN
6	3.00	78	50.00	32.00	31.00	26	1.00
7	10.00	115	NaN	NaN	35.30	29	NaN
8	2.00	197	70.00	45.00	30.50	53	1.00
9	8.00	125	96.00	NaN	NaN	54	1.00

Fonte: IFSP, 2022.

Pela figura, é possível observar que existem diversos valores ausentes, identificados por *NaN*. Ao analisar essas informações, o cientista de dados decide substituir os valores ausentes em cada coluna, da seguinte forma:

- **Gravidez:** substituir *NaN* pelo valor 0;
- **Pressão Sanguínea:** substituir *NaN* pela média dos valores da coluna;
- **Espessura da Pele:** substituir *NaN* pela moda dos valores da coluna;
- **IMC:** substituir *NaN* pela mediana dos valores da coluna;
- **Resultado:** substituir *NaN* pelo valor 0.

O cientista de dados armazenou os valores dentro de um *dataframe* do Pandas, chamado *df*. Para realizar a substituição dos valores ausentes, ele decide executar o seguinte comando:

```
df.fillna(value=valores, inplace=True)
```

Observe que o argumento “value” recebe um dicionário do Python. Esse dicionário contém as instruções para atualizar os valores das colunas, tal como o cientista de dados deseja. Assinale a alternativa correta, que indica o trecho de código relacionado a esse dicionário de dados.

(A) valores = {
 Gravidez: 0,
 Pressão Sanguínea: df[‘Pressão Sanguínea’].mean(),
 Espessura da Pele: df[‘Espessura da Pele’].mode()[0],
 IMC: df[‘IMC’].median(),
 Resultado: 0
 }

(B) valores = {
 ‘Gravidez’: 0,
 ‘Pressão Sanguínea’: df[‘Pressão Sanguínea’].mean(),
 ‘Espessura da Pele’: df[‘Espessura da Pele’].mode(),
 ‘IMC’: df[‘IMC’].median(),
 ‘Resultado’: 0
 }

```
(C) valores = {
    'Gravidez': 0,
    'Pressão Sanguínea': df['Pressão Sanguínea'].mean(),
    'Espessura da Pele': df['Espessura da Pele'].mean(),
    'IMC': df['IMC'].mean(),
    'Resultado': 0
}
```

```
(D) valores = {
    'Gravidez': 0,
    'Pressão Sanguínea': df['Pressão Sanguínea'].mean(),
    'Espessura da Pele': df['Espessura da Pele'].mode()[0],
    'IMC': df['IMC'].median(),
    'Resultado': 0
}
```

34 Você recebeu um chamado para colaborar no desenvolvimento de um dos módulos do software acadêmico. O referido módulo deve apresentar um gráfico estatístico para simplificar a interpretação dos dados e facilitar a tomada de decisões. Anexo ao chamado, consta um exemplo de uma matriz de dados extraída do Sistema Gerenciador de Banco de Dados oficial da instituição, possuindo dados relacionados à altura, peso, idade, renda familiar e número de reprovações. Assinale a alternativa que representa o tipo de gráfico mais adequado para análise dos referidos dados.

- (A) Pizza.
- (B) Tendência.
- (C) Dispersão.
- (D) Histograma.

35 *Machine Learning*, ou aprendizado de máquina, é um subcampo da inteligência artificial baseada na ideia de que os sistemas podem aprender a identificar padrões e tomar decisões, por meio da análise de dados. De acordo com o tipo de técnica utilizada no processo de aprendizagem, existem vários algoritmos que podem ser aplicados para gerar o modelo de aprendizado. Os parâmetros desse modelo podem ser atualizados por meio de técnicas de otimização. Com base nessas informações, assinale a alternativa correta.

(A) O gradiente descendente é um algoritmo de otimização utilizado pelos cientistas de dados para encontrar um máximo local de uma função diferenciável, movendo-se iterativamente na direção da descida mais íngreme, conforme definido pelo negativo do gradiente.

- (B) O método do gradiente descendente é utilizado pelos cientistas de dados para encontrar os valores de parâmetros de uma função, que reduzem ao máximo uma função de custo.
- (C) O gradiente descendente é um algoritmo de otimização iterativa, que necessita do cálculo da segunda derivada para poder realizar atualizações dos parâmetros do modelo preditivo.
- (D) No método do gradiente descendente, o tamanho do passo (também conhecido como taxa de aprendizado, ou *learning rate*) não exerce nenhuma influência no resultado.

36 `select payment_id, case when not exists (select 1 from customer where customer_id=1 limit 1) then 0 else payment_id end as payment from payment;`

Com base no comando SQL, considere as afirmativas:

- I – Trata-se de uma consulta que utiliza duas tabelas;
- II – Serão retornadas duas colunas, ambas nomeadas ‘payment_id’;
- III – O comando retornará, no máximo, um registro;
- IV – O comando contém erro de sintaxe.

Assinale a alternativa correta.

- (A) Somente I está correta
- (B) Somente I e II estão corretas
- (C) Somente II e III estão corretas
- (D) Somente IV está correta

37 Uma amostra é um subconjunto da população em estudo. Ela representa a população maior e é usada para fazer inferências sobre essa população. A amostragem probabilística é uma técnica de amostragem; nela, as amostras são recolhidas em um processo que dá a todos os indivíduos da população as mesmas chances de serem selecionados. Nesse contexto, associe a técnica de amostragem probabilística com o exemplo apresentado e selecione a alternativa mais apropriada.

1 - Amostra aleatória simples

A - Para obter uma amostra de estudantes universitários, o pesquisador precisaria dividir em classes de amostras. Primeiro organizaria a população por semestre de graduação e, então, selecionaria determinado número de representantes de calouros, pessoas que estão no meio do curso e formandos. Isso garantiria que o pesquisador tivesse quantidades adequadas de indivíduos de cada classe na amostra final.

2 - Amostra sistemática

B - Numa população de 1000 pessoas, você gostaria de escolher uma amostra aleatória de 50 pessoas. Primeiro, cada pessoa é numerada de 1 até 1000. Então, você gera uma lista de 50 números aleatórios e os números dessa lista serão os únicos que você incluirá na amostra.

3 - Amostra Estratificada

C - Uma população de estudo contém 2000 estudantes do ensino fundamental e o pesquisador quer uma amostra de 100 estudantes. Os estudantes poderiam ser colocados em uma lista e cada 20º estudante seria selecionado para inclusão na amostra. A fim de evitar o viés humano nesse método, o pesquisador deve selecionar o primeiro elemento aleatoriamente.

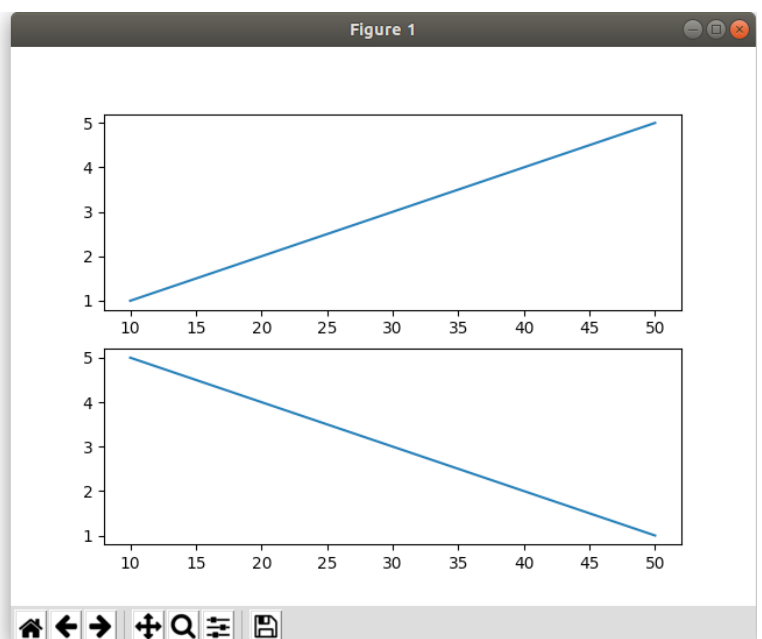
4 - Amostra por aglomerados

D - Suponha que a população-alvo em um estudo seja membros de igrejas nos EUA. Não há uma lista de todos os membros de igrejas no país. O pesquisador poderia, nesse caso, criar uma lista de igrejas nos EUA, escolher uma amostra de igrejas e, então, obter listas de membros dessas igrejas.

- (A) 1C, 2B, 3D e 4A
- (B) 1A, 2B, 3D e 4C
- (C) 1B, 2C, 3A e 4D
- (D) 1D, 2A, 3C e 4B

38 A biblioteca Matplotlib é usada para plotagem de gráficos 2D em Python. A função subplot dessa biblioteca é usada para criação de uma figura com subplots.

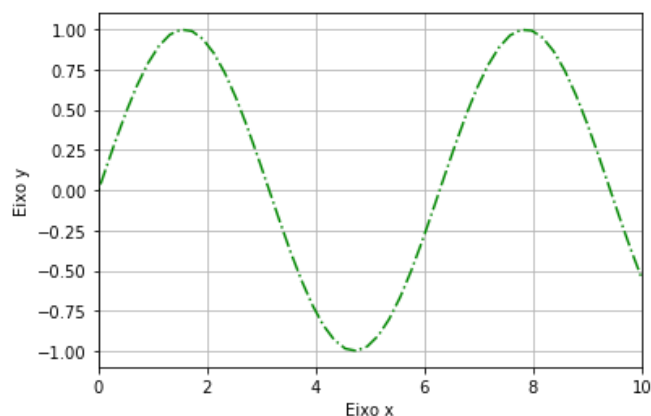
```
1 import matplotlib.pyplot as plt
2 import numpy as np
3
4 x = np.array([10, 20, 30, 40, 50])
5 y = np.array([1, 2, 3, 4, 5])
6
7 plt.subplot(2,1,1)
8 plt.plot(x,y)
9
10 x = np.array([10, 20, 30, 40, 50])
11 y = np.array([5, 4, 3, 2, 1])
12
13 plt.subplot(2,1,2)
14 plt.plot(x,y)
15
16 plt.show()
17
```



Considerando a imagem apresentada, assinale a alternativa que representa qual das sequências deveria ser usada na função subplot nas linhas 7 e 13, respectivamente, para geração dos gráficos apresentados ao lado do código.

- (A) 1, 2, 1 e 1, 2, 2
- (B) 2, 1, 2 e 2, 2, 1
- (C) 2, 1, 1 e 2, 1, 2
- (D) 2, 1, 2 e 2, 1, 1

39 A variedade de recursos para a visualização de dados é um dos diferenciais que atraem o interesse sobre a linguagem de programação Python. Para isso, existem diferentes bibliotecas disponíveis na Internet, com destaque a biblioteca Matplotlib. Dentre os inúmeros recursos dessa biblioteca, o método plot disponibiliza recursos para a visualização bidimensional de dados, como o do gráfico a seguir.



Fonte: IFSP, 2022.

Considerando esse gráfico gerado utilizando programação Python, selecione a opção que possibilita gerá-lo corretamente de acordo com a visualização apresentada nesta imagem. Para isso, considerar que a curva é uma senoide gerada pelo trecho de código a seguir:

```
import numpy as np
import matplotlib.pyplot as plt
x = np.linspace(-10, 10, 100)
y = np.sin(x)
```

(A)

```
plt.plot(x,y,'g-.',label="Seno(x)")
plt.grid(True)
plt.xlabel("Eixo x")
plt.ylabel("Eixo y")
```

(B)

```
plt.plot(y,x,'g-.',label="Seno(x)")
plt.grid(True)
plt.xlabel("Eixo x")
plt.ylabel("Eixo y")
```

(C)

```
plt.plot(y,x,'g-.',label="Seno(x)")
plt.grid(True)
plt.xlim(0,10)
plt.xlabel("Eixo x")
plt.ylabel("Eixo y")
```

(D)

```
plt.plot(x,y,'g-.',label="Seno(x)")
plt.grid(True)
plt.xlim(0,10)
plt.xlabel("Eixo x")
plt.ylabel("Eixo y")
```

40 Entre os modelos de aprendizado de máquina mais comuns, estão as árvores de decisão. Elas são métodos de aprendizado de máquinas muito utilizados em tarefas de classificação e regressão. Em problemas de classificação, os modelos em árvore são designados de árvore de decisão. Para resolver um problema de decisão, esse tipo de método utiliza a estratégia de dividir para conquistar. Uma proposta natural é rotular cada conjunto da divisão por sua classe mais frequente e escolher a divisão que tem menores erros. O conceito fundamental nessa proposta é denominado de entropia. Considerando as árvores de decisão, assinale a alternativa que define corretamente o conceito de entropia.

- (A) A entropia é uma medida que representa a soma de todos os valores, dividida pelo número de valores do conjunto de dados.
- (B) A entropia é uma medida que representa a probabilidade de obter uma ocorrência do evento, a partir de uma seleção aleatória do subconjunto de dados.
- (C) A entropia é uma medida que representa a diferença entre o maior e o menor valor, dentro de um conjunto de dados.
- (D) A entropia é uma medida que representa a categoria, ou o valor, de maior ocorrência em um conjunto de dados.

41 Uma tarefa muito comum durante a etapa de pré-processamento de dados é o tratamento de valores ausentes. Na linguagem R, os valores ausentes são representados por *NA* (*Not Available*). Considere os dados exibidos pela figura abaixo:

	Gravidez	Glicose	Pressão Sanguínea	Idade	Resultado
	<int>	<int>	<int>	<int>	<int>
1	6	148	72	50	1
2	1	85	66	31	NA
3	8	183	64	32	1
4	1	89	66	21	NA
5	NA	137	40	33	1
6	5	116	74	30	NA
7	3	78	50	26	1
8	10	115	NA	29	NA
9	2	197	70	53	1
10	8	125	96	54	1

Fonte: IFSP, 2022.

Esses dados correspondem aos valores que foram carregados e armazenados em um *dataframe* da linguagem R. A linguagem R, assim como a linguagem Python, é muito utilizada na área de Ciência de Dados. Ela oferece diversas bibliotecas que podem ser empregadas para auxiliar nas etapas de pré-processamento e transformação dos dados. Ao analisar as informações exibidas pela figura, o cientista de dados percebe a existência de diversos valores ausentes e decide substituí-los da seguinte forma:

- **Gravidez:** substituir *NA* pelo valor 0;
- **Pressão Sanguínea:** substituir *NA* pela média dos valores da coluna;
- **Resultado:** substituir *NA* pelo valor 0.

Assinale a alternativa correta, que indica o trecho de código escrito em linguagem R, que pode ser utilizado para realizar essa tarefa.

(A)

```
df <- df %>%
  mutate (
    across(c('Gravidez','Resultado'), ~replace_na(.x, 0)),
    across('Pressão Sanguínea', ~replace_na(.x, mean(.x, na.rm=TRUE)))
  )
```

(B)

```
df <- df %>%
  mutate (
    across(c('Gravidez','Resultado'), 0),
    across('Pressão Sanguínea', mean())
  )
```

(C)

```
df <- df %>%
  mutate_all (
    across(c('Gravidez','Resultado'), ~replace_na(.x, 0)),
    across('Pressão Sanguínea', ~replace_na(.x, mean(.x, na.rm=TRUE)))
  )
```

```
(D) df <- df %>%
  mutate_at (
    across(c('Gravidez','Resultado'), ~replace_na(.x, 0)),
    across('Pressão Sanguínea', ~replace_na(.x, mean(.x, na.rm=TRUE)))
  )
```

42 ETL (*Extract, Transform and Load*) consiste num conjunto de técnicas com o objetivo de integração a fim de combinar dados de diversas fontes. Por meio do ETL, é possível definir a qualidade dos dados e a forma como eles são manipulados, a fim de transformá-los em uma informação inteligível e confiável. Nesse contexto, associe a técnica de ETL com a caracterização mais adequada.

- | | |
|-----------------------------------|---|
| 1 - SQL (Standard Query Language) | A - Fornece instruções detalhadas para uma aplicação sobre como obter os dados necessários para processar. Ele também descreve qual campo de origem é correspondente ao campo de destino. Por exemplo, associar um atributo de uma tabela do banco de dados com outro atributo de outra tabela de outro banco de dados. |
| 2 – Web Services | B - É o método mais comum de acessar e transformar os dados de um banco de dados relacional. |
| 3 - Processamento em lote | C - Refere-se a uma operação que envolve a movimentação de grandes volumes de dados entre dois sistemas durante o que é chamado de “janela”. Nesse período determinado, nenhuma ação pode ocorrer com o sistema-fonte, enquanto os dados são sincronizados. |
| 4 – Mapeamento de dados | D – É um método baseado na internet para fornecer dados ou funcionalidades a várias aplicações em tempo quase real. Esse método simplifica os processos de integração de dados e pode entregar, rapidamente, mais valor a partir dos dados. |

- (A) 1A, 2C, 3D e 4B.
 (B) 1B, 2D, 3C e 4A.
 (C) 1C, 2A, 3D e 4B.
 (D) 1D, 2B, 3A e 4C.

43 Quando se inicia os trabalhos de aprendizado em bases de dados é comum identificar problemas na base de dados. Por isso, na etapa de pré-processamento, os dados passam por recursos de limpeza, integração de bases de dados, redução na quantidade de atributos ou dados, transformação nos formatos dos dados ou discretização de alguns. Considerando esses recursos, indique a alternativa **incorreta** que os define.

- (A) Integração: união de múltiplas fontes em um único local, como ocorre em um *data warehouse*.
 (B) Limpeza: é utilizada apenas para eliminar registros que apresentam problemas.
 (C) Redução: reduzir a dimensão da base de dados, como por exemplo, utilizando algoritmos de *clusterização* ou eliminando atributos redundantes.

- (D) Transformação: padronização ou transformação dos dados em um formato passível de serem utilizados por diferentes técnicas de aprendizado de máquina.

44 from scipy import stats

```
...
Y, Z = stats.normaltest(X)
```

Considerando o trecho do algoritmo acima, assinale a alternativa que corresponde ao significado do valor de Z.

- (A) hipótese
 (B) probabilidade
 (C) acurácia
 (D) p-valor

45 A linguagem SQL (*Structured Query Language*, ou Linguagem de Consulta Estruturada) é a linguagem padrão para se trabalhar com bancos de dados relacionais. Por meio de sua utilização, podemos gerar diversos tipos de relatórios, como por exemplo, o relatório exibido pela Tabela 1.

Tabela 1 – Relatório dos dados da tabela COMPRAS

ID	Cliente	Produto	Quantidade
1	Paulo	Macbook	1
2	Marcelo	Pen-drive	5
3	Ana	<NULL>	<NULL>
4	Tiago	Notebook	1
5	Carlos	<NULL>	<NULL>
6	Renata	<NULL>	<NULL>

Tabela 2 – Resultado desejado pelo usuário

ID	Cliente	Produto	Quantidade
1	Paulo	Macbook	1
2	Marcelo	Pen-drive	5
3	Ana	Nada consta	Nada consta
4	Tiago	Notebook	1
5	Carlos	Nada consta	Nada consta
6	Renata	Nada consta	Nada consta

A Tabela 1 contém as informações de um relatório sobre as compras que foram realizadas por alguns clientes. Esses dados correspondem aos valores armazenados em uma tabela denominada COMPRAS. Analisando as informações exibidas, um usuário percebe que alguns clientes não compraram nenhum produto. Para esses casos, ele deseja que o trecho <NULL> seja substituído por “Nada consta”, conforme exibe a Tabela 2.

Assinale a alternativa correta que apresenta o código em linguagem SQL, que deve ser utilizado para realizar essa tarefa.

- (A) SELECT ID,
 Cliente,
 Produto,
 Quantidade
 FROM COMPRAS;
- (B) SELECT ID,
 Cliente,
 CASE WHEN Produto IS NULL
 THEN ‘Nada consta’
 END AS ‘Produto’,
 CASE WHEN Quantidade IS NULL
 THEN ‘Nada consta’
 END AS ‘Quantidade’
 FROM COMPRAS;
- (C) SELECT ID,
 Cliente,
 IF Produto IS NULL
 THEN ‘Nada consta’ END AS ‘Produto’,
 IF Quantidade IS NULL
 THEN ‘Nada consta’ END AS ‘Quantidade’
 FROM COMPRAS;
- (D) SELECT ID,
 Cliente,
 CASE WHEN Produto IS NULL
 THEN ‘Nada consta’
 ELSE Produto
 END AS ‘Produto’,
 CASE WHEN Quantidade IS NULL
 THEN ‘Nada consta’
 ELSE Quantidade
 END AS ‘Quantidade’
 FROM COMPRAS;

46 O Processamento de Linguagem Natural (PLN) é a subárea da Inteligência Artificial responsável por estudar a capacidade e as limitações de uma máquina de entender a linguagem dos seres humanos. Para poder realizar essa modelagem, são necessários pré-processamentos que abstraem e estruturam a língua, deixando apenas aquilo que representa uma informação relevante. Uma das etapas desse processo compreende a normalização. Uma tarefa que pode ser realizada dentro do processo de normalização é denominada de *tokenização lexical*. Considere a seguinte sentença:

A área de Ciência de Dados é muito interessante!

Assinale a alternativa correta que representa o resultado da tokenização lexical para essa sentença.

- (A) ['A', 'área', 'de', 'Ciência', 'de', 'Dados', 'é', 'muito', 'interessante', '!']
- (B) ['A', 'área', 'Ciência', 'Dados', 'interessante', '!']
- (C) ['A área de Ciência de Dados é muito interessante!']
- (D) A ÁREA DE CIÊNCIA DE DADOS É MUITO INTERESSANTE!

47

```
a = {1,2,3}
b = (1,2,3)
c = [1,2,3]
d = {"a": 1,"b": 2,"c": 3}
print(type(a))
print(type(b))
print(type(c))
print(type(d))
```

Assinale a alternativa que corresponde à saída gerada pelo algoritmo acima:

- (A) <class 'array'>
<class 'array'>
<class 'array'>
<class 'array'>
- (B) <class 'array'>
<class 'array'>
<class 'array'>
<class 'dict'>
- (C) <class 'set'>
<class 'tuple'>
<class 'list'>
<class 'dict'>

- (D) <class 'set'>
<class 'tuple'>
<class 'array'>
<class 'dict'>

48 Uma rede neural é um modelo preditivo motivado pela forma como o cérebro funciona. Redes neurais artificiais são formadas por neurônios artificiais, que desenvolvem cálculos similares sobre suas entradas. Elas podem resolver uma variedade de problemas, tais como o reconhecimento de caligrafia e a detecção facial, entre outros. São geralmente representadas por meio de um grafo orientado, onde os vértices representam os neurônios e as arestas representam as sinapses. Podem ser classificadas em três categorias específicas: *Redes Neurais Feed-Forward*, *Redes Recorrentes* e *Redes Conectadas Simetricamente*. Dentro dessas categorias, existem diversos tipos de arquiteturas.

Assinale a alternativa que define corretamente uma Rede Neural Perceptron Multicamadas.

- (A) Uma Rede Neural Perceptron Multicamadas é uma rede neural profunda, que pode ser utilizada para classificar imagens, agrupá-las por similaridade e realizar o reconhecimento de objetos dentro de cenas.
- (B) Uma Rede Neural Perceptron Multicamadas é uma rede neural artificial composta por mais de um *Perceptron*. Elas são compostas por uma camada de entrada, para receber o sinal, e uma camada de saída, que toma uma decisão ou previsão sobre a entrada. Entre essas camadas podem existir um número arbitrário de camadas ocultas.
- (C) Uma Rede Neural Perceptron Multicamadas é um tipo de rede neural artificial projetada para reconhecer padrões em sequências de dados, como texto, genomas, sons, dados de série temporais ou linguagem natural.
- (D) Uma Rede Neural Perceptron Multicamadas é um tipo de rede neural recorrente estocástica, na qual os nós tomam decisões binárias com alguma tendência. Esse tipo de rede neural é capaz de representar e resolver problemas combinatórios difíceis.

49 A biblioteca NLTK (Natural Language Toolkit) engloba ferramentas para processamento de linguagem natural, tais como funções de tokenização e radicalização. Dessa forma, considerando o código apresentado:

```
frase = "Não esqueçam a lista de materiais: 1 lápis  
e 2 canetas!"  
from nltk.tokenize import RegexpTokenizer  
tokenizador = RegexpTokenizer(r'\w+')  
tokens = tokenizador.tokenize(frase)  
print(tokens)
```

Qual o resultado correto?

- (A) ['Não', 'esqueçam', 'a', 'lista', 'de', 'materiais', '1', 'lápis', 'e', '2', 'canetas']
- (B) ['Não', 'esqueçam', 'a', 'lista', 'de', 'materiais', 'lápis', 'e', 'canetas']
- (C) ['Não', 'esqueçam', 'a', 'lista', 'de', 'materiais', ':', '1', 'lápis', 'e', '2', 'canetas', '!']
- (D) ['Não', 'esqueçam', 'a', 'lista', 'de', 'materiais', ':', 'lápis', 'e', 'canetas', '!']

50 Considere o seguinte código em linguagem Python e selecione a alternativa que apresenta a saída do programa.

```
w = []  
for i in range(0, x, 1) :  
    z = []  
    for j in range(0, y, 1) :  
        z.append(i*y+j)  
    w.append(z)  
print(w)
```

- (A) [[0, 1, 2, 3], [4, 5, 6, 7], [8, 9, 10, 11]]
- (B) [[0, 1, 2, 4], [9, 10, 11, 12]]
- (C) [[0, 1, 2], [3, 4, 5], [6, 7, 8]]
- (D) [[0, 1, 2, 3], [4, 5, 6, 7]]