

- Cada um dos itens das provas objetivas está vinculado ao comando que imediatamente o antecede. De acordo com o comando a que cada um deles esteja vinculado, marque, na **Folha de Respostas**, para cada item: o campo designado com o código **C**, caso julgue o item **CERTO**; ou o campo designado com o código **E**, caso julgue o item **ERRADO**. A ausência de marcação ou a marcação de ambos os campos não serão apenadas, ou seja, não receberão pontuação negativa. Para as devidas marcações, use a **Folha de Respostas**, único documento válido para a correção das suas provas objetivas.
- Nos itens que avaliarem **conhecimentos de informática e(ou) tecnologia da informação**, a menos que seja explicitamente informado o contrário, considere que todos os programas mencionados estão em configuração-padrão e que não há restrições de proteção, de funcionamento e de uso em relação aos programas, arquivos, diretórios, recursos e equipamentos mencionados.
- Eventuais espaços livres — identificados ou não pela expressão “Espaço livre” — que constarem deste caderno de provas poderão ser utilizados para rascunho.

-- PROVAS OBJETIVAS --

-- CONHECIMENTOS BÁSICOS --

Texto CB1A1

As plataformas de conectividade são o principal responsável da digitalização da economia, por realizarem a migração dos processos físicos para o mundo digital. Seu avanço tem sido muito expressivo, com uma ampla penetração (93% das pessoas usam aplicativos de mensagens instantâneas) e um aumento significativo do poder concentrado em algumas grandes plataformas, características de um mercado em franca expansão que têm motivado discussões no âmbito regulatório ao redor do mundo. Ainda não há um consenso global em relação à regulação de plataformas de conectividade, estando sua discussão presente em diferentes países.

A União Europeia se mostra como o ente mais avançado nessa discussão, tendo apresentado duas importantes leis, a Digital Services Act (DSA) e a Digital 156 Markets Act (DMA), que poderão servir como *benchmark* para outros países, sob o ponto de vista regulatório, de experiências praticadas, além de explicitar os seus potenciais impactos no mercado e como ele se adaptará.

No Brasil, não há leis nem regulações que atuem especificamente sobre as plataformas de conectividade. O Marco Civil da Internet e a Lei Geral de Proteção de Dados Pessoais, no entanto, abordam temas transversais do ecossistema digital que são relevantes no contexto de plataformas de conectividade. Alguns projetos de lei, como o PL das *Fake News* e o PL n.º 2.768/2022, endereçam parte das questões relacionadas a esse nicho de mercado atualmente e devem concentrar as principais discussões.

ANATEL. Relatório II: Avaliação de cenários futuros da economia digital induzidos pelo desenvolvimento de plataformas de conectividade no Brasil, falhas de mercado e lacunas regulatórias. 2023, p. 156-157. Internet: <www.gov.br> (com adaptações).

Com relação aos sentidos e a aspectos linguísticos do texto CB1A1, julgue os itens que se seguem.

- 1 No segundo período do primeiro parágrafo, seria correto inserir uma vírgula imediatamente depois do trecho entre parênteses — (... **instantâneas**), e.
- 2 Seriam preservadas a correção gramatical e as relações de coesão e coerência do primeiro parágrafo caso se substituísse, no segundo período, a vírgula empregada após “plataformas” por ponto final, desde que feito o devido ajuste de minúscula para maiúscula na letra inicial da palavra “características”.
- 3 O primeiro período do texto poderia ser reescrito, sem prejuízo da correção gramatical e das relações de sentido estabelecidas entre as orações que o compõem, da seguinte forma: **Por realizarem a migração dos processos físicos para o mundo digital, as plataformas de conectividade são as principais responsáveis pela digitalização da economia.**
- 4 Estariam mantidas a correção gramatical do texto e a coerência de suas ideias caso se reescrevesse o segmento “(93% das pessoas usam aplicativos de mensagens instantâneas)” (segundo período do primeiro parágrafo) da seguinte forma: **(93% da população é usuária de aplicativos de mensagens instantâneas).**

- 5 Estariam preservadas a correção gramatical e a coerência das ideias do último período do primeiro parágrafo caso a oração “estando sua discussão presente em diferentes países” fosse reescrita da seguinte maneira: **tema cuja discussão está presente em diferentes países.**
- 6 O vocábulo “ele”, na última oração do segundo parágrafo, retoma “ponto de vista regulatório”.

Espaço livre

Texto CB1A6

De todas as inovações óbvias que surgiram no século XXI, poucas são tão invisíveis como a mudança na maneira de perceber o tempo.

O controle do tempo na comunicação até o final do século XX decorria especialmente de uma característica: a programação. Com o controle dos meios de produção e distribuição da informação, fosse ela qual fosse, o cenário temporal dispunha também de uma estrutura hierárquica de controle. Era possível traçar e agenciar cronogramas complexos e prazos de entrega. O que basicamente mudou?

A comunicação não mudou. Mudaram os meios. Todos têm acesso às redes. Nelas a produção de conteúdo é incessante. O conteúdo é produzido por todos. A estrutura não é mais regida por uma lógica de programação, mas por uma lógica de fluxo. Nesse novo ecossistema, impera uma economia da atenção e, sendo a atenção um recurso escasso, é preciso a todo tempo escolher algo a que dedicar a atenção e, dedicada a atenção, devemos reagir rapidamente com as mudanças necessárias. Aqui aparecem claramente as diferenças do tempo em um ecossistema em fluxo permanente. As mudanças também devem ser permanentes. Isso em comunicação se traduz em não esperar mais longos prazos para dar respostas. Até o século passado, o furo de um veículo de comunicação só poderia ser ultrapassado 24 horas depois: na próxima edição. Hoje um furo de reportagem é imediatamente absorvido em rede, compartilhado, respondido e repercutido por todos os concorrentes. Em uma sociedade que se comunica em fluxo, é preciso avaliar diariamente onde estão os impedimentos e o que deve ser priorizado, e redefinir foco e estratégia. A comunicação também passa a lidar com o provisório, com o possível e com o impermanente.

Saber como equilibrar o tempo nessa pressão imediatista é o grande desafio do século XXI. Há saberes que só se atingem com reflexão, laços que só se formam com experiências reais compartilhadas. Saber desligar o celular, parar de responder *emails* para completar uma tarefa, ler um livro inteiro e principalmente saber conversar com os outros são habilidades cada vez mais valorizadas em uma sociedade digital, exatamente por serem habilidades capazes de dilatar o tempo em que vivemos, formas de transformar o fluxo que consome nossa atenção em momento no qual nós consumimos o mundo.

Margot Pavan. O tempo e a comunicação digital no século XXI.
Internet: <www.jornaldocomercio.com> (com adaptações).

Julgue os itens que se seguem, relativos às ideias veiculadas no texto CB1A6.

- 7 Conclui-se da leitura do terceiro parágrafo do texto que o avanço da tecnologia no âmbito da comunicação e a democratização do acesso à informação são, respectivamente, o foco e a estratégia a serem redefinidos na sociedade do século XXI, que se comunica em fluxo permanente.
- 8 O texto classifica-se como dissertativo-argumentativo e, no parágrafo de conclusão, são apresentados exemplos de ações que representam possíveis respostas ao que a autora identifica como “o grande desafio do século XXI”.
- 9 Em sua análise da transformação ocorrida na transmissão de informação no século XXI, a autora direciona uma crítica à produção e disseminação de conteúdos superficiais e pouco confiáveis pelas mídias digitais.
- 10 Entende-se da leitura do texto que a comunicação no século XXI, marcada pelas redes sociais digitais, em que a produção de conteúdo é incessante, requer respostas imediatas e em fluxo permanente.

No que diz respeito a aspectos linguísticos do texto CB1A6, julgue os seguintes itens.

- 11 Para se evitar a repetição do termo “atenção” no sétimo período do terceiro parágrafo, seria correto e coerente reescrever o segmento “dedicada a atenção” da seguinte forma: **tendo-lhe dedicado**.
- 12 No primeiro período do texto, a substituição do vocábulo “como” por **quanto** preservaria as relações de coesão e coerência textuais, mas, para a garantia da correção gramatical, seria necessário o emprego do acento indicativo de crase no vocábulo “a” que antecede “mudança” — **quanto à mudança**.
- 13 Estariam mantidas a correção gramatical e a coerência das ideias do texto caso se substituisse o ponto final que encerra o segundo período do segundo parágrafo pelo sinal de dois-pontos, feito o devido ajuste de maiúscula e minúscula no período.
- 14 Estariam preservadas a correção gramatical e a coerência das ideias do texto caso o segmento “sendo a atenção um recurso escasso” (sétimo período do terceiro parágrafo) fosse reescrito da seguinte forma: **porquanto ela seja um recurso escasso**.
- 15 O segmento “por uma lógica”, em “mas por uma lógica de fluxo” (sexto período do terceiro parágrafo), poderia ser suprimido sem prejuízo das relações de coesão e coerência textuais.

Text CB1A2-I

Indigenous languages appear to be disappearing at a concerning rate around the world. Current data indicates that at least 43% of the world’s spoken languages are being forgotten, with indigenous languages making a significant part. According to ONIC (National Indigenous Organization of Colombia), there are 65 indigenous languages in the country. Nearly one-third of them are at a critical risk of disappearing — and only three of them, including Nasa Yuwe, are spoken by more than 50,000 people. This is the reality that Juan Pablo Camayo seeks to change.

Two years ago, in Caldon, Colombia, Juan Pablo and other neighbours started a communications network that enabled them to provide Internet access to remote rural areas and disseminate content in their mother tongue. That’s how Jxa’h Wejxia Casil — “Wind’s Net” in Nasa Yuwe — came to be.

Currently, Jxa’h Wejxia Casil has about 200 families subscribed to their Internet service and provides a tool that allows communities to preserve the use of their language. With support from other project partners, Juan Pablo leveraged this new communication tool to produce and disseminate original communication materials in Nasa Yuwe.

Internet: <<https://unsdg.un.org>> (adapted).

According to the text CB1A2-I, judge the following items.

- 16 The word “leveraged”, in the last sentence of the text, can be correctly understood as **took advantage of**.
- 17 According to the text, the disappearance of indigenous languages is a worrying fact because they represent almost 43% of the world’s spoken languages.
- 18 According to the text, Jxa’h Wejxia Casil is a rural communications network that aims to promote the spread of content in Nasa Yuwe.
- 19 Based on the text, it is correct to conclude that, Nasa Yuwe is not classified as a critically endangered language because it is spoken by more than 50,000 people.

Text CB1A2-II

Internet coverage in the European Union (EU) is impressive, standing at 100%; however, numbers on *de facto* usage (85%), broadband take-up (78%), users with at least basic digital skills (58%), next-generation access coverage providing at least 30 Mbps (86%) and 5G readiness (21%) cloud the picture.

The significance of these circumstances does not only lie in the economic implications but also in the severe consequences for the individual and the society. People without adequate Internet access are missing out on means of participation and opportunities that have become part of everyday life. Countless contributions have been published on socioeconomic inequalities relating to access to, use of or impact of ICTs (information and communication technologies), known as the digital divide(s). These divides will only deepen, as disconnected citizens are likely to miss out on long-term benefits of innovation (information society) and modernization.

The coronavirus pandemic exacerbated the social inequalities related to insufficient connectivity: privileged users experienced dropped calls and disrupted downloads, while disadvantaged users were left with no access at all or with makeshift solutions.

Internet: <www.europarl.europa.eu/thinktank> (adapted).

Based on the text CB1A2-II, judge the following items.

- 20 The adjective “makeshift”, in the last sentence of the text, is a synonym for **obsolete**.
- 21 Digital divide can be understood as the lack of long-term stable connection to the Internet.
- 22 The percentage of actual Internet users in the EU is clouded by the numbers on 5G preparedness.
- 23 Lack of satisfactory Internet access results in exclusion of people from benefits related to digital technology.

Em relação às agências reguladoras federais e às abordagens econômicas na regulação, julgue os itens seguintes.

- 24 A teoria econômica da regulação estabelece que, havendo monopólio, deve-se definir uma regulação, por exemplo, em preços, com o objetivo de evitar a obtenção de lucros extraordinários pela firma monopolista.
- 25 As agências reguladoras federais caracterizam-se pela independência em sua gestão e administração e na tomada de decisões relativas aos contratos de concessão, embora se subordinem às pastas ministeriais nos assuntos regulatórios.
- 26 A criação das agências reguladoras federais constituiu um mecanismo delimitador das divisas dos serviços públicos visando a fixação de normas para a definição de tarifas, a fiscalização do cumprimento dos atos regulatórios e o estímulo à competência dos prestadores de serviços.

Julgue os itens subsequentes, a respeito das boas práticas regulatórias e da regulação responsiva.

- 27 O princípio da proporcionalidade aplicado à análise do resultado regulatório diz respeito à correta calibragem dos recursos, dos esforços e do tempo empregados na análise conforme a relevância e complexidade da regulação a ser revista.
- 28 Segundo os pressupostos da regulação responsiva, a abordagem regulatória não pode ser apenas realizada com caráter punitivo, devendo ter, também, cunho persuasivo, no sentido de reorientar a conduta dos regulados.
- 29 A análise de impacto regulatório é um instrumento para avaliação de uma intervenção regulatória adotada ou alterada, com relação ao atingimento dos objetivos originalmente pretendidos.

Julgue os itens a seguir, com base na Lei n.º 9.472/1997, que dispõe sobre a regulação do setor de telecomunicações no Brasil.

- 30 O serviço de valor adicionado constitui modalidade do serviço de telecomunicações, classificando-se seu provedor como usuário do serviço de telecomunicações que lhe dá suporte, com os direitos e deveres inerentes a essa condição.
- 31 O usuário de serviços de telecomunicações tem direito à liberdade de escolha da prestadora do serviço e à inviolabilidade e ao segredo de sua comunicação, salvo nas hipóteses e condições constitucionais e legalmente previstas.
- 32 O Conselho Consultivo da ANATEL é o órgão de participação institucionalizada da sociedade na ANATEL e possui representantes indicados pelas entidades de classe das prestadoras de serviços de telecomunicações.
- 33 O Conselho Administrativo da ANATEL é composto pelo presidente e por quatro conselheiros e suas decisões são tomadas por maioria absoluta.

A respeito de organização administrativa, administração direta e indireta, espécies de atos administrativos e poderes administrativos, julgue os itens a seguir, considerando, no que couber, o entendimento do Supremo Tribunal Federal (STF).

- 34 A competência normativa da agência reguladora qualifica-se pela edição de atos gerais e abstratos, de caráter técnico, necessários à implementação da política pública setorial, devendo estar subordinados aos preceitos constitucionais e às leis de regência.
- 35 A agência reguladora, entidade da administração indireta, insere-se tanto no conceito formal quanto no conceito material de administração pública.
- 36 Em relação aos atos administrativos normativos, as agências reguladoras podem editar decreto regulamentar, regimento interno e resolução para disciplinar a organização e a prestação dos seus serviços regulatórios.
- 37 No desempenho de suas funções, as agências reguladoras independentes detêm regime especial caracterizado por independência administrativa, autonomia financeira, ausência de subordinação hierárquica, mandato fixo e estabilidade de seus dirigentes.

Acerca da concessão de serviços públicos, do processo administrativo no âmbito da Lei n.º 9.784/1999, do controle judicial da administração pública e do que estabelece a Lei de Acesso à Informação (LAI), julgue os itens subsequentes, considerando, no que couber, o entendimento do STF.

- 38 No processo administrativo federal, o ônus da prova é do interessado, que não pode se valer de provas impertinentes e desnecessárias, sob pena de serem recusadas.
- 39 Não cabe ao Poder Judiciário realizar o controle jurisdicional dos atos de uma agência reguladora que resultem da interpretação e da aplicação do seu próprio estatuto legal, quando a legislação de regência define as políticas a serem perseguidas, os objetivos a serem implementados e os objetos de tutela.
- 40 A reversão dos bens do concessionário ao poder concedente ao final do prazo da concessão de serviço público tem por fundamento a indenização da administração pública pelos custos decorrentes de todo o período da concessão.
- 41 Em se tratando de concessão de serviço público, o reajuste do preço da tarifa deve seguir os parâmetros estabelecidos no contrato de concessão, porém, quando o reajuste tarifário for superior ao índice de inflação apurado no período, deverá prevalecer o referido índice inflacionário, em respeito ao princípio da modicidade da tarifa.
- 42 A LAI aplica-se às empresas privadas contratadas pela administração pública, pois são receptoras de recursos públicos; entretanto, em relação a elas, a publicidade ficará limitada à parcela dos recursos recebidos.

No que se refere às previsões da Constituição Federal de 1988 (CF) concernentes aos direitos e garantias individuais e sociais, aos direitos humanos e à nacionalidade, julgue os itens a seguir.

- 43 Ressalvadas as situações que acarretem apatridia, será declarada a perda da nacionalidade do brasileiro que fizer pedido expresso de perda da nacionalidade brasileira perante autoridade brasileira competente, hipótese em que o interessado ficará impedido de readquirir sua nacionalidade brasileira originária.
- 44 O sigilo da correspondência e das comunicações telegráficas, de dados e das comunicações telefônicas é inviolável, salvo se houver ordem judicial, nas hipóteses e na forma que a lei estabelecer, para fins de investigação criminal ou instrução processual penal.
- 45 Os tratados e convenções internacionais sobre direitos humanos que forem aprovados, em sessão unicameral do Congresso Nacional, por dois terços dos votos de seus membros serão equivalentes às emendas constitucionais.

Com base nas disposições da CF acerca da forma e do sistema de governo, bem como da defesa do Estado e das instituições democráticas, julgue os itens a seguir.

- 46 O decreto que instituir o estado de defesa determinará o tempo de sua duração, especificará as áreas a serem abrangidas e indicará, nos termos e limites da lei, as medidas coercitivas a vigorarem, entre as quais estão restrições ao sigilo de correspondência e de comunicação telegráfica e telefônica.
- 47 A forma republicana de governo é uma cláusula pétrea e o sistema presidencialista de governo, um princípio constitucional sensível.

À luz do disposto na CF em relação à ordem econômica e financeira e à ordem social, julgue os itens a seguir.

- 48 A ordem social, fundada na valorização do trabalho humano e na livre iniciativa, tem como objetivo assegurar a todos existência digna, conforme os ditames da justiça social, sendo assegurado a todos o livre exercício de qualquer atividade social que seja previamente autorizada pelos órgãos públicos.
- 49 A sociedade exercerá a função de planejamento das políticas sociais, assegurada, na forma da lei, a participação do Estado nos processos de formulação, de monitoramento, de controle e de avaliação dessas políticas.
- 50 Incumbe ao poder público, na forma da lei, diretamente ou sob regime de concessão ou permissão, sempre mediante licitação, a prestação de serviços públicos.

Espaço livre

-- CONHECIMENTOS ESPECÍFICOS --

Considerando uma distribuição condicional expressa na forma de função de densidade de probabilidade $f(x|y) = ye^{-xy}$, em que e denota a constante de Euler, e x e y , valores reais positivos que representam, respectivamente, os pontos de suporte das variáveis aleatórias contínuas X e Y , julgue os itens a seguir.

- 51 Se Y seguir uma distribuição exponencial com média igual a 1, então, para $x > 0$, a função de densidade da variável aleatória X será $f(x) = (x + 1)^{-2}$.
- 52 $P(X > 1 | Y = 2) = e^{-2}$.
- 53 A média da variável aleatória X , condicionada ao evento $Y = 5$, é igual a 5.
- 54 $P(X = 1) = P(Y = 10)$.
- 55 As variáveis aleatórias X e Y são independentes.

Os valores 2, 3, 1, 0, 2 constituem uma amostra aleatória simples de tamanho 5 retirada de uma distribuição discreta W , na qual $P(W = w) = p(1 - p)^w$, com $w \in \{0, 1, 2, \dots\}$, sendo p um parâmetro que denota uma probabilidade.

Com base nas informações precedentes e no método de estimação por máxima verossimilhança, julgue os próximos itens.

- 56 O estimador de máxima verossimilhança para a variância de W é a variância amostral.
- 57 Não há estimador de máxima verossimilhança para a moda de W , já que o valor da moda não depende da probabilidade p .
- 58 Se $\hat{P}(W = 2)$ denota a estimativa de máxima verossimilhança da probabilidade $P(W = 2)$, então $\hat{P}(W = 2) = 0,4$.
- 59 A estimativa de máxima verossimilhança da probabilidade p é igual a 0,625.
- 60 Pelo método da máxima verossimilhança, a estimativa da média de W é igual a $\frac{8}{5}$.

Uma amostra aleatória simples de tamanho n será retirada, com reposição, de certa população para a estimação de um parâmetro populacional λ . O estimador, representado por T_n , possui as propriedades $E[T_n] = \frac{(n+2)\lambda}{n}$ e $\text{Var}[T_n] = \frac{\lambda^2}{n}$.

No que diz respeito ao estimador hipotético T_n do parâmetro λ , julgue os seguintes itens.

- 61 Se T_n seguir uma distribuição normal, então a razão $\frac{nT_n - (n+2)\lambda}{\sqrt{n\lambda}}$ será normal padrão.
- 62 T_n é estimador de λ assintoticamente não viciado.
- 63 O erro-padrão de T_n é igual a 1.
- 64 T_n é estimador consistente.
- 65 Se $n = 10$, então o erro quadrático médio de T_n será igual a $\frac{\lambda^2}{10}$.

Acerca das abordagens relacional e não relacional, entre outros conceitos relacionados a banco de dados, julgue os itens a seguir.

- 66 O exemplo a seguir exhibe como os dados podem ser armazenados em um banco de dados relacional do tipo chave-valor.

```
{
  nome_orgao: "ANATEL",
  endereco: {estado: "Distrito Federal",
  cidade: "Brasília"},
  numero_telefone: "55-61-1234-5678",
  atuacao: ["telecomunicação"],
}
```

- 67 Nos bancos de dados NoSQL orientados a colunas, é possível organizar os dados em famílias de colunas, o que facilita o agrupamento e o acesso a informações específicas, como, por exemplo, os dados de contato.
- 68 Em SQL, *triggers* são procedimentos automáticos executados em resposta a determinados eventos em uma tabela ou *view* e são disparados exclusivamente em resposta a ações DML, tais como INSERT, UPDATE ou DELETE.
- 69 Em um banco de dados, os índices são armazenados fisicamente na mesma ordem em que são definidos logicamente para permitir acesso eficiente aos dados.
- 70 Se executada, a consulta a seguir retorna a contagem de clientes com mais de 18 anos de idade que realizaram transações com valores superiores a 1.000, agrupados e ordenados de forma decrescente por idade.

```
SELECT COUNT(c.id), c.idade
FROM cliente c
WHERE c.idade > 18
AND EXISTS (
  SELECT 1
  FROM transacao t
  WHERE t.cliente_id = c.id
  AND t.valor > 1000
)
GROUP BY c.idade
ORDER BY c.idade DESC;
```

Julgue os itens que se seguem, relativos a conceitos de *data warehouse*, técnicas de modelagem dimensional e otimização de bases de dados para BI.

- 71 Nos *data warehouses*, os índices de junção associam as linhas da tabela de fatos com as respectivas colunas na tabela de dimensões, o que facilita as operações de consulta e análise de dados.
- 72 Uma tabela de fatos descreve as entidades de um negócio e apresenta uma ou mais colunas-chave que funcionam como um identificador único exclusivo, bem como colunas descritivas adicionais.
- 73 Em regra, as chaves substitutas são incorporadas nas tabelas de dimensões de um *data warehouse* relacional para atribuir um identificador único para cada registro nas tabelas de dimensões.
- 74 Nos *data warehouses*, as colunas de chave de dimensão definem a dimensão de uma tabela de fatos, enquanto os valores das chaves de dimensão estabelecem sua granularidade.
- 75 A relação entre uma tabela de fatos e suas respectivas tabelas de dimensões em um *data warehouse* é estabelecida pela chave primária da tabela de fatos e a chave estrangeira da tabela de dimensões.

- 76** A consolidação de dimensões de floco de neve em um modelo de tabela única pode implicar o armazenamento de dados não normalizados e redundantes, o que resulta em maior tamanho de armazenamento.
- 77** O processo de otimização de desempenho, em bases de dados, consiste em redimensionar o tamanho do modelo semântico, o que resulta em uma maior rapidez na atualização dos dados, cálculos e renderização dos elementos visuais em relatórios.

Em relação ao tratamento e à qualidade dos dados no sistema de gerenciamento de informações, julgue os itens subsequentes.

- 78** De acordo com a Lei Geral de Proteção de Dados Pessoais (LGPD), os dados pessoais devem ser retidos por um período de cinco anos, mesmo após a conclusão do seu processamento, desde que sejam cumpridos os limites técnicos das atividades em questão.
- 79** O agente de tratamento deve demonstrar a adoção de medidas eficazes e capazes de comprovar a observância e o cumprimento das normas de proteção de dados pessoais, bem como demonstrar a eficácia dessas medidas.

A respeito de visualização, análise exploratória de dados e geoprocessamento, julgue os seguintes itens.

- 80** A definição de seções implícitas pode ser realizada pela utilização de formas coloridas e pela sobreposição de elementos visuais alinhados, o que confere uma clara distinção entre diferentes áreas do *layout* dos *dashboards*.
- 81** Os mapas temáticos incluem áreas geográficas delimitadas por um ou mais polígonos, a exemplo dos mapas de uso do solo e mapas que indicam a capacidade agrícola de determinada região.
- 82** A aplicação de hierarquias de atributo em modelos analíticos é uma prática eficiente, que permite a organização de valores pré-agregados em cada nível e otimiza a análise de dados.
- 83** No contexto de um *layout* de relatório, o equilíbrio assimétrico consiste na distribuição do peso uniforme dos objetos em ambas as metades da página, independentemente do tamanho dos objetos.

No que se refere à governança de dados, julgue os próximos itens.

- 84** De acordo com a LGPD, dados pessoais relacionados a convicção religiosa são considerados sensíveis e só podem ser tratados com o consentimento específico e destacado do titular ou de seu representante legal.
- 85** Um dado de referência, como, por exemplo, os dados dos clientes, é uma informação crucial para a operação do negócio.

Julgue os próximos itens, relativos a Naive Bayes e *random forest*.

- 86** Tamanho do nó, número de árvores e número de recursos amostrados, ou número de preditores amostrados, são parâmetros de algoritmos *random forest*.
- 87** Naive Bayes é um algoritmo de classificação baseado na aprendizagem por reforço, em que um agente realiza uma ação e recebe uma recompensa de acordo com o resultado dessa ação por meio da implementação do teorema de Bayes, com o objetivo de encontrar a probabilidade a *posteriori*.
- 88** *Random forest* é um algoritmo de classificação que permite a realização de mineração dos dados por meio da criação de estruturas de aprendizagem a partir de uma base de dados na qual se utiliza uma única árvore de decisão para a classificação dos dados.
- 89** O algoritmo de classificação Naive Bayes pode ser utilizado para o cálculo da probabilidade de ocorrência de um evento, com base em probabilidades obtidas em eventos numéricos passados, e, por isso, não pode ser empregado em atividades de classificação textual.
- 90** Nas árvores de decisão e em *random forest*, são utilizadas técnicas estatísticas com o objetivo de se produzir, a partir de um conjunto de observações, uma predição de valores em função de uma ou mais variáveis independentes contínuas e(ou) binárias.

A respeito de KNN (*k-nearest neighbours*), SVM (*support vector machines*), *deep learning* e técnicas de agrupamento, julgue os itens a seguir.

- 91** *Deep learning* é um tipo de aprendizado de máquina que usa redes neurais artificiais para permitir que sistemas digitais aprendam e tomem decisões com base em dados não estruturados e não rotulados.
- 92** KNN é um algoritmo de aprendizado supervisionado não paramétrico que não pode ser utilizado em problemas de classificação, uma vez que seu objetivo é prever valores numéricos e não valores categóricos.
- 93** SVM é um algoritmo de aprendizado de máquina supervisionado que pode ser usado para desafios de classificação ou regressão.
- 94** Uma das formas de se realizar um agrupamento é por meio de técnicas de agrupamento baseadas em hierarquia, em que se pode criar estrutura hierárquica de acordo com a proximidade entre os indivíduos, o que resulta em uma árvore binária.
- 95** O SVM classifica os dados encontrando uma linha ou hiperplano ideal; essa linha de separação é encontrada entre duas classes distintas pela análise dos dois pontos, um de cada grupo, mais próximos da outra classe.

Acerca da avaliação de modelos de classificação, julgue os itens que se seguem.

- 96** A acurácia é uma métrica adequada para a avaliação de modelos quando não há desbalanceamento de classes, pois reflete com precisão a capacidade geral do modelo de fazer previsões corretas em todas as classes.
- 97** A matriz de confusão, em problemas de classificação multiclases, é uma tabela com duas linhas e duas colunas; na diagonal principal dessa matriz quadrada, estão os valores corretos e, na matriz secundária, os erros cometidos pelo modelo.
- 98** Um modelo de classificação que apresenta alta revocação é útil em contextos em que seja crucial identificar a maior quantidade possível de casos positivos, mesmo que isso resulte em um número maior de falsos positivos.
- 99** A área sob a curva ROC (*receiver operating characteristic*) é uma métrica de qualidade útil para avaliar um modelo: quanto mais próxima a curva estiver do canto superior direito do gráfico, melhor será a predição do modelo.

A respeito de técnicas de redução de dimensionalidade, julgue os itens subsecutivos.

- 100** Quando da configuração dos parâmetros do *autoencoder*, o tamanho do espaço latente é uma informação crucial, pois determina o tamanho do espaço onde os dados de entrada serão comprimidos.
- 101** Para utilizar de forma adequada a análise de componentes principais (PCA, na sigla em inglês), é essencial normalizar os dados; se as variáveis não estiverem na mesma escala, aquelas com maior variância terão maior impacto, distorcendo o resultado da PCA.

Julgue os próximos itens, referentes ao processamento de linguagem natural.

- 102** A saída do Word2Vec consiste em vetores densos de baixa dimensão que representam palavras em um espaço contínuo, onde cada palavra é mapeada para um vetor numérico no qual cada dimensão captura uma característica da palavra.
- 103** A similaridade de cosseno é uma métrica pela qual se avalia a similaridade entre dois vetores com base no ângulo entre eles em um espaço vetorial, de forma que, à medida que os vetores se aproximarem, aumentará a similaridade de cosseno.
- 104** A lematização prescinde do POS *tagging* para que as palavras sejam reduzidas corretamente, pois todas as palavras são reduzidas ao mesmo *lemma*, independentemente de sua classe gramatical.
- 105** Na redução de palavras ao radical, ocorre *under-stemming* quando duas palavras separadas são reduzidas erroneamente à mesma raiz e, com isso, ocorre a perda de distinção semântica entre palavras com significados diferentes.

A respeito de *Big Data*, julgue os próximos itens.

- 106** *Pipelines* de dados apresentam uma única estrutura para o recebimento dos dados originados de uma fonte não confiável.
- 107** Em *Big Data*, ruídos consistem em informações extras que acabam deturpando as análises, enquanto *overfitting* designa a interpretação equivocada dos ruídos como dados legítimos.
- 108** A coleta de dados por meio de aplicativos é considerada explícita, porque o usuário a autoriza.
- 109** O YARN (*Yet Another Resource Negotiator*) é um sistema de arquivos distribuídos que faz parte do *framework* Hadoop.
- 110** No modelo SaaS (*software as a service*) da computação em nuvem utilizado para *Big Data*, a aplicação e os dados são gerenciados pelo provedor da nuvem.
- 111** Projetos de *Big Data*, quando necessário, crescem horizontalmente, com a inclusão de novos nodos, e verticalmente, com o acréscimo de mais memória.
- 112** Em um *data lake*, os dados são depositados em estado bruto, sem terem passado por qualquer análise e mesmo sem terem uma governança.
- 113** Subconjunto de um *data warehouse*, o *data mart* é especializado em uma área específica de uma organização.
- 114** O processo de ELT, devido às suas etapas, exige maior definição de regras, estruturas e relações do que a abordagem ETL.
- 115** No processamento ROLAP, bancos de dados relacionais são utilizados como local de armazenamento para agregação, enquanto, nos processamentos MOLAP e HOLAP, utilizam-se bancos de dados multidimensionais.

Acerca do fluxo de *Big Data*, julgue os itens que se seguem.

- 116** Na apresentação de dados, a extração de subcoleções e a consulta de parâmetros permitem a navegação em diversos cenários da visualização.
- 117** *Streaming processing* é uma tecnologia de *Big Data* exclusiva para atender processamentos de serviços de *streaming* de áudio e vídeo.
- 118** O serviço Elasticsearch utiliza índices divididos em fragmentos, de maneira que cada nó armazena diversos fragmentos e atua na coordenação das operações nos vários fragmentos.
- 119** As funções do MapReduce transformam um volume grande de dados em grupamentos segmentados, mantendo na saída a mesma quantidade de dados da entrada.
- 120** Na etapa de captura de *Big Data*, grandes volumes de dados são armazenados em bancos de dados NoSQL, devido à sua escalabilidade e à sua flexibilidade.