

COLETÂNEA CONCURSOS PÚBLICOS



GLOSSÁRIO DO EDITAL DO CONCURSO PARA SEFAZ PR

TECNOLOGIA DA INFORMAÇÃO PARA O CARGO AUDITOR FISCAL



**COM QUESTÕES INÉDITAS COMENTADAS
PÓS-EDITAL 2025**

APRESENTAÇÃO

A ideia central deste ebook preparatório para o concurso da Secretaria da Fazenda do Paraná (SEFAZ PR) é que o estudante possa aprender de forma rápida todo o conteúdo relacionado a tecnologia da informação (TI) do edital que estão nos tópicos ANÁLISE DE DADOS e PRÁTICA EM ANÁLISE DE DADOS, portanto, para um maior aprofundamento em cada assunto é necessário utilizar materiais complementares, porém em concurso de cargos gerais, que não exigem formação específica, as bancas não costumam aprofundar tanto nesses assuntos e, também, não adianta aprofundar tanto em TI e deixar as outras disciplinas de lado.

Este ebook foi elaborado com auxílio de inteligência artificial e revisado e complementado pelo autor Izaias Batista dos Santos.

Siga a Kuasarnex nas redes sociais para receber notificações sobre conteúdos, cupons e oportunidades.



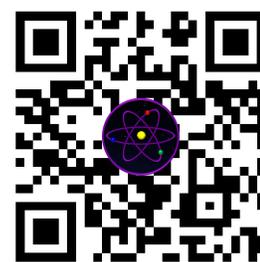
[Seguir no Instagram](#)



[Inscrição no Canal](#)



[Seguir no LinkedIn](#)



[Site Oficial](#)

AUTOR: IZAIAS BATISTA DOS SANTOS

Olá, seja muito bem-vindo(a)!

Sou Izaias Batista dos Santos, autor deste material preparatório especialmente desenvolvido para auxiliar você em sua jornada rumo à aprovação no concurso da SEFAZ PR. Tenho orgulho de compartilhar um pouco da minha trajetória com você, pois acredito que conhecer quem está por trás dos conteúdos reforça a confiança no estudo e nos resultados.

Sou mestre em **Tecnologias Computacionais para o Agronegócio** pela **Universidade Tecnológica Federal do Paraná (UTFPR)**, especialista em **Engenharia de Software** pela **Pontifícia Universidade Católica de Minas Gerais (PUC-Minas)**, possuo **MBA em Gerenciamento de Projetos de TI** pelo **Instituto de Gestão em Tecnologia da Informação (IGTI)** e sou bacharel em **Sistemas de Informação** pelo **Centro Universitário Dinâmica das Cataratas (UDC)**.

Minha trajetória profissional foi construída com dedicação e propósito em diversas áreas da Tecnologia da Informação. Atuei como **coordenador de TIC no Exército Brasileiro**, **analista de sistemas sênior** na **Fundação Parque Tecnológico Itaipu**, **coordenador de projetos de software** na empresa **Eits Prognus Software Livre** e também fui **professor substituto de informática e suas tecnologias** e **técnico em tecnologia da informação e comunicações** no **Instituto Federal do Paraná**.

Atualmente, sou **analista de sistemas e processos** no **Conselho Federal de Química**, onde atuo com **gestão por processos** e exerço a função de **gestor técnico de sistemas**, sempre buscando alinhar boas práticas, inovação e eficiência para garantir resultados sustentáveis e soluções de alto impacto.

Com base em minha experiência acadêmica e profissional, preparei este e-book para oferecer a você um conteúdo claro, atualizado e voltado aos temas mais cobrados nos concursos públicos da área de Tecnologia da Informação. Meu objetivo é ajudá-lo(a) a dominar os assuntos e encurtar o caminho até a aprovação.

E, nos momentos em que bater a dúvida sobre sua capacidade de aprender e vencer esse desafio, lembre-se sempre da poderosa frase:

“Tudo posso naquele que me fortalece!”

Ah, e caso queira acompanhar mais conteúdos, dicas e novidades sobre concursos públicos e Tecnologia da Informação, será um prazer ter você comigo nas redes sociais!

Siga-me e vamos juntos nessa jornada:



[Seguir no Instagram](#)



[Inscrição no Canal](#)



[Seguir no LinkedIn](#)

Estarei sempre compartilhando conhecimento, informações atualizadas e motivação para que você se mantenha firme no seu propósito. Será uma honra contar com sua presença por lá!

Bons estudos e sucesso em sua caminhada!

OUTROS MATERIAIS

Conheça outros materiais acessando o seguinte



[Acessar Conteúdos](#)

Conheça também os materiais gratuitos acessando o seguinte



[Material Gratuito](#)

SIMBOLOGIAS UTILIZADAS

As simbologias estão estrategicamente posicionadas em nossos materiais a fim de destacar alguns assuntos

MARCADOR	EXPLICAÇÃO
 Foco para discursivas!	Conteúdo com maior probabilidade de ser cobrado em provas discursivas por ter alta incidência em provas anteriores
 Memorize!	Conteúdo que é explorado em quase todas as provas e que pode cair em discursivas
 Resumo!	Explicação sintetizada
 Destaque!	Destacar um ponto
 Despenca nas provas!	Sempre cai nas provas e em mais de uma questão
 Cai muito!	É cobrado em muitas provas
 Exercício essencial!	Treinamento para fixação do conteúdo apresentado
 Atenção!	Hora de ficar atento se estiver lendo no automático
 Importante!	Conteúdo relevante
 Ponto chave!	Ponto mais relevante do assunto
 Sugestão de leitura!	Recomendação de conteúdo complementar

CARGO: AUDITOR FISCAL

SUMÁRIO

CAPÍTULO I - ANÁLISE DE DADOS - CONCEITOS E APLICAÇÕES

1. Estatística descritiva. Média, média aparada, moda, mediana, quartis e percentis, variância, desvio padrão, coeficiente de variação
2. Probabilidade, distribuições de probabilidade
3. Método de Monte Carlo: conceito, aplicações
4. População, amostra, amostragem
5. Inferência: estimação pontual e intervalar, intervalo de confiança
6. Testes de hipóteses
7. Séries temporais: decomposição, média móvel. Outliers
8. Normalização: mudança de escala, z-score, min-max
9. Regressão linear simples: conceito, suposições, linearização, coeficientes, R-quadrado, resíduos, valor-p, elasticidades, interpretação. Regressão linear múltipla: aplicações, riscos e limitações, interpretação
10. Dados: conceitos, atributos, métricas, transformação de dados
11. Análise de dados. Agrupamentos. Tendências. Projeções
12. Análise exploratória: aplicação de conceitos e técnicas; visualização
13. Conceitos e técnicas de transformação e remodelagem de dados: pivotamento, melting e outros
14. Gráficos: linha, base 100, barra, barra empilhada e normalizada, rosca, dispersão, histograma, radar, box plot, bolhas, mapa de calor, cascata e outros
15. Noções de data warehouse
16. Noções de analytics
17. Noções de business intelligence

CAPÍTULO II - PRÁTICA EM ANÁLISE DE DADOS

1. Análise de dados com Microsoft Excel. Análise e manipulação de dados com as funções SE, E, OU, SOMASES, CONT.SES, PROCV, PROCX, ÚNICO, SERRRO, funções de tratamento de texto e data, e outras funções básicas e avançadas e suas combinações
2. Identificação e tratamento de valores faltantes ou duplicados
3. Aplicações da ferramenta na manipulação e transformação de dados, análise exploratória, limpeza de dados, relacionamento entre conjuntos de dados, resolução de problemas de estatística, matemática e lógica
4. Manipulação, preparação e análise de dados com tabelas e gráficos dinâmicos
5. Análise de dados estruturados utilizando SQL. Estruturas de dados, tipos de dados, relacionamento entre tabelas, joins, condicionais, agrupamento, sumarização, filtragem, manipulação de valores, textos e datas, subconsultas, identificação e tratamento de valores faltantes ou duplicados, produto cartesiano, erros comuns em consultas
6. Lógica de programação: aplicações com pseudocódigo utilizando expressões lógicas, estruturas condicionais, laços de repetição, manipulação de variáveis (escalares, vetores, matrizes e outros); algoritmos; resolução de problemas de lógica

7. Noções de expressões regulares utilizando software Notepad++: caracteres e quantificadores básicos, grupos de captura, asserções simples, busca e substituição de texto

8. Desenvolvimento de RPA com Microsoft Power Automate Desktop: leitura e gravação de arquivos, interação com páginas da web, interação com programas da área de trabalho, agendamento de scripts

CAPÍTULO I - ANÁLISE DE DADOS - CONCEITOS E APLICAÇÕES

- 1.1. Estatística descritiva. Média, média aparada, moda, mediana, quartis e percentis, variância, desvio padrão, coeficiente de variação
- 1.2. Probabilidade, distribuições de probabilidade
- 1.3. Método de Monte Carlo: conceito, aplicações
- 1.4. População, amostra, amostragem
- 1.5. Inferência: estimação pontual e intervalar, intervalo de confiança
- 1.6. Testes de hipóteses
- 1.7. Séries temporais: decomposição, média móvel. Outliers
- 1.8. Normalização: mudança de escala, z-score, min-max
- 1.9. Regressão linear simples: conceito, suposições, linearização, coeficientes, R-quadrado, resíduos, valor-p, elasticidades, interpretação. Regressão linear múltipla: aplicações, riscos e limitações, interpretação
- 1.10. Dados: conceitos, atributos, métricas, transformação de dados
- 1.11. Análise de dados. Agrupamentos. Tendências. Projeções
- 1.12. Análise exploratória: aplicação de conceitos e técnicas; visualização
- 1.13. Conceitos e técnicas de transformação e remodelagem de dados: pivotamento, melting e outros
- 1.14. Gráficos: linha, base 100, barra, barra empilhada e normalizada, rosca, dispersão, histograma, radar, box plot, bolhas, mapa de calor, cascata e outros
- 1.15. Noções de data warehouse
- 1.16. Noções de analytics
- 1.17. Noções de business intelligence

1.1. Estatística descritiva. Média, média aparada, moda, mediana, quartis e percentis, variância, desvio padrão, coeficiente de variação

 **Memorize! Todo conteúdo deste tópico é muito cobrado nas provas quando previsto no edital do certame.**

A **Estatística Descritiva** é a área da Estatística responsável por coletar, organizar, resumir, descrever e apresentar dados de forma clara e objetiva, sem tirar conclusões generalizadas além do conjunto analisado.

Utilidades:

- Resumir grandes volumes de dados de maneira compreensível.
 - Identificar padrões e tendências.
 - Apoiar na tomada de decisão com base em informações quantitativas.
 - Servir de base para análises estatísticas mais avançadas (como a Inferencial).
-

Exemplos de Aplicação

- Análise de desempenho de sistemas computacionais (tempo médio de resposta).
 - Resumo de notas de concursos públicos.
 - Monitoramento de tráfego em redes de computadores.
 - Estudos sobre consumo de recursos em ambientes cloud.
 - Avaliação de tempos de execução de algoritmos.
-

Tecnologias Relacionadas

Ferramentas e linguagens usadas para cálculo e visualização de estatísticas descritivas:

- **Excel / LibreOffice Calc** (funções estatísticas).
 - **Python (bibliotecas Pandas, NumPy e SciPy)**.
 - **R (linguagem estatística)**.
 - **SQL** (uso de funções de agregação).
 - **Power BI / Tableau** (dashboards com indicadores estatísticos).
-

Tópicos mais cobrados em concursos públicos

✓ Média (Aritmética)

Despenca nas provas!

É a soma dos valores dividida pelo número total de observações.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

✓ Exemplo:

Notas: 7, 8, 9

Média = $(7 + 8 + 9) / 3 = 8$

✓ Média Aparada

É a média calculada após eliminar uma porcentagem fixa dos menores e maiores valores, reduzindo a influência de outliers.

✓ Exemplo:

Valores: 1, 2, 3, 4, 100

Eliminando 20% inferior e superior (1 e 100), a média aparada é $(2 + 3 + 4) / 3 = 3$.

✓ Moda

Despenca nas provas!

Valor que mais aparece em um conjunto de dados.

✓ Exemplo:

Valores: 2, 3, 3, 4

Moda = 3

♦ Pode não existir moda (se não houver repetições) ou haver múltiplas modas (bimodal, multimodal).

✓ Mediana

📢 Despenca nas provas!

Valor central de um conjunto de dados ordenados.

✓ Exemplo:

Valores: 1, 3, 5

Mediana = 3

Para números pares, a mediana é a média dos dois valores centrais.

✓ Quartis e Percentis

📢 Despenca nas provas!

- **Quartis** dividem os dados ordenados em quatro partes iguais:
 - Q1 (25% dos dados abaixo).
 - Q2 (50%, que coincide com a mediana).
 - Q3 (75% dos dados abaixo).
- **Percentis** dividem os dados em 100 partes. O percentil 90 indica que 90% dos dados estão abaixo dele.

✓ Exemplo:

Valores ordenados: 10, 20, 30, 40, 50

Q1 = 20, Q2 = 30, Q3 = 40.

✓ Variância (σ^2)

📢 Despenca nas provas!

Mede a dispersão dos dados em relação à média.

É a média dos quadrados dos desvios dos valores em relação à média.

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

✓ Exemplo:

Valores: 2, 4, 6

Média = 4

Variância = $[(2-4)^2 + (4-4)^2 + (6-4)^2] / 3 = (4 + 0 + 4) / 3 = 2,67$

✓ Desvio Padrão (σ)

📢 Despenca nas provas!

É a raiz quadrada da variância. Indica o quão dispersos estão os dados em relação à média.

$$\sigma = \sqrt{\sigma^2}$$

✓ Exemplo anterior:

Desvio padrão = $\sqrt{2,67} = 1,63$

✓ Coeficiente de Variação (CV)

📢 Despenca nas provas!

É a razão entre o desvio padrão e a média, expressa em percentual. Mede a dispersão relativa dos dados.

$$CV = \frac{\sigma}{\bar{x}} \times 100$$

✓ Exemplo:

Média = 4, Desvio padrão = 1,63

CV = $(1,63 / 4) \times 100 = 40,75\%$

- ◆ Quanto maior o CV, maior a variabilidade dos dados.
-

✓ Outliers (Valores Atípicos)

📢 Despenca nas provas!

- ◆ Definição:

Outliers são **valores que se distanciam significativamente** dos demais dados de um conjunto. Eles podem indicar erros, variabilidade incomum ou fenômenos especiais e podem afetar drasticamente medidas estatísticas.

◆ Utilidades:

- Identificar erros de digitação ou coleta.
- Detectar comportamentos anômalos (ex.: intrusões em redes, picos de consumo).
- Auxiliar em decisões sobre limpar, corrigir ou estudar fenômenos incomuns.

◆ Impacto nas Medidas:

Medida	Sensibilidade aos Outliers
Média	Alta (muito afetada)
Mediana	Baixa (pouco afetada)
Moda	Geralmente não afetada
Variância / Desvio Padrão	Alta

◆ Como identificar Outliers:

1) Regra dos Quartis (Boxplot)

Usa os limites inferior e superior:

$$\text{Limite Inferior} = Q1 - 1,5 \times AIQ$$

$$\text{Limite Superior} = Q3 + 1,5 \times AIQ$$

Onde:

- $Q1$ = Primeiro quartil (25%).
- $Q3$ = Terceiro quartil (75%).
- $AIQ = Q3 - Q1$ (Amplitude Interquartílica).

Valores fora desses limites são considerados outliers.

✓ Exemplo:

$Q1 = 10, Q3 = 30$

$AIQ = 20$

Limite Inferior = $10 - (1,5 \times 20) = -20$

Limite Superior = $30 + (1,5 \times 20) = 60$

→ Dados abaixo de -20 ou acima de 60 são outliers.

2 Regra dos Z-Scores

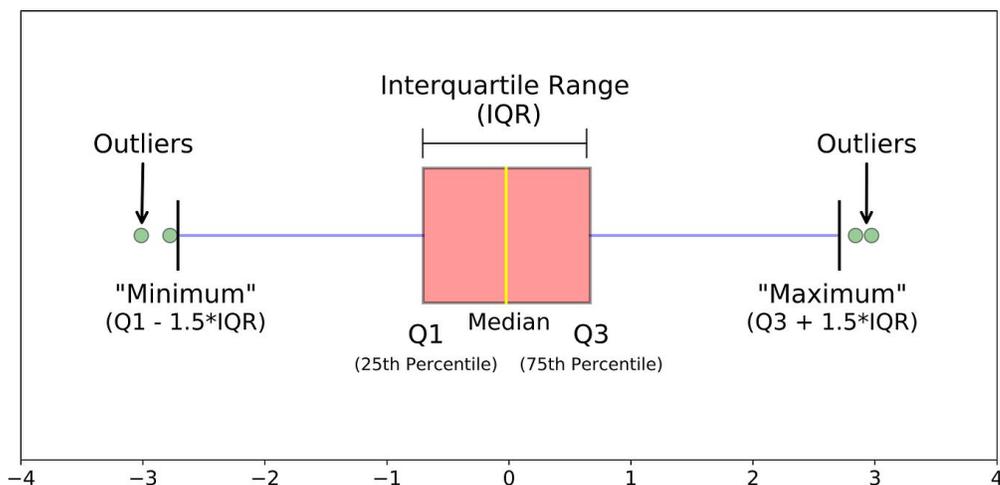
Para dados com distribuição aproximadamente normal:

$$Z = \frac{(x - \bar{x})}{\sigma}$$

Se $|Z| > 3$, o valor x é considerado outlier (ou seja, está a mais de 3 desvios padrão da média).

♦ O que fazer com outliers:

- **Corrigir** se for erro evidente.
- **Excluir** se distorcer a análise sem justificativa.
- **Manter** se for informação relevante (ex.: identificar ataques DDoS).



Resumo!

- **Média** é sensível a valores extremos (outliers).
- **Mediana** é resistente a outliers.
- **Moda** é útil para dados qualitativos.
- **Desvio padrão** e **variância** indicam dispersão: quanto maiores, mais espalhados estão os dados.
- **CV** é muito cobrado para comparar variabilidade de conjuntos com médias diferentes.

Memorize!

Medida	Descrição	Sensível a Outliers?
Média	Soma dos valores / n° total	Sim
Média Aparada	Média sem extremos	Menos sensível
Moda	Valor mais frequente	Não
Mediana	Valor central	Não
Variância	Dispersão quadrática	Sim
Desvio Padrão	Raiz da variância	Sim
Coefficiente de Variação	Dispersão relativa (%)	Sim
Quartis / Percentis	Dividem dados em partes iguais	Não

Outliers (Valores Atípicos)

Outliers são **valores que se distanciam significativamente** dos demais dados de um conjunto. Eles podem indicar erros, variabilidade incomum ou fenômenos especiais e podem afetar drasticamente medidas estatísticas.

Lista de Questões no Estilo FGV para estatística descritiva. Média, média aparada, moda, mediana, quartis e percentis, variância, desvio padrão, coeficiente de variação

1. Qual das medidas abaixo indica o valor que ocorre com maior frequência em um conjunto de dados?

- Mediana
- Moda

- c) Média
- d) Percentil
- e) Quartil

💡 Comentário: A moda é a medida de tendência central que corresponde ao valor mais frequente em um conjunto de dados. Um conjunto pode ser amodal (sem moda), unimodal (uma moda), bimodal (duas modas) ou multimodal (mais de duas).

✅ Resposta: b

2. Quando calculamos a média excluindo uma porcentagem dos menores e maiores valores, estamos utilizando qual conceito?

- a) Moda
- b) Mediana
- c) Média aparada
- d) Primeiro quartil
- e) Variância

💡 Comentário: A média aparada é útil para reduzir a influência de valores extremos (outliers), pois descarta uma porcentagem pré-determinada dos menores e maiores valores antes de calcular a média.

✅ Resposta: c

3. Qual medida de dispersão é calculada como a raiz quadrada da variância?

- a) Coeficiente de variação
- b) Desvio padrão
- c) Moda
- d) Percentil
- e) Média aparada

💡 Comentário: O desvio padrão é a medida de dispersão que expressa o grau de variação dos dados em torno da média e é obtido pela raiz quadrada da variância.

✅ Resposta: b

4. Ao dividir um conjunto de dados ordenados em quatro partes iguais, identificamos:

- a) Percentis
- b) Moda
- c) Quartis

- d) Variância
- e) Média

💡 Comentário: Os quartis dividem um conjunto ordenado em quatro partes iguais. O primeiro quartil (Q1) separa os 25% menores dados, o segundo quartil (Q2) é a mediana e o terceiro quartil (Q3) marca 75% dos dados.

✅ Resposta: c

5. A mediana de um conjunto de dados corresponde ao valor:

- a) Da soma dos valores dividida pelo total de elementos
- b) Mais frequente do conjunto
- c) Que ocupa a posição central nos dados ordenados
- d) Obtido pela diferença entre o maior e o menor valor
- e) Que representa 10% dos dados

💡 Comentário: A mediana é o valor central de um conjunto de dados ordenados. Caso o total de dados seja par, a mediana será a média aritmética dos dois valores centrais.

✅ Resposta: c

6. 📌 O coeficiente de variação (CV) serve para:

- a) Medir a frequência de um valor
- b) Comparar a dispersão relativa entre diferentes conjuntos de dados
- c) Calcular a média aparada
- d) Determinar a moda
- e) Calcular a diferença entre quartis

💡 Comentário: O coeficiente de variação é usado para comparar a dispersão relativa de conjuntos de dados com médias diferentes. Quanto maior o CV, maior a variabilidade relativa.

✅ Resposta: b

7. Sobre a variância amostral, assinale a afirmativa correta:

- a) É igual ao quadrado do desvio padrão
- b) Mede a tendência central dos dados
- c) Corresponde ao somatório das frequências dividido pelo total de elementos
- d) Representa o intervalo interquartil
- e) É a média aparada ao quadrado

💡 Comentário: A variância mede o grau de dispersão dos dados em relação à média e é calculada como a média dos quadrados dos desvios em relação à média, utilizando o divisor (n-1) para amostras.

✅ Resposta: a

8. 📌 **Percentis dividem o conjunto de dados ordenados em:**

- a) 2 partes
- b) 4 partes
- c) 10 partes
- d) 100 partes
- e) 50 partes

💡 Comentário: Percentis são medidas que dividem o conjunto ordenado em 100 partes iguais. Por exemplo, o percentil 75 indica que 75% dos dados são menores ou iguais a esse valor.

✅ Resposta: d

9. **Em distribuições simétricas, podemos afirmar que:**

- a) Média e moda são sempre diferentes
- b) A mediana é menor que a média
- c) Média, mediana e moda são iguais ou muito próximas
- d) A variância é zero
- e) O desvio padrão é negativo

💡 Comentário: Em distribuições perfeitamente simétricas, como a normal, média, mediana e moda coincidem.

✅ Resposta: c

10. **O coeficiente de variação (CV) é calculado por meio da seguinte fórmula:**

- a) $(\text{Moda} \div \text{Média}) \times 100$
- b) $(\text{Desvio padrão} \div \text{Média}) \times 100$
- c) $(\text{Variância} \div \text{Média}) \times 100$
- d) $(\text{Mediana} \div \text{Moda}) \times 100$
- e) $(\text{Quartil} \div \text{Média}) \times 100$

💡 Comentário: O CV expressa a dispersão relativa, sendo obtido ao dividir o desvio padrão pela média e multiplicar por 100 para expressar em percentual.

✅ Resposta: b

11. Quando um conjunto de dados apresenta mais de um valor com maior frequência, dizemos que:

- a) A média é multimodal
- b) A moda é múltipla
- c) A mediana é repetida
- d) Os quartis são duplicados
- e) A variância é nula

💡 Comentário: Quando há dois ou mais valores com maior frequência, a distribuição é chamada de bimodal ou multimodal, e a moda é múltipla.

✅ Resposta: b

12. 📌 Um conjunto de dados com alta variabilidade apresentará:

- a) Desvio padrão baixo
- b) Média inferior à mediana
- c) Coeficiente de variação elevado
- d) Quartis concentrados
- e) Moda representativa

💡 Comentário: Alta variabilidade implica maior dispersão dos dados, o que reflete em desvio padrão e coeficiente de variação elevados.

✅ Resposta: c

13. A média aparada é útil em conjuntos com outliers porque:

- a) Amplifica a influência dos extremos
- b) Elimina todos os dados iguais
- c) Reduz o impacto dos valores extremos
- d) Zera a variância
- e) Substitui a mediana

💡 Comentário: A média aparada descarta os dados extremos e calcula a média com os valores restantes, reduzindo o efeito dos outliers.

✅ Resposta: c

14. No cálculo da variância amostral, utiliza-se (n-1) no denominador para:

- a) Aumentar a média



www.kuasarnex.com

 [@kuasarnex](https://www.instagram.com/kuasarnex)