

Moda e sua definição

Transcrição

[0:00] Legal galera, continuando aqui com medidas de tendência central, vamos falar agora da moda, e a moda é o mais fácil da gente definir, porque a Moda é justamente o valor mais frequente dentro de um conjunto de dados, ok?

[0:11] Vamos fazer aqui bem rápido, temos o nosso DF, vou mostrar pra vocês, é bem simples de entender. O nosso DF aqui, a gente pode observar que Beltrano tirou 10, 2, meio, 1, 3, 9.5 e 10. Conjunto pequeno, a gente consegue ver que a moda, o valor que mais aparece é o 10. Aqui no fulano, podemos olhar, 10 aparece duas vezes, 8 aparece 3 vezes, então podemos identificar que a moda do fulano, a nota modal é 8.

[0:45] Você pode ter, obviamente, mais de uma moda dentro de um conjunto de dados, quando a gente tem duas modas, a gente chama que a variável é bimodal. Se tem mais de duas, começamos a chamar de multimodal.

[0:55] Vou dar uns exemplos, vamos fazer com nosso amigo DF, que são as notas, pra obter a moda com pandas, Mode. Bem simples assim, vai te retornar um dataframe com valores modais, no caso só tem um pra esse, um pra esse, e um pra esse, 8, 8, e 8.

[1:16] Esse também, 8 aparece aqui, aqui e aqui. 7 aparece duas vezes, 8, três, então é isso mesmo.

[1:22] Eu dei mole no exemplo do DF, não botei um bimodal, um multimodal. Então vamos fazer um rapidinho, exemplo vai ser igual a PD, vamos criar uma series, aqui dentro, vou passar uma lista com valores dentro, 1, 2, 2 repete, 3 não, 4 repete, 0 5 não, 0 6 repete.

[1:50] o 2 tá repetindo, o 4 e o 6. Vamos ver essa series, exemplo. Tá aqui minha series, esse aqui são os valores, esse é o index. Vou fazer a moda disso aqui, exemplo.mode.

[2:10] Olha o que ele joga pra mim, uma series com os 3 valores modais, esse é multimodal, se mudar pra 7, vai me jogar só 2 valores, o 2 e o 4, moda é isso, o que mais repete. Esse tipo de estatística descritiva é muito útil, muito utilizada com variáveis do tipo qualitativo, porque as continuas, você pode ter muitas modas, a não ser que você consiga alisar uma distribuição pra tentar identificar uma ou duas modas, mas geralmente esse tipo de estatística utilizamos com valores categóricos, valores qualitativos.

[2:58] Um exemplinho do nosso amigo, nosso dataset que temos que fazer uma análise descritiva, vamos fazer com a renda, que não é muito boa pra isso, mas vamos ver que tipo de moda, eu acredito que vai ser aquele valorzinho de salário mínimo.

[3:13] Não deu outra, 788, como tínhamos visto a uns vídeos atrás, é o valor do salário mínimo, quer dizer, a moda, todo mundo, a maioria das pessoas está recebendo salário mínimo. Vamos fazer com outro. Dados, vamos fazer com altura, pra ver que vai dar duas modas. Altura é uma variável que criei, a gente não mexeu com ela. Mode. 9 modas, você vê que num valor continuo é complicado. Não significa muita coisa pra gente.

[3:55] O próximo vídeo, vamos ver uma relação que existe entre as três variáveis, média, estatística e descritiva, que a gente acabou de ver, média, mediana e moda, quando a gente consegue identificar simetrias na distribuição, probleminhas na distribuição, somente olhando as estatísticas, a gente tem as três, começamos olhar se é simétrica direita, esquerda, ou simétrica, no próximo vídeo começamos a falar sobre isso, beleza?

