CENTRO UNIVERSITÁRIO DO PARÁ - CESUPA

ESCOLA DE NEGÓCIOS, TECNOLOGIA E INOVAÇÃO - ARGO

CURSO DE ENGENHARIA DA COMPUTAÇÃO

ERNESTO BERTOLDO BRITO COSTA

DIEGO YUTA SATO

CAIO PEREIRA GILLET MACHADO

**ANÁLISE DA QUALIDADE DO VINHO**

BELÉM

ANO

**1. INTRODUÇÃO**

A indústria vinícola é, atualmente, de potencial crescente em todo o mundo. Novas técnicas e tecnologias têm assegurado a produção de qualidade não apenas em regiões tradicionais na produção do vinho, mas também em novas áreas, onde um conhecimento mais avançado sobre a produção de uvas só adquirido no século passado com o advento de novas tecnologias de medição, higiene e padronização das variedades de uvas permitiu o cultivo nestas regiões.

Países como a França, Portugal e Itália enfrentam concorrência acirrada de vinhos sul-americanos, advindos de regiões remotas do Chile e Argentina. Mais recentemente, grandes críticos mundialmente respeitados têm reconhecido as qualidades de novos produtores emergentes, como Austrália e África do Sul. Estes sommeliers1 têm destacado a presença de vinhos de tons marcantes nestas regiões, produzidos com uvas Shiraz e Cabernet Sauvignon, principalmente. O Brasil talvez seja um dos últimos entrantes nesta terceira geração de países produtores e exportadores de vinhos; sua diversidade de climas, extensão territorial e solos repletos de nutrientes são mais que suficientes para a produção vinícola em escala comercialmente aceitável, ainda mais para uma fruta que só pode ser encontrada na Europa em determinadas regiões tradicionais, onde a demanda alta por vinhos de qualidade acaba culminando em vinhos caríssimos. No entanto, os produtores brasileiros necessitam aderir a critérios e certificações de produção e qualidade mundialmente reconhecida, e talvez até criar barreiras legais para a produção de vinhos de baixa qualidade, já que grandes exportações e reconhecimento internacional são consequência da fama dos vinhos aqui produzidos. Por quanta desses motivos, resolvemos pesquisar e demonstrar através desse artigo a qualidade de vinho.

**2. METODOLOGIA**

Para a realização da coleta de dados usamos o site uci machine learning, Nele retiramos uma data set onde podíamos coletar dados sobre a qualidade, o total sulfur dioxide e o teor de açúcar. Com os dados já coletados criamos um código em Python para analisar os dados solicitados na questão e criar um histograma para melhor visualização dos dados

**3. RESULTADOS**

Neste tópico iremos apresentar os nossos dados retirados do nosso código Python através de imagens, gráficos e histogramas.

**Valores qualitativos:**

**Figura 1 -** Histograma

Interface gráfica do usuário, Texto

Descrição gerada automaticamente Fonte: Diego Yuta,2022

Podemos ver no gráfico que a média está bem no meio entre os dois maiores valores não tendo uma discrepância no resultado gerado.

**Figura 2 –** Média

Padrão do plano de fundo

Descrição gerada automaticamenteFonte: Diego Yuta, 2022

**Valores quantitativos discretos:**

**Figura 3 -** Média e Mediana

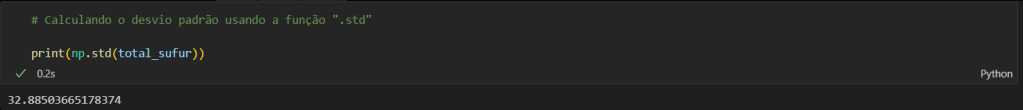
Texto

Descrição gerada automaticamente

Fonte: Diego Yuta,2022

Podemos observar que os valores da média e da mediana estão muito distantes uma das outras. Em outras palavras, a mediana, sendo o número do "meio", se colocado em ordem crescente os valores coletados, estão bem distantes da média geral dos dados coletados

**Figura 4 -** Desvio Padrão



Fonte: Diego Yuta,2022

Ao analisar o desvio padrão, notamos que o desvio dos dados é bem grande, dando um panorama maior da comparação com a média e mediana anteriormente mostrado.

**Figura 5** - Histograma de frequência relativa

Interface gráfica do usuário, Aplicativo

Descrição gerada automaticamente

Fonte: Diego Yuta,2022

Diante dos dados mostrados, podemos ver que ao analisar o histograma, as colunas estão ordenadas de uma forma distorcida a direita quando a distribuição de dados indica a ocorrência de altos valores com baixa frequência.

**Figura 6 -** Diagrama de Caixa

Interface gráfica do usuário, Aplicativo

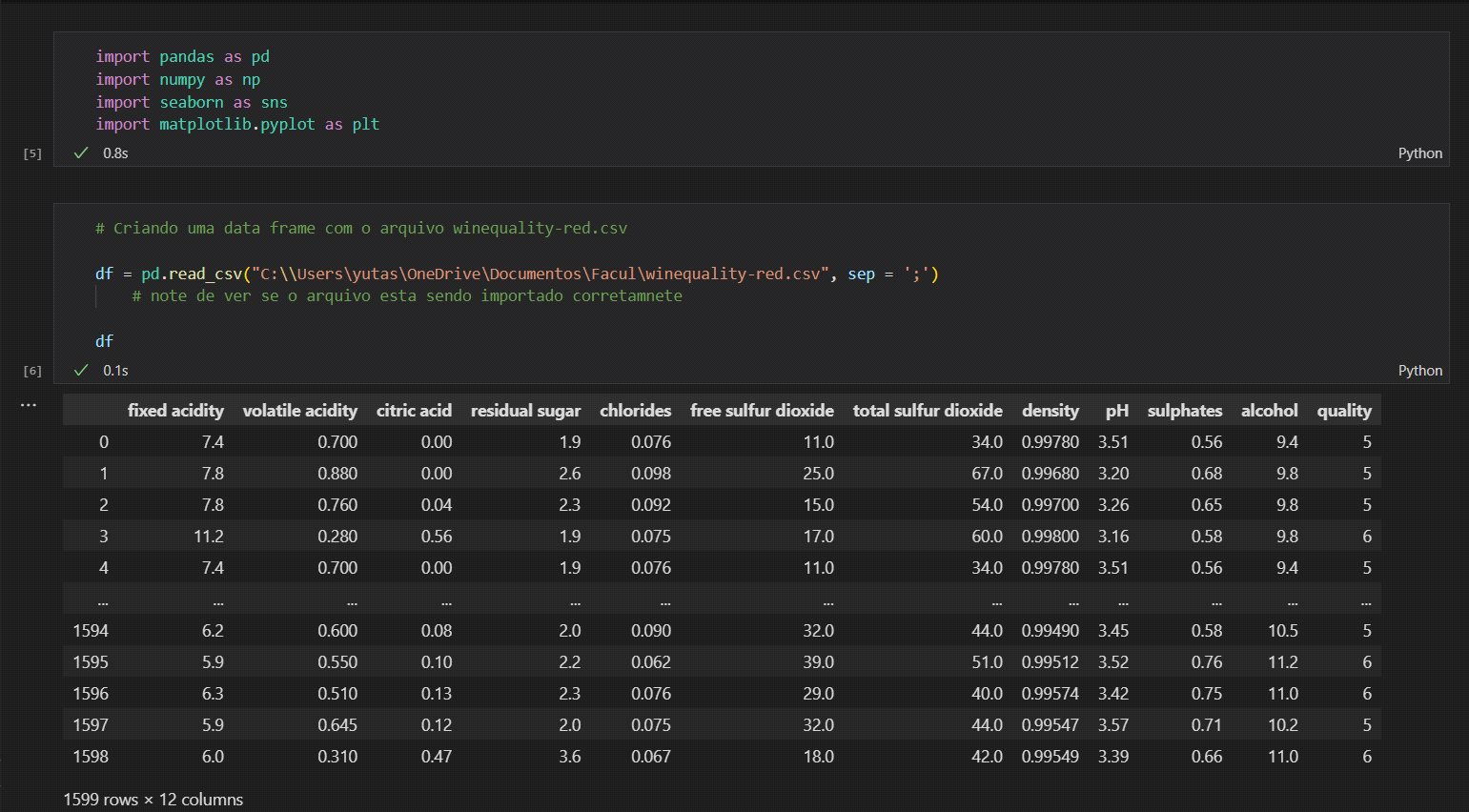
Descrição gerada automaticamente

Fonte: Diego Yuta,2022

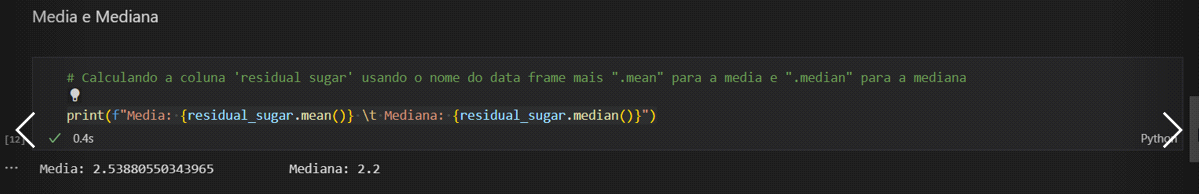
Ao ver o diagrama de caixa, podemos ver com clareza a distribuição onde podemos ver dois pontos bem a direita que fazem os nossos valores retirados terem uma clara discrepância e múltiplos pontos bem perto da caixa podendo ver o grau de dissenção dos dados

**Valores quantitativos contínuos:**

**Figura 7 -** Qualidade do vinho

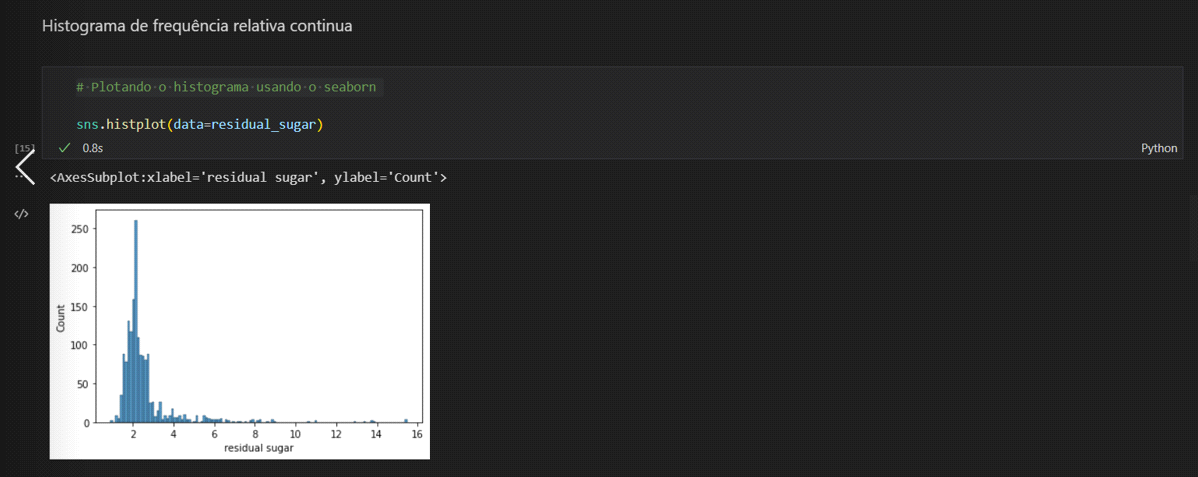
Fonte: Diego Yuta,2022

**Figura 8 -** Média e Mediana

 Fonte: Diego Yuta,2022

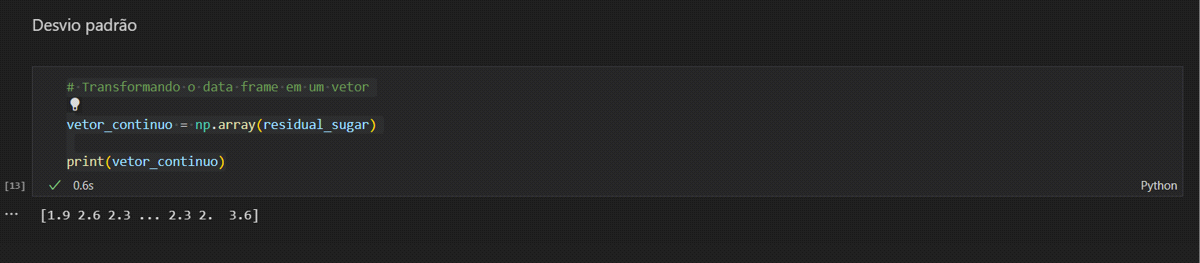
Podemos ver que a média e a mediana estão bem pertos uma da outra não tendo uma grande variação

**Figura 9 -** Histograma de frequência relativa contínuos

Fonte: Diego Yuta,2022

Porém ao ver o histograma, podemos analisar que existem um pico extremamente grande no gráfico, e mais à esquerda, diversas colunas pequenas espalhadas sendo que bem a esquerda podemos ver uma grande concentração dos dados

**Figura 10 -** Desvio Padrão

Fonte: Diego Yuta,2022

Ao ver o desvio padrão, podemos notar que não tem um desvio muito grande dos dados que está sendo analisado. Sendo assim, os dados estão bem uniformes estando perto do 0

**Figura 11 -** Diagrama de Caixa

Fonte: Diego Yuta,2022

No diagrama de caixa, podemos ver essa dispersão com mais clareza ao olhar os pontos fora da caixa. Nota-se, também, que podemos ver um ponto muito distante, sendo esse o ponto 16

**4. REFERÊNCIAS BIBLIOGRÁFICAS**

# NETTO, AMÍLICAR NETTO; Python Para Data Science: E Machine Learning Descomplicado. 1ª edição. Rio de Janeiro - RJ:  Alta Books (8 julho 2021)

# MAGALHÃES, MARCOS NASCIMENTO MAGALHÃES; Noções de Probabilidade e Estatística:  7ªedição. Universidade de São Paulo -SP:  Edusp (1 janeiro 2007)

# PIMENTEL, ROBERTA LIANA PIMENTEL Manual para elaboração de trabalhos acadêmicos: Guia prático do estudante 3ª edição. Petrópolis -RJ: Vozes (1 janeiro 2014)