

Linear regression and metrics

Jupyter e git

Fazer cada atividade no Jupyter (Collab). Poste apenas os links

Para cada parte da atividade, colocar o comando do problema e fazer uma introdução com figuras, equações e texto contextualizando o problema

A

Com o mesmo dataset da Atividade 2, faça **boxplots** de duas features do dataset, conforme procedimento mostrado em sala. Para cada boxplot calcule:

- CV
- Curtose
- Skewness
- Mean
- Moda

Analise os resultados.

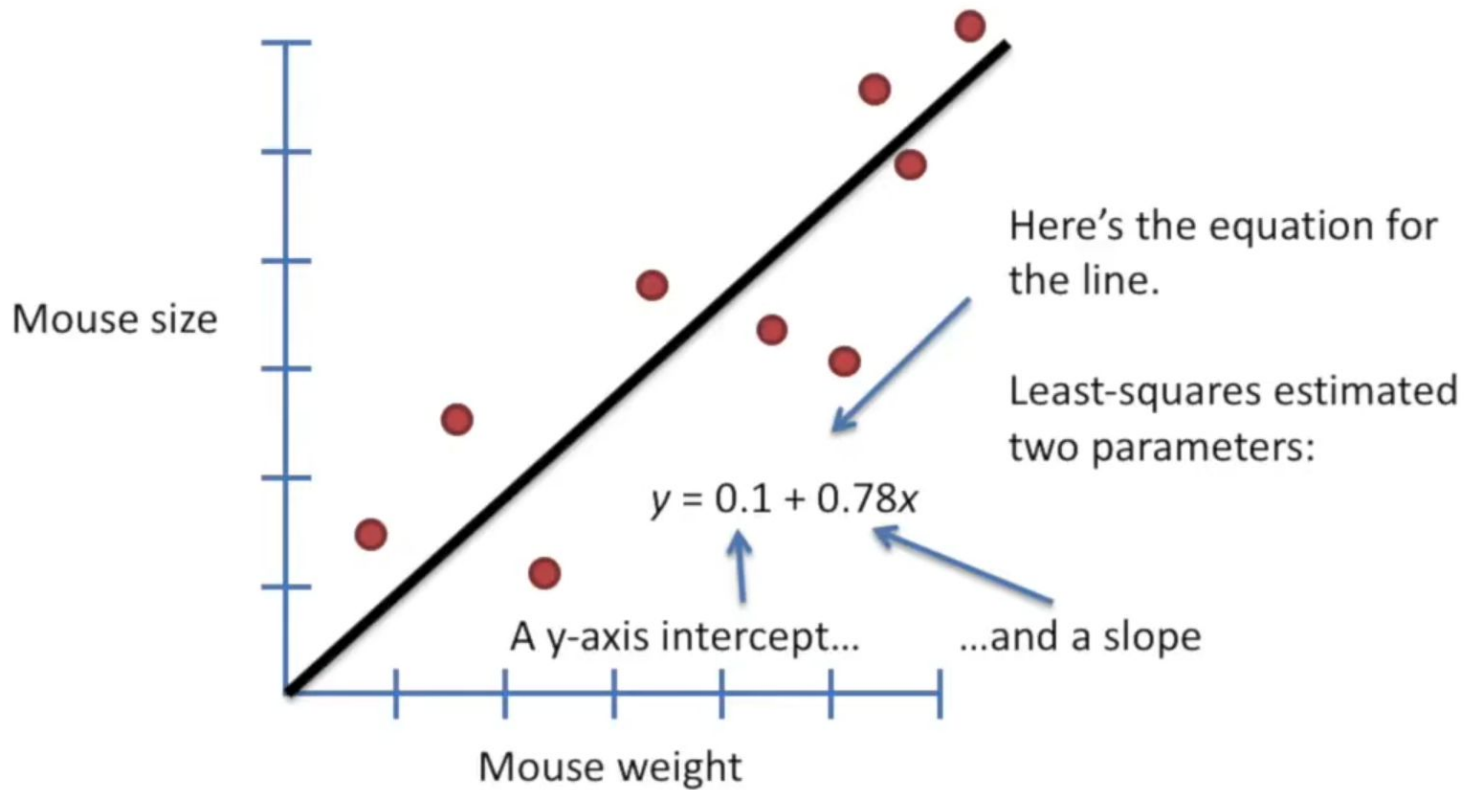
Para um desses boxplots, introduza artificialmente outliers e verifique o que ocorre com o boxplot e essas métricas.

Não esqueça do contexto sobre o dataset usado, com a explicação sobre ele, suas classes e features.

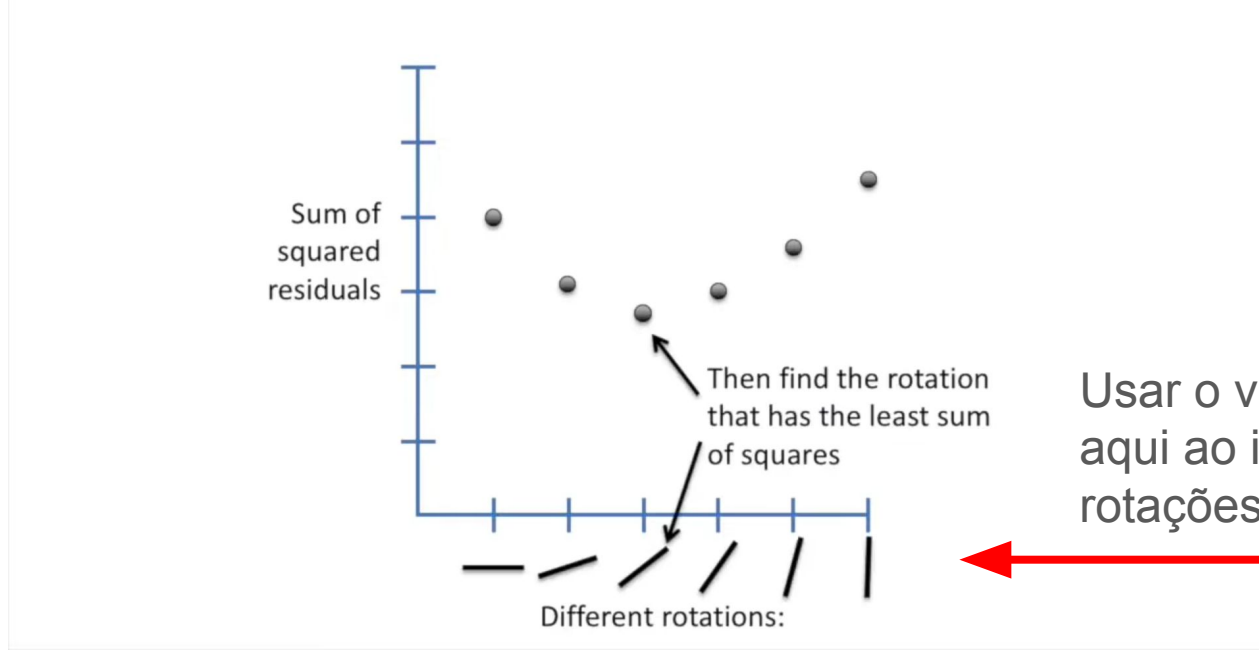
B

[Part 6 - Linear Regression.pptx](#)

Now we have fit a line to the data! This is awesome!



Reproduza o mais próximo possível o exemplo do slide anterior. Em seguida, aplique o *linear regression* do scikit-learn. Use a biblioteca `import statsmodels.api as sm` para gerar a estatísticas da regressão linear. A partir do b obtido, use 6 diferentes a 's para plotar diferentes retas, mostrando a variação na rotação e gere também o gráfico a seguir:



Calcule o R-Squared para cada 'a' e compare com o melhor obtido pelo scikit-learn.

Análise também:

R-squared:

Adj. R-squared:

F-statistic:

Prob (F-statistic):

P-value