

## Dec 5th meeting notes

---

### - Reminder:

- sign up for the Holiday Gift Exchange and to fill out the DS website Form
  - [https://docs.google.com/spreadsheets/d/142u9wDt\\_E1PMYDR5HF2bXrpz-5XOaFyAE4-19ebjSq8/edit?usp=sharing](https://docs.google.com/spreadsheets/d/142u9wDt_E1PMYDR5HF2bXrpz-5XOaFyAE4-19ebjSq8/edit?usp=sharing)

### - Presentation by Ascari on Linear Regression

- Create the linear regression model on **gasoline sales over years** using python through Jupyter notebook

<https://github.com/Cornell-MSBA-DS/docs/blob/main/lessons/linear%20regression/Project%20%20Linear%20Regression%20Ascari.ipynb>

- If you haven't download the cleaned up dataset:  
[https://drive.google.com/drive/u/1/folders/1\\_pgXO-emL62Cc49oZuMlheDldHALQ51Y](https://drive.google.com/drive/u/1/folders/1_pgXO-emL62Cc49oZuMlheDldHALQ51Y)
- Recommend to open up a new Jupyter notebook to plot the codes and understand the codes line by line
  - You would have to change the code "data = pd.ExcelFile('C:/Users/achot/Downloads/data\_science\_club\_data.xlsx')" into where you store the file on your own computer
- The first part of the codes create a linear regression
- The second part of the codes focus on creating a train-test split for ML algorithm
  - More information regarding the train-test split can be found here:  
<https://machinelearningmastery.com/train-test-split-for-evaluating-machine-learning-algorithms/>

### - data science club website: <https://cornell-msba-ds.github.io/website/index.html>

- Complete the bio survey: <https://forms.gle/mPgXCyDuiWwHssLv8> so we could put your bio onto the website for the club

### - deadline on project 2: Jan. 9th 2023