

# Calculating correlation coefficients

INTRODUCTION TO PYTHON IN POWER BI



**Jacob H. Marquez**

Data Scientist

# What is a correlation coefficient?

**Definition:** *a numerical measure of some type of statistical relationship between two variables*

**Range:** -1 to 1

**Range**

-1:

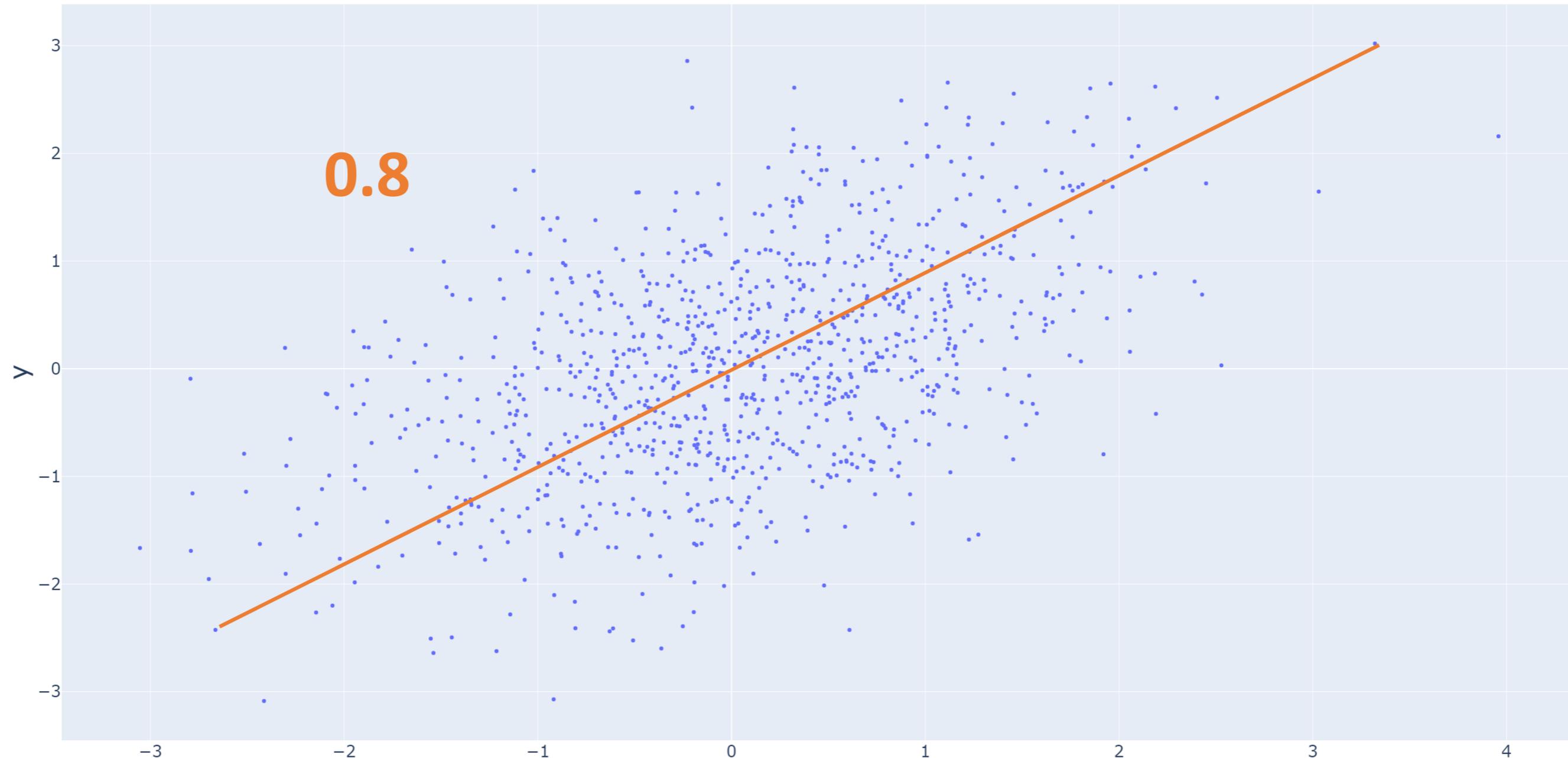
- **strong, negative relationship**
- increase in variable A is associated with a decrease in variable B

1:

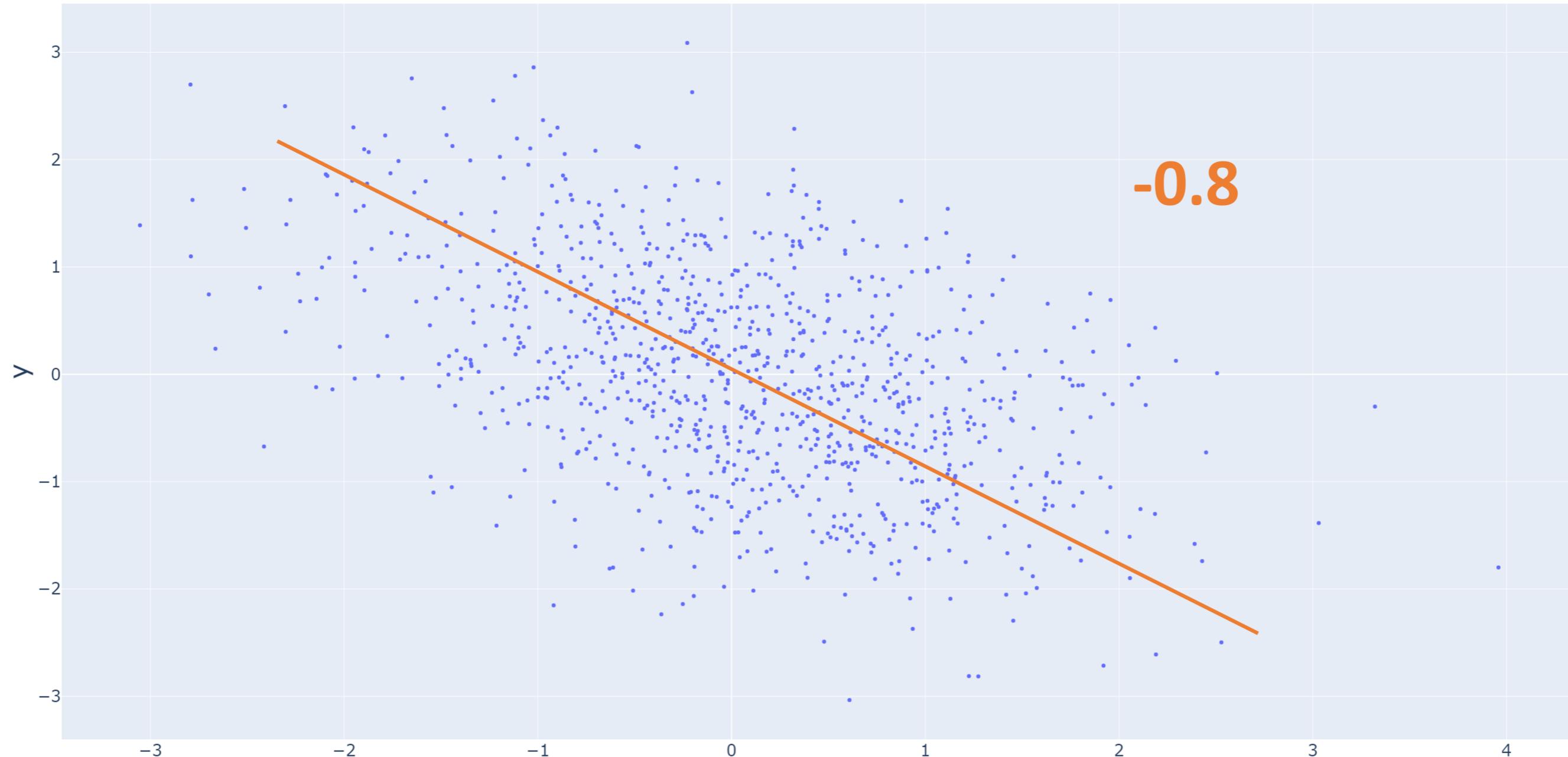
- **strong, positive relationship**
- increase in variable A is associated with an increase in variable B

0: **no relationship**

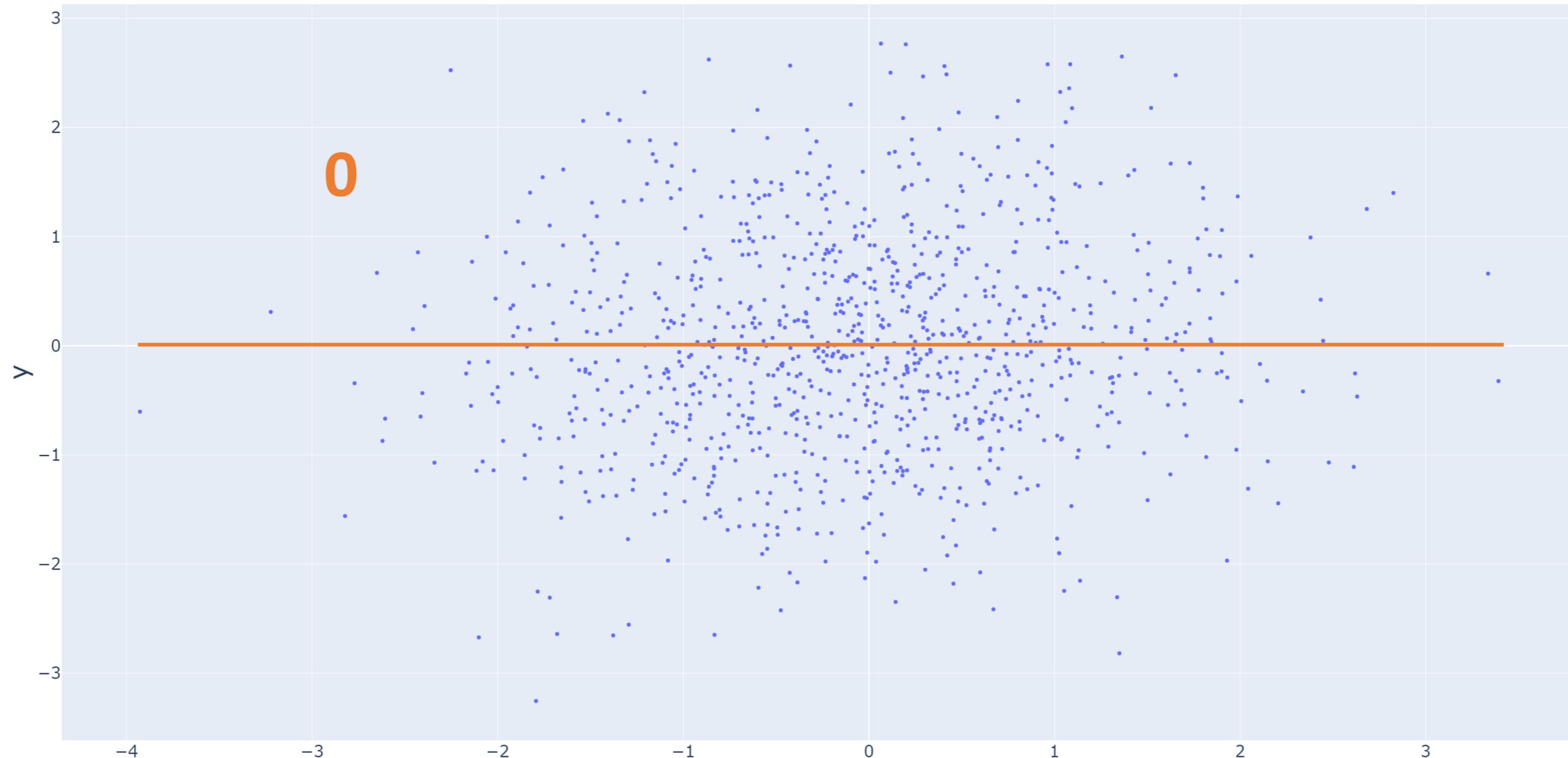
# Correlation coefficient example #1



# Correlation coefficient example #2



# Correlation coefficient example #3



# Correlation matrix

	Income	MntWines	MntRegularProds	MntMeatProducts
Income	1.0	0.73	0.82	0.70
MntWines	0.73	1.0	0.90	0.59
MntRegularProds	0.82	0.90	1.0	0.86
MntMeatProducts	0.70	0.59	0.86	1.0

# Correlation matrix

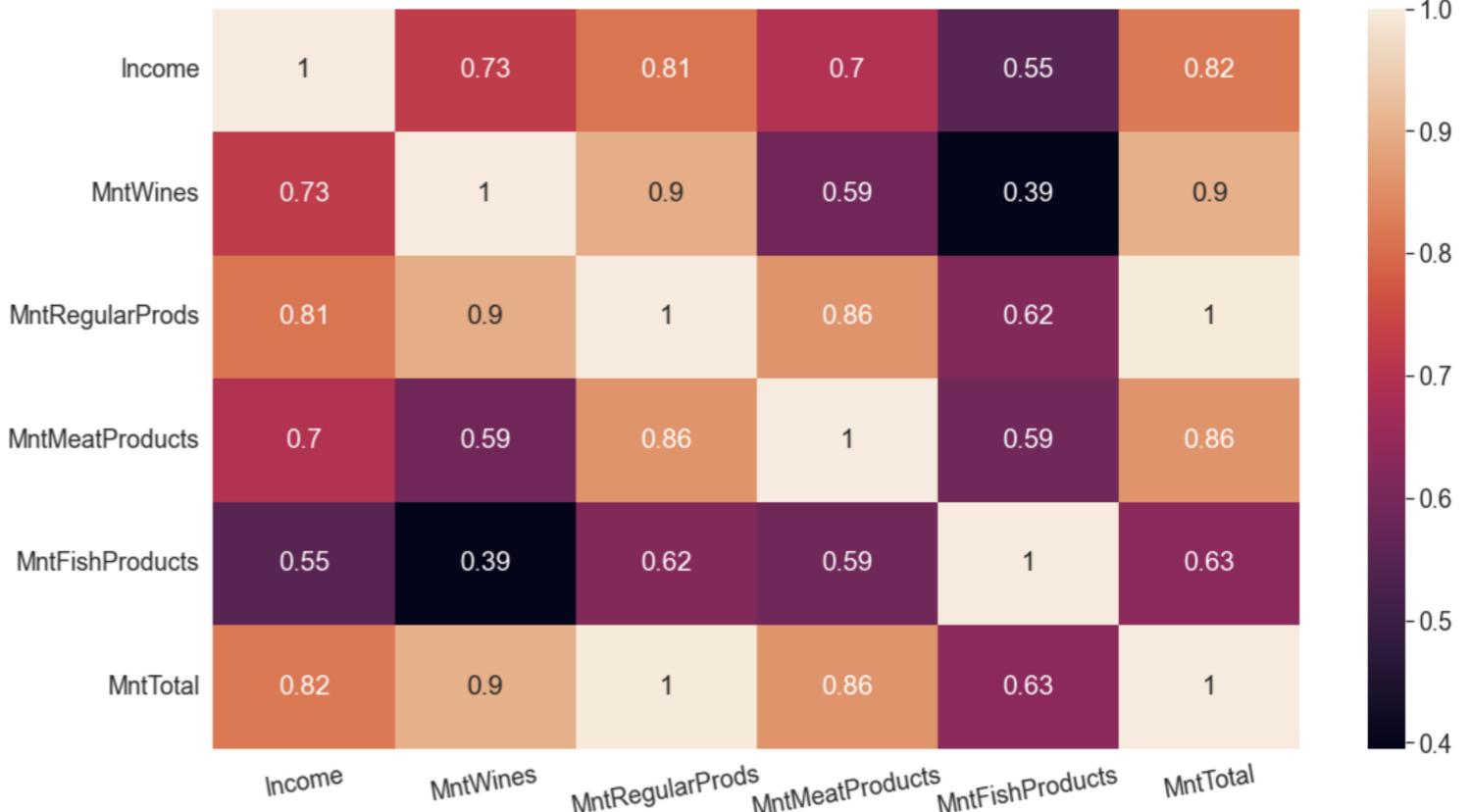
	Income	MntWines	MntRegularProds	MntMeatProducts
Income	1.0	0.73	0.82	0.70
MntWines	0.73	1.0	0.90	0.59
MntRegularProds	0.82	0.90	1.0	0.86
MntMeatProducts	0.70	0.59	0.86	1.0

# Correlation matrix

	Income	MntWines	MntRegularProds	MntMeatProducts
Income	1.0	0.73	0.82	0.70
MntWines	0.73	1.0	0.90	0.59
MntRegularProds	0.82	0.90	1.0	0.86
MntMeatProducts	0.70	0.59	0.86	1.0

# Correlation heatmap

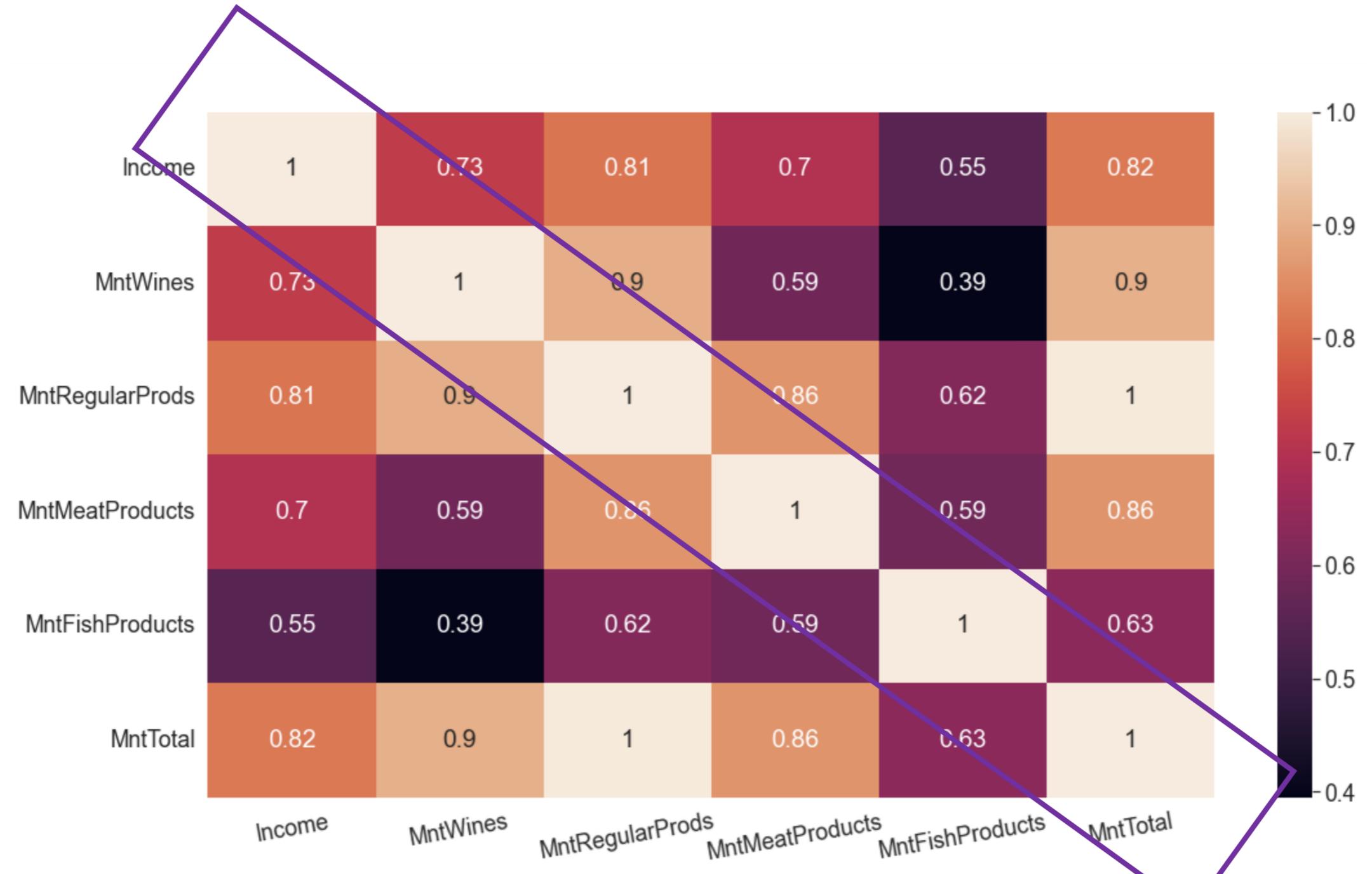
```
import seaborn as sns  
  
corrMatrix = dataset.corr()  
  
sns.heatmap(  
    corrMatrix,  
    annot=True  
)
```



# Correlation heatmap example

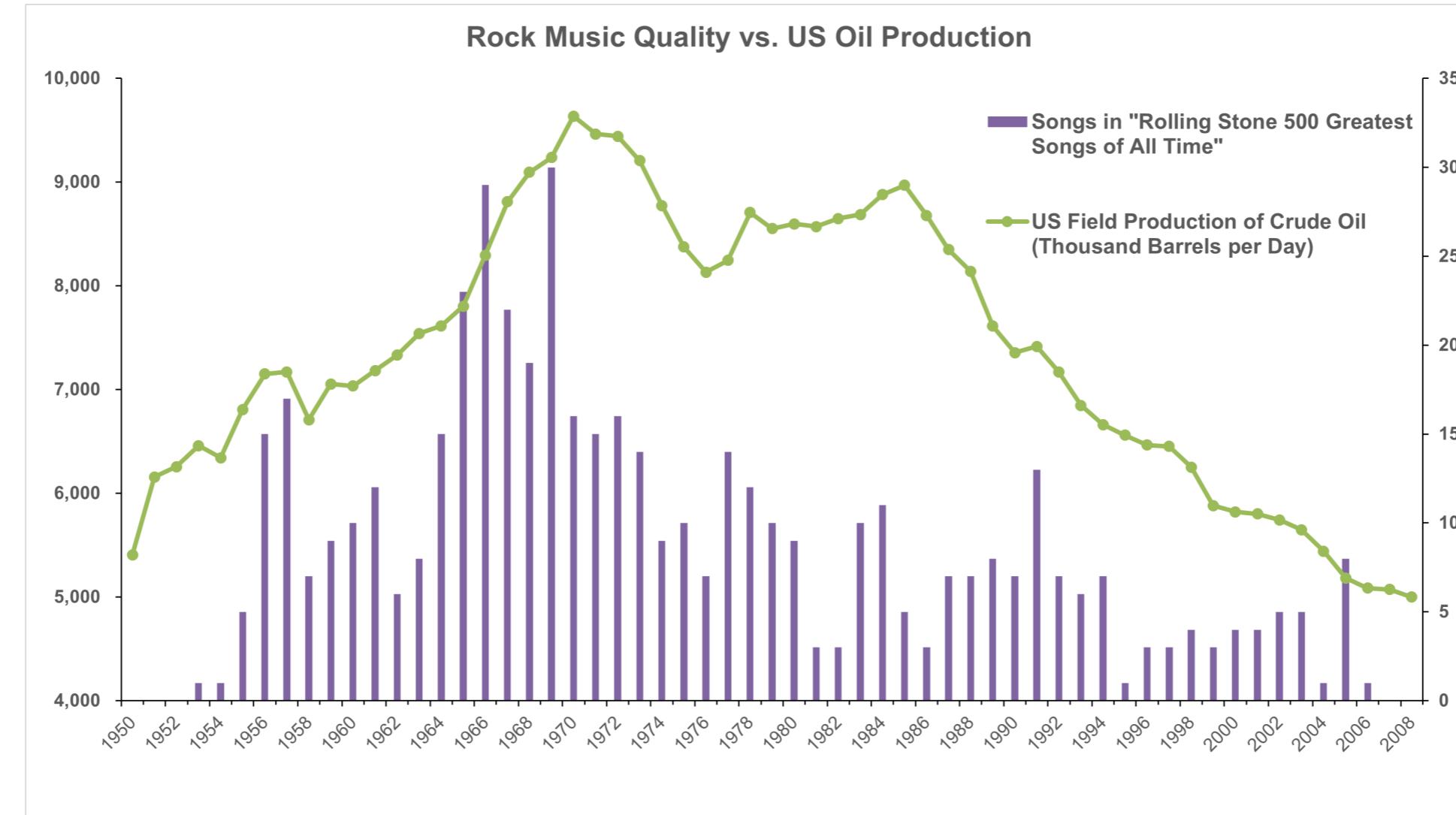


# Correlation heatmap example



# Correlation does not mean causation

- Strong correlative relationship  $\leftrightarrow$  One caused the other
- Causal relationship typically requires experimentation



# **Let's practice!**

**INTRODUCTION TO PYTHON IN POWER BI**

# How to calculate in Power BI and Python

INTRODUCTION TO PYTHON IN POWER BI



**Jacob H. Marquez**

Data Scientist

# **Let's practice!**

**INTRODUCTION TO PYTHON IN POWER BI**

# Wrap-up

## INTRODUCTION TO PYTHON IN POWER BI

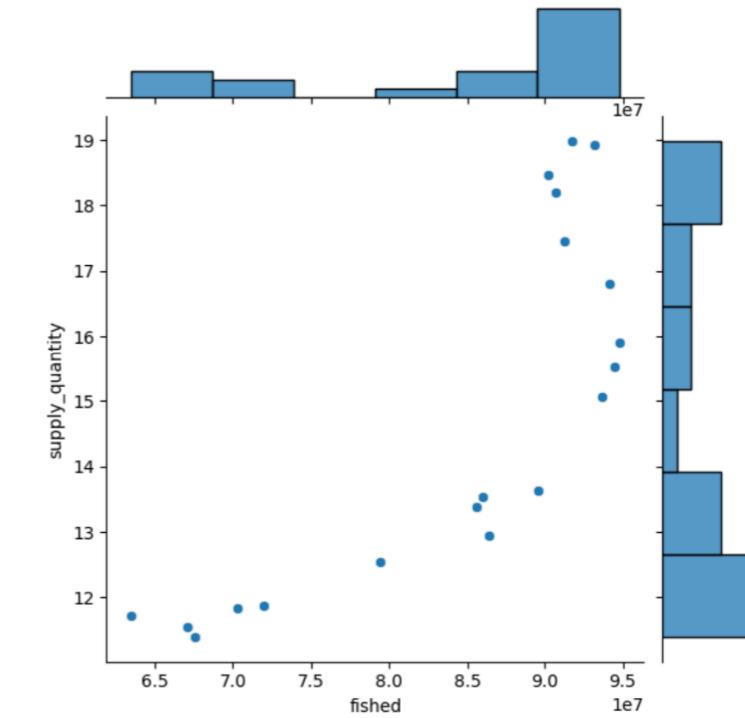


**Jacob H. Marquez**

Data Scientist

# Summary

- Explore the differences between Power BI and Python in performing certain tasks (e.g., data prep and transformations)
- Create custom visualizations using the Seaborn package
- Calculate the correlation coefficient



# Next Steps

## Python

- Seaborn and Matplotlib
- Clean data and manipulating
- Time series and machine learning

## Power BI

- Data visualizations
- Report design

# Congratulations!

INTRODUCTION TO PYTHON IN POWER BI