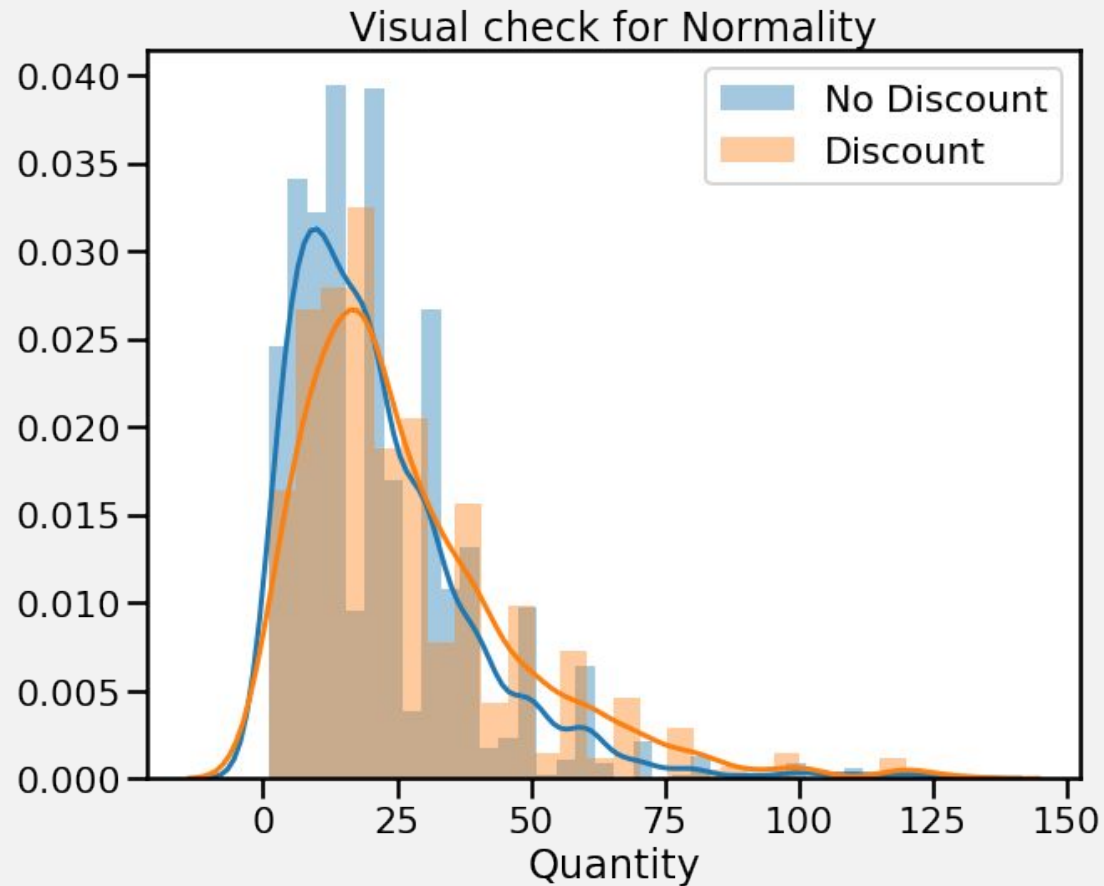# BUSINESS ANALYSES USING THE NORTHWIND DATASET

By: Andy Martin del Campo

WHAT IS THE NORTHWIND DATASET?

- Northwind is a SQL dataset created by Microsoft

- Tables or Categories include:

  - Employees

  - Customers

  - Products

  - Orders

  - Etc.

- This presentation will show a few statistical relationships between tables
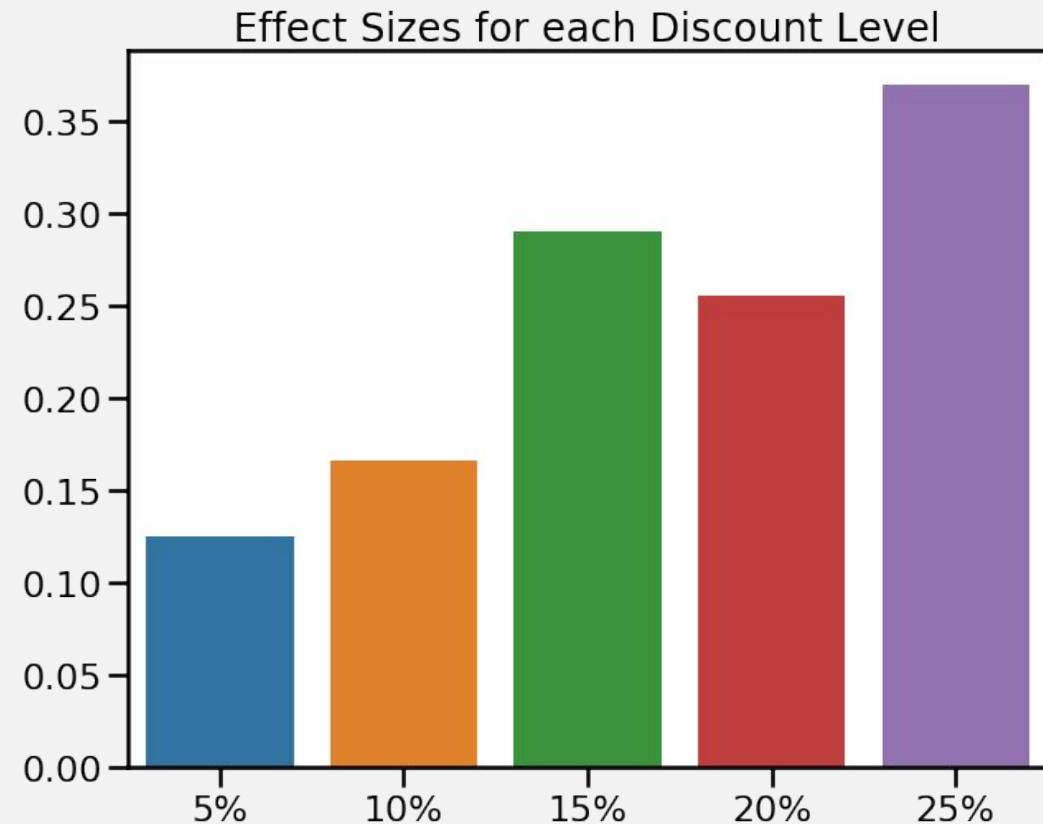
CASE 1: DO DISCOUNT AMOUNTS HAVE A STATISTICALLY SIGNIFICANT EFFECT ON THE QUANTITY OF A PRODUCT IN AN ORDER?

Visual check for Normality

- Discounts assesed 5%, 10%, 15%, 20%, 25%

- Should discounts be offered?

BUSINESS ANALYSIS:

- 15% and 25% had largest effect.
- 5% had lowest effect
- How to use this to your advantage?
  - Help push products
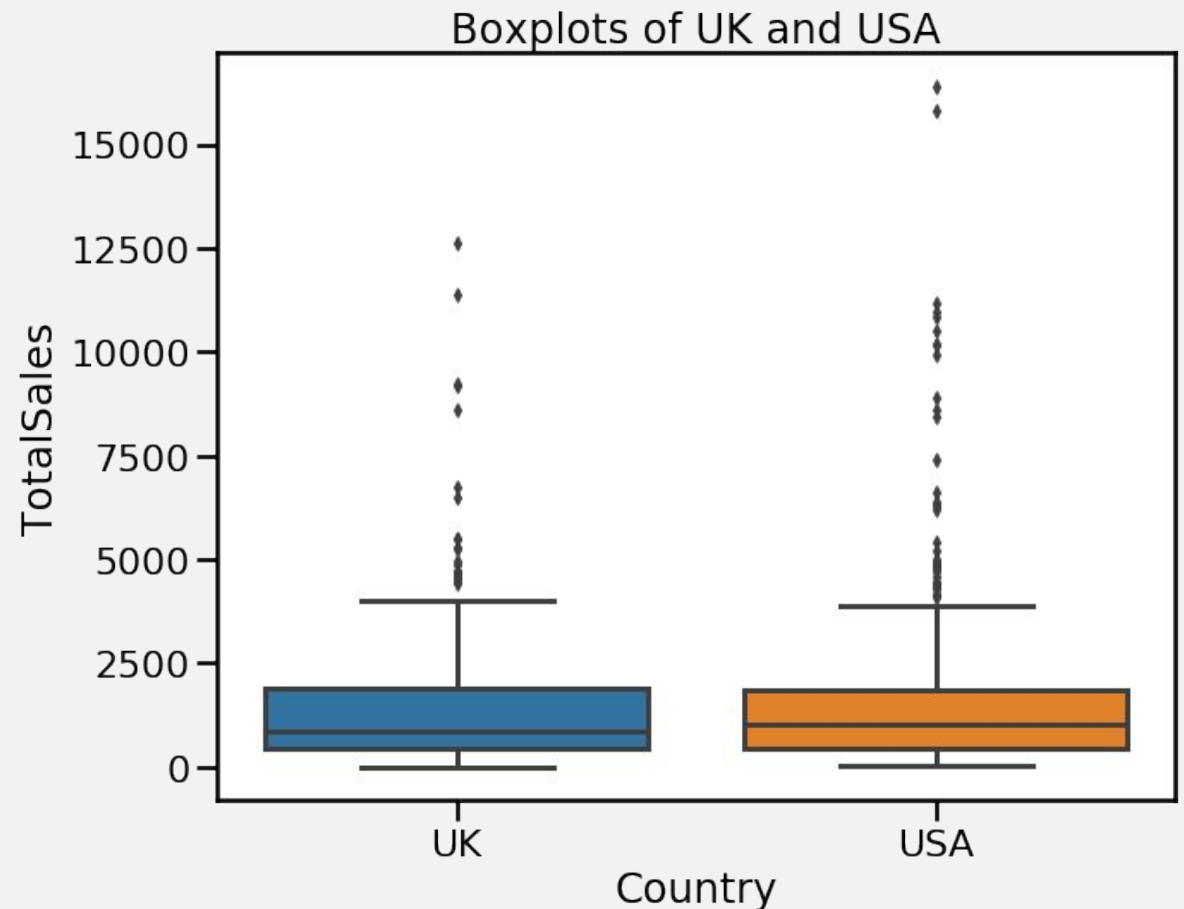 - Boost sales of old/new products
  - Get rid of inventory

Effect Sizes for each Discount Level

FURTHER ACTIONS FOR CASE 1:

- Find exact profit from each discount level

- Compare to no discount

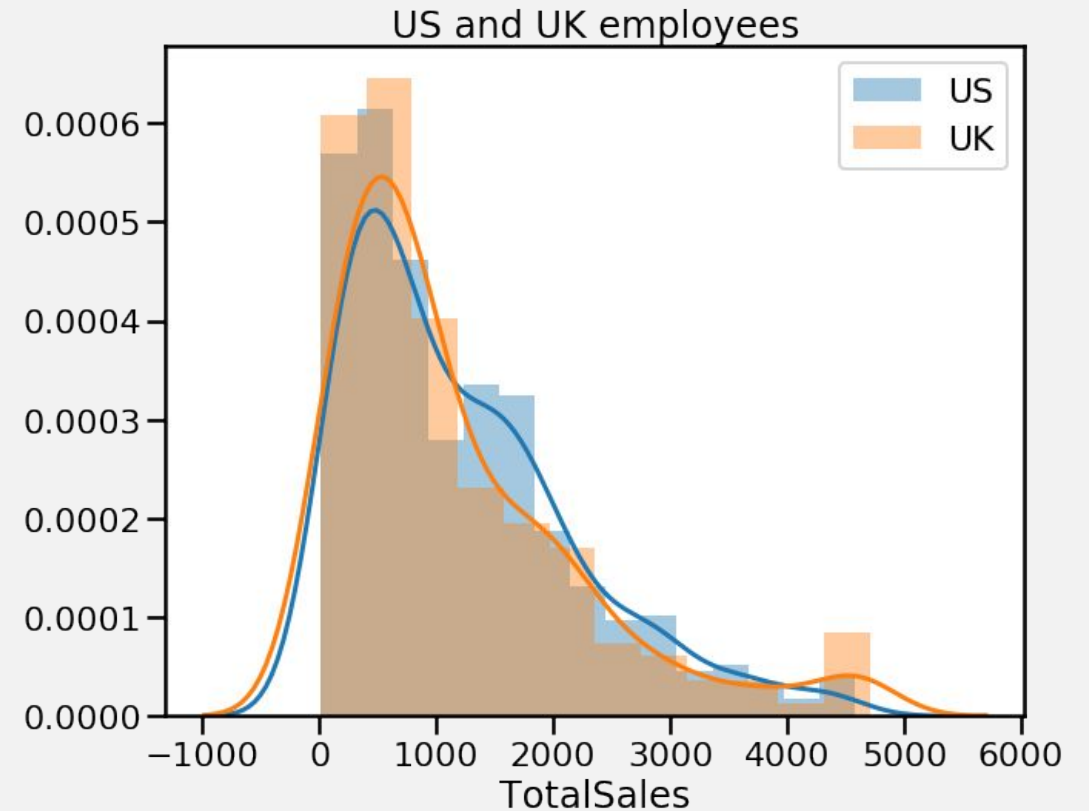- Maximize profit levels using discounts

CASE 2: DO US EMPLOYEES SELL SIGNIFICANTLY MORE PER ORDER THAN UK EMPLOYEES?

- Does one country perform better?

- Very similar average sale price

- Looking at plot:
  - A lot in common
  - Boxes are similar sizes

Boxplots of UK and USA

BUSINESS ANALYSIS:

- After running tests…
- no difference in average sales
- P-value = 0.9, no significance
- US has ~600 sales
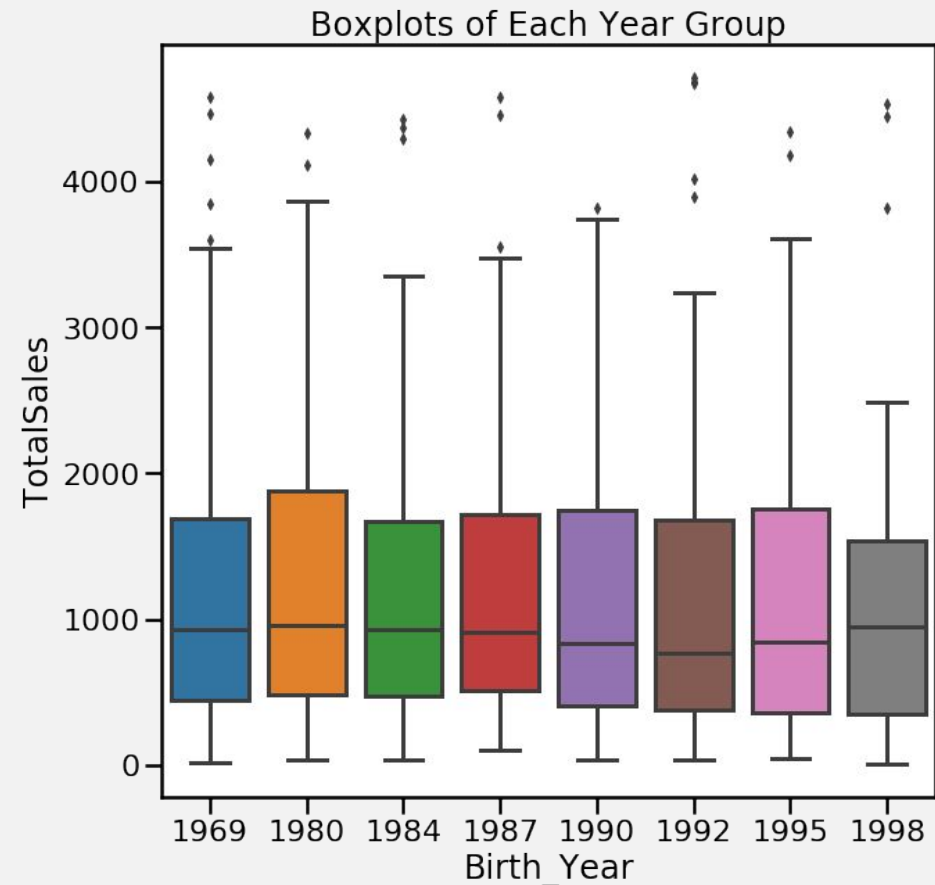- UK has ~200 sales

US and UK employees

FUTURE WORK

- Why does the UK sell less?

- Look to see if products differ?

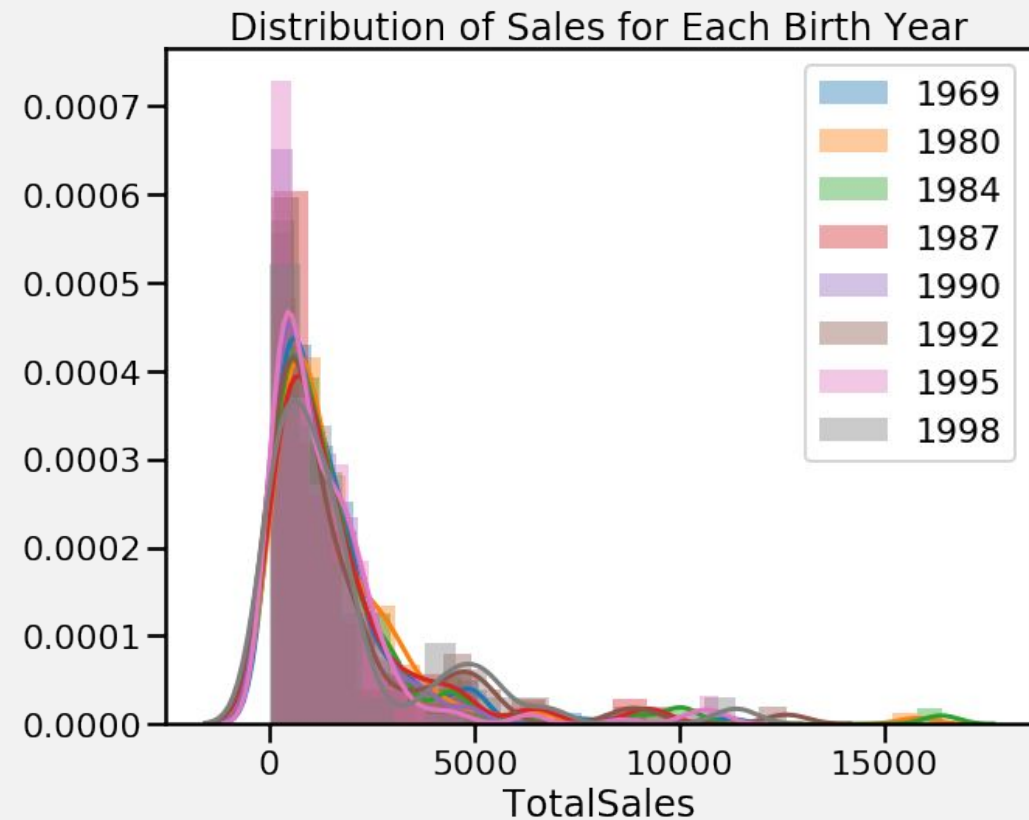- Look to see what is different from US an UK

CASE 3: DOES AN EMPLOYEE'S AGE HAVE A STATISTICALLY SIGNIFICANT EFFECT ON ORDER TOTALS?

- Have a similar average of sales

- No year significantly outperforms another visually

Boxplots of Each Year Group

BUSINESS ANALYSIS:

- After running tests…
- No significant differences in each year
- All employees sell about the same regardless of age

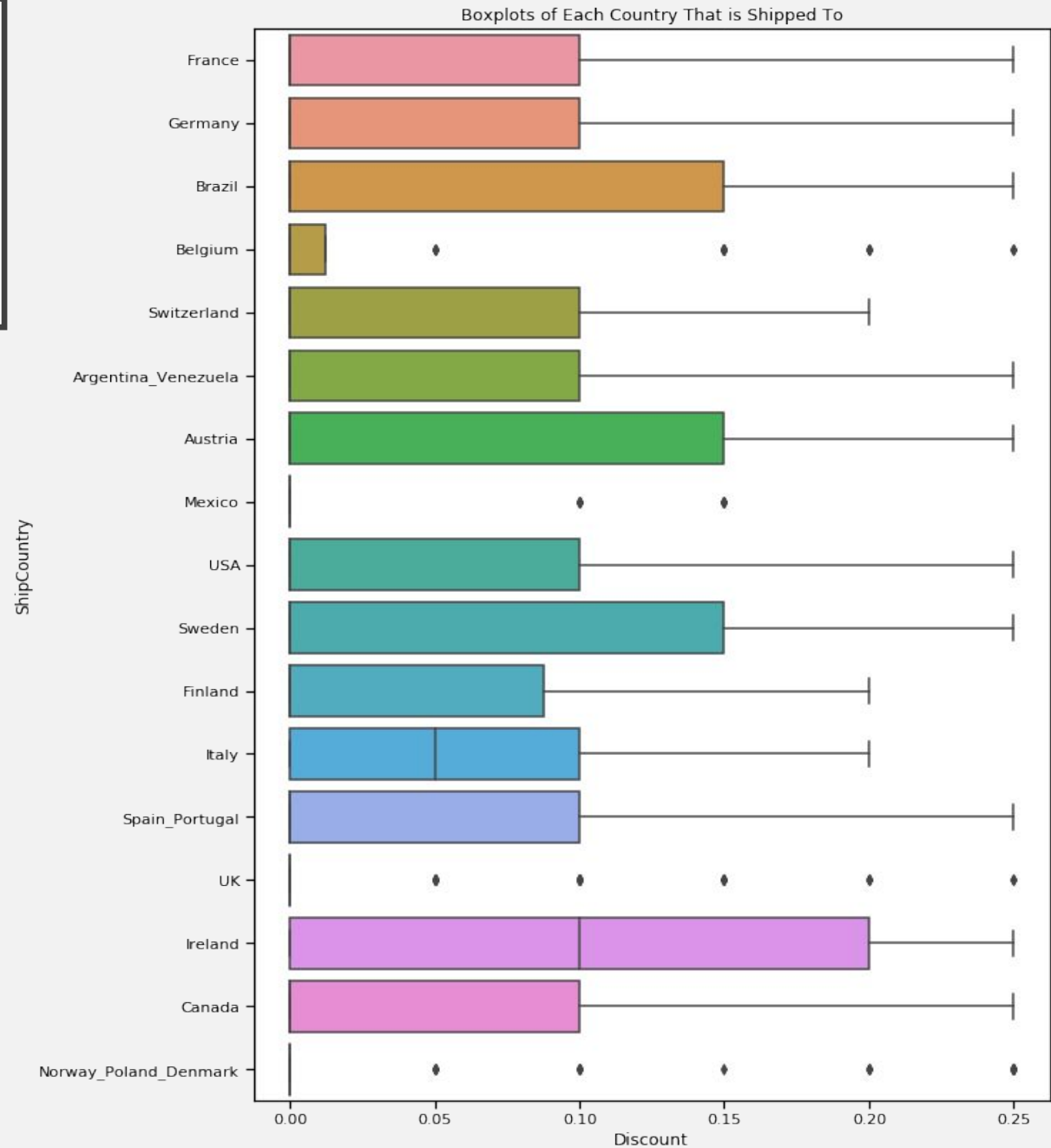Distribution of Sales for Each Birth Year

FUTURE WORK

- Combine case 2 and 3
  - What ages work in what country
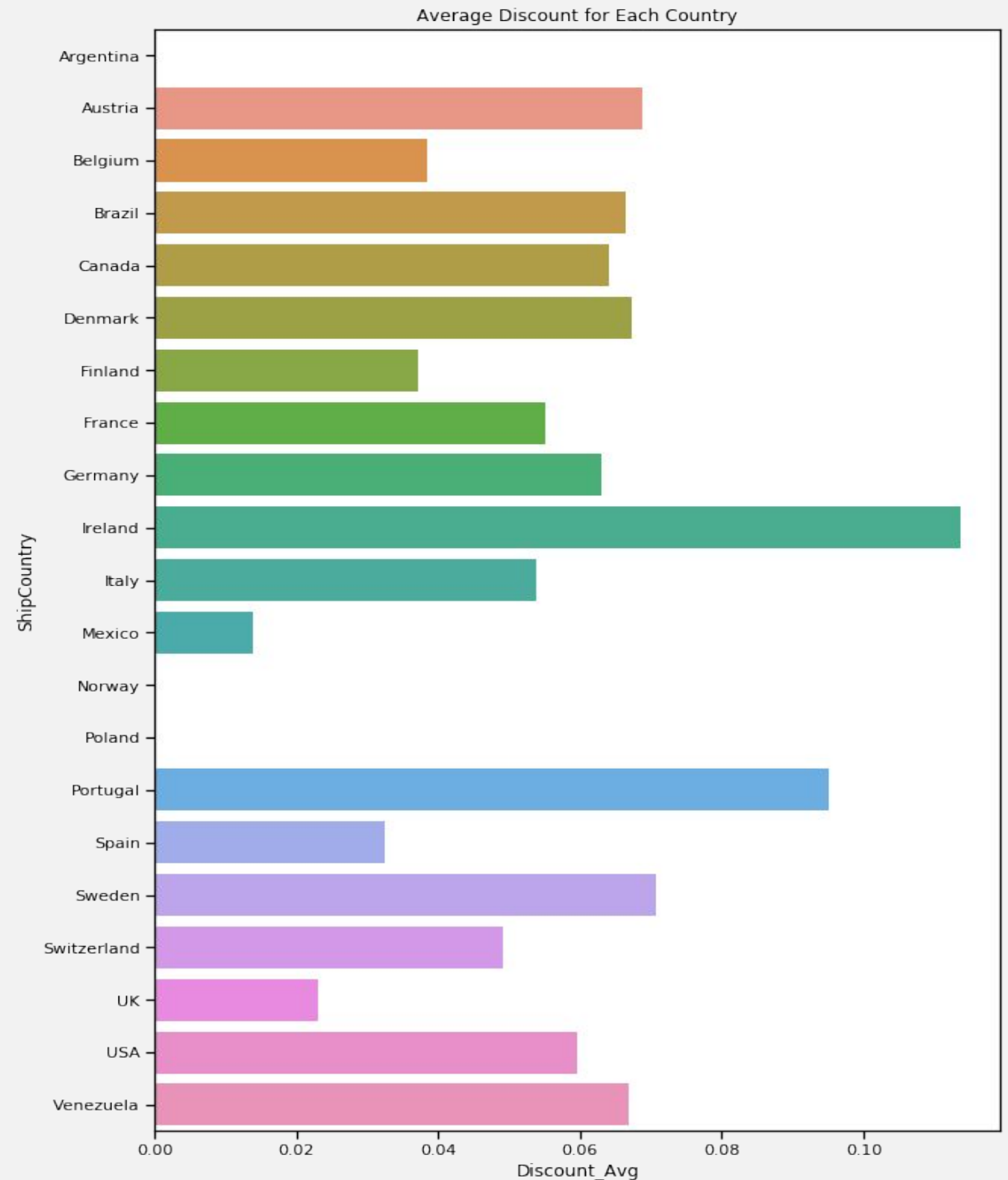 - Difference between ages and country

CASE 4: DO COUNTRIES SHIPPED TO HAVE A STATISTICALLY SIGNIFICANT EFFECT ON DISCOUNTS

- Boxplot of each country and discounts used
- Most have an average of 0
- Some have a lot of discounts

Boxplots of Each Country That is Shipped To

BUSINESS ANALYSIS

- Norway and Poland use no discounts
- Ireland uses the highest and then Portugal
- There is a correlation between ship country and discount
  - Just not sure for which

Average Discount for Each Country

FUTURE WORK

- Discover which countries have highest statistical significance
- Do higher discounts = higher discounts
  - Or vice-versa
- Do countries receive un-neccesarry discounts
  - Do profits suffer from these?

CONCLUSION AND THANK YOU

- This is only the beginning of testing

- Many more questions are raised from each case

- Thank You for your time

- Any questions or comments?