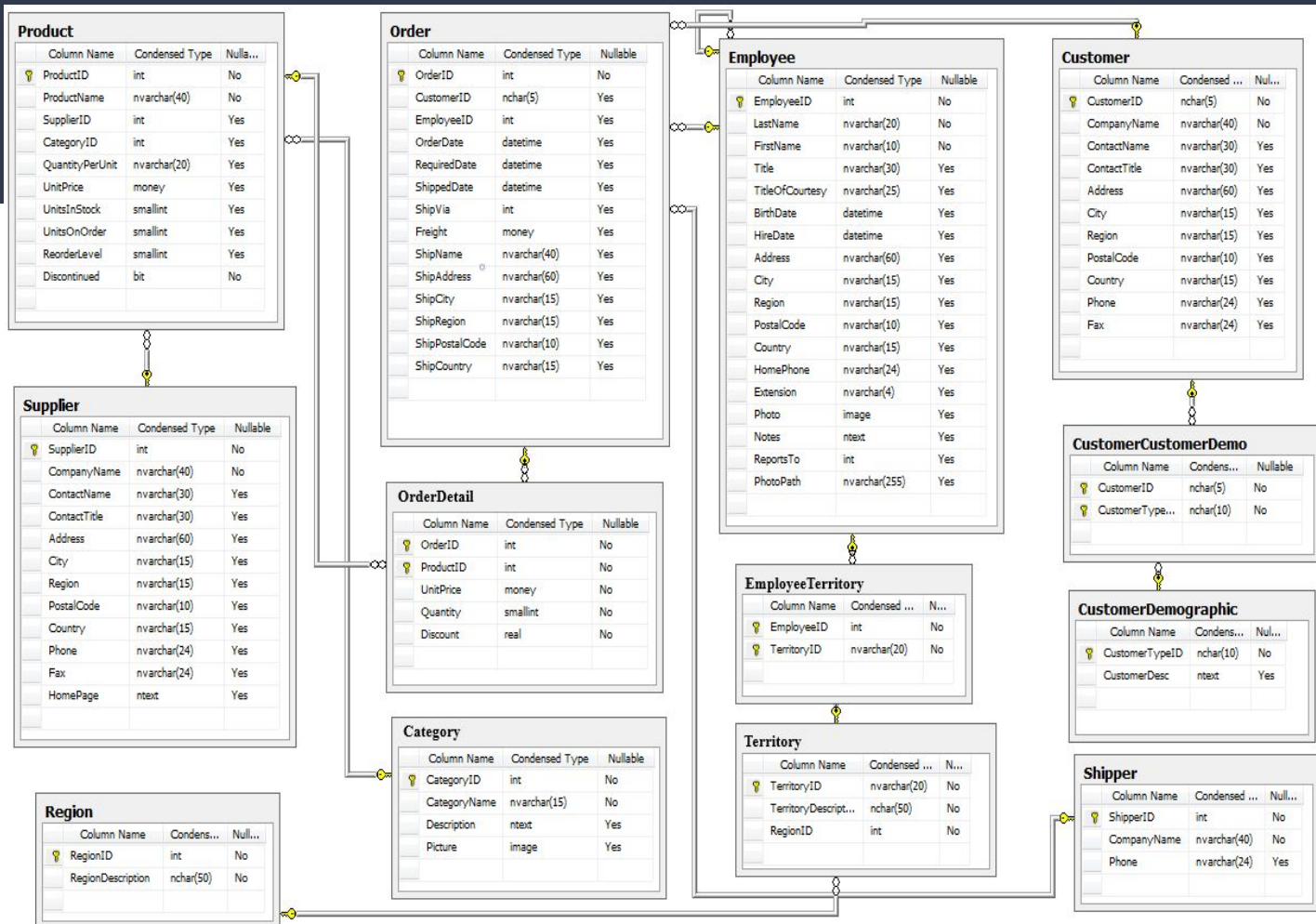# NorthWind Statistical Analysis

Susanna Han

STEPS

- Explored data

- Created 4 hypothesis'

- Connect tables for correlation

- Pulled information

- Ran tests through the grouped data.

- Reject/Accept hypothesis.

Tests applied on the data...

D'Agostino-Pearson's normality test - "is a goodness-of-fit measure of departure from **normality**, that is the test aims to establish whether or not the given sample comes from a **normally distributed population.**"

Levene's Test - "is a statistic used to assess the **equality of variances** for a variable calculated for two or more groups."

Kruskal-Wallis - "used to determine if there are statistically significant differences between two or more groups of an independent variable on a continuous or ordinal dependent variable."

Statistical Power - is the probability that the test correctly rejects the null hypothesis.

Tukey's Pairwise - "used to test differences among sample means for significance."
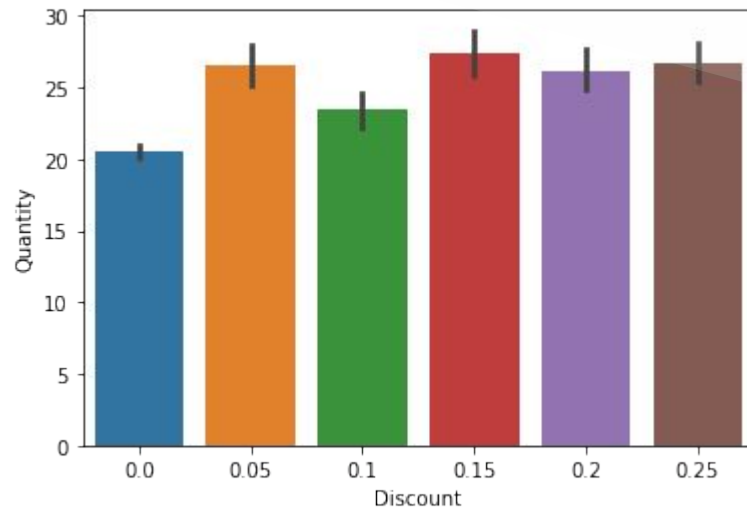
- Does *discount* have a statistically significant effect on the **number** of a product in an order?

- Does *discount* have a statistically significant effect on the **total amount** spent in an order?

- Does the *supplier region* have a statistically significant effect on the **product quantity sold** in an order?

- Does the *supplier region* have a statistically significant effect on the **total amount spent** in an order?

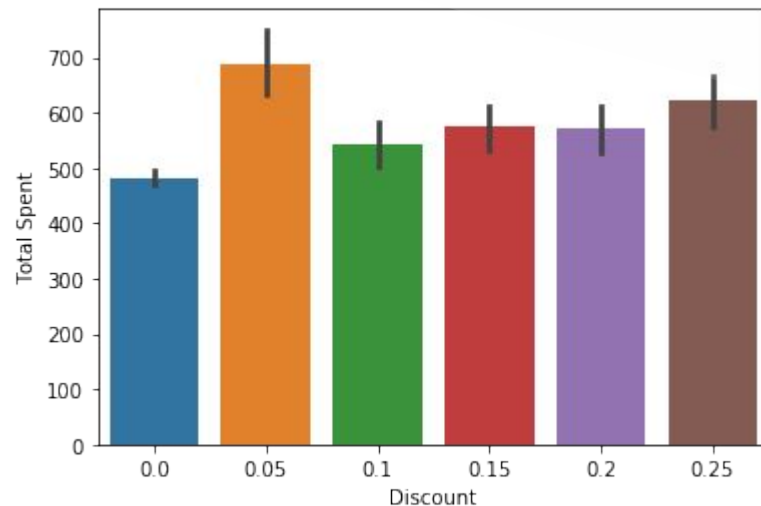- Does *discount* have a statistically significant effect on the quantity amount of a product sold in an order?

yes!

P-value < 0.05

- Does _discount_ have a statistically significant effect on the **total amount** spent in an order?
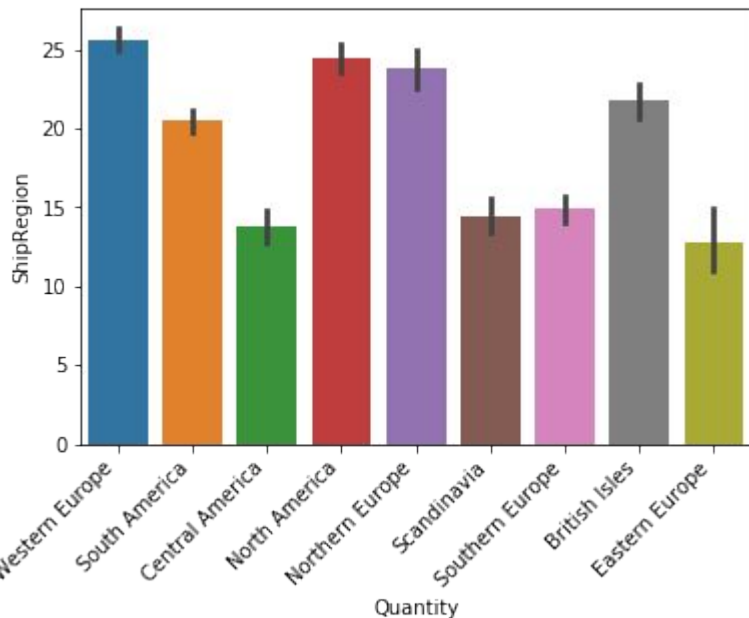
yes!

P-value < 0.05

- Does the *supplier region* have a statistically significant effect on the product quantity sold in an order?
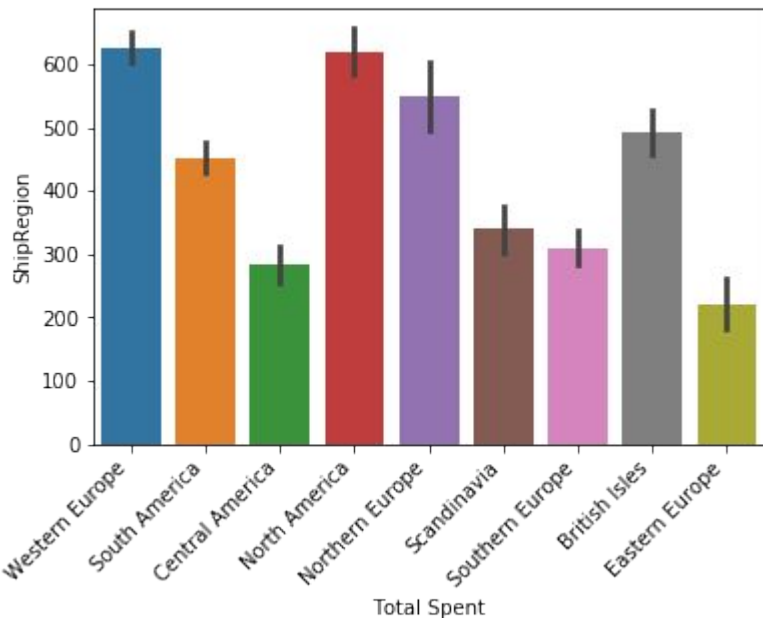
yes!

(Except Eastern Europe)

TOP 3 SUPPLIER REGIONS

1. Western Europe
2. North America
3. Northern Europe

- Does the *supplier region* have a statistically significant effect on the total amount spent in an order?

yes!

Except Eastern Europe

TOP 3 SUPPLIER REGIONS

1. Western Europe
2. North America
3. Northern Europe

Recommendations

- Discounting a product will **<u>increase</u>** the *quantity amount bought* and also increase the *total amount spent* in an order. BUT higher the discount **does NOT** mean an *increase* in money spent/quantity bought. A 5% discount is **NOT** significantly different than a 25% discount.
  - **Discount products at 5%** to *increase* sales.

- The TOP 3 supplier regions with the most sales in total amount spent and quantity of products sold in an order are in *Western/Northern Europe* and *North America*.
  - **<u>Increase</u> product suppliers in those regions.**

- The only supplier region that is **NOT** statistically significant in increasing the total amount spent and quantity of products sold is in *Eastern Europe*.
  - ***Consider discontinuing suppliers in that region.***

Thank You!

Susanna Han