INSIGHTS FROM DATA: HOW DO DISCOUNTS REALLY DRIVE SALES?

Gina Durante
Full-time student
04/15/19 Cohort

Full-Time Online
Data Science Program
Flatiron School

Agenda

- Project overview
- Questions addressed
- Methodology
- Findings / Discussion
- Future work

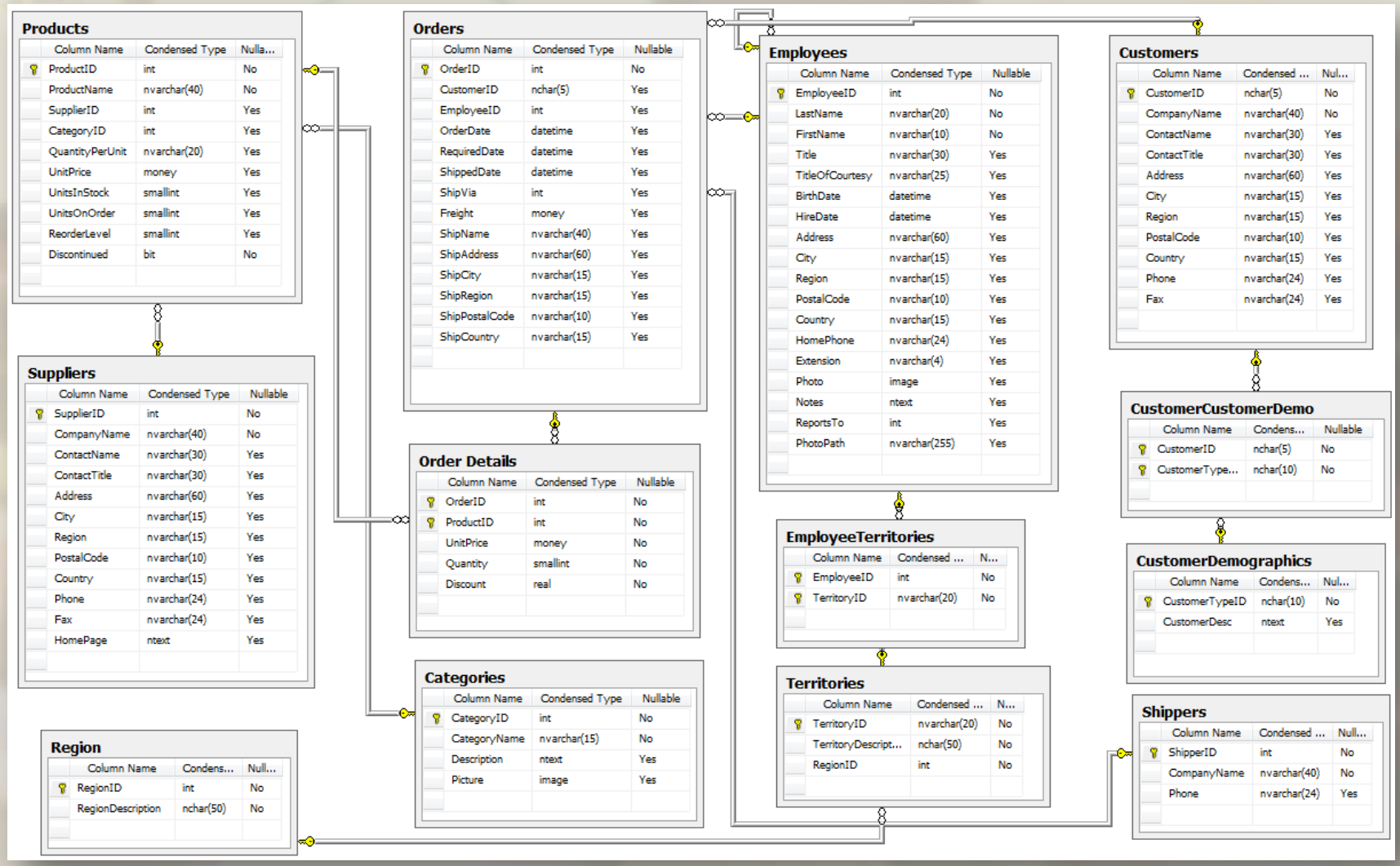
PROJECT OVERVIEW

Answering questions about product sales using hypothesis testing

Overview

- Purpose:
  - Deliver insights by answering business-relevant questions in an analytically rigorous way
  - Identify future questions to consider and analysis to pursue
- Approach: hypothesis testing
- Source of data:
  - Northwind's much-loved SQL database!

Data Source: Northwind SQL database



QUESTIONS CONSIDERED

Questions considered:

1. Do discounts generally result in a statistically significant effect on average order **quantity**?
2. Do discounts generally result in a statistically significant effect on average order **revenue**?
3. Do certain discount amounts (e.g., 5%, 10%, 15%, …) affect average order **quantity** more than others?
4. Do certain discount amounts (e.g., 5%, 10%, 15%, …) affect average order **revenue** more than others?
5. Do **revenue distributions** differ across product categories?

METHODOLOGY

Hypothesis Testing

Hypothesis Testing

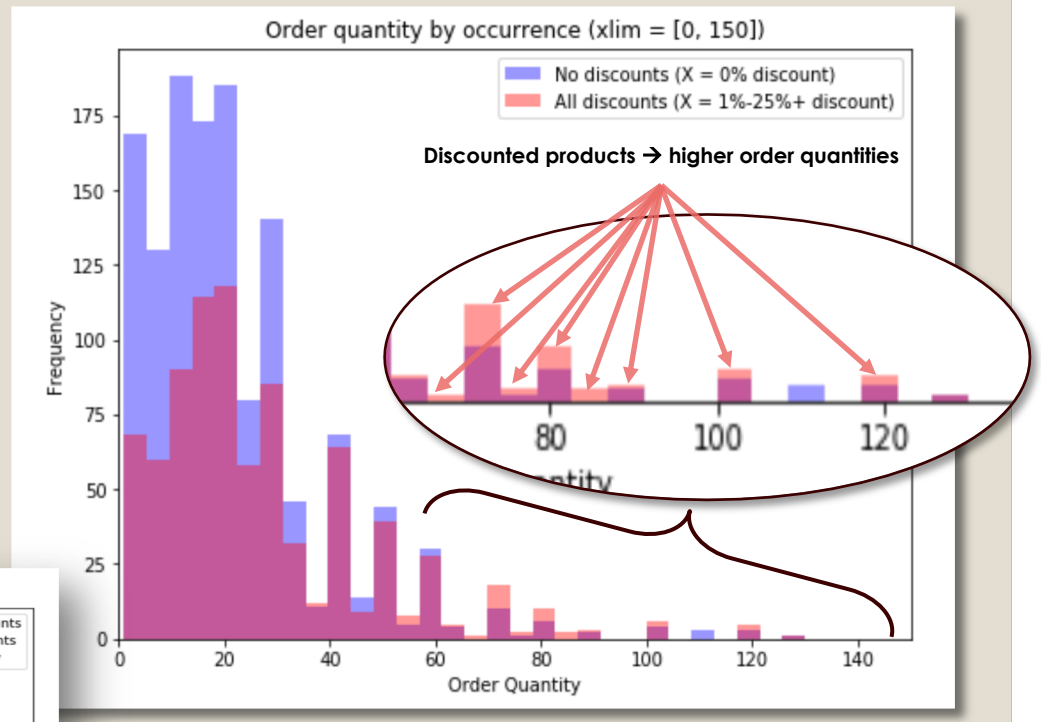
- What is it
  - Conceptually simple approach (differences in results ARE or ARE NOT due to chance)
  - Introduces rigor through powerful statistical methods
- Why it's great
  - Limits conscious / unconscious bias when applied correctly
  - Naturally invites additional questions and areas of exploration

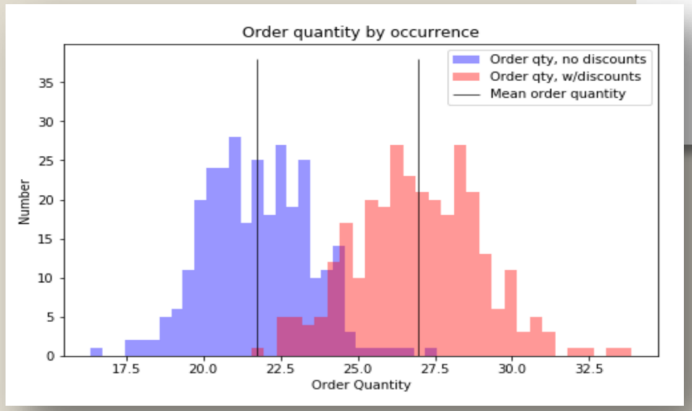
FINDINGS

The results of hypothesis testing on four discount questions, plus an additional question regarding revenues across categories, are coming up next!

Question 1: Do discounts increase **quantity** of products ordered?

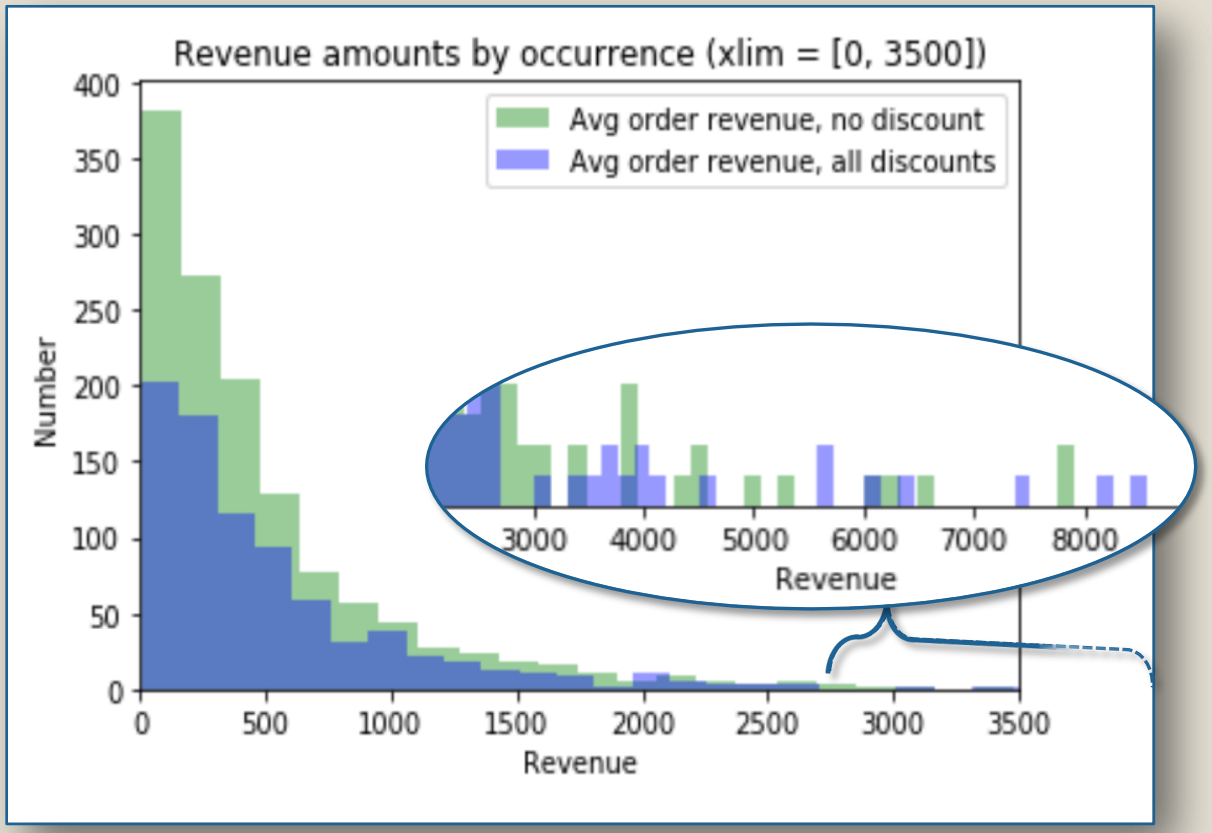
- Average order quantity: discounts vs. no discounts
  - Statistically significant ***increase in order quantity*** for discounted vs. non-discounted products
    - **Accept alternative hypothesis (H_a)!**
  - Avg order qty, no discounts = 21.7
  - Avg order qty, all discounts = 27.0
- Next slide: how does this translate into revenues…





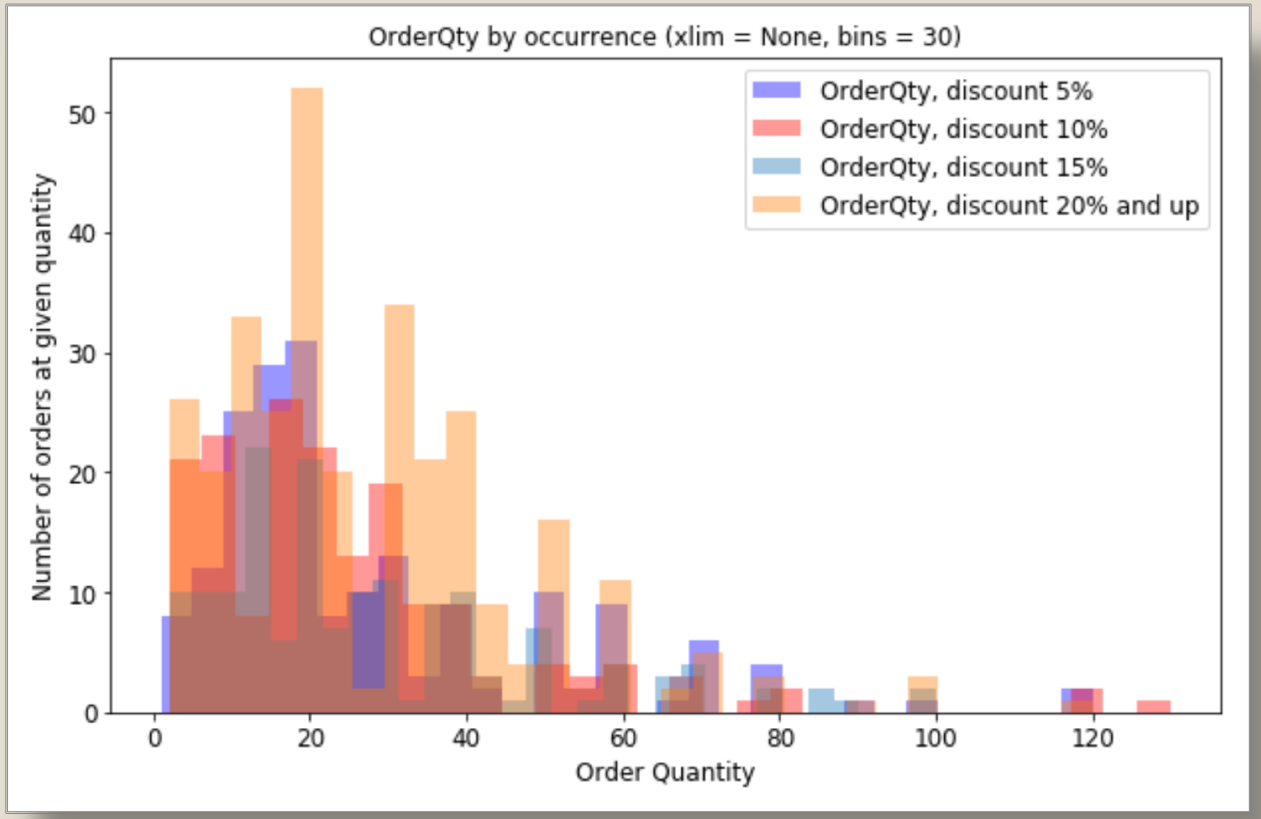
Question 2: Do discounts increase average **revenues** per order?

- Average order revenue: discounts vs. no discounts
  - Yes! → Statistically significant difference in ***revenues per order*** vs. non-discounted products
    - ***Accept alternative hypothesis (H_a)***
- Next: how does discount level affect quantities and revenues?



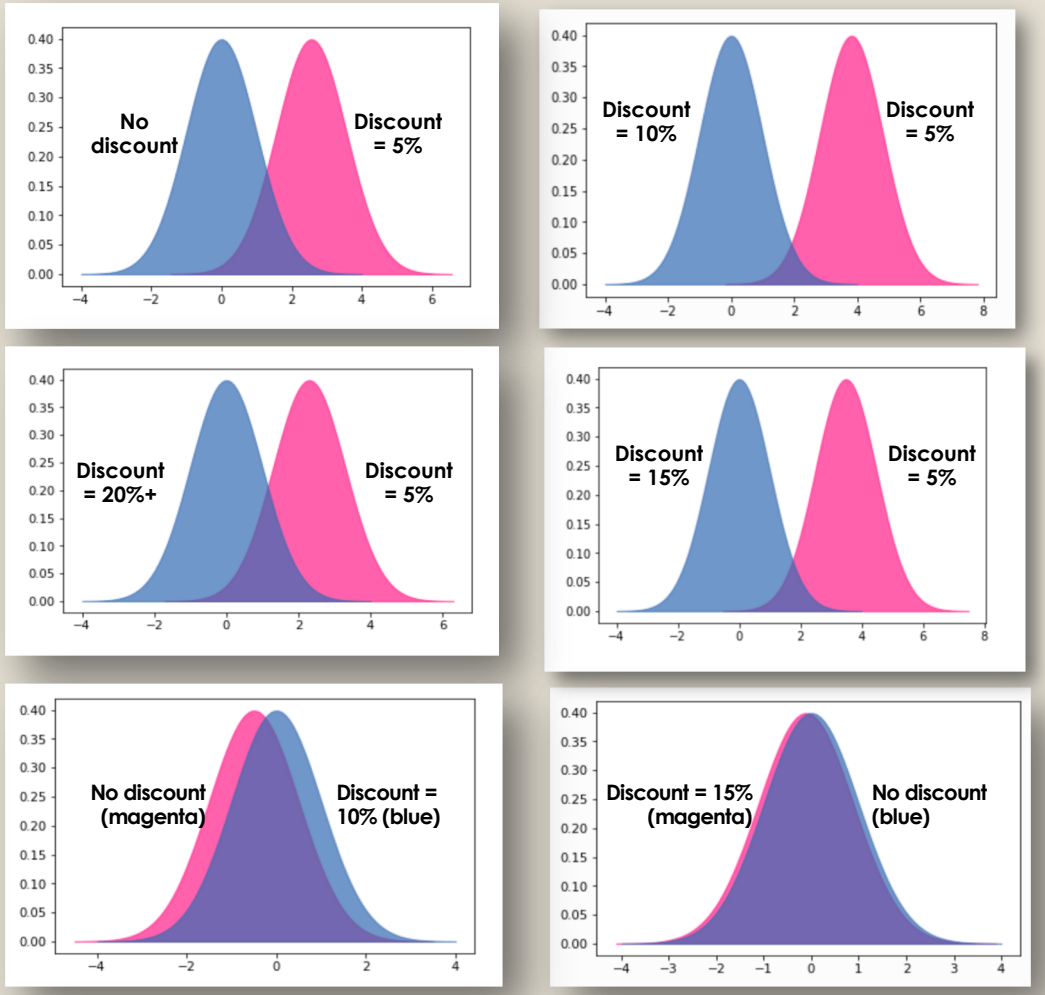
Question 3: Do *some* discounts affect order **quantity** *more* than others?

- Yes! → Statistically significant difference in average order quantity across all discount levels *except* 15-19.9%
  - ***Accept H_a for all (except 15-19.9%)***
- Interpretation of findings:
  - Generally, order quantities go *up* with discount… however…
  - Discounts not applied evenly to products across all categories
  - ***Most utilized discount level is 5%***
    - Could higher discounts drive orders?
- Next: We need to look at how average order *revenues* fare across different discount levels…



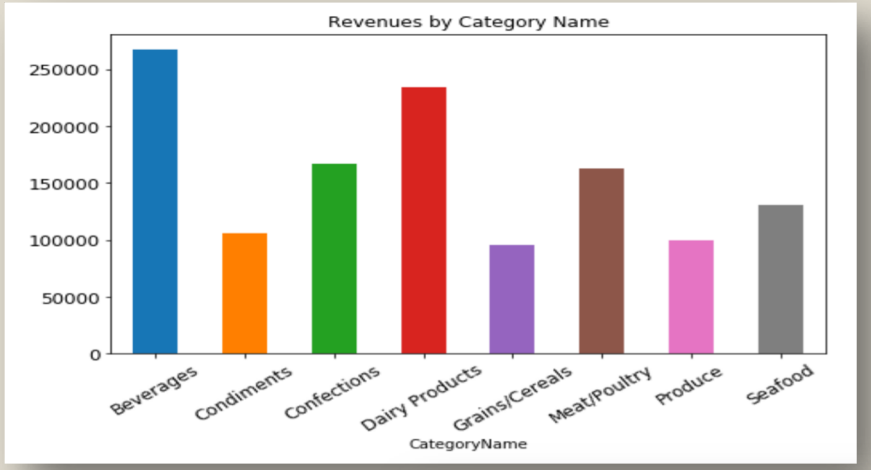
Question 4: Do *some* discounts affect order **revenues** *more* than others?

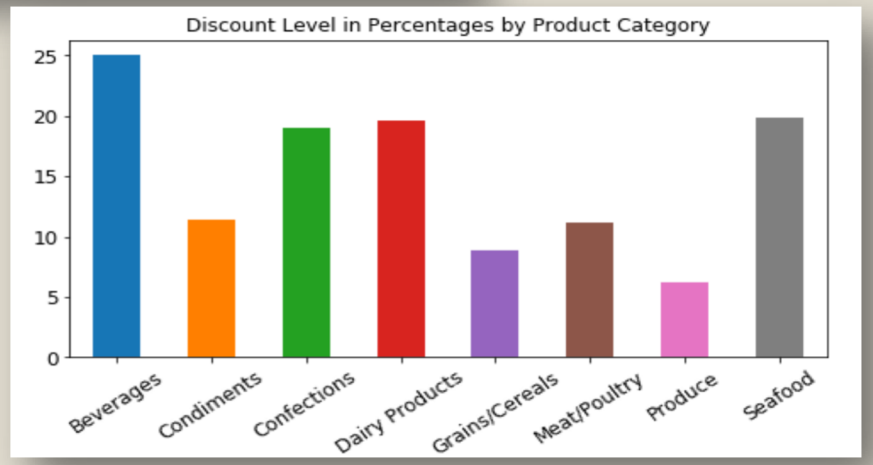
- Yes! → Statistically significant difference in average order revenues across all discounts *except* 15-19.9%
  - ***Accept H_a for all except 15-19.9%*** (same as for order qty in Question 3!)
- Interpretation of findings:
  - **Highest average order revenues: products discounted 5%**
  - Discounts not applied evenly across all categories *(see bar chart at right)*
- Next: Distribution of order revenue across product category…



Question 5: Do product categories differ in **revenue distributions**?

- Hypotheses tested:
  - Null Hypothesis (H_0): revenue distributions ***not*** statistically significantly different
  - * Alternative Hypothesis (H_a): at least one product category will have a significantly different distribution of order revenues
- Findings:
  - Most categories have revenue distributions that are significantly different from each other (**accept H_a**; reject H_0)
  - Ex.: Beverages—high number of small revenue orders, but also has some very large outliers (e.g., $8000, Côte de Blaye)





FUTURE WORK

1. ***Discount level by product category:*** Does the discount level vary significantly by product category?
2. ***Discounts within categories?*** Do revenues from discounts vary in a statistically significant way from revenues on non-discounted products in each product category?
3. ***Discounts on specialty products:*** How are discounts applied on speciality / high-end products (e.g., Cote de Blaye)?
4. ***Sales by geography:*** What are the total revenues by geography and the average revenues per order? Are there opportunities to grow number of orders and / or order revenues in new and existing markets?
5. ***Time of year:*** How do seasonality and business cycles affect sales? Do time of year effects vary significantly by geography?

FUTURE WORK

Answers to the questions at the left may provide additional insights into strategies for increasing revenues in existing and new markets!

Thank you for your time!!!

ADDITIONAL MATERIAL

(would have more material here for a presentation to stakeholders in an organization)

Question 5: Do product categories differ in **revenue distributions**?

- Hypotheses tested:
  - Null Hypothesis (H_0): revenue distributions ***not*** statistically significantly different
  - * Alternative Hypothesis (H_a): at least one product category will have a significantly different distribution of order revenues
- Findings:
  - Most categories have revenue distributions that are significantly different from each other (**accept H_a**; reject H_0)
  - Ex.: Beverages—high number of small revenue orders, but also has some very large outliers (e.g., $8000, Côte de Blaye)