BUAN 3500: Data Visualization and Descriptive Analytics

Introduction

Lauren M. Nelsen, Ph.D.

University of Colorado Colorado Springs

Data Visualization Introduction Activity

Example

Go to the following link, and use the data in the document to create a visualization.: https://bit.ly/47BYPhc to an external site.

Rules:

- You may not use ALL the data. Choose some that you want to highlight.
- You can create your visualization using anything you want. (Draw on paper, use Excel, Desmos, etc.)
- I want you to think about how the following two questions:
  - What do you notice about the data?
  - What do you wonder about the data?

(Here is a link to the checklist referenced in the document that you can look at if you'd like, but you do not need to: https://bit.ly/45cFilX to an external site.)

Business Analytics

> "From the dawn of civilization to 2003, five exabytes of data were created. The same amount was created in the last two days."
>
> –Google CEO Eric Schmidt speaking in the keynote presentation at the Guardian's Activate summit in 2010

Business Analytics: transforming data into insight for making better decisions

- Descriptive analytics (BUAN 3500)
- Predictive analytics (BUAN 3600)
- Prescriptive analytics (BUAN 4000)[1]

[1] All three of these classes use the same textbook.

Descriptive analytics (BUAN 3500):

$\rightarrow$ Techniques to describe what has happened in the past

Predictive analytics (BUAN 3600):

$\rightarrow$ Techniques to predict from past data what will happen in the future

Prescriptive analytics (BUAN 4000):

$\rightarrow$ Using analysis techniques to suggest a certain course of action

Often consists of a prediction plus a rule

- Predict the probability of a loan default
- Rule: If the probability is $> 60\%$, don't approve the loan

Reasons data analytics is important in business:[2]

- Gain greater insight into target markets
- Enhance decision-making capabilities
- Create targeted strategies and marketing campaigns
- Improve operational inefficiencies and minimize risk
- Identify new product and service opportunities

[2] https://lpsonline.sas.upenn.edu/features/5-key-reasons-why-data-analytics-important-business

Big Data

Definition

Big data is "any set of data that is too large or too complex to be handled by standard data processing techniques and typical desktop software".

Example: Kroger (King Soopers) Kroger (King Soopers) sends 11 million mailings each quarter with 12 coupons in each mailing. The analysis and tracking of who uses which coupons means dealing with Big Data!

Business Analytics

Business analytics applies across all discliplines!

Examples:

- Finance
- HR
- Marketing
- Health care
- Supply chain
- Government
- Nonprofits
- Sports
- Internet activity
- your life!

Legal and Ethical Issues in the Use of Data and Analytics:

- Use and misuse of customer data
- Misrepresentation of data
- Biased algorithms

Our Class

In this class we will focus on

- **describing data** and
- **visualizing data**.

> "Thinking carefully about *how* data is presented is just as important as the data itself."
> -Jonathan Schwabish in "Better Data Visualizations"

Main topics we plan on covering in this course

- **Descriptive Statistics**

$$P(z \leq 1.76) - P(z < -0.86)$$

$$= 0.9608 - 0.1949$$

$$= \boxed{0.7659}$$

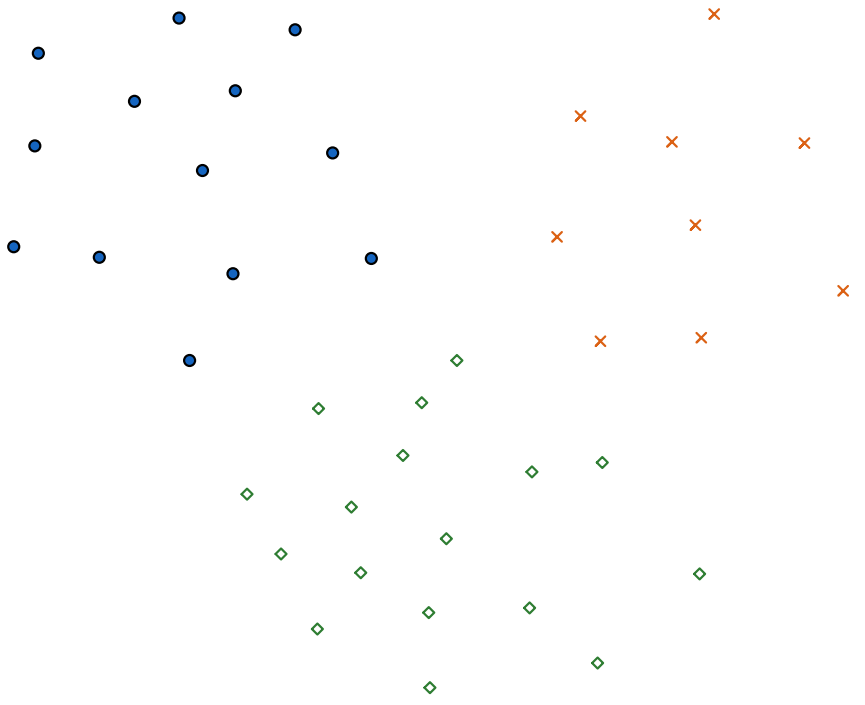
Main topics we plan on covering (continued)

- **Data Visualization**
  - Design techniques and principles
  - Excel charts
  - Creating visualizations using Python
  - **Using Tableau to create data visualizations**
  - Data Dashboards
    - Principles of data dashboard design
    - Using Tableau to create dashboards

    Examples of Tableau dashboards:
    https://public.tableau.com/app/discover/viz-of-the-day

Main topics we plan on covering (continued)

- **Probability and using it to analyze data**
- **Descriptive Data Mining**
  - **Cluster Analysis:**



    - Example of use: market segmentation

Main topics we plan on covering (continued)

- **Descriptive Data Mining**
  - **Association Rules**
    - Examples of use: analyzing transaction data ("market basket analysis"), medical treatments

  - **Text Mining:** extracting useful information from text data
    - Examples of use: analyzing social media data, seeing if reviews are positive or negative
    - Ted Lasso:

`https://public.tableau.com/app/profile/anne.sophie5083/viz/TedLassoPopCultureReferences/TedLassoReferences`

Main topics we plan on covering (continued)

Machine Learning:

- The data mining methods we will discuss are examples of **unsupervised learning** techniques.
- If we have time, we might talk briefly about **supervised machine learning** and how that is different.

- **Statistical Inference**
  - Point Estimation
  - Sampling Distributions
  - Interval Estimation
  - Hypothesis Tests
  - Statistical Inference, and Practical Significance

We want to learn how to use these techniques to summarize data and use it to make well-informed decisions.

Exercise

Example

Open up the file "Dow Jones Industrial Average Historical Data" (in Canvas).

- Create a graphic depicting the price over time so the viewer can gain some insight into the relationship.

If you finish this, here are some other things you can think about:

- Is there a way you could modify your graph that might help the viewer more easily see what is happening?
- Think about if there is a different graphic you could create using this dataset that would provide useful insight.
- Are there any questions that you have about the data? (Are there any calculations you could do that would give you a better understanding of the data?)