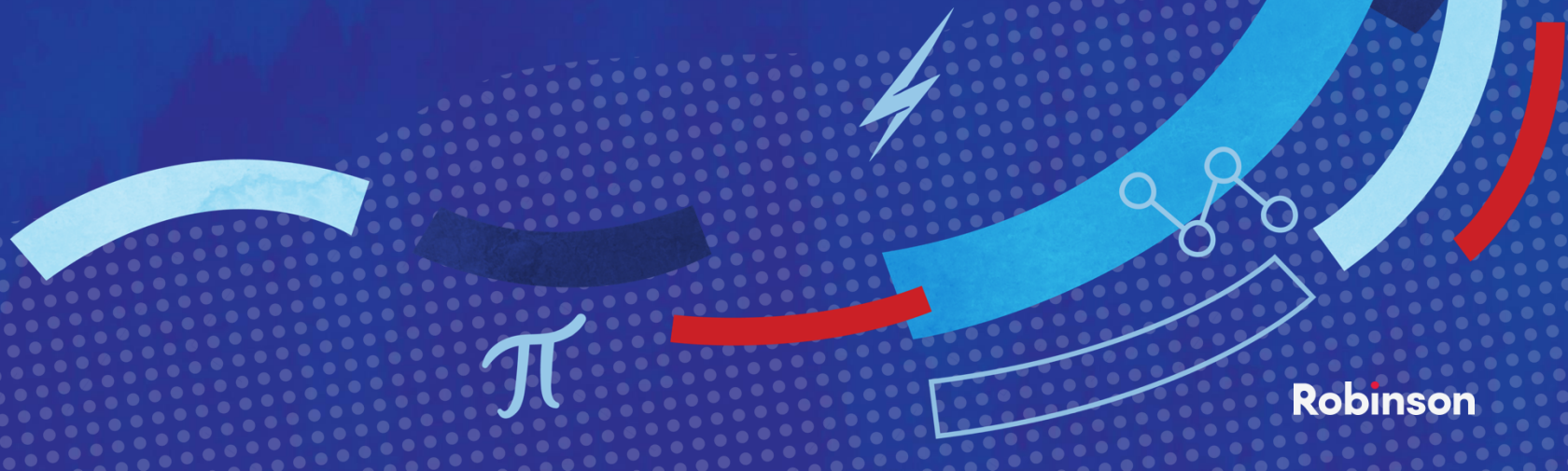
We’re starting soon

Please submit your responses for the Attendance Quiz 2. You will have 15 min to submit your answers once you start. Thanks!

MSA 8030 – Communicating with Data

Mark Jack mark.a.jack@gmail.com

October 21, 2024



Georgia State University | J. Mack Robinson College of Business

Robinson

Week 2 –
Team Presentations: Present Your Data Dictionary
Class Discussions: Define Your Business Problem

Mark Jack mark.a.jack@gmail.com

October 21, 2024

Georgia State University | J. Mack Robinson College of Business

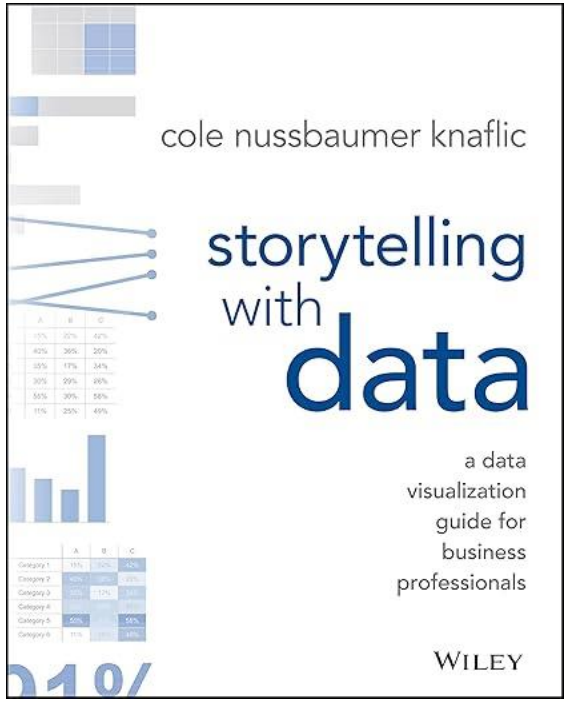
Robinson

Course Resources

Recommended textbook:

***story telling with data - a data visualization guide for business professionals*, 1st edition**

by Cole Nussbaumer Knaflic (Wiley)



Github site:

CommunicatingWithData

https://github.com/mjgrav2001/CommunicatingWithData

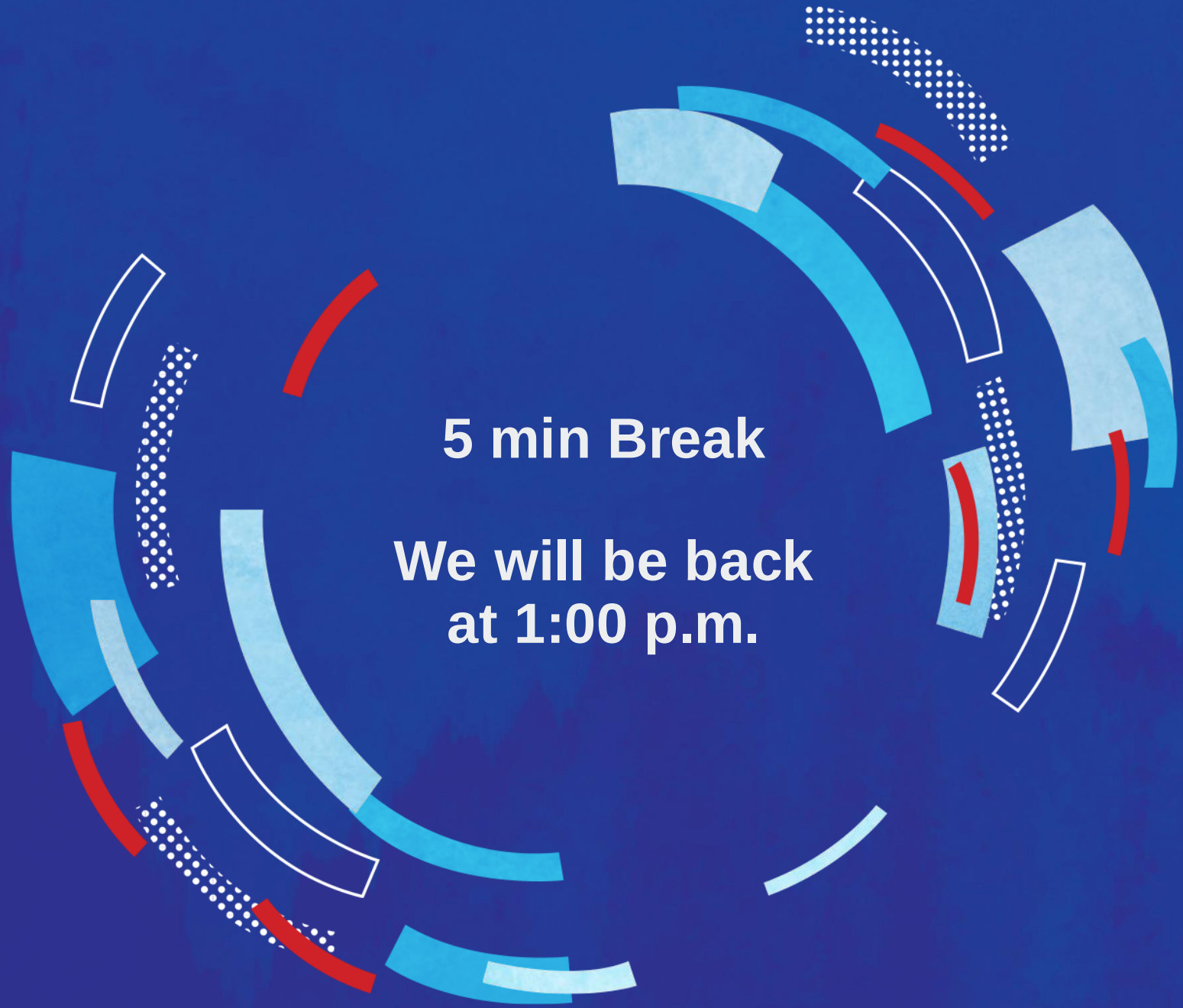
It is 12:10 p.m.

Let’s start
the Team Presentations
…

Team Presentations (45 min)

Team 1 – Wine Quality dataset
Team 2 – Airline Delay and Cancellation dataset
Team 3 – AirBnB (Kaggle) dataset
Team 4 – Wine Quality dataset
Team 5 – Microsoft Malware Prediction dataset
Team 6 – Water Potability of Plants dataset
Team 7 – LAPD Crime dataset

(presentations in no particular order!)



5 min Break

We will be back at 1:00 p.m.

Lecture 1 (30 min): How to extract the business problem

How To Ask The Right Questions As A Data Scientist To Define A Problem Statement:

- **Understand the problem that needs to be addressed and solved:** What is the opportunity that needs to be ascertained? What is the pain point that our stakeholders are facing?
  - We need to see the problem from the perspective of the stakeholders.
  - Our task is to learn the domain knowledge from them.
  - We need to combine our technical knowledge with data to come up with a solution to drive business values.
- **Assess the situation with respect to the problem:**
  - We need to analyze requirements, assumptions, constraints and resources of the situation.
- **Understand the potential risks and benefits of the project:**
  - What are the main costs associated with this project?
  - What are the potential benefits?
  - What risks are there in pursuing the project?
  - What are the contingencies to potential risks?
- **Define success criteria (= quantifiable metrics) to assess the project:**
  - Discuss with the stakeholders what metrics should be used to gauge and evaluate the success of the project.

Lecture 1 (30 min): How to extract the business problem

How To Draft A Problem Statement:

Be SMART!

Follow this SMART methodology checklist when drafting your problem statement:

- **S**pecific, not general
- **M**easurable
- **A**ction-oriented
- **R**elevant (to the key problem)
- **T**ime-bound

Lecture 1 (30 min): How to extract the business problem

Typical Business Problems in the Market Place:

- **Customer-centric problems:**
  - Knowing and understanding your customers in detail
  - Objectives:
    - increasing revenue by improving product recommendations
    - upselling
    - cross-selling
    - reducing churn and improving retention rates
    - personalizing the user experience
    - improving targeted marketing
    - sentiment analysis
    - product or service personalization
    - pricing optimization

Lecture 1 (30 min): How to extract the business problem

Typical Business Problems in the Market Place:

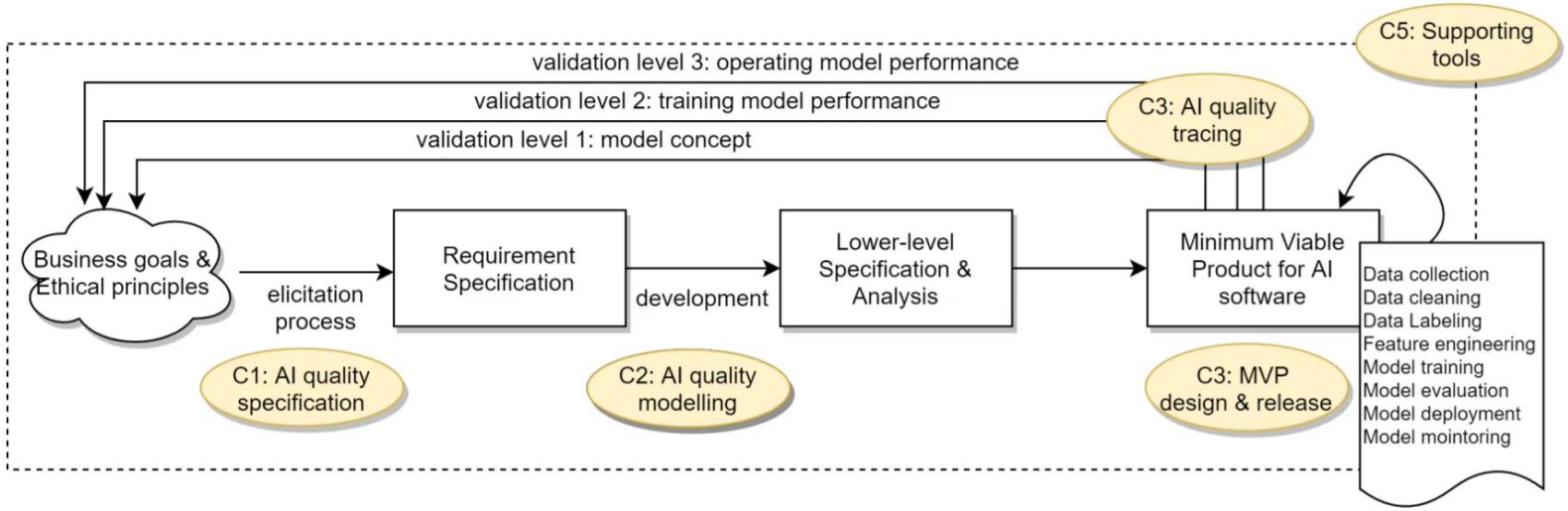
- **Optimization problems:**
  - Maximizing or minimizing factors such as costs, revenues, risks, time etc. within a well-defined quantitative framework and with a given set of constraints
  - Examples:
    - supply chain optimization, logistics and transportation (e.g. delivery routing)
    - finance (e.g. minimizing risk in an investment portfolio for a given target return)
    - scheduling (e.g. retailer optimizing staffing levels per store within shifts, airline optimizing its route network)
- **Demand prediction:**
  - Estimation of demand by product line or business unit based on historical aggregate demand
  - Estimation of demand on a per store, per hour or per customer basis
  - Estimation based on consumer data, macroeconomic data and other open data
- **Fraud detection:**
  - Rare-event detection problem (99.9% of banking transactions are not fraudulent)

Lecture 1 (30 min): How to extract the business problem

Lean Ideology of AI Project Management:

Prototyping and Getting Feedback:

- Create a continuous validation loop by iterating between customer needs, business needs, and the product.
- Perform continuous experimentation with a minimum viable product (MVP).



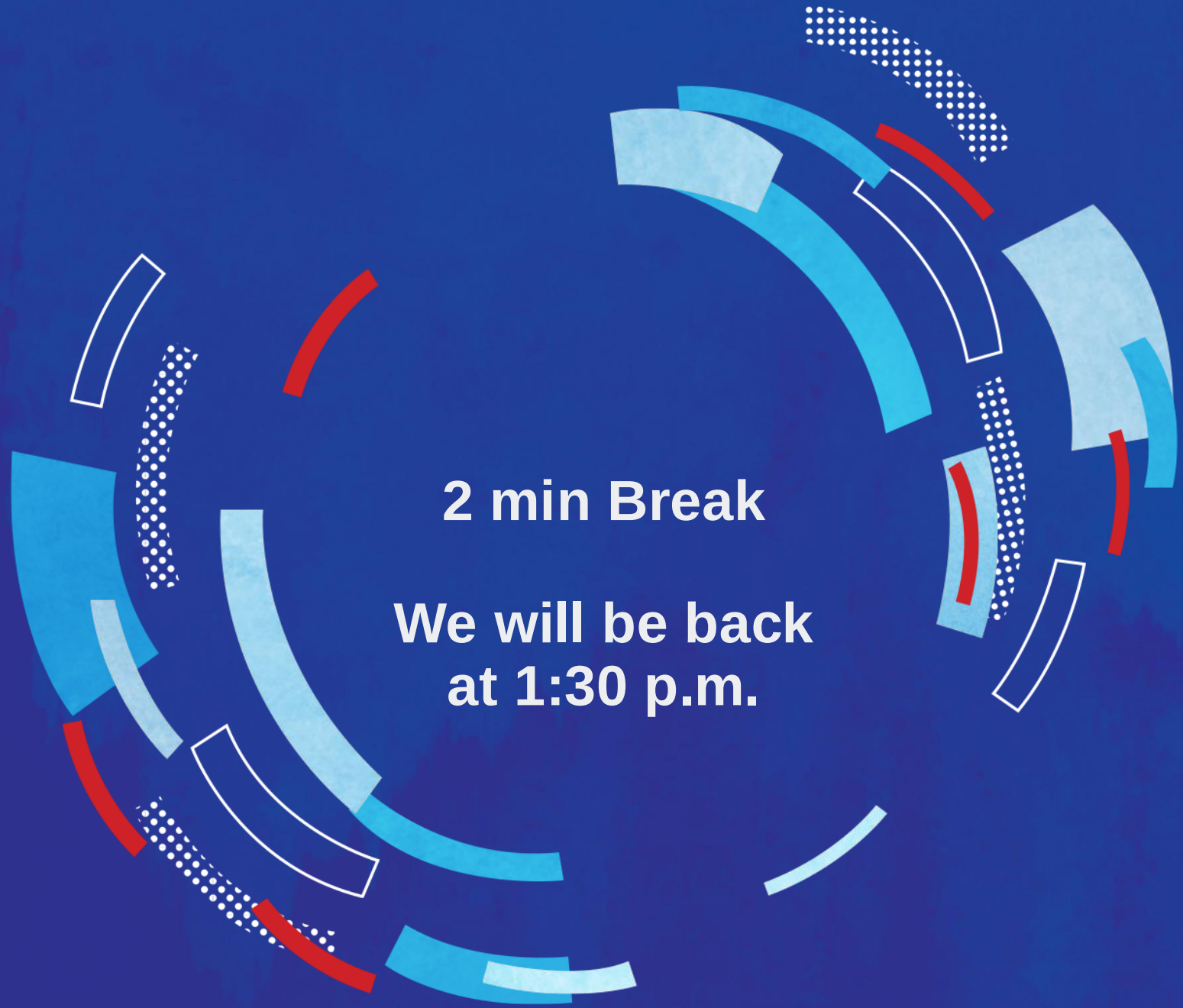
=> Minimum Viable Model, Minimum Viable Platform, Minimum Viable Data Product …

Lecture 1 (30 min): How to extract the business problem

Identify Artificial Intelligence Use Opportunity:

- **Annual Meeting With Business Leaders**
- **Spontaneous Request From Business Leaders**
- **Spontaneous Ideas From Data Scientists**
- **Discussion Between Companies**
- **Use Well Known Use Cases**
- **Try New Research Paper**

=> Proof-Of-Concepts (POCs): POCs are an opportunity to quickly test ideas by demonstrating their feasibility using simple and inexpensive techniques.



2 min Break

We will be back at 1:30 p.m.

Group Activity (30 min): Teams brainstorm on next week's assignment - Present your business problem!

Questions to brainstorm on as a team:

- What is your **business problem** / project idea?
- What is the larger business **context**?
- What is the **scope** at which you are proposing to solve the problem?
- What are the **requirements, assumptions, constraints, resources** of your situation?
- What are the main **costs, potential benefits, risks, contingencies?**
- What are your **metrics** of success for the project?
- Follow the **SMART** methodology to formulate your problem statement.
- Going further: Do you already have possible **use cases** in mind?

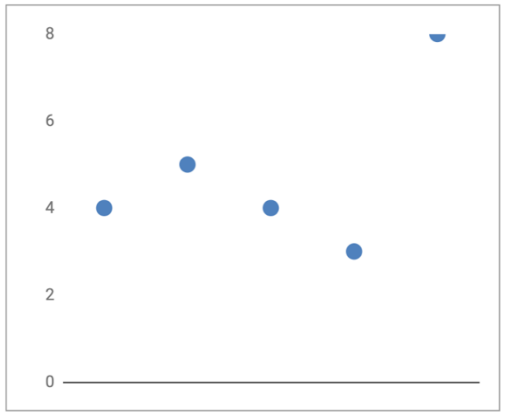
2 min Break

We will be back at 2 p.m.

Lecture 2 (30 min): Storytelling With Data
Topic: Choosing an effective visual (Chapter 2)

91
%

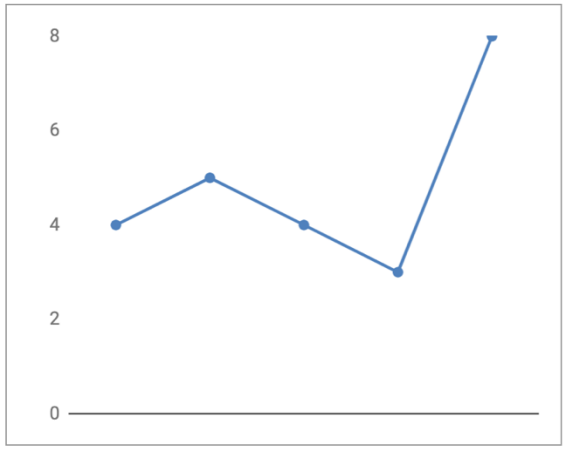
Simple text



Scatterplot

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

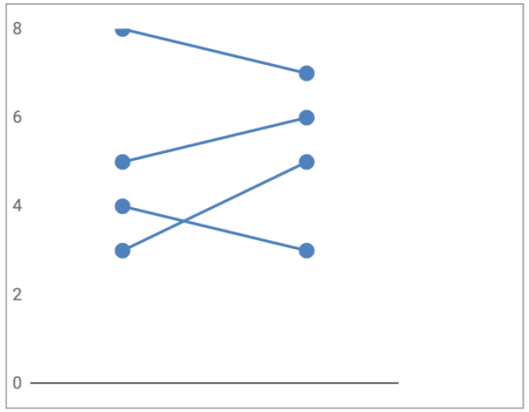
Table



Line

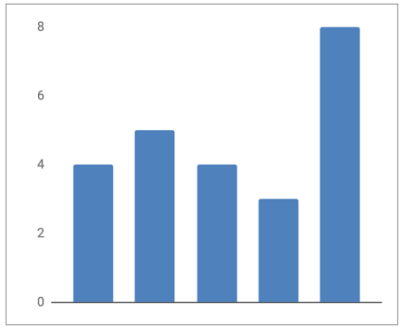
	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

Heatmap

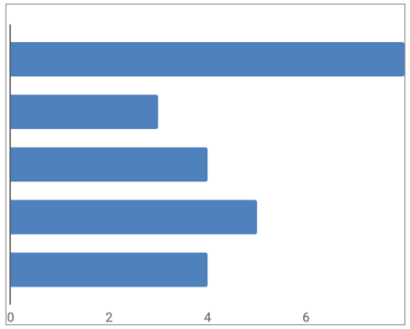


Slopegraph

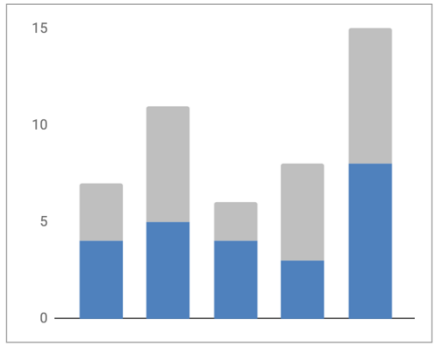
Lecture 2 (30 min): Storytelling With Data
Topic: Choosing an effective visual (Chapter 2)



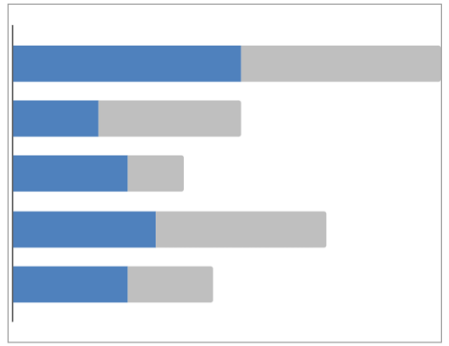
Vertical bar

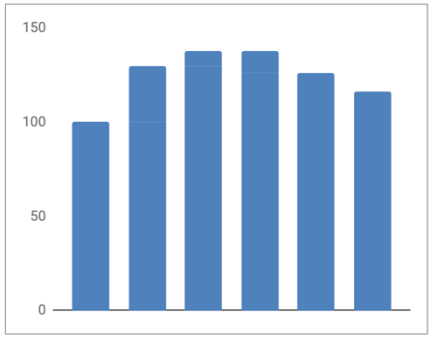


Horizontal bar

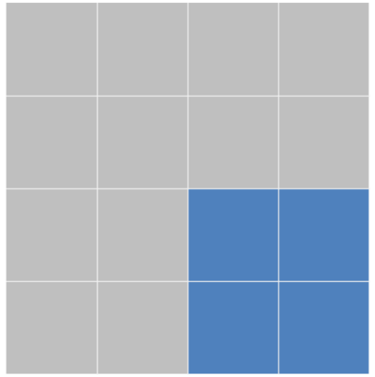


Stacked vertical bar


Stacked horizontal bar



Waterfall

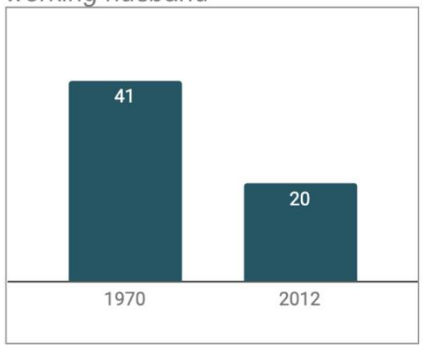

Square area

Lecture 2 (30 min): Storytelling With Data
Topic: Choosing an effective visual (Chapter 2)

FIG0202

Children with a "Traditional" Stay-at-Home Mother

% of children with a married stay-at-home mother with a working husband



Note: Based on children younger than 18. Their mothers are categorized based on employment status in 1970 and 2012.

Source: Pew Research Center analysis of March Current Population Surveys Integrated Public Use MIcrodata Series (IPUMS-CPS), 1971 and 2013

Adapted from PEW RESEARCH CENTER

FIG0203

20%

of children had a **traditional stay-at-home mom** in 2012, compared to 41% in 1970

Lecture 2 (30 min): Storytelling With Data
Topic: Choosing an effective visual (Chapter 2)

FIG0205

Table

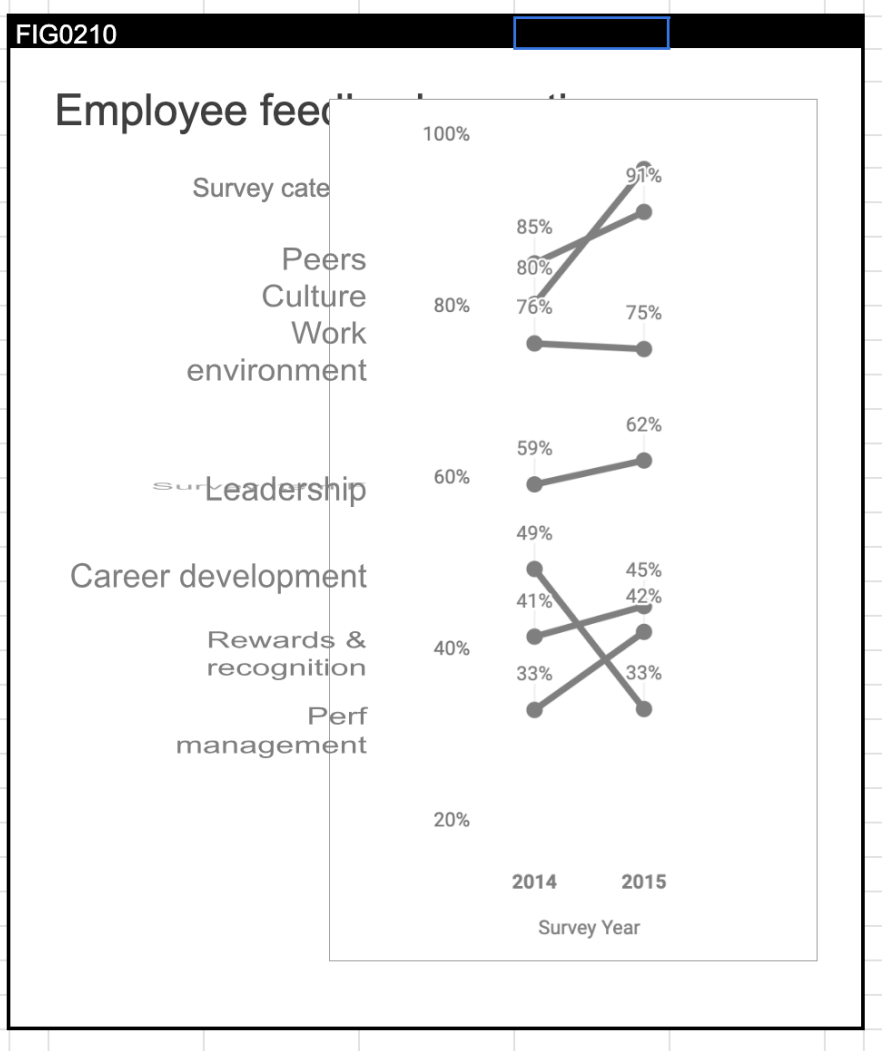
	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

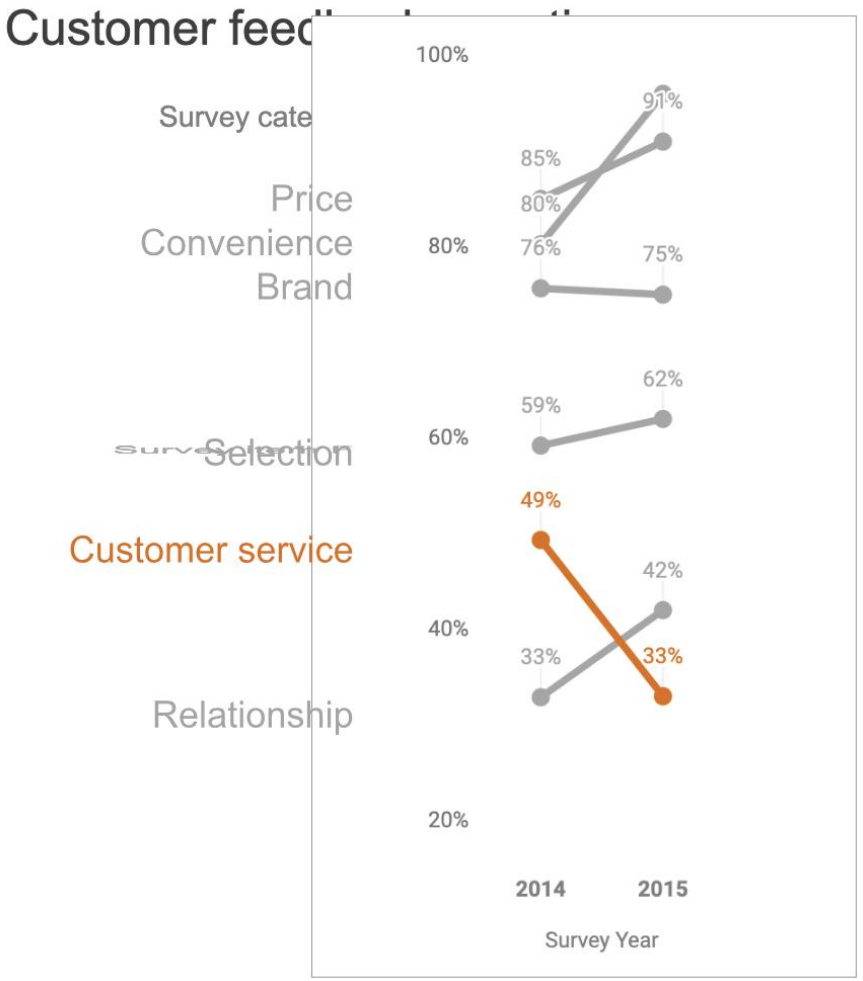
Heatmap

LOW-HIGH

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

Lecture 2 (30 min): Storytelling With Data
Topic: Choosing an effective visual (Chapter 2)

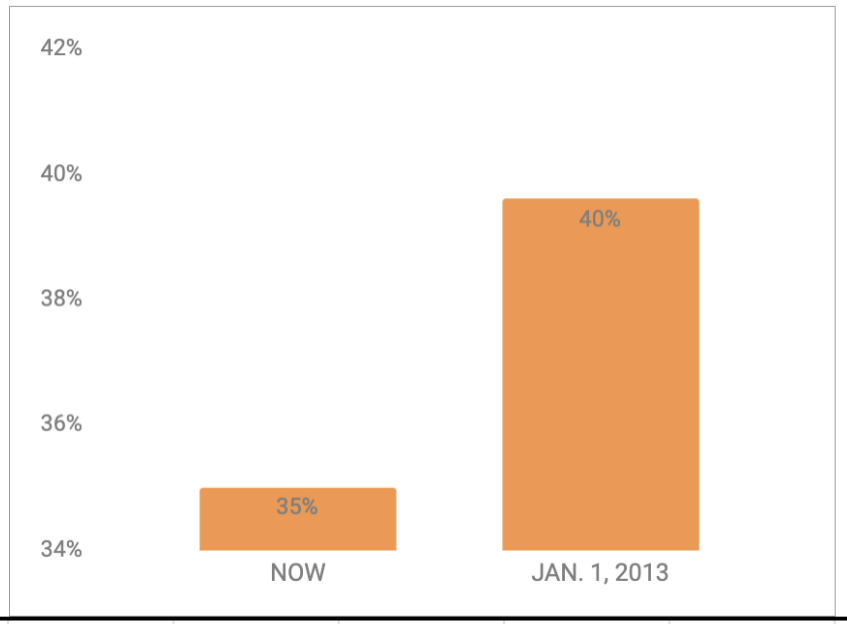




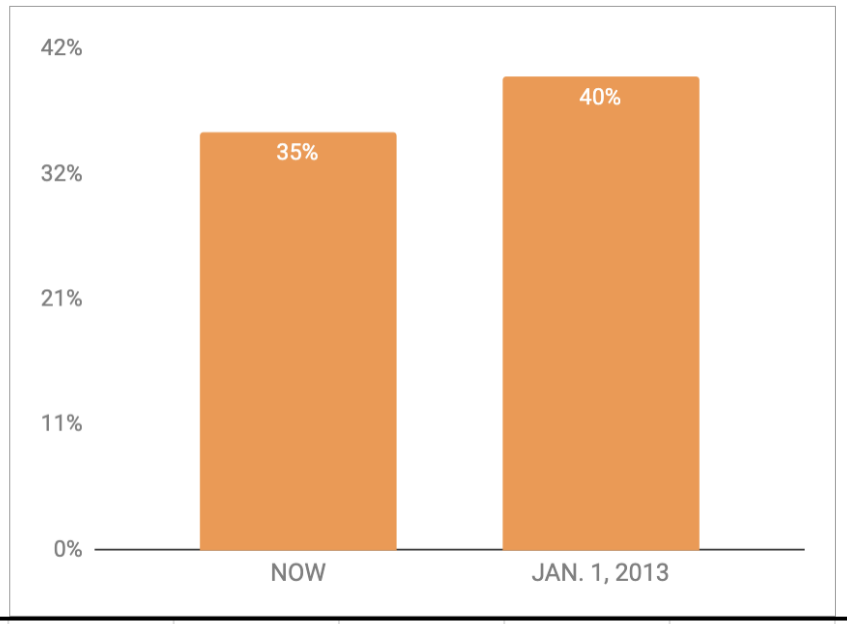
Lecture 2 (30 min): Storytelling With Data
Topic: Choosing an effective visual (Chapter 2)

FIG0213

Non-zero baseline: as originally graphed



Zero baseline: as it should be graphed



Lecture 2 (30 min): Storytelling With Data
Topic: Choosing an effective visual (Chapter 2)

FIG0214

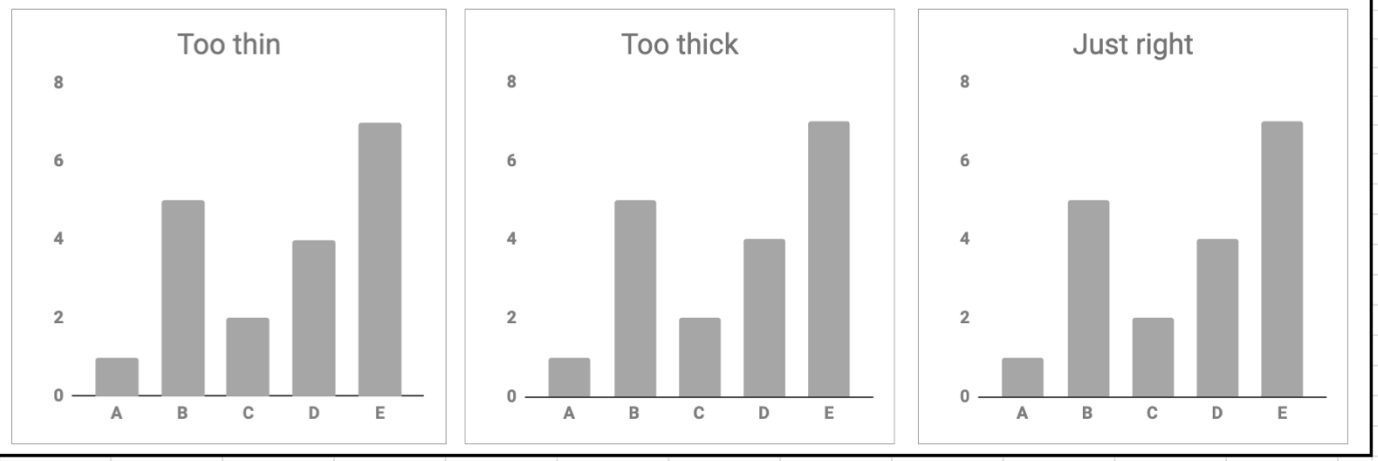
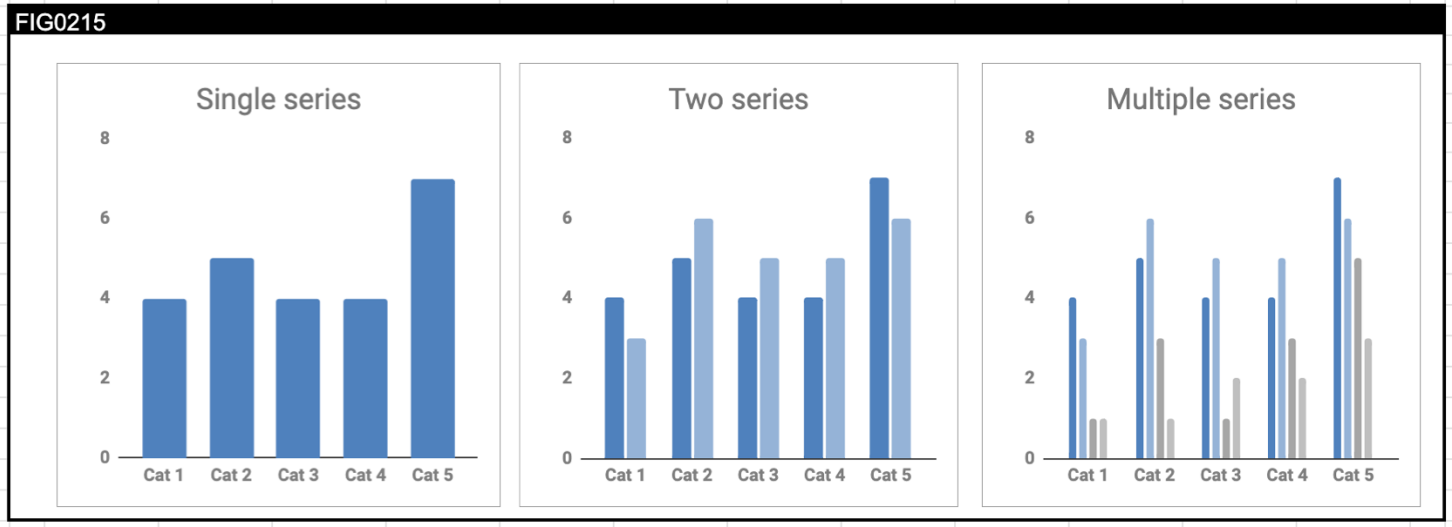


FIG0215



Lecture 2 (30 min): Storytelling With Data
Topic: Choosing an effective visual (Chapter 2)

FIG0216

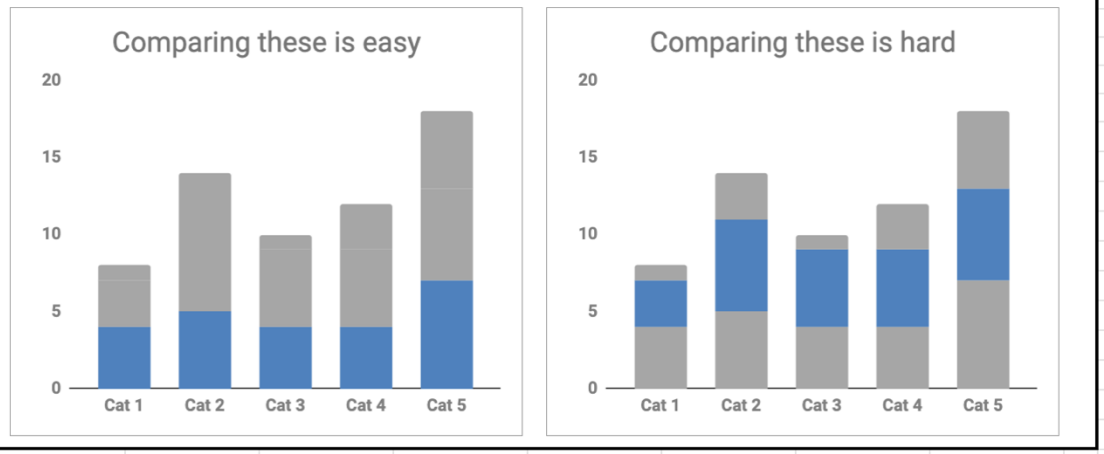
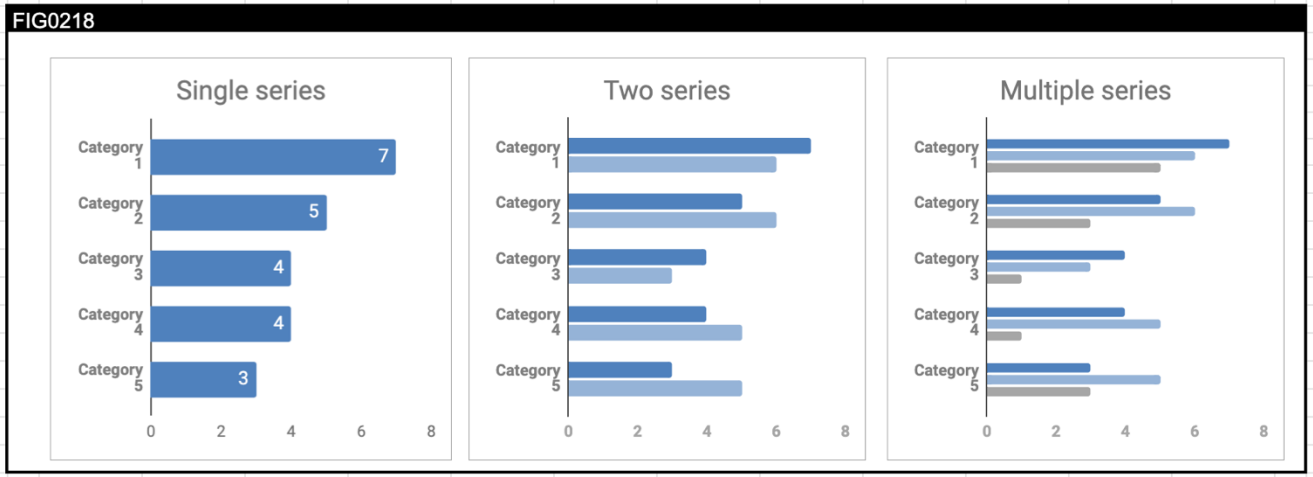
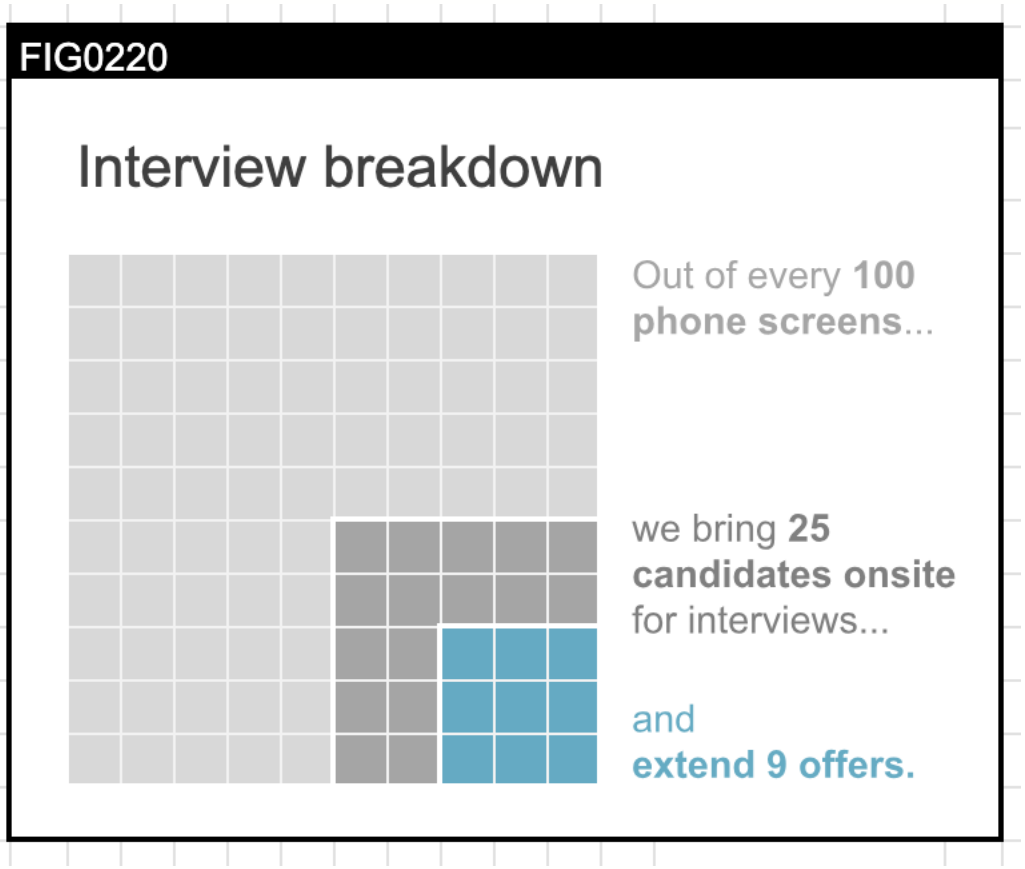


FIG0218



Lecture 2 (30 min): Storytelling With Data
Topic: Choosing an effective visual (Chapter 2)



Lecture 2 (30 min): Storytelling With Data
Topic: Choosing an effective visual (Chapter 2)

FIG0222

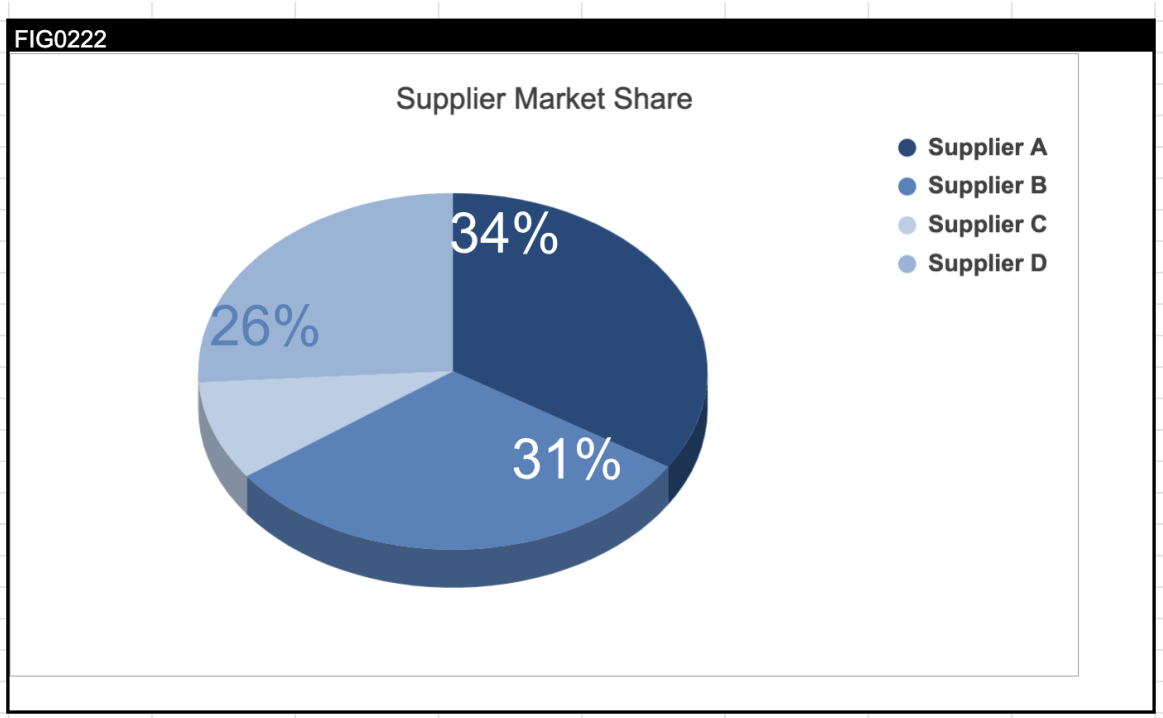
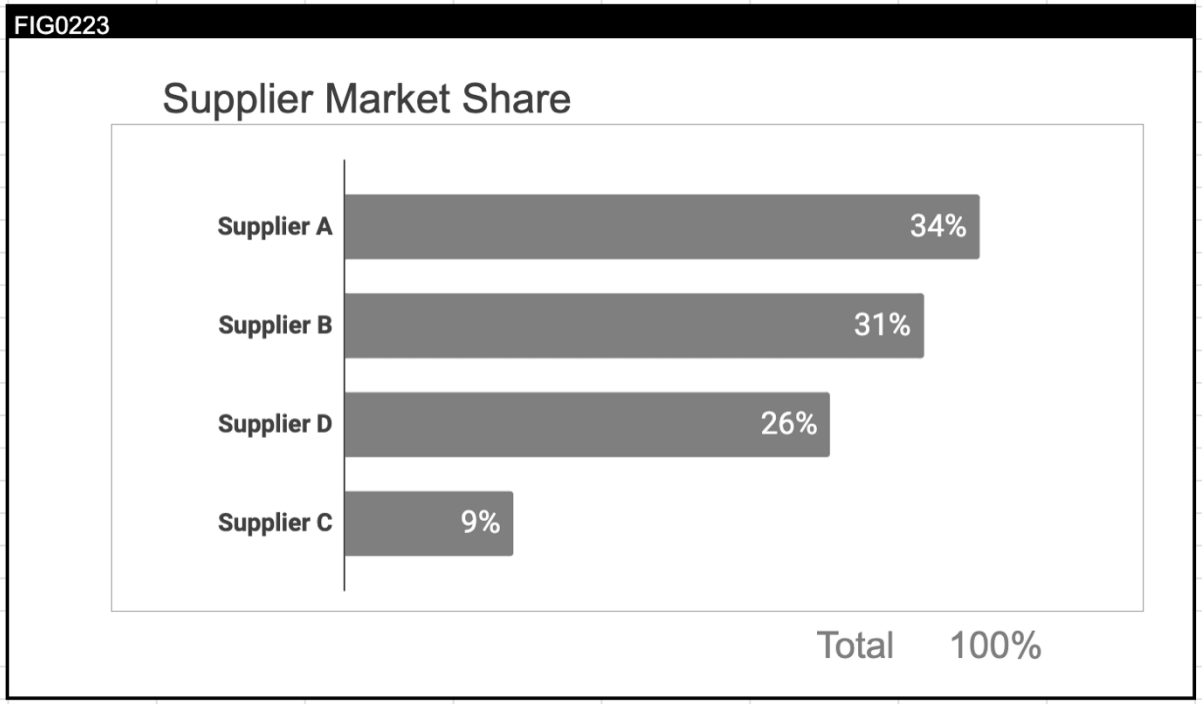


FIG0223



Skewed 3D perspective makes comparison of sector sizes hard!

Better!

Lecture 2 (30 min): Storytelling With Data
Topic: Choosing an effective visual (Chapter 2)

FIG0226

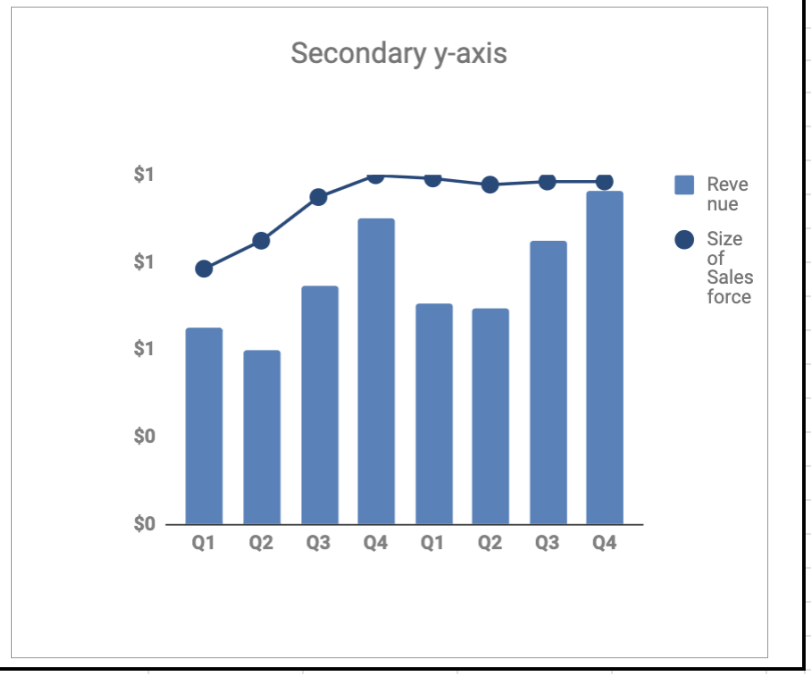
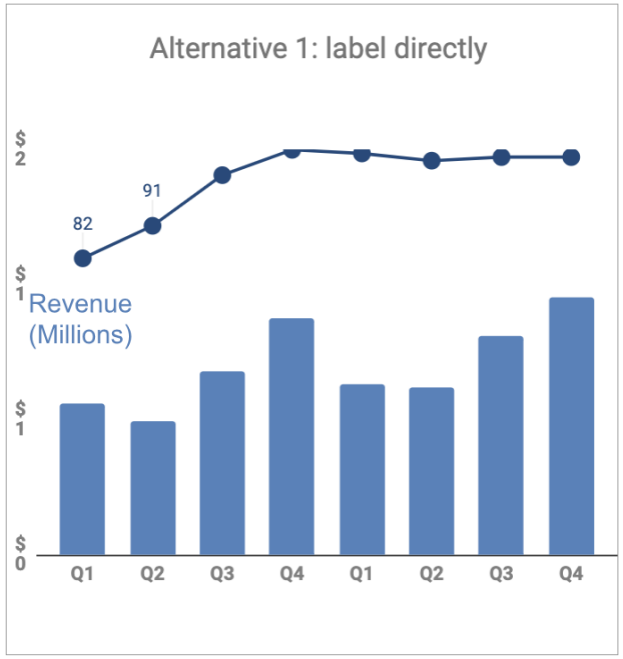
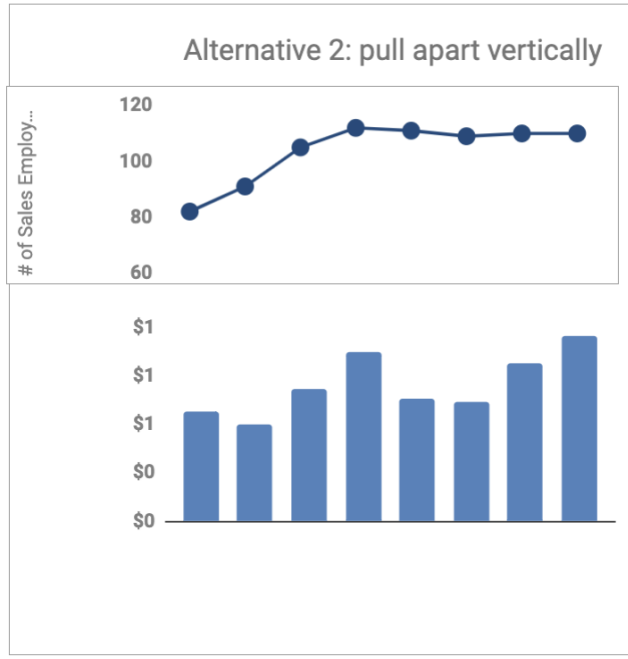


FIG0227





Lecture 2 (30 min): Storytelling With Data
Topic: Choosing an effective visual (Chapter 2)

Helpful online resource:

choose an effective visual
with the
SWD CHART GUIDE

https://www.storytellingwithdata.com/chart-guide

Class ends at 2:30 p.m.

Course Schedule

#	Topic and Objectives
1	**Intro & Getting Started** • Course Overview (relevance, examples, etc.) • Market-Ready-to-do List (MRTDL) by Career Advancement Center • Pick a dataset (Datasets will be provided by instructor on first day of class. If you already have formed a team of 4 students to collaborate and work together and want to use your own dataset, this needs to be vetted and approved by the instructor. Examples: something you are working on from another project, Walmart data on Kaggle, synthetic data from Synthea, etc.) • Explanation of peer-to-peer evaluation of presentations every week • Instruction: Exploratory data analysis • Assignment: **Prepare 1-minute "describe your dataset" presentation**
2	Start with Presentations: (present what was assigned in the previous class) **Understand the Business (and core business processes)** • Activity (for a specific case or example business): Describe the business for an example business (inputs, activities, outputs/metrics), develop a simple flowchart, identify opportunities • Instruction: Understanding the business problem, extracting the use case(s) • Assignment (for your selected business): **Prepare 3-minute presentation that describes the business, core business process(es), and opportunities for your selected business**
3	Start with Presentations: (present what was assigned in the previous class) **Identify a Business Problem (and why it needs to be addressed)** • Activity (for a specific case or example business problem): Developing persuasive arguments; Create tension with a visualization (draft) • Instruction: Story telling with data - visualizations • Assignment (for your selected business problem): **Create a 3-minute "tension" presentation; only 1 visual**

Course Schedule

4	Start with Presentations: (present what was assigned in the previous class) **Develop a Solution Pitch (for solving the identified business problem)** • Activity (for an example business problem): Big idea, exec summary, peer review • Instruction: Feasibility study, selection of final use case (big idea) • Assignment (for your selected business problem): **Create a 1-minute pitch (includes business overview, tension, and solution)**
5	Start with Presentations: (present what was assigned in the previous class) **Provide a Progress Update (for an ongoing project)** • Activity (for your selected solution): Strong visualizations, exploration, status, revisions, issues, lessons learned • Instruction: Data exploration and feedback loops with business stakeholders • Assignment (for your selected solution): **Prepare a 5-minute presentation; 5 slides** (excluding title slide); **3 visualizations**
6	Start with Presentations: (present what was assigned in the previous class) **Planning a Final Presentation and Final Report (for a completed project)** • Activity (for your project): 1st draft of headlines only and main messages per slide; **Python Notebook for technical audience** and **Word document for leadership**: clear connections to final presentation, i.e., same structure/order, etc., including a narrative in the final report • Instruction: Technical writing skills • Assignment (for your project): **Complete final presentation and reports; 7-minute presentation; 5-7 slides** (excluding title slide); appendix if required
7	**Final Presentations** (and final reports, notebook and Word document) are due