Data Analysis Course 6: Share Data through the Art of Visualization

Module 1

Data visualization: the graphic representation and presentation of data.

- putting information into an image to make it easier to understand

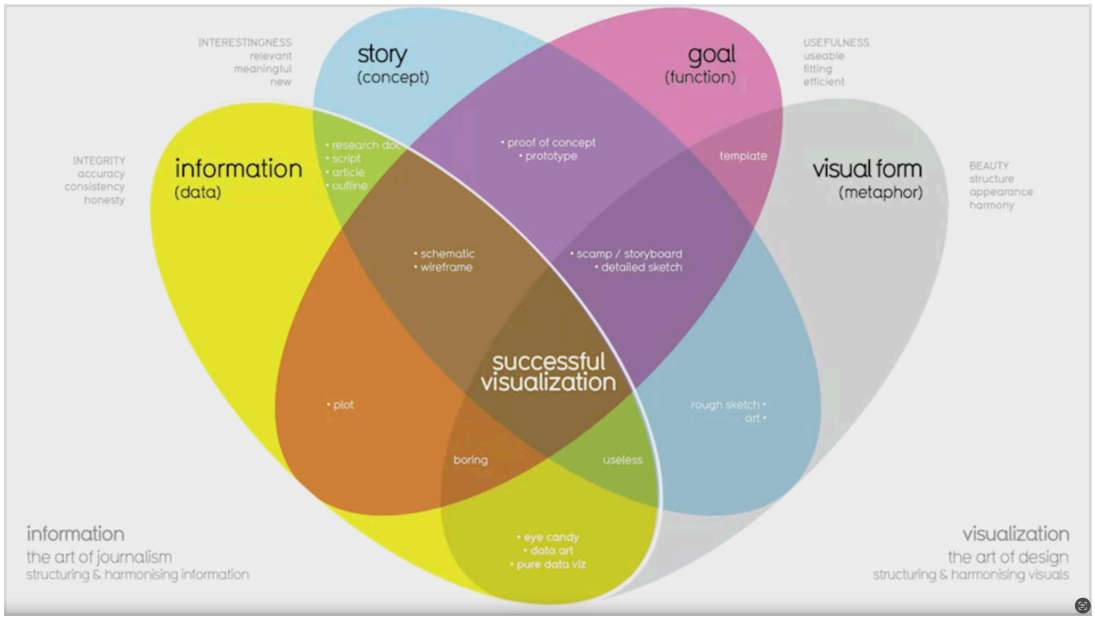
Visualizations began with maps

Scientists and mathematicians began to truly embrace the idea of arranging data visually in the 1700s and 1800s

How data analysts use visualization:

1. Looking at visuals in order to understand and draw conclusions about data
2. Creating visuals using raw data to tell a story

Rule of Thumb for Visualizations:
Creating visualizations: your audience should be able to tell what they're looking at within the first 5 seconds of looking at it.
In the next 5 seconds, they should be able to recognize the conclusion your visualization is making.



Story and data combined provide an outline of what you're trying to show.

Two frameworks for organizing data

- Frameworks help organize your thoughts about data viz and give a useful

checklist to reference as you plan and evaluate your data viz.
- McCandless method: provides terminology to isolate specific elements of a graphic.
  - Information: the data with which you're working
  - Story: a clear and compelling narrative or concept
  - Goal: a specific objective or function for the visual
  - Visual form: an effective use of metaphor or visual expression
- Kaiser Fung's Junk Charts Trifecta Checkup: questions to determine viz's effectiveness.
  - What is the practical question?
  - What does the data say?
  - What does the visual say?

Pre-attentive attributes: the elements of a data visualization that people recognize automatically and without conscious effort.
- Essential, basic building blocks that make visuals immediately understandable are called marks and channels.
  - Marks: basic visual objects such as points, lines, and shapes. Every mark can be broken down into four qualities:
    - 1. Position: Where is a specific mark in space relative to a scale or other marks?
    - 2. Size: How big, small, long, or tall is a mark?
    - 3. Shape: Does the shape of a specific object communicate something abut it?
    - 4. Color: What color is a mark?
  - Channels: visual aspects or variables that represent characteristics of the data in a visualization. Specialized marks that have been used to visualize data. Three elements determine effectiveness:
    - 1. Accuracy: Are the channels helpful in accurately estimating the values being represented?
    - 2. Popout: How easy is it to distinguish certain values from others?
    - 3. Grouping: How effective is a channel at communicating groups that exist in the data?

Bar graphs: use size contrast to compare two or more values.
Line graphs: help your audience understand shifts or changes in your data.
Pie chart: show how much each part of something makes up the whole.
Maps: help organize data geographically

One of your biggest consideration when creating a data visuzliation is where you'd like your audience to focus.

Histogram: a chart that shows how often data values fall into certain ranges.

Correlation charts: show relationships among data.

- Correlation in statistics is the measure of the degree to which two variables move in relationship to each other.
- Causation: occurs when an action/event directly leads to an outcome.
  - Correlation ≠ Causation
  - Be careful not to show causation where it doesn't exist.
  - In your data analysis, remember to:
    - Critically analyze any correlations that you find
    - Examine the data's context to determine if a causation makes sense (and can be supported by all of the data)
    - Understand the limitations of the tools that you use for analysis

Time series charts
Ranked bar charts

Static visualizations: do not change over time unless they're edited.

Dynamic visualizations: visualizations that are interactive or change over time

Tableau: a business intelligence and analytics platform that helps people see, understand, and make decisions with data.

Scatterplots show relationships between different variables. Scatterplots are typically used for two variables for a set of data, although additional variables can be displayed.

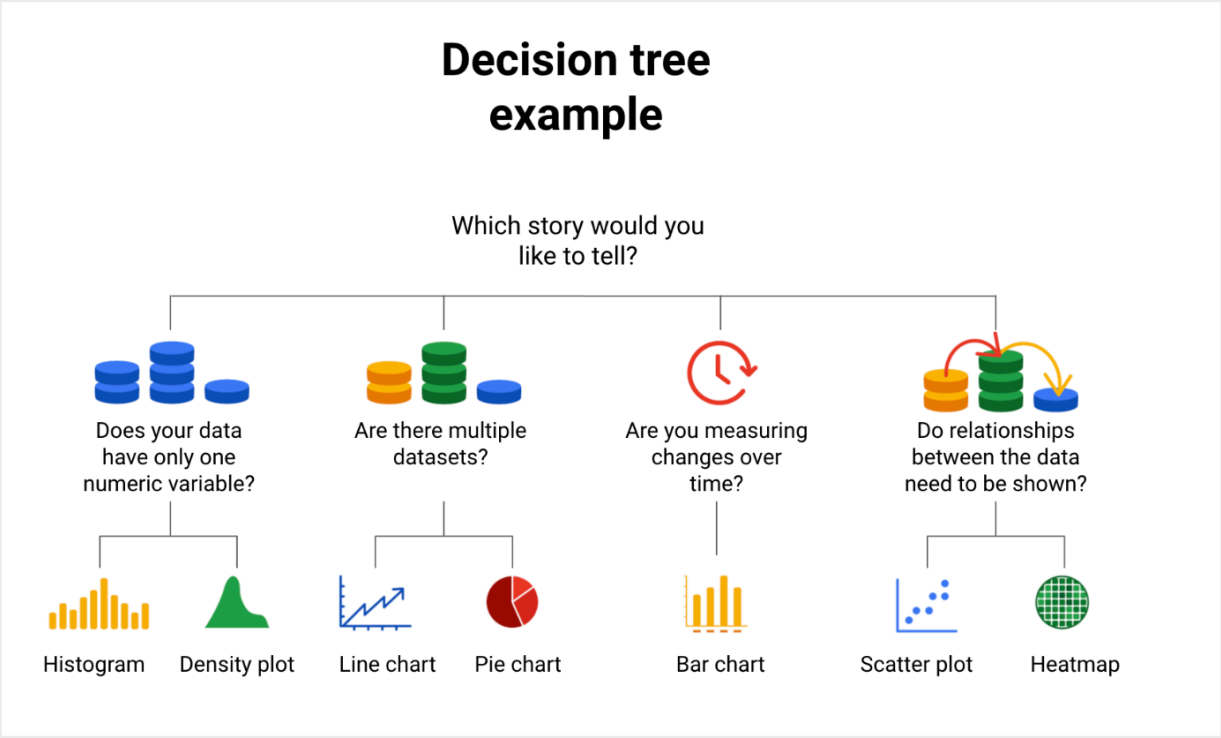
Distribution graph displays the spread of various outcomes in a dataset.

Considerations for what you want to communicate:

- Change: This is a trend or instance of observations that become different over time. A great way to measure change in data is through a line or column chart.
- Clustering: A collection of data points with similar or different values. This is best represented through a distribution graph.
- Relativity: These are observations considered in relation or in proportion to something else. You have probably seen examples of relativity data in a pie chart.
- Ranking: This is a position in a scale of achievement or status. Data that requires ranking is best represented by a column chart.
- Correlation: This shows a mutual relationship or connection between two

or more things. A scatterplot is an excellent way to represent this type of data pattern.

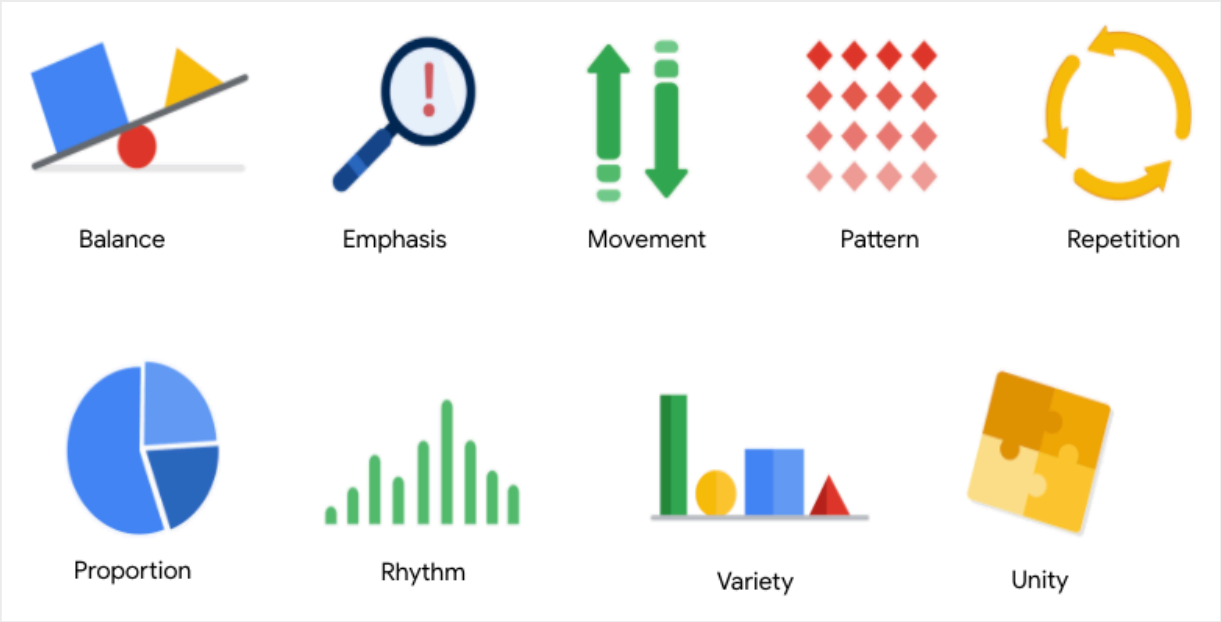
Decision tree: a decision-making tool that allows you to make decisions based on key questions that you can ask yourself. This can help you make decisions about critical features of your visualization.



Selecting the best chart

Elements of art:

- Line
- Shape: always 2D (pie chart, maps)
- Color:
  - Hue: color
  - Intensity: bright or dull
  - Value: light (tints) or dark (shades)
- Space: around, between, and in the objects
- Movement: create a sense of flow or action
  - use sparingly



9 Basic Principles of Design (first 6 are considerations while creating viz, last 3 are useful checks once viz is completed)

1. Balance: making sure key visual elements are distributed evenly
2. Emphasis: having a focal point so the audience knows where to concentrate.
3. Movement: can refer to the path the viewer's eye travels as they look at a data visualization or the literal movement created by animations.
4. Pattern: using similar shapes and colors to create patterns.
5. Repetition: repeating chart types, shapes, or colors adds to the effectiveness of your visualization.
6. Proportion: using various colors and sizes helps demonstrate that you are caring attention to a specific visual over others.
7. Rhythm: creating a sense of movement or flow in your viz.
8. Variety: viz should have some variety in char types, lines, shapes, colors, and values you use.
9. Unity: final viz should be cohesive.

Data composition: combining the individual parts in a visualization and displaying them together as a whole.

Elements for effective visuals:

- clear meaning
- sophisticated use of contrast
- refined execution

Four elements of successful visualizations:

1. information (data)
2. Story (concept)
3. Goal (function)
4. Visual form (metaphor)

Design thinking: a process used to solve complex problems in a user-centric way.

- adopting a user based mindset.

Five phases of the design process:

- Empathize: think about emotions and needs of target audience
- Define: helps you find your audience's needs, their problems, and your insights.
- Ideate: generate your data viz ideas
- Prototype: start putting dashboards, charts, or other visualizations together for testing and feedback
- Test: show them to other team members before sharing them with stakeholders

Headline: a line of words printed in large letters at the top of a visualization to communicate what data is being presented.

Subtitle: supports headline by adding more context and description.

Label: identifies data in relation to other data. Most commonly indicate what the x and y axis show.

Annotation: briefly explains data or helps focus the audience on a particular aspect of the data in a visualization.

Visualization components	Guidelines	Style checks
Headlines	- **Content**: Briefly describe the data - **Length**: Usually the width of the data frame - **Position**: Above the data	- Use brief language - Don’t use all caps - Don’t use italic - Don’t use acronyms - Don't use abbreviations - Don’t use humor or sarcasm
Subtitles	- **Content**: Clarify context for the data - **Length**: Same as or shorter than headline - **Position**: Directly below the headline	- Use smaller font size than headline - Don’t use undefined words - Don’t use all caps, bold, or italic - Don’t use acronyms - Don't use abbreviations
Labels	- **Content**: Replace the need for legends - **Length**: Usually fewer than 30 characters - **Position**: Next to data or below or beside axes	- Use a few words only - Use thoughtful color-coding - Use callouts to point to the data - Don’t use all caps, bold, or italic
Annotations	- **Content**: Draw attention to certain data - **Length**: Varies, limited by open space - **Position**: Immediately next to data annotated	- Don’t use all caps, bold, or italic - Don't use rotated text - Don’t distract viewers from the data

People take in info in a lot of different ways.
Accessibility

- Labeling directly
- Text alternatives
- Text-based content
- Distinguishing
- Simplify

Alternative text: provides a textual alternative to non-text content.

Make a chart in 60 minutes:

- Prep (5 min): Create the mental and physical space necessary for an environment of comprehensive thinking. This means allowing yourself room to brainstorm how you want your data to appear while considering the amount and type of data that you have.

- Talk and listen (15 min): Identify the object of your work by getting to the "ask behind the ask" and establishing expectations. Ask questions and really concentrate on feedback from stakeholders regarding your projects to help you hone how to lay out your data.
- Sketch and design (20 min): Draft your approach to the problem. Define the timing and output of your work to get a clear and concise idea of what you are crafting.
- Prototype and improve (20 min): Generate a visual solution and gauge its effectiveness at accurately communicating your data. Take your time and repeat the process until a final visual is produced. It is alright if you go through several visuals until you find the perfect fit.

Terms and definitions for Course 6, Module 1

Alternative text: Text that provides an alternative to non-text content, such as images and videos

Annotation: Text that briefly explains data or helps focus the audience on a particular aspect of the data in a visualization

AVERAGEIF: A spreadsheet function that returns the average of all cell values from a given range that meet a specified condition

Balance: The design principle of creating aesthetic appeal and clarity in a data visualization by evenly distributing visual elements

Bar graph: A data visualization that uses size to contrast and compare two or more values

Calculus: A branch of mathematics that involves the study of rates of change and the changes between values that are related by a function

Causation: When an action directly leads to an outcome, such as a cause-effect relationship

Channel: A visual aspect or variable that represents characteristics of the data in a visualization

Chart: A graphical representation of data from a worksheet

Cluster: A collection of data points on a data visualization with similar values

CONVERT: A SQL function that changes the unit of measurement of a value in data

Correlation: The measure of the degree to which two variables change in relationship to each other

CREATE TABLE: A SQL clause that adds a temporary table to a database that can be used by multiple people

Data composition: The process of combining the individual parts in a visualization and displaying them together as a whole

Decision tree: A tool that helps analysts make decisions about critical features of a visualization

Design thinking: A process used to solve complex problems in a user-centric way

Distribution graph: A data visualization that displays the frequency of various outcomes in a sample

DROP TABLE: A SQL clause that removes a temporary table from a database
Dynamic visualizations: Data visualizations that are interactive or change over time
Emphasis: The design principle of arranging visual elements to focus the audience's attention on important information in a data visualization
HAVING: A SQL clause that adds a filter to a query instead of the underlying table that can only be used with aggregate functions
Headline: Text at the top of a visualization that communicates the data being presented
Heat map: A data visualization that uses color contrast to compare categories in a dataset
Histogram: A data visualization that shows how often data values fall into certain ranges
Inner query: A SQL subquery that is inside of another SQL statement
Label: Text in a visualization that identifies a value or describes a scale
Legend: A tool that identifies the meaning of various elements in a data visualization
Line graph: A data visualization that uses one or more lines to display shifts or changes in data over time
Map: A data visualization that organizes data geographically
Mark: A visual object in a data visualization such as a point, line, or shape
MAXIFS: A spreadsheet function that returns the maximum value from a given range that meets a specified condition
Mental model: A data analyst's thought process and approach to a problem
Movement: The design principle of arranging visual elements to guide the audience's eyes from one part of a data visualization to another
MINIFS: A spreadsheet function that returns the minimum value from a given range that meets a specified condition
Narrative: (Refer to story)
Ordinal data: Qualitative data with a set order or scale
Pattern: The design principle of using similar visual elements to demonstrate trends and relationships in a data visualization
Pie chart: A data visualization that uses segments of a circle to represent the proportions of each data category compared to the whole
Pre-attentive attributes: The elements of a data visualization that an audience recognizes automatically without conscious effort
Proportion: The design principle of using the relative size and arrangement of visual elements to demonstrate information in a data visualization
R: A programming language used for statistical analysis, visualization, and other data analysis
Ranking: A system to position values of a dataset within a scale of achievement or status
Relativity: The process of considering observations in relation or proportion to something else
Repetition: The design principle of repeating visual elements to demonstrate

meaning in a data visualization
Rhythm: The design principle of creating movement and flow in a data visualization to engage an audience
Scatterplot: A data visualization that represents relationships between different variables with individual data points without a connecting line
SELECT INTO: A SQL clause that copies data from one table into a temporary table without adding the new table to the database
Sort range: A spreadsheet menu function that sorts a specified range and preserves the cells outside the range
Sort sheet: A spreadsheet menu function that sorts all data by the ranking of a specific sorted column and keeps data together across rows
Static visualization: A data visualization that does not change over time unless it is edited
Story: The narrative of a data presentation that makes it meaningful and interesting
Subtitle: Text that supports a headline by adding context and description
Tableau: A business intelligence and analytics platform that helps people visualize, understand, and make decisions with data
Unity: The design principle of using visual elements that complement each other to create aesthetic appeal and clarity in a data visualization
Variety: The design principle of using different kinds of visual elements in a data visualization to engage an audience
Visual form: The appearance of a data visualization that gives it structure and aesthetic appeal
X-axis: The horizontal line of a graph usually placed at the bottom, which is often used to represent time scales and discrete categories
Y-axis: The vertical line of a graph usually placed to the left, which is often used to represent frequencies and other numerical variables

Three essential elements of effective visuals:

1. Clear meaning
2. Sophisticated use of contrast
3. Refined execution

Module 2

Tableau: a business intelligence and analytics platform that helps people see, understand, and make decisions with data.

- Looker and Google DataStudio are similar

Tableau creates dynamic visualizations.

Chart types in spreadsheets:

- Column (vertical bar): a column chart allows you to display and compare

multiple categories of data by their values.

- Line: a line chart showcases trends in your data over a period of time. The last line chart example is a combo chart which can include a line chart. Refer to the description for the combo chart type.
- Pie: a pie chart is an easy way to visualize what proportion of the whole each data point represents.
- Horizontal bar: a bar chart functions similarly to a column chart, but is flipped horizontally.
- Area: area charts allow you to track changes in value across multiple categories of data.
- Scatter: scatterplots are typically used to display trends in numeric data.
- Combo: combo charts use multiple visual markers like columns and lines to showcase different aspects of the data in one visualization. The example below is a combo chart that has a column and line chart together.

Tableau resources:

- Which chart or graph is right for you?
- The Ultimate Cheat Sheet on Tableau Charts

Types of viz in Tableau:

- Highlight tables appear like tables with conditional formatting. Review the steps to build a highlight table.
- Heat maps show intensity or concentrations in the data. Review the steps to build a heat map
- Density maps illustrate concentrations (such as a population density map). Refer to instructions to create a heat map for density
- Gantt charts demonstrate the duration of events or activities on a timeline. Review the steps to build a Gantt chart
- Symbol maps display a mark over a given longitude and latitude. Learn more from this example of a symbol map
- Filled maps are maps with areas colored based on a measurement or dimension. Explore an example of a filled map
- Circle views show comparative strength in data. Learn more from this example of a circle view
- Box plots, also known as box and whisker charts, illustrate the distribution of values along a chart axis. Refer to the steps to build a box plot
- Bullet graphs compare a primary measure with another and can be used instead of dial gauge charts. Review the steps to build a bullet graph
- Packed bubble charts display data in clustered circles. Review the steps to build a packed bubble chart

Different signs in Tableau represent different kinds of data

- #: numeric data
- Abc: string data
- Globe: geographic data
- Calendar: date data
- Calendar w/ a clock: date and time data

In Tableau sheet, on left side there are dimensions of the data above the line

- Measures you can track for these dimensions are below the line

Tableau uses dimensions and measures to generate customized charts.

Diverging color palette: displays two ranges of values using color intensity to show the magnitude of the number and the actual color to show which range the number is from.

Principle	Description
Choose the right visual	One of the first things you have to decide is which visual will be the most effective for your audience. Sometimes, a simple table is the best visualization. Other times, you need a more complex visualization to illustrate your point.
Optimize the data-ink ratio	The data-ink entails focusing on the part of the visual that is essential to understanding the point of the chart. Try to minimize non-data ink like boxes around legends or shadows to optimize the data-ink ratio.
Use orientation effectively	Make sure the written components of the visual, like the labels on a bar chart, are easy to read. You can change the orientation of your visual to make it easier to read and understand.
Color	There are a lot of important considerations when thinking about using color in your visuals. These include using color consciously and meaningfully, staying consistent throughout your visuals, being considerate of what colors mean to different people, and using inclusive color scales that make sense for everyone viewing them.
Numbers of elements	Think about how many elements you include in any visual. If your visualization uses lines, try to plot five or fewer. If that isn't possible, use color or hue to emphasize important lines. Also, when using visuals like pie charts, try to keep the number of segments to less than seven since too many elements can be distracting.

Marks and channels are pre-attentive attributes of visualizations.

The above graphic shows essential design principles.

Max five lines in visualizations or use hue to emphasize important stuff.
Max 7 slices in pie chart, otherwise it's distracting.

Common errors to avoid:

What to avoid	Why
Cutting off the y-axis	Changing the scale on the y-axis can make the differences between different groups in your data seem more dramatic, even if the difference is actually quite small.
Misleading use of a dual y-axis	Using a dual y-axis without clearly labeling it in your data visualization can create extremely misleading charts.
Artificially limiting the scope of the data	If you only consider the part of the data that confirms your analysis, your visualizations will be misleading because they don't take all of the data into account.
Problematic choices in how data is binned or grouped	It is important to make sure that the way you are grouping data isn't misleading or misrepresenting your data and disguising important trends and insights.
Using part-to-whole visuals when the totals do not sum up appropriately	If you are using a part-to-whole visual like a pie chart to explain your data, the individual parts should add up to equal 100%. If they don't, your data visualization will be misleading.
Hiding trends in cumulative charts	Creating a cumulative chart can disguise more insightful trends by making the scale of the visualization too large to track any changes over time.
Artificially smoothing trends	Adding smooth trend lines between points in a scatter plot can make it easier to read that plot, but replacing the points with just the line can actually make it appear that the point is more connected over time than it actually was.

Reminders about making a helpful data viz:

- Five-second rule: A data visualization should be clear, effective, and convincing enough to be absorbed in five seconds or less.
- Color contrast: Graphs and charts should use a diverging color palette to show contrast between elements.
- Conventions and expectations: Visuals and their organization should align with audience expectations and cultural conventions. For example, if the majority of your audience associates green with a positive concept and red with a negative one, your visualization should reflect this.
- Minimal labels: Titles, axes, and annotations should use as few labels as it takes to make sense. Having too many labels makes your graph or chart

too busy. It takes up too much space and prevents the labels from being shown clearly.

Line chart is ideal for highlighting trends over time.
Histogram is ideal for comparing the distribution of two variables by individual grouping.

Terms and definitions for Course 6, Module 2
Area chart: A data visualization that uses individual data points for a changing variable connected by a continuous line with a filled in area underneath
Box plot: A data visualization that displays the distribution of values along an x-axis
Bubble chart: A data visualization that displays individual data points as bubbles, comparing numeric values by their relative size
Bullet graph: A data visualization that displays data as a horizontal bar chart moving toward a desired value
Circle view: A data visualization that shows comparative strength in data
Column chart: A data visualization that uses individual data points for a changing variable, represented as vertical columns
Combo chart: A data visualization that combines more than one visualization type
Density map: A data visualization that represents concentrations, with color representing the number or frequency of data points in a given area on a map
Distribution graph: A data visualization that displays the frequency of various outcomes in a sample
Diverging color palette: A color theme that displays two ranges of data values using two different hues, with color intensity representing the magnitude of the values
Donut chart: A data visualization where segments of a ring represent data values adding up to a whole
Filled map: A data visualization that colors areas in a map based on measurements or dimensions
Gantt chart: A data visualization that displays the duration of events or activities on a timeline
Gauge chart: A data visualization that shows a single result within a progressive range of values
Highlight table: A data visualization that uses conditional formatting and color on a table
Packed bubble chart: A data visualization that displays data in clustered circles
Symbol map: A data visualization that displays a mark over a given longitude and latitude

Module 3
Dashboard: a tool that organizes information from multiple datasets into one

central location for tracking, analysis, and simple visualization.

Dashboard filter: a tool for showing only the data that meets a specific criteria while hiding the rest.

“Numbers have an important story to tell. They rely on you to give them a clear and convincing voice.” -Stephen Few

Data storytelling: communicating the meaning of a dataset with visuals and a narrative that are customized for each particular audience.

3 data storytelling steps

1. Engage your audience
2. Create compelling visuals
    1. Show, not just tell.
3. Tell the story in an interesting narrative
    1. Beginning, middle, end
    2. Organized and concise

Engage your audience: Capture and hold their attention.

- Know your audience
    - What role does this audience play?
    - Wha tis their stake in the project?
    - What do they hope to get from the data insights I deliver?
- Choose your primary message
    - Spotlighting: scanning through the data to quickly identify the most important insights.
        - Sticky notes on a whiteboard, find ideas that keep coming up again and again
    - Explore your discoveries and find the meaning behind the numbers
    - Keep it clear and concise

Create compelling visuals: Show, don’t just tell

Vertical or horizontal layout in Tableau

- Even distribution is important in dashboards
- Tiled items: part of single layered grid
- Floating items: can overlap other objects

Dashboards put storytelling power in the hands of the viewer

Live vs. Static data

- How do ID:
    - How old is the data?

    - How long until the insights are stale or no longer valid to make decisions?
    - Does this data or analysis need updating on a regular basis to remain valuable?
- Static data: provides screenshots or snapshots in presentations or building dashboards using snapshots of data.
    - Pros
        - Can tightly control a point-in-time narrative of the data and insight
        - Allows for complex analysis to be explained in-depth to a larger audience
    - Cons
        - Insight immediately begins to lose value and continues to do so the longer the data remains in a static state
        - Snapshots can't keep up with the pace of data change
- Live data: dashboards, reports, and views connected to automatically updated data.
    - Pros
        - Dashboards can be built to be more dynamic and scalable
        - Gives the most up-to-date data to the people who need it at the time when they need it
        - Allows for up-to-date curated views into data with the ability to build a scalable "single source of truth" for various use cases
        - Allows for immediate action to be taken on data that changes frequently
        - Alleviates time/resources spent on processes for every analysis
    - Cons
        - Can take engineering resources to keep pipelines live and scalable, which may be outside the scope of some companies' data resource allocation
        - Without the ability to interpret data, you can lose control of the narrative, which can cause data chaos (i.e. teams coming to conflicting conclusions based on the same data)
        - Can potentially cause a lack of trust if the data isn't handled properly

Filters can be used in Tableau

Tell the story in an interesting narrative: Organized and concise

- Characters: bring more human context
- Setting
- Plot: creates tension
- Big reveal: how the data has shown you can solve the problem

- Aha moment: share your recs and explain why you think they'll be successful

Themes are great for presentations.

- They control the color, font types and sizes, formatting, and positioning of text and visuals.

Include title and subtitle
Add a date on title slide

Make sure it's easy to read.

- Less than 5 lines and 25 words per slide
- Choose words carefully
  - Avoid abbreviations

Great visuals don't leave room for interpretation because they are easily understood.

Ask yourself, "What's the single most important thing I want my audience to learn from my analysis?"

- Helps decide visuals

Link or images in slideshow?

- Paste, link, or embed all have pros and cons and behave differently
- When you copy and paste a visual into your presentation, you can edit it directly within your slideshow. If your visual or its data points exist in other places, such as a Tableau dashboard, any changes you make will not affect them there.

Terms and definitions for Course 6, Module 3
Data storytelling: Communicating the meaning of a dataset with visuals and a narrative that are customized for an audience
Engagement: Capturing and holding someone's interest and attention during a data presentation
Live data: Data that is automatically updated
Spotlightling: Scanning through data to quickly identify the most important insights

Module 4

Use a strategic framework for your presentation
The framework for your presentation starts with your understanding of the business task.

Use speaker notes in presentations

By showcasing what business metrics you used, you can help your audience understand the impact your findings will have.

Start by helping your audience understand what data was available during data collection.

Establish your initial hypothesis

- Hypothesis: the theory you're trying to prove or disprove with data
- Establish this early in the presentation to help your audience understand the data

Explain the solution to your business task using examples and visualizations.

McCandless Method

- moves from general to specific
  - 1. Introduce graphic by name
  - 2. Answer obvious questions before they're asked
  - 3. State the insight of your graphic
  - 4. Call out data to support that insight
  - 5. Tell your audience why it matters
    - Answer the "so what?"
    - Present the possible business impact of the solution and clear actions stakeholders can take

Ask yourself: "Does this data point or chart support the point I want people to walk away with?"

Evaluating a slide presentation:
Your evaluation criteria
When exploring a slide presentation, use your knowledge of effective presentation practices to evaluate it. This includes reviewing your own work! When you're checking over slide presentations, there are some best practices you can check for:

- **Include a title, subtitle, and date:** Making sure that your slide deck presentation has a title, subtitle, and date makes sure that your audience knows exactly what you are presenting and when the information was from. That way they know it's relevant and current to them!
- **Use a logical sequence of slides:** Organizing your slides in an order that makes sense guides your audience through your narrative, building understanding step by step.

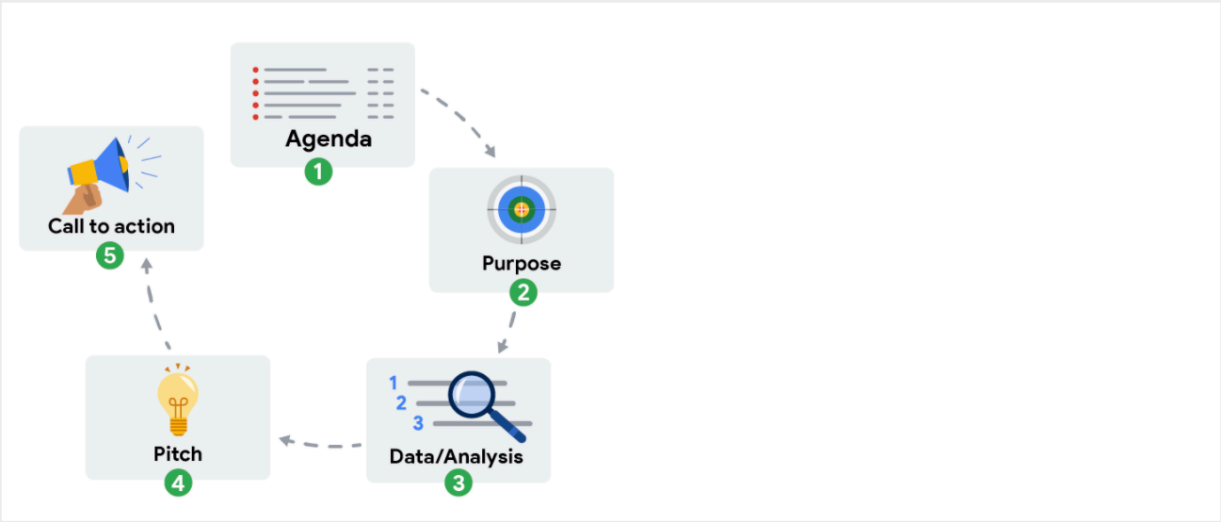
- **Provide an agenda with a timeline:** An agenda offers a roadmap of your presentation, allowing your audience to follow along and anticipate key topics.
- Limit the amount of text on slides: Keeping text brief ensures clarity and retains the audience's attention; aim for your audience to scan it within 5 seconds.
- **Start with the business task:** By immediately relating the content to the business task at hand, you contextualize your information, making it relevant and actionable.
- **Establish the initial hypothesis:** Presenting an initial hypothesis gives your audience a starting point for what to expect and frames the subsequent analysis.
- **Show what business metrics you used:** Clarifying which metrics you're analyzing validates your arguments and helps the audience gauge your presentation's relevance to business outcomes.
- **Use visualizations:** Visual aids can illustrate complex data more effectively than text alone, making your message more accessible.
- **Introduce the graphic by name:** A brief introduction to each graphic aids in understanding and retaining information.
- **Provide a title for each graph:** Titles act as signposts, helping the audience quickly grasp the meaning of each visual.
- **Go from the general to the specific:** Starting with a broad overview before diving into details ensures that all audience members are on the same page.
- **Use speaker notes to help you remember talking points:** Notes act as your cue cards, enabling a smoother delivery and ensuring no critical point is missed.
- **Include key takeaways:** Summarizing the main points at the end of your presentation reinforces the message and ensures the audience leaves with the intended takeaways.

Like so many parts of your job as a data professional, creating presentations is an iterative process. Reviewing your work, making changes as needed, and improving it when you can will make your presentations clearer and more useful for stakeholders.

Avoid eyesore charts
Make your presentation fun
Incorporate storytelling
Make sure you have an ally in the room

Build familiarity throughout the presentation
Titles on every slide

Make sure you have conclusion/recommendation slide



A good slideshow:

- Title slide: Title, name of presenter, and last date updated
- Tables of contents type slide, outline the presentation
- What are we talking about? Share the objective
- Then start presenting the data
  - Have transitional phrases as you end one slide and go to another
  - Explain visuals in depth
- Conclusions: sum up your findings
- Recommendations?
- Final slide, thank you
- Allow time for Q&A

Correlation coefficient: a measure of strength and direction of the linear relationship between two variables.

- Closer to 1: more positively correlated (1 goes up, other also goes up)
- Closer to -1: more negatively correlated (1 goes up, other goes down)
- Closer to 0: not correlated at all

Proven presentation tips:

- Two key responsibilities of data analyst:
  - 1. Analyze data
  - 2. Present your findings effectively.
- Channel your excitement
  - Take deep, controlled breaths
- Start with the broader ideas
- Use the five second rule on all visualizations

    - Wait five seconds after showing a data viz
    - Ask if they understand it
    - Give your audience another five seconds
    - Tell them the conclusion
        - Don’t rush through data visualizations
        - It’ll be many in your audience’s first time seeing your data
- Preparation is key
    - More prepared you are, the better you’ll perform

Assume that everyone in your audience is busy.

Through critical self-review, use this checklist to refine both your verbal and visual communication, turning data into compelling stories. Here are some questions you can incorporate into your checklist:

- Do I use an attention-grabbing opening?
- Do I start with broad ideas and later talk about specific details?
- Do I speak in short sentences?
- Do I pause for five seconds after showing a data visualization?
- Do I pause intentionally at certain points?
- Do I keep the pitch of my sentences level?
- Do I stand still and move with purpose?
- Do I have good posture?
- Do I look at my audience (or camera) while speaking?
- Do I keep my message concise?
- Do I end by explaining to my audience why the data analysis matters?

You can also add checklist items that help you refine your slide deck:

- Do I include a good title and subtitle that describes what I’m about to present?
- Do I include the date of my presentation or the date when my slideshow was last updated?
- Does my font size let the audience easily read my slides?
- Do I showcase what business metrics I used?
- Do I include effective visuals (like charts and graphs)?

Your audience:

- Will not always see the steps you took to reach a conclusion
    - Curse of knowledge: Your work makes sense to you because you did it.
- Has a lot on their mind
- Is easily distracted

How you present is as important as what you present:

- Keep your sentences short
- Build in intentional pauses
- Keep the pitch of your sentences level

Be mindful of nervous habits you have:

- Stay still and move with purpose
- Practice good posture
- Control breathing, pace of speaking
- Make positive eye contact

Preparing for the Q&A:

- Before the Presentation:
    1. Assemble and prepare your questions
    2. Discuss your presentation with your manager, other analysts, or other friendly contacts in your organization.
    3. Ask a manager or other analysts what sort of questions were normally asked by your specific audience in the past.
    4. Seek comments, feedback, and questions on the deck or the document of your analysis.
    5. At least 24 hours ahead of the presentation, try and brainstorm tricky questions or unclear parts you may come across- this helps avoid surprises.
    6. It never hurts to practice what you will be presenting, to account for any missing information or simply to calm your nerves.
- During the presentation:
    1. Be prepared to respond to the things that you find and effectively and accurately explain your findings.
    2. Address potential questions that may come up.
    3. Avoid having a single question derail a presentation and propose following-up offline.
    4. Put supplementary visualization and content into the appendix to help answer questions.

Understand your stakeholders' expectations
Make sure you have a clear understanding of the objective and what the stakeholders wanted

The colleague test: a test run of your presentation to a colleague that has no previous experience of your work.

Start with zero assumptions: don't assume that your audience is already familiar with jargon, acronyms, past events, or other necessary background information.

Work with your team to anticipate questions and draft responses

Be prepared to consider any limitations of your data by:

- Critically analyzing the correlations
- Looking at the context
- Understanding the strengths and weaknesses of the tools you used in your analysis

Types of objections

- About data
  - Where and what systems from?
  - How fresh and accurate?
  - What transformations happened to it?
    - Keep detailed log of data transformations
- About your analysis
  - Is it reproducible?
    - Keep changelog
  - Who did you get feedback from?
    - Use lots of perspectives
- About your findings
  - Do these findings exist in previous time periods?
  - Did you control for the differences in your data?

Responding to possible objections

- Communicate any assumptions
- Explain why your analysis might be different than expected
- Acknowledge that those objections are valid and take steps to investigate further

Q&A best practices

- Listen to the whole question
- Repeat the question
- Understand the context
- Try to involve the whole audience
- Keep your responses short and to the point

Terms and definitions for Course 6, Module 4

Dashboard filter: A tool for showing only the data that meets a specific criteria while hiding the rest

Data blending: A Tableau method that combines data from multiple data sources

Framework: The context a presentation needs to create logical connections that

tie back to the business task and metrics
Professional relationship building: Building relationships by meeting people both in person and online (Refer to Networking)
Static data: Data that doesn’t change once it has been recorded
Statistics: The study of how to collect, analyze, summarize, and present data
WITH: A SQL clause that creates a temporary table that can be queried multiple times