Lesson 20 – Data Storytelling

**Questions for Mentor:**

**Storytelling through EDA:**

* Can find through data that the labels could not be right?
  + i.e. bacteria comparisons with antibiotics
* Visualization has 2 main goals
  + Explanatory
    - Present ideas
    - Explain and inform
    - Provide evidence and support
    - Influence and persuade
  + Exploratory
    - Explore data
    - Assess situation
    - Determine how to proceed
    - Decide what to do
* What makes a visualization effective?
  + Most are not effective
  + Important to be able to show visualization that allows you to see pattern without thinking very hard
* 5 principles to effective visualizations
  + Have graphical integrity
    - People tend to have misleading graphs (i.e. fox news examples)
    - Mistakes can hurt integrity and that’s hard to get back
  + Keep it simple
    - Data-ink ratio = data ink / total ink used in graph
      * Use ink to show data, don’t waste space with fancy crap
      * 3D is generally bad
      * Avoid chartjunk
        + i.e. needless gridlines, over the top background colors, frames etc
  + Use right display
* A close up of text on a white background

  Description automatically generated
  + - Lines imply continuous data, bars imply discrete data
  + Use color strategically
  + Tell story with data
* Perceptual Effectiveness
  + Don’t use color intensity to show value
  + Can do absolute (darker color equals one thing, lighter equals another)
  + Don’t use color for quantitative encoding
  + Efficiency
* A close up of text on a white background

  Description automatically generated
* Color effectiveness
  + Don’t use rainbow
  + Color brewer
    - Color scales that are more effective
    - Anytime you need a color scale, go to this tool
    - Can make it colorblind safe

**Show me the data:**

* Weaving data into presentation
* Always think about what the data is showing

**Storytelling and Effective Communication:**

* Two fundamental Questions
  + What is the goal?
    - Predict future data?
    - Explain/understand a phenomenon?
    - Test hypothesis?
    - Compare two groups?
  + Who cares?
* IMAC
  + **I**nferential goal
    - Scientific question of interest
  + **M**odel
    - All models are wrong, some are useful
  + **A**lgorithms
  + **C**onclusions and checking
* Some key principles
  + Remember the golden rule
  + Know your audience
  + Tell a story
    - Need memorable stories and examples
    - Has beginning, middle and end
      * Introduce interesting characters
      * Put them in predicament
      * Resolve the predicament
      * But leave room for sequels! (limitations and future work)
  + Choose and use notation carefully
  + Read great writers
  + Signposting (heres what we’ve done, heres where we are, heres where we’re going)
* Use a lot of visuals
  + Enhance flow
  + Break up the heavy stuff
  + Visually pleasing
* Story telling is crucial
  + Tell audience something they don’t know yet
  + More importantly, tell them something they didn’t know they needed to know
* Key considerations
  + Who is your audience
    - Try to make it accessible for everybody
    - Interest a wider audience and they can find the nitty gritty in the ipython notebooks
    - People you don’t know are difficult to influence
    - What do they know
    - What motivates them? What do they desire?
    - Don’t make them think
      * No jargin
      * Lead them through major steps of project
      * Point out interesting facts using captions and annotations
    - Make the audience aware that there is something they didn’t know they didn’t know
  + What questions are you answering?
    - From mariano rivera example
      * What are the hotspots?
      * What are different approaches for lefties vs righties
      * What makes a dominant closer?
  + Why should the audience care?
    - If you’re putting together a team, its important
    - Historical context
    - “here’s the right way to think about the problem I was trying to solve”
    - Catch audience’s attention using captions and annotations
    - If done well, insights will seem obvious given this framing – and that’s good
  + What are your major insights and surprises?
    - Insights are related to questions
    - Surprises like different types of pitches, hitters picking up diff pitches
  + What change do you want to affect?
    - Not all affect change – some are informational
    - Affecting change engages audience on an emotional level
* Visual story design
  + Beginning middle and end
    - Beginning
      * Start upper left
    - Middle
      * Somewhere off to the right
    - End
      * Below
  + Headline is key
    - Make it the answer to most important question
  + Add annotations for additional details or call-outs
* Successful data stories
  + Target the audience
  + Engage and are memorable
  + Answer concise questions
  + Are carefully designed
  + Move us to want to change the world

**STAR interview method:**

* Straightforward format to use answering behavioral interview questions
* Situation
  + Set the scene and give necessary details of example
* Task
  + Describe what responsibility was in that situation
* Action
  + Explain exactly what steps you took to address it
* Result
  + Share what outcomes your actions achieved

**How to give a killer presentation:**

* Frame the story
  + Need something worth talking about and must conceptualize and frame it to make it engaging
* Plan the delivery
  + Generally shouldn’t be reading a script or off a teleprompter
  + Best way to go is to memorize
    - Need to get out of “valley of awkwardness” so you’re not spending energy trying to remember and audience will notice
    - If you don’t have time to do this, use the bullet points on note cards
  + Typically best to try to sound conversational
  + Best advice before you speak is to breathe deeply
* Using visuals
  + Don’t default to powerpoint
    - use if it’s necessary
* Presentations rise or fall on the quality of the idea, the narrative, and the passion of the speaker

**Avoid “Death by PowerPoint”:**

* Make it clear
  + Outline first
  + One message per slide
  + Control # of slides – 2-3 mins/slide
* Boring ppts will cause people to forget the material
* 5 guiding principles
  + One message
    - Limit to one message per slide
    - People can’t focus on more
  + Working memory
    - Don’t do redundancy effect
  + Size
    - Headline is always biggest and content is smallest even though content is more important
    - Most important part of powerpoint should also be biggest
  + Contrast
    - Use contrast if you want somebody to focus on something (example of the white lettering on focus point and all else greyed out)
    - Dark background – make presenter most contrast-rich object
  + Objects
    - Massive difference between 5 and 7 objects
      * Difference between just seeing and having to count – 500% different
    - **Magical number is 6**
    - 7 or more objects is too many
* Amount of slides has never been the problem
  + Amount of objects/slide is the problem

**Presenting to Executives:**

* Pyramid organizational pattern
  + 1 – Key idea
  + 2 – Supporting Information
  + 3 – Logic behind conclusion
* BLUF: Bottom Line Up Front

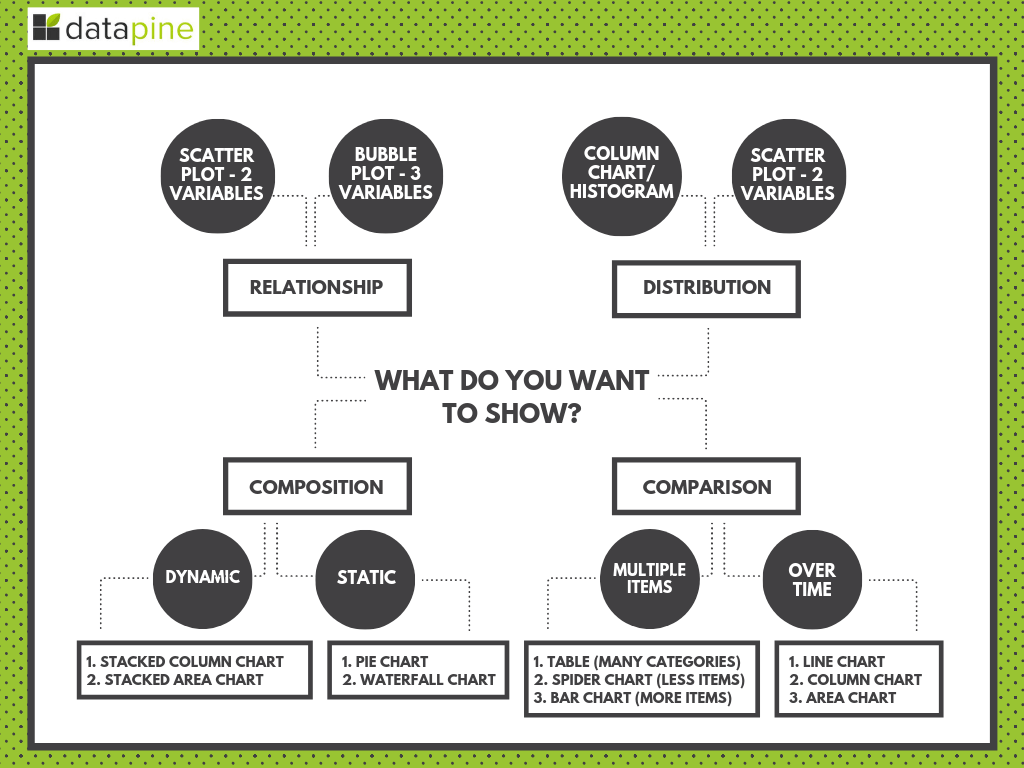
**Presenting Boring Technical information – so it isn’t boring:**

* Not easy to make technical stuff interesting to non-technical people but it can be done – takes passion
* 5 tips:
  + Realize a presentation is about persuasion, not information
  + Ask yourself, what is the problem that the audience has for which my information is the solution?
    - Start with the problem, follow with the solution once they’re interested
  + Don’t give out information, give examples and case studies
  + Use vivid metaphors and analogies
    - If data is abstract, turn into a metaphor, use imagination
  + If all else fails, turn the information into a contest for the audience
    - Jeopardy?
* With creative thought, any topic can become interesting

**Create an effective presentation:**

* Find a story to tell
* Draw audience in quickly
* Explain the threat
* Outline the solution
* Provide an action step

**Choosing the right data visualization types:**

* 
* 7 essential questions:
  + What story do you want to tell?
    - Put it into context that everybody can understand
  + Who do you want to tell it to?
  + Are you looking to analyze particular trends?
  + Do you want to demonstrate the composition of your data?
    - Pie charts
    - Waterfall charts
    - Stacked bars
    - Map based graphs (if geographical)
  + Do you want to compare two or more sets of values?
    - Bubble charts
    - Spider charts
    - Bar charts
    - Columned visualizations
    - Scatter plots
  + Is timeline a factor?
  + How do you want to show your KPIs?
* Most commonly used data visualization types
  + Number chart
    - Essentially just a number with a variance figure
  + Line chart
    - Trending
    - Don’t include too many variables
    - Can be combined with other types of data viz – like bar charts
  + Maps
    - Geographic
  + Waterfall chart
  + Pie charts
  + Gauge charts
    - Like speedometer – assign min and max value and the value will fall somewhere in that range
    - Should only be used for single data points
  + Scatter plots
    - To show correlation in dataset
  + Spider chart
    - Like the NFL combine skills chart – looks like spider web
  + Tables
    - Sometimes data needs to be shown in raw format, but be careful because too much data can be overwhelming
  + Area charts
  + Bubble plots
* Use dynamic text boxes
* Design thinking
  + Biggest keys are
    - Understand reason for visualization
    - Know your audience
    - Know your approaches to get your point across