Data Driven Storytelling

Roshani Abbey

Neueda™

DATA VISUALIZATION SPECIALIST – ROSHANI ABBEY

Roshani Abbey is a self confessed data-nerd, passionate about effective communication and inspirational storytelling. She works in the analytics & reporting space and has served as a data engineer and data analyst, responsible for enabling the exchange of information in an efficient and timely manner.

Prior to working in the financial sector she worked as a research associate at Imperial College London, investigating data science methods such as feature selection and classification using network construction. She is also an actor and appears in theatre shows in London's West End.

Neueda

STORYTELLING FUNDAMENTALS - OVERVIEW

THE ESSENTIAL DATA SCIENCE SKILL EVERYONE NEEDS

Importance of data storytelling

- Communicate
- Building your narrative
- Unlocking insights

The components of an impactful story

- Know your audience – speak their language
- Tell one story at a time
- Simple is better than complex

Visual best practices

- Process & Chart selection
- Checklist
- What not to do & Fallacies

Presenting your story

- Types of stories
- Practice and feedback

Neueda™

IMPORTANCE OF DATA STORYTELLING:

COMMUNICATE

Communication is your #1 job

IMPORTANCE OF DATA STORYTELLING:

BUILDING YOUR NARRATIVE

Relatability

Novelty

Tension

Fluency

IMPORTANCE OF DATA STORYTELLING:

UNLOCKING INSIGHTS

Profit by Sub-Category (Data)

Sub-Category	
Accessories	$41,937
Appliances	$18,138
Art	$6,528
Binders	$30,222
Bookcases	-$3,473
Chairs	$26,590
Copiers	$55,618
Envelopes	$6,964
Fasteners	$950
Furnishings	$13,059
Labels	$5,546
Machines	$3,385
Paper	$34,054
Phones	$44,516
Storage	$21,279
Supplies	-$1,189
Tables	-$17,725

Neueda

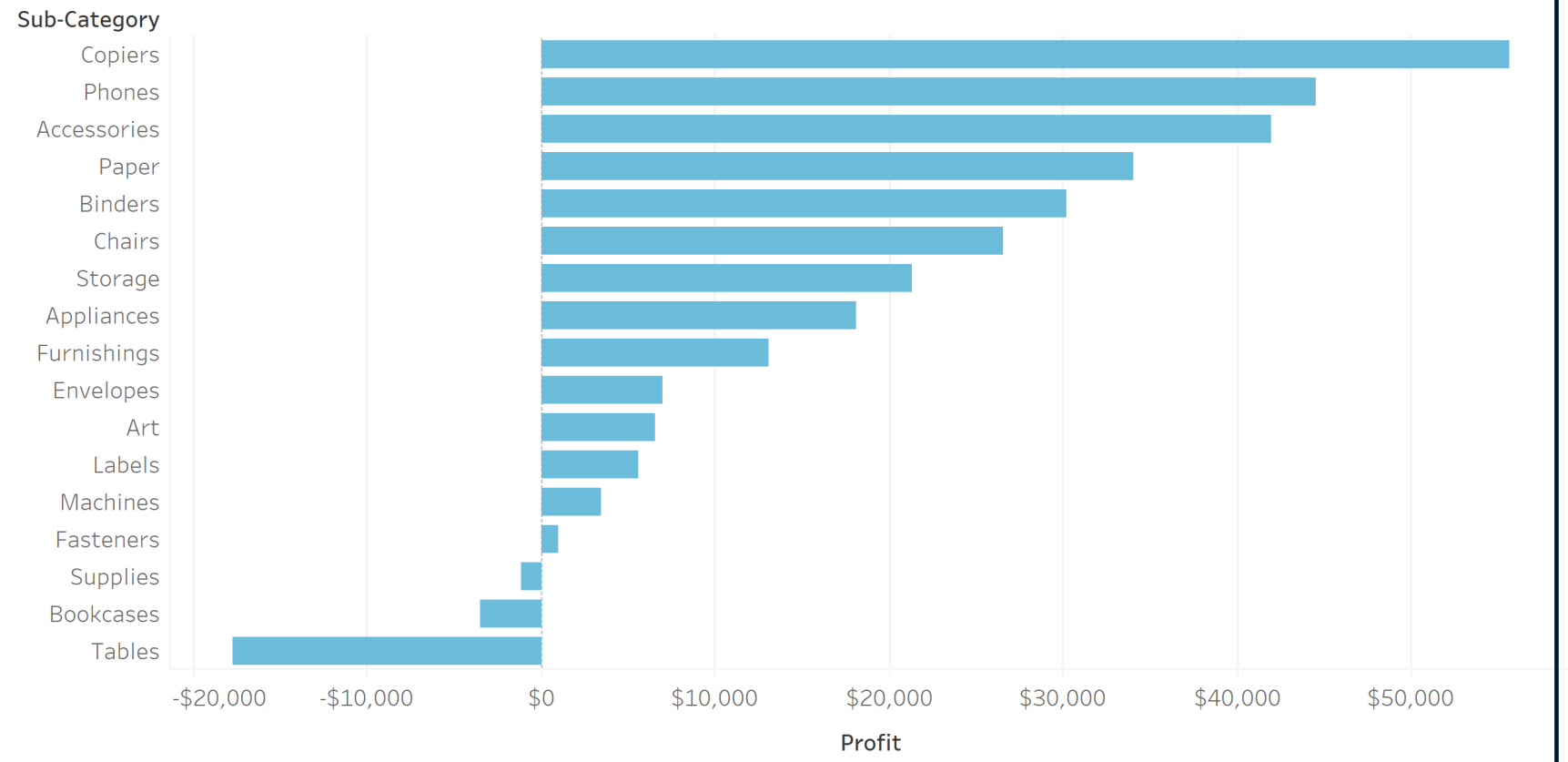
IMPORTANCE OF DATA STORYTELLING:

UNLOCKING INSIGHTS

Profit by Sub-Category (Data)

Sub-Category	
Accessories	$41,937
Appliances	$18,138
Art	$6,528
Binders	$30,222
Bookcases	-$3,473
Chairs	$26,590
Copiers	$55,618
Envelopes	$6,964
Fasteners	$950
Furnishings	$13,059
Labels	$5,546
Machines	$3,385
Paper	$34,054
Phones	$44,516
Storage	$21,279
Supplies	-$1,189
Tables	-$17,725

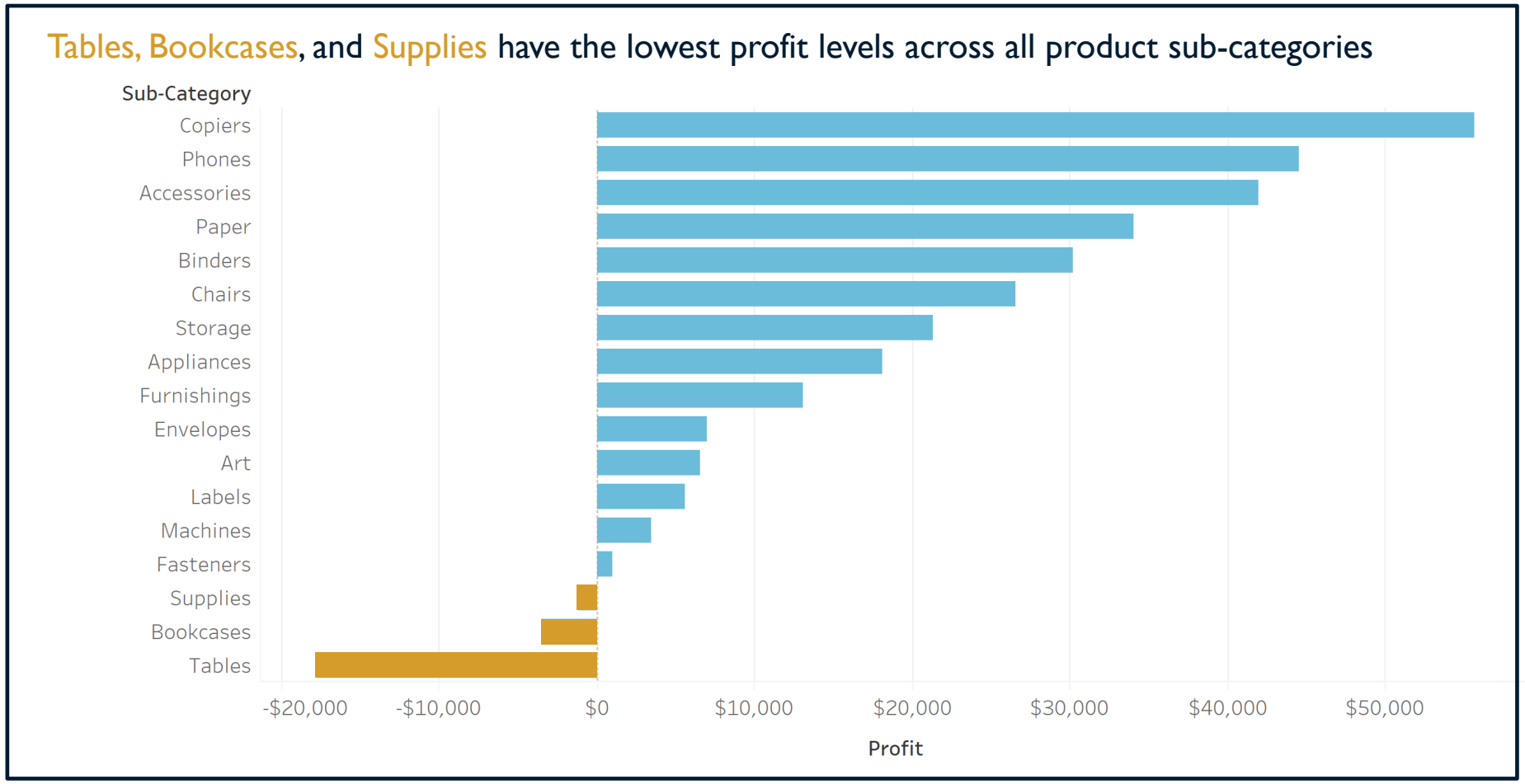
Profit by Sub-Category (Chart)



Neueda

IMPORTANCE OF DATA STORYTELLING:

UNLOCKING INSIGHTS



Neueda

THE COMPONENTS OF AN IMPACTFUL STORY:

KNOW YOUR AUDIENCE

THE COMPONENTS OF AN IMPACTFUL STORY:

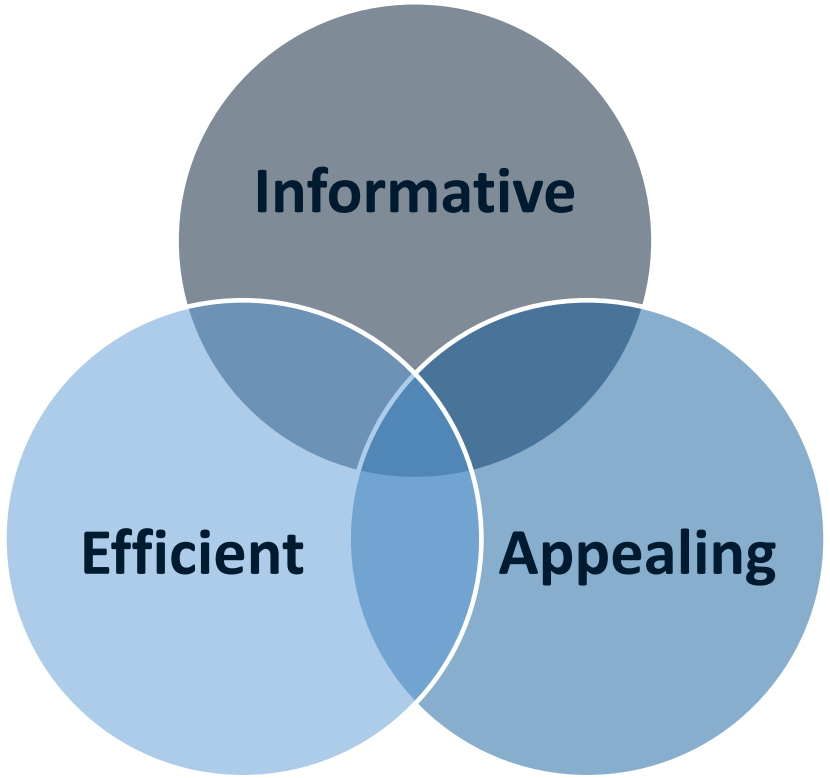
TELL ONE STORY AT A TIME

THE COMPONENTS OF AN IMPACTFUL STORY:

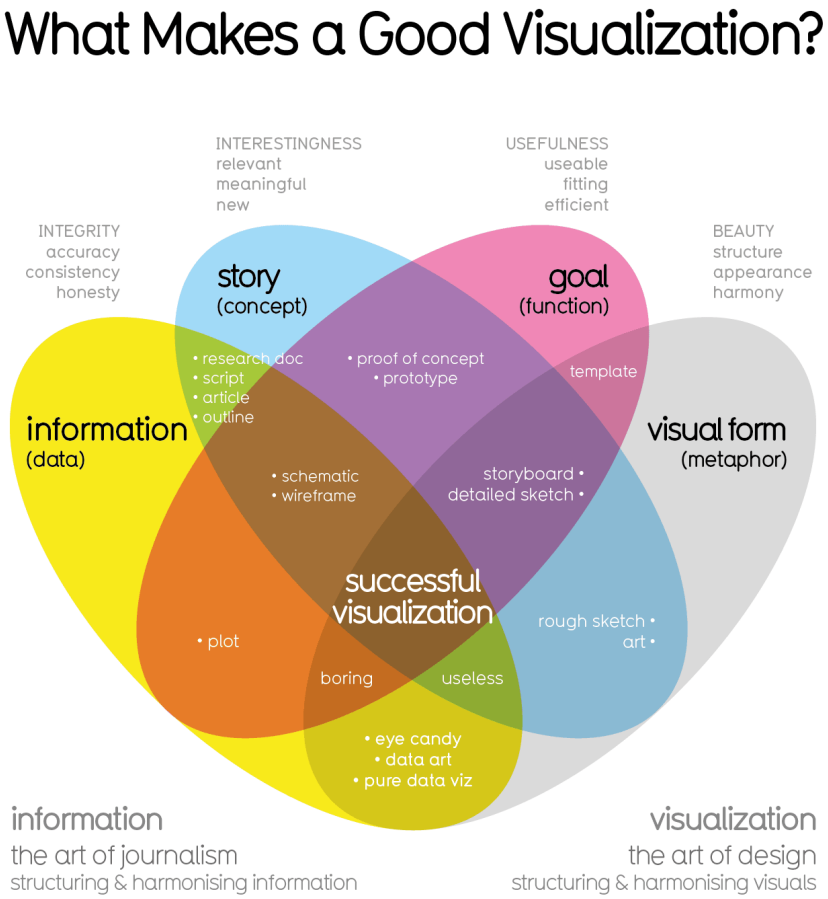
SIMPLE > COMPLEX

Neueda

VISUAL BEST PRACTICES



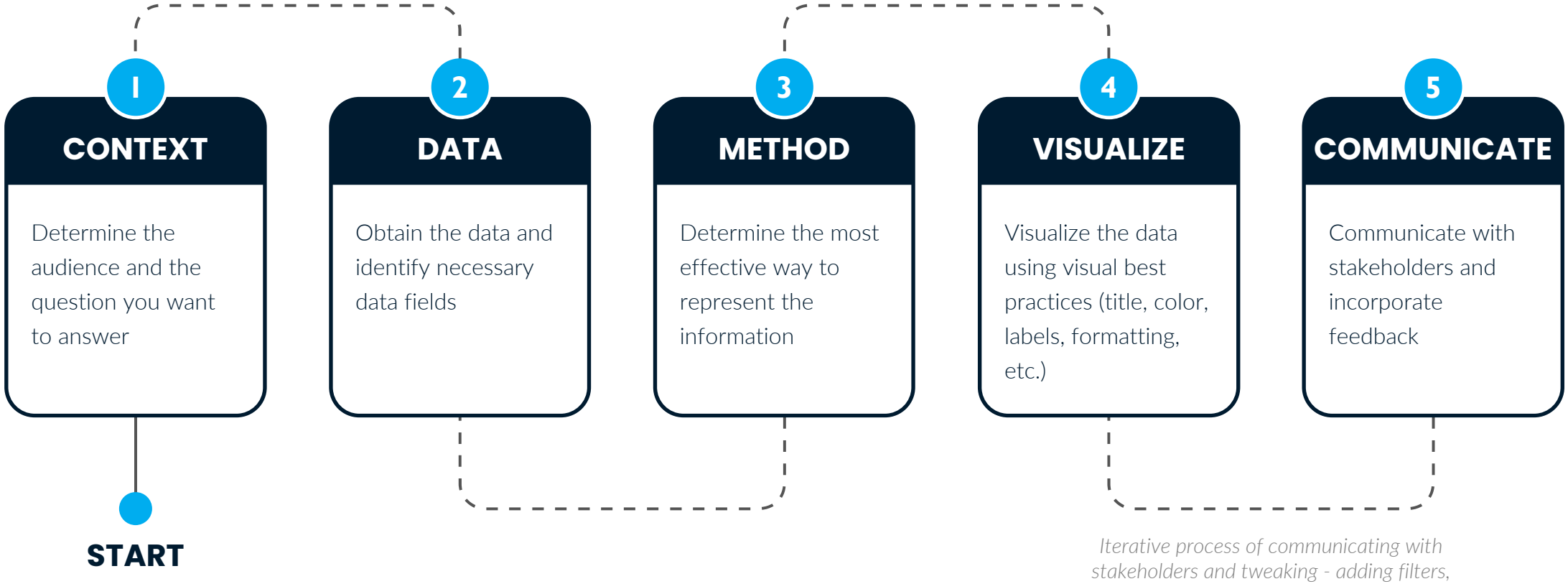
VISUAL BEST PRACTICES



Neueda

VISUAL BEST PRACTICES:

PROCESS



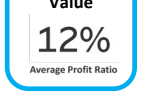
Neueda

VISUAL BEST PRACTICES:

CHART SELECTOR GUIDE

Specific Value

Single Value



Show the raw number prominently displayed

Table

Show the exact values and compare pairs of related values

Highlight Table

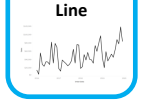
Show the exact values and use color to convey relative magnitude

Heatmap

Compare values by encoding the marks with color and size

Comparison

Single Line

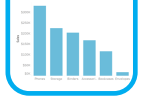


Display trends over a period of time for a single category

Multiple Lines

Display trends over a period of time for multiple categories

Bar Chart



Shows comparisons among discrete categories

Grouped Bar Chart

Shows comparisons among discrete categories and sub-categories

Stacked Bar Chart

Shows comparisons among discrete categories and sub-categories

Dual Axis Chart

Show the relationship between two variables with different magnitudes and scales

Radar

Plots one or more series of values over multiple quantitative variables

Slope Graph

Compares a data point, typically between two points in time

Bullet Graph

Compares data against historical performance or pre-assigned thresholds

Relationship

Scatter Plot

Shows the relationship between two variables

Bubble

Shows relational value without regards to axes

Word Cloud

Shows the relative frequency of words in our data

Composition

Tree Map

Shows a hierarchical part-to-whole relationship

Pie Chart

Shows a part-to-whole relationship

Donut Chart

Shows a part-to-whole relationship

Stacked Bar 100%

Shows a part-to-whole relationship across categories

Stacked Area

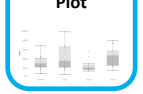
Shows a part-to-whole relationship over a period of time

Waterfall Chart

Shows how a value changes by various factors that either increase the value, or decrease it

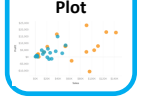
Distribution

Box & Whisker Plot



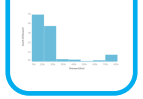
Displays the data distribution through quartiles

Scatter Plot



Shows the relationship between two variables

Histogram



Show the underlying shape of a set of continuous data

Geographic

Filled Map

Shows geographic data using shading on a country or state basis to indicate relationships

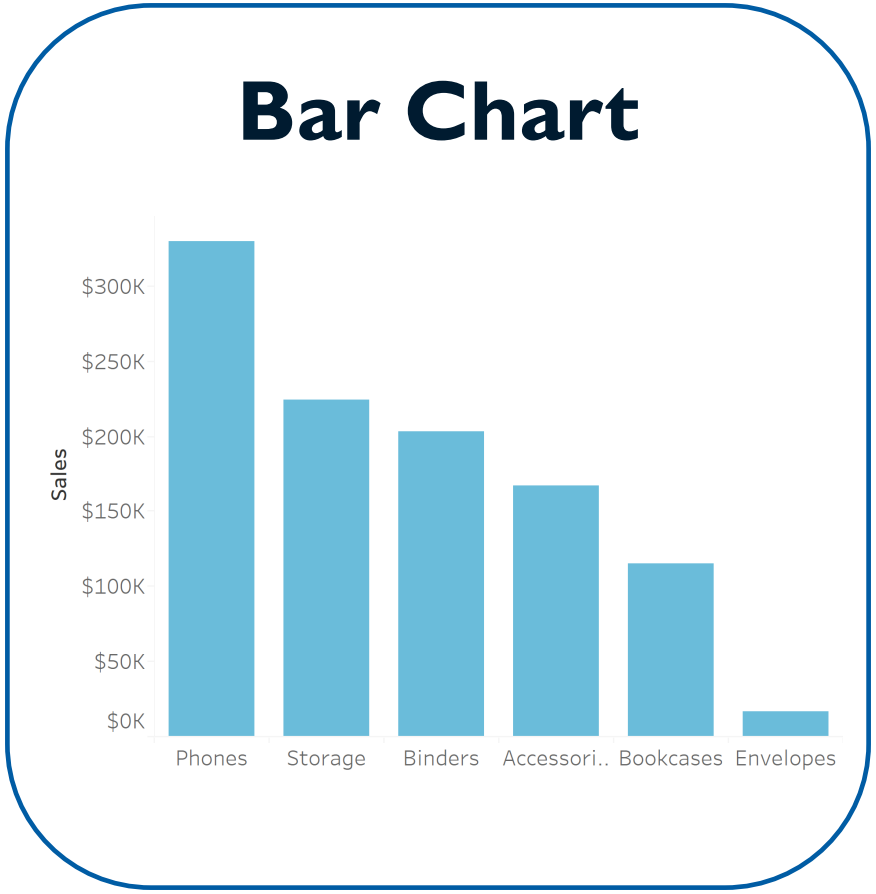
Symbols Map

Shows geographic data using a symbol plotted over a longitude and latitude

Neueda

VISUAL BEST PRACTICES:

CHART SELECTOR GUIDE EXAMPLE #1



Best Practices

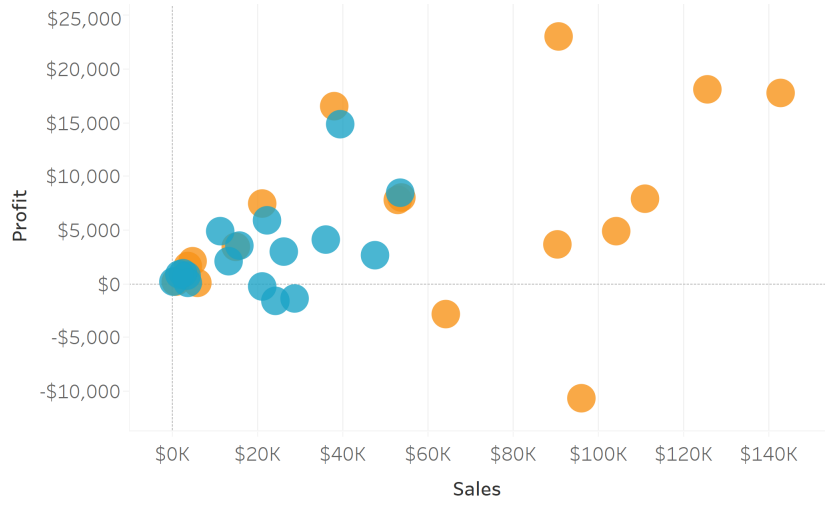
- ✓ Sort your data either from greatest to smallest or the other way around - to help your audience spot the differences.
- ✓ Avoiding using different colors for each bar as it will confuse your audience
- ✓ Use a horizontal bar chart for displaying many categories or if you're using lengthy data labels.
- ✓ Sometimes it makes sense to add data labels directly on the bars
- ✓ Maintain a zero baseline.
- ✓ Ensure there is enough space between the bars.
- ✓ The bar thickness needs to be appropriate for the visualization.
- ✓ Avoid 3D bar charts because they can distort the perception of the data.

Neueda™

VISUAL BEST PRACTICES:

CHART SELECTOR GUIDE EXAMPLE #2

Scatter Plot



Best Practices

- ✓ Formatting the marks in the view can really help in making your scatter plots stand out. You can use color to identify different categories in the view.
- ✓ Make the mark a bit transparent - this allows the audience to see what's behind the marks.
- ✓ Reduce overlapping marks and see more individual plots.
- ✓ Independent variables need to be placed on the x-axis and dependent variables on the y axis.
- ✓ Consider including a trend line to define the correlation.
- ✓ Use custom shapes in place of the typical circles that you are likely familiar with, as applicable

Neueda™

VISUAL BEST PRACTICES:

CHECKLIST

GENERAL

- ❑ **Chart type** is appropriate for the data – e.g. use a line chart for demonstrating trends over time
- ❑ Objects **work together** to **clearly highlight a finding** or takeaway message

LAYOUT

- ❑ The **number of charts** in the view is limited to four
- ❑ **Proper use of real estate;** the more important elements in the view should take up more space
- ❑ Data is displayed in a **logical order** (e.g. chronological, magnitude, etc.)
- ❑ Ensure proper **sizing of elements** in the visualization (avoid scroll-bars, or scrunched up charts)

TEXT

- ❑ Short and **descriptive title** in the upper-left corner
- ❑ Include **clear labels** throughout the visualization
- ❑ **Annotations** highlight specific data points, as needed
- ❑ All **text is horizontal**
- ❑ **Font type and size** is consistent and legible

FORMAT

- ❑ **Gridlines** are not present or muted (light gray)
- ❑ Redundant **borders** are removed
- ❑ Values are formatted to a suitable **level of precision** (round up) and apply applicable display units (e.g. thousands, **currency symbols**)
- ❑ **Graphs are 2 dimensional;** avoid 3D or shape bevels
- ❑ **Data legends** (color, size, or shape) are positioned near the relevant data and used sparingly

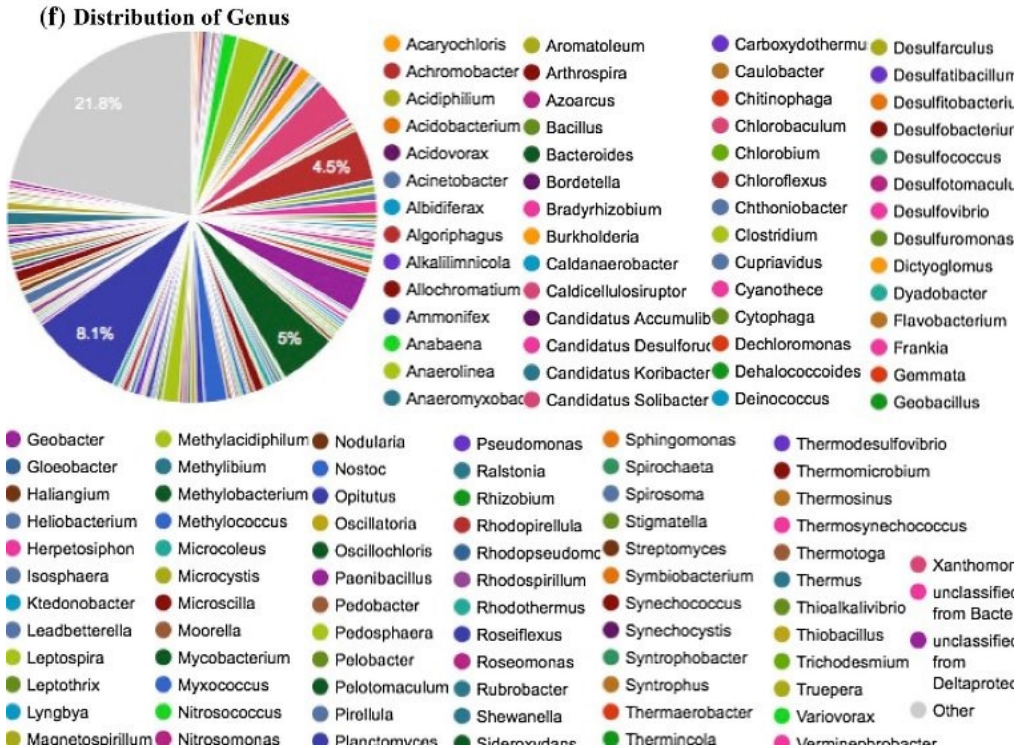
COLOR

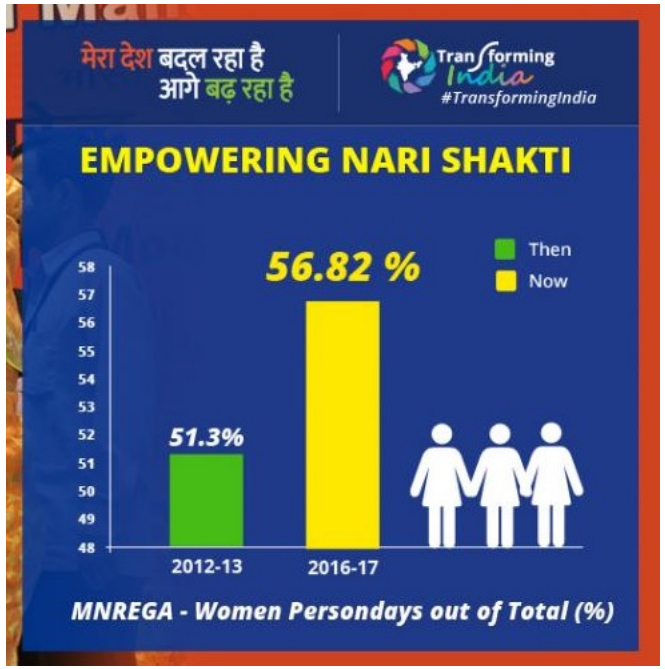
- ❑ Color is used to **highlight key patterns** and guides the viewer
- ❑ **Supporting data is muted** (light gray)
- ❑ Patterns are still viewable when **printed in black and white**
- ❑ Color is legible for people with **colorblindness**
- ❑ **Color scheme** is intentional (e.g. in line with brand)
- ❑ Not more than **5 colors** are used in the visualization
- ❑ **Consistent color** is used for same variables

Neueda™

VISUAL BEST PRACTICES:

WHAT NOT TO DO

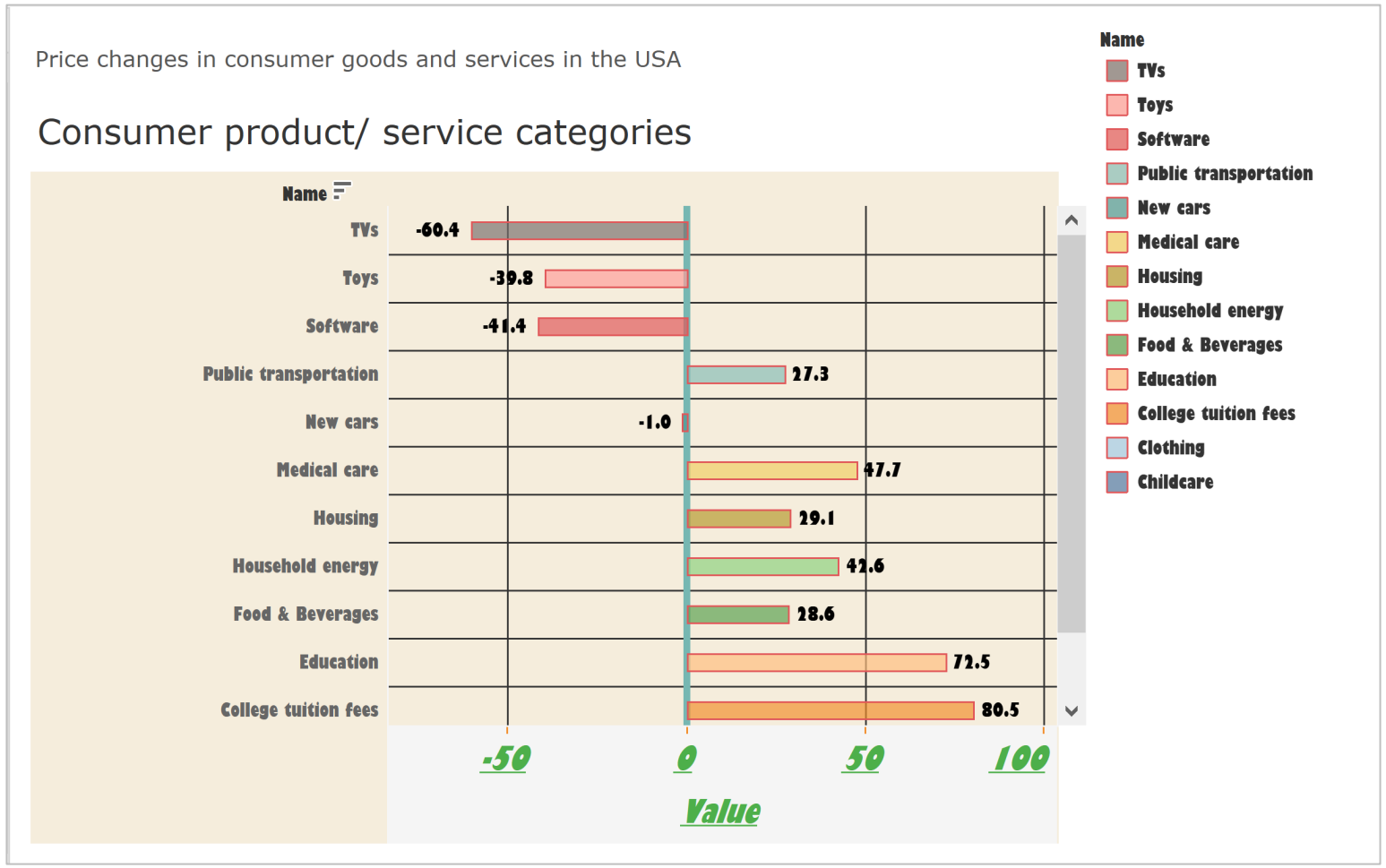




Neueda

VISUAL BEST PRACTICES:

BEFORE



Neueda

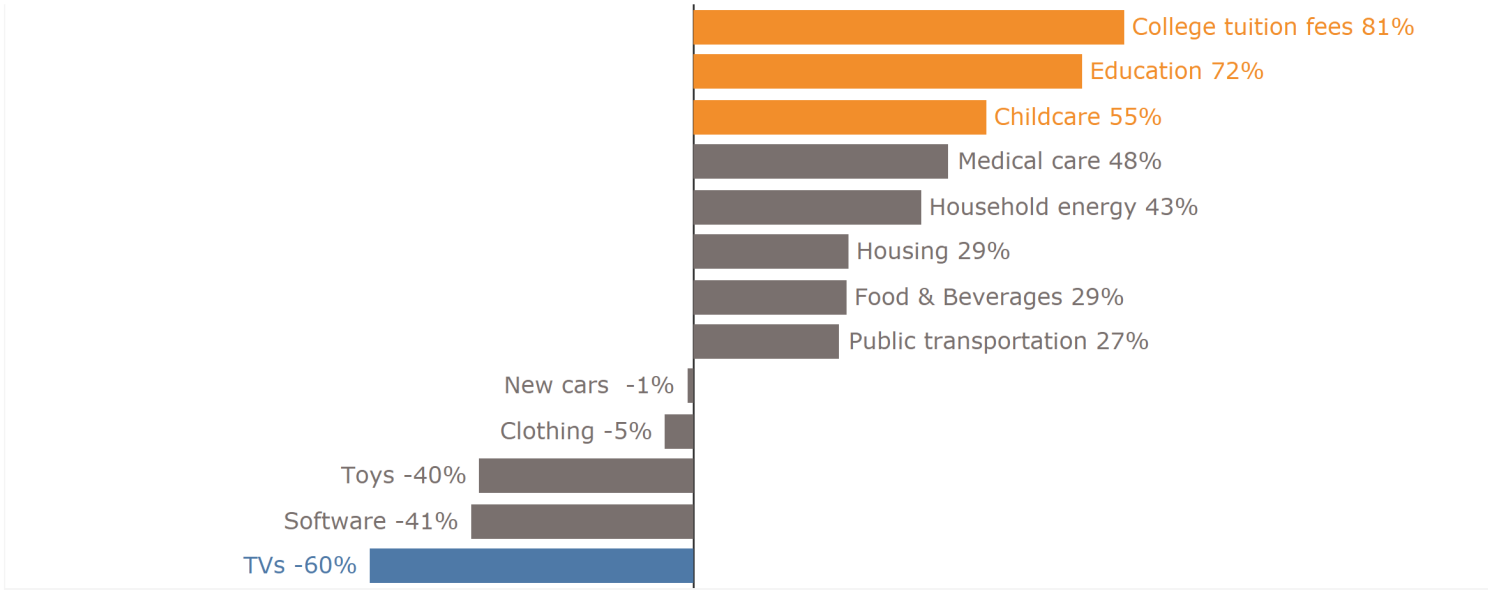
VISUAL BEST PRACTICES:

AFTER

Price changes in consumer goods and services in the USA

Price change is measured as the average percentage change since 1997 - 2017

Consumer product/ service categories that **increased in price by over 50%** and **decreased in price by over 50%**

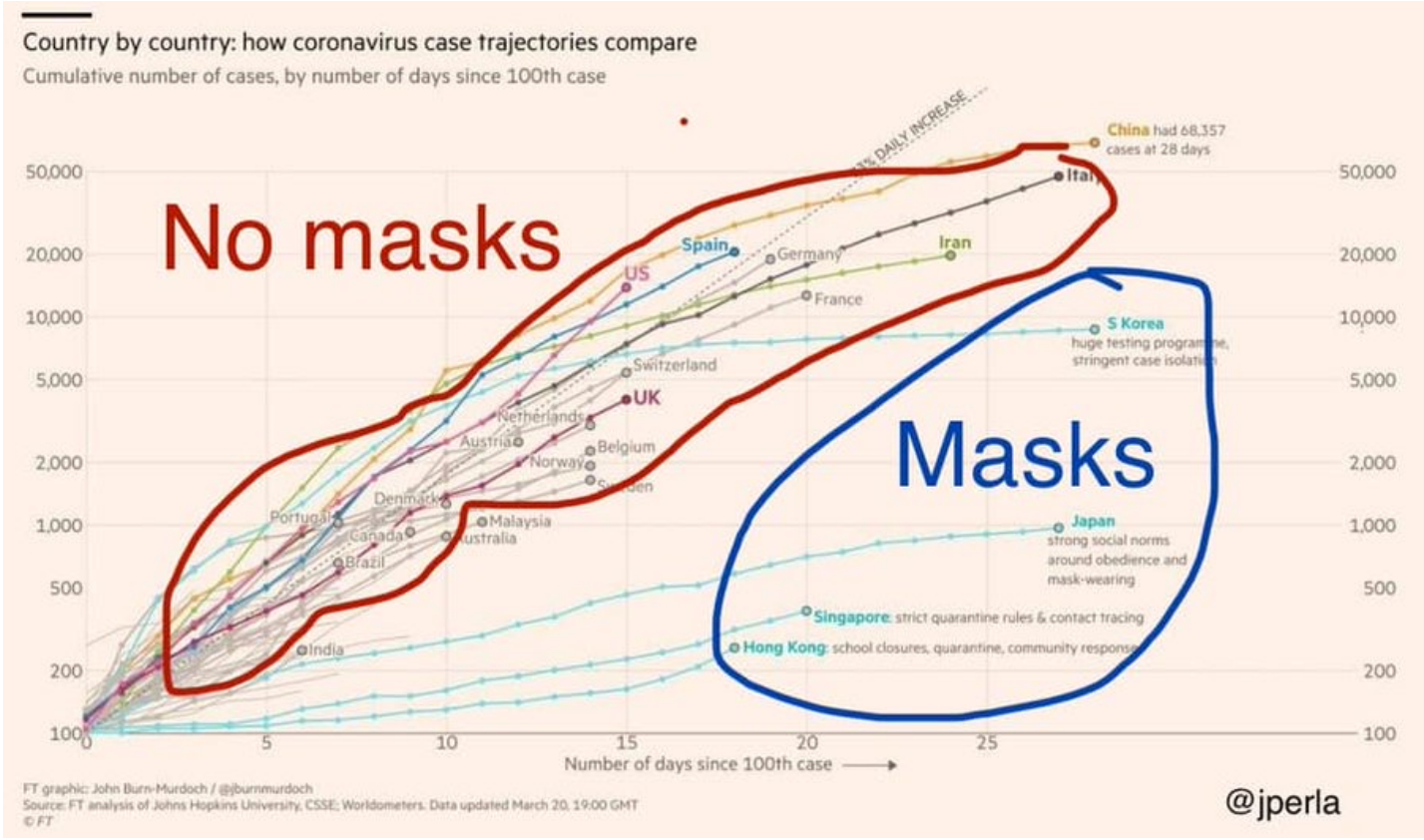


Source: https://ourworldindata.org/grapher/price-changes-in-consumer-goods-and-services-in-the-usa-1997-2017

Neueda

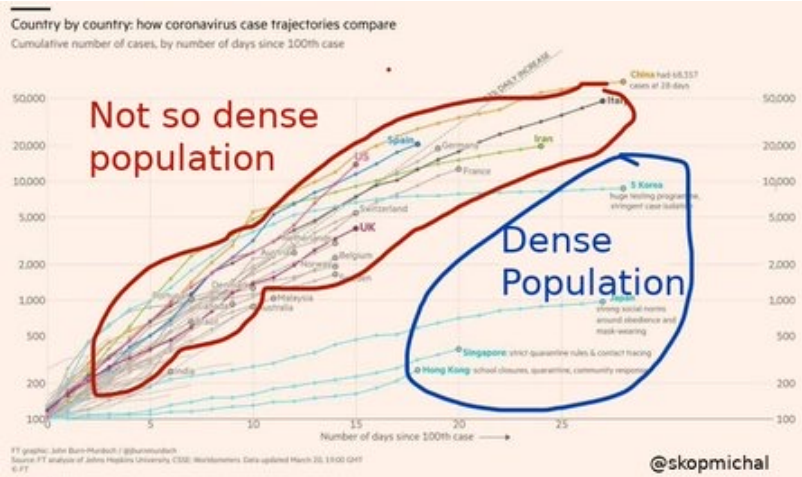
VISUAL BEST PRACTICES:

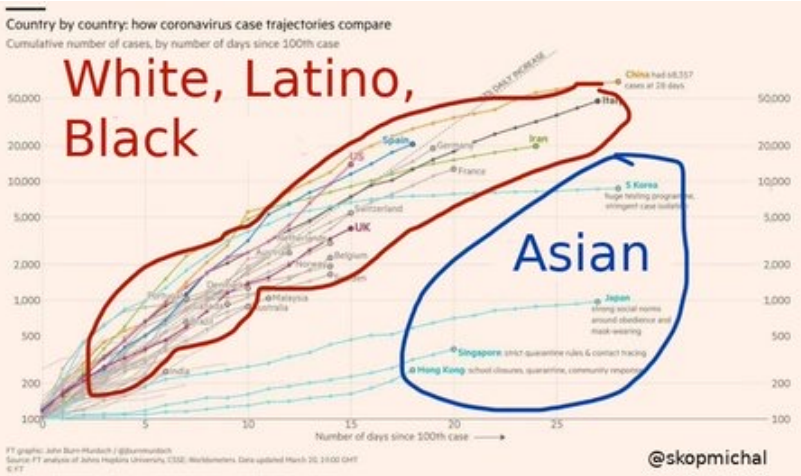
FALLACIES (CORRELATION)

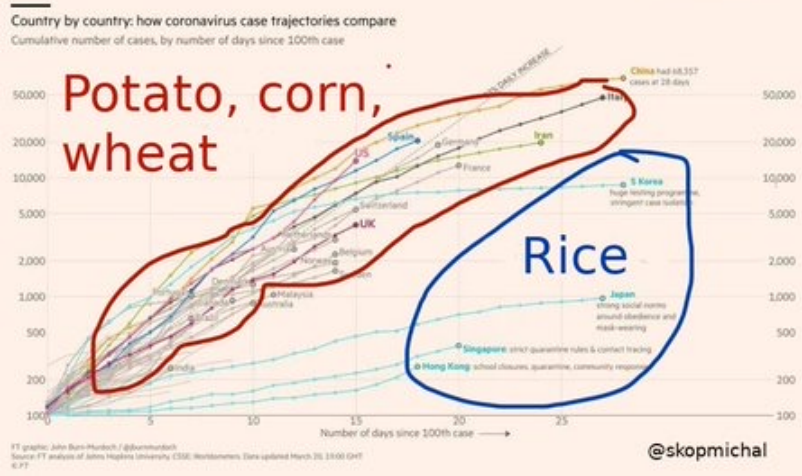


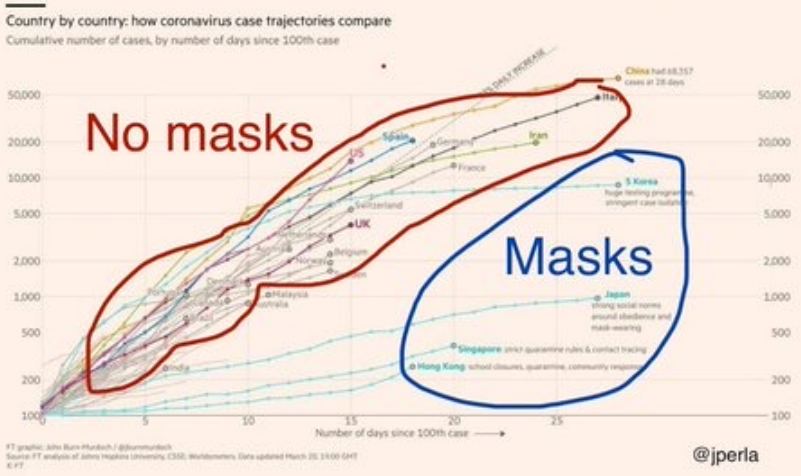
VISUAL BEST PRACTICES:

FALLACIES (CORRELATION)





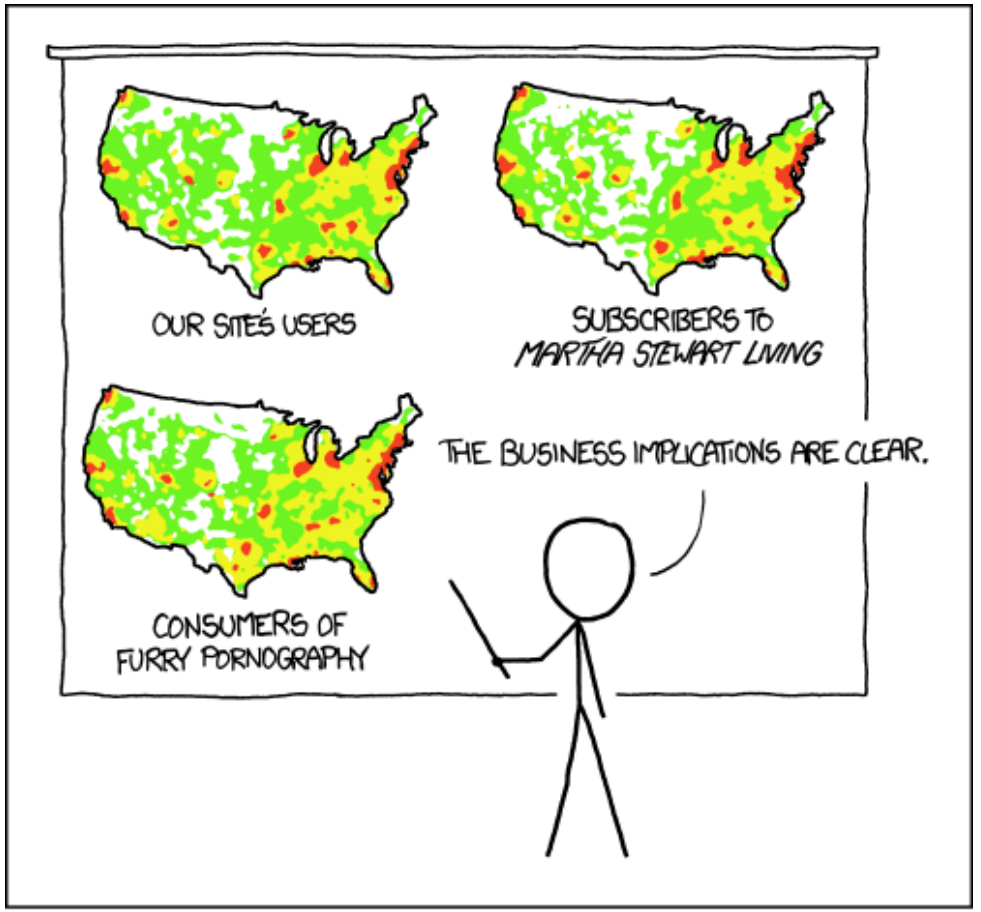




Neueda

VISUAL BEST PRACTICES:

FALLACIES (POPULATION BIAS)

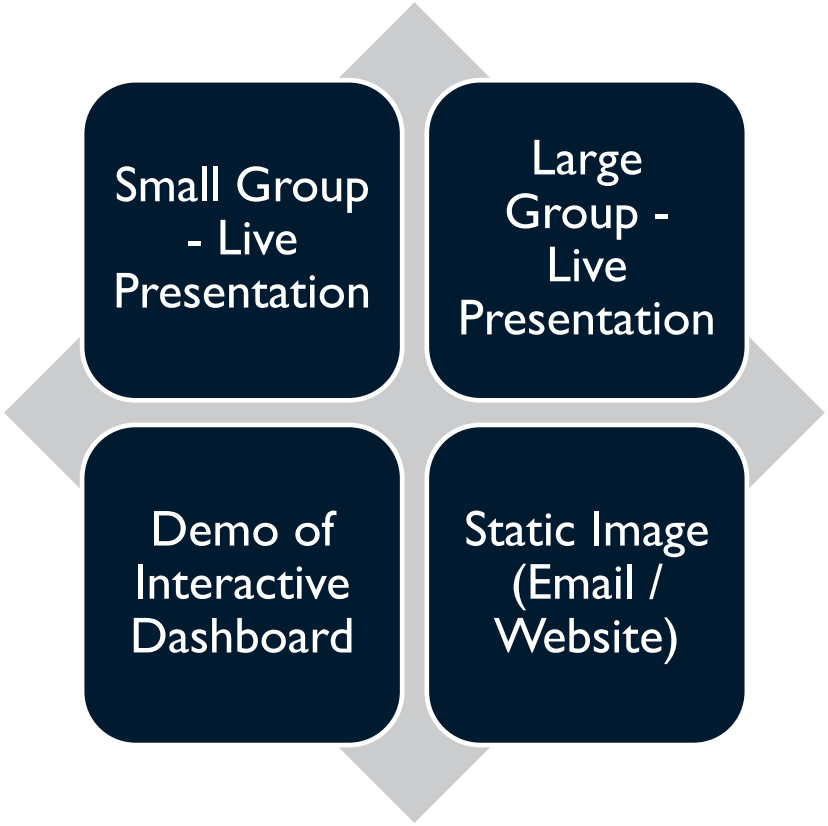


PET PEEVE #208:
GEOGRAPHIC PROFILE MAPS WHICH ARE BASICALLY JUST POPULATION MAPS

Neueda

PRESENTING YOUR STORY:

TYPES OF STORIES



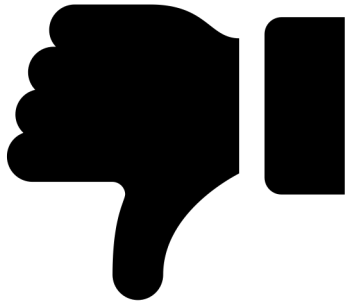
PRESENTING YOUR STORY:

TEST YOUR VISUALS

- ☐ Will the audience understand the goal of the visualization in less than 30 seconds?
- ☐ Is the language used appropriate for your audience?
- ☐ Does the flow and sequence of the objects support the goal?
- ☐ Are the included objects the best choices for the goal?
- ☐ Is it clear on how to interact with the visualization (if it's interactive)?

PRESENTING YOUR STORY:

POWER OF FEEDBACK



DATA SCIENCE:

PORTFOLIO

Best Practices	Data Analysis and Visualization	Machine Learning and AI
▪ Data Driven Story Telling ▪ Exploratory Data Analysis ▪ Using Color Intentionally in Data Visualization ▪ Deep Dive on Chart Selection ▪ Formatting Visualizations for Greatest Impact	▪ Analysis and Charting with Excel ▪ From Data to the Dashboard with Tableau ▪ From Data to the Dashboard with Power BI ▪ Financial Data Analysis Python & Pandas ▪ Statistical Analysis with R	▪ AI Primer for Markets Professionals ▪ Back-testing Trading Strategies ▪ Back-testing Risk Models ▪ Useful Machine Learning Algorithms ▪ Reinforcement Learning ▪ Linear and Logistic Regression ▪ Accelerated Data Classification ▪ Neural Networks ▪ Predictive Analytics ▪ Sentiment Analysis

DATA SCIENCE:

RESOURCES

Tableau Website

- Tableau instruction videos
- Data Viz tips: https://www.tableau.com/en-gb/learn/articles/data-visualization-tips

Code Academy

- Python & R for Data Science

More data visualization principles - Github

- https://rafalab.github.io/dsbook/data-visualization-principles.html

Other Data Visualisation tools

- Data Wrapper
- Flourish Studio – for slick animations