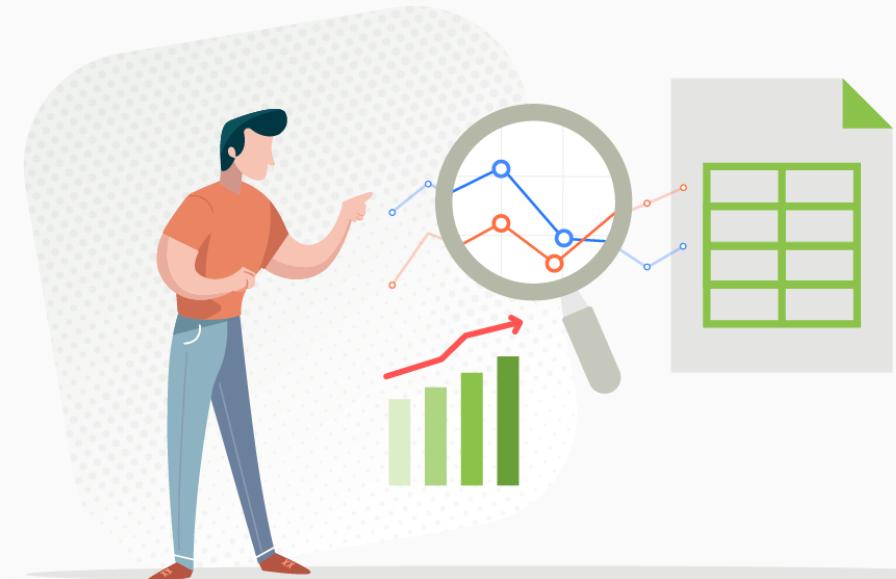


Hands-On Foundational Data Science Bootcamp

Day 2



Instructor
Bastin Robins .J
Chief Data Scientist

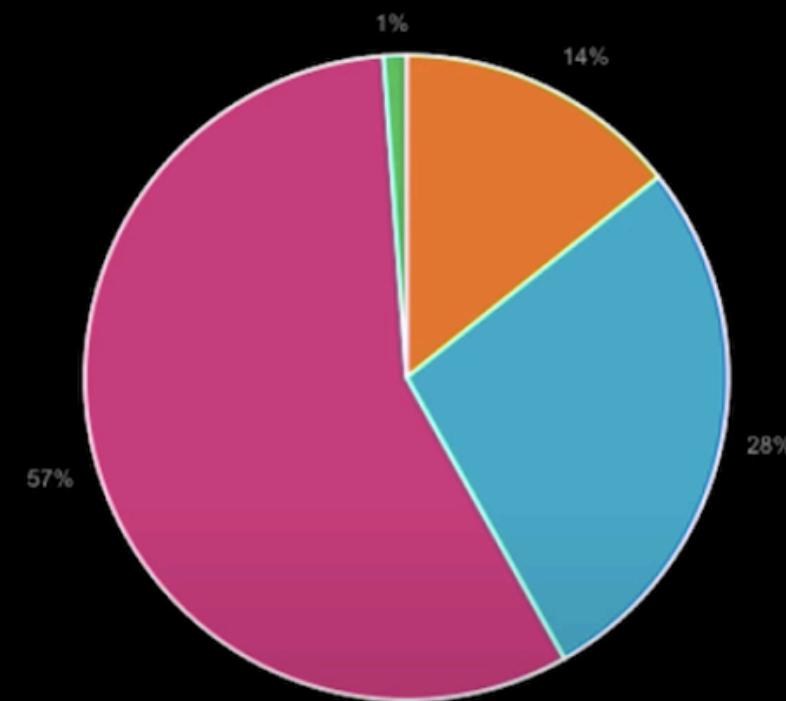


- Data Visualization & Grammar of Graphics
- Data Wrangling & Cleaning
- Exploratory Data Analysis
- Break -
- Machine Learning Models
- How to pick right model

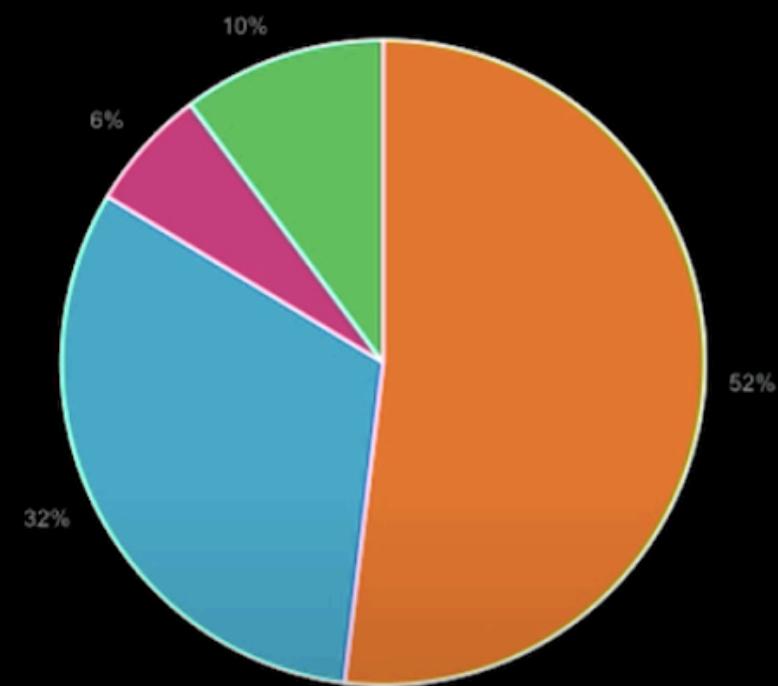
- Pandas – Data Wrangling
- Numpy - Vector
- Sklearn – Model
- Seaborn – Data Visualization

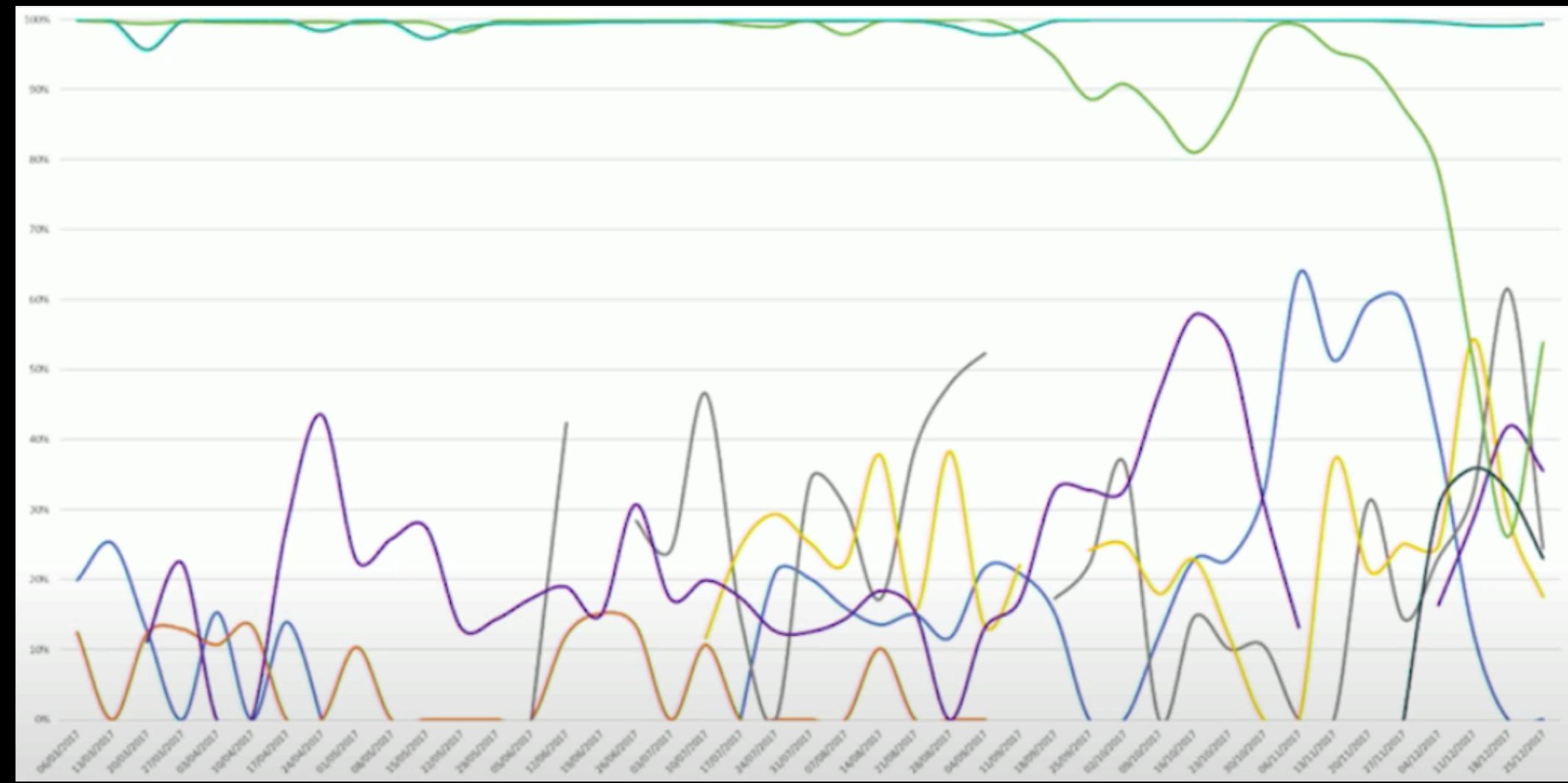
Grammar Of Graphics

2017



2018

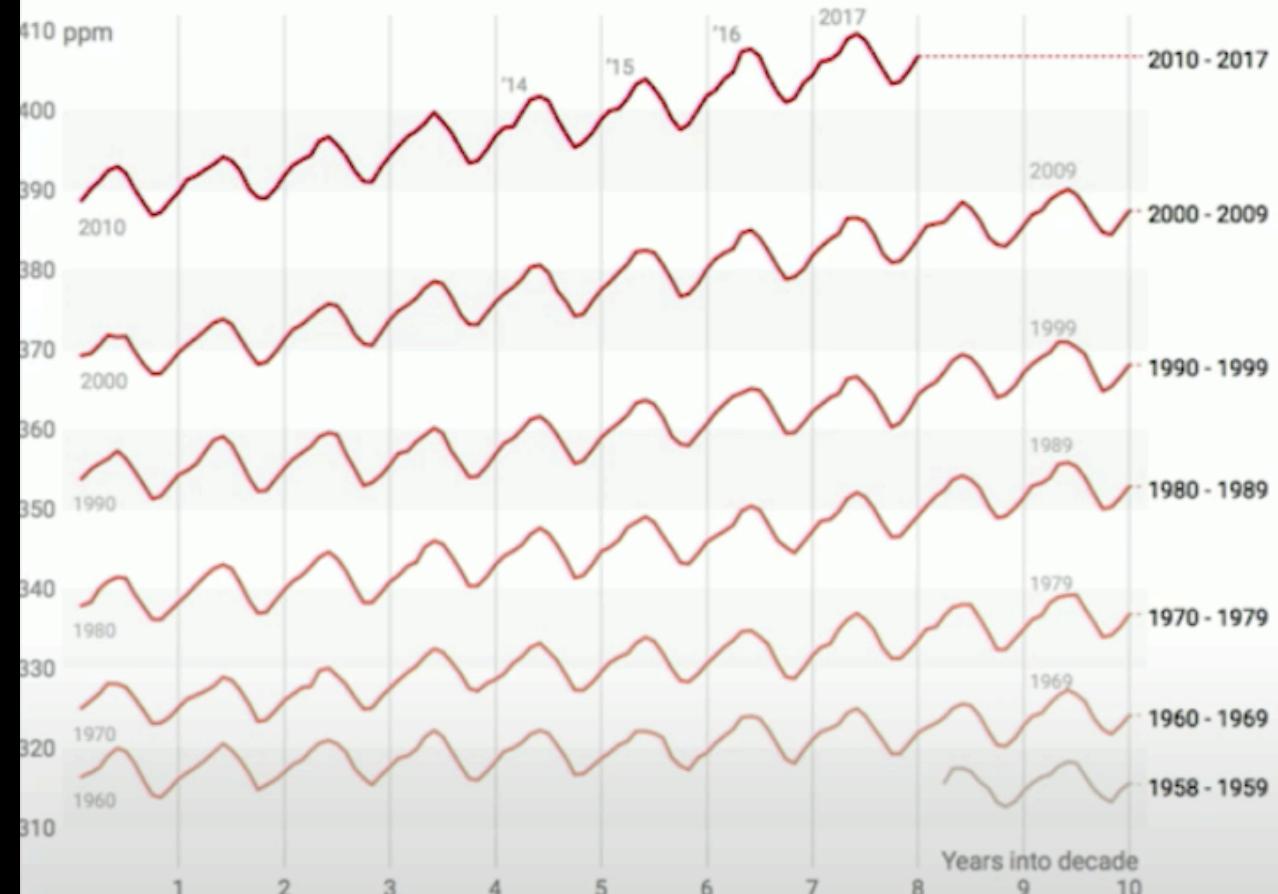




Carbon Dioxide Concentration By Decade

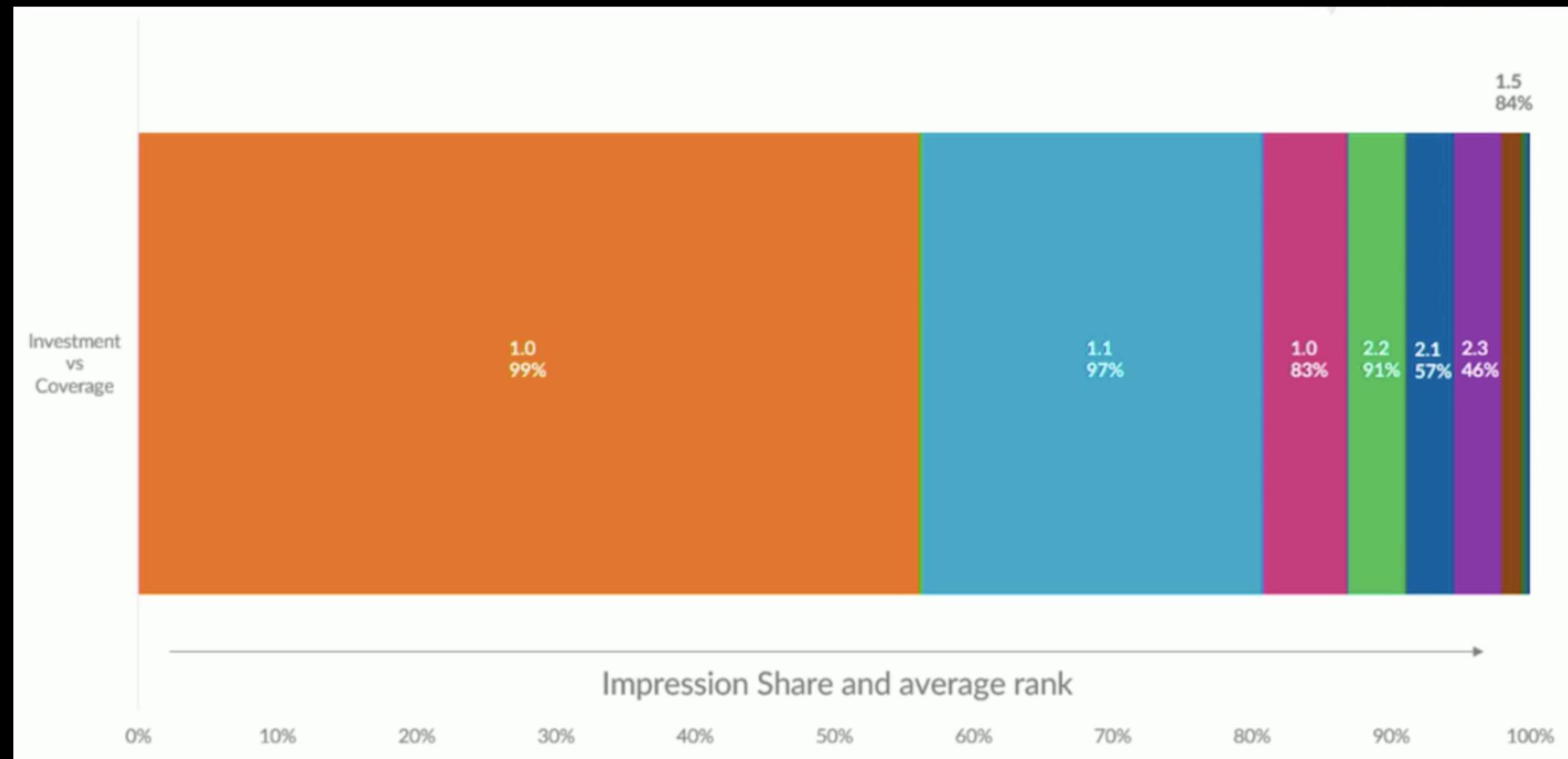
CO₂ concentration in parts per million* (ppm)

Each line represents one decade, from 1958 to 2017.



*The mole fraction of CO₂, expressed as parts per million (ppm) is the number of molecules of CO₂ in every million molecules of dried air (water vapor removed).

Source: National Oceanic & Atmospheric Adm. (NOAA) • Get the data • Created with Datawrapper



Anatomy of a Winning TED Talk

● 1%

Sophisticated Visual Aids

We're not sure who puts the D in TED—most of the best presentations favor tepid PowerPoint slide shows (sorry, Brené Brown). Pictionary-quality drawings (really, Simon Sinek?), or no props at all.

● 5%

Opening Joke

Remember the one about the shoe salesmen who went to Africa in the 1900s? That's how Benjamin Zander opened his talk—which turned out to be about classical music.

● 5%

Spontaneous Moment

Don't overprepare. Tease the guy in the front row ("You could light up a village with this guy's eyes"). Command the stagehand who handles the human brain you brought.

● 5%

Statement of Utter Certainty

People come for answers—give 'em what they want, as Shawn Anchor did: "By training your brain ... we can reverse the formula for happiness and success."

● 12%

Snappy Refrain

The TED equivalent of "I have a dream." Example: "People don't buy what you do; they buy why you do it." Repeat 7x.

● 23%

Personal Failure

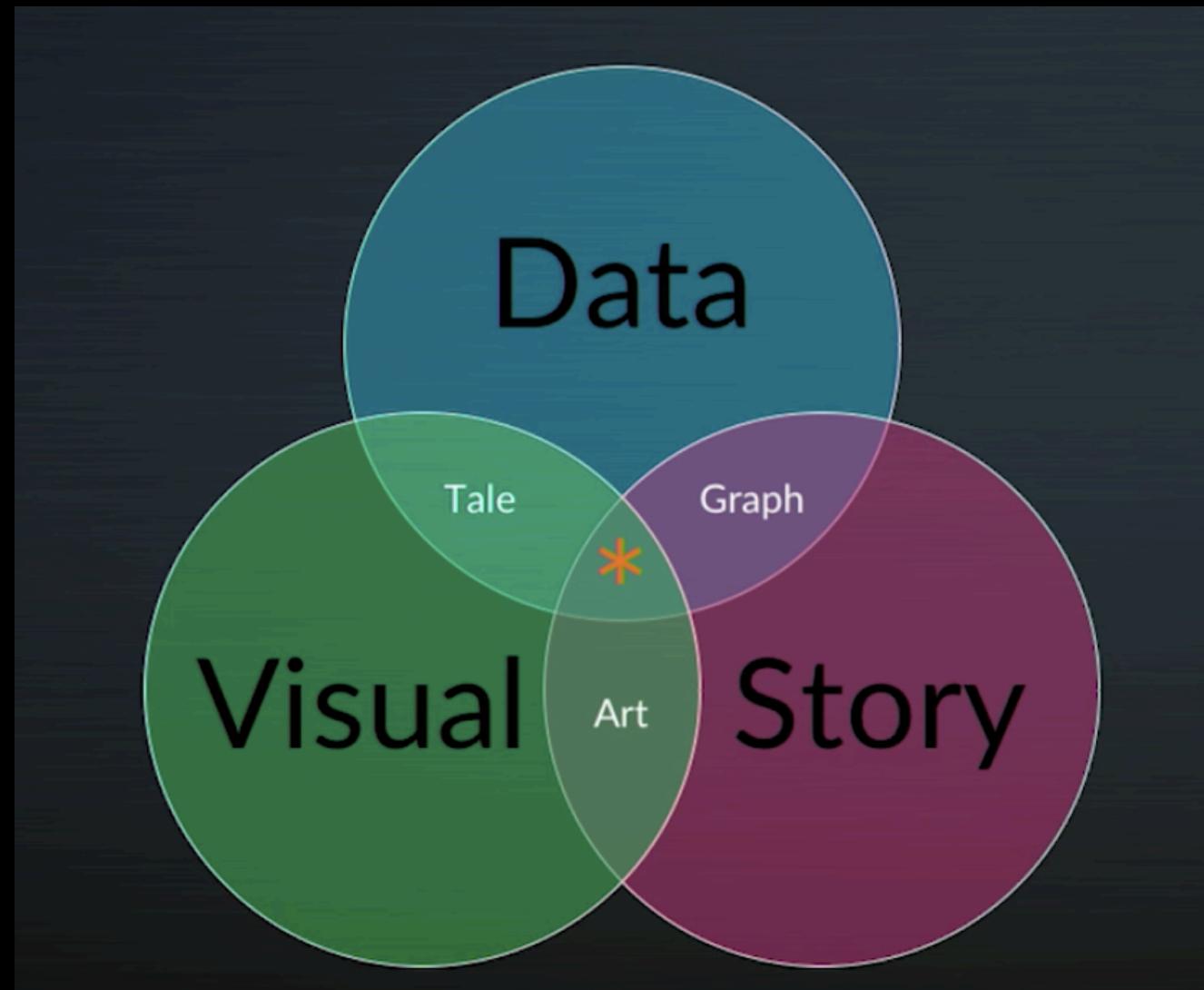
Be relatable. We want to know about that nervous breakdown. Or at least the time you didn't fit in at summer camp.

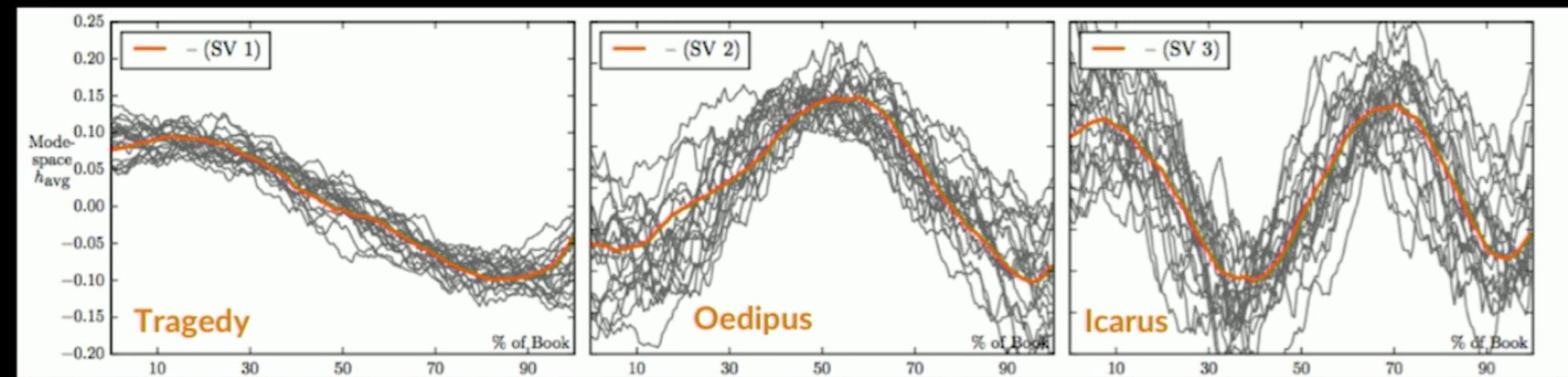
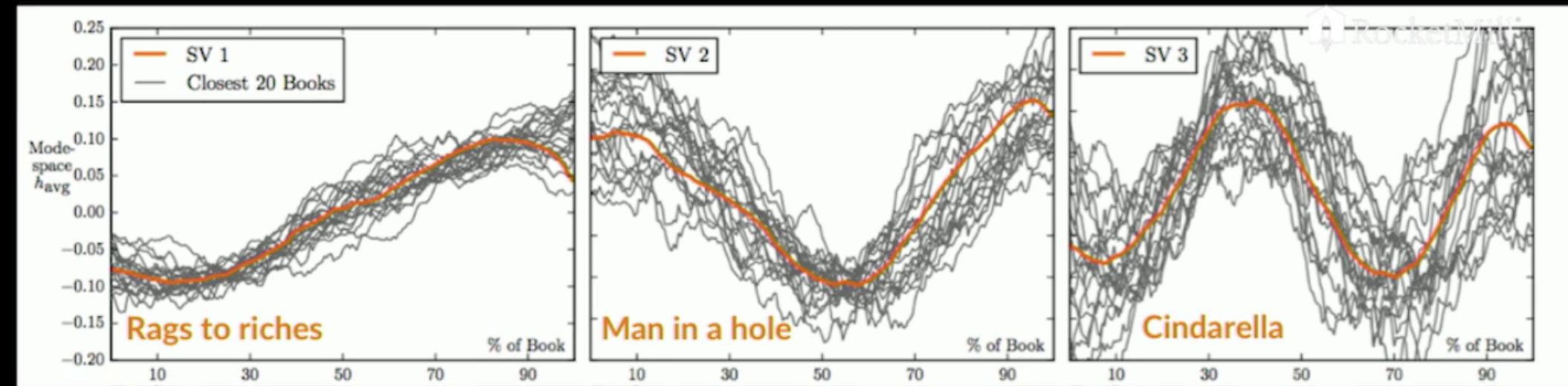
● 49%

Contrarian Thesis

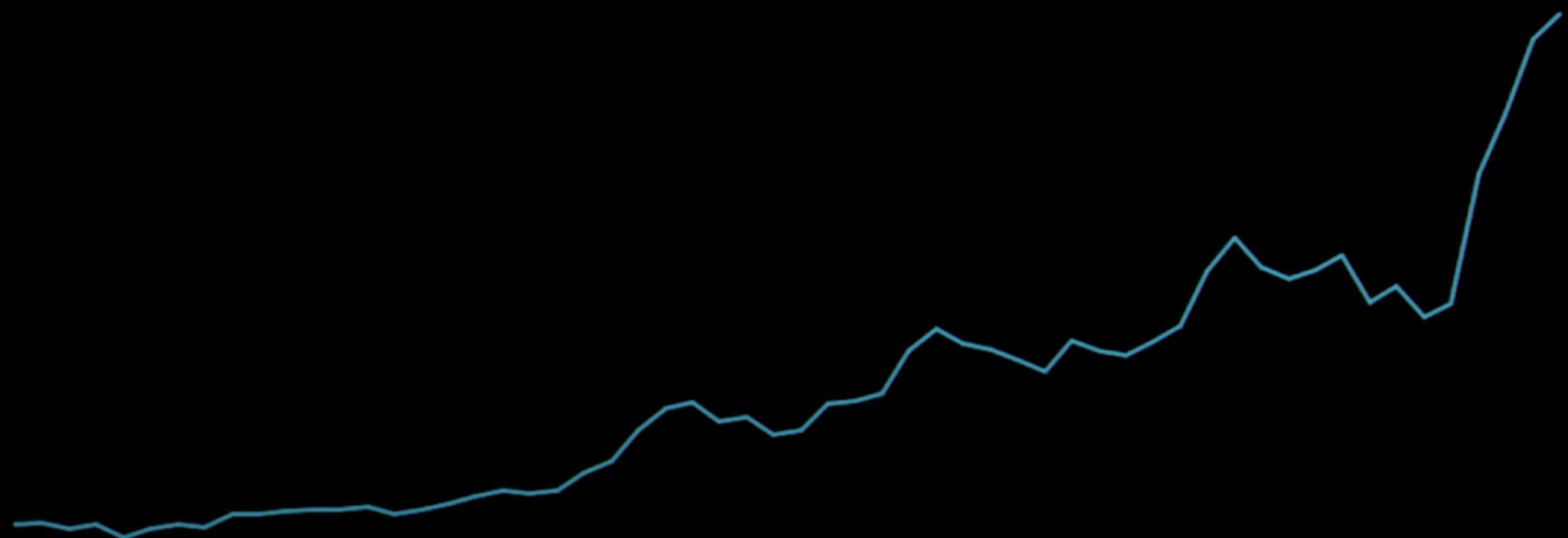
Wait a sec—we should be playing more videogames? The more choices we have, the worse off we are? TED is where conventional wisdom goes to die.



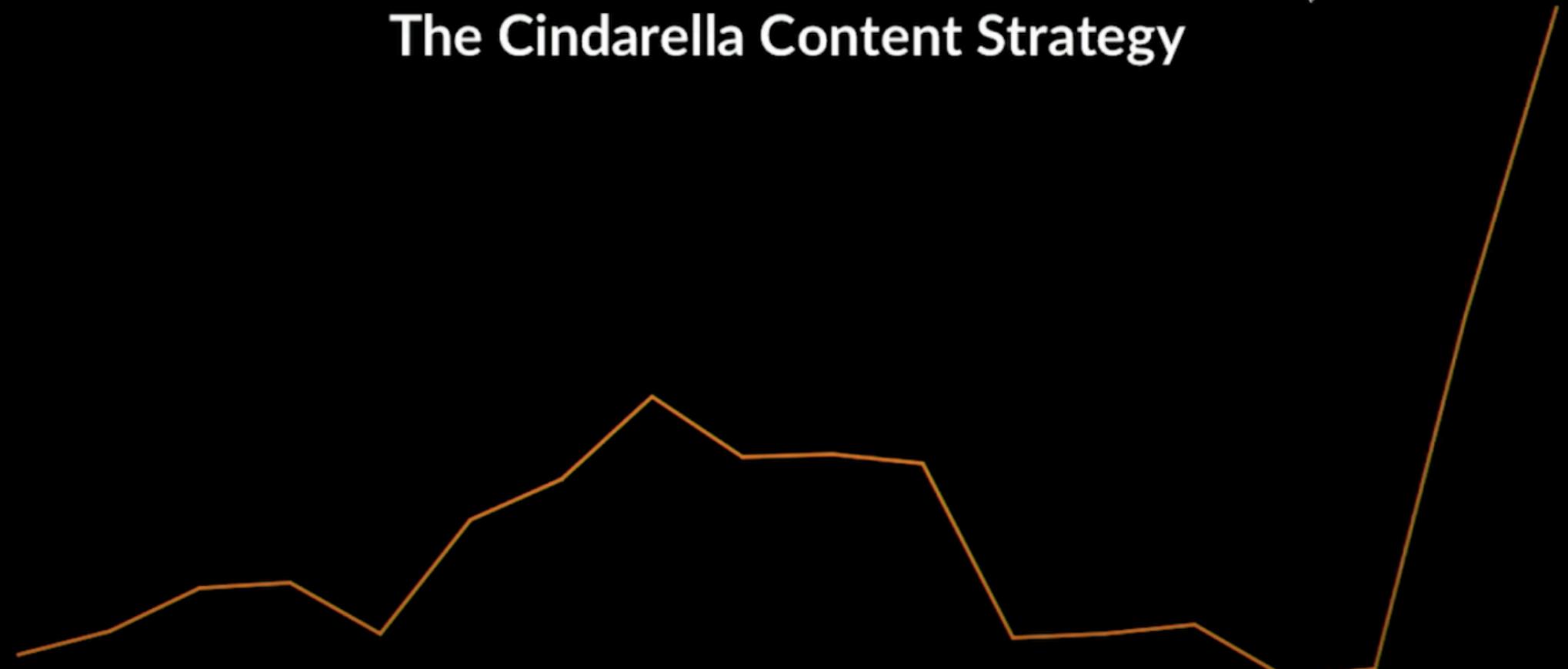




Rags to Riches – The tale of a winning SEO strategy

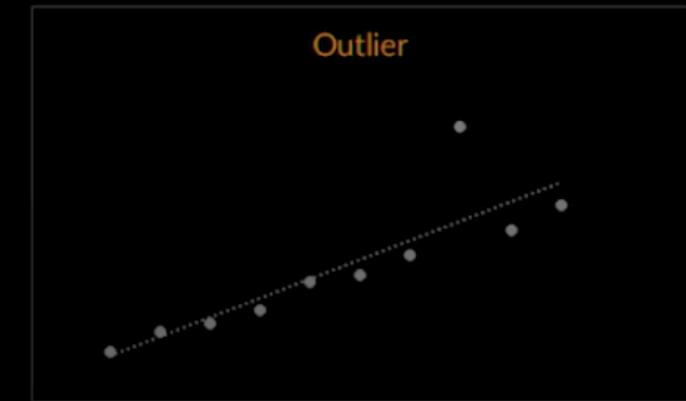


The Cindarella Content Strategy



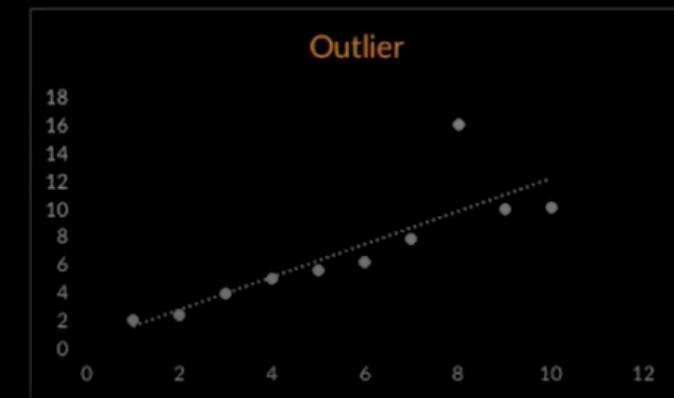
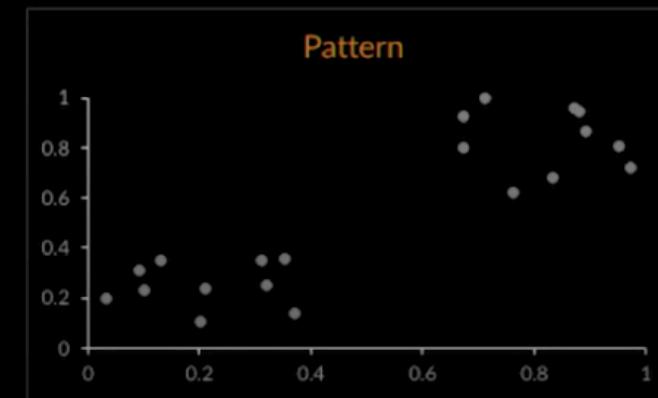
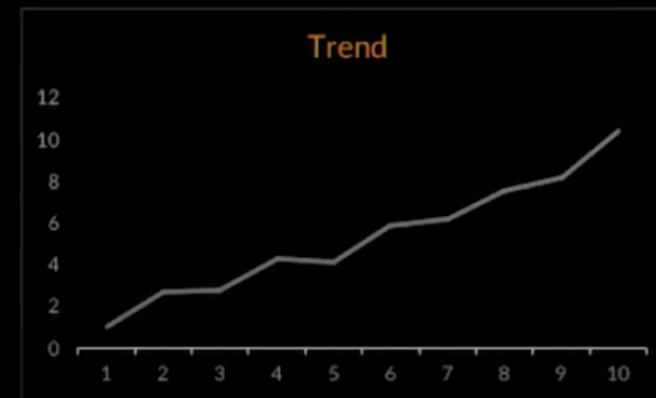
Abstraction: Find the story

Explore the data to find the patterns, trends, or outliers that can form the basis of a story



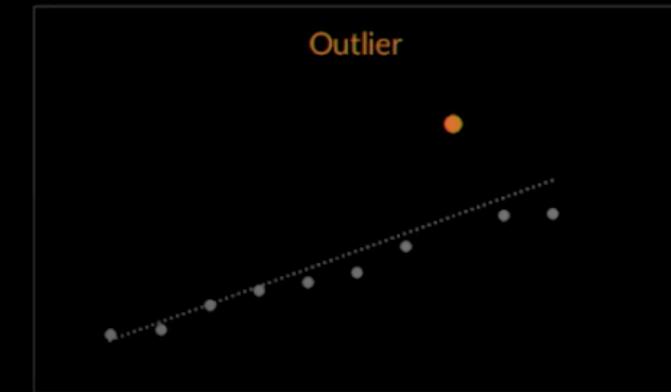
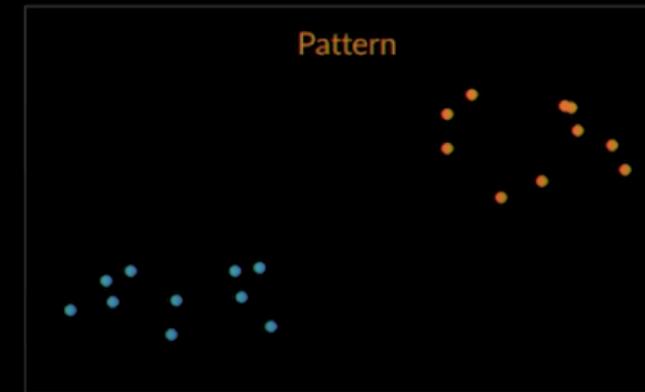
Representation: Choose how to tell it

The aim of visualisation is to improve understanding, so choosing the right representation is critical



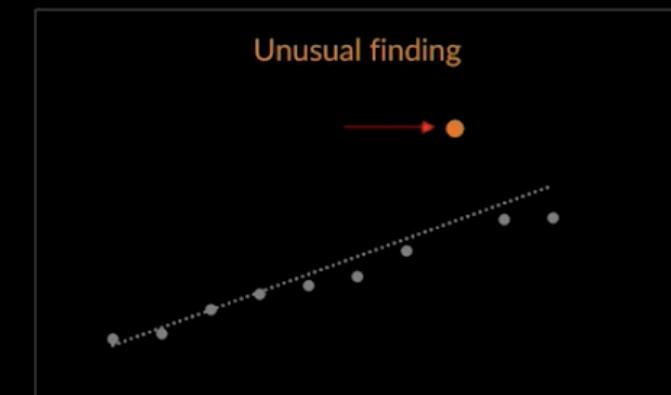
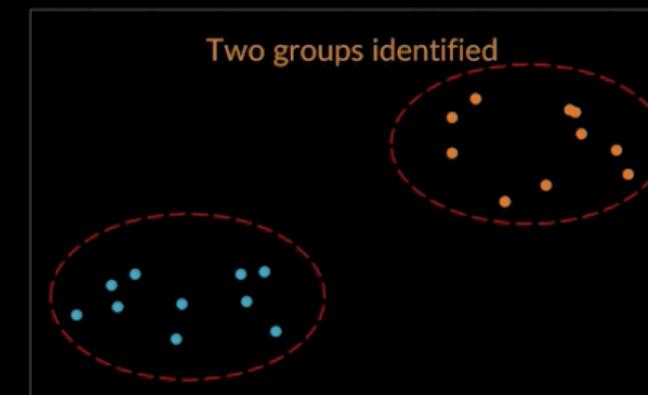
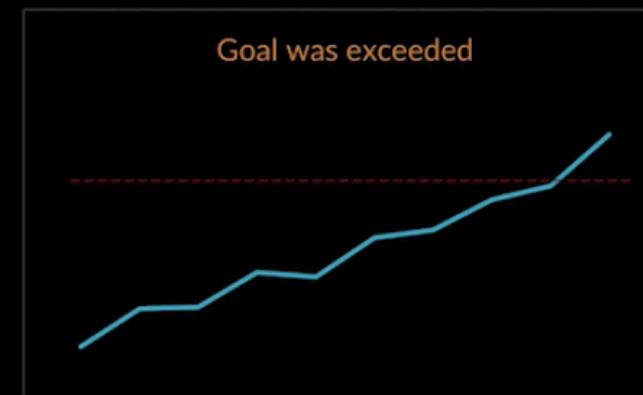
Aesthetics: Make it compelling

Proximity, alignment, repetition, scale, contrast, hue, saturation, and subtraction can all be used to emphasise your findings



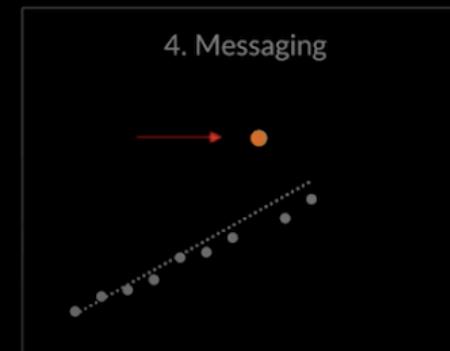
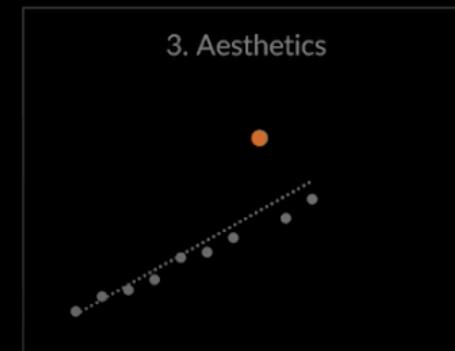
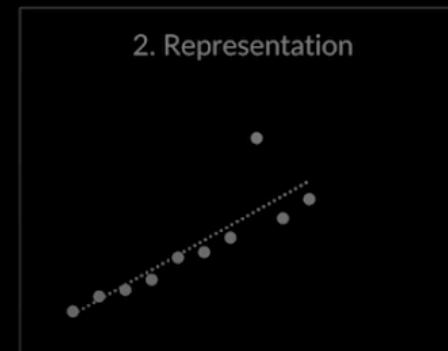
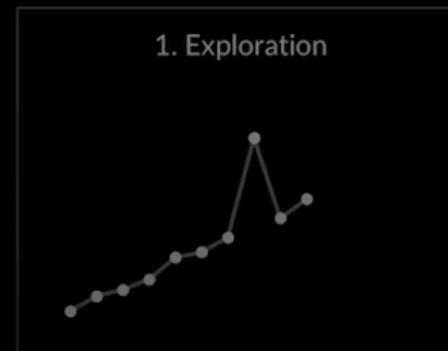
Messaging: Narrate the story

Proximity, alignment, repetition, contrast, and subtraction can all be used to emphasise a finding



Use visualisation to enhance your message

1. Find your story through exploratory visualisation
2. Choose a suitable representation for your story
3. Use aesthetics to highlight your findings
4. Support with appropriate messaging



people | tell stories

words | tell stories

pictures | tell stories

comics | tell stories

movies | tell stories

Stories are emotional

Stories are memorable

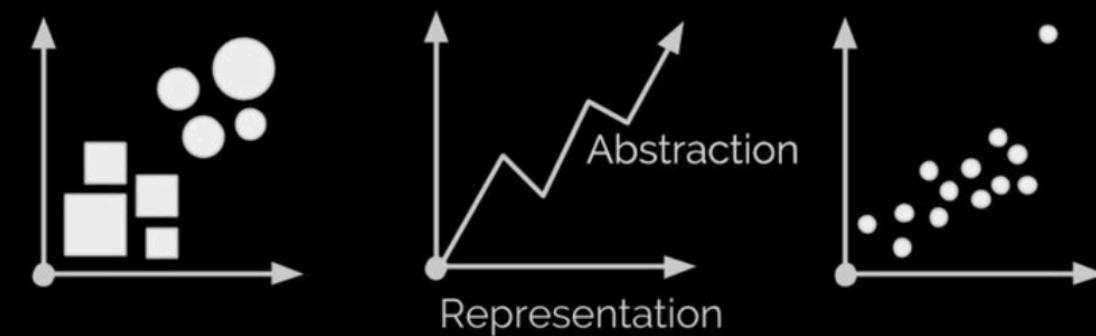
Stories are impactful

- ① Abstraction
- ② Representation
- ③ Aesthetics
- ④ Framing & Transition
- ⑤ Messaging
- ⑥ Flow
- ⑦ Interactivity

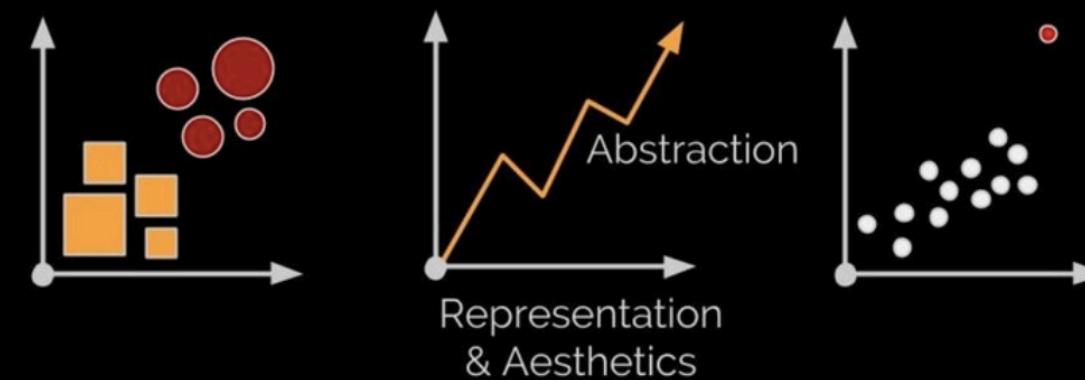
① Abstraction



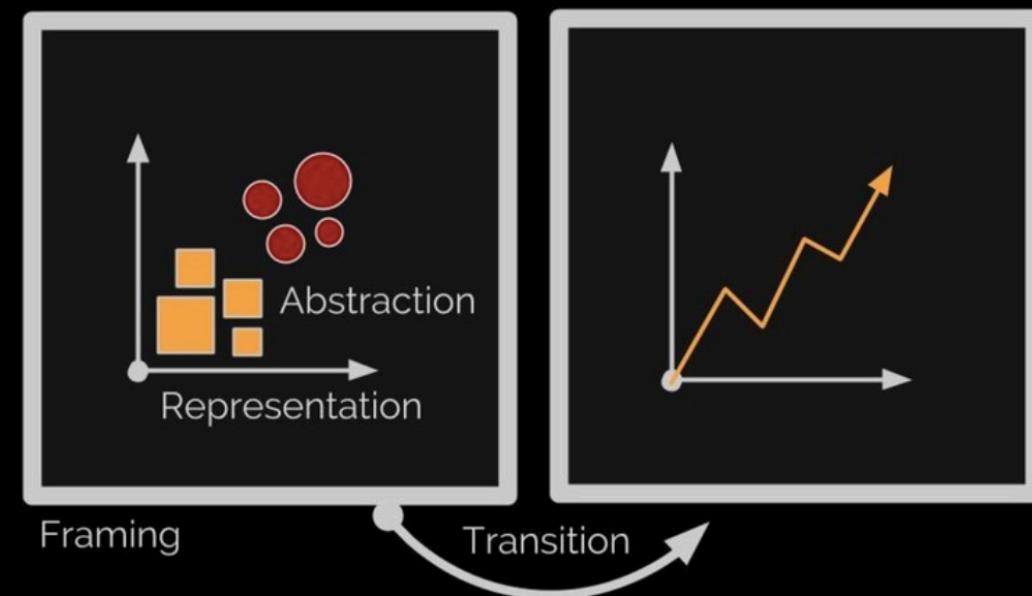
② Representation



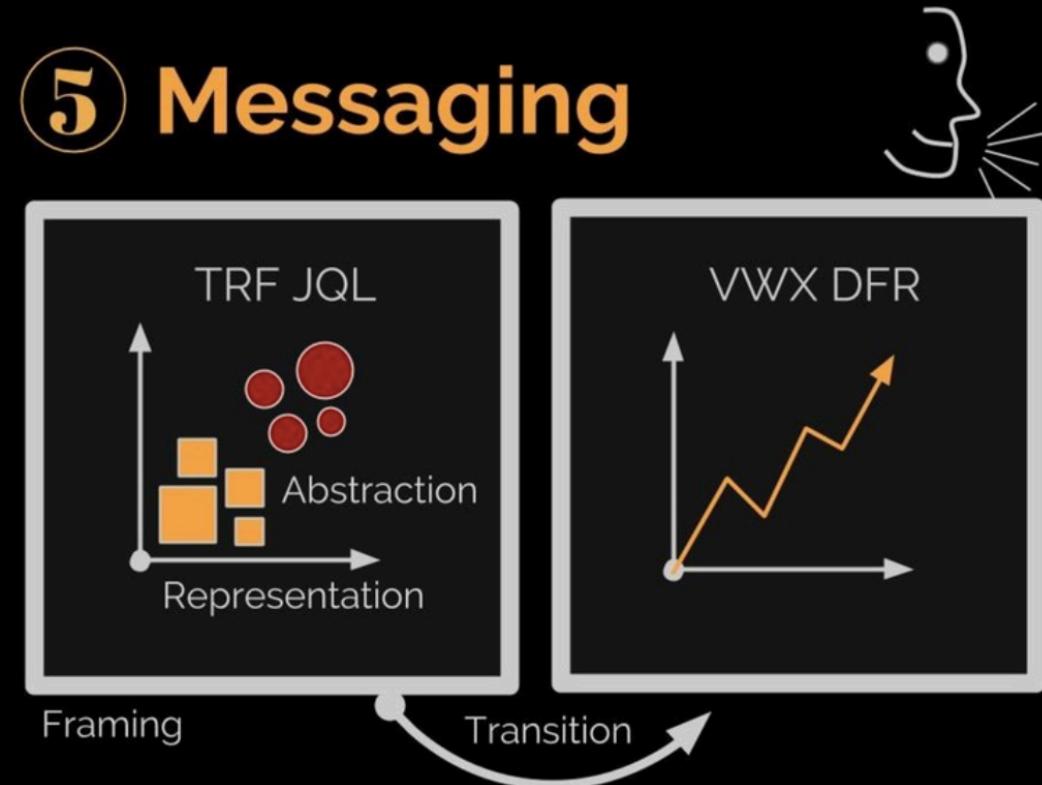
③ Aesthetics



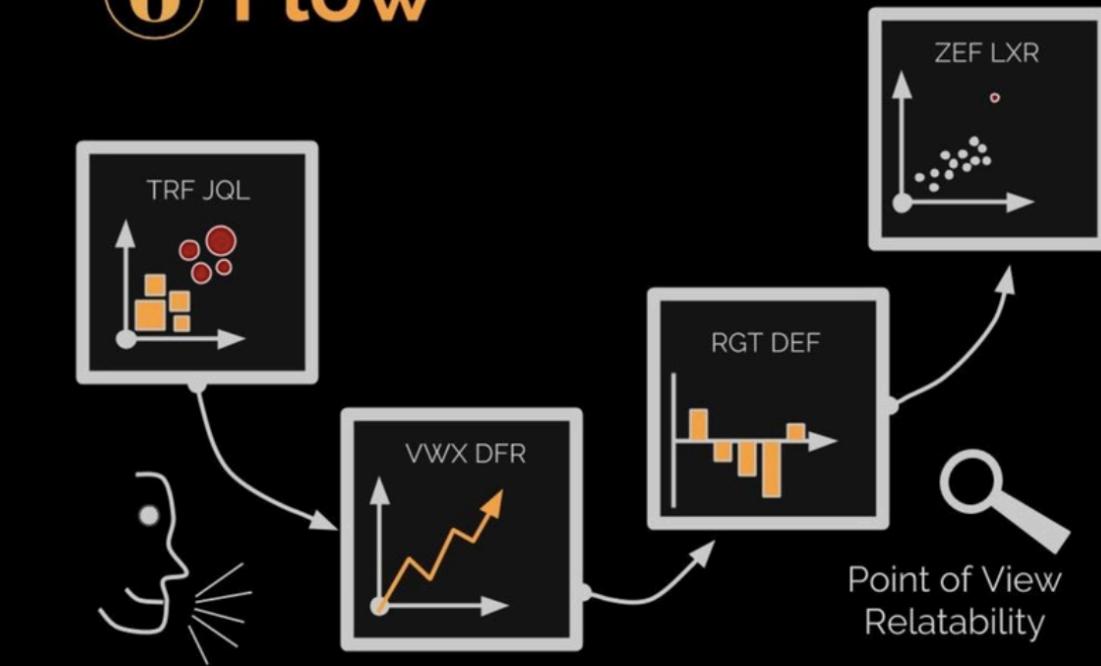
④ Framing & Transition



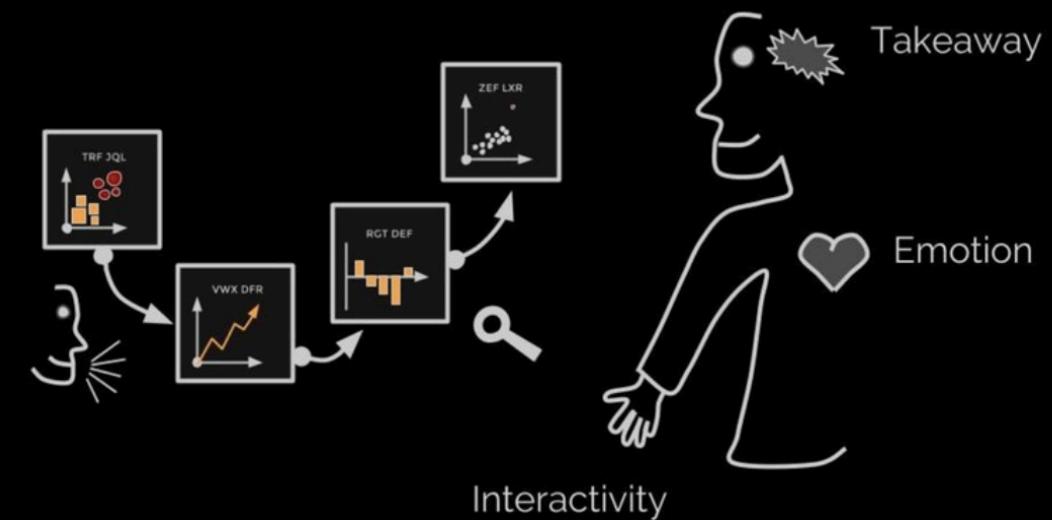
⑤ Messaging



⑥ Flow

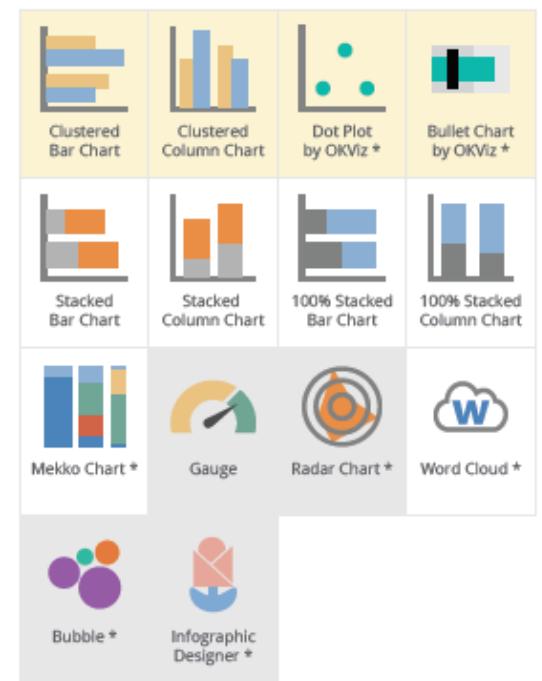


⑦ Interactivity



COMPARISON

Use these visuals when you want to display measures compared by its magnitude.



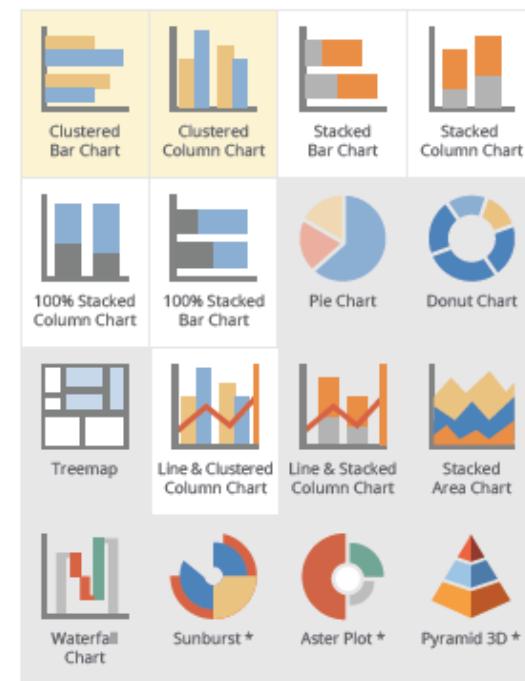
CHANGE OVER TIME

Use these visuals when you want to display the changing trend of measures.



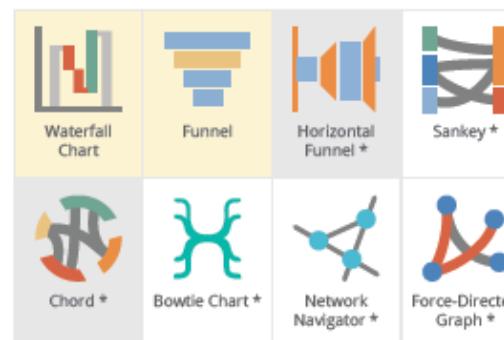
PART-TO-WHOLE

Use these visuals when you want to display parts that compose measures.



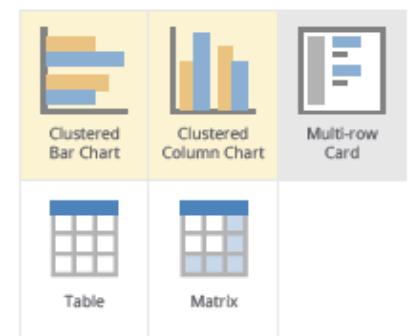
FLOW

Use these visuals when you want to display a flow or dynamic relations between measures.



RANKING

Use these visuals when you want to display measures by its rank order.



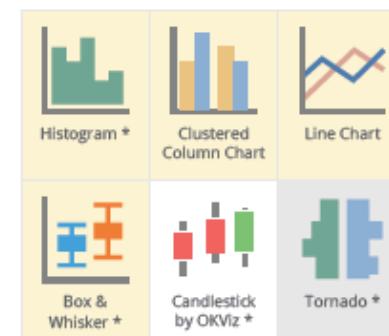
SPATIAL

Use these visuals when you want to display measures over spatial maps.



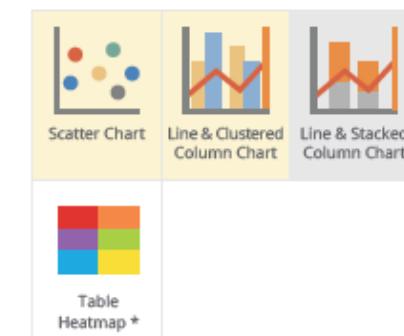
DISTRIBUTION

Use these visuals when you want to display the distribution of a measure.



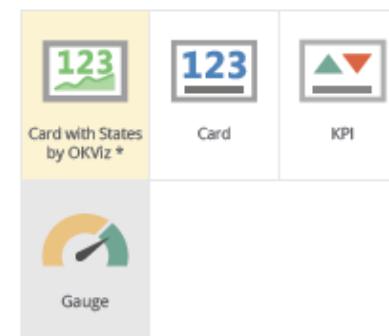
CORRELATION

Use these visuals when you want to display relations between measures.



SINGLE

Use these visuals when you want to display a single value.



Data Wrangling

Let try hands-on

Get the dataset from remote url

```
!wget
```

Install basic dependencies

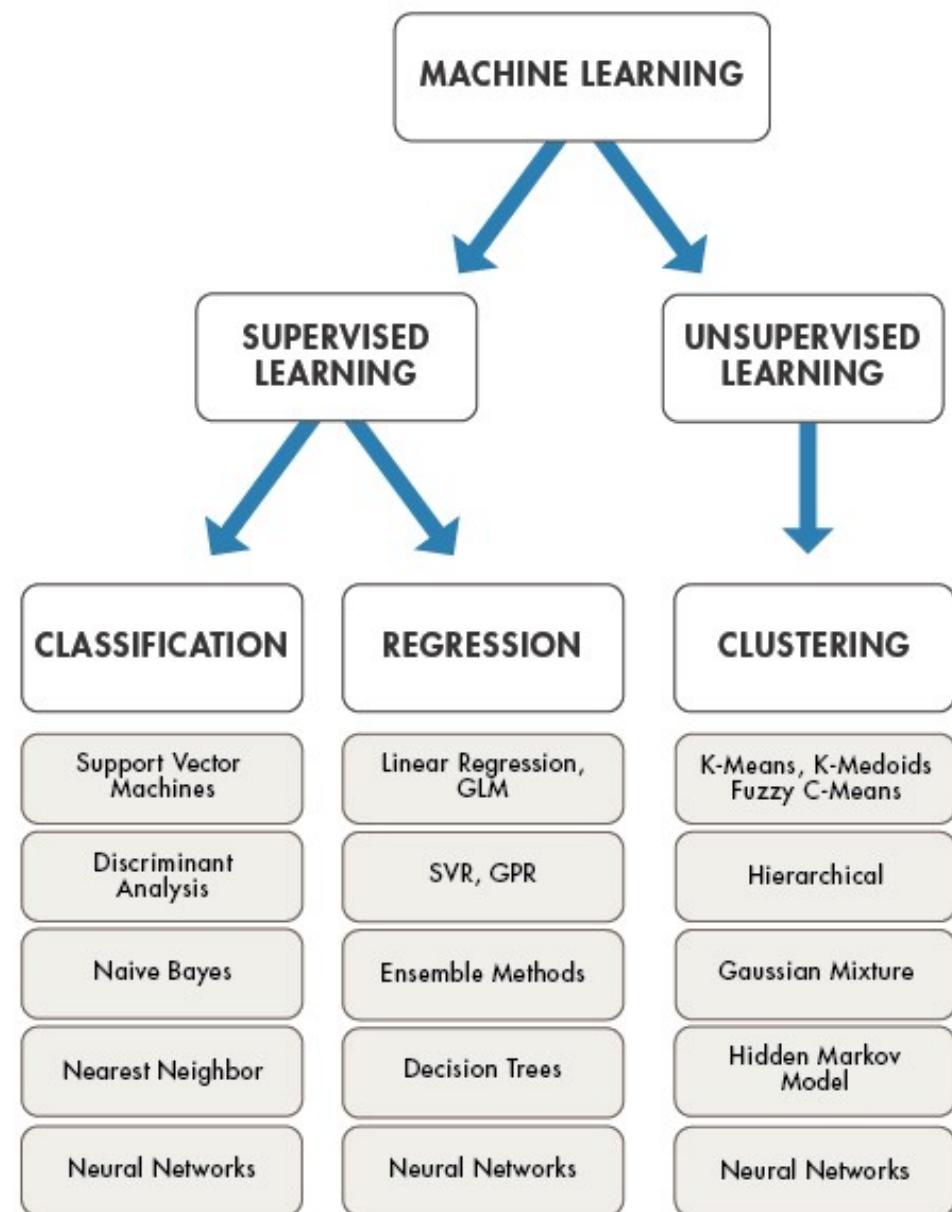
```
!pip install pandas  
!pip install numpy  
!pip install matplotlib  
!pip install seaborn
```

Import all libraries

```
%matplotlib inline  
import pandas as pd  
import numpy as np  
import matplotlib.pyplot as plt  
import seaborn as sns
```

Machine Learning Models





How many lines of code?


```
In [18]: from sklearn.externals.six import StringIO  
from IPython.display import Image  
from sklearn.tree import export_graphviz  
import pydotplus
```

```
DecsTree = DecisionTreeClassifier(criterion='gini', max_depth=6)  
DecsTree.fit(X_train_new, y_train_new)
```

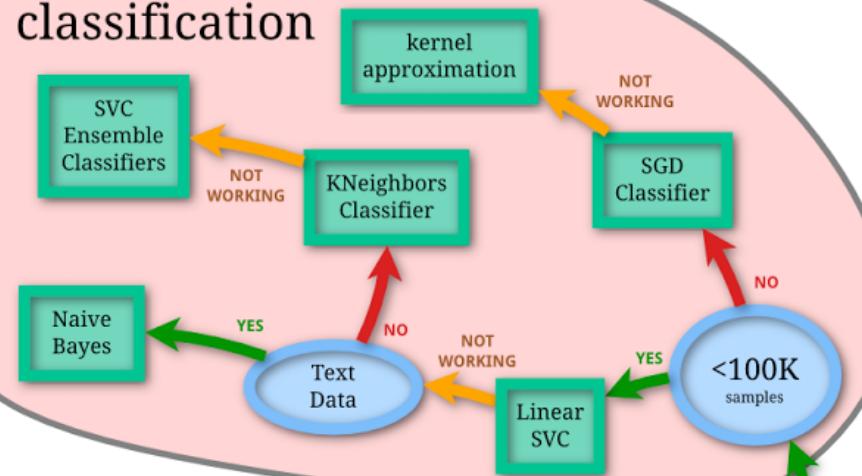
```
dot_data = StringIO()  
  
export_graphviz(DecsTree, out_file=dot_data,  
                filled=True, rounded=True,  
                special_characters=True, feature_names = col_names_list_new, class_names=['0','1'])  
  
graph = pydotplus.graph_from_dot_data(dot_data.getvalue())  
  
graph.write_png('Bank_DecsT_new.png',)  
# graph.set_size('"300, 180!"')  
# graph.write_png('resized_tree.png')  
Image(graph.create_png())
```

Out[18]:

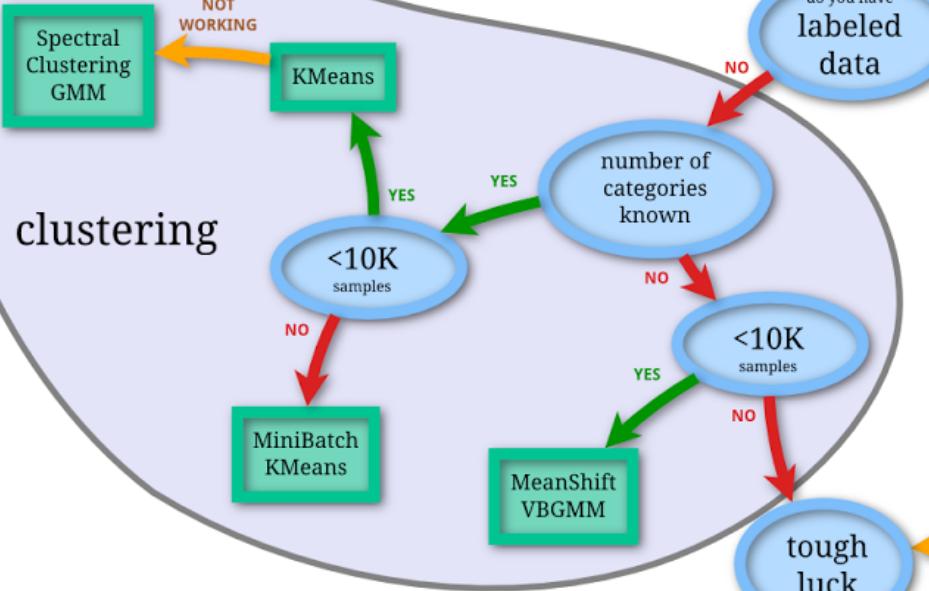


How to Pick Right Model

classification

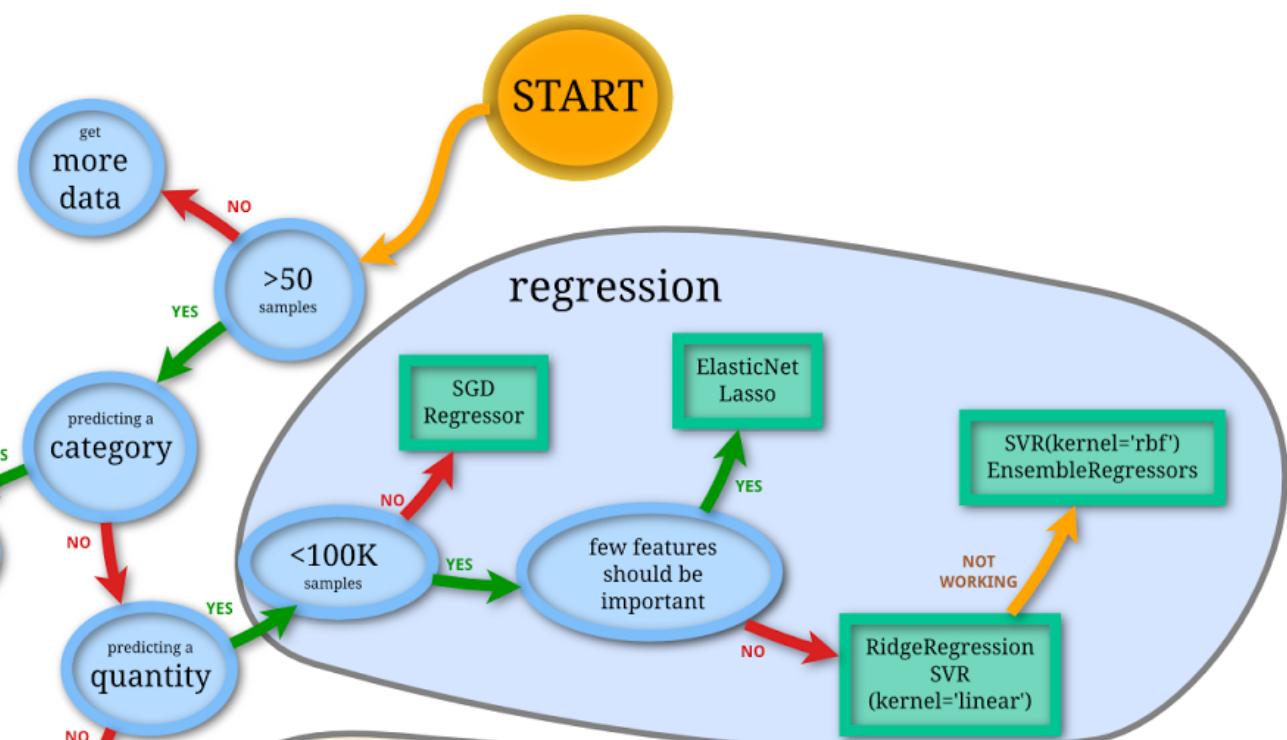


clustering

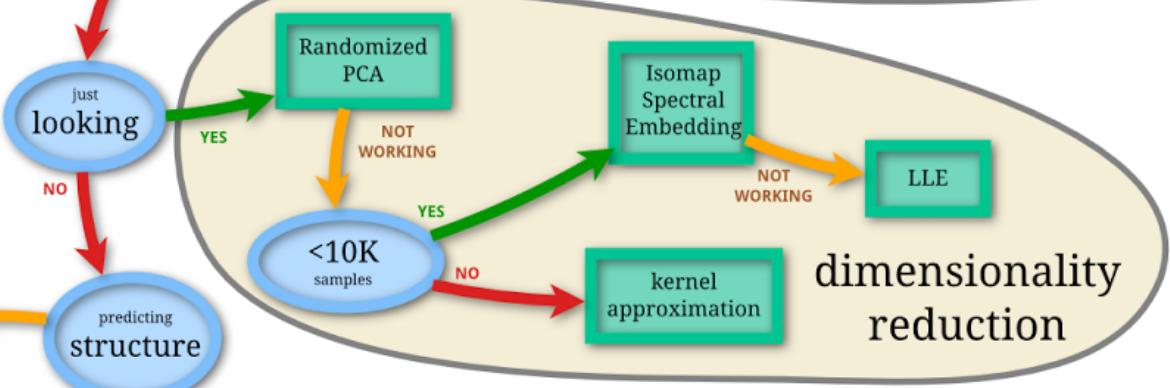


START

regression



dimensionality reduction



Quiz

<https://bit.ly/3fEldtO>