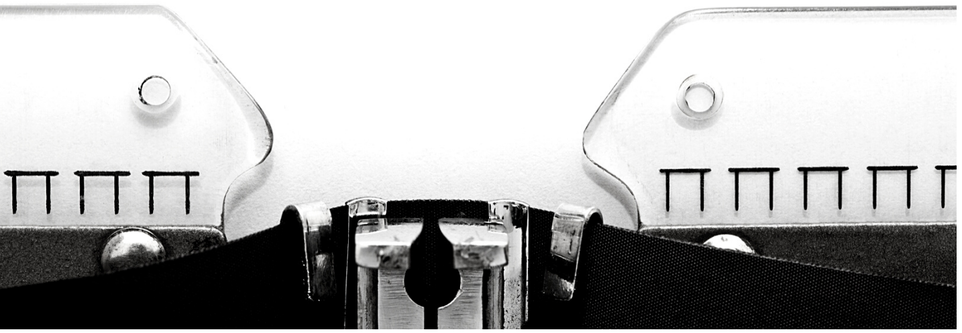
DATA

What's your story?



# CREATE

GOOD CHARTS

FAITH KANE

2021

**"Programs visualize data.**

**People visualize ideas."**

- Scott Berinato


# BIG IDEA

Meaningful context like audience, purpose, and domain knowledge can only be provided by you.

# STEPS TO CREATE YOUR CHARTS SYSTEMATICALLY

PREP → PROTOTYPE

1. PREPARE TO CREATE
2. TALK AND LISTEN
3. SKETCH
4. PROTOTYPE

# WHY USE A SYSTEM?

Going through this process will take your charts from "unthinking, autogenerated charts" to Good Charts.

- Scott Berinato

# 1. PREPARE TO CREATE

**Concept**

Carve out mental space, physical space, and white space.
Get organized in mind and space as you find direction for your project.

**Headspace**

Keep your thinking broad and don't let your data limit you.
You are not using your data at this stage, but you understand it.

**Artifacts**

Notes on Story Framework, Audience, Setting, Brainstorm, Types of Viz.
Document your decisions honing in on the Big Idea for your presentation.

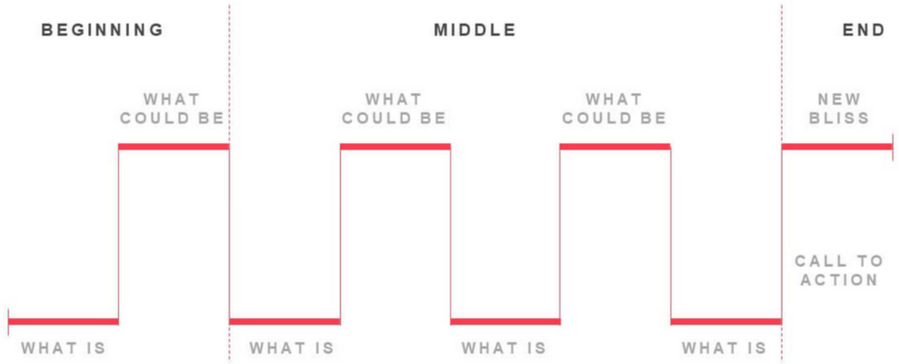
# ASK YOURSELF

1. WHICH STORY FRAMEWORK FITS MY PROJECT?
2. WHO IS THE AUDIENCE FOR MY PRESENTATION?
3. WHAT ARE MY SETTING AND DELIVERABLES?
4. WHAT DID I LEARN FROM MY BRAINSTORM SESSION?
5. WHAT TYPES OF VISUALIZATIONS ALIGN WITH MY GOALS?

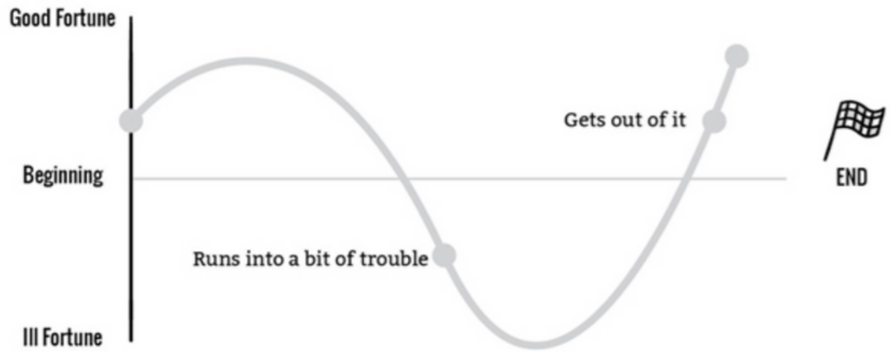
# WHICH STORY FRAMEWORK DO I USE?

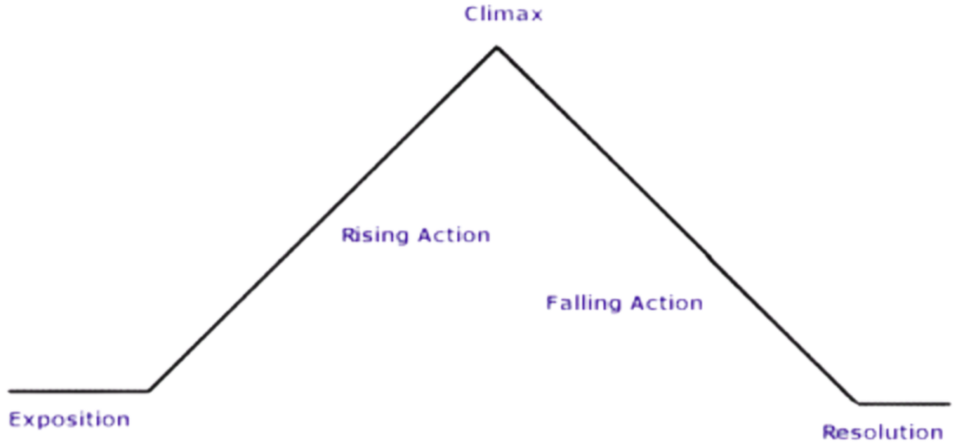
MAKE IT EASY

Persuasive Story Pattern




Man in Hole



# WE USED...

Explored our second problem statement/hypothesis.
Reported that our data proved one part to indeed be a driver!

Explored our second problem statement/hypothesis.
Reported that our data proved one part to not be a driver.

Climax

Demonstrated what we discovered about the driver of performance that we found.

Explored our first problem statement/hypothesis.
Reported that our data proved our initial thoughts to not be a factor in performance.

Went into the significance of what we found.

Pulled back to talk about our tools and process.
Introduced our two problem statements.

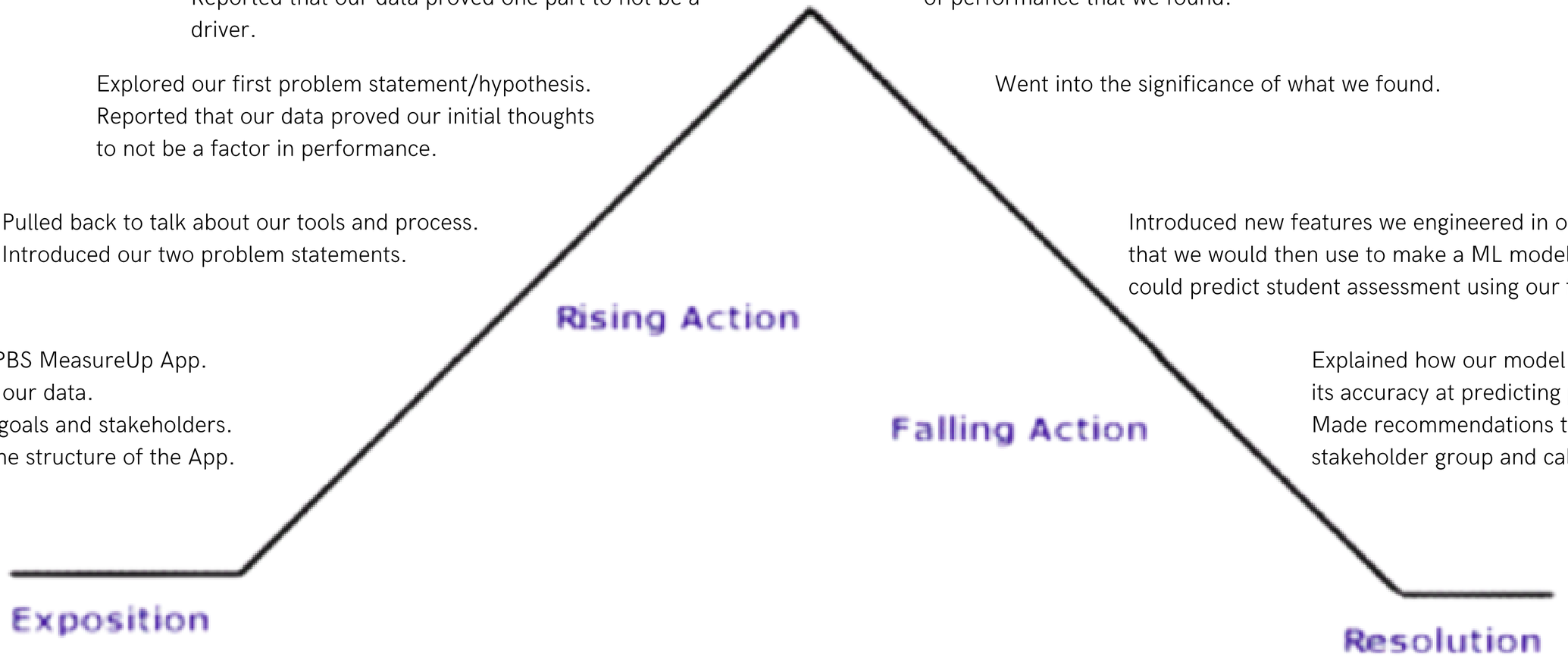
Introduced new features we engineered in our data that we would then use to make a ML model that could predict student assessment using our features.

Rising Action

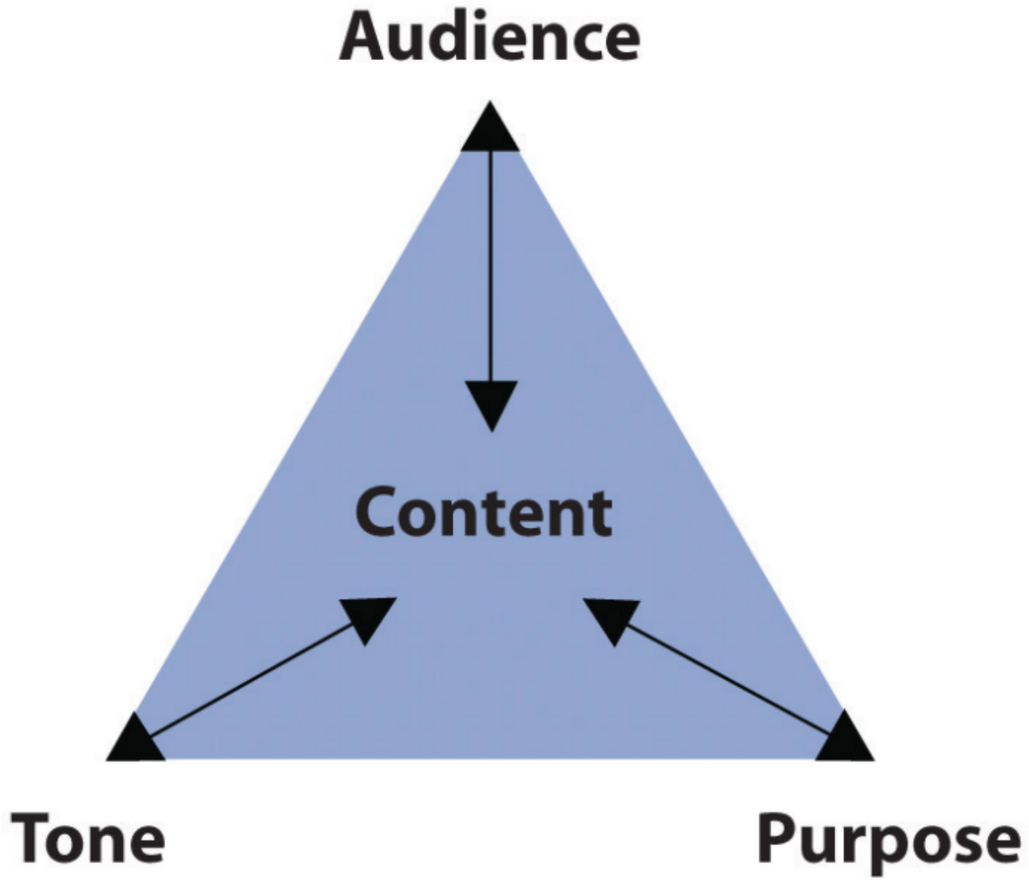
Introduced the PBS MeasureUp App.
Gave context to our data.
Made clear our goals and stakeholders.
Demonstrated the structure of the App.

Falling Action

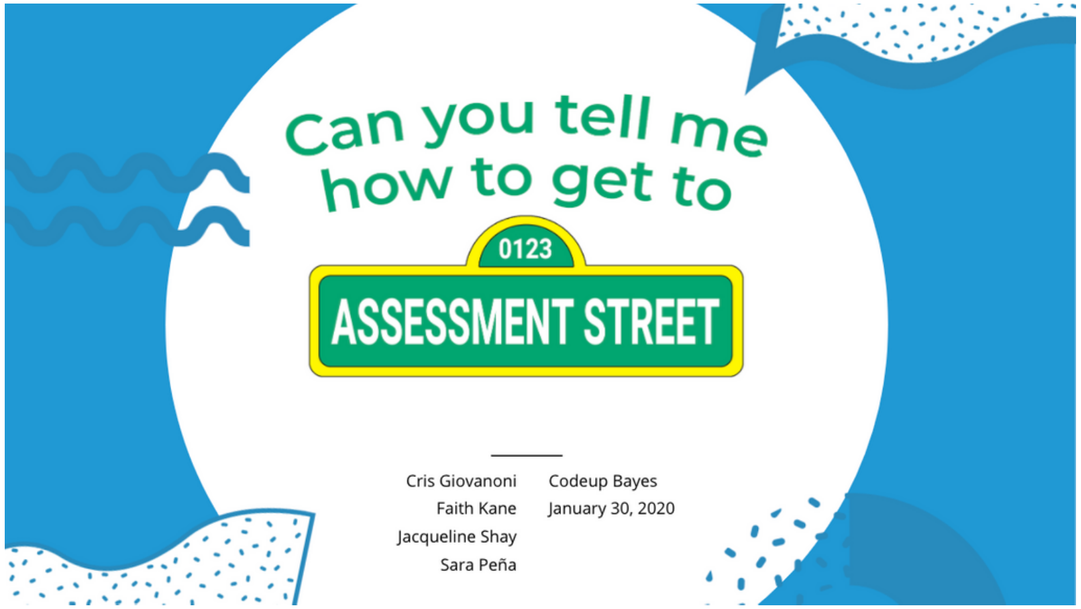
Explained how our model worked and its accuracy at predicting performance.
Made recommendations to each stakeholder group and called it a night!



# WHO IS MY AUDIENCE?



# OUR AUDIENCE...

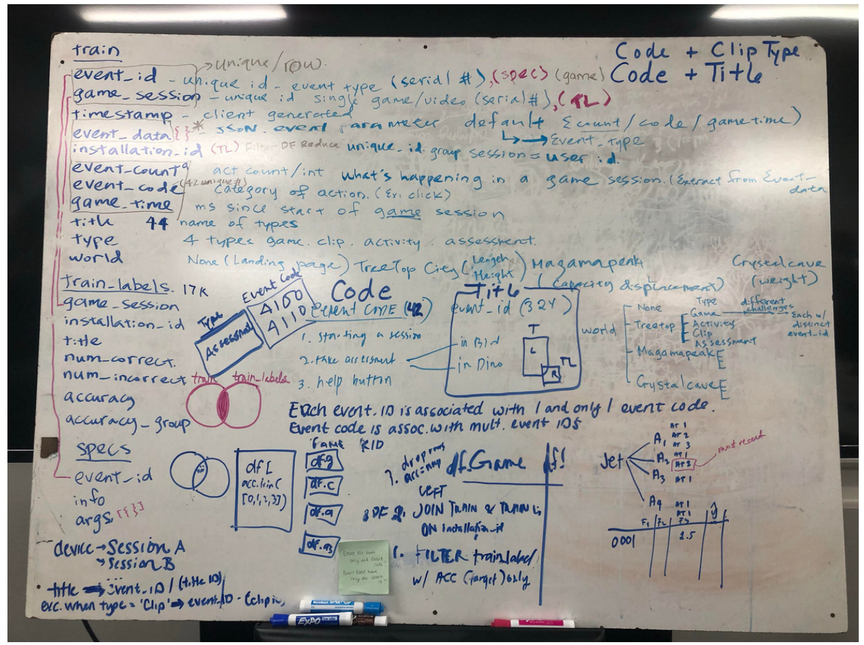



1. LEVEL OF LANGUAGE
2. TYPES OF CHARTS
3. AMOUNT OF TEACHING REQUIRED

# WHAT ARE MY SETTING AND DELIVERABLES?

# YOUR SETTING...

# WHAT DID I LEARN FROM MY BRAINSTORM?



# FIND YOUR BIG IDEA


"A big idea is that one key message you want to communicate. It contains the impetus that compels the audience to set a new course with a new compass heading. Screenwriters call this the “controlling idea.” It’s also been called the gist, the takeaway, the thesis statement, or the single unifying message." - Nancy Duarte

# A BIG IDEA MUST...

NANCY DUARTE

1. CONVEY YOUR UNIQUE PERSPECTIVE.
2. COMMUNICATE WHAT'S AT STAKE.
3. BE STATED AS A COMPLETE SENTENCE.

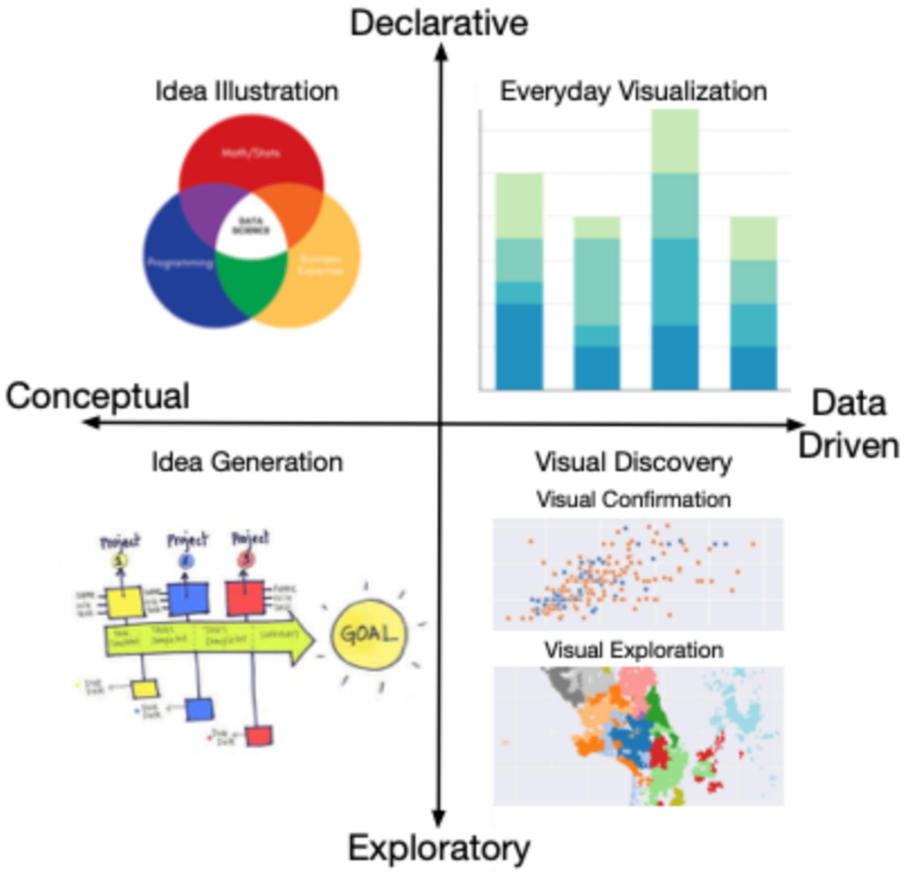
# EXAMPLE BIG IDEA


"For example, “the fate of the oceans,” is only a topic, not a big idea. “Worldwide pollution is killing the ocean and us,” is a big idea with a unique perspective. Your big idea doesn’t need to be so unusual that no presenter has ever expressed it before. But it must be unique to you, not merely a generalization." - Nancy Duarte

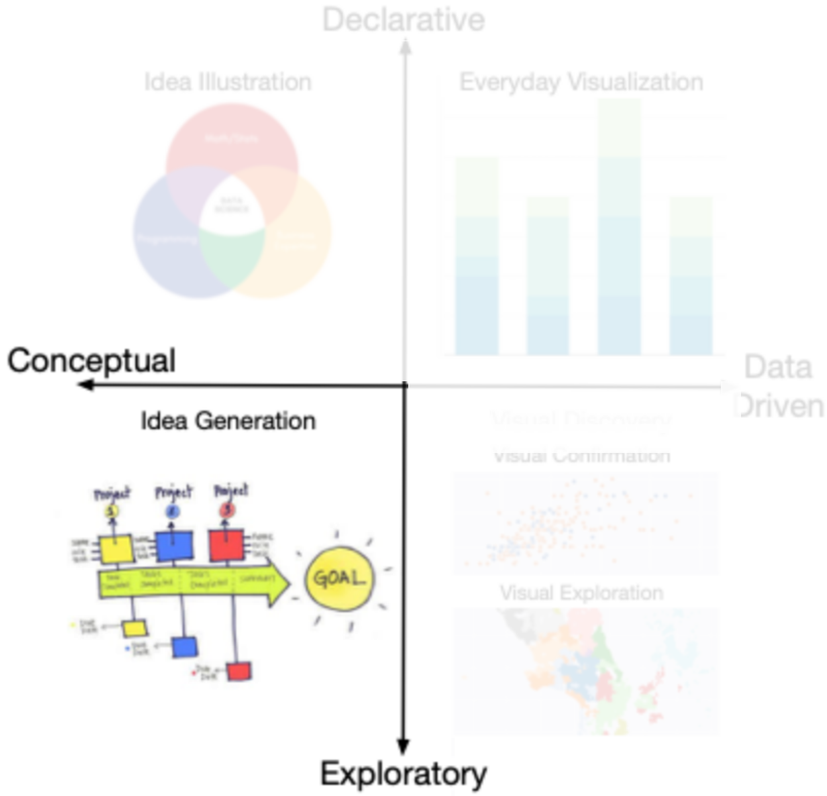
# WHAT TYPE OF VISUALIZATION DO I USE?

"[A]lthough there are a few rules you should know and try to follow, most of them are actually just conventions. When it comes to choosing what kind of chart you'll make, the ends ought to justify the means. If it clearly conveys the idea you want your audience to come away with, use it."

- Scott Berinato, Good Charts Workbook



# IDEA GENERATION




### DATA TYPE

Complex

Undefined

### TYPICAL SETTING

Working Session

Brainstorming

### VIZ TYPE

Creative

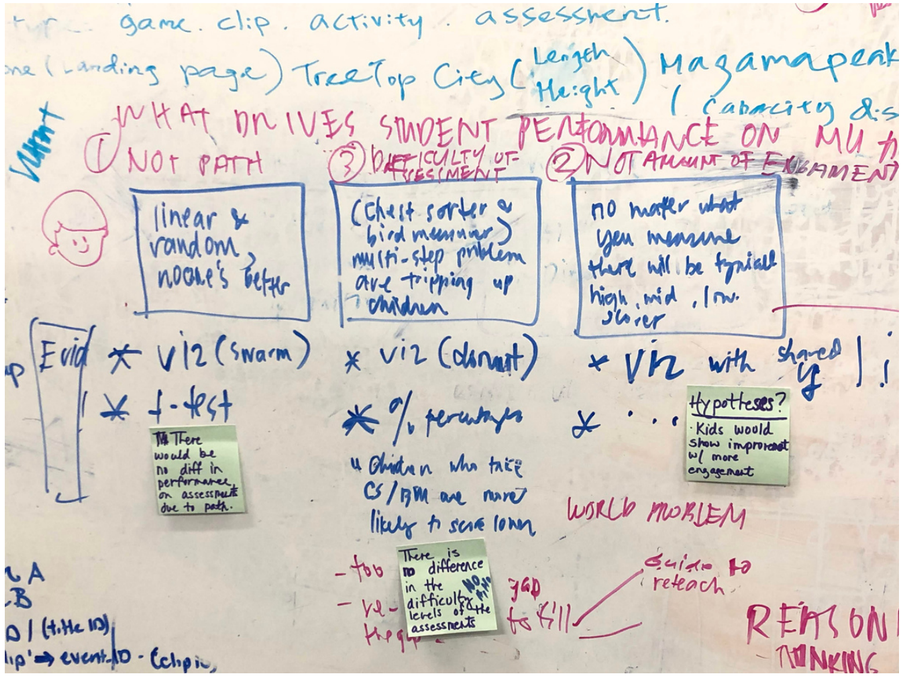
Metaphorical

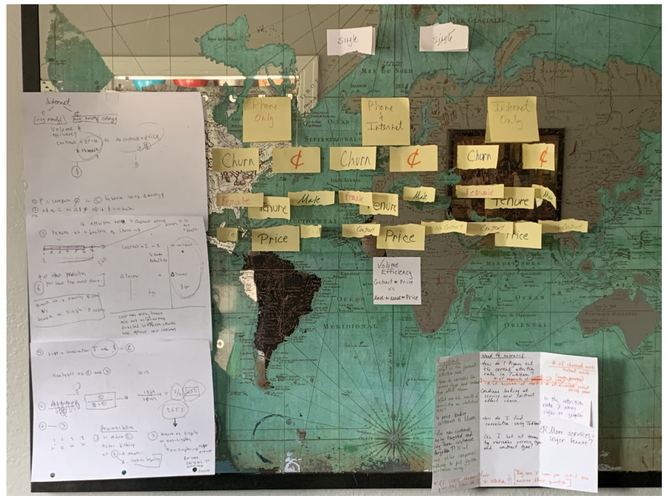
### GOALS

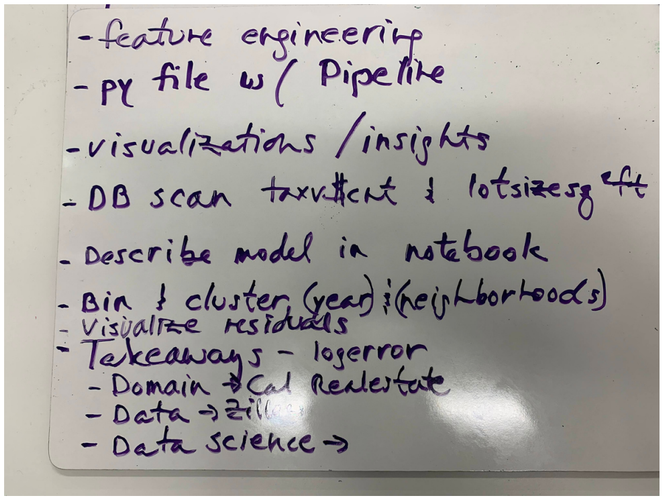
Discovery

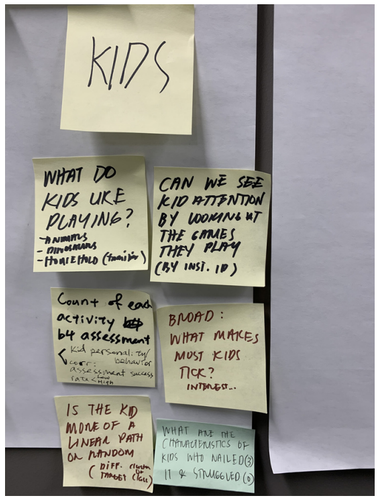
Learning & Simplification

# IDEA GENERATION EXAMPLES

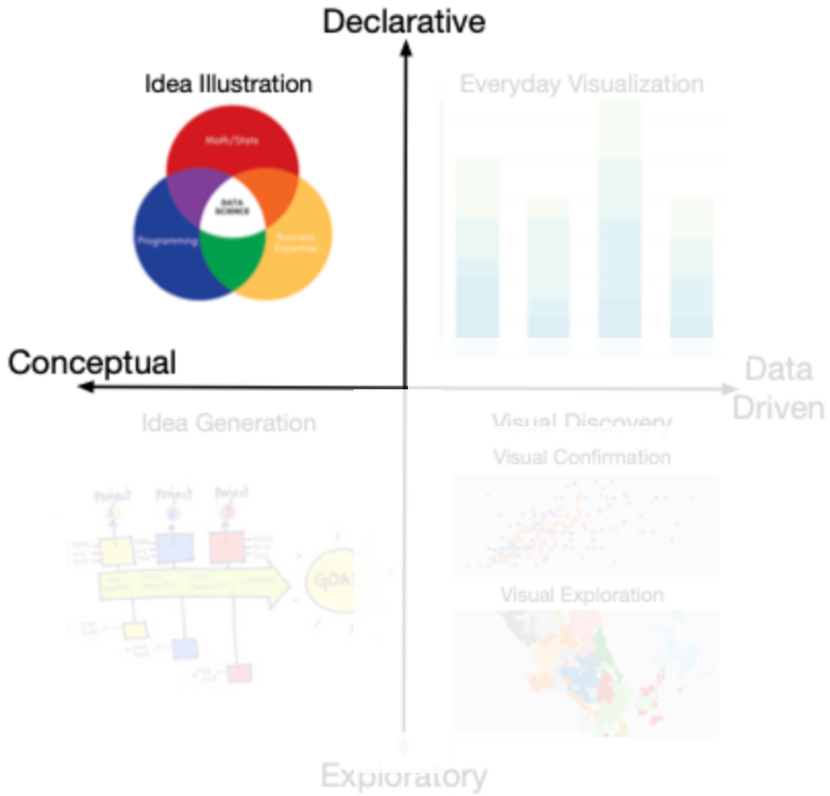


# IDEA ILLUSTRATION




### DATA TYPE

Process Explanation
Concept Demonstration

### TYPICAL SETTING

Presentations
Teaching

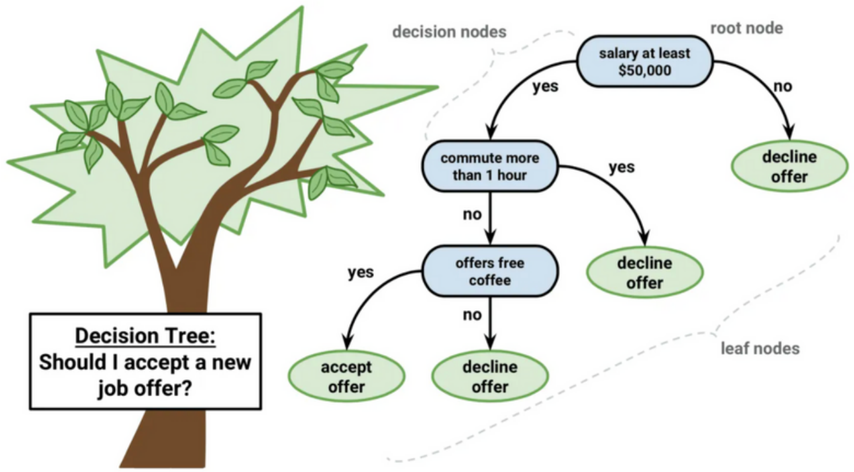
### VIZ TYPE

Simple
Metaphorical

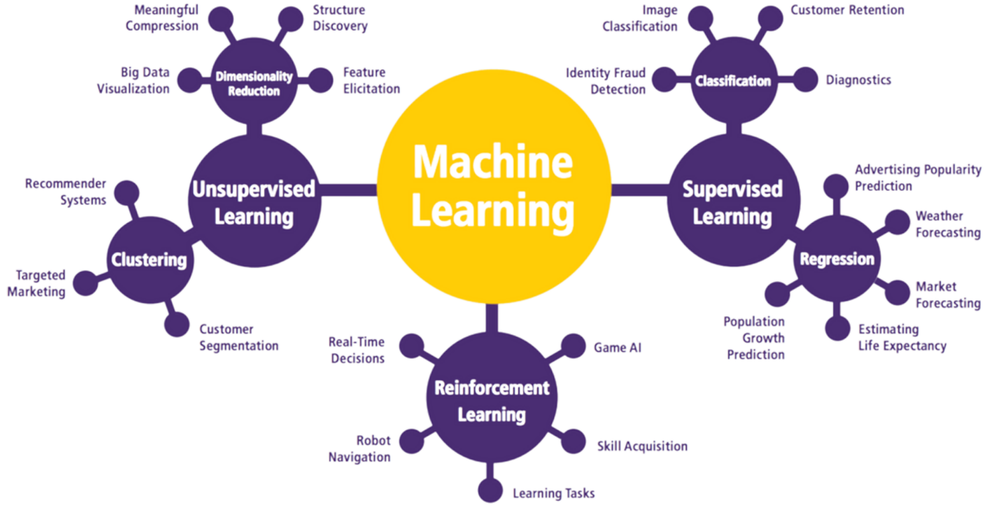
### GOALS

Learning
Simplifying

# IDEA ILLUSTRATION EXAMPLES

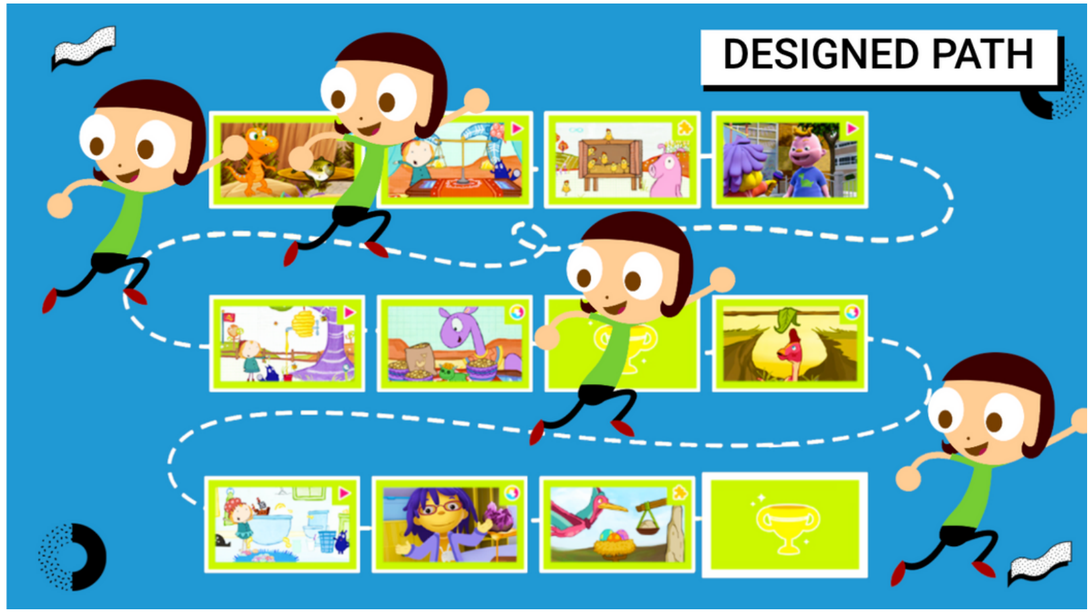

Decision Tree classifier, **Image credit:** www.packtpub.com



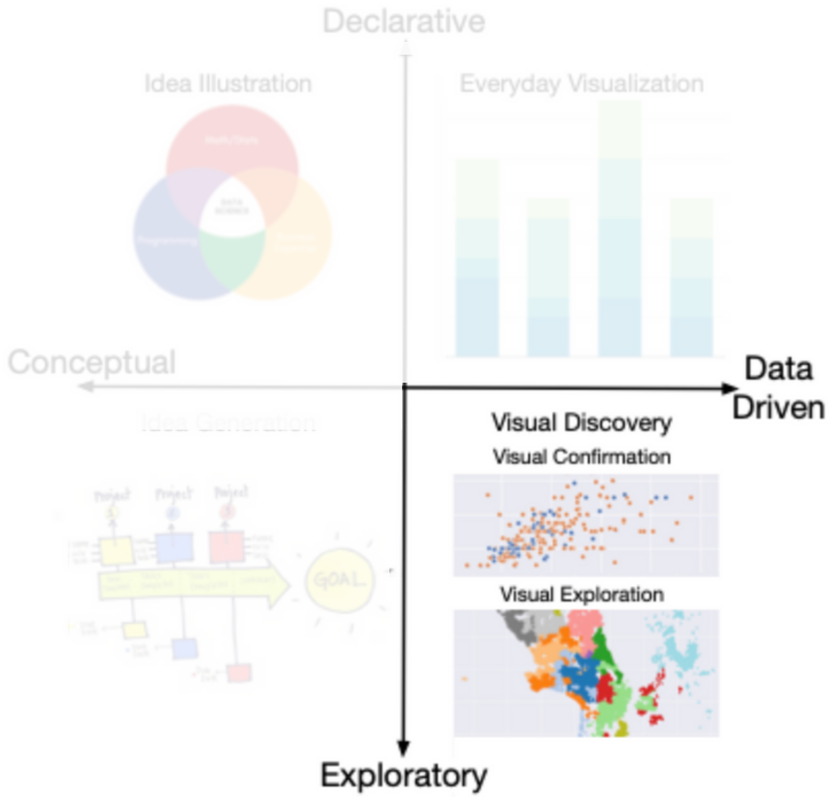
# OUR IDEA ILLUSTRATIONS







# VISUAL DISCOVERY




### DATA TYPE

Big Data

Dynamic

### VIZ TYPE

Advanced

Unconventional

### TYPICAL SETTING

Working Session

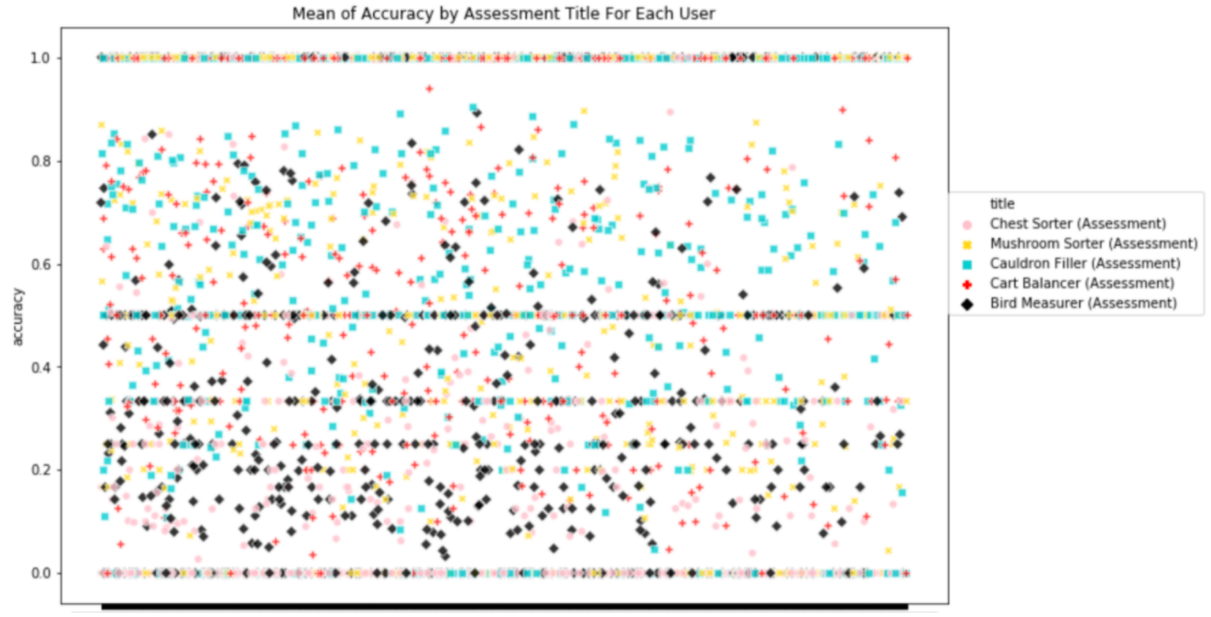
Testing Analysis

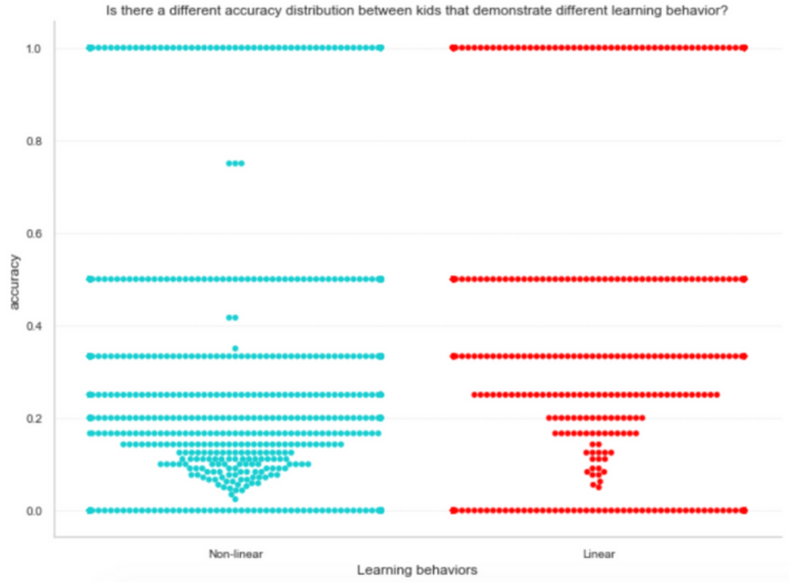
### GOALS

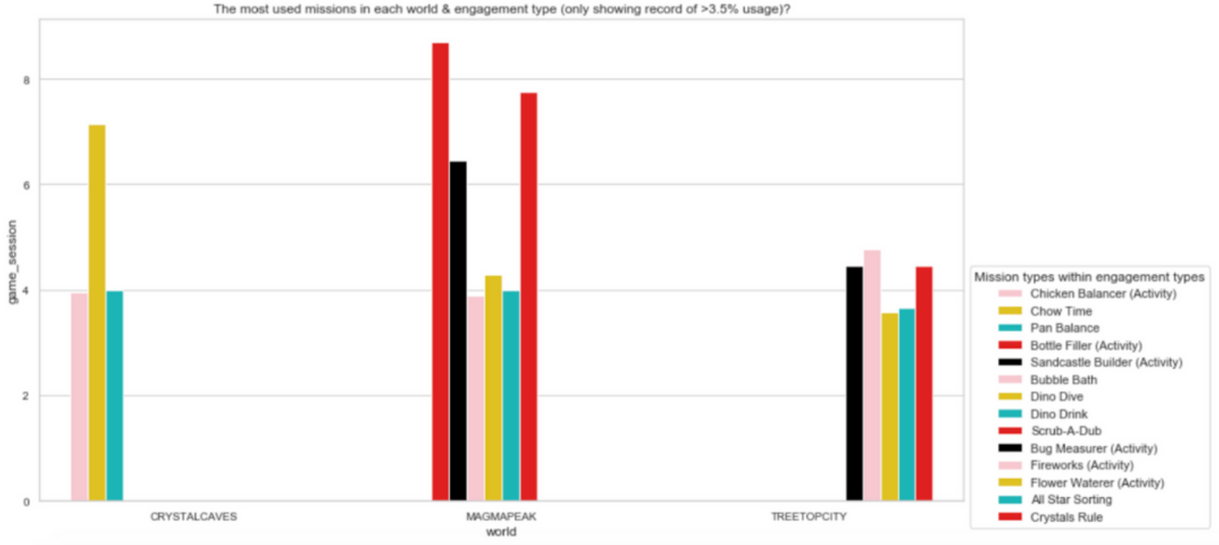
Trend Spotting

Deep Analysis & Sense Making

# VISUAL DISCOVERY EXAMPLES



# EVERYDAY VISUALIZATION


### DATA TYPE

Simple

Low Volume

### VIZ TYPE

Conventional Chart

Static

### TYPICAL SETTING

Formal

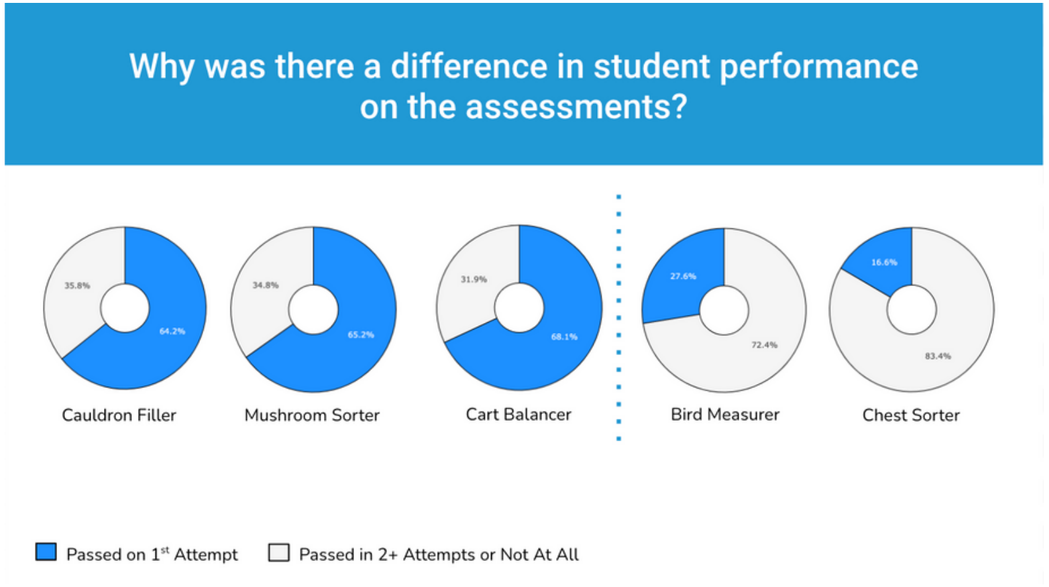
Presentations

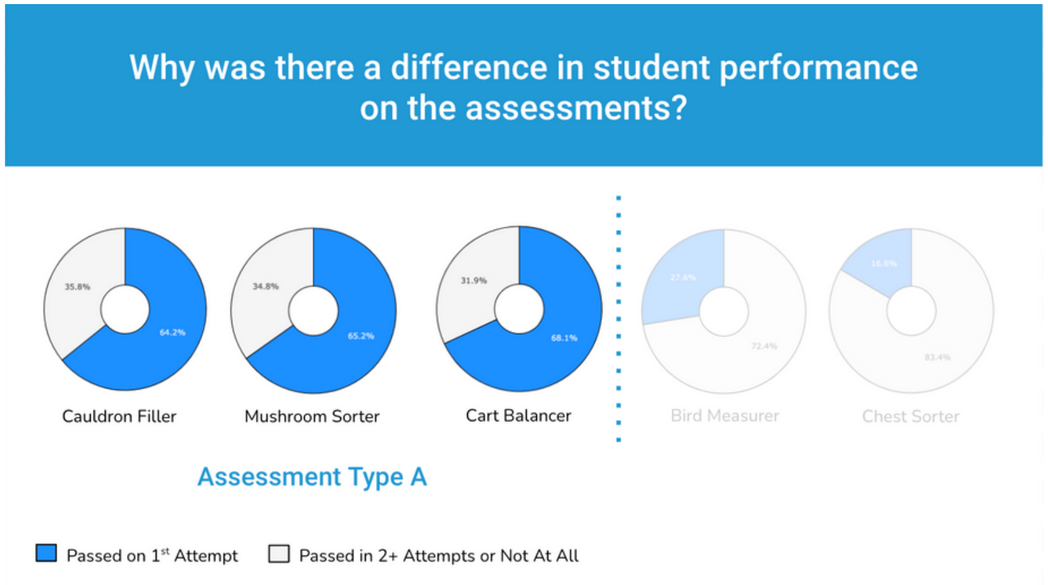
### GOALS

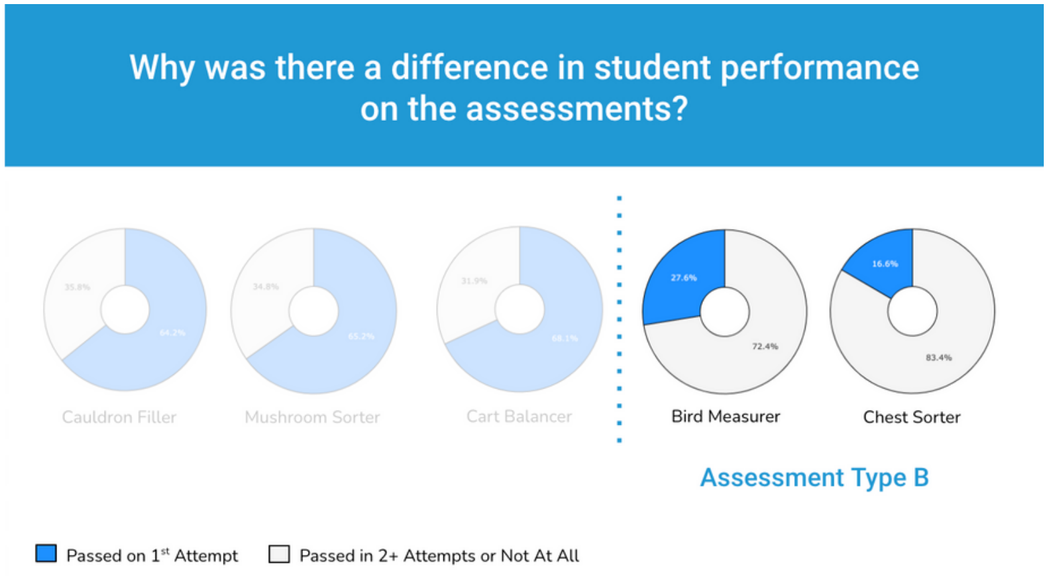
Affirm

Set Context

# EVERYDAY VISUALIZATION EXAMPLES



# 2. TALK AND LISTEN

**Concept**

Explain basic concepts and information about your project and ideas.
Take notes and pay attention to keywords as clues to chart types.

**Headspace**

You may feel a little vulnerable and even awkward at first. It takes practice.
Be attentive to keywords in yours and others' conversations.

**Artifacts**

Notes capturing keywords and identifying gaps in thinking or knowledge.
Questions that need to be answered to close gaps.

# LET'S DISCUSS →

STEM 1 →

What I'm working on is...

01

STEM 2 →

What I'm trying to say or show (confirmatory)/prove or learn (exploratory) is...

02

STEM 3 →

The reason I'm doing this is...

03

STEM 4 →

The Big Idea of my presentation is...

04

# WHAT WE LEARNED

ARTIFACT 1 →

We want to identify the drivers of student performance on in-app assessments by comparing factors like learning path taken, time spent in the app, and assessment taken to assessment performance. Our goal is to communicate our findings to parents, educators, and app designers to improve the usefulness of using the app.

ARTIFACT 2 →

We need to provide context (teach) about:

- the app.
- the student users.
- the assessments.

# 3. SKETCH

**Concept**

Match keywords from Step 2 to different types to visualizations.
Sketch ideas for charts that might best represent what you want to show.

**Headspace**

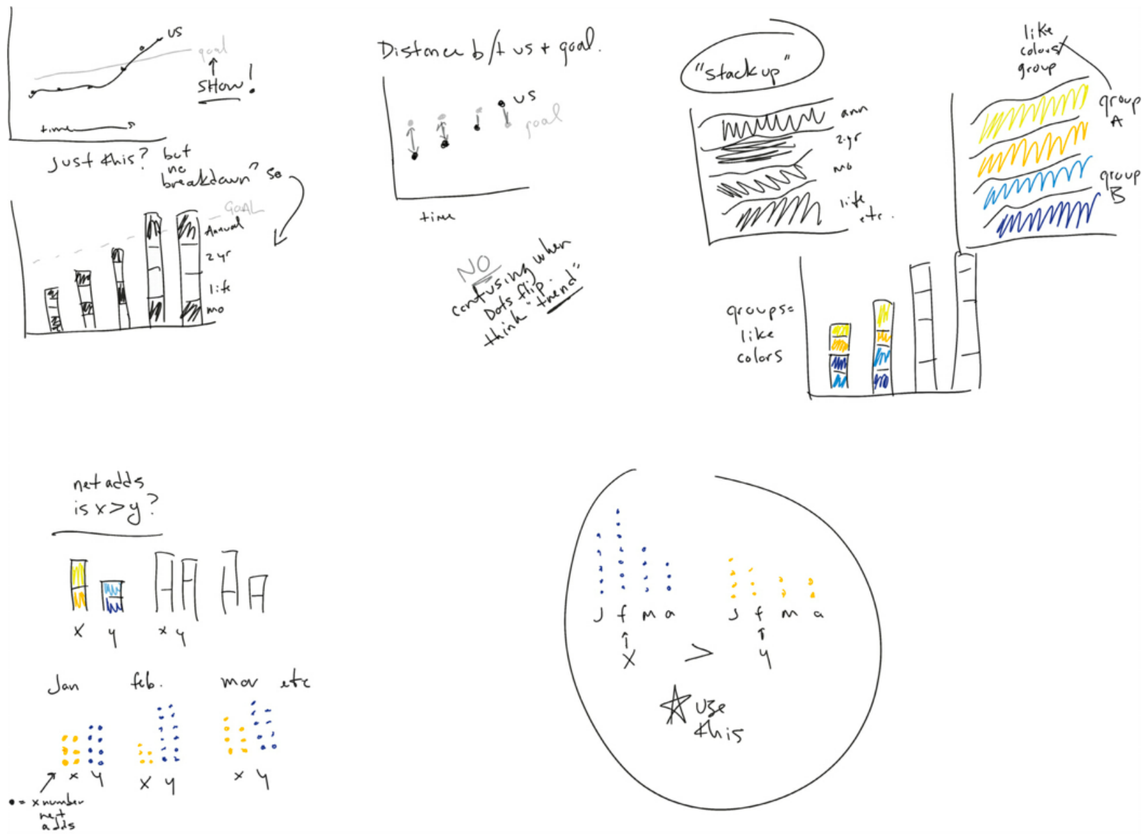
Stay in a creative space here, open-minded about different ideas.
Don't worry about details or accuracy at this point. Generate chart ideas!

**Artifacts**

Quick, messy sketches of charts inspired by your keywords from Step 2.
A rough draft of your charts and presentation that support your Big Idea.

# SKETCH EXAMPLE

## GENERATIVE



# 4. PROTOTYPE

**Concept**

Refine your rough draft from Step 3 using real figures, titles, themes, etc.
"Sketching is generative. Prototyping is iterative." -Scott Berinato

**Headspace**

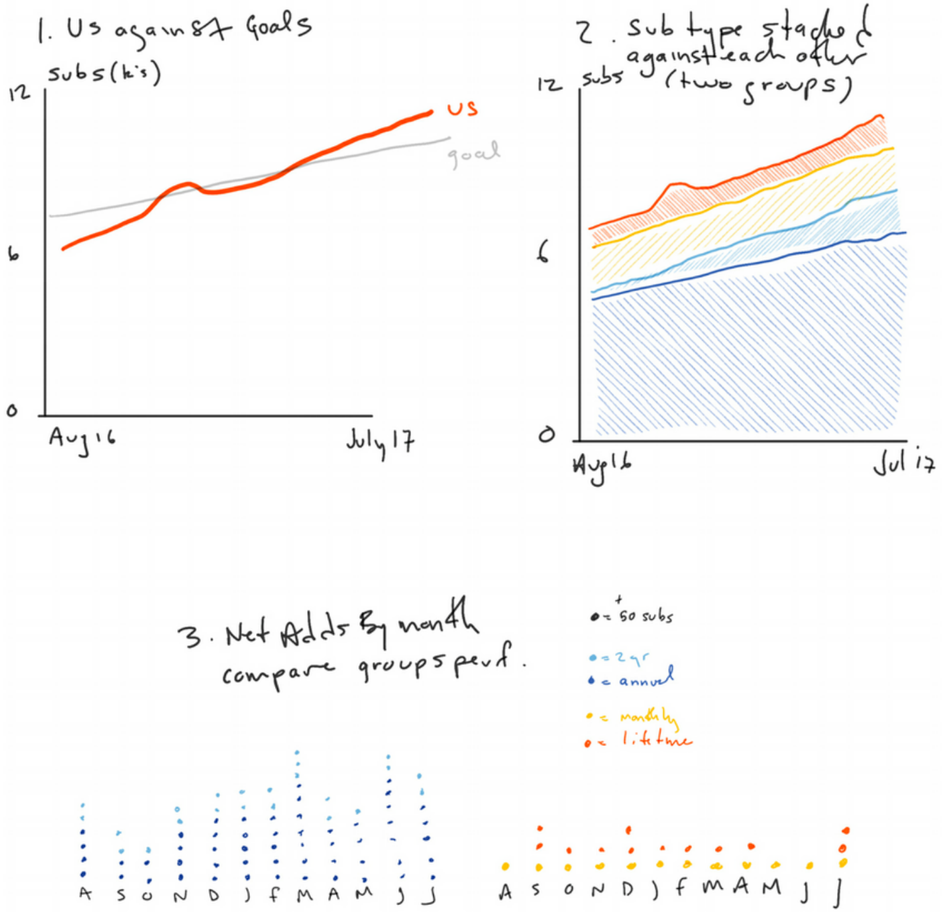
You are ready to focus in on the details like colors, values, and titles.
You are comfortable with your Big Idea, and you're out of ideas.

**Artifacts**

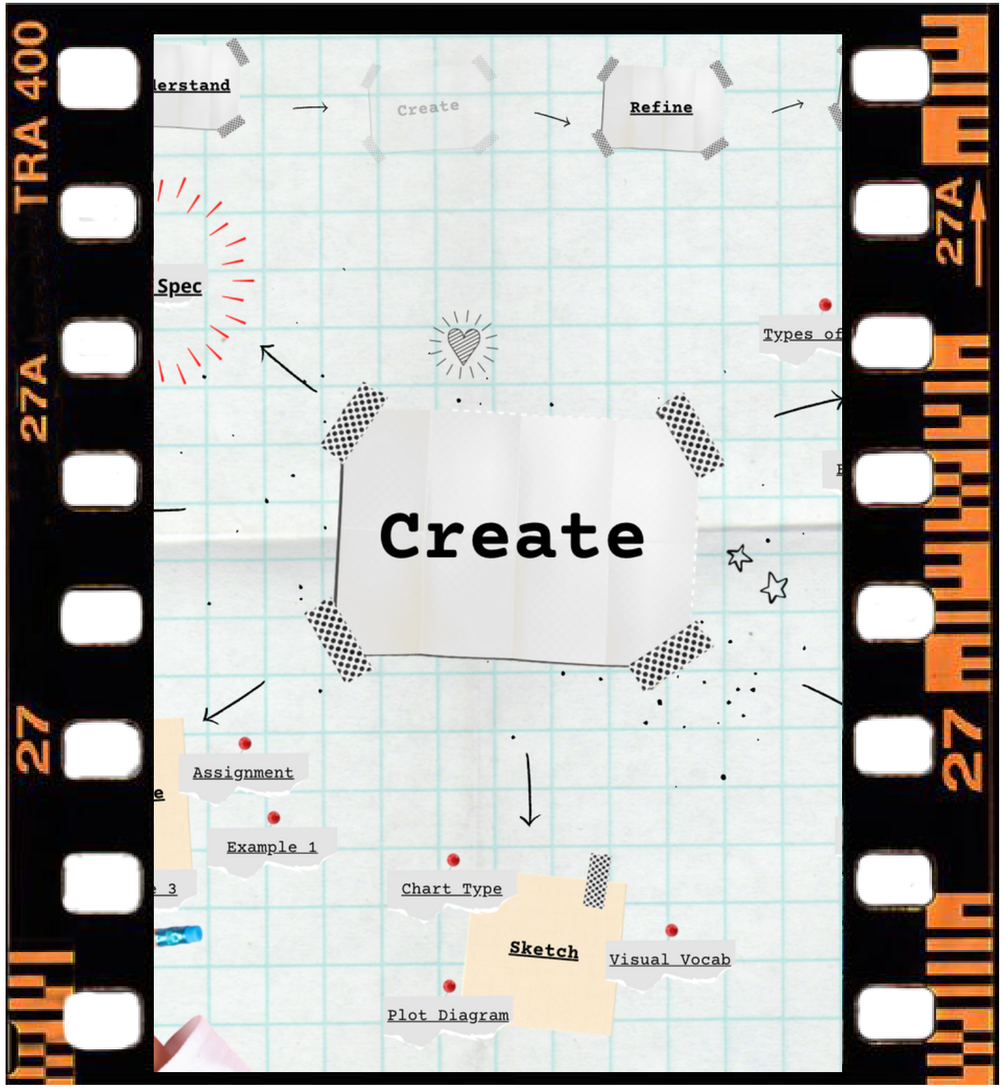
Refined charts on paper, whiteboard, or both.
A map to guide you through creating your charts and presentation.

# PROTOTYPE EXAMPLE

## ITERATIVE

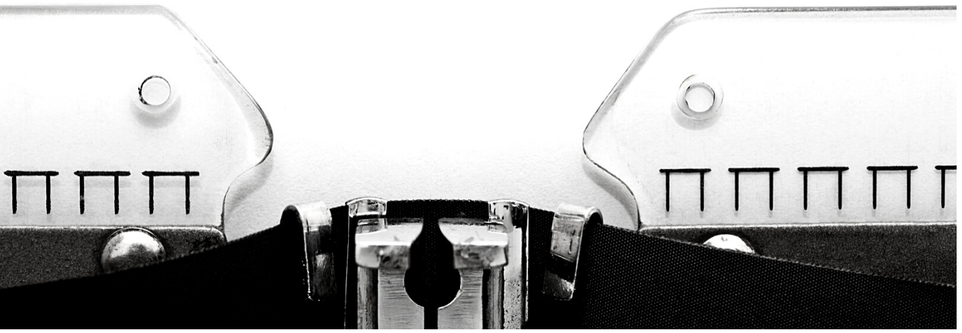


# TIME TO EXPLORE



DATA

What's your story?



# REFINE & PRESENT

GOOD CHARTS

FAITH KANE

2021

**"The less you have to talk about a visualization, and the more you can talk about its ideals, the better."**
- Scott Berinato

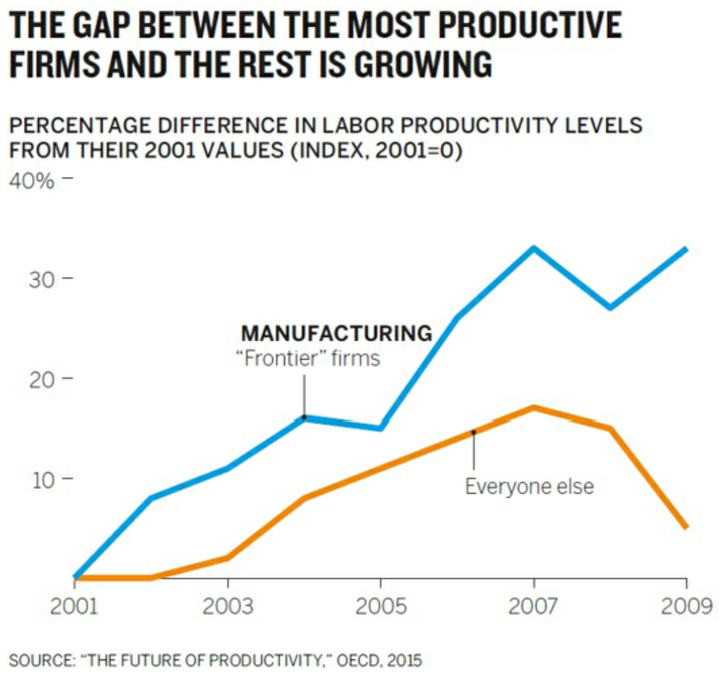
WHAT'S THE

# BIG IDEA

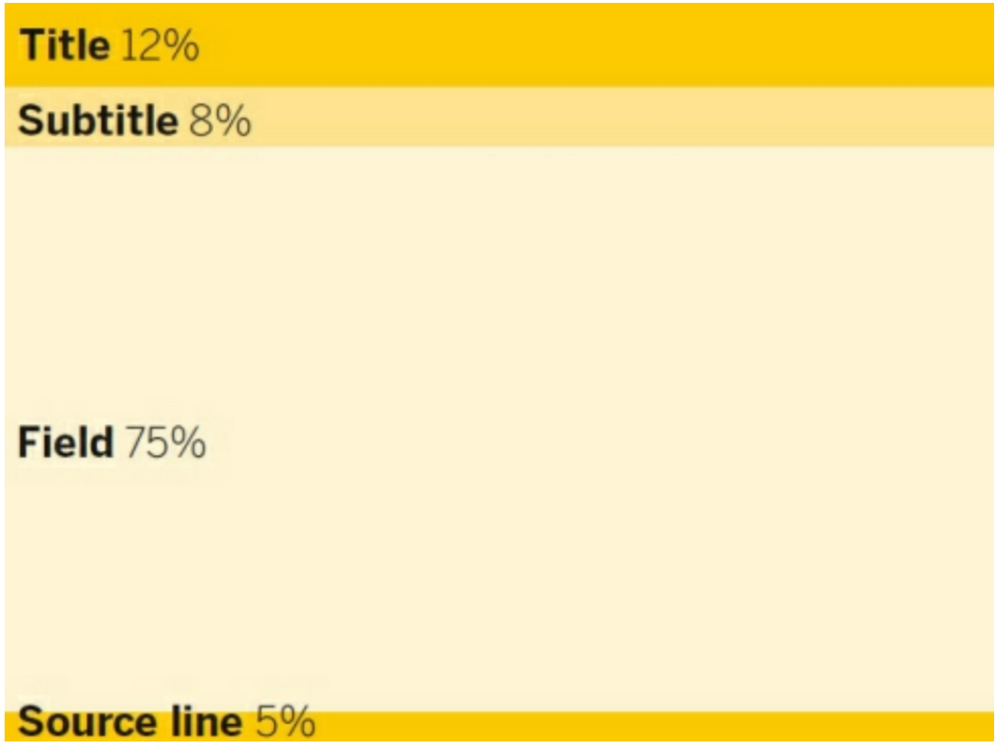
How little can you show and still convey your idea clearly and effectively? "Simplicity is Courageous."

# CHART STRUCTURE

## KEEP CHART COMPONENTS CONSISTENT



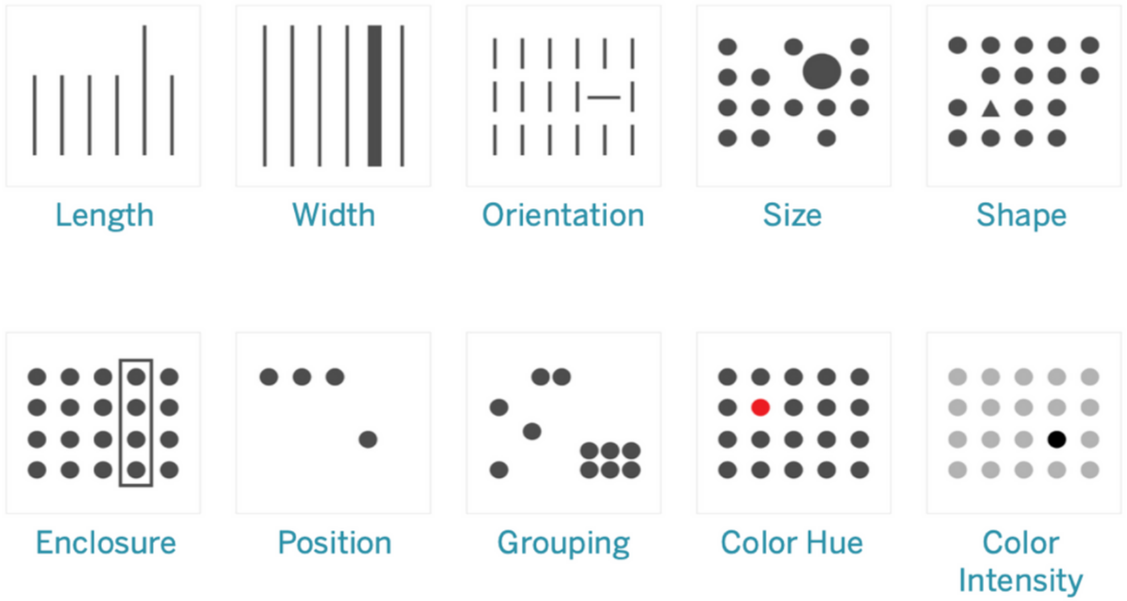

GUIDELINES



# CHART CLARITY

## LET THE IDEA COME THROUGH

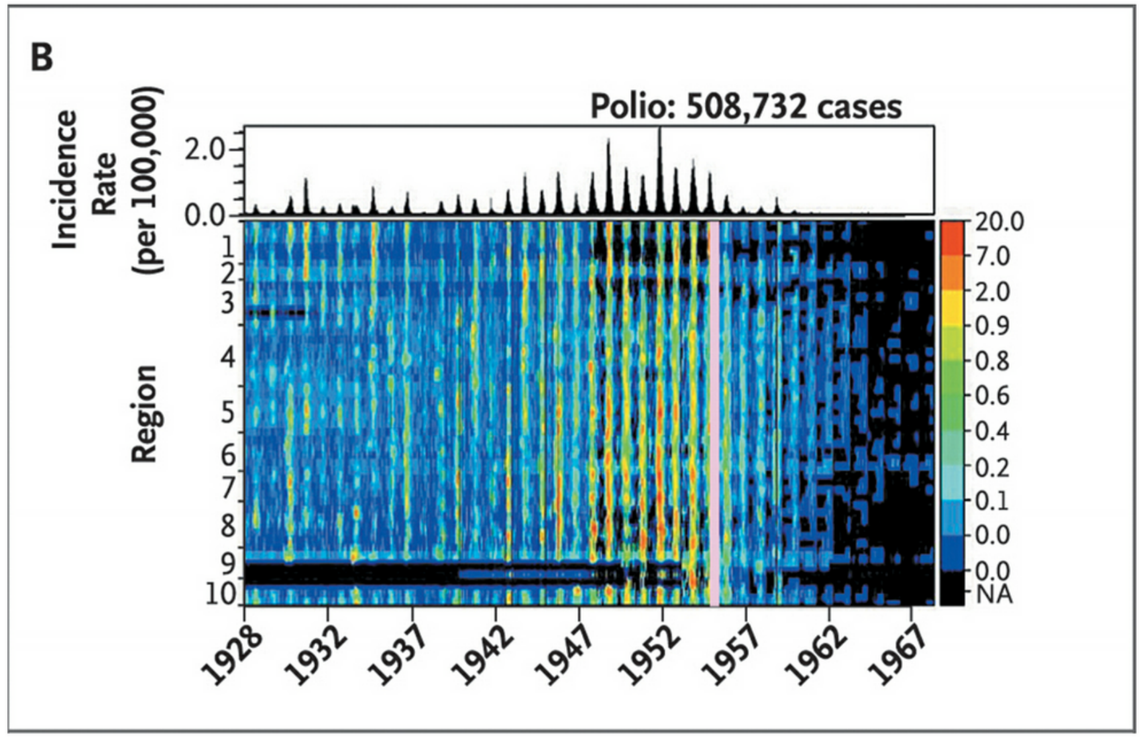
Our brains are designed to identify patterns, differences.




1. PRE-ATTENTIVE ATTRIBUTES
2. VISUAL ENCODING
3. COLOR MATCHES MEANING
4. GUIDE WITH YOUR TITLES

# CLARITY

## LET THE IDEA COME THROUGH



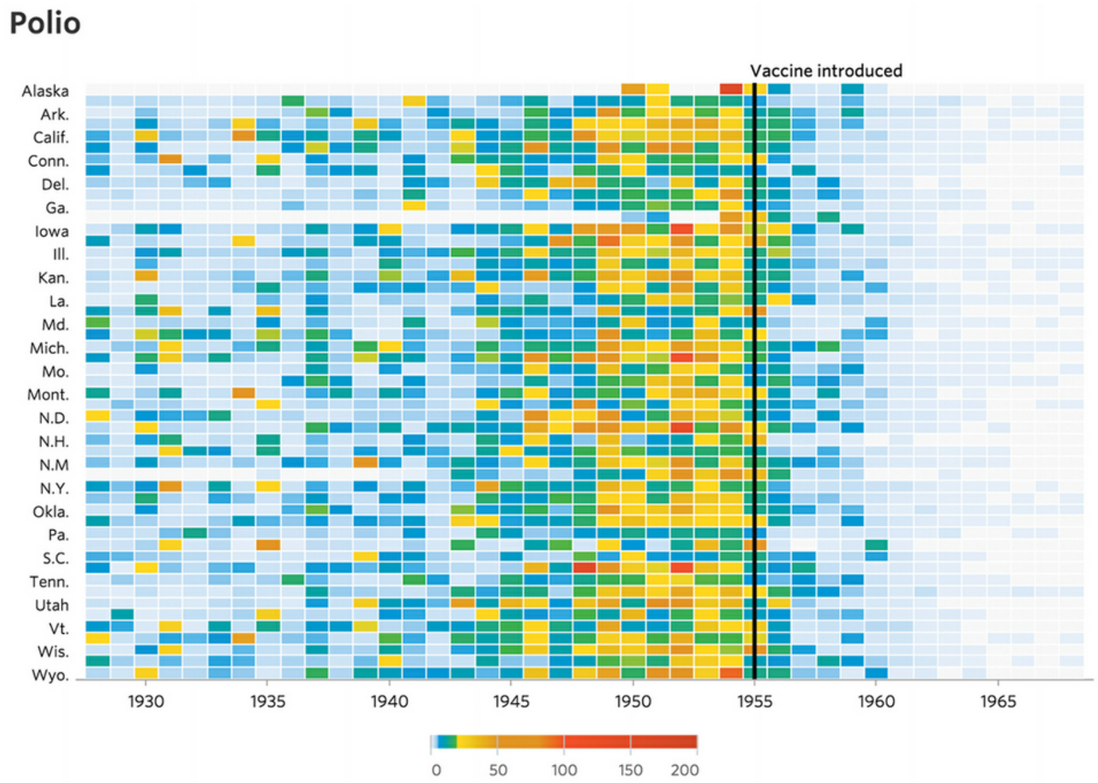

There's a lot of competition for our attention here.

1. PREATTENTIVE ATTRIBUTES
2. VISUAL ENCODING
3. COLOR MATCHES MEANING
4. GUIDE WITH YOUR TITLES

# CLARITY

## LET THE IDEA COME THROUGH

You want your audience to see your idea, not your charts.




1 PREATTENTIVE ATTRIBUTES

2 VISUAL ENCODING

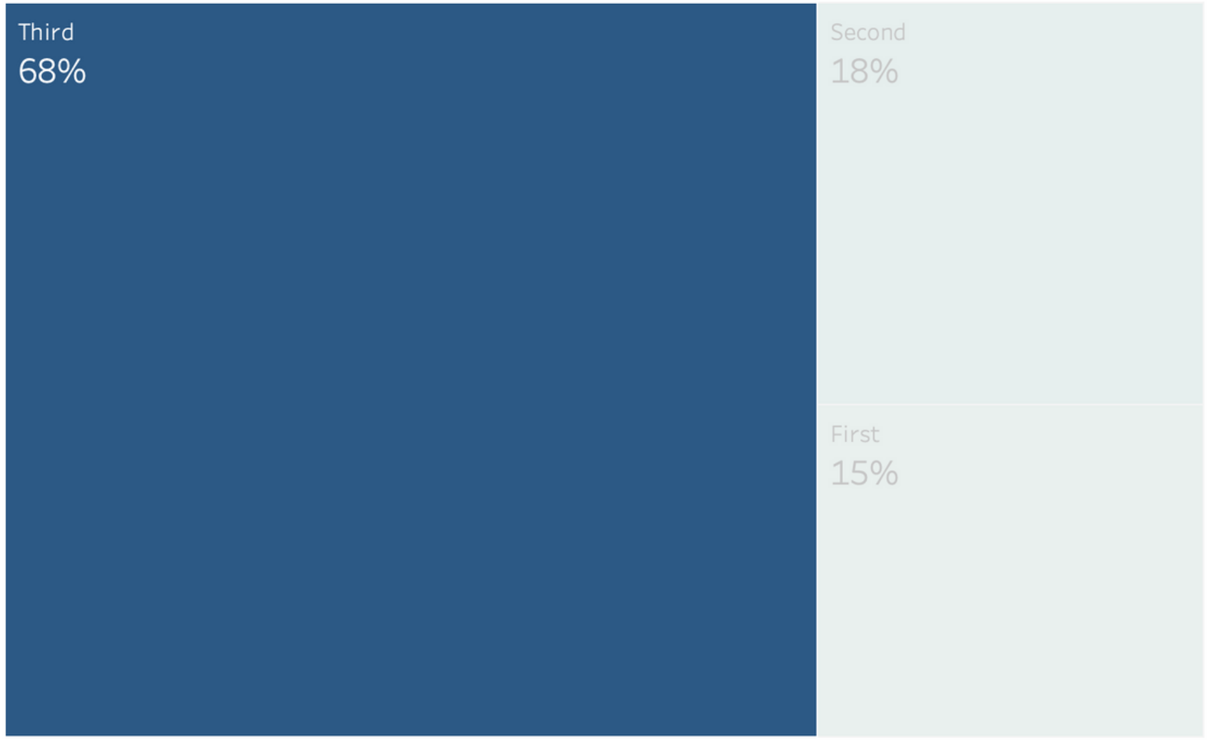
3 COLOR MATCHES MEANING

4 GUIDE WITH YOUR TITLES

# CLARITY

## LET THE IDEA COME THROUGH

Third Class Suffered Far More Loss of Life




Don't waste the opportunity to guide your audience with a meaningful title!

1 PREATTENTIVE ATTRIBUTES

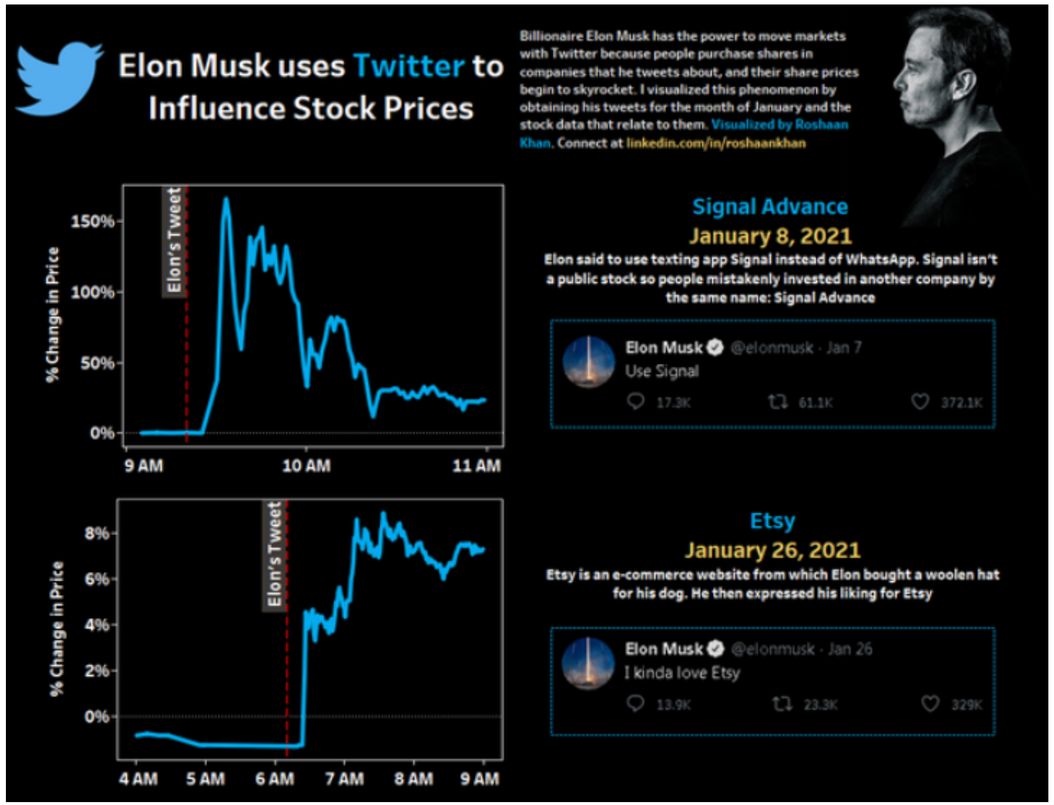
2 VISUAL ENCODING

3 COLOR MATCHES MEANING

4 GUIDE WITH YOUR TITLES

# CLARITY

## LET THE IDEA COME THROUGH




Titles allow you to direct the audiences attention and add valuable context.

1. PREATTENTIVE ATTRIBUTES
2. VISUAL ENCODING
3. COLOR MATCHES MEANING
4. GUIDE WITH YOUR TITLES

# SIMPLICITY

## CREATE EFFECTIVE DESIGNS

**"A visualization is an abstraction. labeling every value is a concretization."**
- Scott Berinato

**"Perfection is achieved, not when there is nothing more to add, but when there is nothing left to take away."**
- Antoine de Saint-Exupery

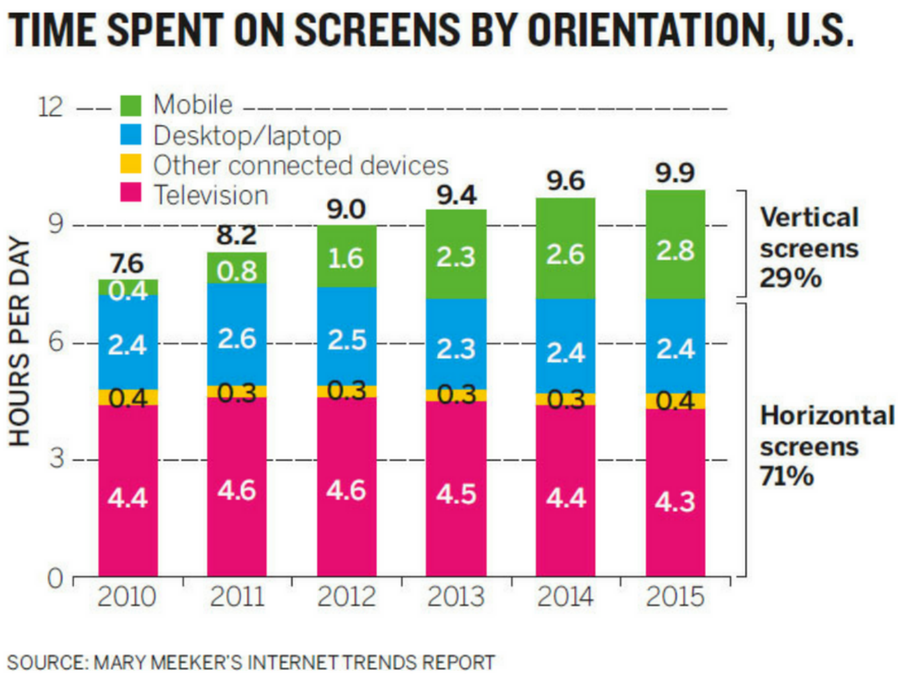
1. IS IT NECESSARY?
2. IS IT UNIQUE?
3. IS IT SIMPLE?

* "BELT AND SUSPENDERS DESIGN"

# SIMPLICITY

## CREATE EFFECTIVE DESIGNS

If you need every label, maybe what you need is a table, not a chart.



# SIMPLICITY

## CREATE EFFECTIVE DESIGNS

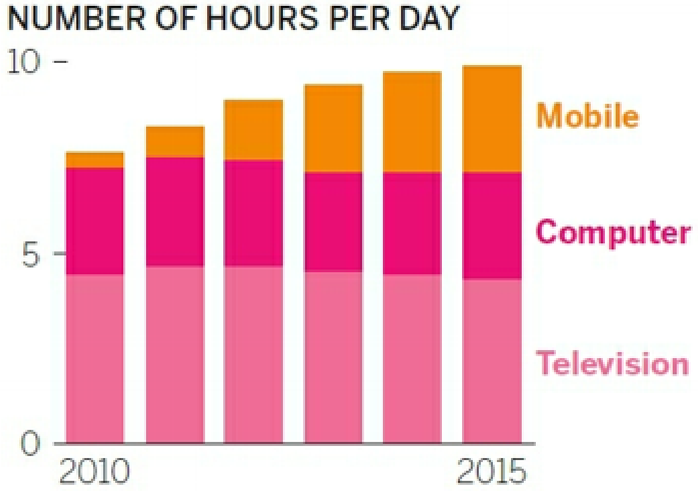
Or maybe you need a series of charts…

### HOURS SPENT ON SCREENS

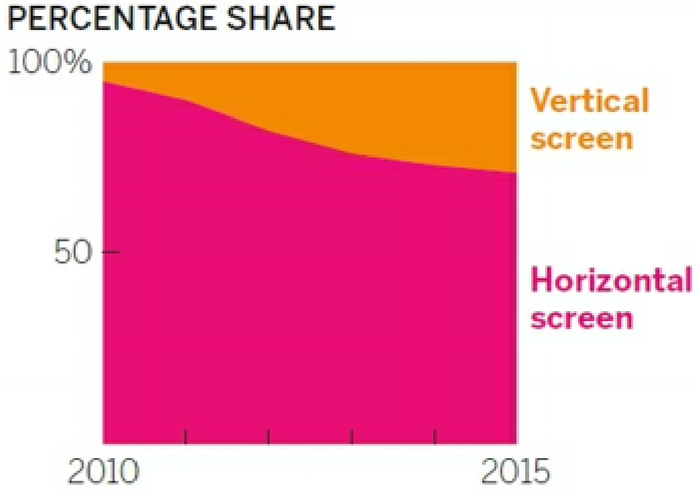
HOURS PER DAY SPENT ON SCREENS, U.S.

| | 2010 | 2015 |
|---|---|---|
| Television | 4.4 | 4.3 |
| Desktop/laptop/other | 2.8 | 2.8 |
| Mobile | 0.4 | 2.8 |
| **Total** | **7.6** | **9.9** |
| % Horizontal screens | 95 | 71 |
| % Vertical screens | 5 | 29 |

### SCREEN TIME IN THE U.S.



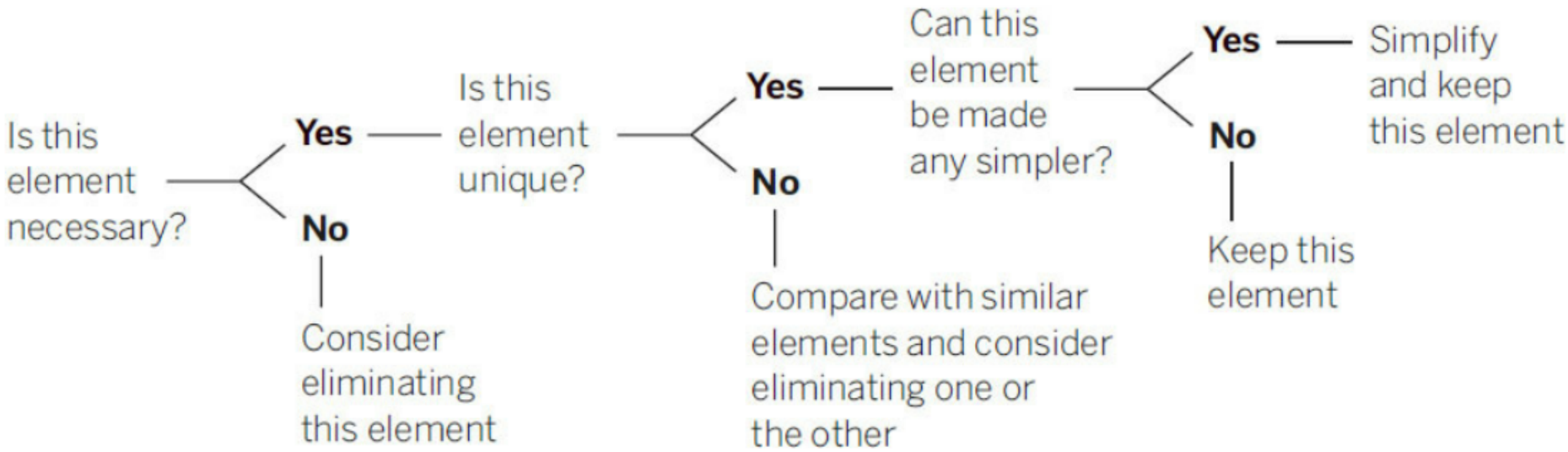

### TIME BY ORIENTATION



# SIMPLICITY

## CREATE EFFECTIVE DESIGNS

## WHICH ELEMENTS SHOULD YOU KEEP?



# SIMPLICITY

## CREATE EFFECTIVE DESIGNS





What do you need to get your idea across quickly, clearly?

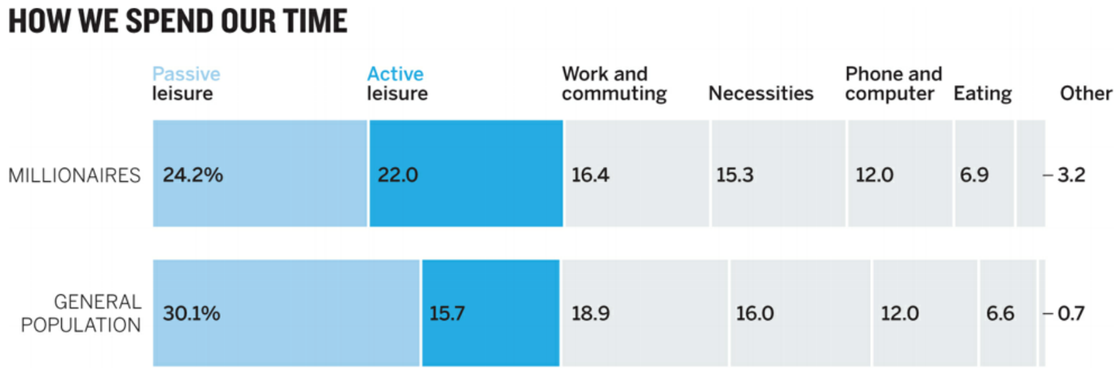
1. REMOVE UNNECESSARY DECIMALS
2. LIMIT COLORS
3. DO YOU NEED GRIDLINES?
4. MINIMIZE TICK MARKS, LABELS

# SIMPLICITY

## CREATE EFFECTIVE DESIGNS

What do you need to get your idea across quickly, clearly?



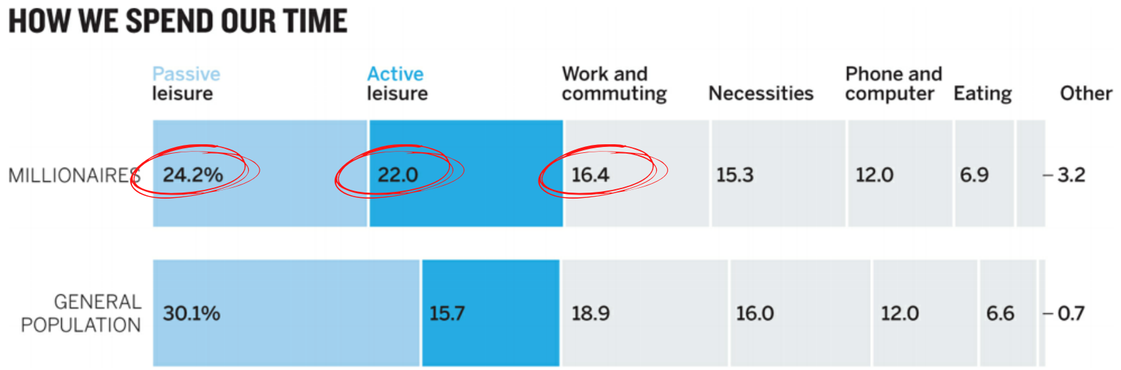

1. REMOVE UNNECESSARY DECIMALS
2. LIMIT COLORS
3. DO YOU NEED GRIDLINES?
4. MINIMIZE TICK MARKS, LABELS

# SIMPLICITY

## CREATE EFFECTIVE DESIGNS

What do you need to get your idea across quickly, clearly?



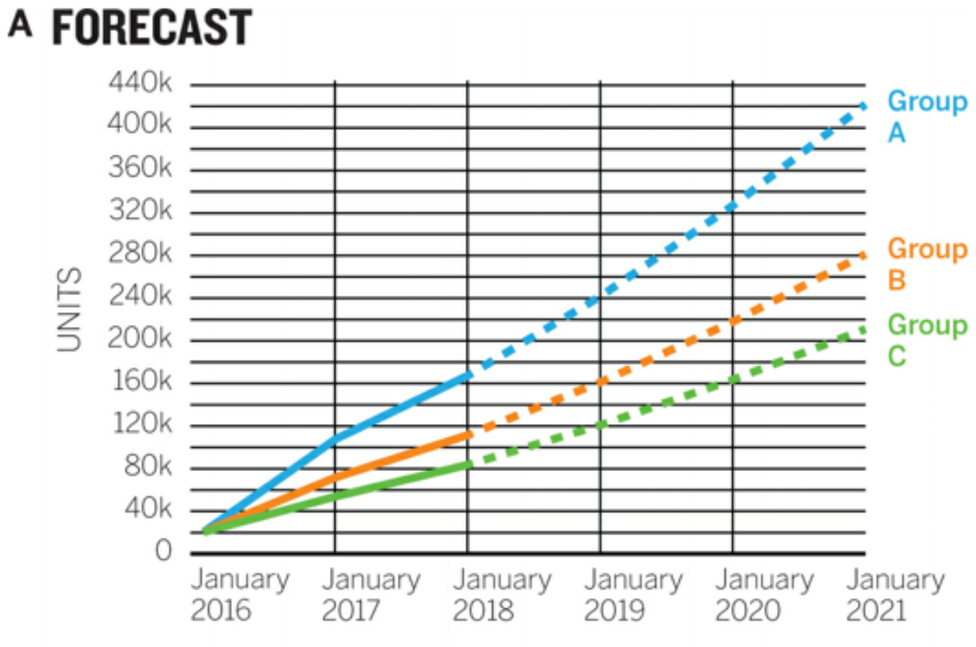




1. REMOVE UNNECESSARY DECIMALS
2. LIMIT COLORS
3. DO YOU NEED GRIDLINES?
4. MINIMIZE TICK MARKS, LABELS

# SIMPLICITY

## CREATE EFFECTIVE DESIGNS

What do you need to get your idea across quickly, clearly?

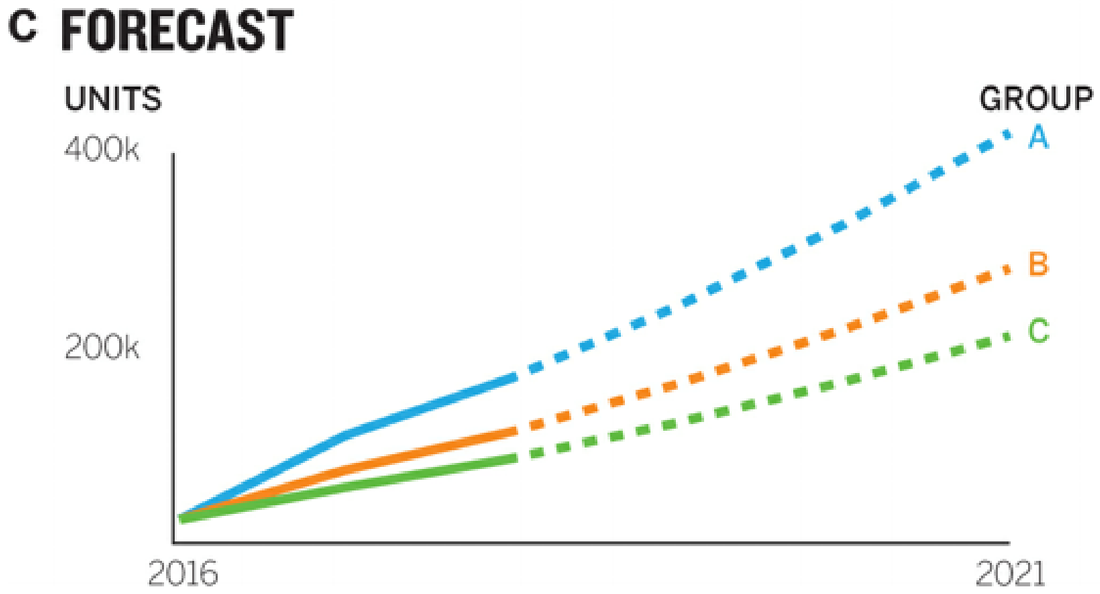



1. REMOVE UNNECESSARY DECIMALS
2. LIMIT COLORS
3. DO YOU NEED GRIDLINES?
4. MINIMIZE TICK MARKS, LABELS

# SIMPLICITY

## CREATE EFFECTIVE DESIGNS

What do you need to get your idea across quickly, clearly?




1. REMOVE UNNECESSARY DECIMALS
2. LIMIT COLORS
3. DO YOU NEED GRIDLINES?
4. MINIMIZE TICK MARKS, LABELS

# SLIDE DECK STRUCTURE

1. TITLE SLIDE
2. AGENDA SLIDE
3. EXECUTIVE SUMMARY SLIDE
4. FINDINGS VISUALIZATION SLIDE
5. RECOMMENDATION SLIDE
6. "WHAT IF" VISUALIZATION SLIDE
7. CONCLUSION AND NEXT STEPS SLIDE
8. APPENDIX SLIDES

# TITLE SLIDE ↘

ZILLOW TALK

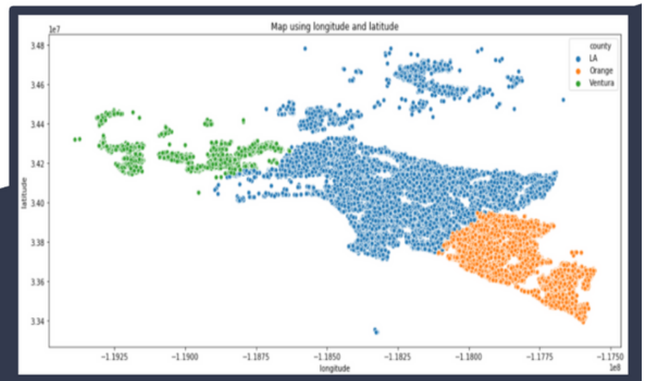
## Finding Value Around the Clock

*By Matthew Mays, Gilbert Noriega*
*Date: 10/08/2020*

## Property Tax Value Prediction Model

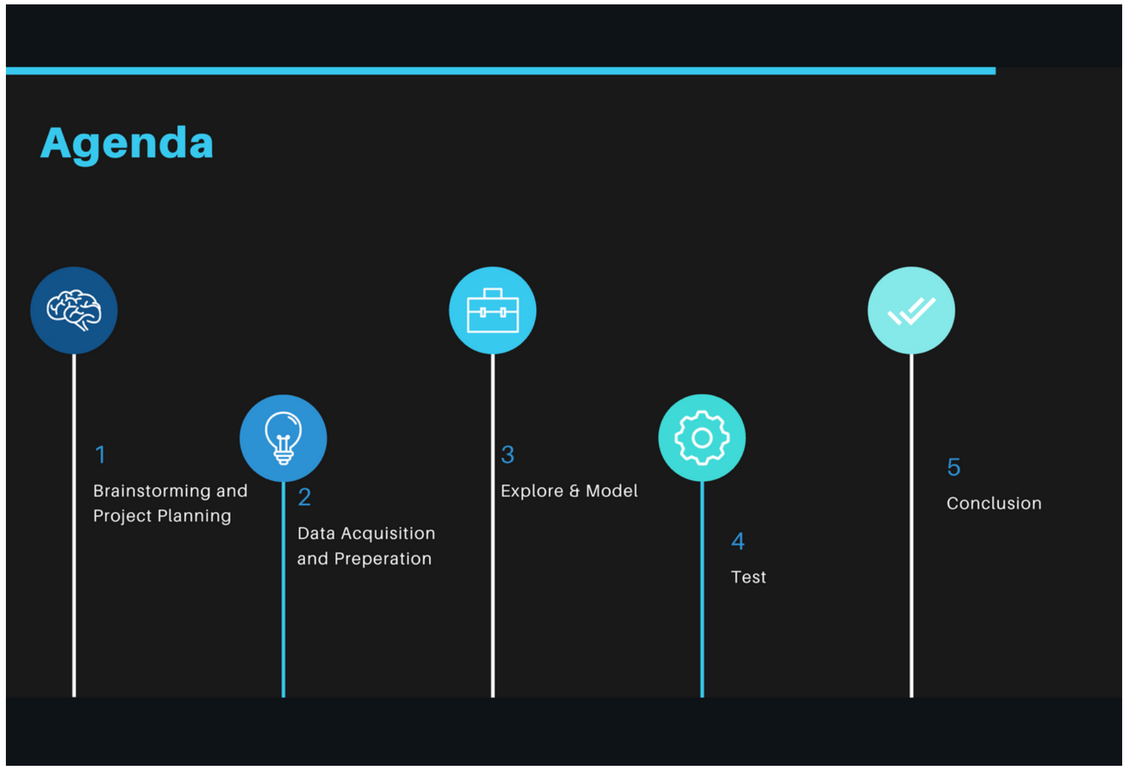
Zillow Regression Project

10/8/2020




CY Data Services Corey Solitaire, Rwvyn Young

# AGENDA SLIDE




## Agenda

EXECUTIVE SUMMARY

Quick summary of goals and analysis.

EXPLORATION

Our process to find the drivers of single unit property values.

CONCLUSION

Concluding our analysis as well as thoughts for next steps with the data set.

In the next 5 minutes...

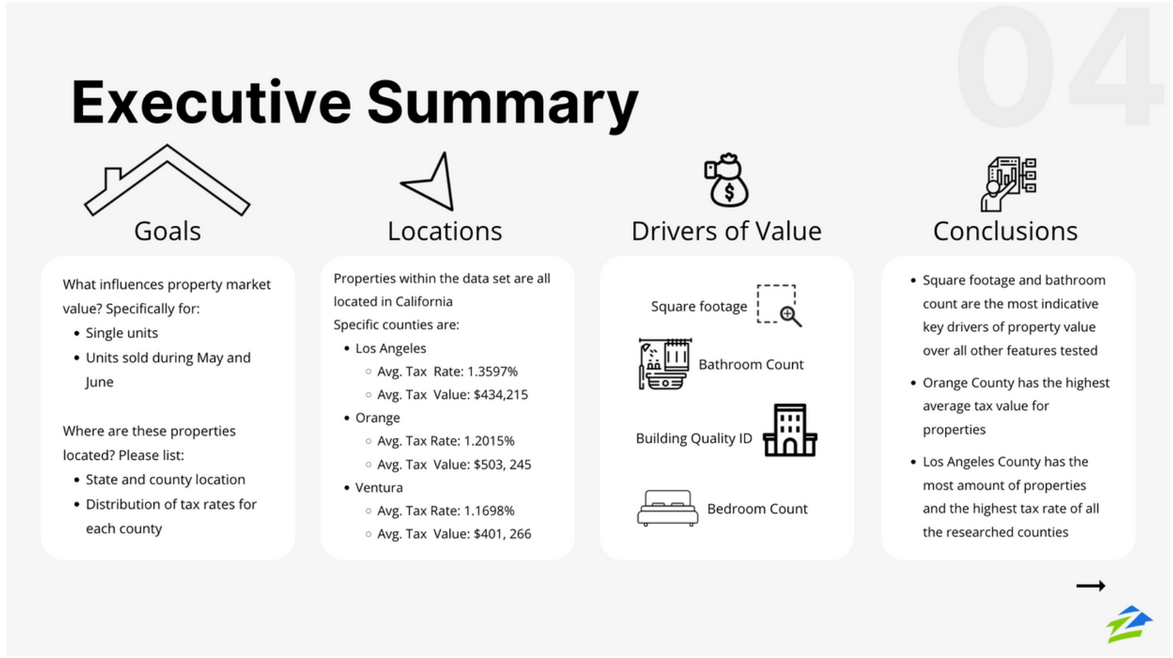
## OUR AGENDA:

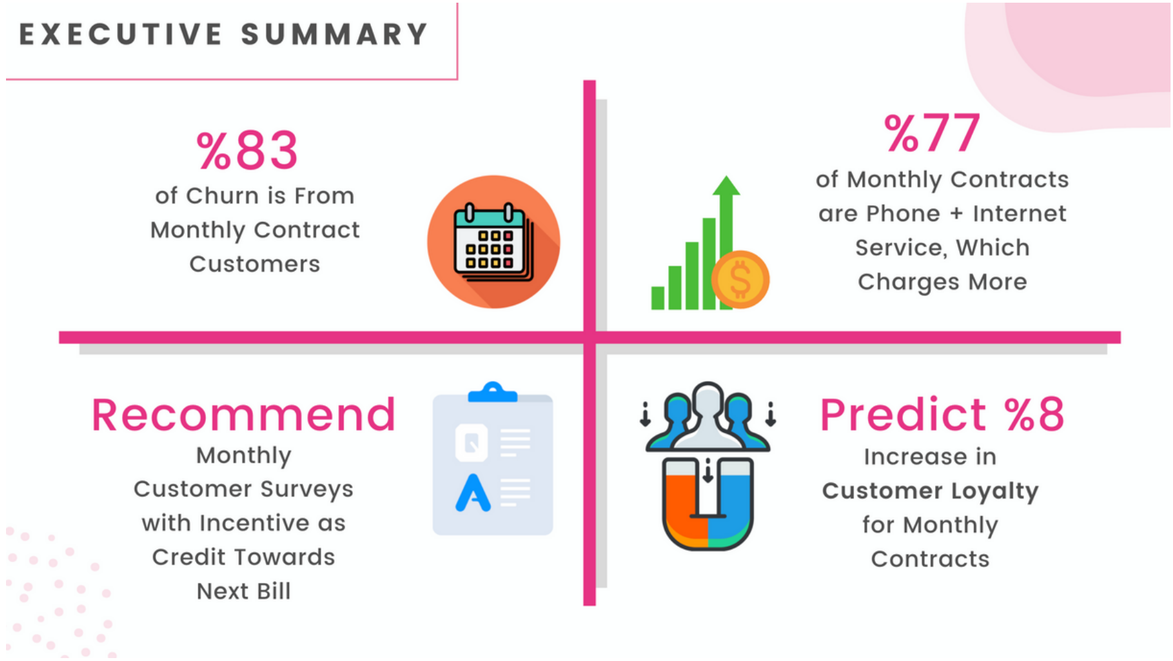
1. Define the **goals** of the project
2. Discuss **tax rate distribution** among the counties involved
3. Review the **data acquisition**, **preparation**, **exploration**, and **modeling**
4. Compare our model to the current '**Zestimate**' model being used
5. Cover key **takeaways** and **recommendations** for future improvements

An **appendix** is included at the end of the slides.

# EXECUTIVE SUMMARY SLIDE

1. YOUR BIG IDEA
2. GOAL(S)
3. KEY FINDINGS SUMMARIZED
4. RECOMMENDATION(S)



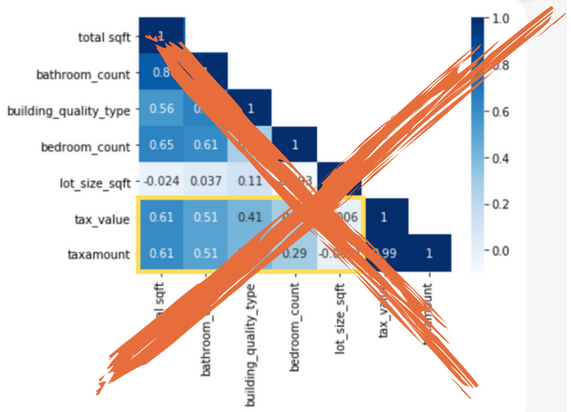
# FINDINGS VISUALIZATION SLIDE

06

## Key Drivers

- **Total Square Feet**
  - One Sample, Two Tailed T-test and Correlation Test
  - Rejected the null hypothesis
  - Tax value is dependent on square feet
- **Bathroom Count**
  - One Sample, Two Tailed T-test
  - Rejected the null hypothesis
  - Tax value is dependent on bathroom count
- **Building Quality Type**
  - Correlation heat map
- **Bedroom Count**
  - Correlation heat map

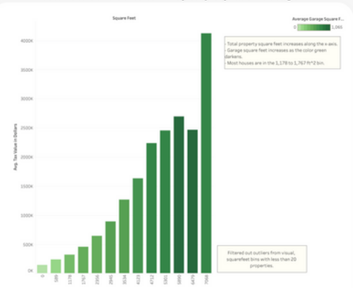
Correlation Heat Map




07

## Visuals of Top Drivers

Tax Value Amount vs. Property Square Footage



Click image to be directed to Tableau Workbook

**Property Square Feet**
- **As square footage increase, so does tax value**
- **Additional emphasis on garage square feet**

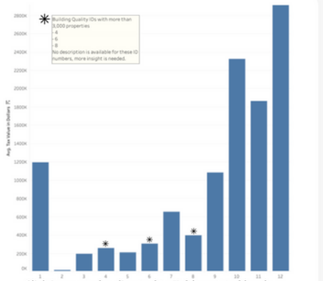
Bedroom + Bathroom Count
- Increase number of rooms correlates with increase of square feet
- Additional emphasis on total square feet

Building Quality Type
- Different quality ids have an influence on tax value
- No information on id descriptions

07

## Visuals of Top Drivers

Tax Value Amount vs. Building Quality Ids



Click image to be directed to Tableau Workbook

Property Square Feet
- As square footage increase, so does tax value
- Additional emphasis on garage square feet

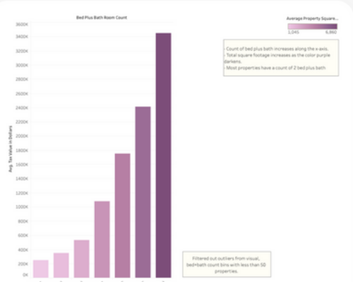
Bedroom + Bathroom Count
- Increase number of rooms correlates with increase of square feet
- Additional emphasis on total square feet

**Building Quality Type**
- **Different quality ids have an influence on tax value**
- **No information on id descriptions**

07

## Visuals of Top Drivers

Tax Value Amount vs. Bed+Bath Count



Click image to be directed to Tableau Workbook

Property Square Feet
- As square footage increase, so does tax value
- Additional emphasis on garage square feet

**Bedroom + Bathroom Count**
- **Increase number of rooms correlates with increase of square feet**
- **Additional emphasis on total square feet**

Building Quality Type
- Different quality ids have an influence on tax value
- No information on id descriptions

# RECOMMENDATIONS SLIDE

## Recommendation

Give current and future month to month customers who pay for both our phone and internet services a $20 price reduction on their monthly charges as a bundle deal.


This puts it in line with the cost of only paying for phone and only paying for internet combined, while also keeping customers under the average payment that they would start becoming more likely to churn.

## Let's get this bread!

My official recommendation to Telco is this:

Offer a trial price of 40% off (similar price to what customers are already paying for DSL) for the first month of upgrading to Fiber Optic from DSL if a customer signs up with a family member.

# "WHAT IF" VISUALIZATION SLIDE

| Expectation Level | Rate of Churn | Customers Retained (of 7043) |
|---|---|---|
| *Current* | *20%* | 5634 |
| *Low* | *15%* | 5986 (+352) |
| *Average* | *10%* | 6339 (+705) |
| *High* | *5%* | 6691 (+1057) |

Expected Result

## Predictions

USING WORST BEST-CASE

+19%

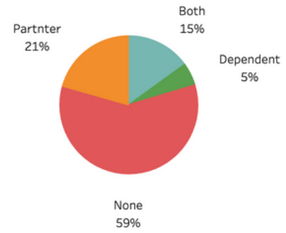
Percentage of Monthly Customers that Churn: Current to Prediction

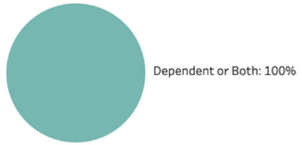

Taking %19 of customers that churn (as the "unhappy customers" from previous slides), and predicting they remain loyal when able to complain though surveys.

## What if we optimize our market for month-tomonth contract?

### Recommendation

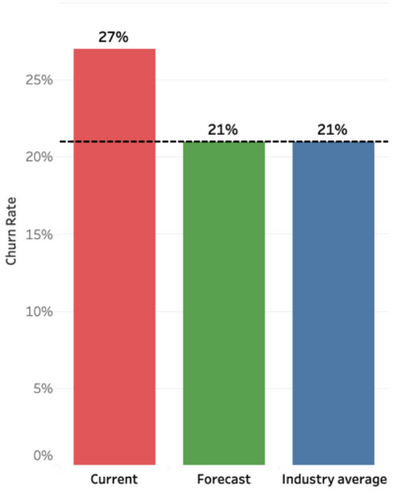
Target customers with dependents or both for month-to-month contracts




### Forecast

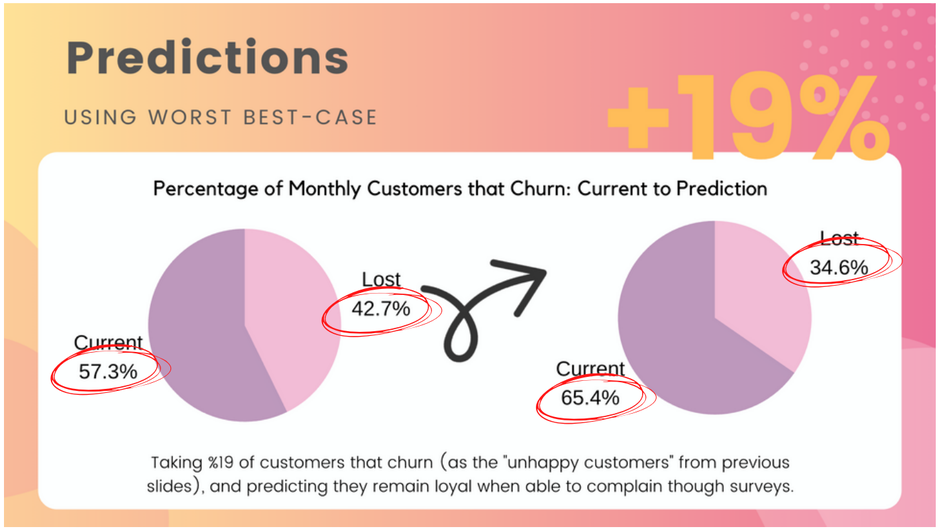
Telco's churn rate will reduce by 6% and will level with industry average



# "WHAT IF" VISUALIZATION SLIDE

| Expectation Level | Rate of Churn | Customers Retained (of 7043) |
|---|---|---|
| *Current* | *20%* | 5634 |
| *Low* | *15%* | 5986 (+352) |
| *Average* | *10%* | 6339 (+705) |
| *High* | *5%* | 6691 (+1057) |

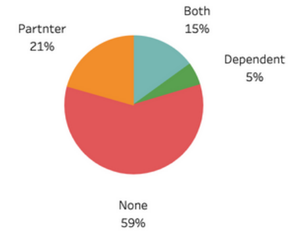
Expected Result

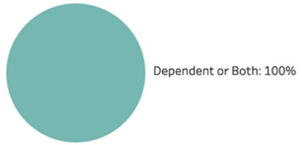



## What if we optimize our market for month-tomonth contract?

### Recommendation

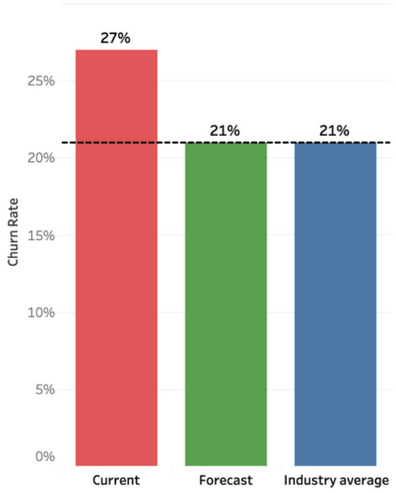
Target customers with dependents or both for month-to-month contracts




### Forecast

Telco's churn rate will reduce by 6% and will level with industry average

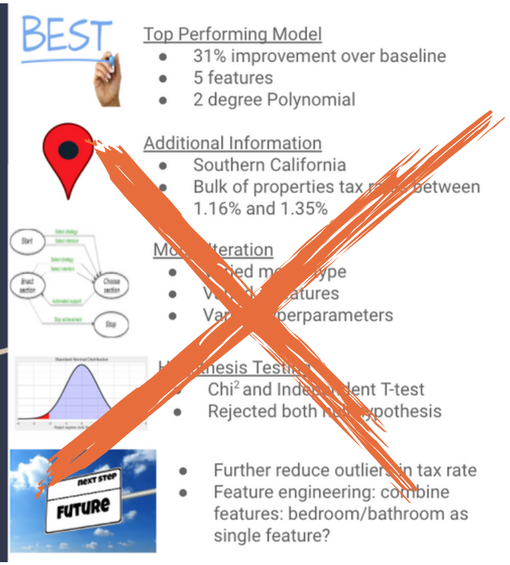


# CONCLUSIONS AND NEXT STEPS SLIDE

## Conclusions

| BATHROOM/ BEDROOM COUNTS | SQUARE FEET | ROOM COUNT | NEXT STEPS |
|---|---|---|---|
| The count of bathrooms and bedrooms in the property. | Total square footage of the property. | A count of all rooms including bed/bathroom. (Counts rooms other than bed/bath) | A large portion of the data had to be scrapped due to the lack of information. |

## Conclusions & Next Steps




08

## Conclusions

### Key Drivers

Listed from high to low correlation

1. Square Footage
2. Bathroom Count
3. Building Quality
4. Bedroom Count

### Observations

- Most units are 3 bedrooms/ 2 bath
- Whether a property has a garage, it greatly affects that total square footage
- Orange County has the highest average tax value for properties
- Los Angeles County has the most amount of properties and the highest tax rate of all the researched counties

### Next Steps

Do further investigation into 'Building Quality Type ID'

Continue research of price value changes over time

# APPENDIX SLIDES

## Appendix

| Column | Description |
|---|---|
| id | Auto-incremented unique index id for each property |
| bathroomcnt | Number of Bathrooms; Includes half baths as 0.5 |
| bedroomcnt | Number of Bedrooms |
| calculatedbathnbr | Precise meaning unknown, but appears to be redundant with bathroomcnt and bedroomcnt |
| calculatedfinishedsquarefeet | Total square feet of home; doesn't include property square feet |
| finishedsquarefeet12 | Unknown, but appears to be redundant with calculatedfinishedsquarefeet |
| fips | Federal Information Processing System codes used to identify unique geographical areas |
| fullbathcnt | Number of full bathrooms |
| latitude | The latitude of the property, not accurate to a global map, but relative difference is preserved |
| longitude | The longitude of the property, not accurate to a global map, but relative difference is preserved |
| lotsizesquarefeet | The size of the total property lot |
| propertycountylandusecode | Unknown, but represents categorical government code |
| propertylandusetypeid | Categorical variable describing the general type of property |
| rawcensustractandblock | Government id for each property linked to geographic location |
| regionidcity | Categorical variable identifying geographic location |
| regionidcounty | Categorical variable identifying geographic location |
| roomcnt | Number of rooms |
| yearbuilt | The year the house was built |
| structuretaxvaluedollarcnt | The tax assessed value of only the property structure in USD |

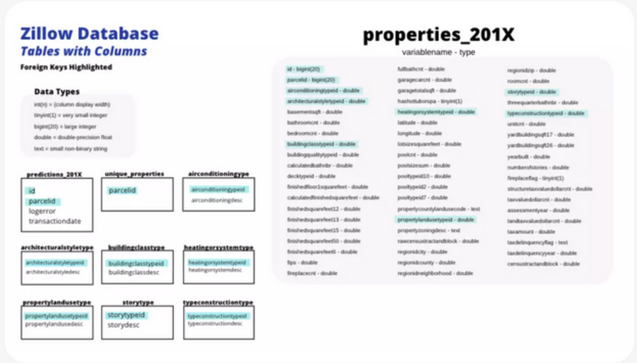
## Appendix: 1

Data analysis was done with the Zillow database for Los Angeles, CA properties in 2017

- **Total database:** 2,985,217
- **Variable columns:** 52

Acquired through Python by:

- Connecting to SQL database
- Query selected single-unit properties sold in 2017 during May and June
- Saved to .csv file




## Appendix: 2

- Using Tableau, we plotted properties by location
- To find location, used longitude and latitude
- Confirmed accuracy with FIPS code and Tableau map lines

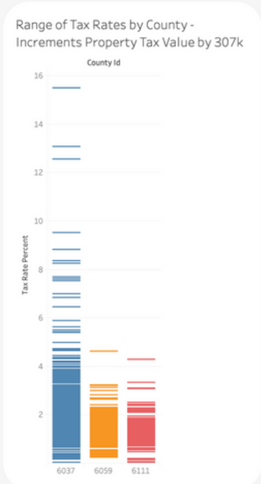
FIPS Codes
Los Angeles County : #6037
Orange County : #6059
Ventura County : #6111

- Calculated tax rate percentage and its distribution by county

$$\text{Tax Rate \%} = \left\{ \frac{\text{tax paid \$}}{\text{property value \$}} \right\} * 100$$

- Determined tax rates by county
- Grouped by FIPS code
- Used average of taxvaluedollarcnt

**Distribution of Tax Rates within Counties by Property Tax Rates**



# PRACTICE
# PRACTICE
# PRACTICE
# PRACTICE

1. PLAN FOR YOUR AUDIENCE
2. SCRIPT IT OUT - DON'T READ
3. TIME IT
4. LISTEN FOR FILLER WORDS