DATA ANALYTIC SERVICES - TRAINING

DATA NARRATIVES

Pivot Tables
Intro to Statistics
Data Narratives

Matthew Morris
Git: DasTraining
Msmorris@costco.com

Previously in Data Analytics

Data Analytic Summary
Iseries Options Source
Validating Data
Understanding your data

Matthew Morris
Git: Morrisdata
Msmorris@costco.com

PIVOT TABLE DEMO

DATA ANALYTICS

INTRO TO STATISTICS

Data

Describe, Categorize, Interpret Data

Data

Personal budgets
Events in town
Likes Dislikes
Home or Office inventories
Diet & Exercise
Etc.

Data Analyst

Data Set

Observational Units
Variables

Dimensions
Measures

Analysis

Data Set

Population
Sample

Analysis

STATISTICS

Collecting Analyzing Interpreting

Descriptive

↓

Know all the data

Inferential

↓

Use a Sample of Data
to make guesses about
all the data

Analysis

Descriptive

Count: 40
Karaoke: 30
Bowling: 10
Karaoke: 75%
Bowling: 25%

Inferential

Count: 210
Karaoke: 141
Bowling: 69
Karaoke: 67%
Bowling: 33%

Analysis

Inferential

Count: 210
Karaoke: 141
Bowling: 69
Karaoke: 67%
Bowling: 33%

How close?

Karaoke: 67% + 3%

70% ⟷ 64%

How confident?

95%

"95% certainty that Karaoke has an approval rating among students of 67% plus or minus 3% "

Analysis

MEAN (average) add up all the numbers and divid by the number of numbers you added up. 1,3,3,5
(1+3+3+5)/4 = 12/4=3

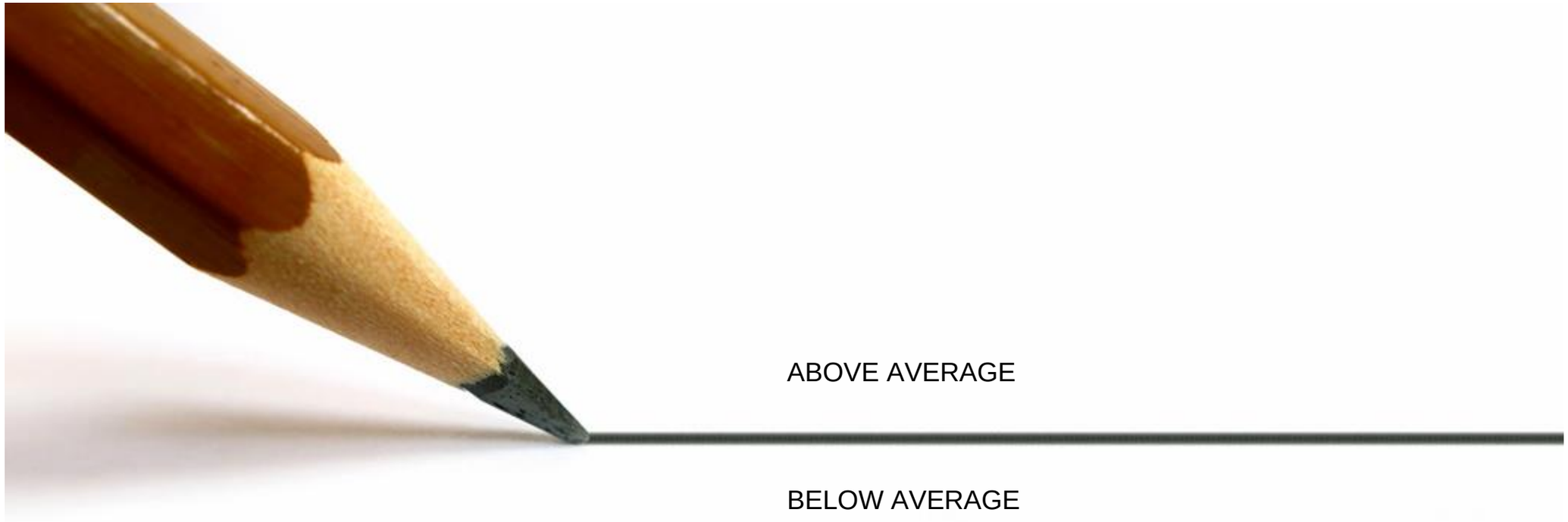
MEDIAN (middle) sort numbers highest to lowest. Then take the middle number as the Median

Odd Numbers 1,2,5,7,8
Even Numbers - Take the average of the two middle numbers 1,2,5,7,8,10
(5+7)/2= 12/2=6

MODE (most often) count the number that appears most often. If there are no repeats, there is no mode.

13, 18, 13, 14, 16, 13, 20, 15, 18, 13 The mode = 13

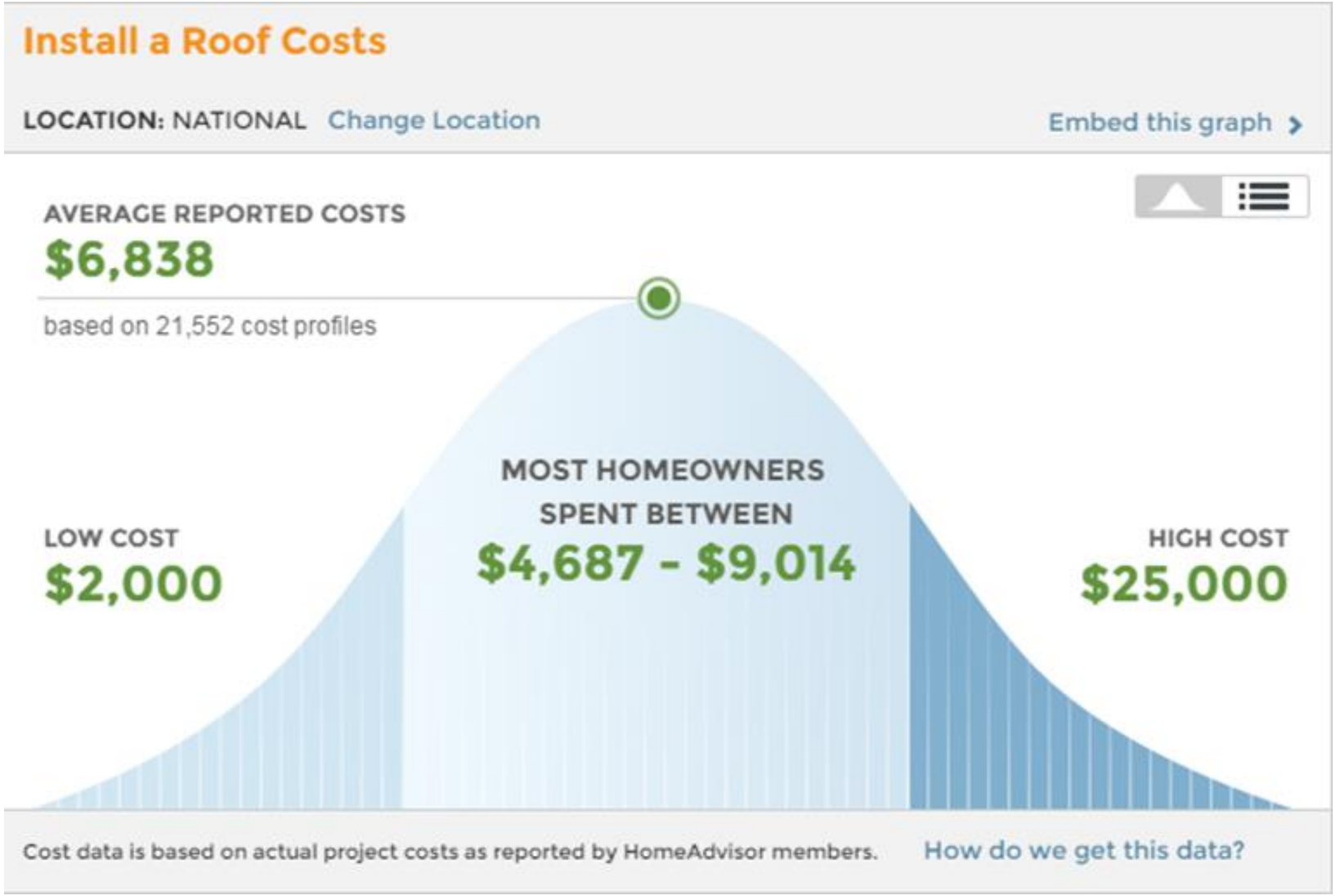
Data Analyst



Data Distribution

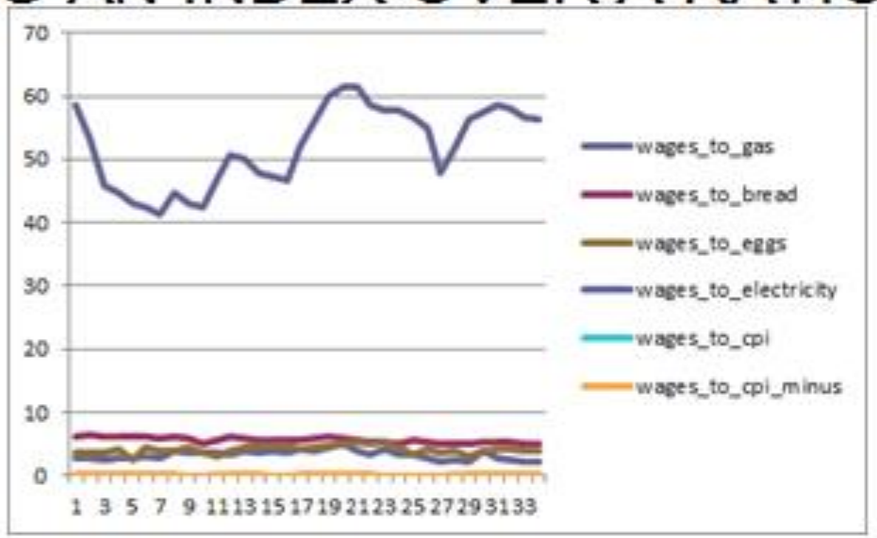
		FREQUENCY			
Score Bins	TALLY	FREQUENCY	RELATIVE FREQUENCY	CUMULATIVE FREQUENCY	RELATIVE AND CUMULATIVE FREQUENCY
90-99	IIII	4	4/20=20%	4	4/20 = 20%
80-89	~~IIII~~	5	5/20=25%	9	9/20 = 45%
70-79	~~IIII~~ I	6	6/20=30%	15	15/20 = 75%
60-69	IIII	4	4/20=20%	19	19/20 = 95%
50-59	I	1	1/20=5%	20	20/20 = 100%

Data Analyst



Data Analyst

CREATING AN INDEX OVER A RATIO

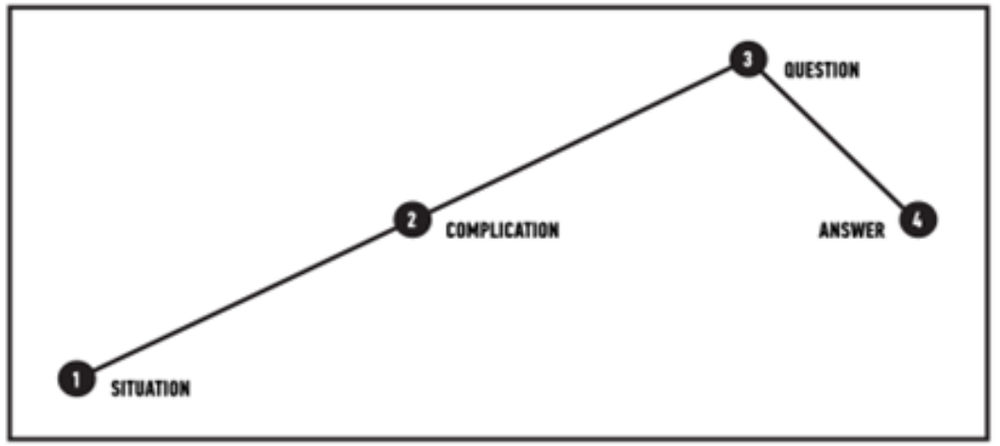


DATA NARRATIVES

Data Analyst

The traditional narrative arc is a linear story, consisting of four elements:

Situation	Complication	Question	Answer
Explains where we are now.	Creates tension in the story you're telling; triggers the Question you will ask.	Asks what we should do now given the Complication.	The Answer to the Question is the substance of your presentation.



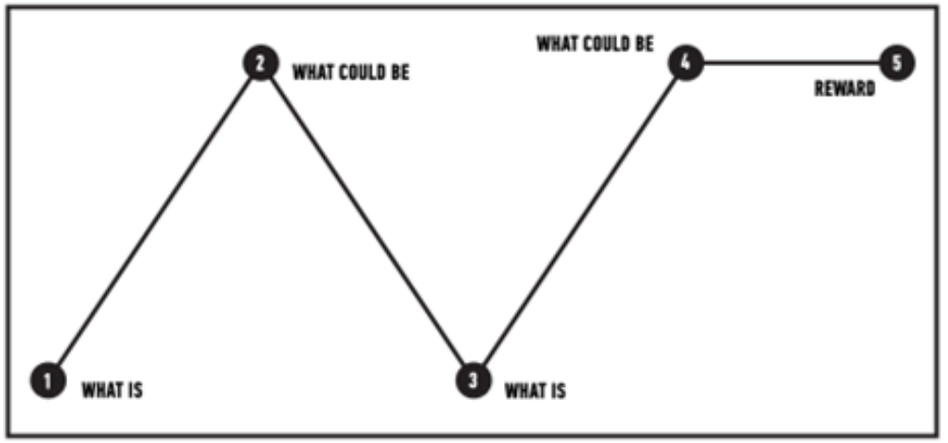
- As you can see, this pattern maps pretty closely to our Story Map:
  - Situation to setting
  - Complication to problem
  - Answer to resolution

Data Analyst

Data Analyst

The "What is vs. What could be" is a non-linear story, consisting of three elements:

What Is	What Could Be	Reward
The current undesirable situation.	A utopian future where the original problem of "what is" no longer exists.	The future situation that could exist if we all believe in it. This is your call to action.



- Here your findings identify "What Is."
- Your motivation, and your selling point to your audience, is "What Could Be."
- Your next steps will allow the "What Could Be"
  - to become a reality, leading to this "Reward."

Data Analyst

PROBLEM

APPROACH

SOLUTION/NEXT STEPS

Data Analyst

PRESENTATION OBJECTIVES What does your presentation need to accomplish?		PRESENTATION CONTENT How will your presentation fit both needs?
AUDIENCE SEGMENTS What describes your audience & their enrollment?	AUDIENCE OBJECTIVES What does your audience need from your presentation?	

Q & A

“The goal is to turn data into information, and information into insight.”

-Carly Fiorina, prior CEO of Hewlett-Packard

Conclusion

- Identifying a problem will dictate the value of the answer
- Just cleaning data and creating percentages can add immense value to a dataset.
- Prepping data is a major part of data analytics.

FEEDBACK FORM

CLASS : DATA NARRATIVES

QUESTION: What are the steps of the Data Analytics Workflow?