# THE VALUE OF DATA + DATA NARRATIVES



# **LEARNING OBJECTIVES**

In today's lesson, we will:

- 1. Explain our Data Analysis course objectives and tools.
- 2. Explore the Data Framework and iterative problem-solving.
- 3. Discuss the value and impact of data-driven decision making.
- 4. Survey common presentation strategies.
- 5. Practice applying techniques to a sample business case.

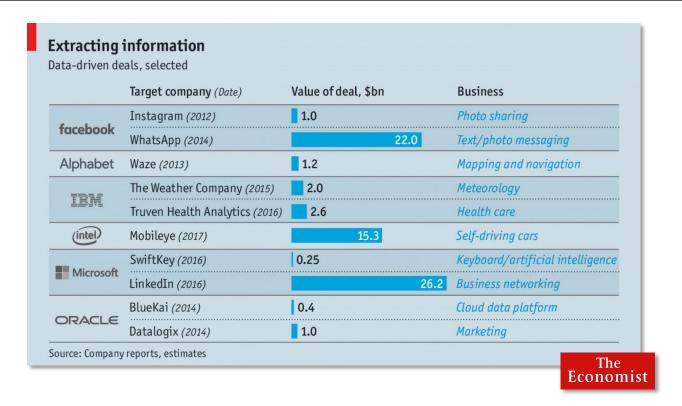
# THE VALUE OF DATA

# INTRODUCTION: DIVE INTO DATA

# THE VALUE OF DATA



# THE VALUE OF DATA: DEAL-MAKING



# **EXAMPLE: PEDESTRIAN ACCIDENT DATA**

# cambridge open data **CAMBRIDGE, MA** Unsaved View Save As... ased on Accident Map 2010-2013 List of accidents involving motor vehicles, bicycles and/or HOW ARE YOU TRAVELING? 1079 Commonwealth Ave. Boston MA > 45 Chestnut St, Cambridge MA HEADS UP! The accident risk for this route is MODERATE Be careful on Brookline St, especially the intersection of Brookline & WHEN ARE YOU Cambridgeport PLANNING TO GO? Insights? Actions? HEADS UP! The frequency of pedestrian accidents between 12pm & 1pm is HIGH CAMBRIDGE PEDESTRIAN ACCIDENTS (2010-2013) Question(s)?

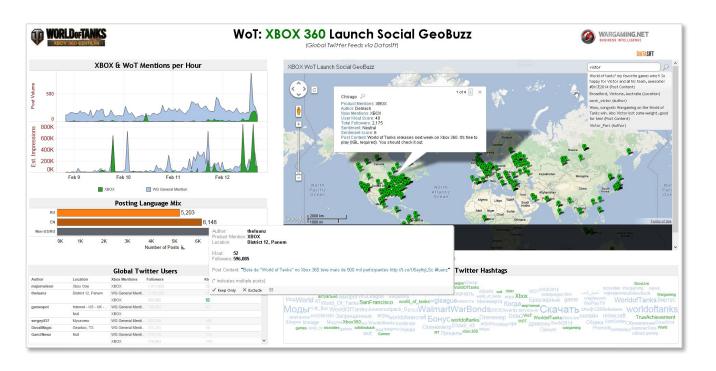
# **EXAMPLE: MOBILE DASHBOARDS**







# **EXAMPLE: SOCIAL MEDIA BUZZ**



# THE VALUE OF DATA: DATA-DRIVEN DECISIONS

Issue: Google AdWords spending analysis.



- CEO's question: "What is the ad spend ROI on 3,200 keywords?"
- Ad agency analysis: 98% of the revenue is coming from only 22 words.
- ▶ 22 words = "Restoration Hardware" + 21 misspelled variations.
- CEO's reply: "Cancel no need to buy our own name from Google."





# THE VALUE OF DATA: DATA-DRIVEN DECISIONS

Issue: American Express needs a data model to predict customer churn.





# THE VALUE OF DATA: DATA-DRIVEN DECISIONS

# **Goals & Outcomes:**



- Preserve 100M global cards, \$1T annual charges.
- Biz model: Targets affluent customers (\$150/charge vs. Visa \$50/charge).
- Developed big data churn forecast model with 115 variables.
- Flagged 24% of accounts to churn next qtr  $\rightarrow$  retention marketing.
- Also reduced personalization model from three days to 20 minutes.
- Machine learning increased online customer acquisition by 40%.

# **ACTIVITY: DATA-DRIVEN DECISION MAKING**



## **DIRECTIONS**

- 1. Pair up with the person next to you. Take five minutes to...
- 2. Discuss if and how your department views data as a strategic asset (either currently or at a former place you've worked).
- 3. Describe an example of a data-driven decision (*or the opportunity for one*) in a current or former business unit that you've been a part of.

BONUS: List 1–2 current blockers to making data-driven decisions.

# **DELIVERABLE**



Share out various examples of data-driven decisions from the class.



# THE VALUE OF DATA

# INTRODUCTION: HOW TO THINK ABOUT DATA



times from 2005

### It's estimated that 2.5 QUINTILLION BYTES

[ 2.3 TRILLION GIGABYTES ]











Most companies in the U.S. have at least

## 100 TERABYTES

100,000 GIGABYTES T of data stored

The New York Stock Exchange captures

**WORLD POPULATION: 7 BILLION** 

### 1 TB OF TRADE INFORMATION

during each trading session



**ANALYSIS OF** STREAMING DATA



### 18.9 BILLION NETWORK CONNECTIONS

- almost 2.5 connections per person on earth



Modern cars have close to 100 SENSORS

that monitor items such as fuel level and tire pressure

# **Velocity**

# FOUR V's of Big **Data**

The

break big data into four dimensions: Volume. Velocity, Variety and Veracity

customer needs, of mize operations and

### 4.4 MILLION IT JOBS

alue

As of 2011, the global size of data in healthcare was estimated to be

### 150 EXABYTES

[ 161 BILLION GIGABYTES ]



**Variety** 

DIFFERENT **FORMS OF DATA** 



are shared on Facebook every month







4 BILLION+ HOURS OF VIDEO

By 2014, it's anticipated

**HEALTH MONITORS** 

WEARABLE, WIRELESS

there will be

420 MILLION

are watched on YouTube each month



100 MILLION TWEETS

are sent per day by about 200 million monthly active users



don't trust the information they use to make decisions



Poor data quality costs the US economy around



27% OF

in one survey were unsure of how much of their data was inaccurate

**Veracity** 

**UNCERTAINTY** OF DATA

Gartner.



# THE FIVE V's OF DATA

- 1. Volume: Consider the scale of the data (big & small, structure).
- 2. Velocity: Understand data sources(s), timing, and flow.
- 3. Variety: What forms and types are required to answer questions?
- 4. Veracity: Verify quality, accuracy, and reliability of source(s).
- 5. Value: What are the metrics or measures for desired outcomes?

# INTRODUCTION: PRESENTATION STRATEGIES

# **DATA AS THE STORY**

Ask great questions...





...to reveal key findings.

A **story map** can be used to diagram your presentations ahead of time.

This can help you consider all of the relevant elements you may want to discuss, including:

- Setting / Time / Place
- o People Involved
- o Problem
- o Events
- Resolution

# **STORY MAP**

SETTING:	IIME:	PLACE:
	<u> </u>	
CHARACTERS:		
PROBLEM:		
·! _		
	PLOT/EVENTS:	
<del>+</del>		
RESOLUTION:		

# **Setting / Time / Place**

The specific business context:

- o Location of problem
- Where was data collected?
- What locations are involved?

# **STORY MAP** SETTING: TIME: PLACE: CHARACTERS: PROBLEM: PLOT/EVENTS: RESOLUTION:

# **Characters**

People and data involved:

- o Data dictionary?
- o Business unit owner?
- Metadata description?
- Issues of data governance?

Audience stakeholder(s) involved:

- Who is the presentation for?
- o Identify all key decision makers.

# **STORY MAP**

SETTING:	TIME:	PLACE:
		l

CHARACTERS:			

PROBLEM:		_
		_
I		

1	
! 1	PLOT/EVENTS:
 	PLUI/LULATIS:
1	
i	
•	

RESOLUTION:

# **Problem**

The issues and opportunities at-hand:

- Motivations? 0
- Pain points? Ο
- Opportunities? Ο
- Hypothesis?

# **STORY MAP**

SETTING:	TIME:	PLACE:
CHARACTERS:		
PROBLEM:		
!		
   <del> </del>	PLOT/EVENTS:	
1		
*		
RESOLUTION:		

# **Plot/Events**

Process, considerations, actions taken:

- What did you do to solve these problems?
- What was the timeline of your approach?
- Explain your process in terms your audience can easily understand

# STORY MAP

SETTING:	TIME:	PLACE:
CHARACTERS:		
PROBLEM:		
 	PLOT/EVENTS:	
RESOLUTION:		

# Resolution

What should be done?

- •What are your recommendations?
- •Include your assumptions.
- ∘Is any additional data needed?

Communicate your results

- Use visuals.
- Customize this for your audience.
- Make your data the focus.

# **STORY MAP**

I DIACE

I TIME

CETTING.

ETTINO:	III-L:	PLACE:
HARACTERS:		
ROBLEM:		
	PLOT/EVENTS:	
<b>← &gt;</b>	PLUI/LVERIS:	
100		
' '		
ESOLUTION:		
ESULUTION:		

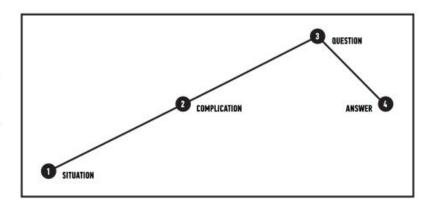
# **PRESENTATION CANVAS**

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?	

# TRADITIONAL NARRATIVE: INFORMING AN AUDIENCE

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

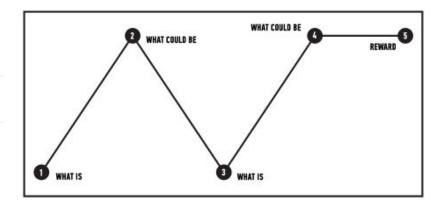


- "Situation:" This is the current state; define the problem(s).
- "Complication:" Contextualize the problem with details.
- "Question:" Given these barriers, what should be done?
- "Answer:" Your call to action or methods, framed as the solution.

# **NONLINEAR NARRATIVE: PERSUADING AN AUDIENCE**

The "What is vs. What could be" is a nonlinear 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.



- "What is:" Establish the problem at-hand.
- "What could be:" Communicate a vision, the desired future state.
- "Reward:" New opportunities that would be unlocked by this future state.
- Point: This structure frames your goals as an opportunity; connecting a specific (data-driven) future state to larger goals or initiatives.

# **Data Analytis Life Cycle**



# PRESENTATION STRATEGIES

- Point: Presentations = stories. Stories are a cultural framework that most audiences are already familiar with: setting, characters, problem, solution.
- Planning presentations in this manner can help you remember to describe and focus on the people, problems, and goals involved.
- Different story formats (linear vs non-linear) should be experimented with, depending on the results you want to achieve.
- Finally, the Data Analytics Life Cycle is not just a framework for solving problems, but can also be used to help you clearly organize your data presentations.

# DATA NARRATIVES: TELLING THE STORY WITH DATA

**Gapminder: Hans Rosling's 200 Countries, 200 Years, 4 Minutes — The Joy Of Stats** 



# INDEPENDENT PRACTICE: BUSINESS CASE

# **ACTIVITY: BUSINESS CASE QUESTIONS**

# **DIRECTIONS**



- 1. In small groups, review the "Data Narrative Business Case" file.
- 2. Collaborate and brainstorm on the key questions.
- 3. Prepare your findings for class discussion.

# **DELIVERABLE**

Identify the costs, benefits, and risks of the scenario.

# CONCLUSION

Q&A

# RESOURCES



# **RESOURCES**

For additional information on chart type selection, look for the following handouts in the <u>student resource directory</u>:

- "Whitepaper: Which Chart Type is Right for You? (Tableau)"
- "How to Choose the Right Data Visualization (Sisense)"