The **Power** of Data Visualization

Perception and Knowledge

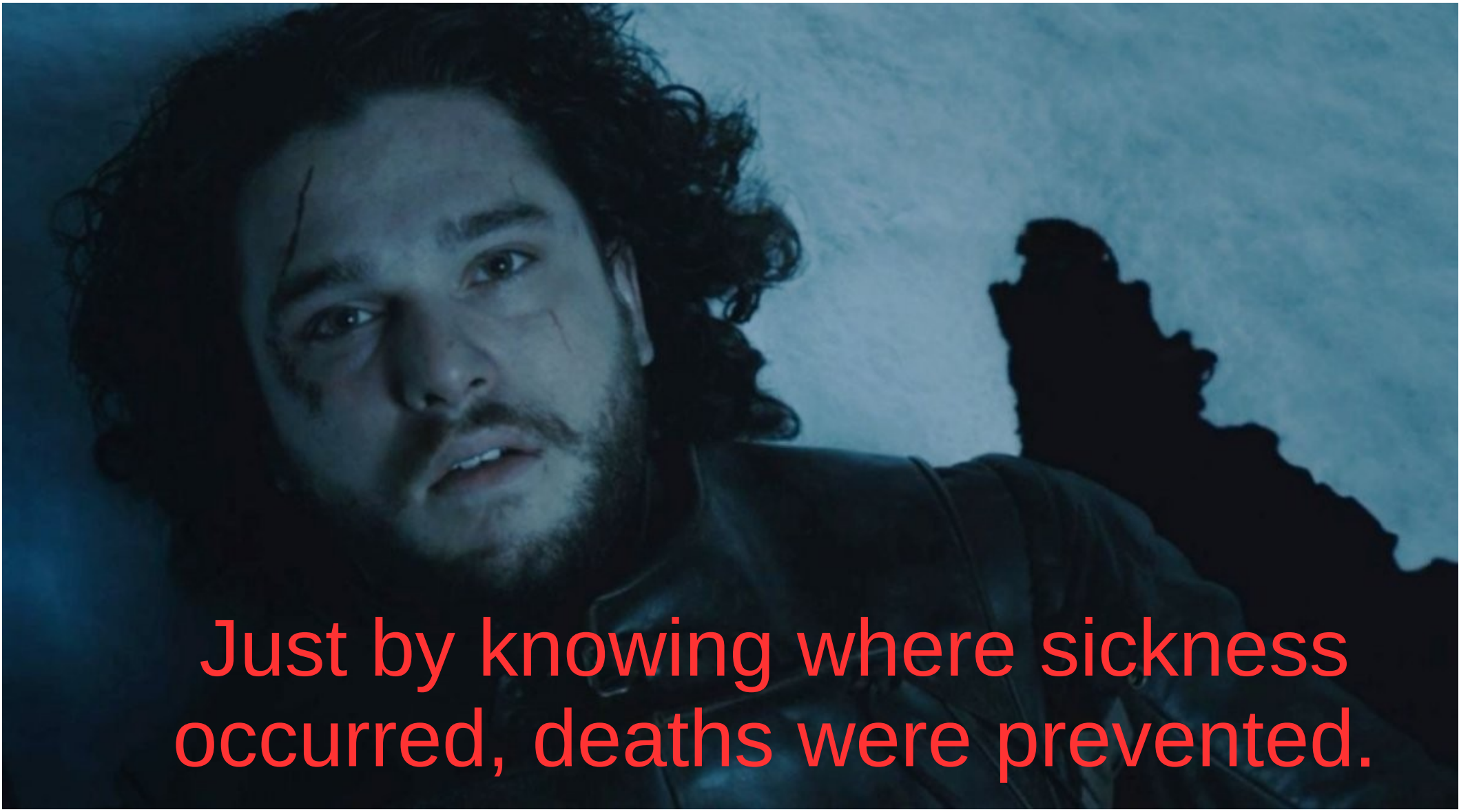
Why “Data Viz”?

- Two components to every lab so far:
  - Analytics
  - Visualization
- Remember what we did for each?
  - Why W&M
  - Island Villas
  - Business Analytics

Why bother with Data Visualization?

- Let it be known: Ya'll *hate* being forced to be creative.
  - *“More instructions on what to include vs. say "be creative"*
  - *“Ambiguity; not a lot of instruction is given”*
  - *“Visual appeal requirements for labs..”*
  - *“The Q mapmaker was pretty terrible to make a good looking map...”*

But, Data Viz Alone can Change the World


Just by knowing where sickness occurred, deaths were prevented.

Hans Rosling

- www.bit.ly/1Q3J0l1
- www.bit.ly/1fsOp6b

The World's 100 Most Influential People: 2012

They entertain us, lead us and challenge us. TIME honors the icons who are defining the world in 2012

Tweet 0 G+1 40 Share

TIME 100: THE LIST

Hans Rosling

Statistician

By Nicholas A. Christakis | Wednesday, Apr. 18, 2012

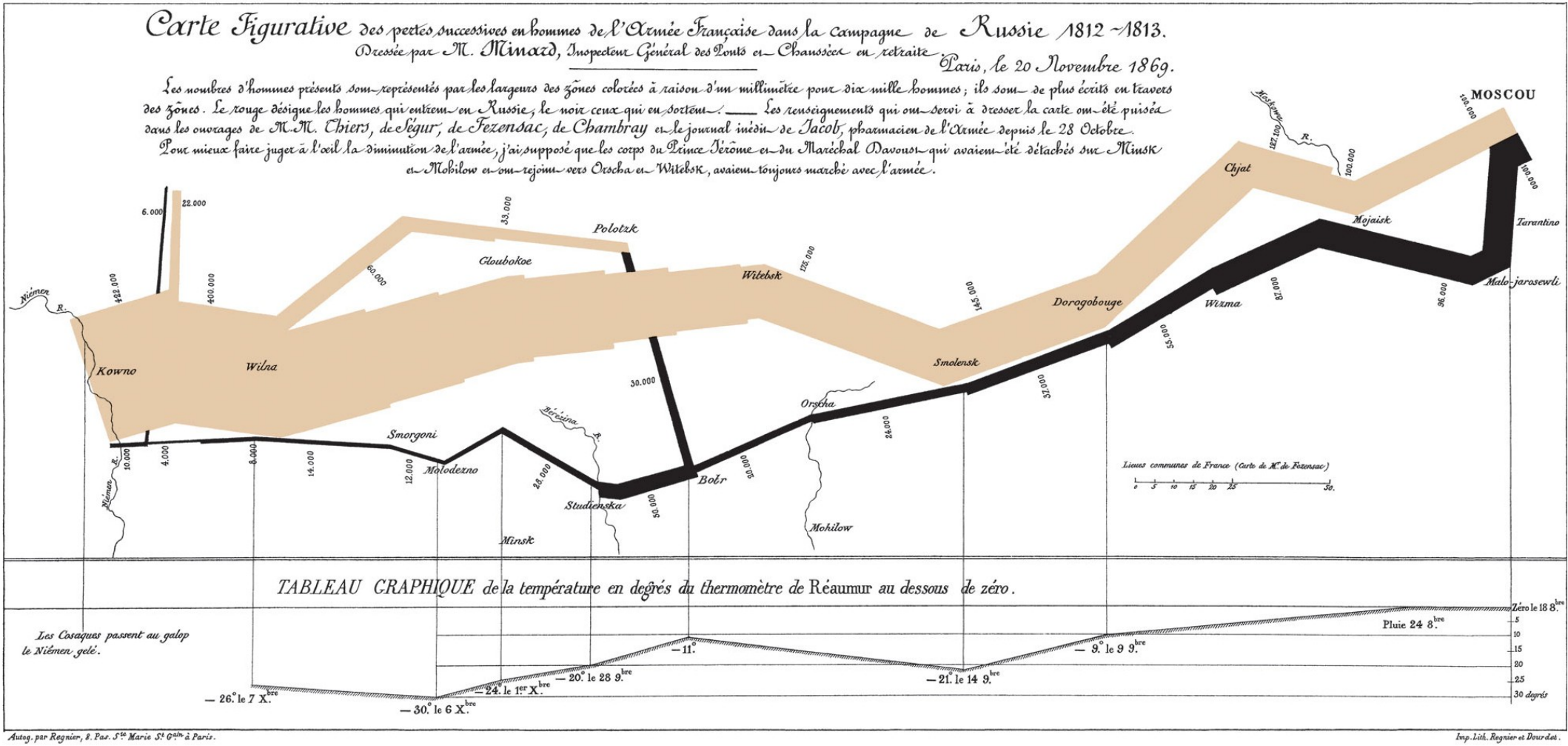
Hans Rosling trained in statistics and medicine and spent years on the front lines of public health in Africa. Yet his greatest impact has come from his stunning renderings of the numbers that characterize the human condition.

His 2006 TED talk, in which he animated statistics to tell the story of socio-economic development, has been viewed over 3.8 million times and translated into dozens of languages. His subsequent talks have moved millions of people worldwide to see themselves and our planet in new ways by showing how our actions affect our health and wealth and one another across space and time.

When you meet Rosling, 63, you are struck by his energy and clarity. He has the quiet assurance of a sword swallower (which he

Hans' graphics are regularly seen at UN conferences, and his work is frequently attributed to changing global policies.

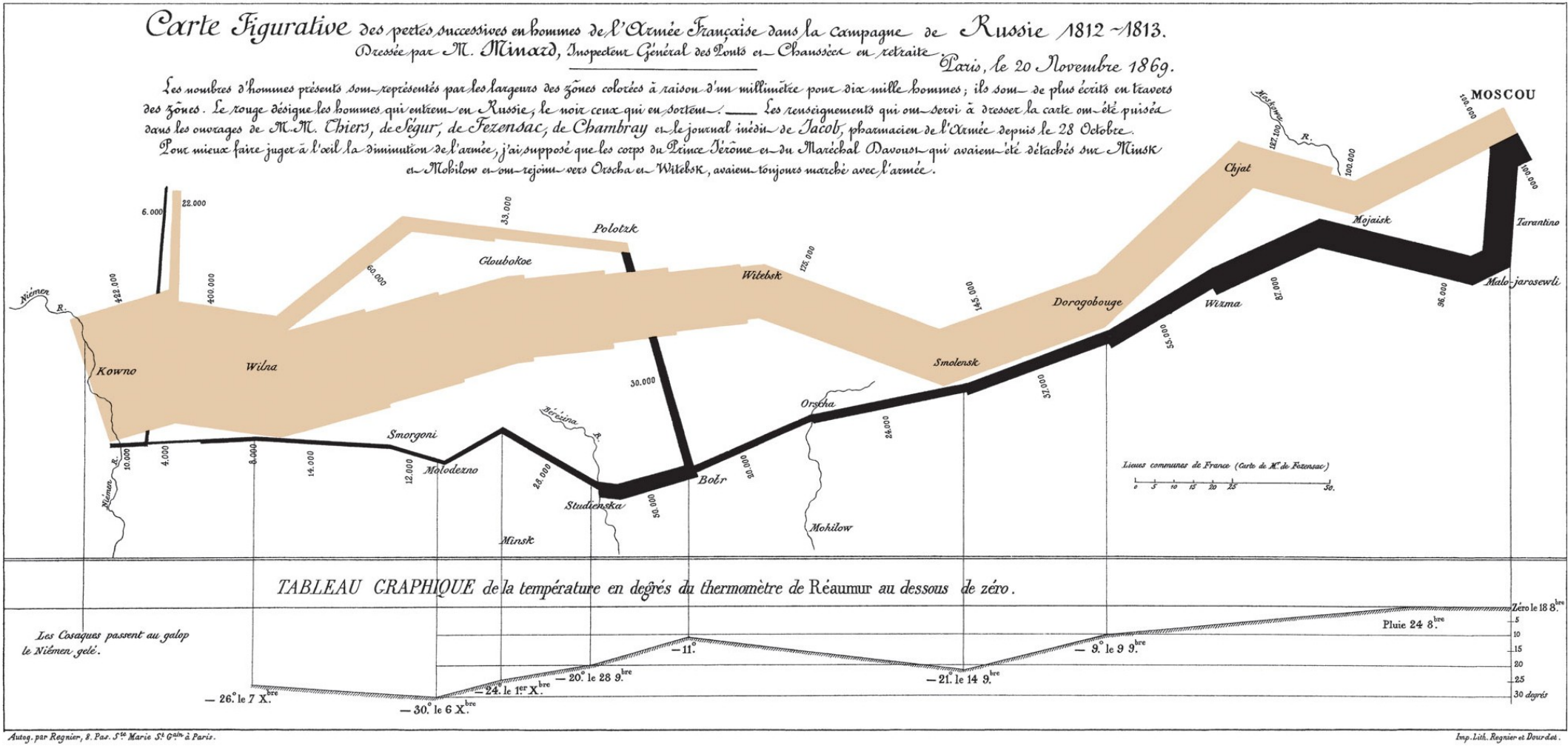
1812: Napoleon



Autog. par Regnier, 8. Pas. Ste Marie St Gain à Paris.

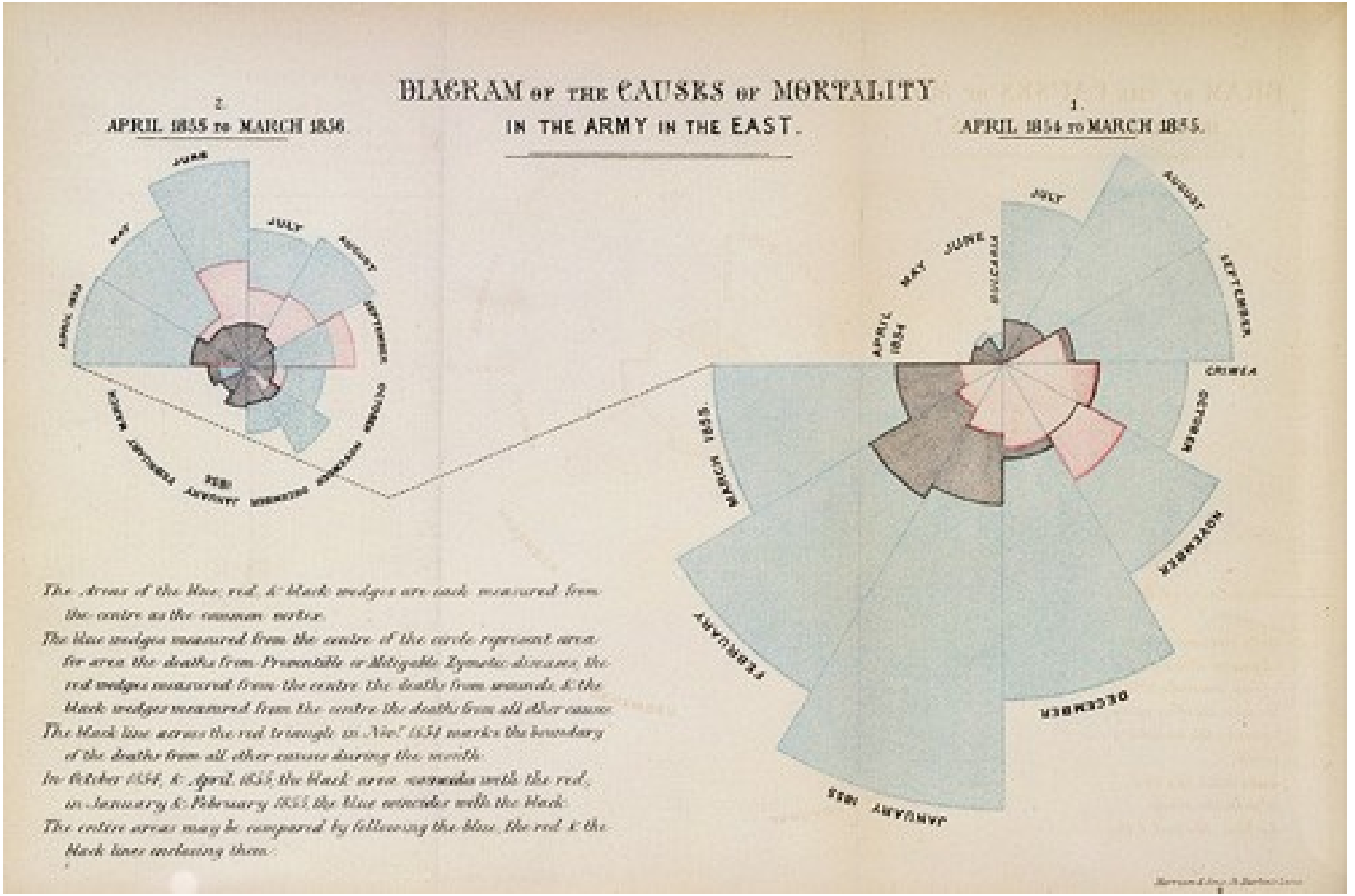
Imp. Lith. Regnier et Dourdet.

1812: Napoleon



Charles Minard illustrated that data visualization can be used to overcome limits of human perception.

Florence Nightingale



The Areas of the blue, red, & black wedges are each measured from the centre as the common vertex.

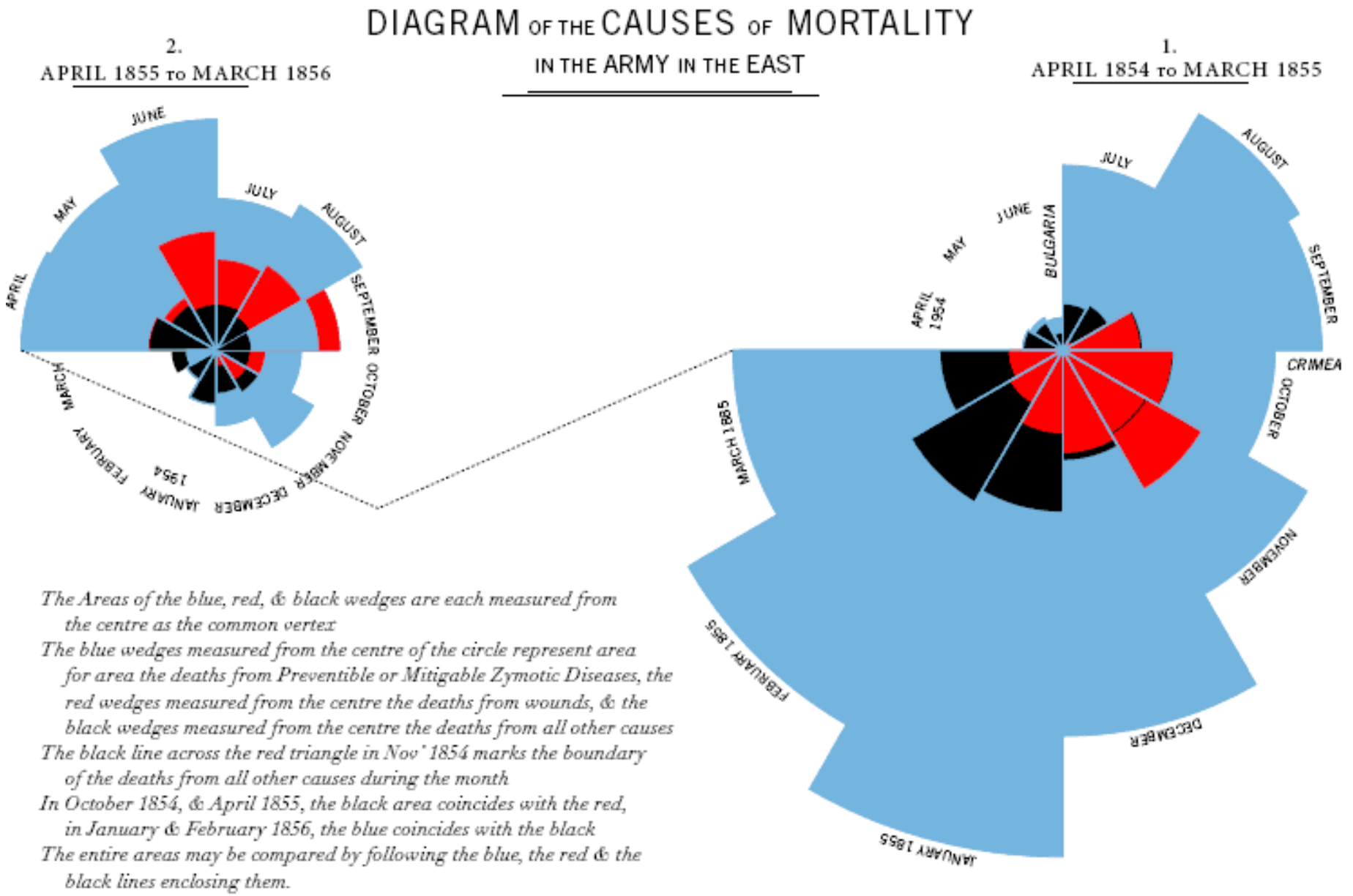
The blue wedges measured from the centre of the circle represent area for area the deaths from Preventible or Mitigable Zymotic diseases, the red wedges measured from the centre the deaths from wounds, & the black wedges measured from the centre the deaths from all other causes.

The black line across the red triangle in Nov.r 1854 marks the boundary of the deaths from all other causes during the month.

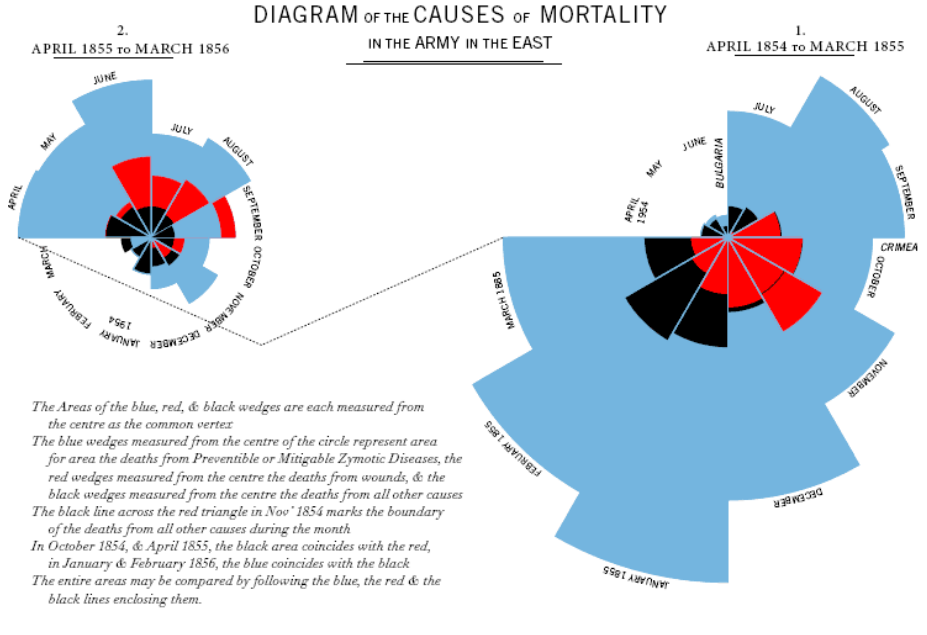
In October 1854, & April 1855, the black area coincides with the red, in January & February 1855 the blue coincides with the black.

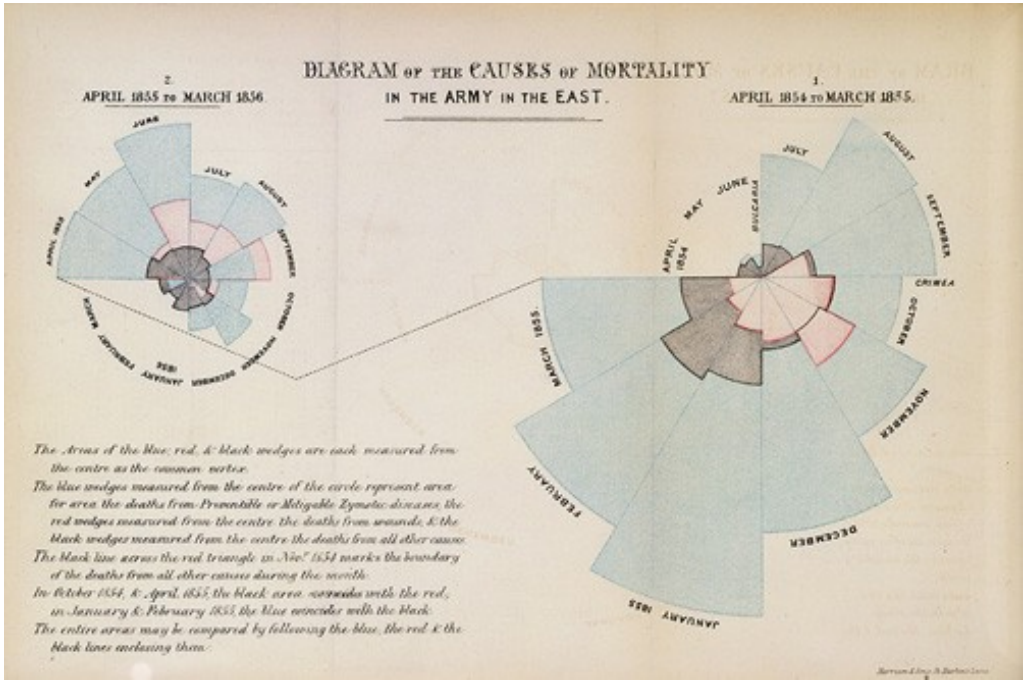
The entire areas may be compared by following the blue, the red & the black lines enclosing them.

Florence Nightingale, 1855

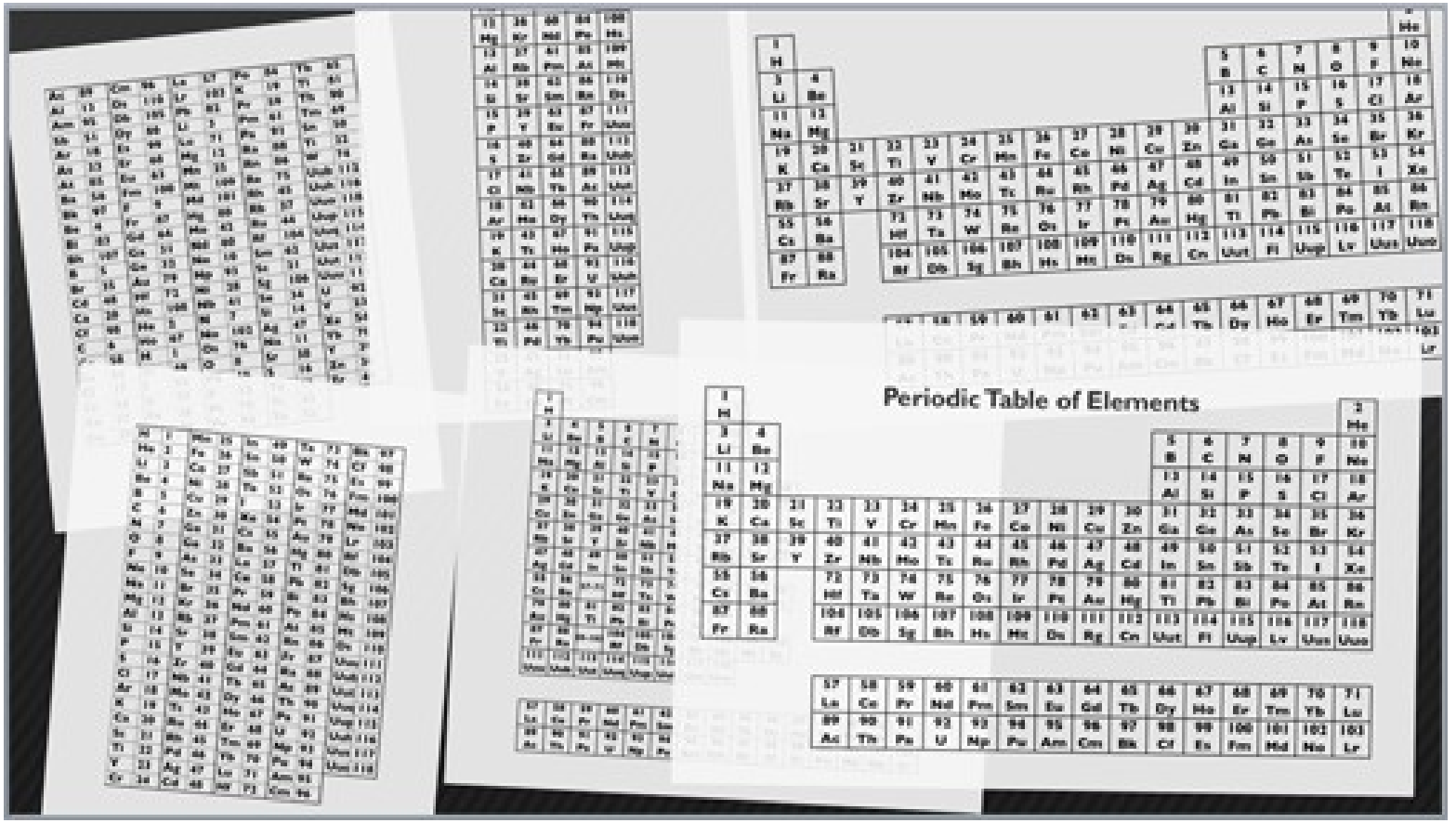


Florence Nightingale

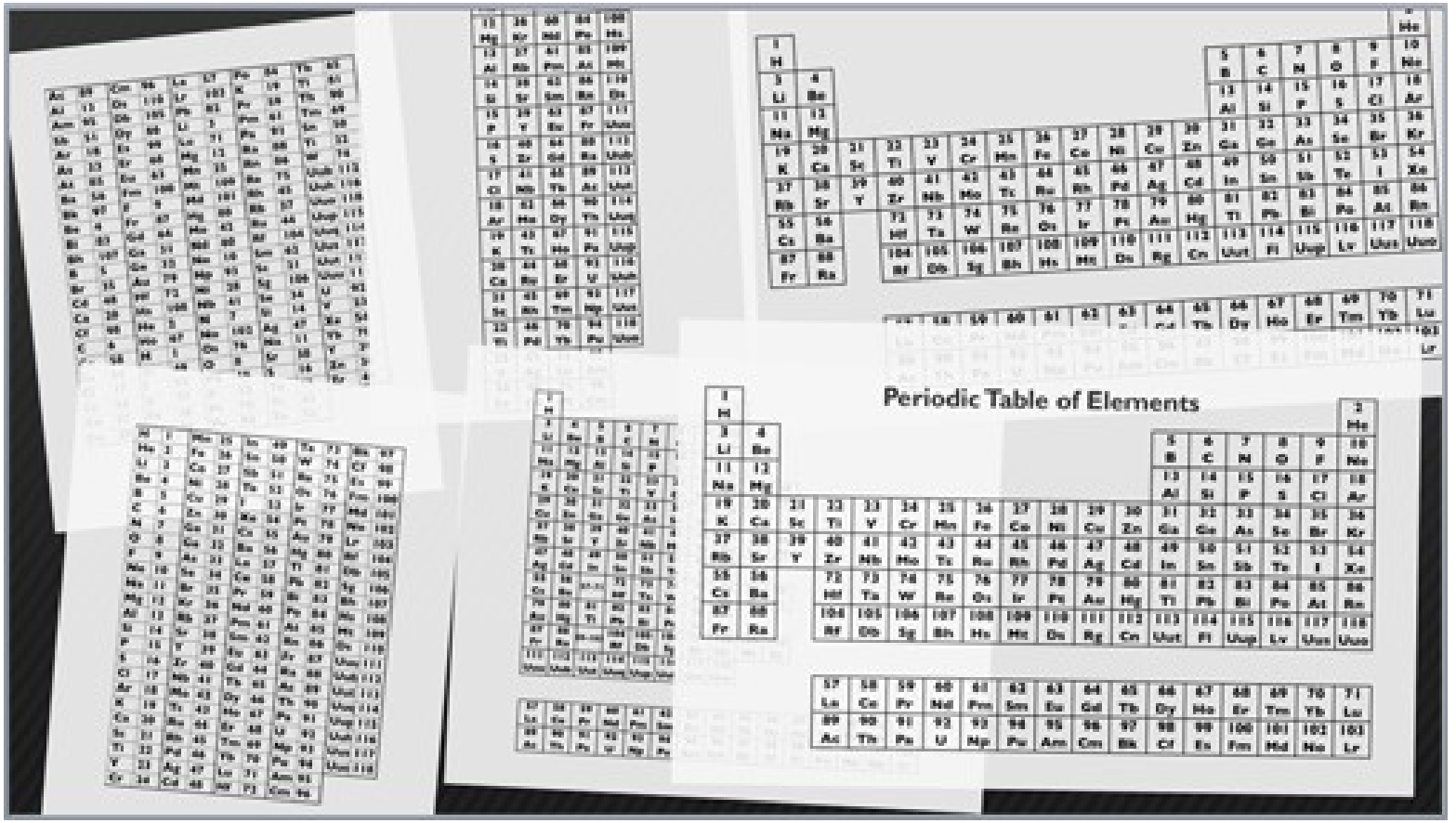




Florence Nightingale is credited with convincing a nation to invest in hospitals and medical care, rather than guns, to win a war.



1869, Dmitri Mendeleev



Dmitri Mendeleev predicted “blanks” in his early table would be filled in later – a trend that continues today.

LaToya Egwuekwe - Unemployment

- https://www.youtube.com/watch?v=hwWGzQ_FUtQ

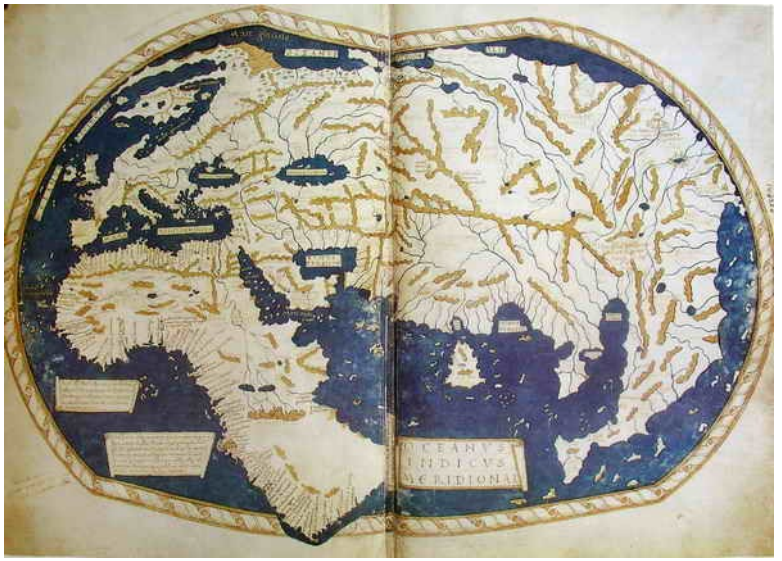
LaToya Egwuekwe - Unemployment

-

Harry Beck – Electrical Engineer

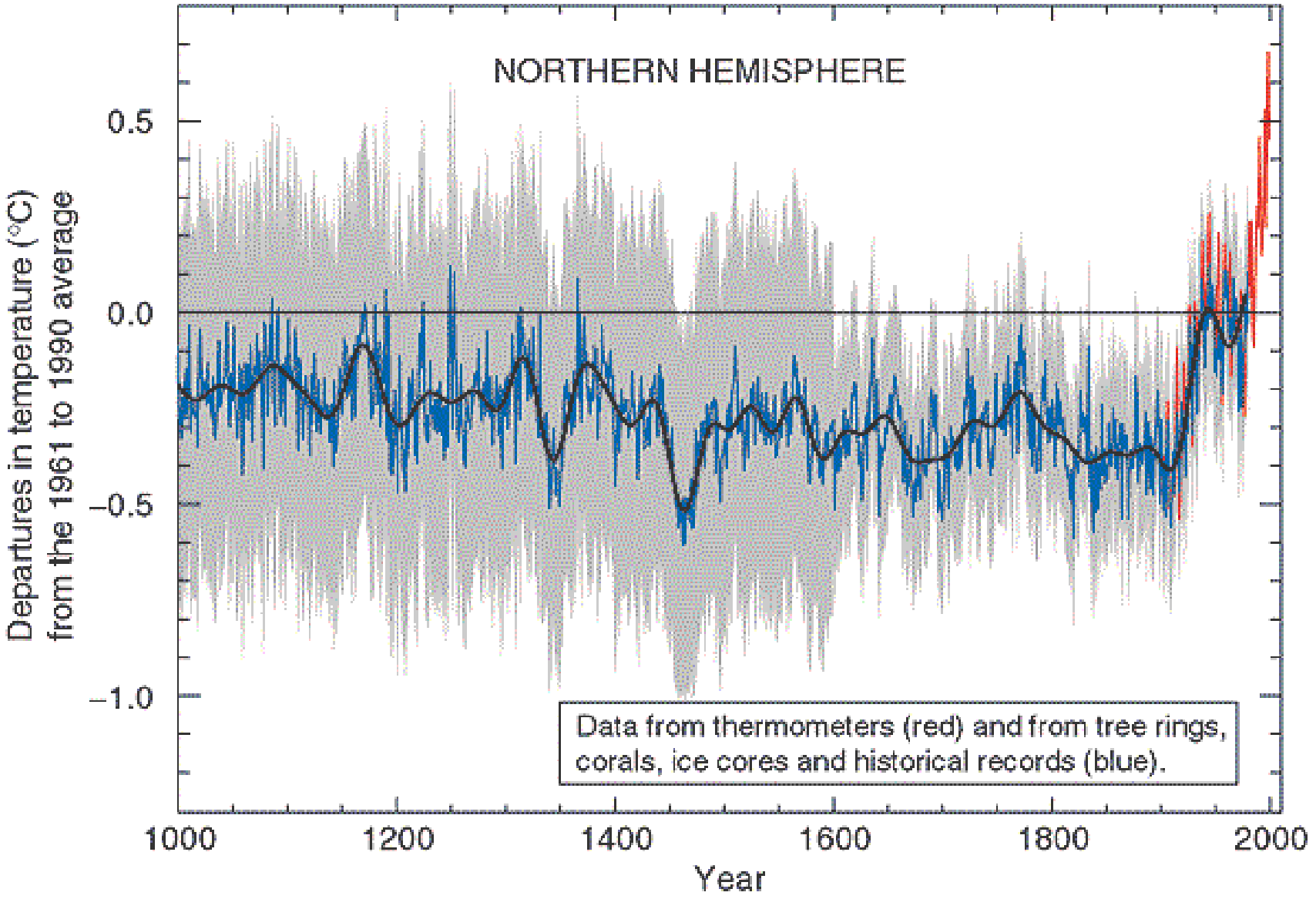
Henricus Martellus (1490)

Henricus Martellus (1490)

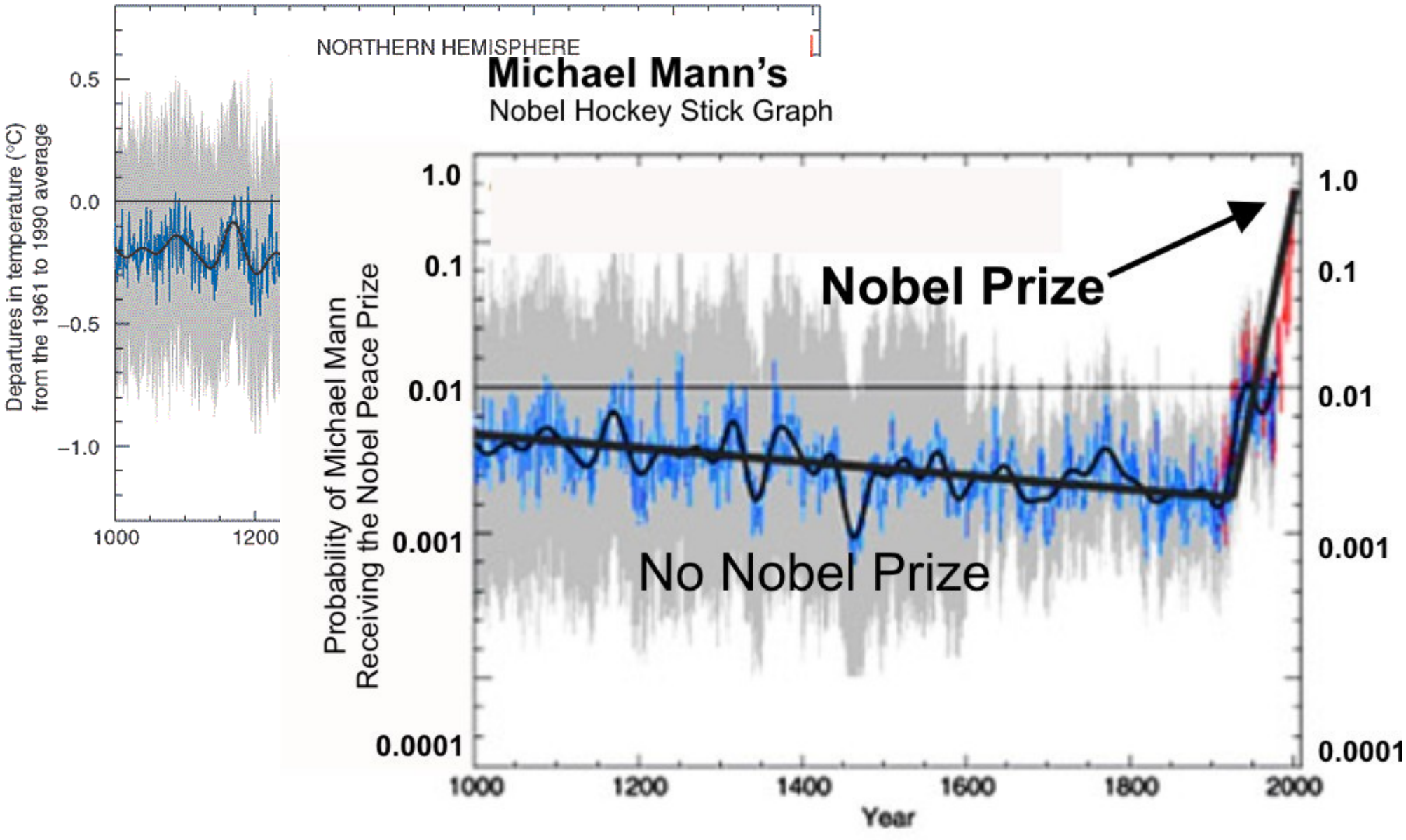


Though some controversy exists, this map was used by Columbus (you know, fall break Columbus) to convince Ferdinand of Aragon that Europe wasn't really that far from China.

Michael Mann – Climate Change



Michael Mann – Climate Change



“The Process” - According to McCandless

- Have a question! Don't let the data run your day.
- Collect and Clean the data
- Describe the Data
- Analyze the Data (...if relevant!)
- Visualize the Data – iterate, iterate, iterate.

So, let's do it:

- Have a question!
  - Are students having a positive experience in Breaking Intuition?
- Collect and Clean the data
  - It just so happens we did a little survey...
- Describe the Data
- Analyze the Data (...if relevant!)
- Visualize the Data – iterate, iterate, iterate.

The Survey

On a five-point scale, where 5 is the best and 1 is the worst, how would you rate the following (N/A indicates you haven't participated in that aspect of the course):

The “Why William and Mary” Lab (Intro to R): 1 2 3 4 5 N/A

The “Choosing Your Island Villa” Lab (Intro to GIS): 1 2 3 4 5 N/A

The “Business Analytics” Lab (Intro to Data Mining): 1 2 3 4 5 N/A

Course Lectures to Date (Not including guest lectures): 1 2 3 4 5 N/A

What is the one thing you enjoy most about this course so far?:

_____

_____

What is the one thing you would change about this course so far?:

_____

_____

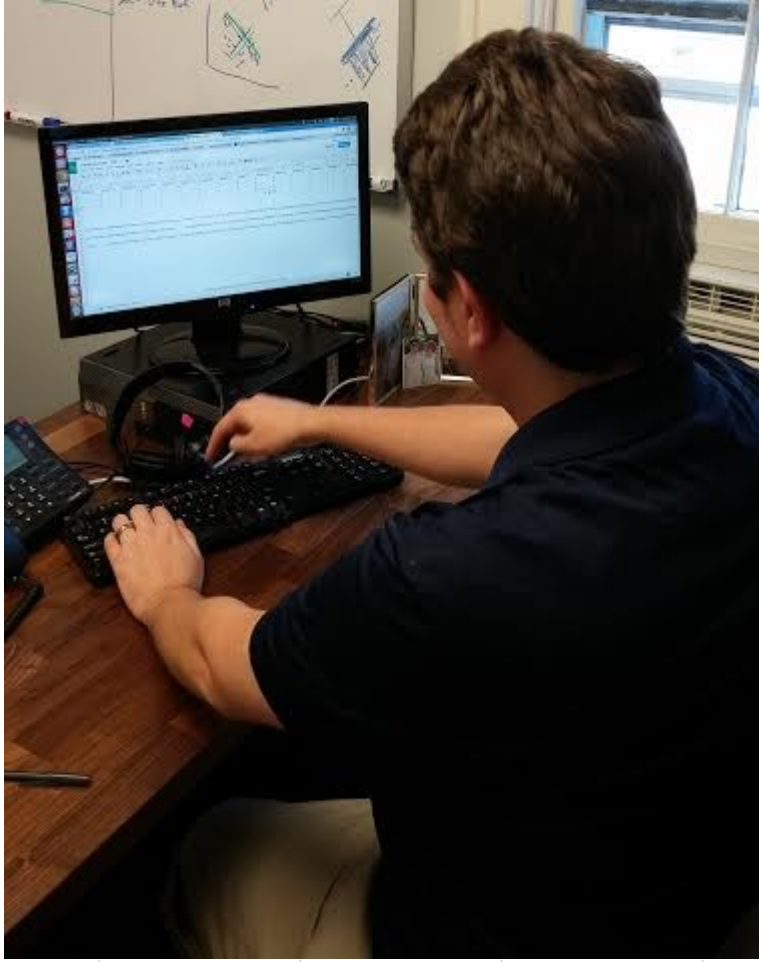
What type(s) of things would you like to learn more about?:

_____

_____

Cleaning the Data...

(a picture of a scientist in his natural habitat)



How would you rank each of the following:	Average (5 is best, 1 is worst)	Minimum	Max	Student 1	Student 2	Student 3	Student 4
Why William and Mary (Lab 1)	3.76	1	5	2	2	5	4
Human Habitation & Your Villa (Lab 2)	3.76	2	5	5	5	2	4
Business Analytics (Lab 3)	3.89	2	5	4	4	3	5
Lectures	4.69	3	5	5	5	4	5
Notes:							
What is the one thing you enjoy most about this course so far?							
				The lectures are	The lectures are	How we learn ho	The challenge o
What is the one thing you would change about this course so far?							
				more lectures	nothing	more instruction	I don't enjoy doi
What types of things would you like to learn more about?							
				hurricanes	Financial analysi	the uses/advantages of working in	

Describe the Data

- Basic Statistics: Averages, Medians, Modes

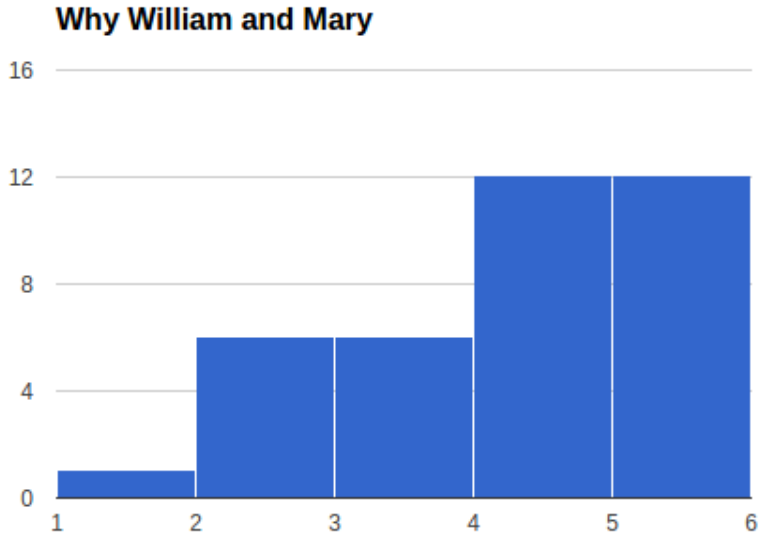
How would you rank each of the following:	Average (5 is best, 1 is worst)	Mode	Median	Minimum	Max
Why William and Mary (Lab 1)	3.76	5	4	1	5
Human Habitation & Your Villa (Lab 2)	3.76	4	4	2	5
Business Analytics (Lab 3)	3.89	4	4	2	5
Lectures	4.69	5	5	3	5

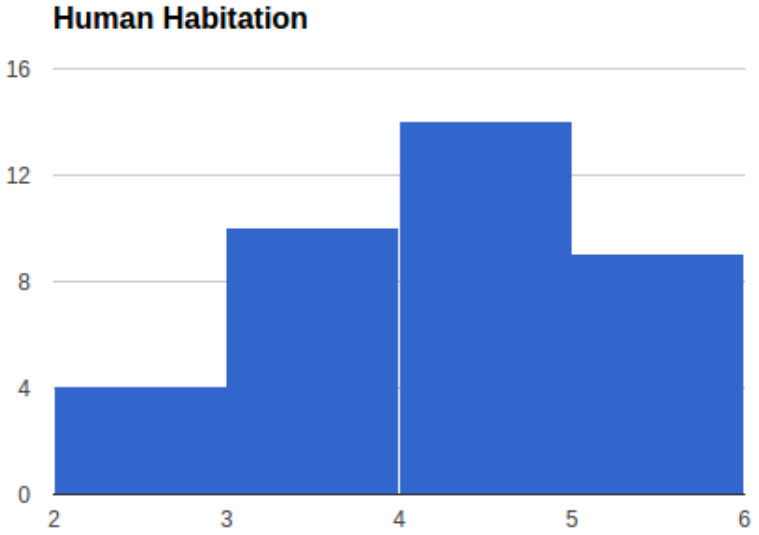
Describe the Data

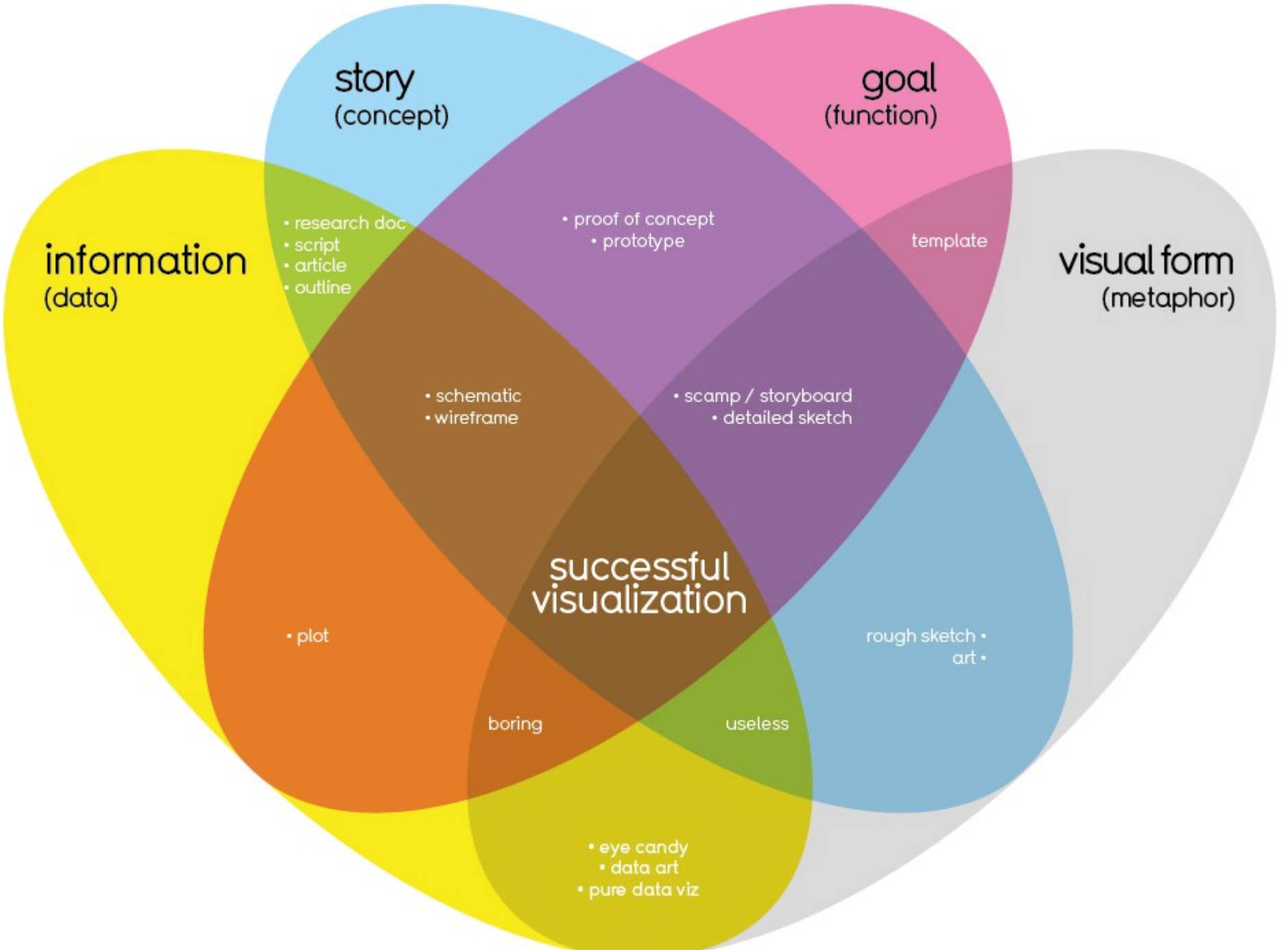
- Basic Statistics: Averages, Means, Modes

How would you rank each of the following:	Average (5 is best, 1 is worst)	Mode	Median	Minimum	Max
Why William and Mary (Lab 1)	3.76	5	4	1	5
Human Habitation & Your Villa (Lab 2)	3.76	4	4	2	5
Business Analytics (Lab 3)	3.89	4	4	2	5
Lectures	4.69	5	5	3	5

- Let's check out the distributions...







David McCandless
InformationisBeautiful.net

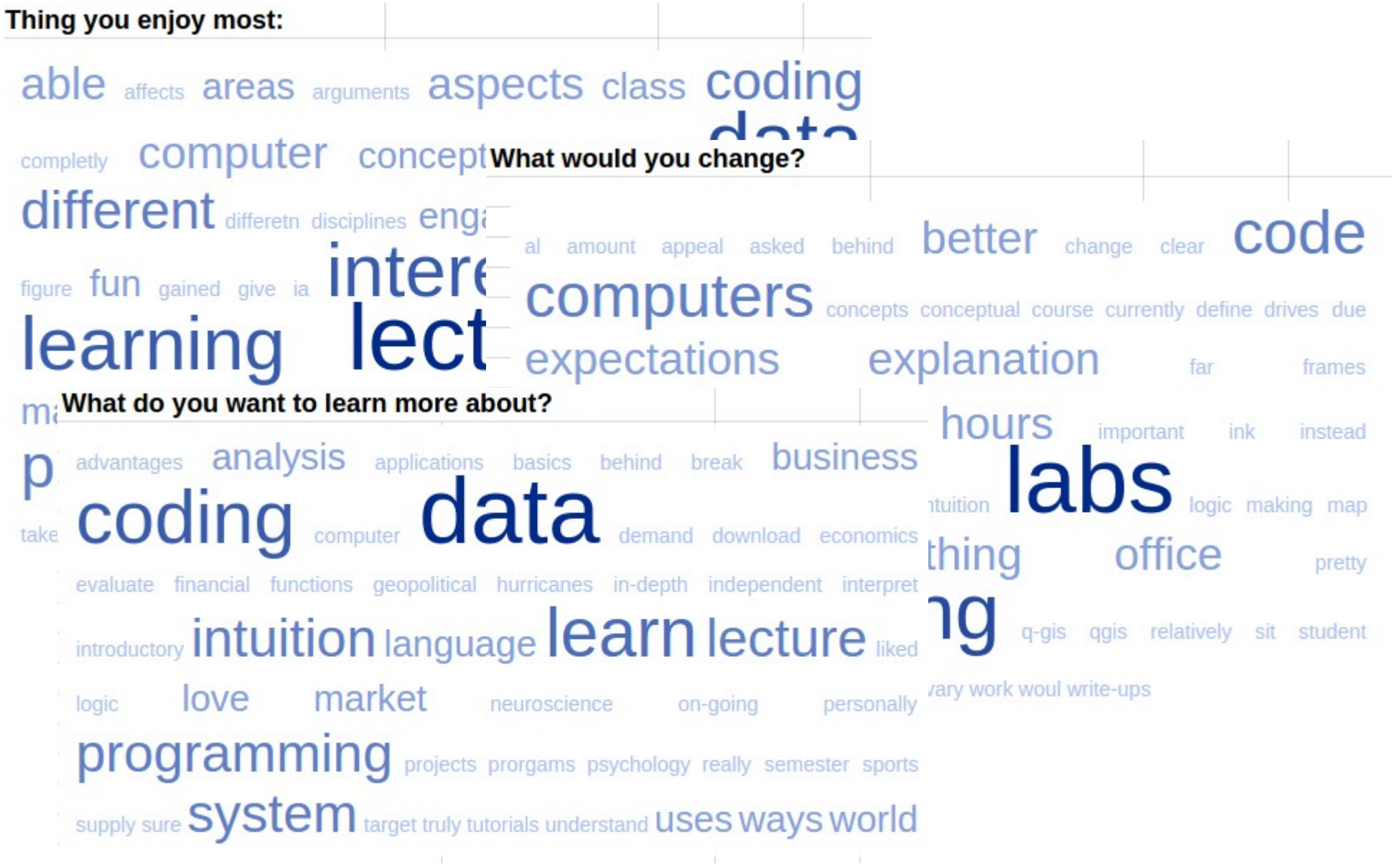
taken from new book
Knowledge is Beautiful

find out more
bit.ly/KIB_Books

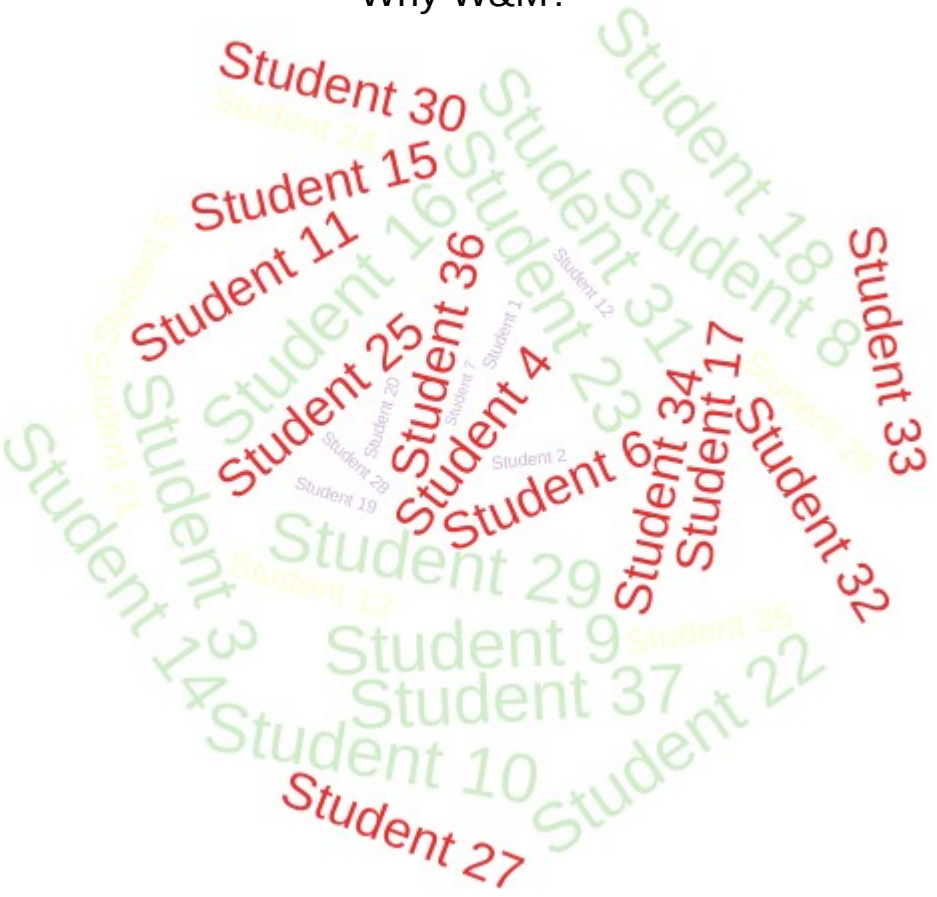
Viz Example: Course Reviews

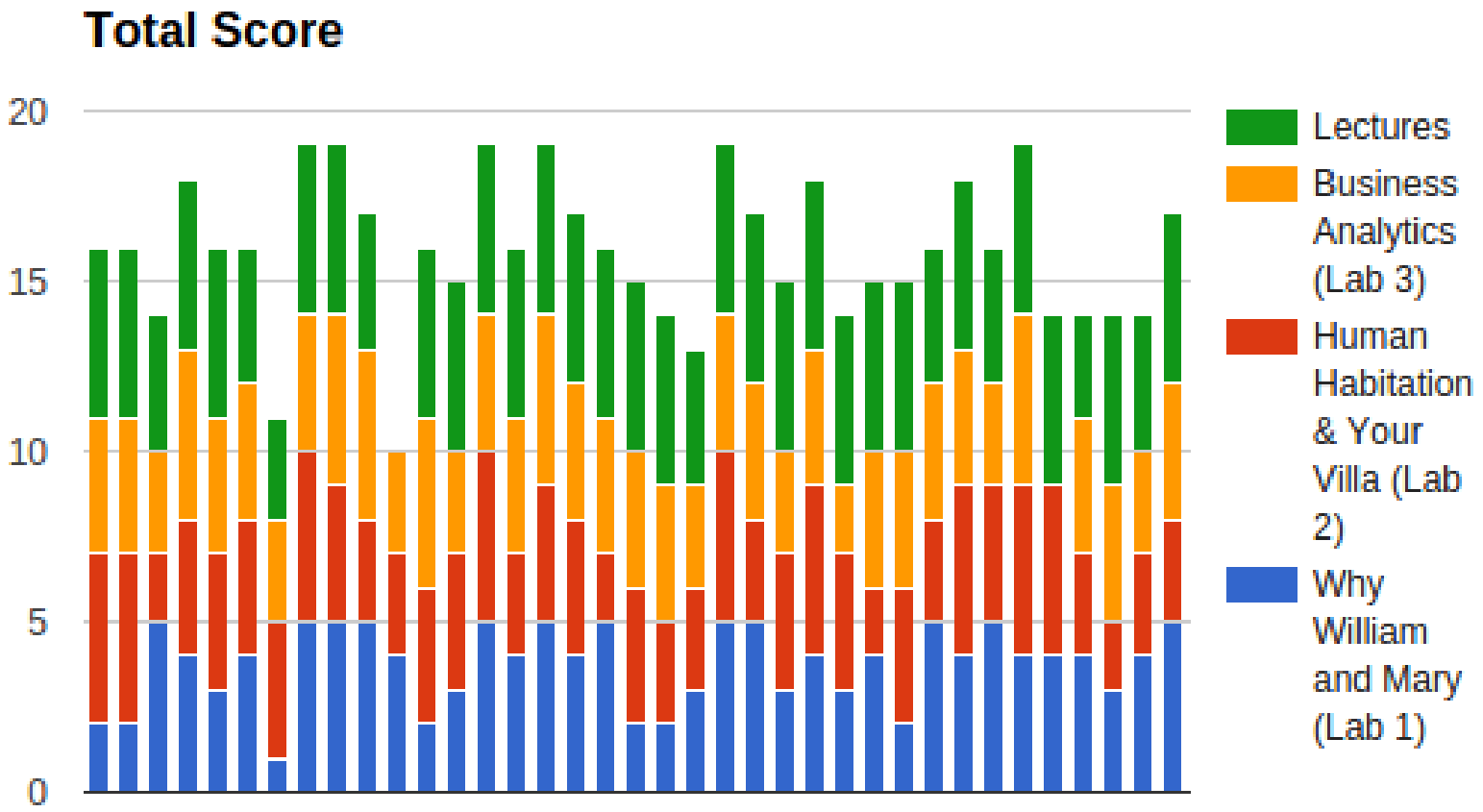
- Information
  - Data collected from COLL 100 students on how much they enjoy this course.
- Story
  - What are the key things we can learn from the student surveys?
- Goal
  - Learn to improve lectures and labs to get better student scores on the final surveys.
- Visual Form
  - This is what you all hate doing. “Creativity” and such nonsense.

Iterate, Iterate, Iterate



Why W&M?




Total Score
20
15
10
5
0
Lectures
Business Analytics (Lab 3)
Human Habitation & Your Villa (Lab 2)
Why William and Mary (Lab 1)

But, does data tell the entire story?

- Pair with qualitative findings:
  - Dislike of ambiguity
  - Desire for harder “R” skills (...but not too much because we don't like programming).
  - More interaction in labs
  - Getting to know one-another
  - ...more!

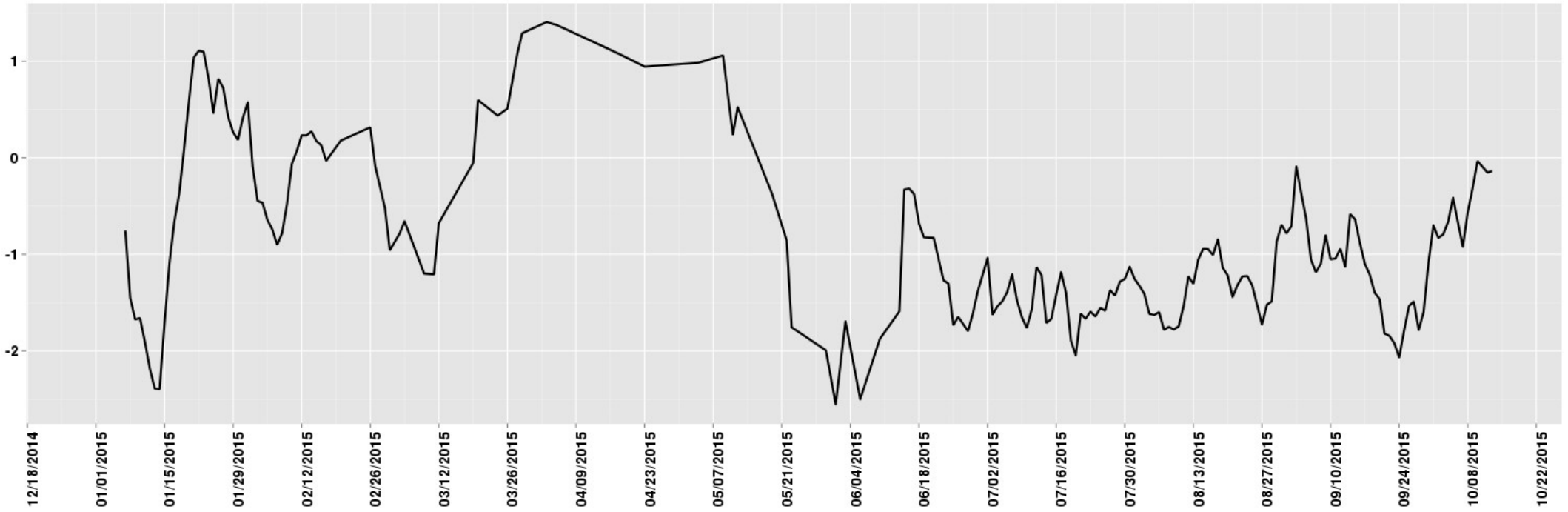
Quick ways YOU can visualize data

(i.e., really cool things too easy for a lab).

GDELT (http://analysis.gdeltproject.org/)

- News Articles from Dem Debate: http://data.gdeltproject.org/analytics_user/20151015150501.2938.heatmap.html

Donald Trump Sentiment



NOAA

- NTL: http://www.nnvl.noaa.gov/view/#BLUE?timespan=yearly&date=2001-12-31&lat=36.988353179083&lon=-71.971991991222&zoom=6

United Nations

- http://unctadstat.unctad.org/wds/ReportFolders/reportFolders.aspx?sCS_ChosenLang=en

What are the Ramifications of Data?

- What if you can't access data?
- What if you can't interpret it?
- What if you can't visualize it to make an argument?

  At the end of the day, “big data” is shifting power relations around the world.