

Data Literacy for Court Staff

OR Data 101 for Court Staff

Sit back, relax. The training is about to begin!

Sufjan Stevens, "Should Have Known Better" (Official Audio)



Monday, July 23 2018

1 / 39

Icebreaker Activity

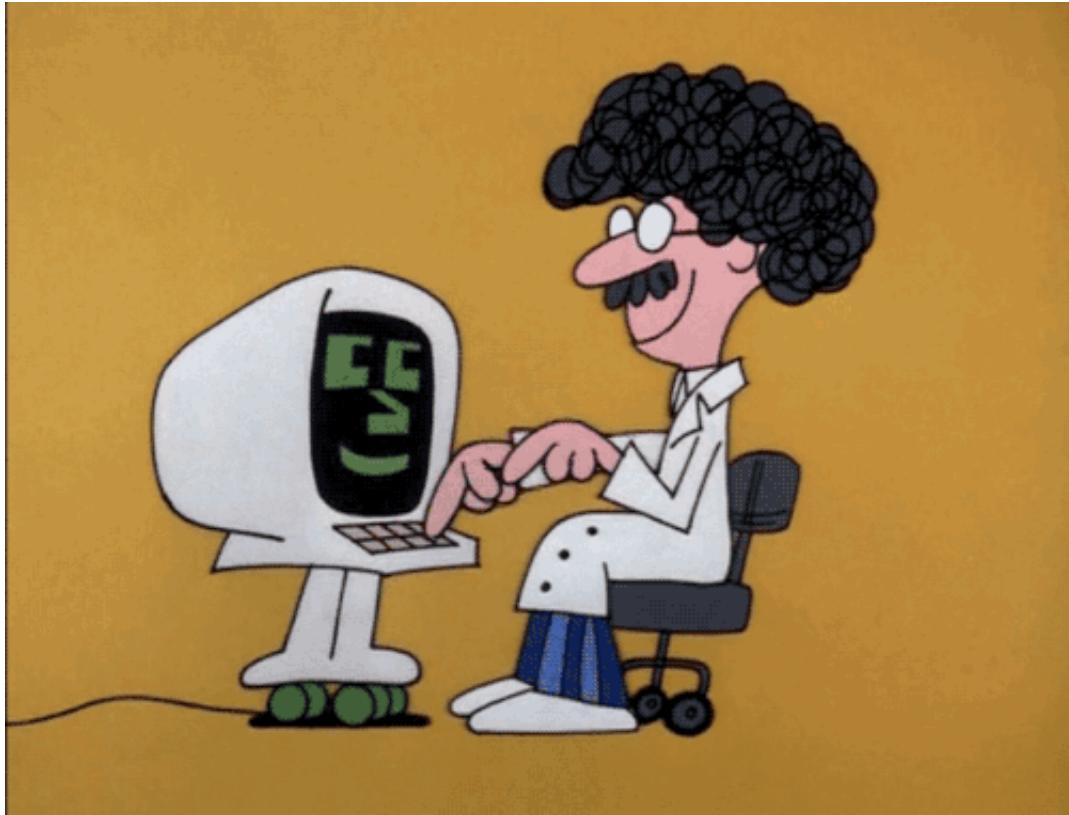
- Line up according to birthday: Month, Day (no need for age)
- No VERBAL or WRITTEN communication allowed



Learning Objectives:

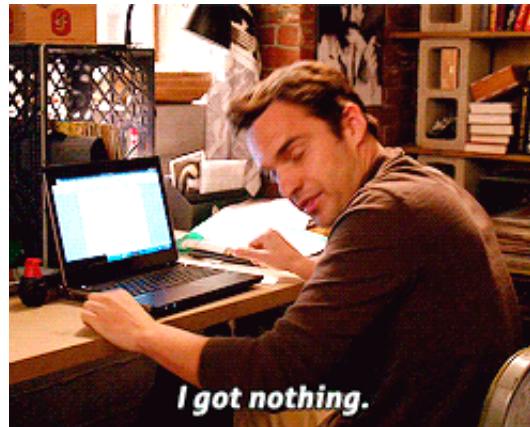
- To become fluent with the language of data
- To think critically about data
- To become familiar with the fundamentals of OUR data
 - What we collect
 - How we collect it
 - Why we need it
- What we can do to improve our Data Quality
- To discover what you want to learn next

Expectations



What do **you** want to learn today?

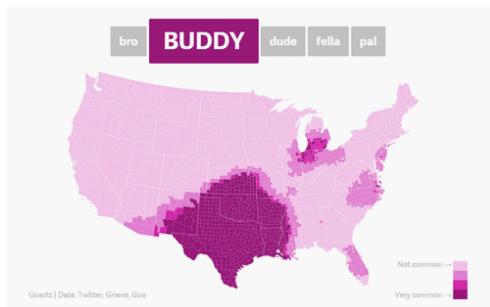
What Is Data?



I got nothing.

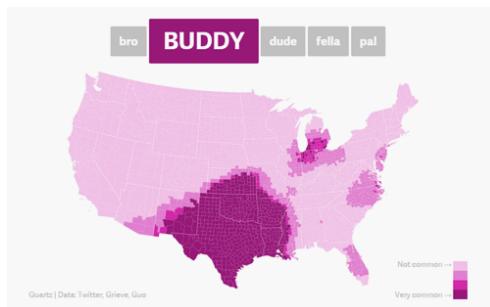
What questions do you have about data?

The Bro-Map



Exercise 1: How do you use data?

The Bro-Map



Exercise 1: How do you use data?

- At your table.
- Take the allotted time to note how you use personal **or** probation data.
- Write it down!

Sharing is Caring



Who wants to share how they use data?

Data or Information

Data:

- Facts, figures, and other abstract representations of the world
 - Height
 - Weight
 - Number of Charges
 - Length of Term of Probation
 - Demographics

Data or Information

Information:

- Data that has been processed, given structure, and context
 - Charts and graphs
 - Risk Level
 - Credit
 - Census
 - Body Mass Index (BMI)
 - Far From Perfect Model

Data or Information

- Exercise: Revisit the *How do you use data.*
- Reclassify items as "Data or Information"



How much data do we need?

How much data do we need?

- Pros and cons of Personal Experience

How much data do we need?

- Pros and cons of Personal Experience
- Professional/Clinical Judgment

How much data do we need?

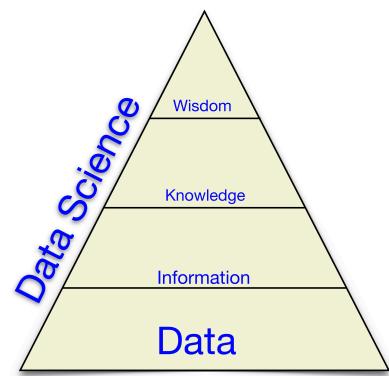
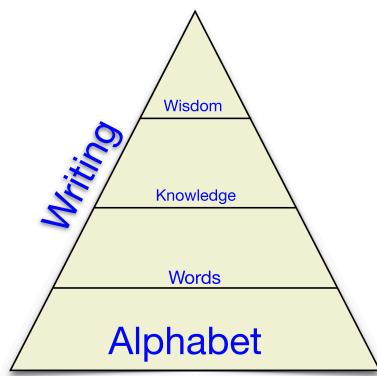
- Pros and cons of Personal Experience
- Professional/Clinical Judgment
 - *Challenging either can be easily interpreted as a personal attack*

How much data do we need?

- Pros and cons of Personal Experience
- Professional/Clinical Judgment
 - *Challenging either can be easily interpreted as a personal attack*
 - *Why not just collect all the data?*
- **A better question is: What data do we need?**



The Language of Data



Think of data and information as a new language that requires fluency

Word Cloud: Your Registration Data²



From a total of 9 registrants.

13 / 39

Types of Data

Quantitative

- Discrete: only specific values matter for that data
- Continuous: All values
- *Examples?*
 - Height

Types of Data

Quantitative

- Discrete: only specific values matter for that data
- Continuous: All values
- *Examples?*
 - Height
 - Weight

Types of Data

Quantitative

- Discrete: only specific values matter for that data
- Continuous: All values
- *Examples?*
 - Height
 - Weight
 - Age

Types of Data

Quantitative

- Discrete: only specific values matter for that data
- Continuous: All values
- *Examples?*
 - Height
 - Weight
 - Age
- Income Data

Types of Data

Quantitative

- Discrete: only specific values matter for that data
- Continuous: All values
- *Examples?*
 - Height
 - Weight
 - Age
- Income Data
- Blood Pressure

Types of Data

Qualitative

- Binary: Yes/No
- Nominal: Names with no particular meaning
- Ordinal: Scores or Ranks
- Count/Ratio: Items per Area/Volume
- *Examples?*

Types of Data

Qualitative

- Binary: Yes/No
- Nominal: Names with no particular meaning
- Ordinal: Scores or Ranks
- Count/Ratio: Items per Area/Volume
- *Examples?*
 - Binary
 - Ordinal
 - Nominal
 - Geographic Location?

Types of Data

Probation Specific Examples

- Quantitative Examples?
 - Number of youth on probation
 - Number of Violations
 - Number of service agencies

Types of Data

Probation Specific Examples

- Quantitative Examples?
 - Number of youth on probation
 - Number of Violations
 - Number of service agencies
- Qualitative Examples
 - What services is the young person ordered to attend?
 - Violations per district
 - Termination Status

The Coffee Example



Morning Beverage Excercise

- What do you drink every morning?
- How much do you drink every morning
 - In discrete terms.
 - In continuous terms.



Descriptive Statistics

"Descriptive statistics are brief descriptive coefficients that summarize a given data set, which can be either a representation of the entire or sample population. Descriptive statistics are broken down into measures of central tendency and measures of variability (spread)."

--[Investopedia](#)

Measures of Central Tendency and Variability

- Mean: Average

Measures of Central Tendency and Variability

- Mean: Average
- Median: Equal point in the distribution

Measures of Central Tendency and Variability

- Mean: Average
- Median: Equal point in the distribution
- Mode: Most frequent number

Measures of Central Tendency and Variability

- Mean: Average
- Median: Equal point in the distribution
- Mode: Most frequent number
- Range: Minimum number to maximum number

Measures of Central Tendency and Variability

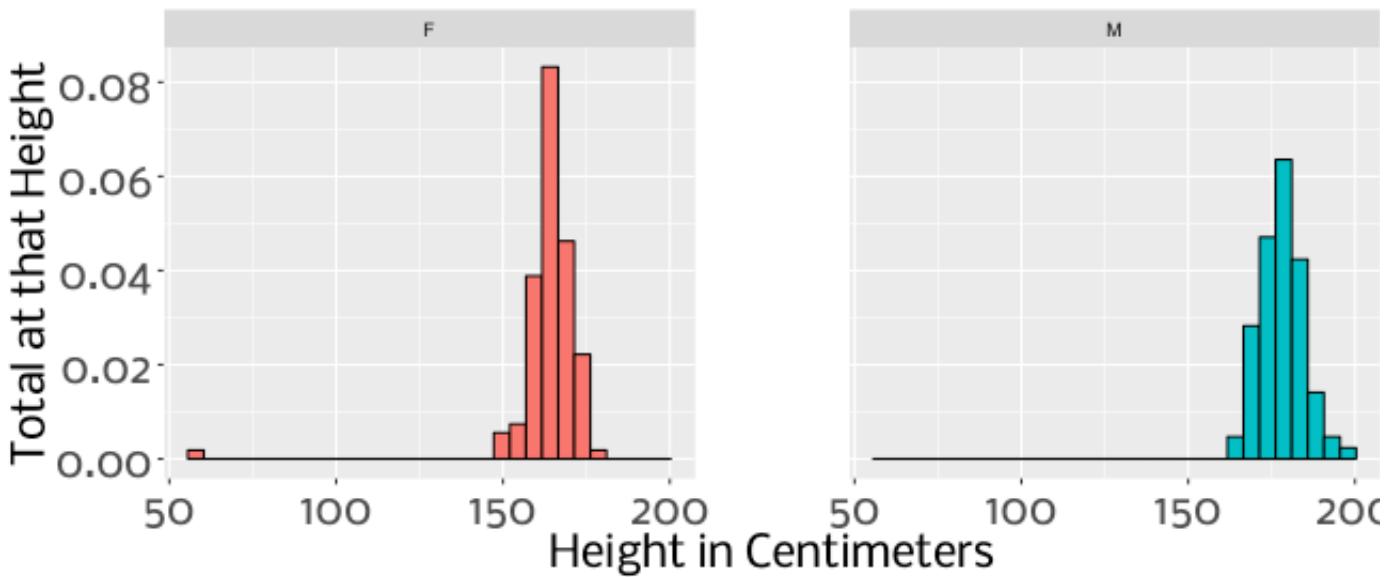
- Mean: Average
- Median: Equal point in the distribution
- Mode: Most frequent number
- Range: Minimum number to maximum number
- Other measures of central tendency include:
 - Variance
 - Standard Deviation
 - Standard error of the Mean
 - Beyond the scope of today's training

Terms in Practice: With Height!

- One volunteer¹ to document how tall everyone is
- In inches
- We will calculate:
 - Mean
 - Median
 - Mode
 - Range

[1] Or voluntold, I'm not picky.

Best Practices: Visualization



All Measurements in cm

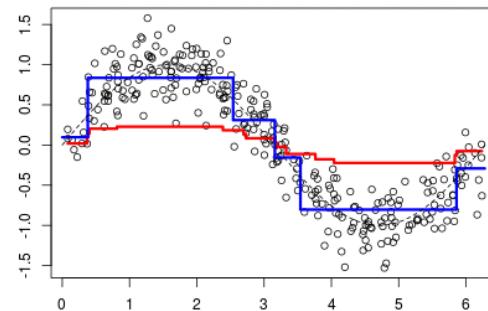
| sex | Average Height | Median Height | Shortest | Tallest | Count | Standard Deviation |
|-----|----------------|---------------|----------|---------|-------|--------------------|
| F | 163.7411 | 165 | 57 | 178 | 112 | 11.643925 |
| M | 178.0114 | 178 | 163 | 197 | 88 | 6.440701 |

The Stastical Model

- A representation of what the data tells us
- The goal of:
 - Big Data
 - Machine Learning
 - AI

The Statistical Model

- A representation of what the data tells us
- The goal of:
 - Big Data
 - Machine Learning
 - AI



Four Kinds of Analytics

- Descriptive: What?
 - Very similar to descriptive statistics.
 - Examples of probation descriptive statistics?
- Diagnostic: Why?
- Predictive: What if?
- Prescriptive: Best course?

Diagnostic Analytics



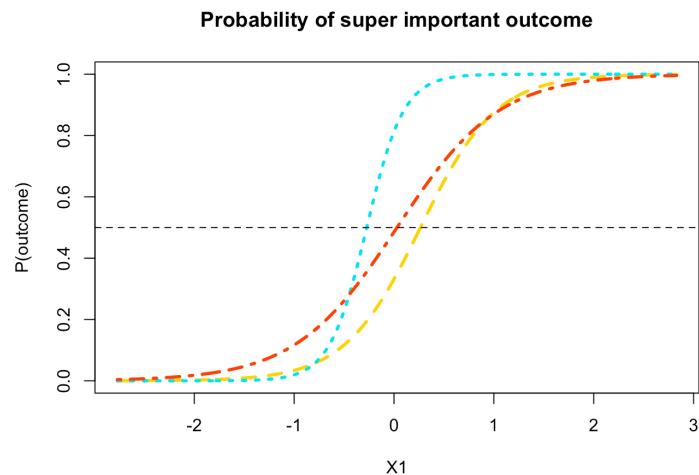
- Why did this happen?
- What kind of diagnostic questions should we ask?

Predictive Analytics

- What could happen
 - Chances of winning an election
 - Chances of developing an illness
 - Chance at re-offending

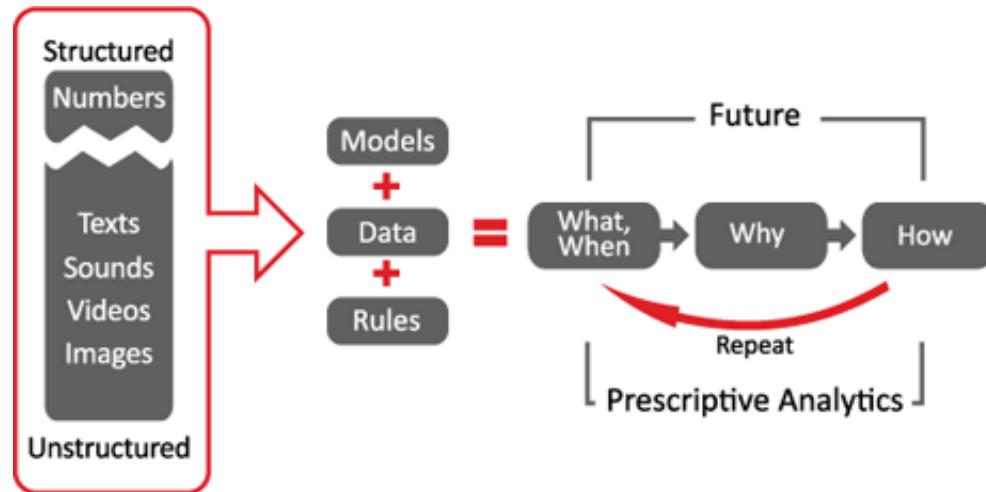
Predictive Analytics

- What could happen
 - Chances of winning an election
 - Chances of developing an illness
 - Chance at re-offending



Perscriptive Analytics

- What is the best course of action?



Exercise 3: Discussion of Prescriptive Statistics

Exercise 3: Discussion of Prescriptive Statistics



- This is what all of our models are attempting to answer

How do we collect data or information?

How do we collect data or information?

- Personal Experiences
- TV
- Newspapers
- Radio
- Conversations
- Comparison shopping
- Google
- Facebook

How does the *Court* collect data?

- Forms

How does the *Court* collect data?

- Forms
- Reports

How does the *Court* collect data?

- Forms
- Reports
- Studies

How does the *Court* collect data?

- Forms
- Reports
- Studies
- JEMS/C5/Supervisor
- Assessment Tools

How does the *Court* collect data?

- Forms
- Reports
- Studies
- JEMS/C5/Supervisor
- Assessment Tools
- Other Systems
 - I-Clear
 - Clerks
 - Crimes (State's Attorneys)
 - Schools
 - Hospital
 - Promise

How do we store Data

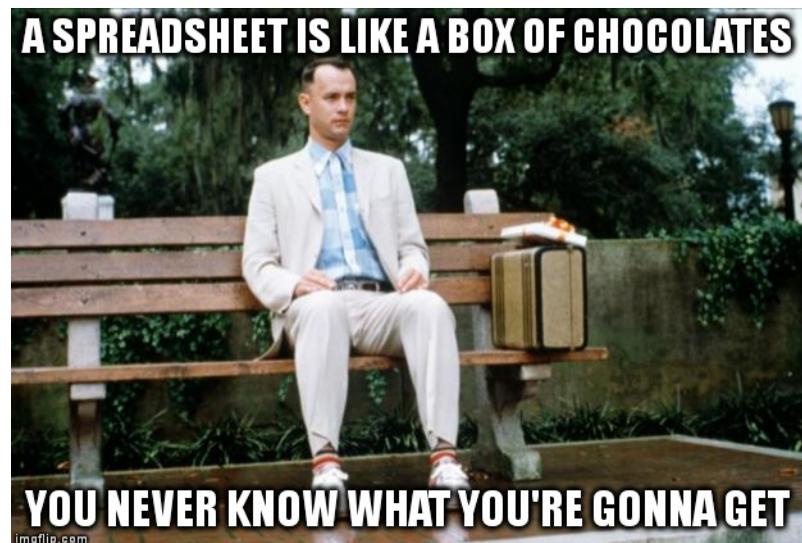
- Word Documents
- Databases

How do we store Data

- Word Documents
- Databases
- *Spreadsheets*

How do we store Data

- Word Documents
- Databases
- *Spreadsheets*



Tidy Data

- Each variable forms a column.
- Each observation forms a row.
- Each cell is atomic.
- Each type of observational unit forms a table.

Tidy Data

- Each variable forms a column.
- Each observation forms a row.
- Each cell is atomic.
- Each type of observational unit forms a table.



Spreadsheet Use

- We use spreadsheets for:
 - Collection, cleaning, results, analysis
 - Provide structure to our data
 - Share data/information
 - Data Analysis
- In our office, we use Excel
 - Google Sheets
 - And, occasionally, tables in Word
 - What should we use Spreadsheets for?

Exercise 4: Critique This Spreadsheet

| | A | B | C | D | E |
|----|-------------------------|---------|---------------|-----------|--------------|
| 1 | Name | Age | Work Location | Hire Date | Phone number |
| 2 | Cambell, Aaron | 25 | 3rd Floor | 2/1/1990 | x3-1234 |
| 3 | Arrington, Buford | 65 | Retired | 18-Sep-89 | 773.444.1212 |
| 4 | Gleason, Martin A | ##### | Eighth Floor | 6.1.2000 | x3-5511, |
| 5 | Flanagan, Kelly "Judge" | ageless | Court Room A1 | 7/20/2001 | na |
| 6 | Amanda Halawa-Mahdi | | 3rd Floor | 1/8/1900 | 312.433.651 |
| 7 | | | | | |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |
| 11 | | | | | |
| 12 | | | | | |
| 13 | | | | | |
| 14 | | | | | |
| 15 | | | | | |
| 16 | | | | | |
| 17 | | | | | |
| 18 | | | | | |
| 19 | | | | | |

Exercise 4: Takeaways

- Cells are not atomic
- Missing values
- Variable formats within entries

Why Do We Need Data: Takeaways

- Evidence Based Practices
- Performance based budgeting
- Risk Level and Classification

Where Do We Go From Here

- We have the language, now what are *our* goals?

Where Do We Go From Here

- We have the language, now what are *our* goals?
- With your table, discuss where you want to go from here.
 - Start with a question you want to answer
 - What descriptive and diagnostic analyses do you need?
 - Everyone must have a role
 - Report back to the group
 - No longer than 60 seconds!
- Use your SMART goals!

Thank you



- Please fill out your evaluation forms
- Email martin.gleason@cookcountil.gov if you have questions, comments, or vague misgivings

Sources and Further Reading

- [Tidy data Paper](#)
- [Data Literacy](#)
- [Algorithmic Sentencing](#)
- [Descriptive, predictive, prescriptive](#)
- <https://twitter.com/nlj/status/1004066012063559686>
- [4 kinds of analytics](#)
- [Super Important Outcome](#)
- [Minitab: Flavors of Data](#)
- [Statistical Model](#)
- [BMI](#)
- [Coffee Info Graphic](#)