Data Ethics

Why Data Science?

- Quantitative Fields
  - Math
  - Probability
  - Statistics
- On-Demand Job Market
  - LinkedIn Workforce Report: Shortage in All Cities (2018)
  - Article by Gregory Piatetsky
- Applicable Across All Areas
- Bigger Data Requires Better Skills
- People are Scared Because Analytics is Coming

Why Data Science?

Why Data Science?

- Data Science is Not a Fad
- People Have Data and Need Us to Analyze the Information to Answer Their Questions
- Decisions Based on Our Analyses Can Effect Others
- We Must Operate Ethically
- There is Power in the Information

Some Dope Quotes

With Great Power Comes Great Responsibility.

- Uncle Ben

No One Man Should Have All That Power.

- Kanye West

Hippocratic Oath for Data Scientists

I solemnly pledge to practice my profession with conscience and dignity;

To respect the privacy of the people whose data is confided in me;

To maintain the utmost respect for the individuals whose data I am analyzing;

To be transparent, open, and honest about the type of analysis I am applying to their data;

To never use my knowledge to violate human rights and civil liberties, even under threat.

-- Marie Wallace

Data Fraud

- Appropriate Data Analysis
  - Let the Data Speak for Itself
  - Hypotheses Before Analyses
- Transparency
  - Data Imputation
  - Handling of Outliers
  - Data Transformations
  - Statistical Methods
  - Necessary Assumptions
  - Important Caveats
  - Variables Included and Excluded

Data Fraud

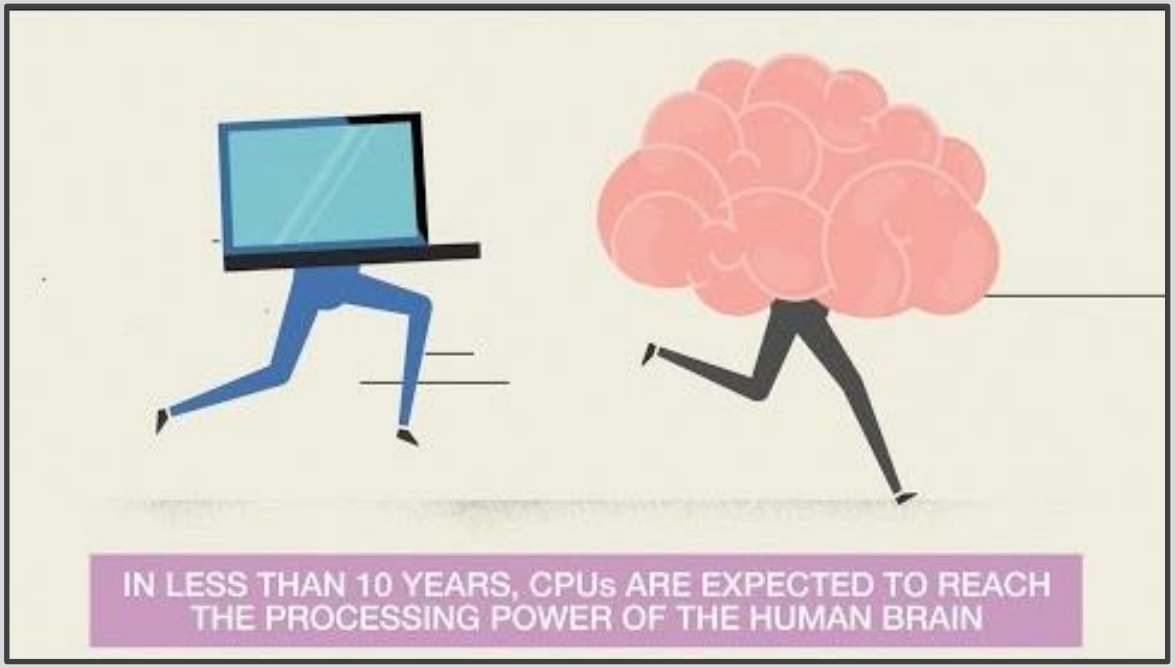
- Incident at Harvard (2018)
  - Dr. Piero Anversa
  - 31 Publications
  - Stem Cell Research
  - Showed the Heart Has the Ability to Regenerate
  - Fraudulent Data
  - Article by NY Times

Data Fraud

- Thought Provoking Questions
  - Why Does This Happen?
  - Are You Capable of This?
  - How Does This Apply to You in Your Group Projects?

Data is Coming

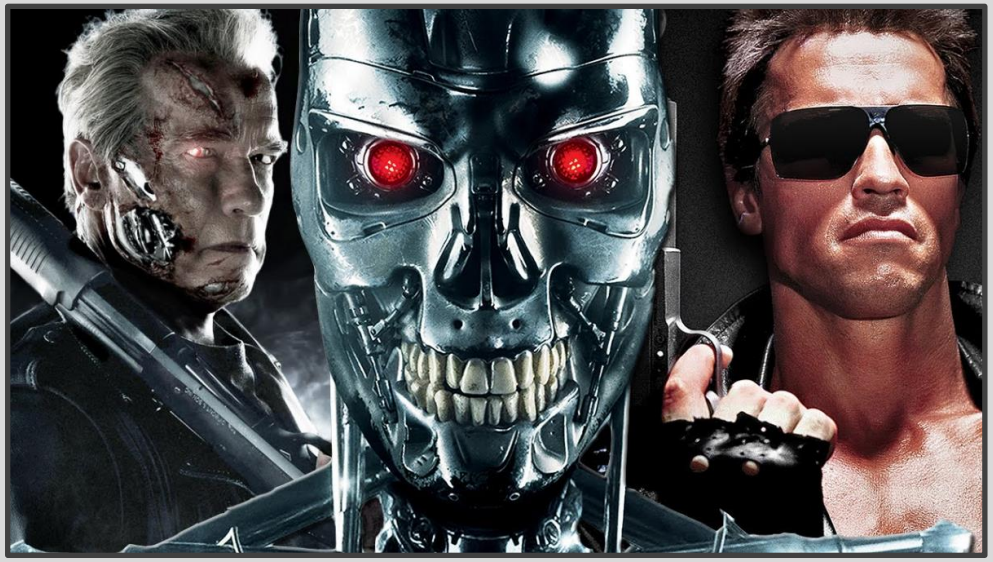
- The Expansion of Data



- Thoughts and Concerns?

Data is Coming

- How About Now?



Value of Data

- How Much is Data Worth?
  - In 2016, Microsoft Purchased LinkedIn for $26.2 Billion
  - In 2016, Caesar's Total Rewards Program Was Valued at $1 Billion During Bankruptcy
  - Article in MIT Sloan Magazine
- Project Nightingale (Google)
- Thought Provoking Questions
  - Are You Worried About Data Privacy?
  - How Should We Address This?

Modern Data Ethics

- Private Customer Data and Identity Should Remain Private
- Shared Private Information Should Be Treated Confidentially
- Customers Should Have a Transparent View
- Big Data Should Not Interfere with Human Will
- Big Data Should Not Institutionalize Unfair Biases

Article by Pedro Uria-Recio

Modern Data Ethics

- Weapons of Math Destruction
  - Author: Cathy O'Neil
  - Blog: mathbabe.org
  - Worked on Wall Street
  - Saw the Dark Side of Using Algorithms on Big Data to Make Business Decisions
  - Leader in Data Ethics

Closing

Disperse and Make Reasonable Decisions