

Introduction: Social Data Analytics

INF1005/6, Prof. Alex Hanna

January 12, 2017

Class Description

- “Big (Social) Data”
- Data Ethics
- Computer programming

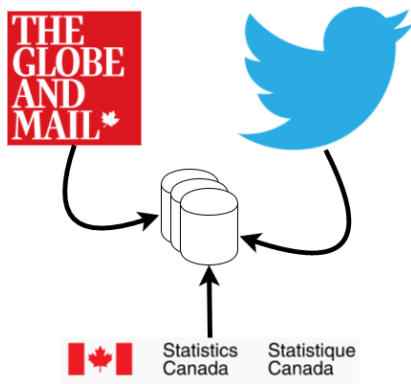
Class Goals

- Identify
- Collect
- Transform
- Analyze
- Visualise



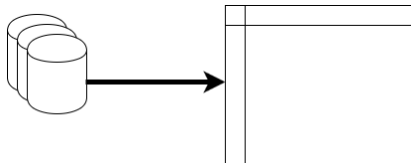
Class Goals

- Identify
- Collect
- Transform
- Analyze
- Visualise



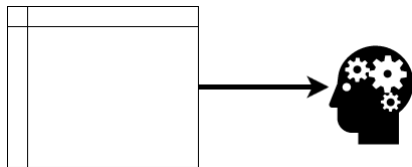
Class Goals

- Identify
- Collect
- Transform
- Analyze
- Visualise



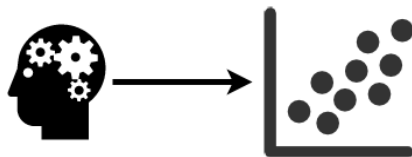
Class Goals

- Identify
- Collect
- Transform
- Analyze
- Visualise



Class Goals

- Identify
- Collect
- Transform
- Analyze
- Visualise



Class Requirements

- Attendance (20%)
- Assignment 1 (40%)
- Assignment 2 (40%)

Class Requirements

- Attendance (20%)
- Assignment 1 (40%)
- Assignment 2 (40%)

Class Requirements

- Attendance (20%)
- Assignment 1 (40%)
- Assignment 2 (40%)

Procedures and Rules

- E-Culture
 - Academic integrity
 - Equity and diversity

Procedures and Rules

- E-Culture
- Academic integrity
- Equity and diversity

Procedures and Rules

- E-Culture
- Academic integrity
- Equity and diversity

About Me

- Professor Alex Hanna
(Prof. Hanna, or Dr.
Hanna)
- PhD and MS in Sociology
- BS in Computer Science
- Transgender and queer
- Pronouns: she/her/hers

About Me

- Professor Alex Hanna
(Prof. Hanna, or Dr.
Hanna)
- PhD and MS in Sociology
- BS in Computer Science
- Transgender and queer
- Pronouns: she/her/hers

About Me

- Professor Alex Hanna
(Prof. Hanna, or Dr.
Hanna)
- PhD and MS in Sociology
- BS in Computer Science
- Transgender and queer
- Pronouns: she/her/hers

About Me

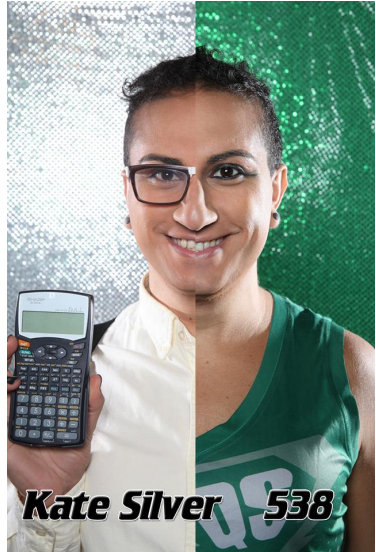
- Professor Alex Hanna
(Prof. Hanna, or Dr.
Hanna)
- PhD and MS in Sociology
- BS in Computer Science
- Transgender and queer
- Pronouns: she/her/hers

About Me

- Professor Alex Hanna
(Prof. Hanna, or Dr. Hanna)
- PhD and MS in Sociology
- BS in Computer Science
- Transgender and queer
- Pronouns: she/her/hers

About Me

- Professor Alex Hanna
(Prof. Hanna, or Dr. Hanna)
- PhD and MS in Sociology
- BS in Computer Science
- Transgender and queer
- Pronouns: she/her/hers
- Roller derby
player/announcer/coach



About You!

`http://bit.ly/inf1005-survey`

What are Social Data Analytics?



500M+

Monthly Actives

300M+

Daily Actives

80%+

Outside The U.S.

4.2B

Likes Daily

95M+

Photos/Videos Per Day



313M

Monthly active
users

1B

Unique visits
monthly to sites
with embedded
Tweets

82%

Active users on
mobile

3,860

Employees
around the world



SHOPIFY PLATFORM RELEASED

2006

EMPLOYEES

1500+

MERCHANTS

300,000+

TOTAL SALES ON SHOPIFY

\$20 Billion

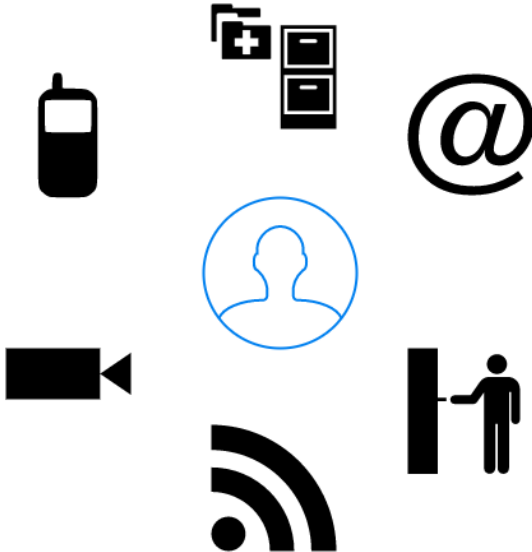


Statistics
Canada

Statistique
Canada

Five-year Census of every Canadian (>35 million)
350+ active surveys on Canadian life

Big Data is “Trace” Data



Big Data Fields

“Data Science”
“Business intelligence and analytics”
“Computational social science”



DATA

Data Scientist: The Sexiest Job of the 21st Century

Computational Social Science (Shah et al. 2015)

- the use of large, complex datasets, often... measured in terabytes or petabytes
- the frequent involvement of 'naturally occurring' social and digital media sources and other electronic databases
- the use of computational or algorithmic solutions to generate patterns and inferences from these data
- the applicability to social theory in a variety of domains from the study of mass opinion to public health, from examinations of political events to social movements

Computational Social Science (Shah et al. 2015)

- the use of large, complex datasets, often... measured in terabytes or petabytes
- the frequent involvement of 'naturally occurring' social and digital media sources and other electronic databases
- the use of computational or algorithmic solutions to generate patterns and inferences from these data
- the applicability to social theory in a variety of domains from the study of mass opinion to public health, from examinations of political events to social movements

Computational Social Science (Shah et al. 2015)

- the use of large, complex datasets, often... measured in terabytes or petabytes
- the frequent involvement of 'naturally occurring' social and digital media sources and other electronic databases
- the use of computational or algorithmic solutions to generate patterns and inferences from these data
- the applicability to social theory in a variety of domains from the study of mass opinion to public health, from examinations of political events to social movements

Computational Social Science (Shah et al. 2015)

- the use of large, complex datasets, often... measured in terabytes or petabytes
- the frequent involvement of 'naturally occurring' social and digital media sources and other electronic databases
- the use of computational or algorithmic solutions to generate patterns and inferences from these data
- the applicability to social theory in a variety of domains from the study of mass opinion to public health, from examinations of political events to social movements

Computational Social Science

- Network analysis
- Automated content (e.g. text and image) analysis
- Massive Online Experiments

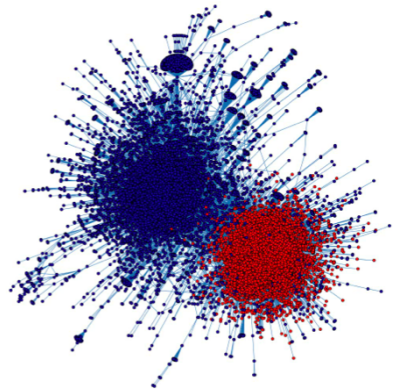
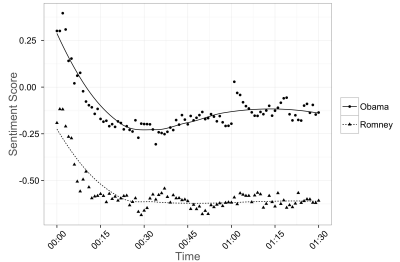


Fig. 2. The political retweet network, laid out using a force-directed algorithm. Node colors reflect cluster assignments (see text).

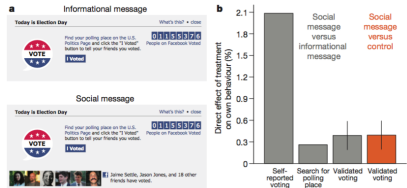
Computational Social Science

- Network analysis
- Automated content (e.g. text and image) analysis
- Massive Online Experiments

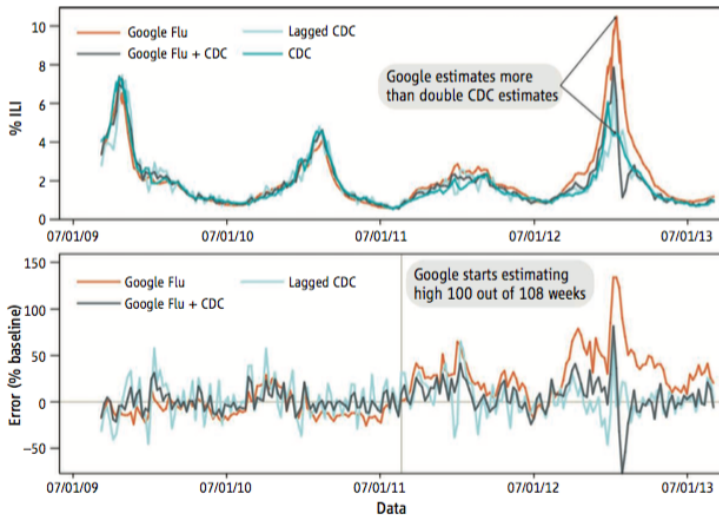


Computational Social Science

- Network analysis
- Automated content (e.g. text and image) analysis
- Massive Online Experiments



Big Data Gone Wrong: Google Flu Trends



Big Data Gone Wrong: Facebook and Emotions



Coding as Thinking

- Programming languages get old, thinking does not
- When you program, you open doors
 - DIY
 - Troubleshooting
 - Career



Programming (thinking) exercise:
How to program a computer