BSAD 8700 - Business Analytics

Welcome to Business Analytics!

Instructor: Dusty White, PhD

Who am I?

- Economics Faculty Member
  - Undergrad in Economics, from Brigham Young University
  - PhD in Economics, from Washington State University
  - Data and Data Analytics are my passion!

Contacting me

Email: drwhite@unomaha.edu

Slack: datascienceuno.slack.com

- Use the `#businessanalytics` channel for this class
- Be sure to sign up with your UNO email

I am also happy to arrange meetings by Zoom if you would like a more personal interaction. We can and will find a time that works for both of us!

Course material

The entire course is accessible through Canvas!

- You will find lectures here
- You will submit assignments and projects here
- You will find links to all documents here

Software needed for class

- Tableau Desktop (we will provide a license for the duration of the course!)
- Orange Data Mining
- SQLectron, a great and free SQL interface
- Microsoft Excel

What do you hope to get out of this course?

Let me know in the Slack channel!

How will this course help you?

- Learn the principles of analytics
- Learn to communicate with Data Analysts
- Learn when Data Analytics should be used
- Understand when Analytics is not being correctly applied.

I'm a Manager, not a Data Scientist...

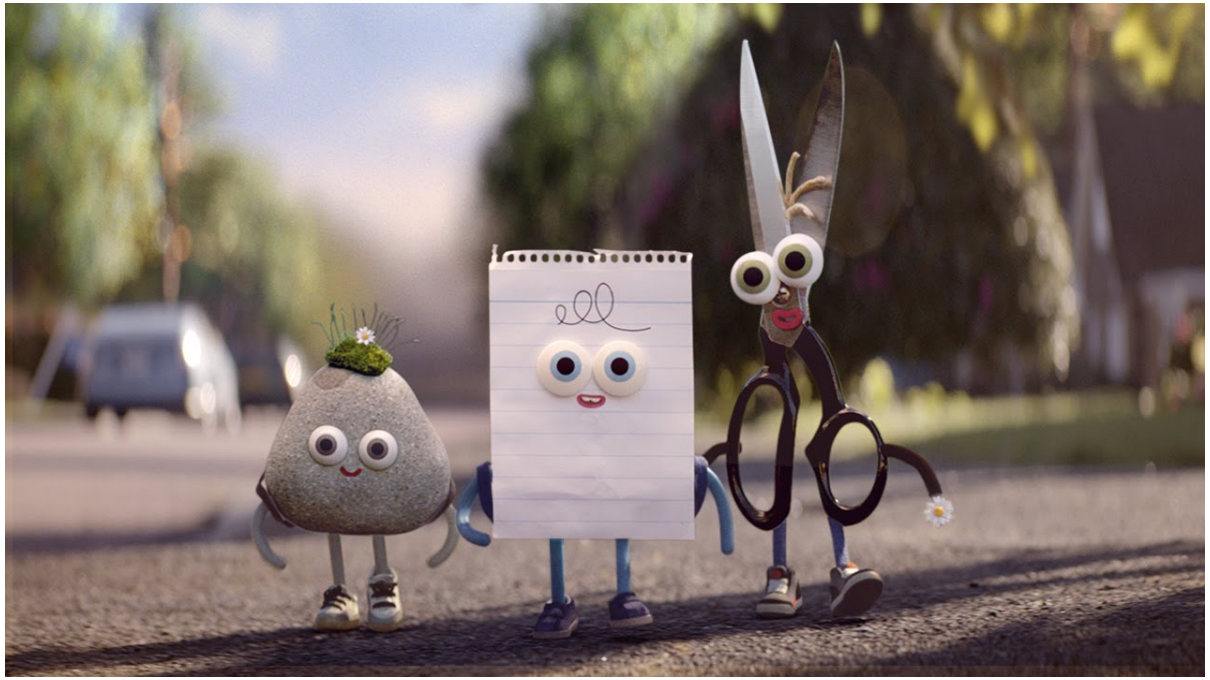
"It is important to understand data science even if you never intend to apply it yourself. Data-analytic thinking enables you to evaluate proposals for data mining projects. For example, if an employee, a consultant, or a potential investment target proposes to improve a particular business application by extracting knowledge from data, you should be able to assess the proposal systematically and decide whether it is sound or flawed. ... [Y]ou should be able to spot obvious flaws, unrealistic assumptions, and missing pieces."
-- Data Science for Business

How can you be successful in this class?

- Do the required readings, and then be active in discussions on Slack!
- Be ready to “get your hands dirty” in lab
- Collaborate with classmates (it's required!)
- Always ask questions!
    - I only know what is interesting to ME
    - Tell me what is interesting to YOU, and I can talk about it

Let's Get Started!

Play a Game



Start by playing a simple game (I recommend about 20 rounds) of Rock, Paper, Scissors

Where do we actually use models like this?

The Economist gives us a few examples:

- Predict Repeat Offenders
- Predict Heart Attacks in Hospital Patients
- Predict at-risk Students, Police Officers
- Find children at risk of lead poisoning

Why are these models so valuable to us?

Computers have the ability to "see" lots of data points at once, so if we use algorithms carefully, we can utilize vast amounts of data to inform our decisions.

What is data?

Data is a representation of objects or information from the real world

What is data?

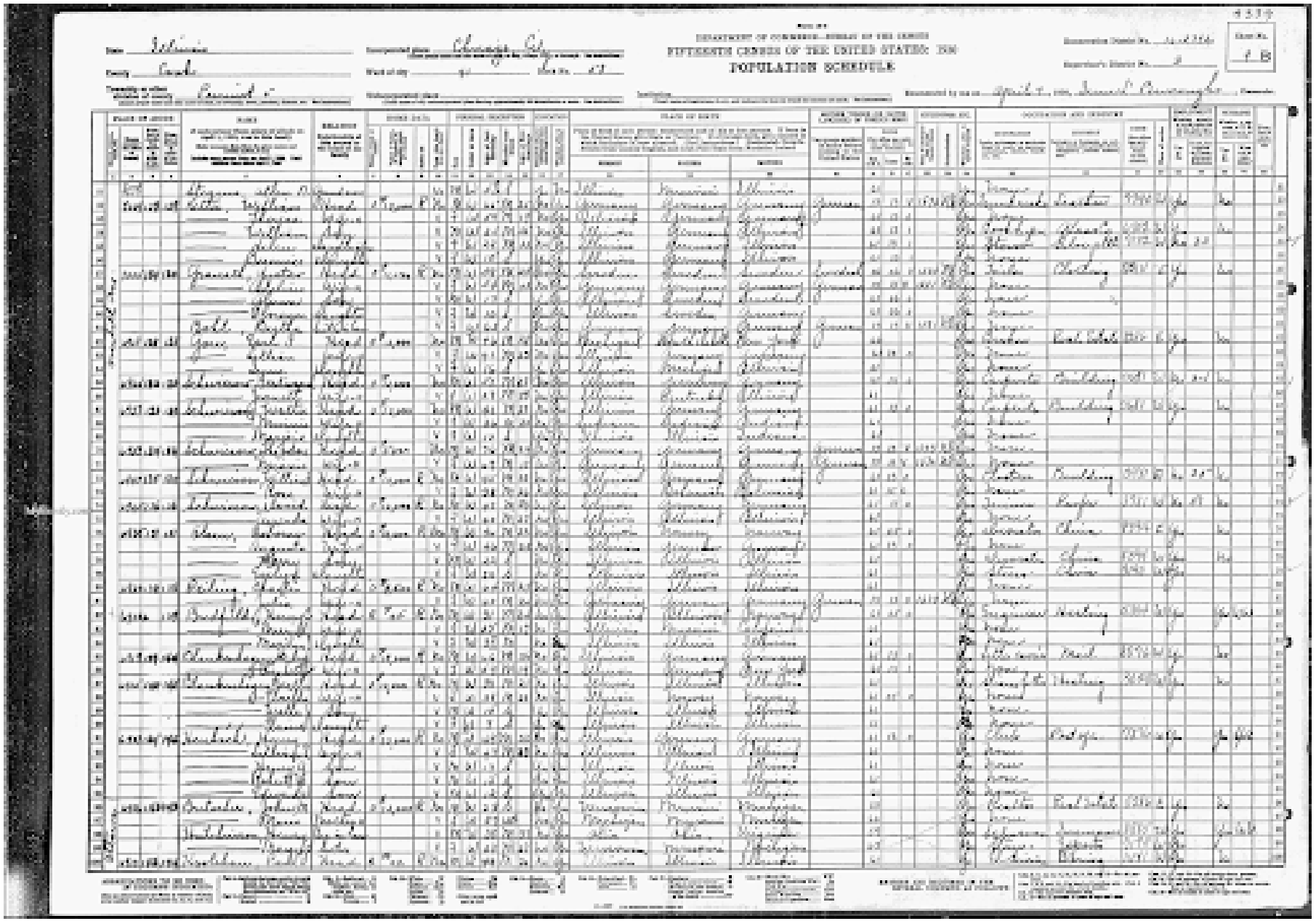
Data is a representation of objects or information from the real world

- Census responses are data

DEPARTMENT OF COMMERCE—BUREAU OF THE CENSUS

FIFTEENTH CENSUS OF THE UNITED STATES: 1930

POPULATION SCHEDULE



State: Illinois

County: Cook

Township or other division of county: Precinct 5

Incorporated place: Chicago City

Ward of city: 41

Enumeration District No.: 14-2326

Supervisor's District No.: 8

Sheet No.: 4 B

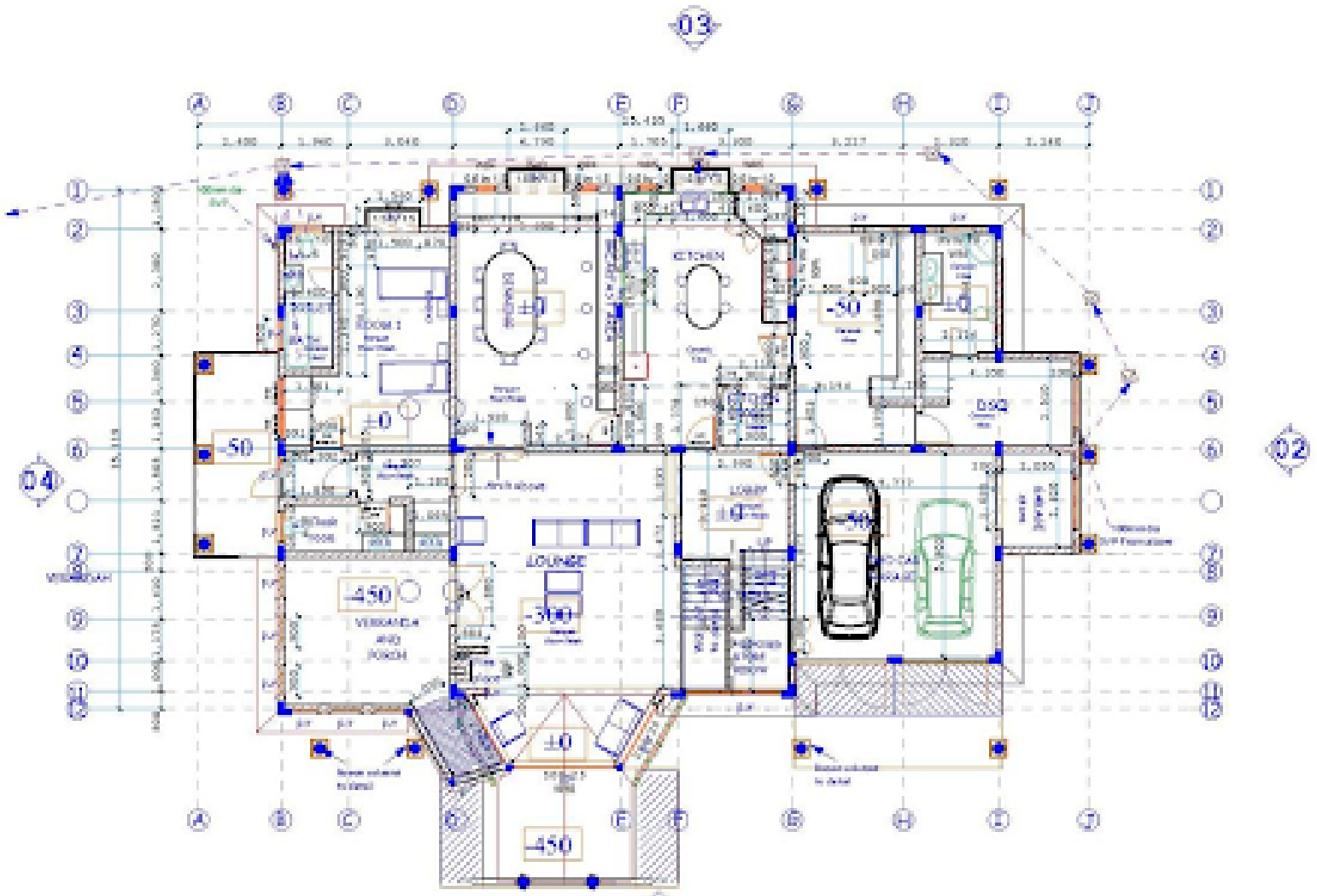
Enumerated by me on April 7, 1930. Jennie C. Burroughs, Enumerator.

Line	Street	House No.	Dwelling No.	Family No.	Name	Relation	Home owned or rented	Value or rent	Radio	Sex	Color or race	Age	Marital status	Age at first marriage	Attended school	Can read	Can write	Place of birth: Person	Father	Mother	Mother tongue	Year of immigration	Naturalized or alien	Year of naturalization	Occupation	Industry	Employer, salary or wage worker, or own account	Hours worked	Weeks worked	Veteran	Farm schedule
1	[illegible]				Stigum, Allen D.	Lodger				M	W	28	S					Illinois	Norway	Illinois					None						
2	[illegible]				Lette, William	Head	O	7500	R	M	W	44	M	22				Germany	Germany	Germany	German	45			Bookkeeper	Leather	[illegible]	[illegible]		No	
3					— Theresa	Wife				F	W	44	M	22				Illinois	Germany	Germany		45			None						
4					— William	Son				M	W	20	S					Illinois	Germany	Illinois		45			Cashier	Bank	[illegible]	[illegible]		No	
5					— Allen	Son				M	W	15	S					Illinois	Germany	Illinois		45			Student	High school	[illegible]	[illegible]		No	22
6					— Bernice	Daughter				F	W	12	S					Illinois	Germany	Illinois		45			None						
7	[illegible]				Grannell, Peter	Head	O	7500	R	M	W	58	M	22				Germany	Germany	Germany	Germany	45			Tailor	Clothing	[illegible]	[illegible]		No	
8					— Julia	Wife				F	W	54	M	22				Germany	Germany	Germany	German	45			None						
9					— Thomas	Son				M	W	22	S					Illinois	Germany	Germany		45			None						
10					— Theresa	Daughter				F	W	20	S					Illinois	Germany	Germany		45			None						
11					Bohl, Bertha	Lodger				F	W	22	S					Illinois	Germany	Germany	German	45			None						
12	[illegible]				Gow, Earl F.	Head	O	7500		M	W	36	M	26				Illinois	Middle Atlantic	New York		45			Broker	Real Estate	[illegible]	[illegible]		No	
13					— Lillian	Wife				F	W	30	M	22				Illinois	Norway	Illinois		45			None						
14					— Jane	Daughter				F	W	16	S					Illinois	Norway	Illinois		45			None						
15	[illegible]				Schuerman, Gertrude	Head	O	7500		M	W	57	M	21				Illinois	Germany	Germany		41			Carpenter	Building	[illegible]	[illegible]		No	
16					— Jennie	Wife				F	W	51	M	21				Illinois	Germany	Germany		41			Carpenter	Building	[illegible]	[illegible]		No	
17					— Minnie	Niece				F	W	24	S					Illinois	Indiana	Indiana		41			None						
18					— Herman	Nephew				M	W	12	S					Illinois	Illinois	Indiana		41			None						
19	[illegible]				Schuerman, Nicholas	Head	O	7500		M	W	70	M	18				Germany	Germany	Germany	German	45			None						
20					— Rosina	Wife				F	W	62	M	18				Germany	Germany	Germany	German	45			None						
21	[illegible]				Schuerman, William	Head	O	7500	R	M	W	38	M	25				Illinois	Germany	Germany		45			Clerk	Building	[illegible]	[illegible]		No	
22					— Rose	Wife				F	W	34	M	21				Illinois	Germany	Illinois		45			None						
23	[illegible]				Schuerman, David	Head	O	7500	R	M	W	41	M	21				Illinois	Germany	Germany		41			Foreman	Roofing	[illegible]	[illegible]		No	
24					— Jennie	Wife				F	W	41	M	21				Illinois	Germany	Illinois		41			None						
25	[illegible]				Olsen, Antonia	Head	O	7500	R	M	W	50	M	20				Illinois	Norway	Norway		41			Operator	China	[illegible]	[illegible]		No	
26					— Augusta	Wife				F	W	46	M	20				Illinois	Norway	Norway		41			None						
27					— Mary	Daughter				F	W	24	S					Illinois	Illinois	Illinois		41			Operator	China	[illegible]	[illegible]		No	
28					— Lloyd	Son				M	W	22	S					Illinois	Illinois	Illinois		41			None						
29	[illegible]				Reiling, Charles	Head	O	7500	R	M	W	64	M	24				Illinois	Illinois	Illinois		41			None						
30					— Julia	Wife				F	W	60	M	24				Illinois	Illinois	Illinois		41			None						
31	[illegible]				Budfield, Herman F.	Head	O	7500	R	M	W	40	M	21				Illinois	Illinois	Norway	German	45			Engineer	Heating	[illegible]	[illegible]		No	
32					— Bernice	Wife				F	W	42	M	21				Illinois	Norway	Norway		45			None						
33					— Marion	Daughter				F	W	22	S					Illinois	Illinois	Illinois		45			None						
34	[illegible]				Clinkenbeard, Ralph	Head	O	7500	R	M	W	36	M	22				Illinois	Germany	Germany		45			Letter carrier	Mail	[illegible]	[illegible]		No	
35					— Ethel	Wife				F	W	36	M	22				Illinois	Germany	Germany		45			None						
36	[illegible]				Blankenberg, Fred	Head	O	7500	R	M	W	50	M	21				Illinois	Illinois	Illinois		45			Stamp fitter	Heating	[illegible]	[illegible]		No	
37					— Bertha	Wife				F	W	48	M	21				Illinois	Norway	Norway		45			None						
38					— Ruth	Daughter				F	W	18	S					Illinois	Illinois	Illinois		45			None						
39					— Frank	Son				M	W	7	S					Illinois	Illinois	Illinois		45			None						
40	[illegible]				Kenbeck, Henry	Head	O	7500	R	M	W	48	M	20				Illinois	Germany	Germany		45			Clerk	Postoffice	[illegible]	[illegible]		No	
41					— Lillian	Wife				F	W	47	M	20				Illinois	Illinois	Illinois		45			None						
42					— Henry	Son				M	W	13	S					Illinois	Illinois	Illinois		45			None						
43					— Robert	Son				M	W	12	S					Illinois	Illinois	Illinois		45			None						
44					— Edward	Son				M	W	10	S					Illinois	Illinois	Illinois		45			None						
45	[illegible]				Outerbridge, John M.	Head	O	7500	R	M	W	50	M	22				Michigan	Michigan	Michigan		45			Broker	Real Estate	[illegible]	[illegible]		No	
46					— Anna	Wife				F	W	47	M	22				Michigan	Michigan	Michigan		45			None						
47					Hutchinson, Henry	Lodger				M	W	30	M	22				Illinois	Illinois	Michigan		45			Salesman	Insurance	[illegible]	[illegible]		No	
48					— Margaret	Lodger				F	W	24	S					Michigan	Michigan	Michigan		45			Teacher	School	[illegible]	[illegible]		No	
49	[illegible]				Kohlman, Paul	Head	O	7500	R	M	W	48	M	20				Illinois	Illinois	Illinois		45			Printer	Library	[illegible]	[illegible]		No	

What is data?

Data is a representation of objects or information from the real world

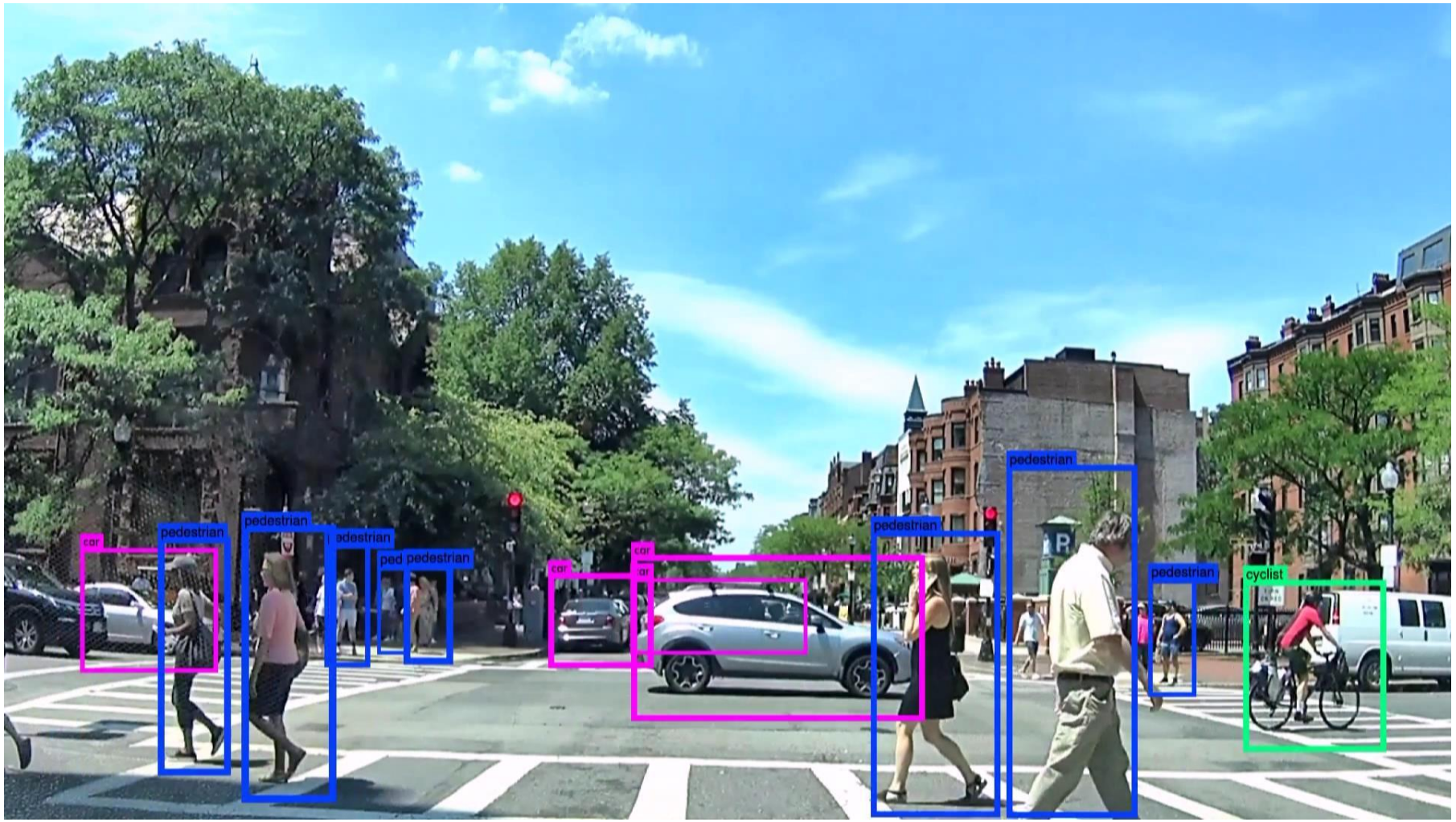
- Census responses are data
- Blueprints are data


03
02
04
01
KITCHEN
LOUNGE
VERANDAH AND PORCH
-50
±0
-450
-300
DSQ
GROUND FLOOR
SCALE 1:150

What is data?

Data is a representation of objects or information from the real world

- Census responses are data
- Blueprints are data
- Images are data


car
pedestrian
pedestrian
edestrian
ped pedestrian
car
car
car
pedestrian
pedestrian
pedestrian
cyclist

What is data?

Data is a representation of objects or information from the real world

- Census responses are data
- Blueprints are data
- Images are data
- Financial records are data

Cash Flow

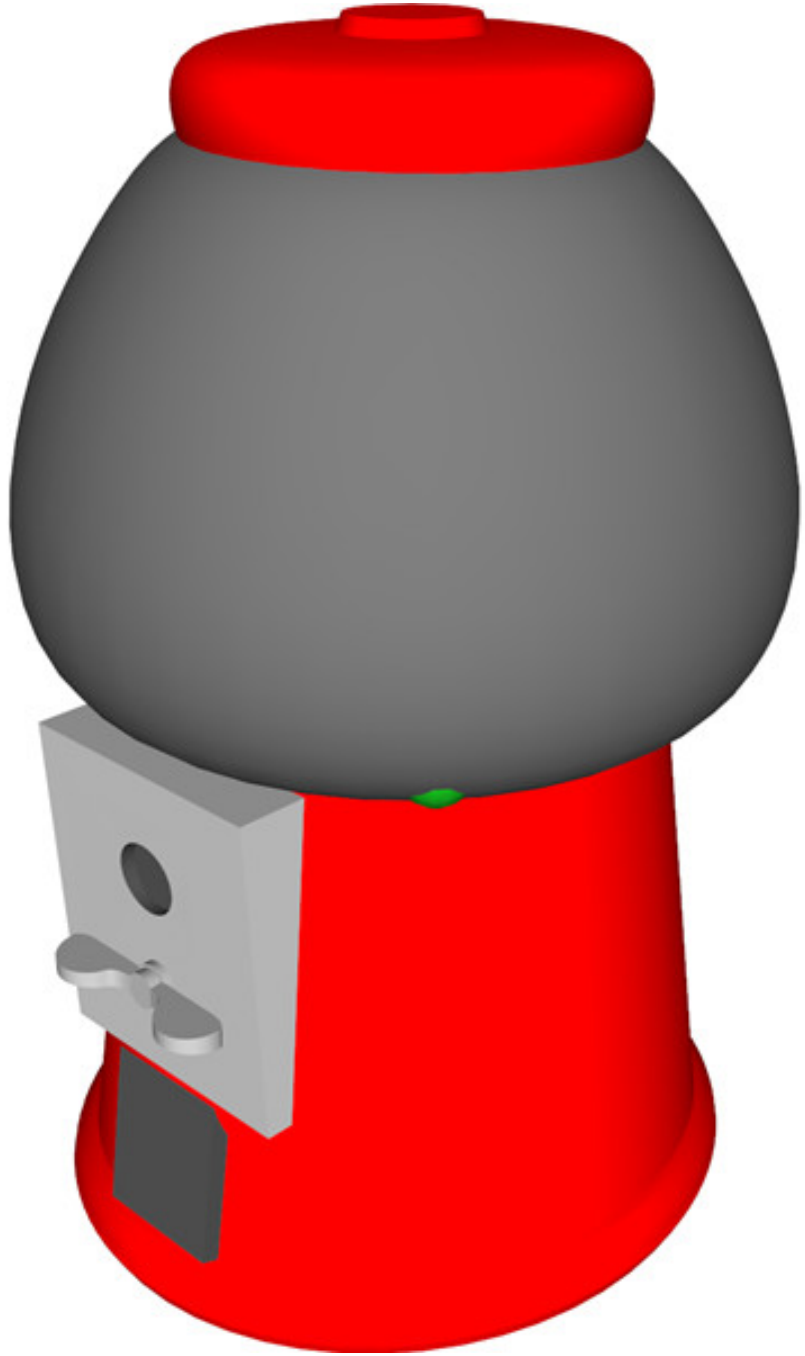
Date:	February 2013

Account	Description	March 2000	February 2000	Current Change	YTD March 2000
Cash Flow From Operations					
0	Net Income	365,102	361,218	3,884	3,884,060
1100	Net change in Accounts receivable - trade	202,826	(299,201)	502,028	(3,067,762)
1110	Net change in Accounts receivable - other	1,310	1,179	131	14,938
1199	Net change in Allowance for bad debts	4,373	4,327	47	46,523
1200	Net change in Inventory	(5,386)	(10,027)	4,642	(208,522)
1900	Net change in Other current assets	(1,804)	(2,035)	231	(22,825)
3000	Net change in Accounts payable	40,428	40,086	342	595,083
3100	Net change in Accrued payroll	(491)	(406)	(85)	(5,492)
3400	Net change in Other current liabilities	12,819	12,146	673	141,702
Total	**Cash Flow From Operations**	**619,177**	**107,286**	**511,891**	**1,377,705**
Cash Flow From Investments					
2900	Increases in Other assets	(398)	(420)	22	(5,126)
2000	Decreases in Fixed assets	354	346	8	4,027
2100	Decreases in Accumulated depreciation	23,107	22,862	246	245,823
Total	**Cash Flow From Investments**	**23,063**	**22,788**	**276**	**244,724**
Cash Flow From Financing					
3500	Increases in Long-term notes	279	318	(39)	4,163
3900	Increases in Other liabilities	261	189	72	2,506
4000	Increases in Capital stock	211	203	8	2,501
Total	**Cash Flow From Financing**	**751**	**710**	**41**	**9,170**
Total Cash Flow		**642,991**	**130,783**	**512,208**	**1,631,599**

What is data?

Data is a representation of objects or information from the real world

- Census responses are data
- Blueprints are data
- Images are data
- Financial records are data
- CAD designs are data



Defining Data Analytics

Good luck! 😛

But seriously, **Data Analytics** is a notoriously broad term describing pretty much anything that uses some data with any amount of analysis of that data to make decisions. Let's try some better terminology...

Definitions

Predictive Analytics - Using data to predict outcomes on an individual level

Forecasting - Using data to predict macro-level outcomes

Why now? Why not when computers were invented?

- More Data
  - There has been a tremendous increase in the amount of data available to companies
- More Computer Power
  - In 2005, my laptop had ¾GB of RAM, and cost over $2000
  - As of 2014, my laptop (still at just over $2000) had 16GB of RAM, nearly a 2000% increase! The same was true of the increase in processor performance.
  - Today, a similar laptop can be purchased for closer to $1000

Why now?

- Synergy between Data and Computation Power
  - We now have enough data to make meaningful conclusions
  - We have enough computational power!
  - OmniSci, more OmniSci
  - NVIDIA GTX 1080 Ti: ~11 TFlops (11 trillion operations per second)
    - Only costs $599

So where can we get the data we need?

Anywhere that records are kept, if the owner is willing to share.

- Facebook, Twitter, and LinkedIn all collect very robust information on users and their interactions
- We could also record our own data if needed: surveys, usage statistics, credit card transactions, etc.

So where can we get the data we need?

Some of my all-time favorites:

- IPUMS.org
- Data.gov
- ICPSR@UMich

So the data will magically analyze itself, right?!

- Just like a book that is never read, data that is only recorded and never used is not that useful
- We need to find a way to UTILIZE data
- That is what we will learn about during the rest of this class!

We need to ask questions!

Exploring data helps us to find the questions we care about

Let's use Tableau to explore our data

For lab today

- Complete part 1 of the first project
  - Find a group (use Slack!)
  - Play with the example data in Tableau
    - Variable descriptions are available here
  - Choose a research question

If you need, remember that Tableau provides all sorts of great training videos. I have posted some of my favorites on Canvas.