COGS 1: Spring 2019

Section **H**, Week 9

Professor Boyle	mboyle@ucsd.edu	Friday, 2-4 pm	CSB 130
Lauren	lcurley@ucsd.edu	Tuesday, 10-11 am	CSB 225
Lexi D.	adalenco@ucsd.edu	Tuesday, 12:30-1:45 pm	The Loft
Elena	edreisba@ucsd.edu	Thursday, 1-2 pm	CSB 114
Adrian	ajm033@ucsd.edu	Wednesday, 5-6 pm	CSB 114
Audrey	aberardi@ucsd.edu	Tuesday, 4-5 pm	CSB 114
Devansh	d4agarwa@ucsd.edu	Monday, 4-5 pm	CSB 114
Lori	rol044@ucsd.edu	Monday, 10-11 am	CSB 114
Lexi F.	adfrankl@ucsd.edu	Thursday, 4-5 pm	CSB 114

Important Information

- **COGS 1 Website completely updated - all readings for the rest of the quarter, syllabus, etc.**

- **Week 9 Quiz: Take Home** 😀
  - ONLINE: 5/31 (Opens Friday @ 4pm) - 6/2 (Closes Sunday @ 6pm), time is STILL LIMITED but you can take it whenever convenient.
  - Make sure to check TritonEd for exact times - give yourself **lots** of extra time to allow for internet/technical issues.

- **EXTRA CREDIT QUIZ**: June 3 @ 4PM - June 4 @ 10AM on *Gut Instinct*

Last Week's Topics

- **[Lecture 14 | Dr. Voytek: Introduction to Data Science](#)**

Lecture 14 | Review Questions (1 of 2)

1. What is data science?

2. Why are models important in data science, according to Dr. Voytek?

3. What is cognitive science according to Dr. Voytek?

4. What is "cognition"? Do other animals have cognition? Provide examples.

5. What is the link between data science and cognitive science? How does this apply to Parkinson's disease?

Lecture 14 | Review Questions (2 of 2)

6.	How does UBER make use of data? What predictions can be derived from data collected? How does that impact and inform their business model?

7.	Compare and contrast the Brown Corpus with the Google Corpus. Why is "big data" so powerful?

8.	What are n-grams?

9.	What is the Semantic Web? What are the main orthogonal problems?
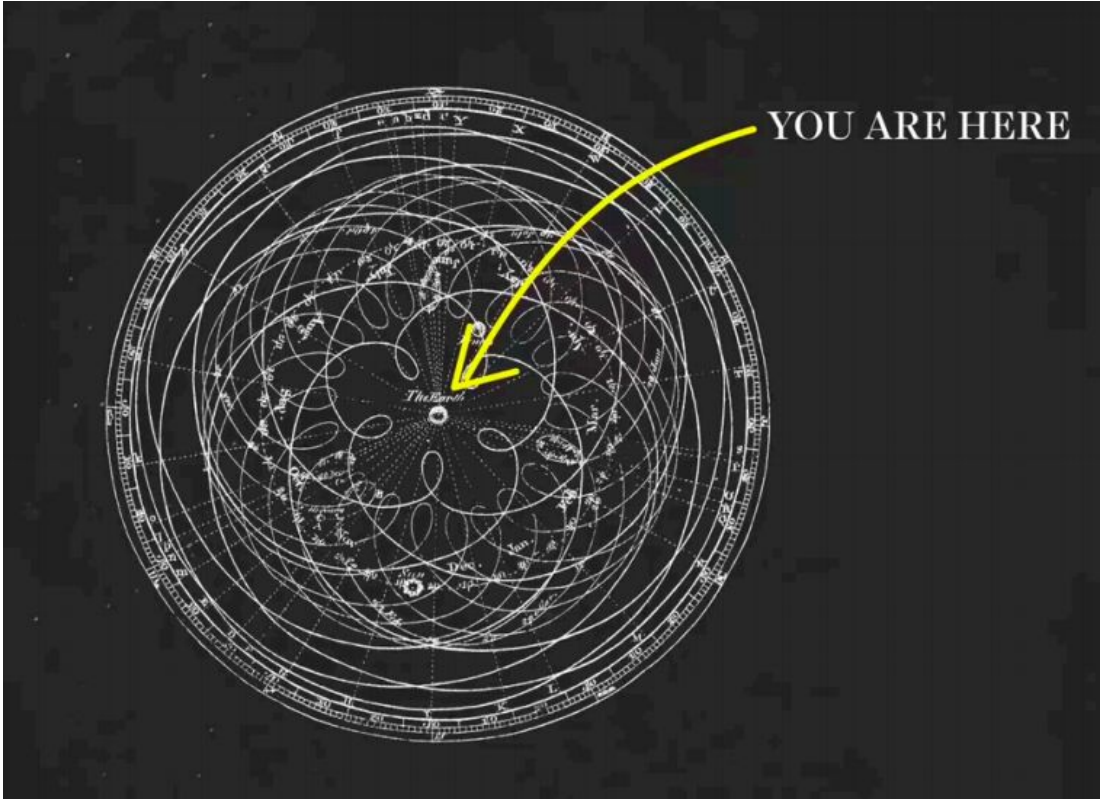
1. What is data science?

Utilizing Data to Elicit Truths

- **Ptolemaic Model of Solar System** - Incorrect (Geocentric) but predictive
- **Scientific Method** - Hypothesis driven, experimental technique, resulting in a conclusion.
- **Rigor and Intuition** - Not only computing math but also applying it in real life.
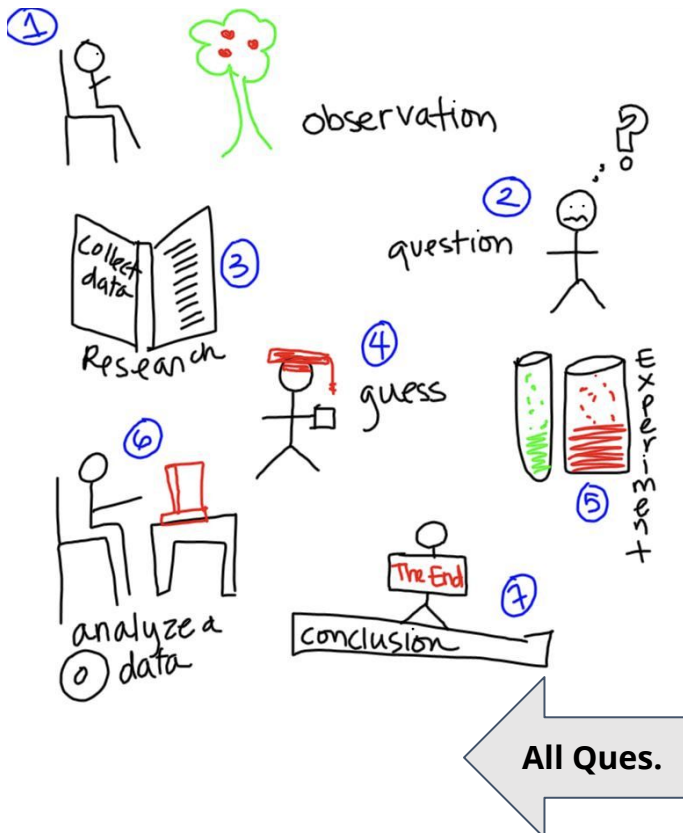
1. What is data science?

2. Why are models important in data science, according to Dr. Voytek?

Purpose of Modeling

- **Prediction** - Use previous data to forecast future events.
- **Classification** - Categorize multiple data points into meaningful sets.
- **Knowledge Discovery** - Uncover new and unknown facts from data.
- **Doing Useful Stuff** - Create better drugs, design a better product, arrive at accurate conclusions, etc
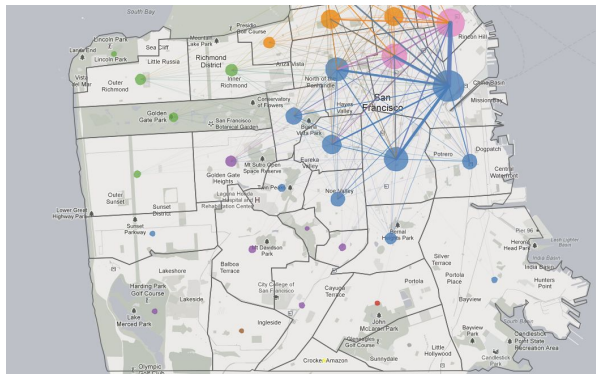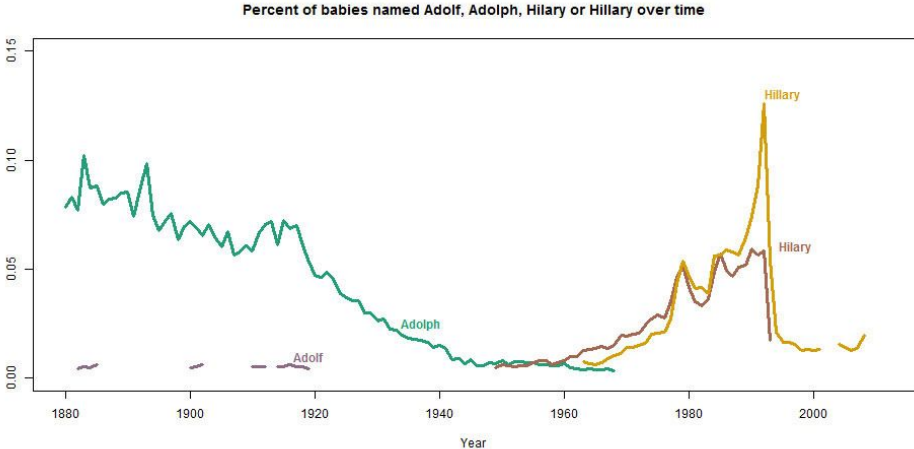
All models are wrong but some are useful!

All Ques.
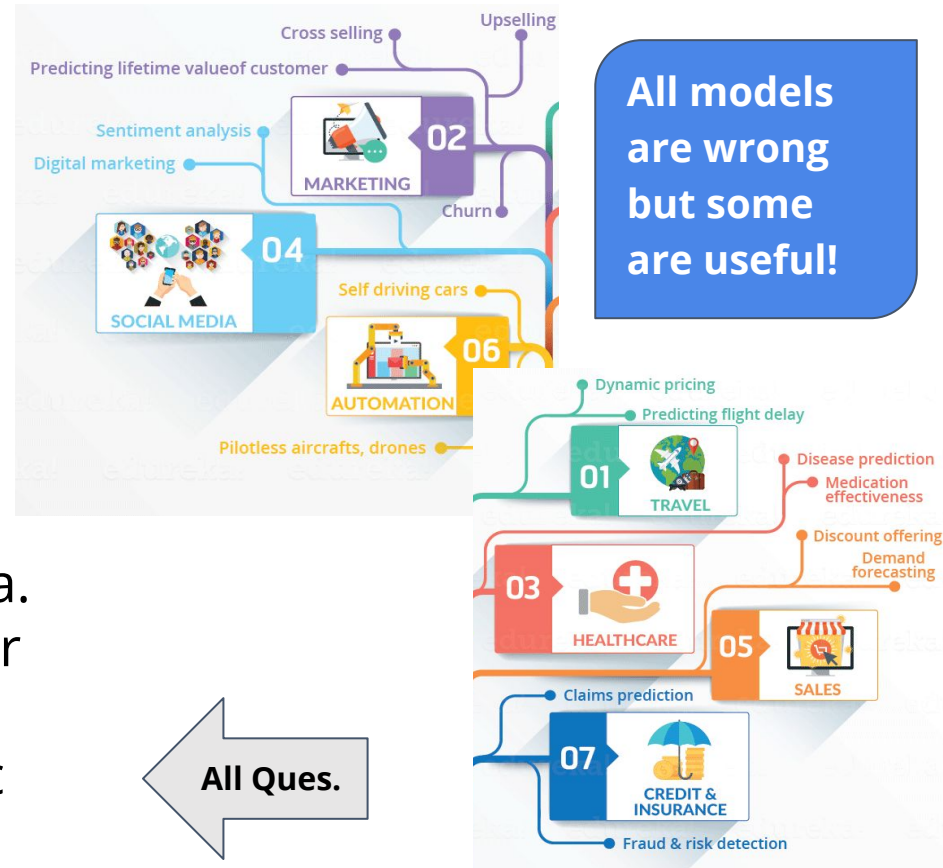
3. What is cognitive science according to Dr. Voytek?

Study of Intelligences

- **Human (Natural) Side -** Instances include changes in baby names over time (toxicity and popularity), verb regularization (tense forms) and culturomics (word usage).

- **Computer (Artificial) Side -** Parametrization of data to convert it into actionable business ideas (Uber).

Percent of babies named Adolf, Adolph, Hilary or Hillary over time

Hillary

Hillary

Adolph

Adolf

Percent

Year

All Ques.

4. What is "cognition"? Do other animals have cognition? Provide examples.

Multi Faceted, Observed in Various Organisms

- **Involves:**
  - Communication
  - Computation
  - Reasoning
  - Memory
  - Planning/Decision Making
- **Seen in (not just humans)**:
  - Mammals, Primates
  - Crows
  - Octopuses
  - Slime Molds
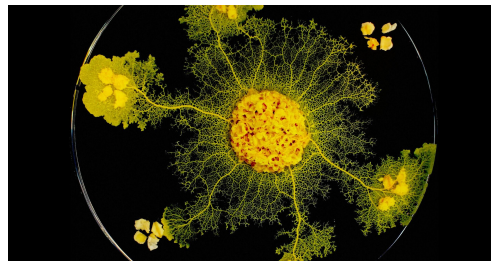
5. What is the link between data science and cognitive science? How does this apply to Parkinson's disease?
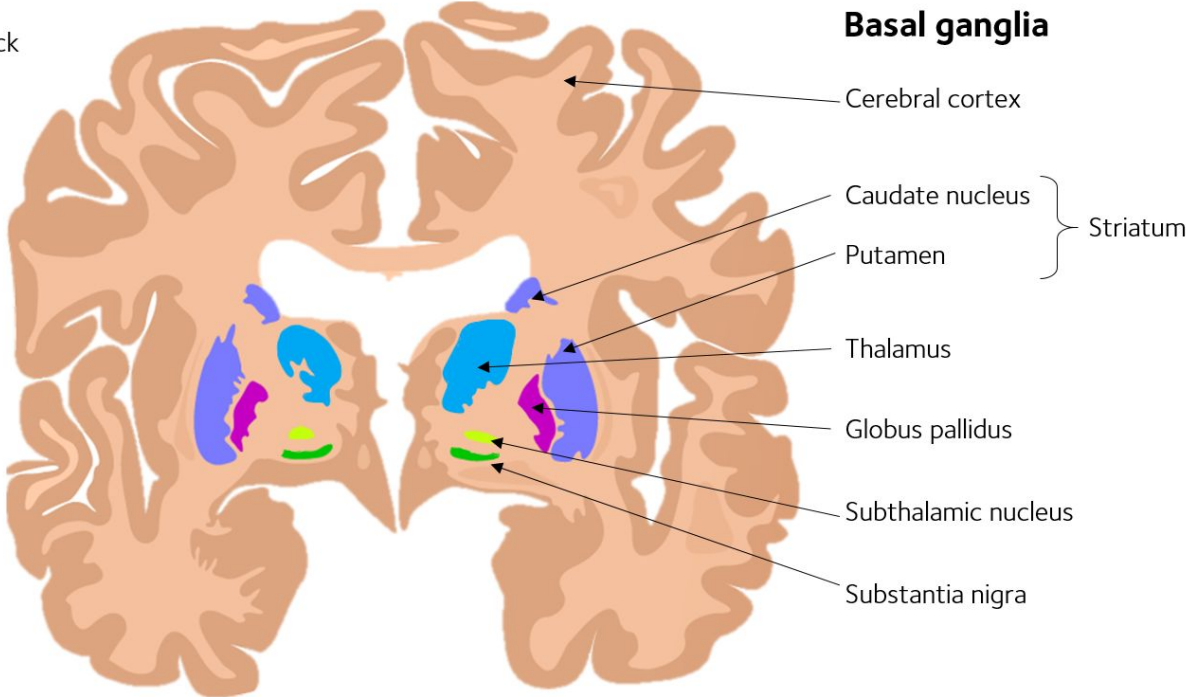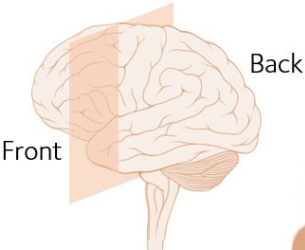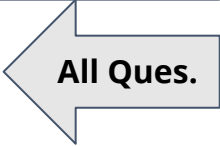

Basal ganglia

Cerebral cortex

Caudate nucleus — Striatum
Putamen

Thalamus

Globus pallidus
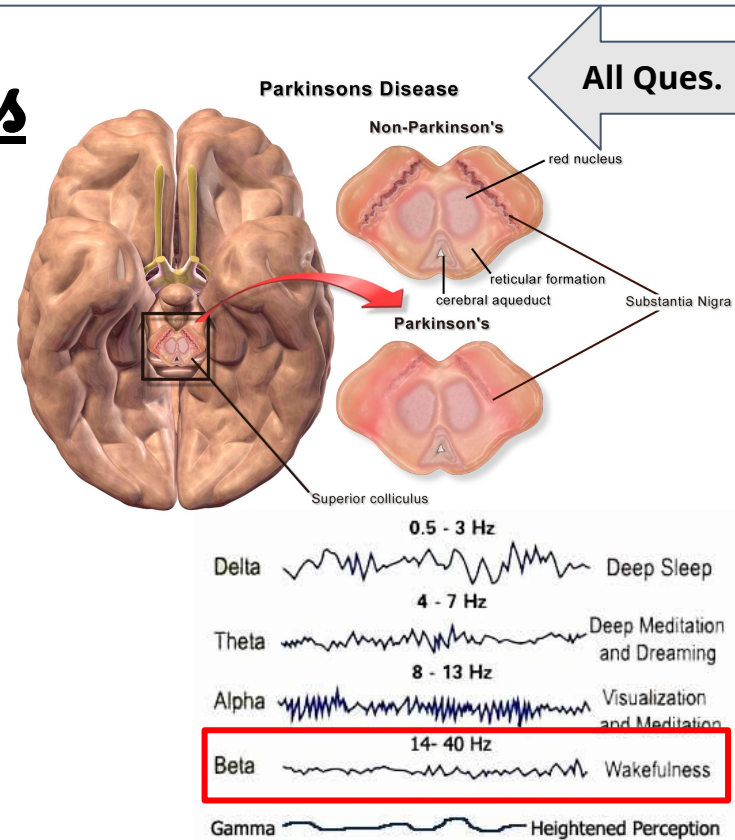
Subthalamic nucleus

Substantia nigra

Back

Front

All Ques.

5. What is the link between data science and cognitive science? How does this apply to Parkinson's disease?

Using Data to Treat Disorders

- **Parkinson's Disease:**
  - Degeneration of Dopaminergic Neurons
  - Located in Substantia Nigra
  - Symptoms: Slow movement, walking/balance impaired
- **Data science application:**
  - Analyze abnormal beta oscillations
  - Conduct Deep Brain Stimulation to reduce some symptoms.

Parkinsons Disease

Non-Parkinson's

red nucleus

reticular formation

cerebral aqueduct

Substantia Nigra

Parkinson's

Superior colliculus

	0.5 - 3 Hz	
Delta		Deep Sleep
	4 - 7 Hz	
Theta		Deep Meditation and Dreaming
	8 - 13 Hz	
Alpha		Visualization and Meditation
	14- 40 Hz	
Beta		Wakefulness
Gamma		Heightened Perception

6. How does UBER make use of data? What predictions can be derived from data collected? How does that impact and inform their business model?

Spatiotemporal Dynamics

- **Aggregate user data:** Across time and distance to infer neighborhood type within cities.

- **Understand demand for taxi:** Correlate neighborhoods within and between cities for e.g. there exists greater correlation of taxi demand in financial districts of two different cities than the taxi demand in two different districts of the same city.
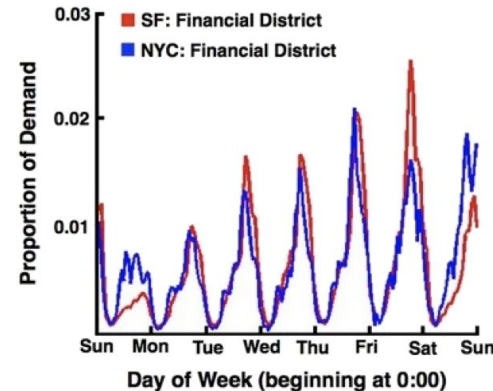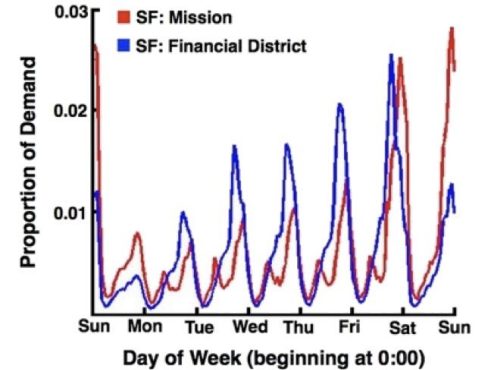
All Ques.

7. Compare and contrast the Brown Corpus with the Google Corpus. Why is "big data" so powerful?

Google Corpus

- Comprised of a trillion words
- All sequences up to five words long
- Incomplete sentences, spelling errors and grammatical errors.

Brown Corpus

- One million english words.

All Ques.

Categorize every rare aspect of human behavior

8. What are *n*-grams?

Definition

- Short sequences of consecutive words.

This is Big Data AI Book

Uni-Gram	This	Is	Big	Data	AI	Book

Bi-Gram	This is	Is Big	Big Data	Data AI	AI Book

Tri-Gram	This is Big	Is Big Data	Big Data AI	Data AI Book

Usage

- Used for speech recognition and machine translation.
- Present in a huge database of probabilities of n-grams.

All Ques.

9. What is the Semantic Web? What are the main orthogonal problems?

Orthogonal Problems

- Choose a representation language.
- Encode a model in the language.
- Perform inference on the model.

Semantic Web

- A convention for formal representation languages.
- Permits software services to interact with one another.

9. What is the Semantic Web? What are the main orthogonal problems?

Semantic Web versus Semantic Interpretation

The Semantic Web is a convention for formal representation languages that lets software services interact with each other "without needing artificial intelligence."[11] A software service that enables us to make a hotel reservation is transformed into a Semantic Web service by agreeing to use one of several standards for representing dates, prices, and locations. The service can

service interoperability. Semantic interpretation deals with imprecise, ambiguous natural languages, whereas service interoperability deals with making data precise enough that the programs operating on the data will function effectively. Unfortunately, the fact that the word "semantic" appears in both "Semantic Web" and "semantic interpretation" means that the two prob-