

# Class Connections

John Lee

# I want to min/max my classes

- What class fulfil prerequisites for the most classes?
  - APPM 2360 (Introduction to Differential Equations with Linear Algebra): 75
- What class takes the most prerequisites?
  - ATLS 4889 (Interactive Machine Learning for Customizable and Expressive Interfaces): 19
- Which classes are connected to another?
  - About half
- Which classes are similar to another?
  - We'll see



# Data

- BeautifulSoup for Web Scraping in Python
  - Get prerequisite, equivalent and recommended classes
  - Get description, class name, credits, requirements
- 
- 8680 classes
  - 10078 total connections
  - 4868 classes with connections



# HITS

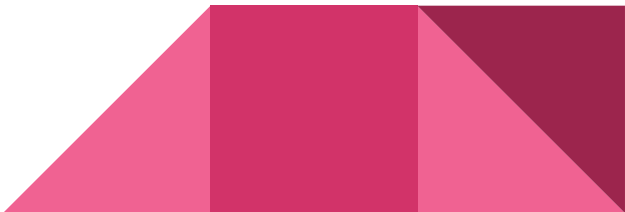
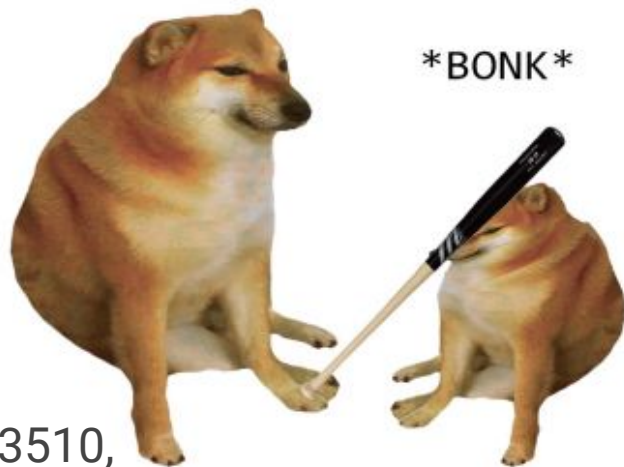
HIT it with the data science stick

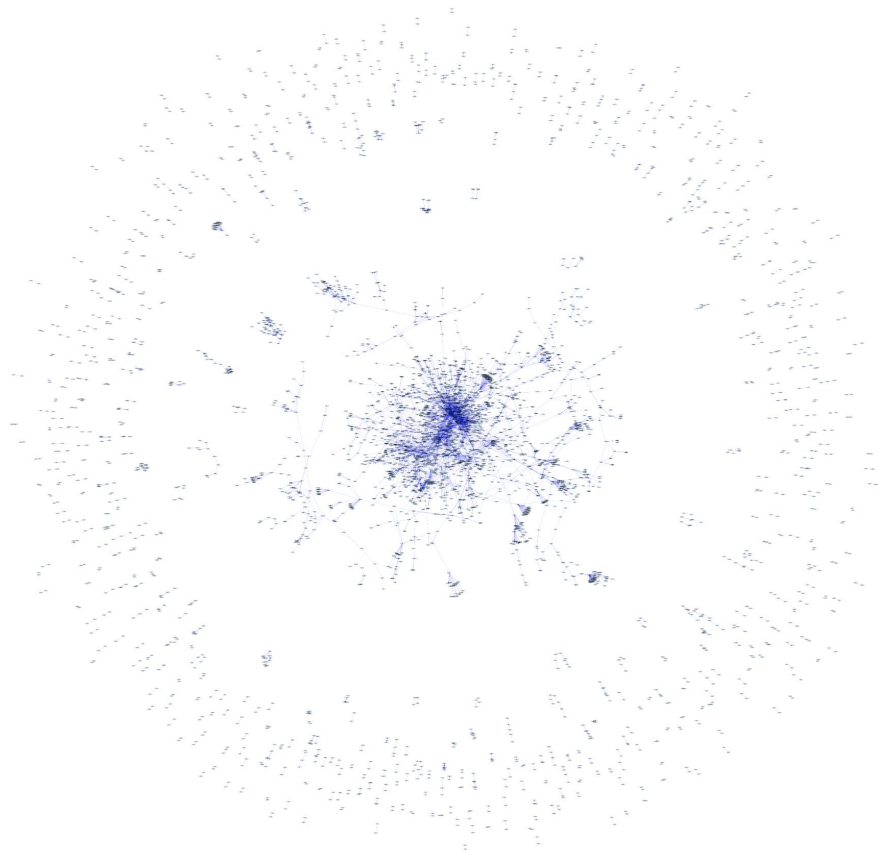
Hubs:

[ECON3818, STAT3100, CVEN3227, CSCI3022, MATH3510,  
CHEN3010, MATH2130, STAT4000, STAT4520, APPM3310]

Authorities:

[CSCI4802, CSCI4022, CSCI4622, CSCI4889, ATLS4889,  
CSCI3202, CSCI4200, CSCI5622, ECON4858, ECON4211]





## Second Stick: TF-IDF

Run sklearn's TfidfVectorizer

Words with highest TF-IDF sums:

[students, topics, research, course, examines,  
design, studies, social, explores, introduces]

Classes with highest TF-IDF sums:

[EBIO5050, EBIO4050, ASEN5067, ASEN4067, CESR4130,  
MGMT4130, HIST3020, CSCI5160, ORGL5405, ECEN5154]



## Second Stick: TF-IDF

Cosine distance

Classes with highest cosine distance sum

[EDUC8103, MCDB3000, ATLS7500, MDST1001, INFO5651,  
INFO 4651, ATLS4000, ENGL3830, FYSM1000, GEEN4400]



# Future Work

Calculate an aggregated similarity score using department, college, course number, fulfilments, description for clustering, combine with FCQ data

Weighted pagerank with teleportation sets

Third 4022 stick of clustering

Given previous classes, recommend next classes

Interactive graph on web application using D3

