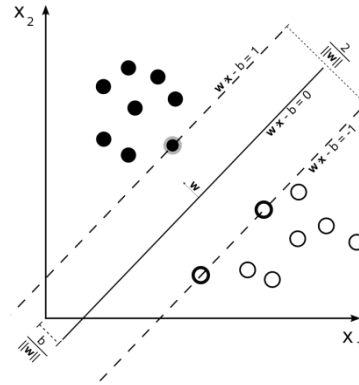


WHILE YOU'RE WAITING

MEET PEOPLE

ENJOY FOOD



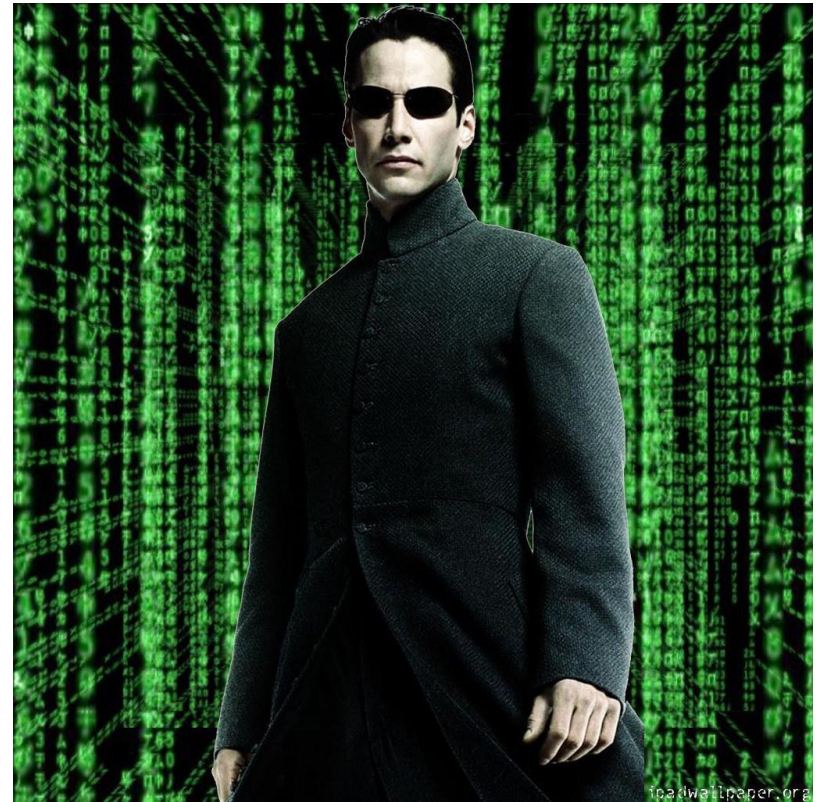
PRINCETON DATA SCIENCE

EVAN
FRANCINE
EMILY
ANNA

THE PLAN FOR TODAY

GROUP DATA PROJECT

FREE HACKING TIME



GROUP ASSIGNMENTS

SHOWER THOUGHTS



When Sweden is playing Denmark, it is SWE-DEN.



The remaining letters, not used, are DEN-MARK.

THE GOOGLE SELF-DRIVING CAR SHOULD
HAVE AN "I'M FEELING LUCKY" BUTTON
THAT DRIVES YOU TO A RANDOM LOCATION.



**Given a shower thought, can I
predict its popularity on Reddit?**

THE DATASET

1	created_utc	score	title
2	1361920466	2387	a tv channel where all the shows are set in the same universe

also have:

comments

upvotes

downvotes

EXPLORE

ANALYZE

PREDICT

SECTION 0-1

setup, import packages
read in data, some preprocessing

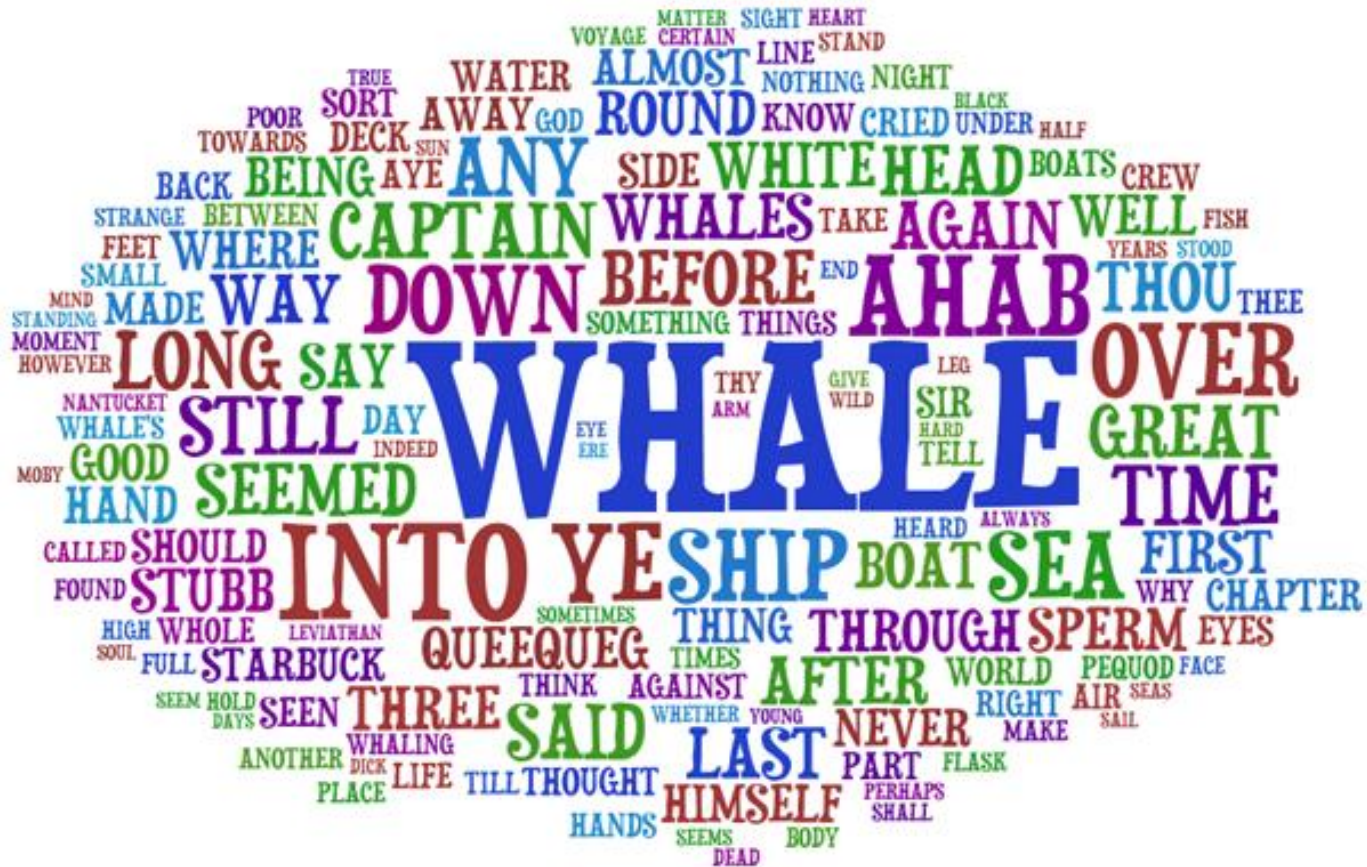
SECTION 2

exploring the numbers

SECTION 3

exploring the words

Wordcloud for Moby Dick



SECTION 4

some machine learning,
and a contest

Playing with SVM parameters

Example 1 (current version in code)

```
svm.model <- svm(x, y, # <-- don't change the data!  
  ## ADD/CHANGE YOUR PARAMETERS HERE  
  kernel="linear", cost=0.5, epsilon=0.2)
```

Example 2

```
svm.model <- svm(x, y, # <-- don't change the data!  
  ## ADD/CHANGE YOUR PARAMETERS HERE  
  kernel="radial", cost=1, epsilon=0.5, cross=5)
```

#Example 3

```
svm.model <- svm(x, y, # <-- don't change the data!  
  ## ADD/CHANGE YOUR PARAMETERS HERE  
  kernel="polynomial", C=1, cross=10)
```

SECTION 5

some extra fun generating shower thoughts

Questions?

Post-event questions:
datasci@princeton.edu

Free Hacking Time

- Further info, feedback, Q&A
 - PDS email: datasci@princeton.edu
 - Data science listserv: datascience@princeton.edu
 - Eddie Zhou '16, Evan Chow '16, Daway Chou-ren '16, Francine Camacho GS, Emily Huang '17, Anna Matlin '17
- Next steps
 - Go through the free book: “An Introduction to Statistical Learning, with Applications in R” (Gareth et al.)
 - Play around with datasets!
 - rs.io/100-interesting-data-sets-for-statistics/
 - <http://www.reddit.com/r/datasets>