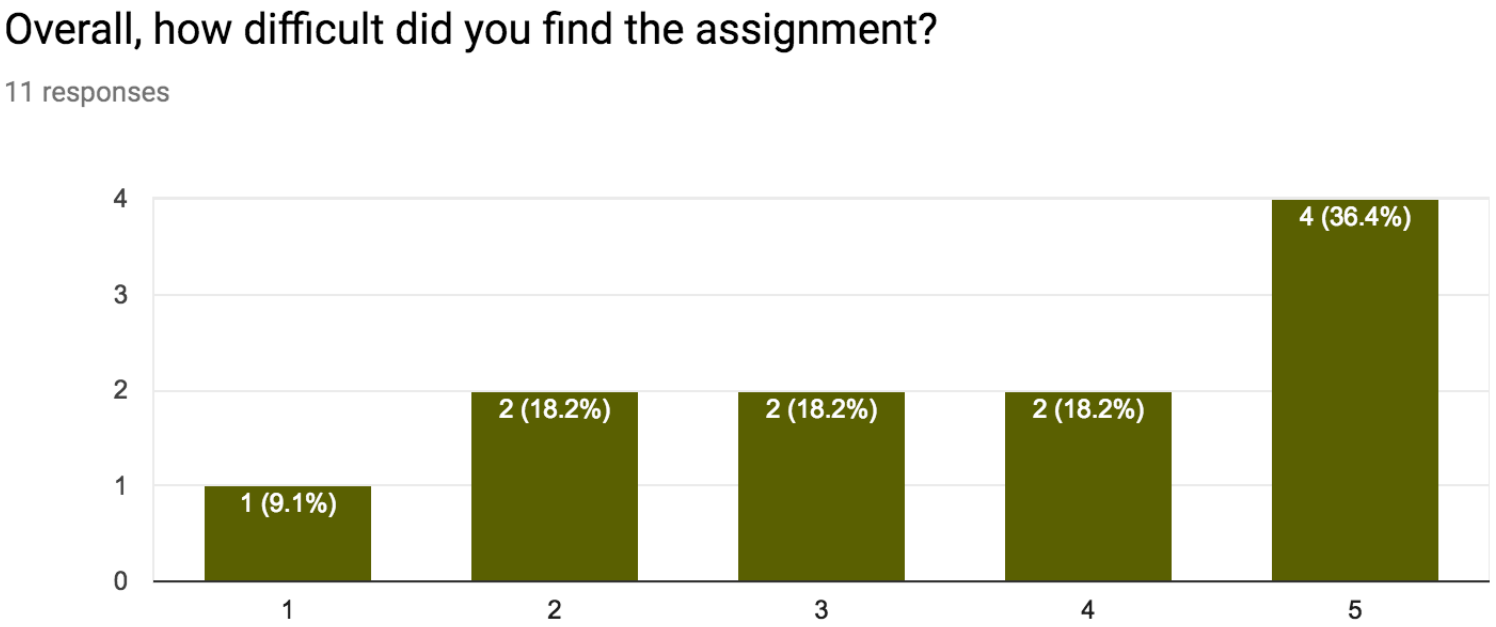
# Assignment 1 Postmortem

CSCI 4360/6360 Data Science II

Tuesday, September 5, 2017

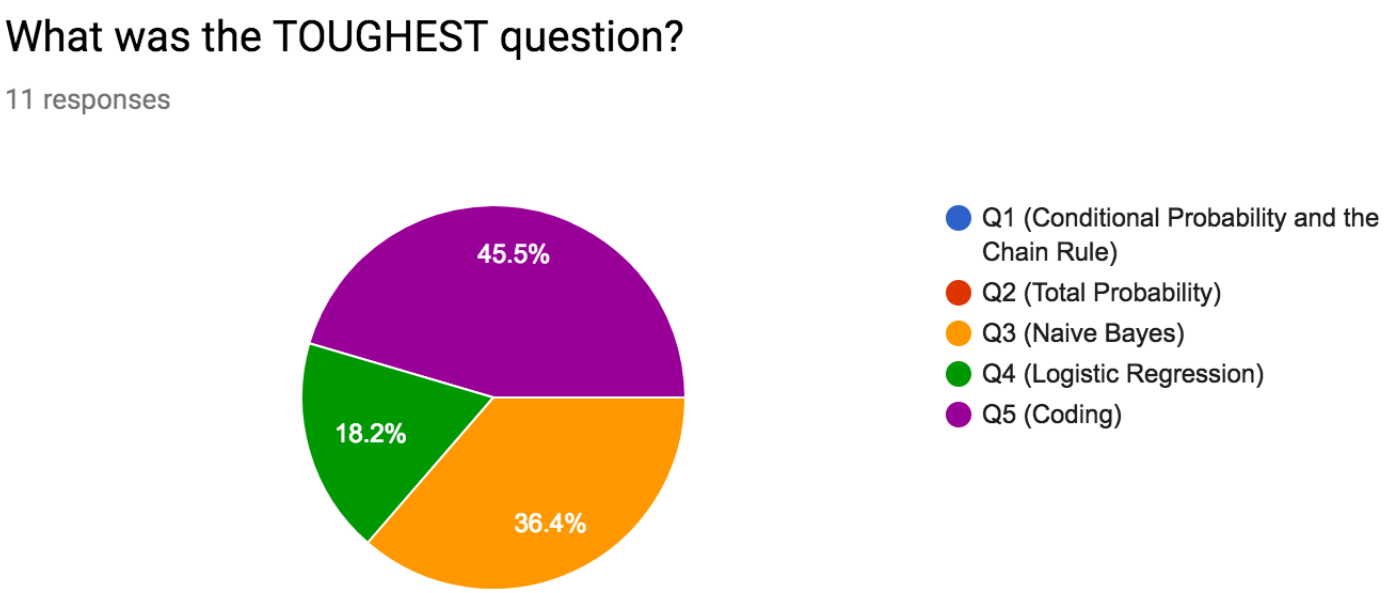
# Poll Review

- 11 responses (as of writing)



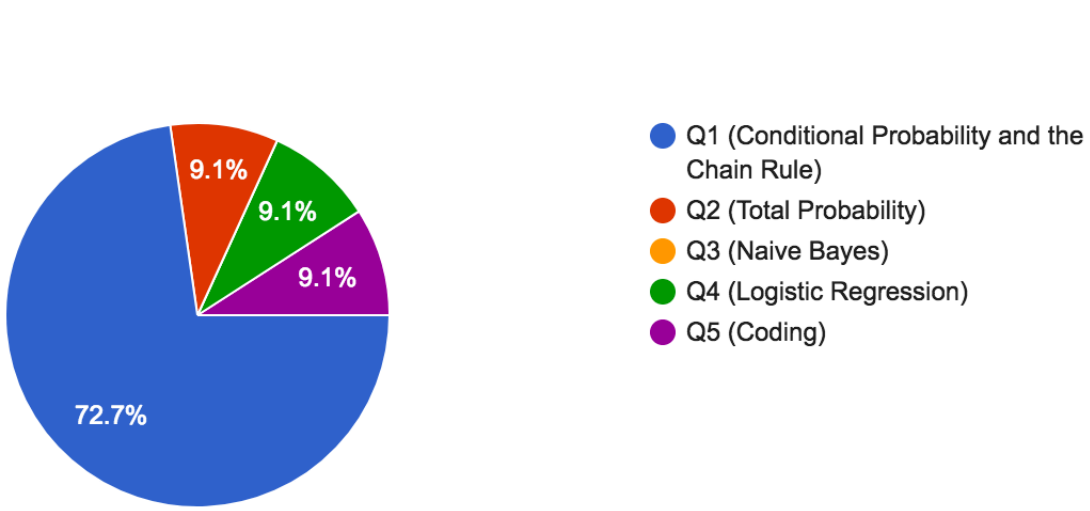
# Poll Review

- Q3: 4 votes
- Q4: 2 votes
- Q5: 5 votes



# Poll Review

- Clear winner: Q1
- (No opinion on Q2)



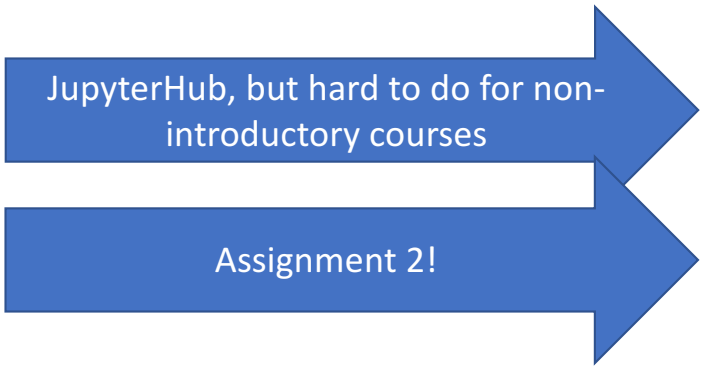
# Poll Review

- What should be **removed**?

- Nothing (7 votes)
- Command line / reading input (2 votes)
- Part of Q5 (1 vote)
- Q2 (1 vote)

# Poll Review

- What should be **added**?




- Nothing (3 votes)
- Clarifications (3 votes)
- Additional background review (2 votes)
- Additional proofs (1 vote)
- Jupyter notebook submissions (1 vote)
- Regularization or K-means (1 vote)

# Poll Review

- What should be **kept**?

- Everything (5 votes)
- Q5, i.e. coding (4 votes)
- Practice of concepts (1 vote)
- General probability (1 vote)

# Poll Review

- Comments

- Coding instructions
  - I/O details
  - Expected output format
- Background materials
  - Probability theory
  - LR / NB theory versus practice
- Other
  - LR decision boundary question
  - Prior knowledge
  - Intimidating

# Assignment 1 Review

**Fourth**, get ready for some math! If we have

$$l(W) = \sum_l Y^l \ln P(Y^l = 1|X^l, W) + (1 - Y^l) \ln P(Y^l = 0|X^l, W)$$

Expand the last term:

$$l(W) = \sum_l Y^l \ln P(Y^l = 1|X^l, W) + \ln P(Y^l = 0|X^l, W) - Y^l \ln P(Y^l = 0|X^l, W)$$

Combine terms with the same $Y^l$ coefficient (first and third terms):

$$l(W) = \sum_l Y^l \left[\ln P(Y^l = 1|X^l, W) - \ln P(Y^l = 0|X^l, W)\right] + \ln P(Y^l = 0|X^l, W)$$

Recall properties of logarithms--when subtracting two logs with the same base, you can combine their arguments into a single log dividing the two:

$$l(W) = \sum_l Y^l \left[\ln \frac{P(Y^l = 1|X^l, W)}{P(Y^l = 0|X^l, W)}\right] + \ln P(Y^l = 0|X^l, W)$$

Now things get interesting--remember earlier where we defined exact parametric forms of $P(Y = 1|X)$ and $P(Y = 0|X)$? Substitute those back in, and you'll get:

$$l(W) = \sum_l \left[Y^l (w_0 + \sum_i^d w_i X_i^l) - \ln(1 - \exp(w_0 + \sum_i^d w_i X_i^l))\right]$$

# Take-aways

- Clarity of the assignments absolutely needs to be improved
- A big part of data science is data preparation and munging… **BUT** outside of an assignment on that exact topic, time frame is a bit short
- Course prerequisites should be noted and not ignored… **BUT** I will absolutely strive to include background notes
- This is a 4000/6000 course. **It's supposed to be hard!**
- …but you can still **ask for help**!