

# Predicting End-User Adoption and Engagement

DAT 7 Course Project

Mike Yea

# Background and Project Question

- Project Background
  - Implemented web-based, FB-like tool in 2012 in a large (90K+ ) federal agency
  - Opt-in participation with adoption rate of 20% within 6 months.
  - Would like to revamp user engagement “lift” campaign
  - Messages (user posts) is building block of engagement

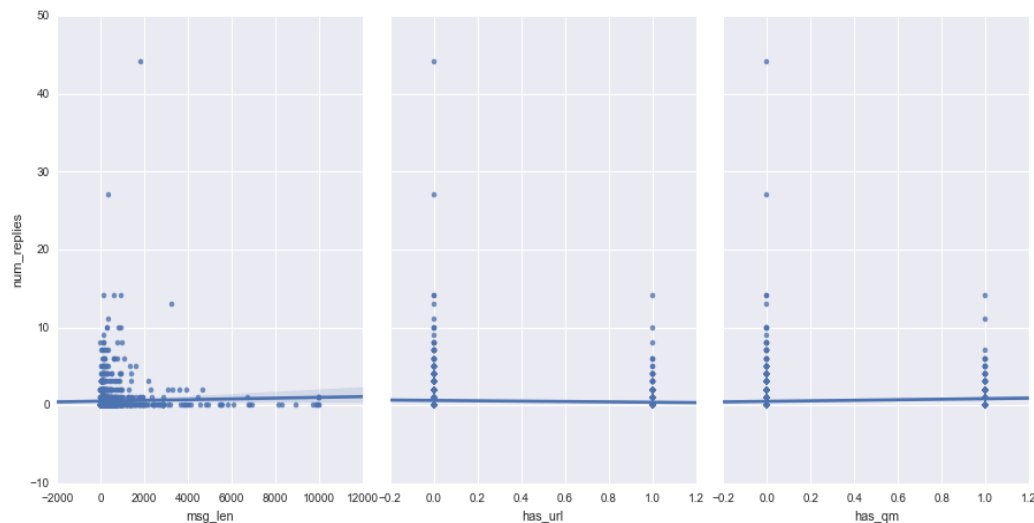
- Project Question:

**Can I predict a “lift” in user engagement (i.e., replies to messages) from message attributes (e.g., form of question, hyperlink, tag, key words in body of message, message tone/sentiment)?**

# Data

- Web interface has data export (in .csv): 6 months worth
  - Magic of “sys” to solve encoding errors
  - Only interested in public messages
  - Only interested in “top-level” messages or engagement “trigger”
- Extract data from un-normalized table to a data frame
  - Used two “for loops” on about 6,000 rows of data
  - String operations and Pandas “merge”

Figure: Scatter Plot of Response vs. Candidate Features



# Model and Still to Come...

Response	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6
num_replies	msg_len	has_attachment	has_url	in_group	has_qm	tone*

\*anticipated, but not yet implemented

Initial approach: Regression model

Next:

1. Features with more explanatory power?
2. 3-class classification model?
3. Use NLP to detect tone, key phrases?
4. Add more data (e.g., user profile)