Fitness Enthusiasts: How happy are they, really?

Team Post-Modern Bash:
Nameyeh Alam, Emre Celik, Madeleine Cieslak, Jessica Gutierrez, Stefan Sampaleanu

What questions did we ask?

- Main Question:
  - Are “fitness enthusiasts” happier than the average person on twitter?
- Sub-questions:
  - Which hashtags are popular in a random sample of “fitness enthusiasts?”
  - Is there a link between “fitness enthusiasts” and healthy eating?

Data Sources/ Analysis Tools

- Data Source
  - Twitter API
    - Generated two unique groups of users based on specific criteria
- Analysis Tools
  - VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media, and works well on texts from other domains (Hutto & Gilbert, 2014).
    - Performed sentiment analysis on each user using VADER Sentiment Analysis package
  - T-test
    - Compared group means of sentiment compound score

Data Exploration and Cleanup

- Identified search terms for both groups
- Determined “real person” parameters
  - Limited accounts by followers count, following count, tweet count etc
  - Exclude business/ abnormal accounts (gyms, trainers, celebs, etc.)
- Limitations:
  - Exceeding Twitter API rate limit
  - Determining criteria for normal population
  - Fitness data bias towards certain hashtags

“Fitness Users” Definition	“Normal Users” Definition
• Search terms: trending fitness hashtags • Following_count minimum of 50 • Followers_count maximum of 1500 • Tweet count between 100 and 20,000 • Excluded if “gym” in twitter handle	• Search term: “today” • Followers_count between 15 and 1500 • Tweet count between 100 and 20,000

Analysis: Tweet Sentiments by Group

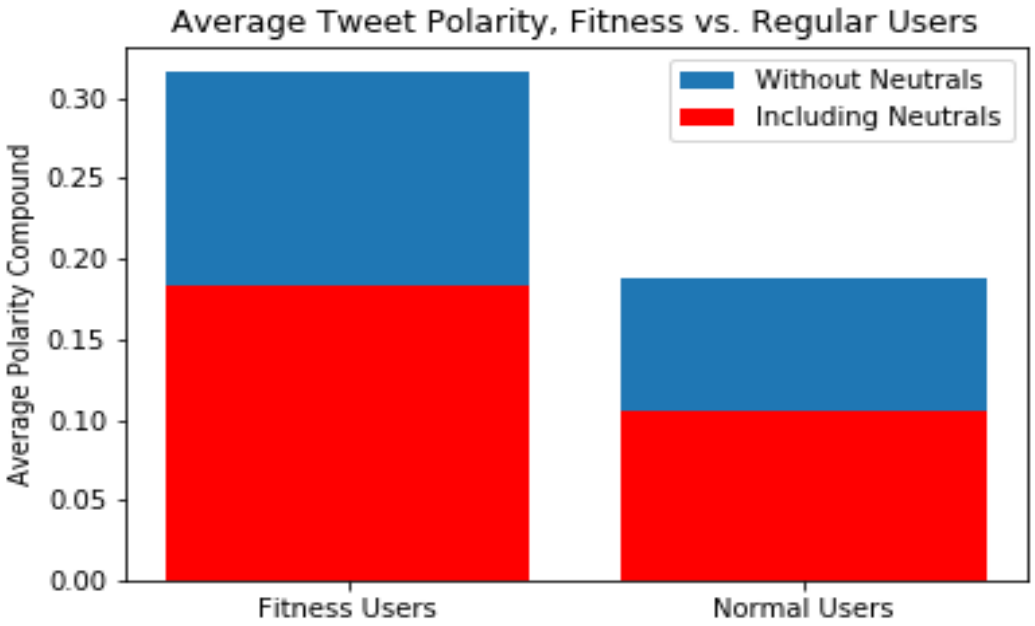
- Sentiments including neutrals

Value	Fitness User	Normal User
Compound	0.183	0.106
Positive	0.117	0.105
Negative	0.033	0.054
Neutral	0.851	0.841

- Sentiments without neutrals

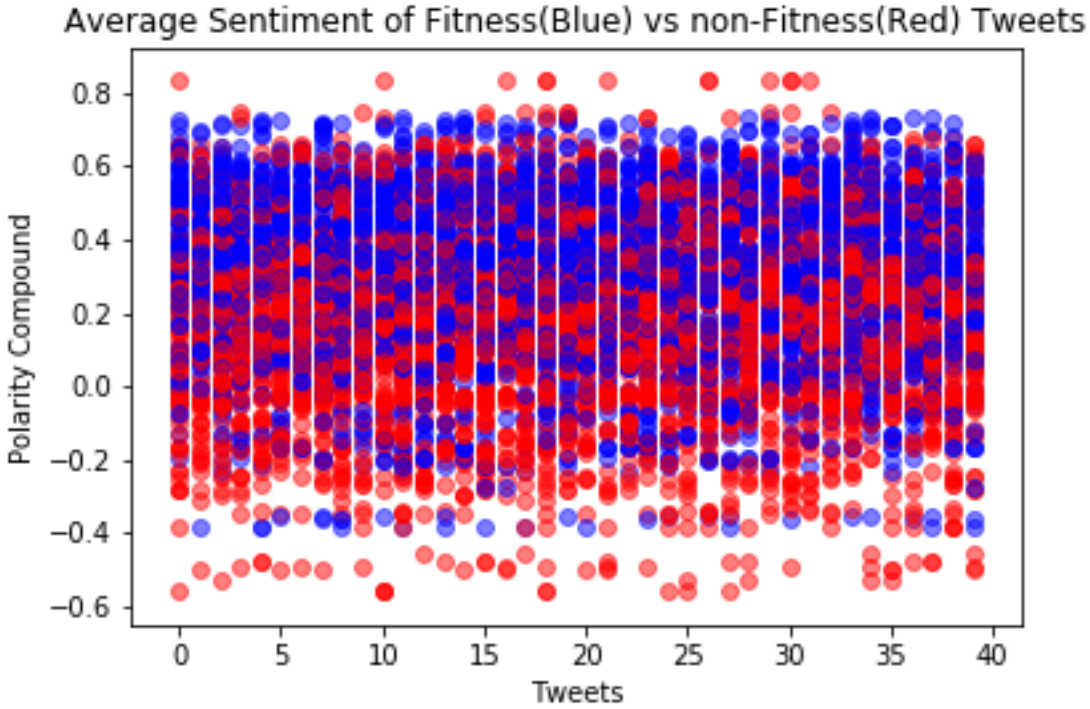
Value	Fitness User	Normal User
Compound	0.316	0.187
Positive	0.207	0.185
Negative	0.058	0.094
Neutral	0.736	0.721

Analysis: Tweet Polarity

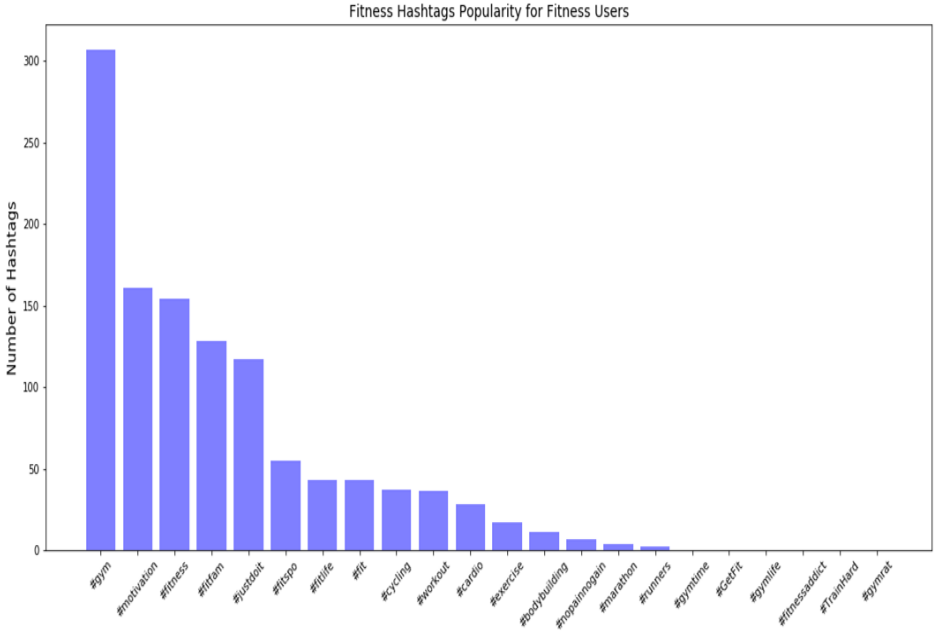


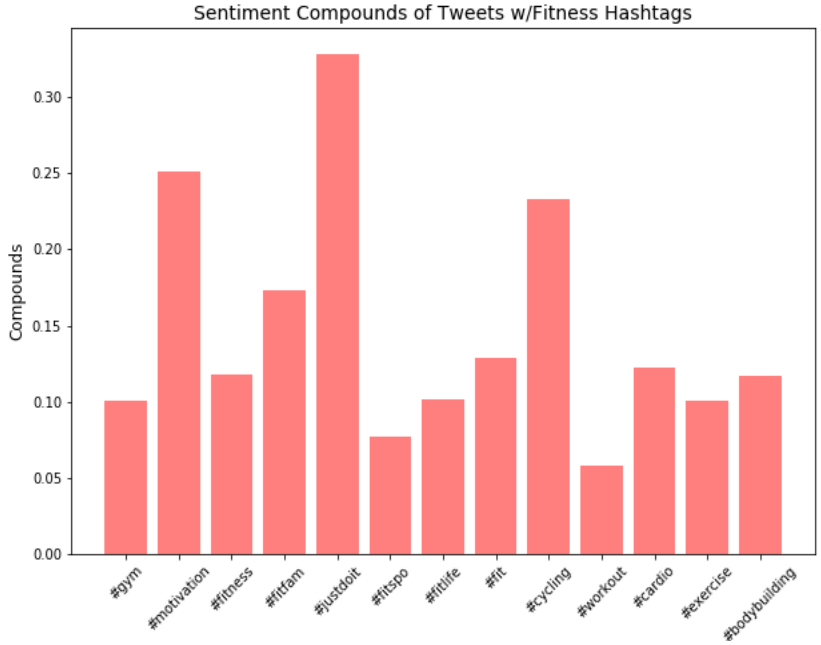
Neutrals?	N	Y
Mean Fit Sent	0.316	0.183
Mean Normal Sent	0.187	0.106
P-Value	1.023e-17	3.307e-16
T-Score	-8.740	-8.303

Analysis: Tweet Sentiments



Analysis: Fitness Users

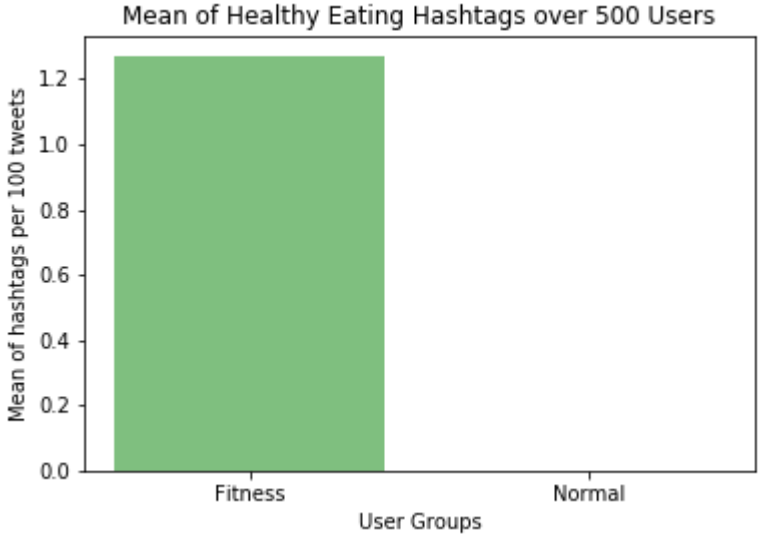


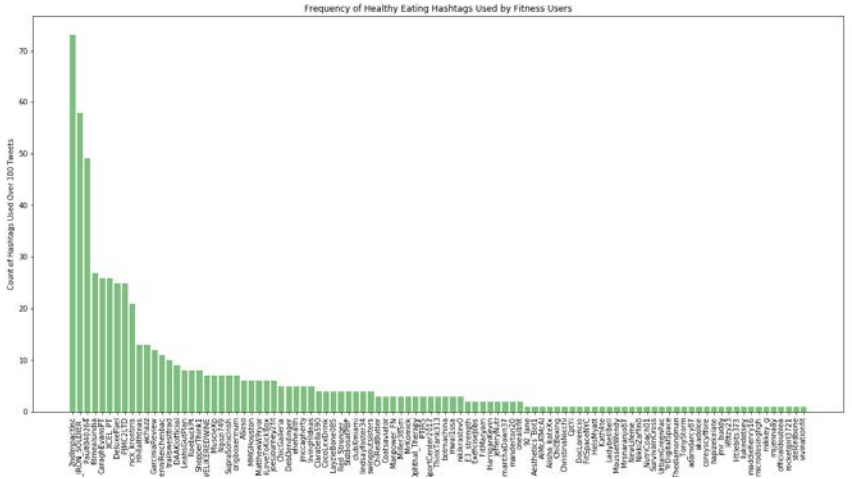


Analysis: Healthy Eating Comparison

- Is there a link between fitness and healthy eating?
- Defined healthy eating hashtags as EatClean, FitFood, GlutenFree, HealthyEating, HealthyRecipes, Nutrition, Protein, CleanEating, FitnessFood, Diet, PlantBased
- Looked at past 100 tweets for each of our users
- Counted how many times a user tweeted using at least one of the hashtags
- Results...

Analysis: Healthy Eating Comparison





	Avg of H.E. hashtags for fitness group	Avg of H.E. hashtags for normal group
0	1.270541	0.0

	Avg amt of tweets with H.E. hashtags	Avg amt of tweets with H.E. hashtags (less top 3)
0	6.40404	4.729167

Post Mortem

- Would we see a difference in fitness tweet trends based on the month/season?
- Would we see a difference in fitness tweet trends based on location?
- Did we subconsciously skew results to be more positive for our fitness users based on the way we picked our hashtags?

Questions?