STAT 139 Project Proposal

Group Members:
Yiming Xu, Zizheng Xu, Phoebe Wong, Nick Stern

Project Title:
What Determines Amazon Product Rating?

Motivation:
There exists ambiguity as to how people rate products. One line of logic may be that expensive products are higher quality, and thus have higher ratings. On the other hand, if the product is pricey, the reviewer may be more critical in their judgement of its worth. Our project aims to clarify this psychological quandary by using Amazon product reviews, prices, and ratings to elucidate the question of whether users prioritize ease over aesthetic.

Hypotheses of Interest:
1. Is there a correlation between product ratings on Amazon and various product-related features such as price, number of comments, and the superlatives used in comments?
2. Is there a significant difference in mean rating between product categories?
3. Is there a temporal effect on the mean rating within product categories?

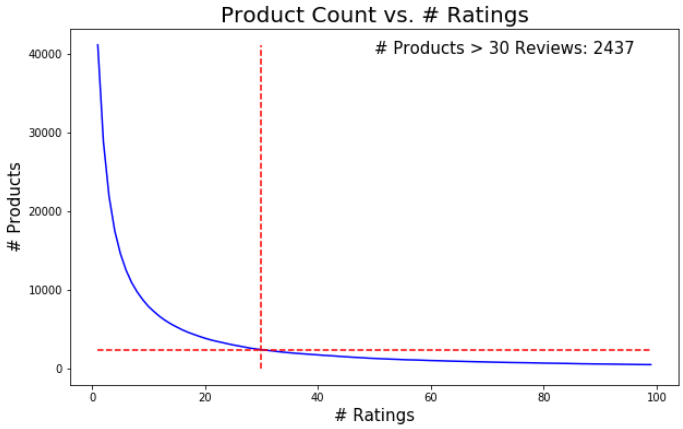
Variables of Interest:

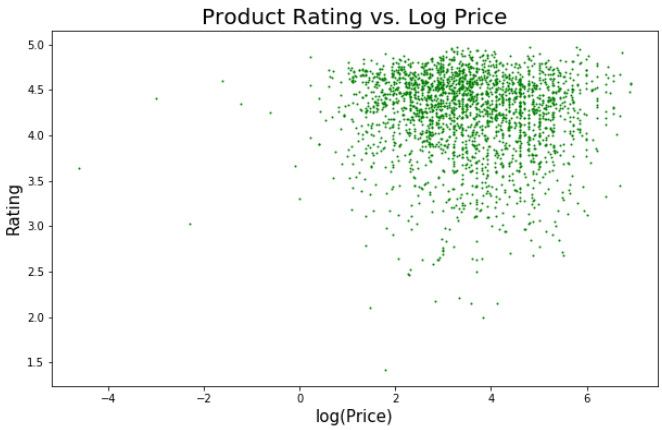
Response:
- Mean of user ratings of the product (each individually on a scale from 1-5)

Predictors: * for each distinct product
- Listed price
- Product category
- Number of comments
- Some NLP features of the comment (e.g., count of positive words)
- UNIX timestamp of review

*Note: We obtained the Amazon ratings/metadata from http://jmcauley.ucsd.edu/data/amazon/links.html with the permission from Professor Julian McAuley. The dataset spans May 1996 - July 2014.

Analysis Plan:
- We plan to run a multiple linear regression model to predict Amazon product rating based on user features (e.g., number of ratings, comment text) and product features (e.g., categories). Given the ratings are discrete integers, we plan to exclude products with < 30 ratings to treat the average ratings as a continuous distribution in our regression model. A scatter plot of mean rating vs. product price is provided below to demonstrate the impact of this choice for the musical instrument product category:





- We would like to run an ANOVA test and then comparison t-tests to investigate whether there are significant differences in average ratings between product categories. If the assumptions of these tests are not met, we may transform our data to meet the assumptions of the t-test, or defer to using a non-parametric test such as a Wilcoxon signed rank sum test.

- We will explore a basic time series analysis to understand whether there are significant temporal trends in how users rate products on Amazon.