



ANALYSIS OF TRIPADVISOR REVIEWS

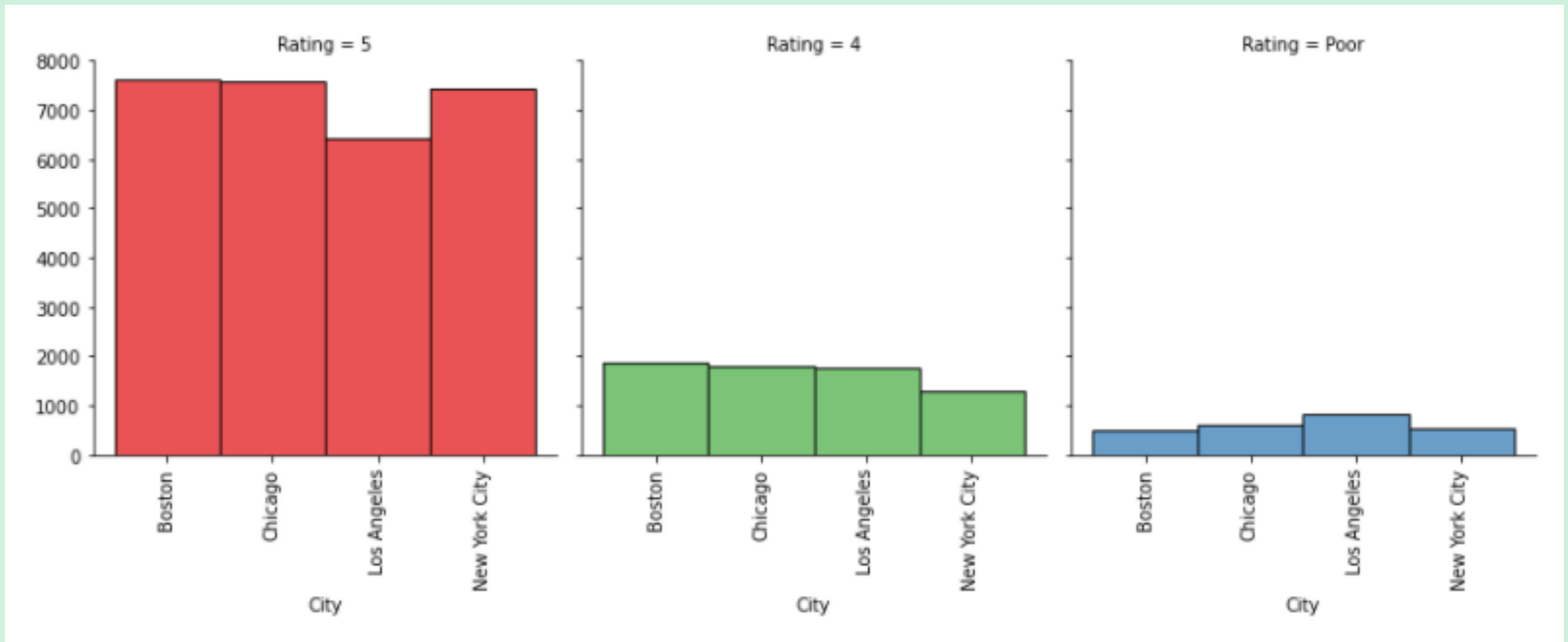
A WEBSCRAPING PROJECT BY THEODORE CHOI

BACKGROUND

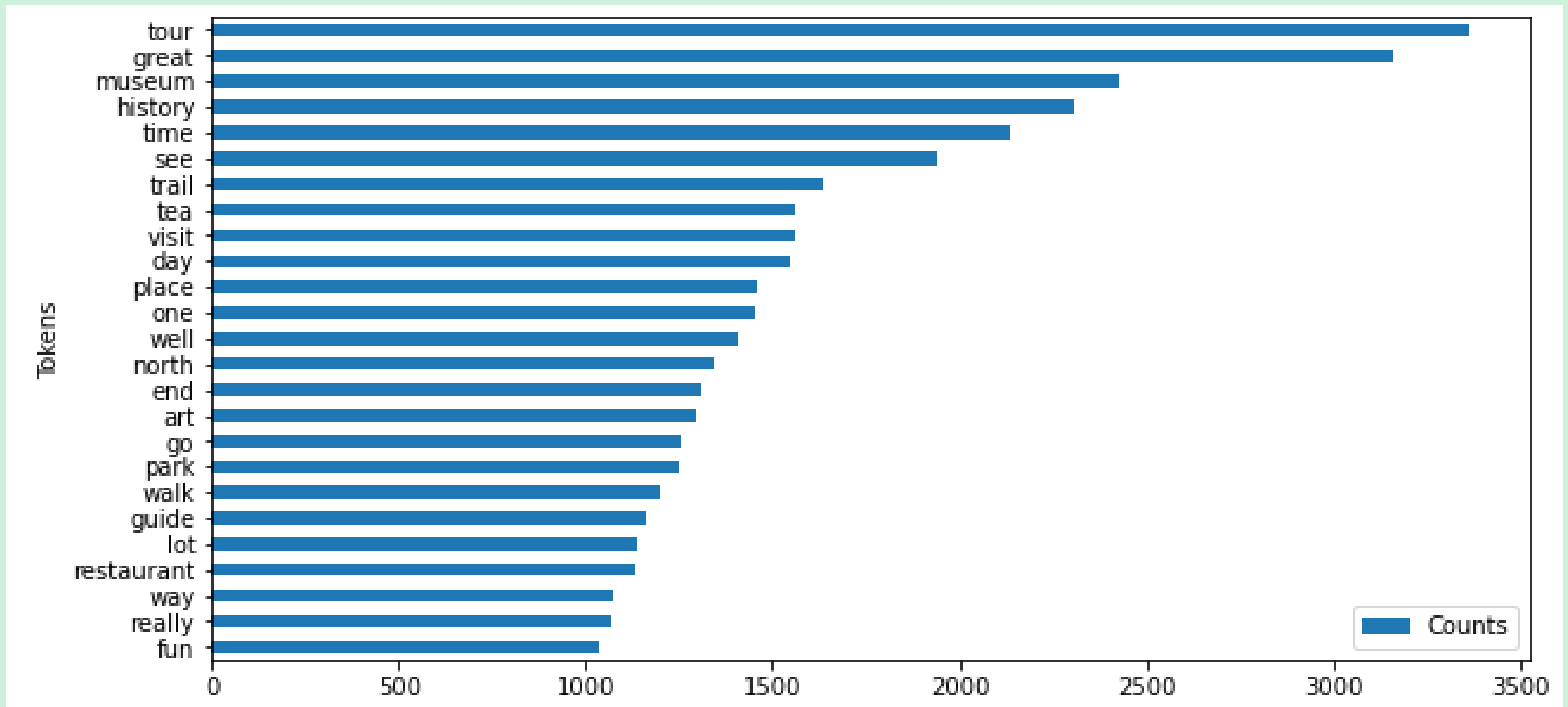
- Reviews scraped on 12/12/2020
- Scrapy to grab simple info
- Selenium to grab text
- Scraped across 4 cities (BOS, ORD, LAX, JFK)
 - Scraped top 5 Attractions per city
 - Collected ~2,000 reviews per city
 - Unless there were fewer reviews
- Initial Dataframe: 38,294 rows x 12 columns
 - Expanded to 16 columns to accommodate multiple location formats

BACKGROUND

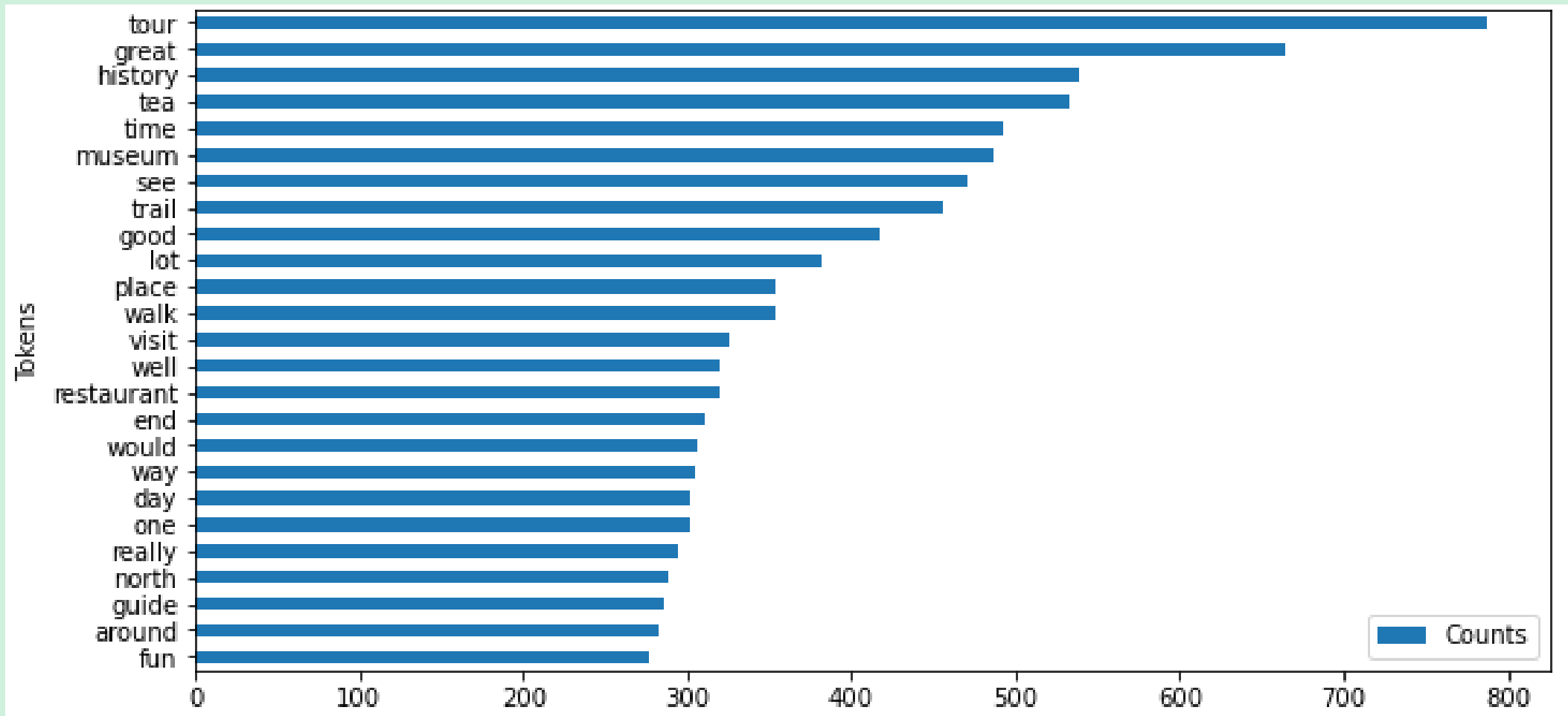
- Analyzed Reviews by 5 stars, 4 stars, 'Poor'
 - Verbiage and sentiment analysis based on factors
- How are reviews influenced?
 - Which factors seem to affect ratings?
 - Is there a potential to market to certain locations?
 - How do the reviews influence others?
- Will primarily focus on Boston as the same analysis can be repeated across all cities



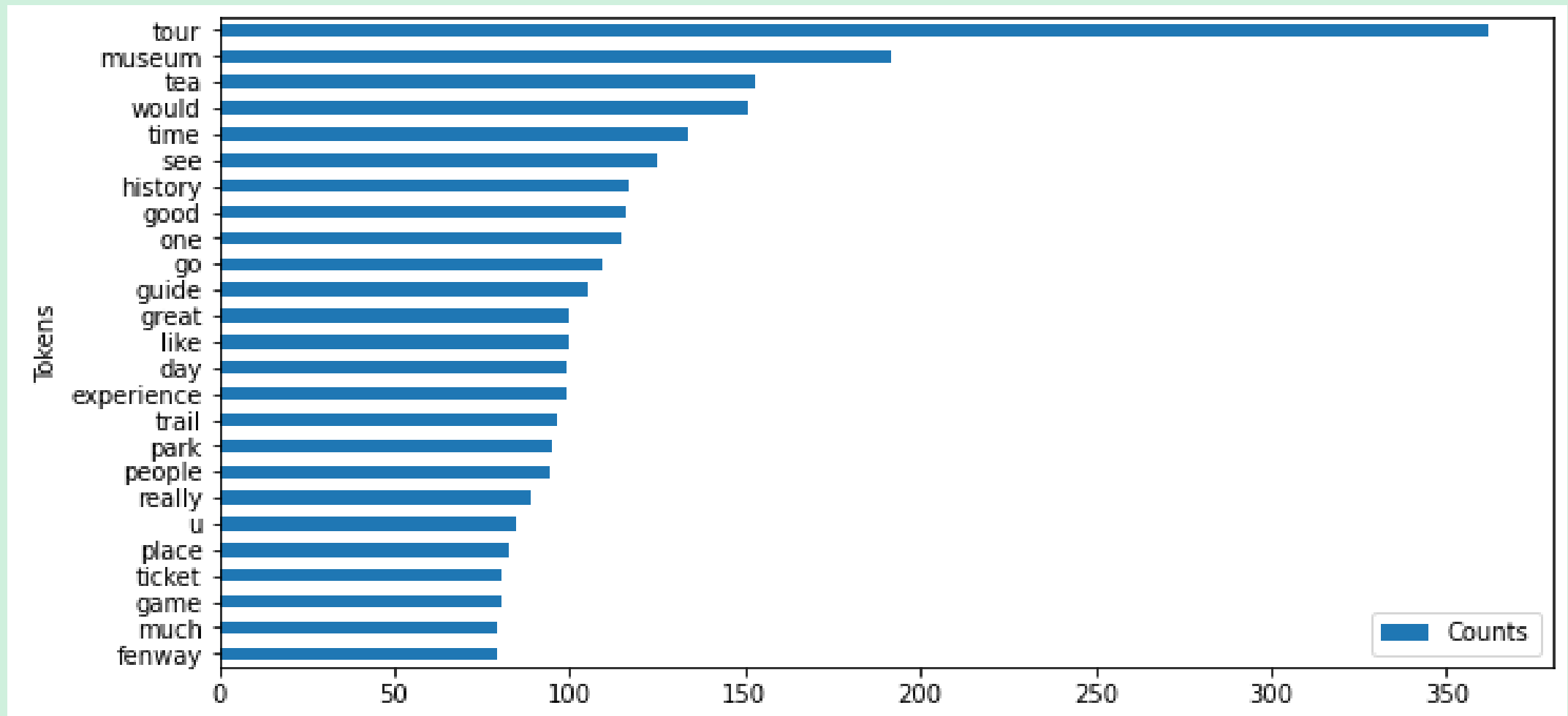
COMPARISON OF CITIES



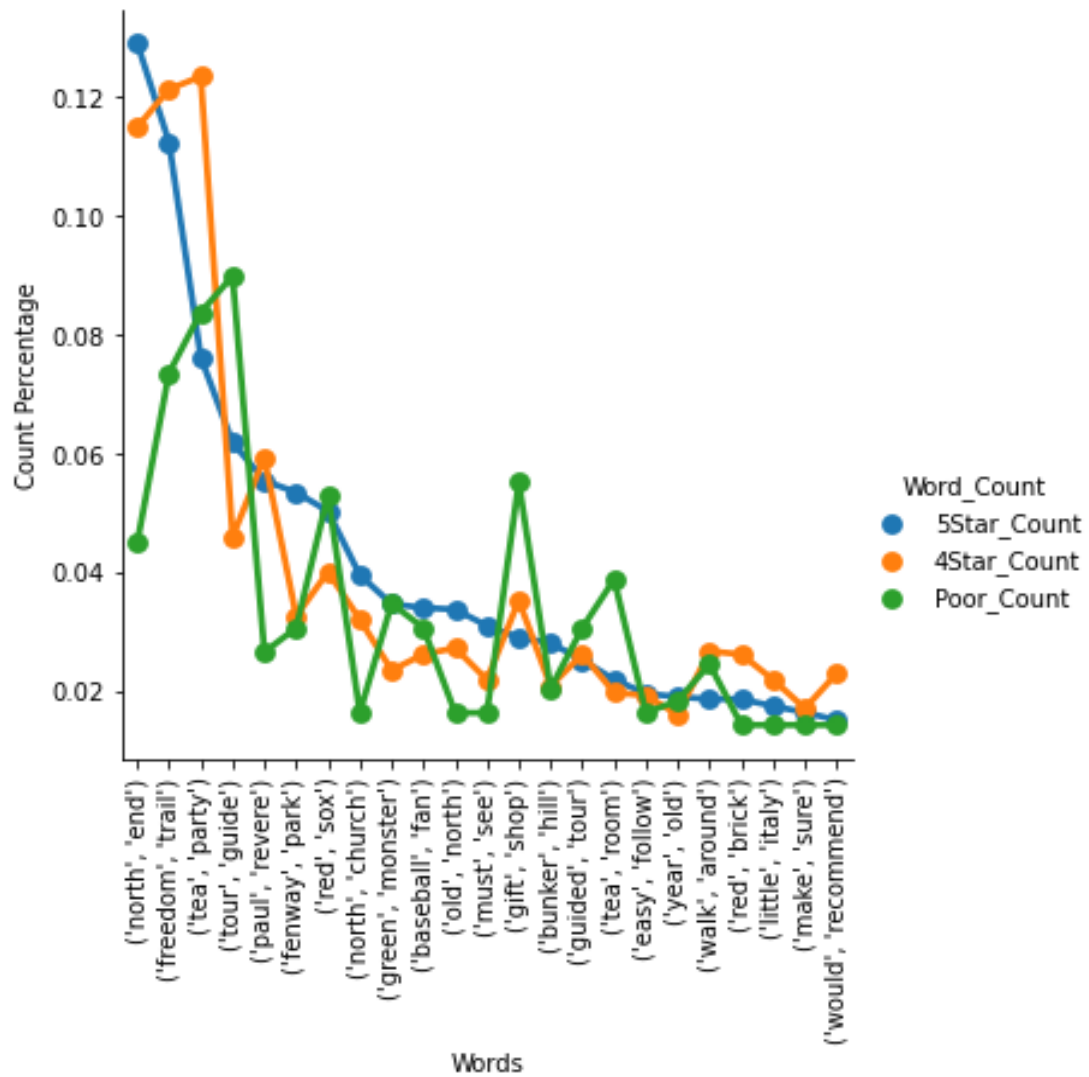
SINGLE WORDS ANALYSIS FOR 5-STAR REVIEWS



SINGLE WORDS ANALYSIS FOR 4-STAR REVIEWS



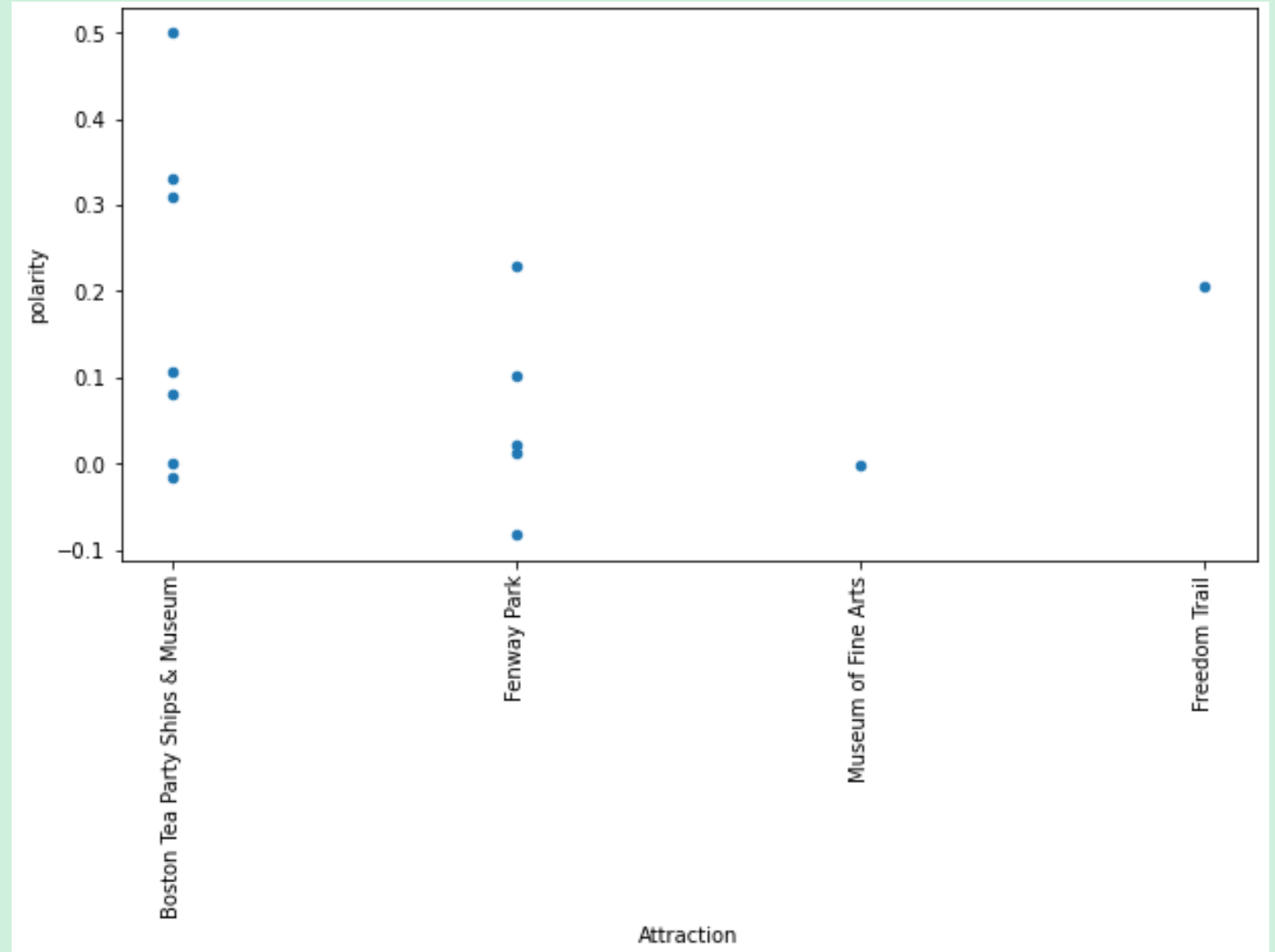
SINGLE WORDS ANALYSIS FOR POOR REVIEWS

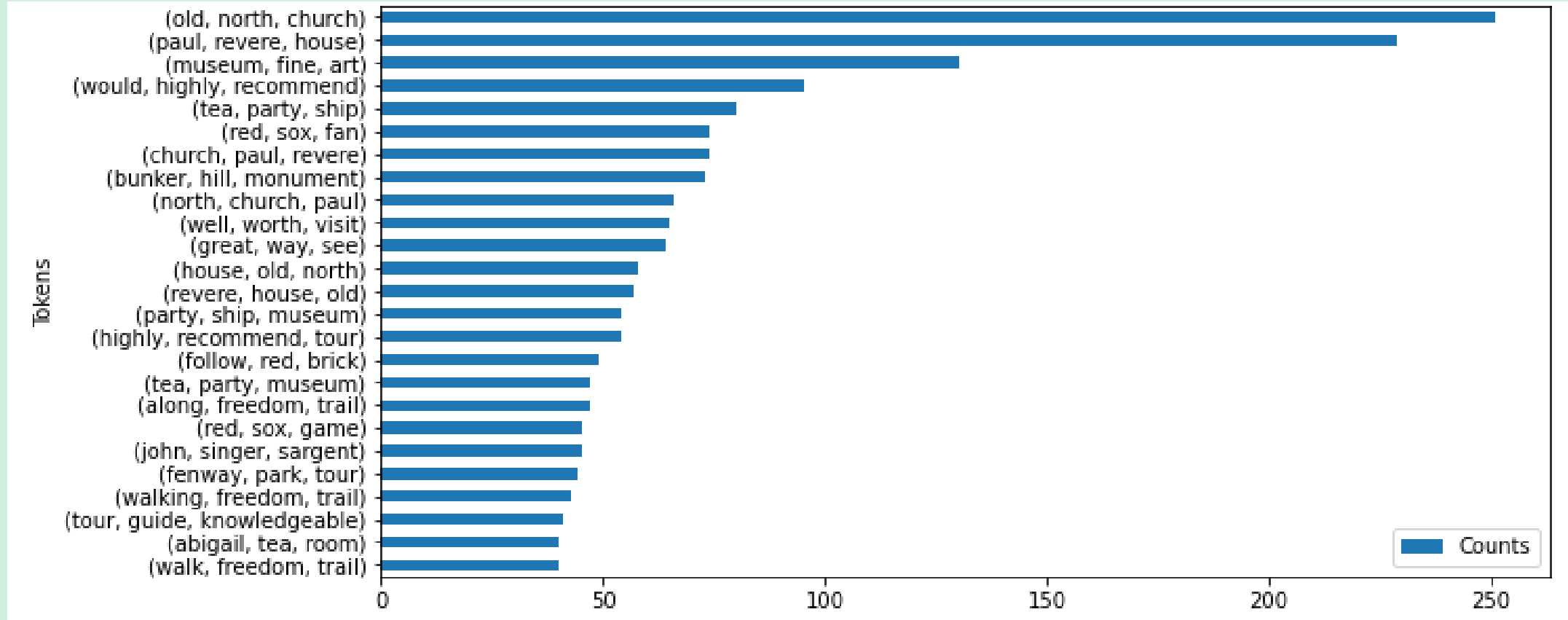


- Frequency of 2-combination words
- Gift Shop
 - Mainly 1 attraction

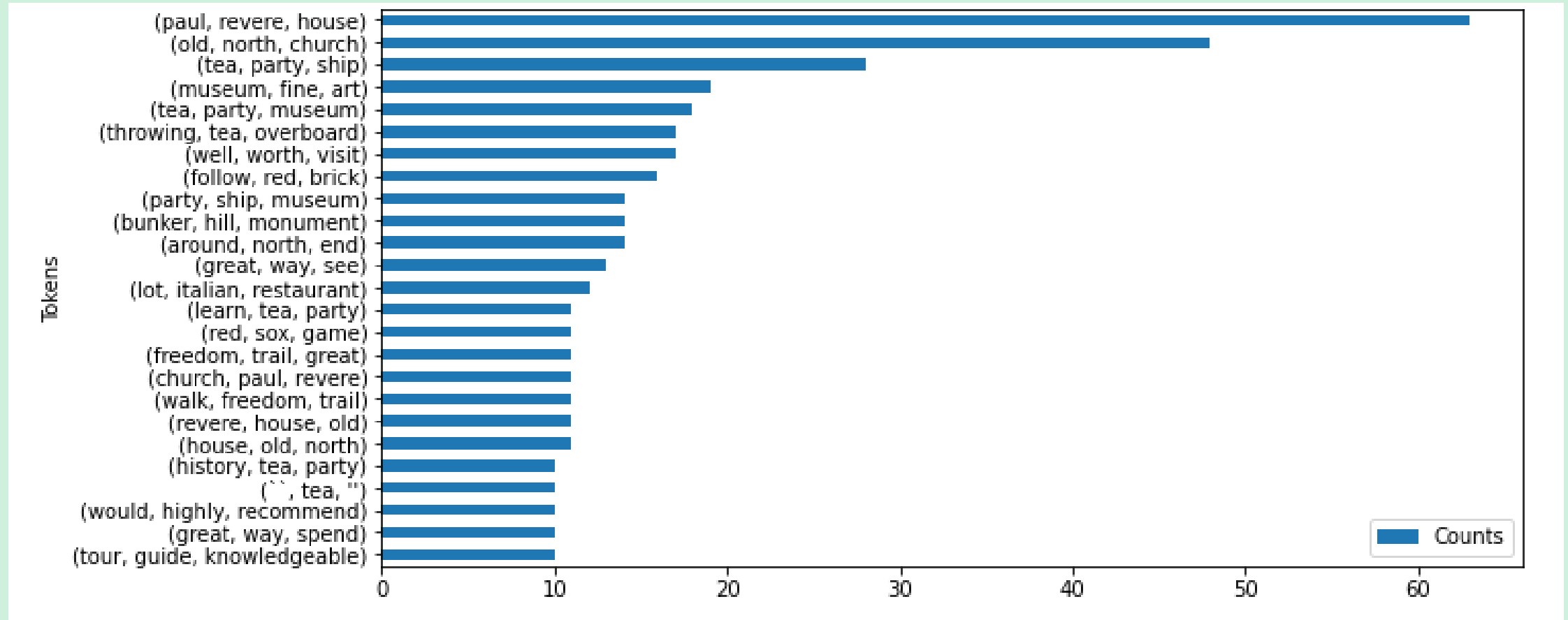
BIGRAM ANALYSIS

‘TIME’, ‘MONEY’ ANALYSIS FOR POOR REVIEWS

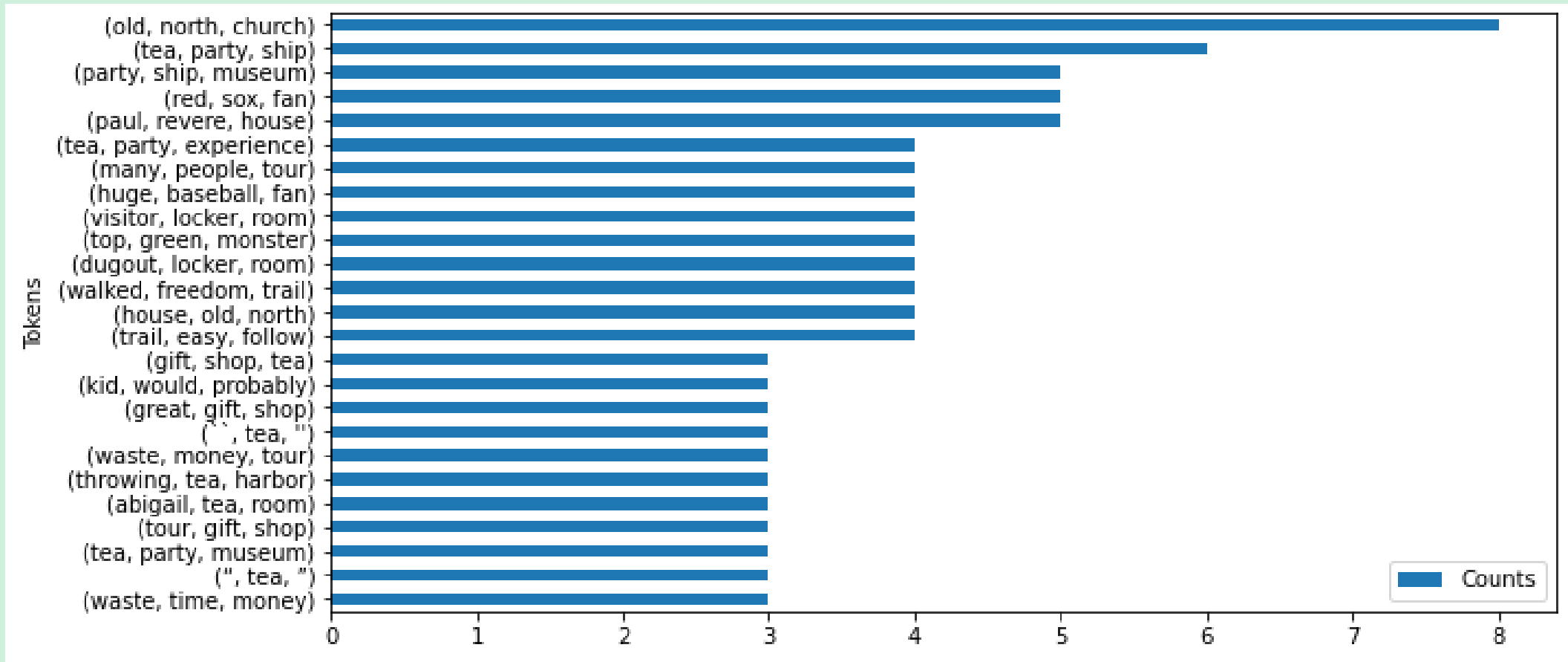




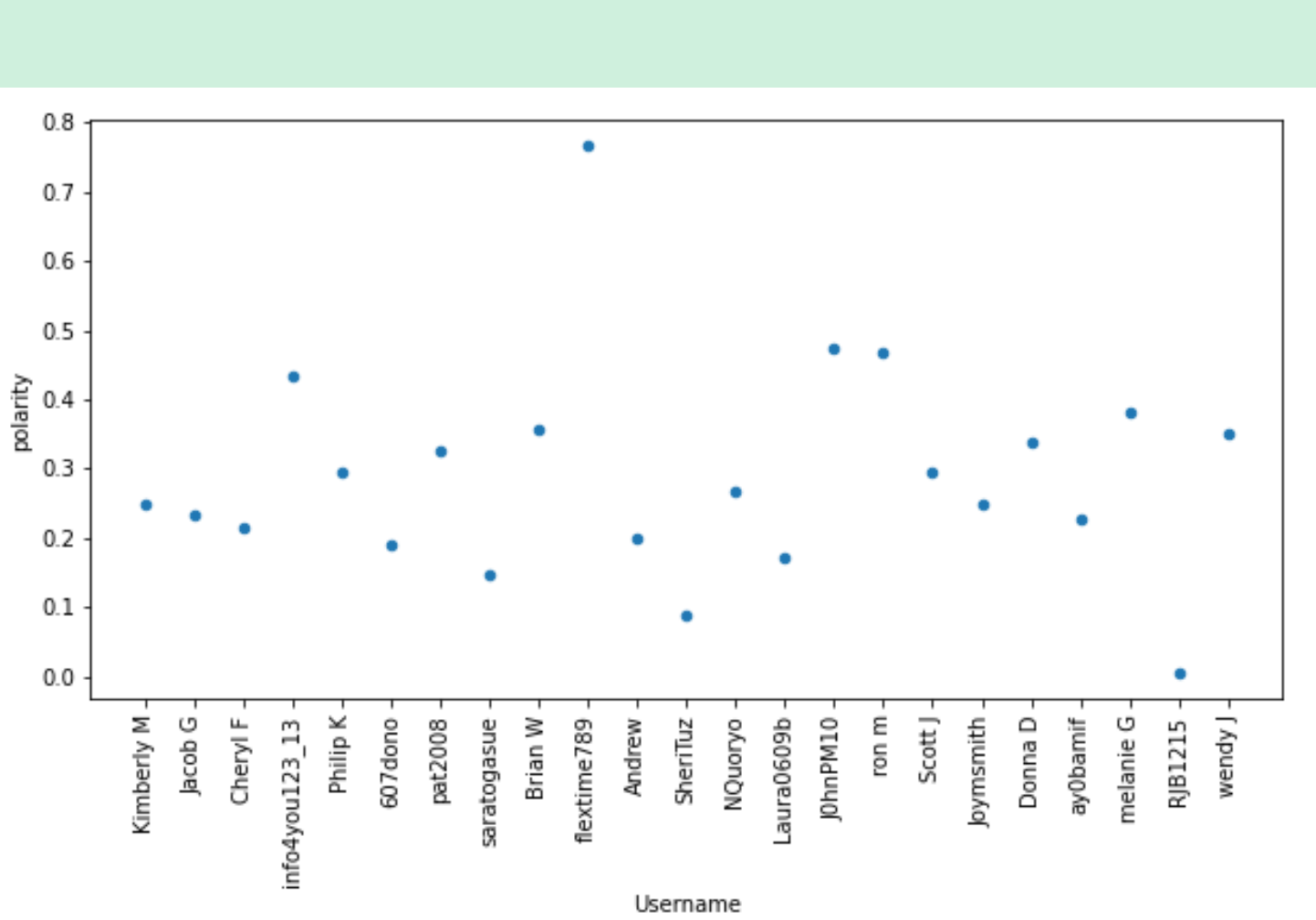
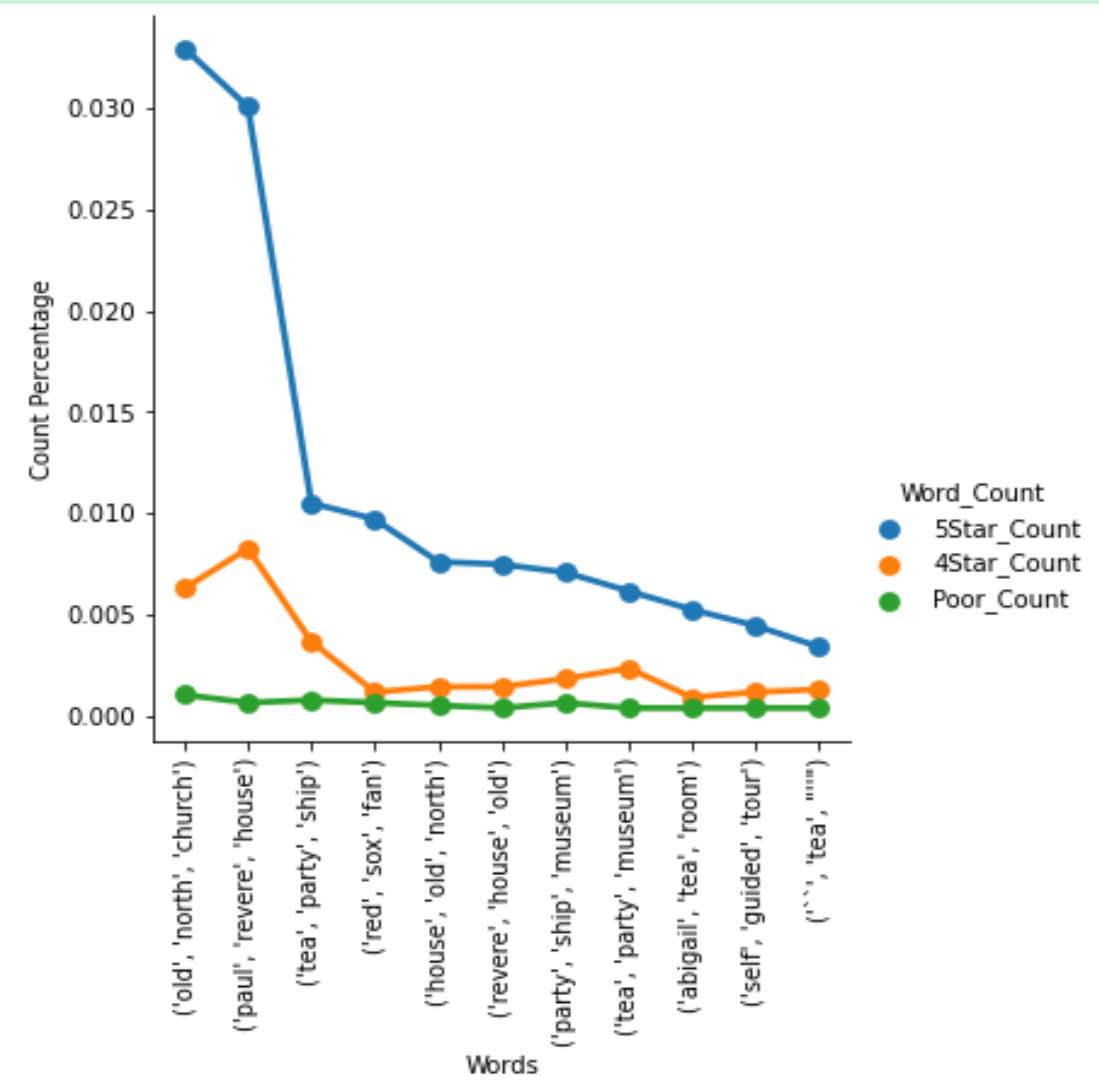
TRIGRAMS FOR 5-STAR REVIEWS



TRIGRAMS FOR 4-STAR REVIEWS

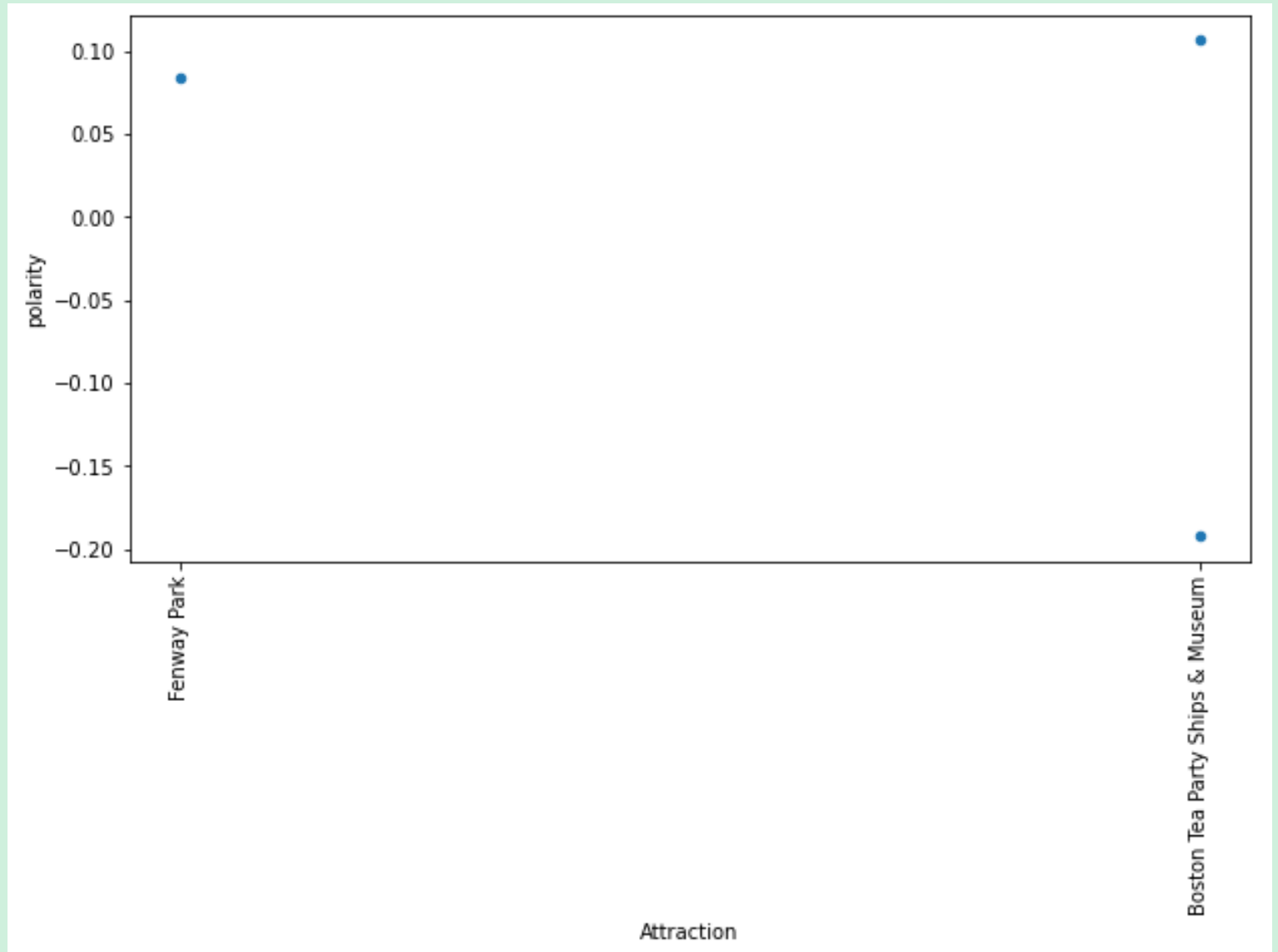


TRIGRAMS FOR POOR REVIEWS



TRIGRAMS

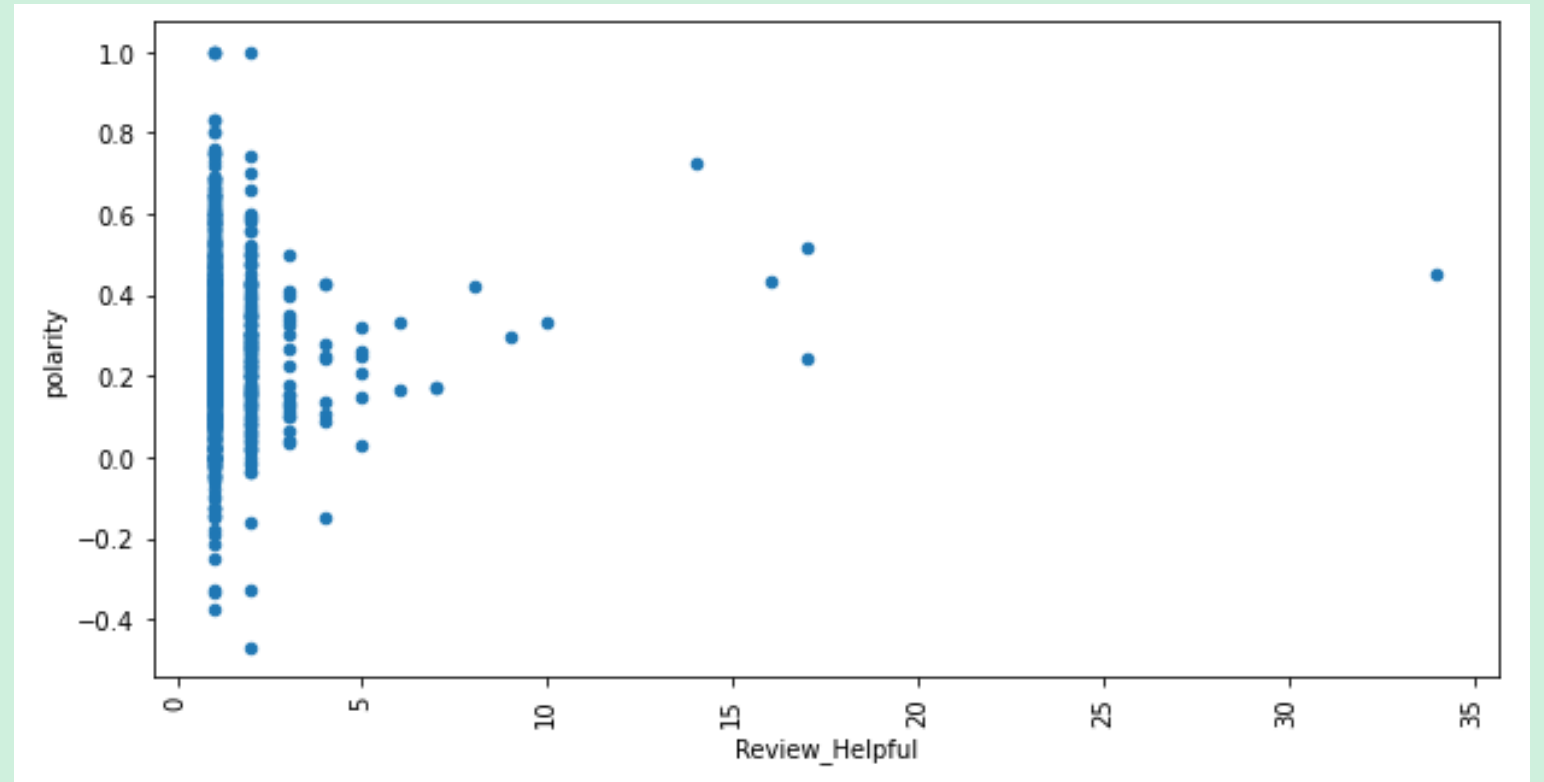
‘WASTE’, ‘MONEY’, ‘TIME’ ANALYSIS



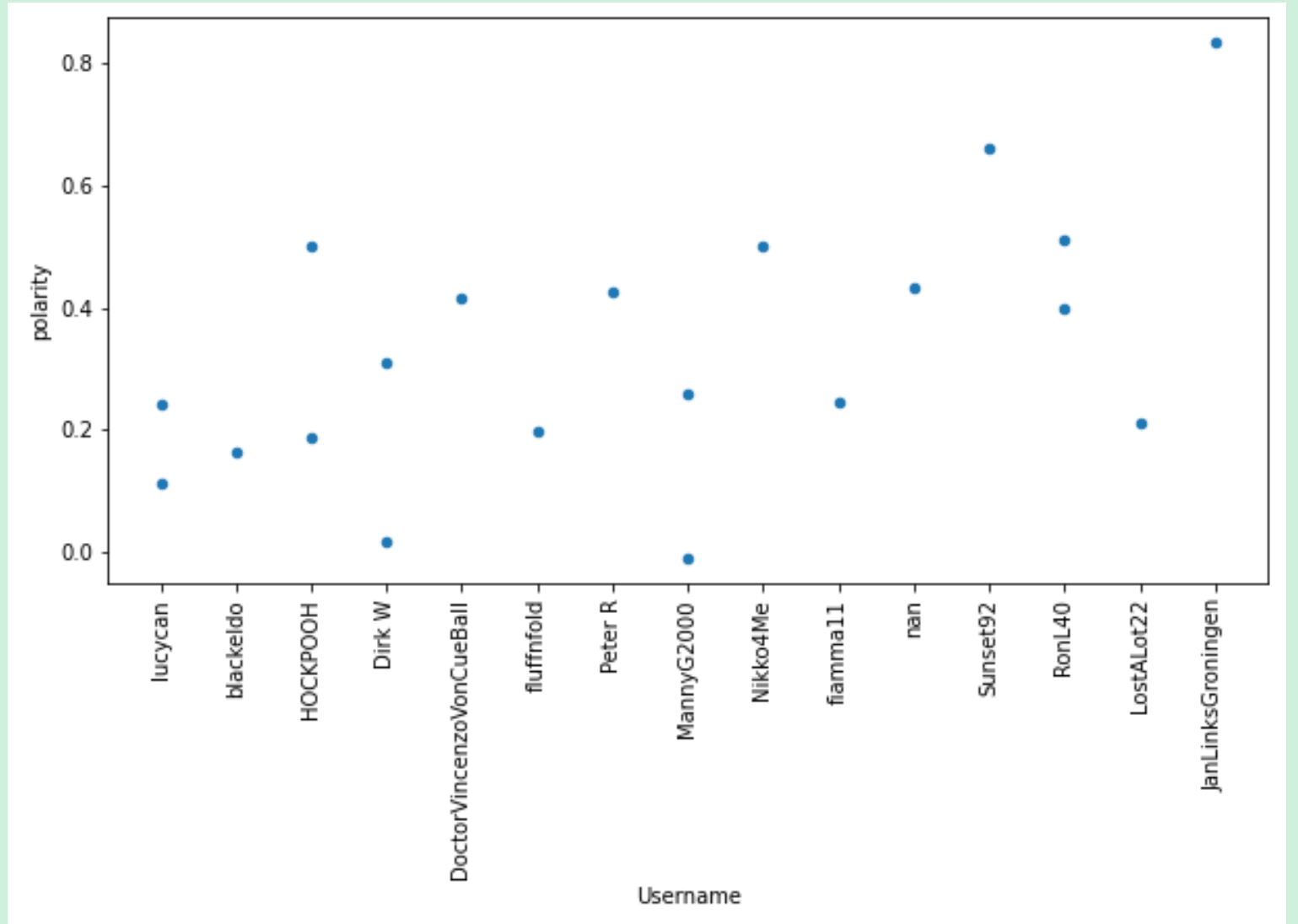
HELPFUL

- 5 Stars
 - 10.64% 5 Star Reviews
- 4 Stars
 - 10.03% 4 Star Reviews
- Poor
 - 18.37% Poor Reviews

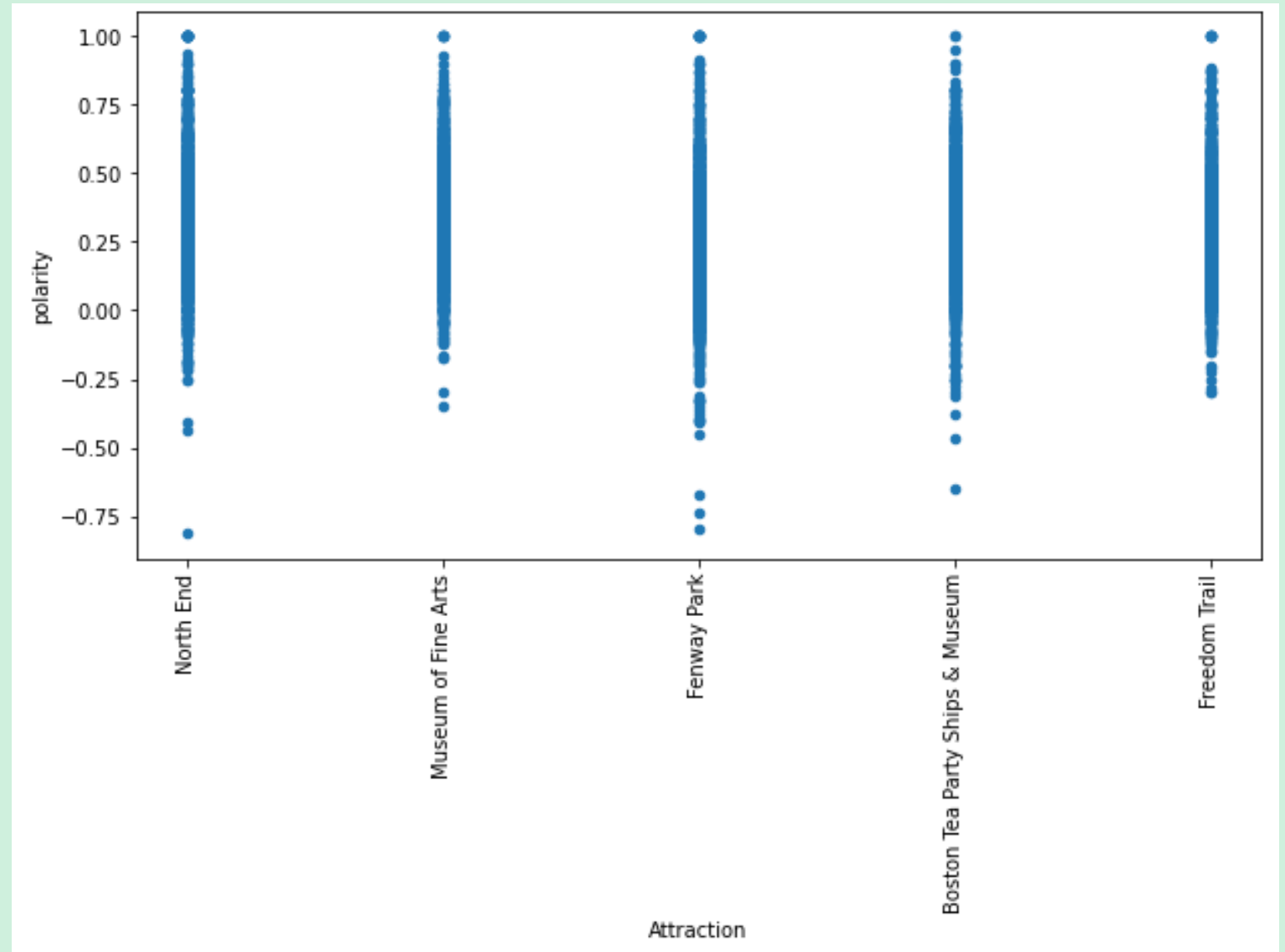
HELPFUL REVIEWS

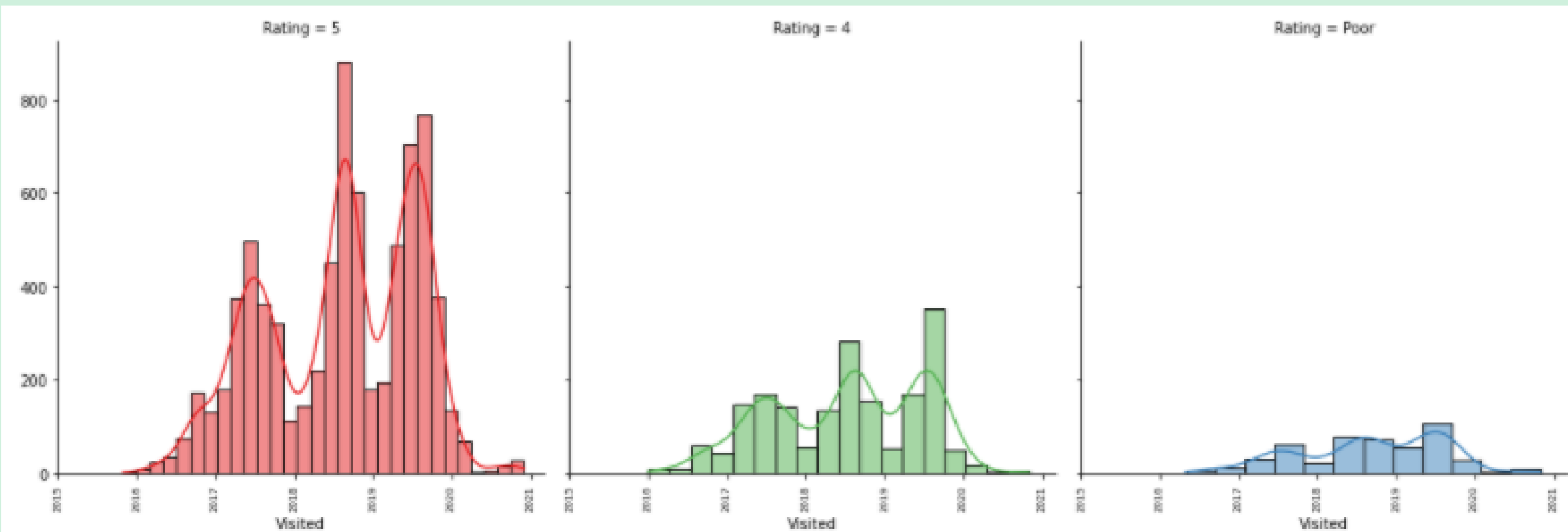


HELPFUL USERS



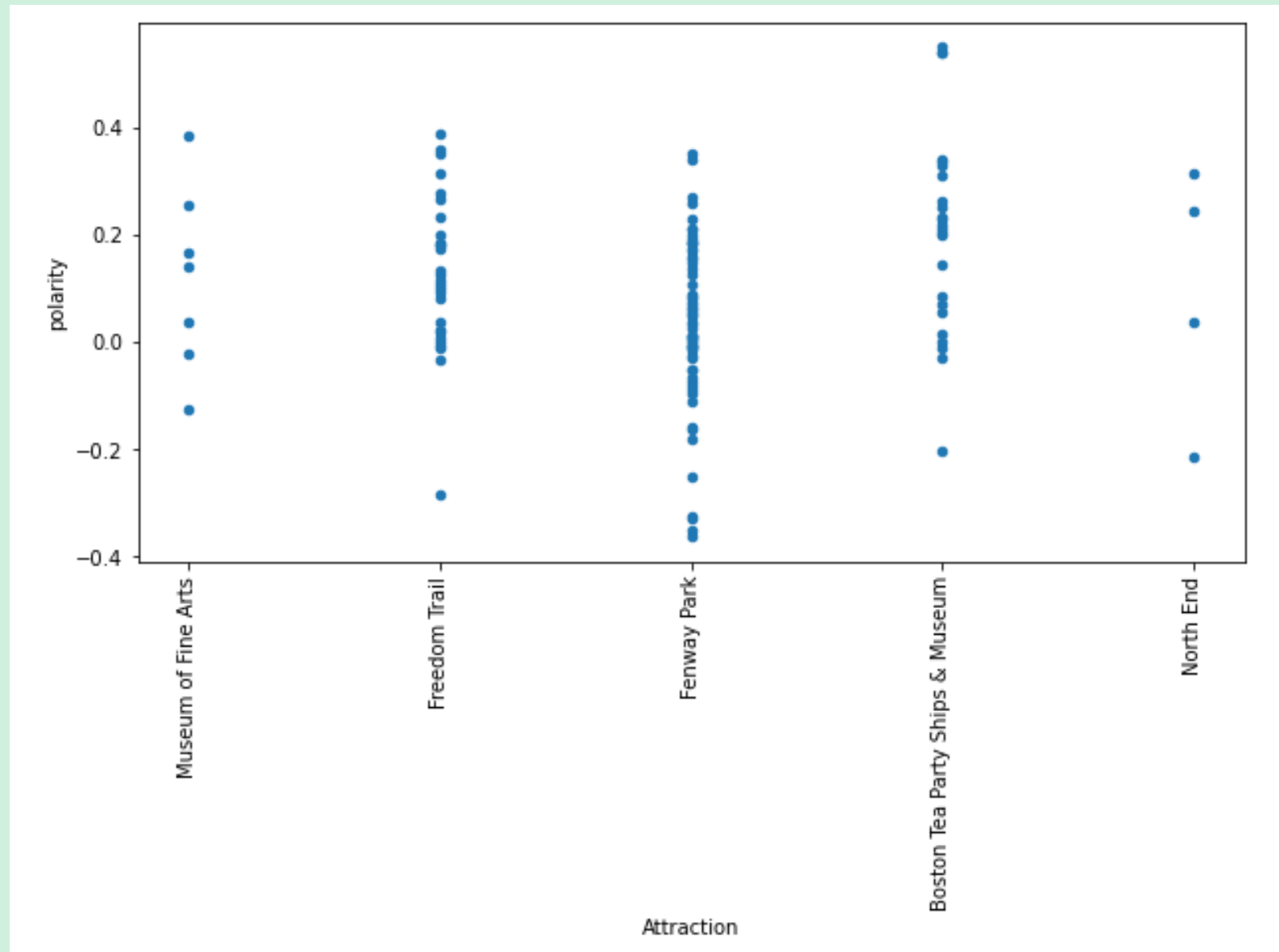
POLARITY OF REVIEWS BASED ON ATTRACTION



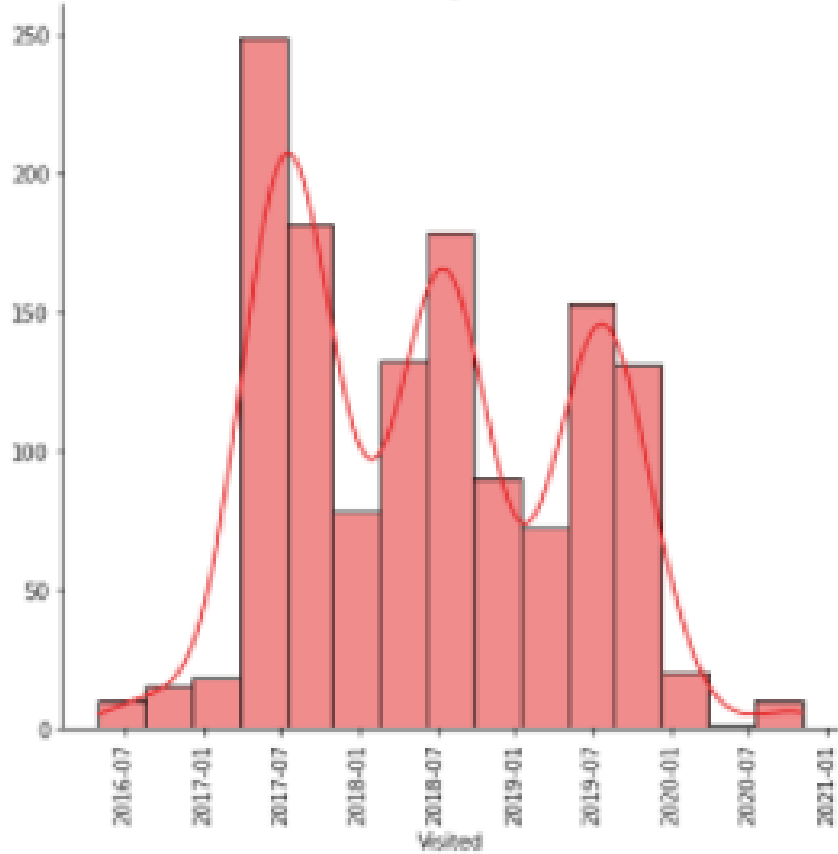


**RATINGS BASED ON
MONTH AND YEAR**

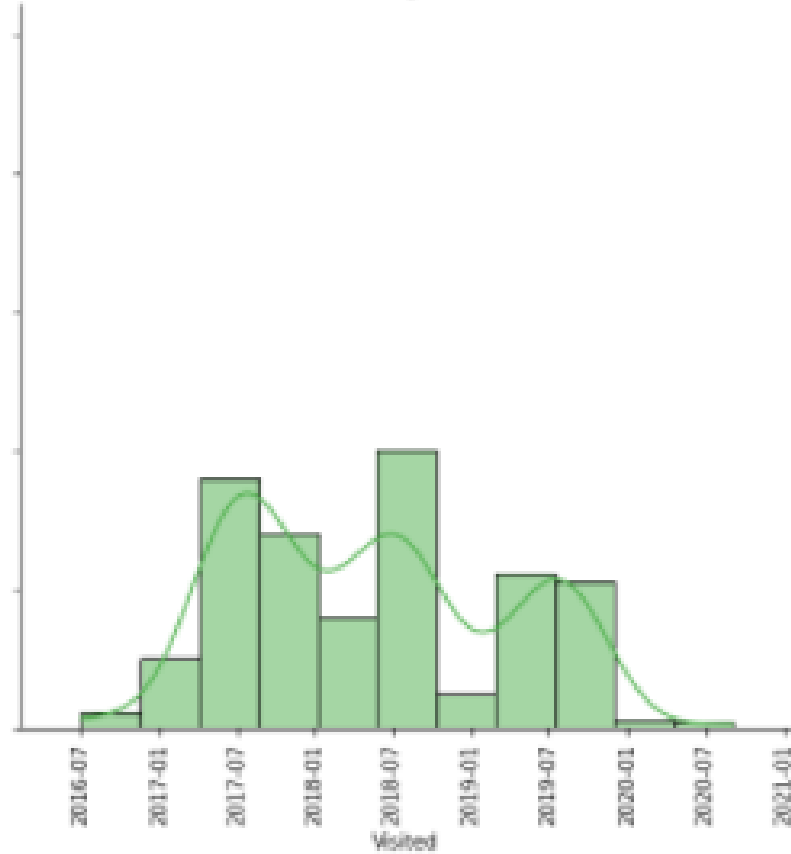
WHY THE POOR REVIEWS MID-2019?



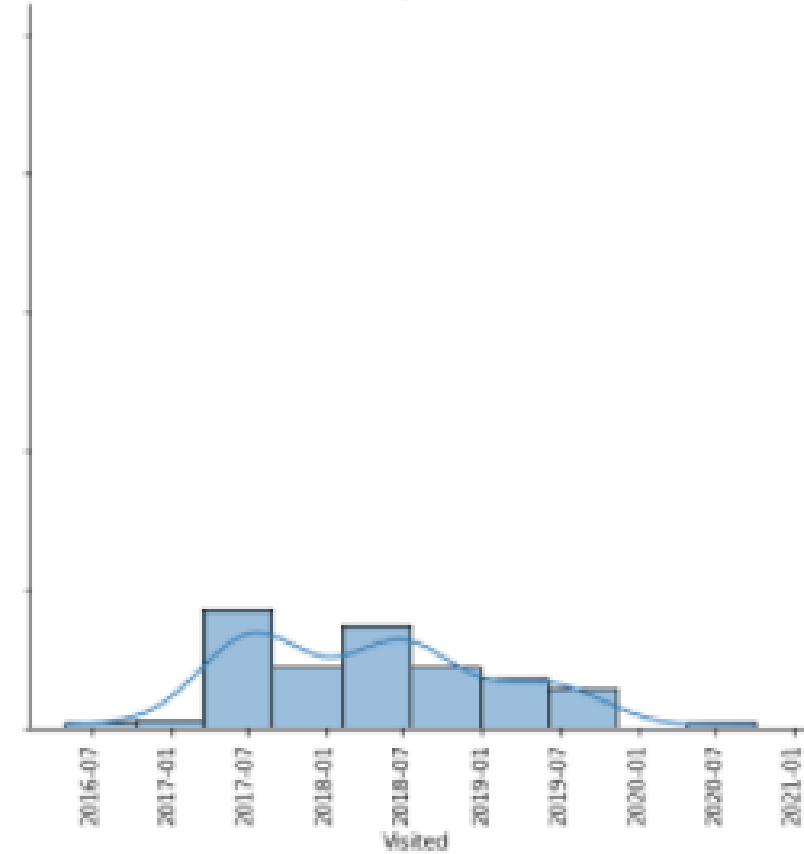
Rating = 5



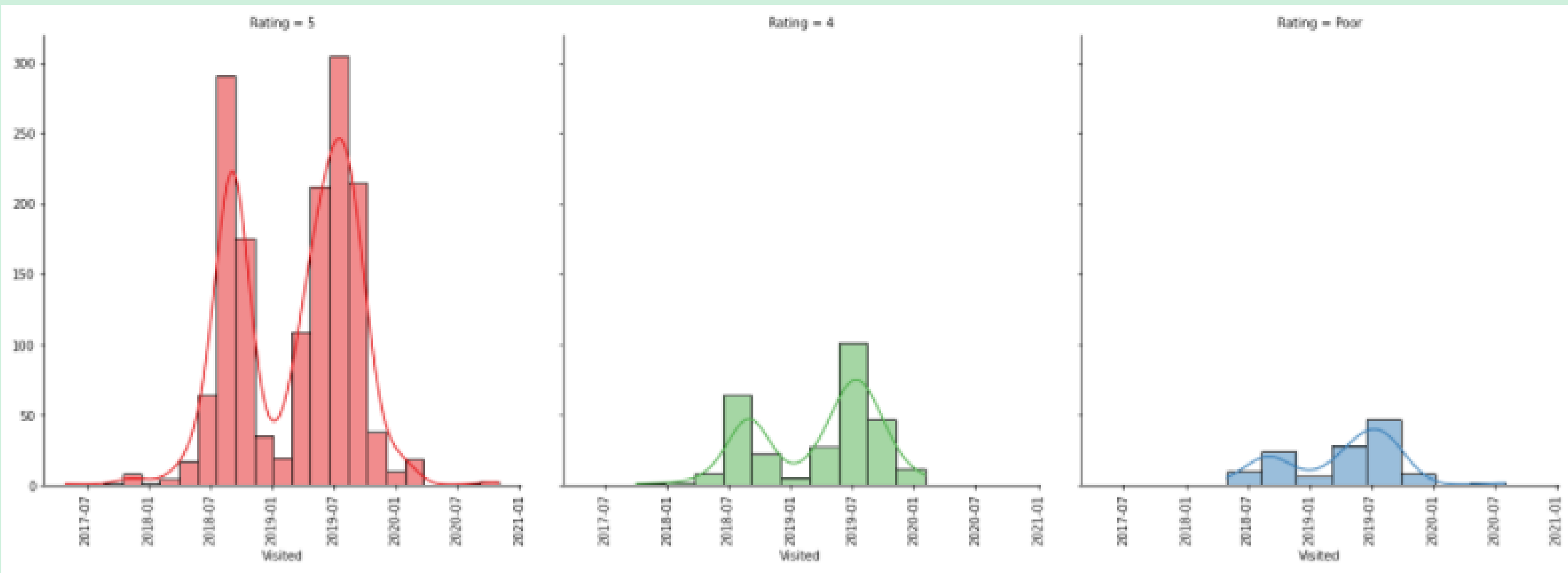
Rating = 4



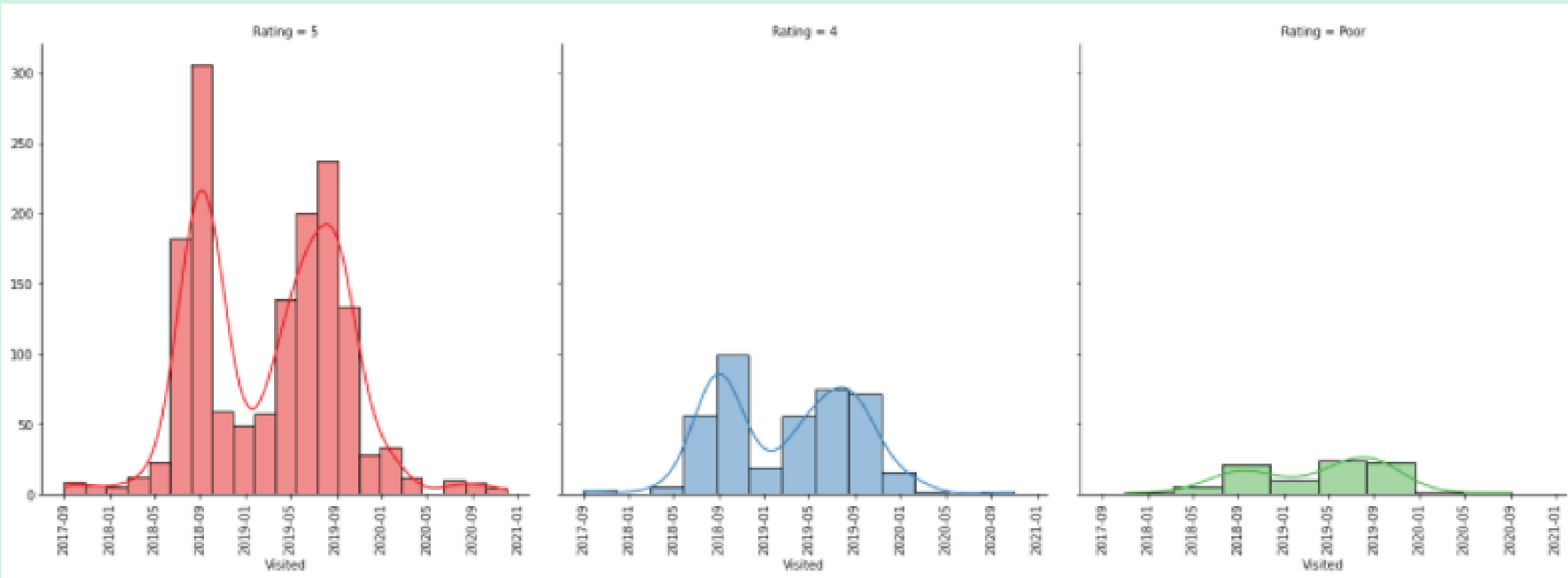
Rating = Poor



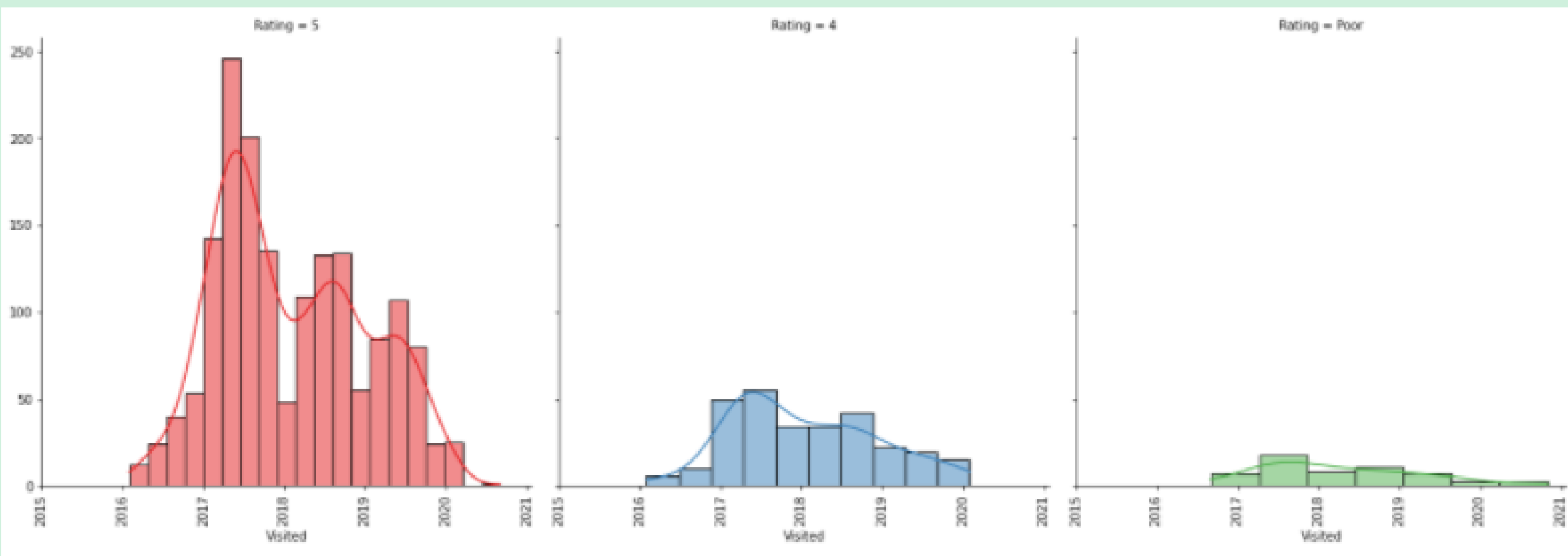
RATINGS BASED ON
MONTH AND YEAR
(TEA PARTY)



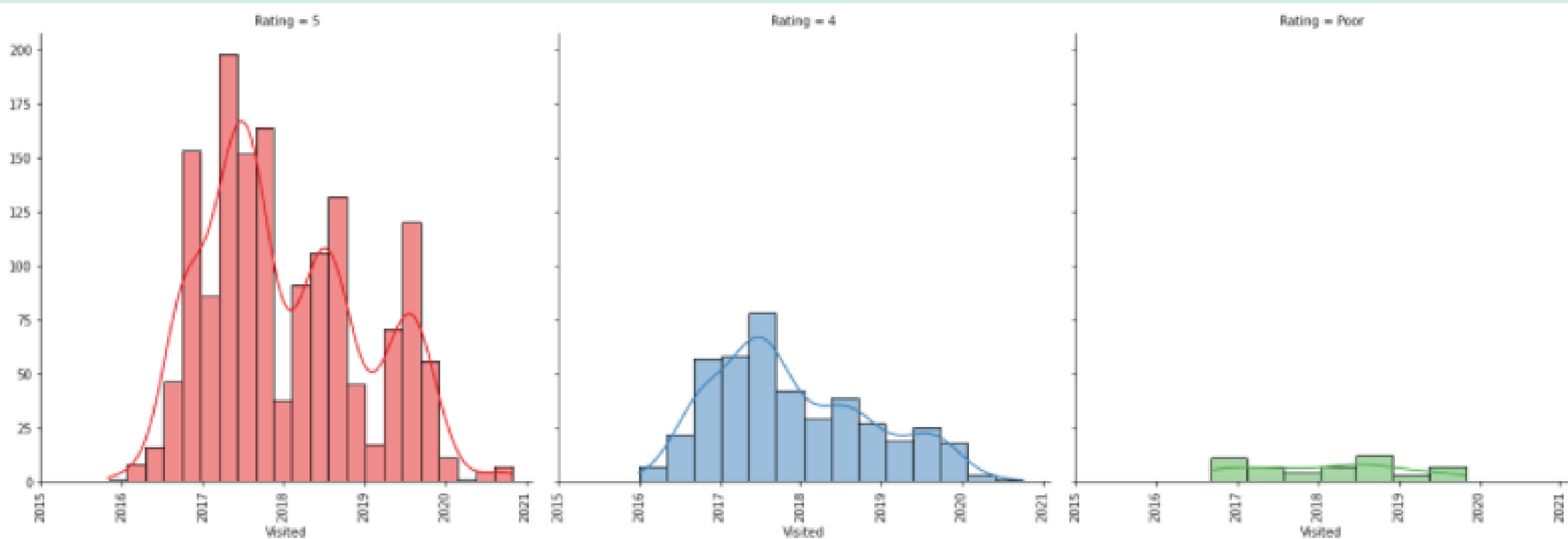
**RATINGS BASED ON
MONTH AND YEAR
(FENWAY)**



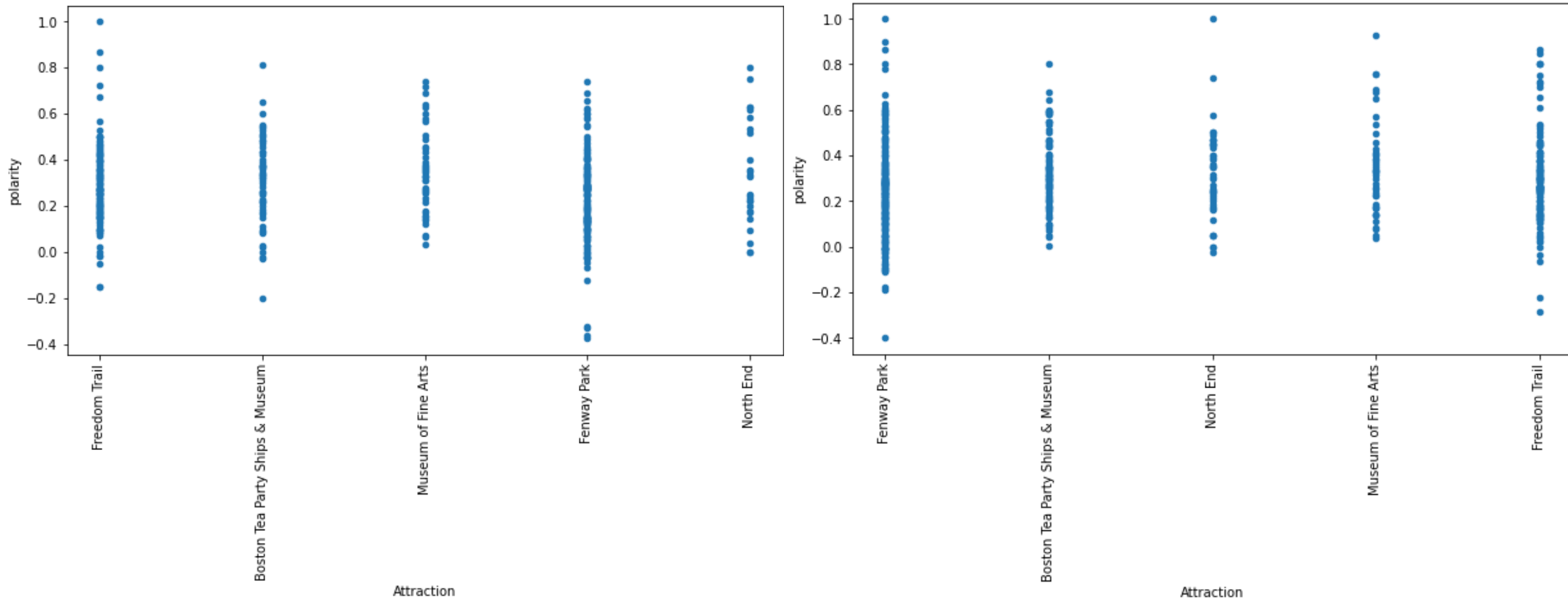
**RATINGS BASED ON
MONTH AND YEAR
(FREEDOM TRAIL)**



**RATINGS BASED ON
MONTH AND YEAR
(MFA)**

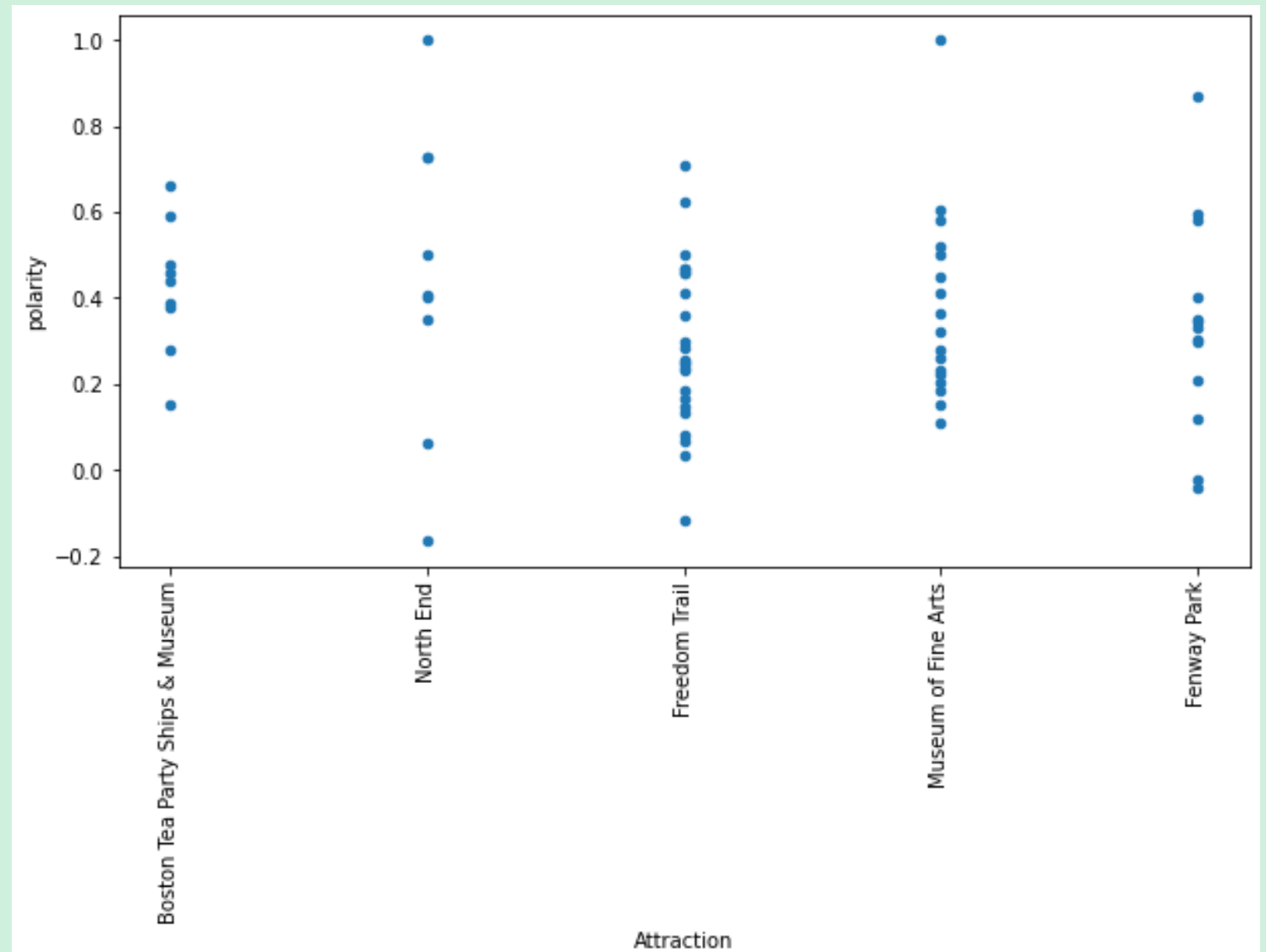


**RATINGS BASED ON
MONTH AND YEAR
(NORTH END)**

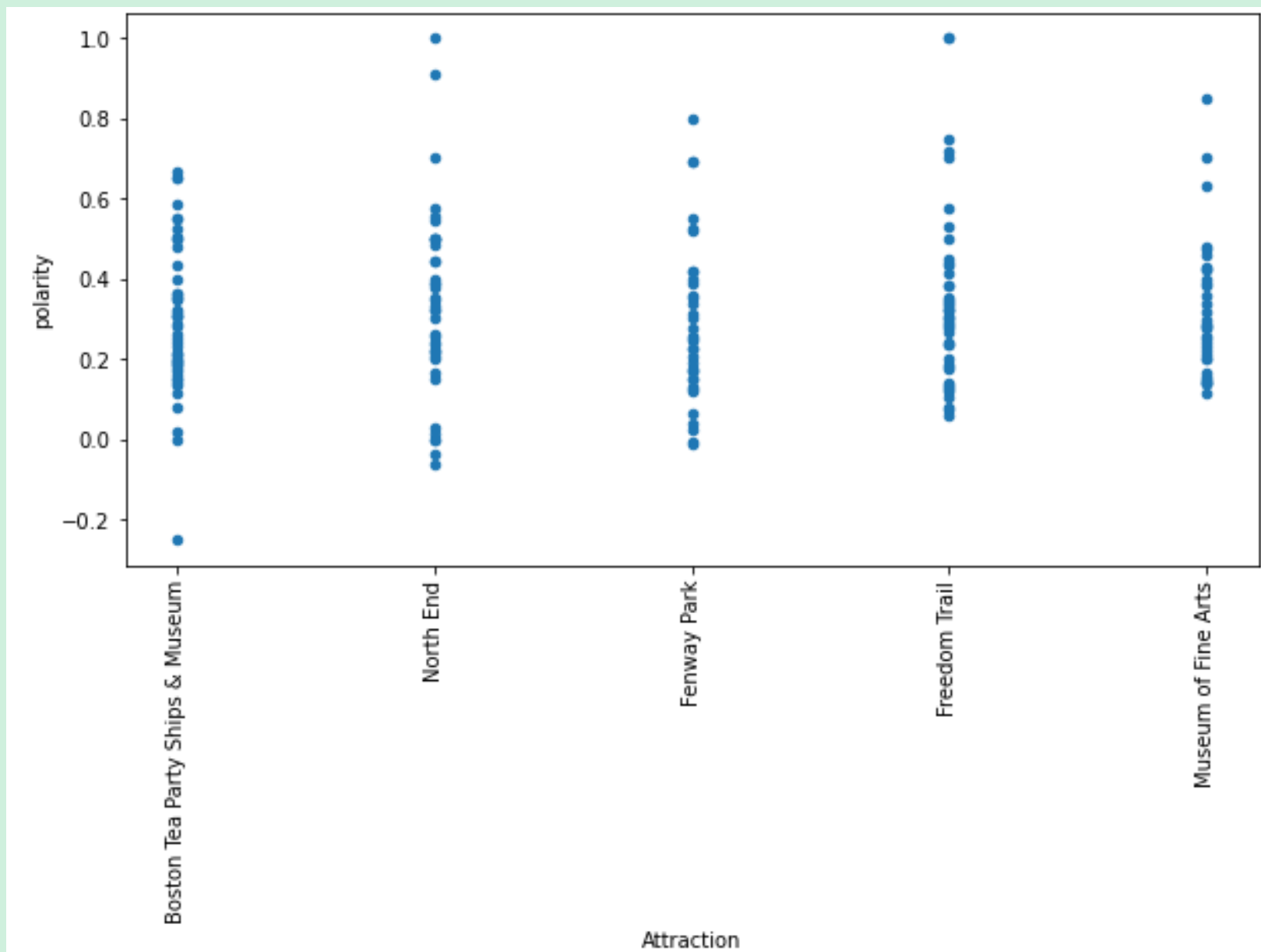


POLARITY COMPARISON
BETWEEN JUNE/JULY

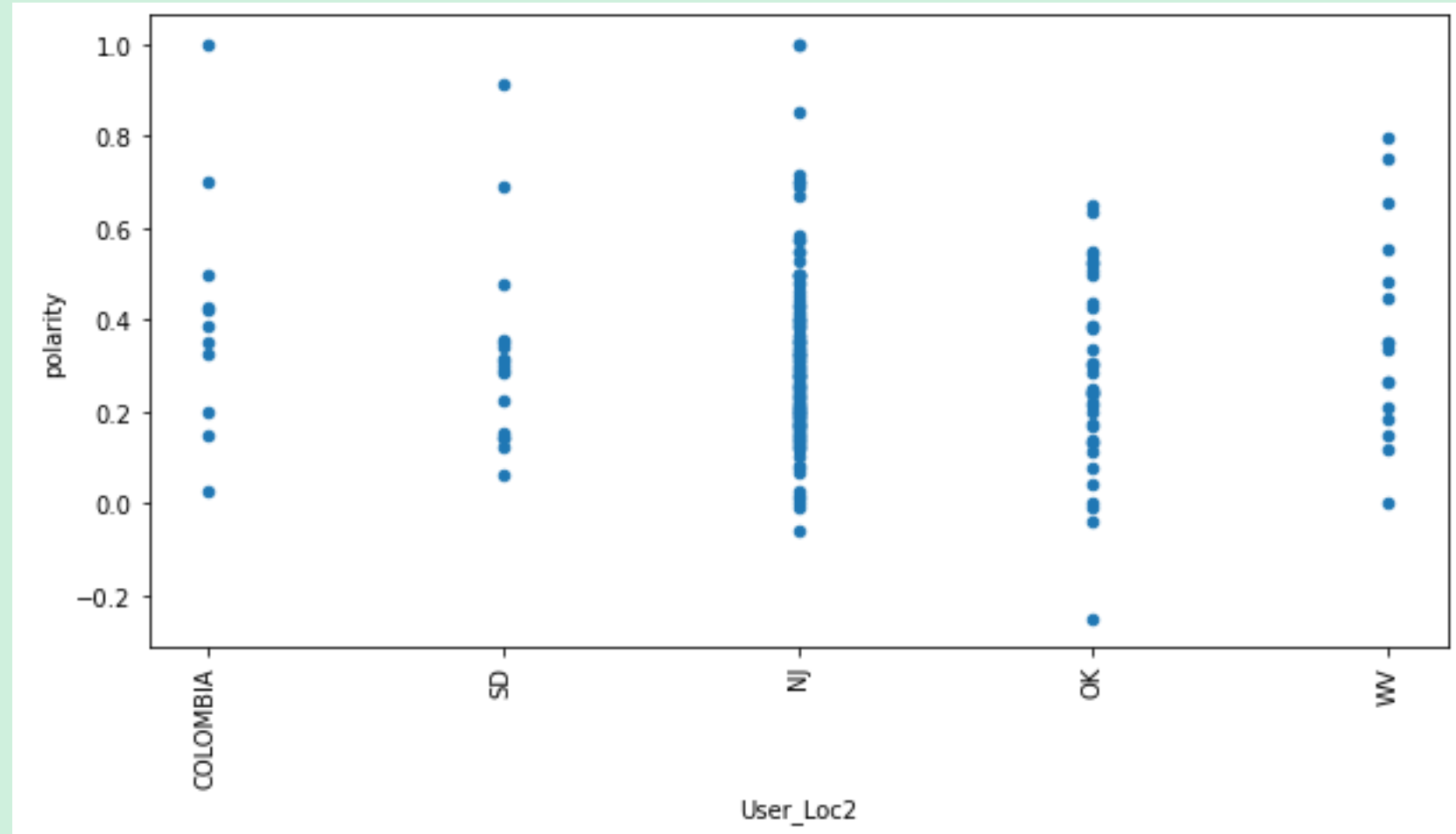
POLARITY OF REVIEWS IN JANUARY 2020



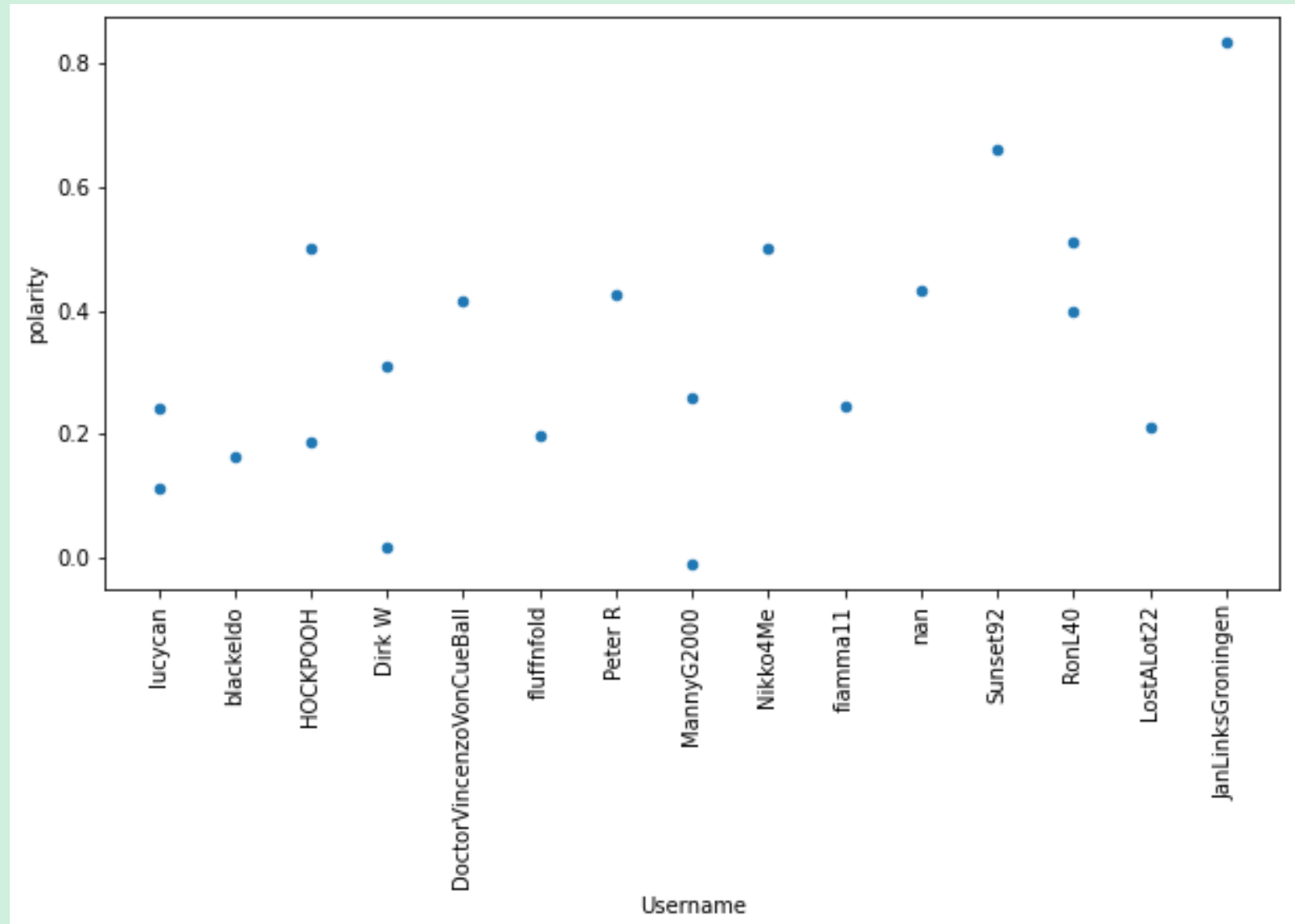
POLARITY
OF REVIEWS
BASED ON
LOCATION:
WV, OK, SD,
COLOMBIA,
NJ



POLARITY
OF REVIEWS
BASED ON
LOCATION:
WV, OK, SD,
COLOMBIA,
NJ



POLARITY MOST HELPFUL USERS



- Improve Boston Tea Party attraction with better acting and more informative museum to complement the good gift shop
- Fenway to provide better customer service