Airbnb data analysis

Identifying most common amenities that give Airbnb hosts a competitive advantage over other hosts in Gauteng & Cape Town, South Africa

~Henrick Aduda

The Data Incubator

Motivation

- Airbnb hosts took a massive hit in 2020
- Airbnb revenues decreased by 48.6% ($35 billion in Jan to $18 billion in May)
- Opportunity: hosts in rural areas of the U.S. earned over USD 200 million in June 2020
- Motivation : find approaches/means for airbnb hosts to get more clients.
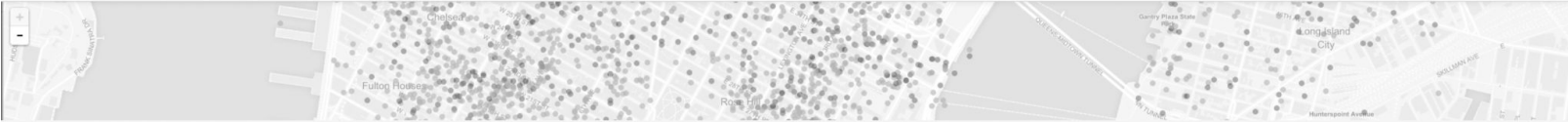
Data source description

- Original dataset came from Inside Airbnb web database
- Inside Airbnb is an independent, non-commercial set of tools and data that allows you to explore how Airbnb is really being used in cities around the world.

http://insideairbnb.com/get-the-data.html

Analysis

Deliverables

1. Identify the most common amenities affecting review ratings - correlation analysis
2. Perform a sentiment analysis - monkeylearn web tool - aim for ML model accuracy >80%
3. Identify the most common amenities affecting review comments - Yet to figure out plan
- **Opportunity:** Scale it up to look at world wide data. Analysis per continent

Airbnb sentiment analysis reviews
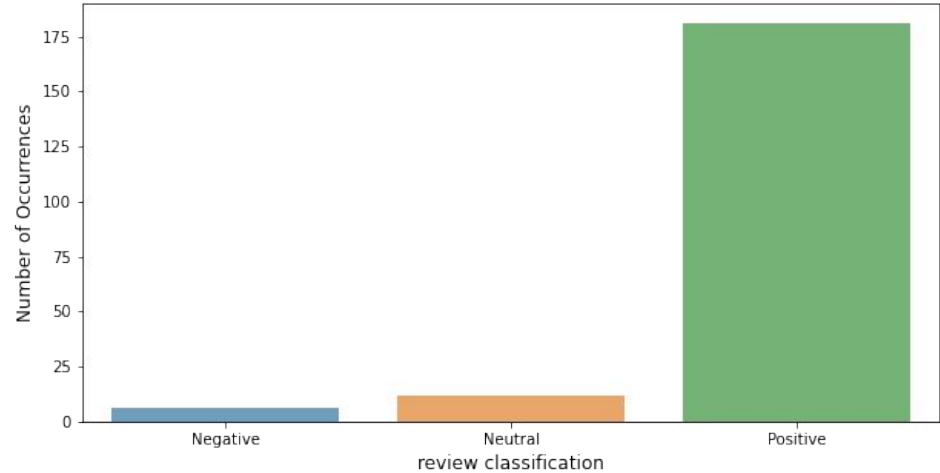
200 out of 305085 review entries

Confidence mean scores
Positive : 0.97
Negative: 0.83
Neutral: 0.69
All entries: 0.97