NLP for Digital Videogame Reviews

Jason Kim

Metis

Background

- Videogame Industry: over $200 billion within 2020-2025 (marketwatch)
- Amazon: 150 million prime subscribers in 2021 + other customers
- Billions of people who play videogames, worldwide
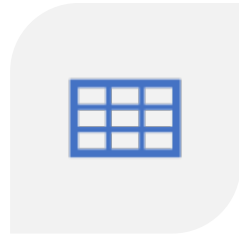
Goal

- Explore language of Amazon videogame reviews to understand the most important qualities for videogame products
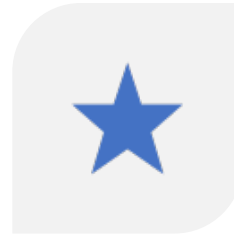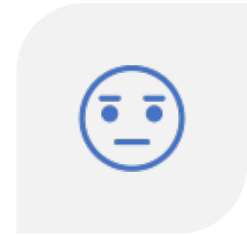
Process

DATA: 144,724 REVIEWS

PREPROCESSING: REMOVING NON-ENGLISH WORDS, STOP WORDS, ETC.

TOPIC MODELING: FINDING TOPICS BASED ON REVIEW WORDS

SENTIMENT ANALYSIS: BASED ON STAR RATING

FINDING TOPICS ASSOCIATED WITH POSITIVE/NEGATIVE SENTIMENT

Topics

Ease/Speed of Transaction

Good Price

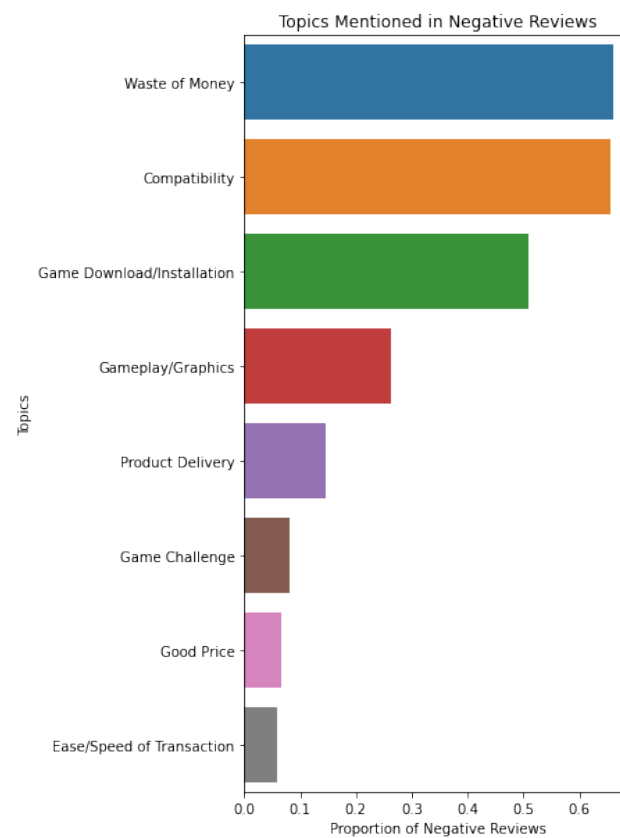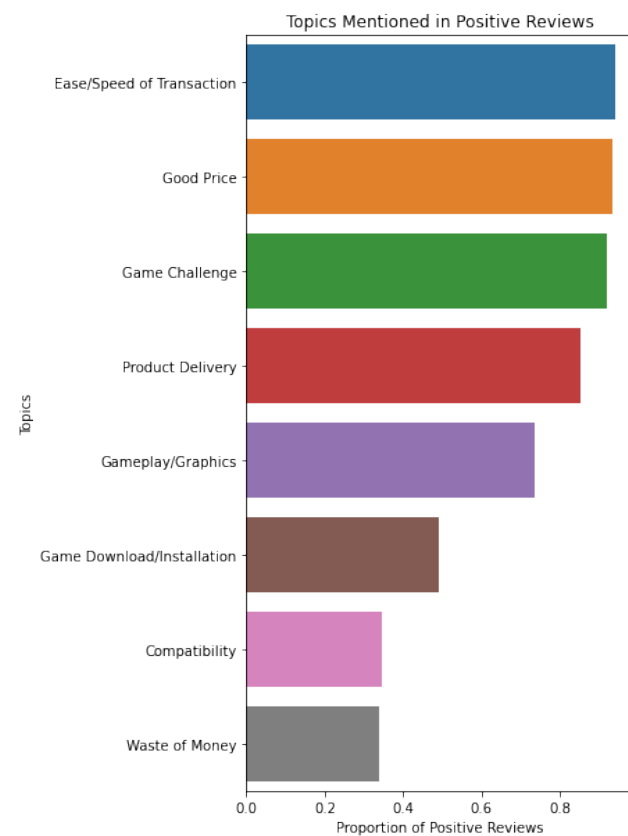Game Challenge

Product Delivery
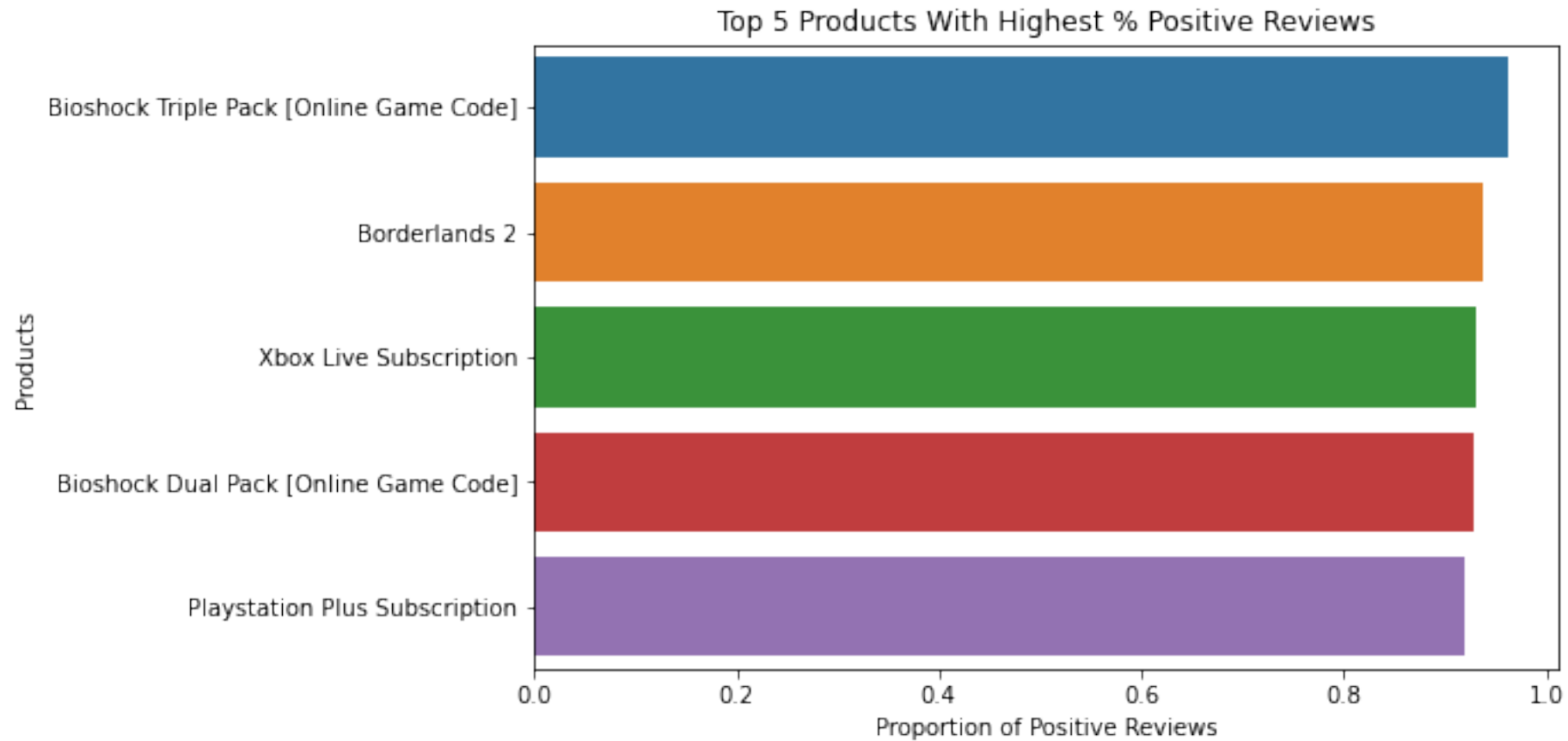
Gameplay/Graphics

Download/Installation
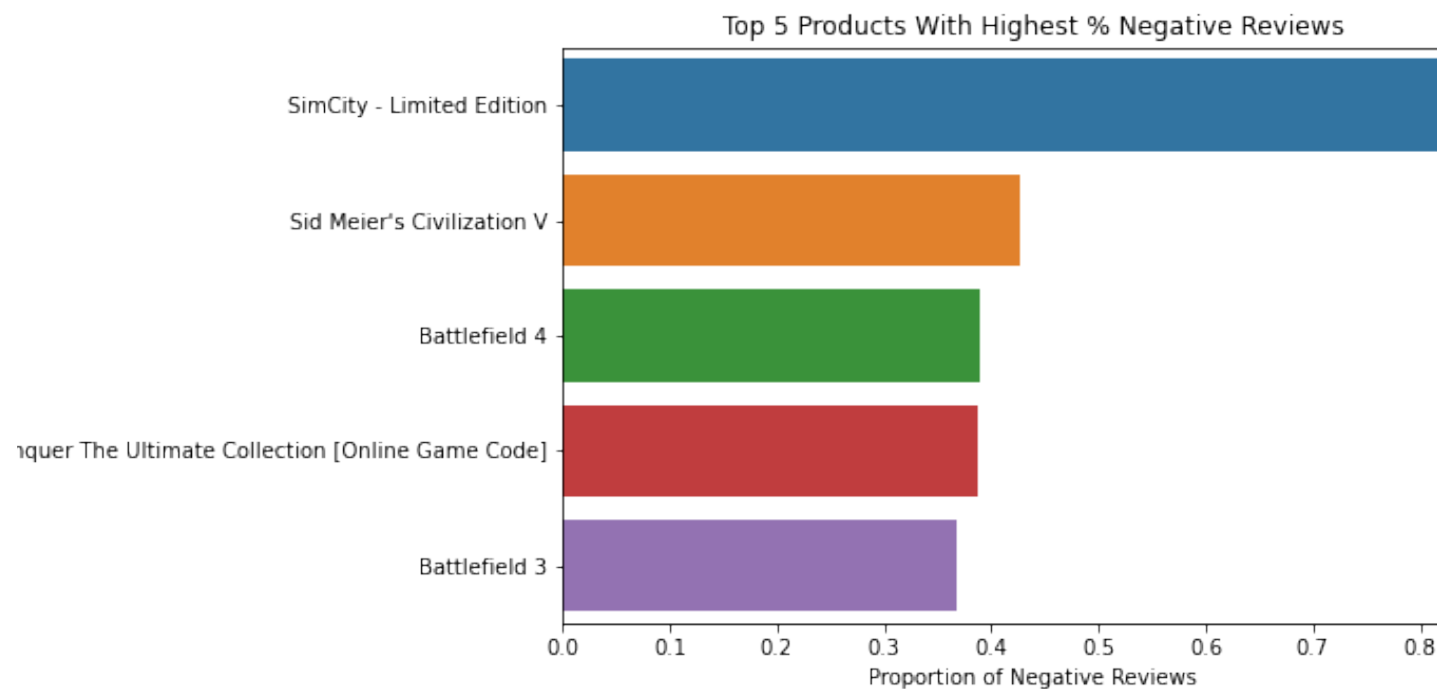
Compatibility

Waste of Money

Topics by Sentiment

Topics Mentioned in Positive Reviews

Positively Reviewed Products

Top 5 Products With Highest % Positive Reviews

Negatively Reviewed Products

Top 5 Products With Highest % Negative Reviews

- SimCity - Limited Edition
- Sid Meier's Civilization V
- Battlefield 4
- nquer The Ultimate Collection [Online Game Code]
- Battlefield 3

Proportion of Negative Reviews

A Closer Look at Bioshock Triple Pack (Positive)

A Closer Look at SimCity (Negative)

Future Work

- Larger Dataset
- Better preprocessing:
    - Lemmatization
    - SpaCy
- More Advanced Sentiment Analysis