YouTube Views Analysis

Exploring Video Titles To Predict Views
By: Nelson Genao

Background and Overview

YouTube is one of the largest platforms
Videos can reach millions of people worldwide

Used to promote whatever you want
Share important information and educate others

Monetize your videos
With enough views, videos can provide a steady income

Sources

LowkoTV's YouTube Channel - Publicly Available Data Only

- Lowko's channel primarily features StarCraft 2 (SC2), a military science fiction strategy game and popular E-sport
- Casts matches between professional SC2 players
- Used YouTube's API to download video data from 2018 - 2021, approximately 1,200 videos
- Data includes the video's title, thumbnail image, duration, and number of views

Aligulac - StarCraft 2 Pro-gaming Statistics

- Database containing professional SC2 player statistics active at different points in time
- Used Aligulac's API to download country, preferred SC2 race, and activity

Exploratory Data Analysis - P1

Strategies:

- Titles contain engaging words with some emphasized in all capital letters

- Game titles are usually behind the " : "

- Filtering special numbers related to technology, years, or matchups

- Use professional player names to get matchups and other statistics
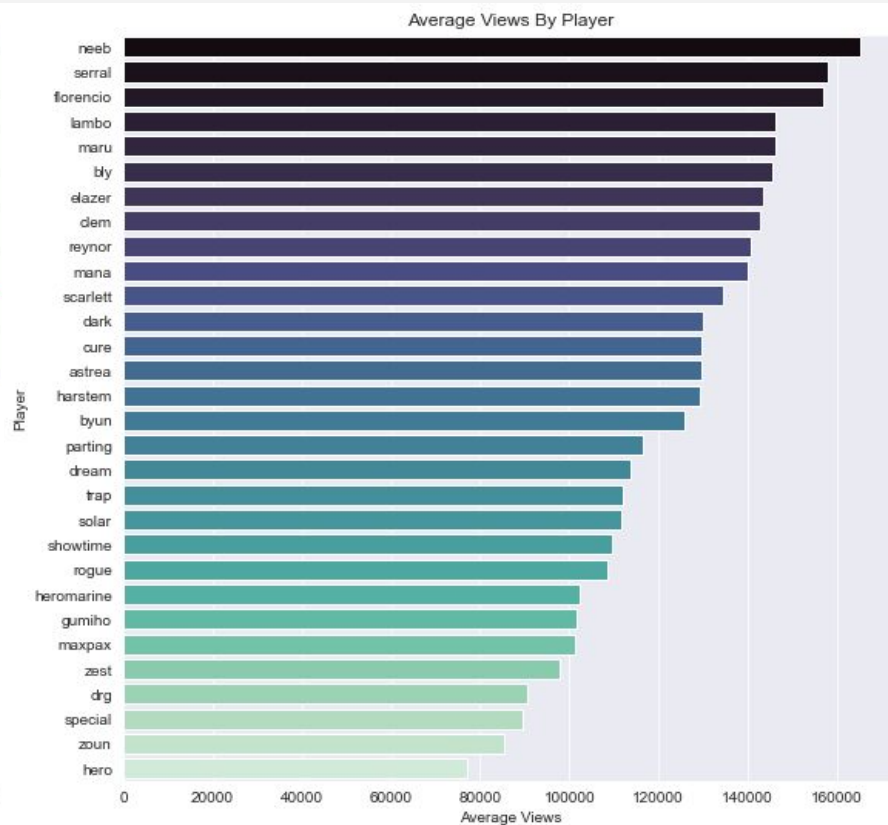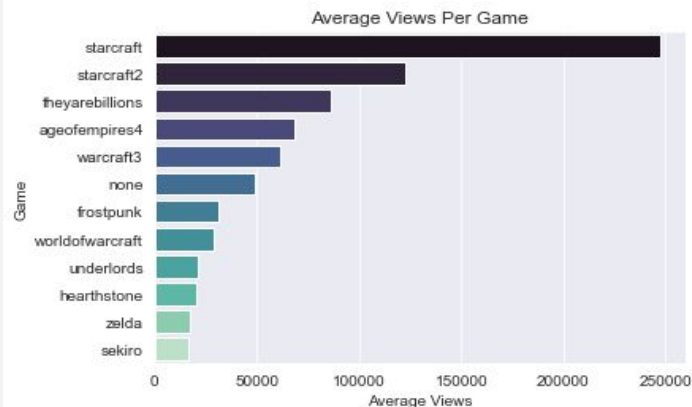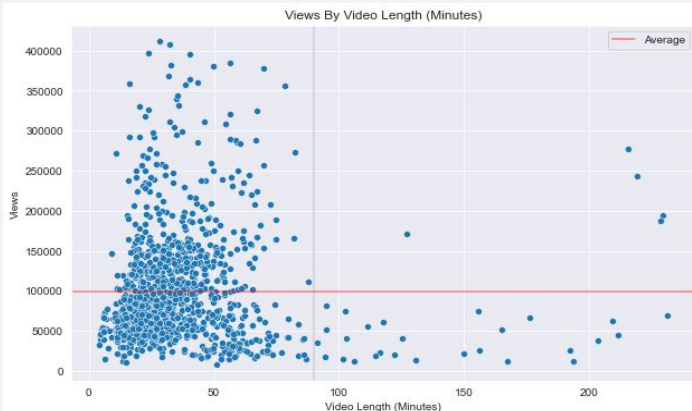
Example Titles Include:

Starcraft 2: The MOST PROMISING Pro Gamer of 2022! (HeroMarine vs MaxPax)

Age of Empires 4: StarCraft Grandmasters Play 2v2! (Lowko & Winter)

Frostpunk: Endless Mode - FINAL EPISODE!
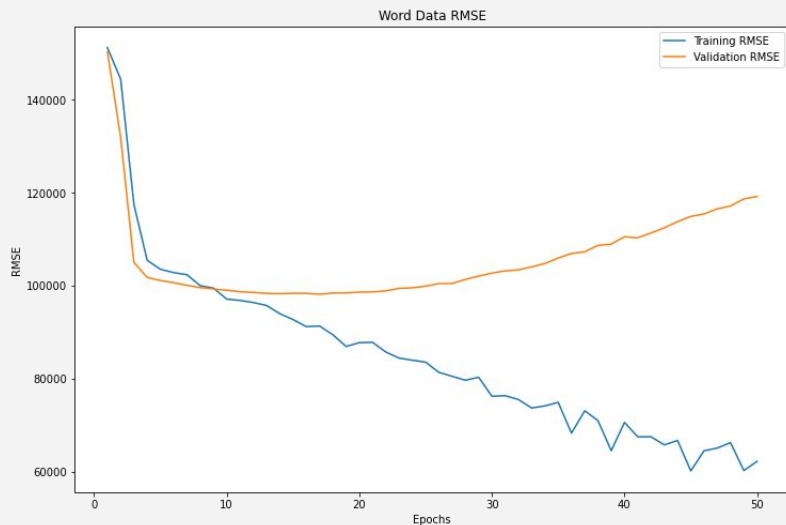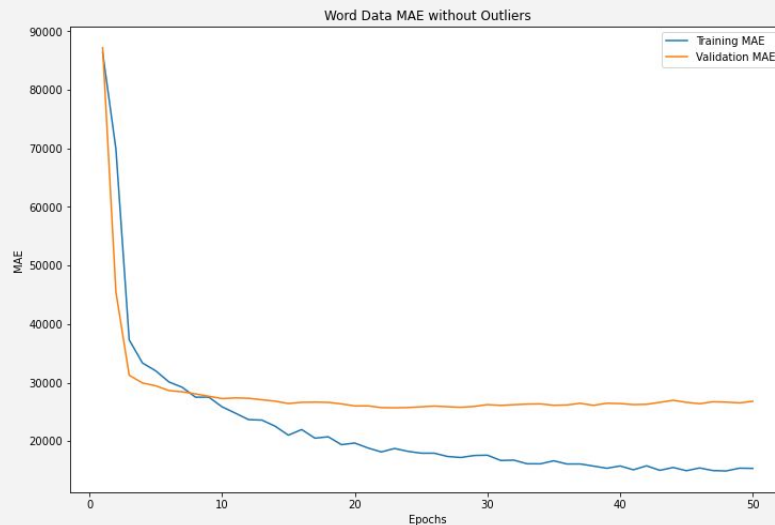
Exploratory Data Analysis - P2

Modeling

Base Model - RMSE With Outliers

Word Data RMSE

Avg Views: 107,394
Train RMSE: 92,474
Test RMSE: 80,133

Best Model - MAE Without Outliers

Word Data MAE without Outliers

Avg Views: 86,653
Train MAE: 24,838
Test MAE: 30,935

App Demo!

https://yt-views-predictor-ltv.herokuapp.com/

- Enter a sample title and predict the number of views with average views per word
- Look a pre-built visualizations for general data or SC2 specific data
- More features to be added upon request

Note: App takes about 30-45 seconds to load

Next Steps

Investigate Thumbnails
Explore effectiveness of different neural network architectures for thumbnail analysis

Continue Model Training
Improving accuracy with additional data and evolve the model to more recent videos

Add Significant Dates
See the effects of video game launch dates, large tournaments, or major content update releases on views for that period of time

Use Different N-grams
Investigate popularity of word combinations or matchups

THANK YOU

My GitHub: https://github.com/NelGen/NG-YouTube-Views-Analysis
App Link: https://yt-views-predictor-ltv.herokuapp.com/
LowkoTV's Channel: https://www.youtube.com/c/LowkoTV
Aligulac SC2 Player Data: http://aligulac.com/