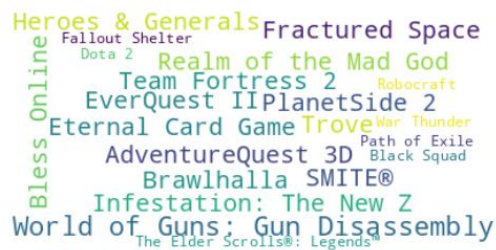# Final Project Report
Parin Pandya

Main Results

For the project, the results of my analysis was finding games, which had the most number of positive reviews. To demonstrate this, I decided to use a word cloud, to show the top 25 games, with the most number of positive reviews. The top 5 games are:

1. World of Guns: Gun Disassembly
2. Dota 2
3. Heroes & Generals
4. War Thunder
5. SMITE®

This shows which games are the most popular, and the most likely to be recommended.

Revised Results Figure

Cross Validation

For this project, I used cross validation to observe/discover some results. I used cross validation to compare between positive and negative reviews. To classify a review as positive or negative, I first extracted out the positive and negative words from each review and gave them a score. With this score, a positive word is classified using the sentiment analyzer, to be greater than 0, and negative to be less than 0. After getting each individual score, I took the sum of the positive and negative word scores, for each review. This way I was able to see the differences in the scores.

Additional Results Figure

Some Roadblocks

Some roadblocks that I faced were pretty early on, when I was trying to clean and filter my data. At the beginning, I was not able to remove reviews which were in a different language, or any words not in English. But using the NLTK stopwords library helped me remove those. Another issue I encountered was finding a way to enumerate my data. With all text data, I wanted a way to calculate which reviews were more positive, based on a score. For this, I had two options: TD-IDF or the Sentiment Analyzer. I chose to use the analyzer, because it cooperated better with the data.

Revised Dataset Figure

↑	user_review
0	hearing creepy pause moment write review wait ...
1	best game better sam pepper account need play ...
2	iffy know play easy master made floor due cert...
3	great game fun colorful side note though getti...
4	many cute tag right next horror tag steam firs...