The title of the project is Video Game Sales Analysis. The purpose for the data was to analyze video game sales across different genres, companies, and systems. Using the data, I was able to load, transform, aggregate, and visualize.

The data was loaded into the tool HDInsight using a Hadoop cluster. Using Hive, I was able to query and inspect the data.

One of the big insights I looked at was global sales. Using queries I can look at clean, round numbers or data that fits a certain criteria. I can check which games sold the most globally. Another insight was to look at critic scores. This data can be sorted in descending order to see the highest rated games.

Munge – the file in this folder contains documentation for data transformations done with the data.

Src - the file in this folder contains documentation for data aggregations done with the data.

Reports – the file in this folder contains data visualizations.