**ETL Project Report**

By: Austin Williams, Gregory Parker, Micah West, and Ryan Brown

For our project we used two tables that contained sales data for two different video game consoles ending in 2018. We acquired these tables from Kaggle and both were formatted as CSVs. To clean this data we had to delete a generic position column that was only present in one of the tables, merge the two tables together, drop all of the empty rows, created a column for the total of both consoles global sales, converted the year column from a float to an int, genre, and publisher columns generic and deleted the duplicates. We also renamed a few of the columns to make more sense. After this we laid out an ERD for how we wanted our tables to look and created the tables in PostGresSQL and joined them together to add all of the data together in one view. The ERD consisted of three tables, one for information about the Games, one for Playstation sales data, and another for Xbox sales data. The final view included all columns from the three tables and joined them on the Game Id column that they all shared. We chose to show the data this way because we felt that it was a gave us the most information we could have without sacrifice any of the extra sales data from other countries.