**ETL Project Proposal**

*Group 5: Christine Mitchell, Byron Allen, Kyle Peterson, Nick Allen*

For this project we have identified a SQLite database containing Super Mario Maker data, along with a CSV file holding general country data for the world and a JSON file which we will use to convert country abbreviations to country names or vice-versa. The SQLite file was found on data.world. Our CSV was found on Kaggle. Lastly, our JSON file was found on datahub.

Our SQLite file contains two tables which will need to be joined and cleaned before using them further. The CSV file we have will need to be cleaned down to only the columns we are interested in using. As for our JSON file, it will be used as a bridge to assist in joining our SQLite and CSV files as they hold only country abbreviations and full country codes, respectively. Once we get our files joined they will be ready for analysis.

This data was chosen to help identify the popularity of Super Mario Maker across populous nations by comparing populations with number of levels created in particular countries.