Matt Weiper

Alivia Varvel

Nick Sheets

**Unit 13 – ETL Case Study Project – Project Proposal**

**Overview and Objective**

We have decided to focus on 2 data sets we found on Kaggle, one analyzing video game sales and the other focuses on the rating of the game. We will include links at the bottom of this document for reference. Both datasets are csv files and we plan on pulling title, console, esrb rating from the rating dataset and joining that with name, platform, Global sales from sales dataset.

**Data Destination**

PgAdmin

**Requirements**

Our approach is extracting csvs, then cleaning them by creating new dataframes (Pandas) that we can align headers and join them, converting to sql, and then creating connection using SQL Alchemy.

**Data Sources**

Rating URL

<https://www.kaggle.com/imohtn/video-games-rating-by-esrb?select=Video_games_esrb_rating.csv>

Sales URL

<https://www.kaggle.com/gregorut/videogamesales>