**Project Proposal / Tech Report**

For Project 2, we worked as a team to complete the ETL process on a dataset.

**Extract**

**Transform**

**&**

**Load**

After considering multiple datasets our group decided on a dataset focusing on gender representation in video games across various platforms.

With three datasets (x amount of rows)…..

* We used three datasets focused on PlayStation games info sources from Kaggle.com. These datasets were on games data, character data & sexualisation info.
* The types of data wrangling that you performed (such as cleaning & joining,). The types of functions we used as part of our transform element of the process is we cleaned select columns to ensure the Datatable in our jypyter is similar to our tables within SQL over the 3 datasets and data tables, we joined the data set on…….as well cleaned up column names to either simplify names or renaming columns to shorter names.
* The schemata that you used in the final production database.

Attached is the scheme that we used in our final production database.

* Decide how you’re going to communicate with your group members when you begin. Create a Slack channel, exchange phone numbers, and ensure that the group knows each group member’s available working hours.
* Create internal milestones to ensure that your group is on track. Set due dates for these milestones so that you have a timeline for completing the project. Some of these milestones might include:
  + Project ideation
  + Data fetching/API integration
  + Data analysis
  + Testing
  + Creating documentation
  + Creating the presentation