**Technical Report**

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

**Extract**

**Transform**

**&**

**Load**

After considering multiple topic areas and datasets our group decided on the topic of gender representation within modern day video games across multiple platforms.

With three datasets (638 character data across 65 modern day popular games)

* We used three datasets focused on gender representation within modern day video 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 or more relevant names.
* The Schemata that is used in the final production of our database is saved into our group Git Hub Repository at; [https://github.com/kass173/Project2/blob/main/Gen\_Rep\_%20Games.sql](https://github.com/kass173/Project-2/blob/main/Gen_Rep_%20Games.sql)

**Group Communication**

* To be able to complete this project effectively we ensure we had communication within our group. We put in place a few mediums firstly being opening a slack channel just with our group members in it.

Graphical user interface, text, application, chat or text message

Description automatically generated

Secondly we also created a whatapp group again just with our group members init which we discuss the project during workin hours and ew also booked in extra session where we work together and on our respective parts using Team as the platform.

* After gaining all the information needed about the project we managed to break it down into managble work chunks but also having the abvaliliblility of the team to work on our work targets which we identified and is show in the process map on our git repository.
* Process Map - <https://github.com/kass173/Project-2/blob/main/Process%20Map/Colorful%20Process%20Prjt%202.png>
* We created workload and separated then between the group and set deadline for review so that we had a realistic timeline to complete the project.
  + Project ideation
  + Data fetching
  + Data analysis
  + Cleaning & Transforming
  + Data loading into final database
  + Creating documentation (Technical report and Project Visuals)