**ETL Project – Social Media**

Raven Thornton, Ramsha Zaman, Allison Gibbons

**E**xtract: your original data sources and how the data was formatted (CSV, JSON, pgAdmin 4, etc)

The extraction of our data was done from CSV files.

**T**ransform: what data cleaning or transformation was required

As we began to look at our social media related data, there were a few aspects we wanted to clean up or transform. We deleted irrelevant data to not confuse the reader and to make the data cleaner to answer necessary questions. We renamed columns to align with the datasets and merged with aligning data depending on the analysis. We transformed dataset to correspond with social media dataset and make the data more meaningful. We merged similar data points to make it easier to compare relevant data. There were times where we had to add further columns to better identify certain data types. Lastly, all the data frames were transported as CSV format.

**L**oad: the final database, tables/collections, and why this was chosen

As we continued to organize our data, all the CSVs were loaded into PostgreSQL through Pandas. This was chosen to organize the data smoothly and better understand the effects of social media usage.

**Works Cited**

Bhatia, Ruchi [Ruchi Bhatia]. (2021). Stress Analysis in Social Media: Dreaddit: A Reddit Dataset (Version 1)  [Data file and study]. Available from Web site: <https://www.kaggle.com/ruchi798/stress-analysis-in-social-media>

Chaturvedi, Kunal [Kunal Chaturvedi]. (2020). COVID-19 and its Impact on Students: Impact of Covid-19 Outbreak on Education, Mental health, and Daily Routine (Version 1).   [Data file]. Available from Web site: <https://www.kaggle.com/kunal28chaturvedi/covid19-and-its-impact-on-students>

Lucif3r. (2021). Social Network Ads: Predict whether the product has been purchased or not (Version 1). [Data file]. Available from Web site: <https://www.kaggle.com/d4rklucif3r/social-network-ads>

MorganMazer. (2017). How News Appears on Social Media: Comparing What’s on Twitter and Reddit to What’s Happening in the World (Version 1). [Data file]. Available from Web site: <https://www.kaggle.com/socialmedianews/how-news-appears-on-social-media>

Statcounter Global Stats. (2021). StatCounter. <https://gs.statcounter.com/social-media-stats#monthly-200903-202112>