Project Title
Data Science Employment Opportunities in Australia

Team Members
1: Anam Khalid
2: John Antony
3: Nishant Patel

Project Description/Outline
Find out the Data Science job opportunities using ETL process
(Extract, Transform, Load)

Research Questions to Answer
1: Is it worth it to be a Data Scientist in Australia?
2: Which IT company category is more popular in Australia in
Data Science?
3: Which job titles are getting the highest salaries per annum?

Note: The main purpose of the project is ETL.

Datasets to Be Used
[Australia Data Science Jobs | Kaggle](#)
[Data Science Job Listings - Australia - 2019-2021 | Kaggle](#)

Rough Breakdowns of Task
1: Create a new repository and clone to local. Add contributors
and their respective branches to keep everything organized.
2: Create files and folders and push to github.

3: Create a new database in pgAdmin. Then create two empty tables for two datasets, inner join them on job location.

4: For extraction, import dependencies in Jupiter notebook. Store CSV files into a Dataframe. For this, set the path to the CSV files and read and display the dataframes using pandas.

5: For transformation, clean the dataframe by keeping a copy of the columns that you're interested in.

6: For loading, connect to the local database. Refer to README.md for the details.

7: Check for the tables

8: Use pandas to load csv converted DataFrames into a database.

9: Confirm data has been added by querying the tables in both pandas and postgreSQL.

10: Join the two tables in pgAdmin or join the two tables in with pandas and SQLAlchemy.