Traveling the world on a mission to discover new data

In this assignment we explored the Python pandas library in depth, performing aggregations (<https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.aggregate.html>), data cleansing, and merge (<https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.merge.html>). We used the seaborn (<https://seaborn.pydata.org>) Python library do plot different chars to visualise and analyse the relationship among the data set features.

The data set used was obtained from:

<https://www.kaggle.com/lava18/google-play-store-apps>

It contains two csv files:

* googleplaystore.csv: details of the applications on Google Play. There are 13 features that describe a given app.
* googleplaystore\_user\_reviews.csv: This file contains the first 'most relevant' 100 reviews for each app. Each review text/comment has been pre-processed and attributed with 3 new features - Sentiment, Sentiment Polarity and Sentiment Subjectivity.

We tried to analyse the relationship among the Sentiment Polarity, Sentiment Subjectivity and Rating and other features of an app.