GOOGLE PLAY APPS

In this notebook, we analyzed over ten thousand apps from the Google Play Store. We can use our findings to inform our decisions should we ever wish to create an app ourselves.

Let's take a look at the data, which consists of two files:

- apps.csv: contains all the details of the applications on Google Play. There are 13 features that describe a given app.
- user_reviews.csv: contains 100 reviews for each app, most helpful first. The text in each review has been pre-processed and attributed with three new features: Sentiment (Positive, Negative or Neutral), Sentiment Polarity and Sentiment Subjectivity.

In [47]:
```
#Import Libraries
import plotly
plotly.offline.init_notebook_mode(connected=True)
import plotly.graph_objs as go
import numpy as np
import os
os.chdir('D:\Personal Project\project2\The Android App Market on Google Play\datasets')
import pandas as pd

%matplotlib inline
import seaborn as sns
sns.set_style("darkgrid")
import warnings
warnings.filterwarnings("ignore")
```

In [48]:
```
# Read in dataset
apps_with_duplicates = pd.read_csv('apps.csv')

# Drop duplicates from apps_with_duplicates
apps = apps_with_duplicates.drop_duplicates()

# Print the total number of apps
print('Total number of apps in the dataset = ', apps.count())

# Have a look at a random sample of 5 rows
print(apps.sample(5))
```

```
Total number of apps in the dataset =  Unnamed: 0        9659
App               9659
Category          9659
Rating            8196
Reviews           9659
Size              8432
Installs          9659
Type              9659
Price             9659
Content Rating    9659
Genres            9659
```