Google Play Store Analysis

Introduction:

This report was created to determine which Content Rating has the highest number of installs and how many categories it belongs to.

For my purposes, the focus will be on just the Category, Content Rating and Install columns and to determine if there is any correlation between them.

Body:

Data:

In order to get the data to a state that it will be able to answer the above question I needed to remove the columns that are not relevant to this analysis. I chose to do so as follows;

After looking through the unique values in each of the columns I removed '1.9' from the Category column as I could not determine what is was used for. There was also an item in the Content Rating column that did not have a value so that was dropped as well. Further cleaning of the data showed that the Everyone content rating had the highest number of installs.
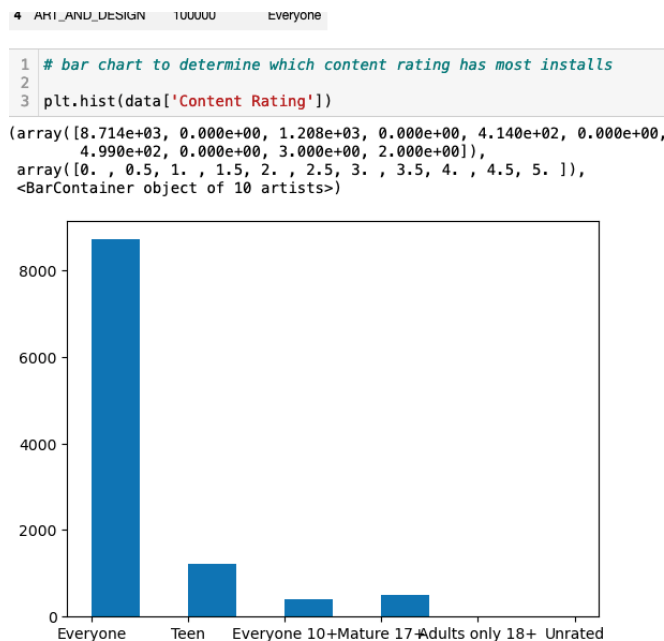
Results:

The total number of installs for the Everyone content rating was 8,714. Within that content rating there was 33 unique categories.

```
1  # total installs for Everyone data
2
3  everyoneData.Installs.count()
```

8714

```
1  # Number of values in Category
2
3  len(everyoneData.Category.unique())
```

33

Is there a correlation between the Content Rating of an app on the Google Play Store and the number of installations? I did not find any correlation as exhibited below. If there were any significant correlation the plot below would have a very visible pattern between the axis.
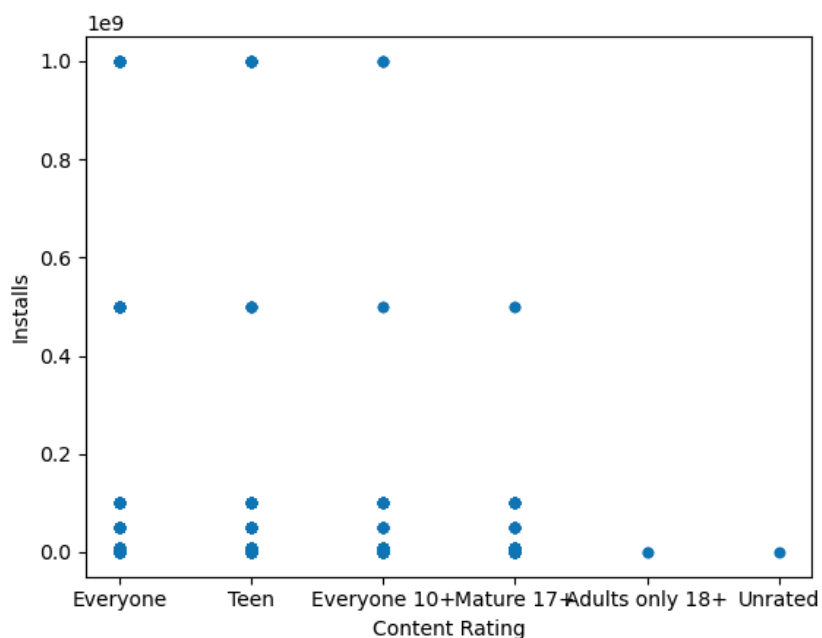
Conclusion:

The 'Everyone' content rating had the highest number of overall installations. There was not a significant correlation between what content rating and number of installations based upon my analysis.