Google Play Store Art & Design Report and Analysis

1. Introduction

This report will include data and information regarding the Google Play Store and will be used to answer a user story. The data utilized for answering the user story consists of 65 rows and 13 columns all related to the Google Play Store statistics and metrics regarding their Art and Design category.

The question we are tasked with is, “User would like to know how many installs there are per genre based on the ART\_AND\_DESIGN category.” In this report I will be discussing more about the following sections about the data, methods used, analysis performed, and the overall results. The hypothesis I will be trying to test in this report will be that the more generically named genre, “Art and Design”, will have the most installs due to the amount of apps in that genre compared to the other sub-genres of Art and Design.

1. Body

Data:

To start off this analysis I initially gathered all of the data into a Data Fram for inspection of rows which will be relevant to the user story and the hypothesis which will be tested. After the initial inspection I noticed numerous rows and columns would be unnecessary for this analysis and thus removed them. I decided I will be needing to keep 2 columns Genre, and Installs.

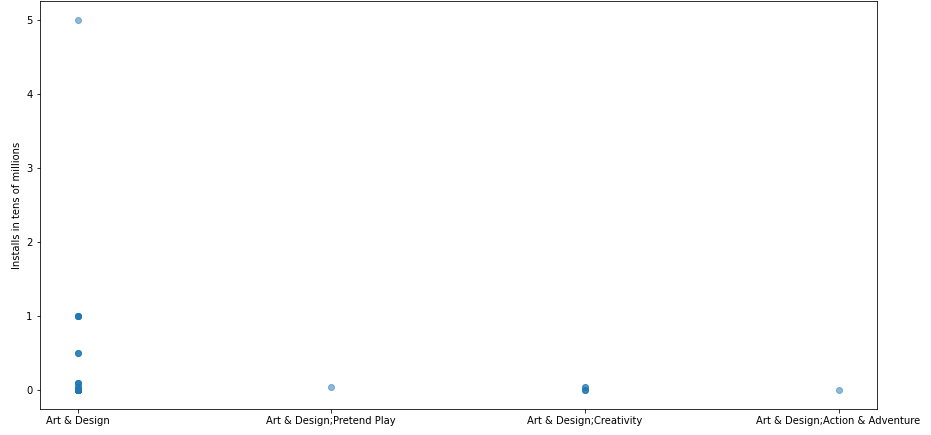
The next step taken was to start cleaning the data and preparing it for later analysis. One of the first things I noticed was an issue in the Installs column. This is a numerical column and would be needed in operations and thus is changed to the integer data type as well as removing the symbols.

I now have the essential data needed to complete the request of User Story 116. I will discuss in the section below my method for analysis.

Method:

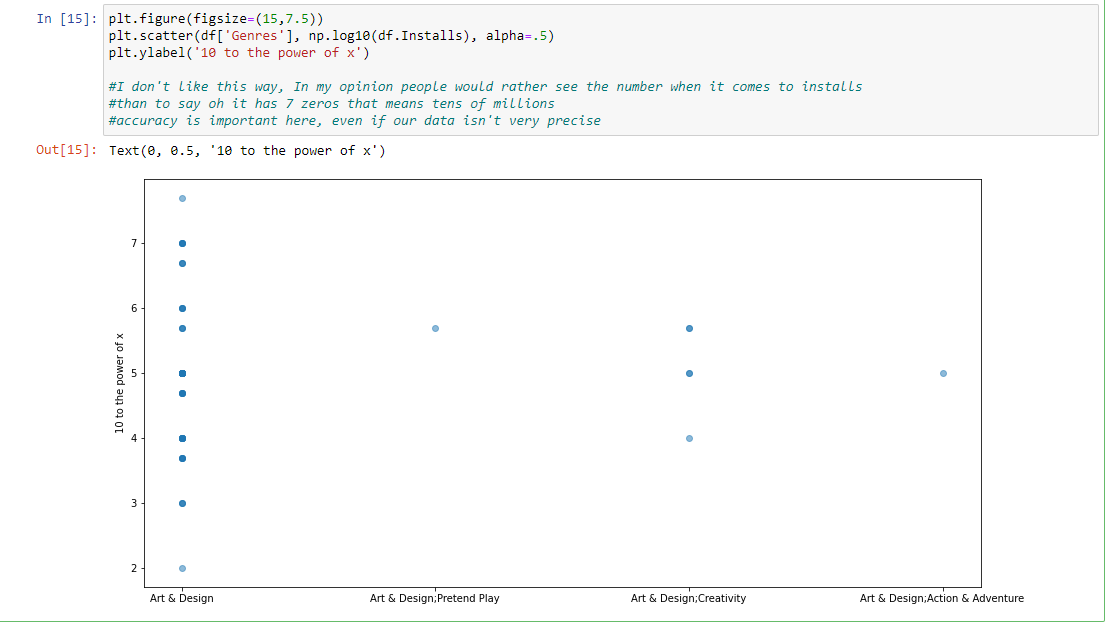
Once I had just the Installs and Genres columns it was an easy step to getting the visual we wanted.

A simple scatter chart would show an outlier in the Art & Design column. This particular app itself, “Sketch - Draw & Paint”, is more than the other categories combined.

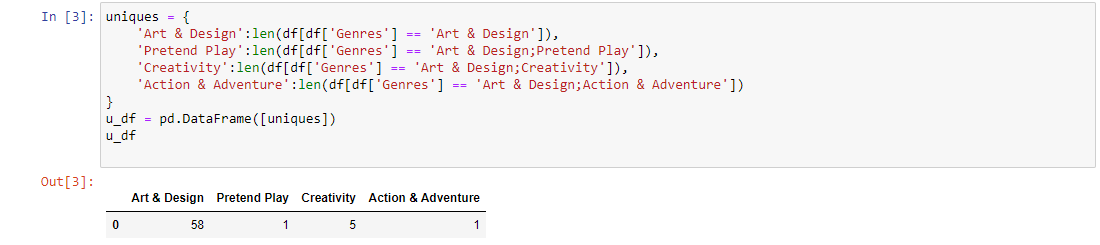


While this scatter plot shows the number of installs the “Art & Design” genre has over the other genres, it does not accurately show how many apps there are in this category.

In this next image you will see a different way to represent the diversity of applications in this category.



Now, here is a table showing the number of applications per genre:

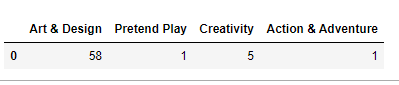


As you can see, the more generically named genre “Art & Design” has the majority of the category installs, which will be discussed more in-depthly in the next section.

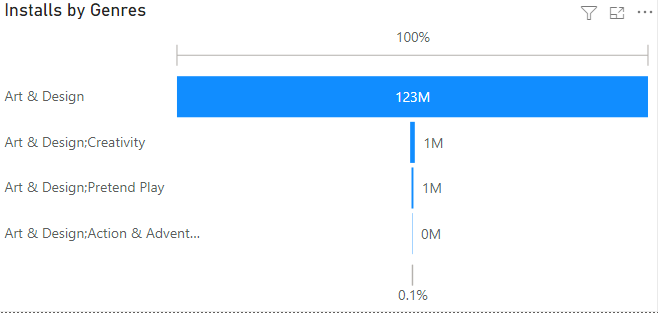
Analysis:

Gathering the information the way shown above immediately displays the gap between the genres of the ART\_&\_DESIGN category. One can see on first glance that the top installed application for the Art & Design genre has more installs than all other genres combined.

The Art & Design genre also holds 89% of apps for the overall category :



And holds 99% of the downloads for the overall category:



Results:

After analyzing the past few visuals you can see how the conclusion next is a simple one to reach. Even removing the top installed from the leading genre does not change the rankings of the genres for installations.

Conclusions:

To conclude this report for number of apps installed per genre of the ART\_&\_DESIGN category with the hypothesis of “The more generically named genre, “Art and Design”, will have the most installs due to the amount of apps in that genre compared to the other sub-genres of Art and Design.”. I will say that the hypothesis was validated, but with one caveat. The name of the genres do matter, and if there were more diversity and precision with the naming convention for this particular category then our dataset would show a much more accurate representation of install per genre for the ART\_&\_DESIGN category.