

TDS 3401 DATA VISUALIZATION

PROJECT

**TITLE OF PROJECT:**

**What is the factor that affect popularity of an app?**

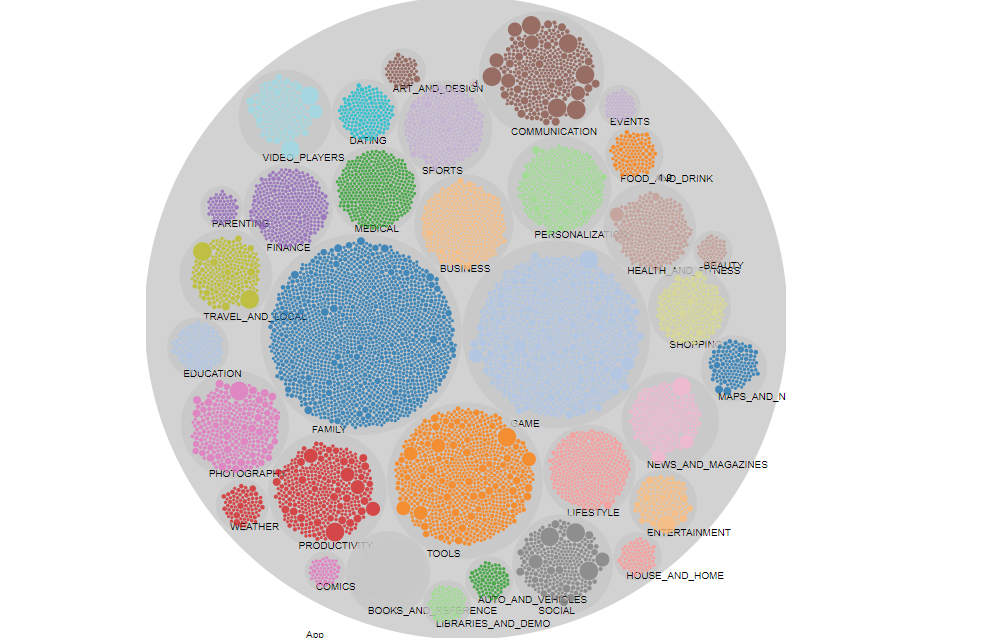
|  |  |  |
| --- | --- | --- |
| **NO** | **STUDENT ID** | **STUDENT NAME** |
| 1 | 1161100641 | ALEX WOON JUN RONG |
| 2 | 1161304221 | CHEW SHI LIANG |

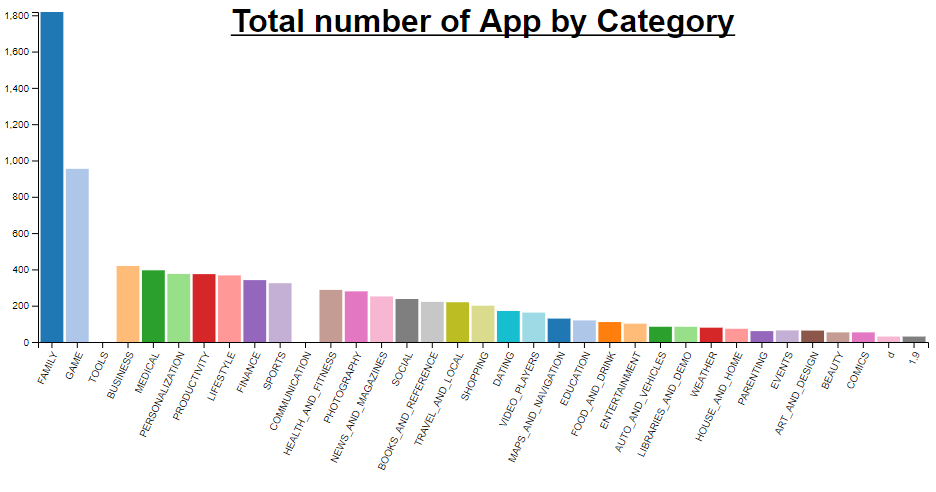
**Introduction**

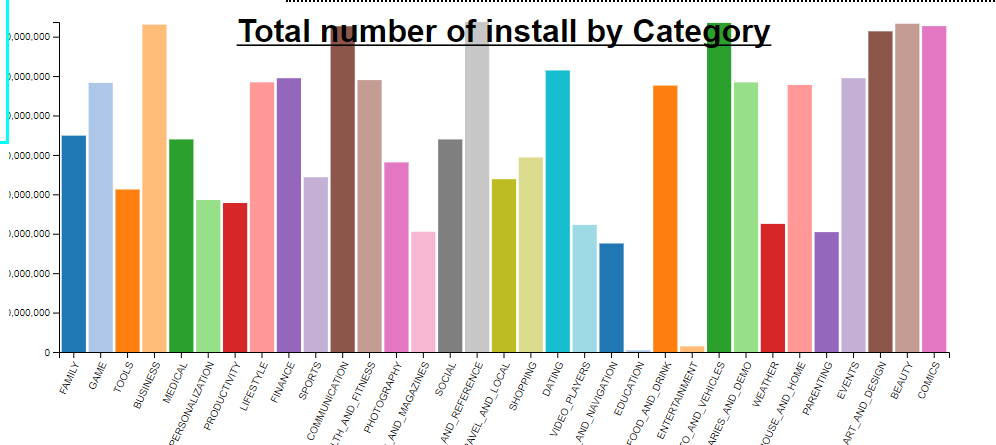
"There is an app for that", we hear this phrase very frequently nowadays. It is true that there is a mobile application for almost everything these days. People are migrating from computers to mobile devices and as a result, mobile applications are becoming that important in our daily lives. According to a study, an average person in the US spends over two hours browsing over his mobile applications. Therefore, mobile application development is definitely a huge potential to make big money .

In this project, the visualization will show that how each field affects the number of installed of an apps. Besides, the visualization will also show the correlation between each fields and the number of installed. At the end of project, we will conclude that what is the main key to design an popular apps.

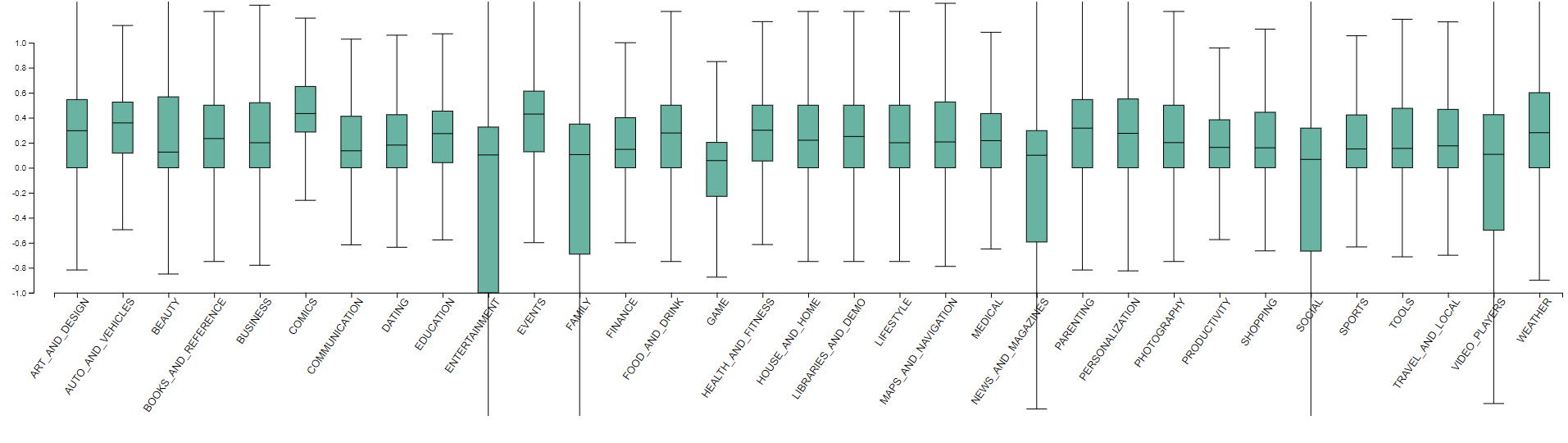
**Storyboard**

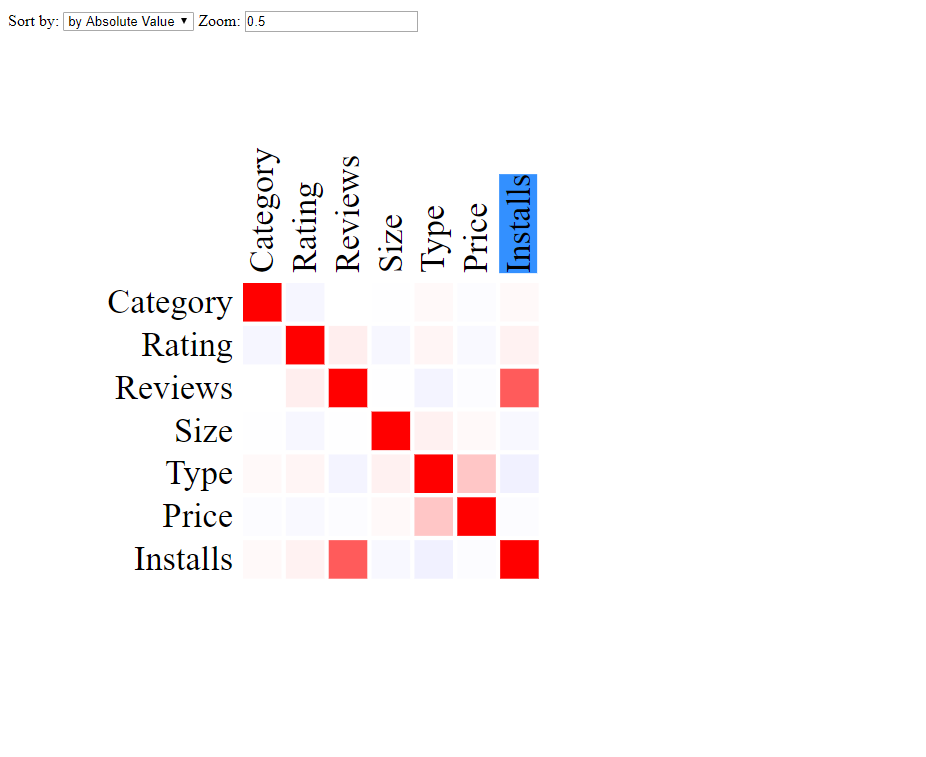
The dashboard will include four visualization, which include one bubble chart, one box-plot diagram, one sankey diagram and one correlation diagram. The chart on the top left is the number of installed by category. By visualizing the bubble chart, we can know that which of the category will be most popular. Based on the visualization, we found that category of family contains most number of app in playstore. The second higher of number of app in playstore is Game category while third is Tools Category.





The top right box-plot diagram shows that the sentiment result of reviews for each category. We can visualize the maximum, minimum and the average of sentiment result for each category in this diagram. Based on the diagram, we can observe that the category of COMICS and WEATHERS obtain overall higher sentiment score compare to the others. While the entertainment category obtain the lowest sentiment score.

Next, the bottom left diagram is a bar chart which represents the number of installed by each category. We can observe that which category of apps has highest popularity. Based on the visualization, we found that



Lastly, the bottom right diagram is the correlation for each fields. We can observe that how strong the relationship between number of installed with each field. There are 7 fields listed in this diagram. By comparing the number of installs which means the popularity of app, we can observe that the reviews has the highest correlation to the number of installs. Hence, the reviews of app is the biggest factor that affect the popularity of an app. The second high of correlation will be Rating of app, while the third high of correlation is Category which affect the popularity of an app.

In the nutshell, we can conclude that what is the factor that affect the popularity of an apps.

**Development Process**

We have spent around 3 days for this project. Both of us have been divided two visualization task. The aspect which has spent most our time is data preprocessing. Besides, we spent a lot of time on discovering the interesting visualization. However, it spent us many time and we fail to do so. There are many versions of d3 available in google. However, it is very hard for us to look for all the exactly same version’s d3 graph visualization.

As a conclusion, we have learnt a lot of knowledge in d3.js which definitely will be beneficial to us in the future.