**App level insights**

* Through top 10 most installed and most reviewed/popular apps, we can generate revenue by increasing advertisement in them. We can direct users towards our revenue generating apps. Also, paid versions of such free apps can also be introduced including some exclusive features.
* With least popular and least installed free apps, we probably don’t need to invest resources in them, so they can be pushed towards decommissioning.
* From the most profitable apps, we can conclude that more apps with low price and high profitability must be promoted.
* Apps with least ratings can also be decommissioned or improved.

**Categorical Analysis**

* Most popular categories amongst developers include Family, Game and Tools; while the users prefer Games, Communication, Social and Productivity above others, followed by Tools and Family.
* These above categories are the ones generating the most revenue as depicted in graph 5, so it’s natural to proceed in such direction. App size doesn’t seem to have much effect on the number of installs.
* Paid apps seem to occupy 7.4 % of the market share while 92.6% of the apps are free.
* The correlation plot between columns shows slight correlation between number of installs and the number of reviews, but no significant correlation between any other pair.
* The Rating plot seems to be a skewed normal distribution left skewed around average 4.2 stars.
* We plotted the rating in a boxplot distribution to check if any discrepancies and differences exist between categories. Rating of an app signifies the level with which the user preferred the app, and is an important metric in understanding the success or failure of the same. Through this graph, we found that some categories have performed better than others according to users. ‘Arts and Design’, ‘Beauty’, ‘Comics’, ‘Entertainment’, ‘Events’, ‘House and home’, and ‘Parenting’ categories seem to be doing much better than others. We can recognize this by the lack of many outliers and a more uniform and balanced distribution. The results can also be attributed to low number of installs and apps, but it also means that these apps have a narrow audience or demographic base. This is opposed to categories like ‘Family’, ‘Tools’, ‘Games’, ‘Medical’, ‘Finance’, ‘Business’, etc. who have a much broader base and a huge number of bad ratings, indicating a need to improve/remove the lower rated apps.
* Higher number of reviews whether positive or negative signifies a higher engagement of users with the apps, and hence a higher chance of success when developing apps in most reviewed categories.
* The sentiment distribution shows the degree of polarity of the distributions. In general, we find 63.6% of reviews to be positive, 25% negative and 11.4% neutral. Graphs also show a distribution of sentiment subjectivity and polarity, which can be further used to make deeper inferences about the apps.
* Next, we used sentiment subjectivity to filter reviews of the apps and found the apps found most engaging by the users. This was given by filtering subjectivity > 0.5 and watching apps with highest number of reviews.
* Lastly, we checked the last update history of the apps to figure which apps might be discontinued or decommissioned. Apps with last update before 2012 were said to be discontinued, with last update between 2012 and 2017 were said to be obsolete, and last update in 2018 were classified as active.
* A graph of content rating was also plotted, but not much inferences were found.