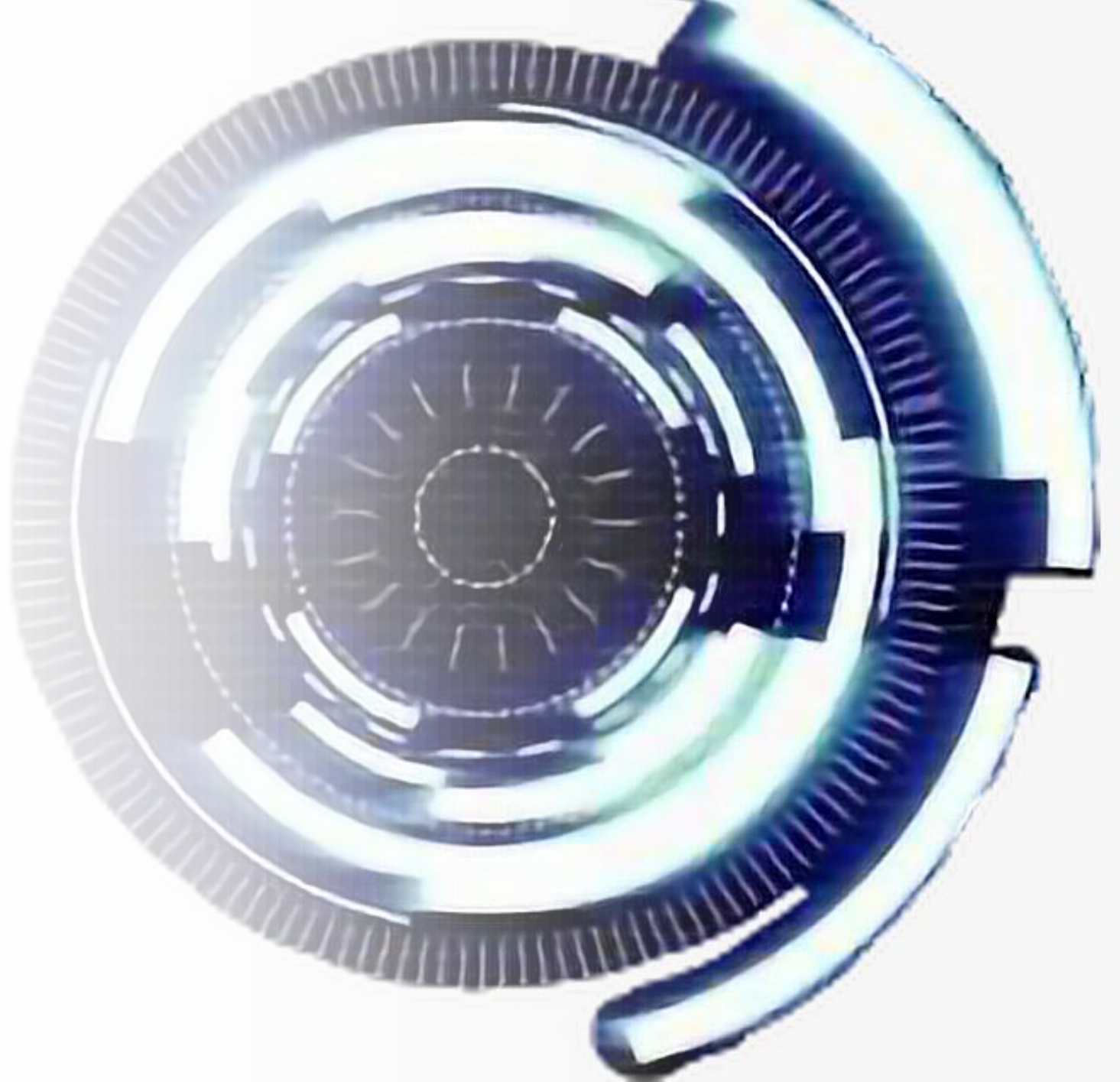




Foresight*

Implementing data to inform business decisions



Mobile Phones Applications

Have you ever had a great idea for an app?

The Mobile phone Apps industry:

- Generates annually 71.4 billion US Dollars, worldwide*
- 3.5 billion potential customers within reach*
- Great idea, little investment, marketing research, timely launch?

Not so fast!

- 6,140 mobile apps were released through the Google Play Store every day*
- 8 out of 10 Apps do not even make to the App Stores
- Apps for making Apps

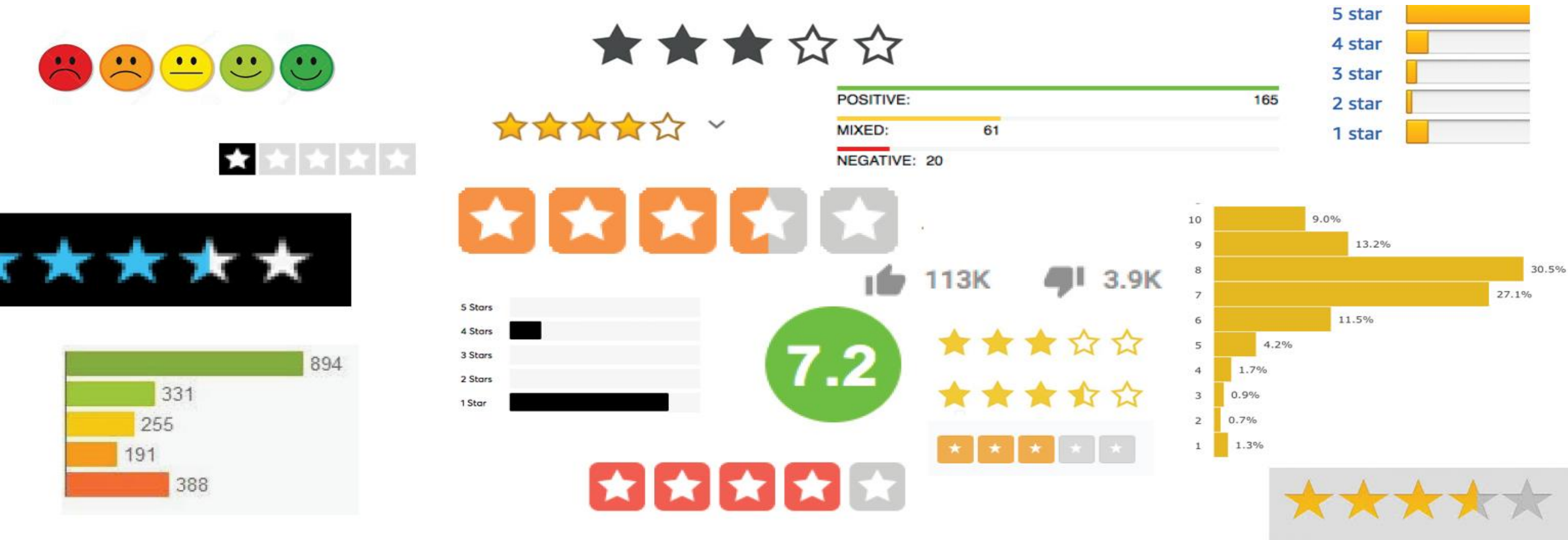


WELCOME
TO
OBLIVION !



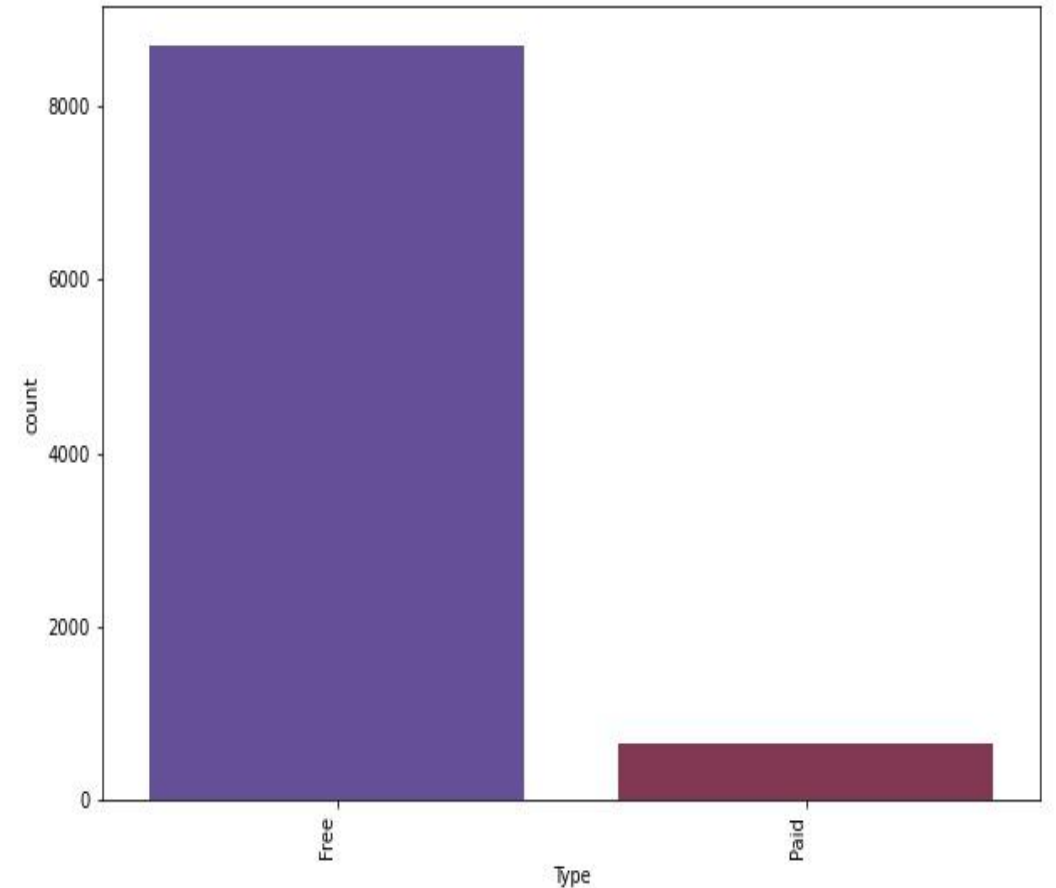
Buying Behaviour

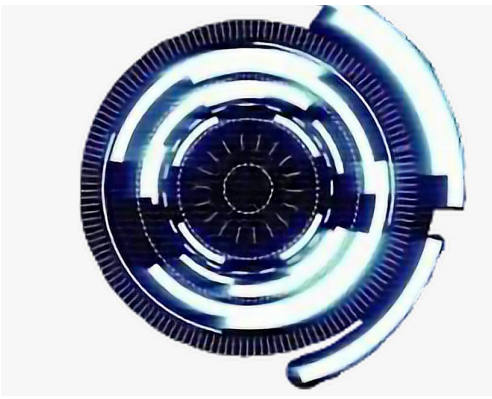
- Feedback from their fellow consumers, inform decisions
- Learn from others' experiences - Hindsight for Foresight
- Short cuts such as stars, points, ticks, download numbers and ratings have been invented on top of the standard written reviews – a wealth of data.



What do we know about Mobile Apps' Ratings?

- They are mostly designed for the Family and Games
- With Content Rated as for Everyone
- The great majority are Free of charge
- Most of the apps in the data frame have 1 million installs or more, followed by 10 million and 100k installs
- The Distribution of Rating scores are clearly skewed to the left, with most of the Apps presenting scores between 4.0 and 4.7.

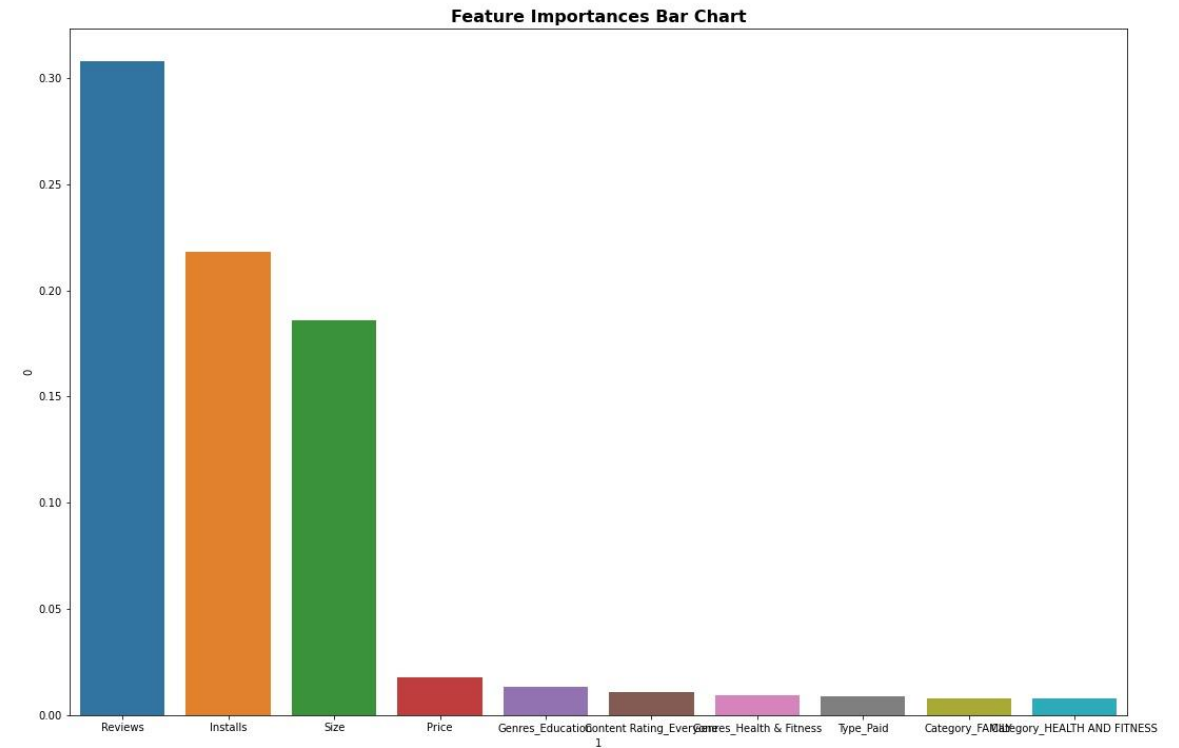




Foresight is a tool which uses machine learning techniques to help developers to anticipate potential issues with their mobile app design and execution.

After extensive analysis, our investigation has showed that in order to obtain the so important ratings, anyone considering developing an App should focus on the following features:

- Reviews
- Number of Installs
- App size
- Price



Thank you ! Any questions ?

- References:

<https://www.geeksforgeeks.org/confusion-matrix-machine-learning/>

<https://medium.com/bigdatarepublic/feature-importance-whats-in-a-name-79532e59eea3>

<https://www.statista.com/statistics/276703/android-app-releases-worldwide/>

<https://kharshit.github.io/blog/2018/03/23/scaling-vs-normalization>

<https://en.wikipedia.org/wiki/AdaBoost>

[https://en.wikipedia.org/wiki/Random forest](https://en.wikipedia.org/wiki/Random_forest)

<https://towardsdatascience.com/understanding-hyperparameters-and-its-optimisation-techniques-f0debba07568>

https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html

<https://towardsdatascience.com/logistic-regression-model-tuning-with-scikit-learn-part-1-425142e01af5>

<https://www.statisticshowto.com/regularization/>

<https://towardsdatascience.com/introduction-to-logistic-regression-66248243c148>