Student name : Michael Gokhler

Group : Tuesday – Thursday data science

Homework 1 report :

1. What are three conclusions we can make about Kickstarter campaigns given the provided data?

Theater is the most popular category (34%) and plays is by far the most popular (26%) sub-category.

Journalism is the least popular category (0.6%) and **all of them (24)** eventually cancelled.

Music is the category with highest success rate (77%)

Food is the category with highest fail rate (70%)

Projects tend to be most successful when lunched during April-May (60%-61%) and least successful when lunched in December (44%)

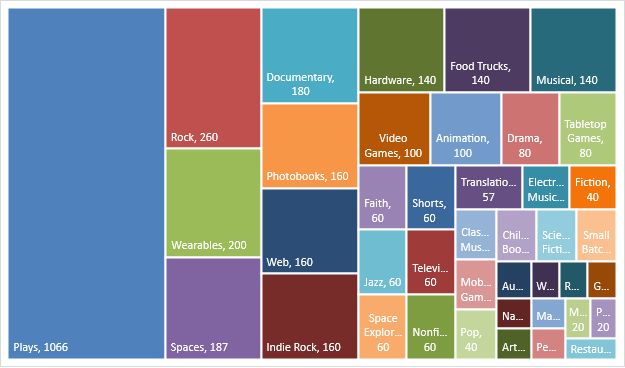
1. What are some of the limitations of this dataset?

The sheet is missing data about:

* The reasons for failure / cancelation of a project.
* The distribution origin of the donors / backers
* The profile of the donors

1. What are some other possible tables/graphs that we could create?

* Proportion of each state over the years (shows that there is a significant decline starting 2014)
* Treemap of all subcategories



* Different metrics comparing the average donation (per category / sub category / over years etc. )
* Different metrics on the country (success rates / most popular categories etc. )