

NAME: VICTOR KABUI KARANJA

PRESENTATION OF FINDINGS

PHASE 1 PROJECT

DATA SCIENCE PART TIME (dsfpt01)



BUSINESS BACKGROUND:

- Microsoft a company dealing with operating systems among many other deliverables intends to venture into the film industry
- Having no prior knowledge on the film industry, consultation and adequate research had to be conducted to ensure they did a good investment.



PROBLEM STATEMENT

- Lack of Knowledge:
Microsoft wills to venture into the movie creation sector however, having not been in the trade before they risk failing.
- Resource Wastage:
Microsoft is not willing to go into losses or drain their resources into a money pit
- Need to Keep with the Trend:
The other big companies 'competitors' having moved into the film industry, Microsoft has to ensure it is not left behind in the old ways of running, else they will be out of the market with no point .of return

SOLUTION



In order, to mitigate through the problem statements, I was hired to perform Exploratory data analysis from the Web and give insights on the way forward to be taken.

Some of the graphical representations used included bar charts, line graphs, histograms among many others to arrive at the solution:

- Microsoft should invest in the ImdbDisplay types of films because of their high profits compared to other types of films for a period of 7 years continuously
- Microsoft should not move after originality of the video's title:
very small negative correlation in the revenue influence on the total gross revenue and profits, this could be because of the poor naming of the films.
- ImdbDisplay is the most common type of movie types :
This forms a ready platform with a ready market. Having the ImdbDisplay type of videos will grant Microsoft a step forward since it is already well known by the vast majority of the people.



MVP

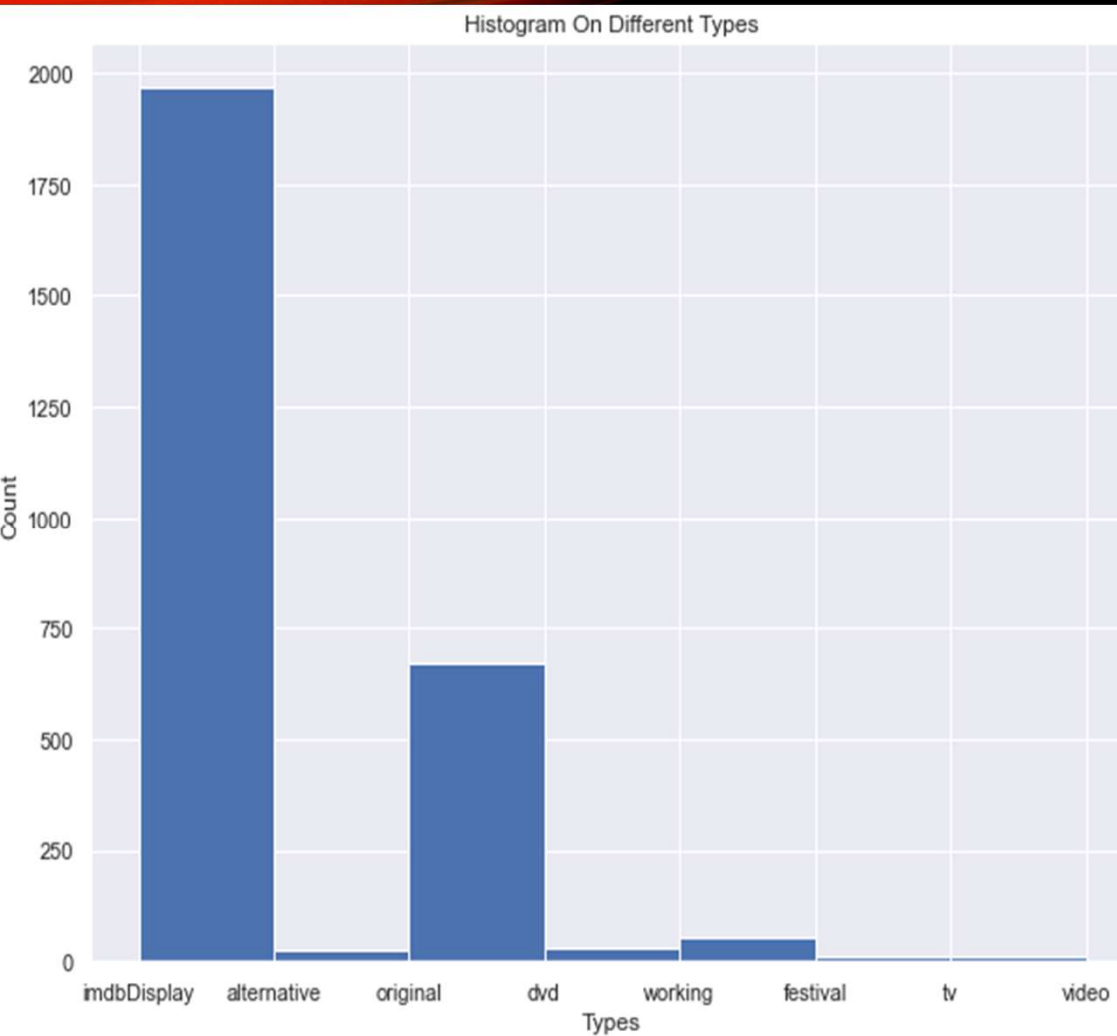
In order to have the most impact, the movies should be introduced to targeted groups of people through application ads based on their previous web histories



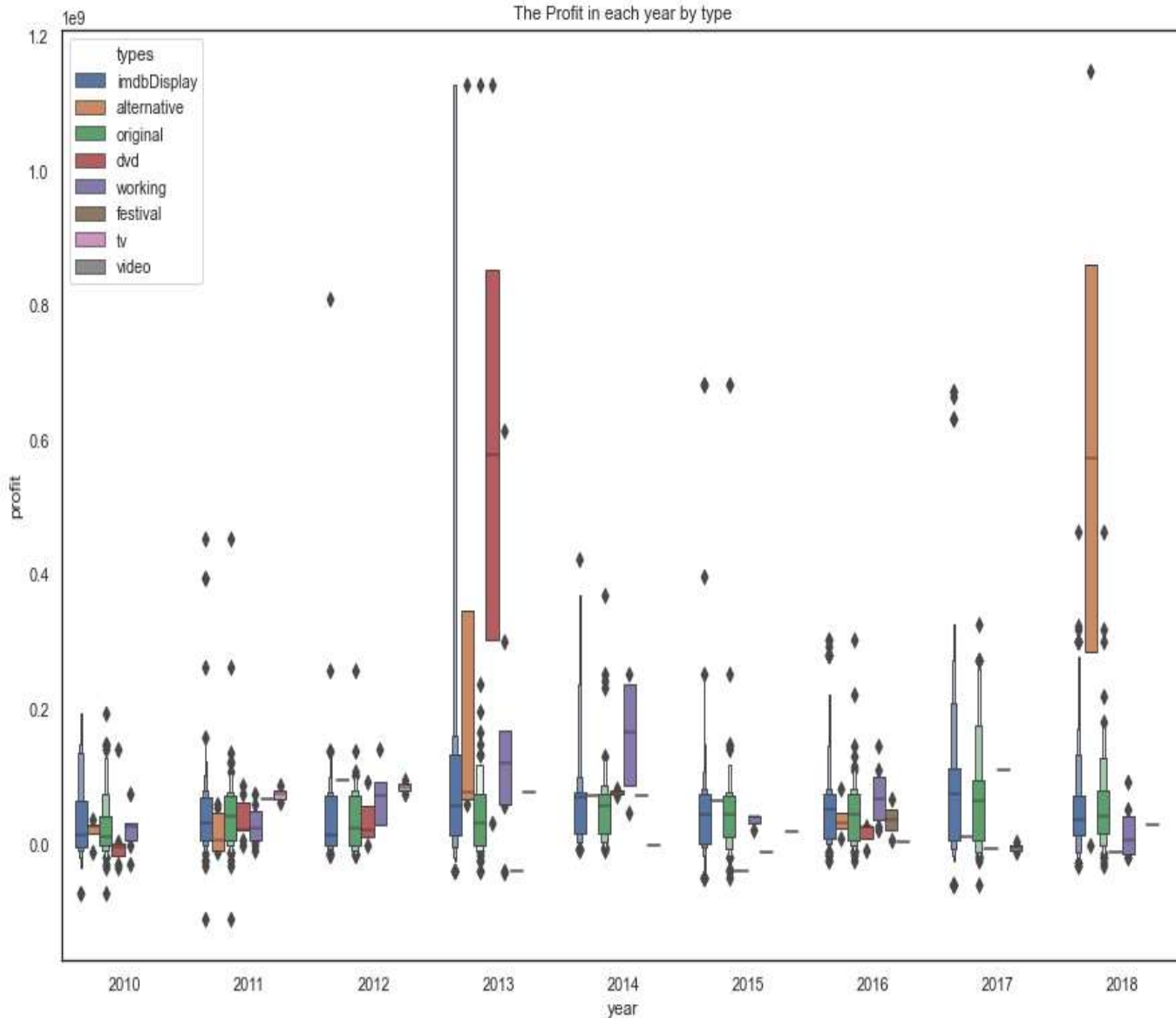
DEVELOPMENT PROCESS

1. Research and data analysis
2. Discussion of the findings
3. Installation of the Movie studio
4. Releasing of sample movies to a targeted group for feedback
5. Discuss any feedbacks obtained and make necessary corrections
6. Release to the larger population
7. Continuously receive feedback and improve
8. Research on what to improve so as to stay top of the market

DATA VISUALIZATION AND INSIGHTS

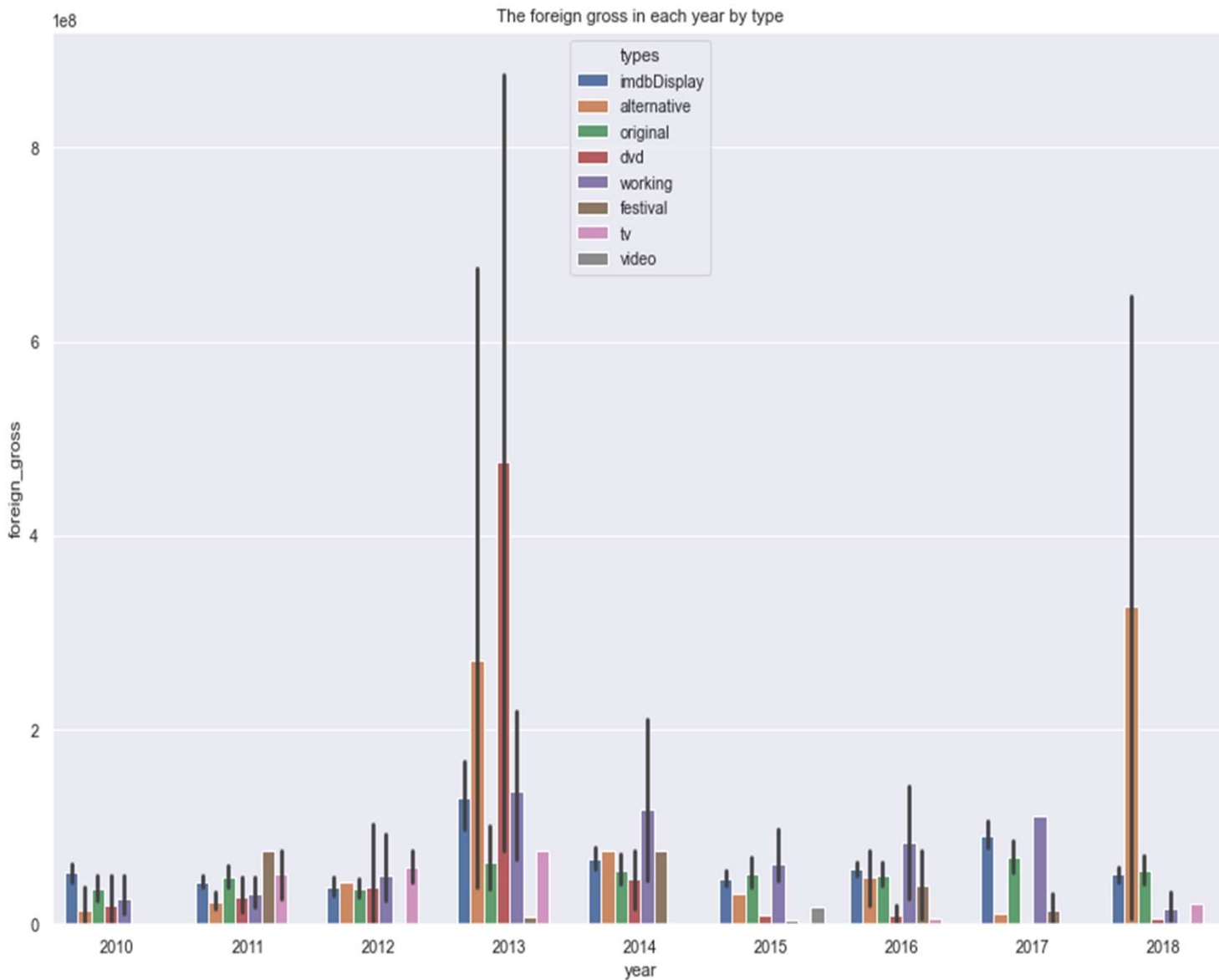


- A histogram indicating the most preferred type of movies.
- ImdbDisplay has the highest frequency followed by original



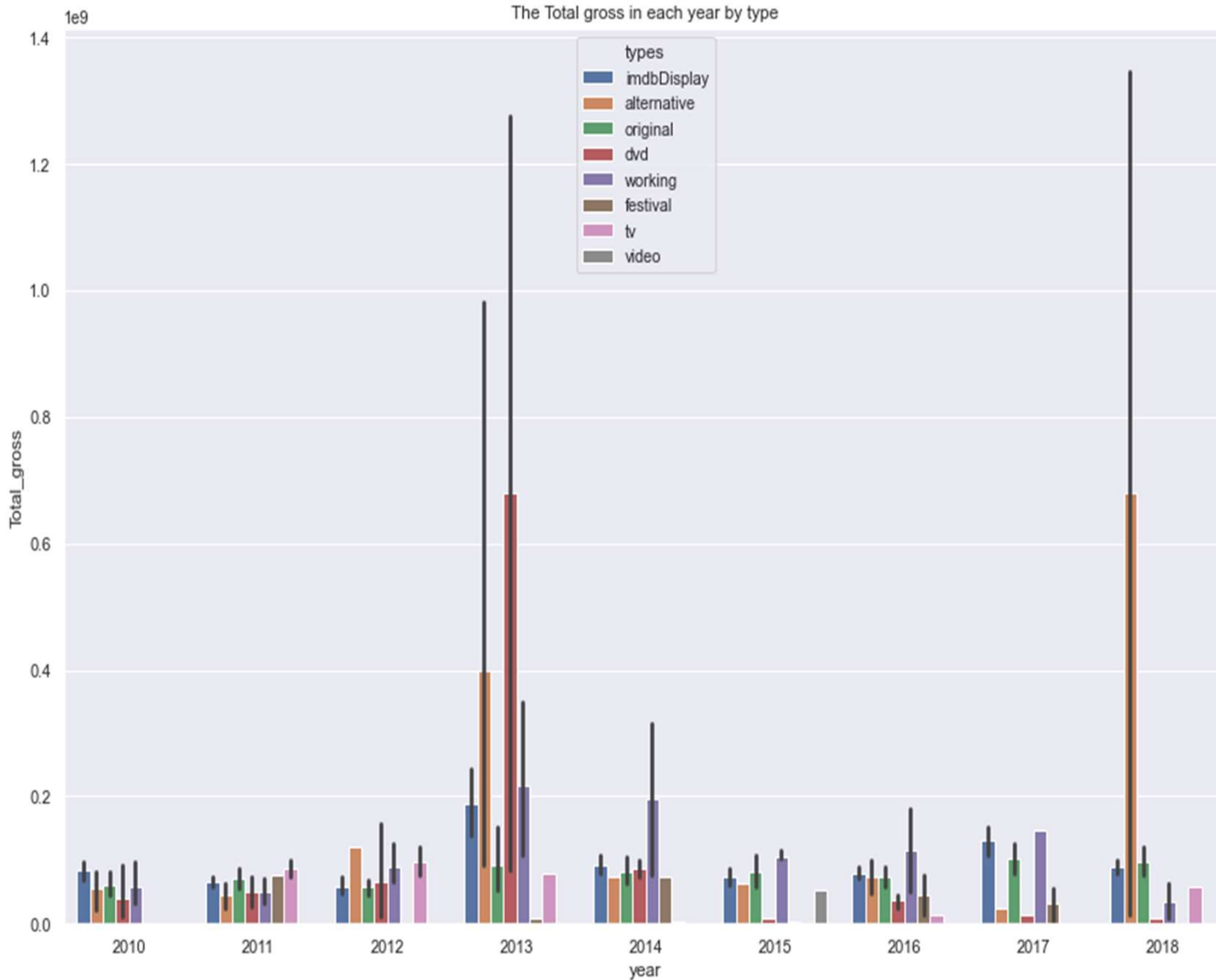
- A boxplot showing the profit distribution among the years when the data is grouped by the type of movies
- The Boxplot shows imdbDisplay had the highest profits from 2010 until 2018 when alternative types took the lead





- A bar-plot showing the relationship of foreign gross income of the various types across the years





- A bar-plot of the total gross income of the different types during the 8 years
- Where,
Total gross income = domestic + foreign gross income
- Alternative types gave the record highest in the 8 years but had its lows unlike imdbDisplay that maintained an average of about 1e8 dollars

Technologies used



The software used to perform the analysis included:

- Python programming libraries
 - ✓ Matplotlib.pyplot
 - ✓ Pandas
 - ✓ Numpy
 - ✓ Seaborn
- Jupyter notebook
- Github
- Git bash

CONCLUSION AND RECOMMENDATION



The Analysis above leads to the following recommendations to Microsoft:

- ImdbDisplay is the most common type of movie types -> The frequency on the histogram(last plot) shows that it is more than twice as common as the other types of movie types. This would minimize the risk on investment compared to the other types
- Originality of the video title has no big influence on the total gross revenue -> There was only one type of original that had its original titles and other types were performing better in terms of revenue generation. It is thus clear that the business should not venture fully into providing original titled movies, but rather aim at renaming the titles to create more traffic. This would increase their market. It is advisable that Microsoft invites stakeholders to invest their movies in their studio
- ImdbDisplay has a higher profit margin -> From the analysis it is evident that apart from the last year 2018, ImdbDisplay has been taking the lead in terms of the profits made. This gives a higher probability that the investment made through ImdbDisplay would be returned at a faster rate compared to investments made on other types of movies. Thus i recommend publishing ImdbDiagram type of movies on the Microsoft studio that you intend to start.