Long Term Digital Investment Strategy

As prepared for XYZ Limited

Business Problem

- XYZ Limited requires a longer term, more focused, digital investment strategy and presents experimental data in that regard.

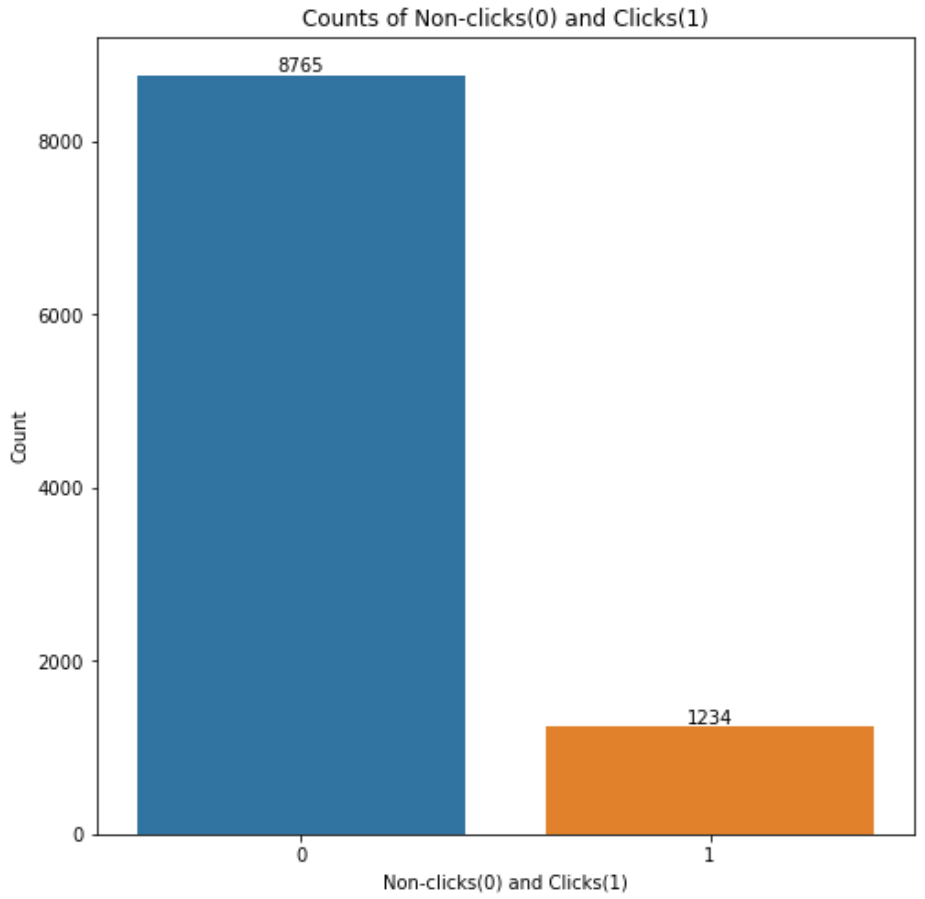
- To solve this problem
  - This project analyses the available data to find insights that helps us better understand click through rate improvement
  - This project makes use of the available data to train ‘click’ prediction models

Data and data quality issues

- Data provided was obtained by XYZ. It was collated by carrying out an experiment, advertising their service for the last 10 weeks across a wide variety of social, blog, and user content driven websites.
- Data quality issues faced were:
  - Presence of a major outlier handled by removal
  - Missing values handled by imputing averages
  - Data inconsistency, Handled by renaming inconsistent data

Exploratory data analysis Insight 1:
Percentage of Clickers

- Averagely, 12% of people who come across digital advertisements click on them



Exploratory data analysis Insight 2:

Analysis of Click through Rate

- Analysis shows a Click through rate of 0.0020

- This means that per 1000 impressions, 2 clicks are gotten.
- Assuming:

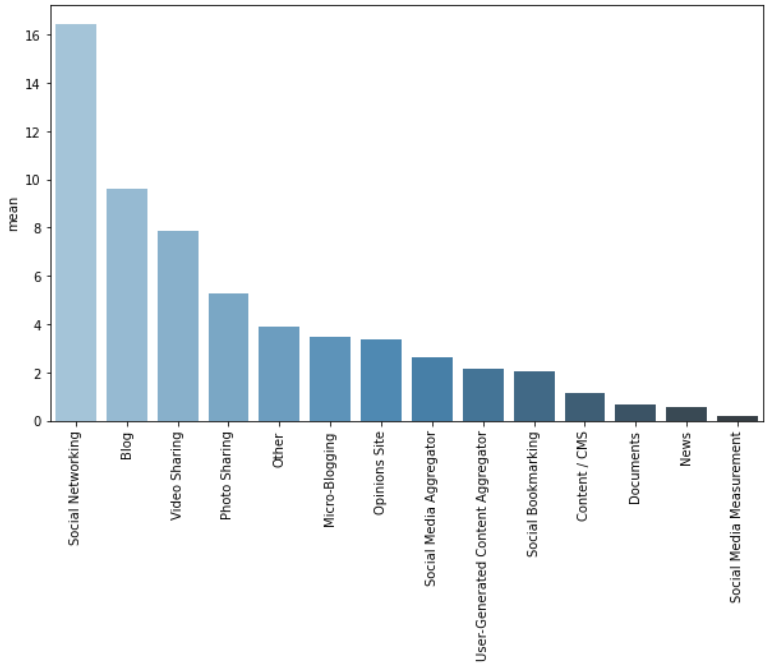
- CPM (cost of media) equal to £2.31 per 1000 impressions
- an average click is worth £1.03 to the client.

- The implication is that £2.31 is spent making £2.06 consequently meaning the present marketing strategy is incurring a £0.25 loss on every 1000 impression

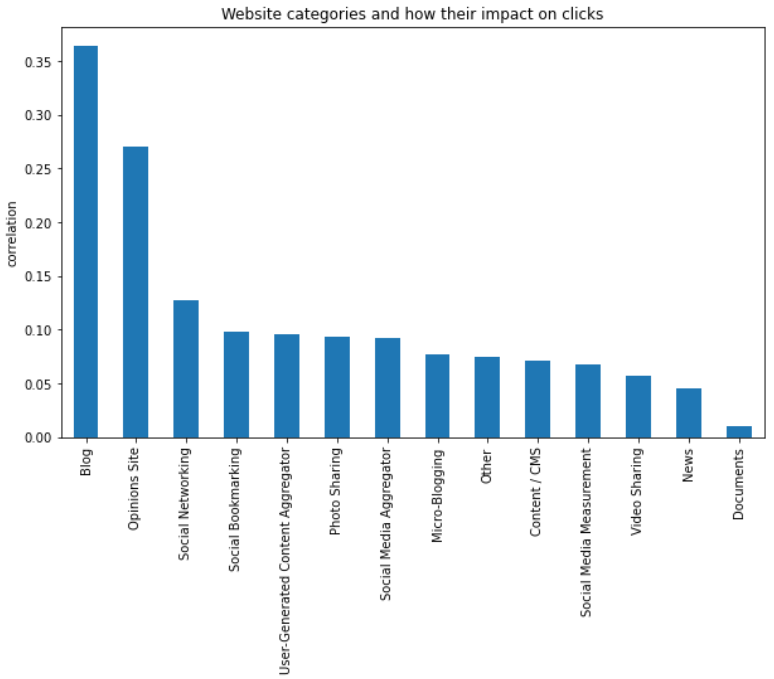
Exploratory data analysis Insight 3:

Analysis of website category

Different categories of websites impact click through rates differently. Using Correlation. It was able to determine how categories affect clicks obtained



Current mean impression on users per category

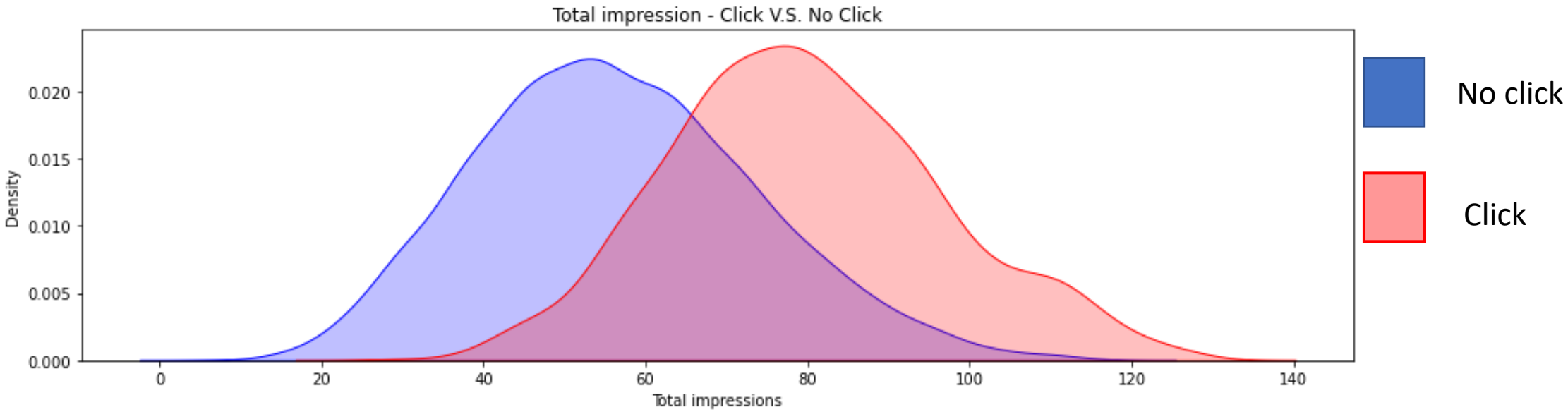


Impact of different categories

Exploratory data analysis Insight 4:

Optimum Clicks

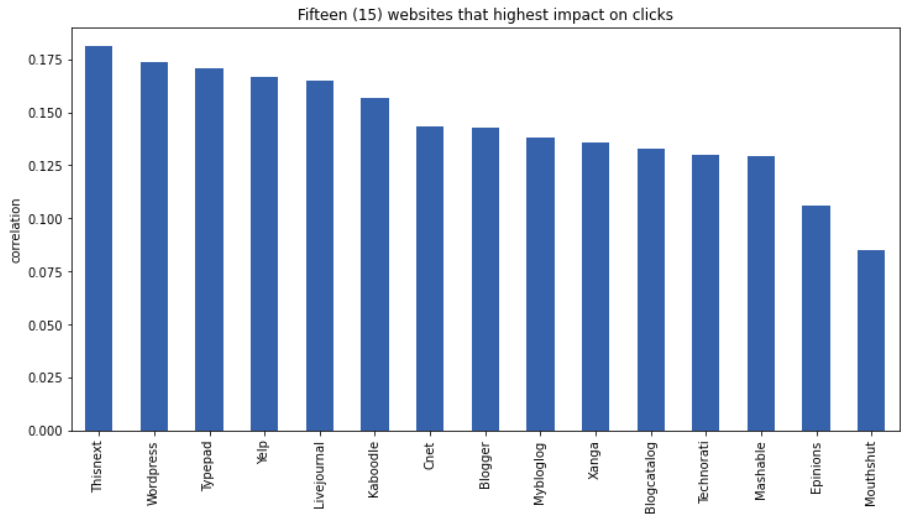
- Generally, more clicks are gotten at a higher engagement.
- When around 80 impressions are made on customers, They are a lot more likely to click compared to when less or more is made.
- Having more impressions than this, comes with a cost and may not have any marginal impact.
- Impressions above 140 are likely to yield no clicks as it shows that customer may be disinterested in what is advert.



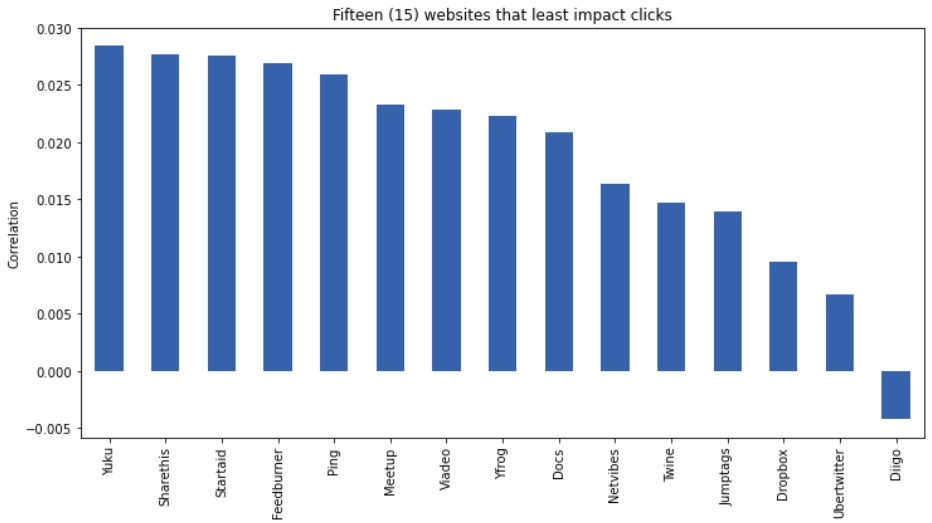
Exploratory data analysis Insight 5:
How website impacts clicks

websites impact click through rates differently. Using Correlation, It was able to determine how categories affect clicks obtained

Top websites that impact Clicks



Websites with least impact on Clicks



Top 15 websites in descending order: Thisnext, Wordpress, Typepad, Yelp, Livejournal, Kaboodle, Cnet, Blogger, Mybloglog, Xanga, Blogcatalog, Technorati, Mashable, Epinions, Mouthshut

Worst 15 websites in descending order: Yuku, Sharethis, Startaid, Feedburner, Ping, Meetup, Viadeo, Yfrog, Docs, netvibes, Twine, Jumptags, Dropbox, Ubertwitter, Diigo

Modelling

- A logistic regression was used to create a model for prediction of clicks.
- The model mathematically relates the number of impressions gotten from all websites to the outcome, 0/1 (Non-Click or Click)
- The model was trained to robustly to handle generalized data and is optimized to perform at an accuracy of 98%

Recommendation

- Focus on websites that have a high impact on click such as Thisnext, Wordpress, Typepad, Yelp and pay less attention to low impact websites
- Focus on website categories that have higher impact on clicks paying most attention on blogs and opinion sites
- Target users with around 80 impressions as more or less will be less effective
- A long-term strategy could include improving Click through rate by considering alternative adverts and experimentally determining the most effective
- Clicks are important but conversions are better. As some websites may present high click, low conversions may be gotten. On that note, it will be important to consider an analysis on websites that present the most conversions as this has longer term benefits.
- The trained model can be used to make predictions before embarking on marketing campaigns for more efficient use of resources
- As the behavior of customers and potential customers change with time, subsequent collation, visualization and modelling will be required. An automated system is recommended.