

ISOM3530 HW1**Due: 5pm 7 Mar**

- You can collaborate with your classmates for the assignments. If you work in groups, please list the names of your group members in the report. Submit by one member only.
- You need to submit both the report and the source code.

Marketing on Facebook is a great spot to develop your brand identity. It's a great space to keep customers informed, develop brand identity, and broaden your reach. Facebook offers its own form of advertising. The number of clicks may increase if we pay the advertising fee on Facebook.

You are given a set of Facebook data (see p.2). Conduct the following tasks:

Preliminary study

1. How many cases pay for advertisement fee to Facebook?
2. 'Type', 'Category', 'Post_time', 'Paid' may have impact on the number of clicks, visualize each factor on 'Click' by side-by-side boxplot. Comment on the results

Missing Value

1. Report which variables contain missing values and the total number of them for each variable
2. Fill in the missing values in 'Page_total_likes' by its mean value.
3. According to Q2 in preliminary study, use the most influential factor on 'Click', fill in the missing values of 'Click' by grouped means.

Feature engineering

1. "Post_day" indicates the day of a week. Classify the day into weekday and weekend may be a better way to observe the effect on number of click. Create a variable corresponds to this classification. Name it as "isWeekend".

Regressing modeling

1. Build a regression model using the details below:
 - The response variable is log("Click");
 - The predictors are "Type", "Page_total_likes", "Post_time", "Category", "Paid", "isWeekend".
 - Use the dataset with all the data preprocessing completed from above tasks
2. Comment on normality, constant variance assumption with graphical supports.
3. Give a short business insight from the regression model.

Clicks	Number of post clicks
Category	A. Action (special offers and contents) B. Product (direct advertisement, explicit brand content) C. Inspiration (non-explicit brand related content)
Type	Link, Photo, Status, Video
Post_day	Day of the week
Post_time	Day time, Night time
Paid	whether the company paid to Facebook for advertising or not
Page_total_likes	Number of people who have liked the company's page