



SMU

SINGAPORE MANAGEMENT
UNIVERSITY

ANLY482 - Analytics Practicum

AY2017-18T2

Project Proposal



Group Members:

Eric Yeo Pu Zhong | Tan Yong Siang Ivan | Tang Shing Hei

1. Overview of Company

TheSmartLocal is an independent media publisher that focuses on travel and lifestyle stories for Singaporeans. Most of their content are publish through TSL's website, YouTube channel, Facebook page and Instagram. The website comprises of editorial based and community based where members can upload their own reviews. Today, TheSmartLocal reaches over 3 million Singaporeans each month.

2. Business Problems & Motivation

Having published various content types over the years, TheSmartLocal knows what posts perform well and what doesn't. However, monitoring content engagement over different platforms such as Facebook, YouTube, Instagram and their blogs can be highly complex due to the nature of the posts such as timing, content type, authors. As such, they would like to perform a holistic, cross-platform analysis to quantify and investigate the causes of virality.

In addition, they learned that Facebook has recently changed ranking algorithms for posts in December 2017. Specifically, posts that encourage sharing and commenting to participate in competitions will be downgraded in rankings. TSL would like to investigate the impact of this change on their Facebook outreach, and if possible, provide recommendations.

3. Project Objectives

1. Gather all available data from TSL across platforms to evaluate possibility of cross-platform social media engagement analysis
2. Conduct preliminary analysis on all datasets to discover possible insights to deliver
3. Synthesize data from all platforms (Facebook, Blog, Instagram, YouTube)
4. Investigate impact of Facebook's algorithm change on TSL's post engagement rates

4. Data Collection & Description

TheSmartLocal will provide us with social media insights data from various platforms (Facebook, Instagram, YouTube, Blog) for the period of January 2017 - December 2017. As of the time of proposal submission, we have only received data for Facebook and YouTube.

4.1 Facebook Insights

Facebook Insights exports to .xls format, giving us the following data:

4.1.1 Facebook Insights Page Data

Metadata	Metadata Description
Date	No Description Required
Lifetime Total Likes	Lifetime: Total number of people who have liked your Page (unique users)
Daily New Likes	Daily: Number of new people who have liked your Page (unique users)
Daily Unlikes	Daily: Number of Unlikes of your Page (unique users)
Daily Page Engaged Users	Daily: Number of people who engaged with your Page. Engagement includes any click or story created. (unique users)
Weekly Page Engaged Users	Weekly: Number of people who engaged with your Page. Engagement includes any click or story created. (unique users)
28 Days Page Engaged Users	28 Days: Number of people who engaged with your Page. Engagement includes any click or story created. (unique users)
Daily Total Reach	Daily: Number of people who have seen any content associated with your Page (unique users)

Weekly Total Reach	Weekly: Number of people who have seen any content associated with your Page (unique users)
28 Days Total Reach	28 Days: Number of people who have seen any content associated with your Page (unique users)
Daily Organic Reach	Daily: Number of people who visited your Page, or saw your Page or one of its posts in News Feed or ticker. These can be people who have liked your Page and people who haven't. (unique users)
Weekly Organic Reach	Weekly: Number of people who visited your Page, or saw your Page or one of its posts in News Feed or ticker. These can be people who have liked your Page and people who haven't. (unique users)
28 Days Organic Reach	28 Days: Number of people who visited your Page, or saw your Page or one of its posts in News Feed or ticker. These can be people who have liked your Page and people who haven't. (unique users)
Daily Paid Reach	Daily: Number of people who saw a sponsored story or ad pointing to your Page (unique users)
Weekly Paid Reach	Weekly: Number of people who saw a sponsored story or ad pointing to your Page (unique users)
28 Days Paid Reach	28 Days: Number of people who saw a sponsored story or ad pointing to your Page (unique users)
Daily Viral Reach	Daily: Number of people who saw your Page or one of its posts from a story shared by a friend. These stories include liking your Page, posting to your Page's Timeline, liking, commenting on or sharing one of your Page posts, answering a question that you posted, responding to one of your events, mentioning your Page, tagging your Page in a photo or checking in at your location. (unique users)

Weekly Viral Reach	Weekly: Number of people who saw your Page or one of its posts from a story shared by a friend. These stories include liking your Page, posting to your Page's Timeline, liking, commenting on or sharing one of your Page posts, answering a question that you posted, responding to one of your events, mentioning your Page, tagging your Page in a photo or checking in at your location. (unique users)
28 Days Viral Reach	28 Days: Number of people who saw your Page or one of its posts from a story shared by a friend. These stories include liking your Page, posting to your Page's Timeline, liking, commenting on or sharing one of your Page posts, answering a question that you posted, responding to one of your events, mentioning your Page, tagging your Page in a photo or checking in at your location. (unique users)
...	...

4.1.2 Facebook Insights Data Export (Post Level)

Metadata	Metadata Description
Post ID	No Description Required
Permalink	No Description Required
Post Message	No Description Required
Type	No Description Required
Countries	No Description Required
Languages	No Description Required
Posted	No Description Required
Audience Targeting	No Description Required
Lifetime Post Total Reach	Lifetime: The total number of people your Page post was served to. (Unique Users)
Lifetime Post organic reach	Lifetime: The number of people who saw your Page post in news feed or ticker, or on your Page's timeline. (Unique Users)
Lifetime Post Paid	Lifetime: The number of people your advertised Page post was

Reach	served to. (Unique Users)
Lifetime Post Total Impressions	Lifetime: The number of impressions of your Page post. (Total Count)
Lifetime Post Organic Impressions	Lifetime: The number of impressions of your post in News Feed or ticker or on your Page's Timeline. (Total Count)
Lifetime Postpaid Impressions	Lifetime: The number of impressions of your Page post in an Ad or Sponsored Story. (Total Count)
Lifetime Engaged Users	Lifetime: The number of unique people who engaged in certain ways with your Page post, for example by commenting on, liking, sharing, or clicking upon particular elements of the post. (Unique Users)
Lifetime Post Consumers	Lifetime: The number of people who clicked anywhere in your post. (Unique Users)
Lifetime Post Consumptions	Lifetime: The number of clicks anywhere in your post. (Total Count)
Lifetime Negative feedback	Lifetime: The number of people who have given negative feedback to your post. (Unique Users)
Lifetime Negative Feedback from Users	Lifetime: The number of times people have given negative feedback to your post. (Total Count)
Lifetime Post Impressions by people who have liked your Page	Lifetime: The number of impressions of your Page post to people who have liked your Page. (Total Count)
Lifetime Post reach by people who like your Page	Lifetime: The number of people who saw your Page post because they've liked your Page (Unique Users)
Lifetime Postpaid Impressions by people who have liked your Page	Lifetime: The number of paid impressions of your Page post to people who have liked your Page. (Total Count)
Lifetime Paid reach of a post by people who like your Page	Lifetime: The number of people who like your Page and who saw your Page post in an ad or sponsored story. (Unique Users)
Lifetime People who have liked your Page and engaged with your post	Lifetime: The number of people who have liked your Page and clicked anywhere in your posts. (Unique Users)
Lifetime Organic views	Lifetime: Number of times your video was viewed to 95% of its

to 95%	length without any paid promotion. (Unique Users)
Lifetime Organic views to 95%	Lifetime: Number of times your video was viewed to 95% of its length without any paid promotion. (Total Count)
Lifetime Paid views to 95%	Lifetime: Number of times your video was viewed to 95% of its length after paid promotion. (Unique Users)
Lifetime Paid views to 95%	Lifetime: Number of times your video was viewed to 95% of its length after paid promotion. (Total Count)
Lifetime Organic Video Views	Lifetime: Number of times your video was viewed for more than 3 seconds without any paid promotion. (Unique Users)
Lifetime Organic Video Views	Lifetime: Number of times your video was viewed for more than 3 seconds without any paid promotion. (Total Count)
Lifetime Paid Video Views	Lifetime: Number of times your video was viewed more than 3 seconds after paid promotion. (Unique Users)
Lifetime Paid Video Views	Lifetime: Number of times your video was viewed more than 3 seconds after paid promotion. (Total Count)
Lifetime Average time video viewed	Lifetime: Average time video viewed (Total Count)
Lifetime Video length	Lifetime: Length of a video post (Total Count)

4.1.3. Facebook Insights Data Export (Video Level)

Metadata	Metadata Description
Post ID	No Description Required
Permalink	No Description Required
Post Message	No Description Required
Type	No Description Required
Countries	No Description Required
Languages	No Description Required

Posted	No Description Required
Audience targeting	No Description Required
Content tags	No Description Required
Custom labels	No Description Required
Is broadcast	No Description Required
Universal Video ID	No Description Required
Lifetime Post Total Impressions	Lifetime The number of impressions of your Page post. (Total Count)
Lifetime Post Total Reach	Lifetime The total number of people your Page post was served to. (Unique Users)
Lifetime Total Video Views	Lifetime Total number of times your video was viewed for more than 3 seconds. (Total Count)
Lifetime Unique Video Views	Lifetime Number of unique people who viewed your video for more than 3 seconds. (Unique Users)
Lifetime Total 30-second Views	Lifetime Total number of times your video was viewed for 30 seconds or viewed to the end, whichever came first. (Total Count)
Lifetime Unique 30-second Views	Lifetime Number of unique people who viewed your video for 30 seconds or to the end, whichever came first. (Unique Users)
Lifetime Total watches at 95%	Lifetime The number of times your video was watched at 95% of its length, including times when viewers skipped to this point. (Total Count)
Lifetime Unique watches at 95%	Lifetime The number of unique people who watched your video at 95% of its length, including people who skipped to this point. (Unique Users)

4.1.4. YouTube Insights (YouTube Watch Time)

Metadata		
Video title	Your YouTube Red rev-en-ue (USD)	Your es-tim-ated rev-en-ue (USD)
Video ID	You-Tube Red watch time (hours)	Your es-tim-ated ad rev-en-ue (USD)
Video length (minutes)	An-nota-tion clicks	Your es-tim-ated Ad-Sense rev-en-ue (USD)
Video created	Click-able an-nota-tions shown	Your es-tim-ated Double-Click rev-en-ue (USD)
Watch time (minutes)	Clicks per click-able an-nota-tion shown	Your trans-ac-tion rev-en-ue (USD)
Views	An-nota-tion closes	Trans-ac-tions
You-Tube Red watch time (minutes)	Clos-able an-nota-tions shown	Your rev-en-ue per trans-ac-tion (USD)
You-Tube Red views	Close rate	Es-tim-ated mon-et-ised play-backs
Av-er-age view dur-a-tion (minutes)	An-nota-tions shown	Play-back-based CPM (USD)
Card clicks	Likes	Ad im-pres-sions
Cards shown	Likes ad-ded	Videos ad-ded to playl-ists
Clicks per card shown	Likes re-moved	Videos re-moved from playl-ists
Card teas-er clicks	Dis-likes	Sub-scribers
Card teas-ers shown	Dis-likes ad-ded	Sub-scribers gained
Teas-er clicks per card teas-er shown	Dis-likes re-moved	Sub-scribers lost

End screen elements shown	Shares	Average percentage viewed
End screen element clicks	Comments	Watch time (hours)
Clicks per end screen element shown	Videos in playlists	
YouTube ad revenue (USD)	CPM (USD)	

4.1.5 YouTube Audience Retention Rate

Metadata		
Video title	Likes	Transactions
Video ID	Likes added	Your revenue per transaction (USD)
Video length (minutes)	Likes removed	Estimated monetised playbacks
Video created	Dislikes	Playback-based CPM (USD)
Watch time (minutes)	Dislikes added	Ad impressions
Average view duration (minutes)	Dislikes removed	Close rate
Average percentage viewed	Shares	Annotations shown
Card clicks	Comments	YouTube Red watch time (hours)
Cards shown	Videos in playlists	Annotation clicks
Clicks per card shown	Videos added to playlists	Clickable annotations shown
Card teaser clicks	Videos removed from	Clicks per clickable

	playlists	annotation shown
Card teasers shown	Subscribers	Annotation closes
Teaser clicks per card teaser shown	Subscribers gained	Closable annotations shown
End screen elements shown	Subscribers lost	Views
End screen element clicks	Your estimated AdSense revenue (USD)	Your estimated Double-Click revenue (USD)
Clicks per end screen element shown	Watch time (hours)	Your transaction revenue (USD)
Your estimated revenue (USD)	CPM (USD)	YouTube Red views
Your estimated ad revenue (USD)	YouTube ad revenue (USD)	
Your YouTube Red revenue (USD)	YouTube Red watch time (minutes)	

4.1.6. YouTube Traffic Sources

Metadata
Traffic source type
Watch time (minutes)
Views
YouTube Red watch time (minutes)
YouTube Red views

Av-er-age view dur-a-tion (minutes)
You-Tube Red watch time (hours)
Av-er-age per-cent-age viewed
Watch time (hours)

Instagram - to be received

Blog - to be received

5. Methodology

5.1. Data Collection

We will use the data provided to us by our project sponsor exported from FaceBook Insights, YouTube Analytics. The data comes in the format of Microsoft Excel.

5.2. Data Preparation

As the exported data comprises of multiple tabs and columns in multiple excel files, such as post level or video post, we will attempt to organize the data into consistent formats that are easier for our analysis especially when evaluating cross-platform social media engagement. We will also need to mask confidential information such as the company's staff names as required by the project sponsor.

5.3. Exploratory Data Analysis

We will examine the entries of the various social media platform for the same campaign. From here, we will be able to discover possible insights to deliver to our sponsors, for example to find out which social media platform is most suitable for video or blog posts.

5.4. Data Cleaning

To ensure accuracy of our model, we will identify missing values and outliers that are observed during the previous stage. We will go through these missing values separately and decide on how we should handle it (whether by replacing with the average value or simply remove the entire row). As for handling outliers, we will try attempt to analyze and come up with a reason for the outlier and see if it will affect our analysis.

5.5. Data Normalization and Transformation

To better cater the data to our needs, we will perform data transformation to transform some of the columns into rows and transforming between categorical and numerical variables so that we can better analyze the data. If the values in certain attributes varies too much, we will normalize these attributes to ensure that the analysis will be accurate.

5.6. Data Modeling (Steps 6-8)

We will develop an analytical software application if necessary, using software like SAS or JMP Pro.

6. Scope of Work

As we have not obtained the complete data and master data file from our project sponsor, we will only share our initial plans on how we intend to tackle the project at this initial stage. We intend to adopt the following steps in our analysis, adopting the approaches as highlighted earlier in the proposed methodology:

- I. Data Collection
- II. Data Preparation
- III. Exploratory Data Analysis
- IV. Data Cleaning
- V. Data Normalization and Transformation
- VI. Data Modeling
- VII. Re-train Prescriptive Model
- VIII. Recommendations and Insights

7. Project Schedule

Tasks		Week 0 (1/1 - 7/1)	Week 1 (8/1 - 14/1)	Week 2 (15/1 - 21/1)	Week 3 (22/1 - 28/1)	Week 4 (29/1 - 4/2)	Week 5 (5/2 - 11/2)	Week 6 (12/2 - 18/2)	Week 7 (19/2 - 25/2)
Project Preparation	Understanding Sponsor Requirements		✓						
	Create and Update Proposal		✓						
	Create and Update Wikipedia		✓						
	Project Acceptance								
Milestone 1: Project Proposal Submission									
Data Gathering & Scoping	Gather Data								
	Finalize Requirements with Sponsor								
	Scope Project								
Research & Preparation	Explore Software								
	Finalize Proposal								
Data Cleaning	Data Collection								
	Data Cleaning								
	Data Transformation								
Data Modelling	Exploratory Analysis								
Interim Preparation	Gather Feedback from Sponsor								
	Prepare Interim Report & Slides								
	Buffer								
Milestone 2: Interim Report & Presentation									

Legend	
	Planned
✓	Completed

8. Deliverables

The final deliverables of this project are:

- Project Proposal
- Midterm Report
- Midterm Presentation
- Final Report
- Final Presentation
- Project Wiki Page
- Project Poster

9. Stakeholders

Main Stakeholders of this project includes

Project Sponsor: Mushi Luke, Chief Technology Officer of The Smart Local

Project Supervisor: Prof. Kam Tin Seong, Associate Professor of Information Systems