



SMU

**SINGAPORE MANAGEMENT
UNIVERSITY**

IS482 Analytics Practicum

***Kolaveri Di* Social Analytics Project**

Project Proposal

Submitted on 10th September, 2014

Prepared by:

Chan Wei Yin

Lee Jaehyun

Introduction

“In a world where over two days of video get uploaded every minute, only that which is truly unique and unexpected can stand out in the way that [viral videos] have.”

- Kevin Allocca, Youtube Trends Manager

While a picture can tell a thousand words, a video can tell a story. Virality of videos has been increasing for the past few years, where more social media channels are tapped on such as Facebook, Instagram, and Twitter. For this project, we will be focusing on videos that went viral among the Twitter users, in particular “Why this *Kolaveri Di*” - a music video that went viral in 2011 on social networking sites for its Tamil and English lyrics. With a combination of both research work and hands-on data analysis, we aim to unveil the implications of social networks in marketing campaigns and feasible recommendations.

Objectives

The main objective of this project is to find out **why** and **how** it went viral. To further supplement the main objective, below is the list of questions that needed to be answered by the end of this project:

1. **Why** did the video go viral?
 - a) What are the factors that made *Kolaveri Di* go viral?
 - b) What are the other possible factors (or *triggers*) that can make anything (not just *Kolaveri Di*) go viral on social network sites?
 - c) What is the impact of social media on the virality of the video, in particular Twitter?

2. **What/Who** are the influencers?
 - a) Define the influencers.
 - b) Identify the influencers.
 - i. Who are the more *influential* influencers?
 - ii. Who are the more *powerful* influencers?
 - iii. What are their characteristics?
 - c) How do they influence their followers?
 - i. What kind of tweets do they make?
 - ii. What kind of keywords do they use?
 - d) Are there any commonalities among the influencers?

3. What implications does this project have?
 - a) How do we engage the influencers in social media marketing?
 - b) Make any other feasible recommendations for another video to be successful/go viral in Twitter.

Approaches to Solve the Problem

We will first attempt to understand and observe the situation in 2011 when the video “Why This *Kolaveri Di*” was first released. A thorough research work will be necessary to see the pattern of the viewership, the impact it has created on the online users, and the factors behind its unexpected success. After the preliminary research, hypotheses will be formed, which will be tested on with the analysis and exploration of the data (of tweets related to *Kolaveri Di*).

Thus, our group attempts to approach this project in two parts:

1. **Research-Driven:** a heavy part of the analysis will be rooted from research. Research work will be also used to form necessary hypotheses, which will be tested on through the actual data analysis. Some main research work would include:
 - a. Similar case studies done on Twitter analytics and network analytics
 - b. Factors that can make videos go viral on social media
 - c. Past academic papers or research papers that have been done on social media influences using Twitter (such as systematic framework and analytics framework)
2. **Data-driven:** given the actual dataset on the users and their tweets during a certain time period, the dataset will be explored and analyzed to test the hypotheses.

The outcomes of two methods will substantiate each other to come up with the implications and recommendations.

Dataset and Programs to be Used

Kolaveri Di.csv dataset will be analyzed. It contains attributes namely:

- ❖ ID
- ❖ AuthorName
- ❖ TweetContent
- ❖ TweetDate

Potential tools to be used:

- ❖ R or NodeXL
- ❖ SAS Enterprise Guide or SPSS for data exploration
- ❖ Enterprise Guide Text Mining

Expected Final Deliverables

At the end of the project, we aim to have achieved the following:

- ❖ The factors affecting the virality of the video
- ❖ Identification of the influencers and characteristics of influencers
- ❖ Key words that are used in the top tweets
- ❖ Marketing recommendations which will be substantiated with data analysis

Weekly Timeline

Week	Objectives and Tasks	Remark
1	<ul style="list-style-type: none"> ❖ Understanding of the details of the project with Prof Seema ❖ Studying the research materials given by Prof Seema ❖ Installation of R 	
2	<ul style="list-style-type: none"> ❖ Taking up R tutorial ❖ Meeting with Prof Srimi to discuss the requirements and expectations of the project ❖ Coming up with the rough structure (outline) of the project 	
3	<ul style="list-style-type: none"> ❖ Complete project proposal ❖ Look into NodeXL ❖ Decide R/NodeXL as analytics tool for this project 	Extended Project Proposal Deadline
4 (Project Proposal) & 5	<ul style="list-style-type: none"> ❖ Research on previous research cases on Twitter ❖ Research on relevant case studies and blog posts ❖ Understand the components of Twitter data ❖ Understand context of each column, explore dataset with NodeXL/R ❖ Learn and understand the components of graph and social network analysis ❖ Make any assumptions and hypotheses we may want to test through the analytics tool ❖ Explore and analyze the Twitter dataset 	10th Sept - Project proposal deadline
6	<ul style="list-style-type: none"> ❖ Prepare for presentation and progress report ❖ Revise weekly timeline 	
7 (Midterm)	<ul style="list-style-type: none"> ❖ Midterm presentation ❖ Adjust the direction of the project based on the feedback 	1st Oct - Midterm Presentation
8 & 9	<ul style="list-style-type: none"> ❖ Explore and analyze the Twitter dataset 	
10	<ul style="list-style-type: none"> ❖ Test out the hypothesis generated from our understanding of research cases and see if it's applicable on the Twitter dataset 	
11	<ul style="list-style-type: none"> ❖ Come up with feasible marketing recommendations and conclusion of the project ❖ Do final revision of the data analysis 	

12	❖ Prepare for presentation and the completion of report	
13	❖ Final presentation ❖ Revise final report based on feedback	
14	❖ Submission of the final report	