

YouTube Videos Folk Theories, Too Good to be True?

CS 6474 - Social Computing Project

Arvind Akpuram Srinivasan (arvind_s@gatech.edu)
Karthik Nama Anil (kanil3@gatech.edu)
Miasia J Jones (mjones386@gatech.edu)
Prithvi Alva Suresh (al.prithvi@gatech.edu)

Motivation

- YouTube is among the largest video sharing platforms
- Content creators want to **increase** the reach of their videos
- Curation algorithm is **opaque**
- **Folk theories** are developed about the curation algorithms and how to increase the popularity of the video
- **Efficacy** of these theories is not usually tested

Goal

To test the efficacy of popular folk theories relating to methods to increase the popularity of content on the video sharing platform YouTube.

Folk Theories on Social Media Feeds

A **folk theory** is an idea or assumption developed by ordinary people which is used to predict or explain a certain behavior, and does not involve a high level of technical knowledge.

“While the operation of [curation] algorithms is typically opaque to users, users often develop and sometimes share theories about how these curation algorithms work in order to plan their behavior”

First I “like” it, then I hide it: Folk Theories of Social Feeds - Motahhare Eslami et al. (2016)

Example: “I should always click ‘like’ on my own status updates in order to tell the algorithm to start sharing them”.

Related Work

- **I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System**
Meeyoung Cha, et al. (2007)
- **“Yes We Can”: How Online Viewership, Blog Discussion, Campaign Statements, and Mainstream Media Coverage Produced a Viral Video Phenomenon**
Kevin Wallsten (2010)
- **Virality over YouTube: an empirical analysis**
Gohar Feroz Khan, Sokha Vong (2014)
- **“I always assumed that I wasn't really that close to [her]”: Reasoning about invisible algorithms in the news feed**
Motahhare Eslami et al. (2015)

Data & Social Media Platform

Dataset for Analysis

Trending YouTube Video Statistics

Daily statistics for trending YouTube videos

(<https://www.kaggle.com/datasnaek/youtube-new>)

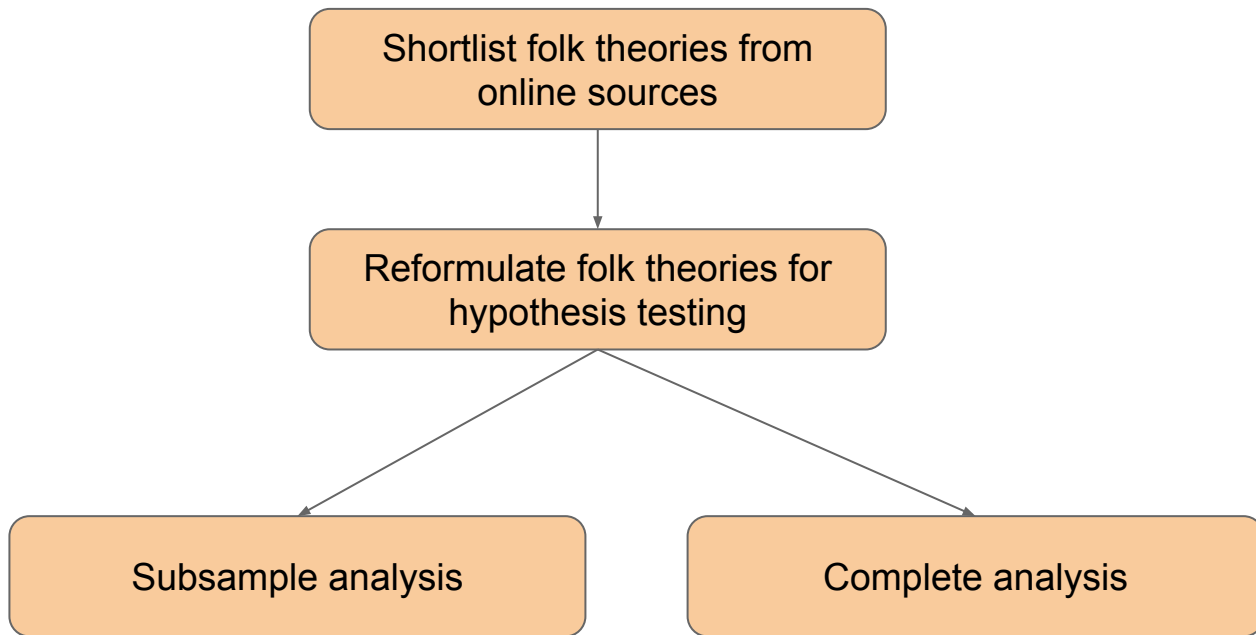
- 7 months of daily top 200 trending videos as per the data provided by YouTube Trending videos API.
- Our analysis restricted to USA.
- Dataset has metadata about videos like likes, comments, title, channel published, etc.



Folk Theories



Methodology



Methodology: Hypothesis Testing

Hypothesis testing refers to the process of inferring about the effect of a particular parameter/attribute.

All folk theories mentioned are treated as hypotheses and tested accordingly -

- Identify a test metric that can be used to assess the truth of the hypothesis.
- Test the hypothesis using the test metric.
- Determine if the hypothesis can be supported or rejected.

Complete Analysis

For the hypothesis where the attribute in question is present as a feature in the dataset and is not subjective -

- Analysed the complete dataset
- Visualised the results.

Example hypothesis -

- Significance of the day of the week for video release (analysis uses publishing date of the video)

Subsampled Analysis

For the hypothesis where the feature was either not present in the dataset or was subjective in nature.

Example hypothesis where this method was applied -

- Viral videos are rehearsed and not candid.

Selection of videos for qualitative analysis -

- Top 10 videos from each month based on view count
- Every video published after 1st December 2017
- If the video is viral during several months, consider it in the month of its last appearance

Hypothesis Testing

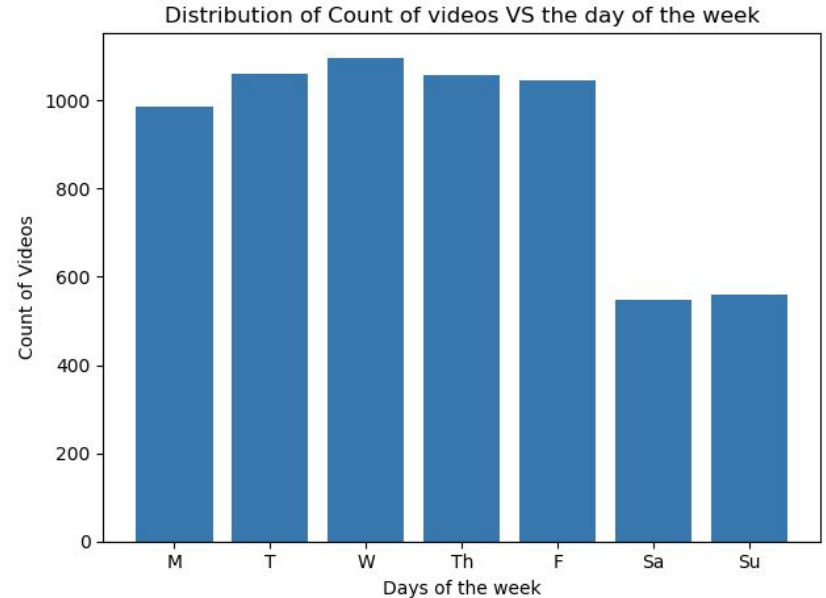


Complete Analysis

Day of The Week

Hypothesis -

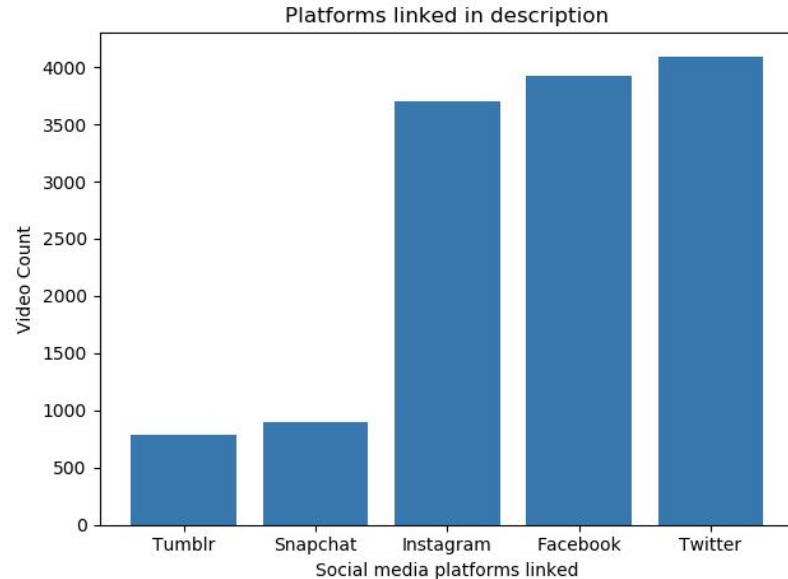
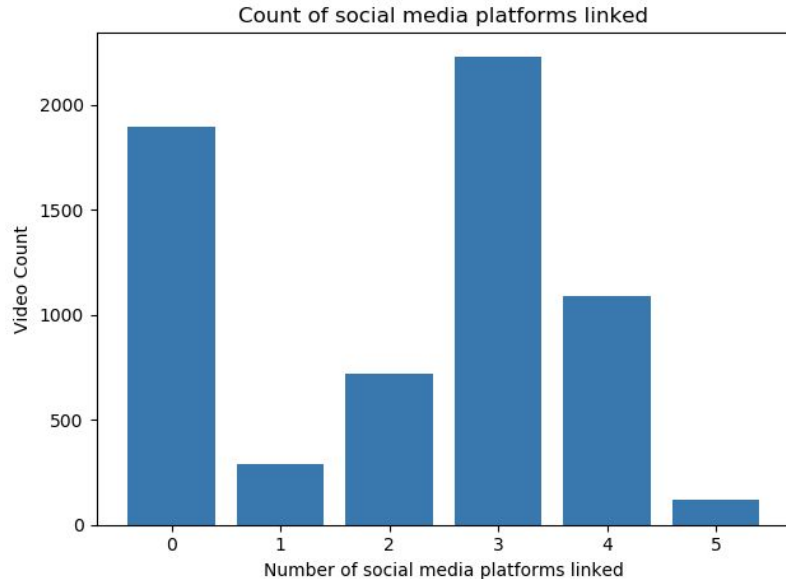
Viral videos are more likely to be published on a Monday or Tuesday(start of the week).



Use Multiple Social Media Platforms

Hypothesis -

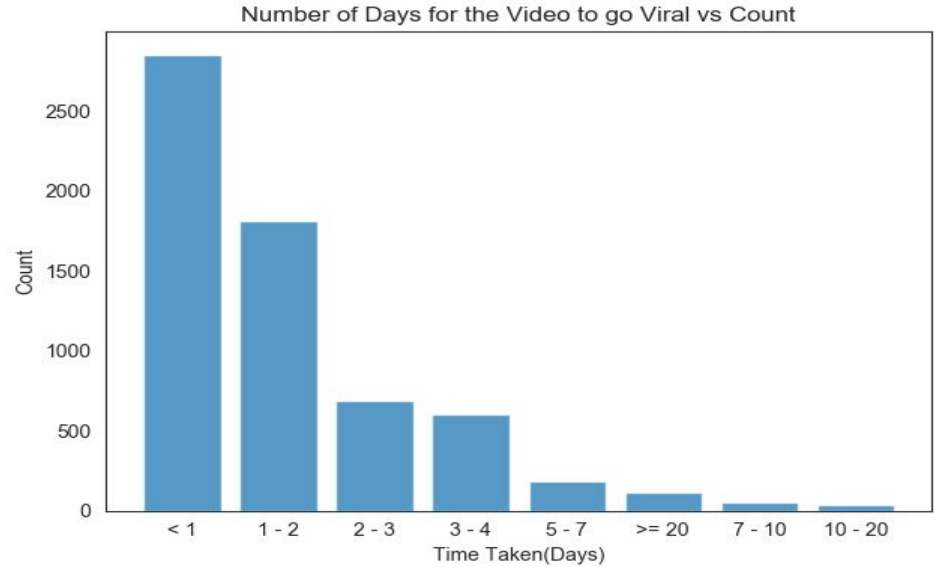
Video's chances of virality increases if the content creator links the video to their social media accounts.



Viral in a Day

Hypothesis -

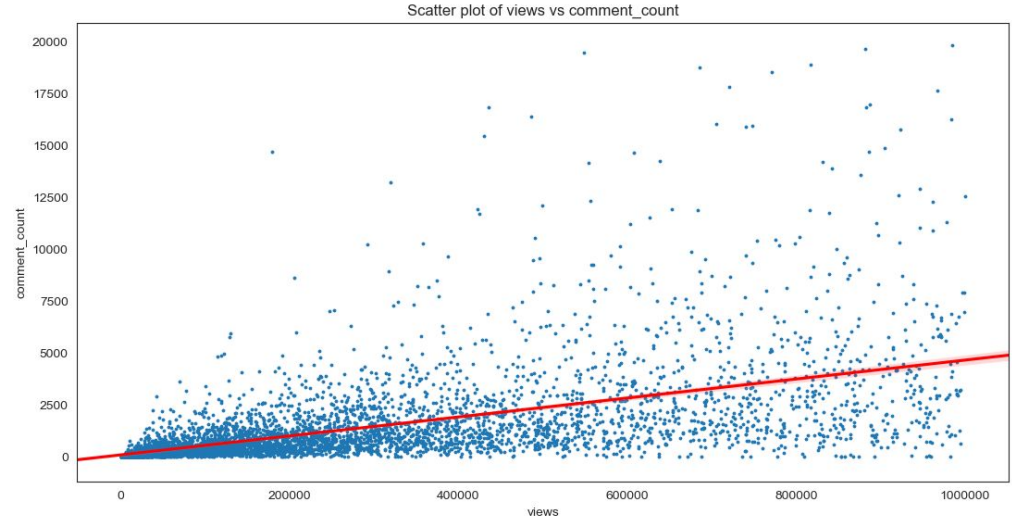
Videos go viral within a day of being published.



Comments vs Views

Hypothesis -

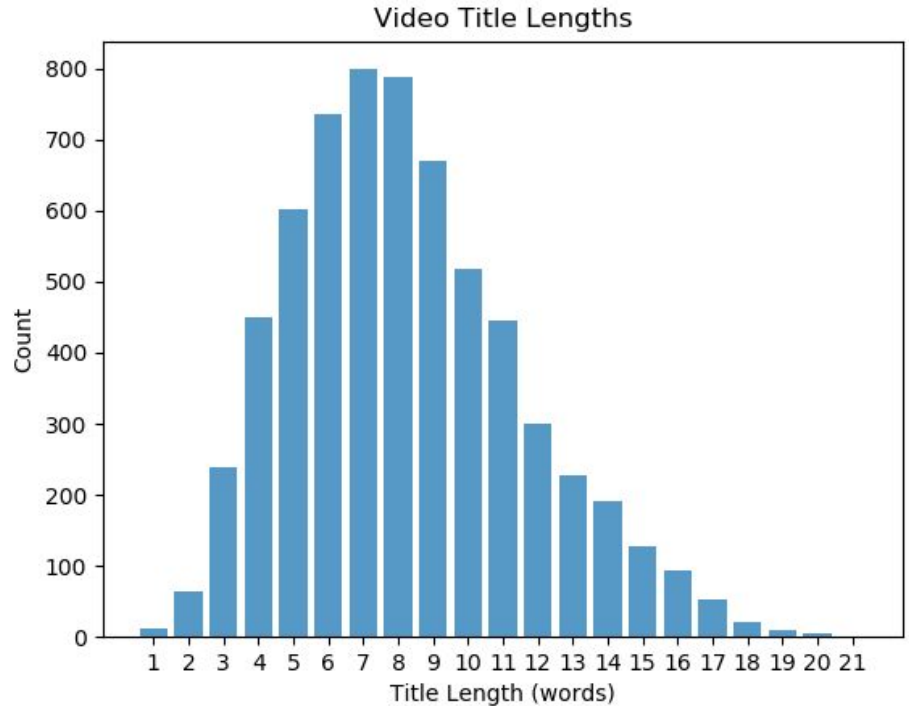
The number of comments and the number of views for a video exhibit high levels of correlation.



Short Video Titles

Hypothesis -

Length of the title is short (≤ 4 words).



Hypothesis Testing



Subsampled Analysis

Short Videos

Hypothesis -

Viral videos are short(≤ 300 seconds).

76%
(51/67)

Exception -



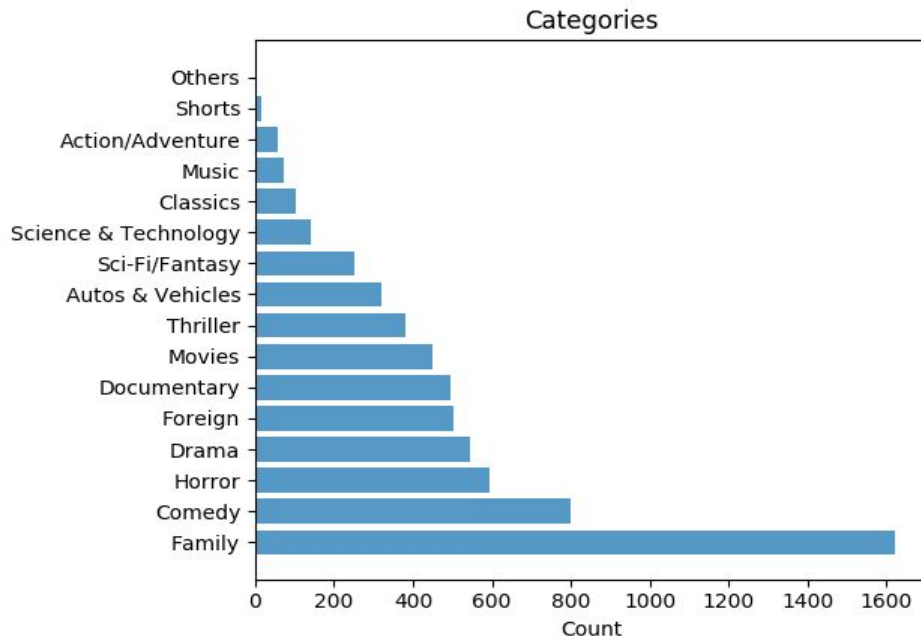
YouTube Rewind: The Shape of 2017 | [#YouTubeRewind](#)

Comedy is King ?

Hypothesis -

Comedy genre videos are more likely to go viral.

12%
(8/67)



Candid

Hypothesis -

Viral videos are rehearsed and not candid.

91%
(61/67)



Musical Elements

Hypothesis -

Viral videos have at least few musical elements embedded in the video(can be background music or the entire video can be a music video).

91%
(61/67)

Exception -



Do You Hear "Yanny" or "Laurel"? (SOLVED with SCIENCE)

No Country for The Old

Hypothesis -

Viral videos have majority of individuals in the video under the age of 35 years.

82%
(54/67)

Exception -



Jennifer Lopez - El Anillo (Official Video)

No I, You, We ...

Hypothesis -

Viral videos titles are written in third person.

96%
(64/67)

Exception -



To Our Daughter

No Lag or Distortion

Hypothesis -

Viral videos run smoothly on all platforms and devices(especially mobile phones).

100%
(67/67)



Entertaining & Engaging

Hypothesis -

Viral videos are generally engaging or entertaining.

94%
(63/67)

Exception -



#BeThe1To

Suicide: Be Here Tomorrow.

YouTube Channel is Key

Hypothesis

Viral videos are generally published by popular YouTube channels.

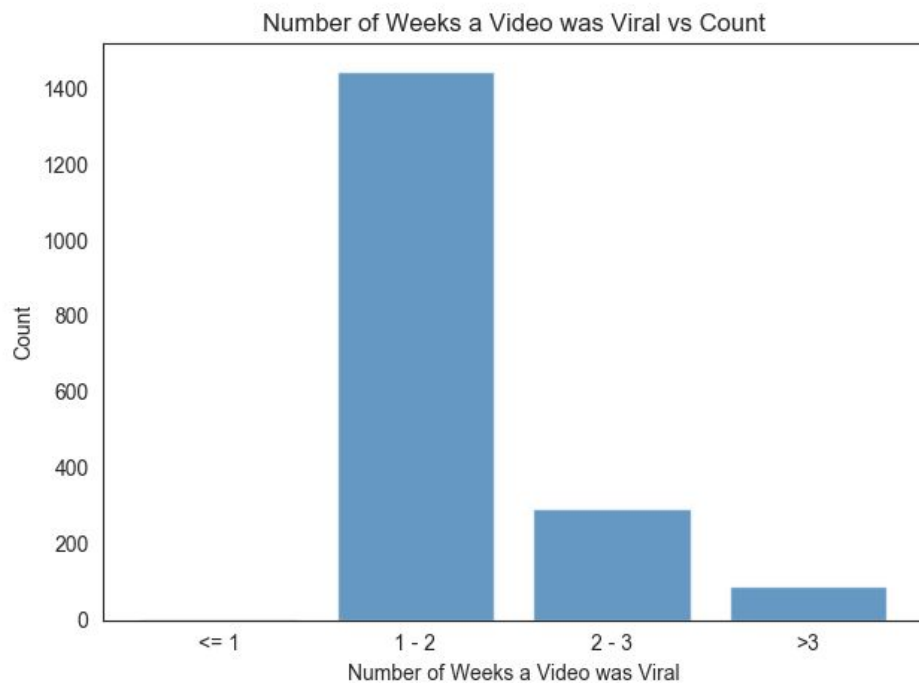
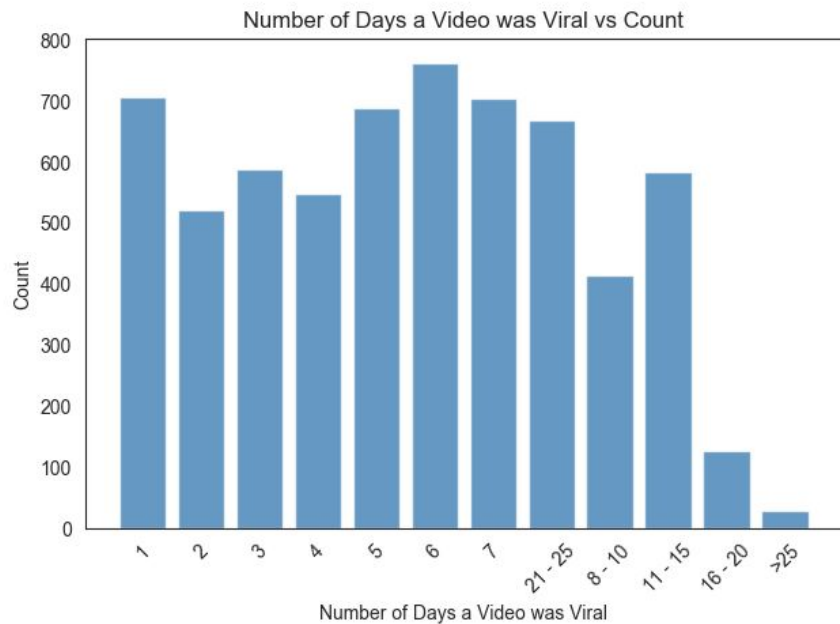
98%
(66/67)

Exception -



#WidenYourWorld #TurkishAirlines #SuperBowl
5 Senses with Dr. Oz - Turkish Airlines

Longevity of Video Virality

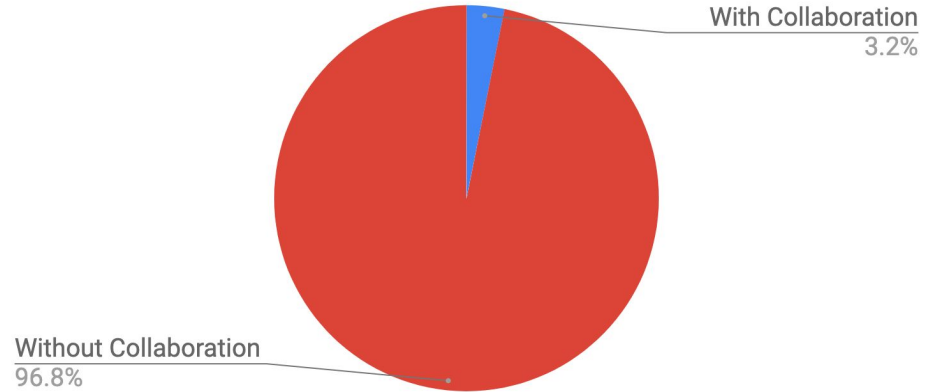


YouTuber Collaborations

Hypothesis -

Two or more YouTubers in a collaboration will bring together subscribers from each of their channels and the chances of the video going viral will be higher.

Collaboration Videos



Learnings

- Many of the popular folk theories that are present on the internet do indeed turn out to be true.
- Some of the theories do not play out in reality and therefore, these theories should not be blindly trusted
- Subsampled analysis can help provide interesting insights which can give more holistic picture of the phenomena at hand

Thanks

We are open to questions & suggestions