

UGCA FALL 2019

CLICKBAITED!

Ritika Mungal | Rutwick Bhawsar | Sharang
Wakankar | Shikha Singh | Shivang Arya



TABLE OF CONTENTS

- Problem Statement
- Approach and Workflow
- Findings
- Recommendations





Welcome to Bali | Travel Vlog | Priscilla Lee

Priscilla Lee • 3.6M views • 2 years ago

I had the chance to fly out to Bali with my whole family this Thanksgiving for our first ever trip abroad! I wanted to document my ...

PROBLEM STATEMENT

Can we quantify impact of the Thumbnail and the Title on the Views of the Travel video?

MAJOR POINTS

- Travel
- Title
- Thumbnail
- Views

OUR APPROACH AND WORKFLOW



DATA PREPARATION

DATA SCRAPING

Got the thumbnail, title and views of YouTube videos using YouTube API calls and beautiful soup.

DATA LABELS

Using Google Vision API, we got data labels for our thumbnails.

ENGAGEMENT METRIC

Generated a metric called Engagement using number of Views.

0 - If Views were greater than median views

1 - Else

DATA TRANSFORMATION

TF-IDF VECTORIZER

Combined the labels and the title and performed TF-IDF vectorization on the data

SENTIMENT SCORE

Appended the Sentiment score from the analysis to the data in order to enhance the performance

TRAIN-TEST SPLIT

As the name suggests, we split the data into training and testing in a 3:1 ratio.

DATA ANALYSIS

SENTIMENT ANALYSIS
LOGISTIC REGRESSION
TOPIC MODELLING AND
LATENT DIRICHLET ANALYSIS
RECOMMENDATIONS AND
SUGGESTIONS

LOGISTIC REGRESSION

Finding Impact of X variables. Various other models tested.

Findings:

- High negative sentiment score
- Negative Title leads to higher engagement.

	#Views	#Likes	#Dislikes	engagement	sentiment_score
#Views	1.000000	0.619144	0.779035	0.387379	-0.085882
#Likes	0.619144	1.000000	0.496856	0.400807	-0.089684
#Dislikes	0.779035	0.496856	1.000000	0.313101	-0.072817
engagement	0.387379	0.400807	0.313101	1.000000	-0.246056
sentiment_score	-0.085882	-0.089684	-0.072817	-0.246056	1.000000

```
model.coef_
array([[ 0.72611896,  0.          , -0.39045609, ... ,  0.25357534,
       1.08027513, -0.80293881]])
```

TOPIC MODELING AND LATENT DIRICHLET ANALYSIS

WHY TOPIC MODELING

Needed to explain the effect of thumbnail labels on engagement.

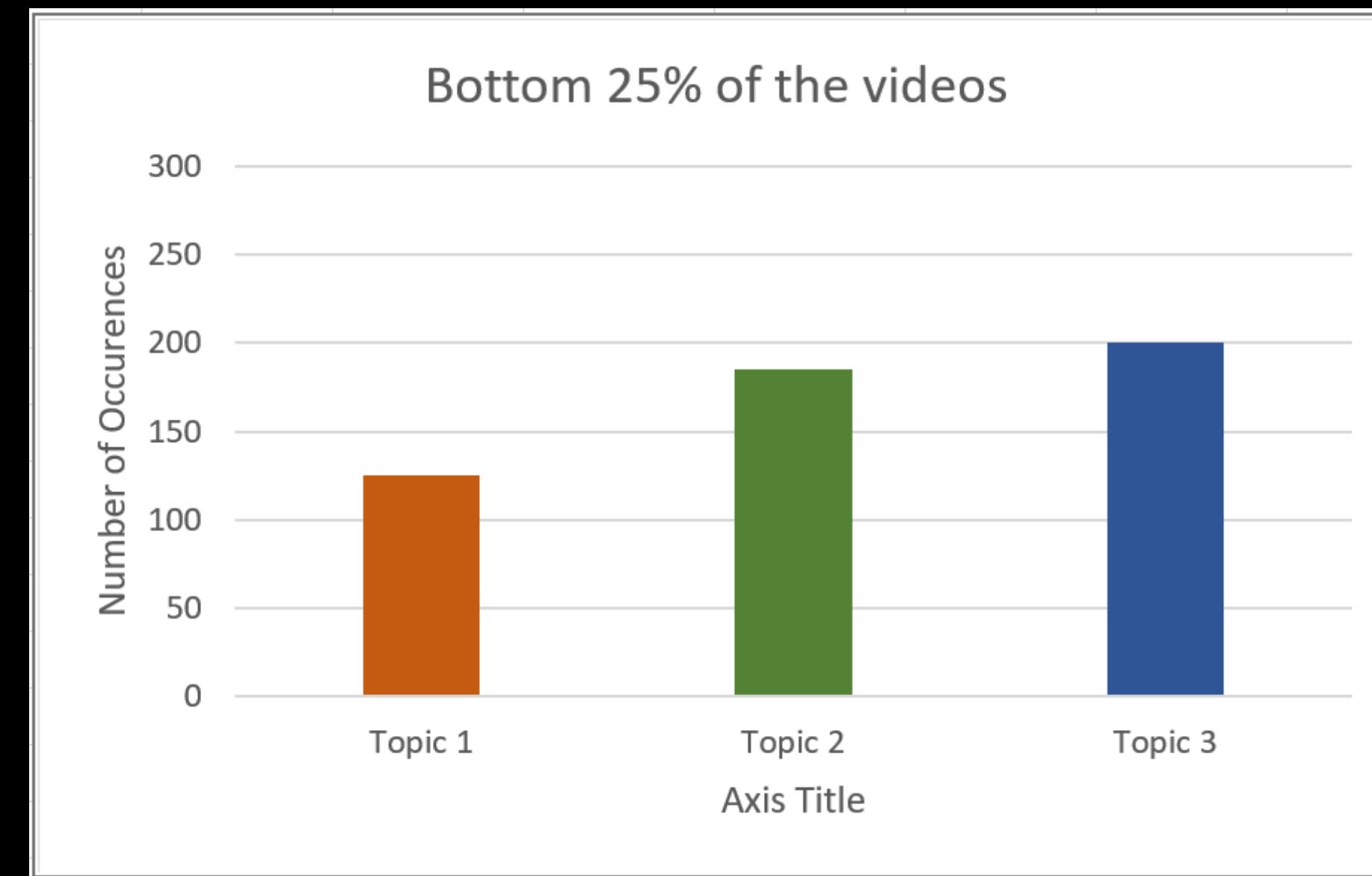
LDA MODEL

Split all the words from image labels and created a model with three topics.

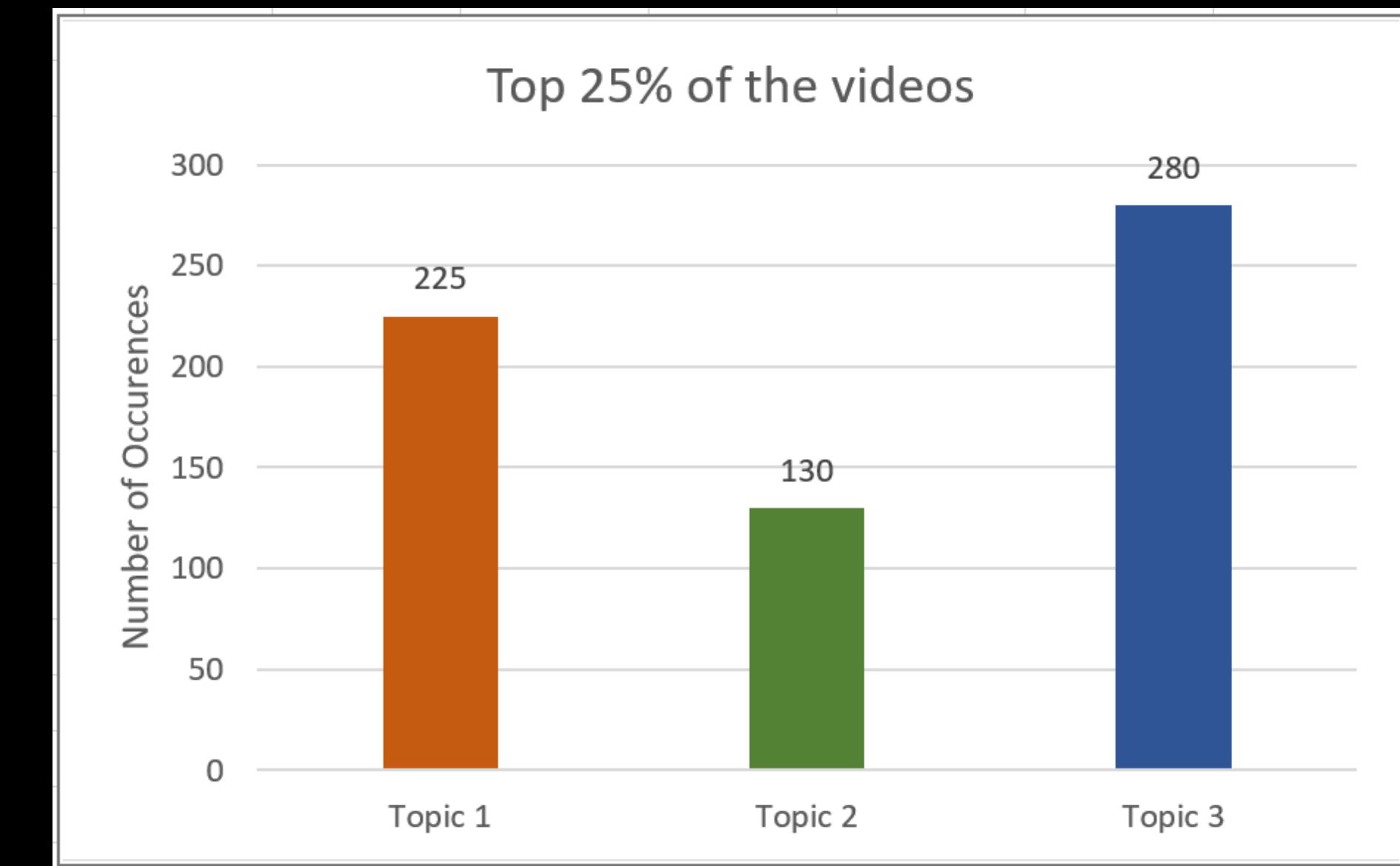
But, the number of terms associated with travel is limited. So to avoid overlapping, we decided to limit the amount of topics.

Assigned a topic with the maximum weight among all topics for each thumbnail image i.e. each video.

LOWEST ENGAGEMENT



HIGHEST ENGAGEMENT



RESULTS OF LDA

WHAT DO THESE TOPICS MEAN?



1M views • 3 months ago



Slime School Field Trip ! Pretend Teacher FAIL - New Toy School

Naiah and Elli • 4M views • 10 months ago

The slime school teacher Miss CrayCray decides to take the students on a field trip to shop for slime supplies. they find lots of ...



WHAT MAKES A GOOD THUMBNAIL?

TOPIC 1



Dubai Vacation Travel Guide | Expedia

Expedia 3.2M views • 5 years ago

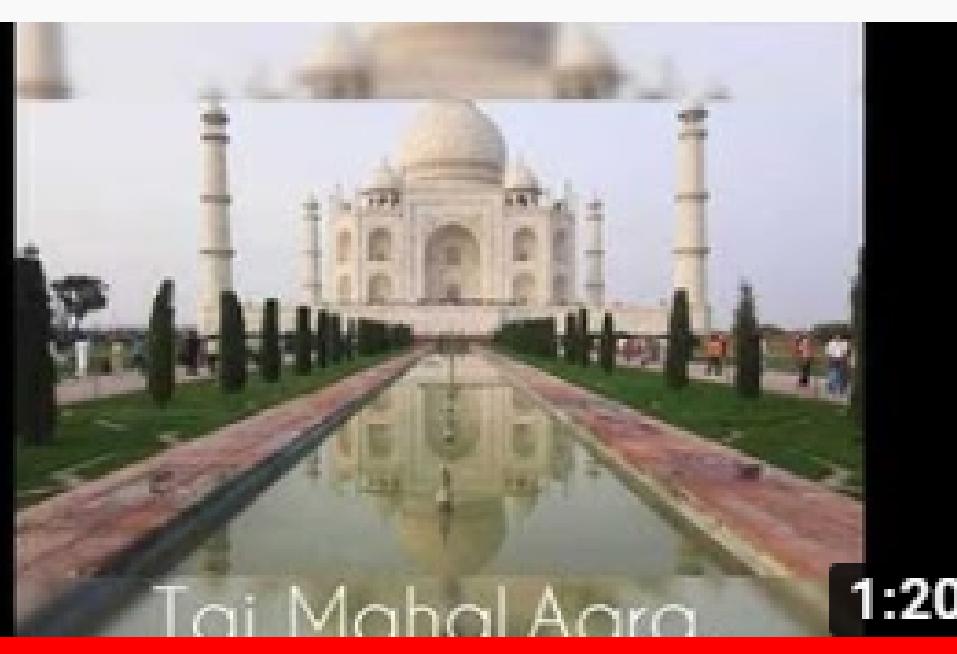
Dubai – a city that defies its extreme environment and redefines cityscape. Check out the footage of this incredible destination and ...

CC



WHAT MAKES A GOOD THUMBNAIL?

TOPIC 3



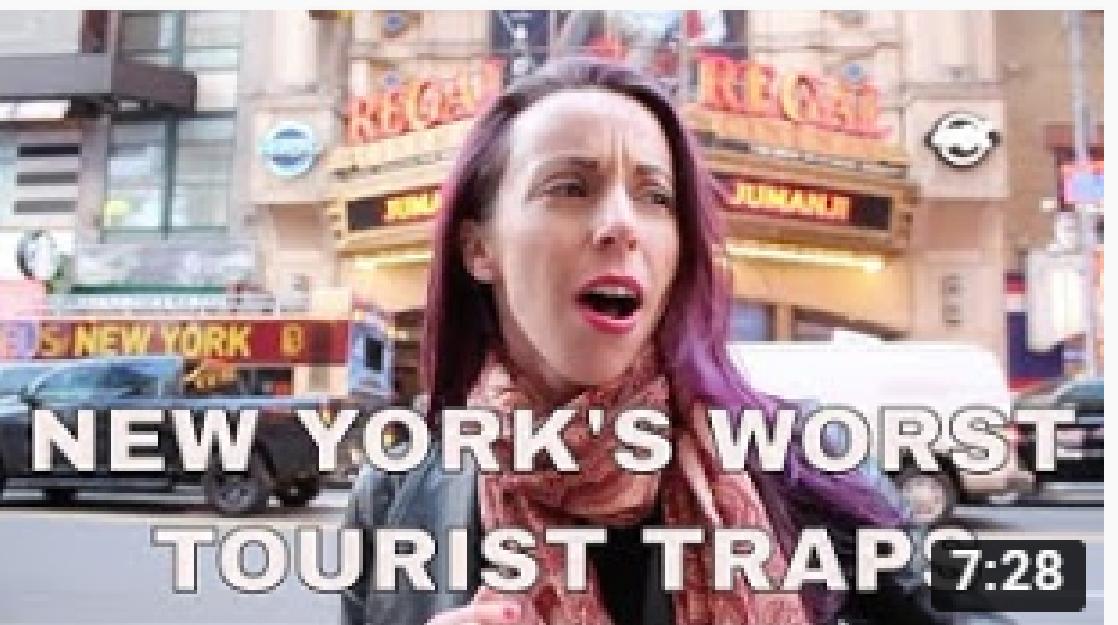
Top Travel Destination of india

SP Dude • 26 views • 3 years ago

Top Travel Destination of india you should visit at least once in your life... Are you looking to plan a
hassle free summer holidays in ...



WHAT MAKES A
BAD THUMBNAIL?
TOPIC 2



New York's Worst Tourist Traps: Times Square, Scams, Frauds, and More

Sarah Funk • 1.2M views • 1 year ago

New: "NYC GUIDE | CHINATOWN: Your PERFECT One-Day Itinerary" https://www.youtube.com/watch?v=4_wiBD6QbVg ---- BOOK A TOUR ...

CC

WHAT MAKES A GOOD TITLE?

NEGATIVE SENTIMENT

- According to our sentiment analysis and the results of the logistic regression we can safely conclude that a more negative title will get you higher engagement.
- 25% Higher Engagement!

RECIPE FOR A VIRAL VIDEO

- Include negative sentiment in the video titles.
- Choose Thumbnails with:
Slice of life and during the journey pictures or clear text about travel destination.
- Avoid Thumbnails with:
Typical tourist landmarks in urban or rural settings



Tasty 14M views • 1 year ago



THANK YOU!