ANALYSIS ON PODCAST REVIEWS

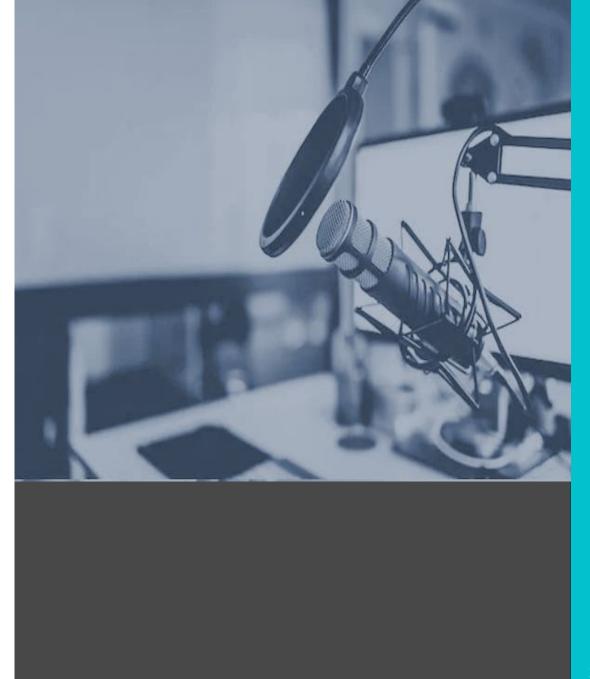
Group 8A: Site Bai, Kevin Cheung, Mira Daya, Ankit Jain, Yu Hsin (Kathy) Lee

Introduction

We chose to focus our analysis on podcast reviews and ratings from iTunes. Podcasts are becoming ever increasingly popular, ranging from a wide variety of topics such as the arts, society, business, career, religion, and so on.

We explore the business implications and the effects of reviews on various categories of podcast through text, sentiment, and causal analysis.

- 1. What factors (time, category, and sentiment) influence the rating and review content?
- 2. How have podcast reviews changed over time in terms of sentiment and rating?
- 3. How certain historic events happening at the time might have affected the number of reviews and the ratings of the podcast? Do these events have a causal impact on reviews and ratings?



Process Overview



Data

- Data Source: Apple Podcast Kaggle
- Datasets: Podcasts, Categories, Reviews
 - Number of Podcasts: 46665
 Number of Categories: 26
 Number of Reviews: 1162840

Cleaning:

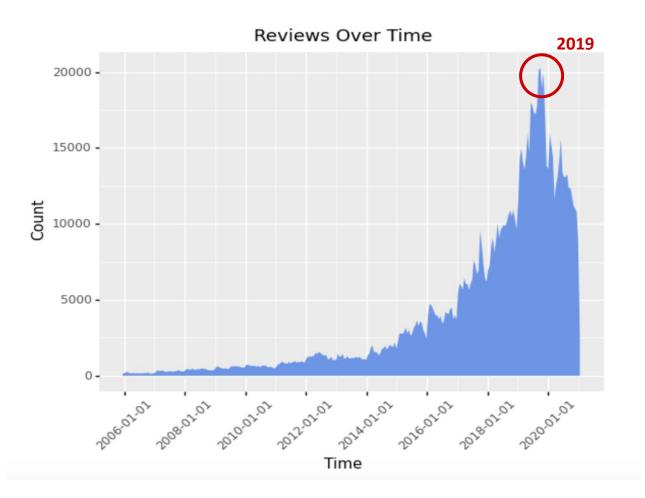
Checked for Missing Values in Rows

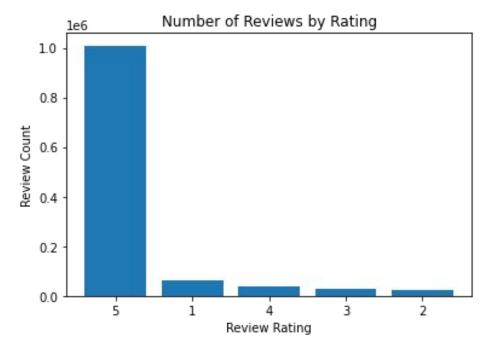


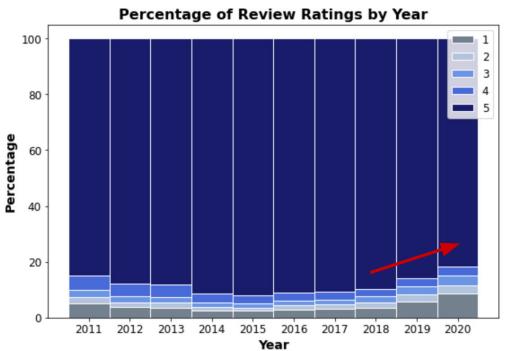
 Transformed Review Created At Variable to DateTime Object, and created columns for Year, Month, Weekday, Hour for Reviews Table

Data Tables	Attributes
Categories	podcast_id
	category
Podcasts	podcast_id
	itunes_id
	slug
	itunes_url
	title
Reviews	podcast_id
	Review title
	Review Description
	rating
	created_at

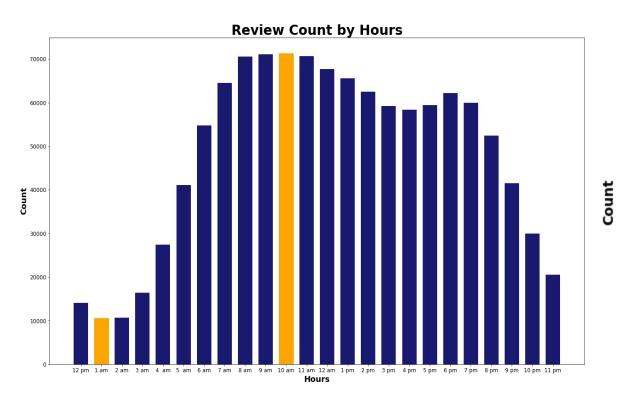
Exploratory Analysis

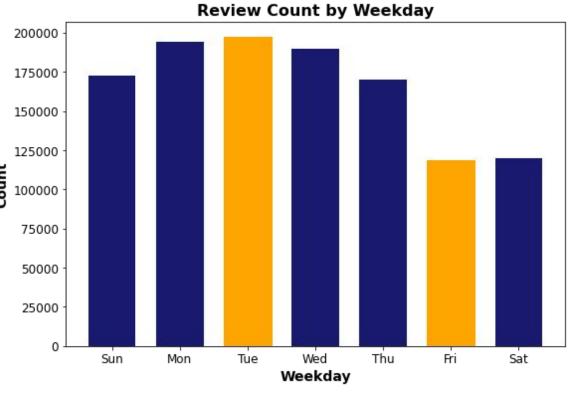






Exploratory Analysis



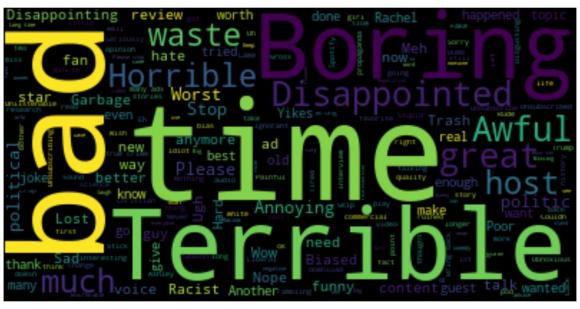




Text Modeling

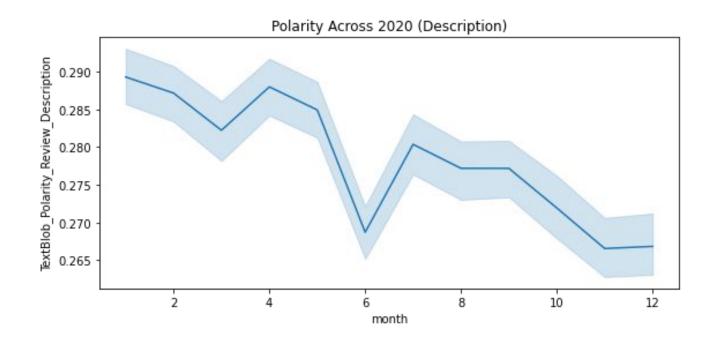
tfidf

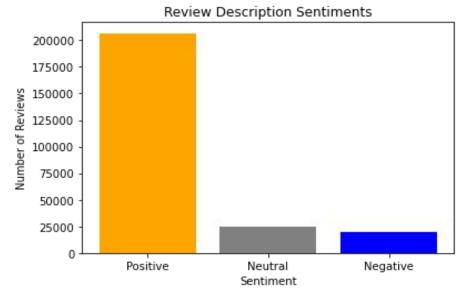
love podcast	3.699773
listen podcast	3.925193
look forward	4.324125
feel like	4.398882
podcast listen	4.732202
true crime	4.738565
love listen	4.760473
everi episod	4.805424
favorit podcast	4.812092
love show	4.835136
start listen	4.881524
great podcast	4.896952
highli recommend	4.916428
everi week	5.073896

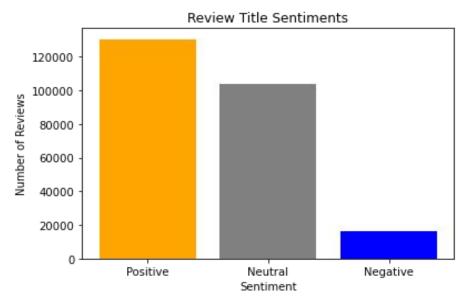




Sentiment Analysis



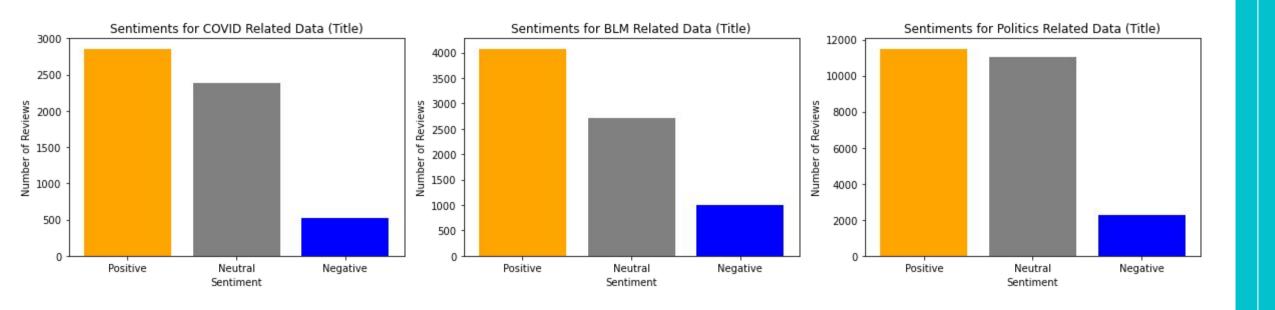




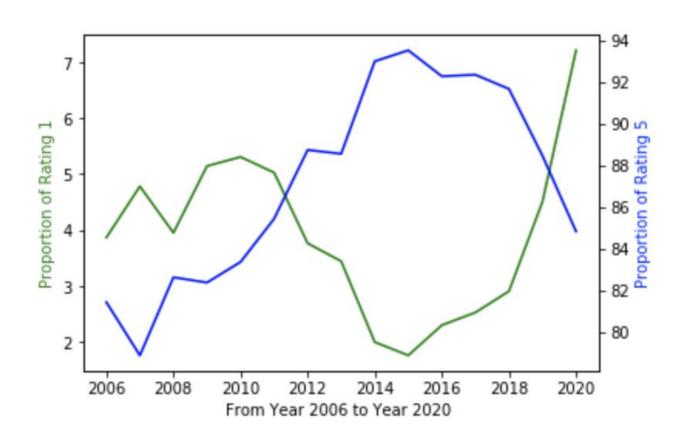
Sentiment Analysis



Sentiment Analysis: Historic Events

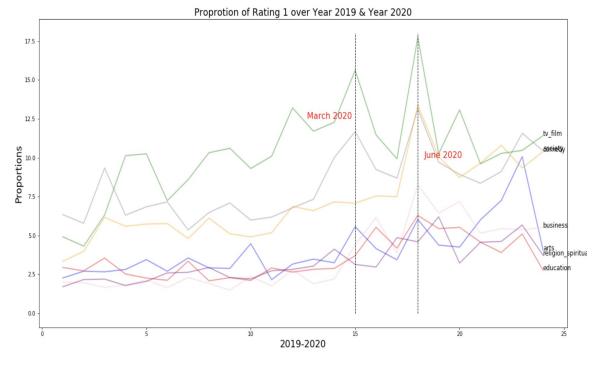


Sentiment Analysis: Historic Events

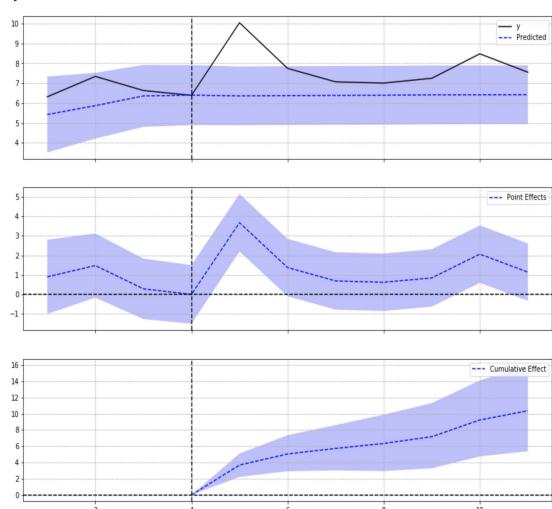


Causal Impact Analysis: 2020 Historic Event

Effect of Protests in June 2020 on Proportion of 1 Star Ratings

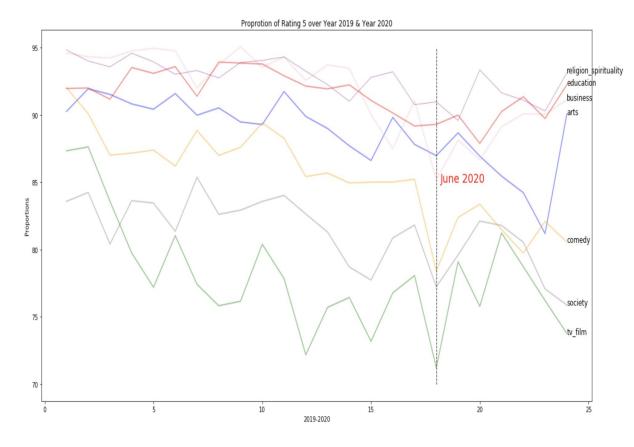


The probability of obtaining this effect by chance is very small (Bayesian one-sided tail-area probability p = 0.0). This means the causal effect can be considered statistically significant.

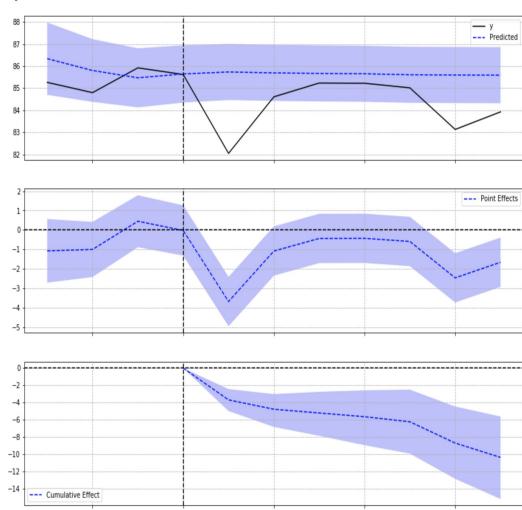


Causal Impact Analysis: 2020 Historic Event

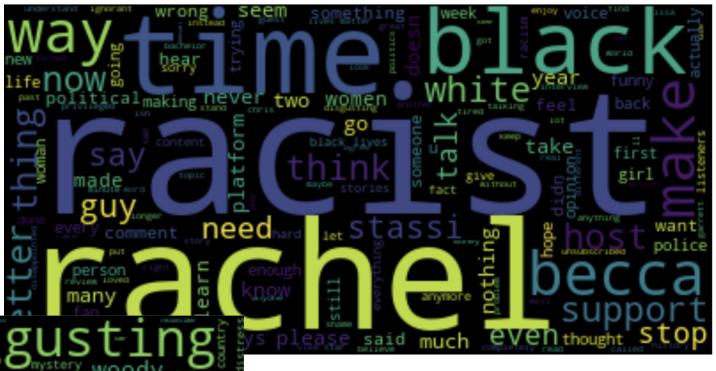
Effect of Protests in June 2020 on Proportion of 5 Star Ratings



The probability of obtaining this effect by chance is very small (Bayesian one-sided tail-area probability p=0.0). This means the causal effect can be considered statistically significant.

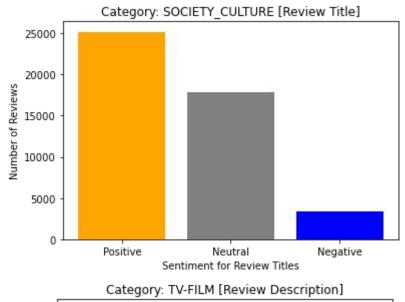


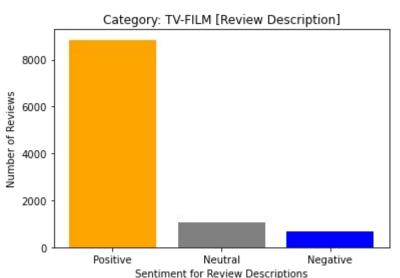
June Keywords

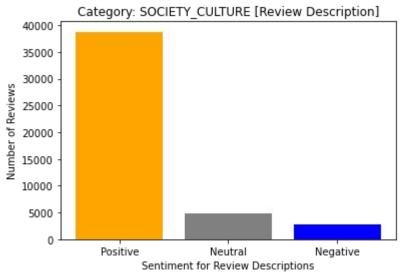


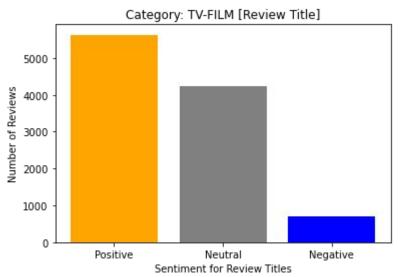


Sentiment for Different Categories

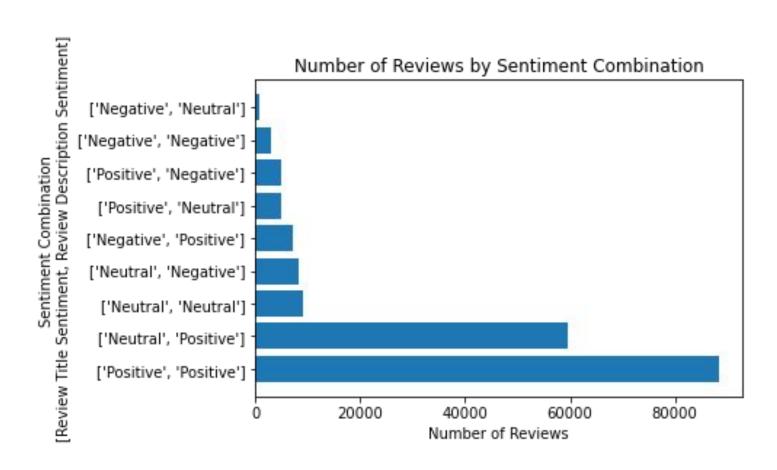


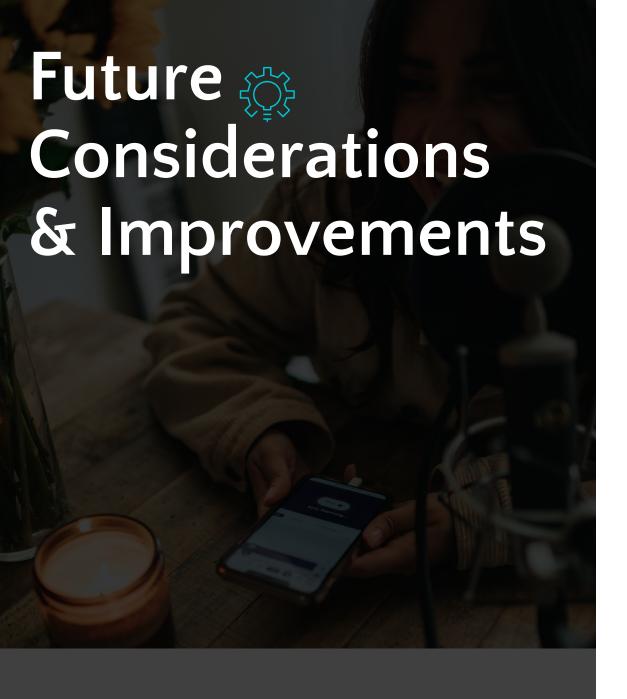






Sentiment for Title and Descriptions





Social Network Analysis with Listener Data

- Understanding the seasonality behind sentiment and review ratings
- Further understanding of causal impact of historic events and possibly predicting what podcasts should do in case of such events
- Establishing relationships between popularity of a podcast and reviews

Key Takeaways

Be aware while talking about any possible historic events happening at the time because it can have extreme effects on ratings

Listeners are more critical during specific time periods in a year

 Plan content efficiently during specific seasons

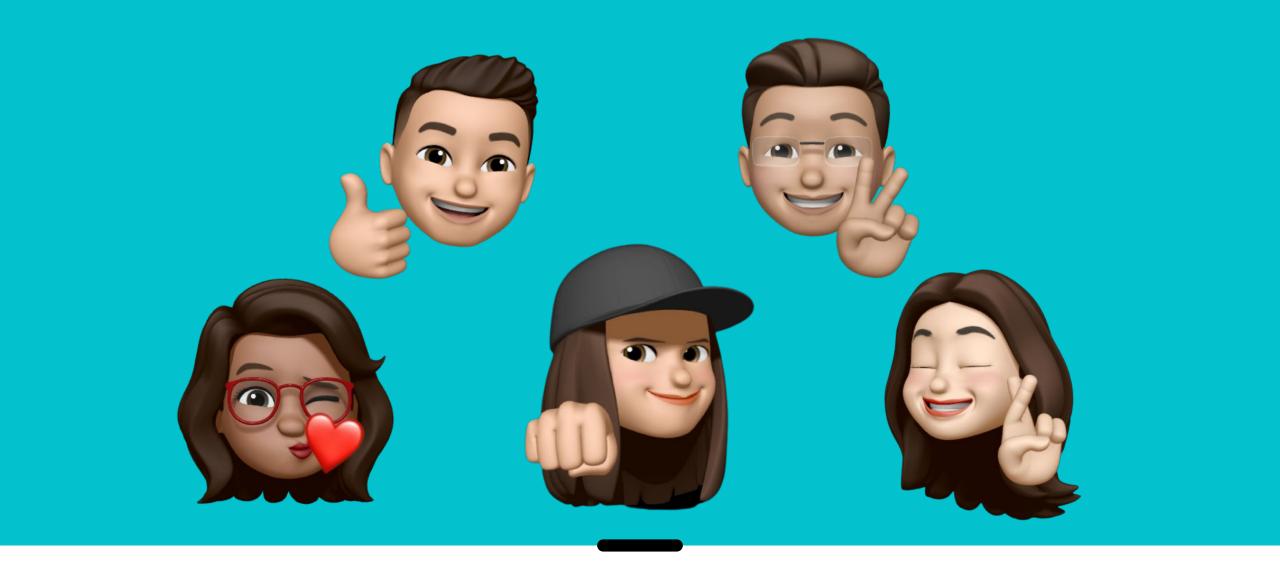


Apple Podcasts can leverage on the fact that they are the only platform that allows for ratings and reviews of podcasts.

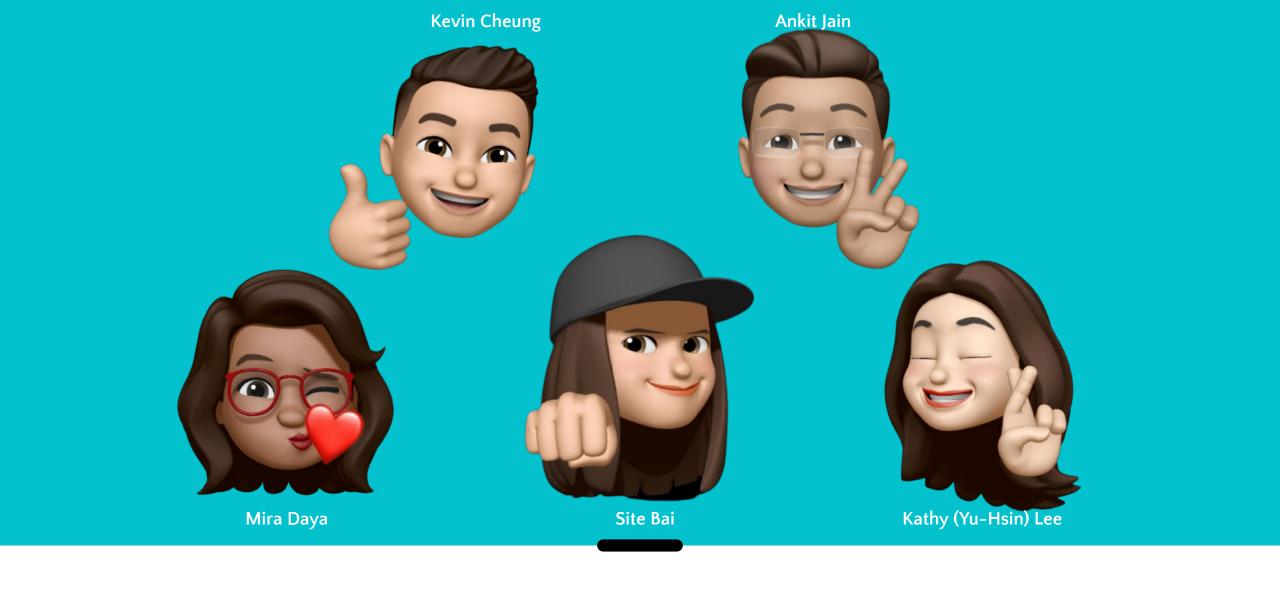
The best time for advertisements and podcast releases seems to be between early weekdays during commuting hours to get the most feedback and reviews

Implementation





THANK YOU!



Apple Podcast Analysis