



University of New Haven

Team: Data Ops

Team Members:

- Team Lead & Data Scientist: Naga Swetha Nallaninnall10@unh.newhaven.edu
- Data Engineer: Govardhan Namala – gnama2@unh.newhaven.edu
- Data analyst: Amani Ponnamm – aponn6@unh.newhaven.edu

Research Question:

How do consumer sentiments in product reviews influence brand loyalty and perception of Apple's product ecosystem?

The Dataset we chose is the Apple Product Review Dataset which is available in Kaggle at this [link](#)

Exploration Techniques Used:

The following exploration techniques from the Data Mining Map were applied:

- Data Cleaning: To normalize the data for analysis, we removed special characters, converted text to lowercase, and eliminated stopwords.
- Summary Statistics: We calculated the distribution of reviews to understand consumer sentiment patterns.
- Text Mining: Extracted keywords using WordClouds to identify common themes in positive, negative, and neutral reviews.
- Correlation Analysis: Explored the relationship between product mentions and sentiment to assess how specific Apple products are perceived.
- Visualization: Used pie charts, bar graphs, and WordClouds to present the data insights

Data Exploration Findings:

1. Sentiment Distribution:

- 51.5% of the reviews were neutral, 32.0% were negative, and 16.5% were positive.
- This suggests that while Apple products receive many neutral reactions, a significant proportion of reviews are negative, indicating areas for improvement.

- Visualization: *Pie chart* showing the proportion of neutral, negative, and positive reviews.

2. Keyword Analysis with WordClouds:

- Positive Reviews:
 - Common words include “apple,” “thank,” “great,” “iphone,” and “love.” These words suggest that customers appreciate product quality, functionality, and updates.
- Negative Reviews:
 - Words such as “fix,” “icloud,” “update,” “time,” and “issue” point to frustrations with software bugs, updates, and service issues.
- Neutral Reviews:
 - Neutral reviews frequently mention “apple,” “iphone,” “siri,” and “ipod,” indicating general mentions without strong positive or negative sentiment.
- Visualization: *WordClouds* for positive, negative, and neutral reviews showing frequently used keywords.

Product Mentions:

- iPhone is the most frequently mentioned product, reflecting its central role in customer discussions.
- MacBooks and iPads are also discussed, though less frequently, with more favorable sentiments compared to the iPhone.
- Visualization: *Bar chart* showing the number of mentions for iPhone, MacBook, and iPad in the reviews.

GitHub repository link:

A GitHub repository has been created for this final project and all the source code used for this phase -4 is uploaded in a different branch named Phase-4

You can access the repository at this [link](#). The code can be seen in the said branch.