**Email Campaigns**

**Problem Statement:**

The current email marketing strategy faces challenges in engaging recipients, targeting diverse audience segments, ensuring content relevance, and optimizing automation. Low engagement rates and a need for improved segmentation and personalization are key concerns.

**About dataset:**

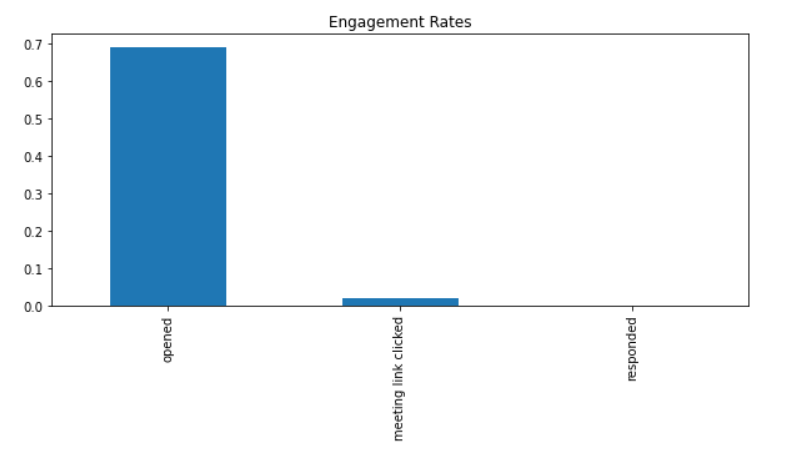
The Dataset consists of :-

'example1', 'marketingAnalytics0', 'HRConsultingSeries', 'marketingAnalyticsSeries', 'MarketingAnalyticsSeries', 'series\_legal', 'IT\_Solutions\_Series', 'Marketing\_Analytics\_Series', 'series\_marketing\_analytics', 'series1', 'HR\_Consulting\_Series', 'financial\_advisory\_series', 'series\_IT\_Solutions', 'Series1\_HR\_Consulting', 'Series\_IT\_Solutions', 'email\_series\_marketing\_analytics', 'legal\_services0'

The most of the columns except the example1 where have majority of null values so I have combined all the columns into a single column which is called as ‘**email\_campaigns**’

**Feature Insights (Before Pre-Processing) :-**

* 68% of the email messages are **opened**
* 0.3% of the email messages are **Meeting link clicked**
* There is no **responses**

****

**Pre-processing Techniques: -**

The pre-processing techniques applied to the text involve several steps to enhance the quality of the data for analysis. These steps include:

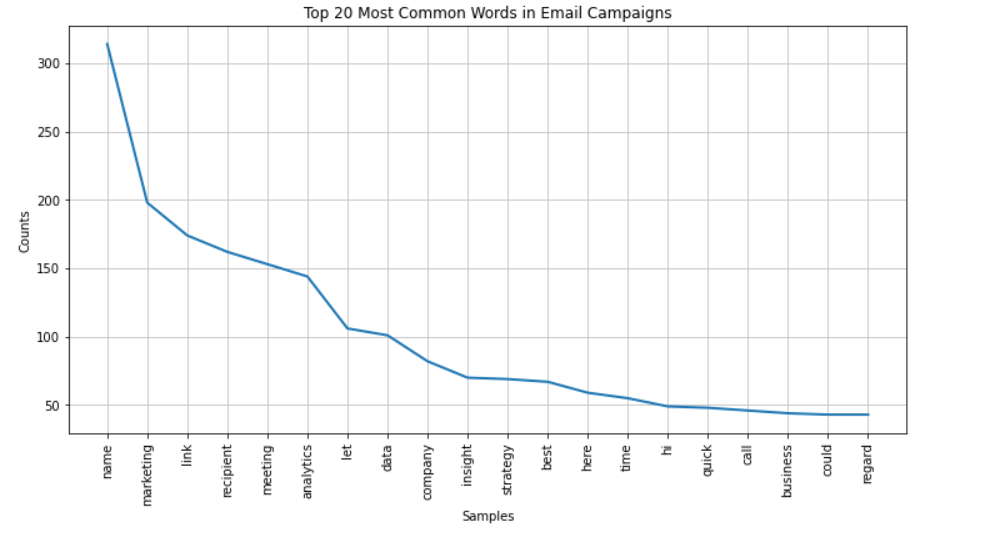
* **Lowercasing**: Converting all text to lowercase ensures uniformity and simplifies subsequent analysis by treating words in a case-insensitive manner.
* **Special Character** and Number Removal: Removing special characters and numerical values eliminates noise from the text, focusing on the linguistic content.
* **Tokenization**: Tokenization breaks the text into individual words or tokens, facilitating the analysis of word-level patterns.
* **Stop word Removal**: Eliminating common stop words, such as "and," "the," or "is," helps reduce noise in the data by focusing on more meaningful terms.
* **Lemmatization**: Lemmatization reduces words to their base or root form, ensuring that different inflections or derivations of a word are treated as the same term.
* **Joining Tokens**: The final step involves reassembling the processed tokens into a coherent text, ready for further analysis.

**EDA**

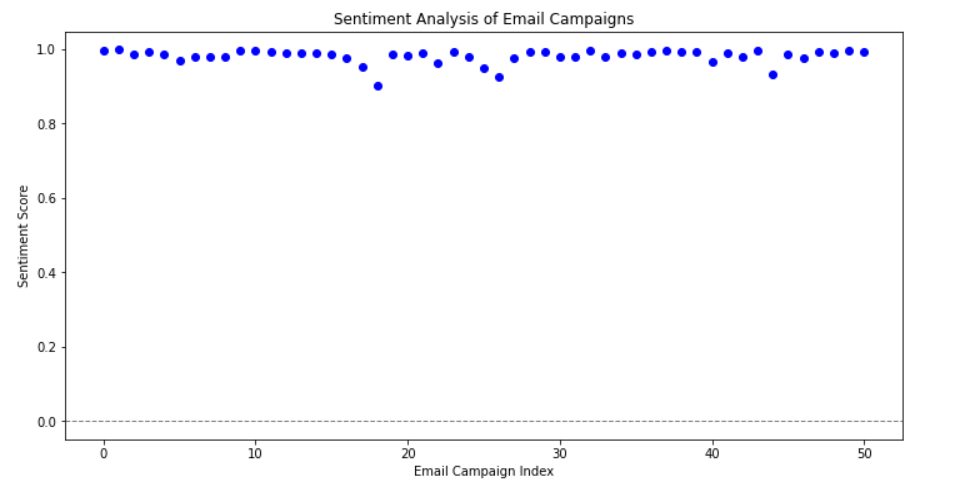
* The **word cloud** visually represents the prominent terms utilized in the emails, highlighting recurring themes such as 'Marketing,' 'Meeting,' 'Recipient,' and others.



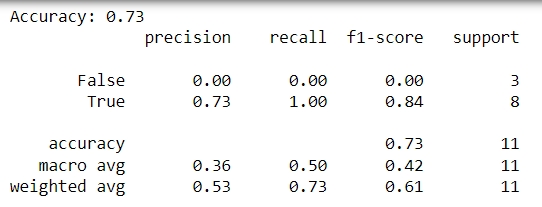
* The **word frequency** analysis reveals the top 20 words that frequently appear in the emails.



* The **Sentiment Analysis** for the email campaigns Column .



**Model Training:**

****

**Conclusion:**

**Accuracy (Overall Performance):**

Overall accuracy is 73%, meaning the model's predictions are correct 73% of the time. Precision, Recall, F1-Score (Performance Metrics for each Class):

**True Class (Opened):**

Precision (Accuracy of positive predictions): 73% Recall (Ability to capture all positive instances): 100% F1-Score (Balance between precision and recall): 84% Support (Number of actual occurrences): 8 instances False Class (Not Opened):

Precision: 0% (indicates no correct positive predictions for this class) Recall: 0% (indicates the model misses all actual instances of this class) F1-Score: 0% Support: 3 instances Analysis and Recommendations:

While the model performs well in predicting opened emails, it struggles with emails that are not opened. Consider analyzing misclassifications to identify patterns or issues. Experiment with different models, hyperparameters, or feature engineering to improve performance. Address class imbalances in the data, if present.