Email Marketing Campaign Success Prediction - Business Objective and Dataset Description

# Business Objective

The objective of this project is to predict whether a customer will open an email in an email marketing campaign. By building a classification model, the marketing team can identify high-potential customers, segment audiences, and tailor campaigns to optimize engagement.   
Predicting the open rate can help improve targeting, reduce marketing costs, and enhance overall campaign performance.

# Dataset Variable Descriptions

Numerical Variables:

### Customer\_Age

Type: Integer  
Description: Age of the customer in years, generated with a mean of 35 and a standard deviation of 10.

### Emails\_Opened

Type: Integer  
Description: Number of emails the customer has opened in the past, with an average of 5 and a standard deviation of 2.

### Emails\_Clicked

Type: Integer  
Description: Number of emails the customer clicked on after opening them, with an average of 2 and a standard deviation of 1.

### Purchase\_History

Type: Float  
Description: Total amount of money (in dollars) the customer has spent in previous purchases. The data is normally distributed with a mean of $1500 and a standard deviation of $500.

### Time\_Spent\_On\_Website

Type: Float  
Description: Average time (in minutes) the customer spends on the website after receiving marketing emails, with an average of 5 minutes and a standard deviation of 2.

### Days\_Since\_Last\_Open

Type: Integer  
Description: Number of days since the customer last opened an email, with an average of 30 days and a standard deviation of 15.

### Customer\_Engagement\_Score

Type: Float  
Description: A score reflecting overall customer engagement, calculated based on interactions with emails and website activity. It’s normally distributed with an average score of 70 and a standard deviation of 15.

## Categorical/Binary Variables:

### Opened\_Previous\_Emails (Target Variable)

Type: Binary (0, 1)  
Description: Whether the customer has opened any previous emails (1 for opened, 0 for not opened).

### Clicked\_Previous\_Emails

Type: Binary (0, 1)  
Description: Whether the customer has clicked on links in previous emails (1 for clicked, 0 for not clicked).

### Device\_Type

Type: Binary (0, 1)  
Description: The type of device used by the customer to open the email. 1 for mobile and 0 for desktop.

**Acceptance Criterion:** Need to deploy the end results using Streamlit etc.

|  |  |  |
| --- | --- | --- |
| **Milestone** | **Duration** | **Task start - End Date** |
| Kick off and Business Objective discussion | 1 day |  |
| EDA | 1 Week |  |
| Model Building | 1 Week |  |
| Model Evaluation | 1 Week |  |
| Feedback | 1 week |  |
| Deployment |
| Final presentation | 1 Day |  |

Protocols:

1. All participants should adhere to agreed timelines and timelines will not be extended.
2. All the documentation – Final presentation and python code to be submitted before the final presentation day.
3. All the participants must attend review meetings.

# \*\*Description\*\*

1. \*\*Numerical Variables:\*\*

    \* Customer\_Age:

      \* Type \*integer\*

      \* Age of the customer

      \* mean : 35

      \* Standard Deviation : 10

    \* Email\_Opened:

      \* Type \*integer\*

      \* Emails opened in the past

      \* mean : 5

      \* Standard Deviation : 2

    \* Emails\_Clicked:

      \* Type \*integer\*

      \* Emails clicked after opening them

      \* mean : 2

      \* Standard Deviation : 1

    \* Purchase\_History:

      \* Type \*float\*

      \* Amount(USD) spent on previous purchase

      \* mean : 1500

      \* Standard Deviation : 500

      \* Normally Distribured

    \* Time\_Spent\_On\_Website:

      \* Type \*float\*

      \* Time spent on website after receving email

      \* mean : 5

      \* Standard Deviation : 2

    \* Days\_Since\_Last\_Open:

      \* Type \*integer\*

      \* Days since email opened

      \* mean : 30

      \* Standard Deviation : 15

    \* Customer\_Engagement\_Score:

      \* Type \*float\*

      \* email interactions and website activity

      \* mean : 70

      \* Standard Deviation : 15

      \* Normally Distribured

2. \*\*Categorical/Binary Variables:\*\*

    \* Opened\_Previous\_Emails:

      \* Target Variable

      \* Type \*binary\*

      \* Opened previous emails(0 no, 1 yes)

    \* Clicked\_Previous\_Emails:

      \* Type \*binary\*

      \* Clicked on previous links(1 clicked, 0 not clicked)

    \* Device\_Type:

      \* Type \*binary\*

      \* User device (0 mobile, 1 Desktop)