ONE PAGE REPORT

Name: B. Pooja
Role: Internship Applicant
Project: Lead Generation Data Cleaning Tool

OBJECTIVE: The main aim was to clean the leads dataset in order to analyze and enhance a simple lead generation process. I decided to concentrate on enhancing the dependability and quality of data by developing a straightforward tool that:

- Eliminates erroneous or absent email addresses
- Removes redundant leads
- Gets the dataset ready for more effective sales outreach.

TOOLS USED:

- Python
- Pandas for data manipulation
- Regular Expressions (import re) for email validation
- Google Colab for implementation

METHODOLOGY:

- Open the CSV file: brought in the **leads.csv** file with the email addresses and names.
- Make the data clean:
- Rows with missing emails were eliminated.
- Regex was used to filter only legitimate email formats, such as pooja@example.com.
- To make sure every lead is distinct, duplicate emails were eliminated.
- The cleaned data was saved for later use in cleaned_leads.csv.

IMPORTANCE:

- Reduces email bounce backs
- Prevents contacting the same lead multiple times
- Saves time for sales teams
- Enables faster and targeted outreach
- Easily integrates with scraping tools

FUTURE IMPLEMENTATION:

- Email domain analysis (for example: company mail vs personal mail)
- Lead scoring features
- CRM integration