used to predict the likelihood of a default. Here is a brief description of each column in the dataset:

- **TARGET:** Binary indicator where '1' represents a default on a loan and '0' represents a non-default. This is the label for our predictive modeling.
- **NAME_CONTRACT_TYPE:** Type of loan contracted. Categorical variable (e.g., 'Cash loans', 'Revolving loans').
- **CODE_GENDER:** Gender of the applicant. Categorical variable ('M' for male, 'F' for female).
- **FLAG_OWN_CAR:** Indicates whether the applicant owns a car ('Y' for yes, 'N' for no, missing in some records).
- **FLAG_OWN_REALTY:** Indicates whether the applicant owns real estate ('Y' for yes, 'N' for no).
- **CNT_CHILDREN:** Number of children the applicant has.
- **AMT_INCOME_TOTAL:** Total annual income of the applicant. Some records are missing.
- **AMT_CREDIT:** Credit amount of the loan taken.
- **AMT_ANNUITY:** Loan annuity.
- **DAYS_BIRTH:** Applicant's age in days at the time of application (negative values indicating the age).
- **YEARS_EMPLOYED:** Number of years the applicant has been employed.

Your Task:

1. **Download the Dataset**

2. **Preliminary Work Reference:** For some preliminary work already done with the dataset, you can access the following Google Colab notebook. This resource will help you understand the initial steps and give you a head start on the data cleaning process.

   Access the Colab Notebook

3. **Data Cleaning Objectives:**

   - **Handle Missing Values:** Identify any missing data and determine the best approach to handle it (e.g., imputation, deletion).
   - **Detect and Treat Outliers:** Examine the dataset for any outliers that could affect analysis and decide on an appropriate treatment.
   - **Correct Data Inconsistencies:** Check for inconsistencies in data types, formatting, or