Understand the Dataset

Review the dataset description to understand the meaning of each column and its data type.
Identify the target variable (Approval Status) and the independent variables.

==========

Data Loading

Load the dataset using Python and libraries such as Pandas.
Display the first few rows of the dataset to get an overview of the data structure.

==========

Data Cleaning

Check for missing values in the dataset and decide how to handle them.

==========

Slice Rows and Columns

Extract the first 10 rows of the dataset to get a quick overview of the data.
Select only the columns Applicant ID, Gender, Age, Annual Income, and Approval Status to create a smaller dataframe for analysis.

==========

Filter Rows Based on Conditions

Extract all rows where the Approval Status is 1 (approved).
Filter all applicants who have an Annual Income greater than $50,000 and a Credit Score above 700.

==========

Index-Based Selection

Use .iloc to select the first 5 rows and columns 1 to 5 (based on index position).
Use .iloc to retrieve rows 50 to 60 and columns corresponding to Credit Score and Loan Amount

Requested.

==========

Access Specific Values

Retrieve the value of Annual Income for the applicant in the 25th row.
Retrieve the Approval Status for a specific Applicant ID (e.g., "A12345") using boolean indexing.

==========

Subset Rows Using Multiple Conditions

Extract all rows where:
- Property Area is "Urban".
- Gender is "Female".
- Approval Status is 0 (rejected).

Combine these conditions to analyze subsets of data.

==========

Set Index and Access Data

Set the Applicant ID column as the dataframe index and then retrieve the information for a specific applicant using .loc.

==========

Create Custom Columns with Slicing

Create a new column Total Income that combines Annual Income and Co-applicant Income.
Create another column Loan-to-Income Ratio by dividing Loan Amount Requested by Total Income.

==========

Advanced Indexing

Use slicing to extract all applicants who are older than 30, live in a "Semi-Urban" area, and have a Credit History of "Good".

==========

Feature Engineering

Encode categorical variables like Gender, Marital Status, and Property Area using techniques such as one-hot encoding or label encoding.

==========

Data Splitting

Split the dataset into training and testing sets (e.g., 80% training, 20% testing) using train_test_split from Scikit-learn.

==========

Build the Logistic Regression Model

Train a logistic regression model on the training data.
Use the Scikit-learn library to fit the model.

==========

Model Evaluation

Plot the confusion matrix to visualize the performance of the model.

==========

Model Interpretation

Interpret the coefficients of the logistic regression model to understand the importance of

features.
Discuss which factors contribute the most to loan approval or rejection.