Country GDP Analysis

Project Overview

This project involves the analysis of country-level data, focusing on key indicators such as birth rate, internet usage, and income group classification. The dataset is cleaned, sliced, and analyzed using **Pandas** to gain insights into different countries' economic and social indicators. The analysis aims to ensure data quality and provide a better understanding of global trends.

Dataset Columns

The dataset contains the following columns:

- **CountryName**: The name of the country.
- **CountryCode**: A unique code representing the country.
- **BirthRate**: The birth rate per country.
- **InternetUsers**: Percentage of the population using the internet.
- **IncomeGroup**: The income group classification (e.g., High income, Low income, etc.).

Key Objectives

1. **Slicing Data**: Filter and extract relevant subsets of the dataset for analysis based on specific conditions like income groups or internet usage.
2. **Data Overview**: Use **info()** and **describe()** functions to understand the dataset's structure, summary statistics, and data types.
3. **Handling Missing Values**: Identify and treat null or missing values to maintain the integrity of the analysis.

Libraries Used

- **Pandas**: For comprehensive data manipulation and analysis, including slicing, treating missing values, and generating summaries.

Features

- **Slicing Data**: The project demonstrates how to slice and filter data using various conditions, such as countries with high birth rates or high internet usage.
- **Data Overview**: Use the Pandas **info()** function to get detailed information about the dataset, including data types and non-null counts, and the **describe()** function to generate summary statistics like mean, max, min, and quartiles.
- **Null Value Treatment**: Methods are applied to detect and treat null values, ensuring that missing data does not skew the analysis results.

Data Analysis Process

1. **Loading and Exploring the Data**: The dataset is loaded, and initial exploratory analysis is performed using **info()** and **describe()** functions to get a sense of the data distribution and structure.
2. **Slicing and Filtering**: Data slicing is done based on various conditions like filtering countries by income group or internet usage.
3. **Handling Missing Values**: Missing values in the dataset are detected and treated using Pandas techniques like filling, dropping, or imputing missing data.