

## Data Collection and Preprocessing Phase

Date	29 July 2025
Project Title	Inflation Analysis
Maximum Marks	3 Marks

### Data Quality Report

Data Source	Data Quality Issue	Severity	Resolution Plan
Inflation Dataset	Missing values for some years (especially in 2020-2021)	High	Handled using unpivoted columns and filled values based on trends
Inflation Dataset	Inconsistent data types for numeric columns (Date as text)	Moderate	Converted text to numeric format using Power Query
Inflation Dataset	Multiple entries for same country-year pair	Low	Standardized country names using Power Query transformation
Inflation Dataset	Outliers in inflation values post-COVID for certain countries	Moderate	Flagged and reviewed manually; excluded unverified anomalies

Inflation Dataset	Nulls in calculated measures like YoY% or Inflation Category	Moderate	Recalculated post-cleaning; ensured dependent columns were accurate
Continent Dataset	Many unnecessary/extra columns apart from region and country name	Moderate	Removed all irrelevant columns and retained only Region and Country