

DATA DICTIONARY

This document describes the variables used in the dataset for this project.

I. General Information

- **DATESET NAME:** List of Variables for NPL Ratio Prediction – final (consolidated from original sources)
- **ORIGINAL SOURCE:** Data retrieved from the following websites:
 - [Banko Sentral ng Pilipinas](#)
 - [Philippine Statistics Authority OpenSTAT](#)
- **NUMBER OF OBSERVATIONS:** 132 monthly observations
- **NUMBER OF VARIABLES:** 30 (1 Historical NPL ratio and 29 macroeconomic indicators)
- **DATE RANGE:** January 2014 – December 2024

II. Variable Descriptions

The following information about the variables/macroeconomic indicators collected from the original sources are listed in the table below:

| Variable | Variable Code | Definition |
|---|-------------------------|--|
| NPL Ratio | NPL.Ratio | Share of loans that are 90+ days past due or unlikely to be repaid. |
| Gross Domestic Product (2018=100) | GDP.Constant | Inflation-adjusted GDP using 2018 as base year. |
| Gross Domestic Product (current prices) | GDP.Current | Total value of goods/services without inflation adjustment. |
| Gross Capital Formation (2018=100) | Gross.Capital.Formation | Investment in construction, equipment, and intellectual assets (2018=100). |
| Government Expenditure (2018=100) | Government.Expenditure | Public sector spending on goods/services (2018=100). |
| Household Consumption (2018=100) | Household.Consumption | Total household spending on goods/services (2018=100). |
| Unemployment Rate | Unemployment | % of labor force without jobs but actively seeking work. |
| Labor Force Participation Rate | LaborForce.ParRate | % of working-age population employed or seeking work. |
| CPI (2018=100) | CPI.All.Item | Overall consumer price index based on 2018 basket. |

| | | |
|---------------------------------|----------------------|--|
| CPI – Housing | CPI.Housing | CPI for housing, utilities, and fuels. |
| CPI – Furnishing | CPI.Furnishing | CPI for household equipment and maintenance. |
| CPI – Transport | CPI.Transport | CPI for transport-related goods and services. |
| CPI – Restaurants & Misc. | CPI.Goods | CPI for restaurants and miscellaneous services. |
| Inflation Rate (2018=100) | Inf.All.Item | Annual % change in CPI (2018 base). |
| Inflation – Transport | Inf.Transport | Inflation for transport commodities. |
| Inflation – Housing | Inf.Housing | Inflation for housing-related items. |
| Inflation – Furnishing | Inf.Furnishing | Inflation for household furnishings. |
| Inflation – Restaurants & Misc. | Inf.Goods | Inflation for restaurant and miscellaneous goods. |
| Producer Price Index (2018=100) | PPI.2018 | Price change of goods at producer level (2018 base). |
| PHP to USD Exchange Rate | PHP.USD.rate | Number of PHP needed to buy 1 USD. |
| Brent Crude Oil Spot Price | Brent.Crude.Oil.Spot | Monthly Brent crude oil price per barrel (USD). |
| Consumer Confidence Index | CCI | Outlook of consumers on future economic conditions. |
| Business Confidence Index | BCI | Business sentiment on current and future economy. |
| RPI – Manufactured Goods | RPI.Goods | Retail prices of manufactured goods. |
| RPI – Machinery | RPI.Machinery | Retail prices of machinery and transport equipment. |
| WPI – Manufactured Goods | WPI.Goods | Wholesale prices of manufactured goods. |
| WPI – Machinery | WPI.Machinery | Wholesale prices of machinery and transport equipment. |
| Bank Average Lending Rates | Bank.Ave.Lend.Rate | Average interest rate charged by banks on loans. |
| Key Interest Rate / RRP | Interest.Rate | Benchmark rate set by BSP for bank borrowing. |
| WTI Crude Oil Spot Price | WTI.Crude.Oil.Spot | Monthly WTI crude oil price per barrel (USD). |

III. Additional Notes:

- **Missing Data Handling:** The historical NPL ratio and the macroeconomic indicators need alignment in the time series information as the time series data involve quarterly and monthly observations. To address the alignment, this project converts the quarterly information into monthly observations using **interpolation techniques**:
 - **Linear Interpolation** – for quarterly data with obvious linear trends
 - **Kalman Smoothing** – for more complex trends
- **Data Transformations:** Data transformations are done to the historical NPL ratio and indicators as means to improve stationarity and variance of the data. The following data transformations are applied:
 - **Logarithmic Transformation** – The main data transformation used on the historical NPL ratio and other macroeconomic indicators to reduce effect of variance and force non-negative point forecasts and prediction intervals
 - **Inverse Hyperbolic Sine (ASINH) Transformation** – Alternative to the logarithmic transformation for the following time series data with zero and negative values
 - **Differencing** – Removes the trend of the time series data. Due to some of the time series data originally being quarterly data, this project opts in using quarterly seasonal differencing as an alternative for interpolated quarterly data to remove quarterly trends.
- **Feature Engineering** – To also account for the time series and serial correlation aspect of the data, this project involves lagged variables of macroeconomic indicators, included in the model selection process. The method of determining how many months of lag from the macroeconomic indicators should be included is based on the results of each of their PACF plots. The ones with autocorrelated lags from the plots will be included for the model selection process.