**Regression Project Time Series Analysis**

Time series analysis and modeling is a statistical technique used to analyze and model data that is collected over time. It involves examining the patterns, trends, and behavior of the data over time to uncover insights and make predictions.

On the other hand, regression analysis is a statistical method used to examine the relationship between a dependent variable and one or more independent variables. It is often used to determine the relationship between two or more variables and to make predictions about future outcomes.

While both time series analysis and regression analysis are statistical techniques used to make predictions, they differ in their approach. Time series analysis and modeling focus specifically on analyzing data that is collected over time, while regression analysis can be used to analyze any type of data.

For the purposes of this project, I will look at how regression is used in time-series analysis and modeling. Let us look at the various types of regression techniques that are commonly used, as well as the benefits thereof in using regression in time-series analysis. Types of Regression Techniques Used in Time-Series Analysis

Linear Regression: One of the most used regression techniques in time-series analysis is linear regression. It aids in the prediction of the dependent variable using the independent variables. The dependent variable and independent variable are assumed to have a linear relationship. This means that the dependent variable changes linearly as the independent variable changes.

Exponential Smoothing: Another popular time-series model for forecasting is exponential smoothing. It is a statistical method for smoothing out data noise by assigning different weights to data points based on their distance from the current time point.

ARIMA (Autoregressive Integrated Moving Average): ARIMA is a time-series model that combines three key components: Autoregression (AR), Integrated (I), and Moving Average (MA) (MA). ARIMA is a model for non-stationary time series data. It aids in the identification of patterns, trends, and seasonality in data.

What Benefits are there in using Regression in Time-Series Analysis

Forecasting: By analyzing the relationship between the dependent and independent variables, a regression can help predict future trends and values.

Identifying Trends and Patterns: Regression helps in identifying the trends and patterns in the time-series data by analyzing the relationship between the dependent variable and the independent variable.

Understanding the Impact of Past Events: Regression analyzes the relationship between the dependent variable and the independent variable to help understand the impact of past events on current data.

Data Used and Structure

We will predict store sales using data from Corporation Favorita, which is a leading retail corporation in Ecuador, operating various supermarket chains, hypermarkets, and convenience stores. A time series analysis of Favorita's sales data over time could provide insights into patterns, trends, and seasonality of its sales.

The dataset contained a series of CSV files from the store-sales-time-series-forecasting folder. The following datasets were used for the purpose of this project.

***Train.csv****: Training data with time series for “store\_nbr,” “family,” “onpromotion,” and “sales.”*

***Test.csv:*** *Similar data for prediction purposes.*

***Sample\_submission.csv****: A sample submission file in the correct format.*

***Transaction.csv****: Contains dates, store\_nbr, and transactions.*

***Stores.csv:*** *Store metadata.*

***Oil.csv:*** *Daily oil prices.*

**Holidays\_events.csv:** Information on holidays and events.

Exploratory data analysis is an integral part of this project in that it allows us to analyze and understand the data before modeling it. Based on the analysis, it helps us in identifying patterns, relationships, and trends in the data thereby providing insights into the analysis of the entire work. In this project, we take a cursory look at the EDA process using Python libraries and applied it to analyze Corporation Favorita’s sales data.

Overall, time series analysis and modeling and regression analysis are complementary techniques that can be used together to provide a more complete analysis and prediction of data. It helps in predicting future trends, identifying patterns, and evaluating the impact of past events on the current data. Data enthusiast can make more informed decisions and improve their data analytic techniques by understanding the benefits of using regression in time-series analysis.