**Project Proposal: Retail Analysis with Walmart Data**

Introduction:

The aim of this project is to analyze the retail sales data of Walmart stores and build a predictive model to accurately forecast sales and demand. The project will involve exploring the dataset, performing statistical analysis, and developing a statistical model using linear regression. The dataset includes historical sales data from 45 Walmart stores, covering the period from 2010-02-05 to 2012-11-01.

Data Description:

The dataset consists of the following fields:

Store: Store number

Date: Week of sales

Weekly\_Sales: Sales for the given store

Holiday\_Flag: Indicator for special holiday week (1) or non-holiday week (0)

Temperature: Temperature on the day of sale

Fuel\_Price: Cost of fuel in the region

CPI: Prevailing consumer price index

Unemployment: Prevailing unemployment rate

Additionally, there are specific holiday events such as Super Bowl, Labour Day, Thanksgiving, and Christmas, which will be taken into account during the analysis.

Analysis Tasks:

The following analysis tasks will be performed:

Basic Statistics tasks:

Identify the store with the maximum sales.

Determine the store with the maximum standard deviation in sales and calculate the coefficient of mean to standard deviation.

Identify store/s with a good quarterly growth rate in Q3'2012.

Find holidays with higher sales than the mean sales in the non-holiday season for all stores together.

Monthly and Semester Sales Analysis:

Provide a monthly and semester view of sales in units and extract insights from the data.

Statistical Model Development:

For Store 1, build a prediction model to forecast demand using linear regression.

Hypothesize the impact of variables such as date, CPI, unemployment, and fuel price on sales.

Convert dates into days by creating a new variable for regression analysis.

Methodology:

The project will follow the following methodology:

Data preprocessing: Handle missing values, convert data types, and perform necessary transformations.

Exploratory data analysis: Visualize the data and derive insights from the various variables.

Statistical analysis: Calculate basic statistics, identify trends, and perform seasonal analysis.

Model development: Build a linear regression model for Store 1, considering relevant variables.

Model evaluation: Assess the model's performance and analyze the impact of different factors on sales.

Presentation of results: Summarize the findings, insights, and recommendations.

Deliverables:

The project will deliver the following outcomes:

Detailed analysis of the Walmart sales dataset.

Identification of stores with maximum sales and standard deviation.

Determination of stores with good quarterly growth rate in Q3'2012.

Identification of holidays with higher sales than the mean sales in the non-holiday season.

Monthly and semester view of sales in units with insights.

Statistical model for Store 1's sales prediction using linear regression.

Evaluation of the impact of variables on sales.

Final report summarizing the findings, analysis, and recommendations.

By undertaking this project, we aim to provide valuable insights to Walmart for better sales forecasting and decision-making processes.