Assignment - 3

Problem Statement

You are provided with historical sales data for 45 stores of a Retail chain located in different regions. Each store contains a number of departments, and you are tasked with predicting the department-wide sales for each store.

The data is provided in 4 different CSVs as below:

stores.csv
This file contains anonymized information about the 45 stores, indicating the type and size of store.

train.csv
This is the historical training data, which covers to 2010-02-05 to 2012-11-01. Within this file you will find the following fields:

- Store - the store number
- Dept - the department number
- Date - the week
- Weekly_Sales - sales for the given department in the given store
- IsHoliday - whether the week is a special holiday week

test.csv
This file is identical to train.csv, except we have withheld the weekly sales. You must predict the sales for each triplet of store, department, and date in this file.

features.csv
This file contains additional data related to the store, department, and regional activity for the given dates. It contains the following fields:

- Store - the store number
- Date - the week
- Temperature - average temperature in the region

- Fuel_Price - cost of fuel in the region
- MarkDown1-5 - anonymized data related to promotional markdowns that the Retail chain is running. MarkDown data is only available after Nov 2011, and is not available for all stores all the time. Any missing value is marked with an NA.
- CPI - the consumer price index
- Unemployment - the unemployment rate
- IsHoliday - whether the week is a special holiday week

Output

You are expected to predict the weekly sales with a time horizon of a week and share the results in the form of CSVs. Along with that you would be required to do a through EDA on the given dataset and bubble up insights.
Also along with the result submission the implementation in the form of a python notebook is also expected to be submitted.

General Guidelines:

- Fork provided repository and submit a pull request along with the result and python notebook - **https://github.com/AnjnaBhati12/assignments.git**
- Access data from the folder assignments/data/saleforecasting