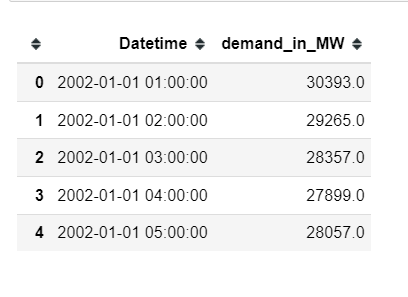
About Dataset (Energy Demand)



PJM Interconnection LLC (PJM) is a regional transmission organization (RTO) in the United States. It is part of the Eastern Interconnection grid operating an electric transmission system serving all or parts of Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia, and the District of Columbia.

The hourly power consumption data comes from PJM's website and are in megawatts (MW).

The regions have changed over the years so data may only appear for certain dates per region.



How to use Data Set :-

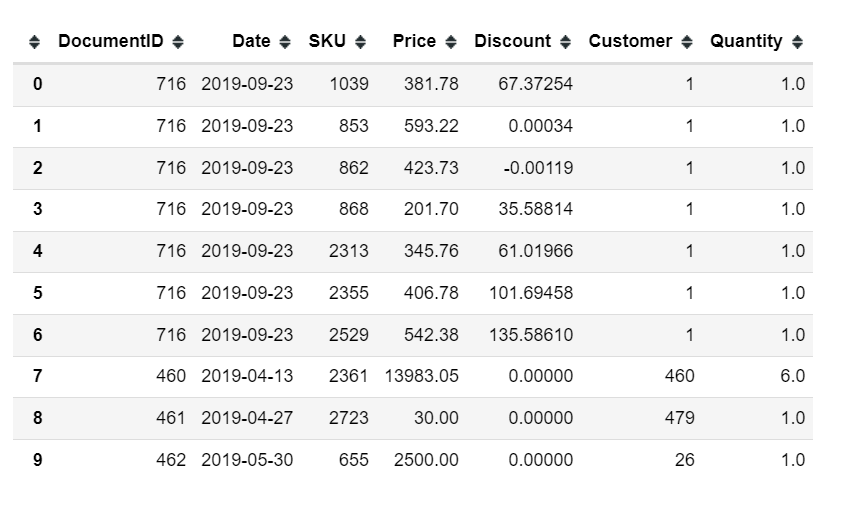
* Split the last year into a test set- can you build a model to predict energy consumption?
* Find trends in energy consumption around hours of the day, holidays, or long term trends?
* Understand how daily trends change depending of the time of year. Summer trends are very different than winter trends.

About Dataset ( Retail Demand)



Data holds the basic information about sales data. The company have sales agencies / resellers and branches and the data file holds only the branch/reseller information in the customer field.

**Data Snapchat :**



**Data Description :**

* **DocumentID** : ID of the transaction. A transaction might hold multiple records for the same customer at the same date with multiple products (SKU). DocumentID might be useful for combining the transactions and detecting the items sold together.
* **Date** : Date of transaction / sell. In the date time format.
* **SKU** : Item / Product code. The unique code for each item sold.
* **Price** : Sales price for the transaction. The price for the product for the customer for the date.
* **Discount** : discount amount for the transaction.
* **Customer** : Unique customer id for each customer. For the data set, customer can be a reseller or a branch of the company.
* **Quantity** : Number of items sold in the transaction.

How to Use Data Set ?

Some of the potential questions from the data might be listed as below, please add more to discussion if you have any idea:

* What is the optimum price for each item?
* Demand prediction for each item on next months.
* Customer Lifetime Value for each Customer.
* Customer Segmentation (an easy approach of RFM or more complex segmentations)
* Customer <> Product Recommendation (what are the best products for the customers).

Our main focus is going to be **“Demand Prediction For Each Item On Next Months”**

Feature Engineering for retail

Feature engineering plays a crucial role in time series analysis, especially in the context of retail sales forecasting. By extracting and creating meaningful features from raw data, such as temporal attributes, customer demographics, order characteristics, promotional activities, and economic indicators, analysts can uncover hidden patterns and relationships that contribute to sales fluctuations over time. These engineered features provide valuable insights into consumer behavior, market dynamics, and external factors influencing sales performance. Moreover, techniques like lag features, rolling window statistics, seasonality indicators, interaction features, and difference features enable analysts to capture complex temporal patterns and dependencies within the data, ultimately enhancing the accuracy and robustness of predictive models. In retail, these features can help businesses understand customer preferences, optimize inventory management, devise targeted marketing strategies, and adapt to changing market conditions. Thus, feature engineering is an indispensable step in harnessing the predictive power of time series data for informed decision-making and competitive advantage in the retail industry.

* **Sales Acceleration:**  This metric measures the rate of change in sales velocity over a specific time interval, such as days, weeks, or months. It indicates how quickly sales are increasing or decreasing over time.



Where Sales Velocity could be units sold per day/week/month

* **Seasonality Strength:** Seasonality strength measures the intensity of sales fluctuations between peak and average periods within a seasonal cycle. It provides insights into the impact of seasonal trends on sales performance.



Peak Sales is the maximum sales volume observed in a season, and Average Sales is the average sales volume over the same period.

* **Demand Variability:** This metric calculates the variability in sales volume over a specified historical period. It helps identify fluctuations and trends in demand, which can inform inventory planning and marketing strategies.



Calculated over a period of time

* **Lag Features, Rolling Window Statistics, Difference Indicators**: These categories involve various techniques for analyzing historical sales data and capturing patterns over time, including lagged demand, rolling averages, seasonality indicators, interaction effects with other variables like promotions or economic indicators, and differences in sales quantities or economic indicators between different time periods. These features help identify trends, seasonality, and relationships between different variables in the sales data, aiding in forecasting and decision-making processes.
* **Economic Indicator ( External Data ) :** These features include direct use or changes in economic indicators such as GDP, inflation rate, interest rate, PPI (Producer Price Index), and CPI (Consumer Price Index). Changes in economic indicators can impact consumer spending behaviour and overall market conditions.

