MEASURE ENERGY CONSUMPTION

Team members

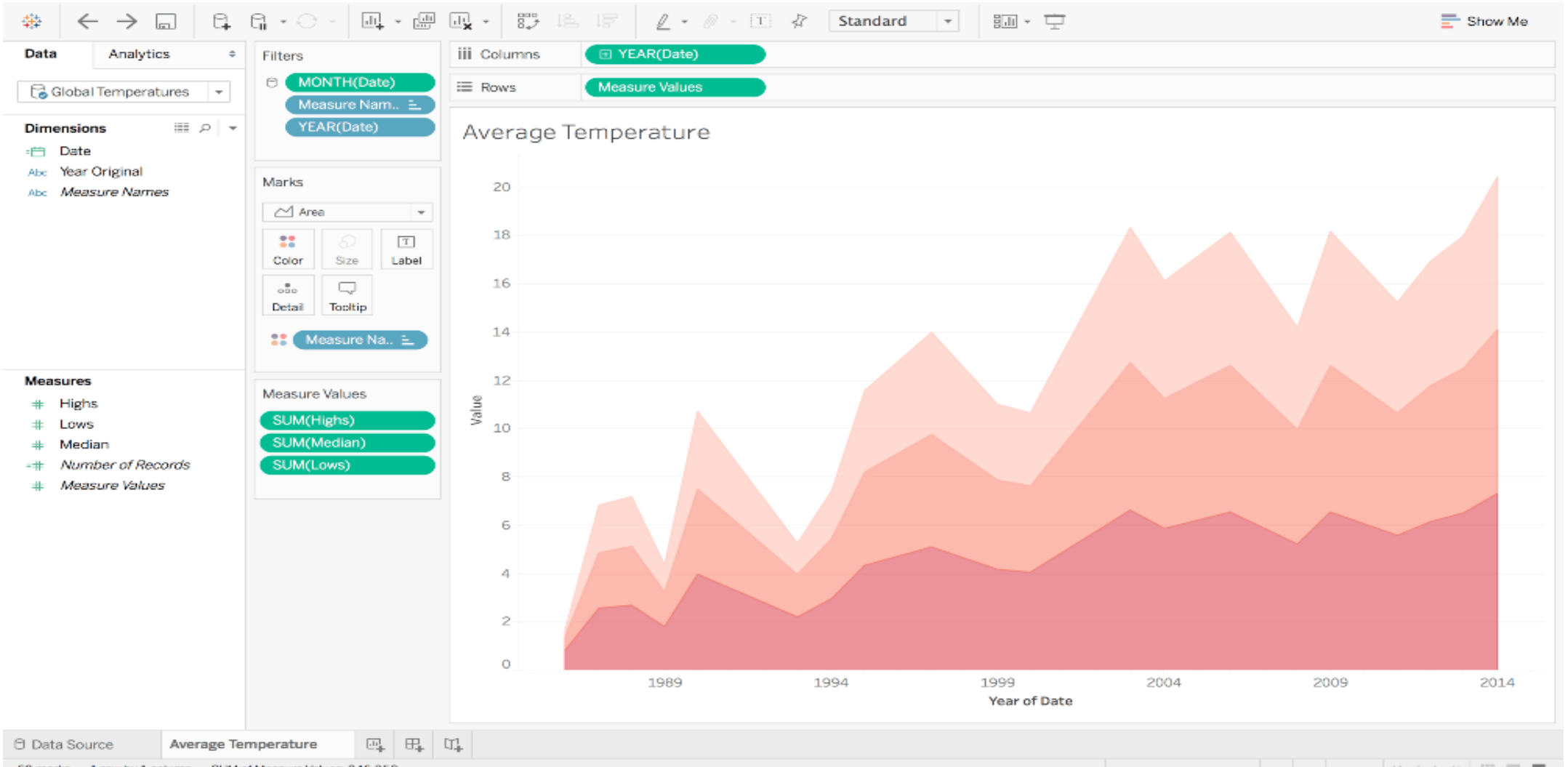
510521104039 : SASIKUMAR R

PHASE 2 : DOCUMENT SUBMISSION

PROJECT TITLE :

Explore innovative techniques such as time series analysis and machine learning models to predict future energy consumption patterns.

Time series analysis



🔻 Time series analysis is a specific way of analyzing a sequence of data points collected over an interval of time. In time series analysis, analysts record data points at consistent intervals over a set period of time rather than just recording the data points intermittently or randomly. However, this type of analysis is not merely the act of collecting data over time.

🔻 What sets time series data apart from other data is that the analysis can show how variables change over time. In other words, time is a crucial variable because it shows how the data adjusts over the course of the data points as well as the final results. It provides an additional source of information and a set order of dependencies between the data.

🔻 Time series analysis typically requires a large number of data points to ensure consistency and reliability. An extensive data set ensures you have a representative sample size and that analysis can cut through noisy data. It also ensures that any trends or patterns discovered are not outliers and can account for seasonal variance. Additionally, time series data can be used for forecasting—predicting future data based on historical data.

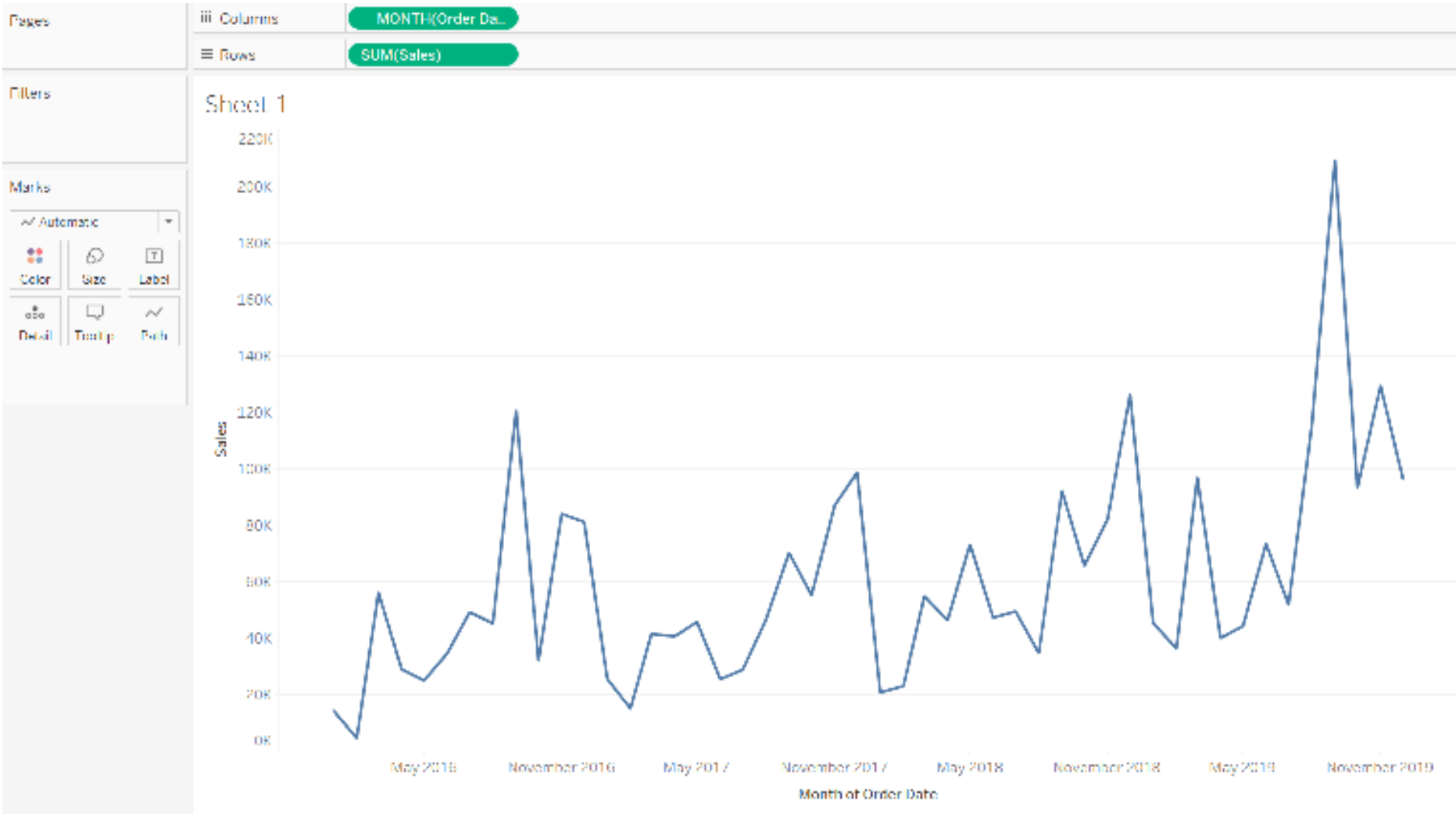
Why organizations use time series data analysis

🔻 Time series analysis helps organizations understand the underlying causes of trends or systemic patterns over time. Using data visualizations, business users can see seasonal trends and dig deeper into why these trends occur. With modern analytics platforms, these visualizations can go far beyond line graphs.

🔻 When organizations analyze data over consistent intervals, they can also use time series forecasting to predict the likelihood of future events. Time series forecasting is part of predictive analytics. It can show likely changes in the data, like seasonality or cyclic behavior, which provides a better understanding of data variables and helps forecast better.

🔻 For example, Des Moines Public Schools analyzed five years of student achievement data to identify at-risk students and track progress over time. Today's technology allows us to collect massive amounts of data every day and it's easier than ever to gather enough consistent data for comprehensive analysis.

Time series analysis examples



Time series analysis is used for non-stationary data—things that are constantly fluctuating over time or are affected by time. Industries like finance, retail, and economics frequently use time series analysis because currency and sales are always changing. Stock market analysis is an excellent example of time series analysis in action, especially with automated trading algorithms. Likewise, time series analysis is ideal for forecasting weather changes, helping meteorologists predict everything from tomorrow's weather report to future years of climate change. Examples of time series analysis in action include:

- Weather data
- Rainfall measurements
- Temperature readings
- Heart rate monitoring (EKG)
- Brain monitoring (EEG)

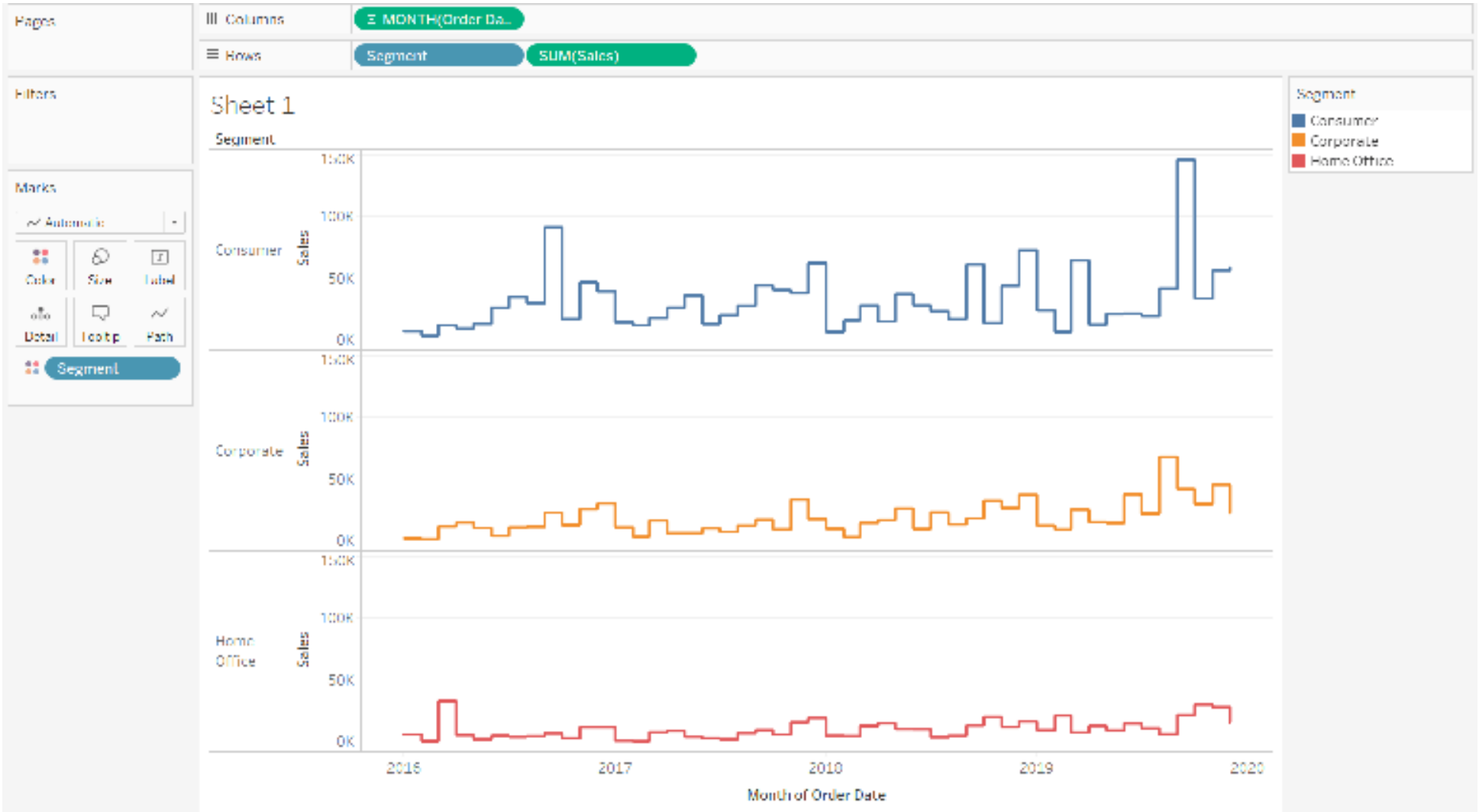
- Quarterly sales
- Stock price
- Automated stock trading
- Industry forecasts
- Interest rates

Time Series Analysis Types

Because time series analysis includes many categories or variations of data, analysts sometimes must make complex models. However, analysts can't account for all variances, and they can't generalize a specific model to every sample. Models that are too complex or that try to do too many things can lead to a lack of fit. Lack of fit or overfitting models lead to those models not distinguishing between random error and true relationships, leaving analysis skewed and forecasts incorrect.

Models of time series analysis include:

Classification: Identifies and assigns categories to the data.
Curve fitting: Plots the data along a curve to study the relationships of variables within the data.
Descriptive analysis: Identifies patterns in time series data, like trends, cycles, or seasonal variation.
Explanative analysis: Attempts to understand the data and the relationships within it, as well as cause and effect.
Exploratory analysis: Highlights the main characteristics of the time series data, usually in a visual format.
Forecasting: Predicts future data. This type is based on historical trends. It uses the historical data as a model for future data, predicting scenarios that could happen along future plot points.
Intervention analysis: Studies how an event can change the data.
Segmentation: Splits the data into segments to show the underlying properties of the source information.


Pages
Columns
MONTH(Order Da...
Rows
Segment
SUM(Sales)
Filters
Marks
Color
Size
Label
Detail
Tooltip
Path
Segment
Sheet 1
Segment
Consumer
Corporate
Home Office
Sales
150K
100K
50K
0K
2016
2017
2018
2019
2020
Month of Order Date

Data classification

Further, time series data can be classified into two main categories:

Stock time series data means measuring attributes at a certain point in time, like a static snapshot of the information as it was.
Flow time series data means measuring the activity of the attributes over a certain period, which is generally part of the total whole and makes up a portion of the results.

Data variations

In time series data, variations can occur sporadically throughout the data:

Functional analysis can pick out the patterns and relationships within the data to identify notable events.

Trend analysis means determining consistent movement in a certain direction. There are two types of trends: deterministic, where we can find the underlying cause, and stochastic, which is random and unexplainable.

Seasonal variation describes events that occur at specific and regular intervals during the course of a year. Serial dependence occurs when data points close together in time tend to be related.

Time series analysis and forecasting models must define the types of data relevant to answering the business question. Once analysts have chosen the relevant data they want to analyze, they choose what types of analysis and techniques are the best fit.

Important Considerations for Time Series Analysis

While time series data is data collected over time, there are different types of data that describe how and when that time data was recorded. For example:

Time series data is data that is recorded over consistent intervals of time.

Cross-sectional data consists of several variables recorded at the same time.

Pooled data is a combination of both time series data and cross-sectional data.

Time Series Analysis Models and Techniques

Just as there are many types and models, there are also a variety of methods to study data. Here are the three most common.

Box-Jenkins ARIMA models: These univariate models are used to better understand a single time-dependent variable, such as temperature over time, and to predict future data points of variables. These models work on the assumption that the data is stationary. Analysts have to account for and remove as many differences and seasonalities in past data points as they can. Thankfully, the ARIMA model includes terms to account for moving averages, seasonal difference operators, and autoregressive terms within the model.

Box-Jenkins Multivariate Models: Multivariate models are used to analyze more than one time-dependent variable, such as temperature and humidity, over time.

Holt-Winters Method: The Holt-Winters method is an exponential smoothing technique. It is designed to predict outcomes, provided that the data points include seasonality.

Books about time series analysis

▼ Time series analysis is not a new study, despite technology making it easier to access. Many of the recommended texts teaching the subject's fundamental theories and practices have been around for several decades. And the method itself is even older than that. We have been using time series analysis for thousands of years, all the way back to the ancient studies of planetary movement and navigation.

▼ Because of this, there are thousands of books about the study, and some are old and outdated. As such, we created a list of the top books about time series analysis. These are a mix of textbooks and reference guides, and good for beginners through to experts. You'll find theory, examples, case studies, practices, and more in these books.

Times series analysis and R

▼ The open-source programming language and environment R can complete common time series analysis functions, such as plotting, with just a few keystrokes. More complex functions involve finding seasonal values or irregularities. Time series analysis in Python is also popular for finding trends and forecasting.

▼ Time series analysis is a technical and robust subject, and this guide just scratches the surface. To learn more about the theories and practical applications, check out our time series analysis resources and customer stories.

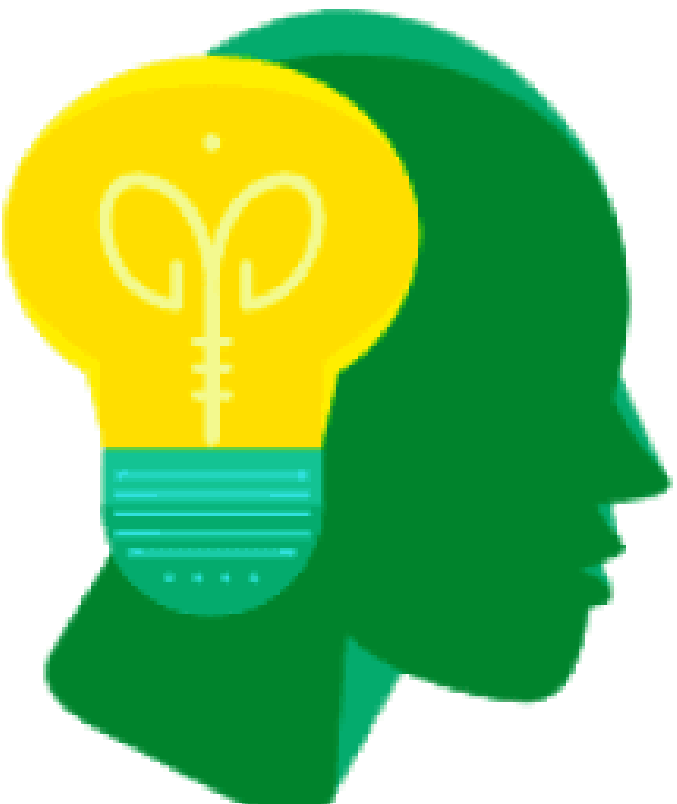
Machine Learning Models:

🔻 A machine learning model is defined as a mathematical representation of the output of the training process. Machine learning is the study of different algorithms that can improve automatically through experience & old data and build the model. A machine learning model is similar to computer software designed to recognize patterns or behaviors based on previous experience or data. The learning algorithm discovers patterns within the training data, and it outputs an ML model which captures these patterns and makes predictions on new data.

🔻 Let's understand an example of the ML model where we are creating an app to recognize the user's emotions based on facial expressions. So, creating such an app is possible by Machine learning models where we will train a model by feeding images of faces with various emotions labeled on them. Whenever this app is used to determine the user's mood, it reads all fed data then determines any user's mood.

🔻 Hence, in simple words, we can say that a machine learning model is a simplified representation of something or a process. In this topic, we will discuss different machine learning models and their techniques and algorithms.

Machine Learning Models

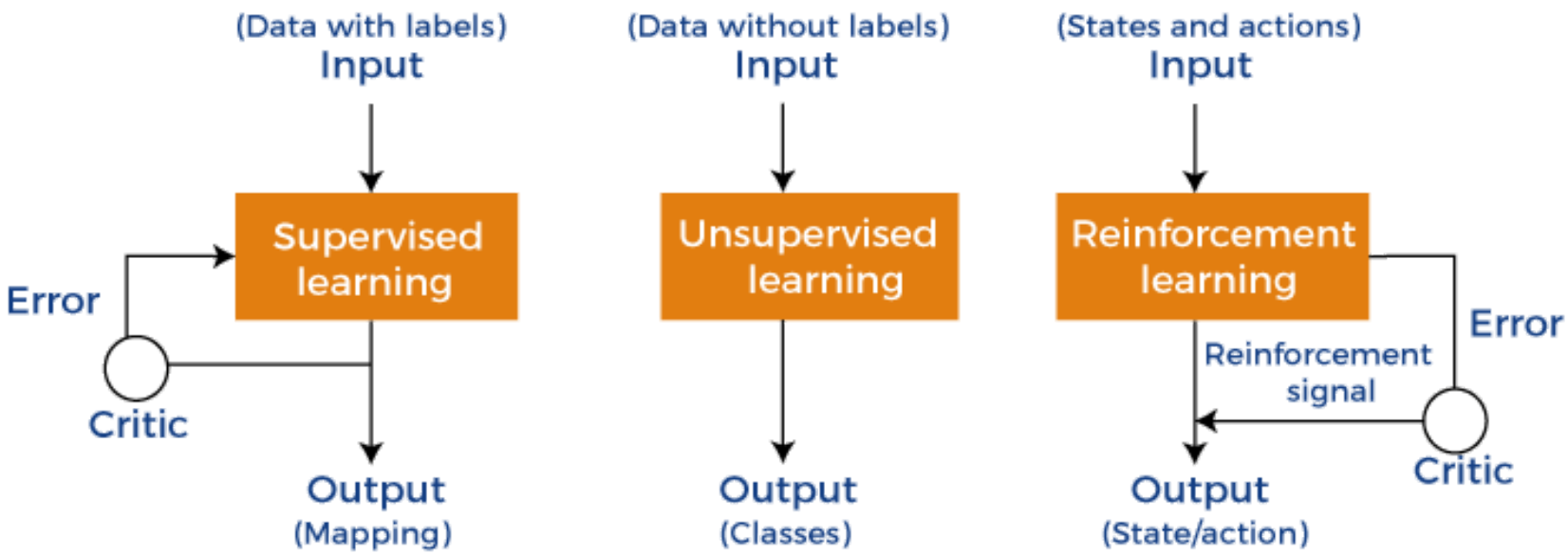


- Classification Models
- Clustering
- Regression Models
- Dimensionality Reduction
- Deep Learning etc.

Classification of Machine Learning Models:

Based on different business goals and data sets, there are three learning models for algorithms. Each machine learning algorithm settles into one of the three models:

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning



ELECTRICITY CONSUMPTION PREDICTION USING MACHINE LEARNING

Introduction

Electricity is expected to replace other energy sources as the main source for usage in homes, businesses, and transportation in the near future[9]. This emphasises how important it is to predict power consumption properly because it has a big impact on a lot of operational and business operations. Electricity demand is frequently referred to as load in the context of electrical engineering; both terms will be used here. Electricity is becoming a major aspect of our daily lives.In today's society, it has evolved into a basic requirement. The amount of power consumed is rising quickly

Methodology

Problem Statement

Predicting electricity usage is a major issue in energy management. For effective energy management, accurate electricity consumption forecasting is crucial because it enables energy suppliers to optimise energy distribution, cut down on energy waste, and avoid overloading the power system.

The formula for KNN regression is as follows:

y_hat = (1/K) * ∑(yi), i=1 to K Where:

y_hat is the predicted value of the target variable for a given observation.

K is the number of nearest neighbors that will be used to make the prediction. yi is the value of the target variable for the i-th nearest neighbor to the observation

```
df['day']=df['date'].dt.day
```

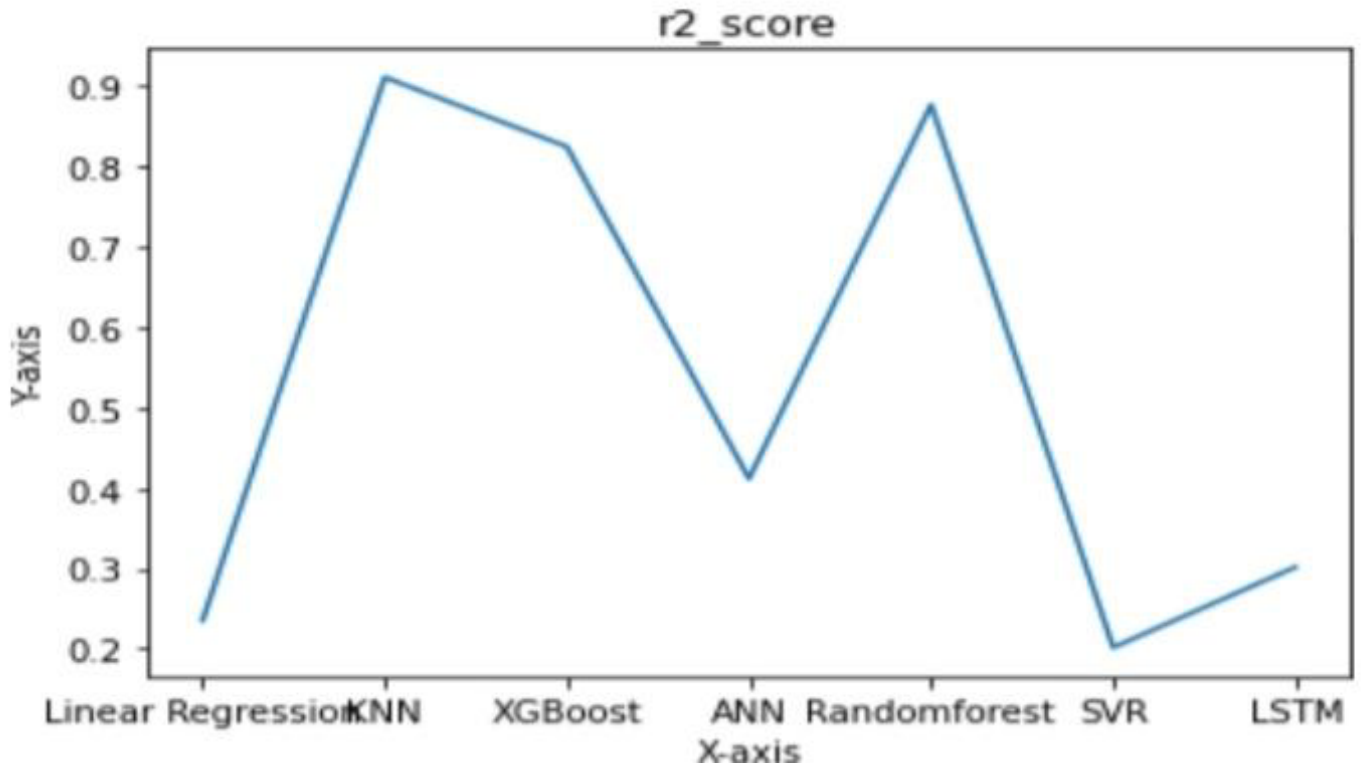
```
df['month']=df['date'].dt.month
df['year']=df['date'].dt.year
```

```
df.head()
```

	datetime	id	name	geoid	geoname	value	date	day	month	year
0	2014-01-01 23:00:00	600	Precio mercado SPOT Diario ESP	3.0	España	25.280833	2014-01-01 23:00:00	1	1	2014
1	2014-01-02 23:00:00	600	Precio mercado SPOT Diario ESP	3.0	España	39.924167	2014-01-02 23:00:00	2	1	2014
2	2014-01-03 23:00:00	600	Precio mercado SPOT Diario ESP	3.0	España	4.992083	2014-01-03 23:00:00	3	1	2014
3	2014-01-04 23:00:00	600	Precio mercado SPOT Diario ESP	3.0	España	4.091667	2014-01-04 23:00:00	4	1	2014

Experiment Results

Coefficient of Determination(R2)



KNN result :

```
[18] from sklearn.neighbors import KNeighborsRegressor
     knn = KNeighborsRegressor(n_neighbors=5)
     knn.fit(X_train, y_train)
```

```
KNeighborsRegressor()
```

```
predi=knn.predict(X_test)
```

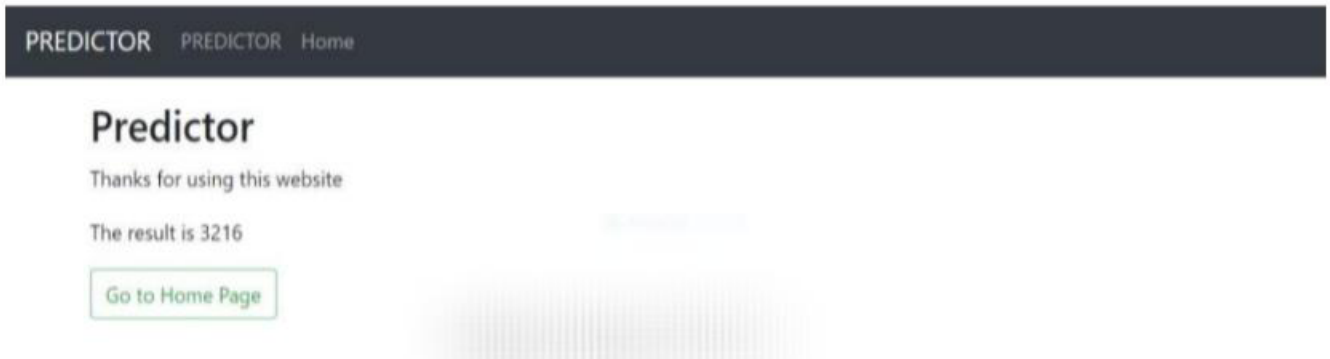
+ Code + Text

```
[20] knn.score(X_test, y_test)
```

```
0.909202964846872
```

Input :

Output screen :



Conclusion :

In recent years, forecasting electricity usage using machine learning approaches has gained popularity as a study topic. Accurately projecting future power consumption isessential for effective energy management, cost savings, and environmental sustainability given the rising demand for energy. It is important to keep in mind that forecasting electricity consumption is a challenging process that calls for careful consideration of a number of variables, including seasonality, time of day, and weather. To make accurate forecasts, it is essential to choose the right characteristics and models.Additionally, predicting energy consumption is a continual process that has to be updated and monitored often to account for changes in consumer behaviour, environmental conditions, and other pertinent variables.

References:

- A Review of Machine Learning Models for Forecasting Electricity Consumption
- Jui-Sheng Chou, Duc-Son Tran,Forecasting energy consumption time series usingmachine learning techniques based on usage patterns of residentialhouseholders,Energy,Volume 165, Part B,2018,Pages709-726,ISSN0360-5442
- A. González-Briones, G. Hernández, J. M. Corchado, S. Omatu and M. S. Mohamad,"Machine Learning Models for Electricity Consumption Forecasting: A Review,"2019 2nd International Conference on Computer Applications & Information Security(ICCAIS), Riyadh, Saudi Arabia, 2019, pp. 1-6
- A 2019 study by Alsmadi, Al-Madi, and Al-Turjman. a short-term powerconsumption forecasting method based on machine learning. IEEE Access, 7,141329–141338.
- In 2019, Wang, X., Li, Y., and Wang, K. an innovative extreme learning machinealgorithm-based prediction model for power use. Energy 12, page 3283.
- Zhang, J., Gao, F., Liu, L., and Zou, Y. (2021). A thorough examination of short-termload prediction utilising machine learning methods. 14, no. 10 (energy), 2958.
- In 2020, Xu, X., Yang, Y., Jiang, L., and Wang, W. Forecasting electricity use basedon upgraded extreme learning machine and meteorological data. 51162-51172. IEEEAccess,

- Almutairi, A. H., Alqahtani, A. M., & Alqahtani, M. A. (2020). Alotaibi, M. R. Aoverview of machine learning-based methods for forecasting power use. SustainableDevelopment, 12(8), 3381.
- Hu, M., Huang, and Chen, Y. (2020). a hybrid strategy combining machine learningand wavelet transform for estimating power usage. 20632-20644. IEEE Access,
- 2019: Yin, J., Li, H., Zhang, X., and Liu, X. a method based on deep learning forpredicting power usage in a smart grid. 12, 703 Energy
- Li, W., Li, J., Li, Z., and Li, Y. (2021). Prediction of electricity usage using anupgraded extreme learning machine optimised by adaptive chaotic particle swarmoptimisation. Energy, 14, 82.
- Nair, N., and Chen, Z. (2021). A thorough evaluation of machine learning-basedelectricity demand forecasting. Applied Energy, 286. 116505.
- Zhang, J., Wang, L., Wang, W., and Wu, Y. (2019). a cutting-edge hybrid modelintegrating deep learning and ensemble learning for predicting power use. Energy, 12,4082.
- Hong, Y., Chen, W., and Chen, J. (2021). employing cutting-edge feature extractiontechniques and machine learning algorithms to anticipate short-term power demand.Applied Energy, 296. 117511.

▼ Gandomi, A. H., Alavi, A. H., and Siahkolah (2019). models of hybrid machinelearning for predicting power use. 12, 1362, Energies.
▼ In 2020, Wang, Y., Zhao, L., Shi, Y., and Li H. using hybrid machine learningmethods to anticipate electricity use. IEEE Access, 8 (58283-58294).
▼ Zhang, Y., Li, Y., Wu, J., and Yang, X. (2019). Energy, 12(2), 208. Machinelearning-based short-term forecasting of power use.
▼ Y. Liu, T. Lin, D. Zhao, and J. Chen (2019). LSSVM-based electricityconsumption forecasting that is hybrid algorithm-optimized. 36(6), 5921–5932,Journal of Intelligent & Fuzzy Systems.
▼ Wang, J.; Yu, J. (2020). Short-term load forecasting using ensemble learning andmachine learning techniques. Energy 13(3), 60 ▼ Y. Sri Lalitha, G. V. Reddy, K. Swapnika, et al., "Analysis of Customer Reviewsusing Deep Neural Network," International Conference on Artificial IntelligenceTrends and Pattern Recognition (ICAITPR), 2022, pp. 1-5,