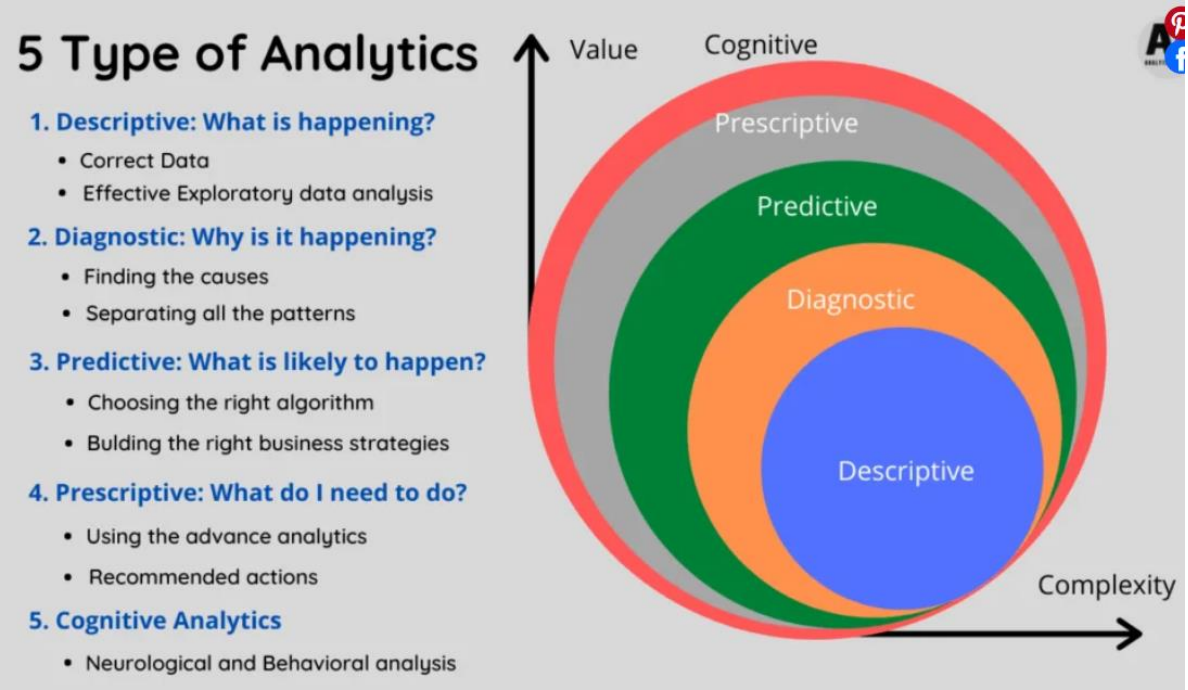
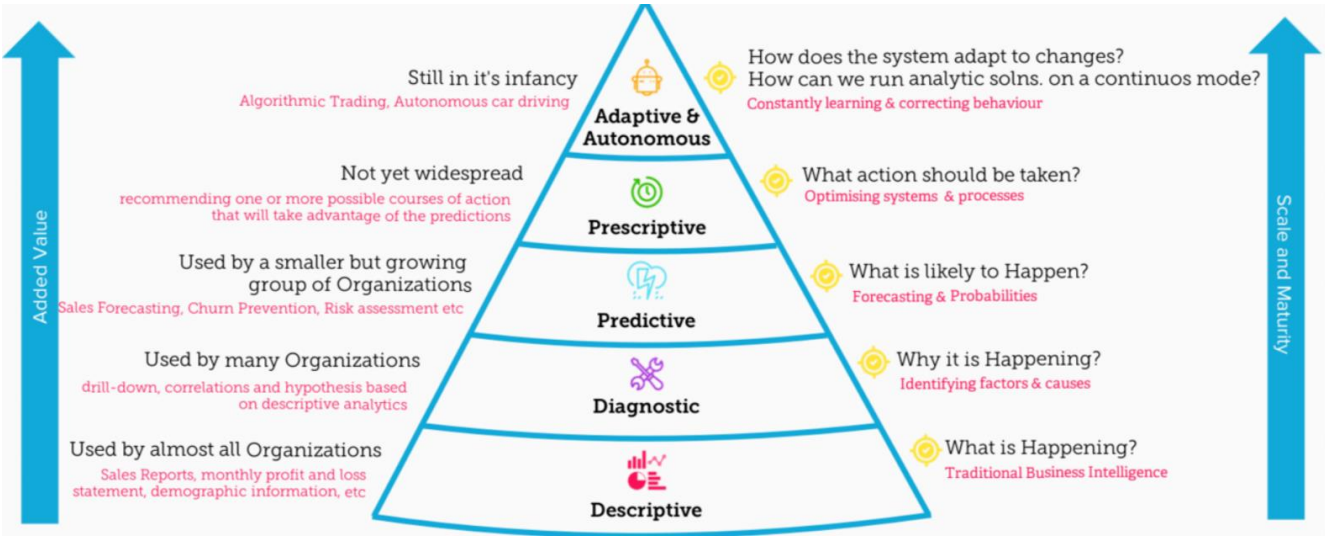
Data Analytics

- Data Analytics Uncovers trends and answers questions from raw data.
- Data Analytics Used across various fields with different techniques and goals.
- Data Analytics Helps businesses analyze past, predict future, and make informed decisions.
- Data Analytics Draws meaningful conclusions from complex data.
- Data Analytics Improves decision-making and business success.

Types Of Data Analytics?





1. Descriptive Analytics: What is happening?

- Descriptive analytics helps explain past performance through historical data analysis.
- It provides insights into the causes of success and failure.
- Management reports across various departments rely heavily on descriptive analytics.
- Unlike predictive models focused on individual behavior, descriptive analytics identifies broader customer-product relationships.

Focus: Analyze past data and events

Objective: Understand what happened and why

Applications: Management reporting, classification, customer segmentation

Common Tools: Data queries, Reports, Descriptive statistics, Data dashboards

2. Diagnostic Analytics: Why is it happening?

- Diagnostic analytics delves deeper than descriptive analytics, aiming to explain the reasons behind past occurrences.
- Companies benefit from existing historical data to diagnose problems and refine future strategies.
- Techniques like data discovery, mining, and correlation analysis are crucial for uncovering valuable insights.

Focus: Uncover the "why" behind past events

Data Source: Primarily historical data

Goals:

- Identify patterns and dependencies
- Gain insights into problems and solutions
- Improve future performance

Common Techniques:

- Data discovery (exploratory analysis)
- Data mining (uncovering hidden patterns)
- Correlation analysis (finding relationships between variables)

3. Predictive Analytics: What is likely to happen?

- ❖ Predictive analytics goes beyond understanding the past, aiming to anticipate the future and inform decision-making.
- ❖ Its diverse techniques leverage data to make informed predictions about various outcomes.
- ❖ Building predictive models, optimizing decisions based on predictions, and understanding transaction patterns are essential components.

Focus: Turning data into actionable insights by predicting future outcomes and events.

Data Source: Both historical and current data.

Methods: Employs various statistical and analytical techniques:

- Modeling: Building statistical models to represent relationships between variables and predict future values.
- Machine Learning: Algorithms that learn from data to identify patterns and make predictions.
- Data Mining: Uncovering hidden trends and relationships within large datasets.
- Game Theory: Analyzing strategic interactions between different entities to predict outcomes.

Common Techniques:

- **Linear Regression**: Predicting continuous values based on linear relationships between variables.
- **Time Series Analysis and Forecasting**: Predicting future values in a time series based on past trends and patterns.
- **Data Mining Techniques**: Association rule mining, clustering, anomaly detection, etc.

Core Components:

- ➢ Predictive Modeling: Developing models to predict future values or events.
- ➢ Decision Analysis and Optimization: Choosing the best course of action based on predicted outcomes and risks.
- ➢ Transaction Profiling: Identifying patterns and behaviors in transactions to predict future activities.

4. Prescriptive Analytics: What do I need to do?

- Prescriptive analytics is the pinnacle of data-driven decision-making.
- It empowers organizations to make proactive and informed choices.
- Its impact spans various industries, from healthcare to finance and beyond.

Focus: Optimize decision-making by combining prediction and recommendation.

Tools: Big data, mathematical science, business rules, machine learning.

Unique Value: Goes beyond prediction to suggest actions and implications.

Capabilities:

- Anticipates what, when, and why events will happen.
- Suggests decision options to exploit opportunities or mitigate risks.
- Illustrates the consequences of each decision option.

Example: Healthcare strategic planning:

Leverages operational, usage, economic, and demographic data.

Identifies patterns and trends to anticipate future needs.

Recommends resource allocation, staffing, and service adjustments.

Usage of Data Analytics:

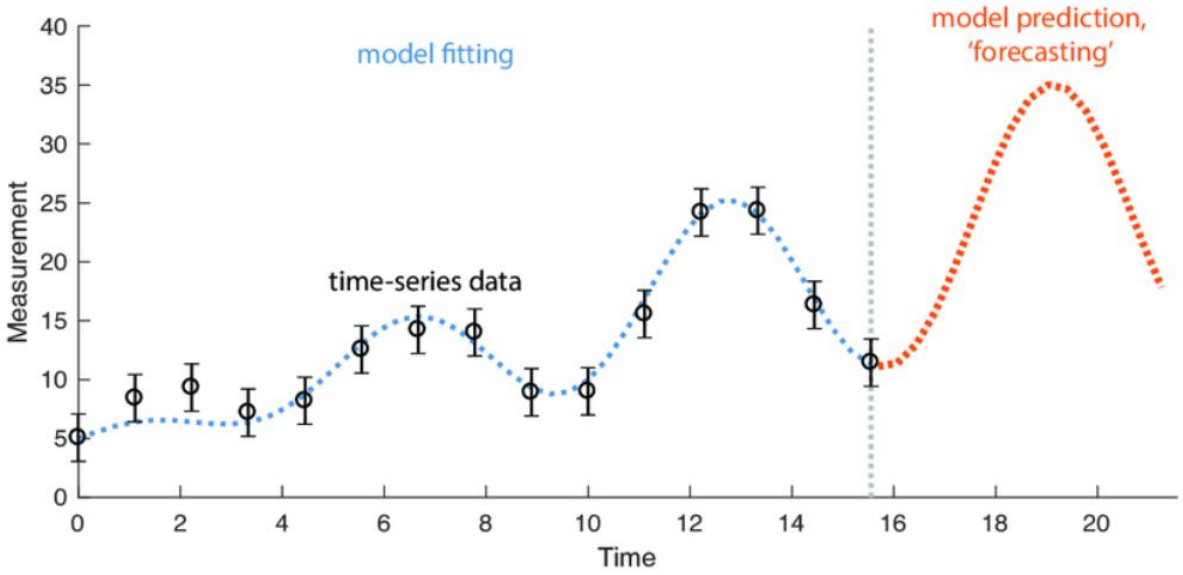
- **Improved Decision-Making:** Data supports effective and confident decision-making.
- **Better Customer Service**: Churn modeling helps retain customers by identifying churn risks and addressing them.
- **Efficient Operations:** Data analysis optimizes processes for better results and streamlined operations.
- **Effective Marketing:** Market segmentation targets the right audience for improved marketing strategies.

Future Scope of Data Analytics:

- **Retail:** Studying sales patterns, consumer behavior, and optimizing inventory management.
- **Healthcare:** Personalized treatment, patient outcome prediction, and cost reduction.
- **Finance:** Informed investment decisions, risk analysis, and performance optimization.
- **Marketing:** Personalized marketing, campaign efficiency improvement, and greater impact.
- **Manufacturing:** Enhanced production efficiency, cost reduction, and improved product quality.
- **Transportation**: Optimized logistics, efficient route planning, and reduced transportation costs.

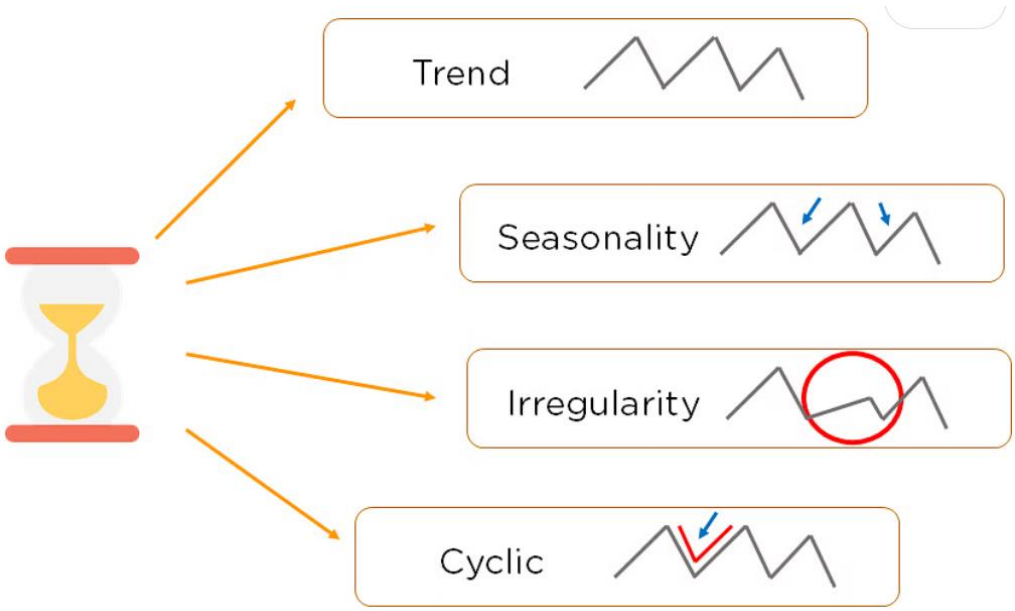
Time Series Analysis

- A Time-Series represents a series of time-based orders. It would be Years, Months, Weeks, Days, Horus, Minutes, and Seconds
- A time series is an observation from the sequence of discrete-time of successive intervals.
- A time series is a running chart.
- The time variable/feature is the independent variable and supports the target variable to predict the results.
- Time Series Analysis (TSA) is used in different fields for time-based predictions – like Weather Forecasting, Financial, Signal processing, Engineering domain – Control Systems, Communications Systems.
- Since TSA involves producing the set of information in a particular sequence, it makes a distinct from spatial and other analyses.
- Using AR, MA, ARMA, and ARIMA models, we could predict the future.



Terms in Time Series Analysis:

- **1.Time Series:** Sequence of data points collected at regular intervals.
- **2. Stationarity:** A time series with constant statistical properties over time.
- **3. Trend:** General direction of the data over a long period (upward, downward, or stationary).
- **4. Seasonality:** Repeating pattern within a fixed time period.
- **5. Cyclical Component:** Long-term, non-seasonal patterns (e.g., business cycles).
- **6. Irregular Variation:** Unpredictable fluctuations not explained by trend, seasonality, or cycles.
- **7. ETS Decomposition:** Separating time series into error, trend, and seasonality components.
- **8. Dependence:** Association between observations of the same variable at different time points.
- **9. Differencing:** Transforming a non-stationary series into a stationary one.
- **10. Specification:** Identifying linear or non-linear relationships using models like ARIMA.
- **11. ARIMA**: Autoregressive Integrated Moving Average model for time series forecasting.



Time Series Forecasting

Forecasting refers to the future predictions based on the time series data analysis. Below are the steps performed during time series forecasting

- Step 1: Understand the time series characteristics like trend, seasonality etc
- Step 2: Do the analysis and identify the best method to make the time series stationary
- Step 3: Note down the transformation steps performed to make the time series stationary and make sure that the reverse transformation of data is possible to get the original scale back
- Step 4: Based on data analysis choose the appropriate model for time series forecasting
- Step 5: We can assess the performance of a model by applying simple metrics such as residual sum of squares(RSS). Make sure to use whole data for prediction.
- Step 6: Now we will have an array of predictions which are in transformed scale. We just need to apply the reverse transformation to get the prediction values in original scale.
- Step 7: At the end we can do the future forecasting and get the future forecasted values in original scale.

Models Used For Time Series Forecasting

- Autoregression (AR)
- Moving Average (MA)
- Autoregressive Moving Average (ARMA)
- Autoregressive Integrated Moving Average (ARIMA)
- Seasonal Autoregressive Integrated Moving-Average (SARIMA)
- Seasonal Autoregressive Integrated Moving-Average with Exogenous Regressors (SARIMAX)
- Vector Autoregression (VAR)
- Vector Autoregression Moving-Average (VARMA)
- Vector Autoregression Moving-Average with Exogenous Regressors (VARMAX)
- Simple Exponential Smoothing (SES)
- Holt Winter's Exponential Smoothing (HWES)

ARIMA

- ARIMA stands for Auto-Regressive Integrated Moving Averages. It is actually a combination of AR and MA model.
- ARIMA has three parameters 'p' for the order of Auto-Regressive (AR) part, 'q' for the order of Moving Average (MA) part and 'd' for the order of integrated part.

Auto-Regressive (AR) Model:

- As the name indicates, its the regression of the variables against itself. In this model linear combination of the past values are used to forecast the future values.
- To figure out the order of AR model we will use PACF function

Integration(I):

- Uses differencing of observations (subtracting an observation from observation at the previous time step) in order to make the time series stationary. Differencing involves the subtraction of the current values of a series with its previous values d number of times.
- Most of the time value of d = 1, means first order of difference.

Moving Average (MA) Model:

- Rather than using past values of the forecast variable in a regression, a moving average model uses linear combination of past forecast errors
- To figure out the order of MA model we will use ACF function