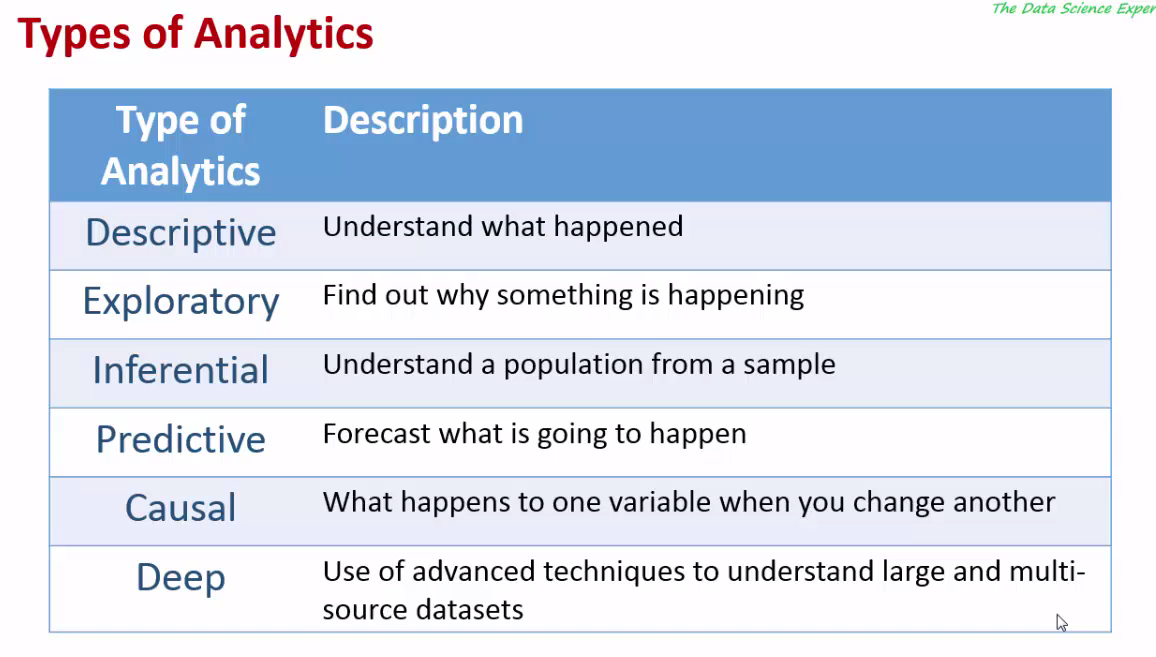
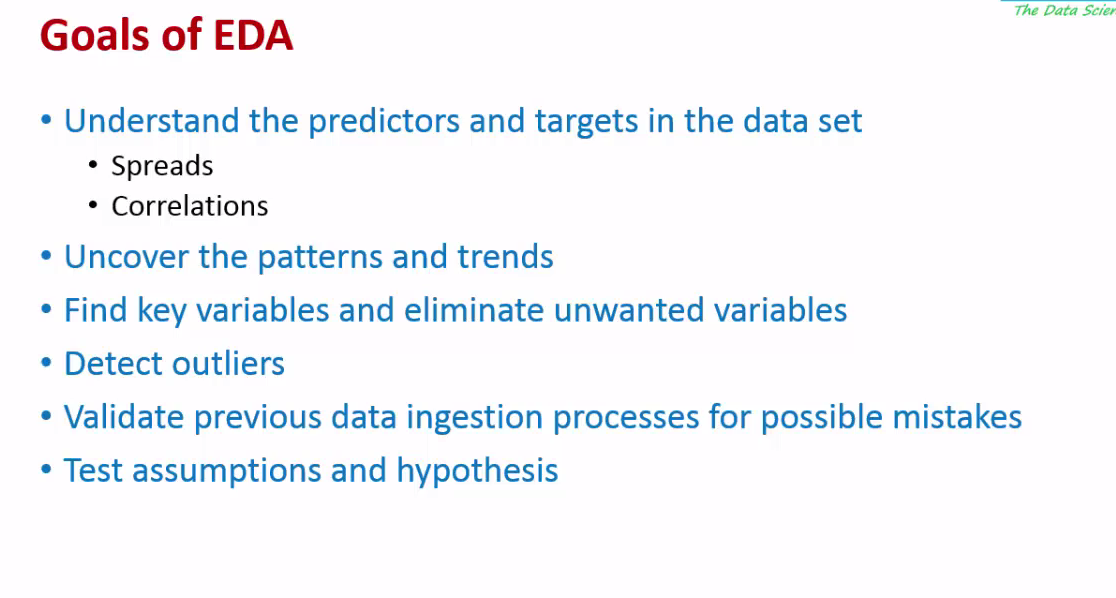
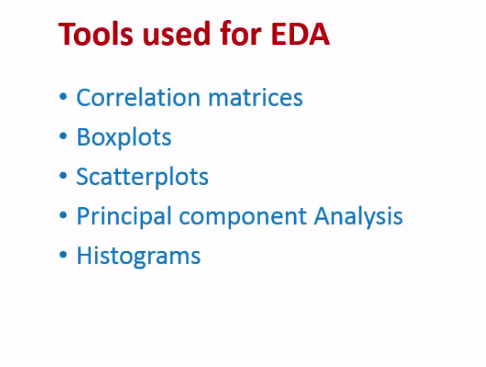
1. What are the types of Analytics?



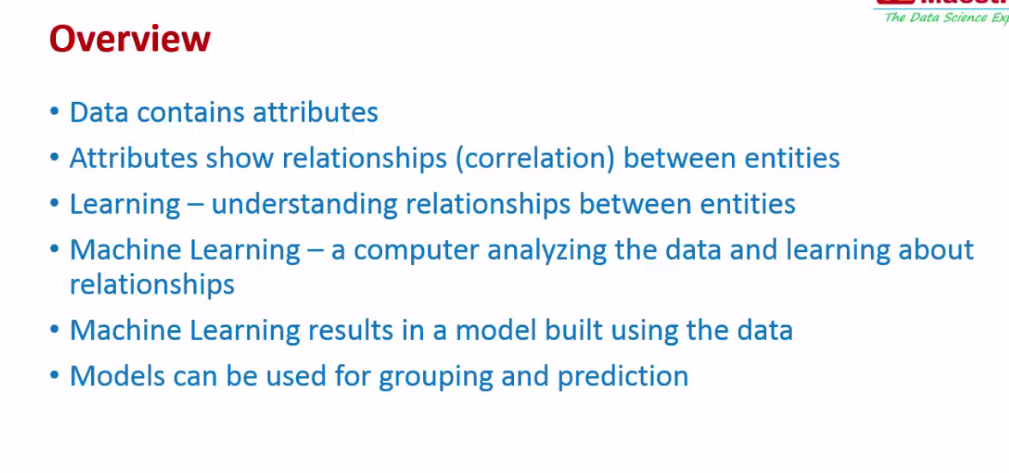
1. Goals of Exploratory DataAnalysis (EDA)?



1. Tools used for EDA?



1. Over view of machine learnings?



1. What does **learning** means in Machine Learning?

Learning: Understanding relationship between the entities.

1. What is machine learning?

A computer analyzing the data and learning about the relationships.

1. What is the input for machine learning? (This is at high level)

Data.

1. What is the output of machine learning?

Model which is built using data.

1. What is Model in Machine learning?

Definition or explanation of the relationships between the various attributes.

*Model can be an equation* as to specify how you can derive one variable from the other.

*Model can be a decision tree* by using this decision tree on these variable values how can you derive at final target.

1. For what purpose we can use model?

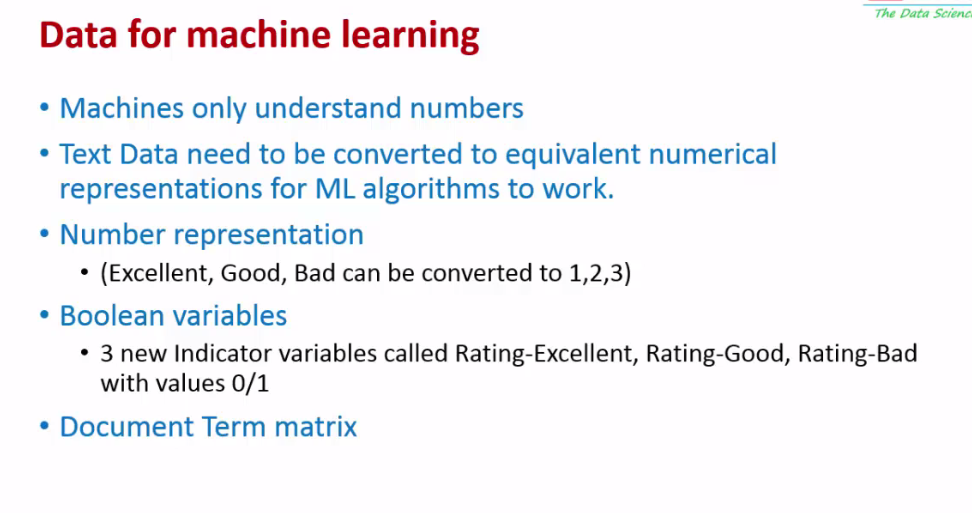
We can use model for

1. Grouping of similar Data like similar customer(One type of Recommendation is done based on this), similar products
2. Predictions
3. **What is the Actual Data / input for machine learning?**

Machines can understand only numbers. So we need convert the text data to numerical representations for Machine Learning (ML) algorithms.

For Example:

The classifications like Excellent, Good or Bad can be converted to 1, 2, and 3.



1. **What are the types of learning?**

There are 2 types of learning.

1. Supervised learning: In case of supervised learning there is a target variable and that target variable is specified by external person.
2. Unsupervised learning: In case of unsupervised learning there is no target variable. Instead finding hidden structure of data, finding similarities in the attributes and grouping them
3. **What are the types of unsupervised learning?**
4. Clustering:

Group based on data

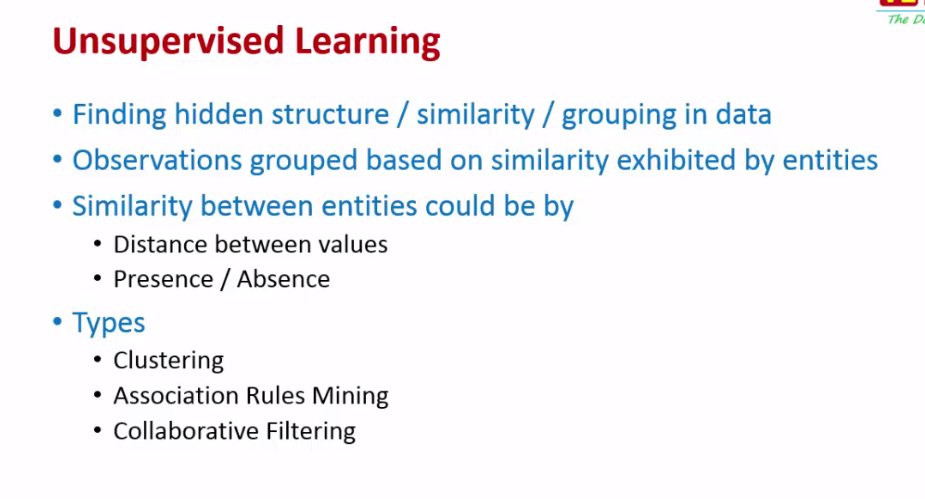
1. Association rules mining:

How things are used together. For example in case of Super market analysis what kind of products are bought together.

1. Collaborative Filtering:

To find similarity between people. Similarity between items based on usage.

For example: If you go to amazon then finding the people who were trying to do similar thing like who by similar kind of products, give similar kind of comments



1. **What is supervised learning?**

In case of supervised learning, we try to predict unknown attributes (outcomes) based on known attributes (predictors) for an entity.

038 Types of Machine Learning.mp4

1. **How model is built?**

Models are built on training data (History data). In History data outcomes and predictors are known. Model is mainly used to predict future outcomes.

1. **What are the types of Supervised Learning?**
2. Regression: (Continuous outcome values)

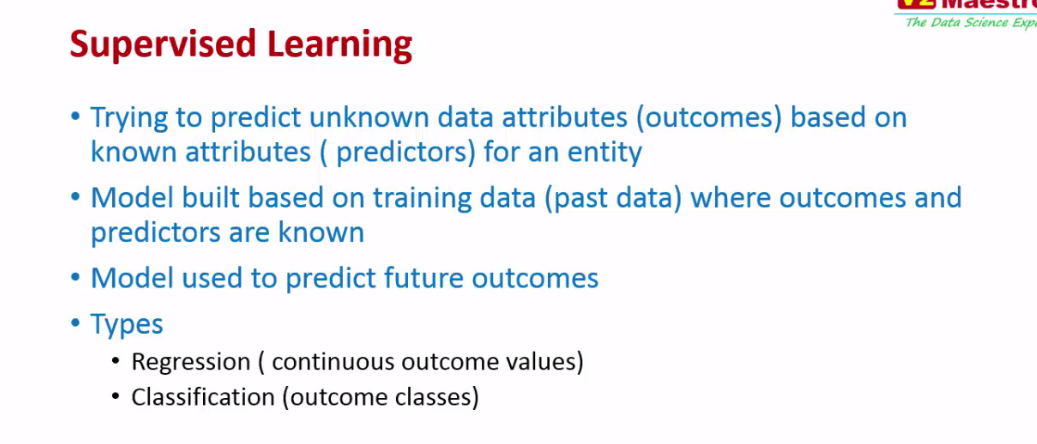
In case of Regression we try to predict values.

For example: Predict age of a person, predict price of something, Predict total value of something

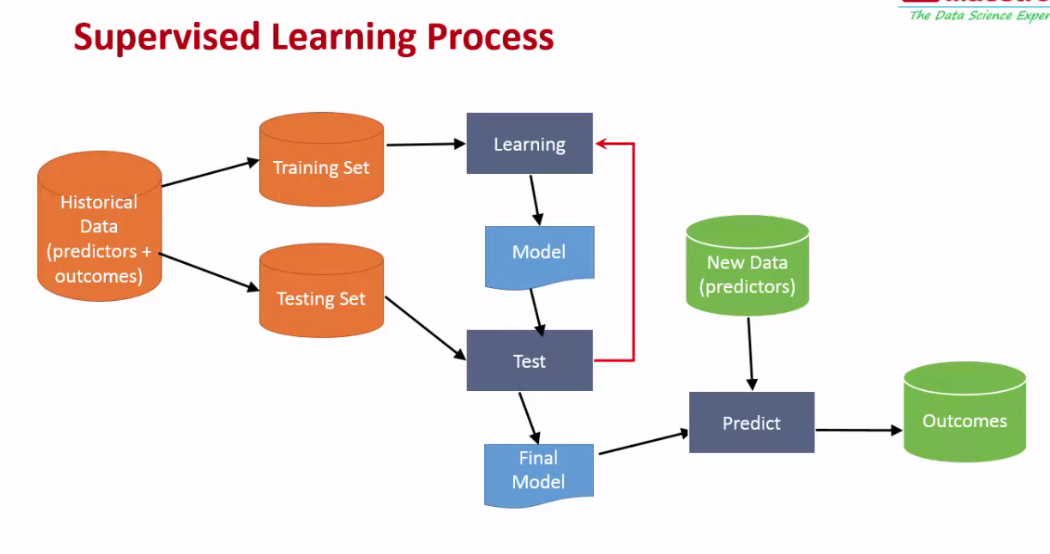
1. Classification: (Outcome classes)

In case of classification we predict class or group the person might belong to.

For example: Whether person will buy something or will not buy. Predict If bank customer can be gold customer or silver customer.



1. **Supervised learning process? How supervised learning process is performed?**



1. **Historical Data:**

We should use significantly large amount of data for this if predictions needs to be great and quality of data also should be good.

1. **Training Set (data):**

This is part of Historical data, using which machines learns and build model.

This data will have both predictors and outcomes.

1. **Testing Set:**

This is part of Historical data

This data will have both predictors and outcomes.

1. **Learning: (Machine learning algorithm)**

Training data set will be input for Machine learning algorithm. Then Machine learning algorithm will come out with a model

1. **Testing:**

Using the Testing data set and Model outcome from learning will be tested for predictions.

The predictions are matched with outcomes of testing data because testing data already knows the outcome.

In this case we can know how good our prediction is if prediction is not good then need to go back and tweak learning process. It is iterative process can be performed till we get quality predictions from the model.

How can I tweak learning process?

Eliminate some variable. Add some new variable. Try some technique like creating indicator variable. Or Try different machine learning algorithm or for the same machine learning algorithms tweak some parameters.

1. **Final model:**

After learning, testing and tweaking data and algorithm we get model which gives quality predictions and this model is used during predictions.

1. **New Data:**

This is the incoming data other than history data in which there will be only predictors and we need to predict outcomes from that

1. **Predict:**

Predict can be program/algorithm which validates the new data with Model and gives prediction.

1. **Outcome:**

The prediction done using model and new data is outcome, which is actual prediction.

1. **What is the good practice to split the Historical data to Training data and testing data?**

Need to split data to 70-30

70% of data for training and 30% of data for testing.

Need to randomly select the data but need to make sure data is spread equally between both the data sets

We should divide Training set and Testing set such that both should retain characteristics of Historical data.

For example: In the historical data say 30% of the customer bought the product and 70% did not. So when we split the data to training and testing data set they individually should have same ratio.

1. **Confusion Matrix?**

039 Analyzing Results and Errors.mp4

1. **Where machine learning is used?**
2. Machine learning is used in predictive analysis.