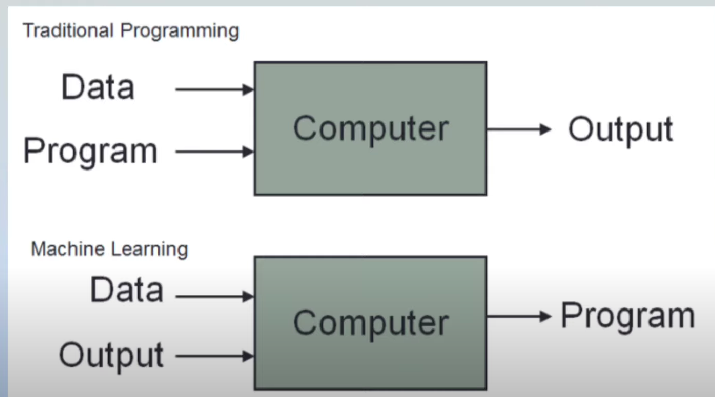
**Machine Learning**

Machine learning is a field of computer science that uses statistical techniques to give computer systems the ability to "learn" with data, without being explicitly programmed.



Types

**Supervised ML**- input and output both.

Regression

Classification

**Unsupervised ML-** Only input

Clustering

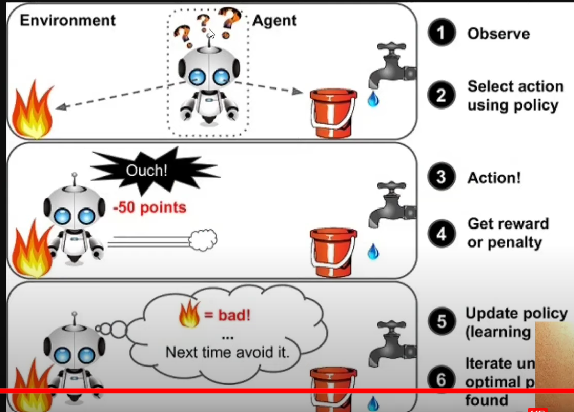
Dimensionality Reduction

Anamoly Detection

Association Rule Learning

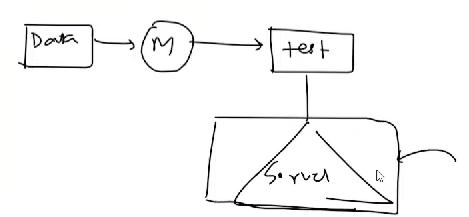
**Semi-supervised-** Ex Google photo label

**Reinforcement-** self driving car

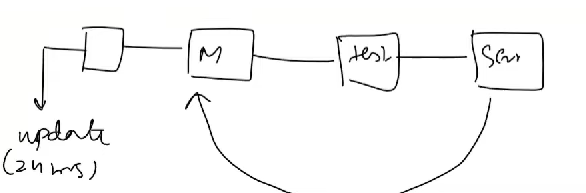


**4**.**Batch Machine Learning (offline) vs Online Learning**

**Offline Learning**



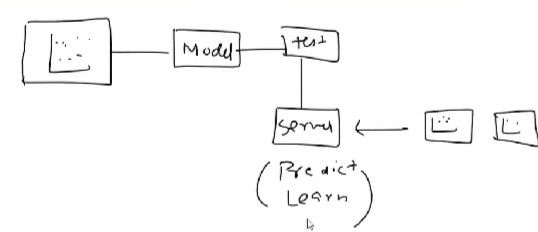
The problem



Disadvantage



**5. Online ML Learning**



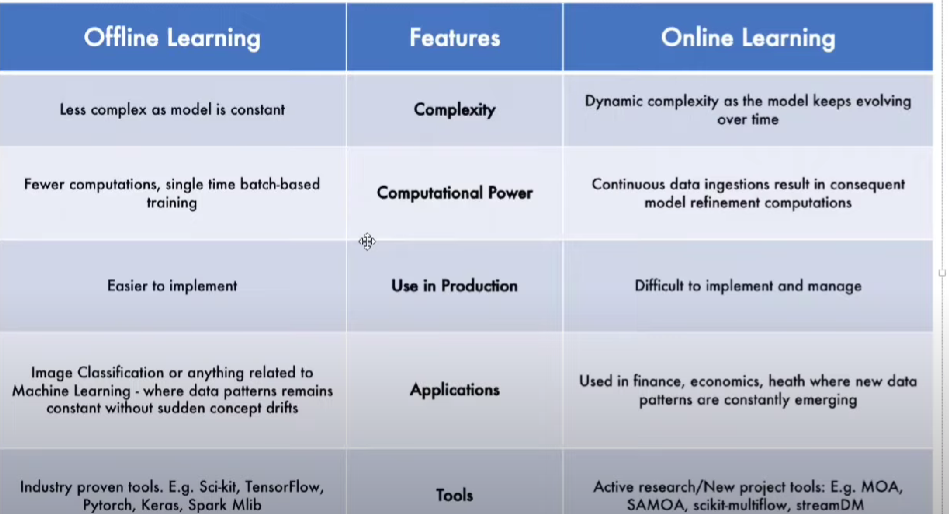
Example: Chatbot, youtube.

When to use Online Learning?

1. Where there is a concept drift
2. Cost Effective
3. Faster Solutions

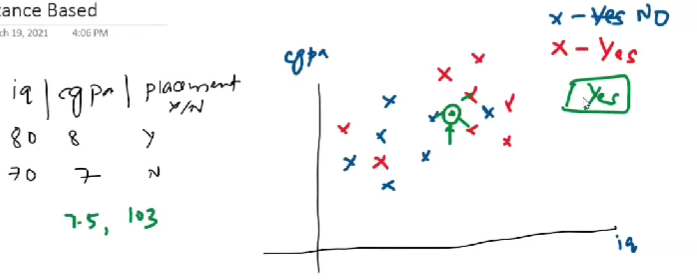
Disadvantage:

* Tricky to use
* Risk



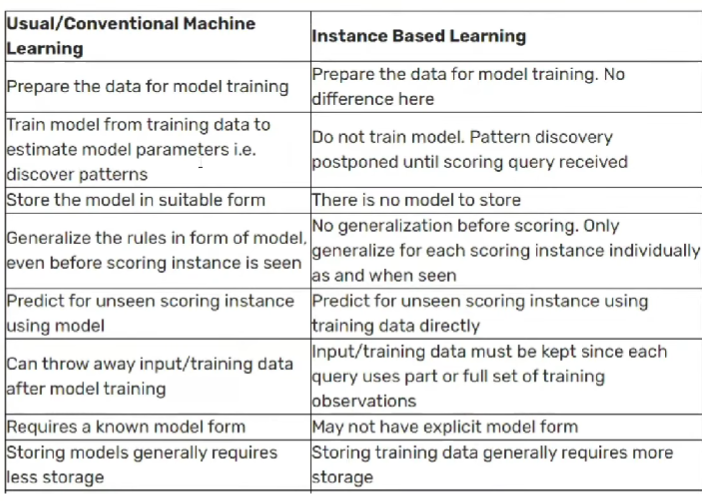
**6. Instance – Based Vs Model-based Learning**

Instance Based

Example KNN,

Model -Based Learning

Example – Linear regression



**7. Challenges in ML**

1) **Data Collection:**

2) **Insufficient Data/Labelled:**

3) **Non-Representation Data**: Sampling noise, sampling bias

4) **Poor Quality Data:**

5) **Irrelevant Features:** Garbage In Garbage Out

6) **Overfitting:**

7) **Underfitting:**

8) **Software Integration:**

9) **Offline Learning/Deployment:**

10) **Cost Involved:**

**8. Application of Machine Learning**

* **Retail – Amazon/Big Bazaar:**
* **Banking and Finance:**
* **Transport – QLA:**
* **Manufacturing – Tesla:**
* **Consumer Internet – Twitter:**

**9. Machine Learning Development Life Cycle (MLDC)**

**1) Frame the Problem:**

**2) Gathering Data:** CSV, API, web scraping, Database (Data warehouse)

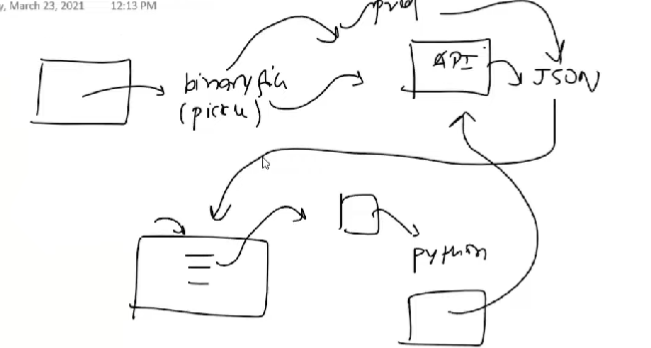
**3) Data Preprocessing:**

**4) Exploratory Data Analysis**: visualization, univariate / Bivariate, outlier detection, Imbalance

**5) Feature Engineering and selection:**

**6) Model Training Evaluation and selection:** Ensemble learning

**7)Model Deployment:**

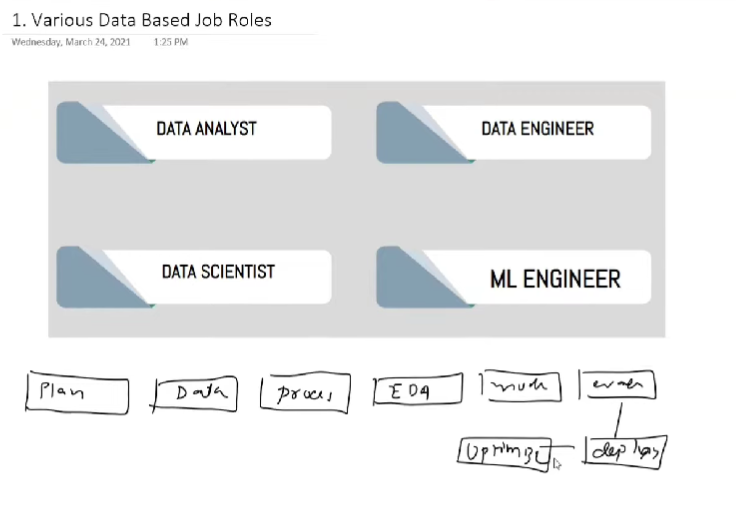


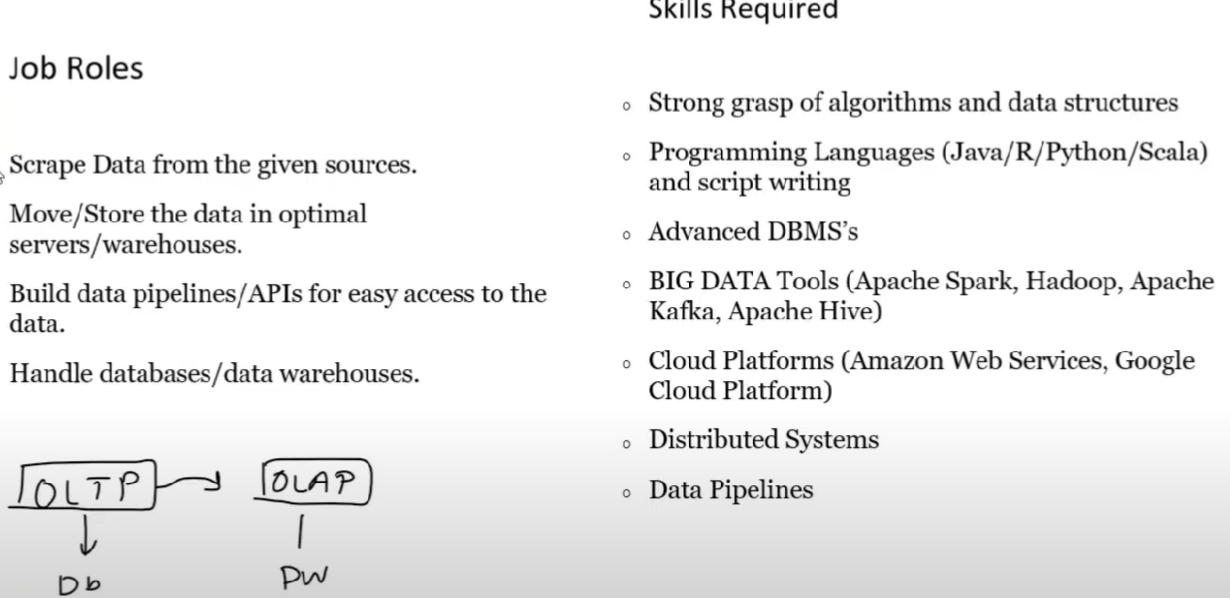
**8) Testing:** A/B testing

**9) Optimize:** Backup, Data, Load balancing

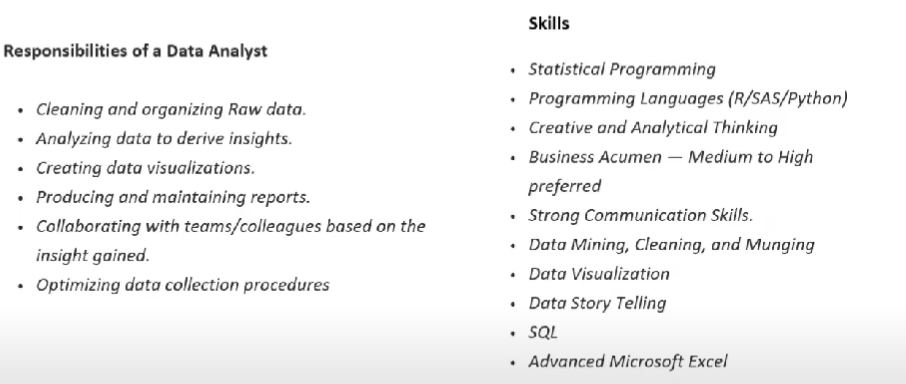
**10 Data Engineer vs Data Analyst vs Data Scientist Vs ML Engineer**

**Various Data based**





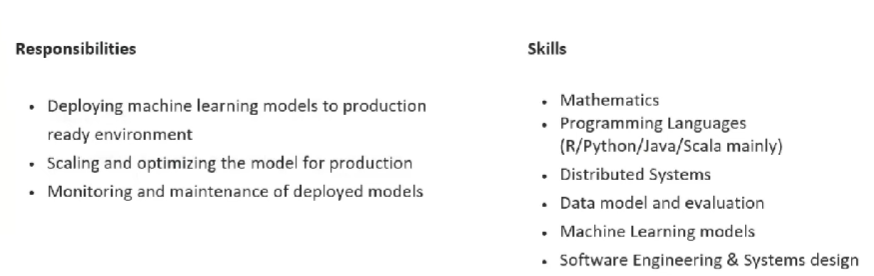
**Data Analyst**

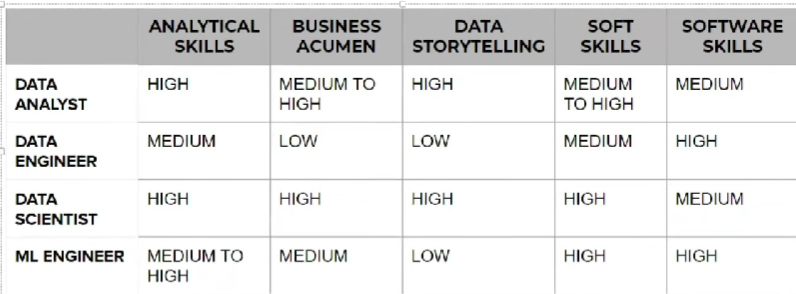


**Data Scientist:**

"A data scientist is someone who is better at statistics than any software engineer and better at software engineering than any statistician".

**ML Engineer**





**11. What are Tensors**

**One Dimension Tensor/vector**: [1,2,3,4]

**Two Dimension Tensor/Matrices: [1,2,3] [4,5,6] [7,8,9]**

**N- Dimension Tensor:**

**Rank, Axis and Shape:**

**Example of 1D Tensors:**

**Example of 2D Tensors**:

**Example of 3D Tensors**: Time series Data (Highest | lowest)

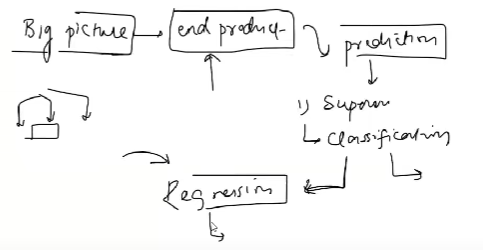
**Example of 4D Tensors**: Images

**Example of 5D Tensors**: Videos

**14. How to Frame a Machine Learning Problem**

1) Business Problem to ML Problem:

2) Type of Problem:



3) Current Solution:

4) Gathering Data:

* Watch time
* Search but did not find
* Content left in the middle
* Clicked on recommendations (order of recommendation)

5) Metrics to measure:

6) Online vs Batch:

7) Check Assumptions

15. Working with CSV files

Data Gathering: CSV, JSON/SQL, Fetch data from API, Web scraping

<https://pandas.pydata.org/docs/reference/api/pandas.read_csv.html>

16 Working with JSON/SQL

<https://pandas.pydata.org/docs/reference/api/pandas.read_json.html>

17. Fetching Data from an API

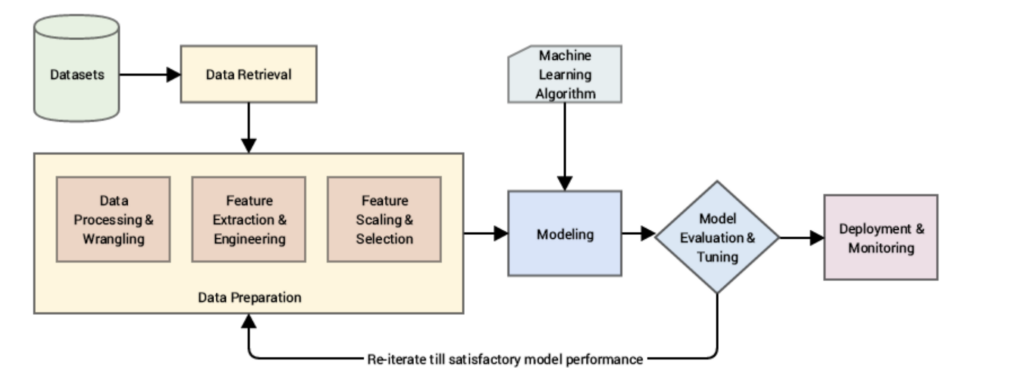
What is an API?

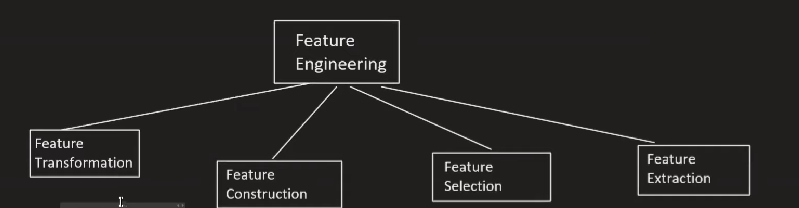
<https://developer.themoviedb.org/docs/getting-started>

<https://rapidapi.com/collection/list-of-free-apis>

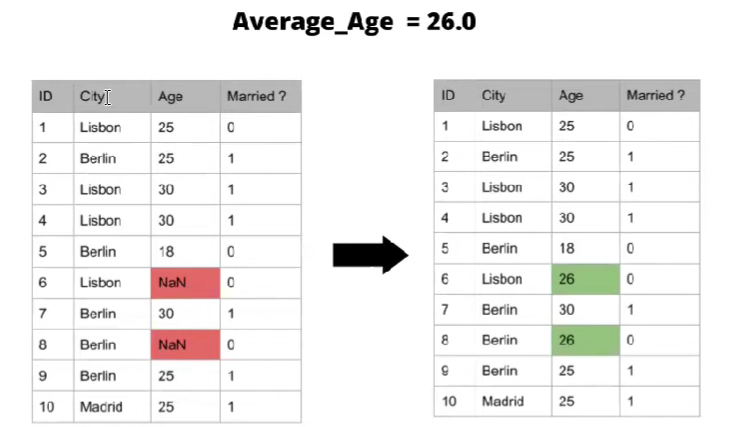
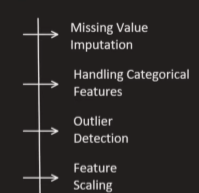
<https://jsonviewer.stack.hu/>

23 What is Feature Engineering

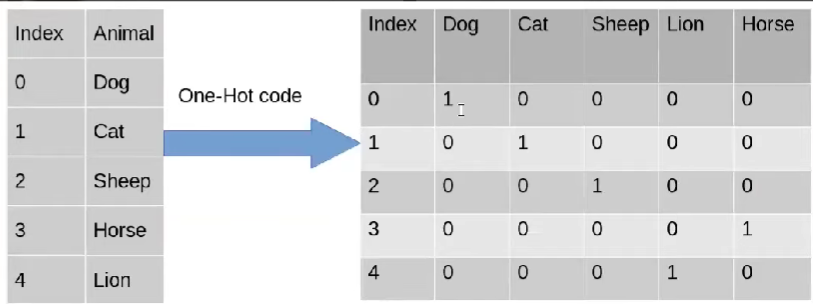
Feature engineering is the process of using domain knowledge to extract features from raw data. These features can be used to improve the performance of machine learning algorithms.  




1.1 Missing values imputation

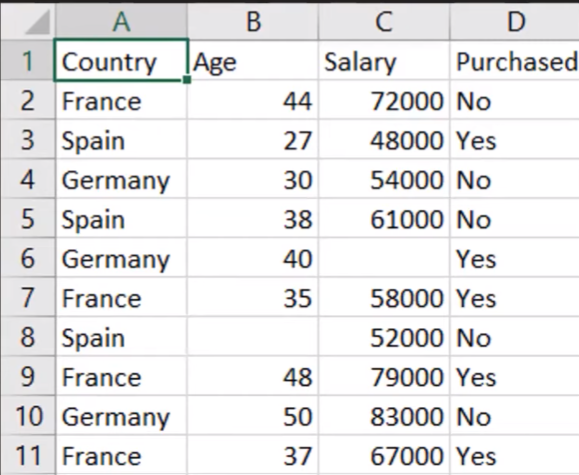


1.2 Handling categorical Values

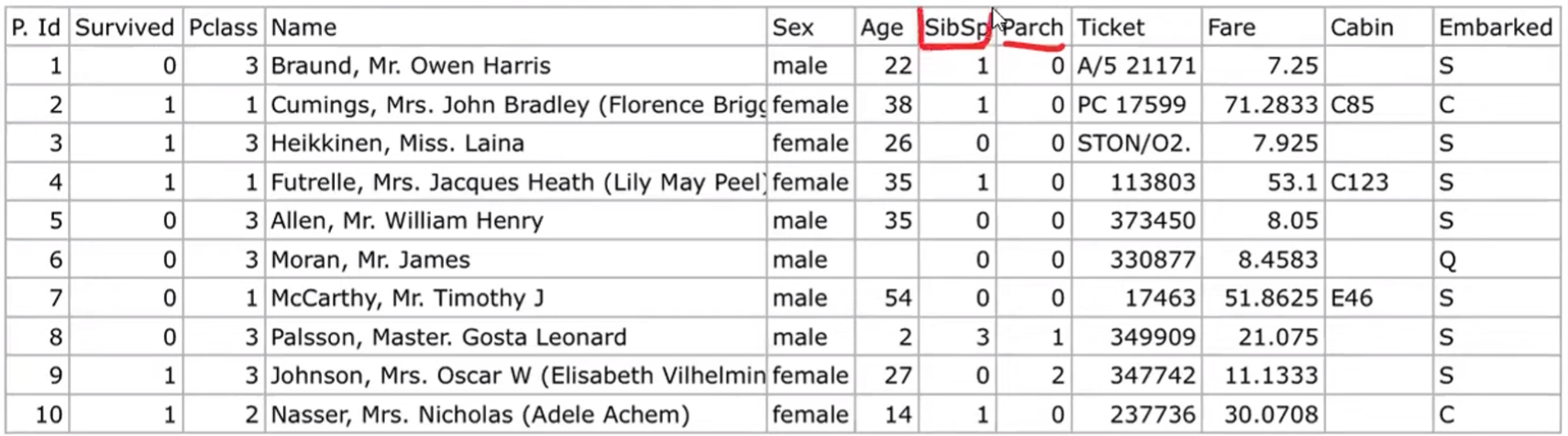


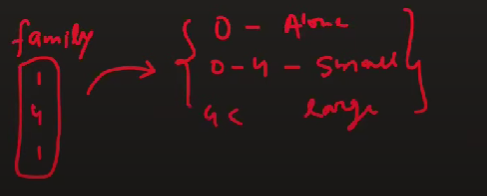
1.3 Outlier Detection

1.4 Feature Scaling

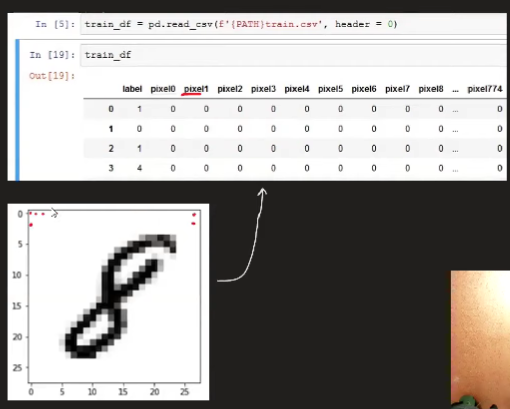


2. Feature Construction

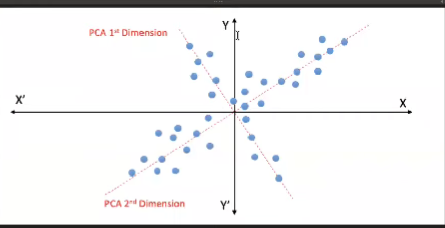




3) Feature Selection:

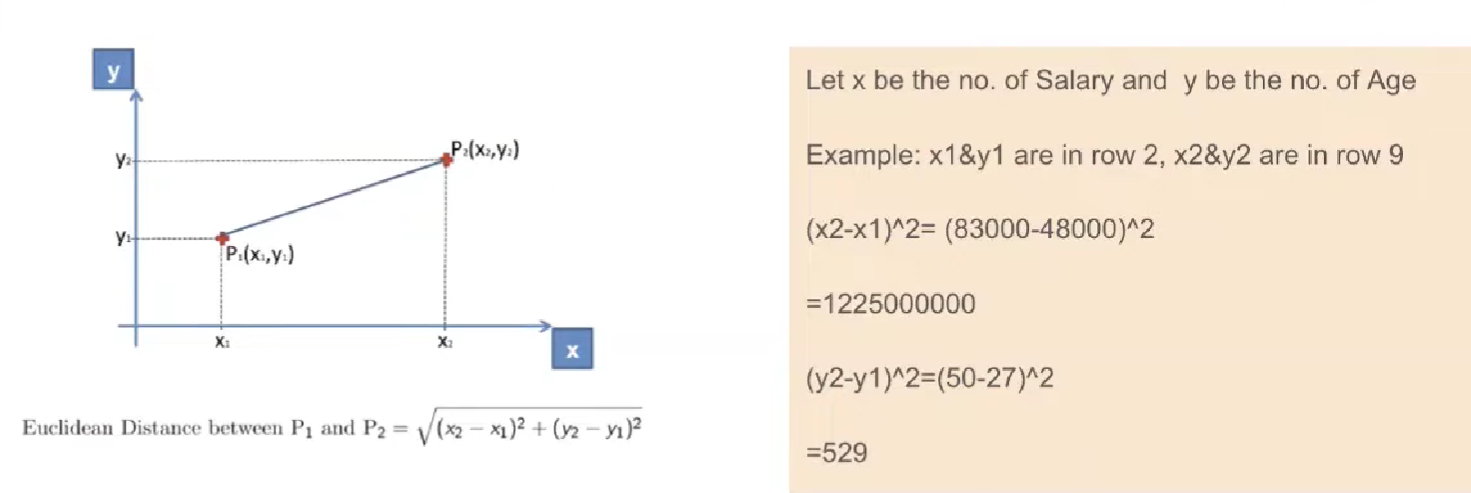


4) Feature Extraction:

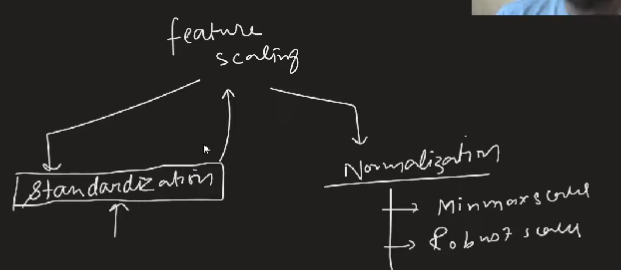


# 24 Feature Scaling - Standardization

Why do we need Feature Scaling?



Types of Feature Scaling



Standardization – Intuition

Also called as Z – score Normalization

Example:

