Data Science

-We require understanding different things to understand the world. This we do through data.

-We Uncover the insights and hands that are hiding behind the data and does find suitable and relevant explanation from it.

(First point is to prove any data)

-Generally there are lot of languages and frameworks available for this field but R and PYTHON are mainly used.

**Why Data Science got this popular now?**

-Availability of data

-Algorithms

-Software

-Problems

-Open Source

-Kaggle(as a great platform)

**To be a Data scientist one must be :-**

-Curious

-Judgemental

-Story Telling

-Argumentative

**Prerequisites:-**

-Probability -Statistics -Linear Algebra -Programming(R or Python) -Data sets

**Steps:-**

-Data Cleaning.

-Analyze Data.

-Organizing data.

-Draw out inferences from the data.

**Technologies:-**

-Regression(Predict coming data by help of previous data.)

-Visualization

-Neural Networks

-Nearest Neighbour

suppose:-

f(x)=kx^2+kx+p where k and p are constant .

This this helps us to predict the nature of data using constants and variable.

**Questions which must be asked as a data scientist:-**

-What you are trying to measure?

-What do you require?

-How will you do?

-What?

**Some knowledge which helps you to work with data:-**

-Cloud

-Advanced computing algorithms

**Big Data**

Big data refers to data sets that are too large or complex for traditional data-processing application software to adequately deal with. Data with many cases offer greater statistical power, while data with higher complexity may lead to a higher false discovery rate.

Never overwrite a data

(archive never data)