G. H. Raisoni College Of Engineering And Management, Wagholi Pune Assignment no :- 1 2021- 2022			
Department	CE [SUMMER 2022 (Online)]		
Term / Section	III/B	Date Of **submission**	01-10-2021
Subject Name /Code	Python for Data Science / UCSP204		
Roll No.	SCOB77	Name	Pratham Rajkumar pitty
Registration Number	2020AC0E1100107		

Assignment No 1

Aim :→

Introduction to data Science and various python Spyder Tool with its some basic operation and various Phases in data Science

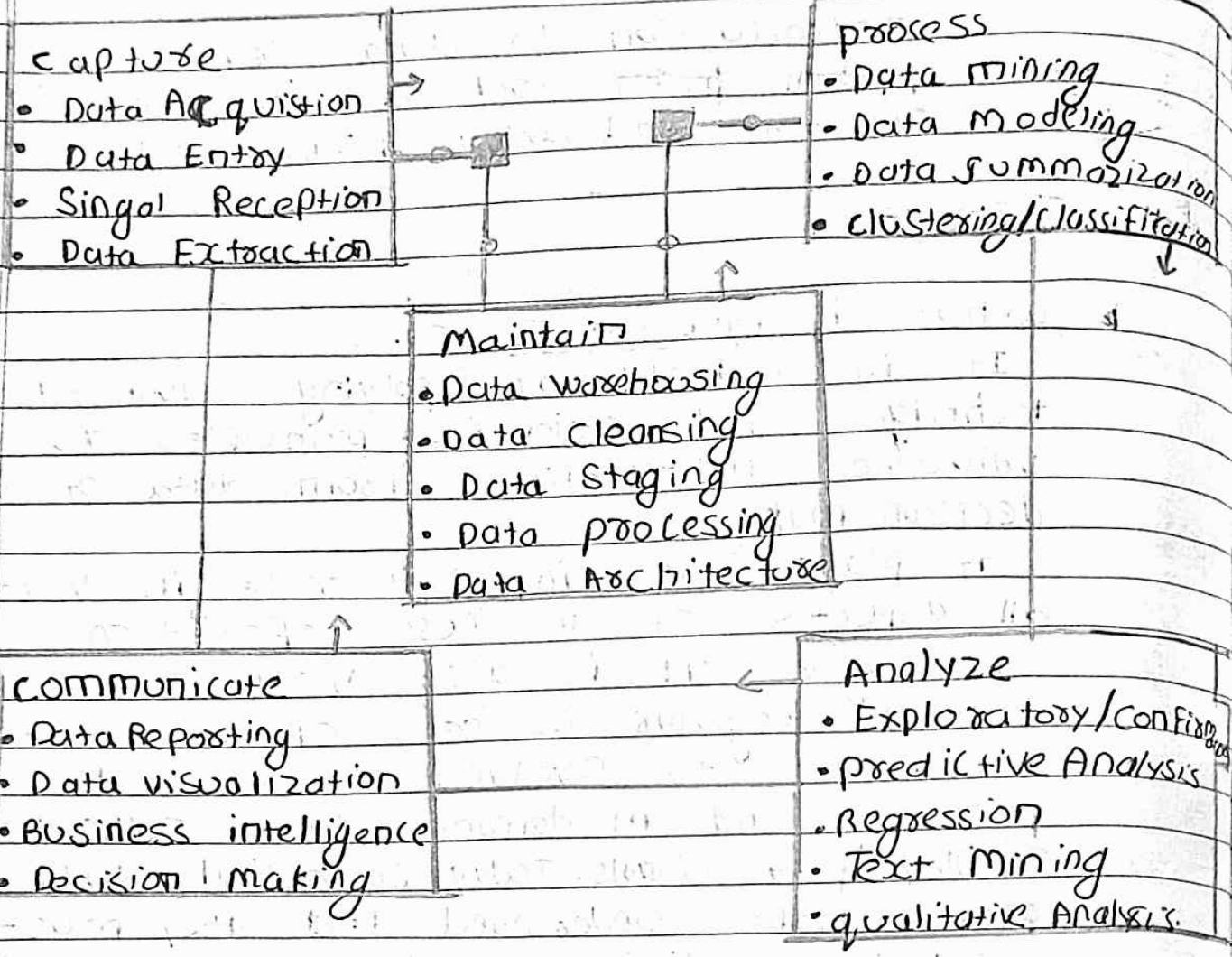
Theory :→

What is Data Science?

It is a field of applying advanced analytics technigies and scientific principles to extract valuable information from data of bussiness decision making.

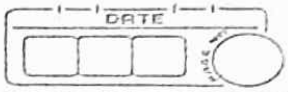
It plays an important role in virtually all aspects of bussiness operations and strategies. It is also vital in areas beyond regular bussiness operations. It cons continues to evolve as one of the most promising and in demand carrer paths for skilled professionals. Today successful data professionals understand that they must advance past the traditional skill of analysing large amount of data, data mining, and programming skills in order to uncover useful intelligence for their organization.

The Data Science life cycle



* (1) Capture :→ The very first step of a data science project is straightforward. We obtain/capture the data that we need from available resources. In this step we gather data with the help of data acquistom, data entry, data Extraction, etc.

* (2) maintain :→ After obtaining the data, the next thing to do is processing /scrabbing data. This process helps us to clean and filter the data.



If the results of the analysis will not mean anything we can process data with the help of mining, modelling and summerising the data.

* (3) process :→ After obtaining data we maintaing the data thereafter we have to process the data according to our needs.

* (4) Analyze :→ After processing the data we have to analyise the data by the means of various techniques i.e regression, text mining, qualitative analysis etc

* (5) Communicate :→ After Analyzing the data we have to communicate with the bussiness intelligence data report to know about the data and now we can grow our bussiness.

▶ Data Scientist :-

Data scientist examine which questions need answering and where to find the related data.
They have bussiness acumen and analytical skills as well as the ability to mine, clean and present data. Business used data scientist to sourse manage and analyze large amounts of ~~data~~ unstructured data.

TOOLS

* Skills needed For Data Scientist :→
- programming skills (SAS, R, Python)
- Statistic and mathematical skills
- Storytelling and data visualization
- Hadoop
- SQL
- machine learning.

▶ Data Analyst :→

Data analysts bridge the gap between data scientist and business analysts They are provided with the questions that needs answering from an organization and then organize and analyze data to find results that align with high-level business strategy Data analysts are responsible for translating technical analysis to qualitative action items and effectively communicating their findings to diverse Stake holders.

Tools :

* Skills needed :→ • programming Skills (SAS, R, Python) statical. Statistical and mathematical Skills • data wrangling, • data visualization.

▶ Data Engineer :→

Data Engineers manage exponential amounts of rapidly changing data. They focus on the development, deployment, management, and optimization of data pipelines and infrastructure to transform and transfer data to data scientists for querying

TOOLS

* Skills needed :→ • programming languages (Java, Scala), • NoSQL database (MongoDB, Cassandra DB), • frame works (Apache Hadoop).

* There are various applications of Data Science like
- In Education
- Airline root planning
- Healthcare industry
- Delivery Logistics
- Banking and Finance
- Filtered internet search
- Product Recommendation System
- Digital Marketing and Advertising

Airline root planning

In Now a days Data Science is revolutionizing business activities in the Airline industry

previously due to competition, the Airline companies used to provide Huge discount to customers to attract them. Also due to the high rates of fuels and the lack of proper analysis for the delay of flight, destination, Rates, etc airlines were spending a lot on extra fuel. Now,

With the help of data science, the airlines have started analyzing the data of their business to strategically improve the sectors, due to which they were bearing Heavy losses.

* Data Science Helps airlines in following ways
- Identifying potential customers to offer calculated discounts, instead of providing discounts to all everyone.

- Deciding on the optimized route by analyzing the traffic on different routes. It helps in saving expensive fuel that gets unnecessarily exhausted otherwise.
- Predicting delays in flight
- Setting the cost of flight as per seasons, festivals, and ~~different routes~~. number of travelers This is done by analyzing ~~delays in flight~~ the number of potential travelers and frequent travelers.

Conclusion: Hence conclude that we learn the basics of data science.