ASSIGNMENT:2

Submited By,

SANIL BABU .P

Roll No: 50

S6-IT

ITT 306 - Data science

Q. Why Python Data analysis

1. Ease of Learning

Python has a simple and readable syntax, making it accessible for beginners. It's easy to understand and write code, which reduces the learning curve for data analysis tasks.

2. Rich Ecosystem

Python has a vast ecosystem of libraries and frameworks specifically designed for data analysis, such as Pandas, Numpy, matplotlib, and scipy. These libraries provide powerful tools for data manipulation, visualization, and statistical analysis.

3. versatility

Python is a versatile language that can be used for various purposes beyond data analysis, including web development, automation, machine learning, and more. This versatility makes it a valuable skill for professionals in various fields.

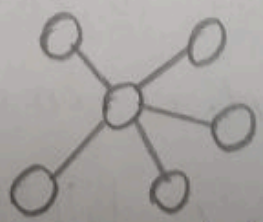
4. community support

Python has a large and scientists who contribute to opensource projects and provide support through forums, tutorials, and doccumentation. this community-driven approach ensures that there are resources available for solving almost any data analysis problem.

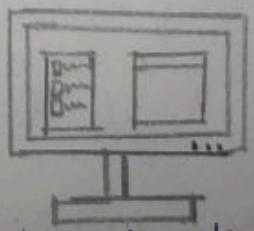
5. Integration

Python integrates well with other technologies and tools commonly used in data analysis, such as databases, cloud services, and big data platforms. This interoperability allows data scientists to seamlessly incorporate Python into their existing workflows.

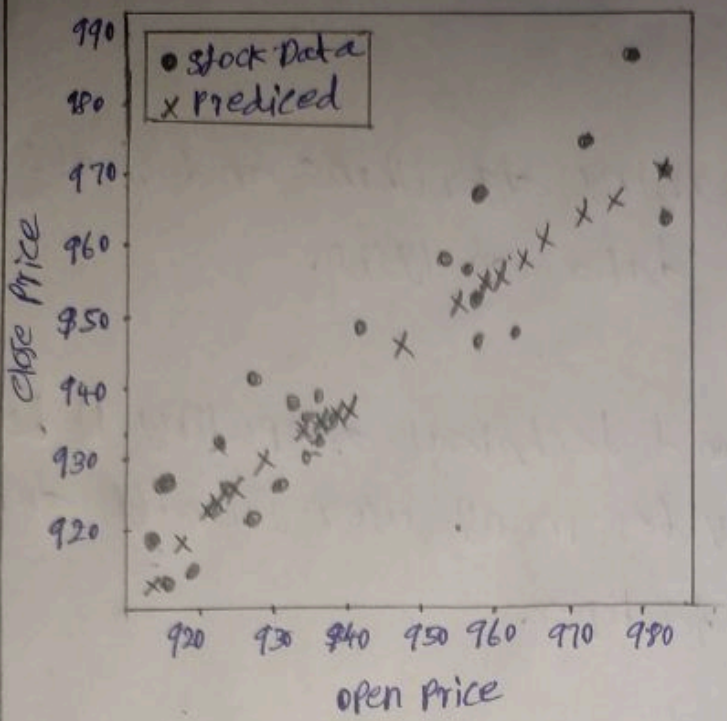
Data mining



Data modeling



Data visualization



Q. Fundamental Python libraries for data science.

1. Numpy

For numerical computing and handling arrays and matrices.

2. Pandas

For data manipulation and analysis, especially with structured data.

3. matplotlib

For creating static, interactive, and animated visualizations in Python.

4. seaborn

Built on top of matplotlib, it provides a high-level interface for drawing attractive and informative

statistical graphics.

5. Scikit-learn

For machine learning algorithms and tools for data mining and data analysis.

6. scipy

For scientific and technical computing. It builds on numpy and provides many user-friendly and efficient numerical routines.

7. Statsmodels

For statistical modeling and hypothesis testing.