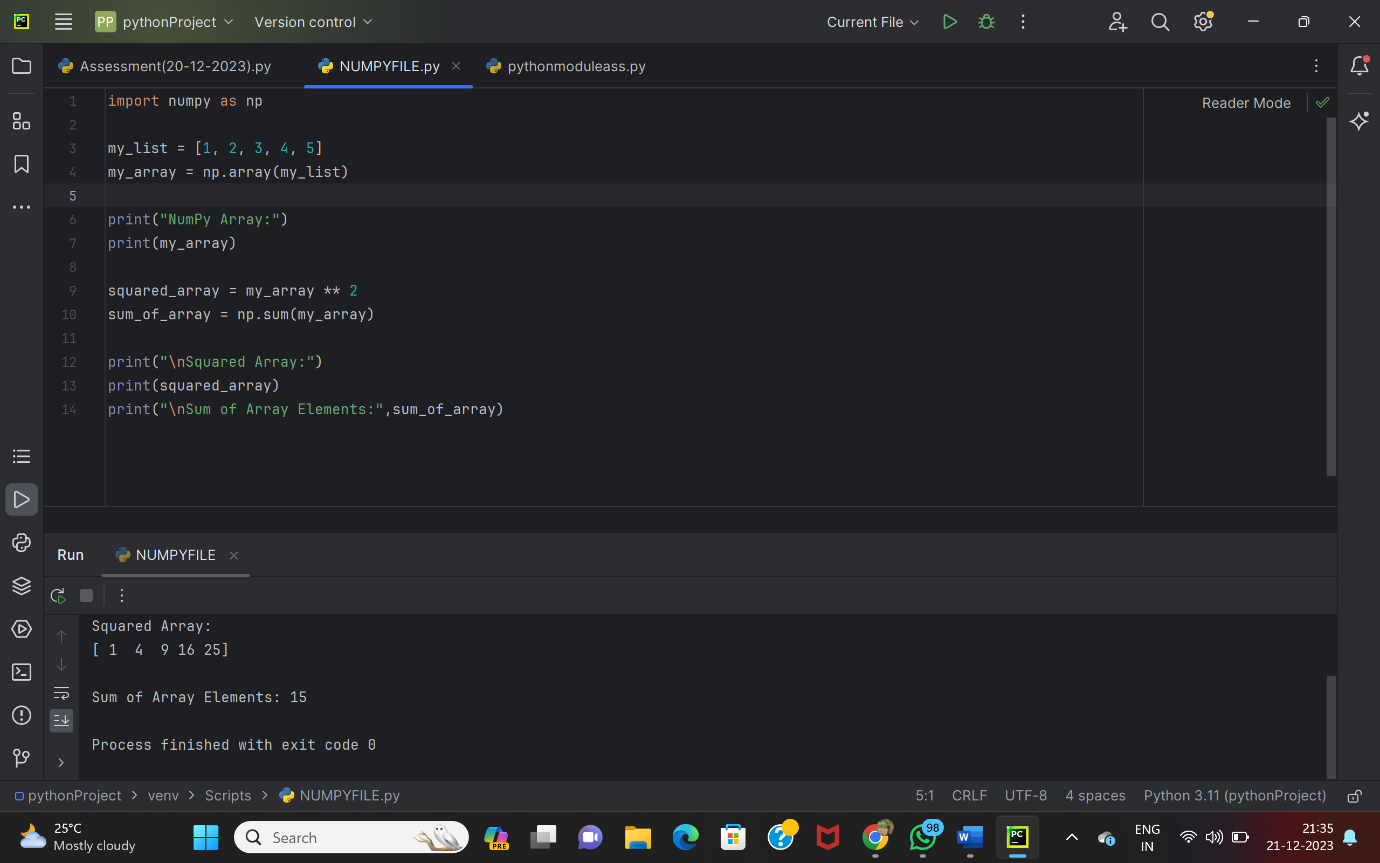
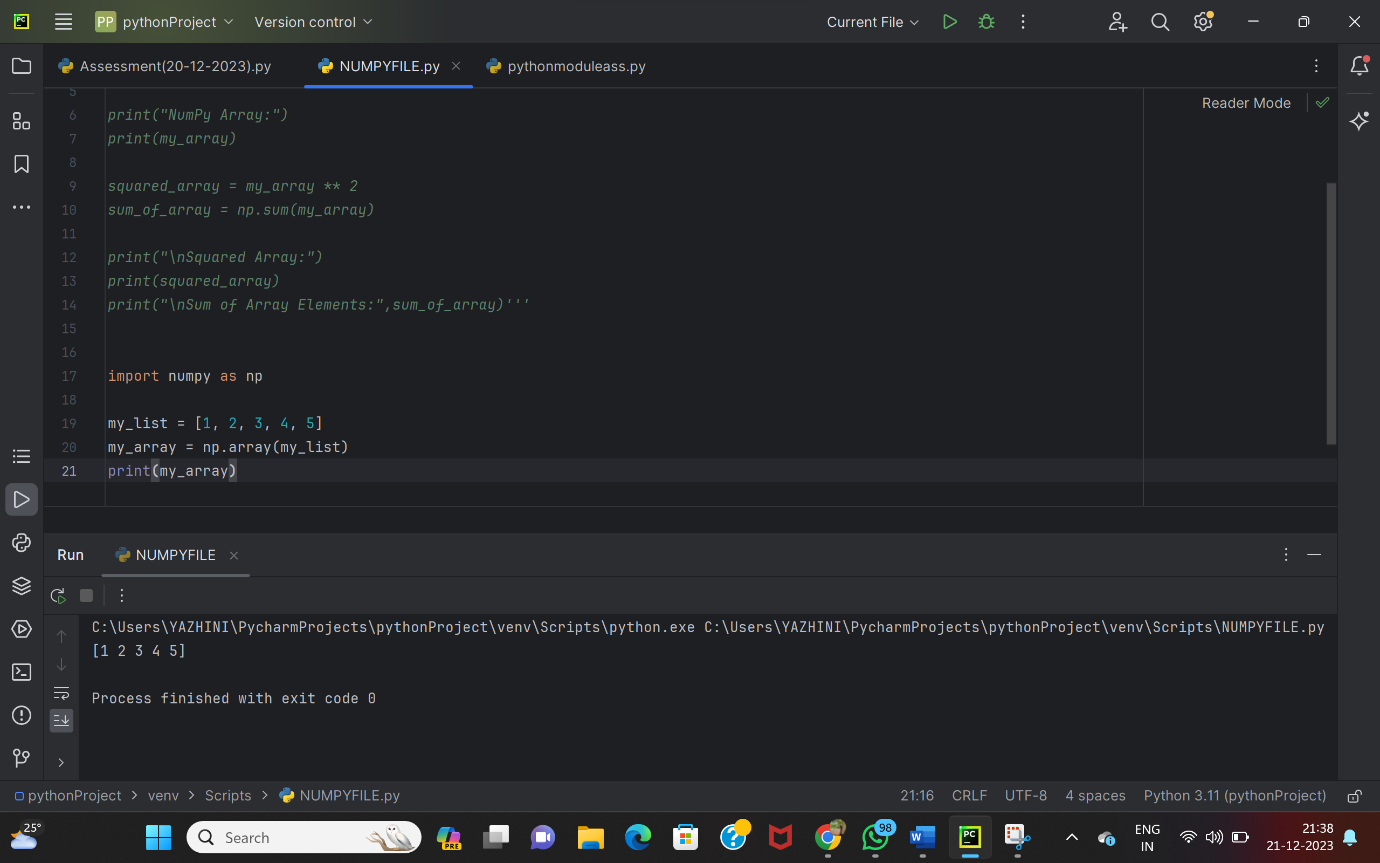
ASSIGNMENT (20-12-2023)

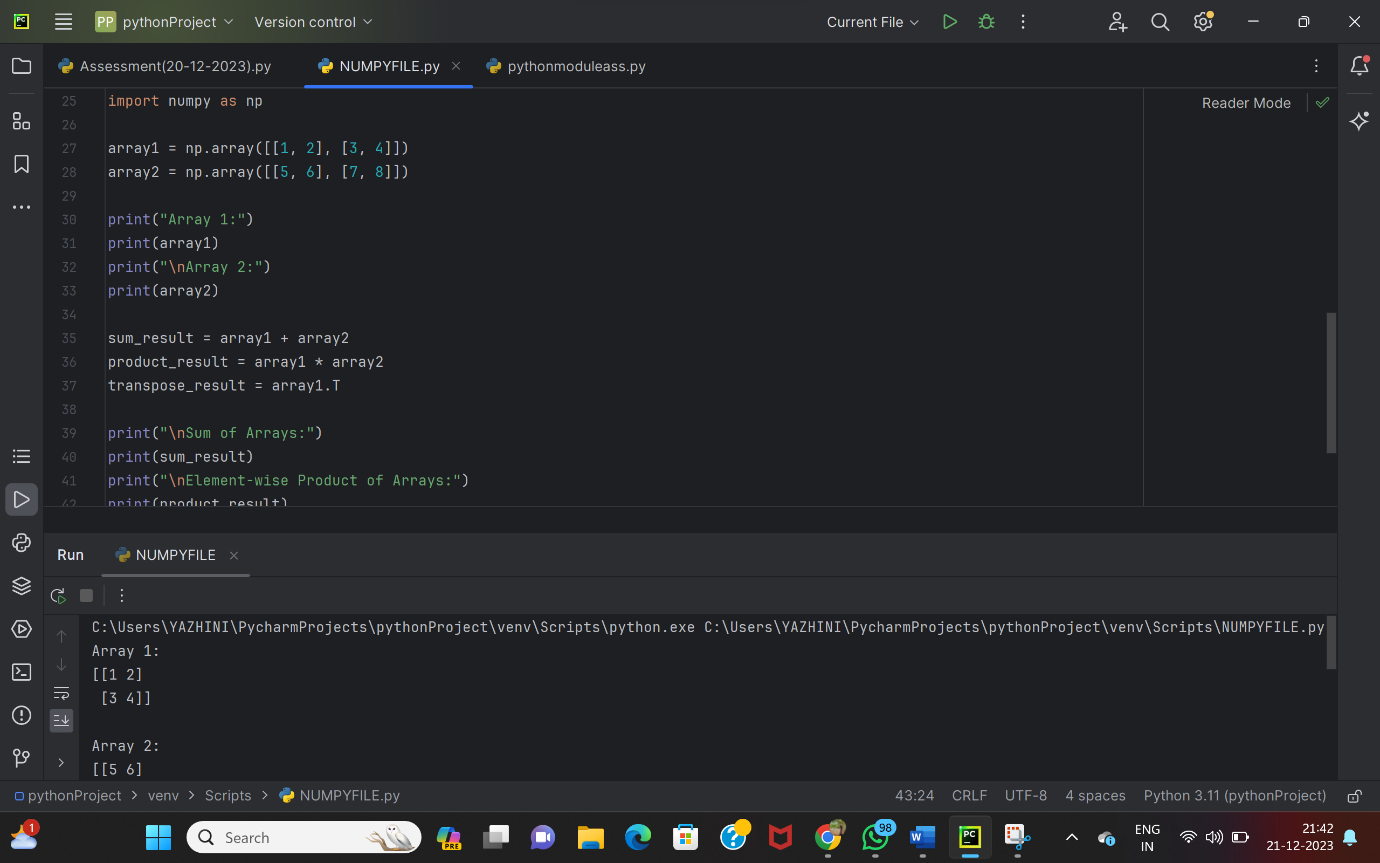
ARRAYS IN NUMPY



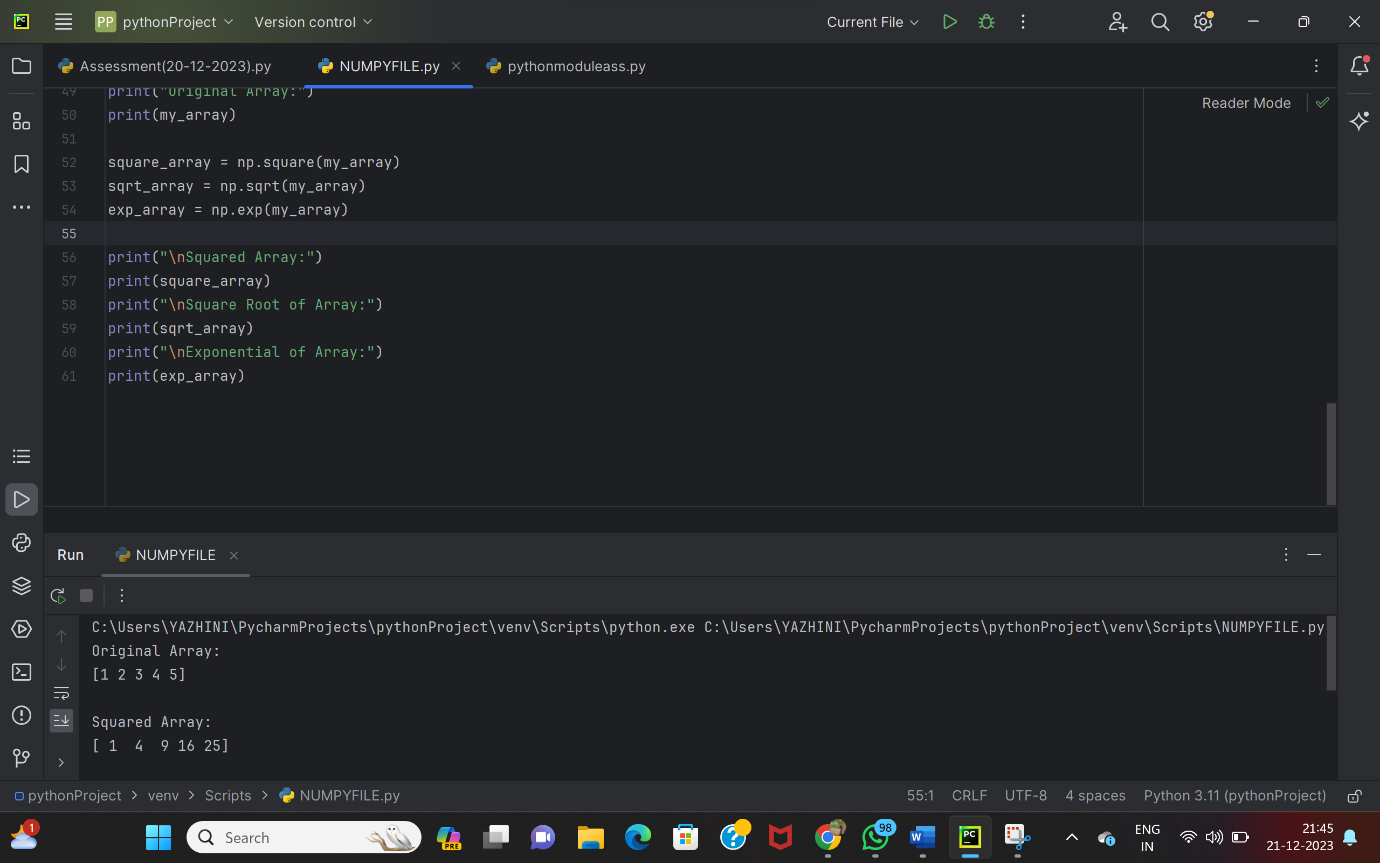
NUMPY ARRAY CREATION



NUMPY BASIC OPERATIONS



UNARY OPERATORS



BIG DATA

🡪Big data refers to the large and complex sets of information that traditional data processing methods struggle to handle efficiently.

🡪It encompasses vast volumes of data, often characterized by the four Vs: volume, velocity , variety and veracity.

**VOLUME :**

* + - Refers to the sheer size of the data generated and collected.
    - Big data involves handling vast amounts of information that exceed the capabilities of traditional databases.
    - The challenge is managing and processing this immense volume efficiently.

**VELOCITY :**

* + - Relates to the speed at which data is generated, processed, and updated.
    - Big data often involves real-time or near-real-time processing.
    - Examples include social media updates, financial transactions, and sensor data.
    - Managing data at such high speeds is a critical aspect of big data systems.

**VARIETY :**

* + - * Encompasses the different types of data—structured, semi-structured, and unstructured.
      * Big data includes diverse forms of information, such as text, images, videos, and more.
      * Traditional databases are not well-suited for handling this variety.
      * Big data technologies must accommodate this diversity in data types.

**VERACITY:**

* Refers to the reliability and trustworthiness of the data.
* With the increasing volume and variety of data, ensuring data quality becomes paramount.
* Veracity is about addressing the uncertainty and potential inaccuracies in the data.
* It involves assessing and maintaining the integrity of information for effective decision-making.

DATA SOURCES:

Structured, semi-structured, and unstructured data are classifications based on the organization and format of information.

**STRUCTURED DATA:**

🡪Data that has a well-defined and organized structure.

**SEMI-STRUCTURED DATA:**

🡪Data that doesn’t fit neatly into a structured format but has some level of organization.

**UNSTRUCTURED DATA:**

**🡪** Data that lacks a predefined data model or structure.