A white robot hand with red and blue stripes

Description automatically generated



**Module-5   
AI & ML**

**LAB MANUAL**

**Disclaimer: The content is curated from online/offline resources and used for educational purpose only**

**Analyzing Annual Rainfall Variability Across Five Indian States Using Variance and Standard Deviation with NumPy and Pandas.**

**Objective:**

* To **calculate and analyze** the **variance** and **standard deviation** of annual rainfall across five Indian states.
* To understand **rainfall variability** and its implications for agricultural planning and water management.

**Duration** 2 Hrs

**Problem Statement:**

Rainfall variability significantly impacts agriculture, water resources, and disaster management in India. Analyzing **annual rainfall data** across five Indian states can help in understanding rainfall distribution and fluctuations. Using **variance and standard deviation**, we will measure how rainfall differs among the states.

**Requirement:**

Python with libraries:

* + NumPy (for statistical calculations)
  + Pandas (for handling data)

Dataset: <https://www.kaggle.com/datasets/vanvalkenberg/historicalweatherdataforindiancities>

**Procedure:**

**Step 1: Install Required Libraries**

**Step 2: Import Necessary Libraries**

* numpy is used for calculating **mean and median**.
* pandas is used to **store and process tabular data**.

A black rectangular object with white text

AI-generated content may be incorrect.

**Step 3: Load the Rainfall Dataset**

**A screen shot of a computer

AI-generated content may be incorrect.**

**Step 4: Replaces errors with Nat (Not a Time) and drop the rows with Nat value.**

A black screen with text

AI-generated content may be incorrect.

pd.to\_datetime(df['time']): This function converts the column 'time' into datetime format, which allows for easier date-based operations.

errors='coerce':

* If any value in the 'time' column is not a valid date format, it will be replaced with NaT (Not a Time).
* This avoids errors that would normally crash the script.

**Step 5: Extract the year from Time column**

A black screen with colorful text

AI-generated content may be incorrect.

**Step 6: Group by Year and sum of total rainfall**

**A black screen with white text

AI-generated content may be incorrect.**

**Step 7: Compute Variance and Standard Deviation**

**A screen shot of a computer program

AI-generated content may be incorrect.**

**Variance\_Explanation:**

* np.var() computes the **variance**, which measures the spread of the data.
* ddof=0 calculates **population variance** (if ddof=1, it calculates **sample variance**).

**Formula for Variance:**

A math equations and numbers

AI-generated content may be incorrect.

**Std\_Explanation:**

* **np.std()** calculates the standard deviation, which represents the average deviation from the mean.
* Formula for Standard Deviation:

A black text with black letters

AI-generated content may be incorrect.

**Step 8: Print the Results**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Conclusion**

By calculating the variance and standard deviation of annual rainfall, we evaluated how much rainfall fluctuates over the years. A high variance and standard deviation indicate significant annual rainfall variability, while lower values suggest consistency in rainfall patterns. Understanding this variability helps in water resource management, flood prediction, and drought preparedness. This analysis is crucial for agriculture, urban planning, and climate studies.

.