21-10-2024

# Training Day – 23

*Daily Diary on Data Analysis Topics*

## WHAT IS DATA ANALYTICS?

sets to extract meaningful insights and support decision-making. This process helps businesses and individuals identify patterns, trends, and actionable conclusions from raw data.

- **Mathematics and Statistics**: Basics of probability, linear algebra, and hypothesis testing.
- **Programming**: Learn languages like Python which are widely used for data analysis.
- **Data Manipulation and Visualization**: Master libraries like:

- *Python: Pandas, NumPy, Matplotlib, Seaborn*
- *R: ggplot2,*

## Work on Real-Life Projects

- Start small, such as analyzing public datasets on Kaggle or Google Dataset Search.
- Gradually tackle more complex datasets, like financial records or social media metrics

## Tools for Data Analytics

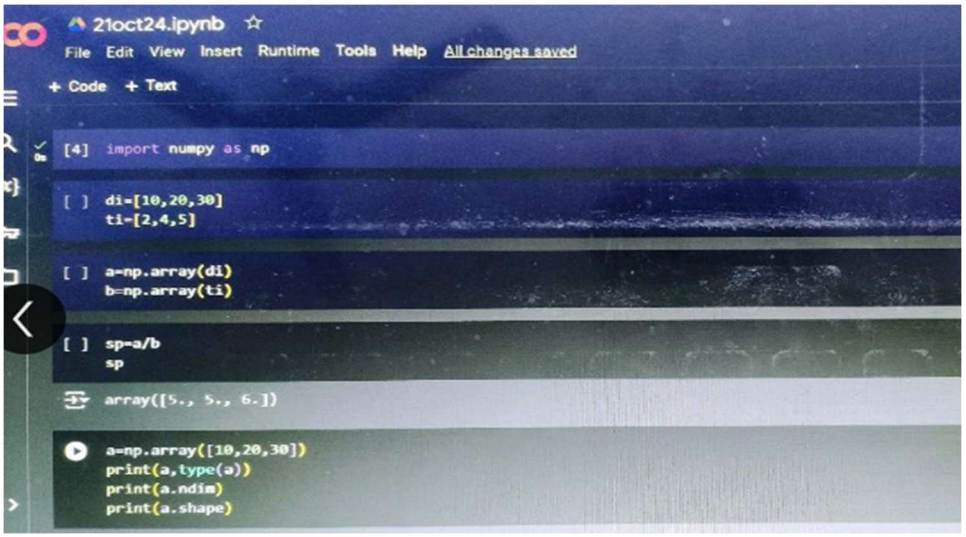
1. **Data Processing and Analysis**
   - **Excel**: Good for small-scale analysis.
   - **Python**: Libraries like Pandas, NumPy, and Scikit-learn.
2. **Data Visualization**
   - **Tableau**: Easy-to-use for creating interactive dashboards.
   - **Power BI**: Microsoft's tool for creating visual reports.
   - **Matplotlib and Seaborn**: Python-based libraries for visualization.

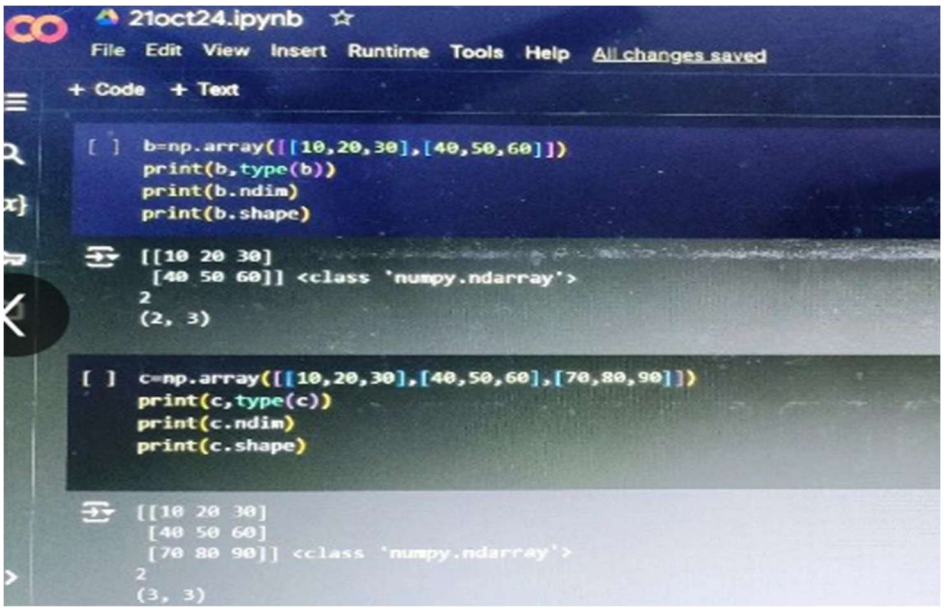
## *Topic:* Introduction to NumPy Variables

- Learned about numpy.ndarray, its creation, and basic properties.
- Example: Created arrays using np.array() and explored their dimensions, shapes, and data types learned Numpy (np) library and some in-built functions of numpy.

  **Functions like ->** *astype, size, ndim, dtype, shape, type()*

  •And also practiced about indexing and slicing of arrays in numpy.


21oct24.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[4] import numpy as np
[ ] di=[10,20,30]
ti=[2,4,5]
[ ] a=np.array(di)
b=np.array(ti)
[ ] sp=a/b
sp
array([5., 5., 6.])
a=np.array([10,20,30])
print(a,type(a))
print(a.ndim)
print(a.shape)


21oct24.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[ ] b=np.array([[10,20,30],[40,50,60]])
print(b,type(b))
print(b.ndim)
print(b.shape)
[[10 20 30]
[40 50 60]] <class 'numpy.ndarray'>
2
(2, 3)
[ ] c=np.array([[10,20,30],[40,50,60],[70,80,90]])
print(c,type(c))
print(c.ndim)
print(c.shape)
[[10 20 30]
[40 50 60]
[70 80 90]] <class 'numpy.ndarray'>
2
(3, 3)