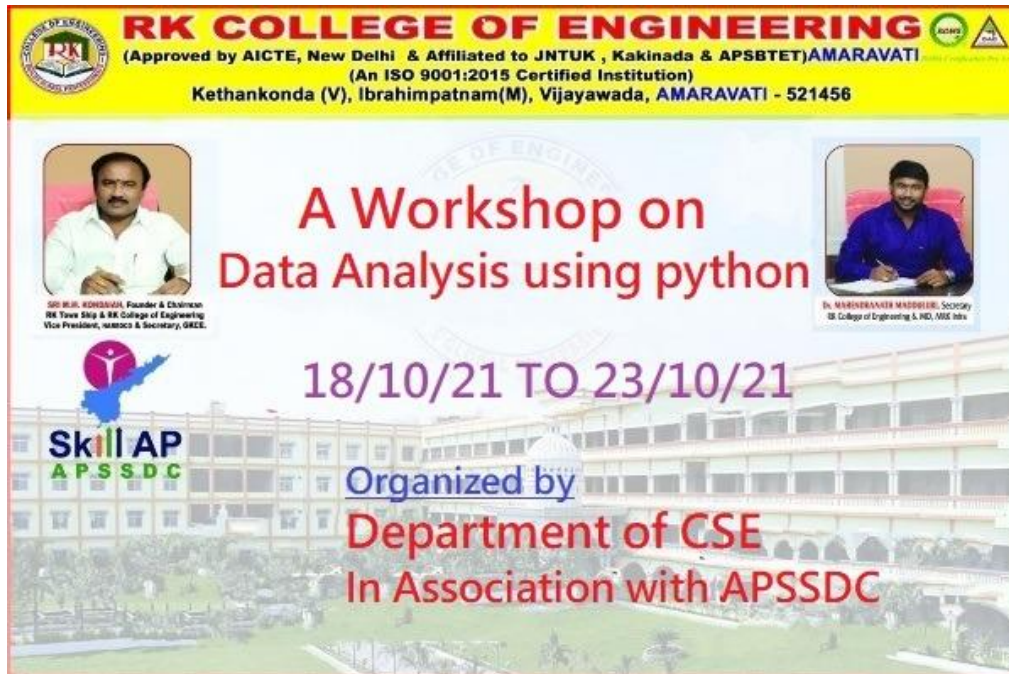


A Workshop on “Data Analysis Using Python”

Dates:18-10-2021 to 23-10-2021



Our Department of Computer Science and Engineering has organized A workshop on “Data Analysis Using Python” from **18-10-2021 to 23-10-2021** and the resource persons are **Lokesh Uppugunduru and Gopi Mandepudi**. This workshop is conducted for III CSE –I Semester students.

No. of Students:96

Faculty Co-ordinators:

Mr.Sk.Shabaaz, Asst.Prof

Miss.D.SudhaRani, Asst.Prof

Workshop Objectives: The main goal of this workshop is to help Students to learn, understand, and practice data analysis and machine learning approaches, which include the

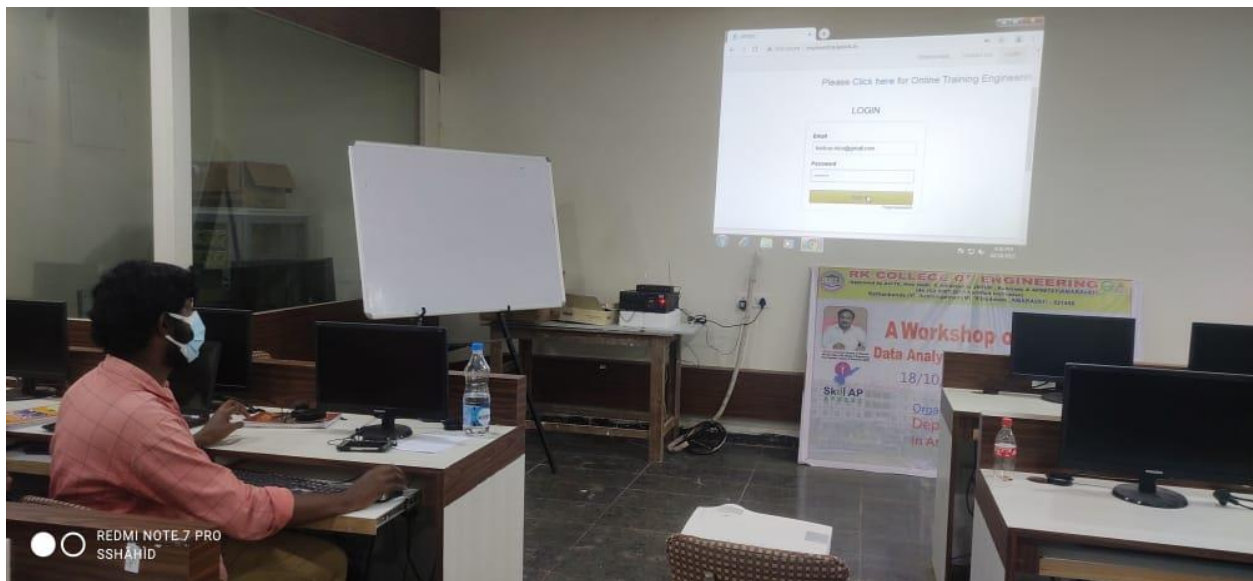
study of modern computing data technologies and scaling up machine learning techniques focusing on industry applications. Mainly the course objectives are conceptualization and summarization of Data Analysis and machine learning computing technologies, machine learning techniques, and scaling up machine learning approaches.

DAY 1: 18th October

Introduction to python and students were taught that python is simple programming language compared to other programming languages, and the following topics were covered.

1. Introduction to Data Analysis (6 Hr):

- a. Introduction to Data
- b. Types of Data in Statistics (Numerical & Categorical) .
- c. Overview of Python Concepts



Day-II: 19th October 2021

The following topics were covered in second day of workshop:

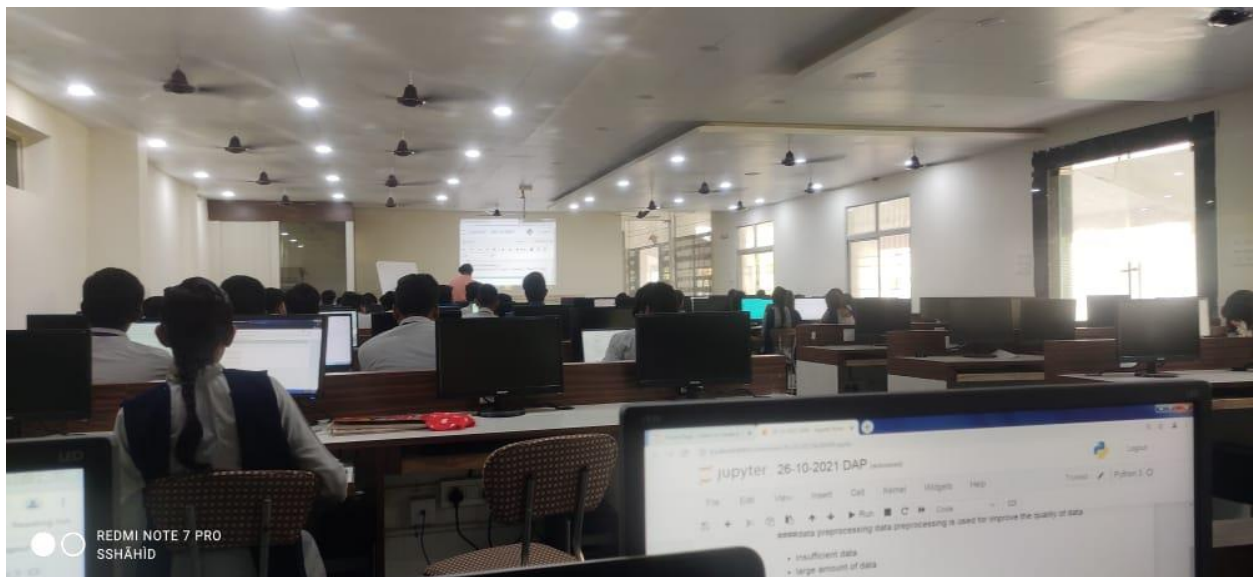
- #### **2. Data Manipulation with NumPy (6 Hr):**
- a. Introduction
 - b. NumPy Arrays
 - c. NumPy Basics
 - d. Math
 - e. Indexing
 - f. Random
 - g. Filtering
 - h. Statistics
 - i. Aggregation
 - j. Saving Data



Day-III: 20th October 2021

The following topics were covered in third day of workshop:

3. Data Analysis with Pandas(6 Hr): a. Introduction to Data Analysis using Pandas b. Pandas Series c. Pandas DataFrame d. Combining e. Indexing f. File I/O g. Grouping h. Features i. Filtering j. Sorting k. Statistical l. Plotting



Day-IV: 21st October 2021

The following topics were covered in fourth day of workshop:

4. Data Cleaning with Pandas & Data Preprocessing with Scikit Learn (6 Hr): a. Introduction b. Standardizing Data c. Data Range d. Robust Scaling e. Normalizing Data f. Label and One Hot Encoding g. Polynomial Features h. Working with Duplicates and Missing Values i. Which values should be replace with missing values based on data j. Identifying and Eliminating Outliers k. Dropping duplicate data l. Filling missing data using Data Imputation



Day-V: 22nd October 2021

The following topics were covered in third day of workshop:

5. Introduction to Data Visualization with Matplotlib (6 Hr): a. Introduction to Visualization and Python packages b. Matplotlib history c. Introduction to plotting d. Line Plot e. Scatter Plot f. Bar Graph g. Histogram h. Pie Chart i. Box Plot j. Tasks



Day-III: 23rd October 2021

The following topics were covered in last day of workshop:

6. Data Visualization With Seaborn (6 Hr): a. Using Seaborn Styles b. Setting the default style c. Color Palettes d. Creating Custom Palettes e. stripplot() and swarmplot() f. box plots, violin plots, barplots, point plots and counterplots g. Regression Plots h. Binning data i. Creating heatmaps j. Applying on raw dataset and introduction to Kaggle and other data sources