# DIGITAL NOTES
## ON
# DATA VISUALIZATION


# B.TECH IV YEAR
## (2022-23)



**Edited by Jogi John**


# DEPARTMENT OF COMPUTER TECHNOLOGY

**LOKMANYA TILAK JANKALYAN SHIKSHAN SANSTHA'S**
**PRIYADARSHINI COLLEGE OF ENGINEERING**
**(Approved by A.I.C.T.E., New Delhi & Govt. Of Maharashtra,**
**Affiliated to R.T.M. Nagpur University, Nagpur)**
**Priyadarshini Campus, Near CRPF, Hingna Road,**
**Nagpur – 440 019, Maharashtra (India)**

### Elective IV : Data Visualization (TH)

**Course Objectives:**

To learn different statistical methods for Datavisualization.

To understand the basics of R and Python.

To learn usage of Watson studio.

To understand the usage of the packages like Numpy, pandas and matplotlib.

To know the functionalities and usages ofSeaborn.

**UNIT I**
Introduction to Statistics : Introduction to Statistics, Difference between inferential statistics and Descriptive statistics, Inferential Statistics- Drawing Inferences from Data, Random Variables, Normal Probability Distribution, Sampling, Sample Statistics and Sampling Distributions.
R overview and Installation- Overview and About R, R and R studio Installation, Descriptive Data analysis using R, Description of basic functions used to describe data in R.

**UNIT II**
Data manipulation with R: Data manipulation packages-dplyr,data.table, reshape2, tidyr, Lubridate, Data visualization with R.
Data visualization in Watson Studio: Adding data to datarefinery, Visualization of Data on Watson Studio.

**UNIT III**
Python: Introduction to Python, How to Install, Introduction to Jupyter Notebook, Python Scripting basics, Numpy and Pandas-Creating and Accessing Numpy Arrays, Introduction to pandas, read and write csv, Descriptive statistics using pandas, Working with text data and datetime columns, Indexing and selecting data, groupby, Merge / Joindatasets

**UNIT IV**
Data Visualization Tools in Python- Introduction to Matplotlib, Basic plots using matplotlib, Specialized Visualization Tools usingMatplotlib, Advanced Visualization Tools using Matplotlib Waffle Charts, Word Clouds.

**UNIT V**
Introduction to Seaborn: Seaborn functionalities and usage, Spatial Visualizations and Analysis in Python with Folium, Case Study.

**TEXT BOOKS:**
1. Core Python Programming - Second Edition,R. Nageswara Rao, DreamtechPress.
2. Hands on programming with R by Garrett Grolemund,Shroff/O'Reilly; Firstedition
3. Fundamentals of Mathematical Statistics by S.C. Gupta, Sultan Chand &Son9

**REFERENCE BOOKS:**
1. Learn R for Applied Statistics: With Data Visualizations, Regressions, and Statistics by Eric Goh Ming Hui,Apress
2. Python for Data Analysis by William McKinney, Second Edition, O'Reilly Media Inc.\
3. The Comprehensive R Archive Network-https://cran.r-project.org