

## About the College

Ram Lal Anand College is a University Maintained Institution founded in the year 1964 by Late Shri Ram Lal Anand, a senior advocate in the Supreme Court of India. The college is located in the picturesque surroundings against the backdrop of the Aravali ranges in the neighbourhood of the South Campus of the University of Delhi and several other educational institutions. The college has excellent infrastructure, with state of the art Laboratories, Seminar room, Amphitheatre, Library, Playground and Cafeteria. The campus is Wi-Fi enabled. Being a multi-disciplinary, co-educational institution it has approximately 1500 students pursuing 14 undergraduate courses in Arts, Commerce and Science streams. The college boasts of a highly learned and committed teaching faculty of about 80 teachers. Apart from their traditional role of disseminating knowledge, the teachers inspire and guide the students to manage different activities such as seminars, workshops, debates, theatre, cultural activities including classical music and dance programmes.

## About StepUp Analytics

StepUp Analytics is a community of creative, high energy data science and analytics professionals and data enthusiasts. It aims at bringing together influencers and learners from industry to augment knowledge and to provide as many resources as possible for analytics and data science learning.

## Instructional Methods

1. Instructor Led Training
2. Audience Participation/Hands-on Exercises
3. Assignments or Problem Solving Strategies
4. Opportunity to learn one of the popular and applicable programming language.
5. Small Groups & Individual Practice

## Target Audience

UG students, PG students and Research Scholars

Registration Fee: Rs 4000/-

## Contact Details

Programme Coordinator: Dr. Seema Gupta  
Associate Professor, Department of Statistics

**Contact no. 9891664133**  
**seemag\_66@yahoo.com**  
**For online registration [Click Here](#)**

## Important Notes

1. The fee will include the Participation fee, Certificate and Refreshment.
2. Participants are requested to submit registration form along with the fee in the form of NEFT/DD drawn in favour of "Principal, Ram Lal Anand College" payable at New Delhi on or before May 7 2018, UBI, Bank SSF A/c No. 403502010005997 IFSC UBIN0540358.  
**NO SPOT REGISTRATION**

# Certificate Course in Data Science and Programming with Python

May 26th - June 1st 2018



Organised by



**Ram Lal Anand College**  
University of Delhi, South Campus  
Benito Juarez Road, New Delhi 110021  
rlac.du@gmail.com  
with



Collaborative learning | Skill Up Scale Up  
C-510, Siddha Pines Opp Derozio college Kolkata 700136.

## Objectives of Course

1. To help the participants to familiarize with Python Programming language and data analysis.
2. To train the participants in accessing data from various domains.
3. To develop confidence as independent Python Programmer. and data analyst

## Why Python

Easy to use, open source and versatile programming language specially popular among those learning programming for the first time. It is portable and cross platform

It is concise and easy to read, and can be used for a variety of industry needs including web development, data analytics, core software development and a wide range of scientific and mathematical applications.

Python also has numerous libraries for data manipulation and analysis as well as a very active development community that continually updates and creates new packages. It has been adopted by a wide variety of industries and applications including data science, machine learning, data analytics, predictive analytics, business intelligence and web analytics.

SAGE, NumPy, SciPy, Matplotlib and Cython make Python an open source alternative to MATLAB, Mathematica, Magma and Maple.

# Curriculum

9.00 am - 5.00 pm

## Introduction to Data Science

- What is Data Science?
- Data Science Vs. Analytics vs. Data warehousing, OLAP, MIS Reporting
- Relevance in industry and need of the hour
- Type of problems and objectives in various industries
- How leading companies are harnessing the power of Data Science?
- Different phases of a typical Analytics/Data Science projects

## Python – Introduction & Essentials

- Introduction to Python Editors & IDE's (Canopy, pycharm, Jupyter, Ipython etc...)
- Custom Environment Settings
- Python Basic Rules in Python 3
- Concept of Packages/Libraries - Important packages (NumPy, SciPy, scikit-learn, Pandas, Matplotlib, etc)
- Data Types & Data objects/structures List and Dictionary Comprehensions
- Variable & Value Labels– Date & Time Values
- Basic Operations - Mathematical - string - date
- Reading and writing data
- Simple plotting/Control flow/Debugging/Code profiling

## Python: Data Manipulation – cleansing

- Cleansing Data with Python
- Data Manipulation steps(Sorting, filtering, duplicates, merging, appending, subsetting, derived variables, sampling, Data type conversions, renaming, formatting etc)
- Data manipulation tools (Operators, Functions, Packages, control structures, Loops, arrays etc)
- Python Built-in Functions
- User Defined Functions in Python
- Stripping out extraneous information
- Normalising data and Formatting data
- Important Python Packages for data manipulation (Pandas, Numpy etc)

## Python - Data Analysis – Visualization

- Introduction exploratory data analysis
- Descriptive statistics, Frequency Tables
- Univariate Analysis (Distribution of data & Graphical Analysis)
- Bivariate Analysis(Cross Tabs, Distributions & Relationships, Graphical Analysis)
- Creating Graphs-
- Important Packages for Exploratory Analysis (NumPy Arrays, Matplotlib, Pandas and scipy.stats etc)

## Python: Basic Statistics

- Basic Statistics
- Building blocks - Probability Distributions
- Inferential Statistics -Sampling - Concept of Hypothesis Testing
- Statistical Methods - Z/t-tests, ANOVA, Correlation and Chi-square

## Python: Introduction to Machine learning

- Statistical learning vs. Machine learning
- Iteration and evaluation
- Major Classes of Learning Algorithms -Supervised vs Unsupervised
- Different Phases of Predictive Modelling Concept of Overfitting and Under fitting (Bias-Variance Trade off) & Performance Metrics
- Types of Cross validation (Train & Test, Bootstrapping, K-Fold validation etc)

## Python: Predictive Modelling – Basics Introduction to Predictive Modelling

- Types of Business problems - Mapping of Techniques
- Linear Regression
- Logistic Regression
- Segmentation - Cluster Analysis
- Decision Trees (CHAID/CART/CD 5.0)
- Time Series Forecasting

- **Case study on frequently purchased items for a large retailer**
- **Case study to access a twitter account and create several visualizations to draw interesting insights**
- **Case study on predicting sales for a large retailer using real data**