

CETPA INFOTECH PVT. LTD.
CURRICULUM OF DATA SCIENCE USING SAS

Getting Familiar with DATA Science

- Importance of Analytics industry & need of the hour
- Types of analytics – Fraud, Marketing, Risk, Operations, etc
- Business & Technology drivers for analytics
- Analytics Tool Kit - Popular tools & Techniques in the Industry
- Future of analytics & How to fulfill requirement.
- Types of problems and business objectives in various industries
- Different phases of Analytics Project

Analytics using Advanced Excel

- Introduction to Excel
- Working with Formulas and functions
- Conditional Formatting
- Advance Filtering, Sorting, Paste special 15 different functions.
- Functions (Logical & Text, Mathematical, Statistical)
- Data Manipulation & Data Consolidation
- Data Analysis using Analysis Tool Pack (Histogram, Correlation, etc)
- Analyzing Data using Pivot Tables
- Use in Quality Industry using Pareto Analysis
- Creating Charts & Graphs
- Data analytical tool (What -if analysis, Goal seek, Data Table, Solver)
- Protecting Workbooks, worksheets and formulas

Analytics using SAS

Getting into SAS

- Introduction to SAS, GUI
- Creation and Analysis of SAS Dataset (Data & Proc Steps - Statements & options)
- Debugging SAS Codes
- Importing different types of data & connecting to data bases

- Data Understanding(Variable attributes(format, informat, length, label etc))
- Concepts of Libraries, PDV, data execution, etc
- SAS Procedures for data import /export / understanding(Proc import/Proc Export)

Data Manipulation using SAS

- Data Manipulation steps(Sorting, filtering, duplicates, merging, appending, subsetting, derived variables, sampling, Data type conversions renaming, formatting, etc)
- Data manipulation tools (Operators, Functions, Procedures, conditional , Loops, arrays, etc)
- SAS Functions (Text, numeric, date, utility functions)
- SAS Procedures for data manipulation (Proc sort, proc format, etc)
- SAS Options (System Level, procedure level)
- Debugging SAS Codes
- Importing different types of data & connecting to data bases
- Data Understanding(Variable attributes(format, informat, length, label etc))
- Concepts of Libraries, PDV, data execution, etc
- SAS Procedures for data import /export / understanding(Proc import/Proc Export)

ANNOVA

- Verify the assumptions of ANOVA
- Analyze differences between population means using the GLM and TTEST procedures

Linear Regression Model

- Linear Regression - Introduction - Applications
- Assumptions of Linear Regression
- Building Linear Regression Model
- Understanding standard metrics (Variable significance, R-square/Adjusted R-Square)
- Validation of Linear Regression Models (Re running Vs. Scoring)
- Standard Business Outputs (Decile Analysis, Error distribution (histogram), Model equation, drivers etc)

Advanced SAS

- Introduction to Advanced SAS - Proc SQL & Macros
- Understanding select statement (From, where, group by, having, order by etc)
- Proc SQL - Data creation/extraction
- Proc SQL - Data Manipulation steps
- Proc SQL - Summarizing Data
- Proc SQL - Concept of sub queries, indexes etc
- SAS Macros - Creating/defining macro variables
- SAS Macros - Defining/calling macros
- SAS Macros - Concept of local/global variables
- SAS Macros - Debugging techniques

Descriptive Analysis via SAS

- Types of Data (Metric and Non Metric)
- Descriptive statistics, Inferential Statistics
- Introduction to data analysis
- Frequency Tables and summarization
- Univariate Analysis (Distribution of data & Graphical Analysis)
- Bivariate Analysis (Cross Tabs, Distributions & Relationships, Graphical Analysis)
- SAS Procedures for Data Analysis (proc freq/Proc means/proc summary/proc tabulate/Proc univariate etc)
- SAS Procedures for Graphical Analysis (Proc Sgplot, proc gplot etc)

T-Test Interpretation

- Concept of Hypothesis Testing
- P-Value Approach in Accepting and Rejecting Hypothesis
- DV and IV

- Interpretation of Results - Business Validation - Implementation on new data

Chi Square Test

- Working on NOMINAL data
- Cramer's V importance

Logistic Regression Model

- Logistic Regression - Introduction - Applications
- Linear Regression Vs. Logistic Regression Vs. Generalized Linear Models
- Building Logistic Regression Model
- Understanding standard model metrics (Concordance, Variable significance, Hosmer Lemeshov Test, Gini, KS, Misclassification etc)
- Validation of Logistic Regression Models
- Interpretation of Results - Business Validation - Implementation on new data

HEAD OFFICE: 200 Purwawali , 2nd Floor, (Opp. Railway Ticket Agency), Railway Road , Ganeshpur,
Roorkee – 247667, Ph.No.: 09219602769, 01332-270218 Fax - 1332 – 274960

CORPORATE OFFICE: D-58, Sector-2, Near Red FM. Noida -201301, Uttar Pradesh
Contact Us: +91-9212172602 , 0120-4535353

BRANCH OFFICE: 401 A, 4th Floor, Lekhraj Khazana, Faizabad Road, Indira Nagar,
Lucknow-220616 (U.P.) Ph. No: +91-522-4233162, +91-9258017974

BRANCH OFFICE: 105, Mohit Vihar, Near Kamla Palace, GMS Road, Dehradun-248001, UK
Contact: +91-9219602771, 0135-6006070

Toll Free- 1800-8333-999 (from any network)

CETPA[®]

Because Knowledge Matters

ISO 9001:2015 Certified