

# **SYLLABUS FOR Ph.D. ENTRANCE TEST**

## **SUBJECT: DATA SCIENCE**

### **PART-A**

Part-A shall consist of 50 objective type compulsory questions of 1 mark each based on research methodology. It shall be of generic nature, intended to assess the research aptitude of the candidate. It will primarily be designed to test reasoning ability, data interpretation, and quantitative aptitude of the candidate.

### **PART-B**

Part-B shall also consist of 50 objective type compulsory questions of 1 mark each based on the syllabus of the subject at Masters Level as follows:

#### **1- FORECASTING METHODS:**

An overview of forecasting methods. Basic steps in forecasting. Basic forecasting tools: time series and cross-sectional data, graphical and numerical summaries, forecasting accuracy, prediction intervals, transformations and adjustments. Time series decomposition: principles of decomposition, moving averages, classical decomposition, census bureau methods, forecasting and decomposition. Smoothing methods: averaging methods, exponential smoothing methods, comparison of methods, general aspects of smoothing methods. Qualitative forecasting methods.

#### **2- STATISTICAL RESEARCH METHODS:**

Measures of variation, Skewness, moments and kurtosis, Index numbers. Interpolation and Extrapolation. Sampling and test of significance: Procedure of testing a hypothesis. Sampling of attributes, sampling of variables (large and small samples). Association of Attributes, Chi-square test and goodness of fit. Analysis of variance, Design of experiments. Introduction Two-variable linear Regression model: Basics, Problem of Estimation, Classical linear Regression Model.

#### **3: DECISION ANALYSIS:**

Types of decision; structuring the decision problem: payoff tables, decision trees; decision making without probabilities: optimistic approach, conservative approach, minimax regret approach; decision making with probabilities-expected value of perfect information. Decision analysis with sample information: expected value of sample information, efficiency of sample information; computing branch probabilities; utility and decision making. Game Theory, Two Person Zero-sum Games, Graphical Solution of  $(2 \times n)$  and  $(m \times 2)$  games.

#### **4- MATHEMATICAL MODELLING AND SIMULATION:**

Operations Research: History, Phases, Applications and Limitations. Introduction to Linear Programming, Formulation of LP models, Solution of LP models-Graphical Method, Simplex Method, M-Method, Two-phase method, Duality & Dual-simplex method. Transportation model, Assignment model, Goal programming, Integer programming. Networks Models: CPM, PERT. Simulation.

#### **5: DATA ANALYSIS:**

Machine learning and data mining algorithms. Review of probability and probability distributions, Bayes Rule. Big data; Prediction vs. description; exploratory data analysis; data visualization, data journalism, dashboards. Classification, ranking, logistic regression. Ethics, time series, advanced regression.