RAJALAKSHMI ENGINEERING COLLEGE
An AUTONOMOUS Institution
Affiliated to ANNA UNIVERSITY, Chennai

Minor Degree in Business Analytics

Sl. No.	COURSE CODE	COURSE TITLE	CONTACT PERIODS	L	T	P	C
1.	MCB1901	Fundamentals of Business Statistics	3	3	0	0	3
2.	MCB1941	Data Science for Business Analytics	5	3	0	2	4
3.	MCB1942	Programming for Data Analytics	6	2	0	4	4
4.	MCB1902	Digital Marketing and Web Analytics	3	3	0	0	3
5.	MCB1903	Operations and Supply Chain Analytics	3	3	0	0	3
6.	MCB1911	Mini Project	2	0	0	2	1
TOTAL			**22**	**14**	**0**	**8**	**18**

Subject Code	Subject Name (Theory Course)	L	T	P	C
MCB1901	**FUNDAMENTALS OF BUSINESS STATISTICS**	3	0	0	3

Objectives:	
•	To gain the basic knowledge on statistics.
•	To understand the relation among the data and selecting the highly influential data for the process.
•	To learn the role, need and selection approaches of the factors.
•	To explore the various types of possible errors and error reduction approaches.
•	To analyze the time series data with statistical models.

UNIT-I	**DATA PROJECTION, LINEAR REGRESSION AND CORRELATION**	**9**
Graphical representation - histogram, boxplots, and scatterplots, Numerical data: Mean, Median, Quantiles, Variance and standard deviation, Normal Distribution. Categorical data: Two-way tables, bar graphs, segmented bar graphs. Linear regression: Least Squares, residuals, outliers and influential observations, extrapolation. Inference in Linear Regression: Confidence intervals for intercept and slope, significance test, mean response and predication intervals. Correlation: Correlation coefficient, Rank correlation. Multiple Linear Regression: Confidence intervals, test of significance, squared multiple correlation, ANOVA: Analysis of variance for simple and multiple regression, F statistics.		
UNIT-II	**PRINCIPAL COMPONENT ANALYSIS**	**9**
Principal components, Algorithm for conducting principal component analysis: Data standardization, covariance matrix evaluation, Eigen vector and Eigen value calculation, deciding on how many principal components to retain, H-plot.		
UNIT-III	**FACTOR ANALYSIS**	**9**
Factor Analysis: Definition, exploratory factor analysis and confirmatory factor analysis. Factor analysis model, Extracting common factors, determining number of factors, Transformation of factor analysis solutions, Factor scores.		
UNIT-IV	**ERROR ANALYSIS**	**9**
Introduction, Accuracy vs Precision, Types of Errors / Uncertainties: Random errors, Systematic errors. Type I error (false positive), Type II error (false negative), Bias, Regression Error, Standard Error, Sampling and Non-Sampling error.		
UNIT-V	**ANALYSIS & FORECASTING ON TIME SERIES DATA**	**9**
Stationary, ARIMA Models: Identification, Estimation and Forecasting.		
Contact Hours :		**45**

Course Outcomes: On completion of the course, students will be able to	
•	Gain the basics of statistics knowledge
•	Expertise in analyzing and selecting the highly influential features.
•	Understand the factors' role for further analysis.
•	Proficiency in finding and optimizing the errors to improve the model performance.
•	Able to construct an optimal model on time series data for the required analysis.

Text Books:	
1.	Bruce, Peter, Andrew Bruce, and Peter Gedeck,"Practical statistics for data scientists: 50+ essential concepts using R and Python", O'Reilly Media, 2020.
2.	Johnson, Richard A., Irwin Miller, and John E. Freund, "Probability and statistics for engineers", Vol. 2000. London: Pearson Education, 2000.
3.	Cowpertwait, Paul SP, and Andrew V. Metcalfe,"Introductory time series with R", Springer Science & Business Media, 2009.

Reference Books / Web links:	
1.	Montgomery, Douglas C., Elizabeth A. Peck, and G. Geoffrey Vining, “Introduction to linear regression analysis”, John Wiley & Sons, 2021.
2.	Mood, A. M., Franklin A. Graybill, and D. C. Boes. "Introduction to the theory of statistics”, McGraw-Hill series in probability and statistics, 1974.
3.	Draper, Norman R., and Harry Smith, “Applied regression analysis”,Vol. 326. John Wiley & Sons, 1998.
4.	Goon, A. M., M. K. Gupta, and B. Dasgupta, "Fundamentals of Statistics, Vol. I & II.” 2002.
5.	Chatfield, Chris, “The analysis of time series: an introduction”, Chapman and Hall, CRC, 2003.
6.	http://www.stat.yale.edu/Courses/1997-98/101/stat101.htm
7.	https://journals.physiology.org/doi/pdf/10.1152/ajpendo.00484.2003

Subject Code	Subject Name (Lab Oriented Theory Course)	L	T	P	C
MCB1941	Data Science for Business Analytics	3	0	2	4

Objectives:	
•	To understand the basic concept, process, and practice of data analytics and how it helps to develop a solution for decision-making.
•	To understand various methods and models for the evaluation of business problems.
•	To learn various model fitting and machine learning of overfitting and underfitting techniques for accurateprediction.
•	To design an analytical solution based on the available data, tools, and frameworks.
•	To learn data visualization and various types of visual charts and techniques used in real-time problem-solving with case studies.

UNIT-I	**Introduction to Business Analytics**	**9**
Introduction to Data Science - Business Analytics-The science of data-driven decision making-Descriptive, Predictive and Prescriptive Analytics Techniques-Big Data Analytics-Framework for Data-Driven decision making-Analytics capability building-Roadmap for Analytics Capability Building-Challenges in Data-Driven Decision Making.		
UNIT-II	**Business Problems and Model Building**	**9**
From business problems to data mining tasks – Supervised vs Unsupervised methods-The data mining process-Business understanding – Data understanding- Data preparation - Preprocessing - Modeling-Evaluation-Deployment-Implications for managing the data science team-Other analytics techniques and technologies-Predictive model building- Models, Induction and Prediction-Supervised segmentation-Clustering as similarity based segmentation-Nearest Neighbor Reasoning-Hierarchical Clustering-Visualizing segmentation-Trees as a set of rules-Probability Estimation.		
UNIT-III	**Model fitting and Overfitting**	**9**
Finding optimal model parameters based on data-Classification via mathematical functions-Regression via mathematical functions-Class probability estimation and Logistic Regression-Nonlinear Functions, Support Vector Machines and Neural Networks-Overfitting Examined-Overfitting avoidance and complexity control.		
UNIT-IV	**Decision Analytic Thinking**	**9**
Key evaluation framework – Evaluating classifiers- Key analytical framework – Evaluation, Baseline performance and implications for investments in Data – Visualizing model performance – Techniques – Profit curves – Cumulative Response Curves – Lift curves – ROC curves - Co-occurrences and Associations- Measuring Surprise-Profiling: Finding Typical Behavior- Link Prediction and Social Recommendation- Data Reduction, Latent Information, and Movie Recommendation- Bias, Variance, and Ensemble Methods.		
UNIT-V	**Data Visualization Tools and Techniques**	**9**
Access, merge, and transform all of your data-Make sense of your data with the tools - Support enterprise-wide data science practices-Leverage insights gained from your data - Visualization Basics- Data Visualization with Analytics– Histogram-Bar Chart-Pie Chart-Scatter Plot-Coxcomb Chart-Box Plot (or Box and Whisker Plot) - Treemap - Business Case Studies: Sales Forecasting - Customer Segmentation- Fraud Detection- Inventory Management- Stock Market Analysis.		
`	**Contact Hours : 45**	

List of Experiments	
1	Build a predictive model for a data set and predict the result. Read the training data from a .CSVfile.
2	For a given set of training data examples stored in a .CSV file, implement and demonstrate the clustering algorithm to output similarities.
3	Write a program to demonstrate the working of the decision tree technique. Use an appropriatedata set for building the decision tree and apply this knowledge to classify a new sample.
4	Write a program and implement a regression technique for predicting the sales revenue of acompany based on data such as the previous sales of the company.
5	Write a program to build an effective email spam classification
6	Implement a program for analyzing goodness of Fit (Overfit or Underfit) for a given data set.
7	Write a program to compare Cumulative Response and Lift Curves of Classifiers built on givendata set.

8	Implement and apply a variety of link prediction methods to each of the ego networks containedwithin the SNAP Facebook dataset and SNAP Twitter dataset.			
9	For a given data set access, blend, analyze, and visualize data using KNIME analytics platform open source software.			
10	Implement the data visualization for a given data set a. Find the data distributions using box and scatter plot. b. Find the outliers using plot. c. Plot the histogram, bar chart and pie chart on sample data.			
		Contact Hours	**:**	**30**
		Total Hours	**:**	**75**

Course Outcomes:
On successful completion of the course, students will be able to

- Understand the basic concept, process, and practice of data Analytics and how it helps to develop a solutionfor decision-making.
- Understand various methods and models for the evaluation of business problems.
- Learn various model fitting and machine learning of overfitting and underfitting techniques for accurate prediction.
- Design an analytical solution based on the available data, tools, and frameworks.
- Learn various tools and techniques used in real-time problem-solving with case studies.

Text Books:

1. U Dinesh Kumar, "Business Analytics The Science of Data-driven Decision Making", First Edition, Wiley Publishers, 2017.
2. Foster Provost and Tom Fawcett, "Data Science for Business" First Edition 2013, O'Reilly Media,Inc.

Reference Books / Web links:

3. R.N.Prasad, Seema Acharya, "Fundamentals of Business Analytics", Second Edition, Wiley Publishers, 2016.
4. Regi Mathew, "Business Analytics for Decision Making", First Edition, Pearson India, 2020.
5. Jeffrey D Camm, James J.Cochran, Michael J.Fry, Jeffrey W.Ohlmann, David R.Anderson, Dennis J.Sweeney, "Essentials of Business Analytics", First Edition, Cengage Learning, 2015.
6. https://www.kaggle.com/datasets
7. https://archive.ics.uci.edu/ml/index.php
8. https://data.gov/
9. https://aws.amazon.com/marketplace/solutions/data-analytics/data-sets#

Subject Code	Subject Name (Lab Oriented Theory Course)	L	T	P	C
MCB1942	Programming for Data Analytics	2	0	4	4

Objectives:
• To learn the basics of python programming and File Formats.
• To Understand about the Arrays and Objects Using Python Packages.
• To Understand the basic concepts of Data Processing and Visualization.
• To learn the basics of R programming structures.
• To learn the exploratory data analysis using R.

UNIT-I	**Introduction to Python Programming**	**6**
Python Concepts, Data Structures, Classes: Interpreter, Program Execution, Statements, Expressions, Flow Controls, Functions, Numeric Types, Sequences and Class Definition, Files and Operating Systems, Text & Binary Files - Reading and Writing-Data Loading and Storing File Formats.		
UNIT-II	**Arrays and Objects in Python**	**6**
Understanding Data Types in Python-The Basics of NumPy Arrays-Computation of NumPy Arrays, - vectorized Computation-Introducing Pandas Data Structures: Essential functionality, Summarizing and Computing descriptive Statistics, Data Indexing and Selection, Operating on Data in pandas		
UNIT-III	**Data Processing and Visualization**	**6**
Data Wrangling: Combining and Merging Datasets, Reshaping and Pivoting, Data Transformation, String Manipulation, Regular Expressions-Data Aggregation, Group Operations, Time series: GoupBy Mechanics, Data Aggregation, Groupwise Operations and Transformations, Pivot Tables and Cross Tabulations, Time Series Basics, Data Ranges, Frequencies and Shifting- Visualization in Python: Matplotlib package, Plotting Graphs, Controlling Graph, Adding Text, More Graph Types, Getting and setting values, Patches.		
UNIT-IV	**Introduction to R programming**	**6**
Overview – Data Types, Variables, Operators, Decision Making, Loops, Functions, Strings, Vector, List, Matrices, Arrays, Factors, Data Frames, Packages- Data Interfaces – Reading and writing of CSV files,		
UNIT-V	**R programming for Data Analytics**	**6**
Data Wrangling - The Jupyter and PyDev development environments- Exploratory Data Analysis using R– Association – Classification – Clustering – Time series Analysis – Prediction Models- Data Visualization.		
`	**Contact Hours : 30**	

List of Experiments	
1	To perform operations using Operators and Flow Controls in python.
2	Demonstrate the Built -in and User defined functions in the following using python. a) Strings b) Lists c) Tuples
3	File Operation using python d) To read and write operations on a file. e) To copy the contents of a file to another file. f) To count the frequency of characters in a given file. g) To print each line of a file in reverse order.
4	To perform arrays and vectored computation using NumPy.
5	Implementation of data loading, storage and file formats of CSV, XLS and JSON using pandas.
6	To perform data wrangling operations using pandas. a) Merging b) Grouping c) Concatenating
7	Read and write different types of datasets in R. a) Reading different types of data sets (.txt, .csv) from web and disk and writing in file in specific disk location. b) Reading Excel data sheet. c) Reading XML dataset.
8	Implementation of vector data object operation using R.
9	Implementation and use of data frames in R. a) To select the rows where the score is missing, i.e. is NaN. b) To select the rows where the number of attempts in the examination is less than 2 and score greater than

	15. c) To calculate the mean score for each different student in the Data Frame. d) To sort the Data Frame first by 'name' in descending order, then by 'score' in ascending order. e) To count the city wise number of students from a given data set (city, name of the student).
10	Descriptive Statistics in R a) To find basic descriptive statistics using summary, str, quartile function on mtcars & cars datasets. b) To find subset of dataset by using subset (), aggregate () functions on iris dataset.
11	Build a regression model for the following. a) Financial forecasting (like house price estimates or stock prices, Beta and CAPM). b) Weather analysis and prediction. c) Time series forecasting.
12	Implementation of the clustering algorithms for the following. a) Fraud Detection. b) Website Recommendation. c) Market Basket Analysis.
13	Implementation of data visualization with ggplot2. a) Data Layer b) Aesthetic Layer c) Geometric Layer d) Facet Layer e) Statistics Layer f) Coordinates Layer

Contact Hours	**:**	**60**
Total Hours	**:**	**90**

Course Outcomes: On successful completion of the course, students will be able to	
●	Learn the basics of python programming and File Formats.
●	Understand about the Arrays and Objects Using Python Packages.
●	Understand the basic concepts of Data Processing and Visualization.
●	Apply programming for Data Wrangling and learn the basics of Visualization Concepts
●	learn the basics of R programming and implement Data Science algorithms using R

Text Books:	
1	Wes Mckinney “Python for Data Analysis”, Third Edition, Publisher O’Reilly Media, 2022.
2	Hadley Wickham & Garrett Grolemund, “R for DataScience”, Second Edition, O’Reilly Media, 2017.
3	Jake Vendor plus,”Python Data Science Handbook”, First Edition, 2016.

Reference Books:	
1	Frank kane, “Hands on Data Science and Python Machine Learning”, Second Edition, Packt Publishing, 2017.
2	David Taieb,” Data Analysis with Python: A Modern Approach “, First Edition, Packt Publishing, 2018.

Subject Code	Subject Name (Theory Course)	L	T	P	C
MCB1902	Digital Marketing and Web Analytics	3	0	0	3

Objectives:	
•	To gain marketing advantage by learning digital marketing fundamentals to achieve better user.
•	To increase brand awareness and visibility.
•	To develop customer engagement and loyalty.
•	To perform quantitative and qualitative analysis to give business that extra advantage.
•	Improve website usability and increase website traffic.

UNIT-I	**ONLINE MARKET SPACE**	**9**
Digital Marketing Strategy- Components -Opportunities for building Brand Website - Planning and Creation- Content Marketing. Case study: Build a digital branding strategy for a multinational apparel shop to help the brand establish itself as a new product in the market.		
UNIT-II	**TERMINOLOGY USED IN DIGITAL MARKETING**	**9**
PPC and online marketing through social media, Social Media Marketing, SEO techniques, Keyword advertising, Google web-master and analytics overview, Affiliate Marketing, Email Marketing, Mobile Marketing. Case study: Social media marketing using Facebook Ads Manager.		
UNIT-III	**DIGITAL MARKETING TECHNOLOGY**	**9**
Technology behind digital marketing - Evolution of digital marketing- Digital Marketing Strategy-10Ps of digital marketing-Choosing web designer / developer- Trust in Internet Marketing- Ethical and Legal Issues- Future of digital marketing. Case study: Application of Google Ads Manager in any Healthcare, Finance or Banking tracks.		
UNIT-IV	**WEB ANALYTICS**	**9**
Present and Future, Data Collection - Importance and Options, Overview of Qualitative Analysis, Business Analysis, KPI and Planning, Critical Components of a Successful Web Analytics Strategy, Web Analytics Fundamentals, Concepts, Proposals & Reports, Web Data Analysis. Case study: Application of Google Analytics in E-commerce track.		
UNIT-V	**SEARCH ANALYTICS**	**9**
Search engine optimization (SEO), non-linear media consumption, user engagement, user generated content, web traffic analysis, navigation, usability, eye tracking, online security, online ethics, content management system, data visualization, RSS feeds, Mobile platforms, User centered design, Understanding search behaviors.		
Contact Hours :		**45**

Course Outcomes: On completion of the course, students will be able to	
•	Know how to improve website visits and sales.
•	Develop a mass strategy and guide campaigns to increase sales and revenue.
•	Apply digital marketing strategy to increase customer lifetime value.
•	Perform web analytics process for better optimization.
•	Effectively use the search analytics insights to support brand recognition and ROI

Text Books:	
1.	Ryan Deiss & Russ Henneberry, “Digital Marketing for Dummies”, John Wiley Sons, Inc., 2020
2.	Dave Chaffey & Fiona Ellis-Chadwick, “Digital Marketing: Strategy, Implementation & Practice”, Sixth edition, Pearson, 2016.
3.	Dr.Anil Maheshwari, “Data Analytics Made Accessible”, 2023.

Reference Books / Web links:	
1.	K. M. Shrivastava, “Social Media in Business and Governance”, Sterling Publishers Private Limited, 2013.
2.	Christian Fuchs, “Social Media a Critical Introduction”, SAGE Publications Ltd, 2014.
3.	Bittu Kumar, “Social Networking”, V & S Publishers, 2013.
4.	Avinash Kaushik, “Web Analytics - An Hour a Day”, Wiley Publishing, 2007.
5.	T. Peterson, “Web Analytics Demystified”, Celilo Group Media and Café Press, 2004.
6.	TakeshiMoriguchi, “Web Analytics Consultant Official Textbook”, 7th Edition, 2016.

Subject Code	Subject Name (Theory Course)	L	T	P	C
MCB1903	**Operations and Supply Chain Analytics**	3	0	0	3

Objectives:	
•	To familiarize the fundamentals of operations and supply chain.
•	To learn techniques to optimize inventory levels and minimize waste.
•	To learn to create a responsive supply chain by reducing lead times.
•	To enhance decision-making capabilities by providing better data-driven insights through fundamentals of supply chain analytics with analytical tools.
•	To estimate possible outcomes which reduce costs associated with the supply chain.

UNIT-I	**INTRODUCTION**	**9**
Introduction to Operations management, Introduction to Demand Forecasting - Demand & Planning: Matching demand & supply, Inventory management Models. Role of analytics in supply chain, Supply chain strategies, Tools/Drivers for supply chain management, Framework for structuring Drivers.		
UNIT-II	**LOGISTICS & GLOBAL SUPPLY CHAIN MANAGEMENT**	**9**
Freight transportation: Selection & its impact on inventory, Warehousing: Design, Operations heuristics, Material handling, Customs, Duties, Tariffs, INCO terms, Rules of origin, Letter of credit etc. International transportation, Trading blocks, Trade zones, Bonded warehouses, Currency fluctuations, Exchange rate risks, Transfer pricing, Permanent establishment.		
UNIT-III	**SUPPLY CHAIN PERFORMANCE**	**9**
Lack of supply Chain Coordination and BULLWHIP, Managing uncertainties in a supply chain, Cycle Service Level, Monte Carlo Simulation, Tools and methods in supply chain risk management.		
UNIT-IV	**SUPPLY CHAIN ANALYTICS**	**9**
Introduction to analytics – descriptive, predictive and prescriptive analytics, Data Driven Supply Chains – Basics, transforming supply chains, Barriers to implementation, Road Map.		
UNIT-V	**MCDM MODELS**	**9**
Analytic Hierarchy Process (AHP), Data Envelopment Analysis (DEA), Fuzzy Logic and Techniques, the analytical network process (ANP), TOPSIS-Application in SCM.		
Contact Hours :		**45**

Course Outcomes: On completion of the course, students will be able to	
•	Understand supply chain fundamentals which improve customer service and satisfaction.
•	Enhance supply chain performance which increases profitability and cost savings.
•	Analyze models and strategies in inventory management which improves supplier relationships and better collaboration.
•	Reduce inventory costs and improve inventory management through analytics.
•	Make decision using multi-criteria in applications of SCM for better visibility and control of supply chain activities.

Text Books:	
1.	B. Mahadevan , “Operations Management Theory & Practice”, Third Edition , Pearson,2022.
2.	Sunil Chopra and Peter Meindl, “Supply Chain Management: Strategy, Planning, and Operation”, Pearson Education, 2019.
3.	Nada R. Sanders, “Big Data Driven Supply Chain Management: A Framework for Implementing Analytics and Turning Information Into Intelligence”, Pearson, 2015.
4.	Muthu Mathirajan, Chandrasekharan Rajendran, Sowmyanarayanan Sadagopan, Arunachalam Ravindran, Parasuram Balasubramanian, “Analytics in Operations/Supply Chain Management” , I.K. International Publishing House Pvt. Ltd., 2016.

Reference Books(s) / Web links:	
1	Nada R. Sanders, “Big data driven supply chain management: A framework for implementing analytics and turning information into intelligence”, Pearson Education, 2014.
2	Michael Watson, Sara Lewis, Peter Cacioppi, Jay Jayaraman, “Supply Chain Network Design: Applying Optimization and Analytics to the Global Supply Chain”, Pearson Education, 2013.
3	Anna Nagurney, Min Yu, Amir H. Masoumi, Ladimer S. Nagurney,”Networks Against Time: Supply Chain Analytics for Perishable Products”, Springer, 2013.
4	Muthu Mathirajan, Chandrasekharan Rajendran, Sowmyanarayanan Sadagopan, Arunachalam Ravindran, Parasuram Balasubramanian, “Analytics in Operations/Supply Chain Management” , I.K. International Publishing House Pvt. Ltd., 2016.
5	Gerhard J. Plenert, “Supply Chain Optimization through Segmentation and Analytics”, CRC Press, Taylor & Francis Group, 2014.