DATA VISUALIZATION
23MBPE664

(COURSE HANDBOOK)

MBA

COURSE HEADS: Ramya Bharadwaj B S/ Suchithra

1. GENERAL INFORMATION

Welcome to Data Visualization!

This course provides a comprehensive introduction to the principles, techniques, and tools of data visualization, a critical skill for making sense of complex data in today's information-driven world. Designed for aspiring data analysts, business professionals, and decision-makers, this course focuses on transforming raw data into insightful visual representations that drive understanding and decision-making.

The curriculum is structured into five interconnected modules, beginning with the fundamentals of data visualization, including its importance, history, and ethical considerations. Subsequent modules will cover the design principles for effective visualizations, exploration of various chart types, and hands-on experience with leading visualization tools such as R,Tableau, Power BI, and Python libraries like Matplotlib and Seaborn. You will also learn to interpret visualizations critically and communicate data stories effectively to diverse audiences.

Throughout the course, you will engage in practical, project-based activities aimed at enhancing your ability to design and present data visualizations. These activities will involve working with real-world datasets, analysing trends, and creating dashboards that provide actionable insights. Interactive discussions and group projects will further hone your ability to collaborate and interpret data in meaningful ways.

This course emphasizes experiential learning and practical application. You will be encouraged to explore creative visualization techniques, solve real-world business problems using data, and reflect on the impact of your visualizations. By the end of this module, you will not only have a strong foundation in data visualization theory but also the ability to create compelling visuals that tell impactful stories.

We are excited to embark on this data-driven journey with you and are confident that this course will enhance your analytical capabilities and visualization expertise. Be sure to review the course handbook, as it includes essential information on learning outcomes, assessment criteria, and additional resources. Let's make this semester an engaging and transformative learning experience!

1.1. Course Objectives
- **Equip students with foundational knowledge of data visualization** tools to effectively analyse and present data.

- **Provide Insight into Power BI and R Studio** to enable students to perform statistical analysis and create interactive dashboards using these mediums.

- **Impart the skills to design meaningful and visually compelling data visualizations** for business data exploration.

- **Familiarize with descriptive statistics and visualization techniques** to summarize and interpret business data.

1.2. Course Outcomes

- **CO1**: Summarize different data visualization tools.
- **CO2**: Perform statistical analysis and hypothesis testing using R for data-driven decision-making.
- **CO3**: Apply interactive and dynamic visualizations in Power BI and Tableau to drive actionable business insights.
- **CO4**: Apply descriptive statistics and data manipulation techniques to analyse large datasets and solve real business problems.

1.3. Set Text and Suggested Sources

All the below mentioned books are available in the 1st Floor Library.

Key Text Books:
1. Mark Gardner, "Beginning R: The Statistical Programming Language", 1st Edition, Wiley, 2017
2. Joshua N. Milliga "Learning Tableau 2020: Create effective data visualizations, build interactive visual analytics, and transform your organization", 4th Edition, 2020.
3. Dr. Shirshendu Roy "Data Visualization: Using Power BI, Orange and Excel", 1st Edition Notion Press, 2021.

Reference Books:
1. Selva Prabhakaran, "Mastering R Programming", 1st Edition, PAKT Publishers, 2017.
2. Chandraish Sinha "Mastering Power BI", 2nd Edition, BPB Publications, 2024
3. Alexander Loth "Visual Analytics with Tableau", Wiley, 2019.

1.4.Self-Study Course

In this course, students can take up the "Zero to Hero in Tableau: Data Visualization using Tableau" offered by Udemy, which would be great use for the learning of the syllabus. This digital literacy program is essential as it provides students with valuable skills that complement the traditional learning methods employed in our course. By engaging with these certifications students can enhance their understanding of digital tools and analytical techniques that are increasingly important in today's academic and professional environments.

Note: The above title is hyperlinked and directly leads to the course to be undertaken by the students. Please register using the same links or stay logged in to be able to access the courses.

2. THE COURSE

2.1. Course Description

DATA VISUALIZATION			
Semester	**III**	CIE Marks	**50**
Course Code	**23MBPE664**	SEE Marks	**50**
Teaching Hrs/Week (L:T:P)	**4:0:0**	Exam Hrs	**03**
Total Hrs	**52**	Credits	**04**

The Data Visualization course designed to provide students with foundational knowledge in visualizing and interpreting data effectively. This course emphasizes the principles of creating impactful and accurate visual representations of data. The course will run for 13 weeks during Semester III and consists of 5 modules that cover essential topics in data visualization. Each week includes 4 lectures, delivered by Professor of Practice, focusing on theoretical concepts, practical applications, and course-related activities. Spanning a total of 52 hours, this 4-credit course is assessed through Continuous Internal Evaluation (CIE) for 50 marks and a Semester-End Examination (SEE) for 50 marks. This structure ensures a balanced and engaging learning experience for students.

2.2. Initiating Contact with Staff and Other Students

We encourage open communication and value your inquiries about the Course. The most effective means are to directly enquire during class hours and faculty office hours (with appointments). Given the large number of students in the course, you are advised to go through the online materials i.e. this handbook and the official website to check if your query can be answered. Additionally, we encourage you to engage with your peers for discussions and collaborative learning, as this will enhance your understanding of the course material and foster a supportive academic community.

2.3. Resources

Along books, resources include dynamic tools like digital libraries, e-learning platforms, and research databases. Students can access these through the college website. Some of the key resources include the VTU Consortium, e-learning platforms, and additional sources like open-access repositories, government portals (e.g., NPTEL, NDLI). These digital tools provide access to e-books, research papers, video lectures, and interactive tutorials, offering flexible and comprehensive learning environments.

E-learning and digital library can be accessed via the college website https://mite.ac.in/ (Campus Life section > Library > VTU Consortium/e-learning platforms/additional sources).

Apart from this, students opting for Business Analytics as their MBA specialization are encouraged to explore the following independent resources to enhance their knowledge:
- International Institute of Business Analytics
- Digital Analytics Association
- Analytics Society of India
- CompTIA
- Institute for Operations Research and the Management Sciences

2.4.Staff

Course Faculty: Suchithra
Cabin: 3rd Floor, PG Block
Email: suchithra@mite.ac.in

Course Faculty: Ramya Bharadwaj B S
Cabin: 3rd Floor, PG Block
Email: ramya@mite.ac.in

2.5.Topics and Reading materials for each module

Module 1   *No. of Hours: 11*

- **Topic: Data Visualization with Power BI**
    o Data sources in Power BI, Power BI Desktop - loading data, views, query editor, transform, clean, shape, and model data, manage data relationship, editing a relationship, cross filter direction, saving work file, measures.
    o Visualization in Power BI: visualization charts in Power BI, matrixes and tables, slicers and map visualizations, gauges and single number cards, modifying colours in charts and visuals, shapes, text boxes, and images, custom visuals, page layout and formatting, bookmarks and selection pane, KPI visuals, Z-order, grouping and Binning.

- **Activities:**
    o Students are given with assignments to use Power BI to analyse and present relatively huge data within a deadline.

- **Essential Readings:**
    o Dr. Shirshendu Roy "Data Visualization: Using Power BI, Orange and Excel", 1st Edition Notion Press, 2021, Chapter 1,2and 3.

- **Additional Reading:**
    o Chandraish Sinha "Mastering Power BI", 2nd Edition, BPB Publications, 2024, Chapter 1 and 2.

Module 2   *No. of Hours: 9*

- **Topic: R Studio interface**
    o Basic operations, working with R script and R Markdown. Basic data types: Vectors, matrices, data frames, lists. Importing data from different formats: CSV, Excel, and databases, Data inspection. Data cleaning: Handling missing values, outliers, and duplicates. Data transformation: Renaming columns, sub setting, merging data frames, Basic data manipulation with dplyr package.

- **Activities:**
    o Students will be given with assignments on data cleaning and manipulation using R, with a deadline for submission

- **Essential Reading:**
  - o Mark Gardner, "Beginning R: The Statistical Programming Language", 1stEdition, Wiley, 2017, Chapter 1 and 2.

- **Additional Reading:**
  - o Selva Prabhakaran, "Mastering R Programming", 1st Edition, PAKT Publishers, 2017, Chapter 1 and 2.

Module 3   *No. of Hours: 12*

- **Topic: Descriptive Statistics and Data Visualization**
  - o Measures of central tendency: Mean, median, mode. Measures of dispersion: Range, variance, standard deviation, inter-quartile range. Visualizing distributions: Histograms, boxplots, bar charts Data Visualization with ggplot2.
    Common probability distributions: Normal, Binomial, Poisson. Formulating null and alternative hypotheses. Types of errors: Type I and Type II errors. One-sample and two-sample tests: t-tests, z-tests, Chi-square test for categorical variables. Conducting hypothesis tests in Regression and Correlation Analysis.

- **Activities:**
  - o Students will be given with assignments on statistical data analysis (Descriptive) using ggplot2, with a deadline for submission.

- **Essential Reading:**
  - o Mark Gardner, "Beginning R: The Statistical Programming Language", 1stEdition, Wiley, 2017, Chapter 4, 5 and 6.

- **Additional Reading:**
    Selva Prabhakaran, "Mastering R Programming", 1st Edition, PAKT Publishers, 2017, Chapter 2 and 3.

Module 4   *No. of Hours: 10*

- **Topic: Data Exploration and Basic Visualizations**
  - o Understanding data types in Tableau (dimension vs. measure), Creating and managing data connections: Connecting to files and databases. Data blending and joining tables, using Data Interpreter for cleaning and preparing data. Best practices for organizing data for analysis.

- **Activities:**
  - o Students will be given with assignments on data visualization using Tableau pertaining to tables.
- **Essential Reading**:
  - o Joshua N. Milliga "Learning Tableau 2020: Create effective data visualizations, build interactive visual analytics, and transform your organization", 4th Edition, 2020, Chapter1, 2, 4 and 5.

- **Additional Reading:**
    o Alexander Loth "Visual Analytics with Tableau", Wiley, 2019, Chapter 1.

Module 5   *No. of Hours: 10*

- **Topic: Tableau Worksheets**
    o Understanding the structure of Tableau worksheets: sorting, filtering, and grouping data, creating calculated fields and basic aggregations, using filters and parameters to refine data views, Sorting, grouping, and creating hierarchies in dimensions
    Basic visualizations and charts, types of charts in Tableau: bar charts, line charts, pie charts, histograms, etc. Creating simple visualizations using dimensions and measures, understanding how to choose the right visualization for data, formatting and customizing visualizations.

- **Activities:**
    o Students will be given with assignments on data visualization using Tableau pertaining to charts.

- **Essential Reading:**
    o Joshua N. Milliga "Learning Tableau 2020: Create effective data visualizations, build interactive visual analytics, and transform your organization", 4th Edition, 2020, Chapter 6, 7, 9 and 10.

- **Additional Reading:**
    o Alexander Loth "Visual Analytics with Tableau", Wiley, 2019, Chapter1, 2 and 3.

3. **ASSESSMENT**

The assessment for Data Visualization course is divided into two components: Continuous Internal Evaluation (CIE) and Semester End Examination (SEE), each accounting for 50% of the total marks.

Continuous Internal Evaluation (CIE) consists of two internal tests, scheduled for the 8th and 14th weeks, contributing a total of 30% to the overall marks. Students can earn the other 20% through assignments, which are allocated as follows:
- 10 marks for assignment data cleaning and manipulation using R
- 10 marks for assignment on data visualization using Tableau pertaining to charts.

Semester End Examination (SEE) constitutes the remaining 50% of the total marks. Key information regarding examination dates and related details can be accessed via the college website (Academics and Courses section > Calendar of Events > PG Odd Sem).

Rubrics for Assignment Evaluation (Total: 20 Marks / 40% of CIE)

Assignments on R & Tableau (10 Marks Each for 2 assignments)				
Criteria	**10 Marks (Excellent)**	**9 Marks (Good)**	**8 Marks (Fair)**	**5-6 Marks (Poor)**
Completion	Completed on time.	Completed slightly late.	Completed much later than due.	Missed the deadline.