



Data Visualization and Analytics (CAP401)



VII Semester (CSE[AIML])



Data Visualization



- Data visualization is the **graphical representation** of information and data.
- Caters to **non-technical audience**
- Uses **visual elements** like charts, graphs, and maps
- Data visualization tools provide an accessible way to **see and understand trends, outliers, and patterns in data.**

Traditional Tools: PowerPoint, Excel, Dashboard and business intelligence platforms

Latest: Power BI, Tableau, google charts, Datawrapper, Zoho Reports, FusionCharts, Infogram, Ploymaps, RapidMiner, Gephi, etc.

Data Analysis

- Process of cleaning, analyzing, and interpreting data.
- To inform decision-making and discover advantages using data
- Gain insight from your data
- Companies use data analysis to gain a better understanding of their internal processes, products, customers, the market, and competitors

Data Analytics

- Data analysis is an important step in data analytics strategy.
- The term “data analytics” its not only analysis.
- Data analytics includes how data is stored, governed, organized, then analyzed
- It also includes tools, techniques, and strategies involved in those efforts.
- Although analysis is a core focus of every effective data analytics strategy, it is just a part

Course Objectives

1. To understand data analytics life cycle for solving challenging business problems.
2. To adopt appropriate statistical procedures for analysis based on goals and nature of data.
3. To employ best practices in data visualization to develop charts, maps tables and
other visual representations of data.

SYLLABUS

Lecture: 03 Hrs/per, Credits: 03

UNIT-I:

Importance of analytics and visualization, data preprocessing, Data categorization, Levels of Measurement, Data management and indexing, Descriptive statistics , Data Analytics Lifecycle, and different Phases

UNIT-II:

Basic Analysis Techniques : Statistical hypothesis generation and testing, Chi-Square test, t-Test, Analysis of variance, Correlation analysis, Maximum likelihood test.

SYLLABUS

UNIT-III:

Basic concepts of probability, random variables, probability distributions, sampling and estimation, statistical inference.

Unit IV:

Time Series Analysis: Box-Jenkins Methodology, ARIMA (Auto Regressive Integrated Moving Average) Model, Choice of a Model, Overview of ARMAX, Spectral Analysis and GARCH.

SYLLABUS

Unit V:

Understanding Data Visualization Principles, Mapping Data onto Aesthetics, Visualizing - Distributions, Proportions, Time Series, Trends and Uncertainty; Commonly used File Formats and Software.

Unit VI:

Why Planning?, Creating Interesting Stories with Data – Reader-driven Narratives, Author-driven Narratives; Perceptions and Presentation Methods, Best Practices in Visualization, Interactive Visualization, Event Listeners and Layouts, Case Studies for Visualization.

Course Outcome

After successful completion of this course, the student will be able to:

1. Apply data pre-processing and basic data analysis techniques.
2. Conduct data analytics using scientific methods.
3. Analyze time series data.
4. Create presentations and visualizations.

Text Books

1. David Dietrich, Barry Heller and Beibel Yang, - Data Science and Big Data Analytics – Discovering, Analyzing, Visualizing, and Presenting Data, John Wiley and Sons [EMC Education Services], 2015.
2. Claus O. Wilke, - Fundamentals of Data Visualization – A Primer on Making Informative and Compelling Figures, O'Reilly, 2019.
3. Python: Data Analytics and Visualization, Packt Publishing, 2017.

Reference Books

1. Jiawei Han, Micheline Kamber and Jian Pei, - Data Mining Concepts and Techniques, 3rd edition; Morgan Kaufmann Publishers, 2011.

Happy Learning...!!

