

INT350:DATA VISUALISATION

L:2 T:0 P:3 Credits:4

Course Outcomes: Through this course students should be able to

- CO1 :: understand the basic concepts of database design.
- CO2 :: determine the data modelling techniques and apply their use on databases
- CO3 :: apply the best practices used in SQL
- CO4 :: demonstrate the python concepts to visualize the data
- CO5 :: understand and apply the basic functions of tableau on the data
- CO6 :: analyze the data using dashboards and apply advanced tableau functions

Unit I

Database design and introduction to MySQL : Data warehouse, ERD, star and snowflake schemas, OLAP vs OLTP, entity constraints, referential constraints, semantic constraints, comprehension: ERD, introduction to SQL, DDL statements, DML statements, SQL basic statements and operators, aggregate and inbuilt functions, string and date-time functions and ordering, regular expressions, nested queries, views, venn diagrams and inner and outer joins, left and right join, cross join join, views with join, intersect, minus, union and union all

Unit II

Data modelling : introduction to data modelling, a data model vs a floor model, database design - creation - manipulation cycle, relational schemas, relational vs non-relational schemas, database design, DDL statements syntax, database creation, DML statements syntax, database manipulation, database querying

Unit III

Advanced SQL and best practices : rank functions, partitioning, frames, lead and lag functions, case statements, UDFs, stored procedures, cursors, best practices, indexing, clustered vs non-clustered indexing, order of query execution, joins vs nested queries, profitability analysis, profitable customers, customers without orders, fraud detection

Unit IV

Data visualization in python : the necessity of data visualization, data handling and cleaning, sanity checks, outliers analysis with boxplots, histograms, distribution plots, styling options, pie - chart and bar chart, scatter plots, pair plots, bar graphs and box plots, heatmaps, line charts, stacked bar charts, plotly

Unit V

Basic visualization using tableau : introduction to data analytics, why data visualization?, what is tableau?, why tableau?, tableau vs excel and power BI, exploratory vs explanatory analysis, getting started with tableau, bar charts, line charts and filters, area charts, box plots and pivoting, maps and hierarchies, pie charts, treemaps and grouping

Unit VI

Advanced visualization using tableau : dashboards - I, joins and splits, numeric and string functions, logical and date functions, histograms and parameters, scatter plots, dual axis charts, top n parameters and calculated fields, stacked bar charts, dashboards -II and filter actions, storytelling

List of Practicals / Experiments:

List of practical

- write SQL queries to create and manipulate tables.
- write SQL queries to retrieve specific details from a given table.
- write SQL queries to demonstrate the use of aggregate functions, inbuilt functions, string, datetime function, and ordering.
- write SQL queries to use the nested queries.
- creating tableau workbook containing multiple dashboards pertaining to the multiple categories
- demonstrating the match, player and team statistics using tableau

- Demonstrating connecting different data sets in tableau
- fetching the data from different sources and prepare it for further analysis.
- Visualization to understand the hierarchies in data and drill down approaches.
- visualization of data using bar and stacked bar charts, line and area charts, box and pie charts.
- visualization of data using scatter plots
- visualization of data using grouping and tree maps
- creating histograms using tableau.
- demonstrating the power BI

References:

1. SQL FOR DATA ANALYSIS: ADVANCED TECHNIQUES FOR TRANSFORMING DATA INTO INSIGHTS by CATHY TANIMURA, SHROFF/O'REILLY
2. SQL COOKBOOK by ANTHONY MOLINARO, SHROFF/O'REILLY
3. HANDS-ON DATA VISUALIZATION: INTERACTIVE STORYTELLING FROM SPREADSHEETS TO CODE by JACK DOUGHERTY, SHROFF/O'REILLY