**B.Sc., Data Science Honours**

**MAJOR**

**w.e.f AY 2024-25 onwards**

| Year | Semester | Paper | Title of the Course | No. of Hrs./ Week | No. of Credits |
|---|---|---|---|---|---|
| II | IV | 9 | Data visualization using Tableau | 3 | 3 |
| | | | Data visualization using Tableau Practical Course | 2 | 1 |
| | | 10 | Data visualization using python | 3 | 3 |
| | | | Data visualization using python Practical Course | 2 | 1 |
| | | 11 | Introduction to SQL & Advanced Tableau | 3 | 3 |
| | | | Introduction to SQL & Advanced Tableau Practical Course | 2 | 1 |

**SRI VENKATESWARA UNIVERSITY::TIRUPATI**
**B.Sc. Data Science Honours**
**II Year IV Semester**
**Course 9: Data visualization using Tableau**
**(w.e.f. 2024-25)**

**Theory** **Credits: 3** **3 hrs/week**

**Aim and objectives of Course:**

- To know the importance of data Visualization in the world of Data Analytics and Prediction
- To know the important libraries in Tableau
- To get equipped with Tableau Tool

**Learning outcomes of Course:**

- Students should be able to visualize data through seven stages of data analysis process
- Should be able to do explanatory and hybrid types of data visualization
- Should be able to understand various stages of visualizing data

**UNIT I:**

Creating Visual Analytics with tableau desktop, connecting to your Data-How to Connect to your data, what are generated Values? Knowing when to use a direct connection, joining tables with tableau, blending different data sources in a single worksheet.

**UNIT II:**

**Building your first Visualization-** How Me works- Chart types, Text Tables, Maps, bar chart, Line charts, Area Fill charts and Pie charts, scatter plot, Bullet graph, Gantt charts, Sorting data in tableau, Enhancing Views with filters, sets groups and hierarchies.

**UNIT III:**

**Creating calculations to enhance your data-** What is aggregation, what are calculated values and table calculations, Using the calculation dialog box to create, Building formulas using table calculations, Using table calculation functions

**UNIT IV:**

**Using maps to improve insights-**Create a Standard Map View, Plotting your own locations on a map, Replace Tableau's standard maps, Shaping data to enable Point-to-Point mapping.

**UNIT V:**

**Developing an Adhoc analysis environment-** generating new data with forecasts, providing self evidence adhoc analysis with parameters, Editing views in tableau Server.

**TEXT BOOKS:**


1. Tableau your data-Daniel G. Murray and the Inter works BI team, Wiley Publications
2. Tableau Data Visualizaton Cookbook, Ashutosh Nandeshwar, PACKT publishing.
3. Storytelling with Data: A Data Visualization Guide for Business Professionals by Cole Nussbaumer Knaflic (2014)
4. ggplot2: Elegant Graphics for Data Analysis by Hadley Wickham (2009)


**REFERENCE BOOKS:**


1. Designing Data Visualizations: Representing Informational Relationships by Noah Iliinsky, Julie Steele (2011)
2. Alexandru C. Telea – "Data Visualization principles and practice" Second Edition, CRC Publications
3. Joshua N. Millign–" Learning Tableau -2019" – Third Edition- Packt publications

Student Activity

Create a sample super store data set and visualize the following requirements

General Requirements

1. Dashboard size is 1250px wide by 750px tall.
2. Prefer using containers
3. The dashboard has a total of 5 containers (no more, no less)
4. The Filter Pane
5. Each filter has some padding

**1.** Charts Pane Requirement

1. All 3 charts must be in one vertical container
2. Do proper formatting
3. Each chart has some padding between them and other objects
4. Each chart has a grey border, slightly darker than the Pane background color.
5. The Pane under the Title has a border

**2.** The second graph should have the title as "Sales" and should show monthly sales per year. Make sure it is an area chart with proper formatting.

**3.** The third graph should the title as "Profit" and should show monthly profit per year. Make sure it is an area chart with proper formatting.

Continuous assessment:

Let the students be tested in the following questions from each unit

1. What are generated values? Join tables using Tableau
2. Create any visualization charts using Chart types, Text Tables, Maps, bar chart, Line charts, Area Fill charts and Pie charts, scatter plot etc.,
3. What is aggregation, what are calculated values and table calculations?
4. Using Standard Map View, Plot your own locations on a map
5. Develop an Adhoc analysis environment.

**B.Sc. Data Science Honours**
**II Year IV Semester**
**Course 9: Data visualization using Tableau**
**(w.e.f. 2024-25)**

**Practicals** **Credits: 1** **2 hrs/week**

**List of Experiments:**

1. Connect to data Sources
2. Create Univariate Charts
3. Create Bivariate and Multivariate charts
4. Create Maps
5. Calculate user-defined fields
6. Create a workbook data extract
7. Save a workbook on a Tableau server and web
8. Export images, data.

**SRI VENKATESWARA UNIVERSITY::TIRUPATI**
**B.Sc. Data Science Honours**
**II Year IV Semester**
**Course 9: Data visualization using Tableau**
**(w.e.f. 2024-25)**
MODEL QUESTION PAPER

Time: **3 Hours** Max. Marks: **75**

**PART-A**

Answer any **FIVE** of the following. Each Question Carries 5 marks. (5X5=25)

1. **Short answer question from unit-I**
2. **Short answer question from unit-I**
3. **Short answer question from unit-II**
4. **Short answer question from unit-II**
5. **Short answer question from unit-III**
6. **Short answer question from unit-III**
7. **Short answer question from unit-IV**
8. **Short answer question from unit-IV**
9. **Short answer question from unit-V**
10. **Short answer question from unit-V**

**PART-B**

Answer any **FIVE** of the following. Each Question Carries 10 marks. (5X10=50)

**11. a) Essay question from Unit-I**
**OR**
**b) Essay question from Unit-I**

**12. a) Essay question from Unit-II**
**OR**
**b) Essay question from Unit-II**

**13. a) Essay question from Unit-III**
**OR**
**b) Essay question from Unit-III**

**14. a) Essay question from Unit-IV**
**OR**
**b) Essay question from Unit-IV**

**15. a) Essay question from Unit-V**
**OR**
**b) Essay question from Unit-V**

**SRI VENKATESWARA UNIVERSITY::TIRUPATI**
**B.Sc. Data Science Honours**
**II Year IV Semester**
**Course 10: Data visualization using Python**
**(w.e.f. 2024-25)**

**Theory** **Credits: 3** **3 hrs/week**

**Course Objective:**

This course introduces students to data analysis and visualization in the field of exploratory data science using Python.

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

- Use data analysis tools in the pandas library.
- Load, clean, transform, merge and reshape data.
- Create informative visualization and summarize data sets.
- Analyze and manipulate time series data.
- Solve real world data analysis problems.

**Unit 1**

Introduction: Introduction to Data Science, Exploratory Data Analysis and Data Science Process. Motivation for using Python for Data Analysis, Introduction of Python Jupyter Notebook. Essential Python Libraries: NumPy, pandas, matplotlib, SciPy, scikit-learn, stats models, seaborn.

**Unit 2**

Getting Started with Pandas: Arrays and vectorized computation, Introduction to pandas Data Structures, Essential Functionality, Summarizing and Computing Descriptive Statistics. Data Loading, Storage and File Formats. Reading and Writing Data in Text Format, Web Scraping, Binary Data Formats, interacting with Web APIs, Interacting with Databases Data Cleaning and Preparation. Handling Missing Data, Data Transformation, String Manipulation

**Unit 3**

Data Wrangling: Hierarchical Indexing, Combining and Merging Data Sets Reshaping and Pivoting. Data Visualization matplotlib: Basics of matplotlib, plotting with pandas and seaborn, other python visualization tools. Advanced categorical and numeric plots.

**Unit 4**

Data Aggregation and Group operations: Group by Mechanics, Data aggregation, General split-apply-combine, Pivot tables and cross tabulation

Time Series Data Analysis: Date and Time Data Types and Tools, Time series Basics, date Ranges, Frequencies and Shifting, Time Zone Handling, Periods and Periods Arithmetic, Resampling and Frequency conversion, Moving Window Functions.

**Unit 5 Advanced Pandas:**

Categorical Data: cleaning data and visualization techniques, Advanced Group By methods, Use Techniques for Method Chaining.

**Textbook:**

1. McKinney, W.(2017). Python for Data Analysis: Data Wrangling with Pandas, NumPy and IPython. 2nd edition. O'Reilly Media. Reference:
1. O'Neil, C., & Schutt, R. (2013). Doing Data Science: Straight Talkfrom the FrontlineO'Reilly Media.

**B.Sc. Data Science Honours**
**II Year IV Semester**
**Course 10: Data visualization using Python**
**(w.e.f. 2024-25)**

**Practicals** **Credits: 1** **2 hrs/week**

**List of Experiments:**

1. Practicals based on NumPy and array
2. Practicals based on Pandas Data Structures
3. Practicals based on Data Loading, Storage and File Formats
4. Practicals based on Interacting with Web APIs
5. Practicals based on Data Cleaning and Preparation
6. Practicals based on Data Wrangling
7. Practicals based on Data Visualization using matplotlib
8. Practicals based on Data Aggregation
9. Practicals based on Time Series Data Analysis

**SRI VENKATESWARA UNIVERSITY: : TIRUPATI**
**B.Sc. Data Science Honours**
**II Year IV Semester**
**Course 10: Data visualization using Python**
**(w.e.f. 2024-25)**
MODEL QUESTION PAPER

Time: **3 Hours** Max. Marks: **75**

**PART-A**

Answer any **FIVE** of the following. Each Question Carries 5 marks. (5X5=25)

1. **Short answer question from unit-I**
2. **Short answer question from unit-I**
3. **Short answer question from unit-II**
4. **Short answer question from unit-II**
5. **Short answer question from unit-III**
6. **Short answer question from unit-III**
7. **Short answer question from unit-IV**
8. **Short answer question from unit-IV**
9. **Short answer question from unit-V**
10. **Short answer question from unit-V**

**PART-B**

Answer any **FIVE** of the following. Each Question Carries 10 marks. (5X10=50)

**11. a) Essay question from Unit-I**
**OR**
**b) Essay question from Unit-I**

**12. a) Essay question from Unit-II**
**OR**
**b) Essay question from Unit-II**

**13. a) Essay question from Unit-III**
**OR**
**b) Essay question from Unit-III**

**14. a) Essay question from Unit-IV**
**OR**
**b) Essay question from Unit-IV**

**15. a) Essay question from Unit-V**
**OR**
**b) Essay question from Unit-V**

**SRI VENKATESWARA UNIVERSITY: : TIRUPATI**
**B.Sc. Data Science Honours**
**II Year IV Semester**
**Course 11: Introduction to SQL & ADVANCED TABLEAU**
**(w.e.f. 2024-25)**

**Theory** **Credits: 3** **3 hrs/week**

**Learning Objectives:**

- Design a database using DBMS softwares.
- Perform SQL queries on database.
- Use Tableau's visualization tools to conduct data analysis, especially exploration of an unfamiliar dataset.

**Course Outcomes:**

- Design a database by their own and perform simple and adhoc queries.
- Employ best practices in data visualization to develop charts, maps, tables, and other visual representations of data.
- Create compelling, interactive dashboards to combine several visualizations into a cohesive and functional whole.
- Utilize advanced Tableau features including parameters, data blending, custom SQL, very large data

**Unit I:**

Overview of Database Management System: Introduction to data, information, database, database management system. The building blocks of an entity relationship diagram, classification of entity sets, attribute classification, relationship degree, relationship classification, reducing ER diagram to tables,

**Unit 2:**

Structured Query Language: Introduction, History of SQL Standard, Commands in SQL, Data Types in SQL, Data Definition Language, Data Manipulation Language, database constraints, Aggregate functions, Join Operation, Set Operations, Views. SQL queries, sub queries and corelated queries,

**Unit 3:**

Optimal visualization types – bar chart, pie chart, gantt chart, bubble chart, bullet chart, scatter plot, line chart, heat map, tree map Maps- geographical locational plotting Binning values, Calculated fields, Table calculations, Level of Detail calculations.

**Unit 4:**

Dashboard development, Dashboard design principles, dashboard interactivity, connected "drill-down" dashboards Best Practices, Creating visualizations with Tableau.

**Unit 5:**

Advanced Tableau, Large datasets, Fiscal Year Calculations, Parameters, tableau scripting, tableau server, integration of tableau with R programming.

**Textbooks:**

1. Show me the Numbers: Designing Tables and Graphs to Enlighten by Stephen Few
2. The Data Loom: Weaving Understanding by Thinking Critically and Scientifically with Data by Stephen Few

**Reference Books:**

1. The Big Book of Dashboards: Visualizing your Data using Real-World Business Scenarios by Steve Wexler, Jeffrey Shaffer, and Andy Cotgreave

**B.Sc. Data Science Honours**
**II Year IV Semester**
**Course 11: Introduction to SQL & ADVANCED TABLEAU**
**(w.e.f. 2024-25)**

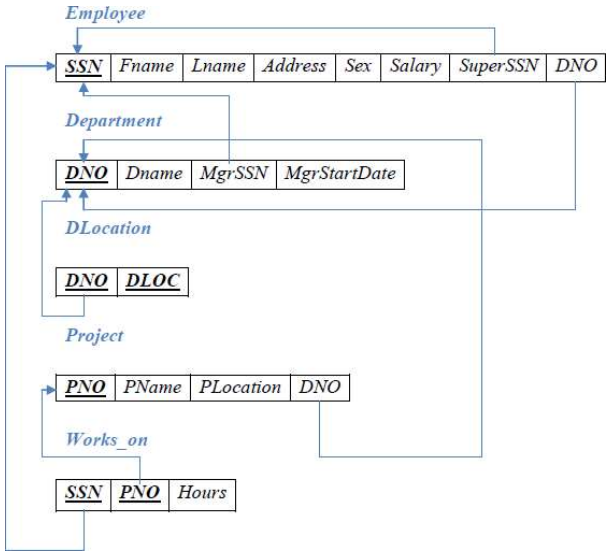
**Practicals** **Credits: 1** **2 hrs/week**

**List of Experiments:**

**DATABASE MANAGEMENT SYSTEM LAB**

**Consider following databases convert entities and relationships torelation table for a given scenario.**

1. **COLLEGE DATABASE:**
   STUDENT (stno, SName, Address, Phone, Gender)
   COURSE(courseid, Sem, Sec)
   CLASS (stno, courseid)
   SUBJECT (Subcode, Title, Sem, Credits)
   MARKS (stno, Subcode, courseid, Test1, Test2, Test3, total)
2. **COMPANY DATABASE:**
   EMPLOYEE (SSN, Name, Address, Sex, Salary, SuperSSN, DNo)
   DEPARTMENT (DNo, DName, MgrSSN, MgrStartDate) DLOCATION (DNo,DLoc)
   PROJECT (PNo, PName, PLocation, DNo)WORKS_ON (SSN, PNo, Hours)
3. **Consider a college database schema**




1. Create above tables with relevant Primary Key, Foreign Key and other constraints
2. Populate the tables with data

4. **Perform queries to generate outputs:**

1. Display all the details of all employees working in the company.
2. Display ssn, lname, fname, address of employees who work in department no 7.

3. Retrieve the Birthdate and Address of the employee whose name Is 'Franklin T.Wong'
4. Retrieve the name and salary of every employee.
5. Retrieve all distinct salary values
6. Retrieve all employee names whose address is in ‘Bellaire’
7. Retrieve all employees who were born during the 1950s
8. Retrieve all employees in department 5 whose salary is between 50,000 and 60,000 (inclusive)

**5. Perform the following queries**

1. Retrieve the names of all employees who do not have supervisors
2. Retrieve SSN and department name for all employees
3. Retrieve the name and address of all employees who work for the 'Research' department
4. For every project located in 'Stafford', list the project number, the controlling department number, and the department manager's last name, address, and birth date.
5. For each employee, retrieve the employee's name, and the name of his or her immediate supervisor.
6. Retrieve all combinations of Employee Name and Department Name
7. Make a list of all project numbers for projects that involve an employee whose last name is 'Narayan’ either as a worker or as a manager of the department that controls the project.
8. Increase the salary of all employees working on the 'ProductX' project by 15%. Retrieve employee name and increased salary of these employees.
9. Retrieve a list of employees and the project name each works in, ordered by the employee's department, and within each department ordered alphabetically by employee first name.
10. Select the names of employees whose salary does not match with salary of any employee in department.

**6. Perform following queries :**

1. Retrieve the employee numbers of all employees who work on project located in Bellaire, Houston, or Stafford.
2. Find the sum of the salaries of all employees, the maximum salary, the minimum salary, and the average salary. Display with proper headings
3. Find the sum of the salaries and number of employees of all employees of the ‘Marketing’ department, as well as the maximum salary, the minimum salary, and the average salary in this department.
4. Select the names of employees whose salary is greater than the average salary of all employees in department 10.
5. Delete all dependents of employee whose ssn is ‘123456789’.
6. Perform a query using alter command to drop/add field and a constraint in Employee table.
7. Format your data using filters with colors
8. create dashboards and stories.
9. Distribute and publish your visualization.
10. create advanced mapping –
    1. point-to-point map
    2. Dual axis map
11. Calculate distance between two points on a map.

**SRI VENKATESWARA UNIVERSITY: : TIRUPATI**
**B.Sc. Data Science Honours**
**II year IV Semester**
**Course 11: Introduction to SQL & ADVANCED TABLEAU**
**(w.e.f. 2024-25)**
MODEL QUESTION PAPER

Time: **3 Hours** Max. Marks: **75**

**PART-A**

Answer any **FIVE** of the following. Each Question Carries 5 marks. (5X5=25)

1. **Short answer question from unit-I**
2. **Short answer question from unit-I**
3. **Short answer question from unit-II**
4. **Short answer question from unit-II**
5. **Short answer question from unit-III**
6. **Short answer question from unit-III**
7. **Short answer question from unit-IV**
8. **Short answer question from unit-IV**
9. **Short answer question from unit-V**
10. **Short answer question from unit-V**

**PART-B**

Answer any **FIVE** of the following. Each Question Carries 10 marks. (5X10=50)

**11. a) Essay question from Unit-I**
**OR**
**b) Essay question from Unit-I**

**12. a) Essay question from Unit-II**
**OR**
**b) Essay question from Unit-II**

**13. a) Essay question from Unit-III**
**OR**
**b) Essay question from Unit-III**

**14. a) Essay question from Unit-IV**
**OR**
**b) Essay question from Unit-IV**

**15. a) Essay question from Unit-V**
**OR**
**b) Essay question from Unit-V**