Avinashilingam Institute for Home Science and Higher Education for Women

(Deemed to be University under Category 'A' by MHRD, Estd. u/s 3 of UGC Act 1956)
Re-accredited with 'A+' Grade by NAAC.Recognised by UGC under Section 12 B
Coimbatore - 641 043, Tamil Nadu, India

Bachelor of Vocation Degree Examination – June 2021
IV Semester

Class : II B.Voc **Time: 3 Hrs**
Major : Artificial Intelligence and Machine Learning **Max. Marks: 100**

19VAIC12 Data Visualization Techniques

Part A **10 x 1 = 10**
Choose the correct answer

1. _____ method shows hierarchical data in a nested format. CO2K1
   a.Treemaps b. Scatter plots c.Population pyramids d. Area charts

2. The function used to query and edit graphical settings is _____. CO2K2
   a.anova() b.par() c.plot() d. cum()

3. The lattice command used for producing boxplots is _____. CO2K1
   a.plot() b.bwplot() c.xyplot() d.barlm()

4. The purpose of **correct.cor** is to correct _____ in values. CO3 K1
   a.difference b.reliability c.error d.similar

5. The method used to find the position of a quantile in a dataset is _____ CO1K2
   a.quantile() b.barplot() c.barchart() d.rep()

6. Function used for linear regression in R is _____ CO3 K2
   a.lm(formula, data. b.lr(formulat,data.
   c.lrm(formula,data. d.legression.linear(formula,data.

7. The parameter **β0** and **β1**are usually called as CO3 K2
   a. Regressionists b. Coefficients c. Regressive d.Regression Coefficients

8. A mosaic plot is used when graphing CO3K1
   a. relationship between two continuous variables
   b. relationship between one continuous and one categorical variable
   c. relationship between two categorical variables
   d. data that are not normally distributed by group

9. _____ function opens the ggplot2 library. CO1K2
   a. Install.packages("ggplot2") b. Library(package = "ggplot2")
   c. Summary(object = ggplot2) d. Open(x=ggplot2)

10. _____ level plotting function starts a new plot. CO4K1
    a. High b. Low c. Medium d. Simple

Part B **5x6 = 30**

Answer the following

Answer should not exceed 400 words or two pages

11.a. Illustrate with examples the methods used for importing data from text files and excel spreadsheets into data frames. CO1 K3

(or)

11.b. Discuss the methods used for variable selection, observation and summarization during data cleaning process. CO1K2

12.a. Discuss with examples the difference between Univariate and Bivariate Graphs. CO2 K2

(or)

12.b.Explain the features and characteristics of Box plots and Mean/SEM plots with suitable examples. CO2 K3

13.a. Consider a crime dataset with following fields (date, offense, address, latitude, longitude). Show how it can be represented using Dot Density Maps. CO3 K4

(or)

13.b. Discuss the usage of Dumbbell charts and slope graphs with suitable examples. CO3 K2

14.a. Justify the need for customization in graphs. Illustrate with examples the process of customizing axes, gridlines and colors in a graph. CO4 K5

(or)

14.b. Discuss the features and characteristics of Bubble charts and Radar charts with suitable examples. CO4 K2

15.a. Discuss some open-source packages that supports creation and deployment of Interactive Graphs on web pages. CO5 K2

(or)

15.b. Explain the options available for saving the graphs in RStudio. CO5 K3

Part C **5x12 = 60**

Answer the following

Answer should not exceed 800 words or four pages

16.a. Discuss the impact of placing data and mapping options in a ggplot. Show how graphs can be saved as objects and used as per the requirements. CO1 K2

(or)

16.b. Discuss the approaches used to deal with missing data and show how it is applied in ggplot2. Also illustrate the methods used for reshaping the data. CO1 K2

17.a. Consider a fuel dataset containing information such as manufacturer, model, engine displacement, year, transmission type, city mileage/litre, highway mileage/litre, fuel type. Represent this dataset using Pie chart and Tree Maps. CO2 K5

(or)

17.b. Use strip plot to represent salary distribution by rank in an academic environment. Discuss its pros and cons. Show how the usage of jitter will vary its outcomes. CO2 K5

18.a. Discuss with suitable dataset the usage of choropleth maps and area charts. CO3 K3

(or)

18.b. Illustrate the application of Correlation plots and Survival plots with suitable dataset. CO3 K3

19.a. Discuss the usage of scatterplot matrix, waterfall charts and word clouds with suitable examples. CO4 K2

(or)

19.b .Illustrate with an example the usage of Biplots, Heat Maps and scatter plots. CO4 K3

20.a. Discuss some open-source initiatives to perform customization and external editing of graphs. CO5 K3

(or)

20.b. Discuss the process involved in modifying various axes (quantitative, categorical, date) in a graph. Illustrate with an example. CO5 K3

****