**Student Outcomes**

Graduates of the Computer Engineering program will have an ability to:

1. Identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics.

2. Apply engineering design to produce solutions that meet specified needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors.

3. Communicate effectively with a range of audiences.

4. Recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts.

5. Function effectively on a team whose members together provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives.

6. Develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions.

7. Acquire and apply new knowledge as needed, using appropriate learning strategies.

**Course Outcomes:**

|  |  |
| --- | --- |
| CO-1 | Explain steps in data science lifecycle |
| CO-2 | Describe ETL process and its significance |
| CO-3 | Obtain, clean and transform data |
| CO-4 | Analyse and interpret data using ethical approach |

**Mapping of Course Outcomes with student Outcomes:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Mapping | SO-1 | SO-2 | SO-3 | SO-4 | SO-5 | SO-6 | SO-7 |
| CO-1 | M |  |  |  |  |  |  |
| CO-2 |  | H |  |  |  |  |  |
| CO-3 |  |  |  |  |  | H |  |
| CO-4 |  |  |  | H |  |  |  |

**Lab Experiments Details with mapping to Course Outcomes:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment No.** | **Title** | **Prerequisite\*** | **CO#** |
| 1 | Identify different Data collection techniques and tools | Software Programming | CO1 |
| 2. | Introduction to Pre-processing tool WEKA | Knowledge of handling CSV file | CO2 |
| 3 | SPSS |  | CO2 |
| 4. | Reading and writing different types of datasets  a. Reading different types of data sets (.txt, .csv) from Web and disk and writing in file in specific disk location.  b. Reading Excel data sheet in R | R environment | CO3 |
| 5 | DESCRIPTIVE STATISTICS IN R a. Write an R script to find basic descriptive statistics using summary, str, quartile function on mtcars& cars datasets. b. Write an R script to find subset of dataset by using subset (), aggregate () functions on iris dataset. |  | CO3 |
| 6 | CORRELATION AND COVARIANCE a. Find the correlation matrix. b. Plot the correlation plot on dataset and visualize giving an overview of relationships among data on iris data. c. Analysis of covariance: variance (ANOVA), if data have categorical variables on iris data. |  | CO3 |
| 7 | REGRESSION MODEL Import a data from web storage. Name the dataset and now do Logistic Regression to find out relation between variables that are affecting the admission of a student in a institute based on his or her GRE score, GPA obtained and rank of the student. Also check the model is fit or not. Require (foreign), require (MASS) |  | CO4 |
| 8 | CLASSIFICATION MODEL a. Install relevant package for classification. b. Choose classifier for classification problem. |  | CO4 |
| 9 | CLUSTERING MODEL a. Clustering algorithms for unsupervised classification. b. Plot the cluster data using R visualizations. |  | CO4 |
| 10 | Micro Project, using all the experiments. |  | CO4 |

**Experiment No.01**

PART A

(PART A: TO BE REFFERED BY STUDENTS)

**A.1 Aim:** Identify different Data collection techniques and tools

**A.2 Prerequisite:** Software Programming

**A.3 Outcome:**

**After successful completion of this experiment students will be able to**

1. Identify data collection method as per the requirement/research
2. Plan the data collection procedure
3. Collect the data
4. Store the data

**A.4 Theory:**

**Data collection** is a systematic process of gathering observations or measurements. Whether you are performing research for business, governmental or academic purposes, data collection allows you to gain first-hand knowledge and original insights into your [research problem](https://www.scribbr.com/research-process/research-problem/).

While methods and aims may differ between fields, the overall process of data collection remains largely the same. Before you begin collecting data, you need to consider:

* The aim of the research
* The type of data that you will collect
* The application used to collect the data

**A.5 Task to be completed in PART B**

**A.5.1. Task 1:**

**Every student needs to follow following steps and record the findings in appropriate section of PART B**

1. Identify any research topic
2. Identify the requirements of the research
3. Collect the data. (Real-time or already available online)
4. Store the data in Excel file. Perform sort, mean, median for the data collected.
5. Identify the type of data (Quantitative or Qualitative)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

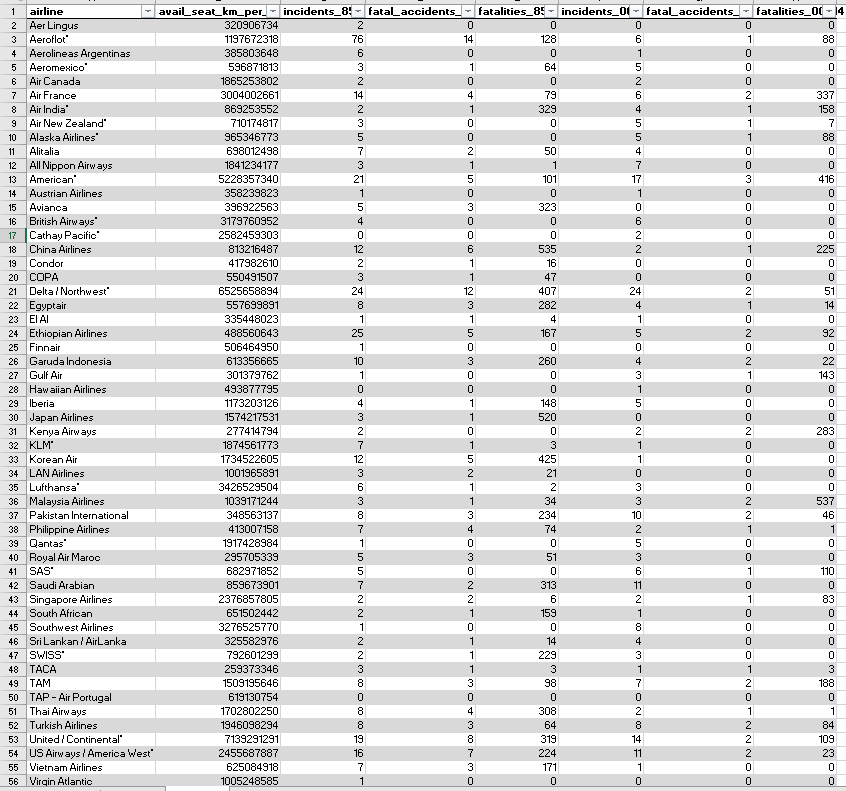
**PART B**

(PART B: TO BE COMPLETED BY STUDENTS)

**(Students must submit the soft copy as per following segments within two hours of the practical. The soft copy must be uploaded on the Blackboard or emailed to the concerned lab in charge faculties at the end of the practical in case the there is no Black board access available)**

|  |  |
| --- | --- |
| Roll No.N021 | Name:Saransh Singh Dhapola |
| Program :MBA.Tech(cs) | Division:D |
| Batch:B2 | Date of Experiment: 21-07-22 |
| Date of Submission: 21-07-22 | Grade : |

**B.1 Tasks given in PART A to be completed here**





**Part A answers:**

1. Identify any research topic  
   Ans- Airline safety.
2. Identify the requirements of the research  
   Ans- the requirements are: airline names, fatalities, seats available per km, incidents
3. Collect the data. (Real-time or already available online)   
   Ans-Pasted above, dataset was available online.

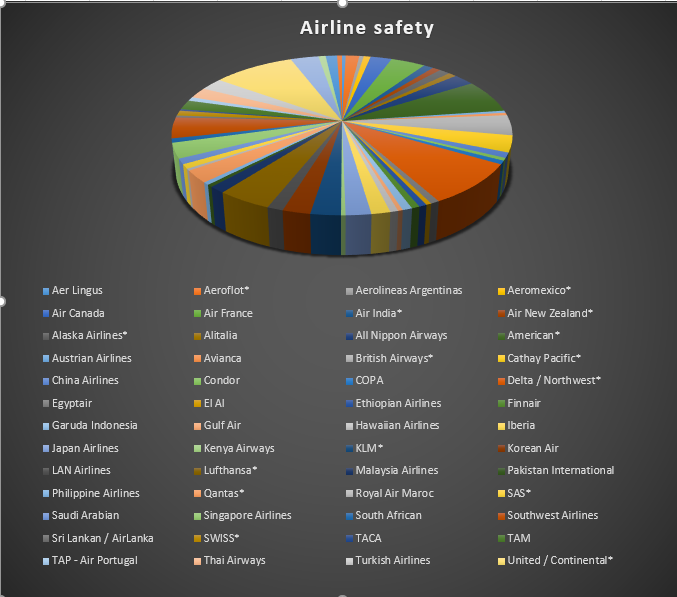
Store the data in Excel file. Perform sort, mean, median for the data collected.

Identify the type of data (Quantitative or Qualitative)

Ans- Data is Quantitative.

**B.2 Observations and Learning:**

*(****Students must write the observations and learning based on their understanding built about the subject matter and inferences drawn)***

Analysis of dataset done, airlines were rated based on their safety. Parameters included the number of incidents and fatalities. The average incidents were 7 per year and the average number of fatalities 112 per year. United/Continental was the safest airline.  


**B.3 Conclusion:**

*(****Students must write the conclusive statements as per the attainment of individual outcomes listed above and learning/observation noted in section B.2)***

Ans- Basic concept of database analysis and its importance acquired.

**B.4 Question of curiosity:**

1. Why data extraction is important?  
   Ans-Data extraction is important because it allows us to know the values pointing to a particular parameter which is helpful in analysis and predictions.
2. What is qualitative and quantitative data? Explain in terms of data collected in this experiment.  
   Ans- Quantitative data is data collected in form of numeric values, while qualitative data is data collected in from of “type”. It can be represented using characters or signs.
3. What is missing value, redundancy? From the data collected identify redundant data in excel file.   
   Ans- Missing values are the numeric values that are not defined for a certain parameter in a dataset. Redundancy is storing the data in multiple locations i.e. having different numeric values for the same record/variable at different positions.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*