VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI

A R PROGRAMMING MINI PROJECT REPORT ON

Exploratory Data Analysis and Customer Segmentation in a Mall: A Comprehensive R Programming Approach

IN

COMPUTER SCIENCE AND DESIGN

By

DHANYASHREE	4AL22CG016
KATHANIKA MOGERAYA	4AL22CG026
SHAYANASHREE N	4AL22CG053

UNDER THE GUIDANCE OF

Dr . Shivaprasad D J

ALVA'S

Education Foundation

DEPARTMENT OF COMPUTER SCIENCE AND DESIGN

ALVA'S INSTITUTE OF ENGINEERING AND TECHNOLOGY

MOODBIDRI-574225,KARNATAKA

2023-2024

ALVA'S INSTITUTE OF ENGINEERING AND TECHNOLOGY
MIJAR, MOODBIDRI, D.K. -574225

ALVA'S
Education Foundation

DEPARTMENT OF COMPUTER SCIENCE AND DESIGN

CERTIFICATE

This is to certify that the R Progrraming Mini Project entitled "Exploratory Data Analysis and Customer Segmentation in a Mall: A Comprehensive R Programming Approach" has been successfully completed by

DHANYASHREE	4AL22CG016
KATHANIKA MOGERAYA	4AL22CG026
SHAYANASHREE N	4AL22CG053

The bonafide students of Department of computer science and design , Alva's institute of Enginnering and technology in **DEPARTMENT OF COMPUTER SCIENCE AND DESIGN** of the **VISVESVARAYA TECHNOLOGICAL UNIVERSITY,**

BELAGAVI during the year 2023-2024. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report deposited in the departmental library. The Mini project report has been approved as it satisfies the academic requirements in respect of Mini Project work prescribed for the Bachelor of Engineering Degree.

Dr . Shivaprasad D J	**Mr . JayanthKumar A. Rathod**
Mini Project Guide	**HOD CSD**

EXTERNAL VIVA

Name of the Examiners	**Signature with Date**
1.	

ABSTARCT

This mini project employs the R programming language to conduct an in-depth analysis of a mall dataset, focusing on exploratory data analysis (EDA) and customer segmentation. The project begins with the extraction and preprocessing of a large dataset containing information on mall customers. The exploratory data analysis encompasses a thorough examination of customer demographics, annual income, and spending scores, utilizing statistical measures and visualizations.

The data visualization section employs various techniques, such as bar plots, pie charts, histograms, and box plots, to illustrate patterns and trends in customer gender distribution, age distribution, and annual income. Additionally, scatter plots and line graphs depict the relationship between annual income and spending scores.

Furthermore, the project delves into customer segmentation using the K-Means clustering algorithm, exploring different cluster solutions and evaluating their effectiveness. The results provide valuable insights into distinct customer segments based on their annual income and spending behaviors.

The mini project concludes with a comparative analysis of clustering techniques, emphasizing the practical implications of the findings. This comprehensive exploration showcases the power of R programming for extracting meaningful insights from large datasets and informing strategic decision-making in a mall setting.