

Data Analysis 1

COURSE INSTRUCTOR

(DR. Idrees Alsolbi)

SUBMITTED BY:

XXXX XXXXXX XX

XXXX

XXXX

XXXX

STUDENT ID:

42XXXXXXX

XXXX

XXXX

XXXX

XXXX

DEPARTMENT OF (XXXXX)

COLLEGE OF COMPUTING

UMM AL-QURA UNIVERSITY

2023

**TABLE OF CONTENT**

Table of Contents

[INTRODUCTION 2](#_Toc161007671)

[TASKS DESCRIPTIONS 2](#_Toc161007672)

[THE EXPERIENCES AND SKILLS ACQUIRED BY THE TEAM MEMBERS 5](#_Toc161007673)

[CONCLUSION 6](#_Toc161007674)

[REFERENCES 7](#_Toc161007675)

# INTRODUCTION

# TASKS DESCRIPTIONS

**TASK 1: Data Acquisition and Preparation (Week 3):**

* Acquire a dataset containing customer-related data from multiple sources, such as CRM systems, transaction records, and website logs.
* Preprocess the data by handling missing values, encoding categorical variables, and normalizing numerical features.

**TASK 2: Exploratory Data Analysis (EDA) (Week 4):**

* Perform exploratory data analysis to gain insights into the dataset's characteristics, distributions, and correlations.
* Visualize key features using histograms, scatter plots, and correlation matrices.

**TASK 3: Linear and Nonlinear Regression (Week 5):**

* Implement linear regression models to predict customer spending based on demographic and behavioral factors.
* Explore the use of nonlinear regression techniques, such as polynomial regression, to capture complex relationships in the data.

**TASK 4: Logistic Regression and Classification (Week 6):**

* Apply logistic regression to predict binary outcomes, such as customer churn or purchase likelihood.
* Utilize classification algorithms, such as decision trees or random forests, to categorize customers into segments based on their characteristics.

**TASK 5: Clustering (Week 7):**

* Employ clustering techniques, such as K-means or hierarchical clustering, to group customers with similar attributes or behaviors.
* Evaluate the effectiveness of different clustering algorithms using appropriate metrics.

**TASK 6: Anomaly Detection (Week 8):**

* Identify unusual patterns or outliers in customer behavior data using anomaly detection methods, such as isolation forests or density-based clustering.

**TASK 7: Principal Component Analysis (PCA) (Week 9):**

* Perform PCA to reduce the dimensionality of the dataset while preserving the most important features.
* Visualize the principal components and analyze their contributions to the variance in the data.

**TASK 8: Data Summarization and Visualization (Week 10):**

* Summarize key insights and findings from the analysis, including significant trends, correlations, and clusters.
* Create visualizations, such as heatmaps, scatter plots, and bar charts, to communicate the results effectively.

# THE EXPERIENCES AND SKILLS ACQUIRED BY THE TEAM MEMBERS

|  |  |  |
| --- | --- | --- |
| Student Name | Experiences | Skills |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

# CONCLUSION

# REFERENCES