Department of Computer Science

Summative Coursework Set Front Page

Module Title: Data Integration and Visualization
Module Code: CS3DV20
Lecturer responsible: Dr Nisha Singh
Type of Assignment: Written assignment
Individual / Group Assignment: Individual assignment
The weighting of the Assignment: 70% of the overall assessment
Report page limit: 6 pages (excluding the front page and a list of references)
Expected hours spent for this assignment: 14 hours
Items to be submitted online through Blackboard Learn by: 1/12/2023, 12:00 Noon
Work will be marked and returned 15 working days after the submission deadline.

NOTES

By submitting this work, you are certifying that you have read the assessment guidelines which are displayed at the top of the Assessment Folder on the Blackboard course for this module, and that you have conformed to the associated policies and practices, including those on:

- Submitting your own work, not that of other people or systems, and the associated penalties for Academic Misconduct.
- Submitting by the specified deadline, and the penalties associated with late submission (if allowed).
- The exceptional circumstances system (for applying for extensions).
- The use of a green sticker for students with relevant needs.

Assessment classifications

The table shows what is typically expected of the work to obtain a given mark.

Marks Criteria	
First Class (>= 70%)	Outstanding/excellent work with a complete solution, a good presentation of the workflows, and a critical analysis of the results. An outstanding work will present fully automated solutions based on advanced techniques. • deep & insightful analysis of the data, • helpful & precise comments, • excellent & and compelling presentation of the work.
Upper Second (60-69%)	Very good work with complete solution but partially correct results: most work has been carried out correctly. Discussions with superior quality. The presentation is good, well-structured, clear, and complete with respect to the work done.
Lower Second (50-59%)	Decent work but missing some significant part of the assignment, and/or with partially correct results. Some discussions but not in-depth ones. The presentation is accurate, complete, and clear with few technical errors, and written in plain English.
Third (40-49%)	Achievement of the minimum requirements. Some parts of the assignment are missing or incorrect. Little discussions. The presentation is, in general, accurate and complete, though it may lack some clarity and quality.
Pass (35-39%)	Partial solutions to a limited part of the assignment. Some tasks have not been carried out. Some results may not be complete or technically sound. The presentation is not accurate, or complete and lacks clarity.
Fail (<35%)	Incomplete and unsatisfactory solutions. The results contain major defects. The presentation is not accurate/complete and lacks clarity and quality.

Assignment specification

This assignment is expected to be carried out using Tableau - the Data Integration and visualization platform used in the module. In this assignment, you are required to explore the given dataset for further analysis and present visualizations of the results to highlight insights related to the business areas of concern.

Case Study: Integrate and Visualize E-Commerce Sales Dataset

This dataset provides an in-depth look at the profitability of e-commerce sales. It contains data on a variety of sales channels, including Shiprocket and INCREFF, as well as financial information on related expenses and profits. The columns contain data such as SKU codes, design numbers, stock levels, product categories, sizes, and colors. In addition to this we have included the MRPs across multiple stores like Ajio MRP, Amazon MRP, Amazon FBA MRP, Flipkart MRP, Limeroad MRP Myntra MRP, and PaytmMRP along with other key parameters like amount paid by the customer for the purchase, rate per piece for every individual transaction Also we have added transactional parameters like Date of sale months category fulfilled by B2b Status Qty Currency Gross amt.

Dataset Link: E-Commerce Sales Dataset (kaggle.com) -Amazon Sale Report data only.

Outline for the report:

Module Code:
Assignment report Title:
Student Enrollment Number:
Date (when the work is completed):
Actual hours spent on the assignment:

1. **Introduction** (5 Marks)
   - Introducing the problem domain where your assignment falls in.
   - For this given context and resources, briefly describe how your knowledge of data visualization might be helpful in extracting information from raw data.

2. **Analyzing the sales data by adopting Exploratory Data Analysis (EDA)** (25 Marks)
   - Using the **Amazon Sale Report.xls** datasheet, define a story (i.e., a context) for analyzing the dataset, and
   - Design a star schema for populating all the dimensions.

3. **Create a dashboard using Amazon Sale Report.xls** (25 Marks)
   Design a dashboard that visualizes:
   - Quantity of shipment of a particular category by month,
   - Quantity of sales by category in a particular month,
   - Amount of sales in category by month,
   - Top 10 states about the amount of sales in category for ship service level.

4. **Recommendation to improvement of sales performance** (20 marks)
   Analyze the dashboard and provide recommendations to the business areas of concern, e.g., how sales performance can be improved.

5. **Conclusion and Future Improvement** (5 Marks)
   - Concluding remarks about your findings and discussions in this report, and
   - Proposing future improvements.

Presentation of Report (20 Marks)

Presenting the report in a clear structure, providing references and their citations in the text, and using citations for figures and tables.

Report formatting requirements

The report is suggested to use the writing style and format as follows:

Discussions in the technical report

a) focusing on technical aspects
b) using technical terms
c) containing relevant content succinctly

Heading with label – e.g., 1. Introduction
Page number
Front: Calibri
Font size: 11
Line spacing: 1.5 lines
Use references where needed.

In-text citation for references:

… (Booch, et al., 1999).

In-text citation and caption for figures/tables:
Table 1: The comparison results … place it on top of the table.
Table in-text citation e.g., Table 1 presents … or … (Table 1)
Figure 1: The class model for … place it below the figure
Figure in-text citation e.g., Figure 1 shows … or … (Figure 1).