# Ngee Ann Polytechnic Rebrand: Child's Play

# School of InfoComm Technology

**Data Exploration & Analysis Assignment**

Diploma in DS

April 2024 Semester

**ASSIGNMENT 2**

(40% of DEA Module)

8th July 2024 – 11th August 2024

**Submission Deadline:**

**Presentation: Week 17 during class**

**Report and files: 11th August 2024 (Sunday), 11:59PM**

|  |  |  |
| --- | --- | --- |
| **Tutorial Group** | **:** | **T01 / T02 / T03** |
| **Student Name** | **:** |  |
| **Student Number** | **:** |  |

**Penalty for late submission:**

Marks will be deducted every calendar day after the deadline.

**NO** submission will be accepted after 18th August 2024 (Sunday), 11:59PM.

1. **Assignment Overview**

This assignment will assess your ability to apply the knowledge and skills acquired from the Data Exploration and Analysis (DEA) module. You will be provided with a detailed company profile, sales data, and related datasets. Your task is to utilize analytics techniques learned during the module to assist management in making informed sales and marketing decisions.

Based on the findings and trends discovered through your analysis, you will provide recommendations to improve business performance. A comprehensive report, detailing your analysis, findings, and recommendations, is required.

**2. Assignment Background**

**Company Profile**

Founded in 2007, Flipkart is a prominent e-commerce company in India, offering a diverse range of products including electronics, fashion, and home essentials. Headquartered in Bangalore, Flipkart has established itself as a household name, continually expanding its operations and adapting to the digital marketplace.

**Project Context**

In response to the rapid growth of e-commerce, Flipkart recognizes the necessity of leveraging data analytics to stay competitive. The CEO has initiated the formation of a Business Analytics team to provide business intelligence and support decision-making. The IT department is considering the implementation of an Enterprise Data Warehouse (EDW) for centralizing sales data. Currently, various departments have consolidated their data into different CSV files, with descriptions provided in DataDictionary.csv.

Your role as a Business Analyst is to Integrate these datasets, perform data modelling, and create dashboards to support managerial decision-making using the dataset provided.

**3. Assignment Tasks**

***Part 3.1: Creating dashboards to address business questions (70 Marks)***

You are to **create separate dashboards** for each of the following questions that the CEO of Flipkart are interested using the dataset provided. These dashboards will enable stakeholders to enhance the company's sales performance, product offerings, operations, and more:

1. **Which period (month) has the most sales?**  
   Analyse sales seasonality to optimize inventory management, staffing, and the timing of marketing campaigns.
2. **What are the top 5 most popular products (by Category & another one by Sub-category) across different sales channels?**  
   Identify key products to effectively target marketing activities and engage customers.
3. **What are the top and bottom 5 locations in terms of profit?**  
   Evaluate branch performance to make informed decisions, optimize resources, and enhance overall profitability.
4. **Propose and address an additional business question with rationale.**  
   Explain why this question is critical for business improvement.

**Dashboard Design**

Design a dashboard using Power BI Desktop that highlights critical Key Performance Indicators (KPIs) to answer the above questions. Ensure the application of dashboard best practices in your design.

**Report Section**

Write a report section (maximum 2,000 words with supporting visualization) detailing your approach in completing the tasks in Part 3.1. Include the rationale behind your choices, and clearly state any additional DAX measures, columns, and their corresponding formulas used.

***Part 3.2: Further Analytics (30 marks)***

Perform advanced analytics using techniques such as univariate, bivariate, and multivariate analysis, Pivot Tables, and Data Mining. Identify any critical trends or patterns to help Flipkart enhance its business performance.

For example, analyse if a particular product shows significant clustering towards a specific geographical locality or customer purchasing mode.

**Report Section**

Write a report section (maximum 2,000 words) detailing your approach in Part 2. Elaborate on the rationale behind your choices and how the analysis provides deeper insights into the business.

|  |  |  |
| --- | --- | --- |
|  | **Suggested Report Sections & Content Guidelines** | **Word Count** |
| 1. | Table of Contents | NA |
| 2. | Abstract/Overview | Max: 500 words |
| 3. | Creating dashboard to answer business question (for Part 3.1) | Min: 1000 words  Max: 2000 words |
| 4. | Further analysis (for Part 3.2) | Max: 2000 words |
| 5. | Summary   * Summarize your findings * Reflections on Data modeling, Data exploration and Analysis process | Min: 500 words  Max: 1000 words |

1. **Assignment Deliverables**

**Files submission**

Submit all required solution files in the assignment submission folder in POLITEMall. Ensure that all files are organized into a single folder, which should then be zipped and named as [student\_name]\_[student\_ID].zip.

**Face-to-Face Presentation**

Present your work face-to-face in Week 17. The presentation will be assessed based on clarity, organization, and understanding of the work done. Each presentation should last no more than 5 minutes. Your presentation will be based on the dashboard constructed, report (and excel/other tools you used to create visualizations for your report).

**If you are unable to attend the scheduled class on Week 17, you must make an appointment with your lecturer for an alternative presentation time.**

**Note: DO NOT PLAGIARIZE (https://www1.np.edu.sg/clte/antiplagiarism/policy.htm for more information)**

1. **Assessment Rubrics**

Please refer to Appendix B for the Assessment Criteria.

**Please refer to the appendices following this section.**

**Appendix A (Datasets)**

**Raw Files**

**1. List of Orders**

This dataset contains purchase information, including ID, date of purchase, and customer details.

**Columns:**

* **Order ID:** ID of the purchase order
* **Order Date:** Date of purchase
* **Customer Name:** Name of the customer
* **State:** State of the customer
* **City:** City of the customer

**2. Order Details**

This dataset contains order details including the order ID, price, quantity, profit, category, and subcategory of the product.

**Columns:**

* **Order ID:** ID of the purchase order
* **Amount:** Price of the order
* **Profit:** Profit made by the purchase
* **Quantity:** Quantity of the purchase
* **Category:** Category of the product
* **Sub-Category:** Sub-category the product belongs to

**3. Sales Target**

This dataset contains sales target amounts and dates for each product category.

**Columns:**

* **Month of Order Date:** Month of the order
* **Category:** Category of the product
* **Target:** Sales target of the category

**Appendix B (Assessment Rubrics)**

|  |  |  |
| --- | --- | --- |
| Section | Assessment Criteria | Marks |
| 1 | Dashboard Quality  Assessed based on:   * + Quality of visualization design   + Use of appropriate charts/table   + Dashboard and visualization interactivity (such as filters, slices, drill up/down)   + Clear concise visualization goals with the use of appropriate data for visualization | 30 |
| Report Quality  Assessed based on:   * + Clear and concise business definition.   + Data modelling and preparation process   + Data analysis   + Insights and Conclusion | 20 |
| Presentation  Assessed based on:   * + Clear, concise, and well-organized presentation   + Clear understanding of work done (thought process, settings, whether the work fulfils the requirements) | 20 |
| 2 | Report  Assessed based on:   * + Clear and concise problem statement and approaches used in analysis.   + Clear understanding of univariate, bivariate and multivariate analysis.   + Clear understanding in time-series analysis and data mining techniques   + Clear description of insights and conclusion | 20 |
| Presentation  Assessed based on:   * + Clear, concise, and well-organized presentation   + Clear understanding of work done (thought process, settings, whether the work fulfils the requirements) | 10 |

**End of Document**