

**Data Science Fundamentals**

Year 1 (2023/24), Semester 1

## SCHOOL OF INFOCOMM TECHNOLOGY

Diploma in Cyber Security & Forensics

Diploma in Data Science

Diploma in Immersive Media

Diploma in Information Technology

Common ICT Programme

**ASSIGNMENT**

|  |  |
| --- | --- |
| **Weightage:** | 50% |
| **Individual/Team/Both:** | Both |
| **Format:** | Font Arial  Font Size 11, 1.5 spacing |
| **Deadline:** | **Report:**  **04 Aug 2023, 2359hrs**  Softcopy via Brightspace Assessment  Hardcopy – to be submitted in class during Week 17  **BI Files:**  **04 Aug 2023, 2359hrs**  Softcopy via Brightspace Assessment  **Presentation:**  **04 Aug 2023, 2359hrs**  Softcopy via Brightspace Assessment |

**Penalty for late submission**:

10 marks per day (including Sunday and public holiday).

No report will be accepted after **11 Aug 2023, 2359hrs**.

**Assignment Deliverables**

**Section A: Report and Power BI Dashboard (80 marks)**

Your report should include visuals and diagrams where appropriate to support and enhance their readability. The structure for your report is as follows:

1. **Stakeholder Background (Individual: 10 marks)**

**Purchasing Manager**

**Purchasing Manager refers to the head of a team that oversees the purchase and procurement of products or items that are to be resold or used in manufacturing.**

**Q1. How can purchasing managers use the data of sales orders to evaluate and decide on which items to purchase?**

**Q2. How can purchasing managers make sure that there are no surplus or shortages of items?  
Q3. How can purchasing managers make sure that we are making profit on items instead of losing money?**

**Q4. How can purchasing managers make sure items what to purchase based on states.  
Q5. How can purchasing managers make sure which items to have more sizes in to accommodate to allow more sales as well consider the colour of products.**

1. **Actionable Statements (Individual: 5 marks)**

**As a purchasing manager, we want to make sure that items that are procured or purchased are in good quality as well as in trend so that the company can profit. We also want to make sure that all our products can accommodate all types of customers to have a wider spread of them. We also must make sure that we purchase items at an optimal amount so we can have enough stock to sell but also don’t have a surplus to keep.**

1. **Data Analysis and Visualisations (Individual: 25 marks)**

**Figure I)**

**A screenshot of a computer

Description automatically generated**This figure is to allow the purchasing manager to investigate and analyse the profit, revenue and cost of items which can be sorted into states and categories. This solves the exploratory questions 1 and 4 because they can analyse the profit of specific items with the slicers provided to see which items are making more profit then others and so allow them to make more purchases towards that certain product for that state.

**Figure II)**

**A colorful circle with text

Description automatically generated**

This figure is a drill through of Figure I, it allows the manager to go in depth on what colour the product make. As shown on the image provided, the main category is Jacket, and it shows the profit of all the states and the colours. This helps with exploratory question 5, the jackets can be sorted into their item type, colour and states which allows the manager to see which colour is more wanted then others in that state so as to acquire more of that item to sell.

**Figure III)**

A graph with a curved line

Description automatically generated

This is a drill through of figure I where it shows profit by item size, category, and states. This will help exploratory question 5 because the manager can sort the states to view what sizes are making a profit and can decide to expand upon that size for other items.

**Figure IV)**

**A screenshot of a graph

Description automatically generated**

This figure shows the profit by quantity and Revenue vs Profit by quantity and name of that type. It allows the manager to analyse and view what items and types are making above average with the red line provided as the average profit of all items as well as the quantity of items.

This chart will solve exploratory questions 2 and 3 by allowing the manger to view which items are earning more then others even if their quantity is on the lower amount. They can also sort it by the name of the item or the category so if an item is making a profit, but the quantity is lower than others, they could purchase more of that item and upsell.

1. **Dashboard (Individual: 15 marks)**

**Figure IV)**

A screenshot of a graph

Description automatically generated

This dashboard collates the information needed to view the profit and revenue of the types of items the company sells. For the first image starting from the left, it compares the profit revenue and cost of types of items, with this information it allows the manager to make decision on what items to buy if it is making more then what it cost to purchase them. In the next figure, it shows the revenue vs cost per unit based on the quantity and type, it allows the manager to view the amount earned compared to purchasing of that quantity, for the figure shown if the manager is to purchase 80 of all types, they would make a profit. Lastly, the figure below shows the in-depth view of the profit based on quantity of item names including an average line on profit made. It will allow the manage to view which items are making more than average or less than average based on the quantity set which is 80 in the figure. So, the manager can conclude that if they made 80 purchases for ‘Denim’ product, they will make more then average compared to others.

1. **Individual Reflection (10 marks)**

**DSF has taught me the foundation of Power BI which can help me in collating or creating charts for future use in a work environment. It also taught me the importance of having a convenient way of viewing certain type of data via bar charts and pie chart as there can be a lot of data especially in large corporations such data companies or product companies such as “Coca Cola” or “Apple”. Having a way to view and analyse large amount of customer or information data can help companies to view trends and changes within adjustable ranges that can allow them to make informed decisions on their marketing, targeted audience, and customer segmentation.**

**With the word documents provided by Power BI, I learned about the massive amount of functionality provided within the software. For example, within a few clicks, it can create calculations, transform data, and merge datasets, and can also assist users who may not be tech-savvy in creating charts with their datasets. These functions are beneficial for users who need to manage and analyse big data without knowledge in computer expertise.**

**During the hands-on experience in DSF, one feature that I learnt is that Power Bi can connect multi-sources of data. Whether it’s a databases, spreadsheet, online services, or cloud-based platforms. It has seamless integration, removing the need of manually inputting and extracting data. Moreover, DSF has also thought me how to create relationships between tables and creating hierarchies.**

**I felt that DSF has helped me gain more appreciation for applications such as Power BI, it allows the use of sharing data and viewing data more convenient for users and companies as well as help them get more insight into their data. It helps people view their data not just in numbers but visually and to be honest some people can’t handle numbers so by using Power BI to showcase the data in a more picturesque way even old people could understand.**

**During my study in DSF, our teacher ‘Mr Chee’ has really helped with the understanding and use of Power BI, he provides explanations that are easy to absorb and goes in depth on how Power BI can help us right now and, in the future, as well. He also provides us with his own examples which makes the lesson enjoyable, entertaining, and more fun to learn without his help Power BI might just be another boring module but with him teaching its more enjoyable.**

**To summarize, DSF lessons in Power BI and the multiple types of data and functions, has taught me the necessary skills to proficiently analyse and visualize data. With its user-friendly interface as well as wide arrays of functionality and features. Power BI has become a useful software that can be used in the future for my work life and learning experience in other modules that may require data visualizing or data inspection. With this knowledge I’ll be able to analyse and work with data that may be provided to me to visualize and present.**

1. **Group Reflection (10 marks)**

As a group, share 5 challenges, problems encountered or suggest areas of improvement while your team is doing the analysis.

**Section B: Dashboard Storytelling Presentation (Group: 20 marks)**

Each group will share your analysis using dashboards and visualisations in a 20 minutes formal **dashboard storytelling presentation** including Q&A. Your slides should cover the following:

1. An introduction
2. Individual stakeholder to do storytelling presentation for any 2 visualisations from your dashboard
3. Conclusion

**Submission Guideline**

1. Naming convention
   * Power BI file – “Your Full Name DSF - Analysis.pbix”
   * Report in MS Word – “Leader’s Full Name DSF – Assignment.docx”
   * Dashboard Storytelling Presentation in MS PowerPoint – “Leader’s Full Name DSF – Assignment Storytelling Presentation.pptx”
2. All submissions via Brightspace
3. **Leader** to compile the following:
4. One Group Report in MS Word consists of following:
   1. Cover Page
   2. Table of Content
   3. Introduction
   4. Content

e.g.

Member 1 Full Name

* Stakeholder Background
* Actionable Statements
* Data Analysis and Visualisatisations
* Dashboard
* Individual Reflection

Member 2 Full Name

* Stakeholder Background
* Actionable Statements
* Data Analysis and Visualisatisations
* Dashboard
* Individual Reflection

….

Member 5 Full Name

* Stakeholder Background
* Actionable Statements
* Data Analysis and Visualisatisations
* Dashboard
* Individual Reflection
  1. Group Reflection

1. One Presentation Slides in MS PowerPoint
   * Follow the requirements in Section B: Dashboard Storytelling Presentation
2. **Member** including leader to submit individual Power BI file.

**Marking Criteria**

|  |  |  |
| --- | --- | --- |
| Component | Grade | Criteria |
| * Report * Visualisation * Dashboard Storytelling Presentation | A+ | * In-depth analysis coupled with a good list of exploratory questions tailored for targeted stakeholder. * Excellent mix of visualisations with accurate interpretation of all visualisations and dashboards. * Content is well-organised in a logical structure with excellent use of vocabulary. * Reflection moves beyond simple description of experience and learning. * Excellent dashboard storytelling presentation flow with detailed sharing of visualisations and dashboard that are highly relevant to targeted stakeholder. |
| A – B | * Good level of analysis with a suitable list of exploratory questions for targeted stakeholder. * Good use of visualisations with adequate consistency to build the dashboards supported by appropriate interpretation. * Content is fairly organised with appropriate use of vocabulary. * Reflection shows attempt to analyse experience and learning. * Good dashboard storytelling presentation with adequate sharing of visualisation and dashboard that provides relevant insights to targeted stakeholder. |
| C – D | * Minimal analysis and listing of exploratory questions and poor justification of targeted stakeholder. * Content is somewhat organised with some use of vocabulary. * Minimal variety of visualisations with poor or wrong interpretations and mediocre design of dashboards. * Reflection does not move beyond description of learning experiences. * Brief dashboard storytelling presentation showing little understanding of stakeholder’s requirements and lack of suitable visualisations. |
| F | * Poor or wrong analysis for targeted stakeholder. * Report is skimpy with unclear content, showing no attempt in consideration of target stakeholder. * Minimal or largely wrong visualisations. * Minimal effort made on reflection. * Skimpy dashboard storytelling presentation with no clear flow and difficult to understand. |

**Definition of Attribute/Field**

1. **Customer File**

* **Customer\_Code** is a unique number which identifies a particular customer
* **Customer\_Age** refers to customer’s age as of 2021

1. **Item File**

* **Item\_Code** is a unique number to identify each item
* **Item\_Selling\_Price** is the standard selling price of the item which the company sells it to customers. The standard selling price may be different from Order Details File 🡪 Selling\_Price\_Per\_Unit if the company offered seasonal promotions, special discounts or further marked up the standard selling price to customers.
* **On\_Hand\_Quantity** refers to the balance of stocks as of 20 Nov 2021

1. **Order File**

* **Order\_No** is a unique number to identify each sales order from one customer
* **Customer\_Code** refers to a particular customer who placed an order (i.e. Order\_No) to the company

1. **Order Details File** (Keep track of the order details for each sales order placed by a customer)

* **Sales\_No** is a system generated number that uniquely identify each row
* **Order\_No** is an order placed by a customer

Example:

Customer\_Code = 64 has placed a sales order with Order\_No = 1

The Order\_No = 1 consists of 6 products in the Order Details file (Sales\_No = 0 to 5)

In this scenario, the Customer “64” has made an order with order\_id “1” for 6 products

|  |  |  |  |
| --- | --- | --- | --- |
| **Order File** | | | |
| **Order\_No** | **Customer\_Code** | **Order\_Date** | **Order\_Delivery\_Date** |
| 1 | 64 | 30/08/2021 | 24/09/2021 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Order Details File** | | | | | | |
| **Sales\_No** | **Order\_No** | **Item\_Code** | **Selling\_Price\_Per\_Unit** | **Cost\_Per\_Unit** | **Order\_Quantity** | **Total\_Amount** |
| 0 | 1 | 218 | 106 | 77 | 2 | 212 |
| 1 | 1 | 481 | 118 | 95 | 1 | 118 |
| 2 | 1 | 2 | 96 | 50 | 3 | 288 |
| 3 | 1 | 1002 | 106 | 77 | 2 | 212 |
| 4 | 1 | 691 | 113 | 51 | 3 | 339 |
| 5 | 1 | 981 | 106 | 62 | 3 | 318 |

* **Selling\_Price\_Per\_Unit** refers to the actual selling price that the company transacted including any discount given to customer
* **Cost\_Per\_Unit** refers to the purchase price per unit that the company paid to supplier to bring in the product
* **Order\_Quantity** refers to the order quantity that placed by the customer

**References:**

1. Refer to the training materials from <https://learn.microsoft.com/>.
2. You may refer to the supplementary materials to help you in this assignment:

* Creating Measures Using DAX In Power BI
* Optimize Model Performance

**Plagiarism and Copyright Issues**

Ngee Ann Polytechnic (NP) develops students to uphold good academic practices, academic integrity and respect for the works of others. These practices are guided by the NP Plagiarism Policy which mandates that students assume full responsibility for the content and integrity of the academic work they submit.

Plagiarism means, “copying any part of a source, and then submitting it, claiming that it is your own work.”

Please ensure that all the works submitted by you are not copied from other sources. Any attempt to plagiarize will be dealt with severely, and it may result in your failing the module. If you have made any references to certain materials, make sure you cite the sources by acknowledging and providing the information necessary to find the source (e.g. Title and author of book, Internet links, etc.)

If a student is found to have submitted work not done by him/her, he/she will not be awarded any marks for this assignment. Disciplinary action may also be taken.

Similar action will be taken for the student who allows other student(s) to copy his/her work.

Please refer to the following URL for more details:

<http://www.np.edu.sg/tlc/antiplagiarism/policy.htm>