SLT-Mobitel Nebula Institute of Technology
Pearson BTEC Level 5 Higher National Diploma in Digital Technologies (RQF)

Assignment Brief

Unit Number and Title	**Unit 5 - Big Data and Visualization**
Academic Year	2025/2026
Unit Tutor	Dr. Tilani Gunawardena Miss. Sakunika Adhikari
Assignment Title	**Assignment 1**
IV Name and Date	Mr. Dilan Sooriyamudalige
Issue Date	2025/08/06
Submission Date	➢ Portfolio of fact sheet: 2025/08/12 ➢ Technical Report: 2025/08/12 ➢ Power BI file: 2025/08/12 ➢ Python Code files : 2025/08/12 ➢ Dataset : 2025/08/12 ➢ PPT: 2025/08/12

Submission Format

- The written technical report must be submitted in a PDF format.
- The PowerBI file should be uploaded to the cloud, and a link to the file must be provided in a notepad document.
- The Python code should be well-documented and submitted either as .py files or .ipynb files
- The datasets used for the tasks should be included with the submission. If the dataset is too large, please upload it to the cloud and include the link in the notepad.
- Portfolio of Fact Sheet: Create a portfolio of fact sheet that explore the core principles of big data. The fact sheet should be concise, summarizing and highlighting the key concepts, tools, and techniques in big data analytics and visualization.
- Formal Presentation: Prepare a 10–15 minute presentation (around 10–15 slides, with supporting speaker notes), demonstrating the application of big data tools and

the roles of data specialists in analyzing a provided dataset. The presentation should include:

- 10–15 slides in PowerPoint format.
- Video evidence of the presentation. Upload the video to your YouTube channel, Dropbox, or Google Drive with public access. Include the link to the video on the final slide of your presentation.

You are required to make use of headings, paragraphs and sub-sections as appropriate, and all work must be supported with research and referenced using the Harvard referencing system (or an alternative system). You will also need to provide a bibliography using the Harvard referencing system (or an alternative system). Inaccurate use of referencing may lead to issues of plagiarism if not applied correctly.

Referencing your content is mandatory.

Unit Learning Outcomes

LO1 Examine big data and visualisation for decision making

LO2 Investigate statistical and graphical techniques, tools and industry software solutions for big data and visualisation

LO3 Demonstrate the use of industry software to manipulate data and prepare visual presentations for a given data set

LO4 Assess the role, responsibilities and challenges for data specialists

Transferable skills and competencies developed

- Understanding of the scientific method and its applications to problem-solving
- Demonstrate knowledge and understanding of essential facts, concepts, principles and theories
- Recognise the professional, economic, social, environmental, moral and ethical issues involved in the sustainable exploitation of computer technology
- Recognise and analyse criteria and specifications appropriate to specific problems
- Specify, design and construct reliable, secure and usable computer-based systems
- Evaluation of systems in terms of quality attributes
- Plan and manage projects
- Recognition of any risks and safety aspects that may be involved in the deployment of computing systems
- Deploy the tools used for the construction and documentation of computer applications
- Critical evaluation and analysis of complex problems
- Intellectual skills
- Self-management
- Reflection and communication
- Contextual awareness

- Sustainability

Vocational scenario

BigCloud Pvt. Ltd. is a highly reputable big data analytics company that has provided data-driven insights for leading organizations across Asia. BigCloud specializes in processing and visualizing data from client-provided datasets or those selected by BigCloud itself. Clients typically approach BigCloud with requests regarding how to leverage and interpret the information within large datasets to enhance their business opportunities.

As a recent graduate from Nebula Institute of Technology, you have applied for the position of Data Analyst and are expected to demonstrate your capabilities in handling data for effective decision-making. BigCloud expects you to highlight your expertise in areas such as data preparation, data manipulation, data visualization, and decision-making based on your findings.

Additionally, you will be required to demonstrate your understanding of the roles, responsibilities, and challenges faced by data specialists, along with strategies to ensure data compliance.

To complete the task, You must find and download a real datasets suitable. Recommended sources:

- UCI Machine Learning Repository
- Kaggle Datasets (free login required)
- Data.gov.lk (for Sri Lankan datasets)
- World Bank Data
- OpenML

Make sure that the datasets that you have selected:

- Have at least 500 records.
- Include a target attribute (class label).-optional
- Include a mix of numerical and categorical features.
- Include missing values

Assignment activity and guidance

Task 01 – Factsheet Portfolio

You are required to create a portfolio of fact sheets aimed at a semi-technical audience. These fact sheets will explore how big data and visualization can be leveraged for decision-making. Your fact sheets should cover a range of statistical and graphical tools, as well as software solutions used in the field of big data and visualization.

Although the portfolio targets a semi-technical audience, you are expected to incorporate the appropriate technical terminology where necessary.

Your fact sheet portfolio should include the following:

- An explanation of the fundamental concepts of big data.
- An investigation into the value of data for decision-making, from the perspective of both end users and the organization itself.
- An analysis of the advantages and challenges of data-driven decision-making in the context of the selected dataset.
- A final evaluation of the potential impact of using data for decision-making on both users and organizations.

Additionally, your fact sheets should highlight your investigation into the range of tools available for data analytics.

Your fact sheets should include:

- A description of the statistical and graphical techniques used in industry for big data analysis and visualization.
- A review of various industry-leading tools and software solutions available for data analysis and visualization.
- A comparison of how industry-leading tools and software solutions are utilized for data analysis and visualization with real life examples.

Task 02 – Technical Report

Following the completion of your factsheet portfolio, you are required to prepare a comprehensive technical report. Select one or more datasets from the recommended data sources, and develop **at least five analytical questions** based on the dataset(s). Use your knowledge of **statistical and graphical techniques** to explore and answer these questions.

Your technical report must include the following components:

I. Data Impact and Decision-Making

Evaluate the potential impact of data on both users and organizations in the decision-making process. Discuss how data manipulation and visualization can drive strategic business decisions and how this aligns with the needs of various stakeholders. Reflect on the dataset you selected and describe the specific tools and statistical techniques used to prepare, process, and visualize the data.

II. Industry Practices in Big Data Analysis

Explain the statistical and graphical techniques commonly used in the industry for analyzing and visualizing big data. Provide a comparative review of leading tools such as Power BI, Tableau, and Python for data analysis and visualization. Highlight the strengths and limitations of each tool, and choose one to demonstrate its application on your selected dataset. Show how it can be used to summarize, group, and analyze data through queries.

III. Visualization and Stakeholder Needs

Select an industry-leading tool or software solution to manipulate your chosen dataset and demonstrate the use of queries to effectively summarize and group the data. Create a visual presentation that highlights the key insights derived from your analysis, using at least five different chart visualizations to clearly communicate patterns, trends, and comparisons. Justify your choice of statistical methods and visualization tools by explaining how they address stakeholder requirements and illustrate the advantages of data analytics for organizations. Evaluate how your data preparation and manipulation process supports strategic decision-making and aligns with overall business objectives.

IV. Ethics, Compliance, and the Role of Data Specialists

Explain the different roles, responsibilities, and challenges faced by data specialists when working with data in an organizational context. Analyze how data professionals contribute to building an ethical, data-driven culture while navigating the complexities of modern data ecosystems. Review the strategies used by data specialists to ensure legal and regulatory compliance, including practices aligned with frameworks such as GDPR. Evaluate the challenges encountered by them when embedding ethics into data-driven decision-making processes

Upon completing your report, summarize your key findings in a **presentation** that effectively communicates your insights, methods, and conclusions.

Recommended Resources

Please note that the resources listed are examples for you to use as a starting point in your research – the list is not definitive.

Textbooks

Dietel, P. (2020) Intro to Python for Computer Science and Data Science: Learning to Program with AI, Big Data and The Cloud. London: Pearson.

Franks, B. (2020) 97 Things About Ethics Everyone in Data Science Should Know. USA: O'Reilly Media.

Freeman, M., Ross, J. (2019) Data Science Foundations Tools and Techniques: Core Skills for Quantitative Analysis with R and Git. London: Addison-Wesley Professional.

Graesser, L. and Keng, W. L. (2020) Foundations of Deep Reinforcement Learning: Theory and Practice in Python. London: Addison-Wesley Professional.

Kirk, A. (2019) Data Visualisation: A Handbook for Data Driven Design. London: Sage Publications.
Knaflic, C. N. (2015) Storytelling with Data: A Data Visualization Guide for Business Professionals. USA: John Wiley & Sons.

Loukides, M., Mason, H. and Patil, D. J. (2018) Ethics in Health Data Science. USA: O'Reilly Media.

Marr, B. (2017) Data Strategy: How to Profit from a World of Big Data, Analytics and the Internet of Things. London: Kogan Page.

McCormick, K. and Salcedo, J. (2017) SPSS Statistics for Data Analysis and Visualization. USA: John Wiley & Sons.

Viescas, J. L. (2018) SQL Queries for Mere Mortals: A Hands-On Guide to Data Manipulation in SQL. 4th Edition. London: Addison-Wesley Professional.

Wilke, C. O. (2019) Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures. USA: O'Reilly Media

Journals
Big Data & Society, SAGE Journals
International Journal of Computer Applications (IJCA)
Journal of Big Data, SpringerOpen
Journal of Data Science, Statistics and Visualisation, International Association for Statistical Computing (IASC)

Websites
gov.uk UK government (Data Ethics Framework)
ukdataservice.ac.uk UK Data Service(General reference)

Pearson BTEC

SLTMOBITEL NEBULA INSTITUTE OF TECHNOLOGY

Learning Outcomes and Assessment Criteria

Learning Outcomes and Assessment Criteria		
Pass	Merit	Distinction
LO1 Examine big data and visualization for decision making		**D1.** Evaluate the potential impact of data on both users and organizations when using data for decision-making.
P1. Explain the fundamental concepts of Big data. **P2** Investigate the value of data for decision making to both end users and organizations.	**M1.** Analyze the advantages and challenges of data-driven decision making to an organization.	
LO2 Investigate statistical and graphical techniques, tools and industry software solutions for big data and visualization		**LO2 and LO3** **D2** Evaluate own data preparation and manipulation, justifying your choice of statistical techniques, to show how this meets the needs of stakeholders for a given data set
P3. Describe statistical and graphical techniques for big data and visualization used in industry. **P4** Review different industry-leading tools and software solutions available for analyzing and visualizing data	**M2.** Compare how different industry-leading tools and software solutions are used to analyses and visualize data, with examples	
LO3 Demonstrate the use of industry software to manipulate data and prepare visual presentations for a given data set		
P5. Select an industry-leading tool and software solution to manipulate data for a given data set. **P6** Demonstrate the use of queries to summarize and group data for a given data set.	**M3.** Prepare a visual presentation to summarize data for a given data set	
LO4 Assess the role, responsibilities and challenges for data specialists		**D 3** Analyze the role, responsibilities and challenges faced by data specialists when building ethics into a data-driven culture.
P7 Explain the different roles, responsibilities and challenges faced by data specialists.	**M4.** Review the different strategies used by data specialists to ensure data compliance.	