

|  |  |
| --- | --- |
| **Student Name/ID Number:** | Muhammad Kemal / BDSE07-0922-084 |
| **Academic Year:** | 2023 |
| **Unit Assessor:** | Ei Thandar Khaing |
| **Project Title:** | Data Science and Business Intelligence |
| **Issue Date:** | November 4, 2023 |
| **Submission Date:** |  |

|  |
| --- |
| **Learner declaration** |
| I certify that the work submitted for this assignment is my own and research sources are fully acknowledged.  Student signature:  Date: |

# Task 1

**Data Science tools**

* **Apache Spark:** Apache Spark is an open-source distributed processing system used for big data workloads. It leverages in-memory caching and optimized query execution for fast analytic queries against data of all sizes. It provides development APIs in Java, Scala, Python, and R, and supports code reuse across many workloads, namely batch processing, interactive queries, real-time analytics, machine learning, and graph processing.
* **Jupyter Notebooks:** Jupyter notebook is a free web application used by data scientists to create and share documents containing code, calculations, analysis, visualizations, and other data processing.
* **Python:** Python is a computer programming language commonly used to build websites, software/applications, automate tasks and perform data analysis.

**BI Tools**

* **Power BI :** this is a business analytics tool that offers business intelligence features and interactive visualizations. Its user interface is designed to be straightforward enough for end users to generate reports and dashboards on their own.
* **Tableau**: Users can connect, visualize, and share data in an understandable manner with this data visualization application.
* **Microstrategy:** This business intelligence program offers strong data visualization and analytics features.

# Task 2

**Data Science tools**

* **Apache Spark** : Apache Spark: that use this tool such as Amazon, EBay, NASA JPL, dan Yahoo, Pinterest.
* **Jupyter Notebooks:** Microsoft, Google, NASA, Netflix, Spotify, Facebook are using this tool.
* **Python:** using this tool are YouTube, Google Maps, Instagram, Spotify, etc.

**BI Tools**

* **BI Tools:** Power BI: Customers that utilize this product include Apple, Walmart, and Exxon Mobil.
* **Tableau:** The US Department of Defense, Walmart, and CVS Health Corporation are a few of the clients that make use of this product.
* **Microstrategy:** Exxon Mobil, Walmart, and the US Department of Defense are a few of the clients that utilize this application.

# Task 3

**Data Science tools**

* **Apache Spark** : Apache Spark is a potent data processing platform that facilitates the effective management of massive data analysis. Spark is used by Pinterest for many different things, such as analyzing user behavior, proposing tailored content, and exploring data. Pinterest is able to extract relevant insights from very huge data sets thanks to the tool's quick and distributed computing capabilities, which enhances the platform's user experience overall. The platform dynamically adjusts to user preferences by utilizing Spark's real-time analytics, guaranteeing timely and pertinent content recommendations.
* **Jupyter Notebooks** : An open-source web tool called Jupyter Notebooks makes collaborative and interactive data science projects possible. With it, users may produce papers with live code, explanation text, and visualizations. Jupyter Notebooks are a flexible tool for teams working on a variety of data-driven projects since they support many programming languages and are widely used for data exploration, analysis, and insight sharing.
* **Python** : Python is used extensively by Google's internal applications in web development, machine learning, data analysis, and other domains. Because of Python's libraries and adaptability, Google is able to develop complex algorithms, analyze user behavior, and enhance a variety of services and goods.

**BI Tools**

* **Power BI:** One customer, Walmart, tracks sales and inventory in real time across all its stores using Power BI. The tool's interactive dashboards and reports help Walmart make informed decisions about product promotions, stock levels, and overall store performance.
* **Tableau:** One of Tableau's clients is the US Department of Defense, which uses it for defense-related data processing and visualization. Military analysts can gain insights into complex data sets with Tableau's powerful visualization capabilities, which aids strategic decision-making in the areas of defense and national security.
* **Microstrategy:** To achieve powerful data analysis and visualization, Exxon Mobil is one of the clients using Microstrategy. Exxon Mobil uses this application to evaluate complex operational data, optimize processes, and make strategic decisions to improve productivity and overall business performance.