ASSIGNMENT 2

Q1) What is power Query Editor?

Ans:- Power Query is a business intelligence tool available in Excel that allows you to import data from many different sources and then clean, transform and reshape your data as needed.
It allows you to set up a query once and then reuse it with a simple refresh. It's also pretty powerful. Power Query can import and clean millions of rows into the data model for analysis after. The user interface is intuitive and well laid out so it's easy to pick up. It's an incredibly short learning curve when compared to other Excel tools like formulas or VBA.
The best part about it, is you don't need to learn or use any code to do any of it. The power query editor records all your transformations step by step and converts them into the M code for you, similar to how the Macro recorder with VBA.

Q2) When was Power BI launched?

Ans:- July 2015

Q3) What is the difference between a data & a business analyst?

Ans:- here are some key differences between data and a business analyst:-

1:-

Data Analysts primarily work with large datasets to extract valuable insights and patterns, focusing on analyzing and interpreting data to provide actionable recommendations. On

the other hand, Business Analysts have a broader focus, analyzing business processes, identifying problems or opportunities, and proposing solutions to improve operational efficiency and profitability.

2:- Data Analysts require strong analytical skills, proficiency in statistical analysis tools, data querying languages, and data visualization tools. They need a deep understanding of data structures, statistical methods, and programming. Business Analysts, on the other hand, need a combination of analytical, communication, and problem-solving skills. They should have a solid understanding of business processes, requirements gathering techniques, and proficiency in tools like Microsoft Excel and project management software.

3:-

The primary objective of a Data Analyst is to extract insights from data for data-informed decision-making. They focus on answering specific questions or solving specific problems through data analysis. Business Analysts aim to improve business performance and address organizational challenges by understanding business needs, identifying opportunities for improvement, and proposing solutions aligned with business goals.

4:-

Data Analysts deliver insights through reports, data visualizations, and statistical models, providing data-driven recommendations to support decision-making processes. Business Analysts deliver artifacts like business requirement documents, process models, use cases, and feasibility studies. They provide actionable recommendations, impact assessments, and implementation plans to drive organizational improvements and support strategic decision-making.

Q4) What is Data Mining?

Ans:- Data mining, also known as knowledge discovery in data (KDD), is the process of uncovering patterns and other valuable information from large data sets. Given the evolution of data warehousing technology and the growth of big data, adoption of data mining techniques has rapidly accelerated over the last couple of decades, assisting companies by transforming their raw data into useful knowledge.

Data mining usually consists of four main steps: setting objectives, data gathering and preparation, applying data mining algorithms, and evaluating results.

1. Set the business objectives: This can be the hardest part of the data mining process, and many organizations spend too little time on this important step. Data scientists and business stakeholders need to work together to define the business problem, which helps inform the data questions and parameters for a given project. Analysts may also need to do additional research to understand the business context appropriately.

2. Data preparation: Once the scope of the problem is defined, it is easier for data scientists to identify which set of data will help answer the pertinent questions to the business. Once they collect the relevant data, the data will be cleaned, removing any noise, such as duplicates, missing values, and outliers. Depending on the dataset, an additional step may be taken to reduce the number of dimensions as too many features can slow down any subsequent computation. Data scientists will look to retain the most important predictors to ensure optimal accuracy within any models.

3. Model building and pattern mining: Depending on the type of analysis, data scientists may investigate any interesting data relationships, such as sequential patterns, association rules, or

correlations. While high frequency patterns have broader applications, sometimes the deviations in the data can be more interesting, highlighting areas of potential fraud.

Deep learning algorithms may also be applied to classify or cluster a data set depending on the available data. If the input data is labelled (i.e. supervised learning), a classification model may be used to categorize data, or alternatively, a regression may be applied to predict the likelihood of a particular assignment. If the dataset isn't labelled (i.e. unsupervised learning), the individual data points in the training set are compared with one another to discover underlying similarities, clustering them based on those characteristics.

4. Evaluation of results and implementation of knowledge: Once the data is aggregated, the results need to be evaluated and interpreted. When finalizing results, they should be valid, novel, useful, and understandable. When this criteria is met, organizations can use this knowledge to implement new strategies, achieving their intended objectives.

Q5) What is data profiling?

Ans:- Data profiling, is the process of reviewing and cleansing data to better understand how it's structured and maintain data quality standards within an organization. The main purpose is to gain insight into the quality of the data by using methods to review and summarize it, and then evaluating its condition. The work is typically performed by data engineers who will use a range of business rules and analytical algorithms.

- Data profiling helps in the understanding of data and its characteristics, whereas data mining is the process of discovering patterns or trends by analyzing the data.

- Data profiling focuses on the collection of metadata and then using methods to analyze it to support data management.
- Data profiling, unlikely data mining, produces a summary of the data’s characteristics and enables use of the data.