Experiment No.7

Title: Data Warehouse simulation

Batch: A2   Roll No.: 16010421059   Experiment No.:7

Aim: To run Data Warehouse simulation

Resources needed: Different RDBMS such as MySQL, Postgres and Excel, CSV, Rapidminer 5.3

Theory

Data Warehouse :

A data warehouse is a type of data management system that is designed to enable and support business intelligence (BI) activities, especially analytics. Data warehouses are solely intended to perform queries and analysis and often contain large amounts of historical data. The data within a data warehouse is usually derived from a wide range of sources such as application log files and transaction applications.

A data warehouse centralizes and consolidates large amounts of data from multiple sources. Its analytical capabilities allow organizations to derive valuable business insights from their data to improve decision-making. Over time, it builds a historical record that can be invaluable to data scientists and business analysts. Because of these capabilities, a data warehouse can be considered an organization's "single source of truth**.**

ETL :

Extract, Transform, Load (ETL) refers to a process in database usage and especially in data warehousing. Data extraction is where data is extracted from homogeneous or heterogeneous data sources; data transformation where the data is transformed for storing in the proper format or structure for the purposes of querying and analysis; data loading where the data is loaded into the final target database, more specifically, an operational data store, data mart, or data warehouse.

By using an established ETL framework, one may increase one's chances of ending up with better connectivity and scalability. A good ETL tool must be able to communicate with the many different relational databases and read the various file formats used throughout an organization. ETL tools have started to migrate into Enterprise Application Integration, or even Enterprise Service Bus, systems that now cover much more than just the extraction, transformation, and loading of data. A common use case for ETL tools include converting CSV files to formats readable by relational databases. A typical translation of millions of records is facilitated by ETL tools that enable users to input csv-like data feeds/files and import it into a database with as little code as possible. ETL tools in most cases contain a GUI that helps users conveniently transform data, using a visual data mapper, as opposed to writing large programs to parse files and modify data types.
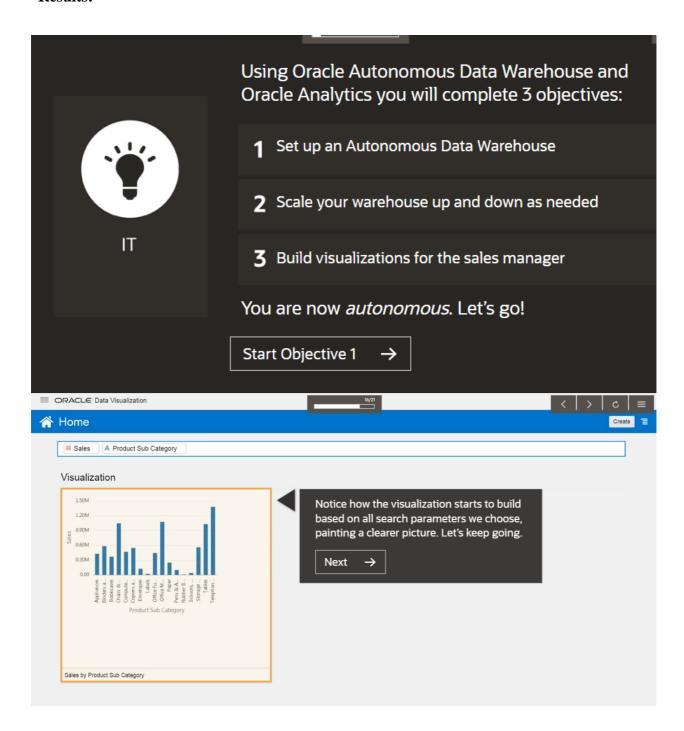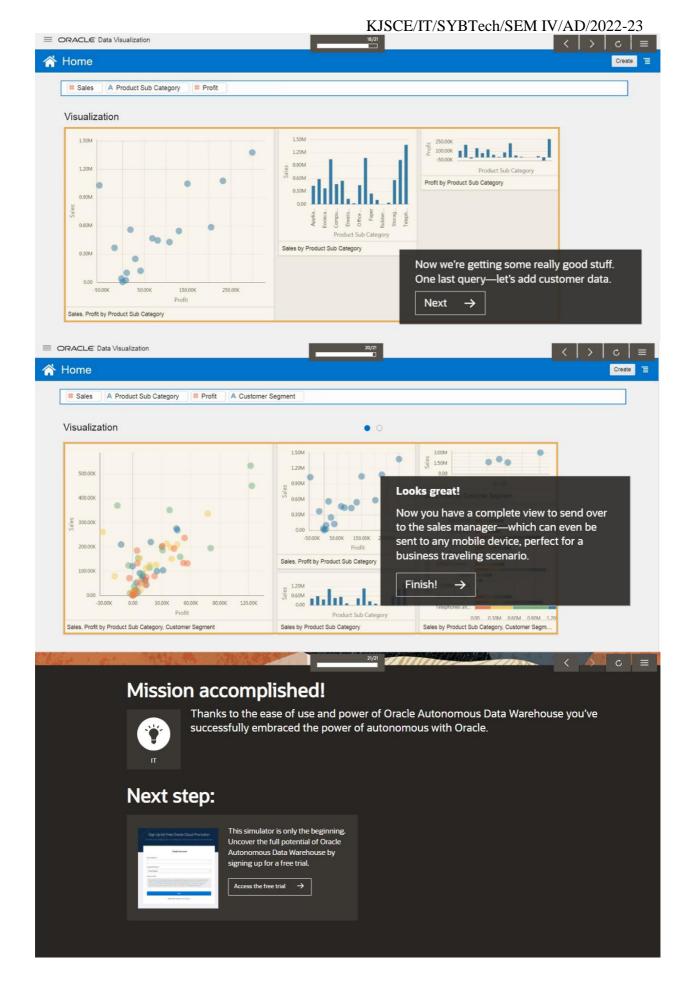

Activities:

For Data Warehouse:
1. Visit _https://www.oracle.com/webfolder/s/assets/demo/adw-quicktour-na/index.html#step1_
2. Go through the demo of Autonomous Data Warehouse for different businesses
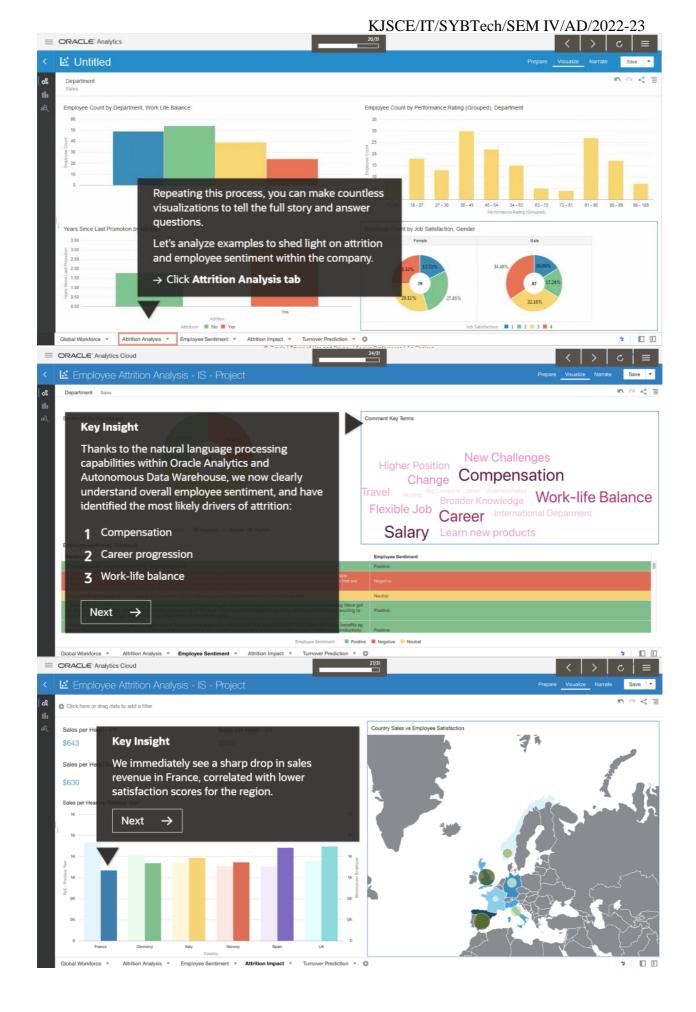3. Prepare a report with following points

a. The nature of analytics for different businesses given
b. Comparison between traditional analysis and analysis with Data Warehouse
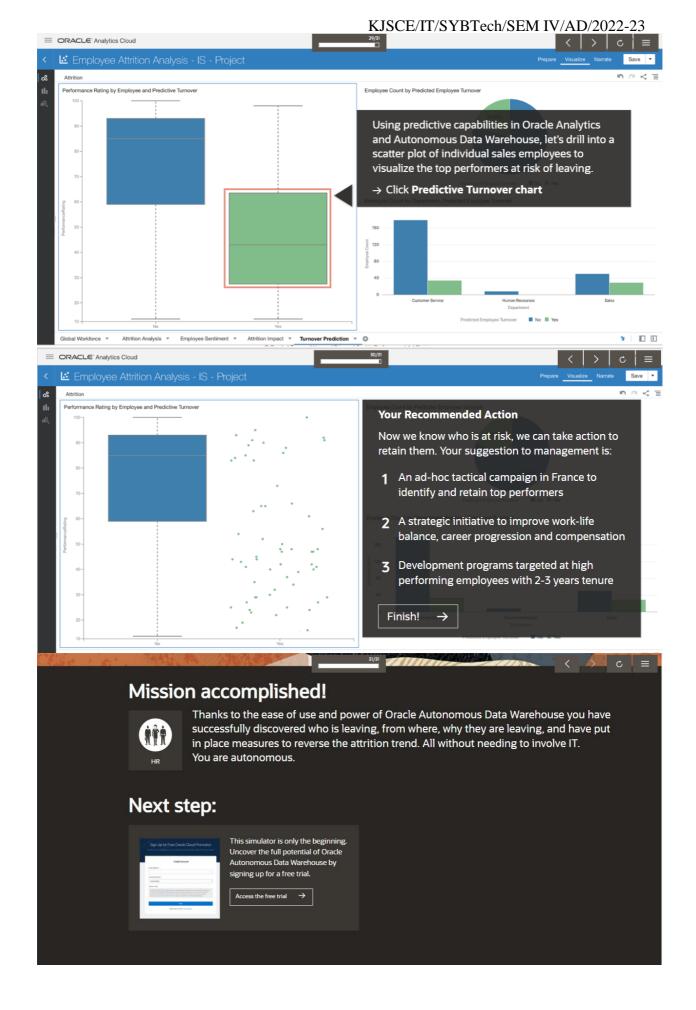c. For any two business type given, specify at least two different scenarios wherethe tool can be useful

---

Results:

Repeating this process, you can make countless visualizations to tell the full story and answer questions.

Let's analyze examples to shed light on attrition and employee sentiment within the company.

→ Click **Attrition Analysis tab**

Key Insight

Thanks to the natural language processing capabilities within Oracle Analytics and Autonomous Data Warehouse, we now clearly understand overall employee sentiment, and have identified the most likely drivers of attrition:

1 Compensation

2 Career progression

3 Work-life balance

Next →

Key Insight

We immediately see a sharp drop in sales revenue in France, correlated with lower satisfaction scores for the region.

Next →

Report for Data Warehouse

Page No.:

a. The nature of analytics for different businesses given

Ans.The nature of analytics for IT and HR in the context of data warehouse can vary depending on the specific needs and goals of each department. Generally speaking, analytics for IT and HR within a data warehouse involve using data to gain insights into various aspects of the organization.

For IT, analytics within a data warehouse can include:

- Performance monitoring: Analyzing system usage patterns, detecting bottlenecks, predicting resource requirements, and optimizing system performance.
- Security analytics: Monitoring security events, identifying potential vulnerabilities, and detecting and responding to security threats.
- User behavior analytics: Analyzing user behavior patterns to identify anomalies or suspicious activities that may indicate a security threat.
- Service level analytics: Analyzing service level data to identify trends and patterns, measure performance against SLAs, and identify areas for improvement.

For HR, analytics within a data warehouse can include:

- Talent management: Analyzing employee data to identify high-performing employees, track employee performance over time, and identify areas for improvement in talent management strategies.
- Diversity and inclusion: Analyzing employee demographics, performance evaluations, and employee engagement surveys to track diversity metrics over time, identify areas for improvement, and measure the impact of diversity and inclusion initiatives.
- Recruitment analytics: Analyzing recruitment data to identify trends in the hiring process, measure the effectiveness of recruitment strategies, and identify areas for improvement.
- Employee engagement analytics: Analyzing employee feedback and engagement data to identify areas of improvement in employee satisfaction, measure the effectiveness of engagement initiatives, and track progress over time.

In summary, analytics for IT and HR within a data warehouse involve using data to gain insights into various aspects of the organization, such as system performance, security, talent management, and employee engagement. By analyzing this data, organizations can make data-driven decisions to optimize their processes and improve their performance.

b. Comparison between traditional analysis and analysis with Data Warehouse.

Traditional analysis and analysis with Data Warehouse are two different approaches to data analysis, with their own strengths and weaknesses. Here are some of the key differences between the two:

- Data Sources: Traditional analysis typically involves analyzing data from a single source, such as a spreadsheet or a database. Data Warehouse, on the other hand, integrates data from multiple sources into a central repository, allowing for more comprehensive analysis.
- Data Volume: Traditional analysis can be limited by the amount of data that can be processed efficiently. Data Warehouse, however, is designed to handle large volumes of data, enabling more in-depth analysis.
- Data Structure: Traditional analysis usually involves analyzing data in its original format, which may not be optimal for analysis. Data Warehouse, on the other hand, transforms data into a structured format that is optimized for analysis, enabling faster and more accurate insights.
- Analysis Speed: Traditional analysis can be time-consuming, especially when dealing with large datasets. Data Warehouse, on the other hand, is designed for fast querying and analysis, allowing for more efficient analysis.
- Accessibility: Traditional analysis is typically limited to analysts with specialized skills and access to data sources. Data Warehouse, on the other hand, can be accessed by a wider range of users,

including business analysts and executives, who can use self-service analytics tools to access and analyse data.

c. For any two business type given, specify at least two different scenarios wherethe tool can be useful.

- IT Scenario: Monitoring System Performance IT departments often need to monitor the performance of their systems to identify issues and optimize their performance. A data warehouse can be useful in this scenario by integrating data from various sources such as servers, network devices, and application logs into a central repository. This allows IT analysts to perform more comprehensive analysis on the data, such as identifying patterns of system usage, detecting bottlenecks, and predicting future resource requirements. With a data warehouse, IT can gain valuable insights into system performance and make data-driven decisions to optimize their infrastructure.
- IT Scenario: Managing Customer Interactions IT departments are also responsible for managing customer interactions across various channels such as email, chat, and social media. By integrating data from these different channels into a data warehouse, IT can gain a 360-degree view of customer interactions, allowing them to identify trends and patterns in customer behavior. This can help IT teams improve the customer experience by identifying areas of improvement in their customer service processes, optimizing their response times, and providing more personalized support.
- HR Scenario: Talent Management HR departments are responsible for managing talent acquisition, retention, and development. By integrating data from various sources such as resumes, job applications, employee surveys, and performance evaluations into a data warehouse, HR teams can gain a more comprehensive view of their talent pool. This allows them to identify high-performing employees, track employee performance over time, and identify areas for improvement in their talent management strategies. With a data warehouse, HR can gain valuable insights into their workforce and make data-driven decisions to optimize talent management.
- HR Scenario: Diversity and Inclusion HR departments are increasingly focused on improving diversity and inclusion in the workplace. A data warehouse can be useful in this scenario by integrating data from various sources such as employee demographics, performance evaluations, and employee engagement surveys. This allows HR to track diversity metrics over time, identify areas where diversity and inclusion can be improved, and measure the impact of their diversity and inclusion initiatives. With a data warehouse, HR can gain valuable insights into their workforce and make data-driven decisions to improve diversity and inclusion in the workplace.

Outcomes:

CO3: Understanding of data warehouse and its multi-dimensional modeling

Conclusion: (Conclusion to be based on the outcomes achieved)

Successfully ran Data Warehouse simulation through oracle simulation for HR and IT.

Grade: AA / AB / BB / BC / CC / CD /DD

Signature of faculty in-charge with date

References:

- https://www.oracle.com/in/database/what-is-a-data-warehouse
- Paulraj Ponniah, "Data Warehousing: Fundamentals for IT Professionals", Wiley India