|  |  |  |
| --- | --- | --- |
| **Students Details** |  | nibm LOGO |
| **Module Name:** | Software Engineering |
| **Module Lecturer/ Course Coordinator:** | l.S. Chathurika |
| **Department:** | School of Computing |
| **Submission Due on:** | 13th June ,2025 | |
| **Type of Coursework:** | Report | |
| **Title of the Coursework:** | Data Warehouse and Business Intelligence | |

|  |  |  |
| --- | --- | --- |
| ***Students Details:*** | | |
|  | **Student No.** | **Student Name** |
| 01 | COHNDSE24.2F - 026 | P H N Sethulya |
| 02 | COHNDSE24.2F - 005 | M K D Gangadara |
| 03 | COHNDSE24.2F - 040 | W A U S Weerasooriya |
| 04 | COHNDSE24.2F - 008 | A H V M Jayaratine |
| 05 | COHNDSE24.2F - 085 |  |

|  |
| --- |
| **Office use only:** |
| ***Date Stamp Required of the Department*** |

**Introduction**

This project focuses on designing and making an Electronic and Home appliances retail e-commerce platform. The main purpose of this data warehouse is to combine data from different sources, analyze past performances, and generate useful reports using the Business Intelligence tools.

**Part 1: Data Warehouse Design**

The three types of data sources are,

* CSV
* JSON
* SQL

**Data Sources**

These are the source names and the format types.

|  |  |  |
| --- | --- | --- |
| Source Name | Format | Description |
| Sales Data | SQL | Contains transactional records of customer purchases, including date, product ID, quantity, and total price. |
| Customer Data | CSV | Contains customer information such as customer ID, name, email, address, and phone number. |
| Product Data | JSON | Contains product catalog details including product ID, name, category, price, and stock availability. |

**Data Sets and Data Fields**

**Sales Data**

This is the sakes data using the SQL format.

|  |  |  |
| --- | --- | --- |
| Field Name | Data Type | Description |
| Order ID | VARCHAR(14) | Unique identifier for each sales transaction. |
| Order Date | DATE | The date when the order was placed. |
| Ship Date | DATE | The date when the product was shipped to the customer. |
| Customer Id | VARCHAR(15) | Unique identifier for the customer who placed the order. |
| Product Id | VARCHAR(10) | Unique identifier for the product being sold. |
| Actual Price | DECIMAL(10,2) | The original listed price of the product before any discounts. |
| Discounted Price | DECIMAL(10,2) | The discounted price offered at the time of sale. |
| Sold Price | DECIMAL(10,2) | Final price at which the product was sold (after discount and any adjustments). |
| Quantity | NUMERIC(5,2) | Number of product units sold in this order. |
| Bill Amount | DECIMAL(10,2) | Total billing amount for the order (Sold Price × Quantity). |

**Customer Data**

This is the customer data using the CSV format.

|  |  |  |
| --- | --- | --- |
| Field Name | Data Type | Description |
| Index | **Integer** | **Row number in the dataset (may be used as a simple index).** |
| Customer Id | **String** | **Unique identifier for each customer (alphanumeric).** |
| First Name | **String** | **Customer's first name.** |
| Last Name | **String** | **Customer's last name.** |
| Company | **String** | **Company or organization the customer is associated with.** |
| City | **String** | **City of residence.** |
| Country | **String** | **Country of residence.** |
| Phone 1 | **String** | **Primary contact number.** |
| Email | **String** | **Customer’s email address.** |

**Product Data**

Product data using JSON format.

|  |  |  |
| --- | --- | --- |
| Field Name | Data Type | Description |
| product\_id | String | Unique identifier for the product. |
| product\_name | String | Full product name including specifications. |
| category | String | Hierarchical category path separated by ` |
| discounted\_price | String | Current selling price (may include currency symbol). |
| actual\_price | String | Original price before discount (may include currency symbol). |
| discount\_percentage | String | Discount percentage as a string (e.g., "64%"). |
| rating | Float | Average user rating out of 5. |
| rating\_count | String | Total number of ratings (may include commas). |
| about\_product | String | Bullet-point style product description, delimited by ` |
| img\_link | String (URL) | Link to the product image. |
| product\_link | String (URL) | Direct link to the product page. |

**Data Warehouse Schema**

Using Galaxy Schema, we have created the design for the project. These are the fact and dimensional tables we created.

We created three fact tables:

* **Sales Fact** – stores information about customer purchases.
* **Purchased Fact** – stores details of products bought from suppliers.
* **Inventory Fact** – tracks stock levels over time.

We also used several dimension tables to describe the facts:

* **Customer** – who made the purchase.
* **Date** – when the activity happened.
* **Supplier** – who supplied the product.
* **Satisfaction Levels** – how customers felt after purchase.
* **Product** – what was sold or stocked.
* **Location** – where the product was sent or stored.

This design helps us easily answer questions like:

* What are the best-selling products?
* Which suppliers deliver the most?
* How is customer satisfaction changing over time?
* What location does products sell the most?
* During what time(month, year) does products sell the most?

**Galaxy Schema**

A diagram of a computer

AI-generated content may be incorrect.