Data Analysis Assignment

# Objective

The objective of this assignment is to review the provided historical billing data, upload the data into a MySQL database as separate tables, and perform analysis to extract meaningful insights.

# Data preparation

Task: Review the provided data files and structure them for upload into MySQL.

Deliverables:

1. Create a MySQL database and define appropriate tables to match the data

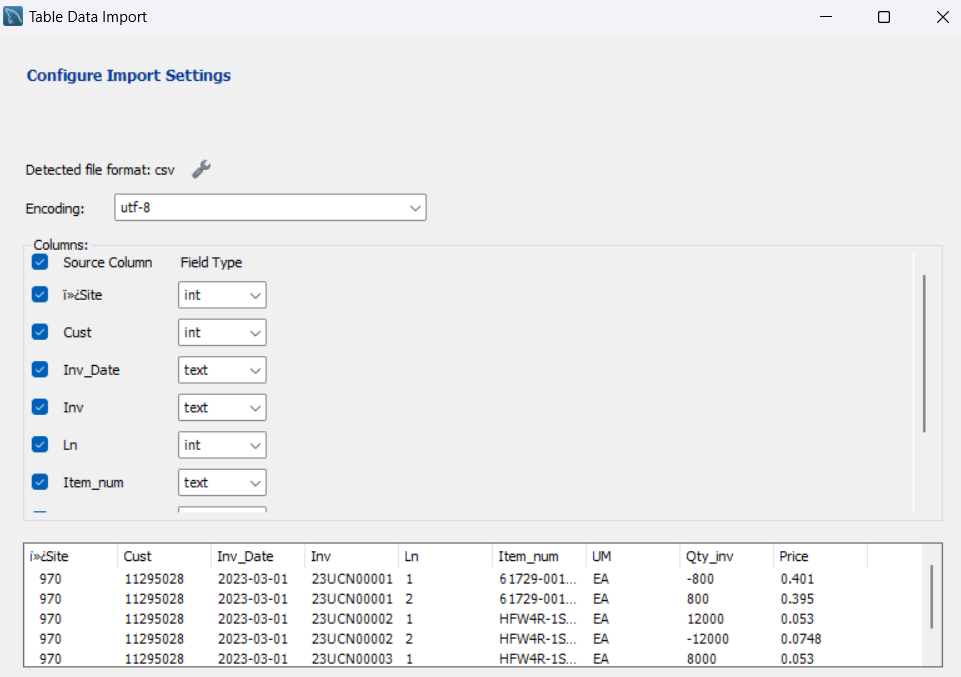
provided in the Excel sheets.

1. Upload the data into the respective MySQL tables using appropriate scripts or

import tools.

I have used Table Data Import Wizard to upload data into the respective MySQL tables.





# Tables to create in MySQL

## Invoice Sheet

|  |  |  |  |
| --- | --- | --- | --- |
| COLUMN NAME | DATA TYPE | CONSTRAINTS | DESCRIPTION |
| Inv\_Site | INT | PRI KEY, NOT NULL | Unique identifier for each transaction |
| Inv\_CustNbr | INT | FOREIGN KEY REFERENCES  Customer\_master(custmast\_CustNbr),  NOT NULL | Identifier linking to the Customer\_Master table. |
| Inv\_Nbr | VARCHAR(255) | PRI KEY, NOT NULL | Unique identifier for each transaction |
| Inv\_Line | INT | PRI KEY, NOT NULL | Unique identifier for each transaction |
| Inv\_ItemNbr | VARCHAR(255) | FOREIGN KEY REFERENCES  Part\_Master (Partmast\_ItemNbr),  NOT NULL | Identifier linking to the Part\_Master table. |
| Inv\_UM | TEXT | IS NULL |  |
| Inc\_Qty | BIGINT | IS NULL |  |
| Inv\_Price | DOUBLE | IS NULL |  |
| Inv\_Date | DATE | IS NULL |  |

## Customer\_Master

|  |  |  |  |
| --- | --- | --- | --- |
| COLUMN NAME | DATA TYPE | CONSTRAINTS | DESCRIPTION |
| Custmast\_CustNbr | INT | PRI KEY, NOT NULL | Unique identifier for each Customer |
| Custmast\_Name | TEXT | NULL |  |
| Custmast\_GenAcc | TEXT | NULL |  |
| Custmast\_CtryCode | TEXT | NULL |  |
| Custmast\_type | INT | FOREIGN KEY REFERENCES  Cust\_Type(Cust\_Type) | Identifier linking to the Cust\_Type table. |

## Part\_Master

|  |  |  |  |
| --- | --- | --- | --- |
| COLUMN NAME | DATA TYPE | CONSTRAINTS | DESCRIPTION |
| Partmast\_Itemnbr | VARCHAR(255) | PRI KEY, NOT NULL | Unique identifier for each Part |
| artmast\_HPLcode | INT | FOREIGN KEY REFERENCES  HPL(HPLNbr),  NOT NULL | Identifier linking to the HPL table. |
| Partmast\_Status | TEXT | FOREIGN KEY REFERENCES  Part\_Status(Partstatus\_Status),  NOT NULL | Identifier linking to the Part\_Status table. |
| Partmast\_SOP | VARCHAR(255) | FOREIGN KEY REFERENCES  SOP(SOP\_nbr),  NOT NULL | Identifier linking to the SOP table. |
| Partmast\_UM | TEXT | IS NULL |  |
| Partmast\_Stdcost | DOUBLE | IS NULL |  |

## Part\_Status

|  |  |  |  |
| --- | --- | --- | --- |
| COLUMN NAME | DATA TYPE | CONSTRAINTS | DESCRIPTION |
| Partstatus\_Status | VARCHAR(255) | PRIMARY KEY NOT NULL | Unique identifier for each description |
| Partstatus\_desc | VARCHAR(255) | NULL |  |

## SOP

|  |  |  |  |
| --- | --- | --- | --- |
| COLUMN NAME | DATA TYPE | CONSTRAINTS | DESCRIPTION |
| SOP\_nbr | VARCHAR(255) | PRIMARY KEY NOT NULL | Unique identifier for each SOP description |
| SOP\_desc | TEXT | NULL |  |

## HPL

|  |  |  |  |
| --- | --- | --- | --- |
| COLUMN NAME | DATA TYPE | CONSTRAINTS | DESCRIPTION |
| HPL\_nbr | VARCHAR(255) | PRIMARY KEY NOT NULL | Unique identifier for each SOP description |
| HPL\_desc | TEXT | NULL |  |

## Cust\_Type

|  |  |  |  |
| --- | --- | --- | --- |
| COLUMN NAME | DATA TYPE | CONSTRAINTS | DESCRIPTION |
| Cust\_Type | INT | PRIMARY KEY NOT NULL | Unique identifier for each customer description |
| Cust\_Desc | TEXT | NULL |  |
|  |  |  |  |

# Basic Statistics

## Total records count from Invoice\_sheet: 48,057

## Count of unique part numbers in Invoice\_sheet: 7,761

## Count of unique customers in Invoice\_sheet table: 207

## Count of unique invoices in Invoice\_sheet table: 35,514

## Total invoice value: $80,207,836

## Count of unique HPL code in partmaster: 134

## Count of unique SOP code in partmaster: 495

## count of unique generic account in customer master: 72

## count of unique country codes in customer master: 9

## count of distinct parts transacted (in the invoice table) for each HPL\*: HPL-578has 1212 Transactions

## count of distinct parts transacted (in the invoice table) for each SOP\*: SOP-5NPF has 719 Transaction

**\* Has multiple outputs, giving single example.**

# Analysis Scenarios

## Total Sales by Top 10 Customer

The bar chart displays sales figures for the top 10 customers, with **AVNET ASIA PTE LTD (TAIWAN BR)** leading at **$7.91M**.

**TTI INC** follows closely with sales amounting to **$6.62M**.

**ALLIANCE AEI COMMERCIAL WHSE** registers the third-highest sales at **$6.25M**.

## Monthly Sales Trend

The line chart illustrates the monthly sales trend across the year, showing significant variability.

**January** marks a high point with **$12.76M**, while **December** shows a substantial dip to **$5.10M**.

**Key insights:**

* Another peak occurs in June with **$10.14M**, suggesting a mid-year sales increase.
* The latter part of the year shows a general **declining trend** from September onward, with minor fluctuations.
* The sales trajectory indicates **seasonality**, with peaks likely driven by specific market activities or demand cycles in January and June.

## Top 10 Products

**Key insights:**

* The sales distribution shows a **steep decline after the top two products**, suggesting they might be flagship or highly demanded items in the market.
* The considerable sales gap between the top products and others may indicate niche markets or specialized uses for the top sellers.

## Regional Sales Distribution

The bar chart presents sales figures across various regions, with the **United States (US)** dominating the chart at **$54.7M**.

**Canada (CA)** and **Taiwan (TW)** follow with **$12.1M** and **$9.8M** respectively.

**Key insights:**

* The **US market** is by far the most significant, suggesting a strong market presence and potentially favourable market conditions or customer base.
* The drastic drop-off in sales after the top three regions indicates high regional concentration of sales in **North America and Taiwan**.

## Customer Segmentation by Type

**Key insights:**

* **Domestic sales** more than double the revenue of **export sales**, highlighting a stronger market penetration or demand within the domestic sector.
* Despite the lower figure, **export sales** still represent a substantial market, suggesting potential for growth or strategic focus.

## Cost Analysis of Products

Very High Profitability (>75% Margin): 69 products, representing 1% of the total product count.

High Profitability (>50% Margin): 144 products, making up 3% of the total.

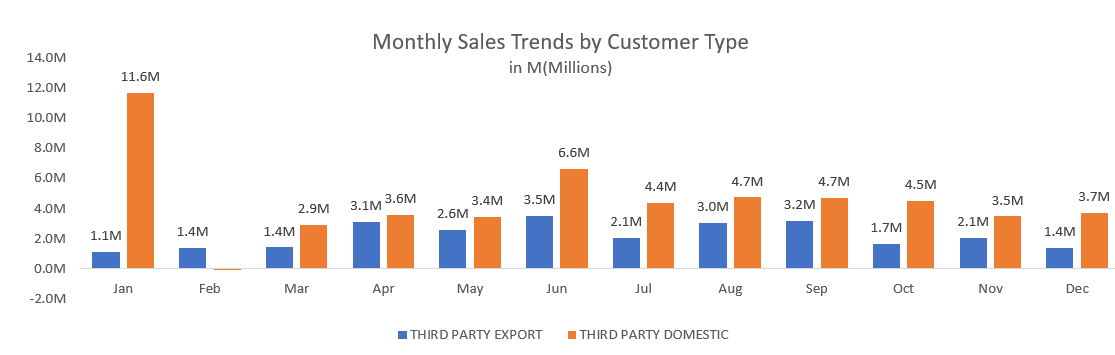
Medium Profitability (>25% Margin): 2,078 products, accounting for 35% of the total.

Low Profitability (≤25% Margin): 3,611 products, dominating the distribution with 61% of total products.

**Key Insights:**

* The **majority of products (61%) fall within the low profitability range**, indicating a significant portion of the inventory generates limited profit margins.
* Products with **very high profitability are minimal** but could potentially offer niche market leverage or premium pricing opportunities.
* **Medium profitability products form a substantial part of the inventory**, suggesting a balanced approach in the product lineup.

## Monthly Sales Trends by Customer Type



**January** records a significant peak for Third Party Domestic sales at **$11.6M**, substantially higher than any other month.

Export sales show more consistency month-to-month, generally staying below **$4M.**

**Key Insights:**

* The sales trend indicates a **strong reliance on domestic activities**, especially in January and June, which could be tied to seasonal demand or specific market dynamics.
* Export sales, while lower in volume, offer a stable revenue stream across the year, which could serve as a financial buffer against the variability in domestic markets.

## Top 10 Customer by Revenue

**AVNET ASIA PTE LTD TAIWAN BR:** Leads the chart with the highest total revenue of **$7.91M**, contributing **9.86%** to overall sales.

**MOUSER ELECTRONICS** and **AVNET EM** are also significant contributors, with revenues exceeding **$5M**, but their sales contributions slightly drop below **7%**.

**Key Insights:**

* The top two customers, **AVNET ASIA PTE LTD TAIWAN BR** and **TTI INC**, together account for over **18%** of total sales, highlighting their strategic importance.
* The revenue and contribution trend shows a moderate decline, suggesting a relatively **high dependency on the top few customers** for a significant portion of total sales.

## Assess the contribution of each HPL Code to total sales revenue & assess the contribution of each HPL Code to total Profit

* Identify the top HPL’s contributing to 30% of the sales
* Identify the top HPL’s contributing to 30% of the profit

**Sales Revenue Contribution:**

* The top HPL code, **627 MEG-ARRAY R+**, significantly drives sales, contributing 19% of total revenue.
* Other key contributors include **581 LOT CHARGES MACHINES-LOT (5%)**, **578 Bergstik (4%)**, and **522 IDC QUICKIE (4%)**.
* A large portion (68%) of the sales comes from all other HPLs combined, indicating a dispersed contribution across smaller players.

**Profit Contribution:**

* **627 MEG-ARRAY R+** leads in profit generation, accounting for 28% of the total profit.
* **549 MEG-ARRAY** contributes 8%, while the remaining 64% of profits are distributed among all other HPLs.

**Key Insights:**

* **627 MEG-ARRAY R+** is a critical driver for both sales and profit, showcasing its importance for business strategy.
* There is a concentrated profit share compared to a more dispersed sales share, implying that profitability varies significantly by HPL.

## Identify the top 10 HPL’s having the highest average monthly qty sold.

**MICROSPACE** leads with the highest monthly quantity sold at **3.0M units**, followed by **Bergstik (2.5M units)** and **FFC/FPC SYSTEMS (1.6M units)**.

Other notable contributors, all with quantities ranging from **0.6M to 0.3M units**, include **PV**, **USB**, **MINITEK 2.0MM**, **GRIPLET**, **IDC QUICKIE**, and **D-Sub Professional**.

The dominance of **MICROSPACE** and **Bergstik** indicates they are the most significant drivers of volume among HPLs.

**Key Insights:**

1. **MICROSPACE** and **Bergstik** are far ahead of the others in terms of average monthly sales quantity, making them key products for scaling volume-driven revenue.
2. Products like **FFC/FPC SYSTEMS** and **PV** also show strong volume performance and could be strategically leveraged.
3. The remaining HPLs, while significant in this top 10 list, show comparatively lower volumes and may require a targeted approach to grow sales.