End-to-End Business Data Analysis for MIVI

From Creation to Insightful Reporting Using MySQL, Python, and Power BI

Project file:

-End-to-End-Business-Data-Analysis-for-MIVI-MySQL-Python-and-Power-BI-Project-/Mivi's internship task at main · Sandeep200-bit/-End-to-End-Business-Data-Analysis-for-MIVI-MySQL-Python-and-Power-BI-Project-

Report By:

Sandeep Kotipalli

Kotipallisandeep9@gmail.com

Linked id - http://www.linkedin.com/in/sandeep-kotipalli-a94921203

1. About Myself

My name is Sandeep Kotipalli. I am a detail-oriented data analyst with a Bachelor of Science in Computer Science from Dr. B.R. Ambedkar University, Srikakulam, Andhra Pradesh. I have solid skills in statistical analysis, data cleaning, data processing, and creating visual insights using modern tools. My technical skillset includes Python, MySQL, Excel, Tableau, and Power BI. I am passionate about finding insights in real-life data to support better business decisions. I am eager to apply my academic training and hands-on experience to real-world problems, and I am seeking opportunities that will let me prove my abilities through practical tasks.

2. Assignment Rules and Choice of Tools

Rules to Follow

For this MIVI internship task, these rules were given:

- **Identify** a real business problem from a company like MIVI.
- **Create** or gather data related to that problem.
- **Clean** and process the data to prepare it for analysis.
- **Analyze** the data to get practical insights.
- **Present** the results using clear visualizations and a final report.
- **Complete** the whole process with proof within 3 days.

Choice of Tools

- **MySQL:** Used for storing and organizing data efficiently in tables, making it easy to work with large, structured datasets and perform quick data manipulations.
- **Python:** Selected for its powerful libraries (like pandas and matplotlib) to perform advanced data cleaning, handle complex formats, automate repetitive work, and carry out deeper analysis.
- **Power BI:** Used to turn processed data into attractive dashboards and visuals, helping present key insights to business users in a clear and interactive way.

Each tool was chosen for its ability to handle one part of the process well, together covering the complete journey from raw data to actionable business insight.

3. Why MIVI and Business Problems Selected

Why MIVI

MIVI is a well-known Indian electronics brand focused on audio products like earphones and speakers. The company stands for "Made in India" innovation and faces real-world challenges such as manufacturing at scale, managing after-sales service, and building its brand in a competitive tech market. Studying MIVI allows me to apply data analysis where it matters—solving manufacturing, sales, and marketing issues that impact both the business and its customers.

Selected Business Problems

After reviewing common issues for companies like MIVI, I chose these three focus areas:

- **Supply Chain Challenges:** Difficulties in sourcing components and maintaining efficient production lines.
- **After-Sales Service Issues:** Patterns and delays in customer complaints, defects, and warranty processing.
- **Marketing & Brand Awareness:** Effectiveness and reach of varied marketing campaigns across regions and platforms.

Project Roadmap

My roadmap was:

- **Create** messy, realistic datasets for each problem using Python.
- **Import** them into MySQL, set up tables, and check for structure.
- **Use Python** to clean and analyze the data, connecting directly to MySQL for efficiency.
- **Export the results** as cleaned datasets for visualization.
- **Visualize** all findings in Power BI dashboards tailored to each business area.
- **Compile** a final report with visuals and business recommendations.

4. Data Creation with Python

Since real MIVI data is confidential, I created three large datasets in Python, each with 10+ columns and at least 500 rows, to mimic real business data—with intentional "messiness" such as inconsistent dates, blanks, duplicates, typos, and mixed formats.

a) Supply Chain Dataset

- **Purpose:** Track parts, suppliers, costs, wastage, quantity produced, production dates, and inspection status.

- **Columns:** Part, Cost, Wastage, Quantity, Date, Supplier, Batch, Notes, Production Line, Inspection Status.
- **Rows:** 500+. Example entries show a wide range of actual supplier and manufacturing scenarios.

b) After-Sales Defect Dataset

- **Purpose:** Log customer sales, complaints, regions, dates, types of problems, how fast they were solved, and status.
- **Columns:** Sale ID, Product Type, Region, Sale Date, Customer Code, Complaint Date, Complaint Type, Status, Resolution Time, Comments.
- **Rows:** 500+. Scenarios cover multiple problems, products, regions, and outcomes.

c) Marketing Brand Awareness Dataset

- **Purpose:** Analyze marketing campaigns, platforms, spend, engagement, and response.
- **Columns:** Campaign ID, Platform, Name, Start/End Date, Budget, Impressions, Clicks, Engagement, Notes, Audience.
- **Rows:** 500+. Entries represent many platforms and campaign types with natural data issues.

These datasets set the stage for realistic data cleaning and analysis work.

5. MySQL Work for Data Gathering

In this stage, I:

- **Created a database** called mivi_assignment in MySQL.
- **Built three tables**—one for each dataset—with column types that match real data.
- **Imported each CSV** (generated from Python) into its respective MySQL table.
- **Checked** the import using SQL queries to count rows and view sample data.
- **Ensured data consistency**, making it easier to run further queries or connect the data with Python.

6. Python Work: Data Connection, Cleaning, EDA, and Export

SQL to Python Connection

Using libraries like pymysql or sqlalchemy, I established a connection from Python to MySQL. I used SQL queries to directly access my data, loaded the tables into pandas DataFrames, and began the cleaning process.

Data Cleaning

This included:

- **Fixing inconsistent dates:** Applied parsing logic to handle various formats and convert to a standard.
- **Handling missing values:** Replaced NAs with medians/means for numeric columns, or with clear markers for categories.
- **Removing duplicates:** Ensured only one row per record.
- **Text standardization:** Trimmed spaces, fixed capitalization, and standardized category names.
- **Type conversions:** Changed columns to numeric or categorical types wherever appropriate.

Data Processing and EDA

- **Exploratory analysis:** Counted types, calculated averages/medians, examined relationships, and checked for outliers.
- **Visual checks:** Created preliminary plots using matplotlib and seaborn to discover patterns.

Export

After cleaning and analysis, I exported all three cleaned DataFrames as CSVs for loading into Power BI.

7. Power BI Work: Further Cleaning and Dashboard Creation

Data Cleaning in Power BI

- Used **Power Query** to do final cleaning, such as:
    - Removing duplicates by key columns.
    - Ensuring date columns were recognized and formatted.
    - Trimming and cleaning text for slicers and visuals.
    - Fixing data types for numbers and dates.

Dashboards

A. Production Dashboard ("Production Pulse")

- **Card:** Shows total units produced.
- **Slicers:** Filter by production line and part.
- **Table:** Top suppliers by production volume.
- **Visuals:** Monthly trend (line), wastage percentage by part (bar), inspection status breakdown (donut).

B. Sales Dashboard ("Service & Sales Insights")

- **Card:** Total complaints logged.
- **Slicers:** Filter by product type and region.
- **Table:** Top complaint types.
- **Visuals:** Complaint resolution times (histogram), complaint status (pie), complaints by month (column).

C. Marketing Dashboard ("Marketing Momentum")

- **Card:** Total campaigns run.
- **Slicers:** Select by platform and audience.
- **Table:** Top campaigns by engagement rate.
- **Visuals:** Budget vs impressions (scatter), clicks by platform (bar), monthly engagement trend (line).

Each dashboard used modern, clear titles and visual types that match business user needs.

8. Final Thoughts and Recommendations

Achievements:

- Demonstrated the ability to create, clean, process, and visualize data end-to-end.
- Tackled relevant business problems for a real-world Indian company.
- Sharpened skills in MySQL, Python, and Power BI—all highly valued by employers.

Improvements Suggested:

- **Supply Chain:** Improve real-time tracking to reduce wastage and quickly detect cost increases.
- **After-Sales:** Speed up complaint resolution by analyzing which products or regions have the most issues.

- **Marketing:** Focus spending on platforms and audiences with the highest engagement and best ROI.

Next steps:

- Automate data updates for continuous insights.
- Integrate real, live data sources for production usage.
- Expand analysis to include predictive modelling for greater business value

This structured report shows my full workflow, thought process, and technical choices, making it clear how I turned a challenging business problem into a data-driven action plan—ready to present and discuss for the internship assessment.

The End