**WideWorldImporters**  
**Data Analysis Project**

**Prepared by:**  
• AbdelRahman AbdelMoez Anwar  
• Fatma Ali Khaled  
• Youssef Mohamed Farag  
• Noha Soliman Mohamed

**Supervised by:** Eng. Ahmed Alaa  
**Group:** ALX2\_DAT1\_G1 – Google Data Analyst Specialist

**Table of Contents**

1. Project Planning & Management
2. Literature Review
3. Requirements Gathering
4. System Analysis & Design
5. Data Analysis Track
6. Conclusion

**1. Project Planning & Management**

**1.1 Project Overview**

This project utilizes the WideWorldImporters database to generate actionable insights in the domains of Sales, HR, Supply Chain, and Marketing. The primary objective is to enhance decision-making through structured data analysis and visualization.

**1.2 Objectives**

* Provide data-driven insights for improved business decisions.
* Develop interactive dashboards and reports using SQL Server, Power BI, Python, and Excel.
* Ensure data integrity and optimize analytical methodologies.

**1.3 Project Scope**

* Exclusive focus on data analysis methodologies.
* Use of SQL Server for data extraction.
* Python and Excel for data processing.
* Power BI for dashboard creation and visualization.

**1.4 Project Timeline**

**Week | Task**

1. Database familiarization & task assignments
2. Data extraction & initial cleaning
3. Exploratory Data Analysis (EDA)
4. Dashboard design in Power BI
5. Insights refinement & dashboard usability testing
6. Drafting documentation
7. Final reporting & recommendations
8. Presentation preparation & submission

**1.5 Task Assignments & Roles**

* **AbdelRahman AbdelMoez Anwar:** Data Visualization (Power BI), Business Insights.
* **Fatma Ali Khaled:** Data Cleaning & Processing (Python, Excel).
* **Youssef Mohamed Farag:** SQL-based Data Extraction.
* **Noha Soliman Mohamed:** Documentation & Presentation Preparation.

**1.6 Risk Assessment & Mitigation Plan**

* **Data Integrity Issues:** Use robust cleaning techniques in Python (Pandas) & SQL validations.
* **Time Constraints:** Adhere to timeline & conduct regular progress reviews.
* **Visualization Complexity:** Focus dashboards on Key Performance Indicators (KPIs) for clarity.

**1.7 Key Performance Indicators (KPIs)**

* **Sales:** Revenue growth, best-selling products, customer retention rate.
* **HR:** Employee turnover rate, performance analysis.
* **Supply Chain:** Inventory turnover, supplier efficiency.
* **Marketing:** Campaign effectiveness, customer acquisition cost.

**2. Literature Review**

**2.1 Feedback & Evaluation**

* **Strengths:** Effective SQL queries and interactive dashboards.
* **Areas for Improvement:** Enhanced data cleaning techniques, deeper narrative insights.

**2.2 Suggested Improvements**

* **Advanced Analytics:** Implement predictive modeling.
* **Enhanced Dashboards:** Add interactive filters & drill-down capabilities.
* **Improved Documentation:** Provide in-depth business impact explanations.

**2.3 Grading Criteria**

* **Documentation:** Clarity, structure, and reporting depth.
* **Implementation:** Effective SQL, Python, and Power BI utilization.
* **Testing:** Accuracy in extraction, cleaning, and visualization.
* **Presentation:** Delivery quality, storytelling, and stakeholder relevance.

**3. Requirements Gathering**

**3.1 Stakeholder Analysis**

* **Sales Managers:** Revenue trends & product distribution.
* **HR Specialists:** Employee turnover & retention strategies.
* **Supply Chain Teams:** Supplier reliability & inventory management.
* **Marketing Executives:** Campaign performance & customer acquisition costs.

**3.2 Functional & Non-Functional Requirements**

| **Type** | **Requirements** |
| --- | --- |
| **Functional** | Data Extraction (SQL), Data Cleaning (Python, Excel), Visualization (Power BI), Reporting (Excel). |
| **Non-Functional** | Performance (fast report generation), Security (restricted access), Usability (user-friendly UI). |

**4. System Analysis & Design**

**4.1 System Architecture**

* **Backend:** SQL Server for data storage & querying.
* **Processing:** Python (Pandas, Matplotlib, Seaborn) for data manipulation.
* **Visualization:** Power BI for interactive dashboards.
* **Reporting:** Excel for detailed insights.

**4.2 Data Flow & System Behavior**

1. **Data Extraction:** Retrieve raw data using SQL.
2. **Data Cleaning:** Apply preprocessing using Python & Excel.
3. **Data Analysis:** Perform trend analysis & derive insights.
4. **Visualization & Reporting:** Develop dashboards & generate reports.

**4.3 UI/UX Design Principles**

* **Consistency:** Standardized color schemes & typography.
* **Accessibility:** Intuitive navigation & tooltips.
* **Interactivity:** Drill-down capabilities & filters.

**4.4 Deployment Strategy**

* **Hosting:** Power BI Service for dashboards, scheduled SQL & Python scripts.
* **Security Measures:** User-based access restrictions.

**5. Data Analysis Track**

**5.1 Data Cleaning & Preprocessing**

* Handle missing values using imputation techniques.
* Remove duplicates using SQL DISTINCT and Pandas drop\_duplicates().
* Standardize data formats (dates, currencies).

**5.2 Exploratory Data Analysis (EDA)**

* **Sales:** Top products, revenue trends.
* **HR:** Employee retention, salary distributions.
* **Supply Chain:** Supplier efficiency, inventory turnover.
* **Marketing:** Campaign performance & customer insights.

**5.3 Data Visualization & Reporting**

* Develop interactive Power BI dashboards.
* Create Excel reports with pivot tables & charts.

**5.4 System Deployment & Automation**

* Schedule SQL queries and Python scripts for real-time data updates.
* Secure dashboard access for stakeholders.

**5.5 Final Deliverables**

* **Executive Summary Report:** Key insights & business recommendations.
* **Stakeholder Presentation:** Storytelling through data visualizations.

**6. Conclusion**

This project provides a structured approach for analyzing WideWorldImporters data using industry-standard tools. The combination of SQL, Python, Power BI, and Excel ensures data integrity and enables effective decision-making in Sales, HR, Supply Chain, and Marketing. The insights generated will help stakeholders optimize business strategies and drive growth.

**GitHub Links:**  
https://github.com/youssefm0/DEPI-Project-Documentation