Business Requirements Document (BRD)

*Project Title:*  
**Implementation of Data Warehouse for Northwind Traders Dataset**

**Prepared by:**Ahmed Fawzy

**Date:**September 30, 2024

**Table of Contents**

[1. Introduction 3](#_heading=h.gjdgxs)

[2. Business Objectives 3](#_heading=h.30j0zll)

[3. Project Scope 3](#_heading=h.1fob9te)

[4. Project Stages 4](#_heading=h.3znysh7)

[4.1. Dataset Selection 4](#_heading=h.2et92p0)

[4.2. Business Requirements 4](#_heading=h.tyjcwt)

[4.3. Data Modeling 4](#_heading=h.3dy6vkm)

[4.3.1. Conceptual Model 4](#_heading=h.1t3h5sf)

[4.3.2. Logical Model 4](#_heading=h.4d34og8)

[4.3.3. Physical Model 4](#_heading=h.2s8eyo1)

[4.4. ETL Process 4](#_heading=h.17dp8vu)

[4.4.1. ODS (Operational Data Store) 4](#_heading=h.3rdcrjn)

[4.4.2. Stage 5](#_heading=h.26in1rg)

[4.4.3. Data Warehouse 5](#_heading=h.lnxbz9)

[4.5. Administration 5](#_heading=h.35nkun2)

[4.6. Analytics 5](#_heading=h.1ksv4uv)

[5. Non-Functional Requirements 6](#_heading=h.44sinio)

[6. Stakeholders 6](#_heading=h.2jxsxqh)

[7. Assumptions 6](#_heading=h.z337ya)

[8. Constraints 6](#_heading=h.3j2qqm3)

[9. Risks 6](#_heading=h.1y810tw)

[10. Deliverables 7](#_heading=h.4i7ojhp)

[11. Timeline and Milestones 7](#_heading=h.2xcytpi)

# 1. Introduction

This Business Requirements Document (BRD) outlines the business needs and requirements for the implementation of a data warehouse for the Northwind Traders dataset. Northwind Traders is a fictitious company that imports and exports specialty foods worldwide. The project involves setting up a data warehouse using the Northwind dataset, implementing ETL processes, and building an analytics solution with Power BI and an SSAS tabular model.

# 2. Business Objectives

The objectives of this project include:

* **Improved Reporting:** Enable dynamic reporting on customers, suppliers, orders, and products.
* **Better Decision-Making:** Facilitate data-driven decision-making by consolidating business data into a centralized data warehouse.
* **Advanced Analytics:** Provide tools for forecasting, trend analysis, and performance metrics using SSAS tabular models and Power BI.
* **Efficient Data Management:** Streamline data extraction, transformation, and loading (ETL) processes to ensure consistency and accuracy of business data.

# 3. Project Scope

This project covers the following stages:

* **Dataset Selection:** Identify and prepare the Northwind dataset for use.
* **Business Requirements:** Define the key business needs for data integration and analytics.
* **Data Modeling:** Create the data warehouse's conceptual, logical, and physical models.
* **ETL Process:** Implement an ETL pipeline using SSIS to load data from source to staging, and from staging to the data warehouse.
* **Administration:** Automate and manage SSIS package executions and monitoring.
* **Analytics:** Build an SSAS tabular model for analysis and a Power BI dashboard for data visualization.

# 4. Project Stages

## 4.1. Dataset Selection

The Northwind dataset includes:

* **Suppliers:** Information about vendors supplying products.
* **Customers:** Customer records, including contact details and location.
* **Employees:** Employee data is used to track employee details.
* **Products:** Product information, including categories and suppliers.
* **Shippers**: Data on shippers used to deliver products.
* **Orders and Order Details:** Sales transactions, including order dates and products ordered.

## 4.2. Business Requirements

* **Reporting:** Detailed reporting on product sales, customer orders, and supplier performance.
* **KPI Tracking:** Track KPIs such as total sales by product category, sales growth, and customer retention.
* **Analytics:** Enable advanced analytics such as trend forecasting, sales performance, and inventory optimization.

## 4.3. Data Modeling

### 4.3.1. Conceptual Model

* High-level understanding of business entities such as **Customers**, **Orders**, **Products**, **Employees**, **Shippers**, and **Suppliers**.
* Relationships between these entities will define how data is structured in the warehouse.

### 4.3.2. Logical Model

* **Dimensional Modeling:** Star schema to structure data for analytical queries.
  + **Fact Tables:** Store quantitative data (e.g., sales transactions).
  + **Dimension Tables:** Store descriptive data (e.g., products, customers, employees).

### 4.3.3. Physical Model

* Create the actual tables, columns, and relationships in the database system.
* Optimize the database for query performance, indexing, and partitioning.

## 4.4. ETL Process

### 4.4.1. ODS (Operational Data Store)

* Data extracted from the Northwind dataset is loaded into the Operational Data Store (ODS) for real-time data processing and validation.

### 4.4.2. Stage

* Intermediate storage where raw data is transformed before loading it into the data warehouse.
* Perform data cleaning, standardization, and validation in this layer.

### 4.4.3. Data Warehouse

* The final storage is where data is loaded from staging after transformation.
* Organized into Fact and Dimension tables using a star schema for fast querying.

## 4.5. Administration

**Automation**

* **Automate SSIS Package Runs:** Configure scheduled execution of SSIS packages to automate ETL processes. The SSIS packages will be scheduled to run at 9:00 PM daily to ensure timely updates of the data warehouse.

**Job Monitoring and Alerts**

* **Tracking Progress of Scheduled Jobs:** Implement monitoring to track the progress of ETL jobs and ensure successful completion.
* **Alert System:** Set up an alert system to notify users via email or SMS when scheduled ETL jobs fail or experience delays.

## 4.6. Analytics

**SSAS Tabular Model**

* **Build SSAS Tabular Model:** Create a tabular model in SQL Server Analysis Services (SSAS) for advanced analytical processing.
* The model will allow business users to slice and dice data based on dimensions like **Product**, **Customer**, **Supplier**, and **Time**.

**Power BI Dashboard**

* **Create a Power BI Dashboard:** Design a Power BI dashboard to visualize key metrics, including sales performance, customer orders, and supplier efficiency.
* The dashboard will support drill-downs for detailed analysis and customizable reporting options for end-users.

# 5. Non-Functional Requirements

* **Performance:** Ensure the data warehouse supports fast query execution, even with increasing data volumes.
* **Scalability:** Design the data warehouse and ETL processes to handle future growth.
* **Data Security:** Protect sensitive data (e.g., customer and employee data) and ensure that only authorized users have access.
* **High Availability:** Ensure minimal downtime for the data warehouse and the automation of ETL jobs.

# 6. Stakeholders

* **Digital Egypt Pioneers Initiative, NTI:** Requested the Northwind Traders data warehouse project and will be monitoring its progress to ensure alignment with strategic goals.

# 7. Assumptions

* The Northwind dataset is reliable and comprehensive.
* The project team has access to all required tools (SQL Server, SSIS, SSAS, Power BI).
* The infrastructure for the data warehouse is in place and supports scaling.

# 8. Constraints

* The project must be completed within the allocated time.
* The BI tool chosen must align with the existing technical ecosystem.

# 9. Risks

* **Data Quality Issues:** Poor data quality might require additional cleaning.
* **System Performance:** Inadequate infrastructure could impact performance.
* **Job Failures:** If scheduled ETL jobs fail, data could be delayed.

# 10. Deliverables

* **Data Warehouse Design:** Schema and architectural design.
* **ETL Processes:** Fully functional ETL pipelines.
* **SSAS Tabular Model:** Analytical model for data slicing.
* **Power BI Dashboard:** Interactive reports and dashboards for end-users.
* **Documentation:** User guides and technical documentation.

# 11. Timeline and Milestones

| **Milestone** | **Target Date** |
| --- | --- |
| Project Kickoff | September 30, 2024 |
| Data Modeling Completion | October 1, 2024 |
| ETL Process Setup | October 3, 2024 |
| SSAS Tabular Model Creation | October 3, 2024 |
| Power BI Dashboard Completion | October 4, 2024 |
| Final Delivery & Go-Live | October 5, 2024 |