



Requirements for Power BI Data and Reporting Automation

Version History:

Version	Approved By	Revision Date	Description of Change	Author
1.1		6/5/2020	Requirements creation	Andray Gorban

Table of Contents

1. Stakeholders	3
2. Purpose.....	3
3. General	3
4. Current vs. Future State Comparison	5
5. Assumptions.....	6
6. Risk and Mitigation	6
7. Constraints	8
8. Functional Requirements	10
9. Non-Functional Requirements	12
10. Appendices	12

Section 1 Stakeholders

Name	Role	Contact
Patrick Traphagen	IT Manager	Ptraphagen@customseng.com
Mike Whelan	Sales Manager	Mwhelan@customseng.com
Andray Gorban	IT Intern	agorban@customseng.com

Section 2 Purpose

The objective of creating a Power BI dashboard is to provide a centralized and interactive platform for measuring and analyzing various metrics that are critical to organizational performance. By consolidating data from multiple sources into a visually dynamic interface, the dashboard enables stakeholders to monitor key performance indicators (KPIs), identify trends, and gain actionable insights in real-time. This tool empowers decision-makers to make data-driven decisions, optimize processes, and track progress toward business goals, ensuring improved efficiency and accountability across all levels of the organization.

Section 3 General

3.3 General

- Project number: P164502
- What technologies will be used to consume data: SQL, PBI, Python
- The reports will be generated using data sourced from:

- CE_DCPRESHUB01 and CE_INFOR03A
- Is the raw data needed or an aggregate view: No
- When is data expected to be available for consumption (Daily, weekly, monthly, ect): Daily
- What is the specific time of day that data will need to be available: 8 am and 5pm

3.2 Inscope:

- Data Integration and Preparation:
 - Consolidate and standardize data from key databases to create a unified, reliable data model for reporting and analysis.
- Dashboard Development:
 - Design and implement interactive Power BI dashboards to provide insights into key operational areas, such as efficiency, performance, quality, and costs.
- Reporting Automation:
 - Automate reporting processes to ensure timely and accurate data updates while minimizing manual effort.
- Security and Access Management:
 - Establish secure, role-based access controls to protect data and ensure appropriate visibility for stakeholders.
- Testing and Validation:
 - Validate the accuracy and performance of dashboards to ensure they meet organizational needs and handle large datasets effectively.
- User Enablement:
 - Provide training and resources to empower stakeholders to effectively use and derive value from the dashboards.

3.3 Out of Scope:

- Modifications to the ERP System:
 - Any changes to the ERP system or its data generation processes.
- Non-ERP Data Sources:
 - Integration of non-ERP data sources or third-party applications.
- IT Infrastructure Setup:
 - The setup of Power BI licensing, network permissions, or server infrastructure is assumed to be already in place.

Section 4 Current vs. Future State Comparison

Feature	Current State	Future State
Reporting Tools	Static Crystal Reports with predefined templates.	Interactive Power BI dashboards.
Types of Reports	<p>Reports include:</p> <ul style="list-style-type: none"> • Indirect Time Report by Employee (tracks non-production activities). • Work Orders by Employee (tracks work orders completed). • Machinist Utilization (measures idle time and bottlenecks). • Revised Efficiency (evaluates performance trend recalculations). 	<p>Reports will include:</p> <ul style="list-style-type: none"> • Machinist Job Efficiency (real-time tracking of productivity and task deviations). • Production and OOE (Overall Equipment Effectiveness) metrics (availability, performance, quality). • Operational Quality (defect rates, rework percentages, trends).
Report Refreshes	Reports are generated individually for each dataset and manually compiled in spreadsheets.	Automated report refreshes in Power BI.
Data integrity Issues	<p>The ERP system exports data into three separate tables: Labor Ticket, and Work Order. Merging the tables required manual work to address issues such as:</p> <ul style="list-style-type: none"> • No direct match for ID in the Labor Ticket table. • Estimated hours and cost calculations had to be figured out manually. • Joins created incorrect matches due to multiple labor tickets per operation. • Fields imported as strings instead of numbers, causing conversion errors and rounding issues. • Merged tables had missing decimal precision, creating zero values that skewed efficiency calculations. 	<p>Direct merging of ERP data in Power BI without manual intervention. Power BI's data model and transformation capabilities will handle table joins and ensure correct data types. Queries and measures will maintain decimal precision, avoiding zero-value errors and manual fixes.</p>
Report Timeframe	Monthly reports created from scratch.	Historical and real-time data available.
Data Access	Disconnected files for different time periods.	Centralized access to all data via Power BI dashboards.
Report Customization	Minimal customization due to reliance on static reporting templates and predefined structures.	Customizable dashboards with filters and slicers (e.g., Site, Shift, Employee, Work Order Type) for tailored reporting.

Section 5 Assumptions

- **Data Accessibility:** The necessary data from the CE_DCPRESHUB01 and CE_INFOR03A databases is readily accessible and up to date.
- **Licensing and Tools:** All relevant stakeholders have access to Power BI licenses and any necessary permissions to view and interact with the dashboards.
- **Stakeholder Readiness:** Key stakeholders and end-users will be available for feedback, testing, and training as needed throughout the project. The team has the foundational knowledge and skills required to adopt Power BI and related processes efficiently.
- **Data Structure:** The structure of the data will remain consistent during the project to avoid disruptions in data integration and report accuracy.
- **System Infrastructure:** The IT infrastructure (including database connections and network settings) is already in place to support the data refresh and dashboard access.
- **Security Compliance:** Role-based access and data sharing within Power BI will comply with organizational security policies.
- **Device Compatibility:** The dashboards will be accessed primarily on standard web browsers and mobile devices without requiring special configurations.
- **Training Resources:** A single training session or brief documentation will be sufficient for most users to understand how to navigate and use the dashboards, documentation will be shared, and the session will be recorded.

Section 6 Risk and Mitigation

Process/Task	Risk	Mitigation
Automate Data Extraction and ETL Transformation	<ul style="list-style-type: none">○ Data quality issues in the database tables, such as missing values or inconsistent formats, may cause errors in ETL processes.○ Failure in automated scripts or Power Automate configurations could lead to outdated or incomplete data.	<ul style="list-style-type: none">○ Perform an initial data quality assessment and clean the data before starting the automation process.○ Implement robust error-handling mechanisms in Python scripts and Power Automate workflows to log and notify errors during ETL.○ Test ETL processes on small datasets before scaling to full production.
Automate Power BI Report Refreshes	<ul style="list-style-type: none">○ Power BI gateways may encounter configuration or connectivity issues, disrupting the refresh process.	<ul style="list-style-type: none">○ Set up redundancy in gateway configurations to avoid a single point of failure.

	<ul style="list-style-type: none"> ○ High data volume could slow down or fail data refreshes. 	<ul style="list-style-type: none"> ○ Optimize data models and queries in Power BI to handle large datasets efficiently. ○ Monitor refresh logs to detect and resolve refresh issues proactively.
Develop Customizable Power BI Dashboards	<ul style="list-style-type: none"> ○ Stakeholders may request excessive customization, leading to scope creep and delays. ○ Dashboards might not meet user expectations due to unclear requirements. 	<ul style="list-style-type: none"> ○ Prioritize features based on stakeholder input and align expectations early in the project. ○ Use an iterative approach to gather feedback and refine dashboards incrementally. ○ Maintain clear documentation of scope and customization limits.
Non-Functional Performance Requirements	<ul style="list-style-type: none"> ○ Dashboards may not meet the performance goal of loading within 5 seconds for large datasets. 	<ul style="list-style-type: none"> ○ Use indexing, aggregation, and optimized queries to improve performance. ○ Leverage Power BI's built-in performance monitoring tools to identify bottlenecks. ○ Use Power BI premium features for larger data models if necessary.
Data Security	<ul style="list-style-type: none"> ○ Misconfigured role-based access control (RBAC) may expose sensitive data to unauthorized users. 	<ul style="list-style-type: none"> ○ Perform regular security audits and tests to validate RBAC configurations. ○ Train administrators on Power BI security best practices. ○ Implement data classification and encryption for sensitive information.
Usability and User Training	<ul style="list-style-type: none"> ○ Limited user training may result in low adoption or improper usage of dashboards. 	<ul style="list-style-type: none"> ○ Provide multiple training sessions tailored to different user roles. ○ Create comprehensive but concise user guides and video tutorials for self-learning. ○ Offer ongoing support for dashboard usage and troubleshooting.
Data Refresh Constraints	<ul style="list-style-type: none"> ○ Limited refresh frequency (twice daily) may not meet real- 	<ul style="list-style-type: none"> ○ Clearly communicate the refresh frequency to stakeholders and set expectations.

	time reporting needs for some stakeholders.	<ul style="list-style-type: none">○ Explore the possibility of increasing refresh frequency if infrastructure permits.
Scalability	<ul style="list-style-type: none">○ Adding new data sources or metrics might require reworking existing dashboards and data models.	<ul style="list-style-type: none">○ Design dashboards with scalability in mind, using modular data models and flexible queries.○ Test integration of dummy data sources during the initial phase to ensure compatibility.
Project Timeline	Delays in data access, stakeholder feedback, or training sessions could affect project deadlines.	<ul style="list-style-type: none">○ Develop a detailed project plan with buffer time for unforeseen delays.○ Schedule regular progress updates and ensure active stakeholder engagement.○ Escalate unresolved dependencies early to avoid timeline disruptions.

Section 7 Constraints

- **Data Refresh Frequency:** Data extraction and ETL processes are limited to twice daily (8:00 AM and 5:00 PM).
 - Python job gathering machine data via pyserial runs twice daily (8:00 AM and 5:00 PM).
 - Real-time updates are not available outside these scheduled refresh intervals.
- **Data Source Dependence:** The project relies on consistent and accurate data from database tables.
 - Any inconsistencies or changes in data structures may disrupt the ETL process and dashboard accuracy.
- **Licensing:** Power BI licenses and necessary permissions must be secured for all stakeholders accessing the dashboards.
 - Lack of licenses could limit dashboard accessibility for intended users.
- **Gateway Configuration:** Power BI refreshes depend on properly configured data gateways for automation.
 - Incorrect or delayed gateway configuration could disrupt scheduled data updates.

- **ETL Automation Tools:** Python and Power Automate are the designated tools for ETL automation, with no scope for additional software integration.
 - Limits flexibility to explore alternative ETL tools during the project.
- **IT Infrastructure:** The project assumes existing infrastructure supports Python, Power Automate, and Power BI integration.
 - Upgrading or reconfiguring the infrastructure could delay the implementation.
- **User Training:** End-user training is limited to a single session or minimal documentation due to time constraints.
 - Users with limited technical knowledge may require additional support post-implementation.
- **Dashboard Scalability:** The scope covers dashboards for existing tables; adding new data sources or metrics requires separate approval.
 - Expansion of scope could impact the project timeline and require additional resources.
- **Data Security Policies:** Role-based access control must comply with organizational security and compliance standards.
 - Delays in aligning security configurations with compliance standards could delay deployment.
- **Project Timeline:** The implementation must adhere to the predefined schedule, with minimal allowances for scope or timeline changes.
 - Delays in data access, stakeholder feedback, or testing may impact the overall timeline.

Section 8 Functional Requirements

ID	Requirement	Priority	Success Criteria
1	<p>Automate data extraction and ETL transformation using python and power automate:</p> <ul style="list-style-type: none"> Consolidating data from the following tables: <ul style="list-style-type: none"> Labor Ticket Work Order Inventory Operation Tables will be accessed via: <ul style="list-style-type: none"> CE_DCPRESHUB01 and CE_INFOR03A databases Data must be extracted and processed twice daily, at 8:00 AM and 5:00 PM, ensuring timely updates for stakeholders. All ETL processes must be fully automated to guarantee data consistency, accuracy, and reliability across all reports. 	High	ETL processes and Power BI refreshes are completed automatically and error-free twice daily, ensuring stakeholders access up-to-date and accurate data.
2	<p>Automate Power BI report refreshes:</p> <ul style="list-style-type: none"> Power BI dashboards will be refreshed twice daily via configured on premise gateway to ensure users have access to the latest data without manual intervention. 	Medium	Reports are generated automatically in Power BI without manual intervention and contain up-to-date data.
3	<p>Develop customizable Power BI dashboards with interactive slicers for date ranges, sites, and employee names, allowing users to analyze machinist and plant productivity, operational quality, time utilization, and manufacturing costs.</p> <p>The dashboards must include the following key metrics:</p> <ul style="list-style-type: none"> Plant and Machinist Utilization: Measure overall productivity and identify idle time or inefficiencies. Machinist Performance Splits: Breakdown of indirect time, setup, and run time to assess task-level efficiency. Overall Operational Effectiveness (OOE): Metrics for availability, performance, and quality to evaluate plant efficiency. Scrap Rate: Percentage of material lost during production. 	High	Users can filter data using slicers for dates, sites, and employee names in Power BI dashboards.

	<ul style="list-style-type: none">• First Pass Yield: Percentage of products meeting quality standards without rework (goal: 99.75%).• Lead Time: Average time to complete work orders, including setup and production. <p>The dashboards should enable users to drill down into detailed insights, filter data dynamically, and visualize trends to support data-driven decision-making.</p>		
--	--	--	--

Section 9 Non-Functional Requirements

ID	Requirement	Priority	Success criteria
1	Performance: Dashboards must load within 8 seconds for queries involving up to 1 million records.	High	Dashboards and reports load quickly (< 5 seconds) even with large datasets during testing.
2	Scalability: The system should support the addition of new data sources without requiring major rework.	High	Additional data sources (e.g., new company sites) can be integrated with minimal configuration in Power BI.
3	Data Security: Implement role-based access control to restrict sensitive data access.	High	Different user roles (e.g., manager, operator) have appropriate access to data based on their permissions.
4	Availability: Ensure 99.9% system availability during business hours.	Medium	Dashboards are available and accessible during regular working hours with minimal downtime or disruptions.
5	Maintainability: System updates and changes must be completed without breaking existing reports.	High	Changes to data schemas or report layouts do not result in errors or data loss.
6	Usability: Dashboards should be intuitive for end users with minimal training required (less than 1 day).	Medium	End users can navigate and interpret dashboards effectively after a brief training session.
7	Portability: Dashboards should be accessible from desktops, tablets, and smartphones.	Medium	Reports and dashboards render correctly on different devices and screen resolutions without performance degradation.
9	Auditability: Logs must track when reports are generated and by whom for auditing purposes.	Medium	Access and changes to reports are logged, with the ability to generate audit reports when needed.
10	Interoperability: The solution must work seamlessly with Microsoft SQL Server, Access, and Power Query.	High	Reports integrate smoothly with SQL databases and other Microsoft tools without connection or compatibility issues.

Section 10 Appendices

- **Hardware:** According to Microsoft documentation, the minimum requirements for a Microsoft "on-premises data gateway" include:
 - 64-bit version of Windows 10 or Windows Server 2019
 - .NET Framework 4.8
 - At least 4GB of disk space for performance monitoring logs
 - 8-core CPU and 8GB of RAM

- **Software:** Users must have Power BI 2022 and Python installed with the pandas library. For python the IDE doesn't matter unless troubleshooting, in that case PyCharm was used.