ANALYTICS REPORT

ABDELRAHMAN YAHIA
Data Analyst

Supplier Quality Analysis
EYouth X DEPI Tech Challenge
03.03.2025

Table of Contents

Introduction

About the dataset:

Description Page: This industry sample dataset focuses on one of the typical supply chain challenges: supplier quality analysis. Two primary metrics are at play in this analysis: total number of defects and the total downtime that these defects caused.

Supplier quality analysis is a critical aspect of supply chain management. It focuses on evaluating and monitoring suppliers' performance to ensure consistent product quality and timely deliveries. Effective supplier analysis helps businesses reduce defects, enhance operational efficiency, and improve customer satisfaction. By leveraging data analytics and visualization tools like Power BI, companies can gain actionable insights into supplier performance, detect patterns, and implement improvements.

1. Category Classification

- Includes **6 distinct categories**, with a corresponding **Subcategory and Subcategory ID** for classification.

2. Defect Information

- **Defect Table:** Contains **272 distinct defect types**, with **245 unique entries** and **305 total records**.
- **Defect Report:** The core dataset, comprising **6,145 records** across **10 columns**, including:
    - **Defect Quantity (0 – 487,008):** Total defective units recorded.
    - **Downtime (0 – 999 minutes):** Production time lost due to defects.
    - **7 ID Columns:** Unique identifiers for defect tracking, with **one Material ID unrelated to dimension tables**.

3. Defect Type Classification

Categorizes defects into **three impact levels: (No Impact, Impact & Rejected)**

4. Material, Plant & Supplier Data

- **Material Type:** Includes **22 unique materials** tracked for defects.
- **Plant Locations:** Covers **24 distinct production sites** where materials were used.
- **Vendors:** Consists of **326 distinct supplier entries** (**324 unique vendors across 328 total records**).

Business Questions

Defect Analysis

Which defects and materials drive the **highest failure rates**?

Which material categories are **most prone to defects**?

Supplier Performance

Which suppliers have the **highest defect rates and downtime impact**?

How do suppliers **rank** in **defect trends over time**?

Are certain suppliers **consistently underperforming** across materials and plants?

Plant Efficiency

Which plants **struggle with the highest defect rates and downtime**?

Are defect rates **higher at specific plant locations**?

Operational Impact

What is the **financial cost** of supplier defects and downtime?

Is there a **link between high downtime and specific suppliers or materials**?

Data Cleaning Documentation

Overview

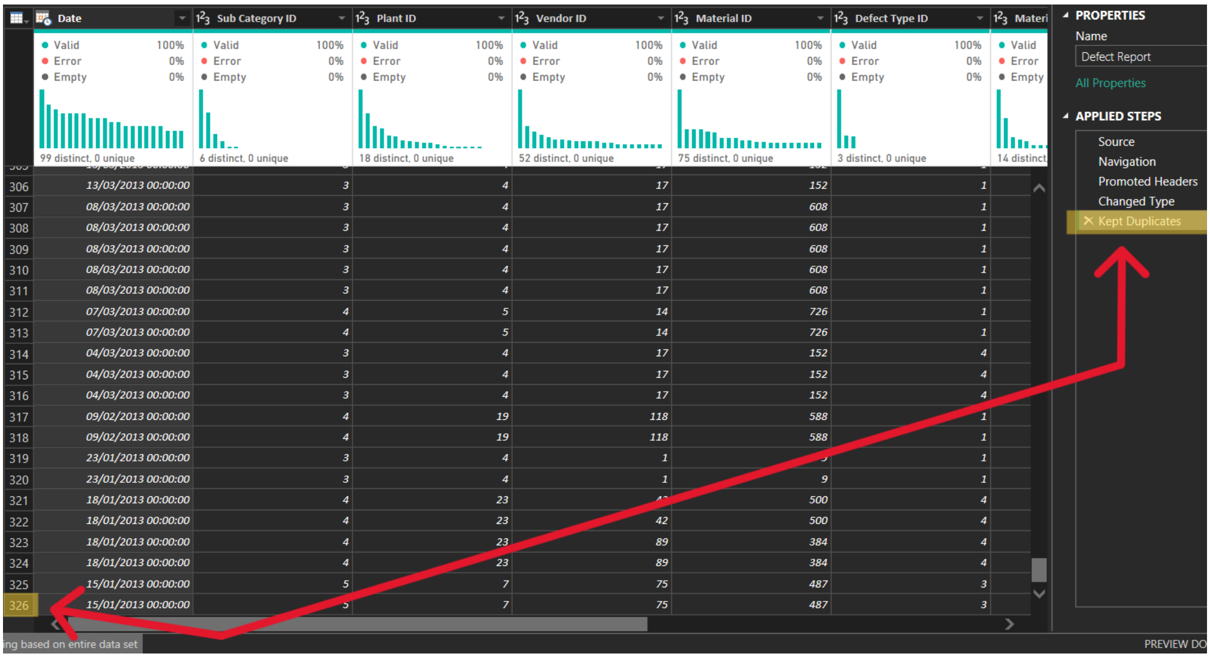
Data cleaning is a crucial step in ensuring accurate analysis and reporting. This document outlines the **cleaning steps, issues encountered, and solutions applied** to prepare the dataset for analysis.

Issues Identified & Cleaning Steps

1. Removing Duplicate Records in Fact Table

Issue:

- The Fact Defect Report table contained **326 duplicate rows** (identical across all columns).



Solution:

- Used **"Keep Duplicates"** to identify duplicate records.
- Found **pure duplicates** (same values in all columns, including timestamps).
- Removed duplicates to ensure **accurate defect counts**.

Impact:

- The total defect count dropped from **56M to 55M**, confirming that duplicate records were inflating values.

2. Handling Redundant Columns

Issue:

- The Sub Category column in **both Fact and Dimension tables** was an **exact duplicate** of Category.

Solution:

- Removed the Sub Category column.
- Renamed Sub Category ID to **Category ID** in both **Fact and Dim tables** for consistency.

Impact:

- **Improved data model efficiency** and reduced unnecessary columns.

---

3. Referential Integrity Issue (Orphaned Records in Fact Table)

Issue:

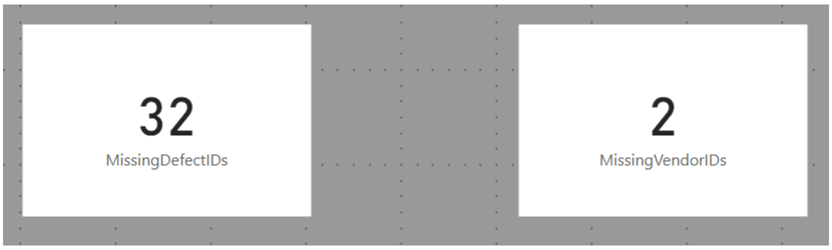
- When removing duplicates in Dim Defect and Dim Vendor, their **corresponding IDs (Defect ID, Vendor ID) were also removed**.
- This created **32 missing Defect IDs** and **2 missing Vendor IDs** in the Fact Defect Report, causing **referential integrity issues** (fact table pointing to missing dimension records).

Root Cause:

- The **same Defect/Vendor name** had multiple unique IDs (e.g., "Vendor A" had IDs **2, 4, and 81**).
- Removing duplicates **removed some IDs still referenced in the Fact Table**, causing null values.

```
MissingDefectIDs =
VAR UnmatchedDefects =
    EXCEPT(
        DISTINCT('Defect Report'[Defect ID]),
        DISTINCT('Defect'[Defect ID])
    )
RETURN
    COUNTROWS(UnmatchedDefects)
```

```
MissingVendorIDs =
VAR UnmatchedVendors =
    EXCEPT(
        DISTINCT('Defect Report'[Vendor ID]),
        DISTINCT('Vendor'[Vendor ID])
    )
RETURN
    COUNTROWS(UnmatchedVendors)
```



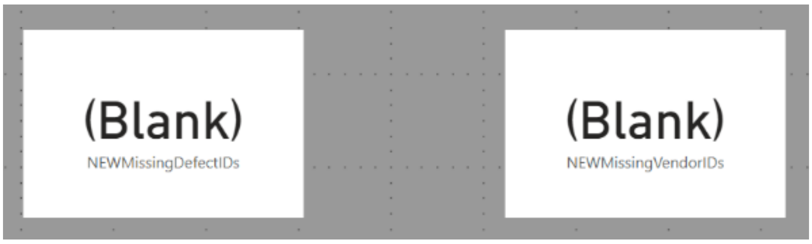
Solution: Rebuilding Dimension Tables with Consolidated IDs

Step-by-step fix:

1) **Merged Defect ID in Fact Defect Report with Defect ID in Dim Defect** to pull in defect names.
2) **Expanded only the Defect column** (ignoring duplicate IDs for now).
3) **Created "New Dim Defect" Table** from the fact table:
    - Kept only Defect and Defect ID columns.
    - Removed duplicates, ensuring **each defect has a single unique ID**.
4) **Updated the Fact Table:**
    - Merged Defect in Fact Defect Report with Defect in New Dim Defect.
    - Expanded only the **New Defect ID** column.
    - Removed the **old Defect ID** column.

```
NEWMissingDefectIDs =
VAR UnmatchedDefects =
    EXCEPT(
        DISTINCT('Defect Report'[New Defect ID]),
        DISTINCT('Defect'[New Defect ID])
    )
RETURN
    COUNTROWS(UnmatchedDefects)
```

```
NEWMissingVendorIDs =
VAR UnmatchedVendors =
    EXCEPT(
        DISTINCT('Defect Report'[New Vendor ID]),
        DISTINCT('Vendor'[New Vendor ID])
    )
RETURN
    COUNTROWS(UnmatchedVendors)
```



Impact:

- **Preserved referential integrity** while eliminating duplicate IDs.
- Ensured every Defect in the Fact Table **maps to a valid Defect ID** in the Dimension Table.

4. Standardizing Data Types

Issue:

- Inconsistent data types for ID columns across tables:
    - Some IDs were stored as **whole numbers**, others as **text**.
    - This could cause **relationship mismatches** in Power BI.

Solution:

- Converted **all ID columns** to **Text** for consistency.
- **Exception:** Kept **old Vendor ID and old Defect ID** as Whole Numbers in **Old Dim tables** for tracking.

Impact:

- **Avoided data type mismatches in relationships** between Fact and Dimension tables.

---

5. Renaming & Organizing Tables

Final Table Names:

Old Table Name	New Table Name
Defect Report	Fact Defect Report
Category	Dim Category
Defect	Dim Defect
Defect Type	Dim Defect Type
Material Type	Dim Material Type
Plant	Dim Plant
Vendor	Dim Vendor
Defect Report (2)	New Dim Defect
Defect Report (3)	New Dim Vendor

Impact:

- **Clearer naming** ensures better readability and usability in Power BI.

Summary of Data Cleaning Impact

Issue	Solution Applied	Impact on Data
Duplicate records in Fact Table	Removed 326 duplicates	Defect count corrected (from 56M → 55M)
Redundant Sub-Category Column	Removed & renamed IDs	Improved model efficiency
Referential Integrity Issue	Rebuilt Dim Defect & Dim Vendor with consolidated IDs	Eliminated orphaned records (32 missing Defect IDs, 2 Vendor IDs fixed)
Inconsistent Data Types	Converted all IDs to text (except old IDs)	Prevented relationship mismatches
Unclear Table Names	Renamed tables for clarity	Improved readability

Data Modelling Documentation

Overview

The data model is designed using a **Star Schema** approach to ensure **efficient data retrieval, optimized relationships, and high-performance reporting** in Power BI. This section outlines the **structure, efficiency optimizations, and rationale** behind the chosen model.

Data Model Structure

1. Creating a Date Table (Dim Calendar)

Since time-based analysis is crucial, I created a **DAX-generated Date Table**:

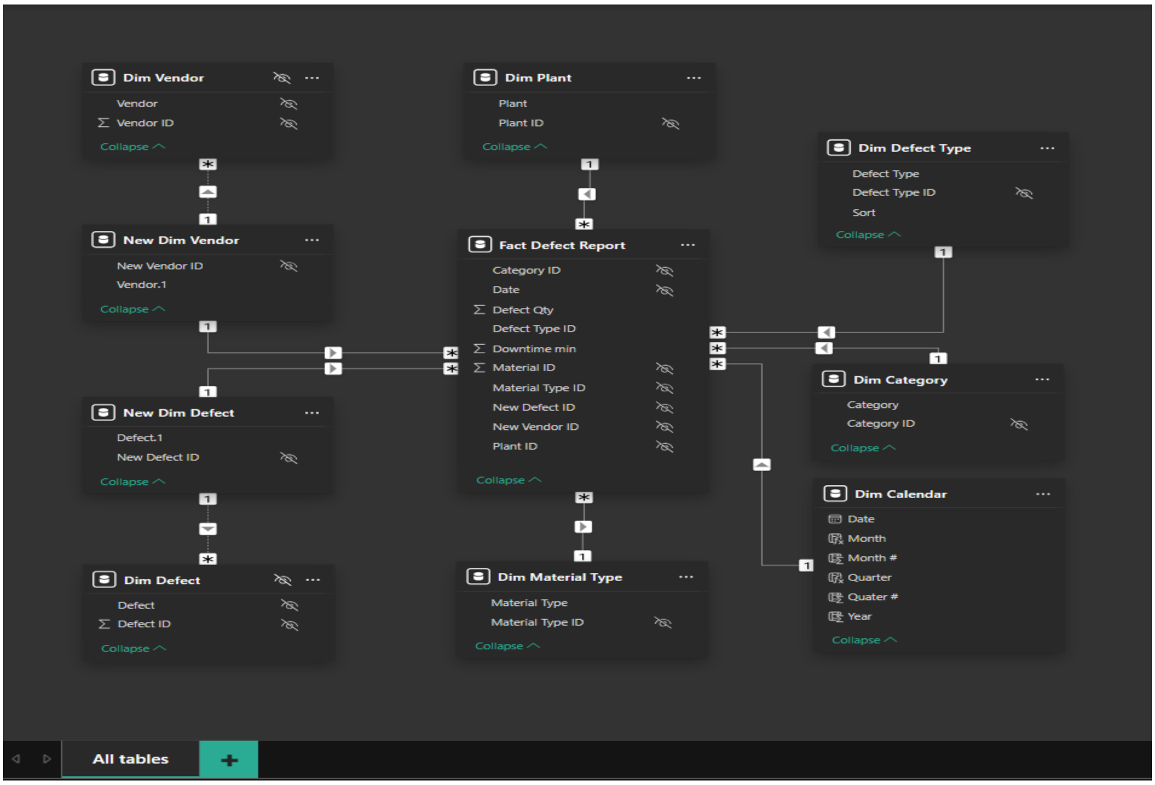
```
Dim Calendar =
ADDCOLUMNS(
    CALENDARAUTO(),
    "Year", YEAR([Date]),
    "Quarter", "Q" & FORMAT([Date], "Q"),
    "Quarter #", QUARTER([Date]),
    "Month", FORMAT([Date], "mmm"),
    "Month #", MONTH([Date]),
    "Day", FORMAT([Date], "dddd"),
    "Day #", WEEKDAY([Date])
)
```

- **Marked Dim Calendar[Date] as the official Date Table** to ensure accurate **Time Intelligence DAX functions like** `SAMEPERIODLASTYEAR()`, `TOTALYTD()`, etc.
- **Manually created a Date Hierarchy**:
    - **Hierarchy:** Year → Quarter → Month → Day
    - **Sorting applied:** Month sorted by Month #, Quarter sorted by Quarter # to ensure proper ordering.

Why? Ensures that **date-based calculations work properly** while giving users easy access to drill-down reports.

2. Implementing a Star Schema Data Model

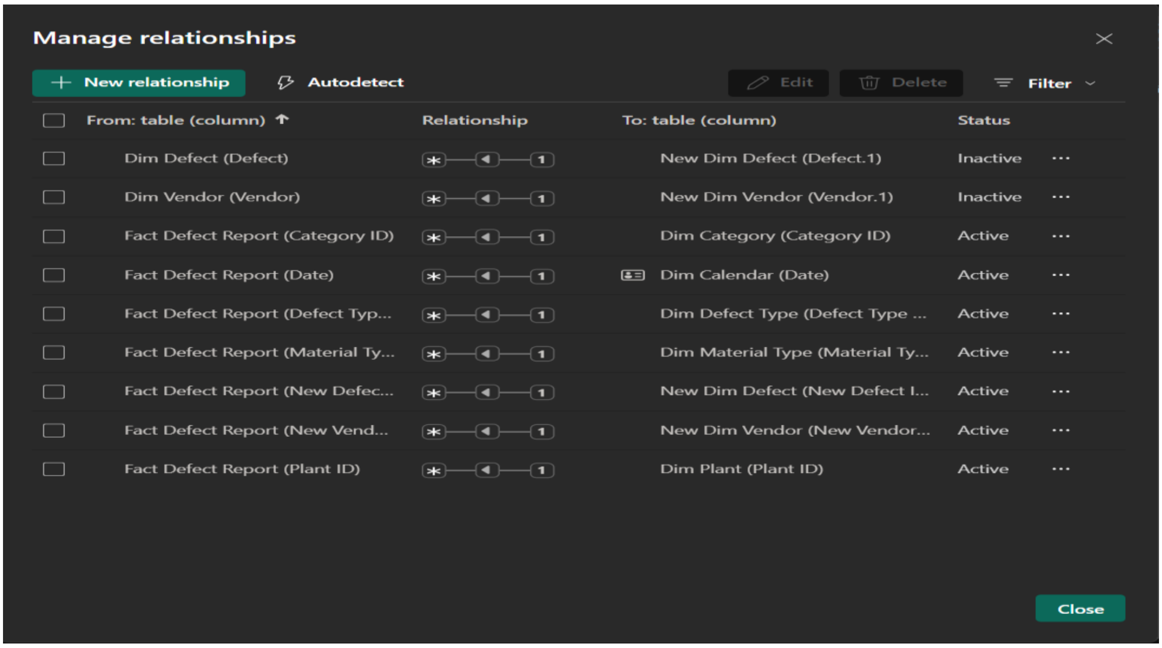
The **Star Schema** was chosen for **better performance, maintainability, and scalability**.



---

3. Relationship Design (One-to-Many)

- **Fact Table (Fact Defect Report)** is at the **center** of the model, linking to **Dimension Tables** through **one-to-many (1:*) relationships**.



Why? This structure follows **best practices for Power BI modeling**, enabling **fast queries, easy filtering, and accurate aggregations**.

4. Performance Optimization

1. Hide Unnecessary Columns & Tables

2. Used Proper Data Types

- Converted **all ID columns** to **Text** for consistency in relationships, also since mathematical operations are not performed on IDs, there is no need to store them as numeric data types.
- Numeric fields like **Defect Qty & Downtime (Minutes)** were kept as **Whole Numbers** for **accurate aggregations**.

3. Optimized Date Table for Fast Time Intelligence

- Marked **Dim Calendar as the official Date Table** to avoid **auto-generated date tables**, which slow performance, which impacts on faster **Time Intelligence calculations (YTD, YoY, Rolling Averages, etc.)**.

Summary of Data Modelling Impact

Optimization	Impact on Performance & Usability
Star Schema Design	Faster queries, simpler relationships
Created Dim Calendar Table	Better time intelligence & trend analysis
Removed Unnecessary Columns	Reduced model size & improved efficiency
Converted IDs to Text for Consistency	Prevented relationship mismatches
Used Single-Directional Filters	Optimized cross-filtering performance
Marked Date Table	Enabled advanced DAX date functions

DAX Documentation

Overview

This document provides a detailed explanation of the **key DAX measures** used across different report pages. Each measure is documented based on **its purpose, logic, and performance optimization**.

Key DAX Measures

1. Total Defect Quantity

Purpose: Calculates the total number of defects across all records.

```
Total Defect Qty = SUM('Fact Defect Report'[Defect Qty])
```

Why It's Important?

- Used in KPI cards to show **overall defect count**.
- A fundamental measure for other calculations like **YoY changes and rankings**.

---

2. Previous Year Defect Quantity

Purpose: Retrieves the defect quantity from the previous year for **Year-over-Year (YoY) analysis**.

```
Previous Year Defect Qty =
CALCULATE(
    SUM('Fact Defect Report'[Defect Qty]),
    SAMEPERIODLASTYEAR('Dim Calendar'[Date])
)
```

Why It's Important?

- Helps compare defect performance **across years**.
- Supports trend analysis and YoY percentage changes.

---

3. Defect YoY Change (%)

Purpose: Calculates the percentage change in defect quantity compared to the previous year.

```
Defect YoY Change % =
VAR CurrentYearDefects = SUM('Fact Defect Report'[Defect Qty])
VAR PreviousYearDefects =
    CALCULATE(
        SUM('Fact Defect Report'[Defect Qty]),
        SAMEPERIODLASTYEAR('Dim Calendar'[Date])
    )
RETURN
IF(
    ISBLANK(PreviousYearDefects),
    BLANK(),
    DIVIDE(CurrentYearDefects - PreviousYearDefects, PreviousYearDefects) * 100
)
```

Why It's Important?

- Displays **how defect count has changed** over time.
- Ensures **null handling** to avoid errors when there's no previous year data.

---

4. Defect Status Indicator

Purpose: Dynamically assigns a **status label** (🔴 Increase, 🟢 Decrease, 🟡 No Change) based on defect trends.

```
Defect Status =
VAR Change =
    DIVIDE(
        SUM('Fact Defect Report'[Defect Qty]) -
        CALCULATE(SUM('Fact Defect Report'[Defect Qty]), SAMEPERIODLASTYEAR('Dim Calendar'[Date])),
        CALCULATE(SUM('Fact Defect Report'[Defect Qty]), SAMEPERIODLASTYEAR('Dim Calendar'[Date]))
    )
RETURN
SWITCH(
    TRUE(),
    ISBLANK(Change), "No Data",
    Change > 0, "🔴 Increase",
    Change < 0, "🟢 Decrease",
    "🟡 No Change"
)
```

Why It's Important?

- **Enhances storytelling** by visually indicating defect trends.
- **Switch function optimizes performance** by evaluating only the necessary condition.

---

5. Total Loss Calculation

Purpose: Estimates the **total financial impact** of defects based on a selectable minutely loss value.

```
Total Loss = [Total Downtime Min] * Loss[Loss Value]
```

Why It's Important?

- Converts **downtime into monetary loss**.
- Provides a **business impact perspective** on defect performance.

(The next formulas apply similarly to the **Plant Performance Page** as well.)

6. Average Defect Quantity Per Vendor

Purpose: Calculates the average number of defects per vendor.

```
Avg Defect Qty Per Vendor =
DIVIDE(
    SUM('Fact Defect Report'[Defect Qty]),
    DISTINCTCOUNT('Dim Vendor'[Vendor ID])
)
```

Why It's Important?

- Helps compare **supplier quality performance**.
- Ensures division by zero is handled properly.

---

7. Average Downtime Per Vendor

Purpose: Calculates the **average downtime per vendor** to assess supplier reliability.

```
Avg Downtime Per Vendor =
DIVIDE(
    SUM('Fact Defect Report'[Downtime min]),
    DISTINCTCOUNT('Dim Vendor'[Vendor ID])
)
```

Why It's Important?

- Helps identify **vendors causing significant production delays**.
- Used in visualizations to compare vendor efficiency.

---

8. Top Vendor (Highest Defect Count)

Purpose: Identifies the **vendor with the highest defect count dynamically**.

```
Top Vendor =
TOPN(1, VALUES('New Dim Vendor'[Vendor.1]),
    CALCULATE(SUM('Fact Defect Report'[Defect Qty])), DESC)
```

Why It's Important?

- Highlights the **worst-performing vendor** automatically.
- Used in KPI visuals and drill-through reports.

---

9. Top Vendor by Downtime

Purpose: Identifies the vendor causing the **most downtime** due to defects.

```
Top Vendor by DTime =
TOPN(1, VALUES('New Dim Vendor'[Vendor.1]),
    CALCULATE(SUM('Fact Defect Report'[Downtime min])), DESC)

```

Why It's Important?

- Helps in **supplier negotiations & quality control actions**.
- Used in decision-making for **supplier audits**.

Summary of DAX Measures & Their Impact

Measure Name	Purpose	Optimization Strategy
Total Defect Qty	Aggregates total defects	Simple SUM(), ensuring high performance
Previous Year Defect Qty	Fetches last year's defect data	Uses SAMEPERIODLASTYEAR() for efficiency
Defect YoY Change %	Calculates YoY change	Uses VAR for cleaner, optimized DAX
Defect Status Indicator	Assigns trend labels (🔴 🟢 🟡)	SWITCH() avoids redundant calculations
Total Loss	Calculates financial impact	Uses parameter selection for user control
Avg Defect Qty Per Vendor	Measures supplier defect rate	DIVIDE() prevents division errors
Avg Downtime Per Vendor	Measures supplier downtime impact	Ensures smooth aggregation
Top Vendor	Identifies worst vendor	Uses TOPN() for ranking
Top Vendor by Downtime	Identifies vendor causing most downtime	Uses TOPN() on Downtime min

Key Findings

Where Are We Losing the Most?   Quality & Downtime Insights

This next insight pages provides a **business-friendly summary** of key trends across all dashboard pages, identifying **risks, opportunities, and actionable recommendations** for stakeholders.

Executive Summary – High-Level Trends

Defect quantities **surged to 55 million units**, driving a **138,000-minute (2,300-hour) increase in downtime**, which in turn **led to an estimated financial impact of $1.037 million**. These findings highlight the **significant cost of defective materials on production efficiency** and operational performance.

1. Financial Loss

The analysis reveals a **direct correlation between downtime and financial losses**. With an estimated **downtime cost of $7.5 per minute**, the company **incurred over $1.037 million** in downtime-related expenses. Notably, **March and October experienced the highest spikes**, indicating **critical periods of operational disruption**. This financial insight underscores the **urgent need for supplier quality improvements** to mitigate revenue loss and enhance overall efficiency.

2. Quarterly Defect Trends

- **Q4 had the highest defect count** at **15.1 million**, followed by **Q2, Q3, and Q1**, indicating a **year-end quality decline**.
- **Defect quantity and maximum allowed quantity are negatively correlated**, meaning as **defects increase, quality compliance weakens**.
- **2014 Q4 alone contributed 17.84% of total defects**, making it a key period for investigation.

3. Plant Performance Insights

- **Detroit, MI recorded the highest defect quantity (6.6M defects)**, significantly **outpacing Davenport, IA (105K defects)**.
- **Springfield, IL accounted for 12.5% of total downtime**, the **highest among all plants**, signaling a **bottleneck in operations**.
- Most plants had **downtime below 5,764 minutes**, but Springfield far exceeded this benchmark.

4. Category Impact

- **Mechanical defects** accounted for **31.14% of total defects**, making them the **top contributor to quality issues**.
- **Logistics-related issues caused the most downtime (40.41%)**, indicating **supply chain inefficiencies** affecting operations.

5. Supplier Performance

- **Reddoit was responsible for 18.93% of total downtime**, the **highest among 318 vendors**, suggesting **supplier-driven production delays**.
- **Solholdings had the highest defect count (7.29% of all defects),** making it the **largest contributor to quality failures among vendors**.

6. Defect Analysis

- **The "Misc" defect category accounted for 8.40% of all defects**, making it the **most frequent defect type**.
- **"Bad Seams" defects were responsible for 11.77% of total downtime**, pointing to **material or manufacturing issues**.

7. Defect Type Impact

- **Rejected defects accounted for 35.54% of all defects**, highlighting **a high rejection rate across operations**.
- **Impact-type defects caused 75.71% of total downtime**, making them the **most disruptive defect type**.

Supplier Performance Analysis – Vendor Quality Risks

1. Key Findings:

- **Plustax (under the "Impact" defect type) was responsible for 13.03% of total defects**, requiring immediate supplier review.
- **Impact-type defects caused 75.71% of all downtime**, suggesting that **defect severity is a major issue, not just quantity**.
- **Plustax also contributed to 25.51% of downtime,** making it a **high-risk supplier**.

2. Business Actions Needed:

a) **Review Plustax and Solholdings as priority suppliers** due to high defect rates.
b) **Work with Reddoit to identify supply chain inefficiencies causing downtime.**
c) **Negotiate stronger quality control terms with vendors to reduce high-defect shipments.**

Plant Performance Analysis – High-Risk Locations

1. Key Findings:

- **Detroit, MI had the highest defect count (6.6M), while Springfield, IL had the highest downtime (17,296 minutes).**
- **Most plants had total downtime below 5,764 minutes**, but Springfield far exceeded this threshold.
- **Packaging defects accounted for the highest defect count at plant level.**
- **Logistics-related defects caused 40.41% of all downtime.**

2. Business Actions Needed:

a) **Implement additional quality controls at Detroit, MI to reduce defect rates.**
b) **Investigate Springfield, IL's downtime issues—whether it's machine failures or operational inefficiencies.**
c) **Streamline logistics operations to improve supply chain efficiency.**

Material Performance Analysis – Supply Chain Weaknesses

1. Key Findings:

- **Corrugate materials had the highest downtime impact (52,726 minutes), followed by Raw Materials and Carton.**
- **Raw Materials had the highest defect quantity (13.2M defects), while Wires had zero reported defects.**
- **Corrugate contributed to 38.11% of all downtime, making it a bottleneck material.**

2. Business Actions Needed:

a) Investigate why Corrugate materials cause the highest downtime—production, storage, or supplier-related?
b) **Optimize Raw Material usage to prevent defects early in the supply chain.**
c) **Analyze why Wires have zero defects—are they underreported, or is this a best-practice area?**

Overall Business Recommendations

---

1. Focus on High-Risk Suppliers & Plants:
   - **Plustax, Reddoit, and Solholdings require urgent quality reviews** to mitigate defects and downtime.
   - **Detroit, MI (highest defects) and Springfield, IL (highest downtime) need intervention to stabilize operations.**
2. Improve Supply Chain Efficiency:
   - **Logistics-related downtime is a major issue (40.41%)**, meaning supply chain improvements could significantly reduce production delays.
   - **Corrugate materials cause the highest downtime, requiring supplier audits or alternative sourcing.**
3. Data-Driven Quality Control:
   - **High rejection rates (35.54%) suggest quality control improvements at the manufacturing stage.**
   - **Impact-type defects cause 75.71% of downtime—focus on resolving these critical issues.**

Conclusion

Supplier Quality & Downtime Analysis Report

🔍 Key Takeaways

1. Defect Trends:

- **Q4 had the highest defect rates**, signaling potential year-end quality control challenges.
- **Defect quantity and compliance levels were negatively correlated**, meaning more defects reduce adherence to quality standards.

2. Supplier Performance:

- **Reddoit contributed to the highest downtime (18.93%)**, and **Solholdings had the highest defect count (7.29%)**, identifying them as **top risk suppliers**.
- **Impact-type defects caused 75.71% of downtime**, making them the **biggest disruptor to operations**.

3. Plant Efficiency:

- **Detroit, MI had the highest defect count (6.6M defects)**, while **Springfield, IL had the highest downtime (17,296 minutes)**.
- **Packaging defects were the most frequent issue, while logistics problems caused the most downtime**.

4. Material Impact:

- **Corrugate materials caused the highest downtime (52,726 minutes)**, suggesting issues in **supplier reliability or production processes**.
- **Raw Materials had the highest defect quantity (13.2M defects)**, warranting further investigation into **procurement and storage conditions**.

Next Steps

To improve supplier quality and reduce operational inefficiencies, we recommend the following:

1. Strengthen Supplier Quality Control

a) Conduct **audits and quality assessments** for **high-risk suppliers (Reddoit, Solholdings, Plustax)**.
b) Implement **stricter compliance requirements** to minimize defect shipments.
c) Establish a **vendor performance scoring system** to track improvement.

2. Optimize Plant Operations

a) Implement **targeted defect reduction programs** at **Detroit, MI and Springfield, IL**.
b) Address **logistics bottlenecks** to improve plant efficiency.
c) Deploy **real-time defect tracking** to catch issues earlier in production.

3. Reduce Downtime & Material-Related Delays

a) Review **Corrugate material supply chains** to identify delays and defects.
b) Introduce **preventive maintenance programs** to minimize machine downtime.
c) Implement **Lean Manufacturing techniques** to optimize resource use.

Final Thoughts

This **comprehensive analysis report** of supplier performance, defect trends, plant efficiency, and material impact has provided **critical insights** into the key drivers of **quality issues and operational inefficiencies**.

Centralizing procurement and performance data through Power BI will enable the management team to make more **informed decisions** regarding vendor performance, plant operations, and material selection.

After leveraging Power BI's advanced analytics, we identified **high-risk suppliers, defect-prone materials, and plants with excessive downtime,** allowing for targeted intervention strategies **to improve overall performance**. The analysis not only highlighted key performance gaps but has **also paved the way for long-term process improvements.**

By implementing these recommendations, the organization can **reduce downtime, lower financial losses, enhance supplier collaboration, and ultimately drive higher product quality and customer satisfaction**.