SUPPLIER QUALITY ANALYSIS
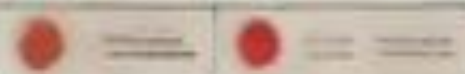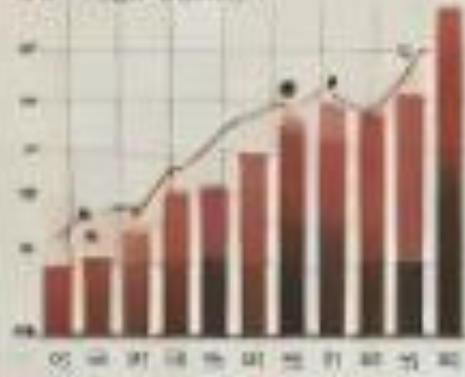
Suppliers Quality Analysis Final Project
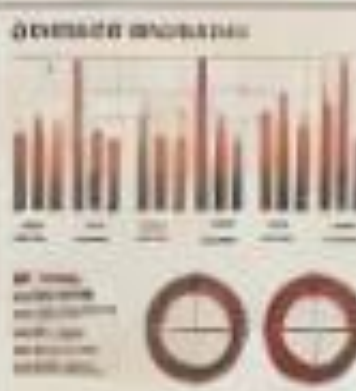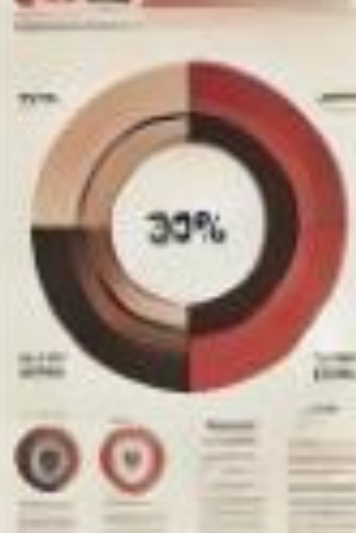
DEPI

Track: Data Analyst Specialist

Class Group: CAI_DAT1_G4e

Presented to Mr. Yasser Abdel Rahman

Group Members:

Aya Sayed Zayed

Marwa Ibrahim Waked

Doaa Abdalaal Gadallah

Noha Abdllah Muhammad

Lamya Muhammad Elhosseiny

Marwa Hamdi Muhammad

Table of Contents

1. Executive Summary

This project analyzed manufacturing defect data to identify key issues impacting product quality and operational efficiency. The objective was to uncover trends in defect rates, assess vendor performance, and offer recommendations to reduce defects and enhance production processes.

Key Findings:

- **Defect Impact**: A total of **54.54 million defective units** resulted in **138,311 minutes of downtime**, significantly affecting production efficiency and increasing operational costs.
- **Underperforming Vendors**: Vendors like **Solholdings** and **PlanetHouse** consistently contributed to high defect rates, particularly in key defect categories, which presented significant risks to production quality.
- **Improving Vendors**: Vendors such as **Tampquote**, **Instrip**, and **Plustax** showed notable improvements, highlighting the success of recent corrective measures in reducing defect rates.

Recommendations:

- **Improve High-Defect Vendors**: Take immediate action with **Solholdings** and **PlanetHouse** by enforcing stricter quality control and regular audits.
- **Enhance Data Insights**: Prioritize collecting data on production volumes, financial impacts, and root causes to better focus improvement efforts and track defect rates.
- **Sustain Vendor Success**: Keep monitoring vendors like **Tampquote** and **Plustax** to maintain their progress, and share their best practices with underperforming vendors to boost overall quality.
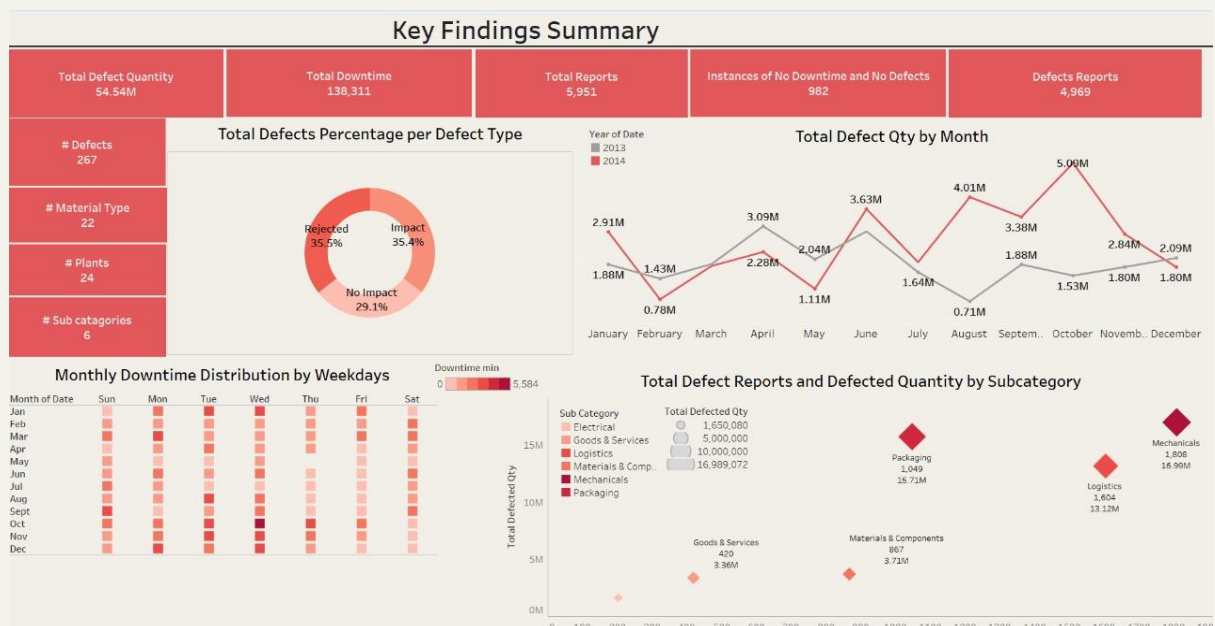
2. Introduction

The purpose of this project is to analyze defects reports to identify trends and insights related to vendors, plants, materials, and categories. The objective is to understand the frequency and distribution of defects, which vendors are responsible for the most defects, and whether there are correlations between material types, defect types, and defect quantities. Ultimately, this analysis aims to support decision-making processes around vendor management, material sourcing, and process improvements to reduce defects and enhance product quality and quality management inspections.

Problem Definition

The project aims to assess and benchmark vendor quality performance by analyzing defect quantities, defect types, and their impact on production efficiency. Multiple vendors, particularly **Solholdings** and **PlanetHouse**, have consistently exhibited elevated defect numbers across various categories, leading to significant operational inefficiencies and downtime. The goal is to identify the root causes of these high defect trends and evaluate the effectiveness of corrective measures for underperforming vendors, while also highlighting vendors that have improved over time.

Key Questions to Answer

1. Which vendors are contributing the most defects across various categories?
2. What are the key defect types that contribute to the highest defect rates, and how do they impact production efficiency?
3. Which subcategories show the highest defect frequencies, and which vendors are the biggest contributors?
4. Are there any seasonal trends or time-related spikes in defects?
5. Which vendors have shown improvements over time, and what factors contributed to their reduced defect rates?
6. What is the impact of defects on operational efficiency, particularly in terms of downtime?
7. What actionable recommendations can be made to improve vendor performance and reduce defect rates?

Project Phases

No.	Phase	Tools used
1	Data Exploration	Excel, Power Query & Python
2	Data Cleaning	Power Query & Python
3	Data Analysis & Visualization	Tableau & Python
4	Insights and Recommendations	Tableau
5	Storytelling	Tableau

Data Sources

Data Sources and Characteristics:

The dataset is stored in an Excel file with seven distinct sheets, each containing different categories of information:

Dimension Sheets:

1. **Vendor**: Contains information about different vendors (Vendor ID, Vendor Name).
2. **Plant**: Details about the plants involved in production (Plant ID, Plant Name).
3. **Material Type:** Information about the material types involved in production (Material Type ID, Material Type).
4. **Defects:** A list of defect types and their corresponding defect IDs.
5. **Defect Type**: Information on different defect types (Defect Type ID, Defect Type).
6. **Category:** Includes additional categorization details.

Fact Sheet:

Defected Items: The main data contains defect records, including Vendor ID, Material Type ID, Defect Qty, Date, etc.

The data varied in size and format across the sheets. For example, the "Defected Items" sheet was the largest, containing thousands of rows of data, while some lookup tables like "Category" and "Defect Type" contained fewer rows.

---

3. Data Exploration

Descriptive Statistics:

Tools used: Python and Excel

```
# Data Description
description = merged_df_cleaned[['Downtime min', 'Defect
Qty']].describe()
print("Data Description for 'Downtime min' and 'Defect Qty':")
display(description)
```

Phase 1. Data Exploration

```
Data Description for 'Downtime min' and 'Defect Qty':
              Downtime min      Defect Qty
count        5952.000000      5952.000000
mean           23.237735      9163.179772
std            76.344141     30585.908677
min             0.000000         0.000000
25%             0.000000         9.000000
50%             0.000000       445.000000
75%            20.000000      5173.250000
max           999.000000    487008.000000
```

	Defect Qty		Downtime min
Mean	9,164.71	Mean	23.24
Standard Error	396.52	Standard Error	0.99
Median	445.00	Median	0.00
Mode	0.00	Mode	0.00
Standard Deviation	30,588.25	Standard Deviation	76.35
Sample Variance	935,641,014.46	Sample Variance	5,829.32
Skewness	8.02	Skewness	7.88
Range	487,008.00	Range	999.00
Minimum	0.00	Minimum	0.00
Maximum	487,008.00	Maximum	999.00
Sum	54,539,216.00	Sum	138,311.00
Count	5,951.00	Count	5,951.00

The data reveals that most defects result in little to no downtime (median downtime is 0 minutes), The mean is 23 mins, although some defects cause extended downtime (up to 999 minutes). Regarding defect quantities, the data shows significant variability: while 50% of records report fewer than 445 defects, there are substantial outliers, with some cases having up to 487,008 defects. These outliers suggest that a few vendors or plants may have major quality issues, warranting further investigation.

Histogram:

```python
import matplotlib.pyplot as plt
# Part 3:Data Visualization
# Creating and displaying Histograms of Defect Quantity and Downtime
# Create a figure with 1 row and 2 columns
fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(12, 6))

# Plot the histogram for Defect Quantity
```

Phase 1. Data Exploration

```
ax1.hist(merged_df_cleaned['Defect Qty'], bins=30, color='grey',
edgecolor='black')
ax1.set_title('Histogram of Defect Quantity')
ax1.set_xlabel('Defect Quantity')
ax1.set_ylabel('Frequency')
ax1.grid(False)

# Plot the histogram for Downtime
ax2.hist(merged_df_cleaned['Downtime min'], bins=30, color='red',
edgecolor='black')
ax2.set_title('Histogram of Downtime')
ax2.set_xlabel('Downtime (min)')
ax2.set_ylabel('Frequency')
ax2.grid(False)
# Adjust spacing for better readability
plt.tight_layout()
# Display the plot
plt.show()
```
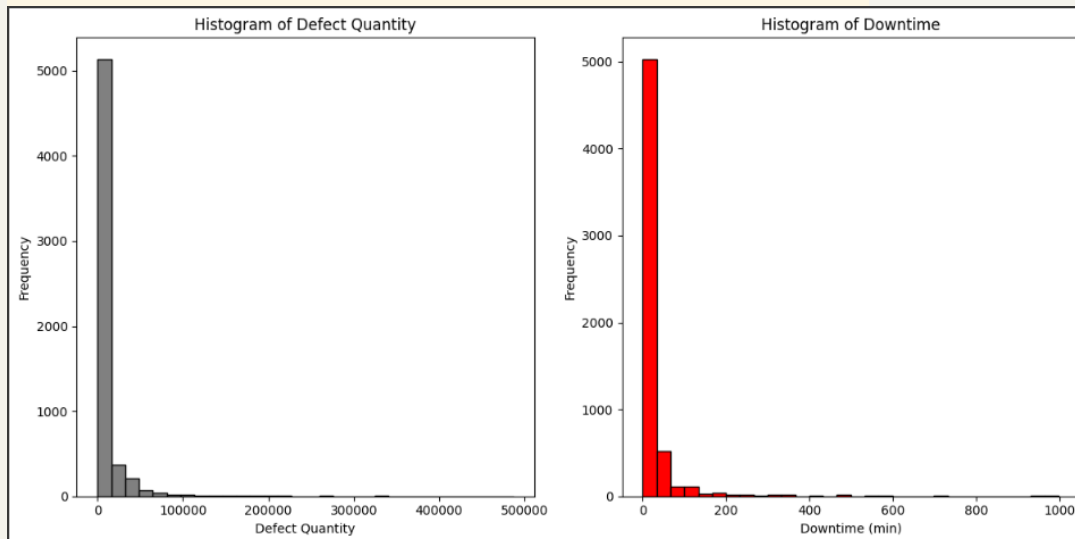

Key Observations:

1. **Defect Quantity:** The distribution is highly right-skewed, with most defect records concentrated below 100,000, but a small number of records extend to much higher values, up to around 487,000. This suggests that while most defect events involve smaller quantities, there are a few extreme outliers with very large defect counts.

2. **Downtime (minutes):** The downtime data also shows a right-skewed distribution, with the majority of downtimes clustered near 0-100 minutes, and a few outliers reaching up to 999 minutes. Most records indicate little to no downtime, but some cases have significant downtime.

Both distributions suggest that while the majority of the data lies at the lower end of the spectrum, there are rare but impactful cases where defect quantities or downtimes are unusually high, potentially driving the overall operational inefficiencies.

Scatter Plot:

```
# Create a figure with a suitable size for the scatter plot
fig, ax = plt.subplots(figsize=(10, 6))
# Visualize the relationship between Downtime and Defect Quantity
ax.scatter(merged_df_cleaned['Downtime min'], merged_df_cleaned['Defect Qty'], alpha=0.5,
color='red')
# Add a title and labels to describe the plot
ax.set_title('Scatter Plot of Downtime vs. Defect Quantity')
ax.set_xlabel('Downtime (min)')
ax.set_ylabel('Defect Quantity')
# Display the scatter plot
plt.show()
```
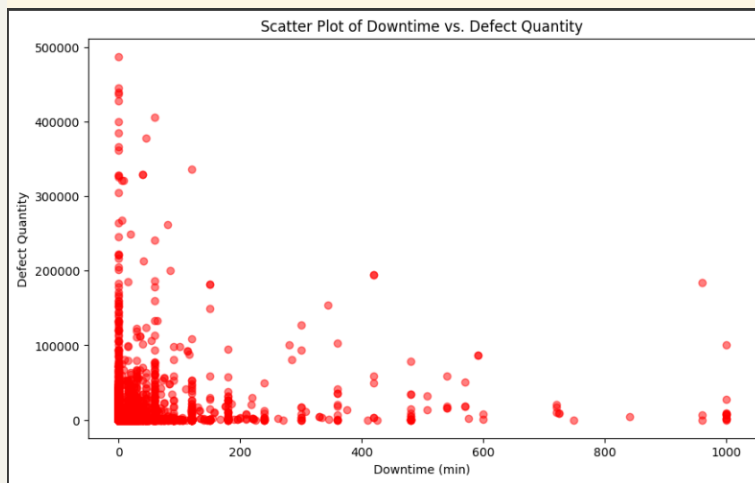

This scatter plot illustrates the relationship between Downtime (minutes) and Defect Quantity.

Key Observations:

1. **Clustered Data:** Most of the data points are clustered in the lower-left corner, indicating that many defects have low downtime (0-100 minutes) and low defect quantities (below 100,000).

2. **Outliers:** There are a few outliers with high defect quantities (above 300,000) and a few cases with high downtimes (over 600 minutes). However, high defect quantities do not always correspond with high downtime.

3. **General Pattern:** The scatter plot shows that there isn't a strong correlation between defect quantity and downtime. Some points show high defect quantities with low downtime, and vice versa, suggesting that other factors may influence downtime beyond defect quantity.

The relationship between downtime and defect quantity is weak, as higher defect counts don't necessarily result in longer downtimes. Most incidents involve smaller defects and minimal downtimes, but a few critical outliers significantly impact overall performance.

Data Cleaning

Tools used: Python and Power Query

To ensure the accuracy and reliability of the data, several steps were taken to clean and preprocess the dataset:

1. **Merging Tables:**
   The "Defected Items" sheet was merged with the other relevant sheets (e.g., Vendor, Plant, Material Type, Defect Type, Category) to enrich the dataset. This merging provided additional context to the defect records, enabling deeper analysis based on vendor performance, material types, defect types, and plant locations.

2. **Handling Missing Values:**
   In the dataset, Downtime mins and Defected Qty columns: contained values that were zeros rather than nulls. These zero values represent records where no defects were identified or recorded.

   Reasoning: Since the absence of defects is represented by zero values rather than null, these values were left as they are, indicating defect-free records. This approach maintains data integrity and avoids introducing bias by misinterpreting the zeros.

3. **Splitting the "Plant" Column:**
   The Plant Name column contained combined information for both the city and state of each plant (e.g., "Houston, TX"). To facilitate location-based analysis, this column was split into two.
   This separation allows us to analyze defects by geographic location more easily, enabling targeted quality control efforts in specific cities or states.

4. **Removing Duplicate Records:**
   The dataset contained some duplicate rows, which could skew the analysis results. These duplicates were identified and removed to ensure each defect record was unique and accurately represented in the analysis. This step is crucial to eliminate redundant information and ensure each row represents a unique observation.

5. **Removing the Duplicate "Category" Column:**
   After merging multiple sheets, a duplicate Category column was discovered in the dataset. This redundant column was removed to avoid confusion and streamline the dataset.

Phase 2. Data Cleaning

6. **Data Type Conversion:**
   - **Date Column:** To enable accurate time-based tracking, we converted this column from DateTime to date format. This conversion allows us to effectively analyze defect trends over time, such as identifying monthly or yearly fluctuations in defect quantities.

   - **Numeric Columns:** To ensure data consistency and facilitate calculations, we converted any relevant numeric columns, such as defect quantities or severity ratings, to appropriate numeric data types (whole number or integer). This step is crucial for performing quantitative analysis and generating meaningful insights.
   - **Categorical Columns:** For categorical columns like defect types or product categories, we converted them to a suitable categorical data type (text or string).

7. **Standardizing Columns:**
   Some columns contained inconsistencies, such as spaces or varying capitalization. To improve readability and maintain consistency, the columns were standardized using Trim, Clean, and functions in Power Query.

8. **Capitalization and Spell Check:**
   - **Capitalization:** To standardise text data, a transformation was applied to capitalize the first letter of each word in relevant text columns (Vendor Name, Plant City, Plant State, Material Type, Defect Type). This improved consistency and readability across the dataset.

   - **Spell Check on "Defect":** A spell check was performed exclusively on the "Defects" column to ensure that defect descriptions are correctly spelled. Correcting spelling mistakes in this column reduced the risk of misinterpretation and ensured that defect types were consistently labelled.

   (Refer to Appendix A for Python code and Power Query screenshots)

Phase 3. Data Analysis & Visualization

4.Data Analysis

Vendor Analysis

The below "Vendors Analysis," provides several critical insights about vendor performance based on different metrics:

Geographic Vendor Distribution

Key Findings:

1. **Illinois Leads in Vendor Activity:**

   o **Illinois** has the highest vendor transaction volume with 3.18M from the vendor **Plustax**, significantly ahead of other states.

2. **Michigan and Indiana Show Strong Vendor Presence:**

   o **Michigan** reports 1.81M transactions via **Bamtechnology**, while Indiana follows with 1.64M from **Dentocity,** both contributing majorly to the overall activity.

3. **Wisconsin and Ohio Have Mid-Range Vendor Activity:**

   o **Wisconsin** records 1.48M transactions by **Quotelane,** and **Ohio** shows 0.73M from Sanin, both in a moderate range compared to other states.

4. **Iowa Has the Least Activity:**

   o **Iowa** has the lowest vendor activity, with just 0.04M transactions from **Blueit.**

Statistical Evidence:

- Darker map shades (like **Illinois)** indicate higher vendor transaction volumes, while lighter shades (like Iowa) reflect much lower volumes.

- Transaction volume ranges from 0.04M in Iowa to 3.18M in Illinois, highlighting a large disparity.

Interesting Discoveries:

- **Disparity in Vendor Distribution:** Illinois has nearly 80 times more vendor transactions than Iowa, emphasizing unequal distribution.

- **Midwest Focus**: Vendor activities are concentrated in the Midwest, suggesting a regional focus for vendor relations.

Phase 3. Data Analysis & Visualization

Top Vendors by Total Defected Items:

Top 10 Vendors by Total Defected Items
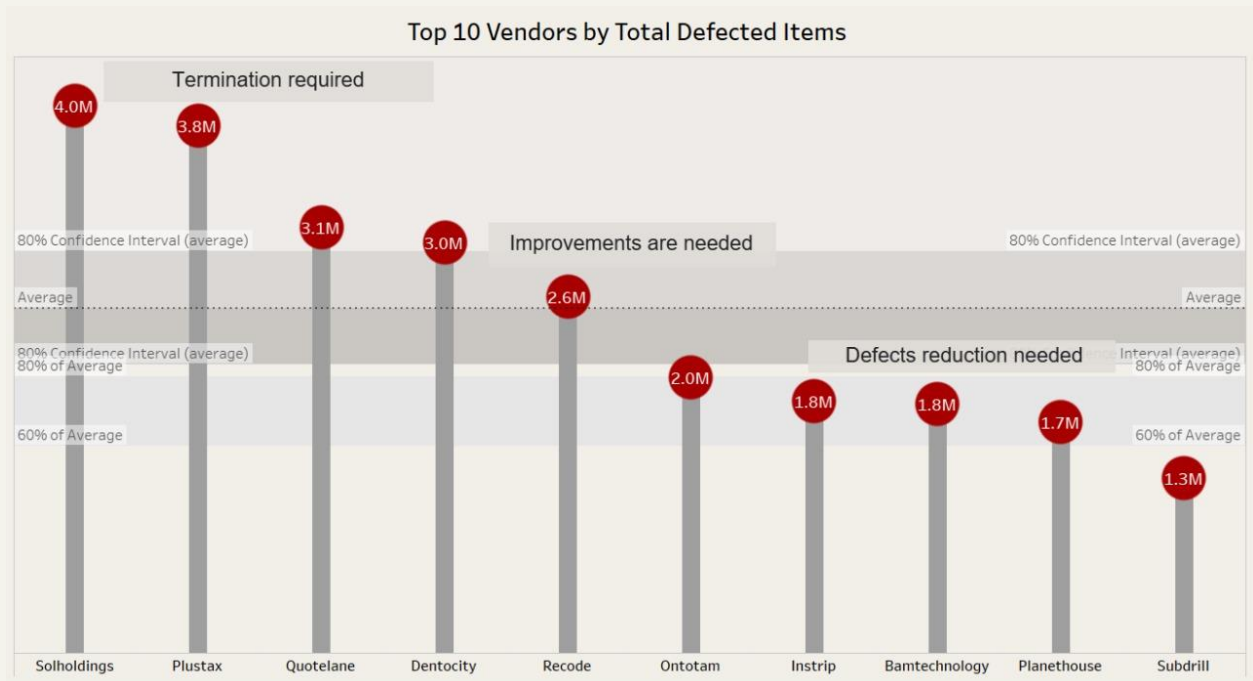
Key Findings:

- **Solholdings** and **Plustax** have the highest defect counts at 4.0M and 3.8M, respectively. These vendors exceed the average by a significant margin, warranting immediate action, possibly even termination.

- **Quotelane**, **Dentocity**, and Recode, with defects ranging between 2.6M and 3.1M, require improvements as they perform above the average but not as critically as the top two.

- **Ontotam**, **Instrip**, **Bamtechnology**, and **Planethouse** contribute between 1.7M and 2.0M defects and need to focus on reducing defects to improve quality.

- **Subdrill**, with 1.3M defects, is performing closer to acceptable levels but still requires improvement.

Unexpected Discoveries:

- **Solholdings** and **Plustax** stand out with drastically higher defect rates, suggesting systemic quality issues.

- The distribution of vendors in the middle-showed variability, with some, like **Subdrill**, closer to acceptable defect levels.

- Wide confidence intervals for vendors like **Solholdings** indicate inconsistent performance, suggesting further investigation is needed.

Vendor Downtime per Incident:
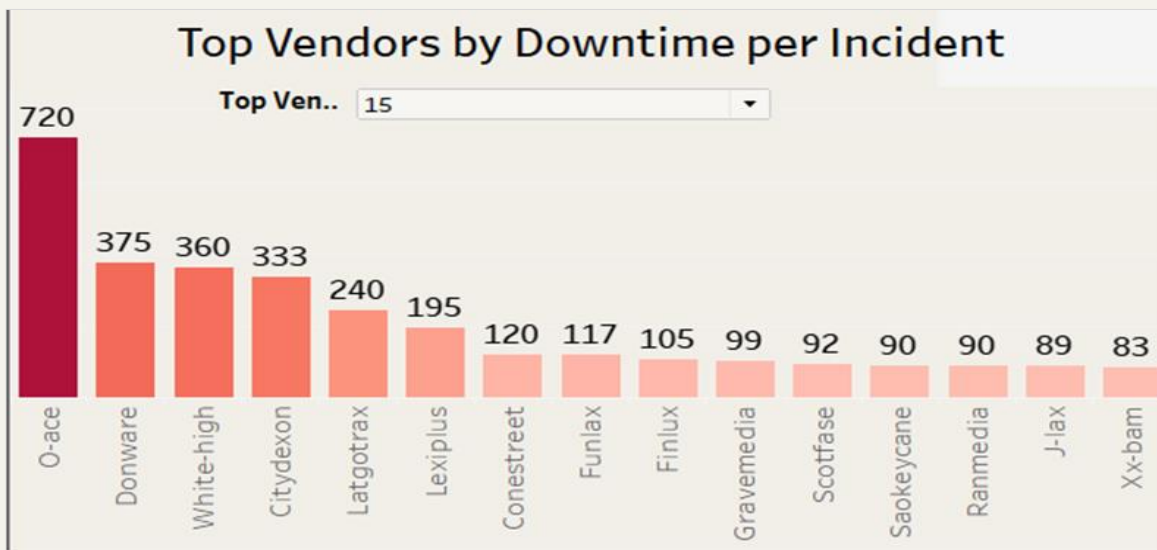
Key Findings:
1. **O-ace** Leads Significantly in Downtime per Incident:
    - **O-ace** has the highest downtime per incident at 720 minutes, which is drastically higher than the other vendors.
2. **Donware**, **White-high**, and **Citydexon** Show High Downtime:
    - **Donware** (375 minutes), **White-high** (360 minutes), and **Citydexon** (333 minutes) also have high downtime but are far below **O-ace**.
3. **Mid-Range Vendors**:
    - **Latgotrax** (240 minutes) and **Lexiplus** (195 minutes) are in the mid-range of downtime per incident.
4. **Lower Downtime Vendors in the Top 15:**
    - Vendors like **Conestreet, Funlax, Gravemedia**, and others report lower downtimes, ranging from 120 minutes to 83 minutes (with Xx-bam being the lowest).

Statistical Evidence:
- The downtime per incident ranges from 720 minutes (O-ace) to 83 minutes (**Xx-bam), with O-ace** being nearly double the downtime of the second-highest vendor, **Donware**.

Interesting Discoveries:
1. **Extreme Downtime from O-ace:**
    - **O-ace's 720 minutes** per incident is a significant outlier, nearly twice as much as the next vendor, indicating potential operational issues.
2. **Gradual Decline in Downtime Among Remaining Vendors:**

Phase 3. Data Analysis & Visualization

- o After the top four vendors, the downtime decreases more gradually, with mid-range vendors having similar downtimes compared to those at the lower end of the top 15.

Vendor Defects Across Subcategories:

Vendor..	Logistics	Materials & Components	Packaging	Mechanicals	Goods & Services	Electrical
Dentocity	3	107	1	6	70	
Ontotam			49		21	82
Plustax	308					
Quotelane	79	102	9	190		
Reddoit	546	80	19	9		52
Sanlab	60				110	
Scotquote	148					
Subdrill				251		
Trio-dax		239				
Xx-way			402			

Vendor Defects Frequency Across Subcategories — Sub Category

Key Findings:

1. **Reddit Leads in Defects in Logistics:**

   - o **Reddit** has the highest number of defects in the Logistics subcategory, **totaling 546**, which is the largest defect count across all vendors and subcategories.

2. **Xx-way Has the Most Defects in Packaging:**

   - o **Xx-way** reports the most defects in the Packaging subcategory, with 402 defects, indicating a significant issue in this category.

3. **Plustax and Subdrill Show High Defect Frequencies:**

   - o **Plustax records 308** defects in Logistics, while **Subdrill** shows 251 defects in Mechanicals, reflecting major defect concerns in these subcategories.

4. **Quotelane's Mechanical Defects Are Significant:**

   - o **Quotelane** has 190 defects in the **Mechanical** subcategory, the highest defect count for mechanical issues among all vendors.

Statistical Evidence:

- Darker shades in the heatmap represent higher defect frequencies, with Logistics and Packaging showing high concentrations of defects for certain vendors.

- **Reddit** (546 in Logistics) and **Xx-way** (402 in Packaging) dominate defect counts in their respective subcategories.

Interesting Discoveries:

1. **Concentration of Defects in Specific Categories:**

   o Vendors such as **Reddit** and **Plustax** are heavily focused on **Logistics** defects, while **Xx-way** and **Quotelane** stand out for defects in **Electrical and Packaging**, respectively.

2. **Scotquote and Sanlab Have Relatively Balanced Defects:**

   o **Scotquote** and **Sanlab** display a more balanced spread of defects across subcategories, contrasting with other vendors that show concentrated issues.

3. **Varied Defect Distribution Across Vendors:**

   o Some vendors, like **Trio-dax**, focus their defect frequencies in specific subcategories (e.g., **Materials & Components**), while others like **Dentocity** and **Ontotam** have a more dispersed distribution of lower defect counts.

Defects Analysis

The below dashboard provides an in-depth look at defect quantities, their distribution across states, the types of defects, and the downtime associated with them. Below are the key findings:
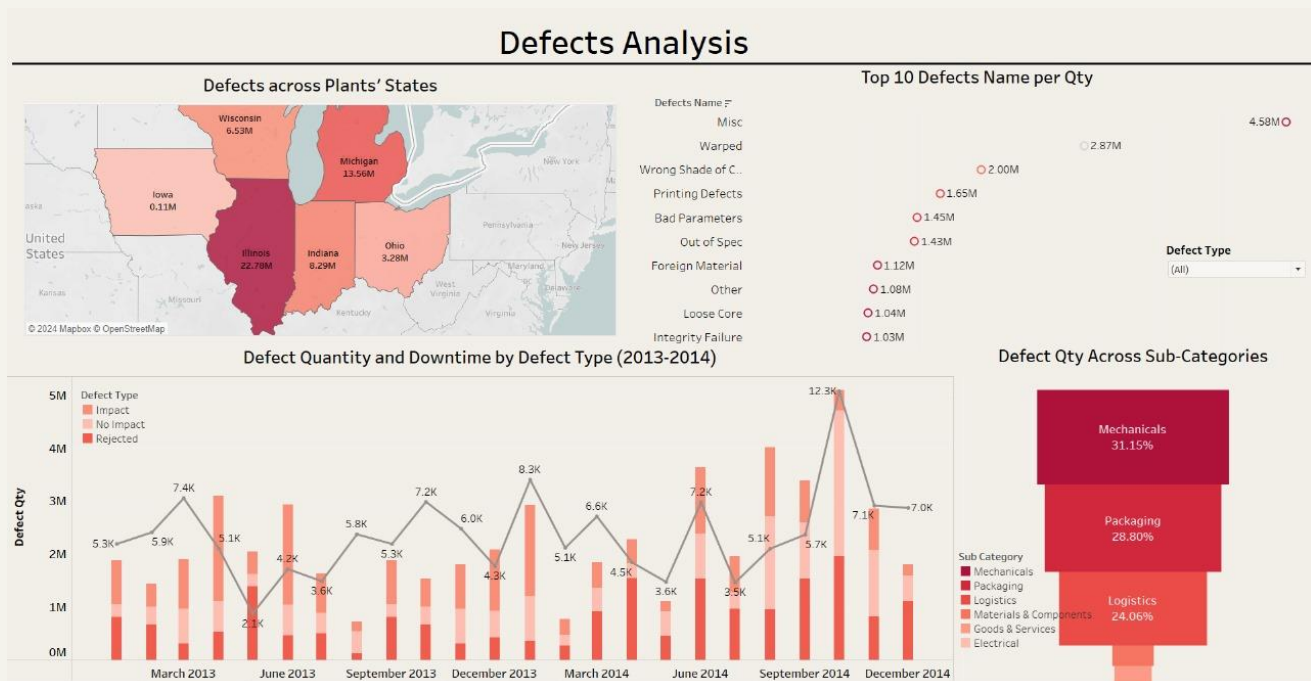
Defects Analysis

Phase 3. Data Analysis & Visualization

Defects Across Plants' States:

Key Findings

1. **Illinois Leads in Defects:**

   o **Illinois** reports the highest number of defects at **22.78M**, making it the state with the most significant defect issues.

2. **Michigan Shows a Large Volume of Defects:**

   o **Michigan** follows with **13.56M defects**, a substantial number but significantly lower than Illinois.

3. **Indiana and Wisconsin Also Report High Defect Volumes:**
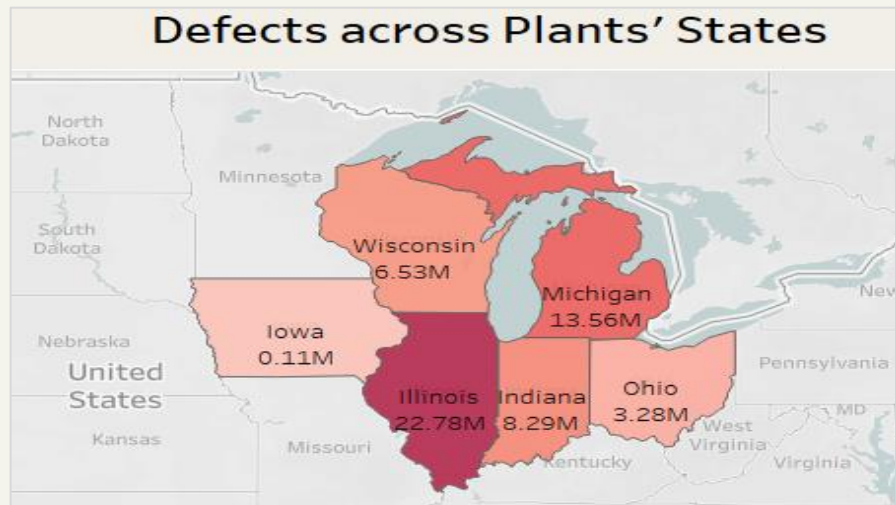
   o **Indiana** has **8.29M defects**, while **Wisconsin** shows **6.53M**, representing mid-range defect volumes compared to Illinois and Michigan.

4. **Lower Defect Volumes in Ohio and Iowa:**

   o **Ohio** reports 3.28M defects, and **Iowa** has the lowest at 0.11M, significantly trailing behind the other states.

Statistical Evidence:

- Darker shades on the map correspond to higher defect volumes, with Illinois having the darkest shade. The range of defects spans from 0.11M in Iowa to 22.78M in Illinois, illustrating the disparity in defect counts between the states.

Interesting Discoveries:

Phase 3. Data Analysis & Visualization

1. **Illinois as a Major Outlier:**

   o   Illinois's 22.78M defects far exceed those of other states, creating a considerable gap of nearly 9M defects between Illinois and Michigan, the second-highest state.

2. **Regional Clustering of Defects:**

   o   The Midwest region, particularly **Illinois, Michigan, Indiana, and Wisconsin,** accounts for the majority of defects, with over 50M defects combined, indicating a regional focus for improvement.

Defect Quantity and Downtime by Defect Type (2013-2014)
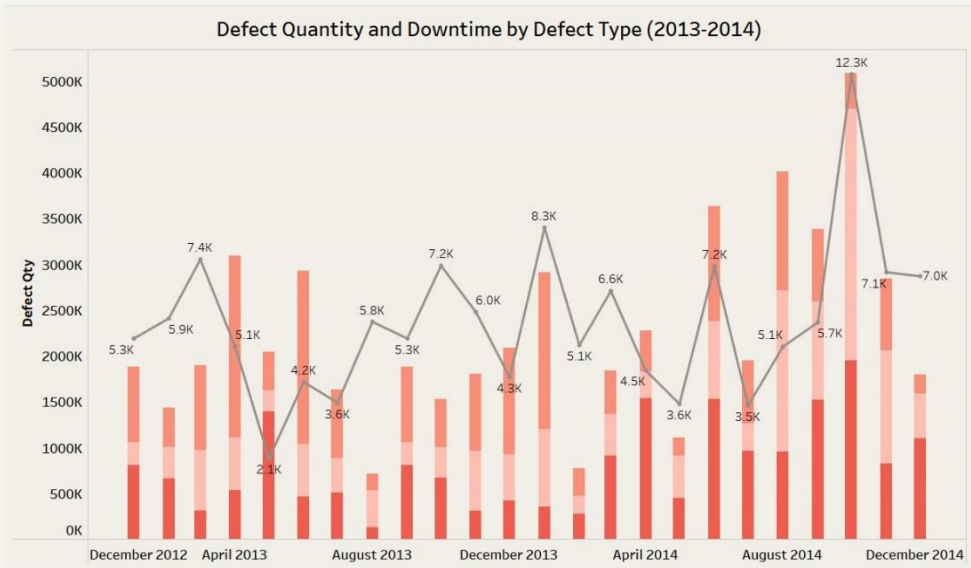
3. **Iowa's Minimal Defects:**

   o   **Iowa's** 0.11M defects are remarkably low compared to other states, suggesting fewer plants or better-quality control processes in the state.

Defect Quantity and Downtime by Defect Type (2013-2014):

Key Findings:

1. **Significant Defect Spikes in April 2013 and October 2014:**

   o   Defect quantities saw major spikes in April 2013 and October 2014, with October with the highest downtime, marking it as a critical period.

2. **Fluctuations in Defect Quantity:**

   o   There were continuous fluctuations throughout the observed period, with peaks in August 2013 (712K) and April 2014 (2.28M). These spikes were followed by notable declines, showing dynamic shifts in defect occurrences.

Phase 3. Data Analysis & Visualization

3. **General Trend of Increasing Defect Quantity Toward the End of 2014:**

   o A clear upward trend in defect quantities is observed toward the end of 2014, particularly in August 2014 and October 2014, suggesting a worsening defect trend during that period.

Statistical Evidence:

- The most significant peak occurred in October 2014 with 12.3K downtime minutes, while other peaks (such as April 2013 with 5.1K minutes) show similar trends but with less severe downtime impacts.

Interesting Discoveries:

1. **Mismatch Between Defect Quantity and Downtime:**

   o Although October 2014 had the highest combined defect quantity and downtime, other periods like August 2013 and February 2014 had high defect counts but relatively lower downtime, indicating the possibility of less operational impact from those defects.

2. **Frequent Defect Cycles:**

   o There is a recurring cycle of defect spikes every few months, with noticeable peaks followed by drops, suggesting potential cyclical operational or external factors influencing defect trends.

3. **Steep Declines Follow High Spikes:**

   o Significant declines often follow major defect spikes, implying corrective actions or changes may have been implemented after high defect periods.

Defect Quantity Across Subcategories:

Defect Qty Across Sub-Categories
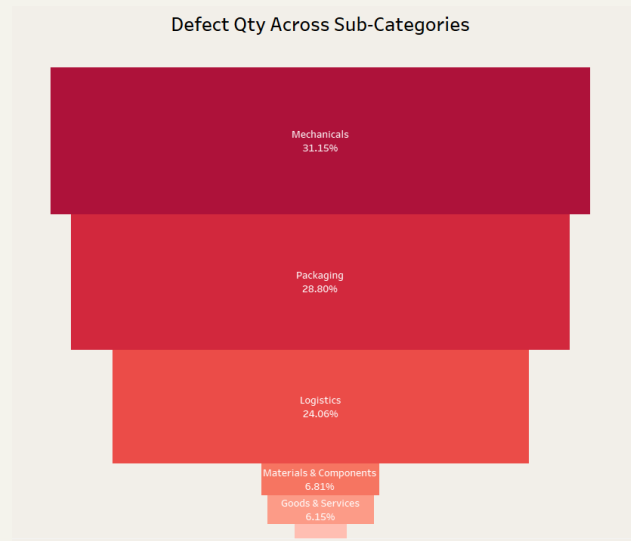
Key Findings:

1. **Mechanical Defects Dominate**:

   o **Mechanical defects** contribute the largest portion, accounting for **31.15%** of the total defects, making this the top category for defects.

2. **Packaging Defects Are the Second-Largest**:

   o **Packaging defects** make up **28.80%** of the total, placing them as the second-most frequent defect category.

3. **Logistics Defects Are Also Significant**:

   o **Logistics defects** represent **24.06%**, indicating that while slightly less than the top two categories, it still contributes substantially to the total defect count.

4. **Materials & Components and Goods & Services Have Lower Defect Quantities**:

   o **Materials & Components** account for **6.81%**, while **Goods & Services** contribute **6.15%**, both making up a smaller portion of the overall defects.

Statistical Evidence:

- **Mechanical defects** lead at **31.15%**, followed by **Packaging** at **28.80%**, and **Logistics** at **24.06%**.

- **Materials & Components** and **Goods & Services** combined make up about **12.96%** of the total defect share.

Phase 3. Data Analysis & Visualization

Interesting Discoveries:

1. **Mechanical Defects as the Primary Issue**:

   o   With **31.15%** of defects coming from the Mechanical category, addressing these defects could have the most significant impact on reducing overall defect levels.

2. **Small Discrepancy Between Top Three Categories**:

   o   The defect percentages for **Mechanical**, **Packaging**, and **Logistics** are fairly close, with only about **7%** separating the top (Mechanical) from the third-largest category (Logistics), indicating that all three are key areas to address.

3. **Relatively Minor Contributions from Materials & Components and Goods & Services**:

   o   The smaller percentage of defects from **Materials & Components** and **Goods & Services** shows that these categories contribute far less to the overall defect count, suggesting fewer quality issues in these areas.

Top 10 Defects by Quantity:

Top 10 Defects Name per Qty

Defects Name	Qty
Misc	4.58M
Warped	2.87M
Wrong Shade of C..	2.00M
Printing Defects	1.65M
Bad Parameters	1.45M
Out of Spec	1.43M
Foreign Material	1.12M
Other	1.08M
Loose Core	1.04M
Integrity Failure	1.03M
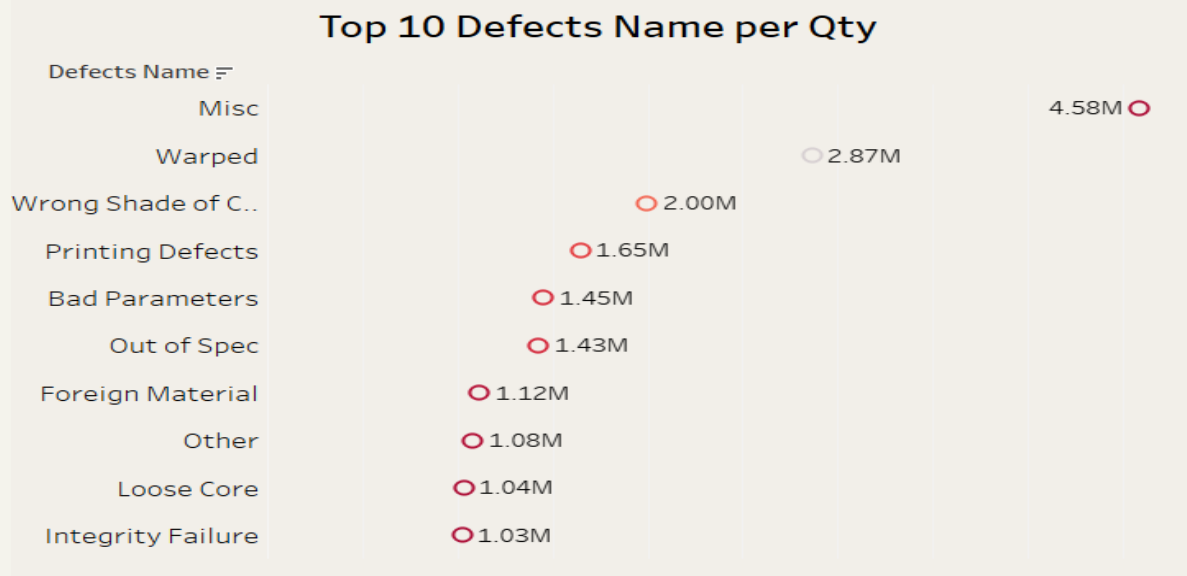
Key Findings:

Phase 3. Data Analysis & Visualization

1. **Misc Defects Dominate**:

    o The **Misc** defect category leads by a significant margin with **4.58M** defects, making it the most frequent defect type in the dataset.

2. **Warped Defects Follow**:

    o **Warped** defects are the second most common, with **2.87M** defects, significantly lower than the Misc category but still substantial.

3. **Lower Frequency Defects**:

    o Defects like **Foreign Material** (1.12M), **Other** (1.08M), **Loose Core** (1.04M), and **Integrity Failure** (1.03M) round out the list, contributing slightly over a million defects each.

    Statistical Evidence:

- The defect count for other issues like **Wrong Shade of Color** (2.00M) and **Printing Defects** (1.65M) shows these issues are also quite common.

- The lowest categories, such as **Integrity Failure**, still represent over **1.00M** defects, reflecting the broad impact of various defect types.

    Interesting Discoveries:

1. **Large Discrepancy Between Misc and Other Defects**:

    o **Misc** defects contribute nearly **double** the number of defects compared to **Warped** defects, indicating that a large portion of issues may be generalized into this category, potentially signaling a need for better defect classification.

2. **Relatively High Incidence of Color and Printing Defects**:

    o The prominence of the **Wrong Shade of Color** (2.00M) and **Printing Defects** (1.65M) highlights the importance of quality control in the aesthetic or labeling aspects of production.

3. **General Consistency Among Lower Defect Types**:

    o The bottom four categories—**Foreign Material**, **Other**, **Loose Core**, and **Integrity Failure**—each contribute just over **1.00M** defects, suggesting that while these defects are less frequent than the top contributors, they still represent significant quality issues.

Material Analysis

The "Material Analysis" dashboard below offers insights into defect quantities and downtime contributions by various material types, as well as their distribution across different plants and states:
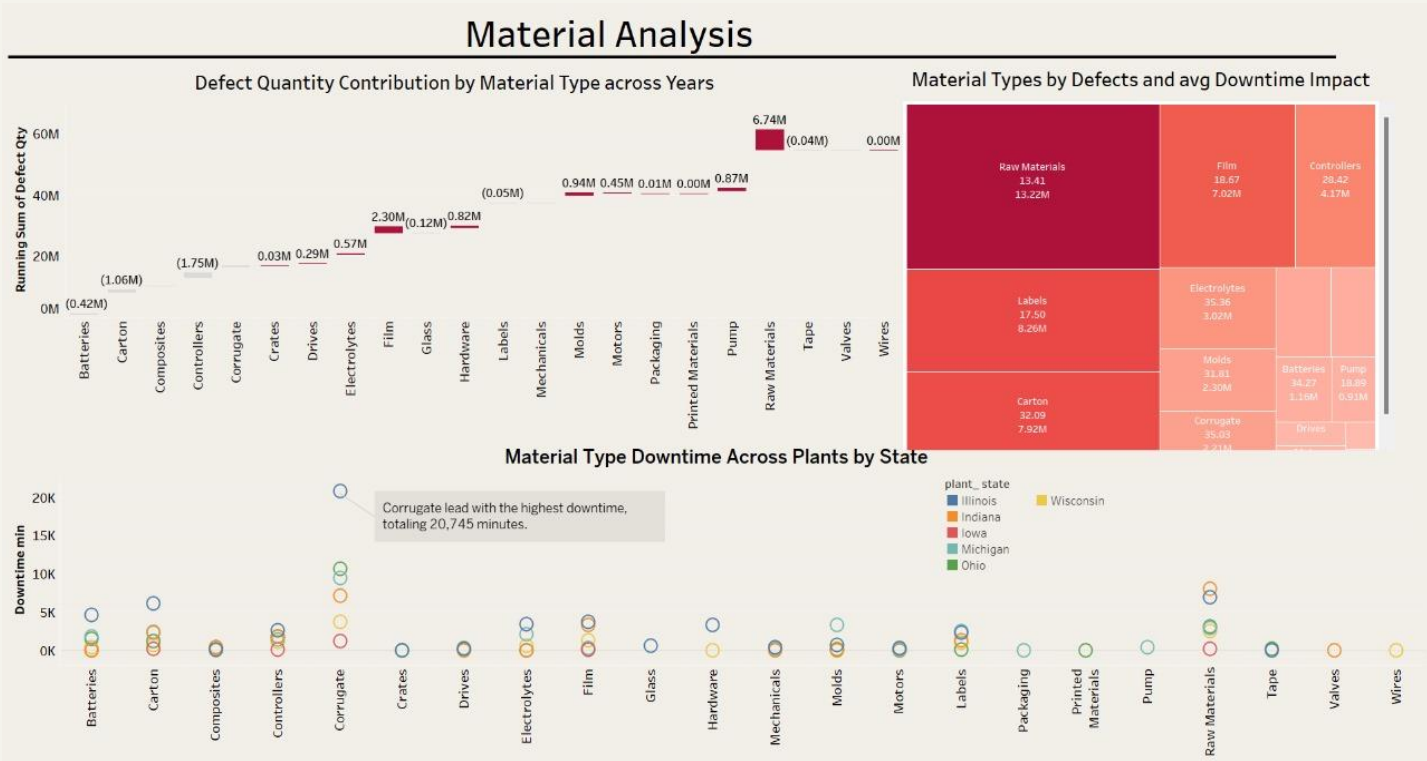
Defect Quantity Contribution by Material Type:
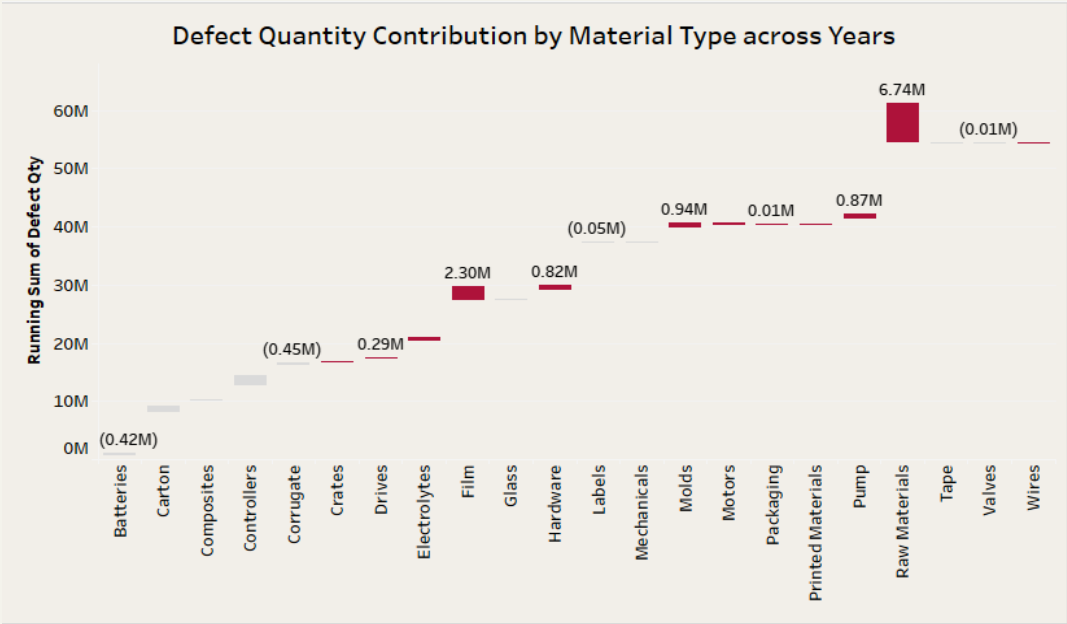
Key Findings:

1. **Raw Materials Dominate Defect Contributions**:

    o **Raw Materials** contribute the highest increase number of defects, totaling **6.74M**, which is significantly more than any other material type.

2. **Film and Electrolytes Follow in Defect Contribution**:

    o **Film** ranks second with **2.30M** defects, while **Electrolytes** come in third with **0.87M**, making them significant contributors as well.

3. **Decrease in Contributions from a Wide Range of Materials**:

    o Materials like **Glass** (0.82M), **Labels** (0.94M), and **Packaging** (0.01M) contribute smaller, yet notable, amounts of defects when compared across years.

Statistical Evidence:

- **Raw Materials** lead with **6.74M** defects, followed by **Film** with **2.30M**, and **Electrolytes** with **0.87M**.

- Other materials like **Glass** and **Labels** contribute over **0.80M** defects, while **Tape**, **Valves**, and **Wires** show much lower defect quantities, between **0.01M** and **0.05M**.

Interesting Discoveries:

1. **High Contribution from Raw Materials**:

    o The significant increase from **Raw Materials** 6.74M suggests this is a critical area for process improvement, as it far surpasses other material types in defect totals.

2. **Film and Electrolytes as Key Contributors**:

    o The defect quantity increased for **Film** 2.30M and **Electrolytes** 0.87M highlights potential issues related to these materials, signaling that they could be areas requiring attention.

3. **Negligible Defect Quantities for Certain Materials**:

    o Materials like **Tape**, **Valves**, and **Wires** contribute very small quantities of defects, indicating they are not major areas of concern in terms of defects.

Material Types by Defects and Avg Downtime

Material Types by Defects and avg Downtime Impact
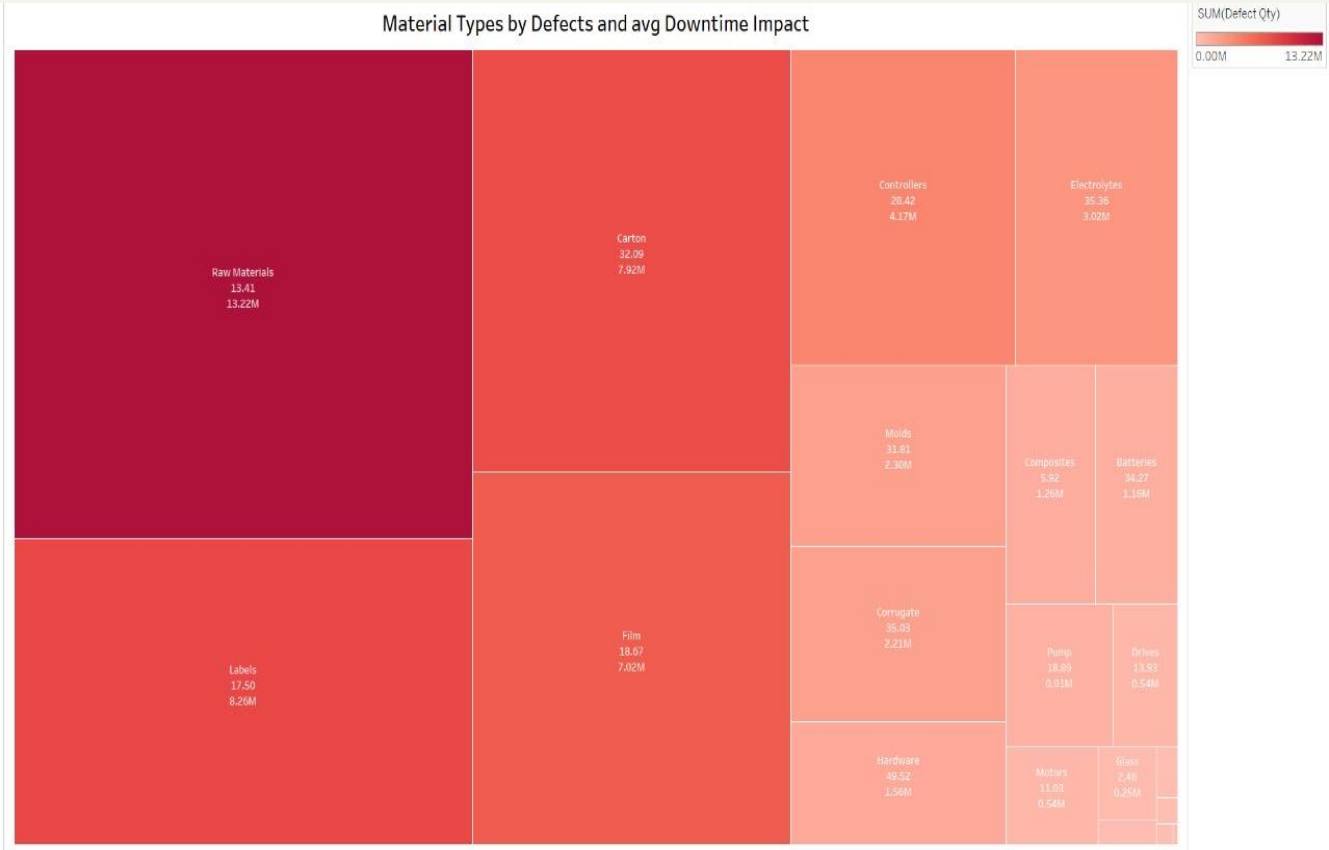
Key Findings:

1. **Raw Materials Have the Highest Defects and Downtime**:

    o **Raw Materials** lead in defect quantity (**13.22M**) while average downtime impact (**13.41** minutes), making them the most significant contributor to overall defects and operational downtime.

2. **Carton and Film Also Contribute Substantial Defects and Downtime**:

    o **Carton** reports **7.92M** defects and **32.09** average downtime minutes, while **Film** contributes **7.02M** defects and **18.67** average downtime minutes, highlighting these materials as major contributors to defects and downtime.

3. **Labels, Controllers, and Electrolytes as Significant Contributors**:

    o **Labels** (8.26M defects, 17.5 average downtime), **Controllers** (4.17M defects, 28.42 average downtime), and **Electrolytes** (3.02M defects, 35.36 average downtime) are also key contributors, significantly affecting both defect rates and downtime.

Statistical Evidence:

- **Raw Materials** dominate defect quantity (**13.22M**) with average downtime impact (**13.41** minutes), while **Carton** and **Film** contribute **7.92M** and **7.02M** defects respectively.

- **Electrolytes** 3.02M defects, 35.36 average downtime, and **Controllers** (4.17M defects, 28.42 average downtime) also impact both defect and downtime rates significantly.

Material Type Downtime Across Plants by State:

Material Type Downtime Across Plants by State
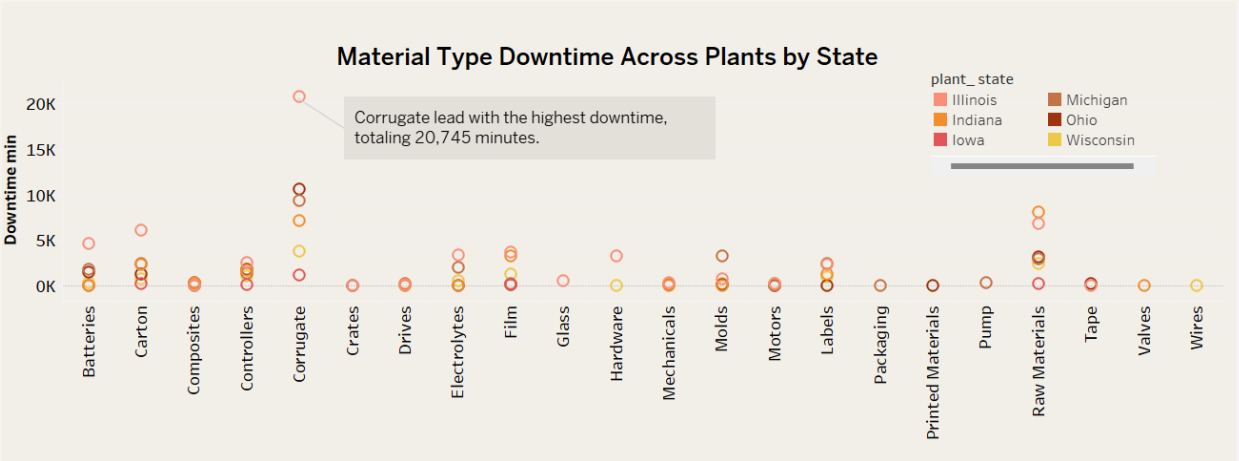
Key Findings:

1. **Corrugate Leads in Downtime**:

   o **Corrugate** is the material type with the highest downtime across all states, with a total of **20,745 minutes**, making it the most significant contributor to delays.

2. **Raw Materials and Labels Show High Downtime**:

   o **Raw Materials** show substantial downtime, approaching **15,000 minutes**.

   o **Labels** also contribute significantly, with downtime of around **10,000 minutes**, indicating a major cause of delays.

3. **Moderate Downtime for Film, Packaging, and Controllers**:

   o **Film**, **Packaging**, and **Controllers** experience moderate downtime, each ranging between **5,000 and 7,000 minutes**, which is notable but not as severe as the top contributors.

4. **Smaller Downtime for Batteries, Valves, and Wires**:

Phase 3. Data Analysis & Visualization

- Batteries, **Valves**, and **Wires** have relatively low downtime, all below **2,000 minutes**, indicating these materials cause fewer operational delays.

Statistical Evidence:

- **Corrugate** leads with **20,745 minutes** of downtime, followed by **Raw Materials** and **Labels** with around **15,000** and **10,000 minutes** of downtime, respectively.

- **Film**, **Packaging**, and **Controllers** show downtime between **5,000 to 7,000 minutes**.

- **Batteries**, **Valves**, and **Wires** have downtime totals below **2,000 minutes**, indicating minimal delays.

Interesting Discoveries:

1. **State Contributions Vary Across Material Types**:

   o Different states contribute varying levels of downtime for each material type. States like **Illinois** and **Michigan** have higher downtime contributions for materials like **Corrugate** and **Raw Materials**.

   o States like **Iowa** and **Wisconsin** contribute less to overall downtime, indicating better operational performance with certain materials.

2. **Corrugate's Significant Impact on Downtime**:

   o **Corrugate** stands out with **20,745 minutes** of downtime, far exceeding other materials. This suggests that Corrugate-related issues have a significant impact on operations.

3. **Materials with Minimal Downtime**:

   o **Valves**, **Tape**, and **Wires** have very minimal downtime, showing that these materials are less problematic in terms of causing delays in operations.

Phase 3. Data Analysis & Visualization

Downtime Analysis

The "Downtime Analysis" dashboard below provides a comprehensive view of the average downtime across various plants, as well as the total downtime minutes contributed by different states and locations.
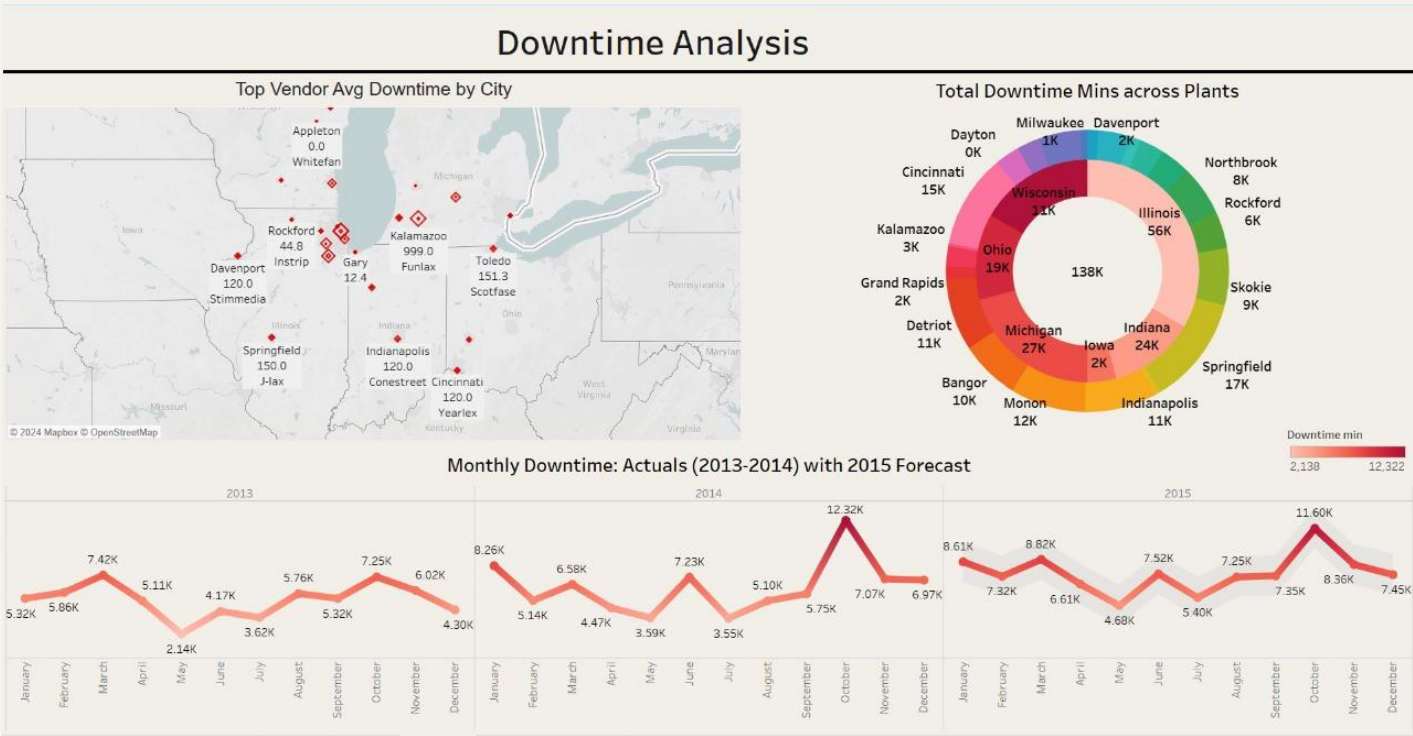
Top Vendor Avg Downtime by City

Phase 3. Data Analysis & Visualization

Key Findings:

1. **Kalamazoo Has the Highest Average Downtime:**

   o **Kalamazoo** reports the highest vendor downtime at 999 minutes, caused by the vendor **Funlax**, which is notably higher than any other city on the map.

2. **Milwaukee and Lansing Also Show High Average Downtime:**

   o Milwaukee reports 288 minutes of downtime from Lazap, while Lansing shows 360 minutes of downtime from **Viatom**, making these cities significant contributors to downtime.

3. **Toledo and Springfield Show Moderate Average Downtime:**

   o Toledo (151.3 minutes from Scotfase) and Springfield (150 minutes from J-lax) have moderate downtime levels compared to other cities.

4. **Cities with Lower Average Downtime:**

   o Cities like Grand Rapids (3.5 minutes) and Gary (38 minutes) show much lower downtime, with vendors Funlax and Solholdings, respectively, indicating fewer operational delays in these locations.

Statistical Evidence:

- Kalamazoo leads with 999 minutes of average downtime, which is nearly three times that of Milwaukee (288 minutes) and Lansing (360 minutes).

- Toledo and Springfield contribute 151.3 and 150 minutes, respectively, placing them in the mid-range.

- Gary (38 minutes), Rockford (44.8 minutes), and Grand Rapids (3.5 minutes) show minimal average downtime.

Interesting Discoveries:

1. **Kalamazoo's Exceptionally High Average Downtime:**

   o The 999 minutes of downtime in Kalamazoo is a significant outlier, indicating potential operational challenges tied to the vendor Funlax.

2. **Notable Gaps in Average Downtime Between Cities:**

- A substantial difference exists between cities with high downtime, like Kalamazoo, and those with minimal downtime, such as Grand Rapids, suggesting variations in vendor efficiency or operational conditions.

3. **Similar Average Downtime Across Multiple Cities:**

- Cities such as Monon, Indianapolis, Cincinnati, and Davenport all show 120 minutes of downtime, which may indicate uniform operational issues or similar downtime response times across vendors.

Total Downtime Minutes Across Plants:

Total Downtime Mins across Plants
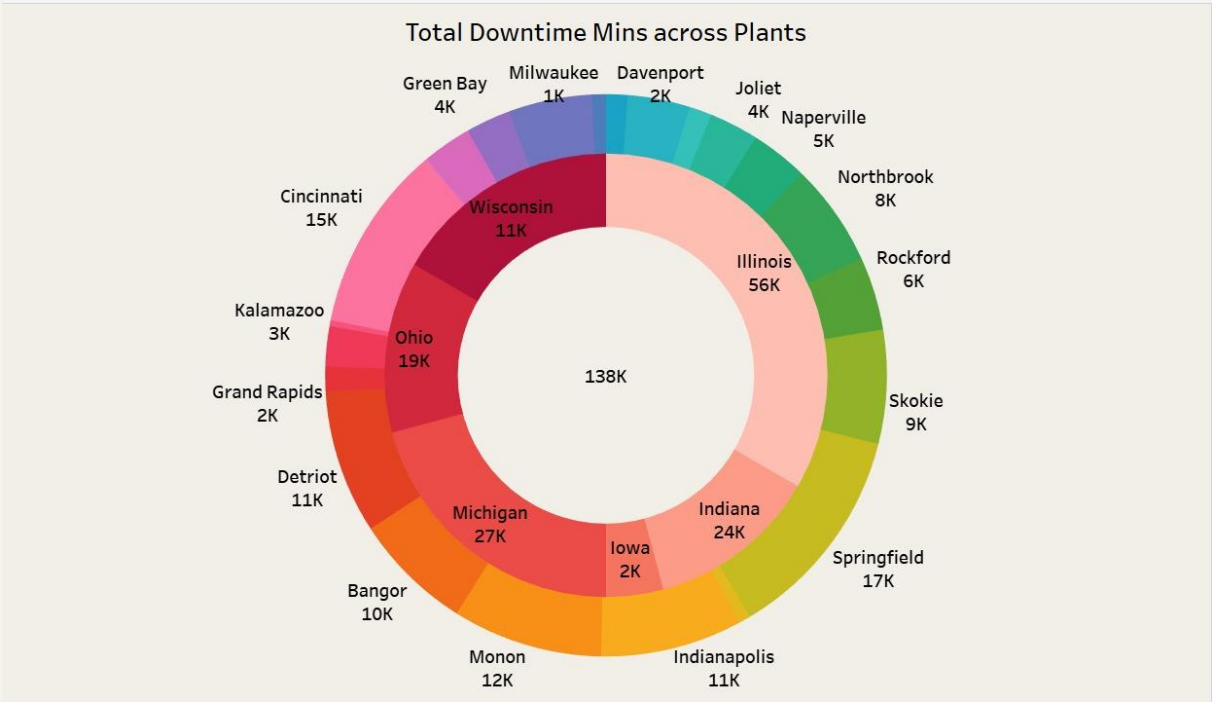
Key Findings:

1. **Illinois Has the Highest Downtime:**

- Illinois leads in downtime, with a total of 56K minutes, making it the largest contributor by a significant margin compared to other states.

2. **Michigan and Indiana Show Notable Downtime:**

- Michigan contributes 27K minutes of downtime, ranking second overall.

- Indiana follows closely with 24K minutes, making it the third-highest state in downtime contribution.

3. **Ohio and Springfield Report Moderate Downtime:**

   o Ohio shows 19K minutes of downtime, while Springfield City contributes 17K minutes, representing a moderate level of downtime compared to Illinois and other high-contributing states.

4. **Cities with Lower Downtime:**

   o Locations such as Milwaukee (1K minutes), Davenport (2K minutes), and Kalamazoo (3K minutes) have significantly lower downtime, indicating fewer operational delays.

Statistical Evidence:

- Total downtime across all plants amounts to 138K minutes, with Illinois contributing 56K minutes, followed by Michigan with 27K and Indiana with 24K minutes.
- Ohio and Springfield City contribute 19K and 17K minutes, respectively.
- On the lower end, Milwaukee, Davenport, and Kalamazoo cities report minimal downtime, ranging from 1K to 3K minutes.

Interesting Discoveries:

1. **Illinois as a Major Contributor:**

   o Illinois accounts for nearly 41% of the total downtime, with 56K minutes, marking it as a significant outlier compared to other locations.

2. **High Downtime Spread Across Multiple Locations:**

   o In addition to Illinois, Michigan, Indiana, and Ohio show considerable downtime, indicating that operational challenges are distributed across several states.

3. **Low Downtime in Certain Locations:**

   o Cities like Milwaukee, Davenport, and Kalamazoo exhibit minimal downtime (less than 3K minutes), suggesting these plants are operating more efficiently or face fewer disruptions.

Phase 3. Data Analysis & Visualization

Monthly Downtime: Actuals (2013-2014) with 2015 Forecast

Monthly Downtime: Actuals (2013-2014) with 2015 Forecast
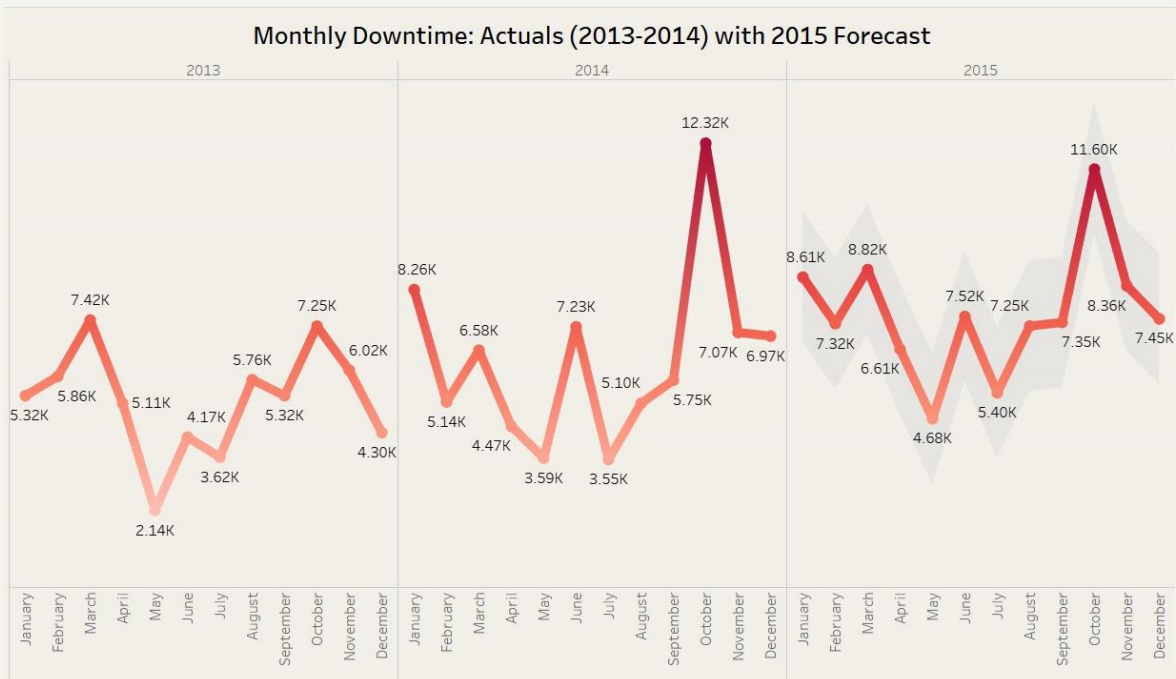
Key Findings:

1. **Significant Downtime Peaks in October 2014 and 2015:**
   o October 2014 shows the highest actual downtime, reaching 12.32K minutes.
   o The forecasted peak for October 2015 is 11.60K minutes, indicating similar downtime trends in the fall months.

2. **Recurring Monthly Downtime Trends:**
   o Downtime decreases significantly in February and September across the years.
   o February 2013 recorded 5.86K minutes, while February 2014 saw 5.40K minutes, showing a repeating spike in early-year downtime.

3. **Lower Downtime in Early Summer:**
   o The lowest downtime consistently occurs in the summer months, with the lowest points in May 2013 (2.14K minutes) and May 2014 (3.55K minutes).

4. **Steady Downtime Decline Towards the End of 2014:**
   o After the peak in September 2014, downtime declined steadily, reaching 6.97K minutes by December 2014.

Statistical Evidence:

- The highest actual downtime was in October 2014 with 12.32K minutes, while October 2015 is forecasted to reach 11.60K minutes.

- Downtime during summer is significantly lower, with May 2013 at 2.14K minutes and May 2014 at 3.55K minutes.
- Recurring early-year spikes are seen in March 2013 (7.42K minutes) and January 2014 (8.26K minutes).

Interesting Discoveries:

1. **Recurring Seasonal Trends:**
   - The consistent spikes in downtime during March and October indicate potential seasonal factors influencing operations, such as external disruptions or resource limitations.
2. **Forecast Similarities to Past Trends:**
   - The 2015 forecast follows the same pattern of downtime spikes in fall and dips in summer, aligning closely with actual data from 2013-2014, suggesting consistent operational patterns.

Phase 4. Insights

5. Insights and Recommendations

Key Findings and Conclusions:

1. **Persistently High Defect Rates Across Multiple Vendors**:
   - Vendors such as **Solholdings** and **PlanetHouse** showed consistently high defect rates throughout Q4. This indicates ongoing quality issues that pose a significant risk to overall production performance.
   - **Conclusion**: These vendors are underperforming, and their sustained high defect rates highlight the need for immediate intervention.
2. **Critical Defect Types**:
   - **Rejected Defects**: Vendors like **PlanetHouse**, **Vaiazozice**, and **Dentocity** reported high levels of rejected defects, indicating significant quality control failures that directly impact production.
   - **Impacted Defects**: **Yearlex**, **Ontotam**, and **Dentocity** consistently experienced high defect rates in categories that led to direct operational disruptions.
   - **Non-Impacted Defects**: Vendors like **Quotelane** and **Dentocity** recorded elevated non-impacted defect counts, which, while less critical, indicate inefficiencies in their processes.
   - **Conclusion**: Addressing highly rejected and impacted defect types should be a priority as these defects have the greatest impact on operational efficiency. Non-impacted defects, while less critical, still require attention to eliminate inefficiencies.

3. **Seasonal and Sub-Category Trends**:
   - A noticeable spike in defects occurred in **October 2014**, particularly in the **Electrical** and **Mechanical** sub-categories, which were most problematic for vendors like **Ontotam** and **Fax-ex**.
   - **Logistics** also proved to be a widespread issue, with several vendors (e.g., **Sanlab, Quotelane,** and **Reddoit**) showing high defect rates in this category.
   - **Conclusion**: These sub-categories are particularly prone to quality issues, requiring more focused quality control efforts, especially during high-risk periods like Q4.
4. **Vendor Improvement**:
   - Vendors such as **Tampquote**, **Instrip**, **Plustax**, and **Sanin** showed notable improvement in impacted defect type rates during Q4, reflecting the positive impact of corrective measures and better-quality management.
   - **Conclusion**: Continued application of these quality control measures will be essential to maintain progress, and these vendors could serve as models for others.

Phase 4. Recommendations

Actionable Recommendations

1. **Vendor Contract Review**:
   - **Solholdings** and **Planethouse** should undergo a thorough contract review. Given their persistently high defect rates, renegotiating or even reconsidering their contracts may be necessary to mitigate the ongoing quality risks.
   - **Action**: Focus on developing vendor performance improvement plans, with clear defect reduction targets and tighter monitoring protocols.
2. **Targeted Quality Control Interventions**:
   - **Planethouse**, **Vaiazozice**, **Yearlex**, **Ontotam**, and **Dentocity** need targeted interventions to address their high rejected and impacted defect rates. These measures should focus on root cause analysis and corrective actions.
   - **Action**: Implement stricter quality monitoring and develop targeted improvement initiatives for vendors with high defect rates in critical categories (e.g., Electrical, Mechanical, Logistics).
3. **Proactive Seasonal Quality Measures**:
   - The spike in defects in October suggests potential seasonal factors affecting vendor performance. Proactive quality control measures should be introduced in advance of high-risk periods like Q4.
   - **Action**: Establish preemptive monitoring and inspection protocols during these periods to minimize defect spikes.

4. **Ongoing Monitoring for Improvement**:
    o Vendors like **Tampquote**, **Instrip**, **Plustax**, and **Sanin** demonstrated significant improvement, showing that corrective measures can be effective when implemented properly.
    o **Action**: Continue monitoring these vendors to ensure sustained performance. Share best practices from these vendors with other underperforming suppliers.

---

6. Limitations and Future Work

Phase 5. Limitations

Limitations

- **Data Quality and Missing Values**

A key limitation in the dataset is the absence of data on the total number of products produced by each vendor. While the dataset provides the Defect Qty for each vendor, it does not allow for the calculation of defect rates, which are essential for a fair assessment of vendor quality. Defect rates, calculated as the percentage of defective items relative to total production, are a crucial metric for benchmarking vendors. Without this data, it becomes difficult to compare vendors effectively and identify the most problematic suppliers, as we cannot account for differences in production volumes.

Additionally, the dataset lacks comprehensive data on the costs associated with rejected products and downtime. Rejected products carry significant financial implications, including material waste and the costs of rework or replacement. Downtime, similarly, can result in lost production time, missed deadlines, and increased operational expenses. However, incomplete downtime data in the dataset limits the ability to estimate the full economic impact of these defects on production, making it challenging to prioritize vendor improvements based on financial costs. Moreover, while the dataset contains a reference to a missing dimension, Material, its absence further limits the ability to analyze material-specific defect trends, though we are unable to determine exactly what this dimension is associated with or how it influences overall defect rates.

- **Lack of Root Cause Data:**

While the dataset contained information on defect quantities and categories (e.g., rejected, impacted, and non-impacted defects), it did not provide data on the root causes of these defects. This restricts the ability to propose targeted

recommendations for defect reduction, as understanding the underlying reasons behind vendor underperformance is essential for crafting effective improvement strategies.

- **Limited Time Range:**
  The dataset appears to focus on a relatively short timeframe (2013-2014). This limits the ability to detect long-term trends in vendor performance or seasonal variations in defect rates. A longer dataset would provide greater insights into whether certain vendors consistently perform poorly or improve over time.
- **Vendor and Subcategory Context:**
  While the dataset provides information on vendors, subcategories, and defect types, it lacks more detailed context, such as vendor capacity, production volume, or the criticality of certain subcategories. These factors could help provide more targeted recommendations by highlighting which vendors or materials are the most vital to operations.

Phase 5. Future Work

Future Work

1. **Root Cause Identification**: Collecting data on the underlying causes of defects will help implement more targeted corrective measures. Collaborating with vendors to identify process failures or material issues will allow for more effective defect reduction strategies.
2. **Long-Term Vendor Performance Tracking**: Expanding the time range of data collection will enable better detection of trends, seasonal variations, and long-term performance improvements or declines. This continuous tracking will help address issues before they escalate.
3. **Incorporating Vendor and Material Criticality**: Future studies should focus on identifying which vendors and materials are most critical to production. Prioritizing improvements for high-impact vendors will help mitigate risks that can significantly affect operations.
4. **Collecting Data on Total Production Volumes**: Gathering information on the total number of products produced by each vendor will allow for the calculation of defect rates. This will provide a fairer comparison of vendor performance.
5. **Financial Impact Analysis**: Understanding the costs of rejected products and downtime is crucial for prioritizing vendor improvements based on economic impact. Capturing this data will help optimize resources and improve vendor negotiations.

7. Conclusion

The analysis of manufacturing defects across vendors has identified key areas where quality issues persist and where significant improvements have been made. Vendors such as Solholdings and PlanetHouse consistently displayed high defect rates, contributing to significant operational inefficiencies and downtime. These vendors require immediate attention to address their quality control failures. In contrast, vendors like Tampquote, Instrip, and Plustax have shown measurable improvements in defect rates, reflecting the effectiveness of their corrective actions and quality management practices.

Key defect types such as Rejected Defects and Impacted Defects have been found to heavily contribute to production inefficiencies, particularly in subcategories like Electrical, Mechanical, and Logistics. Addressing these high-defect categories is crucial for reducing downtime and enhancing overall production quality.

Key Recommendations

- Target High-Defect Vendors: Immediate interventions should focus on high-defect vendors, such as Solholdings and PlanetHouse, to improve their quality performance and reduce the operational impact of their defects.
- Enhance Data Collection: Future data collection should focus on capturing total production volumes, financial impacts, and the root causes of defects. This will enable the calculation of defect rates and allow for a more precise prioritization of improvements.
- Sustained Monitoring: Vendors who have shown improvements, such as Tampquote and Plustax, should continue to be monitored to ensure sustained performance, while their best practices should be shared with underperforming vendors.

Importance of the Project

This project has laid the foundation for making data-driven decisions regarding vendor management and quality control. By focusing on the most problematic vendors and defect types, the organization can significantly reduce defect rates, minimize downtime, and improve production efficiency. Implementing the recommended actions will have a substantial impact on product quality, operational costs, and overall supply chain reliability, leading to long-term success and competitiveness.