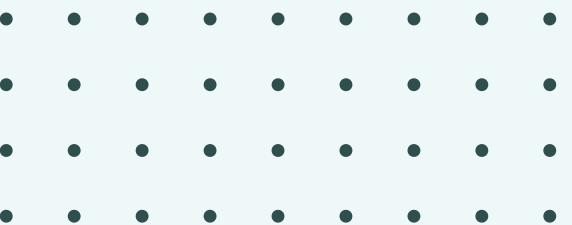


# Supplier Quality Analysis



# Problem Overview

- **Benchmark** supplier quality by evaluating the number and type of defects across various materials.
- **Identify** which suppliers or materials are consistently underperforming.
- **Provide** actionable insights for improving supplier selection and quality control processes.



# OUR TEAM



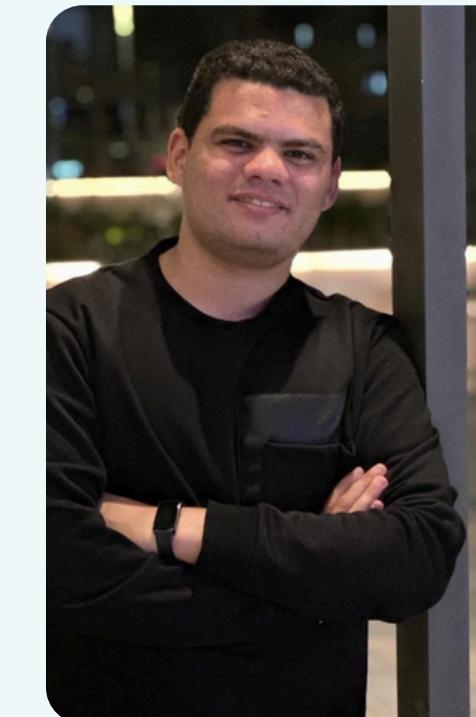
**Mostafa ALi**



**Mostafa Aly**



**Karim Abdelhamed**



**Mohamed Haitham**



**Sayed Eladawy**

**01**

Benchmark  
Supplier  
Quality

**02**

Find  
Weak  
Suppliers

**03**

Improve  
Quality  
Control

## Analysis Goals

**04**

Analyze  
Defect  
Trends

**05**

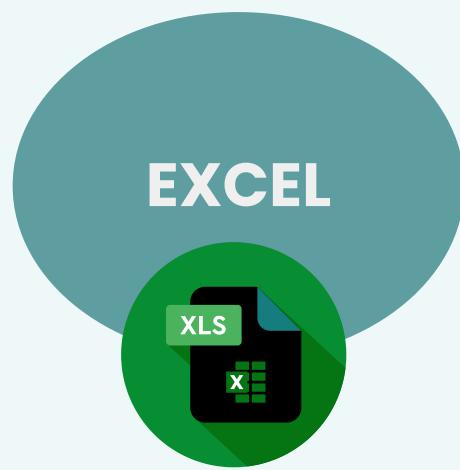
Optimize  
Supplier  
Management

**06**

Enhance  
Supply  
Chain  
Efficiency

# Process





• • • • •

**Material ID**

**Plant ID**

**Defect Type ID**

**Material Type ID**

**Defect QTY**

**Vendor ID**

**Sub Category ID**

**Defect ID**

**Downtime (min)**

**Date**

**10  
Columns**

**6,145 rows**

# Data Exploration

## Data Characteristics

### Date:

- Dates are available for use in analysis.

### Defect Quantity & Downtime:

- Quantitative data for defects and downtime in integer and float formats.

### ID Columns:

- Identification columns linking vendors, materials, and defect types.

• • • • •  
• • • • •  
• • • • •

# Data Exploration

## Data Quality

### **Missing Values:**

- 1 missing value in the Downtime column; all other fields are complete.

### **Weird Record:**

- This weird record has no Downtime minutes and contains all the missing references.

### **Defect Quantities:**

- Significant range with extremes (e.g., max defect quantity of 42,275), requiring closer examination for outlier handling.

• • • • •  
• • • • •  
• • • • •

# Data Exploration

## Data Quality

### Duplicates:

- Found duplicates in the Vendor and Defects dimensions.

### Plant Column Issue:

- Plant column in the Plant dimension has two values separated by a comma (",").

### Material ID:

- Material ID column in the Defected Items Fact table doesn't have reference data for each material.

• • • • •  
• • • • •  
• • • • •

# Data Exploration

## Data Quality

### Date Format:

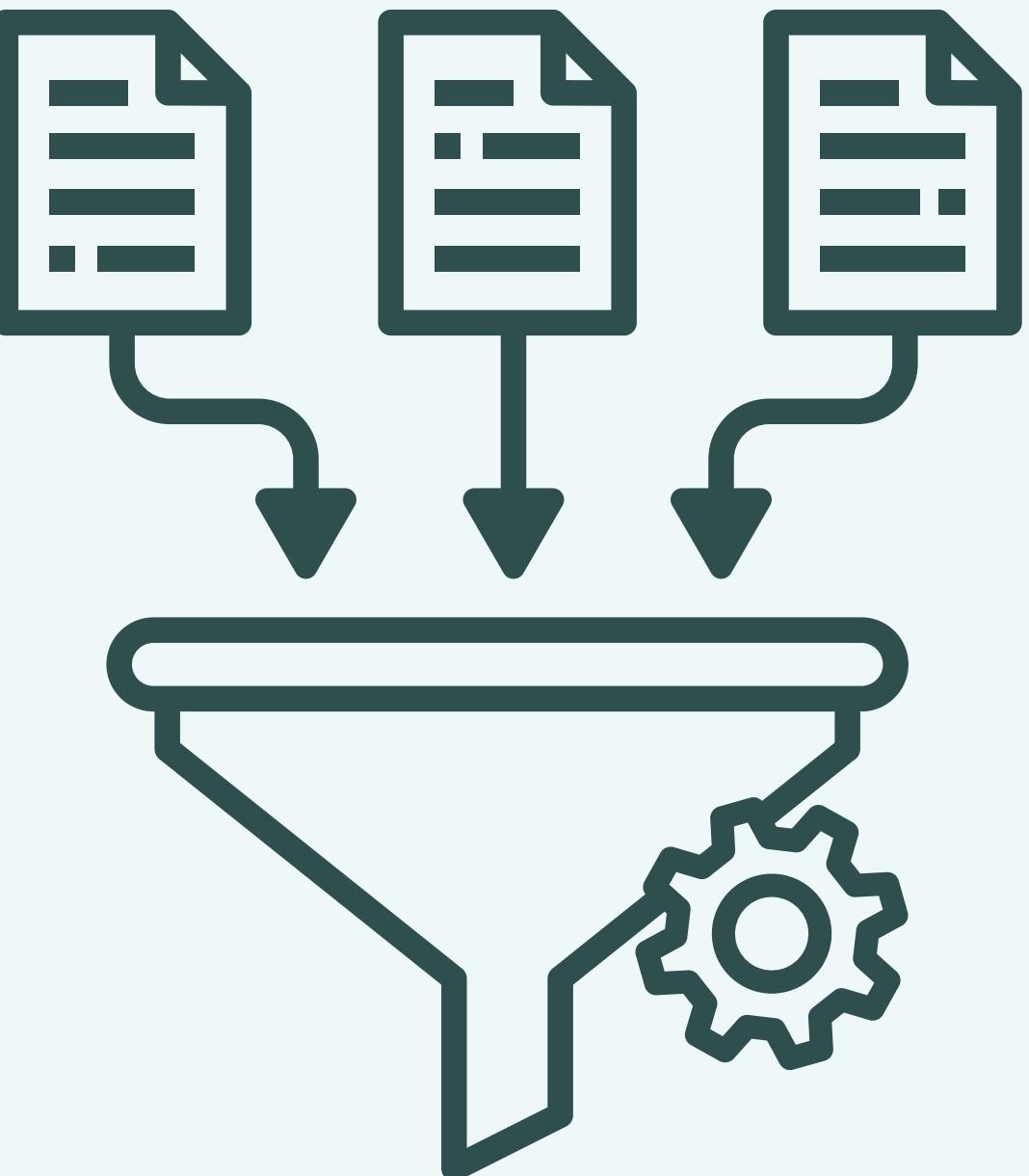
- Date column in Defected Items Fact is stored as a timestamp, but there are no changes in hours, minutes, and seconds.

### Unreferenced IDs:

- Vendor ID "810" is missing from the Vendor dimension.
- Defect Type ID "8" is missing from the Defect Type dimension.
- Material Type ID "69" is missing from the Material Type dimension.
- These unreferenced IDs are all found in one record without any repeats in other records.

• • • • •  
• • • • •  
• • • • •

# Cleaning



# Data Cleaning

## Removals

- Found and removed duplicates in the Vendor dimension.

```
# drop the duplicates without the first occurrence from vendor  
  
vendor = vendor.drop_duplicates(subset='Vendor', keep='first')
```

- Found and removed duplicates in the Vendor dimension.

```
defects = defects.drop_duplicates(subset='Defect', keep='first')
```

• • • • •  
• • • • •  
• • • • •

# Data Cleaning

## Removals

- Removed the "weird" record that has all the incorrect data.

```
# Delete row that have Defect Type ID = 8 in defected_items  
  
defected_items = defected_items[defected_items['Defect Type ID'] != 8]
```

• • • • •  
• • • • •  
• • • • •

# Data Cleaning

## Replacements

- Found and replaced all records in the Defected Items Fact that have duplicate Vendor IDs with the unique Vendor ID.

```
# Step 1: Identify the first occurrence of each duplicated vendor
first_occurrences = vendor.drop_duplicates(subset='Vendor', keep='first')

# Step 2: Create a mapping of all vendor names to their first occurrence Vendor ID
vendor_mapping = dict(zip(vendor['Vendor'], vendor['Vendor ID']))

# Step 3: Create a reverse mapping for duplicates (replace duplicate Vendor IDs with first occurrence)
for vendor_name in first_occurrences['Vendor']:
    first_id = first_occurrences[first_occurrences['Vendor'] == vendor_name]['Vendor ID'].values[0]
    duplicate_ids = vendor[vendor['Vendor'] == vendor_name]['Vendor ID'].tolist()

    # Update the mapping for all duplicate IDs
    for dup_id in duplicate_ids:
        vendor_mapping[dup_id] = first_id

# Step 4: Apply this mapping to update the 'Vendor ID' in defected_items DataFrame
defected_items['Vendor ID'] = defected_items['Vendor ID'].replace(vendor_mapping)
• • • • •
• • • • •
• • • • •
```

# Data Cleaning

## Replacements

- Found and replaced all records in the Defected Items Fact that have duplicate Defect IDs with the unique Defect ID.

```
# Step 1: Identify the first occurrence of each duplicated vendor
first_occurrences_defects = defects.drop_duplicates(subset='Defect', keep='first')

# Step 2: Create a mapping of all vendor names to their first occurrence Vendor ID
defects_mapping = dict(zip(defects['Defect'], defects['Defect ID']))

# Step 3: Create a reverse mapping for duplicates (replace duplicate Vendor IDs with first occurrence)
for defect_name in first_occurrences_defects['Defect']:
    first_id = first_occurrences_defects[first_occurrences_defects['Defect'] == defect_name]['Defect ID'].values[0]
    duplicate_ids = defects[defects['Defect'] == defect_name]['Defect ID'].tolist()

    # Update the mapping for all duplicate IDs
    for dup_id in duplicate_ids:
        defects_mapping[dup_id] = first_id

# Step 4: Apply this mapping to update the 'Vendor ID' in defected_items DataFrame
defected_items['Defect ID'] = defected_items['Defect ID'].replace(defects_mapping)
. . . . .
. . . . .
. . . . .
```

# Data Cleaning

## Updates

- Updated the Plant table by separating the Plant column into two columns: Plant and State.

```
# seperate the Plant column in plant dataframe to 2 columns by comma , and drop the Plant column  
  
plant[['Plant', 'state']] = plant['Plant'].str.split(',', expand=True)  
  
plant
```

• • • • •  
• • • • •  
• • • • •

# Data Cleaning

## Updates

- Added three more columns to the Plant table: State Name (derived from the State column), Longitude, Latitude.
- Changed the Date column type in the Fact table from timestamp to just date.

```
plant['State Name'] = plant['State'].str.strip().str.upper().map(  
    {'MI': 'Michigan',  
     'WI': 'Wisconsin',  
     'IL': 'Illinois',  
     'IN': 'Indiana',  
     'OH': 'Ohio',  
     'IA': 'Iowa'})
```

```
plant
```

• • • • •

```
plant['Longitude'] = plant['State Name'].map(  
    {'Michigan': -84.506836,  
     'Wisconsin': -89.500000,  
     'Illinois': -89.000000,  
     'Indiana': -86.126976,  
     'Ohio': -82.996216,  
     'Iowa': -93.581543})  
plant['Latitude'] = plant['State Name'].map(  
    {'Michigan': 44.182205,  
     'Wisconsin': 44.500000,  
     'Illinois': 40.000000,  
     'Indiana': 40.273502,  
     'Ohio': 40.367474,  
     'Iowa': 42.032974})
```

```
plant
```

# Data Cleaning

## Additions

- Created a Calendar dimension to help in visualizing data containing six columns: Date, Year, Quarter, Month, Name of Month, Day.

```
# Define the date range
date_range = pd.date_range(start='2012-01-01', end='2024-12-31', freq='D')

# Create a DataFrame
calendar_df = pd.DataFrame({
    'Date': date_range,
    'Year': date_range.year,
    'Quarter': date_range.quarter,
    'Month': date_range.month,
    'Name of Month': date_range.strftime('%B'), # Full month name
    'Day': date_range.day
})
...
...
...
...
```

# Data Cleaning

## Additions

- Created a Group By table to help in the benchmark story between vendors

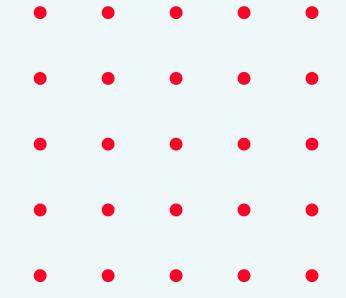
```
# Group the data by Vendor ID
grouped_by_vendor = defected_items.groupby('Vendor ID')

# You can also aggregate the data, for example, summing up defect quantities and downtime minutes
aggregated_data = grouped_by_vendor.agg({
    'Defect Qty': 'sum',
    'Downtime min': 'sum'
}).reset_index()

# Display the aggregated data
aggregated_data
```

• • • • •  
• • • • •  
• • • • •

# Our Insights



# Suppliers Quality Dashboard



Next

# Defected Items

Qty

Total Downtime..

Max. Downtime ..

# of Vendors

# of Defect

6,144

36,942,501

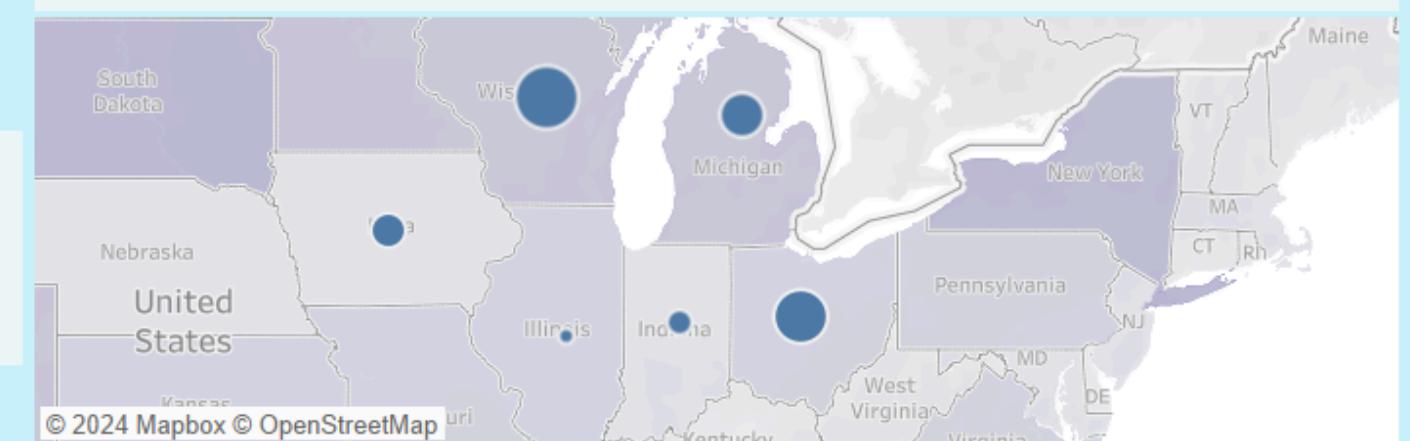
87,121

90

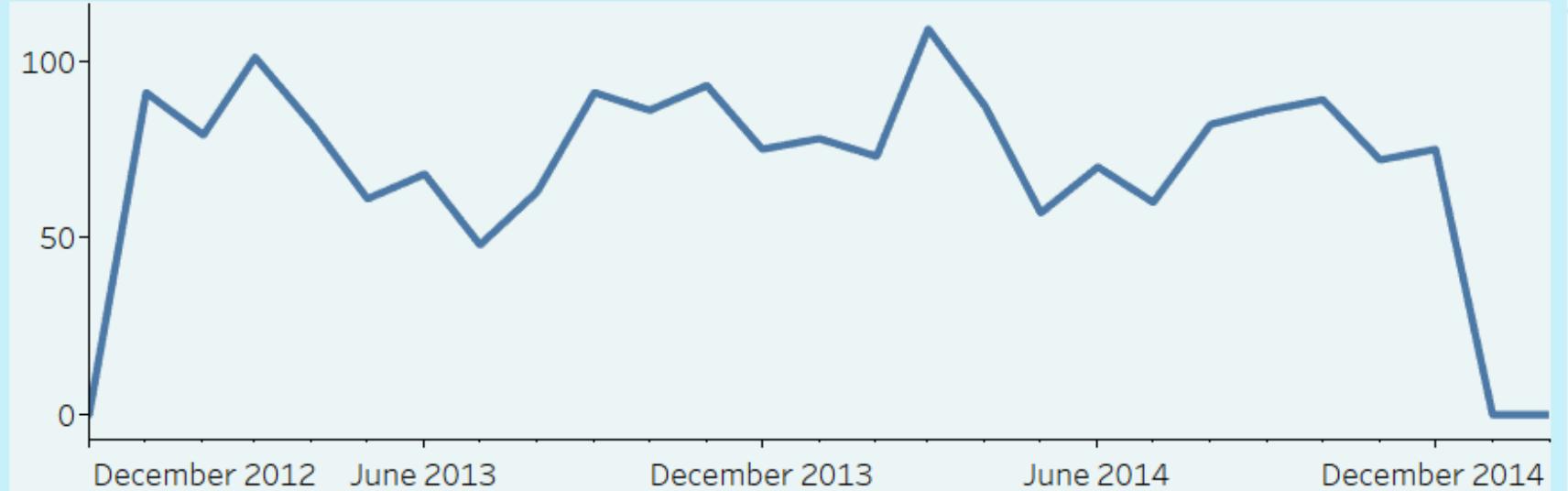
326

268

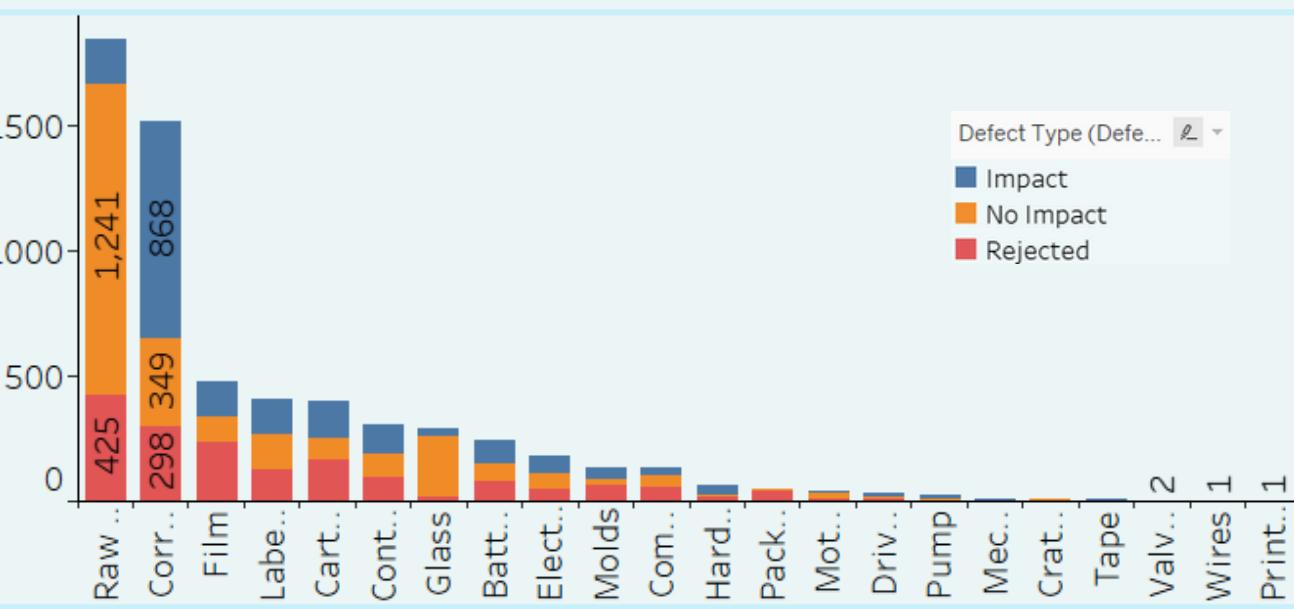
## Total Qty Per Plants Geo



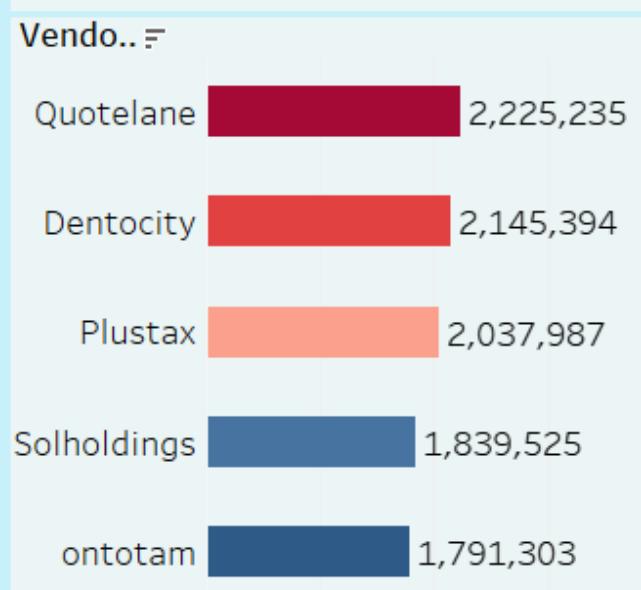
### No. of Defects Over Time



### Defect Type by Material Type



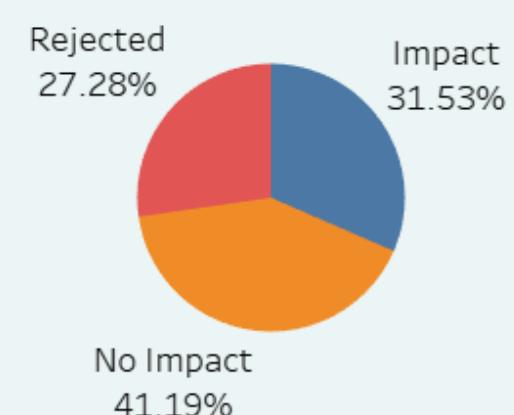
### Total Qty Per Vendor



### Total Quantity Over Years



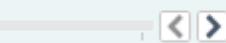
### % Of Defected Items



### # Of Defected Items With Total QTY



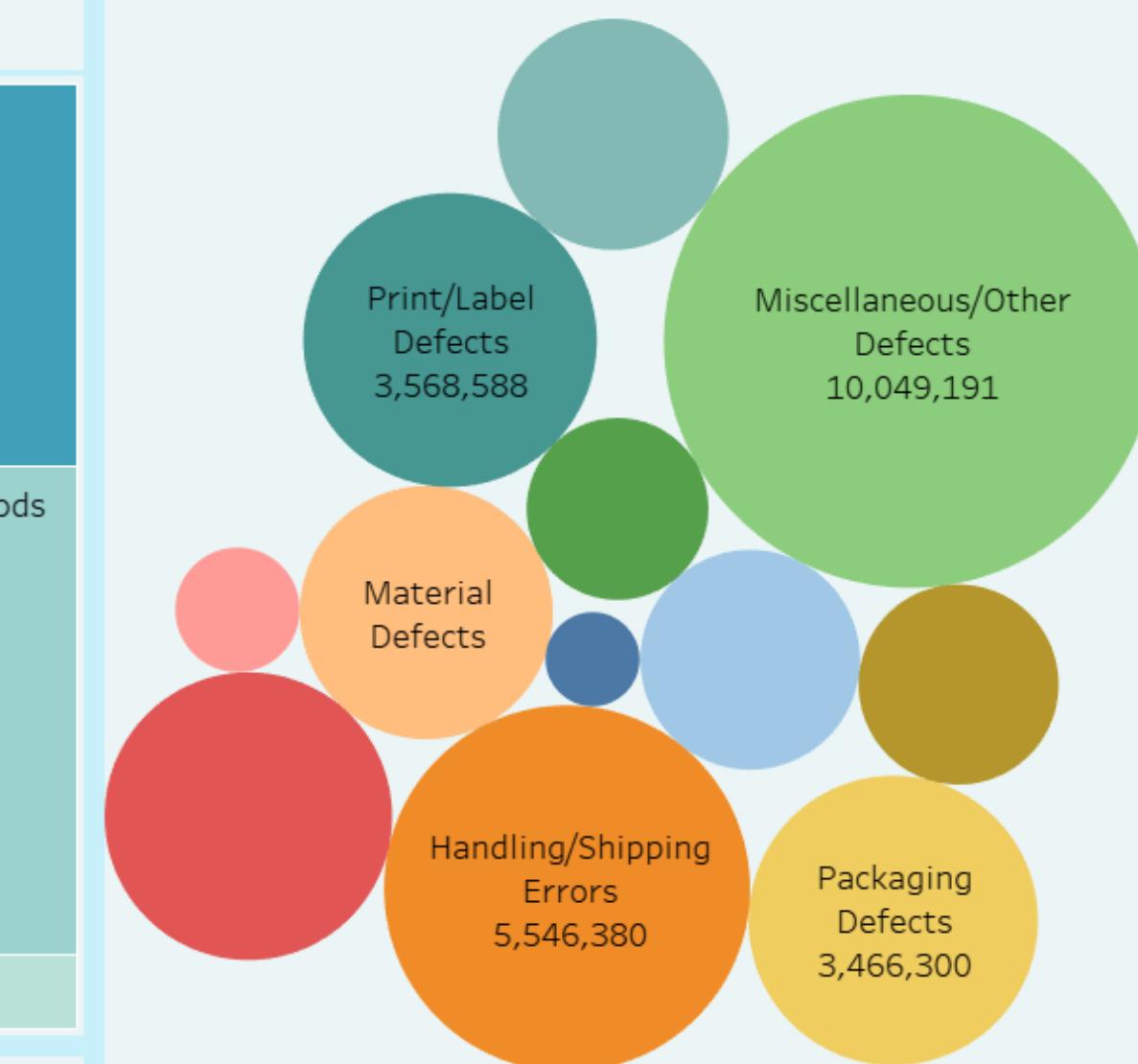
## Avg. DownTime mins Over Years

2013 

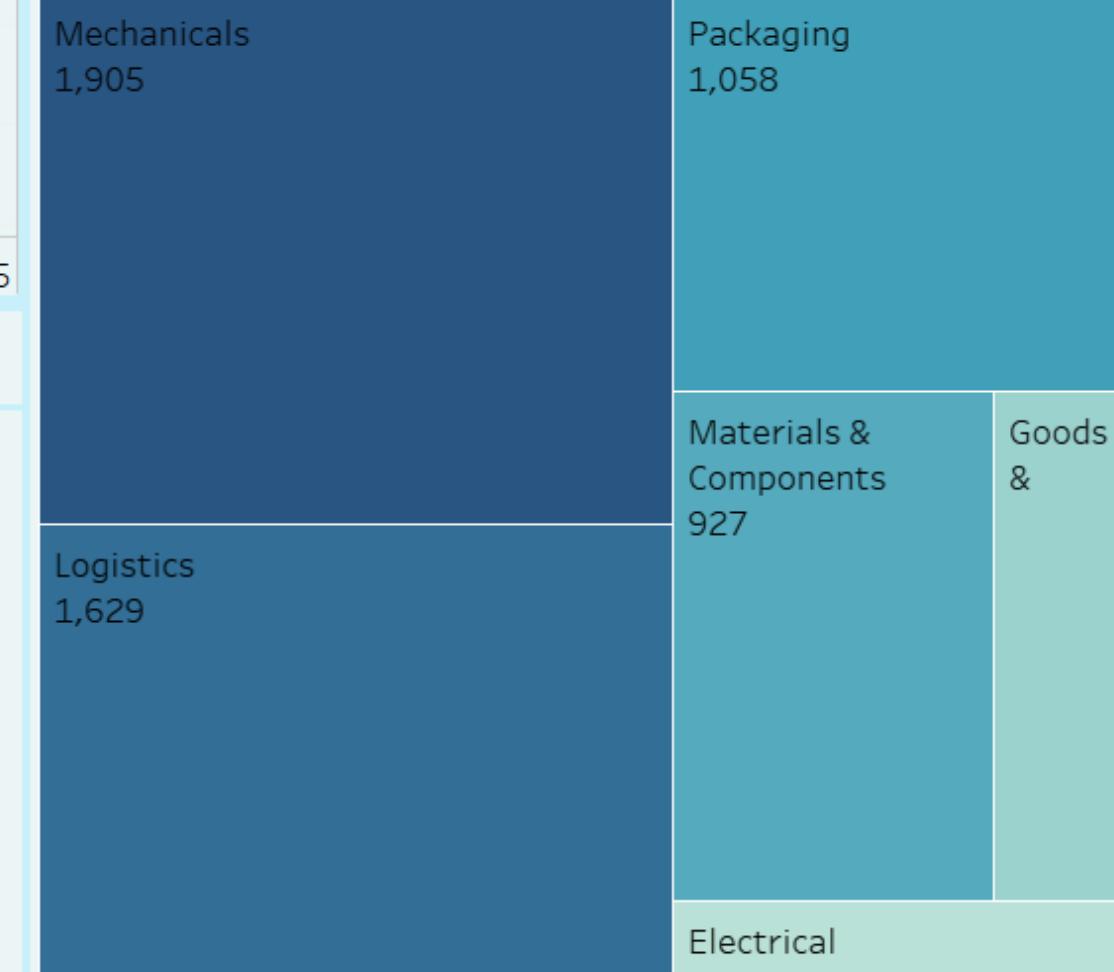
Previous

Next

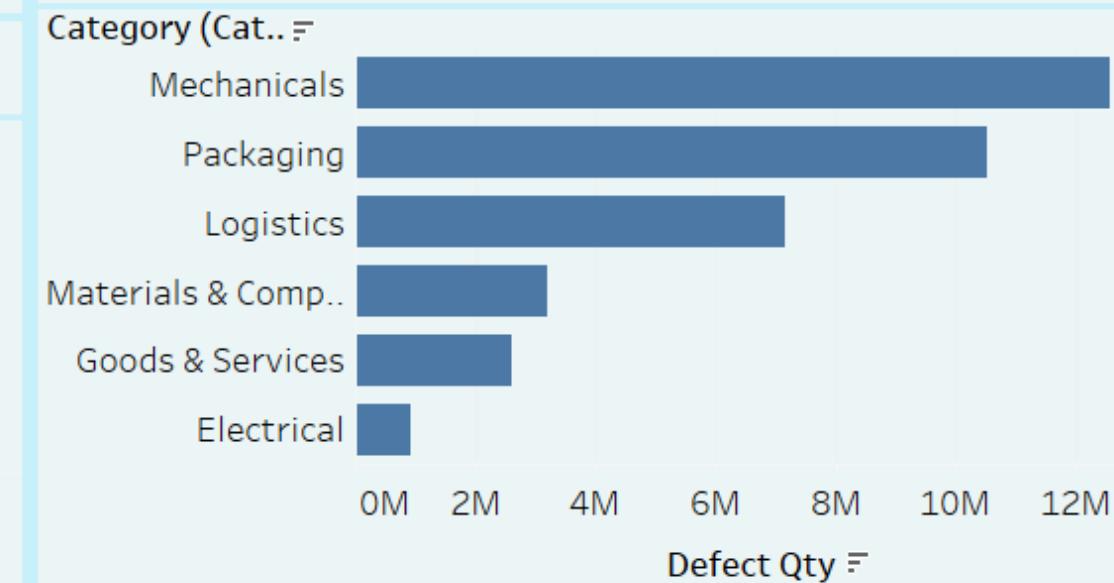
## Qty For Each Defect Type



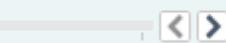
## # OF Defected Items By Category



## The Most Material Category Per QTY

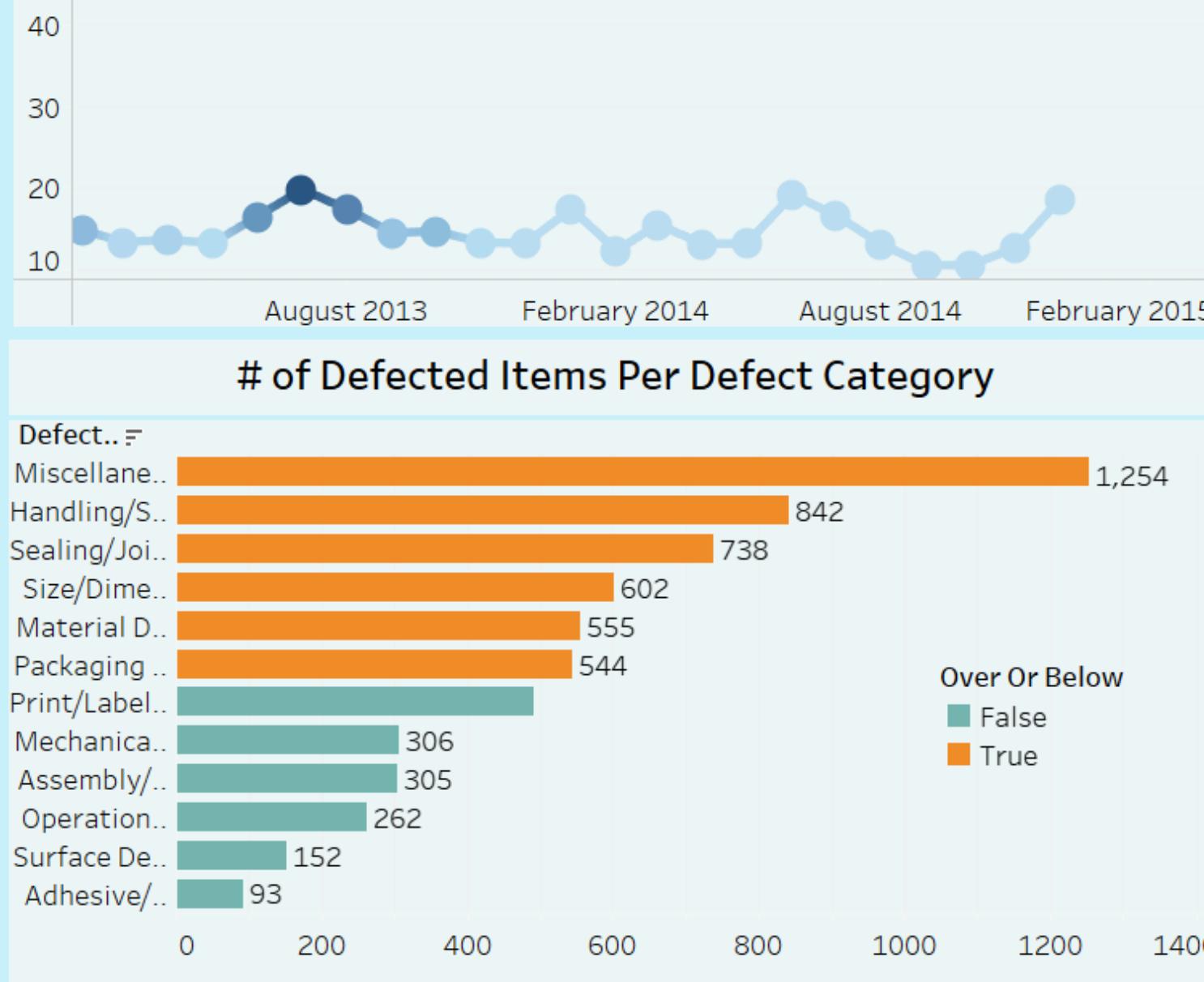


## Avg. DownTime mins Over Years

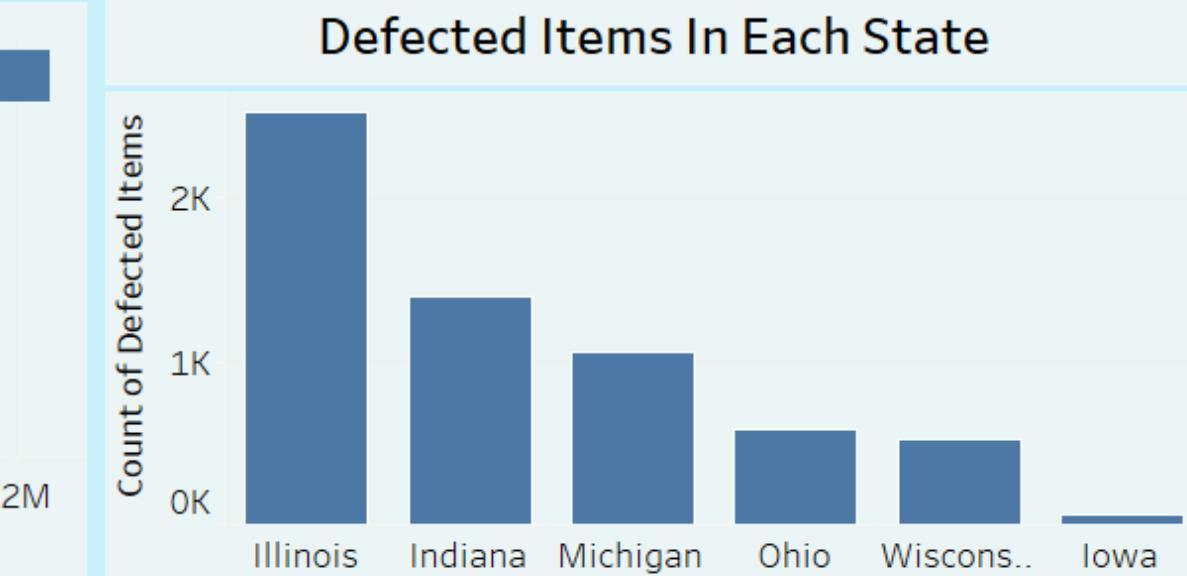
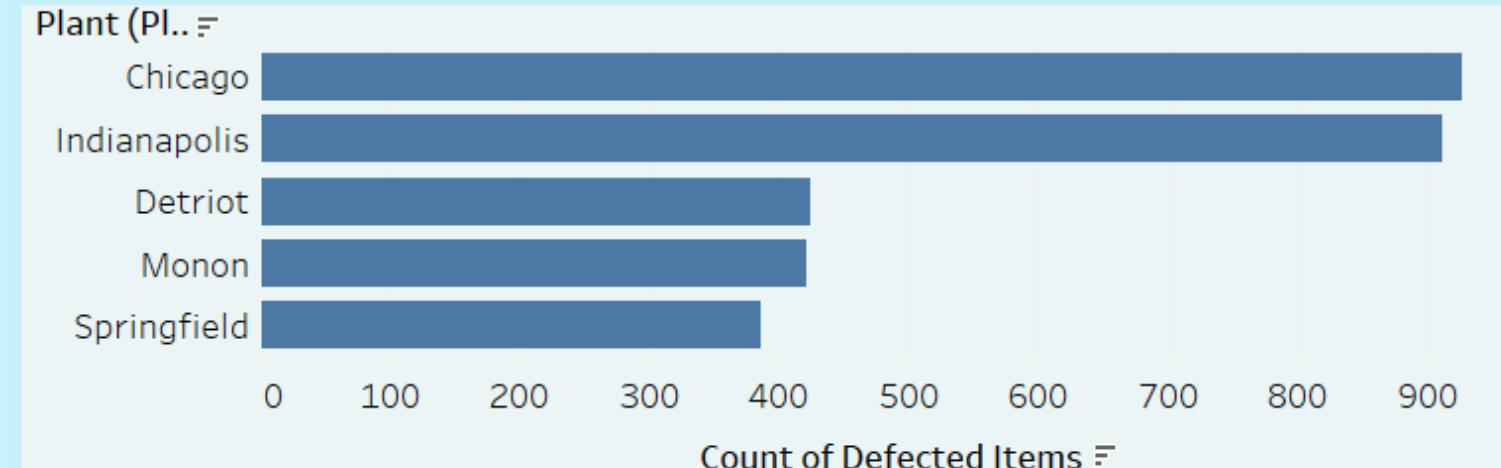
2013 

Previous

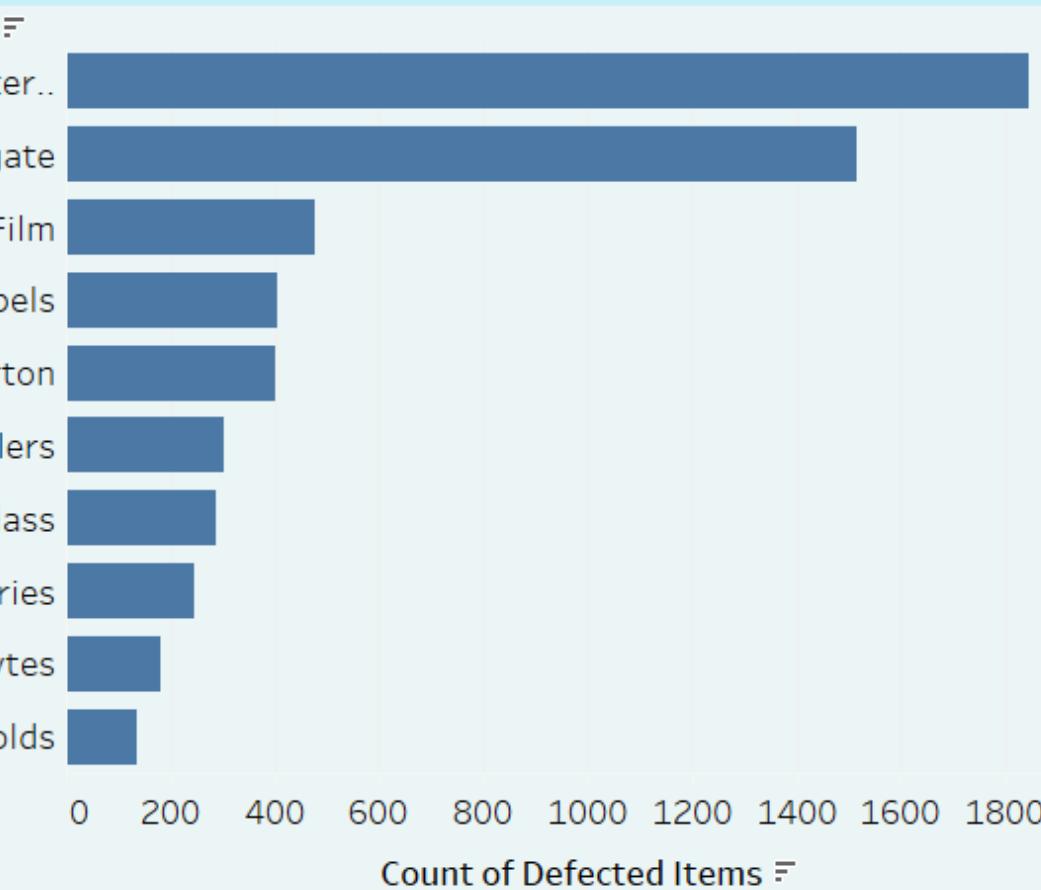
Next



## Top 5 Plants in # of Defected Items

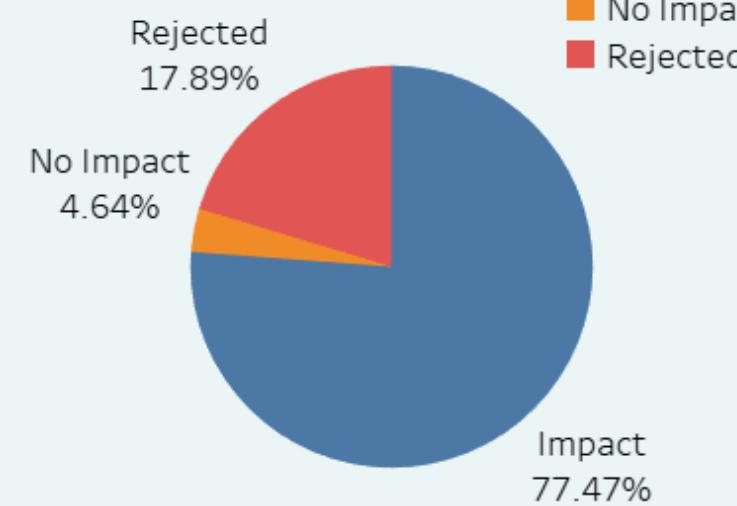


## # Of Defected Items Per Material Type

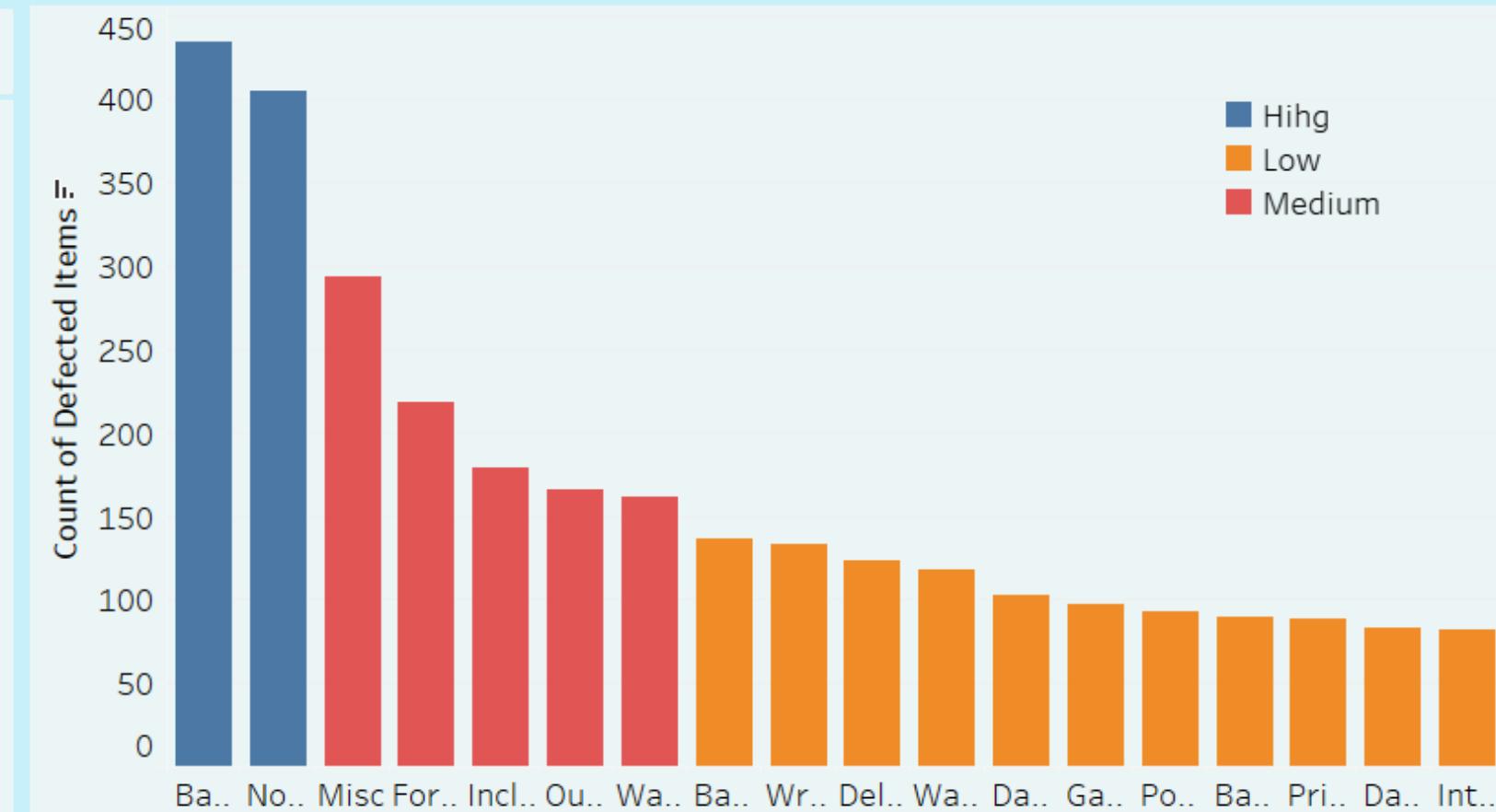


Previous

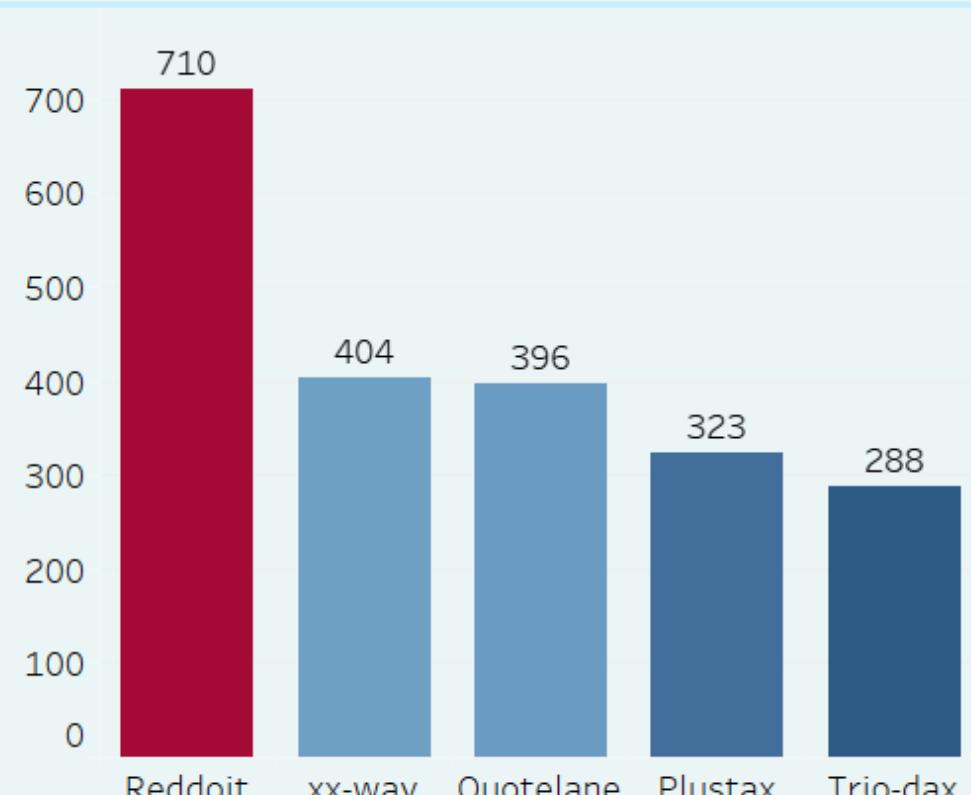
## AVG. DownTime



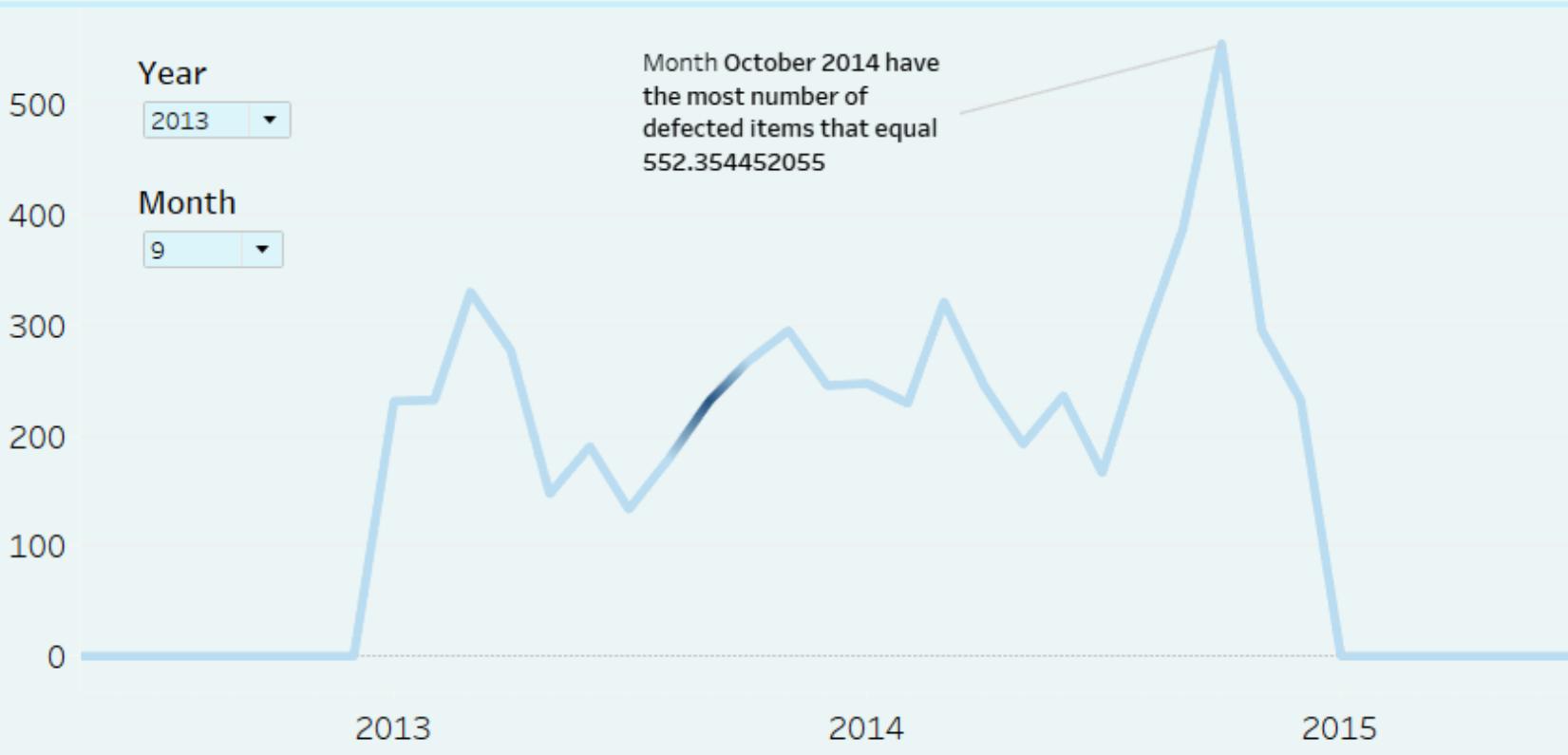
## Defect Severity



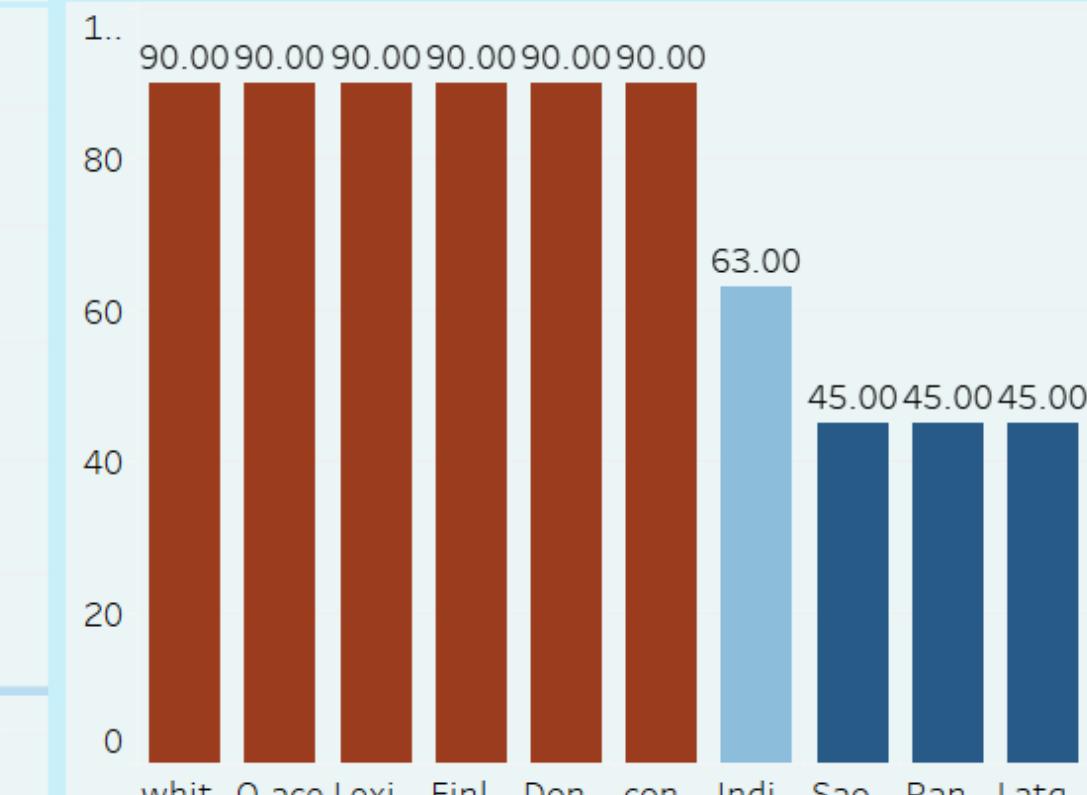
## The Top Worst 5 Vendors



## # of Defect Items Over Months



## Top 10 Vendors





Total Defected QTY

**36 M**

Total DownTime (hr)

**1453 h**

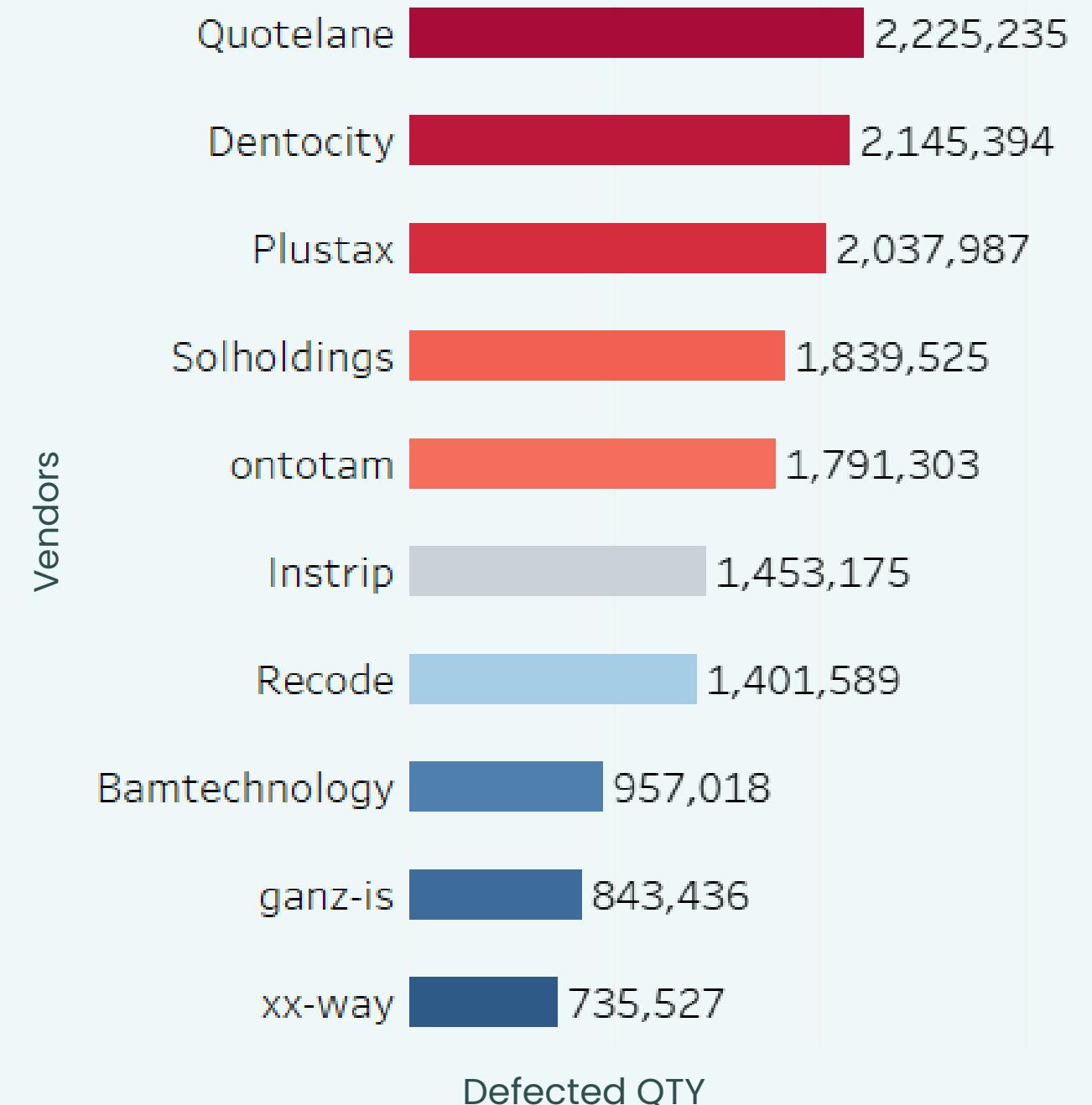
Nom Of Defected Items

**6144**

# Vendors

- **Quotelane** has the highest total quantity of defective items, surpassing all other vendors in terms of overall defects.
- The final vendor in the top 5 worst performers is **Ontotam**, rounding out the group.

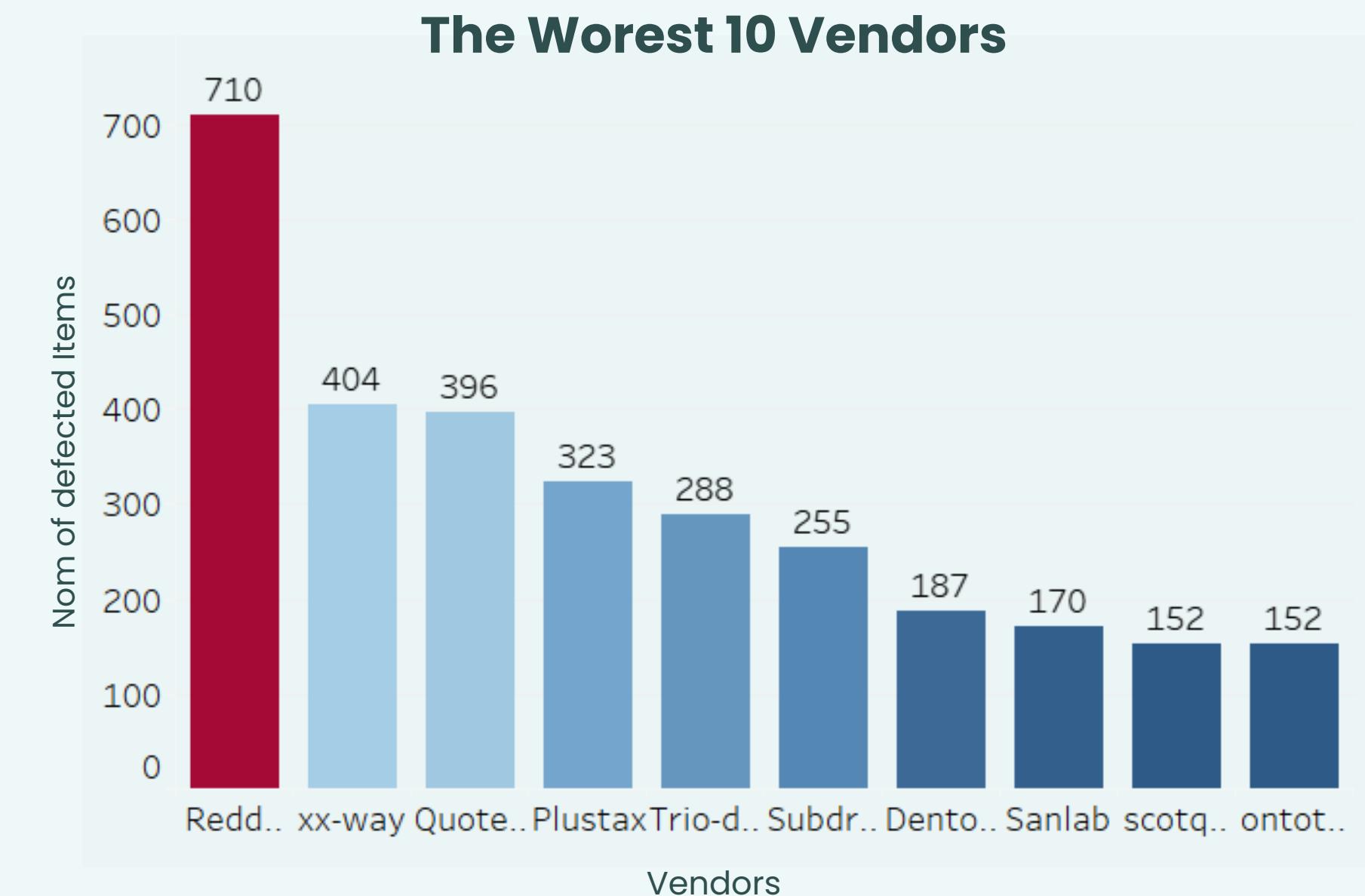
**Total Defected QTY Per Vendor**



• • • • •  
• • • • •  
• • • • •

# Vendors

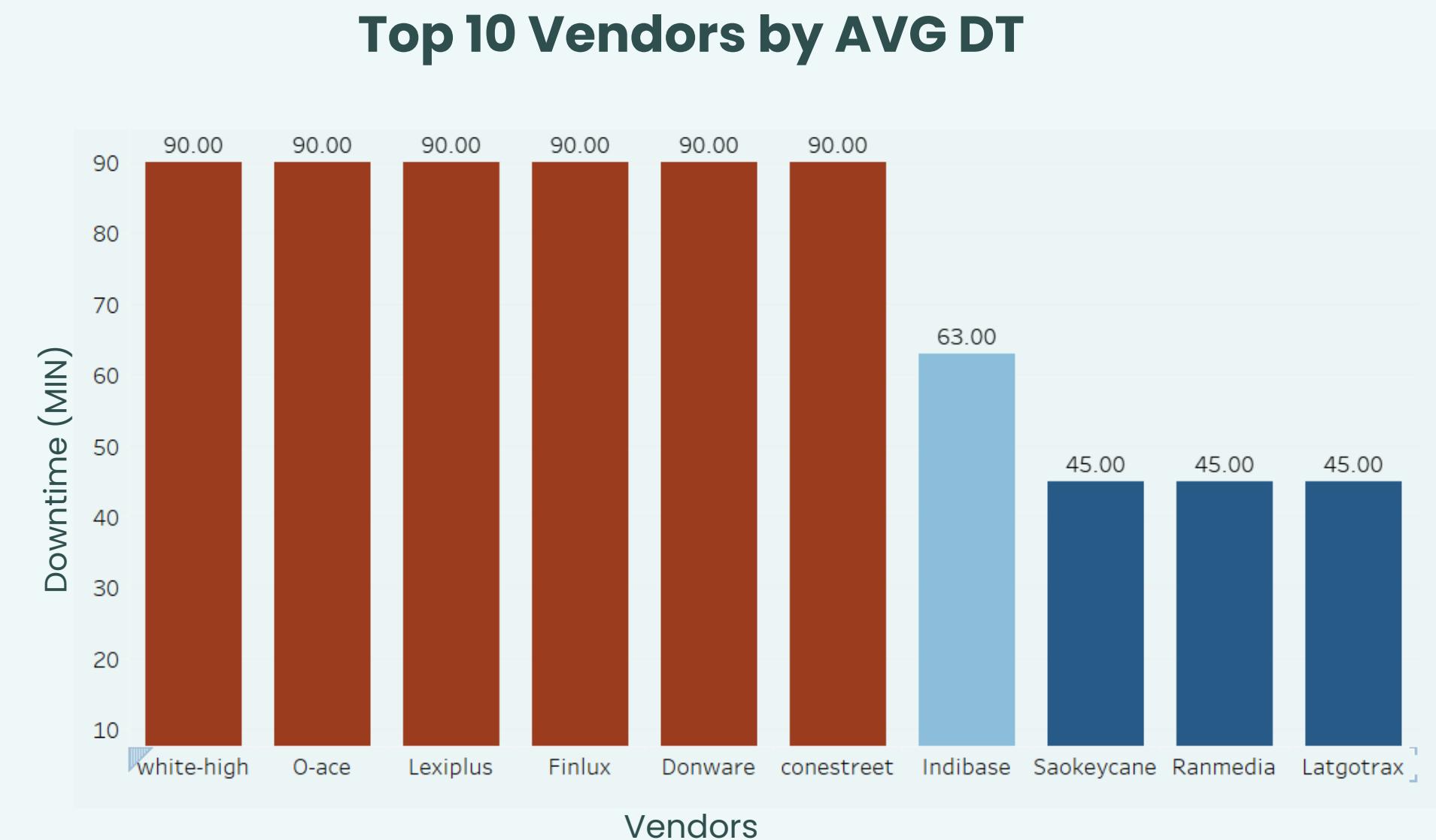
- **Reddoit** is the worst-performing vendor, with a notably higher number of defects—710 defective items—making it stand out significantly from other vendors.
- The vendors **Plustax** and **Trio-dax**, with 323 and 288 defective items respectively, are also among the top 5 worst performers, though their defect counts are considerably lower than Reddoit.



• • • • •  
• • • • •  
• • • • •

# Vendors

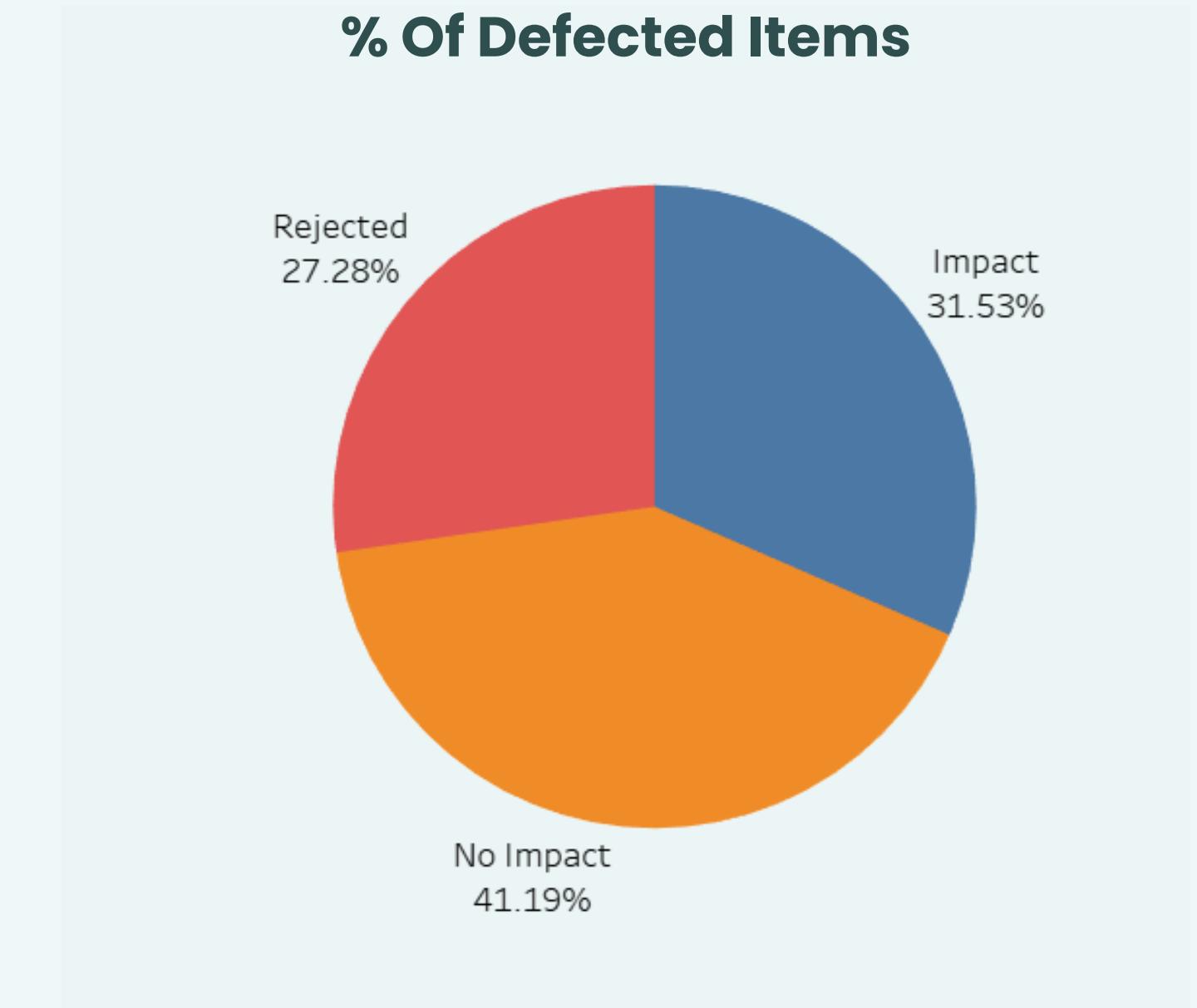
- The first six vendors—**White-High, O-Ace, Lexiplus, Finlux, Donware, and Conestreet**—exhibit an exceptionally high average downtime of approximately **90 minutes**, significantly exceeding that of other vendors.
- In contrast, the last three vendors, **Saokeycane, Ranmedia, and Latgotrax**, have average downtimes of **45 minutes**, which is notably lower than the top five vendors.



• • • • •  
• • • • •  
• • • • •

# Defect Type

- The **No Impact** type accounts for the highest percentage of defective items, making it the most common category.
- The **Impact** and **Rejected** types have similar proportions; however, **Rejected** represents the smallest percentage among them.

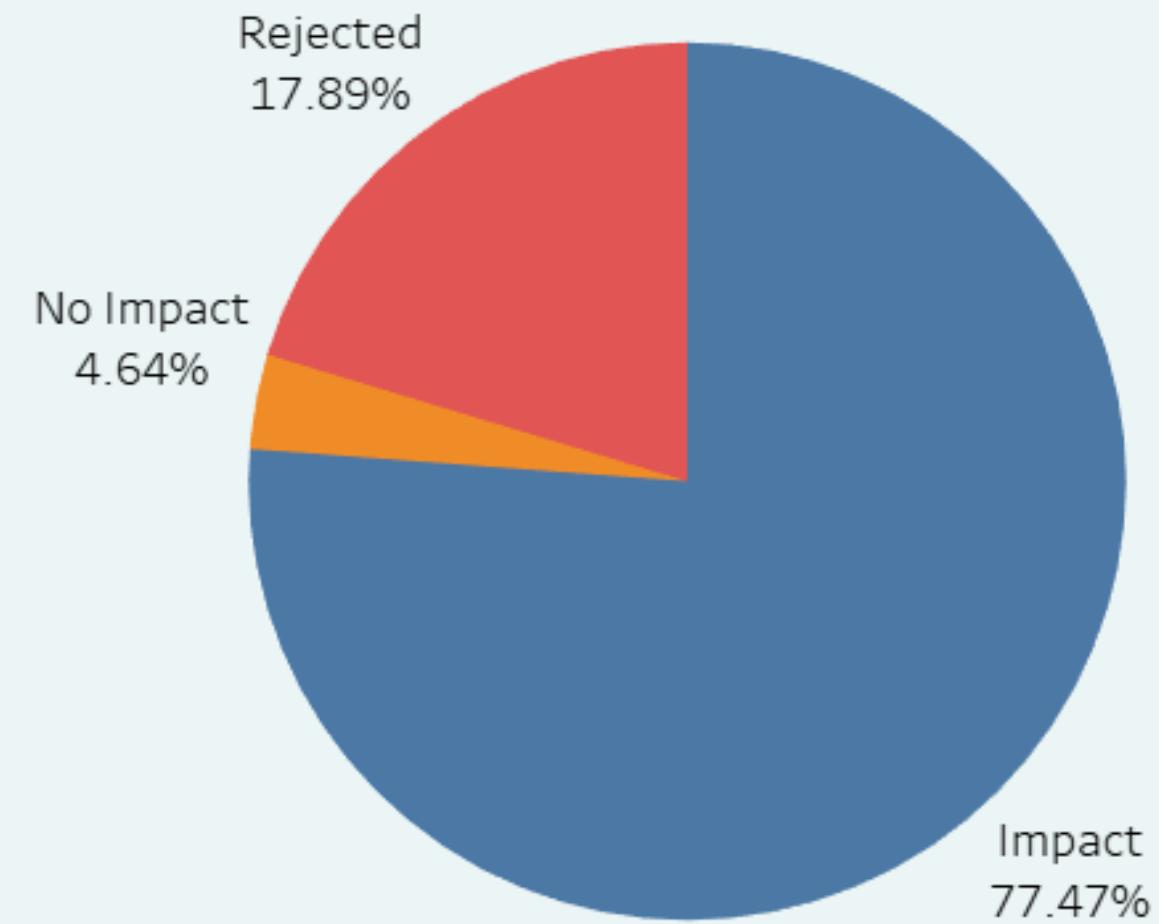


• • • • •  
• • • • •  
• • • • •

# Defect Type

- The **Impact** category has the highest percentage of average downtime in minutes, indicating it contributes significantly to overall downtime.

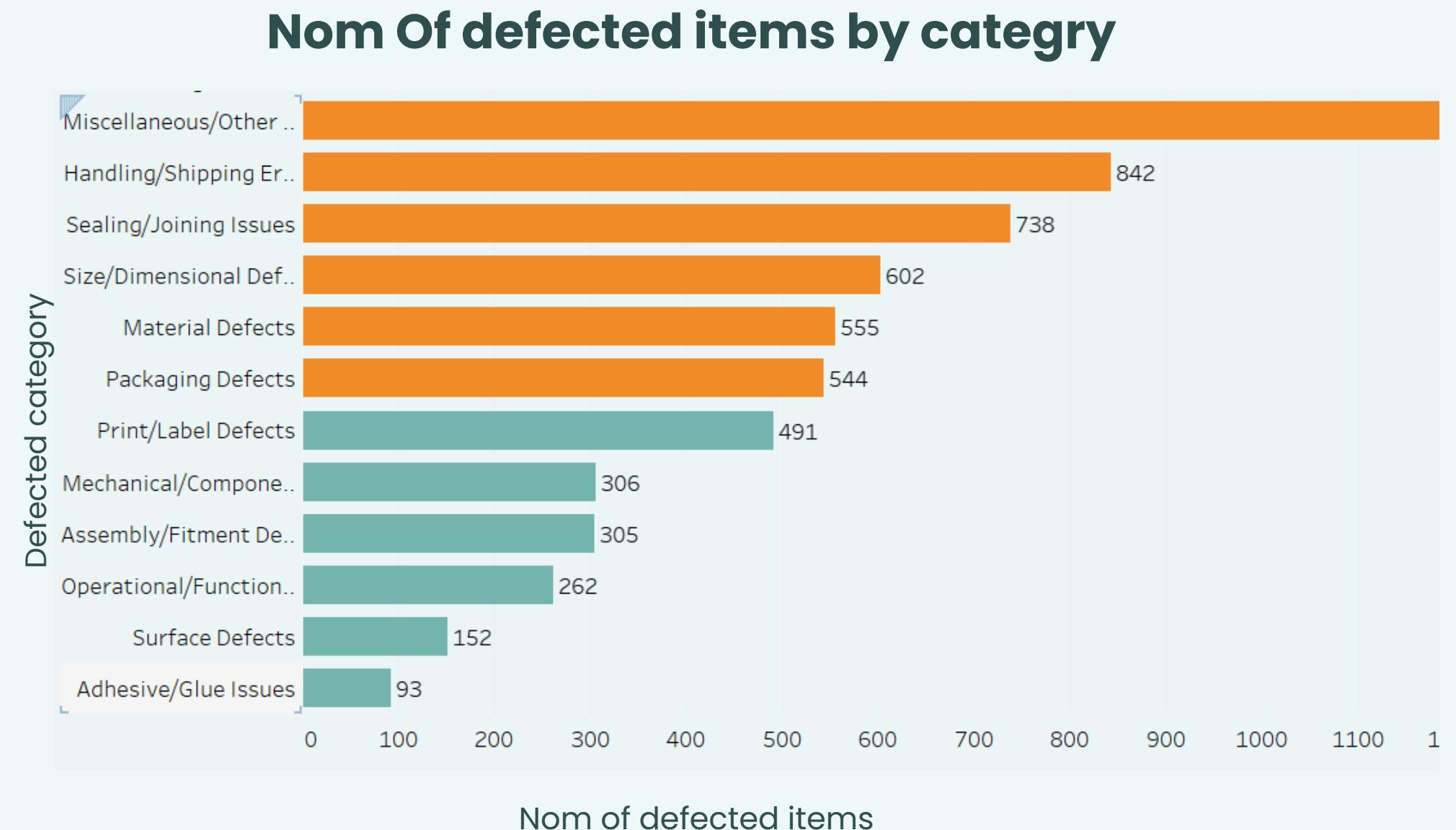
**AVG DownTime**



• • • • •  
• • • • •  
• • • • •

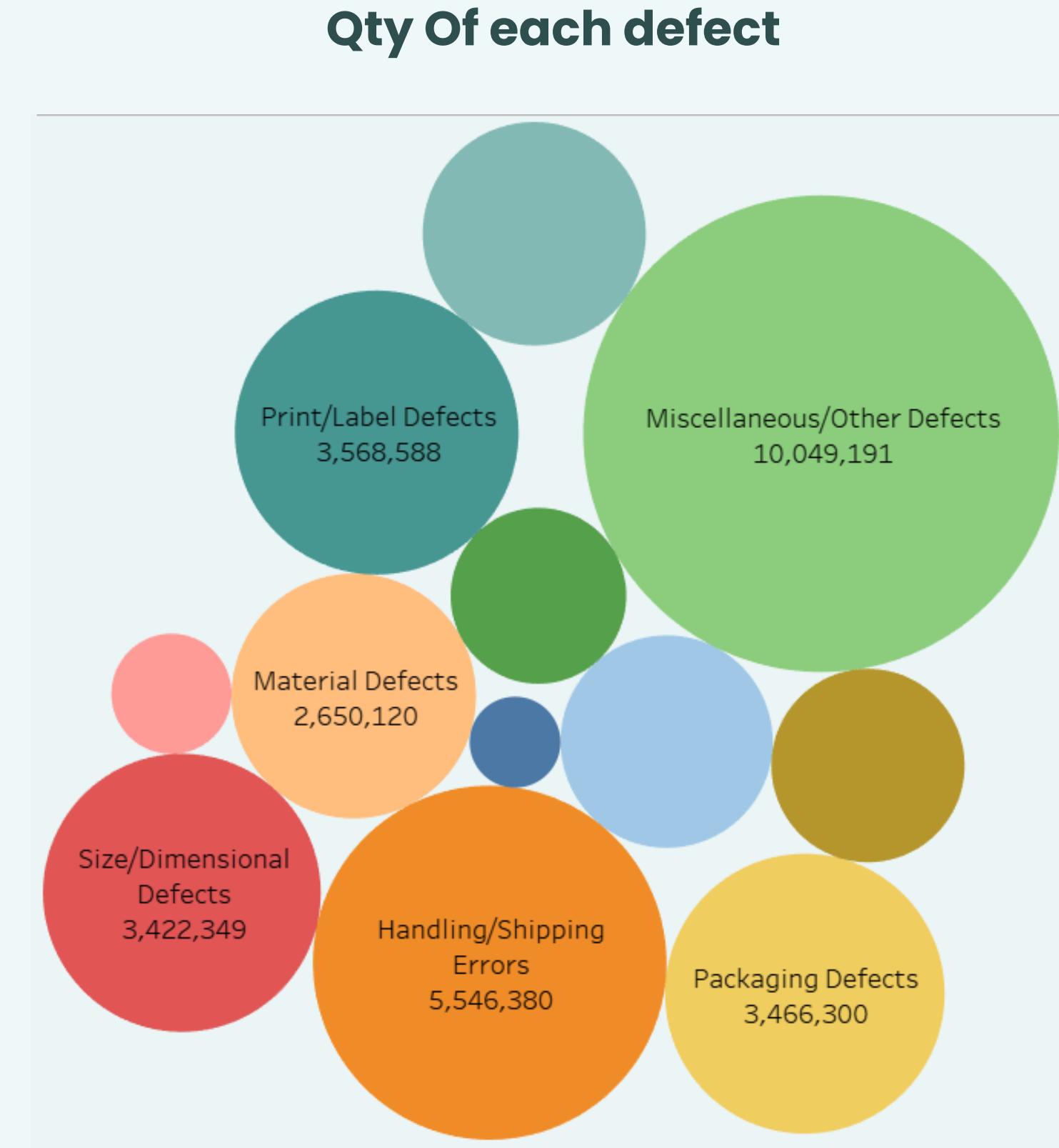
# Defects

- The data illustrates the number of defective items for each defect category, indicating whether the count is above or below the average.
- It shows that the **Miscellaneous/Other Defects** category is the most affected, with the largest number of defective items.



# Defects

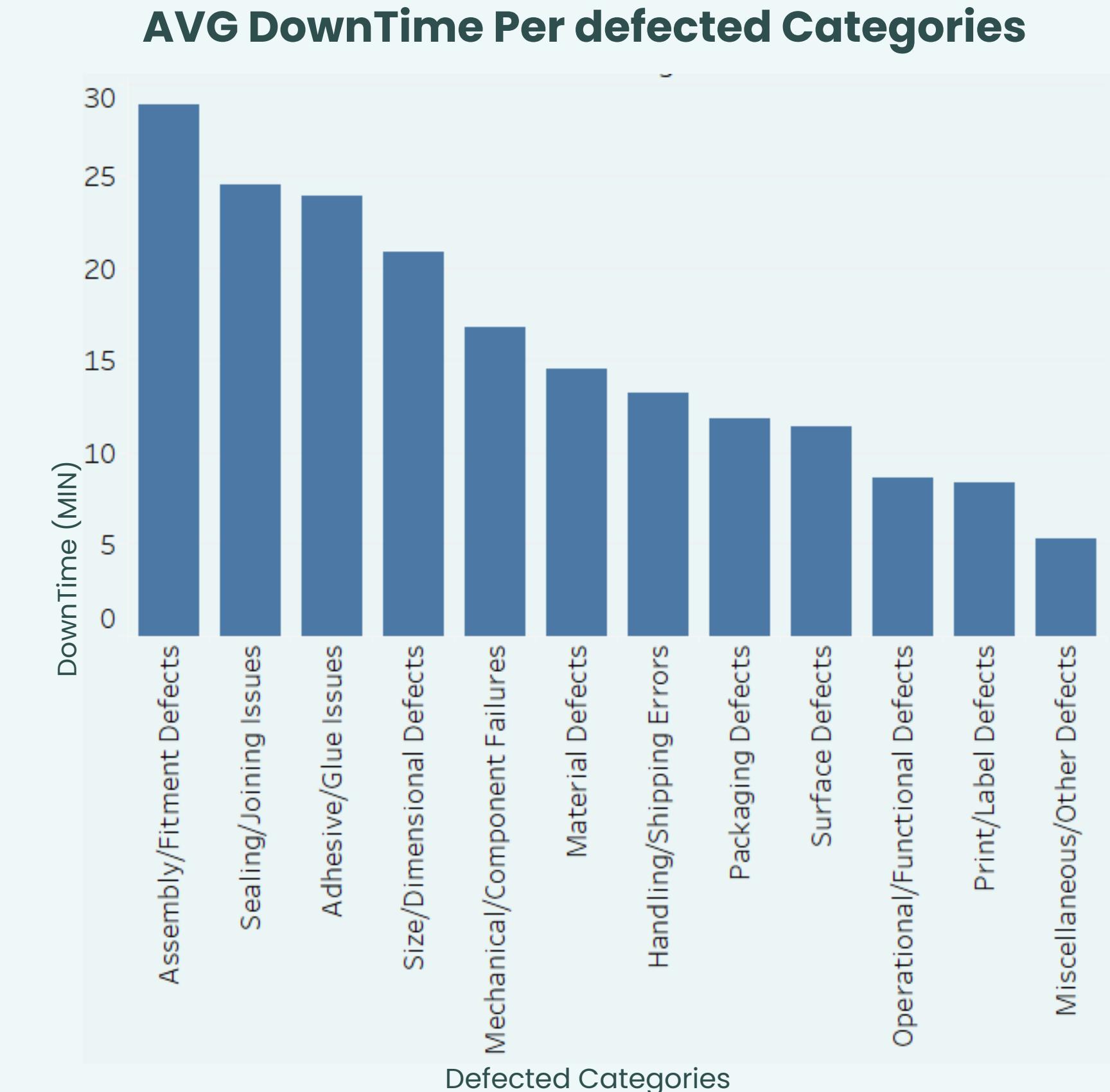
- The **Miscellaneous/Other** Defects category has the highest number of defective items, totaling 10 million.
- Conversely, the **Adhesive/Glue Issues** category has the lowest count, with only **366,000** defective items.



• • • • •  
• • • • •  
• • • • •

# Defects

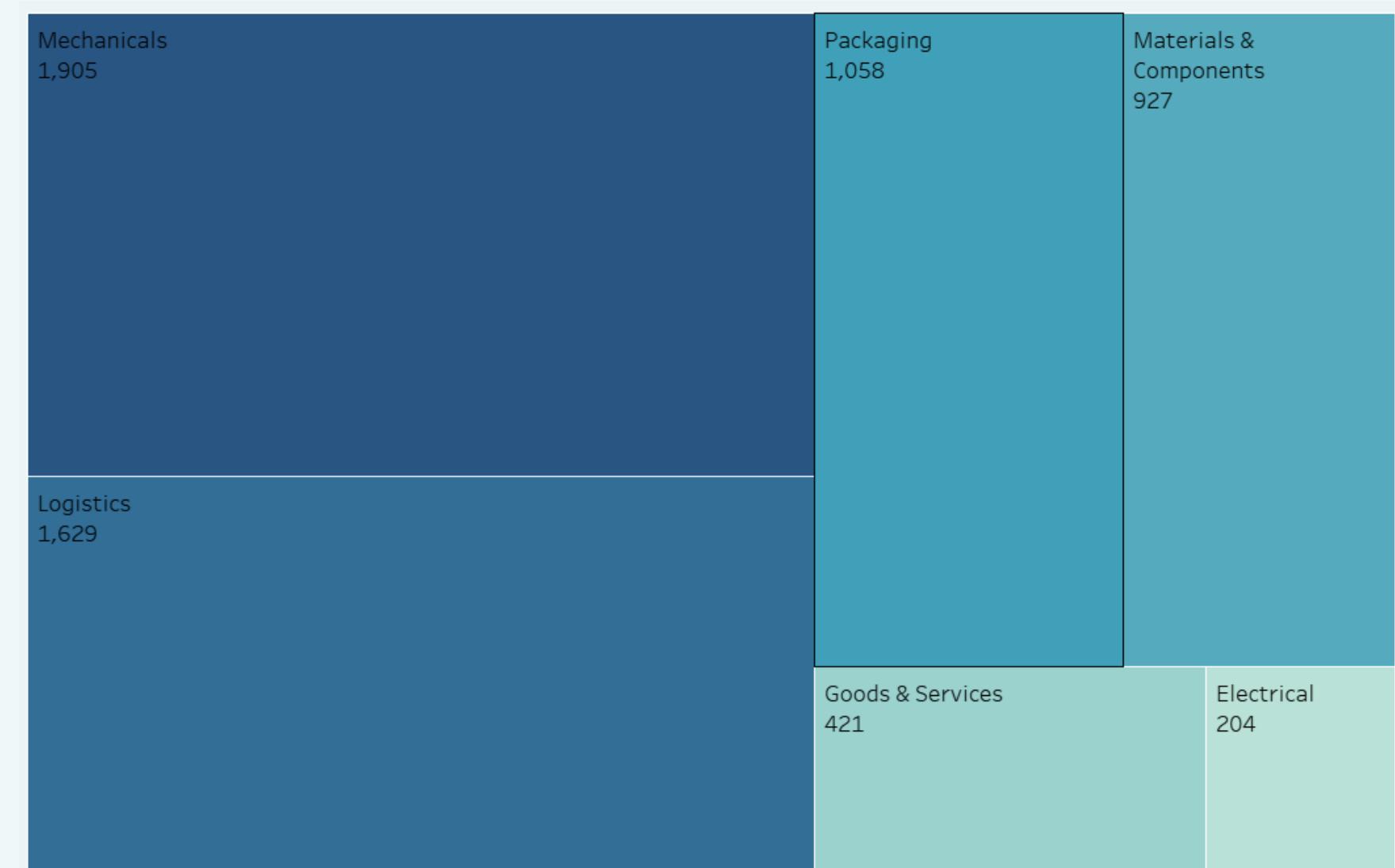
- While all defect categories exhibit similar proportions, **Assembly/Fitment Defects** still have the highest average downtime in minutes, indicating a greater impact on operations.



# Category

- The **Mechanical** category has the highest number of defective items, making it the most prevalent category.
- In contrast, the **Electrical** category has the fewest defective items, indicating it is the least problematic area.

**Nom Of defected items by category**

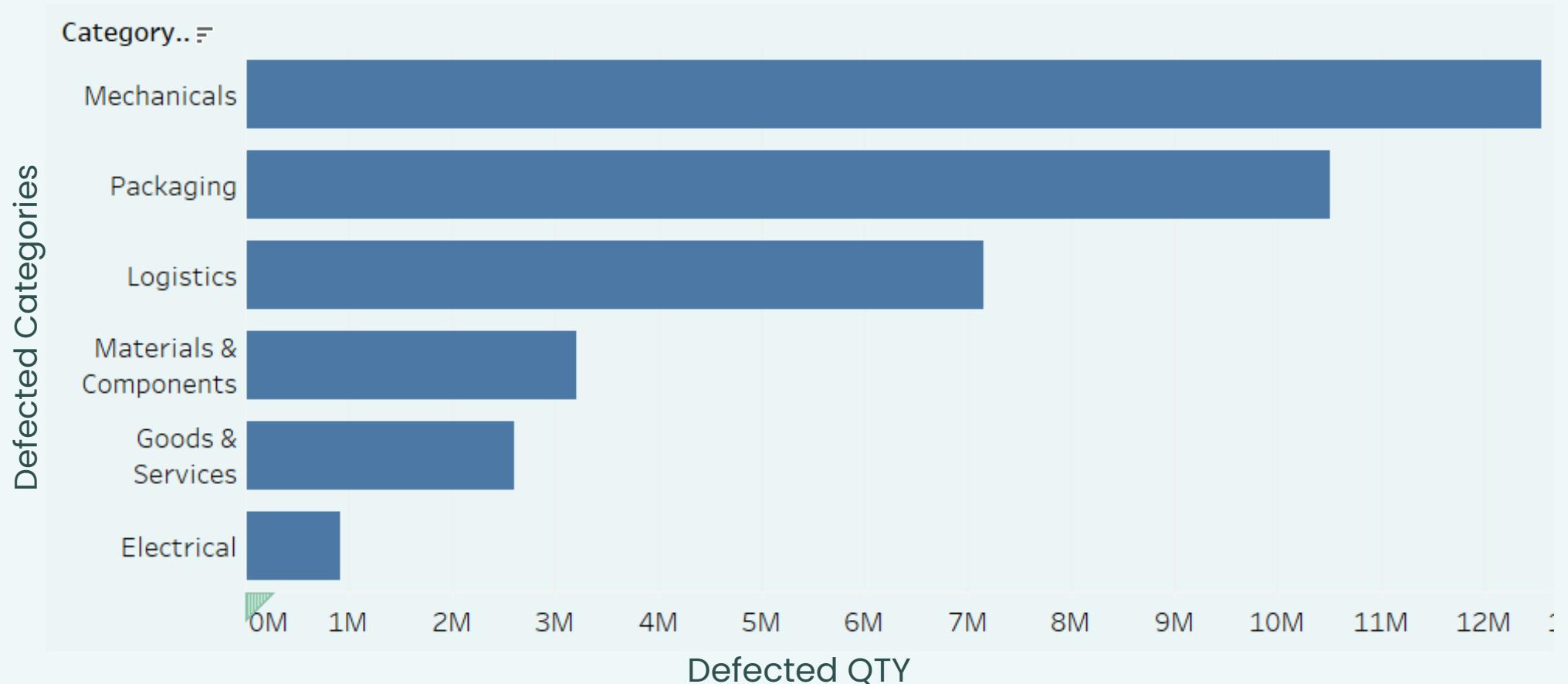


• • • •  
• • • •  
• • • •

# Category

- The data confirms that the **Mechanical** category not only has the highest quantity of defective items but also supports the conclusion that the **Electrical** category has the lowest quantity.

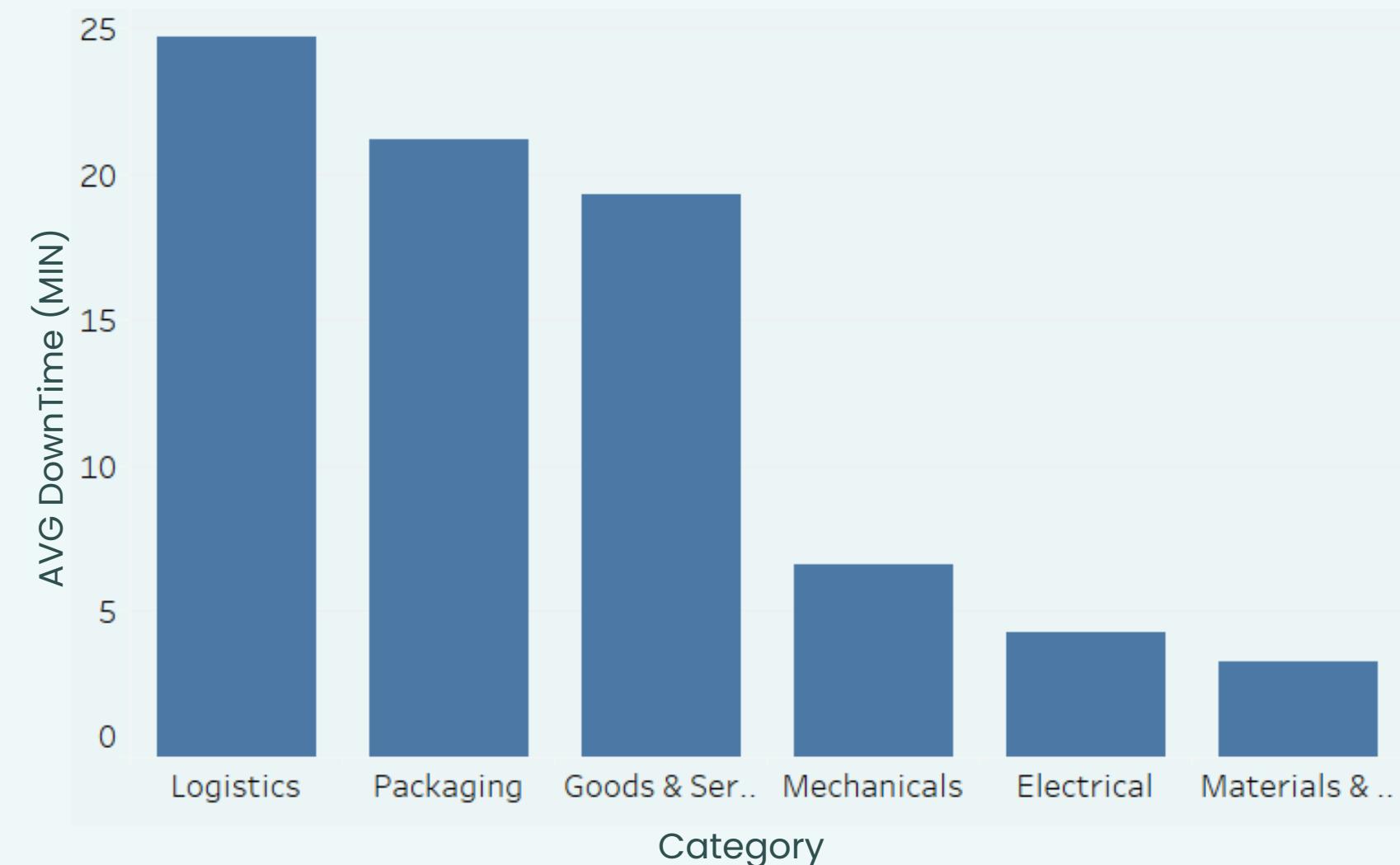
**The Most Material Category Per QTY**



# Category

- However, the chart indicates that the **Logistics** category experiences the highest average downtime in minutes among all material categories.

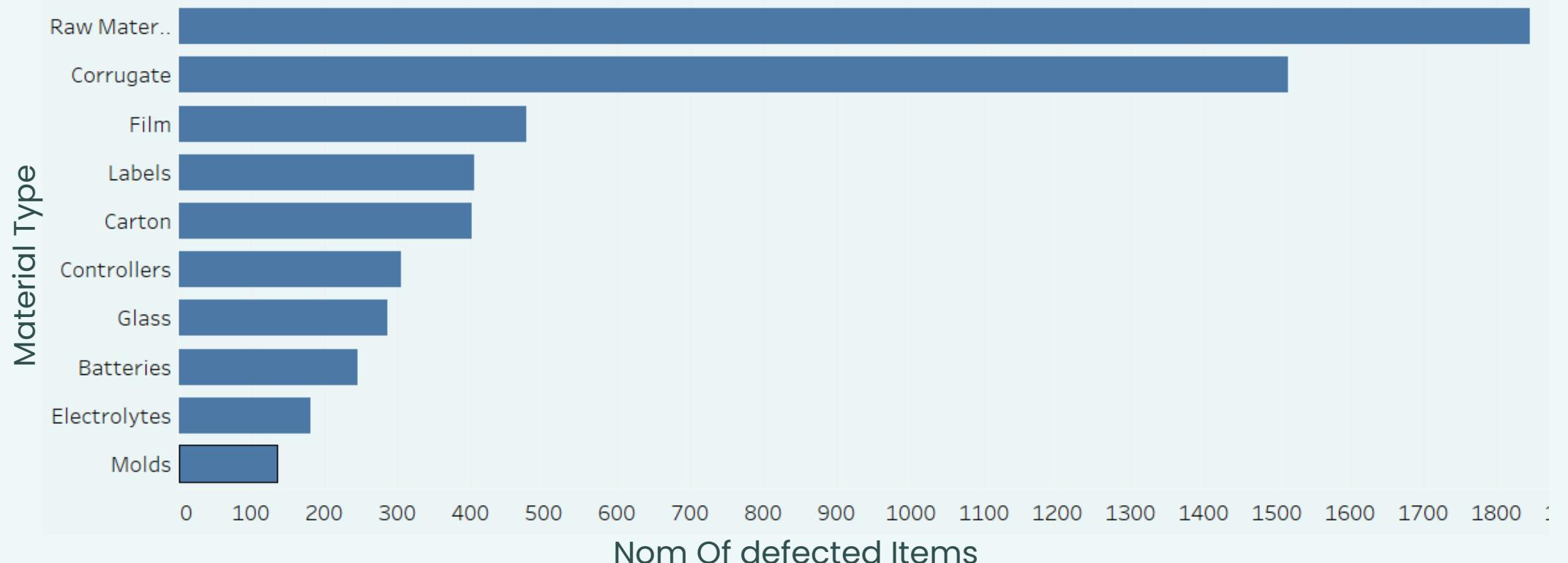
**AVG DownTime PEr Material Category**



# Material Type

- **Raw Materials** and **Corrugate** have the highest number of defective items, indicating significant issues in these categories.
- In contrast, **Electrolytes** and **Molds** have the fewest defective items, suggesting they are less problematic.

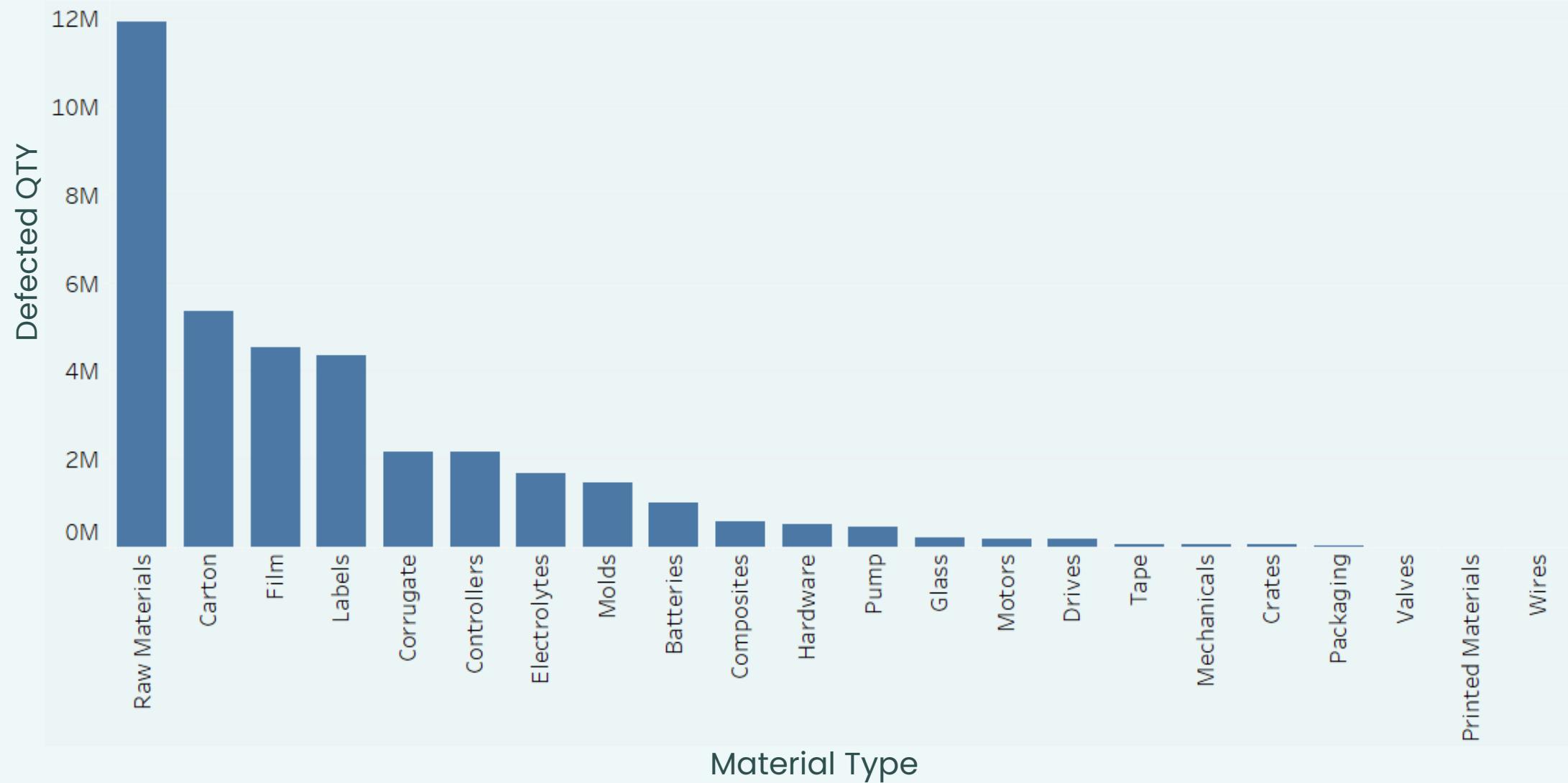
**Nom Of Defected Items by material Type**



# Material Type

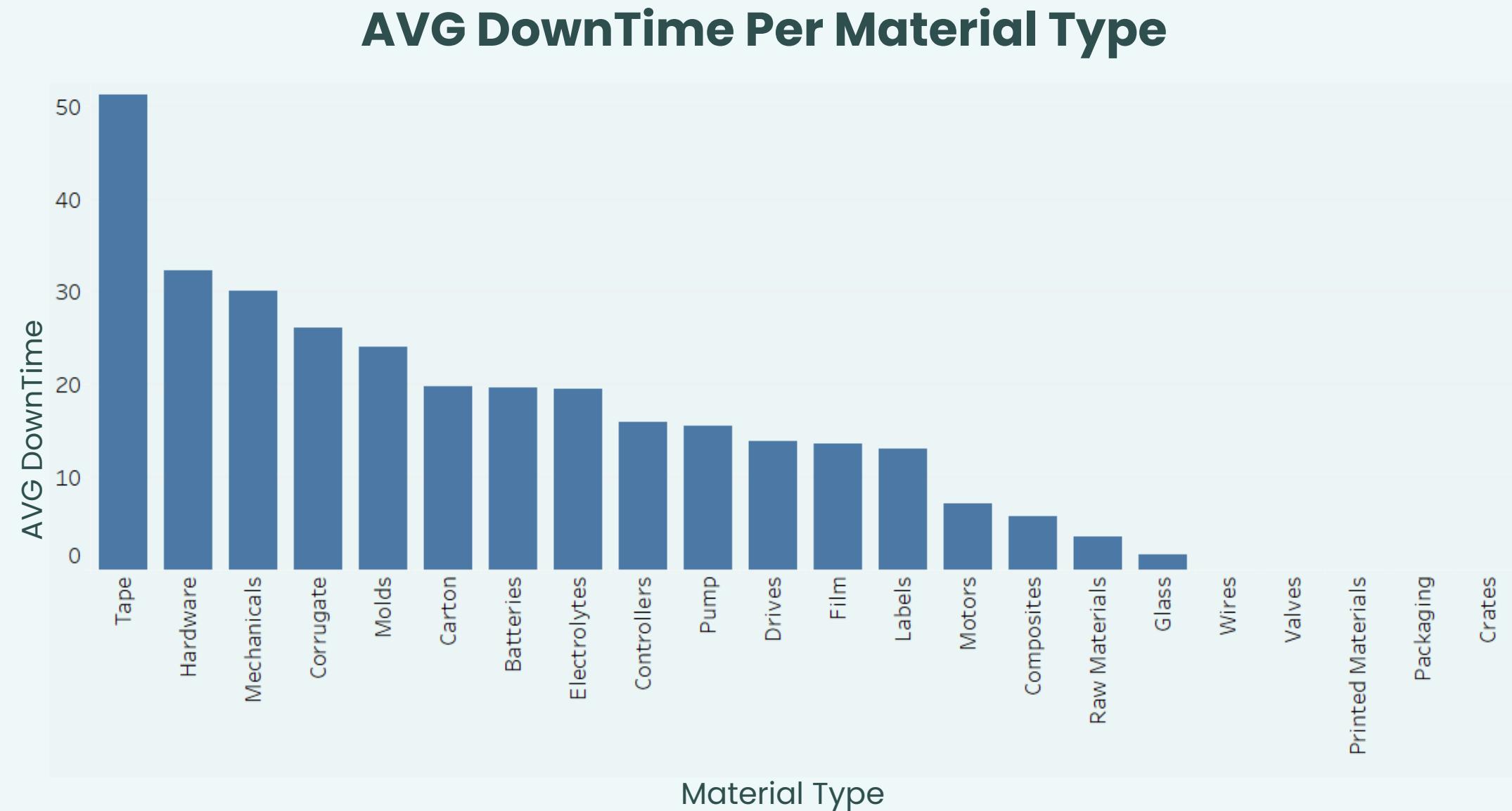
- The chart confirms that **Raw Materials** have the highest quantity of defective items as well.
- Conversely, **Wires** have an exceptionally low quantity of defective items, nearly reaching zero.

**QTY Of Each Material Type**



# Material Type

- The **Tape Type** category has the highest average downtime in minutes for items, indicating significant delays associated with it.
- In contrast, the **Crates Type** category has the lowest average downtime in minutes for items, suggesting minimal impact on operations.



# Trends

- **October 2014** marks the peak period, with the highest number of defective items, totaling approximately **555** items.
- Following this peak, there is a significant decline in the number of defective items until the end of December.

## Defected Items Over Months



# Trends

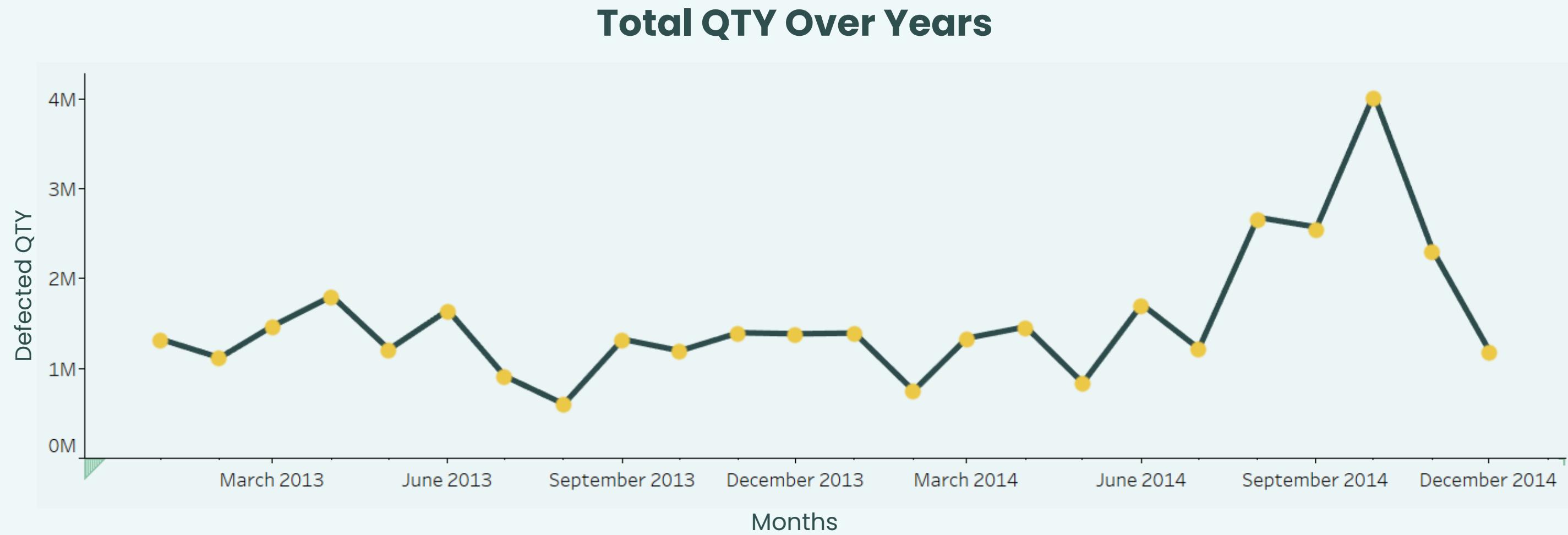
- The data reveals fluctuations between **2013** and **2014**, where the number of defective items experiences several rises and falls but generally trends upward leading up to the peak.

## Defected Items Over Months



# Trends

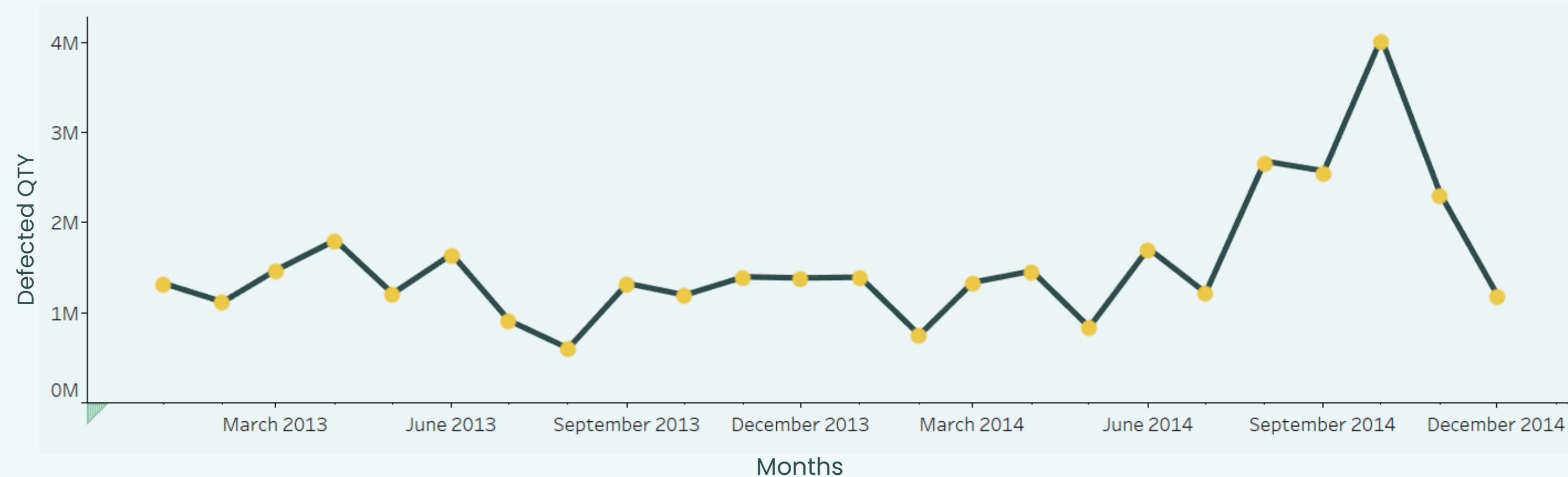
- There is a noticeable upward trend in defective quantities, especially as the chart progresses toward the latter months of **2014**.
- The highest defect quantity is recorded in **October 2014**, reaching **4,046,971** defective items.



# Trends

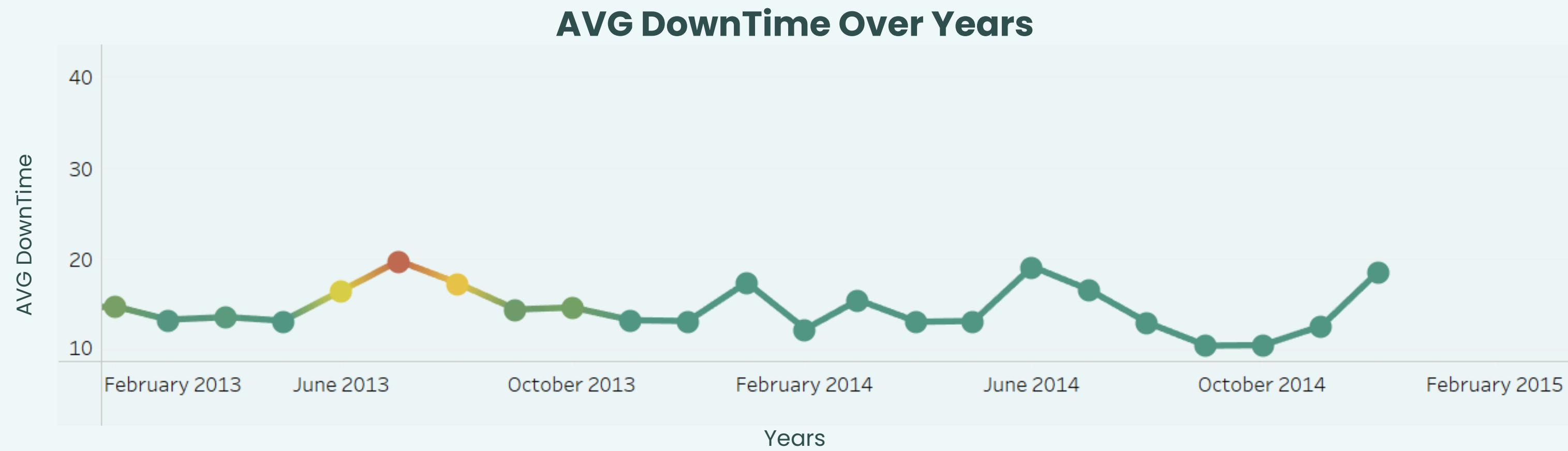
- Throughout **2013** and **2014**, the data shows significant fluctuations in the defective quantity, characterized by sharp rises and drops. The chart illustrates a cyclical pattern, but the overall volume of defects appears to increase as **2014** comes to a close.
- The lowest point occurs at the end of **2013**, specifically in **August**, with a value of **604,357** defective items.

**Total QTY Over Years**



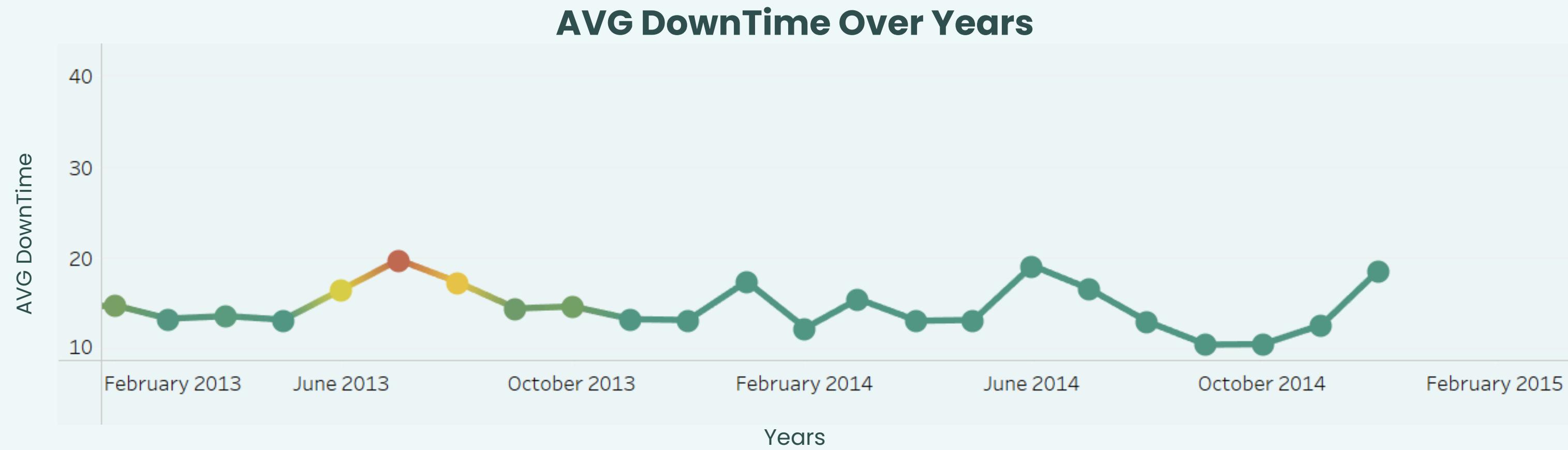
# Trends

- The chart reveals a cyclical pattern of peaks and valleys in downtime, with no consistent upward or downward trend over the years. This indicates fluctuating issues with machine or process efficiency rather than a continuous improvement or decline.
- The peak downtime is recorded in **July 2013**, with an average of **19.83 minutes**, marking the highest downtime in the observed timeframe.



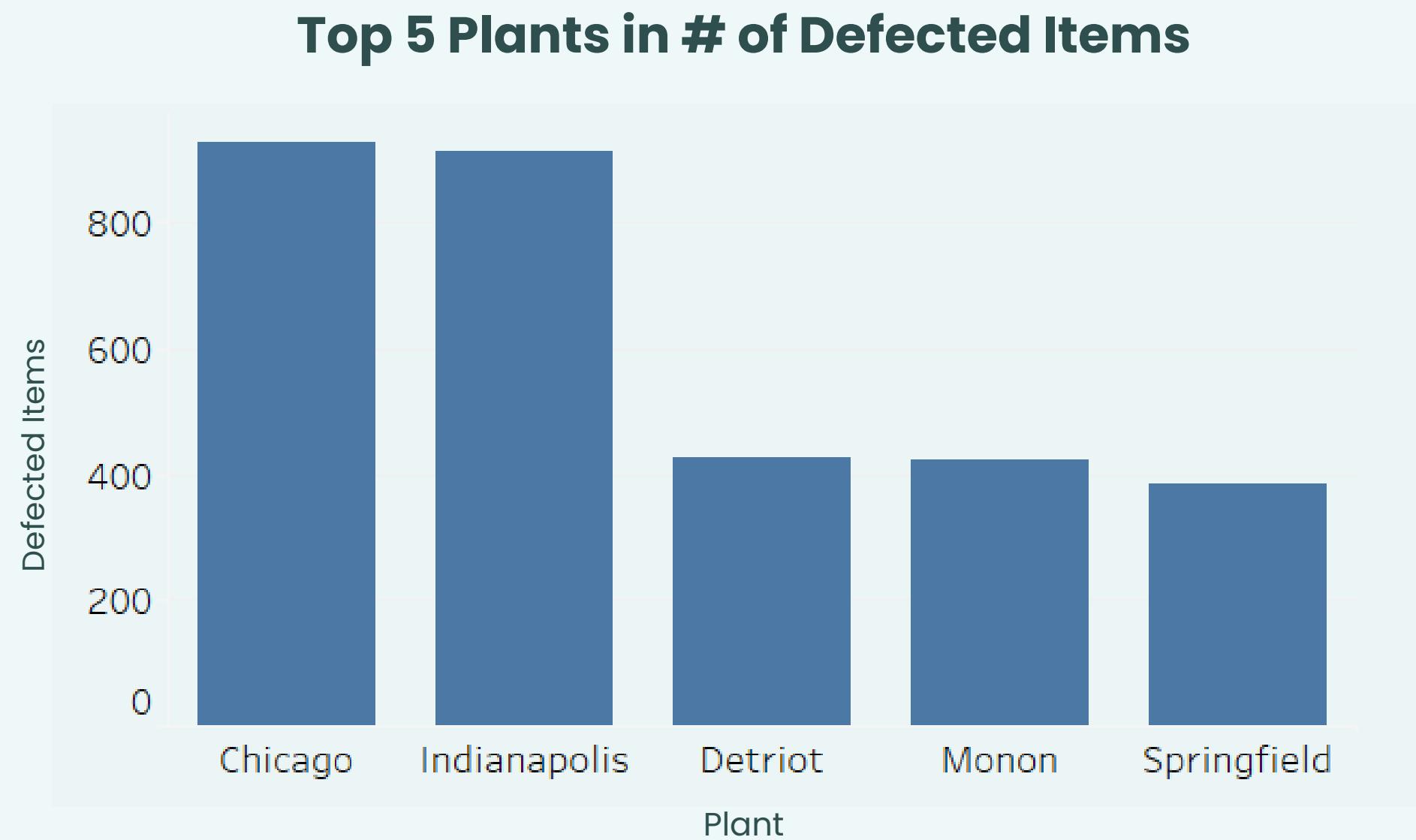
# Trends

- Several months show a significant drop in average downtime, falling below **15 minutes**, particularly in **May 2013, December 2013, February 2014, and September 2014**.



# Plants

- The **Chicago** and **Indiana** plants have the highest number of defective items, with similar counts of approximately **900** each.
- In contrast, the **Monon** and **Spring** plants show significantly lower numbers of defects, each recording around **400** defective items.

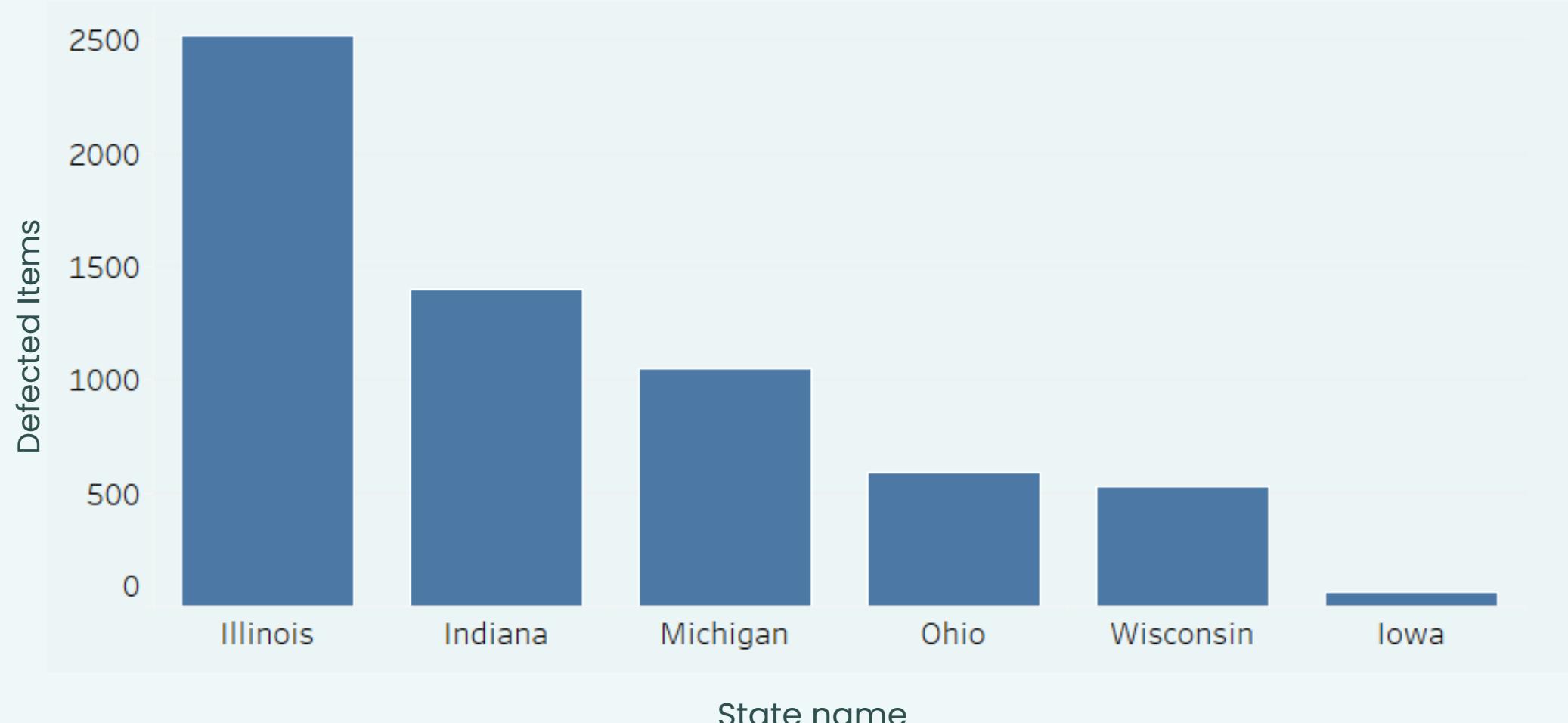


• • • •  
• • • •  
• • • •

# Plants

- **Illinois State** has the highest quantity of defective items, with similar counts of approximately **2,500**.
- Conversely, **Iowa State** exhibits significantly lower numbers of defects, with around **50** defective items.

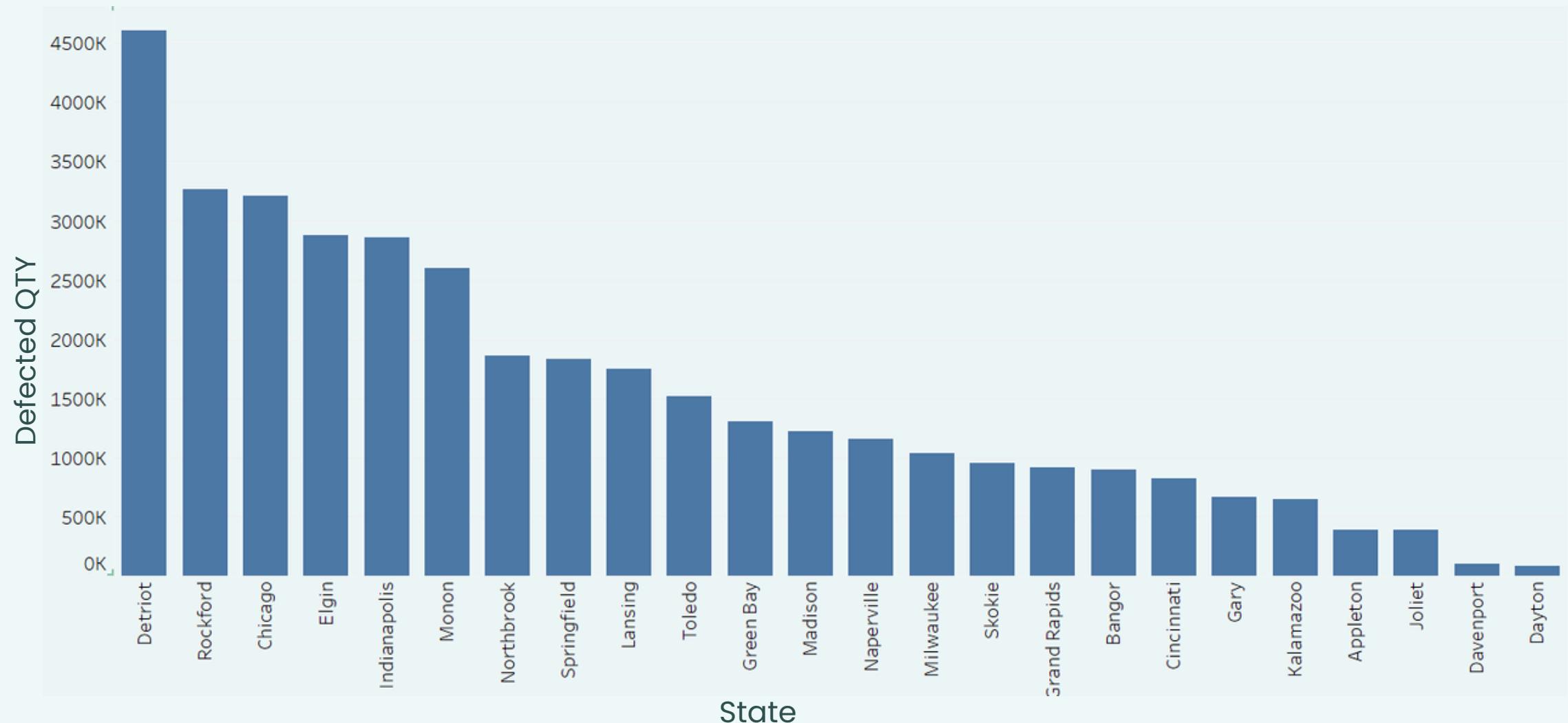
**Defected Items In Each State**



# Plants

- **Detroit** stands out with a significantly higher quantity of defects, nearing **4.5 million** defective items, compared to all other plants.
- There is a gradual decrease in defect quantities among the remaining plants, with **Dayton** and **Davenport** showing the lowest numbers of defects.

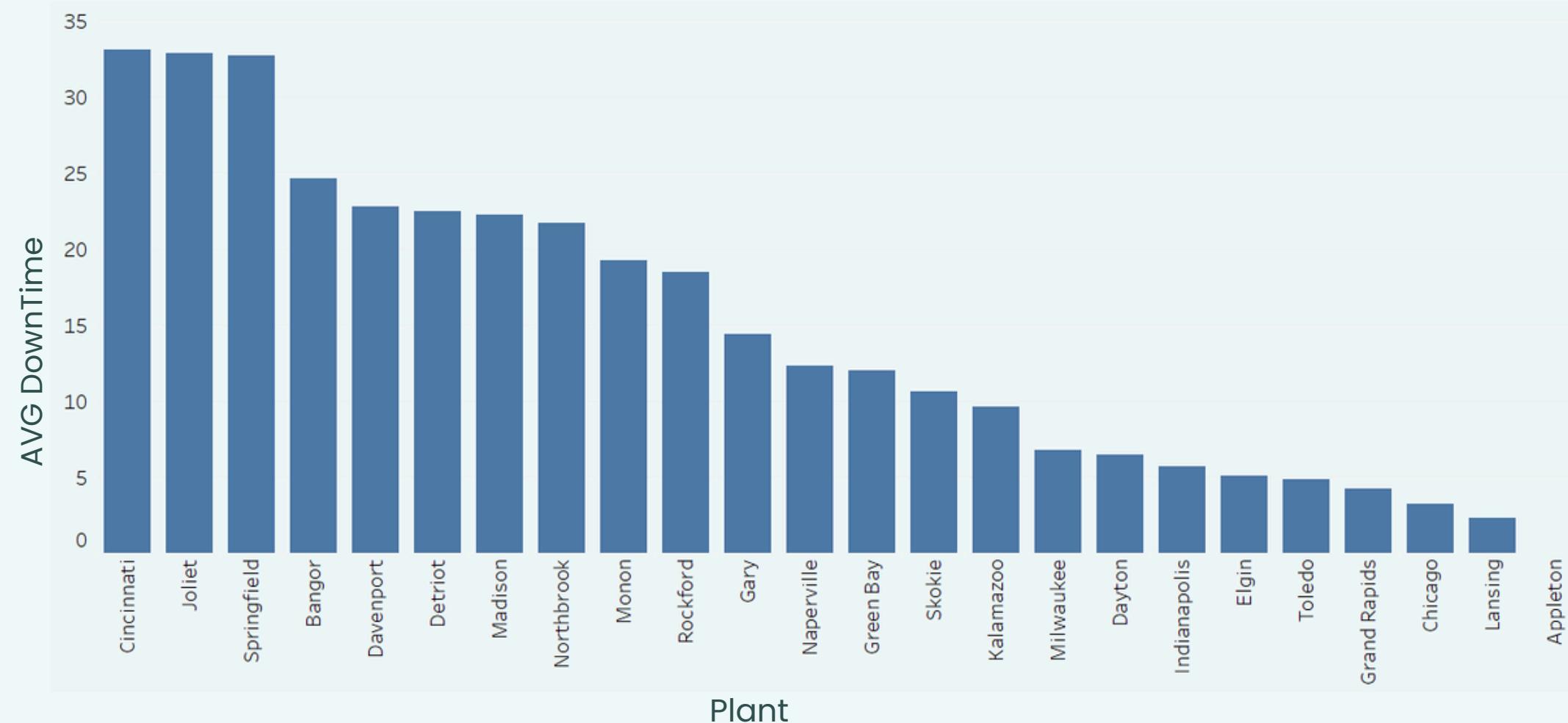
**Total QTY in each Plant**



# Plants

- **Cincinnati** has the highest average downtime, exceeding **33 minutes**, indicating significant operational issues or inefficiencies at this plant.
  - In contrast, downtime decreases steadily across the remaining plants, with the Appleton plant reporting the least downtime, at nearly **0 minutes**.
- • • • •

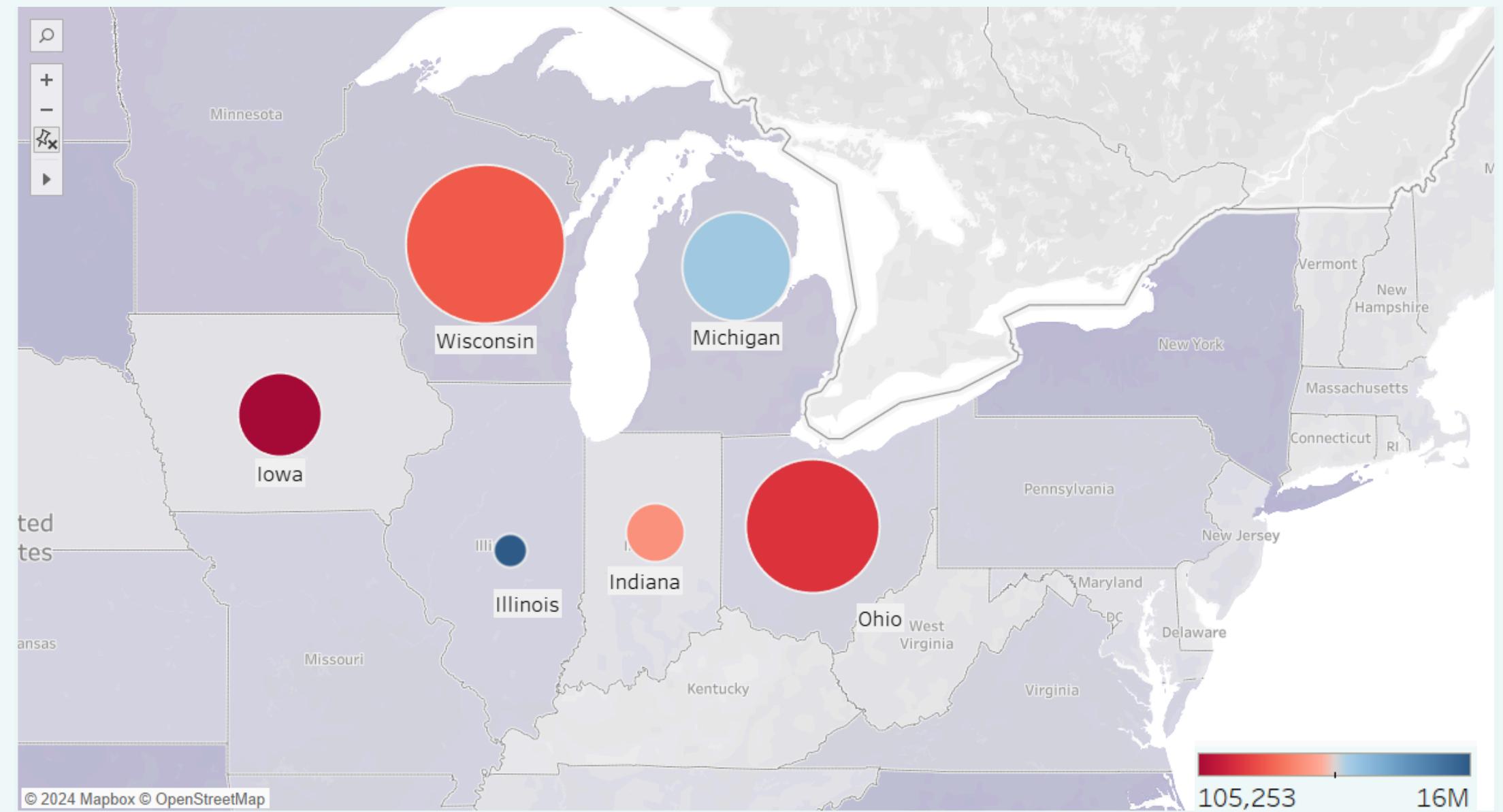
**AVG DownTime In Each Plant**



# Plants

- **Wisconsin** stands out with the highest quantity of defects, surpassing all other states by a significant margin.

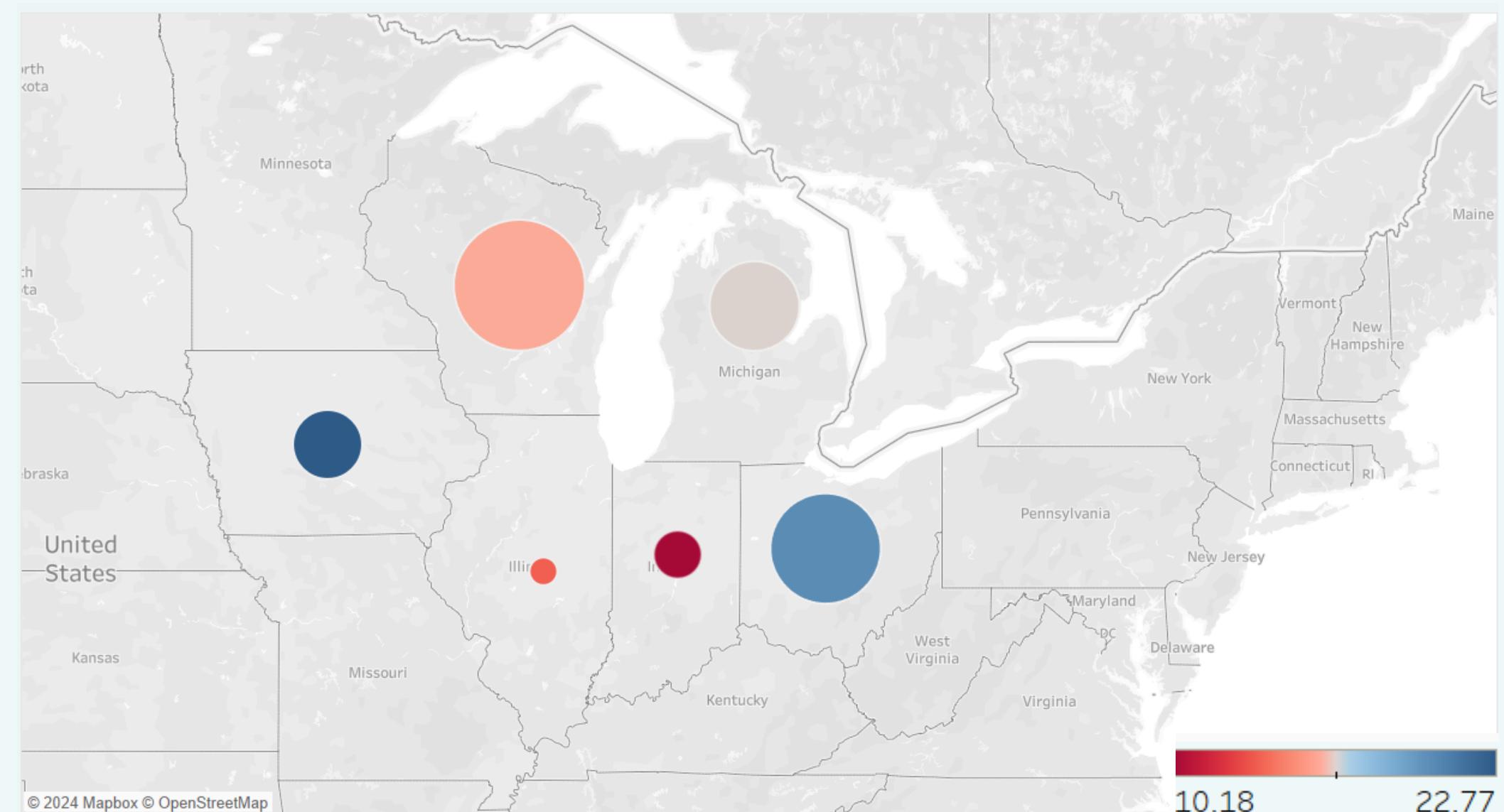
**Total QTY Per Plants Geo**



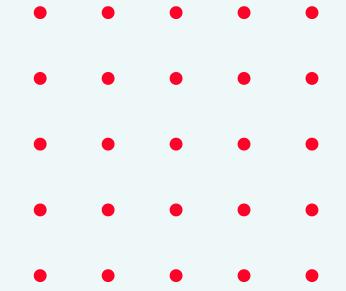
# Plants

- **Iowa** has the highest average downtime, exceeding **22 minutes**, which indicates significant operational issues or inefficiencies in this state.
  - Conversely, the downtime decreases steadily across the remaining states, with the **Indiana** plant recording the least downtime, at nearly **10 minutes**.
- • • • •

**AVG DownTime In Each State**



# Key Findings





Vendors



Defects



Material Types



Defects Types



Plants



States



Material Categories



Over Time

• • • • •

# Recommendations



# What strategies can reduce defect rates and improve product quality?

Focus on reducing defects in raw materials that currently show no impact.

1

Target rejected defects in raw and corrugated materials to enhance quality and reduce costs.

2

Review sourcing and production processes to address defects in materials like film and labels.

3

Focus on improving quality control in plants like Chicago, Indiana, and Detroit with high defect rates.

4

# What strategies can reduce defect rates and improve product quality?

- 5 Illinois needs more attention due to having the highest number of defective items across its plants.
- 6 Address high downtime at the Cincinnati plant to improve productivity.
- 7 Work closely with major vendors like Reddoit and Quotelane to reduce overall defect rates.
- 8 Investigate the spike in defects from October 2014 to identify potential production or quality control issues.

# **Next Steps: Advancing Insights in Future**



# Key Enhancements for Future Analysis

- 1 Add Material Dimension
- 2 Enrich Defected Items Fact Table: adding (Material Costs, Cost of Defects, Downtime Cost per Minute)
- 3 Enhance Vendor Dimension: Add Vendor Rating feature
- 4 Introduce Machine and Maintenance Data
- 5 Capture Environmental Factors
- 6 Integrate Logistics and Supply Chain Data
- 7 Add Customer Feedback and Satisfaction Metrics
- 8 Incorporate Process Variables





THANK  
YOU!!