DEPI Graduation Project

Suppliers Chain Quality Project

وزارة الاتصـــــالات
وتكنولوجيا المعلومات

CLS Learning Solutions

Our Professor:
Eng. Yasser Abdulrahman

Mohamed Selim

Youssef Ismail

Ahmed Yahia

Ahmed Hany

Ahmed Ayman

وزارة الاتصـــــالات
وتكنولوجيا المعلومات

CLS Learning Solutions

Table of contents

01
Data Overview

About The Data

This is an analysis project of Suppliers "Vendors" Quality and Manufacturing Downtime. Here we deal with a data set of how much some vendors' materials quality is.

We aim for two main objectives:

- To understand who the best and worst suppliers are.
- Identify which plants do a better job finding and rejecting defects, to minimize downtime.

Downtime is the Time during which production is stopped, during setup for an operation, when making repairs, or when stoppages occur due to supply or labor shortages. Defects are imperfections or errors of any type that affect materials.

Entity Relationship Diagram

Plants	
Plant_ID	**int**
Plant	varchar

Defect_Items	
Date	date
Category_ID	int
Plant_ID	int
Vendor_ID	int
Defect_Type_ID	int
Matrial_Type_ID	int
Defect_ID	int
Defect_qty	int
Downtime_min	int

Defects	
Defect_ID	**int**
Defect	varchar

Vendors	
Vendor_ID	**int**
Vendor	varchar

Materials_Type	
Material_Type_ID	**int**
Material_Type	varchar

Defect_Types	
Defect_Type_ID	**int**
Defect_Type	varchar
Sort	int

Category	
Sub_Category	varchar
Sub_Category_ID	**int**
Category	varchar

The raw data

02

Technical Report

Table Preparation

Plant	Plant ID
Grand Rapids, MI	1
Milwaukee, WI	2
Springfield, IL	3
Chicago, IL	4
Indianapolis, IN	5
Northbrook, IL	6
Detroit, MI	7
Gary, IN	8
Joliet, IL	9

City	State	Plant ID
Grand Rapids	MI	1
Milwaukee	WI	2
Springfield	IL	3
Chicago	IL	4
Indianapolis	IN	5
Northbrook	IL	6
Detroit	MI	7
Gary	IN	8
Joliet	IL	9

Plants	
Plant_ID	**int**
Plant	varchar

Plants	
Plant_ID	**int**
State	varchar
City	varchar

```sql
CREATE TABLE NewPlants (
    City VARCHAR(255),
    State CHAR(2),
    Plant_ID INT PRIMARY KEY
);

INSERT INTO newplants (City, State, Plant_ID)
SELECT
    SUBSTRING_INDEX(Plant_Name, ',', 1) AS City,
    CASE
        WHEN Plant_Name REGEXP '.*, [A-Z]'
         THEN SUBSTRING_INDEX(Plant_Name, ', ', -1)
        WHEN Plant_Name REGEXP '.* [A-Z]'
         THEN SUBSTRING_INDEX(Plant_Name, ' ', -1)
        ELSE RIGHT(Plant_Name, NULL)
    END AS State,
    Plant_ID
FROM Plants;
```

SQL

Table Preparation

Sub Category	Sub Category ID	Category
Electrical	1	Electrical
Logistics	2	Logistics
Materials & Components	3	Materials & Components
Mechanicals	4	Mechanicals
Packaging	5	Packaging
Goods & Services	6	Goods & Services

Category	Category ID
Electrical	1
Logistics	2
Materials & Components	3
Mechanicals	4
Packaging	5
Goods & Services	6

Categories	
Sub	Category
Sub_Category_ID	**int**
Category	varchar

Categories	
Cat_ID	**int**
Category	varchar

```sql
ALTER TABLE Categories
DROP COLUMN Sub_Category,
RENAME COLUMN Sub_Category_ID TO Category_ID;
```
SQL

Table Preparation

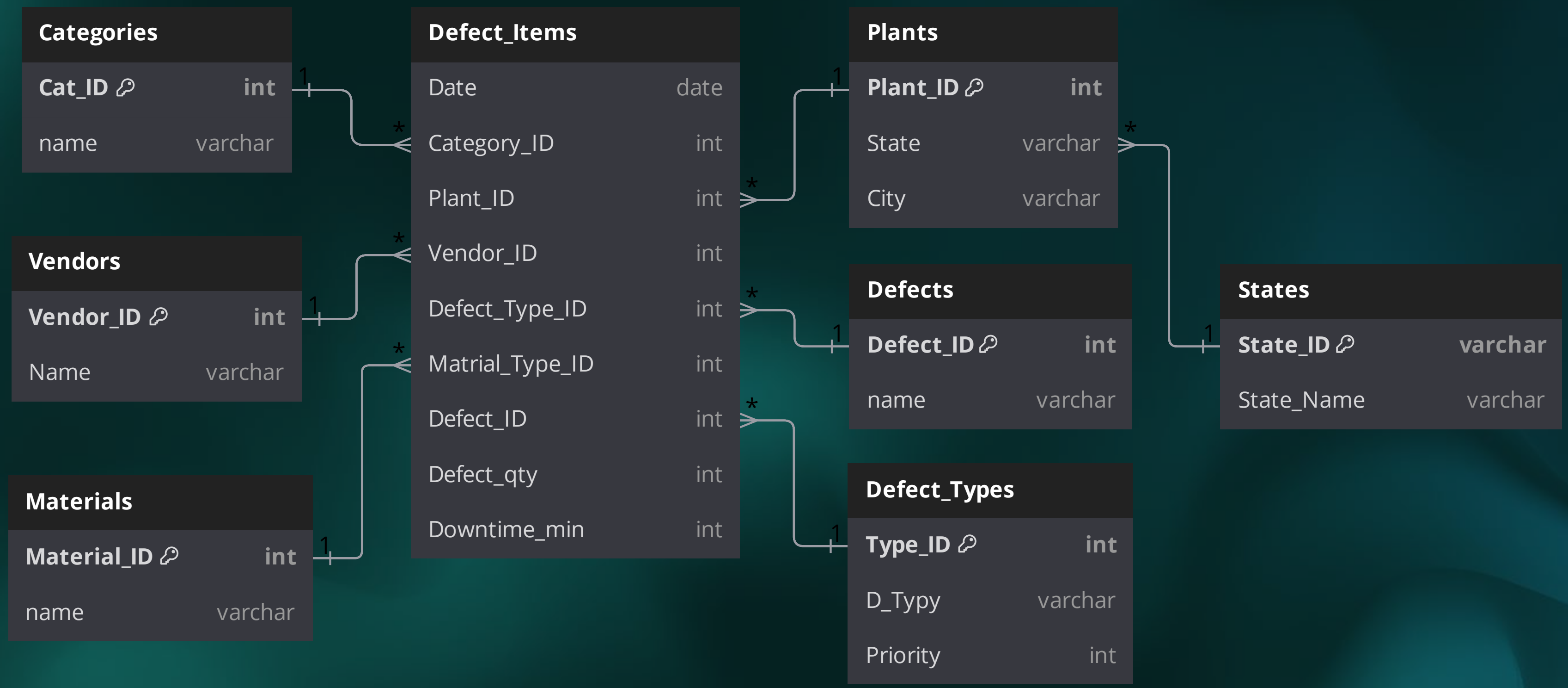
State	Name
MI	Michigan
WI	Wisconsin
IL	Illinois
IN	Indiana
OH	Ohio
IA	Iowa

States	
State_ID	**varchar**
State_Name	varchar

```sql
CREATE TABLE IF NOT EXISTS States(
    State_ID VARCHAR(2) PRIMARY KEY,
    State_Name VARCHAR(125)
);

INSERT INTO States (State_ID,State_Name)
VALUES
    ('MI','Michigan'),
    ('WI','Wisconsin'),
    ('IL','Illinois'),
    ('IN','Indiana'),
    ('OH','Ohio'),
    ('IA','Iowa');
```

SQL

Entity Relationship Diagram



Data after the preparation

Data Cleaning

Remove duplicates

Vendor	Vendor ID
roundphase	80
roundphase	113
Quotefix	125
Quotefix	144

Remove duplicates

```python
print(df.duplicated())
df.loc[[80,113,125,144]]
df.drop_duplicates(inplace = True)
```

Python

Data Cleaning

Remove duplicates

Step #01

Main	Replaced
Crack	Cracked
Printing defect	Printing defects
Scratch	Scratches
Wrapping	Wrapped
Wrong coloring	Wrong colors
Wrong Label	Wrong Labeling
Wrong Spec	Wrong Specifications
Out of Spec	Out of Specifications

Step #02

Dimensions - Bad Finishing	Wrinkles / Scratches/ Scuffing	Bowed/Warped	Flaps - Incorrect Gap
Bad Finish	Creases / Wrinkles	Bowed	Bad Flaps
Incorrect Dimensions	Scratches	Warped	Wrong Flaps
Lay Flat Dimension	Scratched Glass	Warped Sheets	Gaps
Dimensions Wrong	Scratch	Warping	Gap Variation
--	Scuffed Packaging	--	--

Step #03

Defect	Main ID	Replaced ID
Bad Seams	2	3, 4, 99, 206
Scrap attached	15	20
Wrong Size	13	14
Wrong Core	43	48
Foreign objects found	30	69, 207
Out of Specifications	41	64
Bowed/Warped	29	66, 281
Incorrect Dimensions	22	37
Excessive Grease	36	237
Damaged in Transit	38	265
No Adhesive	51	298
Wrinkles / Scratches/ Scuffing	16	53
Too Stiff	24	164
Film Not Sealing	46	74
Roll Tension	44	50
Wrong Labeling	57	105
Wrong Registration	60	182
Off-set	31	200

Defect	Main ID	Replaced ID
No Liner	62	249
Slitting Errors	49	236
Damaged Parts	180	260
Dents	111	168, 279
Holes	112	284
Leaking Packaging	188	240
Missing Components	163	212
Missing Labels	144	178
Not Certified	227	247
Not Cleaned	217	303
Odor	192	304
Other	166	210
Out of Specifications	41	64
Packaging Issues	110	137
Wrinkles / Scratches/ Scuffing	83	175, 244
Split Seams	45	101
Water Damage	79	246
Wrong Cut	153	179

Data Cleaning

Remove duplicates

Step #01 & Step #02

```python
import pandas as pd

df = pd.read_excel(r"C:dataset.xlsx", sheet_name="Defects")

replacements = {
    'Crack': 'Cracked','Printing defect': 'Printing defects','Scratch': 'Scratches',
    'Wraped': 'Wraping','Wrong coloring': 'wrong colors’,
    'Wrong Label': 'Worng Labeling','Wrong Spec': 'Wrong Specifications',
    'Out of Spec': 'Out of Specifications','Creases / Wrinkles': 'Wrinkles / Scratches / Scuffing',
    'Scratches': 'Wrinkles / Scratches / Scuffing','Scratched Glass': 'Wrinkles / Scratches / Scuffing',
    'Scratch': 'Wrinkles / Scratches / Scuffing','Scuffed Packeging': 'Wrinkles / Scratches / Scuffing',
    'Bowed': 'Bowed/Warped','Warped': 'Bowed/Warped',
    'Warped Sheets': 'Bowed/Warped','Warping': 'Bowed/Warped',
    'Bad Flaps': 'Flaps - Incorrect Gap','Wrong Flaps': 'Flaps - Incorrect Gap',
    'Gaps': 'Flaps - Incorrect Gap','Gap Variation': 'Flaps - Incorrect Gap',
    'Bad Finish': 'Dimensions - Bad Finishing','Incorrect Dimensions': 'Dimensions - Bad Finishing',
    'Lay Flat Dimension': 'Dimensions - Bad Finishing','Dimensions Wrong': 'Dimensions - Bad Finishing'
}
```

Python

Data Cleaning

Remove duplicates

Step #01 & Step #02

```python
replaced_values = []

for old_value, new_value in replacements.items():
    if old_value in df['Defect'].values:
        replaced_values.append((old_value, new_value))

df['Defect'] = df['Defect'].replace(replacements)

for old_value, new_value in replaced_values:
    print(f"'{old_value}' replaced with '{new_value}'")

output_file_path = r"C:dataset.xlsx"
df.to_excel(output_file_path, index=False)

print(f"Replacements done and saved to {output_file_path}")
```

Python

Data Cleaning

Remove duplicates

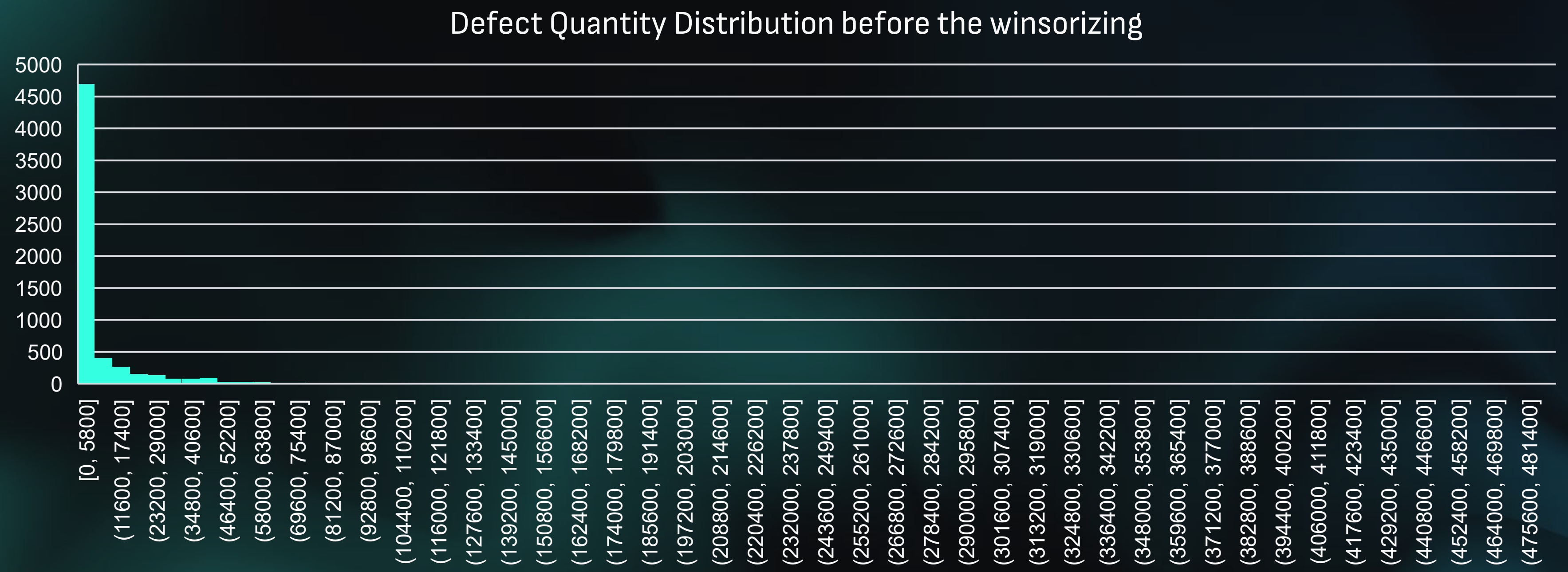
Step #03

```python
import pandas as pd
file_path = r" C:dataset.xlsx "
duplicates_info = df[df.duplicated(subset=['Defect'], keep=False)]
removed_ids = []
for defect in duplicates_info['Defect'].unique():
    ids = duplicates_info[duplicates_info['Defect'] == defect]['Defect ID'].tolist()
    if len(ids) > 1:
        removed_ids.extend(ids[1:])

df_unique = df.drop_duplicates(subset=['Defect'])
output_file_path = r"C:dataset.xlsx"
df_unique.to_excel(output_file_path, index=False)

if removed_ids:
    print("Removed the following IDs due to duplicates:")
    for removed_id in removed_ids:
        print(removed_id)
else: print("No duplicates found.")

print(f"File saved without duplicates to: {output_file_path}")
```

Python

Dealing with outliers

Defect quantity outliers



Dealing with outliers

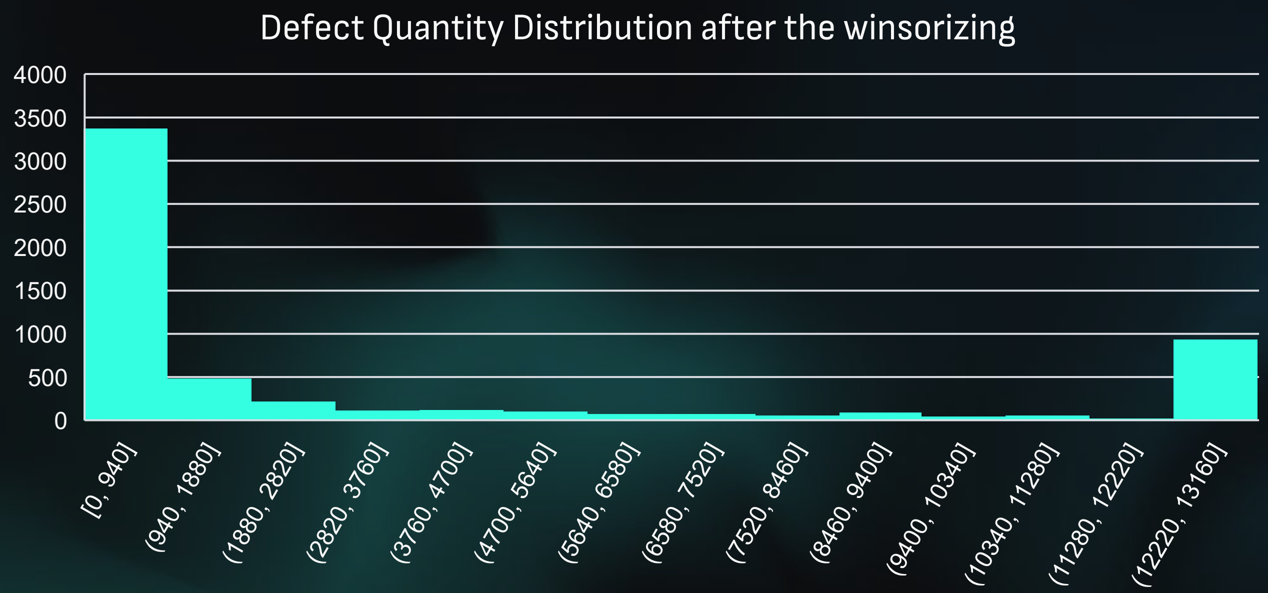
Defect quantity outliers

```python
import pandas as pd
from scipy.stats.mstats import winsorize
df = pd.read_excel("C:dataset.xlsx")
Q1 = df['Defect Qty'].quantile(0.25)
Q3 = df['Defect Qty'].quantile(0.75)
IQR = Q3 - Q1
lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

print(f"Lower bound for Winsorizing: {lower_bound}")
print(f"Upper bound for Winsorizing: {upper_bound}")
original_count = df['Downtime min'].count()
df.loc[df['Defect Qty'] < lower_bound, 'Defect Qty'] = lower_bound
df.loc[df['Defect Qty'] > upper_bound, 'Defect Qty'] = upper_bound
adjusted_count = df.loc[(df['Defect Qty'] < lower_bound) | (df['Defect Qty'] > upper_bound),
'Defect Qty'].count()
print(f"Handled Defect Qty outliers with Winsorizing.")
df.to_excel("C:dataset.xlsx", index=False)
print("Data saved to 'dataset.xlsx'.")
```

Python

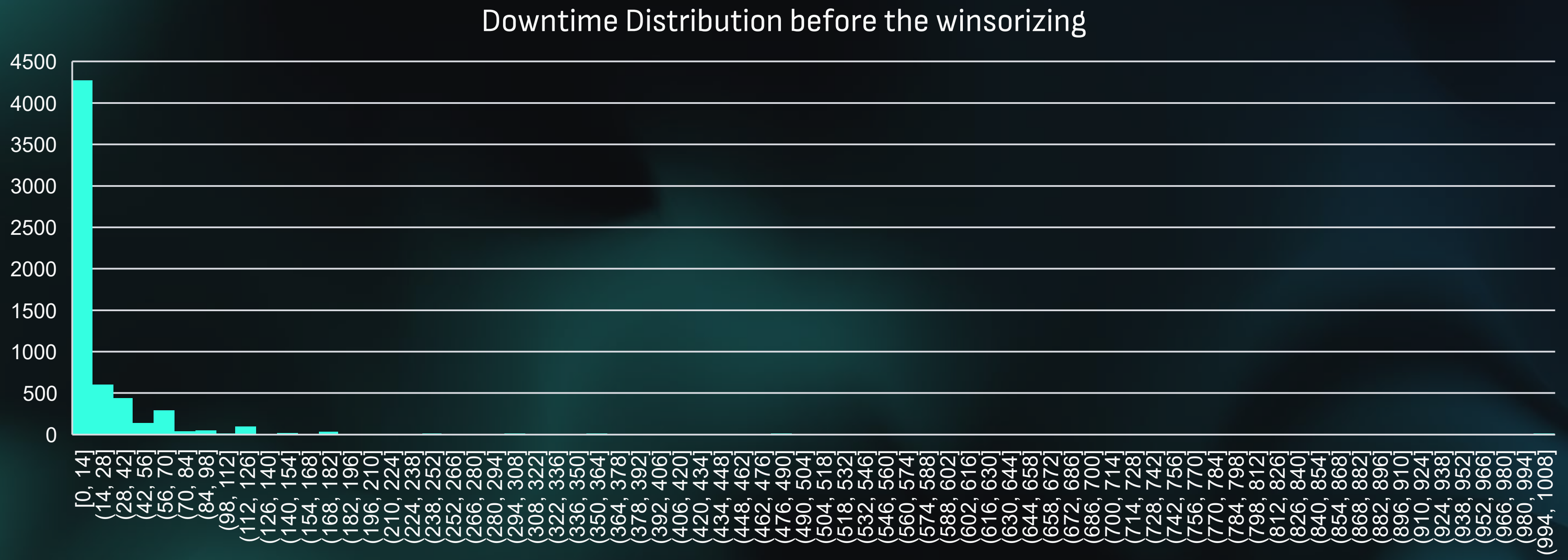
Dealing with outliers

Defect quantity outliers



Dealing with outliers

Downtime outliers



Dealing with outliers

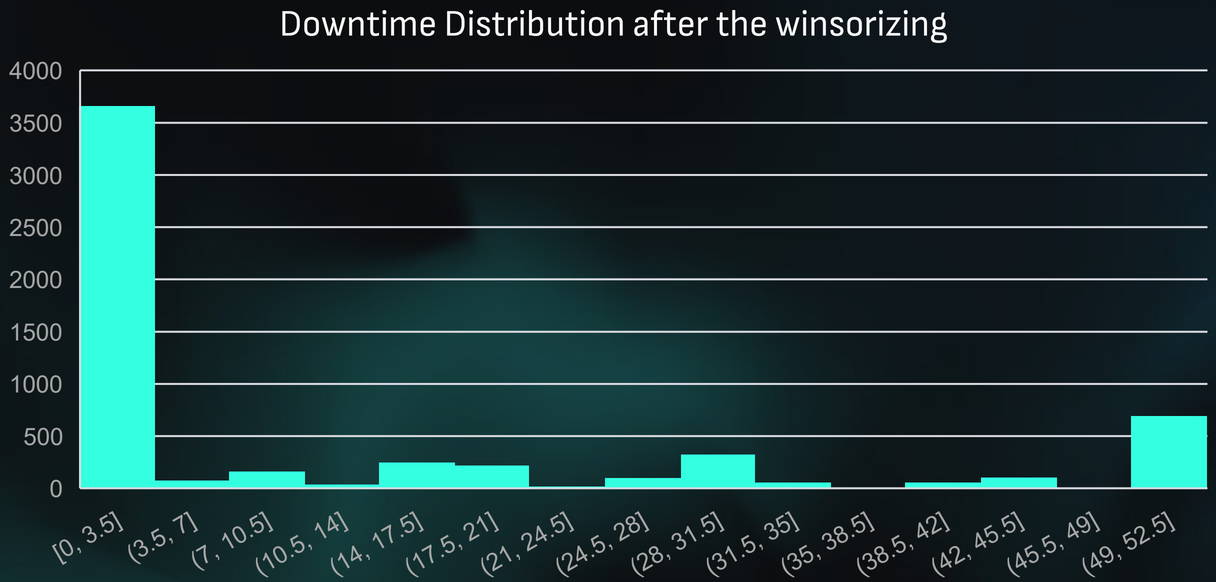
Downtime outliers

```python
import pandas as pd
from scipy.stats.mstats import winsorize

df = pd.read_excel("C:dataset.xlsx")
Q1 = df['Downtime min'].quantile(0.25)
Q3 = df['Downtime min'].quantile(0.75)
IQR = Q3 - Q1
lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

print(f"Lower bound for Winsorizing: {lower_bound}")
print(f"Upper bound for Winsorizing: {upper_bound}")

original_count = df['Downtime min'].count()
df.loc[df['Downtime min'] < lower_bound, 'Downtime min'] = lower_bound
df.loc[df['Downtime min'] > upper_bound, 'Downtime min'] = upper_bound
adjusted_count = df.loc[(df['Downtime min'] < lower_bound) | (df['Downtime min'] >
upper_bound), 'Downtime min'].count()
print(f"Handled Downtime outliers with Winsorizing.")
df.to_excel("C:dataset.xlsx", index=False)
print("Data saved to 'dataset.xlsx'.")
```

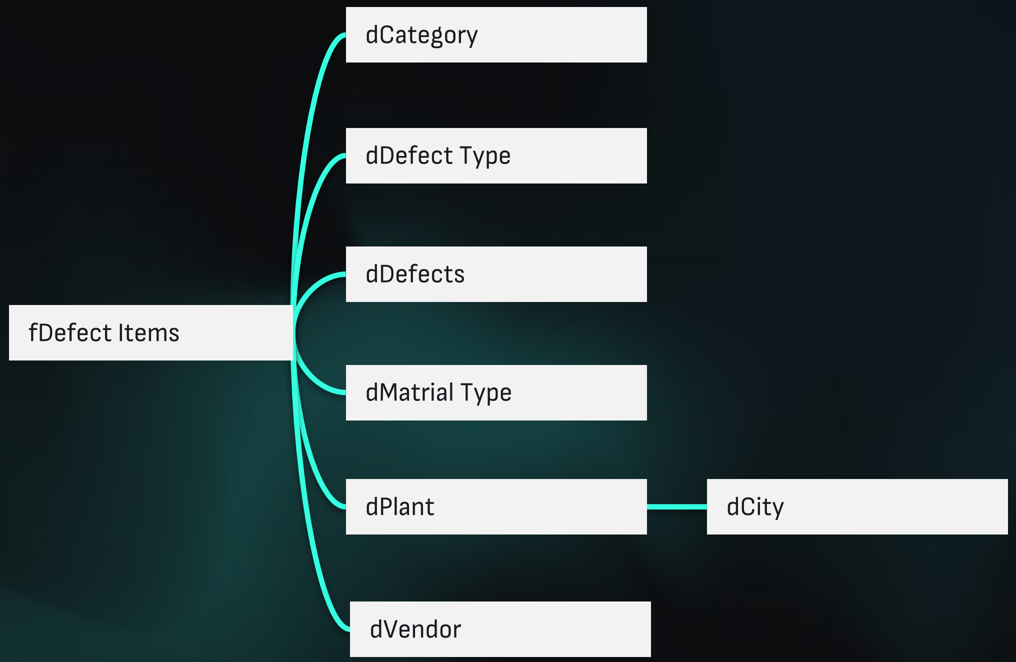
Python

Dealing with outliers

Downtime outliers



Data Modeling



03
Analytical Report

5951 Inspections

Number of Inspections

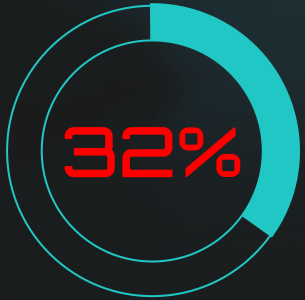
11.46 mins

AVG down time in Mins

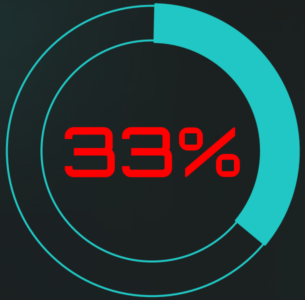
3293 defect

AVG Defects Quantity

Defect Quantity By Defect Type



Rejected

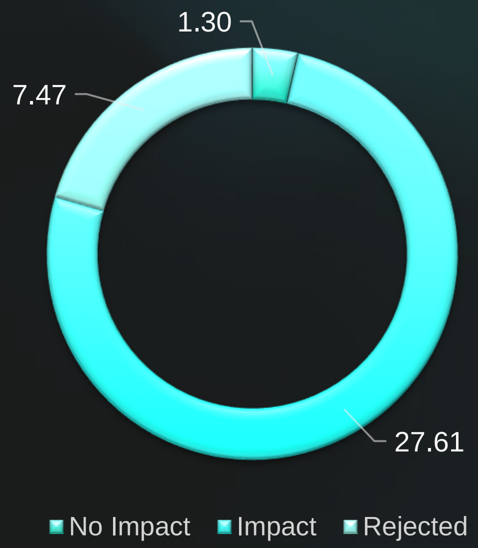


Impact

No Impact

Rejected and Impact product has about 65% of the total defects

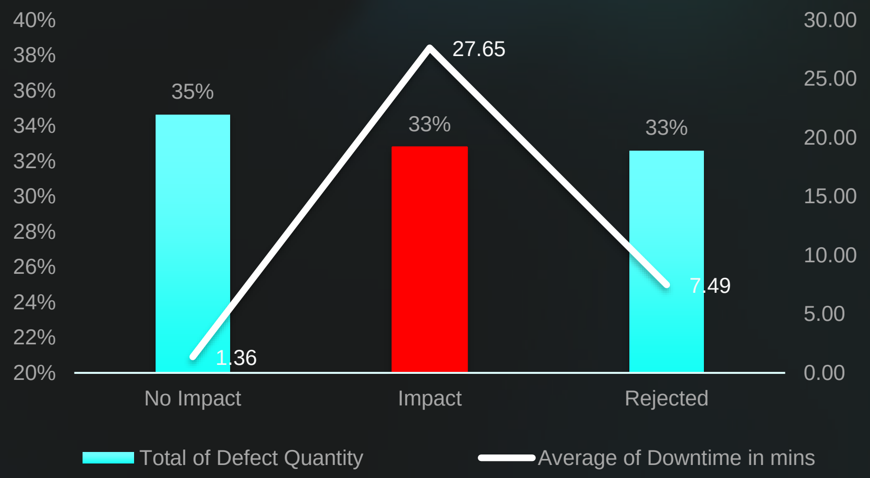
AVG Downtime by Defect type



- No Impact: Low downtime
- Rejected: Downtime is slightly above 8 minutes, comparable to "No Impact."
- Impact: The highest downtime (close to 30 minutes on average)

indicating that defects labeled as "Impact" cause significant operational delays.

AVG Downtime & Defect Qty by Defect type



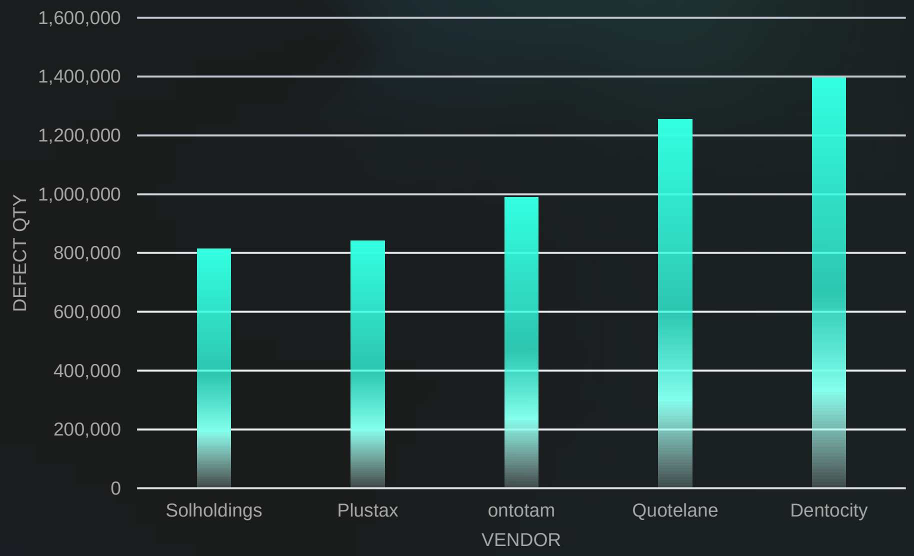
<u>The highest downtime occurred in the "Impact" category with 27.65 minutes.</u>

"No Impact" has the lowest downtime at 1.36 minutes despite having the highest defect quantity (35%).

Both "Impact" and "Rejected" categories have similar defect quantities, but the downtime is much higher for "Impact."

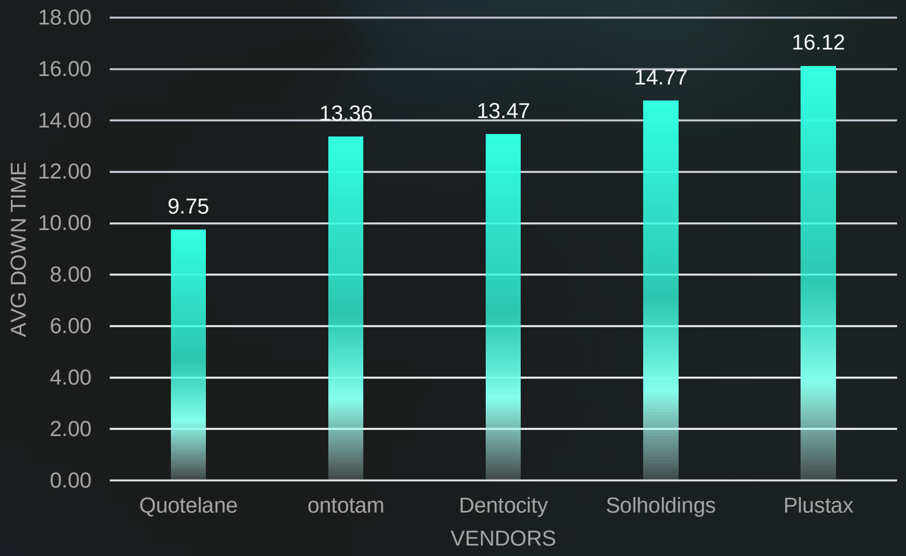
Vendors

Defect quantity by Top 5 vendors



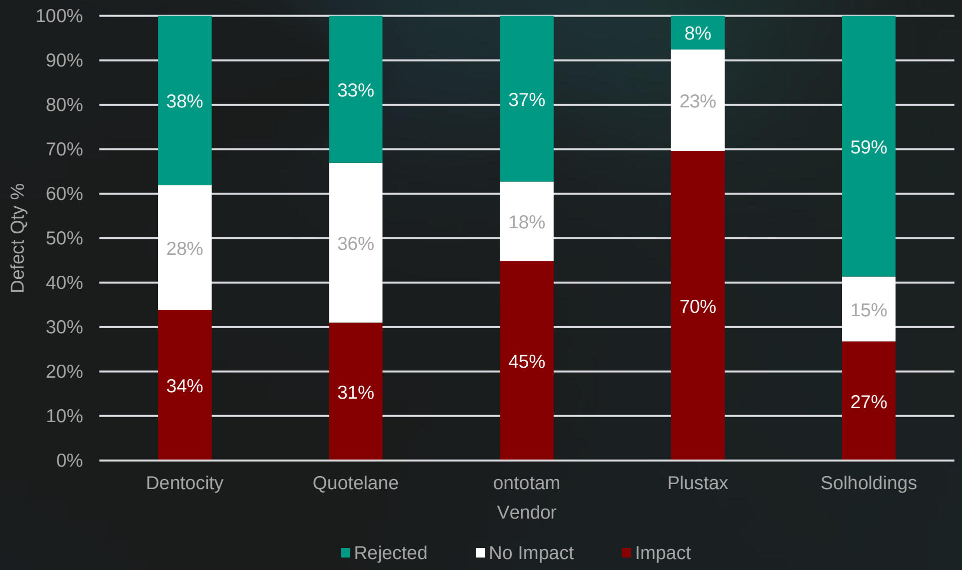
- Dentocity leads with the highest defect quantity, close to 1.4 million, indicating they are a significant contributor to defects among the top vendors.
- Quotelane follows closely behind with over 1.2 million defects, showing a similarly high defect contribution.
- ontotam shows a moderate defect count, just under 1 million.
- Plustax and Solholdings have relatively lower defect quantities, both around 800,000, though still considerable compared to the top two vendors.

Downtime for Top 5 vendors



- Plustax has the highest average downtime, exceeding 16 minutes, indicating they may be causing significant operational delays.
- Solholdings follows closely with an average downtime of about 15 minutes.
- Dentocity and ontotam have moderate downtime levels, both around 13 to 14 minutes, showing some impact but not as severe as the top two.
- Quotelane shows the lowest average downtime (around 11 minutes), making them the least disruptive vendor in terms of downtime.

Defect Types of Top 5 Vendors

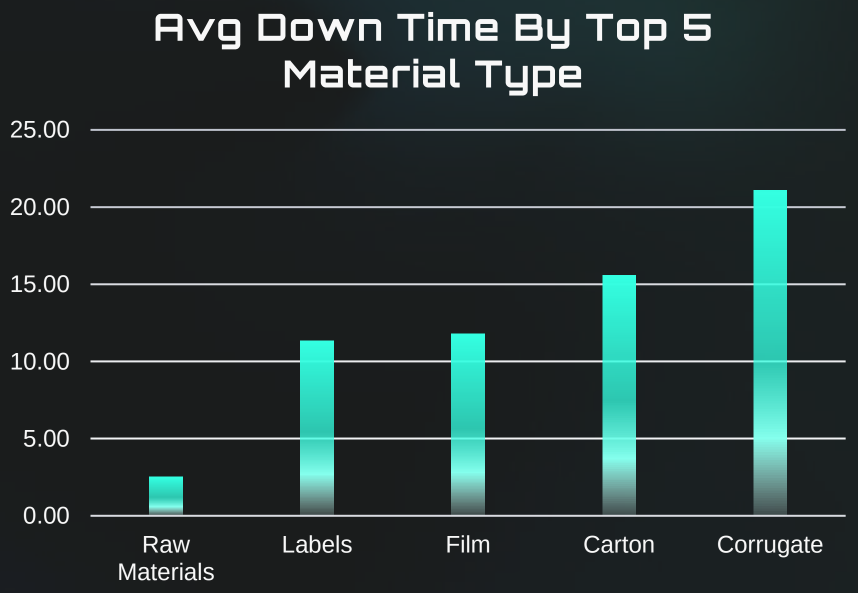


Vendors	Impact	No Impact	Rejected	Grand Total
Dentocity	34%	28%	38%	100%
Quotelane	31%	36%	33%	100%
ontotam	45%	18%	37%	100%
Plustax	70%	23%	8%	100%
Solholdings	27%	15%	59%	100%
Grand Total	40%	25%	35%	100%

- 40% of the top 5 vendor's defective quantities are of the impact type. Because of that, their average downtime is high.

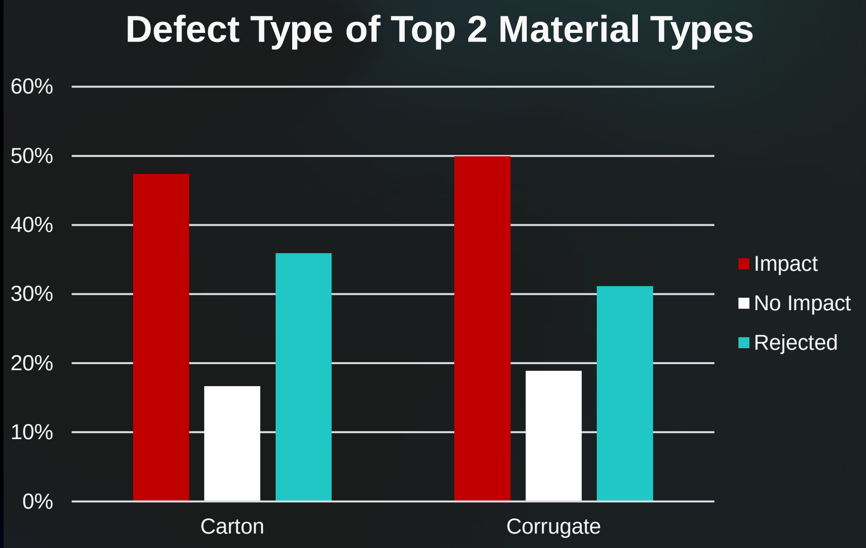
Material Type

Avg Down Time by Top 5 Material Type



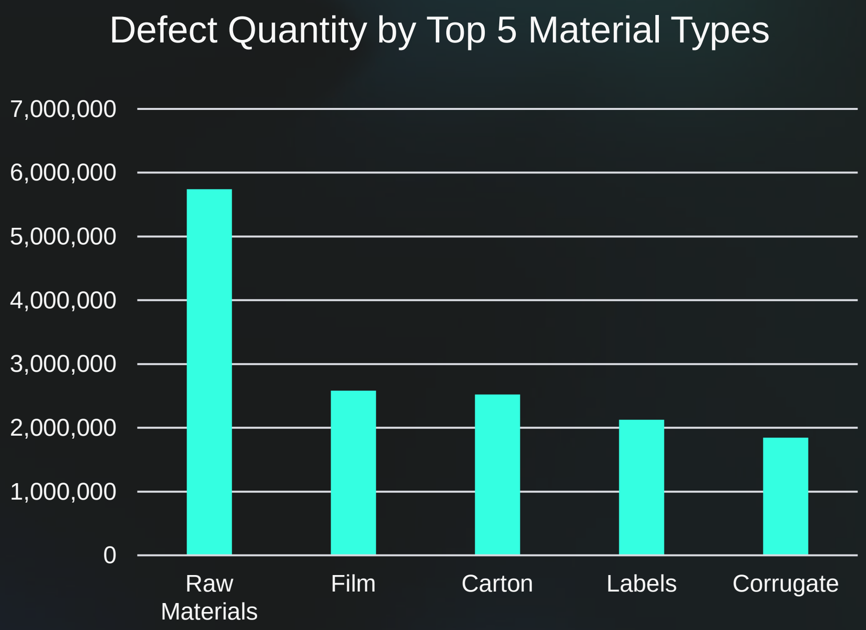
- **Corrugated Materials Dominate Downtime:** Corrugated materials experience the highest average downtime among the five categories, significantly surpassing the others.
- **Consistency in Downtime for Carton, Film, and Labels:** Carton, film, and labels exhibit relatively similar levels of average downtime, suggesting a moderate degree of consistency in their operational disruptions.
- **Raw Materials Lag Behind:** Raw materials have the lowest average downtime, indicating a potentially more stable or efficient supply chain or production process compared to the other materials.

Defect Type Comparison for Top 2 Materials



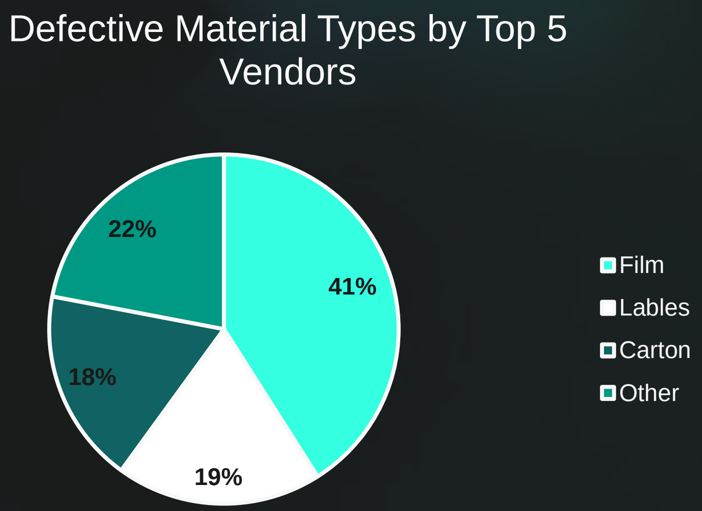
- **Corrugated and Carton** are the top 2 material types in average downtime. This is because most of their defective quantities are of the impact type.

Material Matters: Defect Quantity Disparity



- ❑ Raw materials experience the least defect quantity, while corrugated materials have the highest.
- ❑ Film, carton, and labels exhibit similar levels of defect quantity.

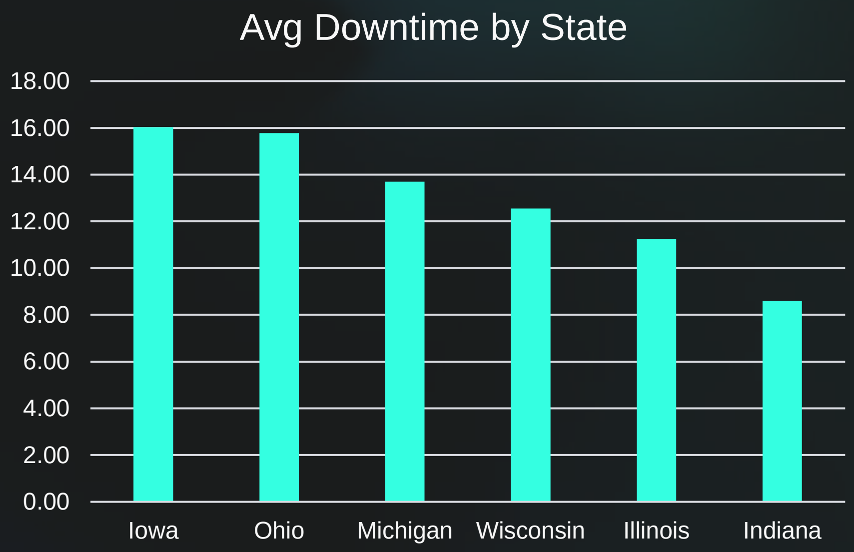
The Impact of Top 5 Vendors on Material Types



- The most defective material types of the top 5 vendors are labels, cartoons, and films. This makes these four material types have high defective quantities.

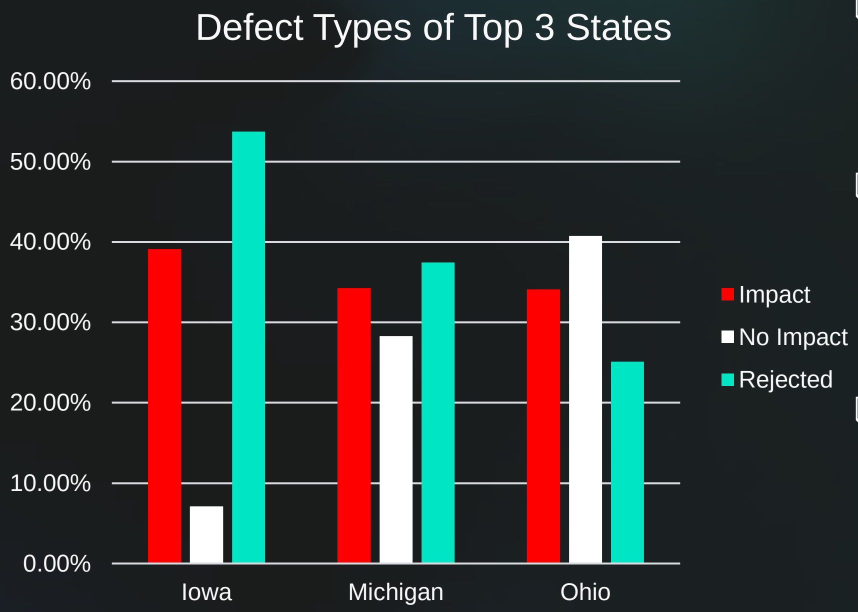
State

Iowa and Ohio Outliers: State-Level Downtime Disparity



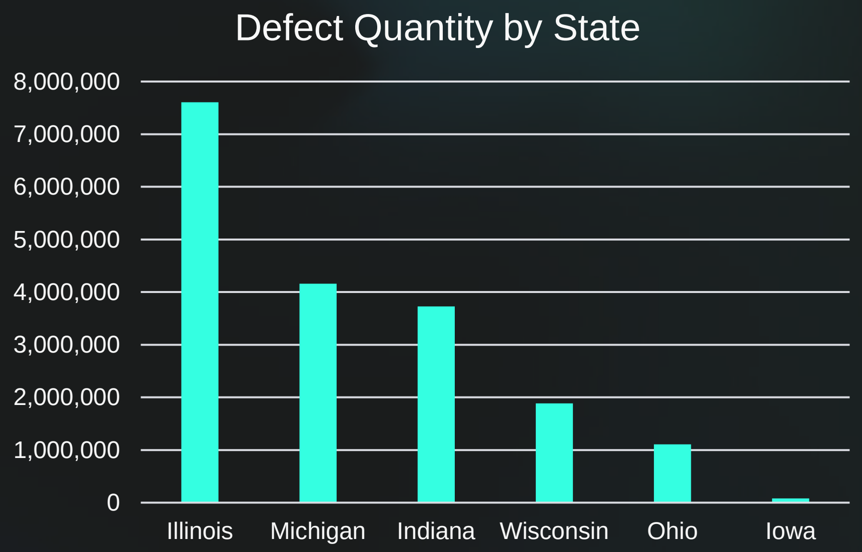
- **Iowa and Ohio Lead in Downtime:** Iowa and Ohio experience the highest average downtime, suggesting potential operational challenges or disruptions in these states.
- **Michigan, Wisconsin, and Illinois Follow:** These states exhibit moderate levels of downtime, indicating a need for further investigation into potential causes.
- **Indiana Has Lowest Downtime:** Indiana demonstrates the lowest average downtime, suggesting more efficient operations or fewer disruptions compared to the other states.

Defect Type Disparity in Top States



- **Iowa's Dominance in Rejected Defects:** Iowa has the highest percentage of rejected defects, indicating potential quality control issues or non-compliance with standards.
- **Michigan's Balanced Defect Types:** Michigan exhibits a more balanced distribution of defect types, with a significant portion having no impact, suggesting that while defects occur, they may not always result in major consequences.
- **Ohio's Impact-Focused Defects:** Ohio has a higher proportion of defects that have an impact, suggesting that while the overall defect rate may be lower, the defects that do occur are more likely to cause problems.

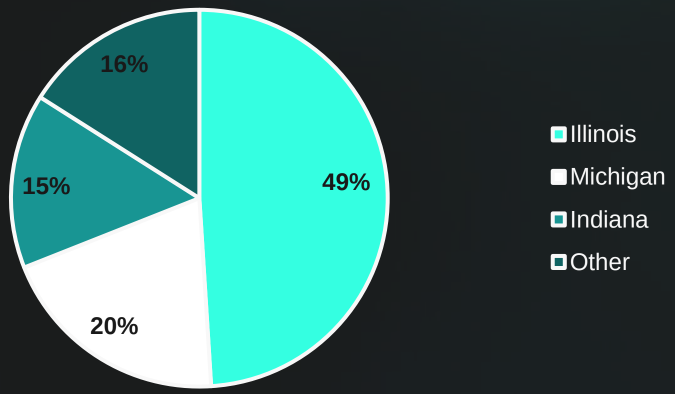
State-by-State Defect Disparity



- **Illinois Leads in Defects:** Illinois experiences a significantly higher number of defects compared to the other states, indicating a potential quality control issue or other underlying factors.
- **Michigan and Indiana Follow:** Michigan and Indiana report substantial defect quantities, suggesting similar challenges in quality or production processes.
- **Wisconsin, Ohio, and Iowa Show Lower Defects:** These states exhibit considerably fewer defects, potentially due to more effective quality management practices or different production conditions.

Vendor State Drives Defect Dominance

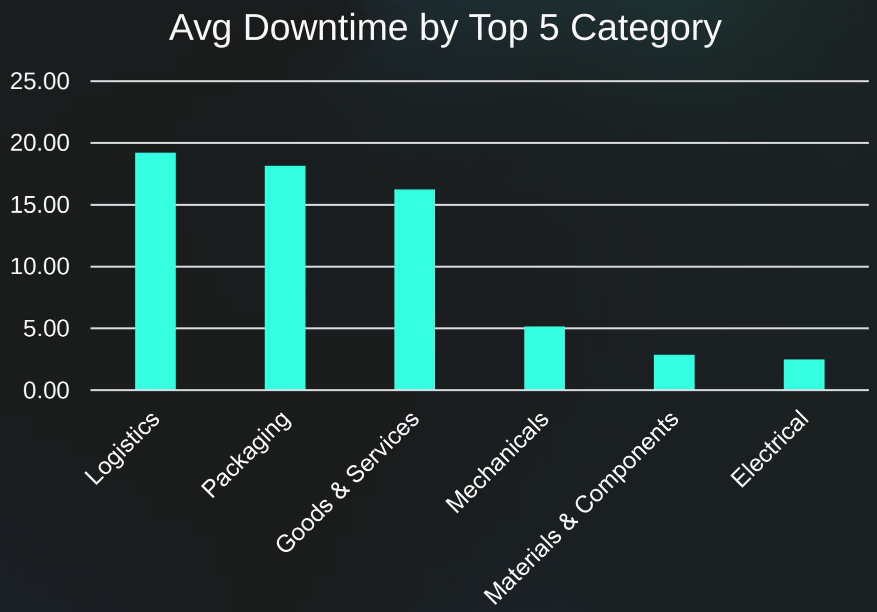
Defect Quantity by Top 5 Vendors' State



- **Michigan Dominates Defects:** Michigan accounts for nearly half (49%) of the total defects, indicating a significant quality issue or other underlying problem with vendors from this state.
- Illinois and Indiana Follow: Illinois and Indiana contribute substantial portions of the defects, suggesting that quality challenges may also exist among vendors in these states.
- Other States' Influence: The remaining 20% of defects are attributed to vendors from other states, highlighting a more diverse range of quality issues across the vendor base.

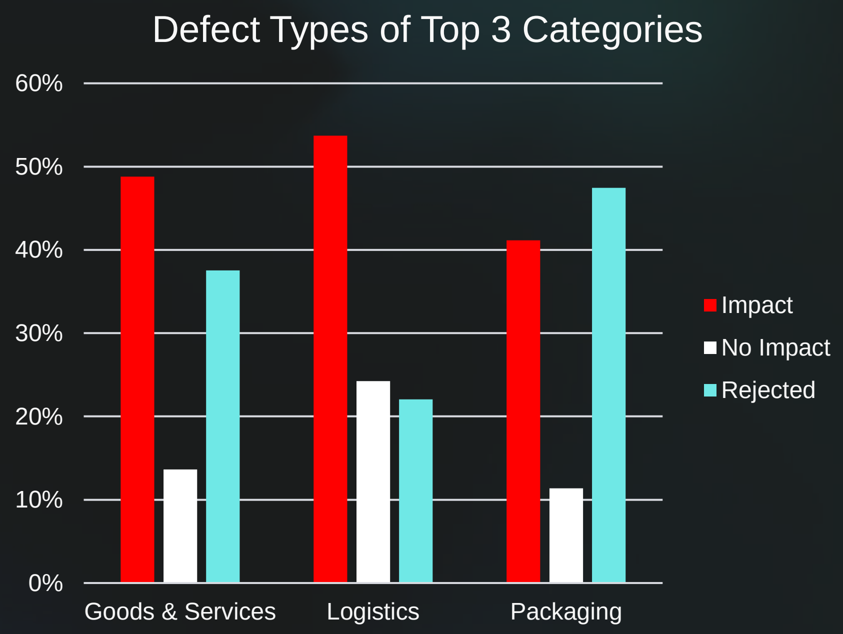
Category

Logistics and Packaging Lag: Category-Based Downtime Disparity



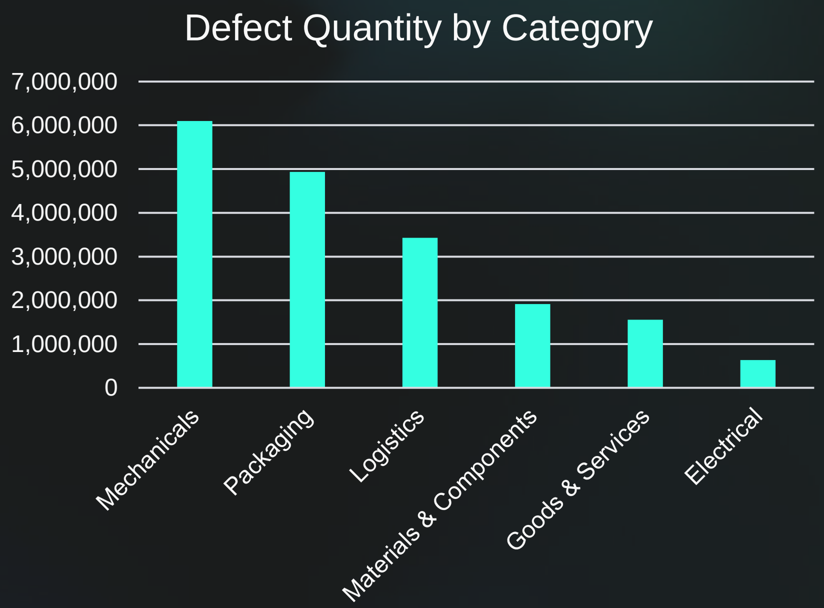
- **Logistics and Packaging Lead in Downtime:** Logistics and packaging experience the highest average downtime, indicating potential challenges or disruptions in these areas.
- **Goods & Services Follow:** Goods & Services have a moderately high level of downtime, suggesting that issues in this category may also contribute to overall operational inefficiencies.
- **Mechanical and Materials & Components Lag Behind:** Mechanicals and Materials & Components exhibit significantly lower downtime, suggesting more stable or efficient processes in these areas.

Defect Type Disparity in Top 3 Categories



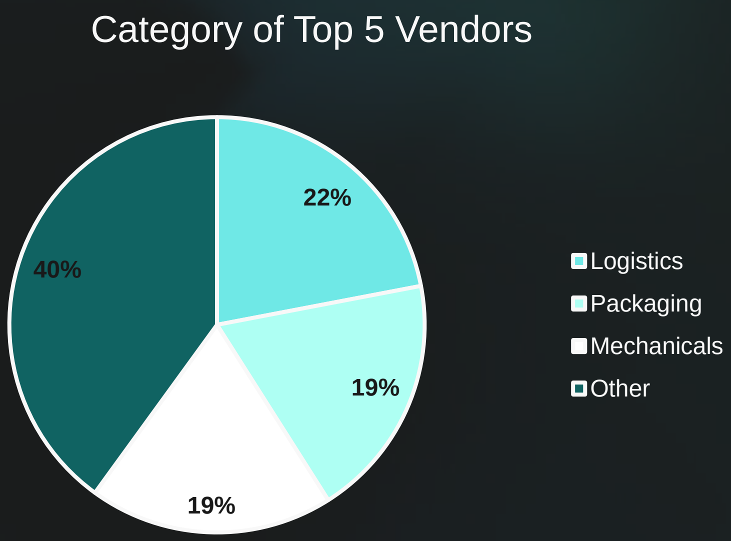
- **Goods & Services Lead in Rejected Defects:** Goods & Services have the highest percentage of rejected defects, indicating potential quality control issues or non-compliance with standards.
- **Logistics and Packaging Exhibit Similar Defect Profiles:** Logistics and packaging show comparable distributions of defect types, with a higher proportion of defects having no impact.
- **Impactful Defects in Logistics and Packaging:** While rejected defects are higher in Goods & Services, Logistics and Packaging still have a significant percentage of defects that have an impact, suggesting potential operational or quality issues.

Defect Type Disparity in Top 3 Categories



- **Mechanical Defects Dominate:** Mechanical defects far outnumber defects in other categories, indicating a significant quality issue or underlying problem in this area.
- **Packaging and Logistics Follow:** Packaging and logistics also experience a relatively high number of defects, suggesting potential challenges in these areas.
- **Materials & Components, Goods & Services, and Electrical Lag Behind:** These categories exhibit significantly fewer defects, potentially due to more effective quality management or different production processes.

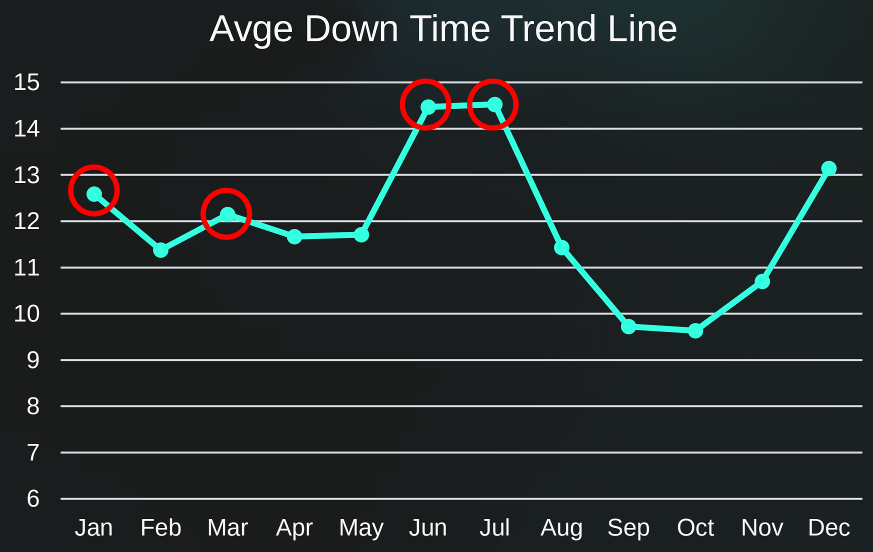
Logistics Leadership: Vendor Category Dominance



- **Logistics Dominates Vendor Categories:** Logistics accounts for a significant portion (40%) of the top 5 vendors, suggesting that this category plays a crucial role in the supply chain or overall business operations.
- **Packaging and Mechanics Follow:** Packaging and Mechanics each represent 22% of the top 5 vendors, indicating their importance in the supply chain as well.
- **Other Categories Contribute:** The remaining 19% is distributed among other categories, suggesting a diverse range of vendors involved in the business.

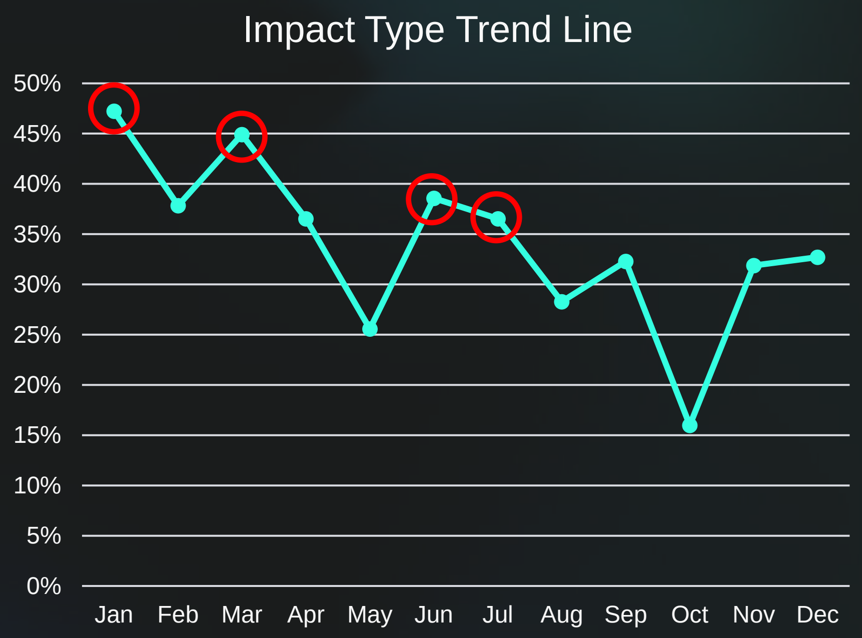
Date

Summer Surge: Downtime Fluctuations Throughout the Year



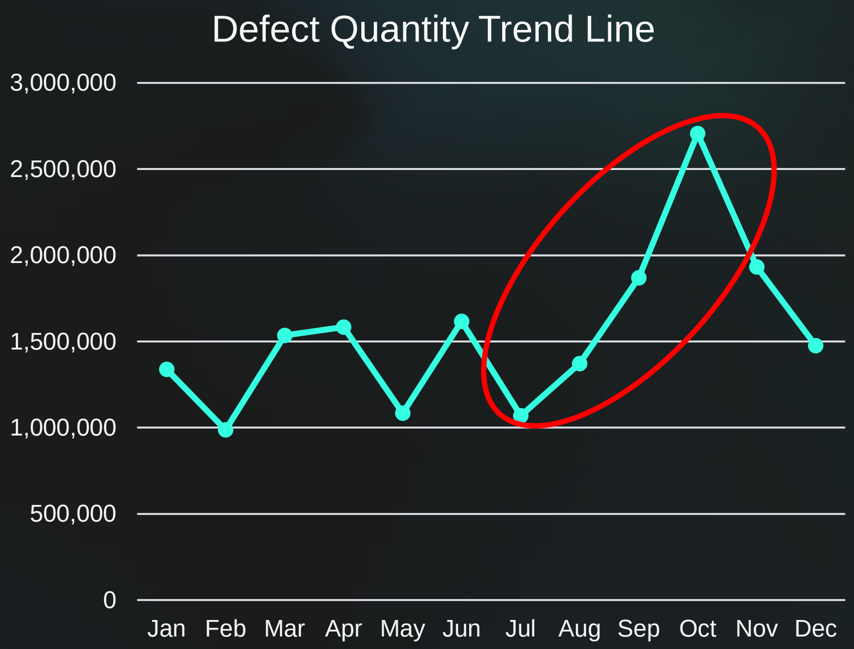
- **Initial Decrease:** The trend line starts with a decline from January to February, suggesting an improvement in operational efficiency or a reduction in disruptions during that period.
- **Mid-Year Spike:** A significant increase in downtime occurs between June and July, indicating potential challenges or disruptions during this timeframe.
- **Summer Recovery:** The line then dips back down from July to September, suggesting a recovery in operational performance.
- **Late-Year Upsurge:** A final spike in downtime is observed from October to December, indicating potential issues or challenges towards the end of the year.

Impact Type Trend Analysis



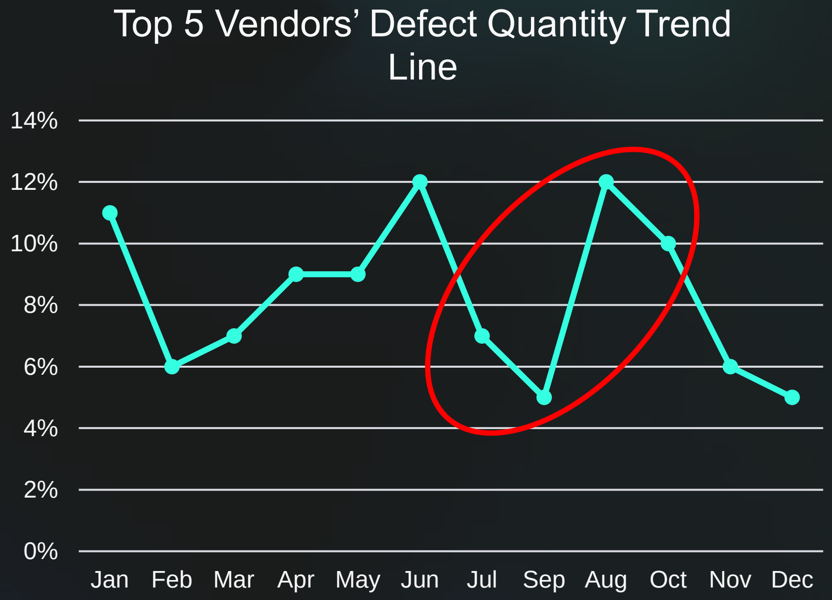
- **Initial Decline:** The trend line starts with a decrease from January to February, suggesting a reduction in the impact of defects or issues during that period.
- **Mid-Year Fluctuations:** The line exhibits fluctuations between March and August, indicating varying levels of impact throughout this timeframe.
- **Late-Year Stability:** The trend line becomes relatively stable from September to December, suggesting a more consistent level of impact.

Impact Type Trend Analysis



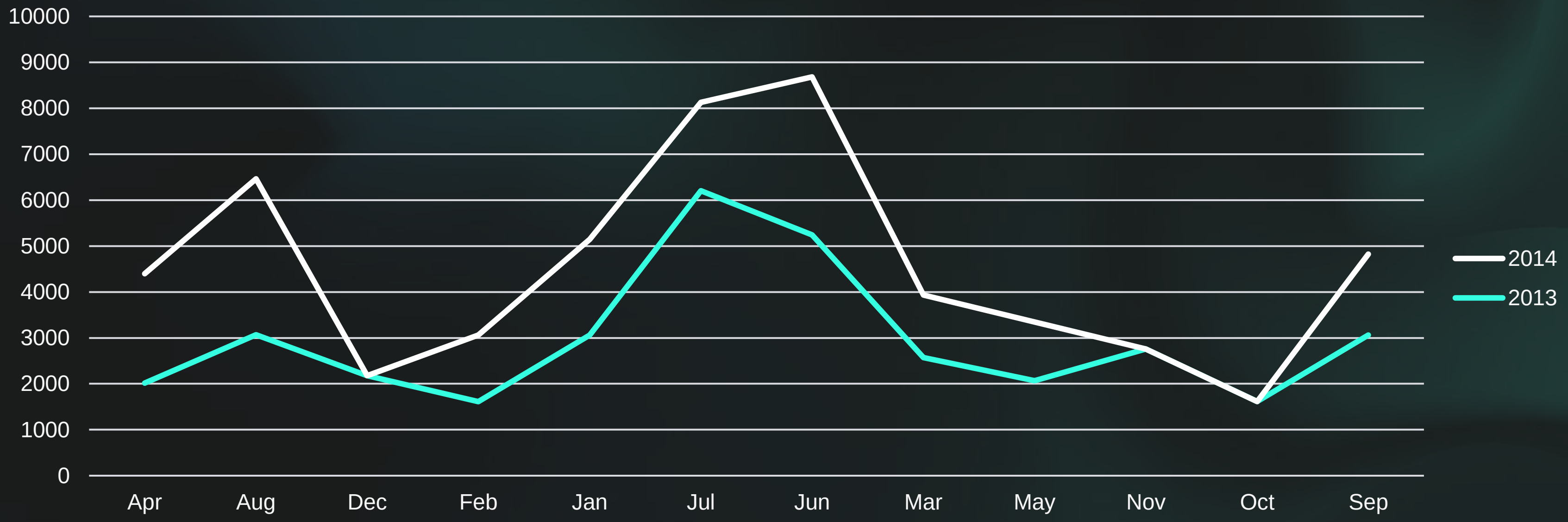
- **Initial Decline:** The trend line starts with a decrease from January to February, suggesting a reduction in defect quantity during that period.
- **Mid-Year Fluctuations:** The line exhibits fluctuations between March and October, indicating varying levels of defects throughout this timeframe.
- **Late-Year Surge**: A significant increase in defects occurs from October to November, followed by a sharp decline in December.

Early Improvement: Defect Quantity Trend for Top 5 Vendors



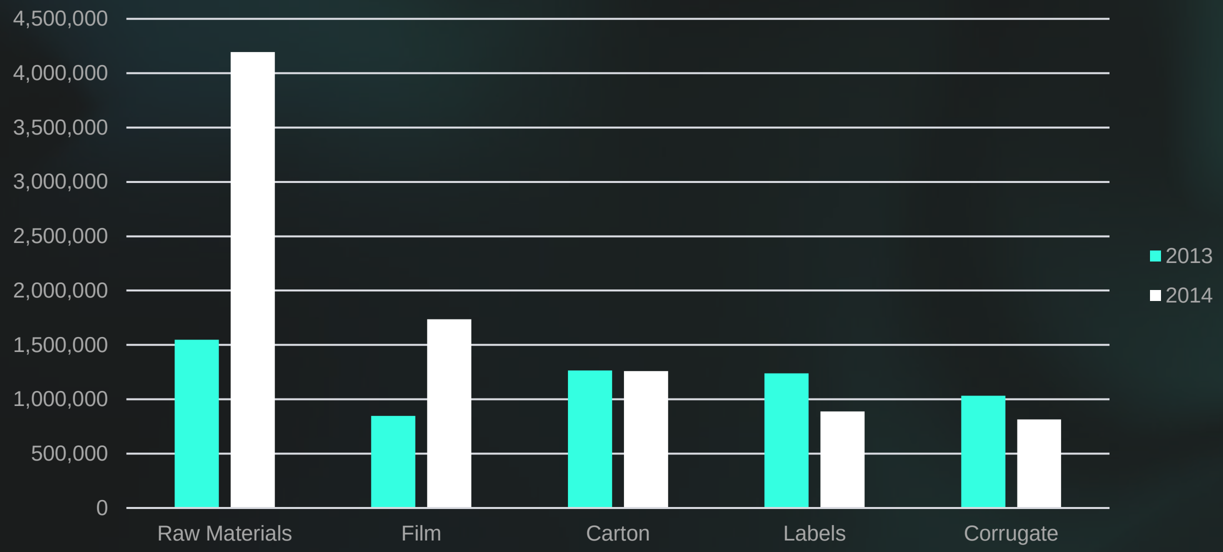
- **Initial Decline:** The trend line starts with a decrease from January to February, suggesting an improvement in quality or a reduction in defects from the top 5 vendors during that period.
- **Mid-Year Fluctuations:** The line exhibits fluctuations between March and August, indicating varying levels of defect quantity among the top 5 vendors.
- **Late-Year Stability:** The trend line becomes relatively stable from September to December, suggesting a more consistent level of defects from these vendors.

Avg Defect Quantity Timeline



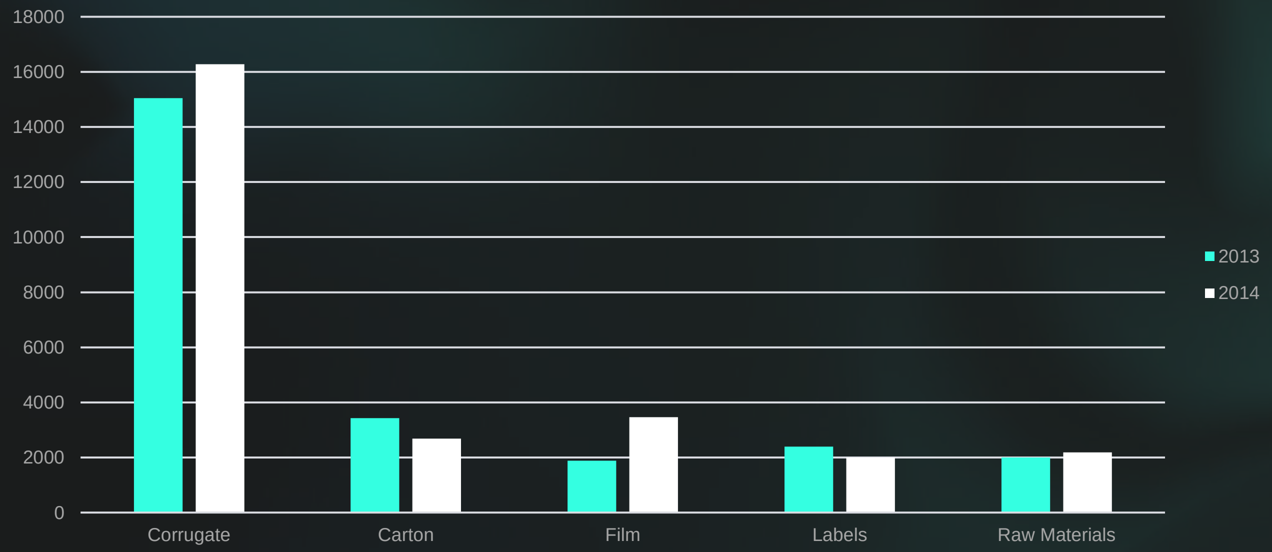
2014 experiences more volatility with significant peaks and drops, while 2013 has a more gradual trend. The highest point in both years occurs in June, but 2014 generally displays more variability throughout the period.

Defect Category by years



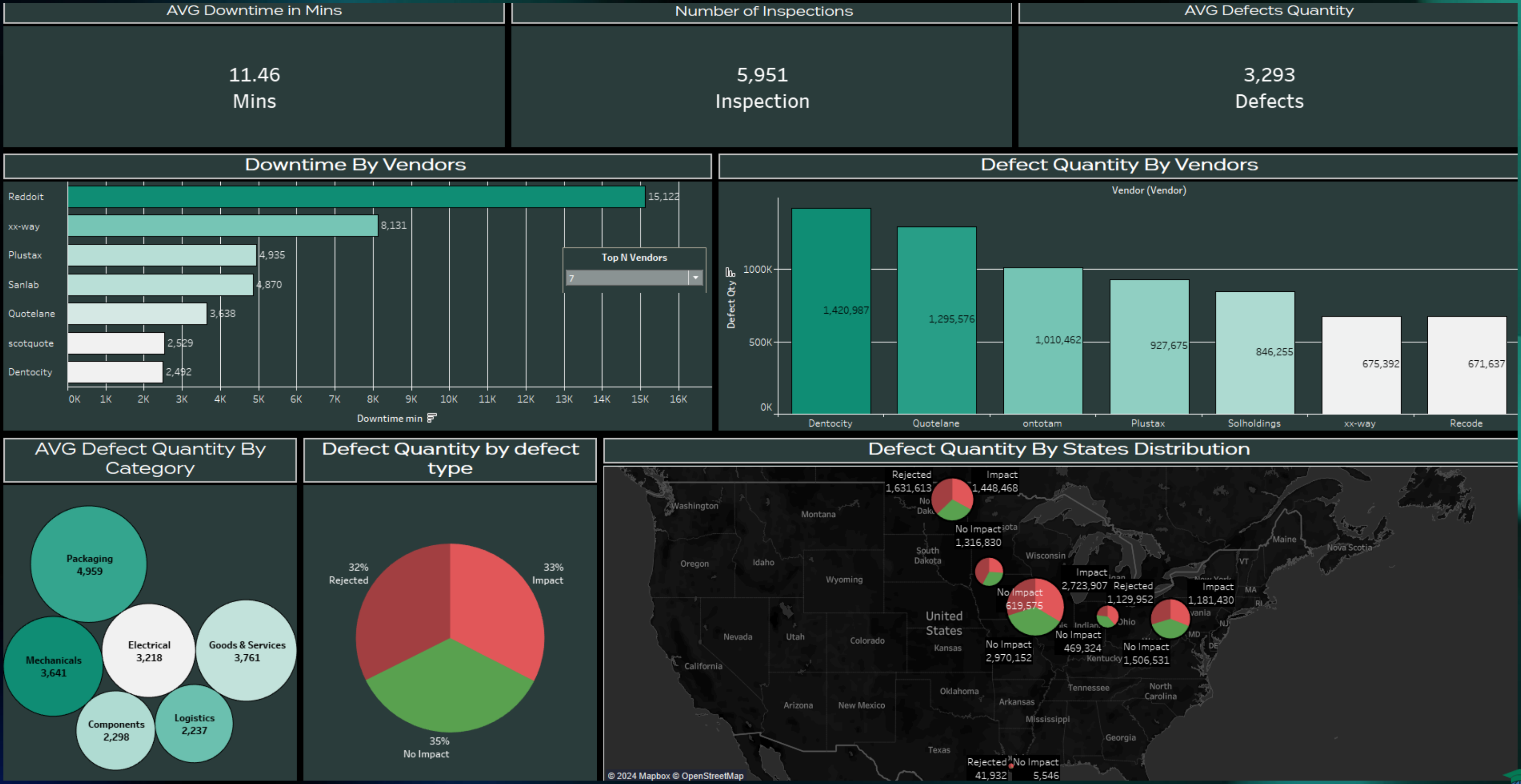
- 2014 shows large increases in **Raw Materials** and Film compared to 2013.
- Labels had higher value in 2013.
- Carton and Corrugate categories show similar performance across both years

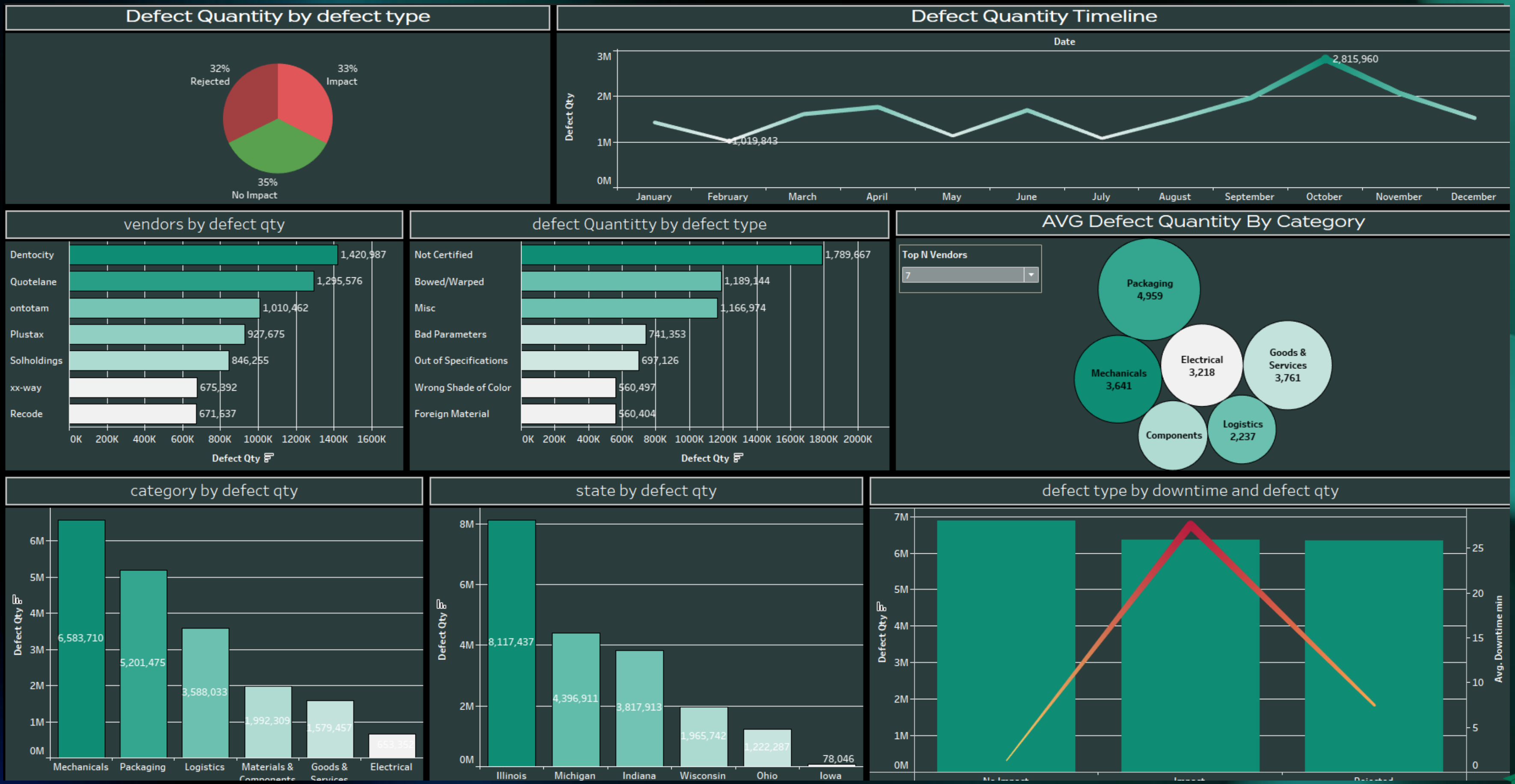
Downtime of Categories by years

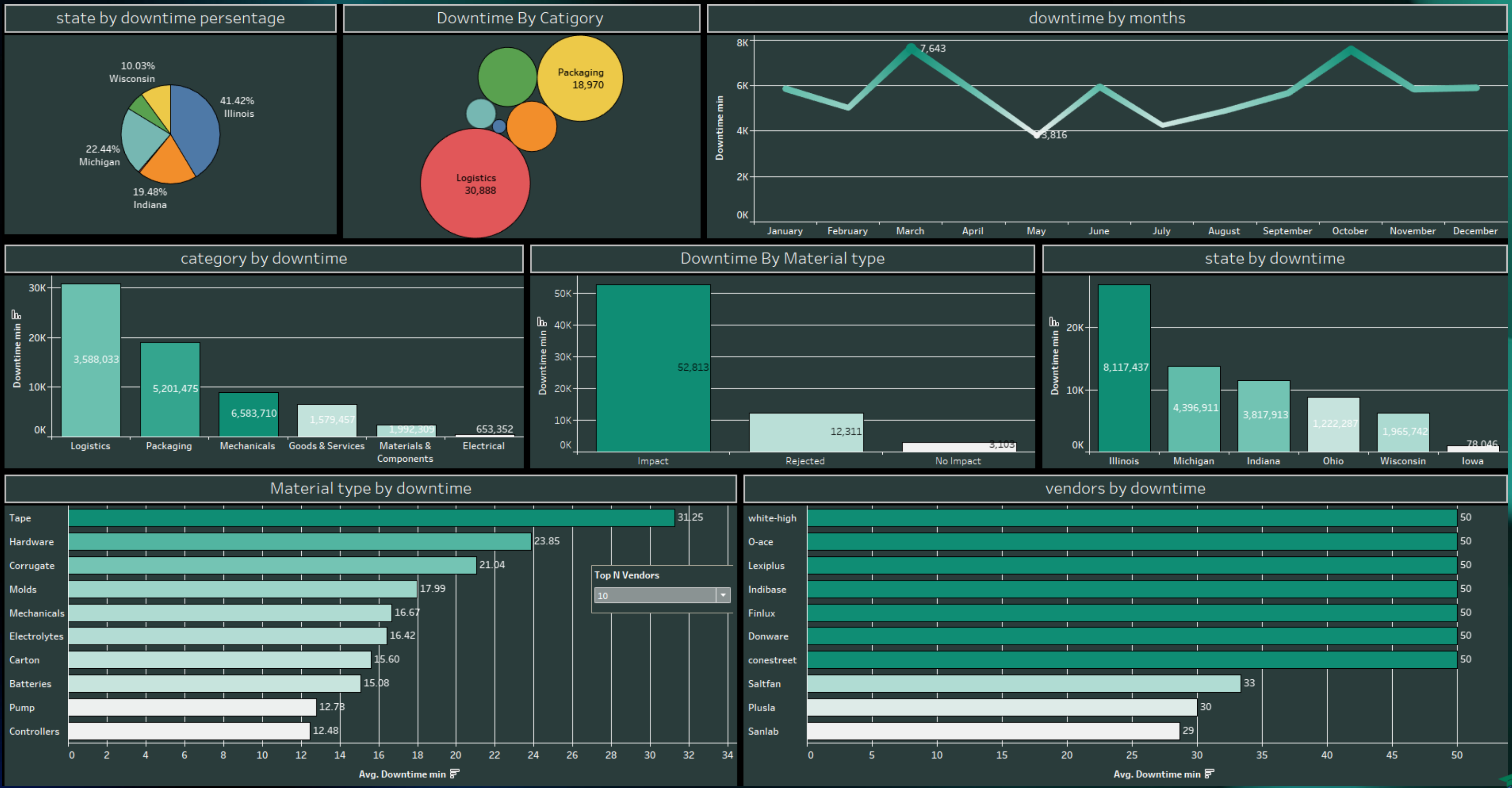


- 2014 shows large increases in Corrugate and Film compared to 2013.
- Corrugate had higher value in 2013.
- **Raw Materials** and labels categories show similar performance across both years

Dashboards


AVG Downtime in Mins
11.46
Mins
Number of Inspections
5,951
Inspection
AVG Defects Quantity
3,293
Defects
Downtime By Vendors
Reddoit 15,122
xx-way 8,131
Plustax 4,935
Sanlab 4,870
Quotelane 3,638
scotquote 2,529
Dentocity 2,492
Top N Vendors
7
0K 1K 2K 3K 4K 5K 6K 7K 8K 9K 10K 11K 12K 13K 14K 15K 16K
Downtime min
Defect Quantity By Vendors
Vendor (Vendor)
Defect Qty
1000K
500K
0K
Dentocity 1,420,987
Quotelane 1,295,576
ontotam 1,010,462
Plustax 927,675
Solholdings 846,255
xx-way 675,392
Recode 671,637
AVG Defect Quantity By Category
Packaging 4,959
Electrical 3,218
Goods & Services 3,761
Mechanicals 3,641
Components 2,298
Logistics 2,237
Defect Quantity by defect type
32% Rejected
33% Impact
35% No Impact
Defect Quantity By States Distribution
Rejected 1,631,613
Impact 1,448,468
No Impact 1,316,830
Impact 2,723,907
Rejected 1,129,952
Impact 1,181,430
No Impact 619,575
No Impact 2,970,152
No Impact 469,324
No Impact 1,506,531
Rejected 41,932
No Impact 5,546
Washington
Montana
Oregon
Idaho
Wyoming
South Dakota
Wisconsin
Maine
Nova Scotia
Nevada
Utah
Colorado
United States
Kansas
California
Oklahoma
Arkansas
Tennessee
North Carolina
Mississippi
Georgia
Arizona
New Mexico
Texas
Ohio
Kentucky
© 2024 Mapbox © OpenStreetMap


Defect Quantity by defect type
32% Rejected
33% Impact
35% No Impact
Defect Quantity Timeline
Date
Defect Qty
3M
2M
1M
0M
2,815,960
1,019,843
January February March April May June July August September October November December
vendors by defect qty
Dentocity 1,420,987
Quotelane 1,295,576
ontotam 1,010,462
Plustax 927,675
Solholdings 846,255
xx-way 675,392
Recode 671,637
0K 200K 400K 600K 800K 1000K 1200K 1400K 1600K
Defect Qty
defect Quantitty by defect type
Not Certified 1,789,667
Bowed/Warped 1,189,144
Misc 1,166,974
Bad Parameters 741,353
Out of Specifications 697,126
Wrong Shade of Color 560,497
Foreign Material 560,404
0K 200K 400K 600K 800K 1000K 1200K 1400K 1600K 1800K 2000K
Defect Qty
AVG Defect Quantity By Category
Top N Vendors
7
Packaging 4,959
Mechanicals 3,641
Electrical 3,218
Goods & Services 3,761
Components
Logistics 2,237
category by defect qty
Defect Qty
6M 5M 4M 3M 2M 1M 0M
Mechanicals 6,583,710
Packaging 5,201,475
Logistics 3,588,033
Materials & Components 1,992,309
Goods & Services 1,579,457
Electrical
state by defect qty
Defect Qty
8M 6M 4M 2M 0M
Illinois 8,117,437
Michigan 4,396,911
Indiana 3,817,913
Wisconsin 1,965,742
Ohio 1,222,287
Iowa 78,046
defect type by downtime and defect qty
Defect Qty
7M 6M 5M 4M 3M 2M 1M 0M
Avg. Downtime min
25 20 15 10 5 0


state by downtime persentage
10.03%
Wisconsin
41.42%
Illinois
22.44%
Michigan
19.48%
Indiana
Downtime By Catigory
Packaging
18,970
Logistics
30,888
downtime by months
Downtime min
8K
6K
4K
2K
0K
7,643
3,816
January
February
March
April
May
June
July
August
September
October
November
December
category by downtime
Downtime min
30K
20K
10K
0K
3,588,033
5,201,475
6,583,710
1,579,457
1,992,309
653,352
Logistics
Packaging
Mechanicals
Goods & Services
Materials & Components
Electrical
Downtime By Material type
Downtime min
50K
40K
30K
20K
10K
0K
52,813
12,311
3,103
Impact
Rejected
No Impact
state by downtime
Downtime min
20K
10K
0K
8,117,437
4,396,911
3,817,913
1,222,287
1,965,742
78,046
Illinois
Michigan
Indiana
Ohio
Wisconsin
Iowa
Material type by downtime
Tape
31.25
Hardware
23.85
Corrugate
21.04
Molds
17.99
Mechanicals
16.67
Electrolytes
16.42
Carton
15.60
Batteries
15.08
Pump
12.78
Controllers
12.48
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
Avg. Downtime min
Top N Vendors
10
vendors by downtime
white-high
50
O-ace
50
Lexiplus
50
Indibase
50
Finlux
50
Donware
50
conestreet
50
Saltfan
33
Plusla
30
Sanlab
29
0
5
10
15
20
25
30
35
40
45
50
Avg. Downtime min

04
Take aways

1. Supplier Performance Evaluation:

Top Contributors to Defects: Vendors like Dentocity and Quotelane contribute the highest defect quantities, indicating they need closer scrutiny and possible renegotiation of contracts or improvements in quality control processes.

Plustax and Solholdings also show high downtime, suggesting that these suppliers could be causing operational inefficiencies, even if their defect quantities are lower.

2. Operational Efficiency through Downtime Reduction:

Impact of Defect Types: Defects classified as "Impact" lead to the highest downtime, averaging around 27 minutes. This suggests a need for better preventative measures in identifying and reducing high-impact defects to improve overall operational efficiency.

No Impact Defects: While "No Impact" defects are most frequent, they have minimal effect on downtime, suggesting that the business can tolerate such defects if prioritized correctly.

3. Material and Vendor Strategy:

Material-Type Focus: Corrugated materials lead to average downtime, signaling potential issues in the handling or sourcing of this material. Investing in better quality or optimizing processes related to corrugated materials could reduce downtime significantly.

State Disparities: States like Iowa and Ohio face the highest downtime, meaning-focused improvements on operations in these regions can lead to overall efficiency gains.

4. Supply Chain Optimization:

Logistics and Packaging Downtime: Categories such as Logistics and Packaging experience the highest average downtimes, indicating that these are the areas in most need of process improvement.
Material Management: Raw materials exhibit the least downtime and defect quantity, suggesting this part of the supply chain is functioning efficiently. It could be used as a model for improving other material types.

5. Seasonal Trends and Preparation:

Mid-Year and Late-Year Challenges: The business experiences downtime spikes in the mid-year (June-July) and at the end of the year (October-December). This suggests potential seasonal challenges, such as labor shortages or higher defect rates during peak production periods. Preparing for these challenges in advance with resource allocation and process improvements could mitigate their impact.

6. Focus on Key States:

State-Specific Issues: Michigan and Illinois lead in defect quantity, with Michigan accounting for nearly 50% of total defects. These states may benefit from targeted quality control interventions and supplier reviews to reduce defects and improve efficiency.

7. Vendor and Material Type Interactions:

Vendor-Material Relationships: Labels, cartons, and films are the most defective material types associated with top vendors. Focusing on these specific vendor-material relationships could yield improvements in reducing overall defect rates.
These insights provide clear directions for optimizing supplier quality, reducing downtime, and strategically managing high-impact defects.

05

Next Step

Data Grouping

An important future work is to classify "defects" into data groups so they'll be easy to use for analysis. Plus, the information extracted from them will be really useful for a high-level and public audience. It's a perfect step to using existing GenAI models or to design a model that is fed by related business-domain data, which is the supply-chain business domain, vendors' materials data and the defects commonly occurring to them.

For example, we can notice that many "defects" data records have a common problem or that describes nearly the same thing, as clarified in the attached sample of data.

Defect ID	Defect
33	No Bar Code
129	Wrong Codes
142	Wrong Specifications
158	Expired
195	No Shipment Info

Defect ID	Defect	Class ID
33	No Bar Code	3
129	Wrong Codes	3
142	Wrong Specifications	3
158	Expired	3
195	No Shipment Info	3

ID	Defect Class
1	Contamination and Foreign Material
2	Dimensional or Structural Defects
3	Documentation and Labeling Errors
4	Environmental and Handling Damage
5	Functional Defects
6	Manufacturing and Assembly Defects
7	Material Defects

Data Prediction

Defect Classification using Support Vector Machine (SVM) with Categorical Encoding and Standardization

This code classifies defects in manufacturing data using an SVM model. It preprocesses categorical features, standardizes the data, and evaluates the model's accuracy and classification performance.

	Precision	Recall	F1-Score	Support
False	0.90	0.91	0.90	857
True	0.78	0.77	0.78	372
Accuracy			0.87	1229
Macro Avg	0.84	0.84	0.84	1229
Weighted Avg	0.87	0.87	0.87	1229

Data Prediction

```python
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder, StandardScaler
from sklearn.svm import SVC
from sklearn.metrics import classification_report, accuracy_score
file_path = '/content/drive/MyDrive/Power Query.csv'
df = pd.read_csv(file_path)
X = df[['Category', 'State', 'City', 'Vendor', 'Material Type', 'Defect Type', 'Defect']]
y = df[['Defect Qty', 'Downtime min']]
y['Label'] = (y['Defect Qty'] > 0) & (y['Downtime min'] > 0)  # You can adjust this logic based on your needs
y = y['Label']  # Use this binary label for classification
label_encoders = {}
for column in X.columns:
    le = LabelEncoder()
    X[column] = le.fit_transform(X[column])
    label_encoders[column] = le
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)
svc_model = SVC(kernel='linear')  # You can try other kernels like 'rbf' or 'poly'
svc_model.fit(X_train, y_train)
y_pred = svc_model.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print(f"Accuracy: {accuracy:.2f}")
print("Classification Report:")
print(classification_report(y_test, y_pred))
```

Python

06

Executive Summary

The Summary

- **Defects and Downtime**: Significant downtime is caused by materials from certain vendors, with defects categorized as "Impact" causing the highest operational delays, averaging 27 minutes.

- **Vendor Performance**: Vendors such as Dentocity and Quotelane contribute the highest number of defects, while Plustax and Solholdings have the highest downtime, indicating their materials lead to longer operational interruptions.

- **Material Types**: Corrugated materials contribute the most to downtime, while raw materials show minimal disruption.

- **State-Level Analysis**: Iowa and Ohio exhibit the highest downtime, while Illinois leads in defect quantity, indicating regional differences in quality and efficiency.

- Data Cleaning and Preparation:
  - Python: Used for data manipulation tasks, including handling missing values and standardizing data.
  - Pandas: Employed for reading, processing, and exporting datasets.
- Database and SQL:
  - PostgreSQL/SQL: SQL queries were used to clean and transform data, create new tables, and insert data into the database.
- Outlier Detection:
  - Winsorization (with Python and Scipy) was applied to handle outliers in defect and downtime data.
- Data Modeling:
  - ERD (Entity Relationship Diagrams): Used to visualize relationships between data entities like defects, vendors, and plants.
- Data Prediction:
  - Support Vector Machine (SVM): A machine learning model used for defect classification, supported by Scikit-Learn for preprocessing tasks like label encoding and standardization.
- Data Visualization:
  - Dashboards (e.g., Tableau): Likely used for interactive reporting on key performance metrics.

Special shoutout

At the end we want to extend our heartfelt gratitude to Eng. Yasser Abdulrahman for being an exceptional instructor. Your passion for teaching has made a tremendous impact on our learning journey. Your dedication and enthusiasm have truly inspired us, and we are grateful for your shared knowledge and skills.

" Thanks for being such an extraordinary mentor! "

Thanks!

میم سین