November 20, 2024

MSBA Capstone Presentation

Raunak Sharma, McKay Flake, Hunter Harmer

Team Introduction

Raunak Sharma

McKay Flake

Hunter Harmer

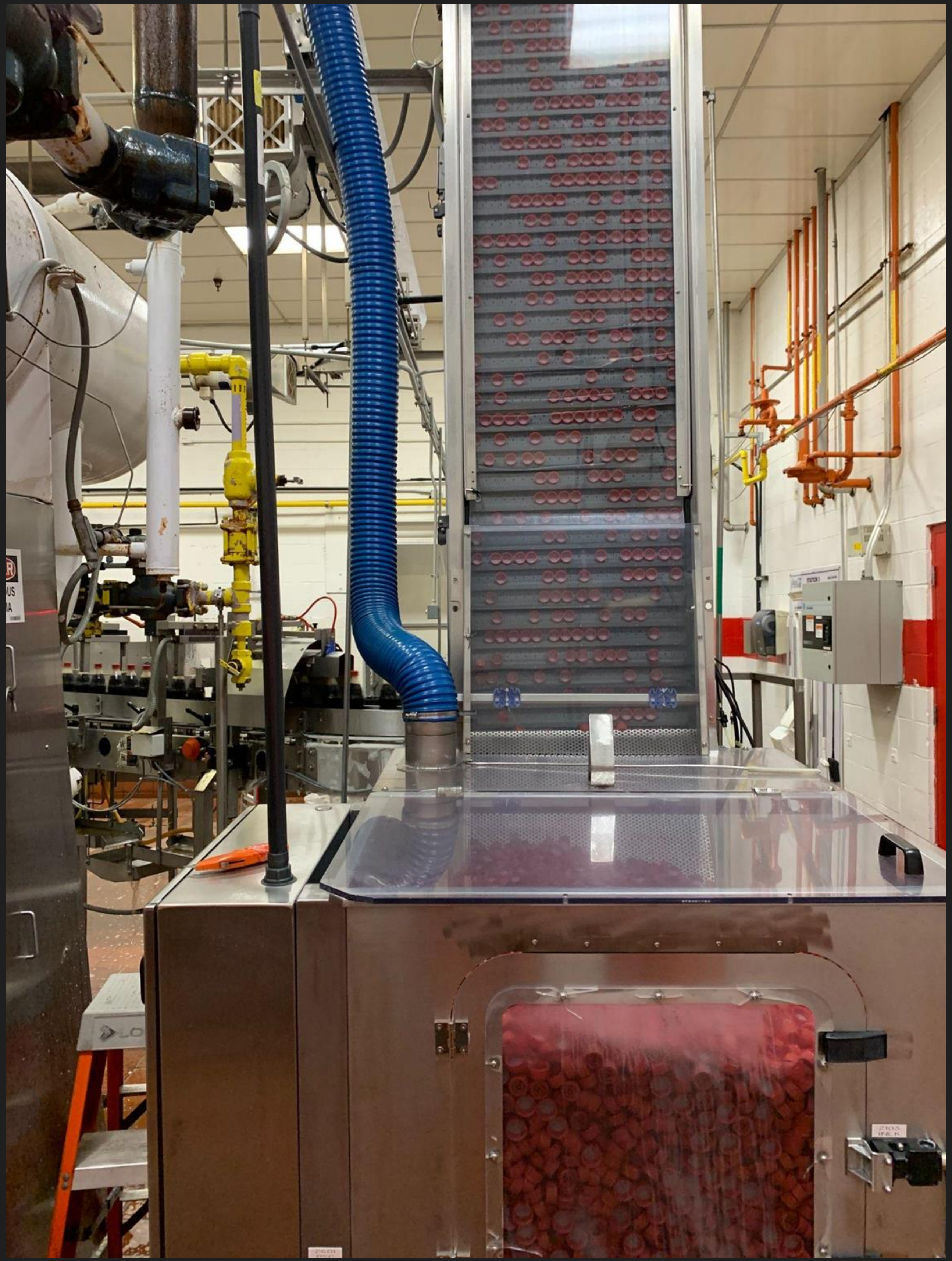
1. Introduction
2. Problem Description
3. EDA
4. Modeling Process
5. Break-even Analysis Tool

Problem Description

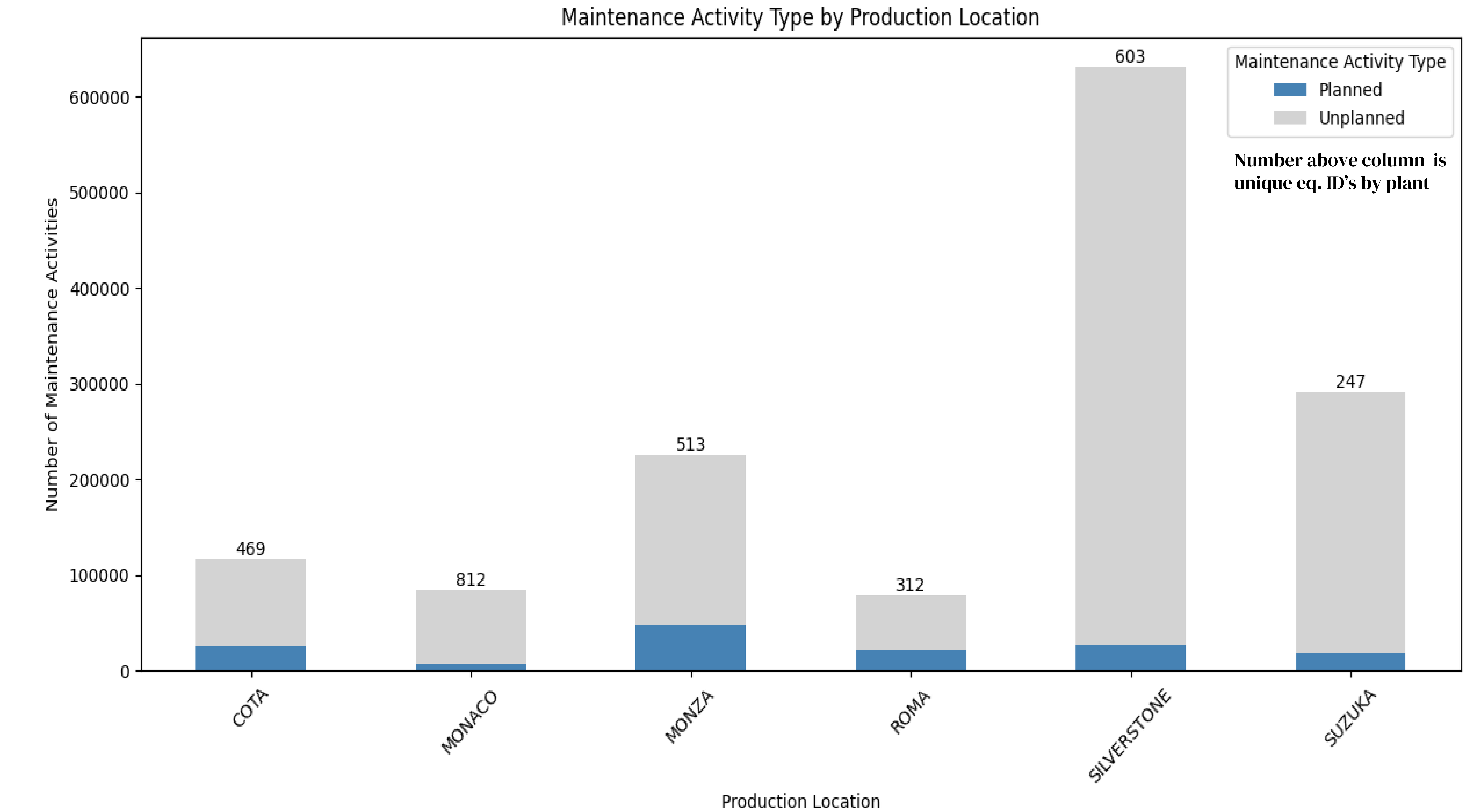
Coca-Cola
SWIRE COCA-COLA

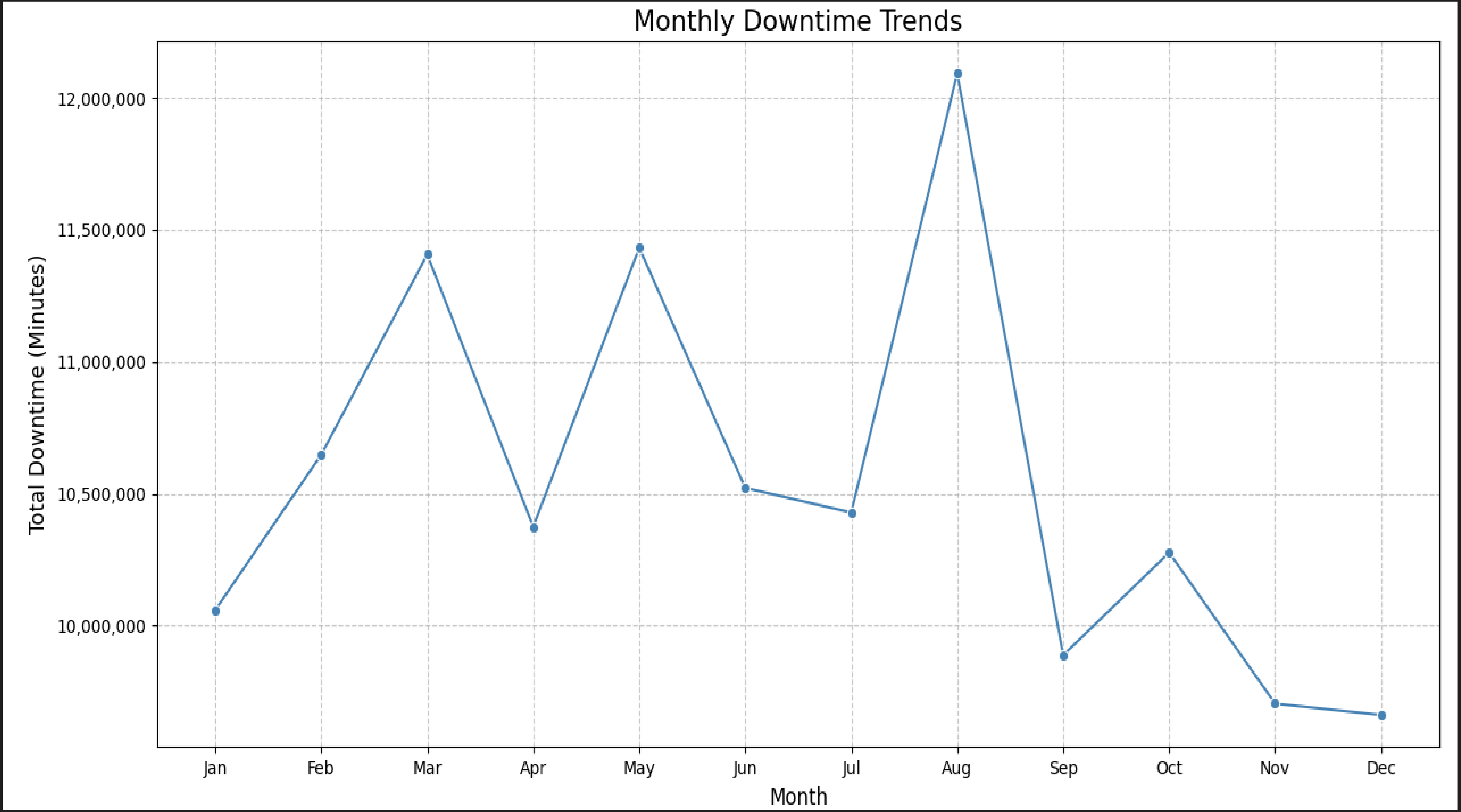
Fill the cans and bottles!

The ultimate goal is to predict machine breakdowns, identify common part failures, optimize maintenance, reduce downtime, and boost production efficiency

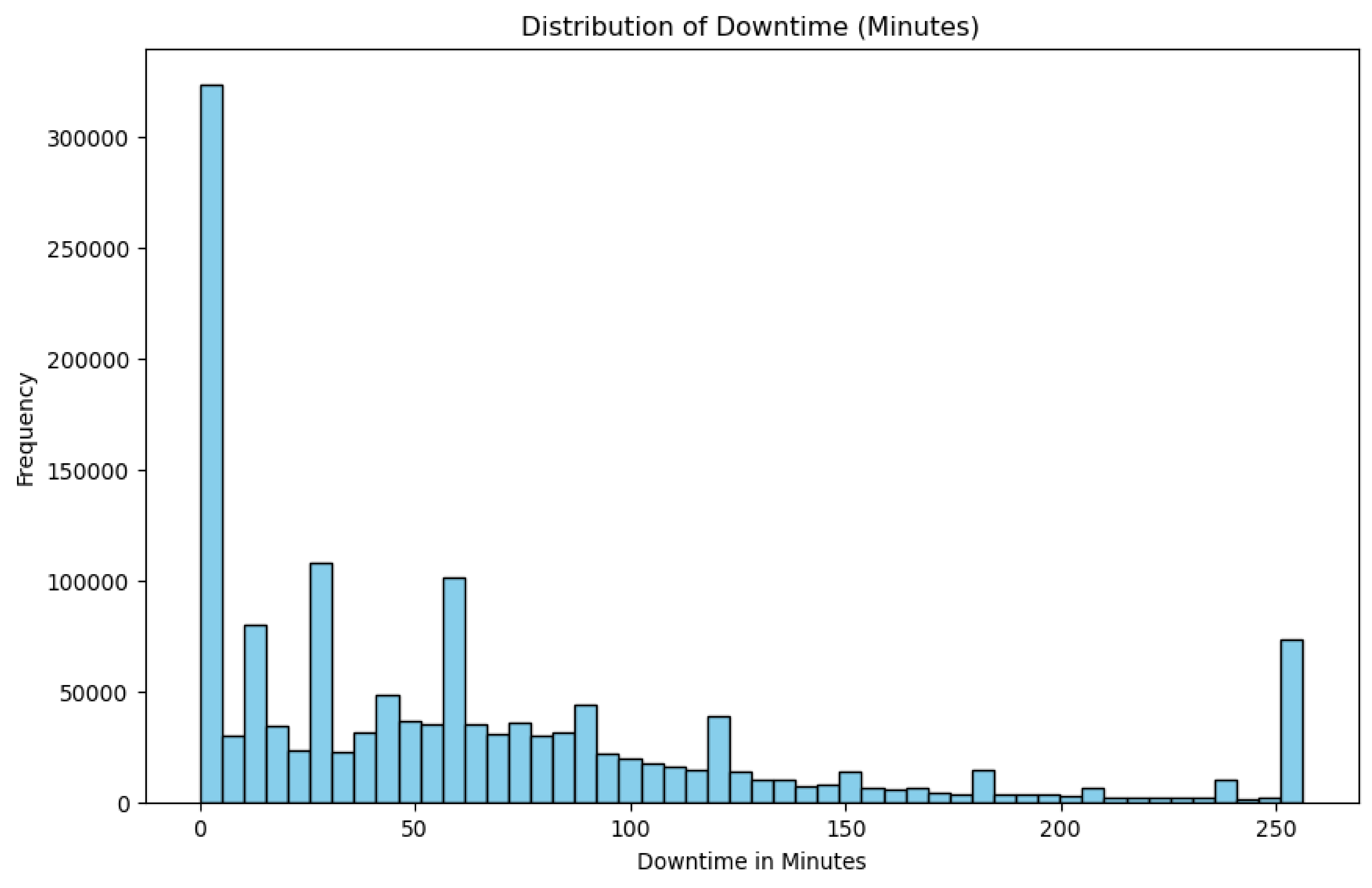


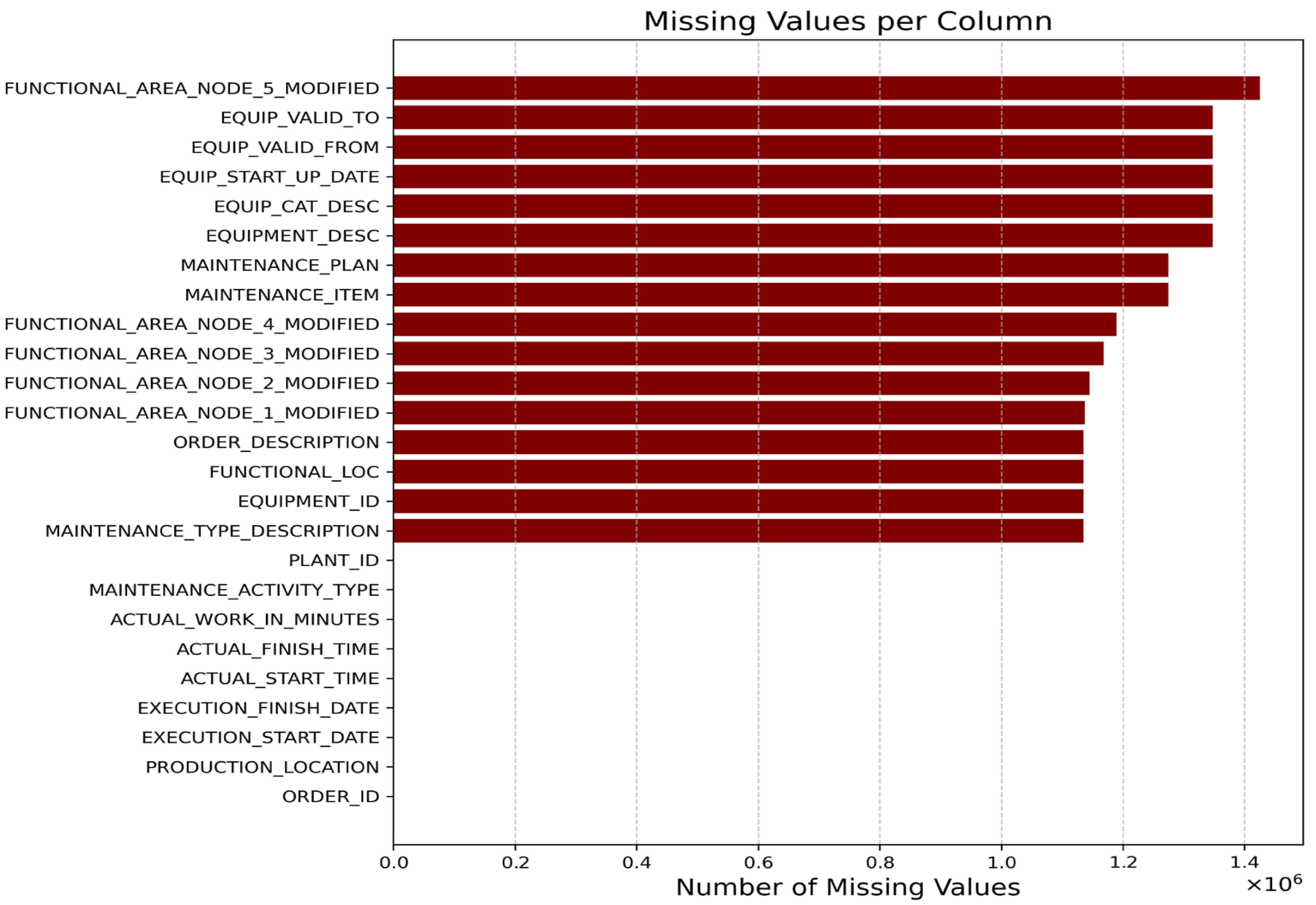
Exploratory Data Analysis


Maintenance Activity Type by Production Location
Maintenance Activity Type
Planned
Unplanned
Number above column is unique eq. ID's by plant
Number of Maintenance Activities
600000
500000
400000
300000
200000
100000
0
469
812
513
312
603
247
COTA
MONACO
MONZA
ROMA
SILVERSTONE
SUZUKA
Production Location

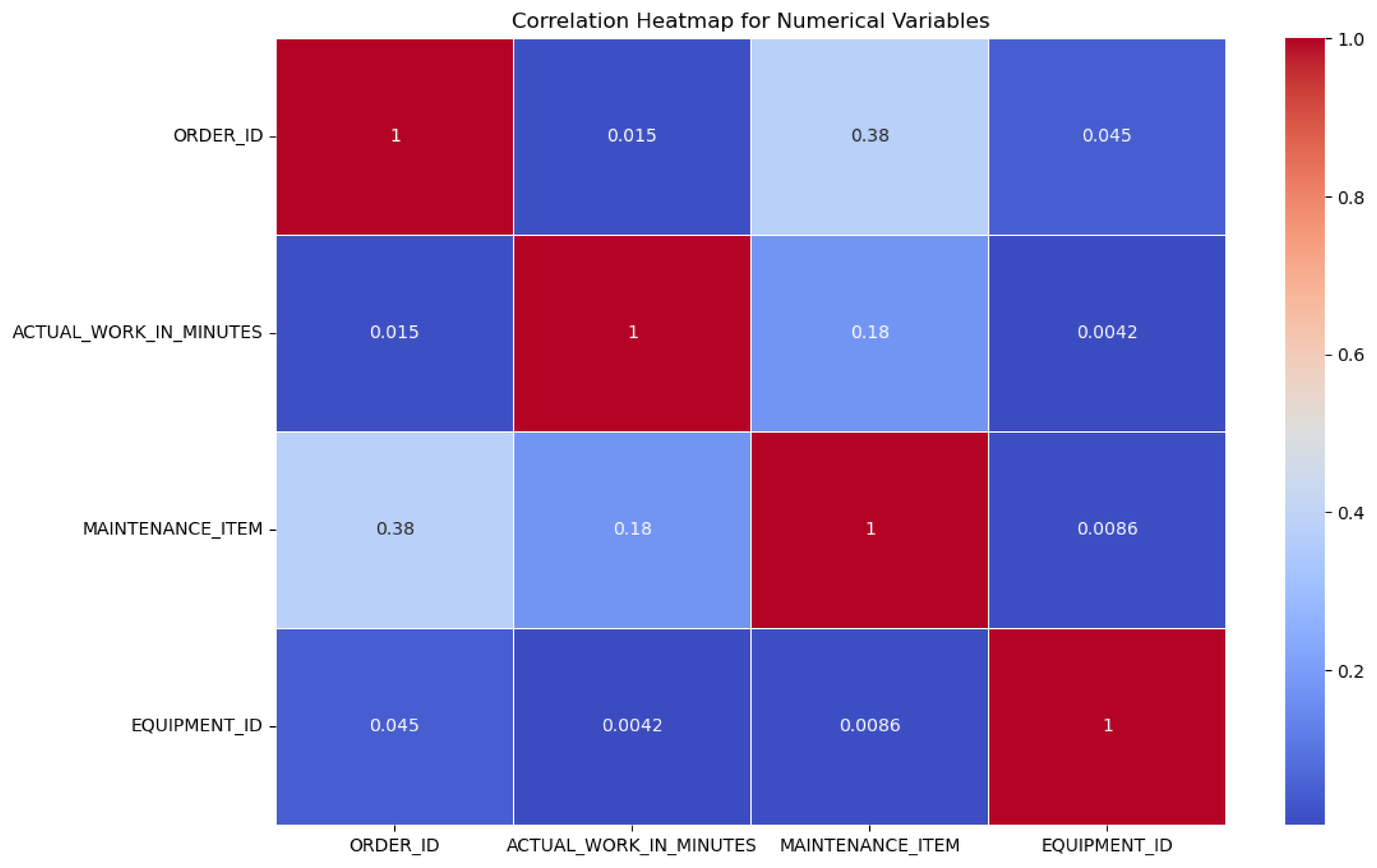


Monthly Downtime Trends
Total Downtime (Minutes)
12,000,000
11,500,000
11,000,000
10,500,000
10,000,000
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month


Distribution of Downtime (Minutes)
Frequency
300000
250000
200000
150000
100000
50000
0
0
50
100
150
200
250
Downtime in Minutes


Missing Values per Column
FUNCTIONAL_AREA_NODE_5_MODIFIED
EQUIP_VALID_TO
EQUIP_VALID_FROM
EQUIP_START_UP_DATE
EQUIP_CAT_DESC
EQUIPMENT_DESC
MAINTENANCE_PLAN
MAINTENANCE_ITEM
FUNCTIONAL_AREA_NODE_4_MODIFIED
FUNCTIONAL_AREA_NODE_3_MODIFIED
FUNCTIONAL_AREA_NODE_2_MODIFIED
FUNCTIONAL_AREA_NODE_1_MODIFIED
ORDER_DESCRIPTION
FUNCTIONAL_LOC
EQUIPMENT_ID
MAINTENANCE_TYPE_DESCRIPTION
PLANT_ID
MAINTENANCE_ACTIVITY_TYPE
ACTUAL_WORK_IN_MINUTES
ACTUAL_FINISH_TIME
ACTUAL_START_TIME
EXECUTION_FINISH_DATE
EXECUTION_START_DATE
PRODUCTION_LOCATION
ORDER_ID
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
×10^6
Number of Missing Values

Correlation Matrix: Identify relationships between numerical features, which can guide feature selection and engineering.



Modeling Roadblocks

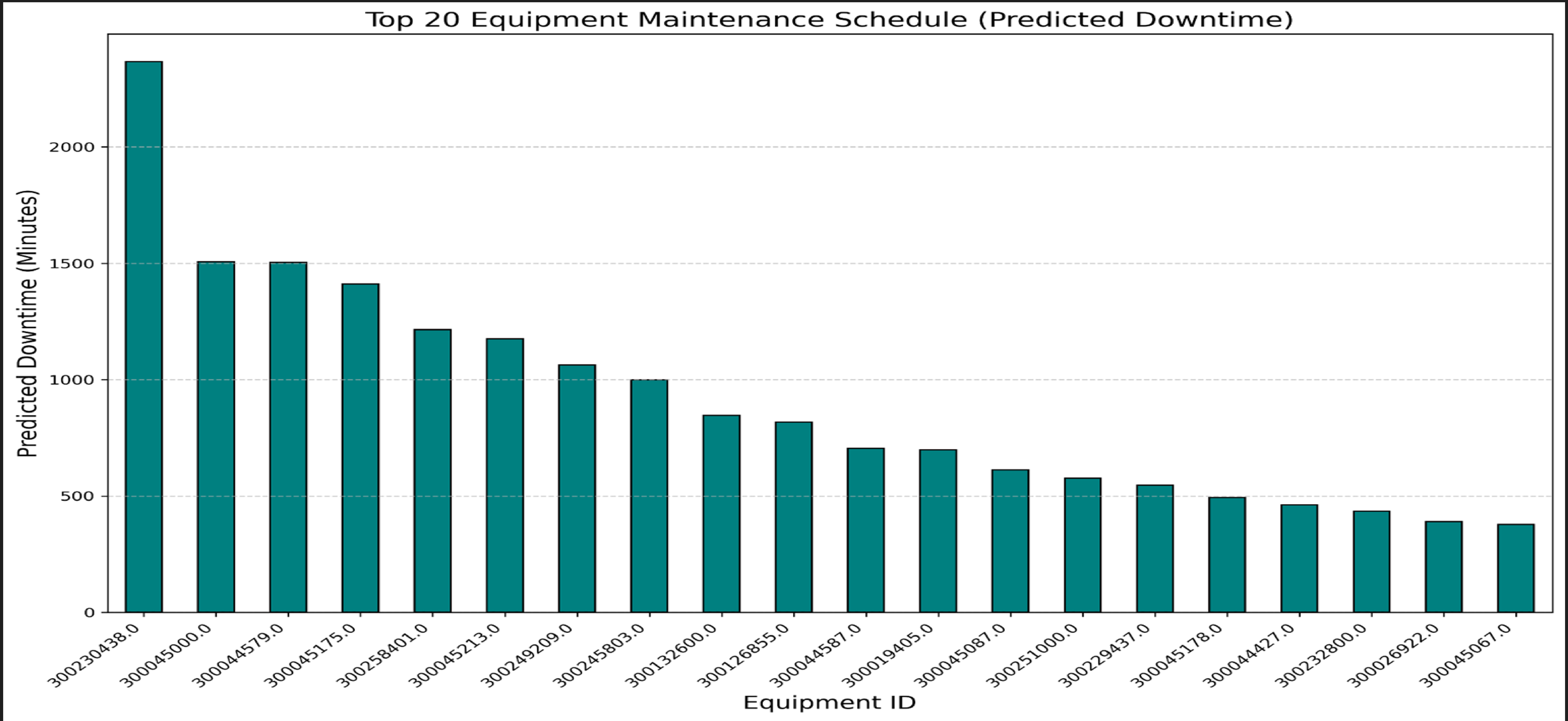
- Given the nature of the data, our initial efforts to produce a true predictive maintenance model fell short
- The most promising method was an XGBoost model, which could be implemented with the introduction of financial components to the dataset
- RMSE: indicates the average difference between the predicted and actual downtime in minutes. It reflects the inherent variability and complexity of downtime data
- R^2 : shows that the model explains 49.2% of the variance in the downtime, demonstrating moderate predictive power

Model Performance Summary by Sample Size

	Model	MSE (1,000 Samples)	R^2 (1,000 Samples)	MSE (10,000 Samples)	R^2 (10,000 Samples)	MSE (Full Dataset)	R^2 (Full Dataset)
0	Linear Regression	4654.31	0.02	4640.27	0.02	670.52	0.07
1	Random Forest	4266.11	0.10	3767.45	0.21	693.57	0.00
2	XGBoost	3471.42	0.27	3471.42	0.27	494.42	0.49
3	Gradient Boosting	3915.21	0.18	3541.29	0.25	-	-

Prioritizing maintenance to minimize operational losses

Analysis identified high-risk equipment like ID 300230438 with over 2000 predicted downtime minutes. Prioritizing preventive maintenance for these assets can significantly reduce downtime and associated costs.



Break-even Analysis Tool

Breakeven Analysis

- We created a calculator to determine the optimal amount of planned maintenance needed to minimize costs
- This calculator will work best when used in a granular way as it works off the ratio of planned to unplanned maintenance

Demo

- For this demo we will use the FUNCTIONAL_AREA_NODE_1_MODIFIED "COTA PRODUCTION" as it is likely to give us the most applicable results due to the large sample size

- Note that the ratio is 1.6

FUNCTIONAL_AREA_NODE_1_MODIFIED	Planned	Unplanned	Ratio
COTA PRODUCTION	26338.0	43849.0	1.664857
SUZUKA PRODUCTION	15512.0	21201.0	1.366748
SILVERSTONE PRODUCTION	27803.0	34822.0	1.252455
PRODUCTION	29074.0	31238.0	1.074431
MONZA PRODUCTION	47849.0	9609.0	0.200819
COOLER SERVICE	2698.0	1.0	0.000371
FLEET	NaN	123.0	NaN

Demo

- Since we do not have access to financial data on a per-hour basis, we will use the following assumptions:
  - **$45.5 per hour for planned maintenance**
  - **$555 per hour for unplanned maintenance**, which includes losses from stopping production

- We need to enter the **ratio of planned to unplanned work**, which in our case is **1.6**

```
Enter the monetary cost per hour of planned maintenance: $45.5
Enter the monetary cost per hour of unplanned maintenance (including compensation for lost production): $555
Enter the ratio of planned to unplanned maintenance events (e.g. 170.33 for 3 planned and 511 unplanned events): 1.6
The break-even point is: 7.62 hours of planned maintenance per hour of unplanned maintenance.
The number of hours of planned maintenance needed is: 12.20 hours.
The cost of planned maintenance is: $555.00.
The value of unplanned maintenance avoided is: $888.00.
```

SWIRE COCA-COLA

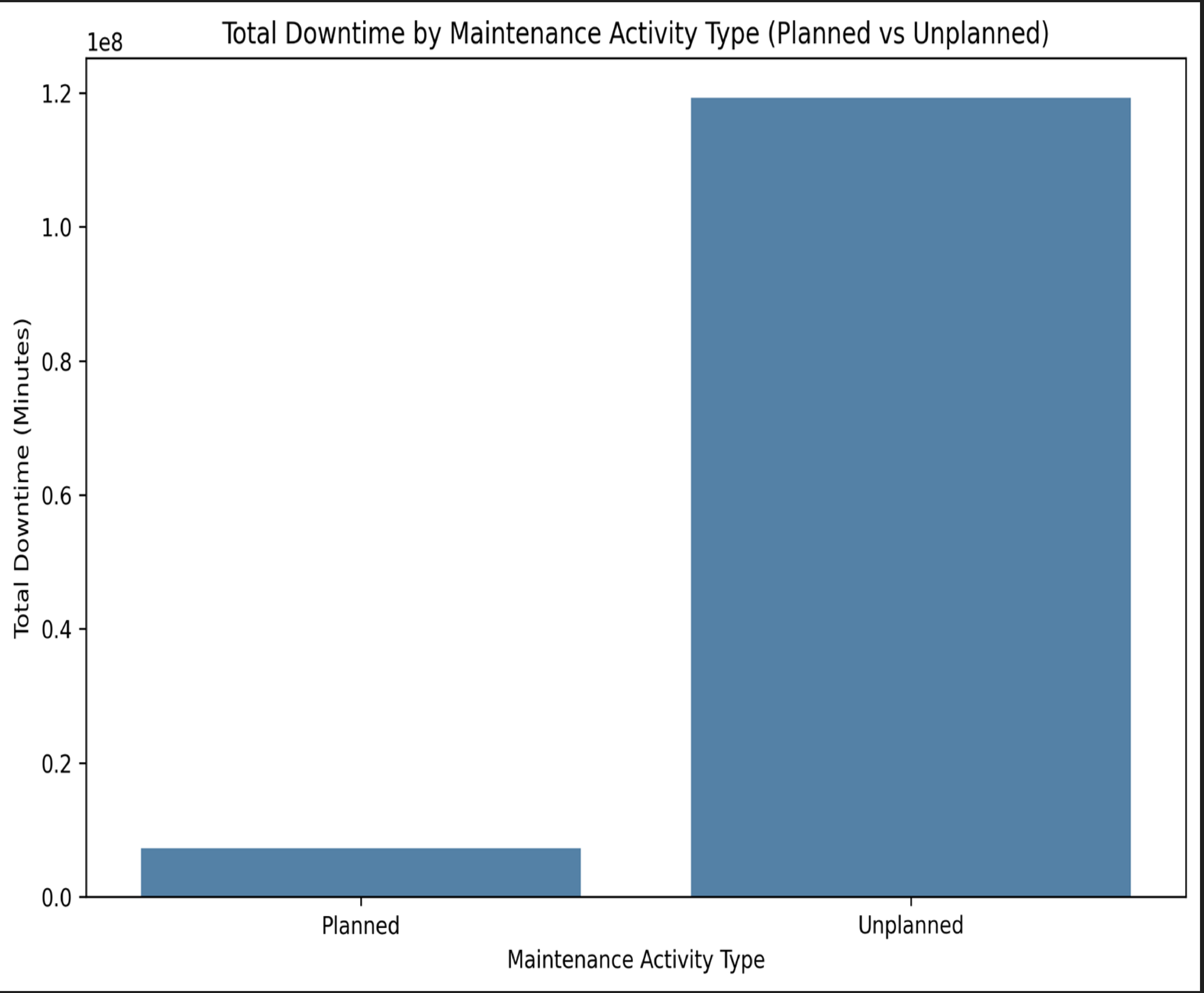
Demo

- We have 12.2 hours to spend on planned maintenance for the COTA PRODUCTION sector, but what are the best ways to spend those 12.2 hours?
- Counts show frequency
- September though the end of the year are the key times to do planned maintenance

	EQUIPMENT_ID	Count	Avg_ACTUAL_WORK_IN_MINUTES
375	300198402.0	3	1655.20
140	300026922.0	40	517.32
264	300027066.0	1	360.00
276	300060046.0	18	315.10
120	300026893.0	3	291.60
..	...	...	...
316	300120013.0	1	15.00
218	300027018.0	1	15.00
58	300026818.0	1	12.00
359	300192002.0	1	6.00
2	300025758.0	1	3.00

[409 rows x 3 columns]

Thank You



Unplanned maintenance cost : 120 M minutes

Planned : 10 M minutes

Preventive maintenance can help us reduce overall downtime