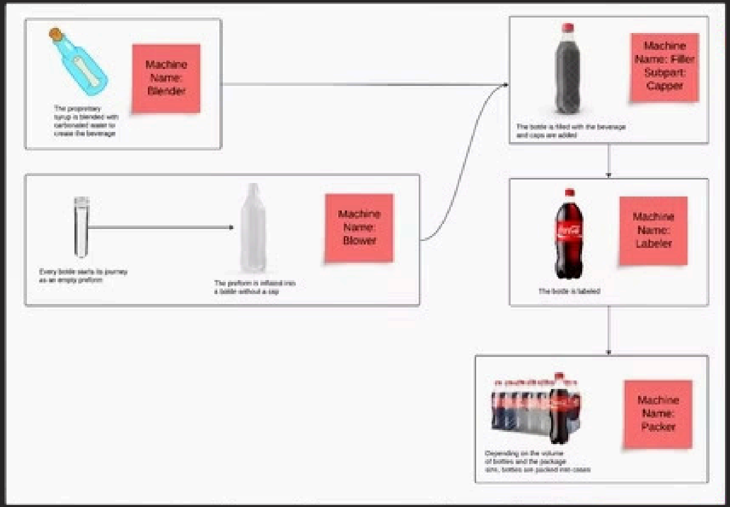
Predictive Maintenance Analysis for Swire Coca-Cola

Group 2

Richard Lim, Ketki Kulkarni, Anusha Vivekanand, Vedika Garg

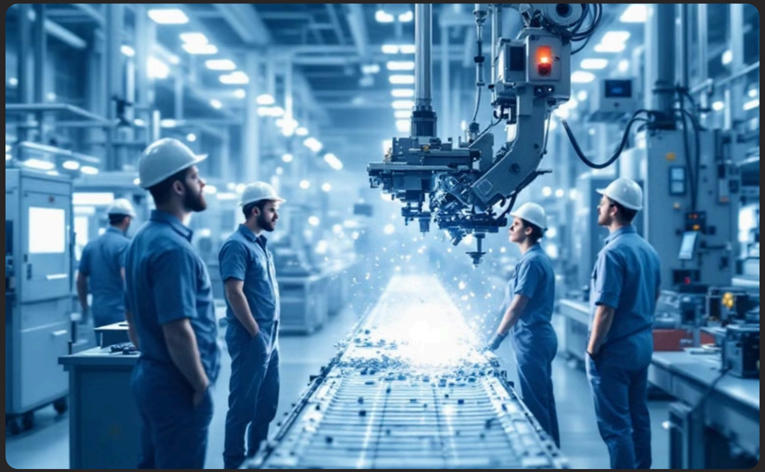
Background



Our capstone project focuses on identify and analyze breakdown patterns in machine downtime, and find major factors influencing breakdowns using IWC dataset across the 6 plants in worldwide.

The goal is to understand the breakdown frequency, timing, and locations, as well as the root causes and possible resolutions.

Objective



Analytical Goal

Develop a predictive model to forecast machine downtime frequency and duration.

Business Goal

Provide actionable items for proactive maintenance, cost reduction, and efficiency improvement.

Assumptions

Predictable Pattern

Machine failure patterns are consistent and predictable across equipment types and plant location.

Preventive repair

Equipment parts causing downtime can be pre-stocked for timely repairs.

Limitations

Unrecorded IWC Tracking

Only significant breakdowns are recorded, potentially missing patterns from less severe issues.

Large Missing Data

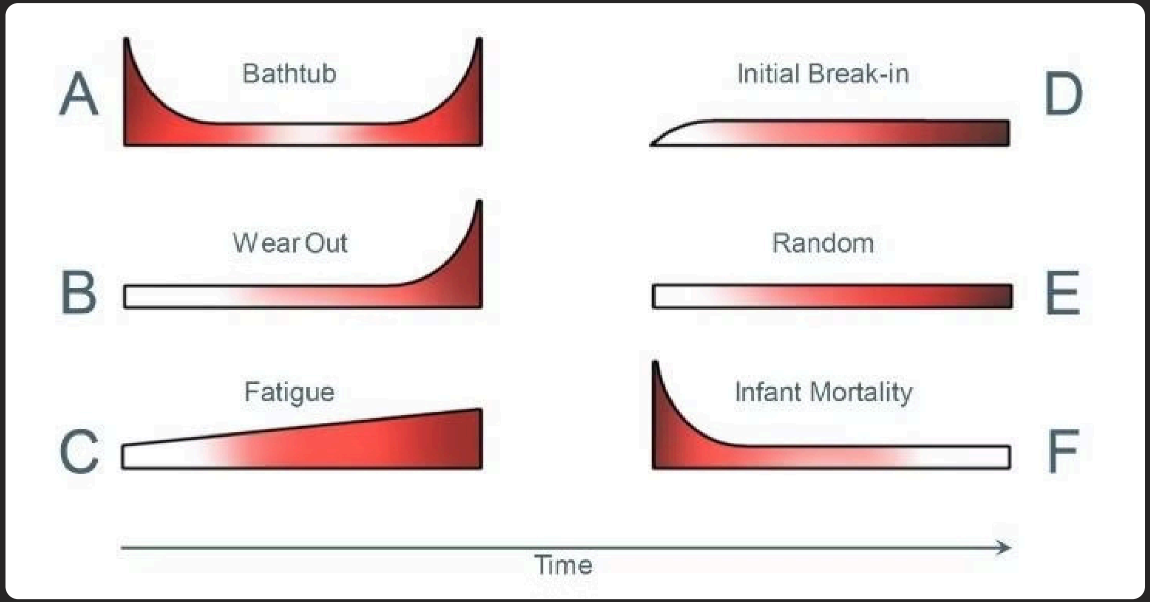
Some work orders have over 70% missing data, impacting analysis accuracy.

Unidentified Machine Age

Machine age and usage patterns may not be suitable for model development, affecting prediction accuracy.

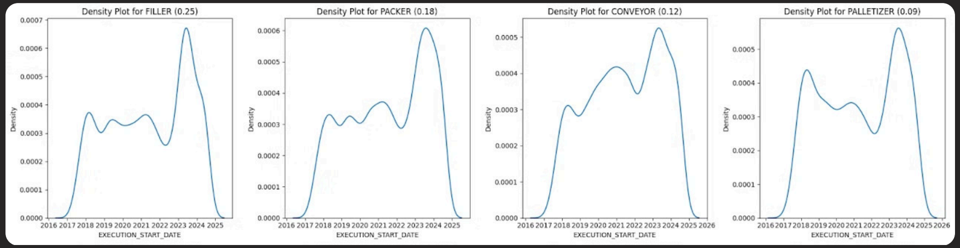
Dataset Overview

The dataset contains 1,427,264 rows and 27 columns. Each row represents a work order recorded by IWC.



EDA

Top 4 Unplanned Equipments Breakdown [FUNCTIONAL_AREA_NODE_4_MODIFIED]



It Consists of 64% of total unplanned breakdowns **[FILLER] [PACKER] [CONVEYOR][PALLETIZER]**

Takeaways

There are certain breakdown patterns per plant location in that we are collecting **unplanned equipment breakdown data only.** ([FILLER] [PACKER] [CONVEYOR][PALLETIZER]). For example, SILVERSTONE shows the consistent pattern as a (A) bathhub, which requires improving quality control to reduce early-life failures and designing for durability to extend the useful life.

Analysis Metrics

50.0

Unplanned

Minutes (Median)

32.0

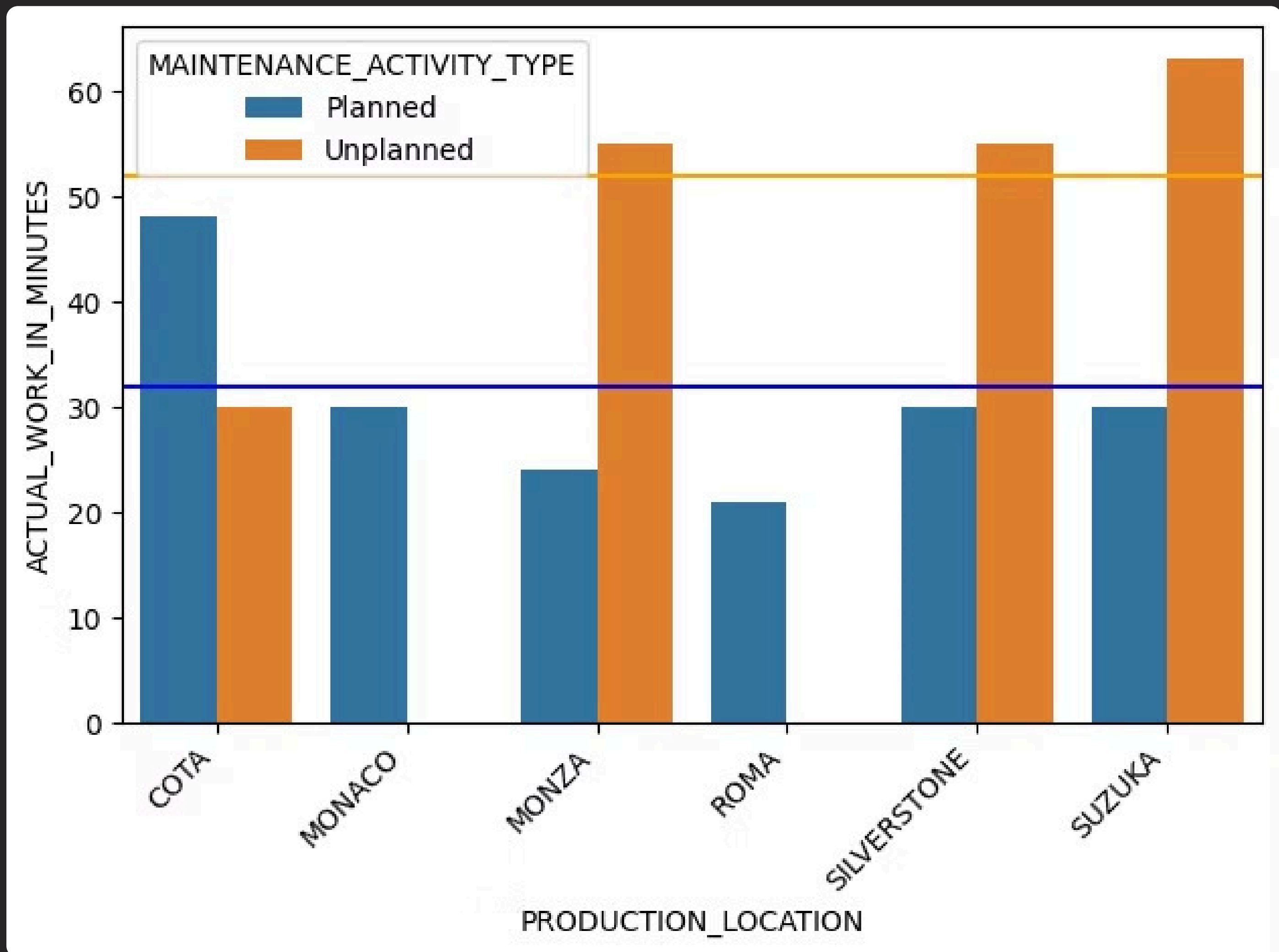
Planned

Minutes (Median)

36%

Reduction

Potential for improvement



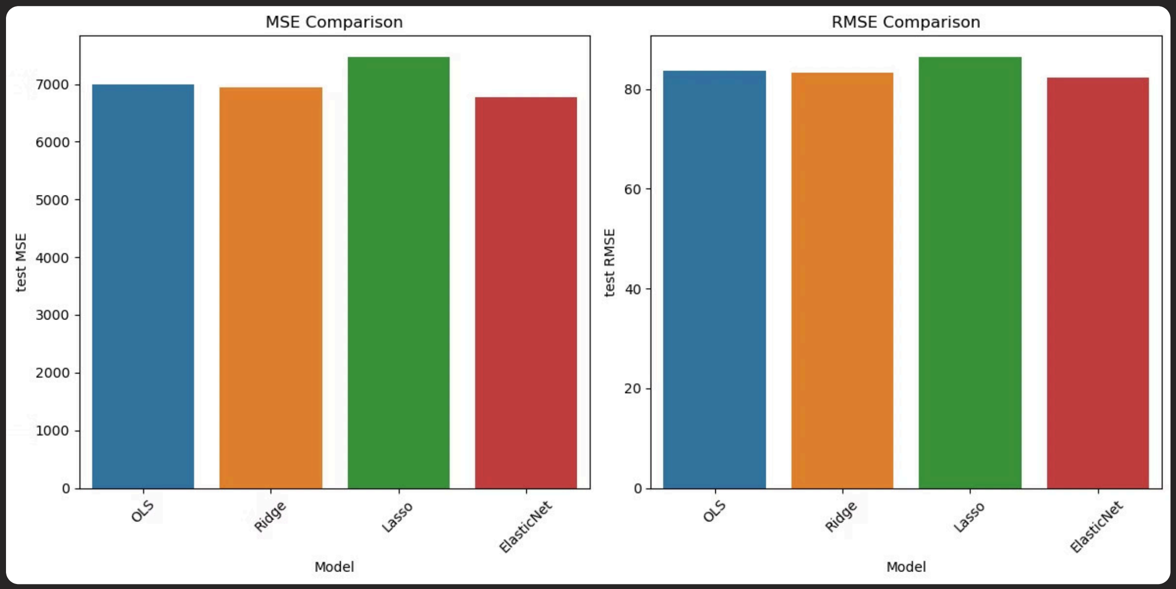
Modeling & Evaluation

Regression Model

Data preprocessing

- Group by given data with [df['PLANT_ID'] + '-' + df['LINE_ID'] + '-' + df['SUBPROCESS_ID'] and predict when is the next breakdown duration based on historical data.

The model comparison provides valuable insights into performance across different regression approaches.



Model	RMSE	MSE	R-squared
OLS	83.6577	6998.6032	0.6086
Ridge	83.2756	6934.8299	0.6122
Lasso	86.3950	7464.0931	0.5826
Elastic-Net	82.2674	6767.9204	0.6215

Takeaways

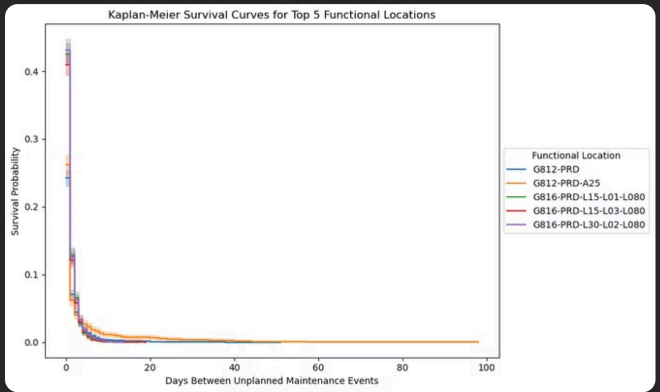
The ElasticNet model stands out, achieving the lowest test RMSE (82.27) and MSE (6767.92), suggesting it strikes the best balance between bias and variance, making it highly suitable for predicting maintenance outcomes.

Modeling & Evaluation

Survival Analysis

The **equipment threshold production capacity** defines how much equipment can be maintained until the machine is working at full capacity before needing maintenance.

- Analyze time between equipment failures across functional locations [FUNCTIONAL_LOC]



Survival Summary Statistics for Top 10 Functional Locations (Sorted by Mean Time Between Replacements):

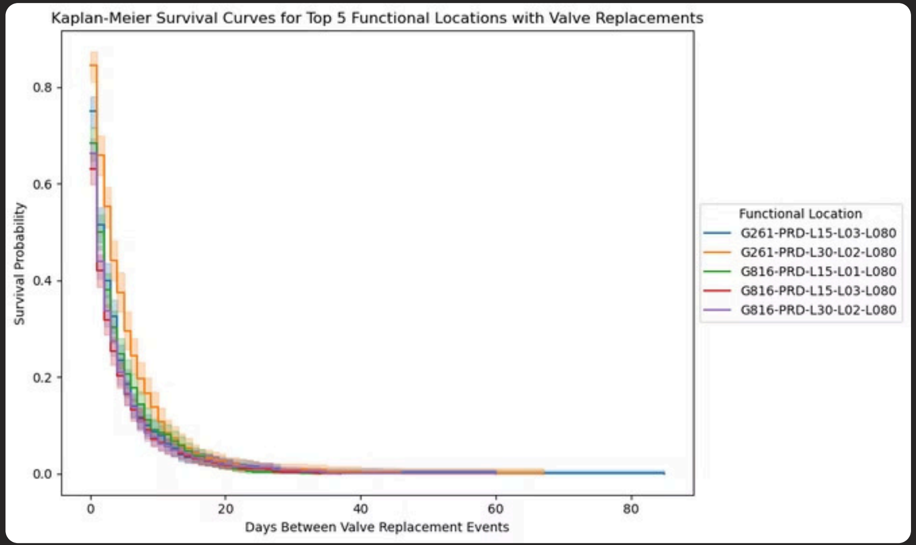
FUNCTIONAL_LOC	count	mean	std	min	25%	50%	75%	max
G812-PRD	5406.0	0.473918	1.617539	0.0	0.0	0.0	0.0	51.0
G816-PRD-L15-L03-L080	3800.0	0.656579	1.149846	0.0	0.0	0.0	1.0	19.0
G816-PRD-L15-L01-L080	3639.0	0.685078	1.167064	0.0	0.0	0.0	1.0	17.0
G812-PRD-A25	3641.0	0.702005	3.541655	0.0	0.0	0.0	1.0	98.0
G816-PRD-L30-L02-L080	3554.0	0.702026	1.204813	0.0	0.0	0.0	1.0	14.0
G816-PRD-L30-L02-L100	2474.0	1.006063	1.776300	0.0	0.0	0.0	1.0	18.0
G816-PRD-L15-L01-L120	2353.0	1.061198	1.899605	0.0	0.0	0.0	1.0	19.0
G816-PRD-L15-L03-L120	2190.0	1.131050	2.102381	0.0	0.0	0.0	1.0	34.0
G816-PRD-L30-L02-L030	2028.0	1.228797	1.853521	0.0	0.0	1.0	2.0	27.0
G221-PRD-L30-L02-L080	2144.0	1.248134	2.200411	0.0	0.0	1.0	1.0	27.0

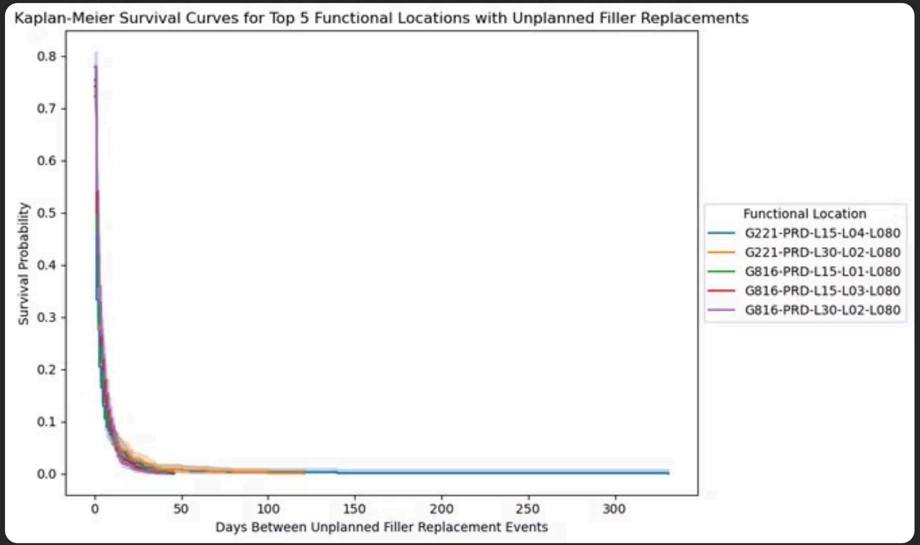
Takeaways

An equipment in Plant Location : Roma (G812) would fail on average at 0.47 days, Cota (G816) in 0.65 days and Suzuka (G221) at 1.24 days.

Survival Analysis

Survival Summary Statistics for Top 5 Functional Locations





Survival Summary Statistics for Top 5 Functional Locations (Sorted by Mean Time Between Valve Replacements):

FUNCTIONAL_LOC	count	mean	std	min	25%	50%	75%	max
G816-PRD-L15-L03-L080	867.0	2.877739	4.764796	0.0	0.0	1.0	4.0	37.0
G816-PRD-L30-L02-L080	812.0	3.061576	5.263573	0.0	0.0	1.0	4.0	60.0
G816-PRD-L15-L01-L080	735.0	3.390476	4.861142	0.0	0.0	2.0	4.0	34.0
G261-PRD-L15-L03-L080	727.0	3.519945	6.027652	0.0	0.5	2.0	4.0	85.0
G261-PRD-L30-L02-L080	541.0	4.713494	6.066391	0.0	1.0	3.0	6.0	67.0

Survival Summary Statistics for Top 10 Functional Locations (Sorted by Mean Time Between Unplanned Replacements):

FUNCTIONAL_LOC	count	mean	std	min	25%	50%	75%	max
G816-PRD-L15-L01-L080	766.0	3.253264	5.351057	0.0	0.0	1.0	4.0	46.0
G816-PRD-L15-L03-L080	715.0	3.488112	5.158639	0.0	0.0	1.0	4.0	45.0
G816-PRD-L30-L02-L080	682.0	3.645161	4.977206	0.0	1.0	2.0	5.0	42.0
G221-PRD-L15-L04-L080	722.0	3.681440	14.785603	0.0	1.0	1.0	3.0	331.0
G221-PRD-L30-L02-L080	671.0	3.934426	9.139313	0.0	0.0	1.0	4.0	121.0
G291-PRD-L30-L02-L080	495.0	5.284848	6.963008	0.0	1.0	4.0	7.0	69.0
G291-PRD-L30-L01-L080	408.0	6.659314	8.128218	0.0	2.0	5.0	8.0	92.0
G812-PRD-L15-L01-L080	277.0	9.227437	11.959877	0.0	1.0	4.0	12.0	78.0
G291-PRD-L30-L03-L080	259.0	9.691120	33.550107	0.0	1.0	2.0	6.0	422.0
G261-PRD-L30-L01-L080	258.0	9.755814	10.988076	0.0	1.0	6.0	14.0	56.0

Analytical Interpretation

1. **Survival Analysis**

   Survival probabilities drop sharply within the first few days after maintenance in key locations like Suzuka, Sliverstone and Monza.

2. **Component-Level Insights**

   Our analysis shows that fillers, packers, conveyers and palletizers were the most parts identified with frequent breakdowns.

3. **Regression Models**

   Regression models provides the stability of modeling for predicting the last breakdown duration with 0.6 R-squared correlation across multiple models such as OLS, ridge, lasso and elastic-net.

4. **Risk Prioritization**

   The analysis helps rank functional locations by their maintenance needs, allowing for precise allocation of resources.

Business Interpretation

1 Operational Efficiency

Proactive maintenance reduces downtime, ensuring smooth production cycles and safeguarding revenue.

2 Cost Reduction

Predictive maintenance minimizes repairs and replacements, cut the labour cost leading to cost savings in maintenance budgets.

3 Strategic Resource Allocation

Prioritizing high-risk locations optimizes resource allocation, ensuring critical parts are readily available.

4 Lean Inventory Management

Insights into high-demand parts enable better inventory planning, reducing overstocking and minimizing repair delays.

Call to Action

1 Pilot Program

We recommend piloting this strategy at high-risk locations to quantify its impact.

2 Scalability Measurement

Once validated, the strategy can be scaled across Swire's operations.

3 Operational Roadmap

This is not a one-time solution, but a roadmap to establish operational excellence.

Questions & Answers