

Use Case for MSBA Capstone

University of Utah – Fall 2024

Contents

Use Case: Predictive Maintenance 2-10

Use Case: Predictive Maintenance

Every year, Swire Coca-Cola produces close to 192 million cases of beverages to sell in its markets in 13 states from 6 production plants. One of the shortcomings of mechanical limitations across its plants is that it is currently able to produce 94.4% of the total number of cases that need to be delivered basis orders. Some reasons for the same could be machines getting taken down for unforeseen mechanical issues, maintenance, wear and tear and so on.

The production and supply chain team at Swire Coca-Cola wishes to bridge this gap by understanding why and when these downtimes occur, so that planning can be set in place in advance to ensure the downtimes don't cost us too much output.

What currently happens is if a machine is broken or a part wore out prematurely, the line is shut down. If the repairs are extensive, there is a work order issued on a system called Internal Warehouse Controller, IWC for short (Fictitious name used to respect the privacy and integrity of the actual system in place). The work order is reviewed, and the necessary parts are ordered, following which the parts get delivered, the parts are replaced, and the line is back up and running. The problem with this approach is that quite a few of the downtimes are unplanned, i.e. nobody sees them coming and in the process of getting the machines up and running again, the time taken costs the company almost \$60m annually in losses.

Swire Coca-Cola now needs to understand how to study and analyze the downtimes recorded in IWC. In doing so, Swire Coca-Cola also needs to understand how gravely their output is affected (in terms of time and productivity per line, plant or package type) owing to these downtimes.

The ask is to get the downtime records across locations on the same page, saying 'X' downtime recorded in IWC needed 'Y' fix, and any further inputs on understanding the severity of the downtime, what factors recorded from the descriptions can be attributed to that breakdown, and the opportunity cost.

Definition of a predictive maintenance pipeline: A predictive maintenance pipeline would let the users know that 'X' downtimes were severe enough to be recorded in the IWC system and the time it was down cost us how much in productivity. This will let the team understand when to flag upcoming downtime, stock up with the repair parts well in advance and the downtime will be lesser than it is now. Analysis perspectives can include, but not restricted to:

- Understanding what the conditions of the machine were at the time of the downtime: For example, machine X always breaks down at this time and this part is most commonly breaking, etc.

- Quantifying how much time the machine could run for before the wear and tear or breakdown sets in, requiring specific repairs: This could be defined as the threshold production capacity, like how cars come with a threshold of needing servicing and oil changes every X miles, this can be a prescriptive insight on how much output can be produced by this machine working at full capacity before needing maintenance. (Bonus food for thought: If the machine runs on lower capacity, like in easier shifts, does the downtime get delayed?)
- Guardrails for production capacity: If a plant had let's say 4 such machines, and a demand of X cases, would the demand be met by lowering the threshold cases production in every machine to maximize machine life and at the same time meet the demand 100%. Basically, instead of all machines running 24x7, if we were to run them in smaller shifts, would they break down less often?
- Prescriptive inventory: If a certain machine requires the same part to be replaced often, would the company benefit from having the spare part in stock well in advance? If so, how many of these parts must be stocked per year per machine?

Problem Statement: Your mission, should you choose to accept it, is to develop a model that is able to describe the current landscape of machine maintenance for the plant, when a breakdown could come up, if it is possible to know what parts are most commonly required to fix these machines and how much time would we save if the downtimes could be predicted well in advance. The primary objective is, help us understand what causes a machine to break and how effectively we can foresee that. You can leverage any attribute available to you from and set ground expectations or assumptions from IWC.

Step-by-Step Approach:

1. Data Exploration and Preprocessing

- o IWC data: Note that only severe breakdowns make it to having a work order issued against them. Therefore, not all breakdowns recorded on the factory floor will have IWC work orders issued against them. An analysis of location, description and package line might be a great place to get started.

Refer to the Data Dictionary for detailed information on the columns in each dataset.

2. Feature Engineering

- o **What:** What broke down. The description is key here
- o **When:** You have timestamps of breakdowns

- o **Why:** Can be inferred from conditions of the breakdown or descriptions in those cases
- o **Where:** The locations of every machine are available, in different grains for you to either analyze by machine or by specific part residing in the machine
- o **How:** The nature of the fix is available in the IWC dataset and can be used to infer the method of correction used to get the machine back up and running
- o **How much:** The machines' downtime (actual work in minutes). Will be useful to infer the opportunity costs and the financial sizing of the solution.
- o **Why not:** Highly encourage you to provide alternative explanations/insights that may help the company better understand its downtimes and the benefits of predicting downtime.

3. **Model Selection**

(Warning: Let this not serve as a restriction to your creativity in any way whatsoever. These are mere suggestions/'Good-to-have' solutions. You are free to let your mind run wild and come up with more solutions that you see helpful)

- o Utilize methods to map the IWC work orders to their corresponding plants/machines as aggregations.
- o Indicate the threshold (Visually or numerically) or break points and how you identified the limit to which a machine can be pushed before it gives up

4. **Model Training and Validation**

- o Train the model on the prepared dataset, ensuring it accurately predicts the time that a machine would break. You are free to use simulations or any similar approaches to prove a point
- o Validate the model to assess its performance and refine it as needed, ensuring the match between the two data sources is as accurate as can get, focusing especially on what each downtime entails, in terms of fixes.

5. **Predictive Maintenance**

- o Develop an end-to-end comprehensive pipeline that narrows down breaks and their reasons with other factors, so the production team knows what factors and/or thresholds need to be flagged and predict when the machine is going to break.

6. **Dashboard Development or Presentation**

- o You've thought of something totally awesome; it is important that the insights and thoughts are presented in a concise and crisp manner for the business stakeholders, who understand machines very well, but are not particularly fond of technical details like model accuracy. They just want insights. What do I look out for, when will the machine break and how much money or time will it make/save me.

7. **Documentation**

- o Provide comprehensive documentation detailing the model development process, feature selection rationale, and the strategy behind incorporating specific predictors for the downtime and how much of an impact it has on the actual prediction (You can go nuts with the numbers here, this is more technical than the presentation).

This structured approach will guide you in building a sophisticated system that enhances Swire Coca-Cola's predictive maintenance pipelines, ensuring we save money by not succumbing to unforeseen downtimes and are able to scale up our optimal production levels to meet the demand and make more money. A deeper understanding of the data and the user story is available later in this dossier for analysis.

Metadata + Context:

To understand the data, the problem statements and potential use cases better, let's imagine there is a Plant Supervisor named George, and view the problem as a user story.

Let's assume there is a plant of Swire Coca-Cola in Monza, and George is the plant supervisor at the Tempe plant. There are multiple lines in this plant producing cans, bottles and syrup.

In the bottle line, the flow of machines is as follows (Public domain information- The Coca-Cola manufacturing process, visually depicted [here](#), time stamps 2:45 to 5:41)

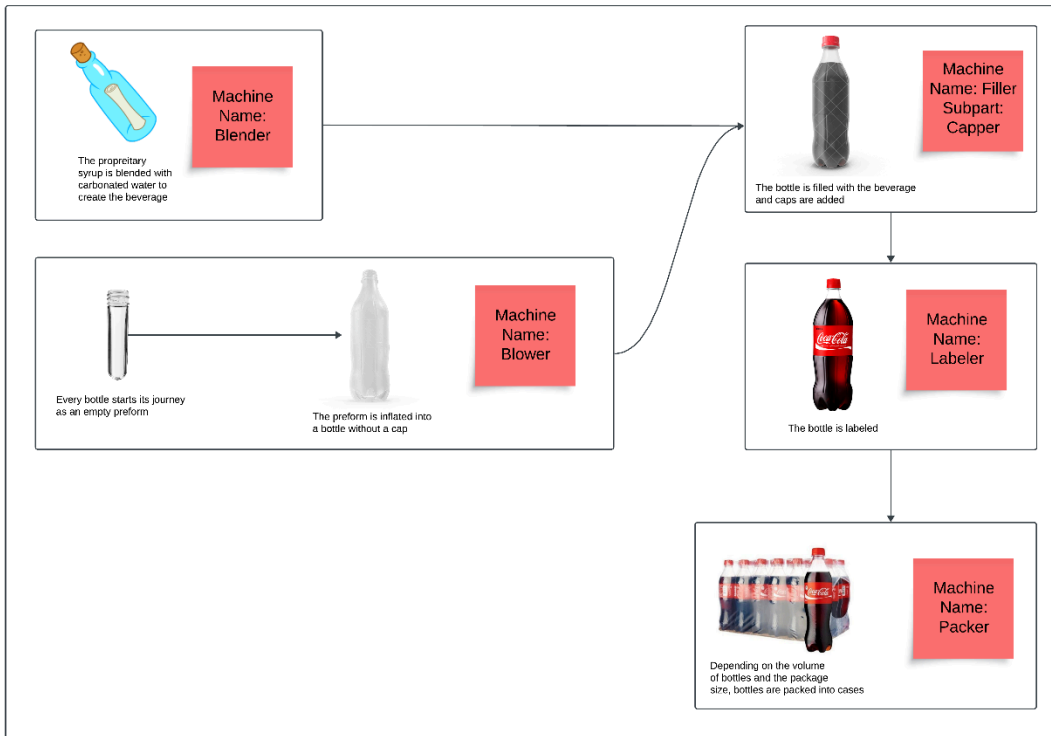


Figure 1. The Coca-Cola Bottling Process

At any given point of time, the process or line may be stopped. This can happen for the following reasons:

1. George notices some issue and pushes a big red button. This may or may not be an issue serious enough to have an IWC work order issued to it. Some minor manual adjustments might get the machine up and running in no time. Note that these downtimes will not be recorded in IWC.
2. The machines need regular maintenance just like your car needs oil changes at regular mileages. These are called 'Planned maintenance' orders and are reported in IWC. So, let's say a machine needs a certain bearing to be changed once every six weeks, this will be scheduled on IWC which will remind George to turn off the machine

once every six weeks and swap out this bearing. IWC records this as 'Preventive Maintenance'.

3. Machines may turn off automatically if an error occurs. This causes George to issue a work order in IWC, but the machine troubleshoots and self resumes, and the work order is no longer needed, and therefore closed immediately, however, the fact that the work order was issued is a red flag and needs to be observed or studied to know why and when it happened and how it can be prevented in the future.
4. Machines can break down. Wear and tear are a natural part of any machine's life cycle. When this happens, an IWC work order is issued as 'Unplanned Maintenance' or 'Corrective Maintenance' and a description and resolve are recorded. Examples provided later in the dossier.
5. In some cases, the fix itself is barely a couple of minutes, however, the plant managers might get busy with other work and not record the work order until the end of the shift retroactively. In such cases you will notice no difference in execution start time and execution end time, but the actual work in minutes should be a reliable metric.

The data: You have access to the IWC dataset, along with a functional location mapping dataset, containing records of the IWC work orders and where the orders link to, respectively.

Here's a look at an exploded view or a data dictionary for each column.

IWC_Work_Orders_Extract.csv

This data contains 1,427,264 rows exclusively of all the information that has been fed into IWC. Below are the columns and what they mean:

Column Name	Definition
ORDER_ID	An identifier used to uniquely map every order raised in IWC
PLANT_ID	Helps uniquely identify all factories of Swire Coca-Cola, between Monza, Cota, Silverstone, Monaco, Roma and Suzuka
PRODUCTION_LOCATION	Text description of Plant_ID
EXECUTION_START_DATE	The date on which work began on the said issue
EXECUTION_FINISH_DATE	The date on which work ended to resolve said issue
ACTUAL_START_TIME	The time at which work began on resolving said issue*
ACTUAL_FINISH_TIME	The time at which work ended on resolving said issue*
ACTUAL_WORK_IN_MINUTES	The time it took, in minutes, to resolve said issue*

MAINTENANCE_PLAN	If the maintenance task was planned, this would have the plan ID. If this column is null, that implies that the maintenance was unplanned
MAINTENANCE_ITEM	If the maintenance task was planned, for example, this bearing needs to be changed every six weeks, then the plan in the IWC system will have code allocated to it, saying this bearing code a maintenance will repeat at X interval and the serial number you see in this column maps to bearings (Mapping not provided for security reasons)
MAINTENANCE_ACTIVITY_TYPE	Lets you know if the maintenance was planned or unplanned, as an additional flag
ORDER_DESCRIPTION	The description entered by the mechanic as to what was the actual work that was carried out in said fix
MAINTENANCE_TYPE_DESCRIPTION	Tells you if the work order was preventive in nature, i.e. done to prevent breakdowns in future, or corrective in nature, i.e. done to correct something that caused a downtime
FUNCTIONAL_LOC	Works like an IP address. Helps locate each machine by plant – process – sub process – product line – machine. Is broken down into nodes for better analysis
FUNCTIONAL_AREA_NODE_1_MODIFIED	The functional area within said plant – Production, fleet, HVAC, etc.
FUNCTIONAL_AREA_NODE_2_MODIFIED	The region-based subset of the functional area – Is it the blender room, the assembly lines – can, bottle or bibs
FUNCTIONAL_AREA_NODE_3_MODIFIED	Within each region, which subset of equipment does this work order belong to
FUNCTIONAL_AREA_NODE_4_MODIFIED	Within line groupings, which line is it specifically referring to – contains line number
FUNCTIONAL_AREA_NODE_5_MODIFIED	Further splits the line into individual machine types
EQUIPMENT_ID	Within a filler, there are close to 60 smaller components. This will help us identify which component the work was done on within that filler, for example
EQUIPMENT_DESC5	Helps us identify what that component is. Mind you, we won't always have a description because some components are either too small to be classified or proprietary and hence undisclosed

EQUIPMENT_CAT_DESC	Identifies if the fixed part was on a machine, a vehicle, an HVAC system, etc.
EQUIP_START_UP_DATE	The date on which this machine was acquired **
EQUIP_VALID_FROM	The date on which this machine became functional **
EQUIP_VALID_TO	The date until which this machine shall be used **

*If ACTUAL_START_TIME = ACTUAL_END_TIME, note that this is a work order where either it was created, feared to be redundant and then immediately closed, or the work order was created after the actual work was done, in which cases the ACTUAL_WORK_IN_MINUTES would be a more accurate representation of the scale of work undertaken. As a rule of thumb, favor the ACTUAL_WORK_IN_MINUTES when in doubt. Dates can just be used for periodicity analyses, if any.

**A machine is acquired on the EQUIP_START_UP_DATE. It becomes functional on EQUIP_VALID_FROM. It will be in service until EQUIP_VALID_TO. If the EQUIP_VALID_TO is '9999-12-31', that means it is currently in use. Note that these columns will be extremely relevant if someone were to choose to correlate the age of a machine with the frequency or severity of breakdowns.

Annexures and additional references:

A temporal analysis of machine failures could look like below:

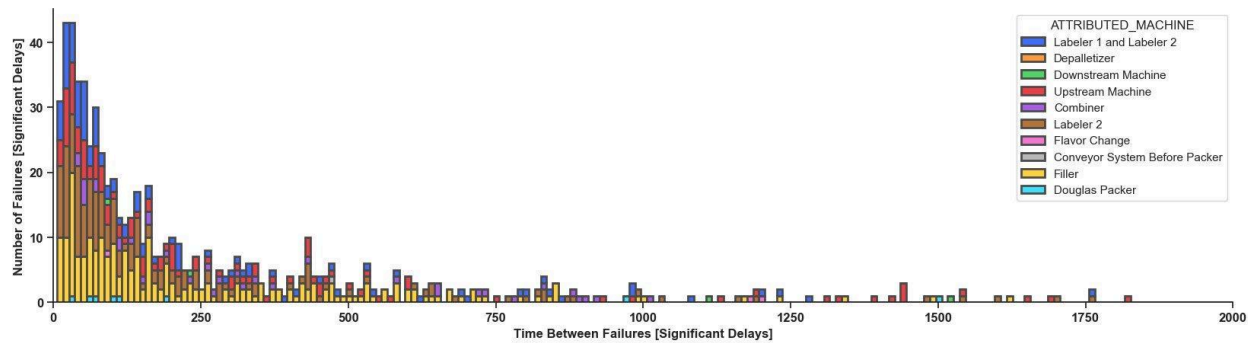
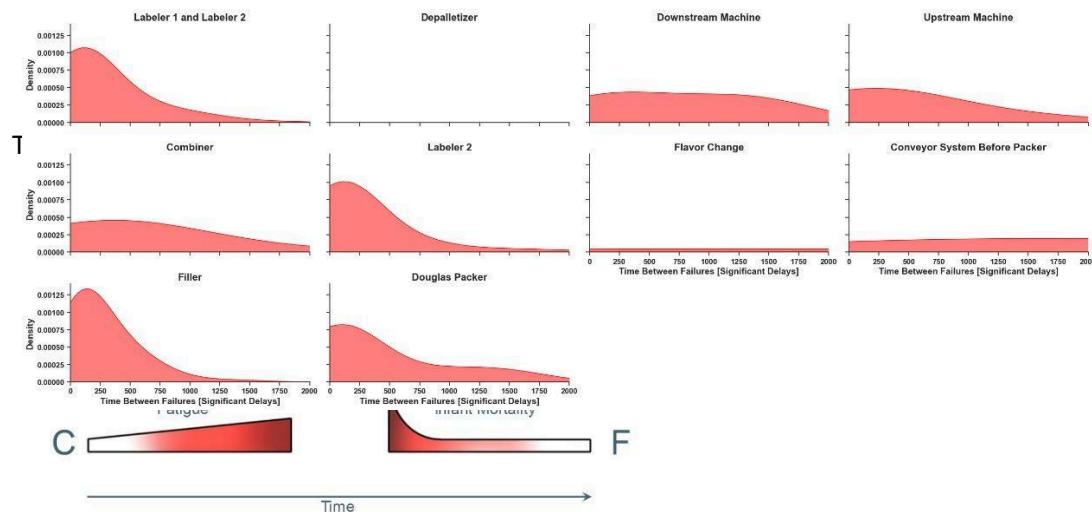


Figure 2. The chart shows that the most significant failures occur within short intervals (0-250 minutes). The frequency of failures decreases as the time between failures increases.

When you smooth the data of each failure mode to reveal underlying trends, the below patterns emerge:



And the shape of the same indicates the below understanding of anticipating and managing product failures:

Bathtub Curve (A): You might focus on improving quality control to reduce early-life failures and designing for durability to extend the useful life.

Wear-Out (B): You could plan for maintenance or replacements as the product ages to avoid unexpected failures.

Fatigue (C): Testing and improving resistance to repeated stresses can extend product life.

Initial Break-In (D): Implementing burn-in procedures can help eliminate early failures before products reach customers.

Each curve gives insights into how a product might fail over time, helping businesses to plan maintenance, improve design, and manage product lifecycles more effectively.

