# Manufacturing Downtime

In manufacturing, **downtime** refers to any period of time in which production has stopped, either facility-wide or on one piece of equipment.

We're members of **Insightopia** team, focused on analyzing the factors contributing to downtime in drinks manufacturing line.

## Three phases/components of the project:

**Building Data Model** | **Analysis and forecasting questions** | **Visualization Dashboard**

**Due Date (all components): 11/04/2025**

**Brainstorm** → **Plan** → **Indicate KPIs** → **Review literature** → **Gather Requirements** → **Analyze Data & design** → **Implement & execute code**

---

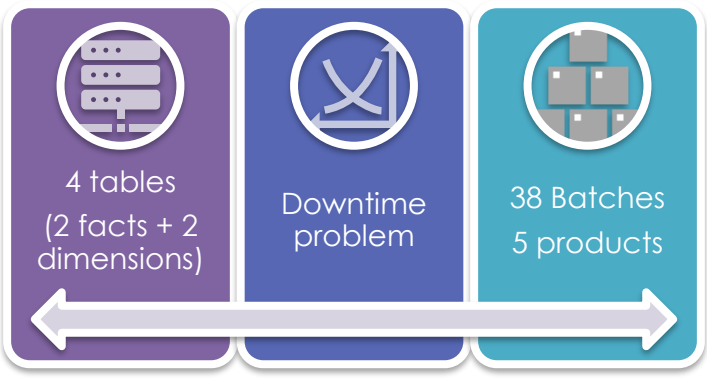
## Project Proposal

### The overview and the objective of the project:

A dataset of manufacturing line productivity of drinks needs to be cleaned and analysed to extract insights in order to achieve the following objective: "***addressing the most downtime factors affecting the productivity in August and September 2024 in order to reduce manufacturing downtime by 20% at least***".

To reach this objective we need to answer the following questions through the project:

- How does batch processing time compare to the minimum batch time for each product?
- What are the most common downtime factors?
- What is the total production time per operator?
- Which operators have the highest and lowest downtime incidents?
- What is the overall efficiency of the manufacturing line (productive time vs. downtime)?



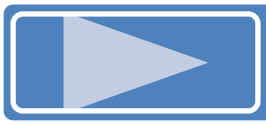
## The scope of the project:

- **In-scope:**
  - clean, preprocess and build a data model for the dataset.
  - Determine all the possible analysis and forecast questions to be answered.
  - Analyse the downtime keys/factors and predict the downtime in the next period.
  - Highlight the number of batches to be produced after predicting the downtime.
  - Build a dashboard to visualize the answers of the questions.
  - Prepare a report summarizing the work of the project.
- **Out-of-scope:**
  - Real time monitoring the system of manufacturing.
  - Changes the workflow or scheduling of production.
- **Roles of Data analyst Team members:**
  - **The leader**: *Mohamed Abdelghaffar*.
  - **Data governors**:
    - Mohamed Abdelghaffar:
      - Follow up the procedures of the project plan.
      - Preprocess the dataset and prepare the analysis questions.
    - Madonna Wadeaa):
      - Clean the data and handle the missing or null value in the dataset.
      - Answer the analysis questions using SQL.
  - **Data storyteller:**
    - Yasser Mabrouk:
      - Answer the analysis questions using Python.
      - Prepare the dashboard using PowerBi.
    - Saif Tarek:
      - Answer the analysis questions using Python.
      - prepare the forecasting questions
- **Deliverables**:
  - Cleaned dataset ready for analysis.
  - Preprocessing notebook.
  - Set of analysis and forecasting questions.
  - Visualization dashboard for answering the questions.
  - Final report and presentation.

---

# Project Plan

## Key Milestones:



**Project Kickoff:**
- Define the scope and objective.
- Brainstorm the roles of the members.

**Preprocessing stage:**
- clean and handle the data.
- Build data model.

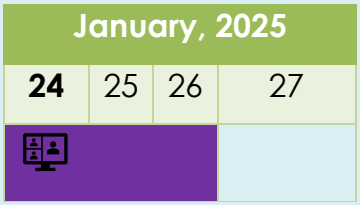
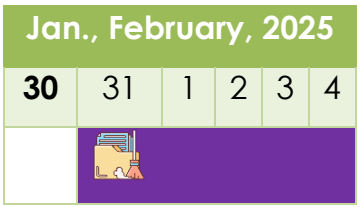
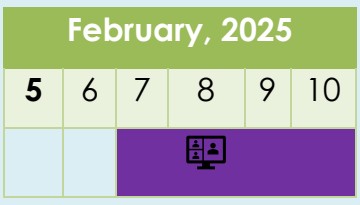
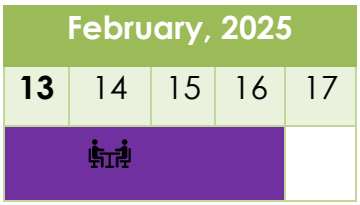
**analysis questions:**
- determine set of questions to be analysed.
- answering the questions in sql/python.

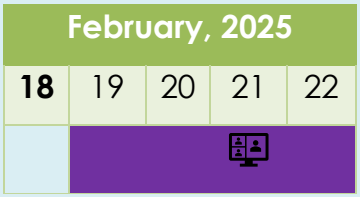
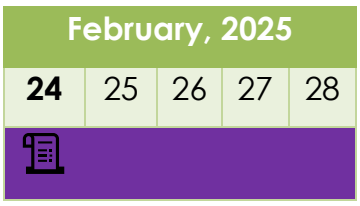
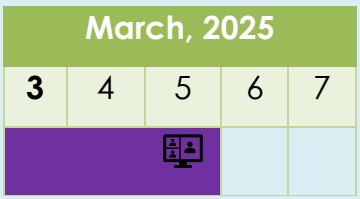
**forecasting questions:**
- determine set of questions to predict the downtime.
- answering the questions using the trends of the data.

**build a dashboard:**
- visualize the answers on a dashboard.

**prepare the final report:**
- summarize the project work on a report.
- prepare a presentation.

## Project Timeline (Gantt Chart):

| Task | Assigned to | Period | Notes |
|---|---|---|---|
| Project kick off | Mohamed Abdelghaffar |   | online meeting with all team members |
| Preprocess and clean the data | Madonna Wadeaa Saif Tarek |   | The leader follows up work and gathers the cleaned dataset |
| Build the data model | Mohamed Abdelghaffar |   | Meeting with the instructor to show him our work. |
| Set the analysis questions | Yasser Mabrouk Madonna Wadeaa |   | Meeting at Eljazera club to discuss the questions |

| Task | Assigned to | Period | Notes |
|---|---|---|---|
| Answer the analysis question using tool (sql or Python) | Mohamed Abdelgh.<br>Saif Tarek |  | Meeting with the instructor to discuss the progress of the project. |
| Set the forecasting questions | Saif Tarek<br>Yasser Mabrouk |  | The due date of submitting the project planner on GitHub. |
| Answering the questions using the trends of the data | Yasser Marbouk<br>Mohamed Abdelgh. |  | Online meeting with the leader to discuss the project progress. |
| Build the dashboard | Madonna Wadeaa<br>Saif Tarek |  | The leader is following up work and gathering the dashboard. |
| Prepare the final report | Mohamed Abdelghaffar |  | online meeting with all team members |
| Prepare the final presentation | Mohamed Abdelghaffar<br>Yasser Marbouk |  | Assign the final steps and methods of presentation |
| **Closing the project** | All the team members |  | The due date for the final presentation, testing and reports |

## Risk Assessment and Mitigation Plan

Potential risks that could impact on the success of the analysis and its implementation:

a) Data collection issues:

- Risk: incomplete data or too many missing values in the dataset.
- Impact: poor analysis leading to incorrect downtime cause identification.
- Mitigation: validate, clean and preprocess the data before the analysis.

b) Analysis challenges:

- Risk: incorrect statistical and data model leading to misleading conclusions.
- Impact: ineffective decision making and wasted time.
- Mitigation: use multiple analysis tools (SQL, python, Power BI) and continuously refine models with the instructor reflection.

c) Lack of actionable insights:

- Risk: analysis results are too complex or do not translate into clear improvements.
- Impact: difficulty in implementing recommendations leading to no improvements in downtime
- Mitigation: present the findings in a clear visual dashboard (e.g. charts) and provide actionable recommendations.

d) Lack of meeting time:

- Risk: team members have their work hours or it's hard to agree on specific time for all members.
- Impact: delay in following the project plan procedures leading to poor willingness to work.
- Mitigation: schedule online meetings via MS. Teams and send the instructions on WhatsApp group of the team.

**Risk Assessment Matrix:**

| Risk category | Likelihood | Impact | Priority |
|---|---|---|---|
| Data collection issues | Medium | High | High |
| Analysis challenges | Medium | Medium | Medium |
| Lack of actionable insights | Medium | High | Medium |
| Lack of meeting time | High | High | Medium |

# Key performance indicators (KPIs)

**Production Efficiency KPIs**

1. **Total Production Time (hours)** – Sum of all batch production times.
2. **Average Cycle Time (minutes per batch)** – Average time taken to complete one batch.
3. **Production Throughput (batches per shift/day)** – Number of batches produced per shift or day.
4. **Operator Efficiency (%)** – Production output per operator per shift.

**Downtime Analysis KPIs**

5. **Total Downtime (minutes/hours per day)** – Sum of all downtime across batches.
6. **Mean Downtime per Batch (minutes)** – Average downtime per production batch.
7. **Downtime as % of Production Time (%)** – Ratio of downtime to total available production time.
8. **Top Downtime Causes (%)** – Percentage contribution of different downtime factors.

# Literature Review

**Lecturer's feedback and evaluation:**

| Date | Feedback | Team's Response |
|---|---|---|
| **8/2/2025** | Add the duration time per product.<br>Answer some analysis questions with Python and the others with SQL. | **Doing the requirements** |
| **22/2/2025** | Write stakeholder's analysis in the project planner's file<br>Modify the code for some analysis questions to make their answer clear | **Finish the requirements** |

**Final Grading Criteria – Breakdown of Marks:**

---

**Final Grading Criteria – Breakdown of Marks**

| Category | Weight (%) | Evaluation Criteria |
|---|---|---|
| Documentation | 30% | Clarity, completeness, and structure of reports (e.g., risk assessment, KPIs, methodology). |
| Implementation | 30% | Quality of analysis, data accuracy, downtime reduction strategies, and process improvements. |
| Testing & Validation | 20% | Effectiveness of downtime predictions, accuracy of models, and validation against historical data. |
| Presentation | 20% | Clarity, engagement, visual representation of data (charts, dashboards), and ability to answer questions. |

**Detailed Breakdown**

1. Documentation (30%)
   - 10% – Clearly defined objectives, problem statement, and project scope.
   - 10% – Detailed methodology, including data collection and analysis techniques.
   - 10% – Well-structured reports with risk assessment, KPIs, and conclusions.
2. Implementation (30%)
   - 15% – Correct use of downtime analysis techniques and data-driven insights.

- 15% – Effectiveness of proposed mitigation strategies and recommendations.

3. Testing & Validation (20%)
   - 10% – Accuracy of downtime predictions and effectiveness of implemented solutions.
   - 10% – Validation against historical downtime records and real-world conditions.
4. Presentation (20%)
   - 10% – Clear and engaging explanation of findings, recommendations, and impact.
   - 10% – Use of visuals (graphs, dashboards) and ability to answer project-related questions.

# Requirement Gathering

## Stakeholder's analysis:

is a process of identifying people responsible for the project of the manufacturing line; grouping them according to their levels of participation, interest, and influence in the project; and determining how best the insights of the analysis impact on each of these stakeholders.

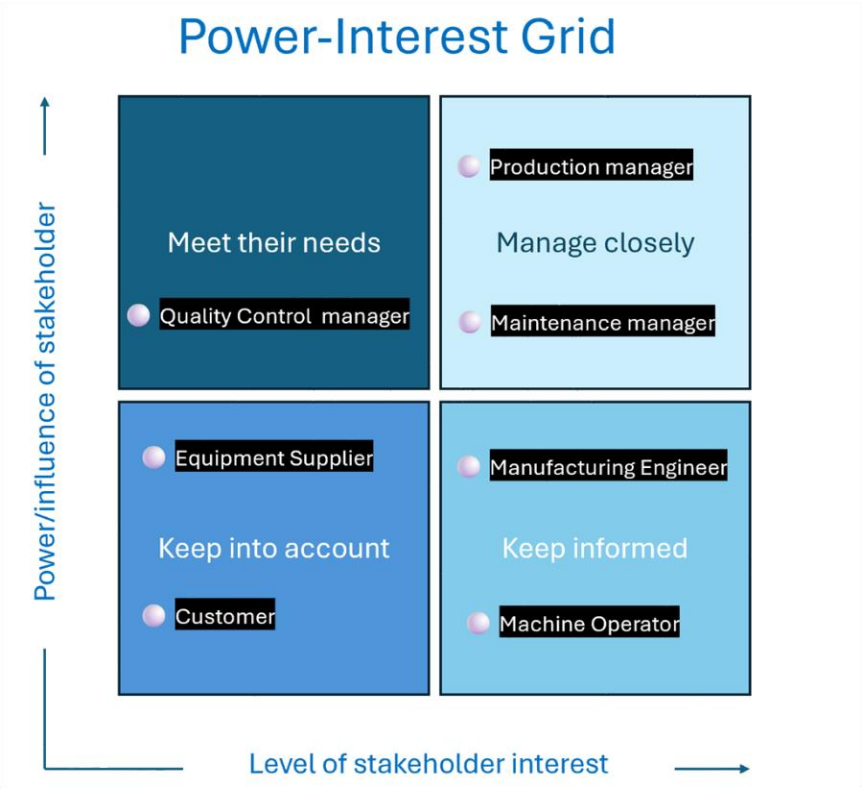
**Introduction:**

The purpose of the analysis [1]:

- to enlist the help of key organizational players.
- To gain early alignment among all stakeholders on goals and plans.
- Proactively address potential conflicts or concerns to mitigate risks and improve collaboration.

**Overview of the stakeholders:**

| Internal stakeholders | External stakeholders |
|---|---|
| Production Manager | Equipment supplier |
| Quality control manager | Customer |
| Maintenance manager | |
| Machine operator | |
| Manufacturing engineer | |




[1]) What is Stakeholder Analysis? | Definition and Overview

Stakeholders' needs and expectations:

| Stakeholder | Need | expectation |
|---|---|---|
| Production Manager | Data-driven insights on downtime causes and trends | Clear reports with actionable recommendations. |
| Maintenance manager | Accurate downtime root cause analysis. | Reliable maintenance scheduling based on data insights. |
| Quality control manager | Analysis of defects or inconsistencies linked to machine stoppages. | Integration of downtime analysis with quality control data. |
| Manufacturing engineer | Insights into how automation or process adjustments can reduce downtime. | Use of real-time monitoring systems. |
| Machine operator | Clear procedures for responding to downtime events | Timely maintenance support. |
| Equipment supplier | Feedback on machine performance and failure rates. | Timely reports on equipment failures. |
| Customer | High-quality products despite operational challenges | Strong quality control despite production issues |

**Impacts of findings on stakeholders:**

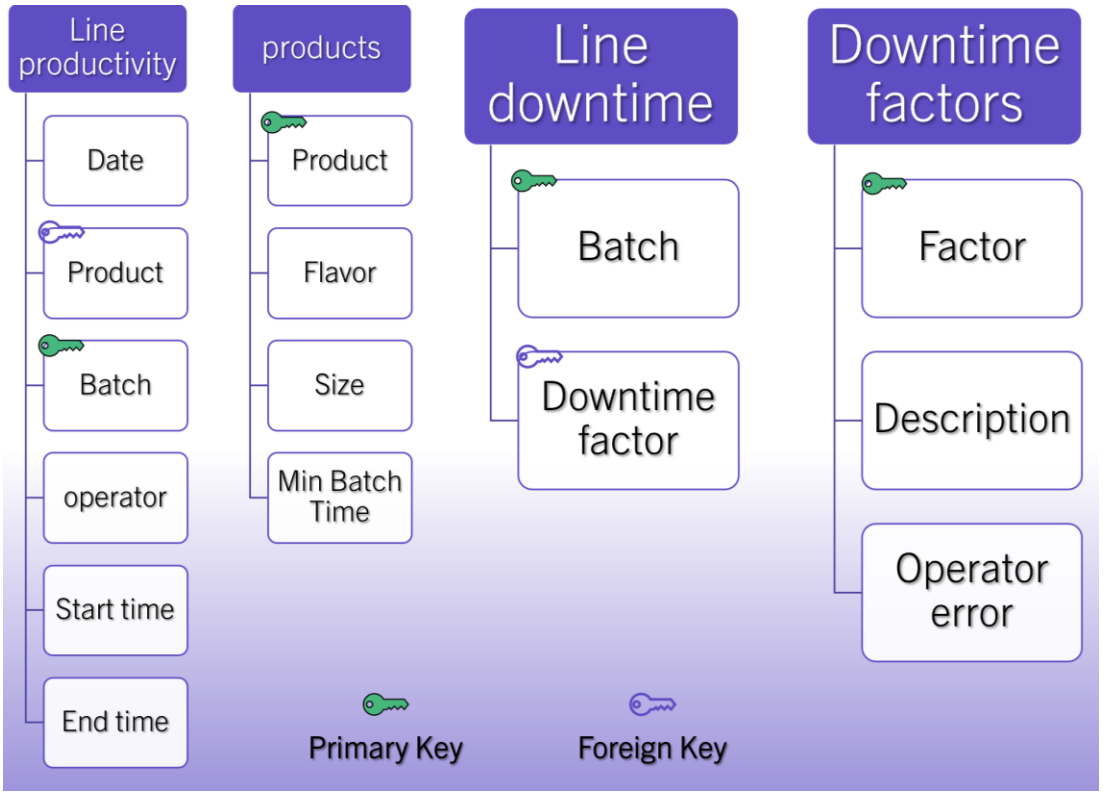
| The analysis question | the stakeholder/s | The impact/s |
|---|---|---|
| **How does batch processing time compare to the minimum batch time for each product?** | Production manager<br>Quality control manager | |
| **Are there batches with no downtime?** | Manufacturing engineer | |
| **How Many Batches were Produced Per Day?** | Maintenance manager<br>Machine operator | |
| **What is the total production time per operator?** | Production manager<br>Machine operator | |
| **Which product has the highest number of batches?** | Machine operator | |

| The analysis question | the stakeholder/s | The impact/s |
|---|---|---|
| **What are the avg/sum/max production time per product type?** | Machine operator | |
| **Are there significant differences in batch processing time between operators in different days?** | Production manager | |
| **How does productivity vary throughout the day?** | Production manager<br>Quality control manager | |
| **What are the most common downtime factors?** | Maintenance manager | |
| **is there any downtime factor that has no impact on the productivity?** | Production manager | |
| **Which products experience the most downtime?** | Manufacturing engineer | |
| **What is the average downtime per batch?** | Manufacturing engineer | |
| **Which operators have the highest and lowest downtime incidents?** | Quality control manager | |
| **Are certain downtime factors linked to specific operators?** | Manufacturing engineer | |

**Overview schema (Tables and attributes)**:

The problem is to find the downtime keys and try to reduce them by 20%, in the following lines the tables and their attributes of the dataset.

- **Before modification:**



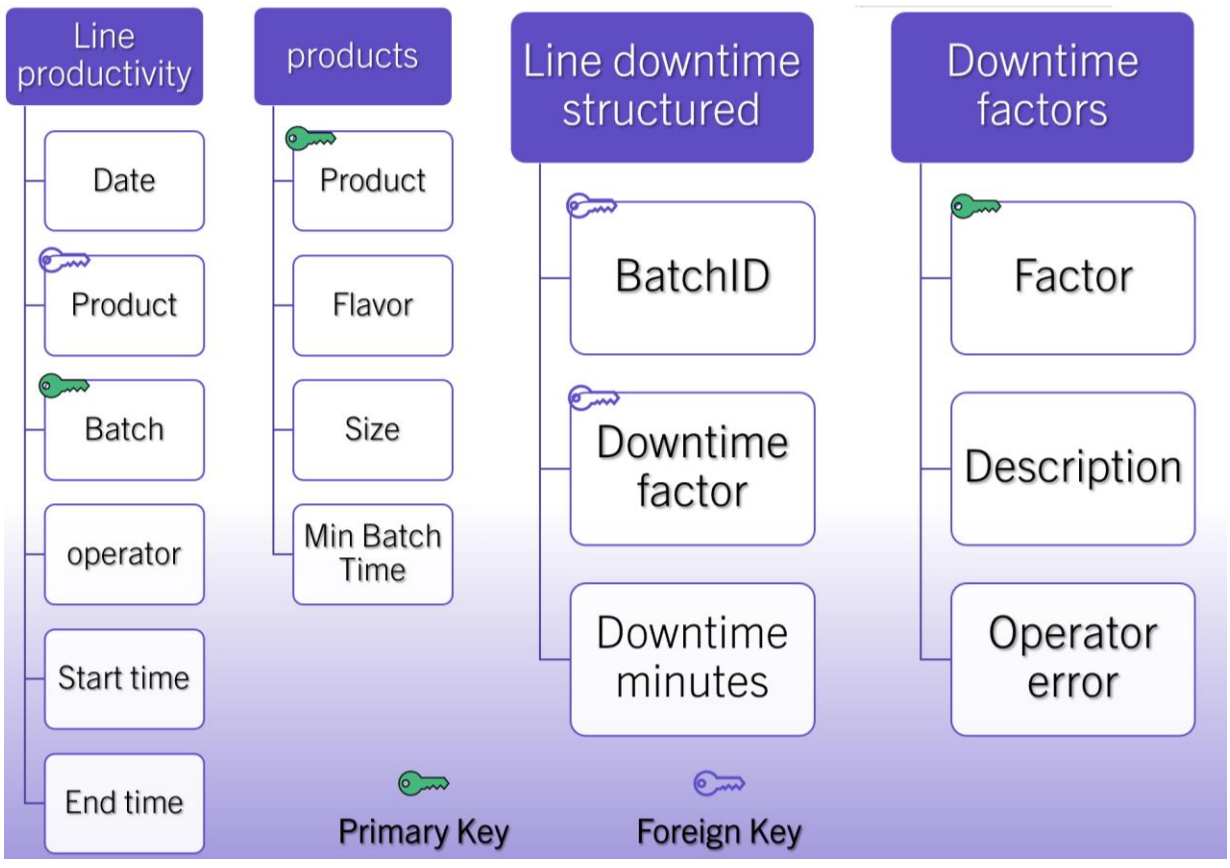

**We found that the table of "line downtime" is not structured very well because it has empty cells as shown:**

| | Downtime factor | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Batch** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
| 422111 | | 60 | | | | | 15 | | | | | |
| 422112 | | 20 | | | | | | 20 | | | | |
| 422113 | | 50 | | | | | | | | | | |
| 422114 | | | | 25 | | 15 | | | | | | |
| 422115 | | | | | | | | | | 24 | | |
| 422116 | | | | | | | | | | | | |
| 422117 | | 10 | | | | 5 | | | | | | |
| 422118 | | | | | | 14 | 16 | | | | 10 | 20 |
| 422119 | | | | 25 | | | | | | | | |
| 422120 | | | | 20 | 15 | | | | 17 | | | |

**So, we had to modify this table into a structured one in order to prepare it for the analysis.**

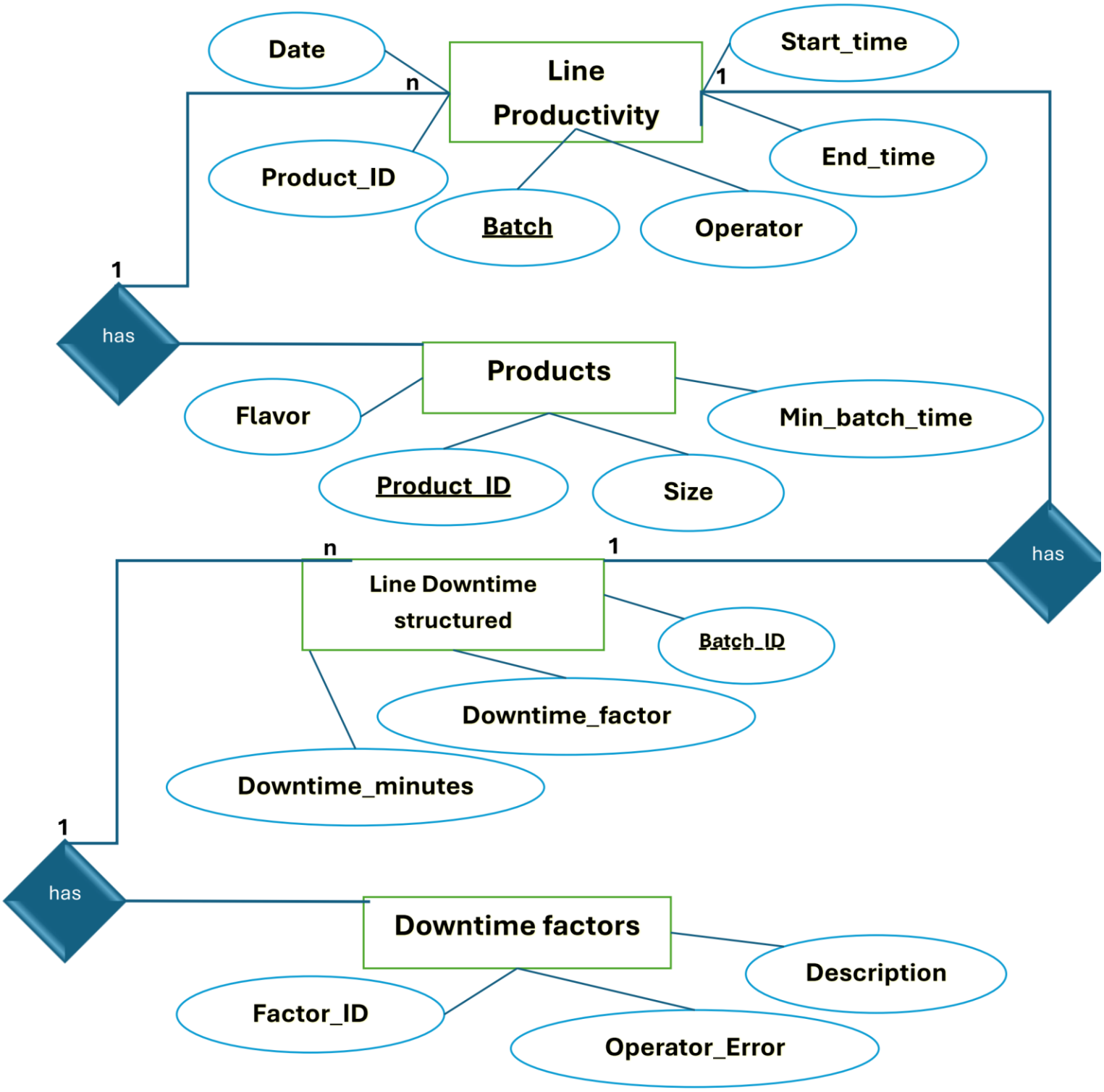
| BatchID | Downtime factor | downtime minutes |
|---|---|---|
| 422111 | 2 | 60 |
| 422111 | 7 | 15 |
| 422112 | 2 | 20 |
| 422112 | 8 | 20 |
| 422113 | 2 | 50 |
| 422114 | 4 | 25 |
| 422114 | 6 | 15 |
| 422115 | 10 | 24 |
| 422117 | 2 | 10 |
| 422117 | 6 | 5 |
| 422118 | 6 | 14 |
| 422118 | 7 | 16 |
| 422118 | 11 | 10 |
| 422118 | 12 | 20 |

- **Overview of schema after modification:**

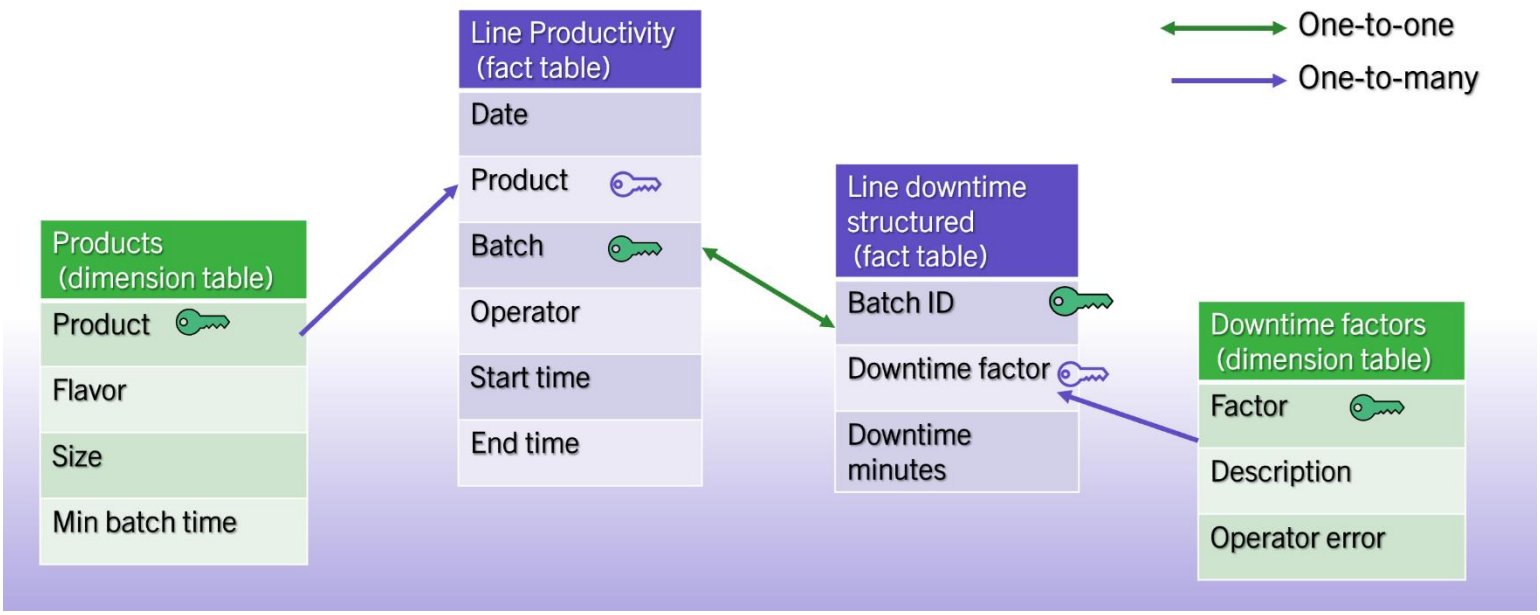


**Database design & Data modeling:**

- ERD (after modification):



## Data Model (Galaxy Schema):




## Data model (from SQL):