



Problem Statement

A washing machine manufacturing company makes a lot of washing machines every year. There are different models and types. And for each type of washing machine, the technology and logistics are different and these items are being manufactured in different parts of the world. All of this is aligned in a supply chain management process with the targeted dates of delivery planned for next 5 years.

As this is a complex process, each department produces a lot of data related to the logistics, supply chain, planning, execution and forecasting of orders and other details. This often means, not only the department that owns a certain type of data produces it, but also the other departments who are either the direct or indirect consumers of the data also produce a possible forecast data. These departments produce the forecast data to keep up with the planning and their day to day activities, instead of waiting for it when it finally reaches them. At the same time, the data owning departments also keep updating the data based on everyday changes, which is an overhead for the consuming teams who have already consumed it and now need to recalibrate their data.

Also, all this data is finally consolidated when the official manufacturing process has achieved them(real-time data).

The intermediate process ends up creating a lot of data by each department, which are then consumed by other departments or sub manufacturing units to plan their logistics and supply Chain.

This intermediate data which gets generated based on milestones achievement in the production process, is mostly a redundant data without any authenticity. This lies in the system consuming a lot of space and memory and in long term and creates sustainability issues.

In the dataset provided, 3 stages in the manufacturing are put-up, fabrication, sub-assembly and assembly. (Dataset provided is just an example. Use it to extend the dataset)

Provide a possible solution and approach to reduce this underlying intermediate data in the system and make the approach of departments using the data more sustainable in the long term.





Requirements

- Create a data lake with a normalized DB to reduce the redundancy.
- Identify the current redundant data from the forecasted data.
- Create an automation process for data stamping(approval) the real time data.
- Create a dashboard for the users in each domain to access the data required for their domain and also allow the forecast and real time data creation.
- Create a dashboard for the data officer to monitor the data stamping process.

Tech stack

Team can choose any language or framework to develop the application.

Restriction

- Use the latest open source technologies.
- Use of proprietary technologies will not be considered .

Submission format

- Source code in any GIT version control system.
- Details about Tech Stack used and in the README.md file in the code
- A Design diagram of the solution and an E-R diagram of the data lake
- Cloud - for hosting and data (preferable)

Evaluation Criteria:

- Platform should meet the requirements and be fully functional for the users.
- Platform should have authentication and authorization for customers like:
 - User and their department.
- Performance and scalability.
- Creativity and innovation.
- Clarity and quality of the presentation or demonstration.





Resources

- Washing machine manufacturing data: [Database](#)

NOTE: Evaluation will start on 13th of May at **10 AM**. There will be **3 to 4 rounds** of evaluation throughout the event and the team is expected to be present for the evaluation in the respective table throughout the event

