

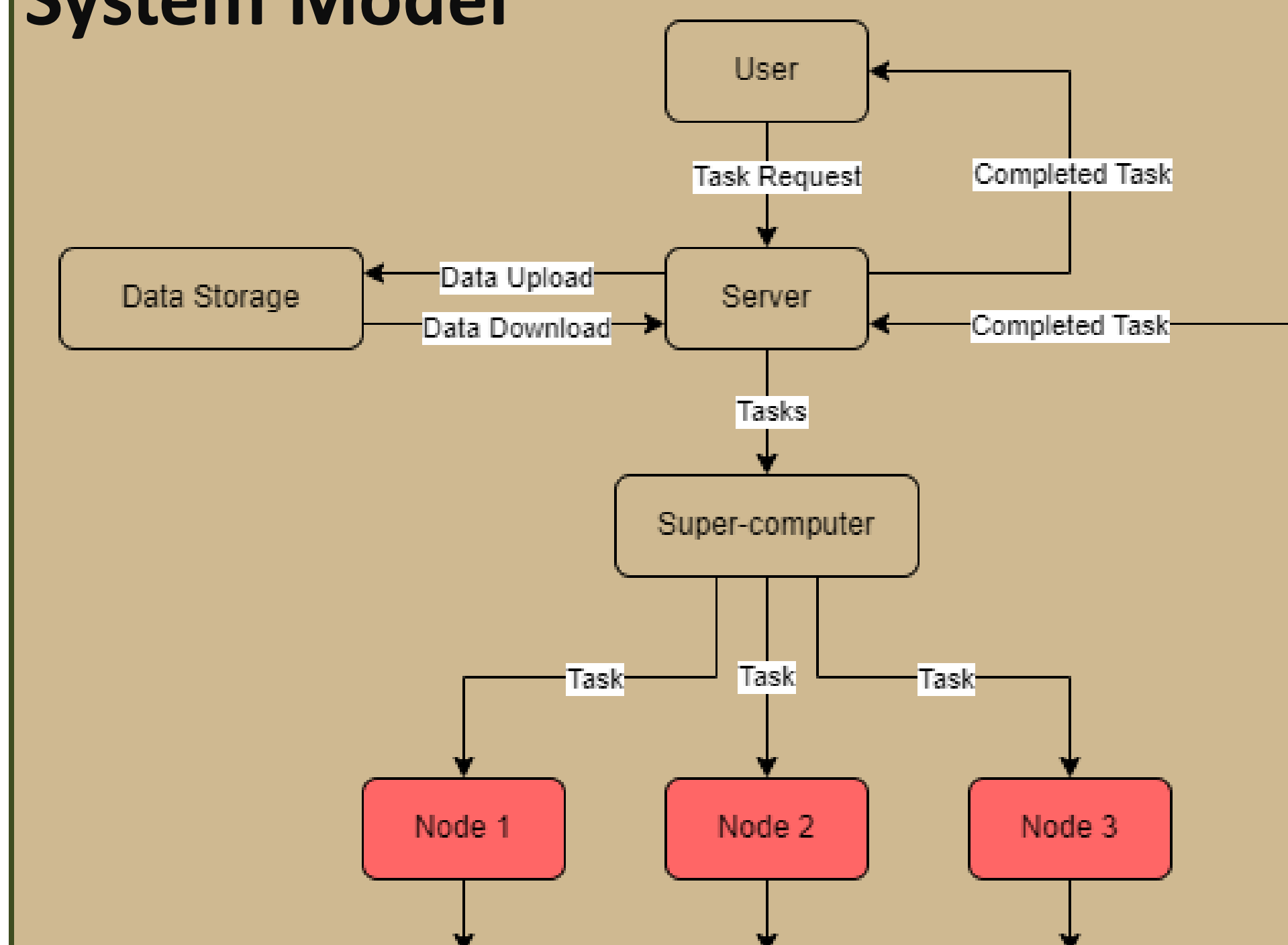
Client Background

The Rosen Center for Advanced Computing (RCAC) at Purdue University provides advanced computational resources and support, including high-performance computing systems and large storage solutions. RCAC also offers training and consultation services to help researchers effectively utilize these technologies in their work.

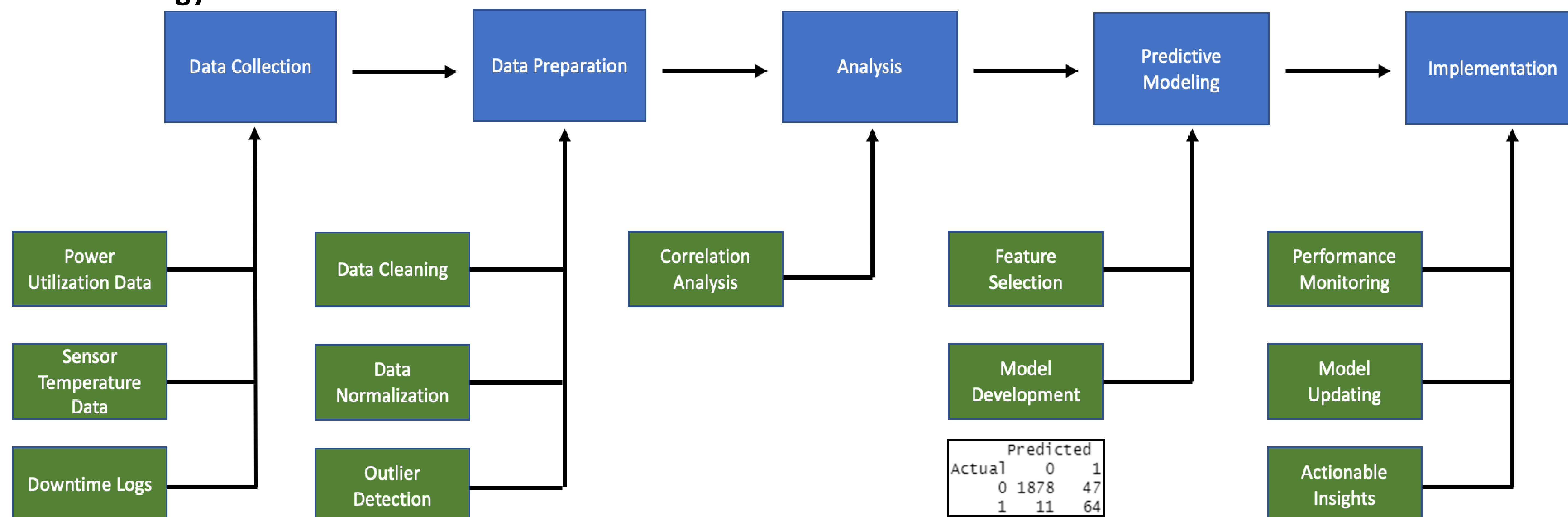
Problem Statement

The Purdue University's Rosen Center for Advanced Computing (RCAC) is experiencing downtime due to unanticipated downtime in individual computational nodes, despite the overall system's high uptime. This downtime disrupts the productivity of research staff, faculty members, and leads to inefficient and costly maintenance efforts. The RCAC lacks the analytical tools necessary to proactively identify and predict factors leading to downtime. Consequently, there is a need for a solution that will allow them to effectively analyze these data through machine learning techniques and predict the nodes' downtimes through an interactive dashboard to facilitate data driven operational decision making.

System Model

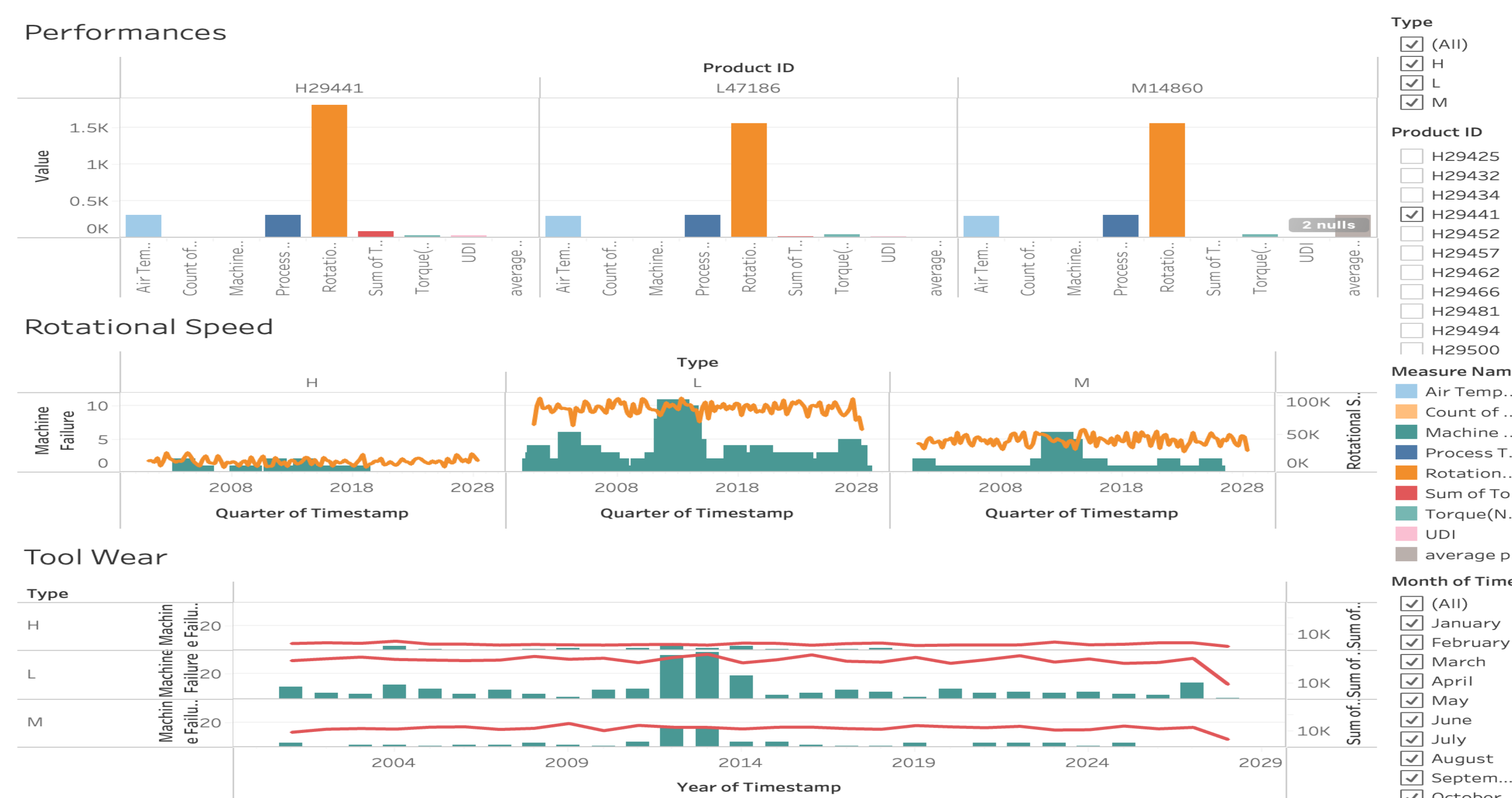
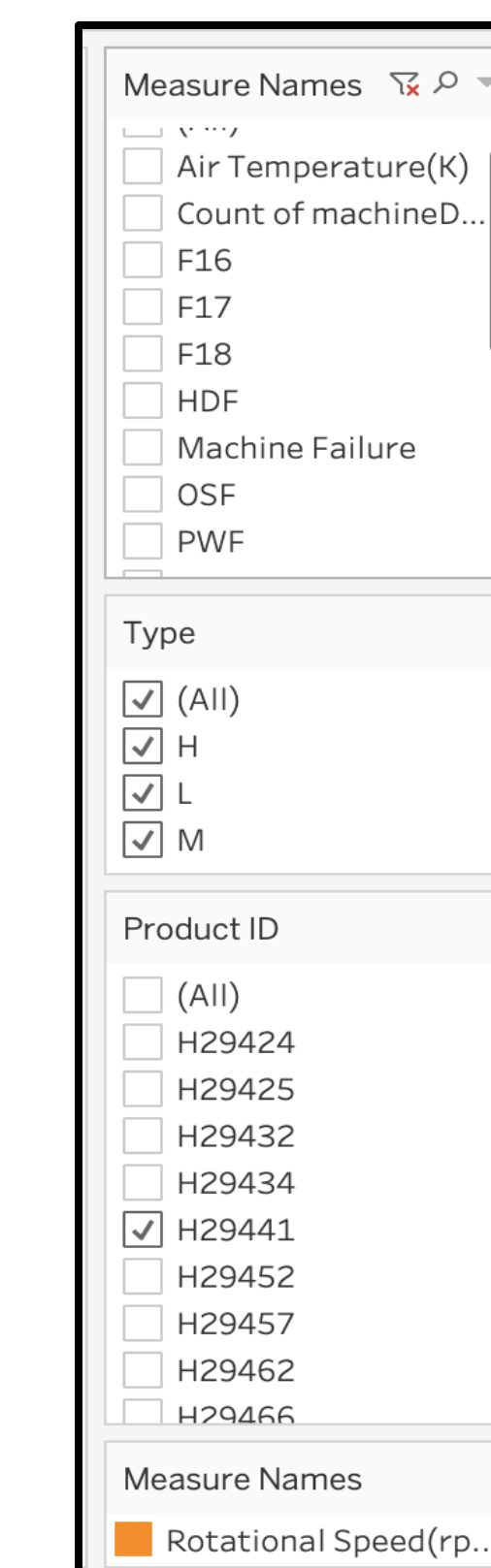
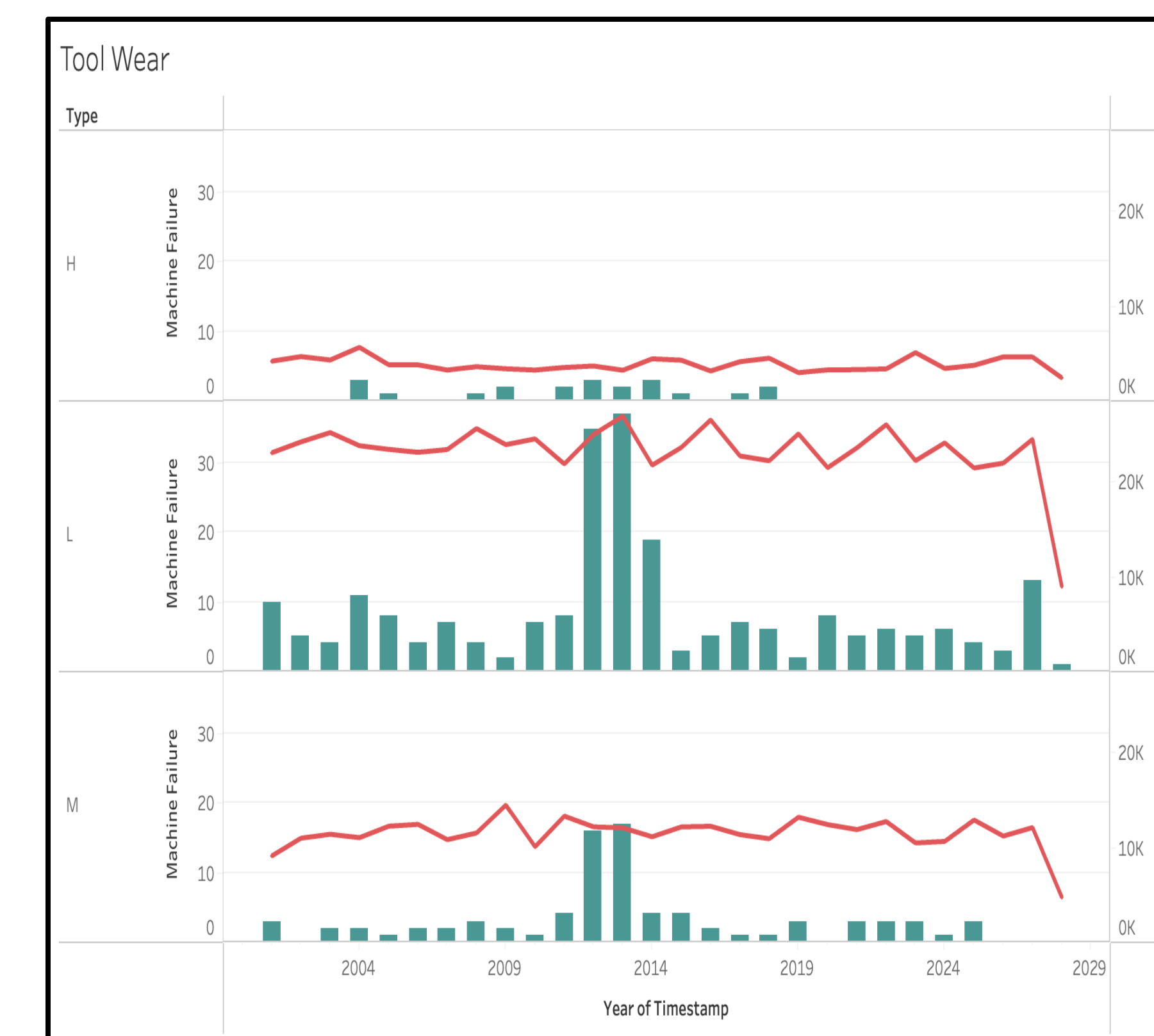
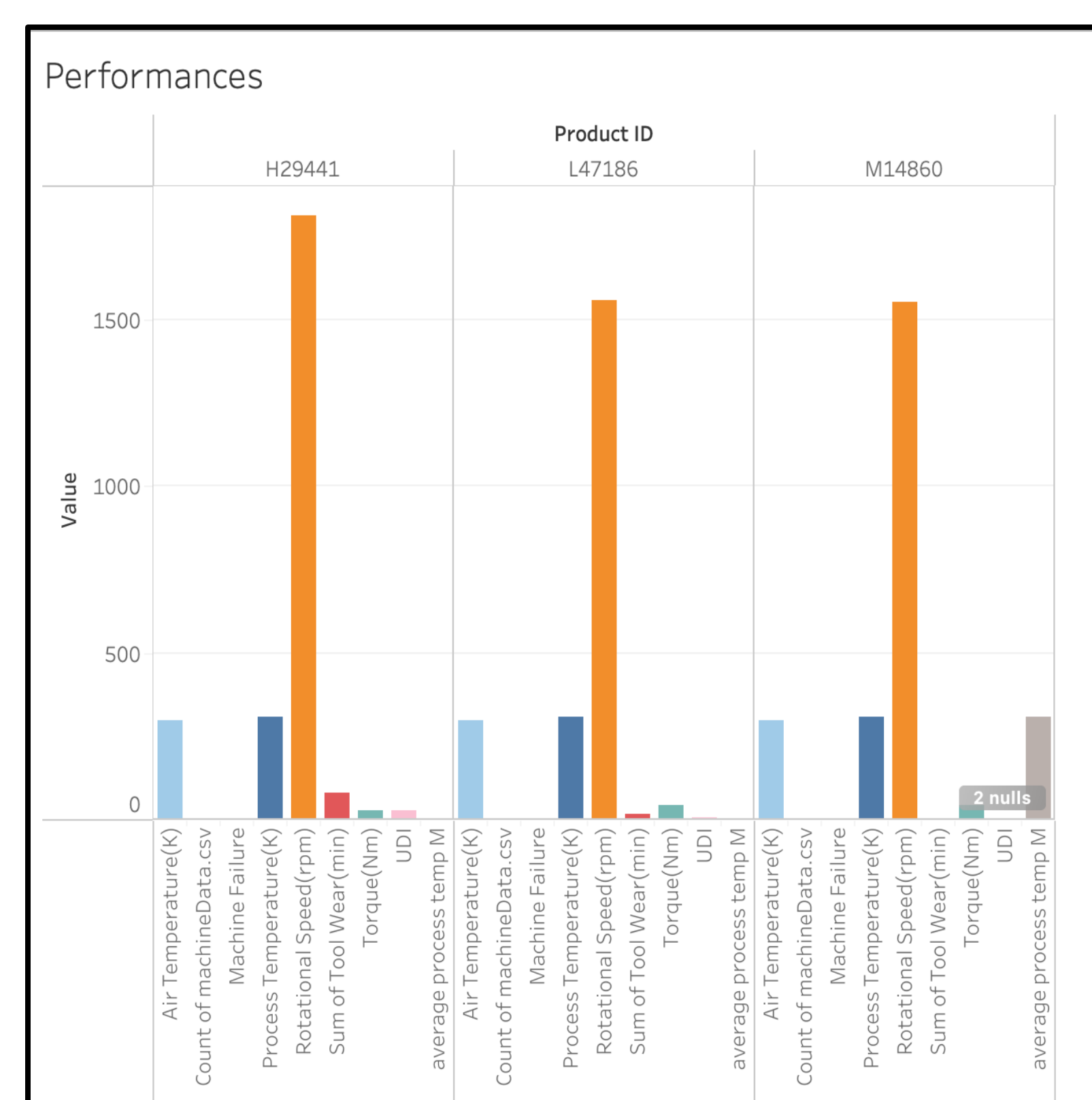


Methodology



Discussion

This project was developed with frequent adjustments to the project scope and expected deliverables due to permissions and size related to data transfers and access to real time data. As a result, the group decided to set realistic expectations, creating an interactive dashboard, giving the client a tool to visualize trends and relationships of critical machine metrics in addition to the machine learning algorithm specialized for RCAC's data. The team was also able to shed light on areas for potential improvement in terms of data collection and monitoring to the RCAC staff. Building upon the current deliverables, the RCAC team or future students can connect the current system to live data to implement a fully automated alert system, effectively notifying staff when action can be taken in real time to prevent unnecessary downtime and losses in revenue.



Results

- **Focus**
 - From analyzing entire system downtime to examining factors that contribute to individual node failures.
- **Machine Learning Model**
 - Accurately predict machine downtime based on critical factors.
 - Further testing with a larger dataset underway for reliability and accuracy validation.
- **Tableau Dashboard:**
 - Visualizes factors impacting node outages.
 - Allows monitoring of thresholds and trends for preventative action