Pranav Bollineni

Auto Defect Insights

Github repo: https://github.com/pbollineni04/auto-defect-insights.git

Job Description: I selected the Data Analyst Intern position at Sanmina because it had a wide range of fields in which the company is a part. This job allows me to apply SQL, data collection, and visualization skills in a real-world industrial context. I am particularly interested in the intersection of manufacturing, quality control, and the automotive sector.

Problem: I plan to solve the problem of identifying patterns in automotive recalls and component failures. This problem is relevant to the job, as it involves analyzing operational data and aligning it with strategic decisions. It is feasible to solve it with a combination of SQL queries, data pipelines (API + web scrape), and dashboards.

Data Sources: NHTSA Vehicle API (https://www.nhtsa.gov/nhtsa-datasets-and-apis#recalls): Provides recall, defect, and safety complaint data across U.S. automotive brands. CarComplaints.com: I will web scrape consumer-submitted complaints and defect reports.

Solution: I will integrate the API and web-scraped data into a database, clean and structure the data, and then perform SQL-based analysis. Key queries will: Aggregate and rank defect types across years and manufacturers, Join government recall data with user complaint data, Use CTEs and window functions to track recurring problems and spot emerging patterns. These insights can help companies like Sanmina identify critical areas for improvement in manufacturing processes.