EDA / Descriptive Statistics

## Introduction:

## An organization that produces fuel pumps. The fuel pump that is used to move the fuel more quickly makes the car drive more quickly. However, there was a hazy downtime throughout manufacturing. Therefore, this could reduce the company's profitability and lengthen delivery times. To make sense of this massive volume of data, compelling insights require effective visualization. Several graphs and charts will be included in our visualization project.

## Overall design strategy

The 2500 records (provided data) that made up the data that was retrieved to learn more about the reason for the failure included details about the machine's operating conditions, including pressure, temperature, speed, and others. There are 165 days between 2021-08-12 and 2022-12-05. Both SQL and Python are used to pre-process this data and perform some coding-based analysis.

Even if there were no null values, the data needs to be cleaned. Overall, to execute the visualization, we used both the raw data and a customized SQL output.

Data visualization color codes are chosen to illustrate the failure rate and no failure rate after the data has been prepared. All things considered; we have also created a model that can anticipate machine failure even before it occurs.

## Data Overview

## Sample Data, which contains a variety of metrics factors related to the stages of machine production, is gathered from the internet. The values were recorded by machines and used in visualization. For machine production level data, about 2500 records are taken; thereafter, the failure rate data is extracted.

## To utilize in SQL, data is extracted into CSV files. Once the data has been loaded, a second extract using custom SQL is made to summarize the production level data.

## Users

Users of the visualization dashboard will be:

* Production manager: Analysis insights enable production managers to optimize resource allocation, identify bottlenecks, and improve efficiency.
* Maintenance team: Analysis helps maintenance teams implement proactive strategies, reduce downtime, and increase machine lifespan.

## Questions

Questions which will be answered by this visualization:

* Are there any specific patterns or trends in the Hydraulic Pressure, Coolant Pressure, or Air System Pressure that could indicate potential issues with the machine's performance?
* How does the Coolant Temperature correlate with the Hydraulic Oil Temperature? Are there any temperature ranges where the machine operates most efficiently?
* Is there a relationship between the Spindle Vibration and the Tool Vibration? Does higher spindle vibration correspond to increased tool vibration, indicating a potential issue?
* Is there any correlation between the Coolant Pressure or Coolant Temperature and the frequency of machine maintenance needs? Can we identify specific coolant conditions that contribute to increased maintenance requirements?
* Are there any seasonal or periodic variations in the machine's performance or maintenance needs based on the timestamps in the "Date" column?

## Describe Visualization and how it answers the questions

* Are there any specific patterns or trends in the Hydraulic Pressure, Coolant Pressure, or Air System Pressure that could indicate potential issues with the machine's performance?
  + Patterns and Trends: By plotting the Hydraulic Pressure, Coolant Pressure, or Air System Pressure over time, line charts or area charts can reveal patterns or trends. Any fluctuations, recurring patterns, or abnormal values can be visually identified, helping to understand potential issues with the machine's performance
* Is there any correlation between the Coolant Pressure or Coolant Temperature and the frequency of machine maintenance needs? Can we identify specific coolant conditions that contribute to increased maintenance requirements?
  + Correlations and Relationships: Scatter plots or bubble charts can visualize the relationship between Coolant Temperature and Hydraulic Oil Temperature, Spindle Vibration and Tool Vibration, or any other parameter pairs. The plotted points can indicate the strength and direction of the correlation, allowing for a quick assessment of the relationship between variables.
* What are the primary reasons for machine downtime based on the Downtime feature? Are certain issues occurring more frequently than others, and how do they impact production schedules?
  + Optimal Operating Conditions: Through line charts or scatter plots, the Spindle Speed can be plotted against Voltage and Torque. The visual representation can reveal any specific combinations or ranges of parameters where the machine demonstrates improved performance or higher cutting quality.
* Are there any seasonal or periodic variations in the machine's performance or maintenance needs based on the timestamps in the "Date" column?
  + Maintenance Needs: By visualizing coolant-related parameters such as Coolant Pressure and Coolant Temperature alongside maintenance events, patterns or trends can be identified. This visualization can help understand if specific coolant conditions are associated with increased maintenance requirements. This graph also compares against Conclusion

## Conclusion

Based on the analysis of the machine features, we can draw a simple conclusion:

The machine's performance and efficiency are influenced by major factors such as **Hydraulic Pressure, Torque, Cutting**. Monitoring and analyzing these parameters can help identify potential issues, optimize maintenance schedules, and improve overall machine performance.

Implementing proactive maintenance strategies based on the analyzed data can reduce unplanned downtime, optimize resource allocation, and enhance product quality. By visualizing the data, we can identify patterns, correlations, and trends that provide valuable insights.

By leveraging the power of visualization, we can gain a deeper understanding of the machine's performance and make data-driven decisions to optimize production processes, reduce costs, and improve overall profitability.