EDA / Descriptive Statistics

## Introduction:

Machine downtime refers to periods when a machine or system is not operational or not producing output. Downtime can significantly impact productivity, efficiency, and profitability in various industries, especially manufacturing, IT, and service sectors. Understanding the types, causes, and implications of machine downtime is crucial for businesses to implement strategies for minimizing its occurrence and mitigating its effects.

## Overall design strategy

Gather data on machine downtime. This data might include the date and time of downtime events, duration, causes, machine types, maintenance logs, and operational parameters.

The data which was extracted to get insights into machine performance consisted of more number of records which covers information related coolent temperature, vibrations, pressure, voltage, e.t.c.,

Ensure the data is clean by handling missing values, removing duplicates, and correcting errors.

Convert data into a suitable format for analysis. This might include converting time formats, categorizing causes, or creating new features.

Calculate basic statistics to get an overview of the data.

* Mean, median, mode of downtime duration.
* Standard deviation and variance.
* Frequency of downtime events.

Examine which causes are most frequent and result in the longest downtimes.

Look for patterns over time, such as whether downtime events are more frequent during certain periods.

Analyze which machines are most prone to downtime.

Use regression models to predict future downtime based on historical data.

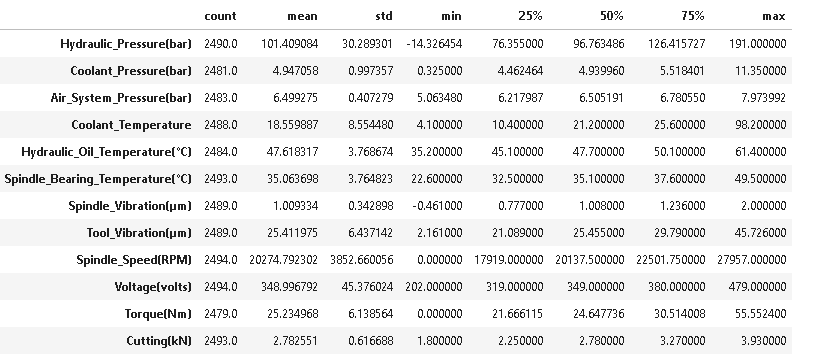
## Data Overview

Data contains machine parameters and downtimemachine by machine data will be collected to use in the visualization. A large no. of records are extracted for machine parameters which cause downtime, downtime.

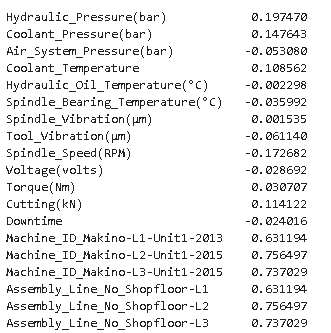
Data is extracted into text files and when uploading into tableau all these files are joined based on the keys. Once the data is loaded another extract summarizing the machine downtime data is created using custom SQL.

**Overview of four Business moments:**

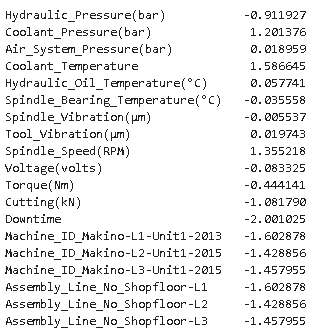
1. **Measures of central Tendency**
2. **Measures of Dispersion**

****

1. **Measures of skewness**

****

1. **Measures of Kurtosis**

****

## Users

Users of the visualization dashboard will be:

* Maintenance Managers: Insights into machine performance, frequent issues, and maintenance planning.
* Operations Managers: Understanding the impact of downtime on production schedules and efficiency.
* Engineers: Detailed analysis of machine performance metrics to identify root causes of downtime

## Questions

1. What is the Failure Rate
2. Which machine unit is most prone to failure
3. Which machine unit is least prone to failure
4. Which component plays key role in the performance
5. Which component if maintained properly would result in reduced down time.
6. How much money can be saved if the machines are maintained in a proper way.

## Answers

1. Failure rate = 49.99%
2. Makino-L1-Unit1-2013 : 436 failures
3. Makino-L2-Unit1-2015 : 382 failures
4. The spindle and the cutting tool.
5. Spindle and cutting tool.
6. Spindle

## Questions

**Maintenance Managers:**

* Downtime Trends: What are the patterns in machine downtime over time?
* Key Issues: What are the most frequent causes of downtime?
* Performance Impact: How does downtime impact overall production efficiency?
* Preventive Maintenance: Which machines require more frequent maintenance?

#### **Operations Managers:**

1. **Production Impact**: How does machine downtime affect production schedules?
2. **Efficiency Analysis**: What is the correlation between downtime and production efficiency?

#### **Engineers**:

1. **Root Cause Analysis**: What are the primary factors contributing to machine downtime?
2. **Metric Correlations**: How do different machine performance metrics correlate with downtime events

## Describe Visualization and how it answers the questions

#### **Downtime Trends:**

1. **Patterns Over Time**:
   * Line charts showing downtime occurrences over time.
   * Insights into seasonal or time-based patterns.
2. **Frequent Causes**:
   * Bar charts displaying the most common downtime reasons.
   * Helps prioritize maintenance activities.

#### **Performance Impact:**

1. **Efficiency Analysis**:
   * Scatter plots correlating downtime with production metrics.
   * Identification of high-impact downtime events.
2. **Preventive Maintenance**:
   * Heatmaps showing downtime frequency by machine and time.

**Manufacturer:**

* What is the trend in the performance of the Machines?
  + The machine downtime data visualization give the information and machine by machine downtime and whether the machines failed or not.
  + This visualization also gives the information on how many failures, assists, steals and fouls the machines have made (we need to select the parameter measure to get each of the metric in the same line graph)
  + The data is also compared against the coolent temperature average to get the insight on where they stand compared to other machines with nearby cooling temperature
* Against what parametersthe machines perform well and against what they perform bad
  + The graph prepared will give the information on the % of failure and not failure against each of the parameters .

**Customers:**

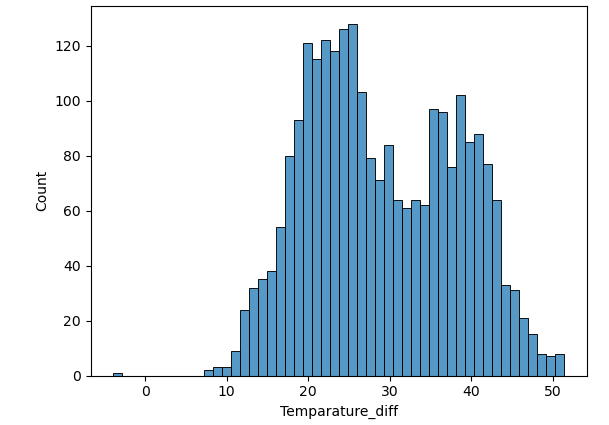
* How is the performance from machine to machine
  + This visualization gives information on the trends in machines at particular parameters. This visualization also gives the information on how many failures, assists, steals and fouls the machines have made (we need to select the parameter measure to get each of the metric in the same line graph)
  + This graph also compared against the coolent temperature average to get the insight on where they stand compared to other machines with nearby cooling temperature
* Against what parametersthe machines perform well and against what they perform bad
* The graph prepared will give the information on the % of failure and not failure against each of the parameters

## The graph prepared will give the information on the % of failure and not failure against each of the parameters

Insights

**Statistical Insights**

* Majority of machine\_temperature(hydraulic oil temperature - coolant temperature) between 19 to 25.

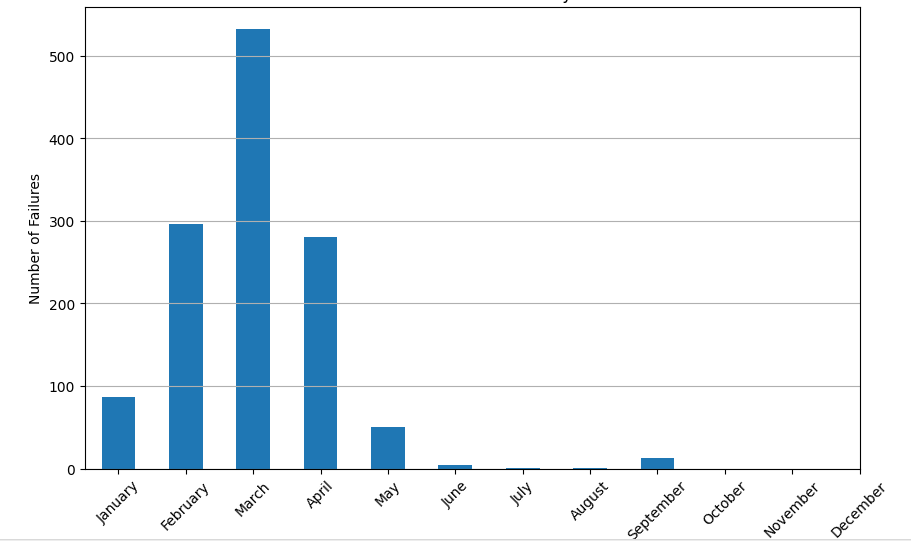


* Mean values:

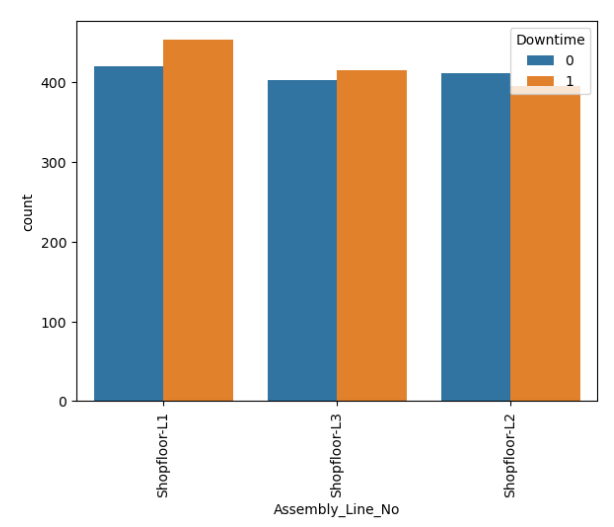
|  |  |
| --- | --- |
| Hydraulic\_Pressure(bar) | 101.409084 |
| Coolant\_Pressure(bar) | 4.947058 |
| Air\_System\_Pressure(bar) | 6.499275 |
| Coolant\_Temperature | 18.559887 |
| Hydraulic\_Oil\_Temperature(Â°C) | 47.618317 |
| Spindle\_Bearing\_Temperature(Â°C) | 35.063698 |
| Spindle\_Vibration(Âµm) | 1.009334 |
| Tool\_Vibration(Âµm) | 25.411975 |
| Spindle\_Speed(RPM) | 20274.792 |
| Voltage(volts) | 348.99679 |
| Torque(Nm) | 25.234968 |
| Cutting(kN) | 2.782551 |
|  |  |

**Business Insights**

* March month has the highest machine failures compared to remaining months and February and march months also have considerable failures. This is due to may be temperature difference in environment and usage time of machine.



* shopfloor-L1' type of machine no is mostly used. when compare to others types. This is may be a fault in manufacturing.



## Conclusion

It is very difficult to analyze the machine downtime which is having more number of records even after aggregating the information. Aggregated numerical data is good to some extent but doesn’t give the connecting information and doesn’t give the flexibility of multiple view of the data. Visualizing this data with various parameters will give good insights quickly and can help taking actions to improve the performance.

With the latest machine tracking technology where each of the machine parameters are tracked, we will have a good opportunity to build visualization to analyze ever further and get more insights on machine performance.