**Introduction:**

This report presents a detailed Exploratory Data Analysis (EDA) and Descriptive Statistical insights derived from machine downtime data. The primary goal is to understand the behaviour of various machine parameters under different conditions, particularly when a machine experiences failure versus when it functions normally.

**Overall Design Strategy:**

The analysis was structured into four major phases:

1. **First Movement - Descriptive Central Tendency Analysis:**
   * Computation of mean, median, and mode to understand the centre of the data distribution.
2. **Second Movement - Variability Analysis:**
   * Assessment of standard deviation, variance, and range to measure spread and dispersion.
3. **Third Movement - Distribution Visualization:**
   * Histogram plots to examine the distribution pattern of each equipment parameter.
4. **Fourth Movement - Higher-Order Statistics:**
   * Kurtosis measurement to assess the tailedness of each distribution.

Further, a comparative analysis of machine condition under failure vs. non-failure was done across **mean, median, and max** values.

**Data Overview:**

* The dataset was cleaned by removing missing values.
* It contains 16 columns and multiple rows of time-series based machine data.
* Each row represents a machine's performance reading at a given time.

**Dataset Columns:**

1. **Date** – Timestamp of the recorded data
2. **Shift** – Production shift (A/B/C)
3. **Operator\_ID** – Identifier for the machine operator
4. **Coolant\_Temperature** – Temperature of the coolant in Celsius
5. **Tool\_Wear** – Duration of tool use in minutes
6. **Feed\_Rate** – Feed rate in mm/rev
7. **Cutting\_Speed** – Speed of the cutting tool in m/min
8. **Depth\_of\_Cut** – Depth in mm
9. **Width\_of\_Cut** – Width in mm
10. **Vibration\_Level** – Vibration magnitude
11. **Power\_Consumption** – Power used in kW
12. **Machine\_Load** – Percentage of machine load
13. **Ambient\_Temperature** – Room temperature in Celsius
14. **Humidity\_Level** – Ambient humidity percentage
15. **Spindle\_Speed\_RPM** – Speed of spindle rotation
16. **Downtime** – Categorical column indicating "Machine\_Failure" or "No\_Machine\_Failure"

The **Spindle\_Speed\_RPM** column was dropped because it contained outlier data which could distort the visualization and interpretation of other attributes. Additionally, Spindle\_Speed\_RPM analysis for both machine failure and non-failure conditions was conducted separately.

**Summary Statistics:**

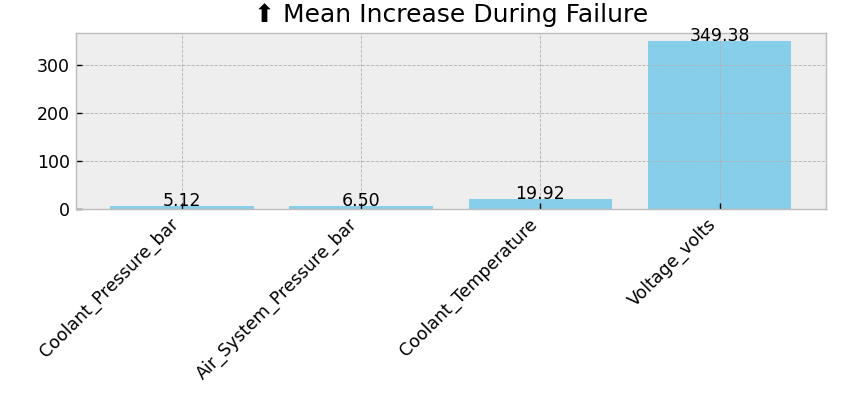
* **Mean** and **Median** were computed separately for Machine\_Failure and No\_Machine\_Failure scenarios.
* The **difference** between them helped identify which features tend to increase or decrease during failure.
* **Histograms** provided an overview of distribution shapes for all operational parameters.
* **Kurtosis** indicated the presence of outliers or sharp peaks in distribution.

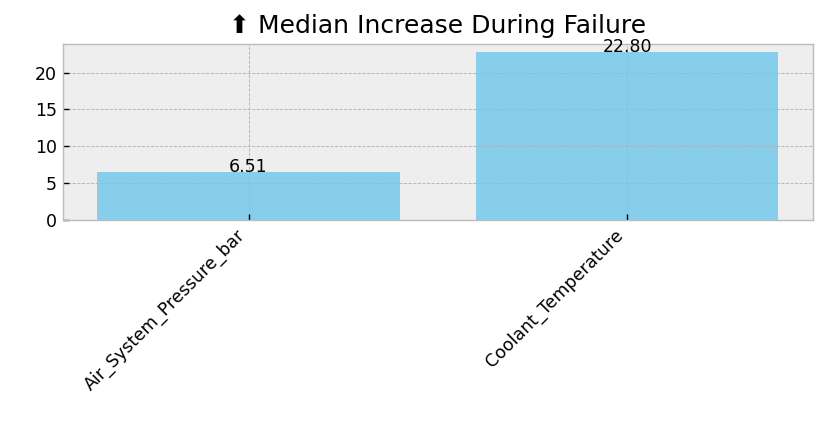
**Key Findings:**

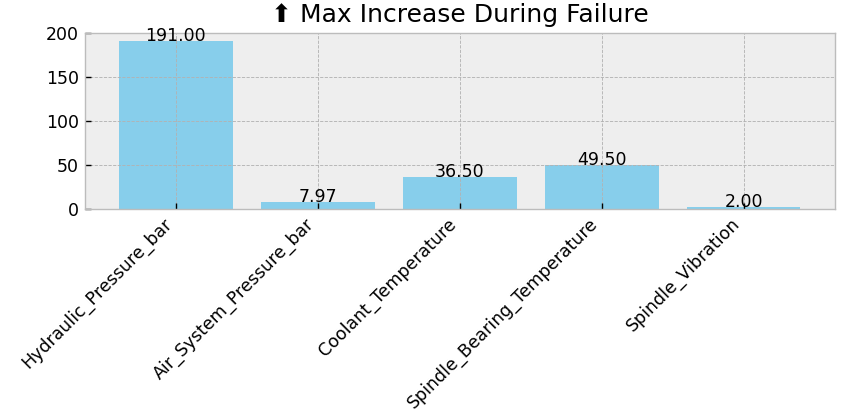
* Several parameters showed **consistent increases or decreases** in value when the machine experienced downtime.
* These findings were visualized using bar charts, separately highlighting:
  + Good vs. Bad Machine Condition
  + Mean/Median/Max values
  + Parameters that increase or decrease during failure

**Common Attribute Change Summary:**

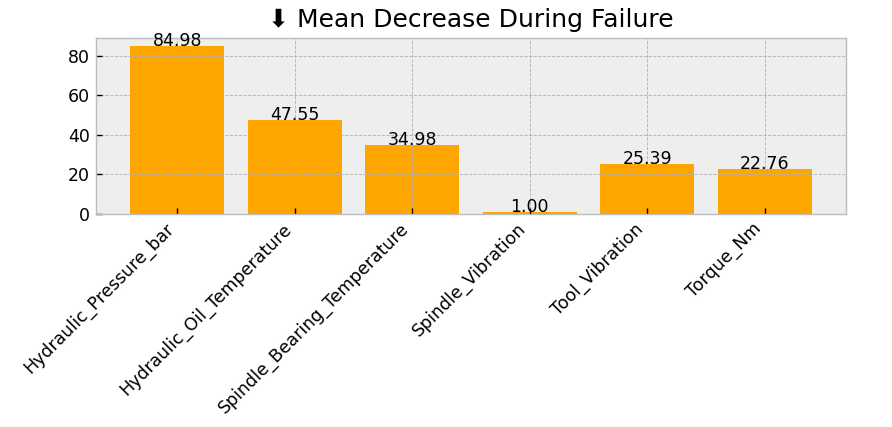
* **Consistently Increased Parameters During Failure:**

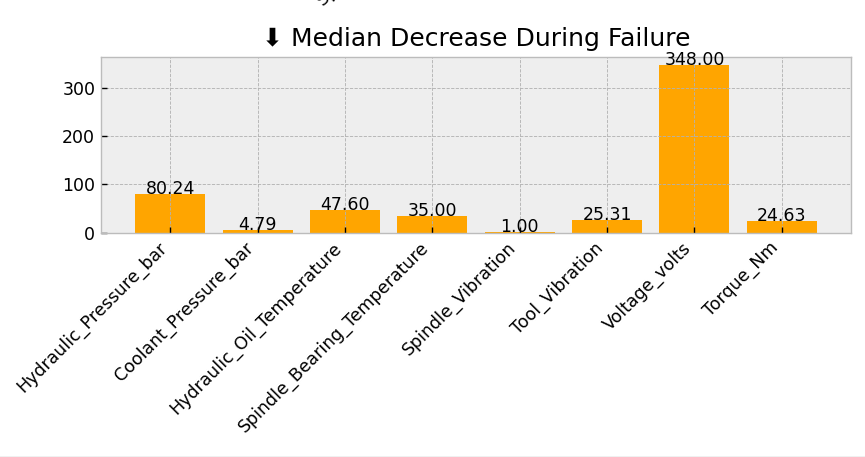


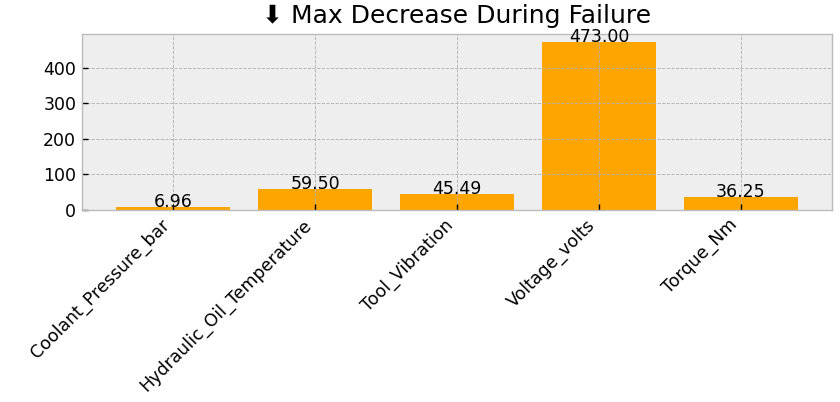




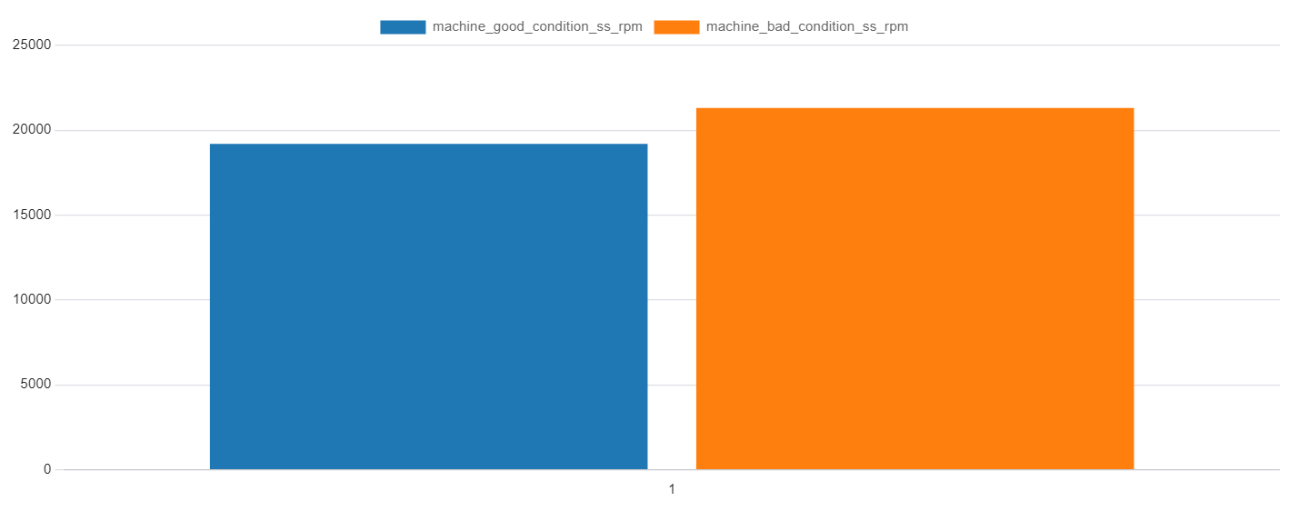
* **Consistently Decreased Parameters During Failure:**







**Spindel Speed Analysis:**

****

**Conclusion:**

This EDA helps in pinpointing key operational parameters that are directly or inversely associated with machine failure. It empowers decision-makers to:

* Set thresholds for early failure detection
* Optimize parameters to prevent downtime
* Understand underlying patterns within machine behaviour

The use of multiple statistical perspectives (mean, median, max) enriches the insight and supports robust business decision-making in predictive maintenance and manufacturing optimization.