1. UDI (Unique Device Identifier):

- A unique identifier for each individual piece of equipment or component. This helps in tracking and distinguishing different units in the dataset.

2. Product ID:

- An identifier for the specific product or item being produced by the machine. It helps in linking the machine's operational data to the particular product it was working on.

3. Type:

- The type of machine or equipment. This could categorize machines by model, function, or any other classification relevant to the dataset. consisting of a letter L, M, or H for low (50% of all products), medium (30%) and high (20%) as product quality variants and a variant-specific serial number.

4. Air temperature [K]:

- The ambient air temperature around the machine, measured in Kelvin (K). This can affect the machine's performance and is a critical factor in predictive maintenance.

5. Process temperature [K]:

- The temperature within the machine’s process, also measured in Kelvin (K). This is crucial for understanding the operating conditions of the machine and detecting potential overheating issues.

6. Rotational speed [rpm]:

- The speed at which the machine’s components are rotating, measured in revolutions per minute (rpm). This metric is important for assessing the machine's operational load and wear.

calculated from a power of 2860 W, overlaid with a normally distributed noise

7. Torque [Nm]:

- The amount of torque being applied by the machine, measured in Newton-meters (Nm). Torque is a key factor in understanding the machine's operational strain and potential failure points.

8. Tool wear [min]:

- The amount of wear experienced by the machine's tool, measured in minutes. This helps in predicting when a tool might need to be replaced or serviced. The quality variants H/M/L add 5/3/2 minutes of tool wear to the used tool in the process.

9. Machine failure:

- A binary indicator (0 or 1) showing whether the machine has failed (1) or not (0). This is the target variable for predictive maintenance models.

**The machine failure consists of five independent failure modes**

1. TWF (Tool Wear Failure):

- A binary indicator showing if the failure was due to tool wear. This provides specific information about the cause of the machine failure.

2. HDF (Heat Dissipation Failure):

- A binary indicator showing if the failure was due to issues with heat dissipation. This helps in identifying temperature-related failures.

3. PWF (Power Failure):

- A binary indicator showing if the failure was due to power-related issues. This can help in diagnosing failures related to electrical problems.

4. OSF (Overstrain Failure):

- A binary indicator showing if the failure was due to overstrain. This indicates failures caused by mechanical overloading or excessive force.

5. RNF (Random Failure):

- A binary indicator showing if the failure was due to random, unforeseen factors. This helps in categorizing failures that don't fit into the other predefined categories.

Understanding these columns is crucial for building a predictive maintenance model as they represent the key variables that can influence machine performance and potential failure.