Peculiarities of the given dataset are mentioned below:

- The original data set consist of 2519 rows and 8 columns
- There was one observation missing in column “part_id”
- If we see the observation with index 2514, we can observe that the elements of that particalr row are shifted to left side as shown below:

2514	2021-06-17 20:05:36.582425088;ACME123	234	STATUS	OK	234	ACME123	ca6a62c8-1d93-43d6-821d-8e4f80f450ff	NaN

- Unique values in prod_id =[“ACME 123”].
- Unique values of sensor = [“A”, “C”, Tmax”, “STATUS”, “Rückführkrft ballist.”, “r_null”, “tol”].
- Unique values of station_id = [234, 235].
- Unique values of part_id = 360 unique part id in complete dataset.
- There are total 360 number of unique part_id and each part id have min 4 to 8 sensors coming from two different station id that is station id = 234 and station id = 235.

- Peculiar observation between column “station_id” and “sensors”.

There are two station id as 234 and 235

1) The station id = 234, consist of total four categories of sensors as [“A”, “C”, Tmax”, “STATUS”].
   That means, If sation id == 234, then sensors = [“A”, “C”, Tmax”, “STATUS”].

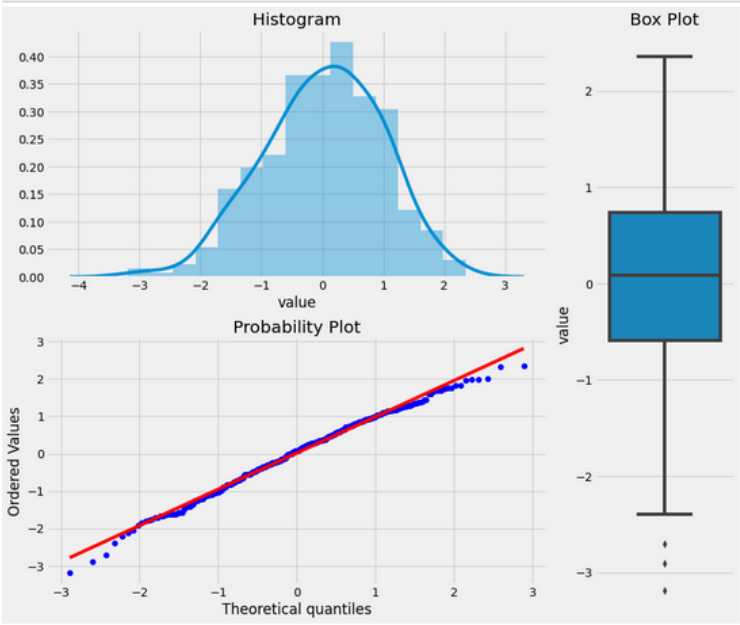
2) Similary the station id = 235, consist of total four categories of sensors as [“STATUS”, “Rückführkrft ballist.”, “r_null”, “tol”].
   That means, If sation id == 235, then sensors = [“STATUS”, “Rückführkrft ballist.”, “r_null”, “tol”].

- The data types for original dataset:
  timestamp : object
  prod_id : object
  station_id : object
  sensor: object
  value : object
  station_id.1 : object
  prodid : object
  part_id : object

- Columns like "station_id" vs "station_id.1" are similar.
- Columns "prod_id" vs "prodid" are similar.

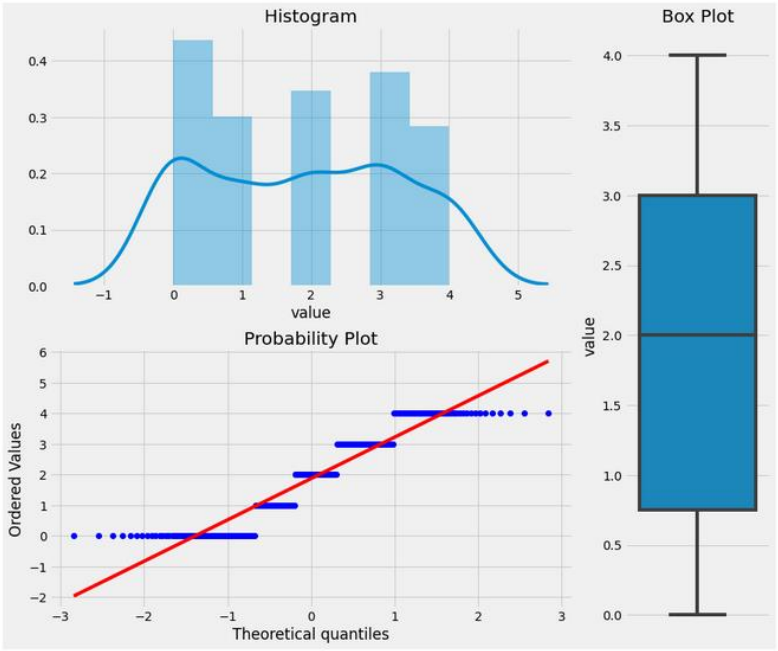
- In depth analysis between the columns “value” and “sensor”.

1) When the sensor = “A”, the column “value” consist of float numbers that lie between -4 to +3, and their probability density distribution is as follows:



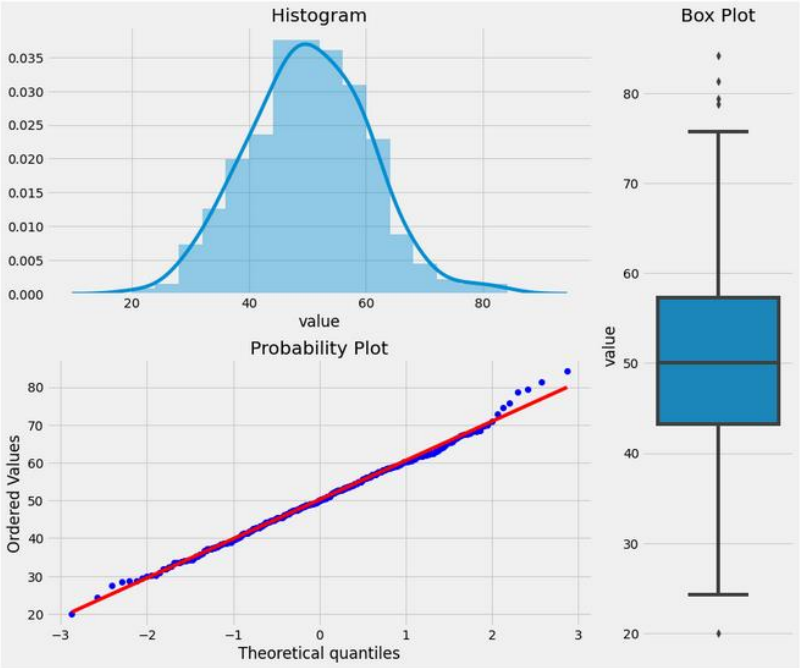
When sensor = “A”, then the numbers in column “value” are float numbers between -4 to +3, with the following distribution.

2) When the sensor = “C”, the column “value” consist of absolute integer numbers that lie between 0 to 4 (all are absolute integers), and their probability density distribution is as follows:



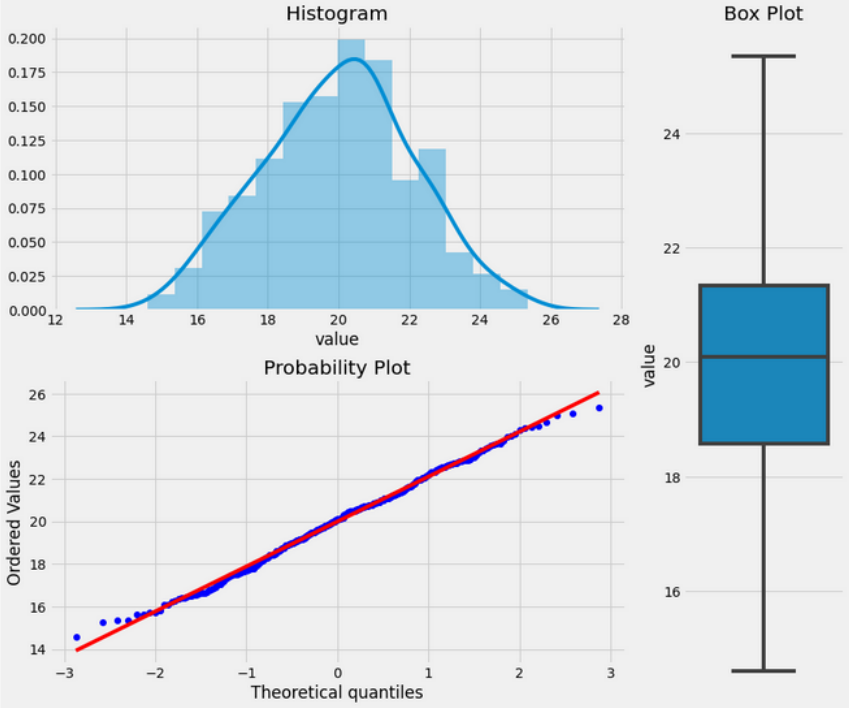
When sensor = “C”, then the numbers in column “value” are integer numbers between [0, 1, 2, 3, 4] (all are absolute integers), with the following distribution.

3) When the sensor = “T_max”, the column “value” consist of float numbers that lie between +20 to +80, and their probability density distribution is as follows:



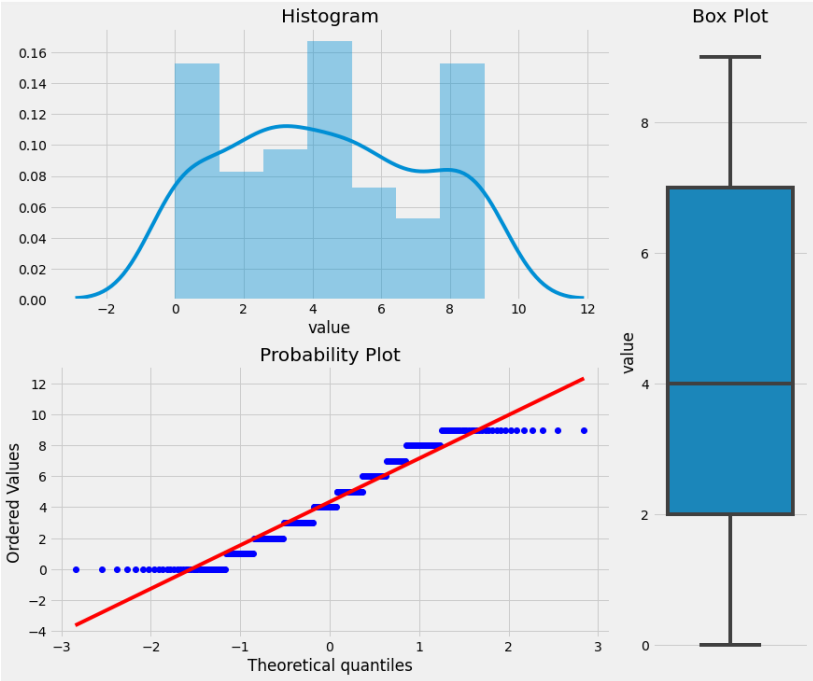
When sensor = “T_max ”, then the numbers in column “value” are float numbers between +20 to +80, with the following distribution.

4) When the sensor = “Rückführkrft ballist.”, the column “value” consist of float numbers that lie between +14 to +26, and their probability density distribution is as follows:



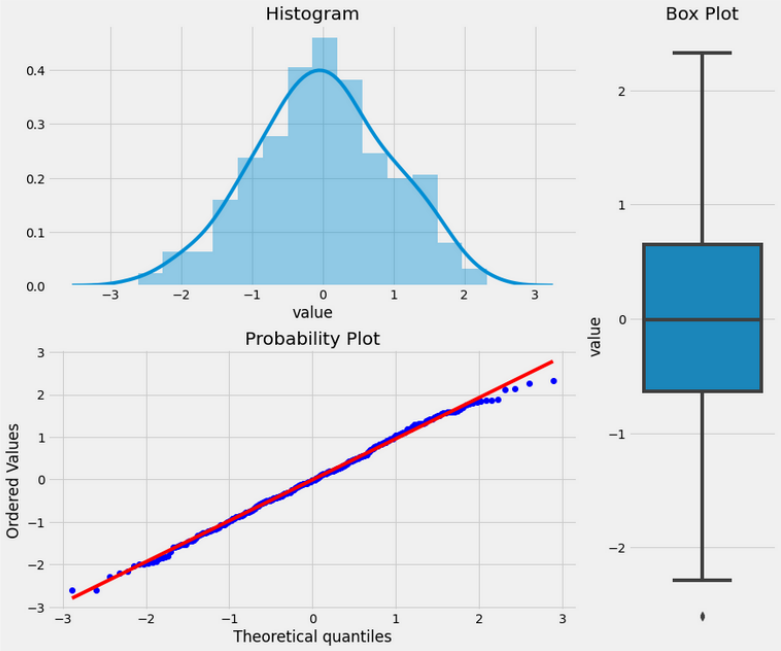
When sensor = “Rückführkrft ballist.”, then the numbers in column “value” are float numbers between +14 to +26, with the following distribution.

5) When the sensor = “r_null”, the column “value” consist of absolute integer numbers that lie between 0 to 9 (all are absolute integers), and their probability density distribution is as follows:



When sensor = “r_null”, then the numbers in column “value” are integer numbers between [0, 1, 2, 3, 4, 5, 6, 7, 8, 9] (all are absolute integers), with the following distribution.

6) When the sensor = “tol”, the column “value” consist of float numbers that lie between -3 to +3, and their probability density distribution is as follows:



When sensor = “tol”, then the numbers in column “value” are float numbers between -3 to +3, with the following distribution.

7) When sensor = “STATUS”, then the strings in column “value” are “OK” and “NOT OK”.