Data Dimension

Columns: 28

Rows: 5505

Data Dictionary

Date & Time: The data and time of recorded data.

Mill TPH: It show Mill in Ton per Hour.

Clinker TPH: It show Clinker in Ton per Hour.

Gypsum TPH: It show Gypsum in Ton per Hour.

DFA TPH: It show Dried Fine Aggregates in Ton per Hour.

WFA TPH: It show Wed Fine Aggregates in Ton per Hour.

Mill KW: The power consumption of the mill in Kilowatts.

Mill I/L Temp: The Temperature of mill inlet.

Mill O/L Temp: The temperature of mill outlet.

Mill O/L BE Amp: Mill outlet belt ampere recording.

Mill Vent Fan RPM: The mill ventilation fan in revolution per minutes.

Mill Vent Fan KW: The mill ventilation fan in kilowatts.

Mill Vent BF I/L Draft: The air pressure created at inlet of back filter in mill ventilation system.

Mill Vent BF O/L Draft: The air pressure created at outlet of back filter in mill ventilation system.

Reject: The amount of material reject by machine.

Sep RPM: The speed of separator in revolution per minutes.

Sep KW: The speed of separator in kilowatts.

Sep Amp: The speed of separator in ampere.

CA Fan RPM: The speed of cooler air fan in revolution per minutes.

CA Fan KW: The speed of cooler air fan in kilowatts.

Mill Folaphone: It is a sensor record sound level of the mill

Mill I/L Draft: The air pressure or airflow at the inlet of mill.

Mill O/L Draft: The air pressure or airflow at the outlet of mill.

Sep. Vent I/L Draft: The air pressure or airflow at the inlet of separator ventilation.

Sep. Vent O/L Draft: The air pressure or airflow at the outlet of separator ventilation.

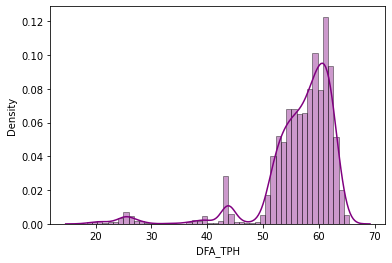
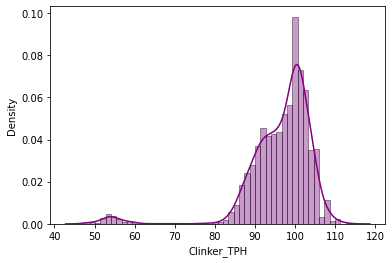
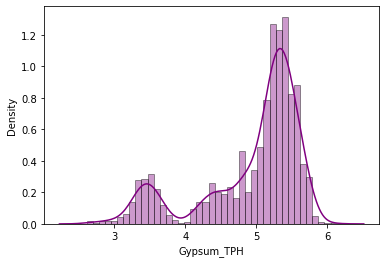
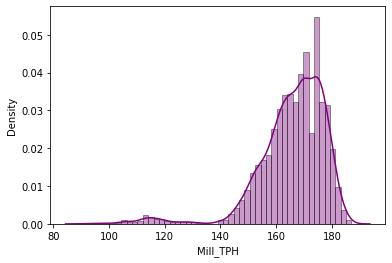
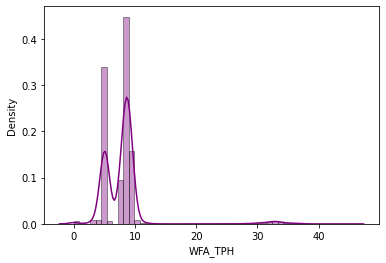
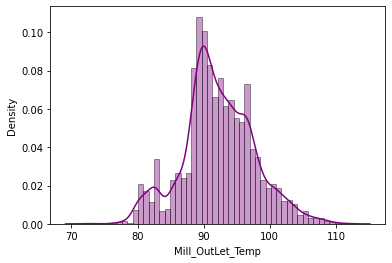
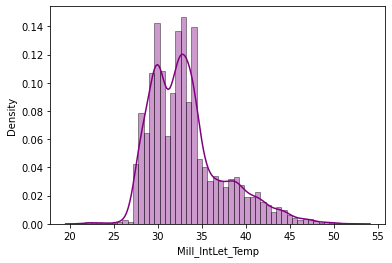
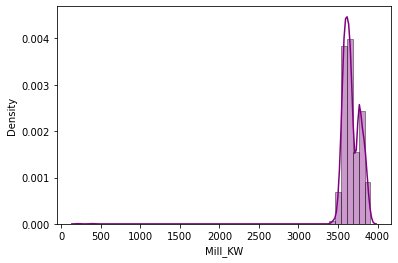
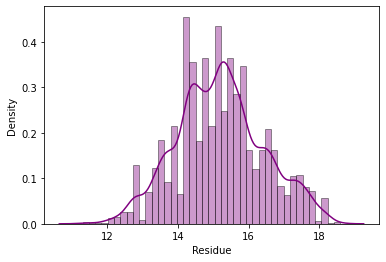
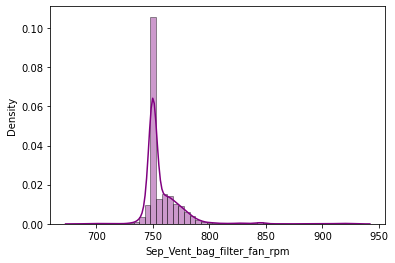
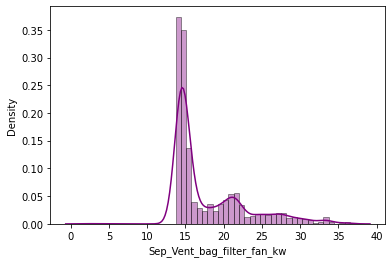
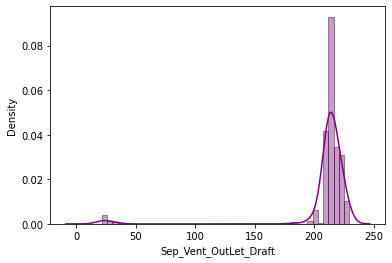
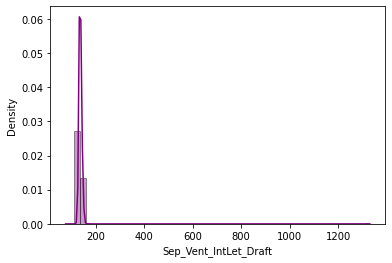
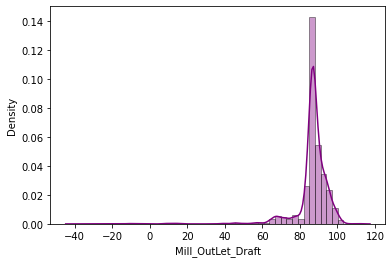
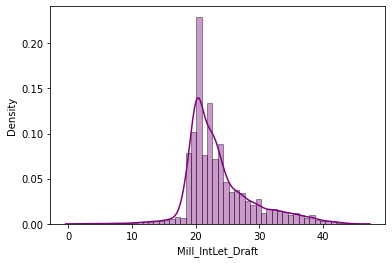
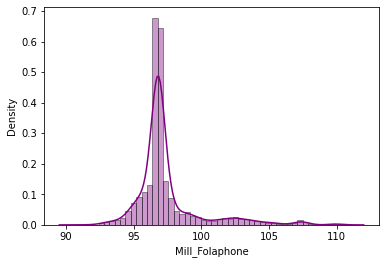
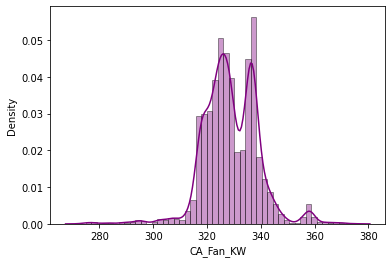
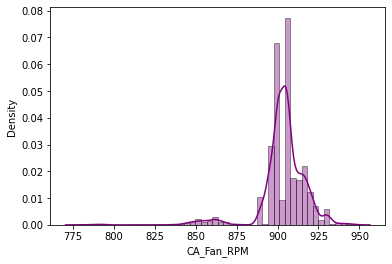
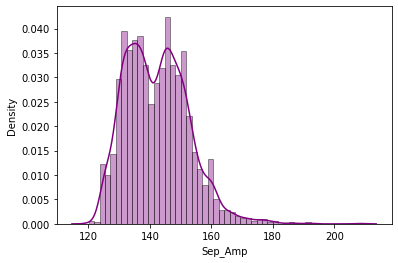
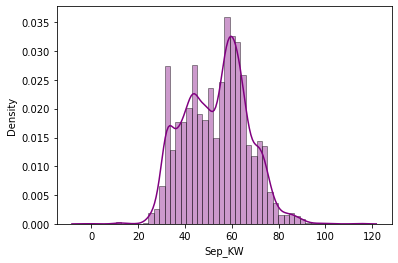
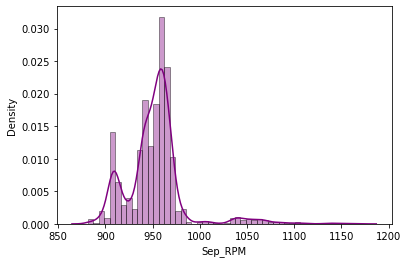
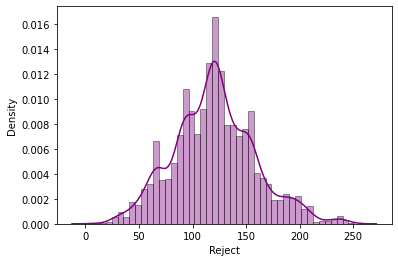
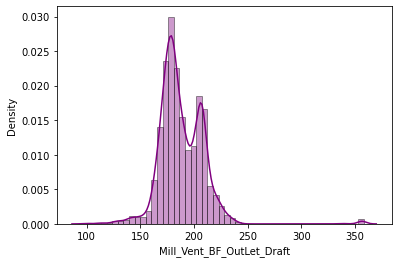
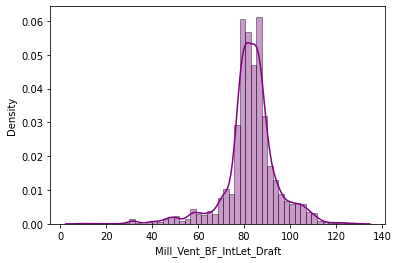
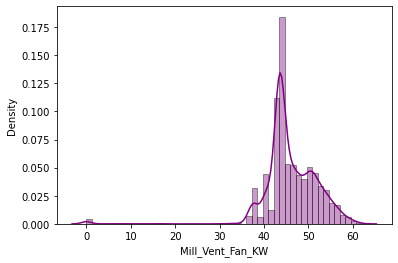
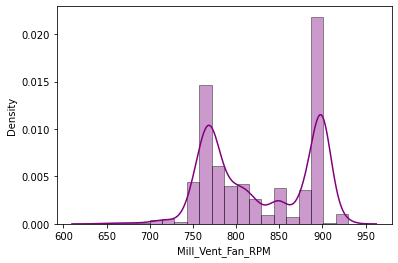
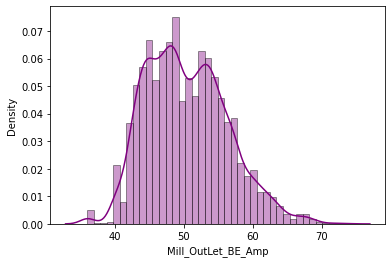
Sep. Vent bag filter fan KW: The power consumption of the separator ventilation bag filter fan in kilowatts.

Sep. Vent bag filter fan rpm: The power consumption of the separator ventilation bag filter fan in resolution per minutes.

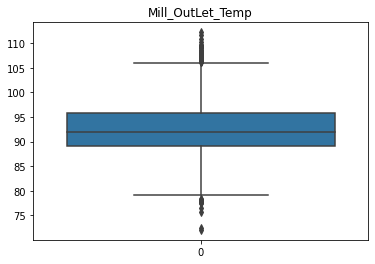
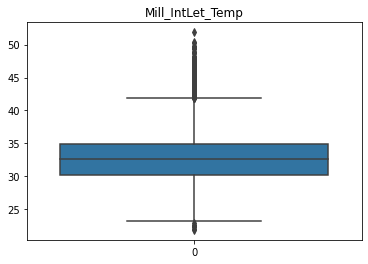
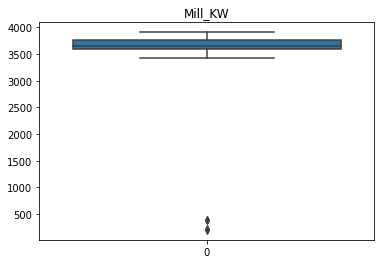
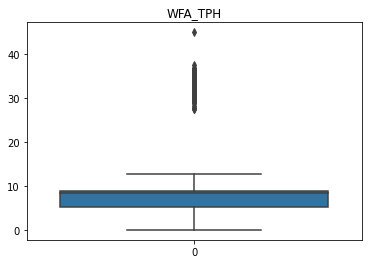
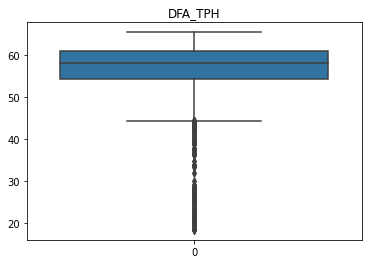
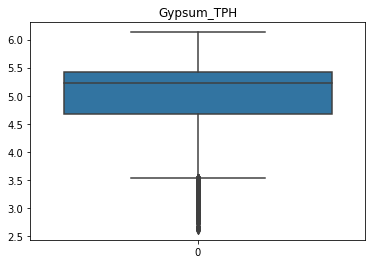
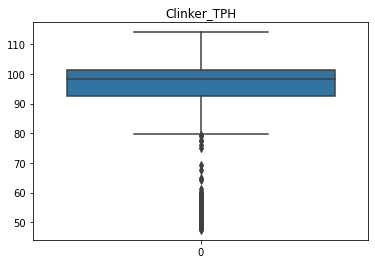
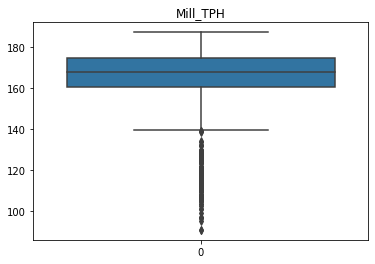
Residue: The size of solid particle is 45 micrometer.

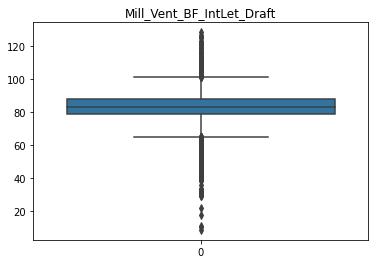
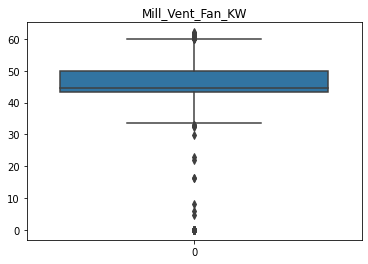
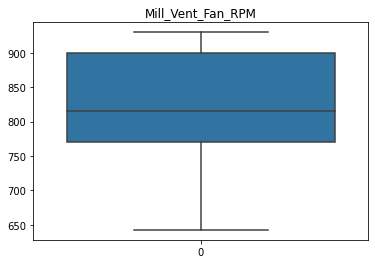
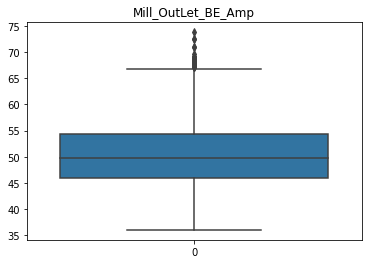
[Districtive Statistic.xlsx](Districtive%20Statistic.xlsx)

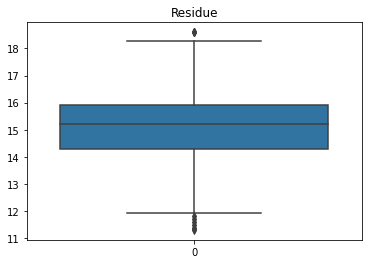
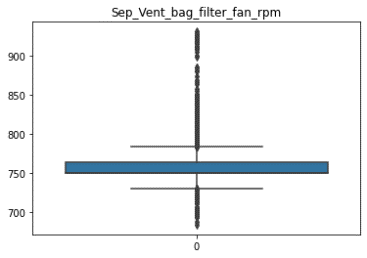
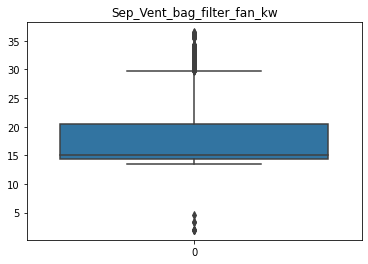
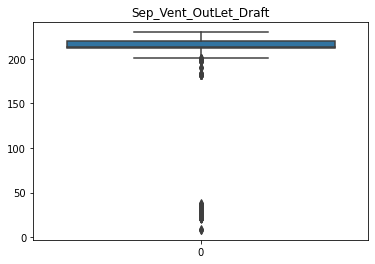
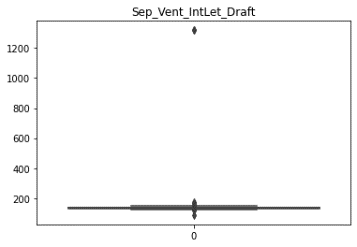
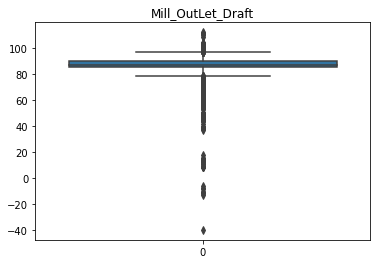
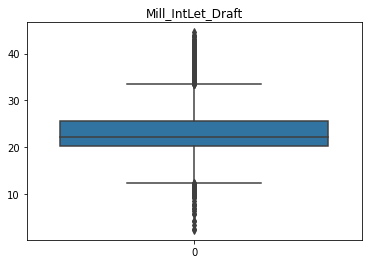
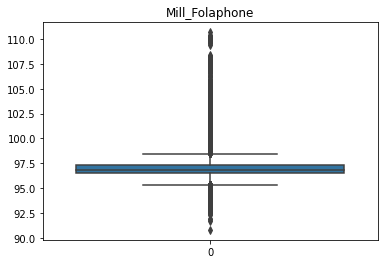
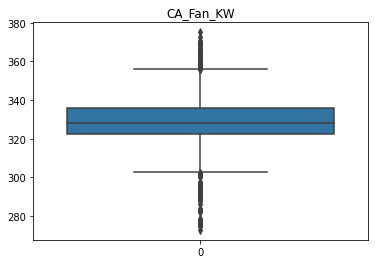
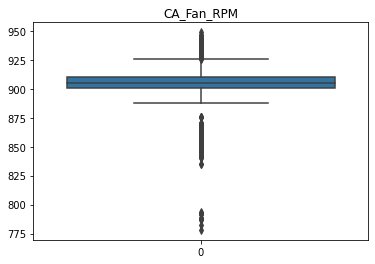
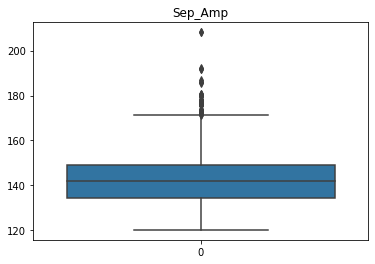
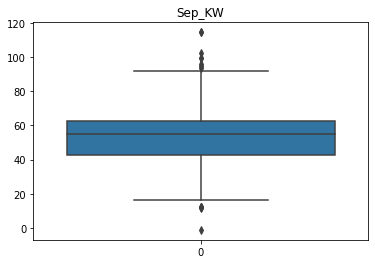
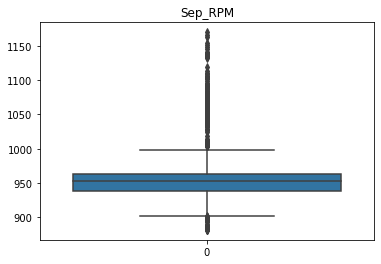
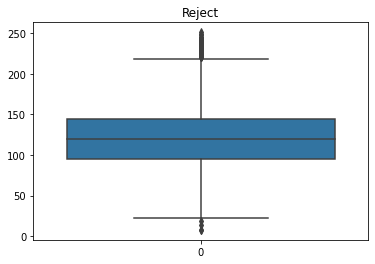
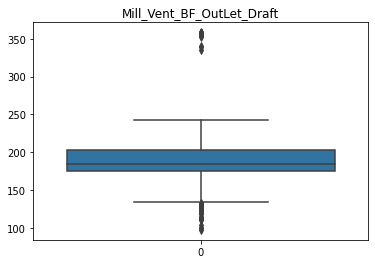
Histogram

Boxplot







Insight

**Mill TPH:**

**Shape of distribution**- Left-skewed, skewness = -1.97, data are negatively Skewed. It indicate most of the data are on higher value lie from 160 to 174.

Median>mean

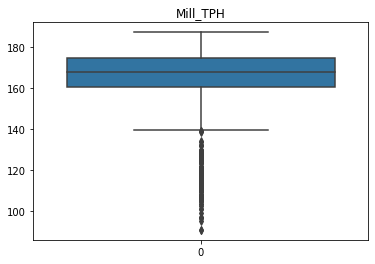
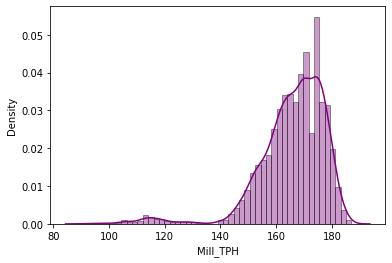
167.95>165.77(Mean is influence by lower value)

**Range**- Range from 90.98 to 187.04, SD = 12.37, variance = 153.04. (It indicate data has moderate variance, fluctuation is low.)

**Peak**- Sharp peak and long tail, Kurtosis = 6.45. It indicate that most of the outlier have moderate impact on the overall distribution and it lie on the left side from 90 to 160.

**Outliers**- 152 outlier, 2.16%

Pre-processing Technique – Data Transformation (Box-Cox transformation)



**Clinker TPH:**

**Shape of distribution**- Left-skewed, skewness = -2.77, data are negatively Skewed. It indicate most of the data are on higher value lie from 85 to 110.

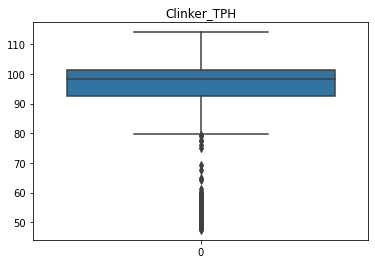
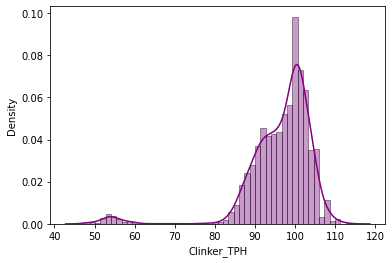
Median>mean

98.23 > 96.26 (Mean is lesser than median, so it indicate the data is left skewed)

**Range**- Range from 47.44 to 114.04, SD = 8.72, variance = 76.11(It indicate data has moderate variance, means fluctuation is moderate.)

**Peak**- Sharp peak and long tail, Kurtosis = 11.14. It indicate that most of the outlier have strong impact on the overall distribution and it lie on the left side from 47 to 80.

**Outliers**- 146 outlier, 2.08% Data Transformation (Box-Cox transformation)



**Gypsum TPH:**

**Shape of distribution**- Left-skewed, skewness = -1.26, data are negatively Skewed. It indicate most of the data are on higher value lie from 4 to 5.5.

Median>mean

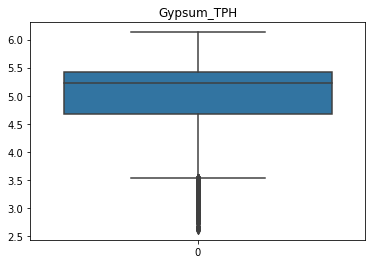
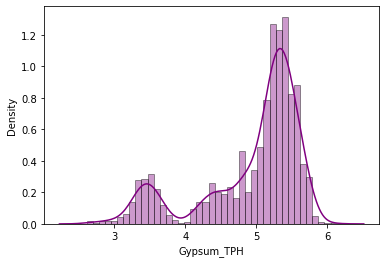
5.22> 4.93(Here median and mode are slightly same. Mean is influence by outlier)

**Range**- Range from 2.61 to 6.13, SD = 0.712, variance = 0.50(It indicate data has low variance, means fluctuation is moderate.)

**Peak**- Sharp peak and slightly long tail, Kurtosis = 0.49 It indicate that most of the outlier have moderate impact on the overall distribution and it lie on the left side from 3 to 4.

**Outliers**- 539 outlier, 7.66%

Here we can see two peak called bimodal which indicate the present of cluster.



**DFA TPH**

**Shape of distribution**- Left-skewed, skewness = -2.69, data are negatively Skewed. It indicate most of the data are on higher value lie from 50 to 67.

Median>mean

58.13 > 56.54 (median is greater than mean it means data is left skewed, mean is influence by outlier)

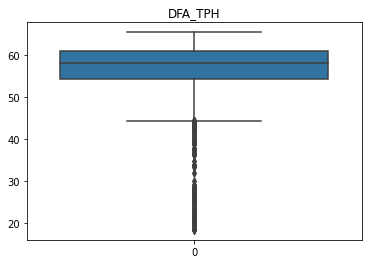
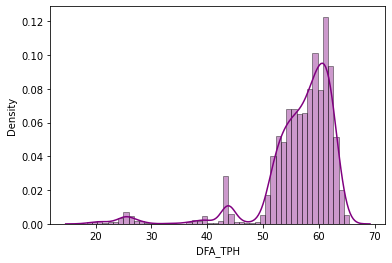
**Range**- Range from 18.35 to 65.46, SD = 6.97, variance = 48.61, (It indicate data has moderate variance, means fluctuation is moderate.)

**Presence of Missing value:** 1, Median imputation

**Peak**- Sharp peak and long tail, Kurtosis = 9.30. It indicate that most of the outlier have strong impact on the overall distribution and it lie on the left side from 20 to 45.

**Missing value:** 1 , **Treatment:** Median Imputation

**Outliers**- 382 outlier



**WFA TPH**

**Shape of distribution**- Right-skewed, skewness = 4.47, data are positively Skewed. It indicate most of the data are lower value lie from 4 to 11.

Median>mean

8.4>8.03(median is greater than mean, it indicate data is right skewed, mean value is not effected by

outlier because median and mean are slightly same.)

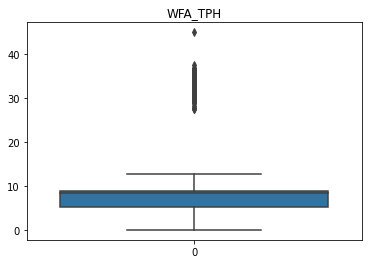
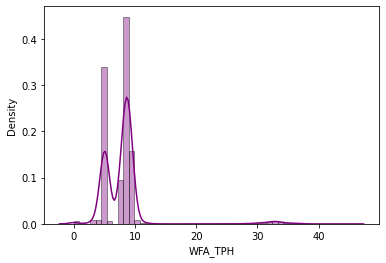
**Range**- Range from 0 to 45.02, SD = 4.35, variance = 18.95 (It indicate data has moderate variance, means fluctuation is moderate.)

**Peak**- Sharp peak and long tail, Kurtosis = 24.16. It indicate that most of the outlier have strong impact on the overall distribution and it lie on the left side from 13 to 40.

**Outliers**- 134 outlier, 1.90%

By looking at graph we can say that is bimodal which indicate the present of cluster.

Winsorization



**Mill KW**

**Shape of distribution**- Left-skewed, skewness = -11.13, data are negatively Skewed. It indicate most of the data are on higher value lie from 3500 to 4000.

Mean>median

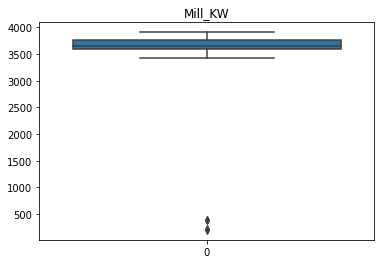
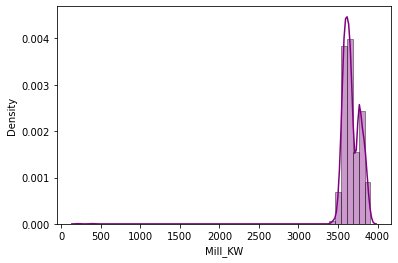
3669.41> 3648.61(Here mean is influence by outliers, because mean is greater than median which indicate that data is right skewed but value of skewness is negative which mean data is left skewed )

**Range**- Range from 203.36 to 3910.5, SD = 135.008, variance = 18227.38(It indicate data has high variance, means fluctuation is very high.)

**Peak**- Sharp peak and long tail on left side, Kurtosis = 281.24. It indicate that most of the outlier have very strong impact on the overall distribution and it lie on right side from 500 to 3480.

**Outliers**- 4 outlier

Here we can see two peak called bimodal which indicate the present of cluster.



**Mill I/L Temp**

**Shape of distribution**- Right-skewed, skewness = 1.06, data are positively Skewed. It indicate most of the data are lower value lie from 27 to 45.

Mean>Median

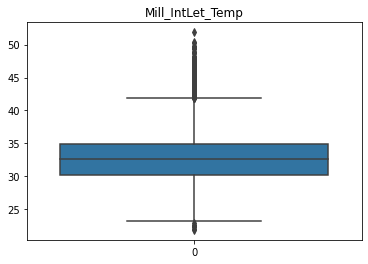
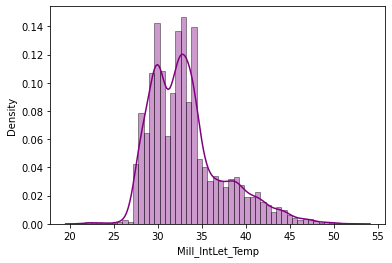
33.25> 32.58 (Mean is greater than median, means data is moderate skewed to right side. Mean is effected by outlier.)

**Range**- Range from 21.8 to 51.89, SD = 4.181, variance = 17.48 (It indicate data has moderate variance, means fluctuation is moderate.)

**Peak**- Sharp peak and slightly long tail, Kurtosis = 1.059. It indicate that most of the outlier have moderate impact on the overall distribution and it lie on right side from 46 to 50.

**Outliers**- 273 outlier, 3.88%

Here we can see two peak called bimodal which indicate the present of cluster.



**Mill O/L Temp**

**Shape of distribution**- NO skewed, skewness = 0.107, the value is approx. equal to zero. Which means data is normally distributed

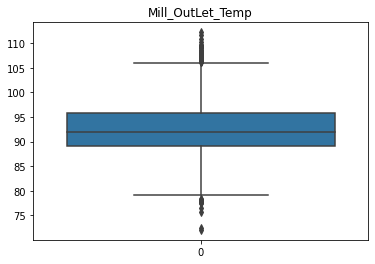
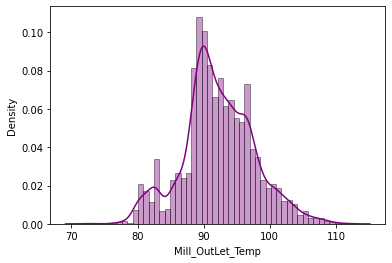
Mean>median (As data is slightly skewed because mean and median are slightly same but

92.18> 91.85 mode value is not same which show that mean is not influence by outliers)

**Range**- Range from 72.04 to 112.22, SD = 72.04, variance = 29.77 (It indicate data has moderate variance, means fluctuation is moderate.)

**Peak**-Data is normally distributed, Kurtosis = 0.297. It indicate that we have very less outliers

**Outliers**- 68 outlier, 0.97%, outlier on both side.



**Mill O/L BE Amp**

**Shape of distribution**- NO skewed, skewness = 0.44, the value is approx. equal to zero. Which means data is normally distributed.

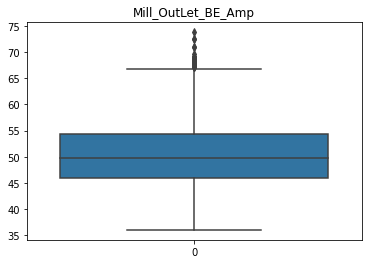
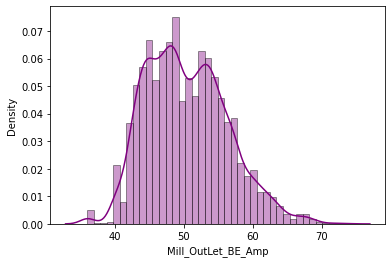
Mean>median (As data is slightly skewed because mean and median are slightly same but

50.40> 49.84 mode value is not same which show that mean is not influence by outliers)

**Range**- Range from 36 to 73.81, SD = 5.85, variance = 34.24(It indicate data has moderate variance, means fluctuation is moderate.)

**Peak**- Wider peak and small tail, Kurtosis = -0.65. It indicate that most of the outlier have less impact on the overall distribution.

**Outliers**- 37 outlier, 0.53%, outlier on right side



**Mill Vent Fan RPM**

**Shape of distribution**- No skewed, skewness = 0.0053, Data is normally distributed due to skewness is equal to 0.

Mean>mean (The value of skewness is zero but mean and median value are not equal, it

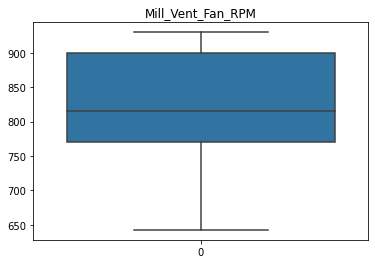
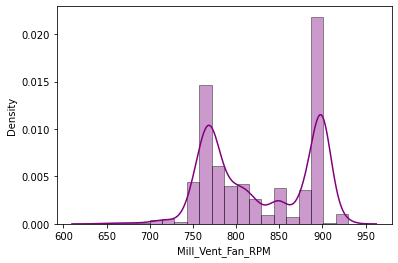
827.84> 814.67 suggest that the data is symmetric but not necessarily normally distributed)

**Range**- Range from 930.01 to 642.37, SD = 60.514, variance = 3662(It indicate data has high variance, means fluctuation is high.)

**Peak**- Wider peak and lighter tail, Kurtosis = -1.464. The value of kurtosis is negative but zero outlier, which indicate that it has lighter tail and less peak as compare to normal distribution.

**Outliers**- 0 outlier, 0%, no outlier.

Here we can see two peak called bimodal which indicate the present of cluster.



**Mill Vent Fan KW**

**Shape of distribution**- Left-skewed, skewness = -2.23, data are negatively Skewed. It indicate most of the data are on higher value lie from 35to 60.

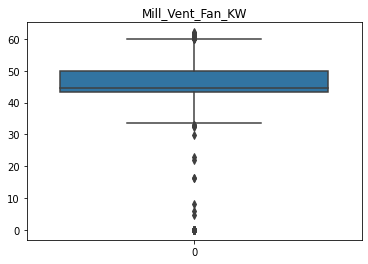
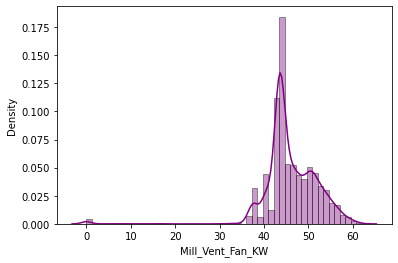
Median>mean

44.64> 46.06

**Range**- Range from 0 to 62.09, SD = 6.08, variance = 36.97(It indicate data has moderate variance, means fluctuation is moderate.)

**Peak**- Sharp peak and long tail, Kurtosis = 17.54. It indicate that data has outlier on right side from 0 to 30 and may have moderate impact of outlier on the overall distribution.

**Outliers**- 62 outlier, 0.88%, outlier on both side.



**Mill Vent BF I/L Draft**

**Shape of distribution**- Left-skewed, skewness = -0.89, data are negatively Skewed or the value is approx. equal to zero. Which means data is normally distributed.

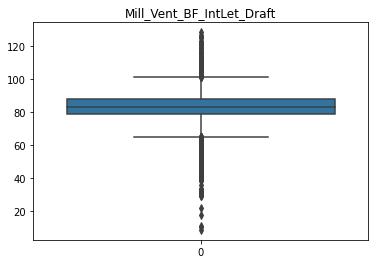
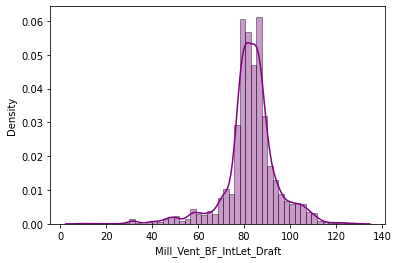
Median>mean (The mean, median and mode value is not equal it mean data is not normally

83.11> 82.96 distributed)

**Range**- Range from 8.55 to 128.51, SD = 11.59, variance = 134.53(It indicate data has moderate variance, means fluctuation is moderate.)

**Peak**- Sharp peak and long tail, Kurtosis = 4.452. It indicate that data has outlier on right side from 20 to 60 and have moderate impact of outlier on the overall distribution.

**Outliers**- 642 outlier, 9.13%, outlier on both side.



**Mill Vent BF O/L Draft**

**Shape of distribution**- Right-skewed, skewness = 1.95, data are positively Skewed. It indicate most of the data are lower value lie from 350 to 380.

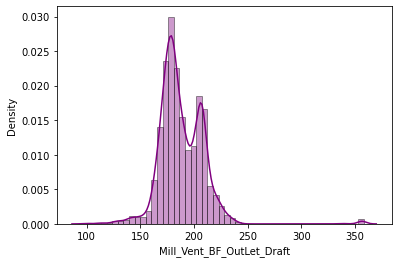
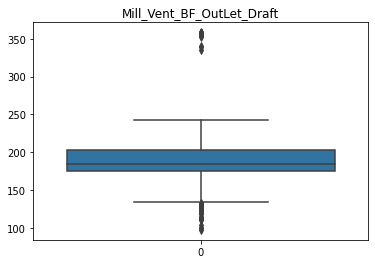
Mean>median

188.23>184.08

**Range**- Range from 97.54 to 358.55, SD = 21.30, variance = 453.77 (It indicate data has high variance, means fluctuation is high.)

**Peak**- Sharp peak and long tail, Kurtosis = 15.16. It indicate the moderate impact of outlier on the overall distribution.

**Outliers**- 61 outlier, 0.87%

**Reject:**

**Shape of distribution**- No skewed, skewness = 0.267. Data is normally distributed due to skewness is equal to 0

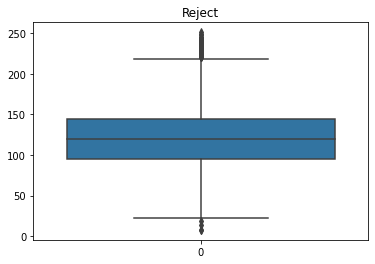
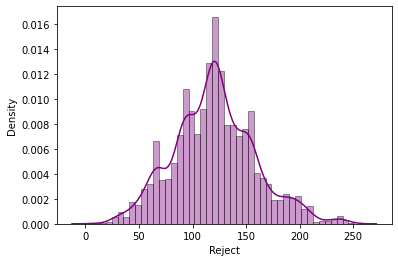
Mean>median (The value of skewness is zero but mean, median and mode value are not equal, it

120.22>119.54 suggest that the data is symmetric but not necessarily normally distributed)

**Range**- Range from 7.81 to 251.16, SD = 38.16, variance = 1456.86(It indicate data has high variance, means fluctuation is high.)

**Peak**- Sharp peak and slightly long tail, Kurtosis = 0.2314. It indicate that very less impact of outlier on the overall data.

**Outliers**- 58 outlier, 0.82%



**Sep RPM:**

**Shape of distribution**- Right-skewed, skewness = 1.86 data are positively Skewed. . It indicate most of the data are lower value lie from 900 to 1000.

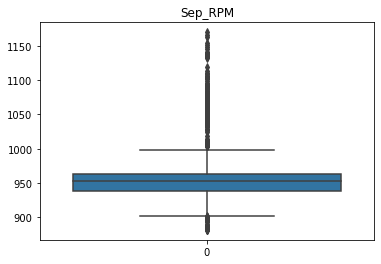
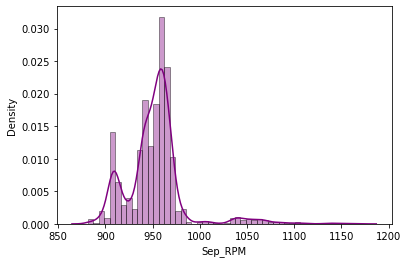
Median>mean (Mean is smaller than median which means mean is influence by outliers)

952.52>952.05

**Range**- Range from 881.43 to 1170, SD = 31.72, variance = 1006.73 (It indicate data has high variance, means fluctuation is high.)

**Peak**- Sharp peak and long tail, Kurtosis = 7.64. It indicate that outlier have strong impact on overall distribution.

**Outliers**- 357 outlier, 5.07%



**Sep KW:**

**Shape of distribution**- No skewed, skewness = 0.065, Data is normally distributed due to skewness is equal to 0.

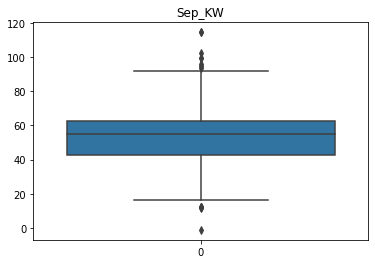
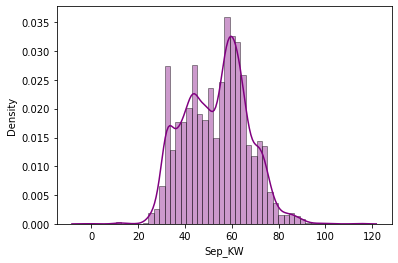
Median>mode (The value of skewness is zero but mean, median and mode value are not equal, it

54.99> 53.38 suggest that the data is symmetric but not necessarily normally distributed)

**Range**- Range from -1.12 to 114.57, SD = 13.42, variance = 180.18 (It indicate data has moderate variance, means fluctuation is moderate.)

**Peak**- Wider peak and small tail, Kurtosis = -0.389. It indicate that outlier have less impact on overall distribution.

**Outliers**- 14 outlier, 0.2%



**Sep Amp:**

**Shape of distribution**- Right-skewed, skewness = 0.683, data are positively Skewed or Data is normally distributed due to skewness approx. equal to 0

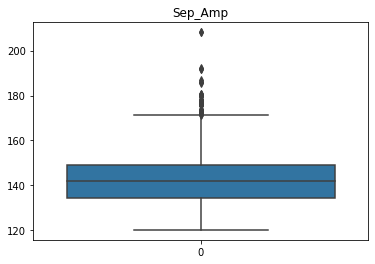
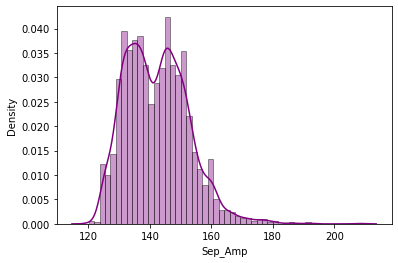
Mean>median

142.30> 141.87

**Range**- Range from 120.12 to 208.19, SD = 10.212, variance = 104.29 (It indicate data has moderate variance, means fluctuation is moderate.)

**Peak**- Sharp peak and slightly longer, Kurtosis = 1.0541. It indicate that outlier have moderate impact on overall distribution.

**Outliers**- 55 outlier, 0.78%



**CA Fan RPM:**

**Shape of distribution**- Left-skewed, skewness = -2.315, data are negatively Skewed

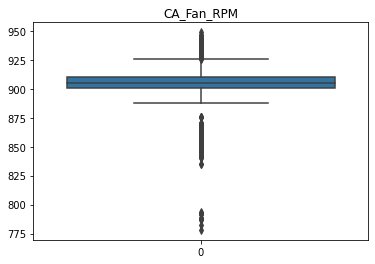
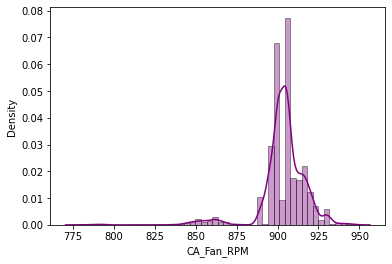
Median>mean

905.29> 903.77

**Range**- Range from 778.15 to 948.77, SD = 14.25, variance = 203.07 (It indicate data has moderate variance, means fluctuation is moderate.)

**Peak**- Sharp peak and longer tail, Kurtosis = 12.69. It indicate that outlier have strong impact on overall distribution.

**Outliers**- 413 outlier, 5.87%



**CA Fan KW:**

**Shape of distribution**- No skewed, skewness = -0.13, data are negatively Skewed or is approx. equal to 0, it mean data is normally distributed.

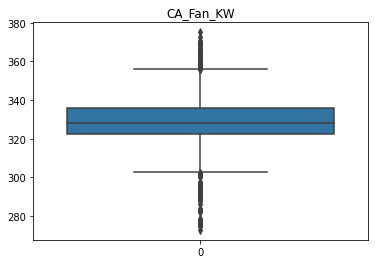
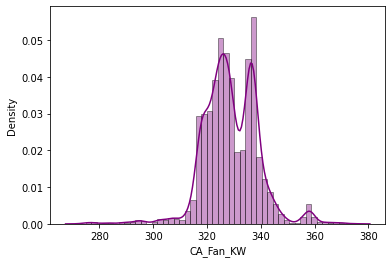
Median>mean (value of mean and median is almost same so we can say that mean is not influence by

327.96>328.89 outliers.)

**Range**- Range from 272.9 to 375.08, SD = 10.26 (It indicate data has moderate variance, means fluctuation is moderate.)

**Peak**- Sharp peak and long tail, Kurtosis = 3.106. It indicate that outlier have moderate impact on overall distribution.

**Outliers**- 181 outlier, 2.57%



**Mill Folaphone:**

**Shape of distribution**- Right-skewed, skewness = 2.28, data are positively Skewed. It indicate most of the data are lower value lie from 90 to 100.

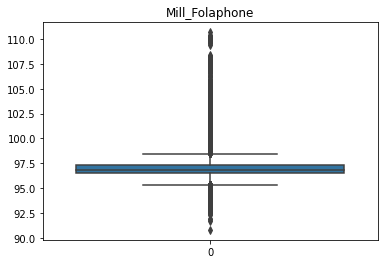
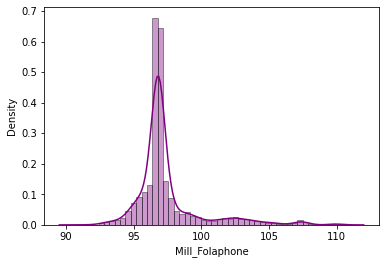
Mean>median

97.42> 96.81

**Range**- Range from 90.98 to 187.04, SD = 2.453, variance = 6.02(It indicate data has low variance, means fluctuation is low.)

**Peak**- Sharp peak and long tail, Kurtosis = 6.356. It indicate that outlier have strong impact on overall distribution.

**Outliers**- 1343 outlier, 19.09%



**Mill I/L Draft:**

**Shape of distribution**- Right-skewed, skewness = 1.098, data are positively Skewed. It indicate most of the data are lower value lie from 18 to 30.

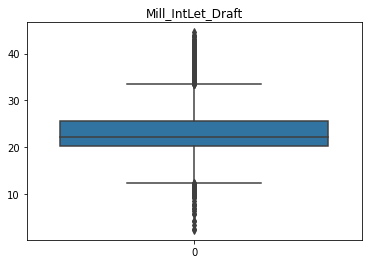
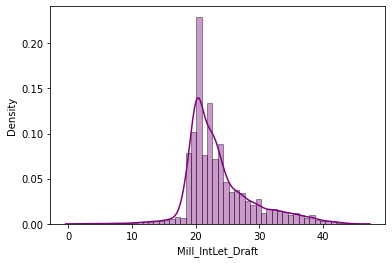
Mean>median

23.60> 22.3

**Range**- Range from 2.37 to 44.57, SD = 5.161, variance = 26.64 (It indicate data has moderate variance, means fluctuation is moderate.)

**Peak**- Sharp peak and moderate long tail, Kurtosis = 1.775. It indicate that outlier have moderate impact on overall distribution.

**Outliers**- 401 outlier, 5.7%



**Mill O/L Draft:**

**Shape of distribution**- Left-skewed, skewness = -4.58, data are negatively Skewed. It indicate most of the data are higher value lie from 80 to 100.

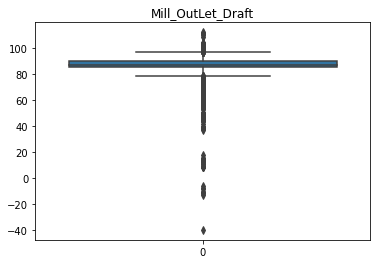
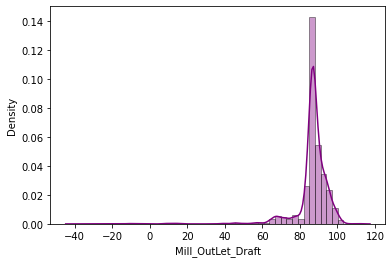
Median>mean

87.35>86.71

**Range**- Range from -39.73 to 112.37, SD = 9.37, variance = 87.83

**Peak**- Sharp peak and longer tail, Kurtosis = 36.612. It indicate that outlier have strong impact on overall distribution.

**Outliers**- 711 outlier, 10.11%



**Sep. Vent I/L Draft:**

**Shape of distribution**- Right-skewed, skewness = 49.81, data are positively Skewed. It indicate most of the data are lower value.

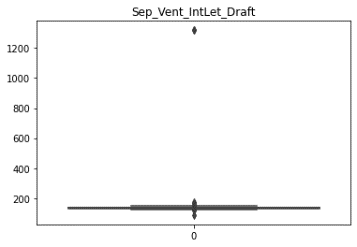
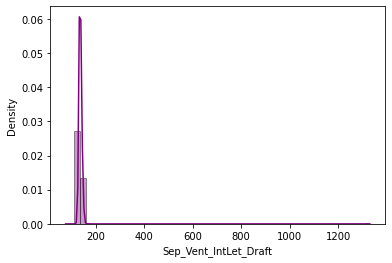
Mean>median

136.32> 134.3

**Range**- Range from 87 to 1319, SD = 22.94, variance = 526.44 (It indicate data has high variance, means fluctuation is very high.)

**Peak**- Sharp peak and longer tail, Kurtosis = 2564.466. It indicate that outlier have very strong impact on overall distribution.

**Outliers**- 117 outlier, 1.66%



**Sep. Vent O/L Draft:**

**Shape of distribution**- Left-skewed, skewness = -5.38, data are negatively Skewed. It indicate most of the data are higher value lie from 200 to 240.

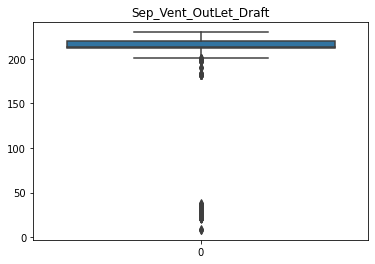
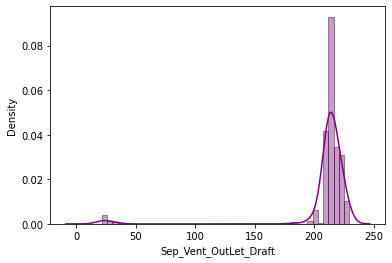
Median> mean

212.8> 209.59

**Range**- Range from 8.16 to 229.69, SD = 32.19, variance = 1036.77 (It indicate data has high variance, means fluctuation is very high.)

**Peak**- Sharp peak and longer tail, Kurtosis = 28.038. It indicate that outlier have very strong impact on overall distribution.

**Outliers**- 321 outlier, 4.56%, data transformation technique (power transform)



**Sep. Vent bag filter fan KW:**

**Shape of distribution**- Right-skewed, skewness = 1.46, data are positively Skewed. It indicate most of the data are lower value lie from 14 to 22.

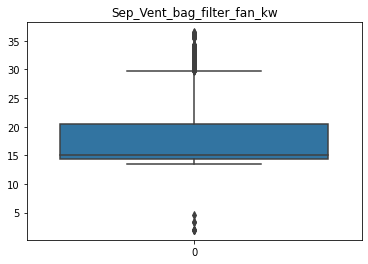
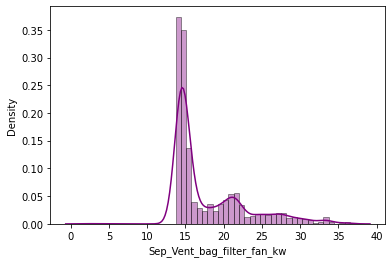
Mean>median

17.78> 15.07

**Range**- Range from 1.98 to 36.5, SD = 4.93, variance = 24.32 (It indicate data has moderate variance, means fluctuation is moderate.)

**Peak**- Sharp peak and long tail, Kurtosis = 1.53. It indicate that outlier have moderate impact on overall distribution.

**Outliers**- 209 outlier, 2.97%, data transformation technique



**Sep. Vent bag filter fan rpm:**

**Shape of distribution**- Right-skewed, skewness = 3.96, data are positively Skewed

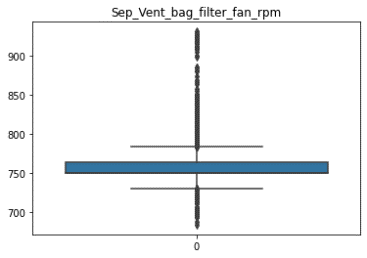
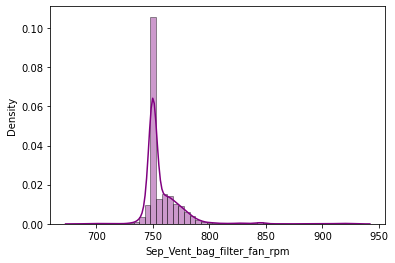
Mean>Median (Mean is greater than median so we can say that data is right skewed

758.46 >750 mean is influence by outliers)

**Range**- Range from 683 to 931, SD = 19.76, variance = 390.51 (It indicate data has high variance, means fluctuation is high.)

**Peak**- Sharp peak and long tail, Kurtosis = 25.21. It indicate that outlier have strong impact on overall distribution.

**Outliers**- 344 outlier, 4.89%, data transformation technique



**Residue:**

**Shape of distribution**-No skewed, Skewness = 0.120. Data is normally distributed due to skewness is equal to 0.

Mean>median (The mean and median value are similar so we can say that our data is

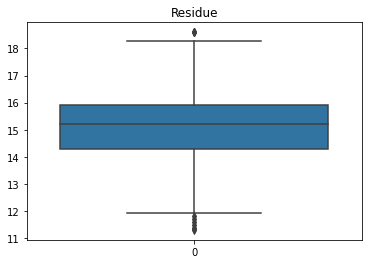
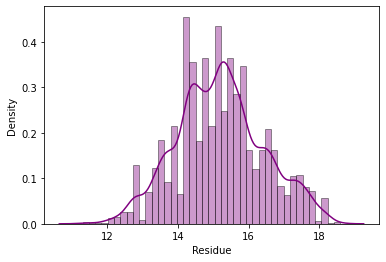
15.19> 15.2 normally distributed)

**Range**- Range from 11.3 to 18.6, SD = 1.23, variance = 1.52(It indicate data have low variance, means fluctuation is low.)

**Peak**- Thick peak and slightly short tail, Kurtosis = -0.24, It indicate that outlier may have less impact on the overall distribution.

**Outliers**- 11 outlier, 0.2%

Winsorization



**QUESTIONS**

1. What is the current power consume by mill and mill ventilation fan in kilowatts? Card
2. What is the current revolution per minute of mill ventilation fan? Card
3. What is relationship between mill inlet temperature and mill outlet temperature? Scatter plot
4. What is different between mill inlet temperature and mill outlet temperature? Histogram
5. What is relationship between residue and mill (ton per hour)? Scatter plot
6. How many percentage of residue data lies between the desire ranges? Card
7. What is relationship between mill inlet draft and mill outlet draft? Scatter plot