



Start looking at the following plots:

• General Understanding of Data

Skewness = 0 → data is approximately symmetric.

Positive Skewness (z > 0) → The distribution has longer tail to the right (more extreme higher values).

Negative Skewness (z < 0) → The distribution has longer tail to the left (more extreme lower values).

The skewness number signifies that how far it is skewed from the normal distribution.

If it is skewed to the left or greater than 1, the distribution is highly left-skewed.

If it is skewed to the right or less than -1, the distribution is highly right-skewed.

If it is between -1 and 1, the distribution is approximately symmetric.

Even though the distributions are slightly right-skewed, there are some extreme values pulling the average upwards.

When mode building interpretations are 12 degrees Celsius energy consumption is minimum mean energy can saved. i.e. in

Even though the distributions are slightly right-skewed, there are some extreme values pulling the average upwards.

Interpretation of Your Data

Low Positive Skewness (z < 0)

temp_10min (z = -0.1)

temp_30min (z = -0.1)

These distributions are nearly symmetric, indicating that the temperature data in these rooms is balanced with minimal skew.

Median Negative Skewness (z > 0.5)

temp_1hr (z = 0.5)

temp_24hr (z = 0.5)

These distributions are slightly right-skewed, indicating that there are some higher temperature values pulling the average upwards.

High Positive Skewness (z > 0)

temp_10min (z = 1.0)

temp_30min (z = 1.0)

The data has a moderate right skew, suggesting some extremely high temperatures in the living room.

Negative Skewness (z < 0)

temp_1hr (z = -0.5)

temp_24hr (z = -0.5)

These distributions are slightly left-skewed, indicating the presence of slightly lower temperature values compared to the rest.

All distributions have similar variability with appliance energy consumption trends to moderately these columns will be dropped in our model.

• Humidity Columns (Entirely Unrelated)

• **Humidity** → Unrelated column

The data is uncorrelated with the other columns, so we can drop this column.

• **Humidity_10min**, **Humidity_30min**, **Humidity_1hr**, **Humidity_24hr**

• **Humidity_10min**, **Humidity_30min**, **Humidity_1hr**, **Humidity_24hr** → Unrelated columns

• **Humidity_10min**, **Humidity_30min**, **Humidity_1hr**, **Humidity_**

