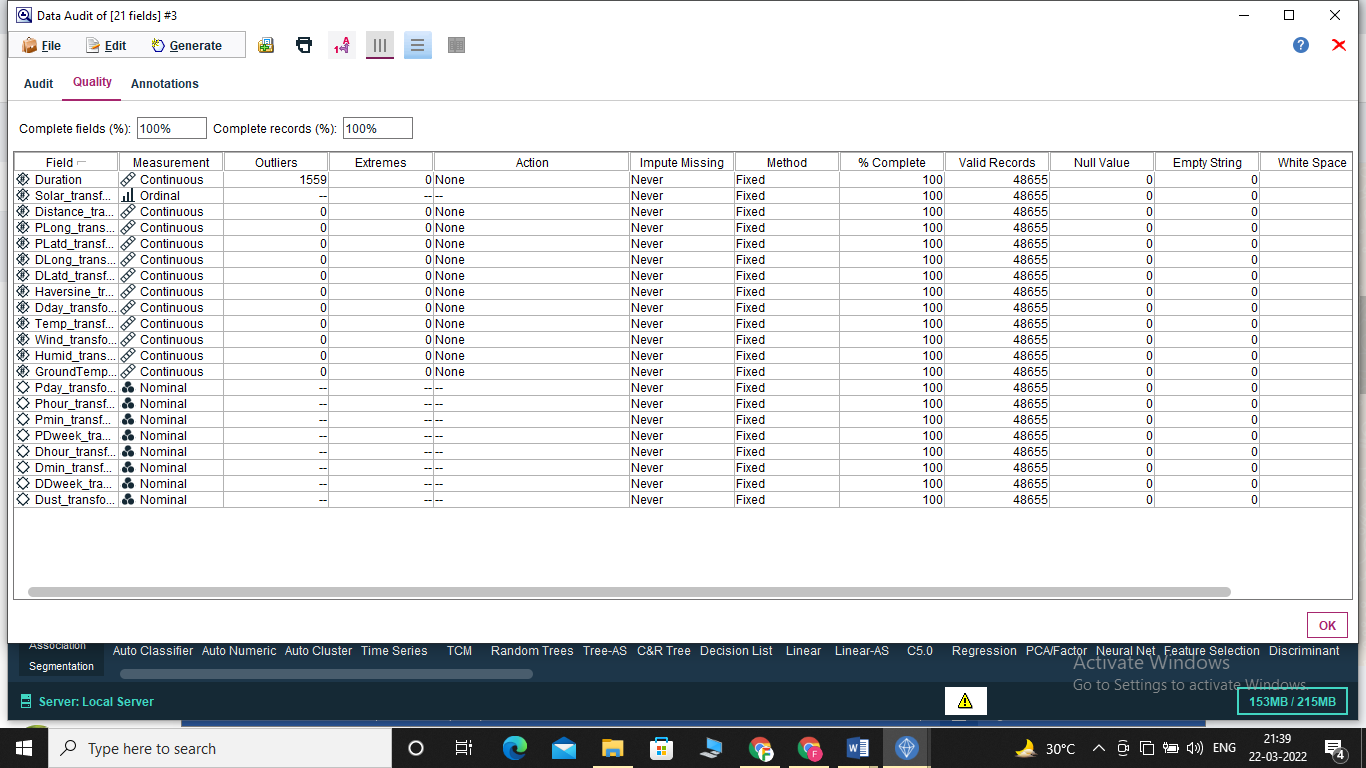
**Data quality:**

Checking the data audit node and checking for outliers, extremes, null values etc.



From the output of this quality tab of the audit node, we can infer the following for each parameters.

**-duration**

(i) It has 1559 Outliers and 0 extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-solar**

(i) It has -- Outliers and -- extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-distance**

(i) It has 0 Outliers and 0 extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-pLong**

(i) It has 0 Outliers and 0 extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-pLatd**

(i) It has 0 Outliers and 0 extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-DLong**

(i) It has 0 Outliers and 0 extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-DLatd**

(i) It has 0 Outliers and 0 extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-haversine**

(i) It has 0 Outliers and 0 extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-Dday**

(i) It has 0 Outliers and 0 extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-Temp**

(i) It has 0 Outliers and 0 extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-Wind**

(i) It has 0 Outliers and 0 extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-humid**

(i) It has 0 Outliers and 0 extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-groundtemp**

(i) It has 0 Outliers and 0 extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-Pday**

(i) It has -- Outliers and -- extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-phour**

(i) It has -- Outliers and -- extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-pmin**

(i) It has -- Outliers and -- extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-PDweek**

(i) It has -- Outliers and -- extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-Dhour**

(i) It has -- Outliers and -- extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-Dmin**

(i) It has -- Outliers and -- extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-DDweek**

(i) It has -- Outliers and -- extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

**-Dust**

(i) It has -- Outliers and -- extremes.

(ii) It has all the records Valid and has no Null values, Empty strings, white spaces and blank values.

* **Have you identified missing attributes and blank fields? If so, is there meaning behind such missing values?**

Missing data are defined as not available values, and that would be meaningful if observed. Missing data can be anything from missing sequence, incomplete feature, files missing, information incomplete, data entry error etc.

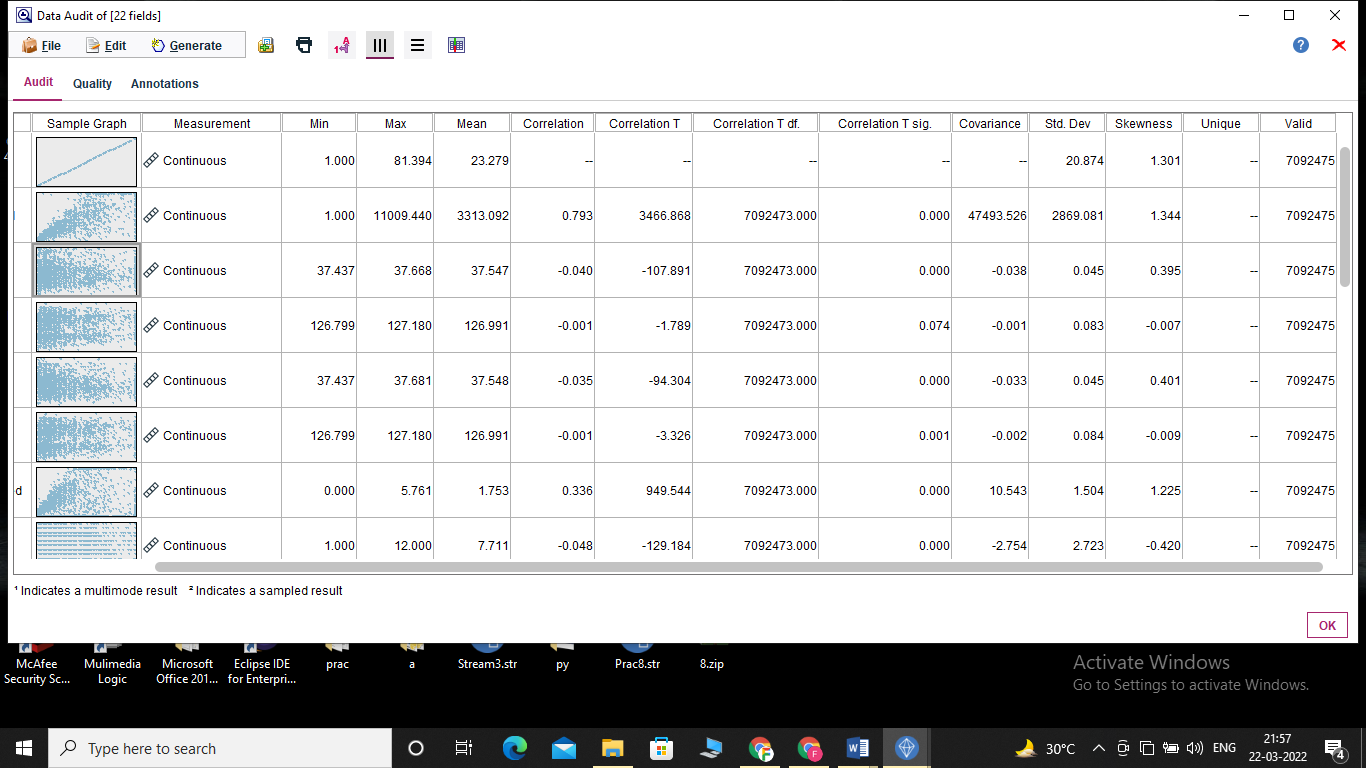
As we can see in the Quality tab of data audit node that there are no missing attributes and blank fields.

* **Are there spelling inconsistencies that may cause problems in later merges or transformations?**

After taking an overview of the .csv file, it can be confirmed that there is no spelling inconsistencies that could possibly cause problems in later merges or transformations.

And since we have only single .csv file so we need not to merge the data at the moment but we will be mindful so that there are no inconsistencies in case this is the need for the future.

* Have you explored deviations to determine whether they are "noise" or phenomena worth analyzing further?



* **Have you considered excluding data that has no impact on your hypotheses?**

Our dataset contains 26 fields and 25 fields are important and useful for the model that we have to create. This can be further validated using the exploratory analysis graphs.