Data Analysis Report

For this report:

I am using a CSV(Comma Separated File) file for data analysis. This file is available on W3schools pandas. This data analysis describes and provide information of the data. What type of value are in there, how many null values are present, does this dataframe have unique values, etc.

CSV DataFrame Overview:

This is the dataframe that I have used in this file. Here, 1'st column is the index, 2'nd is Duration, 3'rd is Pulse. 4'th Maxpulse, 5' is Calories.

	Duration	Pulse	Maxpulse	Calories
0	60	110	130	409.1
1	60	117	145	479.0
2	60	103	135	340.0
3	45	109	175	282.4
4	45	117	148	406.0
..	...	...	...	...
164	60	105	140	290.8
165	60	110	145	300.0
166	60	115	145	310.2
167	75	120	150	320.4
168	75	125	150	330.4

[169 rows x 4 columns]

You can see the complete dataframe here, Dataframe

Data Analysis Report

Descriptive Analysis Data:

This is the description of dataframe. Here, 1'st column is descriptive fields, 2'nd is Duration, 3'rd is Pulse. 4'th Maxpulse, 5' is Calories.

	Duration	Pulse	Maxpulse	Calories
count	169.000000	169.000000	169.000000	164.000000
mean	63.846154	107.461538	134.047337	375.790244
std	42.299949	14.510259	16.450434	266.379919
min	15.000000	80.000000	100.000000	50.300000
25%	45.000000	100.000000	124.000000	250.925000
50%	60.000000	105.000000	131.000000	318.600000
75%	60.000000	111.000000	141.000000	387.600000
max	300.000000	159.000000	184.000000	1860.400000

*** Lets breakdown:**

'COUNT' - Number of all the values present i.e, 169,164

'MEAN' - Mean value of the of that column values

'STD' - Standard deviation value of the of that column values

'MIN' - Min value of the of that column values

'25%' - 25% of the Quartile range of that column values

'50%' - 50% of the Quartile range of that column values

'75%' - 75% of the Quartile range of that column values

'MAX' - Max value of that column values

Note: Quartile Range is method that help to find Outliers in the dataset, that are used in ML for prediction models.

Data Analysis Report

Informational Analysis Data

Information of values present in the dataframe. Like - values count, unique values, null values, etc.

*** Null value analysis**

Check Null Values in column and return sum : Total null values present

Null Values in 'Duration' : 0 values present.

Null Values in 'Pulse' : 0 values present.

Null Values in 'Maxpulse' : 0 values present.

Null Values in 'Calories' : 5 values present.

Note: Here, In 'Calories' sum is 5 stats total 5 null values are present in this column and 0 means no null value.

*** Unique values present**

Values count in column and return sum : Total of no. of values

Values count in 'Duration' : 1830 sum of unique values present.

Values count in 'Pulse' : 5424 sum of unique values present.

Values count in 'Maxpulse' : 7898 sum of unique values present.

Values count in 'Calories' : nan sum of unique values present.

Note: This data is provided on the basis of how many unique values are present and the sum of all the values.

*** Values Counts in the field**

Unique records in column : Total of no. of values

Unique records in 'Duration' : 169 sum of values present.

Unique records in 'Pulse' : 169 sum of values present.

Unique records in 'Maxpulse' : 169 sum of values present.

Unique records in 'Calories' : 164 sum of values present.

Note: There are total count of record is 169 and in 'Calories' count if 164, means it consists null values.

Data Analysis Report

*** DataType of the values**

Values in dataframe : Type of values

Datatype of values in 'Duration' : 'int64' type.

Datatype of values in 'Pulse' : 'int64' type.

Datatype of values in 'Maxpulse' : 'int64' type.

Datatype of values in 'Calories' : 'float64' type.

Note: int64 stats Integer value, float64 stats Floating Point values, and object32 stats String/object value.