**DATA ANALYSIS AND FORECASTING**

Table of Contents

[Introduction 2](#_Toc133245654)

[Main Body 2](#_Toc133245655)

[Conclusion 6](#_Toc133245656)

[References 7](#_Toc133245657)

# Introduction

Data analysis and forecasting provide a clear overview by the means of analysing the raw datasets for a particular time period. In this context, data analysis is mainly integrated by the means of analyzing the raw data sets from various perspectives. Considering this, the assessment is to evaluate the humidity percentage of the city of London for the past ten consecutive days. With respect to this, the raw data sets of the humidity rate in London have been collected and presented in a tabular presentation. After that, the same has been analysed from various perspectives by the means of different statistical methods.

# Main Body

**1. Arrangements of data in table format**

|  |  |
| --- | --- |
| **Humidity Of London For the Past Ten Days** | |
| 14 | 80.9 |
| 15 | 80 |
| 16 | 68.4 |
| 17 | 75.8 |
| 18 | 74.4 |
| 19 | 71.4 |
| 20 | 67.1 |
| 21 | 81.7 |
| 22 | 69.7 |
| 23 | 80 |

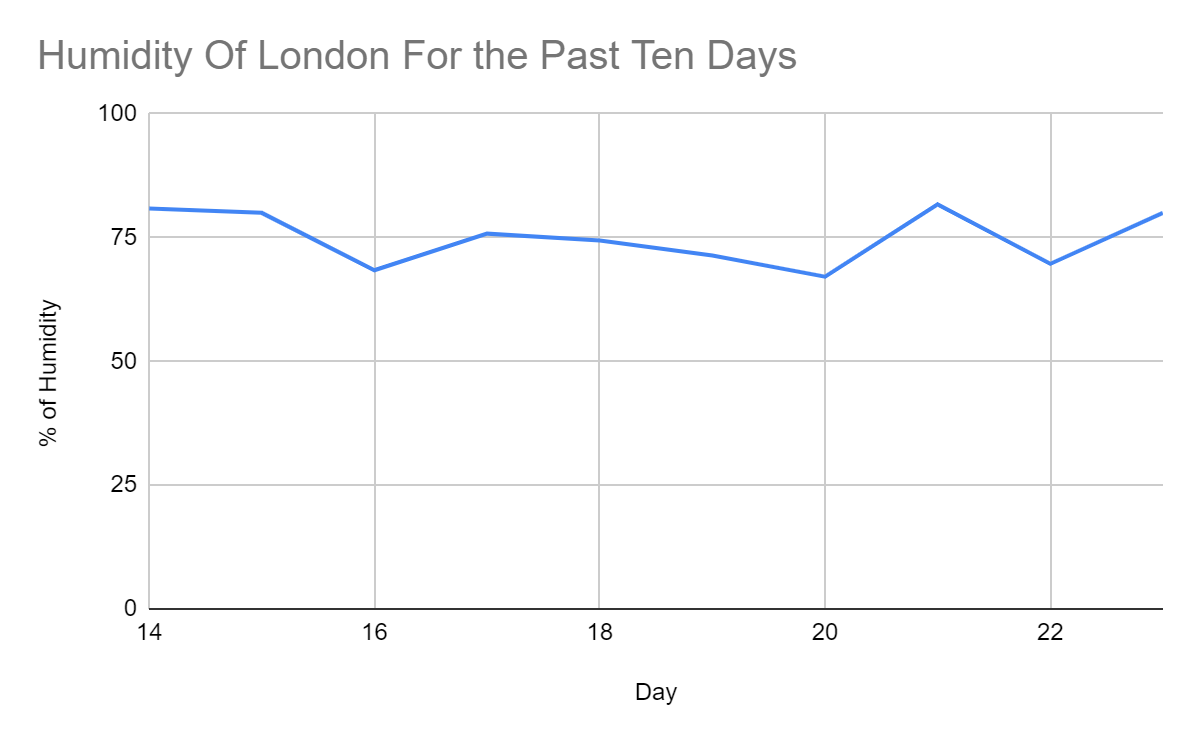
**Table No 1: Table format of the data**

(Source: created by self)

The dataset of the daily percentage of humidity within the entire city of London has been presented over here in a table format in order to analyse them from different statistical perspectives. The dataset has been collected for a period of ten consecutive days in London to analyse them by the means of different statistical tools and techniques.

**2. Presentation of two appropriate charts**

*Line Chart*

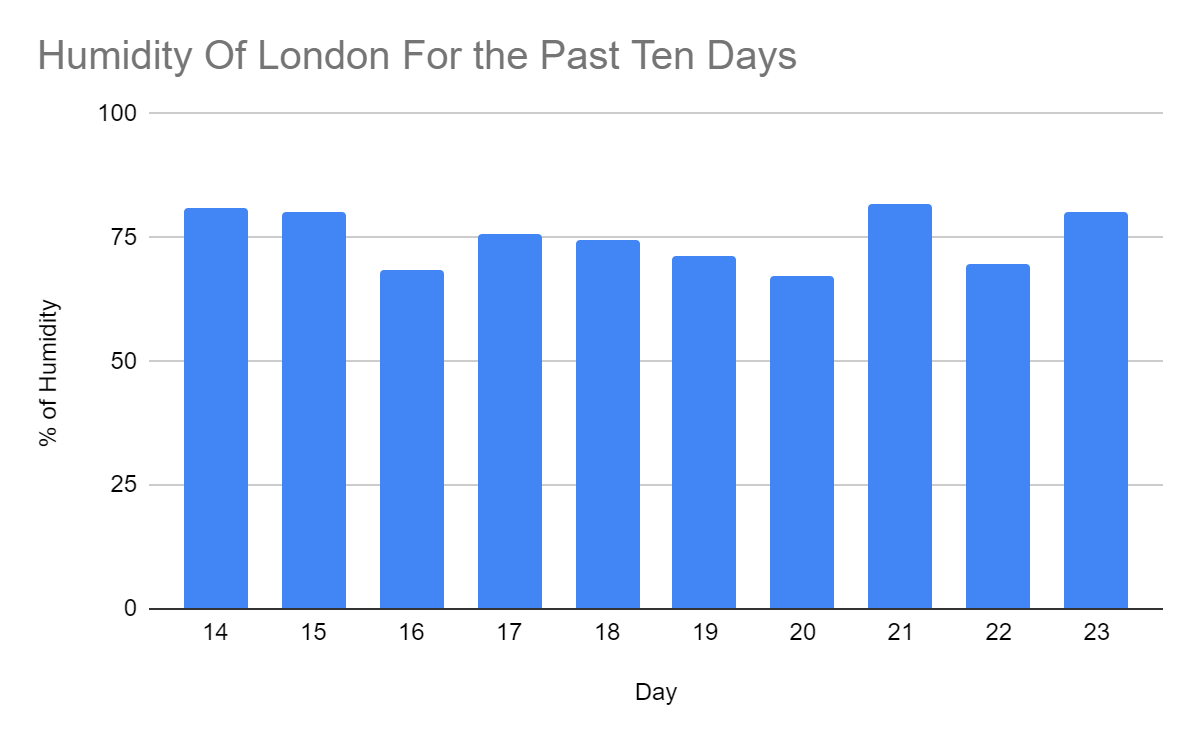
**

**Figure No 1: Line Chart**

(Source: created by self)

The figure presented above that is figure no 1 determines the line chart of the daily humidity rate of London for the concerned ten days.

*Bar Chart*



**Figure No 2: Bar Graph**

(Source: created by self)

The figure presented above (figure no 2) determines the graphical presentation of th bar graph analysisng the dataset of the humidity of London for the past ten days.

**3. Calculation and Discussions**

**(i) Mean**

|  |  |
| --- | --- |
| Mean | **74.94** |

**Table No 2: Valuation of Mean**

(Source: created by self)

**Formula:** Sum of the Total Data Points/Total Number of Data Points

Mean is known to be the average value of a particular set of data (Orcan, 2020). In this context, the mean value of the humidity of London in the last ten days is computed to be 74.94

**(ii) Median**

|  |  |
| --- | --- |
| Median | **75.1** |

**Table No 3: Valuation of Median**

(Source: created by self)

**Formula:** [n+1]th term

2

The median is known to be the middle value of any particular dataset (Fallah, 2020). Considering this, the median computed for London for the past ten days is about 7.51.

**(iii) Mode**

|  |  |
| --- | --- |
| Mode | **80** |

**Table No 4: Valuation of Mode**

(Source: created by self)

**Formula:** l+h (fm-f1)

2fm-f1-f2

Mode is considered to be that particular value which is having the highest level of frequency as compared to the other digits (Sherwood, 2022). For the humidity of London, the Mode is computed to be 80 in the last ten days.

**(iv) Range**

|  |  |
| --- | --- |
| Range | **14.6** |

**Table No 3: Valuation of Range**

(Source: created by self)

**Formula:** Maximum Value - Minimum Value

In any statistical data, the range is known to be the difference between the maximum and minimum values of a particular set of data. Considering this, the range computed for the last ten days of the humidity of London is 14.6.

**(v) Standard Deviation**

|  |  |
| --- | --- |
| Standard Deviation | **5.55** |

**Table No 4: Valuation of Standard Deviation**

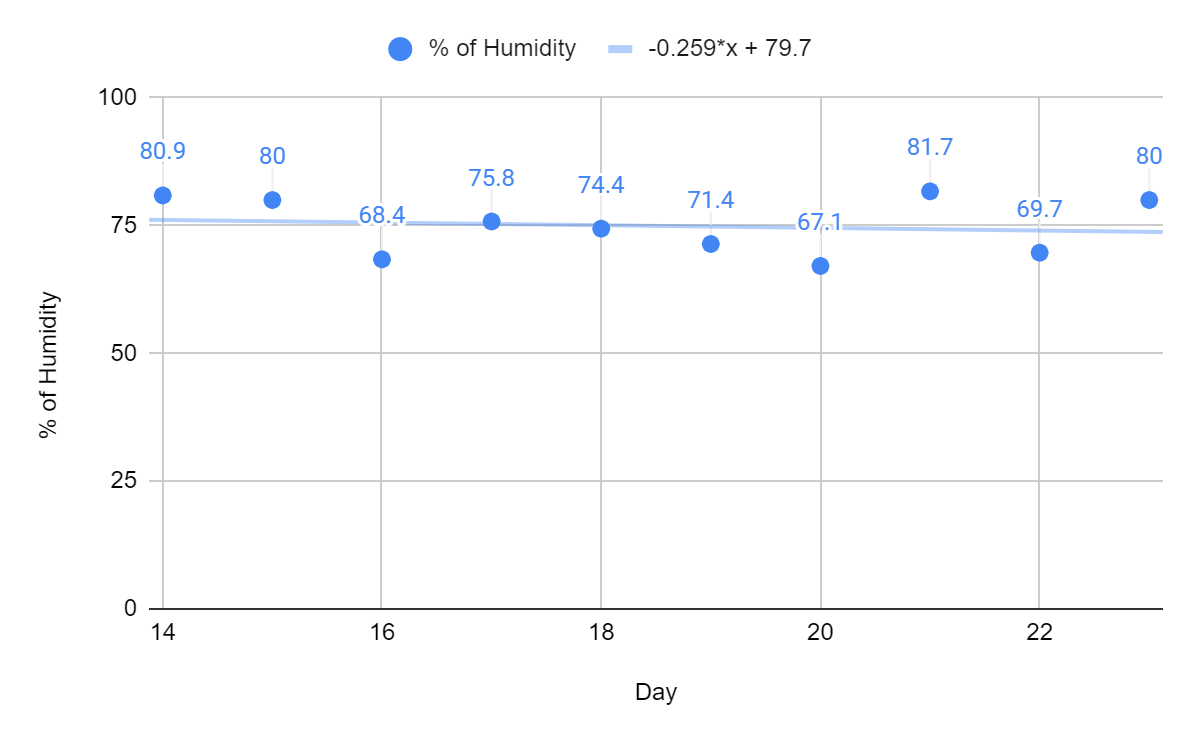
(Source: created by self)

**Formula:** √ ∑i=1n​ (xi ​– x̅)2​ / N

Standard Deviation is a common statistical tool that measures the dispersion level of a particular set of data in relation to the aggregate level of the average value of the same (Rottmann, 2020). With respect to this, the standard deviation of the data set of the humidity of London is computed to be 5.55 over here.

**4. Utilisation of Linear Forecasting model**

Linear Forecasting Model is considered to be a statistical tool that determines the future valuation of a particular set of data in terms of the past valuation of the same. The linear forecasting model also comes with a trend line that reflects the overall formulation of a particular set of data.



**Figure No 3: Trend in the Linear Forecast**

(Source: created by self)

The figure presented above that is figure no 3 determines the graphical presentation of the trend line that has been developed by the means of matching all the schated data of the percentage of humidity of London for the past ten days. With respect to this, it can clearly be observed over here that there is a straight trend line withinn the graph which implies that there was a moderate level of humidity throughout the past ten days in London. In a broader sense, it can be stated that the p]ercentage of the humidity has not been changed much in the past ten days in London.

Moreover, the equuition that has been computed by the means of the concerned regression is as follows:

-0.259\*x + 79.7

# Conclusion

The assessment has covered a wide range of statistical techniques in order to analyse the percentage of humidity in London. In this respect, it can be observed here that there are different types of graph and tabular presentations developed over here which provide a clear scenario of the humidity of London. Different types of statistical tools and techniques has also been used over her in order to understand the entire dataset adequately. Considering this, it can be stated that London has observed a moderate rate of humidity in the last ten days throughout its various areas.

# References

Fallah, A., Rakhshandehroo, G.R., Berg, P., O, S. and Orth, R., 2020. Evaluation of precipitation datasets against local observations in southwestern Iran. International Journal of Climatology, 40(9), pp.4102-4116.

Orcan, F., 2020. Parametric or non-parametric: Skewness to test normality for mean comparison. International Journal of Assessment Tools in Education, 7(2), pp.255-265.

Rottmann, M., Colling, P., Hack, T.P., Chan, R., Hüger, F., Schlicht, P. and Gottschalk, H., 2020, July. Prediction error meta classification in semantic segmentation: Detection via aggregated dispersion measures of softmax probabilities. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-9). IEEE.

Sherwood, C. and Makar, K., 2022. Students making sense of statistics through storytelling: a theoretical perspective based on Bruner's narrative mode of thought. Mathematics Education Research Journal, pp.1-35.