**AIR QUALITY INDEX**

**AND**

**AIR POLLUTION IN BENGALURU**

**ACKNOWLEDGEMENT**

We give praise and thanks to the Almighty God.

This paper is ostensibly supported by the Christ (Deemed to be University), Bangalore.

We would also like to express our sincere gratitude towards the Head of the Department of Statistics, Prof. Joy Paulose for giving us such an opportunity.

We would also like to thank Prof. Nagaraja M S for his constant support throughout this study.

Our team members have made valuable comment suggestions on this project which gave us an inspiration to improve our project.

We would like to thank all those who have been instrumental in our work.

**TABLE OF CONTENTS**

Abstract--------------------------------------------------------------------------------------4

Keywords------------------------------------------------------------------------------------4

Introduction---------------------------------------------------------------------------------5

Materials-------------------------------------------------------------------------------------6

Methodology--------------------------------------------------------------------------------7

Results and Analysis-----------------------------------------------------------------------9

Findings-------------------------------------------------------------------------------------12

Conclusions---------------------------------------------------------------------------------13

References----------------------------------------------------------------------------------14

**Abstract**

This research is an attempted study to analyze the effect of vehicular pollution and variable weather conditions on the **Air Quality Index (AQI)** of Bengaluru city. The present environmental conditions in the city demands an initiative to overcome the issue of depleting nature. The vehicular emissions from urban areas have the major share of contribution towards the environmental destruction in the city. The increasing number of vehicles leads to inconsistent values of AQI. The prediction of air quality by measuring AQI is noted as a major tool to identify the present living conditions in the city. The AQI of study region can be estimated based on seven eminent air pollutants viz., **PM10, PM2.5, SO2, NO2, CO, O3 and NH3.** The monthly AQI calculation during the research revealed that the quality of air in the city can be classified into various categories ranging from low, moderate, satisfactory to unacceptable and hazardous classes. The air quality results also state that lower the value of AQI more is the favorability of living in the city. The increasing rate of AQI can outturn into uncertain living circumstances and health concerns for the population in the city. The research discovered that the prominent index range of AQI for the chosen time period and region falls under the category of moderate class. This interprets that the favorability of living in the city as unaffected or minor health concerns for general public. However, the moderate index of AQI can have serious effects on the health of sensitive and elderly group of the population. The outcomes of the research depicted that it is an essential requirement to conduct periodic testing of vehicular emissions and weather changes in order to resist the plight of hazardous living conditions in the city.

**Keywords**

Air Quality Index (AQI), Vehicular emissions, Temperature, Humidity, Rainfall, Wind, Ambient air quality, Environmental factors, Health concerns, Multiple Linear Regression model, Dependent Variables, Independent Variables, ANOVA, Skewness, Kurtosis, Regression Sum of Squares, Residual Sum of Squares, Dependency, Processes, Bengaluru.

**Introduction**

Air pollution has always been a global cause of worry. The city of Bengaluru is in the grip of rising air pollution. It is ranked second among the highest polluted cities in India after Delhi. The Garden city has almost 14% of school going children suffering due to air pollution. Air pollution has emerged as a major concern for Bangalore, with more than 75% of its areas where air quality monitoring was carried out violated the prescribed standards. The extremely populated areas of Bengaluru such as Rail Wheel factory, Yelahanka, Export promotional park ITPL, Yeshwanthpura and Central Silk Board depict severe impacts of air pollution. Due to the presence of numerous vehicles and industries such densely populated areas have experienced serious consequences of pollution. The city has the largest concentration of IT parks, electronic and bio-tech industries and is popularly known as the ‘Silicon Valley of India. The rapid development and changing landscape of the city over the past few years has led to a rise in the population approximately to 8.4 million inhabitants and number of vehicles were increased by approximately 4 million. This has impacted on the infrastructure of the city with tremendous challenges and pressure with air pollution being the most crucial one. A significant increase in air pollutants and environmental deterioration has been witnessed by the city. The air pollutants have also impacted of public health causing breathing disorders, lung cancer, pulmonary cancer, COPD and several other illnesses. All such hazardous effects of air pollutants reflects the need of testing and estimating the quality of air periodically.

The key factors affecting the air quality of the city included factors such as large number of vehicles, raise in temperature, humidity, rainfall and wind along with many other climate changes.

The most commonly used technique for estimating the pollutants in the air is Air Quality Index (AQI). This process involves detection of poisonous pollutants such as PM10, PM2.5, SO2, NO2, CO, O3 and NH3 in the air. The government agencies use AQI method to communicate with public regarding the health concerns due to air pollutants. The risk on public health increases with the rise in AQI. There is an air quality reference index corresponding to certain air quality standards. This reference criteria can be used to predict the impact and severity of pollutants on public health. The value of AQI ranging from 0-50 represents a low AQI with no health implications, 50-100 range of AQI represents a moderate level of AQI with health concerns for the sensitive and elderly group of population. Whereas, if the AQI exceeds the range of 100 it depicts that air quality of the area under consideration is extremely hazardous for the health of general public. The paper reflects thoroughly upon the major air pollutants sources of air pollution and their causes, by considering a common technique of determining the quality of air in city of Bengaluru. It also provides the general impact of this pollutants on public health and the environment.

**Materials**

The materials and framework used for the research analysis includes the monthly data collected for the time span of 30 months. The considered time period is from the month of April 2016 to September 2018. The recordings of this chosen 30 months is utilized in the research to predict the air quality of Bengaluru city. The data and statistics assembled for the analyzing the air quality of Bengaluru city is recorded as follows:

* The AQI for all the months from April 2016 to September 2018 is calculated. The AQI is predicted based upon the average composition of harmful pollutants such as PM10, PM2.5, SO2, NO2, CO, O3 and NH3 in the air.
* The number of vehicles in Bengaluru city during the months of April 2016 to September 2018. It particularly involves the total number transport and non- transport vehicles along with 2 wheelers, 3 wheelers, 4 wheelers, other transports and LMV goods.
* The Climate change during the chosen time span of 30 months is recorded. The change in climate of Bengaluru involves variations in factors such as amount of rainfall, wind speed, temperature and humidity.
* The data for temperature changes from April 2016 to September 2018 is recorded in Celsius.
* The average rainfall (in mmm) and the number of rainy days during the chosen span is obtained.
* The average speed of wind from April 2016 to September 2018 is calculated in km/hr.
* The humidity changes in the climate during the chosen months is also recorded.

**Methodology**

According to the assembled data and parameters, it is observed that the technique of Multiple Linear Regression model is best suitable for the research analysis. The chosen categories of factors affecting the air quality along with AQI depicts the characteristics of variables required for analysis using **Multiple Regression**. It is noted from the gathered data that the quality of air which is predicted using AQI shows a variation due to extraneous factors such as number of vehicles, temperature, humidity, wind speed and rainfall. Such a variation in AQI is due to the dependency of air quality on this external factors. By involving the concept of Multiple Linear Regression in the prediction of air quality it is observed that the monthly AQI calculated using the average composition of major air pollutants is the Dependent Variable and all other factors involving vehicles, temperature, humidity, wind, rainfall and rainy days are the Independent Variables. Thus, the chosen data reflects one Dependent Variable and six Independent Variable for the research analysis using Multiple Regression.

**The concept of Multiple Regression** describes the relationship between multiple independent variables and one dependent variable. The independent and dependent variables are also known as predictor and criterion variables respectively. A dependent variable is represented as a function of several independent variables with corresponding coefficients, along with the constant term. Multiple Regression is possible only with the involvement of at least two or more predictor variables. Generally, the Dependent Variable is symbolized using Y and Independent Variables are represented using X1, X2, X3…….Xk. The model for multiple regression is given by:

**Y = b0 + b1 X1 + b2X2 + …………………… + bk Xk**

Here b0 is the intercept and b1, b2, b3… bk are analogous to the slope in [linear regression](https://explorable.com/linear-regression-analysis) equation and are also called regression coefficients. The appropriateness of the multiple regression model is tested by the F-test in the [ANOVA](https://explorable.com/anova) table. A significant F indicates a linear relationship between Y and at least one of the X's.

The constructed Multiple Regression model can be used to check predictive ability by examining the coefficient of determination (R2). R-Square, also known as the Coefficient of determination is a commonly used statistic to evaluate model fit. R-square is 1 minus the ratio of residual variability. When the variability of the residual values around the regression line relative to the overall variability is small, the predictions from the regression equation are good. For example, if there is no relationship between the X and Y variables, then the ratio of the residual variability of the Y variable to the original variance is equal to 1.0. Then R-square would be 0. If X and Y are perfectly related then there is no residual variance and the ratio of variance would be 0.0, making R-square = 1. In most cases, the ratio and R-square will fall somewhere between these extremes, that is, between 0.0 and 1.0. This ratio value is immediately interpretable in the following manner. If we have an R-square of 0.4 then we know that the variability of the Y values around the regression line is 1-0.4 times the original variance; in other words we have explained 40% of the original variability, and are left with 60% residual variability. Ideally, we would like to explain most if not all of the original variability. The R-square value is an indicator of how well the model fits the data. An R-square close to 1.0 indicates that we have accounted for almost all of the variability with the variables specified in the model.

Multiple regression technique does not test whether data are [linear](https://explorable.com/linear-relationship). It proceeds by assuming that the relationship between the Y and each of Xi's is linear. Hence as a rule, it is prudent to always look at the scatter plots of (Y, Xi), i= 1, 2… k. If any plot suggests [non linearity](https://explorable.com/non-linear-relationship), one may use a suitable transformation to attain linearity. Another important assumption is non-existence of multicollinearity that is the independent variables are not related among themselves). At a very basic level, this can be tested by computing the correlation coefficient between each pair of independent variables. Other assumptions include those of homoscedasticity and normality.

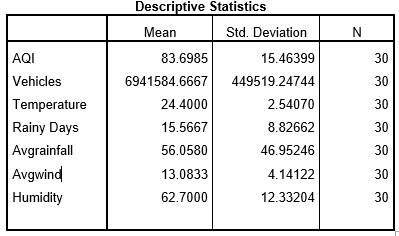
Homoscedasticity describes a situation in which the error term that is, the “noise” or random disturbance in the relationship between the independent variables and the dependent variable is the same across all values of the independent variables and Multivariate Normality assumes that residuals are normally distributed.

Multiple regression analysis is used when we are interested in predicting a continuous dependent variable from a number of independent variables. If dependent variable is dichotomous, then logistic regression should be used.

* The beta value is used in measuring how effectively the predictor variable influences the criterion variable, it is measured in terms of standard deviation.
* R is the measure of association between the observed value and the predicted value of the criterion variable.
* R Square or R2, is the square of the measure of association which indicates the percent of overlap between the predictor variables and the criterion variable.
* Adjusted R2 is an estimate of the R2if you used this model with a new data set.

**Analysis and Results**

The obtained dataset containing the AQI levels and the variations in impacting factors from April 2016 to September 2018 was aggregated and analyzed for the process of research work. Initially, the basic frame of descriptive statistics is acquired for the dataset. It indicates that the mean and standard deviation statistics for all the parameters in the research dataset. For time span of chosen 30 months, it suggests that the average AQI level is 83.6985 with a standard deviation of 15.46399. It depicts that the average air quality is acceptable. However, for some pollutants there may be a moderate health concern for a very small number of people during the sensitive and elderly group. The average number of vehicles are 6941584.6667 suggesting a huge impact of vehicular exposure on the air quality. The average climatic changes including Temperature, Rainy days, Rainfall, Wind, and Humidity were attained as 24.4000, 15.5667, 56.0580, 13.0833 and 62.7000 respectively. This mean quantities depicts a huge impact of weather related factors on air quality. The values of standard deviations for the dependent and independent variables are highly deviated from the mean suggesting that the data points are spread over a wider range of values. The maximum observed level of AQI during the study period is 125.583 and the minimum value is 58.2. In case of the Independent Variables, the maximum number of registered vehicles in Bengaluru is 6078289 and minimum number was 7683753. The maximum recorded temperature is 29 C and the minimum is 20 C. The number of rainy days ranges from 0 to 29 for various months. The maximum amount of rainfall determined is 144.36mm, while the minimum value is 0.03mm. The highest observes speed of wind was found to be 22.8km/hr and the minimum speed is 8.4km/hr. The humidity at 80% is the maximum and 20% is minimum during the study period of 30 months. The data also verifies that AQI, Temperature, Wind and Rainfall are positively skewed whereas, rainy days, vehicles and humidity are negatively skewed. The kurtosis obtained for all the variables are less than 3, which signifies that the data is platykurtic. It means the distribution produces fewer and less extreme outliers than does the normal distribution.



The Correlation statistics directing the mutual relationship or interdependency between the factors affecting AQI is calculated. It reflects that AQI is negatively associated with extrinsic causes such as Rainy Days, Average Rainfall, Average Wind and Humidity. Whereas, it is positively associated with factors such as Temperature and Vehicles. The negative correlation suggests an inverse relationship between the two variables, higher values of one variable tend to be associated with lower values of the other. This leads to the interpretation that the increasing climatic factors such as rainy days, rainfall, wind speed at moderate level and humidity in the environment will decrease the amount of AQI. The lowering the values of AQI depicts a satisfactory air quality and possess no risk and harm to the surroundings or public health. On the other hand, the AQI has a positive correlation with factors such as Vehicles and Average Temperature. The statistical association suggests a linear relationship between this variables. It interprets that the raising temperature and vehicles can lead to extreme levels of AQI causing a severely unhealthy living conditions even for the general public. The AQI levels beyond the range of 100 outturns into acute illnesses and has tremendous threat on the surrounding environment.

The Multiple Linear Regression Model for the dataset is constructed and interpreted with one dependent variable that is AQI and six predictor variables viz., vehicles, temperature, humidity, wind, rainfall and rainy days. The value of R2 obtained by summarizing the Multiple Regression model was attained as 0.757. The coefficient of determination, denoted by R2, is the proportion of the variance in the dependent variable that is predictable from the independent variables. Therefore, the data states 75.7% variation in the levels of AQI during the time period of April 2016 to September 2018 due to vehicles, temperature, humidity, wind, rainfall and rainy days. The remaining 24.3% variation in AQI is due to other extrinsic factors. The Adjusted R-squared that adjusts the statistic based on the number of independent variables in the model is obtained as 0.694.

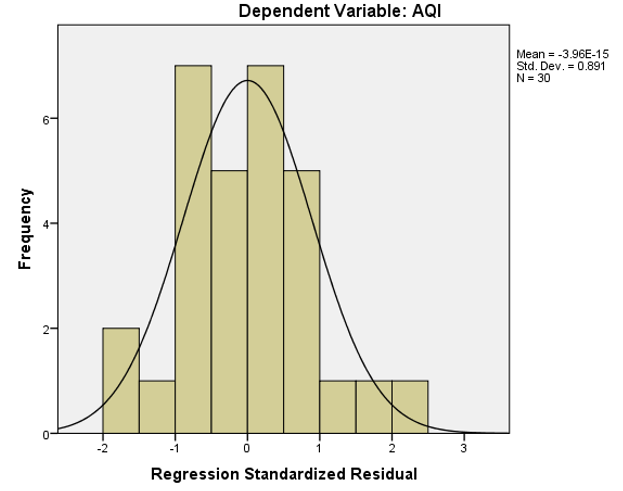
|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model Summary | | | | | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | |
| R Square Change | F Change | df1 | df2 | Sig. F Change |
| 1 | .870a | .757 | .694 | 8.55899 | .757 | 11.944 | 6 | 23 | .000 |
| a. Predictors: (Constant), Humidity, Vehicles, Temperature, Avgwind, Avgrainfall, Rainy Days | | | | | | | | | |
| b. Dependent Variable: AQI | | | | | | | | | |

The Analysis of Variance (ANOVA) is calculated, that depicts the Regression and Residual Sum of Squares of Air Quality dataset. The Regression model is Significant at 1% Level of Significance with a Sum of Squares 5250.020 and F-value of 11.944. This Explained Sum of Squares or the Model Sum of Squares describes that the acquired data and variables are well-fitted into the Multiple Regression Model. The Residual sum of squares (RSS) is obtained as 1684.896. It indicates the amount of variance in a data set that is not explained by the regression model. It gives the measure of error remaining between the regression function and the dataset.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 5250.020 | 6 | 875.003 | 11.944 | .000b |
| Residual | 1684.896 | 23 | 73.256 |  |  |
| Total | 6934.916 | 29 |  |  |  |
| a. Dependent Variable: AQI | | | | | | |
| b. Predictors: (Constant), Humidity, Vehicles, Temperature, Avgwind, Avgrainfall, Rainy Days | | | | | | |

The Coefficients table is obtained for the Independent Variables. By increase in one vehicle the AQI level is decreased by -1.288E-5 that is almost a negligible value. As temperature increases by 1 C the AQI level is decreased by -4.301. With the increase in one Rainy day AQI is increased by 1.717. With the increase in Rainfall the AQI is reduced by -.130. The increasing wind causes the AQI to be reduced by -2.290. Also, the increasing humidity decreases the AQI by -1.208.

A visual representation for the normality of the Dependent Variable AQI is procured that illustrates the Normal Distribution of AQI in terms of its Frequency and Regression Standard Residuals.



**Findings**

1. Air Quality Index (AQI) of Bangalore during the time period of April 2016 to September 2018 was affected widely by factors such as number of vehicles, temperature, rainy days, rainfall, wind and humidity.
2. The comparison of obtained AQI values with the air quality reference range depicts moderate level of AQI for most of the months during study period.
3. The analytical study of the dataset verified the correlation and interdependency amongst the variables.
4. AQI is negatively associated with extrinsic factors Rainy Days, Average Rainfall, Average Wind and Humidity. This states that increasing climatic factors such as rainy days, rainfall, and wind speed at moderate level and humidity in the environment will decrease the amount of AQI.
5. The increasing Temperature and Vehicles can raise the AQI levels by causing serious harm to environment and public health.

**Conclusions**

Bengaluru experienced varying levels of air pollution from April 2016 to September 2018. Many of the densely populated regions the city experienced either high or critical levels of pollution as analyzed during the study. The Air Quality Index (AQI) determined by the composition several gases such as PM10, PM2.5, SO2, NO2, CO, O3 and NH3 depicts that the Reduction in SO2 and NO2 levels in the city may be an outcome of a combing of various interventions such as implementing stringent emission norms for vehicles and improvement in the quality of fuel. Also, most of the vehicles are switching to LPG for automobile fuel. Meanwhile, the levels of PM10 and PM2.5 remains high or critical during time period of study. Uncontrolled exposures to elevated concentrations of PM10 and PM2.5 can trigger health risks and can reduce the life expectancy. The critical levels of PM are likely to have a damaging effect on the health of the citizens in Bangalore that may result in a tremendous burden on the public health system. Also if the pollutants affect skilled young human resources, it would pose a threat to the city’s economy that is growing exponentially. It is the need of the hour to address the public about the impacts of air pollution and adhering to stringent measures of pollution control. The AQI levels in the city must be estimated periodically and variations in the air quality levels should be communicated with public. Also, the awareness about raising number of vehicles and high temperature causing air pollution must be alerted in the city. The time period chosen for the study reflects ‘Moderate’ levels of AQI for most of the months. But even the Moderate level of AQI can cause serious health concerns among the sensitive group. The KSPCB has taken many initiatives to improve the quality of air, such as improving fuel quality, introducing alternate fuels for auto rickshaws, improved traffic management, implementation of emission norms, promoting use of green fuel by industries for DG sets, etc. In a nutshell, there’s is a need for wider research and publications in the area of Bengaluru air pollution and its effects on public health.

**References**

[1] Air Pollution in Bangalore, India: A Six-Year Trend and Health Implication

Analysis

[2] A STUDY ON AIR POLLUTION BY AUTOMOBILES IN BANGALORE CITY

[3] ANALYSIS OF AMBIENT AIR QUALITY USING AIR QUALITY INDEX – A CASE STUDY

[4] Air Quality Index – A Comparative Study for Assessing the Status of Air Quality

[5] A Review on Air Quality Indexing System

[6]<https://www.researchgate.net/publication/292643438_Air_Quality_Index_-_A_Comparative_Study_for_Assessing_the_Status_of_Air_Quality>

[7]<https://www.researchgate.net/publication/282222215_A_Review_on_Air_Quality_Indexing_System>

[8] <http://idl.iscram.org/files/binxuzhai/2017/1466_BinxuZhai+JianguoChen2017.pdf>

[9]<http://www.technicaljournalsonline.com/ijeat/VOL%20I/IJAET%20VOL%20I%20ISSUE%20II%20JULY%20SEPTEMBER%202010/Article%2011.pdf>

[10]<https://www.indiastat.com/table/health-data/16/acute-respiratory-infection/379780/1095746/data.aspx>

[11] <https://airnow.gov/index.cfm?action=aqi_brochure.index>

[12]<https://www.researchgate.net/publication/256481805_Air_pollution_in_Delhi_Its_Magnitude_and_Effects_on_Health>

[13]<https://www.researchgate.net/publication/257816315_Air_Pollution_in_Bangalore_India_A_Six-Year_Trend_and_Health_Implication_Analysis>

[14] <https://www.tandfonline.com/doi/abs/10.1080/00968205609344395?journalCode=aihq20>

[15]<https://www.researchgate.net/publication/283072534_Air_Pollution_and_India_Current_Scenario>

[16]<https://scholar.harvard.edu/files/rpande/files/ipf_2017_greenstone_harish_pande_sudarshan_submittedmanuscript.pdf>

[17] <http://www.teriin.org/projects/teddy/pdf/air-pollution-health-discussion-paper.pdf>

[18]<https://pdfs.semanticscholar.org/1e4d/72a0d05e07c4604c5015164ecb82d63ac2f7.pdf>

[19] <http://www.aaqr.org/files/article/1035/2_AAQR-11-11-OA-0220_1049-1061.pdf>

[20] Vehicles: <http://transport.karnataka.gov.in/index.php/information/details/vehicle_statistics>

[21]Rainfall:<https://www.worldweatheronline.com/lang/en-in/bangalore-weather-averages/karnataka/in.aspx>

[22] AQI: <https://airnow.gov/index.cfm?action=airnow.calculator>

[23] Pollution: <https://www.kspcb.gov.in/ambient_air_quality.html>