

# **Air Quality of Karachi and Lahore: How does the energy industry of Pakistan affect it?**

Muzammil Tariq

Quantitative Research Methods Spring 2021

May 26, 2021

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Section I: Literature Review</b>	<b>2</b>
<b>3</b>	<b>Section II: Data</b>	<b>4</b>
<b>4</b>	<b>Section III: Analytical Method, Results and Interpretation</b>	<b>6</b>
<b>5</b>	<b>Conclusion</b>	<b>8</b>

## List of Figures

1	Scatter Plot of PM2.5 readings from June 2019 to March 2021 . . . . .	4
2	Pie Chart of average gas and electricity usage from 2002 to 2016 . . . . .	5
3	Proportion Test for Karachi . . . . .	7
4	Proportion Test for Lahore . . . . .	7

## List of Tables

1	Yearly total usage of gas and electricity in Sindh and Punjab . . . . .	6
2	Difference in AQI readings between Feb-20 and Feb-21 . . . . .	6

## **Abstract**

Air Quality Index (AQI) is an important measurement of the quality of air we breathe in. In Pakistan the Air Quality Index has worsened over the past few decades and between Pakistan's two major financial hubs Karachi and Lahore there is a noticeable difference between this reading. Meanwhile, the energy sector of Pakistan has increased too over the past few decades and has been producing more than ever given the increasing need for electricity and gas due to an increase in population. This paper will attempt to do a study on this worsening rate of the AQI readings and find a correlation between the increasing energy sector of Pakistan and this increasing rate of AQI. It will then try to understand why the AQI readings of Lahore are much worse than Karachi.

## **1 Introduction**

Air pollution has been a cause of concern around the globe since quite some time, specifically since the industrial revolution. The 2019 World Air Quality Report by IQAir states that every year air pollution causes around 7 million premature deaths while 92% population of the earth is reported to breathe toxic air one way or another. Ever since the industrial revolution the energy industry has been booming and much work on sustainable energy sources has not been done. This is more apparent in Pakistan as there exist very few sustainable methods of energy generation. Past history and general observance in Karachi and Lahore have shown a worrying drop in the quality of air in these two cities when compared to the general air quality of the province of Sindh and Punjab.

The quality of air is observed using a measure called Air Quality Index (AQI) which is a reading of several chemical particles in the air such as ozone(O<sub>3</sub>), particulate mass (PM<sub>10</sub>) and fine particulate matter (PM<sub>2.5</sub>). The most common method of measuring AQI is measuring the amount of PM<sub>2.5</sub> in the air. PM<sub>2.5</sub> is the name given to all the particles in air which measure up to 2.5 microns in size. They are regarded as the most harmful pollutants for humans as their small size allows them to enter the human body and cause respiratory issues among other health

concerns. Readings of AQI in Pakistan have been the highest in Lahore and then significantly lower in Karachi. Even though both these cities share similarities between each other such as being the provincial capital and financial hubs it is strange that there is such a large difference between their AQI readings.

This paper will try to understand the reasons for the difference in the AQI readings of Karachi and Lahore. It will do so by establishing a causal connection between these AQI readings and the continuous increase in the energy industry due to the general increase in human population in the province of Sindh and Punjab. Data taken from various official government sources will be used and STATA will be used to summarize this data and generate figures to understand trends in it. After this, statistical methods will be applied to this data to produce a model which describes how AQI could change in the near future. The paper will be divided into 3 sections. Section I will do an analysis of similar studies which were conducted in the past, section II will describe the data being used and section III will run statistical methods on the data and then run interpret the results.

## **2 Section I: Literature Review**

Air quality is managed at the national level by Pakistan Environmental Protection Agency. In 2001 the National Environmental Action Plan was established which set air pollution as one of its fundamental programs (Qurratulain, Shams, & Khwaja, 2019). Ever since the early 2000s research has been going on on air pollution in Pakistan. Although AQI has not been monitored since then there have been other investigations which lead to some conclusion. As researched by Qurratulain (2019) lead levels in blood of school going children of provincial capitals was monitored in the early 2000s which indicated that due to an increasing use of leaded petrol and increasing vehicular traffic on roads these lead levels had reached an alarming rate, it was reported that the carbon monoxide levels had passed the WHO safety standards in Karachi and Lahore. After these researches it was established back then that even though measurements like this helped understand the severity of air

pollution in the major cities of Pakistan there still was no set standard monitoring system on which people could rely on to understand the trends of air pollution. Based on the limited set of data that was available some guidelines were set in 2006 by Aziz (2006) for 6 major pollutants of air. Brick kilns were reported to be a major cause of this deteriorating quality of air. They continuously produced harmful pollutants into the air.

Later in 2007 Ghauri et.al recognized pollutants which were seasonal such as CO which was more common in Quetta and SO<sub>2</sub> and NO<sub>x</sub> which were more common in Lahore. The highest level of pollutants was found in Lahore when compared to other major hubs like Karachi and Peshawar. In another research Mehmood (2020) discussed the viability of monsoon rain being the solution to the problem of air pollution in Pakistan. The region in Punjab around Lahore faces heavy monsoon rainfall every year. They performed a sampling study of both PM<sub>2.5</sub> and PM<sub>10</sub> in pre and post monsoon seasons. They recorded AQI readings which showed that the air quality went from good to slightly polluted in this shift in weather. In their studies they noticed that there was very little change in the readings of PM<sub>2.5</sub> because of the monsoon but there was a large change in the PM<sub>10</sub> readings where the readings increased after the monsoon. Their collected evidence showed that the pre monsoon period was relatively less polluted when compared to the post monsoon period. They conducted their study in Islamabad which is quite close to Lahore.

Rehana Siddiqui did a study on the economy related to the growth of energy industry in Pakistan in 2004. Although a little bit outdated her paper discusses the rising energy uses in the country at the start of the 21st century. The growth rates have shown that there has been a constant increase and decrease in the consumption of petroleum-based products. Statistics also show that there has been a constant increase in electricity creation and usage in all over Pakistan. Rehana then ran several hypothesis tests to check if several factors have caused the growth in the energy sector or not. Similarly, Muhammad Asif and Naila Saleh did a case study in 2019 on the energy security condition of Pakistan. In their study they formulated that oil and gas made up over 75% of the supply of energy in Pakistan. This makes the entire energy sector highly non sustainable and prone

to decreasing the quality of air in the country specially in the financial hubs where majority of the province's work takes place. While discussing the health security of humans due to this increasing energy use, they touched upon the emission of carbon dioxide (CO<sub>2</sub>) which is very harmful for health. The research stated that steam-based power plants counted as the largest donor of CO<sub>2</sub> to the environment. This steam is generated by either burning oil or gas so it connects with the worsening air quality.

### 3 Section II: Data

This paper will use several datasets to understand the research question it posed at the beginning. The World Air Quality Index is an open-source project which records the AQI readings of all major cities in the world. Karachi and Lahore are 2 cities whose AQI measurements are done on a monthly basis by them and these readings are regularly updated on their websites. As such, they have publicly available datasets of AQI readings of Karachi and Lahore. Figure 1a and 1b show the scatter plot of the daily AQI readings measured from the month of June in 2019 to March 2021. It can be clearly observed that the readings are exceptionally higher in Lahore (reaching up to 500) when compared to Karachi.

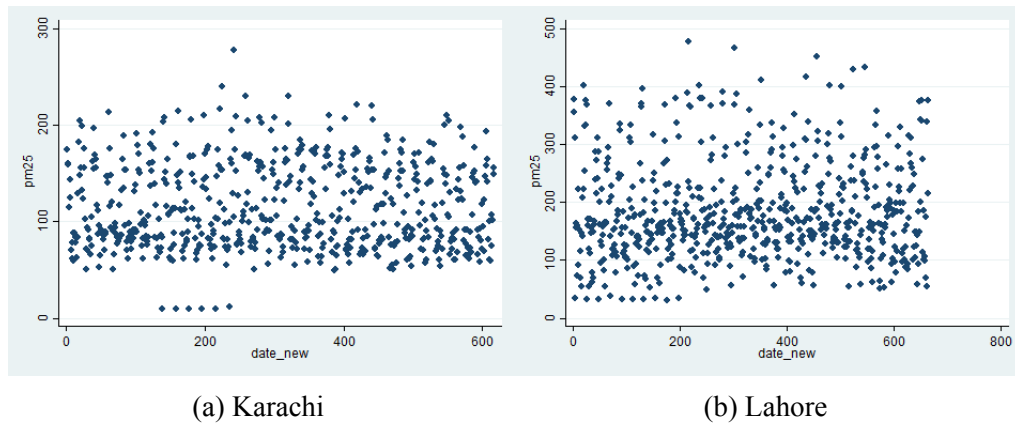


Figure 1: Scatter Plot of PM2.5 readings from June 2019 to March 2021



Besides these datasets showing the AQI readings of Karachi and Lahore this paper will use a dataset detailing the consumption of gas and electricity in Punjab and Sindh to understand the increasing usage of gas and electricity. The Bureau of Statistics Punjab was established in 1957 and has been collecting statistics regarding all activities in Punjab and other provinces when the data is connected to Punjab. It cleans and organizes data after collecting it and this data is later made public to be used in publications. This will paper will use a dataset that they have been maintaining on the gas and electricity usage in Punjab and Sindh. This dataset categorizes the usage in both the provinces into groups such as domestic, commercial and so on. Figure 2a shows a pie chart displaying the collective consumption of both gas and electricity in Sindh and figure 2b shows the usage for Punjab. These pie charts indicate that majority of the energy is consumed in domestic environments with industries being the second largest user.

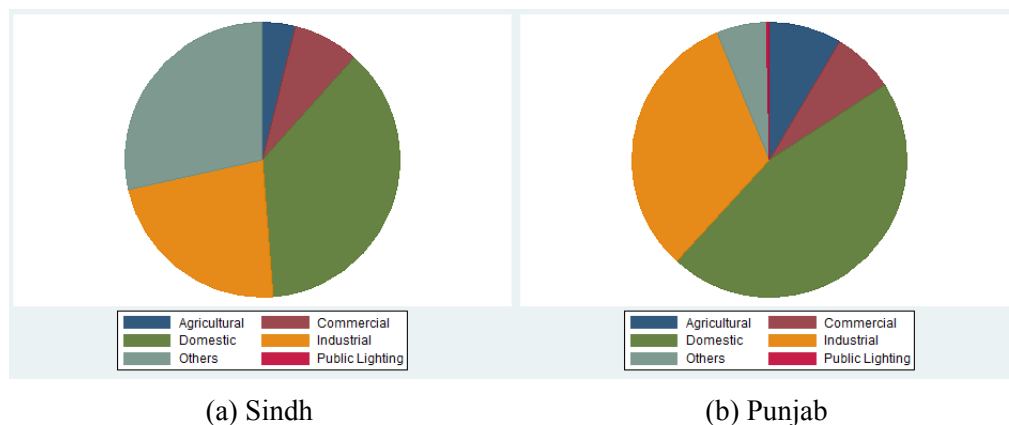


Figure 2: Pie Chart of average gas and electricity usage from 2002 to 2016

These data trends show that on average the average household consumer has been consuming gas and electricity the most. Table 1 shows the total usage of gas and electricity for the past few years. We can see that there has been an increasing trend of gas and electricity usage in Punjab but the measurements have remained almost the same in Sindh. This trend implies that there is a causal relation between the increasingly high AQI readings of Lahore and the usage of gas and electricity in Punjab. Since gas and electricity is a creation of the energy industry this implies that the energy industry affects the air quality of the provinces.

Year	Punjab (Kwh Million)	Sindh (Kwh Million)
2013	52087	23325
2014	53249	24417
2015	57245	24270
2016	60940	24557

Table 1: Yearly total usage of gas and electricity in Sindh and Punjab

## 4 Section III: Analytical Method, Results and Interpretation

Karachi		Lahore	
Month	No. of Days with PM2.5>200	Month	No. of Days with PM2.5>200
February 2020	1	February 2020	4
February 2021	3	February 2021	22

Table 2: Difference in AQI readings between Feb-20 and Feb-21

In this section we will go over the statistical methods this study will use to verify the possible causal relationship between the AQI readings of Karachi and Lahore and the energy usage of Sindh and Punjab. Since data related to AQI was not recorded earlier in Pakistan we are limited to use the data generated after 2018. We can not setup a regression model for the PM2.5 readings as the recorded data is quite limited and has only one variable. Instead, we will run a hypothesis test on the AQI readings and try to see if they have statistically changed in the last year. We use the STATA proportion test calculator to perform a hypothesis test on the change in AQI readings between February 2021 and February 2020. We only take into account February as we are limited by the scarcity of the data which was completely available for only February for both 2021 and 2020. For the proportions we will take those days which had a PM2.5 reading of more than 200 as the air becomes *very unhealthy* after that according to air quality standards, this data is summarized in table 2. We will have separate hypothesis tests for Karachi and Lahore to test individually if there has been a change in the readings of the city. Our hypothesis for Lahore would then be:

- Null Hypothesis: Number of *very unhealthy* days in Lahore did not change between Feb-2020 and Feb-2021

- Alternative Hypothesis: Number of *very unhealthy* days in Lahore did change between Feb-2020 and Feb-2021

The hypothesis for Karachi would be:

- Null Hypothesis: Number of *very unhealthy* days in Karachi did not change between Feb-2020 and Feb-2021
- Alternative Hypothesis: Number of *very unhealthy* days in Karachi did change between Feb-2020 and Feb-2021

Figure 3 and 4 show the results of the hypothesis test for Lahore and Karachi. As the result shows we get a p value of 0 for Lahore which means we will reject the null hypothesis of Lahore in favor of its alternative hypothesis. The point estimate for Karachi is  $3 - 1 = 2$  which shows there has been a change but we have still failed to reject the null hypothesis of Karachi as the p value that we got is more than the significance level of 0.05.

Two-sample test of proportions

x: Number of obs = 29

y: Number of obs = 28

Variable	Mean	Std. Err.	z	P> z	[95% Conf. Interval]
x	.0344828	.033883			
y	.1071429	.0584512			
diff	-.0726601	.0675619			
	under Ho:	.0676788	-1.07	0.283	
diff = prop(x) - prop(y)			z = -1.0736		
Ho: diff = 0					
Ha: diff < 0		Ha: diff != 0		Ha: diff > 0	
Pr(Z < z) = 0.1415		Pr( Z  <  z ) = 0.2830		Pr(Z > z) = 0.8585	

Figure 3: Proportion Test for Karachi

Two-sample test of proportions

x: Number of obs = 29

y: Number of obs = 28

Variable	Mean	Std. Err.	z	P> z	[95% Conf. Interval]
x	.137931	.0640329			.0124289 .2634332
y	.7857143	.0775443			.6337302 .9376983
diff	-.6477833	.100565			-.8448871 -.4506794
	under Ho:	.131963	-4.91	0.000	
diff = prop(x) - prop(y)			z = -4.9088		
Ho: diff = 0					
Ha: diff < 0		Ha: diff != 0		Ha: diff > 0	
Pr(Z < z) = 0.0000		Pr( Z  <  z ) = 0.0000		Pr(Z > z) = 1.0000	

Figure 4: Proportion Test for Lahore

Statistically these results mean that we can be sure that Lahore has had more *very unhealthy* days in February 2021 when compared to February 2020. For Karachi since we failed to reject the null hypothesis we can not claim that the number of *very unhealthy* days did change between February 2020 and February 2021 but our point estimate tells us that there has been an increase in the number of days. The result of the test shows that statistically this point estimate is no different than 0.

For the dataset about gas and electricity consumption we can not perform regression or hypothesis

tests as the data is again quite limited. As the readings are simple addition making a regression model on it would just result in all the coefficients being 1 which does not help us understand much. Reading the trends in the data is enough to understand that the energy industry plays its part in the worsening air quality specially in Punjab.

## **5 Conclusion**

Our analysis showed us that Lahore has been facing constantly increasing AQI readings when compared to Karachi but we were unable to establish a causal connection between these readings and energy usages of Punjab because of limitations in data. For any future research if an organization is established which conducts regular AQI readings based on multiple pollutants then that would help a lot as there is a severe lack of such monitoring in Pakistan. It is quite obvious from the data that Punjab has a significantly higher energy usage when compared to Sindh, this implies that this could be the reason for a higher AQI reading. As this research could not be conducted completely because of several limitations other reasons could also be a factor in these high readings. One of them could be that since Karachi is a sea city it faces some advantage over Lahore in this regard. Another reason Lahore faces abundant air pollution could be the high level of brick kilns in Punjab. There is a very large industry of brick creation in Punjab which is not regularly maintained and creates many pollutants which are harmful to humans. Regardless of these reasons it is imperative that work needs to be done on sustainable energy in Pakistan and specifically in Punjab to improve the air pollution situation in this country.

## References

- [1] IQAir. (2020). 2019 World Air Quality Report. IQAir.
- [2] Mehmood, T., Ahmad, I., Bibi, S., Mustafa, B., and Ali, I. (2020). Insight into monsoon for shaping the air quality of Islamabad, Pakistan: Comparing the magnitude of health risk associated with PM10 and PM2.5 exposure. *Journal of the Air and Waste Management*, 1340-1355.
- [3] Qurratulain, Shams, T., and Khwaja, M. A. (2019). Assessment of Pakistan National Ambient Air Quality Standards (NAAQS'S). Sustainable Development Policy Institute.
- [4] Siddiqui, R. (2004). Energy and Economic Growth in Pakistan Voume 43 Number 2. *The Pakistan Development Review*, 175-200.
- [5] Asif, Muhammad., Saleh, Naila. (2019). Human Security and Energy Security: A Case Study of Pakistan. *Policy Perspectives Volume 16 Number 1*. 99-116
- [6] Development Statistics of Sindh, Punjab Development Statistics, Statistical Supplement 2018. (2020). Gas and Electricity Consumption of Sindh and Punjab. Retrieved from <https://opendata.com.pk/dataset/gas-and-electricity-consumption-of-sindh-and-punjab>
- [7] AQICN: The World Air Quality Index Project. (2008). The World Air Quality Index. Retrieved from <https://aqicn.org/data-platform/register/>