**TERM PROJECT – CSIS 3860**

**Topic: Air Quality Index (AQI) of Vietnamese cities/provinces**

**Data visualization of**

**AQI data in some cities/provinces of Vietnam**

Student name: Phu An Doan

Student ID: 300311334

1. **Introduction:**
2. **The visualization tool used:**

In this project, Tableau is used to make some visualizations to represent for the dataset I chose. This tool is chosen because this is the tool that I learned a lot from the class. One more reason is that Tableau is one of the most well-known tools for visualization which most of everyone know about. Therefore, before the semester started, I spent some time to learn about it via LinkedIn Learning. That’s why I am familiar with this tool rather than others among many visualization tools. In addition, I think Tableau is easy to use rather than other tools and I can use it to import maps and websites as well as actions and filters, which is very useful to make visualizations to represent for my dataset. From all above reasons, I decided to choose this tool for my term project.

1. **What is the dataset I chose and what are the reasons for my choice?**

The selected dataset includes 8 csv files (already converted to Excel workbook files in my submission as required) of some popular air quality indexes of 8 cities/provinces in my home country (Vietnam). They are some statistics recorded over several years, in which some are from 2014 to 2021 and some are from 2019 or 2020 until now.

I decided to choose this dataset because I am really interested in the issue that this dataset can depict. As I remember, before I came to Canada to study, in around 2018 or 2019, Vietnamese citizens raised the air pollution problem. They know about it through AirVisual mobile app, and this issue quickly became a scandal in Vietnam at that time. In some years ago, we always wore a mask when we went out to work, to play, or to study. However, the situation is not serious until AirVisual revealed some statistics about the air quality indexes in Vietnam. In addition, some people at two largest cities in Vietnam detected some air pollutants near where they live, and some newspaper also mention this problem. From that time, Vietnamese government started to research and figure out how they can solve the problem.

Being able to live in Canada and breath the fresh air is really a wonderful experience for me and many other Vietnamese people. In overall, the air pollution in Vietnam is not as serious as some other countries like China. However, two largest cities in Vietnam are considered as two among the most polluted cities in the world. Because Vietnam is a developing country with the high rate of urbanization, most people must make their living at two Vietnamese largest cities, which is really a big problem. People living there have only two options. They can move to rural areas to work, where it is difficult for them to make more money to help their family. Another option is that they must immigrate to another country like Canada to live. I am not sure if the government can solve the problem immediately. This situation encouraged me to research about the pollution problem in my country. That’s why I decided to discover the Air Quality Index of Vietnamese cities/provinces. In this report, the trend, the distribution, and the forecast of those indexes will be discussed by two dashboards of visualizations.

**Note:** This dataset comes from <https://aqicn.org/>

8 csv files are downloaded from this website, then they are put in 1 Excel file as 8 sheets. After that, Tableau is used to create a union of 8 worksheets to generate a dataset which is easy to use in this visualization tool.

How to get this dataset: download the data for each city/province each time by:

* Access <https://aqicn.org/data-platform/register>
* Scroll down to “Free Downloadable Database”
* Search the name of the city/province needed, click them
* Scroll down and click “Download the CSV data for <city name>”
* Enter the full name, email and the organization, then click “I agree…”
* Click “Submit”, then the file will be downloaded automatically.

In this dataset, there are many indexes. However, the dashboards only focus on two main pollutants which are PM25 and PM10. Especially for Ho Chi Minh City, there is only PM25 because of the limitation of the data sources. In general, when AQI is mentioned, it mainly means **PM25** index in this report.

1. **The visualization’s story:**
2. **The goals of the dashboards:**

As mentioned above, by these visualizations, there are 3 main points discussed: the trend of PM10 and PM25 indexes in some past years, which cities have the highest pollution indexes, and how these indexes will be in the future.

1. **Overall visualizations:**

Generally, there are two dashboards in the project’s Tableau workbook. The main dashboard demonstrates the overall situation in Vietnamese cities/provinces. From this dashboard, we can conclude the trend and the future forecast of indexes. On the other hand, the second dashboard describes the Air Quality Index (especially PM25) details of each city/province. The distribution of PM25 can be discussed by this dashboard. In addition, users can read the real-time Air Quality Index on a website imported to this dashboard.

**Screenshots**:

Main dashboard (when PM25 is selected):

Chart

Description automatically generated

Main dashboard (when PM10 is selected):

Chart

Description automatically generated

Second dashboard:

A picture containing chart

Description automatically generated

1. **Insights and Analysis:**

From the trend line visualization, we can see that both indexes have experienced an upward trend. Although the average number for both indexes are just moderate (not high), but the increasing trend of these pollutants are really a warning sign for Vietnamese governments and citizens.

From the monthly data visualization, the figure for AQI decreases every summer from June to September. In general, the figure is below 50 for this period, which is good for people’s health. But for the other months, the average number of AQI is over 50, which has moderate effects on people’s health. It is very lucky for us not to have over 100 for this figure, which is unhealthy for people. However, the interesting point here is the maximum AQI in May, which is up to over 800 for PM25, the magical figure. It can be because Vietnam has the tropical weather so in summer, it is very hot and polluted for some days.

From the Q4 visualization, we can see that 2018 is the year in which the indexes are the lowest one. After that, the figures are very high for 2019, and especially for 2020, the figure is over 40,000. It can be because of the high rate of urbanization and economic development in Vietnam in recent years. However, in Q4 of 2021, Vietnam has just experienced the fourth wave of Covid-19, people do not go outside much this time, so both indexes significantly decreased a lot this time.

Let’s move to the second dashboard. From the city map of Vietnam, it can be noticed that most of cities with high AQI are in the northern of Vietnam.

From the average PM25 chart, the most polluted cities/provinces in my home country are Ho Chi Minh City and Cao Bang, and Hanoi, Thai Nguyen for next positions. Quang Ninh is the province with the least number of indexes in the list. However, according to the sparkline chart, there is no data for Quang Ninh, Thai Nguyen, Cao Bang and Bac Ninh in several years in the past. The reason can be those cities/provinces are only polluted in recent years, which is not concerned much in the past. Therefore, the comparison can be not quite accurate.

1. **Solution and suggestion:**

In general, the situation is not very bad in Vietnam. All average indexes are at moderate for people’s health according to the general criteria of the world. However, the upward trend of these indexes in recent years should be concerned more by the government. In some months, the figure is extremely high, very dangerous for Vietnamese citizens. This year, Covid-19 lock down had a positive impact on the air quality, which is a good sign, but we cannot predict what happens after it is over, when people return to the normal life.

Urbanization and economic development are two of the reasons. The government should develop every area at the same rate and reduce the number of gasoline vehicles in two of the largest cities, so the situation can be better. Obviously, this situation cannot be improved in the short term. Vietnamese government can have many other solutions to get people to have more fresh air to breath in the future.