**Object:**

* Dựa vào dữ liệu khí tượng ở Việt Nam đã thu thập đc trong vòng hơn 10 năm (2011-2022) (source detail refer to: “PM25.pptx” page 4) đưa ra góc nhìn chung về kiểu khí hậu và biến đổi theo mùa.
* Đồng thời áp vào dữ liệu thu thập về mật độ hạt bụi mịn (PM 2.5) của khu vực Thành Phố Hồ Chí Minh để tiến hành phân loại, thống kê để có thế thấy đc độ tương quan (thuận, nghịch hoặc không tương quan) của từng yếu tố khí tượng đến sự phân bố của mật độ hạt bụi mịn trong không khí của khu vực.
* Từ đó có thể đề xuất các phương pháp giảm mật độ hạt bụi mịn (PM 2.5) trong không khí góp phần cải thiện chất lượng không khí, cải thiện môi trường sống.

**Tech:**

* Web data crawling: using BeatifulSoup, Selenium Webdriver
* Python language: library Pandas, data cleaning, data labeling (Pandas.qcut)
* Visualization: library Matplotlib.pyplot, library Seaborn

**Class’s Material Support:**

|  |  |  |  |
| --- | --- | --- | --- |
| 1st Semester | Python | Basic (list, dictionary, loops, define function, library Pandas) | Crawling web data, cleaning data, labeling data, output data to csv file ... |
| Statistic (Pandas.qcut, ...) | Transform data, prepare data for visualization |
| Visualization (library Matplotlib, library Seaborn) | Visualization data on Jupyter Notebook |
| Power Bi | Power Query (language M) | Generate table of datetime to support Crawling web data |