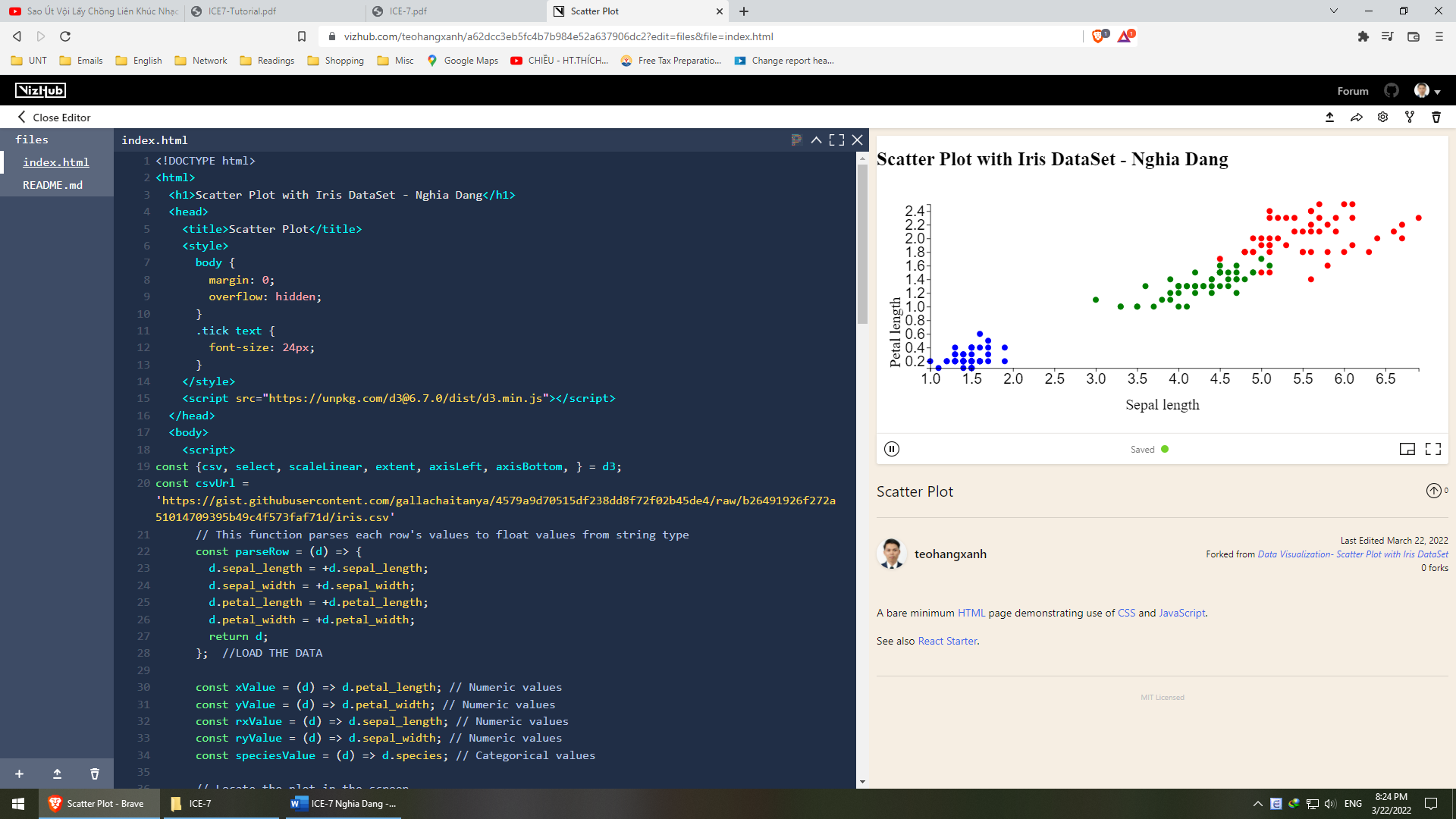
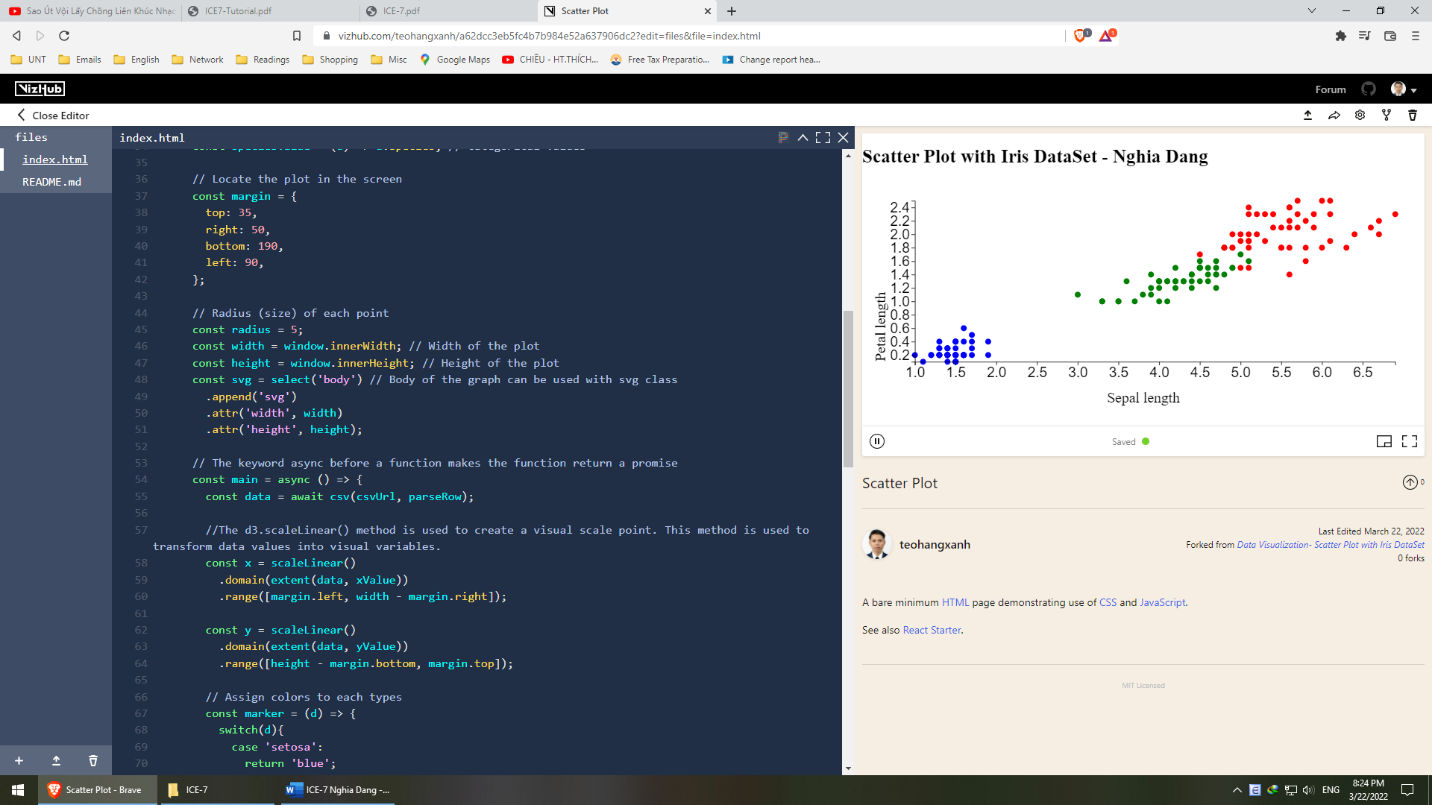
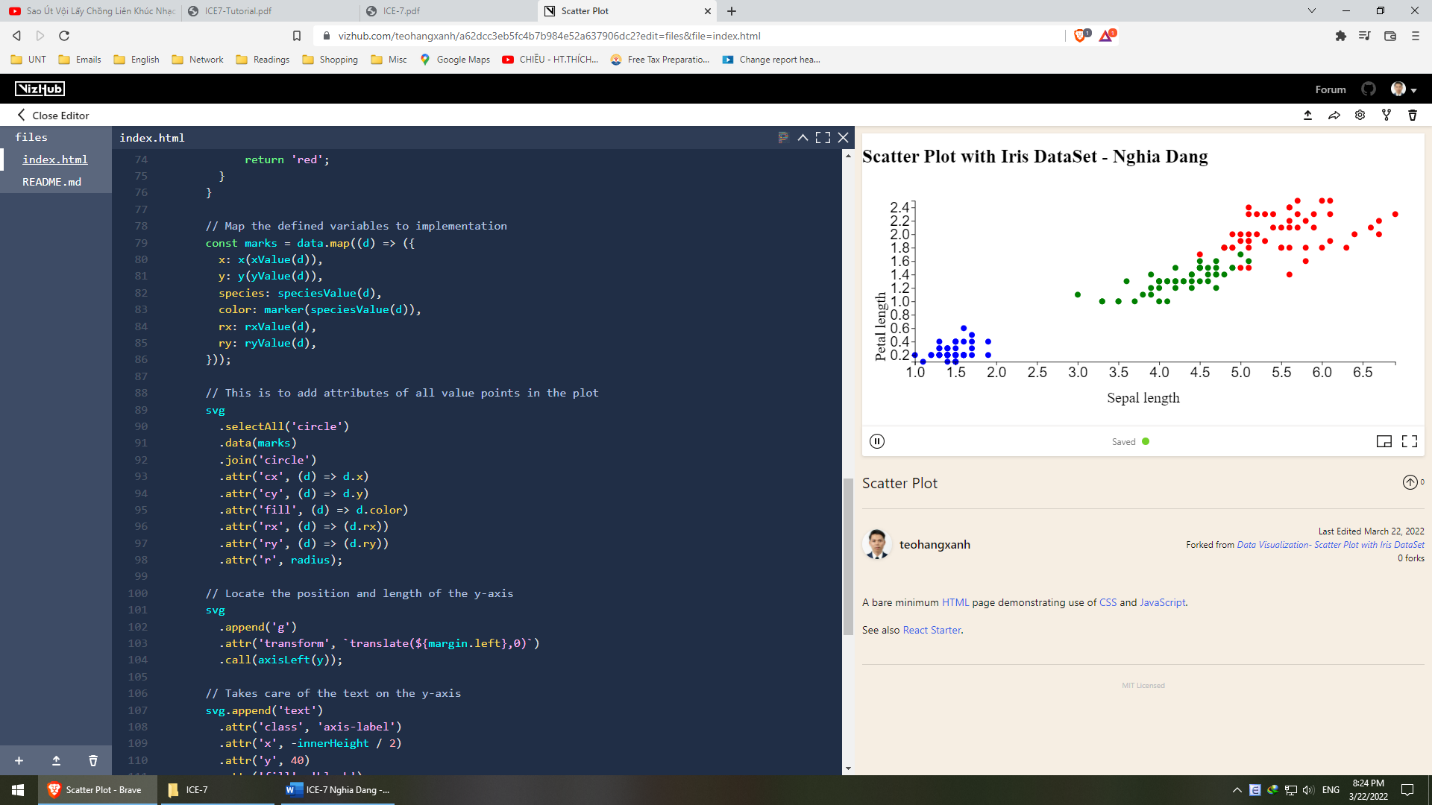
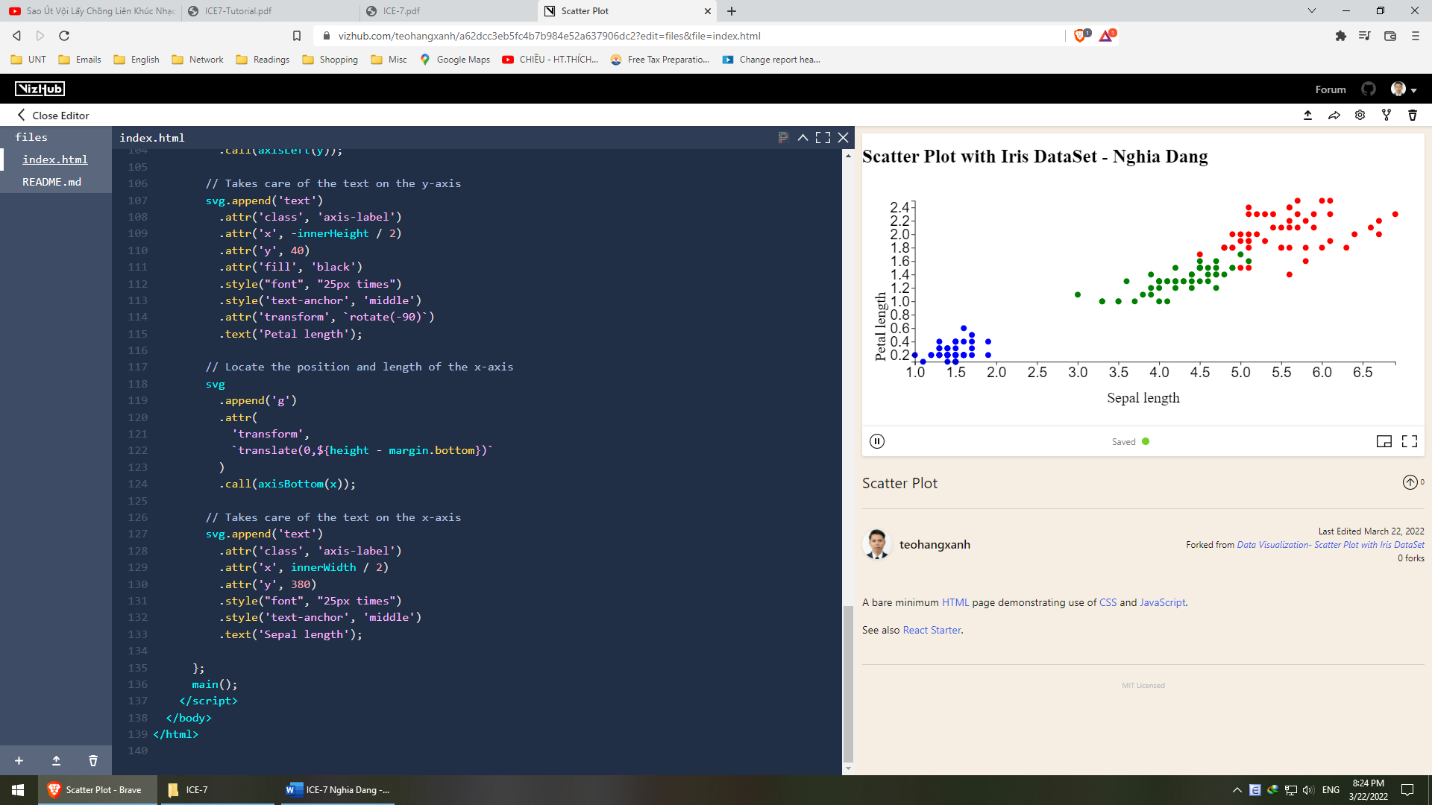
1. Scatter plot of Iris dataset









<https://vizhub.com/teohangxanh/a62dcc3eb5fc4b7b984e52a637906dc2?edit=files&file=index.html>

**Describe how different iris species distributed on your chart**:

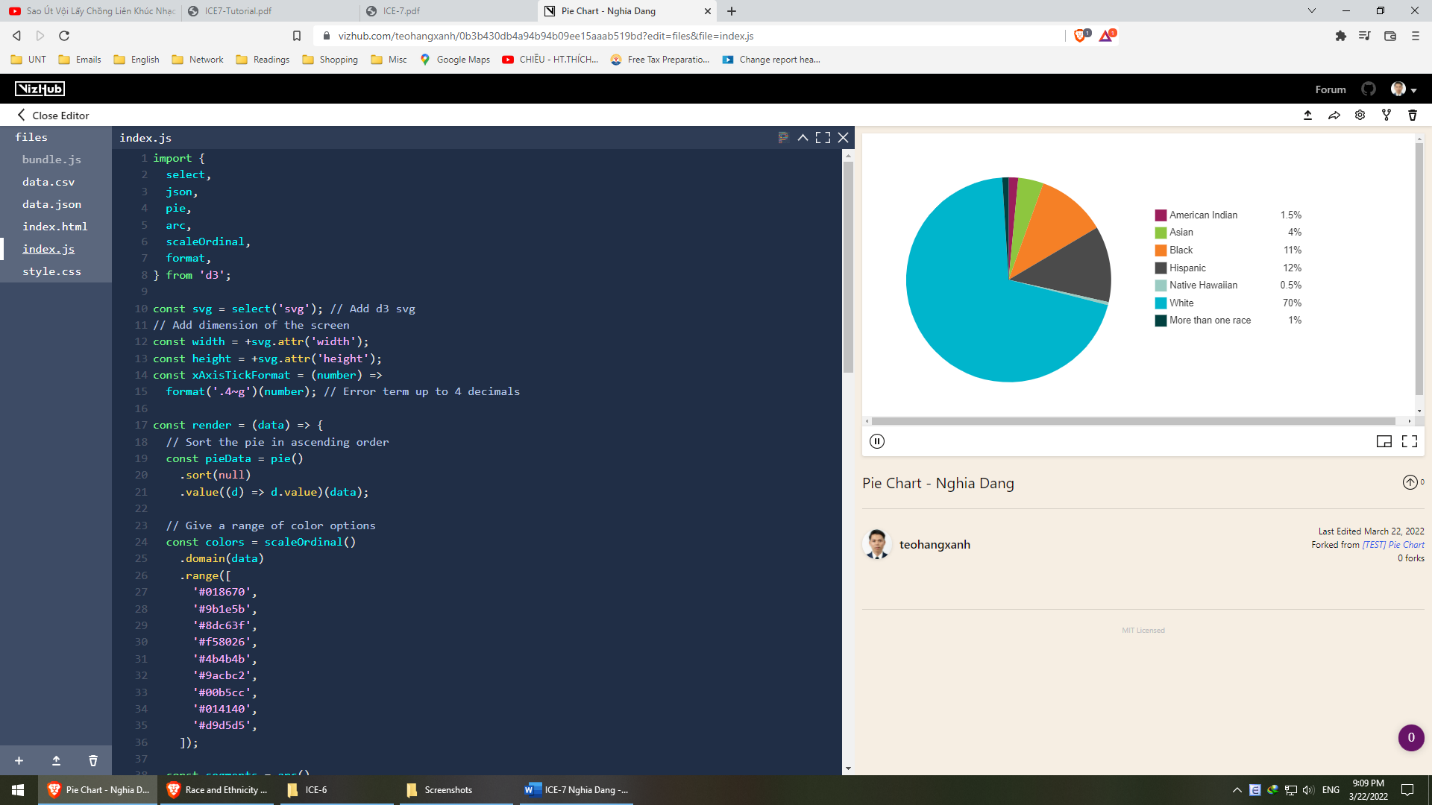
While it is easy to cluster Setosa since their sizes are fairly small against others, it is more complicated to clearly separate Versicolor and Virginica around (1.8, 5.0). Generally, Setosa ranges from (0, 0) to (2.0, 1.6), Versicolor ranges from (3.0, 1.2) to (5.0, 1.6) and Virginica ranges from (5.0, 1.6) to larger.

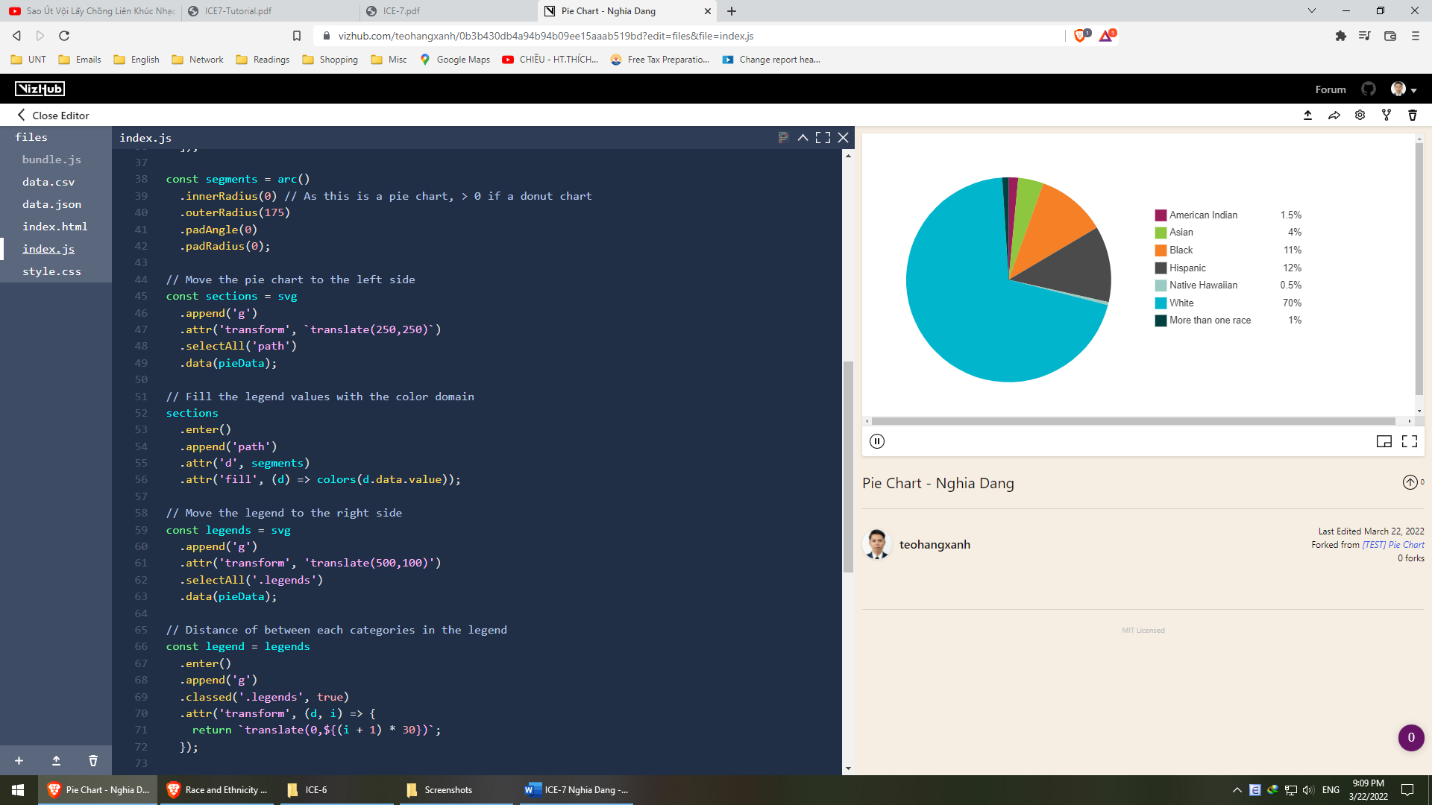
**Explain the reason why we need to build a reusable scatter plot for the Iris**

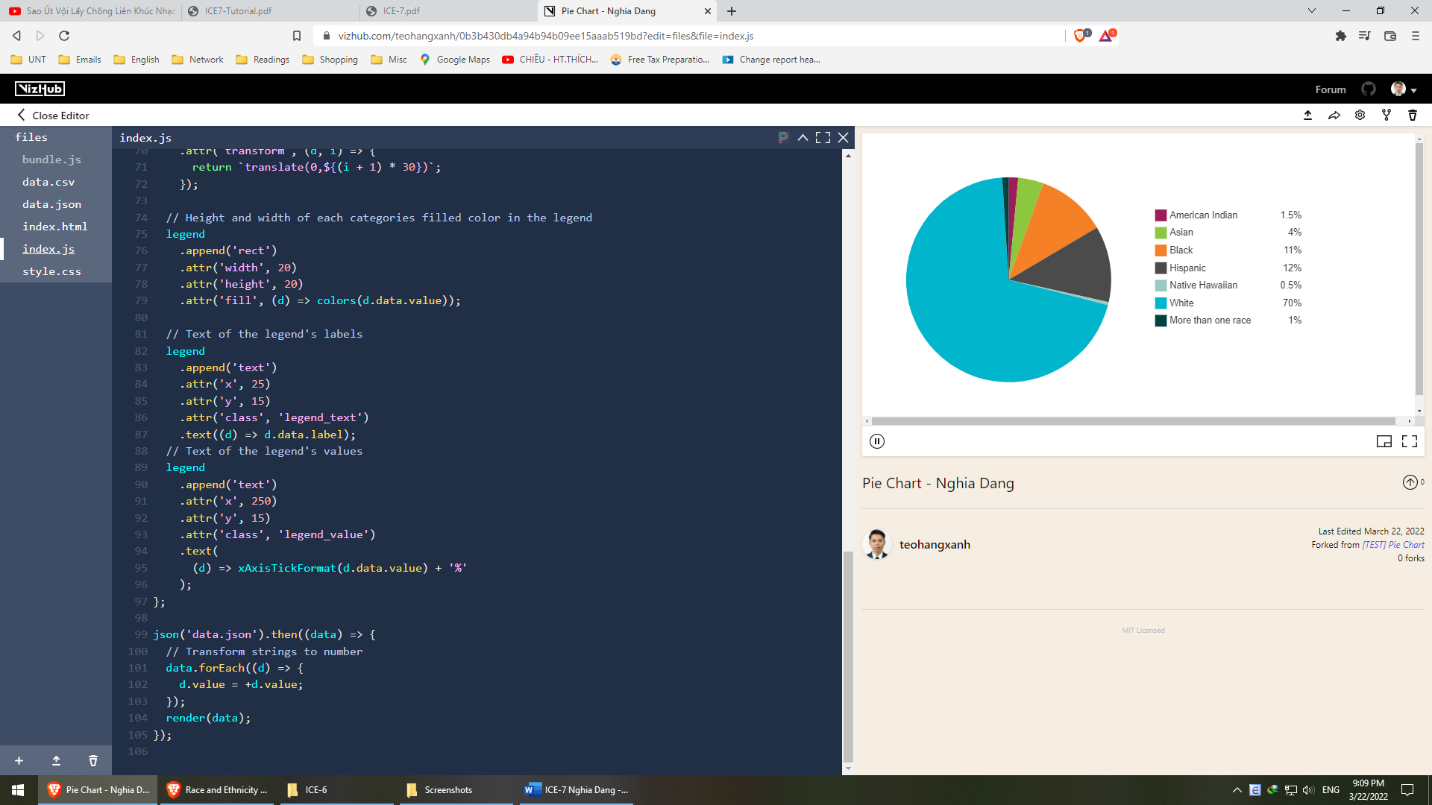
**data.**

It is useful in visualizing the difference among those species in terms of dimensions, which plays a significant role in species classification.

1. Pie chart







<https://vizhub.com/teohangxanh/0b3b430db4a94b94b09ee15aaab519bd?edit=files&file=index.js>

**Analyze the chart: provide explanation on the visualization. What are the pros and cons of using pie chart on this dataset?**

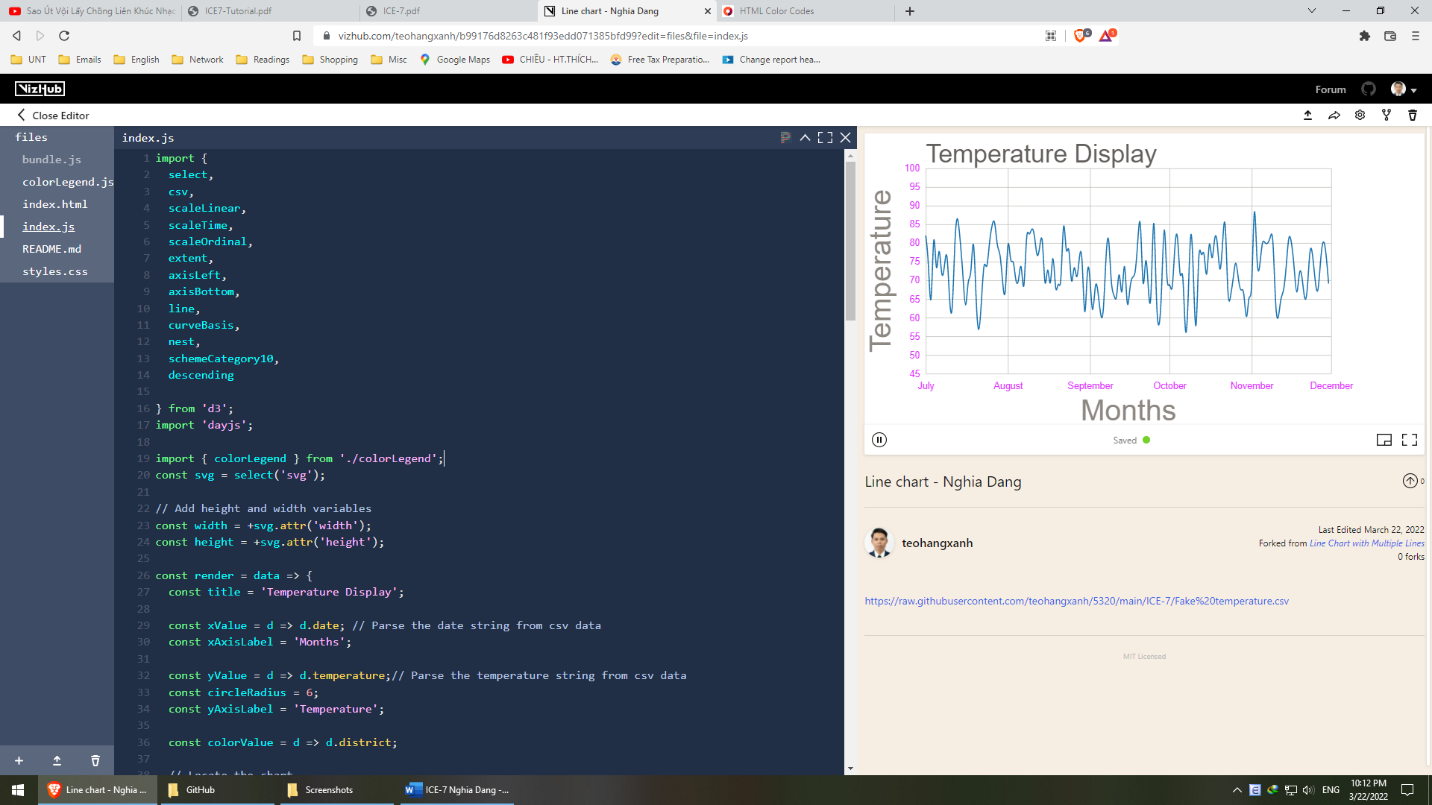
Pros:

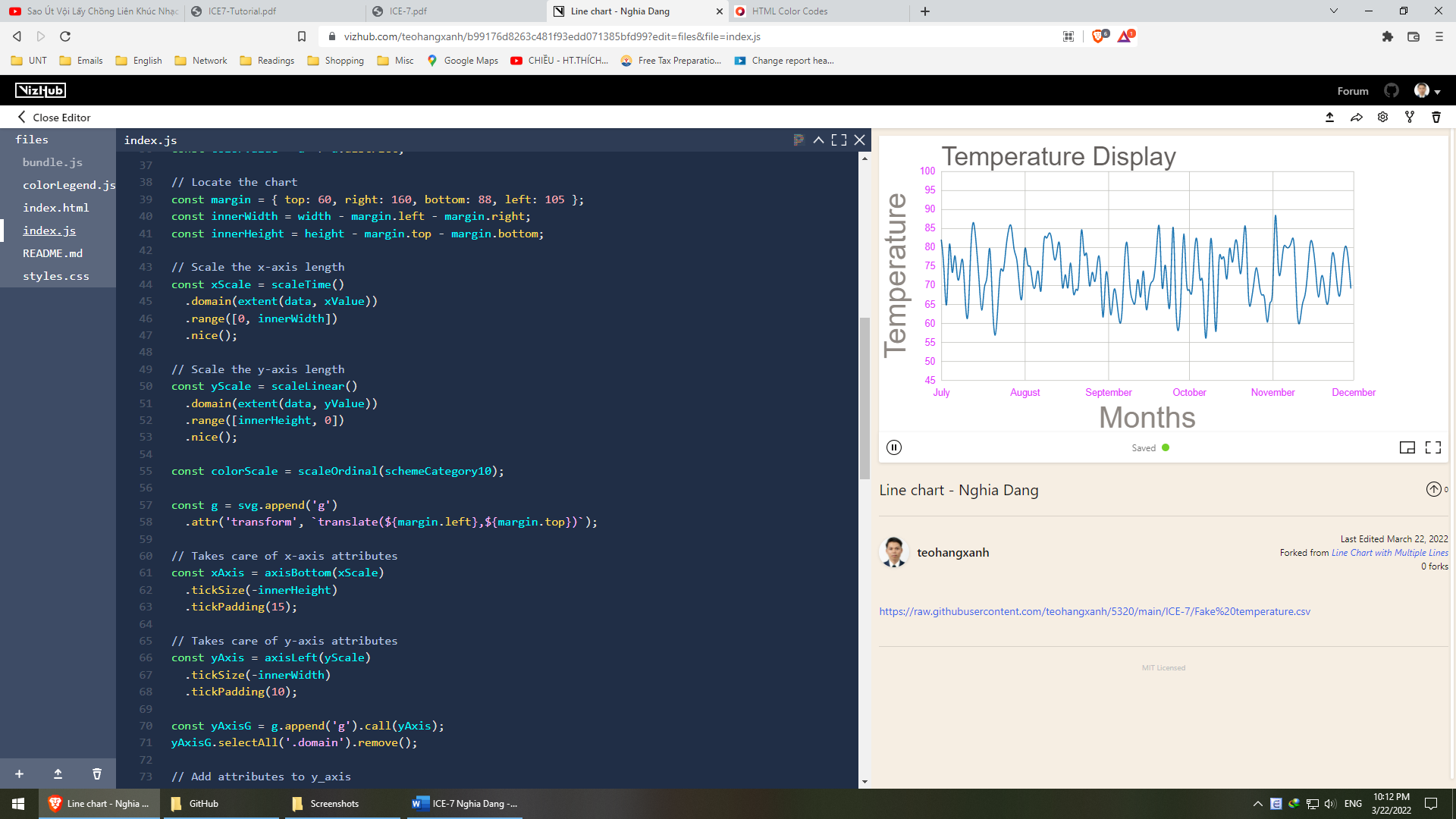
* It’s easy to see the big difference in data
* It can emphasize the data with few instances (less than 5)

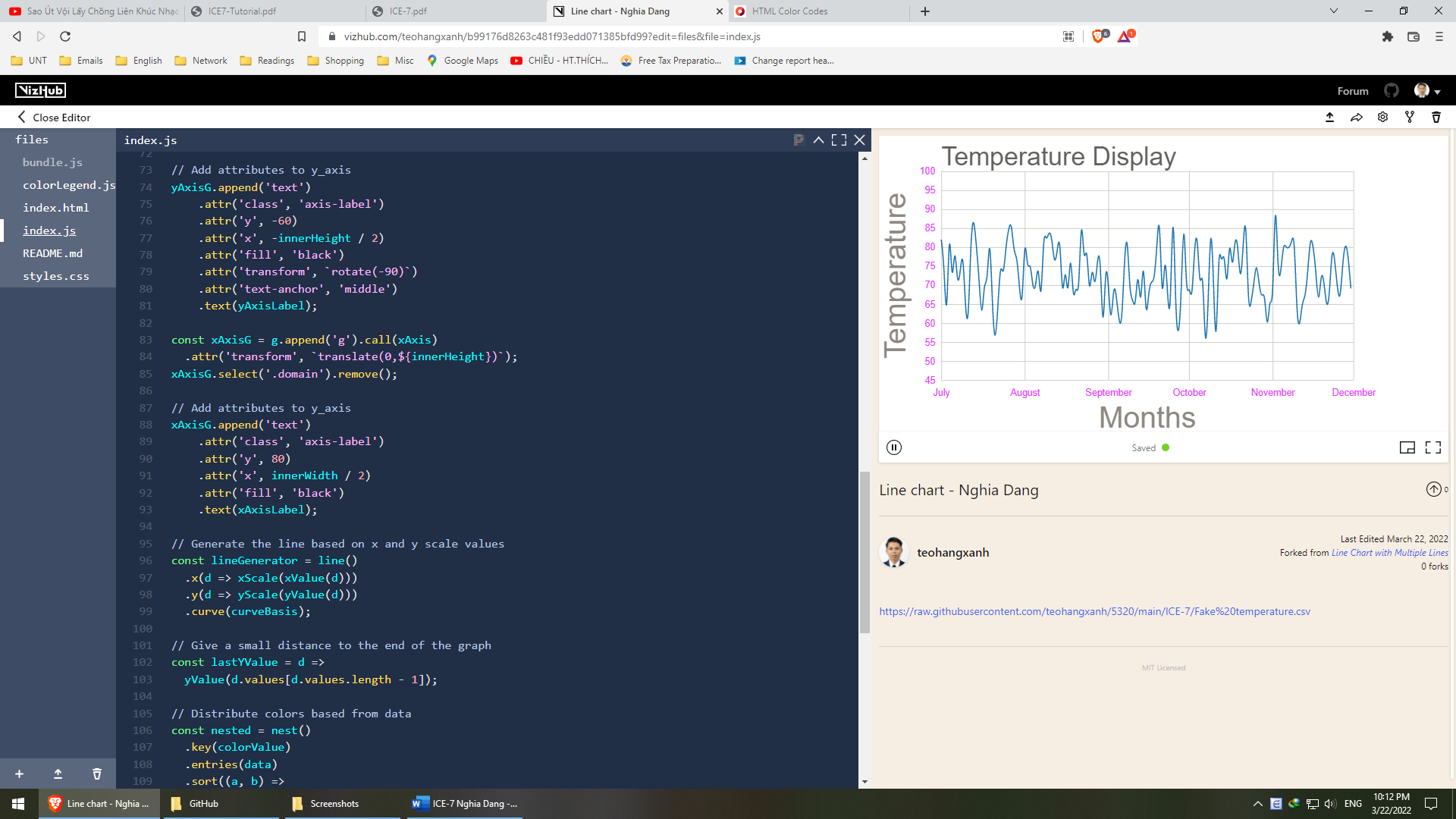
Cons:

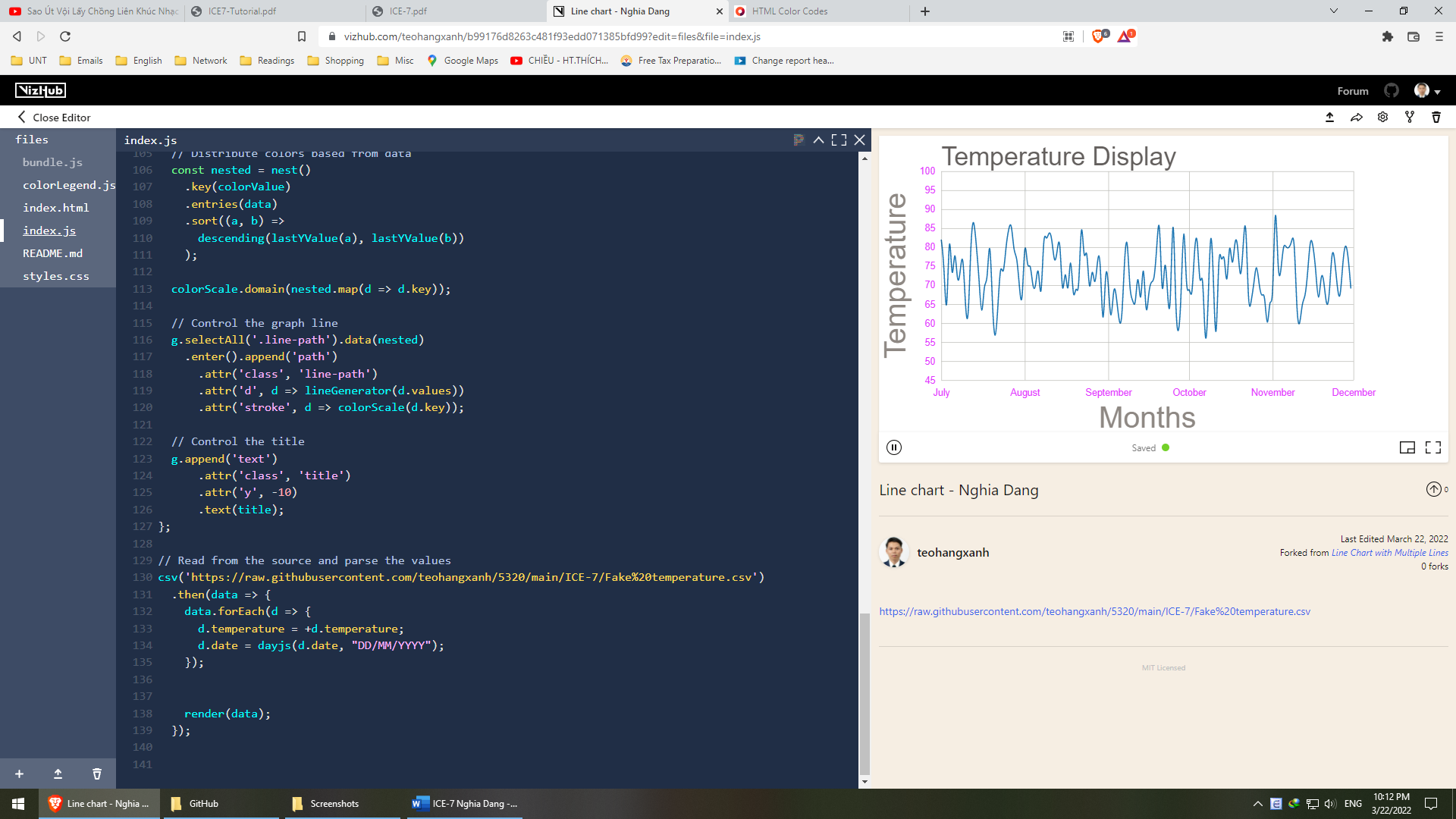
* It becomes less useful if instances start to become over 4
* It cannot show the trend of the data

1. Line chart









<https://vizhub.com/teohangxanh/b99176d8263c481f93edd071385bfd99?edit=files&file=index.js>

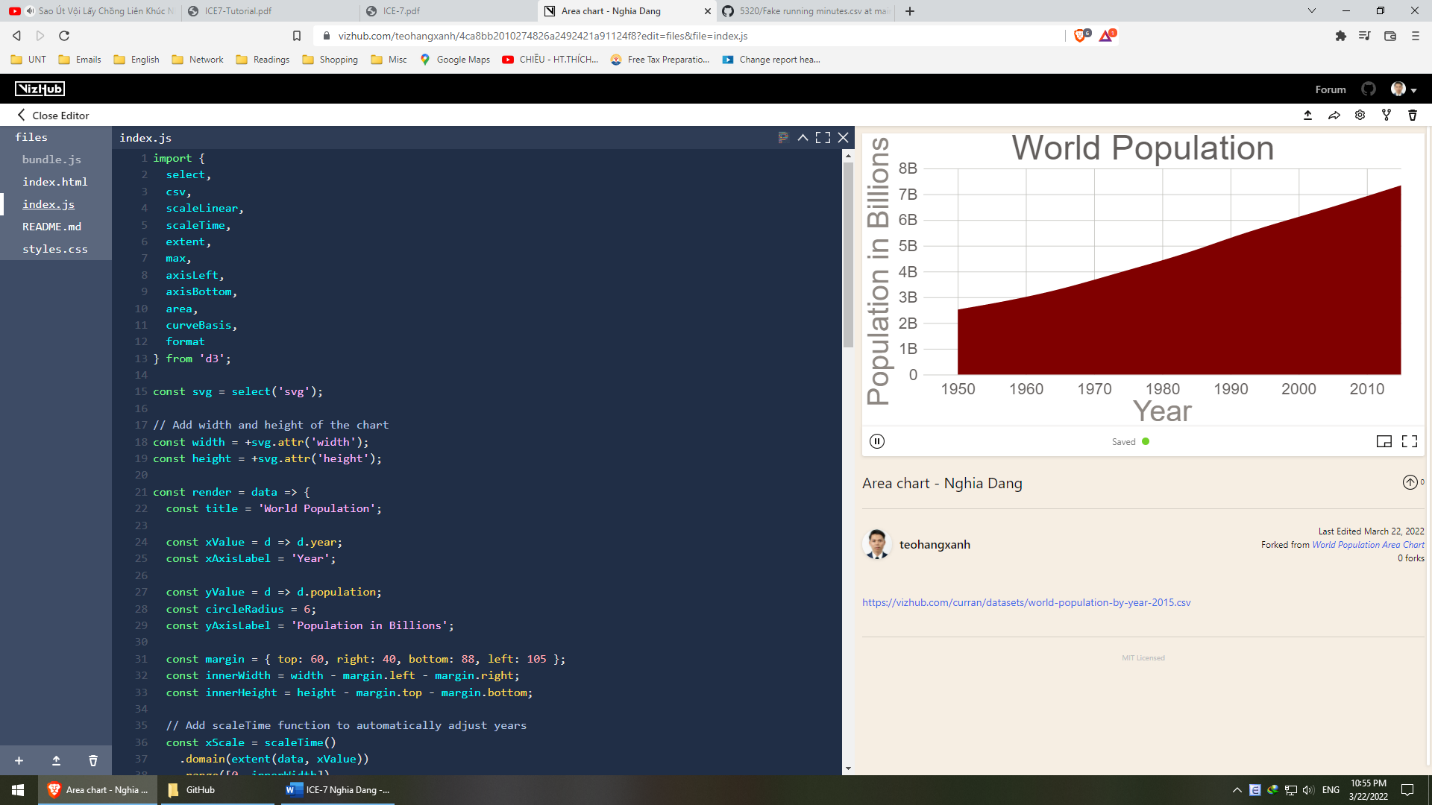
**Analyze the chart: provide explanation on the visualization. What are the pros and cons of using line chart on this dataset?**Pros:

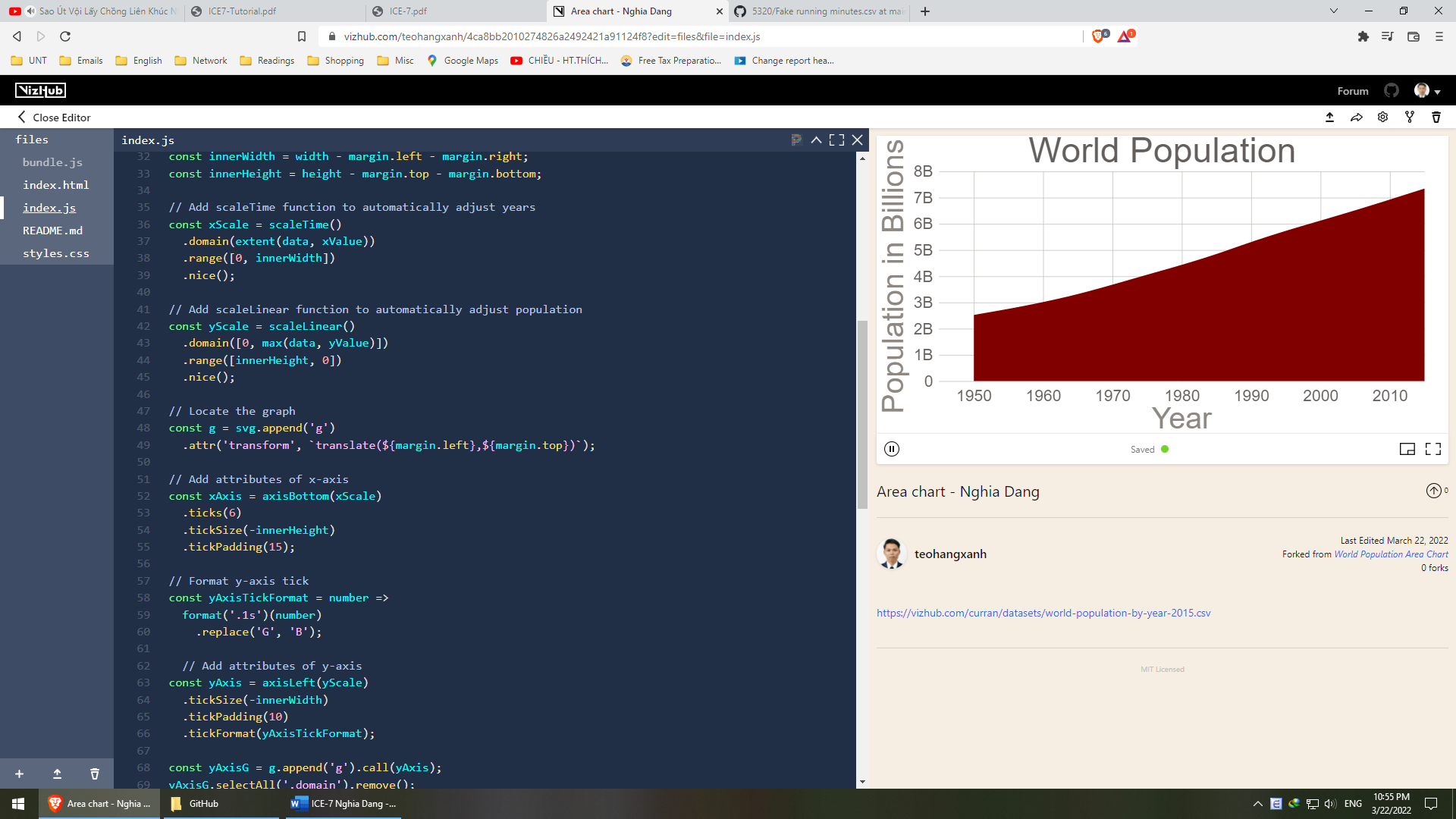
* It’s easy to point out the range, min, max, gaps, clusters
* It’s easy to tell the changes / differences over time

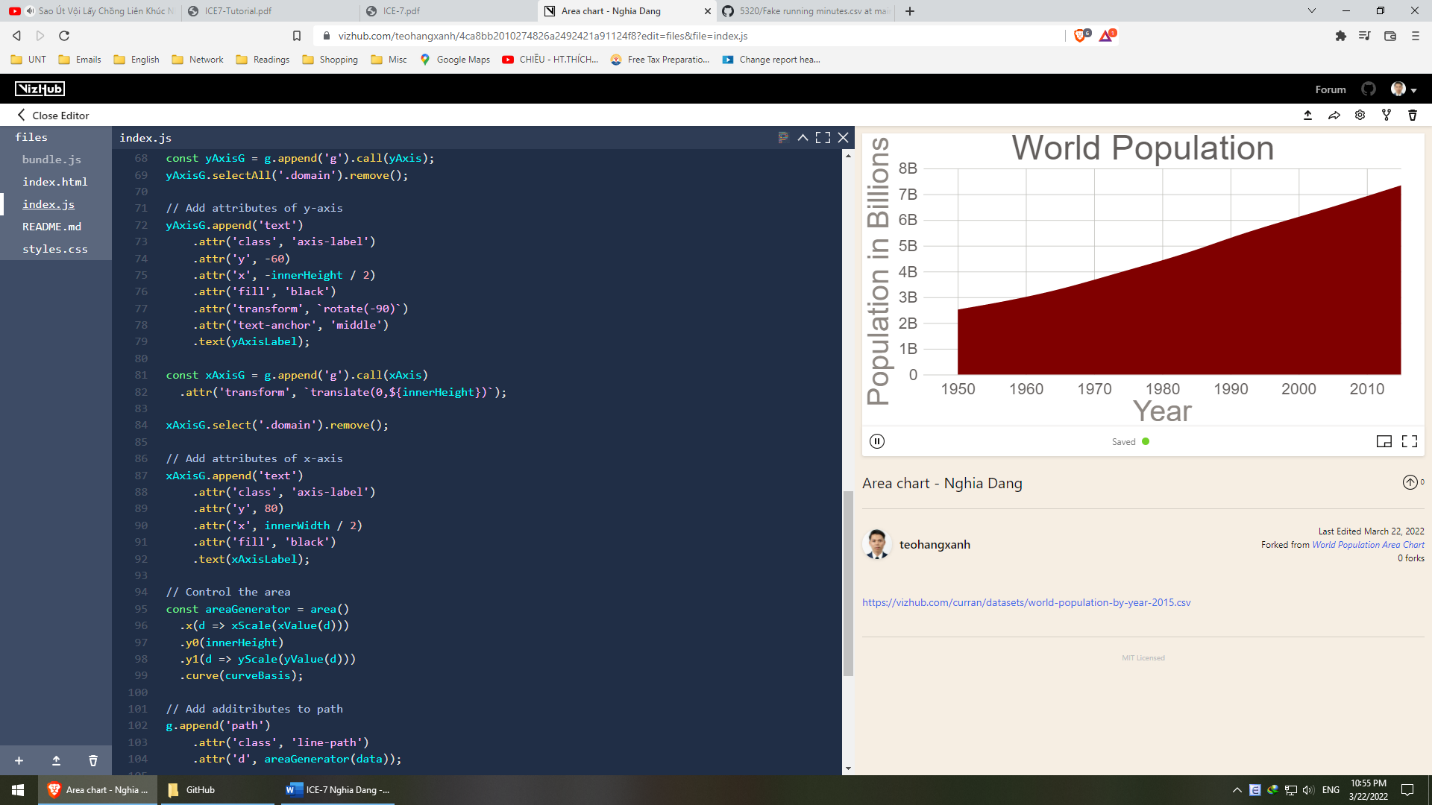
Cons:

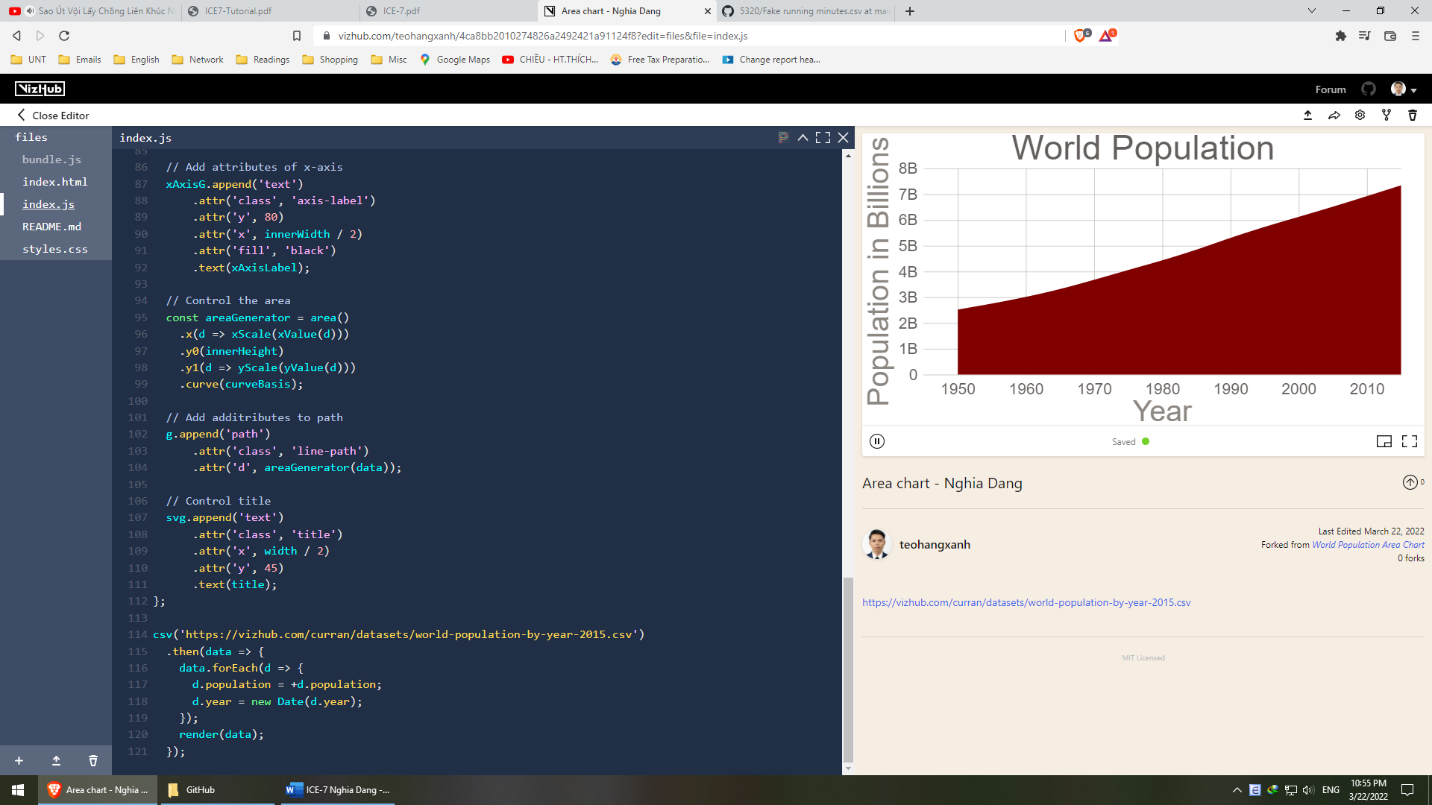
* Data values should not be over 50
* The discrepancy between data values should not be too large

1. Area chart









<https://vizhub.com/teohangxanh/4ca8bb2010274826a2492421a91124f8?edit=files&file=index.js>

**Analyze the chart: provide explanation on the visualization. What are the differences between area chart and line chart on this dataset?**

The area that are filled represents the values from the data where the no-filled area stands for the possible maximum values. We can clearly see the trend of data by year. The difference between area and line charts includes

* Line charts are better in comparison
* Line charts are better in showing distribution or trend
* If the area charts have multiple entities, they can show the accumulated values. However, it is nearly as same as line charts with this dataset.