

CSCE 5320 Scientific Data Visualization

Submission Guidelines:

1. Submission is individual.
2. Code must be properly commented.
3. **You are required to submit by answering questions (1 and 2) given at the end of the tutorial.**
4. **You need to submit the tutorial. It is for your reference(for both D3.js and Python).**
5. The documentation should include the screenshots and links of your code/results.
6. Submit the documentation (.pdf/.doc) with screenshots and Viz hub links and include the explanation.
7. Everyone should submit separate files for D3.JS and python libraries related Questions.
8. **Submit .pdf/.doc file for D3.JS in which you should include the explanation, screenshots of code, visualizations and vizhub links.**
9. **Submit the IPYNB file along with the PDF version of IPYNB for questions related to python libraries.**
10. The similarity score for your document should be less than 15%.
11. **There should be no AI plag for your documentation.**
12. Use Google Colab to complete the Python Tasks.
13. Give your last 2 digits of Student ID for the variable name as shown in the IPYNB file.
14. Submission after the deadline is considered as late submission.

D3.JS:

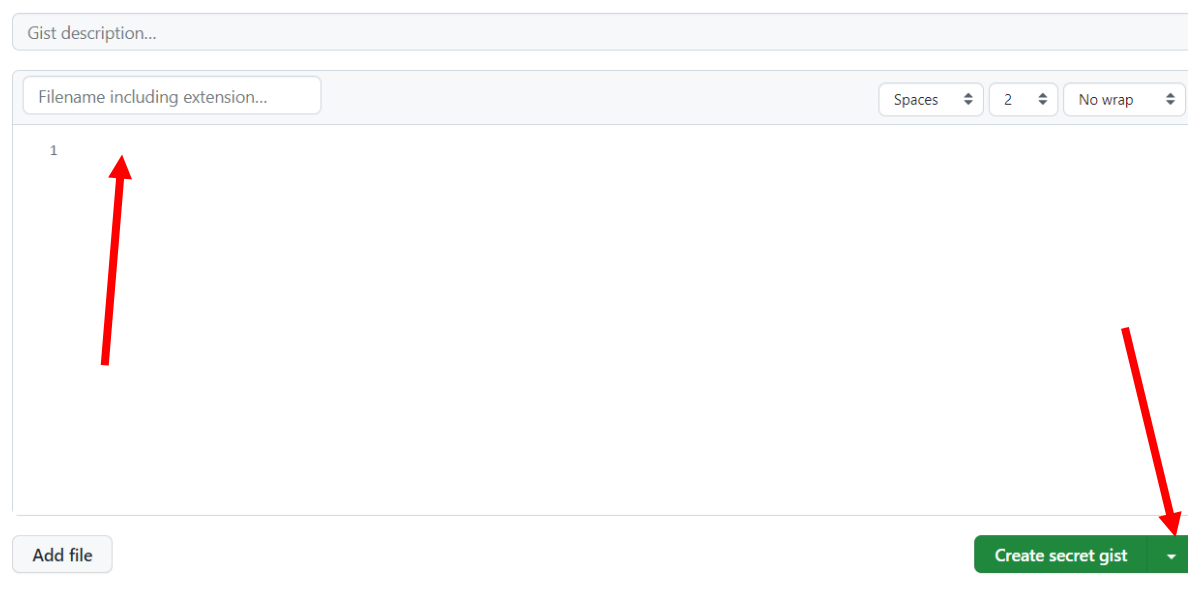
Tutorial:

You don't need to submit the tutorial. Just Practice the tutorial for your understanding.

Reusable Scatter Plot:

Download the given dataset named "Iris.csv" and create a dataset link by using [gist.github](https://gist.github.com)

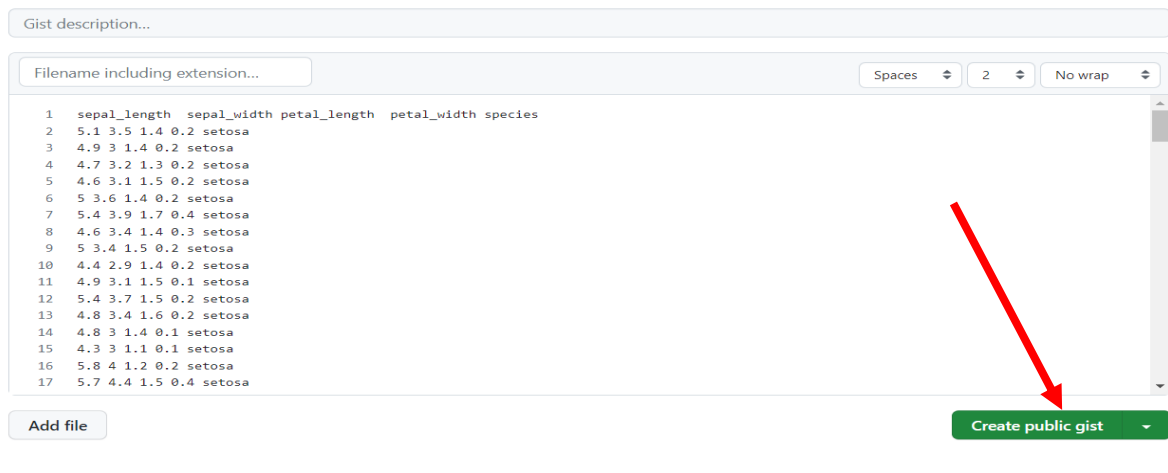
Go to <https://gist.github.com/>. Upload/copy the given csv file and create public gist.



The screenshot shows the GitHub Gist creation page. At the top is a text input for "Gist description...". Below it is a header bar with a "Filename including extension..." input, and dropdown menus for "Spaces" (set to 2), "2" (lines), and "No wrap" (line wrapping). The main area is a large text editor with line numbers on the left. A red arrow points to the first line of the editor. At the bottom left is an "Add file" button. At the bottom right is a green button labeled "Create secret gist" with a downward arrow icon.

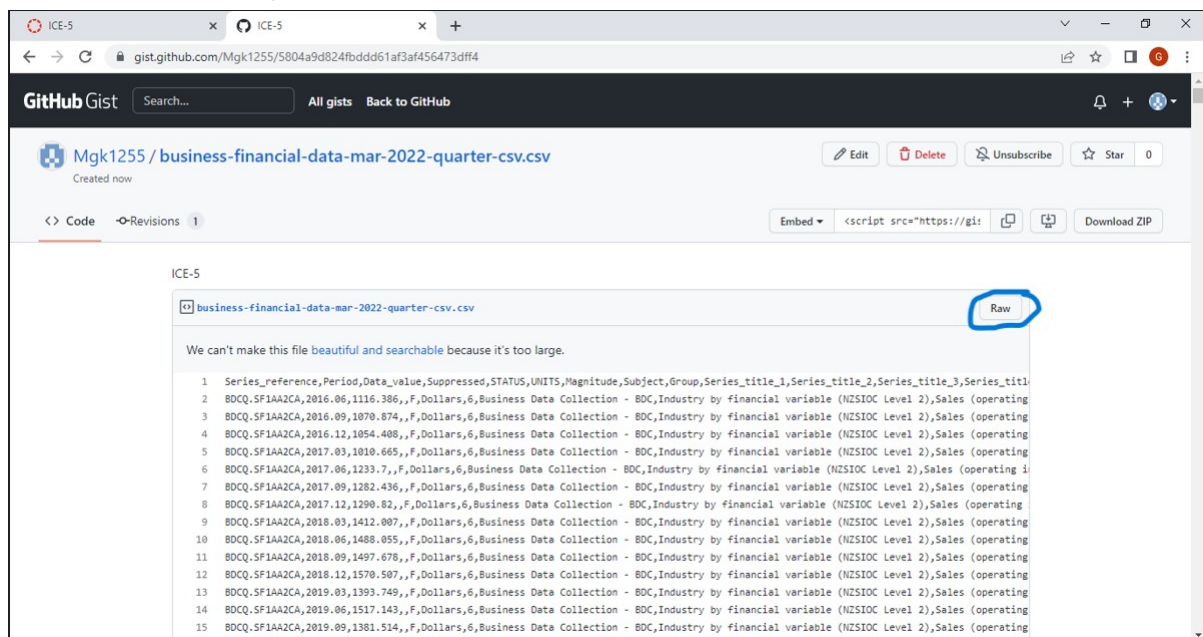
Select the entire data and copy it in gist hub workspace. Click on the downward arrow and select create public gist and click on it.

We are selecting as create public gist by default it will show as the secret gist. We need to select it as create as public gist once after clicking on it.

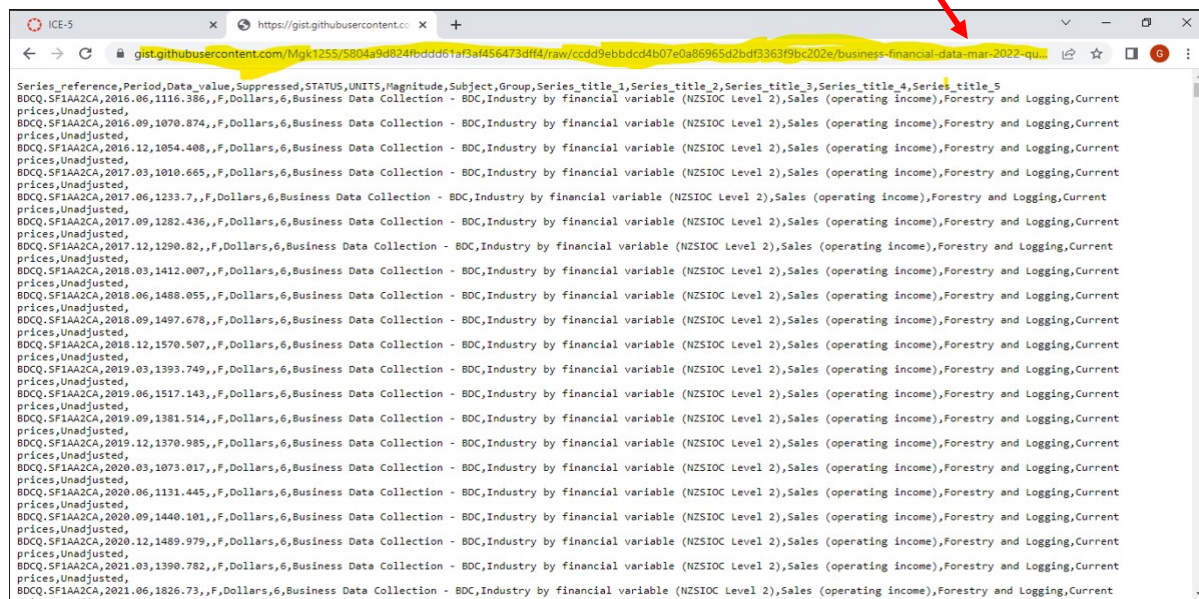


Once after creating the gist we can able to see as a button as raw once if we click on the button raw we can able to see there a link will be created and in that link we can see the entire data that was present in the csv dataset.

Once after clicking on the raw button there will be a link will be created that was



the gist link that was been created once after the creating as below, we can able to see that the data will be available as below. Here we can access the entire data by using the link.



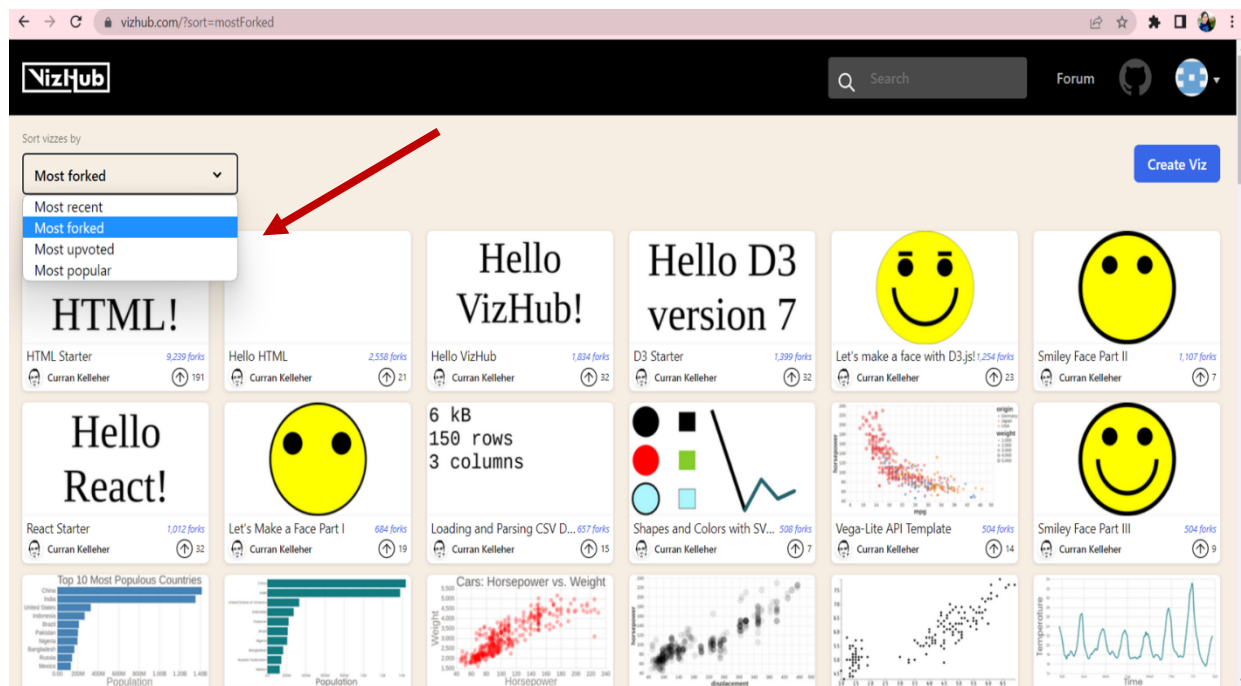
```
Series reference,Period,Data value,Suppressed,STATUS,UNITS,Magnitude,Subject,Group,Series_title_1,Series_title_2,Series_title_3,Series_title_4,Series_title_5
BDCQ.SF1AA2CA,2016.06,1116.386,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2016.09,1070.874,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2016.12,1054.408,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2017.03,1010.665,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2017.06,1233.7,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2017.09,1282.436,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2017.12,1290.82,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2018.03,1412.007,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2018.06,1488.055,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2018.09,1497.678,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2018.12,1570.507,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2019.03,1393.749,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2019.06,1517.143,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2019.09,1381.514,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2019.12,1370.985,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2020.03,1073.017,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2020.06,1131.445,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2020.09,1440.101,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2020.12,1489.979,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2021.03,1390.782,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
prices,Unadjusted,
BDCQ.SF1AA2CA,2021.06,1826.73,,F,Dollars,6,Business Data Collection - BDC,Industry by financial variable (NZSIOC Level 2),Sales (operating income),Forestry and Logging,Current
```

Below is the data link that will be generated :

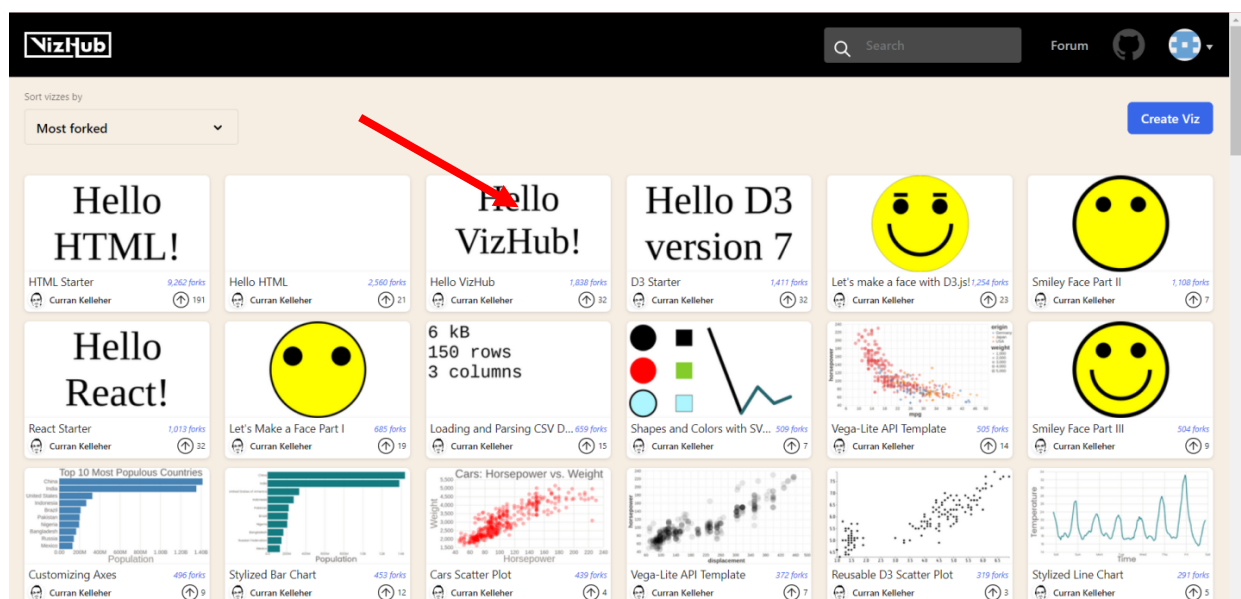
<https://gist.githubusercontent.com/INDU1998-GIT/27d356a6eb8513e0146a91a997d27aa6/raw/ba3c2ada2696991aacf6cb2323b97b8d2c40e46e/iris.csv>

Here you can create your own data link by using your own gist hub account with the given csv file to display the reusable scatter plot.

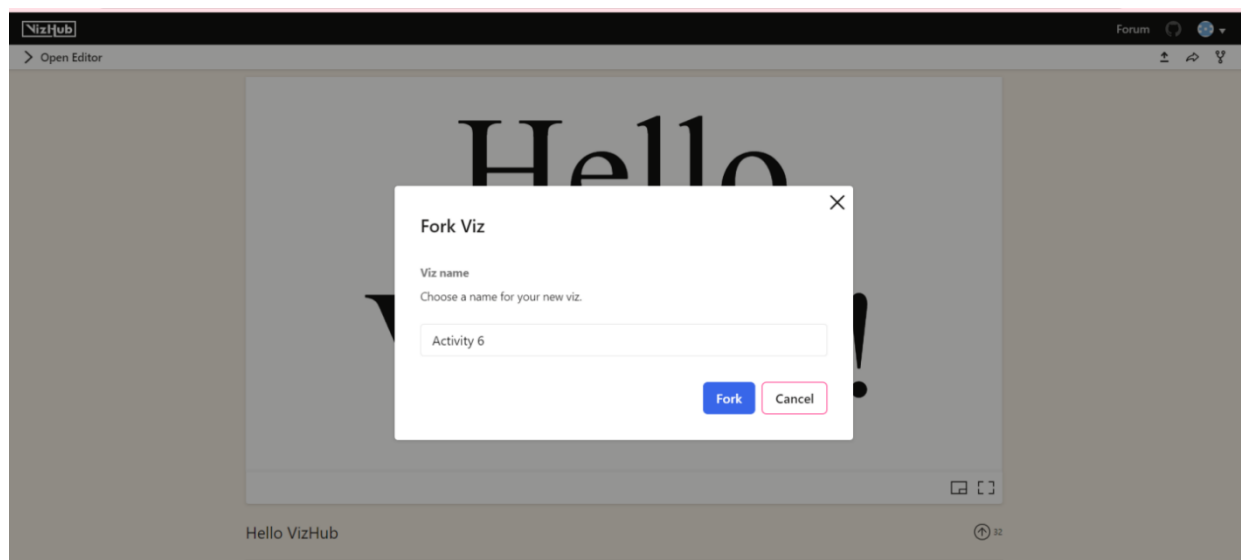
Now Click on <https://vizhub.com/>, log in to it and connect it with your github. Then select a project from most forked/ most popular.



Select “Hello VizHub!” project or any other most forked project.



Click on fork symbol in order to save your changes.



You can give any another name and click on fork.

Then click on open editor, and add index.html, index.js, package. Json, scatterplot.js and styles.css files and add the code in the files accordingly such that it should display reusable scatter plot with X label as sepal width and Y label as sepal length.

Here in the screenshot below, we can see it displays the reusable scatter plot by using iris dataset.



Understanding the dataset:

➤ The dataset contains of the fields namely: sepal length, sepal width, petal length, petal width with different kind of species names.

➤ Totally there are 151 rows information about different kinds of species and they are specified using different colors:

- Setosa
- Versicolor
- Virginica

➤ So we consider values of sepal length, sepal width, petal length, petal width while plotting the graph.

Understanding the Scatter plot:

➤ The graph is color coded according to the species, that is each of the species is given a color.

- Setosa - light blue
- Versicolor - light green
- Virginica - orange

➤ We have labeled both the x-axis and y-axis according to the requirement and also gave an heading to the graph.

➤ From the graph we can come to conclusion of whereabouts of the different species settled on

the scatter plot graph:

- Setosa species are mostly settled in the region between 0 to 0.6 of

petal length and spread through 1.0 to 2.0 of sepal length.

- Versicolor species are mostly settled in the region between 1 to 1.6

of petal length and spread through 3.0 to 5.2 of sepal length.

- Virginica species are mostly settled in the region between 1.4 to

2.4 of petal length and spread through 4.5 to 6.8 of sepal length.

VIZHUB LINK:

<https://vizhub.com/INDU1998-GIT/f47109d26bb242a5825ada34ce89c9da?edit=files&file=index.js>

Please refer to the above vizhub link for any doubts.

Questions:

Question 1(25%)	Points
<ol style="list-style-type: none">1. Create a Pie Chart for any suitable dataset by using d3.JS dataset. Display the legend icons next to the chart and include the labels on X-axis and y-axis.2. Explain about the dataset you have chosen in terms of data types and attribute types and submit the dataset link. Make sure to use the dataset link while generating the visualization in Vizhub.3. Analyze the chart and provide an appropriate description for the generated visualization.4. Submit the screenshots of visualization and your code and provide the VizHub link to your code.	<ul style="list-style-type: none">• Creation of Pie chart with labels and legends in vizhub.(8 Marks)• Explanation about the dataset and creation of the dataset link by using Git hub or Gist hub. (7 Marks)• Explanation of the generated visualization (5 Marks)• Screenshots and Viz hub link (5 marks)
	Total 25 Marks

Question 2(25%)	Points
<p>5. Create an Area Chart for any suitable dataset by using d3.JS dataset. Add the labels on X-axis and y-axis.</p> <p>6. Explain about the dataset you have chosen in terms of data types and attribute types and submit the dataset link. Make sure to use the dataset link while generating the visualization in Vizhub.</p> <p>7. Analyze the chart and provide an appropriate description for the generated visualization.</p> <p>8. Submit the screenshots of final generated visualization and code and provide the VizHub link to your code.</p>	<ul style="list-style-type: none"> • Creation of Area chart with labels in viz hub. (8 Marks) • Explanation about the dataset and creation of the dataset link by using Git hub or Gist hub. (7 Marks) • Explanation of the generated visualization (5 Marks) • Screenshots of generated Visualization and code and Viz hub link (5 marks)
	Total 25 Marks