

(Tutorial) Dataset Link:

<https://raw.githubusercontent.com/nikkiray309/sdv-act6/refs/heads/main/iris.csv>

The screenshot shows a web browser window with the URL <https://vizhub.com/nikkiray309/b7766d485b43452a89d58ce8a23d3f65?edit=files>. The page displays a code editor for 'index.js' and a scatter plot titled 'Iris Scatterplot'. The code in 'index.js' imports 'csv' and 'select' from 'd3', and 'scatterPlot' from './scatterPlot'. It defines a 'csvUrl' pointing to the raw GitHub file, a 'parseRow' function, and a 'main' function that fetches the data and renders the scatter plot. The scatter plot shows 'Sepal Length' on the y-axis and 'Petal Width' on the x-axis, with data points colored by species: Setosa (blue), Versicolour (green), and Virginica (orange). The plot is titled 'Iris Scatterplot' and has a legend for 'Species'.

Vizhub link: <https://vizhub.com/nikkiray309/b7766d485b43452a89d58ce8a23d3f65>

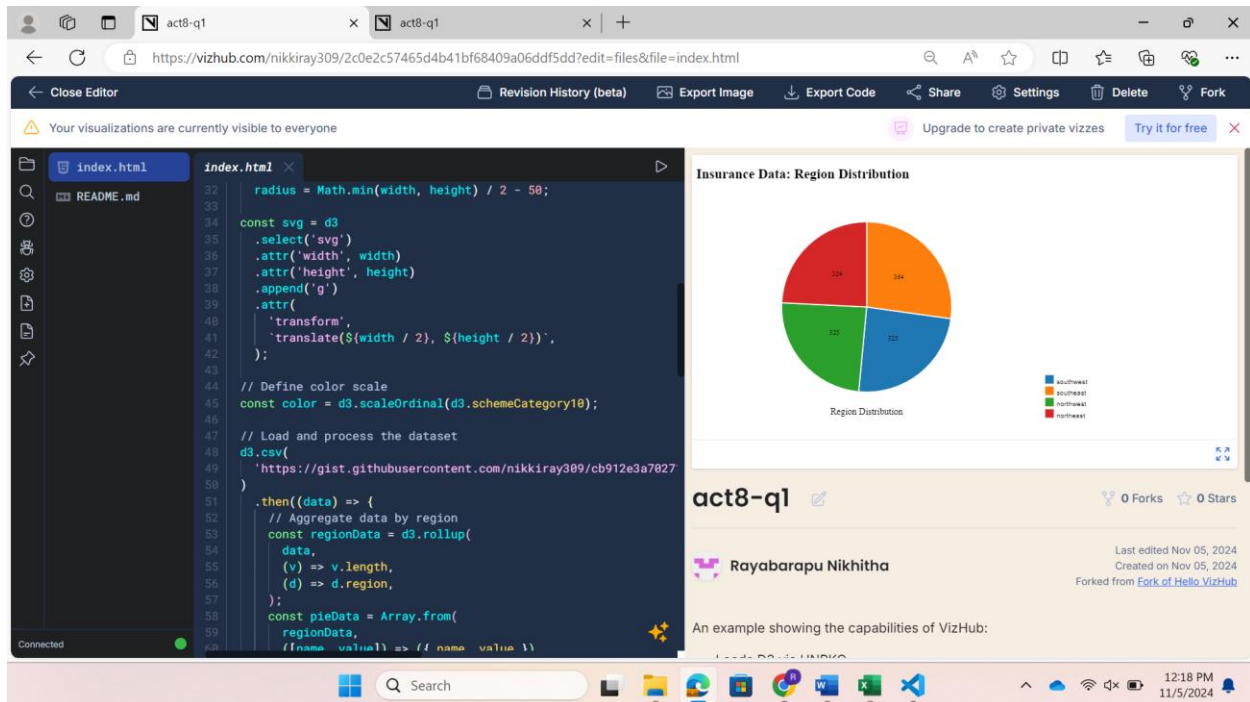
Question 1:

The screenshot shows a web browser window with the URL <https://gist.githubusercontent.com/nikkiray309/cb912e3a70271ab1df380b95f858c7f1/raw/7a60c8fec1613c615025ecffc21d3cdcbd13be29/insurance.csv>. The page displays a list of data rows for the 'insurance.csv' dataset. The columns are: age, sex, bmi, children, smoker, region, charges. The data rows are as follows:

age	sex	bmi	children	smoker	region	charges
19	female	27.9	0	yes	southwest	16884.924
18	male	33.77	1	no	southeast	1725.5523
28	male	33.3	no	southeast		4449.462
33	male	22.705	0	no	northwest	21984.47061
32	male	28.88	0	no	northwest	3866.8552
31	female	25.74	0	no	southeast	3756.6216
46	female	33.44	1	no	southeast	8240.5896
37	female	27.74	3	no	northwest	7281.5056
37	male	29.83	2	no	northeast	6406.4107
60	female	25.84	0	no	northwest	28923.13692
25	male	26.22	0	no	northeast	2721.3208
62	female	26.29	0	yes	southeast	27808.7251
23	male	34.4	0	no	southwest	1826.843
56	female	39.82	0	no	southeast	11090.7178
27	male	42.13	0	yes	southeast	39611.7577
19	male	24.6	1	no	southwest	1837.237
52	female	30.78	1	no	northeast	10797.3362
23	male	23.845	0	no	northeast	2395.17155
56	male	40.3	0	no	southwest	10602.385
30	male	35.3	0	yes	southwest	36837.467
60	female	36.005	0	no	northeast	13228.84695
30	female	32.4	1	no	southwest	4149.736
18	male	34.1	0	no	southeast	1137.011
34	female	31.92	1	yes	northeast	37701.8768
27	male	28.935	2	no	southwest	6303.00175

Dataset Link:

<https://gist.githubusercontent.com/nikkiray309/cb912e3a70271ab1df380b95f858c7f1/raw/7a60c8fec1613c615025ecffc21d3cdcbd13be29/insurance.csv>



Vizhub link: <https://vizhub.com/nikkiray309/2c0e2c57465d4b41bf68409a06ddf5dd>

Dataset Explanation:

The dataset contains the following columns:

- age:** Age of the policy holder.
- sex:** Gender of the policy holder.
- bmi:** Body Mass Index.
- children:** Number of children.
- smoker:** Smoking or not?
- region:** Residential region.
- charges:** Insurance charges.

Chart Explanation:

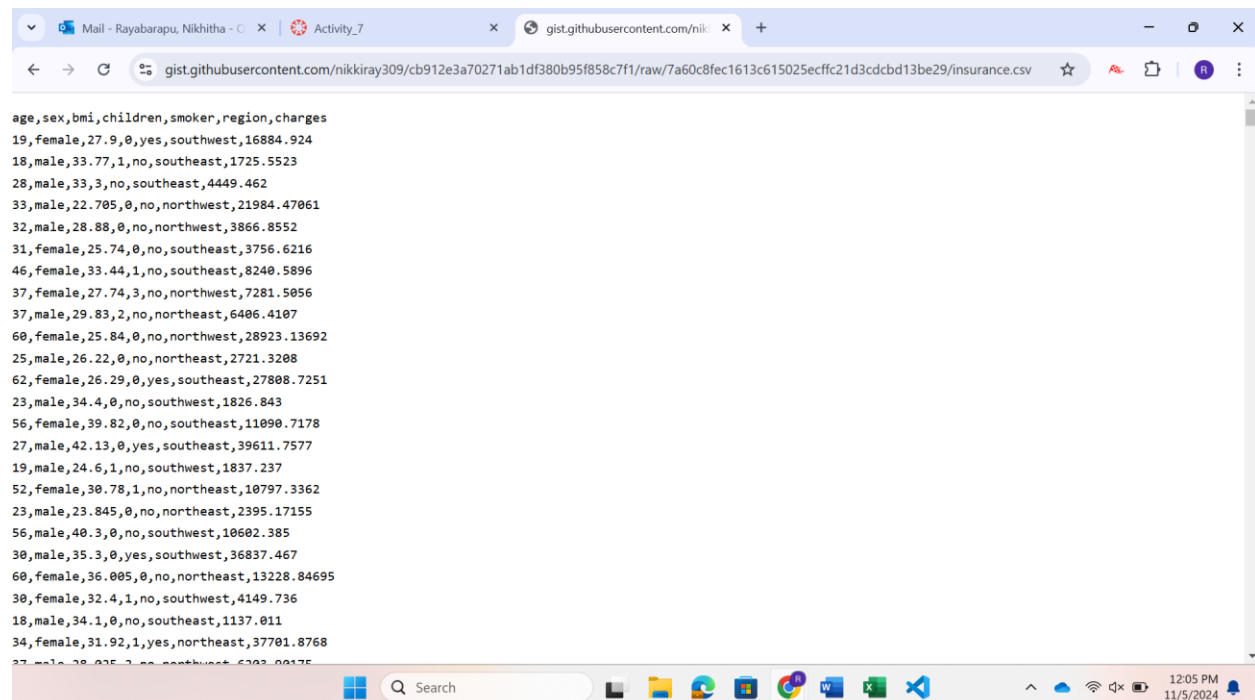
The pie chart shows the distribution of insurance charges across four different regions namely - southwest, southeast, northwest, and northeast. The southeast region has done

highest charges With a total of 364 records. The northeast has the lowest count with 324 records. The southwest and northwest have similar number of records.

Question 2:

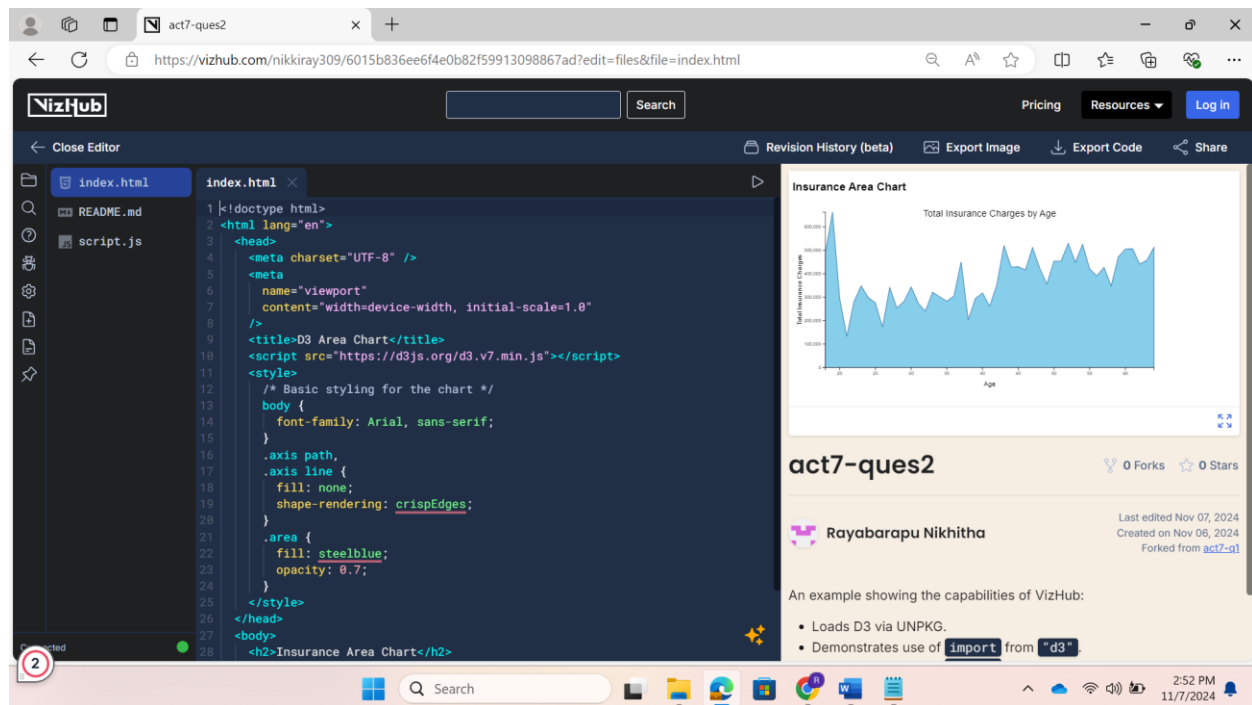
Dataset Link:

<https://gist.githubusercontent.com/nikkiray309/cb912e3a70271ab1df380b95f858c7f1/raw/7a60c8fec1613c615025ecffc21d3cdcbd13be29/insurance.csv>



```
age,sex,bmi,children,smoker,region,charges
19,female,27.9,0,yes,southwest,16884.924
18,male,33.77,1,no,southeast,1725.5523
28,male,33,3,no,southeast,4449.462
33,male,22.785,0,no,northwest,21984.47061
32,male,28.88,0,no,northwest,3866.8552
31,female,25.74,0,no,southeast,3756.6216
46,female,33.44,1,no,southeast,8240.5896
37,female,27.74,3,no,northwest,7281.5056
37,male,29.83,2,no,northeast,6406.4107
60,female,25.84,0,no,northwest,28923.13692
25,male,26.22,0,no,northeast,2721.3208
62,female,26.29,0,yes,southeast,27808.7251
23,male,34.4,0,no,southwest,1826.843
56,female,39.82,0,no,southeast,11090.7178
27,male,42.13,0,yes,southeast,39611.7577
19,male,24.6,1,no,southwest,1837.237
52,female,30.78,1,no,northeast,10797.3362
23,male,23.845,0,no,northeast,2395.17155
56,male,40.3,0,no,southwest,10602.385
30,male,35.3,0,yes,southwest,36837.467
60,female,36.005,0,no,northeast,13228.84695
30,female,32.4,1,no,southwest,4149.736
18,male,34.1,0,no,southeast,1137.011
34,female,31.92,1,yes,northeast,37701.8768
27,male,38.035,2,no,southwest,6203.00175
```

Vizhub Link: <https://vizhub.com/nikkiray309/6015b836ee6f4e0b82f59913098867ad>



Dataset Explanation:

The dataset contains the following columns:

- age:** Age of the policy holder.
- sex:** Gender of the policy holder.
- bmi:** Body Mass Index.
- children:** Number of children.
- smoker:** Smoking or not?
- region:** Residential region.
- charges:** Insurance charges.

Chart Explanation:

The chart displays the relation between insurance charges and age of policyholders. The chart suggests that charges are particularly high at younger ages, but it also tend to increase again in older ages. This relates to risk assessments in insurance calculations, with younger and older groups having higher risk.