gn417126 - Christian Gang Liu - B000415613

**Assignment 4**

on

**CSCI 6610 Visual Analytics**

Submitted by

**Christian Gang Liu (B000415613/gn417126@dal.ca)**

**…………………………**

In partial fulfillment of the requirements for the Course CSCI 6610 Human Computer Interaction

**Computer Science**

**2019-11-09**

# DEPARTMENT OF Computer Science

**Present to Teaching Assistant : Mateus Pereira; Leonardo Milhomem Franco Christino**

**&**

**Professor : Fernando Paulovich, PhD**



# DEPARTMENT OF Computer Science

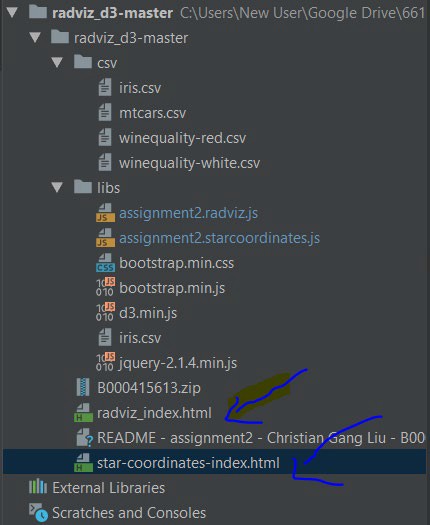
Implement a RadViz OR Star Coordinates visualization [1] using the D3 library. Your implementation should use HTML, CSS, and JS only! Do not use any other library besides D3 (https://d3js.org/).

Use one of the wine datasets provided in brightspace: winequality-red.csv or winequality- white.csv You can make these assumptions about the dataset:

* First row will contain the attribute names
* Last column is a class or regression (prediction) and can be numeric or nominal.
* All other columns are numeric.

How to execute program:

Using webstorm, directly run html, or execute npm http-server start for other IDE:





Code will be marked based on functionality, structure, reusability, best practices, and documentation.

1. [10 Marks] Create a backend that will provide the data and metadata that can be used to display the visualization

a. Use an HTTP request to retrieve data from the back-end and use it to generate the

visualization on the front-end.

b. Tip: You can return, along with the data, some metadata like column names or other

information that could be useful to handle and/or display the data in the front-end.

2. [20 Marks] Add an option on the interface to choose a different dataset (iris or winequality)

a. The backend will return the new dataset

b. You can keep a state on the front-end and send it on every request to identify the current

dataset in use.

3. [20 Marks] When hovering an instance of a given cluster, show (as a tooltip or in other

available space) the correlation matrix for instances of that cluster.

a. The correlation matrix should be calculated and returned by the back-end.

b. This should be displayed on the front-end through a color matrix. See example of such

matrix below (Tip: feel free to use libraries to help you):

4. [40 Marks] Implement a button that requests the backend to clusterize the data using one of:

K-Means or DBScan

a. You should color the instances using the clustering information

b. Add a switch button to choose between the color modes: cluster colors or class-based

colors.

c. The clusterization should be performed on the same dataset currently seem in the

visualization.

d. You just need to implement for one clusterization algorithm.

e. You may use existing implementations of the clustering algorithms.

5. [10 Marks] Add one (or more) options to configure the parameters of the clustering algorithm

a. Clicking the button should make a new clusterization with the new parameters and

update the colors on the visualization.

6. [+30 Bonus Marks] Add an option on the interface to choose to see the preprocessed dataset

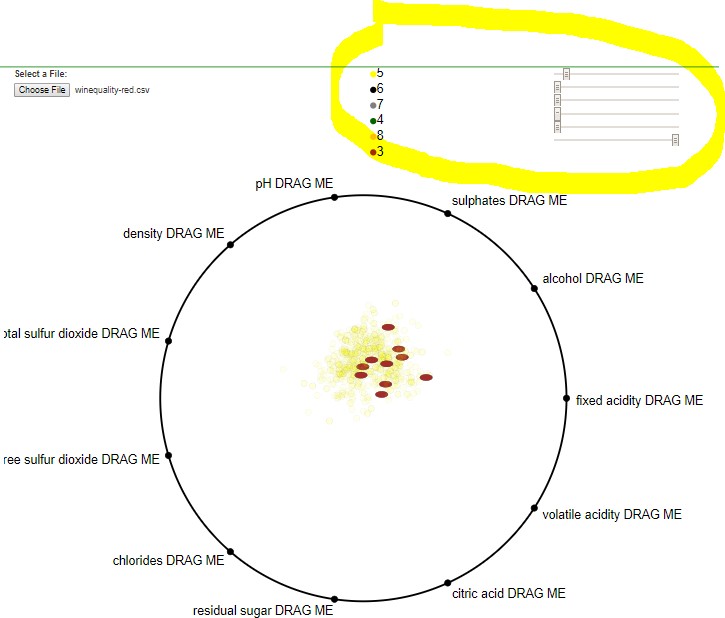
generated by your A1 assignment.

a. RadViz/StarCoordinates should only show the numerical columns as anchor points

b. The Categorical columns should be shown as the color of the plot. Make an input box

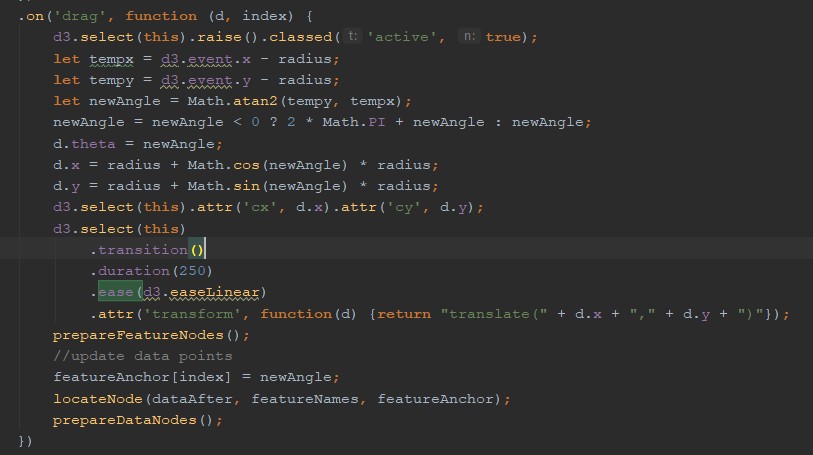
selector in the interface to choose the categorical column to be shown as the color.





2. [15 Marks] Smooth transition while dragging anchors (animation).

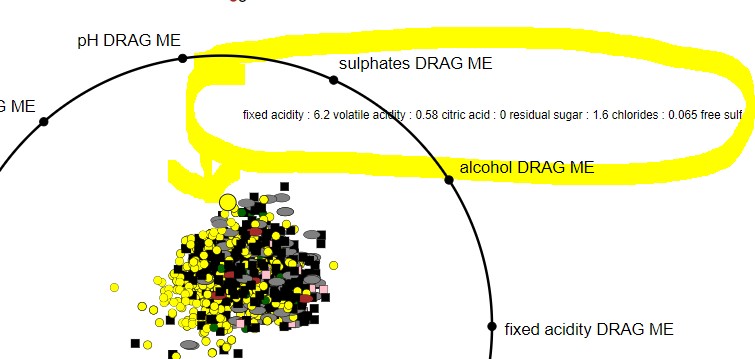
I leveraged “transition” and “transform” to achieve the drag action smoothly.



1. [15 Marks] Add some more unique features. Ideas:
   1. Add more information to nodes (tooltip).

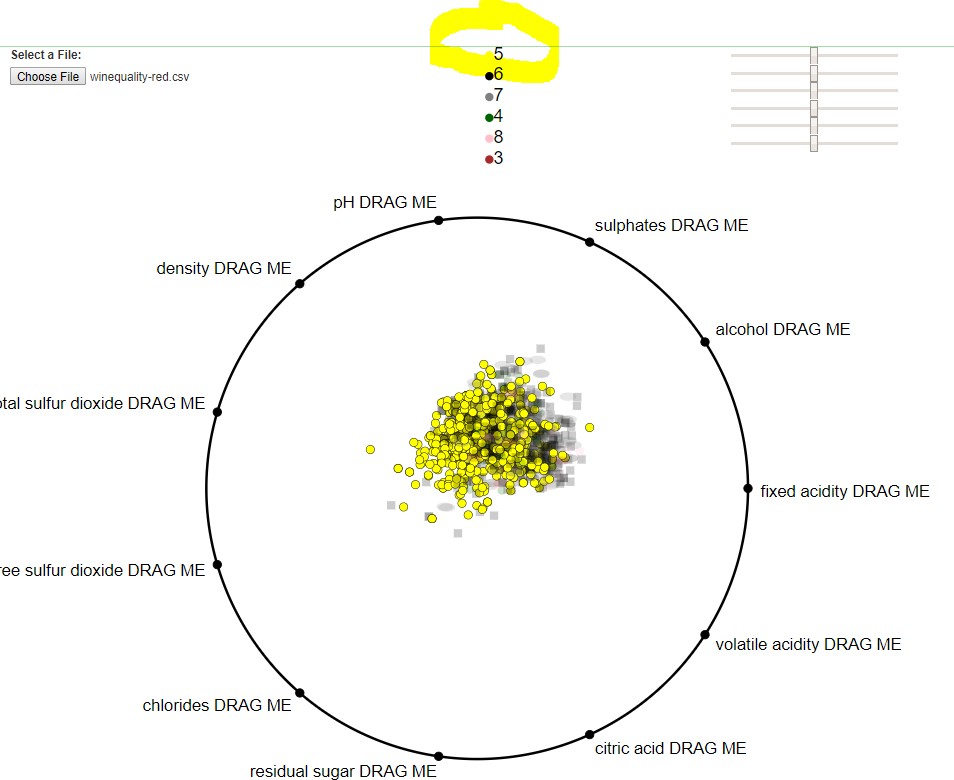
When you hover one data node, this node will enlarged automatically , and the tooltips will show up to show the dimensions about that node.





* 1. Select which anchors to show.

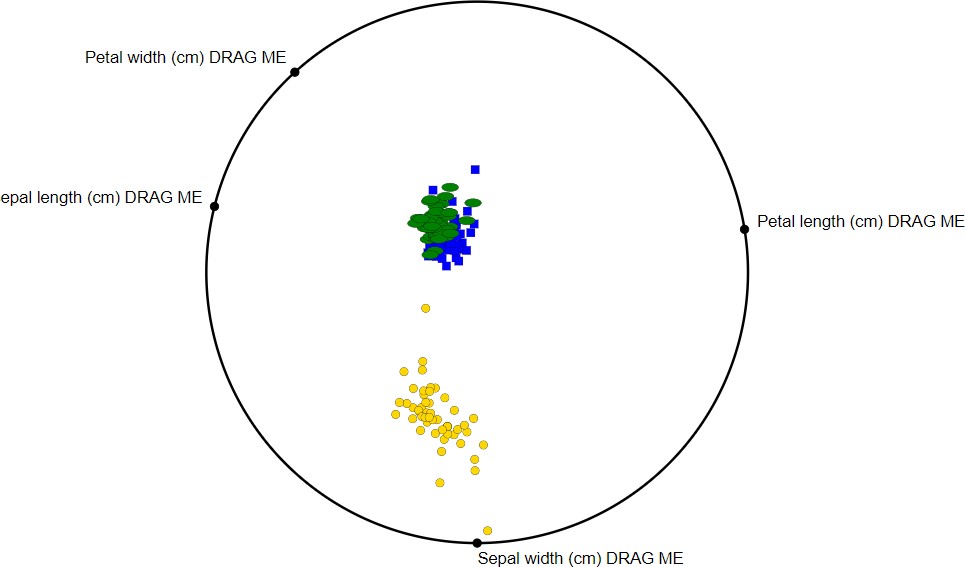
When you hover over the different ‘quality’ values text, it will increase the opacity value, and diminish others. For example on our screenshot, we only show quality 5 nodes.



* 1. Automatically define anchors position to better represent data clusters.

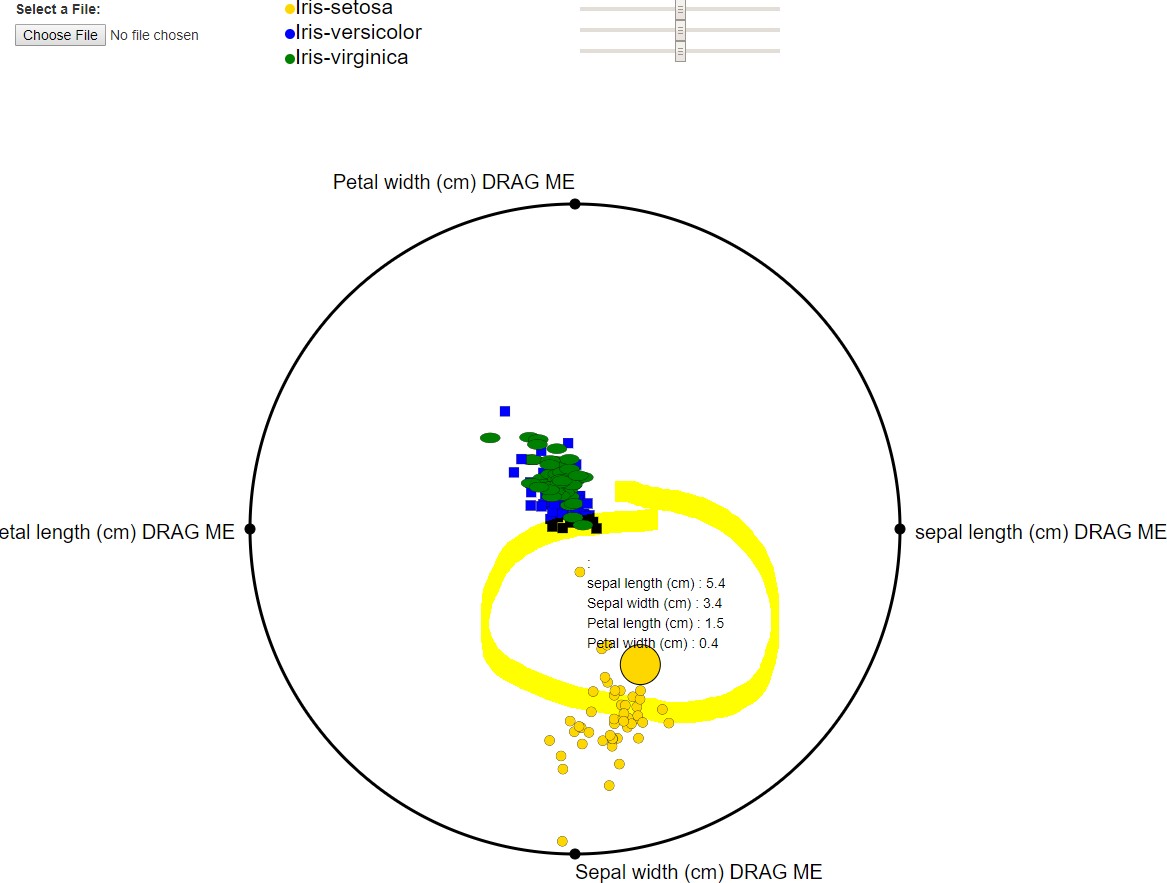
As we can see, the code algorithm can optimize the best representation of clusters, which maximize the intra-variance between different nodes.



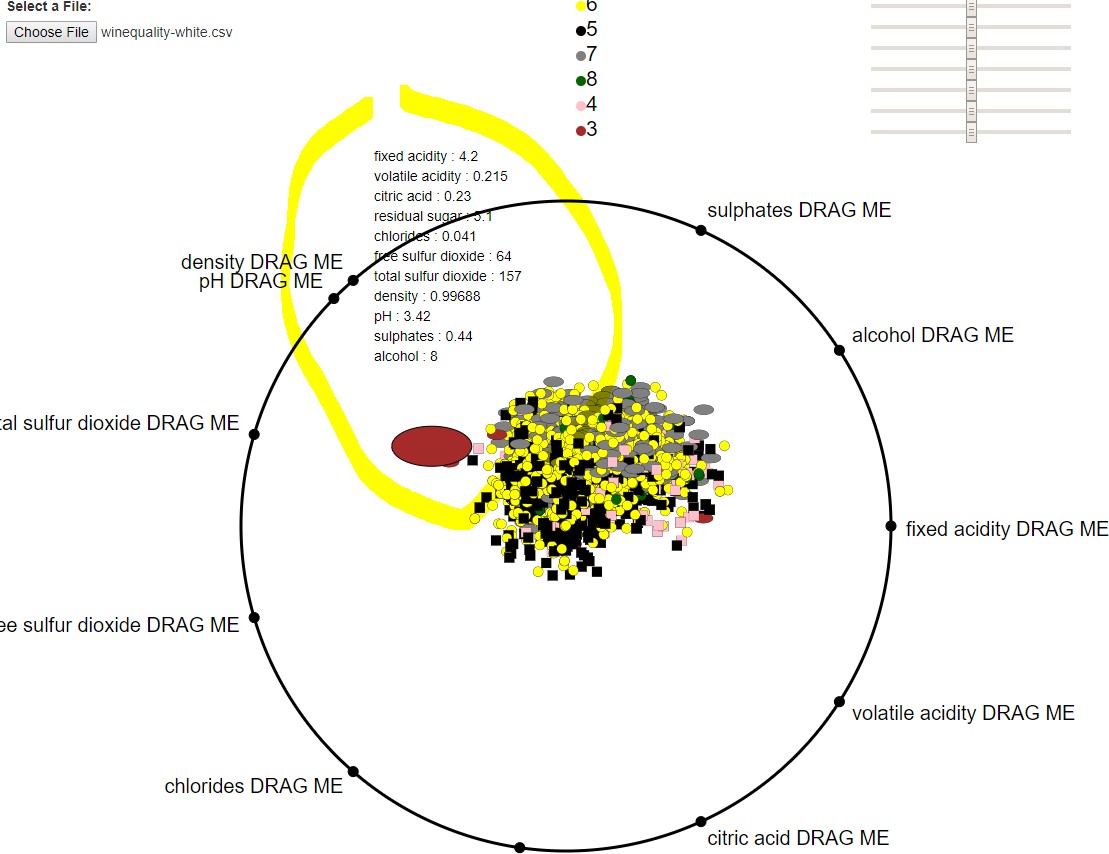


* 1. Some kind of data highlight/selection.

As we can see, when you hover center node, that node will be highlighted to be outstanding over others.

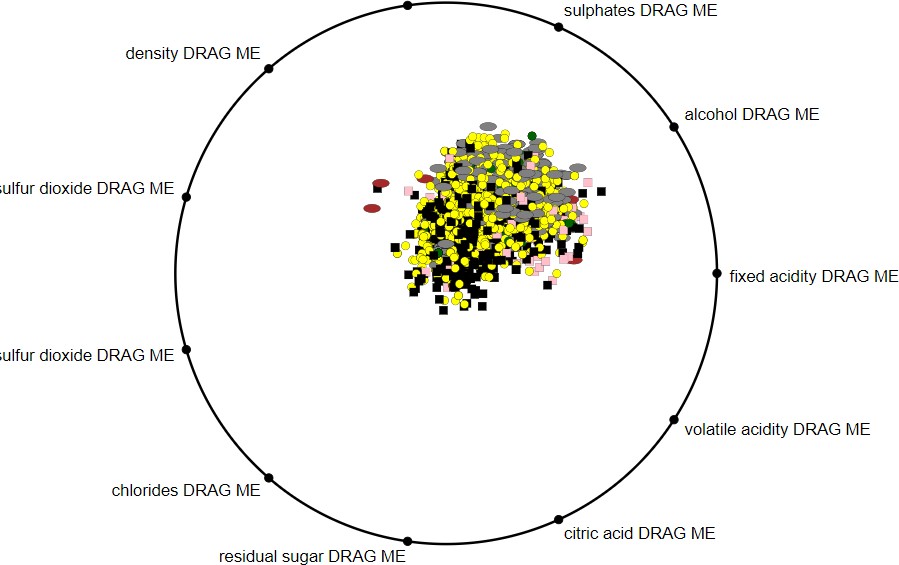






The grade will be based on the complexity and usefulness of the feature. You may need to create some extra interface to control one or more of your unique features. Feel free to add it the way you prefer, as long as it is still contained in the webpage.

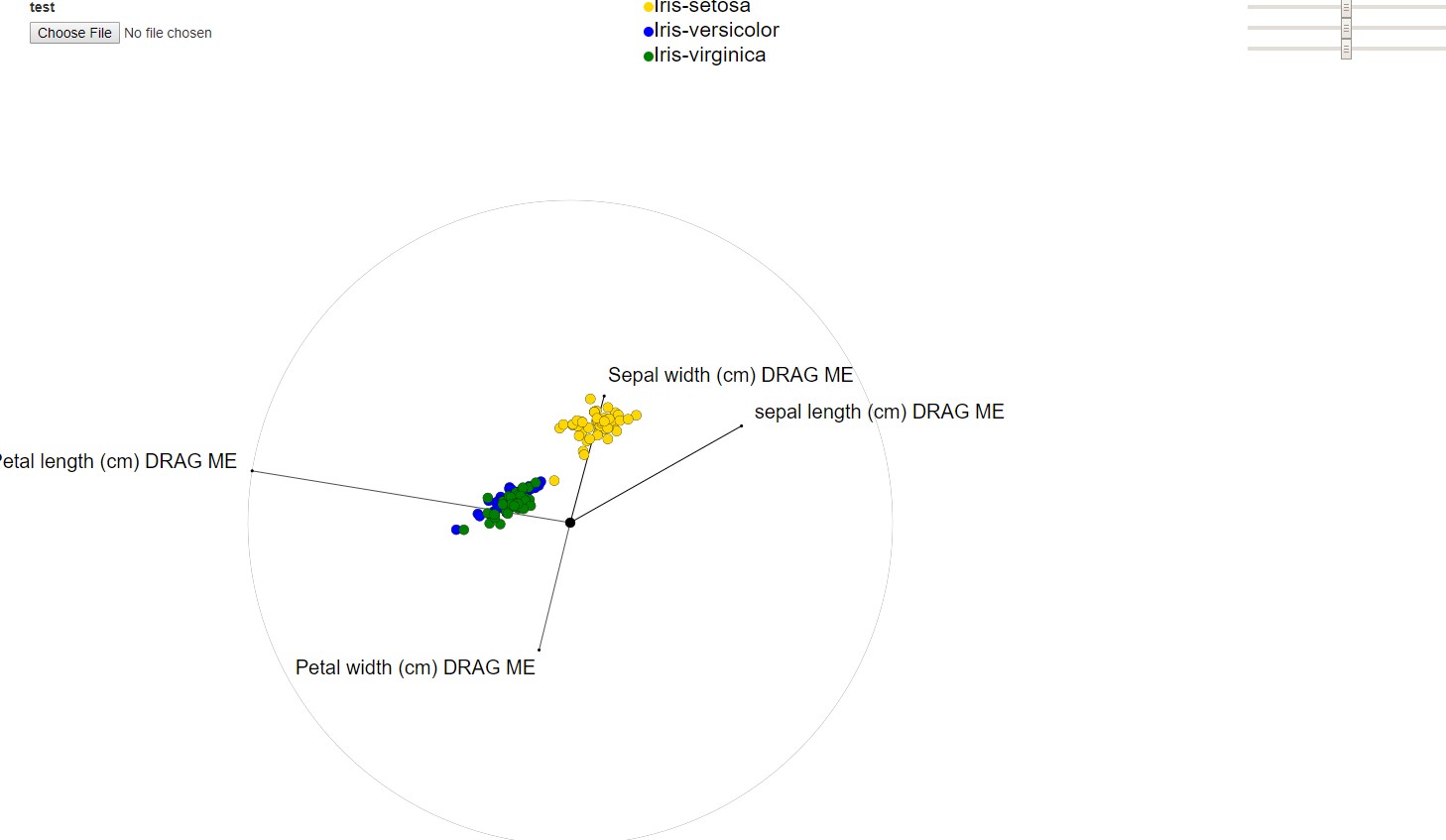
Moreover, my code also can assign different shapes and colors to different target/labels/clusters:



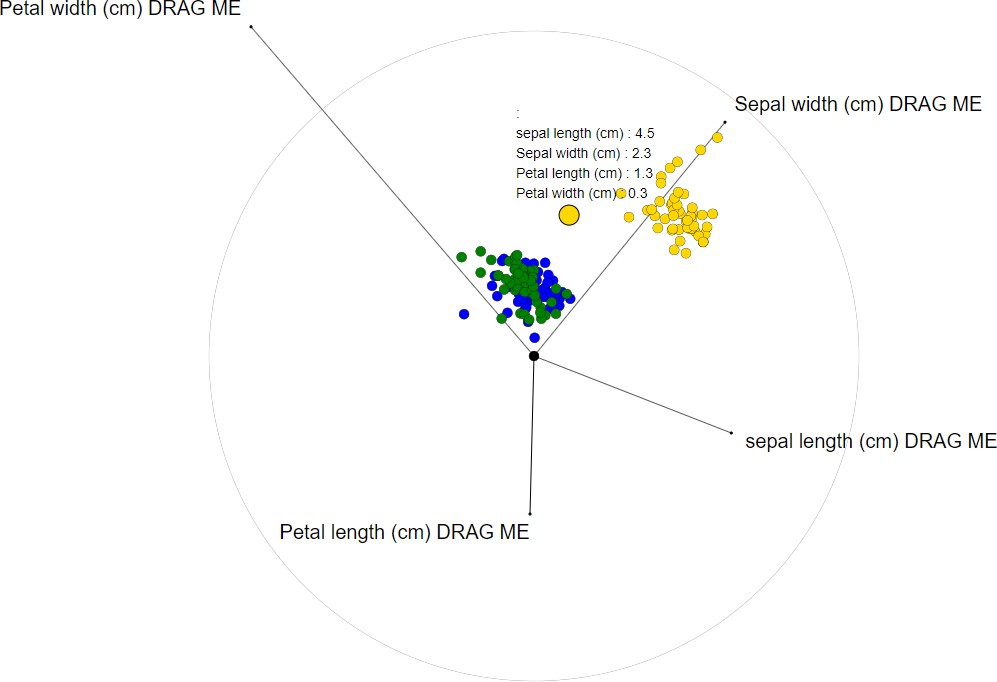


1. [+30 bonus marks] Implement both visualizations (RadViz and Start Coordinates) and include the functionality of selecting a new .csv file and show it in the visualization.
   1. Also, test using the iris dataset (iris.csv). Version uploaded to brightspace contains column names.

Star Coordinates:

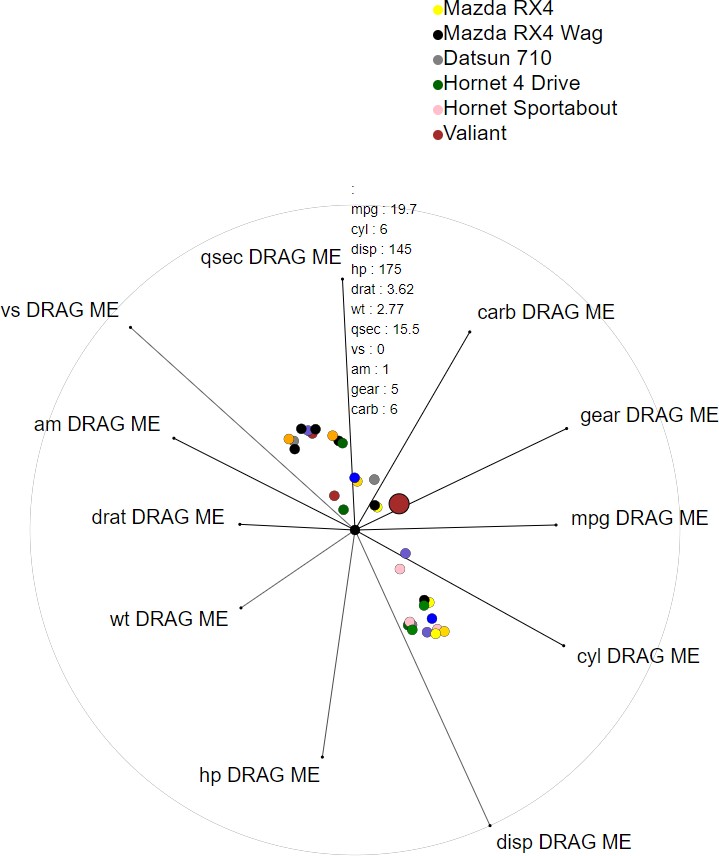


We can also drag the features nodes to longer length to assign different weight than original setup, like this:



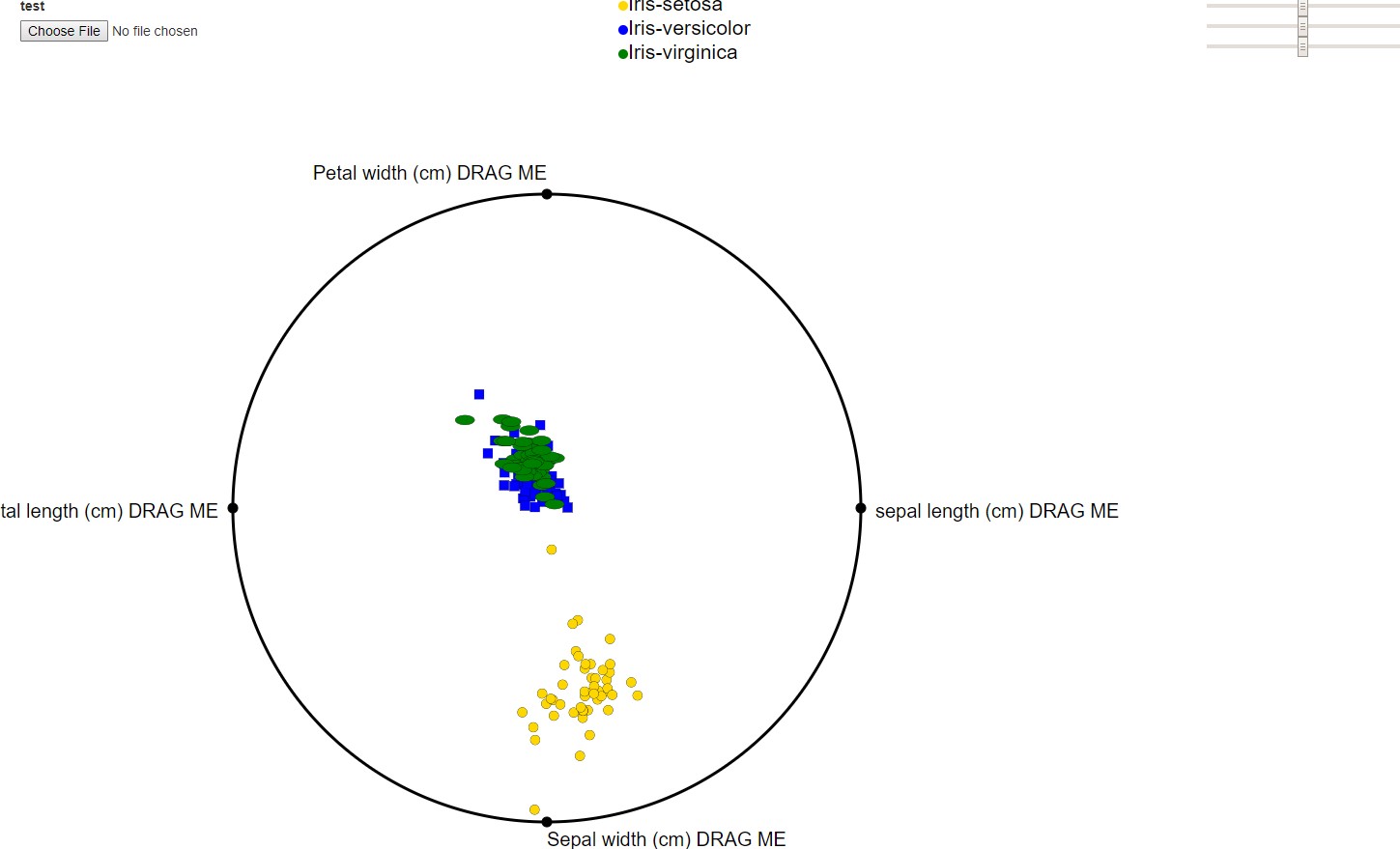


For the new data set: mtcars.csv, similarly:

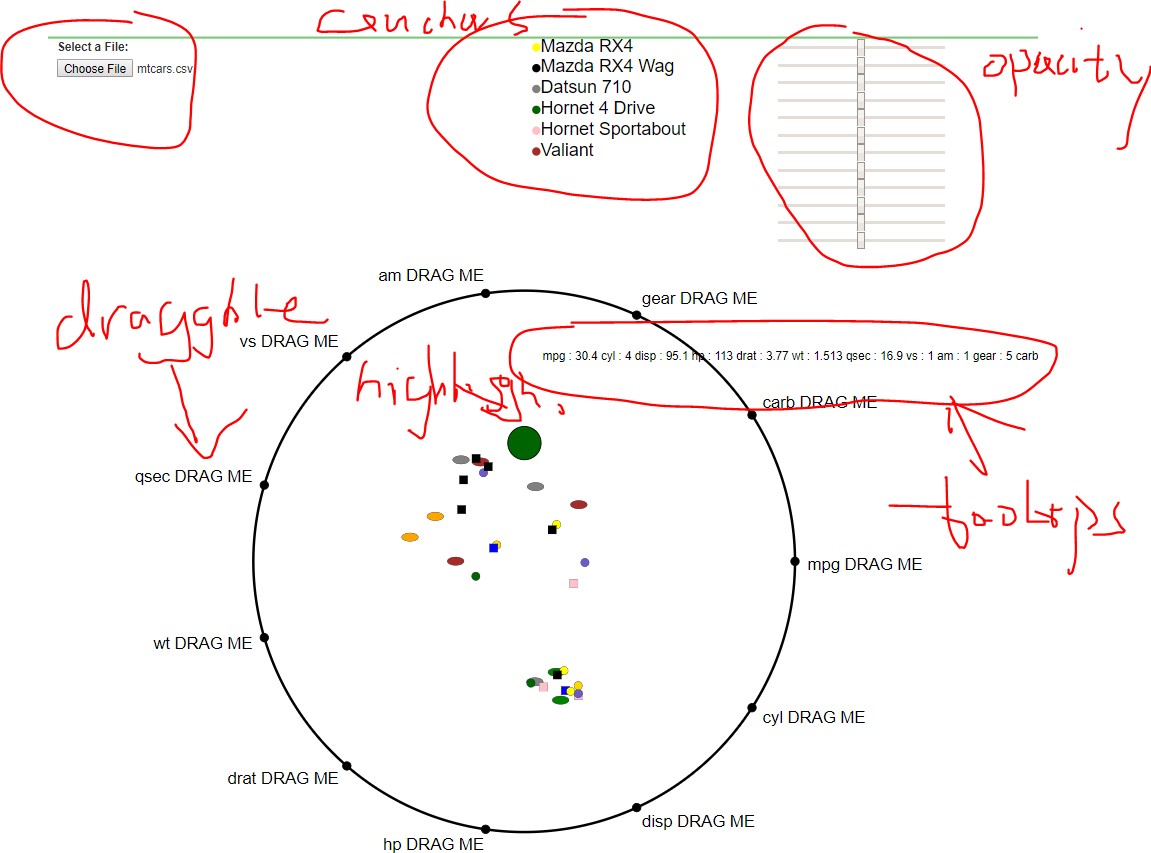


Radviz:

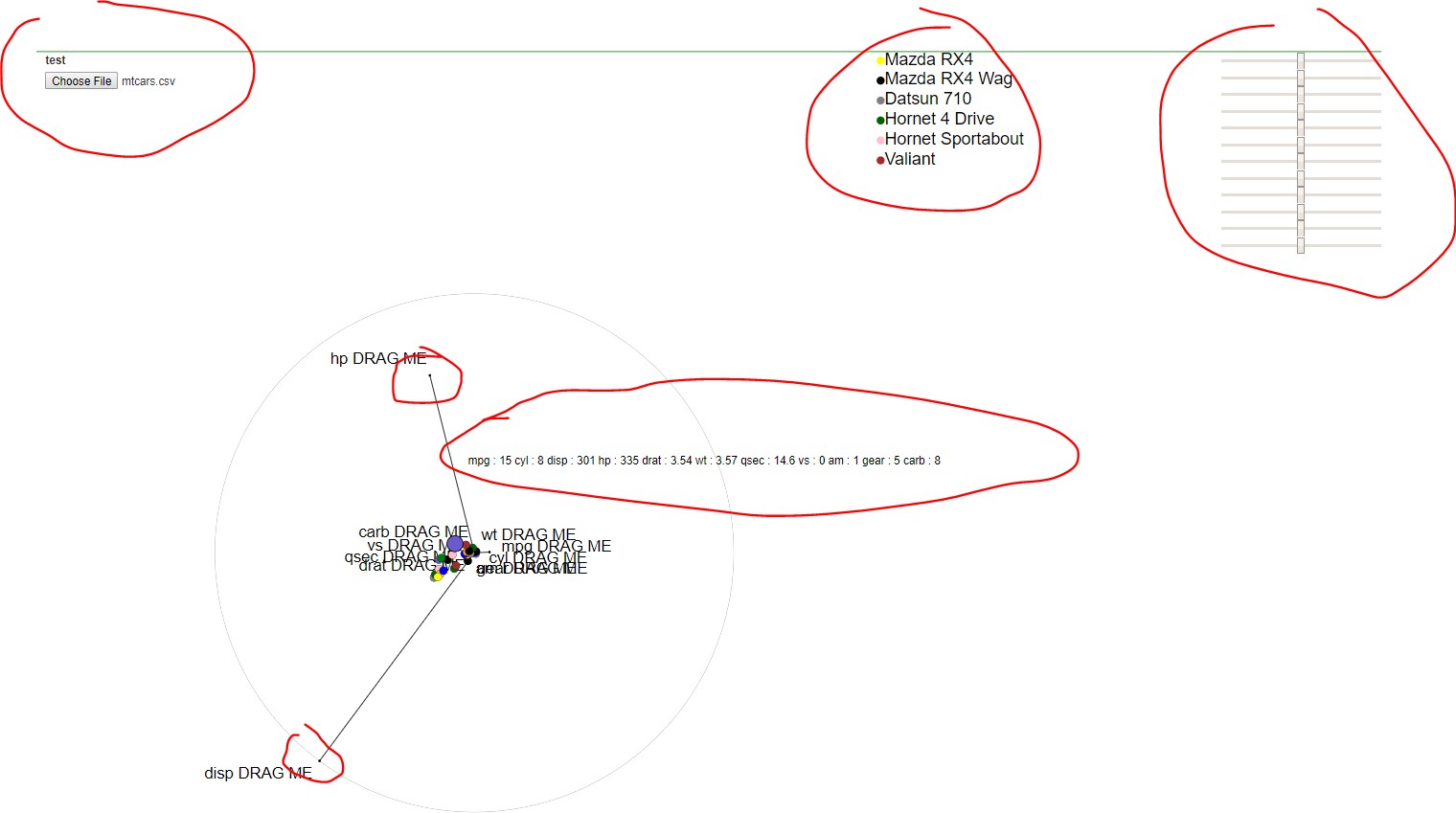




* 1. Grading will follow the same requirements previously stated on items 1 to 4 with marks: 14, 5, 2, 2 and 7 for the dataset changing functionality (selecting a new csv). Here I am choosing famous dataset “mtcar.csv”







* 1. You can do the same assumptions for the datasets as stated previously.

Answers as above screenshots

* 1. Bonus marks will be added to this or previous assignment with the lowest grade. Maximum grade is still 100.
  2. You can show both visualizations at the same time or provide a way to switch between them.
  3. For changing the CSV, you may use the standard browser file dialog to access the file.