Date: 11/02/2022

Walchand College of Engineering, Sangli

Department of Computer Science and Engineering

Software Engineering Tools Lab

Assignment No 1

(Module 1- Introduction to OSS)

Name : Akanksha Bajirao Chaudhari.

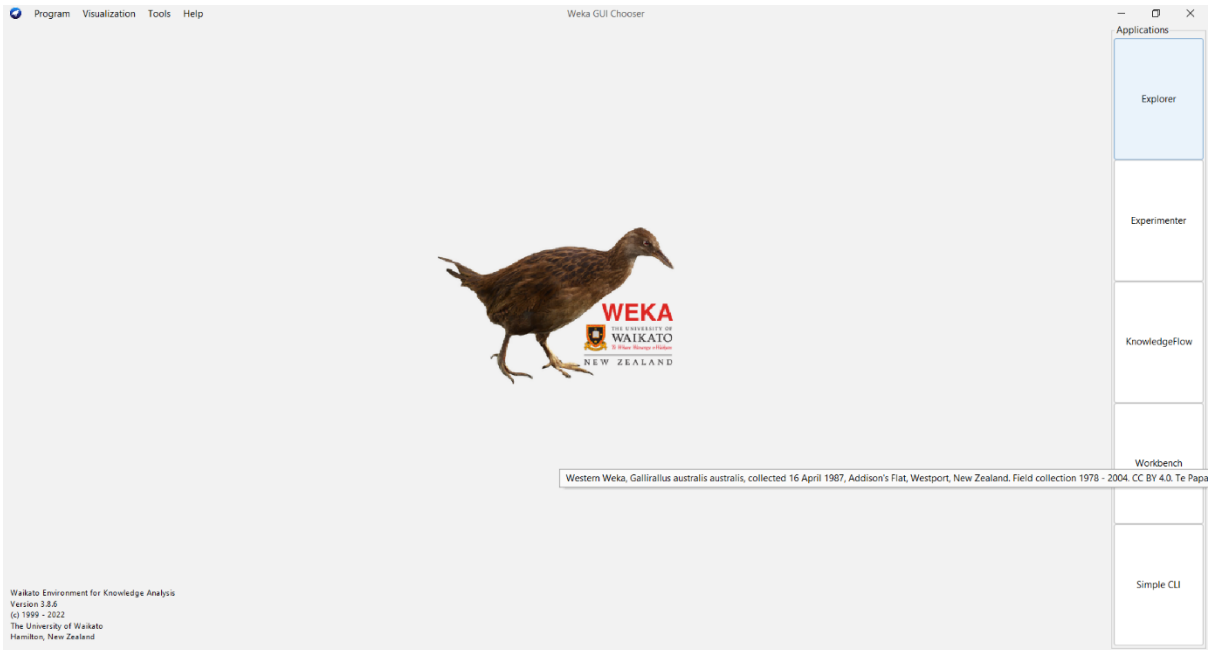
PRN No : 2019BTECS00093

Batch : T7

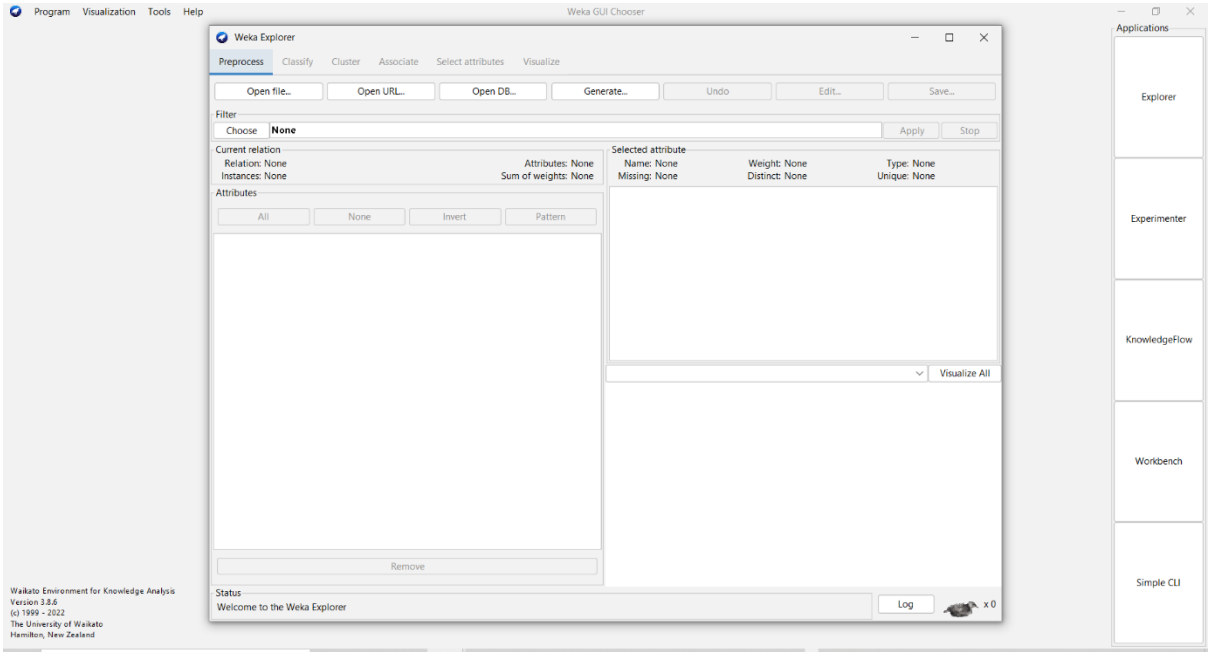
---

1. **Weka** is a GUI workbench that empowers data wranglers to assemble machine learning pipelines, train models, and run predictions without having to write code. Using Weka tool perform below tasks such as data preprocessing, data classification (use any appropriate ML algorithm) and data visualization efficiently on given dataset. Use the Iris dataset given-https://drive.google.com/file/d/1A3Fxsfzm6BSfhFZGDrjI47RTe45bSgYP/view Note-provide screen shots for every task.
   Create a report which will illustrate the details of tasks performed (for e.g to perform preprocessing of data provide details of navigation and selection of appropriate parameters)

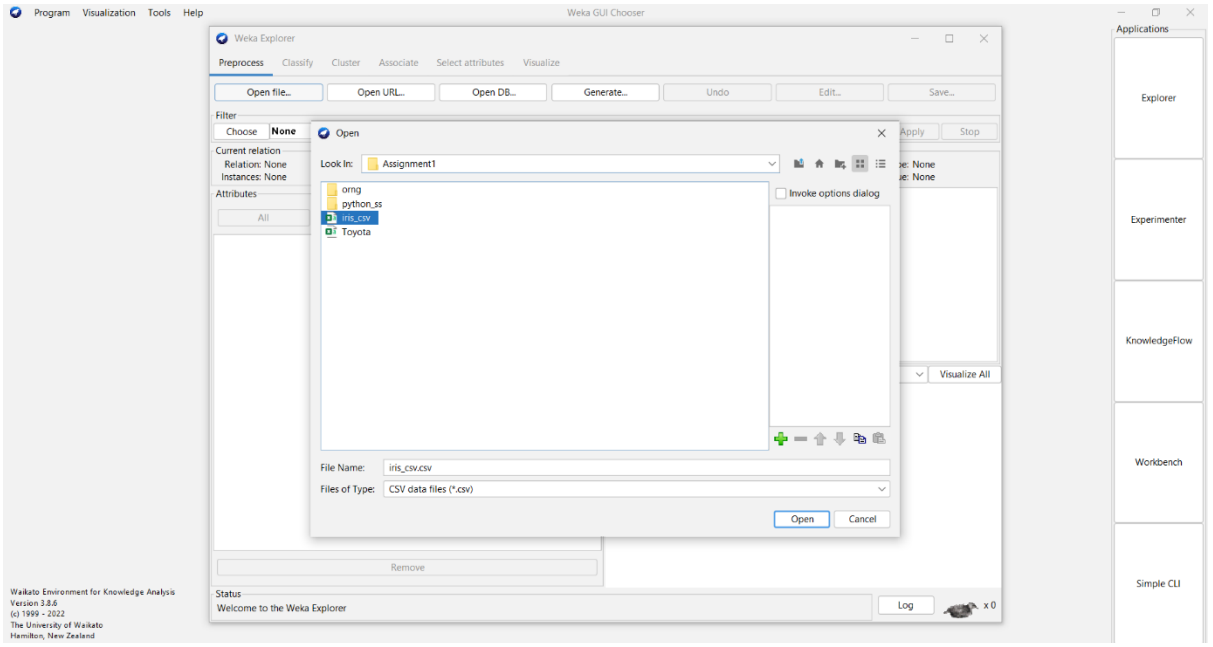
Downloaded and installed Weka Tool:



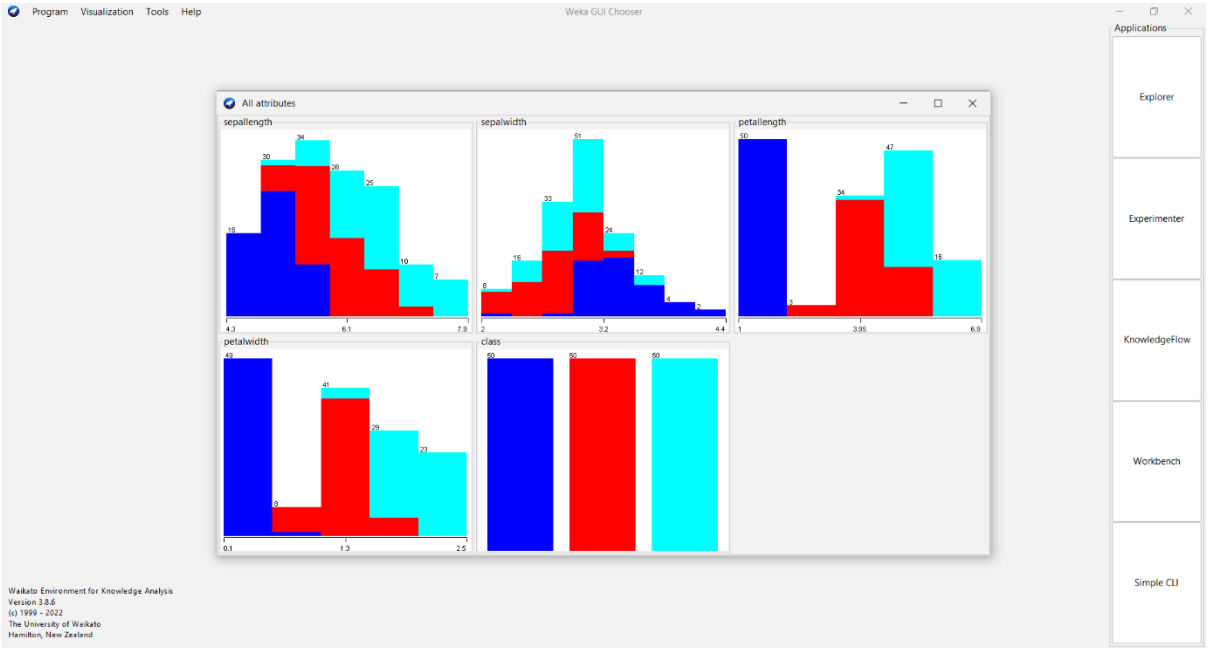
Select Explorer:



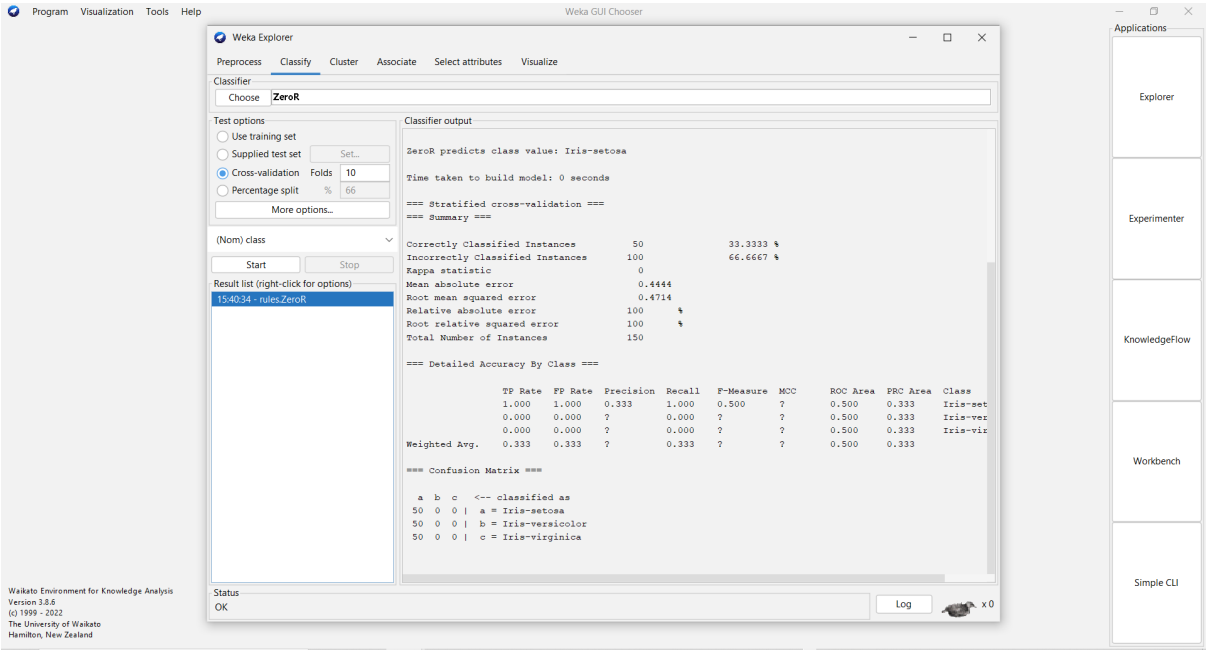
Open iris.csv file.



Select Visualize all:



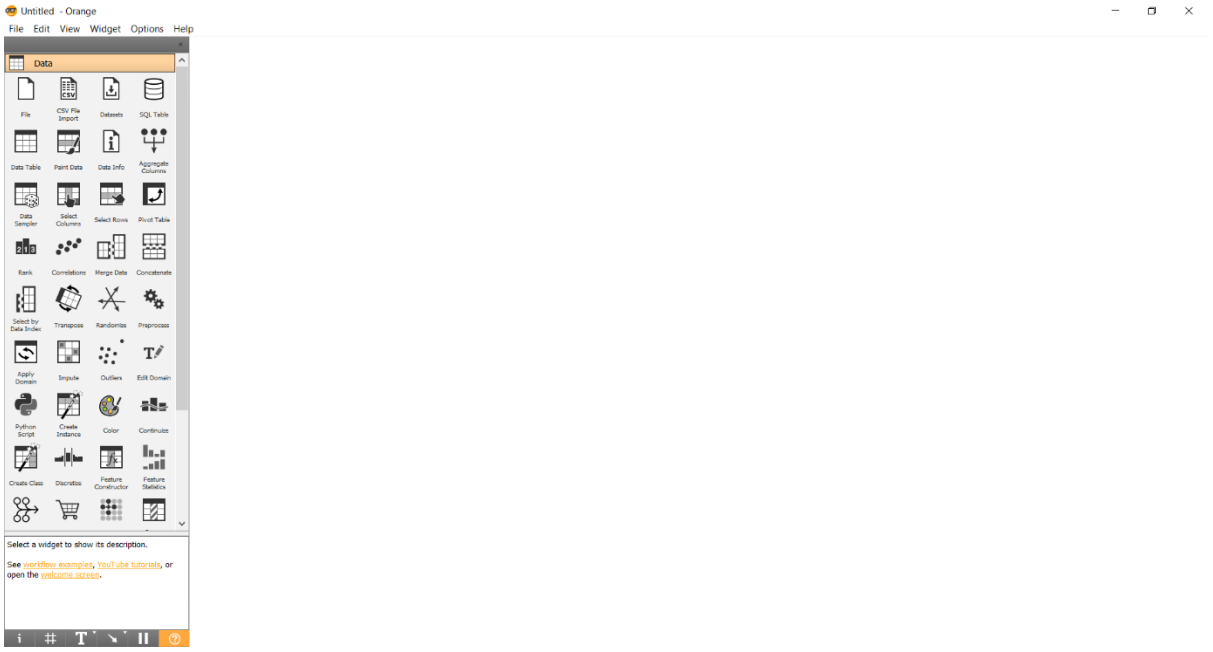
After selecting Classify:



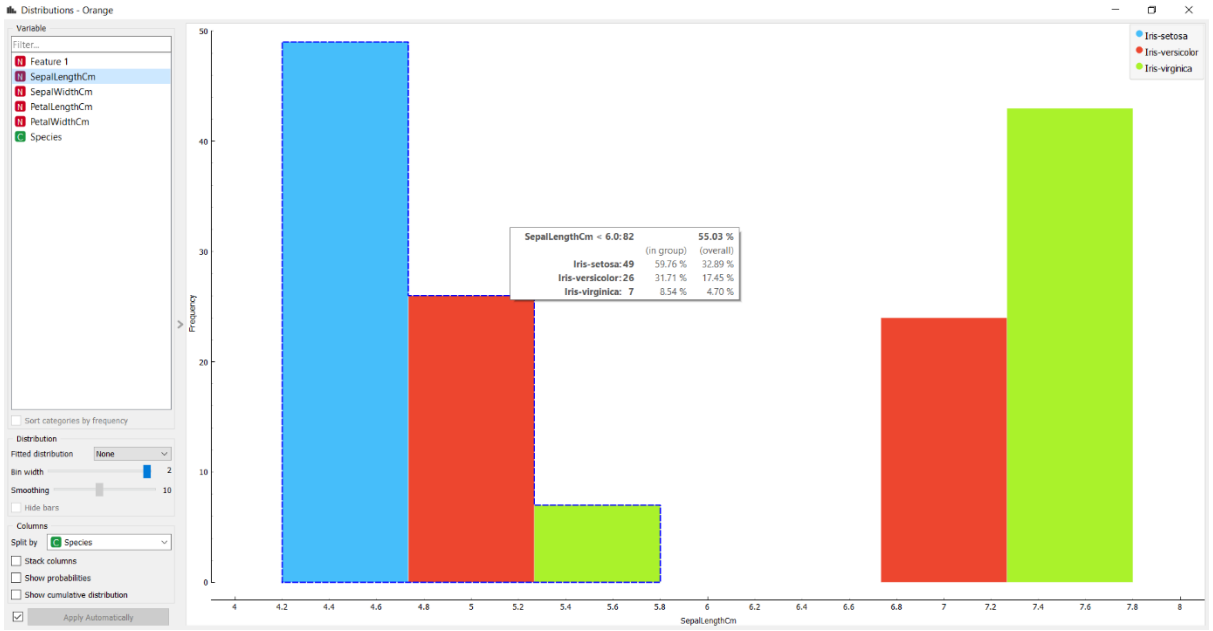
2. **Orange** is an easy to use data visualization tool with a large toolkit. In spite of being a GUI-based beginner-friendly tool, you mustn't mistake it for a light-weight one. It can do statistical distributions and box plots as well as decision trees, hierarchical clustering and linear projections.
   a. Install orange
   b. Show data distribution
   c. Show linear projection
   d. Show FreeViz

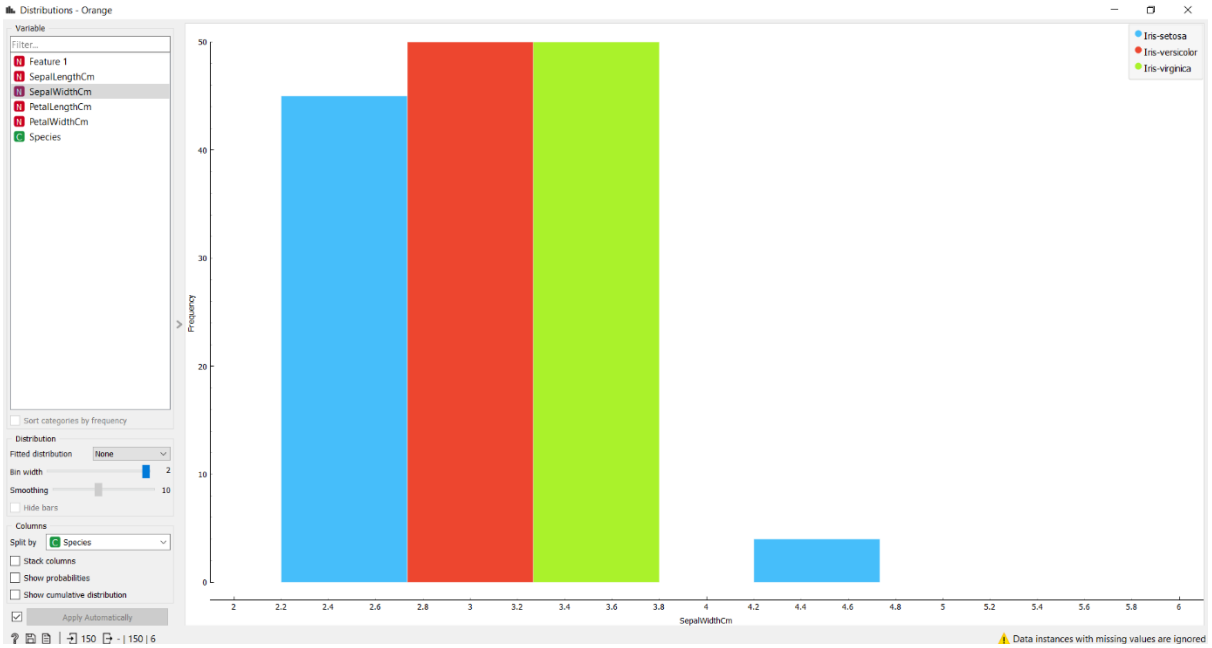
Use dataset
https://drive.google.com/file/d/1m6sKI1Dap0XK6Bw1edUd5PohwpPwXnd9/view
Create a report for this task and upload screenshots for the same.

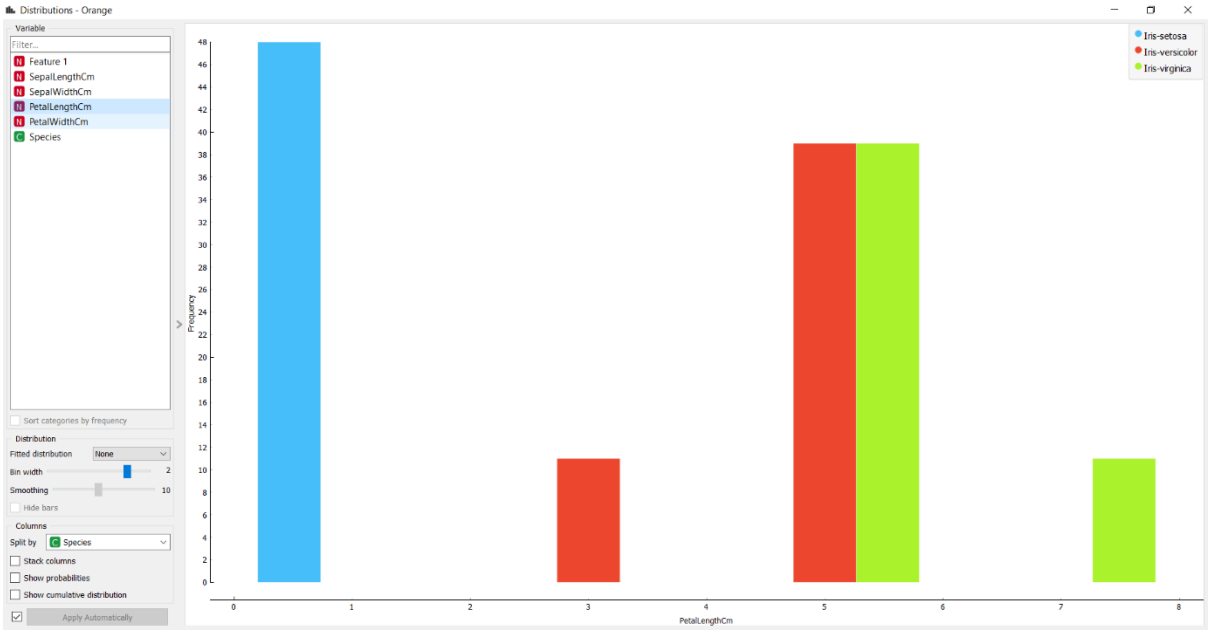
- a. Installation of orange tool.

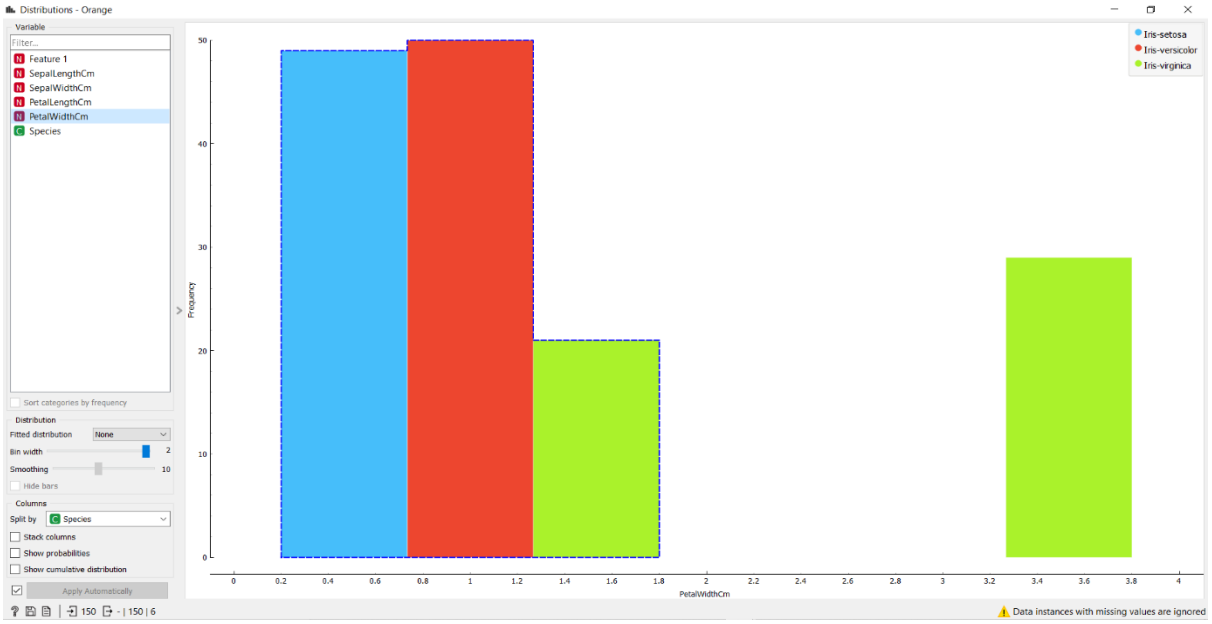


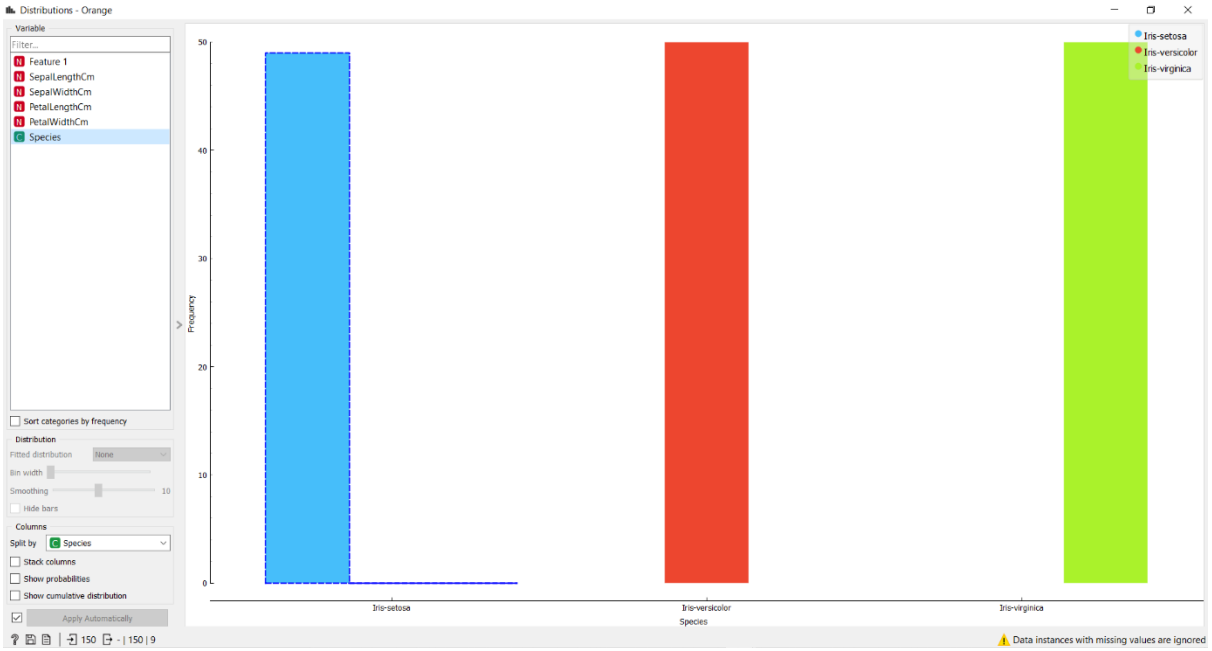
b. Data distribution of given dataset.



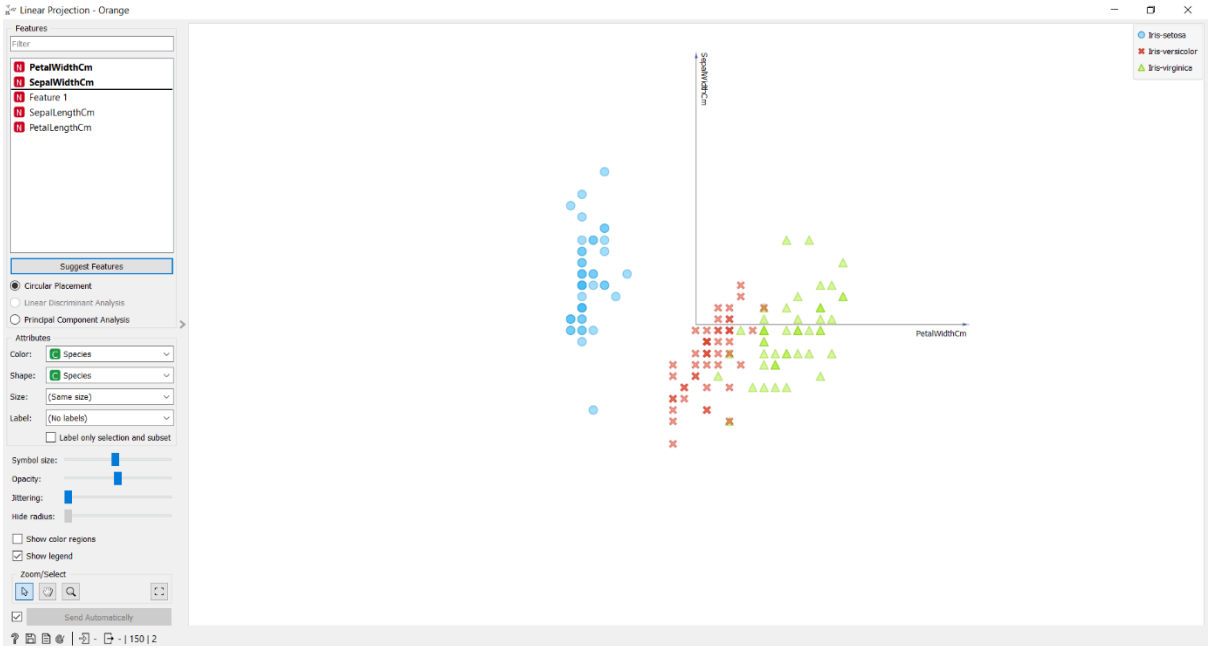

Distributions - Orange
Variable
Filter...
Feature 1
SepalLengthCm
SepalWidthCm
PetalLengthCm
PetalWidthCm
Species
Sort categories by frequency
Distribution
Fitted distribution
None
Bin width
2
Smoothing
10
Hide bars
Columns
Split by
Species
Stack columns
Show probabilities
Show cumulative distribution
Apply Automatically
150 | - | 150 | 6
Iris-setosa
Iris-versicolor
Iris-virginica
Frequency
SepalWidthCm
Data instances with missing values are ignored

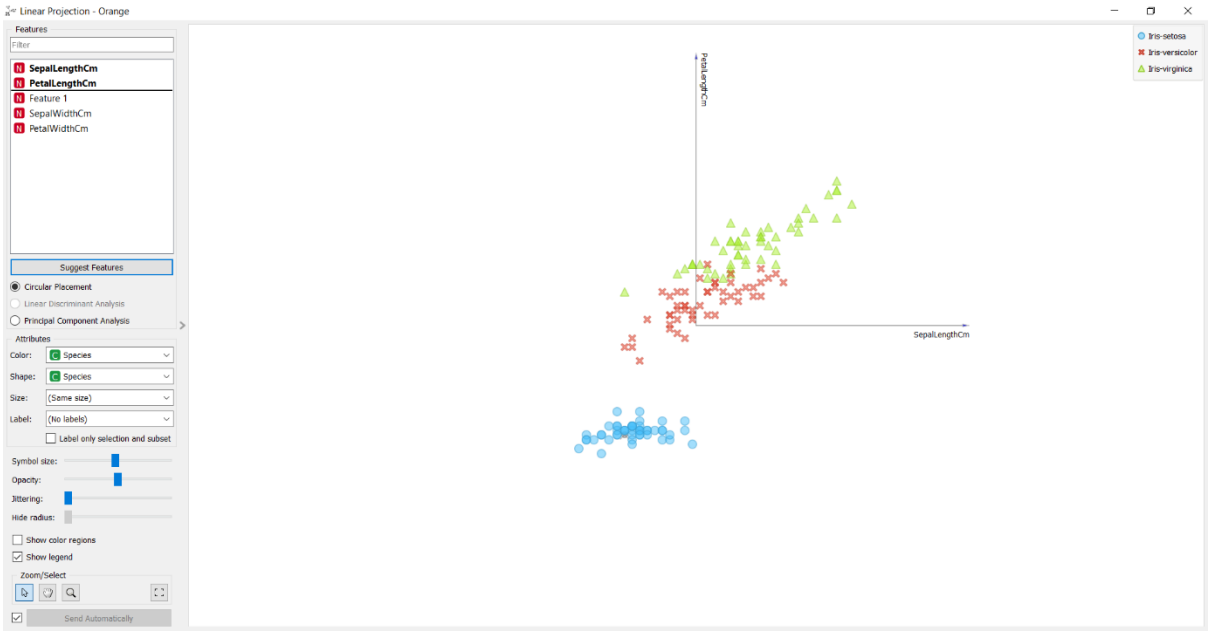

Distributions - Orange
Variable
Filter...
Feature 1
SepalLengthCm
SepalWidthCm
PetalLengthCm
PetalWidthCm
Species
Sort categories by frequency
Distribution
Fitted distribution
None
Bin width
2
Smoothing
10
Hide bars
Columns
Split by
Species
Stack columns
Show probabilities
Show cumulative distribution
Apply Automatically
Iris-setosa
Iris-versicolor
Iris-virginica
Frequency
PetalLengthCm

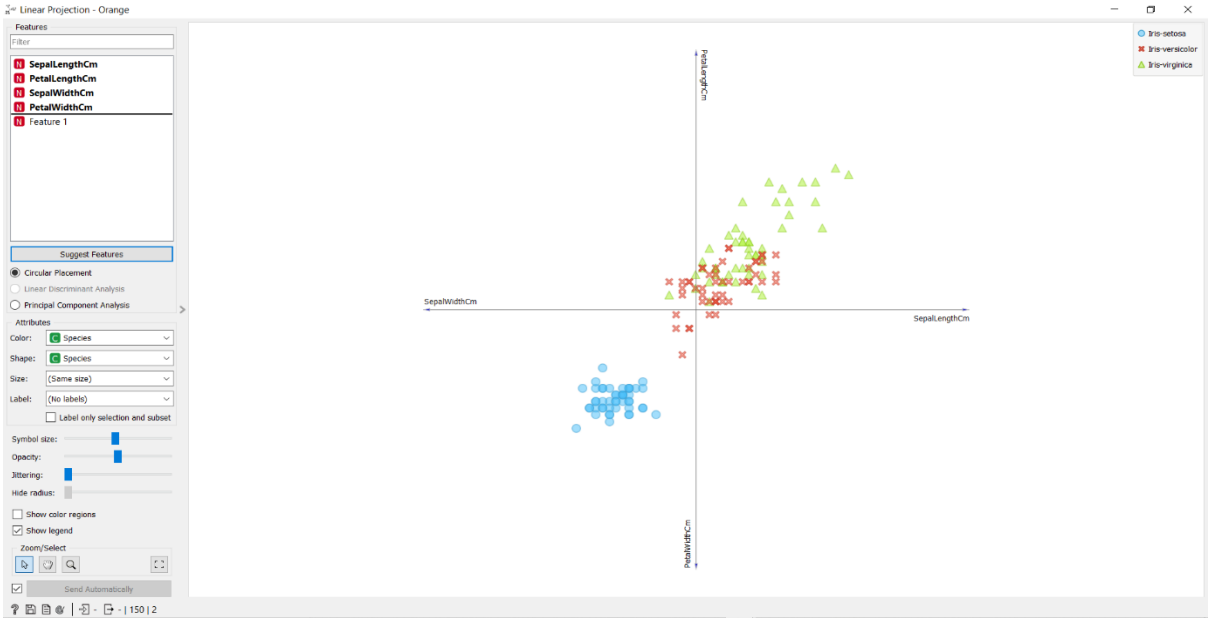

Distributions - Orange
Variable
Filter...
Feature 1
SepalLengthCm
SepalWidthCm
PetalLengthCm
PetalWidthCm
Species
Sort categories by frequency
Distribution
Fitted distribution
None
Bin width
2
Smoothing
10
Hide bars
Columns
Split by
Species
Stack columns
Show probabilities
Show cumulative distribution
Apply Automatically
150 | - | 150 | 6
Iris-setosa
Iris-versicolor
Iris-virginica
Frequency
PetalWidthCm
Data instances with missing values are ignored


Distributions - Orange
Variable
Filter...
Feature 1
SepalLengthCm
SepalWidthCm
PetalLengthCm
PetalWidthCm
Species
Sort categories by frequency
Distribution
Fitted distribution
None
Bin width
Smoothing
10
Hide bars
Columns
Split by
Species
Stack columns
Show probabilities
Show cumulative distribution
Apply Automatically
150 | - | 150 | 9
Iris-setosa
Iris-versicolor
Iris-virginica
Frequency
Species
Data instances with missing values are ignored

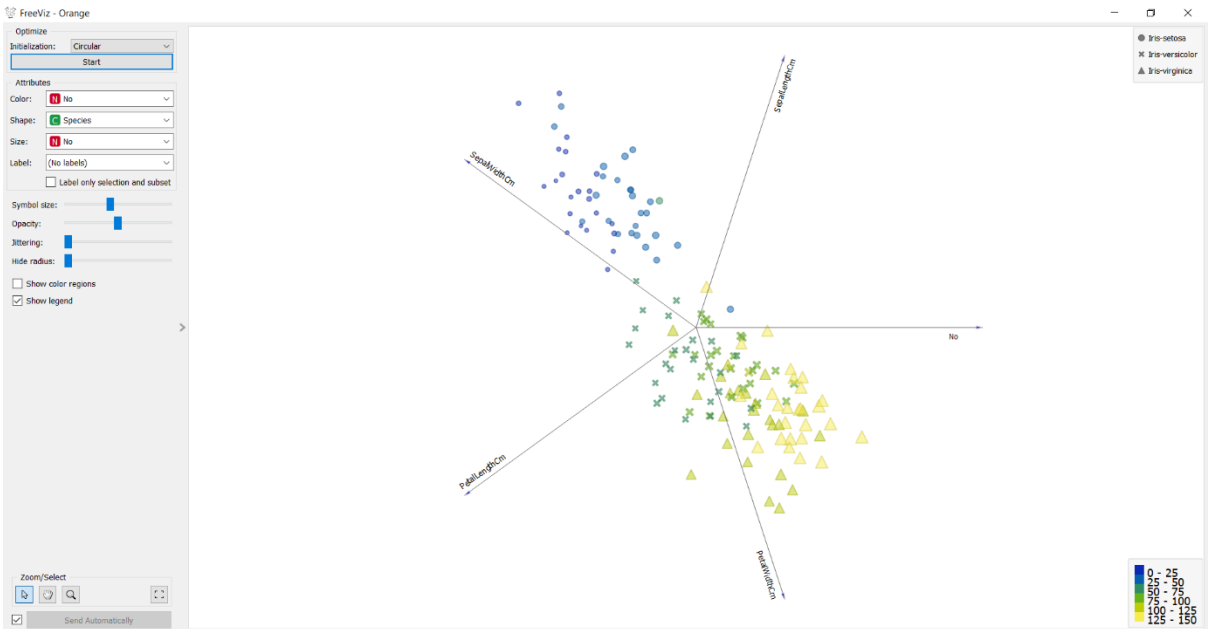
c. Linear projection of given dataset

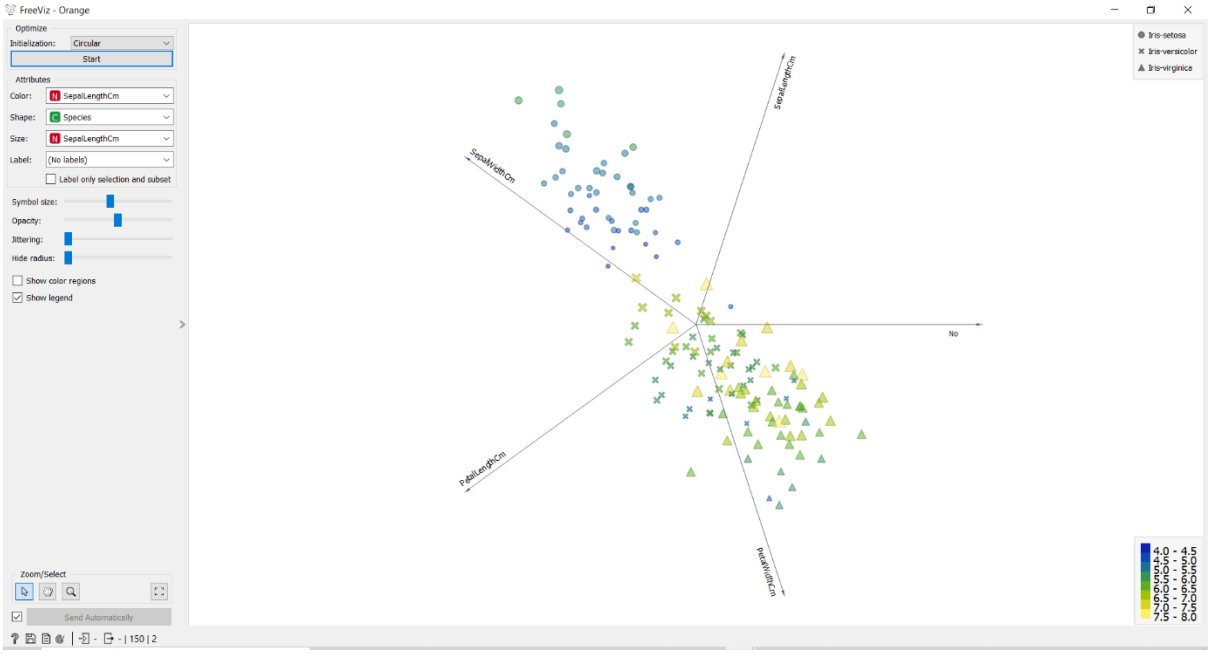


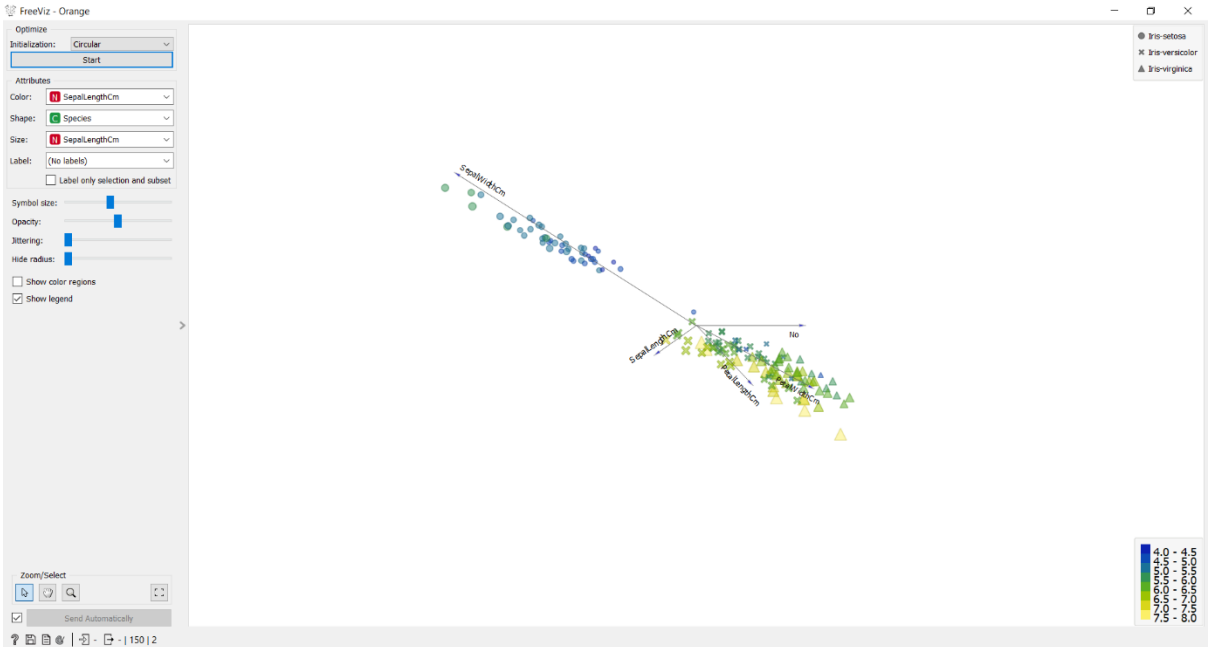


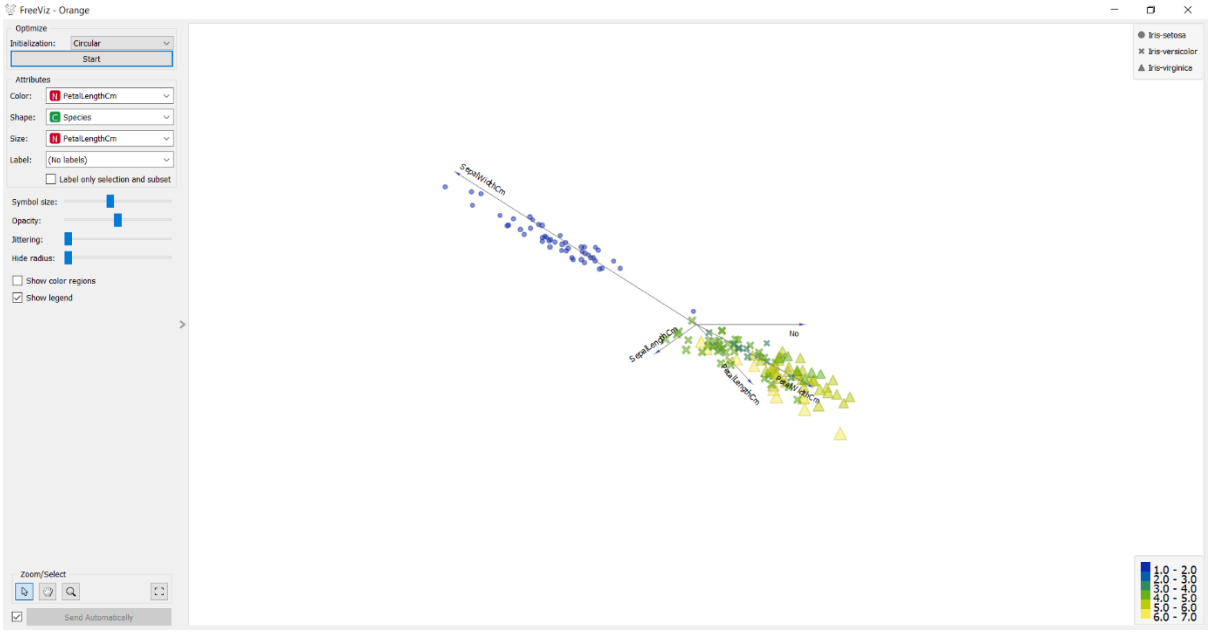


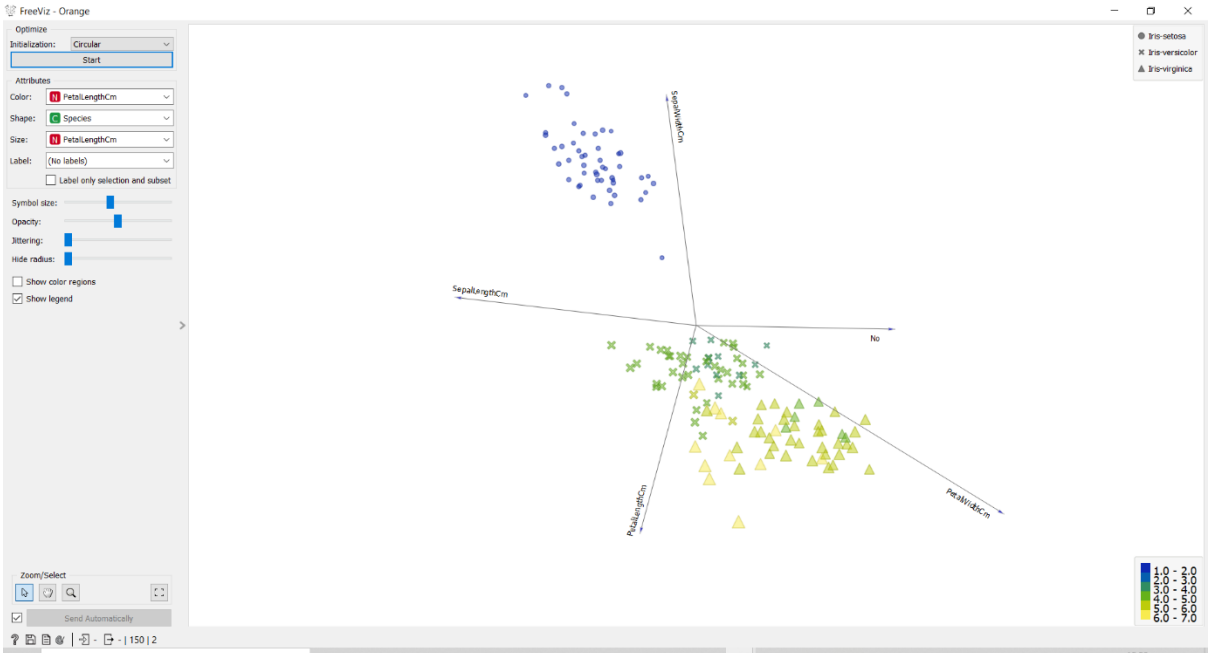
c. Free Viz of given dataset




FreeViz - Orange
Optimize
Initialization: Circular
Start
Attributes
Color: SepalLengthCm
Shape: Species
Size: SepalLengthCm
Label: (No labels)
Label only selection and subset
Symbol size:
Opacity:
Jittering:
Hide radius:
Show color regions
Show legend
Zoom/Select
Send Automatically
150 | 2
Iris-setosa
Iris-versicolor
Iris-virginica
SepalLengthCm
SepalWidthCm
PetalLengthCm
PetalWidthCm
No
4.0 - 4.5
4.5 - 5.0
5.0 - 5.5
5.5 - 6.0
6.0 - 6.5
6.5 - 7.0
7.0 - 7.5
7.5 - 8.0


FreeViz - Orange
Optimize
Initialization: Circular
Start
Attributes
Color: SepalLengthCm
Shape: Species
Size: SepalLengthCm
Label: (No labels)
Label only selection and subset
Symbol size:
Opacity:
Jittering:
Hide radius:
Show color regions
Show legend
Zoom/Select
Send Automatically
150 | 2
Iris-setosa
Iris-versicolor
Iris-virginica
SepalWidthCm
SepalLengthCm
PetalLengthCm
PetalWidthCm
No
4.0 - 4.5
4.5 - 5.0
5.0 - 5.5
5.5 - 6.0
6.0 - 6.5
6.5 - 7.0
7.0 - 7.5
7.5 - 8.0


FreeViz - Orange
Optimize
Initialization: Circular
Start
Attributes
Color: PetalLengthCm
Shape: Species
Size: PetalLengthCm
Label: (No labels)
Label only selection and subset
Symbol size:
Opacity:
Jittering:
Hide radius:
Show color regions
Show legend
Zoom/Select
Send Automatically
Iris-setosa
Iris-versicolor
Iris-virginica
SepalWidthCm
SepalLengthCm
No
PetalLengthCm
PetalWidthCm
1.0 - 2.0
2.0 - 3.0
3.0 - 4.0
4.0 - 5.0
5.0 - 6.0
6.0 - 7.0


FreeViz - Orange
Optimize
Initialization: Circular
Start
Attributes
Color: PetalLengthCm
Shape: Species
Size: PetalLengthCm
Label: (No labels)
Label only selection and subset
Symbol size:
Opacity:
Jittering:
Hide radius:
Show color regions
Show legend
Zoom/Select
Send Automatically
Iris-setosa
Iris-versicolor
Iris-virginica
SepalWidthCm
SepalLengthCm
No
PetalLengthCm
PetalWidthCm
1.0 - 2.0
2.0 - 3.0
3.0 - 4.0
4.0 - 5.0
5.0 - 6.0
6.0 - 7.0

3. Differentiate in between free software, Open source software and proprietary software with respect to its properties.

Ans: Difference between Free Software and Open Source Software

No.	Free Software	Open Source Software
1	Software is an important part of people's lives.	Software is just software. There are no ethics associated directly to it.
2	Software freedom translates to social freedom.	Ethics are to associated to the people not to the software.
3	Freedom is a value that is more important than any economical advantage.	Freedom is not an absolute concept. Freedom should be allowed, not imposed.
4	Examples: Linux kernel, GNU Compiler Collection, C library etc.	Examples: Apache HTTP Server, Android, VLC Media Player etc.

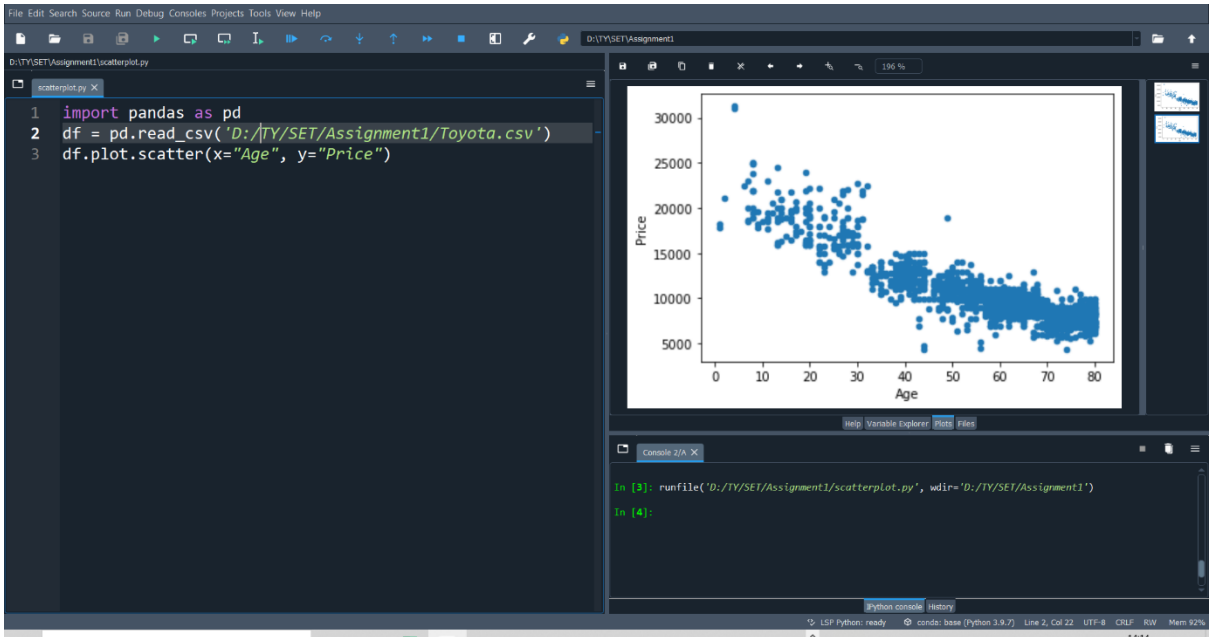
Difference between Open Source Software and Proprietary Software

No.	Open Source Software	Proprietary Software
1	Open source software is a computer software whose source code is available openly in internet and programmers can modify it to add new features and capabilities without any cost.	Proprietary software is a computer software where the source codes are not publicly available, only the company which has created can modify it.
2	Here the software is developed and tested through open collaboration.	Here the software is developed and tested by the individual or organization by which it is owned not by public.
3	In open source software the source code is public.	In proprietary software the source code is protected.
4	Open source software can be installed into any computer.	Proprietary software can be installed into any computer with valid license.
5	Users do not need to have any authenticated license to use this software.	Users need to have a valid and authenticated license to use this software.
6	Open source software is managed by an open source community of developers.	Proprietary software is managed by an closed team of individuals or groups that developed it.
7	It is more flexible and provides more freedom which encourages innovation.	It is not much flexible so here is very limited innovation scope with the restrictions.
8	Users can get open software for free of charge.	Users must have to pay to get the proprietary software.
9	In open source software faster fixes of bugs and better security is availed due to the community.	In proprietary software the vendors is completely responsible for fixing of malfunctions.

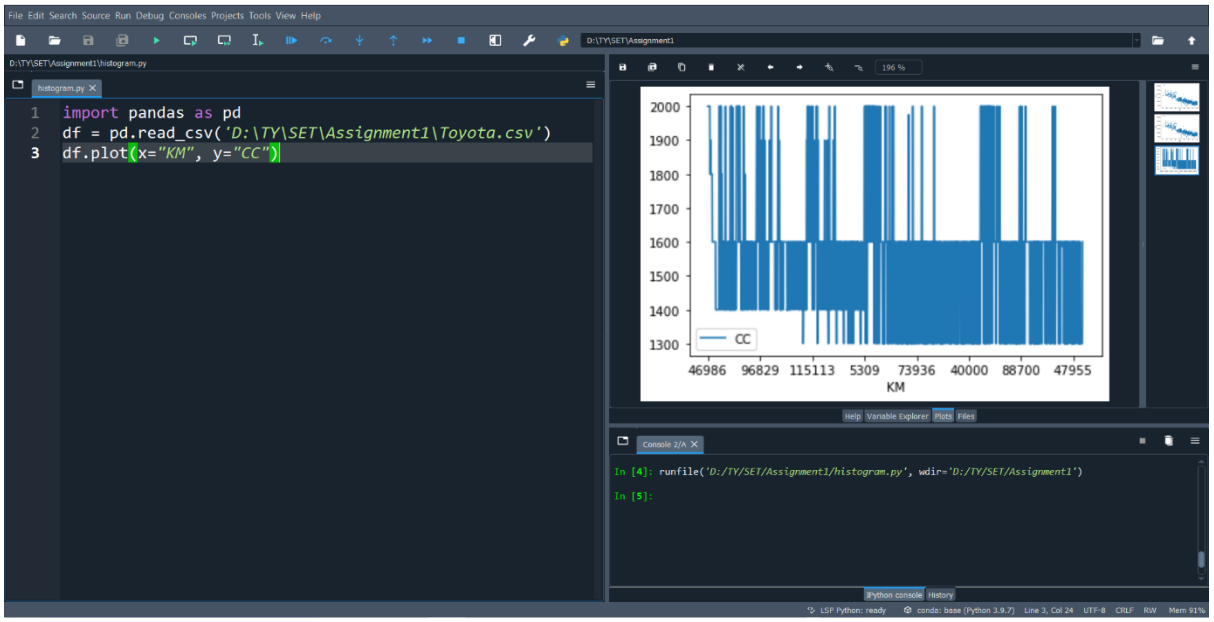
10	Examples: Android, Linux, Firefox, VLC media player etc.	Examples: Windows, MacOS, Internet Explorer, Microsoft Office etc.

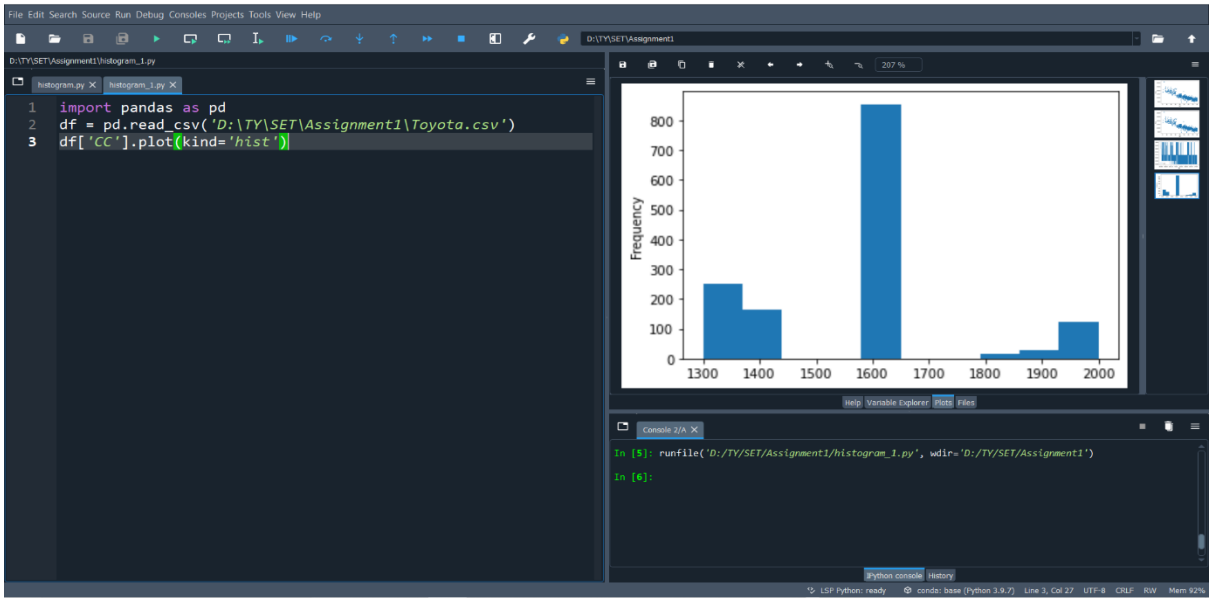
4. Using Anaconda Python create Histogram, Scatter plot and Bar plot for the dataset given below.
   Dataset-
   https://drive.google.com/file/d/1i11BZFe8Xj9kNq7eeE9KOa_Iz1KhEdXJ/view

   a. Scatter plot- Scatter plot of Price Vs Age

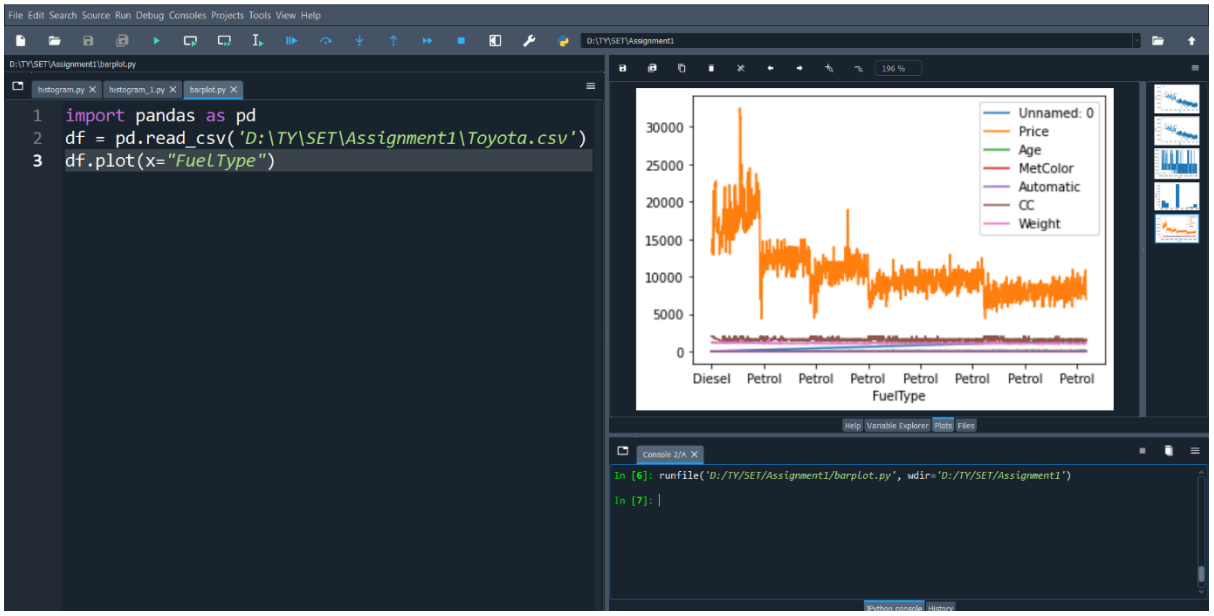


   b. Histogram- for Kilometer and CC





c. Bar plot- Bar plot for different fuel types



5. Enlist some examples along with its purpose and properties (at least 10) of FOSS and proprietary software with respect to database.

FOSS:

Free and open-source software (FOSS) is software that can be classified as both free software and open-source software. ... This is in contrast to proprietary software, where the software is under restrictive copyright licensing and the source code is usually hidden from the users.
Some examples of FOSS software include: the Linux kernel, GNOME Desktop and Free Berkeley Software Distribution (FreeBSD).
Properties of FOSS:

i. **Expansive licensing:** Proprietary software licenses are usually quite restrictive in terms of use—limits on number of users, type or number of machines the software may be installed on, and, of course, there is usually a fee associated with obtaining a license. By contrast, open source has very expansive license conditions that encourage widespread use.

ii. **Development transparency:** Open source development is carried out in the open. In most cases, product decisions are discussed extensively on mailing lists or in forums. All code may be examined. Reported bugs are listed and available for inspection. The development process itself is carried out in public, with all code check-ins also available for inspection

iii. **Ability to inspect source code:** It is often extremely helpful to review the source code of a product to enable better integration with another product or merely to better understand how the product operates so as to ease use in production. Because open source licenses mandate source code availability, it is easy to study the product's code and learn from it.

iv. **Ability to modify source code:** Not only is open source code available for inspection, licenses also allow users to modify the source code. Anyone can add new functionality that better meets user needs. Furthermore, the code can be "contributed" back to the mainstream code base, which means that code modifications are automatically carried forward in subsequent releases, thereby reducing downstream maintenance efforts.

v. **Community:** One of the most important characteristics of open source, and the foundation for successful open source projects, is community. Community is the combined pool of product developers and users; in essence, everyone concerned with a product. Free and honest communication is typical of community, with many peer user interactions typical of a vibrant community. Users can easily share their thoughts about the product with developers, leading to improved functionality and ease of use. The community also enables "corner case" use cases (i.e., unusual product applications that only a subset of users confront) to be exercised, with feedback about product capability in corner-case conditions directly fed back to developers.

vi. **Redistribution rights:** Open source licenses allow users (recipients of open source products) to distribute open source products to third-parties as part of the license conditions, without requiring permission from the original product distributor; this is referred to as redistribution. Redistribution can be of the original form of the product, or a modified form that contains code modified by the original code recipient.

Proprietary Software:

Proprietary software can only be bought, obtained through a software license or leased from its publisher or developer. The company that owns it retains intellectual-property rights and may pose restrictions on its use and distribution. Its source code is usually handled as a trade secret and cannot be modified by its end users. The restrictions imposed by vendors are listed in the product's terms-of-service agreement or end-user license agreement.

Properties of Proprietary Software:

i. Increased Functionality and Convenience

Proprietary systems are easier to use and learn, leading to faster work processes. Skype, for example, is used by organizations worldwide. It takes minutes to sign up for an account and make international phone calls or conduct video interviews online. On top of that, your customers, suppliers and employees may already have a Skype account, so they know how to use it.

ii. Superior Customer Support

Open-source software can be difficult to install and set up. Customizing it isn't easier either. Plus, your staff may not be familiar with the program and may need additional training.

The average employee lacks the expertise to use open-source programs. Therefore, your team members may need help with most tasks. They will spend hours trying to figure things out instead of focusing on the tasks at hand.

iii. Lower Maintenance Costs

As a small-business owner, you may prefer open-source software due to its low cost. Most programs are free or cost next to nothing. The downside is that you may end up paying a lot more for setup, maintenance and customization than you'd pay with proprietary software.

Some open-source programs are difficult to install and set up, so you may need to call an expert to do the job. In some cases, new hardware may be necessary to use the software. If your employees are not familiar with the program, they will need support and training, which

may further increase the costs. Updating the software, testing new versions and applying patches isn't cheap either.

iv. Stronger Competitive Advantage

Proprietary technology enables organizations to be more profitable, productive and innovative. This is particularly true for software-development companies, which often use custom programs at the core of their business model.

By creating your own proprietary software, you can fill a gap in the market and provide customers with a valuable and unique offering. Over time, you may gain market share and build customer loyalty. Open-source software, on the other hand, may diminish your competitive advantage.

v. Secure Financing for Your Business

Nearly one-third of startups close their doors because they run out of capital. Developing proprietary technology doesn't guarantee success, but it could make it asier to secure financing for your small business. Plus, you will be able to charge higher prices because no other company offers the same product as you do.

As it turns out, big data investors prefer to put their money in companies selling proprietary software — or at least something other than open-source software, such as proprietary add-ons. This kind of technology isn't restricted by what already exists in the market.