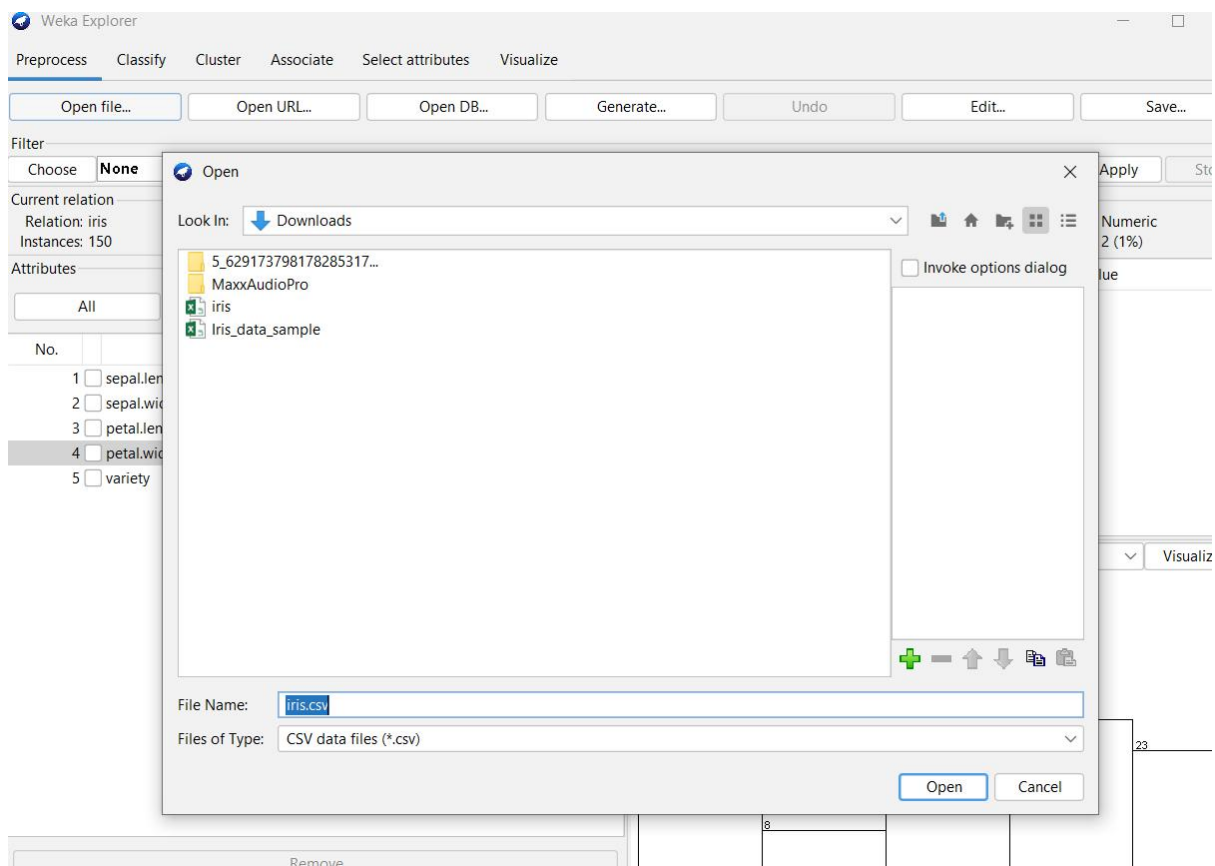**Date: 3/2/2022**

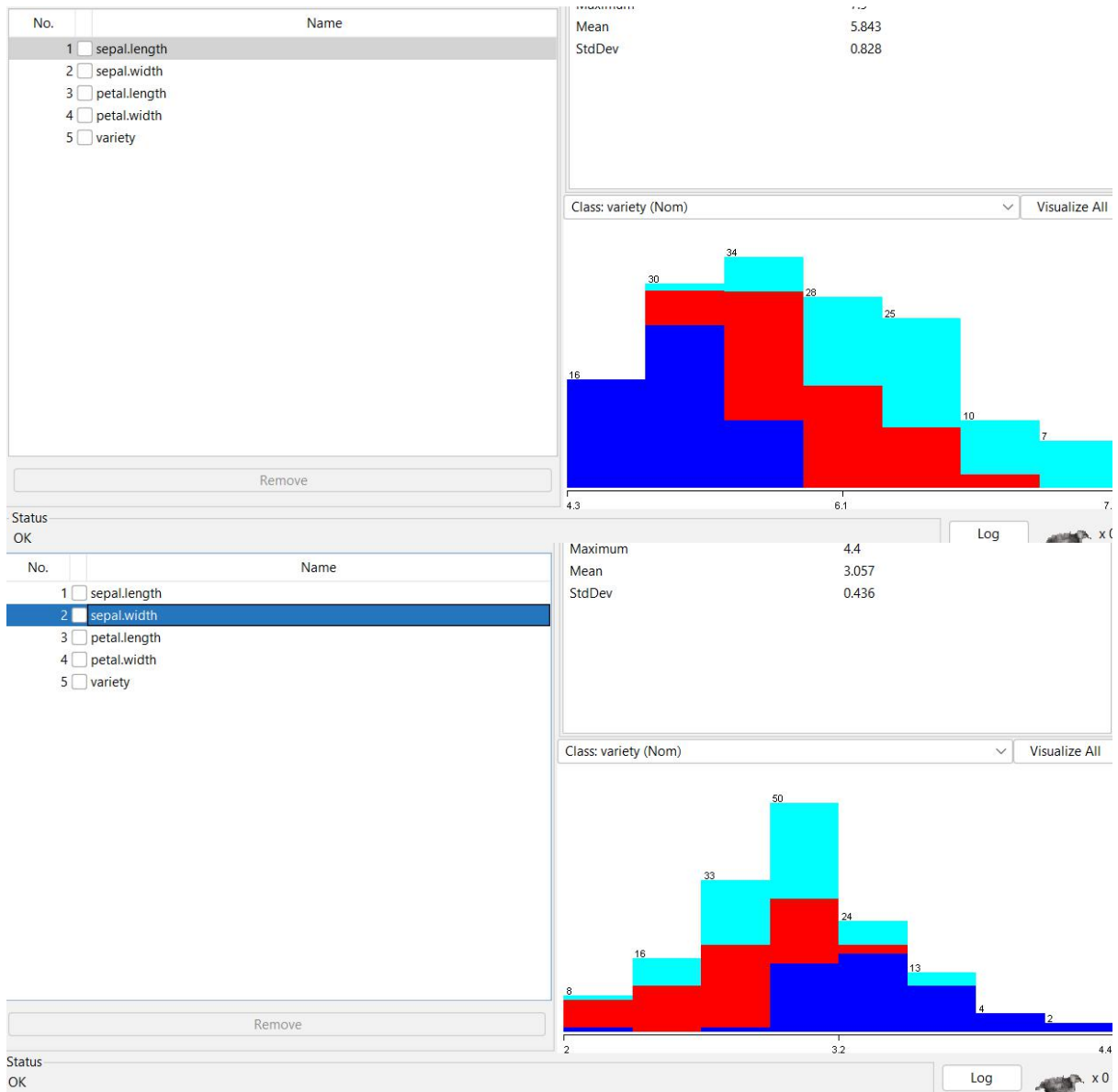# Third Year B.Tech. Sem VI 2021-22
Software Engineering Tool Lab

T1-2019BTECS00024 - Kshitij Amit Sabale

Practical No. 1

Q1.Weka

No.		Name
1	☐	sepal.length
2	☐	sepal.width
3	☐	petal.length
4	☐	petal.width
5	☐	variety

Remove

Status

OK

Maximum   7.3
Mean   5.843
StdDev   0.828

Class: variety (Nom)   ⌄   Visualize All

34
30
28
25
16
10
7
4.3   6.1   7.

Log   x

No.		Name
1	☐	sepal.length
2	☑	sepal.width
3	☐	petal.length
4	☐	petal.width
5	☐	variety

Remove

Maximum   4.4
Mean   3.057
StdDev   0.436

Class: variety (Nom)   ⌄   Visualize All

50
33
24
16
13
8
4
2
2   3.2   4.4

Status

OK

Log   x 0

No.	Name
1	sepal.length
2	sepal.width
3	petal.length
4	petal.width
5	variety

Remove

Status
OK

Maximum 6.9
Mean 3.758
StdDev 1.765

Class: variety (Nom) ▾   Visualize All

50

47

34

16

3

1   3.95   6.9

Log   🐦 x 0

No.	Name
1	sepal.length
2	sepal.width
3	petal.length
4	petal.width
5	variety

Remove

Status
OK

Maximum 2.5
Mean 1.199
StdDev 0.762

Class: variety (Nom) ▾   Visualize All

49

41

29

23

8

0.1   1.3   2.5

Log   🐦 x 0

No.	Name
1	sepal.length
2	sepal.width
3	petal.length
4	petal.width
5	variety

2	Versicolor	50	50
3	Virginica	50	50

Class: variety (Nom)   Visualize All

50   50   50

Remove

Status
OK

Log   x 0

No.	Name
1	sepal.length
2	sepal.width
3	petal.length
4	petal.width
5	variety

Maximum   6.9
Mean   3.758
StdDev   1.765

Class: sepal.width (Num)   Visualize All

50

47

34

16

3

1   3.95   6.9

Remove

Status
OK

Log   x 0

Q2. 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.

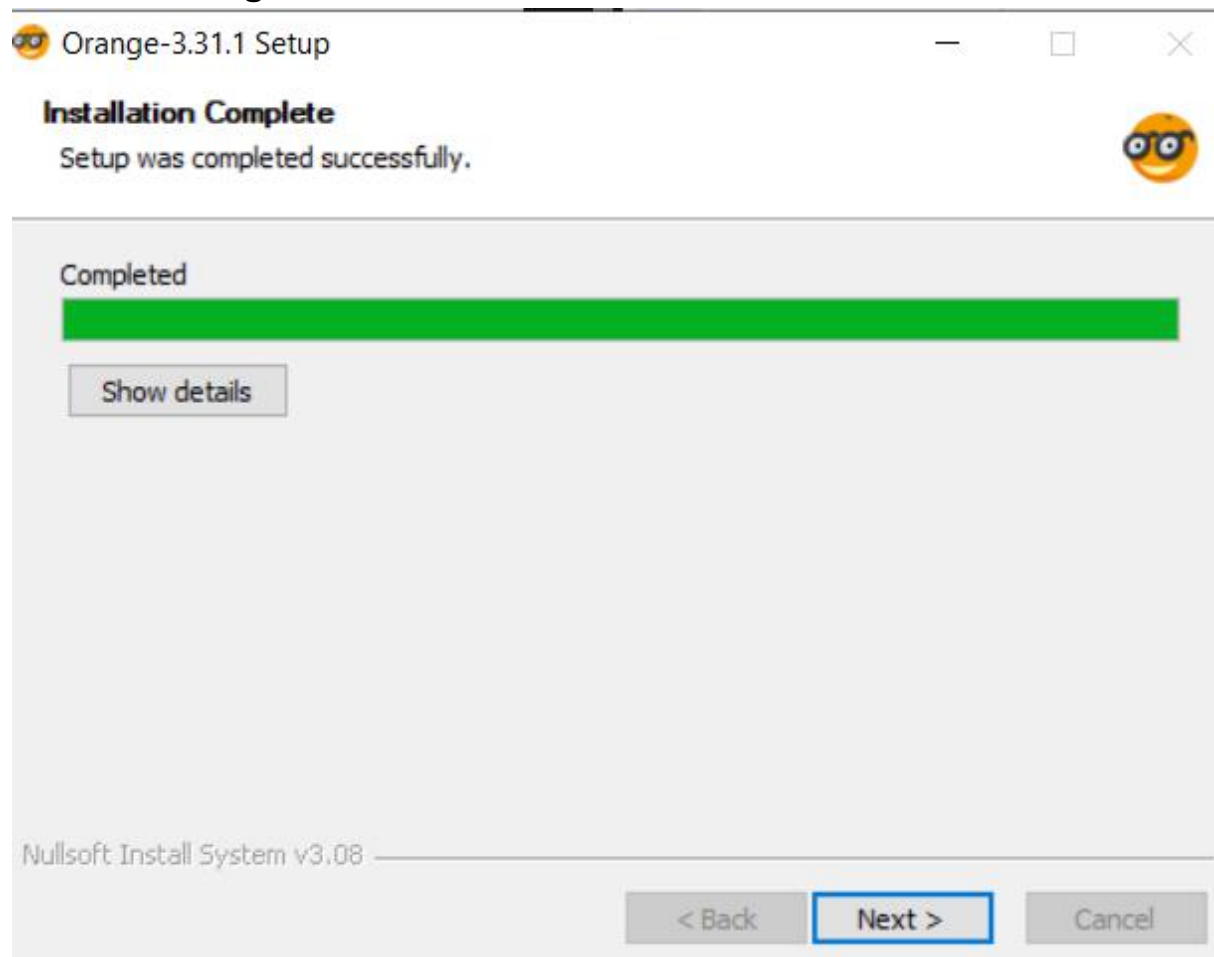**a. Install orange**

b. Show data distribution

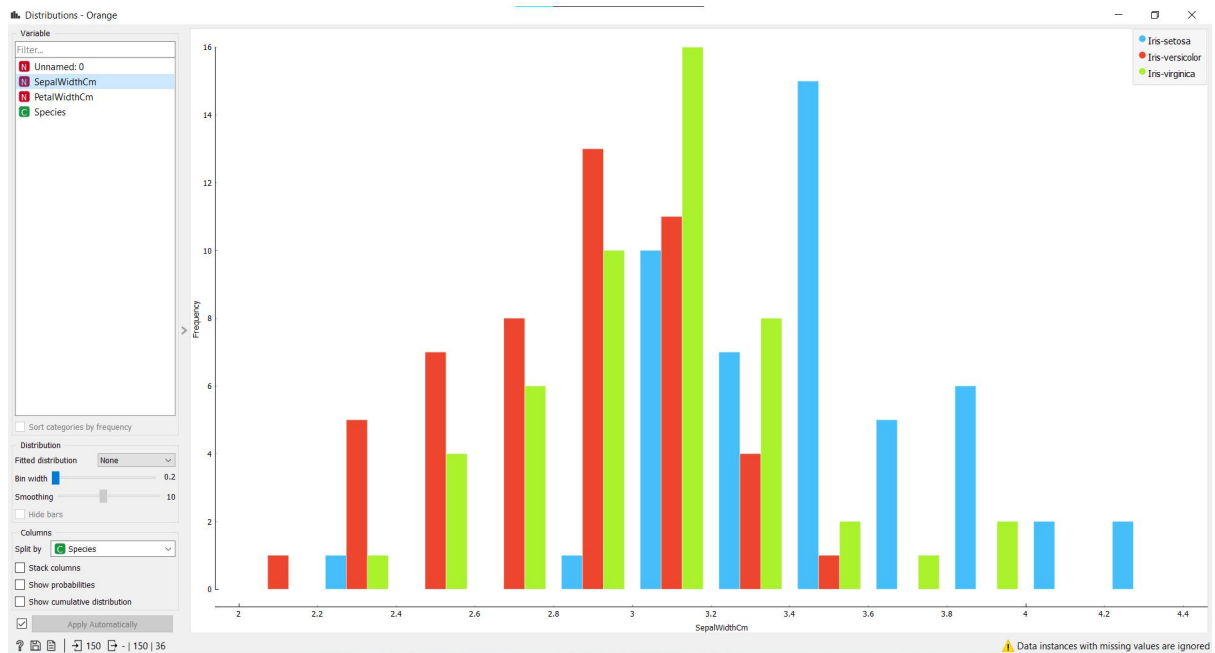**c. Show linear projection**

d. Show FreeViz

Dataset :
https://drive.google.com/file/d/1m6sKI1Dap0XK6Bw1edUd5PohwpPwXnd9/view
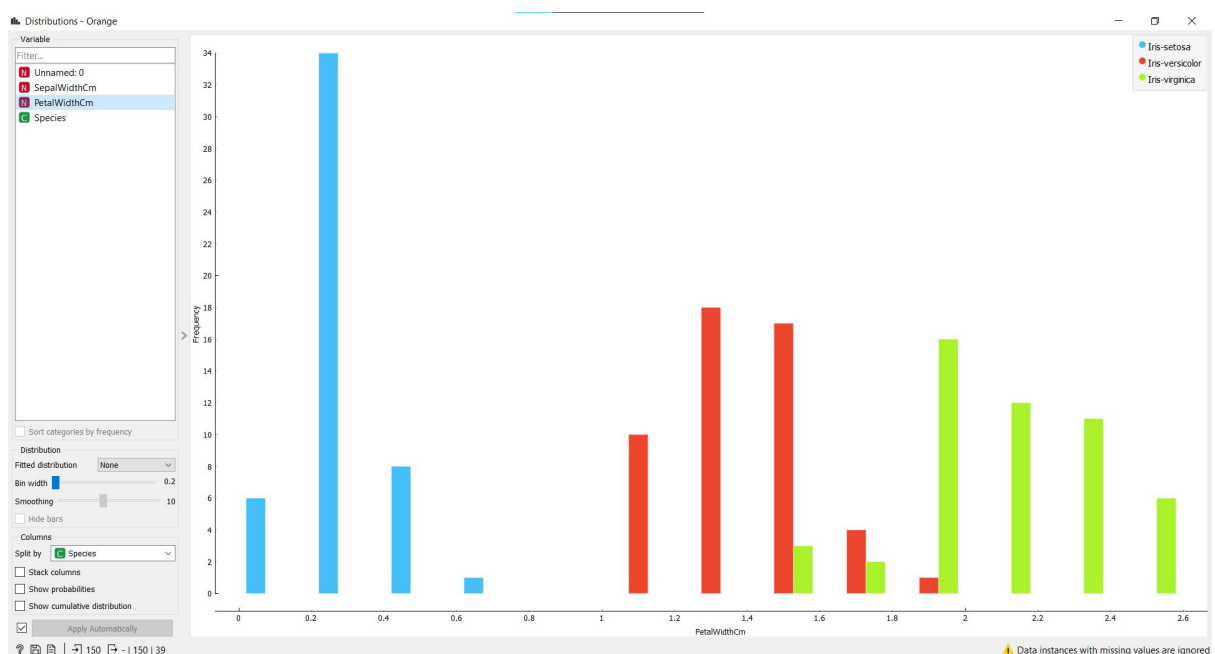
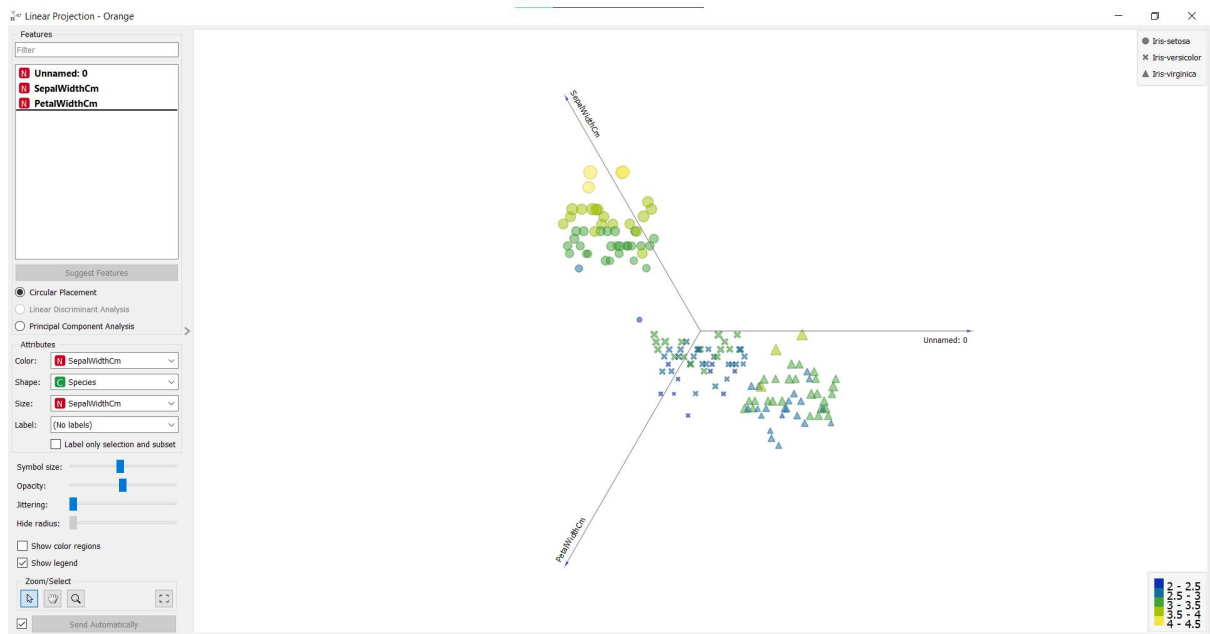**a.Install orange**

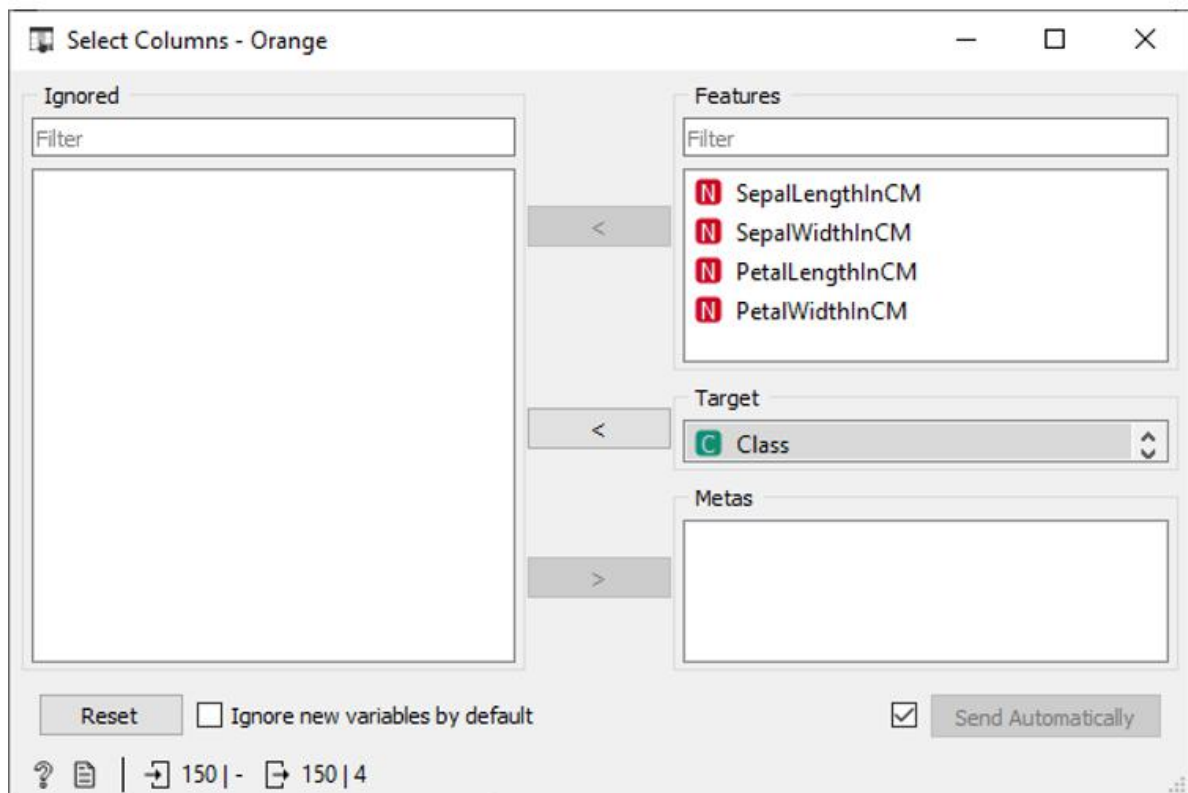**b.Show data distribution**

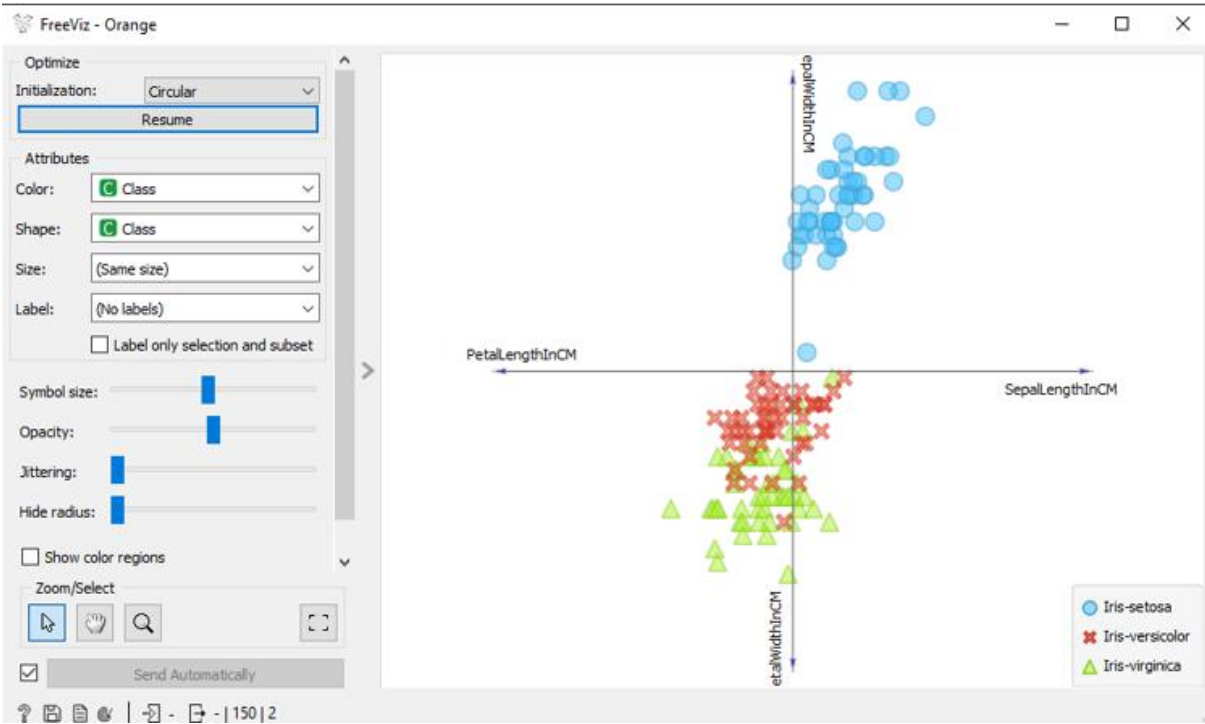On the basis of SepalWidth:

On the basis of PetalWidth:

c. Show linear projection

d. Show FreeViz

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

1. Free Software:

Free software (or libre software is computer software distributed under terms that allow users to run the software for any purpose as well as to study, change, and distribute it and any adapted versions. Free software is a matter of liberty, not price; all users are legally free to do what they want with their copies of a free software (including profiting from them) regardless of how much is paid to obtain the program. Computer programs are deemed "free" if they give end-users (not just the developer) ultimate control over the software and, subsequently, over their devices.

The right to study and modify a computer program entails that source code—the preferred format for making changes—be made available to users of that program. While this is often called "access to source code" or "public availability", the Free Software Foundation (FSF) recommends against thinking in those terms, because it might give the impression that users have an obligation (as opposed to a right) to give non-users a copy of the program.

2. Open source Software :
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. Here the software is developed and tested through open collaboration. This software is managed by an open source community of developers. It provides community support as well as commercial support if available for maintenance. We can get it for free of cost. This software also sometimes comes with license and sometimes does not. This license provides some rights to users like.
- Software can be used for any purpose
- Allows to study how software works
- Freedom to modify and improve the program
- No restrictions on redistributions

Some **examples of Open source software** includes Android, Ubuntu, Firefox, Open Office etc.

3. Proprietary Software :
Proprietary software is a computer software where the source codes are not publicly not available only the company which has created can modify it. Here the software is developed and tested by the individual or organization by

which it is owned not by public. This software is managed by an closed team of individuals or groups that developed it. We have to pay to get this software and it commercial support if available for maintenance. The company gives a valid and authenticated license to the users to use this software. But this license put some restrictions on users also like.

- Number of installations of this software into computers
- Restrictions on sharing of software illegally
- Time period up to which software will operate
- Number of features allowed to use

| S.No.	OPEN SOURCE SOFTWARE	PROPRIETARY SOFTWARE
01.	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 not available only the company which has created can modify it.
02.	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.
03.	In open source software the source code is public.	In proprietary software the source code is protected.
04.	Open source software can be installed into any computer.	Proprietary software can be installed into any computer without valid license.
05.	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.
06.	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.
07.	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.
08.	Users can get open software for free of charge.	Users must have to pay to get the proprietary software.
09.	In open source software faster fixes of bugs and better security is availed due to the community.	In proprietary software the vendor is completely responsible for fixing of malfunctions.
10.	Examples are Android, Linux, Firefox, Open Office, GIMP, VLC Media player etc.	Examples are Windows, MacOS, Internet Explorer, Google earth

Q4.Anaconda

Dataset :
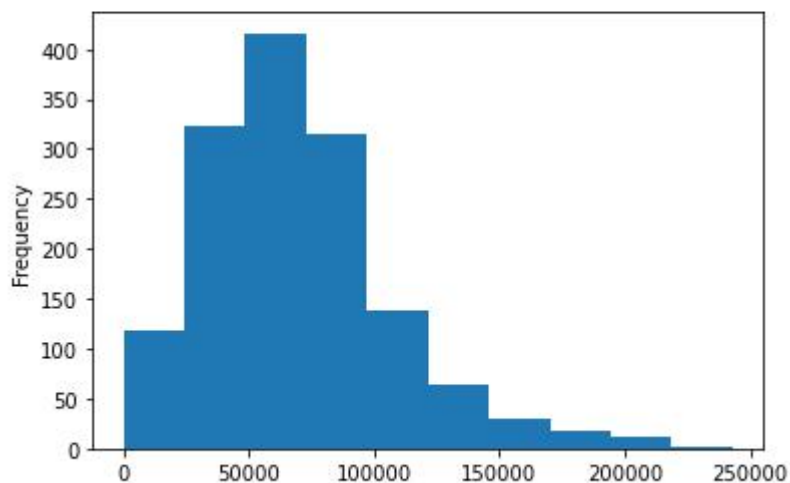https://drive.google.com/file/d/1i11BZFe8Xj9kNq7eeE9KOa_Iz1KhEdXJ/view
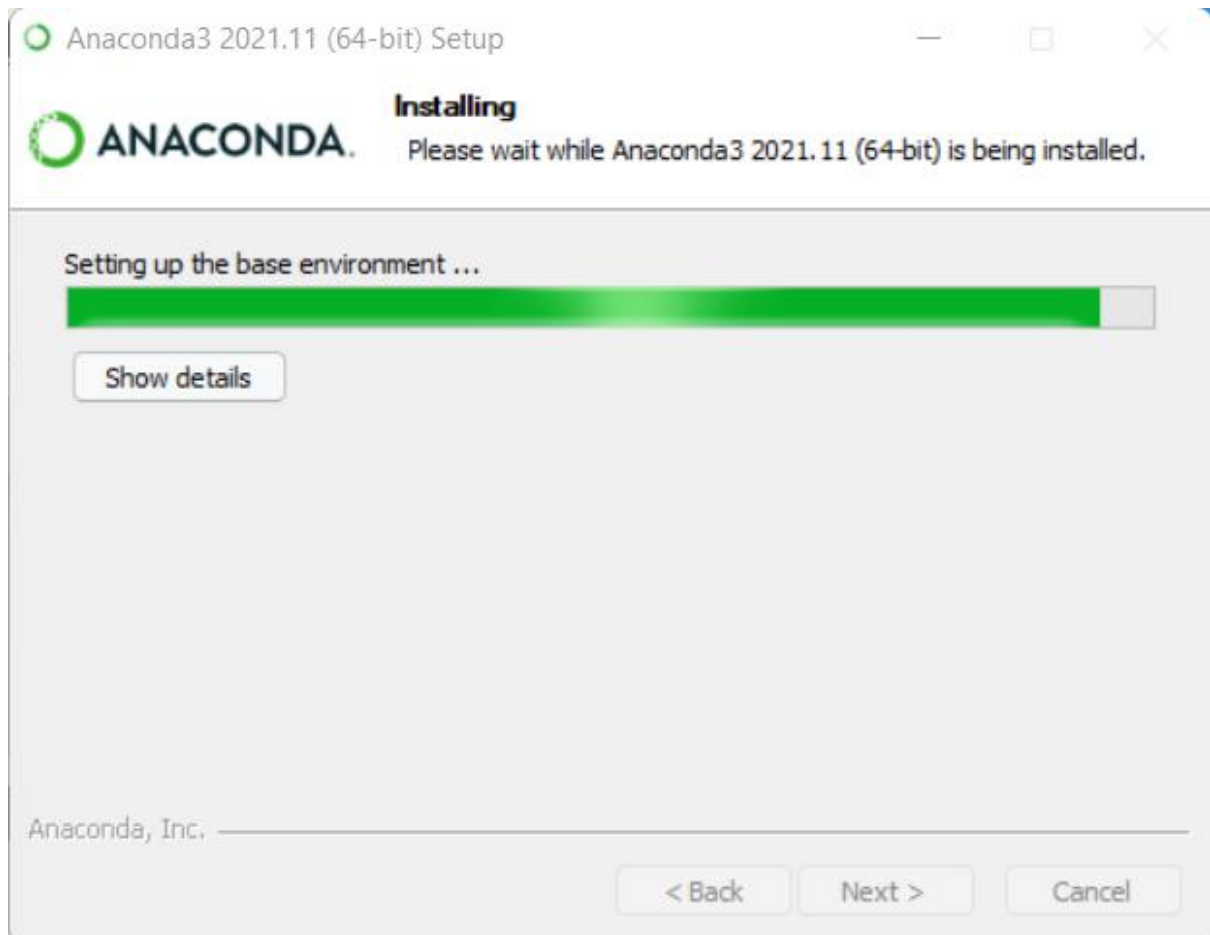
A. Scatter plot- Scatter plot of Price Vs Age

```python
import pandas as pd

df = pd.read_csv('Toyota.csv')

df.plot.scatter(x="Age",y="Price")
```
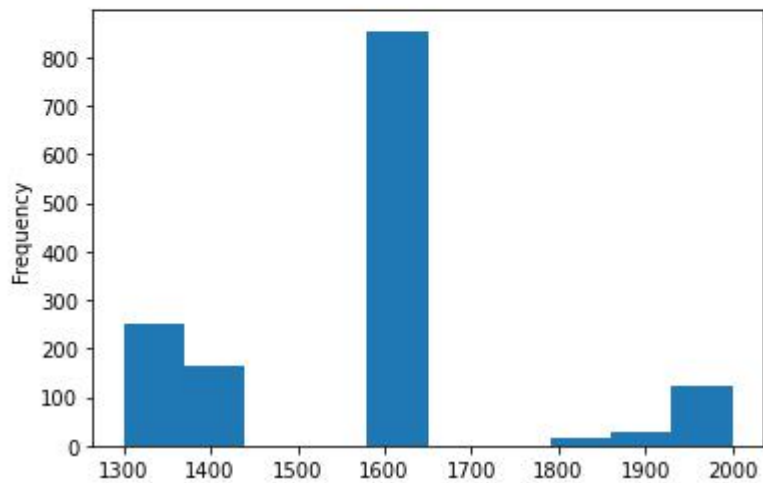
```
<AxesSubplot:xlabel='Age', ylabel='Price'>
```
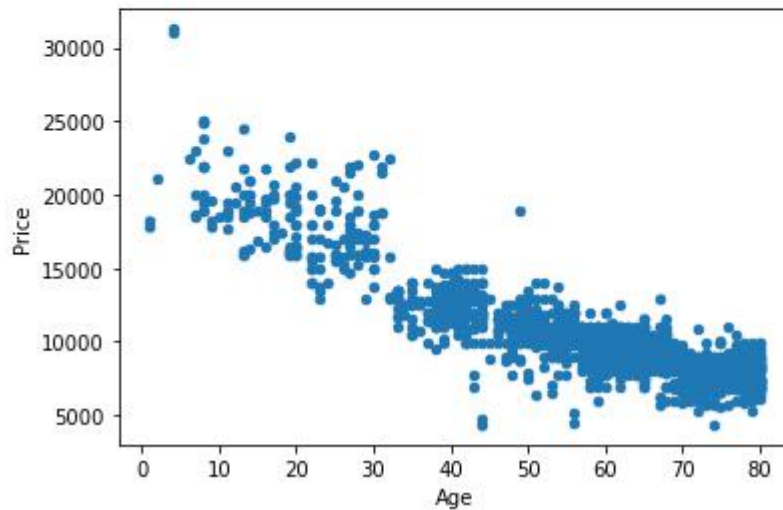

B. Histogram – Kilometer and CC

```python
import pandas as pd

df = pd.read_csv('Toyota.csv')
df['KM'].plot(kind='hist')
```

```
<AxesSubplot:ylabel='Frequency'>
```
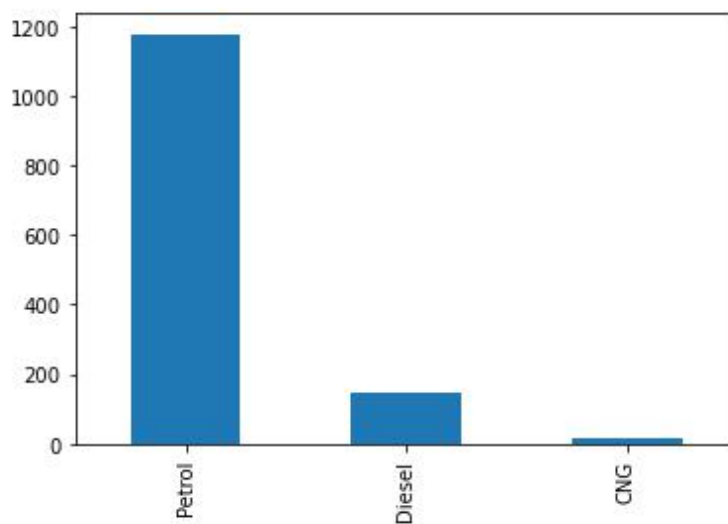
```
: import pandas as pd

  df = pd.read_csv('Toyota.csv')

  df['CC'].plot(kind='hist')
```

: `<AxesSubplot:ylabel='Frequency'>`

C. Bar Plot – Different Fuel Types

```
: import pandas as pd

  df = pd.read_csv('Toyota.csv')

  df['FuelType'].value_counts().plot(kind='bar')
```

: `<AxesSubplot:>`

Q5. Enlist some examples along with its purpose and properties (at least 10) of FOSS and proprietary software with respect to databases.

1. PostgreSQL

This relational database software has been around since 1997 and is the top choice in communities like Ruby, Python, Go, etc.

2. MariaDB

MariaDB was created as a replacement for MySQL by the same person who developed MySQL.

3. CockroachDB

The idea behind "cockroach" is that it's an insect built for survival. No matter what happens — predators, floods, eternal darkness, rotting food, bombing, the cockroach finds a way to survive and multiply.

4. ClickHouse

It uses every hardware to its maximum potential to approach each query faster. The peak performance of processing a query usually remains more than two terabytes each second.

5. Neo4j

Support for transactional applications and graph analytics. Data transformation abilities for digesting large-scale tabular data into graphs. Specialised query language (Cypher) for querying the graph database Visualisation and discovery features

6. Redis

When it comes to databases, it's almost too easy to overlook the existence of Redis. That's because Redis is an in-memory database and is mostly used in support functions like caching.

7. SQLite

SQLite is a lightweight C library that provides a relational database storage engine. Everything in this database lives in a single file (with a .sqlite extension)

that you can put anywhere in your filesystem. And that's all you need to use it! Yes, no "server" software to install and no service to connect to.

8. Cassandra

Cassandra belongs to what's known as the "columnar" family of databases. The storage abstraction in Cassandra is a column rather than a row. The idea here is to store all the data in a column physically together on the disk, minimising seek time.

9. Timescale

The timescale is a type of what's called a "time series" database. It's different from a traditional database in that time is the primary axis of concern, and the analytics and visualisation of massive data sets is a top priority

10. CouchDB

is a neat little database solution that sits quietly in a corner and has a small but dedicated following. It was created to deal with the problems of a net loss and eventual resolution of data, which happens to be a problem so messy that developers would instead switch jobs than deal with it