

DATA VISUALISATION

PL

PRELIMINARIES

We will be deploying two Python modules dedicated to data visualisation: `Matplotlib` and `Seaborn`, so we will install them first.

```
1 pip install matplotlib
```

```
2
```

```
3 pip install seaborn
```

```
1 import pandas as pd
```

```
2
```

```
3 from matplotlib import pyplot as plt
```

```
4
```

```
5 import seaborn as sns
```

LOCAL DATA

Fisher's dataset is publicly available from [Scikit-learn](#) as well as a comma-separated variables file from a myriad of online sources.

We use the [Pandas](#) module to upload the data into a suitable Python data structure.

```
1 # github users
2 FILE = "../data/iris.csv"
3
4 # colab and general users
5 #FILE = "/iris.csv"
6
7
8 iris = pd.read_csv(FILE)
9
10 iris
```

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalW
0	1	5.1	3.5	1.4	0.2

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalW
1	2	4.9	3.0	1.4	0.2
2	3	4.7	3.2	1.3	0.2
3	4	4.6	3.1	1.5	0.2
4	5	5.0	3.6	1.4	0.2
...
145	146	6.7	3.0	5.2	2.3
146	147	6.3	2.5	5.0	1.9

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalW
147	148	6.5	3.0	5.2	2.0
148	149	6.2	3.4	5.4	2.3
149	150	5.9	3.0	5.1	1.8

150 rows × 6 columns

SEABORN

Seaborn supplies visualisation functions that are oriented to Python data structures.

Data becomes an input parameter to a Seaborne object.

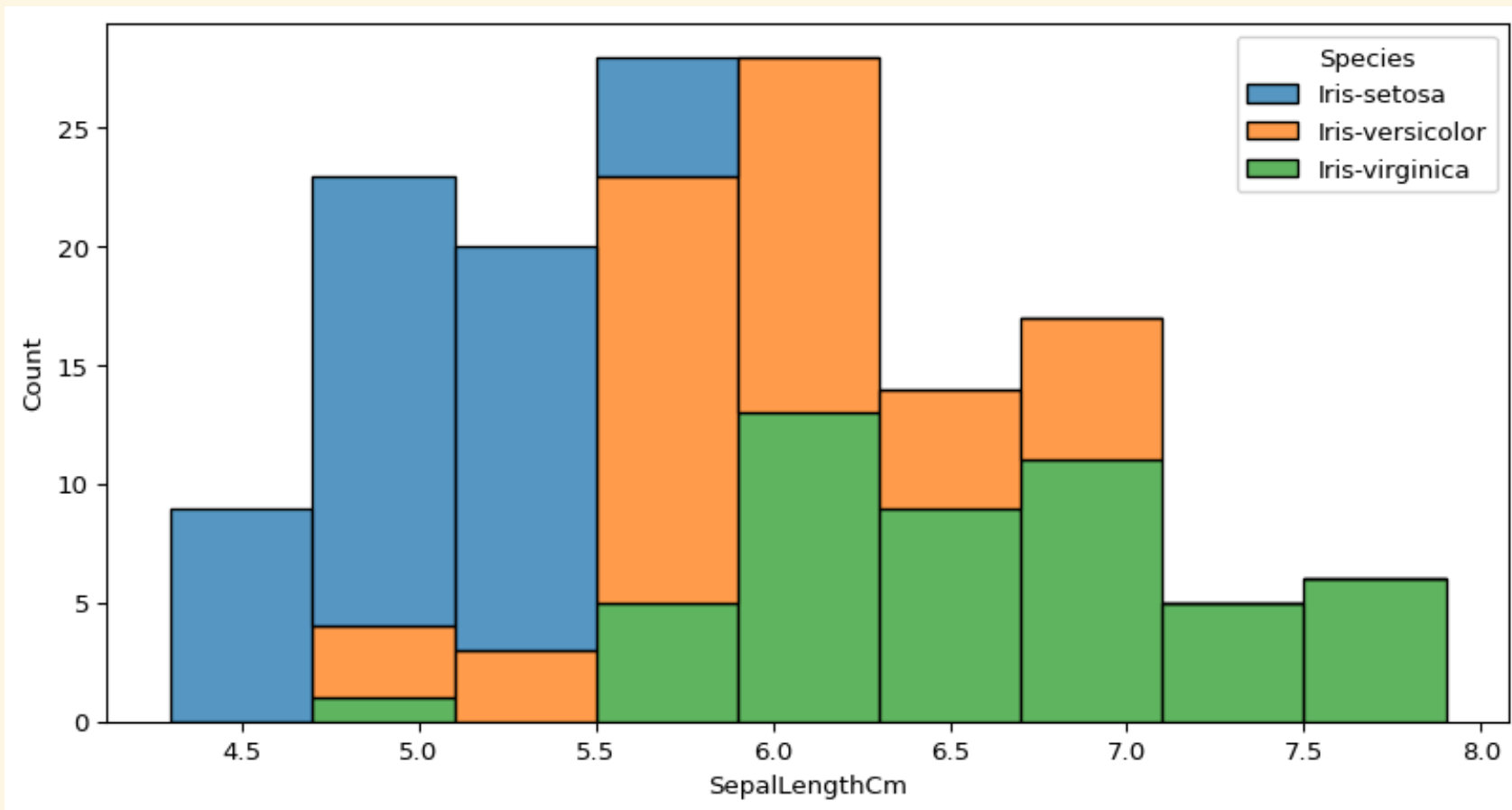
Manuals: seaborn.pydata.org/

Paper: [10.21105/joss.03021](https://doi.org/10.21105/joss.03021)

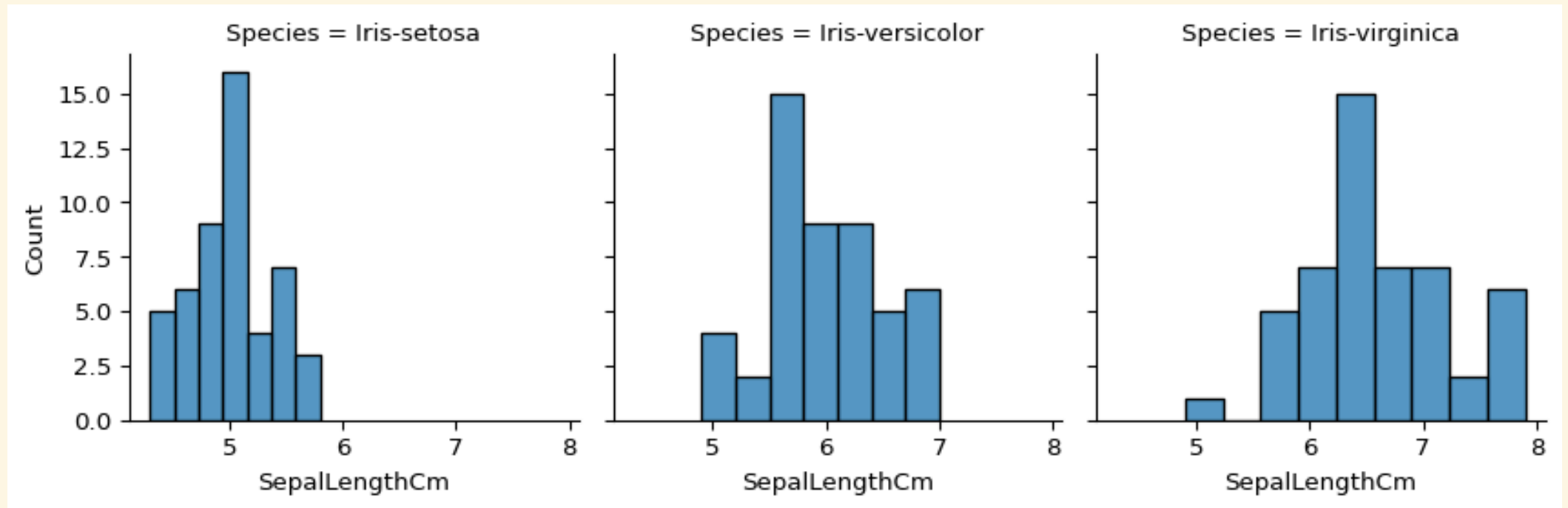
PLOT SEPAL-LENGTH IN STACKED BARS

The `hue` parameter controls class separation

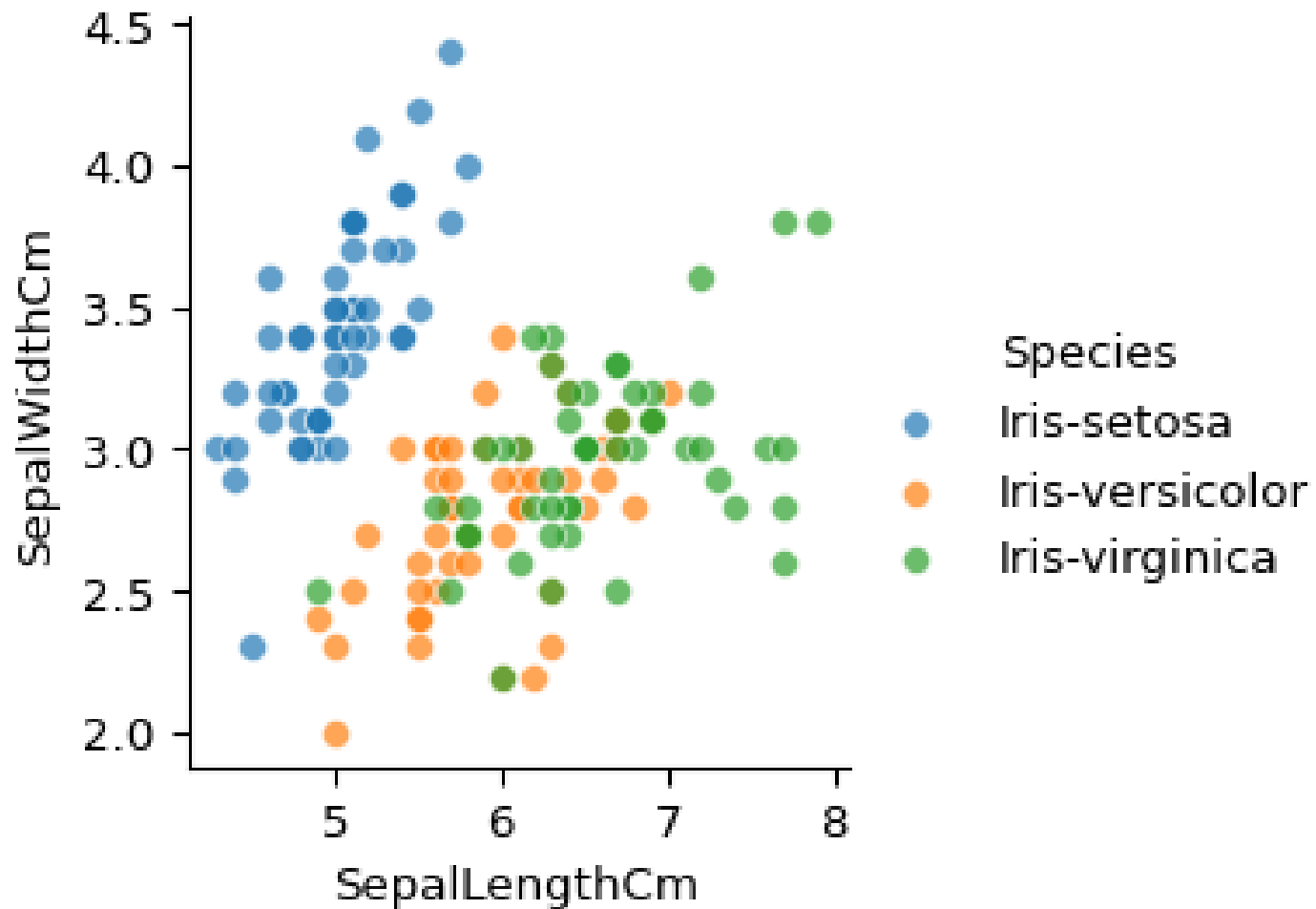
```
1 sns.histplot(data=iris, x="SepalLengthCm", hue=iris["Species"], multiple="stack")  
2  
3 plt.show()
```



```
1 # Plot histogram of classes of sepal length dimension in three sub plots
2 g = sns.FacetGrid(iris, col="Species")
3
4 g.map(sns.histplot, "SepalLengthCm")
5
6 # plt.show()
```




```
1 # Scatter plot between sepal length vs petal length
2 g = sns.FacetGrid(iris, hue="Species")
3 g.map(sns.scatterplot, "SepalLengthCm", "SepalWidthCm", alpha=.7)
4 g.add_legend()
5
6 plt.show()
```



PLOT HISTOGRAMS AND SCATTERPLOTS FOR ALL DIMENSIONS IN ONE SINGLE FIGURE

Work with Seaborn's `add_legend` to personalise the data presentation.

```
1 g = sns.PairGrid(iris, hue="Species")
2
3 g.map_diag(sns.histplot)
4 g.map_offdiag(sns.scatterplot)
5 g.add_legend()
6
7 plt.show()
```

