

Seaborn_Guide

September 13, 2018

0.1 Seaborn Guide für die Visualisierung von Daten

Seaborn ist ein Python Paket zur Darstellung von Daten im Bereich Data Science und Machine Learning. In diesem Beispiel soll am Datensatz zur Klassifizierung von Tulpen die Verwendung der Bibliothek erläutert werden. Zunächst werden die benötigten Pakete eingebunden.

Weitere Informationen und Beispiele

1. <https://elitedatascience.com/python-seaborn-tutorial>
2. <https://seaborn.pydata.org/>

```
In [12]: import numpy as np
import scipy as sci
import seaborn as sea
import matplotlib.pyplot as plt
import pandas as pd
```

Importieren des Beispieldatensatzes mittels pandas

```
In [13]: df = pd.read_csv('/home/jme/PycharmProjects/sci_kit_learn_test/Iris_data_set.txt', sep=
```

0.1.1 Anlegen der HS EL Farbpalette

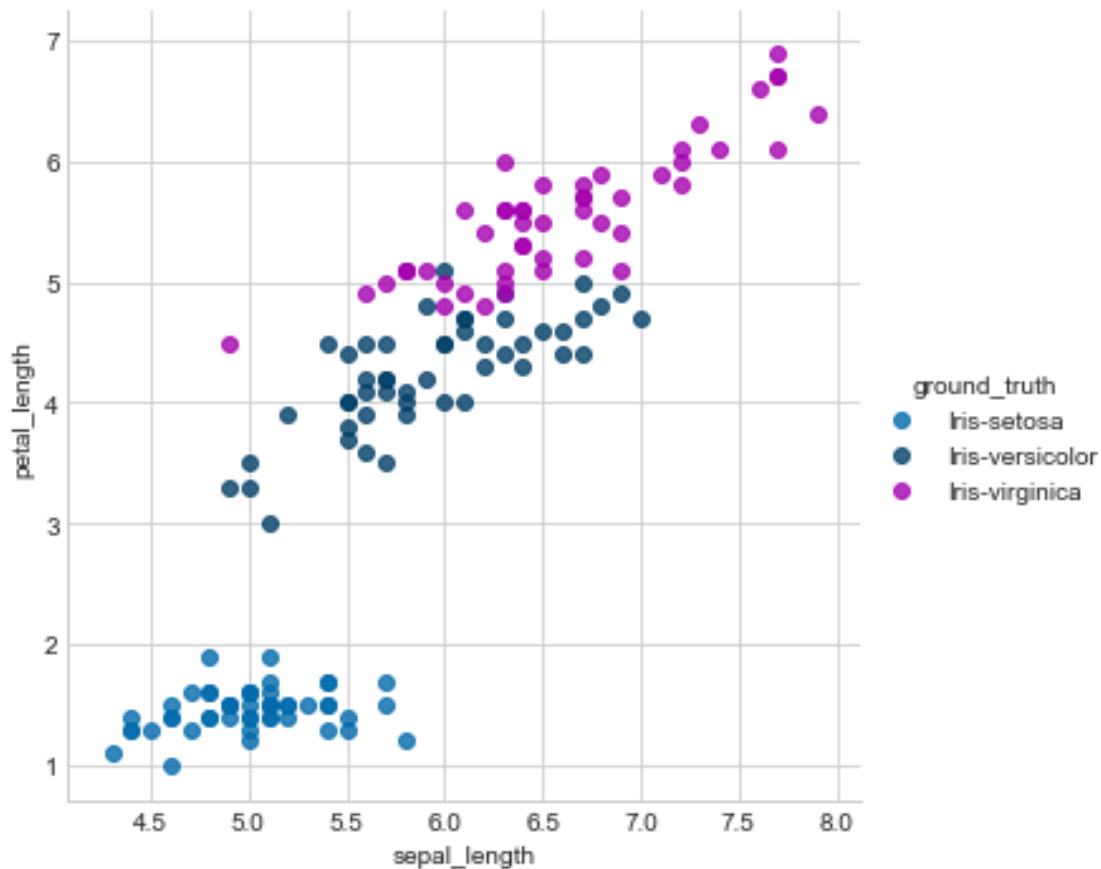
```
In [14]: hs_el_colors = ['#016AAC', # Light Blue
                        '#013F68', # Dark Blue
                        '#A400AF', # Light Purple
                        '#63006A', # Dark Purple
                        '#D3F800', # Light green
                        '#869D00', # Dark green
                        '#FF8E00', # Light gray
                        '#A25A00', # Dark gray
                        ]
id3as_colors = ['#e85511', # orange
                '#005a9f', # blue
                '#63b338', # green
                '#8e340b', # Dark orange
                '#0091ff', # Dark blue
                '#417425', # Dark green
                ]
```

0.1.2 Einfacher XY Scatter Plot mit Seaborn

- `x='sepal_length'` : Feature für die X-Achse aus df
- `y='petal_length'` : Feature für die y-Achse aus df
- `data=df` : Verweis auf das zuvor geladene Datenframe
- `fit_reg=False` : Keine lineare Fitting Linie darstellen
- `hue='ground_truth'` : Verweis auf die Zeile in der die Labels für den Datensatz stehen
- `palette=hs_el_colors` : Verwenden der zuvor angelegten HS EL Farbpalette

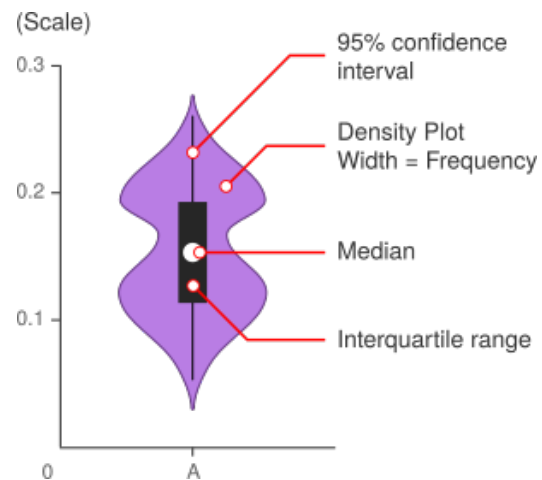
```
In [15]: sea.lmplot(x='sepal_length', y='petal_length', data=df, fit_reg=False, hue='ground_truth')
```

```
Out[15]: <seaborn.axisgrid.FacetGrid at 0x7f4b5ebf62b0>
```



0.1.3 Violinen Plot

https://datavizcatalogue.com/methods/violin_plot.html



Violine Plot

```
In [16]: # set whitegrid theme
         sea.set_style('whitegrid')

         # Violin plot
         sea.violinplot(data=df,palette=hs_el_colors)

         #Customize with Matplotlib
         plt.title('Violine Plot')

         plt.xlabel('Häufigkeit')

         plt.ylabel('Feature')

         plt.show()
```

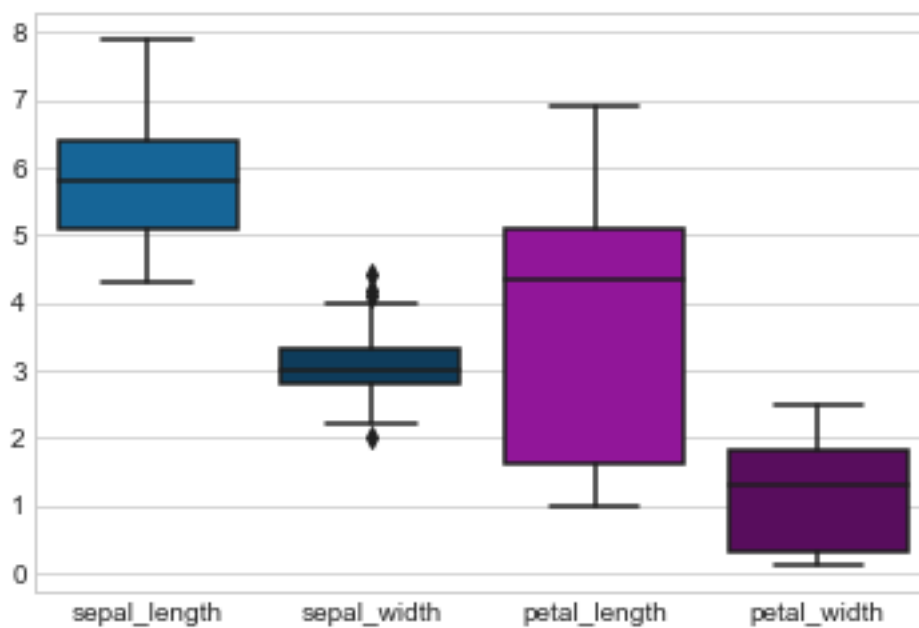


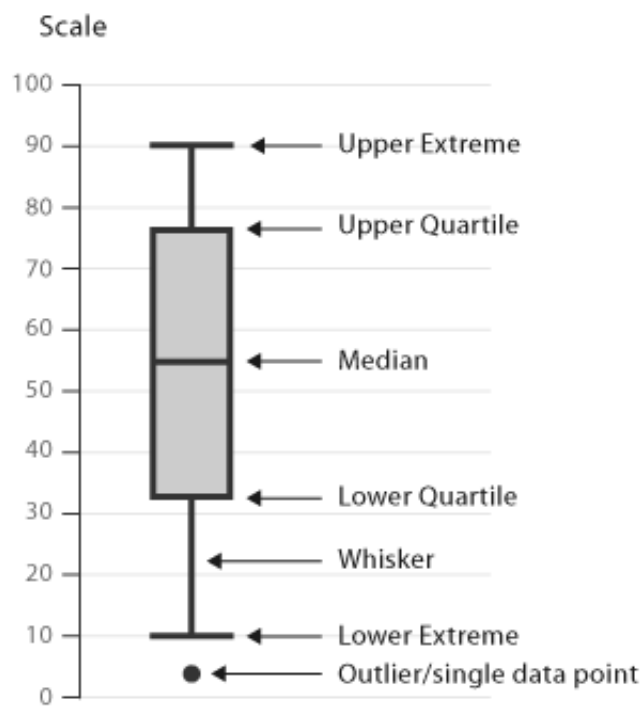
0.1.4 Box & Whisker Plot

https://datavizcatalogue.com/methods/box_plot.html

In [17]: `sea.boxplot(data=df,palette=hs_el_colors)`

Out[17]: `<matplotlib.axes._subplots.AxesSubplot at 0x7f4b5a494668>`



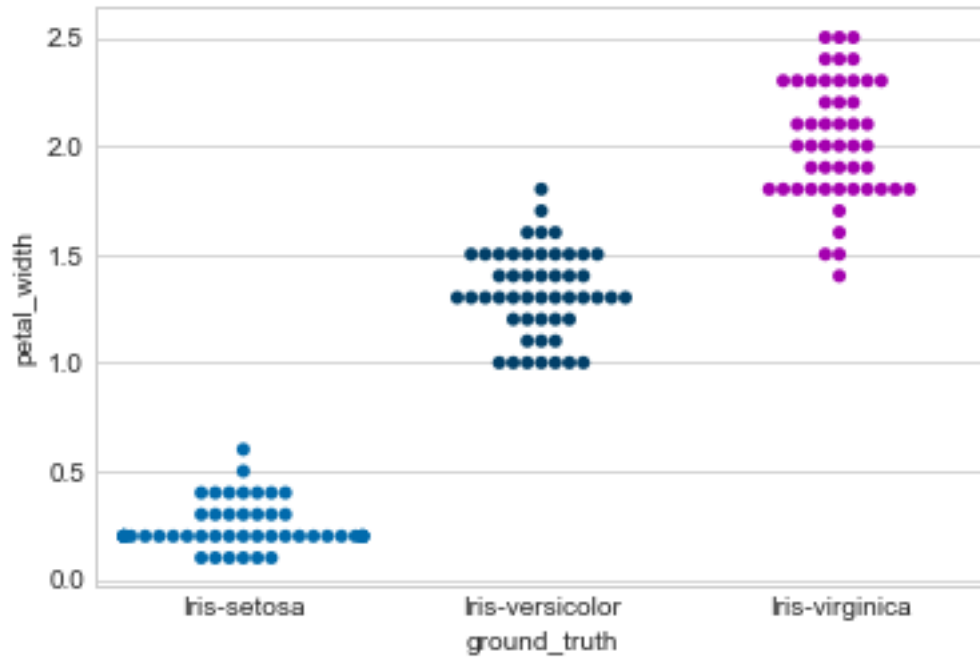


Box Plot

0.1.5 Swarm Plot

```
In [18]: sea.swarmplot(x='ground_truth', y='petal_width', data=df, palette=hs_el_colors)
```

```
Out[18]: <matplotlib.axes._subplots.AxesSubplot at 0x7f4b589410f0>
```

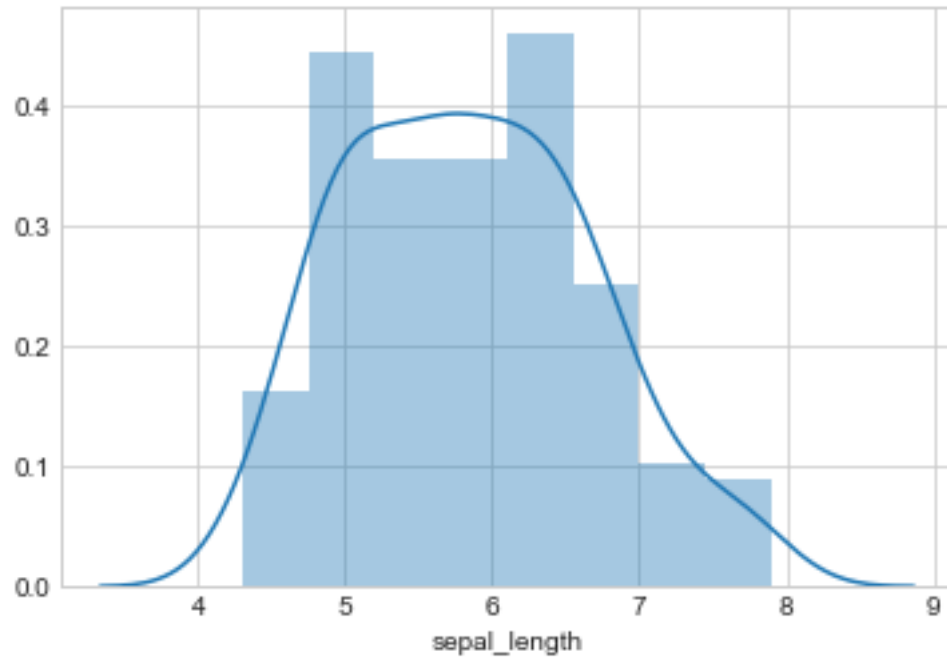


0.1.6 Histogramm

```
In [19]: # Distribution Plot (a.k.a. Histogram)
         sea.distplot(df.sepal_length)
```

/home/jme/anaconda3/lib/python3.6/site-packages/matplotlib/axes/_axes.py:6462: UserWarning: The warnings.warn("The 'normed' kwarg is deprecated, and has been "

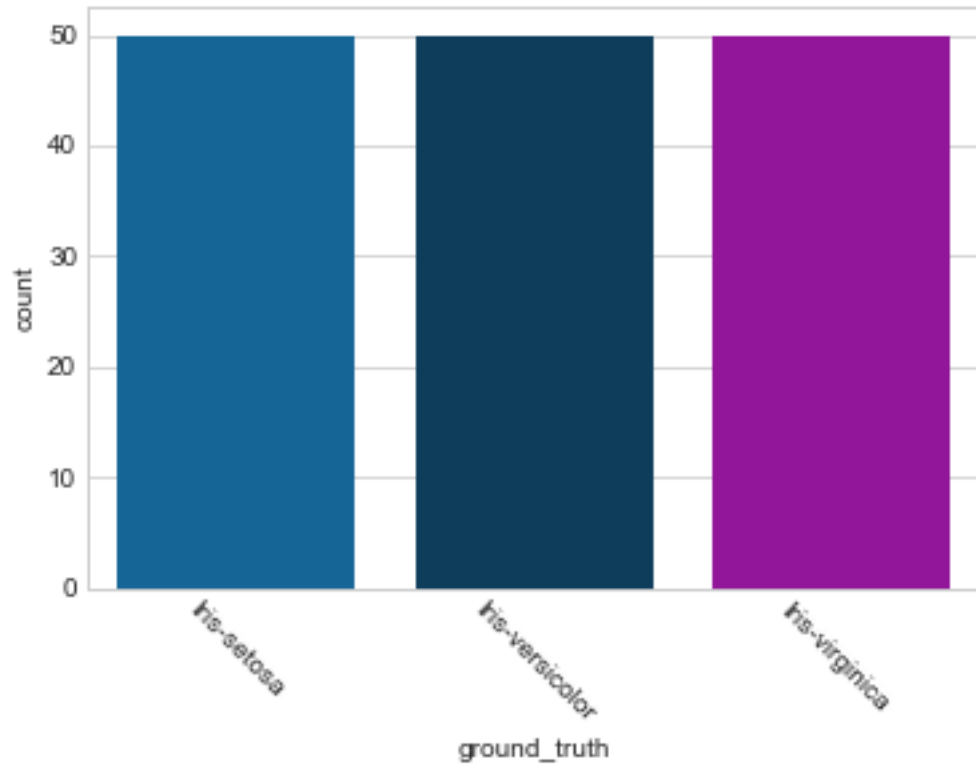
```
Out[19]: <matplotlib.axes._subplots.AxesSubplot at 0x7f4b58a0beb8>
```



0.1.7 Bar Plot

```
In [20]: sea.countplot(x='ground_truth', data=df, palette=hs_el_colors)
         # Rotate x-labels
         plt.xticks(rotation=-45)
```

```
Out[20]: (array([0, 1, 2]), <a list of 3 Text xticklabel objects>)
```



0.1.8 Pair Plot

Darstellung der Beziehungen zwischen Features

```
In [21]: iris = sea.load_dataset("iris")  
         sea.pairplot(iris);
```