session8

September 17, 2024

[1]:
```
import numpy as np
```

[2]:
```
import pandas as pd
```

[3]:
```
import seaborn as sns
```

[4]:
```
import matplotlib.pyplot as plt
```

[5]:
```
df =pd.read_csv('iris.csv')
```

[6]:
```
import warnings
warnings.filterwarnings("ignore")
```

[7]:
```
df.head(10)
```

[7]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa
5	5.4	3.9	1.7	0.4	setosa
6	4.6	3.4	1.4	0.3	setosa
7	5.0	3.4	1.5	0.2	setosa
8	4.4	2.9	1.4	0.2	setosa
9	4.9	3.1	1.5	0.1	setosa

[8]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
 #   Column        Non-Null Count  Dtype  
---  ------        --------------  -----  
 0   sepal_length  150 non-null    float64
 1   sepal_width   150 non-null    float64
 2   petal_length  150 non-null    float64
 3   petal_width   150 non-null    float64
 4   species       150 non-null    object 
```

```
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
```

[9]:
```
plt.style.use("fivethirtyeight")
plt.figure(figsize=(10,8))
sns.pairplot(df,hue="species",height=2.5,)
plt.show()
plt.tight_layout()
```

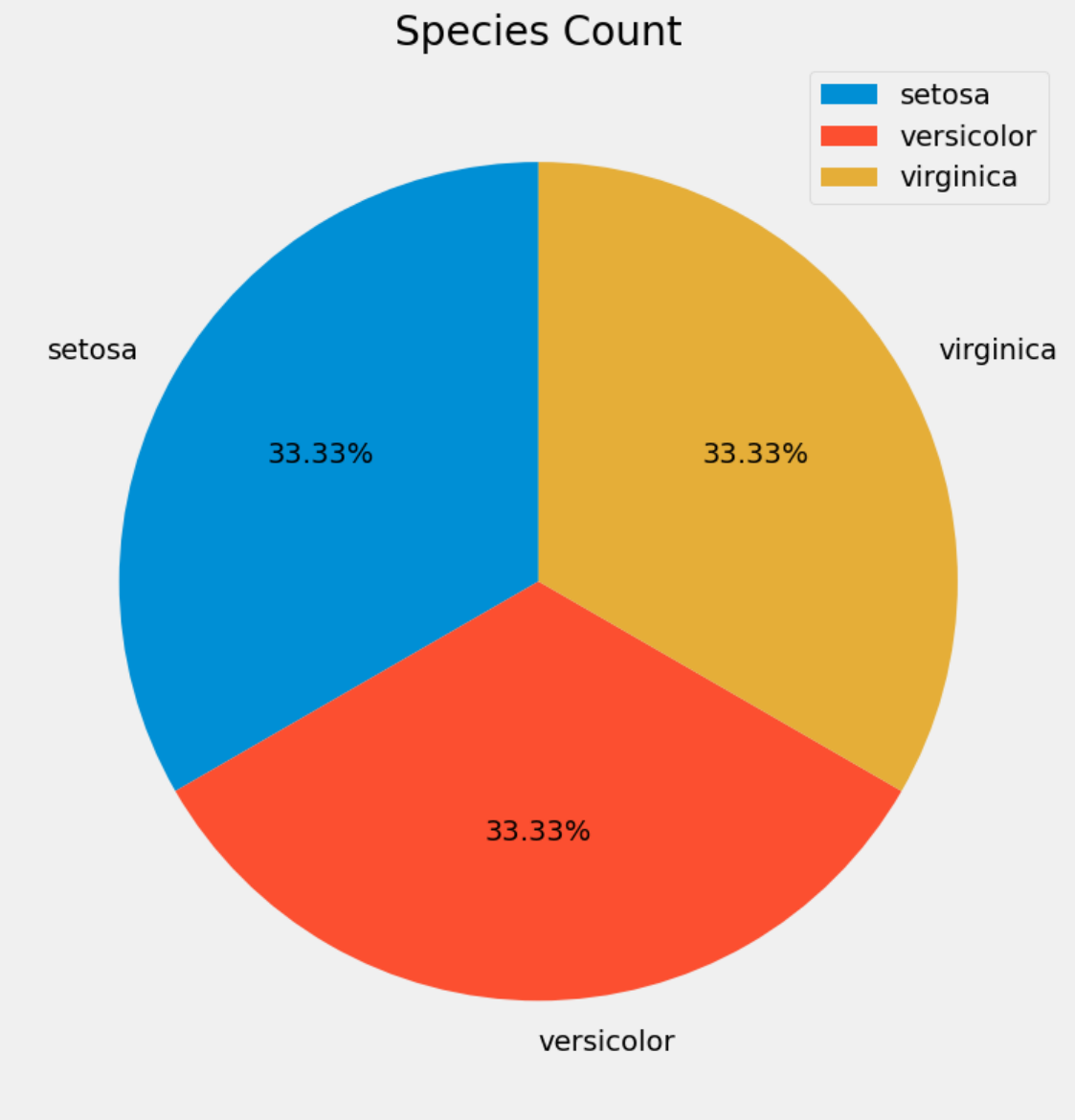
```
<Figure size 1000x800 with 0 Axes>
```


```
<Figure size 640x480 with 0 Axes>
```

[10]:
```
species=df["species"].value_counts()
label=df["species"].unique()
```

[11]:
```
plt.style.use("fivethirtyeight")
plt.figure(figsize=(8,8))
plt.pie(species,labels=label,autopct='%1.2f%%',startangle=90)
```

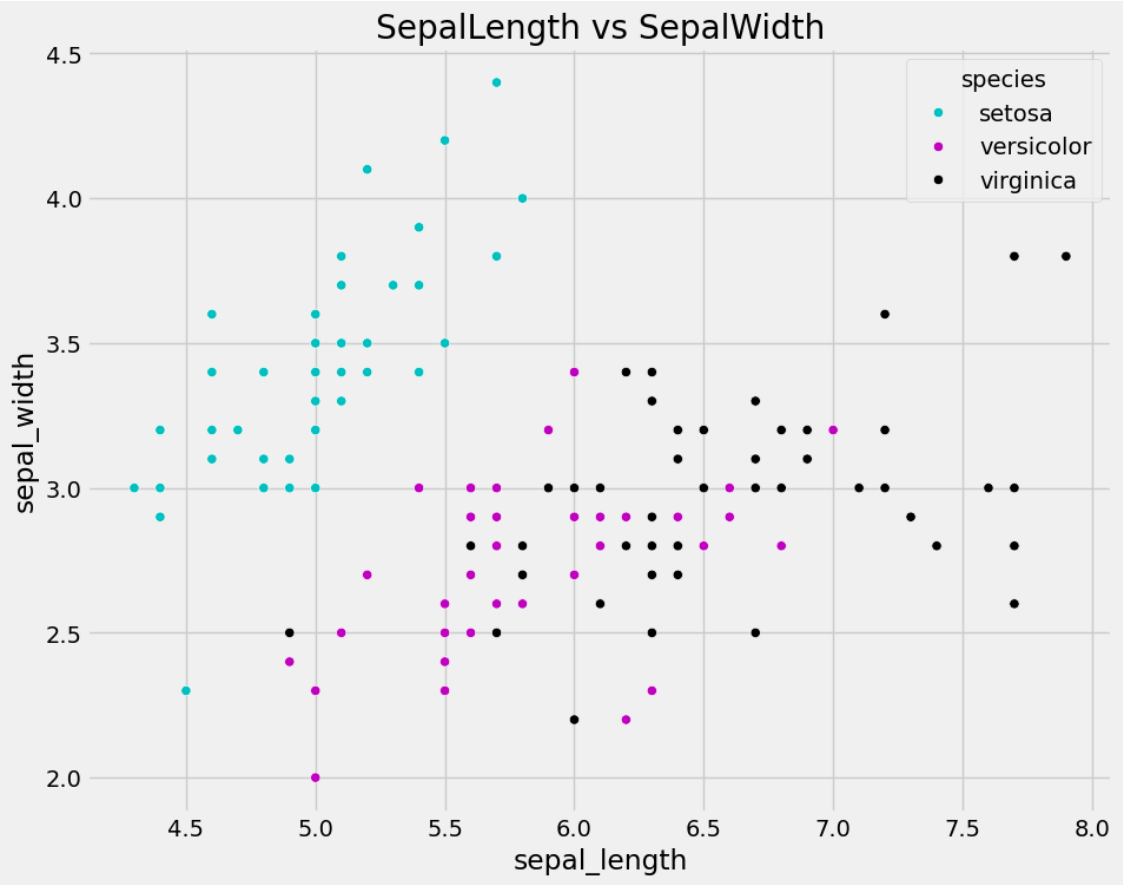
```
plt.title("Species Count")
plt.legend()
plt.tight_layout()
```



[12]:
```
plt.style.use("fivethirtyeight")
plt.figure(figsize=(10,8))
sns.scatterplot(
  x='sepal_length',
  y='sepal_width',
  hue='species',
  data=df,
```

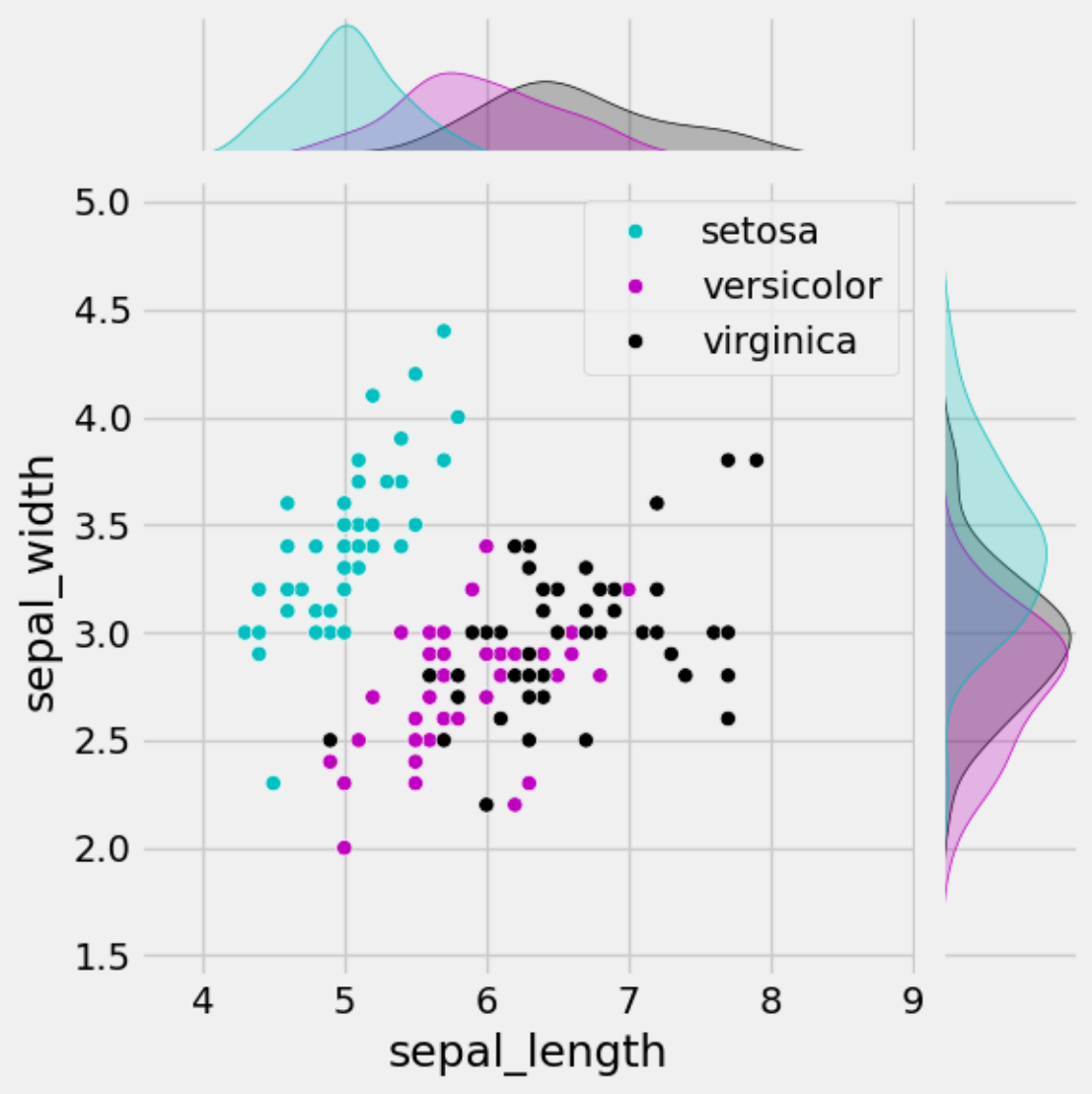
```
  palette={'setosa': 'c', 'versicolor': 'm', 'virginica': 'k'}
 )
plt.title('SepalLength vs SepalWidth')
```

[12]: Text(0.5, 1.0, 'SepalLength vs SepalWidth')



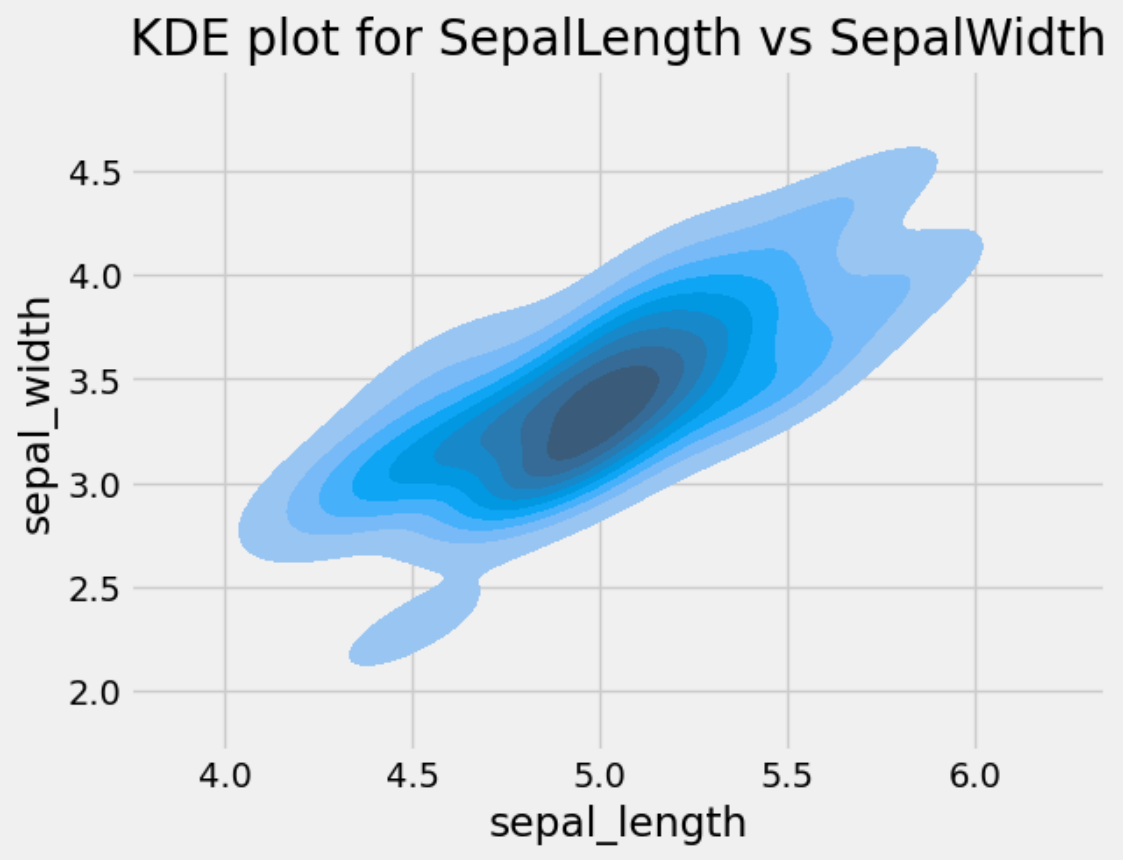
[13]:
```
plt.style.use("fivethirtyeight")
sns.
 ↪jointplot(x='sepal_length',y='sepal_width',data=df,kind='scatter',hue='species',
palette={'setosa': 'c', 'versicolor': 'm', 'virginica': 'k'})
plt.legend()
```

[13]: <matplotlib.legend.Legend at 0x227d3f49d60>

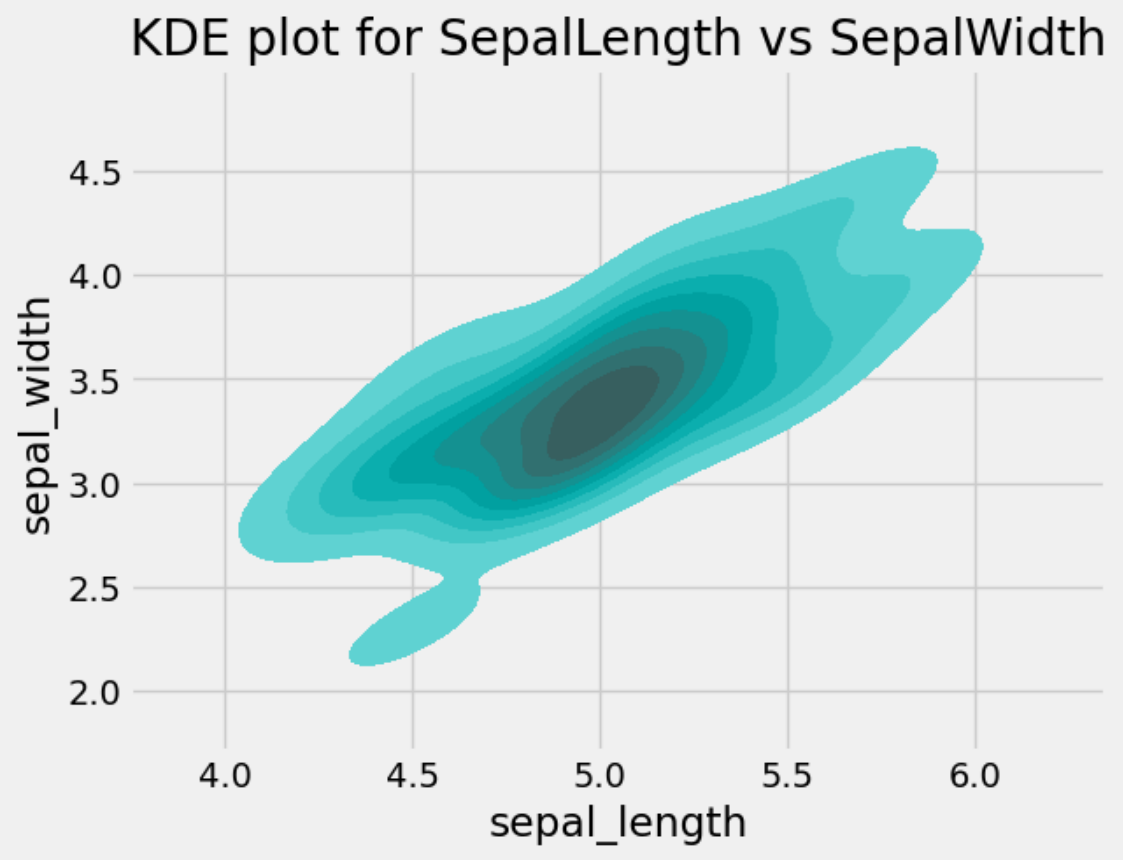


[14]:
```
setosa=df[df['species']=="setosa"]
sns.kdeplot(x='sepal_length',y='sepal_width',data=setosa,fill=True)
plt.title("KDE plot for SepalLength vs SepalWidth")
```

[14]: Text(0.5, 1.0, 'KDE plot for SepalLength vs SepalWidth')



[15]:
```
sns.kdeplot(x='sepal_length',y='sepal_width',data=setosa,fill=True,color='c')
plt.title("KDE plot for SepalLength vs SepalWidth")
plt.show()
```


KDE plot for SepalLength vs SepalWidth
sepal_width
4.5
4.0
3.5
3.0
2.5
2.0
4.0
4.5
5.0
5.5
6.0
sepal_length

[ ]: