In [33]:
```
#import libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns


import warnings
warnings.filterwarnings('ignore')


plt.style.use("fivethirtyeight")
%matplotlib inline
```

In [34]:
```
df=pd.read_csv('Iris.csv')
df.head()
```

Out[34]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa

In [35]:
```
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
```

In [36]:
```
df.describe()
```

Out[36]:

	sepal_length	sepal_width	petal_length	petal_width
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.054000	3.758667	1.198667
std	0.828066	0.433594	1.764420	0.763161
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.350000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

In [37]:
```
df.shape
```

Out[37]:
```
(150, 5)
```

In [38]:
```
df.head()
```

Out[38]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa

In [39]:
```
df['species'].value_counts()
```

Out[39]:
```
Iris-setosa        50
Iris-versicolor    50
Iris-virginica     50
Name: species, dtype: int64
```

In [40]:
```
df.isnull().sum()
```

Out[40]:
```
sepal_length    0
sepal_width     0
petal_length    0
petal_width     0
species         0
dtype: int64
```

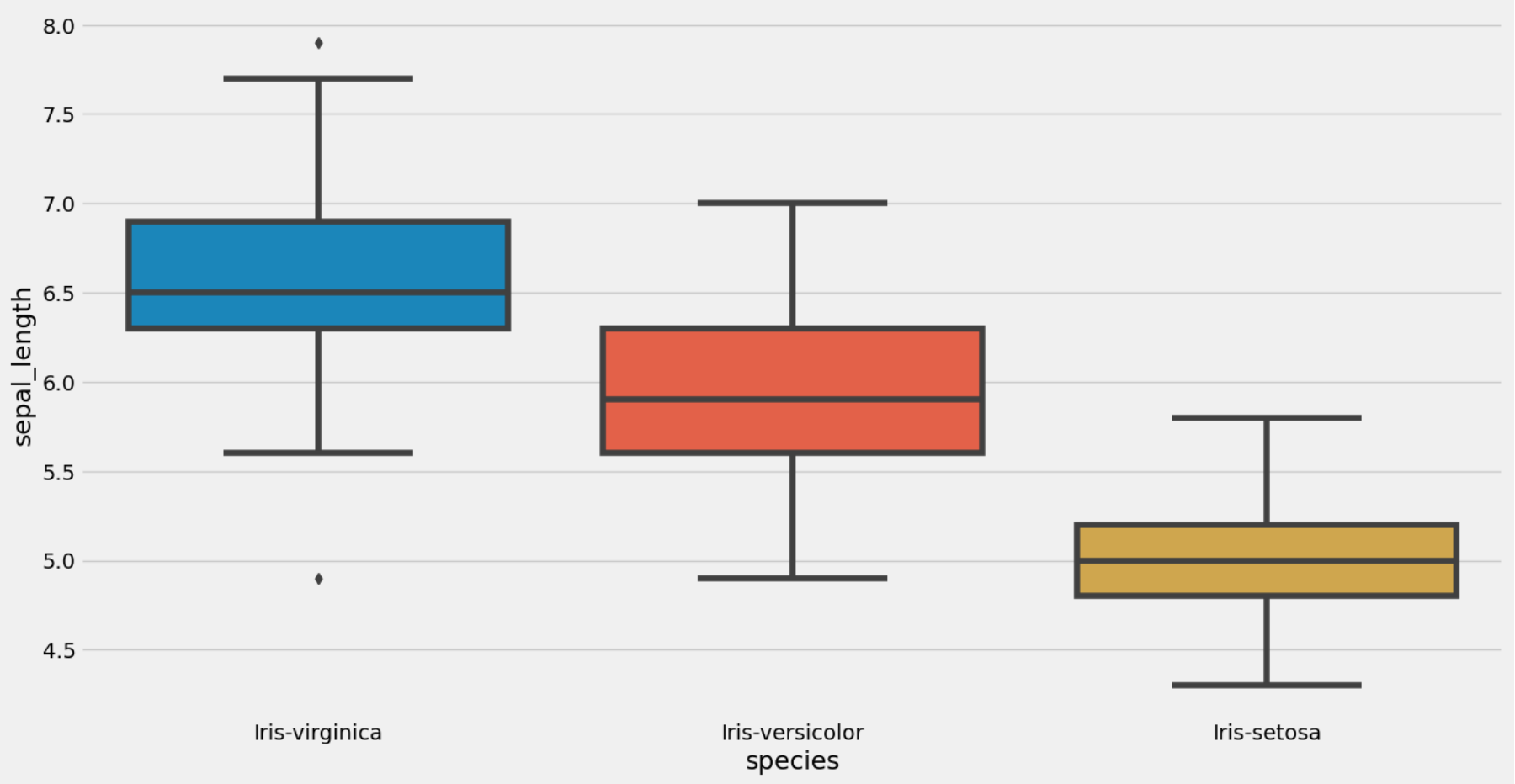
In [ ]:
```

```

In [42]:
```
df.drop_duplicates(inplace=True)
```

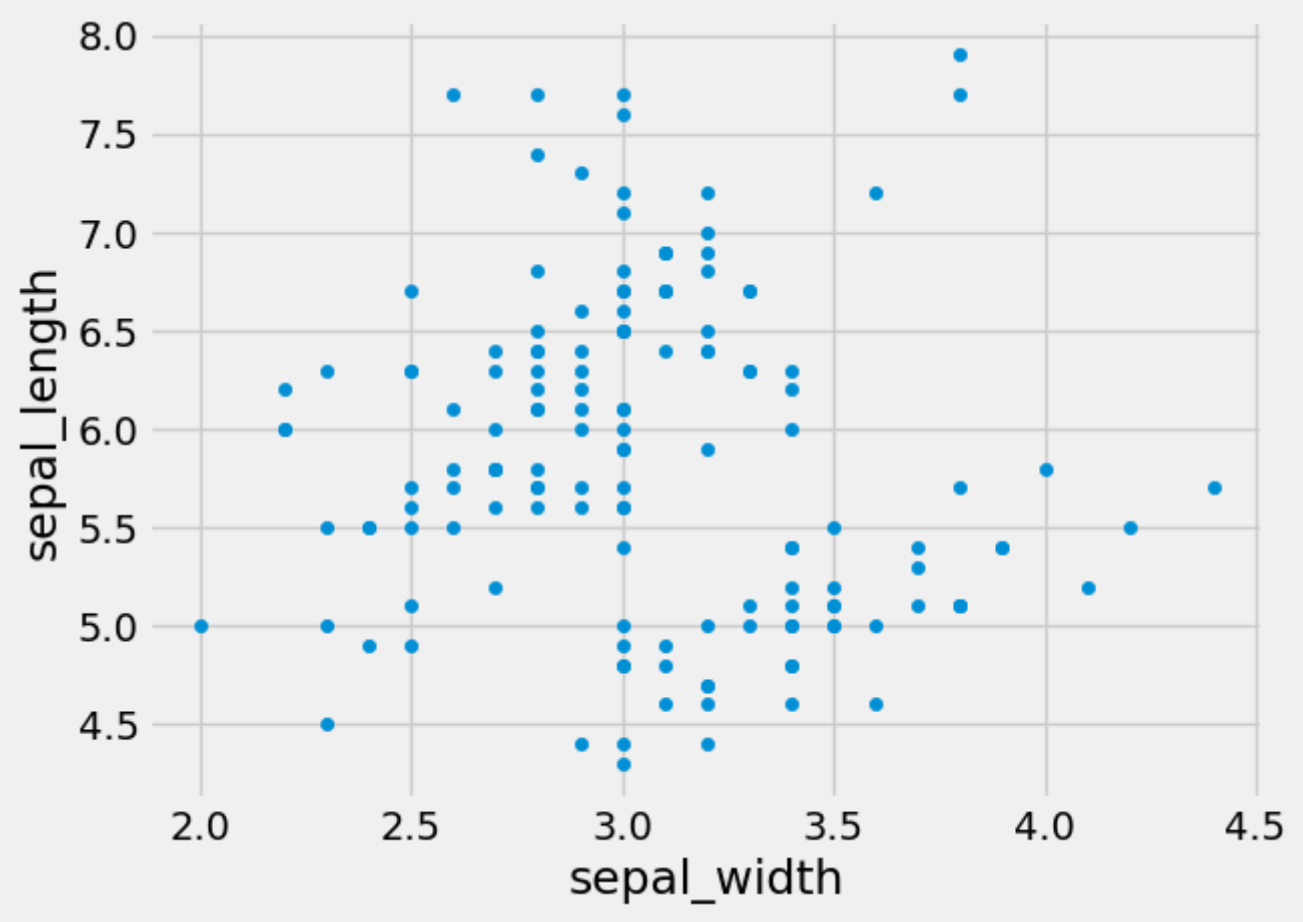
In [43]:
```
plt.figure(figsize=(15,8))
sns.boxplot(x='species',y='sepal_length',data=df.sort_values('sepal_length',ascending=False))
```

Out[43]:
```
<AxesSubplot:xlabel='species', ylabel='sepal_length'>
```



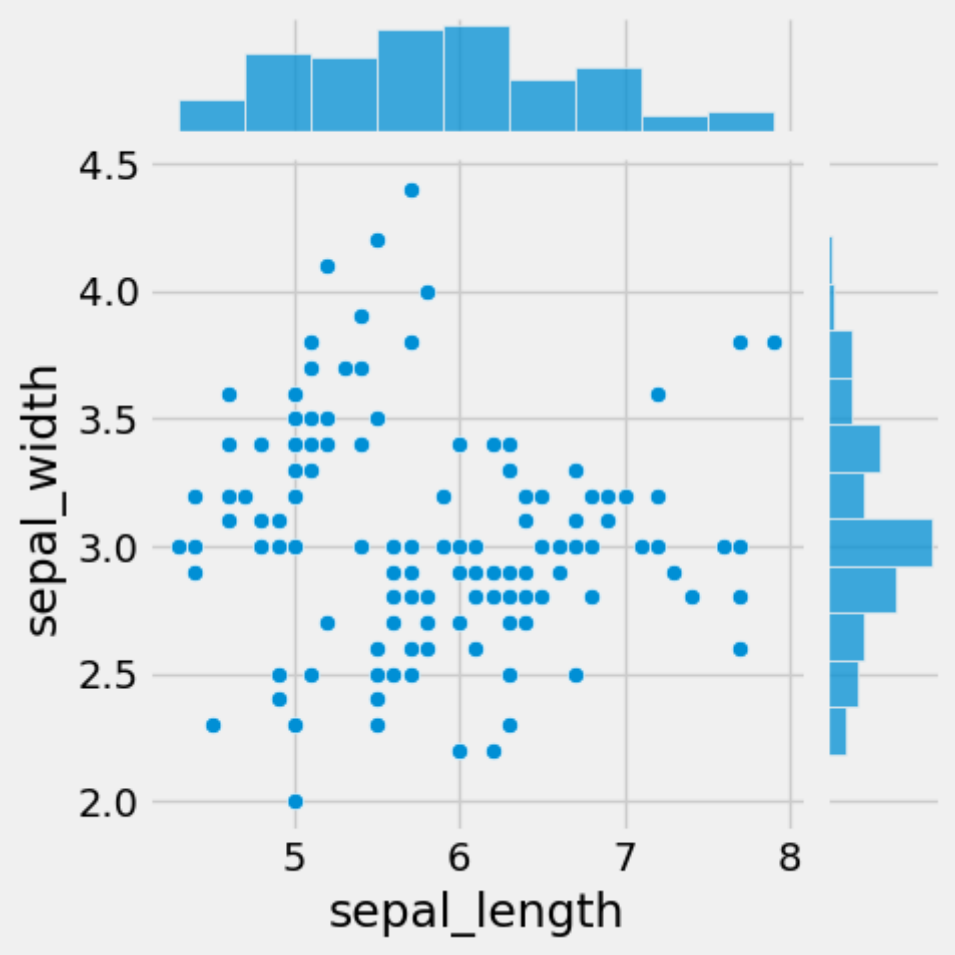
In [44]:
```
df.plot(kind='scatter',x='sepal_width',y='sepal_length')
```

Out[44]:
```
<AxesSubplot:xlabel='sepal_width', ylabel='sepal_length'>
```



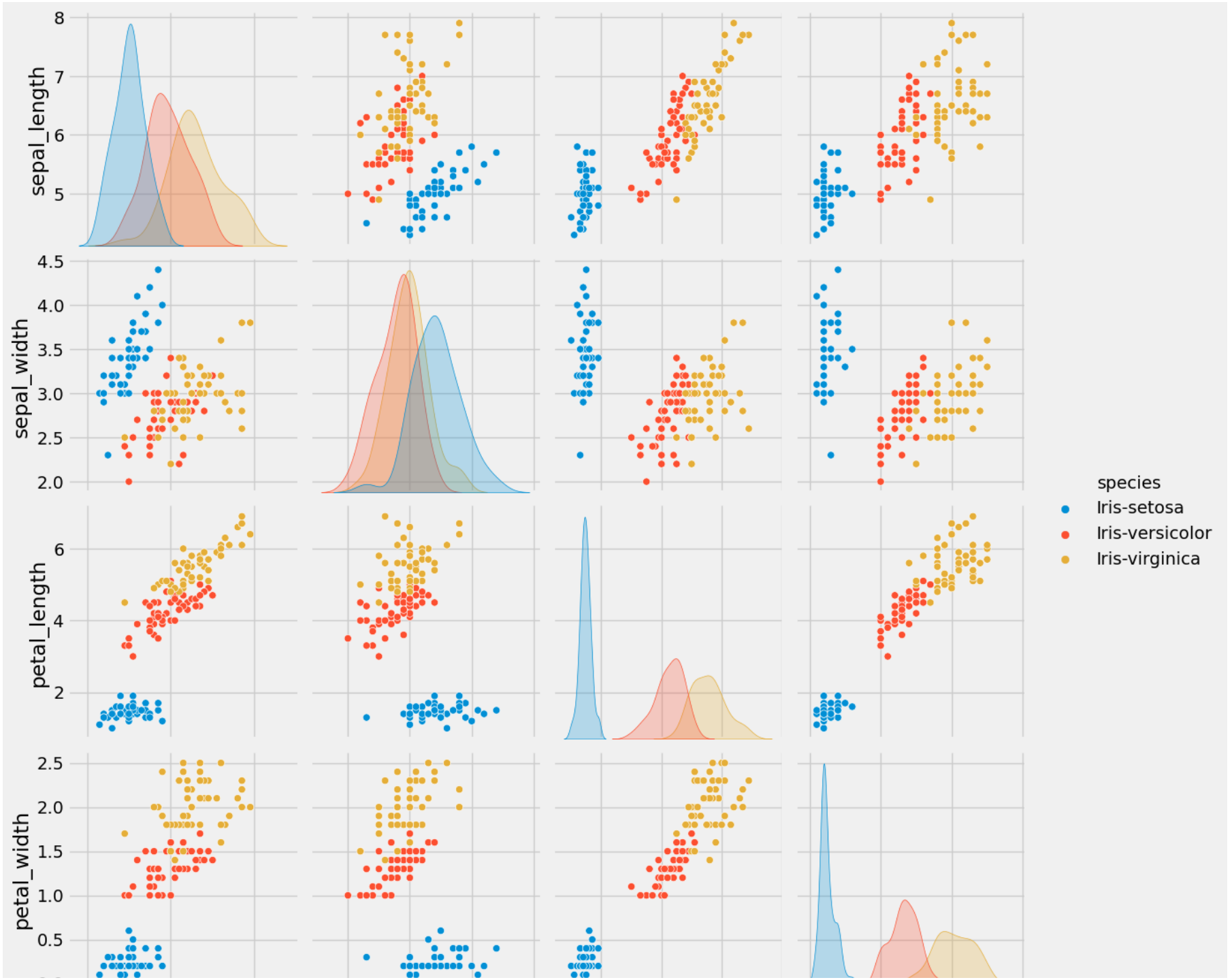
In [45]:
```
sns.jointplot(x="sepal_length", y="sepal_width", data=df, size=5)
```

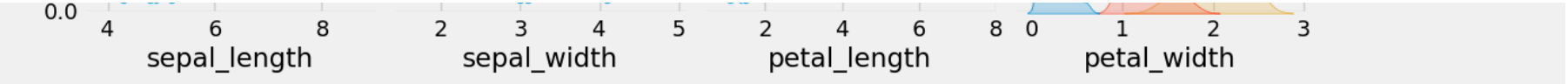
Out[45]:
```
<seaborn.axisgrid.JointGrid at 0x24be8aac6d0>
```



In [46]:
```
sns.pairplot(df, hue="species", size=3)
```

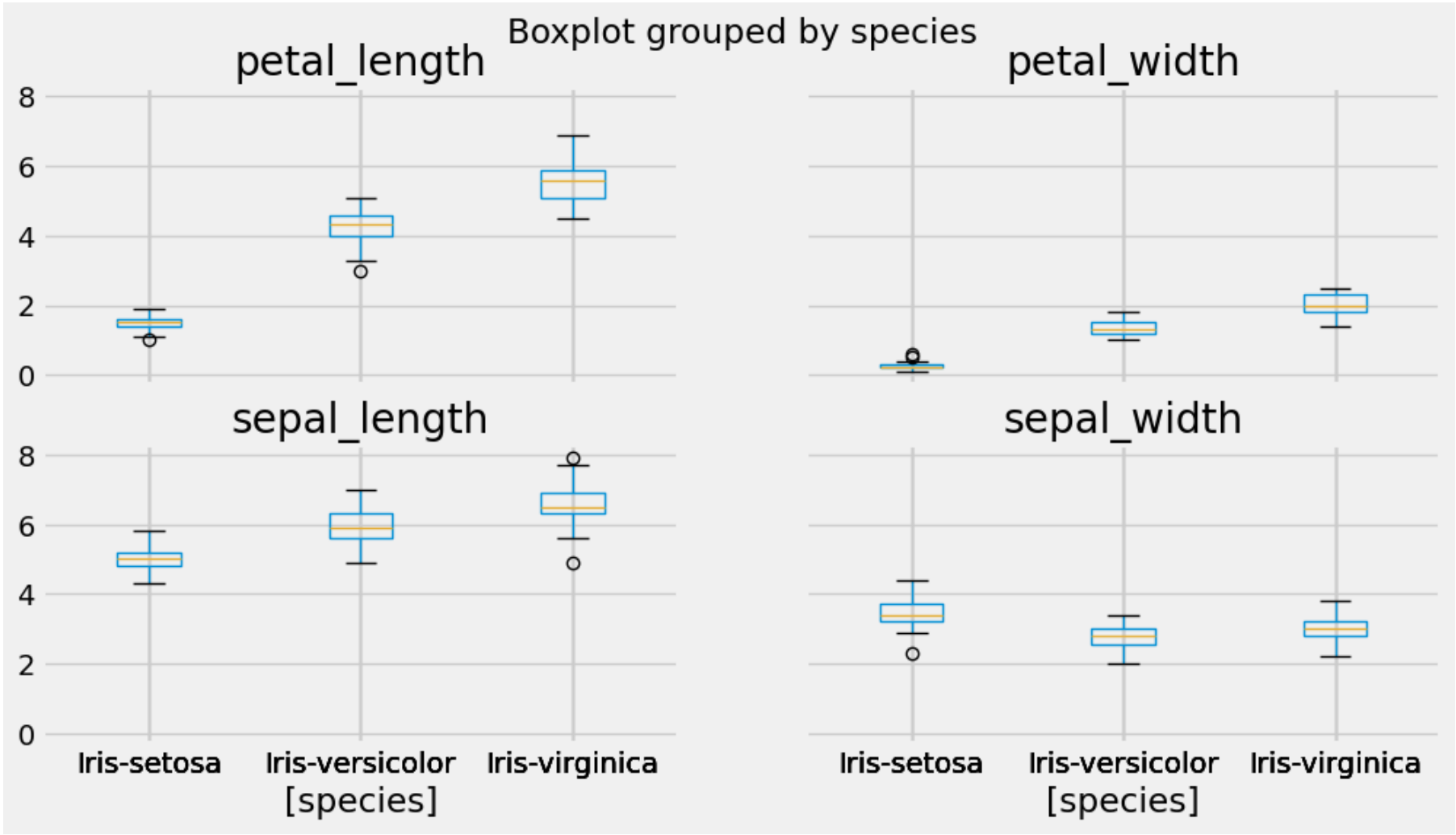
Out[46]:
```
<seaborn.axisgrid.PairGrid at 0x24be8b47e50>
```


sepal_length
sepal_width
petal_length
petal_width
species
Iris-setosa
Iris-versicolor
Iris-virginica


0.0
4
6
8
sepal_length
2
3
4
5
sepal_width
2
4
6
8
petal_length
0
1
2
3
petal_width

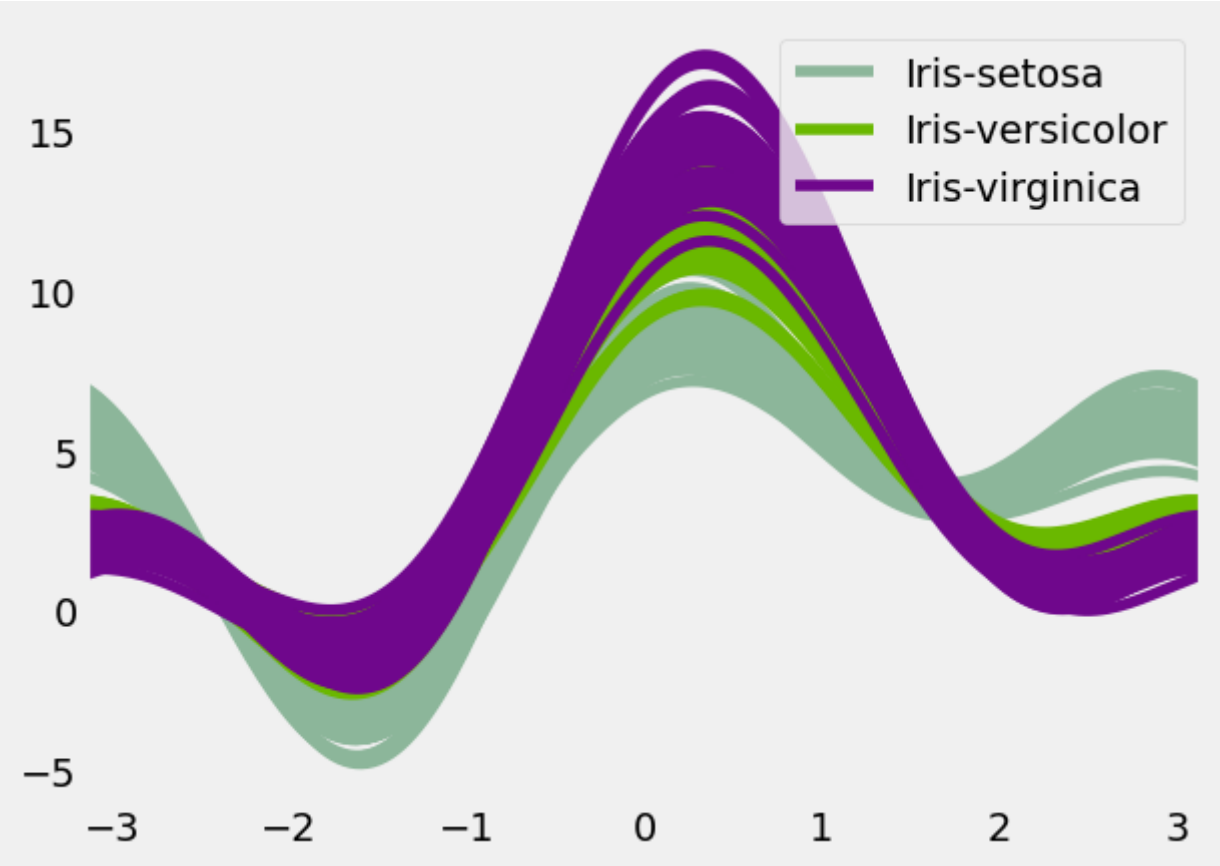
In [47]:
```
df.boxplot(by="species", figsize=(12, 6))
```

Out[47]:
```
array([[<AxesSubplot:title={'center':'petal_length'}, xlabel='[species]'>,
        <AxesSubplot:title={'center':'petal_width'}, xlabel='[species]'>],
       [<AxesSubplot:title={'center':'sepal_length'}, xlabel='[species]'>,
        <AxesSubplot:title={'center':'sepal_width'}, xlabel='[species]'>]],
      dtype=object)
```



In [48]:
```
import pandas.plotting
from pandas.plotting import andrews_curves
andrews_curves(df, "species")
```

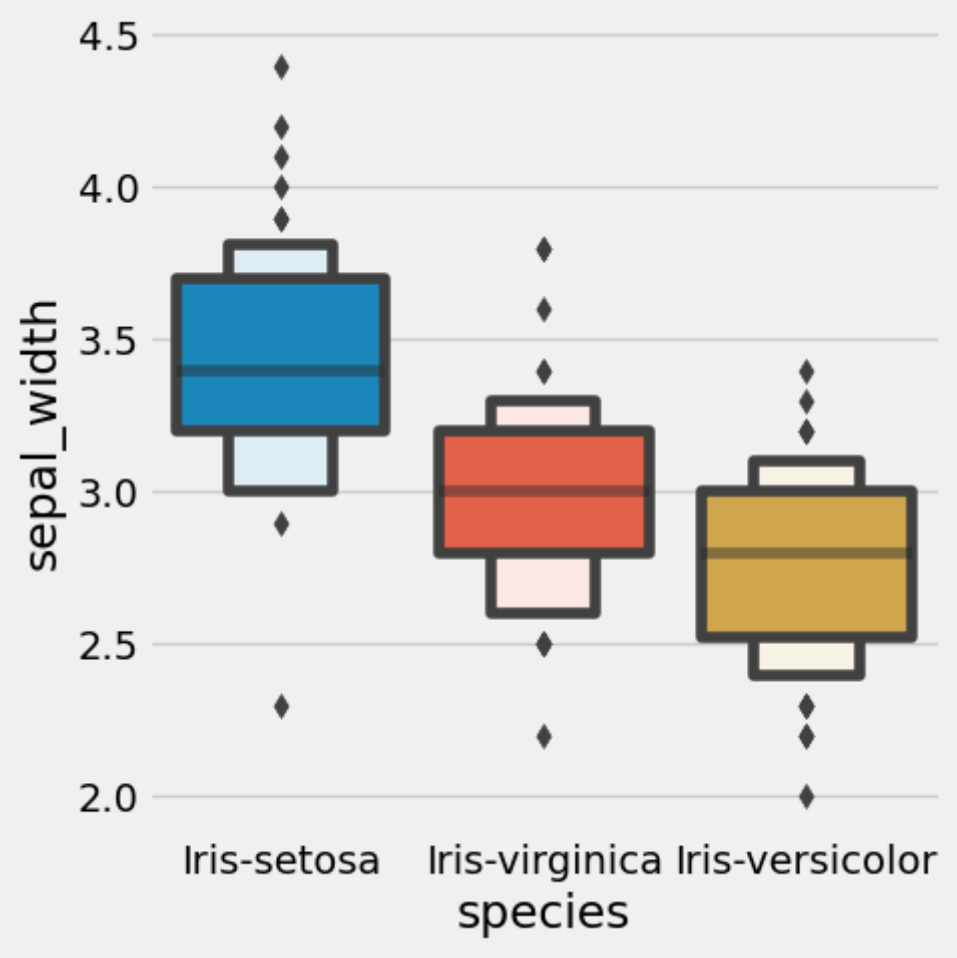
Out[48]: <AxesSubplot:>



In [49]:
```
plt.figure(figsize=(15,15))
sns.catplot(x='species',y='sepal_width',data=df.sort_values('sepal_width',ascending=False),kind='boxen')
```

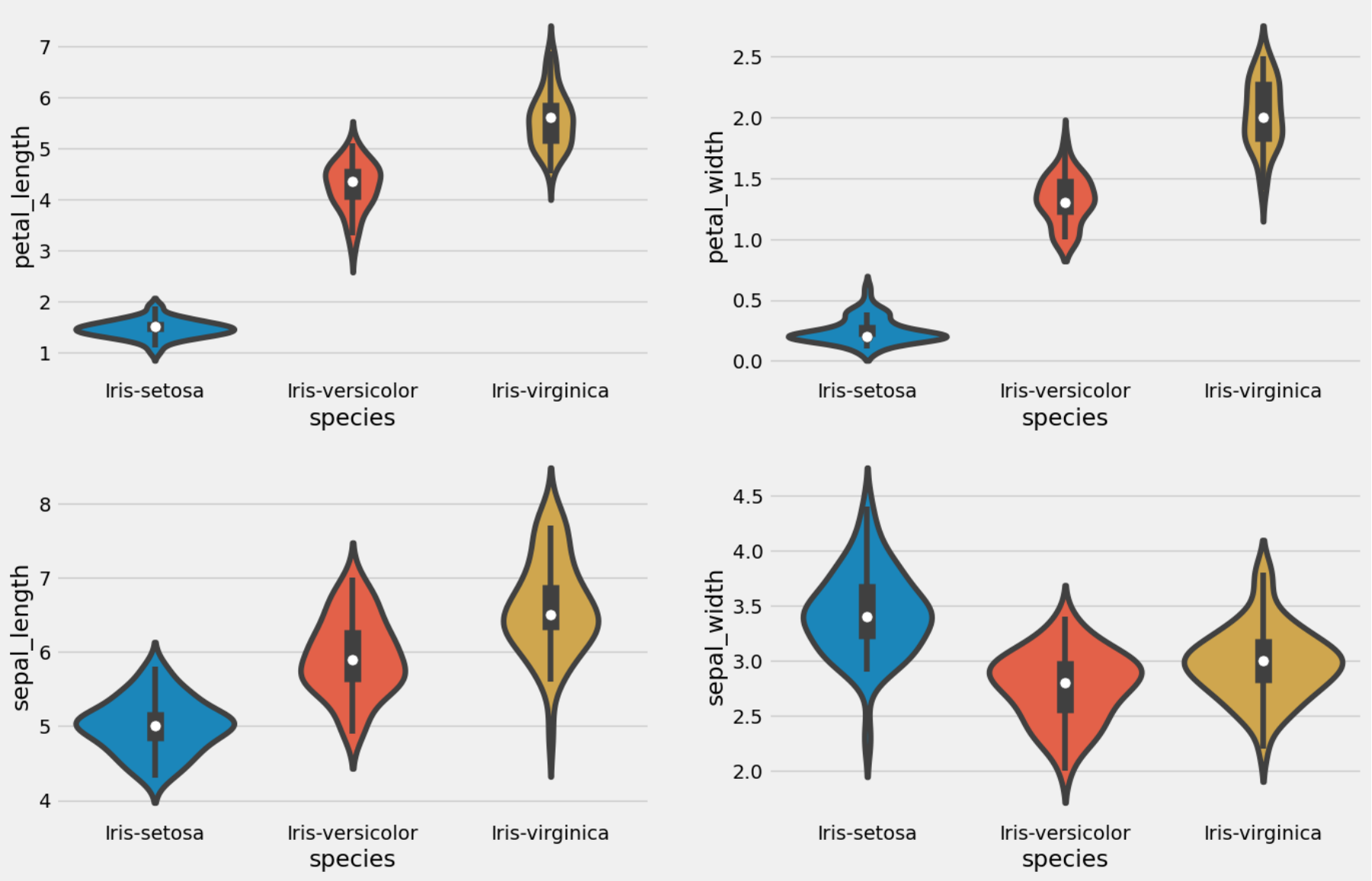
Out[49]:
```
<seaborn.axisgrid.FacetGrid at 0x24beb6239a0>

<Figure size 1500x1500 with 0 Axes>
```



In [50]:
```
plt.figure(figsize=(15,10))
plt.subplot(2,2,1)
sns.violinplot(x='species',y='petal_length',data=df)
plt.subplot(2,2,2)
sns.violinplot(x='species',y='petal_width',data=df)
plt.subplot(2,2,3)
sns.violinplot(x='species',y='sepal_length',data=df)
plt.subplot(2,2,4)
sns.violinplot(x='species',y='sepal_width',data=df)
```

Out[50]:
```
<AxesSubplot:xlabel='species', ylabel='sepal_width'>
```


petal_length
7
6
5
4
3
2
1
Iris-setosa
Iris-versicolor
Iris-virginica
species
petal_width
2.5
2.0
1.5
1.0
0.5
0.0
Iris-setosa
Iris-versicolor
Iris-virginica
species
sepal_length
8
7
6
5
4
Iris-setosa
Iris-versicolor
Iris-virginica
species
sepal_width
4.5
4.0
3.5
3.0
2.5
2.0
Iris-setosa
Iris-versicolor
Iris-virginica
species

In [31]:

In [ ]:

In [ ]: