IMPORT LIBRARIES

```
In [1]:  import pandas as pd
         import seaborn as sns
         import matplotlib.pyplot as plt
         import plotly.express as px
```

LOAD DATASET

```
In [2]:  df= sns.load_dataset('iris')
```

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

Out[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

```
In [4]:  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
```

SUMMARY STATISTICS

```
In [5]:  df.describe()
```

Out[5]:

	sepal_length	sepal_width	petal_length	petal_width
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.057333	3.758000	1.199333
std	0.828066	0.435866	1.765298	0.762238
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 [6]:
```python
df['species'].value_counts()
```
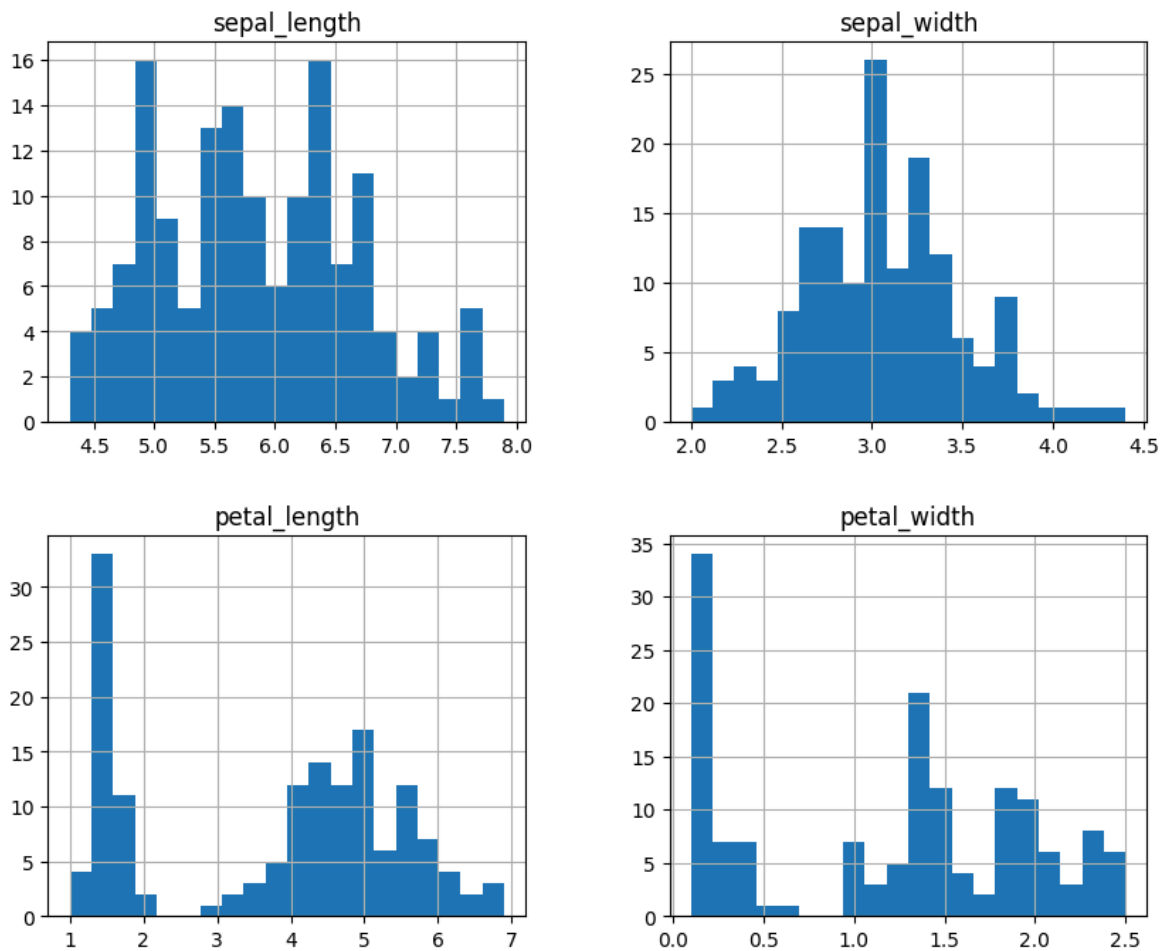
Out[6]:
```
species
setosa        50
versicolor    50
virginica     50
Name: count, dtype: int64
```

BOXPLOTS AND HISTOGRAMS

In [8]:
```python
df.hist(figsize=(10,8),bins=20)
plt.suptitle('Histogram of numeric features')
plt.show()
```

Histogram of numeric features

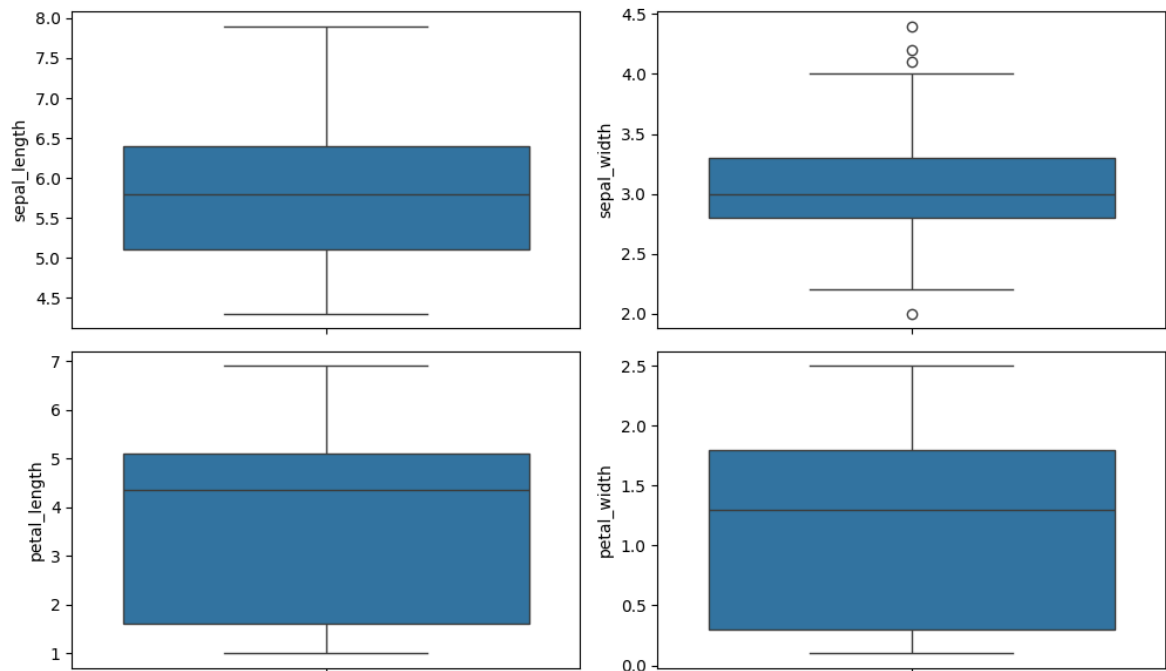```
In [9]: plt.figure(figsize=(10,6))
         for i, col in enumerate(df.columns[:-1]):
             plt.subplot(2,2,i+1)
             sns.boxplot(data=df,y=col)
         plt.tight_layout()
         plt.show()
```

C:\Users\Gnanesh\AppData\Local\Programs\Python\Python312\Lib\site-packages\seabor
n\categorical.py:640: FutureWarning: SeriesGroupBy.grouper is deprecated and will
be removed in a future version of pandas.
  positions = grouped.grouper.result_index.to_numpy(dtype=float)
C:\Users\Gnanesh\AppData\Local\Programs\Python\Python312\Lib\site-packages\seabor
n\categorical.py:640: FutureWarning: SeriesGroupBy.grouper is deprecated and will
be removed in a future version of pandas.
  positions = grouped.grouper.result_index.to_numpy(dtype=float)
C:\Users\Gnanesh\AppData\Local\Programs\Python\Python312\Lib\site-packages\seabor
n\categorical.py:640: FutureWarning: SeriesGroupBy.grouper is deprecated and will
be removed in a future version of pandas.
  positions = grouped.grouper.result_index.to_numpy(dtype=float)
C:\Users\Gnanesh\AppData\Local\Programs\Python\Python312\Lib\site-packages\seabor
n\categorical.py:640: FutureWarning: SeriesGroupBy.grouper is deprecated and will
be removed in a future version of pandas.
  positions = grouped.grouper.result_index.to_numpy(dtype=float)

CORRELATION MATRIX FOR FEATURE RELATIONSHIP

In [12]:
```python
plt.figure(figsize=(8,6))
numeric_df=df.select_dtypes(include='number')
sns.heatmap(numeric_df.corr(), annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()
```
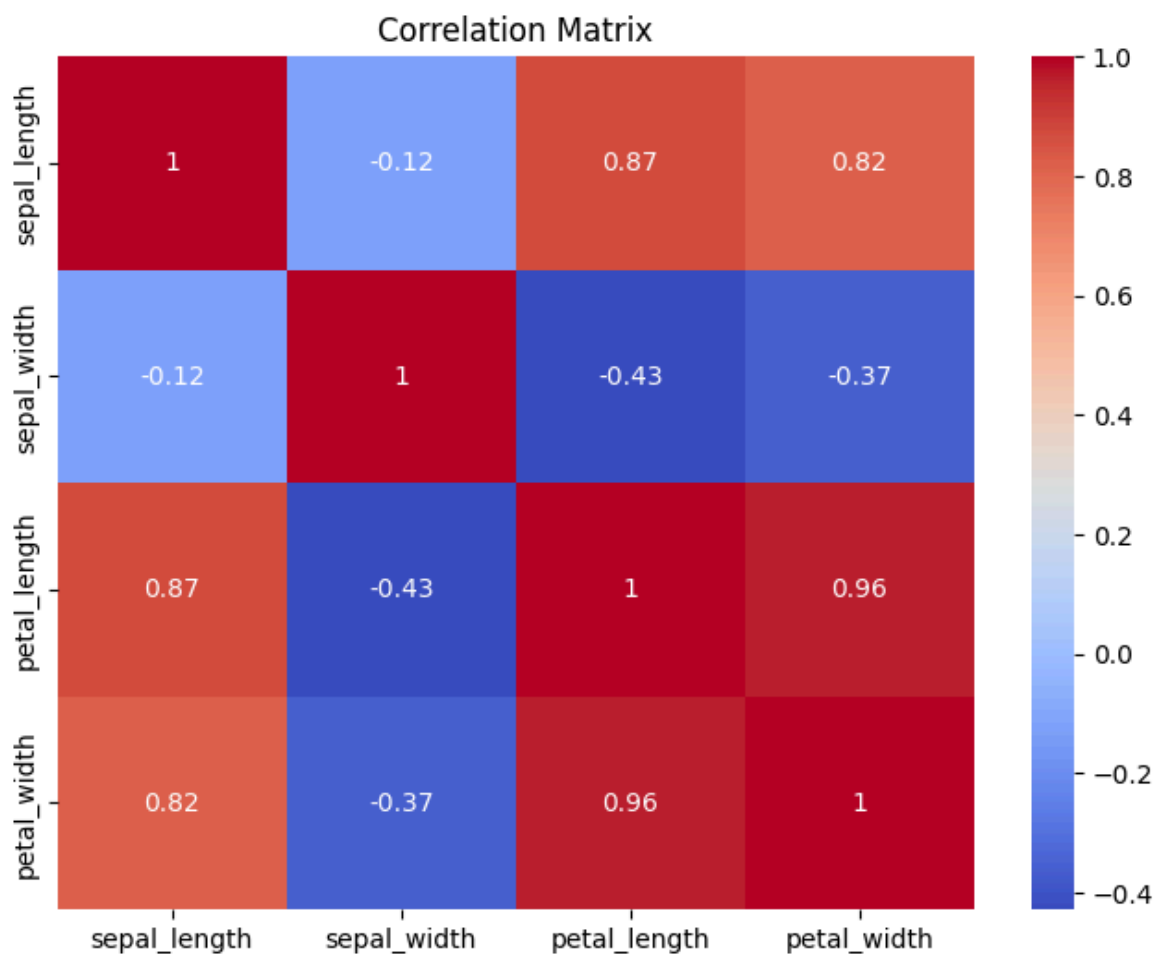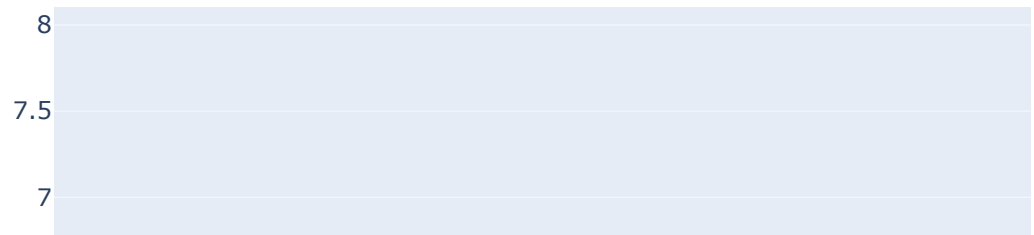

INTERACTIVE PLOTLY VISUALIZATIONS

In [14]:
```python
fig=px.scatter_matrix(df, dimensions=df.columns[:-1],color='species', title="Int
fig.show()
```

Interactive Scatter Matrix
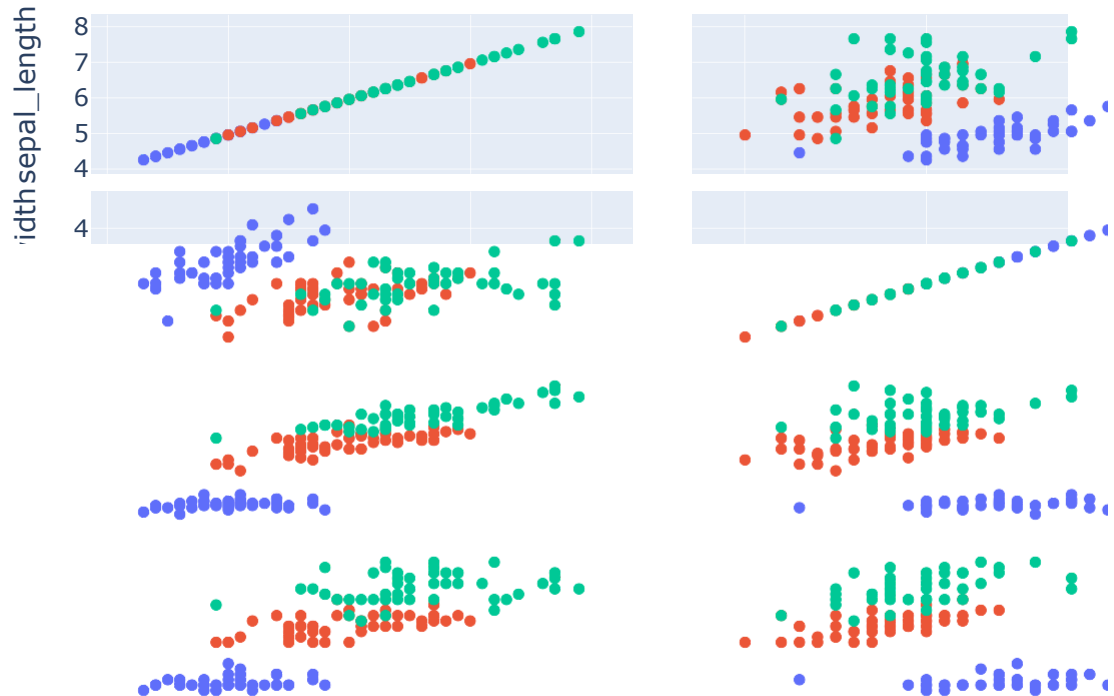
In [15]:
```python
fig2= px.box(df, x='species', y='sepal_length', title='Sepal lengthby species')
fig2.show()
```

Sepal lengthby species

INFERENCES FROM THE VISUALS

- There are no null values in the dataset as it is clean.
- Boxplots show a few mild outliers as Sepal length/width are moderately correlated whereas petal length/width are highly correlated.
- Setosa species is linearly seperable from the other two species.
- Versicolor and Virginica overlap more.
- Petal length and width are strong discriminators for species classification.