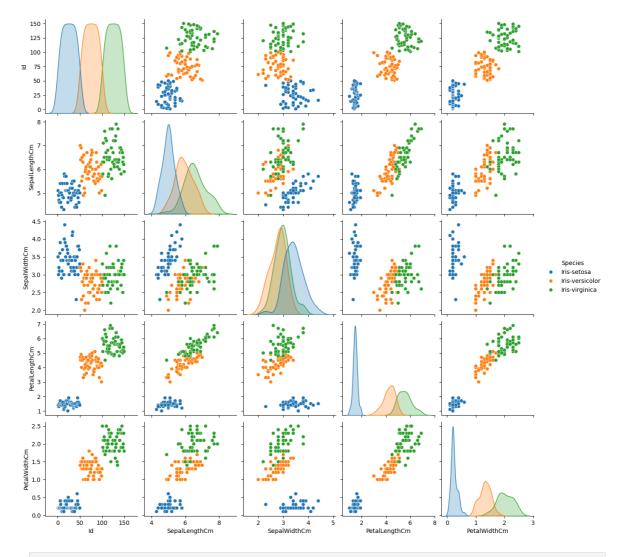
19/04/2025, 16:58 eda iris

```
In [1]:
        import pandas as pd
        import seaborn as sns
        import matplotlib.pyplot as plt
In [2]: df = pd.read_csv(r"C:\Users\Shreyas\OneDrive\Desktop\Dataset\Iris.csv")
In [3]: print(df.info())
       <class 'pandas.core.frame.DataFrame'>
       RangeIndex: 150 entries, 0 to 149
       Data columns (total 6 columns):
            Column
                           Non-Null Count Dtype
            -----
                           -----
        0
            Ιd
                                           int64
                           150 non-null
            SepalLengthCm 150 non-null
        1
                                           float64
        2
           SepalWidthCm
                           150 non-null
                                           float64
        3
            PetalLengthCm 150 non-null
                                           float64
            PetalWidthCm
                           150 non-null
                                           float64
        5
            Species
                           150 non-null
                                           object
       dtypes: float64(4), int64(1), object(1)
       memory usage: 7.2+ KB
       None
In [4]:
        print(df.describe())
                         SepalLengthCm
                                        SepalWidthCm
                                                       PetalLengthCm
                                                                      PetalWidthCm
                      Ιd
                             150.000000
       count 150.000000
                                           150.000000
                                                          150.000000
                                                                         150.000000
       mean
               75.500000
                               5.843333
                                             3.054000
                                                             3.758667
                                                                           1.198667
       std
               43.445368
                               0.828066
                                             0.433594
                                                            1.764420
                                                                           0.763161
       min
                1.000000
                               4.300000
                                             2.000000
                                                             1.000000
                                                                           0.100000
       25%
               38.250000
                                                                           0.300000
                               5.100000
                                             2.800000
                                                            1.600000
       50%
               75.500000
                               5.800000
                                             3.000000
                                                            4.350000
                                                                           1.300000
       75%
              112.750000
                               6.400000
                                             3.300000
                                                             5.100000
                                                                           1.800000
       max
              150.000000
                               7.900000
                                             4.400000
                                                             6.900000
                                                                           2.500000
In [5]:
        print(df.isnull().sum())
       Ιd
                        0
       SepalLengthCm
       SepalWidthCm
                        0
       PetalLengthCm
                        0
       PetalWidthCm
                        0
       Species
                        0
       dtype: int64
In [6]: sns.pairplot(df, hue='Species')
Out[6]: <seaborn.axisgrid.PairGrid at 0x218e79ad490>
```

19/04/2025, 16:58 eda_iris

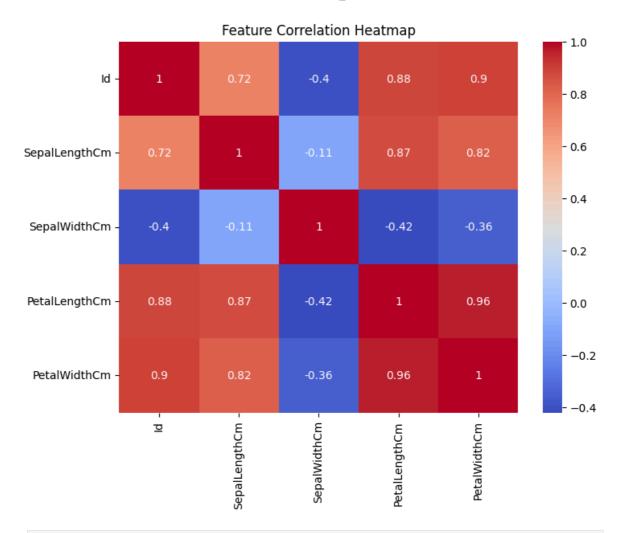


In [7]: plt.savefig("pairplot.png")

<Figure size 640x480 with 0 Axes>

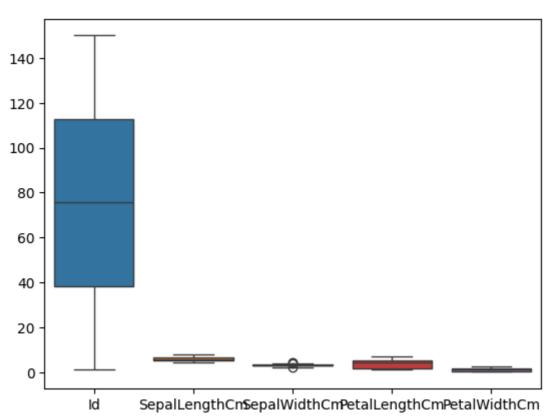
```
In [8]: numeric_df = df.select_dtypes(include=['float64', 'int64'])
In [9]: plt.figure(figsize=(8, 6))
    sns.heatmap(numeric_df.corr(), annot=True, cmap="coolwarm")
    plt.title("Feature Correlation Heatmap")
    plt.savefig("correlation_heatmap.png")
    plt.show()
```

19/04/2025, 16:58 eda_iris



In [17]: sns.boxplot(data = df)

Out[17]: <Axes: >



19/04/2025, 16:58 eda_iris

In []