Iris Dataset Analysis

Python Program

```
import pandas as pd
import numpy as np
from sklearn.datasets import load_iris

# Load dataset
iris = load_iris()
df = pd.DataFrame(data=iris.data, columns=iris.feature_names)
df['species'] = pd.Categorical.from_codes(iris.target, iris.target_names)

# Rename columns
df.columns = ['sepal_length', 'sepal_width', 'petal_length', 'petal_width', 'species']

# Number of rows and columns
num_rows, num_cols = df.shape

# Average petal length
avg_petal_length = df['petal_length'].mean()

# Average of all numerical columns
avg_all_numerical = df.select_dtypes(include=np.number).mean()

# Global petal length outliers
outliers_global = df[df['petal_length'] > 1.5 * avg_petal_length]

# Standard deviation by species
std_by_species = df.groupby('species').std(numeric_only=True)

# Petal length outliers by species
def get_species_outliers(group):
    mean_petal_length = group['petal_length'].mean()
    return group[group['petal_length'] > 1.5 * mean_petal_length]

species_outliers = df.groupby('species').apply(get_species_outliers).reset_index(drop=True)

# Group-wise outliers using merge
avg_by_species = df.groupby('species')['petal_length'].mean().reset_index()
avg_by_species['threshold'] = avg_by_species['petal_length'] * 1.5
```

```
df_merged = df.merge(avg_by_species, on='species')
groupwise_outliers = df_merged[df_merged['petal_length_x'] > df_merged['threshold']]
```

Output Summary

1. Number of rows: 150, Number of columns: 5

2. Average petal length: 3.7580

3. Average of all numerical columns:

```
sepal_length    5.843333
sepal_width     3.057333
petal_length    3.758000
petal_width     1.199333
```

4. Number of global petal length outliers: 19

5. Standard deviation by species:

```
            sepal_length  sepal_width  petal_length  petal_width
species
setosa          0.352490     0.379064      0.173664     0.105386
versicolor      0.516171     0.313798      0.469911     0.197753
virginica       0.635880     0.322497      0.551895     0.274650
```

6. Number of species-wise petal length outliers: 0

7. Number of group-wise petal length outliers (using merge): 0