

# polars

July 19, 2024

```
[1]: !pip install polars
```

```
Collecting polars
  Downloading
polars-0.18.4-cp37-abi3-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata
(14 kB)
Collecting typing_extensions>=4.0.1 (from polars)
  Downloading typing_extensions-4.7.1-py3-none-any.whl.metadata (3.1 kB)
Downloading
polars-0.18.4-cp37-abi3-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (18.9 MB)
18.9/18.9 MB
127.1 MB/s eta 0:00:0000:0100:01
Downloading typing_extensions-4.7.1-py3-none-any.whl (33 kB)
Installing collected packages: typing_extensions, polars
  Attempting uninstall: typing_extensions
    Found existing installation: typing-extensions 3.7.4.2
    Uninstalling typing-extensions-3.7.4.2:
      Successfully uninstalled typing-extensions-3.7.4.2
Successfully installed polars-0.18.4 typing_extensions-4.7.1

[notice] A new release of pip is
available: 23.3.1 -> 24.0
[notice] To update, run:
pip install --upgrade pip
```

```
[7]: import polars as pl

# Create DataFrame
data = [{"fruit": "apple", "count": 10, "price": 0.50},
        {"fruit": "banana", "count": 20, "price": 0.25},
        {"fruit": "orange", "count": 10, "price": 0.75},]
df = pl.from_dicts(data)

# Expressions to select, filter, aggregate
sel = df.select(["fruit", "count"]) # Select columns
filt = sel.filter(pl.col("fruit") == "apple") # Filter rows
agg = filt.groupby("fruit").agg(pl.col("count").sum()) # Aggregate
```

```
print(agg)
```

shape: (1, 2)

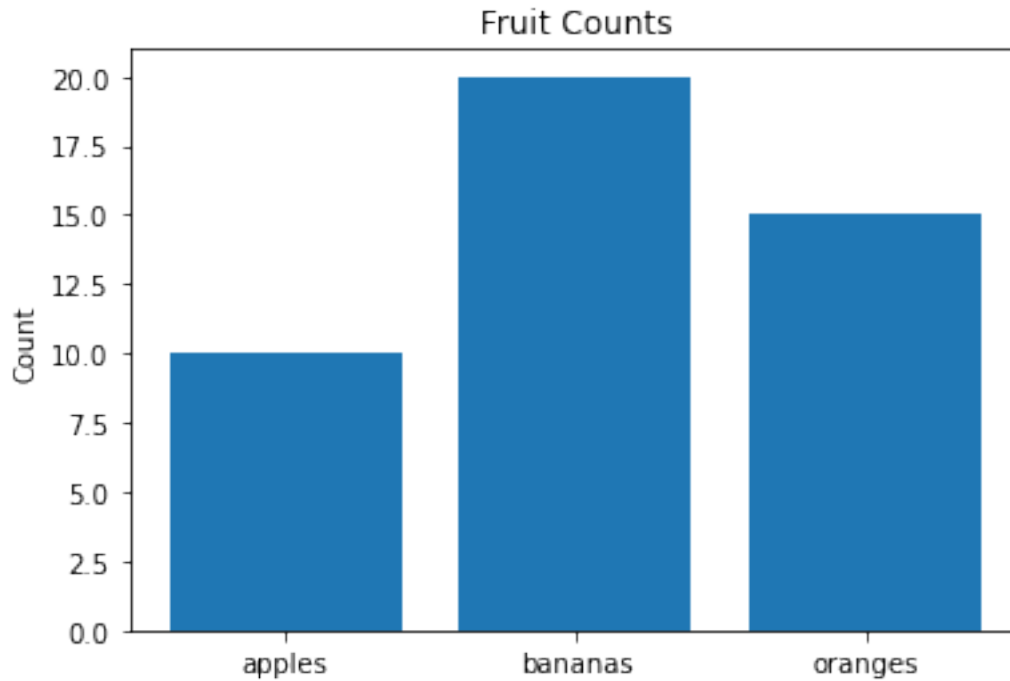
fruit	count
---	---
str	i64
apple	10

```
[3]: import polars as pl
import matplotlib.pyplot as plt

# Create sample data
data = [{"fruit": "apples", "count": 10},
        {"fruit": "bananas", "count": 20},
        {"fruit": "oranges", "count": 15}]
df = pl.from_dicts(data)

# Extract columns
fruits = df['fruit']
counts = df['count']

# Plot bar chart
fig, ax = plt.subplots()
ax.bar(fruits, counts)
ax.set_ylabel('Count')
ax.set_title('Fruit Counts')
plt.show()
```



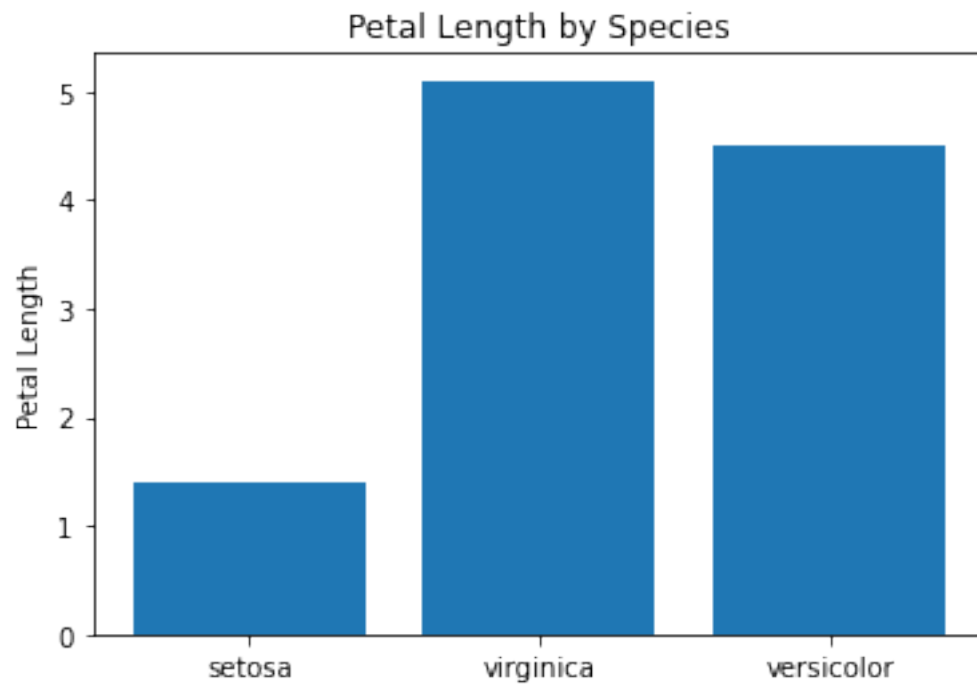
```
[3]: import polars as pl
import matplotlib.pyplot as plt

data = [{"species": "setosa", "petal_length": 1.4},
        {"species": "virginica", "petal_length": 5.1},
        {"species": "versicolor", "petal_length": 4.5}]

df = pl.from_dicts(data)

# Extract columns
species = df['species']
lengths = df['petal_length']

# Plot bar chart
fig, ax = plt.subplots()
ax.bar(species, lengths)
ax.set_ylabel('Petal Length')
ax.set_title("Petal Length by Species")
plt.show()
```



[ ]: