GroupBy Workflow in pandas

1 **Select Column to Group By**

```
grouped = df.groupby('Sector')
```

Creates GroupBy object (lazy evaluation)

2 **Select Column to Aggregate**

```
grouped['Return']
```

Choose which column(s) to operate on

3 **Apply Aggregation Function**

```
.mean()  # or sum(), count(), std()...
```

Actually performs the computation

4 **Get Results**

```
# Returns Series or DataFrame
```

Indexed by group keys

Complete Example

```
# Group by sector and calculate

# mean return for each


result = df.groupby('Sector')\
            ['Return']\
            .mean()


# Output:

# Sector

# Finance    0.052

# Tech       0.078

# Energy     0.034
```

Common Aggregation Functions:

```
mean()  |  sum()  |  count()  |  std()  |  min()  |  max()  |  first()  |  last()
```