pandas

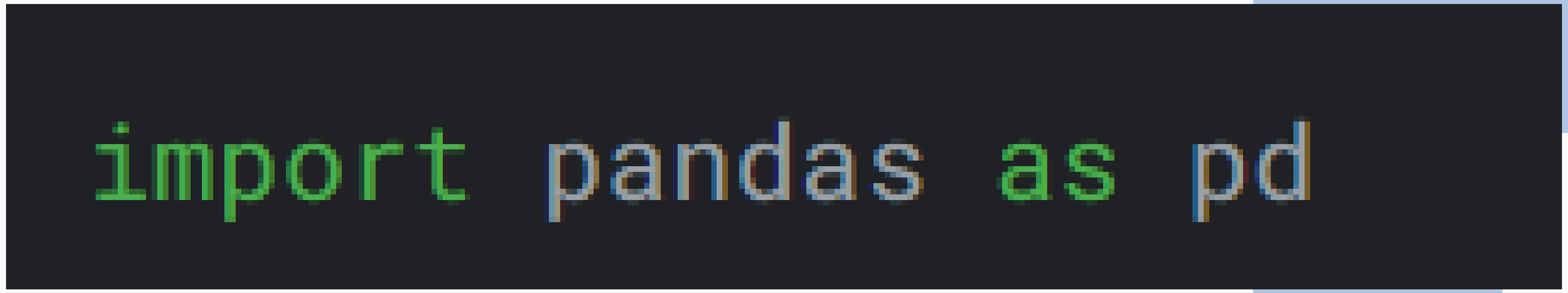
Creating, Reading and Writing

1

By Abdul Rauf Jatoi

Intro

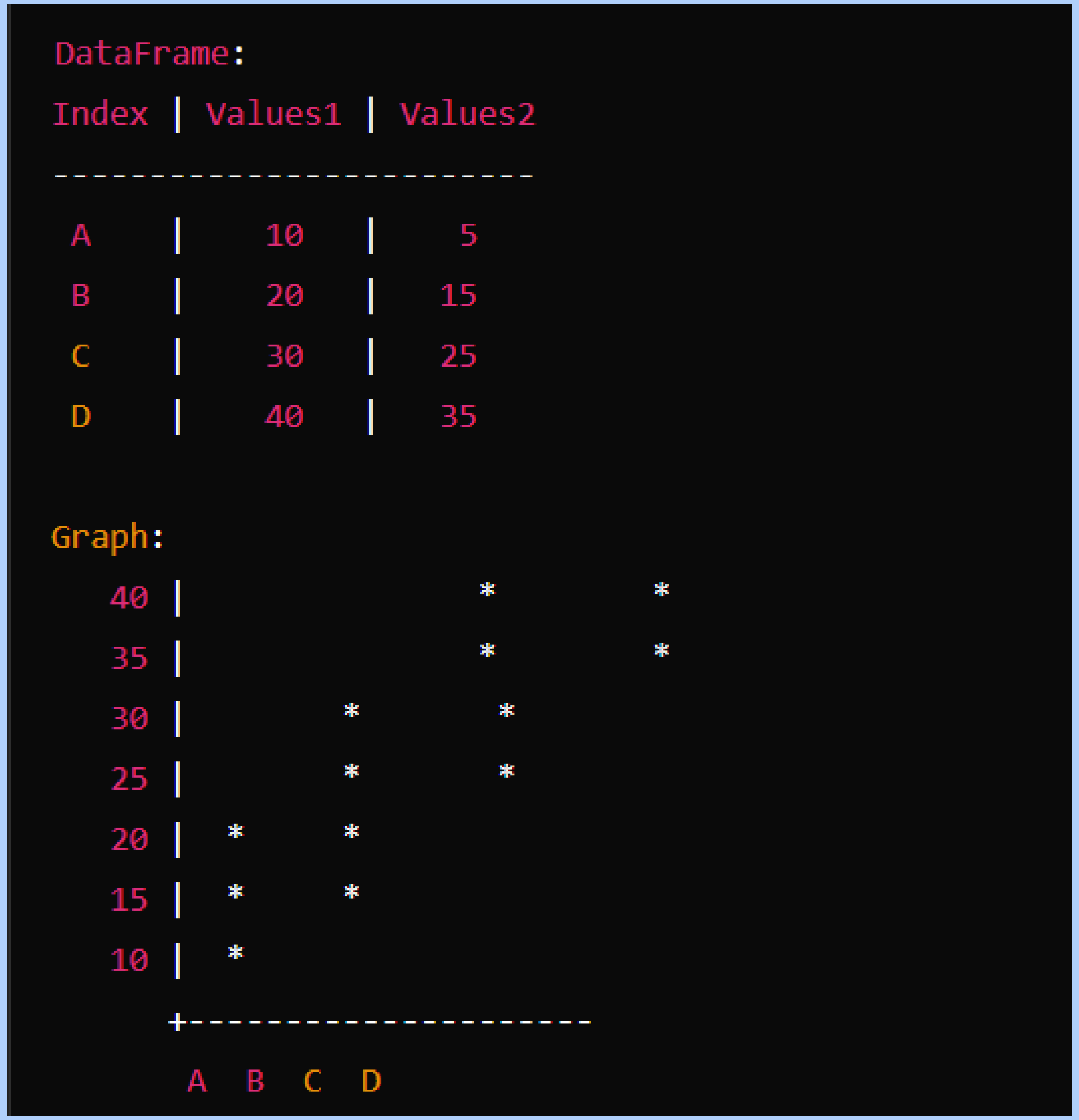
Pandas is a Python library for data manipulation, offering tools to handle, clean, and analyze structured data efficiently.



```python
import pandas as pd
```

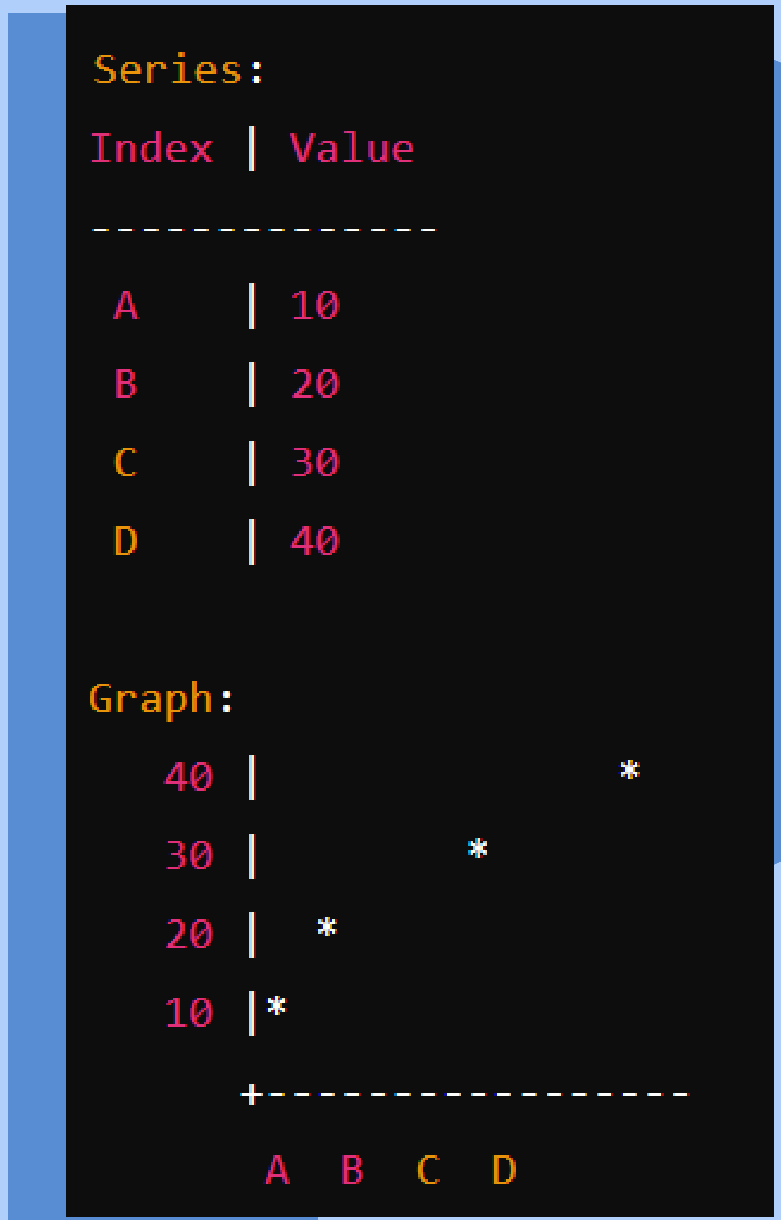
Creating

1. **DataFrame : 2D table with labeled rows and columns in pandas**
2. **Series : 1D labeled array, holds data with an index in pandas.**



```
DataFrame:
Index | Values1 | Values2
-------------------------
 A    |    10   |    5
 B    |    20   |   15
 C    |    30   |   25
 D    |    40   |   35

Graph:
   40 |               *        *
   35 |               *        *
   30 |        *       *
   25 |        *       *
   20 |  *     *
   15 |  *     *
   10 |  *
      +---------------------
       A  B  C  D
```

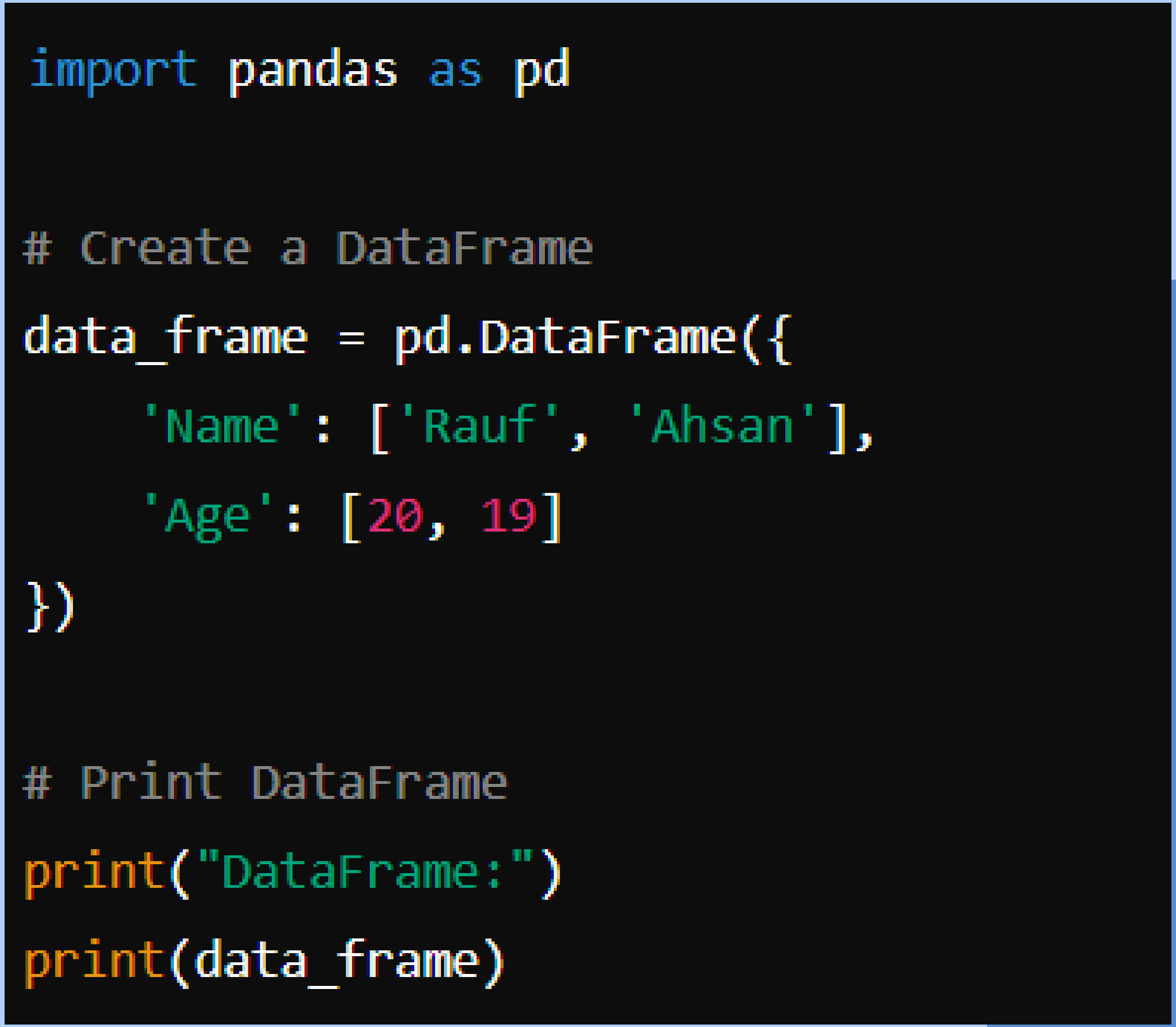


```
Series:
Index | Value
--------------
 A    | 10
 B    | 20
 C    | 30
 D    | 40

Graph:
   40 |               *
   30 |         *
   20 |   *
   10 |*
      +------------------
        A  B  C  D
```

DataFrame

code

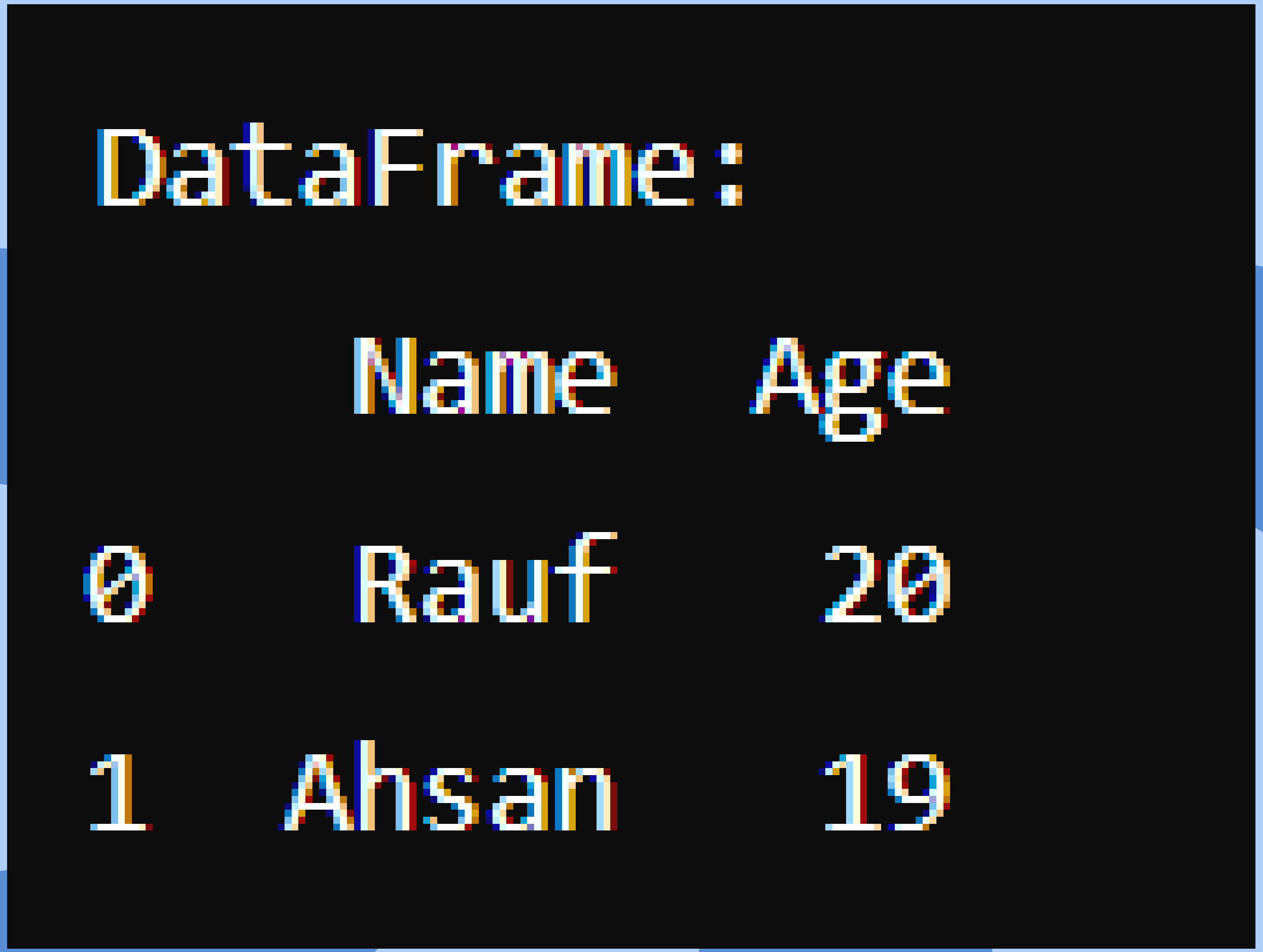


```
import pandas as pd


# Create a DataFrame
data_frame = pd.DataFrame({
    'Name': ['Rauf', 'Ahsan'],
    'Age': [20, 19]
})


# Print DataFrame
print("DataFrame:")
print(data_frame)
```

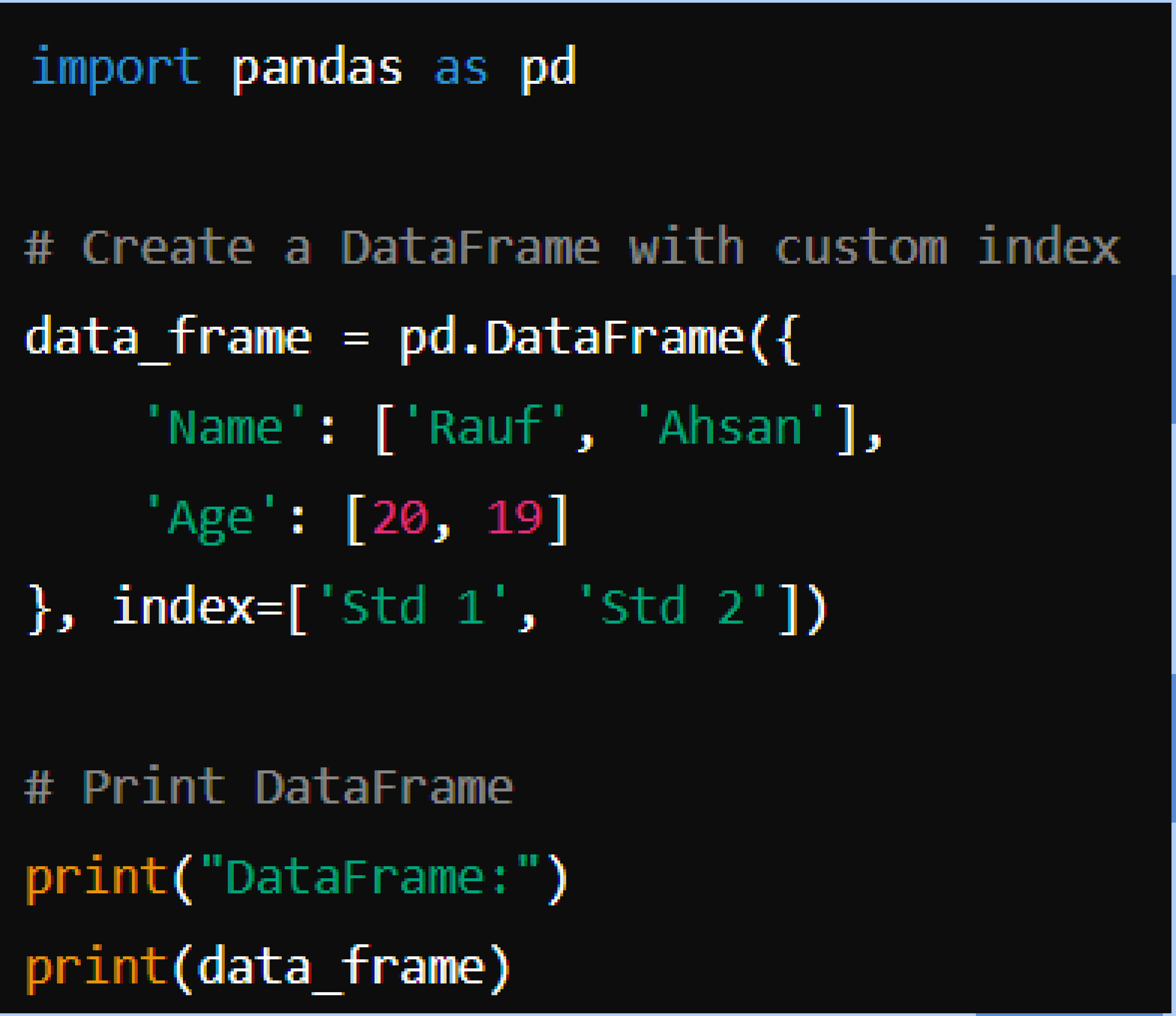
output



```
DataFrame:
    Name  Age
0   Rauf   20
1  Ahsan   19
```

DataFrame

code

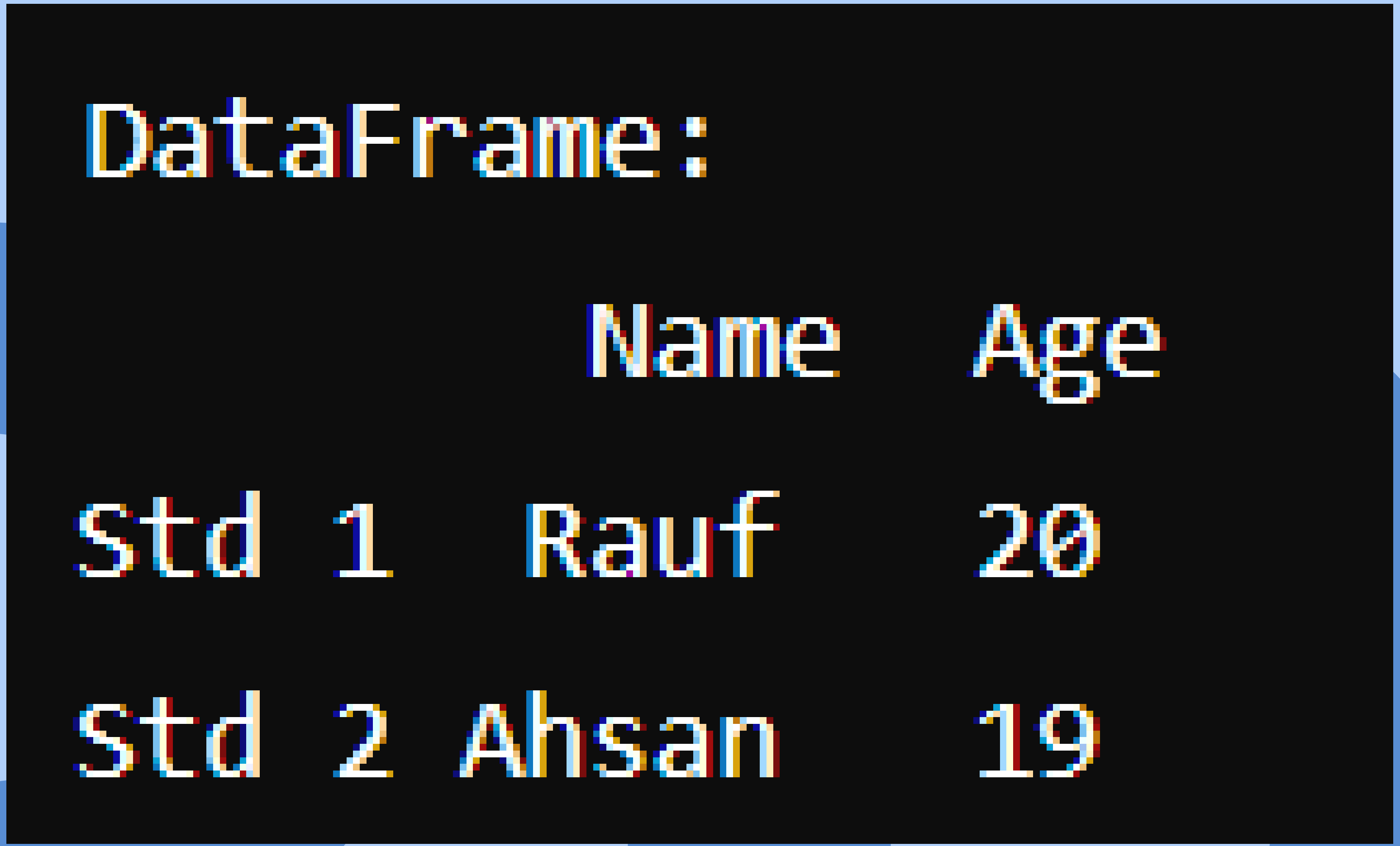


```python
import pandas as pd

# Create a DataFrame with custom index
data_frame = pd.DataFrame({
    'Name': ['Rauf', 'Ahsan'],
    'Age': [20, 19]
}, index=['Std 1', 'Std 2'])

# Print DataFrame
print("DataFrame:")
print(data_frame)
```

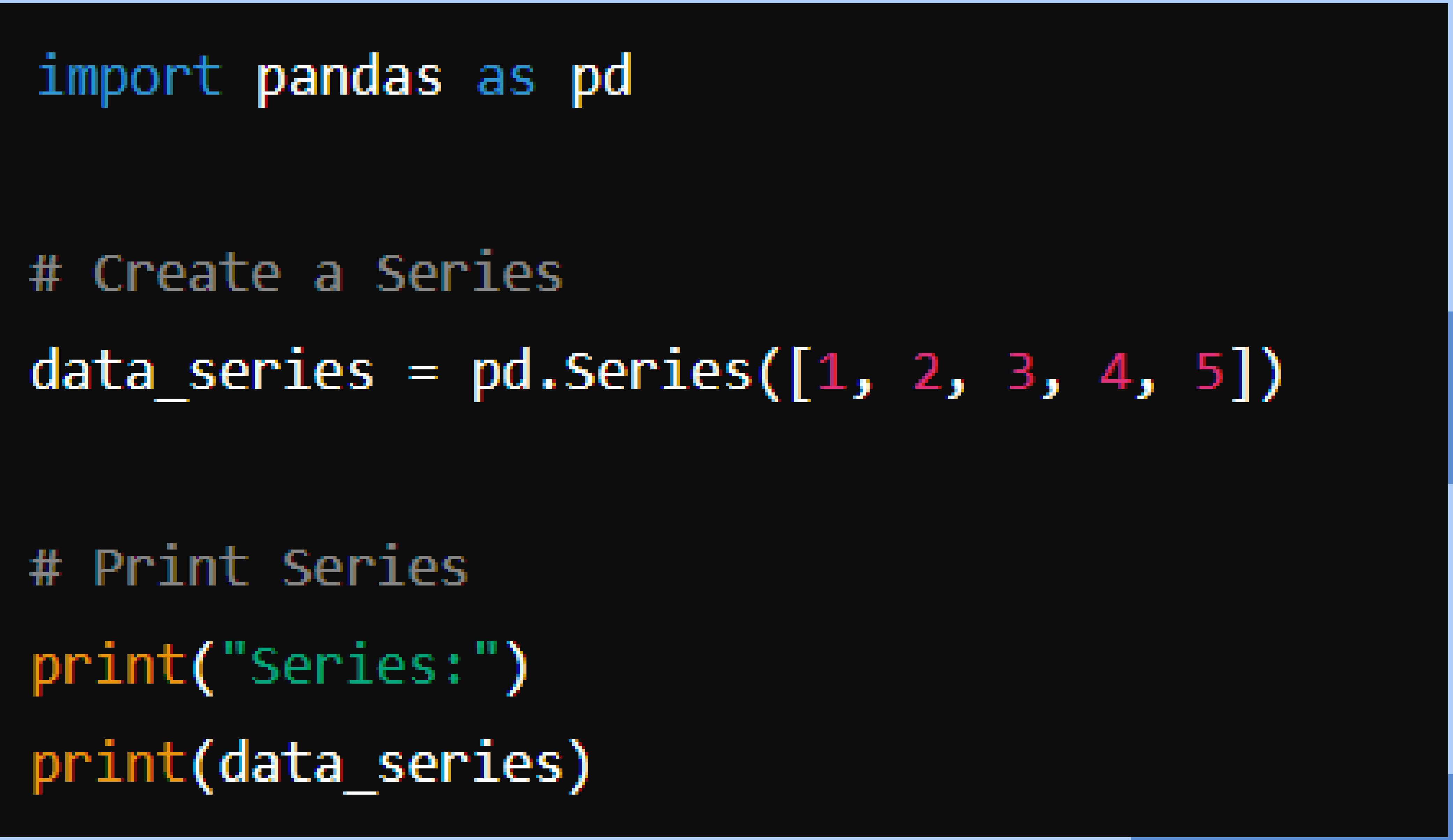
output



```
DataFrame:
        Name  Age
Std 1   Rauf   20
Std 2  Ahsan   19
```

Series

code

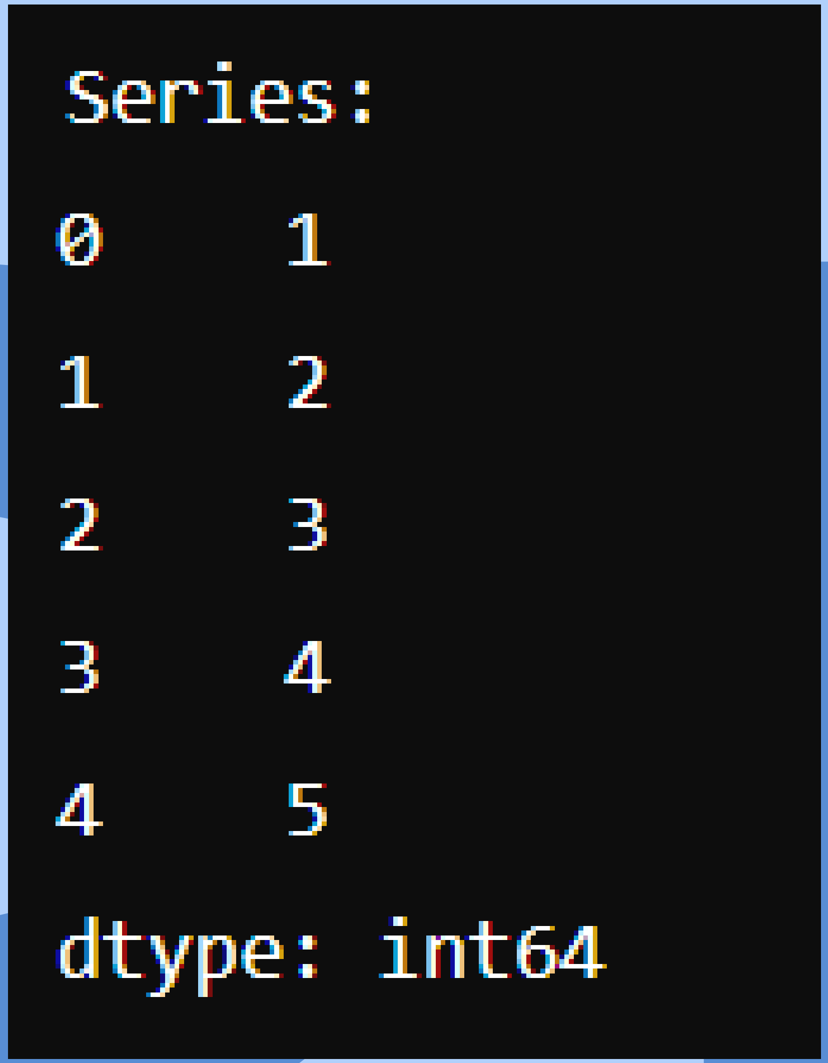


```
import pandas as pd

# Create a Series
data_series = pd.Series([1, 2, 3, 4, 5])

# Print Series
print("Series:")
print(data_series)
```

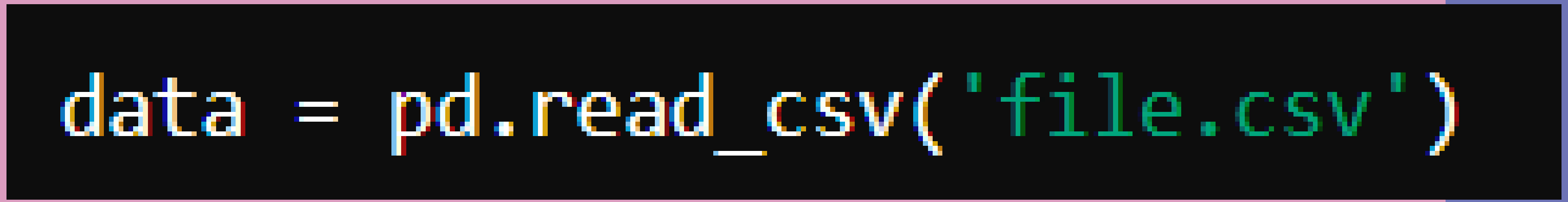
output



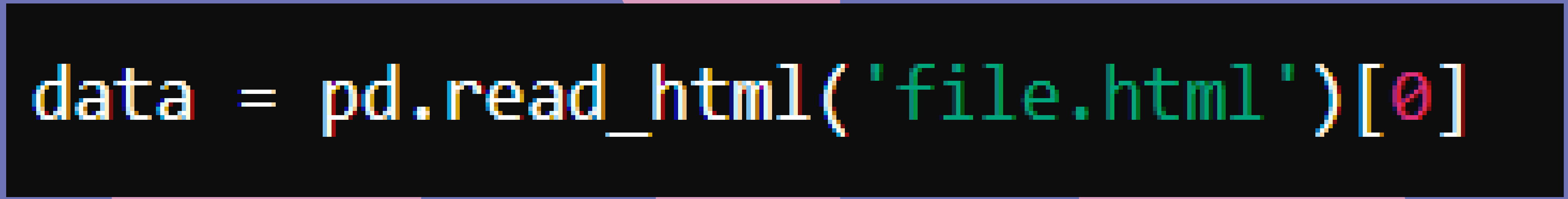
```
Series:
0    1
1    2
2    3
3    4
4    5
dtype: int64
```

Reading

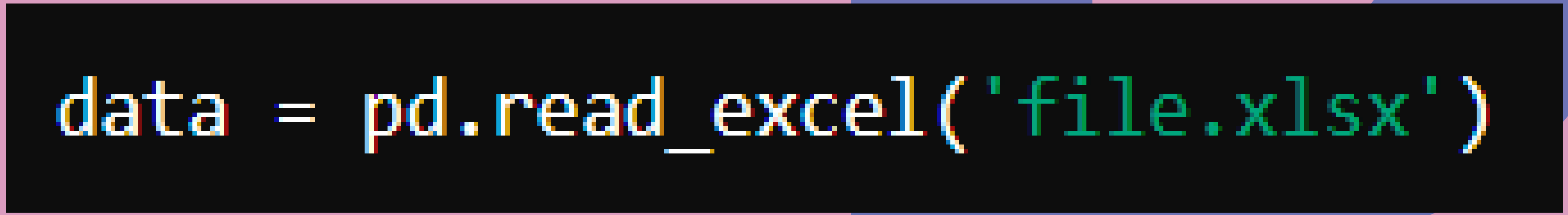
With pandas, reading data means loading files (like CSV, Excel) or databases into DataFrames for analysis and manipulation.



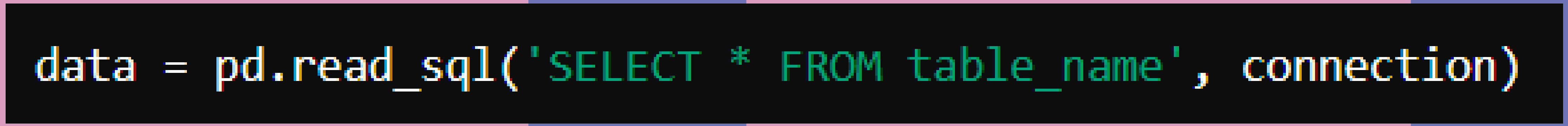
```
data = pd.read_csv('file.csv')
```



```
data = pd.read_html('file.html')[0]
```

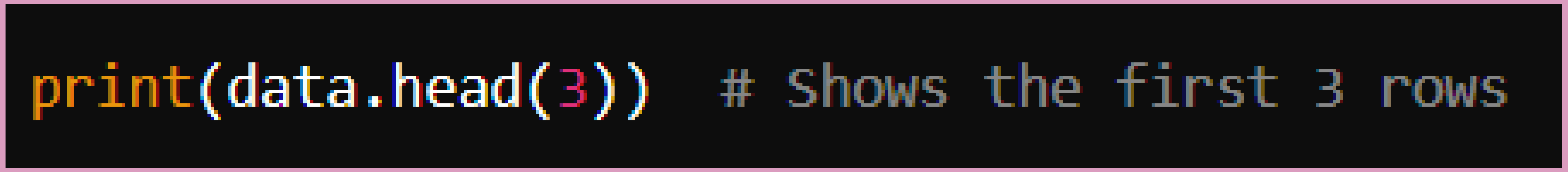


```
data = pd.read_excel('file.xlsx')
```

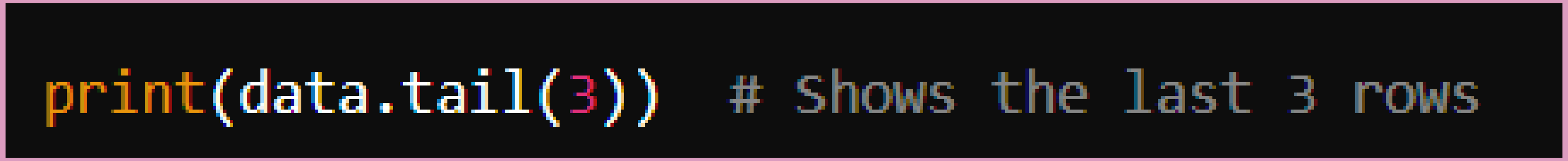


```
data = pd.read_sql('SELECT * FROM table_name', connection)
```

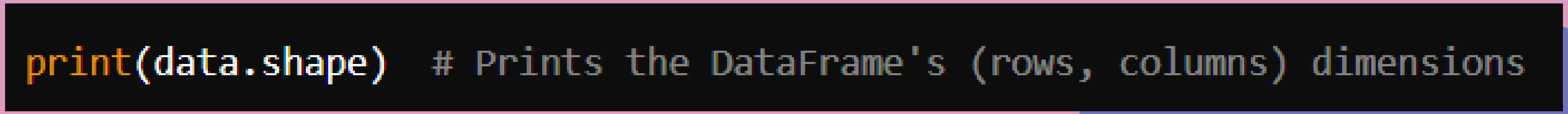
Reading



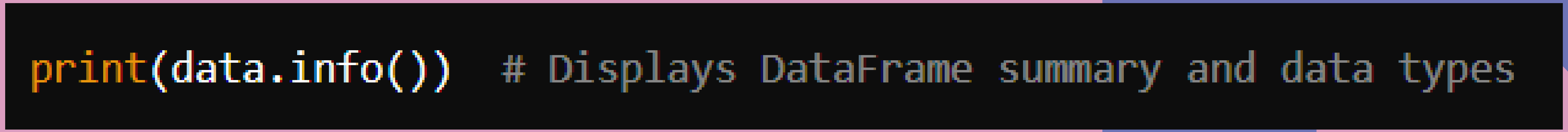
```python
print(data.head(3))  # Shows the first 3 rows
```



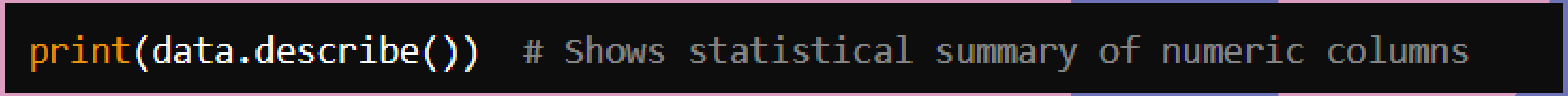
```python
print(data.tail(3))  # Shows the last 3 rows
```



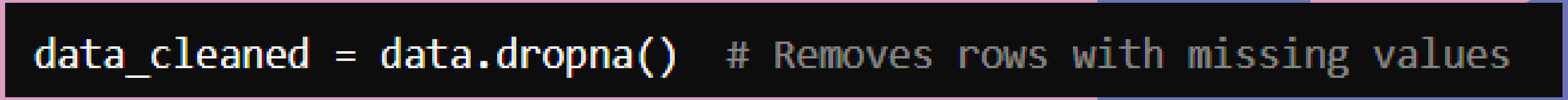
```python
print(data.shape)  # Prints the DataFrame's (rows, columns) dimensions
```



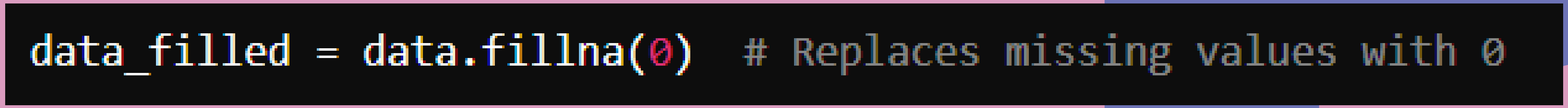
```python
print(data.info())  # Displays DataFrame summary and data types
```



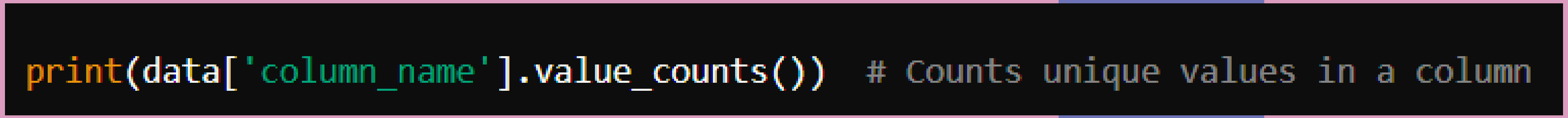
```python
print(data.describe())  # Shows statistical summary of numeric columns
```



```python
data_cleaned = data.dropna()  # Removes rows with missing values
```



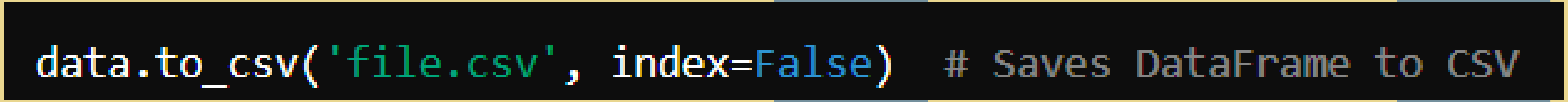
```python
data_filled = data.fillna(0)  # Replaces missing values with 0
```



```python
print(data['column_name'].value_counts())  # Counts unique values in a column
```

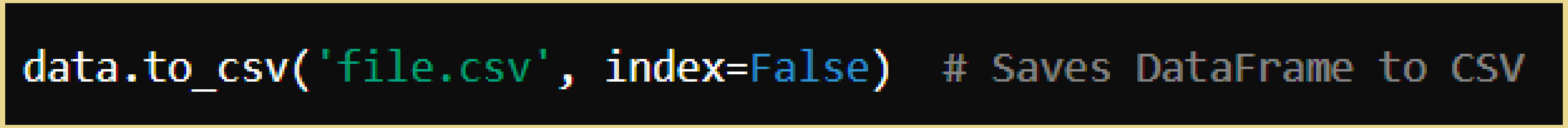
Writing

Writing in pandas involves saving DataFrames to files like CSV or Excel using functions such as `to_csv()` or `to_excel()`.

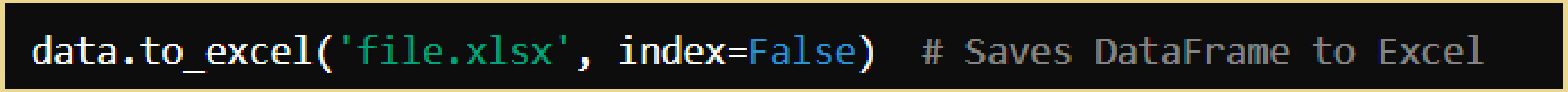


```python
data.to_csv('file.csv', index=False)  # Saves DataFrame to CSV
```

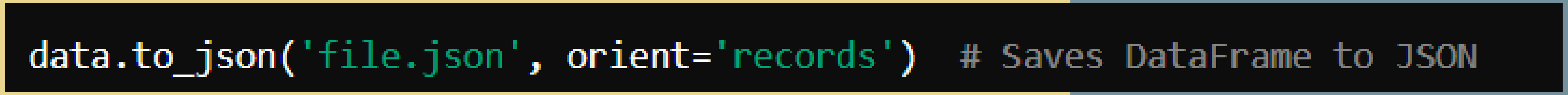
Writing



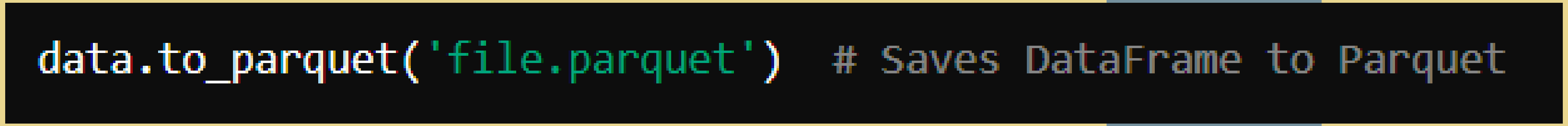
```
data.to_csv('file.csv', index=False)  # Saves DataFrame to CSV
```



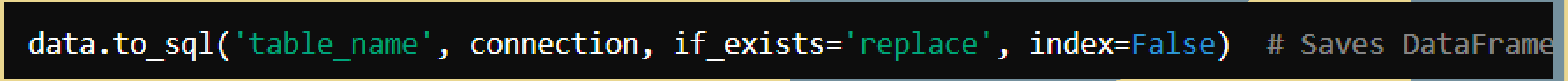
```
data.to_excel('file.xlsx', index=False)  # Saves DataFrame to Excel
```



```
data.to_json('file.json', orient='records')  # Saves DataFrame to JSON
```



```
data.to_parquet('file.parquet')  # Saves DataFrame to Parquet
```

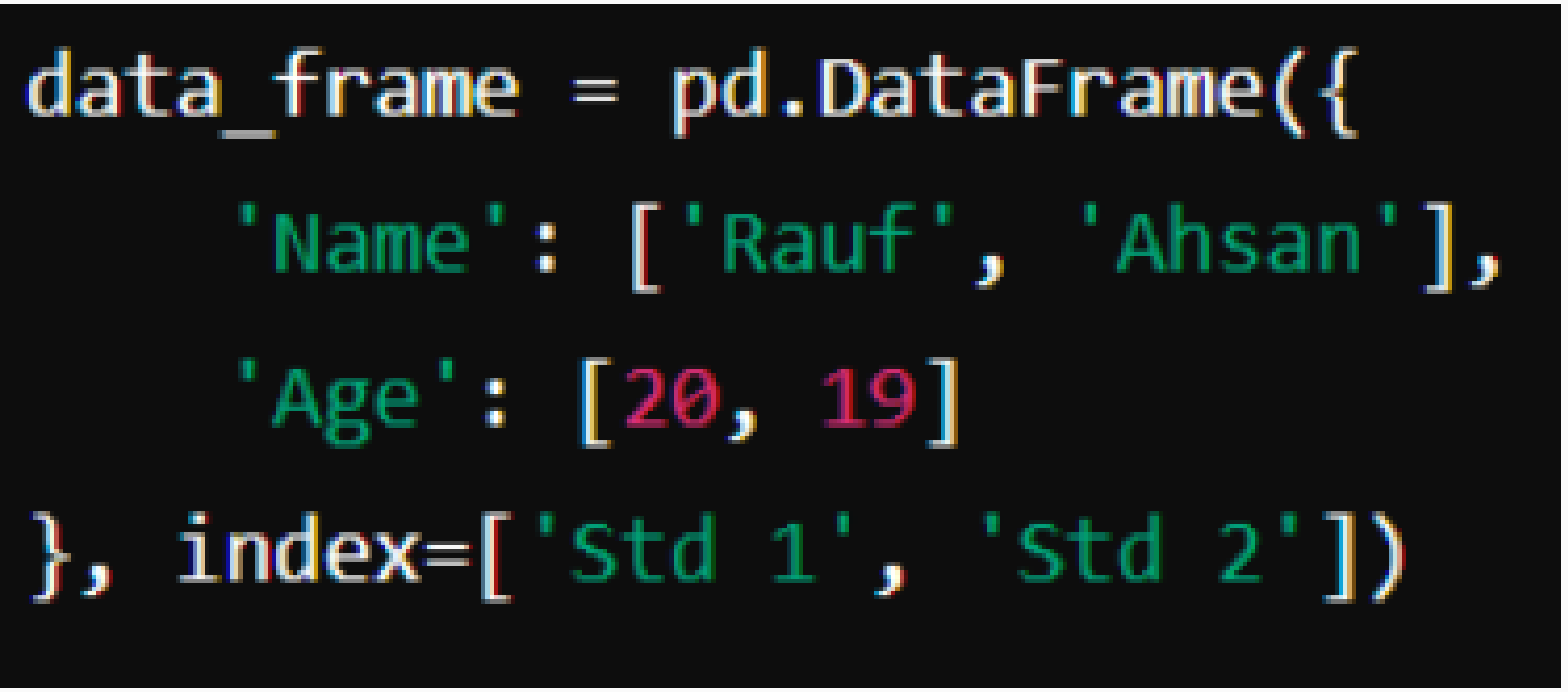


```
data.to_sql('table_name', connection, if_exists='replace', index=False)  # Saves DataFrame
```

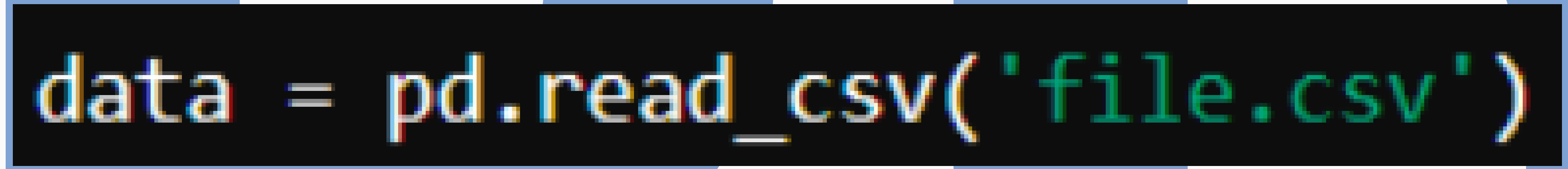
1. **Creating**

2. **Reading**

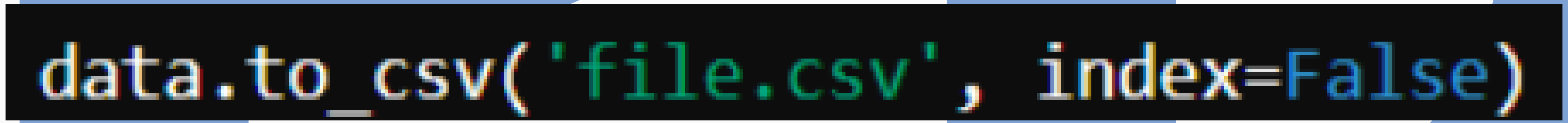
3. **Writing**



```
data_frame = pd.DataFrame({
    'Name': ['Rauf', 'Ahsan'],
    'Age': [20, 19]
}, index=['Std 1', 'Std 2'])
```



```
data = pd.read_csv('file.csv')
```



```
data.to_csv('file.csv', index=False)
```

pandas

Thank you

By Abdul Rauf Jatoi