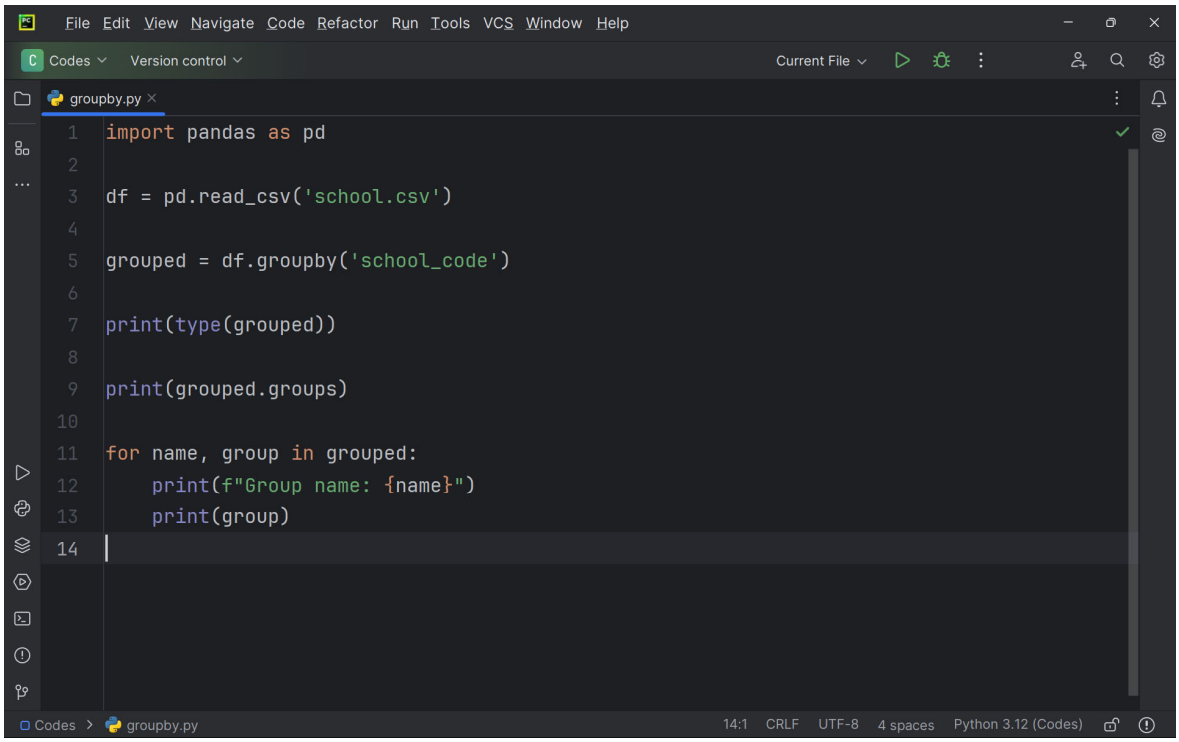
Experiment 16

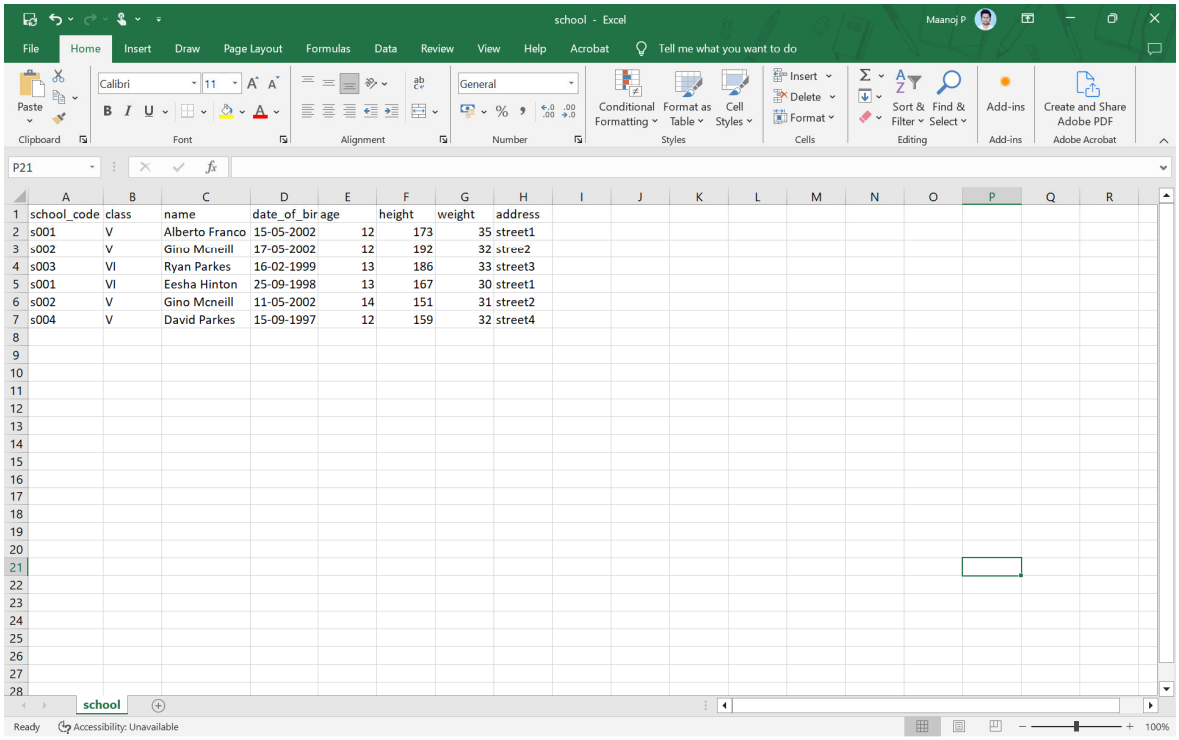
Aim:

To develop a Pandas program to split the dataset into groups based on school code. Also check the type of GroupBy object

Code:

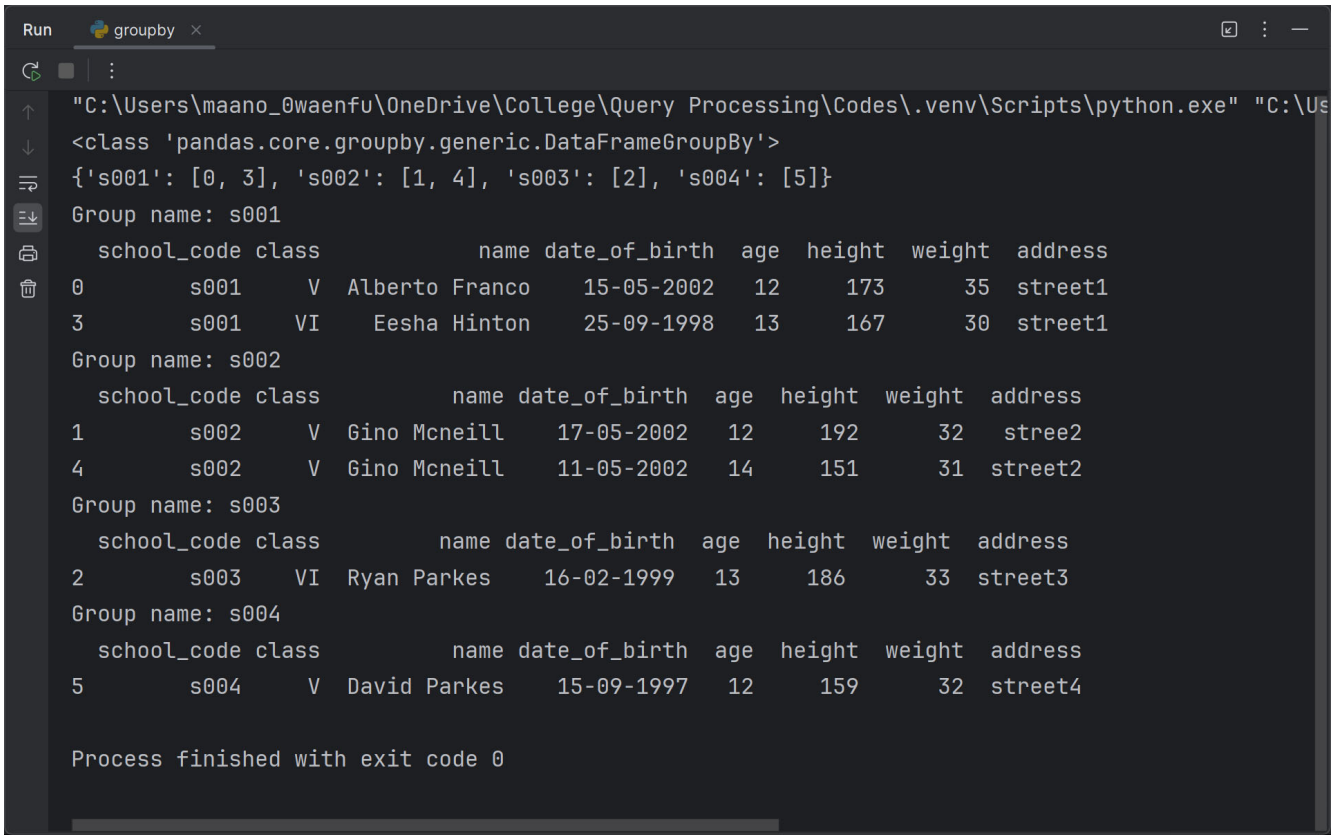


Input:



school_code	class	name	date_of_bir	age	height	weight	address
s001	V	Alberto Franco	15-05-2002	12	173	35	street1
s002	V	Gino Mcneill	17-05-2002	12	192	32	street2
s003	VI	Ryan Parkes	16-02-1999	13	186	33	street3
s001	VI	Eesha Hinton	25-09-1998	13	167	30	street1
s002	V	Gino Mcneill	11-05-2002	14	151	31	street2
s004	V	David Parkes	15-09-1997	12	159	32	street4

Output:

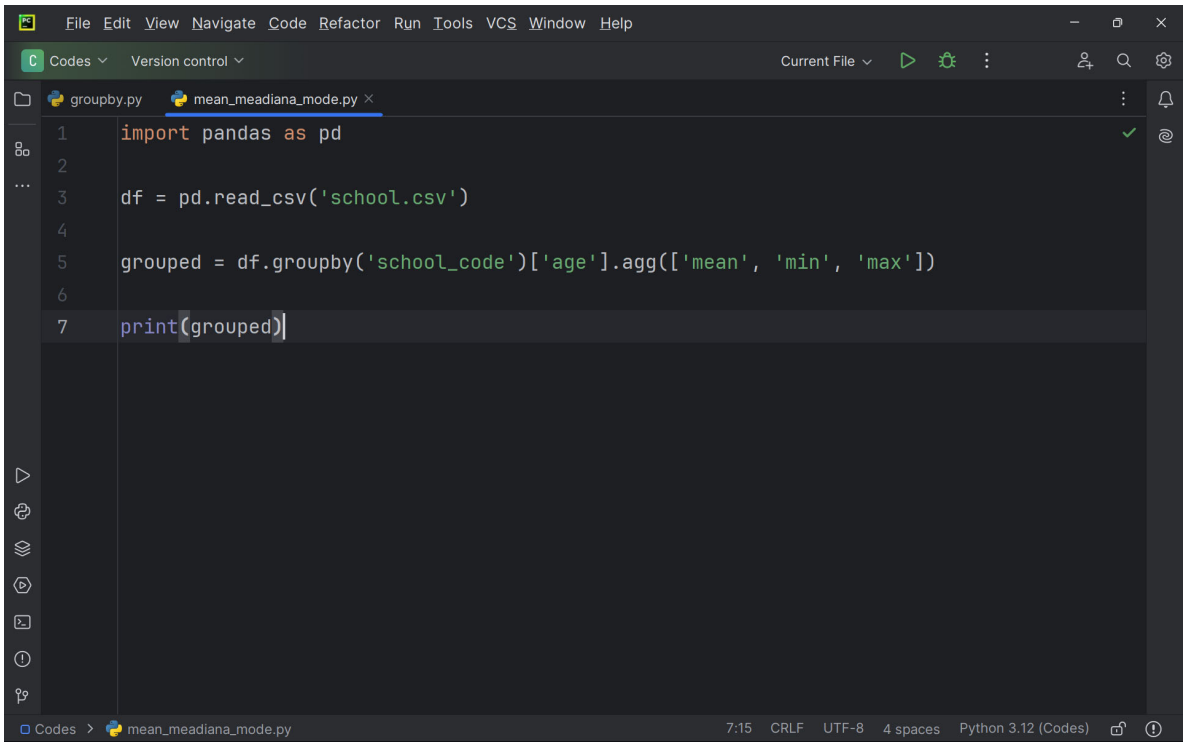

Run
groupby
"C:\Users\maano_0waenfu\OneDrive\College\Query Processing\Codes\.venv\Scripts\python.exe" "C:\Us
<class 'pandas.core.groupby.generic.DataFrameGroupBy'>
{'s001': [0, 3], 's002': [1, 4], 's003': [2], 's004': [5]}
Group name: s001
school_code class name date_of_birth age height weight address
0 s001 V Alberto Franco 15-05-2002 12 173 35 street1
3 s001 VI Eesha Hinton 25-09-1998 13 167 30 street1
Group name: s002
school_code class name date_of_birth age height weight address
1 s002 V Gino Mcneill 17-05-2002 12 192 32 stree2
4 s002 V Gino Mcneill 11-05-2002 14 151 31 street2
Group name: s003
school_code class name date_of_birth age height weight address
2 s003 VI Ryan Parkes 16-02-1999 13 186 33 street3
Group name: s004
school_code class name date_of_birth age height weight address
5 s004 V David Parkes 15-09-1997 12 159 32 street4
Process finished with exit code 0

Experiment 17

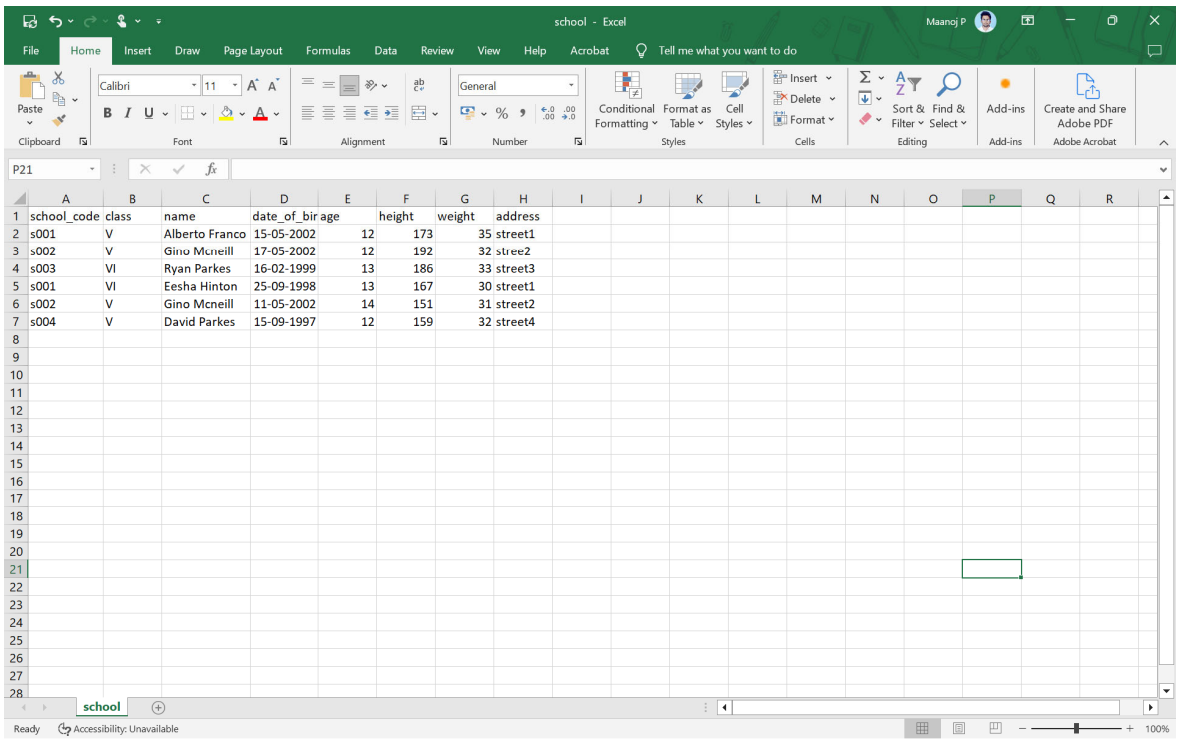
Aim:

To develop a Pandas program to split the dataset into groups based on school code and get mean, min, and max value of age for each school

Code:

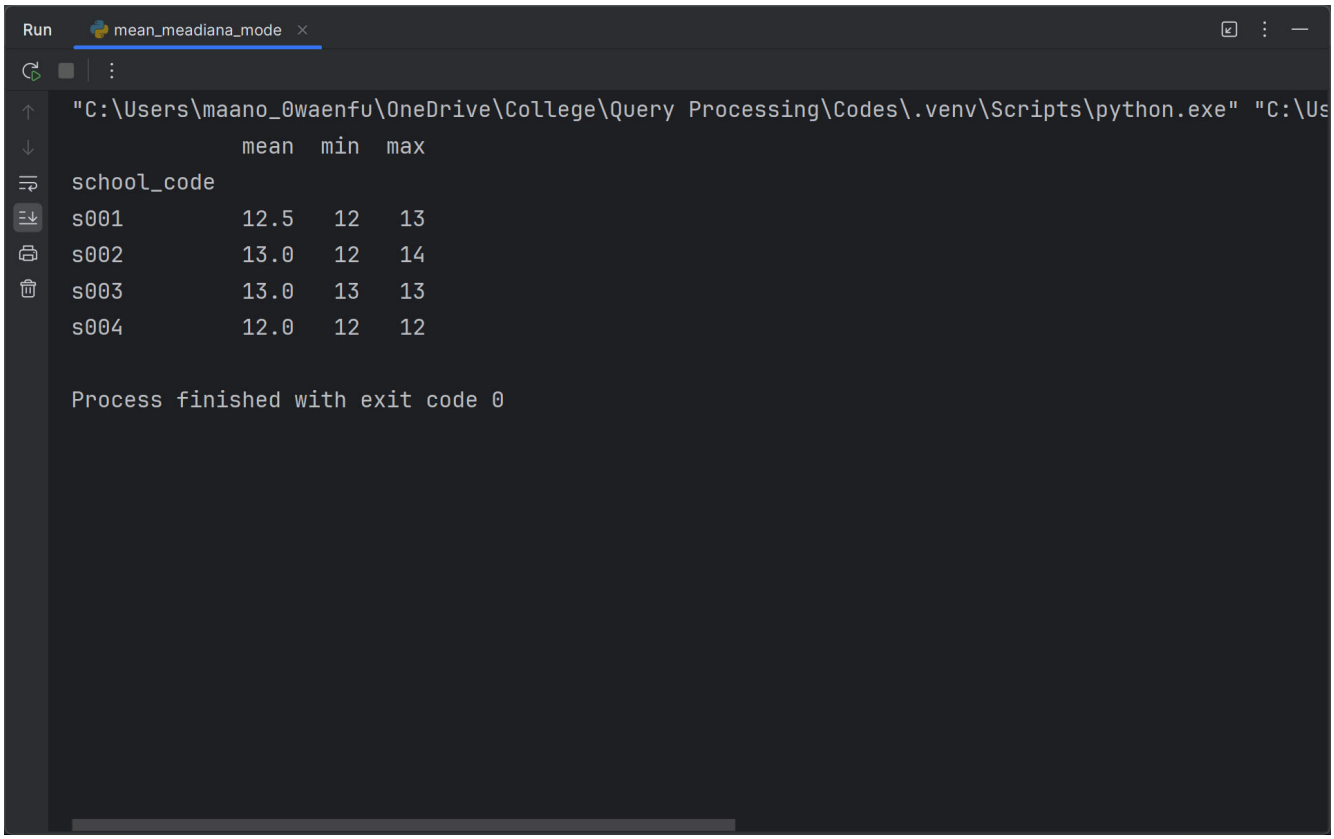


Input:



school_code	class	name	date_of_bir	age	height	weight	address
s001	V	Alberto Franco	15-05-2002	12	173	35	street1
s002	V	Gino Mcneill	17-05-2002	12	192	32	stree2
s003	VI	Ryan Parkes	16-02-1999	13	186	33	street3
s001	VI	Eesha Hinton	25-09-1998	13	167	30	street1
s002	V	Gino Mcneill	11-05-2002	14	151	31	street2
s004	V	David Parkes	15-09-1997	12	159	32	street4

Output:

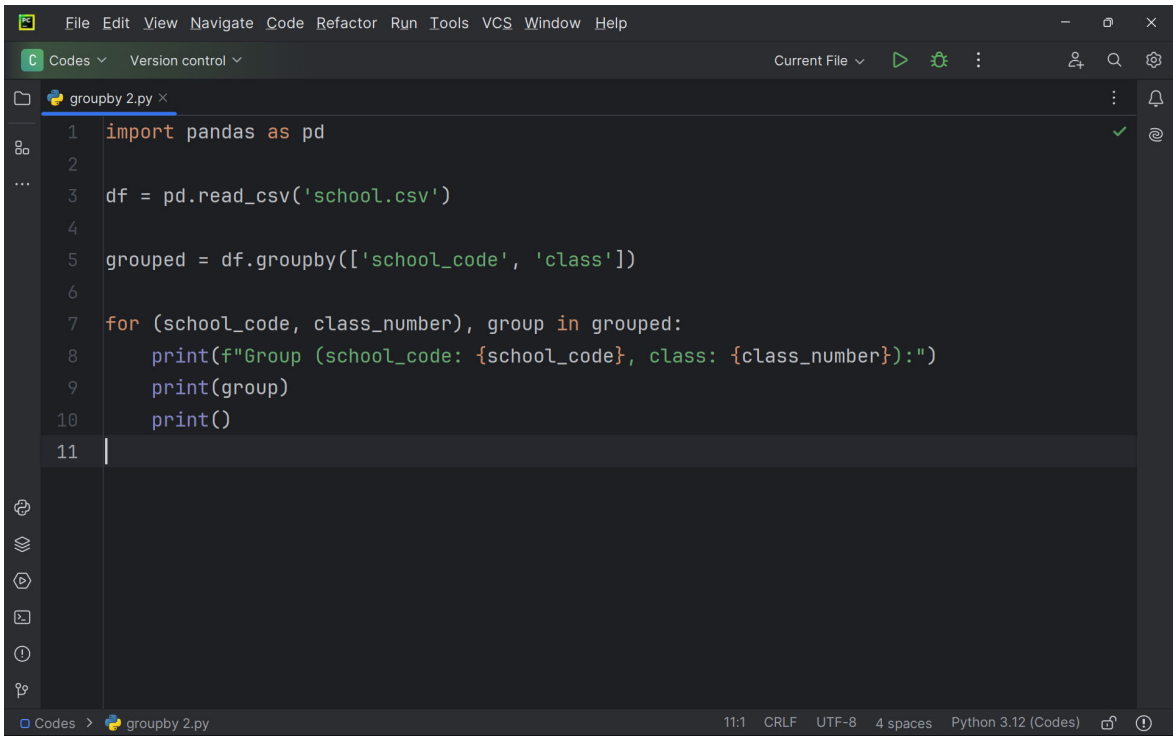

Run
mean_meadiana_mode
"C:\Users\maano_0waenfu\OneDrive\College\Query Processing\Codes\.venv\Scripts\python.exe" "C:\Us
mean min max
school_code
s001 12.5 12 13
s002 13.0 12 14
s003 13.0 13 13
s004 12.0 12 12
Process finished with exit code 0

Experiment 18

Aim:

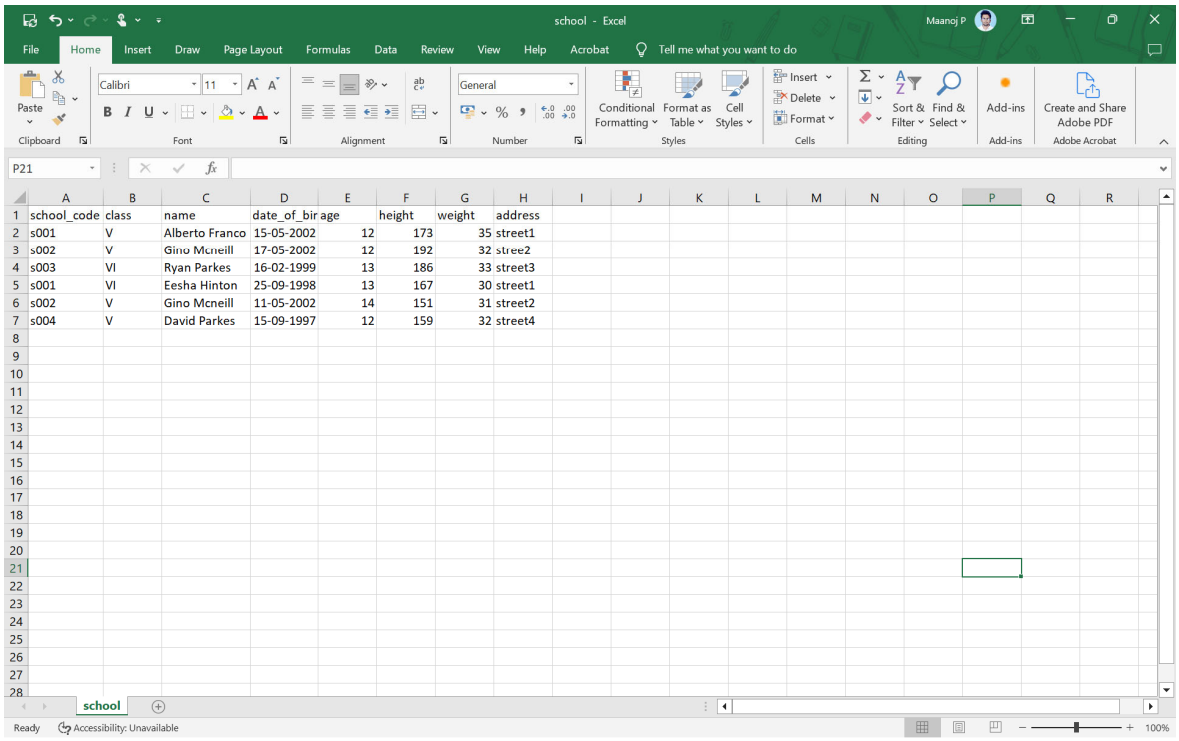
To develop a Pandas program to split the dataset into groups based on school code and class.

Code:



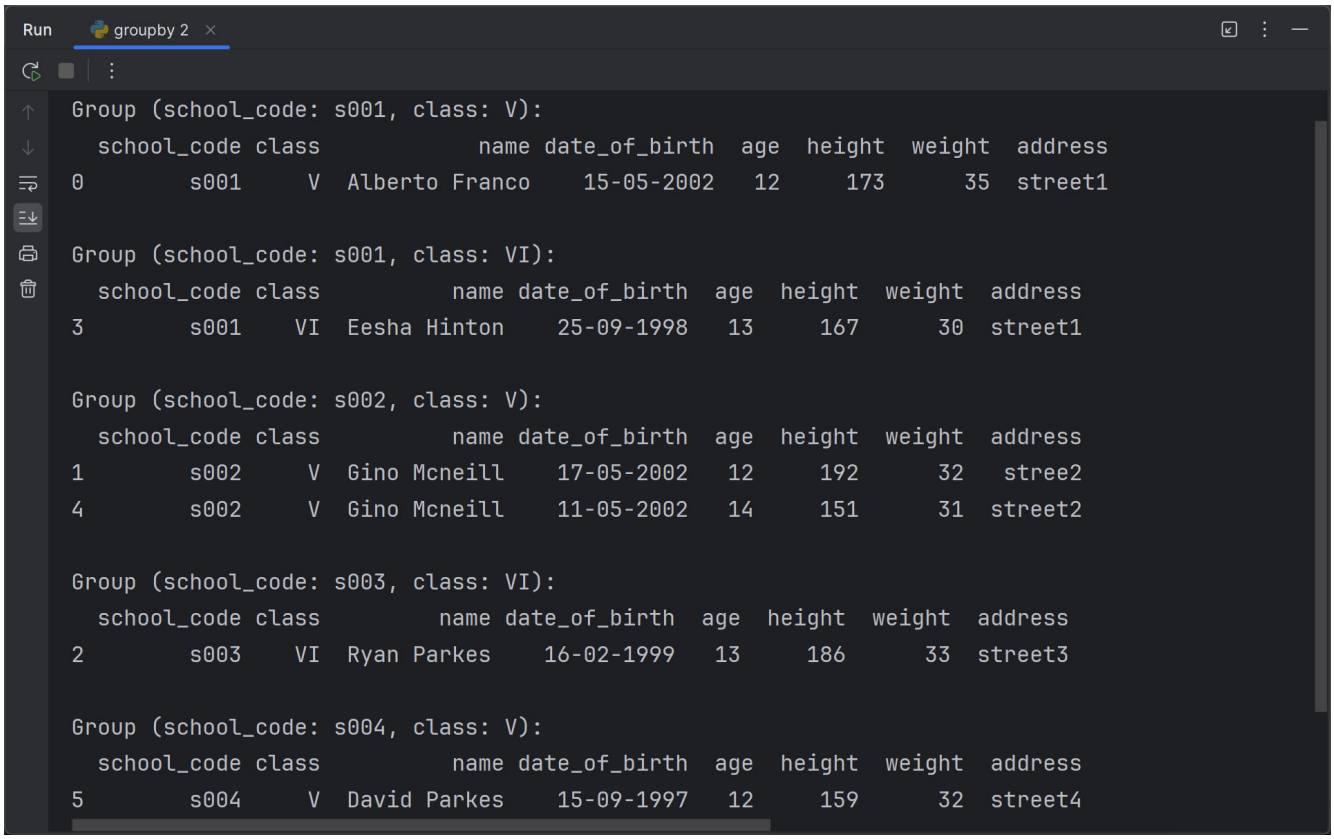
```python
import pandas as pd

df = pd.read_csv('school.csv')

grouped = df.groupby(['school_code', 'class'])

for (school_code, class_number), group in grouped:
    print(f"Group (school_code: {school_code}, class: {class_number}):")
    print(group)
    print()
```

Input:



	A	B	C	D	E	F	G	H
1	school_code	class	name	date_of_bir	age	height	weight	address
2	s001	V	Alberto Franco	15-05-2002	12	173	35	street1
3	s002	V	Gino Mcneill	17-05-2002	12	192	32	street2
4	s003	VI	Ryan Parkes	16-02-1999	13	186	33	street3
5	s001	VI	Eesha Hinton	25-09-1998	13	167	30	street1
6	s002	V	Gino Mcneill	11-05-2002	14	151	31	street2
7	s004	V	David Parkes	15-09-1997	12	159	32	street4

Output:

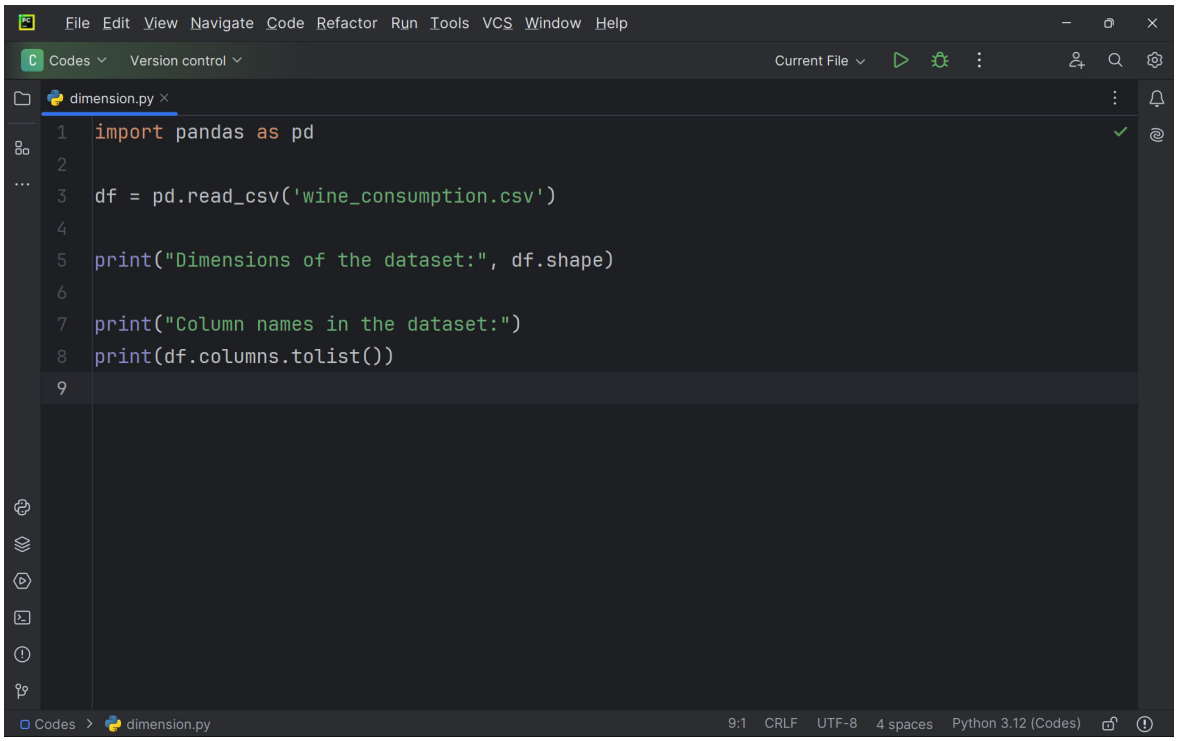

Run
groupby 2
Group (school_code: s001, class: V):
school_code class name date_of_birth age height weight address
0 s001 V Alberto Franco 15-05-2002 12 173 35 street1
Group (school_code: s001, class: VI):
school_code class name date_of_birth age height weight address
3 s001 VI Eesha Hinton 25-09-1998 13 167 30 street1
Group (school_code: s002, class: V):
school_code class name date_of_birth age height weight address
1 s002 V Gino Mcneill 17-05-2002 12 192 32 stree2
4 s002 V Gino Mcneill 11-05-2002 14 151 31 street2
Group (school_code: s003, class: VI):
school_code class name date_of_birth age height weight address
2 s003 VI Ryan Parkes 16-02-1999 13 186 33 street3
Group (school_code: s004, class: V):
school_code class name date_of_birth age height weight address
5 s004 V David Parkes 15-09-1997 12 159 32 street4

Experiment 19

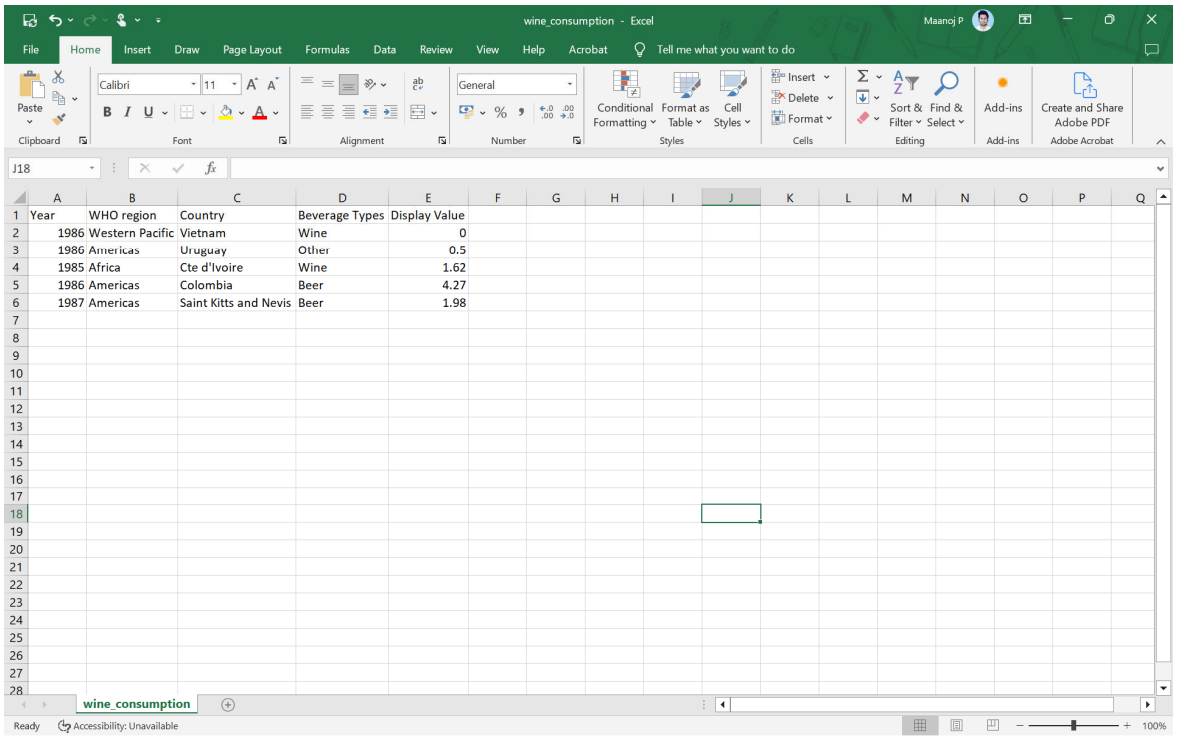
Aim:

To develop a Pandas program to split the dataset into groups based on school code and class.

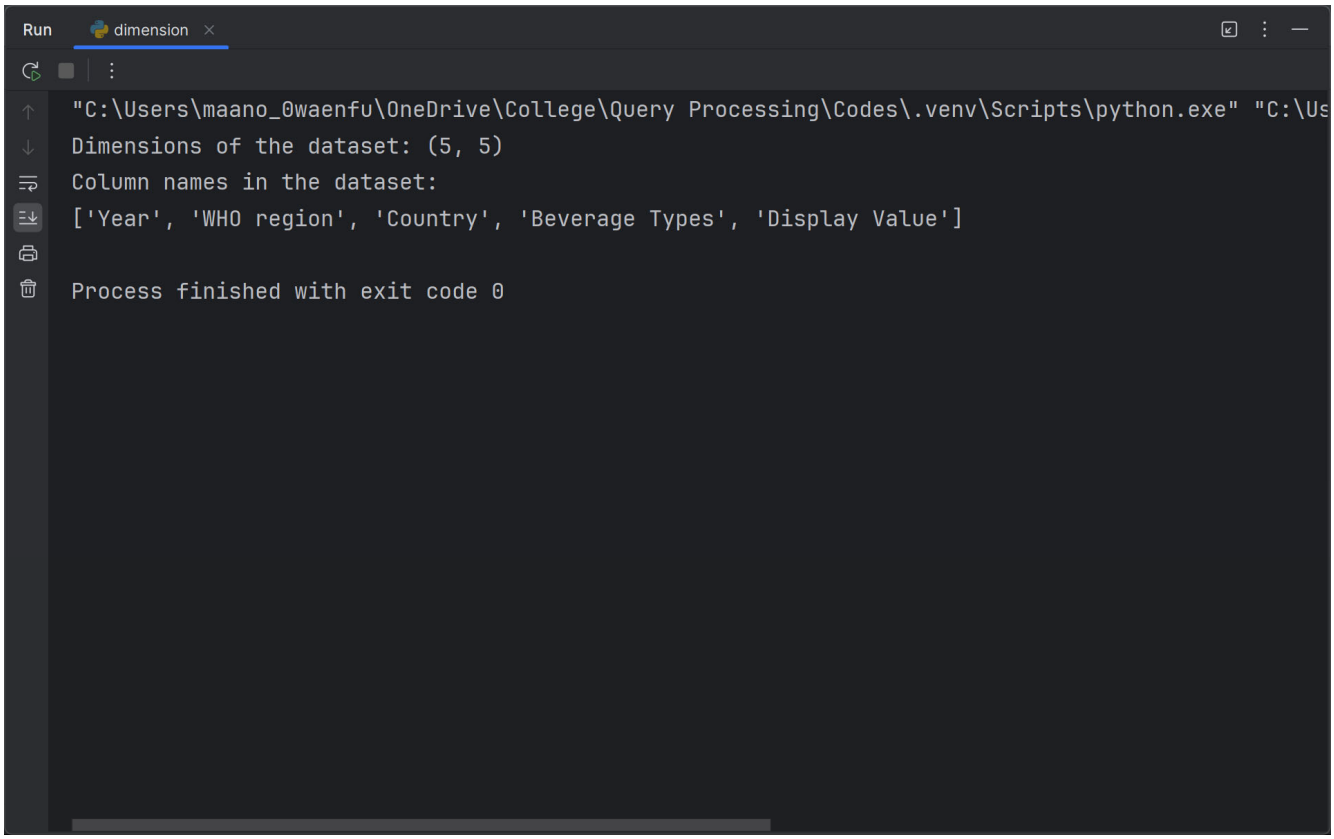
Code:



Input:



Output:

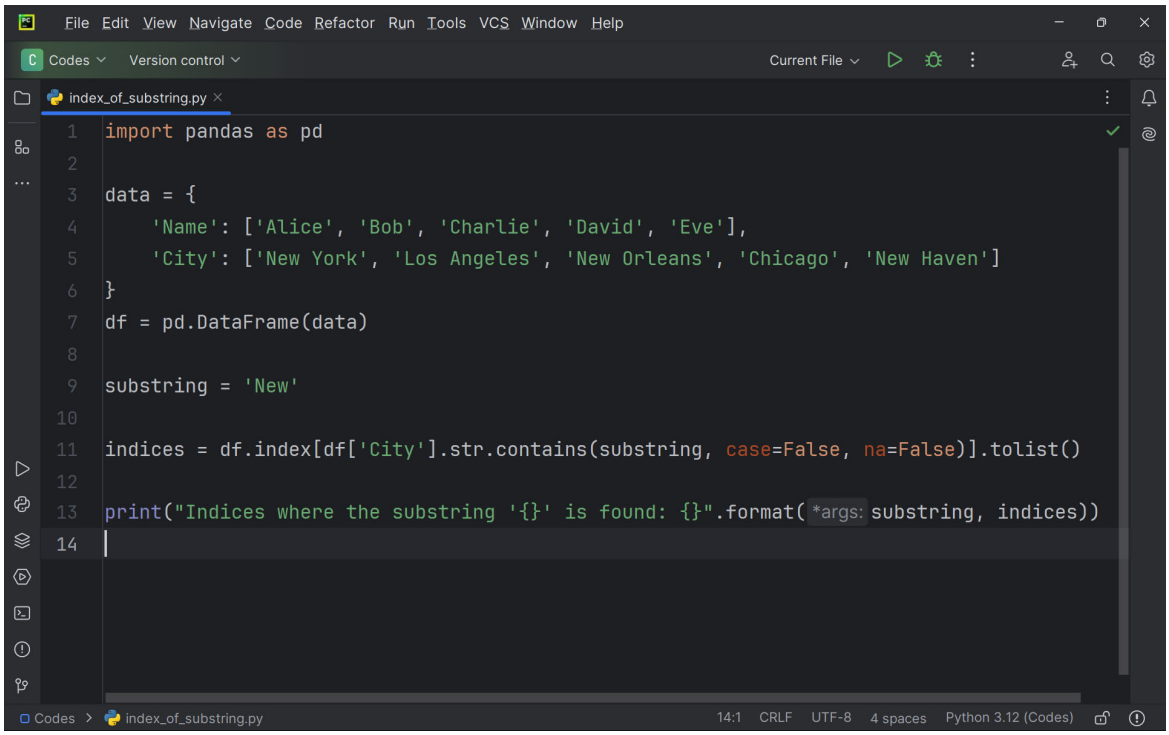

Run
dimension
"C:\Users\maano_0waenfu\OneDrive\College\Query Processing\Codes\.venv\Scripts\python.exe" "C:\Us
Dimensions of the dataset: (5, 5)
Column names in the dataset:
['Year', 'WHO region', 'Country', 'Beverage Types', 'Display Value']
Process finished with exit code 0

Experiment 20

Aim:

To develop a Pandas program to split the dataset into groups based on school code and class.

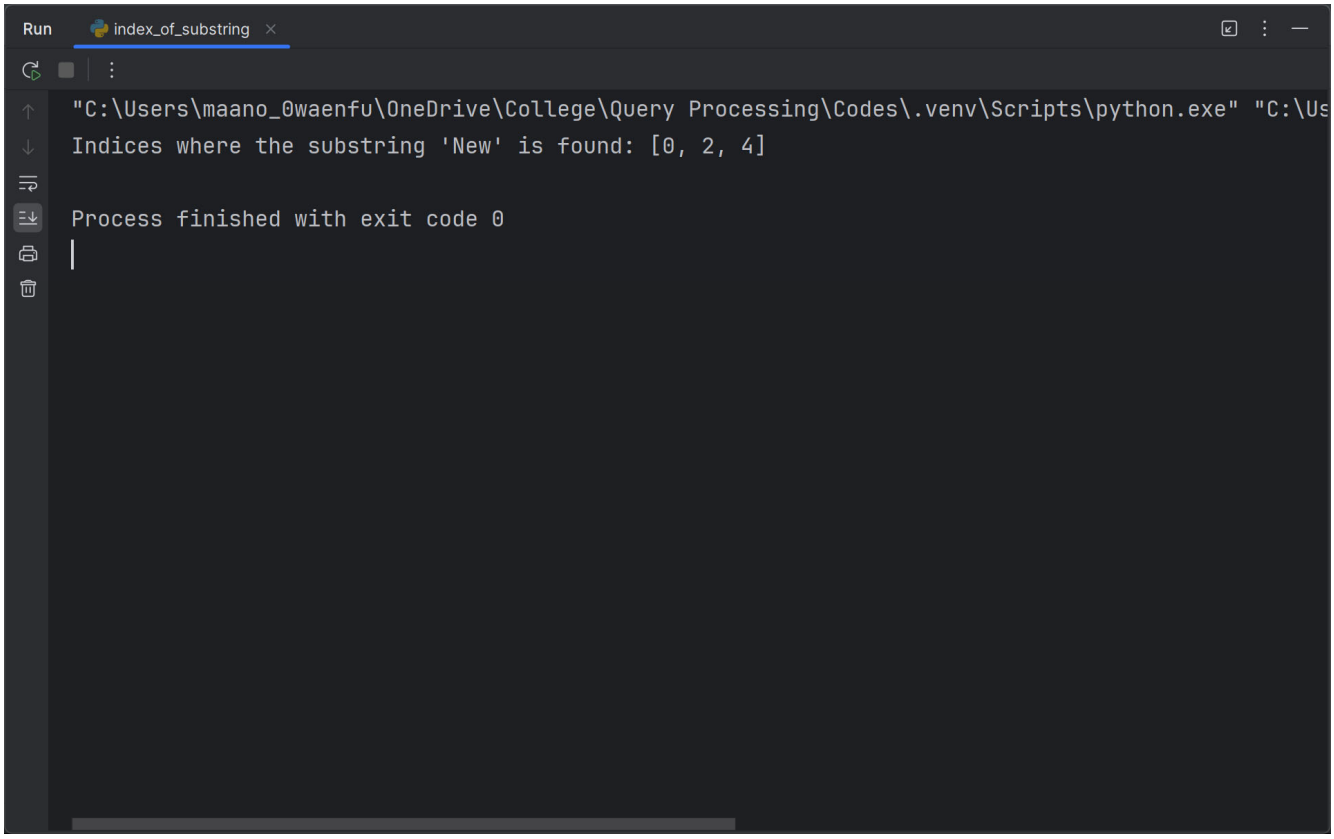
Code:



Input:

```
data = {
    'Name': ['Alice', 'Bob', 'Charlie', 'David', 'Eve'],
    'City': ['New York', 'Los Angeles', 'New Orleans', 'Chicago', 'New Haven']
}
df = pd.DataFrame(data)

substring = 'New'
```

Output:


Run
index_of_substring
"C:\Users\maano_0waenfu\OneDrive\College\Query Processing\Codes\.venv\Scripts\python.exe" "C:\Us
Indices where the substring 'New' is found: [0, 2, 4]
Process finished with exit code 0