

Міністерство освіти та науки України  
Львівський національний університет імені Івана Франка

Факультет електроніки та  
комп'ютерних технологій

## Звіт

Про виконання лабораторної роботи №9  
з курсу “ Основи аналітики даних”  
«Pandas: Структурна конкатенація операцій»

Виконав:

Студент групи ФЕІ-43

Масендич Максим

Перевірив:

асис. Лозинський В. М.

Львів-2024

**Мета роботи:** отримати навички у об'єднанні багатовимірних масивів. **Хід роботи:**

1. Створіть новий файл. Підключіть бібліотеку Pandas.
2. Виконайте наступні завдання:
  - 2.1. Напишіть програму Pandas, щоб об'єднати два фрейми даних уздовж рядків і призначити всі дані.

```
student_data1:
  student_id      name marks
0    S1 Danniella Fenton  200
1    S2  Ryder Storey    210
2    S3  Bryce Jensen   190
3    S4    Ed Bernal    222
4    S5  Kwame Morin    199
student_data2:
  student_id      name marks
0    S4  Scarlett Fisher   201
1    S5  Carla Williamson   200
2    S6   Dante Morse     198
3    S7  Kaiser William   219
4    S8  Madeeha Preston   201
```

```
# Об'єднання двох фреймів даних уздовж рядків
student_data1 = pd.DataFrame({
    'student_id': ['S1', 'S2', 'S3', 'S4', 'S5'],
    'name': ['Danniella Fenton', 'Ryder Storey', 'Bryce Jensen', 'Ed Bernal', 'Kwame Morin'],
    'marks': [200, 210, 190, 222, 199]
})
student_data2 = pd.DataFrame({
    'student_id': ['S4', 'S5', 'S6', 'S7', 'S8'],
    'name': ['Scarlette Fisher', 'Carla Williamson', 'Dante Morse', 'Kaiser William', 'Madeeha Preston'],
    'marks': [201, 200, 198, 219, 201]
})

merged_data_rows = pd.concat([student_data1, student_data2], ignore_index=True)
print("\nDataFrame після об'єднання уздовж рядків:\n", merged_data_rows)
```

DataFrame після об'єднання уздовж рядків:

	student_id	name	marks
0	S1	Danniella Fenton	200
1	S2	Ryder Storey	210
2	S3	Bryce Jensen	190
3	S4	Ed Bernal	222
4	S5	Kwame Morin	199
5	S4	Scarlette Fisher	201
6	S5	Carla Williamson	200
7	S6	Dante Morse	198
8	S7	Kaiser William	219
9	S8	Madeeha Preston	201

2.2. Напишіть програму Pandas, щоб об'єднати два фрейми даних уздовж стовпців і призначити всі дані. Використовуйте фрейми student\_data1 та student\_data2

```
# Об'єднання двох фреймів даних уздовж стовпців
merged_data_cols = pd.concat([student_data1, student_data2], axis=1)
print("\nDataFrame після об'єднання уздовж стовпців:\n", merged_data_cols)
```

DataFrame після об'єднання уздовж стовпців:

	student_id	name	marks	student_id	name	marks
0	S1	Danniella Fenton	200	S4	Scarlette Fisher	201
1	S2	Ryder Storey	210	S5	Carla Williamson	200
2	S3	Bryce Jensen	190	S6	Dante Morse	198
3	S4	Ed Bernal	222	S7	Kaiser William	219
4	S5	Kwame Morin	199	S8	Madeeha Preston	201

2.3. Напишіть програму Pandas для додавання рядків до існуючого DataFrame і відображення об'єднаних даних. Використовуйте фрейми student\_data1 та student\_data2

```
student_id    S6
name          Scarlett Fisher
marks         205
```

```
# Додавання рядка до існуючого DataFrame
merged_data_rows.loc[len(merged_data_rows.index)] = ['S6', 'Scarlette Fisher',
205] print("\nDataFrame після додавання рядка:\n", merged_data_rows)
```

```
DataFrame після додавання рядка:
   student_id  name  marks
0         S1 Danniella Fenton  200
1         S2   Ryder Storey  210
2         S3   Bryce Jensen  190
3         S4     Ed Bernal  222
4         S5   Kwame Morin  199
5         S4  Scarlettte Fisher  201
6         S5   Carla Williamson  200
7         S6    Dante Morse  198
8         S7   Kaiser William  219
9         S8   Madeeha Preston  201
10        S6  Scarlettte Fisher  205
```

2.4. Напишіть програму Pandas, щоб додати список dictionaries або series до існуючого DataFrame і відобразити об'єднані дані.

Використовуйте фрейми student\_data1 та student\_data2

```
student_id    S6
name          Scarlettte Fisher
marks         205
```

```
# Додавання списку dictionaries або Series до існуючого DataFrame
new_data = {'student_id': 'S6', 'name': 'Scarlettte Fisher', 'marks': 205}
merged_data_rows.loc[len(merged_data_rows.index)]=new_data
print("\nDataFrame після додавання списку dictionaries або Series:\n",
merged_data_rows)
```

```
DataFrame після додавання списку dictionaries або Series:
   student_id  name  marks
0         S1 Danniella Fenton  200
1         S2   Ryder Storey  210
2         S3   Bryce Jensen  190
3         S4     Ed Bernal  222
4         S5   Kwame Morin  199
5         S4  Scarlettte Fisher  201
6         S5   Carla Williamson  200
7         S6    Dante Morse  198
8         S7   Kaiser William  219
9         S8   Madeeha Preston  201
10        S6  Scarlettte Fisher  205
11        S6  Scarlettte Fisher  205
```

2.5. Напишіть програму Pandas для об'єднання двох фреймів даних уздовж рядків і об'єднання з іншим фреймом даних уздовж загального ідентифікатора стовпця. Використовуйте фрейми student\_data1 та student\_data2

```

exam_data:
  student_id exam_id
0      S1      23
1      S2      45
2      S3      12
3      S4      67
4      S5      21
5      S7      55
6      S8      33
7      S9      14
8     S10      56
9     S11      83
10     S12      88
11     S13      12

```

# Об'єднання двох фреймів даних уздовж рядків та об'єднання з іншим фреймом даних уздовж загального ідентифікатора стовпця

```

exam_data = pd.DataFrame({
    'student_id': ['S1', 'S2', 'S3', 'S4', 'S5', 'S7', 'S8', 'S9', 'S10', 'S11',
                  'S12', 'S13'],
    'exam_id': [23, 45, 12, 67, 21, 55, 33, 14, 56, 83, 88, 12]
})

merged_data_rows = pd.concat([student_data1, student_data2],
                              ignore_index=True)
merged_data = pd.merge(merged_data_rows, exam_data,
                        on='student_id')
print("\nDataFrame після об'єднання уздовж рядків та об'єднання з іншим фреймом уздовж ідентифікатора стовпця:\n", merged_data)

```

DataFrame після об'єднання уздовж рядків та об'єднання з іншим фреймом уздовж ідентифікатора стовпця:

	student_id	name	marks	exam_id
0	S1	Danniella Fenton	200	23
1	S2	Ryder Storey	210	45
2	S3	Bryce Jensen	190	12
3	S4	Ed Bernal	222	67
4	S5	Kwame Morin	199	21
5	S4	Scarlette Fisher	201	67
6	S5	Carla Williamson	200	21
7	S7	Kaiser William	219	55
8	S8	Madeeha Preston	201	33

2.6. Напишіть програму Pandas для об'єднання двох фреймів даних за допомогою спільного стовпця обох фреймів даних.

Використовуйте фрейми student\_data1 та student\_data2

```

# Об'єднання двох фреймів даних за допомогою спільного стовпця
merged_data_shared_col = pd.merge(student_data1, student_data2, on='student_id')
print("\nDataFrame після об'єднання за допомогою спільного стовпця:\n",
      merged_data_shared_col)

```



```
DataFrame після об'єднання за допомогою спільного стовпця:
```

	student_id	name_x	marks_x	name_y	marks_y
0	S4	Ed Bernal	222	Scarlette Fisher	201
1	S5	Kwame Morin	199	Carla Williamson	200

2.7. Напишіть програму Pandas, щоб об'єднати два фрейми даних із відповідними записами з обох сторін, де це можливо.

Використовуйте фрейми student\_data1 та student\_data2

```
# Об'єднання двох фреймів даних із відповідними записами
merged_data_common_rows = pd.merge(student_data1, student_data2,
how='outer') print("\nDataFrame після об'єднання із відповідними
записами:\n", merged_data_common_rows)
```

```
DataFrame після об'єднання із відповідними записами:
```

	student_id	name	marks
0	S1	Danniella Fenton	200
1	S2	Ryder Storey	210
2	S3	Bryce Jensen	190
3	S4	Ed Bernal	222
4	S4	Scarlette Fisher	201
5	S5	Carla Williamson	200
6	S5	Kwame Morin	199
7	S6	Dante Morse	198
8	S7	Kaiser William	219
9	S8	Madeeha Preston	201

2.8. Напишіть програму Pandas для об'єднання (з'єднання зліва) двох кадрів даних, використовуючи ключі лише з лівого кадру даних.

```
data1:
  key1 key2 P Q
0 K0 K0 P0 Q0
1 K0 K1 P1 Q1
2 K1 K0 P2 Q2
3 K2 K1 P3 Q3
data2:
  key1 key2 R S
0 K0 K0 R0 S0
1 K1 K0 R1 S1
2 K1 K0 R2 S2
3 K2 K0 R3 S3
```

```
# Об'єднання (з'єднання зліва) двох кадрів даних, використовуючи ключі лише
з лівого кадру даних
```

```
data1 = pd.DataFrame({'key1': ['K0', 'K0', 'K1', 'K2'],
                      'key2': ['K0', 'K1', 'K0', 'K1'],
                      'P': ['P0', 'P1', 'P2', 'P3'],
                      'Q': ['Q0', 'Q1', 'Q2', 'Q3']})
```

```
data2 = pd.DataFrame({'key1': ['K0', 'K1', 'K1', 'K2'],
                      'key2': ['K0', 'K0', 'K0', 'K0'],
                      'R': ['R0', 'R1', 'R2', 'R3'],
                      'S': ['S0', 'S1', 'S2', 'S3']})

merged_data_left = pd.merge(data1, data2, on=['key1', 'key2'], how='left')
print("\nDataFrame після зліва з'єднання двох кадрів даних:\n", merged_data_left)
```

```
DataFrame після зліва з'єднання двох кадрів даних:
   key1 key2  P  Q  R  S
0  K0  K0  P0  Q0  R0  S0
1  K0  K1  P1  Q1  NaN NaN
2  K1  K0  P2  Q2  R1  S1
3  K1  K0  P2  Q2  R2  S2
4  K2  K1  P3  Q3  NaN NaN
```

2.9. Напишіть програму Pandas для об'єднання двох фреймів даних, використовуючи ключі лише правого фрейму даних. Використовуйте фрейми data1 та data2

```
# Об'єднання двох фреймів даних, використовуючи ключі лише правого фрейму
даних merged_data_right = pd.merge(data1, data2, on=['key1', 'key2'],
how='right') print("\nDataFrame після з'єднання, використовуючи ключі лише з
правого кадру даних:\n", merged_data_right)
```

```
DataFrame після з'єднання, використовуючи ключі лише з правого кадру даних:
   key1 key2  P  Q  R  S
0  K0  K0  P0  Q0  R0  S0
1  K1  K0  P2  Q2  R1  S1
2  K1  K0  P2  Q2  R2  S2
3  K2  K0  NaN NaN  R3  S3
```

2.10. Напишіть програму Pandas для об'єднання двох заданих наборів даних за допомогою кількох ключів об'єднання. Використовуйте фрейми data1 та data2

```
# Об'єднання двох заданих наборів даних за допомогою кількох ключів об'єднання
merged_data_multikey = pd.merge(data1, data2, on=['key1', 'key2'])
print("\nDataFrame після об'єднання за допомогою кількох ключів
об'єднання:\n", merged_data_multikey)
```

DataFrame після об'єднання за допомогою кількох ключів об'єднання:

	key1	key2	P	Q	R	S
0	K0	K0	P0	Q0	R0	S0
1	K1	K0	P2	Q2	R1	S1
2	K1	K0	P2	Q2	R2	S2

2.11. Напишіть програму Pandas для створення нового DataFrame на основі існуючих Series, використовуючи вказаний аргумент і замінивши існуючі назви стовпців. Використовуйте фрейми data1 та data2

```
# Створення нового DataFrame на основі існуючих Series
series_P = pd.Series(['P0', 'P1', 'P2', 'P3'], name='P')
series_Q = pd.Series(['Q0', 'Q1', 'Q2', 'Q3'], name='Q')

new_dataframe_from_series = pd.concat([series_P, series_Q],
axis=1) print("\nНовий DataFrame на основі існуючих Series:\n",
new_dataframe_from_series)
```

Новий DataFrame на основі існуючих Series:

	P	Q
0	P0	Q0
1	P1	Q1
2	P2	Q2
3	P3	Q3

2.12. Напишіть програму Pandas для створення комбінації з двох фреймів даних, де комбінація ідентифікатора стовпця з'являється більше одного разу в обох фреймах даних. Використовуйте фрейми data1 та data2

```
# Створення комбінації з двох фреймів даних
data1_combo = pd.DataFrame({'key': ['K0', 'K1', 'K2'],
                             'A': ['A0', 'A1', 'A2'],
                             'B': ['B0', 'B1', 'B2']})

data2_combo = pd.DataFrame({'key': ['K0', 'K1', 'K2'],
                             'C': ['C0', 'C1', 'C2'],
                             'D': ['D0', 'D1', 'D2']})

combo_data = pd.merge(data1_combo, data2_combo, on='key', how='outer')
print("\nКомбінація з двох фреймів даних:\n", combo_data)
```



Комбінація з двох фреймів даних:

	key	A	B	C	D
0	K0	A0	B0	C0	D0
1	K1	A1	B1	C1	D1
2	K2	A2	B2	C2	D2

2.13. Напишіть програму Pandas, щоб об'єднати стовпці двох потенційно різноіндексованих DataFrame в один результат DataFrame.

```
data1:
  A B
K0 A0 B0
K1 A1 B1
K2 A2 B2
data2:
  C D
K0 C0 D0
K2 C2 D2
K3 C3 D3

# Об'єднання стовпців двох потенційно різноіндексованих DataFrame
data1_multiindex = pd.DataFrame({'A': ['A0', 'A1', 'A2'],
                                  'B': ['B0', 'B1', 'B2']},
                                  index=['K0', 'K1', 'K2'])

data2_multiindex = pd.DataFrame({'C': ['C0', 'C2', 'C3'],
                                  'D': ['D0', 'D2', 'D3']},
                                  index=['K0', 'K2', 'K3'])

merged_multiindex = pd.concat([data1_multiindex, data2_multiindex], axis=1)
print("\nDataFrame після об'єднання стовпців з різноіндексованими DataFrame:\n",
merged_multiindex)
```

DataFrame після об'єднання стовпців з різноіндексованими DataFrame:

	A	B	C	D
K0	A0	B0	C0	D0
K1	A1	B1	NaN	NaN
K2	A2	B2	C2	D2
K3	NaN	NaN	C3	D3

## 2.14. Напишіть програму Pandas для об'єднання двох фреймів даних із різними стовпцями

```
data1:
  key1 key2 P Q
0 K0 K0 P0 Q0
1 K0 K1 P1 Q1
2 K1 K0 P2 Q2
3 K2 K1 P3 Q3
data2:
  key1 key2 R S
0 K0 K0 R0 S0
1 K1 K0 R1 S1
2 K1 K0 R2 S2
3 K2 K0 R3 S3
```

# Об'єднання двох фреймів даних із різними стовпцями

```
data1_diff_cols = pd.DataFrame({'key1': ['K0', 'K0', 'K1', 'K2'],
                                'key2': ['K0', 'K1', 'K0', 'K1'],
                                'P': ['P0', 'P1', 'P2', 'P3'],
                                'Q': ['Q0', 'Q1', 'Q2', 'Q3']})
```

```
data2_diff_cols = pd.DataFrame({'key1': ['K0', 'K1', 'K1', 'K2'],
                                'key2': ['K0', 'K0', 'K0', 'K0'],
                                'R': ['R0', 'R1', 'R2', 'R3'],
                                'S': ['S0', 'S1', 'S2', 'S3']})
```

```
merged_diff_cols = pd.merge(data1_diff_cols, data2_diff_cols, on=['key1', 'key2'])
print("\nDataFrame після об'єднання із різними стовпцями:\n", merged_diff_cols)
```

```
DataFrame після об'єднання із різними стовпцями:
  key1 key2 P Q R S
0 K0 K0 P0 Q0 R0 S0
1 K1 K0 P2 Q2 R1 S1
2 K1 K0 P2 Q2 R2 S2
```

## 3. Оформіть звіт

**Висновок:** на лабораторній роботі я отримав навички у об'єднанні багатовимірних масивів.