

Аналіз даних Titanic та Apple з GUI на Python



Мета:

- Показати приклад аналізу реальних даних за допомогою Python.
- Об'єднати статистичний аналіз, візуалізацію та графічний інтерфейс для зручного перегляду результатів.
- Продемонструвати застосування Pandas, Matplotlib, Seaborn, Scikit-Learn та Tkinter у практичному проекті.

Владислав Бакало

ПМАМ-11

Результат запуску програми: вкладки з аналізом

Статистика по даних Titanic і Apple

Основна статистика			
Групування по статі та віку			
Групування: клас, стать та вік			
Виживання по класу та статі			
Виживання за віковими групами			
Apple: ціна, тренд, середня та прогноз			
Apple: волатильність			
Apple: денна зміна			
Виживання по статі (Titanic):			
Sex	Survived	Count	Percent
female	0	81	25.80
female	1	233	74.20
male	0	468	81.11
male	1	109	18.89
Кількість людей по класу та статі (Titanic):			
Pclass	Sex	Count	Percent
1	female	94	43.52
1	male	122	56.48
2	female	76	41.30
2	male	108	58.70
3	female	144	29.33
3	male	347	70.67
Відсоток виживших за віковими групами:			
AgeGroup			
0-12		58.0	
13-18		43.0	
19-40		39.0	
41-60		39.0	
61-80		23.0	
Середній вік виживших/невиживших:			
Survived			
0		30.6	
1		28.3	
Аналіз часових рядів Apple (2014-10-05 – 2018-04-01):			
• Середня ціна акцій по закриттю (Close) за період: 127.5			
• Роки, коли торги починалися із ціни вище 110.2: 2017, 2018			
Для аналізу часових рядів використано дані про ціну акцій Apple.			
Середнє значення (mean) розраховано за формулою: Mean = $\Sigma(\text{Close}) / N$, де N – кількість днів у періоді.			
Для пошуку років, коли торги відкривалися в ціні вище 110.2, використано групування за роками в першою ціною 'Open' кожного року.			
Якщо перше значення 'Open' > 110.2 – рік вважається роком високого старту.			

Групування по статі та віку

Sex	Age	Не вижив	Вижив
female	0.75	0	2
female	1.00	0	2
female	2.00	4	2
female	3.00	1	1
female	4.00	0	5
female	5.00	0	4
female	6.00	1	1
female	7.00	0	1
female	8.00	1	1
female	9.00	4	0
female	10.00	1	0
female	11.00	1	0
female	13.00	0	2
female	14.00	1	3
female	14.50	1	0
female	15.00	0	4
female	16.00	1	5
female	17.00	1	5
female	18.00	5	8
female	19.00	0	7
female	20.00	2	0
female	21.00	3	4
female	22.00	2	10

Групування: клас, стать і вік

Pclass	Sex	Age	Count
1	female	2.00	1
1	female	14.00	1
1	female	15.00	1
1	female	16.00	3
1	female	17.00	2
1	female	18.00	3
1	female	19.00	3
1	female	21.00	2
1	female	22.00	4
1	female	23.00	2
1	female	24.00	5
1	female	25.00	1
1	female	26.00	1
1	female	29.00	1
1	female	30.00	5
1	female	31.00	2
1	female	32.00	1
1	female	33.00	3
1	female	35.00	6
1	female	36.00	3

Дані Titanic: ключові змінні та методи аналізу

Огляд основних характеристик пасажирів і підходів до вивчення виживання

 Інформація про пасажирів:

 Стать

 Вік

 Клас

 Виживання

 Аналіз виживання по:

 Статі

 Віку

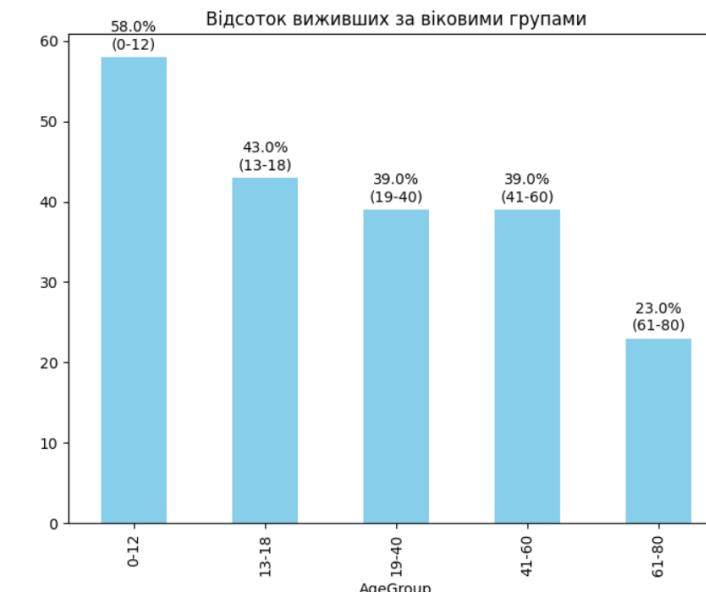
 Класу

 Вікових групах

 Візуалізація результатів:

 Гістограми

 Бар-чарти



Виживання за віковими групами

Аналіз фінансових даних Apple (2012–2018)

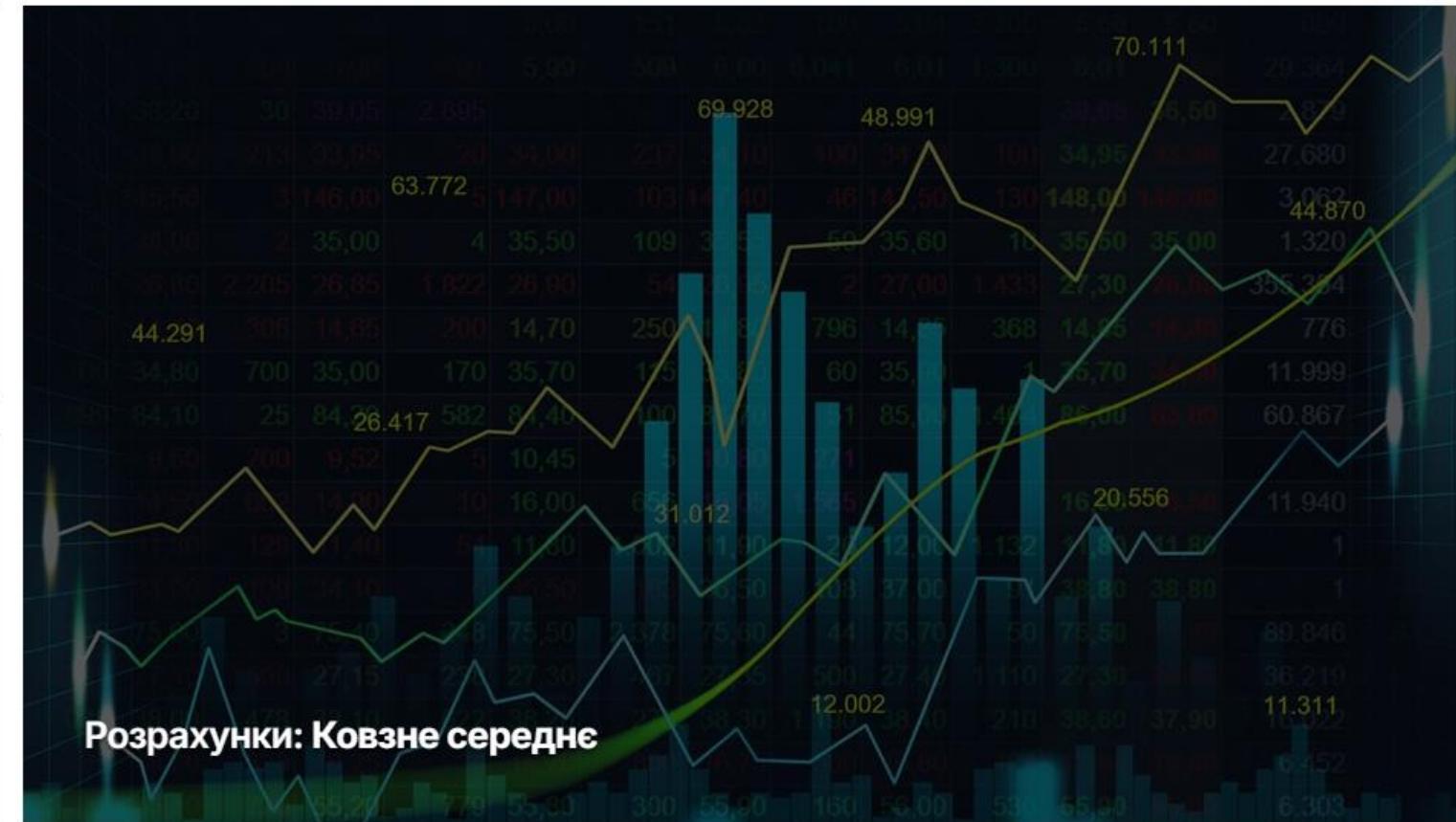
Ключові параметри та методи для щоденних цін акцій, розрахунків, прогнозів і візуалізації

Дані про щоденні ціни акцій:
Open, Close, High, Low

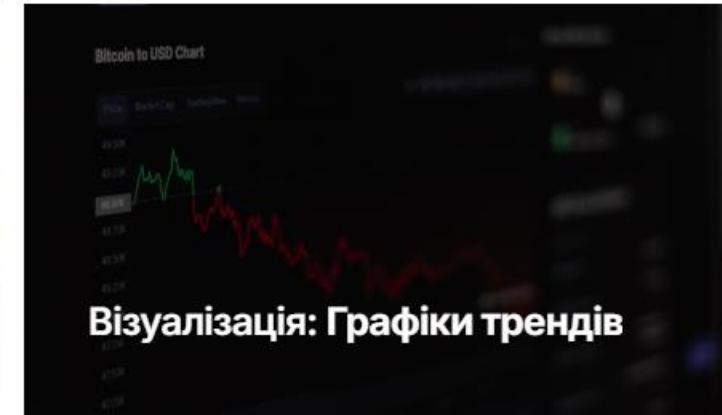


Розрахунки: Щоденна зміна

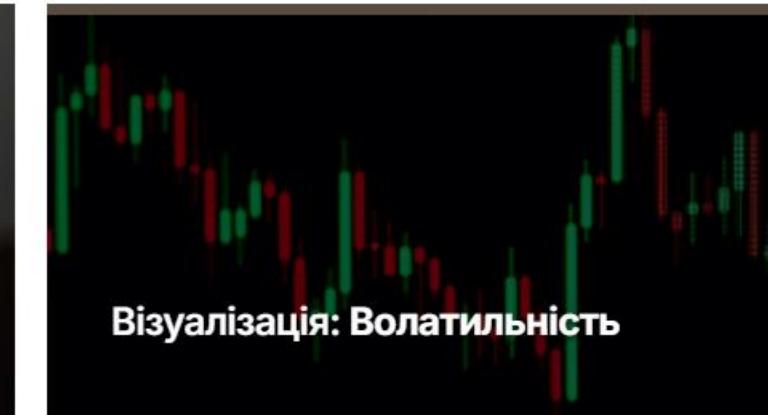
Прогноз: Ціни на 90 днів за
допомогою лінійної регресії



Розрахунки: Ковзне середнє



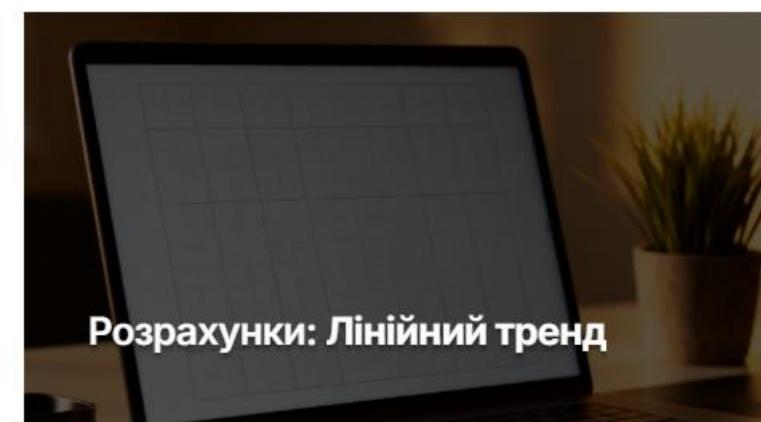
Візуалізація: Графіки трендів



Візуалізація: Волатильність



Розрахунки: Стандартне
відхилення



Розрахунки: Лінійний тренд



Візуалізація: Прогноз

Реалізовані функції:

Огляд ключових можливостей GUI та аналітики для роботи з даними Titanic та Apple



1 Табличний аналіз: Відсоток виживших по статі, віку, класу



2 Візуалізація даних: Графіки для Titanic та Apple



3 Прогнозування: Побудова лінійного тренду та прогнозу ціни акцій Apple



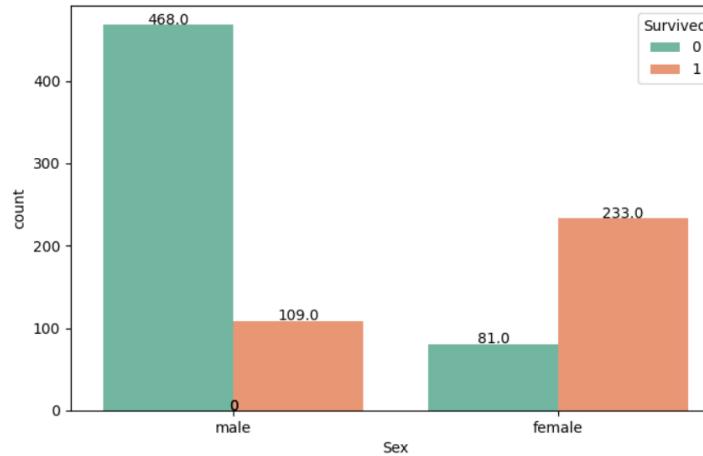
4 Інтерактивний GUI: Вкладки для таблиць та графіків



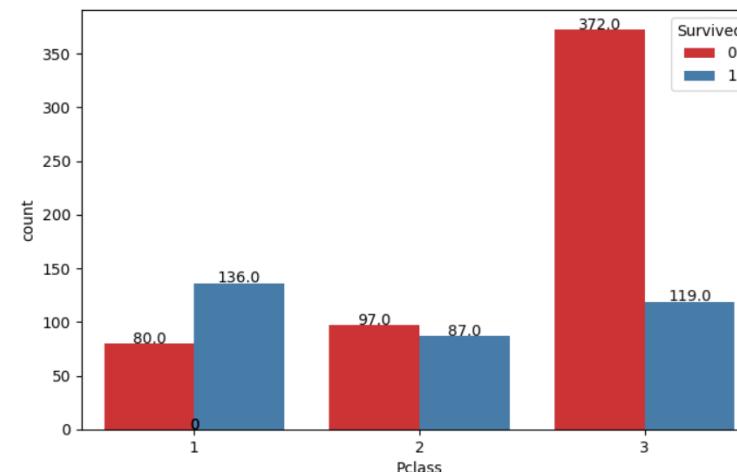
**5 Інтерактивний GUI:
Прокручувані текстові області**



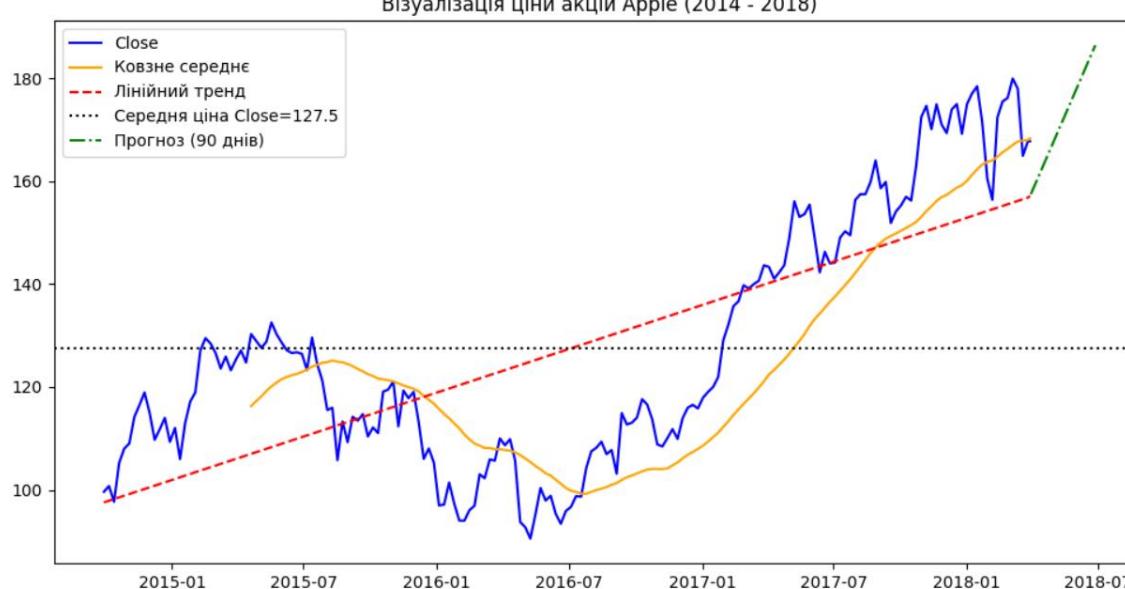
Волатильність акцій Apple



Виживання по статі



Виживання по класу



Apple: ціна, тренд, середня і прогноз



Apple: денна зміна ціни

Реалізаційні деталі: GUI, бібліотеки та аналіз

Ключові елементи інтерфейсу, використані бібліотеки та підходи до часових рядів



1 Tkinter GUI

- Вкладки для основної статистики, вікових груп, класів і графіків
- Прокручувані текстові області для таблиць
- Можливість поєднання графіків Matplotlib у GUI



2 Використані бібліотеки

- Pandas, Numpy, Seaborn, Matplotlib, Scikit-Learn



3 Аналіз часових рядів

- Ковзне середнє для згладження коливань
- Стандартне відхилення для оцінки волатильності
- Лінійний тренд та прогноз на основі регресії