

ПРАКТИЧНЕ ЗАВДАННЯ 2.

КЛАСИФІКАЦІЯ НАБОРІВ ДАНИХ

Завдання

1. Назвіть два різних набори двовимірних даних, які прямо або побічно пов'язані з вашою спеціальністю. У кожному разі визначте характер залежності між двома факторами і встановіть, чи можна і чи корисно мати можливість передбачати один фактор на підставі іншого.

2. Виконайте попередню задачу, але для багатовимірних даних.

3. Виберіть деяку організацію і назвіть дві кількісні змінні, що представляють інтерес для неї. Вкажіть для кожної змінної, чи є вона дискретною або безперервною.

4. Виберіть деяку організацію і назвіть дві якісні змінні, що представляють для неї інтерес. Для кожної змінної вкажіть, чи є вона номінальною або порядковою.

5. Вкажіть три часових ряди, що представляють інтерес особисто для вас. Для кожного ряду визначте:

- а) чи є часовий тренд?
- б) чи є сезонні впливи?

6. Виберіть деяку організацію і опишіть (в загальних термінах) базу даних цієї організації. У середині цієї бази даних визначте три набори даних різних видів. Для кожного з цих трьох наборів вкажіть, що є елементарною одиницею, і визначте, що можна дізнатися в результаті відповідного аналізу.

7. Ваша фірма вирішила подати до суду на ненадійного постачальника. Який вид аналізу можна виконати для оцінки втрат через втрачені можливості, виходячи з результатів конкурентів, економічної ситуації і пори року?

8. Визначте вид (первинні або вторинні) наступних даних.

а) Дані уряду України про поточну економічну ситуацію в кожній області, які використовуються фірмою, яка планує розширення.

б) Дані про собівартість продукції одного з підприємств вашої фірми, зібрані в ході кампанії по зниженню витрат.

в) Дані звіту по галузі, придбані вашою фірмою з метою оцінки свого місця серед конкуруючих фірм.

9. У табл. 1 міститься кілька об'єктів з бази даних співробітників. Інформація дана для 5 осіб станом на 3 квітня 2020 року.

а) Що є елементарною одиницею в цьому наборі даних?

б) Визначте вид даних: одномірні, двовимірні, багатовимірні?

в) Які з цих чотирьох змінних є якісними, а які – кількісними?

г) Які з змінних (якщо такі є) є порядковими якісними змінними?

Поясніть свою відповідь.

д) Це часовий ряд або це дані про один часовий зріз?

Таблиця 1. Дані про п'ятьох службовців

Стать	Зарплата, грн.	Освіта	Стаж, років
М	42300	Вища	9
Ж	31800	Молодший спеціаліст	4
М	29500	Магістр	2
Ж	58100	Магістр	15
Ж	36000	Молодший спеціаліст	7

10. Розглянемо набір даних з табл. 2, що містить інформацію про п'ять виробничих підрозділів (позначаються кодами).

а) Що є елементарною одиницею для цього набору даних?

б) Визначте вид даних: одномірні, двовимірні, багатовимірні?

в) Вкажіть якісні змінні (якщо вони є).

г) Чи є в цих даних порядкова змінна? Якщо так, то вкажіть її.

д) Це часовий ряд або дані про один часовий зріз?

Таблиця 2. Інформація про деякі види продукції п'яти виробничих підрозділів

Код	Деталь	Якість	Кількість працівників
A-235-86	Гальма	Гарна	53
W-186-74	Паливопровід	Відмінна	37
X-937-85	Радіо	Досить гарне	26
Z-447-91	Шасі	Чудова	85
P-258-89	Дріт	Гарна	16

11. У табл. 3 містяться дані про обсяги продажів і розміри доходу (в тисячах гривень); інформація за перші 6 місяців 2019 року.

- Що є елементарною одиницею для цього набору даних?
- Визначте вид даних: одномірні, двовимірні, багатовимірні?
- Яка з двох змінних є кількісною, а яка змінна є якісною?
- Це часовий ряд або дані про один часовий зріз?

Таблиця 3. Продажі і прибуток з січня по червень 2019 року

Продажі	Дохід (збитки)
350	30
270	23
140	(2)
280	14
410	53
390	47

12. Табл. 4 представляє собою частину бази даних про споживачів одного із співробітників відділу продажів.

- Що є елементарною одиницею в цьому наборі даних?
- Визначте вид даних: одномірні, двовимірні, багатовимірні?
- Визначте, які з цих змінних кількісні, а які якісні?
- Які з цих змінних номінальні? Які порядкові?
- Це часовий ряд або дані про один часовий зріз?

Таблиця 4. Деякі споживачі і покупки

Рівень інтересу до нових виробів	Загальний обсяг закупівель в минулому році, грн.	Регіон
Слабкий	88906	Захід
Помірний	396808	Південь
Дуже сильний	438442	Південь
Слабкий	2486	Центр
Слабкий	37375	Захід
Дуже сильний	2314	Центр
Помірний	1244096	Центр
Слабкий	857248	Південь
Сильний	119650	Центр
Помірний	711514	Захід
Слабкий	22616	Захід

13. Щоб спланувати обсяг витрат на рекламу в різних засобах масової інформації (телебачення, радіо, газети та ін.), ви вивчаєте набір даних, що містить минулорічні витрати кожного з ваших конкурентів на телерекламу, радіорекламу і на рекламу в газетах. Дайте повний опис типу для такого набору даних.

14. Обсяги квартальних продажів вашої фірми за останні п'ять років можуть бути корисні для стратегічного планування.

- а) Це часовий ряд чи дані про один часовий зріз?
- б) Визначте вид даних: одномірні, двовимірні, багатовимірні?

15. Розглянемо інформацію про пропоновану і запитувану ціну для 18 різних зобов'язань Казначейства України на момент закриття торгів в конкретний день.

- а) Визначте вид даних: одномірні, двовимірні, багатовимірні?
- б) Це часовий ряд чи дані про один часовий зріз?

16. Розглянемо дані продажів 35 компаній.

- а) Визначте вид даних: одномірні, двовимірні, багатовимірні?
- б) Це якісна або кількісна змінна?
- в) Це порядкова, номінальна або будь-яка інша змінна?

17. Інспектор з контролю якості оцінив кожен з вироблених сьогодні партій продукції за шкалою від А до Е, де А – вищий сорт, Е – нижчий.

- а) Яка це змінна: кількісна або якісна?
- б) Це порядкова, номінальна або будь-яка інша змінна?

18. Одна з колонок електронної таблиці містить назви компаній, що постачають комплектуючі.

- а) Яка це змінна: кількісна або якісна?
- б) Це порядкова, номінальна або будь-яка інша змінна?

19. У табл. 5 містяться рейтинги якості лез для бритв.

- а) Що є елементарною одиницею в цьому наборі даних?
- б) Це одномірні, двовимірні або багатовимірні дані?
- в) Це часовий ряд або дані про один часовий зріз?
- г) Змінна «тип» – кількісна, порядкова або номінальна?
- д) Змінна «ціна» – кількісна, порядкова або номінальна?
- е) Змінна «зручність» – кількісна, порядкова або номінальна?

Таблиця 5. Рейтинги якості лез для бритв

Виріб	Тип	Ціна, дол.	Зручність використання
Gillette Sensor Excel	ріжучий вузол	4,50	Відмінна
Schick Tracer	ріжучий вузол	3,34	Дуже добре
Bic Twin Select for Sensible Skin	одноразові	3,68	Гарна
Gillette Sensor for Women	ріжучий вузол	3,93	Відмінна
Schick Silk Effects for Women	ріжучий вузол	4,64	Дуже добре
Bic Twin Pastel	одноразові	2,03	Гарна

20. Припустимо, що набір даних включає змінну «тип цінних паперів», закодовану таким чином: 1 – звичайна акція; 2 – привілейована акція; 3 – зобов'язання; 4 – ф'ючерсний контракт; 5 – опціон. Це кількісна або якісна змінна?

21. Технологічність складання виробів оцінюється за такою шкалою: 1 – високотехнологічна збірка; 2 – технологічна; 3 – задовільна; 4 – складна; 5 – дуже складна. Це кількісна, порядкова або номінальна змінна?

22. Припустимо, що набір даних включає змінну «економічна організація», закодовану в такий спосіб: 1 = ФОП, 2 = один власник, 3 = товариство, 4 = акціонерна компанія. Це кількісна або якісна змінна?

23. У табл. 6 містяться дані про побутові пилососи.

Таблиця 6. Порівняння пилососів

Ціна, дол.	Вага, фунти	Якість	Тип
170	17	Добре	Жорсткий шланг
260	17	Відносно добре	М'який шланг; самохідний
100	21	Добре	М'який шланг
90	14	Добре	Жорсткий шланг
340	13	Відмінно	М'який шланг
120	24	Добре	М'який шланг; самохідний
130	17	Відносно добре	М'який шланг; самохідний

- а) Що є елементарною одиницею в цьому наборі даних?
- б) Це одномірні, двовимірні або багатовимірні дані?
- в) Які з змінних є якісними, а які кількісними?
- г) Для кожної якісної змінної в цьому наборі даних визначте її тип: порядкова або номінальна?
- д) Це часовий ряд або дані про один часовий зріз?

24. Компанія Dow Jones розраховує ряд індексів фондового ринку, які використовуються на Нью-йоркській фондовій біржі. Найбільш відомий індекс Dow Jones Industrial Average (DJIA), який розраховують на основі даних про акції 30 промислових компаній. Інший індекс, званий Dow Jones Transportation Average (DJTA), розраховують на основі даних про акції 20 компаній, які спеціалізуються на наданні транспортних послуг (залізничний, автомобільний, морський транспорт, авіалінії і ін.). У табл. 7 містяться дані про 20 транспортних компаній, які фігурують в DJTA.

- а) Що є елементарною одиницею в цьому наборі даних?
- б) Це одномірні, двовимірні або багатовимірні дані?
- в) Які з змінних є якісними, а які кількісними?
- г) Якщо в цьому наборі є якісні змінні, то які вони – порядкові чи номінальні?
- д) Це часовий ряд або дані про один часовий зріз?

25. Продовжимо працювати з Dow Jones Transportation Average (DJTA), про який йшлося в завданні 24. У табл. 8 містяться дані про щоденні

спостереження значень DJTA, про зміну значення індексу між двома послідовними спостереженнями і про процентну зміну DJTA між двома послідовними спостереженнями.

- а) Що є елементарною одиницею в цьому наборі даних?
- б) Це одномірні, двовимірні або багатовимірні дані?
- в) Які з змінних є якісними, а які кількісними?
- г) Якщо в цьому наборі є якісні змінні, то які вони – порядкові чи номінальні?
- д) Це часовий ряд або дані про один часовий зріз?

Таблиця 7. Ціна акції і місячна процентна зміна для компаній, що фігурують в DJTA

Компанія	Ціна на момент закінчення торгів на біржі, 30 вересня 2019 року	Процентне зміна ціни в порівнянні з 31 серпня 2019 року
Yellow Corp.	13,5	13,1
Southwest Airlines	20,0	12,3
CSX	42,1	11,4
UAL	64,8	7,5
Union Pacific	42,6	7,1
Ryder Systems	24,9	5,6
Burl. North. SF	32,0	3,2
Norfolk Southern	29,1	3,1
AMR	55,4	1,7
GATXCorp.	33,1	0,2
Delta Airlines	97,3	-4,7
CNF	29,1	-6,8
FDXCorp.	45,1	-9,9
Airborne Freight	17,3	-11,2
US Freightways	19,9	-11,4
Alaska Air	34,1	-12,5
US Airways	50,6	-13,1
Alex & Baldwin	19,9	-15,4
Xtra Corp.	46,6	-20,9
Roadway Express	11,0	-27,3

Таблиця 8. Щоденні значення і зміни DJTA у вересні 2019 року

Дата	DJTA	Зміна значення	Процентна зміна
31 серпня	2752,51	–	–
1 вересня	2753,77	1,26	0,05
2 вересня	2702,55	-51,22	-1,86
3 вересня	2624,47	-78,08	-2,89
4 вересня	2616,75	-7,72	-0,29
8 вересня	2749,30	132,55	5,07
9 вересня	2691,20	-58,10	-2,11
10 вересня	2631,51	-59,69	-2,22
11 вересня	2679,17	47,66	1,81
14 вересня	2805,14	125,97	4,70
15 вересня	2863,70	58,56	2,09
16 вересня	2884,13	20,43	0,71
17 вересня	2813,99	-70,14	-2,43
18 вересня	2814,67	0,68	0,02
21 вересня	2802,64	-12,03	-0,43
22 вересня	2865,29	62,65	2,24
23 вересня	2904,10	38,81	1,35
24 вересня	2831,24	-72,86	-2,51
25 вересня	2795,83	-35,41	-1,25
28 вересня	2803,10	7,27	0,26
29 вересня	2746,13	-56,97	-2,03
30 вересня	2644,67	-101,46	-3,69

26. Зверніться до бази даних про співробітників фірми, що міститься в Табл. 9.

1. Опишіть і класифікуйте цю базу даних і її частини.
 - а) Це одномірні, двовимірні або багатовимірні дані?
 - б) Що є елементарною одиницею в цьому наборі даних?
 - в) Які змінні якісні, а які кількісні?
 - г) Змінна «рівень підготовки» є порядковою чи номінальною? Чому?
 - д) Чи можна виконувати арифметичні дії зі змінною «номер службовця»? Про що це свідчить з точки зору того, чи є ця змінна дійсно кількісною?
 - е) Це часовий ряд або дані про один часовий зріз?
2. Для кожної змінної з Табл. 9 визначте, які із зазначених нижче операцій можна застосовувати до цієї змінної.
 - а) Арифметичні (додавання, віднімання і т.д.).

- б) Підрахунок кількості службовців в кожній категорії.
- в) Ранжування по порядку.
- г) Обчислення відсотка службовців в кожній категорії.

Проекти

1. Знайдіть в Інтернет статті з соціологічного журналу або в газеті таблицю даних. Скопіюйте цю статтю і таблицю.

- а) Класифікуйте набір даних відповідно до кількості змінних.
- б) Що виступає в якості елементарної одиниці?
- в) Це часовий ряд або дані про один часовий зріз?
- г) Визначте тип кожної змінної.
- д) Вкажіть, які операції можна застосовувати до кожної з змінних.
- е) Сформулюйте (в загальних термінах), на які питання, пов'язані з бізнесом, можна знайти відповіді при детальному аналізі набору даних такого типу.

Таблиця 9. База даних працівників

Номер співробітника ¹	Заробітна плата за рік (\$)	Стать	Вік (Років)	Досвід (Років)	Рівень підготовки ²
1	\$ 32 368,00	Жіноча	42	3	В
2	\$ 53 174,00	Чоловіча	54	10	В
3	\$ 52 722,00	Чоловіча	47	10	А
4	\$ 53 423,00	Чоловіча	47	1	В
5	\$ 50 602,00	Чоловіча	44	5	В
6	\$ 49 033,00	Чоловіча	42	10	А
7	\$ 24 395,00	Чоловіча	30	5	А
8	\$ 24 395,00	Жіноча	52	6	А
9	\$ 43 124,00	Чоловіча	48	8	А
10	\$ 23 975,00	Жіноча	58	4	А
11	\$ 53 174,00	Чоловіча	46	4	С
12	\$ 58 515,00	Чоловіча	36	8	С
13	\$ 56 294,00	Чоловіча	49	10	В
14	\$ 49 033,00	Жіноча	55	10	В
15	\$ 44 884,00	Чоловіча	41	1	А

¹ Ці числа вказані з єдиною метою – надати кожному працівнику свій унікальний номер.

² Періодично і в добровільному порядку працівнику пропонують пройти курс підготовки (це не є обов'язковою вимогою). Працівники, що не пройшли підготовку, отримують кваліфікацію «А», після проходження одного курсу працівник отримує кваліфікацію «В», після другого і кінцевого курсу – кваліфікацію «С».

16	\$ 53 429,00	Жіноча	52	5	В
17	\$ 46 574,00	Чоловіча	57	8	А
18	\$ 58 968,00	Жіноча	61	10	В
19	\$ 53 174,00	Чоловіча	50	5	А
20	\$ 53 627,00	Чоловіча	47	10	В
21	\$ 49 033,00	Чоловіча	54	5	В
22	\$ 54 981,00	Чоловіча	47	7	А
23	\$ 6 253,00	Чоловіча	50	10	В
24	\$ 27 525,00	Жіноча	38	3	А
25	\$ 24 395,00	Чоловіча	31	5	А
26	\$ 56 884,00	Чоловіча	47	10	А
27	\$ 52 111,00	Чоловіча	56	5	А
28	\$ 44 183,00	Жіноча	38	5	В
29	\$ 24 967,00	Жіноча	55	6	А
30	\$ 35 423,00	Жіноча	47	4	А
31	\$ 41 188,00	Жіноча	35	2	В
32	\$ 27 525,00	Жіноча	35	3	А
33	\$ 35 018,00	Чоловіча	39	1	А
34	\$ 44 183,00	Чоловіча	41	2	А
35	\$ 35 423,00	Чоловіча	44	1	А
36	\$ 49 033,00	Чоловіча	53	8	А
37	\$ 40 741,00	Чоловіча	47	2	А
38	\$ 49 033,00	Чоловіча	42	10	А
39	\$ 56 294,00	Жіноча	44	6	С
40	\$ 4 718,00	Жіноча	45	5	С
41	\$ 46 574,00	Чоловіча	56	8	А
42	\$ 52 722,00	Чоловіча	38	8	С
43	\$ 51 237,00	Чоловіча	58	2	В
44	\$ 53 627,00	Чоловіча	52	8	А
45	\$ 53 174,00	Чоловіча	54	10	А
46	\$ 56 294,00	Чоловіча	49	10	В
47	\$ 49 033,00	Жіноча	53	10	В
48	\$ 49 033,00	Чоловіча	43	9	А
49	\$ 55 549,00	Чоловіча	35	8	С
50	\$ 51 237,00	Чоловіча	56	1	С
51	\$ 352,00	Жіноча	38	1	В
52	\$ 50 175,00	Жіноча	42	5	А
53	\$ 24 352,00	Жіноча	35	1	А
54	\$ 27 525,00	Жіноча	40	3	А
55	\$ 29 606,00	Жіноча	34	4	В
56	\$ 24 352,00	Жіноча	35	1	А
57	\$ 4 718,00	Жіноча	45	5	В
58	\$ 49 033,00	Чоловіча	54	10	А
59	\$ 53 174,00	Чоловіча	47	10	А
60	\$ 53 429,00	Жіноча	45	7	В
61	\$ 53 627,00	Чоловіча	47	10	А

62	\$ 26 491,00	Жіноча	46	7	A
63	\$ 42 961,00	Чоловіча	36	3	B
64	\$ 53 174,00	Чоловіча	45	5	A
65	\$ 37 292,00	Чоловіча	46	0	A
66	\$ 37 292,00	Чоловіча	47	1	A
67	\$ 41 188,00	Жіноча	34	3	B
68	\$ 57 242,00	Жіноча	45	7	C
69	\$ 53 429,00	Жіноча	44	6	C
70	\$ 53 174,00	Чоловіча	50	10	B
71	\$ 44 138,00	Жіноча	38	2	B