

29	1	4	4	5	3	1	2
30	1	3	3	7	5	1	2

### Завдання для виконання

1. Описати необхідні змінні і ввести дані.
2. Розрахувати **частотний розподіл** для змінної «*стать*», «*використання Інтернет для покупок*» і «*використання Інтернет для банківських операцій*».
3. Розрахувати **статистику** – середнє число годин на тиждень, витрачене користувачами на *використання Інтернет*, а також моду, медіану, стандартне відхилення і мінімальне і максимальне значення для цієї ознаки.
4. Інтерпретувати отримані дані.
5. Зберегти дані і результати розрахунків в особистій теці (мають бути два файли).

#### Вимоги до звіту

Звіт має містити:

- відповіді на контрольні питання;
- файли з результатами розрахунків.

### Контрольні запитання

1. Як відкрити файл з даними в SPSS для Windows і як зберегти результати розрахунків у файлі?
2. Що таке частотний розподіл змінної і за допомогою яких команд меню він будується?
3. Як розрахувати для вибраної змінної моду, медіану, середнє, дисперсію, стандартне відхилення?
4. Яке призначення кнопок у головному діалоговому вікні?
5. Як повернутися з вікна перегляду результатів SPSS «Viewer» до вікна редактора даних SPSS «Data Editor»? Вказати 3 способи.

6. Як відмовитися від показу частотних таблиць у вікні перегляду результатів?
7. Як побудувати гістограму SPSS для Windows?
8. Як відредагувати графік?
9. Як зберегти побудований графік у файлі?
10. Назвіть основні елементи вікна редактора діаграм.

## ЛАБОРАТОРНА ЗАНЯТТЯ № 6

**Тема:** *Використання критерію Пірсона  $\chi^2$  для перевірки узгодженості розподілів за допомогою MS Excel і SPSS для Windows*

**Мета:** навчитися представляти концентровані дані для оброблення в Excel і SPSS для Windows, будувати для них таблиці зв'язаності й обчислювати критерій Пірсона  $\chi^2$ .

### Постановка завдання

Дослідників у сфері інформаційних технологій цікавить, чи залежить обсяг використання інтернету від сімейного стану людей різної статі.

У таблицях 6.1 і 6.2 наведені дані для чоловіків і жінок відповідно.

*Таблиця 6.1*

### Використання інтернету чоловіками

Використання інтернету	Сімейний стан	
	Одружений	Неодружений
Багато	32	12
Мало	70	19
Не використовує	5	5
Підсумки	107	36

**Використання інтернету жінками**

Використання інтернету	Сімейний стан	
	У шлюбі	Не у шлюбі
Багато	24	31
Мало	75	20
Не використовує	7	5
Підсумки	106	56

За допомогою критерію хі-квадрат визначити, чи є зв'язок між сімейним станом і використанням інтернету для чоловіків і жінок.

**Розв'язання за допомогою електронних таблиць EXCEL****Хід роботи:**

Алгоритм розрахунку  $\chi^2$

1. Занести в таблицю найменування розрядів ознаки (стовпець А).
2. Занести емпіричні частоти, що відповідають їм, для одружених і неодружених чоловіків (стовпці В і С).
3. Підрахувати підсумки в рядках і стовпцях (стовпець D).
4. Обчислити теоретичні частоти для одружених і неодружених чоловіків як відношення добутку підсумків у рядку і підсумків у стовпчику до загальної кількості об'єктів (стовпці Е і F).
5. Для розрахунку  $\chi^2$  емп. скористатися функціями ХИ2ТЕСТ і ХИ2ОБР. Функції ХИ2ТЕСТ передаються два параметри: інтервал з емпіричними частотами і інтервал з теоретичними частотами. Функції ХИ2ОБР передаються два параметри: результат, отриманий в ХИ2ТЕСТ, і число ступенів свободи, що визначається за формулою :  $n = (k - 1) * (l - 1)$ , де  $k$  – кількість розрядів ознаки або рядків у таблиці (для нашого прикладу  $k=3$ ),  $l$  – кількість стовпців у таблиці (для нашого прикладу 2).

6. За таблицями критичних значень розподілу  $\chi^2$  для цього числа ступенів свободи визначити  $\chi^2$  критич.

Якщо  $\chi^2_{\text{емп.}} < \chi^2_{\text{критич.}}$ , то розбіжності між розподілами статистично недостовірні (розподіли узгоджуються між собою), а, отже, одружені та неодружені чоловіки не розрізняються за рівнем використання інтернету.

Формули для розрахунків наведені в табл. 6.3, результати розрахунків у таблиці 6.4.

Таблиця 6.3

**Таблиця з формулами**

	A	B	C	D	E	F
1	<b>Зв'язок між використанням інтернету і сімейним станом для чоловіків</b>					
2	Використання інтернету	Емпірична частота		Підсумки	Теоретична частота	
3		Сімейний стан			Сімейний стан	
4		Одружений	Неодружений		Одружений	Неодружений
5	Багато	32	12	=СУМ(B5: C5)	=\$D5*B8/\$D\$8	=\$D5*C8/\$D\$8
6	Мало	70	19	=СУМ(B6: C6)	=\$D6*B8/\$D\$8	=\$D6*C8/\$D\$8
7	Не використує	5	5	=СУМ(B7: C7)	=\$D7*B8/\$D\$8	=\$D7*C8/\$D\$8
8	Підсумки	=СУМ(B5: B7)	=СУМ(C5: C7)	=СУМ(D5: D7)	=СУМ(E5: E7)	=СУМ(F5: F7)
9						
10						
11	хи2тест		=ХИ2ТЕСТ(B5: C7;E5: F7)			
12	<i>хі-квадрат</i>		=ХИ2ОБР(C11;2)			

Таблиця 6.4

**Таблиця з результатами:**

	A	B	C	D	E	F
1	<b>Зв'язок між використанням інтернету і сімейним станом для чоловіків</b>					
2	Використання	Емпірична частота		Підсумки	Теоретична частота	

3	інтернету	Сімейний стан		44	Сімейний стан	
4		Одружений	Неодружений		Одружений	Неодружений
5	Багато	32	12	44	32,9	11,1
6	Мало	70	19	89	66,6	22,4
7	Не використовує	5	5	10	7,5	2,5
8	Підсумки	107	36	143	107	36
9						
10						
11	хі-тест		0,131			
12	<i>хі-квадрат</i>		4,066			

Аналогічно розв'яжіть завдання про використання інтернету жінками.

### Розв'язання задачі в SPSS

#### Хід роботи:

1. Визначити в редакторі 3 змінні: Використання інтернету (1 - багато, 2 - мало, 3 - не використовує), Сімейний стан (1 - одружений, 2 - неодружений), - freq (частота) - вказує частоту кожного поєднання.

2. Ввести дані для заданих змінних згідно наступної таблиці 6.5.

Таблиця 6.5

Використання інтернету	Сімейний стан	freq
Багато	Одружений	32
Мало	Одружений	70
Не використовує	Одружений	5
Багато	Неодружений	12
Мало	Неодружений	19
Не використовує	Неодружений	5

3. Зважити дані, вибравши меню *Data* (Дані) *Weight Cases* (Зважити випадки).

У діалоговому вікні *Weight Cases* (рис. 6.1) *вибрати опцію Weight Cases by і перенести змінну freq в полі Frequency Variable*. Закрити діалогове вікно.

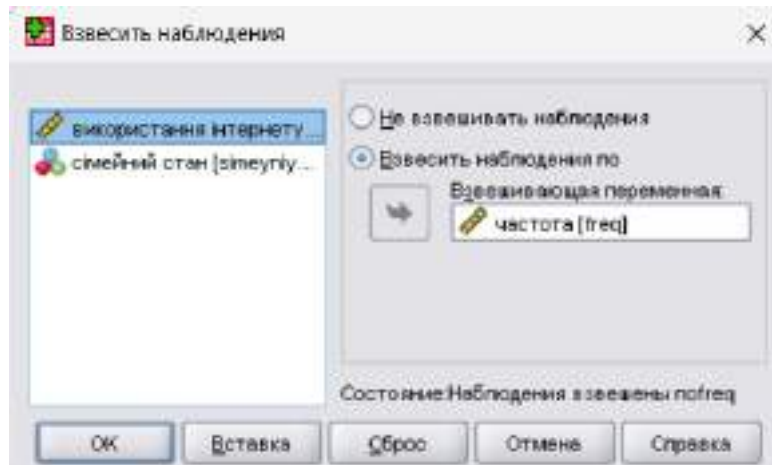


Рис. 6.1. Діалогове вікно «*Weight Cases*»

4. Побудувати таблиці зв'язаності, вибравши меню *Analyze (Аналіз), Descriptive statistics (Описова статистика), Crosstabs* (Таблиці зв'язаності) (рис. 6.2) :

- Перенести змінну «**Використання інтернету**» в список змінних рядків.
- Змінну «**Сімейний стан**» - в список змінних стовпців.

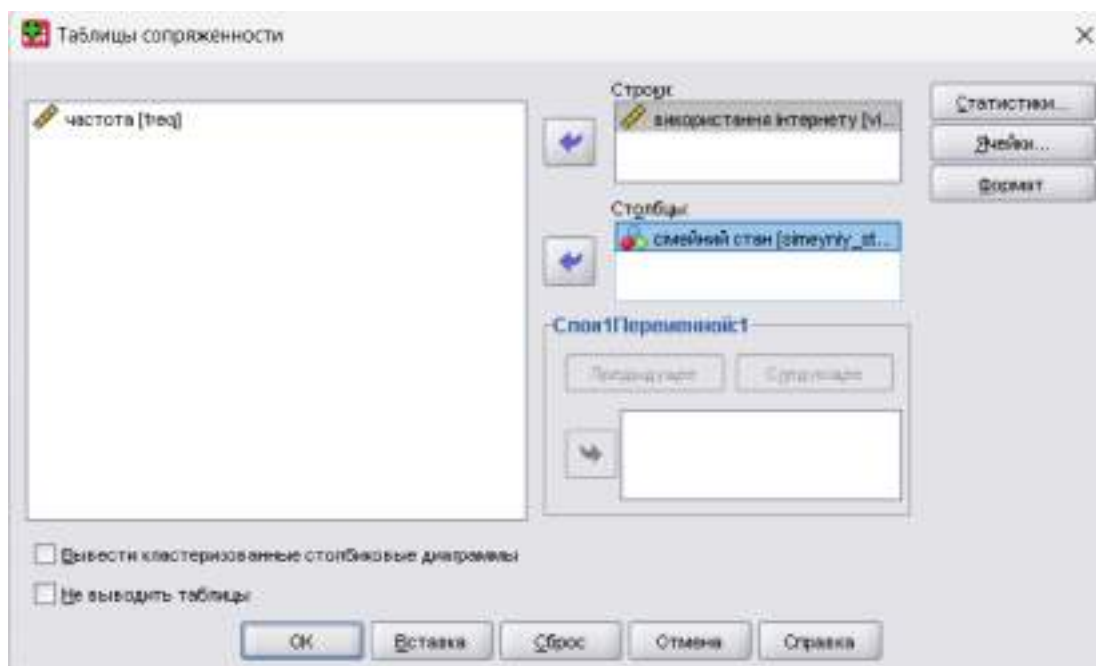


Рис. 6.2. Діалогове вікно «*Crosstabs*»

- Натиснути кнопку *Statistics*.

– У діалоговому вікні **Crosstabs: Statistics ...** (Частоти: Статистика) клацнути на опції **Chi-square** і потім на кнопці **Continue**(Далі) (рис. 6.3).

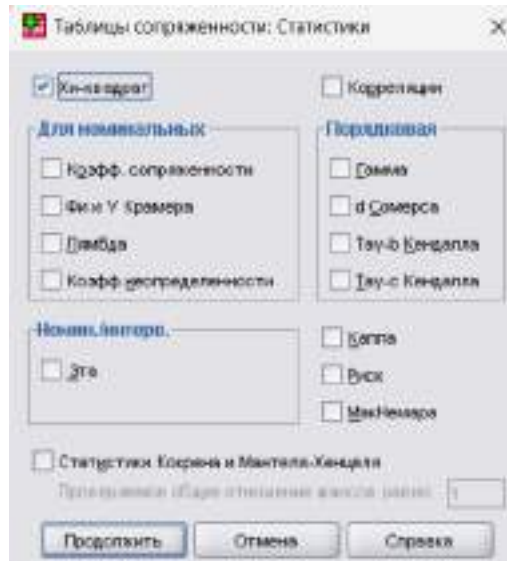


Рис. 6.3. Діалогове вікно «Crosstabs: Statistics»

– Натиснути на кнопці **Cell** (комірка). У групі **Counts** встановити опцію спостережувані і очікувані частоти. У групі **Percentage** поставити галочку біля опції **Column** (Відсоток по колонці). Натиснути **Continue** ОК (рис. 6.4).

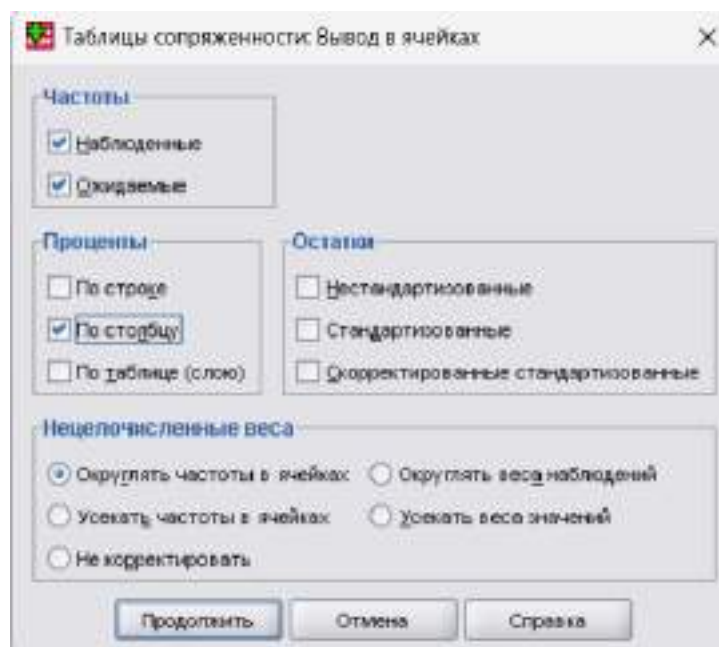


Рис. 6.4. Діалогове вікно «Crosstabs: Cell Display»

5. Розглянути і проаналізувати результати у вікні виведення (таблиця. 6.6 - 6.7).

### ➔ Таблицы сопряженности

Сводка обработки наблюдений

	Наблюдения					
	Валидные		Пропущенные		Итого	
	N	Процент	N	Процент	N	Процент
використання інтернету * сімейний стан	143	100,0%	0	,0%	143	100,0%

Таблица сопряженности использования интернета \* семейный стан

			сімейний стан		Итого
			одружений	не одружений	
використання інтернету	"багато"	Частота	32	12	44
		Ожидаемая частота	32,9	11,1	44,0
		% в сімейний стан	29,9%	33,3%	30,8%
	"мало"	Частота	70	19	89
		Ожидаемая частота	66,6	22,4	89,0
		% в сімейний стан	65,4%	52,8%	62,2%
	не використовує	Частота	5	5	10
		Ожидаемая частота	7,5	2,5	10,0
		% в сімейний стан	4,7%	13,9%	7,0%
Итого	Частота	107	36	143	
	Ожидаемая частота	107,0	36,0	143,0	
	% в сімейний стан	100,0%	100,0%	100,0%	

Таблиця 6.7

Критерии хи-квадрат

	Значение	ст.св.	Асимпт. значимость (2-стор.)
Хи-квадрат Пирсона	4,068 <sup>a</sup>	2	,131
Отношение правдоподобия	3,649	2	,161
Линейно-линейная связь	,279	1	,597
Кол-во валидных наблюдений	143		

a. В 1 (16,7%) ячейках ожидаемая частота меньше 5. Минимальная ожидаемая частота равна 2,52.

Аналогічно **самостійно** розв'язати завдання **про використання інтернету жінками**.

### Завдання для самостійного виконання

1. За даними таблиці 6.8 визначте, чи існує залежність між віком людини і її витратами на навчання.

**Залежність між віком людини і її витратами на навчання**

Вік	Рівень витрат на навчання			Всього
	Високий	Середній	Низький	
Молоді	25	40	31	
Середнього віку	52	63	47	
Літні	12	23	58	
Всього				

2. За даними таблиці 6.9 перевірити припущення про те, що частота звертань до телефонної служби довіри нерівномірно розподіляється по днях тижня (за допомогою критерію  $\chi^2$  порівняти заданий емпіричний розподіл рівномірним).

**Розподіл звертань до телефонної служби довіри по днях тижня**

День тижня	Число звернень
Понеділок	9
Вівторок	6
Середа	7
Четвер	6
П'ятниця	9
Субота	15
Неділя	18
Всього	70

**Вимоги до звіту**

Звіт має містити:

- відповіді на контрольні питання;

- файл з результатами розрахунків.

### Контрольні запитання

1. Призначення критерію Пірсона  $\chi^2$ .
2. Обмеження критерію Пірсона  $\chi^2$ .
3. Що таке емпіричні частоти?
4. Що таке теоретичні частоти?
5. Як обчислюється критерій Пірсона  $\chi^2$  в Excel?
6. Як представити дані для розрахунків критерію Пірсона  $\chi^2$  за допомогою SPSS?
7. Дайте змістовну інтерпретацію результатам, отриманим у розв'язаних завданнях.

### ЛАБОРАТОРНЕ ЗАНЯТТЯ № 7

**Тема:** *Обчислення коефіцієнта кореляції Пірсона-Брава для метричних шкал*

**Мета:** навчитися обчислювати коефіцієнт кореляції Пірсона, будувати діаграму розсіювання, визначати значущість отриманого коефіцієнта.

#### Короткі теоретичні відомості

Для визначення кореляційної залежності між двома змінними, заданими метричними шкалами, слугує *коефіцієнт Пірсона-Брава*. Він характеризує міру лінійного зв'язку між змінними.

Нехай  $X$  і  $Y$  дві змінних, що корелюють між собою. Коефіцієнт Пірсона-Брава обчислюється за формулою:

$$r = \frac{\sum_{i=1}^n Z_i^x \circ Z_i^y}{n-1}, \text{ где} \quad (7.1)$$

$$Z_i^x = \frac{X_i - X_{cp}}{S_x}, \text{ где} \quad (7.2)$$

$S_x$  - стандартне відхилення,

$X_i$  –  $i$ -значення ознаки,

$X_{cp}$  - середнє значення ознаки  $X$ .

$$Z_i^y = \frac{Y_i - Y_{cp}}{S_y}, \text{ где} \quad (7.3)$$

$S_y$  - стандартне відхилення,

$Y_i$  –  $i$ -значення,

$Y_{cp}$  - середнє значення ознаки  $Y$ .

### Постановка завдання

У таблиці 7.1 наведені дані про кількість студентів, що щодня відвідують лекції і лабораторні роботи в університеті.. Знайдіть і інтерпретуйте кореляцію між цими двома величинами.

Таблиця 7.1

#### Кількість студентів, які відвідують заняття в університеті

<i>День тижня</i>	<i>Лекції</i>	<i>Лабораторні</i>
Понеділок	140	180
Вівторок	120	200
Середовище	130	190
Четвер	110	210
П'ятниця	160	160
Субота	135	185

### Розв'язання за допомогою Excel

#### Хід роботи:

1. Ввести дані.

2. Обчислити коефіцієнт Пірсона для цих даних спочатку за формулами, а потім виконати теж саме за допомогою функції КОРРЕЛ (рис. 7.1).

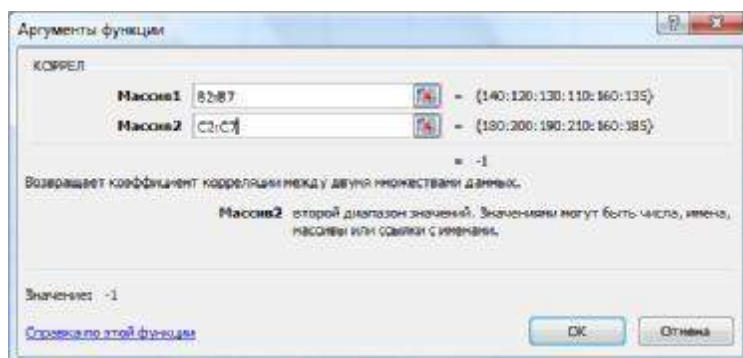


Рис. 7.1. Діалогове вікно «Аргументи функції КОРРЕЛ»

3. Заповнити в зошиті таблицю 7.1 (внести розрахункове і критичне значення коефіцієнта, визначити його значущість по таблиці критичних значень).

Таблиця 7.1

Об'єм вибірки	Емпіричне значення коефіцієнта	Критичне значення	Резюме (означаємо/незначимий)

4. Побудувати діаграму розсіювання (виділити всі дані (без рядка заголовку), викликати майстер діаграм, вибрати тип діаграми - «точкова», підписати усі заголовки) (див. рис. 7.2).

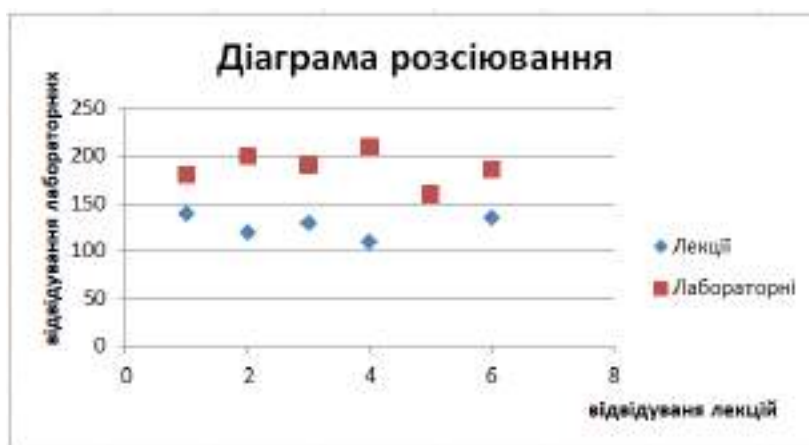


Рис. 7.2. Діаграма розсіювання, побудована за даними таблиці 7.1.

## Розв'язання за допомогою SPSS

### Хід роботи:

1. Створити файл даних в SPSS: описати дві змінні (Лекції – кількість студентів, які відвідують лекції, Laboratorni - кількість студентів, які відвідують лабораторні).
2. Ввести або скопіювати з файлу Excel дані.
3. Обчислити коефіцієнт Пірсона для цих даних: Меню: *Analyze* (Аналіз), *Correlate* (Кореляція), *Bivariate* (Двовимірна).

У вікні *Bivariate Correlation* (рис. 7.3) задати необхідні параметри.

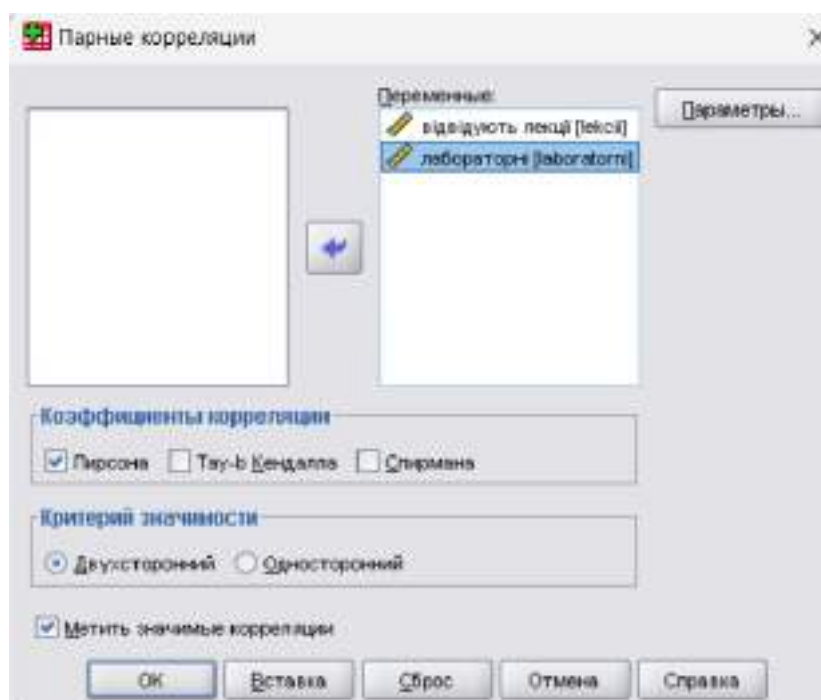


Рис. 7.3. Діалогове вікно «Bivariate Correlation»

4. Ознайомитися з отриманими результатами у вікні перегляду (рис. 7.4). Результати містять: коефіцієнт кореляції (*Pearson Correlation*), рівень значущості (*Sig*), об'єм вибірки (*N*). Записати їх у зошит. Визначити, чи значимий коефіцієнт кореляції.

**Корреляции**

		відвідують лекції	лабораторні
відвідують лекції	Корреляция Пирсона	1	$-1,000^{**}$
	Знач.(2-сторон)		,000
	N	6	6
лабораторні	Корреляция Пирсона	$-1,000^{**}$	1
	Знач.(2-сторон)	,000	
	N	6	6

\*\*. Корреляция значима на уровне 0.01 (2-сторон.).

Рис. 7.4. Вікно перегляду для прикладу з відвідуванням занять

### 5. Побудувати діаграму розсіювання:

Меню: **Graphs** (Графіки), **Scatter...**(Діаграма розсіювання), **Simple** (Проста) (рис. 7.5).

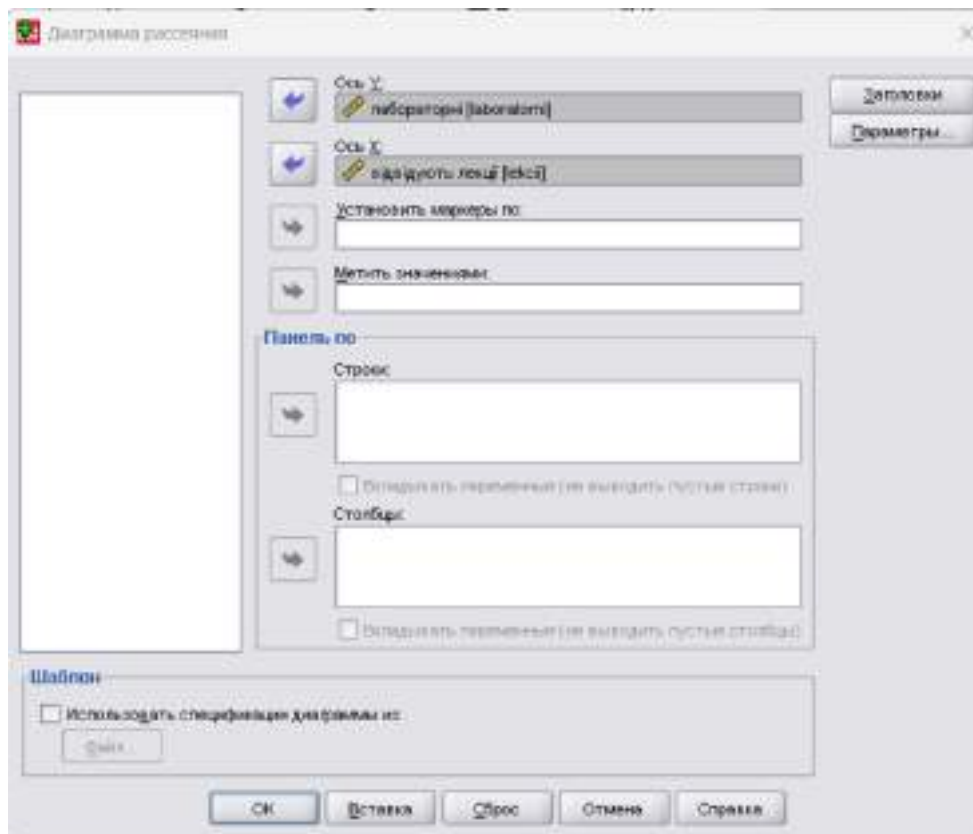


Рис. 7.5. Діалогове вікно «Діаграма розсіювання» (проста точкова діаграм)

Розглянути отриману діаграму розсіювання (рис. 7.6).

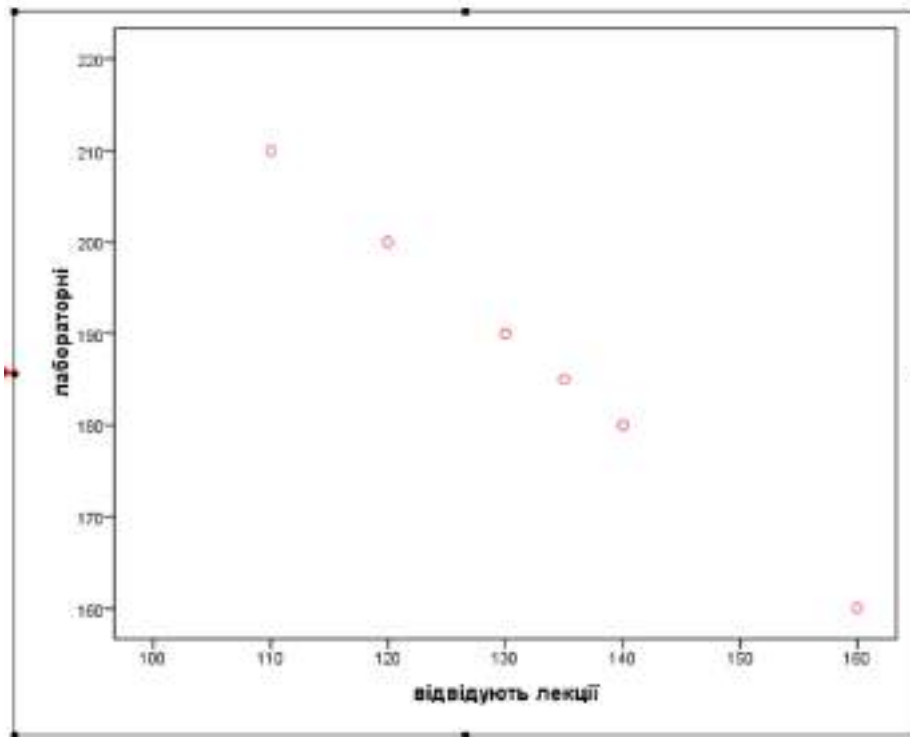


Рис. 7.6. Точкова діаграма для прикладу з відвідуванням занять

6. Зберегти дані розрахунків і початкові дані у файлах.
7. Зробити висновок про наявність або відсутність значимої кореляції між двома ознаками.

### Завдання для самостійного виконання

1. У таблиці 7.2 містяться дані про щоденний об'єм продажів в ум.один. лазерних принтерів, картриджів до принтерів і шкільного приладдя. Знайдіть і інтерпретуйте кореляцію між цими величинами. Розв'яжіть завдання в середовищі Microsoft Excel (+ побудувати діаграму) і в середовищі SPSS (+ побудувати 3 діаграми (попарно: лазерні + картриджі, лазерні + шкільне, картриджі + шкільне), в яких потрібно застосувати різні кольори).

**Щоденний об'єм продажів (ум.один)**

День	Лазерні принтери	Картриджі	Шкільне приладдя
1	17291	4379	3618
2	13734	2258	3514
3	18802	4206	3587
4	12171	2137	3007
5	16402	3775	3850
6	19904	4781	3675
7	14023	1991	3120
8	17513	2663	3345
9	17847	4451	2045
10	12718	1648	3292
11	12292	2342	3405
12	11846	2646	2799
13	17088	3216	2417
14	13523	2184	2405

**Вимоги до звіту**

Звіт має містити:

- відповіді на контрольні питання;
- файл з результатами розрахунків.

**Контрольні запитання**

1. Дайте визначення функціональної залежності.
2. Дайте визначення кореляційної залежності.
3. Які значення можуть приймати коефіцієнти кореляції?
4. Як вчислити коефіцієнт кореляції Пірсона-Браве з допомогою

електронних таблиць?

5. Для яких шкал він використовується?
6. Як побудувати діаграму розсіювання?
7. Які параметри задаються функції КОРРЕЛ?
8. Як розрахувати коефіцієнт кореляції Пірсона-Брава і побудувати діаграму розсіювання в SPSS?
9. Як визначити значущість коефіцієнта кореляції в SPSS?
10. Дайте змістовну інтерпретацію результатам розрахунків.

## ЛАБОРАТОРНЕ ЗАНЯТТЯ № 8

**Тема:** *Обчислення коефіцієнта рангової кореляції Спірмена*

**Мета:** навчитися обчислювати коефіцієнт рангової кореляції Спірмена і визначати його значущість.

### Короткі теоретичні відомості

Нехай  $N$  об'єктів можуть бути впорядковані як за ознакою  $X$ , так і по  $Y$ .

Позначимо

$R_{ix}$  - ранг  $i$ - об'єкту за ознакою  $X$

$R_{iy}$  - ранг  $i$ - об'єкту за ознакою  $Y$

$d_i = R_{ix} - R_{iy}$  - міра неспівпадіння рангів

Для визначення кореляційної залежності між двома змінними, заданими ранговими шкалами слугує коефіцієнт рангової кореляції Спірмена, який обчислюється за формулою:

$$\rho = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)} \quad (8.1)$$

### Постановка завдання

Дослідження, проведене на британському ринку, виявило наступну

значущість чинників успіху нових продуктів, вироблених японськими і британськими фірмами (див. таблиця 8.1, в якій *n* означає число опитаних фірм).

З таблиці 8.1 видно, що найчастіше ключовим чинником успіху називали міру адаптації до вимог покупців. І тільки відносно невелика доля фірм (1/4), як японських, так і британських, вважає ефективний маркетинг важливим критерієм успіху.

Порівняння оцінок, отриманих від японських і британських фірм, виявляє, що перші з них надають більше значення наявності конкурентної переваги, а другі - більшій адаптації до запитів споживачів.

Таблиця 8.1

**Чинники успіху нових продуктів, виявлених на британському ринку**

Чинники успіху	Відсоток фірм	
	Японських (n=116)	Британських (n=86)
Хороша адаптація до потреб	69,8	75,6
Перевага над конкурентами за якістю	79,3	59,3
Перевага над конкурентами по дизайну	69,9	45,3
Перевага над конкурентами по співвідношенню гідність/ціна	58,6	61,6
Перевага над конкурентами по конструкції	55,2	48,8
Дуже конкурентна ціна	41,4	27,9
Адаптація	39,7	34,9
Унікальність	36,2	29,1
Ефективний маркетинг	27,6	25,6
Глибокий аналіз ринку	27,6	18,6
Великий об'єм ринку	20,7	16,3
Синергія виробництво/маркетинг	16,4	18,6
Ухилення від ринків з високою конкуренцією і задоволеними покупцями	7,8	10,5
Ухилення від динамічних ринків з частою зміною товару	2,6	4,7

Визначте, чи корелюють між собою оцінки японських і британських фірм.

### Розв'язання задачі за допомогою Excel

#### Хід роботи

1. Ввести дані в стовпці А, В, С.

2. У колонках D і E записати формули для обчислення рангів.

3. У комірці F3 записати формулу для обчислення різниці рангів.

Скопіювати цю формулу в комірки F4 : F16.

4. У комірку G3 записати формулу для зведення цієї різниці в квадрат.

Скопіювати її в комірки G4 : G16.

5. У комірку G17 підсумувати різниці квадратів.

6. У комірку B18 підрахувати число вимірів(об'єктів) за допомогою функції СЧЕТ (у нашому випадку це число чинників успіху - 14), а в комірку G18 - формулу для обчислення коефіцієнта рангової кореляції Спірмена:

$$=1-6*G17/(B18*(B18*B18 - 1)) \text{ (див. табл. 8.2).}$$

Таблиця 8.2

#### Введення формул для розрахунку коефіцієнта рангової кореляції Спірмена

	A	B	C	D	E	F	G
1	Чинники успіху	Відсоток фірм		Ранги		Різниця рангів	Різниця рангів у квадраті
2		японських	британських	японських	британських		
3	Хороша адаптованість до потреб	69,8	75,6	=РАНГ(B3;B\$3:\$B\$16)	=РАНГ(C3;C\$3:\$C\$16)	=D3 - E3	=F3*F3
4	Перевага над конкурентами за якістю	79,3	59,3	=РАНГ(B4;B\$3:\$B\$16)	=РАНГ(C4;C\$3:\$C\$16)	=D4 - E4	=F4*F4
5	Перевага над конкурентами по дизайну	69,9	45,3	=РАНГ(B5;B\$3:\$B\$16)	=РАНГ(C5;C\$3:\$C\$16)	=D5 - E5	=F5*F5
6	Перевага над конкурентами по співвідношенню гідність/ціна	58,6	61,6	=РАНГ(B6;B\$3:\$B\$16)	=РАНГ(C6;C\$3:\$C\$16)	=D6 - E6	=F6*F6
7	Перевага над конкурентами по	55,2	48,8	=РАНГ(B7;B\$3:\$B\$16)	=РАНГ(C7;C\$3:\$C\$16)	=D7 - E7	=F7*F7

	конструкції						
8	Дуже конкурентна ціна	41,4	27,9	=РАНГ(B8;\$B\$3:\$B\$16)	=РАНГ(C8;C\$3:\$C\$16)	=D8 - E8	=F8*F8
9	Адаптованість до можливостей фірми	39,7	34,9	=РАНГ(B9;\$B\$3:\$B\$16)	=РАНГ(C9;C\$3:\$C\$16)	=D9 - E9	=F9*F9
10	Унікальність	36,2	29,1	=РАНГ(B10;\$B\$3:\$B\$16)	=РАНГ(C10;C\$3:\$C\$16)	=D10 - E10	=F10*F10
11	Ефективний маркетинг	27,6	25,6	=РАНГ(B11;\$B\$3:\$B\$16)	=РАНГ(C11;C\$3:\$C\$16)	=D11 - E11	=F11*F11
12	Глибокий аналіз ринку	27,6	18,6	=РАНГ(B12;\$B\$3:\$B\$16)	=РАНГ(C12;C\$3:\$C\$16)	=D12 - E12	=F12*F12
13	Великий об'єм ринку	20,7	16,3	=РАНГ(B13;\$B\$3:\$B\$16)	=РАНГ(C13;C\$3:\$C\$16)	=D13 - E13	=F13*F13
14	Синергія виробництва/маркетинг	16,4	18,6	=РАНГ(B14;\$B\$3:\$B\$16)	=РАНГ(C14;C\$3:\$C\$16)	=D14 - E14	=F14*F14
15	Ухилення від ринків з високою конкуренцією і задоволеними покупцями	7,8	10,5	=РАНГ(B15;\$B\$3:\$B\$16)	=РАНГ(C15;C\$3:\$C\$16)	=D15 - E15	=F15*F15
16	Ухилення від динамічних ринків з частою зміною товару	2,6	4,7	=РАНГ(B16;\$B\$3:\$B\$16)	=РАНГ(C16;C\$3:\$C\$16)	=D16 - E16	=F16*F16
17							=СУМ(G3:G16)
18	Об'єктів	=СЧЕТ(B3:B17)			<b>Коеф.Спірмена</b>		=1-6*G17/(B18*(B18*B18-1))

У результаті обчислень отримаємо таблицю з результатами (див. табл. 8.3).

Таблиця 8.3

### Результати обчислення коефіцієнта Спірмена

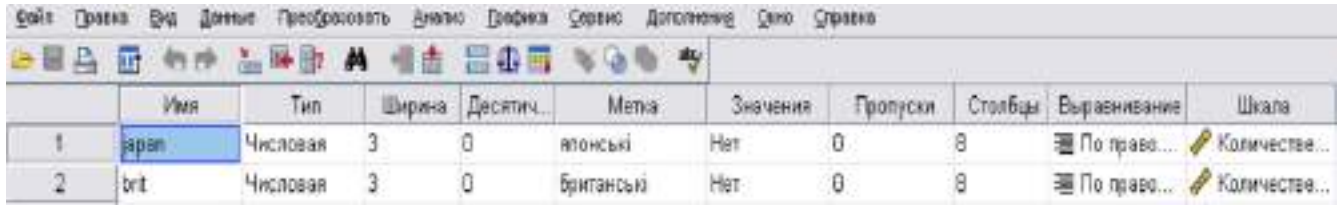
	A	B	C	D	E	F	G
1	Чинники успіху	Відсоток фірм		Ранги		Різниця рангів	Різниця рангів у квадраті
2		японських	британських	японських	британських		
3		Хороша адаптованість	69,8	75,6	3		

	до потреб						
4	Перевага над конкурентами за якістю	79,3	59,3	1	3	-2	4
5	Перевага над конкурентами по дизайну	69,9	45,3	2	5	-3	9
6	Перевага над конкурентами по співвідношенню гідність/ціна	58,6	61,6	4	2	2	4
7	Перевага над конкурентами по конструкції	55,2	48,8	5	4	1	1
8	Дуже конкурентна ціна	41,4	27,9	6	8	-2	4
9	Адаптованість до можливостей фірми	39,7	34,9	7	6	1	1
10	Унікальність	36,2	29,1	8	7	1	1
11	Ефективний маркетинг	27,6	25,6	9	9	0	0
12	Глибокий аналіз ринку	27,6	18,6	9	10	-1	1
13	Великий об'єм ринку	20,7	16,3	11	12	-1	1
14	Синергія виробництво/маркетинг	16,4	18,6	12	10	2	4
15	Ухилення від ринків з високою конкуренцією і задоволеними покупцями	7,8	10,5	13	13	0	0
16	Ухилення від динамічних ринків з частою зміною товару	2,6	4,7	14	14	0	0
17							34
18	Об'єктів	14			<b>Коеф. Спірмена</b>		0,9252747 25

## Розв'язання задачі за допомогою SPSS для Windows

### Хід роботи

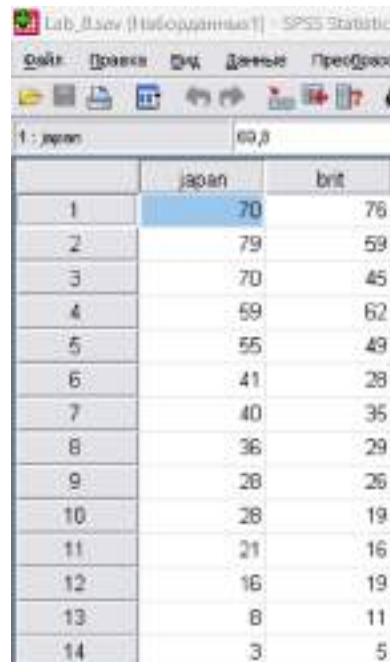
1. Описати в редакторі даних 2 змінних з іменами japan, brit (японські і британські фірми).



	Имя	Тип	Ширина	Десятич...	Метка	Значения	Пропуски	Столбцы	Выравнивание	Шкала
1	japan	Числовая	3	0	японські	Нет	0	8	По право...	Колличество...
2	brit	Числовая	3	0	британські	Нет	0	8	По право...	Колличество...

Рис. 8.1. Опис змінних у SPSS

Ввести дані згідно з таблицею, представленою на рис. 8.2.



	japan	brit
1	70	78
2	79	59
3	70	45
4	59	62
5	55	49
6	41	28
7	40	35
8	36	29
9	28	26
10	28	19
11	21	16
12	16	19
13	8	11
14	3	5

Рис. 8.2. Введення даних у SPSS

2. Обчислити ранги для змінних japan, brit:

Меню: **Transform(Перетворити), Rank Cases.(Ранги)** (див. рис. 8.3)

У результаті з'являться 2 нові змінні:

From New

variable variable Label

-----

JAPAN RJAPAN RANK of JAPAN

BRIT RBRIT RANK of BRIT

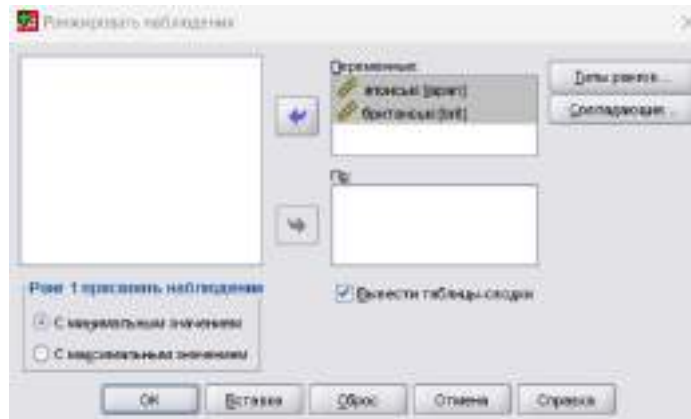


Рис. 8.3. Вікно «Rank Cases»

3. Обчислити коефіцієнти **Спірмена** і **Кенделла** для даних, представлених рангами:

Меню: *Analyze(Аналіз), Correlate(Кореляція), Bivariate(Двовимірна)*.

У вікні *Bivariate Correlation* задати необхідні параметри (див. рис. 8.4).

*Примітка: можна обчислювати коефіцієнти рангової кореляції без попереднього обчислення рангів, оскільки SPSS сам виконає необхідні розрахунки.*

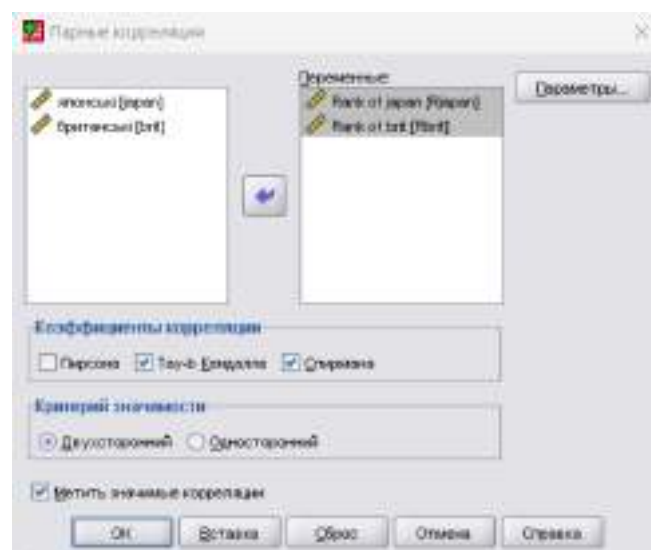


Рис. 8.4. Вікно «Bivariate Correlation»

4. Ознайомитися з отриманими результатами у вікні перегляду (рис. 8.5).

## РАНГ

### Созданные переменные<sup>b</sup>

Переменная источника	Функция	Новая переменная	Метка
japan <sup>a</sup>	Ранг	Rjapan	Rank of japan
brit <sup>a</sup>	Ранг	Rbrit	Rank of brit

a. Ранги по возрастанию.

b. Для совпадающих значений используется средний ранг.

## ➔ Непараметрические корреляции

### Корреляции

			Rank of japan	Rank of brit
tau-b Кендалла	Rank of japan	Кoeffициент корреляции	1,000	,811**
		Знач. (2-сторон)	.	,000
		N	14	14
	Rank of brit	Кoeffициент корреляции	,811**	1,000
		Знач. (2-сторон)	,000	.
		N	14	14
rho Спирмена	Rank of japan	Кoeffициент корреляции	1,000	,928**
		Знач. (2-сторон)	.	,000
		N	14	14
	Rank of brit	Кoeffициент корреляции	,928**	1,000
		Знач. (2-сторон)	,000	.
		N	14	14

\*\* Корреляция значима на уровне 0.01 (2-сторонняя).

Рис. 8.5. Результаты обчислення коефіцієнтів кореляції

### Завдання для самостійного виконання

Керівництво компанії Marston Book Services (Великобританія), що займається наданням комплексу послуг з доставки споживачам друкарської продукції (книги, плакати, листівки і тому подібне), розробило план посилення орієнтації діяльності компанії на запити споживачів. На одному з етапів розробки цього плану було проведено вивчення думок співробітників компанії і споживачів відносно важливості окремих показників якості послуг, що надавалися. У таблиці 8.4 наводяться рейтинги окремих показників якості.

## Рейтинг показників якості послуг

Показники якості послуг	Оцінки співробітників фірми	Оцінки споживачів
Швидкість обслуговування	1	7
Надійність	4	1
Неушкодженість упаковки замовлень	3	8
Задоволення термінових замовлень	8	3
Простота оформлення замовлень	7	6
Низький рівень реєстрацій	2	4
Надання інформації по запитах	6	5
Витримка термінів виконання замовлень	5	2

Визначте, чи узгоджуються оцінки співробітників компанії і споживачів. Розв'яжіть завдання в середовищі Excel і в середовищі SPSS для Windows.

**Вимоги до звіту**

Звіт про роботу має містити:

- постановку завдання, файли з початковими даними і результатами, формули для обчислення, значення обчислених коефіцієнтів;
- відповіді на контрольні питання.

**Контрольні запитання**

1. Які коефіцієнти рангової кореляції ви знаєте?
2. Для чого призначений коефіцієнт кореляції Спірмена?
3. За якою формулою він обчислюється? Яких граничних значень набуває?
4. Як визначити його значущість? Визначте значущість коефіцієнта Спірмена в прикладі «японські і британські фірми».
5. Як обчислити коефіцієнт Спірмена за допомогою електронних таблиць Excel?

6. Призначення функції «РАХУНОК». Для чого вона використовується в цій лабораторній роботі?

7. Призначення функції «РАНГ». Які параметри передаються цій функції? Для чого використовується у формулі для обчислення рангу абсолютна адресація?

8. Як перетворити дані в ранги? Які типи рангів є в SPSS для Windows?

9. Як обчислити коефіцієнти рангової кореляції в SPSS?

10. Як визначити значущість коефіцієнтів в SPSS?

11. Дайте змістовну інтерпретацію отриманим результатам.

## ЛАБОРАТОРНЕ ЗАНЯТТЯ № 9

**Тема:** *Обчислення коефіцієнтів кореляції та прогноз за допомогою лінійної регресії за допомогою Microsoft Excel*

**Мета:** навчитися використовувати електронні таблиці MS Excel для обчислення коефіцієнтів кореляції і прогнозування з використанням лінійної регресії.

### Постановка завдання

У таблиці 9.1 наведені результати тестування 20-ти абітурієнтів на вступному іспиті в ІТ-школу (назвемо його «тест 1» - це ознака X) і результати тестування цих же осіб уже на випускному курсі («тест 2», ознака Y). Побудувати діаграму розсіювання для цих даних, обчислити коефіцієнт кореляції Пірсона-Браве і вивести рівняння регресії, що пов'язує ознаки X і Y.

*Таблиця 9.1*

### Результати тестування в ІТ-школі

Тест 1 (X)	Тест 2 (Y)
66	94
62	100
70	101

80	102
82	103
74	105
73	104
79	106
83	106
80	109
81	110
84	110
80	111
89	112
88	112
95	114
98	114
97	115
94	117
89	118

**Хід роботи:**

1. Ввести початкові дані згідно з таблицею 9.1 (див. рис. 9.1).

	A	B	C	D	E	F	G
1	<b>Результати тестування в ІТ-школі</b>						
2	<b>№ студента</b>	<b>тест 1</b>	<b>тест 2</b>				
3	1	66	94				
4	2	62	100				
5	3	70	101				
6	4	80	102				
7	5	82	103				
8	6	74	105				

9	7	73	104				
10	8	79	106				
11	9	83	106				
12	10	80	109				
13	11	81	110				
14	12	84	110				
15	13	80	111				
16	14	89	112				
17	15	88	112				
18	16	95	114				
19	17	98	114				
20	18	97	115				
21	19	94	117				
22	20	89	118				

Рис. 9.1. Введення результатів тестування

2. Обчислити коефіцієнт кореляції Пірсона-Браве :

а) меню *Дані, Аналіз даних, Кореляція* (див. рис. 9.2);

б) у діалоговому вікні «*Кореляція*» (рис. 9.3) вивести і записати в зошит довідку;

в) задати параметри: вхідний інтервал (обидва ряди даних), групування – по стовпцях. Натиснути ОК. Записати в зошит отриманий коефіцієнт кореляції. Визначити його значущість по таблиці.

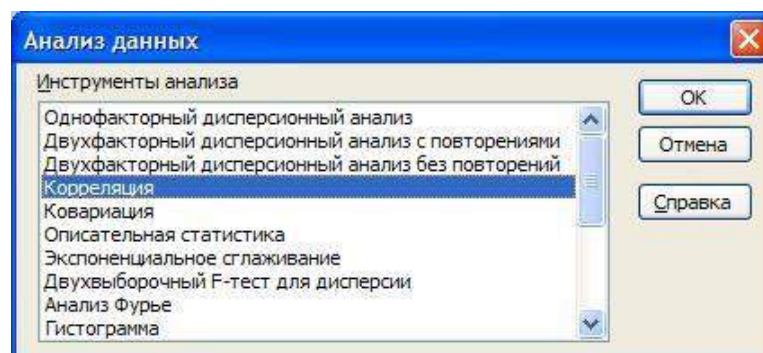


Рис. 9.2. Діалогове вікно «Аналіз даних»

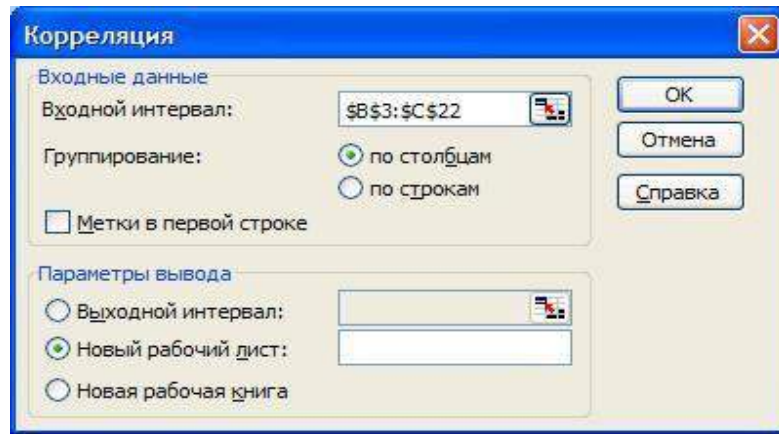


Рис. 9.3. Діалогове вікно «Кореляция»

Отримаємо інформацію в новому робочому аркуші (рис. 9.4):

	A	B	C	D
1		Столбец 1	Столбец 2	
2	Столбец 1	1		
3	Столбец 2	0,874114589	1	
4				

Рис. 9.4. Отриманий коефіцієнт кореляції

3. Побудувати діаграму розсіювання за допомогою *Майстра діаграм* (тип «точкова»). Зробити припущення про наявність або відсутність кореляції між X і Y за зовнішнім виглядом діаграми.

4. Побудувати лінію тренда по методу найменших квадратів (лінійна регресія):

а) натиснути правою кнопкою по точках діаграми;

б) у контекстному меню вибрати «Додати лінію тренду»;

в) у діалоговому вікні «Лінія тренду» на вкладці «Тип» (рис. 9.5) вкажіть - «лінійна», на вкладці «Параметри» - натисніть прапорець у опції «Показувати рівняння на діаграмі» і «Помістити величину достовірності апроксимації».

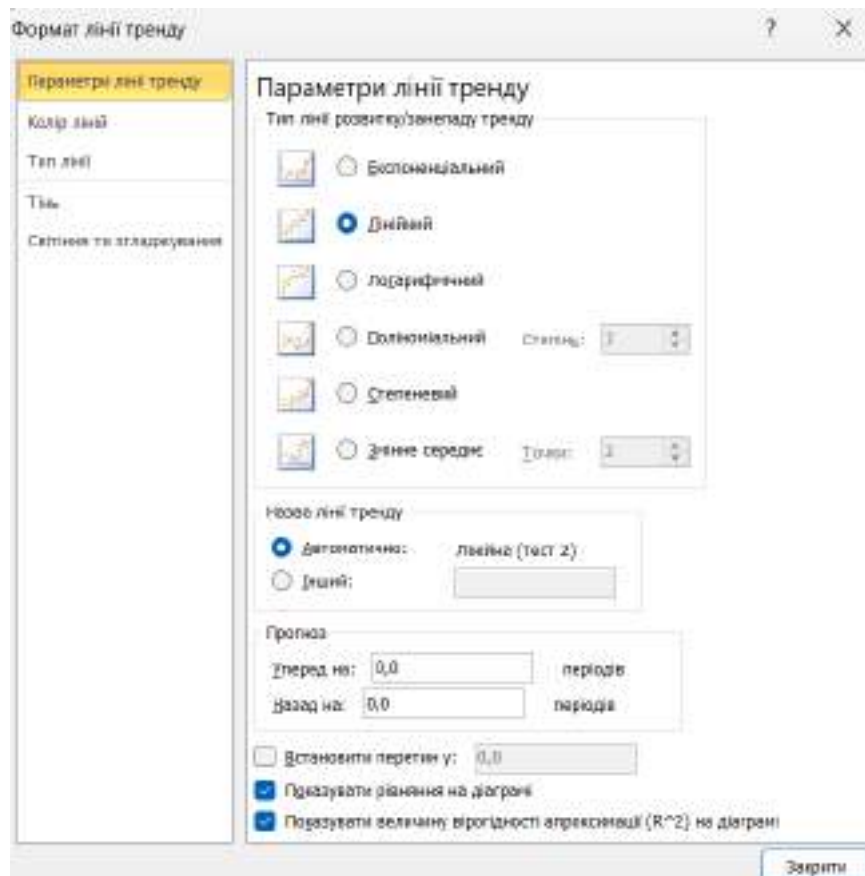


Рис. 9.5. Діалогове вікно «Лінія тренду», вкладка «Тип»

У результаті має вийти точкова діаграма з лінією тренду, рівнянням регресії і коефіцієнтом апроксимації, як це показано на рис. 9.6 (змінити шкалу обох осей згідно рисунку і додати необхідні підписи: назва діаграми, підписи осей, колір, ...).

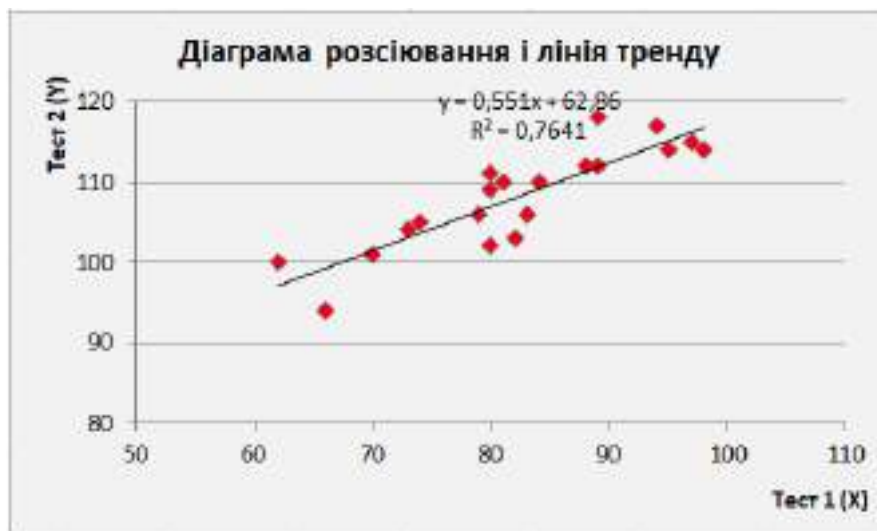


Рис. 9.6. Діаграма розсіювання, лінія тренду і рівняння регресії

Рівняння треба записати в зошит. Користуючись MS Excel, передбачити по ньому значення випускного тесту для осіб, у яких за результатами вступного тесту були наступні бали :

а) 43;

б) 75;

в) 99.

5. Зберегти таблицю в особистій теці.

### Завдання для самостійного виконання

1. У таблиці наведені коефіцієнти інтелекту IQ батьків і дітей десяти сімей. Визначте, чи існує кореляційна залежність між цими ознаками. Побудуйте діаграму розсіювання, лінію тренду і рівняння регресії (повністю аналогічно все згідно попереднього завдання).

IQ батьків	IQ дітей
125	110
120	105
110	95
105	125
105	120
95	105
95	75
90	95
80	90
75	80

2. Побудуйте модель лінійної регресії між роком і чисельністю наукових накопичень в Україні. Змініть повністю всі налаштування діаграми згідно рисунку 9.7.

### Наукові кадри і кількість організацій України

Рік	Кількість організацій, які ведуть наукові дослідження і розробки	Чисельність наукових накопичень, чол.	Чисельність докторів наук в Україні, чол.	Чисельність кандидатів наук в Україні, чол.
2003	1344	295010	8133	8133
2004	1350	248455	8797	8797
2005	1406	222127	9224	9224
2006	1463	207436	9441	9441
2007	1453	179799	9759	57610
2008	1435	160103	9974	58132
2009	1450	142532	10322	59332
2010	1518	134413	10446	59703
2011	1506	126045	10233	59547
2012	1490	120773	10339	58741
2013	1479	113341	10603	60647
2014	1477	107447	11008	62673
2015	1487	104841	11259	64372
2016	1505	106603	11573	65839
2017	1510	105512	12014	68291
2018	1452	100245	12488	71893
2019	1404	96820	12845	74191
2020	1378	94138	13423	77763
2021	1340	92403	13866	81169
2022	1303	89534	14418	84000
2023	1255	84969	14895	84979
2024	1208	89537	14925	85289

Ці дані можна знайти на сайті Держкомстату України: <http://www.ukrstat.gov.ua/>. Побудуйте також експоненціальну і поліноміальну модель. Порівняйте їх (рис. 9.7).



Рис. 9.7. Лінійна, експоненціальна і поліноміальна модель

### Вимоги до звіту

Звіт про роботу має містити:

- постановку завдання, початкові дані, коефіцієнт кореляції, рівняння лінії тренду, результати лінії тренду;
- файл з результатами.

### Контрольні запитання

1. У чому полягає завдання регресійного аналізу?
2. Що таке діаграма розсіювання?
3. Як отримати рівняння лінії тренду?
4. Як здійснюються прогнози за допомогою рівняння лінії регресії (лінії тренду)?
5. Як вичислити коефіцієнт кореляції Пірсона-Брава?
6. Що показує величина достовірності апроксимації?

## ЛАБОРАТОРНЕ ЗАНЯТТЯ № 10

**Тема:** *Обчислення коефіцієнтів кореляції і прогноз за допомогою лінійної регресії в SPSS для Windows*

**Мета:** навчитися використовувати SPSS для обчислення коефіцієнтів кореляції і прогнозування з використанням лінійної регресії.

### Постановка завдання

Керівництво освітнього закладу хотіло б мати інформацію про те, скільки студентів відвідують пари. У таблиці 10.1 наведені дані про відвідуваність занять і відповідні цифри за п'ять тижнів і за семестр, отриманих на початку дослідження. Необхідно представити графічно дані і отримати оцінки найменших квадратів в моделі  $y=b_0+b_1x$ .

Таблиця 10.1.

Відвідуваність у перші п'ять тижнів	Відвідуваність усього за семестр
235	392
122	190
34	74
196	307
121	200
127	185
177	291
125	191
159	314
75	140
140	235
127	188

89	142
139	253
70	96
62	96

### Розв'язання за допомогою SPSS

#### Хід роботи

1. Описати в редакторі дві змінні «Відвідуваність усього» і «Відвідуваність у перші п'ять тижнів» і ввести дані таблиці 10.1.

2. Вибрати в меню *Analyze... (Аналіз) Regression...(Регресія) Linear... (Лінійна)*. З'явиться діалогове вікно Linear Regression(Лінійна регресія) (рис. 10.1).

3. Перенести змінну «Відвідуваність усього» в полі для залежних змінних і присвоїти змінній «Відвідуваність у перші 5 тижнів» статус незалежної змінної. Почати розрахунок натисненням ОК.

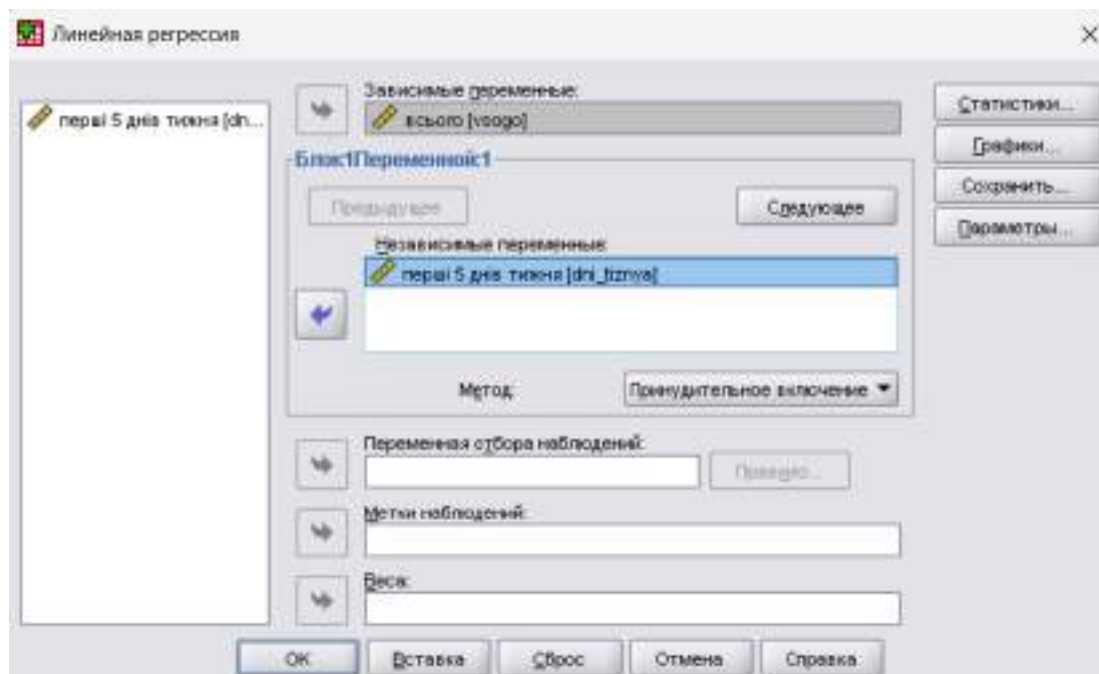


Рис. 10.1. Діалогове вікно «Linear Regression»

4. Виведення основних результатів виглядає так, як представлено на рис. 10.2 і 10.3.

## Регрессия

### Введенные или удаленные переменные<sup>b</sup>

Модель	Включенные переменные	Исключенные переменные	Метод
1	перші 5 днів тижня <sup>a</sup>		Принудительное включение

a. Включены все запрошенные переменные

b. Зависимая переменная: всего

### Сводка для модели

Модель	N	R-квадрат	Скорректированный R-квадрат	Стд. ошибка оценки
1	,974 <sup>a</sup>	,949	,946	20,666

a. Предикторы: (конст) перші 5 днів тижня

### Дисперсионный анализ<sup>b</sup>

Модель		Сумма квадратов	ст. св.	Средний квадрат	Щ	Знач.
1	Регрессия	112114,770	1	112114,770	262,521	,000 <sup>a</sup>
	Остатки	5978,980	14	427,070		
	Всего	118093,750	15			

a. Предикторы: (конст) перші 5 днів тижня

b. Зависимая переменная: всего

Рис. 10.2. Выведення результатів. Сумарна модель

Сумарна модель регресії (рис. 10.2) включає: значення коефіцієнта кореляції Пірсона  $R=0,974$ ; значення коефіцієнта детерміації  $R\text{ Square}=0,949$ ; значення уточненого коефіцієнта детерміації і стандартну помилку оцінки.

### Кoeffициенты<sup>a</sup>

Модель		Нестандартизованные коэффициенты		Стандартизованные коэффициенты	t	Знач.
		B	Стд. Ошибка	Бета		
1	(Константа)	-1,827	13,821		-,132	,897
	перші 5 днів тижня	1,663	,103	,974	16,202	,000

a. Зависимая переменная: всего

Рис. 10.3. Выведення результатів

## Кoeffициенты рівняння регресії

У таблиці «Кoeffициенты» (див. рис. 10.3) вікна виведення наводяться значення коефіцієнтів рівняння регресії, їх стандартної помилки, стандартизовані коефіцієнти, а також значення статистики **Стьюдента**  $t$  і її значущість  $Sig$ .

Коефіцієнт є значущим, якщо  $\text{Sig} < 0,05$ .

Отримане рівняння регресії :  $y = 1,663x - 1,827$ .

**5. Побудувати лінію регресії.** Для цього необхідно виконати наступні кроки:

- Побудувати діаграму розсіювання (згідно лаб\_7).
- Натиснути двічі на цьому графіку, щоб перенести його в редактор діаграм.
- Натиснувши на діаграмі розсіювання вибрати у контекстному меню «Добавить: Линия аппроксимации для итогов» (рис. 10.4):

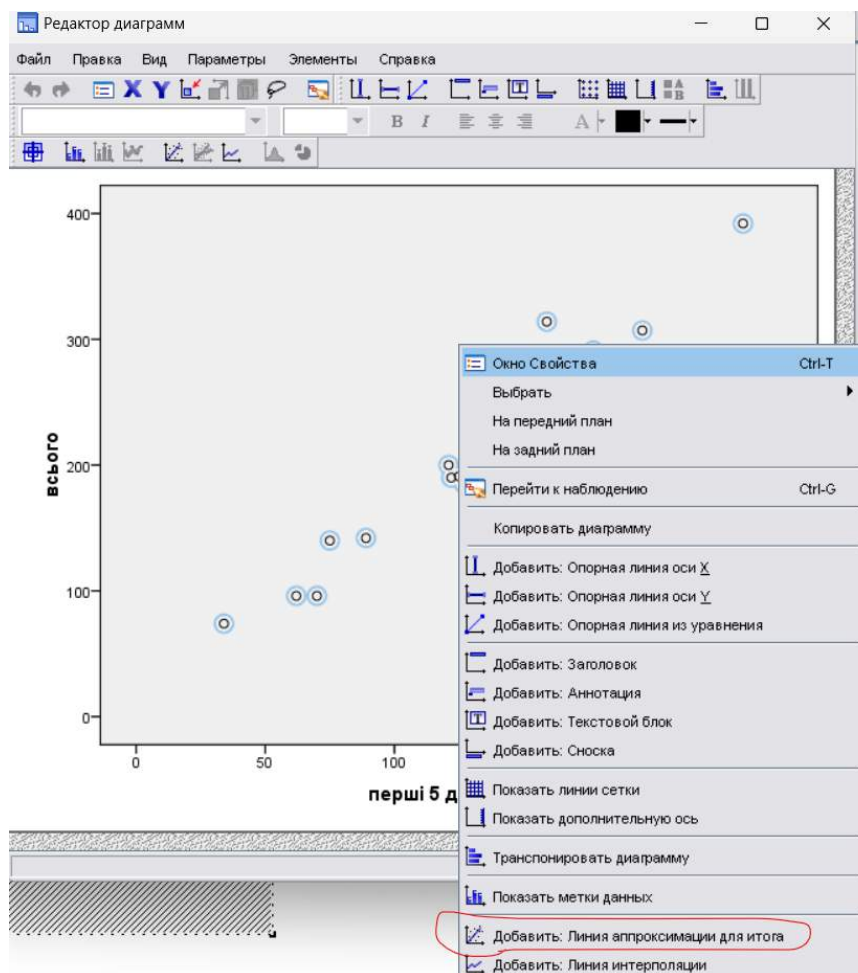


Рис. 10.4. Діалогове вікно налаштування діаграми розсіювання

- Потім вибрати «Лінійна регресія» (див. рис. 10.5).

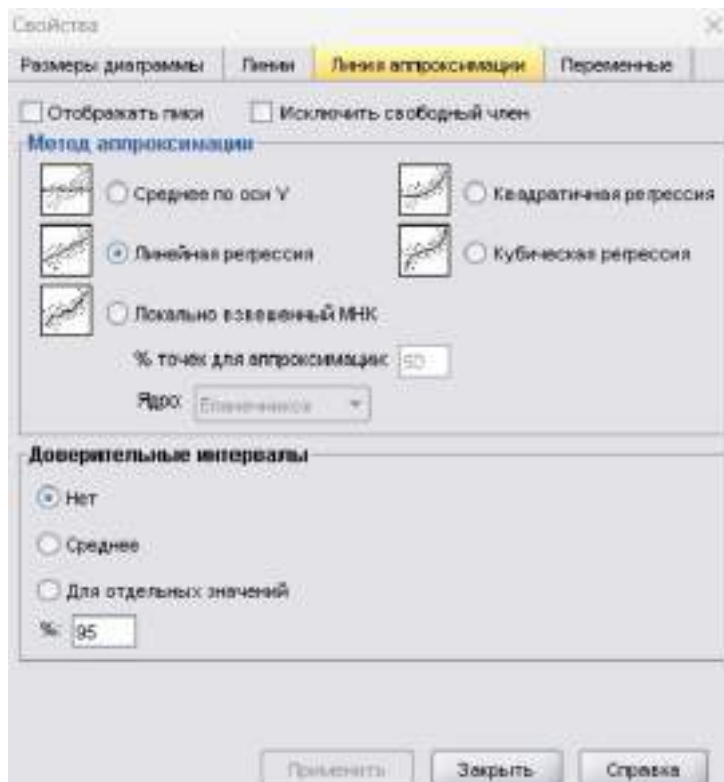


Рис. 10.5. Діалогове вікно властивості «Лінійна регресія»

- Закрити редактор діаграм. У вікні перегляду з'явиться діаграма розсіювання з лінією регресії (див. рис. 10.6).
- Підписати заголовок, осі, змінити колір.

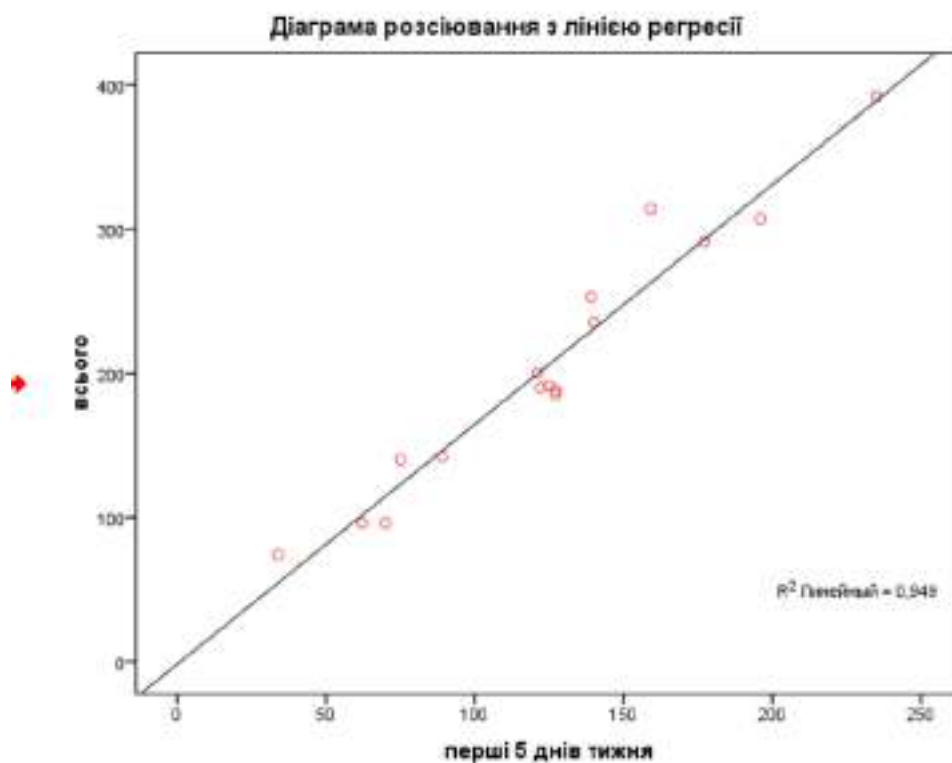


Рис. 10.6. Діаграма розсіювання з лінією регресії

## **Вимоги до звіту**

Звіт про роботу має містити:

- постановку завдання, початкові дані, коефіцієнт кореляції, рівняння лінії пророцтва, результати пророцтва;
- файл з результатами.

## **Контрольні запитання**

1. Яким чином ви описали змінну «Продано в перші 5 тижнів» в редакторі даних?
2. Як описати елементи діалогового вікна «Scatterplot Options»?
3. Як описати елементи діалогового вікна «Scatterplot Options : Fit Line»?
4. Як отримати рівняння лінії регресії в SPSS? Дайте інтерпретацію отриманим результатам.

## **ЛАБОРАТОРНЕ ЗАНЯТТЯ № 11**

**Тема:** *Перевірка гіпотез про значущість відмінностей середніх у Microsoft Excel*

**Мета:** навчитися використовувати пакет статистичного аналізу MS Excel для перевірки статистичних гіпотез.

### **Постановка завдання**

Досліджувалося навчання 40 студентів університету певної спеціальності. Навчання студентів було організовано так, що їх було розділено на дві групи по 20 чоловік. Студенти першої групи – навчалися на очній формі навчання, студенти другої групи – навчалися на дистанційній формі навчання. У таблиці 11.1 наведено кількість студентів, які навчалися в перший місяць навчання. Чи свідчать представлені в таблиці 11.1. дані про відмінність у підготовленості до навчання студентів, які навчалися на різних формах навчання (очній та дистанційній)?

## Кількість студентів залежно від форми навчання

Кількість студентів	
Група 1	Група 2
16	9
17	7
5	5
7	6
7	7
8	8
8	4
6	6
6	6
8	8
8	8
15	12
8	4
10	4
10	10
13	5
4	4
4	3
8	8
9	3

## Хід роботи:

1. Завантажити Excel. Ввести початкові дані згідно з таблицею 11.1.
2. Для кожного стовпця даних обчислити середнє і дисперсію за допомогою функцій **СРЗНАЧ** і **ДИСП** подивитися, чи рівні в обох групах вийшли середні і дисперсії.

3. Запустити пакет аналізу: меню «Дані», «Аналіз даних», «Двохвибірковий t-тест з різними дисперсіями» (див. рис. 11.1). Прочитати про нього довідку і записати її в зошит.

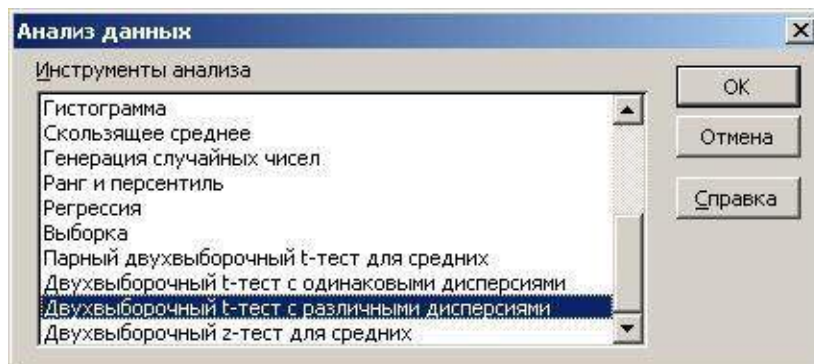
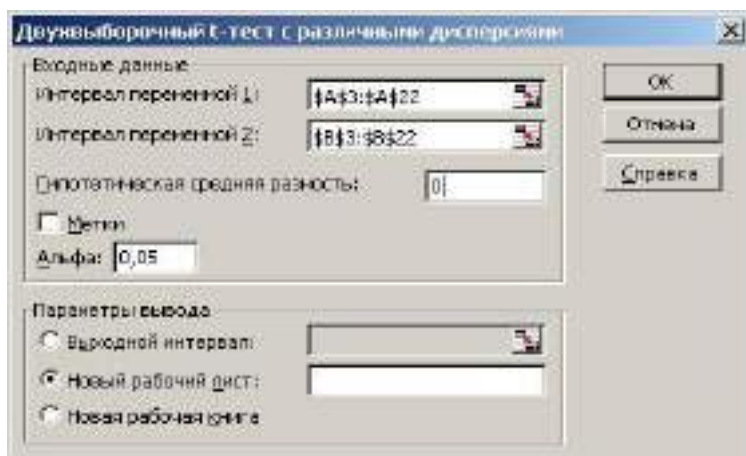


Рис. 11.1. Діалогове вікно «Аналіз даних»

4. У вікні «Двохвибірковий t-тест з різними дисперсіями» задати необхідні параметри: інтервал змінної 1, інтервал змінної 2, гіпотетична різниця середніх – 0, рівень значущості – 0,05. Натиснути кнопку ОК (рис. 11.2).



або

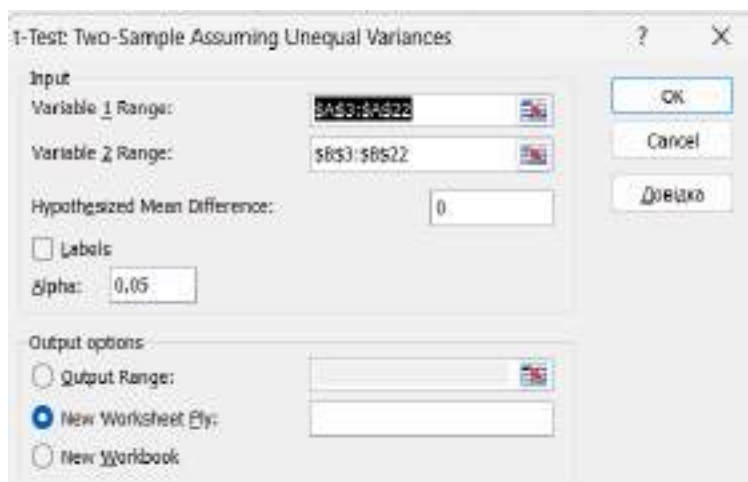


Рис. 11.2. Діалогове вікно «Двохвибірковий t-тест з різними дисперсіями»

5. Виписати таблицю з отриманими результатами (див. рис. 11.3) і зробити висновок відносно значущості відмінностей.

Двохвибірковий t- тест з різними дисперсіями		
	Змінна 1	Змінна 2
Середнє	8,85	6,35
Дисперсія	13,92368421	5,92368421
Спостереження	20	20
Гіпотетична різниця середніх	0	
df	33	
t-статистика	2,509594424	
P(T<=t) одностороннє	0,008586329	
t критичне одностороннє	1,692360309	
P(T<=t) двостороннє	0,017172658	
t критичне двостороннє	2,034515297	

Рис. 11.3. Результати розрахунку двухвибіркового t- тесту з різними дисперсіями

6. Зберегти таблицю в особистій теці.

### Вимоги до звіту

Звіт про роботу має містити:

- постановку завдання, початкові дані, опис двухвибіркового t-тесту з різними дисперсіями;
- файл з результатами, висновки про істинність або хибність гіпотези.

### Контрольні запитання

1. Назвіть види статистичних гіпотез. Наведіть приклади.
2. Сформулюйте нульову і альтернативну гіпотезу для завдання.
3. Охарактеризуйте поняття «Статистичний критерій».
4. У чому полягає відмінність між параметричними і непараметричними критеріями?
5. Які види статистичних критеріїв реалізовані в Excel?
6. Як вони здійснюються?