

## Регрессия

Регрессиялық талдау - бір немесе бірнеше белгілердің (факторлық белгілердің) және салдардың (нәтижелі белгілердің) арасындағы байланысты өлшеуге мүмкіндік беретін статистикалық өңдеу әдісі.

Корреляция (CORREL функциясы)

**Мақсаты:** Екі айнымалы арасындағы байланыстың күшін бағалау.

- **Нәтижесі:** -1 мен +1 аралығындағы сан.
  - **+1** → күшті оң байланыс
  - **0** → байланыс жоқ
  - **-1** → кері байланыс

REGRESSION (Регрессия)

**Мақсаты:** Тәуелді айнымалыны тәуелсіз айнымалыдан болжау.

Нәтижесі: Теңдеу  $y=a+bx$

$y=a+bx$  және  $R^2$ .

Екі әдістiде қолданып, нәтиженi график түрiнде көрсетуге болады.

Қарапайым салыстыру кестесі

Функция	CORREL	REGRESSION
Мақсаты	Байланыстың күшін бағалау	Тәуелді айнымалыны болжау
Нәтиже	-1 ... +1 (сан)	Теңдеу $y = a + bx$ , $R^2$
Байланыс бағысы	Көрінеді (оң/теріс)	Көрінеді + формулада
Болжау	Жоқ	Иә

MS Excel бағдарламасында регрессиялық талдауды орындау:

1 қадам:

Егер сізде “Data Analysis” батырмасы жоқ болса:

File (Файл) → Options (Параметрлер)

Add-ins (Қосымшалар)

Төменнен Excel Add-ins таңдап → Go

Analysis ToolPak белгілеңіз → OK

2 қадам:

Регрессияны орындау

Data (Деректер) қойындысына өтіңіз

Data Analysis (Деректерді талдау) таңдаңыз

Тізімнен Regression (Регрессия) таңдаңыз.

3 қадам:

Параметрлерді толтыру

Input Y Range → тәуелді айнымалы (Y)

Input X Range → тәуелсіз айнымалы (X)

Labels (егер баған атауы болса)

Output Range немесе New Worksheet таңдаңыз

→ ОК басыңыз

4 қадам

Нәтижені түсіну

Excel сізге мыналарды береді:

R Square ( $R^2$ ) → модель сапасы (1-ге жақын болса жақсы)

Coefficients → регрессия теңдеуі

( $y = a + bx$  түрінде)

P-value → айнымалының маңыздылығы

5 қадам

Альтернатив әдіс (график арқылы)

Деректерді белгілеңіз

Insert → Scatter (Нүктелік график)

Графикке оң жақ батырма → Add Trendline

“Display Equation on chart”

→ Сонда теңдеу бірден көрінеді

Мысалы:

оқу уақыты	баға
1	50
2	55
3	65
4	70
5	80

Формулалар мәзірінен =CORREL() формуласын пайдаланып, коррелясы коэффициентін анықтаймыз.  $R=0.99$

Ары қарай, байланыстың бағытын, күшін және нақты болжау формуласын анықтау үшін Деректерді талдау арқылы Regression есептейміз.

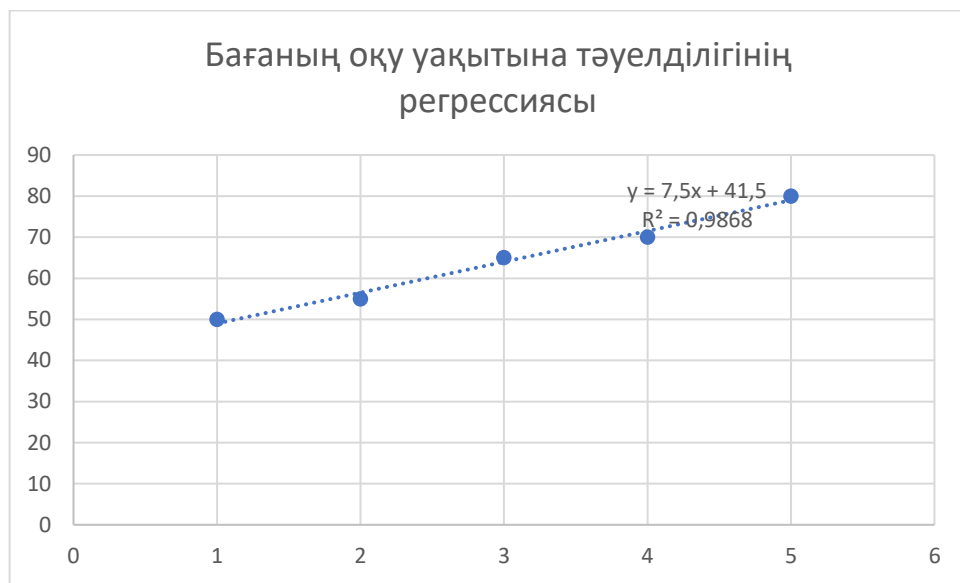
Регрессионная статистика	
Множественный R	0,993399268
R-квадрат	0,986842105
Нормированный R-квадрат	0,98245614
Стандартная ошибка	1,58113883
Наблюдения	5

Дисперсионный анализ						
	df	SS	MS	F	Значимость F	
Регрессия	1	562,5	562,5	225	0,000643119	
Остаток	3	7,5	2,5			
Итого	4	570				

	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
Y-пересечение	41,5	1,658312395	25,0254416	0,000139905	36,22250984	46,77749016	36,22250984	46,77749016
Переменная X 1	7,5	0,5	15	0,000643119	5,908776847	9,091223153	5,908776847	9,091223153



Сурет 1. Бағаның оқу уақытына тәуелділігінің регрессиясы

### Қорытынды:

#### 1. Регрессиялық статистика

- Көптік  $R \approx 0.99 \rightarrow$  байланыс өте күшті
- $R^2 \approx 0.986 \rightarrow$  модель деректердің 98.6%-ын түсіндіреді (өте жақсы нәтиже)
- Қателік  $\approx 1.58 \rightarrow$  қателік аз
- Бақылаулар = 5  $\rightarrow$  дерек саны аз (5 қана)

#### 2. Дисперсиялық талдау (ANOVA)

- $F = 225 \rightarrow$  модель маңызды
- Маңыздылығы (Significance F)  $\approx 0.00064 \rightarrow 0.05$ -тен кіші  $\rightarrow$  модель сенімді

#### 3. Коэффициенттер

- Y қиылысуы = 41.5
- X1 коэффициенті = 7.5

Модель теңдеуі:

$$Y = 41.5 + 7.5X$$

Бұл нені білдіреді:

#### 4. P-мәндері

- X 1-ге артса → Y шамамен 7.5-ке өседі
- Өте кішкентай → айнымалылар маңызды

Нәтиже:

- X пен Y арасында өте күшті байланыс бар
- Модель жақсы жұмыс істейді
- X-тің Y-ға әсері нақты және оң

#### Аудиториялық жұмыс

Жыл	Инвестиция (млн \$)	Экономикалық өсім (%)
2015	100	2.5
2016	120	3.0
2017	150	3.8
2018	180	4.2
2019	200	4.8
2020	220	5.0

#### Тапсырма:

1. Excel-де деректерді енгізіңіз.
2. Регрессияны орындаңыз (Data Analysis → Regression).
3. Экономикалық өсімнің инвестицияға тәуелділігін анықтаңыз.
4. Теңдеуді табыңыз ( $y = a + bx$ ).
5.  $R^2$  көрсеткішін түсіндіріңіз.
6. График салып, регрессия сызығын қосыңыз.

- **Input Y Range** – тәуелді айнымалы (экономикалық өсім).
- **Input X Range** – тәуелсіз айнымалы (инвестиция).

## Жеке үй тапсырмалары

1.

Тыңайтқыш (г)	Өсімдік биіктігі (см)
1	10
2	12
3	15
4	18
5	20
6	22
7	24
8	25
9	27
10	29

**Тапсырма:**

Excel-де деректерді енгізіп, тыңайтқыш мөлшері мен өсімдік биіктігі арасындағы регрессияны анықтаңыз.

- **Input X Range:** A2:A11
- **Input Y Range:** B2:B11
- Регрессия теңдеуін табыңыз және  $R^2$  көрсеткішін түсіндіріңіз.
- График салып, регрессия сызығын қосыңыз.

2.

Оқу уақыты (сағ)	Баға (балл)
1	50
2	55
3	60
4	70
5	75
6	80
7	85
8	87
9	90
10	92

**Тапсырма:**

Оқу уақыты мен емтихан бағасы арасындағы регрессияны Excel-де орындаңыз.

- **Input X Range:** A2:A11
- **Input Y Range:** B2:B11
- Регрессия теңдеуін және корреляция коэффициентін анықтаңыз.
- Графикке тренд сызығын қосыңыз.

3.

Жарнама шығындары (мың \$)	Сатылым көлемі (мың бірлік)
10	200
15	220
20	250

Жарнама шығындары (мың \$)	Сатылым көлемі (мың бірлік)
25	280
30	300
35	330
40	350
45	370
50	390
55	400

**Тапсырма:**

Сатылым көлемі мен жарнама шығындары арасындағы байланысты Excel-де регрессия арқылы талдаңыз.

- **Input X Range:** A2:A11
- **Input Y Range:** B2:B11
- Теңдеуді табыңыз және  $R^2$  мәнін түсіндіріңіз.
- Графикке тренд сызығын қосыңыз.

**4.**

Үй ауданы (м <sup>2</sup> )	Үй бағасы (мың \$)
50	100
60	120
70	140
80	160
90	180
100	200
110	220
120	240
130	260
140	280

**Тапсырма:**

Үй ауданы мен үй бағасы арасындағы регрессияны Excel-де орнатыңыз.

- **Input X Range:** A2:A11
- **Input Y Range:** B2:B11
- Теңдеуді есептеп, регрессияның нәтижесін түсіндіріңіз.
- График құрып, тренд сызығын қосыңыз.

**5.**

Жұмыс өтілі (жыл)	Жалақы (мың \$)
1	30
2	35
3	40
4	45
5	50
6	55
7	60

Жұмыс өтілі (жыл)	Жалақы (мың \$)
8	65
9	70
10	75

**Тапсырма:**

Жұмыс өтілі мен жалақы арасындағы регрессияны Excel-де есептеңіз.

- **Input X Range:** A2:A11
- **Input Y Range:** B2:B11
- Результатты талдап, графикке тренд сызығын қосыңыз.
- Теңдеуді табыңыз ( $y = a + bx$ )

6.

Сағат	Калория
1	200
2	300
3	400
4	500
5	600
6	700
7	800
8	900
9	1000
10	1100

**Тапсырма:**

Спорт уақыты мен калория шығыны арасындағы регрессияны есептеңіз, қорытындыны талдап, графикке тренд сызығын қосыңыз. Теңдеуді табыңыз ( $y = a + bx$ )

**Input X Range:** A2:A11

**Input Y Range:** B2:B11

7.

Уақыт	Баға
1	90
2	85
3	80
4	75
5	70
6	65
7	60
8	55
9	50
10	45

- **Тапсырма:**

Интернет уақыты мен баға арасындағы регрессияны анықтаңыз. Теңдеуді табыңыз ( $y = a + bx$ ). Графикке тренд сызығын қосыңыз.

**Input X Range:** A2:A11

**Input Y Range:** B2:B11

8.

Жаңбыр	Өнім
10	50
20	60
30	65
40	70
50	80
60	85
70	90
80	95
90	100
100	110

- **Тапсырма:**

Жаңбыр мен өнім арасындағы регрессияны есептеңіз. Теңдеуді табыңыз ( $y = a + bx$ ). Графикке тренд сызығын қосыңыз.

**Input X Range:** A2:A11

**Input Y Range:** B2:B11

9.

Жылдамдық	Қашықтық
20	5
30	10
40	18
50	25
60	35
70	50
80	65
90	80
100	95
110	110

- **Тапсырма:**

Жылдамдық пен тоқтау қашықтығы арасындағы регрессияны анықтаңыз. Теңдеуді табыңыз ( $y = a + bx$ ). Графикке тренд сызығын қосыңыз.

**Input X Range:** A2:A11

**Input Y Range:** B2:B11

10.

Ұйқы (сағ)	Өнімділік (%)
4	50
5	60
6	70
7	80
8	90

Ұйқы (сағ)	Өнімділік (%)
9	95
10	97
11	96
12	94
13	90

- **Тапсырма:**

Ұйқы уақыты мен өнімділік арасындағы регрессияны Excel-де есептеңіз. Теңдеуді табыңыз ( $y = a + bx$ ). Графикке тренд сызығын қосыңыз.

**Input X Range:** A2:A11

**Input Y Range:** B2:B11

Результатты талдап, графикке тренд сызығын қосыңыз.

11.

Кофе (кесе)	Тиімділік (%)
0	40
1	55
2	70
3	80
4	85
5	80
6	70
7	60
8	50
9	40

- **Тапсырма:**

Кофе тұтыну мен жұмыс тиімділігі арасындағы регрессияны анықтаңыз. Теңдеуді табыңыз ( $y = a + bx$ ). Графикке тренд сызығын қосыңыз.

**Input X Range:** A2:A11

**Input Y Range:** B2:B11

12.

Көлік саны	Кептеліс (мин)
50	10
70	15
90	20
110	30
130	40
150	55
170	70
190	85
210	100
230	120

- **Тапсырма:**

Көлік саны мен кептеліс уақыты арасындағы регрессияны есептеңіз. Теңдеуді табыңыз ( $y = a + bx$ ). Графикке тренд сызығын қосыңыз.

**Input X Range:** A2:A11

**Input Y Range:** B2:B11

13.

Жұмысшылар	Өнім (дана)
5	100
10	180
15	260
20	330
25	400
30	460
35	510
40	550
45	580
50	600

- **Тапсырма:**

Жұмысшылар саны мен өндіріс көлемі арасындағы регрессияны Excel-де анықтаңыз. Теңдеуді табыңыз ( $y = a + bx$ ). Графикке тренд сызығын қосыңыз.

**Input X Range:** A2:A11

**Input Y Range:** B2:B11

14.

Материал (бет)	Баға
10	50
20	60
30	65
40	70
50	75
60	80
70	85
80	88
90	90
100	92

- **Тапсырма:**

Оқу материалы мен тест нәтижесі арасындағы регрессияны есептеңіз. Теңдеуді табыңыз ( $y = a + bx$ ). Графикке тренд сызығын қосыңыз.

**Input X Range:** A2:A11

**Input Y Range:** B2:B11

15.

Музыка (сағ)	Тиімділік (%)
0	60
1	65

Музыка (сағ)	Тиімділік (%)
2	70
3	75
4	78
5	80
6	79
7	77
8	74
9	70

Музыка тыңдау уақыты мен оқу тиімділігі арасындағы регрессияны Excel-де есептеңіз.

- **Input X Range:** A2:A11
- **Input Y Range:** B2:B11
- Корреляция коэффициентін табыңыз
- Регрессия теңдеуін анықтаңыз
- $R^2$  мәнін есептеңіз
- Scatter (нүктелік) график құрыңыз
- Графикке Trendline қосыңыз
- Нәтижені талдаңыз