

Задача 1. Имеются следующие данные по 10 сельскохозяйственным предприятиям Краснодарского края: выручка от реализации, млн. руб. (Y); среднегодовая численность работников, чел. (X).

Y	258	215	376	218	473	176	130	428	634	371
X	364	275	408	242	453	292	312	434	671	353

1. Оценить тесноту связи с помощью показателей корреляции и детерминации.
 2. Оценить значимость уравнения связи с помощью F-критерия Фишера.
 Сделать выводы по полученным результатам.

Решение

При линейной зависимости теснота связи между переменными X и Y определяется с помощью коэффициента корреляции:

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y}$$

где σ_x и σ_y – средние квадратические отклонения по X и Y.

№	Y	X	x^2	y^2	xy
1	258	364	132496	66564	93912
2	215	275	75625	46225	59125
3	376	408	166464	141376	153408
4	218	242	58564	47524	52756
5	473	453	205209	223729	214269
6	176	292	85264	30976	51392
7	130	312	97344	16900	40560
8	428	434	188356	183184	185752
9	634	671	450241	401956	425414
10	371	353	124609	137641	130963
Сумма	3279	3804	1584172	1296075	1407551
Среднее	327,90	380,40	158417,20	129607,50	140755,10

$$\sigma_x = \sqrt{\overline{x^2} - \bar{x}^2} = \sqrt{158417,2 - 380,4^2} = 117,10$$

$$\sigma_y = \sqrt{\overline{y^2} - \bar{y}^2} = \sqrt{129607,50 - 327,90^2} = 148,62$$

$$r = \frac{140755,10 - 380,4 \cdot 327,9}{117,1 \cdot 148,62} = 0,9206$$

Так как значение полученного коэффициента корреляции довольно близко к единице, то между признаками связь тесная и прямая.

Коэффициент детерминации - это квадрат коэффициента корреляции. Показывает, какая доля дисперсии результативного признака объясняется влиянием независимых переменных.

$$R^2 = 0,8475$$

Коэффициент детерминации показывает, что 84,75% различий в выручке от реализации объясняется изменением в численности работников.

Статистическая надежность уравнения регрессии проверяется с использованием критерия F-Фишера.

$$H_0: r^2=0,$$

$$H_1: r^2 \neq 0.$$

Наблюдаемое (фактическое) значение F – критерия находится по формуле:

$$F_n = \frac{R^2}{1-R^2}(n-2)$$

$$F_n = \frac{0.8475}{1-0.8475}(10-2) = 44.46$$

При уровне значимости $\alpha = 0,05$ и числе степеней свободы $k_1=m=1$, где m – число параметров при факторе x ; $k_2=n-m-1=10-1-1=8$ по таблице находится критическое значение F – критерия.

$$F_{кр}=5,32$$

Так как $F_n > F_{кр}$, то уравнение регрессии статистически значимое

Задача 2.

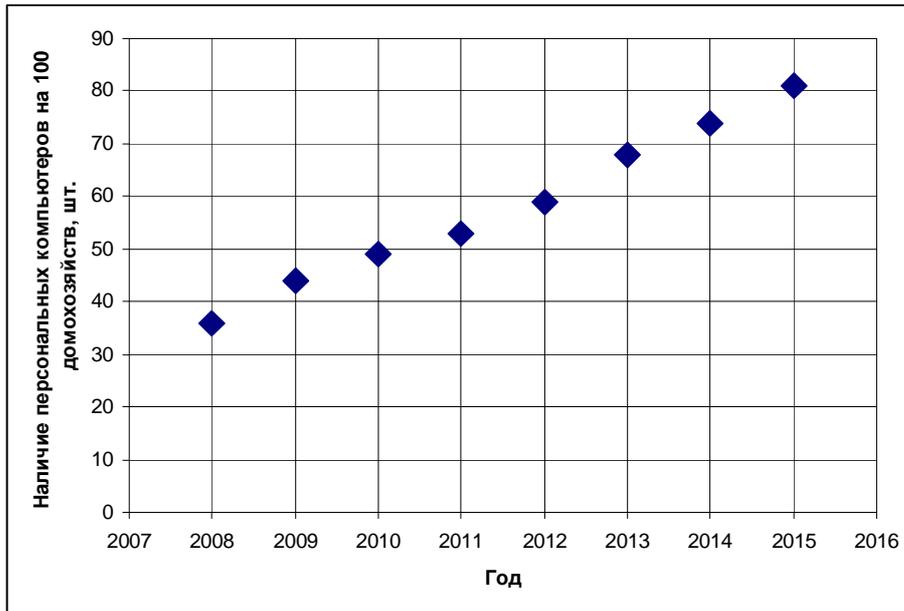
Год	2008	2009	2010	2011	2012	2013	2014	2015
Наличие персональных компьютеров на 100 домохозяйств, шт.	36	44	49	53	59	68	74	81

1. Временной ряд изобразить графически.
2. Подобрать уравнение тренда, отражающее общую тенденцию изменения уровней временного ряда.
3. Определить параметры уравнения тренда.
Сделать выводы по результатам расчетов.



Решение

Изобразим временной ряд графически:



Можно предположить, что в ряду имеется линейный тренд, т.е. $Y=a+bt$
 t – значение дат.

Для определения неизвестных параметров a и b в соответствии с требованием способа наименьших квадратов необходимо решить систему нормальных уравнений:

$$\begin{cases} na + b \sum t = \sum y \\ a \sum t + b \sum t^2 = \sum yt \end{cases}$$

№ года	Год	Наличие персональных компьютеров на 100 домохозяйств, шт., y	Условное обозначение периодов, t	t^2	yt
1	2008	36	1	1	36
2	2009	44	2	4	88
3	2010	49	3	9	147
4	2011	53	4	16	212
5	2012	59	5	25	295
6	2013	68	6	36	408
7	2014	74	7	49	518
8	2015	81	8	64	648
Итого		464	36	204	2352

Получаем:

$$\begin{cases} 8a + 36b = 464 \\ 36a + 204b = 2352 \end{cases}$$

Решаем систему:

$$\begin{cases} a = 58 - 4.5b \\ 36 \cdot (58 - 4.5b) + 204b = 2352 \end{cases}$$

$$\begin{cases} a = 58 - 4.5b \\ -162b + 204b = 2352 - 2088 \end{cases}$$

$$\begin{cases} a = 58 - 4.5b \\ 2088 - 162b + 204b = 2352 \end{cases}$$

$$\begin{cases} a = 58 - 4.5b \\ 42b = 264 \end{cases} \quad \begin{cases} a = 29.7143 \\ b = 6.2857 \end{cases}$$

Получили уравнение тренда:

$$Y = 29,7143 + 6,2857t$$

Т.е. с каждым годом среднее количество персональных компьютеров на 100 домохозяйств увеличивается на 6,3