REGRESSION ANALYSIS

Anarbayeva Fotima

PhD, Associate Professor of the Samarkand branch of Tashkent State University of Economics

Esanov Mirjakhon

Student of the Samarkand branch of Tashkent State University of Economics

e-mail: *esanovmirjaxon25012003@gmail.com*

https://doi.org/10.5281/zenodo.11920616

Abstract. *Regression analysis is a widely used statistical technique that is used in a variety of fields, including the social sciences, engineering, finance, health care, and more. Regression is a tool that helps you understand the relationship between a dependent variable and one or more independent variables.*

Key words: *Regression, regression analysis, construction of single and multi-factor linear and non-linear regression models, types of regression analysis, analysis methods, algorithms, instrumental software tools.*

РЕГРЕССИВНЫЙ АНАЛИЗ

Аннотация. *Регрессионный анализ — это широко используемый статистический метод, который используется в различных областях, включая социальные науки, инженерное дело, финансы, здравоохранение и другие. Регрессия — это инструмент, который помогает вам понять взаимосвязь между зависимой переменной и одной или несколькими независимыми переменными.*

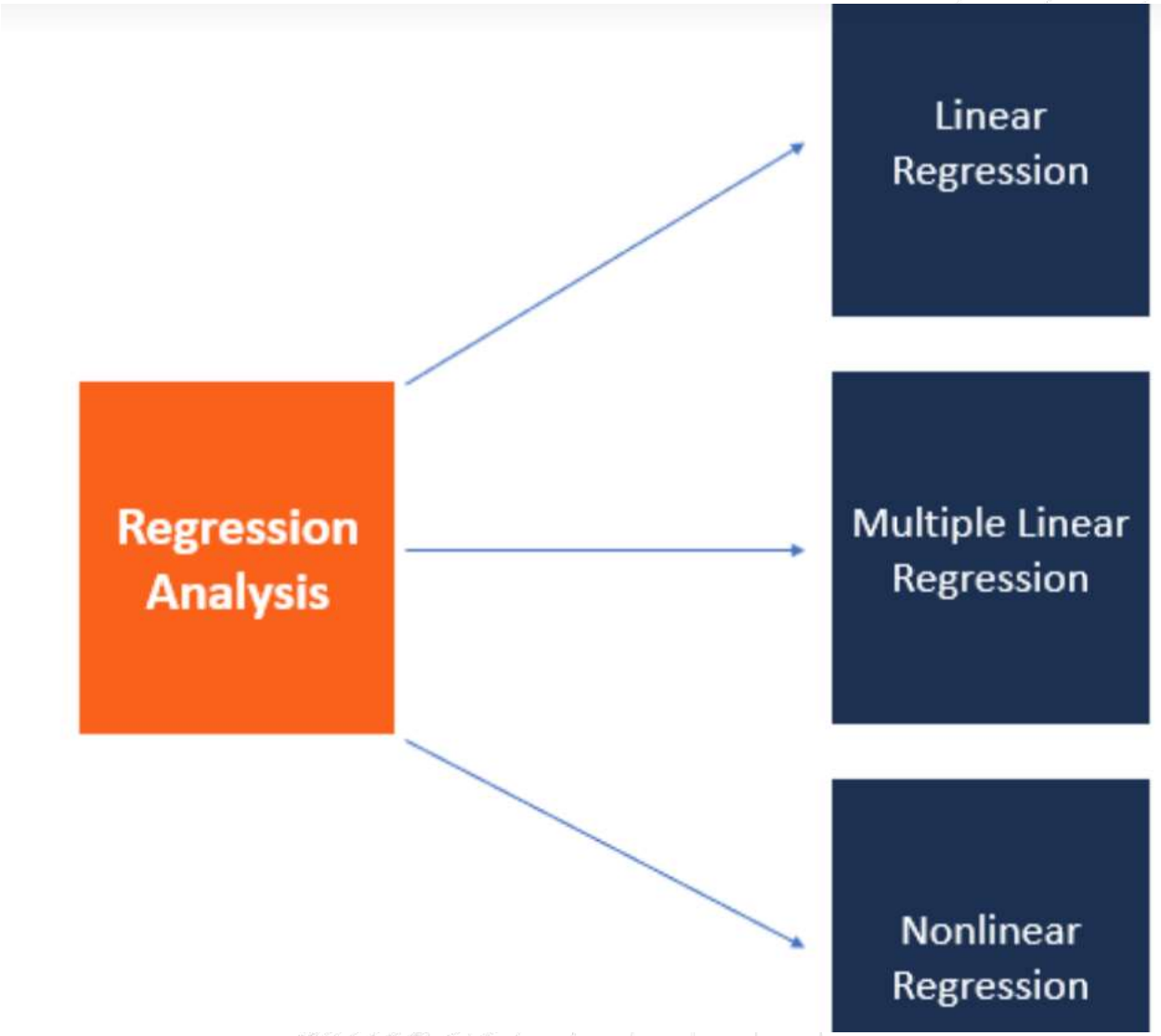
Ключевые слова: *Регрессия, регрессионный анализ, построение одно- и многофакторных линейных и нелинейных регрессионных моделей, виды регрессионного анализа, методы анализа, алгоритмы, инструментальные программные средства.*

Regression is a statistical technique used in finance, investing, and other disciplines that compares the relationship between one dependent variable (usually denoted Y) and a number of other variables (known as the independent variables) tries to determine the strength and character of the relationship.

Regression analysis is a set of statistical methods used to evaluate the relationship between a dependent variable and one or more independent variables. It can be used to estimate the strength of relationships between variables and to model future relationships between them.

Regression analysis includes multiple variables such as linear, multi-linear and non-linear. The most common models are simple linear and multi-linear. Nonlinear regression analysis is

typically used for more complex data sets where the dependent and independent variables exhibit a non-linear relationship.



Regression analysis mainly studies the relationship between two types of variables: the dependent variable and the independent variable. First, the dependent variable, which is affected by one or more independent variables, must be identified by:

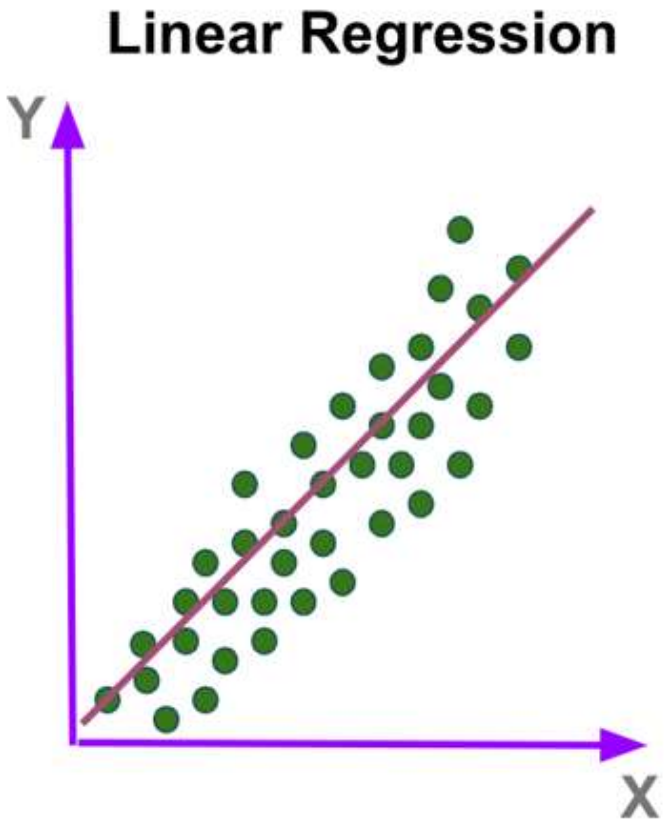
- Identification of variables and data collection;
- Building data graphics;
- Correlation analysis;
- Introduction to the regression line;
- Understanding the regression line;
- Interpretation of the formula;
- Error accounting.

Regression analysis method divided into several types, each of which has its own characteristics, strengths and limitations. When choosing the appropriate method for a particular problem, it is important to understand the different types of regression analysis, which are divided into the following types:

1. Linear regression:

Linear regression regression of analysis the most many p applied type be , he dependent was o ' variable and one or one how many independent o ' variables o ' between relationship modeling for is used .

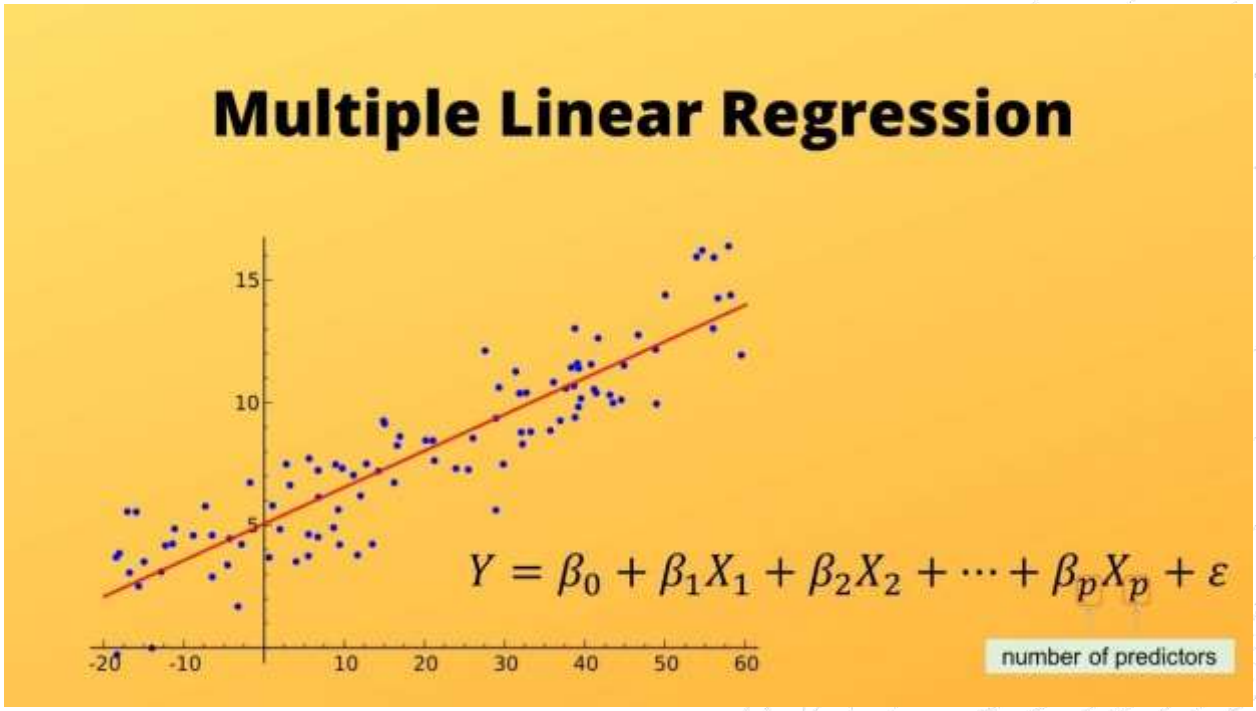
Linear of regression main Mother of God observed and prophecy done values o ' between quadratic of differences gather now which minimizes the most good compatibility line is to find For example , a company price , advertising budget and another factors based on product to sell prophecy to do for linear from regression use can



2. Many linear regression:

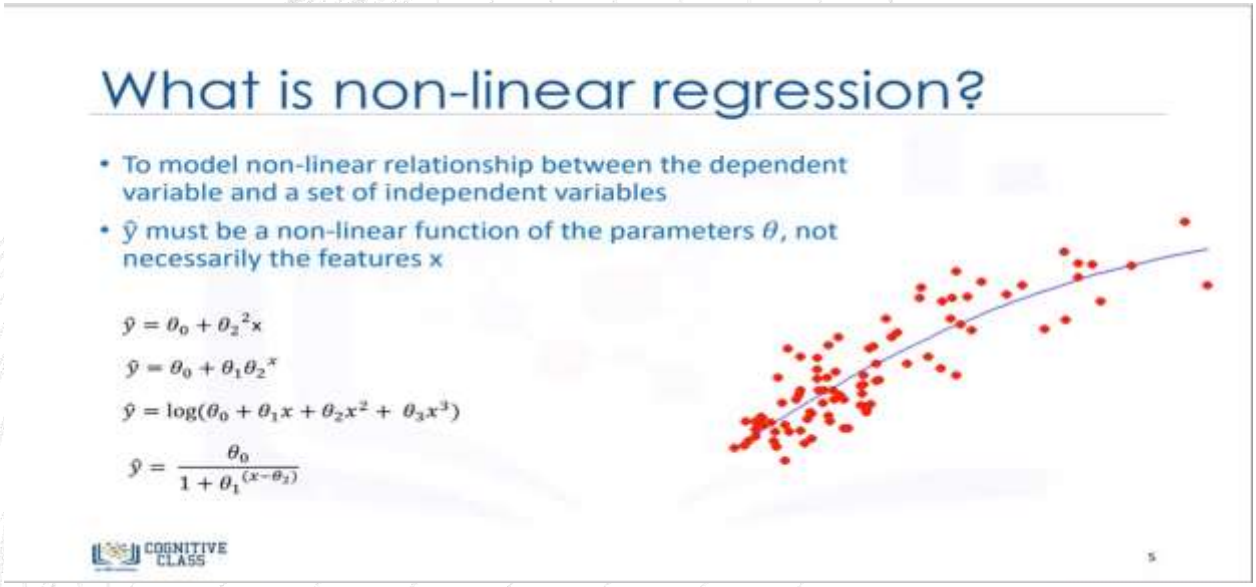
Many linear regression two or from him more than o ' variables to the value of based on o ' variable the result prophecy to do for used statistics technique means This is sometimes known simply as multiple regression and is an extension of linear regression. The variable we want to

predict is known as the dependent variable, while the variables we use to predict the value of the dependent variable are known as the independent or explanatory variables.



3. Nonlinear regression:

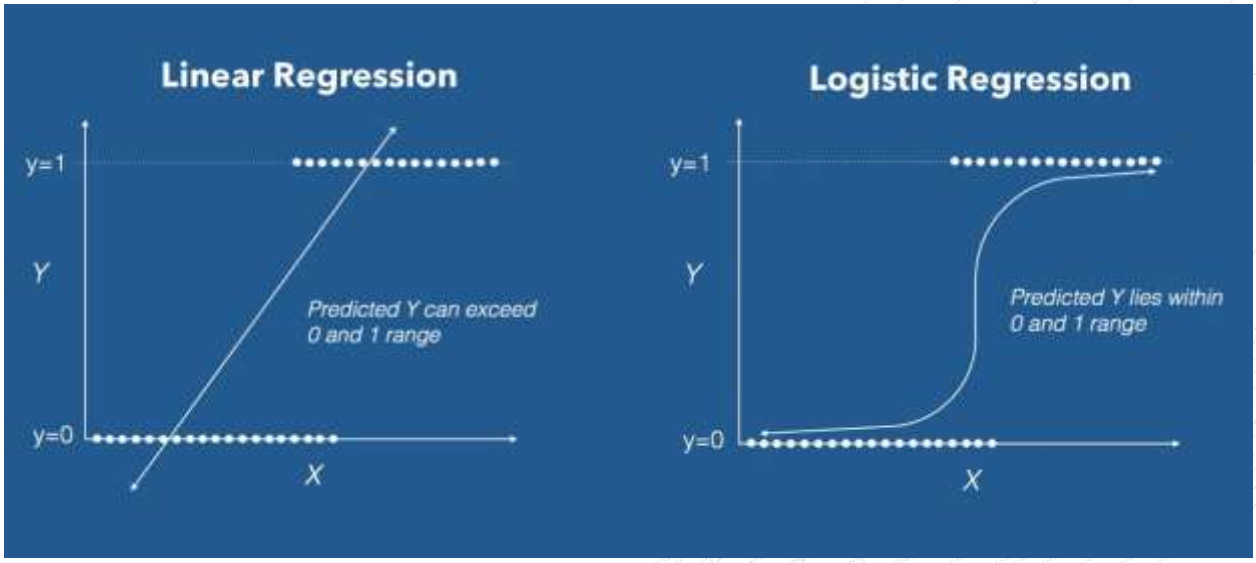
Nonlinear regression is a mathematical model that fits an equation to known data using a generated line. As with linear regression, which uses a straight-line equation (for example, Ỵ=c+mx), nonlinear regression shows the relationship using a curve, making it nonlinear in parameter. Also, nonlinear regression is a statistical technique that helps describe nonlinear relationships in experimental data. Nonlinear regression models are usually considered parametric, where the model is described as a nonlinear equation. Machine learning techniques are commonly used for non-parametric linear regression.



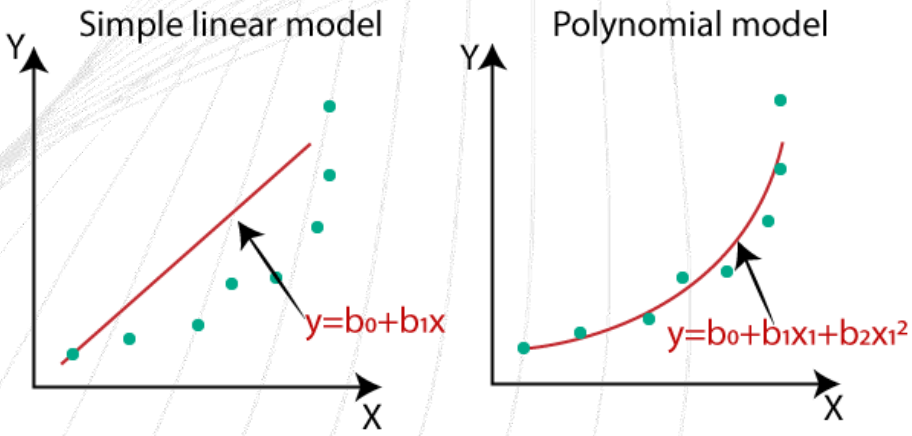
There are also the following types of regression:

Logistic regression is a type of regression analysis technique used when the dependent variable is discrete. Example: 0 or 1 or true or false.

This means that the target variable can have only two values, and the sigmoid curve represents the relationship between the target variable and the independent variable. For example, a hospital might use logistic regression to predict a patient's likelihood of having a heart attack based on a patient's age, gender, blood pressure, and other factors.



Polynomial regression is used when the relationship between the dependent variable and the independent variable is not linear. It is used to model relationships using an n-degree polynomial equation. For example, a scientist might use polynomial regression to model plant growth as a function of the amount of water and sunlight.



Analysis methods, algorithms, instrumental software tools.

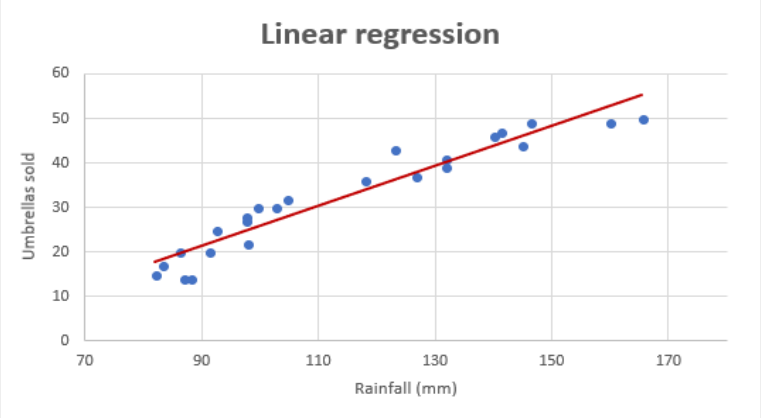
Regression analysis in Excel

In statistical modeling, regression analysis is used to evaluate the relationship between two or more variables:

1. **The dependent variable** (also known as the criterion variable) is the main factor that you are trying to understand and predict.

2. **Independent variables** (called explanatory variables or predictors) are factors that can affect the dependent variable.

For example, take umbrella sales data for the last 24 months and find out the average monthly rainfall for that period. Let this data be plotted and the regression line shows the relationship between the independent variable (rainfall) and the dependent variable (umbrella sales):



Linear regression equation:

Mathematically, linear regression is defined by the following equation:

y = a+ bx + e

Where:

x is the independent variable

y is the dependent variable.

α is the y-intercept, which is the expected mean value of y when all variables of x are equal to 0. In a regression plot, it is the point where the line intersects the Y-axis.

b is the slope of the regression line, which represents the rate of change of y as ux changes.

e is the random error, which is the difference between the true value of the dependent variable and its predicted value.

A linear regression equation always contains error because predictors are never exact in real life. However, some programs, including Excel, perform error calculations "behind the scenes".

Calculate linear regression using the method of least squares in Excel and

y =a+bx.

For our example, the linear regression equation will look like this: **Canopy = b * precipitation + a**. There are several different ways to find a and b. There are three main ways to perform linear regression analysis in Excel:

- The regression tool included in the Analysis ToolPak plugin
- Dot chart with trendline
- Linear regression formula.

Let's see how to perform the regression analysis:

In this example, a simple linear regression is performed in Excel. Column B, the independent variable (predictor), lists the average monthly rainfall for the past 24 months, and column C, the independent variable, lists the number of umbrellas sold. Of course, there are many other factors that can affect sales , but for now we'll focus on just these two variables :

	A	B	C
1	**Oylar**	**Yog'ingarchilik (mm)**	**Soyabonlar soni**
2	yanvar	82	15
3	fevral	92,5	25
4	mart	83,2	17
5	aprel	97,7	28
6	may	131,9	41
7	iyun	141,3	47
8	iyul	165,4	50
9	avgust	140	46
10	sentabr	126,7	37
11	oktabr	97,8	22
12	noyabr	86,2	20
13	dekabr	99,6	30
14	yanvar	87	14
15	fevral	97,5	27
16	mart	88,2	14
17	aprel	102,7	30
18	may	123	43
19	iyun	146,3	49
20	iyul	160	49
21	avgust	145	44
22	sentabr	131,7	39
23	oktabr	118	36
24	noyabr	91,2	20
25	dekabr	104,6	32

With the Analysis ToolPak enabled, perform the following steps to perform regression analysis in Excel:

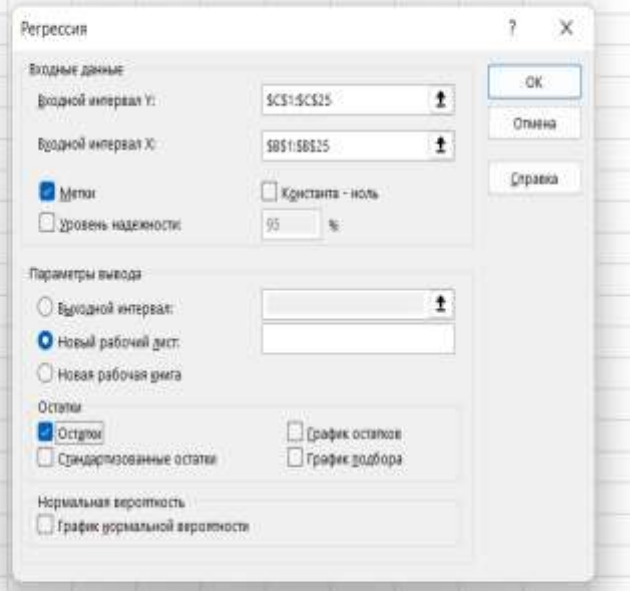
1. **the Dannye** tab, in the **Analysis** group, **Analyze dannyx** is pressed.
2. " Regression " and click " OK ".
3. " Regression " dialog box.

- the input range Y, which is the dependent variable, is selected. In this case, it is an umbrella trade (C1:C25).

- the independent variable X is selected in the desired range. In this case, it is the average monthly precipitation (B1: B25).

If a multiple regression model is being constructed, two or more adjacent columns with different independent variables are selected. If there are labels above the X and Y ranges, the Label box is checked. The output option of the result is selected, in this case a new sheet is selected. check the " Tip " box to get the difference between the predicted and actual values .

	A	B	C
1	**Oylar**	**Yog'ingarchilik (mm)**	**Soyabonlar soni**
2	yanvar	82	15
3	fevral	92,5	25
4	mart	83,2	17
5	aprel	97,7	28
6	may	131,9	41
7	iyun	141,3	47
8	iyul	165,4	50
9	avgust	140	46
10	sentabr	126,7	37
11	oktabr	97,8	22
12	noyabr	86,2	20
13	dekabr	99,6	30
14	yanvar	87	14
15	fevral	97,5	27
16	mart	88,2	14
17	aprel	102,7	30
18	may	123	43
19	iyun	146,3	49
20	iyul	160	49
21	avgust	145	44
22	sentabr	131,7	39
23	oktabr	118	36
24	noyabr	91,2	20
25	dekabr	104,6	32



Click OK and see the regression results generated by Excel:

	A	B	C	D	E	F	G	H	I	J	K
1	ВЫВОД ИТОГОВ										
2											
3	рессионная статистика										
4	Множеств	0,957667									
5	R-квадрат	0,917126									
6	Нормиров	0,913359									
7	Стандартн	3,581414									
8	Наблюден	24									
9											
10	Дисперсионный анализ										
11		df	SS	MS	F	начимость F					
12	Регрессия	1	3122,775	3122,775	243,4623	2,22E-13					
13	Остаток	22	282,1835	12,82652							
14	Итого	23	3404,958								
15											
16	Коэффициен	артная ош	атистика t	-Значение	ижние 95%	ерхние 95%	жние 95,0	рхние 95,0%			
17	Y-пересеч	-19,0741	3,372182	-5,65631	1,09E-05	-26,0676	-12,0806	-26,0676	-12,0806		
18	Yog'ingarc	0,450001	0,02884	15,60328	2,22E-13	0,39019	0,509812	0,39019	0,509812		
19											

Regression analysis: Let's interpret the results:

As you can see, doing regression in Excel is easy because all the calculations are done automatically. Interpreting the results is a bit more difficult because you need to know what is behind each number. Below is a breakdown of the four main parts of the regression analysis results. Regression analysis output: Extracting the results. In this part, it is determined how well the calculated linear regression equation fits the original data.

1	ВЫВОД ИТОГОВ	
2		
3	*Регрессионная статистика*	
4	Множественный R	0,957666798
5	R-квадрат	0,917125697
6	Нормированный R-квадрат	0,913358683
7	Стандартная ошибка	3,58141382
8	Наблюдения	24
9		

Here's what each piece of information means:

Multiple R: It is a correlation coefficient that measures the strength of the linear relationship between two variables. The correlation coefficient can have any value from -1 to 1, and its absolute value indicates the strength of the relationship.

R- square . It is used as a coefficient of determination and an indicator of the degree of compatibility. This shows how many points there are on the regression line.

Normalized R- squared . This is the R square adjusted for the number of independent variables in the model. This value can be used instead of R-squared for multiple regression analysis.

Standard more This is another good indicator of the accuracy of the regression analysis: the lower the number, the more reliable the regression equation.

Nabludeniya . It is simply the number of observations in the model.

Result of Regression Analysis: Analysis of Variance

The second part of the analysis of variance results:

Дисперсионный анализ					
	df	*SS*	*MS*	*F*	*Значимость F*
Регрессия	1	3122,774784	3122,774784	243,4622625	2,21604E-13
Остаток	22	282,1835489	12,82652495		
Итого	23	3404,958333			

Essentially, it breaks down the sum of squares into individual components that provide information about the levels of variability in your regression model:

df is the number of degrees of freedom associated with the scattering sources.

SS is sum of squares. The smaller the residual SS compared to the total SS, the better your model fits the data.

MS is mean square.

F is the F statistic or F test of the null hypothesis. It is used to test the overall significance of the model.

Znachimost is the P-value of FF.

Regression analysis results: coefficients

This section contains specific information about the components of your analysis:

Regression analysis results: coefficients

This section contains specific information about the components of your analysis:

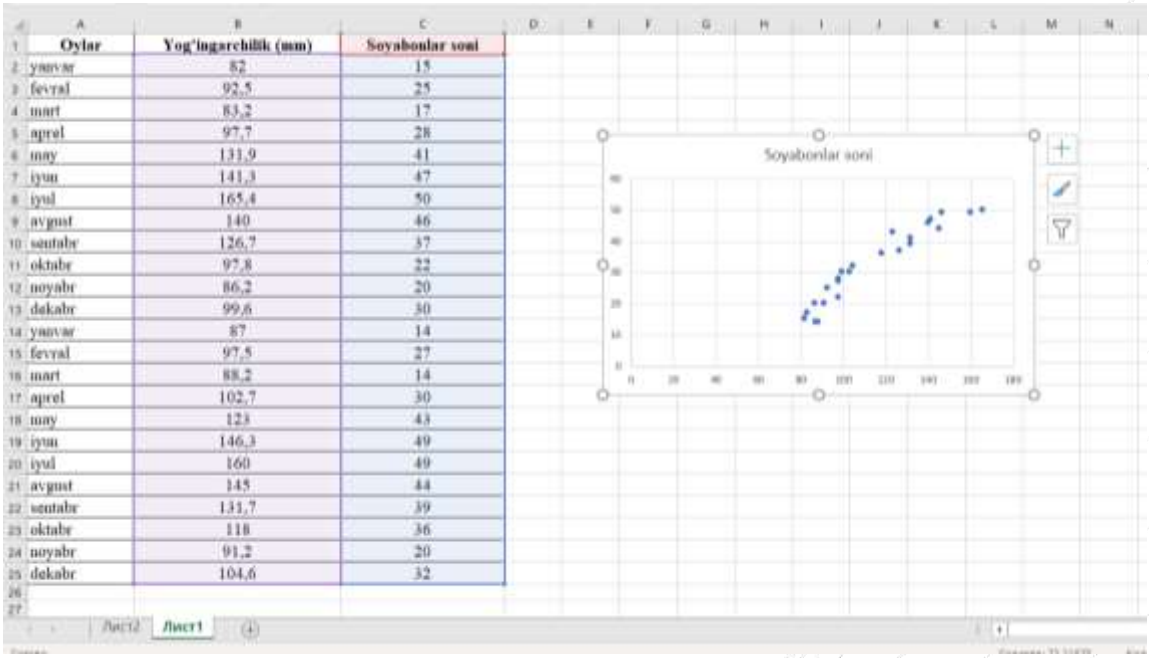
	Коэффициенты	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*	*Верхние 95%*	*Нижние 95,0%*	*Верхние 95,0%*
Y-пересечение	-19,07410899	1,372182168	-5,056310378	1,0929E-05	-26,06758677	-12,08063122	-26,06758677	-12,08063122
Yog'ingan'chlik (mm)	0,45000132	0,02884018	15,6032773	2,21604E-13	0,390190448	0,509812192	0,390190448	0,509812192

The most useful component in this section is the odds. This allows you to create a linear regression equation in Excel: **y = bx + a** . For our data set, where y is the number of umbrellas sold and x is the average monthly rainfall, our linear regression formula is:

Y = Rainfall Coefficient * x + Intercept.

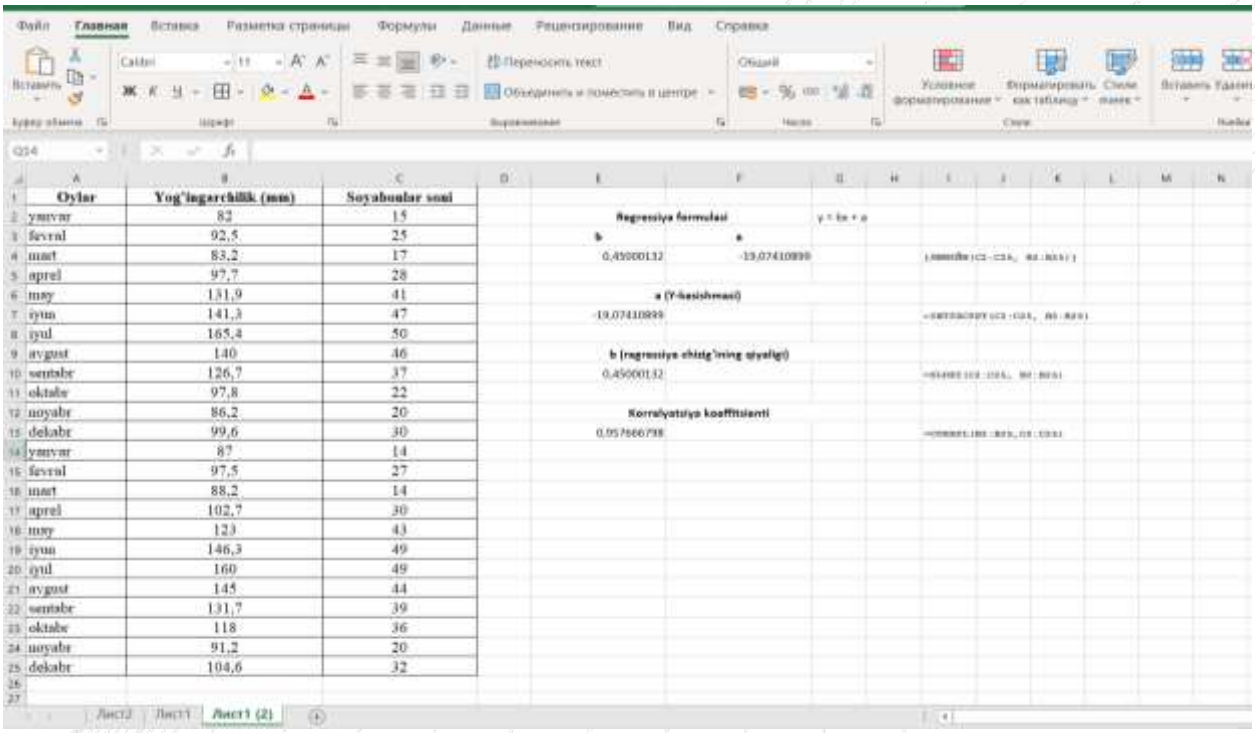
How to draw a linear regression graph in Excel?

To illustrate the relationship between two variables, a linear regression diagram is drawn. As a result, we get the following table:



Oylar	Yog'ingarchilik (mm)	Soyaboular soni
yanvar	82	15
fevral	92,5	25
mart	83,2	17
aprel	97,7	28
may	131,9	41
iyun	141,3	47
iyul	165,4	50
avgust	140	46
sentabr	126,7	37
oktabr	97,8	22
noyabr	86,2	20
dekabr	99,6	30
yanvar	87	14
fevral	97,5	27
mart	88,2	14
aprel	102,7	30
may	123	43
iyun	146,3	49
iyul	160	49
avgust	145	44
sentabr	131,7	39
oktabr	118	36
noyabr	91,2	20
dekabr	104,6	32

After performing the necessary steps, we will have the following:



Oylar	Yog'ingarchilik (mm)	Soyaboular soni
yanvar	82	15
fevral	92,5	25
mart	83,2	17
aprel	97,7	28
may	131,9	41
iyun	141,3	47
iyul	165,4	50
avgust	140	46
sentabr	126,7	37
oktabr	97,8	22
noyabr	86,2	20
dekabr	99,6	30
yanvar	87	14
fevral	97,5	27
mart	88,2	14
aprel	102,7	30
may	123	43
iyun	146,3	49
iyul	160	49
avgust	145	44
sentabr	131,7	39
oktabr	118	36
noyabr	91,2	20
dekabr	104,6	32

CONCLUSION.

Summary, regression analysis is a set of statistical methods used to evaluate the relationship between one or more variables. It can also be used to assess the relationship between variables and to model future relationships between them. In other words, regression analysis is important in every aspect of economics. In particular, it plays an important role in predicting how the future results will be realized in the exchange rate, state budget, taxes, finance, investment and other important areas of the economy, as well as in taking measures in advance. As a result, economists

and others will be limited in their work from various negative consequences. In short, regression analysis is a part of economics.

REFERENCES

1. https://en.wikipedia.org/wiki/Regression_analysis
2. https://corporatefinanceinstitute.com/resources/data-science/nonlinear-regression/
3. https://www.sciencedirect.com/topics/social-sciences/multiple-linear-regression
4. Anarbayeva F. "Regression analysis" textbook.