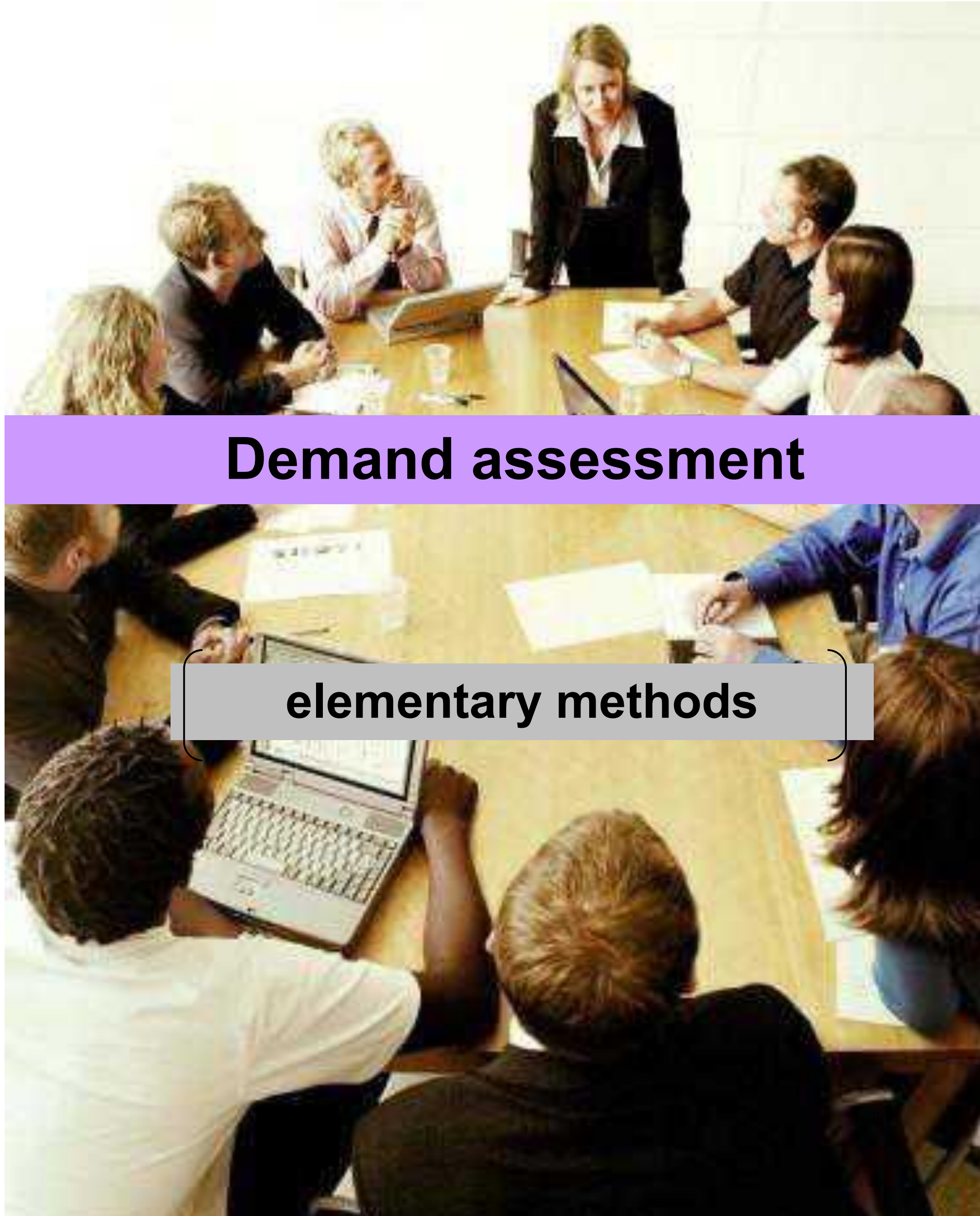


Demand assessment

elementary methods

2 directions in demand assessment

statistical analysis

market intelligence

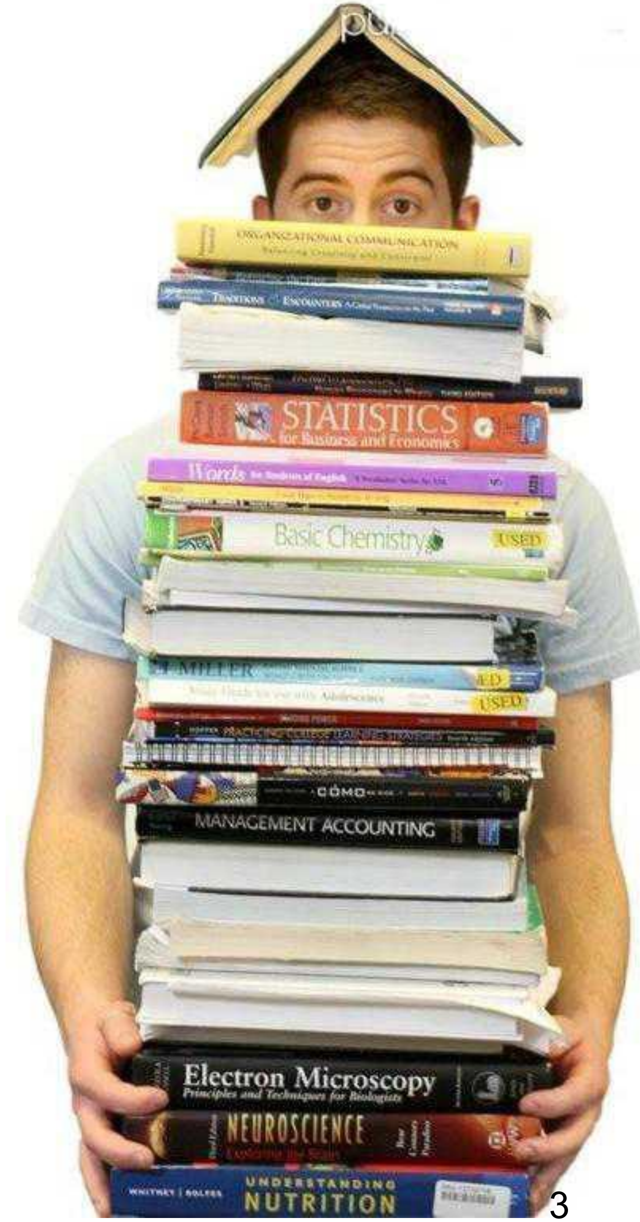
✔Statistical analysis

Steps:

1) Collection, validation and assessment of data

2) The choice of the information curve

3) Verification and evaluation of the selected curve



✔ **Statistical analysis**

1) Collection, validation and assessment of data

- **time series**
- **cross-sectional data**

✔ **Statistical analysis**

1) Collection, validation and assessment of data

 time series

Examine time changes in the demand for certain types of goods or services and the corresponding time changes in pricing, sales volume and other independent variables that affect the demand

time series

Long time period

Adjustment of necessary information

in order to avoid effects such as inflation

Deflationary correction: divide all nominal figures by the consumer price index and multiplied by 100. Get "regular money" base period

And also it is necessary to take into account changes in population, accounting for seasonal and cyclical fluctuations

✔ **Statistical analysis**

1) Collection, validation and assessment of data

cross-sectional data

Considered changing the variables from some set in a particular time

A snapshot of the many variables in one certain time

Ex: In order to determine the effect of prices on demand, as a variable can be selected volume of sales for a particular month,

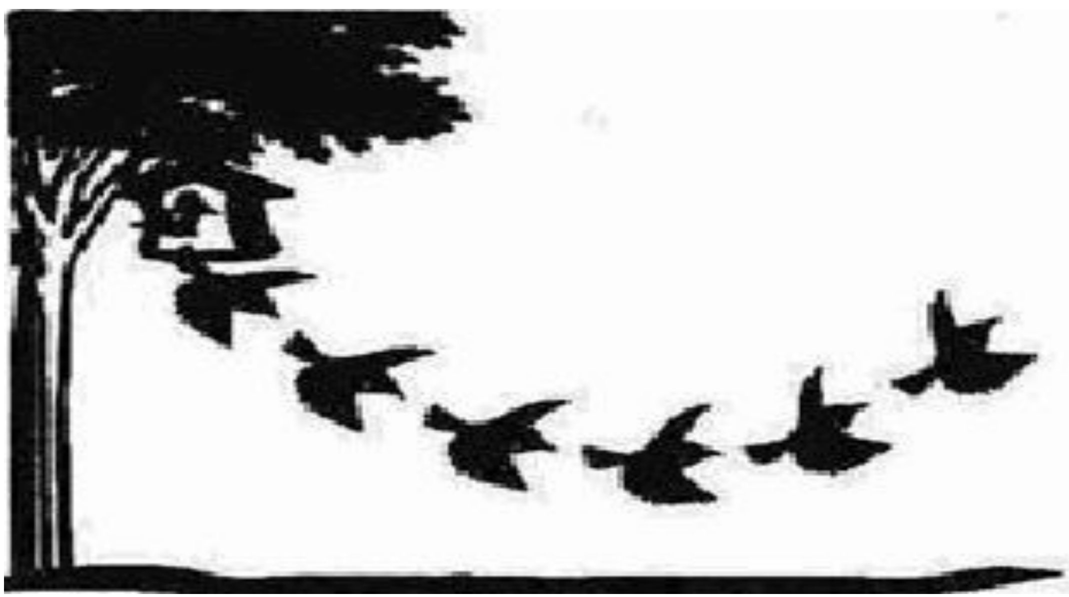
while the set may include a list of firms producing the product

✔ **Statistical analysis**

2) The choice of the information curve

The results of the observations are used to estimate the parameters of demand function

This function can then be used to predict values for the dependent variable for known values of the independent variables

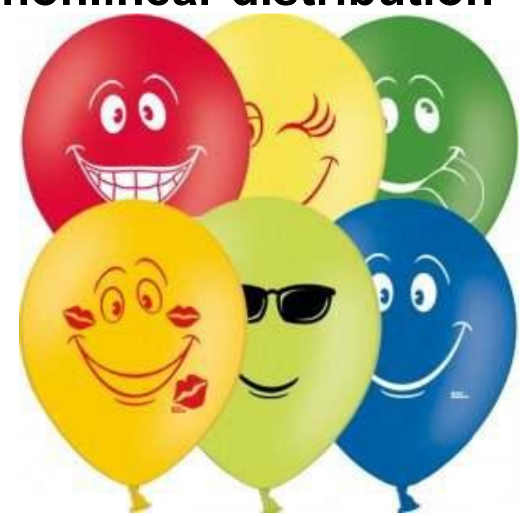


When choosing a curve there are two main questions:

1. **What type of equation it is necessary to use?**
2. **How the selected function fits and predicts the demand?**

The choice of the equation depends on two conditions:

а) the number of independent variables and б) the distribution of the data, i.e. linear or nonlinear distribution



If the trend of the experimental values of the dependent variable is approximately linear, and there are many independent variables, the estimated equation is:

constant value

The coefficients of the independent variables

$$\hat{Q} = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_X X_X$$

The estimated demand for the product

The value of the independent variable

If the data can be reduced to a single independent variable (e.g. price) and the trend is almost linear than to find the formula for this straight line we can use simple (pair) regression analysis

The equation thus is:

A constant value (which determines the point of intersection of the graph of the function with the Y axis)

$$Q_X = a + bP_x$$

The unit price of X (independent variable)

The quantity X,

(dependent variable)

The regression coefficient for Px (defining the slope of a line on the graph of a function)

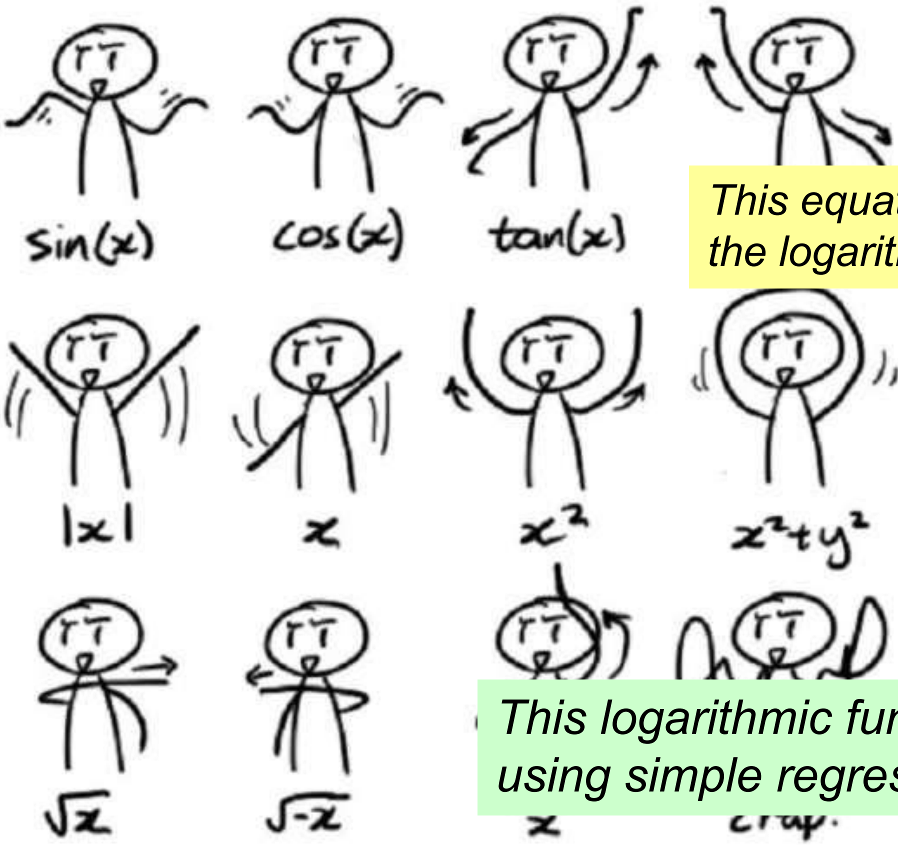
If the trend of the dependent variable is nonlinear and the function has a single independent variable, it is described by the equation:

$$Q_x = aP_x^b$$

This equation can be written as the logarithm, if you find the logarithm of both parts

$$\log Q_x = \log a + b \log P_x$$

This logarithmic function is linear and can be estimated using simple regression analysis



simple linear regression

STEP 1. Data collection

Таблица 7.1

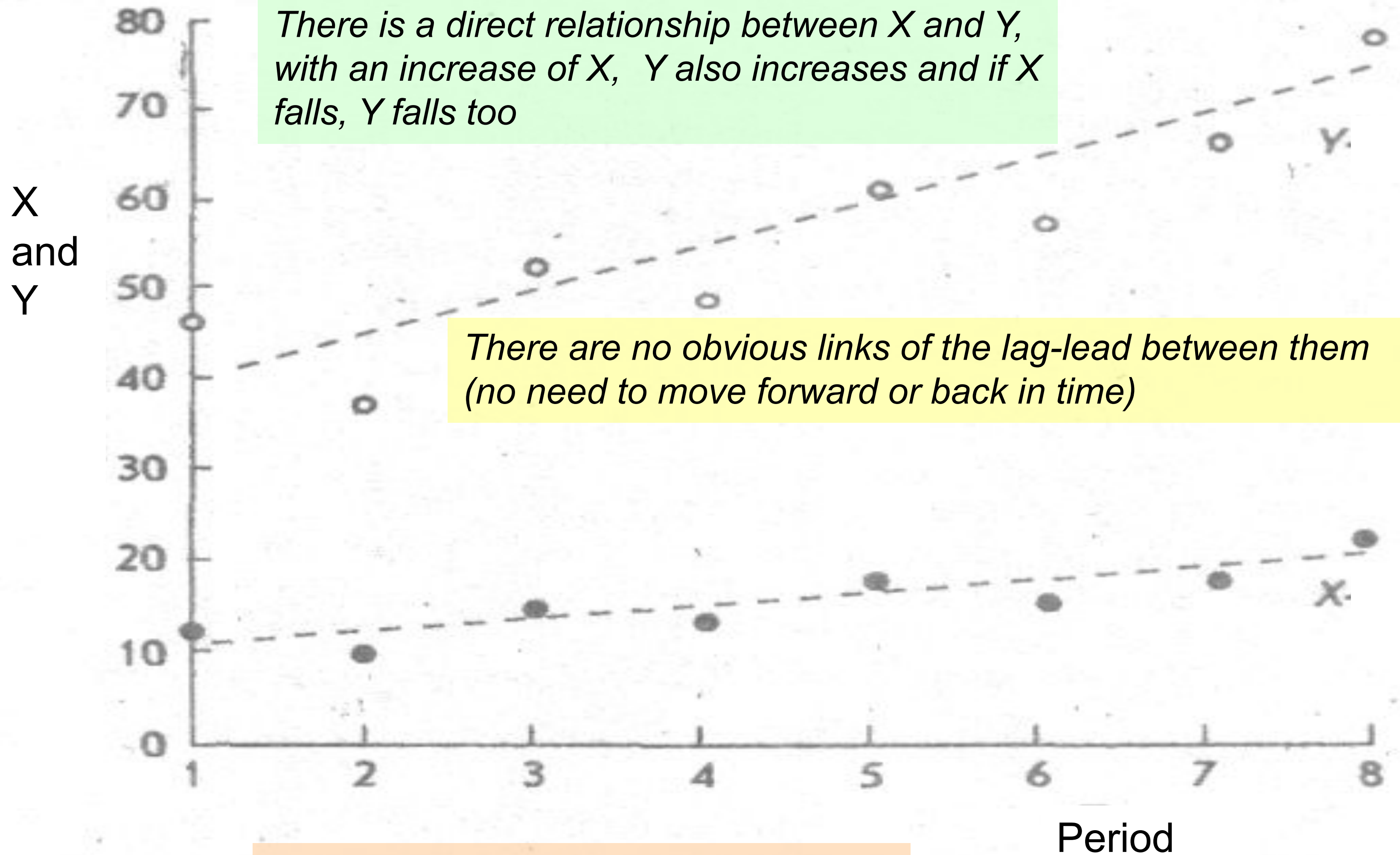
Collect time series data

Period	Observation X	Observation Y
1	12	47
2	10	38
3	15	55
4	14	49
5	19	60
6	17	56
7	20	66
8	25	80

TASK: TO FIND THE REGRESSION FUNCTION for THESE DATA!

STEP 2. Organization variables in time

There is a direct relationship between X and Y, with an increase of X, Y also increases and if X falls, Y falls too



There are no obvious links of the lag-lead between them (no need to move forward or back in time)

the trend, allocated to each series, is linear

STEP 3. Organization of a scatter plot

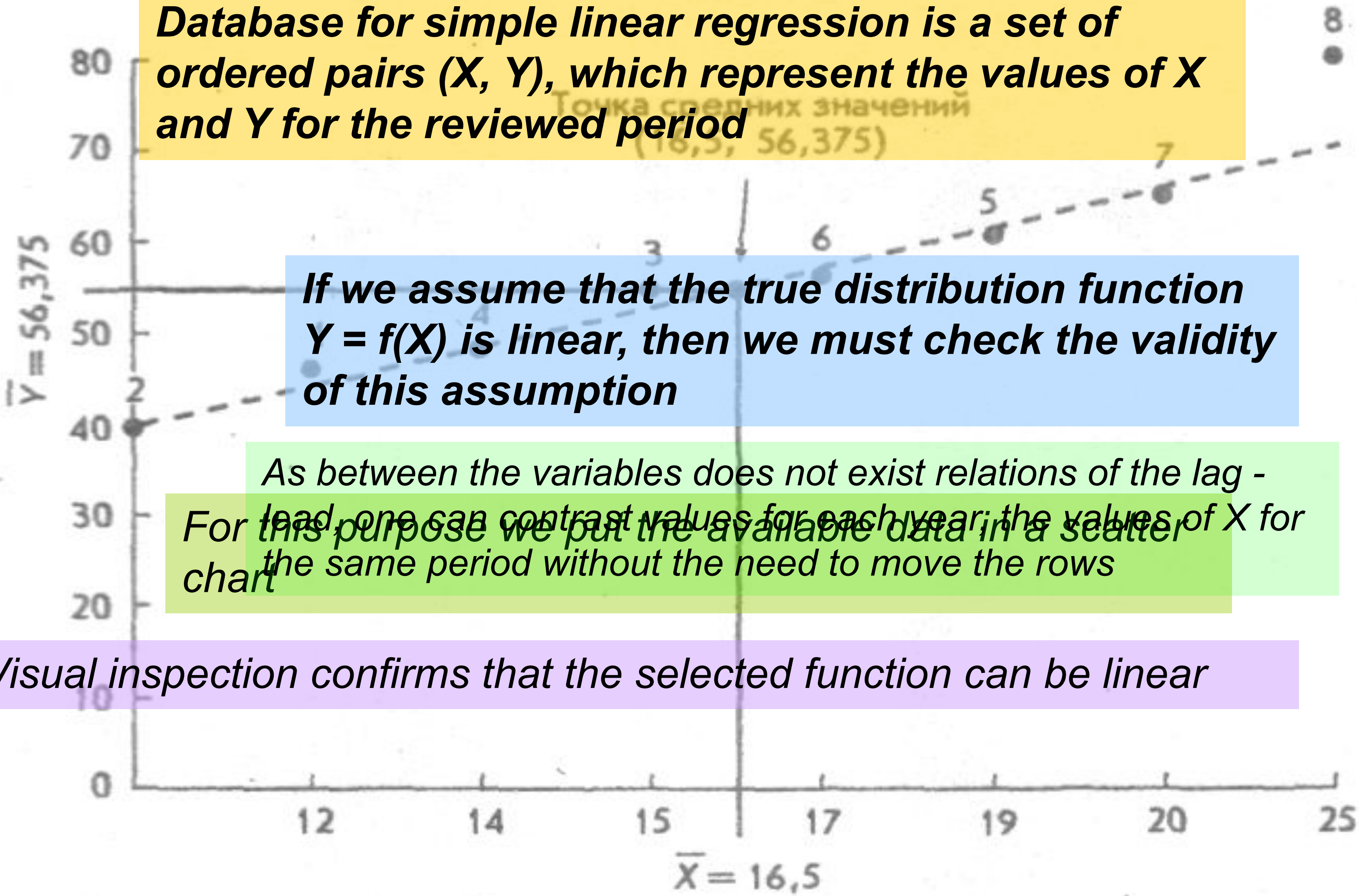
Database for simple linear regression is a set of ordered pairs (X, Y), which represent the values of X and Y for the reviewed period

If we assume that the true distribution function Y = f(X) is linear, then we must check the validity of this assumption

As between the variables does not exist relations of the lag - lead, one can contrast values for each year; the values of X for the same period without the need to move the rows

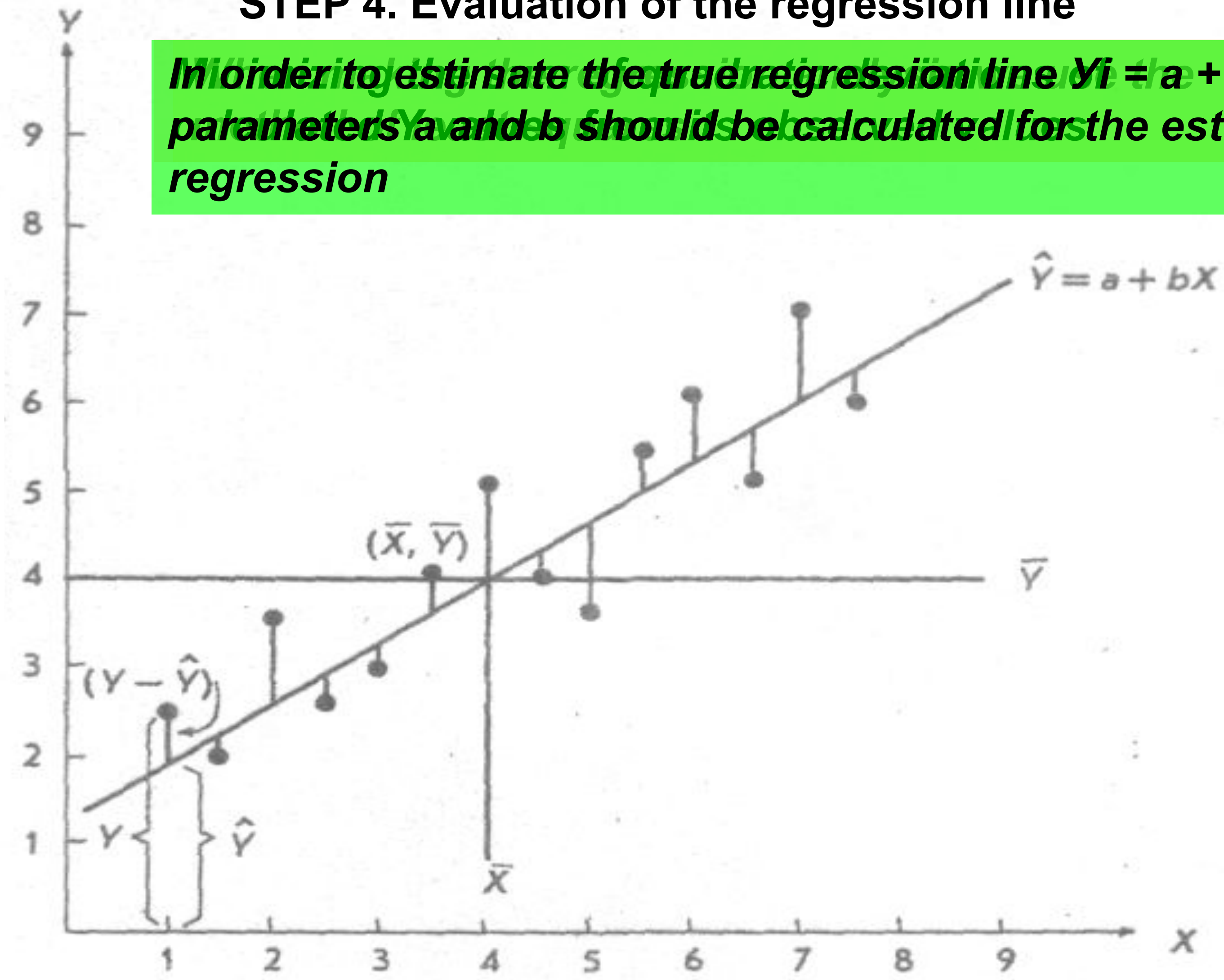
For this purpose we put the available data in a scatter chart

Visual inspection confirms that the selected function can be linear



STEP 4. Evaluation of the regression line

In order to estimate the true regression line Yi = a + b Xi, parameters a and b should be calculated for the estimated regression



simple linear regression

STEP 4. Evaluation of the regression line

$$\hat{Y}_i = 13 + 2{,}6X_i$$

Period	Observation X	Observation Y	XY	X^2	Y^2
1	12	47			
2	10	38			
3	15	55			
4	14	49			
5	19	60			
6	17	56			
7	20	66			
8	25	80			
Sum	132	451			
Average	16,5	56,375			

STEP 5. Comparison of calculated and actual values

How well our estimated regression equation describes Y as a function of X?

Compare the actual and estimated value

Таблица 7.4

Наблюденные и вычисленные значения функции $Y = f(X)$

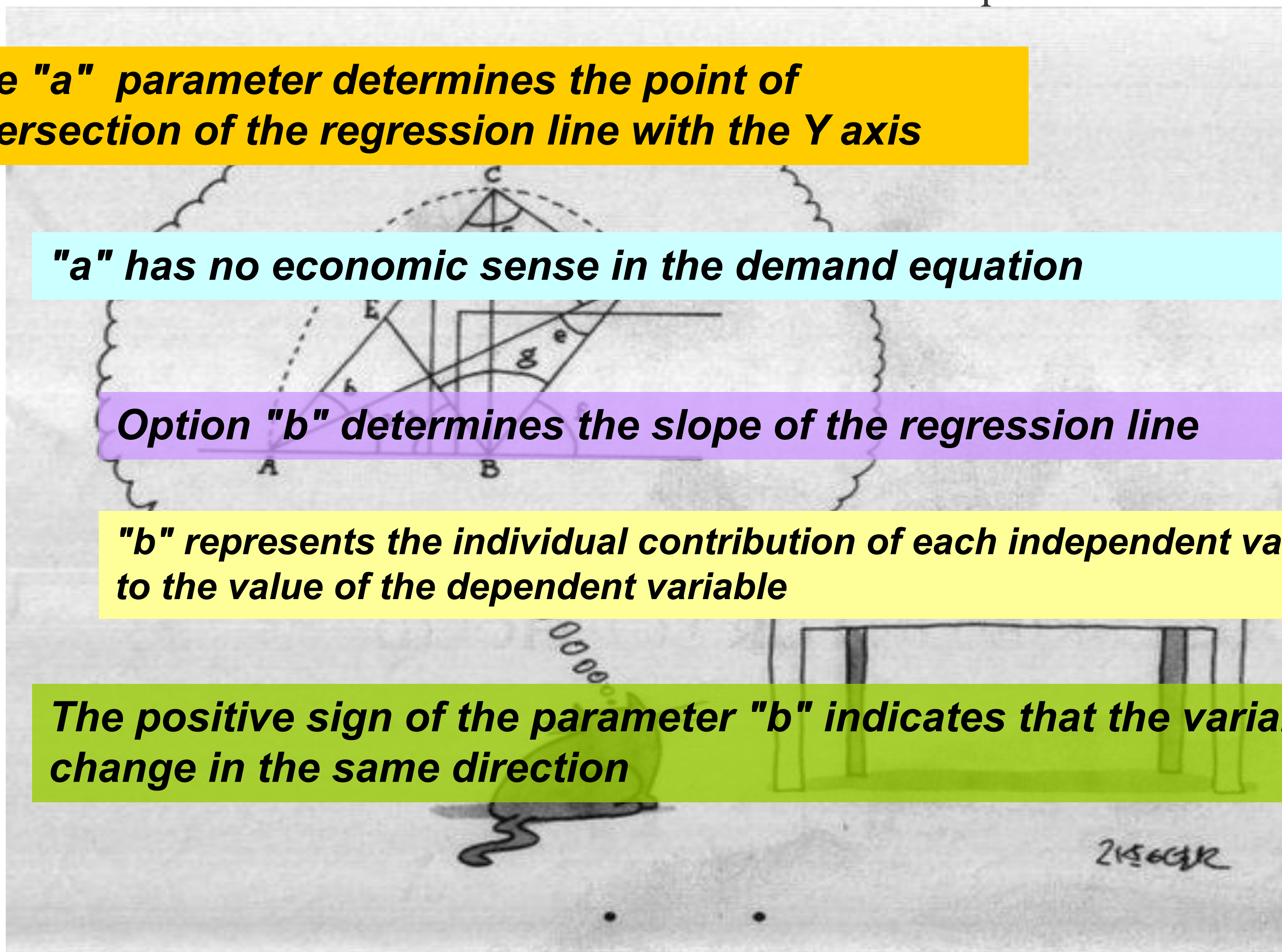
Initial X	Initial Y	Estimated function $Y_i = 13 + 2{,}6X$	Deviation
10	38	39,0	+ 1,0
[illegible]	[illegible]	[illegible]	− 2,8
[illegible]	[illegible]	[illegible]	+ 0,4
[illegible]	[illegible]	[illegible]	− 3,0
[illegible]	[illegible]	[illegible]	+ 1,2
19	60	62,4	+ 2,4
[illegible]	[illegible]	[illegible]	[illegible]

The deviation of the actual values from the calculated values: the results of all observations do not fit on the regression line

The fact that the observations deviate from the regression line indicates that the magnitude of Y is effected also by forces different from X

Interpretation of parameters

$$\hat{Y}_i = 13 + 2{,}6X_i$$



The "a" parameter determines the point of intersection of the regression line with the Y axis

"a" has no economic sense in the demand equation

Option "b" determines the slope of the regression line

"b" represents the individual contribution of each independent variable to the value of the dependent variable

The positive sign of the parameter "b" indicates that the variables change in the same direction

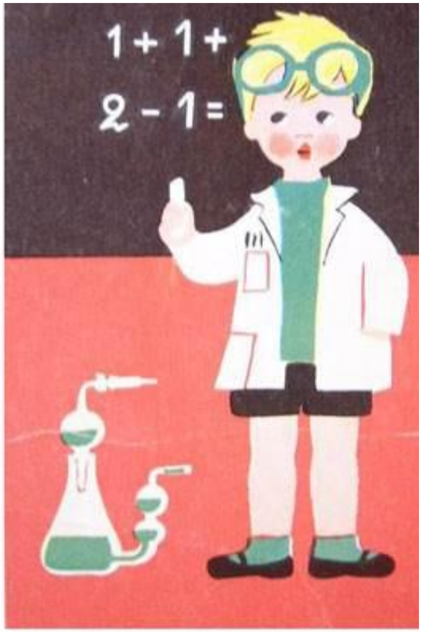
Evaluation of the regression equation

The goal of linear regression evaluation: to get a linear equation, which can be used to determine the values of the independent variable Y on any existing values of the independent variable X

How informative or accurate the determined $\hat{Y}$ is?

When analyzing simple regression use two statistical indicators:

- **The root - mean - square error of the estimation, S_e;**
- **The coefficient of determination, r^2, and its square root, r, which is called the correlation coefficient.**



1) **The root – mean - square error of the estimation, Se;**

Represents the deviation of experimental points from the estimated regression line (determines the variance of random Y values)

1) The root - mean - square error of the estimation, Se;

Observed Y for X_i

Evaluated Y for X_i

$$S_e = \sqrt{\frac{\Sigma(Y_i - \hat{Y}_i)^2}{n-k-1}}$$

Root-mean-square error

Number of observations

Number of independent variables

1) **Root-mean-square error, Se;**

$$S_e = \sqrt{\frac{\Sigma(Y_i - Y_i)^2}{n-k-1}}$$

If Se = 0, than the estimated equation fits perfectly the observed data (all points lie on the regression line)

The more root-mean-square error is, the greater the range of deviations are

2) coefficient of determination, r^2

Shows how well the regression model describes the variation of the dependent variable

EX: if r^2 = 0,975, than approximately 97.5% of the changes in the dependent variable explained by the variation of the independent variable X

Values can range from 0 to 1 or from 0 to 100%

0 - there is no relationship between the variables,

1 - the regression line is perfect (all changes are explained by changes in X)

3) the correlation coefficient, r,

Determines the degree of connection between variables

$$-1 < r > 1$$