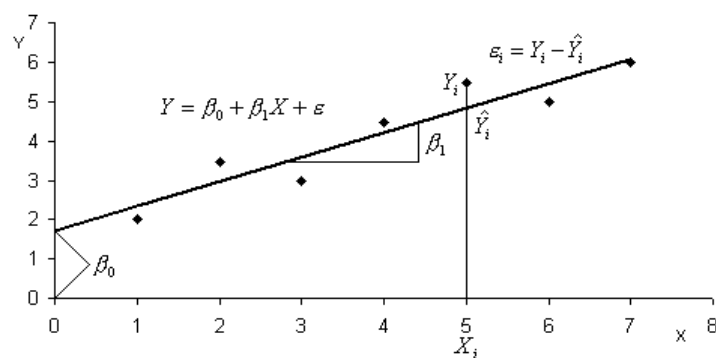


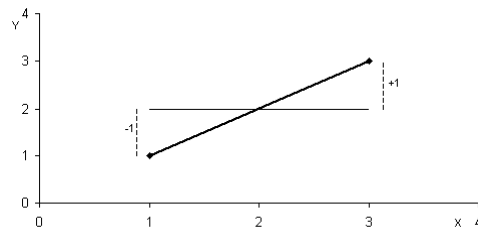
REGRESSÃO E CORRELAÇÃO



AJUSTE DE UMA RETA



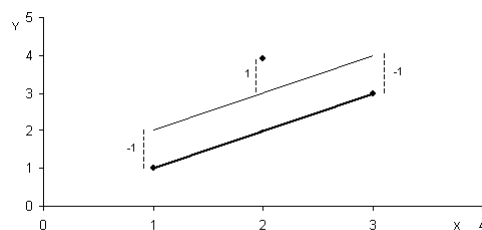
MINIMIZAÇÃO DOS DESVIOS



$$\sum (Y_i - \hat{Y}_i)$$

3

MINIMIZAÇÃO DOS DESVIOS ABSOLUTOS



$$\sum |Y_i - \hat{Y}_i|$$

4



EXEMPLO 1

- Considere o seguinte conjunto de pontos

X	Y
1	1
2	1
3	2
4	2
5	4

5



RETAS DE AJUSTE

R1 $Y = -0.1 + 0.7X$

R2 $Y = 0.5 + 0.5X$

R3 $Y = -0.7 + 0.9X$

6

RETAS



R1	R2	R3
0.6	1	0.2
1.3	1.5	1.1
2	2	2
2.7	2.5	2.9
3.4	3	3.8

7

DESVIOS



Desv1	Desv2	Desv3
0.4	0	0.8
-0.3	-0.5	-0.1
0	0	0
-0.7	-0.5	-0.9
0.6	1	0.2
0	0	0

8

DESVIOS ABSOLUTOS



Desv1	Desv2	Desv3
0.4	0	0.8
0.3	0.5	0.1
0	0	0
0.7	0.5	0.9
0.6	1	0.2
2	2	2

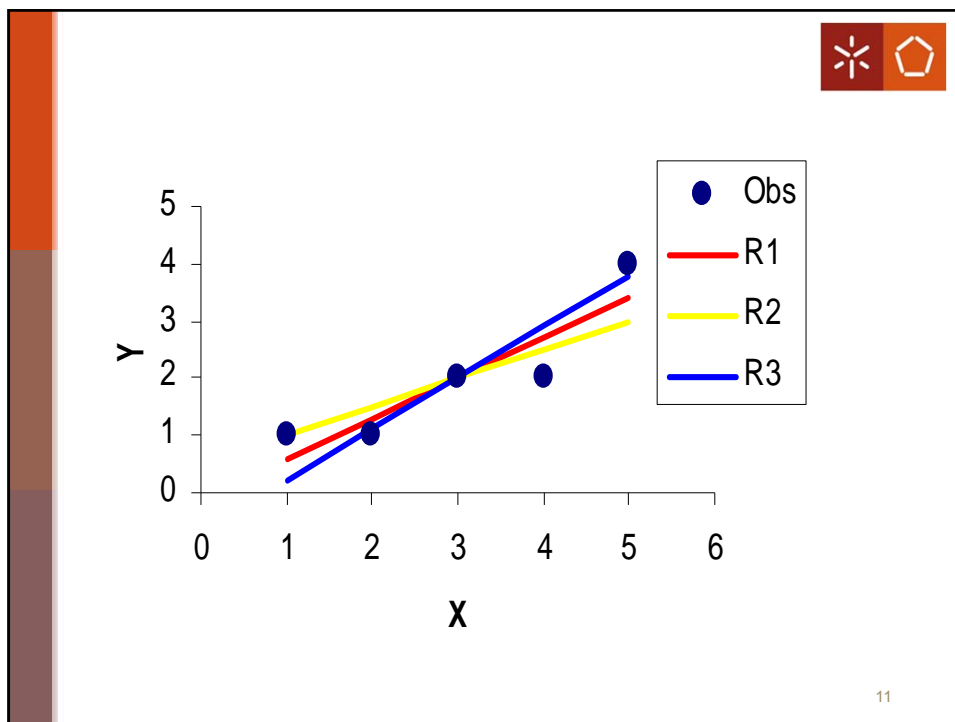
9

QUADRADO DOS DESVIOS



$(\text{Desv1})^2$	$(\text{Desv2})^2$	$(\text{Desv3})^2$
0.16	0	0.64
0.09	0.25	0.01
0	0	0
0.49	0.25	0.81
0.36	1	0.04
1.10	1.50	1.50

10



11

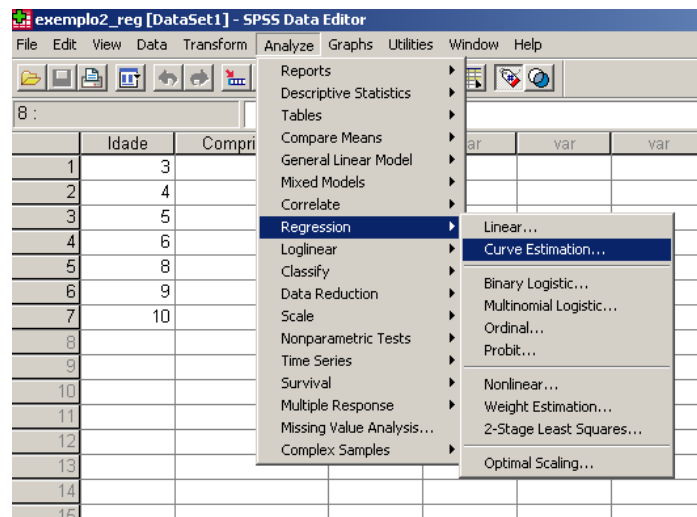
EXEMPLO 2

- Comprimento alar (cm) em função da idade (dias) para andorinhas

Dias	Comp.
3	1,4
4	1,5
5	2,1
6	2,4
8	3,1
9	3,2
10	3,3

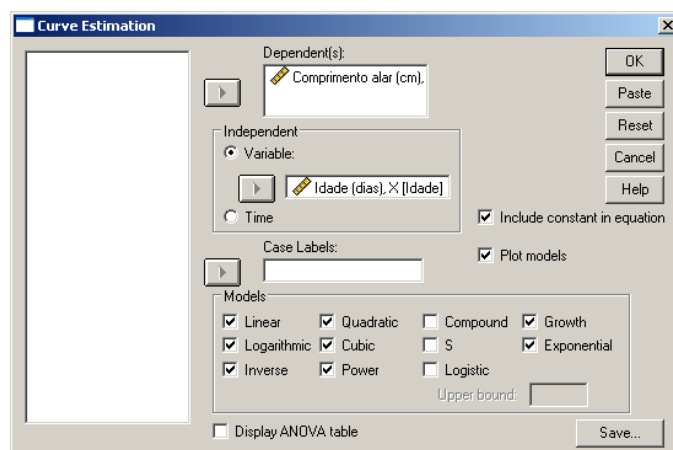
12

EXEMPLO 2



13

EXEMPLO 2



14

EXEMPLO 2

Output 1 - SPSS Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Window Help

Output

- Log
- Curve Fit
- Title
- Notes
- Model Description
- Case Processing Summary
- Variable Processing Summary
- Model Summary and Parameter Estimates
- Curve Fit for Comprimento

* CURVE ESTIMATION.
TSST NEWAR=NONE.
CURVE FIT /VARIABLES=Comprimento WITH Idade
/CONSTANT
/MODEL=LINEAR LOGARITHMIC INVERSE QUADRATIC CUBIC POWER GROWTH EXPONENTIAL
/PLOT FIT.

Curve Fit

Model Description		
Model Name		MOD_2
Dependent Variable	1	Comprimento alar (cm), Y
Equation	1	Linear
	2	Logarithmic
	3	Inverse
	4	Quadratic
	5	Cubic
	6	Power ^a
	7	Growth ^a
	8	Exponential ^a
Independent Variable		Idade (dias), X
Constant		Included
Variable Whose Values Label Observations in Plots		Unspecified
Tolerance for Entering Terms in Equations		,0001

^a The model requires all non-missing values to be positive.

EXEMPLO 2

Output 1 - SPSS Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Window Help

Output

- Log
- Curve Fit
- Title
- Notes
- Model Description
- Case Processing Summary
- Variable Processing Summary
- Model Summary and Parameter Estimates
- Curve Fit for Comprimento

Variable Processing Summary

	Variables	
	Dependent Comprimento alar (cm), Y	Independent Idade (dias), X
Number of Positive Values	7	7
Number of Zeros	0	0
Number of Negative Values	0	0
Number of Missing Values	0	0
User-Missing	0	0
System-Missing	0	0

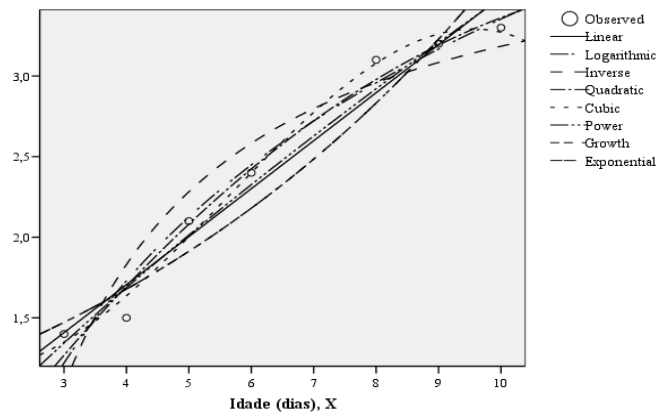
Model Summary and Parameter Estimates

Dependent Variable: Comprimento alar (cm), Y

Equation	R Square	F	Model Summary			Parameter Estimates			
			df1	df2	Sig.	Constant	b1	b2	b3
Linear	,964	132,174	1	5	,000	,515	,298		
Logarithmic	,971	165,753	1	5	,000	-,727	1,772		
Inverse	,915	53,833	1	5	,001	4,087	-,026		
Quadratic	,980	99,695	2	4	,000	-,274	,579	-,021	
Cubic	,981	106,896	3	3	,002	1,471	-,387	,141	-,00
Power	,968	149,638	1	5	,000	,563	,792		
Growth	,931	87,190	1	5	,000	-,006	,131		
Exponential	,931	87,190	1	5	,000	,994	,131		

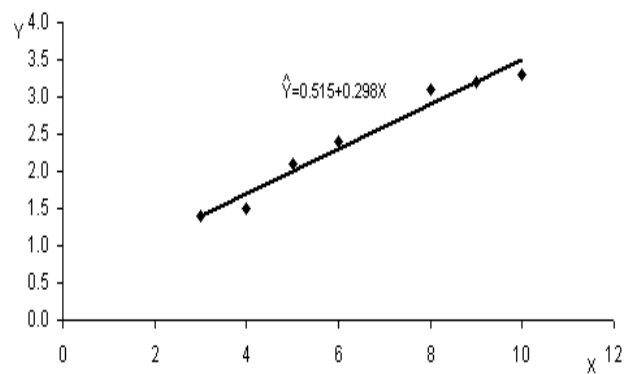
The independent variable is Idade (dias), X.

EXEMPLO 2



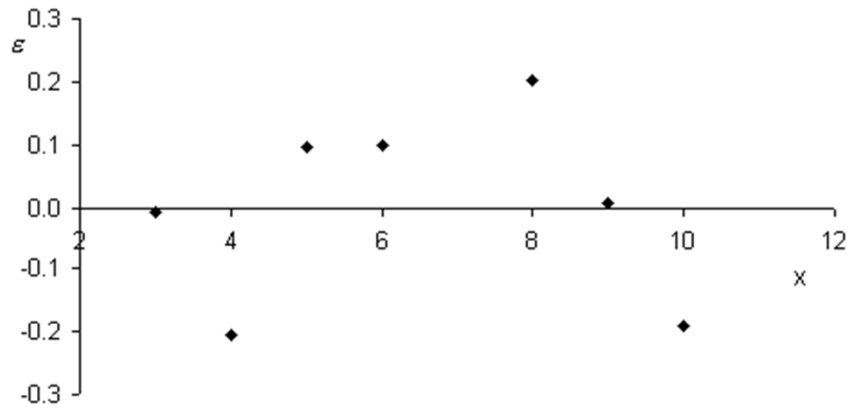
17

RETA DE MÍNIMOS QUADRADOS



18

RESÍDUOS



19

Estimadores

$$Y_i = \beta_0 + \beta_1 (X_i - \bar{X}) + \varepsilon_i \quad i = 1, \dots, n$$

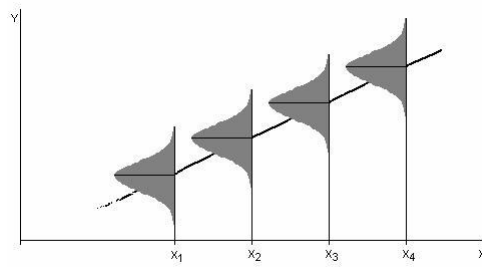
$$\beta_0 \quad \hat{\beta}_0 = \frac{1}{n} \sum_i Y_i = \bar{Y}$$

$$\beta_1 \quad \hat{\beta}_1 = \frac{\sum_i (X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{\sum_i (X_i - \bar{X})^2} = \frac{s_{XY}}{s_{XX}}$$

$$\sigma^2 \quad s^2 = \frac{1}{n-2} \sum_i \hat{\varepsilon}_i^2 = \frac{1}{n-2} \sum_i \left\{ Y_i - \left[\hat{\beta}_0 + \hat{\beta}_1 (X_i - \bar{X}) \right] \right\}^2$$

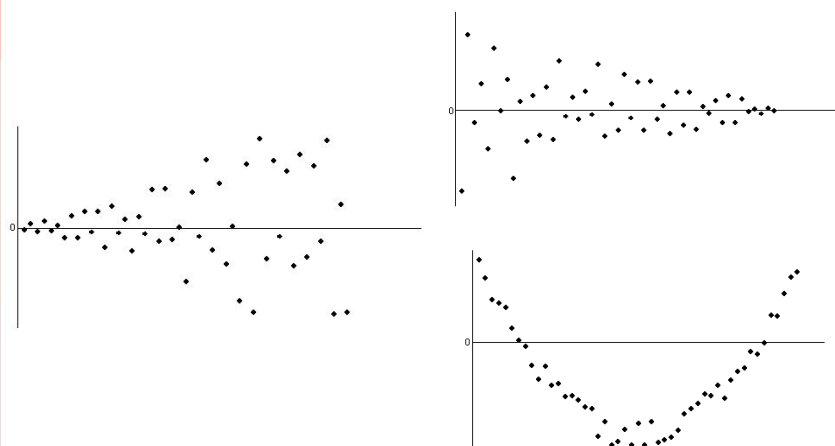
20

DISTRIBUIÇÃO DOS ERROS



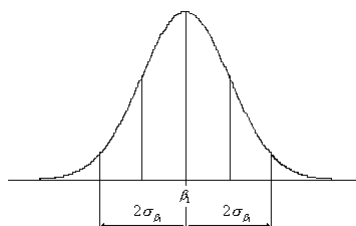
21

RESÍDUOS



22

DISTRIBUIÇÃO DO DECLIVE



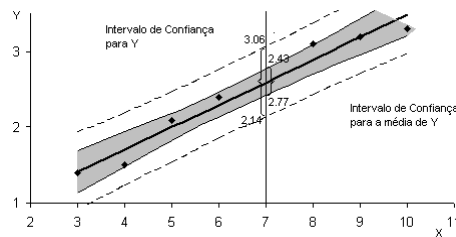
23

IC e Testes de hipóteses

	IC	TH
β_0	$\hat{\beta}_0 \pm t_{n-2, (\alpha/2)} \cdot \frac{s}{\sqrt{n}}$	$H_0 : \beta_0 = b_0$ $H_1 : \beta_0 \neq b_0, \beta_0 > b_0 \text{ ou } \beta_0 < b_0$ $ET = \frac{\hat{\beta}_0 - b_0}{s / \sqrt{n}}$ $H_0 \text{ verdadeira} \Rightarrow ET \sim t_{n-2}$
β_0'	$(\hat{\beta}_0 - \bar{X} \cdot \hat{\beta}_1) \pm t_{n-2, (\alpha/2)} \cdot s \cdot \sqrt{\frac{1}{n} + \frac{\bar{X}^2}{s_{XX}}}$	$H_0 : \beta_0' = b_0'$ $H_1 : \beta_0' \neq b_0', \beta_0' > b_0' \text{ ou } \beta_0' < b_0'$ $ET = \frac{(\hat{\beta}_0 - \bar{X} \cdot \hat{\beta}_1) - b_0'}{s \cdot \sqrt{\frac{1}{n} + \frac{\bar{X}^2}{s_{XX}}}}$ $H_0 \text{ verdadeira} \Rightarrow ET \sim t_{n-2}$
β_1	$\hat{\beta}_1 \pm t_{n-2, (\alpha/2)} \cdot \frac{s}{\sqrt{s_{XX}}}$	$H_0 : \beta_1 = b_{10}$ $H_1 : \beta_1 \neq b_{10}, \beta_1 > b_{10} \text{ ou } \beta_1 < b_{10}$ $ET = \frac{\hat{\beta}_1 - b_{10}}{\frac{s}{\sqrt{s_{XX}}}}$ $H_0 \text{ verdadeira} \Rightarrow ET \sim t_{n-2}$

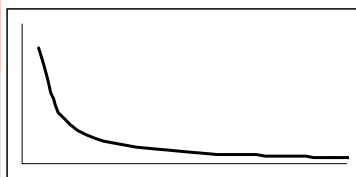
24

INTERVALO DE CONFIANÇA

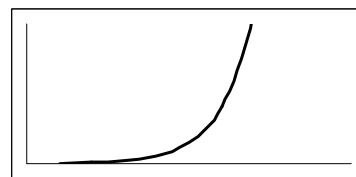


25

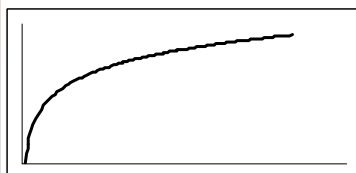
REGRESSÃO NÃO LINEAR



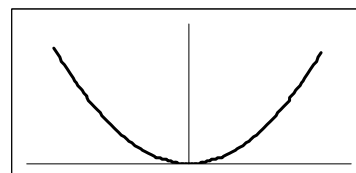
$$\hat{Y} = \beta_0 + \beta_1 \frac{1}{X}$$



$$\hat{Y} = \beta_0 + \beta_1 e^X$$



$$\hat{Y} = \beta_0 + \beta_1 \ln X$$



$$\hat{Y} = \beta_0 + \beta_1 X^2$$

26



REGRESSÃO NÃO LINEAR

Modelo	Transformação
<ul style="list-style-type: none">$Y_i = \alpha' + \frac{\beta}{X_i} + e_i$	$U_i = \frac{1}{X_i}$ $Y_i = \alpha' + \beta.U_i + e_i$
<ul style="list-style-type: none">$Y_i = e^{\alpha' + \beta.X_i + e_i}$	$Z_i = \ln Y_i$ $Z_i = \alpha' + \beta.X_i + e_i$
<ul style="list-style-type: none">$Y_i = e^{\frac{\alpha' + \beta}{X_i} + e_i}$ com $\alpha' > 0, \beta < 0$	$U_i = \frac{1}{X_i}$ $Z_i = \ln Y_i$ $Z_i = \alpha' + \beta.U_i + e_i$

27



COEFICIENTE DE CORRELAÇÃO

Coeficiente de correlação de Pearson

$$R = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} = \frac{s_{XY}}{\sqrt{s_{XX}} \cdot \sqrt{s_{YY}}}$$

28



TESTES DE ASSOCIAÇÃO

Unilateral à direita

$$H_0 : \rho = 0$$

$$H_1 : \rho > 0$$

Unilateral à esquerda

$$H_0 : \rho = 0$$

$$H_1 : \rho < 0$$

Bilateral

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

Estatística de teste

$$t = \frac{r \cdot \sqrt{n-2}}{\sqrt{1-r^2}}$$

Região de Rejeição:

$$t > t_{n-2,(\alpha)}$$

$$t < -t_{n-2,(\alpha)}$$

$$|t| > t_{n-2,(\alpha/2)}$$

29



EXEMPLO

- Índice de Desenvolvimento de Griffiths
- avaliações motora e intelectual para 9 crianças com a idade de 4 anos

Motor	Intelectual
84	77
73	85
101	105
74	86
88	108
100	116
86	96
95	100
82	100

30

exemplo1_reg [DataSet0] - SPSS Data Editor

	Motor	Intellectual	var
1	84	77	
2	73	85	
3	101	105	
4	74	86	
5	88	108	
6	100	116	
7	86	96	
8	95	100	
9	82	100	
10			
11			
12			
13			
14			
15			
16			
17			
18			

Scatter/Dot...

Simple Scatter, Matrix Scatter, Simple Dot, Overlay Scatter, 3-D Scatter

31

Simple Scatterplot

Y Axis: Motor

X Axis: Intellectual

Set Markers by:

Label Cases by:

Panel by

Rows:

Columns:

Template

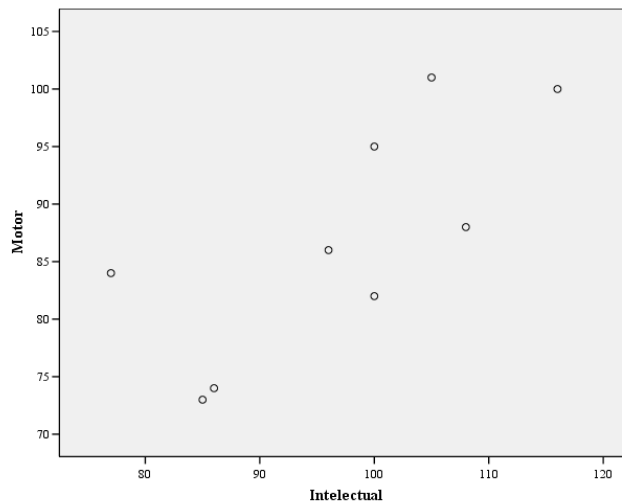
Use chart specifications from:

File...

Titles... Options...

32

DIAGRAMA DE DISPERSÃO



33

exemplo1_reg [DataSet0] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : Intelectual

	Motor	Intelectual
1	84	7
2	73	8
3	101	10
4	74	8
5	88	10
6	100	11
7	86	9
8	95	10
9	62	10
10		
11		
12		
13		

Correlate

- Bivariate...
- Partial...
- Distances...

Bivariate Correlations

Variables:

- Motor
- Intelectual

Correlation Coefficients

☒ Pearson ☐ Kendall's tau-b ☐ Spearman

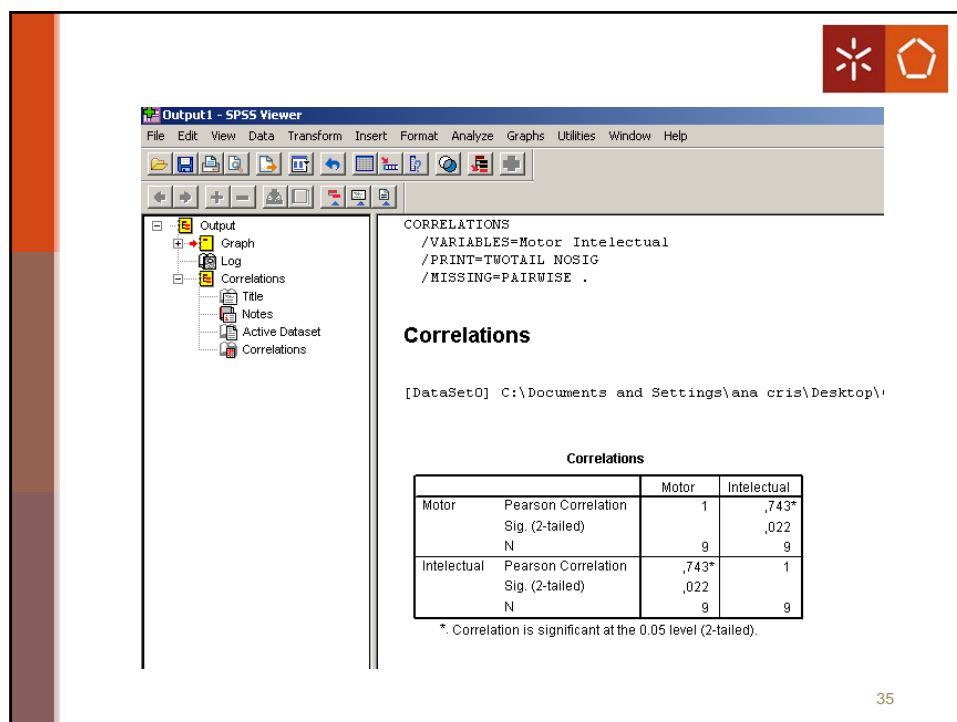
Test of Significance

☒ Two-tailed ☐ One-tailed

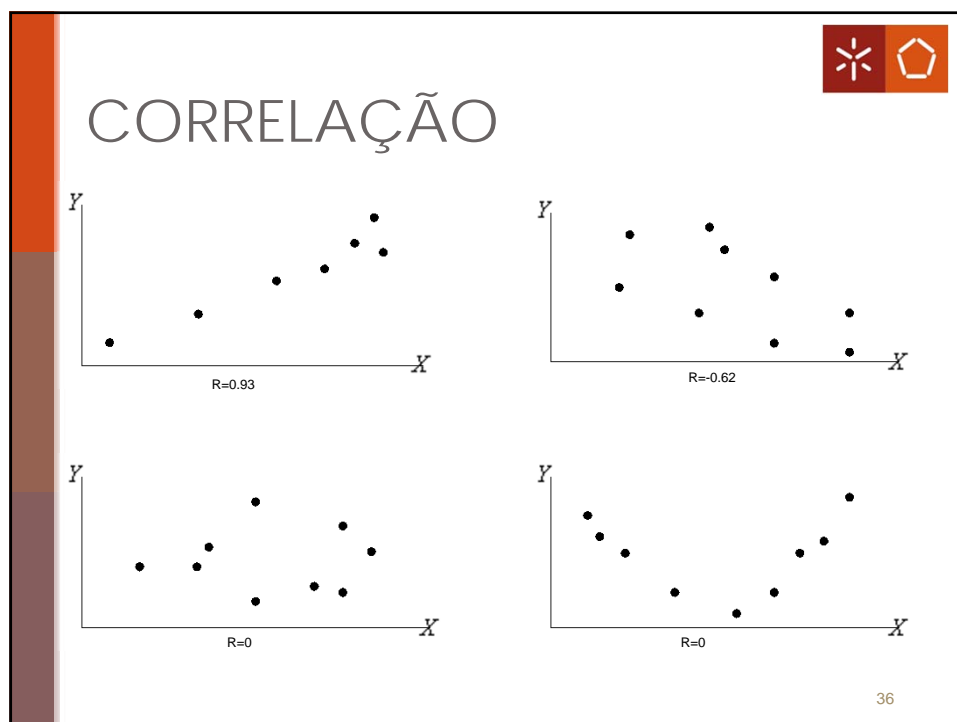
☒ Flag significant correlations

Options...

34



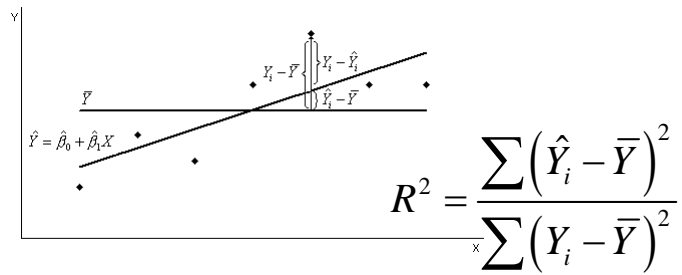
35



36



COEFICIENTE DE DETERMINAÇÃO



37



Coeficiente de determinação (r^2), representa a proporção da variação de Y que é explicada pela regressão

$$r^2 = \frac{\hat{\beta}_1^2 \cdot s_{XX}}{s_{YY}} = \frac{\hat{\beta}_1^2 \cdot \sum_i (X_i - \bar{X})^2}{\sum_i (Y_i - \bar{Y})^2} = \frac{\text{variação de } Y \text{ explicada pela regressão}}{\text{variação total de } Y}$$

38