

## Nové možnosti rozvoje vzdělávání na Technické univerzitě v Liberci

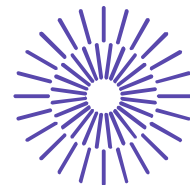
Specifický cíl A3: Tvorba nových profesně zaměřených studijních programů

NPO\_TUL\_MSMT-16598/2022



## Téma 9: Příklad 1 – regresní analýza

Ing. Vladimíra Hovorková Valentová, Ph.D.



### Zadání příkladu:

Data v tabulce udávají poptávku a cenu určitého spotřebního zboží:

Cena	100	110	140	160	170	200
Poptávané množství	120	89	56	41	35	22

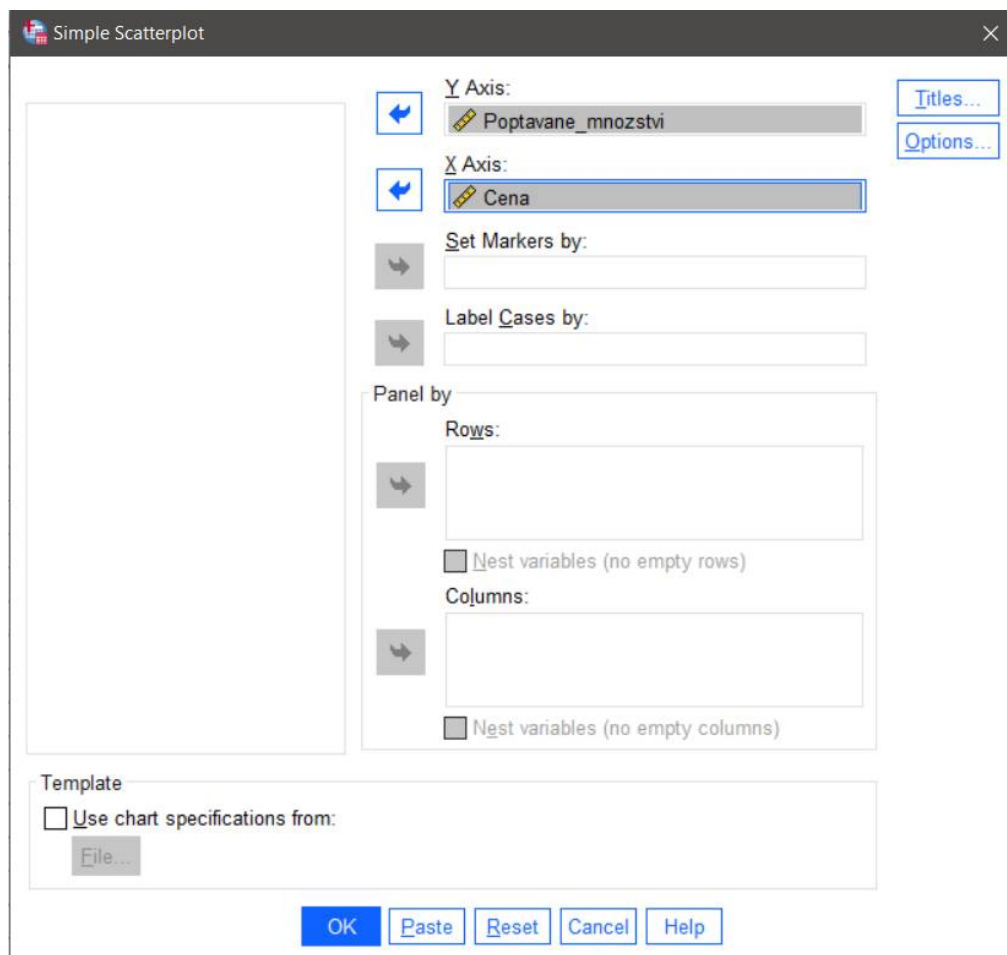
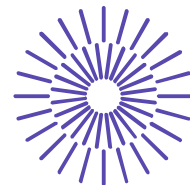
Zvolte vhodný regresní model popisující závislost poptávky na ceně. Ověřte na hladině významnosti 5 %, že vaše volba byla správná.

### Řešení příkladu:

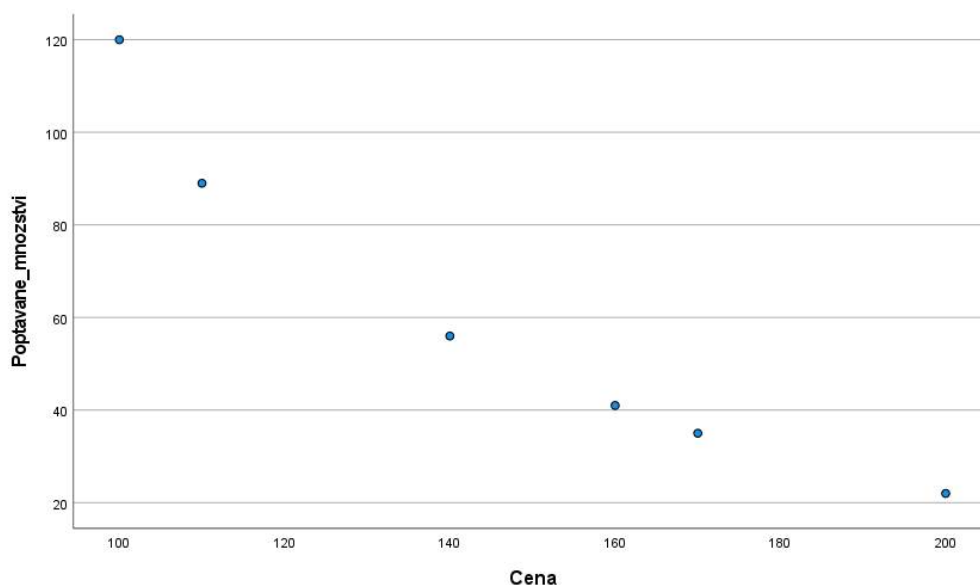
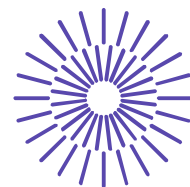
Vzhledem k vyšší časové náročnosti všech výpočtů je příklad řešen jen prostřednictvím programu SPSS. Data o proměnných zadáme do dvou samostatných sloupců:

Cena	Poptavane_ mnozstvi
100	120
110	89
140	56
160	41
170	35
200	22

Abychom si udělali představu o průběhu dané závislosti, zobrazíme si data do bodového diagramu. Ten najdeme v posloupnosti procedur **Graphs – Legacy Dialogs – Scatter/Dot**. Vybereme **Simple Scatter**. Vstupní panel vyplníme podle následujícího vzoru:



Výsledkem zadání dat do vstupního panelu je bodový diagram – viz níže.



Podle uspořádání bodů v bodovém diagramu by se dalo soudit, že vhodnou funkcí může být regresní exponenciální funkce. Následně tedy zvolíme **Analyze – Regression – Curve Estimation**. Vstupní panel vyplníme takto:

Curve Estimation

Dependent(s): Poptavane\_mnozstvi

Independent: Variable: Cena

Case Labels: Include constant in equation Plot models

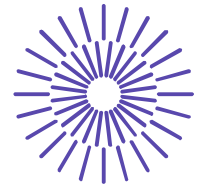
Models: Linear Quadratic Compound Growth Logarithmic Cubic S Exponential Inverse Power Logistic

Upper bound:

Display ANOVA table

OK Paste Reset Cancel Help





Parametry regresní exponenciály najdeme ve výstupu v tabulce *Correlations*:

Coefficients					
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Cena	-,016	,001	-,997	-26,884	<,001
(Constant)	6,360	,092		68,888	<,001

The dependent variable is  $\ln(\text{Poptavane\_mnozsvi})$ .

Rovnice regresní exponenciály je zde ve tvaru:  $Y = e^{b_0 + b_1 x}$ . Konkrétně to tedy bude  $Y = e^{6,360 - 0,016x}$ .

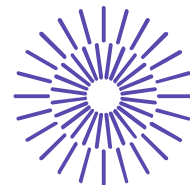
Nyní bychom měli ověřit, že jsme zvolili vhodnou funkci pro popis dané závislosti. Provedeme nejprve celkový F-test, který ověří vhodnost celého modelu a následně individuální t-testy o parametrech regresní exponenciály. Výsledek celkového F-testu najdeme v následujícím výstupu (tabulka ANOVA):

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	1,936	1	1,936	722,773	<,001
Residual	,011	4	,003		
Total	1,946	5			

The independent variable is Cena.

### Celkový F-test

- 1)  $H_0: \beta_0 = c, \beta_1 = 0$  (tj. exponenciála není vhodná k popisu závislosti *poptávky na ceně*)  
 $H_1: \text{non } H_0$
- 2)  $F=722,773$
- 3) Sig. = 0,000011 (konkrétní hodnotu, která je menší než 0,001 získáme dvojitým kliknutím na tabulku ANOVA a dvojitým kliknutím na hodnotu Sig.).
- 4) Sig.  $< \alpha \rightarrow$  zamítáme  $H_0$ , přijímáme  $H_1$ .
- 5) Na hladině významnosti 5 % jsme prokázali, že regresní exponenciála je vhodná pro popis dané závislosti.



Nyní provedeme individuální t-testy o nulové hodnotě regresních parametrů, jejichž výsledky jsou v tabulce Coefficients:

	Coefficients				
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Cena	-,016	,001	-,997	-26,884	<,001
(Constant)	6,360	,092		68,888	<,001

The dependent variable is ln(Poptavane\_mnozstvi).

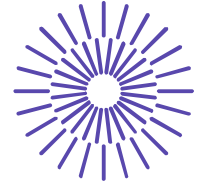
#### t-testy

- 1)  $H_0: \beta_0 = 0$   
 $H_1: \beta_0 \neq 0$
- 2)  $t = 68,888$
- 3)  $\text{Sig.} < 0,001$  ( $\text{Sig.} = 2,6605 \cdot 10^{-7}$ )
- 4)  $\text{Sig} < \alpha \rightarrow$  zamítáme  $H_0$ , přijímáme  $H_1$ .
- 5) Na hladině významnosti 5 % jsme prokázali, že parametr  $\beta_0$  je statisticky významný.

- 1)  $H_0: \beta_1 = 0$   
 $H_1: \beta_1 \neq 0$
- 2)  $t = -26,884$
- 3)  $\text{Sig.} < 0,001$  ( $\text{Sig.} = 0,000011$ )
- 4)  $\text{Sig} < \alpha \rightarrow$  zamítáme  $H_0$ , přijímáme  $H_1$ .
- 5) Na hladině významnosti 5 % jsme prokázali, že parametr  $\beta_1$  je statisticky významný.

Vzhledem k tomu, že je celkový F-test významný a oba t-testy také, můžeme regresní exponenciálu považovat za vhodnou k popisu závislosti poptávaného množství na ceně.

Otázkou ještě je, do jaké míry je tato funkce vhodná. To nám pomůže zodpovědět index determinace, který je mírou kvality regresních funkcí. Jeho hodnotu najdeme v tabulce Model Summary:

**Model Summary**

R	R Square	Adjusted R Square	Std. Error of the Estimate
,997	,994	,993	,052

The independent variable is Cena.

Je zde označený jako R square. Hodnota  $R^2 = 0,994$  nám říká, že 99,4 % z celkové variability poptávaného množství lze vysvětlit pomocí dané regresní exponenciály. Funkce je tedy velmi výstižná.