

Regresná Analýza

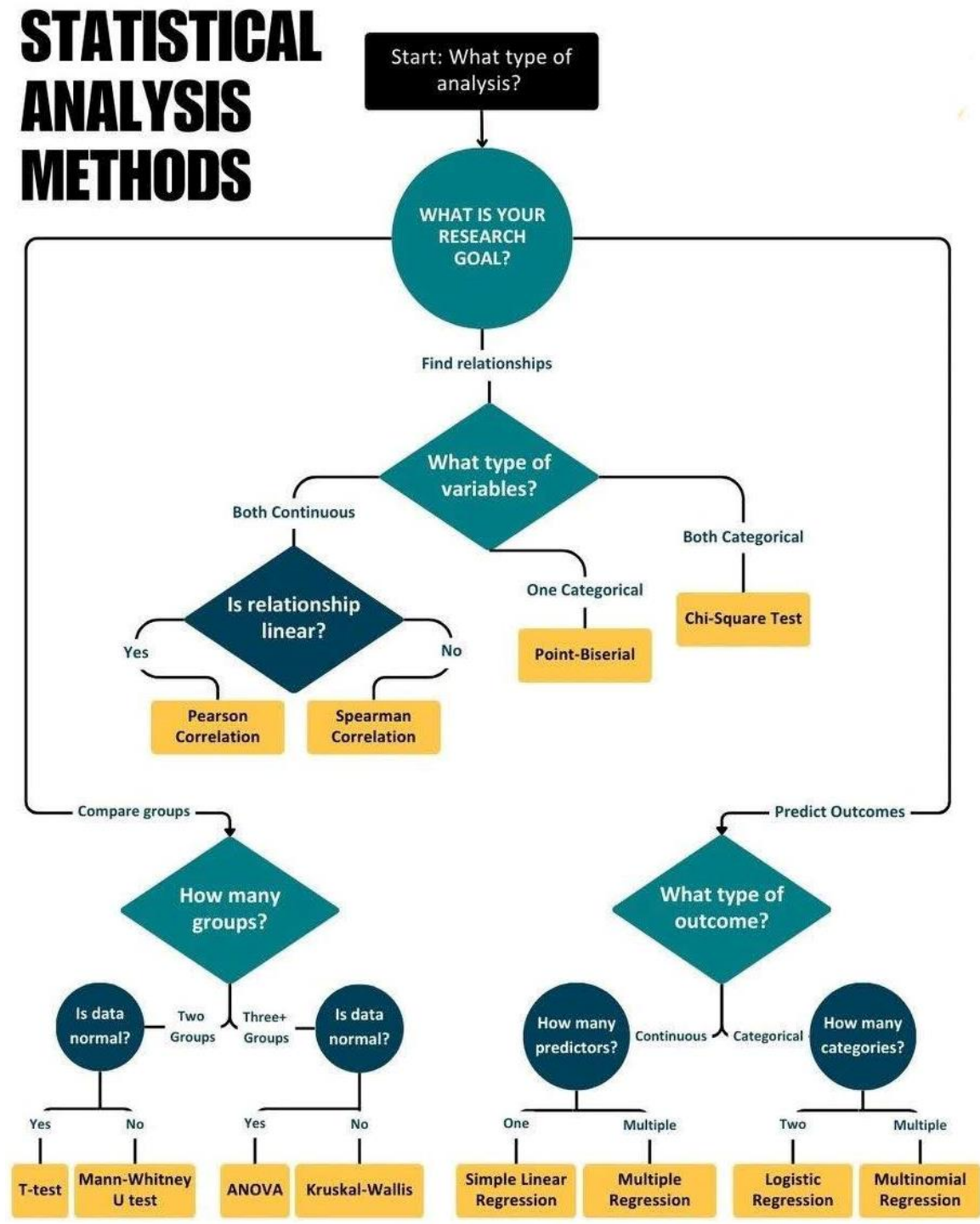




Čo sa naučíme?

1. Čo je to regresia?
2. Čo je to koeficient determinácie?
3. Ako sa používa a interpretuje?
4. Aké máme typy regresií?
5. Ako postupovať?
6. Ako používať regresiu v praxi?
7. Profit

Výber Štatistickej Metódy



Čo je to Regresia?



- **Štatistická technika**
- **Algoritmus** používaný na **predpovedanie** alebo **vizualizáciu** a **vzťah medzi dvoma rôznymi vlastnosťami/premennými**
- Na **modelovanie vzťahu** medzi **závislou premennou (predikovanou)** a **1 alebo viacerými nezávislými premennými (prediktormi)**
- Cieľom je **predpovedať hodnoty závislej premennej na základe hodnôt nezávislých premenných**

Regresný Koeficient Determin. R^2



- **Štatistická metrika**
- Vyjadruje, **ako dobre regresný model vysvetľuje rozptyl závislej premennej**
- **Do akej miery môžeme dôverovať modelu pri predpovedaní?**
- **Vlastnosti:**
 - Rozsah hodnôt: $0 \leq R^2 \leq 1$
 - $R^2 = 1$ model dokonale vysvetľuje všetky odchýlky (ideálny prípad)
 - $R^2 = 0$ model nič nevysvetľuje, predikcie sú ako náhodné
 - $R^2 \approx 0,7$ 70 % variability závislej premennej vysvetlených modelom
 - $R^2 < 0,3$ slabá vysvetľovacia schopnosť modelu
 - Ak $R^2 = 0,85 \rightarrow 85\%$ rozdielov v predajoch vysvetľuje reklama a len 15 % je spôsobených inými faktormi (napr. sezónnosťou, konkurenciou).

Regresný Koeficient Determin. R^2



- **Vlastnosti:**
 - Hodnota sa pohybuje v rozmedzí 0 až 1
 - $R^2 \approx 1 \rightarrow$ model veľmi dobre vysvetľuje variabilitu závislej premennej
 - $R^2 \approx 0 \rightarrow$ model nevysvetľuje žiadnu variabilitu, predikcia je náhodná
 - Vyjadruje podiel rozptylu vysvetleného modelom
 - Citlivý na počet nezávislých premenných a pretrénovanie modelu (overfitting)
 - Platí pre lineárnu aj viacnásobnú regresiu
- **Použitie:**
 - Hodnotenie presnosti predikčných modelov
 - Vyhodnocovanie lineárnych regresných vzťahov v štatistike
 - Odhad účinnosti vysvetľujúcich premenných pri predikcii výsledkov (napr. predaj vs. cena)
- **Príklad:**
 - Ak model predpovedá predaj produktu na základe výšky investície do reklamy a $R^2 = 0,85$, znamená to, že 85 % variability predaja možno vysvetliť výškou reklamy.
 - Zvyšných 15 % ovplyvňujú iné faktory (napr. konkurencia, sezóna).

Vlastnosť	r (korelácia)	R ² (determinácia)
Rozsah hodnôt	-1 až +1	0 až 1
Smer vzťahu	Áno (kladný alebo záporný)	Nie (vždy nezáporné)
Typ analýzy	Vzťah	Presnosť vysvetlenia/predikcie
Lineárnosť	Meria priamo	Odvodzuje sa z r ²
Použitie	Korelačná analýza	Regresná analýza

Regresný Koeficient Determin. R^2

IT a Data Science

✓ Veľkosť databázy a čas zálohovania

- Skúmanie vzťahu medzi veľkosťou databázy (v GB) a časom potrebným na zálohovanie (v min.)
- Ak $R^2 \approx 0,85 \rightarrow$ väčšie databázy majú lineárne vyšší čas zálohovania

✓ Počet commitov a výskyt bugov

- Analýza, či viac commitov (zmien v kóde) vedie k vyššiemu počtu bugov
- Ak $R^2 \approx 0,4 \rightarrow$ slabšia pozitívna lineárna závislosť

Marketing

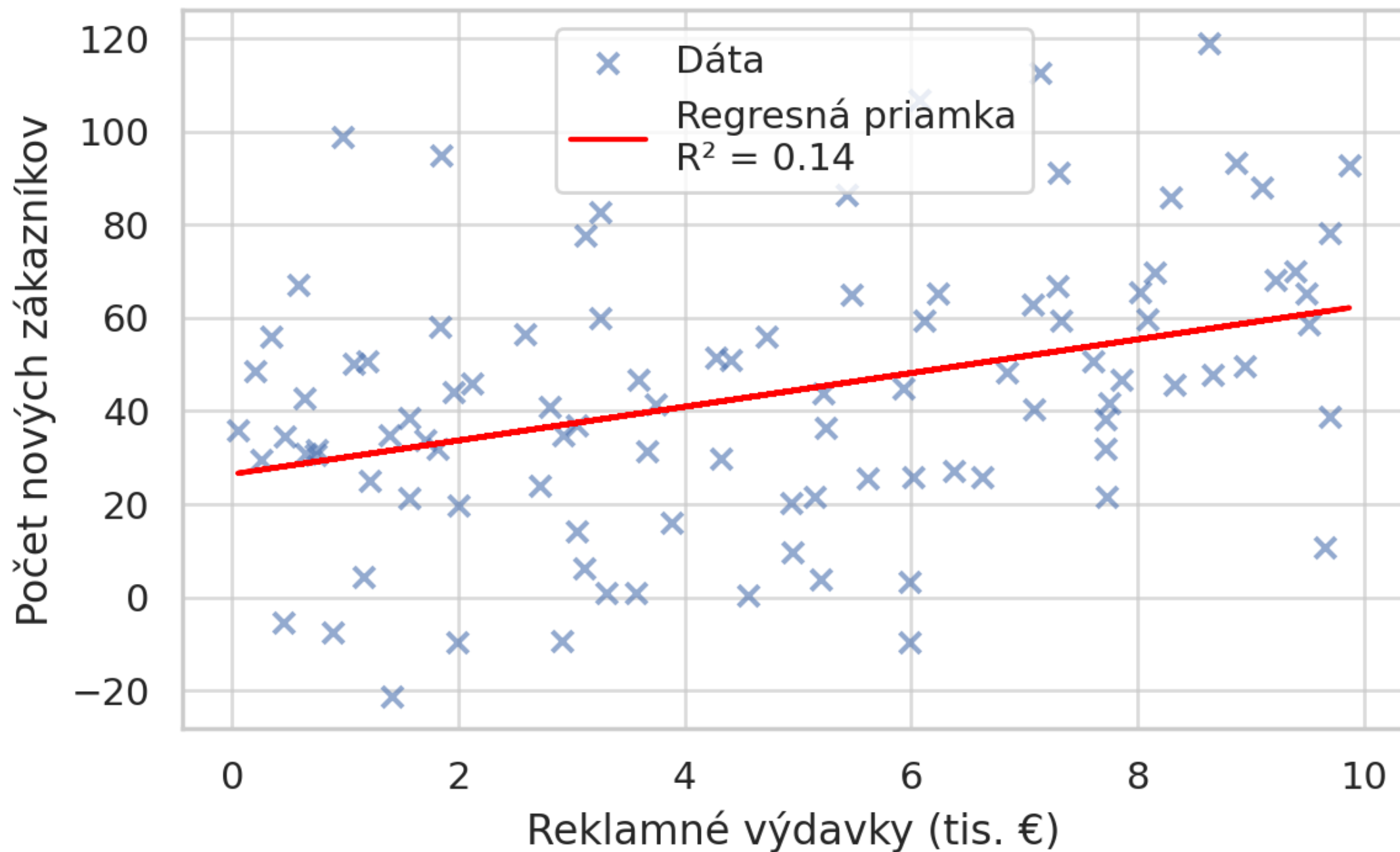
✓ Počet newsletterov a miera odhlásenia

- Skúmanie, či častejšie zasielanie newsletterov zvyšuje mieru odhlásenia
- Ak $R^2 \approx 0,65 \rightarrow$ viac e-mailov vedie k väčšiemu odlivu odberateľov

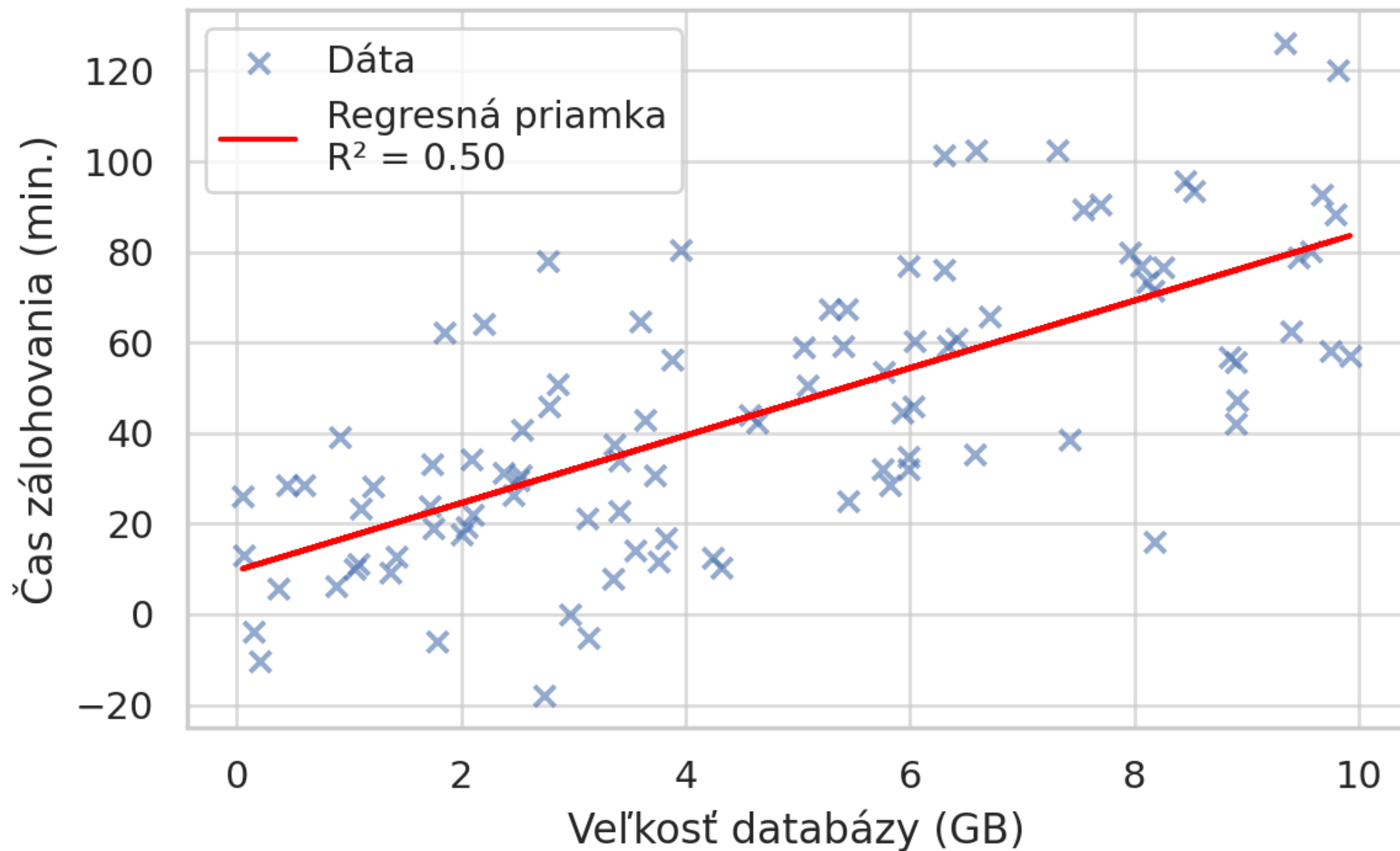
✓ Zľava (%) a objem predaja

- Analýza, či vyššia zľava vedie k vyššiemu predaju v kusoch
- Ak $R^2 \approx 0,9 \rightarrow$ silná lineárna závislosť

Reklamné výdavky vs. počet zákazníkov ($R^2 \approx 0.7$)



Veľkosť databázy vs. čas zálohovania ($R^2 \approx 0.85$)



Regresný Koeficient Determin. R^2

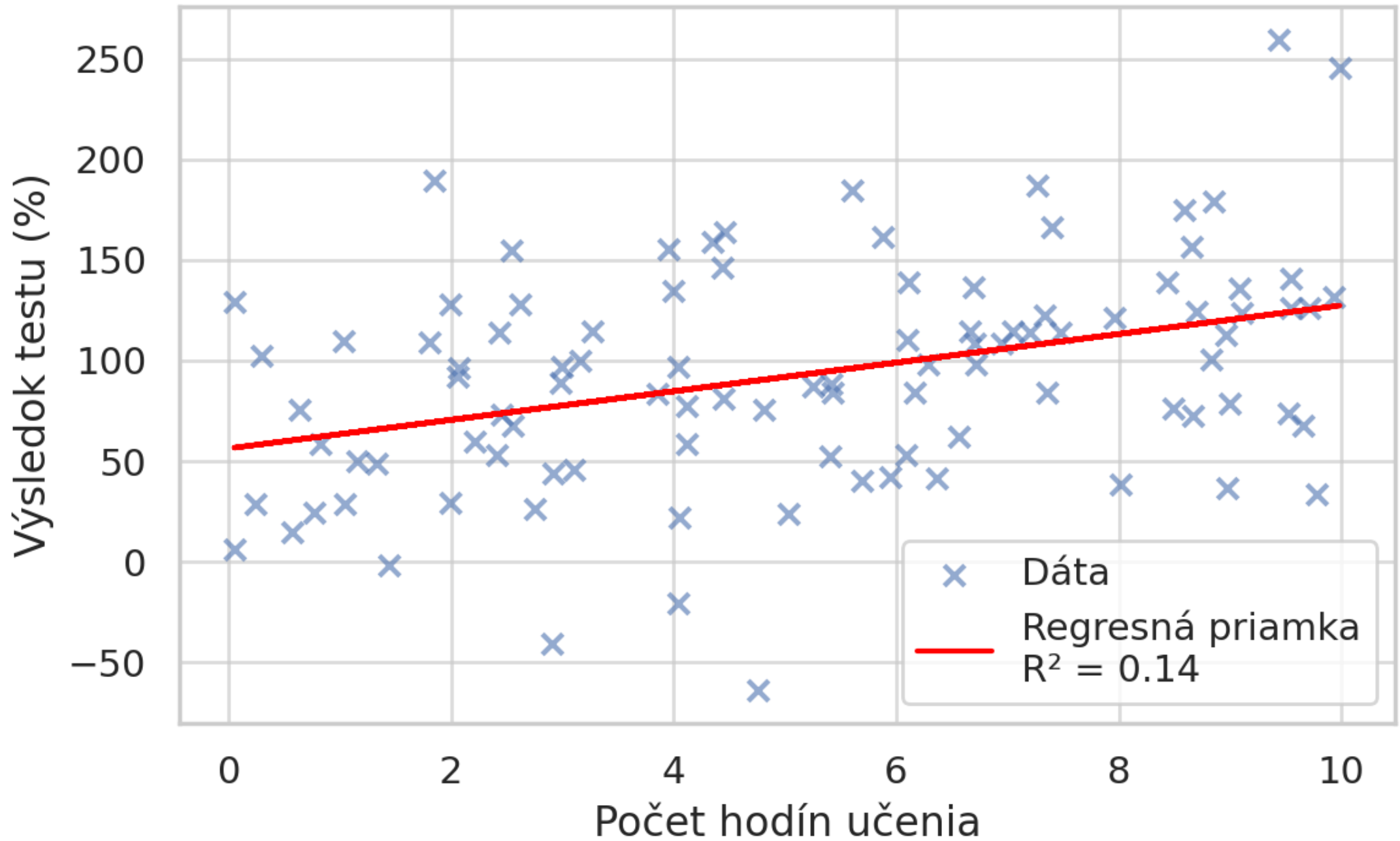
Vzdelávanie

- ✓ **Počet hodín učenia a výsledná známka**
 - Analýza, či viac učenia vedie k lepšiemu hodnoteniu
 - Ak $R^2 \approx 0,7 \rightarrow$ študenti, ktorí sa viac učia, dosahujú lepšie výsledky
- ✓ **Vek študentov a výkonnosť v online kurzoch**
 - Skúmanie, či starší študenti dosahujú iné výsledky ako mladší
 - Ak $R^2 \approx 0,3 \rightarrow$ slabá závislosť

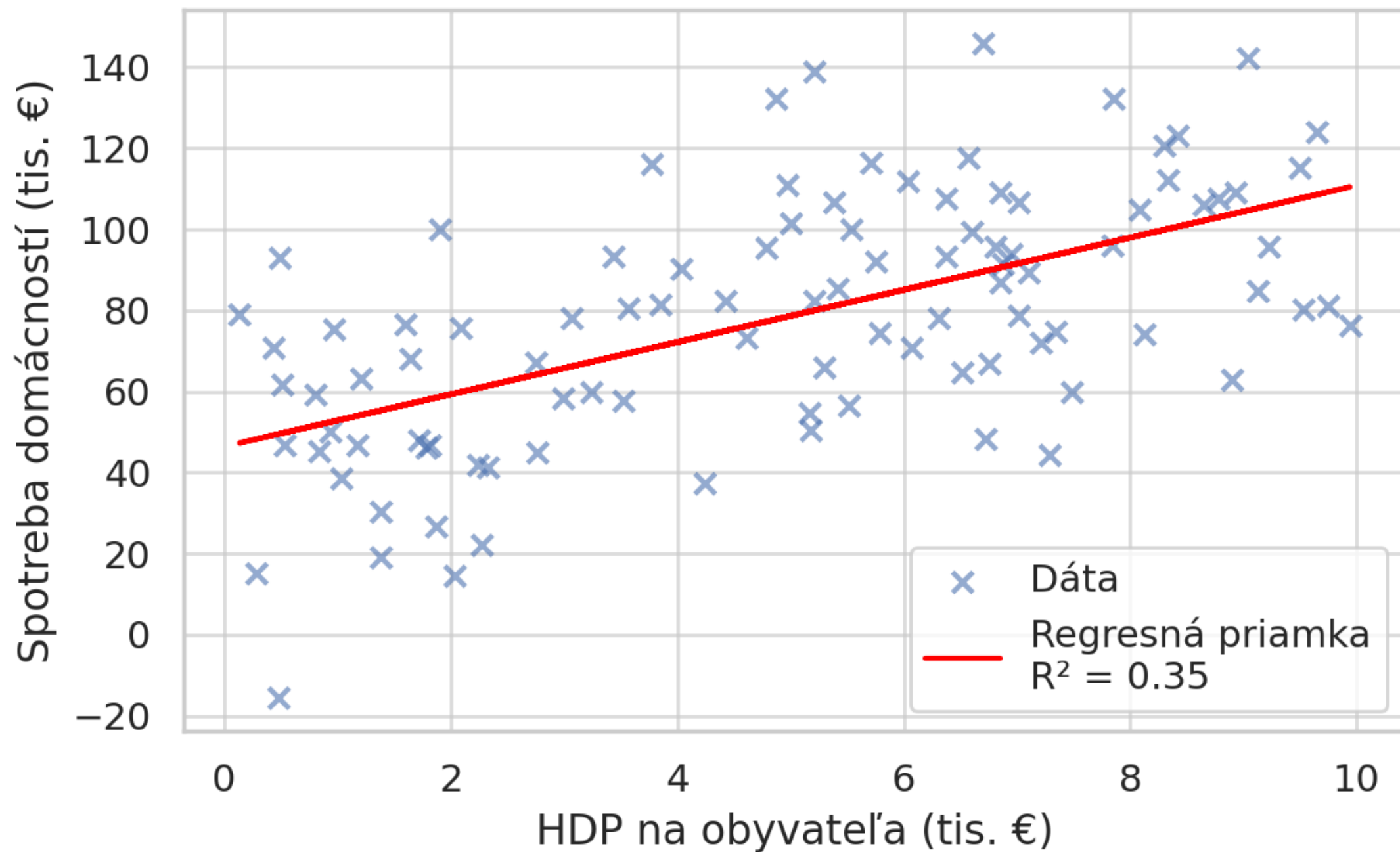
Ekonomika a financie

- ✓ **HDP a spotreba domácností**
 - Analýza, ako rast HDP ovplyvňuje spotrebu domácností
 - Ak $R^2 \approx 0,75 \rightarrow$ vyšší HDP vedie k vyššej spotrebe
- ✓ **Úroková sadzba a dopyt po hypotékach**
 - Skúmanie, či vyššia úroková sadzba znižuje počet žiadostí o hypotéky
 - Ak $R^2 \approx 0,6 \rightarrow$ mierna negatívna lineárna závislosť

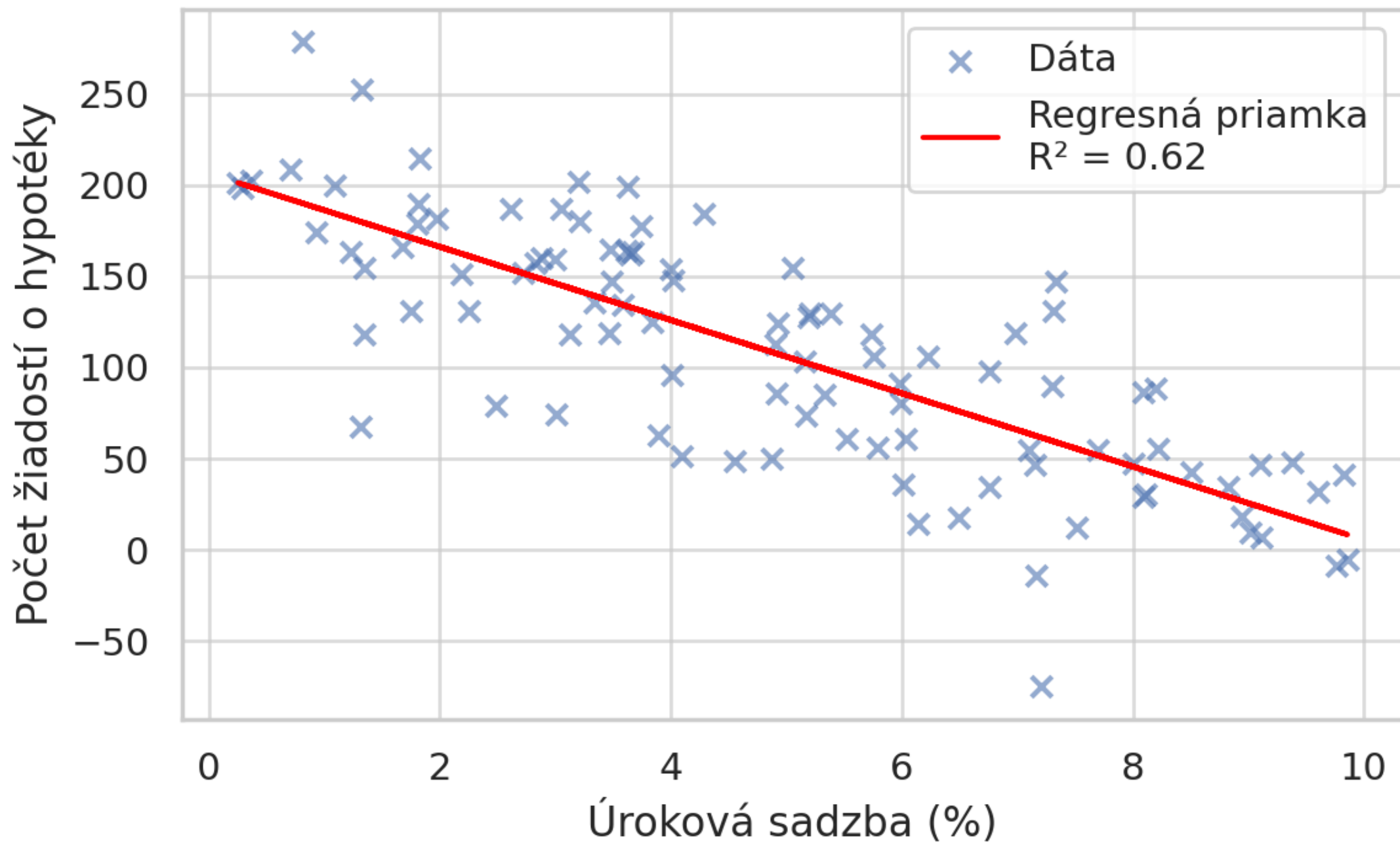
Počet hodín učenia vs. výsledok testu ($R^2 \approx 0.5$)



HDP vs. spotreba domácností ($R^2 \approx 0.75$)



Úroková sadzba vs. dopyt po hypotékach ($R^2 \approx 0.6$)



Regresný Koeficient Determin. R^2

• Manažment



Počet porád a efektivita rozhodovania

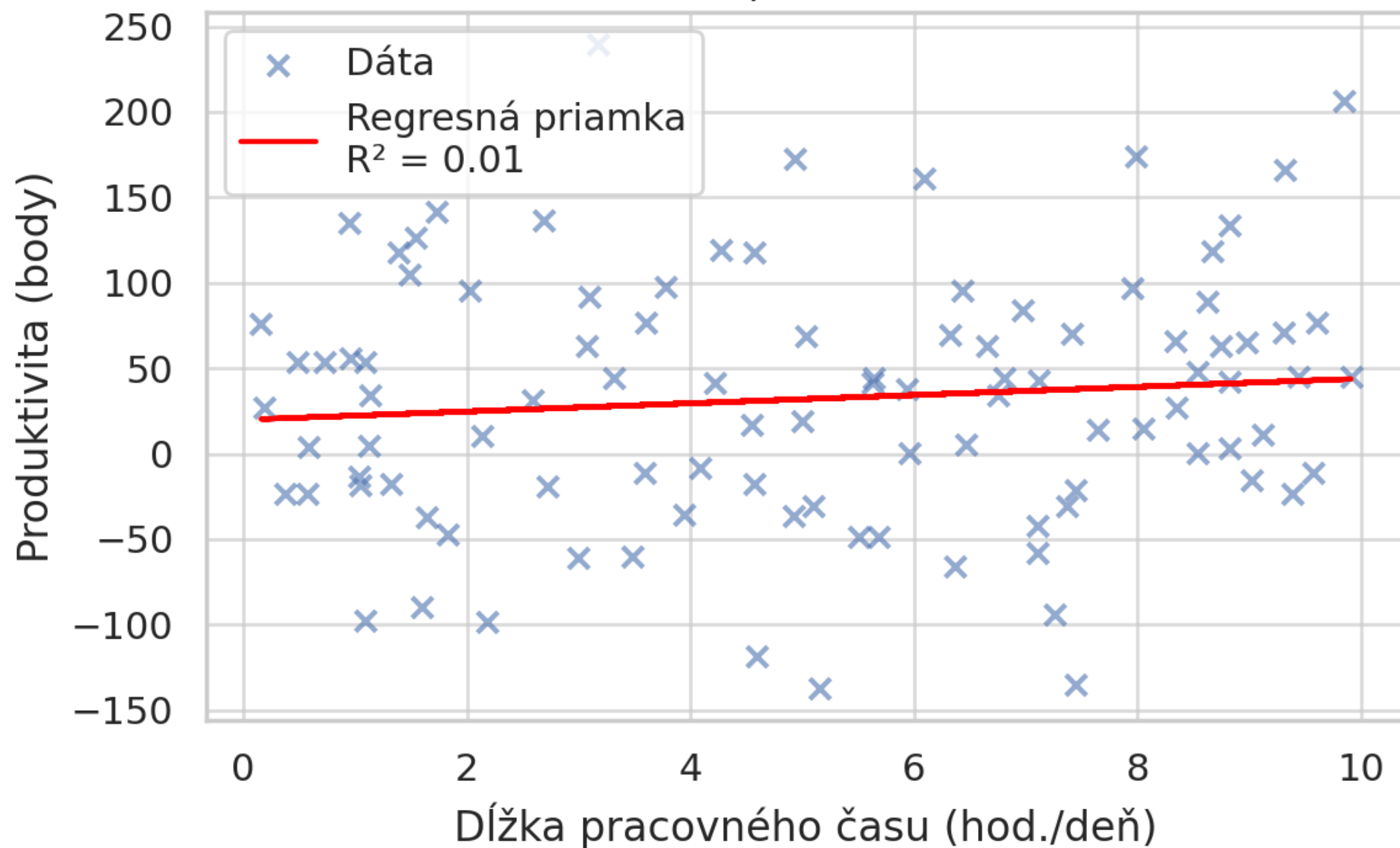
- Skúmanie, či viac porád vedie k lepším rozhodnutiam
- Ak $R^2 \approx 0,3 \rightarrow$ slabá až mierna závislosť (pri priveľa poradách nastáva zdržanie)



Fluktuácia zamestnancov a spokojnosť s vedením

- Analýza vplyvu spokojnosti s vedením na mieru odchodov zamestnancov
- Ak $R^2 \approx 0,6 \rightarrow$ silnejšia negatívna regresná závislosť (nízka spokojnosť vedie k vyššej fluktuácii)

Produktivita vs. dĺžka pracovného času ($R^2 \approx 0.2$)



Základné Typy Regresie

1. Lineárna regresia

- Modeluje vzťah medzi závislou a nezávislou premennou ako priamku
- $Y = \beta_0 + \beta_1 X + \epsilon$
- Predpovedanie ceny domu (Y) na základe jeho veľkosti (X)
- Používa sa, keď existuje očakávaný lineárny vzťah medzi 2 premennými

2. Viacnásobná lineárna regresia

- Modeluje vzťah medzi závislou premennou a viacerými nezávislými premennými
- $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon$
- Predpovedanie ceny domu (Y) na základe viacerých faktorov, ako sú jeho veľkosť (X_1), lokalita (X_2), počet izieb (X_3), a vek domu (X_4)
- Používa sa, keď je viacero faktorov (nezávislých premenných), ktoré ovplyvňujú predikciu závislej premennej

3. Logistická regresia

- Používa sa, keď je závislá premenná kategóriová (napríklad "áno" alebo "nie")
- Predpovedanie pravdepodobnosti, že zákazník vykoná nákup ($Y = 1$) alebo nie ($Y = 0$) na základe jeho demografických údajov (X)

Pokroč. Typy Regresie

4. Polynomiálna regresia

- Používa sa na modelovanie nelineárnych vzťahov medzi nezávislými a závislými premennými, keď lineárny model nedáva uspokojivé výsledky.
- $Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \dots + \beta_n X^n + \epsilon$
- Predpovedanie cenovej závislosti akcie na čase, kde vzťah medzi časom a cenou nie je lineárny.
- Tento typ regresie sa používa, keď vzťah medzi premennými nie je lineárny, ale môže byť prispôsobený nelineárnym polynomiálnym vzorcom

5. Ridge regresia

- Je variantom viacnásobnej lineárnej regresie, ktorý pridáva regulárizačný člen, aby sa zabránilo overfittingu (prílišné prispôsobenie modelu dátam).
- Pri modelovaní s viacerými nezávislými premennými, kde je riziko, že model bude príliš presný (overfitting).
- Používa sa pri veľkých datasetoch, kde sa obávame multikolinearít alebo overfittingu

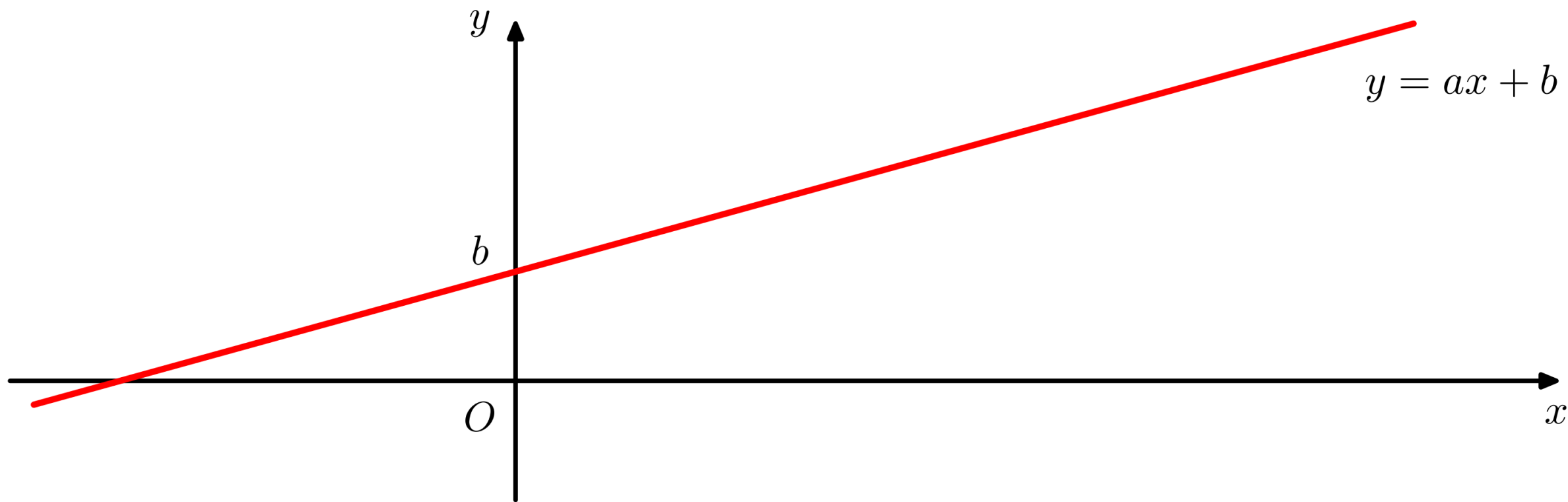
6. Lasso regresia

- Podobná ridge regresii, ale namiesto štvorcov koeficientov používa L1 reguláciu, čo spôsobuje, že niektoré koeficienty regresie môžu byť presne nula. To umožňuje selekciu vlastností.
- Lasso je vhodné, keď máme mnoho premenných a chceme sa zamerať len na najvýznamnejšie.

Lineárna funkcia – $f : y = ax + b, a, b \in \mathbb{R}, a \neq 0$

$\mathcal{D}(f) = \mathbb{R}, \mathcal{H}(f) = \mathbb{R}.$

Grafom je priamka so smernicou a , ktorá na osi y vytína úsek b .



Matematický Model Lineárnej Regresie

1. **Y**: Závislá premenná (predikovaná)

2. **X**: Nezávislá premenná (prediktor)

3. **β_0** : Intercept (priesečník s osou Y)

4. **β_1** : Koeficient regresie (slope)

5. **ϵ** : Chyba modelu

$$Y = \beta_0 + \beta_1 X + \epsilon$$

$$y = a_0 + a_1 x,$$

Regresné Koeficienty

Intercept (β_0)

- **Hodnota**, ktorú **závislá premenná (Y) nadobúda**, keď je **hodnota nezávislej premennej (X) rovná 0**.
 - Inými slovami, intercept je **bod, kde regresná priamka pretína os Y**.
 - Matematicky: $Y = \beta_0$, keď $X = 0$.
- **Interpretácia:**
Predstavuje počiatočnú hodnotu závislej premennej, keď nezávislá premenná nemá žiadny vplyv (napríklad cena produktu pri nulovej hodnote predaja).

Slope (β_1)

- **Koeficient**, ktorý určuje, **ako veľmi sa zmení hodnota závislej premennej (Y) pri zmene nezávislej premennej (X) o jednotku**.
 - Matematicky: Slope (β_1) udáva **zmenu hodnoty Y** pri zmene X o 1.
 - Predstavuje **sklon priamky**: ak je slope pozitívny, priamka stúpa, ak je negatívny, priamka klesá.
- **Interpretácia:**
Ak slope (β_1) je 2, znamená to, že pri zvýšení hodnoty X o 1 sa hodnota Y zvýši o 2. Ak je β_1 -3, znamená to, že pri zvýšení hodnoty X o 1 sa hodnota Y zníži o 3.

Regresná Rovnica

- Regresná rovnica môže vyzerat' napr. takto: **$Y = 5 + 2X$**
 - A. Intercept ($\beta_0 = 5$):** Ak $X = 0$, potom $Y = 5$ (počiatočná hodnota Y).
 - B. Slope ($\beta_1 = 2$):** Každé zvýšenie X o 1 spôsobí, že Y sa zvýši o 2.
- Tento model teda vyjadruje lineárny vzťah medzi X a Y , kde intercept je 5 a slope je 2.

Ako Postupovať?

1. Identifikovať premenné

- Určiť závislú a nezávislé premenné.

2. Vytvoriť model

- Určiť, aký typ regresie je najvhodnejší.











3. Vyhodnotiť model

- Použiť metódy ako R^2 , F-test alebo p-hodnoty na testovanie účinnosti modelu.

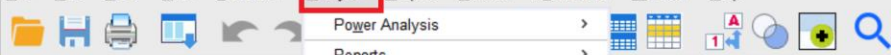
4. Predikcia

- Použiť model na predpovedanie neznámych hodnôt závislej premennej.

Microsoft Excel Regresia

Typ Regresie	Funkcia	Popis
Lineárna regresia (jednoduchá) 	LINEST	Vypočíta regresné koeficienty (slope, intercept) pre lineárnu regresiu.
Predikcia pomocou lineárnej regresie 	TREND	Predpovedá hodnoty Y na základe existujúcich hodnôt X.
Korelačný koeficient medzi X a Y 	CORREL	Vypočíta koreláciu medzi dvoma premennými, ukazuje silu vzťahu.
Koeficient determinácie (R²) 	RSQ	Ukazuje, ako dobre model vysvetľuje variabilitu v závislej premennej.
Sklon regresnej priamky (slope) 	SLOPE	Vypočíta sklon regresnej priamky (koeficient β_1).
Intercept (priesečník s osou Y) 	INTERCEPT	Vypočíta hodnotu interceptu regresnej priamky (priesečník s osou Y).
Štandardná chyba odhadu 	STEYX	Vypočíta štandardnú chybu predikcie regresného modelu.
Predikcia nových hodnôt (lineárna) 	FORECAST.LINEAR	Predpovedá hodnoty Y pre nové hodnoty X na základe lineárnej regresie.
Pokročilá analýza údajov 	Data Analysis Toolpak → Regression	Komplexná analýza regresie s rôznymi štatistikami.
Staršia verzia predikcie 	FORECAST	Predpovedá hodnoty Y pre nové hodnoty X (staršia verzia FORECAST.LINEAR).

	A	B	C	D	E	F	G	H
1	X (Predikátor)	Y (Závislá premenná) OK	Y (Závislá premenná) NG					
2	1	3	3					
3	2	4,1	45661					
4	3	5,2	45693					
5	4	6,3	45722					
6	5	7,4	45754					
7	6	8,5	45785					
8	7	9,6	45817					
9	8	10,7	45848					
10	9	11,8	45880					
11	10	12,9	45912					
12	11	14	14.0					
13	12	15,1	45672					
14	13	16,2	45704					
15	14	17,3	45733					
16	15	18,4	45765					
17	16	19,5	45796					
18	17	20,6	45828					
19	18	21,7	45859					
20	19	22,8	45891					
21	20	23,9	45923					
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								
35								
36								
37								
38								
39								
40								



										Visible: 14 of 14 Variables	
	Age	Education	ra	Average_time_in_course	Gender	Student_type	Educational_category	Graduated	Number_of_badges_earned	V11	Vzdelávacie kategórie
1	39		1%	35,492866648031814	Male	Individual Participant	Financial Management and Accounting	Yes	14		IT Systems and Server Administration
2	33		8%	28,953806774758736	Male	Private Company	Management Skills and Personal Development	No	0		Financial Management and Accounting
3	41		9%	18,630416999475095	Male	Individual Participant	Office Packages (Microsoft Office)	Yes	10		Graphics, Design and Multimedia
4	50		5%	24,625181734233472	Male	Individual Participant	Database Management, Statistics and Data Analysis	Yes	14		Office Packages (Microsoft Office)
5	32		5%	34,181547898393420	Male	Individual Participant	Graphics, Design and Multimedia	Yes	8		Cybersecurity and Data Protection
6	32		8%	29,620680172660960	Male	Individual Participant	Office Packages (Microsoft Office)	Yes	8		Management Skills and Personal Development
7	50					Individual Participant	SAP, ABAP and Enterprise Systems	Yes	10		Marketing, Advertising and Law
8	42					Individual Participant	Management Skills and Personal Development	Yes	8		Operating Systems
9	30					Individual Participant	Cybersecurity and Data Protection	Yes	6		Computer Networks and Cisco
10	40					Individual Participant	Artificial Intelligence and Machine Learning (AI and ML)	Yes	14		Certification Preparation
11	30					Individual Participant	Graphics, Design and Multimedia	Yes	4		Programming and Software Development
12	30					Private Company	SAP, ABAP and Enterprise Systems	No	0		SAP, ABAP and Enterprise Systems
13	37					Individual Participant	Programming and Software Development	Yes	4		Database Management, Statistics and Data Analysis
14	18					Individual Participant	Programming and Software Development	Yes	12		Artificial Intelligence and Machine Learning (AI and ML)
15	18					Individual Participant	Marketing, Advertising and Law	Yes	14		Website Development and Design
16	29					Individual Participant	Management Skills and Personal Development	Yes	14		
17	24					Individual Participant	Marketing, Advertising and Law	Yes	6		
18	38					Private Company	Financial Management and Accounting	Yes	8		
19	25					Individual Participant	Graphics, Design and Multimedia	Yes	12		
20	20					Individual Participant	Management Skills and Personal Development	Yes	16		
21	49					Individual Participant	Artificial Intelligence and Machine Learning (AI and ML)	Yes	6		
22	32					Individual Participant	Certification Preparation	Yes	8		
23	35		5%	28,533874262302916	Female	State or Public Institution	Financial Management and Accounting	Yes	6		
24	20		7%	27,917500108017308	Male	Individual Participant	Management Skills and Personal Development	Yes	16		
25	29		2%	16,929726943942576	Male	Private Company	SAP, ABAP and Enterprise Systems	Yes	2		
26	36		15,37%	34,633146375088170	Male	Individual Participant	SAP, ABAP and Enterprise Systems	Yes	6		
27	23		2	68,37%	30,792714186936550	Male	Individual Participant	Office Packages (Microsoft Office)	Yes	14	
28	38		4	77,65%	26,000836732644288	Male	Individual Participant	Artificial Intelligence and Machine Learning (AI and ML)	Yes	10	
29	28		3	65,11%	31,773546616597194	Male	Individual Participant	Office Packages (Microsoft Office)	Yes	18	
30	32		1	46,43%	25,731107622547796	Male	Individual Participant	Office Packages (Microsoft Office)	Yes	10	
31	28		2	65,74%	36,137046628083100	Male	Individual Participant	SAP, ABAP and Enterprise Systems	Yes	8	
32	53		3	79,85%	50,000000000000000	Male	Individual Participant	Programming and Software Development	Yes	10	
33	34		1	51,34%	26,986080017771805	Male	Individual Participant	Programming and Software Development	Yes	14	
34	24		1	54,87%	21,578967188352344	Male	Individual Participant	Artificial Intelligence and Machine Learning (AI and ML)	No	0	
35	43		4	68,52%	16,863605612096496	Female	Individual Participant	SAP, ABAP and Enterprise Systems	No	0	
36	22		3	80,52%	35,980333600968200	Female	Private Company	Database Management, Statistics and Data Analysis	Yes	14	
37	37		1	50,35%	12,735022491068355	Female	Private Company	Programming and Software Development	Yes	8	
38	18		1	53,96%	17,139341461878427	Female	Private Company	Artificial Intelligence and Machine Learning (AI and ML)	Yes	8	
39	21		1	51,32%	25,819627882849634	Male	Individual Participant	Marketing, Advertising and Law	Yes	8	
40	36		3	84,53%	50,000000000000000	Male	Individual Participant	Artificial Intelligence and Machine Learning (AI and ML)	Yes	6	
41	42		4	78,32%	30,082574620323370	Male	Individual Participant	Management Skills and Personal Development	Yes	8	
42	36		2	61,79%	30,054527388552152	Male	Private Company	SAP, ABAP and Enterprise Systems	Yes	4	



Output
Log
Regression
Title
Notes
Variables Entered
Model Summary
ANOVA
Coefficients
Coefficient Correlations

Regression

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	Number_of_badges_earned, Age, Courses_taken, Education ^b	.	Enter

a. Dependent Variable: Completion_rate

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,879 ^a	,772	,772	4,96301%

a. Predictors: (Constant), Number_of_badges_earned, Age, Courses_taken, Education

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	449605,954	4	112401,488	4563,320	,000 ^b
	Residual	132936,280	5397	24,632		
	Total	582542,233	5401			

a. Dependent Variable: Completion_rate

b. Predictors: (Constant), Number_of_badges_earned, Age, Courses_taken, Education

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95,0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	40,035	,327		122,442	,000	39,394	40,676
	Age	,000	,007	,000	,065	,948	-,014	,015
	Education	5,048	,072	,512	70,582	,000	4,907	5,188
	Courses_taken	1,983	,025	,556	80,147	,000	1,934	2,031
	Number_of_badges_earned	-,008	,014	-,004	-,587	,557	-,035	,019

a. Dependent Variable: Completion_rate

Coefficient Correlations^a






Model			Number_of_badges_earned	Age	Courses_taken	Education
1	Correlations	Number_of_badges_earned	1,000	-,001	-,001	,020
		Age	-,001	1,000	-,001	-,289

Regresia a Časové Rady



Microsoft Excel

Regresia a Časové Rady

Funkcia	Popis	Príklad použitia
LOGEST 	Exponenciálna regresia	<code>=LOGEST(B2:B21, A2:A21, TRUE, TRUE)</code>
FORECAST.ETS 	Predikcia časového radu s ETS	<code>=FORECAST.ETS(A22, B2:B21, A2:A21, 1, TRUE, 1)</code>
FORECAST.ETS.CONFINT 	Interval spoľahlivosti predikcie	<code>=FORECAST.ETS.CONFINT(A22, B2:B21, A2:A21, 1, TRUE, 1)</code>
FORECAST.ETS.STAT 	Štatistika ETS modelu	<code>=FORECAST.ETS.STAT(A22, B2:B21, A2:A21, 1, 1, TRUE, 1)</code>
FORECAST.ETS.SEASONALITY 	Sezónnosť modelu	<code>=FORECAST.ETS.SEASONALITY(B2:B21, A2:A21, 1, TRUE, 1)</code>

E2



fx

=FORECAST.ETS.STAT(B2:B21; A2:A21; 1;;; 1)

	A	B	C	D	E	F	G	H	I	J	K
1	Dátum	Skutočné Hodnoty									
2	1.1.2025	347		0,1							
3	1.2.2025	402									
4	1.3.2025	285									
5	1.4.2025	371									
6	1.5.2025	431									
7	1.6.2025	389									
8	1.7.2025	453									
9	1.8.2025	475									
10	1.9.2025	392									
11	1.10.2025	460									
12	1.11.2025	425									
13	1.12.2025	385									
14	1.1.2024	410									
15	1.2.2024	440									
16	1.3.2024	460									
17	1.4.2024	435									
18	1.5.2024	490									
19	1.6.2024	505									
20	1.7.2024	475									

Argumenty funkcie

FORECAST.ETS.STAT

Values

B2:B21



= {347;402;285;371;431;389;453;475;392;460;425;385;410;440;460;435;490;505;475}

Timeline

A2:A21



= {45658;45689;45717;45748;45778;45809;45840;45871;45902;45933;45964;45995;46026;46057;46088;46119;46150;46181;46212;46243}

Statistic_type

1



= 1

Seasonality



= číslo

Data_completion



= číslo

= 0,1

Vráti požadovanú štatistiku pre prognózu.

Statistic_type

je číslo v rozsahu 1 až 8 označujúce typ štatistiky, ktorú Microsoft Excel vráti pre vypočítanú prognózu.

Výsledok = 0,1

[Pomocník pre túto funkciu](#)

OK

Zrušiť



FORECAST.ETS.STAT (funkcia)

► *Applies To*

Tu je vaša bezplatná skúšobná verzia služby Microsoft 365



Odomknúť >

Dôležité: Funkcia FORECAST. ETS. Stat nie je k dispozícii v Exceli pre web, iOS alebo Android.

Vráti štatistickú hodnotu ako výsledok predpovedaného časového radu.

Typ štatistiky označuje štatistiku, ktorá sa pri tejto funkcii vyžaduje.

Syntax

FORECAST.ETS.STAT(hodnoty; časová_os; typ_štatistiky; [sezónnosť]; [doplňovanie_údajov]; [agregácia])

Syntax funkcie FORECAST.ETS.STAT obsahuje tieto argumenty:

- **Hodnoty** Povinný argument. Hodnoty sú historické hodnoty, pre ktoré chcete predpovedať ďalšie body.
- **Časová os** Povinný argument. Nezávislé pole alebo rozsah číselných údajov. Medzi dátumami na časovej osi musí byť konzistentný krok, nula sa nesmie použiť. Časovú os nie je potrebné zoradiť, keďže funkcia FORECAST.ETS.STAT ju pre výpočty zoradí implicitne. Ak v zadanej časovej osi nie je možné identifikovať konštantný krok, funkcia FORECAST.ETS.STAT vráti chybu #ČÍSLO! . Ak časová os obsahuje duplicitné hodnoty, funkcia FORECAST.ETS.STAT vráti chybu #HODNOTA! . Ak rozsahy na časovej osi nemajú rovnakú veľkosť ako hodnoty, funkcia FORECAST.ETS.STAT vráti chybu #NEDOSTUPNÝ.
- **Typ_štatistiky** Povinný argument. Je to číselná hodnota medzi 1 a 8 označujúca, ktorá štatistika sa

Môžu sa vrátiť nasledujúce voliteľné štatistiky:

- **Alfa parameter algoritmu ETS** Vráti parameter základnej hodnoty – vyššia hodnota dáva aktuálnym údajovým bodom väčšiu váhu.
- **Beta parameter algoritmu ETS** Vráti parameter s hodnotou trendu – vyššia hodnota dáva aktuálnemu trendu väčšiu váhu.
- **Gama parameter algoritmu ETS** Vráti parameter s hodnotou sezónnosti – vyššia hodnota dáva aktuálnemu sezónnemu obdobiu väčšiu váhu.
- **Metrika MASE** Vráti chybu metriky MASE– ide o mieru presnosti prognóz.
- **Metrika SMAPE** Vráti chybu metriky SMAPE – ide o mieru presnosti fungujúcu na základe chýb percentuálnych hodnôt.
- **Metrika MAE** Vráti chybu metriky MAE – ide o mieru presnosti fungujúcu na základe chýb percentuálnych hodnôt.
- **Metrika RMSE** Vráti chybu metriky RMSE – ide o mieru rozdielnosti medzi predpovedanými a pozorovanými hodnotami.
- **Štatistika Step size detected** Vráti štatistiku step size detected na historickej časovej osi.

Stiahnutie vzorového zošita

[Kliknutím na toto prepojenie si môžete stiahnuť zošit s príkladmi excelovej funkcie FORECAST.ETS](#)

Typy Štatistík 1

1. RMSE (Root Mean Square Error)

- Štandardná štatistika, ktorá sa používa na meranie rozdielu medzi predikovanými a skutočnými hodnotami.
- Ukazuje, aký je priemerný štvorcový rozdiel medzi skutočnými a predikovanými hodnotami.
- Nižšia hodnota RMSE znamená lepšiu predikciu.

2. MAD (Mean Absolute Deviation)

- Meria priemerný absolútny rozdiel medzi skutočnými a predikovanými hodnotami.
- Používa sa na vyhodnotenie variability predikcií.
- Čím nižší je MAD, tým presnejšie sú predikcie.

3. MSE (Mean Squared Error)

- Podobný RMSE, ale meria priemerný štvorcový rozdiel medzi skutočnými a predikovanými hodnotami bez aplikácie druhej odmocniny.
- MSE je veľmi citlivý na veľké chyby.
- Vyššia hodnota MSE znamená väčšie rozdiely medzi skutočnými a predikovanými hodnotami.

Typy Štatistík 2

4. SMAPE (Symmetric Mean Absolute Percentage Error)

- Metóda, ktorá hodnotí percentuálnu odchýlku medzi skutočnými a predikovanými hodnotami. Je symetrická, takže penalizuje nadpredikciu aj podpredikciu rovnako.
- Je vhodný na porovnávanie predikcií medzi rôznymi datasetmi. Je to percentuálne vyjadrenie rozdielu, ktoré je upravené na rovnaký rozsah.

5. APE (Absolute Percentage Error)

- Absolútny percentuálny rozdiel medzi skutočnými a predikovanými hodnotami.
- APE je užitočné na vyhodnotenie relatívnej chyby.
- Vyššia hodnota znamená väčší rozdiel v percentách medzi skutočnými a predikovanými hodnotami.

6. Theil's U-statistic

- Štatistika na meranie presnosti predikcií.
- Porovnáva rozdiel medzi skutočnými hodnotami a predikciami z dvoch rôznych
- Tento ukazovateľ je vhodný na porovnávanie rôznych predikčných modelov, pričom hodnoty blízke 0 naznačujú, že model je veľmi presný.