KALIBRAČNÍ KŘIVKA:POSTUP ZPRACOVÁNÍ DAT POMOCÍ LINEÁRNÍ REGRESE

300524 /jp

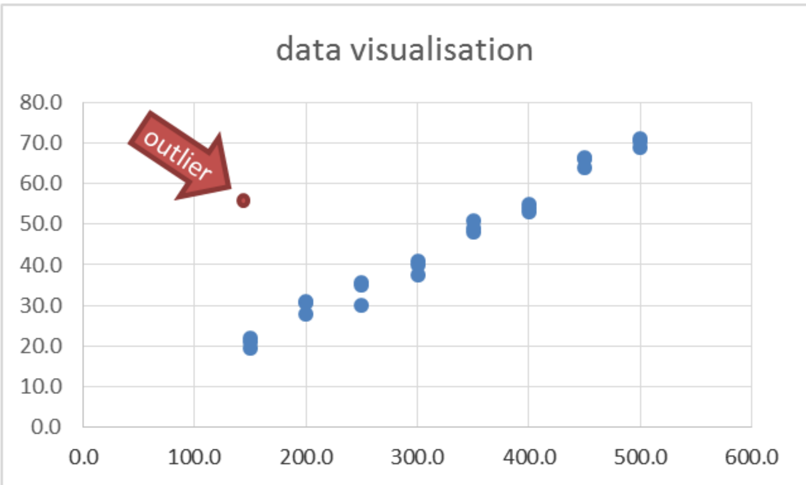
1/ ZOBRAZENÍ DAT

Vstupem je tabulka X-Y dat:

c(mg/L)	signal (mV)		
150.0	22.0	21.0	19.5
200.0	30.5	31.0	28.0
250.0	35.0	35.5	30.0
300.0	40.0	41.0	37.5
350.0	51.0	49.0	48.0
400.0	55.0	53.0	54.0
450.0	64.0	66.5	66.0
500.0	70.5	71.0	69.0

Nezávislá proměnná (X) může být: koncentrace standardu (mg/l, ug/ml nebo mol/l), objem standardu (uL, ml), miligramy nebo milimoly standardu atd. Počet X (různých koncentračních úrovní) je typicky 5-7. Jelikož experimenty by měly být vždy opakovány (aspoň 3x), první otázkou je: kolik bodů se objeví v grafu kalibrační křivky? Odpověď je: v grafu by se měla objevit **všechna** naměřená data (N=24). Neprůměrujeme hodnoty naměřené při stejných koncentracích (řádky)! Takový přístup vede k zamaskování odlehlých bodů a také ke zmenšení stupňů volnosti! Pokud máme opakovaná měření v řádcích, je nejjednodušší přepsat data do dvou sloupců (kvůli Excelu) a vytvořit x-y graf. Grafické zobrazení dat pomůže odhalit zjevně odlehlé body (outliers, často hrubé chyby), které by měly být předem vyloučeny.

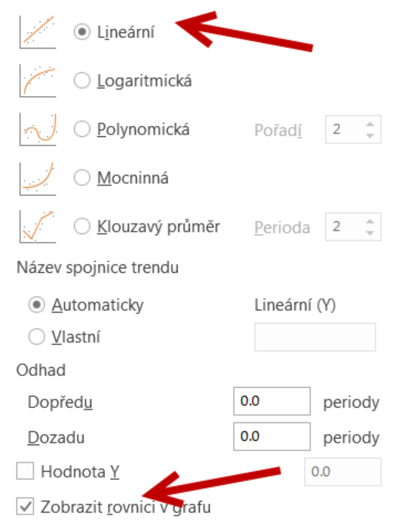
150.0	22.0
200.0	30.5
250.0	35.0
300.0	40.0
350.0	51.0
400.0	55.0
450.0	64.0
500.0	70.5
150.0	21.0
200.0	31.0
250.0	35.5
300.0	41.0
350.0	49.0
400.0	53.0
450.0	66.5
500.0	71.0
150.0	19.5
200.0	28.0
250.0	30.0
300.0	37.5
350.0	48.0
400.0	54.0
450.0	66.0
500.0	69.0

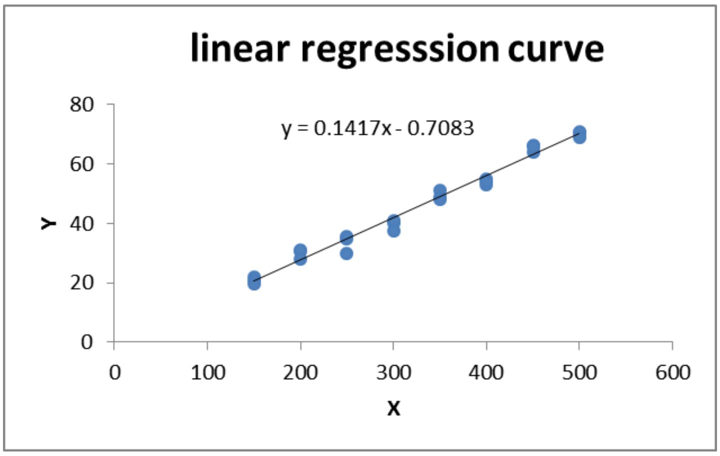


Pozn.: v tomto kroku také vyplyne silná korelace mezi Y a X, což není překvapivé: od počátku jsme totiž zamýšleli vytvořit kalibrační křivku, která těsnou závislost přepokládá! Tuto závislost si tudíž není nutno ověřovat např. Spearmanovým testem pořadové korelace.

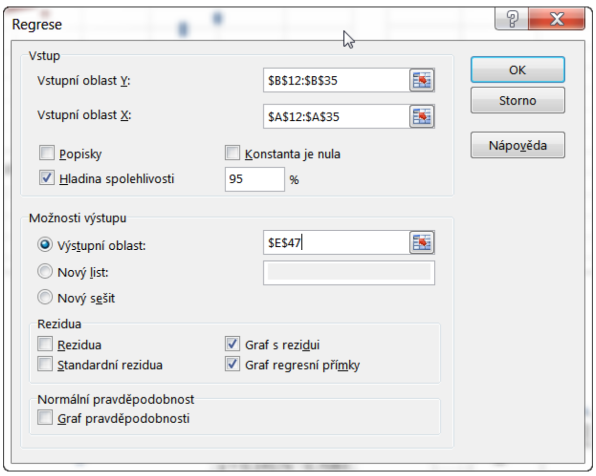
2/ PRVNÍ MODEL LINEÁRNÍ REGRESE Y = b * X + a

Naším konečným cílem je nalézt rovnici přímky, kterou nejlépe proložíme získané body a kterou nazýváme „kalibrační křivka“. Pro první odhad použijeme buď nabídku z pravého kliku na grafu (**Přidej Spojnici trendu + lineární + zobrazit rovnici v grafu** – viz šipky níže)





anebo použijeme dialog Data – Analýza dat – Regrese:



RESULT

Regression statistics	
R	0.991338
R^2	0.982751
Nastavená hodnot	0.981967
Chyba stř. hodnot	2.245787
Pozorování	24

ANOVA

	Rozdíl	*SS*	*MS*	*F*	*ýznamnost F*
Regrese	1	6321.875	6321.875	1253.455	6.82E-21
Rezidua	22	110.9583	5.043561		
Celkem	23	6432.833			

	coefficients	*std. error*	*t Stat*	*P-value*	*lower 95%*	*upper 95%*	*ower 95.(*
intercept	-0.70833	1.378892	-0.5137	**0.613**	-3.56798	2.151[illegible]3	-3.5679
slope	0.141667	0.004001	35.40416	6.82E-21	0.133368	0.149965	0.13336

$t = \frac{|a|}{s_a} = \frac{0.7083}{1.3789} = 0.5137$

t(crit.)=T.INV.2T(0.05;N-2)= 2.07387

t(crit.)=TINV(0.05;N-2)= 2.07387

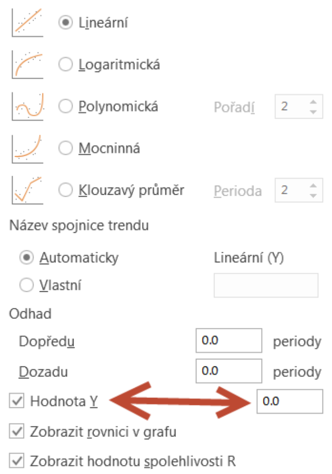
Druhý způsob je vhodnější, jelikož zároveň velmi pomůže **dalšímu** postupu, protože **TOHLE NENÍ KONEC PŘÍBĚHU!!!**

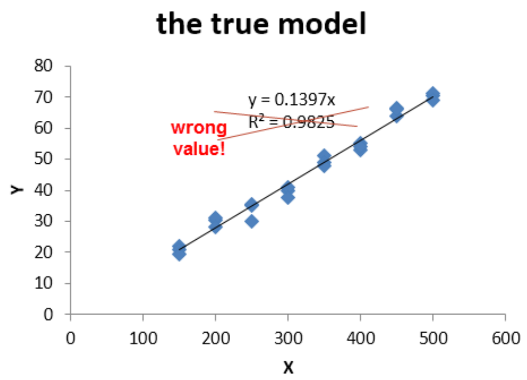
3/ Je úsek na ose Y VÝZNAMNÝ?

Poněvadž naše měření je zatíženo náhodnými chybami, může se stát, že jiná série měření za stejných podmínek, by poskytla rovnici např. Y=0.1417*x **+ 0.001** a další takové měření třeba Y=0.1417*x **– 0.111**. Toto znamená, že hodnota (i znaménko) úseku na ose Y nemají žádný význam, jsou to jen náhodná čísla. Je-li to pravda, učiníme důležitý závěr: ÚSEK BY MĚL BÝT ROVEN NULE. Tabulka, kterou poskytla **Analýza dat** nese informaci, zda je úsek <u>-0.7083</u> našeho regresního modelu <u>Y = 0.1417 * X **– 0.7083**</u> významně odlišný od 0.

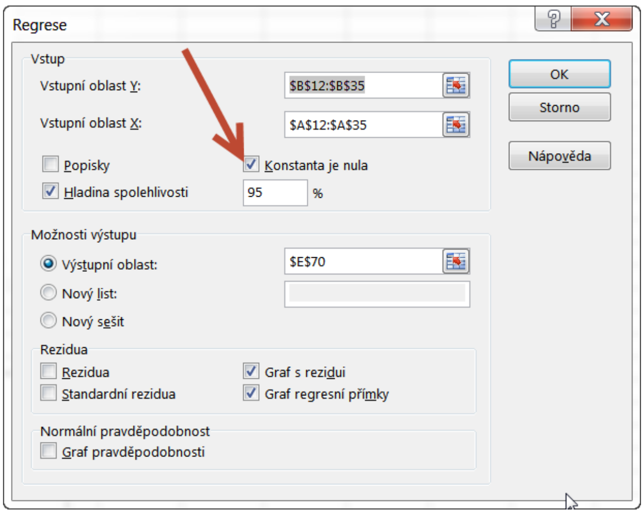
Abychom verifikovali nulovou hypotézu $\boxed{a=0}$ (na hladině významnosti α=0.05), srovnáme poměr **$|a|/s_a$** s **t(krit.)** (viz žlutá políčka výše: **0.532<2.0739**) nebo, jednodušeji, srovnáme *P*-hodnotu s 0,05 (**0.613 > 0.05**). V našem případě, obě hodnoty v tabulce **potvrzují** nulovou hypotézu. Jednoduché pravidlo říká: považujeme úsek za nulový, pokud *P*-hodnota je **větší** než 0.05.

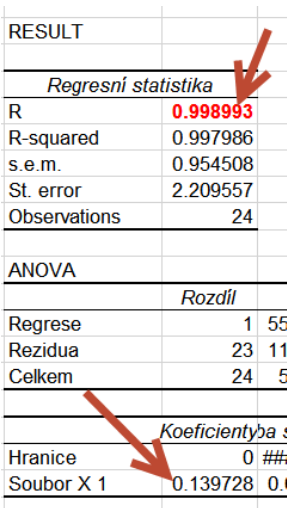
V takovém případě zapomeneme na náš původní regresní model resp. změníme jej na jednodušší tvar s jednodušší rovnicí Y = b*X, což znamená, že postup nalezení nové rovnice musíme zopakovat s podmínkou Y(0)=0 (směrnice se o něco změní, také se změní statistické hodnoty):





Všimněme si, že grafického vyhodnocení **Přidej spojnici trendu vrací nesprávnou hodnoru R^2**. Proto používejme zásadně Data – Analýza dat – Regrese, což poskytne správné hodnoty jak R=0.998993, tak směrnice=0.139728:





RESULT	
Regresní statistika	
R	**0.998993**
R-squared	0.997986
s.e.m.	0.954508
St. error	2.209557
Observations	24
ANOVA	
	Rozdíl
Regrese	1
Rezidua	23
Celkem	24
	Koeficienty
Hranice	0
Soubor X 1	0.139728

4/ VÝPOČET V NEZNÁMÉM VZORKU

Ať už úsek byl vynulován nebo ne, kalibrační křivka se používá k určení neznámé koncentrace (množství) ve vzorku. Pro to použijeme správný regresní model a vypočítáme X pro dané Y:

X = Y / b (byl-li úsek nevýznamný)

anebo, pokud úsek **a** byl významný: **X = (Y – a) / b**

Obvykle je i vzorek měřen opakovaně; po otestování na odlehlé hodnoty pak dosadíme za Y průměr.