MUNI
FI

Lineárna a polynomiálna regresia

Matúš Šikyňa (485591)

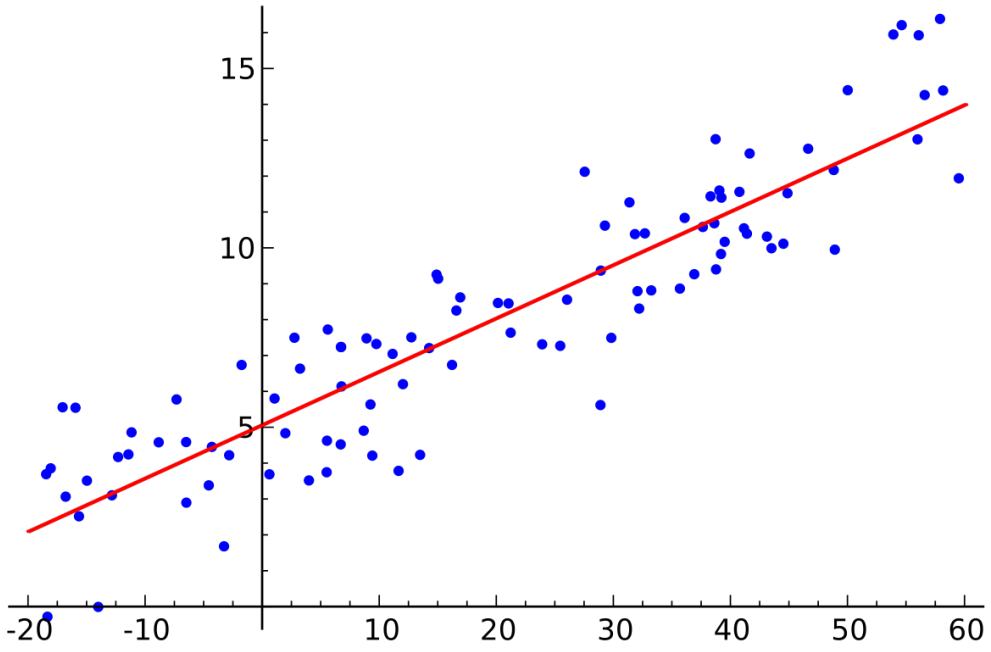
Regresná analýza

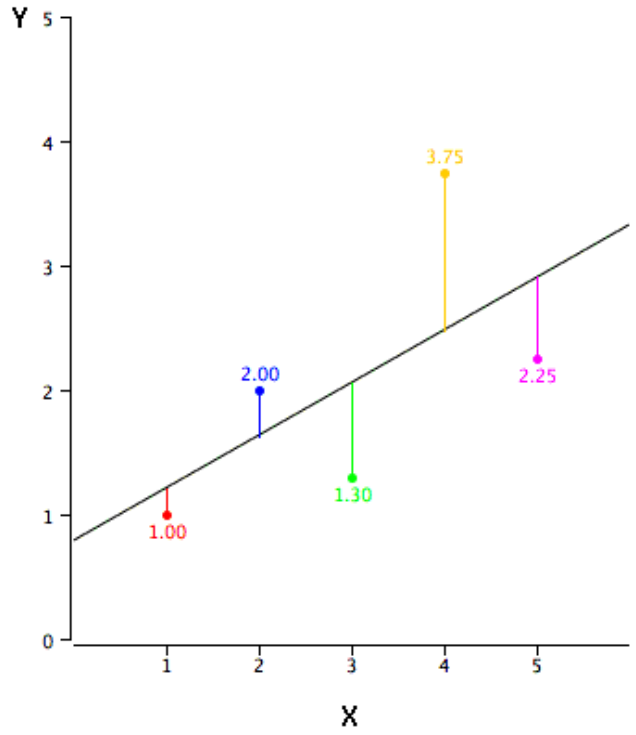
- Opis vzťahu medzi premennými
- Supervised learning

Lineárna regresia

$$Y = a + bX + \varepsilon$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - Y'_i)^2$$





MUNI FI

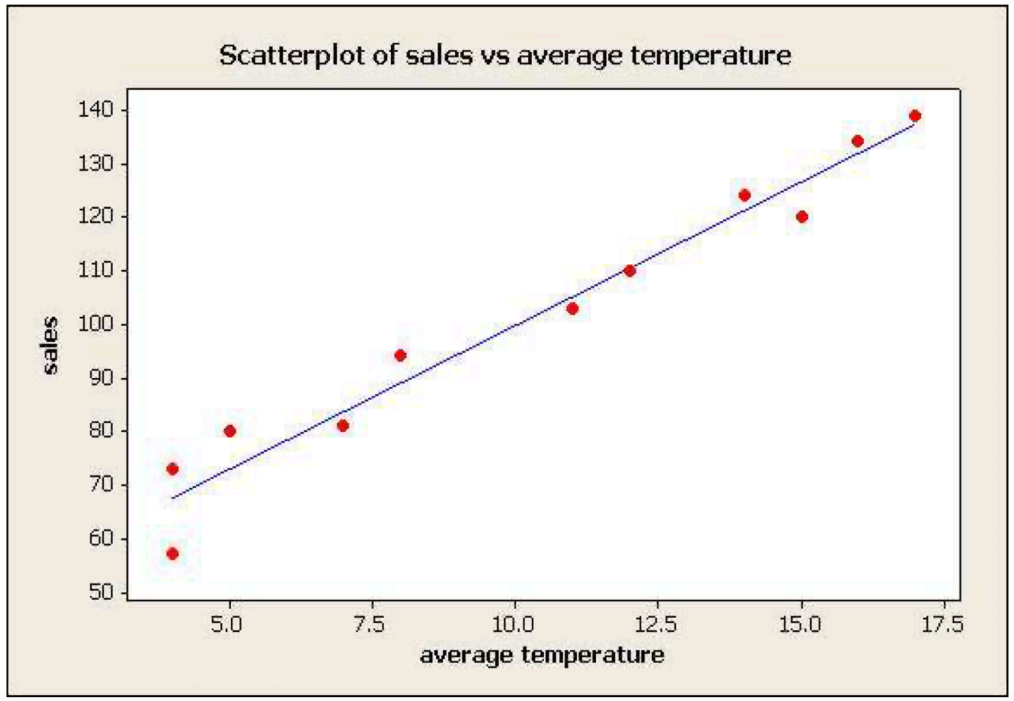
Lineárna regresia

- Predikcia spotreby pre výpočet HDP
- Export a import komodít
- Ponuka a dopyt na pracovnom trhu

Lineárna regresia

– Počet predaných zmrzlín vzhľadom na teplotu $y = 45.52 + 5.448x$

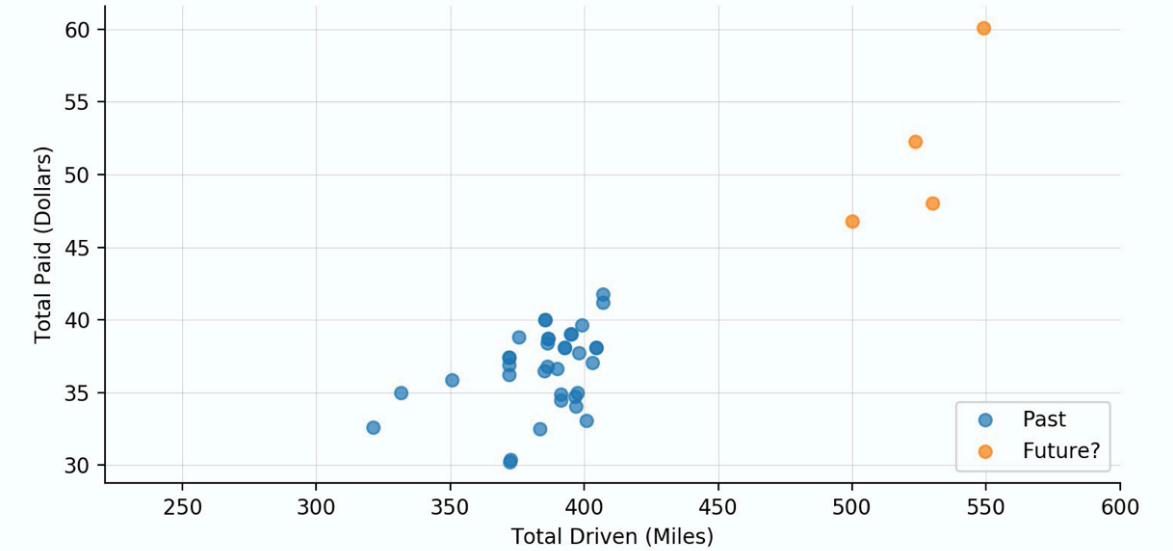
Month	Average Temp (°C)	Sales (£ 000's)
January	4	73
February	4	57
March	7	81
April	8	94
May	12	110
June	15	124
July	16	134
August	17	139
September	14	124
October	11	103
November	7	81
December	5	80



MUNI FI

Lineárna regresia

– Peniaze na benzín

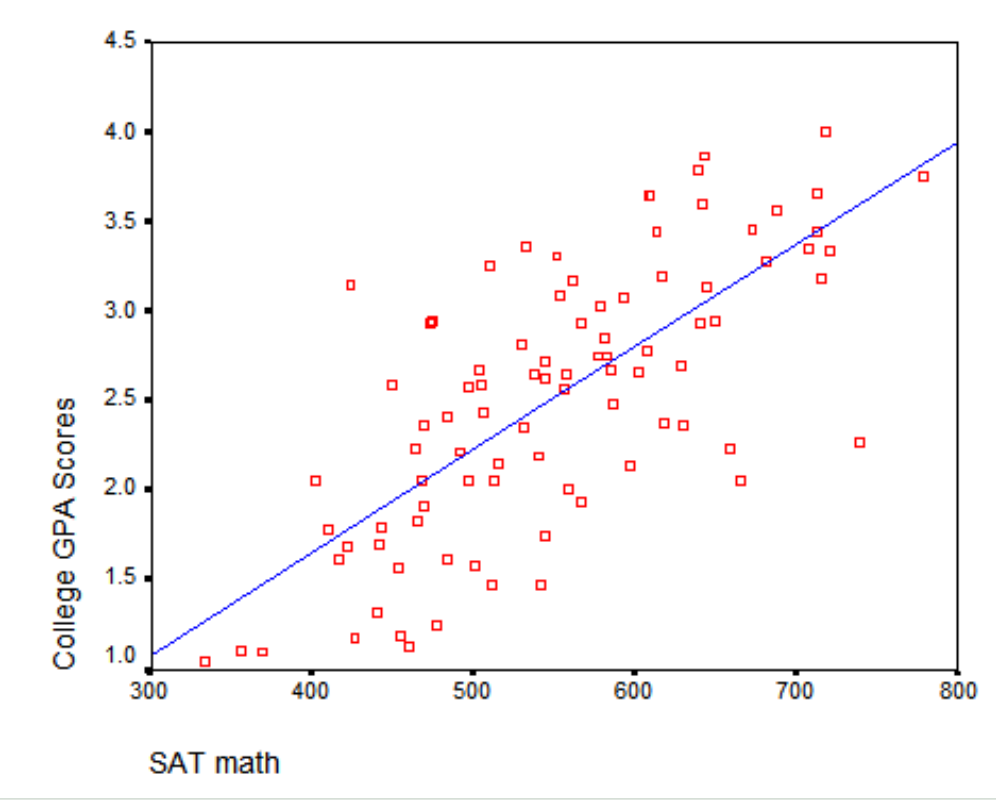


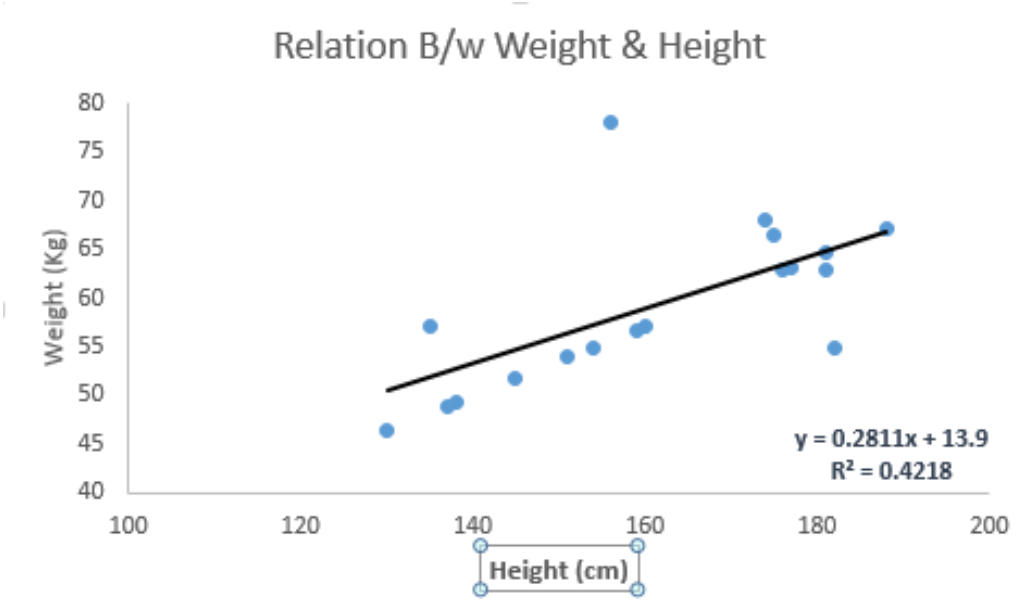
$$Total_{Paid} = 0.0954 * 1200 + 0.001 \equiv 114.5$$

	A	B
1	Total Payed	Total Miles
2	36.66	390
3	37.05	403
4	34.71	396.5
5	32.5	383.5
6	32.63	321.1
7	34.45	391.3
8	36.79	386.1
9	37.44	371.8
10	38.09	404.3
11	38.09	392.6
12	38.74	386.49
13	39	395.2
14	40	385.50
15	36.21	372
16	34.05	397
17	41.79	407
18	30.25	372.33
19	38.83	375.6
20	39.66	399

MUNI FI

Lineárna regresia



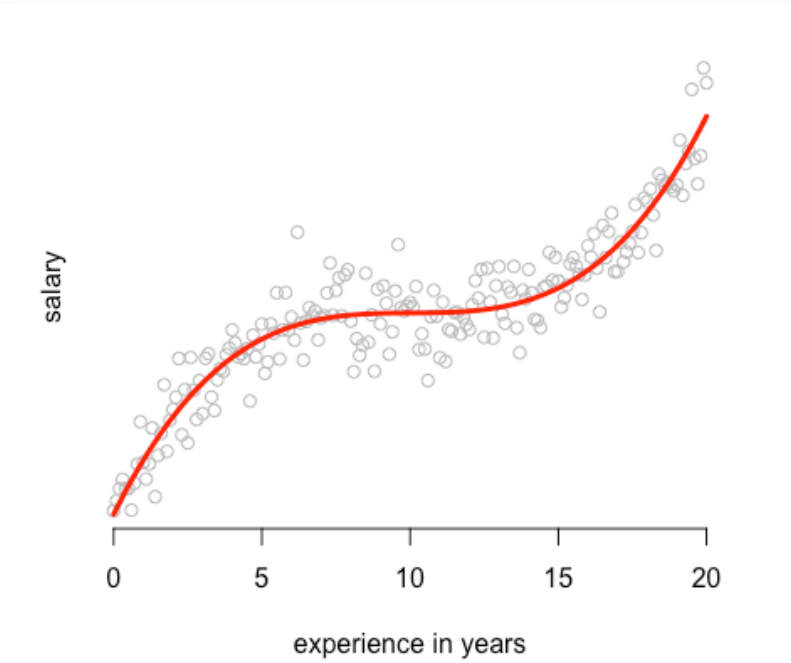


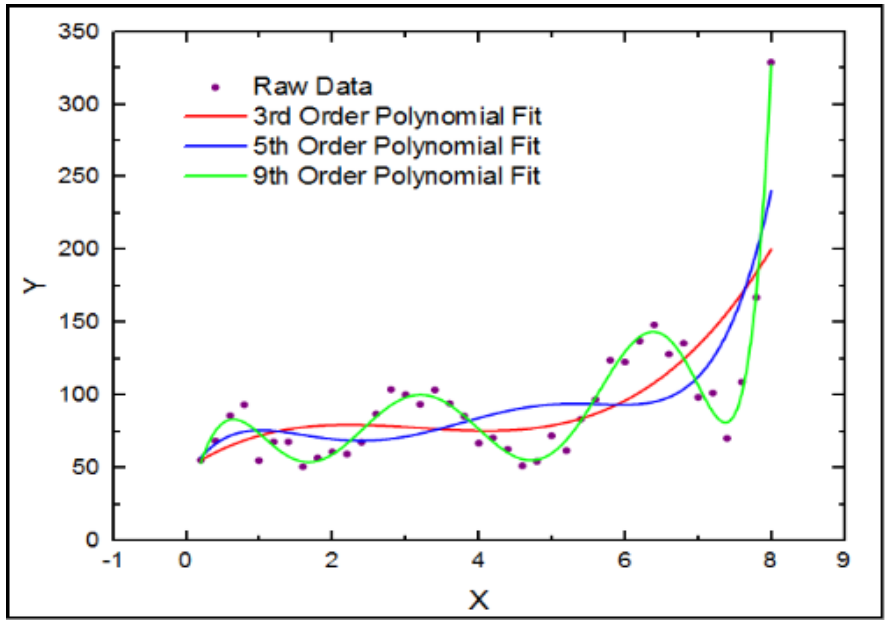
MUNI FI

Polynomiálna regresia

$$Y = a + bX + cX^2 + \varepsilon$$

$$Y = a + bX + cX^2 + dX^3 + \cdots X^n + \varepsilon$$





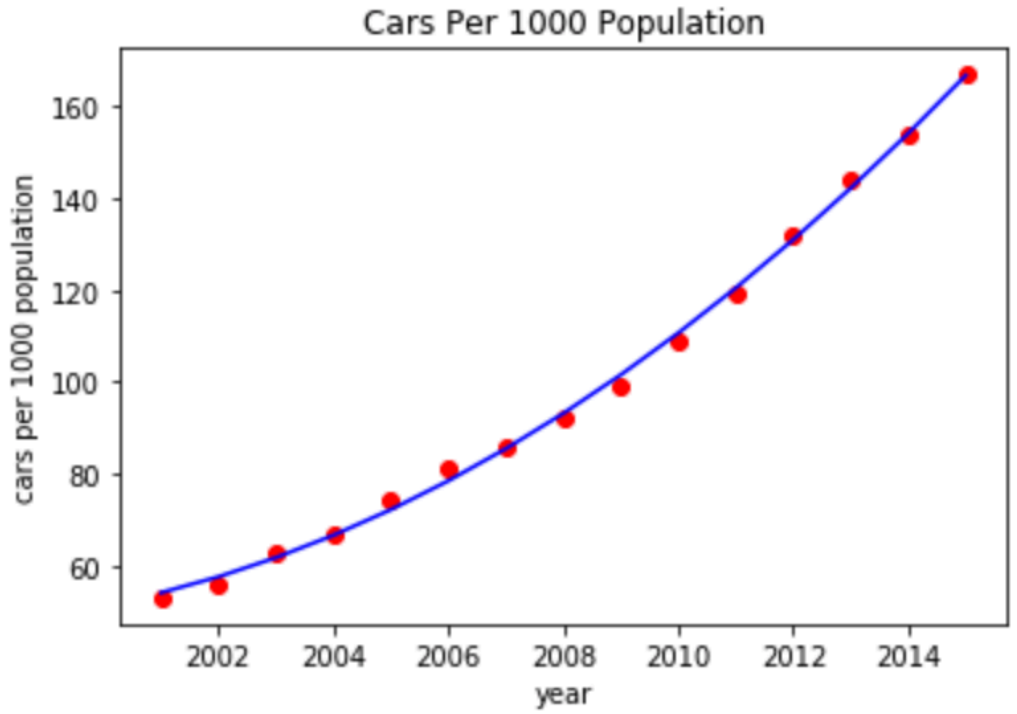
MUNI FI

Polynomiálna regresia

- Meranie progresu epidémie
- Tempo rastu a obnovy tkaniva
- Distribúcia izotopov uhlíka v sedimentoch jazier.

Polynomiálna regresia

– Predaj áut

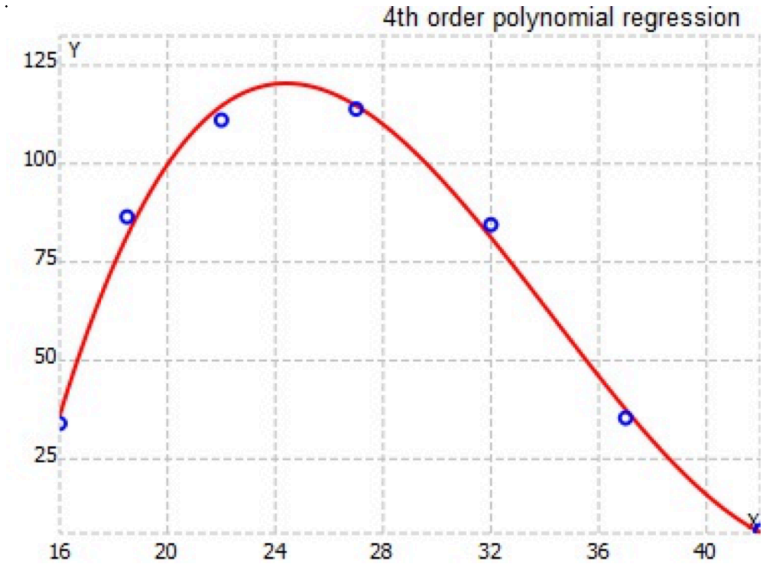


MUNI FI

Polynomiálna regresia

– Priemerný počet narodení detí u žien rôzneho veku

Age	Births/1000 women
16	34
18,5	86,5
22	111,1
27	113,9
32	84,5
37	35.4
42	6,8
45	We will calculate this value



$\text{polyfit}(xz, 4) = -5.368\text{e-}5\, x^4 + 0.038\, x^3 - 3.479\, x^2 + 105.811\, x - 916.689$ $\text{replace symbols}(\text{polyfit}(xz, 4), x, 45) = 2.601$

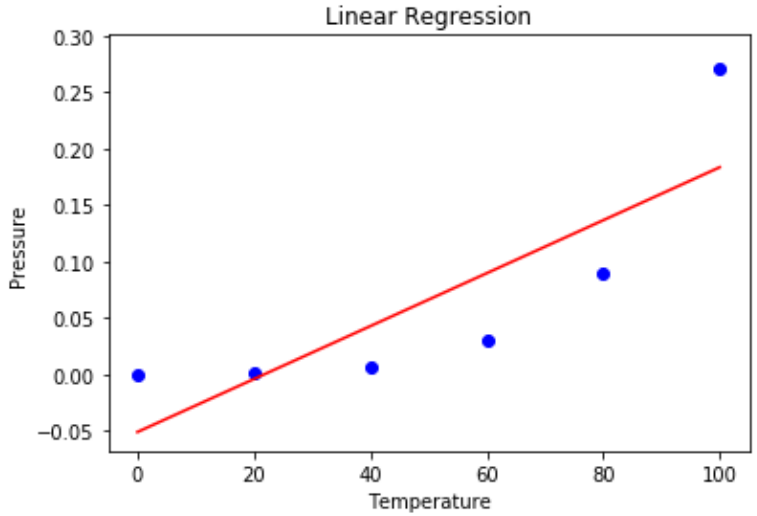
The average number of births per 1000 women of age 45 is 2.6.

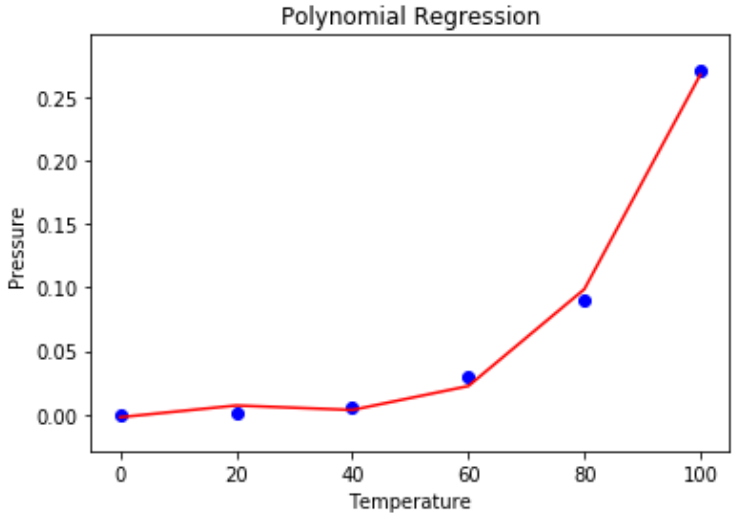
MUNI FI

Polynomiálna regresia

Out[3]:

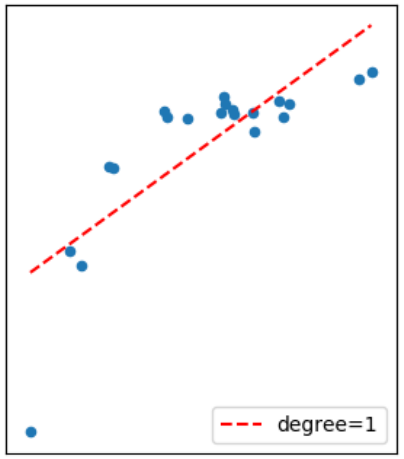
	sno	Temperature	Pressure
0	1	0	0.0002
1	2	20	0.0012
2	3	40	0.0060
3	4	60	0.0300
4	5	80	0.0900
5	6	100	0.2700



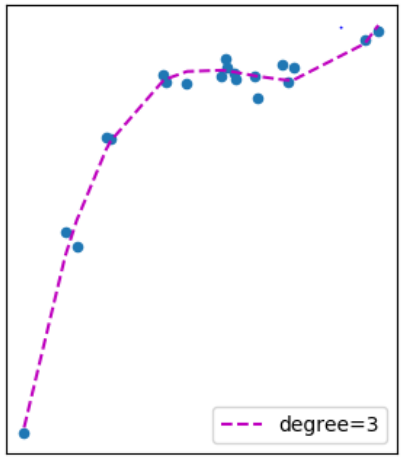


MUNI FI

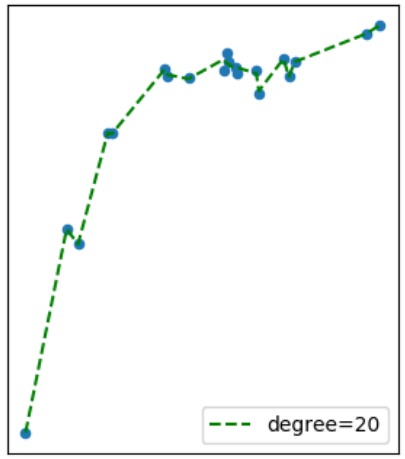
Polynomiálna regresia - overfitting



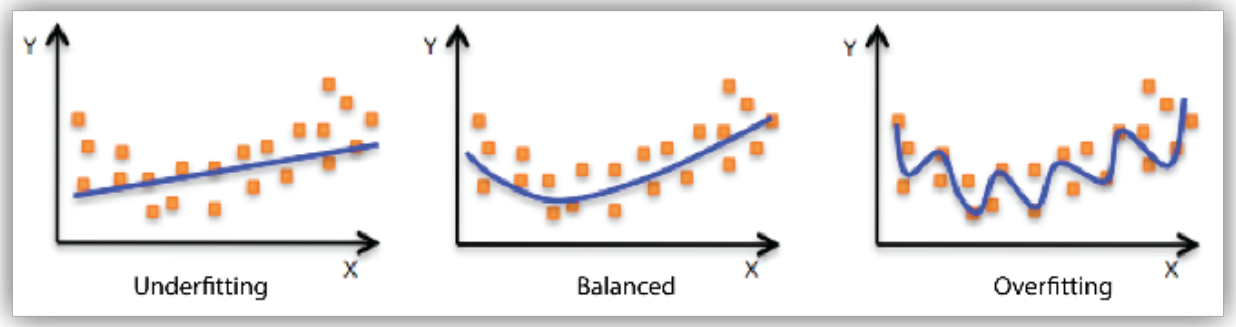
Underfit
High Bias
Low Variance



Correct Fit
Low Bias
Low Variance



Overfit
Low Bias
High Variance



MUNI FI

Zdroje

- https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/
- https://www.listendata.com/2018/03/regression-analysis.html
- https://www.nosimpler.me/polynomial-regression/
- https://www.freecodecamp.org/news/learn-how-to-improve-your-linear-models-8294bfa8a731/
- https://www.theanalysisfactor.com/linear-regression-outcome-boundaries/

Zdroje

- https://towardsdatascience.com/polynomial-regression-bbe8b9d97491
- https://www.geeksforgeeks.org/python-implementation-of-polynomial-regression/
- https://www.quora.com/What-are-the-best-use-cases-of-regression
- http://www.mas.ncl.ac.uk/~nag48/teaching/MAS1403/notes7slr.pdf
- https://statisticsbyjim.com/regression/predictions-regression/
- https://towardsdatascience.com/linear-regression-in-real-life-4a78d7159f16

Zdroje

- https://www.quora.com/Whats-the-point-of-polynomial-regression-if-I-can-just-use-multiple-linear-regression
- https://labdeck.com/examples/curve-fitting/curve-fitting-in-real-life.pdf?39fd30&39fd30