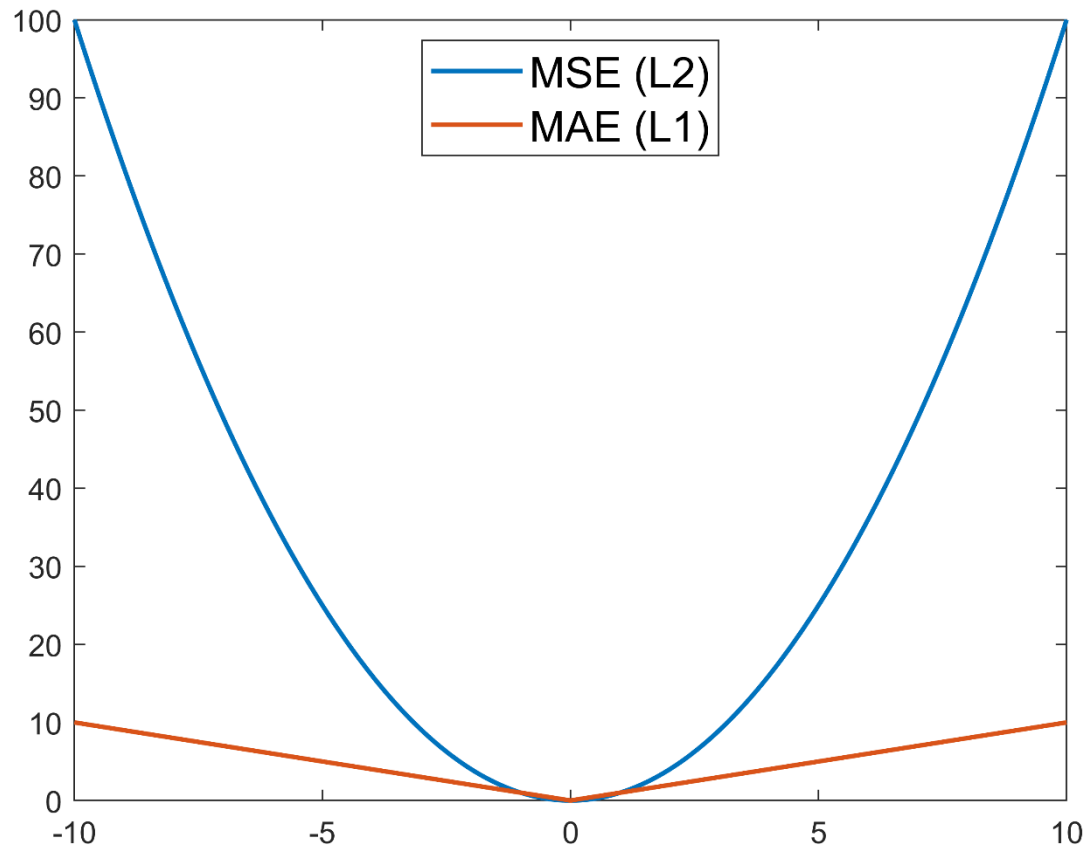


# Funkcija greške (gubitak)

- Mera koliko je model dobar u predviđanju očekivanog ishoda
- Ni jedan gubitak nije univerzalno dobar za sve probleme
- Šta će biti najbolje zavisi od:
  - Prisustva *outlier*-a
  - Modela koji koristimo
  - Vremenska efikasnost optimizacionog algoritma
  - Jednostavnost pronalaska izvoda
  - Samog problema

# Poređenje: L2 i L1 gubitak



$$MSE = \frac{1}{N} \sum_{i=1}^N (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

Jednostavno za minimizaciju

$$MAE = \frac{1}{N} \sum_{i=1}^N |h_{\theta}(x^{(i)}) - y^{(i)}|$$

Robustnije u odnosu na *outlier*-e

# Poređenje: L2 i L1 gubitak

Koren MSE  
(ista jedinica kao MAE)

MAE vs. RMSE for cases with slight variance in data

ID	Error	Error	Error <sup>2</sup>
1	0	0	0
2	1	1	1
3	-2	2	4
4	-0.5	0.5	0.25
5	1.5	1.5	2.25

MAE: 1

RMSE: 1.22

Predikcije bliske tačnim vrednostima.  
Greška (po opservacijama) ima  
manju varijansu

MAE vs. RMSE for cases with outliers in data

ID	Error	Error	Error <sup>2</sup>
1	0	0	0
2	1	1	1
3	1	1	1
4	-2	2	4
5	15	15	225

MAE: 3.8

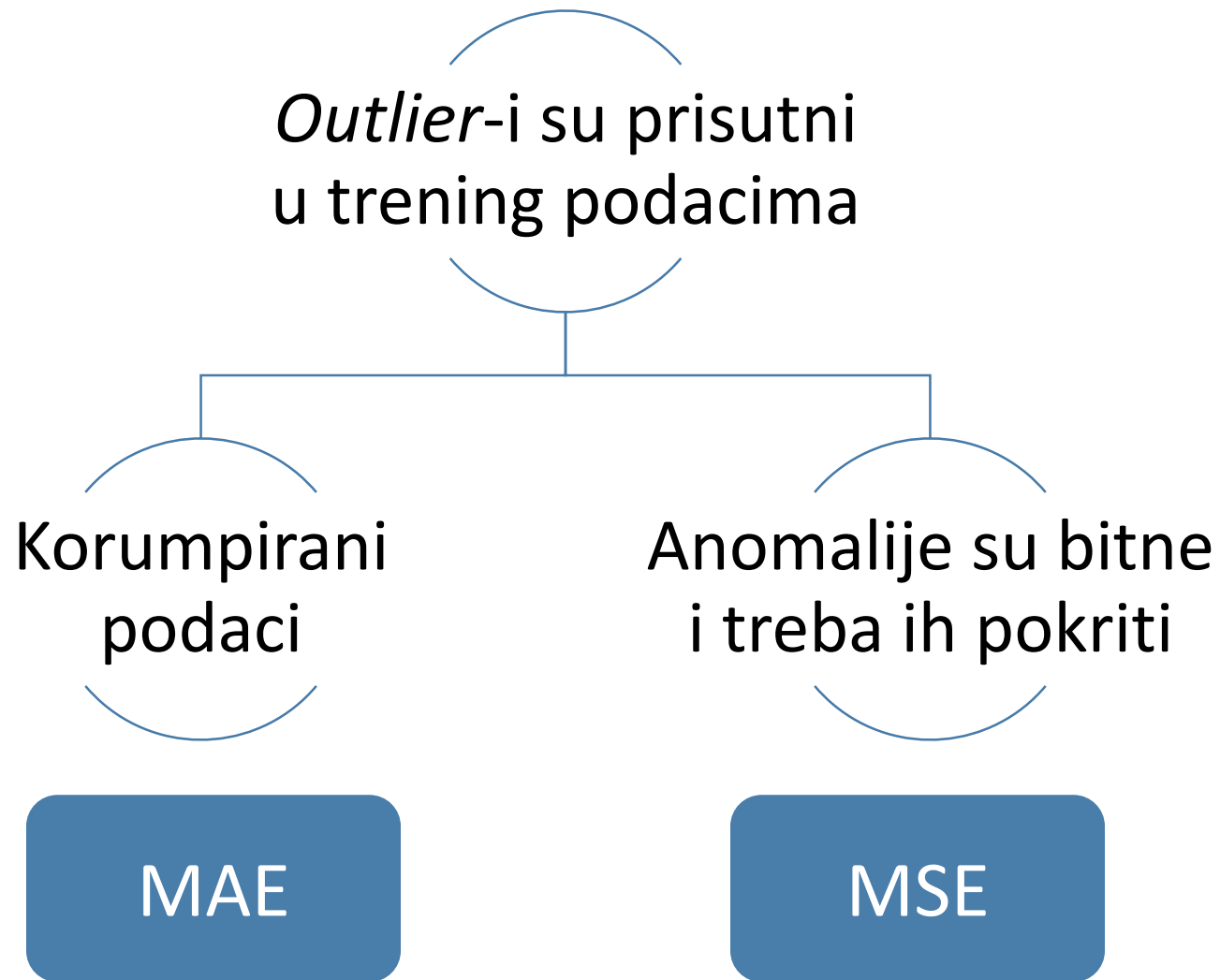
RMSE: 6.79

outlier

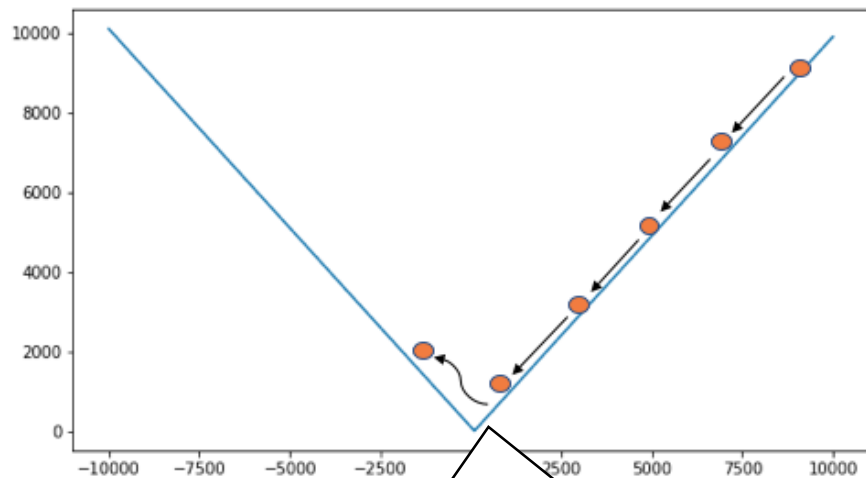
Ogromna greška

MSE daje mnogo veću težinu  
*outlier*-ima u odnosu na MAE

# L2 i L1 gubitak – *outlier*-i

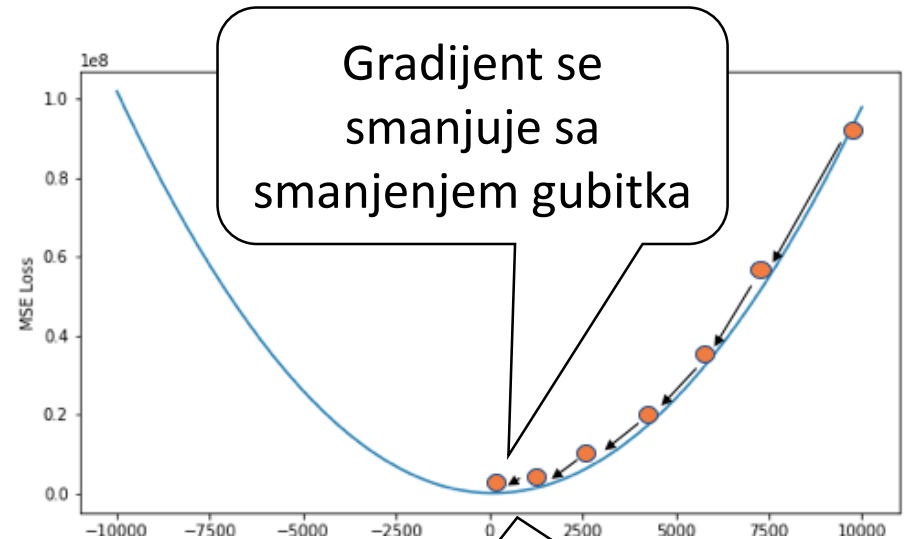


# L2 i L1 gubitak – treniranje modela



Gradijent je uniforman (čak i za male vrednosti gubitka)

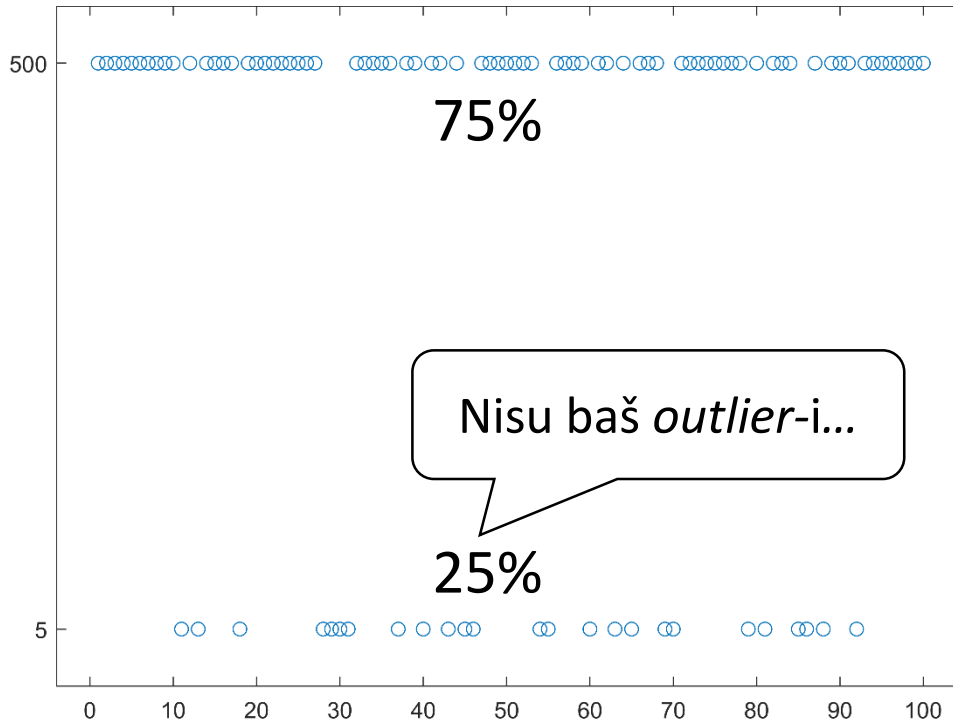
Dinamički smanjivati  $\alpha$  kako se približavamo minimumu



Gradijent se smanjuje sa smanjenjem gubitka

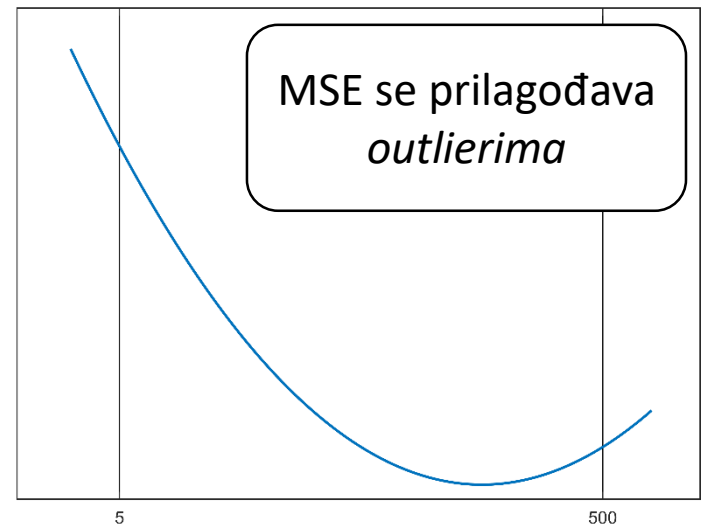
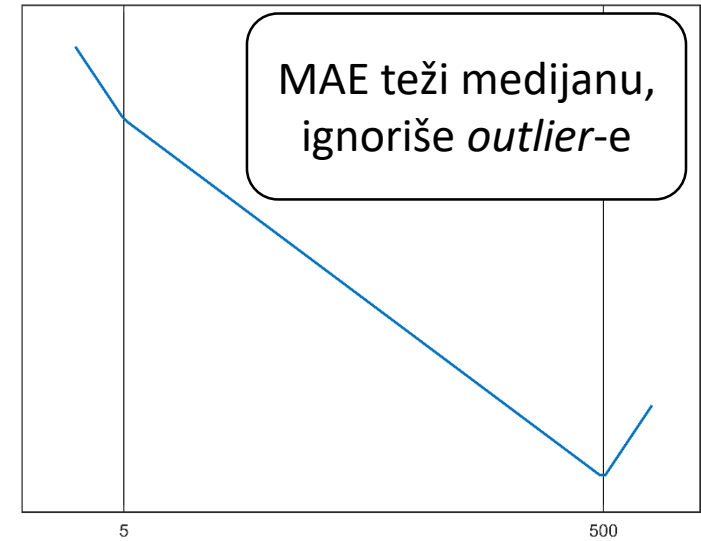
Konvergira za fiksno  $\alpha$

# Poređenje: L2 i L1 gubitak



Ni MAE ni MSE nije dobar izbor

Postoje alternative



# Funkcija greške

- Još jedna stvar koju moramo uzeti u obzir je sam **problem**



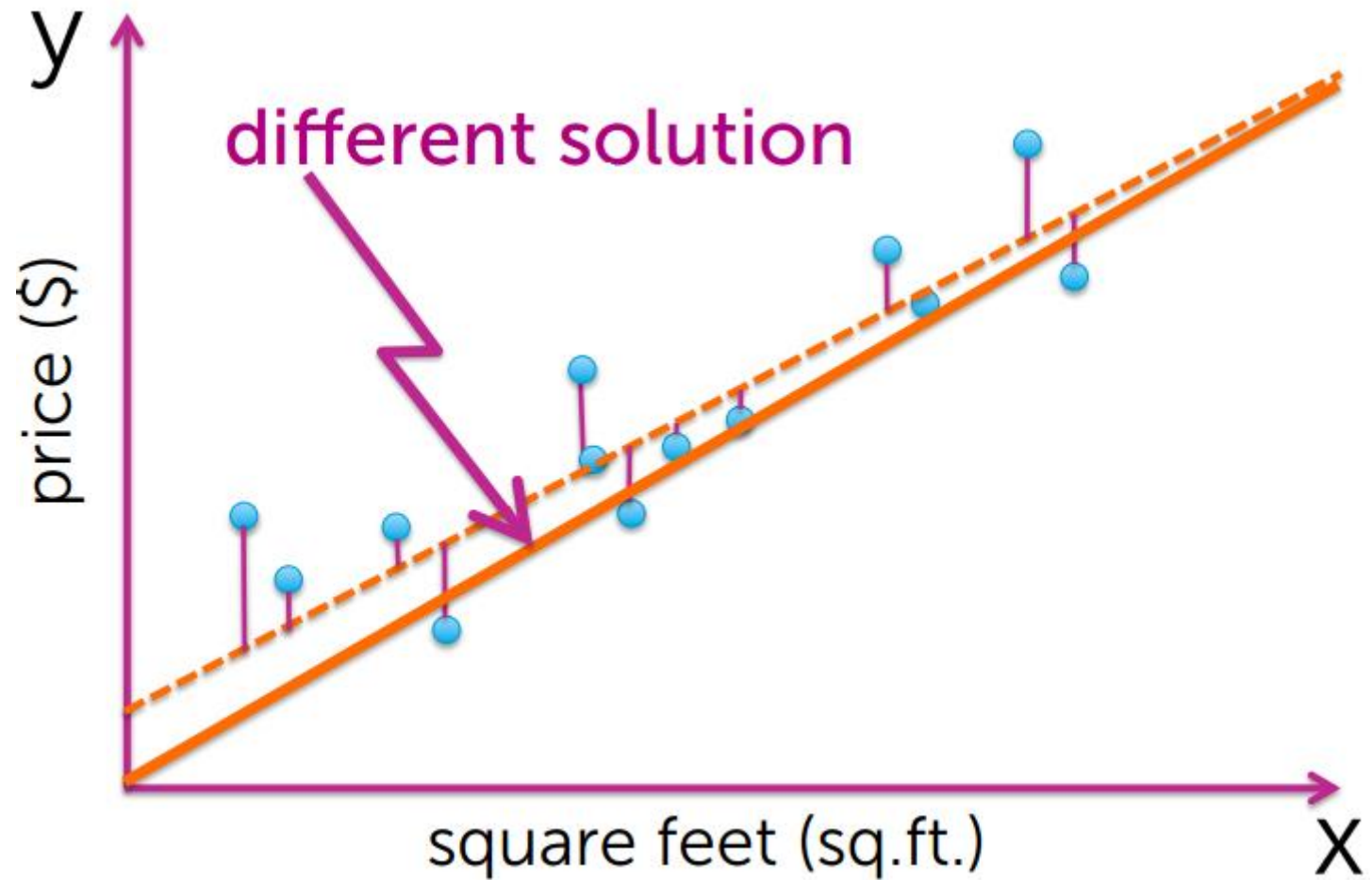
Ako model podbaci...



Ako model prebaci...

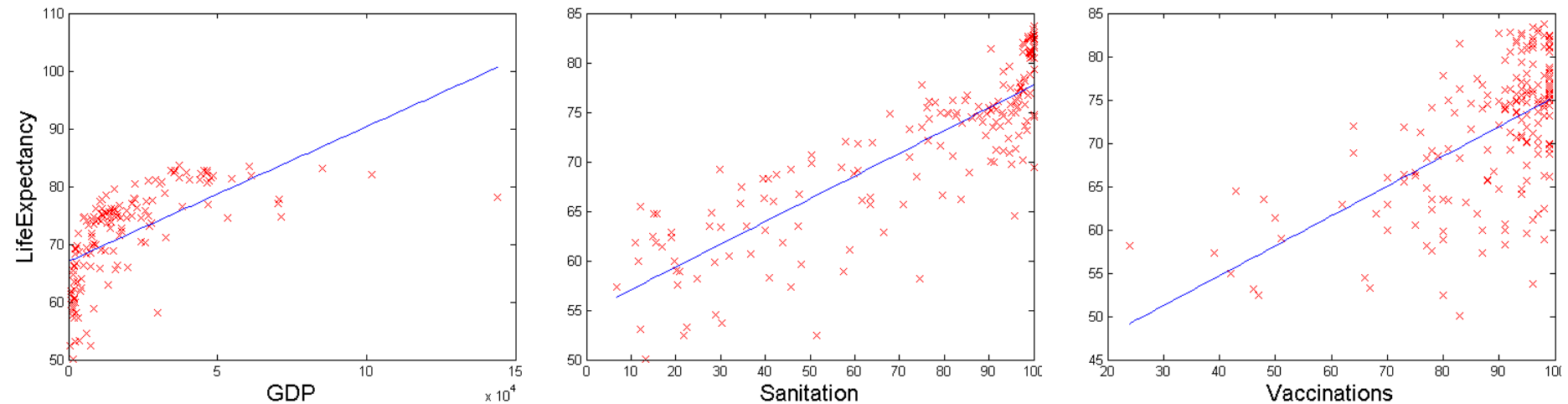
Ovo je možda gore:  
Asimetrična funkcija greške

# Asimetrična funkcija greške





# Modelovanje životnog očekivanja



- Prikazana je zavisnost očekivanog životnog veka u zemlji od:
  1. GDP (bruto domaći proizvod)
  2. Sanitation (% stanovništva koji ima pristup adekvatnim sanitarijama)
  3. Vaccinations (% jednogodišnjaka koji je primio adekvatne vakcine)
- Vidimo da na slikama postoji određen trend (zavisnost) u podacima (sumarizovan regresionim linijama)
  - (Vizuelnom inspekcijom) mogli bismo reći da je trend na 1. i 2. slici izraženiji u odnosu na 3.
- U ovakvim situacijama bismo želeli da znamo udruženu (*joint*) zavisnost između ciljne varijable (životnog očekivanja) i sve 3 ulazne varijable:

$$\text{LifeExpectancy} \approx h(\text{GDP}, \text{Sanitation}, \text{Vaccinations})$$