

HLEDÁNÍ ZNALOSTÍ V DATECH

ROMAN NERUDA

Brno, 18.4.2018

Typy strojového učení

- **Unsupervised:**
 - učení bez učitele
 - hledají se zástupci dat
 - shlukování
- **Supervised:**
 - učení s učitelem
 - data obsahují vstupy a požadované výstupy
- **Reinforcement:**
 - posilované/zpětnovazební učení
 - není k dispozici okamžité ohodnocení
 - učení chování robotů

Datakopectví/Data mining

Hledání zajímavých (nových, netriviálních, užitečných) informací v datech.

- **Predikce:** předvídání některých dat na základě závislosti na jiných datech
 - klasifikace, regrese, detekce anomálií
- **Deskripce:** popis dat tak, aby jim rozuměl člověk
 - pravidla, vzory, ...

Data mining

NENÍ

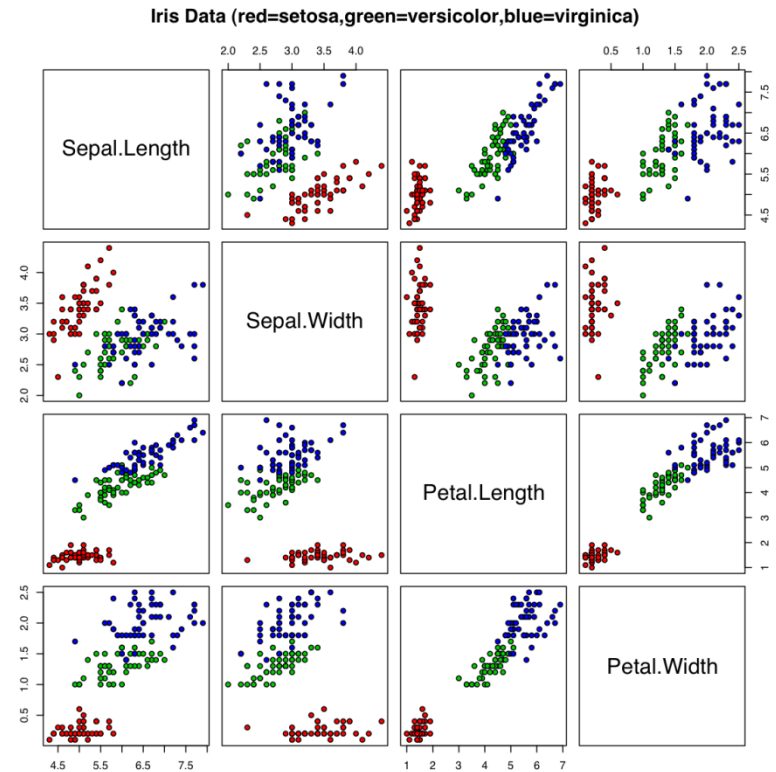
- Najít spolužáka na Facebooku
- Najít na Google, co to znamená kopulace

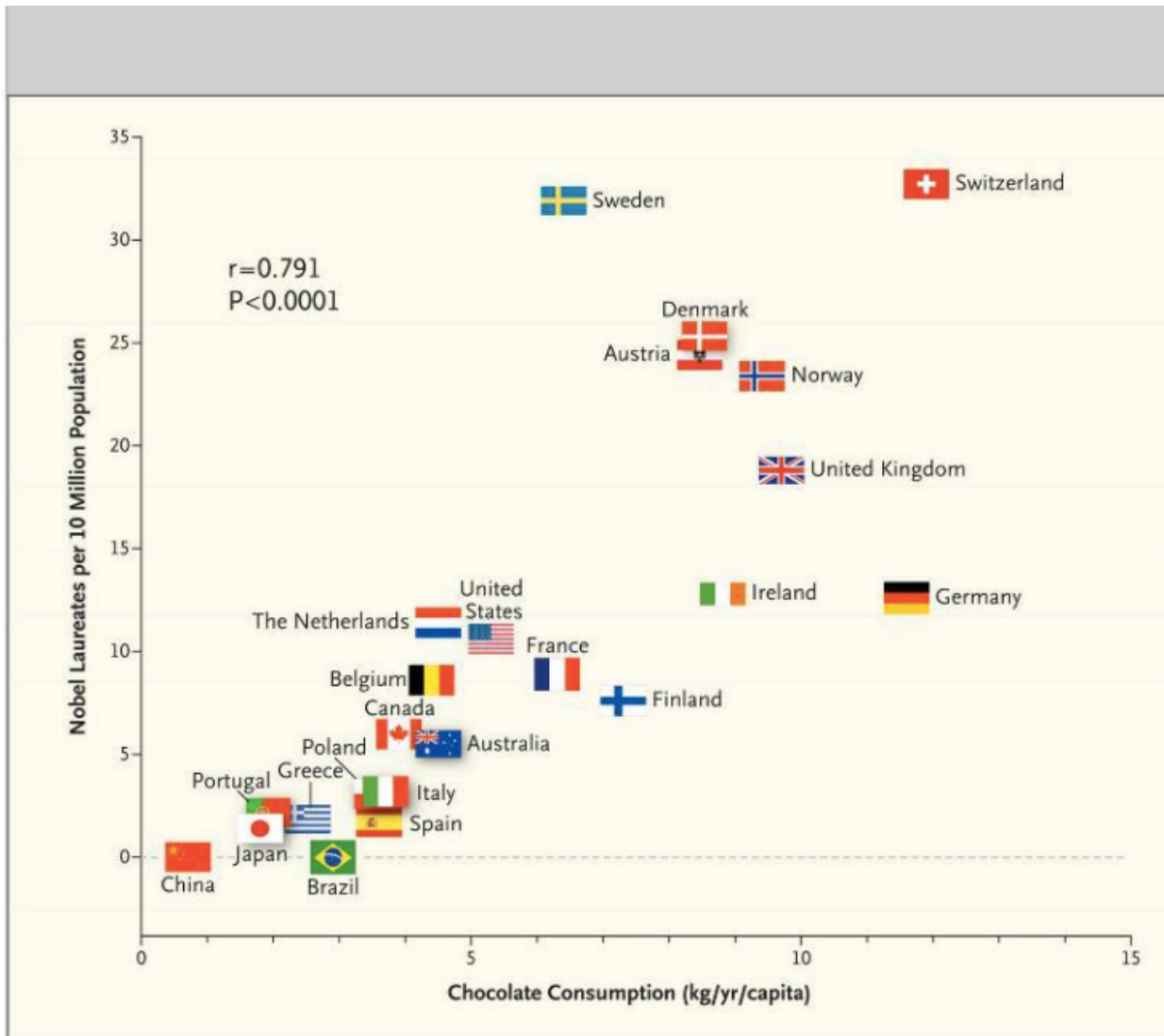
JE

- Najít shluky těsněji propojených uživatelů Facebooku
- Rozlišit dokumenty o kopulaci vrácené Googlem do 3 kategorií:
 - roubování,
 - chemická reakce,
 - rozmnožování

Kosatce

Sepal Length	Sepal Width	Petal Length	Petal Width	Species
5,1	3,5	1,4	0,2	<i>Iris setosa</i>
4,9	3	1,4	0,2	<i>Iris virginica</i>
		...		
5,9	3	5,1	1,8	<i>Iris versicolor</i>





Závisí Nobelova cena na spotřebě čokolády?

playground.tensorflow.org



Data mining party



ITS TIME TO MEASURE UP!

