* Datovy typy
* Distribuce / histogramy features
* Analyza zavislosti : kategoricke : chi-kvadrat, kontingencni tabulka, phi koeficient, point-biserial

Spojite : boxplot pro kazdou s y=0 a y=1

* Stratifikovavy rozdeleni na training a testing
* Training data oversampling and / or undersampling
* Binning
* Feature selection
* Modely (logist. Regrese, forrest, tree, gradient boost ...)
* Vybrat model na zaklade performance