Tisztelt Bioinformatika Tanszék!

A megoldásomhoz Python nyelvet használtam és a **train.csv** file használatát vettem alapul.

Alapvetően a következő pipline elvet követtem:

* Python script *pandas lib*, használatával *dataframe*-be olvassa csv file-t
* Ezt követően a kemoterápia vagy hormonterápia alapján sub dataframeket hoztam létre, amelyeken az elemzéseket végeztem.
  + Összesen 8 sub dataframet hoztam létre.
  + A hormon terápia 1/0 adj. Kemoterápia 1/0 viszonylatában 4-et
  + Illetve individuálisan 2-2-t.
  + Minden elemzést ezen a 8 különböző viszonyban néztem.
* Elemeztem az adatokat
  + „grade”,
  + „pathologic\_stage”,
  + „grade” és „size\_documented”
  + „grade” és „age\_at\_diagnosis” relációjában
* Eredményeimet **train\_statistics.csv** file-ban tettem közzé. Itt számos egyéb oszlop eredményét vettem alapul így pfs, dds,

Készítette Dr. Dul Zoltán,

a Semmelweis EMK 2. Éves hallgatója