

V *datasetu* je 200 pacientů s onemocněním X. Vaše znalost o procedurách léčby není expertní, víte ale, že bývá nějakou kombinací prováděných výkonů 1 a 2 a aplikací léčebných přípravků 1 a 2. Jelikož si ani odborná komunita není zcela jistá, jak je vhodné postupovat a je tak mnoho různých přístupů, vaším úkolem je pokusit se léčbu unifikovat.

Popis souboru:

Indexové proměnné: ID (pouze označení pacienta)

Popisné proměnné: POHLAVÍ, VĚK, VÝŠKA, VÁHA

Parametry prováděné léčby: VÝKON1, VÝKON2 (v počtech provedení), LÉČIVO1, LÉČIVO2 (v množství podané látky)

Výsledek léčby: ÚČINEK (kvantitativní proměnná udávající jak moc se léčba povedla, přičemž vyšší číslo značí lepší výsledek)

1. Podívejte se na data, zda s nimi není něco špatně. Popřípadě je fajn si je nějak pěkně vizualizovat.
2. Pokuste se odhalit, jak si obecně ošetřovatelé počínají.
3. Zhodnoťte, jestli je mezi vámi odhalenými postupy výrazný rozdíl ve výsledcích léčby, popřípadě která se vám zdá nejúčinnější (resp. jestli je některá výhodnější pro určitou subpopulaci atd.)

Co bych očekával:

- Nějakou formu základní vizualizace dat
- Jednoduché slovní nebo grafické odůvodnění toho, jak si myslíte, že ošetřovatelé postupují
- Nějakou výsledkovou tabulku se srovnáním různých druhů léčby
- Libovolný jiný komentář, kterým mě potěšíte
- Snesu i komentář, který mě nepotěší, ale bodu výše dávám přednost

Pokud to budete počítat v Rku, tady jsou vypsány některé užitečné funkce (není nutno je použít, pokud by se vám nehodily):

- *dist()*
- *hclust()*
- *cutree()*
- *aov()*