**1. Gyakorló feladatsor, 2025. március 06-ig (Leíró statisztika, Eseményráta, Regresszió, Korreláció, Power analízis)**

**Elvégzendő, gyakorlandó mindig a következő előadás előtt**. A konzultáció során a feladatmegoldás során felmerülő problémákat beszéljük meg, vagy a feladat megoldását mutatjuk meg.

A **március 13-i első vizsga** hasonló feladatokkal, a tematikában jelzett időpontban, „élőben” (online, saját gépen) kerül megrendezésre, annak a megoldását (statistica output) kell majd emailen megküldeni a [klinbiostat2021@gmail.com](mailto:klinbiostat2021@gmail.com) email címre

A honlapon (https://semmelweis.hu/kutlab/oktatas/klinikai\_biostatisztika\_2025/) található gyakorló adatbázisok (**Ráta\_adatbázis**, **Down\_hipotetikus\_példa**, **MintaAdatbázis2025**) felhasználásával oldd meg az alább feladatokat!

**2025. február 20-ig:**

1. A **RátaAdatbázist** importáld Statistica-ba, számítsd ki az annualizált relapszus rátát, és készíts leíró statisztikát a teljes adatsorról (20 eset), valamint külön-külön az 1-es és 2-es csoportkódnak megfelelő alcsoportokról (10-10 eset). A leíró statisztika tartalmazza a 95%-os konfidencia intervallumot, illetve a 10. és 90. percentilis értékét is! Mit jelent a 95%-os CI tartomány ebben az esetben?
2. Készíts leíró statisztikát és hisztogrammokat a **MintaAdatbázis2025** összes változójáról, keresd meg, hogy vannak-e biológiai nonszensz értékek, ha igen, készíts adatbázis naplót, írd le mit találtál és mit változtattál (helyettesítsd átlaggal a biológiai nonszenszeket, a hiányzó értékeket hagyd üresen).
3. Számold ki egy új változóban a betegek alapbetegségeinek összesített darabszámát (HT, DM, szív, tüdő, malignitás). Adj leíró statisztikát (frequency table) az értékekről! Ábrázold scatterplot és box-and-wishkers ábrázolással a limfocitaszámot az alapbetegségek száma szerinti csoportokban! Számold ki a **MintaAdatbázis2025** „Lázas epizódok száma követés során” változó annualizált esemény rátáját a „követés ideje elbocsátás után” oszlop segítségével (95%-os konfidencia intervallummal).

**2025. február 27-ig:**

1. **MintaAdatbázis2025:** Összefügg-e a CRP (mg/L) a fehérvésejtszám (WBC, G/L) értékkel? Pearson's korreláció számítás a 10-es alapú logaritmizált CRP és WBC értékekkel (új változóban), és nem-paraméteres rank-korrelációszmítás az eredeti CRP és WBC értékek között. Hogy értelmezed a két eredményt? Mennyi az összefüggés powere?
2. **MintaAdatbázis2025:** Meghatározza-e a fehérvérsejtszám a CRP értéket? Készíts regressziós modellt a nyers értékekkel, értékeld az egyenlet együtthatóit, írd le szövegesen az összefüggést! Ismételd meg úgy, hogy a hiányzó értékeket "mean inputation" módszerrel pótlod. Változott az eredmény lényegesen? Mit gondolsz, miért?