

Harjoitus 1

1. Lataa tiedosto `no_mgrs_all_details.csv` R:ään. Laske puuttuvien arvojen lukumäärä jokaiselle aineiston muuttujalle. Katso `(which() is.na())` ja `length()` funktiot sekä `for` silmukka)
2. Muuta ammattinimikkeet (`title`) merkkijonoista faktoreiksi ja laske eri ammattinimikkeiden lukumäärä.
3. Koosta `no_mgrs_all_details.csv` tiedostosta sellainen havaintomatriisi, jossa on mukana pelkästään numeeriset muuttujat.
4. Tutki löytyykö muodostamasi numeerisen aineiston minipalkkaan liittyvistä muuttujista poikkeavia arvoja. (`boxplot`)
5. Muodosta jokaiselle ammattinimikkeille 1 ja 6 omat havaintomatriisit, joissa on mukana pelkästään numeriset muuttujat. Laske muuttujien keskiarvot ja tulosta ne samaan kuvaan. (`colMeans()`, `plot()`, `lines()`)
6. Laske ammattiryhmille 1 ja 6 työkuukausien ja keskipalkan korrelaatio. (`cor()`)