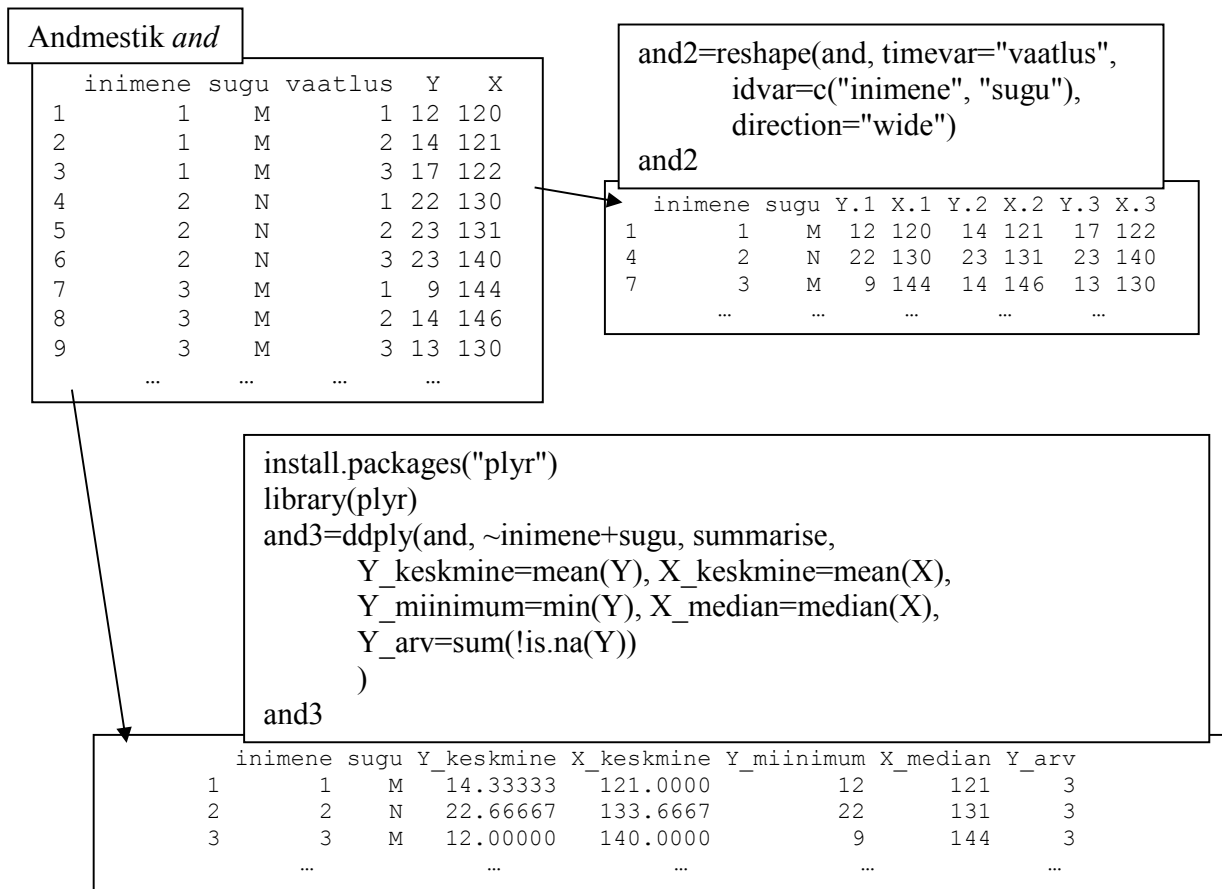


## Nõuded projektile – iseseisvale andmeanalüüsitööle.

1. Kirjelda oma andmestikku. Kui suur on valim, keda uuriti (kes või mis on uurimisobjekt), kes või mis moodustavad uuritava populatsiooni. Kirjelda ka seda, kuidas on valim võetud/moodustatud (katseplaan) ja mis eesmärgil on valim võetud (mida sooviti uurida).

Kui esialgses andmestikus on ühe uurimisobjekti kohta mitu rida, siis kaalu võimalusi koondada sama uurimisobjekti andmed ühele reale. All on toodud ära paar võimalust seda R-i abil teha:



Võiksid iseloomustada paari graafiku või tabeli abil ka andmestikus olevaid tunnuseid – milliseid väärtuseid nad võivad omandada, milline on nende jaotus.

2. Vali andmestikust välja kaks tunnust – soovitatavalt olgu üks neist pidev tunnus ja teine väljavalitu olgu kas binaarne (kahe võimaliku väärtusega), nominaalne või diskreetne (millegi loendamisel saadud) tunnus. Loo mõlema tunnuse jaoks statistiline mudel, mis võimaldaks kas nende tunnuste väärtuseid prognoosida või selline mudel, mis kirjeldaks antud tunnuse väärtust mõjutavaid põhjuslikke mõjusid. Pane leitud mudel kirja ja põhjenda, miks otsustasid just nimelt sellise mudeli kasuks (miks ei lisatud mudelisse pidevate tunnuste kõrgemaid astmeid, miks osad tunnustest jäid mudelist välja, miks pole

lisatud koosmõjusid vms). Võiksid seletada ka seda, kuidas jõudsid valitud mudelini. Põhjenda ka enda poolt valitud seosefunktsiooni ja võimaluse korral arutle mudeli eelduste täidetuse üle.

Juhul kui Sulle hingelähedases andmestikus – selles andmestikus, mida (või millesarnast) loodad kasutada näiteks oma magistritöö kirjutamiseks – pole kõiki nõutavaid tunnuseid (pole näiteks pidevaid tunnuseid), siis ära selle pärast veel loobu oma andmestikust! Kui andmestikus tõepoolest pidevat tunnust pole, siis vaid sellepärast ei jää sa veel maksimumpunktidest ilma – kui ülejäänud analüüs ja esitatud mudelid on mõistlikud!

3. Kirjelda Sinu poolt leitud mudelit (mudeleid). Tee selle mudeli prognooside kirjeldamiseks või tunnustevaheliste seoste iseloomustamiseks graafikuid (või esita tulemusi tabelitena). Näita, milliseid prognoose pakub sinu mudel erinevate argumenttunnuse väärtuste korral, iseloomusta ka seda, kui täpselt on uuritava tunnuse tinglik keskvärtus sulle ikkagi teada (proovi näidata graafikutes või tabelis ka usaldusintervalli – või soovi korral prognoosiintervalli). Kommenteedi leitud seoseid ja prognoose – kas midagi tundub üllatavat? Või on leitavad prognoosid vastavuses Sinu ootustega?

Töö vormistamisest: töö põhiosas peaks võimaluse korral vältima R-i käskude ja pikkade väljundite näitamist, tekstis maini ainult neid näitajaid ja tulemusi, mis otseselt antud kohas vajalikud on. Analüüsitulemusi/grafikuid genereerinud programmi (ja miks mitte ka väljundi) koht on töö lisas. Loomulikult ei tee halba, kui töö on korrektselt vormistatud – tiitellehega varustatud, sissejuhatuse-ülesandepüstitusega jne.