

Korrelaatio-analyysin eteneminen kannattaa hahmottaa askeleittain:

(Huom. ks. tiivistelmät ja kertaava tarvittaessa asiat tiivistelmäpaketista 2)

1. Ensin lomakkeen ja aineiston selailu ja sopivan kysymyksen & muuttujien etsiminen
2. Keksitään ja kirjataan tutkimuskysymys
3. Kirjataan hypoteesit (nollahypoteesi ja tutkimushypoteesi)
4. Tehdään tarvittavat muokkaukset muuttujille: etenkin MISSING/puuttuvan arvon määrittely, jos EOS-vastausvaihto on mukana; mahdollisesti myös muuttujan arvojen kääntäminen, jos kulkevat epäloogiseen suuntaan.
5. Selvitetään millaisia muuttujat ovat eli mietitään, millainen niissä on vastausasteikko ja otetaan tunnusluvut (vinous, huipukkuus) frekvenssi-toiminnolla:
 - jos **molemmat** ovat **mitta-asteikoltaan jatkuvanluonteisia** ja jakaumaltaan riittävän **normaalijakautuneita** → voidaan käyttää parametrista Pearsonin korrelaatiotestiä
 - jos ei (jos toinen tai molemmat järjestysasteikollinen tai ei riittävästi normaalijakautunut) → käytetään Spearmanin epäparametrista korrelaatiotestiä
6. Tehdään korrelaatiotesti → tulkitaan testin tulos → huom. on katsottava erikseen (a) onko havaittu korrelaatio tilastollisesti merkitsevä ja (b) kuinka vahva se on eli mikä on kerroin → todetaan kumpi hypoteesi jää voimaan
7. Tehdään pisteparvikuvio (*scatter*) Chart builderilla (huom. molempien kuvaajassa käytettävien muuttujien tyypiksi määritettävä SPSS:ssä *scale* vaikka olisivatkin oikeasti järjestysasteikkoisia, muuten ohjelma ei anna tehdä muokkauksia) → muokataan kuvaaja → Options: *Bin element* (ja jos halutaan, niin niiden muokkaus *Properties*-ikkunan *Marker*-välilehdestä) → Elements: *Fit line at total* (eli lisätään korrelaation suuntaa kuvaava ns. regressiosuora)
8. Lopuksi lasketaan ns. selitysaste prosentteina, **jos** korrelaatio on ollut tilastollisesti merkitsevä → testin ilmoittama korrelaatiokerroin toiseen potenssiin eli kerrotaan korrelaatio itsellään (R^2 eli $r \cdot r$) → esim. jos korrelaatiokerroin $-0,132$, on selitysaste: $-0,132 \cdot -0,132 = 0,0174$ → kun tämä kerrotaan sadalla, saadaan selitysaste prosentteina: $100 \cdot 0,0174 = 1,7\%$ → tämä on siis prosenttiosuus, jonka verran muuttujien vaihtelu on niille yhteistä, eli minkä verran voidaan ajatella niiden selittävän toisiaan (kuten huomataan, selitysaste jää pienillä korrelaatioilla usein hyvin matalaksi)