**Python ja data-analytiikka**

**Palautettava tehtävä 8**

Tässä harjoituksessa tutustutaan merkitsevyystestaukseen

Palauta kaikki tehtävät Moodleen yhtenä pdf-tiedostona. Liitä tiedostoon myös käytetty Python-koodi. Esitä testien perusteella tehtävät johtopäätöksen selkeästi.

TEHTÄVÄ 1

Tutki t-testin avulla onko kulttuurialan ja tekniikan opiskelijoiden välillä tilastollisesti merkitsevä ero opinnäytetyön tekemiseen käytetyssä ajassa. Käytä muuttujaa *Opinnäytetyön tekemisaika työviikkoina (40 h) aihekuvauksen tekemisestä työn valmistumiseen:työviikkoa*.

Poista aineistosta sellaiset rivit, joissa ko. muuttujan puuttuu.

Palautuksessa on oltava seuraavat asiat:

* Molempien ryhmien keskiarvoaika ja 95 %:n luottamusväli taulukkoon tulostettuna.
* Varianssien testaus Levenen testillä.
* T-testin käyttö. Argumentit on oltava oikein Levenen testin perusteella.
* T-testin testisuureen ja p-arvon tulostus.
* Nollahypoteesi sanallisesti näkyvissä.
* Johtopäätökset nollahypoteesille t-testin perusteella. Käytetään 95 % -merkitsevyystasoa.
* Sanalliset johtopäätökset.

TEHTÄVÄ 2

Tutki Mann-Whitneyn U-testin avulla onko kulttuurialan ja liiketalouden opiskelijoiden välillä tilastollisesti merkitsevä siinä, miten hyödyllisiksi he kokivat seminaarit. Poista aineistosta ne rivit, joista ko. vastaus puuttuu.

Palautuksessa on oltava seuraavat asiat:

* Molempien ryhmien keskiarvovastaus.
* Mann-Whitneyn U-testin käyttö. Argumentit on oltava oikein.
* U-testin testisuureen ja p-arvon tulostus.
* Nollahypoteesi sanallisesti näkyvissä.
* Johtopäätökset nollahypoteesille U-testin perusteella. Käytetään 95 % -merkitsevyystasoa.
* Sanalliset johtopäätökset.

Hyväksyttyyn suoritukseen vaaditaan jompikumpi tehtävä oikein. Viiteen pisteeseen vaaditaan molemmat tehtävät oikein. Kiinnitä erityistä huomioita nollahypoteesien muodostamiseen sekä johtopäätösten oikeellisuuteen.

Kolmeen pisteeseen sallitaan pientä epätarkkuutta. Viiteen pisteeseen on kaiken oltava oikein.