**Python ja data-analytiikka**

**Palautettava tehtävä 3**

Tee opinnäytetyökyselystä seuraavat ristiintaulukoinnit ja palauta ne Moodleen pdf-tiedostona. Liitä tiedostoon myös tarvittava Python-koodi.

**Ristiintaulukointi 1**

* Opiskeluala vs. sukupuoli
* Huomaa, että aineisto on vanhaa. Nykyään sukupuolta kysytään ainoastaan silloin, kun sille on vahva perustelu. Oikeat vaihtoehdot ovat tällöin mies, nainen, muu, en halua kertoa.   
  Kuva, joka sisältää kohteen teksti, kuvakaappaus, Fontti, numero

  Tekoälyn generoima sisältö voi olla virheellistä.

**Ristiintaulukointi 2**

* Miesten ja naisten suhteellinen jakautuminen eri opiskelualoille.
* Huomaa, että myös miesten ja naisten kokonaismäärä pitää esittää taulukossa.

Kuva, joka sisältää kohteen teksti, kuvakaappaus, Fontti, numero

Tekoälyn generoima sisältö voi olla virheellistä.

**Ristiintaulukointi 3**

* Miesten ja naisten suhteellinen jakautuminen eri opiskelualoille.
* Huomaa, että nyt rivisummat ovat 100 %. Saat tämän tehtyä muokkaamalla crosstab-funktion argumenttia *normalize*. Anna sille arvoksi *index*.
* Huomaa myös, että nyt miesten ja naisten kokonaismäärä on esitetty rivillä. Tämä toteutetaan *loc*-indekserin avulla. Käytä tarvittaessa tekoälyä apuna.   
    
    
    
  Kuva, joka sisältää kohteen teksti, Fontti, kuvakaappaus, numero

  Tekoälyn generoima sisältö voi olla virheellistä.

**Dummy-muuttujan käyttäminen**

* Kuten Taanilan matskussa kerrotaan, dummy-muuttujaa kannattaa käyttää silloin, kun monivalintakysymyksessä saa valita useamman kuin yhden vaihtoehdon. Jos käyttävä on valinnut ko. vaihtoehdon, se saa arvokseen 1. Ei-valituille vaihtoehdoille annetaan arvo 0.
* Alkuperäisessä opinnäytetyökyselyssä ei ollut yhtään tällaista muuttujaa, joten muuttujaa Oliko työ parityö on muokattu. Nyt arvo 1 tarkoittaa kyllä ja arvo 0 ei. Nämä muutokset on tehty M365-kansiossa olevaan tiedostoon, joten lataa tiedosto uudelleen. **Huomaa, että tällaista kysymystä ei kannattaisi koodata dummy-muuttujaksi. Tässä se tehdään esimerkin takia.**
* Tee alla olevat taulukko.
* Voit nimetä DataFrame-olion sarakkeen uudelleen komennolla rename(columns={'vanha\_otsikko': 'Uusi\_otsikko'}, inplace = True/False). Inplace argumentti ei ole pakollinen. Tutki, mitä se tekee.

Kuva, joka sisältää kohteen teksti, Fontti, kuvakaappaus, valkoinen

Tekoälyn generoima sisältö voi olla virheellistä.

Hyväksyttyyn suoritukseen (3 p) vaaditaan kolme tehtävää oikein tehdyksi. Viiteen pisteeseen vaaditaan kaikki neljä tehtävää.