**Python ja data-analytiikka**

**Palautettava tehtävä 2**

Tee opinnäytetyökyselydatasta seuraavat kaaviot ja palauta ne Moodleen pdf-tiedostona. Hyväksyttyyn suoritukseen (3 p) vaaditaan kaksi tehtävää oikein tehtynä. Viiteen pisteeseen vaaditaan kolme tehtävää oikein tehtynä.

1. Piirrä alla oleva histogrammi kysymyksen *Sain riittävästi ohjausta* vastauksista.   
     
   Kuva, joka sisältää kohteen diagrammi, kuvakaappaus, Suorakaide, teksti

   Tekoälyn generoima sisältö voi olla virheellistä.  
     
   Vinkkejä:  
   * Hae Excel-data Pythoniin samalla tavalla kuin edellisessä palautettavassa tehtävässä.
   * Viittaa oikeaan sarakkeeseen komennolla df[’sarakkeen nimi’].
   * Muuta x-akselin arvot komennolla: plt.xticks(range(alaraja, yläraja, väli)).
   * Tekoäly on todella hyvä apu, kun haluat muotoilla piirtämiäsi kuvioita. Oletuksena x-akselin arvot eivät ole keskellä palkkia. Alla on ChatGPT:n vastaus tähän ongelmaan.

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

Tekoälyn generoima sisältö voi olla virheellistä.

1. Tee scatterplot sarakkeista Pystyin itse vaikuttamaan opinnäytetyöni ohjaajan valintaan (riippumaton muuttuja) ja Sain riittävästi ohjausta (riippuva muuttuja).   
   Kuva, joka sisältää kohteen teksti, kuvakaappaus, diagrammi, viiva

   Tekoälyn generoima sisältö voi olla virheellistä.  
     
   Kaavio on piirretty oikein, mutta se on käytännössä julkaisukelvoton. Mieti, miksi kaavio on yllä olevan näköinen. Mieti esim, että miksi siinä näkyy vain muutama piste. Kirjoita vastauksesi palautettavaan pdf-tiedostoon.
2. Siivoa datasta pois ne rivit, joissa opinnäytetyön arvosanaa ei ole annettu ja piirrä sen jälkeen scatterplot sarakkeista *Opinnäytetyön tekemisaika työviikkoina (40 h) aihekuvauksen tekemisestä työn valmistumiseen: työviikkoa* (riippumaton muuttuja) ja *Opinnäytetyön arvosana* (riippuva muuttuja).   
     
   *Kuva, joka sisältää kohteen teksti, kuvakaappaus, numero, näyttö

   Tekoälyn generoima sisältö voi olla virheellistä.*

Vinkkejä:

* Datan siivoaminen tehdään komennolla df = df.dropna(subset=['Sarakkeen\_nimi']).
* Voit muuttaa x- ja y-akselin otsikot sekä kaavion otsikon funktioilla xlabel(), ylabel() ja title().