

Tehtävä 1: Veroprosentti

Veroprosentti määräytyy tulojen perusteella. Osoitteessa <https://www.veronmaksajat.fi/luvut/Laskelmat/Palkansaajan-veroprosentti/> on kuvattuna palkansaajan veroprosenttitaulukko.

Tässä tehtävässä sinun tulee toteuttaa ohjelma, joka kysyy ensin käyttäjältä tulotasoa, jonka jälkeen ohjelma tulostaa tulotasoon liittyvän veroprosentin sekä maksetut verot euroissa. Voit olettaa, että veroprosentti muuttuu aina tasatuhansissa -- esimerkiksi 10 999 euroa vuodessa tienaavalla veroprosentti on 7.7%. Ohjelman tarvitsee toimia vain 20 000 euron vuosituloihin asti. Jos käytät liukulukuja, älä välitä mahdollisista pyöristysvirheistä (kts. Esimerkki 3).

Esimerkki 1:

Tulotaso: 3500
Veroprosentti: 7.7%
Veroja maksetaan: 269.5 euroa

Esimerkki 2:

Tulotaso: 10500
Veroprosentti: 7.7%
Veroja maksetaan: 808.5 euroa

Esimerkki 3:

Tulotaso: 20000
Veroprosentti: 17.8%
Veroja maksetaan: 3560.0000000000005 euroa.

Vuositulot, euroa	Veroprosentti
Alle 11 000	7,7 %
11 000 -	8,3 %
13 000 -	8,6 %
14 000 -	10,2 %
15 000 -	11,8 %
16 000 -	13,2 %
17 000 -	14,4 %
18 000 -	15,5 %
19 000 -	16,6 %
20 000 -	17,8 %

Tehtävä 2: Vaalikonevastaukset

Yle julkaisi vuoden 2015 vaalikonevastaukset avoimena datana:

http://yle.fi/uutiset/yle_julkaisee_vaalikoneen_vastaukset_avoimena_datana/7869597

Tässä tehtävässä tehtävänäsi on tehdä sovellus, joka saa syötteen tiedostollisen vaalikonevastauksia. Ohjelma tulostaa erilaisia tilastoja tiedostossa olevista vastauksista. Vaalikonevastauksista on valittu valmiiksi osajoukko käyttöömme, tehtäväpohjassa on tiedostot:

- vaalidata-mooc.csv -- vaalikonevastaukset puolipisteillä (;) eroteltuna
- vaalidata-mooc-lyhyt.csv -- vaalikonevastaukset puolipisteillä (;) eroteltuna, ensimmäiset 100 riviä

Tiedoston ensimmäinen rivi sisältää kunkin sarakkeen otsaketiedot. Indeksissä 0 on vastaajan puolue, indeksissä 1 vastaajan ikä, ja esimerkiksi indeksissä 4 on hakijan perustelu siitä, miksi hänet pitäisi valita eduskuntaan. Seuraavat rivit sisältävät aina hakijakohtaiset vastaukset. Tiedostoa voi tarkastella esimerkiksi tekstieditorilla tai ohjelmallisesti.

Ohjelman tulee kysyä käsiteltävää tiedostoa, jonka jälkeen sen tulee tulostaa tiedoston perusteella seuraavat tiedot:

- Tilastoja lukumääristä
 - Vastaajien lukumäärä
 - Valittujen lukumäärä (sarakeessa "valittu" on arvo 1 jos henkilö valittiin)
 - Ei-valittujen lukumäärä (sarakeessa "valittu" on arvo 0 jos henkilö valittiin)
- Tilastoja kysymykseen "Miksi juuri sinut kannattaisi valita eduskuntaan?" annetuista vastauksista
 - Lyhimmän vastauksen sisältö sekä pituus (tyhjiä vastauksia ei tule huomioida).
 - Vastausten keskimääräinen pituus (tyhjiä vastauksia ei tule huomioida).
- Kymmenen yleisintä perusteluissa käytettyä sanaa sekä kunkin sanan esiintymislukumäärä. Muuta kaikkien sanojen kirjankoko pieniksi, ja poista myös pisteet, pilkut, huutomerkki sekä kysymysmerkit. Tulosta sanat siten, että yleisimmät sanat tulevat ensin. Älä huomioi yhden merkin mittaisia sanoja.
- Mielenpitoisuuden positiivisuus tai negatiivisuus. Sarakkeissa 11-36 on mielipidekysymyksiä, joissa vastaukset on annettu skaalalla täysin eri mieltä - täysin samaa mieltä. Kirjoita ohjelma, joka muuttaa tekstimuotoiset vastaukset numeromuotoon, ja laske mielenpitoisuuden keskiarvo.

Käytä seuraavia muunnoksia:

- täysin eri mieltä = 1
- jokseenkin eri mieltä = 2
- ohita kysymys = 3
- jokseenkin samaa mieltä = 4
- täysin samaa mieltä = 5

Ohjelman tulee tulostaa mielipiteiden numeerinen keskiarvo

- Kaikille ehdokkaille
- Valituille ehdokkaille
- Ei-valituille ehdokkaille
- Puoluekohtaisesti (puolueiden nimen mukaan järjestettynä)

Jos ehdokas ei ole vastannut johonkin kysymykseen, häntä ei tule ottaa huomioon.

- Kymmenen yleisintä "Miksi juuri sinut kannattaisi valita eduskuntaan?" -perustelussa esiintyvää sanaparia sekä niiden lukumäärät. Sanapari muodostuu peräkkäisistä perusteluissa olevista sanoista. Muuta kaikkien sanojen kirjankoko pieniksi, ja poista myös pisteet, pilkut, huutomerkki sekä kysymysmerkit. Tulosta sanaparit siten, että yleisimmät sanaparit tulevat ensin. Älä huomioi yhden merkin mittaisia sanoja.
- Edellisessä osassa rakennettuja sanapareja hyödyntävä ohjelma, joka tulostaa kahdenkymmenen sanan mittaisen vaalipuheen. Vaalipuheen tulee alkaa sanoilla "minulla on" ja sen tulee käyttää kutakin sanaa korkeintaan kerran.

Esimerkki ohjelman toiminnasta, kun käytössä on tiedosto vaalidata-mooc-lyhyt.csv.

Mikä tiedosto luetaan? [vaalidata-mooc-lyhyt.csv](#)

Tilastoja lukumääristä

vastaajia: 99

valittuja: 8

ei-valittuja: 91

Tilastoja perusteluista

lyhin perustelu: Olen arjen asiantuntija

lyhimmän perustelun pituus: 23

keskimääräisen perustelun pituus: 147.77011494252875

Yleisimmät sanat

ja (115)

olen (51)
on (44)
minulla (19)
kokemusta (13)
sekä (12)
asioita (11)
eduskuntaan (11)
kokemus (9)
haluan (8)

Mielipidekeskiarvot

kaikki ehdokkaat: 2.8765652951699465
valitut ehdokkaat: 2.8798076923076925
ei-valitut ehdokkaat: 2.876232741617357

Itsenäisyyspuolue: 2.7615384615384615
Kansallinen Kokoomus: 3.032967032967033
Perussuomalaiset: 2.9615384615384617
Piraattipuolue: 2.8846153846153846
Suomen Keskusta: 2.792307692307692
Suomen Kommunistinen Puolue: 2.769230769230769
Suomen Kristillisdemokraatit (KD): 2.8111888111888113
Suomen Sosialidemokraattinen Puolue: 2.735042735042735
Suomen ruotsalainen kansanpuolue: 2.9743589743589745
Vasemmistoliitto: 2.8197115384615383
Vihreä liitto: 3.08974358974359

Sanaparit

minulla on (17)
eduskunta tarvitsee (5)
eduskuntaan tarvitaan (5)
ja minulla (4)
on kokemusta (4)
20 vuoden (3)
asioihin ja (3)
ja olen (3)
jolla on (3)
kokemusta ja (3)

Vaalipuhe

minulla on kokemusta ja olen ahkera energinen luotettava asioihin hyvin laaja asiantuntemus yhdistettynä palavaan haluun muuttaa