

Havo 5, Wiskunde A

Computertoets beroepsbevolking

Inleiding

In deze toets maak je gebruik van een groot gegevensbestand van het CBS namelijk de enquête beroepsbevolking van 2011. Het bestand bevat 76746 records. In bijlage 1 vind je een beschrijving van de variabelen die zijn opgenomen in dit gegevensbestand. In deze toets nemen we aan dat dit bestand een aselechte steekproef is uit de Nederlandse bevolking tussen 15 en 65 jaar.

In sommige onderdelen van de toets wordt gevraagd een uitspraak te doen over het verschil tussen twee groepen. Je kunt daarbij gebruik maken van bijlage 2 met het formuleblad dat ook in het Centraal Schriftelijk Eindexamen wordt opgenomen.

Tijdens de toets bewaar je je uitwerkingen in een word-document. Vermeld bovenaan in dit document je eigen naam. Aan het einde van de toets lever je dit word-document in via Magister, opdrachten.

De toets duurt 60 minuten en bestaat uit acht onderdelen. Bij elk onderdeel staat hoeveel punten je maximaal kunt behalen. In totaal kun je voor deze toets maximaal 36 punten behalen.

VUStat starten en gegevenbestand inlezen

Ga naar alle programma's, wiskunde, VUStat2015, hoofdmenu, data analyse en dan openen bestand.

Het bestand staat op het bureaublad in de map VUStat-bestanden.

Open het bestand beroepsbevolking_2

Controleer het aantal records en de variabelen

Controleer linksonder of je een bestand hebt met 76746 records.

Maak een overzicht van de variabelen in het bestand via data, variabelen.

Controleer of het bestand de twaalf variabelen bevat zoals die in de bijlage staan vermeld.

Op de volgende bladzijde begint de toets.

Veel succes!

BEGIN VAN DE TOETS

In de eerste twee onderdelen kijk je naar de leeftijd-geslachtssamenstelling van de werkzame beroepsbevolking. Deze wordt door verschillende organisaties scherp in de gaten gehouden, bijvoorbeeld om de uitstroom uit de werkzame beroepsbevolking door pensionering voor de komende jaren te voorspellen.

2p 1 Hoeveel procent van de werkzame beroepsbevolking valt in de leeftijdsgroep 55 t/m 64 jaar?

Een andere reden om de leeftijd-geslachtssamenstelling van de werkzame beroepsbevolking in de gaten te houden is vanuit het oogpunt van gelijkheid tussen mannen en vrouwen. Vanuit dat oogpunt wordt bijvoorbeeld in bepaalde leeftijdsgroepen gekeken naar welk deel van de werkzame beroepsbevolking vrouw is.

3p 2 Hoeveel procent van de werkzame beroepsbevolking in de leeftijdsgroep 15 t/m 24 jaar is vrouw?

Het Centraal Bureau voor de Statistiek publiceerde nog niet zo lang geleden een persbericht met de titel: 3 miljoen Nederlanders willen of kunnen niet werken. Hoogstwaarschijnlijk is deze uitspraak gebaseerd op een betrouwbaarheidsinterval voor de proportie die niet wil of kan werken.

5p 3 Bereken het 95%-betrouwbaarheidsinterval voor de proportie van de bevolking die niet wil of wel wil maar niet kan werken

In de tweede helft van de twintigste eeuw is de gemiddelde arbeidsduur per week flink afgenomen. Vroeger was een werkweek van 50 uur of meer heel gewoon, tegenwoordig is een werkweek vaak 36 uur. De gemiddelde arbeidsduur in ons land is laag als die wordt vergeleken met andere Westerse landen.

5p 4 Bereken het 95%-betrouwbaarheidsinterval voor de gemiddelde arbeidsduur van mensen in de werkzame beroepsbevolking

De gemiddelde arbeidsduur is niet in alle beroepsrichtingen gelijk. In sommige beroepen is parttime werken gebruikelijker dan in andere.

5p 5 Onderzoek hoe groot het verschil is in gemiddelde arbeidsduur voor mensen die werkzaam zijn in de beroepsrichting technisch en mensen die werkzaam zijn in de beroepsrichting medisch/paramedisch

Sommige mensen zouden wel meer uren per week willen werken, anderen willen zouden juist graag minder uren per week willen werken. Omdat daarbij allerlei overwegingen een rol kunnen spelen is het op voorhand niet duidelijk of er een verschil bestaat in de leeftijdssamenstelling van beide groepen.

5p 6 Onderzoek hoe groot het verschil is in leeftijdssamenstelling van de mensen die meer willen werken ten opzichte van mensen die minder willen werken

De overheid heeft een lange tijd een campagne gevoerd om met name meisjes te stimuleren om een technische opleiding te volgen.

5p 7 Onderzoek hoe groot het verschil is tussen mannen en vrouwen in de proportie die als onderwijsrichting techniek heeft

Iemand beweert dat er op de Nederlandse arbeidsmarkt een scheidslijn bestaat tussen aan de ene kant hoog opgeleide autochtone mannen van middelbare leeftijd (35 tot en met 44 jaar) en aan de andere kant laag opgeleide allochtone vrouwen van jonge leeftijd (15 tot en met 24 jaar).

6p 8 Onderzoek of het gegevensbestand aanleiding geeft om het bestaan van een dergelijke scheidslijn te bevestigen of juist niet.

EINDE VAN DE TOETS

Inleveren word-document

Lever je word-document in via Magister, opdrachten.

Sluit pas af als je je word-document hebt ingeleverd.

Bijlage 1. Gegevensbestand enquête beroepsbevolking

1. Geslacht

1=man; 2=vrouw

2. Leeftijdsgroep

Leeftijd wordt berekend n.a.v. geboortedatum en vervolgens ingedeeld in tienjaarsklassen.

1=15-24 jaar; 2=25-34 jaar; 3=35-44 jaar; 4=45-54 jaar; 5=55-64 jaar

3. Herkomstgroepering

De herkomstgroepering wordt bepaald n.a.v. het geboorteland van beide ouders. Als beide ouders in Nederland zijn geboren is de respondent autochtoon. Als een van beide ouders in het buitenland is geboren is de respondent allochtoon. Wanneer beide ouders in een ander land zijn geboren, kijkt het CBS naar het geboorteland van de moeder om de herkomstgroepering te bepalen. Voorbeeld: moeder is geboren in Turkije, vader in Duitsland, dan heeft de respondent Turkije als herkomstland.

10=autochtonen; 20=westerse allochtonen; 31=Turken en Marokkanen; 33=Antillianen/Arubanen en Surinamers'; 35=overig niet-westerse landen; 99=onbekend

4. Beroepsbevolking

Heeft de respondent voor 12 uur per week of meer werk? Dan valt hij onder de werkzame beroepsbevolking.

Wil of kan de respondent niet werken voor 12 uur of meer per week (bijvoorbeeld scholieren en huisvrouwen/-mannen)? Dan valt hij onder de niet-beroepsbevolking.

Wil de respondent voor 12 uur of meer werken en doet hij dat nu niet? En kan hij op korte termijn beginnen? En zoekt hij actief naar werk? Dan valt hij onder de werkloze beroepsbevolking. Anders valt hij onder de niet-beroepsbevolking.

1=werkzame beroepsbevolking; 2=werkloze beroepsbevolking; 3=niet-beroepsbevolking; 7=nvt

5. Arbeidsduur per week

Het antwoord op de vraag: "Hoeveel uur werkt u in totaal gemiddeld per week, overuren en onbetaalde uren niet meegerekend?"

1=minder dan 12 uur; 2=12 tot 20 uur; 3=20 tot 35 uur; 4=35 uur of meer

6. Meer of minder willen werken

Hier wordt aan de respondenten die tot de beroepsbevolking behoren, gevraagd of ze meer of minder willen werken dan hun arbeidsduur per week.

1=meer willen werken; 2=minder willen werken; 3=niet meer/minder willen werken; 7=vraag niet gesteld

7. Bereidheid tot werken

Hier wordt aan de respondenten die tot de niet-beroepsbevolking of werkloze beroepsbevolking behoren, gevraagd of ze wel of niet 12 uur of meer per week zouden willen en kunnen werken. Het minimum van 12 uur wordt voor de Nederlandse definitie van werkloosheid aangehouden als indicatie dat de respondent een substantieel aantal uren per week werkt. Vanaf anderhalve dag per week geeft namelijk de meerderheid van de mensen aan dat betaald werken de belangrijkste bezigheid is.

1=persoon wil niet werken / wel werken, maar kan niet werken;

2=persoon wil minder dan 12 uur per week werken / wil wel 12 uur of meer per week werken maar kan niet

3=persoon wil 12 uur of meer per week werken / persoon heeft werk gevonden van 12 uur of meer per week

7=nvt

8. Beroepsrichting

Hier is voor respondenten die tot de werkzame beroepsbevolking behoren afgeleid in welke categorie zijn beroep of functie het beste past.

00=Geen werkkring; 02=Docenten en staffuncties onderwijs, onderwijskundig;

04=Agrarisch / Exact; 06=Technisch; 08=Transport, communicatie en verkeer;

08=Medisch en paramedisch; 10=Economisch, administratief en commercieel

13=Juridisch, bestuurlijk en openbare orde en veiligheid / Taal en cultuur;

16=Gedrag en maatschappij; 17=Persoonlijke en sociale verzorging; 18=Management / Algemeen; 99=Beroepsrichting onbekend

9. Onderwijsniveau

In deze variabele is het hoogst behaalde onderwijsniveau van de respondent afgeleid:

1=Laag: basisonderwijs, LBO, VBO, VSO, VMBO, MAVO, ULO, MULO;

2=Midden: HAVO, MMS, VWO, HBS, MBO;

3=Hoog: HBO, WO.

10. Publicatie indeling onderwijsrichting

In deze variabele is de richting van het hoogst behaalde onderwijsniveau van de respondent afgeleid.

1=Algemeen; 2=Leraren; 3=Humaniora, soc. wetenschap, communicatie en kunst; 4=Economie, commercieel, management en administratie; 5=Juridisch, bestuurlijk, openb. orde en veiligheid; 6=Wiskunde, natuurwetenschap en informatica; 7=Techniek; 8=Agrarisch en milieu; 9=Gez. zorg, sociale dienstverl. en verzorging; 10=Horeca, toerisme, vrijetijdsbest. transp. en logistiek; 11=Onbekend

11. Leeftijd

Een hercodering van leeftijdsgroep zodat:

Als de leeftijdsgroep 1 is dan is de leeftijd gelijk aan 20;

Als de leeftijdsgroep 2 is dan is de leeftijd gelijk aan 30;

Als de leeftijdsgroep 3 is dan is de leeftijd gelijk aan 40;

Als de leeftijdsgroep 4 is dan is de leeftijd gelijk aan 50;

Als de leeftijdsgroep 5 is dan is de leeftijd gelijk aan 60;

Je mag leeftijd gebruiken alsof dit een kwantitatieve variabele is.

12. Arbeidsduur2

Een hercodering van arbeidsduur zodat:

Als de arbeidsduur minder dan 12 uur is dan is arbeidsduur2 gelijk aan 6;

Als de arbeidsduur tussen 12 en 20 uur is dan is arbeidsduur2 gelijk aan 16;

Als de arbeidsduur tussen 20 en 35 uur is dan is arbeidsduur2 gelijk aan 28;

Als de arbeidsduur 35 uur of meer is dan is arbeidsduur2 gelijk aan 38.

Je mag arbeidsduur2 gebruiken alsof dit een kwantitatieve variabele is.

Bijlage 2. Formuleblad bij Centraal Schriftelijk Eindexamen

Betrouwbaarheidsintervallen

Het 95% -betrouwbaarheidsinterval voor de populatieproportie is $p \pm 2 \cdot \sqrt{\frac{p(1-p)}{n}}$
met p de steekproefproportie en n de steekproefomvang.

Het 95%-betrouwbaarheidsinterval voor het populatiegemiddelde is $\bar{X} \pm 2 \cdot \frac{S}{\sqrt{n}}$
met \bar{X} het steekproefgemiddelde, n de steekproefomvang en S de steekproefstandaardafwijking.

Vuistregels bij de grootte van het verschil van twee groepen

2x2 kruistabel $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ met $phi = \frac{ad - bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$

- $phi < -0,4$ of $phi > 0,4$: groot verschil.
- $-0,4 < phi < -0,2$ of $0,2 < phi < 0,4$: middelmatig verschil.
- $-0,2 < phi < 0,2$: gering verschil.

Maximale verschil in cumulatief percentage ($\max V_{cp}$) (met steekproefomvang $n > 100$)

- $\max V_{cp} > 40$: groot verschil.
- $20 < \max V_{cp} \leq 40$: middelmatig verschil.
- $\max V_{cp} \leq 20$: gering verschil.

Effectgrootte $E = \frac{\bar{X}_1 - \bar{X}_2}{\frac{1}{2}(S_1 + S_2)}$, met \bar{X}_1 en \bar{X}_2 de steekproefgemiddelden ($\bar{X}_1 \geq \bar{X}_2$)

en S_1 en S_2 de steekproefstandaardafwijkingen

- $E > 0,8$: groot verschil.
- $0,4 < E \leq 0,8$: middelmatig verschil.
- $E \leq 0,4$: gering verschil.

Twee boxplots vergelijken

- De boxen¹ overlappen elkaar niet: groot verschil.
- De boxen overlappen elkaar wel en de mediaan van de ene boxplot ligt buiten de box van de andere boxplot: middelmatig verschil.
- Alle andere gevallen: gering verschil.

¹ De 'box' is het interval vanaf het eerste kwartiel tot en met het derde kwartiel.