

Statistiek en Fototaarten bij HEMA



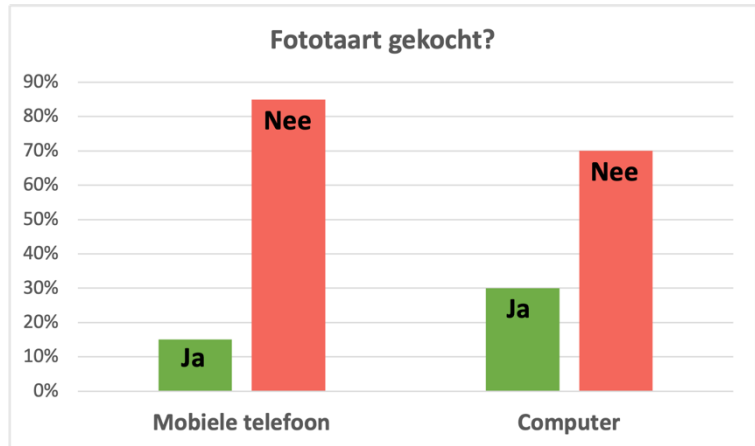
Statistiek en fototaarten bij HEMA

Deze opgaven horen bij de video *Statistiek en fototaarten bij HEMA* van de videoserie *Succesformules in beeld* van Platform Wiskunde Nederland.

Opgave 1

In de video was het staafdiagram hiernaast te zien.

Hieruit werd duidelijk dat HEMA-klanten die via een computer winkelden vaker tot een aankoop overgaan dan via een mobiele telefoon. Deze gegevens kun je ook in een **kruistabel** zetten, zoals hieronder.



	Fototaart gekocht	Geen fototaart gekocht	Totaal
Mobiele telefoon	15%	85%	100%
Computer	30%	70%	100%

- Is dit een kruistabel met rij-percentages of kolom-percentages? Licht toe.
- Waarom noemen we deze getallen **relatieve** aantallen?
- Betekent deze tabel dat er in **absolute** aantallen minder fototaarten worden gekocht via de mobiele telefoon dan via de computer? Licht toe.

In het jaar 2021 bezochten 80232 klanten via een mobiele telefoon de pagina van de fototaart, en 26744 klanten via een computer.

- Vul met bovenstaande gegevens de onderstaande kruistabel met **absolute aantallen**. Rond af op gehelen.

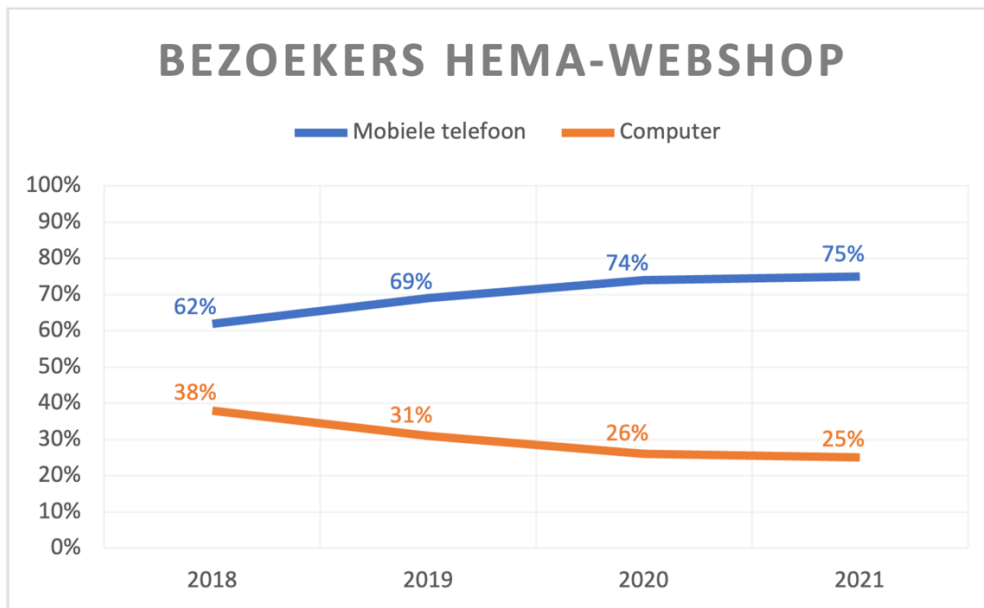
	Fototaart gekocht	Geen fototaart gekocht	Totaal
Mobiele telefoon			
Computer			
Totaal			

Uit de staafdiagram en kruistabel blijkt duidelijk een verschil tussen klanten die via hun mobiele telefoon en computer winkelen in hoe vaak zij tot de aankoop van een fototaart overgaan.

- Bepaal met behulp van het formuleblad in de bijlage of dit verschil groot, middelmatig of gering is.

Opgave 2

In de video was ook onderstaande lijndiagram te zien.



Bepaal met behulp van de gegevens uit 2018 en 2021 wat het percentage bezoekers via de mobiele telefoon in 2023 is. Ga hierbij uit van een lineair verband. Rond af op gehele procenten.

Bijlage

FORMULEBLAD

Vuistregels voor de grootte van het verschil van twee groepen

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

- als $\phi < -0,4$ of $\phi > 0,4$, dan zeggen we “het verschil is groot”,
- als $-0,4 \leq \phi < -0,2$ of $0,2 < \phi \leq 0,4$, dan zeggen we “het verschil is middelmatig”,
- als $-0,2 \leq \phi \leq 0,2$, dan zeggen we “het verschil is gering”.

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

- als $\max V_{cp} > 40$, dan zeggen we “het verschil is groot”,
- als $20 < \max V_{cp} \leq 40$, dan zeggen we “het verschil is middelmatig”,
- als $\max V_{cp} \leq 20$, dan zeggen we “het verschil is gering”.

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

($\bar{X}_1 \geq \bar{X}_2$), S_1 en S_2 de steekproefstandaardafwijkingen

- als $E > 0,8$, dan zeggen we “het verschil is groot”,
- als $0,4 < E \leq 0,8$, dan zeggen we “het verschil is middelmatig”,
- als $E \leq 0,4$, dan zeggen we “het verschil is gering”.

Twee boxplots vergelijken

- als de boxen¹⁾ elkaar niet overlappen, dan zeggen we “het verschil is groot”,
- als de boxen elkaar wel overlappen en een mediaan van een boxplot buiten de box van de andere boxplot ligt, dan zeggen we “het verschil is middelmatig”,
- in alle andere gevallen zeggen we “het verschil is gering”.

noot 1 De ‘box’ is het interval vanaf het eerste kwartiel tot en met het derde kwartiel.

Uitwerkingen

Opgave 1

- De rijen tellen op tot 100%, daarom is dit een kruistabel met rij-percentages.
- Omdat het percentages zijn. Relatief betekent ook wel 'ten opzichte van'. Deze percentages geven aan hoeveel klanten ten opzichte van het totaal aantal klanten tot aankoop overgaat.
- Nee, want het kan zijn dat er veel meer mensen via een mobiele telefoon de HEMA-webshop bezoeken, waardoor het absolute aantal gekochte fototaarten alsnog hoger uit komt dan bij een computer (ondanks dat een lager percentage klanten tot aankoop overgaat).

d.

	Fototaart gekocht	Geen fototaart gekocht	Totaal
Mobiele telefoon	12035	68197	80232
Computer	8023	18721	26744
Totaal	20058	86917	106975

- Hiervoor gebruiken we de phi-coëfficiënt.

$$phi = \frac{8023 \cdot 68197 - 12035 \cdot 18721}{\sqrt{20058 \cdot 86917 \cdot 80232 \cdot 26744}} \approx 0,166$$

Er geldt $-0,2 \leq phi \leq 0,2$ dus het verschil is gering.

Opgave 2

In 2018 was het percentage 62% en in 2021 was het percentage 75%.

Het percentage is dus 13% gestegen in 3 jaar tijd.

Uitgaande van een lineair verband is dat $\frac{13}{3} \approx 4,33\%$ per jaar.

Het jaar 2023 is twee jaar verder dan 2021.

$$75 + 2 \cdot 4,33 \approx 84\%$$

Dus in 2023 is het percentage bezoekers via de mobiele telefoon ongeveer 84%.