

# Aanjaagprogramma Innovatie toerisme



## AANJAAGPROGRAMMA

*Lesmateriaal digitalisering 2023*

## Introductie

In het kader van het Aanjaagprogramma Innovatie is een fictieve dataset ontwikkeld waar basis data-verkenning mee uitgevoerd kan worden. De dataset bestaat uit energiegegevens voor een toeristische accommodatie. Let er op dat dit *fictieve* data is die opgebouwd is op basis van een aantal regels en aannames. Op deze manier zijn er mogelijke patronen in de dataset aangebracht. De dataset en de patronen hierin hoeven dus niet representatief te zijn voor de realiteit.

## Dataset

*data\_per\_uur.xlsx*

Deze dataset beslaat een tijdperiode van 1 april 2022 middernacht tot en met 5 oktober 2022 23.00h en is opgedeeld in een rij per uur (4512 rijen, exclusief de bovenste rij met de kolomnamen). De dataset bevat onder andere informatie over elektriciteitsverbruik (in kwh) en waterverbruik (in L).

De betekenis van de kolommen is als volgt:

<b>tijdstip</b>	Volledig tijdstip van de rij
<b>moment</b>	Moment op de dag: ochtend, middag, avond of nacht
<b>weekdag</b>	Dag van de week
<b>datum</b>	Datum
<b>maand</b>	Maand
<b>dag</b>	Dag
<b>uur</b>	Uur van de dag
<b>temperatuur</b>	Temperatuur per rij
<b>zon_uren</b>	Hoeveelheid zon
<b>neerslag_duur_h</b>	Duur van de neerslagperiode(s)
<b>neerslag_som</b>	Hoeveelheid neerslag in mm
<b>elektriciteitsverbruik</b>	Elektriciteitsverbruik in kwh
<b>waterverbruik</b>	Waterverbruik in L
<b>gem_temp</b>	Gemiddelde dagtemperatuur (van de gehele 24 uur)
<b>max_temp</b>	Maximale dagtemperatuur (tijdens gehele 24 uur)

*augustusw12.xlsx*

Een kopie van de *data\_per\_uur.xlsx* dataset met alleen de eerste twee weken van augustus van 2022. Deze dataset heeft dezelfde kolommen als de vorige dataset, maar de elektriciteitskosten per uur zijn er aan toegevoegd. Dit zijn de kosten per kilowattuur.

## Opdrachten

Het kan zijn dat het voor de opdracht handig is om nog een losse berekening te maken van de gebruikte dataset, voordat er een datavisualisatie gemaakt wordt. Het advies is om een extra tabblad 'Berekeningen' aan te maken voor deze extra berekeningen en een tabblad 'Grafieken' voor de uiteindelijke datavisualisaties.

### Opdracht 1

#### Instructie

Creëer inzicht in het dagpatroon van waterverbruik door een datavisualisatie te maken van het waterverbruik per uur voor de gehele dataset (van april tot en met oktober). Maak hierin gebruik van een lijngrafiek met de uren (0-23) op de horizontale X-as en het verbruik in totalen (som) op de Y-as.

*Tip:* Maak voor deze uitwerking gebruik van het extra tabblad 'Berekeningen' om het totaal waterverbruik per uur te berekenen voordat er een datavisualisatie gemaakt wordt.

#### Uitwerking

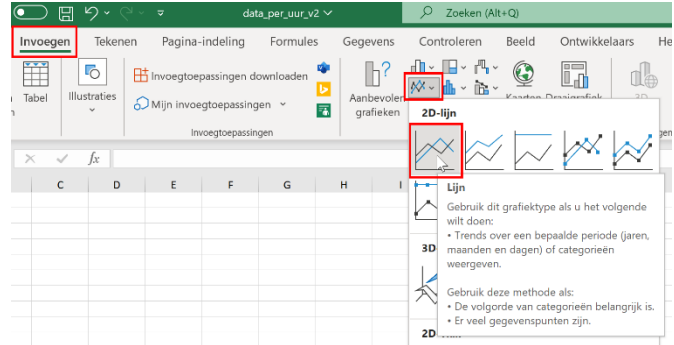
1. Open de dataset *data\_per\_uur*
2. Maak in het extra tabblad 'Berekeningen' een nieuwe kolom door in cel A1 'uur' te typen en nummer onder de naam verticaal van 0 tot en met 23 (van cel A2 tot en met cel A25).
3. Maak een tweede nieuwe kolom door in cel B1 'waterverbruik' te typen. Tel alle waarden van het waterverbruik per uur bij elkaar op. Selecteer cel B2 en typ de volgende functie:

(Nederlands) =SOM.ALS('data\_per uur'!G:G;Berekeningen!A2;'data\_per uur'!M:M)

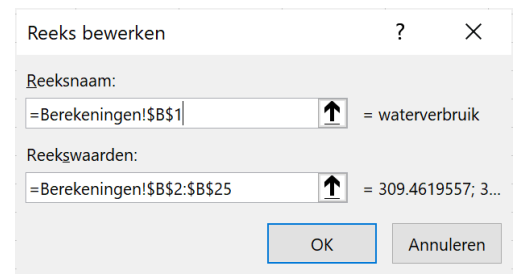
De waarde in B2 kan doorgetrokken worden naar alle andere cellen door de cel te selecteren en met de linkermuisknop op de celhoek van B2 rechtsonderin te klikken (de cursor verandert in een +) en met de muisknop ingedrukt naar onderen te slepen, totdat de cel rechts van '23' (cel A25) gevuld is.

	A	B
1	uur	waterverbruik
2	0	309.4619557
3	1	314.0730824
4	2	317.7379729
5	3	322.7365425
6	4	325.8846834
7	5	304.3096505
8	6	321.9516428
9	7	301.8672241
10	8	421.2723313
11	9	423.9932335
12	10	290.3343399
13	11	289.1456476
14	12	283.109806
15	13	286.5508211
16	14	283.4013252
17	15	284.1833503
18	16	282.4750994
19	17	426.6544735
20	18	428.0953497
21	19	315.813309
22	20	324.8223905
23	21	423.9884195
24	22	423.5740442
25	23	342.3987983

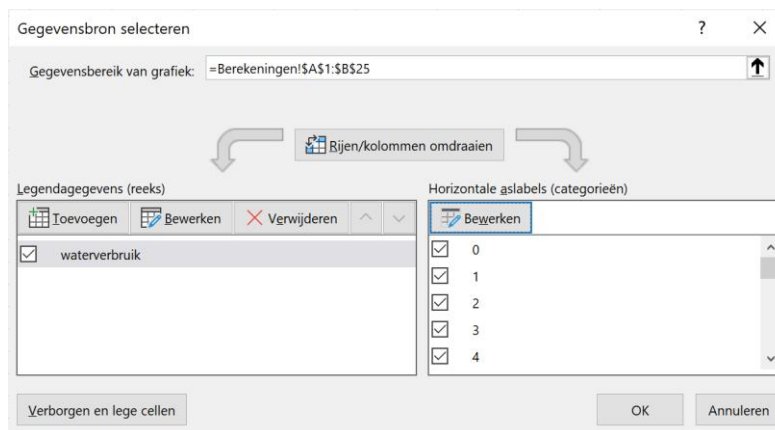
4. Selecteer het tabblad 'Grafieken' en selecteer onder 'Invoegen de lijngrafieken en vervolgens de lijngrafiek.
5. Rechtsklik op het witte vierkant dat zojuist verschenen is en klik op 'Gegevens Selecteren'.
6. Voeg de data toe door op 'Toevoegen' te klikken bij de Legendagegevens (reeks).



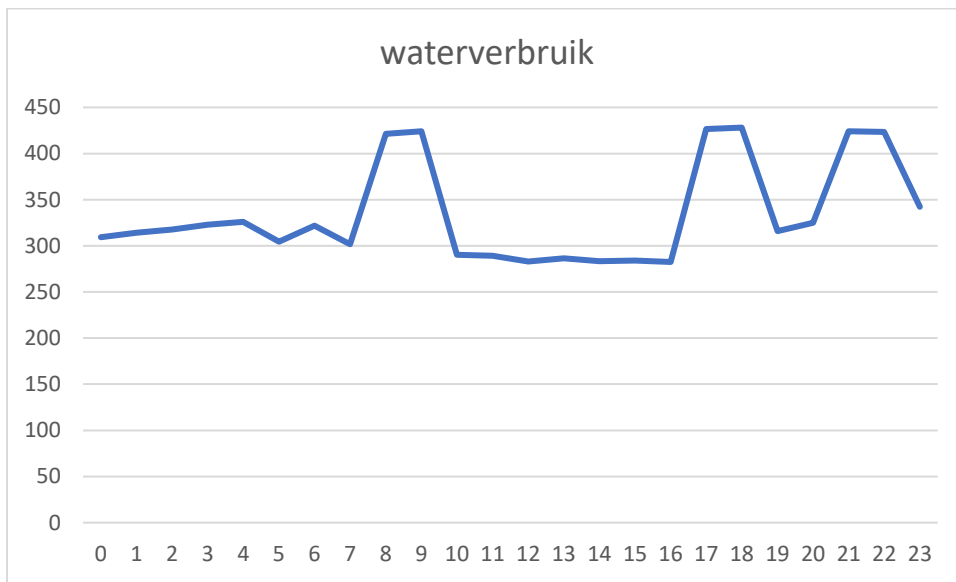
7. Klik op het pijltje naar boven ↑ naast het eerste tekstvak ("Reeksnaam"), klik op het tabblad 'Berekeningen' en selecteer cel B2 met het woord 'waterverbruik'. Klik hierna op het pijltje naar beneden ↓
8. Klik vervolgens op het pijltje naar boven ↑ naast het tweede tekstvak ("Reekswaarden") voor het toevoegen van de seriewaardes. Selecteer in het tabblad 'Berekeningen' alle waardes onder 'waterverbruik' (cel B2 t/m cel B25) en klik hierna op het pijltje naar onderen. Druk op 'OK'.



9. De horizontale aslabels kloppen nog niet. Klik vervolgens op 'Bewerken' onder 'Horizontale aslabels (categorieën)'. Klik op het pijltje naar boven ↑ naast het tekstveld 'Aslabelbereik' en selecteer in het tabblad de waarden 0 t/m 23 in de kolom 'uur'. Klik vervolgens op het pijltje naar onderen ↓ en klik op 'OK' om het venster te sluiten.



De grafiek is nu zichtbaar. Hierin is duidelijk dat er een aantal pieken zijn in het waterverbruik gedurende de dag, namelijk 's ochtends, rond het avondeten en iets later op de avond.



## Opdracht 2

### Instructie

Creëer inzicht in het maandpatroon van het door een datavisualisatie te maken van het elektriciteitsverbruik per maand voor de gehele dataset (van april tot en met oktober). Maak hierin gebruik van een staafdiagram met de maanden (april-oktober) op de horizontale X-as en het verbruik in totalen (som) op de Y-as.

*Tip:* Maak voor deze uitwerking gebruik van het extra tabblad 'Berekeningen' om het totaal elektriciteitsverbruik per maand te berekenen voordat er een datavisualisatie gemaakt wordt.

### Uitwerking

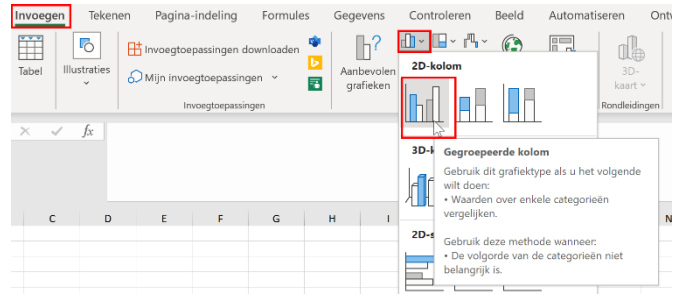
1. Open de dataset *data\_per\_uur*
2. Maak in het extra tabblad 'Berekeningen' een nieuwe kolom door in cel A1 'maand' te typen en nummer onder de naam verticaal de maandnummers 4 t/m 10 (voor april t/m oktober)
3. Maak een tweede nieuwe kolom door in cel B1 'elektriciteitsverbruik' te typen. Tel alle waarden van het elektriciteitsverbruik per maand bij elkaar op. Selecteer cel B2 en typ de volgende functie:

=SOM.ALS(data\_per\_uur!E:E;Berekeningen!A2;data\_per\_uur!L:L)

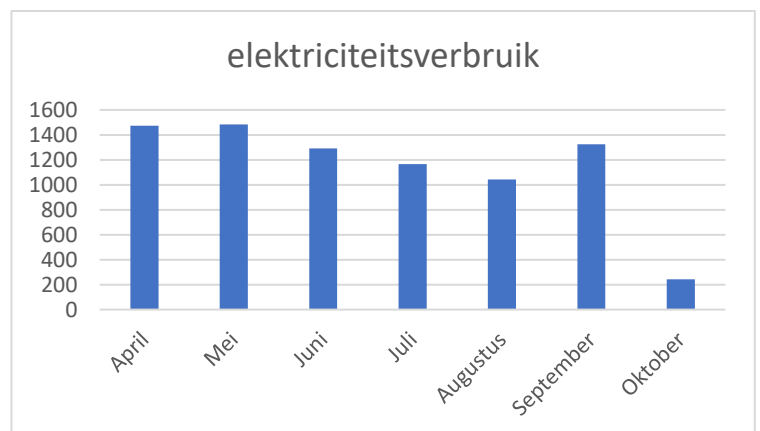
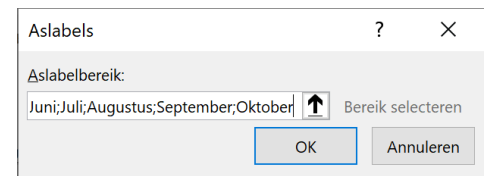
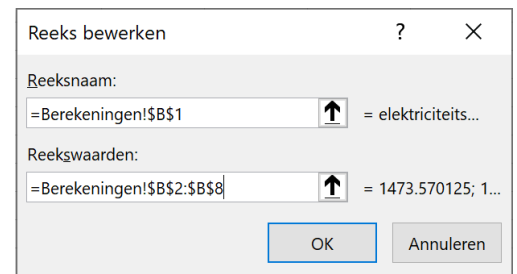
	A	B	C
1	maand	elektriciteitsverbruik	
2	4	1473.57	
3	5	1483.43	
4	6	1291.685	
5	7	1166.18	
6	8	1043.072	
7	9	1323.919	
8	10	243.7066	
9			

De waarde in B2 kan doorgetrokken worden naar alle andere cellen door de cel te selecteren en met de linkermuisknop op de celhoek van B2 rechtsonderin te klikken (de cursor verandert in een +) en met de muisknop ingedrukt naar onderen te slepen, totdat de cel rechts van '10' (cel A8) gevuld is.

- Selecteer het tabblad 'Grafieken' en selecteer onder 'Invoegen' de Kolom- en Staafdiagrammen en vervolgens de gegroepeerde kolom.
- Rechtsklik op het witte vierkant dat zojuist verschenen is en klik op 'Gegevens Selecteren'.
- Voeg de data toe door op 'Toevoegen' te klikken bij de Legendagegevens (reeks).



- Klik op het pijltje naar boven ↑ naast het eerste tekstvak ("Reeksnaam"), klik op het tabblad 'Berekeningen' en selecteer cel B2 met het woord 'elektriciteitsverbruik'. Klik hierna op het pijltje naar beneden ↓
- Klik vervolgens op het pijltje naar boven ↑ naast het tweede tekstvak ("Reekswaarden") voor het toevoegen van de seriewaardes. Selecteer in het tabblad 'Berekeningen' alle waardes onder 'elektriciteitsverbruik' tober (cel B2 t/m cel B8) en klik hierna op het pijltje naar onderen. Druk op 'OK'.
- De horizontale aslabels kloppen nog niet. Klik vervolgens op 'Bewerken' onder 'Horizontale aslabels (categorieën)'. Klik op het pijltje naar boven ↑ naast het tekstveld 'Aslabelbereik' en typ in het tekstvak de maanden April tot en met Oktober, gescheiden door puntkomma's (;) en klik op 'OK' om het venster te sluiten. En druk nogmaals op 'OK' om de grafiek af te maken.



## Opdracht 3

### Instructie

Creëer inzicht in de relatie tussen de gemiddelde dagtemperatuur en het elektriciteitsverbruik. Maak hierin gebruik van een scatterplot met de gemiddelde dagtemperatuur op de X-as en het elektriciteitsverbruik op de Y-as.

*Tip:* Maak voor deze uitwerking gebruik van het extra tabblad 'Berekeningen' om de unieke gemiddelde temperaturen te bepalen en de gemiddelde dagtemperaturen te berekenen.

### Uitwerking

1. Open de dataset *data\_per\_uur*
2. Maak in het extra tabblad 'Berekeningen' een nieuwe kolom met de unieke daggemiddelden door in cel A1 de formule '=UNIEK(' te typen en in het tabblad met de data de kolom met de gemiddelde temperaturen te selecteren (kolom N). Sluit het haakje met ')' en druk op Enter. In kolom A staat nu een kolom met alle unieke daggemiddelden.
3. Maak een tweede nieuwe kolom door in cel B1 'electriciteitsverbruik' te typen. Pak het gemiddelde van het elektriciteitsbereik per unieke gemiddelde dagtemperaturen. Selecteer cel B2 en typ de volgende functie:

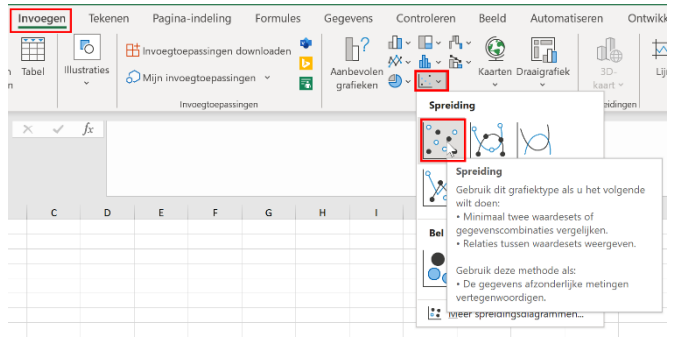
=GEMIDDELDE.ALS(data\_per\_uur\_v2!N:N;Berekeningen!A2;data\_per\_uur\_v2!L:L)

	A	B	C
1	gem_temp	electriciteitsverbruik	
2	3.183333	1.946935	
3	4.475	2.037377	
4	4.933333	2.056681	
5	7.4375	2.065134	
6	9.679167	2.127026	
7	9.633333	2.058824	
8	8.979167	2.071391	
9	6.779167	2.002331	
10	7.120833	2.029827	
11	8.558333	1.992748	
12	10.75417	2.011407	
13	12.34583	2.031994	
14	12	2.016865	
15	11.60417	2.136486	

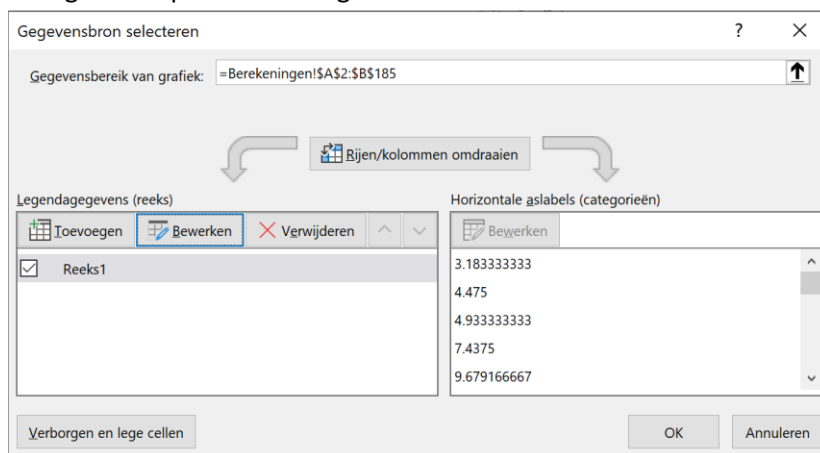
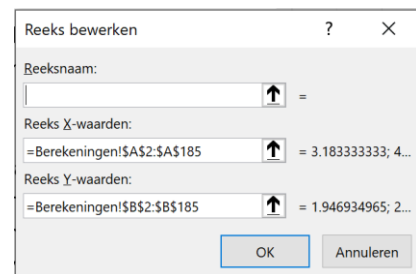
De waarde in B2 kan doorgetrokken worden naar alle andere cellen door de cel te selecteren en met de linkermuisknop op de celhoek van B2 rechtsonderin te klikken (de cursor verandert in een +) en met de muisknop ingedrukt naar onderen te slepen, totdat de laatste cel gevuld is.



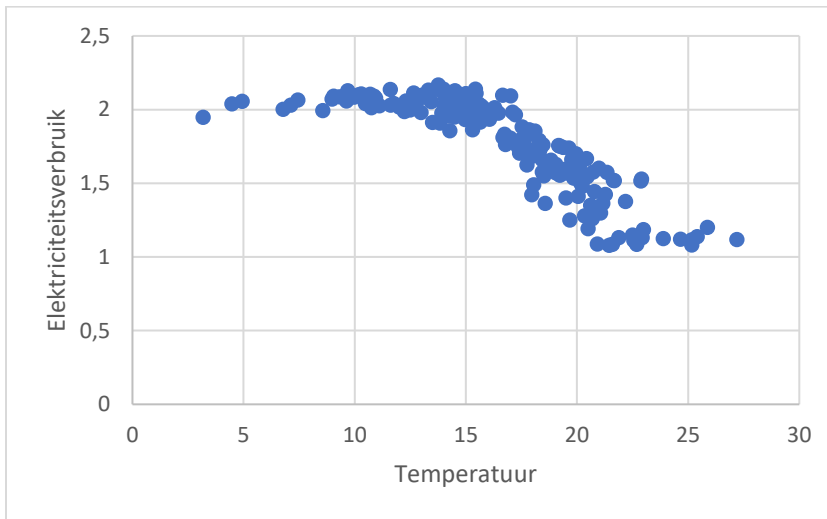
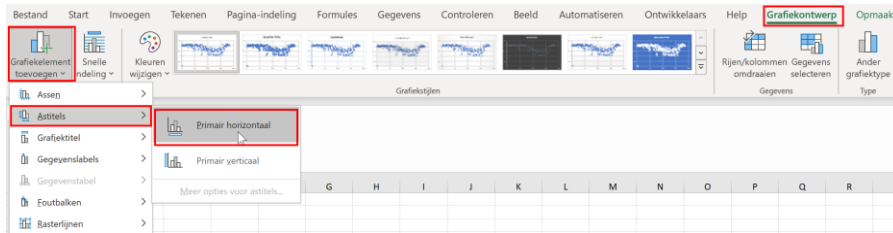
4. Selecteer het tabblad 'Grafieken' en selecteer onder 'Invoegen' de Spreidings- en Bellendiagrammen en vervolgens de Spreiding visualisatie.
5. Rechtsklik op het witte vierkant dat zojuist verschenen is en klik op 'Gegevens Selecteren'.
6. Voeg de data toe door op 'Toevoegen' te klikken bij de Legendagegevens (reeks).



7. Klik op het pijltje naar boven ↑ naast het tekstvak "Reeks X-waarden", klik op het tabblad 'Berekeningen' en selecteer alle valide data in kolom A (rij 2 t/m rij 185). Klik hierna op het pijltje naar beneden ↓. Klik op het pijltje naar boven ↑ naast het tekstvak "Reeks Y-waarden", klik op het tabblad 'Berekeningen' en selecteer alle valide data in kolom B (rij 2 t/m rij 185). . Klik hierna op het pijltje naar beneden ↓. Druk op 'OK'. Druk nogmaals op 'OK' om de grafiek af te maken.



8. We kunnen de grafiek duidelijker maken door astitels toe te voegen. Dit kan gedaan worden door onder 'Grafiekontwerp' te klikken op 'Grafiekelement toevoegen' en zowel een horizontale als een verticale as toe te voegen. De titels kunnen aangepast worden in de grafiek zelf.



## Opdracht 4

### Instructie

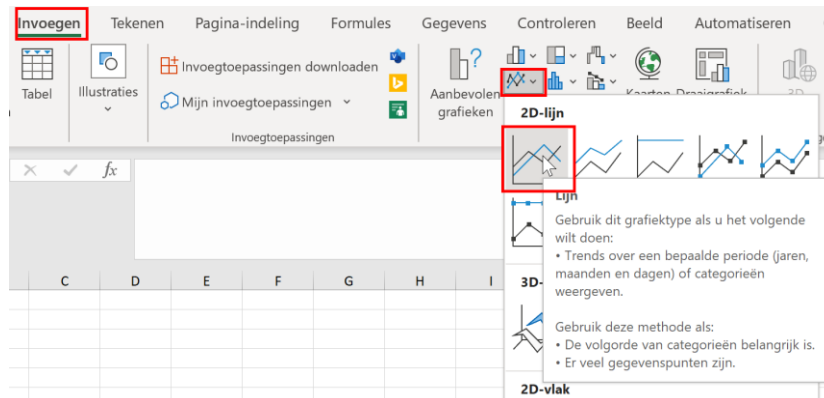
Er is ook een dataset voor de eerste twee weken van augustus van 2022, de dataset heet 'augustusw12'. Deze dataset heeft dezelfde kolommen als de vorige dataset, maar de energiekosten per uur zijn er aan toegevoegd. Dit zijn de kosten per kilowattuur.

Creëer inzicht in de relatie tussen de het elektriciteitsverbruik en de elektriciteitskosten per uur. Maar hierin gebruik van een lijngrafiek met twee lijnen: één voor het elektriciteitsverbruik en één voor de elektriciteitskosten. De tijd in uren staat op de X-as en het elektriciteitsverbruik en de elektriciteitskosten komen op de verticale Y-assen.

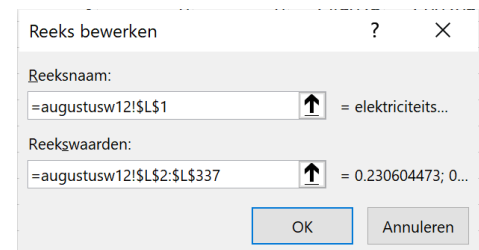
*Tip:* Voor deze opdracht is het niet nodig om gebruik te maken van een extra 'Berekeningen' tabblad.

### Uitwerking

1. Open de dataset *augustusw12*.
2. Selecteer het tabblad 'Grafieken' en selecteer onder 'Invoegen de lijngrafieken en vervolgens de lijngrafiek.

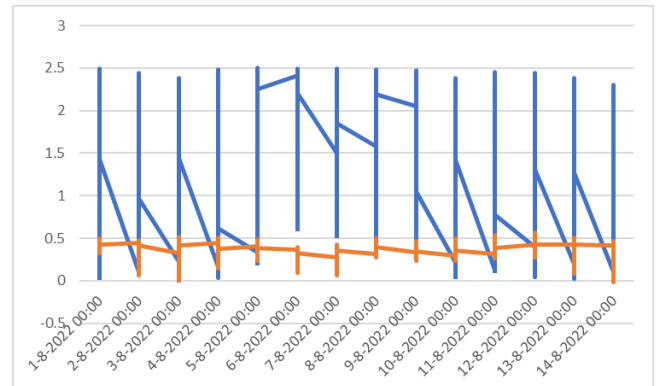
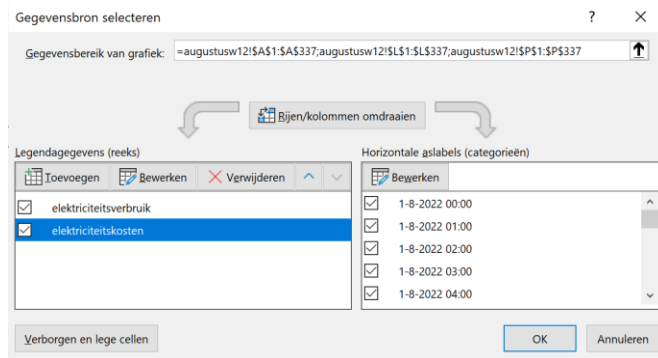


3. Rechtsklik op het witte vierkant dat zojuist verschenen is en klik op 'Gegevens Selecteren'.
4. Voeg de data toe door op 'Toevoegen' te klikken bij de Legendagegevens (reeks).
5. Klik op het pijltje naar boven ↑ naast het eerste tekstvak ("Reeksnaam"), klik op het tabblad 'augustusw12 en selecteer cel L1 met het woord 'electriciteitsverbruik'. Klik hierna op het pijltje naar beneden ↓
6. Klik vervolgens op het pijltje naar boven ↑ naast het tweede tekstvak ("Reekswaarden") voor het toevoegen van de seriewaardes. Selecteer in het tabblad 'Berekeningen' alle waardes onder 'electriciteitsverbruik' (cel L2 t/m cel L337) en klik hierna op het pijltje naar onderen. Druk op

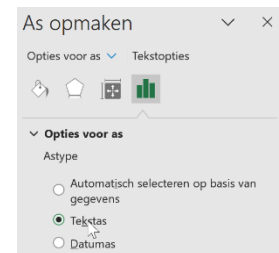


'OK'.

7. Klik op 'Bewerken' onder 'Horizontale aslabels'. Klik op het pijltje naar boven ↑ bij het tekstvak 'Aslabelbereik' en selecteer alle tijdstippen in kolom A (cel A2 t/m cel A37). Druk op 'OK'.
8. Klik nogmaals op 'Toevoegen' onder 'Legendagegevens (reeks)' en voer stappen 5 t/m 7 ook uit voor de elektriciteitskosten. Let er op dat kolom P gebruikt wordt (elektriciteitskosten) in plaats van kolom L (elektriciteitsverbruik). Druk uiteindelijk op 'OK' om de grafiek te maken.



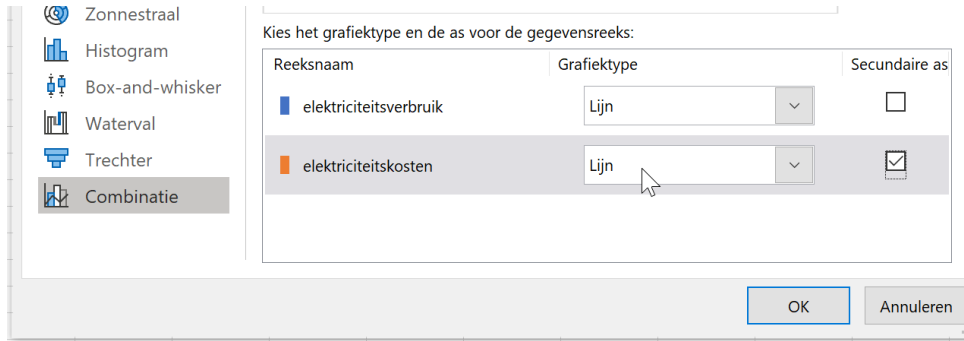
9. De grafiek klopt nog niet helemaal. Klik rechts op de horizontale X-as en selecteer 'As opmaken'. Selecteer onder 'Opties voor as', onder 'Astype' de optie 'Tekstas'. Sluit het menu door op X rechtsboven te klikken.



10. De twee lijnen delen nu nog dezelfde Y-as. Om een betere vergelijking te maken, kan er gebruikt worden gemaakt van een *tweede* Y-as die aan de rechterkant van de grafiek komt te staan. Klik op 'Ander Grafiektype' onder 'Grafiekontwerp'.



11. Selecteer onderaan de lijst de optie 'Combinatie'. Stel het type van zowel elektriciteitskosten als elektriciteitsverbruik in op 'Lijn' en vink 'Secundaire as' aan bij de elektriciteitskosten. Klik op 'OK' om de grafiek af te maken.



12. De grafiek kan verder afgemaakt worden door net als in de eerdere opdracht gebruik te maken van Grafiekelementen. Door zowel een primaire als een secundaire Y-astitel toe te voegen, wordt de grafiek duidelijker.

