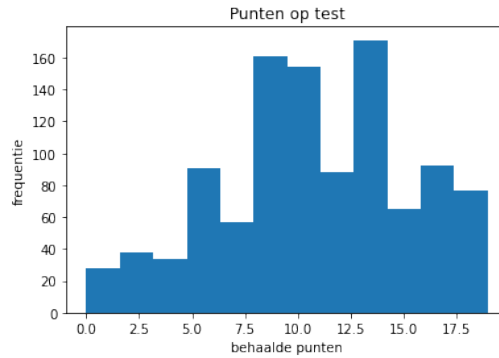
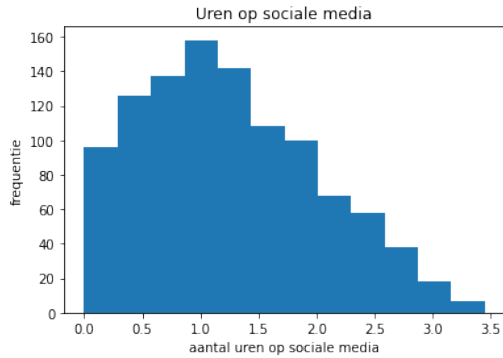


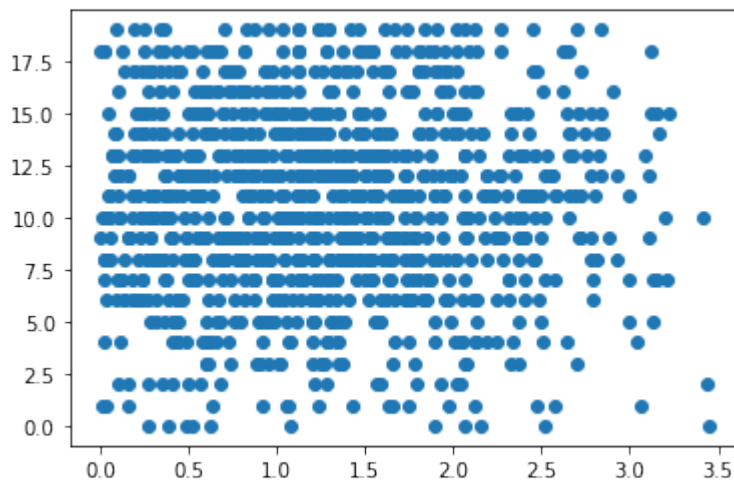
Samenhang Oplossingen

1 Sociale media en studieresultaat

1. 1056
2. 12 klassen:



3. 1.2641
4. 0.7578164. Dit wil zeggen dat de data veel varieert tegenover het gemiddelde
5. 10.8125
6. 4.495904
- 7.

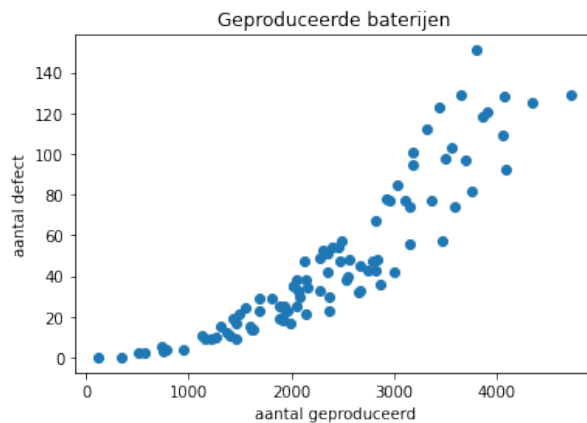


we verwachten een lage correlatie.

8. -0.01466928 Er is geen lineair verband aangetoond tussen de twee variabelen.
9. -0.006172415. Ook deze coëfficiënt geeft aan dat er geen verband is
10. Nee, want er is geen correlatie. Een regressielijn zou maar 0.02152% (R^2) van de variantie verklaren.

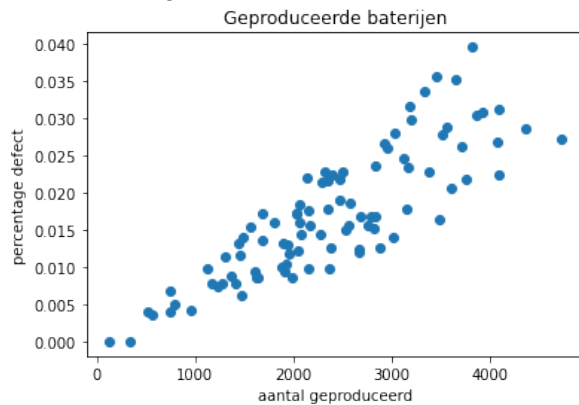
2 Productiefouten batterijen

1. 93
- 2.

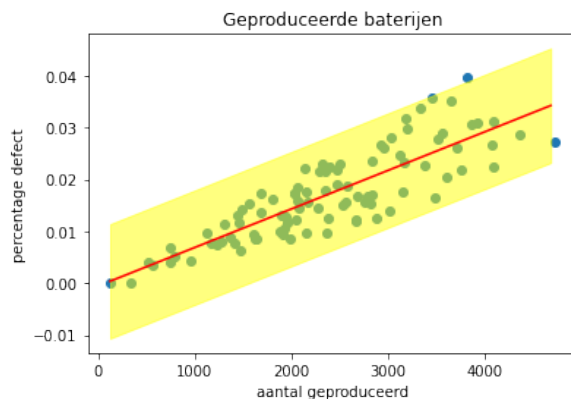


er is geen lineair verband. De punten liggen op een kromme. We mogen hier dus geen lineaire regressie op toepassen.

3. 0.8140
4. 0.01706094 (1,71% dus)
5. de punten liggen niet meer op een kromme maar vormen meer een wolk die van links onder naar rechts boven gaat



6. 0.850309
- er is een zeer hoge correlatie tussen het aantal geproduceerde batterijen en het percentage defecte kaarten. De hogere werkdruk zorgt dus voor een slechtere kwaliteit.
7. $R^2 = 0.7230$ (72,3% van de fouten wordt dus door het tempo van de productie verklaard)
- 8.



intercept = -0.0005587171250994497

slope = 7.439079104402141e-06

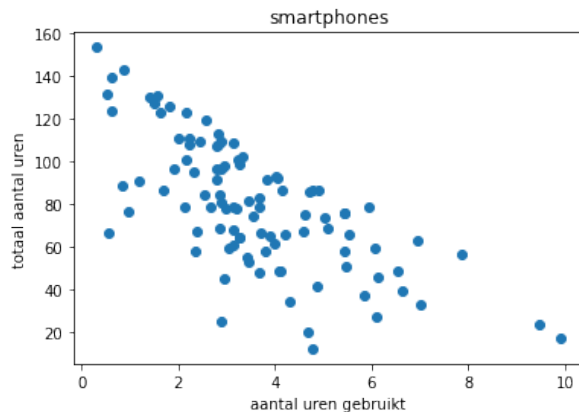
9. 0.011034265073165446 de gevonden lijn zit dus gemiddeld 1,1% naast de echte waarde
10. 0.058953915710117676. We verwachten dus 5,9% defecte kaarten
11. 1419

3 Stress en weersomstandigheden

1. 1090
2. stress: ordinaal, temp: interval, neerslag: ratio, vochtigheid: ratio
3. 10.3211, 79.87523 en 2.370734
4. aangezien temperatuur ordinaal is, kunnen we Pearson niet gebruiken. Een betere maat is Kendall. Dit levert 0.6533315. Er is dus een positieve correlatie tussen deze 2.
5. Kendall levert -0.3230207. Er is dus een licht negatieve correlatie tussen deze 2
6. Kendall levert 0.003270432. Er is dus geen enkele correlatie tussen de neerslag en het stressgevoel
7. Temperatuur speelt dus duidelijk een rol op het stressgevoel. Luchtvochtigheid slechts in heel kleine mate.

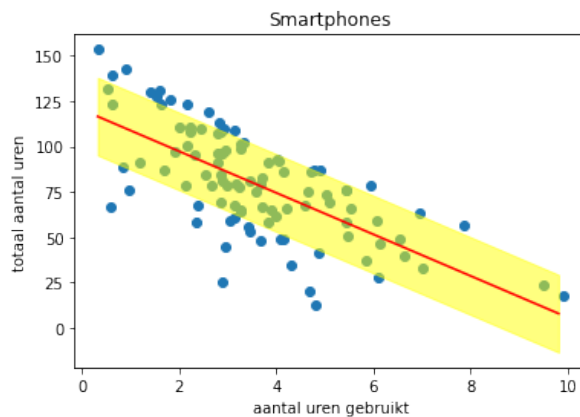
4 Wanneer laad ik mijn smartphone op?

1. 100
2. 96,50573
- 3.



neergaande trend, niet duidelijk of het lineair is

4. -0.696947 en -0.517578
5. er is een redelijk negatief lineair verband tussen de twee

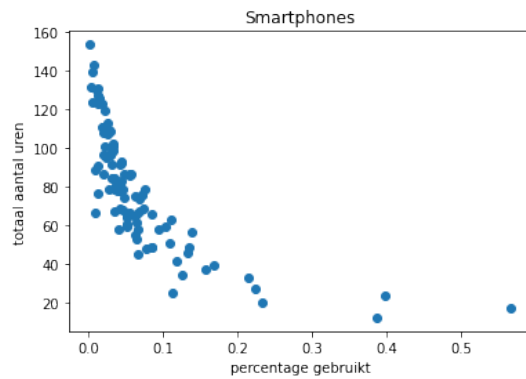


intercept = 120.259

slope = -11.47

6. 0.4857352
48,6% van de variantie in het totaal gebruik wordt verklaard door de variantie in het gebruik van de smartphone. Het gebruik van de smartphone bepaalt dus voor 48,6% de tijd die je tussen 2 laadbeurten hebt. Er is dus ook nog een ander aandeel dat niet aangetoond wordt (kwaliteit batterij, het soort gebruik, welke apps, temperatuur, instellingen, ...).
7. 21.4199. Als je de totale tijd wil voorspellen, zit je er gemiddeld 21,42 uur naast
8. 85.84792

9.



grafiek is duidelijk niet lineair --> geen lineaire regressie mogelijk

je kan niet-lineaire modellen proberen en telkens de verklaarde variantie berekenen:

lineair: 0.499673

kwadratisch: 0.724954

kubisch: 0.780096

logaritmisch: 0.793184

exponentieel: 0.643213

of je kan ook de se berekenen per model:

lineair: 21.127683

kwadratisch: 15.664901

kubisch: 14.006879

logaritmisch: 13.583663

exponentieel: 17.841423

het beste model is dus logaritmisch.

de formule van het model is: $-7.198086 - 27.309926 * \log(x)$

