|  |  |
| --- | --- |
| Naam | Swen Meeuwes |
| Studentnummer | 0887127 |
| Naam | Jesse van Breda |
| Studentnummer | 0902729 |
| Klas | INF2A |
| Docent | Jan Kroon |
| Datum | 29-05-2016 |

De volledige beschrijving van de opdracht vind je hier:

<https://confluence.hr.nl/display/CMIP/Visualisatie+-++inleveropdracht+2>

Opdracht 2A. Oorzaak en gevolg vs. correlatie

2A-1) Als je sterke positieve correlatie tussen twee grootheden A (op x-as) en B (op y-as) vindt wat kan je dan met zekerheid concluderen:

a. A veroorzaakt B

b. B veroorzaakt A

c. A veroorzaakt B niet

d. B veroorzaakt A niet

e. geen van deze antwoorden

|  |
| --- |
| Antwoord: E, omdat er tussen de grootheden A en B een correlatie bestaat, betekent dit niet dat de twee grootheden elkaar veroorzaken (correlatie is namelijk niet hetzelfde als een causaal verband). Echter kunnen we op basis van de gegeven informatie niet met zekerheid concluderen dat er geen oorzakelijk verband bestaat tussen de twee grootheden. Met oorzakelijk verband wordt bedoeld dat de ene grootheid de ander veroorzaakt. |

2A-2) Als je sterke negatieve correlatie tussen twee grootheden A (op x-as) en B (op y-as) vindt wat kan je dan met zekerheid concluderen:

a. A veroorzaakt B

b. B veroorzaakt A

c. A veroorzaakt B niet

d. B veroorzaakt A niet

e. geen van deze antwoorden

|  |
| --- |
| Antwoord: E, zie opdracht 2A-1 |

Opdracht 2B. Scatterplot

Maak een scatterplot van de [data](https://confluence.hr.nl/download/attachments/46466404/scatterplot.txt?version=1&modificationDate=1461319759200&api=v2) van het eerste txt-bestand (EIG1 vs. EIG2).

CAT betekent Categorie, die hoef je in eerste instantie niet mee te nemen. EIG1 en EIG2 zijn fictieve veldnamen, zet de waardes van EIG1 op de x-as en de waardes van EIG2 op de y-as (let op waardenbereik!).

In hoeverre zijn de data gecorreleerd?

|  |
| --- |
| Antwoord: Er is een gematigde positieve correlatie tussen EIG1 (x-as) en EIG2 (y-as). De punten liggen rondom een lijn, maar vormen samen niet een rechte lijn. |

Geef nu elk waarnemingspunt een kleur- of vormaspect afhankelijk van de waarde van CAT en beantwoord de vraag opnieuw: In hoeverre zijn de data gecorreleerd?

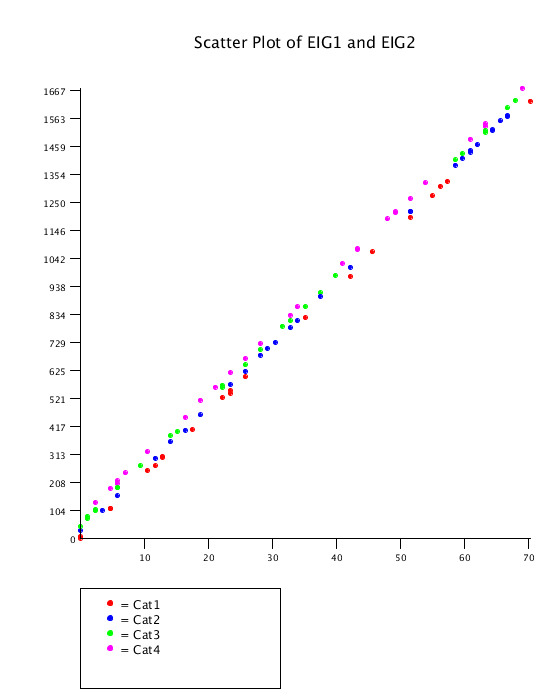
|  |
| --- |
| Antwoord: Als we per categorie kijken, blijkt EIG1 een sterke positieve correlatie te hebben met EIG2. De punten vormen een (bijna) rechte oplopende lijn (van linksonder naar rechtsboven). |

Geef de code van je visualisatie:

|  |
| --- |
| Code van onze git repository: https://github.com/jvanbreda/Dev8-Visualisation-Opdracht2 |

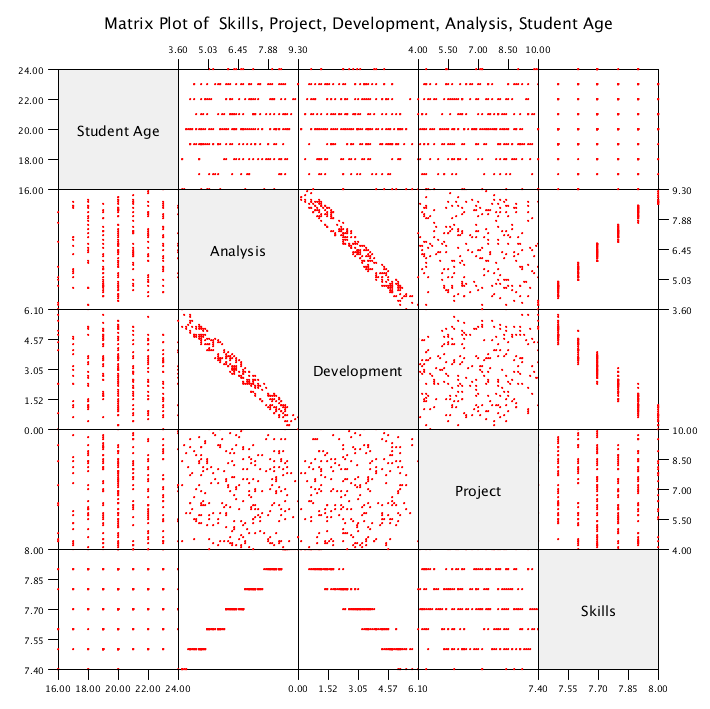
Maak tot slot een screenshot van de visualisatie:

|  |
| --- |
| Screenshot van de visualisatie: |



Opdracht 2C. Scattermatrix

|  |
| --- |
| Screenshot van de visualisatie |



|  |
| --- |
| Code of jouw git repository: https://github.com/jvanbreda/Dev8-Visualisation-Opdracht2 |

Opdracht 2D

Welke conclusies kan je trekken uit de scattermatrix?

|  |
| --- |
| Antwoord: We kunnen het volgende concluderen:   * Er is een gematigde negatieve correlatie tussen de cijfers van Analyse en Development. Dit **zou** kunnen betekenen dat studenten die hoge cijfers halen voor analyse lage cijfers halen voor development. * Tussen de cijfers van Analyse en Skills is er een zwakke positieve correlatie. Dit **zou** kunnen betekenen dat studenten die hoge cijfers halen voor analyse ook hoge cijfers halen voor development. * Ook is er een zwakke negatieve correlatie tussen de cijfers van Development en Skills. Dit **zou** kunnen beteken dat studenten die hoge cijfers halen voor Development lage cijfers halen voor Skills. * We zouden mogen concluderen dat de leeftijd van de student geen invloed heeft op de cijfers die hij haalt. * Verder is er geen correlatie tussen cijfers van Analyse en Project, hetzelfde geldt voor de cijfers van Development en Project. * Tenslotte is er geen correlatie tussen de cijfers van Project en Skills. |