

DE FRIESLAND
ZORGVERZEKERAAR



SOGETI

Arie Twigt

Hoofdvraag

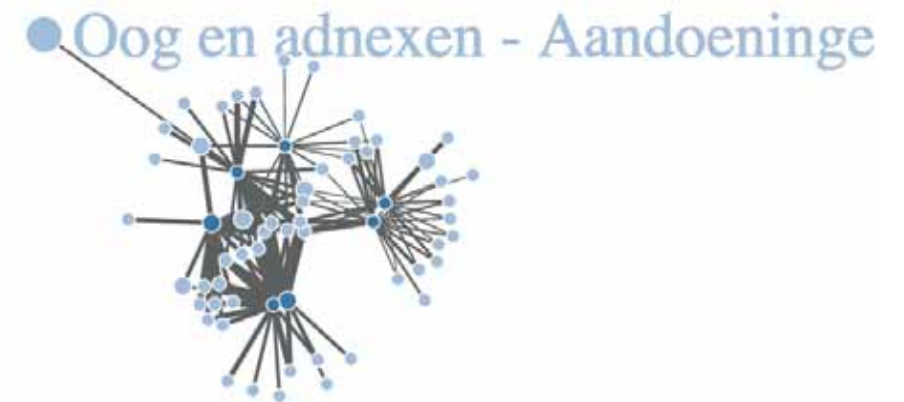
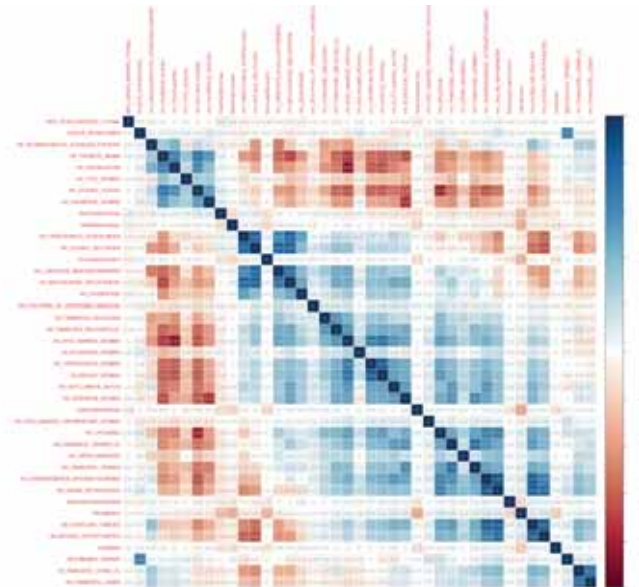
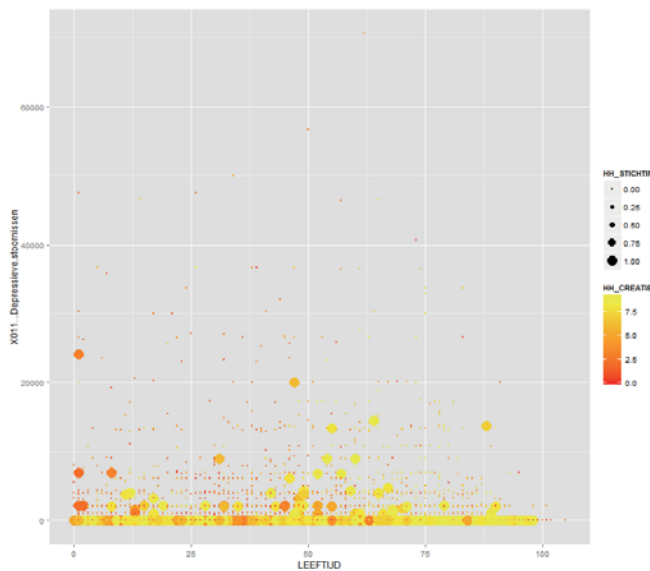
Hoe kan interne en externe data gebruikt worden voor ziektepreventie bij klanten van DFZ?

- Data visualisatie (Grafieken, dashboards);
- Kwantitatieve analyse (cijfers, statistiek);

Software

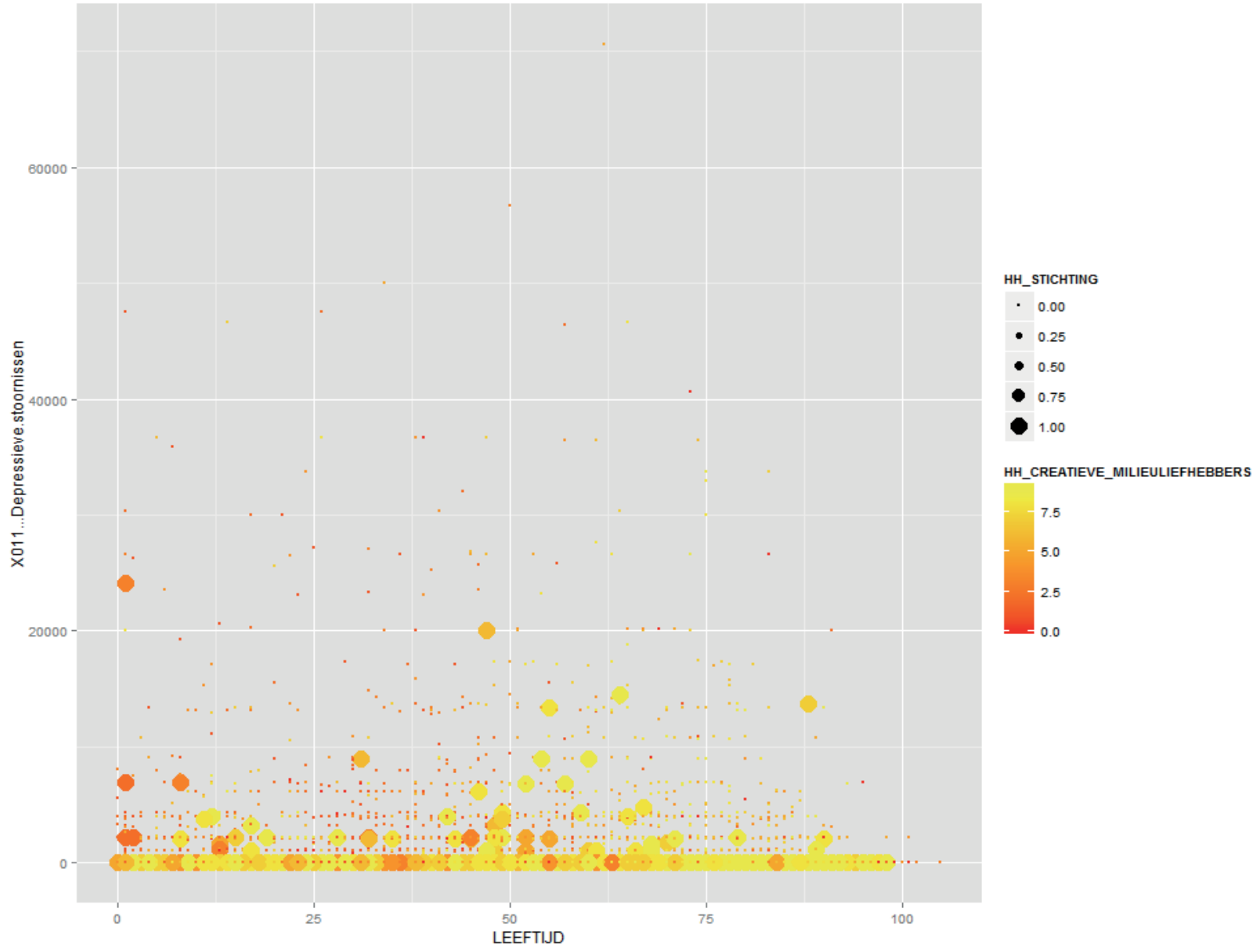


Inzichten krijgen uit visualisaties



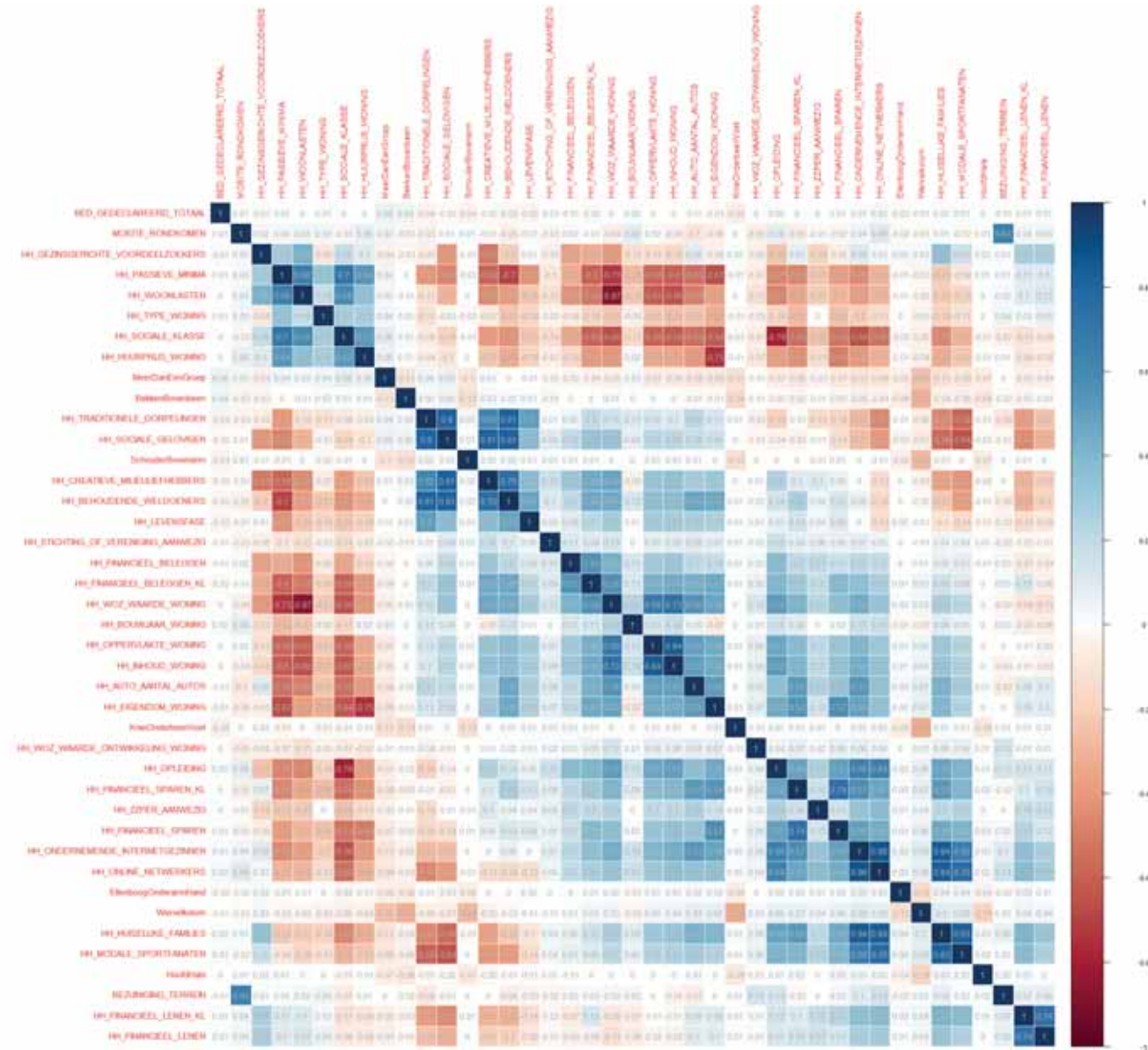


Depressiviteit – Vereniging/stichting



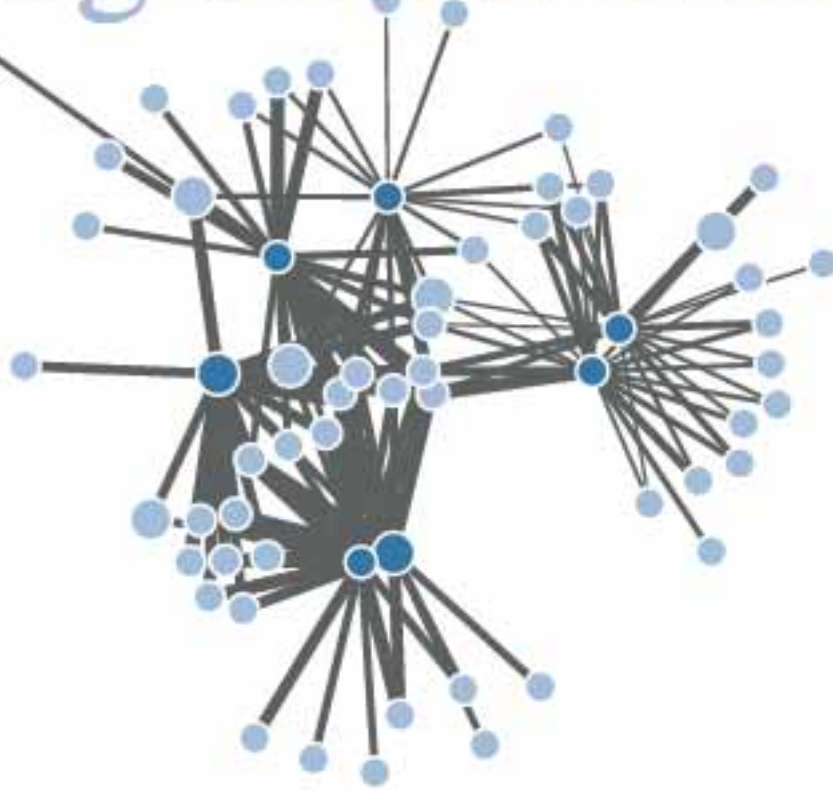
Nieuwe feiten ontdekken: op zoek naar samenhang

Correlatieplot



Aandoeningen binnen leeftijdsgroepen

● Oog en adnaxen - Aandoeninge



Beschrijvend statistisch onderzoek.

Meer botfracturen oplopen tijdens **vakantiedagen** of **niet-vakantiedagen**?

2.2 Hypotheses

Op basis van onderzoeksvragen zijn er volgende hypothesen opgesteld

2.2.1 Hypothese obv Vraag 1

H₀: Het gemiddeld aantal verzekerden dat tijdens een schoolvakantieweek kosten declareert ivm ongeval is **gelijk** aan het gemiddeld aantal verzekerden dat tijdens een niet-schoolvakantieweek kosten declareert ivm ongeval.

H₁: Het gemiddeld aantal verzekerden dat tijdens een schoolvakantieweek kosten declareert ivm ongeval is **niet gelijk** aan het gemiddeld aantal verzekerden dat tijdens een niet-schoolvakantieweek kosten declareert ivm ongeval.

2.2.2 Hypothese obv Vraag 2

H₀: Het gemiddeld gedeclareerde bedrag tijdens een schoolvakantieweek ivm ongeval is **gelijk** aan het gemiddeld gedeclareerde bedrag tijdens een niet-schoolvakantieweek ivm ongeval.

H₁: Het gemiddeld gedeclareerde bedrag tijdens een schoolvakantieweek ivm ongeval is **niet gelijk** aan het gemiddeld gedeclareerde bedrag tijdens een niet-schoolvakantieweek ivm ongeval.

Declaraties voor botfracturen tijdens en buiten de vakanties

Snel en efficiënt een statistisch onderzoeksrapport samenstellen.

Pdf: "Ongevallen schoolvakanties 2016"

Stappen:

1. Data verkenning;
2. Data voorbereiding
3. Hypotheses stellen;
4. Statistische test uitkiezen;
5. Test uitvoeren;
6. Resultaten/output tests bekijken en beoordelen (evt. met visualisaties);
7. Conclusies trekken.

Resultaten

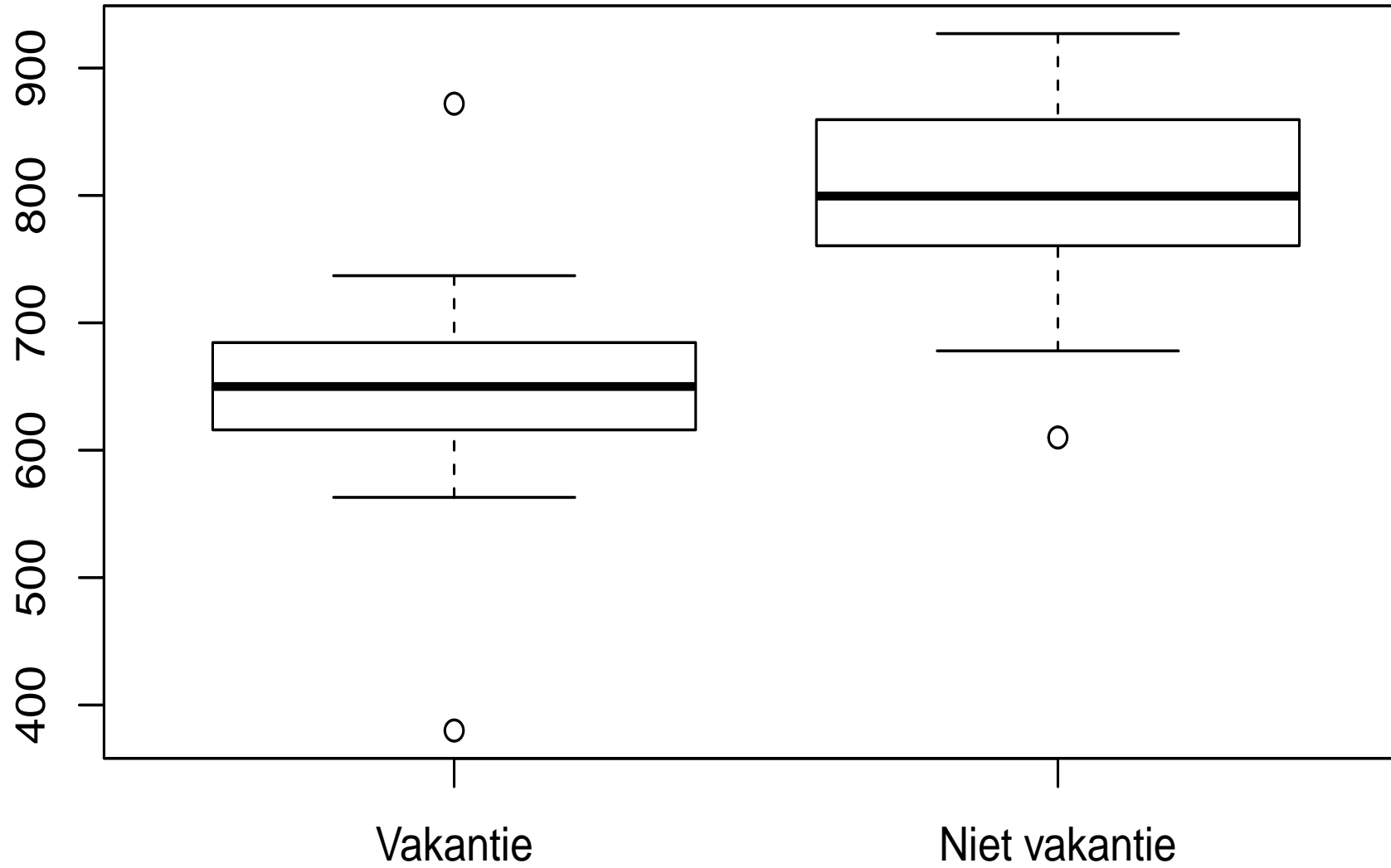
3.6.3 Testen voor significant verschil VakantieAantal en NietVakantieAantal

3.6.4 t-test voor testen op significant verschil

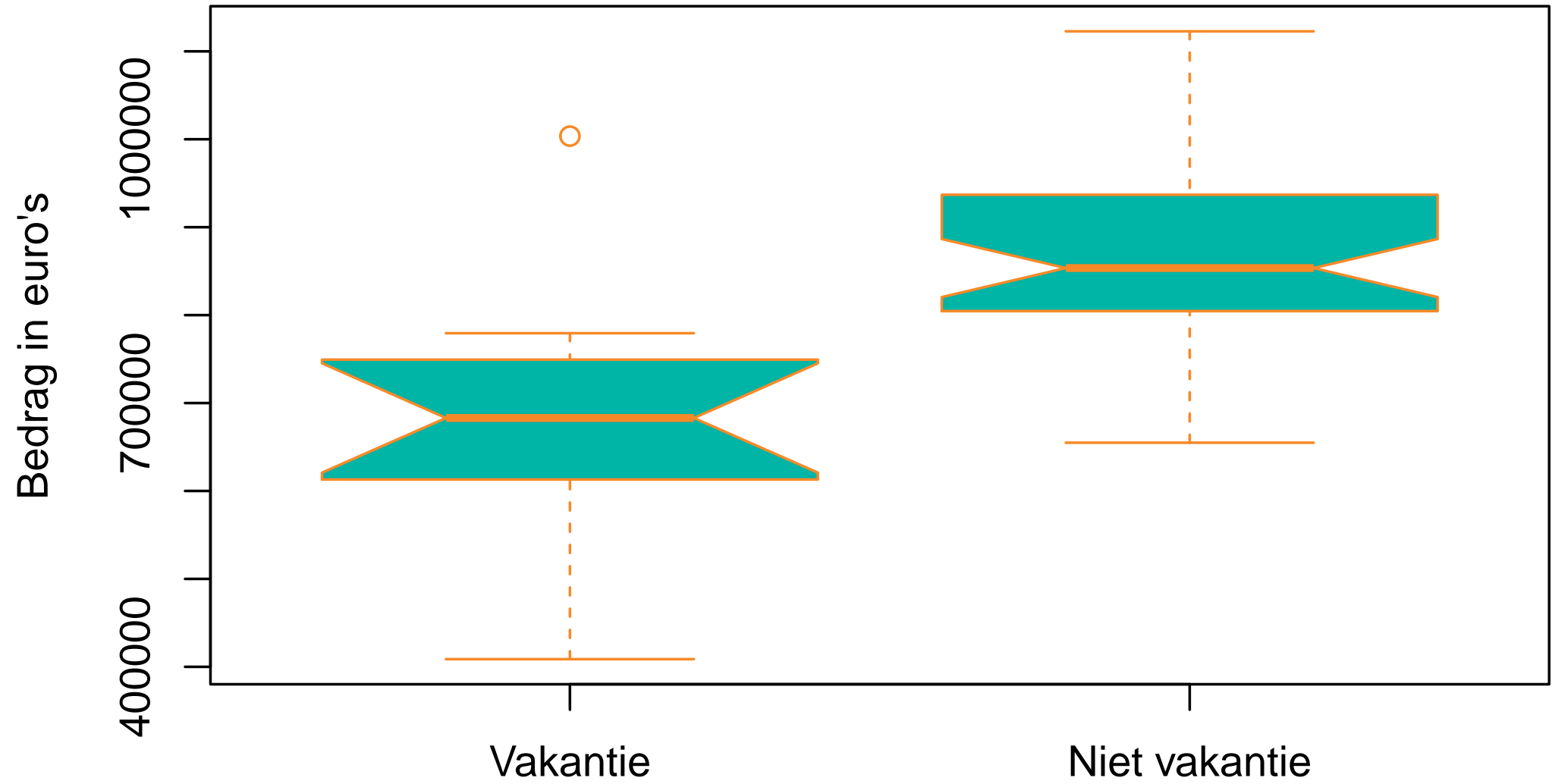
```
t.test(VakantieAantal, NietVakantieAantal, paired = FALSE)
```

```
##  
## Welch Two Sample t-test  
##  
## data: VakantieAantal and NietVakantieAantal  
## t = -4.559, df = 13.579, p-value = 0.000481  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -232.66993 -83.49673  
## sample estimates:  
## mean of x mean of y  
## 645.1667 803.2500
```

Resultaten



Declaratiebedragen ongevallen



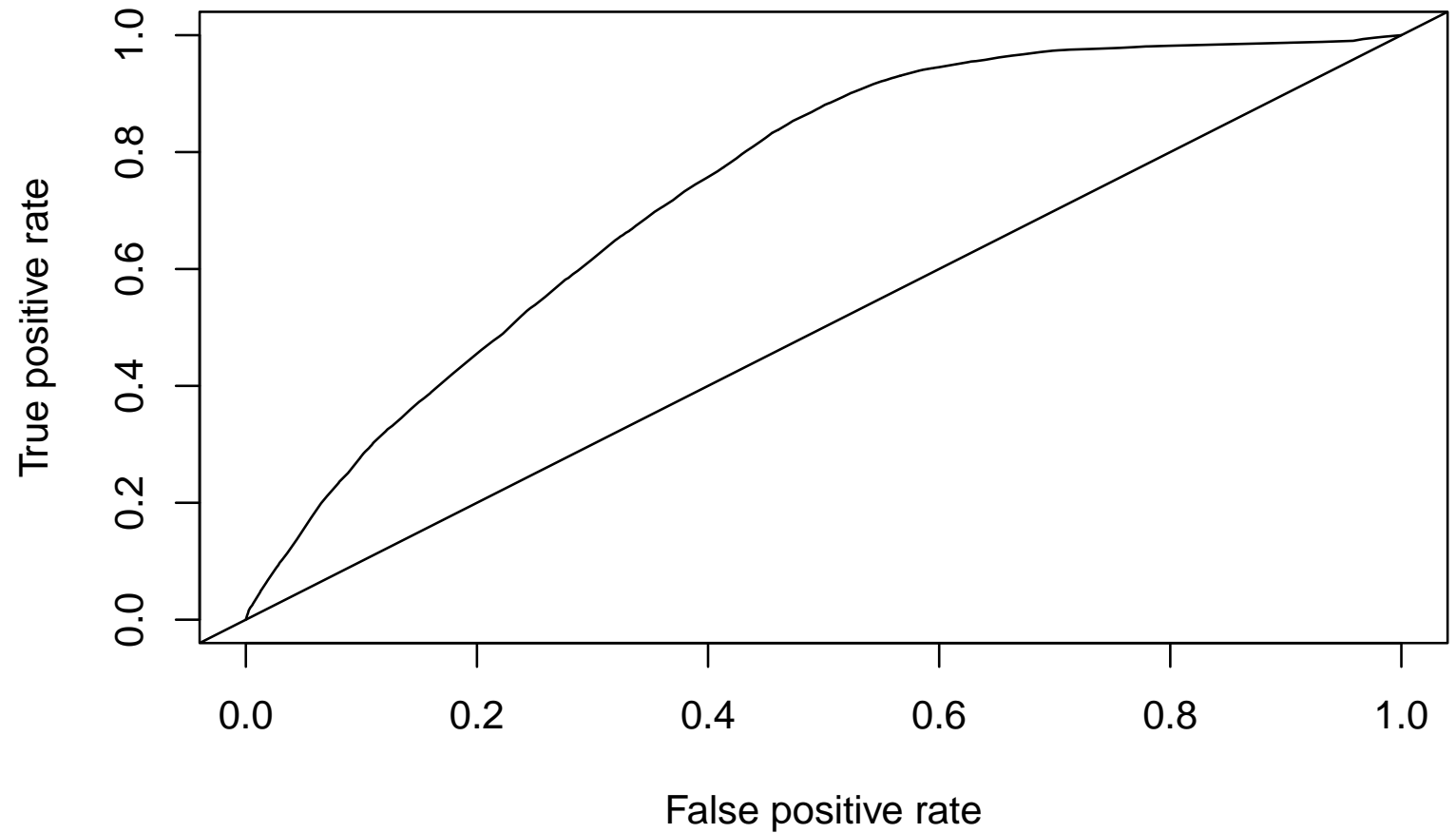
Voorspellingen: “Hoe groot is de kans dat deze klant gaat declareren?”

Pdf: “Voorspelling Fysiologie gebruik”:

1. Data verkenning;
2. Data voorbereiding;
3. Voorspellingsmodel uitkiezen;
4. Variabelen uitkiezen om mee te voorspellen → *Feature vector*;
5. Voorspellingsmodel toepassen/experiment uitvoeren;
6. Resultaten van experiment evalueren en interpreteren;
 1. Scoren met tabel;
 2. Visualiseren.
7. Indien nodig, model aanpassen voor betere resultaten.

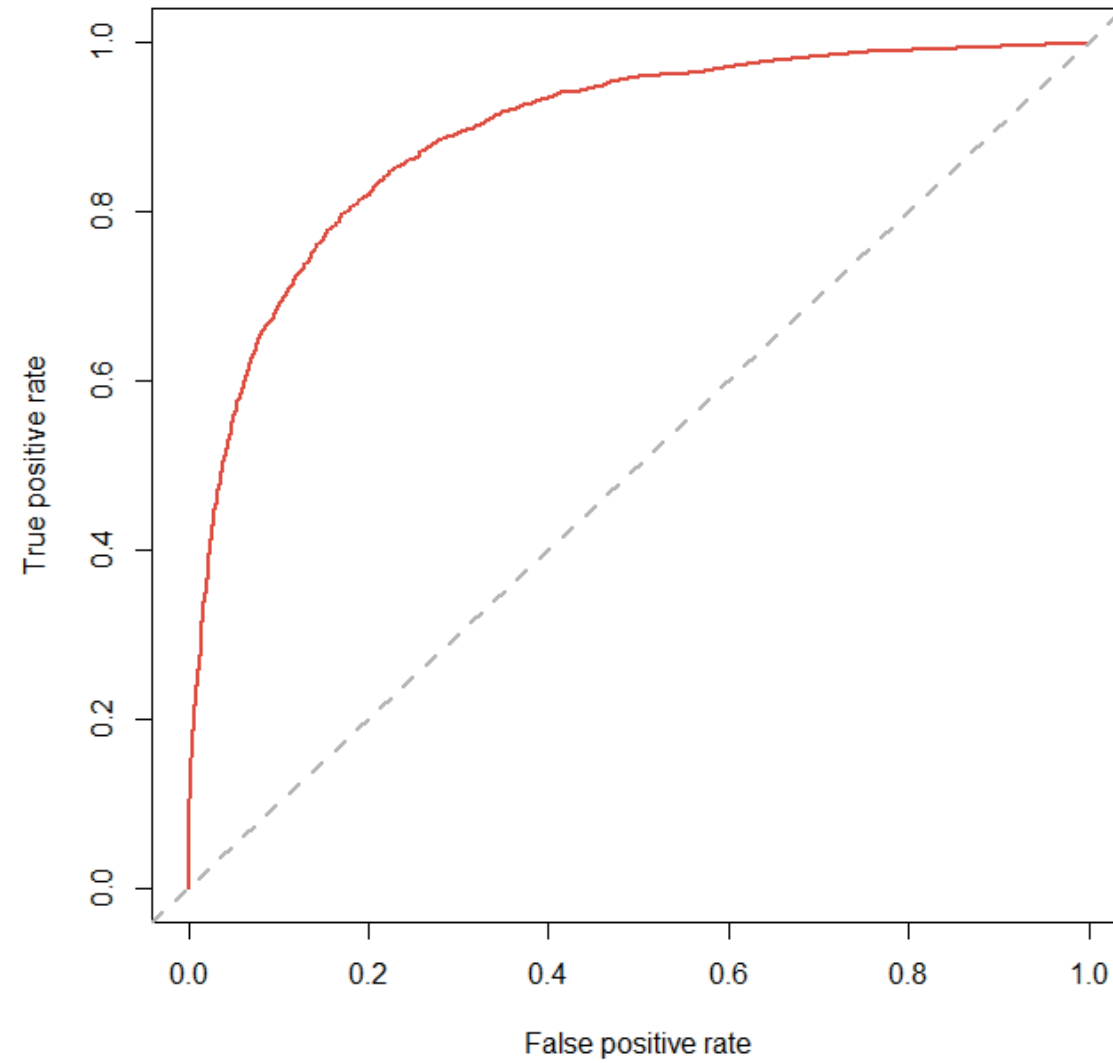
Resultaten/model beoordelen

```
##  
##           0           1  
##  Geen 116251 20807  
##  Wel   18220 11051
```

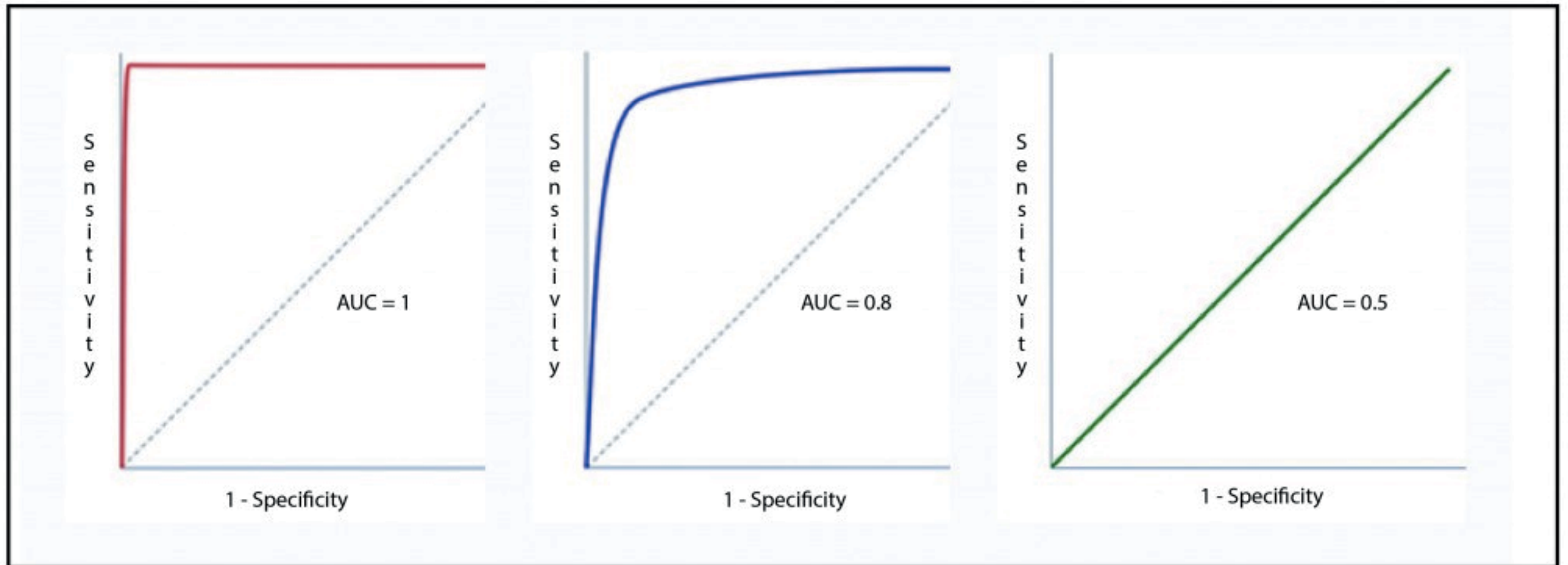


ROC Curve

ROC Curve for Random Forest

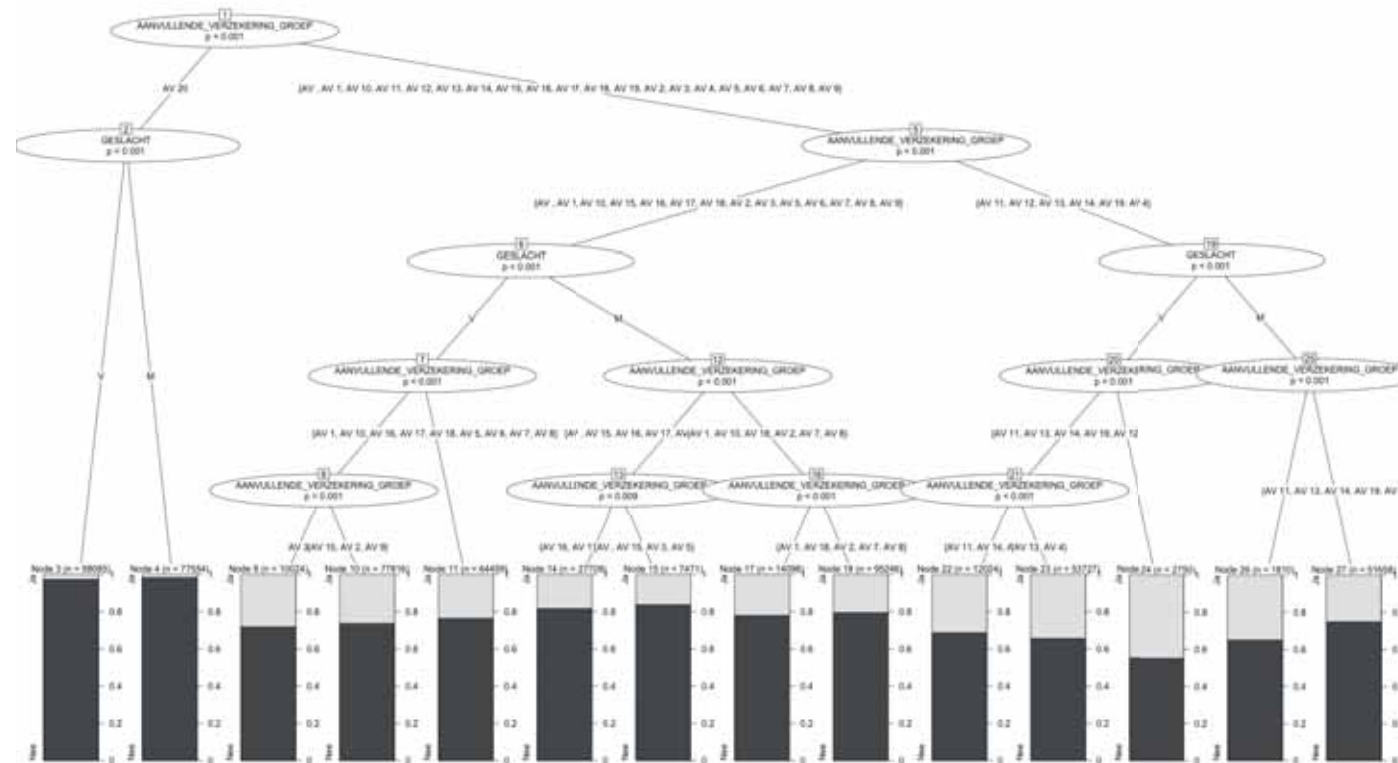


Resultaat van goed naar slecht model



Voorspellingen: "Bij welke type klanten is de kans het grootst dat zij een declaratie indienen?"

Decision Tree → Kansen per profiel met eigenschappen.



Kwantitatieve uitkomsten voorspellen (Bedragen)

Voorspellen uitgave aan Psychologie gerelateerde declaraties op basis van andere declaraties.

Lineaire Regressie

1. Data verkenning;
2. Data voorbereiding;
3. Combinatie van voorspellende variabelen uitkiezen;
4. Voorspellingsmodel toepassen/experiment uitvoeren;
5. Resultaten van experiment evalueren en interpreteren;
 1. Coëfficiënten en significantie per variabele bekijken;
 2. Betrouwbaarheid van het model bekijken.
6. Indien nodig, model aanpassen voor betere resultaten.

Uitkomst eerste model

```
> model<-lm(Botspierstelsel.bindweefsel.Letsel...ziekten.botspierstelsel.bindweefsel.en.late.gevolgen.vroege.compl.trauma..wd._2014~X0...Chirurgie_Bewegingsa  
pparaat  
+ ,data=FZ4)  
> summary(model)
```

```
Call:  
lm(formula = Botspierstelsel.bindweefsel.Letsel...ziekten.botspierstelsel.bindweefsel.en.late.gevolgen.vroege.compl.trauma..wd._2014 ~  
X0...Chirurgie_Bewegingsapparaat, data = FZ4)
```

```
Residuals:  
    Min       1Q   Median       3Q      Max  
-14431   -255    -255    -255   37164
```

```
Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	254.99551	4.63459	55.02	<2e-16 ***
X0...Chirurgie_Bewegingsapparaat	1.62641	0.01846	88.12	<2e-16 ***

```
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1216 on 70794 degrees of freedom
```

```
Multiple R-squared:  0.09884, Adjusted R-squared:  0.09883
```

```
F-statistic: 7765 on 1 and 70794 DF, p-value: < 2.2e-16
```

Uitkomst tweede model

```
> model<-lm(Botspierstelsel.bindweefsel.Letsel...Ziekten.botspierstelsel.bindweefsel.en.late.gevolgen.vroege.compl.trauma..wd._2014~X0...Chirurgie_Bewegingsa
pparaat+HULPMIDDELEN.VOOR.TOILETGEBRUIK+HULPMIDDELEN.OM.ZICH.TE.WASSEN..TE.BADEN.EN.TE.DOCHEN+TAFELS+NIET.GEMOTORISEERDE.ROLSTOELN+TRANSFERHULPMIDDELEN.EN.K
EERHULPMIDDELEN+Botspierstelsel.bindweefsel.Letsel...Ziekten.botspierstelsel.bindweefsel.en.late.gevolgen.vroege.compl.trauma..wd.+X2...Surmenage_Degeneratie
ve.aandoeningen_Dystrofie+Anesthesiologie..Pijnbestrijding
+ ,data=FZ4)
> summary(model)
```

```
Call:
lm(formula = Botspierstelsel.bindweefsel.Letsel...Ziekten.botspierstelsel.bindweefsel.en.late.gevolgen.vroege.compl.trauma..wd._2014 ~
X0...Chirurgie_Bewegingsapparaat + HULPMIDDELEN.VOOR.TOILETGEBRUIK +
HULPMIDDELEN.OM.ZICH.TE.WASSEN..TE.BADEN.EN.TE.DOCHEN +
TAFELS + NIET.GEMOTORISEERDE.ROLSTOELN + TRANSFERHULPMIDDELEN.EN.KEERHULPMIDDELEN +
Botspierstelsel.bindweefsel.Letsel...Ziekten.botspierstelsel.bindweefsel.en.late.gevolgen.vroege.compl.trauma..wd. +
X2...Surmenage_Degeneratieve.aandoeningen_Dystrofie +
Anesthesiologie..Pijnbestrijding, data = FZ4)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-14145   -222   -204   -204   37214
```

Coefficients:

```
(Intercept)
X0...Chirurgie_Bewegingsapparaat
HULPMIDDELEN.VOOR.TOILETGEBRUIK
HULPMIDDELEN.OM.ZICH.TE.WASSEN..TE.BADEN.EN.TE.DOCHEN
TAFELS
NIET.GEMOTORISEERDE.ROLSTOELN
TRANSFERHULPMIDDELEN.EN.KEERHULPMIDDELEN
Botspierstelsel.bindweefsel.Letsel...Ziekten.botspierstelsel.bindweefsel.en.late.gevolgen.vroege.compl.trauma..wd.
X2...Surmenage_Degeneratieve.aandoeningen_Dystrofie
Anesthesiologie..Pijnbestrijding
```

Estimate	Std. Error	t value	Pr(> t)
204.34903	5.05403	40.433	< 2e-16 ***
1.56022	0.01849	84.378	< 2e-16 ***
8.01316	0.64857	12.355	< 2e-16 ***
9.34610	1.26564	7.384	1.55e-13 ***
13.05800	1.36979	9.533	< 2e-16 ***
2.06624	0.72998	2.831	0.00465 **
10.77582	1.74579	6.172	6.76e-10 ***
17.73826	1.42791	12.423	< 2e-16 ***
0.13213	0.01174	11.251	< 2e-16 ***
73.82759	7.34662	10.049	< 2e-16 ***

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1208 on 70786 degrees of freedom
```

```
Multiple R-squared:  0.111,    Adjusted R-squared:  0.1109
```

```
F-statistic: 381.7 on 9 and 70786 DF, p-value: < 2.2e-16
```

Belangrijke aspecten

Kennis van de business:

Aandoeningen;
Producten;
Klanten.

Kennis van de technologie:

Excel;
SQL;
R.

Kennis van statistiek:

Statistische toetsen;
(Voorspellings) modellen;
Interpretatie van
resultaten.

