

Logboek: Publieke Opinie

Floris Menninga

2024-04-22

```
# Libraries
library(data.tree)
library(igraph)
library(tidyverse)
library(viridis)
library(ggplot2)
library(dplyr)
library(rgexf)
library(ggrepel)
library("ggpubr")

# Voor statistische analyses:
library("pwr")
```

Inleiding:

Vragen over het abortus beleid zijn polariserende ethische kwesties. In ons onderzoek trachten we correlaties te vinden tussen verschillende variabelen zoals: leeftijd, geslacht, opleidingsniveau, geloof etc. en de ideologische verschillen in mening tussen deze mensen.

Om de meningen van het publiek te peilen met betrekking tot hun ethische normen en waarden is er een vragenlijst opgesteld die de volgende vragen en stellingen bevat:

Materialen en Methoden

Voor de enquête zijn er een aantal vragen opgesteld die een beeld moeten geven van de meningen van de mensen die hem invullen. Ook zijn er persoonlijke vragen die gebruikt worden om keuzes te correleren met factoren als: geloof, leeftijd, opleidingsniveau en geslacht.

Deze vragen zijn neutraal geformuleerd zodat de invuller niet onbewust geneigd is om voor de optie te kiezen waar de vraagstelling neigt te gaan. De vragen kunnen verdeeld worden in drie categorieën: Persoonlijke informatie, multiple choice en op een schaal van zeer oneens tot zeer eens.

De enquête bestaat uit de volgende vragen:

Persoonlijke vragen:

1. Wat is uw geslacht?
2. Wat is het hoogste onderwijsniveau dat u heeft afgerond?
3. Wat is uw leeftijd?
4. Wat is uw religieuze overtuiging?

5. In welke provincie heeft u het grootste deel van uw jeugd doorgebracht?

Multiple choice vragen:

1. Vindt u dat iemand zelf moet kunnen kiezen of die een abortus wil?
2. Indien u bij de vorige vraag ‘Nee’ heeft ingevuld:
3. Vindt u dat er uitzonderingen zijn waarin iemand dit wel zelf moet kunnen kiezen?
4. Vindt u dat er voor abortus gekozen mag worden als de gezondheid van de vrouw door de zwangerschap in gevaar kan komen?
5. Vindt u dat iemand een abortus mag als de zwangerschap een gevolg is van verkrachting?
6. Vindt u dat een minderjarig kind de keuze moet hebben om voor een abortus te kiezen?
7. Moet het wettelijk verplicht worden voor vrouwen die een abortus willen om een echo van de foetus te zien voordat een abortus toegestaan is?
8. Vindt u dat de wettelijke bedenktijd moet terugkomen?
9. Vindt u dat adoptie een alternatief is voor abortus?
10. Moet de biologische vader zeggenschap hebben over de beslissing voor een abortus?

Schaal zeer oneens tot zeer eens vragen:

1. Hoe goed kunt u zichzelf vinden in de volgende stellingen?
2. Het leven van de vrouw in gevaar is.
3. Het kind met een geestelijke handicap geboren kan worden.
4. Het kind met een levensbedreigende ziekte geboren kan worden.
5. De zwangerschap veroorzaakt is door verkrachting.
6. De vrouw het kind niet wil houden om niet gespecificeerde redenen.
7. De vrouw niet in staat is voor het kind te zorgen vanwege financiële beperkingen.

De vraag: Vindt u dat de wettelijke bedenktijd moet terugkomen? Werd aangevuld met de volgende informatie: Sinds 1 januari 2023 is de verplichte minimale bedenktijd van 5 dagen vervallen. Voorheen moesten er minimaal 5 dagen zitten tussen het gesprek met een arts en de abortus zelf. Op dit moment geldt een flexibele bedenktijd die wordt bepaald door een gesprek tussen de vrouw die abortus wil en een arts. De arts bepaalt nu op basis van dit gesprek hoelang de bedenktijd moet zijn voor de vrouw. Dit kan 0 dagen zijn, maar ook 10.

En de vraag: Vindt u dat de ouders zeggenschap moeten hebben over of hun minderjarige kind wel of niet voor abortus mag kiezen? werd aangevuld met de volgende informatie: In de huidige situatie moet er toestemming zijn van de ouders, er zijn wel uitzonderingen. Nu mag je onder de 15 abortus te plegen met toestemming van de ouders.

Deze vragen zijn neutraal geformuleerd zodat de respondent niet onbewust geneigd is om voor de optie te kiezen waar de vraagstelling naar neigt te gaan. Microsoft Forms is gebruikt om een enquête te maken die deze bovenstaande vragen bevat.

Voor de dataverwerking, zoals beschreven is in de drie logboeken: Logboek_Iris, Logboek_Jasper en Logboek_Floris.

Hiervoor zijn de volgende software versies gebruikt:

Library naam:	Versie	Gebruikt voor:	Referentie:
R	4.4.0 (2024-04-24)	Visualisatie van hiërarchische data	https://www.r-project.org/
ggraph	2.2.1	Circular packing grafieken	https://cran.r-project.org/web/packages/ggraph/
tidyverse	2.0.0	“%>% operator”	https://www.tidyverse.org/

Library naam:	Versie	Gebruikt voor:	Referentie:
viridis	0.6.5	Kleur thema	https://cran.r-project.org/web/packages/viridis/
ggplot2	3.5.1	Alle ggplot grafieken	https://ggplot2.tidyverse.org/
dplyr	1.1.4	Mutate(), select() en filter()	https://dplyr.tidyverse.org/
ggrepel	0.9.5	Pie-chart labels	https://cran.r-project.org/web/packages/ggrepel/
ggridges	0.5.6	Ggridge density plot	https://cran.r-project.org/web/packages/ggridges/
pwr	1.3.0	Power analyse	https://cran.r-project.org/web/packages/pwr/
gridExtra	2.3	Meerdere grafieken in een grid	https://cran.r-project.org/web/packages/gridExtra/

Hypothese Mijn persoonlijke verwachting is dat vragen die alleen op een van de twee geslachten van toepassing zijn voor grotere verdeeldheid onder de respondenten zal zorgen. Bijvoorbeeld de vraag: Moet de vader zeggenschap hebben over de abortus. Bij deze vraag is het waarschijnlijk dat mannen vaker gaan instemmen dan vrouwen. Daarnaast zal er mogelijk een verschil tussen de meningen van mensen met andere leeftijden zitten. Bij de vraag of de ouders zeggenschap moeten hebben over de mogelijke abortus van het kind, hier vermoed ik dat kinderen “nee” kiezen terwijl volwassenen (mogelijk ouders) hier vaker voor “Ja” kiezen.

Het verspreiden van de enquête:

De weblink die verkregen was van Microsoft Forms, is middels digitale communicatiemiddelen zoals: E-mail, whatsapp, discord etc. verspreid.

De makers van deze enquête hebben hem naar klasgenoten, familieleden en vrienden gestuurd met de instructie om hem door te sturen naar andere Familie, kennissen en vrienden. Op deze manier zijn er in totaal 114 respondenten geweest die hem ingevuld hebben.

Voor alle vragen van de enquête zijn hypotheses opgesteld om de onderzoeksvraag te beantwoorden, deze onderzoeksvraag is: Is er een verschil tussen de ethiek gerelateerde meningen van mensen van andere leeftijden, geslachten, opleidingsniveaus en geloven? De onderstaande H0 en H1 hypotheses zijn er voor elke onafhankelijke variabele, hier zijn ze voor geslacht ingevuld maar deze zelfde vraagstelling is ook voor: leeftijd, opleidingsniveau en geloof gebruikt.

Vragen en de bijbehorende H0 en H1 hypotheses:

Vraag: Vindt u dat er voor abortus gekozen mag worden als de gezondheid van de vrouw door de zwangerschap in gevaar kan komen?

H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of een vrouw een abortus mag hebben wanneer haar leven in gevaar is.

H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of een vrouw een abortus mag hebben wanneer haar leven in gevaar is.

Vraag: Vindt u dat iemand een abortus mag als de zwangerschap een gevolg is van verkrachting?

H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze over het toestaan

van een abortus wanneer de zwangerschap gevolg is van verkrachting.

H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze over het toestaan van een abortus wanneer de zwangerschap gevolg is van verkrachting.

Vraag: Vindt u dat een minderjarig kind de keuze moet hebben om voor een abortus te kiezen?

H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of een minderjarig kind de keuze moet hebben om voor een abortus te kiezen.

H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of een minderjarig kind de keuze moet hebben om voor een abortus te kiezen.

Vraag: Vindt u dat de ouders zeggenschap moeten hebben over of hun minderjarige kind wel of niet voor abortus mag kiezen?

H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of de ouders van een minderjarig kind zeggenschap moeten hebben over of hun minderjarige kind wel of niet voor abortus mag kiezen?

H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of de ouders van een minderjarig kind zeggenschap moeten hebben over of hun minderjarige kind wel of niet voor abortus mag kiezen?

Vraag: Moet het wettelijk verplicht worden voor vrouwen die een abortus willen om een echo van de foetus te zien voordat een abortus toegestaan is?

H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of ze vinden dat het wettelijk verplicht worden voor vrouwen die een abortus willen om een echo van de foetus te zien voordat een abortus toegestaan is.

H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of ze vinden dat het wettelijk verplicht worden voor vrouwen die een abortus willen om een echo van de foetus te zien voordat een abortus toegestaan is.

Vraag: Vindt u dat de wettelijke bedenktijd moet terugkomen?

H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of de wettelijke bedenktijd terug moet komen.

H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of de wettelijke bedenktijd terug moet komen.

Vraag: Vindt u dat adoptie een alternatief is voor abortus?

H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of adoptie een alternatief is voor abortus.

H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of adoptie een alternatief is voor abortus.

Vraag: Moet de biologische vader zeggenschap hebben over de beslissing voor een abortus?

H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of de biologische vader zeggenschap moet hebben over de beslissing voor een abortus. H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of de biologische vader zeggenschap moet hebben over de beslissing voor een abortus.

```
abortus <- read.csv("abortus_dataset.csv",  
                  na.strings = "NA")  
#head(abortus)
```

```
#summary(abortus)
#str(abortus)
```

Summary(abortus) en str(abortus) geven veel output omdat de kolom namen erg lang zijn. Indien gewenst kunnen de “#” tekens verwijderd worden om dit toch te laten zien. Het zelfde is gedaan met head(abortus) in de code chunk daar boven.

Omzetten naar factoren

In het onderstaande blok code worden factoren gemaakt van de meerkeuze vragen. Het eerste deel bestaat uit factoren waar geen volgorde in zit, daarna factoren met volgorde.

```
#table(abortus$Vindt.u.dat.iemand.zelf.moet.kunnen.kiezen.of.die.een.abortus.wil., useNA = "always")

# voor de vraag: Vindt u dat iemand zelf moet kunnen kiezen of die een abortus wil?
abortus$eigen_keuze_abortus <- factor(x = abortus$Vindt.u.dat.iemand.zelf.moet.kunnen.kiezen.of.die.een
                                     levels = c("Ja", "Nee"),
                                     labels = c("Wel abortus toestaan", "Geen abortus toestaan"))

# Voor de vraag: Wat is uw religieuze overtuiging?
abortus$geloof <- factor(x = abortus$Wat.is.uw.religieuze.overtuiging.,
                       levels = c("Christendom ", "Islam", "Hindoeïsme", "Boeddhisme", "Niet gelovig", "Geen geloof"),
                       labels = c("Christendom", "Islam", "Hindoeïsme", "Boeddhisme", "Niet gelovig", "Geen geloof"))

abortus$geslacht <- factor(x = abortus$Wat.is.uw.geslacht.,
                         levels = c("Man", "Vrouw", "Wil ik liever niet melden"),
                         labels = c("Man", "Vrouw", "Wil ik liever niet melden"))

# Voor de vraag: Wat is uw religieuze overtuiging?

abortus$afkomst <- factor(x = abortus$In.welke.provincie.heeft.u.het.grootste.deel.van.uw.jeugd.doorgebracht.,
                        levels = c("Drenthe", "Flevoland", "Friesland", "Gelderland", "Groningen", "Limburg", "Noord-Brabant", "Noord-Holland", "Overijssel", "Rijnland", "Zeeland"),
                        labels = c("Drenthe", "Flevoland", "Friesland", "Gelderland", "Groningen", "Limburg", "Noord-Brabant", "Noord-Holland", "Overijssel", "Rijnland", "Zeeland"))

abortus$onderwijsniveau <- factor(x =
                                abortus$Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.,
                                levels = c("Geen diploma ", "Middelbare school", "MBO", "HBO", "Universiteit"),
                                labels = c("Geen diploma", "Middelbare school", "MBO", "HBO", "Universiteit"))

# Voor de vraag: Indien u bij de vorige vraag nee heeft ingevuld...
abortus$sals_nee <- factor(x =
                          abortus$Indien.u.bij.de.vorige.vraag..Nee..heeft.ingevuld...Vindt.u.dat.er.uitz.,
                          levels = c("Ja", "Nee", "Niet van toepassing"),
                          labels = c("Ja", "Nee", "Niet van toepassing"))

abortus.test <- data.frame(abortus$geloof)
```

```

abortus$gezondheid_vrouw <- factor(x = abortus$Vindt.u.dat.er.voor.abortus.gekozen.mag.worden.als.de.gezondheid.van.de.vrouw.is,
                                   levels = c("Ja", "Nee", "Weet ik niet"),
                                   labels = c("Ja", "Nee", "Weet ik niet"))

abortus$gevolg <- factor(x = abortus$De.zwangerschap.veroorzaakt.is.door.verkrachting,
                        levels = c("Ja", "Nee"),
                        labels = c("Ja", "Nee"))

abortus$minderjarig <- factor(x = abortus$Vindt.u.dat.een.minderjarig.kind.de.keuze.moet.hebben.om.voor,
                             levels = c("Ja", "Nee", "Weet ik niet"),
                             labels = c("Ja", "Nee", "Weet ik niet"))

abortus$ouders <- factor(x = abortus$Vindt.u.dat.de.ouders.zeggenschap.moeten.hebben.over.of.hun.minderjarig.kind,
                        levels = c("Ja", "Nee", "Weet ik niet"),
                        labels = c("Ja", "Nee", "Weet ik niet"))

abortus$echo <- factor(x = abortus$Moet.het.wettelijk.verplicht.worden.voor.vrouwen.die.een.abortus.wil,
                      levels = c("Ja", "Nee", "Weet ik niet"),
                      labels = c("Ja", "Nee", "Weet ik niet"))

abortus$bedenktijd <- factor(x = abortus$Vindt.u.dat.de.wettelijke.bedenktijd.moet.terugkomen.,
                             levels = c("Ja", "Nee", "Weet ik niet"),
                             labels = c("Ja", "Nee", "Weet ik niet"))

abortus$adoptie <- factor(x = abortus$Vindt.u.dat.adoptie.een.alternatief.is.voor.abortus.,
                         levels = c("Ja", "Nee", "Weet ik niet"),
                         labels = c("Ja", "Nee", "Weet ik niet"))

abortus$vader <- factor(x = abortus$Moet.de.biologische.vader.zeggenschap.hebben.over.de.beslissing.voor,
                      levels = c("Ja", "Nee", "Weet ik niet"),
                      labels = c("Ja", "Nee", "Weet ik niet"))

# Vragen met een volgorde:
# namen van de variabelen eindigen met "_o" om aan te duiden dat er een volgorde in zit (order)

abortus$levensgevaar_vrouw_o <- factor(x = abortus$Het.leven.van.de.vrouw.in.gevaar.is,
                                       levels = c("zeer oneens", "oneens", "beetje oneens", "neutraal", "heel oneens"),
                                       labels = c("zeer oneens", "oneens", "beetje oneens", "neutraal", "heel oneens"),
                                       order = TRUE)

abortus$geestelijke_handicap_o <- factor(x = abortus$Het.kind.met.een.geestelijke.handicap.geboren.kan.,
                                       levels = c("zeer oneens", "oneens", "beetje oneens", "neutraal", "heel oneens"),
                                       labels = c("zeer oneens", "oneens", "beetje oneens", "neutraal", "heel oneens"),
                                       order = TRUE)

abortus$levensbedreigende_ziekte_o <- factor(x = abortus$Het.kind.met.een.levensbedreigende.ziekte.geboren.kan.,
                                             levels = c("zeer oneens", "oneens", "beetje oneens", "neutraal", "heel oneens"),
                                             labels = c("zeer oneens", "oneens", "beetje oneens", "neutraal", "heel oneens"),
                                             order = TRUE)

abortus$veroorzaakt_o <- factor(x = abortus$De.zwangerschap.veroorzaakt.is.door.verkrachting,
                               levels = c("zeer oneens", "oneens", "beetje oneens", "neutraal", "beetje oneens"),
                               labels = c("zeer oneens", "oneens", "beetje oneens", "neutraal", "beetje oneens"),
                               order = TRUE)

```

```

labels = c("zeer oneens", "oneens", "beetje oneens", "neutraal", "beetje eens", "eens", "zeer eens")
order = TRUE)

abortus$niet_gespecificeerd_o <- factor(x = abortus$De.vrouw.het.kind.niet.wil.houden.om.niet.gespecificeerd_o,
levels = c("zeer oneens", "oneens", "beetje oneens", "neutraal", "beetje eens", "eens", "zeer eens"),
labels = c("zeer oneens", "oneens", "beetje oneens", "neutraal", "beetje eens", "eens", "zeer eens"),
order = TRUE)

abortus$financiële_beperkingen <- factor(x = abortus$De.vrouw.niet.in.staat.is.voor.het.kind.te.zorgen.financiële_beperkingen,
levels = c("zeer oneens", "oneens", "beetje oneens", "neutraal", "beetje eens", "eens", "zeer eens"),
labels = c("zeer oneens", "oneens", "beetje oneens", "neutraal", "beetje eens", "eens", "zeer eens"),
order = TRUE)

df <- data.frame(abortus)

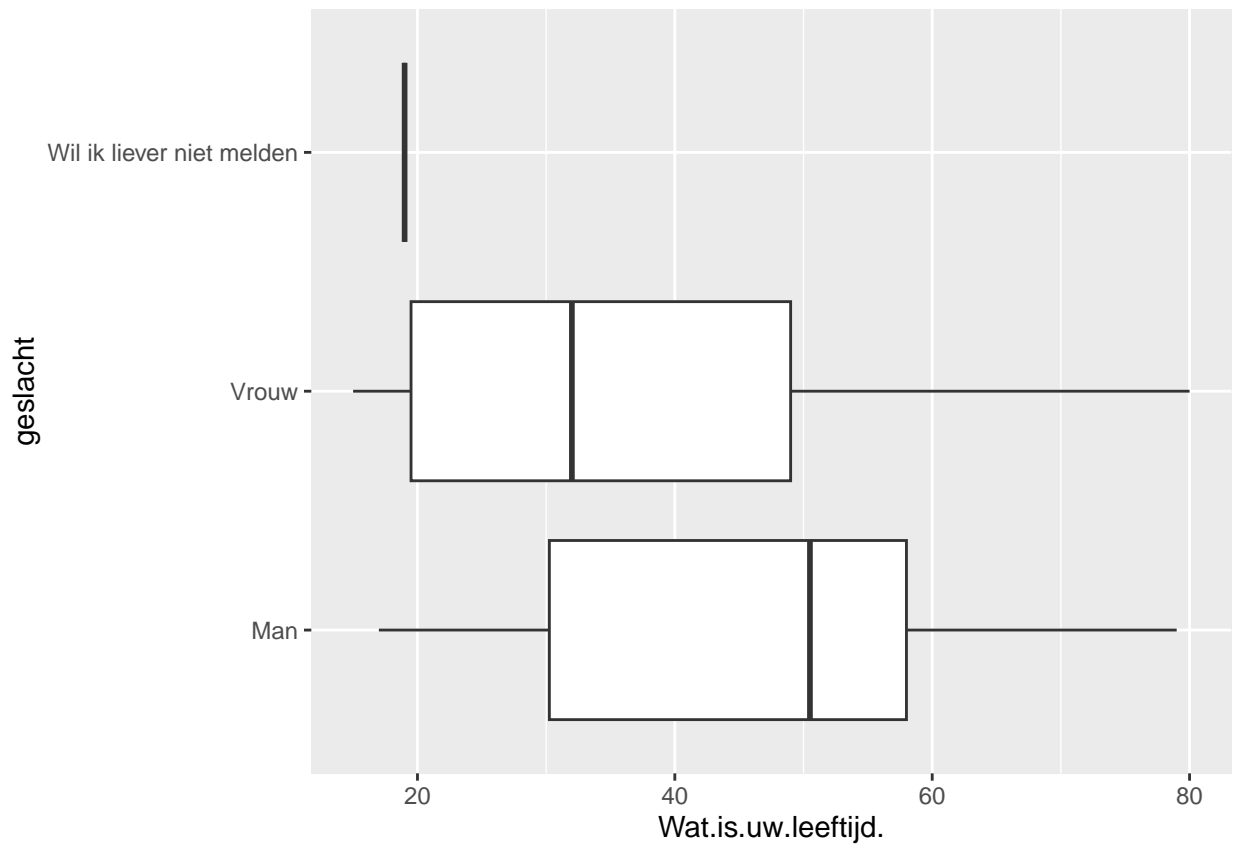
```

Niet alle verkregen data heeft een volgorde, zo is het ene geslacht niet beter dan het andere en “Ja” niet beter of slechter dan “Nee”. Dit is de reden dat alleen de vragen met de volgende antwoorden georderd zijn: “zeer oneens”, “oneens”, “beetje oneens”, “neutraal”, “beetje eens”, “eens”, “zeer eens” In deze antwoorden zit wel een volgorde.

```

ggplot(df, aes(x= Wat.is.uw.leeftijd., y = geslacht)) +
  geom_boxplot()

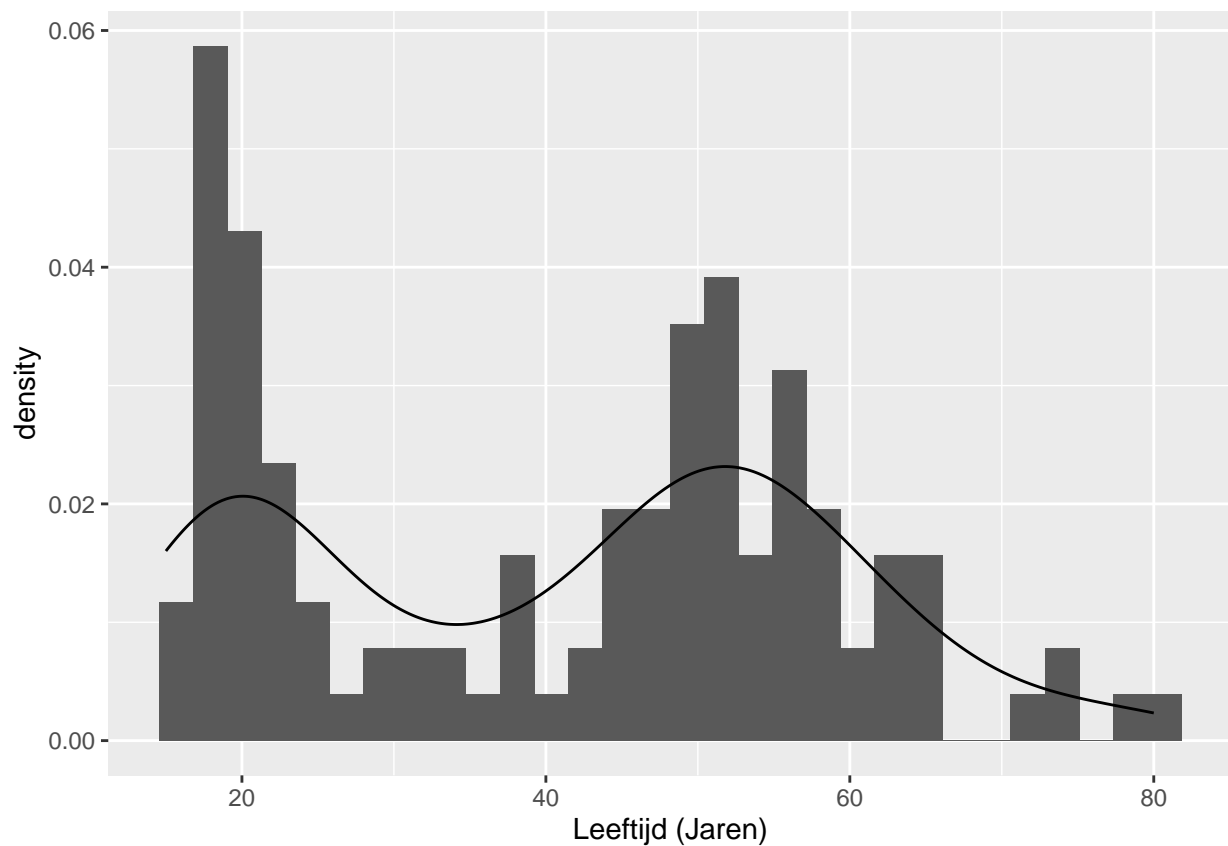
```



```
xlab("Leeftijd (Jaren)")
```

```
## $x  
## [1] "Leeftijd (Jaren)"  
##  
## attr("class")  
## [1] "labels"
```

```
# verdeling van de leeftijden:  
ggplot(df, aes(x= Wat.is.uw.leeftijd.)) +  
  geom_histogram(aes(y = ..density..)) +  
  geom_density() +  
  xlab("Leeftijd (Jaren)")
```



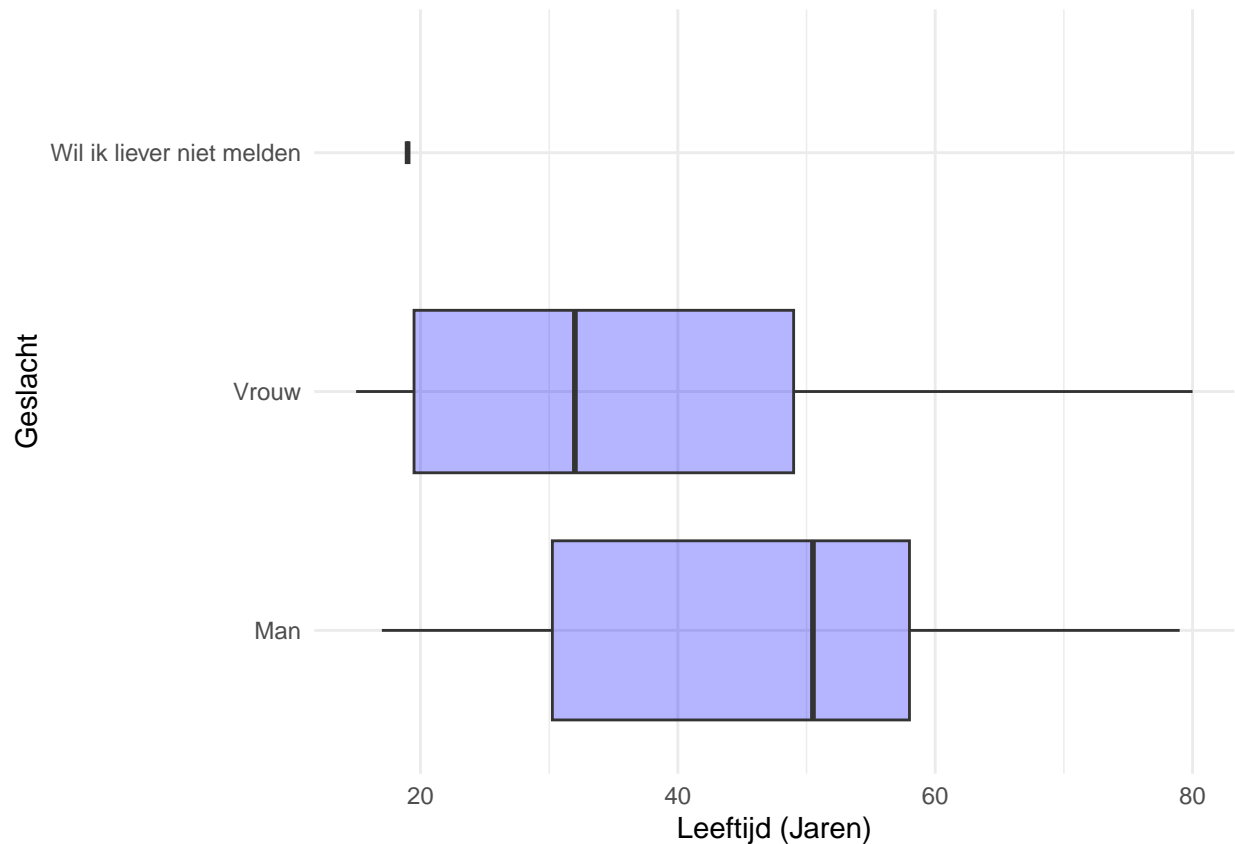
Zoals op het bovenstaande figuur zichtbaar is, is de leeftijddata niet normaal verdeeld maar bimodaal. Er zit een piek bij ~50 jaar en een piek bij ~20 jaar. Deze verdeling van de leeftijd data kan mogelijk verklaard worden door dat de leeftijdsgenoten van de makers van dit artikel rondt de 20 jaar oud zijn en hun ouders in de 50 jaar oud zijn, dit zijn de twee pieken in de leeftijdsverdeling.

```
# test  
ggplot(df, aes(x= df$Wat.is.uw.leeftijd., y = df$geslacht)) +  
  geom_boxplot(fill = "blue", alpha = 0.3, varwidth = T) +  
  xlab("Leeftijd (Jaren)") +  
  ylab("Geslacht") +  
  theme_minimal()
```



```
## Warning: Use of 'df$Wat.is.uw.leeftijd.' is discouraged.
## i Use 'Wat.is.uw.leeftijd.' instead.

## Warning: Use of 'df$geslacht' is discouraged.
## i Use 'geslacht' instead.
```



Voor de bovenstaande grafiek is de $\alpha=0.03$ waarde gekozen voor esthetische doeleinden, niet omdat er overlap kan zijn. Het plot weergeeft de verdeling van de leeftijden van de respondenten van de enquête.

Overzicht: abortus toestaan of niet:

```
table(abortus$eigen_keuze_abortus) # voor de vraag: Vindt u dat iemand zelf moet kunnen kiezen of die e
##
## Wel abortus toestaan Geen abortus toestaan
##          110              4
```

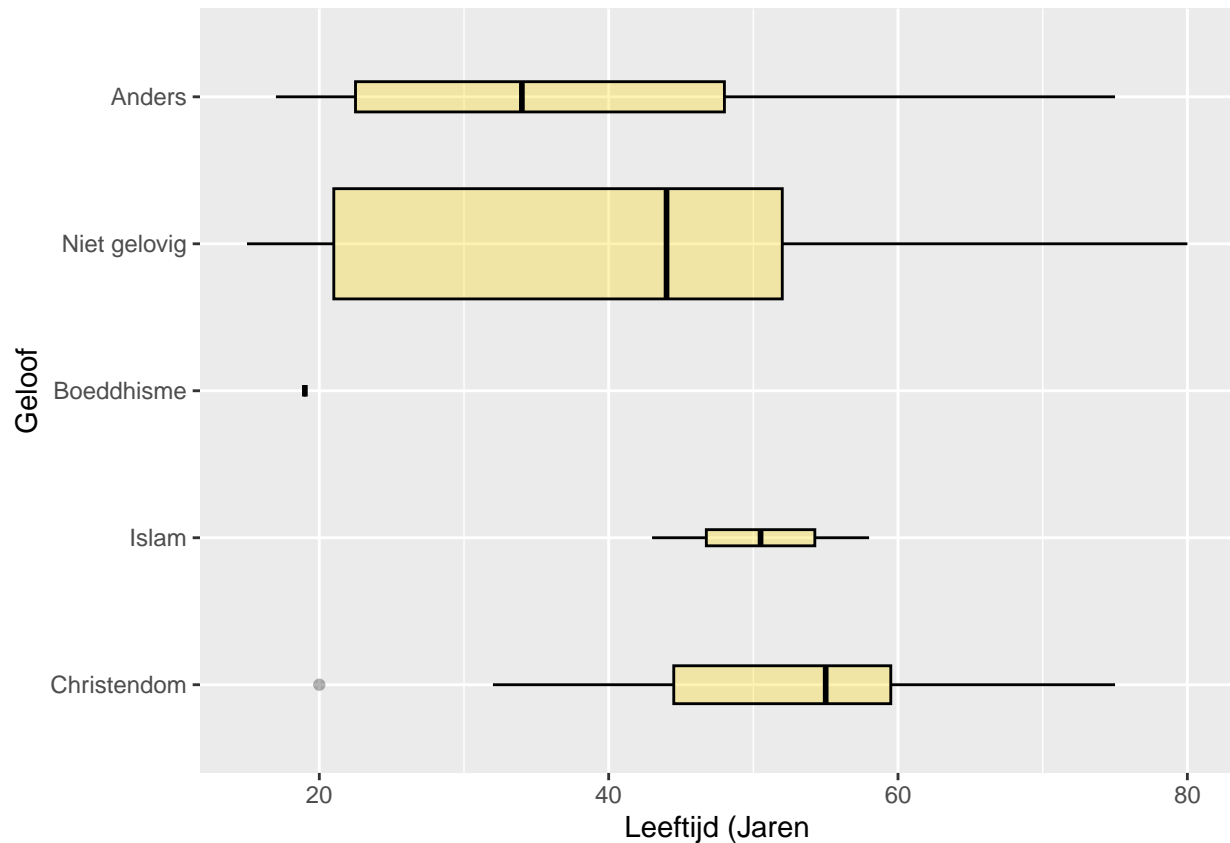
De bovenstaande tabel weergeeft het totaal aantal mensen die voor de twee opties gekozen hebben en wat hun keuze was. De meerderheid kiest met afstand voor het toestaan van de abortus.

```
# Grafiek met de verschillende geloven en de leeftijdsverdeling die er bij hoort.
ggplot(df, aes(y=abortus$gelooft,
               x= abortus$Wat.is.uw.leeftijd.)) +
```

```
geom_boxplot(color = "black", fill = "gold", alpha = 0.3, varwidth=T) +
  xlab("Leeftijd (Jaren)" +
  ylab("Geloof")
```

```
## Warning: Use of 'abortus$Wat.is.uw.leeftijd.' is discouraged.
## i Use 'Wat.is.uw.leeftijd.' instead.
```

```
## Warning: Use of 'abortus$geloof' is discouraged.
## i Use 'geloof' instead.
```

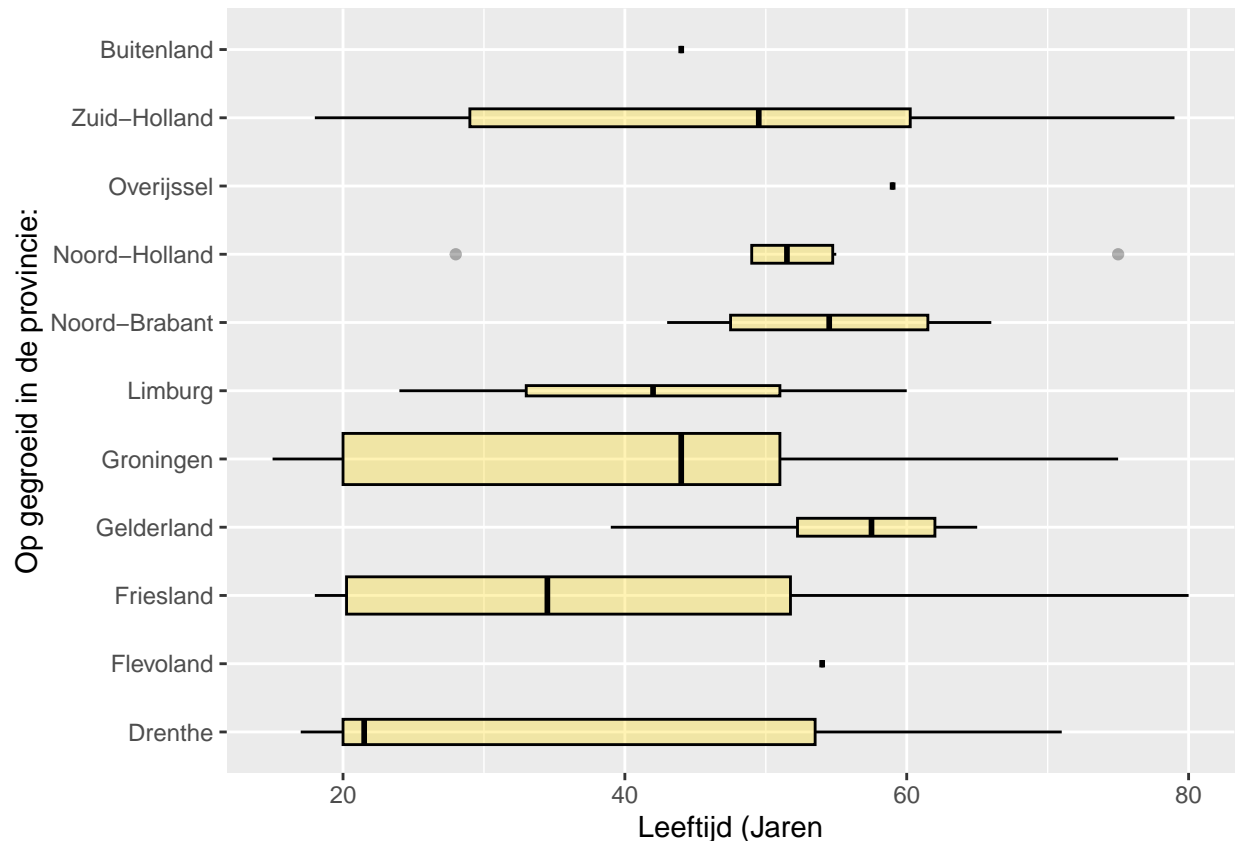


Verdeling van de leeftijd van de mensen die de enquête hebben ingevuld en hun ideologie. De hoogte van het boxplot laat het verschil in het aantal mensen zien die bij deze ideologie horen.

```
# Grafiek met leeftijden uit verschillende provincies.
ggplot(df, aes(y=abortus$afkomst,
  x= abortus$Wat.is.uw.leeftijd.)) +
  geom_boxplot(color = "black", fill = "gold", alpha = 0.3, varwidth = T) +
  xlab("Leeftijd (Jaren)" +
  ylab("Op gegroeid in de provincie:")
```

```
## Warning: Use of 'abortus$Wat.is.uw.leeftijd.' is discouraged.
## i Use 'Wat.is.uw.leeftijd.' instead.
```

```
## Warning: Use of 'abortus$afkomst' is discouraged.
## i Use 'afkomst' instead.
```



De verdeling van de leeftijden en de provincies waar deze mensen vandaan komen wordt weergegeven op deze grafiek. Niet elke provincie heeft genoeg datapunten om een boxplot van te maken.

```
# De invultijd per leeftijd:
begin_tijd <- as.POSIXct(x = (c(abortus$Begintijd)), format = "%m-%d-%y %H:%M:%S") # startpunt van invultijd

#
eind_tijd <- as.POSIXct(x = (c(abortus$Tijd.van.voltoeien)), format = "%m-%d-%y %H:%M:%S") # eindpunt van invultijd

tijd_vector <- c()
# Om de totale tijd uit te rekenen die de mensen nodig hadden om het formulier in te vullen.
# Het verschil tussen de eindtijd en de begintijd.
for (x in seq_along(abortus$Begintijd)) { # Loop door de begintijden heen..
  #print(paste0("Begin tijd: ", begin_tijd[x]))
  #print(paste0("eind tijd: ", eind_tijd[x]))
  tijd_vector[x] <- as.numeric((eind_tijd[x] - begin_tijd[x])) # Voeg toe aan tijdsvector: het verschil
}

# Dit kan veel makkelijker:
tijd_vector <- eind_tijd - begin_tijd
```

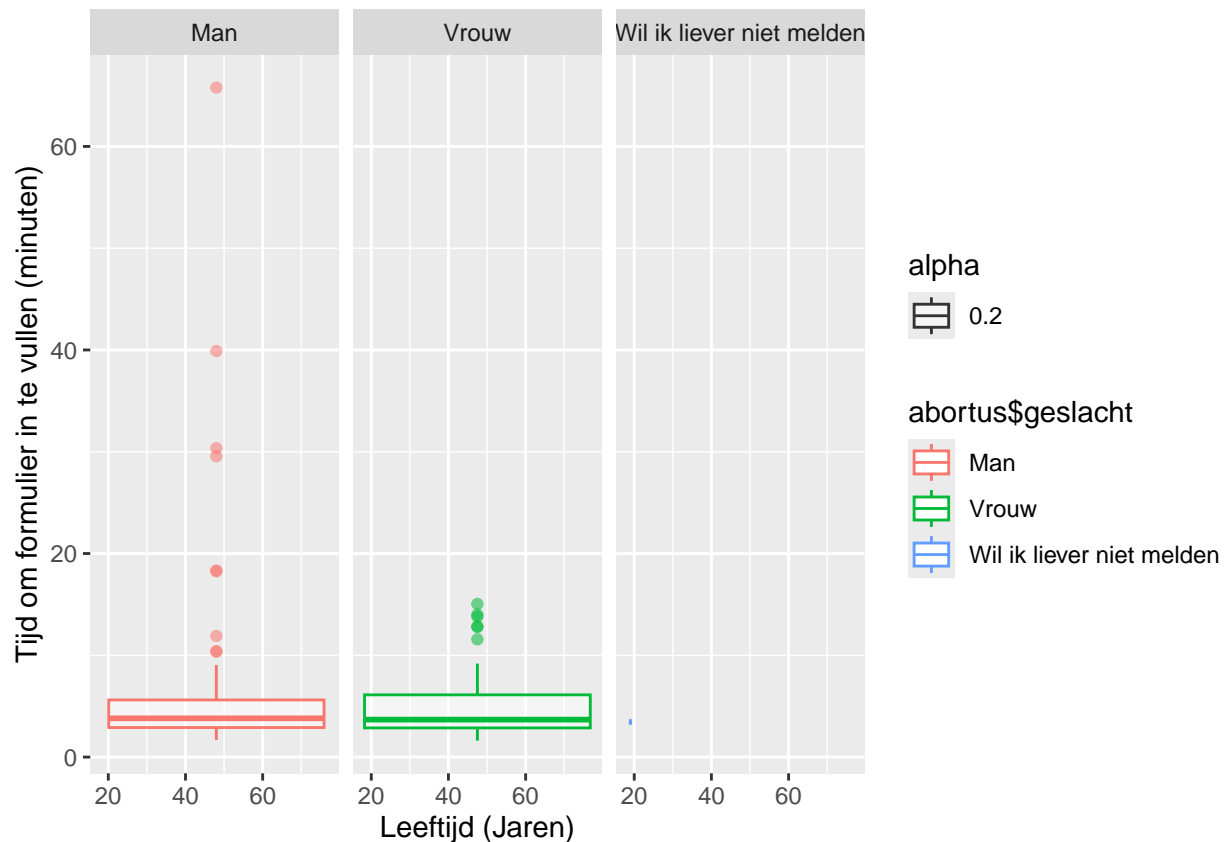
De invultijd is niet belangrijk of relevant voor de onderzoeksvraag maar als test uitgevoerd om meer over R te leren.

Op 11 mei 2024 grafiek gemaakt van leeftijd / tijd nodig om de enquête in te vullen. 12-05-2024: Met facet_wrap() onderscheid gemaakt tussen mannen en vrouwen. (twee plots)

```
tijd_vector <- data.frame(tijd_vector)

ggplot(tijd_vector, aes(y=tijd_vector,
                        x= abortus$Wat.is.uw.leeftijd.)) +
  geom_boxplot(aes(color = abortus$geslacht, alpha=0.2)) +
  facet_wrap(~abortus$geslacht) +
  xlab("Leeftijd (Jaren)") +
  ylab("Tijd om formulier in te vullen (minuten)") #+
```

```
## Don't know how to automatically pick scale for object of type <difftime>.
## Defaulting to continuous.
```



```
#coord_flip()
```

Bovenstaande grafiek laat de relatie tussen leeftijd en tijd nodig om het formulier in te vullen zien. Dit wordt voor mannen en vrouwen gedaan. (links mannen, midden vrouwen en rechts ongespecificeerd.) Er was maar een persoon dat ongespecificeerd opgaf dus kan daar geen boxplot van gemaakt worden maar wordt dit weergegeven met een enkele stip. (links onder in de grafiek)

```
data_frame_invultijd <- data.frame(tijd_vector, abortus$geslacht, abortus$Wat.is.uw.leeftijd.)

# Voor de tijd die de mannen er over doen:
invultijd_mannen <- data_frame_invultijd %>%
  filter(abortus.geslacht == "Man")
```

```
# Voor de tijd die de vrouwen er over doen:
invultijd_vrouwen <- data_frame_invultijd %>%
  filter(abortus.geslacht == "Vrouw")

# Voor de tijd die de "ongespecificeerde geslacht" mensen er over doen:
invultijd_anders <- data_frame_invultijd %>%
  filter(abortus.geslacht == "Wil ik liever niet melden")
```

Gemiddeld doen de mannen er 7.11586021505376 minuten over om de enquête in te vullen. Dit ligt tussen de 2.90416666666667 (25%) en 5.61666666666667 (75%)

Gemiddeld doen de vrouwen er 5.14934640522876 minuten over om de enquête in te vullen. Dit ligt tussen de 2.86666666666667 (25%) en 6.11666666666667 (75%)

Gemiddeld doen de “Geslacht niet opgegeven mensen” er 3.45 minuten over om de enquête in te vullen. Dit ligt tussen de 3.45 (25%) en 3.45 (75%)

Circular Packing test 1:

Test met data van het internet om het hierarchische data format beter te begrijpen. Deze test werkt niet dus is eval=FALSE.

```
# Load dataset from github
data <- read.table("https://raw.githubusercontent.com/holtzy/data_to_viz/master/Example_dataset/11_SeVCo")
data[ which(data$value==1), "value"] <- 1
colnames(data) <- c("Continent", "Region", "Country", "Pop")

# Remove a few problematic lines
data <- data %>% filter(Continent!="") %>% droplevels()

# Change the format. This use the data.tree library. This library needs a column that looks like root/g
library(data.tree)
data$pathString <- paste("world", data$Continent, data$Region, data$Country, sep = "/")
population <- as.Node(data)
data <- data.frame(
  root=rep("root", 15),
  group=c(rep("group A",5), rep("group B",5), rep("group C",5)),
  subgroup= rep(letters[1:5], each=3),
  subsubgroup=rep(letters[1:3], 5),
  value=sample(seq(1:15), 15)
)

data$pathString <- paste("world", data$group, data$subgroup, data$subsubgroup, sep = "/")
population <- as.Node(data)
population

#graph <- tbl_graph(flare$vertices, flare$edges)
set.seed(1)
ggraph(graph, 'circlepack', weight = size) +
  geom_node_circle(aes(fill = depth), size = 0.25, n = 50) +
  coord_fixed()
```

15-05-2024: Om te achterhalen of de leeftijden van de mensen die de enquête ingevuld hebben normaal verdeeld zijn is de Shapiro Wilk test gebruikt.

Shapiro Wilk test:

Deze test wordt uitgevoerd om vast te stellen of leeftijden van de mensen die de enquête ingevuld hebben normaal verdeeld is.

```
shapiro_test <- shapiro.test(abortus$Wat.is.uw.leeftijd.)
shapiro_test
```

```
##
##  Shapiro-Wilk normality test
##
## data:  abortus$Wat.is.uw.leeftijd.
## W = 0.9217, p-value = 5.044e-06
```

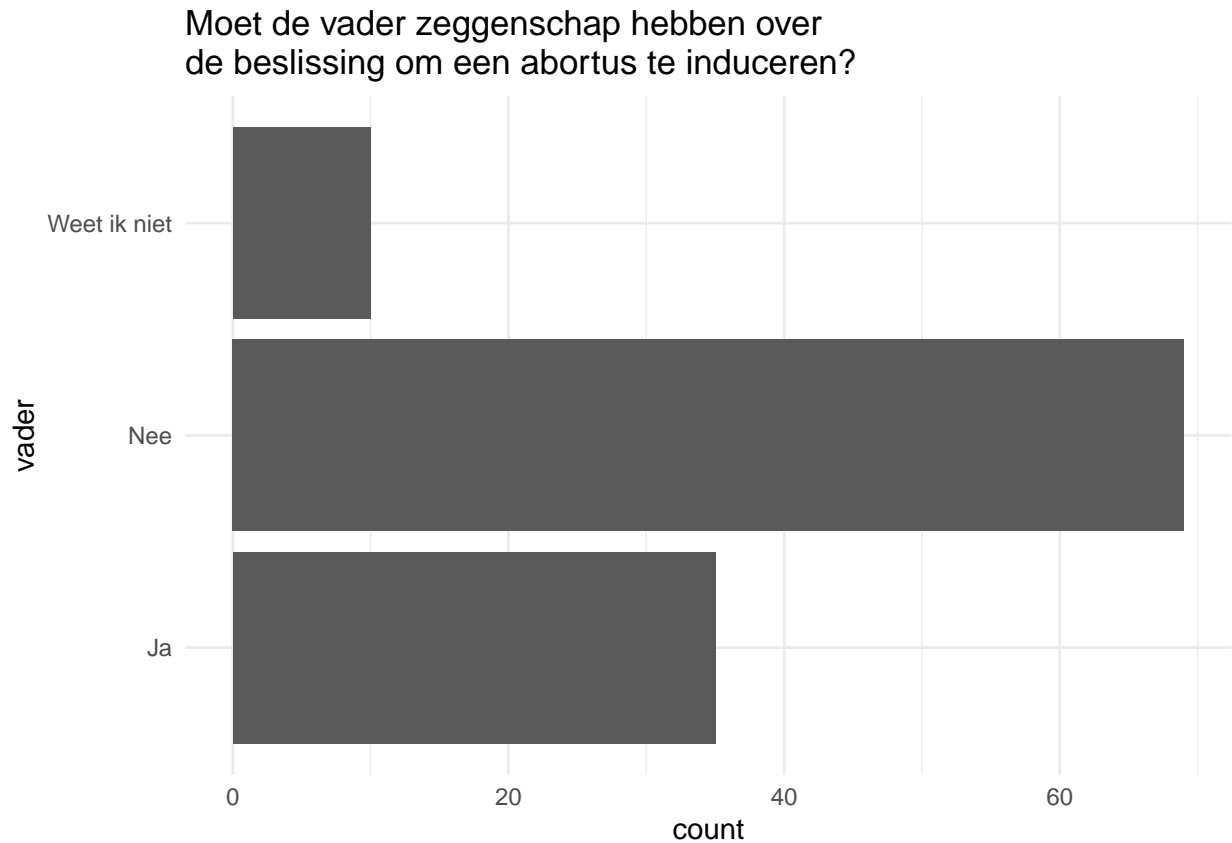
De p-waarde is: 5.044×10^{-6} (5.0444065×10^{-6}) dat betekent dat er wel een significant verschil is tussen een normale verdeling en deze leeftijd data. Omdat de data niet normaal verdeeld is zal een niet-parametrische analyse nodig zijn. Voorbeelden van deze non parametrische zijn: de Kruskal-Wallis test en de Chi-square en de two way z test.

```
proportie_man_vader <- df %>% select(vader) %>% filter(df$geslacht == "Man")
proportie_man_vader_01 <- ifelse(proportie_man_vader == "Ja" , 0,1)

hi <- prop.table(proportie_man_vader_01)

#proportie_vrouw_vader <-

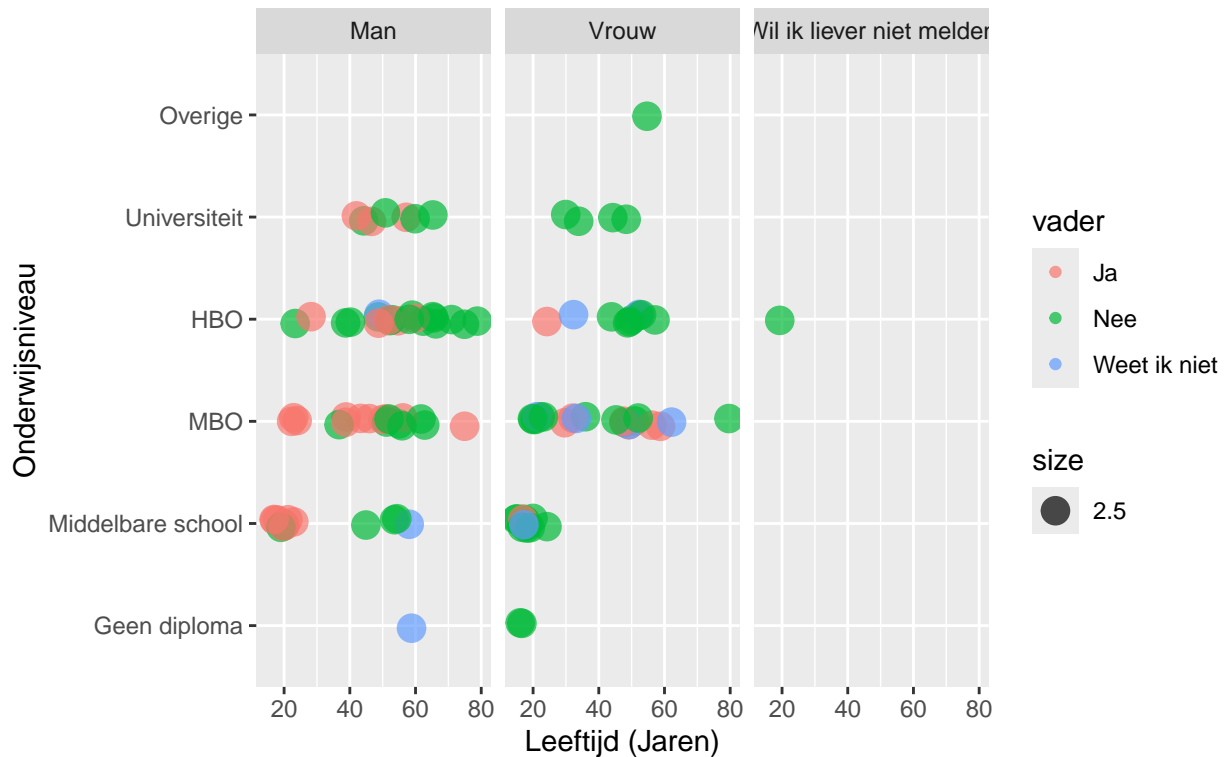
ggplot(df, aes( y = vader)) +
  geom_bar(aes()) +
  labs(color = "test") +
  #scale_color_manual(name = "Cylinders")
# scale_fill_discrete(name = "test") +
  ggtitle("Moet de vader zeggenschap hebben over \nde beslissing om een abortus te induceren? ") +
  theme_minimal()
```



Omdat er niet even veel mannen als vrouwen zijn is het beter om de stemmen (ja/nee) in proporties/percentages te laten zien i.p.v. absolute aantallen. Dit moet later nog gedaan worden..

```
ggplot(df, aes(x= Wat.is.uw.leeftijd., y = onderwijsniveau)) +
  # geom_point(aes(color = vader, size = 2.5), alpha = 0.5) +
  facet_wrap(~geslacht) +
  labs(" ") +
  xlab("Leeftijd (Jaren)") +
  ylab("Onderwijsniveau") +
  geom_jitter(aes(color = vader, size = 2.5), alpha = 0.7, height = 0.05) +
  ggtitle("Moet de vader zeggenschap hebben over \nde beslissing om een abortus te induceren?")
```

Moet de vader zeggenschap hebben over de beslissing om een abortus te induceren?



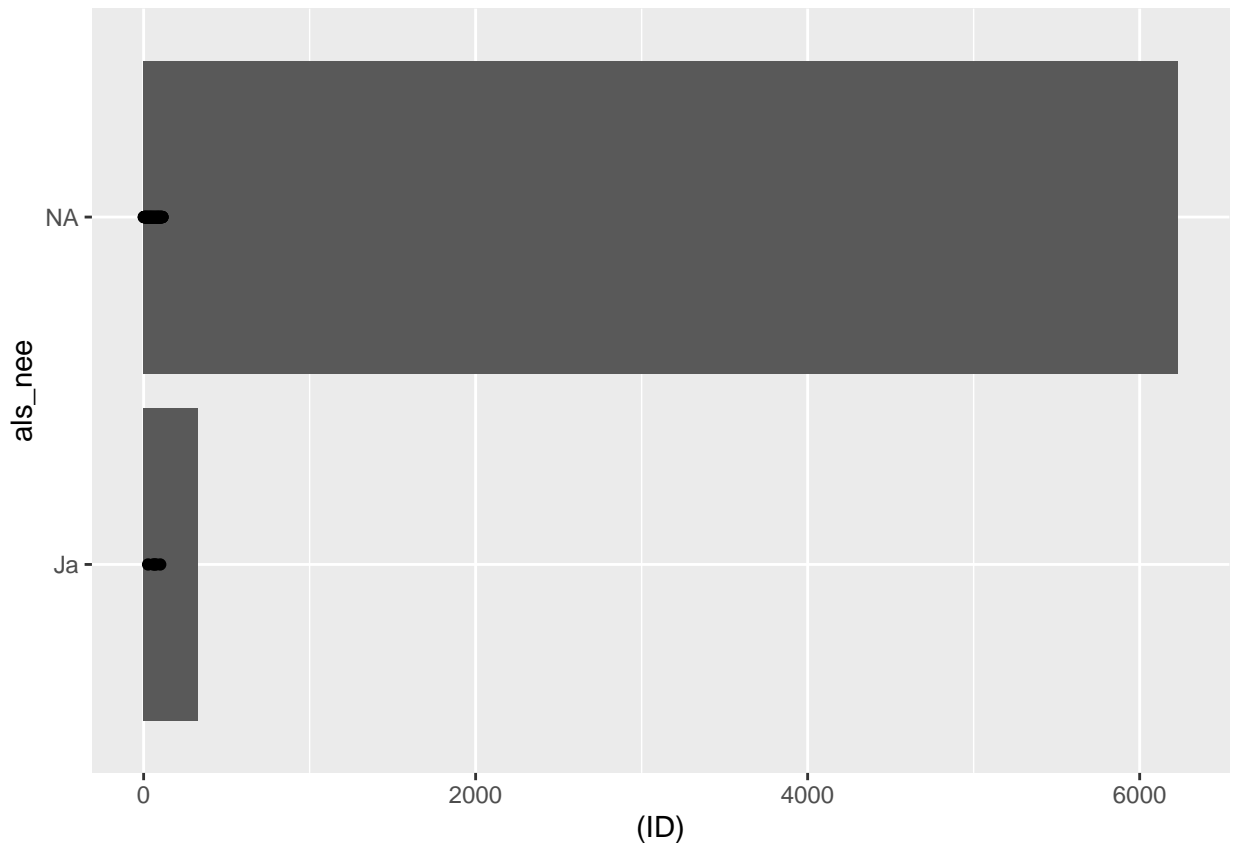
Het opleidingsniveau kan geen ge-orderde factor zijn omdat “Middelbare school” kan leiden tot meerdere andere niveaus. Ook kan “Geen diploma” betekenen dat het persoon nog te jong is om deze te hebben maar later wel in een van de andere categorieën zal horen.

15-05-2024: Bovenstaande grafiek: Moet de vader zeggenschap hebben over de abortus: Ja, Nee of geen mening. Het geslacht van het persoon wordt met de twee kleuren blauw en rood aangegeven en de leeftijd met de grootte van het bolletje. (zie legenda voor schaalverdeling)

20-05-2024:

Voor vraag: Als u geen voorstander bent voor de vrijheid om voor een abortus te kiezen, zijn er geval

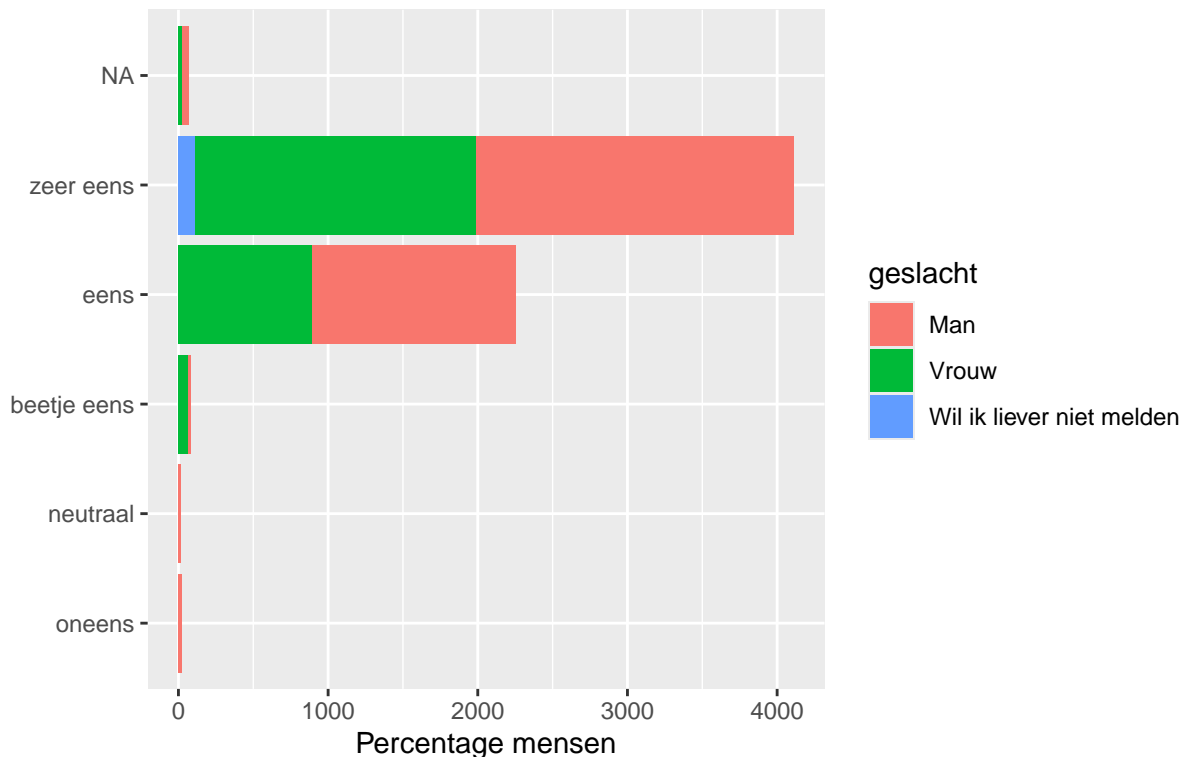
```
ggplot(data = df, mapping = aes(x = (ID), y = als_nee)) +
  geom_bar(stat = "identity") +
  geom_point()
```

Het bovenstaande figuur weergeeft het aantal mensen die Ja/Nee gekozen hebben, dit plot moet later nog aangepast worden zodat er rekening gehouden wordt met het verschil in het aantal respondenten. (de x-as klopt nog niet)

```
# Voor de vraag: Vindt u dat er voor abortusgekozen mag worden als de gezondheid van de vrouw door de zwang
library(ggthemes)
ggplot(data = df, mapping = aes(x = c(1:114), y = levensgevaar_vrouw_o, fill = geslacht)) +
  geom_bar(stat = "identity") +
  geom_text(aes(label = (" "))) +
  ggtitle("Vindt u dat er voor abortusgekozen mag worden als de gezondheid van \nde vrouw door de zwang")
  xlab("Percentage mensen") + # nog aanpassen naar procent.
  ylab(" ") # Geen ylabel nodig omdat er al een beschrijving bij de waardes staan.
```

Vindt u dat er voor abortus gekozen mag worden als de gezondheid van de vrouw door de zwangerschap in gevaar kan komen?



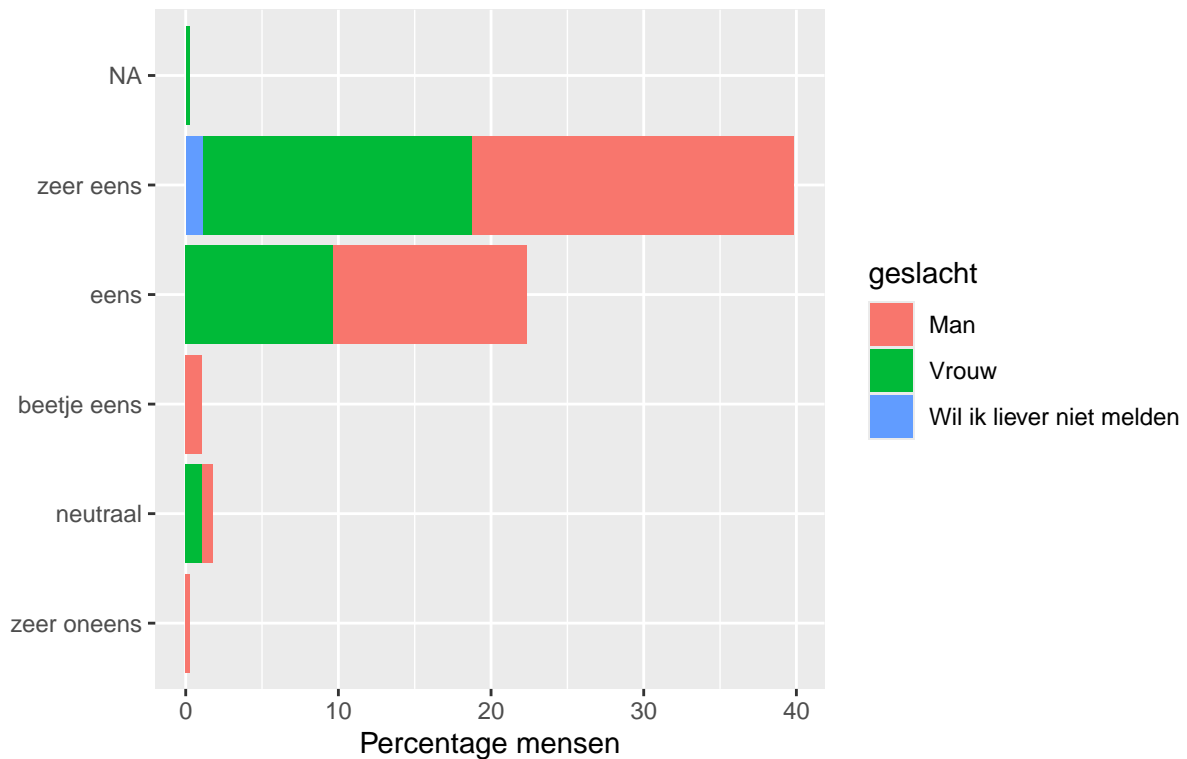
```
# labels beter maken.
# To-do: percentage verbeteren, huidige waarden kloppen nog niet.
```

Dit figuur laat de meningen van de respondenten zien op een schaal van zeer eens tot zeer oneens. Er was hier niemand die voor “zeer oneens” gekozen heeft, dat is waarom deze categorie niet weergegeven wordt.

```
# Voor de vraag: Mag abortus toegestaan zijn als de zwangerschap resultaat is van verkrachting?
```

```
ggplot(data = df, mapping = aes(x = (ID/100), y = veroorzaakt_o, fill = geslacht)) + # geeft procent we
  geom_bar(stat = "identity") +
  ggtitle("Mag abortus als de zwangerschap\nresultaat is van verkrachting?") +
  xlab("Percentage mensen") + # verifiëren of dit wel klopt...
  ylab(" ")
```

Mag abortus als de zwangerschap resultaat is van verkrachting?



Een zelfde plot voor de vraag: Mag abortus toegestaan zijn als de zwangerschap resultaat is van verkrachting? is hierboven afgebeeld.

```
ggplot(data = df, mapping = aes(x = (c(1:114)), y = levensgevaar_vrouw_o)) + # geeft procent weer.
  geom_bar(stat = "identity", aes(fill = Wat.is.uw.leeftijd.)) +
  facet_wrap(~geslacht) +
  ggtitle("Mag abortus als het leven van de vrouw in gevaar komt tijdens\nde zwangerschap?") +
  xlab("Percentage mensen") + # verifiëren of dit wel klopt...
  ylab(" ")
```

Mag abortus als het leven van de vrouw in gevaar komt tijdens de zwangerschap?



To-do: sorteer voorl leeftijd gradient.

Met `facet_wrap()` is hier per geslacht weergegeven wat de meningen van de respondenten zijn. De leeftijdschaal moet aangepast worden omdat het verschil in leeftijd nu per individu weergegeven is.

25-05-2024: # Donut grafiek: Op de onderstaande grafieken wordt het aantal mensen per provincie gevisualiseerd.

```
#library(ggplot2)

# Maak tabel met aantal mensen per provincie.
x <- table(abortus$afkomst)
x <- sort(x, decreasing = T)
provincie_aantal <- data.frame(x)

# Reken de percentages uit:
provincie_aantal$fraction = provincie_aantal$Freq / sum(provincie_aantal$Freq)
provincie_aantal$percentage = round(provincie_aantal$fraction * 100, digits = 1)
provincie_aantal$ymax = cumsum(provincie_aantal$fraction)
provincie_aantal$ymin = c(0, head(provincie_aantal$ymax, n=-1))
provincie_aantal$labelp = (provincie_aantal$ymax+provincie_aantal$ymin)/2
provincie_aantal$label = (paste(provincie_aantal$percentage, "% ", provincie_aantal$Var1))

# Verander de naam van de provincie colom van Var1 naar "provincies":
provincie_aantal <- provincie_aantal %>%
```

```

rename(provincies = Var1)

# Maak het grafiekje:
ggplot(provincie_aantal, aes(ymax=ymax, ymin=ymin, xmax=13, xmin=3)) +
  geom_rect(color = "black", aes(fill=provincies)) +
  geom_label_repel(x = 13, aes(y=labelp, label= label), size=3, max.overlaps = 11, min.segment.length=
# geom_label(x = 13, aes(y=labelp, label= label), size=4)+
  ggtitle("Enquete deelnemers, verdeling over de provincies:") +
  coord_polar(theta="y") +
  xlim(c(-13, 13)) +
  xlab(" ") +
  ylab(" ")

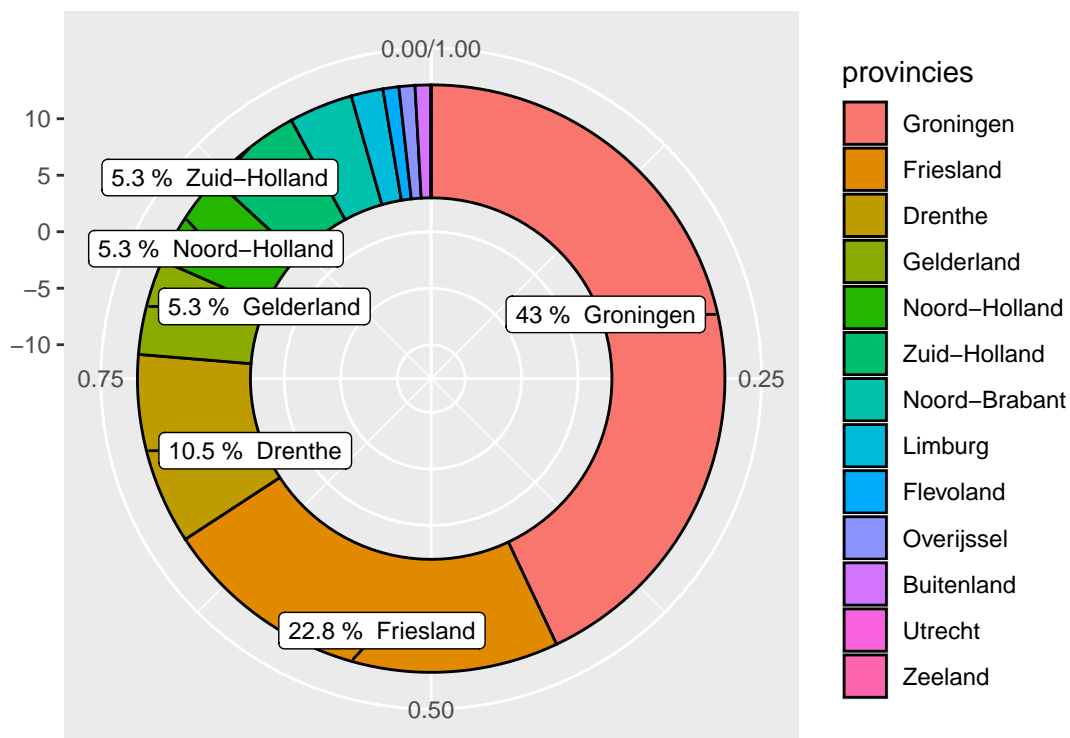
```

```

## Warning: ggrepel: 7 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps

```

Enquete deelnemers, verdeling over de provincies:



De bovenstaande grafiek is voor de presentatie over dontut grafieken bedoeld en visualiseert de verdeling van de respondenten over Nederland, per provincie. ### Tweede variant (stacked-barchart)

```

#library(ggrepel)

# Maak tabel met aantal mensen per provincie.
x <- table(abortus$afkomst)
x <- sort(x, decreasing = F)

```

```

provincie_aantal <- data.frame(x)

# Reken de percentages uit:
provincie_aantal$fraction = provincie_aantal$Freq / sum(provincie_aantal$Freq)
provincie_aantal$percentage = round(provincie_aantal$fraction *100, digits =1)
provincie_aantal$ymax = cumsum(provincie_aantal$fraction)
provincie_aantal$ymin = c(0, head(provincie_aantal$ymax, n=-1))
provincie_aantal$labelp = (provincie_aantal$ymax+provincie_aantal$ymin)/2
provincie_aantal$label = (paste(provincie_aantal$percentage, "% ", provincie_aantal$Var1))

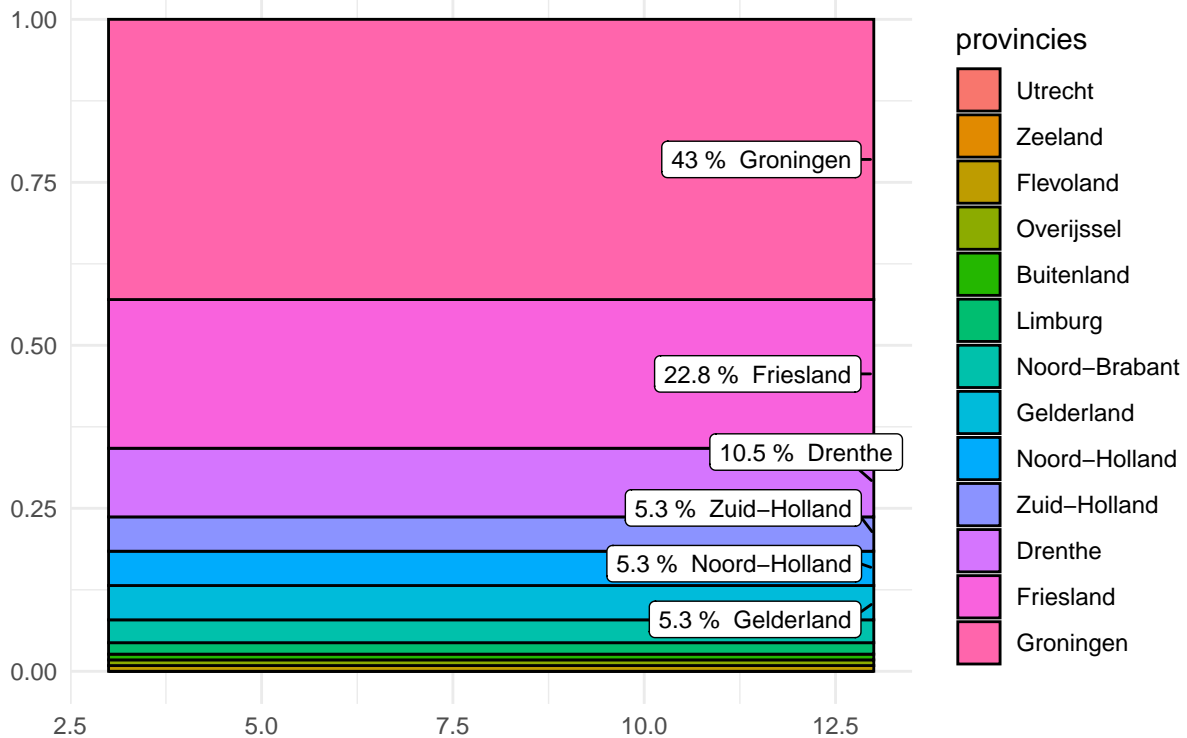
# Verander de naam van de provincie colom van Var1 naar "provincies":
renamed_provincie_aantal <- provincie_aantal %>%
  rename(provincies = Var1)

# Maak het grafiekje:
ggplot(renamed_provincie_aantal, aes(ymax=ymax, ymin=ymin, xmax=13, xmin=3)) +
  geom_rect(color = "black",aes(fill=provincies)) +
  geom_label_repel(x = 13, aes(y=labelp, label= label), size=3, max.overlaps = 10, min.segment.length=0.5)
# geom_label( x = 13, aes(y=labelp, label= label), size=4)+
  ggtitle("Enquete deelnemers, verdeling over de provincies:") +
  theme_minimal() +
  xlab(" ") +
  ylab(" ")

## Warning: ggrepel: 7 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps

```

Enquete deelnemers, verdeling over de provincies:



De bovenstaande grafiek geeft weer hoe een donut grafiek opgebouwd is, het zijn `geom_rect()` blokken die op elkaar gestapeld zijn, de lijn waar een nieuw blok begint wordt bepaald met de cumulatieve som van de vector met proporties. Om er een donut van te maken zoals in chunk 23 weergegeven wordt moet de lijn: `coord_polar(theta="y")` toegevoegd worden.

Cirkel packing grafiek test

```
# Test 2:

edges <- data.frame(c("test", "test.test2", "test.test2.test3", "test0"))
vertices <- data.frame(c("1", "2", "3"), c("1", "2", "3"))

edgelist <- matrix(list(c("1","2","3"), c("1","2","3")),
                    byrow=TRUE, ncol=2)

edge.list(edgelist)

## $nodes
##   id      label
## 1  1 c("1", "2", "3")
##
## $edges
##   source target
```

```
mygraph <- graph_from_data_frame(edgelist)

library(data.tree)
df9 <- data.frame(name = c("land", "water", "air", "car", "bicycle", "boat", "balloon", "airplane", "helicopter"),
                  parent = c("root", "root", "root", "land", "land", "water", "air", "air", "air", "car"))
tree <- FromDataFrameNetwork(df9)

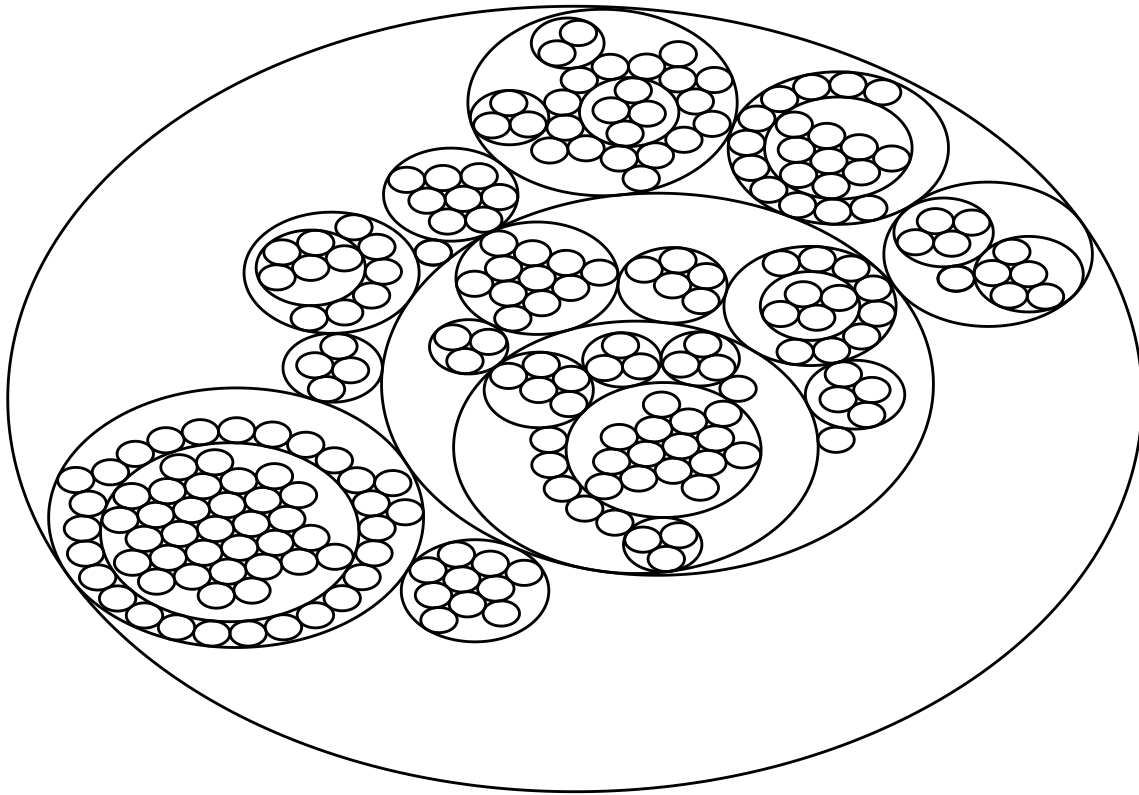
vv <- ToDataFrameTree(tree,
                      level1 = function(x) x$path[2],
                      level2 = function(x) x$path[3],
                      level3 = function(x) x$path[4],
                      level_number = function(x) x$level - 1)[-1,-1])

# Test
#hc <- df %>%
# dist %>%
# hclust
#cc <- tree(c(1:5))
#plot(hc)
```

```
library(ggraph)
library(igraph)
library(tidyverse)
edges <- flare$edges
vertices <- flare$vertices

mygraph <- graph_from_data_frame( edges, vertices=vertices )

ggraph(mygraph, layout = 'circlepack') +
  geom_node_circle() +
  theme_void()
```

De bovenstaande twee grafieken weergeven een visualisatie van hiërarchische data, niet gebaseerd op de data verkregen van de enquête maar van de “flares” dataset die ingebouwd is in R.

Power test: (met leeftijd van de mensen)

17-05-2024: Begin gemaakt met t.test en power test

```
library(dplyr)
library(pwr)
tabel <- table(abortus$geslacht ,abortus$eigen_keuze_abortus) # tabel met mannen en vrouwen en hun keuze
#prop.table(abortus$Wat.is.wv.geslacht..)
tabel2 <- prop.table(tabel)

ja_abortus <- (ifelse(abortus$eigen_keuze_abortus=="Wel abortus toestaan", 1,0)) # 1 voor de mensen die ja zeggen

# Reken de proportie mannen uit die voorstanders van abortus zijn:
man_ja_abortus <- df %>%
  filter(geslacht == "Man")
prop_man_ja_abortus <- prop.table(table(ifelse(df$eigen_keuze_abortus=="Wel abortus toestaan", 1,0)))

# Bereken het zelfde voor de vrouwen:
```

```
vrouw_ja_abortus <- df %>%
  filter(geslacht == "Vrouw")
prop_vrouw_ja_abortus <- prop.table(table(ifelse(df$eigen_keuze_abortus=="Wel abortus toestaan", 1,0)))

# effect sterkte h uitrekenen met de twee ja/nee proporties (tussen mannen en vrouwen)
h <- ES.h(prop_man_ja_abortus[2], prop_man_ja_abortus[1]) # Er werkt iets niet goed...

power1 <- pwr.t.test(sig.level = 0.05, power = 0.8, n = 109 ,alternative = "two.sided") # test, andere h
power2 <- pwr.t.test(d = 0.5,sig.level = 0.05, power = 0.8, alternative = "two.sided") # test, andere h
```

Om met een effect sterkte van 0.5 het verschil tussen de twee geslachten aan te tonen zijn er 64 mensen nodig (van elke groep). En met de 109 mensen die we nu hebben is de effect sterkte: 0.38. Dit is genoeg om een middelgroot verschil waar te nemen.

Chi-square:

En chi-square test voor alle enquête vragen en de invloed die het geslacht op de keuzes heeft. **Vraag:** Vindt u dat iemand zelf moet kunnen kiezen of die een abortus wil? >H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze over het toestaan van de abortus. >H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze over het toestaan van de abortus.

23-05-2024: De H0 en H1 hypothesen worden hier opgesteld voor alle vragen:

Vraag: Vindt u dat er voor abortus gekozen mag worden als de gezondheid van de vrouw door de zwangerschap in gevaar kan komen? >H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of een vrouw een abortus mag hebben wanneer haar leven in gevaar is. >H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of een vrouw een abortus mag hebben wanneer haar leven in gevaar is.

Vraag: Vindt u dat iemand een abortus mag als de zwangerschap een gevolg is van verkrachting? >H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze over het toestaan van een abortus wanneer de zwangerschap gevolg is van verkrachting. >H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze over het toestaan van een abortus wanneer de zwangerschap gevolg is van verkrachting.

Vraag: Vindt u dat een minderjarig kind de keuze moet hebben om voor een abortus te kiezen? >H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of een minderjarig kind de keuze moet hebben om voor een abortus te kiezen. >H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of een minderjarig kind de keuze moet hebben om voor een abortus te kiezen.

Vraag: Vindt u dat de ouders zeggenschap moeten hebben over of hun minderjarige kind wel of niet voor abortus mag kiezen? >H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of de ouders van een minderjarig kind zeggenschap moeten hebben over of hun minderjarige kind wel of niet voor abortus mag kiezen? >H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of de ouders van een minderjarig kind zeggenschap moeten hebben over of hun minderjarige kind wel of niet voor abortus mag kiezen?

Vraag: Moet het wettelijk verplicht worden voor vrouwen die een abortus willen om een echo van de foetus te zien voordat een abortus toegestaan is? >H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of ze vinden dat het wettelijk verplicht worden voor vrouwen die een abortus willen om een echo van de foetus te zien voordat een abortus toegestaan is. >H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of ze vinden dat het wettelijk verplicht worden voor vrouwen die een abortus willen om een echo van de foetus te zien voordat een abortus toegestaan is.

Vraag: Vindt u dat de wettelijke bedenktijd moet terugkomen? >H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of de wettelijke bedenktijd terug moet komen. >H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of de wettelijke bedenktijd terug moet komen.

Vraag: Vindt u dat adoptie een alternatief is voor abortus? >H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of adoptie een alternatief is voor abortus. >H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of adoptie een alternatief is voor abortus.

Vraag: Moet de biologische vader zeggenschap hebben over de beslissing voor een abortus? >H0: Het geslacht van het persoon dat de enquête invult heeft geen invloed op hun keuze of de biologische vader zeggenschap moet hebben over de beslissing voor een abortus. >H1: Het geslacht van het persoon dat de enquête invult heeft invloed op hun keuze of de biologische vader zeggenschap moet hebben over de beslissing voor een abortus.

27-05-2024: Deze zelfde hypotheses zijn ook gebruikt voor de Chi-square test met leeftijd als voorspellende variabele i.p.v. geslacht.

```
#onder21 <- filter(abortus$Wat.is.uw.leeftijd., abortus$Wat.is.uw.leeftijd. < 21)
```

```
#binaire_lijst_keuze_abortus <-
```

```
#ja_abortus <- (ifelse(abortus$eigen_keuze_abortus=="Wel abortus toestaan", 1,0)) # 1 voor de mensen de  
ja_abortus <- table(df$geslacht, df$eigen_keuze_abortus)  
chi_ja_abortus <- chisq.test(ja_abortus, correct = T)
```

```
## Warning in chisq.test(ja_abortus, correct = T): Chi-squared approximation may  
## be incorrect
```

```
ja_abortus_onder_21_df <- df %>%  
  filter(df$Wat.is.uw.leeftijd. <= 21)  
ja_abortus_onder_21 <- (ifelse(ja_abortus_onder_21_df$eigen_keuze_abortus=="Wel abortus toestaan", 1,0))  
#all(duplicated(abortus$geslacht)[-1L])  
  
# boven de 21:  
ja_abortus_onder_21_df <- df %>%  
  filter(df$Wat.is.uw.leeftijd. > 21)  
ja_abortus_onder_21 <- (ifelse(ja_abortus_onder_21_df$eigen_keuze_abortus=="Wel abortus toestaan", 1,0))  
#all(duplicated(abortus$geslacht)[-1L])  
# Resultaat kan bijna niet kloppen.... achterhaal wat er mis is.
```

H0 is verworpen en H1 is aangenomen omdat de p-waarde "r" ja_nee_abortus\$p.value is. Dit is onder de 0.05 en betekent dat met 95% zekerheid gezegd kan worden dat het geen toeval is dat er een verschil tussen de twee geslachten is en hun keuze om de abortus toe te staan of te verbieden.

```
# Chi-square voor de vraag: Vindt u dat iemand een abortus mag als de zwangerschap een gevolg is van ver
```

```
abortus_als_gevolg_van_01 <- table(df$geslacht, df$Vindt.u.dat.iemand.een.abortus.mag.als.de.zwangerschap
```

```
# Chi-square test:
```

```
chi_abortus_als_gevolg_van <- chisq.test(abortus_als_gevolg_van_01, correct = T)
```

```
## Warning in chisq.test(abortus_als_gevolg_van_01, correct = T): Chi-squared  
## approximation may be incorrect
```

```

# Chi-square voor de vraag: Vindt u dat een minderjarig kind de keuze moet hebben om voor een abortus t
abortus_minderjarig_keuze_01 <- table(df$geslacht, df$minderjarig)

# Chi-square test:
chi_abortus_minderjarig_keuze <- chisq.test(abortus_minderjarig_keuze_01, correct = T)

## Warning in chisq.test(abortus_minderjarig_keuze_01, correct = T): Chi-squared
## approximation may be incorrect

# Chi-square voor de vraag: Vindt u dat de ouders zeggenschap moeten hebben over of hun minderjarige ki

# Lijst van 1 en 0 maken voor Ja en Nee.
#abortus_zeggenschap_ouders_01 <- (ifelse(abortus$ouders=="Ja", 1,0)) # 1 voor de mensen de voorstander
abortus_zeggenschap_ouders_01 <- table(df$geslacht, df$ouders)
# Chi-square test:
chi_abortus_zeggenschap_ouders <- chisq.test(abortus_zeggenschap_ouders_01, correct = T)

## Warning in chisq.test(abortus_zeggenschap_ouders_01, correct = T): Chi-squared
## approximation may be incorrect

# Test om leeftijd mee te nemen als factor:
#chi_abortus_als_gevolg_van_leeftijd <- chisq.test(df$Wat.is.uw.leeftijd., abortus_zeggenschap_ouders,

# Chi-square voor de vraag: Moet het wettelijk verplicht worden voor vrouwen die een abortus willen om

# Lijst van 1 en 0 maken voor Ja en Nee.
#abortus_echo_01 <- (ifelse(abortus$minderjarig=="Ja", 1,0)) # 1 voor de mensen de voorstander zijn, 0
abortus_echo_01 <- table(df$geslacht, df$minderjarig)
# Chi-square test:
chi_abortus_echo <- chisq.test(abortus_echo_01, correct = T)

## Warning in chisq.test(abortus_echo_01, correct = T): Chi-squared approximation
## may be incorrect

# Chi-square voor de vraag: Vindt u dat de wettelijke bedenktijd moet terugkomen?

# Lijst van 1 en 0 maken voor Ja en Nee.
#abortus_bedenktijd_01 <- (ifelse(abortus$bedenktijd=="Ja", 1,0)) # 1 voor de mensen de voorstander zij
abortus_bedenktijd_01 <- table(df$geslacht, df$bedenktijd)
# Chi-square test:
chi_abortus_bedenktijd <- chisq.test(abortus_bedenktijd_01, correct = T)

## Warning in chisq.test(abortus_bedenktijd_01, correct = T): Chi-squared
## approximation may be incorrect

# Chi-square voor de vraag: Vindt u dat adoptie een alternatief is voor abortus?

# Lijst van 1 en 0 maken voor Ja en Nee.
#abortus_adoptie_01 <- (ifelse(abortus$adoptie=="Ja", 1,0)) # 1 voor de mensen de voorstander zijn, 0 a
abortus_adoptie_01 <- table(df$geslacht, df$adoptie)
# Chi-square test:
chi_abortus_adoptie <- chisq.test(abortus_adoptie_01, correct = T)

```

```
## Warning in chisq.test(abortus_adoptie_01, correct = T): Chi-squared
## approximation may be incorrect
```

```
# Chi-square voor de vraag: Moet de biologische vader zeggenschap hebben over de beslissing voor een ab

# Lijst van 1 en 0 maken voor Ja en Nee.
#abortus_zeggenschap_vader_01 <- (ifelse(abortus$vader=="Ja", 1,0)) # 1 voor de mensen de voorstander z
abortus_zeggenschap_vader_01 <- table(df$geslacht, df$vader)
# Chi-square test:
chi_abortus_zeggenschap_vader <- chisq.test(abortus_zeggenschap_vader_01, correct = T)
```

```
## Warning in chisq.test(abortus_zeggenschap_vader_01, correct = T): Chi-squared
## approximation may be incorrect
```

Visualisatie van de Chi-square resultaten:

25-05-2024: Om een overzicht te geven van alle chi-square resultaten voor alle vragen is er een grafiek van gemaakt. De rode lijn geeft significantie aan, wanneer een waarde onder deze lijn ligt ($p < 0.05$) is het resultaat statistisch significant en wanneer de waarde er boven ligt ($p > 0.05$) is het resultaat niet statistisch significant. Als het resultaat significant is betekent het voor deze stellingen dat er een verschil is tussen de keuze tussen de twee geslachten.

```
chi_test_vector_geslacht <- c(chi_ja_abortus[3],chi_abortus_als_gevolg_van[3],chi_abortus_minderjarig_k

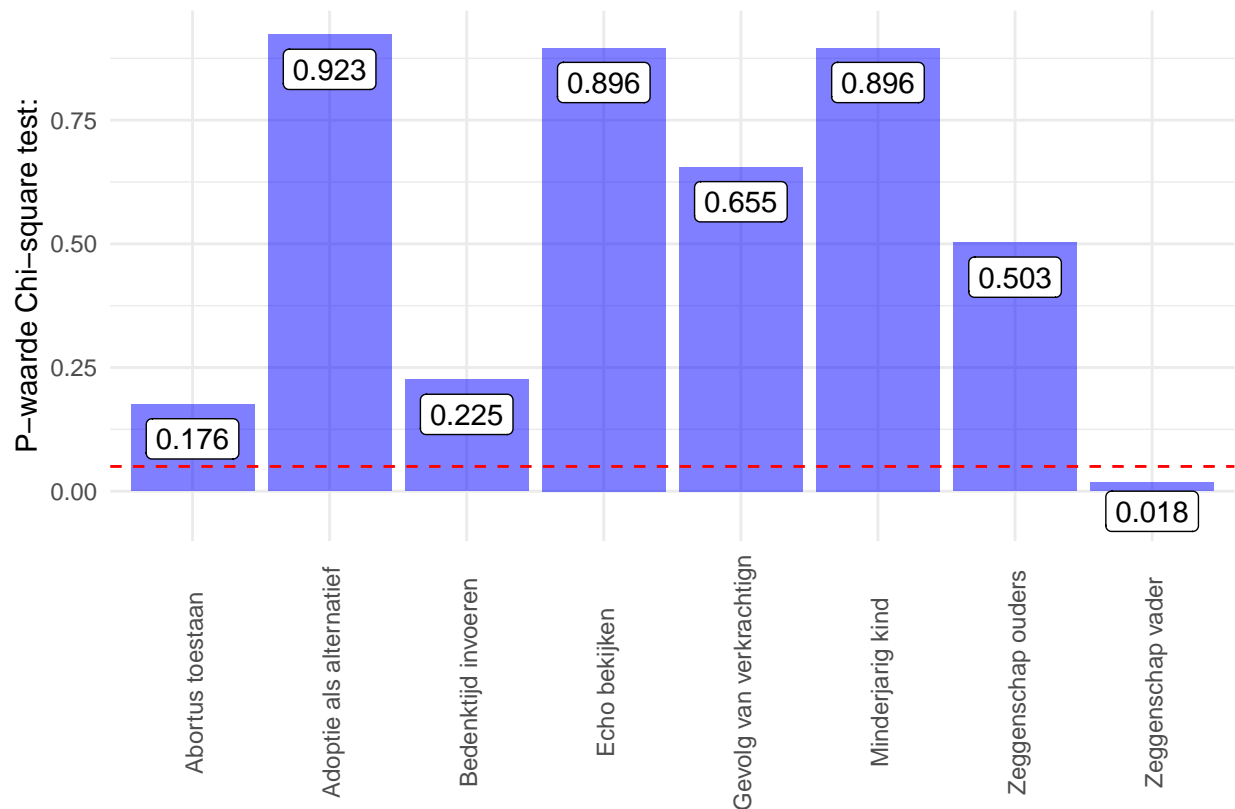
#df_chi_geslacht_test <- data.frame(c(chi_ja_abortus[3],chi_abortus_als_gevolg_van[3]))

df_chi_geslacht <- (data.frame((chi_test_vector_geslacht)) %>% rename(
  `Abortus toestaan` = p.value,
  `Gevolg van verkrachtigen` = p.value.1,
  `Minderjarig kind` = p.value.2,
  `Zeggenschap ouders` = p.value.3,
  `Echo bekijken` = p.value.4,
  `Bedenktijd invoeren` = p.value.5,
  `Adoptie als alternatief` = p.value.6,
  `Zeggenschap vader` = p.value.7
))

df_chi_geslacht %>%
  pivot_longer(cols = c(1:8)) -> df_chi_geslacht
```

In het bovenstaande blok met code wordt er een data-frame gemaakt die alle p-waarden bevat van de hypothesen die opgesteld zijn. Daarna wordt de data veranderd in “tidy data” (een meting per rij en variabelen als kolommen).

```
ggplot(data = df_chi_geslacht, mapping = aes(x = name , y = value)) +
  geom_bar(stat = "identity", fill = "blue", alpha = 0.5) +
  geom_hline(yintercept=0.05, linetype="dashed", color = "red") +
  geom_label_repel(position = position_nudge_repel(y = -0.07), aes(label = round(value,digits = 3)) ) +
  ylab("P-waarde Chi-square test:") +
  xlab(" ") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.6))
```



Aan de bovenstaande grafiek te zien is er alleen een significant verschil tussen de meningen van de verschillende geslachten over de kwestie of de vader ook zeggenschap mag hebben over de abortus, dit is van alle vragen de vraag waar ik ook wel van had verwacht dat het polariserend zou zijn.

Chi² test:

De volgende hypothese is opgesteld: H0: Geslacht is geen bijdragende factor in de keuze of abortus toegestaan mag zijn of niet. H1: Geslacht is een bijdragende factor in de keuze of abortus toegestaan mag zijn of niet.

```
# Test om te kijken hoe het werkt:
```

```
ja_abortus <- table(df$geslacht, df$eigen_keuze_abortus)
# Chi-square test:
chi_ja_abortus <- chisq.test(ja_abortus, correct = F)
```

```
## Warning in chisq.test(ja_abortus, correct = F): Chi-squared approximation may
## be incorrect
```

De p-waarde van de Chi-square test is: 0.1757986 dit is boven de 0.05 dus kan de nul hypothese aangenomen worden, geslacht is geen significant bijdragende factor in de keuze of abortus toegestaan mag zijn of niet. De Chi-squared waarde is: 3.48

```
# Test met andere variabele
```

```

# Lijst van 1 en 0 maken voor Ja en Nee.
abortus_zeggenschap_vader_01 <- (ifelse(abortus$vader=="Ja", 1,0)) # 1 voor de mensen de voorstander zijn.

# Chi-square test:
chi_abortus_zeggenschap_vader <- chisq.test(df$afkomst, abortus_zeggenschap_vader_01, correct = T)

## Warning in chisq.test(df$afkomst, abortus_zeggenschap_vader_01, correct = T):
## Chi-squared approximation may be incorrect

chi_abortus_zeggenschap_vader

##
## Pearson's Chi-squared test
##
## data: df$afkomst and abortus_zeggenschap_vader_01
## X-squared = 8.6173, df = 10, p-value = 0.5688

```

To-do: voor grafieken proporties gebruiken i.p.v. absolute aantallen (om rekening te houden met verschillen en aantal “metingen” tussen mannen en vrouwen.)

21-05-2024:

Grafiek leeftijd/keuze:

Op de onderstaande grafiek wordt het verband tussen leeftijd en keuze bij de vraag: wel of niet abortus toestaan weergegeven.

```

ggplot(data = df, mapping = aes(x = Wat.is.uw.leeftijd., y = ID)) +
  geom_point(aes(color = eigen_keuze_abortus, size = Wat.is.uw.leeftijd.)) +
  facet_wrap(~geslacht) +
  xlab("Leeftijd (Jaren)") +
  ylab("Index nummer (nog aanpassen naar iets anders?)") +
  ggtitle("")

```



Op het bovenstaande figuur wordt het verband tussen leeftijd, geslacht en de keuze of abortus toegestaan zou moeten zijn weergegeven.

29-05-2024: Alle chi-square tests retroactief aangepast, nu met een tabel i.p.v. eerst alle waarden van de vragen omzetten in 1 (Eens) en 0 (Oneens).

30-05-2024:

Logistische regressie:

Om het verband tussen meerdere predictors te zien is de logistische regressie gebruikt.

Omdat er niet genoeg mensen van elk geloof zijn wordt iedereen verdeeld in twee groepen, gelovig en niet gelovig, het is natuurlijk genuanceerder dan dat maar met deze versimpeling zal ook al wel een patroon zichtbaar gemaakt kunnen worden.

```
df <- df %>%
  mutate(gelovig_01 = ifelse(df$geloof == "Niet gelovig", 0,1))
```

Dataframes maken voor de logistische regressie:

In de onderstaande code chunks worden er dataframes gemaakt voor de volgende tests: Keuze van de man met leeftijd en opleidingsniveau als predictors, Keuze van de vrouw met leeftijd en opleidingsniveau als predictors, en Keuze van mensen met ongespecificeerd geslacht met leeftijd en opleidingsniveau als predictors. Van deze laatste categorie is er maar een persoon, dus kan hier de analyse niet op uitgevoerd worden.


```
# Nieuw dataframe met de relevante colommen van df.
#df_logistische_regressie <- data.frame(ID = df$ID, geslacht =df$geslacht, leeftijd = df$Wat.is.uw.leef

# Voor de kolom met 0,1 waarden voor de keuze zijn alle "Weet ik niet" antwoorden vervangen door 0 (tes
df_logistische_regressie <- df %>%
  select(geslacht, Wat.is.uw.leeftijd., gelovig_01, Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afge
  mutate(ouders = (ifelse(ouders == "Ja", 1, 0)), leeftijd_onder_30 = (ifelse(Wat.is.uw.leeftijd. <= 30
```

De leeftijden zijn verdeeld in twee categorieën, mensen onder 30 en mensen boven de 30 jaar oud, de reden voor deze indeling is dat er in de verdeling van leeftijden twee pieken zijn, rond de 20 en rond de 50. Er zijn maar weinig metingen die hier tussen in zitten. Het jongste persoon die niet voor vrijheid van abortus gekozen heeft is 39 jaar oud.

```
# Dataframe voor man:
df_man <- df %>%
  select(geslacht, Wat.is.uw.leeftijd., gelovig_01, Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afge
  filter(geslacht == "Man")
```

Maak het dataframe voor de mannen

```
# Dataframe voor vrouw:
df_vrouw <- df %>%
  select(geslacht, Wat.is.uw.leeftijd., gelovig_01, Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afge
```

Maak het dataframe voor de vrouwen

```
# Dataframe voor ongespecificeerd geslacht:
df_ongespecificeerd_geslacht <- df %>%
  select(geslacht, Wat.is.uw.leeftijd., gelovig_01, Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afge
```

Maak het dataframe voor de “ongespecificeerd geslacht” mensen.

```
# logistische_regressie voor de vraag "moet abortus een eigen keuze zijn?" met als voorspellende factor
logistische_regressie_test <- glm(df$eigen_keuze_abortus ~ df$Wat.is.uw.leeftijd. + df$geslacht + df$Wat
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

Voor de logistische_regressie voor de vraag “moet abortus een eigen keuze zijn?” met als predictors leeftijd, geslacht en opleidingsniveau gaf R een waarschuwing: Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred. Mogenlijk komt dit omdat er niet genoeg metingen zijn voor elk opleidingsniveau.

```
# Logistische regressie met leeftijd, geloof, opleidingsniveau en geslacht als predictors
logistische_regressie <- glm(eigen_keuze_abortus ~ leeftijd_onder_30 + geslacht + gelovig_01 + Wat.is.h
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
lr_man_vrouw_vader <- glm(vader ~ leeftijd_onder_30 + geslacht + Wat.is.het.hoogste.onderwijsniveau.dat
```

```
# Logistische regressie voor de vraag: Moeten minderjarig kind zeggenschap hebben over eigen abortus, m
```

```

lr_man_vrouw_zeggenschap_minderjarig <- glm(minderjarig ~ leeftijd_onder_30 + geslacht + Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO
# Logistische regressie voor de vraag: Moet er een echo van de foetus laten zien worden voor er voor abortus gekozen wordt
lr_man_vrouw_echo <- glm(echo ~ leeftijd_onder_30 + geslacht + Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO
# Bedenktijd:
lr_man_vrouw_bedenktijd <- glm(bedenktijd ~ leeftijd_onder_30 + geslacht + Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO
# Logistische regressie voor de vraag: Mag er voor abortus gekozen worden als de gezondheid van de vrouw goed is
lr_man_vrouw_gezondheid_vrouw <- glm(df_logistische_regressie$gezondheid_vrouw ~ leeftijd_onder_30 + geslacht + Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

lr_man_vrouw_gezondheid_ouders <- glm(df_logistische_regressie$ouders ~ leeftijd_onder_30 + geslacht + Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO

#summary(lr_man_vrouw_zeggenschap_minderjarig) # geloof significant
summary(lr_man_vrouw_vader) # leeftijd en geslacht significant, geloof bijna

##
## Call:
## glm(formula = vader ~ leeftijd_onder_30 + geslacht + Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO,
##      data = df_logistische_regressie, family = binomial)
##
## Coefficients:
##
## (Intercept)                                Estimate
## leeftijd_onder_30                        -1.5508
## geslachtVrouw                             1.9573
## geslachtWil ik liever niet melden         18.8250
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO -15.8463
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO -17.3638
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school -15.7986
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige -0.5294
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit -16.5220
## gelovig_01                               -1.1689
##
## Std. Error
## (Intercept)                        2275.9594
## leeftijd_onder_30                   0.6585
## geslachtVrouw                      0.5739
## geslachtWil ik liever niet melden 3956.1804
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO 2275.9595
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO 2275.9594
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school 2275.9595
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 4564.1379
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit 2275.9595
## gelovig_01                         0.5973
##
## z value
## (Intercept)                        0.008
## leeftijd_onder_30                  -2.355

```

```

## geslachtVrouw 3.411
## geslachtWil ik liever niet melden 0.005
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO -0.007
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO -0.008
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school -0.007
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 0.000
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit -0.007
## gelovig_01 -1.957
## Pr(>|z|)
## (Intercept) 0.993933
## leeftijd_ouder_30 0.018521
## geslachtVrouw 0.000648
## geslachtWil ik liever niet melden 0.996203
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO 0.994445
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO 0.993913
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school 0.994462
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 0.999907
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit 0.994208
## gelovig_01 0.050357
##
## (Intercept)
## leeftijd_ouder_30 *
## geslachtVrouw ***
## geslachtWil ik liever niet melden
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit
## gelovig_01 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 140.61 on 113 degrees of freedom
## Residual deviance: 109.62 on 104 degrees of freedom
## AIC: 129.62
##
## Number of Fisher Scoring iterations: 16

```

```

#summary(logistische_regressie) # niets significant
#summary(lr_man_vrouw_echo) # geloof significant
#summary(lr_man_vrouw_bedenktijd) # geslacht significant
#summary(lr_man_vrouw_gezondheid_vrouw) # niets significant
#summary(lr_man_vrouw_gezondheid_ouders) # Niets significant

```

In het bovenstaande chunk staan de summary()'s uitgecomment, verwijder het “#” om het resultaat te zien, ook is er een samenvatting achter geschreven over welke factoren significant bijdragen. Voor de vraag: Moet een minderjarig kind de keuze hebben om een abortus te nemen? De enige significante factor is geloof, met een p-waarde van 0.00308.

Textuele resultaten logistische regressie:

```
summary(lr_man_vrouw_vader) # leeftijd en geslacht significant, geloof bijna
```

```
##
## Call:
## glm(formula = vader ~ leeftijd_onder_30 + geslacht + Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.,
##      gelovig_01, family = binomial, data = df_logistische_regressie)
##
## Coefficients:
##                                     Estimate
## (Intercept)                        17.3070
## leeftijd_onder_30                  -1.5508
## geslachtVrouw                      1.9573
## geslachtWil ik liever niet melden  18.8250
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO -15.8463
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO -17.3638
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school -15.7986
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige -0.5294
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit -16.5220
## gelovig_01                        -1.1689
##                                     Std. Error
## (Intercept)                      2275.9594
## leeftijd_onder_30                 0.6585
## geslachtVrouw                    0.5739
## geslachtWil ik liever niet melden 3956.1804
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO 2275.9595
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO 2275.9594
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school 2275.9595
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 4564.1379
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit 2275.9595
## gelovig_01                      0.5973
##                                     z value
## (Intercept)                      0.008
## leeftijd_onder_30                 -2.355
## geslachtVrouw                     3.411
## geslachtWil ik liever niet melden 0.005
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO -0.007
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO -0.008
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school -0.007
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 0.000
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit -0.007
## gelovig_01                      -1.957
##                                     Pr(>|z|)
## (Intercept)                      0.993933
## leeftijd_onder_30                 0.018521
## geslachtVrouw                    0.000648
## geslachtWil ik liever niet melden 0.996203
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO 0.994445
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO 0.993913
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school 0.994462
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 0.999907
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit 0.994208
```

```
## gelovig_01                                0.050357
##
## (Intercept)
## leeftijd_onder_30                          *
## geslachtVrouw                             ***
## geslachtWil ik liever niet melden
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit
## gelovig_01                                .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 140.61  on 113  degrees of freedom
## Residual deviance: 109.62  on 104  degrees of freedom
## AIC: 129.62
##
## Number of Fisher Scoring iterations: 16
```

Uit de bovenstaande logistische regressie blijkt dat leeftijd en geslacht significant bijdragende factoren zijn, voor deze analyse is de leeftijd van de deelnemers in twee categorieën verdeeld: boven en onder de 30, wanneer dit niet gedaan wordt is de coefficient van leeftijd nog steeds onder de 0.05 maar wel wat hoger. Geloof was net niet onder de 0.05 met 0.0503.

```
summary(lr_man_vrouw_zeggenschap_minderjarig) # geloof significant
```

```
##
## Call:
## glm(formula = minderjarig ~ leeftijd_onder_30 + geslacht + Wat.is.het.hoogste.onderwijsniveau.dat.u.l
##     gelovig_01, family = binomial, data = df_logistische_regressie)
##
## Coefficients:
##                                     Estimate
## (Intercept)                        -17.5595
## leeftijd_onder_30                   -0.4770
## geslachtVrouw                       0.4671
## geslachtWil ik liever niet melden  -16.3154
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO    14.7302
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO    15.5733
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school 15.0949
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige  -2.5294
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit   14.3919
## gelovig_01                       2.0557
##                                     Std. Error
## (Intercept)                        2284.0720
## leeftijd_onder_30                   0.8130
## geslachtVrouw                       0.6669
## geslachtWil ik liever niet melden  3956.1805
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO    2284.0721
```

```

## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO 2284.0720
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school 2284.0720
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 4568.1887
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit 2284.0723
## gelovig_01 0.6620
## z value
## (Intercept) -0.008
## leeftijd_onder_30 -0.587
## geslachtVrouw 0.701
## geslachtWil ik liever niet melden -0.004
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO 0.006
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO 0.007
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school 0.007
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige -0.001
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit 0.006
## gelovig_01 3.105
## Pr(>|z|)
## (Intercept) 0.9939
## leeftijd_onder_30 0.5574
## geslachtVrouw 0.4836
## geslachtWil ik liever niet melden 0.9967
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO 0.9949
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO 0.9946
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school 0.9947
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 0.9996
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit 0.9950
## gelovig_01 0.0019
##
## (Intercept)
## leeftijd_onder_30
## geslachtVrouw
## geslachtWil ik liever niet melden
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit
## gelovig_01 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 92.477 on 113 degrees of freedom
## Residual deviance: 76.810 on 104 degrees of freedom
## AIC: 96.81
##
## Number of Fisher Scoring iterations: 16

```

```
summary(lr_man_vrouw_bedenktijd) # geslacht significant
```

```

##
## Call:
## glm(formula = bedenktijd ~ leeftijd_onder_30 + geslacht + Wat.is.het.hoogste.onderwijsniveau.dat.u.h

```

```

##      gelovig_01, family = binomial, data = df_logistische_regressie)
##
## Coefficients:
##
##                                     Estimate
## (Intercept)                        0.1209
## leeftijd_onder_30                  -0.1165
## geslachtVrouw                      1.0153
## geslachtWil ik liever niet melden 14.9772
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO 1.4470
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO -0.2780
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school -0.2195
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 15.2925
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit 0.7216
## gelovig_01                        -0.8626
##
##                                     Std. Error
## (Intercept)                        1.3153
## leeftijd_onder_30                  0.5986
## geslachtVrouw                      0.4771
## geslachtWil ik liever niet melden 1455.3978
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO 1.3869
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO 1.3157
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school 1.3284
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 1455.3982
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit 1.4688
## gelovig_01                        0.5694
##
##                                     z value
## (Intercept)                        0.092
## leeftijd_onder_30                  -0.195
## geslachtVrouw                      2.128
## geslachtWil ik liever niet melden 0.010
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO 1.043
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO -0.211
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school -0.165
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 0.011
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit 0.491
## gelovig_01                        -1.515
##
##                                     Pr(>|z|)
## (Intercept)                        0.9268
## leeftijd_onder_30                  0.8457
## geslachtVrouw                      0.0333
## geslachtWil ik liever niet melden 0.9918
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO 0.2968
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO 0.8327
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school 0.8688
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 0.9916
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit 0.6232
## gelovig_01                        0.1298
##
## (Intercept)
## leeftijd_onder_30
## geslachtVrouw
## geslachtWil ik liever niet melden
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO

```

```
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit
## gelovig_01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 147.74  on 113  degrees of freedom
## Residual deviance: 127.56  on 104  degrees of freedom
## AIC: 147.56
##
## Number of Fisher Scoring iterations: 14
```

```
summary(lr_man_vrouw_echo) # Geloof significant
```

```
##
## Call:
## glm(formula = echo ~ leeftijd_onder_30 + geslacht + Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.a
##      gelovig_01, family = binomial, data = df_logistische_regressie)
##
## Coefficients:
##                                     Estimate
## (Intercept)                        18.7212
## leeftijd_onder_30                   0.6646
## geslachtVrouw                      0.9408
## geslachtWil ik liever niet melden  16.9309
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO -15.0888
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO -16.9710
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school -16.4218
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige  1.5659
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit  1.0039
## gelovig_01                      -1.6618
##                                     Std. Error
## (Intercept)                    5953.9161
## leeftijd_onder_30              1.0198
## geslachtVrouw                 0.9246
## geslachtWil ik liever niet melden 10754.0130
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO  5953.9162
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO  5953.9161
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school 5953.9161
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 12292.1891
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit  6702.7014
## gelovig_01                   0.7848
##                                     z value
## (Intercept)                   0.003
## leeftijd_onder_30             0.652
## geslachtVrouw                 1.018
## geslachtWil ik liever niet melden 0.002
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO -0.003
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO -0.003
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school -0.003
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 0.000
```



```
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit      0.000
## gelovig_01                                                                -2.117
##                                                                           Pr(>|z|)
## (Intercept)                                                                0.9975
## leeftijd_onder_30                                                         0.5146
## geslachtVrouw                                                             0.3089
## geslachtWil ik liever niet melden                                         0.9987
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO               0.9980
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO               0.9977
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school 0.9978
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige           0.9999
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit      0.9999
## gelovig_01                                                                0.0342
##
## (Intercept)
## leeftijd_onder_30
## geslachtVrouw
## geslachtWil ik liever niet melden
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit
## gelovig_01                                                                *
```

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

(Dispersion parameter for binomial family taken to be 1)
Null deviance: 62.972 on 113 degrees of freedom
Residual deviance: 48.174 on 104 degrees of freedom
AIC: 68.174
Number of Fisher Scoring iterations: 18

```
# Logistische regressie voor de vraag: Moet abortus een eigen keuze zijn?
```

```
lr_man_eigen_keuze_abortus <- glm(eigen_keuze_abortus ~ Wat.is.uw.leeftijd. + gelovig_01 + Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit)
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
# Logistische regressie voor de vraag:
```

```
lr_man_minderjarig <- glm(minderjarig ~ Wat.is.uw.leeftijd. + gelovig_01 + Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit)
```

```
#coef(lr_man_eigen_keuze_abortus)
```

```
# Logistische regressie voor de vraag: Moet de biologische vader zeggenschap hebben over de abortus
```

```
lr_man_vader <- glm(vader ~ Wat.is.uw.leeftijd. + gelovig_01 + Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit)
```

```
#summary(lr_man_eigen_keuze_abortus)
```

```
#summary(lr_man_vader)
```

De bovenstaande twee chunks worden niet meer gebruikt omdat de modellen met beide geslachten er in tegelijk ook werken.

De odds ratio wordt gebruikt om te laten zien hoeveel invloed de predictors individueel hebben wanneer de anderen niet veranderen.

```
# Zodat het niet in wetenschappelijke notatie weergegeven wordt als het onder de 999 is.
options(scipen = 999)

# Bereken de coëfficiënten (log e)
coefficienten_van_lr_man_eigen_keuze_abortus <- coef(lr_man_eigen_keuze_abortus)
coefficienten_van_lr_man_vader <- coef(lr_man_vader)
coefficienten_van_lr_man_vrouw_vader <- coef(lr_man_vrouw_vader)

# Verander de coëfficiënten naar procentuele verandering:
odds_ratio_lr_man_eigen_keuze_abortus <- ( 1 - (exp(coefficienten_van_lr_man_eigen_keuze_abortus))) * 100
odds_ratio_van_lr_man_vader <- exp(coefficienten_van_lr_man_vader)
odds_ratio_van_lr_man_vrouw_vader <- exp(coefficienten_van_lr_man_vrouw_vader)

#list(odds_ratio_lr_man_eigen_keuze_abortus)
odds_ratio_lr_man_eigen_keuze_abortus <- data.frame(odds_ratio_lr_man_eigen_keuze_abortus) # %>% rename(
#)
odds_ratio_van_lr_man_vrouw_vader <- data.frame(odds_ratio_van_lr_man_vrouw_vader) # maakt de tabel net
#odds_ratio_van_lr_man_vader
#odds_ratio_lr_man_eigen_keuze_abortus

odds_ratio_van_lr_man_vrouw_vader
```

```
##                                odds_ratio_van_lr_man_vrouw_vader
## (Intercept)                    32835420.7421849
## leeftijd_onder_30                0.2120859
## geslachtVrouw                    7.0799520
## geslachtWil ik liever niet melden 149822280.3343318
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.HBO 0.0000001
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.MBO 0.0000000
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Middelbare school 0.0000001
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Overige 0.5889837
## Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.Universiteit 0.0000000
## gelovig_01                      0.3107208
```

Interpretatie van deze odds ratio's: Het percentage geeft aan dat wanneer alle andere variabelen het zelfde zijn, de waarschijnlijkheid dat er een andere uitkomst gekozen wordt (bijvoorbeeld een ander antwoord op de vraag of abortus toegestaan zou moeten zijn). De log-odds van de logistische regressie over de vraag: "Moet de biologische vader zeggenschap hebben over de beslissing voor een abortus?" Zijn als volgt: 7.08 voor geslacht en 0.21 voor leeftijd (verdeeld in twee leeftijdsgroepen). Dit zijn de significante factoren (p-waarde < 0.05) geloof was bijna (maar niet helemaal) significant met een p-waarde van: 0.050357.

```
# Voorspel test met de volgende test data: (van twee personen die niet in de dataset zitten)
newdata1 <- data.frame(geslacht = c("Man", "Vrouw"), Wat.is.uw.leeftijd. = c(20,22), keuze = c("Wel abortus", "Geen abortus"))

# Voorspelling van de bovenstaande twee test observaties aan de hand van de gemaakte logistische regressie
#voorspelde_waarde_1 <- predict.lm(logistische_regressie, newdata = newdata1, type = "response")
#voorspelde_waarde_1
```

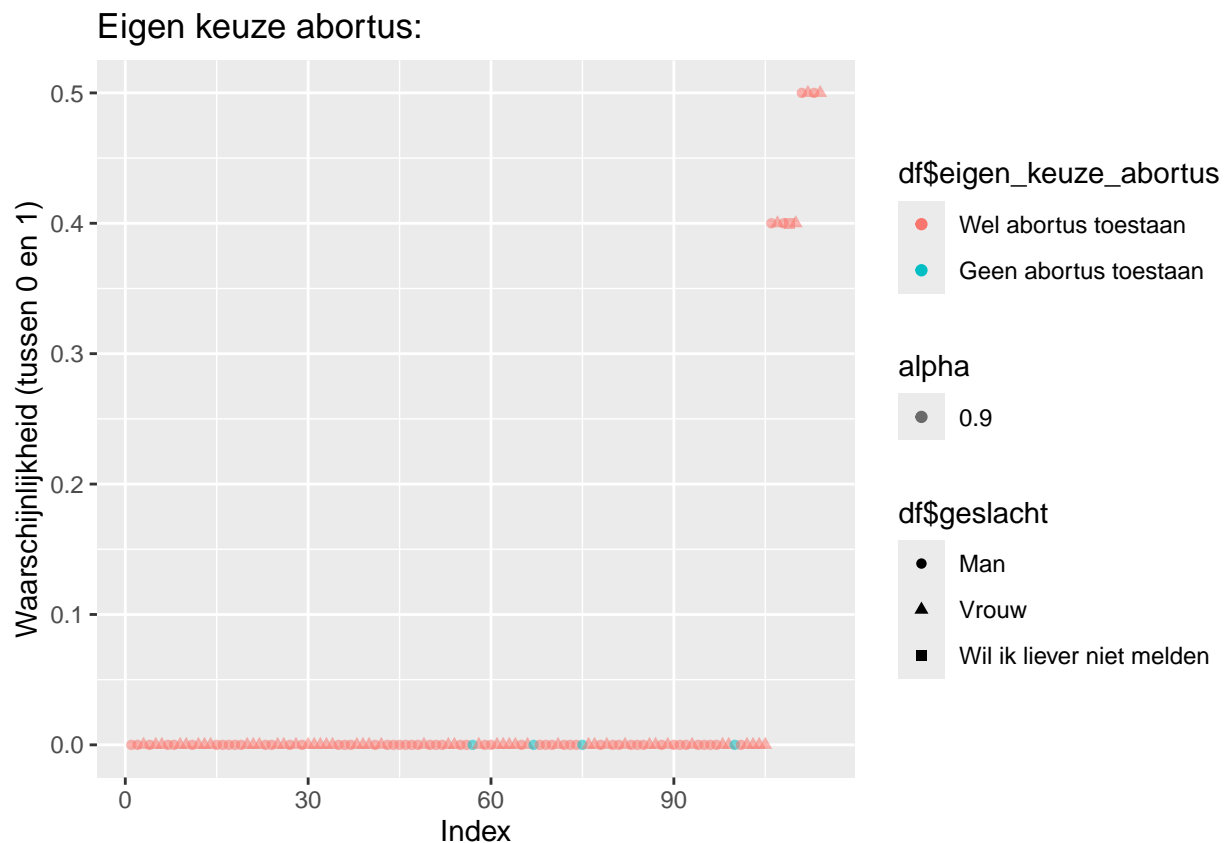
Met behulp van de `predict.lm()` functie worden de twee test observaties (net als in de originele dataset) aan de hand van de gemaakte logistische regressie. De nummers moeten nog omgezet worden naar een Ja/Nee antwoord op de vraag: Abortus toestaan. De waarden die verkregen zijn met de `predict.lm()` functie kunnen vervolgens in de onderstaande grafiek worden geplott om te zien in welke categorie ze waarschijnlijk horen. (Welk antwoord het persoon waarschijnlijk zou geven.)

```
# Maak nieuw dataframe met de waardes van de logistische regressie en de keuze die de mensen die bij de
voorspelde_data <- data.frame(logistische_regressie$fitted.values, df$eigen_keuze_abortus)

voorspelde_data <- voorspelde_data[order(voorspelde_data$logistische_regressie.fitted.values, decreasing=TRUE), ]

voorspelde_data$rank <- 1:nrow(voorspelde_data)

ggplot(data = voorspelde_data, aes(x = rank, y = logistische_regressie.fitted.values)) +
  geom_point(aes(color=df$eigen_keuze_abortus, shape = df$geslacht, alpha = 0.9)) +
  ggtitle("Eigen keuze abortus:") +
  xlab("Index") +
  ylab("Waarschijnlijkheid (tussen 0 en 1)")
```



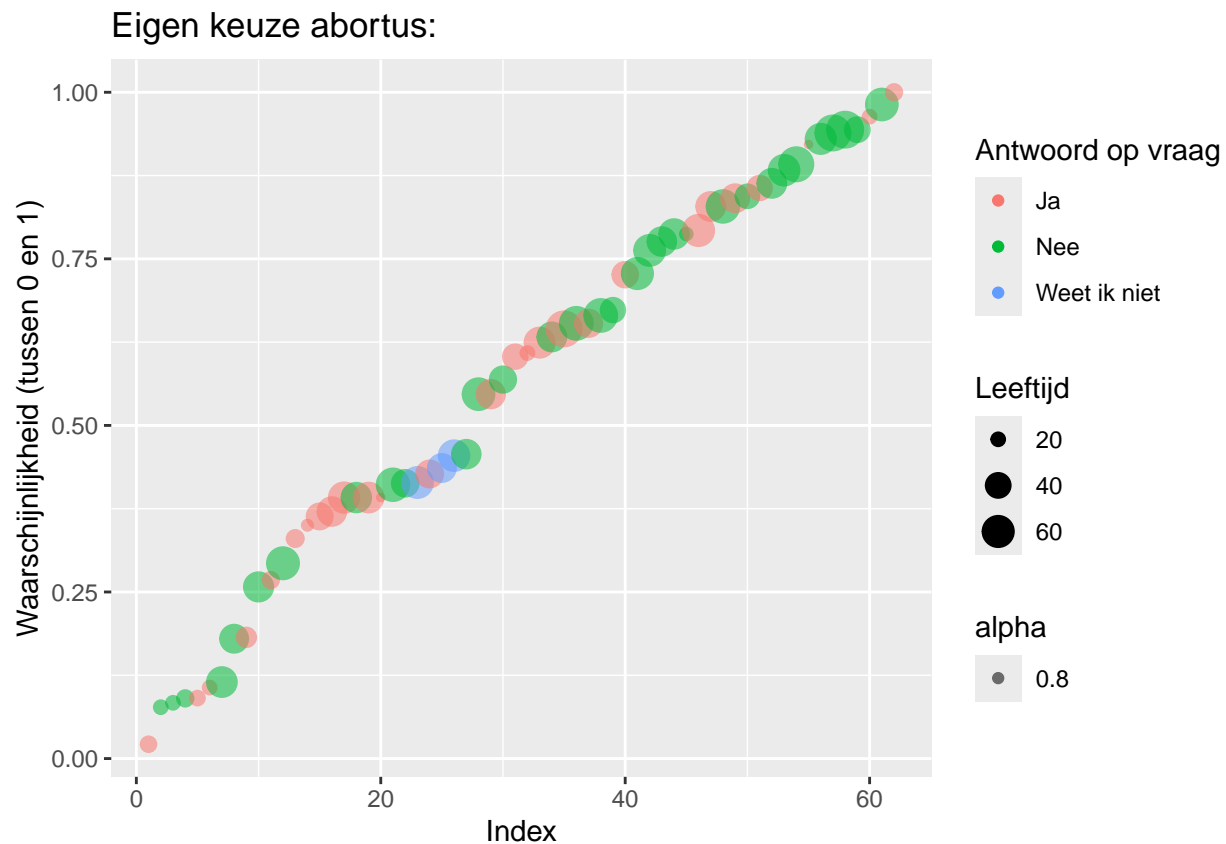
Voor de bovenstaande logistische regressie waren er geen significante predictors, daarom is er geen trend op de grafiek zichtbaar en is deze onbruikbaar. (deze grafiek was een test)

```
# Maak nieuw dataframe met de waardes van de logistische regressie en de keuze die de mensen die bij de
lr_voorspelling_vader <- data.frame(lr_man_vader$fitted.values, df_man$vader)

lr_voorspelling_vader <- lr_voorspelling_vader[order(lr_voorspelling_vader$lr_man_vader.fitted.values, decreasing=TRUE), ]
```

```
lr_voorspelling_vader$rank <- 1:nrow(lr_voorspelling_vader)

ggplot(data = lr_voorspelling_vader, aes(x = rank, y = lr_man_vader.fitted.values)) +
  geom_point(aes(color=df_man$vader, size = df_man$Wat.is.uw.leeftijd., alpha = 0.8)) +
  labs(size = "Leeftijd", color = "Antwoord op vraag") +
  ggtitle("Eigen keuze abortus:") +
  xlab("Index") +
  ylab("Waarschijnlijkheid (tussen 0 en 1)")
```



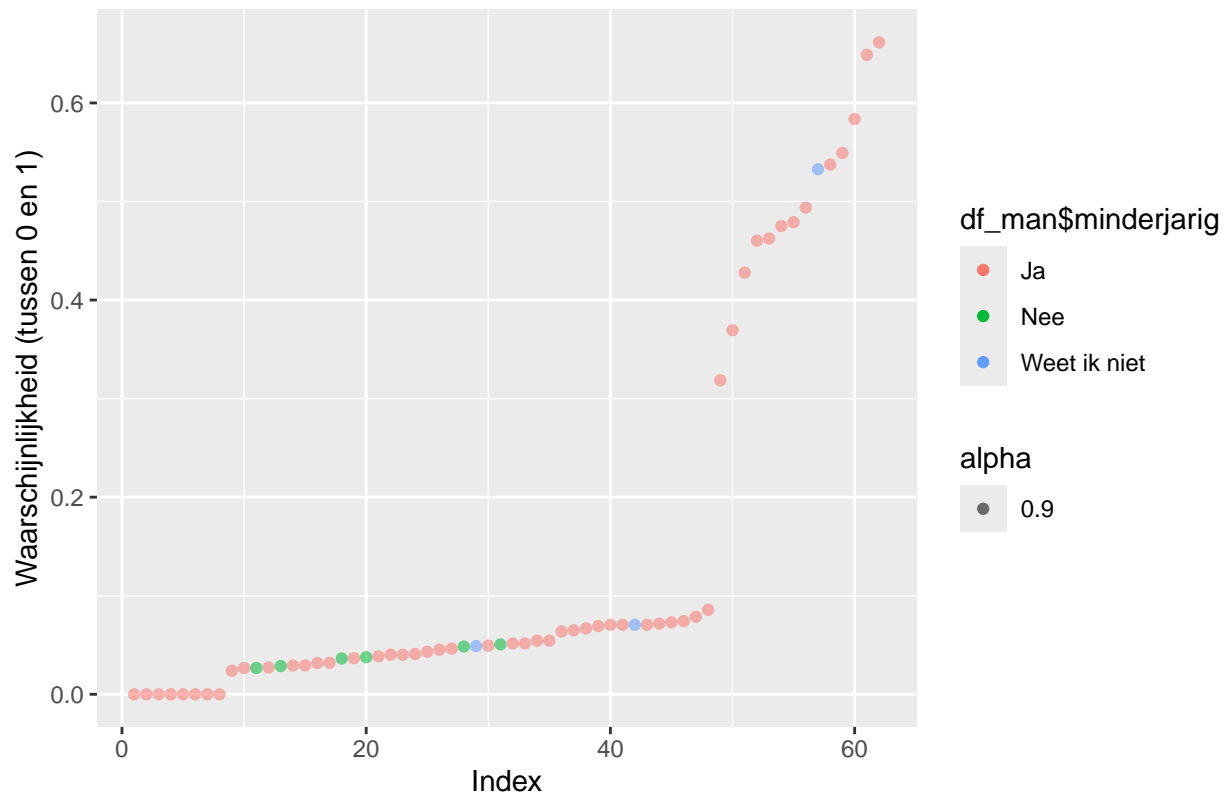
Op de visualizatie van de bovenstaande logistische regressie was leeftijd de enige significante predictor van de keuze voor abortus. (P-waarde: 0.00296) De trend in leeftijd is te zien door de groter wordende cirkels die een hogere leeftijd symboliseren.

```
# Maak nieuw dataframe met de waardes van de logistische regressie en de keuze die de mensen die bij de
lr_man_voorspelling_minderjarig <- data.frame(lr_man_minderjarig$fitted.values, df_man$minderjarig)

lr_man_voorspelling_minderjarig <- lr_man_voorspelling_minderjarig[order(lr_man_voorspelling_minderjarig,
lr_man_voorspelling_minderjarig$rank <- 1:nrow(lr_man_voorspelling_minderjarig)

ggplot(data = lr_man_voorspelling_minderjarig, aes(x = rank, y = lr_man_minderjarig.fitted.values)) +
  geom_point(aes(color=df_man$minderjarig, alpha = 0.9)) +
  ggtitle("Eigen keuze abortus:") +
  xlab("Index") +
  ylab("Waarschijnlijkheid (tussen 0 en 1)")
```

Eigen keuze abortus:



```
lr_man_vrouw_vader
```

```
# Maak nieuw dataframe met de waarden van de logistische regressie en de keuze die de mensen die bij de
lr_voorspelling_man_vrouw_vader <- data.frame(lr_man_vrouw_vader$fitted.values, df_logistische_regressie$
lr_voorspelling_man_vrouw_vader <- lr_voorspelling_man_vrouw_vader[order(lr_voorspelling_man_vrouw_vader
lr_voorspelling_man_vrouw_vader$rank <- 1:nrow(lr_voorspelling_man_vrouw_vader)

test_lr <- ggplot(data = lr_voorspelling_man_vrouw_vader, aes(x = rank, y = lr_voorspelling_man_vrouw_v
  geom_point(aes(color=df_logistische_regressie$geslacht ,size=df_logistische_regressie$Wat.is.uw.leeft
  geom_ribbon(aes(ymin = (rank),
    ymax = (rank), fill = rank), alpha = 0.2) +
    geom_line(aes(colour = rank), size = 1) +
  ggtitle("Eigen keuze abortus:") +
  xlab("Index") +
  ylab("Waarschijnlijkheid (tussen 0 en 1)")
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

Zoals op de bovenstaande afbeelding zichtbaar is, is er minder goed een patroon in te zien wanneer alle predictor factoren (leeftijd, geslacht, opleidingsniveau en geloof) meegenomen worden.

Test visualisatie voor paper:

31-05-2024: Als introductie van de verkregen data moeten er een paar grafieken weergegeven worden waarop de keuzes van de mensen die de enquête ingevuld hebben weergegeven is.

```
x1 <- table(df$adoptie, df$geslacht)
adoptie_plot_data <- data.frame(x1)
adoptie_plot_data$fraction = adoptie_plot_data$Freq / sum(adoptie_plot_data$Freq)
adoptie_plot_data$percentage = round(adoptie_plot_data$fraction *100, digits =1)

#ggplot(data = df, mapping = aes(x = adoptie_plot_data, y = adoptie, fill = geslacht)) +
ggplot(data = adoptie_plot_data, mapping = aes(x = Var2, y = percentage, fill = Var1)) +
geom_bar(stat = "identity" ,
  position ="dodge") +
labs(fill = "Keuze") +
geom_text(aes(label = (""))) +
xlab("Antwoorden") +
ylab(" Percentage antwoord") +
theme_minimal()
```

Net zoals op de bovenstaande afbeelding moeten alle vragen gevisualiseerd worden, ook moeten proporties gebruikt worden i.p.v. absolute getallen omdat er niet even veel mannen als vrouwen zijn.