

Logboek_Iris

Iris Ineke

2024-05-06

Inleiding

In dit onderzoek proberen wij te onderzoeken of er factoren zijn die invloed hebben op iemand zijn mening over abortus. We hebben een enquête met vragen gemaakt en hebben deze vervolgens rondgestuurd. De resultaten die uit deze enquête kwam hebben we met behulp van Rstudio geanalyseerd. Hierin hebben we onderzocht of factoren zoals; geslacht, onderwijsniveau, leeftijd, religieuze overtuiging en de provincie waarin iemand is opgegroeid invloed hebben op hoe mensen de vragen in de enquête hebben beantwoord.

Diclaimer

We dachten eerst dat we een gezamenlijk logboek moesten hebben. Hierdoor begint dit logboek pas op 26 april.

Vrijdag 26 april

Protocol geschreven en introductie voor vraag over bedenktijd: “vindt u dat de wettelijke bedenktijd moet terugkomen?”

Protocol: Materiaal: Een werkende R installatie Een enquête over het door ons gekozen ethische probleem: Abortus.

Methode: Maak een enquête die vragen bevat over de gekozen kwestie: ‘Abortus’. Deze vragenlijst bestaat uit een persoonlijk deel met vragen over onder andere; leeftijd, sekse en geloof. Verder heeft de enquête nog een deel met inhoudelijke vragen over de gekozen ethische kwestie. Deze vragen zullen ons een indicatie geven over hoe de mensen denken over het ethische onderwerp. Zodra deze enquête gemaakt is, zullen we deze laten testen. Dit doen we door familieleden te vragen om de enquête te maken en hun feedback hierna te delen. Waar nodig zullen we vervolgens de vragenlijst verbeteren. Vervolgens gaan we zoveel mogelijk mensen vragen de vragenlijst in te vullen. Met de data die uit deze vragenlijst komt kunnen we vervolgens door middel van R verder onderzoek doen. In R gaan we door middel van grafieken maken onderzoeken of er inderdaad een correlatie is tussen de gevraagde variabele en de mening over abortus.

Achtergrond 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.

Zaterdag 4 mei

Feedback formulier gemaakt voor de enquête vragen. Vragen:

- Hoe zou u de algehele ervaring van het invullen van deze enquête beoordelen?
Slecht, matig, voldoende, goed, uitstekend.
- Welke aspecten van de enquête vond u het meest waardevol en waarom?
- Zijn er specifieke vragen in de enquête die u verwarrend vond of waarvan u denkt dat ze verbeterd kunnen worden? Zo ja, kunt u ons vertellen welke vragen en eventuele suggesties voor verbetering?
- Hoe zou u de lengte van de enquête beoordelen?
te kort, precies goed, te lang.
- Heeft u nog andere opmerkingen of suggesties voor ons?
(Zie document Feedback_enquête.docx).

Zondag 5 mei

Feedback laten invullen door twee mensen en resultaten besproken. Conclusie: De volgorde van de vragen veranderd. Na de vraag over of er uitzonderingen zijn, de vragen met uitzonderingen geplaatst. Verder nog de vraagstelling van sommige vragen aangepast. En de introductie herschreven.

Woensdag 8 mei

Vragenlijst rondgestuurd. Logboek individueel geschreven.

Maandag 13 mei

Inlezen radar chart voor presentatie (met uit welke provincie mensen komen). input: - kolom: quantitative variable (hoeveelheden) - eerste 2 regels: min/max per variable - Library 'fmsb' is nodig, nog installeren!! Je kan het extra customizen, zie r-graph-gallery.com (voor gegevens enquête verwerken gebruiken: 'grouped and stacked barplot')

Dinsdag 14 mei

Omdat een ander groepje al de radar chart gaat doen, hebben we besloten om een andere visualisatie te kiezen. Dus we gaan nu de circular packing uitzoeken. Dit kan met de 'ggraph' package.

Woensdag 15 mei

Data inlezen.

```
# file <- "dataset.csv"
# data <- readr::read_csv(file)
# data <- read.csv(file)
data <- read.csv('dataset_nieuw.csv')
# head(data)
```

```
# summary(data)
```

Kolommen duidelijkere namen gegeven en onnodige kolommen verwijderd. (stellingen nog niet veranderd!)

```
library(dplyr)
```

```
##  
## Attaching package: 'dplyr'  
  
## The following objects are masked from 'package:stats':  
##  
##   filter, lag  
  
## The following objects are masked from 'package:base':  
##  
##   intersect, setdiff, setequal, union
```

```
data_renamed <- data %>%  
  rename(  
    Eindtijd = Tijd.van.voltooien,  
    Gender = Wat.is.uw.geslacht.,  
    Onderwijsniveau = Wat.is.het.hoogste.onderwijsniveau.dat.u.heeft.afgerond.,  
    Leeftijd = Wat.is.uw.leeftijd.,  
    Religie = Wat.is.uw.religieuze.overtuiging.,  
    Provincie = In.welke.provincie.heeft.u.het.grootste.deel.van.uw.jeugd.doorgebracht.,  
    Abortus_keuze = Vindt.u.dat.iemand.zelf.moet.kunnen.kiezen.of.die.een.abortus.wil.,  
    Indien_nee_uitzonderingen = Indien.u.bij.de.vorige.vraag..Nee..heeft.ingevuld...Vindt.u.dat.er.,  
    Gezondheid_vrouw_gevaar_uitzondering = Vindt.u.dat.er.voor.abortus.gekozen.mag.worden.als.de.ge.,  
    Verkrachting_uitzondering = Vindt.u.dat.iemand.een.abortus.mag.als.de.zwangerschap.een.gevolg.i.,  
    minderjarige_keuze_uitzondering = Vindt.u.dat.een.minderjarig.kind.de.keuze.moet.hebben.om.voor.,  
    Ouders_zeggenschap_over_minderjarige_uitzondering = Vindt.u.dat.de.ouders.zeggenschap.moeten.heb.,  
    Echo_verplicht = Moet.het.wettelijk.verplicht.worden.voor.vrouwen.die.een.abortus.willen.om.een.,  
    Wettelijke_bedenktijd = Vindt.u.dat.de.wettelijke.bedenktijd.moet.terugkomen.,  
    Adoptie_als_alternatief = Vindt.u.dat.adoptie.een.alternatief.is.voor.abortus.,  
    Biologische_vader_zeggenschap = Moet.de.biologische.vader.zeggenschap.hebben.over.de.beslissing  
  ) %>%  
  select(- c(Tijd.van.laatste.wijziging, E.mail, Naam))  
  
# head(data_renamed)
```

(later “Anders” en “Wil ik liever niet zeggen” toevoegen!) factor gender toevoegen.

```
library(dplyr)  
data_renamed$Gender <- factor(x = data$gender,  
  levels = c("Vrouw", "Man"),  
  labels = c("female", "male"))
```

Vragen bovenste stukje tekst!: - waarom gender ineens zonder hoofdletter?? - is 'x = Data\$gender' overbodig?? er gaat iets fout.

Oefenen: Gemiddelde leeftijd enquêtes berekend en cat():

```
(gemiddelde_leeftijd <- mean(data_renamed$Leeftijd))
```

```
## [1] 40.86842
```

```
cat(sprintf("de gemiddelde leeftijd van de enquête is: %.2f ik ben iris %.3f/n",
            gemiddelde_leeftijd, gemiddelde_leeftijd))
```

```
## de gemiddelde leeftijd van de enquête is: 40.87 ik ben iris 40.868/n
```

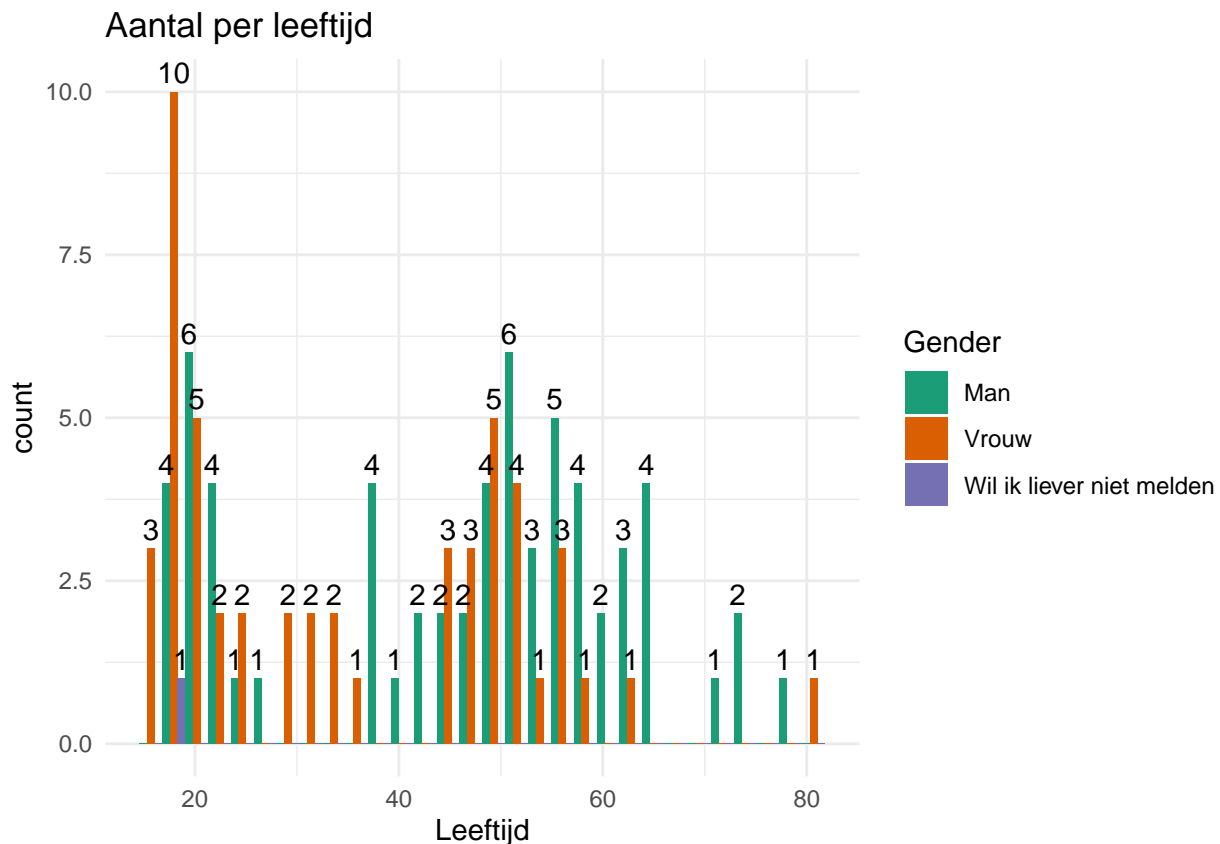
In bovenstaande code zie je dat je met `mean()` het gemiddeld berekent van alle getallen die in de kolom leeftijd staan. Verder kun je met `cat(sprintf())` een variabele in je tekst zetten. Door middel van `'%.2f'` (rond het getal af op 2 decimalen). Hiermee kan je makkelijker variabelen in je tekst plaatsen.

Oefenen histogram maken:

```
library(ggplot2)
ggplot(data_renamed, aes(x=Leeftijd, fill = Gender)) +
  labs(title = "Aantal per leeftijd") +
  scale_fill_brewer(palette = "Dark2") +
  geom_histogram(position = "dodge") +
  stat_bin(aes(label = ifelse(after_stat(count)>0, after_stat(count), "")),
          geom = "text", position = "dodge", vjust = -.4) +
  theme_minimal()
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



In het bovenstaande histogram is er weergegeven hoeveel mensen de enquête hebben ingevuld per leeftijd. 'x = leeftijd' neemt de leeftijd als de x-as. 'fill = Gender' zorgt ervoor dat de gender verdeelt wordt in

twee balkjes per leeftijd. De man en de vrouw. ‘scale_fill_brewer(palette = “Dark2”)’ zorgt ervoor dat de kleuren van de balkjes veranderen naar het gekozen palette: ‘Dark2’ ‘geom_histogram(position = “dodge”)’ maakt het histogram en de position zorgt ervoor dat de balkjes vrouw en man naast elkaar komen te staan ipv boven elkaar. stat_bin(aes(label = ifelse(after_stat(count)>0, after_stat(count),“”), geom = “text”, position = “dodge”, vjust = -.4) deze regel zorgt ervoor dat als er 0 vrouwen of mannen van een leeftijd zijn, dat er dan niet ‘0’ komt te staan in de plot. En als er meer dan 0 mensen zijn dan komt er een aantal boven het balkje te staan. De ‘vjust’ zorgt ervoor dat het aantal niet overlapt met het balkje.

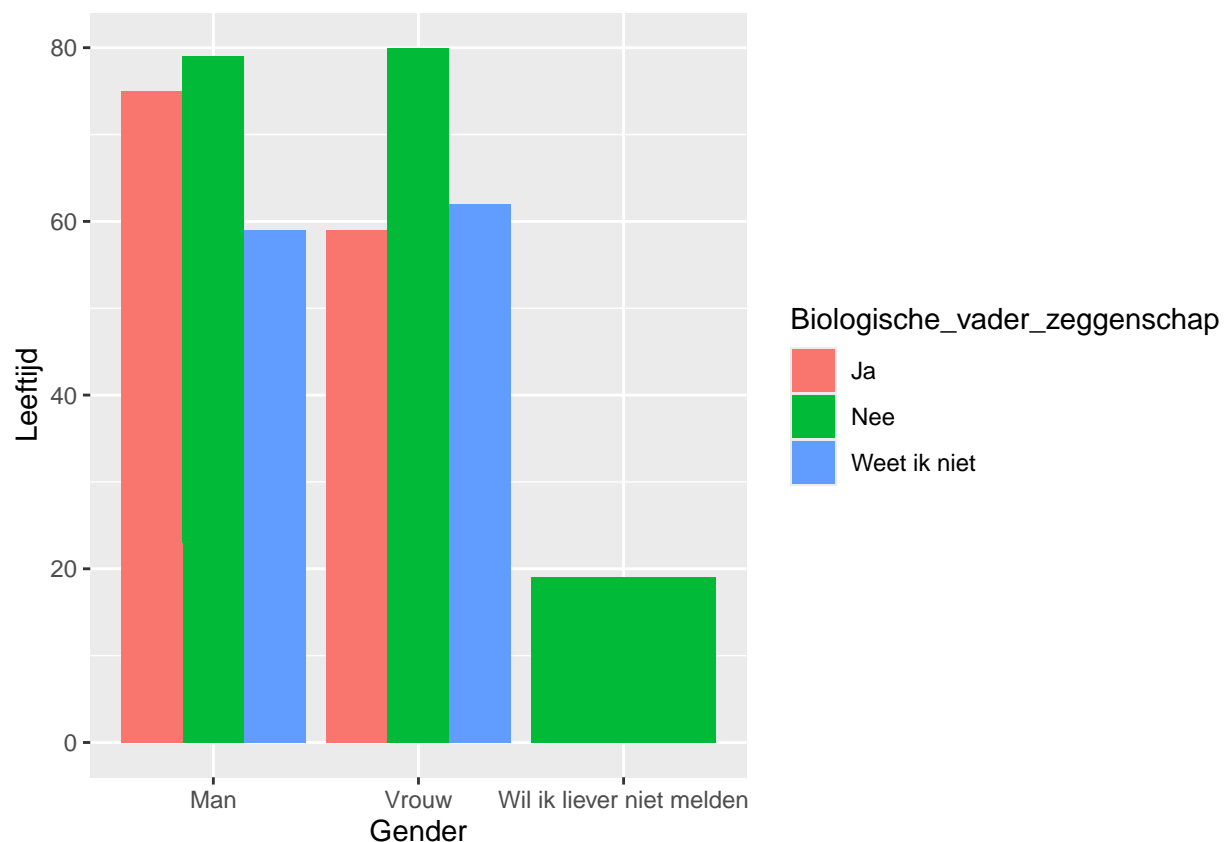
Donderdag 16 mei

kleuren palletten die ook leesbaar is voor mensen die kleurenblind zijn: [http://www.cookbook-r.com/Graphs/Colors_\(ggplot2\)/](http://www.cookbook-r.com/Graphs/Colors_(ggplot2)/)

“Moet de biologische vader zeggenschap hebben over de beslissing voor een abortus?” Hoeveel van deze antwoorden komen van mannen en hoeveel van vrouwen? Zou het zo zijn dat mannen hier eerder ‘Ja’ zullen antwoorden dan vrouwen? (gebruik kolommen: Gender en Biologische_vader_zeggenschap) (Proberen met stacked barplot, y-as: aantal antwoorden, x-as = Mannen, Vrouwen, legenda/kleuren balkjes: Ja, Nee, Weet ik niet)

```
library(ggplot2)

ggplot(data_renamed, aes(fill=Biologische_vader_zeggenschap, x=Gender, y=Leeftijd)) +
  geom_bar(position="dodge", stat="identity")
```



y goed invullen! hoe uitzoeken? Deze plot is nog niet logisch, oefenen.

Vrijdag 17 mei

fill is een soort split. Dus de fill moet de antwoorden worden van de vraag: ‘Moet de biologische vader zeggenschap hebben over de beslissing voor een abortus?’

Ronald zei net terecht dat het aantal mannen en vrouwen niet gelijk zijn. Dus de plot zal niet representatief zijn als ik de data zo invul. Ik kan het gemiddelde nemen van de mannen en de vrouwen per antwoord-optie (dus mannen die ja hebben ingevuld, vrouwen die ja hebben ingevuld, mannen die nee hebben ingevuld, vrouwen die nee hebben ingevuld, mannen die ik weet niet hebben ingevuld vrouwen die ik weet niet hebben ingevuld). hiermee kun ik volgens mij vervolgens een plot maken zoals ik eerst had bedacht. Dan neem je de y-as als gemiddelde aantal antwoorden in plaats van antwoorden.

H1:Er zit een significant verschil in mannen en vrouwen hun mening over het zeggenschap van de vader.

H0:Er zit geen significant verschil in mannen en vrouwen hun mening over het zeggenschap van de vader.

Maakt een tabel van de vraag(biologische vader zeggenschap) en de gender:

```
(tabel <- table(data_renamed$Gender, data_renamed$Biologische_vader_zeggenschap))
```

```
##
##                Ja Nee Weet ik niet
##   Man                27  32         3
##   Vrouw               8  36         7
##   Wil ik liever niet melden  0   1         0
```

Maakt van de tabel hierboven proportie:

```
(prop_tabel <- prop.table(table(data_renamed$Gender, data_renamed$Biologische_vader_zeggenschap)))
```

```
##
##                Ja      Nee Weet ik niet
##   Man          0.23684211 0.28070175  0.02631579
##   Vrouw         0.07017544 0.31578947  0.06140351
##   Wil ik liever niet melden 0.00000000 0.00877193  0.00000000
```

De chisq-test is om te kijken of er een significant verschil zit tussen de mannen en vrouwen in antwoorden:

```
(chisq.test(tabel))
```

```
## Warning in chisq.test(tabel): Chi-squared approximation may be incorrect
```

```
##
##   Pearson's Chi-squared test
##
## data:  tabel
## X-squared = 11.917, df = 4, p-value = 0.01798
```

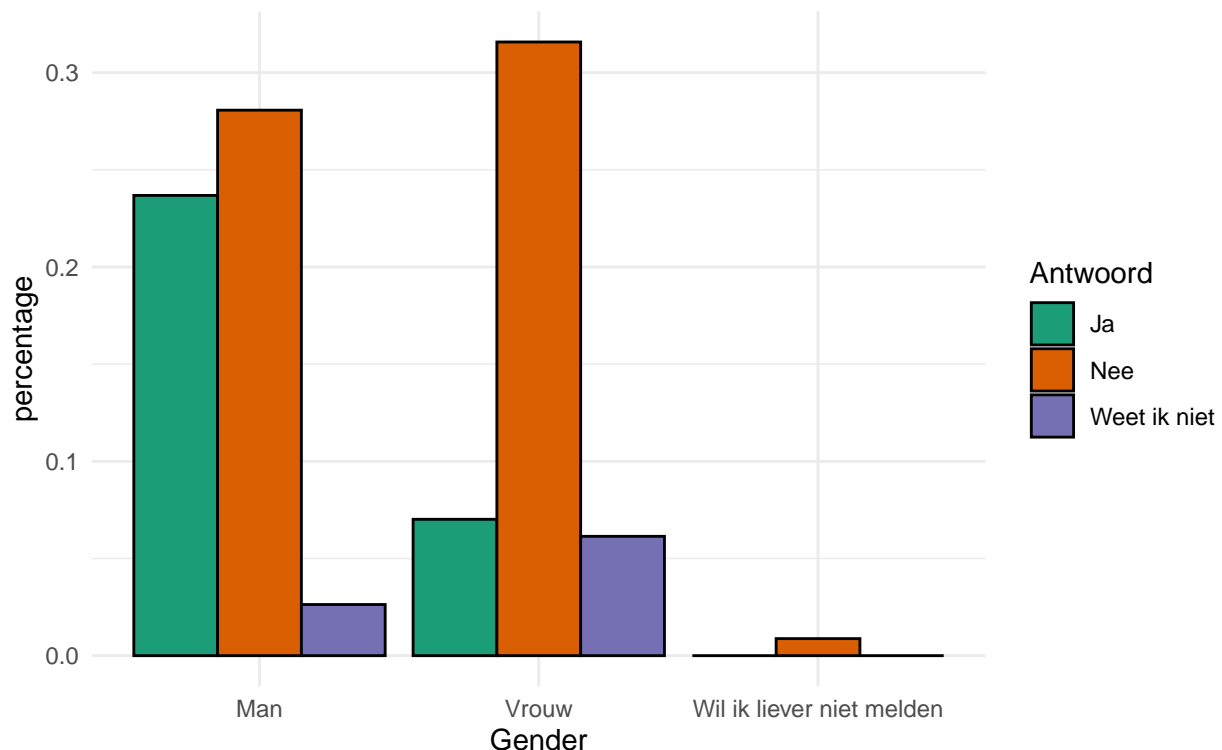
De p-value die hieruit komt is ‘0.01798’. Dit is lager dan 0.05, dus is er een significant verschil aangetoond. Mannen en vrouwen denken dus anders over of de vader zeggenschap heeft in de keuze over een abortus.

```
library(ggplot2)
```

```
ggplot(as.data.frame(prop_tabel), aes(x = Var1, y=Freq, fill=Var2)) +  
  labs(title="Barplot verdeling gender over antwoorden",  
        subtitle = "Moet de biologische vader zeggenschap hebben over de beslissing voor een abortus?",  
        x = "Gender",  
        fill = "Antwoord",  
        y = "percentage") +  
  geom_bar(stat = "identity", position = "dodge", color="black")+  
  theme_minimal() +  
  scale_fill_brewer(palette = "Dark2")
```

Barplot verdeling gender over antwoorden

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



In de bovenstaande plot zie je weergegeven per gender hoeveel procent er wat heeft geantwoord. Je ziet hier dat er gemiddeld meer vrouwen zijn die nee hebben geantwoord. Ook zie je dat er meer mannen ja hebben geantwoord dan vrouwen. Dit kan komen doordat vrouwen zich waarschijnlijk eerder inleven in de vrouw en mannen eerder in de man. Hierdoor zie je dat mannen eerder zullen zeggen dat zij vinden dat de vader ook inspraak moeten hebben.

Dinsdag 21 mei

Inlezen power analyse; power wordt bepaalt door steekproefomvang(hoeveel antwoorden nodig), effectgrootte (verband tussen twee variabelen, bijvoorbeeld verschil in gemiddelden) en significantie niveau (hoe groot is de kans dat een resultaat door toeval is bepaalt). Met een poweranalyse bepaal je de steekproefomvang(dus hoeveel antwoorden er nodig zijn voor een betrouwbaar resultaat). Powerniveau moet boven de 80% zitten.

Voor een poweranalyse zijn drie factoren nodig om de vierde te berekenen. Statistische power(meestal 80% of hoger) Steekproefomvang(de onbekende) Significantieniveau(meestal 5% (of 1%)) Verwachte effectgrootte(?), zelf inschatten dmv eerder onderzoek)

Door het significantie niveau tussen de 5% en de 1% te veranderen heb je invloed op de kans op een type-I fout en een type-II fout. (BRON: Bhandari, P. (2022, 1 september). Statistische power (statistical power) en poweranalyses. Scribbr. <https://www.scribbr.nl/statistiek/statistische-power/>) (In Floris zijn logboek is een power analyse uitgevoerd.)

Ideeën plot voor morgen: - vraag 13: ‘Vindt u dat de wettelijke bedenktijd moet terugkomen?’ vergelijken met ‘gender’ of ‘religieuze overtuiging’

- vraag 10 en vraag 11 vergelijken: ‘Vindt u dat een minderjarig kind de keuze moet hebben om voor een abortus te kiezen?’ ‘Vindt u dat de ouders zeggenschap moeten hebben over of hun minderjarige kind wel of niet voor abortus mag kiezen?’ Zullen mensen die bij vraag 10 ‘Ja’ invullen dan ook ‘Nee’ invullen bij vraag 11? Je kunt namelijk ook zeggen dat een minderjarig kind de keuze moet hebben, maar dat deze keuze dan vervolgens bij de ouders ligt.
- vraag 6: ‘Vindt u dat iemand zelf moet kunnen kiezen of die een abortus wil?’ Vergelijken met geloof. Zal iemand sneller ‘Nee’ zeggen als deze gelooft?

Woensdag 22 mei

Uit eerder onderzoek blijkt dat geloof ervoor zorgt dat de partner de keuze voor abortus minder snel zal steunen. Komt dezelfde conclusie uit onze enquête?

H1: Geloof heeft invloed op de mening over abortus.

H0: Geloof heeft geen invloed op de mening over abortus.

Hoeveel procent van de mensen die gelooft antwoord bij vraag 6: ‘Nee’ uitzoeken:

- hoeveel procent gelooft (selecteer op ‘Niet gelovig’ en trek dat vervolgens af van het geheel)
- hoeveel procent antwoord nee

```
(tabel_geloof_vraag6 <- (table(data_renamed$Abortus_keuze, data_renamed$Religie)))
```

```
##
##      Anders Boeddhisme Christendom  Islam Niet gelovig
##  Ja         6           1           8     2           93
##  Nee         1           0           3     0            0
```

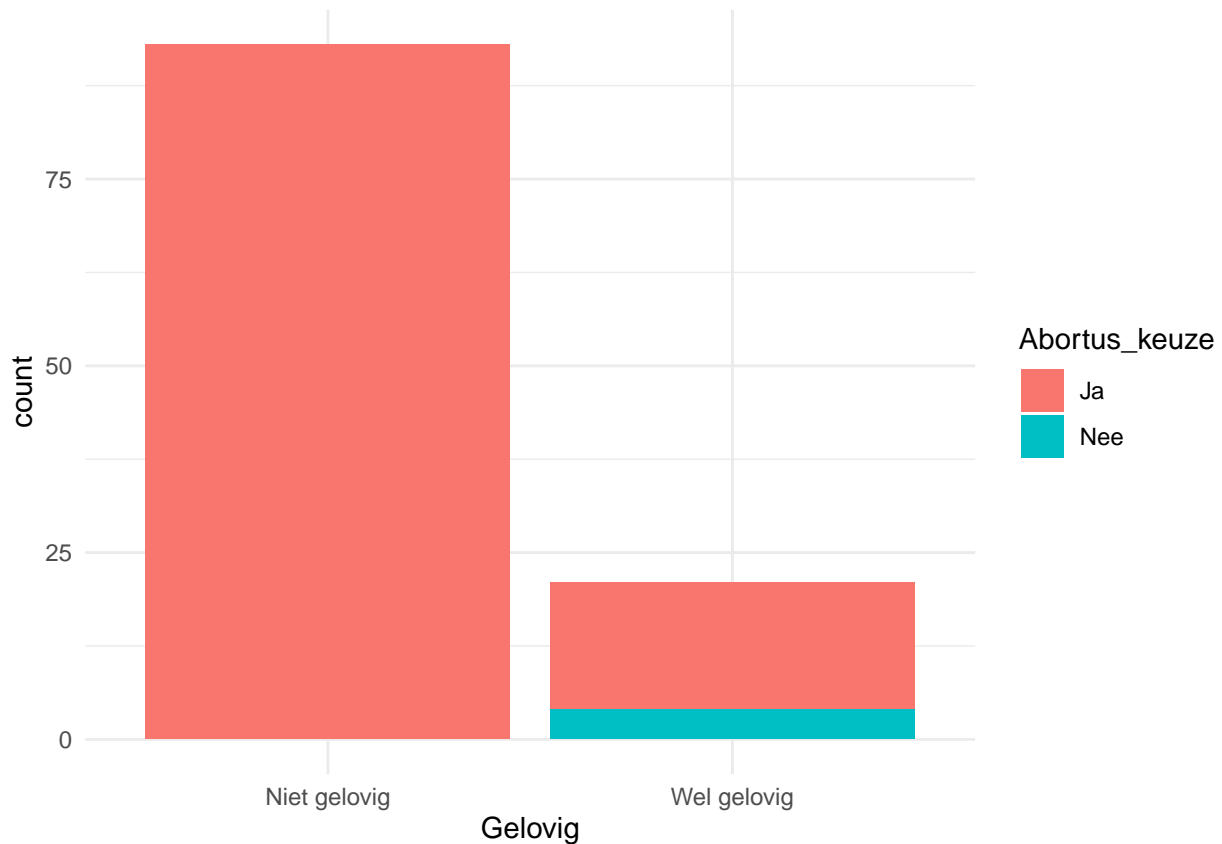
In de tabel hierboven zie je hoeveel mensen die ja/nee hebben geantwoord bij vraag 6 (Vindt u dat iemand zelf moet kunnen kiezen of die een abortus wil?) een religieuze overtuiging heeft.

```
(proptabel_geloof_vraag6 <- prop.table(table(data_renamed$Abortus_keuze, data_renamed$Religie)))
```

```
##
##      Anders Boeddhisme Christendom      Islam Niet gelovig
##  Ja 0.05263158 0.00877193 0.07017544 0.01754386 0.81578947
##  Nee 0.00877193 0.00000000 0.02631579 0.00000000 0.00000000
```

Uit de tabel hierboven kun je zien dat ongeveer 82% van de mensen die de enquête hebben ingevuld niet gelooft. Al deze mensen hebben bij vraag 6 ingevuld dat je zelf voor een abortus mag kiezen.


```
data_renamed %>%
  select(Religie, Abortus_keuze) %>%
  mutate(Gelovig=ifelse(Religie == "Niet gelovig", "Niet gelovig", "Wel gelovig")) %>%
  ggplot(aes(x=Gelovig, fill=Abortus_keuze)) +
  geom_bar() +
  theme_minimal()
```

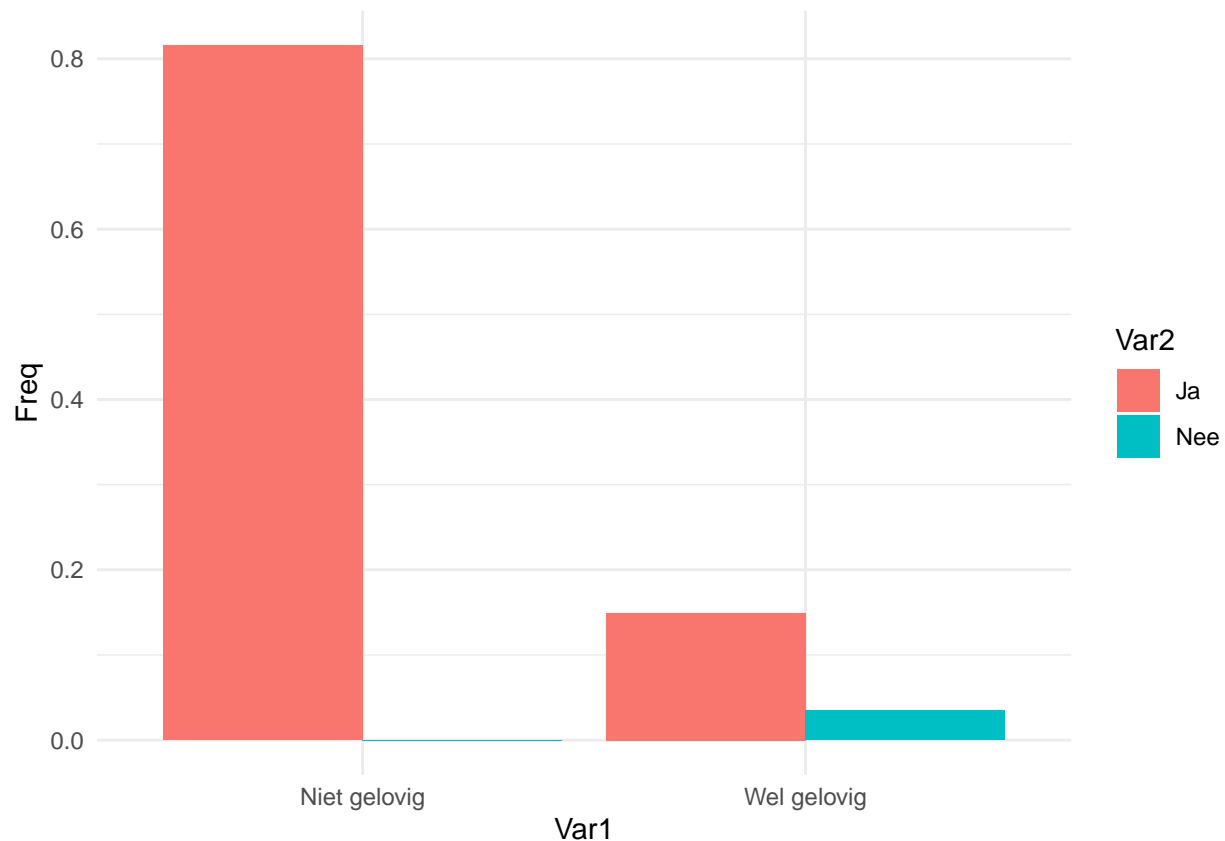


Dit is een mooie plot, maar de proportie klopt nu nog niet.

```
tewmp <- data_renamed %>%
  select(Religie, Abortus_keuze) %>%
  mutate(Gelovig=ifelse(Religie == "Niet gelovig", "Niet gelovig", "Wel gelovig"))
x <- prop.table(table(tewmp$Gelovig, tewmp$Abortus_keuze))
x
```

```
##
##              Ja      Nee
## Niet gelovig 0.81578947 0.00000000
## Wel gelovig  0.14912281 0.03508772
```

```
as.data.frame(x)%>%
  ggplot(aes(x = Var1, y = Freq, fill = Var2)) +
  geom_col(position = "dodge") +
  theme_minimal()
```



In de bovenstaande plot zie je de verdeling weergegeven van mensen die wel geloven en niet geloven en wat deze hebben geantwoord op de vraag of mensen zelf mogen kiezen of ze een abortus willen.

Donderdag 23 mei

Verbeteren plot van gisteren:

```
tijdelijke_data <- data_renamed %>%
  select(Religie, Abortus_keuze) %>%
  mutate(Gelovig=ifelse(Religie == "Niet gelovig", "Niet gelovig", "Wel gelovig"))
x <- prop.table(table(tijdelijke_data$Gelovig, tijdelijke_data$Abortus_keuze))
x
```

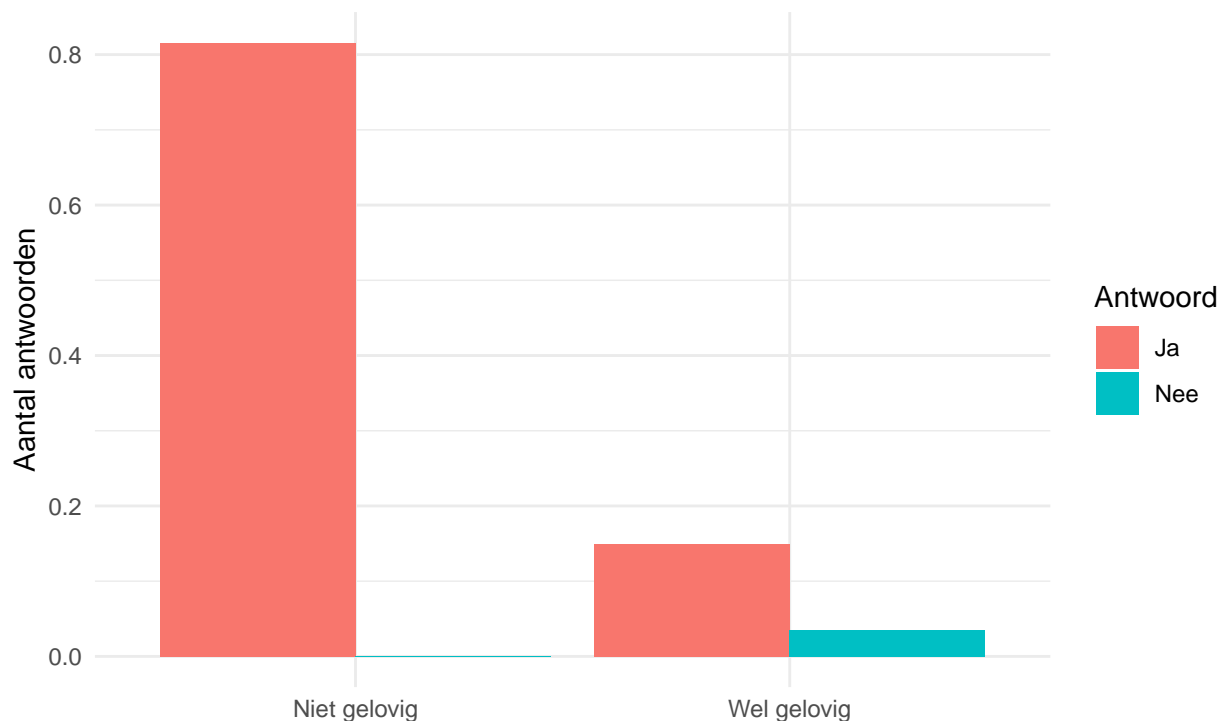
```
##
##              Ja      Nee
## Niet gelovig 0.81578947 0.00000000
## Wel gelovig  0.14912281 0.03508772
```

```
as.data.frame(x)%>%
  ggplot(aes(x = Var1, y = Freq, fill = Var2)) +
  labs(title = "Corrolatie religie en mening abortus",
       subtitle = "Vindt u dat iemand zelf moet kunnen kiezen of die een abortus wil?",
       x = "",
       fill = "Antwoord",
       y = "Aantal antwoorden") +
```

```
geom_col(position = "dodge") +  
theme_minimal()
```

Corrolatie religie en mening abortus

Vindt u dat iemand zelf moet kunnen kiezen of die een abortus wil?



Je ziet in de plot hierboven dat alle mensen die hebben aangegeven niet een religieuze overtuiging te hebben bij vraag 6 ook hebben aangegeven dat iemand zelf moet kunnen kiezen of die een abortus wil. Bij de mensen die hebben ingevuld wel een religieuze overtuiging te hebben is er een deel die heeft aangegeven dat iemand niet zelf moet kunnen kiezen of die een abortus wil. Dus alle mensen die tegen vrije keuze over abortus zijn hebben een religieuze overtuiging. Dit is opvallend.

Achtergrond informatie geloof en abortus: Volgens het christendom is er vanaf het moment van de bevruchting nieuw leven ontstaan. Hierom zijn er veel christenen die tegen abortus zijn. In 'tabel_geloof_vraag6' (zie hieronder nog eens weergegeven) zie je dat van de 11 christenen er 3 tegen abortus zijn. In de koran (bevat de regels van de islam) wordt abortus niet expliciet verboden. Wel wordt er in omschreven dat het verboden is om abortus te plegen zodra er sprake is van leven. Volgens de Islam is dit 120 dagen na de bevruchting. Dit is rond de vier maanden. Een abortus mag je plegen tot de 24e week na de bevruchting. Dit is rond de zes maanden. Dus volgens de islam mag je wel abortus plegen, als dit binnen de vier maanden valt. In de tabel hieronder geven de twee moslims aan tegen abortus te zijn.

tabel_geloof_vraag6

```
##  
##      Anders Boeddhisme Christendom  Islam Niet gelovig  
##   Ja      6          1          8    2          93  
##   Nee      1          0          3    0           0
```

volgens het hindoeïsme mag een vrouw geen abortus plegen, tenzij het leven van de vrouw in gevaar loopt. Volgens het boeddhisme is er niet een voor of tegen abortus. Sommigen boeddhisten zeggen dat je bij abortus

iemands leven neemt en dat de medische ingrijp iemand buiten de kringloop van iemand zijn leven plaatsen. Maar er zijn ook boeddhisten die abortus goed keuren onder bepaalde voorwaarden. Zolang de betrokkenen goed nadenken over de gevolgen van hun keuze.

Vrijdag 24 mei

Op aanraden van Michiel ga ik proberen de 'Barplot verdeling gender over antwoorden' aan te passen. Nu zijn man en vrouw samen honderd procent, maar het is waarschijnlijk beter af te lezen als 'man' honderd procent is en 'vrouw' ook honderd procent.

Data frame maken van de prop_tabel:

```
new_proptabel_dataframe <- as.data.frame(prop_tabel) %>%
  rename(
    gender = Var1,
    antwoord_vraag = Var2,
    aantal = Freq)
```

Het juist verdelen van de procenten:

```
prop_tabel

##
##              Ja              Nee Weet ik niet
##  Man          0.23684211 0.28070175  0.02631579
##  Vrouw        0.07017544 0.31578947  0.06140351
##  Wil ik liever niet melden 0.00000000 0.00877193  0.00000000
```

```
new_proptabel_dataframe <- as.data.frame(prop_tabel) %>%
  rename(
    gender = Var1,
    antwoord_vraag = Var2) %>%
  group_by(gender, antwoord_vraag) %>%
  summarize(count = n(), .groups = "drop") %>%
  group_by(gender) %>%
  mutate(perc = count / sum(count) * 100)
```

Hoe ziet de tabel eruit?:

```
new_proptabel_dataframe

## # A tibble: 9 x 4
## # Groups:   gender [3]
##   gender      antwoord_vraag count  perc
##   <fct>      <fct>          <int> <dbl>
## 1 Man      Ja              1    33.3
## 2 Man      Nee              1    33.3
## 3 Man      Weet ik niet      1    33.3
## 4 Vrouw   Ja              1    33.3
## 5 Vrouw   Nee              1    33.3
## 6 Vrouw   Weet ik niet      1    33.3
## 7 Wil ik liever niet melden Ja              1    33.3
## 8 Wil ik liever niet melden Nee              1    33.3
## 9 Wil ik liever niet melden Weet ik niet      1    33.3
```

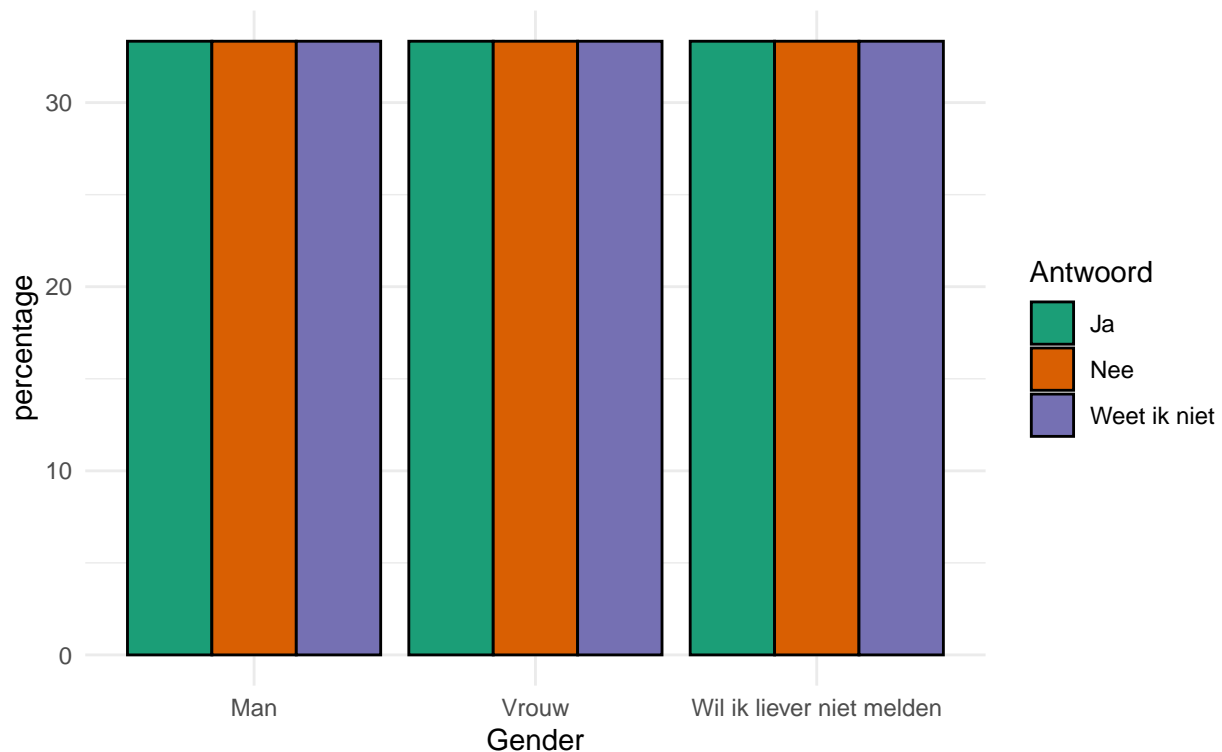
Nieuwe dataframe verwerken in plot:

```
library(ggplot2)

ggplot(new_proptabel_dataframe, aes(x = gender, y=perc, fill=antwoord_vraag)) +
  labs(title="Barplot verdeling gender over antwoorden",
       subtitle = "Moet de biologische vader zeggenschap hebben over de beslissing voor een abortus?",
       x = "Gender",
       fill = "Antwoord",
       y = "percentage") +
  geom_bar(stat = "identity", position = "dodge", color="black")+
  theme_minimal() +
  scale_fill_brewer(palette = "Dark2")
```

Barplot verdeling gender over antwoorden

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



In de bovenstaande plot gaat nog iets fout nu, maar ik snap niet helemaal wat. Waarschijnlijk iets met de y-waarde.

Zondag 26 mei

Aan de presentatie gewerkt voor de Donut chart. Powerpoint gemaakt en voordeel opgeschreven (makkelijk te begrijpen).

Maandag 27 mei

Het gaat waarschijnlijk fout in het maken van de tabel. Morgen even naar kijken.

```
new_proptabel_dataframe
```

```
## # A tibble: 9 x 4
## # Groups:   gender [3]
##   gender          antwoord_vraag count  perc
##   <fct>          <fct>          <int> <dbl>
## 1 Man           Ja                1  33.3
## 2 Man           Nee                1  33.3
## 3 Man           Weet ik niet          1  33.3
## 4 Vrouw         Ja                1  33.3
## 5 Vrouw         Nee                1  33.3
## 6 Vrouw         Weet ik niet          1  33.3
## 7 Wil ik liever niet melden Ja                1  33.3
## 8 Wil ik liever niet melden Nee                1  33.3
## 9 Wil ik liever niet melden Weet ik niet          1  33.3
```

De percentages kloppen inderdaad niet in de tabel.

Begin maken met intro voor bovenaan logboek: In dit onderzoek proberen wij te onderzoeken of er factoren zijn die invloed hebben op iemand zijn mening over abortus. We hebben een enquête met vragen gemaakt en hebben deze vervolgens rondgestuurd. De resultaten die uit deze enquête kwam hebben we met behulp van Rstudio geanalyseerd. Hierin hebben we onderzocht of factoren zoals; geslacht, onderwijsniveau, leeftijd, religieuze overtuiging en de provincie waarin iemand is opgegroeid invloed hebben op hoe mensen de vragen in de enquête hebben beantwoord.

Structuur wetenschappelijk onderzoek:

- inleiding
- materiaal en methode
- resultaten
- discussie/ conclusie

Hoe zit de structuur in elkaar:

- Wat moet er allemaal in de inleiding?: Begin met een aandachtstrekker, waarom doen we dit onderzoek. Omschrijf wat we hebben onderzocht. Als er al kennis is over dit onderwerp, kaart dan eerdere onderzoeken aan. Omschrijf wat de maatschappelijke relevantie is van dit onderzoek. Wat zijn de hypothesen. Wat voor soort onderzoek doen we: hoe verzamelen we data en hoe analyseren we dat vervolgens? (eventueel de structuur van de paper)
- Wat moet er in materiaal en methode?: (Vaak in verleden tijd geschreven)(wanneer? hoelang? hoeveel?) Hoe hebben we die enquête gemaakt en met welk programma. Hoe hebben we de resultaten verwerkt: welk programma gebruikt (grafieken meegeven), hoe betrouwbaar is de data?
- Wat moet er in de resultaten? Doornemen nuttige resultaten. Dus de informatie die antwoord geeft op de onderzoeksvraag (De resultaten moeten je conclusie onderbouwen) (selecteer eerst welke tabellen en figuren je wilt bespreken en schrijf er daarna een tekst omheen).
- Wat moet er in de Discussie/ conclusie? Interpretatie van de resultaten. Hierin zul je de onderzoeksvraag beantwoorden en aangeven of de hypothese klopt (waarom wel/niet?). Komt het resultaat overeen met eerdere onderzoeken? Als de resultaten niet voldoen aan onze verwachtingen, leg dan uit of hier verklaringen voor zijn. Heeft dit onderzoek bijgedragen aan de maatschappij? Wat zou je meegeven voor vervolgonderzoek?

```
new_proptabel_dataframe <- data_renamed %>%
  group_by(Gender, Biologische_vader_zeggenschap) %>%
  summarize(count = n(), .groups = "drop") %>%
  group_by(Gender) %>%
  mutate(perc = count/sum(count) * 100)
```

```
new_proptabel_dataframe
```

```
## # A tibble: 7 x 4
## # Groups:   Gender [3]
##   Gender      Biologische_vader_zeggenschap count  perc
##   <chr>      <chr>                        <int> <dbl>
## 1 Man        Ja                        27  43.5
## 2 Man        Nee                       32  51.6
## 3 Man        Weet ik niet              3   4.84
## 4 Vrouw      Ja                        8  15.7
## 5 Vrouw      Nee                       36  70.6
## 6 Vrouw      Weet ik niet              7  13.7
## 7 Wil ik liever niet melden Nee                        1  100
```

Het lag inderdaad aan de tabel. Michiel heeft me hierbij geholpen. Morgen de grafiek aanpassen.

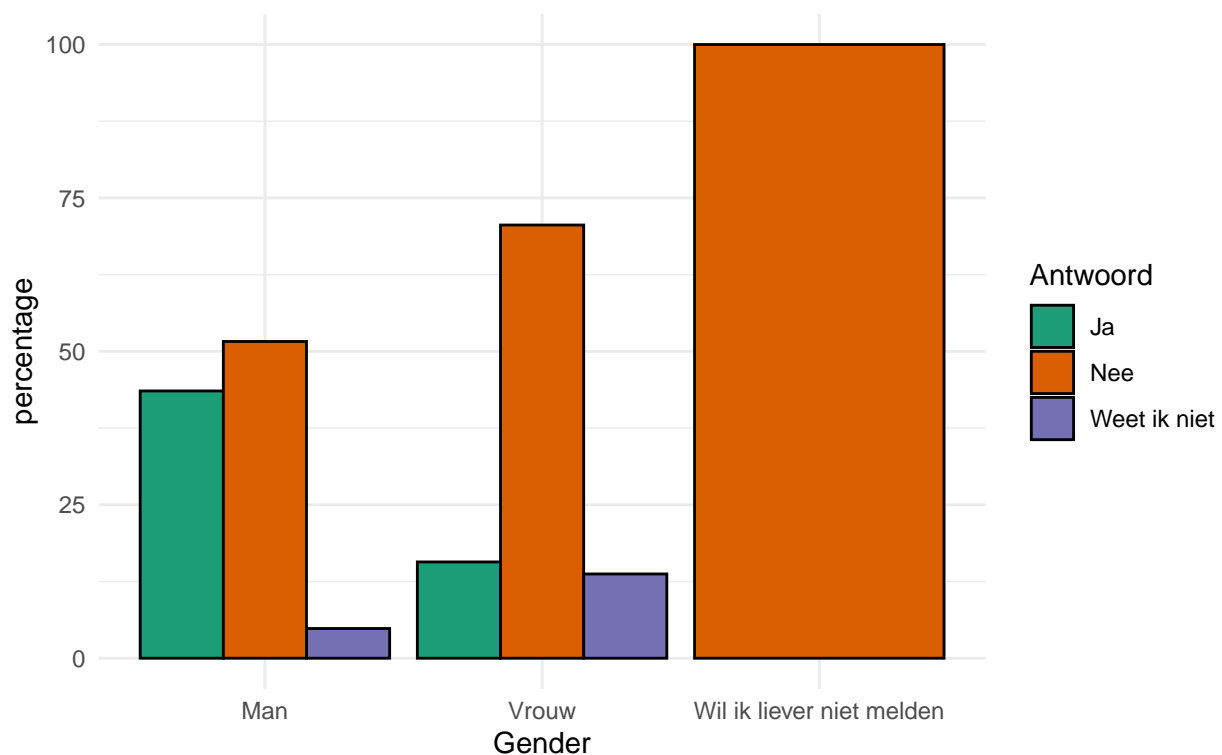
Dinsdag 28 mei

```
library(ggplot2)

ggplot(new_proptabel_dataframe, aes(x =Gender, y=perc, fill=Biologische_vader_zeggenschap)) +
  labs(title="Barplot verdeling gender over antwoorden",
        subtitle = "Moet de biologische vader zeggenschap hebben over de beslissing voor een abortus?",
        x = "Gender",
        fill = "Antwoord",
        y = "percentage") +
  geom_bar(stat = "identity", position = "dodge", color="black")+
  theme_minimal() +
  scale_fill_brewer(palette = "Dark2")
```

Barplot verdeling gender over antwoorden

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



Nu is de grafiek zo dat de antwoorden van de mannen samen honderd procent is en dat de antwoorden van de vrouwen ook samen honderd procent is en dat de grafiek van de mensen die het liever niet wouden zeggen ook 100% is. Nu is de grafiek iets minder misleidend, omdat het nu in proportie is en er iets meer mannen dan vrouwen de enquête hebben ingevuld.

Gemiste kans: Als we het onderzoek opnieuw zouden doen, zouden we een vraag moeten toevoegen die vraagt of mensen een dochter hebben. Dan kunnen we onderzoeken of (voornamelijk) mannen met een dochter sneller voor abortus zullen zijn dan mannen zonder een dochter.

Pie-donut variatie proberen: Ik ga de tabel over de religieuze overtuiging en of mensen voor of tegen de eigen keuze over abortus zijn opnieuw gebruiken hier.

```
data_donut <- as.data.frame(tabel_geloof_vraag6) %>%
  rename(
    antwoord_vraag = Var1,
    religieuze_overtuiging = Var2)
data_donut
```

##	antwoord_vraag	religieuze_overtuiging	Freq
## 1	Ja	Anders	6
## 2	Nee	Anders	1
## 3	Ja	Boeddhisme	1
## 4	Nee	Boeddhisme	0
## 5	Ja	Christendom	8
## 6	Nee	Christendom	3
## 7	Ja	Islam	2
## 8	Nee	Islam	0

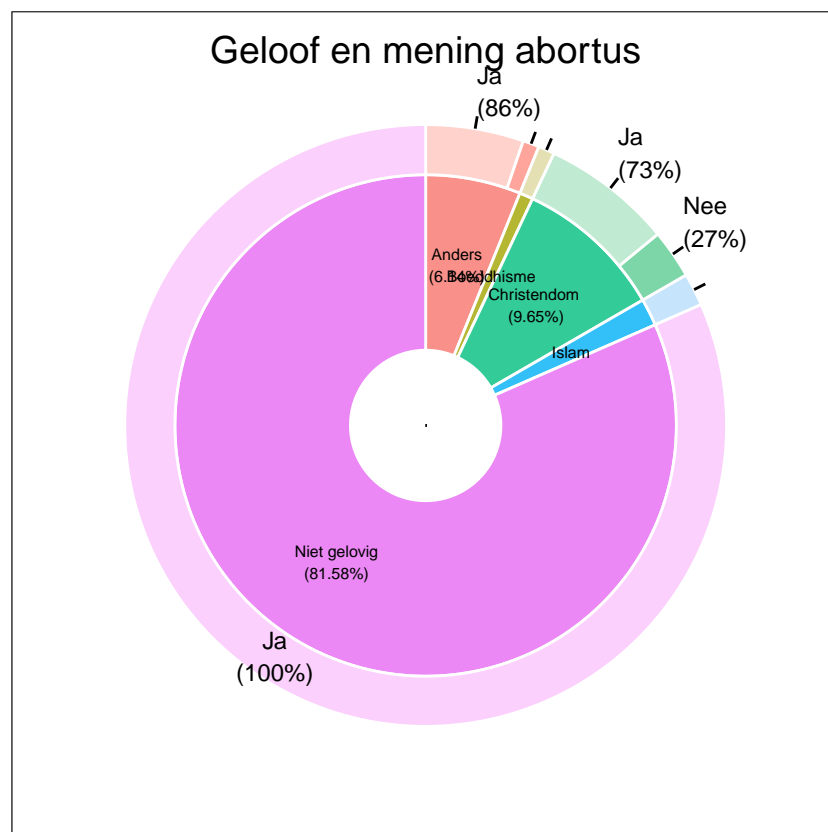

```
## 9          Ja          Niet gelovig  93
## 10         Nee          Niet gelovig   0
```

“webr” installed:

```
library(ggplot2)
library(webR)
library(dplyr)

PieDonut(data_donut, aes(religieuze_overtuiging, antwoord_vraag, count=Freq),
  title = "Geloof en mening abortus",
  showRatioDonut = TRUE, # percentage buitenste cirkel (mening abortus)
  showRatioPie = TRUE, # percentage in de binnenste cirkel (geloof)
  showPieName = FALSE, # haalt de enorme naam in het midden weg, mooier
  pieLabelSize = 2 # grootte tekst binnenste ring (geloof)
)
```

```
## Warning: The '<scale>' argument of 'guides()' cannot be 'FALSE'. Use "none" instead as
## of ggplot2 3.3.4.
## i The deprecated feature was likely used in the webR package.
## Please report the issue at <https://github.com/cardiomoan/webR/issues>.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

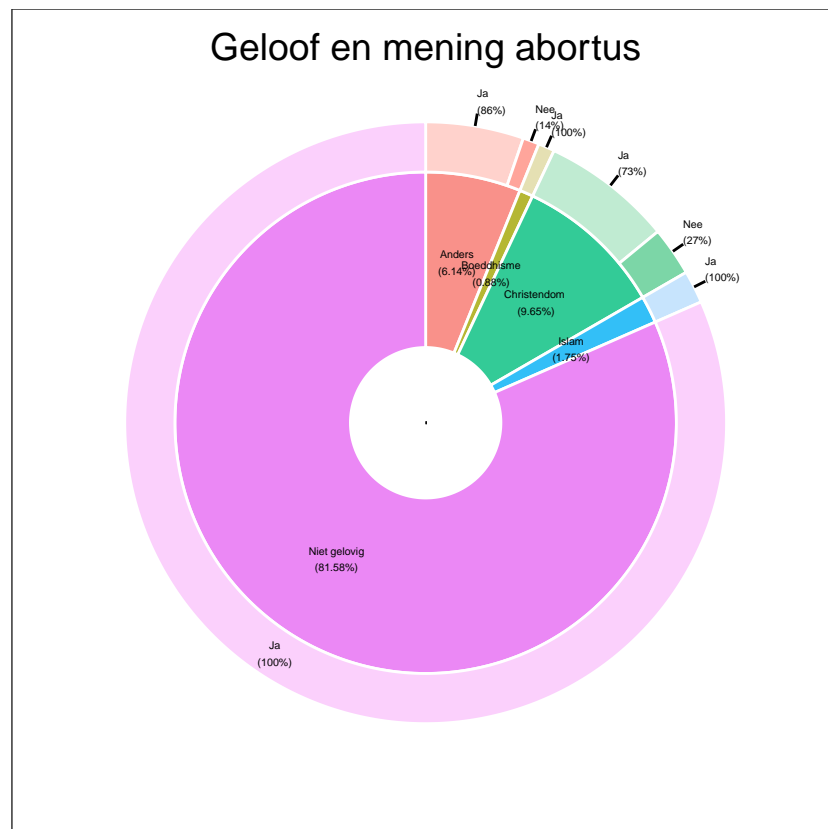


Je ziet in de bovenstaande PieDonut dat van de mensen die geantwoord hebben op de vraag(binnenste ring)

er 81.58% niet gelovig is, 9.65% christelijk, 6.14% anders en de Islam en boeddhisme geeft hij nog niet goed weer. Ik vermoed dat het vak misschien te klein is om het percentage weer te geven? in de buitenste ring zie je dat van de niet gelovige mensen 100% vindt dat abortus een eigen keuze moet zijn en zo voor elk geloof. Ook hier wordt er in de kleinere vakken niet het percentage weergegeven.

```
library(ggplot2)
library(webr)
library(dplyr)

PieDonut(data_donut, aes(religieuze_overtuiging, antwoord_vraag, count=Freq),
  title = "Geloof en mening abortus",
  showRatioDonut = TRUE, # percentage buitenste cirkel (mening abortus)
  showRatioPie = TRUE, # percentage in de binnenste cirkel (geloof)
  showPieName = FALSE, # haalt de enorme naam in het midden weg, mooier
  pieLabelSize = 1.5, # grootte tekst binnenste ring (geloof)
  donutLabelSize = 1.5,
  showRatioThreshold = getOption("PieDonut.showRatioThreshold", 0.0) # percentages gefikst
)
```



In de PieDonut functie staat in de default blijkbaar dat hij vanaf 2% de label niet weergeeft. Deze heb ik op 0 gezet en nu geeft hij alle percentages mooi weer.

In de bovenstaande plot wordt er in de binnenste plot weergegeven hoeveel procent van de respondenten welke religieuze overtuigingen hebben. 81.58% van de respondenten geeft aan niet een religieuze overtuiging te hebben. Verder geeft 9.65% aan in het christendom te geloven, 1.75% in de islam, 0.88% in het boeddhisme en 6.14% geeft aan een religieuze overtuiging te hebben die wij niet als optie hebben meegegeven in de enquête. Vervolgens wordt er in de buitenste ring weergegeven of de respondenten voor of tegen de eigen

keuze over abortus zijn per geloof. Dus van de 81.58% mensen die niet een religieuze overtuiging hebben is 100% voor de eigen keuze over abortus. Verder is van de Christelijke respondenten 73% voor de eigen keuze over abortus en 27% tegen abortus. Van de islamitische en de Boeddhistische respondenten is 100% voor de eigen mening over abortus. En van de respondenten die aangaven ergens anders in te geloven is er 86% voor de eigen keuze van abortus en 14% tegen.

Een oudere versie van de bovenstaande Donut chart staat nu ook als een variatie in onze presentatie.

Woensdag 29 mei

(presentaties waren vandaag) artikel doorlezen (Abortion Attitudes: An Overview of Demographic and Ideological Differences).

Uit dit onderzoek blijkt dat: Er zit geen significant verschil tussen gender en iemands mening over abortus. Uit dit onderzoek blijkt dat religieuze overtuiging ervoor zorgt dat mannen minder support geven aan hun vrouw voor abortus.

Morgen een `chisq.test` doen op de geloof plot.

Donderdag 30 mei

Tabel een naam geven:

```
geloof_tabel_test <- data_renamed %>%
  select(Religie, Abortus_keuze) %>%
  mutate(Gelovig=ifelse(Religie == "Niet gelovig", "Niet gelovig", "Wel gelovig"))
```

Een `chisq.test` moet op cijfers, dus de `geloof_tabel_test` is de verkeerde tabel om te gebruiken in dit geval. Dit is de goede:

```
tabel_geloof_vraag6

##
##      Anders Boeddhisme Christendom  Islam Niet gelovig
##  Ja      6          1          8      2          93
##  Nee     1          0          3      0           0
```

Voer de test erop uit:

```
chisq.test(tabel_geloof_vraag6)

## Warning in chisq.test(tabel_geloof_vraag6): Chi-squared approximation may be
## incorrect

##
##  Pearson's Chi-squared test
##
## data:  tabel_geloof_vraag6
## X-squared = 24.24, df = 4, p-value = 7.149e-05
```

De p-waarde is '7.149e-05'. Dit is lager dan 0.05, dus is er een significant verschil aangetoond?

Maar deze klopt nog niet. De geloven moeten nog samen genomen worden tot een 'wel gelovig'.

```
tabel_geloof_vraag6
```

```
##
##      Anders Boeddhisme Christendom  Islam Niet gelovig
##   Ja      6          1          8      2          93
##   Nee      1          0          3      0          0
```

```
data.frame(tabel_geloof_vraag6)
```

```
##   Var1      Var2 Freq
## 1   Ja      Anders   6
## 2  Nee      Anders   1
## 3   Ja Boeddhisme   1
## 4  Nee Boeddhisme   0
## 5   Ja Christendom   8
## 6  Nee Christendom   3
## 7   Ja      Islam   2
## 8  Nee      Islam   0
## 9   Ja Niet gelovig  93
##10  Nee Niet gelovig   0
```

!!!!!! De tabel moet opgedeeld worden in een kolom met mensen die wel geloven en mensen die niet geloven.

```
poging101 <- data.frame(tabel_geloof_vraag6) %>%
  mutate(gelovig=ifelse(Var2 == "Niet gelovig", resultaat_if, de_else))

#Voorbeeld eerder gebruikt:
#data_renamed %>%
#  select(Religie, Abortus_keuze) %>%
#  mutate(Gelovig=ifelse(Religie == "Niet gelovig", "Niet gelovig", "Wel gelovig"))
```

Het lukt niet.

```
library(tidyr)
# data frame van maken
(tabel_geloof_vraag6 <- (table(data_renamed$Abortus_keuze, data_renamed$Religie)))
tabel_geloof_vraag6 <- data.frame(tabel_geloof_vraag6)
tabel_geloof_vraag6

# tabel maken van dataframe?
tabel_geloof_vraag6 <- tabel_geloof_vraag6 %>%
  pivot_wider(names_from = Var2, values_from = Freq)

# nieuwe kolom maken
tabel_geloof_vraag6$totaal_aantal <- mutate(test=tabel_geloof_vraag6$Anders + tabel_geloof_vraag6$`Christendom` + tabel_geloof_vraag6$Islam + tabel_geloof_vraag6$`Niet gelovig`)
```

Morgen verder.

Vrijdag 31 mei

Ik ga zelf een data frame maken. Zodra er vernieuwde gegevens zijn moeten we dus dit dat frame met de hand aanpassen. (dit is op oude data gedaan, dus nog niet volledig, alleen voor het oefenen)

```
library(dplyr)

# data frame op basis van tabel_geloof_vraag6
data_frame_geloof <- data.frame(
  Anders = c(6, 1),
  Christendom = c(8, 3),
  Islam = c(2, 0),
  Niet_gelovig = c(84, 0)
)
# rij naam veranderen naar 'Ja' en 'Nee'
rownames(data_frame_geloof) <- c("Ja", "Nee")

data_frame_geloof
```

```
##      Anders Christendom Islam Niet_gelovig
## Ja         6           8     2           84
## Nee        1           3     0            0
```

En nu de gekozen samenvoegen tot een.

```
# optellen alle geloven
nieuwe_dataframe_geloof <- tabel_geloof_vraag6 %>%
  mutate(Gelovig = tabel_geloof_vraag6$Anders + tabel_geloof_vraag6$Christendom + tabel_geloof_vraag6$Islam)
# onnodige rijen weghalen
select(-Anders, -Christendom, -Islam)

nieuwe_dataframe_geloof
```

```
table_van_dataframe_geloof <- table(nieuwe_dataframe_geloof)
table_van_dataframe_geloof
```

Ik kwam er net achter dat ik al de juiste data frame heb.. Dit had zoveel sneller gekund, maar goed. Dit is met de juiste nieuwste data. Ik hoef hier alleen nog een tabel van te maken:

```
(nu_echt_tabel_geloof <- table(geloof_tabel_test$Abortus_keuze, geloof_tabel_test$Gelovig))
```

```
##
##      Niet gelovig Wel gelovig
## Ja           93          17
## Nee            0           4
```

En dan nu de test:

```
chisq.test(nu_echt_tabel_geloof)
```

```
## Warning in chisq.test(nu_echt_tabel_geloof): Chi-squared approximation may be
## incorrect
```

```
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  nu_echt_tabel_geloof
## X-squared = 13.163, df = 1, p-value = 0.0002855
```

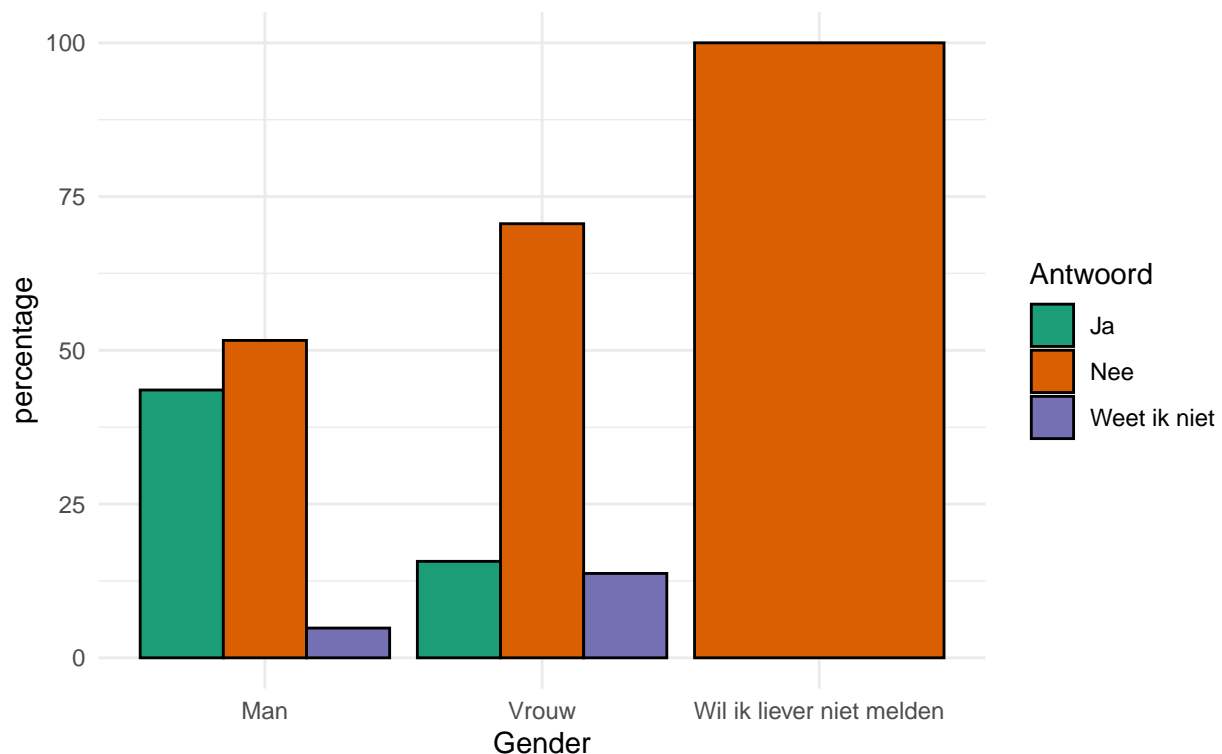
De p-waarde is '0.0002855'. Dit ligt onder de 0.05, dus is er aangetoond dat er een significant verschil zit in mensen hun mening over abortus tussen mensen die wel en niet een religieuze overtuiging hebben.

Zondag 2 juni

Grafieken en omschrijvingen voor de resultaten:

```
library(ggplot2)

ggplot(new_proptabel_dataframe, aes(x = Gender, y = perc, fill = Biologische_vader_zeggenschap)) +
  labs(caption = "Figuur x: Barplot met de verdeling van antwoorden over gender  
op de vraag of de biologische vader zeggenschap moet hebben  
over de beslissing voor een abortus.",
       x = "Gender",
       fill = "Antwoord",
       y = "percentage") +
  geom_bar(stat = "identity", position = "dodge", color = "black") +
  theme_minimal() +
  scale_fill_brewer(palette = "Dark2") +
  # verplaatst de tekst naar links
  theme(plot.caption = element_text(hjust = 0, face = "italic"))
```



*Figuur x: Barplot met de verdeling van antwoorden over gender
op de vraag of de biologische vader zeggenschap moet hebben
over de beslissing voor een abortus.*

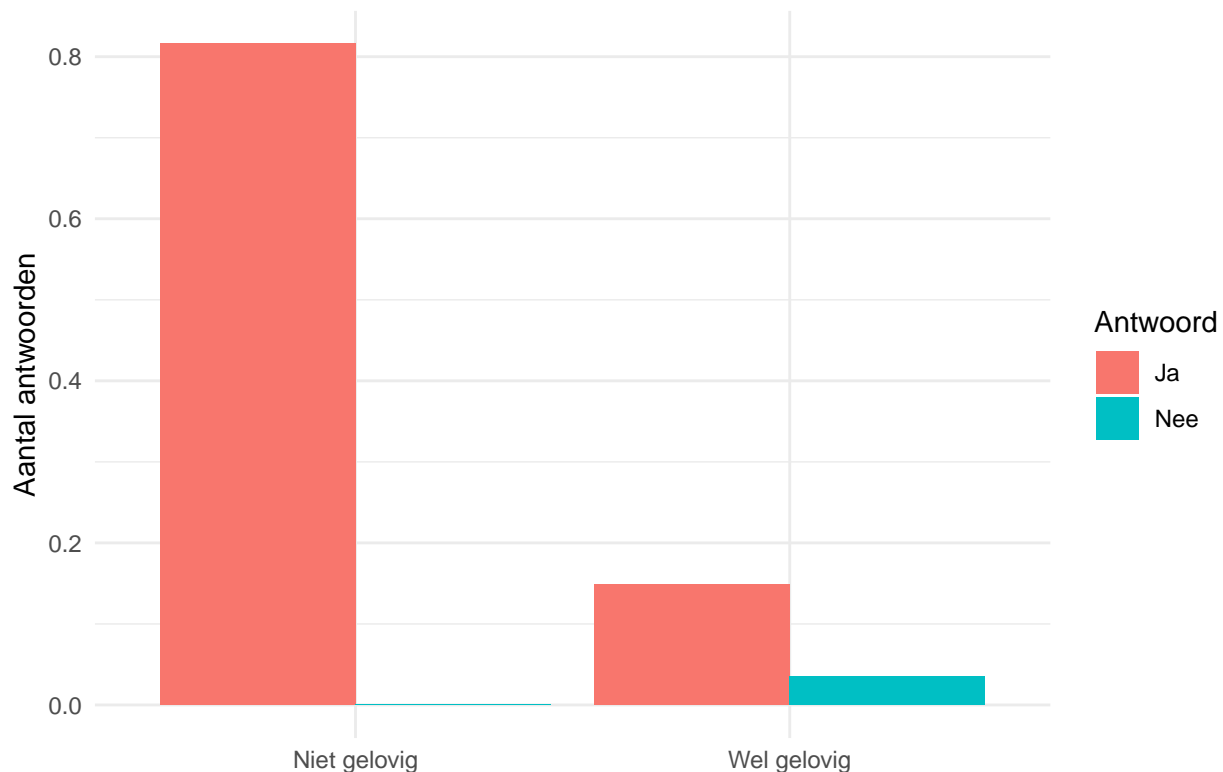
In de bovenstaande plot zie je per gender in procenten weergegeven wat mensen hebben geantwoord op de vraag: 'Moet de biologische vader zeggenschap hebben over de beslissing voor een abortus?' Je ziet hier dat er gemiddeld meer vrouwen zijn die vinden dat mannen hier geen zeggenschap over hebben. Ook zijn er meer mannen die vinden dat de vader hier wel zeggenschap over moet hebben. Uit een Chi-squared

test blijkt dat er een significant verschil zit tussen hoe mannen en vrouwen deze vraag beantwoorden. De p-waarde die hieruit kwam is namelijk '0.01798' en dit ligt onder de '0.05'.

```
tijdelijke_data <- data_renamed %>%
  select(Religie, Abortus_keuze) %>%
  mutate(Gelovig=ifelse(Religie == "Niet gelovig", "Niet gelovig", "Wel gelovig"))
x <- prop.table(table(tijdelijke_data$Gelovig, tijdelijke_data$Abortus_keuze))
x
```

```
##
##              Ja      Nee
## Niet gelovig 0.81578947 0.00000000
## Wel gelovig  0.14912281 0.03508772
```

```
as.data.frame(x)%>%
  ggplot(aes(x = Var1, y = Freq, fill = Var2)) +
  labs(caption = "Figuur x: Corrolatie religie en mening abortus",
       x = "",
       fill = "Antwoord",
       y = "Aantal antwoorden") +
  geom_col(position = "dodge") +
  # verplaatst de tekst naar links
  theme(plot.caption = element_text(hjust = 0, face = "italic")) +
  theme_minimal()
```



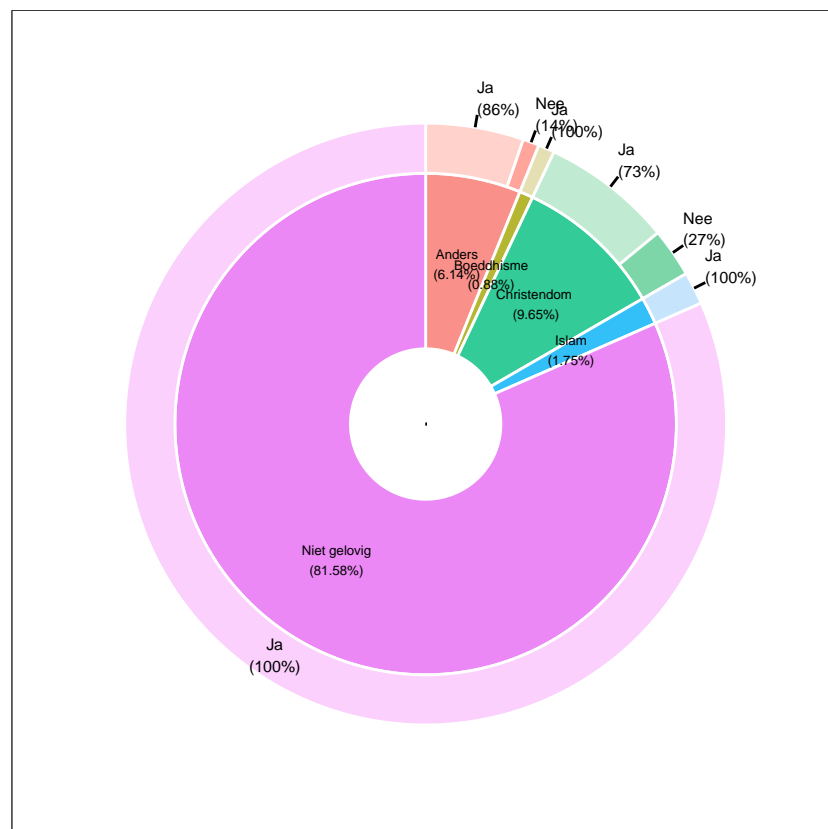
Figuur x: Corrolatie religie en mening abortus

In de bovenstaande plot zie je de verdeling weergegeven van mensen die wel geloven en niet geloven en wat deze hebben geantwoord op de vraag of mensen zelf mogen kiezen of ze een abortus willen.

Je ziet in de plot hierboven dat alle mensen die hebben aangegeven niet een religieuze overtuiging te hebben bij vraag 6 ook hebben aangegeven dat iemand zelf moet kunnen kiezen of die een abortus wil. Bij de mensen die hebben ingevuld wel een religieuze overtuiging te hebben is er een deel die heeft aangegeven dat iemand niet zelf moet kunnen kiezen of die een abortus wil. Uit de Chi-squared test komt dan ook dat er een significant verschil zit tussen mensen die wel aangeven een religieuze overtuiging te hebben en mensen die aangeven dat niet te hebben. De p-waarde die hieruit kwam is '0.0002855' en dit ligt onder de '0.05'.

```
library(ggplot2)
library(webr)
library(dplyr)

PieDonut(data_donut, aes(religieuze_overtuiging, antwoord_vraag, count=Freq),
  # percentage buitenste cirkel (mening abortus)
  showRatioDonut = TRUE,
  # percentage in de binnenste cirkel (geloof)
  showRatioPie = TRUE,
  # haalt de enorme naam in het midden weg, mooier
  showPieName = FALSE,
  # grootte tekst binnenste ring (geloof)
  pieLabelSize = 1.8,
  donutLabelSize = 2,
  # percentages gefikst
  showRatioThreshold = getOption("PieDonut.showRatioThreshold", 0.0)
)
```




```
# + labs(caption = "Figuur x: test101, pls doe het")
```

! zodra ik een caption probeer toe te voegen wordt de binnenste ring niet meer weergegeven, titel op andere manier toevoegen in paper. ! “titel: correlatie geloof en mening abortus”

In de bovenstaande plot wordt er in de binnenste plot weergegeven hoeveel procent van de respondenten welke religieuze overtuigingen hebben. 81.58% van de respondenten geeft aan niet een religieuze overtuiging te hebben. Verder geeft 9.65% aan in het christendom te geloven, 1.75% in de islam, 0.88% in het boeddhisme en 6.14% geeft aan een religieuze overtuiging te hebben die wij niet als optie hebben meegegeven in de enquête. Vervolgens wordt er in de buitenste ring weergegeven of de respondenten voor of tegen de eigen keuze over abortus zijn per geloof. Dus van de 81.58% mensen die niet een religieuze overtuiging hebben is 100% voor de eigen keuze over abortus. Verder is van de Christelijke respondenten 73% voor de eigen keuze over abortus en 27% tegen abortus. Van de islamitische en de Boeddhistische respondenten is 100% voor de eigen mening over abortus. En van de respondenten die aangaven ergens anders in te geloven is er 86% voor de eigen keuze van abortus en 14% tegen.

Maandag 3 juni

```
# alle libraries die gebruikt worden  
(.packages())
```

```
## [1] "webr"      "ggplot2"   "dplyr"     "stats"     "graphics"  "grDevices"  
## [7] "utils"     "datasets"  "methods"   "base"
```

Oude dataset voor nieuwste versie vervangen en hele logboek hierop aangepast. Ook alle code chunks namen gegeven.

Dinsdag 4 juni

Resultaten aangevuld met wat ik zondag geschreven heb. Discussie geschreven. Logboek nagelezen.