Lasma dataset

# Load some packages

pacman::p\_load(tidyverse,   
 sjPlot)

Load the dataset

df <- read\_csv("https://j.mp/2JJYkeN")

Explore the dataset

head(df)

## # A tibble: 6 x 110  
## RespondentID `Power position` Participation Year Gender Language Parties  
## <chr> <chr> <chr> <dbl> <chr> <chr> <chr>   
## 1 6151307314 Oposition yes 1956 Man Russian "Frakc…  
## 2 6150081146 Oposition yes 1980 Man Latvian "Pie f…  
## 3 6144349659 Oposition yes 1982 Man Latvian "Latvi…  
## 4 6139980102 Government Yes 1949 Man Latvian "Frakc…  
## 5 6131571599 Oposition yes 1953 Man Latvian "Latvi…  
## 6 6127580991 Oposition yes 1955 Man Latvian "Frakc…  
## # … with 103 more variables: YearsMP <dbl>, Chairman <chr>,  
## # Specialisation <chr>, `8.LETA/BNS` <chr>, Monitoring <chr>,  
## # Oft\_radio\_TV <chr>, Oft\_newspaper <chr>, Oft\_publication <chr>, FB <chr>,  
## # draugiem.lv <chr>, Instagram <chr>, vkontakte <chr>, Odnoklassniki <chr>,  
## # Twitter <chr>, dont\_use <chr>, Soc.media\_useful <chr>, Lv\_unbiased <chr>,  
## # LV\_truthful <chr>, LV\_complete\_pic <dbl>, LV\_know\_sub <dbl>,  
## # LV\_trustworthy <dbl>, RU\_unbiased <dbl>, RU\_truthful <dbl>,  
## # RU\_comp\_picture <dbl>, RU\_know\_sub <dbl>, RU\_trustworhy <dbl>,  
## # LV\_radio\_pol <dbl>, RU\_radio\_pol <dbl>, LV\_TV\_pol <dbl>, RU\_TV\_pol <dbl>,  
## # LV\_newsp\_pol <dbl>, RU\_newsp\_pol <dbl>, LV\_magaz\_pol <dbl>,  
## # RU\_magaz\_pol <dbl>, LV\_newssite\_pol <dbl>, RU\_newssite\_pol <dbl>,  
## # LV\_radio\_pub <dbl>, RU\_radio\_pub <dbl>, LV\_TV\_pub <dbl>, RU\_TV\_pub <dbl>,  
## # LV\_newsp\_pub <dbl>, RU\_newspap\_pub <dbl>, LV\_magaz\_pub <dbl>,  
## # RU\_magaz\_pub <dbl>, LV\_newssites\_pub <dbl>, RU\_newssites\_pub <dbl>,  
## # Much\_power <dbl>, Media\_owners <dbl>, Motiv\_pol\_power <dbl>,  
## # Journalists\_indep <dbl>, Media\_pub\_agenda <dbl>,  
## # Politicians\_publicity <dbl>, Analitic\_broadcast <dbl>,  
## # Makes\_breakes\_politic <dbl>, Distrust\_politicians <dbl>,  
## # More\_power\_than\_MP <dbl>, More\_power\_elect <dbl>, Leak\_info <dbl>,  
## # Express\_opinion <dbl>, Agenda\_prime <dbl>, Agenda\_ministers <dbl>,  
## # Agenda\_MP <dbl>, Agenda\_parties <dbl>, Agenda\_interest\_groups <dbl>,  
## # Agenda\_TV\_radio <dbl>, Agenda\_written\_press <dbl>, `10y\_parties` <dbl>,  
## # `10y\_MP` <dbl>, `10y\_ministers` <dbl>, `10\_years\_unions` <dbl>,  
## # `10y\_employers` <dbl>, `10y\_municip` <dbl>, `10y\_NGO` <dbl>,  
## # `10y\_media` <dbl>, `10y\_officials` <dbl>, `10y\_EU` <dbl>, `10y\_OECD` <dbl>,  
## # Inf\_Parl\_agenda <chr>, Inf\_Govr\_agenda <chr>, Inf\_disc\_Parliament <chr>,  
## # Inf\_disc\_pub <chr>, Inf\_MP\_success <chr>, Inf\_deb\_Parliament <chr>,  
## # Inf\_Committees <chr>, Inf\_party\_meeting <chr>, Inf\_question\_hour <chr>,  
## # Inf\_coalition\_talks <chr>, Inf\_informal\_gather <chr>,  
## # Act\_initiatives <chr>, Act\_govern <chr>, Act\_question\_govern <chr>,  
## # Act\_plenar\_speach <chr>, Act\_committee\_speach <chr>, Act\_oral\_hour <chr>,  
## # Info\_soc\_prob <chr>, Info\_import\_pub <chr>, Info\_public <chr>,  
## # Inf\_polit\_weapon <chr>, Info\_other\_MP <chr>, Info\_imp\_parties <chr>, …

# Exploratory data analysis

glimpse(df)

## Rows: 100  
## Columns: 110  
## $ RespondentID <chr> "6151307314", "6150081146", "6144349659", "613…  
## $ `Power position` <chr> "Oposition", "Oposition", "Oposition", "Govern…  
## $ Participation <chr> "yes", "yes", "yes", "Yes", "yes", "yes", "Yes…  
## $ Year <dbl> 1956, 1980, 1982, 1949, 1953, 1955, 1951, 1945…  
## $ Gender <chr> "Man", "Man", "Man", "Man", "Man", "Man", "Man…  
## $ Language <chr> "Russian", "Latvian", "Latvian", "Latvian", "L…  
## $ Parties <chr> "Frakcij? \"Saska?a\"", "Pie frakcij?m nepiede…  
## $ YearsMP <dbl> 18, 25, 7, 6, 1, 25, 2, 25, 25, 6, 6, 15, 6, 6…  
## $ Chairman <chr> "No", "No", "No", "Yes", "No", "No", "No", "No…  
## $ Specialisation <chr> "more than two", "more than two", "more than t…  
## $ `8.LETA/BNS` <chr> "No", "No", "No", "yes", "yes", "No", "No", "N…  
## $ Monitoring <chr> "No", "No", "No", "yes", "yes", "yes", "No", "…  
## $ Oft\_radio\_TV <chr> "Few times a year", "Once", "Once", "Once", "n…  
## $ Oft\_newspaper <chr> "Few times a year", "twice a week", "twice a m…  
## $ Oft\_publication <chr> "Few times a year", "Once", "Once", "Once", "n…  
## $ FB <chr> "yes", "twice a week", "twice a month", "twice…  
## $ draugiem.lv <chr> "na", "Many times a week", "Once", "Less or no…  
## $ Instagram <chr> "na", "yes", "twice a month", "yes", "na", "na…  
## $ vkontakte <chr> "na", "na", "yes", "na", "na", "na", "na", "na…  
## $ Odnoklassniki <chr> "na", "yes", "yes", "na", "na", "na", "na", "n…  
## $ Twitter <chr> "na", "na", "yes", "na", "na", "na", "na", "na…  
## $ dont\_use <chr> "na", "na", "na", "na", "na", "yes", "na", "na…  
## $ Soc.media\_useful <chr> "4", "yes", "na", "na", "3", "6", "na", "yes",…  
## $ Lv\_unbiased <chr> "3", "na", "yes", "na", "4", "3", "2", "na", "…  
## $ LV\_truthful <chr> "4", "5", "na", "4", "4", "3", "4", "na", "3",…  
## $ LV\_complete\_pic <dbl> 4, 2, 5, 4, 3, 3, 4, 1, 3, 3, 4, 4, 3, 5, 3, 4…  
## $ LV\_know\_sub <dbl> 3, 2, 3, 3, 3, 3, 4, 3, 2, 3, 3, 3, 3, 2, 3, 4…  
## $ LV\_trustworthy <dbl> 2, 2, 4, 3, 4, 3, 3, 3, 2, 4, 4, 3, 4, 2, 2, 3…  
## $ RU\_unbiased <dbl> 3, 3, 3, 3, 2, 3, 5, 3, 1, 4, 4, 3, 3, 2, 4, 4…  
## $ RU\_truthful <dbl> 4, 2, 3, 3, 2, 3, 6, 3, 5, 2, 3, 3, 2, 2, 2, 4…  
## $ RU\_comp\_picture <dbl> 3, 2, 4, 6, 6, 3, 6, 3, 2, 6, 2, 3, 2, 2, 3, 4…  
## $ RU\_know\_sub <dbl> 2, 2, 2, 6, 2, 3, 6, 3, 2, 6, 2, 3, 3, 2, 3, 3…  
## $ RU\_trustworhy <dbl> 2, 2, 3, 6, 6, 3, 6, 3, 1, 6, 2, 3, 3, 2, 3, 4…  
## $ LV\_radio\_pol <dbl> 8, 3, 2, 6, 5, 7, 6, 3, 2, 6, 2, 3, 2, 2, 4, 4…  
## $ RU\_radio\_pol <dbl> 2, 2, 3, 6, 6, 8, 8, 3, 1, 6, 9, 3, 8, 2, 2, 4…  
## $ LV\_TV\_pol <dbl> 9, 6, 2, 8, 7, 8, 1, 3, 10, 9, 6, 3, 6, 2, 6, …  
## $ RU\_TV\_pol <dbl> 1, 4, 5, NA, 7, 8, 7, 5, 7, NA, 10, 8, 9, 8, 6…  
## $ LV\_newsp\_pol <dbl> 5, 7, 3, 9, 5, 5, 1, 6, 10, 8, 10, 8, 7, 9, 7,…  
## $ RU\_newsp\_pol <dbl> 0, 6, 6, NA, 5, 5, 5, 7, 7, NA, 7, 8, 8, 10, 5…  
## $ LV\_magaz\_pol <dbl> 6, 7, 3, 6, 4, 4, 1, 6, 4, 9, 9, 8, 7, 10, 3, …  
## $ RU\_magaz\_pol <dbl> 0, 5, 2, NA, 3, 4, 5, 6, 6, NA, 8, 8, 8, 10, 3…  
## $ LV\_newssite\_pol <dbl> 7, 6, 2, 4, 8, 8, 1, 5, 5, NA, 6, 8, 7, 10, 4,…  
## $ RU\_newssite\_pol <dbl> 4, 6, 3, NA, 8, 8, 7, 3, 5, NA, 8, 8, 9, 10, 4…  
## $ LV\_radio\_pub <dbl> 8, 8, 2, 7, 5, 7, 1, 2, 10, 10, 8, 8, 8, 10, 8…  
## $ RU\_radio\_pub <dbl> 1, 6, 5, NA, 4, 8, 8, 7, 7, NA, 9, 8, 8, 8, 8,…  
## $ LV\_TV\_pub <dbl> 9, 7, 4, 8, 8, 8, 1, 6, 10, 9, 5, 8, 7, 8, 6, …  
## $ RU\_TV\_pub <dbl> 1, 6, 7, NA, 8, 8, 7, 6, 9, NA, 10, 8, 9, 10, …  
## $ LV\_newsp\_pub <dbl> 6, 8, 6, 9, 3, 5, 1, 6, 10, 8, 9, 8, 8, 10, 7,…  
## $ RU\_newspap\_pub <dbl> 0, 7, 7, NA, 3, 5, 5, 7, 7, NA, 9, 8, 9, 10, 5…  
## $ LV\_magaz\_pub <dbl> 8, 7, 8, 6, 2, 4, 1, 5, 6, 9, 10, 8, 8, 10, 3,…  
## $ RU\_magaz\_pub <dbl> 3, 7, 4, NA, 2, 4, 5, 7, 8, NA, 8, 8, 9, 10, 3…  
## $ LV\_newssites\_pub <dbl> 7, 7, 4, 4, 8, 8, 1, 6, 4, NA, 7, 8, 8, 10, 4,…  
## $ RU\_newssites\_pub <dbl> 5, 7, 5, NA, 8, 8, 7, 3, 6, NA, 9, 8, 9, 10, 4…  
## $ Much\_power <dbl> 3, 9, 5, 4, 3, 3, 1, 2, 10, 10, 9, 8, 8, 10, 8…  
## $ Media\_owners <dbl> 4, 7, 7, NA, 4, 3, 3, 8, 7, NA, 2, 8, 4, 10, 8…  
## $ Motiv\_pol\_power <dbl> 3, 5, 7, 4, 2, 3, 6, 7, 4, 4, 3, 8, 3, 10, 3, …  
## $ Journalists\_indep <dbl> 3, 5, 3, 5, 3, 2, 6, 4, 5, 4, 2, 4, 4, 5, 4, 3…  
## $ Media\_pub\_agenda <dbl> 3, 4, 4, 3, 3, 3, 4, 4, 5, 4, 3, 5, 3, 5, 2, 2…  
## $ Politicians\_publicity <dbl> 5, 1, 6, 3, 0, 2, 3, 4, 2, 5, 4, 4, 3, 4, 3, 4…  
## $ Analitic\_broadcast <dbl> 4, 4, 2, 4, 4, 5, 6, 2, 5, 3, 3, 2, 3, 2, 2, 3…  
## $ Makes\_breakes\_politic <dbl> 5, 5, 4, 4, 3, 4, 3, 4, 4, 3, 3, 4, 4, 5, 3, 4…  
## $ Distrust\_politicians <dbl> 4, 5, 4, 4, 5, 4, 6, 5, 4, 4, 4, 3, 4, 4, 2, 4…  
## $ More\_power\_than\_MP <dbl> 5, 5, 4, 5, 2, 3, 1, 3, 3, 4, 3, 4, 4, 5, 3, 3…  
## $ More\_power\_elect <dbl> 5, 2, 3, 5, 5, 3, 6, 4, 5, 4, 2, 5, 3, 5, 2, 1…  
## $ Leak\_info <dbl> 5, 3, 4, 3, 3, 5, 6, 4, 5, 4, 4, 3, 4, 5, 2, 1…  
## $ Express\_opinion <dbl> 1, 5, 2, 4, 3, 3, 6, 5, 4, 5, 3, 2, 3, 4, 5, 2…  
## $ Agenda\_prime <dbl> 3, 5, 4, 5, 4, 5, 3, 5, 2, 3, 2, 4, 4, 5, 5, 3…  
## $ Agenda\_ministers <dbl> 4, 2, 6, 5, 3, 5, 5, 3, 5, 5, 5, 4, 4, 5, 1, 2…  
## $ Agenda\_MP <dbl> 3, 4, 3, 4, 3, 4, 4, 5, 5, 4, 4, 5, 4, 4, 5, 4…  
## $ Agenda\_parties <dbl> 2, 4, 3, 4, 4, 4, 2, 4, 3, 3, 4, 3, 4, 3, 3, 3…  
## $ Agenda\_interest\_groups <dbl> 4, 4, 3, 3, 3, 4, 2, 4, 3, 3, 4, 3, 3, 3, 2, 4…  
## $ Agenda\_TV\_radio <dbl> 4, 4, 4, 4, 4, 4, 2, 3, 4, 4, 4, 3, 3, 3, 4, 4…  
## $ Agenda\_written\_press <dbl> 3, 3, 3, 3, 4, 4, 5, 3, 4, 4, 5, 4, 4, 3, 3, 4…  
## $ `10y\_parties` <dbl> 4, 4, 3, 4, 6, 6, 3, 3, 5, 4, 4, 3, 4, 4, 3, 4…  
## $ `10y\_MP` <dbl> 3, 2, 3, 3, 5, 5, 6, 4, 3, 4, 7, 2, 4, 2, 2, 3…  
## $ `10y\_ministers` <dbl> 6, 7, 2, 6, 5, 6, 4, 4, 7, 5, 6, 2, 5, 2, 7, 6…  
## $ `10\_years\_unions` <dbl> 2, 7, 5, 5, 3, 3, 6, 6, 5, 4, 6, 6, 5, 6, 6, 5…  
## $ `10y\_employers` <dbl> 5, 7, 4, 5, 3, 6, 2, 2, 3, 4, 4, 5, 3, 4, 6, 6…  
## $ `10y\_municip` <dbl> 5, 4, 4, 3, 3, 6, 4, 6, 5, 5, 5, 6, 3, 4, 2, 4…  
## $ `10y\_NGO` <dbl> 4, 5, 4, 4, NA, 6, 6, 1, 5, 5, 5, 5, 4, 3, 2, …  
## $ `10y\_media` <dbl> 7, 7, 4, 5, 4, 7, 2, 2, 6, 6, 4, 5, 2, 4, 5, 4…  
## $ `10y\_officials` <dbl> 7, 5, 3, 4, 2, 6, 6, 4, 6, 6, 5, 5, 5, 4, 2, 5…  
## $ `10y\_EU` <dbl> 6, 6, 2, 6, 2, 7, 2, 1, 7, 6, 4, 4, 2, 3, 4, 6…  
## $ `10y\_OECD` <dbl> 5, 6, 4, 4, 3, 6, 4, 5, 3, 3, 6, 5, 2, 6, 5, 3…  
## $ Inf\_Parl\_agenda <chr> "to some extent", "6", "5", "4", "to some exte…  
## $ Inf\_Govr\_agenda <chr> "quite a little", "5", "4", "3", "quite a litt…  
## $ Inf\_disc\_Parliament <chr> "quite a little", "to some extent", "3", "to s…  
## $ Inf\_disc\_pub <chr> "very much", "to some extent", "quite a little…  
## $ Inf\_MP\_success <chr> "very much", "to some extent", "to some extent…  
## $ Inf\_deb\_Parliament <chr> "often", "quite a little", "quite a little", "…  
## $ Inf\_Committees <chr> "very often", "very much", "quite a little", "…  
## $ Inf\_party\_meeting <chr> "often", "now and then", "very much", "now and…  
## $ Inf\_question\_hour <chr> "now and then", "often", "often", "often", "no…  
## $ Inf\_coalition\_talks <chr> "na", "now and then", "now and then", "seldom"…  
## $ Inf\_informal\_gather <chr> "very often", "now and then", "very often", "s…  
## $ Act\_initiatives <chr> "Quite small part", "often", "now and then", "…  
## $ Act\_govern <chr> "Quite large part", "often", "na", "now and th…  
## $ Act\_question\_govern <chr> "Quite large part", "Quite large part", "often…  
## $ Act\_plenar\_speach <chr> "Neither big nor small", "Quite large part", "…  
## $ Act\_committee\_speach <chr> "Neither big nor small", "Quite large part", "…  
## $ Act\_oral\_hour <chr> "na", "Very large part", "Quite small part", "…  
## $ Info\_soc\_prob <chr> "Some", "Quite large part", "Very large part",…  
## $ Info\_import\_pub <chr> "Some", "None", "Neither big nor small", "Quit…  
## $ Info\_public <chr> "Quite a lot", "Quite a lot", "Quite large par…  
## $ Inf\_polit\_weapon <chr> "Quite little", "Quite a lot", "Some", "Quite …  
## $ Info\_other\_MP <chr> "Some", "Often", "Quite little", "Quite a lot"…  
## $ Info\_imp\_parties <chr> "Often", "Quite a lot", "Often", "Quite little…  
## $ X108 <chr> NA, "Quite a lot", "Some", "Some", NA, NA, "Qu…  
## $ X109 <chr> NA, "Often", "Some", "Quite a lot", NA, NA, NA…  
## $ X110 <chr> NA, NA, "Quite a lot", NA, NA, NA, NA, "Some",…

Some thoughts: Year should be integer or factor

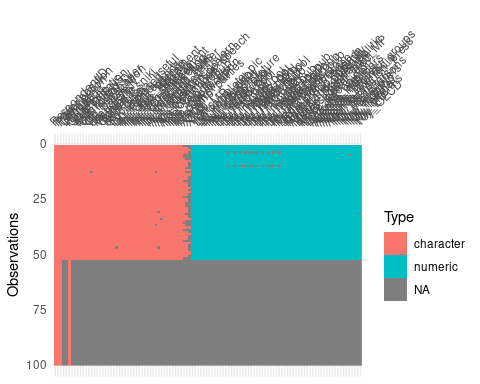
Check the amount of NA data

For this I will use the visdat package

pacman::p\_load(visdat)

and plot

df %>%   
 visdat::vis\_dat()

 Lot of NAs…

Let’s check some variables

table(df$Gender)

##   
## Man Woman   
## 43 9

table(df$`Power position`)

##   
## Government Oposition   
## 61 39

table(df$Gender, df$`Power position`)

##   
## Government Oposition  
## Man 30 13  
## Woman 5 4

Seems to me that the data are divided in

*demographic data*

[1] “RespondentID” “Power position” “Participation” “Year” “Gender”  
[6] “Language” “Parties” “YearsMP” “Chairman” “Specialisation”  
[11] “8.LETA/BNS”

*Some questions about use of social media*

“Monitoring” “Oft\_radio\_TV” “Oft\_newspaper” “Oft\_publication”  
[16] “FB” “draugiem.lv” “Instagram” “vkontakte” “Odnoklassniki”  
[21] “Twitter” “dont\_use”

*some perceptions about their utility*

the rest of the columns

We can use an special package to create summary table ready for publication, it’s called gtsummary

pacman::p\_load(gtsummary)

Learn about the package here: <https://github.com/ddsjoberg/gtsummary>

For example, a table of the demographic data by power position would be:

df %>%   
 select(`Power position`:Specialisation) %>% # here I select the desired columns to summarize  
 gtsummary::tbl\_summary(by = `Power position`)

|  |  |  |
| --- | --- | --- |
| **Characteristic** | **Government**, N = 61 | **Oposition**, N = 39 |
| Participation |  |  |
| No | 26 (43%) | 22 (56%) |
| yes | 0 (0%) | 17 (44%) |
| Yes | 35 (57%) | 0 (0%) |
| Year | 1,963 (1,954, 1,971) | 1,964 (1,956, 1,973) |
| Unknown | 26 | 22 |
| Gender |  |  |
| Man | 30 (86%) | 13 (76%) |
| Woman | 5 (14%) | 4 (24%) |
| Unknown | 26 | 22 |
| Language |  |  |
| Latvian | 32 (91%) | 11 (65%) |
| Russian | 3 (8.6%) | 6 (35%) |
| Unknown | 26 | 22 |
| Parties |  |  |
| Frakcij? “No sirds Latvijai” | 0 (0%) | 7 (18%) |
| Frakcij? “Saska?a” | 0 (0%) | 24 (62%) |
| Frakcij?��Vienot?ba" | 23 (38%) | 0 (0%) |
| Latvijas Re?ionu apvien?bas frakcij? | 0 (0%) | 7 (18%) |
| Nacion?l? apvien?bas �Visu Latvijai!�-�T?vzemei un Br?v?bai/LNNK�frakcij? | 17 (28%) | 0 (0%) |
| Pie frakcij?m nepiedero�s deput?ts | 0 (0%) | 1 (2.6%) |
| Za?o un Zemnieku savien?bas frakcij? | 21 (34%) | 0 (0%) |
| YearsMP | 6 (4, 6) | 7 (3, 18) |
| Unknown | 26 | 22 |
| Chairman | 14 (40%) | 3 (18%) |
| Unknown | 26 | 22 |
| Specialisation |  |  |
| more than two | 25 (71%) | 11 (65%) |
| one or two | 10 (29%) | 6 (35%) |
| Unknown | 26 | 22 |

As you can see, there is *a lot* of data to clean here.

I *strongly suggest* you to learn two of the tidyverse packages:

<https://stringr.tidyverse.org/>

<https://forcats.tidyverse.org/>

# Questions

What are MPs perception on media power compared to other actors (Columns Agenda\_TV\_radio and Agenda\_written\_press compared to Agenda\_prime; Agenda\_ministers; Agenda\_MP; Agenda\_parties; Agenda\_interest\_groups). The survey was in Likert scale 1 - 5);

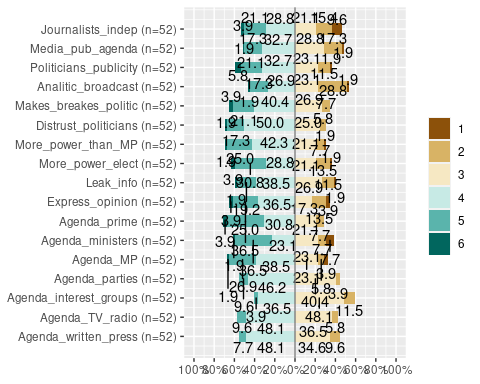
I don’t have much experience with this, so I google: <https://www.google.com/search?q=lickert+data+r&oq=lickert+data+r&aqs=chrome>..69i57.7470j0j1&sourceid=chrome&ie=UTF-8

I will subset the dataset, selecting some columns only

mydf <- df %>%   
 select(Journalists\_indep:Agenda\_written\_press)

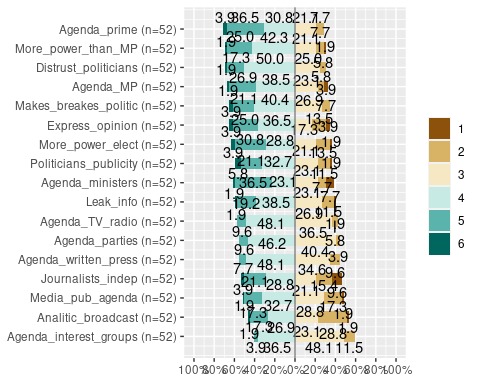
after googling some *hours* found some packages that can help to analyze Likert data

mydf %>%   
 sjPlot::plot\_likert()

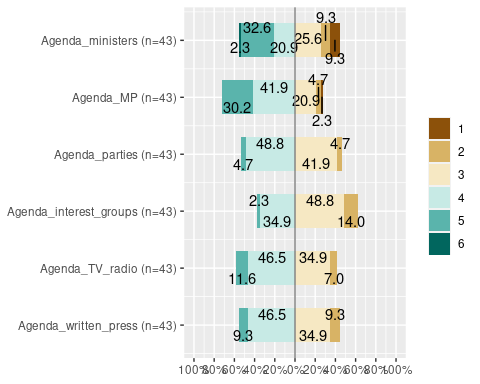
 Please, read the documentation and options of the package here: <https://cran.r-project.org/web/packages/sjPlot/sjPlot.pdf>

For example, if you search the plot\_likert function and read the options, there is one that allows you to sort the answers:

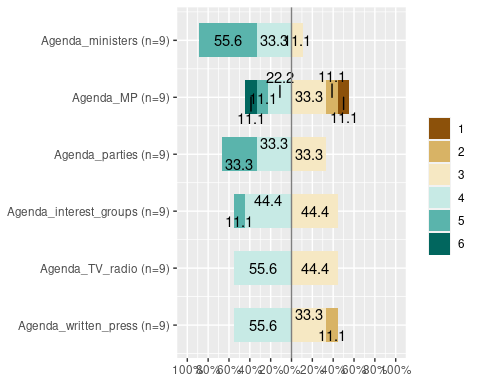
mydf %>%  
 sjPlot::plot\_likert(sort.frq = "pos.asc" ) # see page 25

 and you can create different plot for different groups, for example for men and women

df %>%  
 filter(Gender == "Man") %>% # filter only men  
 select(Agenda\_ministers:Agenda\_written\_press) %>% # and these range of columns  
 sjPlot::plot\_likert() # see page 25



df %>%  
 filter(Gender == "Woman") %>% # filter only men  
 select(Agenda\_ministers:Agenda\_written\_press) %>% # and these range of columns  
 sjPlot::plot\_likert() # see page 25



:)

# Learn more

learn about the sjPlot package here:

<https://strengejacke.github.io/sjPlot/>

# Learn even more!

Check the survey package, specially designed to analyze data from surveys

<http://r-survey.r-forge.r-project.org/survey/>

since is quite an old package, there is a tool that allows you to work with dplyr from tidyverse package:

<https://github.com/gergness/srvyr>