01 Data Wrangling

Thomas J. Brailey

29/09/2019

Contents

Load data

```
# Set working directory
wd <- paste0(here::here(), '/data/')

# Install datasets
# Power-sharing-specific datasets
idc <- rio::import(paste0(wd, 'IDC_country-year_v1_0.RData'))
dtd <- rio::import(paste0(wd,'Democracy Timeseries Data January 2009 Excel2007.csv'))
vdem <- rio::import(paste0(wd,"V-Dem-CY+Others-v8.csv"))
dpi <- rio::import(paste0(wd,'DPI2012.xls'))
qog_ts <- rio::import(paste0(wd,'qog_std_ts_jan19.csv'))
ucdp <- rio::import(paste0(wd, "ucdp-prio-acd-191.xlsx"))
intra <- rio::import(paste0(wd, 'annual_onset.dta'))
p4 <- rio::import(paste0(here::here(), "/data/p4v2017.xls"))
```

Tidy datasets

Set baseline data

```
# Use Strom et al.'s (2017) data as baseline
tb <- idc

# Rename unmatched countries
tb$country[tb$country == "Cent. Af. Rep."] <- "Central African Republic"
tb$country[tb$country == "Dom. Rep."] <- "Dominican Republic"
tb$country[tb$country == "GDR"] <- "German Democratic Republic"
tb$country[tb$country == "PRC"] <- "China"
tb$country[tb$country == "ROK"] <- "South Korea"
tb$country[tb$country == "S. Africa"] <- "South Africa"
tb$country[tb$country == "Serbia and Montenegro"] <- "Montenegro"

tb <- tb[!is.na(tb$country) & !is.na(tb$ifs),]

tb$cowc <- countrycode::countrycode(tb$country, 'country.name', 'cowc')
```

```
## Warning in countrycode::countrycode(tb$country, "country.name", "cowc"): Some values were not matche
```

```
tb$cown <- countrycode::countrycode(tb$country, 'country.name', 'cown')
```

```
## Warning in countrycode::countrycode(tb$country, "country.name", "cown"): Some values were not matche
```

```
# Some values were not matched unambiguously: Serbia

# Select key variables
tb <- tb %>%
  dplyr::select(country,
                cowc,
                cown,
                year,
                ifs,
                mveto,
                gcman,
                gcimp,
                auton,
                jrevman,
                relconstd,
                relconstp,
                milleg,
                partynoethnic,
                jtenure,
                jconst,
                gcseats1,
                gcseats2,
                gcseats3,
                unity,
                resman,
                resseats,
                resseats2,
                resseatsimp,
                miman,
```

```
                subtax,
                subed,
                subpolice,
                fedunits,
                state,
                muni,
                -ifs
                )
```

Clean VDEM

```
# Select key variables
vdem_psp_sub <- vdem %>%
  dplyr::select(country_name,
                year,
                e_miinterc,
                e_Civil_War,
                v2elfrfair) %>%
  dplyr::rename(country = country_name)

vdem_psp_sub$country <- countrycode::countrycode(
  vdem_psp_sub$country, 'country.name', 'country.name'
  )
```

```
## Warning in countrycode::countrycode(vdem_psp_sub$country, "country.name", : Some values were not mat
ncipe, WÃ¼rtemberg
```

```
vdem_psp_sub$cown <- countrycode::countrycode(
  vdem_psp_sub$country, 'country.name', 'cown'
  )
```

```
## Warning in countrycode::countrycode(vdem_psp_sub$country, "country.name", : Some values were not mat
```

```
vdem_psp_sub$cowc <- countrycode::countrycode(
  vdem_psp_sub$country, 'country.name', 'cowc'
  )
```

```
## Warning in countrycode::countrycode(vdem_psp_sub$country, "country.name", : Some values were not mat
```

```
vdem_psp_sub <- vdem_psp_sub %>%
  dplyr::select(-country)
```

Clean QoG

```
# Select key variables and clean.
qog_ts_psp_sub <- qog_ts %>%
  dplyr::select(cname,
                year,
                fe_etfra,
                iaep_ebbp,
                gle_gdp,
                bti_ci,
                cspf_sfi,
                gtm_unit,
                ccp_hr,
```

```
                ffp_hr,
                iiag_phr,
                dpi_housesys,
                jw_bicameral,
                bti_ig,
                vdem_partipdem,
                iaep_nr,
                bti_sop,
                bti_ffe,
                gol_est,
                gol_mt,
                iaep_es,
                no_ef,
                no_ce,
                iaep_eccdt,
                iaep_ecdl,
                iaep_eml,
                iaep_epmf,
                iaep_evp,
                iaep_lcre,
                iaep_lego,
                iaep_lrit,
                wbgi_pve,
                hum_satdem,
                hum_supdem,
                hum_trust,
                wdi_gini,
                gle_pop,
                al_ethnic,
                dpi_auton,
                pt_federal
  ) %>%
  dplyr::rename(country = cname)

qog_ts_psp_sub$country[
  qog_ts_psp_sub$country == "Micronesia"
  ] <- "Federated States of Micronesia"
qog_ts_psp_sub$country[
  qog_ts_psp_sub$country == "Serbia and Montenegro"
  ] <- "Montenegro"
qog_ts_psp_sub$country <- countrycode::countrycode(
  qog_ts_psp_sub$country, 'country.name', 'country.name'
  )
```

```
## Warning in countrycode::countrycode(qog_ts_psp_sub$country, "country.name", : Some values were not ma
```

```
qog_ts_psp_sub$cown <- countrycode::countrycode(
  qog_ts_psp_sub$country, 'country.name', 'cown'
  )
```

```
## Warning in countrycode::countrycode(qog_ts_psp_sub$country, "country.name", : Some values were not ma
```

```
qog_ts_psp_sub$cowc <- countrycode::countrycode(
  qog_ts_psp_sub$country, 'country.name', 'cowc'
  )
```

```
## Warning in countrycode::countrycode(qog_ts_psp_sub$country, "country.name", : Some values were not ma
```

```
qog_ts_psp_sub <- qog_ts_psp_sub %>%
  dplyr::select(-country)
```

Clean DPI

```
# Select key variables
dpi_psp_sub <- dpi %>%
  dplyr::select(countryname,
                year,
                system,
                author,
                pr,
                sensys,
                eiec
                ) %>%
  dplyr::rename(country = countryname) %>%
  dplyr::mutate(year = as.numeric(year))

dpi_psp_sub$country[dpi_psp_sub$country == "Cent. Af. Rep."] <- "Central African Republic"
dpi_psp_sub$country[dpi_psp_sub$country == "Dom. Rep."] <- "Dominican Republic"
dpi_psp_sub$country[dpi_psp_sub$country == "GDR"] <- "German Democratic Republic"
dpi_psp_sub$country[dpi_psp_sub$country == "PRC"] <- "China"
dpi_psp_sub$country[dpi_psp_sub$country == "PRK"] <- "North Korea"
dpi_psp_sub$country[dpi_psp_sub$country == "ROK"] <- "South Korea"
dpi_psp_sub$country[dpi_psp_sub$country == "S. Africa"] <- "South Africa"

dpi_psp_sub$country <- countrycode::countrycode(
  dpi_psp_sub$country, 'country.name', 'country.name'
  )
dpi_psp_sub$cowc <- countrycode::countrycode(
  dpi_psp_sub$country, 'country.name', 'cowc'
  )
dpi_psp_sub$cown <- countrycode::countrycode(
  dpi_psp_sub$country, 'country.name', 'cown'
  )

dpi_psp_sub <- dpi_psp_sub %>%
  dplyr::select(-country)
```

Clean UCDP

```
ucdp_psp_sub <- ucdp %>%
  dplyr::select(location, year,
                side_a, side_b,
                territory_name,
                intensity_level,
                cumulative_intensity,
                type_of_conflict) %>%
  dplyr::filter(stringr::str_detect(location, ",", negate = TRUE)) %>%
  dplyr::rename(country = location)
```

```
ucdp_psp_sub$cowc <- countrycode::countrycode(
  ucdp_psp_sub$country, "country.name", "cowc"
  )
```

```
## Warning in countrycode::countrycode(ucdp_psp_sub$country, "country.name", : Some values were not mat
```

```
ucdp_psp_sub$cown <- countrycode::countrycode(
  ucdp_psp_sub$country, "country.name", "cown"
  )
```

```
## Warning in countrycode::countrycode(ucdp_psp_sub$country, "country.name", : Some values were not mat
```

```
ucdp_psp_sub <- dplyr::select(ucdp_psp_sub, -country)
```

Clean PRIO

```
intra_psp_sub <- intra %>%
  dplyr::select(gwno, year, onset2)

intra_psp_sub$cowc <- countrycode::countrycode(
  intra_psp_sub$gwno, 'gwn', 'cowc'
  )
```

```
## Warning in countrycode::countrycode(intra_psp_sub$gwno, "gwn", "cowc"): Some values were not matched
```

```
intra_psp_sub$cown <- countrycode::countrycode(
  intra_psp_sub$gwno, 'gwn', 'cown'
  )
```

```
## Warning in countrycode::countrycode(intra_psp_sub$gwno, "gwn", "cown"): Some values were not matched
```

```
intra_psp_sub <- intra_psp_sub %>% dplyr::select(-gwno)
```

Clean PolityIV

```
p4_psp_sub <- p4 %>%
  dplyr::select(country, year, polity, fragment)

p4_psp_sub$cowc <- countrycode::countrycode(
  p4_psp_sub$country, 'country.name', 'cowc'
  )
```

```
## Warning in countrycode::countrycode(p4_psp_sub$country, "country.name", : Some values were not matche
```

```
p4_psp_sub$cown <- countrycode::countrycode(
  p4_psp_sub$country, 'country.name', 'cown'
  )
```

```
## Warning in countrycode::countrycode(p4_psp_sub$country, "country.name", : Some values were not matche
```

```
p4_psp_sub <- p4_psp_sub %>%
  dplyr::filter(!is.na(cowc)) %>%
  dplyr::select(-country)
```

Clean RAI

```
rai_psp_sub <- rio::import(paste0(here::here(), "/data/RAI_country_scores_2015.xlsx")) %>%
  dplyr::select(country_name, year, n_RAI) %>%
  dplyr::mutate(cown = countrycode::countrycode(country_name, "country.name", "cown"),
                cowc = countrycode::countrycode(cown, "cown", "cowc"),
                year = as.numeric(year)) %>%
  dplyr::select(cown, cowc, year, n_RAI)
```

```
## New names:
## * n_rep -> n_rep...11
## * n_lawmaking -> n_lawmaking...12
## * n_rep -> n_rep...21
## * n_lawmaking -> n_lawmaking...24
```

```
## Warning in countrycode::countrycode(country_name, "country.name", "cown"): Some values were not matc
```

Clean EPR

```
epr <- rio::import(paste0(here::here(), "/data/EPR-2018.1.1.csv")) %>%
  dplyr::mutate(cown = countrycode::countrycode(gwid, "gwn", "cown")) %>%
  dplyr::select(cown, from, group, reg_aut) %>%
  dplyr::rename(year = from) %>%
  dplyr::group_by(cown, group) %>%
  tidyr::complete(cown, group,
                  year = 1946:2017,
                  fill = list(incidents = 0))
```

```
## Warning in countrycode::countrycode(gwid, "gwn", "cown"): Some values were not matched unambiguously
```

```
epr_wide <- epr %>%
  tidyr::pivot_wider(names_from = group,
                     values_from = reg_aut) %>%
  dplyr::group_by(cown)

epr_wide <- epr_wide %>%
  tidyr::fill_(names(epr_wide[,2:642])) %>%
  dplyr::ungroup()

epr_wide$reg_aut_cont <- rowSums(epr_wide[,3:642] == TRUE, na.rm = TRUE)

epr_psp_sub <- epr_wide %>%
  dplyr::mutate(reg_aut_dum = ifelse(reg_aut_cont >= 1, 1, 0),
                cowc = countrycode::countrycode(cown, "cown", "cowc")) %>%
  dplyr::select(cown, cowc, year, reg_aut_dum, reg_aut_cont)
```

Join data

```
# Join data one-by-one. Check for discrepancies
tb_2 <- dplyr::left_join(tb, qog_ts_psp_sub, by = c("cown", "cowc", "year"))
tb_3 <- dplyr::left_join(tb_2, dpi_psp_sub, by = c("cown", "cowc", "year"))
tb_4 <- dplyr::left_join(tb_3, vdem_psp_sub, by = c("cown", "cowc", "year"))
tb_5 <- dplyr::left_join(tb_4, ucdp_psp_sub, by = c("cown", "cowc", "year"))
```

```
tb_6 <- dplyr::left_join(tb_5, epr_psp_sub, by = c("cown", "cowc", "year"))
tb_7 <- dplyr::left_join(tb_6, intra_psp_sub, by = c("cown", "cowc", "year"))
```

```
## Warning: Column `year` has different attributes on LHS and RHS of join
```

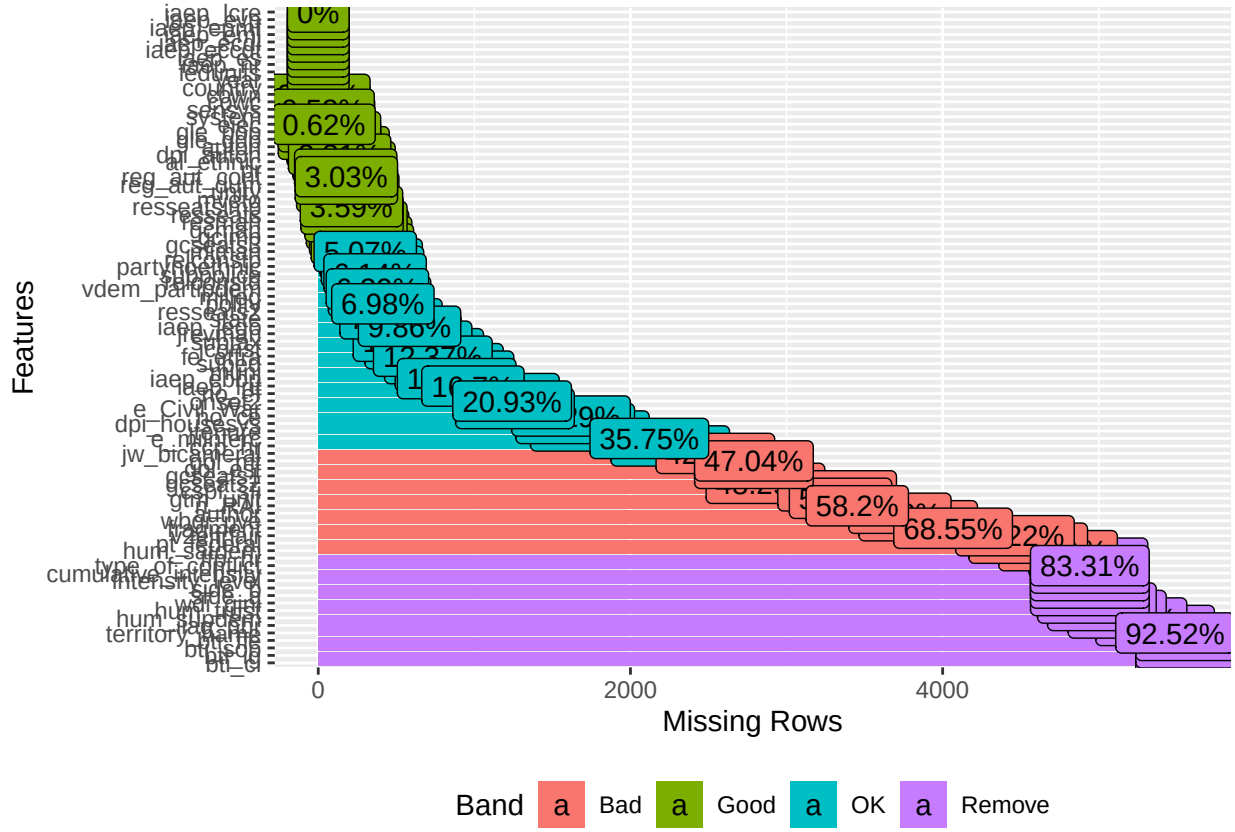
```
tb_8 <- dplyr::left_join(tb_7, p4_psp_sub, by = c("cown", "cowc", "year"))
tb_9 <- dplyr::left_join(tb_8, rai_psp_sub, by = c("cown", "cowc", "year"))

# Collapse duplicate country/years whie retaining values
tb_10 <- tb_9 %>%
  dplyr::group_by(country, year) %>%
  dplyr::summarise_all(dplyr::funs(dplyr::first(na.omit(.))))
```

```
## Warning: funs() is soft deprecated as of dplyr 0.8.0
## Please use a list of either functions or lambdas:
##
##   # Simple named list:
##   list(mean = mean, median = median)
##
##   # Auto named with `tibble::lst()`:
##   tibble::lst(mean, median)
##
##   # Using lambdas
##   list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))
## This warning is displayed once per session.
```

```
# Save data as a .rds file
saveRDS(tb_10, paste0(here::here(), '/data/tjbrailey_psp_not_cleaned.rds'))

# Remove intermediate join files
#rm(tb, tb_1, tb_2, tb_3, tb_4, tb_5, tb_6)
```

Visualize data before recoding

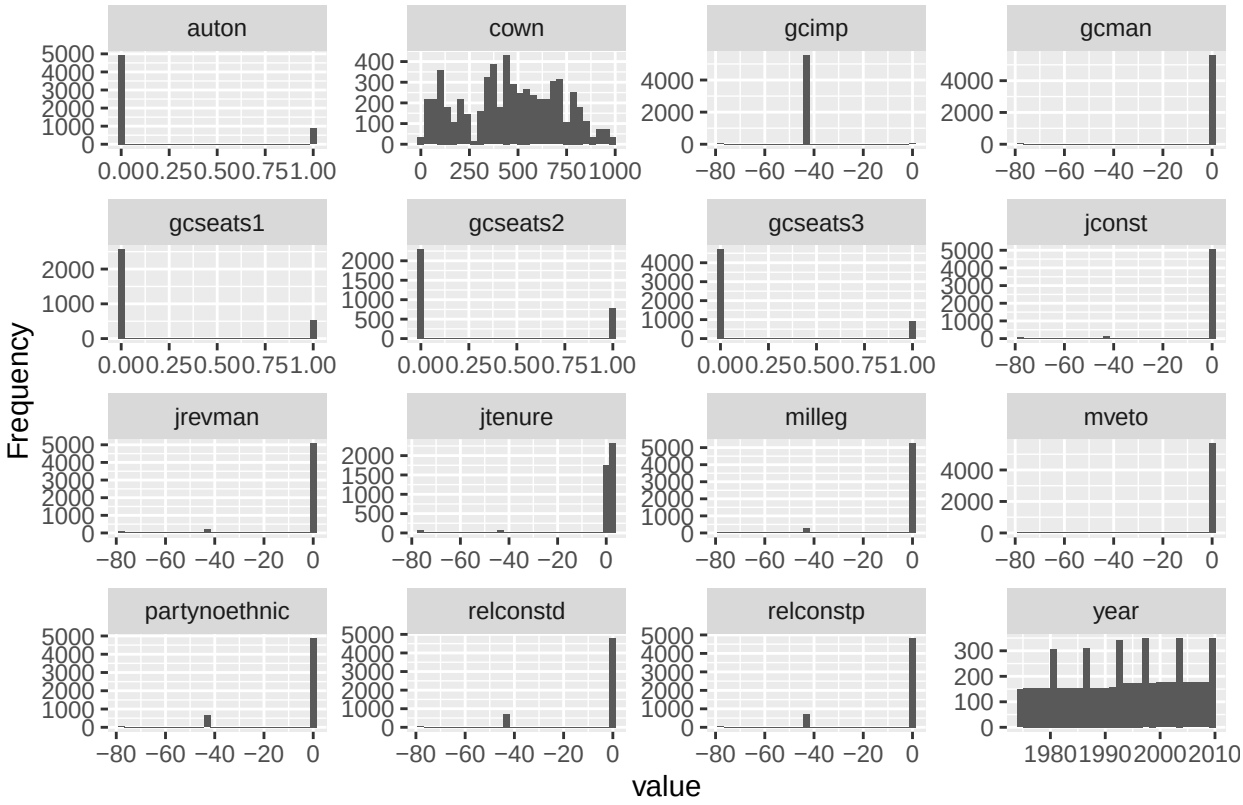
```
# DataExplorer missingness
DataExplorer::plot_missing(tb_10)
```



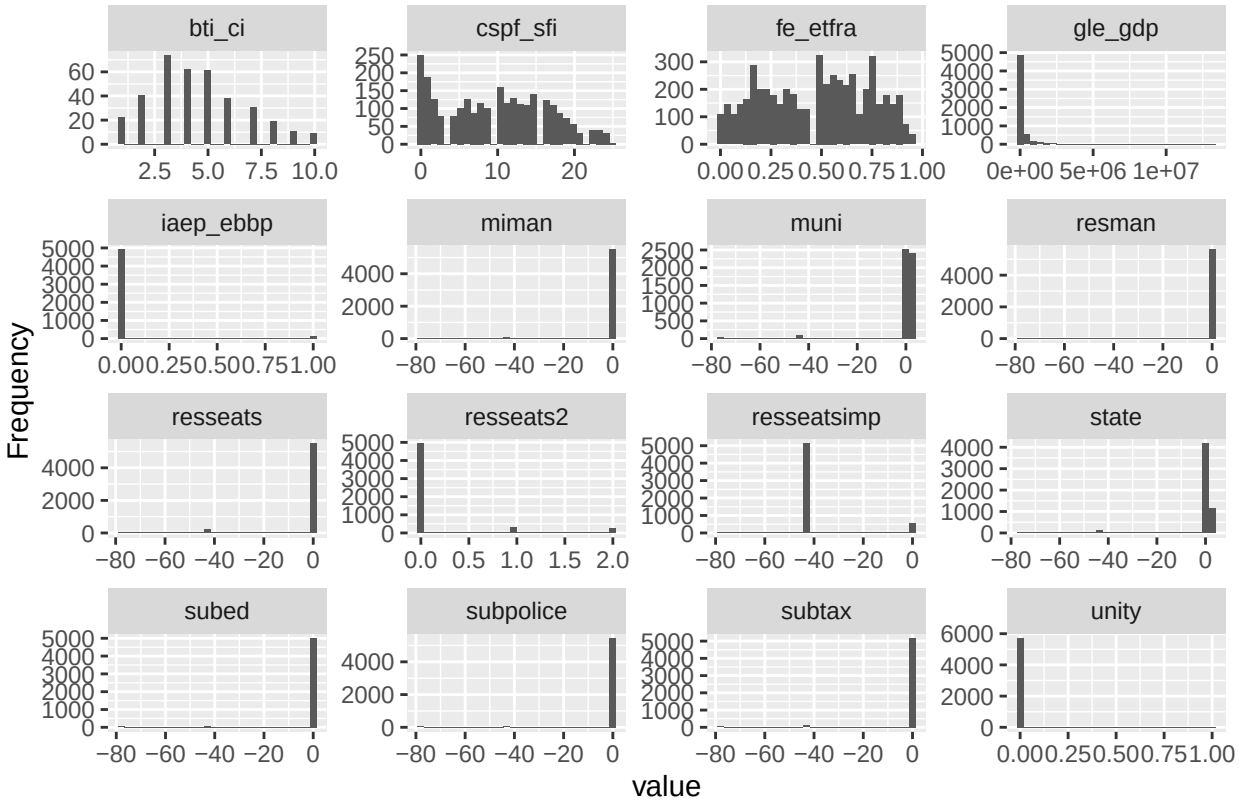

```
# Save image
png(file = paste0(here::here(), '/vis/tjbrailey_psp_datexp.png'),
    width = 2000,
    height = 1000)
DataExplorer::plot_missing(tb_10)
dev.off()
```

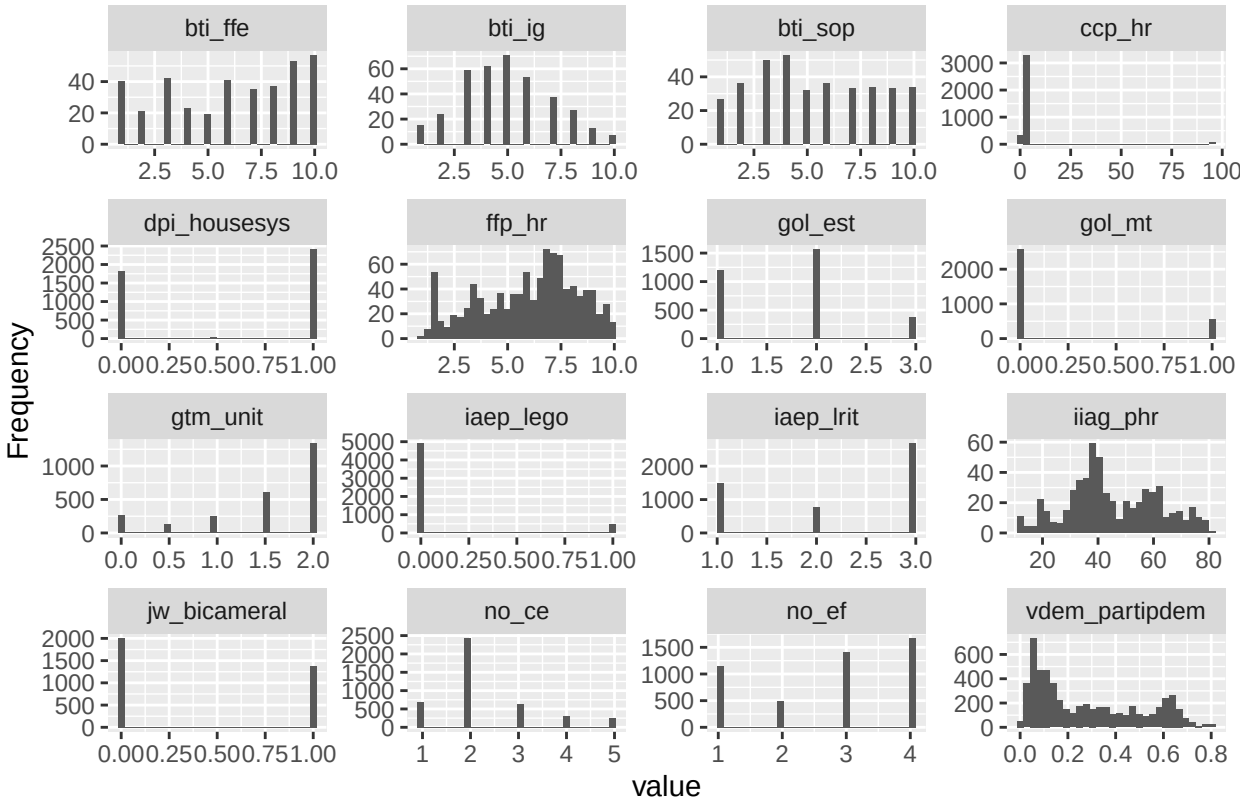
```
## pdf
##   2
```

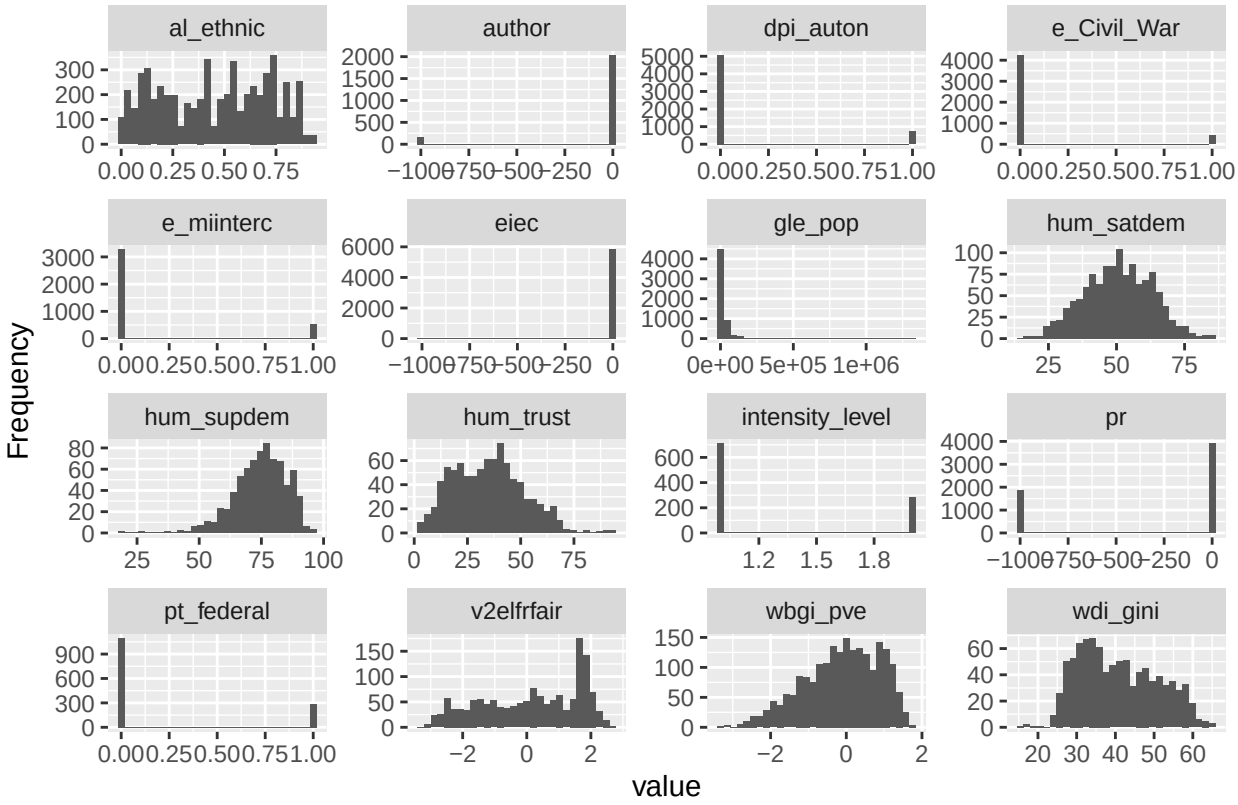
```
# DataExplorer variable histograms
DataExplorer::plot_histogram(tb_10)
```

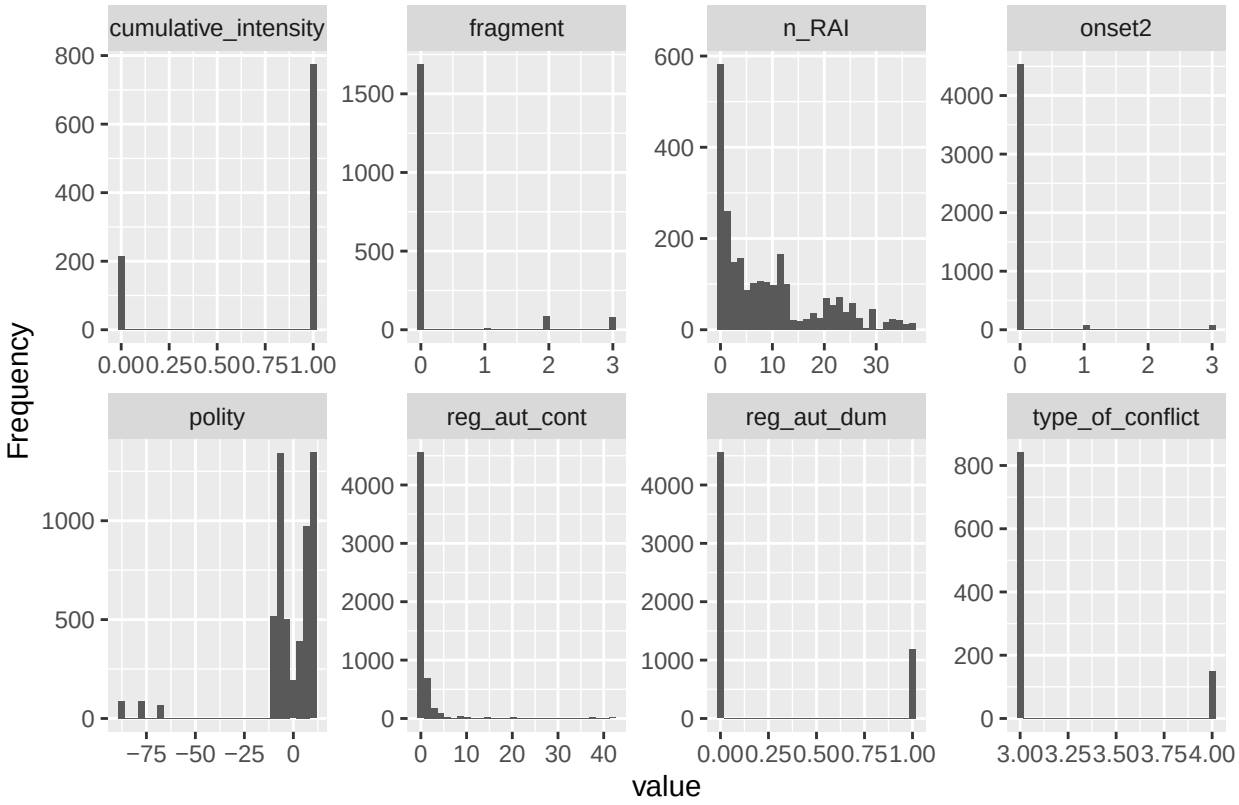

auton
cown
gcimp
gcman
gcseats1
gcseats2
gcseats3
jconst
jrevman
jtenure
milleg
mveto
partynoethnic
relconstd
relconstp
year
Frequency
value

Page 1


bti_ci
cspf_sfi
fe_etfra
gle_gdp
iaep_ebbp
miman
muni
resman
resseats
resseats2
resseatsimp
state
subed
subpolice
subtax
unity
Frequency
value


bti_ffe
bti_ig
bti_sop
ccp_hr
dpi_housesys
ffp_hr
gol_est
gol_mt
gtm_unit
iaep_lego
iaep_lrit
iiag_phr
jw_bicameral
no_ce
no_ef
vdem_partipdem
Frequency
value


al_ethnic
author
dpi_auton
e_Civil_War
e_miinterc
eiec
gle_pop
hum_satdem
hum_supdem
hum_trust
intensity_level
pr
pt_federal
v2elfrfair
wbgi_pve
wdi_gini
Frequency
value




```
# Save images
png(file = paste0(here::here(), '/vis/tjbrailey_psp_datexp.png'),
    width = 2000,
    height = 1000)
DataExplorer::plot_histogram(tb_10)
dev.off()
```

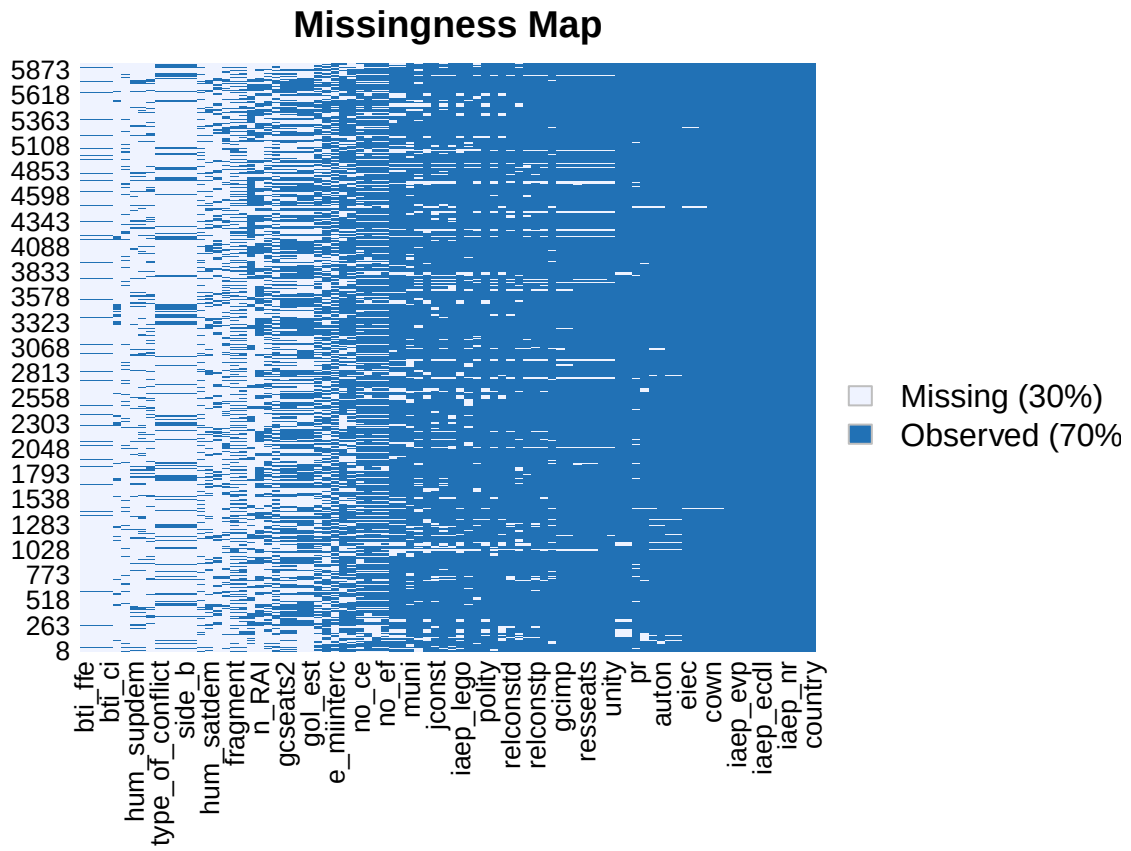
```
## pdf
##   2
```

```
# Amelia missingness
Amelia::missmap(tb_10)
```

```
## Warning: Unknown or uninitialised column: 'arguments'.
```

```
## Warning: Unknown or uninitialised column: 'arguments'.
```

```
## Warning: Unknown or uninitialised column: 'imputations'.
```




```
# Save image
png(file = paste0(here::here(), "/vis/tjbrailey_psp_missingness.png"),
    width = 2000,
    height = 1000)
Amelia::missmap(tb_10)
```

```
## Warning: Unknown or uninitialised column: 'arguments'.
```

```
## Warning: Unknown or uninitialised column: 'arguments'.
```

```
## Warning: Unknown or uninitialised column: 'imputations'.
```

```
dev.off()
```

```
## pdf
##   2
```

Recode and Rename Variables

```
# Install data
psp <- rio::import(paste0(here::here(), "/data/tjbrailey_psp_not_cleaned.rds"))

# Remove codings for NA values
psp_na_recode <- psp

psp_na_recode[psp_na_recode == "-999"] <- NA
```

```r
psp_na_recode[psp_na_recode == "-44"] <- NA

psp_na_recode[psp_na_recode == "-66"] <- NA
psp_na_recode[psp_na_recode == "-77"] <- NA
psp_na_recode[psp_na_recode == "-88"] <- NA
psp_na_recode$ccp_hr[psp_na_recode$ccp_hr == "96"] <- 3

psp_na_recode[psp_na_recode == ".a"] <- NA
psp_na_recode[psp_na_recode == ".b"] <- NA
psp_na_recode[psp_na_recode == ".e"] <- NA

# Rename variables
psp_rename <- psp_na_recode

psp_rename <- psp_rename %>%
  dplyr::rename(idc_mveto = mveto,
                idc_gcman = gcman,
                idc_gcimp = gcimp,
                idc_auton = auton,
                idc_jrevman = jrevman,
                idc_relconstd = relconstd,
                idc_relconstp = relconstp,
                idc_milleg = milleg,
                idc_partynoethnic = partynoethnic,
                idc_jtenure = jtenure,
                idc_jconst = jconst,
                idc_gcseats1 = gcseats1,
                idc_gcseats2 = gcseats2,
                idc_gcseats3 = gcseats3,
                idc_unity = unity,
                idc_resman = resman,
                idc_resseats = resseats,
                idc_resseats2 = resseats2,
                idc_resseatsimp = resseatsimp,
                idc_miman = miman,
                idc_subtax = subtax,
                idc_subed = subed,
                idc_subpolice = subpolice,
                idc_state = state,
                idc_muni = muni,
                idc_fedunits = fedunits,

                qog_fe_etfra = fe_etfra,
                qog_iaep_ebbp = iaep_ebbp,
                qog_gle_gdp = gle_gdp,
                qog_bti_ci = bti_ci,
                qog_cspf_sfi = cspf_sfi,
                qog_gtm_unit = gtm_unit,
                qog_ccp_hr = ccp_hr,
                qog_ffp_hr = ffp_hr,
                qog_iiag_phr = iiag_phr,
                qog_dpi_housesys = dpi_housesys,
                qog_jw_bicameral = jw_bicameral,
```

```
            qog_bti_ig = bti_ig,
            qog_vdem_partipdem = vdem_partipdem,
            qog_iaep_nr = iaep_nr,
            qog_bti_sop = bti_sop,
            qog_gol_est = gol_est,
            qog_gol_mt = gol_mt,
            qog_iaep_es = iaep_es,
            qog_no_ef = no_ef,
            qog_no_ce = no_ce,
            qog_iaep_eccdt = iaep_eccdt,
            qog_iaep_ecdl = iaep_ecdl,
            qog_iaep_eml = iaep_eml,
            qog_iaep_epmf = iaep_epmf,
            qog_iaep_evp = iaep_evp,
            qog_iaep_lcre = iaep_lcre,
            qog_iaep_lego = iaep_lego,
            qog_iaep_lrit = iaep_lrit,
            qog_wbgi_pve = wbgi_pve,
            qog_hum_satdem = hum_satdem,
            qog_hum_supdem = hum_supdem,
            qog_hum_trust = hum_trust,
            qog_wdi_gini = wdi_gini,
            qog_gle_pop = gle_pop,
            qog_al_ethnic = al_ethnic,
            qog_pt_federal = pt_federal,

            dpi_system = system,
            dpi_author = author,
            dpi_pr = pr,
            dpi_sensys = sensys,
            dpi_eiec = eiec,

            vdem_e_miinterc = e_miinterc,
            vdem_e_civil_war = e_Civil_War,

            ucdp_side_a = side_a,
            ucdp_side_b = side_b,
            ucdp_territory_name = territory_name,
            ucdp_intensity_level = intensity_level,
            ucdp_type_of_conflict = type_of_conflict,
            ucdp_cumulative_intensity = cumulative_intensity,

            prio_onset = onset2,

            polity4_polity_score = polity,
            polity4_fragment = fragment,

            epr_reg_aut_dum = reg_aut_dum,
            epr_reg_aut_cont = reg_aut_cont,

            rai_n_RAI = n_RAI
            ) %>%
```

```r
  # Fill out remaining variables
  dplyr::group_by(country) %>%
  tidyr::fill(qog_wdi_gini,
              qog_wbgi_pve,
              qog_hum_satdem,
              qog_hum_supdem,
              qog_hum_trust,
              bti_ffe,
              v2elfrfair)

variable.names(psp_rename)
```

```
##  [1] "country"                    "year"
##  [3] "cowc"                       "cown"
##  [5] "idc_mveto"                  "idc_gcman"
##  [7] "idc_gcimp"                  "idc_auton"
##  [9] "idc_jrevman"                "idc_relconstd"
## [11] "idc_relconstp"              "idc_milleg"
## [13] "idc_partynoethnic"          "idc_jtenure"
## [15] "idc_jconst"                 "idc_gcseats1"
## [17] "idc_gcseats2"               "idc_gcseats3"
## [19] "idc_unity"                  "idc_resman"
## [21] "idc_resseats"               "idc_resseats2"
## [23] "idc_resseatsimp"            "idc_miman"
## [25] "idc_subtax"                 "idc_subed"
## [27] "idc_subpolice"              "idc_fedunits"
## [29] "idc_state"                  "idc_muni"
## [31] "qog_fe_etfra"               "qog_iaep_ebbp"
## [33] "qog_gle_gdp"                "qog_bti_ci"
## [35] "qog_cspf_sfi"               "qog_gtm_unit"
## [37] "qog_ccp_hr"                 "qog_ffp_hr"
## [39] "qog_iiag_phr"               "qog_dpi_housesys"
## [41] "qog_jw_bicameral"           "qog_bti_ig"
## [43] "qog_vdem_partipdem"         "qog_iaep_nr"
## [45] "qog_bti_sop"                "bti_ffe"
## [47] "qog_gol_est"                "qog_gol_mt"
## [49] "qog_iaep_es"                "qog_no_ef"
## [51] "qog_no_ce"                  "qog_iaep_eccdt"
## [53] "qog_iaep_ecdl"              "qog_iaep_eml"
## [55] "qog_iaep_epmf"              "qog_iaep_evp"
## [57] "qog_iaep_lcre"              "qog_iaep_lego"
## [59] "qog_iaep_lrit"              "qog_wbgi_pve"
## [61] "qog_hum_satdem"             "qog_hum_supdem"
## [63] "qog_hum_trust"              "qog_wdi_gini"
## [65] "qog_gle_pop"                "qog_al_ethnic"
## [67] "dpi_auton"                  "qog_pt_federal"
## [69] "dpi_system"                 "dpi_author"
## [71] "dpi_pr"                     "dpi_sensys"
## [73] "dpi_eiec"                   "vdem_e_miinterc"
## [75] "vdem_e_civil_war"           "v2elfrfair"
## [77] "ucdp_side_a"                "ucdp_side_b"
## [79] "ucdp_territory_name"        "ucdp_intensity_level"
## [81] "ucdp_cumulative_intensity" "ucdp_type_of_conflict"
## [83] "epr_reg_aut_dum"            "epr_reg_aut_cont"
```

```
## [85] "prio_onset"                "polity4_polity_score"
## [87] "polity4_fragment"          "rai_n_RAI"
```

Additional variables

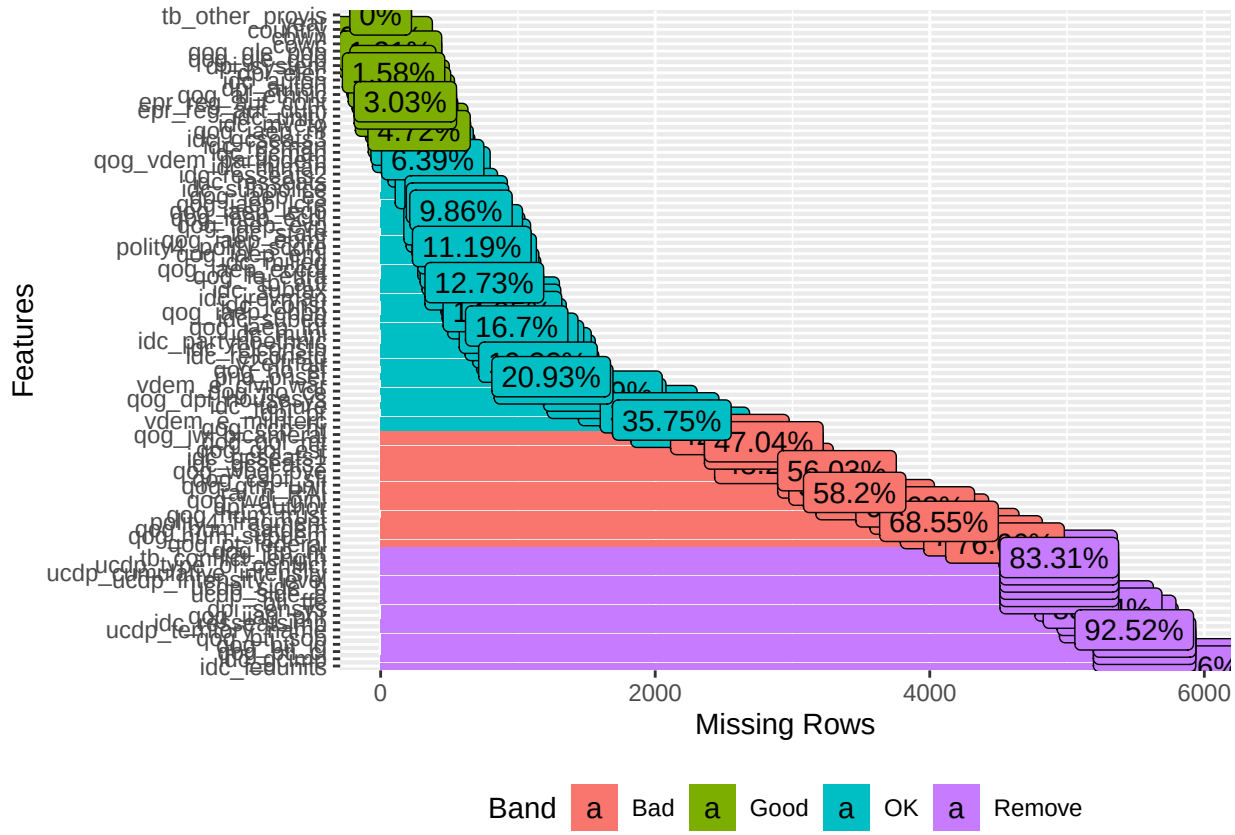
```
psp_add_bin <- psp_rename %>%

  # Other provisions
  dplyr::mutate(tb_other_provis = ifelse(idc_mveto == 1|
                                         idc_gcman == 1 |
                                           idc_gcimp == 1 |
                                           dpi_pr == 1, 1, 0),
  tb_other_provis = ifelse(is.na(tb_other_provis), 0, tb_other_provis)) %>%

  # Alternative measure of autonomy
  dplyr::mutate(tb_aut = ifelse(idc_subtax == 1 |
                                  idc_subed == 1 |
                                  idc_subpolice == 1, 1, 0)) %>%
  dplyr::group_by(ucdp_side_a, ucdp_side_b) %>%

  # Length of conflict
  dplyr::mutate(tb_conflict_length = sum(
    length(ucdp_cumulative_intensity), na.rm = T)
    ) %>%
  dplyr::ungroup()

psp_add_bin$tb_conflict_length[psp_add_bin$tb_conflict_length == 4943] <- NA
```

Save final data

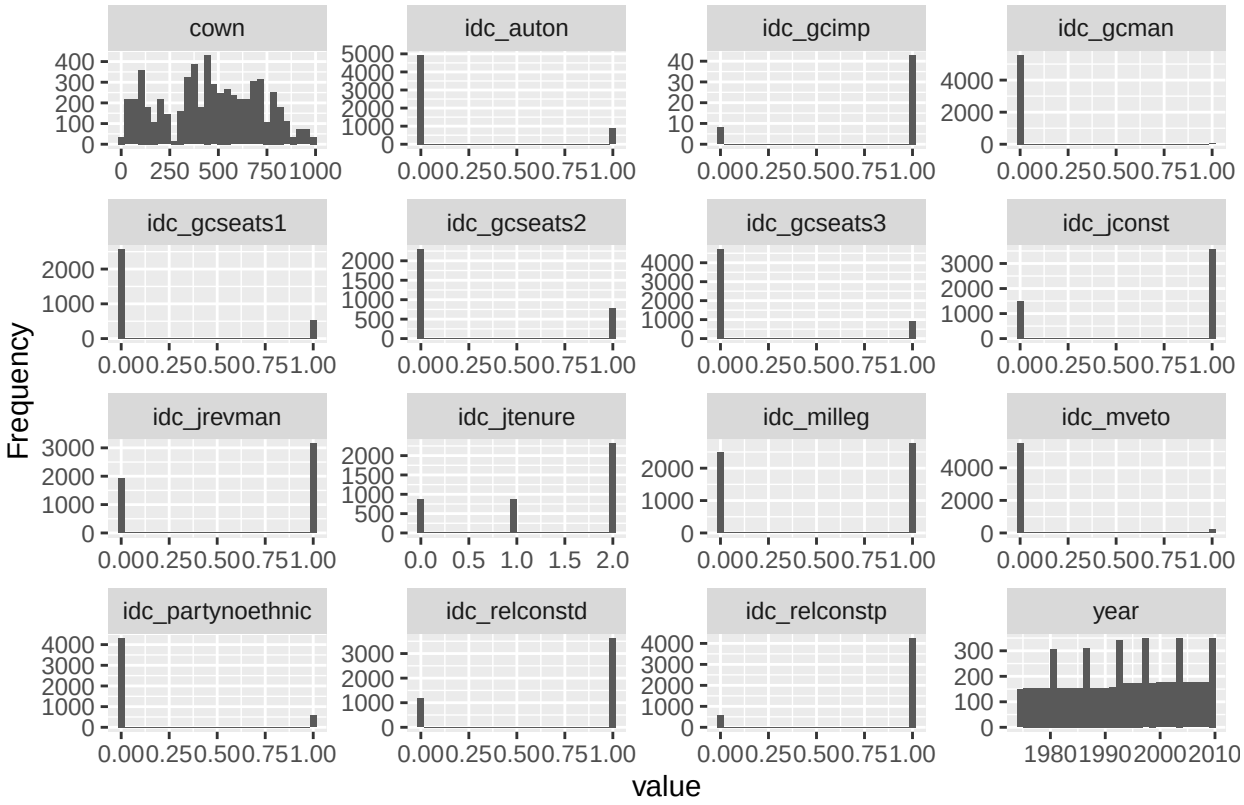
```
# Save as a .csv file
#write.csv(psp_add_bin, file = paste0(here::here(), "/data/tjbrailey_psp_clean.csv"))
```

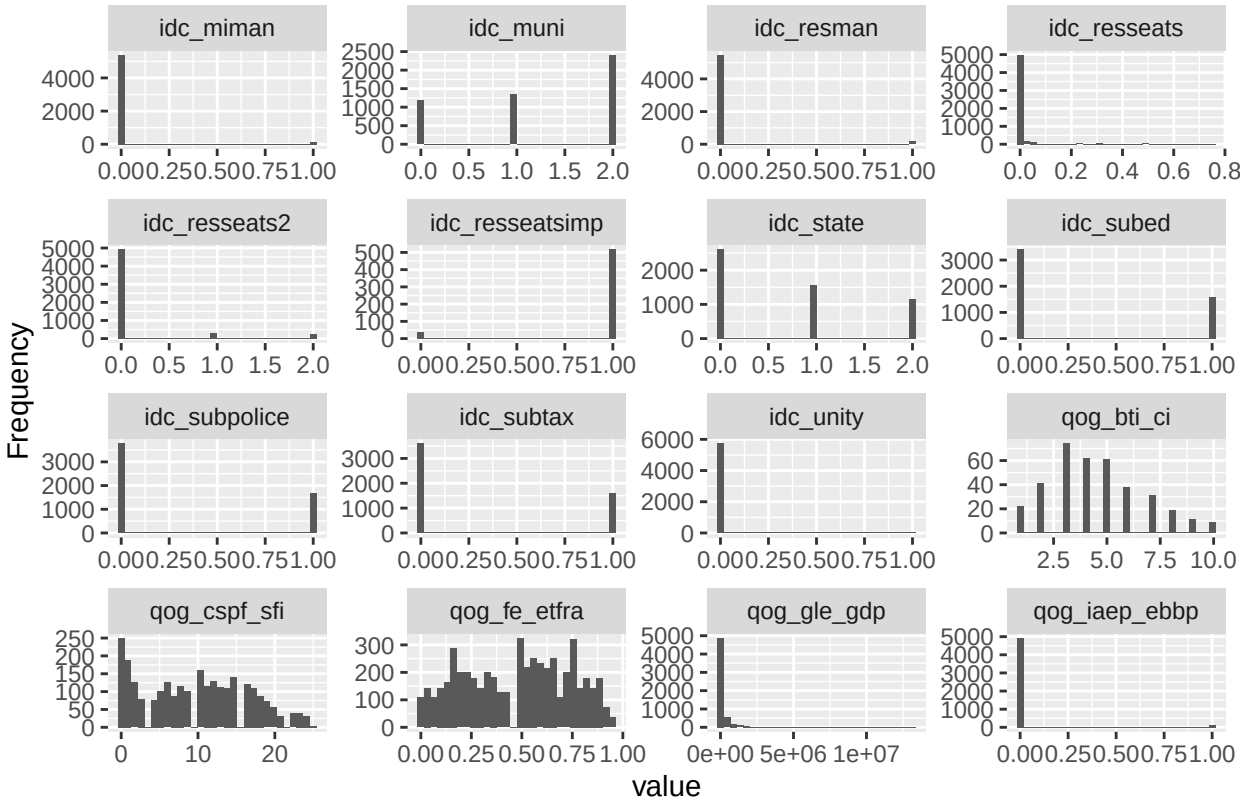
Visualize after recoding

```
DataExplorer::plot_missing(psp_add_bin)
```

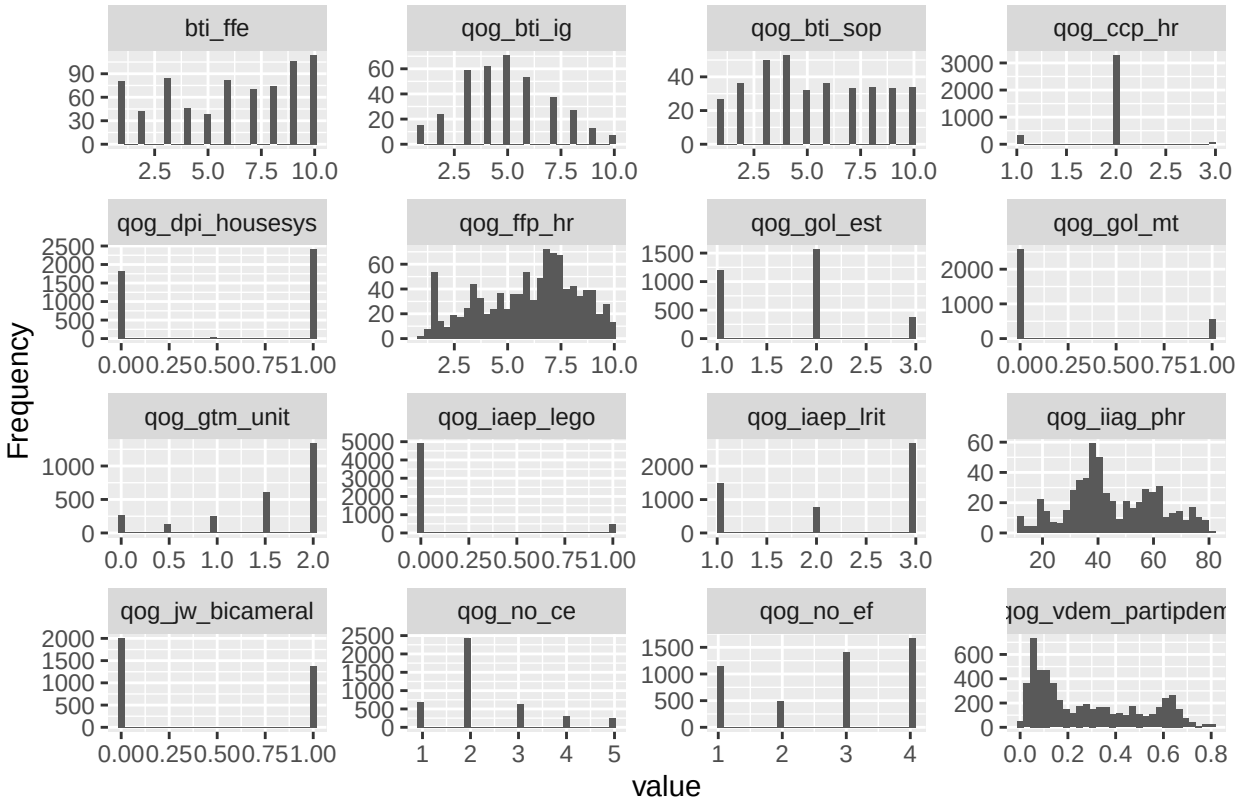


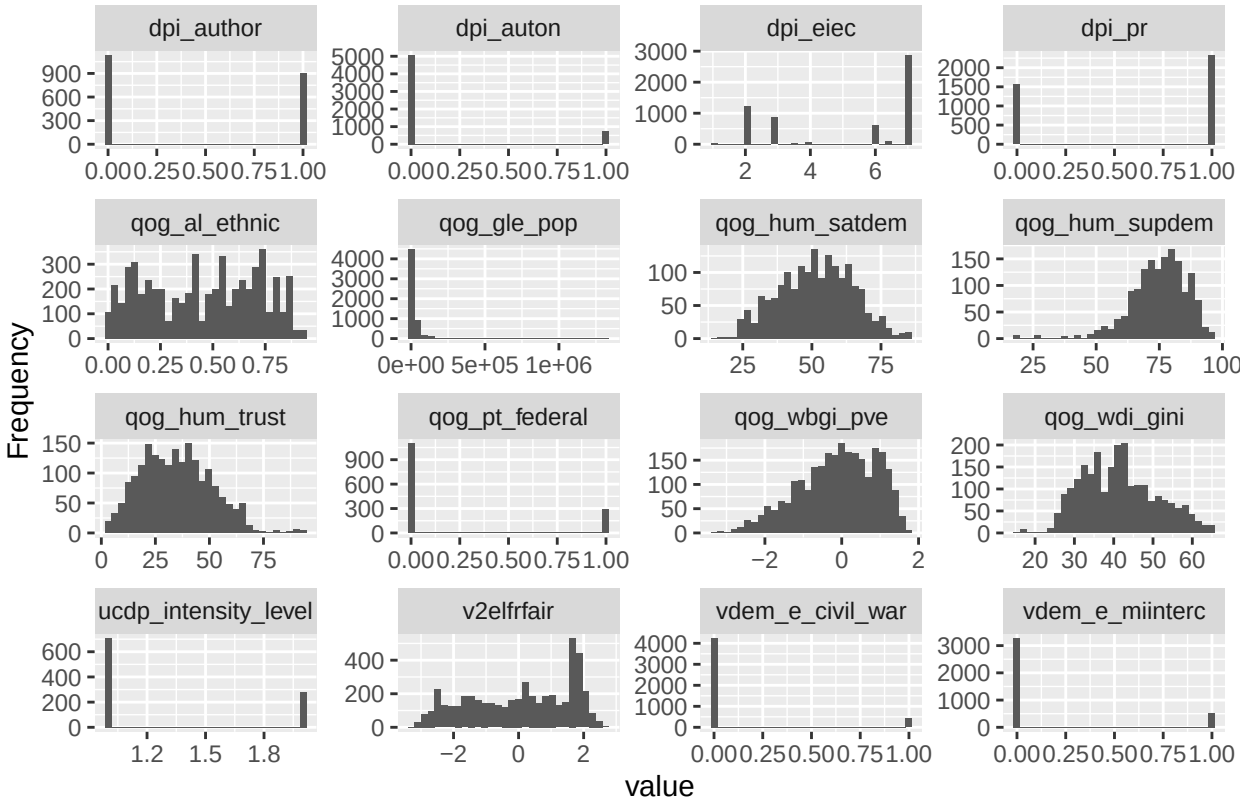

```
DataExplorer::plot_histogram(psp_add_bin)
```

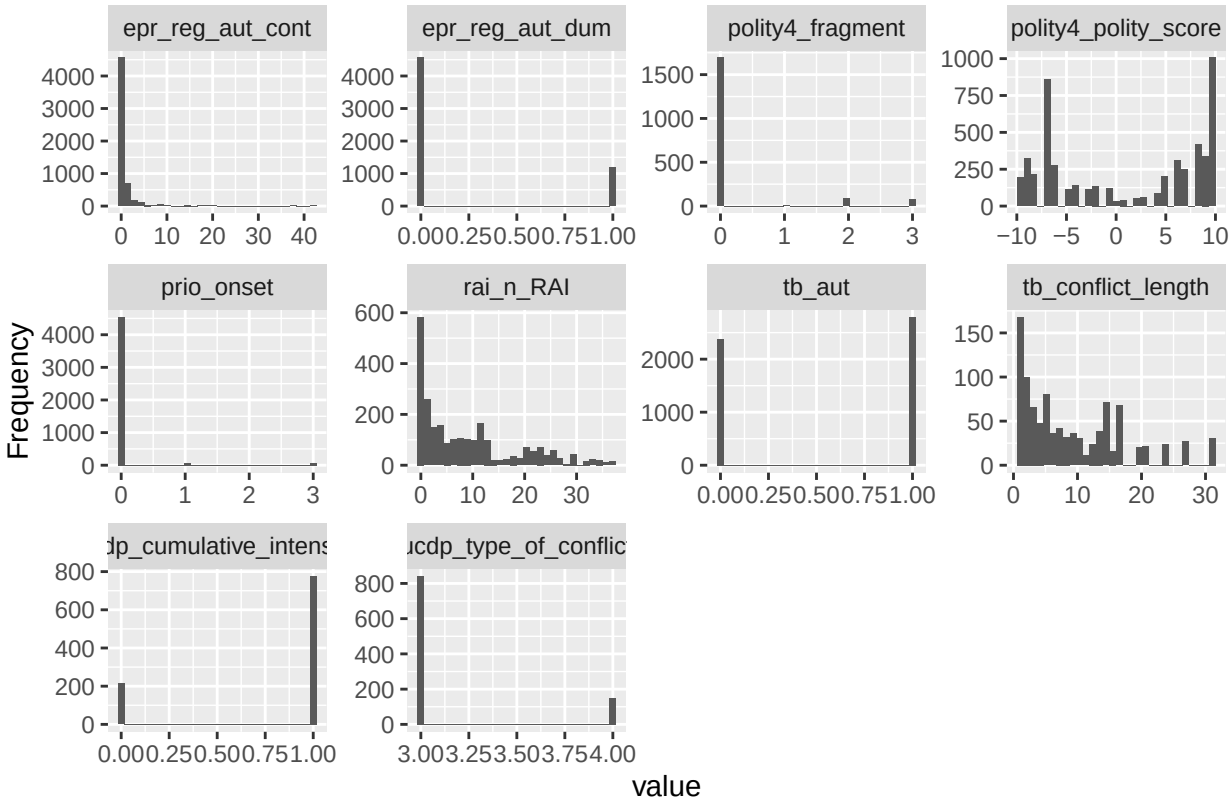

cown
idc_auton
idc_gcimp
idc_gcman
idc_gcseats1
idc_gcseats2
idc_gcseats3
idc_jconst
idc_jrevman
idc_jtenure
idc_milleg
idc_mveto
idc_partynoethnic
idc_relconstd
idc_relconstp
year
Frequency
value


idc_miman
idc_muni
idc_resman
idc_resseats
idc_resseats2
idc_resseatsimp
idc_state
idc_subed
idc_subpolice
idc_subtax
idc_unity
qog_bti_ci
qog_cspf_sfi
qog_fe_etfra
qog_gle_gdp
qog_iaep_ebbp
Frequency
value

Page 2

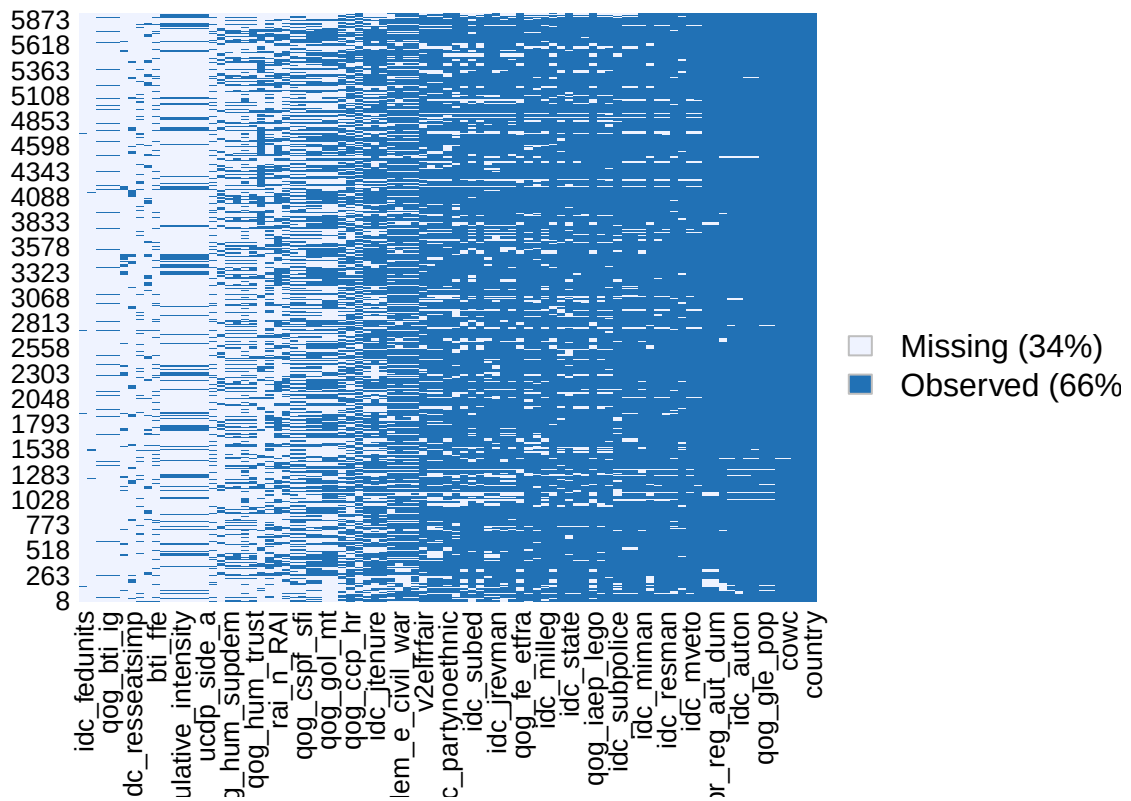

bti_ffe
qog_bti_ig
qog_bti_sop
qog_ccp_hr
qog_dpi_housesys
qog_ffp_hr
qog_gol_est
qog_gol_mt
qog_gtm_unit
qog_iaep_lego
qog_iaep_lrit
qog_iiag_phr
qog_jw_bicameral
qog_no_ce
qog_no_ef
qog_vdem_partipdem
Frequency
value


dpi_author
dpi_auton
dpi_eiec
dpi_pr
qog_al_ethnic
qog_gle_pop
qog_hum_satdem
qog_hum_supdem
qog_hum_trust
qog_pt_federal
qog_wbgi_pve
qog_wdi_gini
ucdp_intensity_level
v2elfrfair
vdem_e_civil_war
vdem_e_miinterc
Frequency
value




```
Amelia::missmap(psp_add_bin)
```

```
## Warning: Unknown or uninitialised column: 'arguments'.

## Warning: Unknown or uninitialised column: 'arguments'.

## Warning: Unknown or uninitialised column: 'imputations'.
```




```
png(file = paste0(here::here(), "/vis/tjbrailey_psp_clean_missingness.png"),
    width = 2000,
    height = 1000)
Amelia::missmap(psp_add_bin)
```

```
## Warning: Unknown or uninitialised column: 'arguments'.
```

```
## Warning: Unknown or uninitialised column: 'arguments'.
```

```
## Warning: Unknown or uninitialised column: 'imputations'.
```

```
dev.off()
```

```
## pdf
##   2
```