

R version 4.1.2 (2021-11-01) -- "Bird Hippie"
 Copyright (C) 2021 The R Foundation for Statistical Computing
 Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
 You are welcome to redistribute it under certain conditions.
 Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
 Type 'contributors()' for more information and
 'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
 'help.start()' for an HTML browser interface to help.
 Type 'q()' to quit R.

```
> path <- "Z:/Workbenches/epguest/seiro_ito/"
> pathprogram <- paste0(path, "outfiles/")
> pathdata <- paste0(path, "data/")
> pathdataCITIRP <- "Z:/Master Data/CIT-IRP5 Panel/"
> pathdataIRP <- "Z:/Master Data/IRP5/Job level/v5/beta/"
> library(qs); library(data.table); library(ggplot2)
qs 0.27.2. Announcement: https://github.com/qsbase/qs/issues/103
Warning message:
package 'qs' was built under R version 4.3.3
data.table 1.14.2 using 16 threads (see ?getDTthreads). Latest news: r-datatable.com
Warning message:
package 'data.table' was built under R version 4.1.3
Want to understand how all the pieces fit together? Read R for Data Science: https://r4ds.had.co.nz/
Warning message:
package 'ggplot2' was built under R version 4.1.3
>
> ipyr <- qread(paste0(pathdata, "irp12.qs"), nthreads = 15)
> ipyr <- ipyr[grepl("A", natureofperson), ]
> #### Not sure if we need this. Test how many obs will be dropped.
> nrow(ipyr[kerr_income != 0, ]); (nrow(ipyr[kerr_income != 0, ])/nrow(ipyr))*100
[1] 9517148
[1] 95.22896
>
> (n0 <- nrow(ipyr[is.na(payreferenceno) | payreferenceno == "", ])); (n0)/nrow(ipyr))*100
[1] 0
Error: unexpected ')' in " (n0)/nrow(ipyr))"
> (n0 <- nrow(ipyr[is.na(payreferenceno) | payreferenceno == "", ])); (n0)/nrow(ipyr)*100
[1] 0
[1] 0
> summary(ipyr[, payreferenceno])
  Length      Class      Mode
9993964 character character
> str(ipyr[, payreferenceno])
chr [1:9993964] "TZKTAKCTBK" "TZKTAKCTBK" "TZKTAKCTBK" "TZKTAKCTBK" "TZKTAKCTBK" "TZKTAKCTBK" "TZKTAKCTBK"
"TZKTAKCTBK" "ZJZXCKCGBK" "ZJZXCKCGBK" "ZJZXCKCGBK" "ZJZXCKCGBK" ...
> length(grep("^NA$", ipyr[, payreferenceno]))
[1] 0
> (n0 <- nrow(ipyr[is.nataxrefno | taxrefno == "", ])); (n0)/nrow(ipyr))*100
Error in .checkTypos(e, names_x) :
  Object 'is.nataxrefno' not found amongst irp5it3aid, ptrs_ind, kerr_ind, a3601, kerr_income
and 164 more
> (n0 <- nrow(ipyr[is.na(taxrefno) | taxrefno == "", ])); (n0)/nrow(ipyr))*100
[1] 1407896
Error: unexpected ')' in " (n0)/nrow(ipyr))"
> (n0 <- nrow(ipyr[is.na(taxrefno) | taxrefno == "", ])); (n0)/nrow(ipyr))*100
[1] 1407896
[1] 14.08746
> ipyr[, NumPrv := as.integer(.N),
+   by = .(busprov_geo, payreferenceno)]
> summary(ipyr[, NumPrv])
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    1      98      658   15965   4646  311648
> #### 1. Use payreferenceno to count the total
> #### 2. Compute the shares of each firms
```

```

> ##### a. In doing so, create a hypothetical "gov entity" to aggregate the entries with NAs
in taxrefno
> ##### b. Compute the shares of each firms including "gov entity" thence HHI
> ipyr[, Txrf := taxrefno]
> ipyr[is.na(taxrefno), Txrf := "GovEntity"]
> ##### Keep only Pretoria and Sandton
> ipSP <- ipyr[grepl("Sandton|Pretoria", busmainplc_geo), ]
> GeoLevel <- c("Prv", "Dis", "Loc", "Mai")
> LShare <- NULL
> for (g in 1:length(GeoLevel)) {
+   ipSP[, EachNum := as.integer(.N), by =
+     c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g], "Txrf")]
+   ipSP[, Total := as.integer(.N), by =
+     c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, Share := round(EachNum/Total, 6)]
+   ipSP[, HHI := sum(Share^2), na.rm = T]
+   , by = c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, AreaLevel := GeoLevel[g]]
+   lshare <- unique(ipyr[, c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g],
+     "taxrefno", "Txrf", "kerr_income", "AreaLevel", "Total", "Share", "HHI"), with = F])
+   LShare <- rbindlist(list(LShare, lshare), use.names = T, fill = T)
+ }
Error in `[.data.table`(ipyr, , c(c("busprov_geo", "busdistmuni_geo", :
column(s) not found: AreaLevel, Total, Share, HHI
> qsave(LShare, paste0(pathdata, "TestData_SandtonPretorial2.qs"), nthreads = 16)
> ipSP <- ipyr[grepl("Sandton|Pretoria", busmainplc_geo), ]
> GeoLevel <- c("Prv", "Dis", "Loc", "Mai")
> LShare <- NULL
> for (g in 1:length(GeoLevel)) {
+   ipSP[, EachNum := as.integer(.N), by =
+     c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g], "Txrf")]
+   ipSP[, Total := as.integer(.N), by =
+     c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, Share := round(EachNum/Total, 6)]
+   ipSP[, HHI := sum(Share^2), na.rm = T]
+   , by = c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, AreaLevel := GeoLevel[g]]
+   lshare <- unique(ipSP[, c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g],
+     "taxrefno", "Txrf", "kerr_income", "AreaLevel", "Total", "Share", "HHI"), with = F])
+   LShare <- rbindlist(list(LShare, lshare), use.names = T, fill = T)
+ }
> qsave(LShare, paste0(pathdata, "TestData_SandtonPretorial2.qs"), nthreads = 16)
> ### Keep only Pretoria and Sandton
> ipSP <- ipyr[grepl("Sandton|Pretoria", busmainplc_geo), ]
> GeoLevel <- c("Prv", "Dis", "Loc", "Mai")
> LShare <- NULL
> for (g in 1:length(GeoLevel)) {
+   ipSP[, EachNum := as.integer(.N), by =
+     c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g], "Txrf")]
+   ipSP[, Total := as.integer(.N), by =
+     c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, Share := round(EachNum/Total, 8)]
+   ipSP[, HHI := sum(Share^2), na.rm = T]
+   , by = c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, nHHI := (HHI-1/Total)/(1-1/Total)]
+   ##### ShareG, HHIG: Share and HHI after dropping GovEntity #####
+   ##### We compute "total without GovEntity" by subtracting GTotal (nrow of Entity==gov)
+   ipSP[grepl("gov", Entity), GTotal := as.integer(.N), by =
+     c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, GTotal := GTotal[!is.na(GTotal)][1], by =
+     c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ##### We define EachNumG by replacing EachNum of GovEntity with 0L
+   ipSP[, EachNumG := EachNum]
+   ipSP[grepl("gov", Entity), EachNumG := 0L]
+   ipSP[, TotalG := Total-GTotal]
+   ipSP[, ShareG := round(EachNumG/TotalG, 8)]
+   ipSP[, HHIG := sum(ShareG^2), na.rm = T]
+   , by = c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, nHHIG := (HHIG-1/TotalG)/(1-1/TotalG)]
+   ipSP[, AreaLevel := GeoLevel[g]]
+   lshare <- unique(ipSP[, c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g],

```

```

+   "taxrefno", "Entity", "Txrf", "kerr_income", "AreaLevel",
+   "Total", "Share", "HHI", "nHHI",
+   "TotalG", "ShareG", "HHIG", "nHHIG",
+   ), with = F))
+   LShare <- rbindlist(list(LShare, lshare), use.names = T, fill = T)
+ }
Error in .checkTypos(e, names_x) :
  Object 'Entity' not found amongst irp5it3aid, ptrs_ind, kerr_ind, a3601, kerr_income and 171
more
> qsave(LShare, paste0(pathdata, "TestData_SandtonPretorial2.qs"), nthreads = 16)
> ##### 1. Use payereferenceno to count the total
> ##### 2. Compute the shares of each firms
> ##### a. In doing so, create a hypothetical "gov entity" to aggregate the entries with NAs
in taxrefno
> ##### b. Compute the shares of each firms including "gov entity" thence HHI
> ipyr[, Txrf := taxrefno]
> ipyr[is.na(taxrefno), Txrf := "GovEntity"]
> ipyr[, Entity := "private"]
> ipyr[is.na(taxrefno), Entity := "gov"]
> ipyr[, Entity := factor(Entity, levels = c("private", "gov"))]
> ##### Keep only Pretoria and Sandton
> ipSP <- ipyr[grepl("Sandton|Pretoria", busmainplc_geo), ]
> GeoLevel <- c("Prv", "Dis", "Loc", "Mai")
> LShare <- NULL
> for (g in 1:length(GeoLevel)) {
+   ipSP[, EachNum := as.integer(.N), by =
+     c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g], "Txrf")]
+   ipSP[, Total := as.integer(.N), by =
+     c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, Share := round(EachNum/Total, 8)]
+   ipSP[, HHI := sum(Share^2), na.rm = T)
+   , by = c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, nHHI := (HHI-1/Total)/(1-1/Total)]
+   ##### ShareG, HHIG: Share and HHI after dropping GovEntity #####
+   ##### We compute "total without GovEntity" by subtracting GTotal (nrow of Entity==gov)
+   ipSP[grepl("gov", Entity), GTotal := as.integer(.N), by =
+     c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, GTotal := GTotal[!is.na(GTotal)][1], by =
+     c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ##### We define EachNumG by replacing EachNum of GovEntity with 0L
+   ipSP[, EachNumG := EachNum]
+   ipSP[grepl("gov", Entity), EachNumG := 0L]
+   ipSP[, TotalG := Total-GTotal]
+   ipSP[, ShareG := round(EachNumG/TotalG, 8)]
+   ipSP[, HHIG := sum(ShareG^2), na.rm = T)
+   , by = c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, nHHIG := (HHIG-1/TotalG)/(1-1/TotalG)]
+   ipSP[, AreaLevel := GeoLevel[g]]
+   lshare <- unique(ipSP[, c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainpl
c_geo")[1:g],
+   "taxrefno", "Entity", "Txrf", "kerr_income", "AreaLevel",
+   "Total", "Share", "HHI", "nHHI",
+   "TotalG", "ShareG", "HHIG", "nHHIG",
+   ), with = F])
+   LShare <- rbindlist(list(LShare, lshare), use.names = T, fill = T)
+ }
Error in c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g], :
  argument 15 is empty
> qsave(LShare, paste0(pathdata, "TestData_SandtonPretorial2.qs"), nthreads = 16)
> ##### Keep only Pretoria and Sandton
> ipSP <- ipyr[grepl("Sandton|Pretoria", busmainplc_geo), ]
> GeoLevel <- c("Prv", "Dis", "Loc", "Mai")
> LShare <- NULL
> for (g in 1:length(GeoLevel)) {
+   ipSP[, EachNum := as.integer(.N), by =
+     c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g], "Txrf")]
+   ipSP[, Total := as.integer(.N), by =
+     c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g]]
+   ipSP[, Share := round(EachNum/Total, 8)]
+   ipSP[, HHI := sum(Share^2), na.rm = T), by =
+     c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g]]
+   ipSP[, nHHI := (HHI-1/Total)/(1-1/Total)]
+   ##### ShareG, HHIG: Share and HHI after dropping GovEntity #####
+   ##### We compute "total without GovEntity" by subtracting GTotal (nrow of Entity==gov)

```

```

+ ipSP[grepl("gov", Entity), GTotal := as.integer(.N), by =
+   c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g]]
+ ipSP[, GTotal := GTotal[!is.na(GTotal)][1], by =
+   c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g]]
+ ##### We define EachNumG by replacing EachNum of GovEntity with 0L
+ ipSP[, EachNumG := EachNum]
+ ipSP[grepl("gov", Entity), EachNumG := 0L]
+ ipSP[, TotalG := Total-GTotal]
+ ipSP[, ShareG := round(EachNumG/TotalG, 8)]
+ ipSP[, HHIG := sum(ShareG^2, na.rm = T), by =
+   c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g]]
+ ipSP[, nHHIG := (HHIG-1/TotalG)/(1-1/TotalG)]
+ ipSP[, AreaLevel := GeoLevel[g]]
+ lshare <- unique(ipSP[, c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")
+   "taxrefno", "Entity", "Txrf", "kerr_income", "AreaLevel",
+   "Total", "Share", "HHI", "nHHI",
+   "TotalG", "ShareG", "HHIG", "nHHIG"
+   ), with = F)])
+ LShare <- rbindlist(list(LShare, lshare), use.names = T, fill = T)
+ }
Error in `[.data.table`(ipSP, , `:=`(Total, as.integer(.N)), by = c("busprov_geo", :
'by' appears to evaluate to column names but isn't c() or key(). Use by=list(...) if you can.
Otherwise, by=eval(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])
should work. This is for efficiency so data.table can detect which columns are needed.
> qsave(LShare, paste0(pathdata, "TestData_SandtonPretorial2.qs"), nthreads = 16)
> ipSP <- ipyr[grepl("Sandton|Pretoria", busmainplc_geo), ]
> GeoLevel <- c("Prv", "Dis", "Loc", "Mai")
> LShare <- NULL
> for (g in 1:length(GeoLevel)) {
+   ipSP[, EachNum := as.integer(.N), by =
+     eval(c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g], "Txrf")))]
+   ipSP[, Total := as.integer(.N), by =
+     eval(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, Share := round(EachNum/Total, 8)]
+   ipSP[, HHI := sum(Share^2, na.rm = T), by =
+     eval(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, nHHI := (HHI-1/Total)/(1-1/Total)]
+   ##### ShareG, HHIG: Share and HHI after dropping GovEntity #####
+   ##### We compute "total without GovEntity" by subtracting GTotal (nrow of Entity==gov)
+   ipSP[grepl("gov", Entity), GTotal := as.integer(.N), by =
+     eval(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, GTotal := GTotal[!is.na(GTotal)][1], by =
+     eval(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ##### We define EachNumG by replacing EachNum of GovEntity with 0L
+   ipSP[, EachNumG := EachNum]
+   ipSP[grepl("gov", Entity), EachNumG := 0L]
+   ipSP[, TotalG := Total-GTotal]
+   ipSP[, ShareG := round(EachNumG/TotalG, 8)]
+   ipSP[, HHIG := sum(ShareG^2, na.rm = T), by =
+     eval(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")[1:g])]
+   ipSP[, nHHIG := (HHIG-1/TotalG)/(1-1/TotalG)]
+   ipSP[, AreaLevel := GeoLevel[g]]
+   lshare <- unique(ipSP[, c(c("busprov_geo", "busdistmuni_geo", "buslocmuni_geo", "busmainplc_geo")
+     "taxrefno", "Entity", "Txrf", "kerr_income", "AreaLevel",
+     "Total", "Share", "HHI", "nHHI",
+     "TotalG", "ShareG", "HHIG", "nHHIG"
+     ), with = F)])
+   LShare <- rbindlist(list(LShare, lshare), use.names = T, fill = T)
+ }
> qsave(LShare, paste0(pathdata, "TestData_SandtonPretorial2.qs"), nthreads = 16)
> lapply(1:length(GeoLevel), function(i) summary(LShare[grepl(GeoLevel[i], AreaLevel), ]))
[[1]]
  busprov_geo      taxrefno      Entity      Txrf      kerr_income
AreaLevel      Total      Share      HHI      nHHI      Tota
lg      ShareG      HHIG
Length:58812      Length:58812      private:58812      Length:58812      Min.      :      0      L
ength:58812      Min.      :93247      Min.      :0.0000107      Min.      :10737      Min.      :10737      Min.      :
NA      Min.      : NA      Min.      :0
Class :character      Class :character      gov      :      0      Class :character      1st Qu.:      28152      C
lass :character      1st Qu.:93247      1st Qu.:0.0005362      1st Qu.:10737      1st Qu.:10737      1st Qu.:
NA      1st Qu.: NA      1st Qu.:0

```

```

Mode :character Mode :character Mode :character Median : 91615 M
ode :character Median :93247 Median :0.0249767 Median :10737 Median :10737 Median :
NA Median : NA Median :0

Mean : 156643
Mean :10737 Mean :10737 Mean :
NaN Mean :NaN Mean :0

3rd Qu.: 203938
NA 3rd Qu.: NA 3rd Qu.:0 3rd Qu.:10737 3rd Qu.:10737 3rd Qu.:

Max. :23801416
NA Max. : NA Max. :0 Max. :10737 Max. :10737 Max. :

NA's :6
NA's :

58812 NA's :58812
nHHIG busdistmuni_geo buslocmuni_geo busmainplc_geo
Min. : NA Length:58812 Length:58812 Length:58812
1st Qu.: NA Class :character Class :character Class :character
Median : NA Mode :character Mode :character Mode :character
Mean :NaN
3rd Qu.: NA
Max. : NA
NA's :58812

```

```

[[2]]
busprov_geo taxrefno Entity Txrf kerr_income
AreaLevel1 Total Share HHI nHHI
TotalG ShareG HHIG
Length:58955 Length:58955 private:58955 Length:58955 Min. : 0 L
ength:58955 Min. :25906 Min. :0.0000148 Min. : 171.6 Min. : 171.6 Min.
: NA Min. : NA Min. :0
Class :character Class :character gov : 0 Class :character 1st Qu.: 28167 C
lass :character 1st Qu.:25906 1st Qu.:0.0012177 1st Qu.: 171.6 1st Qu.: 171.6 1st
Qu.: NA 1st Qu.: NA 1st Qu.:0
Mode :character Mode :character Mode :character Median : 91532 M
ode :character Median :67341 Median :0.0286420 Median :15599.0 Median :15599.2 Medi
an : NA Median : NA Median :0

Mean : 156559
:NaN Mean :NaN Mean :0 Mean :11592.1 Mean :11592.3 Mean

3rd Qu.: 203822
Qu.: NA 3rd Qu.:67341 3rd Qu.:0.6138756 3rd Qu.:15599.0 3rd Qu.:15599.2 3rd

Max. :23801416
: NA Max. :67341 Max. :0.6138756 Max. :15599.0 Max. :15599.2 Max.

NA's :6
NA's

:58955 NA's :58955
nHHIG busdistmuni_geo buslocmuni_geo busmainplc_geo
Min. : NA Length:58955 Length:58955 Length:58955
1st Qu.: NA Class :character Class :character Class :character
Median : NA Mode :character Mode :character Mode :character
Mean :NaN
3rd Qu.: NA
Max. : NA
NA's :58955

```

```

[[3]]
busprov_geo taxrefno Entity Txrf kerr_income
AreaLevel1 Total Share HHI nHHI
TotalG ShareG HHIG
Length:58955 Length:58955 private:58955 Length:58955 Min. : 0 L
ength:58955 Min. :25906 Min. :0.0000148 Min. : 171.6 Min. : 171.6 Min.
: NA Min. : NA Min. :0
Class :character Class :character gov : 0 Class :character 1st Qu.: 28167 C
lass :character 1st Qu.:25906 1st Qu.:0.0012177 1st Qu.: 171.6 1st Qu.: 171.6 1st
Qu.: NA 1st Qu.: NA 1st Qu.:0
Mode :character Mode :character Mode :character Median : 91532 M
ode :character Median :67341 Median :0.0286420 Median :15599.0 Median :15599.2 Medi
an : NA Median : NA Median :0

Mean : 156559
:NaN Mean :NaN Mean :0 Mean :11592.1 Mean :11592.3 Mean

```

```

3rd Qu.:67341 3rd Qu.:0.6138756 3rd Qu.:15599.0 3rd Qu.: 203822
Qu.: NA 3rd Qu.: NA 3rd Qu.:0 3rd Qu.:15599.2 3rd
Max. :67341 Max. :0.6138756 Max. :15599.0 Max. :23801416
: NA Max. : NA Max. :0 Max. :15599.2 Max.
NA's :6 NA's
:58955 NA's :58955
nHHIG busdistmuni_geo buslocmuni_geo busmainplc_geo
Min. : NA Length:58955 Length:58955 Length:58955
1st Qu.: NA Class :character Class :character Class :character
Median : NA Mode :character Mode :character Mode :character
Mean :NaN
3rd Qu.: NA
Max. : NA
NA's :58955

[[4]]
busprov_geo taxrefno Entity Txrff kerr_income
AreaLevel Total Share HHI nHHI
TotalG ShareG HHIG private:58955 Length:58955 Min. : 0 L
ength:58955 Min. :25906 Min. :0.0000148 Min. : 171.6 Min. : 171.6 Min.
: NA Min. : NA Min. :0
Class :character Class :character gov : 0 Class :character 1st Qu.: 28167 C
lass :character 1st Qu.:25906 1st Qu.:0.0012177 1st Qu.: 171.6 1st Qu.: 171.6 1st
Qu.: NA 1st Qu.: NA 1st Qu.:0
Mode :character Mode :character Mode :character Median : 91532 M
ode :character Median :67341 Median :0.0286420 Median :15599.0 Median :15599.2 Medi
an : NA Median : NA Median :0
Mean :56579 Mean :0.2332478 Mean :11592.1 Mean :11592.3 Mean
:NaN Mean :NaN Mean :0
3rd Qu.:67341 3rd Qu.:0.6138756 3rd Qu.:15599.0 3rd Qu.: 203822
Qu.: NA 3rd Qu.: NA 3rd Qu.:0 3rd Qu.:15599.2 3rd
Max. :67341 Max. :0.6138756 Max. :15599.0 Max. :23801416
: NA Max. : NA Max. :0 Max. :15599.2 Max.
NA's :6 NA's
:58955 NA's :58955
nHHIG busdistmuni_geo buslocmuni_geo busmainplc_geo
Min. : NA Length:58955 Length:58955 Length:58955
1st Qu.: NA Class :character Class :character Class :character
Median : NA Mode :character Mode :character Mode :character
Mean :NaN
3rd Qu.: NA
Max. : NA
NA's :58955
```

>