Funds analysis

Jack Ye

August 2, 2016

Contents

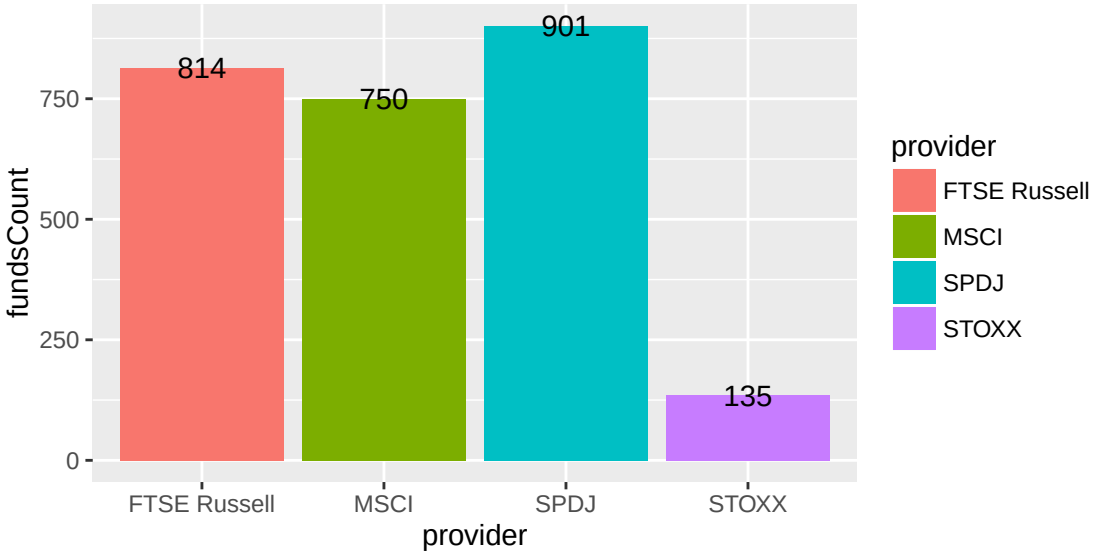
Golbal Configuration

```
filePath <- "~/Desktop/morningstar.csv"
sysKeyword <- "(Summary|Percentile|Sum|Average|Count|Maximum|Minimum|Median|Deviation)"
msciKeyword <-"MSCI"
ftseKeyword <-"(FTSE|Russell)"
stoxxKeyword <-"STOXX"
spdjKeyword <-"(S&P|Dow|DJ)"
esgKeywords <- '(Sustain|ESG|esg|SRI|sri|Social|Governance|Catholic|Ethical)'
envKeywords <-'(Water|Carbon|Climate|Enviro|Green|Energy|Renew|Tech|Fossil|Alternative|Clean|Fuel|Pollu
blackrockKeyword <- '(iShare|BlackRock|BLK|Blackrock)'
sumF <- function(vec) {sum(as.numeric(gsub(",", "", as.character(vec))), na.rm=TRUE)}
"^" <- function(x,y) ifelse(y==0,0,base:::"/"(x,y))
provider <- c("MSCI","FTSE Russell","STOXX","SPDJ")
esgIndexTotal <- c(5,35,81,67)
envIndexTotal <- c(20,20,20,20)
```

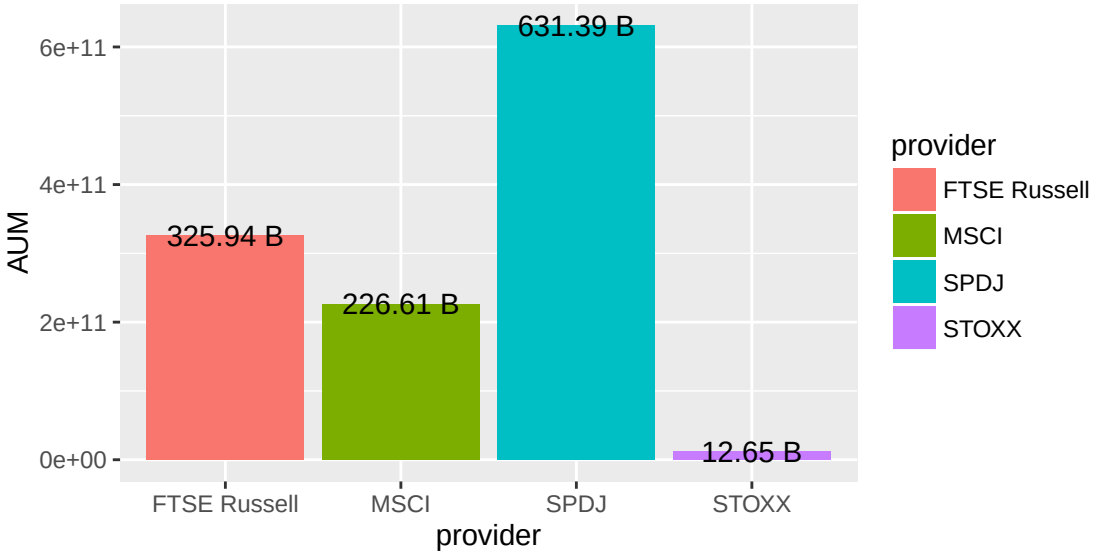
Analysis of funds tracking each provider

provider	fundsCount	fundsPercent	AUM
MSCI	750	0.2884615	226613780269
FTSE Russell	814	0.3130769	325942388219
STOXX	135	0.0519231	12645888332
SPDJ	901	0.3465385	631386740470

Count bar chart

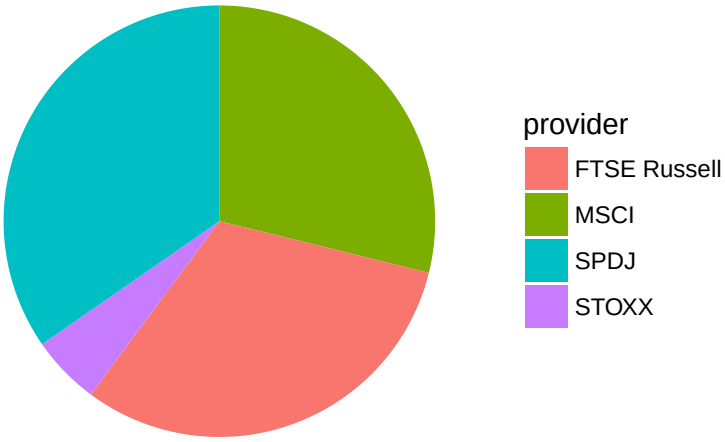


AUM bar chart



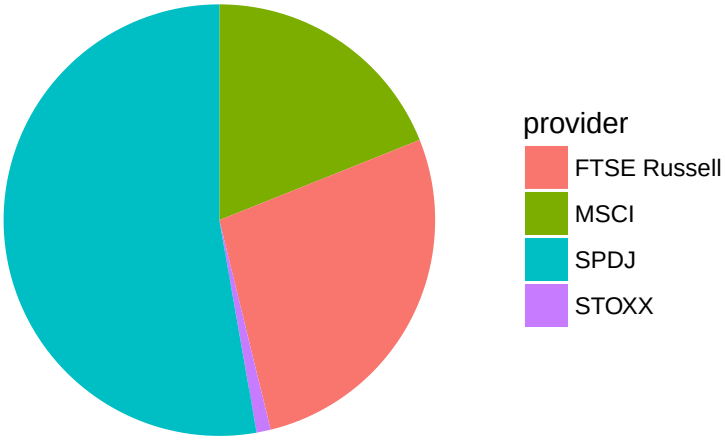
Count pie chart

-



AUM pie chart

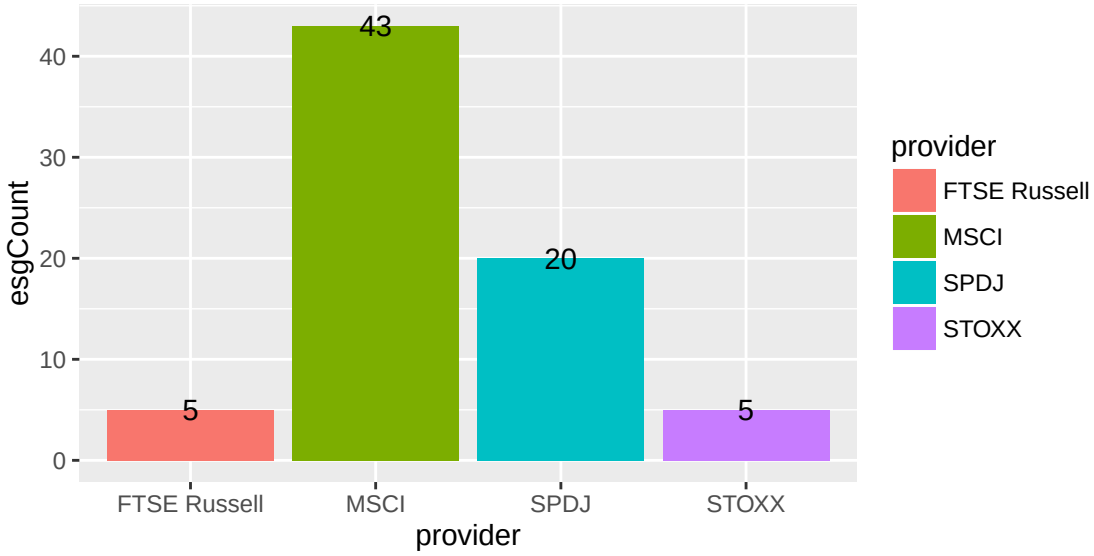
-



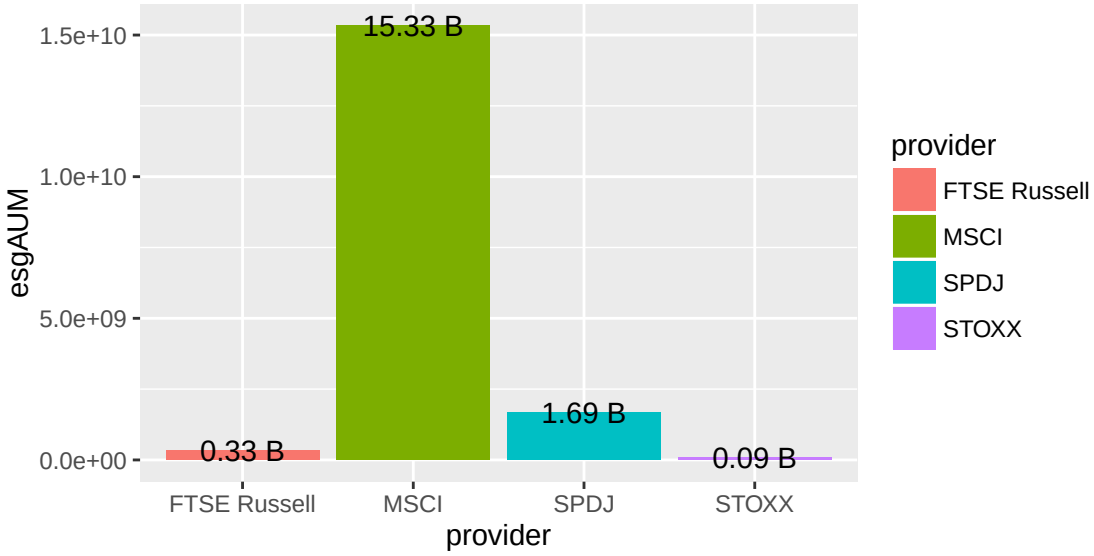
Analysis of funds tracking ESG indexes of each provider

provider	esgCount	esgPercent	esgAUM
MSCI	43	0.5890411	15334725344
FTSE Russell	5	0.0684932	328102278
STOXX	5	0.0684932	92539260
SPDJ	20	0.2739726	1691202927

Count bar chart

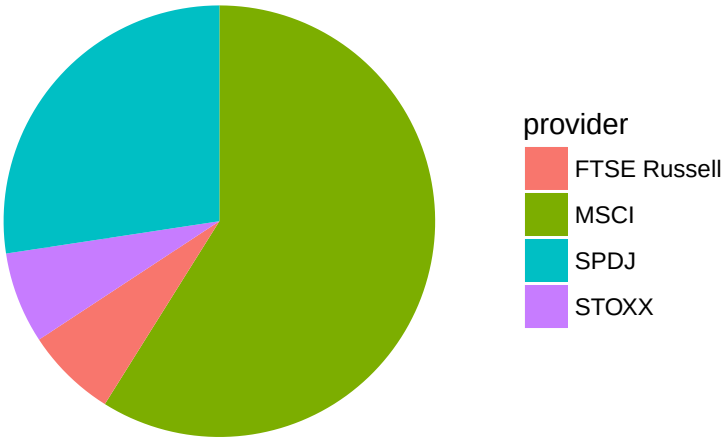


AUM bar comparison



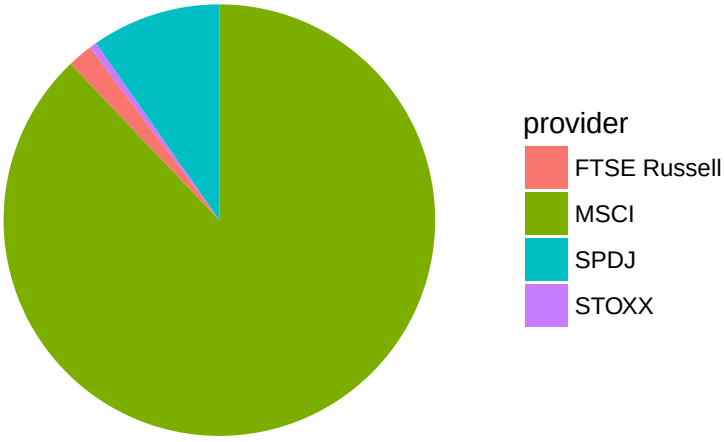
Count pie chart

-



AUM pie chart

-



Percentage of funds tracking ESG indexes for each provider

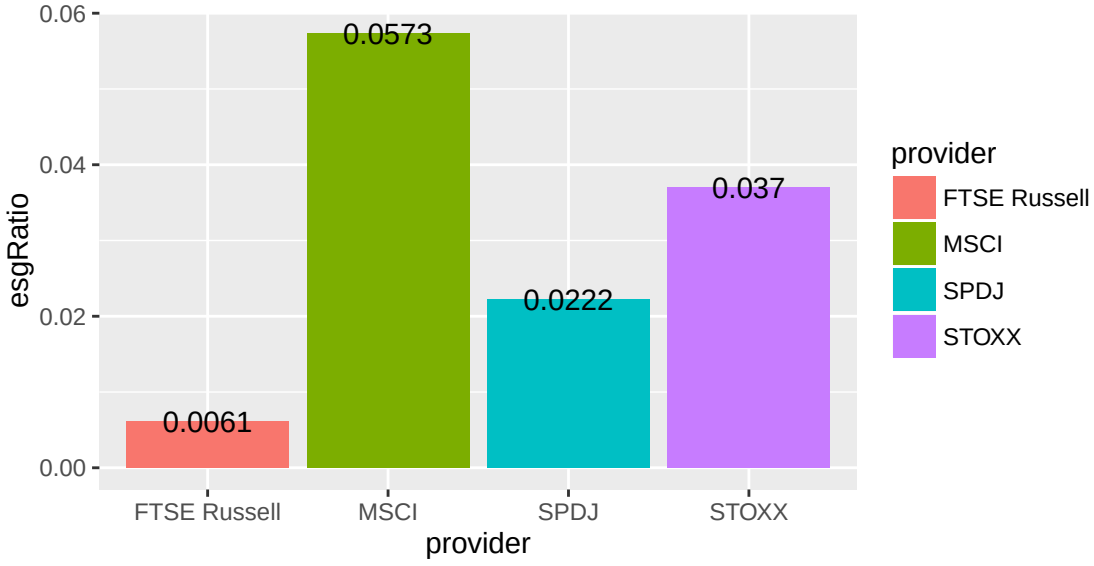
total tracking ESG count: 73

total funds count: 2600

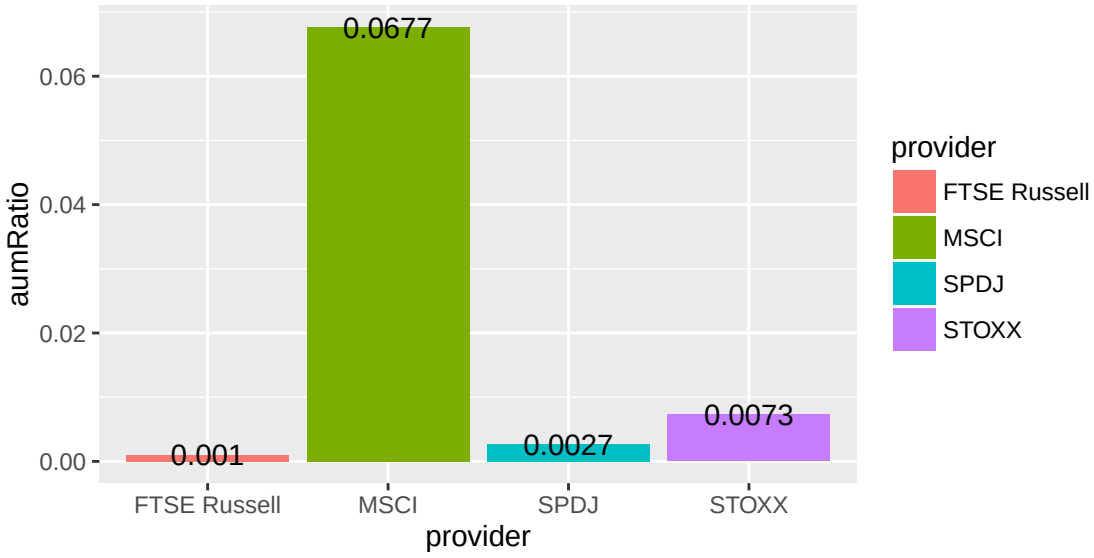
percentage count of total: 0.0280769

percentage AUM of total: 0.0145803

Count ratio for each provider



AUM ratio for each provider

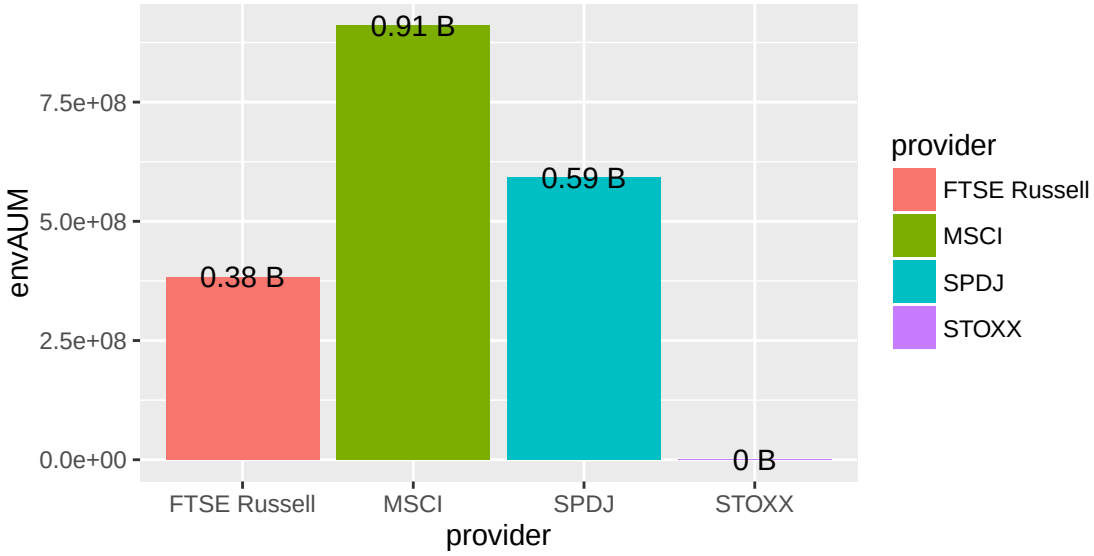


Analysis of funds tracking Environment indexes of each provider

provider	envCount	envPercent	envAUM
MSCI	9	0.4285714	910962321
FTSE Russell	3	0.1428571	383672454
STOXX	0	0.0000000	0
SPDJ	9	0.4285714	591649224

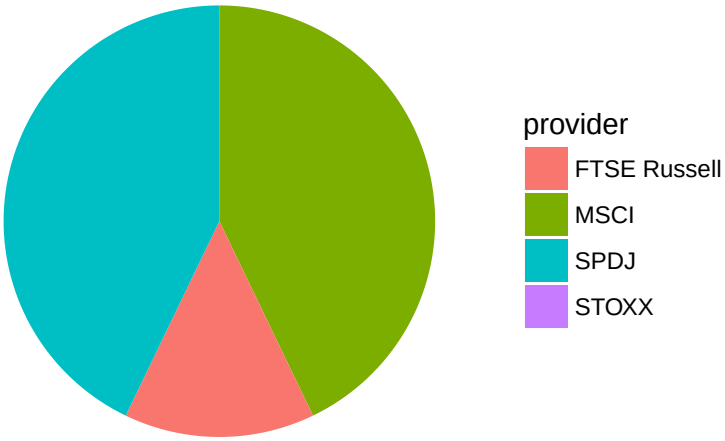
Count bar chart

AUM bar chart



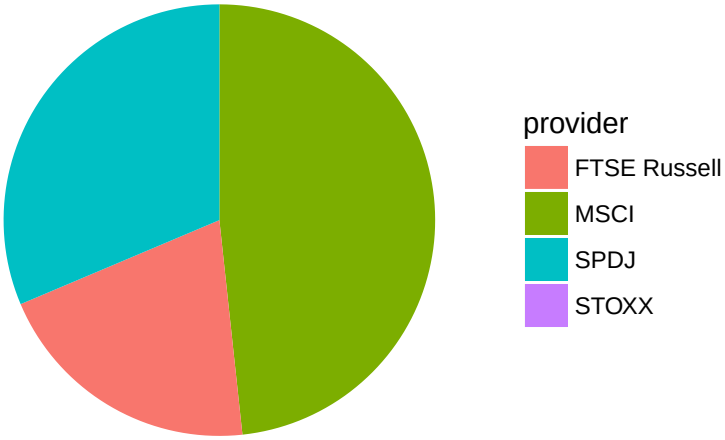
Count pie chart

-



AUM pie chart

-



Percentage of Environment funds tracking Environment indexes for each provider

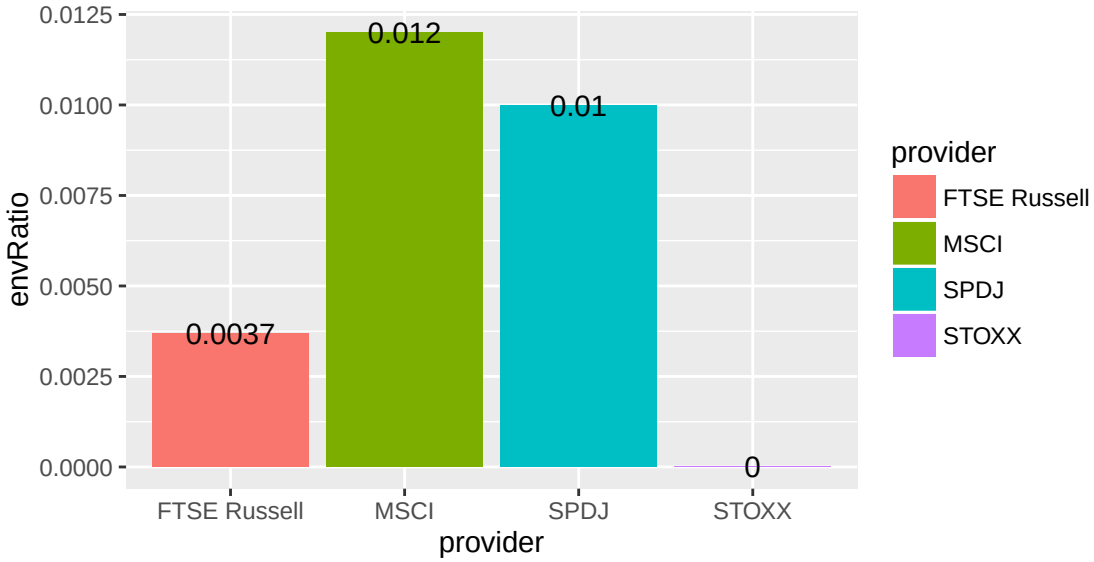
total tracking Environment count: 21

total funds count: 2600

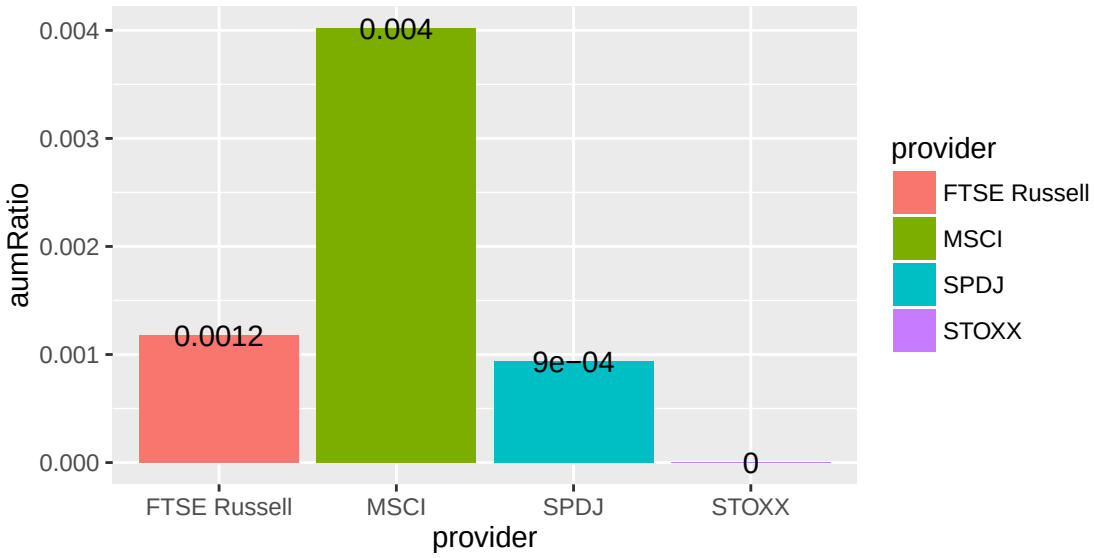
percentage count of total: 0.0080769

percentage AUM of total: 0.0015764

Count ratio for each provider

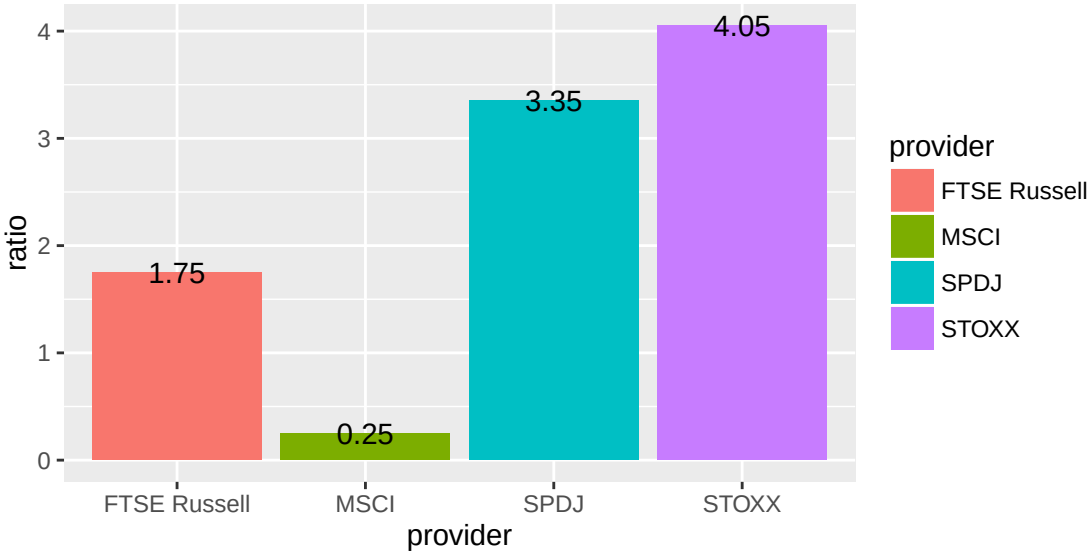


AUM ratio for each provider

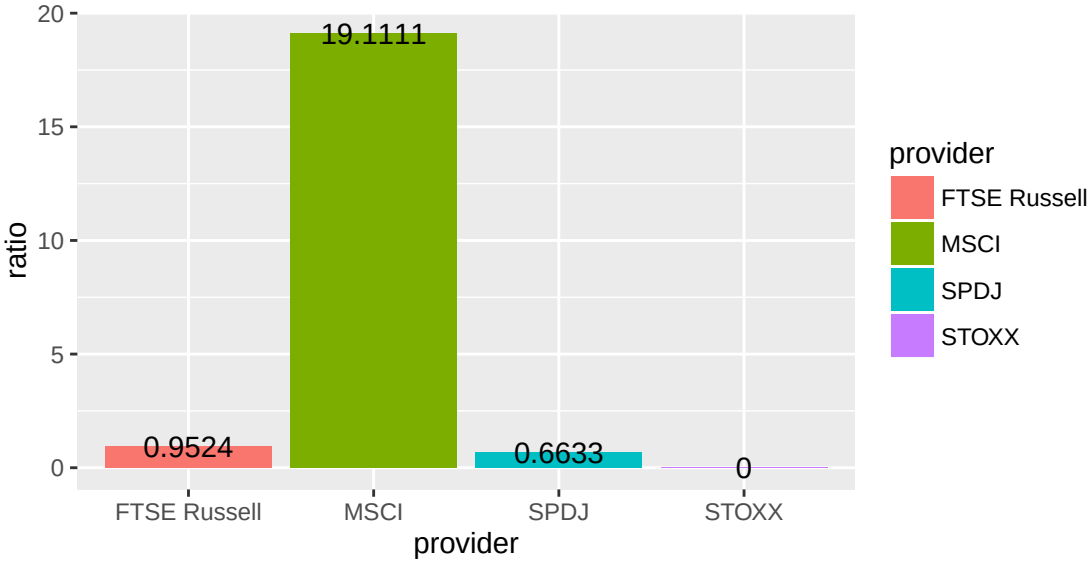


ESG & Environment cross comparison

ESG index count VS Environment index count



Percentage of ESG index in use VS Percentage of Environment index in use

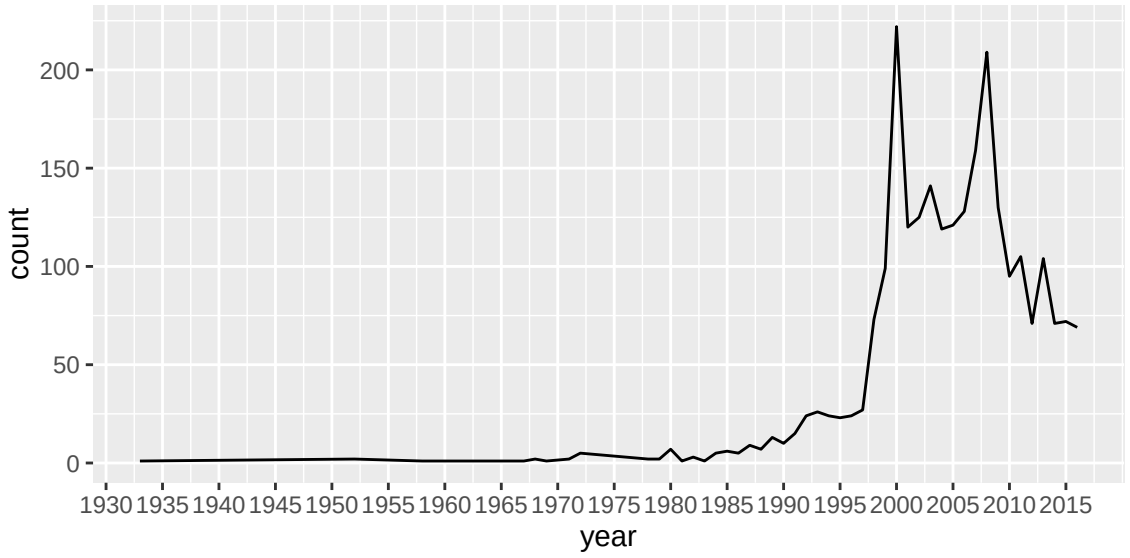


Function 1: get timeseries given any universe and keyword

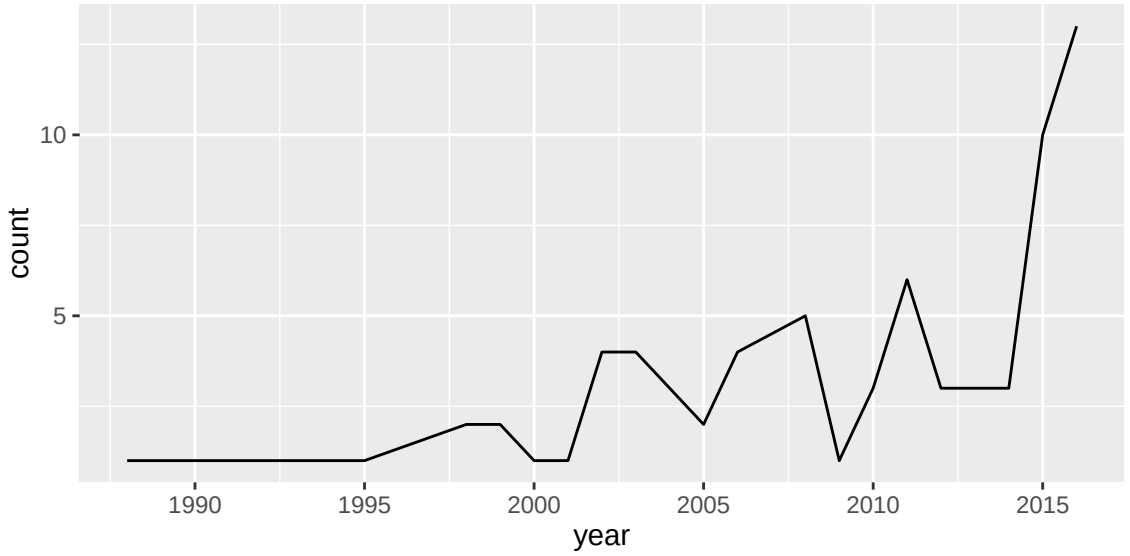
```
getYearCountForUniverse <- function(universe, keyword="", yearLimit=1930) {
  dateU <- universe %>%
    filter(grepl(keyword, Name)) %>%
    filter(grepl("/",as.character(Inception.Date))) %>%
    mutate(date=as.character(Inception.Date))
  yearVec <- dateU$date
  for (i in 1:length(yearVec)) {
    d <- yearVec[i]
    st <- substring(d, nchar(d)-1, nchar(d))
    if (as.numeric(st) < 17) {
      yearVec[i] <- as.numeric(paste("20", st, sep=""))
    } else {
      yearVec[i] <- as.numeric(paste("19", st, sep=""))
    }
  }
  yearVec <- as.numeric(yearVec)
  yearDF <- data.frame(year=yearVec) %>%
    group_by(year) %>%
    filter(year >= yearLimit) %>%
    summarize(count=n())
  return(yearDF)
}

plotTimeSeriesForUniverse <- function(universe, keyword="",  yearLimit=1930) {

  getYearCountForUniverse(universe, keyword, yearLimit) %>%
    ggplot(aes(x=year, y=count)) +
    geom_line() +
    scale_x_continuous(breaks=seq(yearLimit, 2016, 5))
}

plotTimeSeriesForAllScope <- function(universe, esgU, envU, keyword="",  yearLimit=1930) {
  yearDF <- getYearCountForUniverse(universe, keyword, yearLimit)
  esgDF <- getYearCountForUniverse(esgU, keyword, yearLimit)
  envDF <- getYearCountForUniverse(envU, keyword, yearLimit)
  colnames(esgDF) <- c('year', 'esgCount')
  colnames(envDF) <- c('year', 'envCount')
  yearDF %>%
    full_join(esgDF) %>%
    full_join(envDF) %>%
    ggplot() +
    geom_line(aes(x=year, y=count), color="green") +
    geom_line(aes(x=year, y=esgCount), color='red') +
    geom_line(aes(x=year, y=envCount), color='blue') +
    scale_x_continuous(breaks=seq(yearLimit, 2016, 5))
}

plotTimeSeriesForEsgEnv <- function(esgU, envU, keyword="",  yearLimit=1930) {
  esgDF <- getYearCountForUniverse(esgU, keyword, yearLimit)
  envDF <- getYearCountForUniverse(envU, keyword, yearLimit)
  colnames(esgDF) <- c('year', 'esgCount')
```

```
  colnames(envDF) <- c('year', 'envCount')
  esgDF %>%
    full_join(envDF) %>%
    ggplot() +
    geom_line(aes(x=year, y=esgCount), color='red') +
    geom_line(aes(x=year, y=envCount), color='blue') +
    scale_x_continuous(breaks=seq(yearLimit, 2016, 5))
}
```

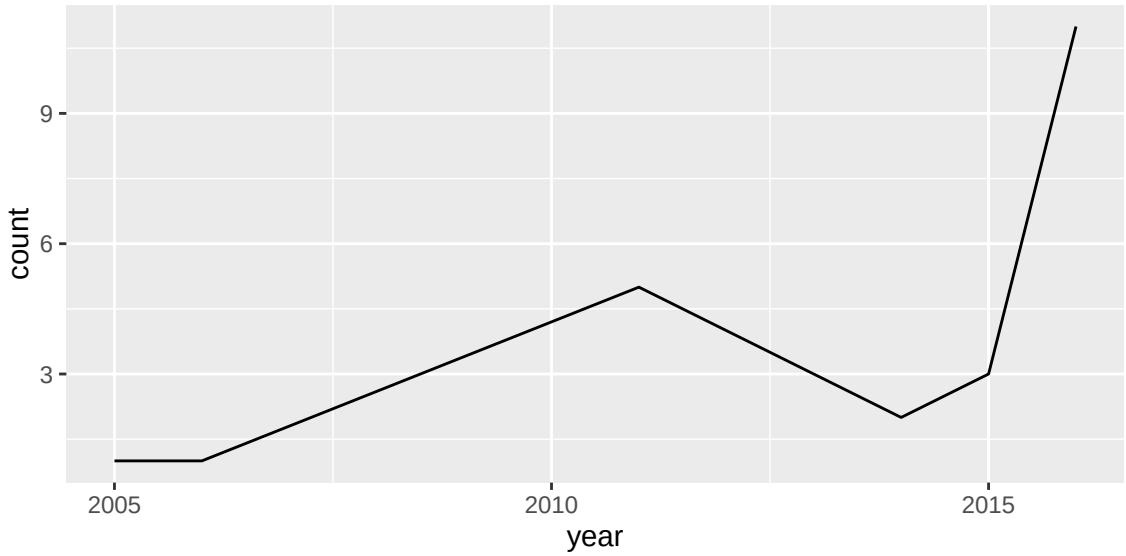
example: universe time series



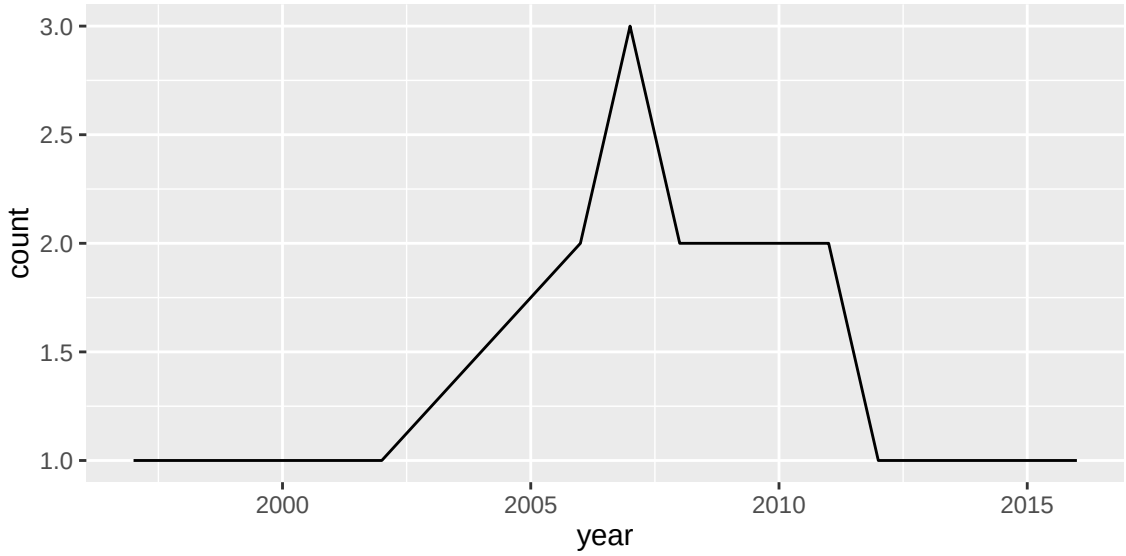
example: ESG time series



example: MSCI ESG time series



example: SPDJ time series



Function 2: check any keyword's distribution of index provider

```
institutionChoice <- function(keyword, universe) {
  companyUniverse <- universe %>% filter(grepl(keyword, Name))
  MSCI <- companyUniverse %>% filter(grepl(msciKeyword,Primary.Prospectus.Benchmark))
  FTSE <- companyUniverse %>% filter(grepl(ftseKeyword,Primary.Prospectus.Benchmark))
  STOXX <- companyUniverse %>% filter(grepl(stoxxKeyword,Primary.Prospectus.Benchmark))
  SPDJ <- companyUniverse %>% filter(grepl(spdjKeyword,Primary.Prospectus.Benchmark))
  fundsCount <- c(length(MSCI$Name),length(FTSE$Name),length(STOXX$Name),length(SPDJ$Name))
  fundsPercent <- fundsCount/sum(fundsCount)
  countTable <- data.frame(provider=provider, fundsCount=fundsCount, fundsPercent=fundsPercent)
  return(countTable)
}

plotInstitutionChoice <- function(keyword, universe) {
  companyUniverse <- universe %>% filter(grepl(keyword, Name))
  MSCI <- companyUniverse %>% filter(grepl(msciKeyword,Primary.Prospectus.Benchmark))
  FTSE <- companyUniverse %>% filter(grepl(ftseKeyword,Primary.Prospectus.Benchmark))
  STOXX <- companyUniverse %>% filter(grepl(stoxxKeyword,Primary.Prospectus.Benchmark))
  SPDJ <- companyUniverse %>% filter(grepl(spdjKeyword,Primary.Prospectus.Benchmark))
  fundsCount <- c(length(MSCI$Name),length(FTSE$Name),length(STOXX$Name),length(SPDJ$Name))
  fundsPercent <- fundsCount/sum(fundsCount)
  data.frame(provider=provider, fundsCount=fundsCount, fundsPercent=fundsPercent) %>%
    ggplot(aes(x=provider, y=fundsCount, fill=provider)) +
    geom_bar(stat="identity") +
    geom_text(aes(label=fundsCount))
}
```

exsample: BlackRock

provider	fundsCount	fundsPercent
MSCI	20	0.4166667
FTSE Russell	18	0.3750000
STOXX	0	0.0000000
SPDJ	10	0.2083333

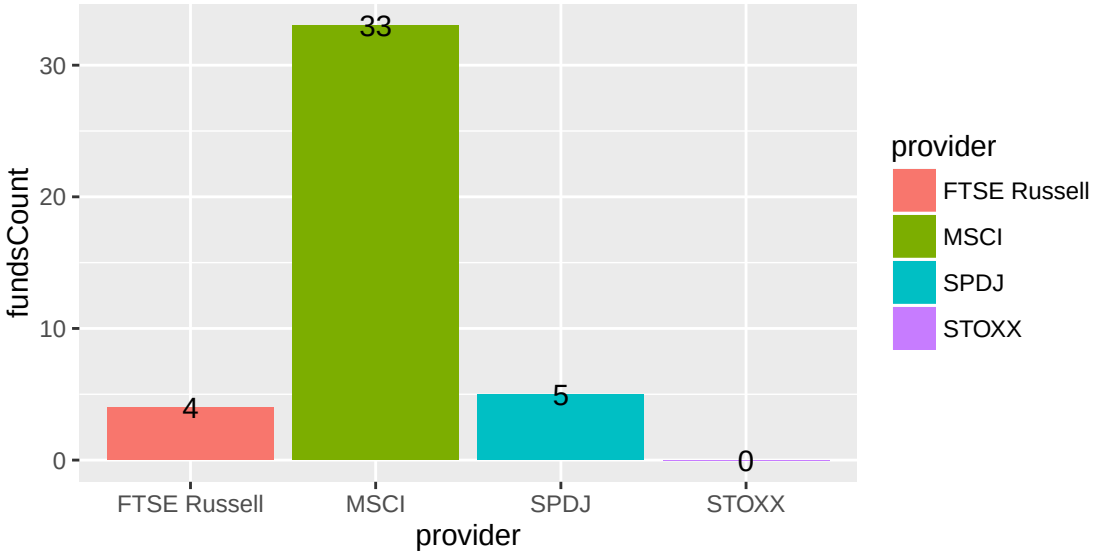
exsample: BlackRock ESG

provider	fundsCount	fundsPercent
MSCI	6	0.6666667
FTSE Russell	0	0.0000000
STOXX	0	0.0000000
SPDJ	3	0.3333333

exsample: ETF

provider	fundsCount	fundsPercent
MSCI	33	0.7857143
FTSE Russell	4	0.0952381

provider	fundsCount	fundsPercent
STOXX	0	0.0000000
SPDJ	5	0.1190476



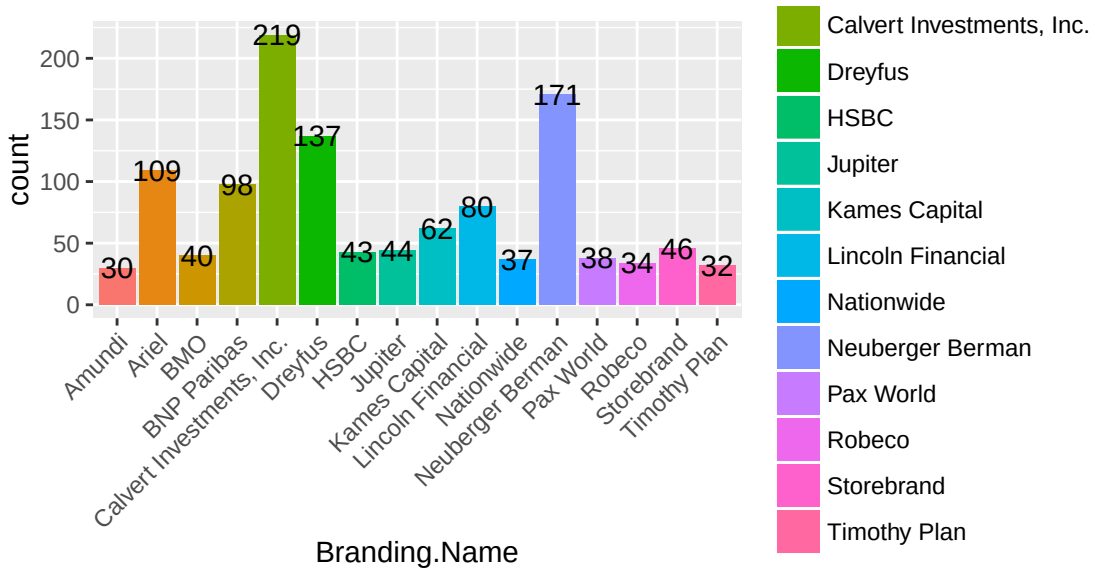
exsample: HSBC

provider	fundsCount	fundsPercent
MSCI	9	0.18
FTSE Russell	0	0.00
STOXX	0	0.00
SPDJ	41	0.82

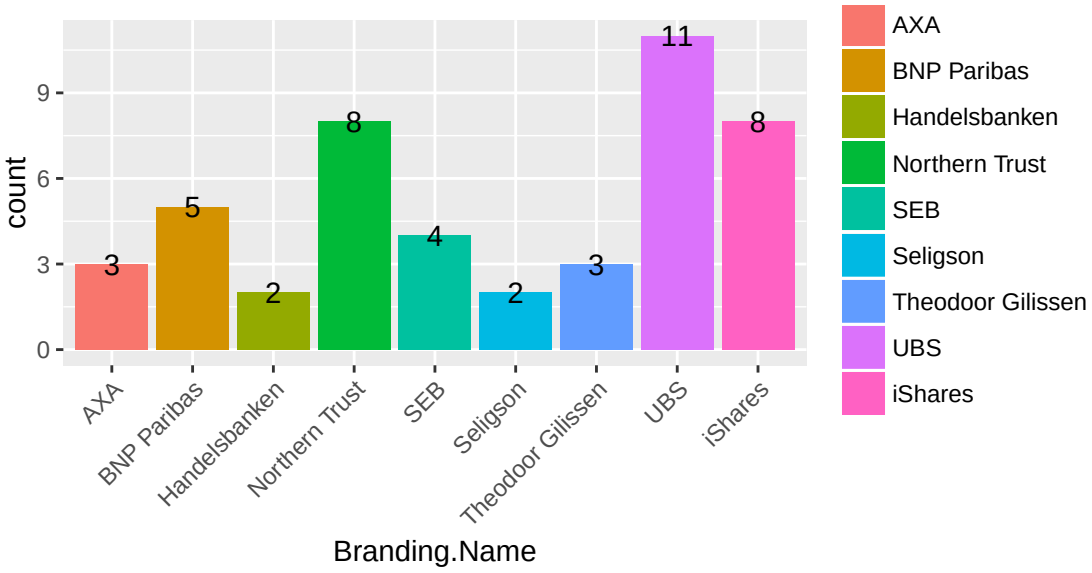
Function 3: check any company's index choice provider popularity

```
plotUniverseByBrandName <- function(universe, floor=5) {
  universe %>%
    group_by(Branding.Name) %>%
    summarize(count=n()) %>%
    filter(count >= floor) %>%
    ggplot(aes(x=Branding.Name, y=count, fill=Branding.Name)) +
    geom_bar(stat="identity") +
    geom_text(aes(label=count)) +
    theme(axis.text.x = element_text(angle = 45, hjust = 1))
}
```

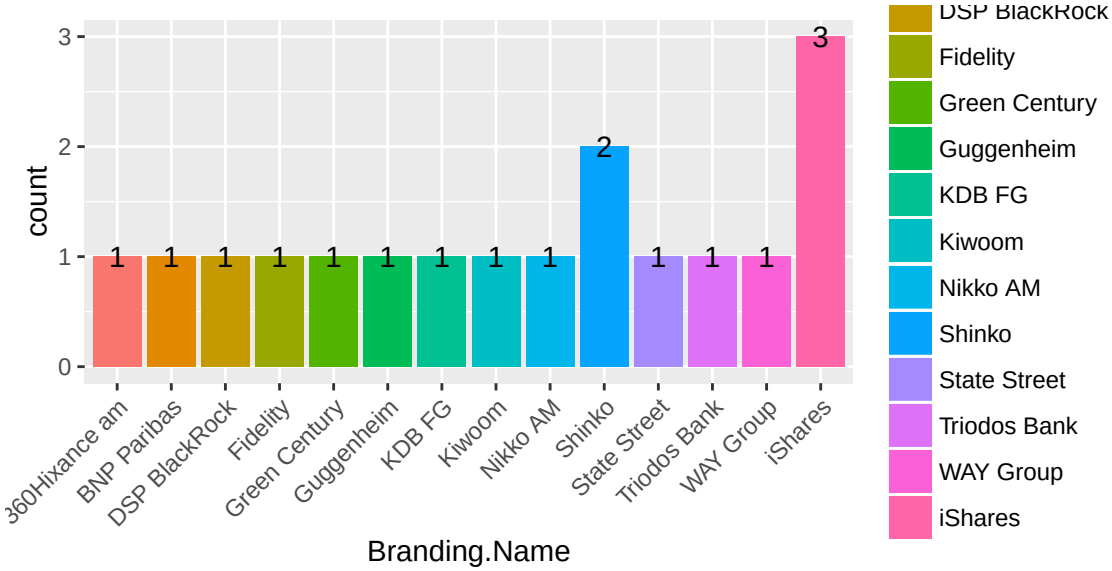
exsample: universe plot



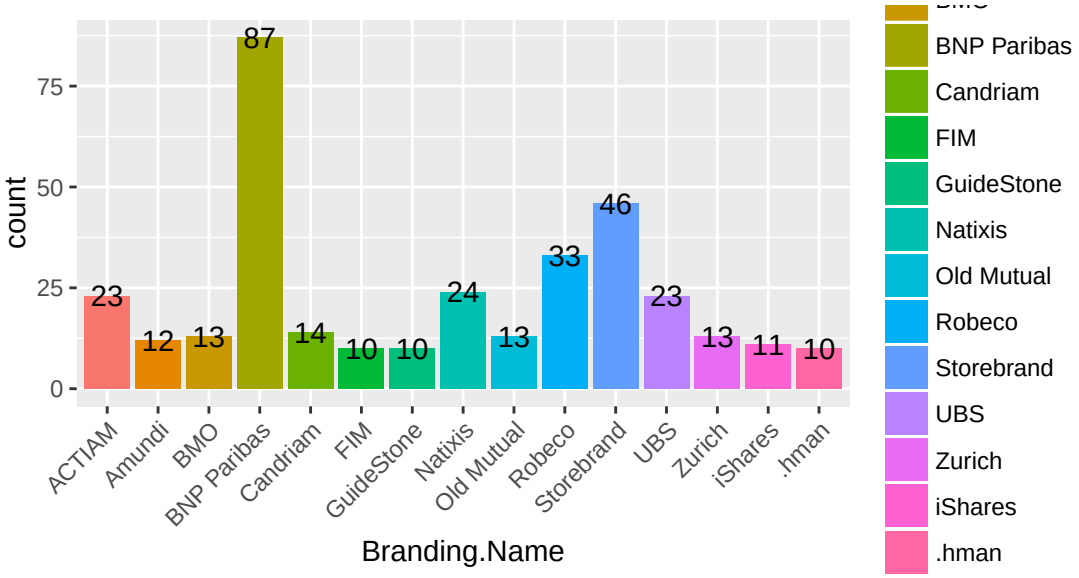
exsample: universe ESG plot



exsample: universe ENV plot



exsample: MSCI plot



exsample: MSCI ESG plot