exercise_2_code+output.R

Emmanuel

Wed Mar 27 14:29:46 2019

```
library(tidyverse)

## -- Attaching packages ---------------------------------------- tidyverse
1.2.1 --

## v ggplot2 3.1.0     v purrr   0.3.0
## v tibble  2.0.1     v dplyr   0.7.8
## v tidyr   0.8.2     v stringr 1.3.1
## v readr   1.3.1     v forcats 0.3.0

## -- Conflicts ------------------------------------------------
tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(tibble)
library(ggplot2)
library(dplyr)

times <- read.csv("timesData.csv")
timesdf <- data.frame(times)

names(timesdf)

##  [1] "world_rank"             "university_name"
##  [3] "country"                "teaching"
##  [5] "international"          "research"
##  [7] "citations"              "income"
##  [9] "total_score"            "num_students"
## [11] "student_staff_ratio"    "international_students"
## [13] "female_male_ratio"      "year"

shangai <- read.csv("shanghaiData.csv")
shangaidf <- data.frame(shangai)

names(shangaidf)

##  [1] "world_rank"      "university_name" "national_rank"
##  [4] "total_score"     "alumni"          "award"
##  [7] "hici"            "ns"              "pub"
## [10] "pcp"             "year"
```

```r
cwur <- read.csv("cwurData.csv")
cwurdf <- data.frame(cwur)

names(cwurdf)
```

```
##  [1] "world_rank"           "institution"          "country"
##  [4] "national_rank"        "quality_of_education" "alumni_employment"
##  [7] "quality_of_faculty"   "publications"         "influence"
## [10] "citations"            "broad_impact"         "patents"
## [13] "score"                "year"
```

```r
#comparing the mean of the total scores from three datasets(cwur, shangai and
times)
#used to determine the world ranking of universities in the year 2014.


timesfilter <- timesdf %>% filter(year == "2014")
timesnumeric <- as.numeric(timesfilter$total_score)
timesomit <- na.omit(timesnumeric)
timesmean <- mean(timesomit)
timesmean
```

```
## [1] 71.3275
```

```r
cwurfilter <- cwurdf %>% filter(year == "2014")
cwurnumeric <- as.numeric(cwurfilter$score)
cwuromit <- na.omit(cwurnumeric)
cwurmean <- mean(cwuromit)
cwurmean
```

```
## [1] 47.27141
```

```r
shangaifilter <- shangaidf %>% filter(year == "2014")
shangainumeric <- as.numeric(shangaifilter$total_score)
shangaiomit <- na.omit(shangainumeric)
shangaimean <- mean(shangaiomit)
shangaimean
```

```
## [1] 36.172
```