Theil

2022-06-23

## R Markdown

library(readxl)  
library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.1 ──

## ✔ ggplot2 3.3.6 ✔ purrr 0.3.4  
## ✔ tibble 3.1.7 ✔ dplyr 1.0.9  
## ✔ tidyr 1.2.0 ✔ stringr 1.4.0  
## ✔ readr 2.1.2 ✔ forcats 0.5.1

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

library(REAT)

##   
## Attaching package: 'REAT'

## The following object is masked from 'package:readr':  
##   
## spec

data <- read\_csv("~/Desktop/theil.csv")

## Rows: 67655 Columns: 8

## ── Column specification ────────────────────────────────────────────────────────  
## Delimiter: ","  
## chr (4): port, state, sp\_name, port\_state  
## dbl (4): year, landed\_lbs, live\_lbs, value  
##   
## ℹ Use `spec()` to retrieve the full column specification for this data.  
## ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

dummy <- data.frame(  
 port = letters[1:6],  
 value = c(10, 10, 10, 20, 500, 7000))  
  
#REAT package  
dummy\_reat <- theil(dummy$value)  
  
#using own equation  
dummy <- mutate(dummy, share = (value / sum(dummy$value)),  
 dummy, share\_mean = share \* n\_distinct(port),  
 t\_value = log(share\_mean)\*share)  
   
dummy\_new <- dummy %>%  
 summarise(sum(t\_value))  
  
#very different values, both over 1?

theil\_2019 <- data %>%  
 filter(sp\_name == "COD, ATLANTIC", year == 2019)   
  
#from REAT  
reat2019 = theil(theil\_2019$value)  
  
  
theil\_2019 <- mutate(theil\_2019,   
 #prop of total landings  
 share = value / sum(value),  
 #prop of mean landings  
 share\_mean = share \* n\_distinct(port\_state),   
 t\_value = log(share\_mean)\*share)  
   
new2019 = summarise(theil\_2019, sum(t\_value))