Spatial clustering luego de venta

# Cargamos librerías

suppressMessages(library(tidyverse))  
library(knitr)  
library(kableExtra)

##   
## Attaching package: 'kableExtra'

## The following object is masked from 'package:dplyr':  
##   
## group\_rows

library(forcats)

# A nivel bandera

## Cargamos archivos

grifos\_sc\_pre <- readRDS(here::here("data","processed","grifos\_con\_sc\_pre\_venta.RDS"))  
grifos\_sc\_post <- readRDS(here::here("data","processed","grifos\_con\_sc\_post\_venta.RDS"))

Hacemos merge de ambas

grifos\_sc <- grifos\_sc\_post %>%  
 rename(sc\_post = sc)  
grifos\_sc$sc\_pre <- grifos\_sc\_pre$sc  
  
grifos\_sc %>%   
 mutate(diff\_sc = sc\_post - sc\_pre) %>%  
 group\_by(bandera) %>%  
 summarise(mean\_diff\_sc = mean(diff\_sc),  
 n = n())

## # A tibble: 4 x 3  
## bandera mean\_diff\_sc n  
## <fct> <dbl> <int>  
## 1 PRIMAX 1.66 169  
## 2 INDEPENDIENTE 0.638 97  
## 3 PETROPERU 0.680 42  
## 4 REPSOL 1.74 124

grifos\_sc %>%  
 filter(stringr::str\_detect(razon\_social, "PERUANA DE ESTACIONES"))

## # A tibble: 28 x 31  
## codificador codigo\_de\_osine~ razon\_social ruc direccion distrito n  
## <chr> <dbl> <chr> <chr> <chr> <chr> <dbl>  
## 1 DU 17944 PERUANA DE ~ 2033~ AV. ARIC~ BREÑA 1764  
## 2 DU 14645 PERUANA DE ~ 2033~ AV. CUBA~ JESUS M~ 1706  
## 3 DUC 6771 PERUANA DE ~ 2033~ JR. ARIA~ MIRAFLO~ 992  
## 4 MC 7170 PERUANA DE ~ 2033~ AV. COLO~ LIMA 2549  
## 5 MC 18830 PERUANA DE ~ 2033~ AV. FAUS~ MAGDALE~ 2504  
## 6 MC 19989 PERUANA DE ~ 2033~ AV. TOMA~ SANTIAG~ 2488  
## 7 MC 8538 PERUANA DE ~ 2033~ JR. SAL~ MAGDALE~ 2391  
## 8 DU 8258 PERUANA DE ~ 2033~ AV. AREQ~ LINCE 2766  
## 9 DU 7310 PERUANA DE ~ 2033~ AV. BOLI~ PUEBLO ~ 2578  
## 10 MC 6873 PERUANA DE ~ 2033~ AV. SAN ~ SANTIAG~ 1995  
## # ... with 18 more rows, and 24 more variables: orden\_original <dbl>,  
## # bandera <fct>, glp <chr>, gnv <chr>, llanteria <chr>,  
## # islas\_comb\_liq <chr>, mecanico <chr>, aceite <chr>, lavado <chr>,  
## # tienda <chr>, cajero <chr>, en\_avenida\_principal <chr>, lat <dbl>,  
## # lon <dbl>, revisar <chr>, tipo <chr>, con\_glp <dbl>, con\_gnv <dbl>,  
## # grifo\_mas\_cercano <chr>, distancia\_min <dbl>, distancia\_avg <dbl>,  
## # num\_grifos\_cerc <int>, sc\_post <dbl>, sc\_pre <dbl>

# A nivel razon social

## Cargamos archivos

grifos\_sc\_pre <- readRDS(here::here("data","processed","grifos\_con\_sc\_pre\_venta\_razon.RDS"))  
grifos\_sc\_post <- readRDS(here::here("data","processed","grifos\_con\_sc\_post\_venta\_razon.RDS"))  
  
grifos\_sc\_pre %>%  
 count(razon\_social, sort = TRUE)

## # A tibble: 255 x 2  
## razon\_social n  
## <chr> <int>  
## 1 REPSOL COMERCIAL S.A.C. 71  
## 2 COESTI S.A. 41  
## 3 PERUANA DE ESTACIONES DE SERVICIOS S.A.C. 28  
## 4 ASESORIA COMERCIAL S.A. 11  
## 5 GRIFO SAN IGNACIO S.A.C. 6  
## 6 GRIFOS ESPINOZA S.A. 5  
## 7 ABA SINGER & CIA. S.A.C. 3  
## 8 ENERGIGAS S.A.C. 3  
## 9 GAZEL PERU S.A.C. 3  
## 10 ABA SINGER & CIA. S.A.C. 2  
## # ... with 245 more rows

Hacemos merge de ambas

# grifos\_sc <- grifos\_sc\_post %>%  
# rename(sc\_post = sc)  
# grifos\_sc$sc\_pre <- grifos\_sc\_pre$sc  
  
grifos\_sc <- grifos\_sc\_pre %>%  
 rename(sc\_pre = sc)  
grifos\_sc$sc\_post <- grifos\_sc\_post$sc  
  
grifos\_sc\_1 <- grifos\_sc %>%   
 mutate(diff\_sc = sc\_post - sc\_pre,  
 razon\_social\_sum = if\_else(razon\_social %in% c("REPSOL COMERCIAL S.A.C.",   
 "COESTI S.A.",   
 "PERUANA DE ESTACIONES DE SERVICIOS S.A.C."),  
 razon\_social,  
 "OTRA RAZÓN"))  
  
grifos\_sc\_1 %>%   
 group\_by(razon\_social\_sum) %>%  
 summarise(mean\_diff\_sc = mean(diff\_sc),  
 sd = sd(diff\_sc),  
 n = n())

## # A tibble: 4 x 4  
## razon\_social\_sum mean\_diff\_sc sd n  
## <chr> <dbl> <dbl> <int>  
## 1 COESTI S.A. 0.155 0.243 41  
## 2 OTRA RAZÓN 0.0415 0.118 292  
## 3 PERUANA DE ESTACIONES DE SERVICIOS S.A.C. 0.200 0.215 28  
## 4 REPSOL COMERCIAL S.A.C. 0.114 0.209 71

df1 <- grifos\_sc\_1 %>%   
 filter(diff\_sc != 0) %>%  
 group\_by(razon\_social\_sum) %>%  
 summarise(Promedio= mean(diff\_sc),  
 SD = sd(diff\_sc),  
 "Mínimo" = min(diff\_sc),  
 "Máximo" = max(diff\_sc),  
 "Numero de Obs." = n()) %>%  
 mutate(razon\_social\_sum= fct\_relevel(razon\_social\_sum,   
 "PERUANA DE ESTACIONES DE SERVICIOS S.A.C.",  
 "COESTI S.A.",   
 "REPSOL COMERCIAL S.A.C.",  
 "OTRA RAZÓN",  
 "Todas las estaciones")) %>%  
 arrange(razon\_social\_sum) %>%  
 select(1, 6, 2, 3, 4, 5)

## Warning: Unknown levels in `f`: Todas las estaciones

df2 <- grifos\_sc\_1 %>%   
 filter(diff\_sc != 0) %>%  
 mutate(razon\_social\_sum = "TODAS LAS ESTACIONES") %>%  
 group\_by(razon\_social\_sum) %>%  
 summarise(Promedio= mean(diff\_sc),  
 SD = sd(diff\_sc),  
 "Mínimo" = min(diff\_sc),  
 "Máximo" = max(diff\_sc),  
 "Numero de Obs." = n()) %>%  
 select(1, 6, 2, 3, 4, 5)  
  
x <- bind\_rows(df1, df2) %>%  
 rename("Razón Social" = razon\_social\_sum) %>%  
 kable(format = "pandoc",   
 digits = c(0, 0, 3, 3, 3, 3)) %>%  
 kable\_styling(full\_width = FALSE, position = "left")

## Warning in bind\_rows\_(x, .id): binding factor and character vector,  
## coercing into character vector

## Warning in bind\_rows\_(x, .id): binding character and factor vector,  
## coercing into character vector

## Warning in kable\_styling(., full\_width = FALSE, position = "left"): Please  
## specify format in kable. kableExtra can customize either HTML or LaTeX  
## outputs. See https://haozhu233.github.io/kableExtra/ for details.

x

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Razón Social | Numero de Obs. | Promedio | SD | Mínimo | Máximo |
| PERUANA DE ESTACIONES DE SERVICIOS S.A.C. | 22 | 0.254 | 0.212 | 0.024 | 0.850 |
| COESTI S.A. | 24 | 0.265 | 0.269 | 0.024 | 1.250 |
| REPSOL COMERCIAL S.A.C. | 29 | 0.279 | 0.249 | 0.014 | 0.861 |
| OTRA RAZÓN | 65 | 0.187 | 0.190 | 0.010 | 1.033 |
| TODAS LAS ESTACIONES | 140 | 0.230 | 0.223 | 0.010 | 1.250 |

pack\_rows(x, "Grupo de tratamiento", 1,5)

## Warning in pack\_rows(x, "Grupo de tratamiento", 1, 5): Please specify  
## format in kable. kableExtra can customize either HTML or LaTeX outputs. See  
## https://haozhu233.github.io/kableExtra/ for details.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Razón Social | Numero de Obs. | Promedio | SD | Mínimo | Máximo |
| PERUANA DE ESTACIONES DE SERVICIOS S.A.C. | 22 | 0.254 | 0.212 | 0.024 | 0.850 |
| COESTI S.A. | 24 | 0.265 | 0.269 | 0.024 | 1.250 |
| REPSOL COMERCIAL S.A.C. | 29 | 0.279 | 0.249 | 0.014 | 0.861 |
| OTRA RAZÓN | 65 | 0.187 | 0.190 | 0.010 | 1.033 |
| TODAS LAS ESTACIONES | 140 | 0.230 | 0.223 | 0.010 | 1.250 |