

# Arbolado Montevideo

Manuel Toledo, Lucas Pescetto y Joaquín Silva

2023-06-28

```
precios <- readRDS(here("precios_actualizada.RDS"))
establecimientos <- read_delim("datos/establecimiento.csv", delim = ";", locale = readr::locale(encoding = "latin1"))

## Rows: 766 Columns: 15
## -- Column specification -----
## Delimiter: ";"
## chr (9): razon.social, nombre.sucursal, direccion, barrio, cadena, long, lat...
## dbl (6): id.establecimientos, ccz, cajas, id.depto, localidad, superficie (m2)
##
## 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.
# Los datos de lat y long del san roque del aeropuerto están mal, así que hay que cambiarlos
establecimientos <- filter(establecimientos, id.establecimientos != 679) |> mutate(lat = as.numeric(sub(',', ' ', lat)))

## Warning: There were 2 warnings in `mutate()`.
## The first warning was:
## i In argument: `lat = as.numeric(sub(",", ".", lat, fixed = T))`.
## Caused by warning:
## ! NAs introducidos por coerción
## i Run `dplyr::last_dplyr_warnings()` to see the 1 remaining warning.

establecimientosCord <- filter(establecimientos, !(is.na(lat) | is.na(long))) %>% filter((long < (-30) & long > (-50) & lat > 34 & lat < 36))
establecimientosCord <- st_as_sf(establecimientosCord, coords = c("lat", "long")) %>% st_set_crs(4326)
df_depto <- st_as_sf(vect(here("Mapas", "ine_depto.shp"))) %>% st_set_crs(5382) %>% st_transform(4326)
df_mvdeo <- st_as_sf(vect(here("Mapas", "ine_ccz_mvd.shp")))
productos <- read_csv2("datos/productos.csv", locale = readr::locale(encoding = "latin1")) %>% filter(id_producto != 1)

## i Using ",", "." as decimal and "'.'" as grouping mark. Use `read_delim()` for more control.
## Rows: 363 Columns: 5-- Column specification -----
## Delimiter: ";"
## chr (4): producto, marca, especificacion, nombre
## dbl (1): id_producto
## 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.
```

## Introducción

Este trabajo busca analizar las diferencias de precios a lo largo del tiempo, en distintos lugares dentro del país y en establecimientos dentro de Montevideo.

Como objetivo adicional se propone ayudar a los consumidores a tomar mejores decisiones a la hora de comprar???? ?0

## Datos

Contamos con los datos del índice de precios del consumidor del INE (cita) a partir del año 2016 y hasta marzo del 2023. Estos contienen 363 productos dentro 766 establecimientos, aunque de entre los productos solo se usaran los 18 productos dentro de la canasta básica más 2 productos que se agregaron por decisión personal. Los productos son:

- Aceite de girasol 900 cc
- Aguja vacuna 1 kg (con y sin hueso)
- Arroz blanco 1 kg
- Arvejas en conserva 300 g
- Azúcar blanco 1 kg
- Carne picada vacuna 1 kg
- Cocoa 500 g
- Dulce de leche 1 kg
- Fideos secos al huevo 500 g
- Galletitas al agua 140 g
- Harina trigo común 0000 1 kg
- Huevos colorados 1/2 docena
- Manteca 200 g
- Pan flauta 215 g
- Papel higiénico hoja simple 4 rollos 30 mts
- Pollo entero fresco con menudos 1 kg
- Pulpa de tomate 1 L
- Sal fina yodada fluorada 500 g
- Yerba mate común 1 kg
- Café (agregado)
- Fideos secos de sémola 500 g (agregado)

Para cada producto se cuenta con una descripción y una marca asociada. A su vez se cuenta con datos de ubicación, tamaño, nombre y cadena de cada establecimiento.

Estos datos se asocian a cada relevación de precios hecha, por producto y dentro de cada establecimiento, así como también se le asocia la fecha en que fue registrado.

## Análisis exploratorio

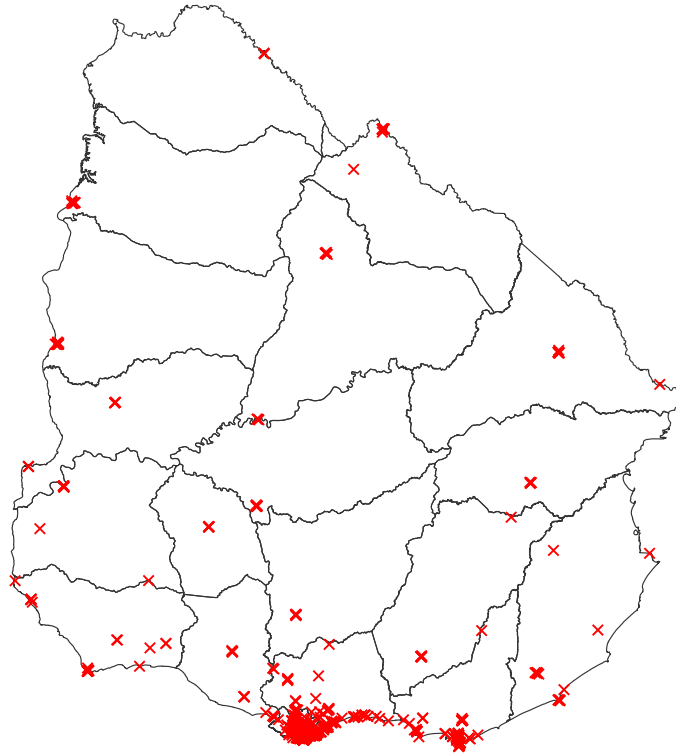
HAY QUE AGREGAR LOS PRODUCTOS 359,361 Y 365 Y QUITAR 74 Y 75

Los productos presentan

```
a <- group_by(productos, producto) %>% mutate(marcas=n()) %>% ungroup()
```

```
ggplot() +  
  geom_sf(data=df_depto, fill=NA, color="gray20") +  
  geom_sf(data=establecimientosCord, pch=4, color="red") +  
  theme_void() +  
  labs(title="Establecimientos") +  
  theme(plot.title=element_text(hjust=1/2))
```

## Establecimientos



## Modelo estadístico ¿?

Posible herramienta para hacer un modelo con bootstrap? <https://towardsdatascience.com/multi-step-time-series-forecasting-with-xgboost-65d6820bec39>

**XGBoost**

## Descripción de la aplicación

### Cosas

mostrar la evolución de los precios ajustando por inflación

Mapa con los supermercados

Hacer una canasta y mostrar el mejor local

Aceite de girasol 900 cc = 1:3,357,359,363 Aguja vacuna 1 Kg = 13 (con hueso) 14 (sin hueso) Arroz blanco 1 Kg = 15:20 Arvejas en conserva 300 grs = 21,23 (300g) y 22 (380g) Azúcar blanco 1 Kg = 24:25 Carne picada vacuna 1 Kg = 29:30 Cocoa 500 grs = 40:41 Dulce de leche 1 Kg = 48:50 Fideos secos al huevo 500 grs = 52:54 Galletitas al agua 140 grs = 61 (140g) 62(120g) Harina trigo común 0000 1 kg = 76:78 Huevos colorados 1/2 docena = 85:87 Manteca 200 grs = 102:104 Pan flauta 215 grs = 121 Papel higiénico hoja simple 4 rollos 30 mts = 122:124 Pollo entero fresco con menudos 1 kg = 130:132 Pulpa de tomate 1 Lt = 133:135 Sal fina yodada fluorada 500 grs = 140:142 Yerba mate común 1 Kg = 149:151

EXTRA: café = 26:28 fideos de sémola = 55:57