



Ciencia de Datos Aplicada al Transporte

Curso de Complementación

## **Trabajo Práctico N° 2: Sistemas de Información Geográfica**

Alumno

---

Mateo Mastelli

## CONSIGNAS

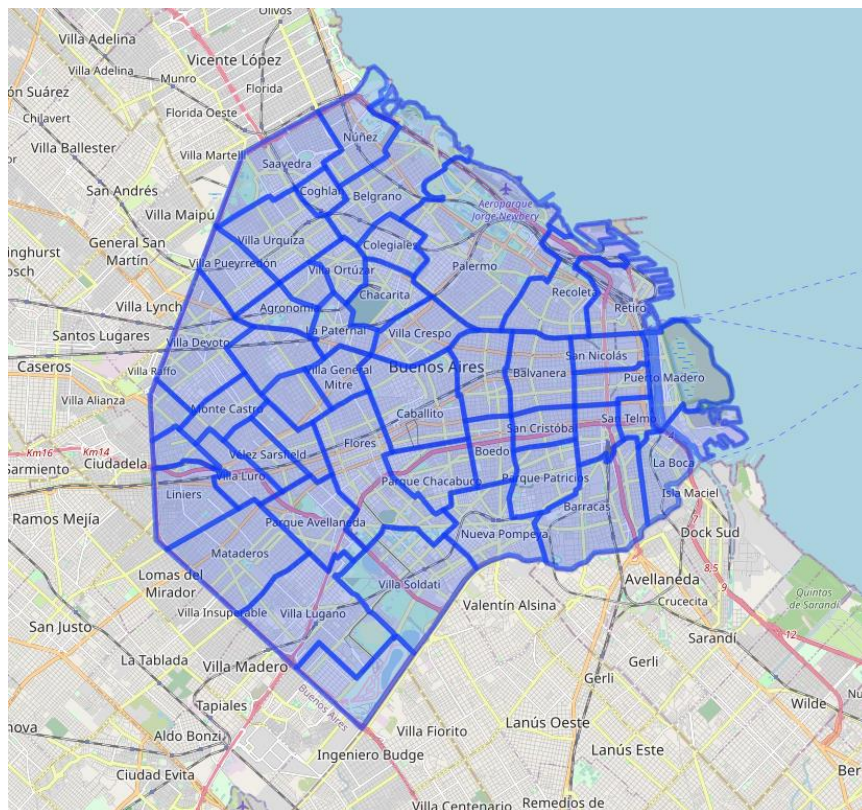
1. Descargar el dataset oficial de los barrios de la Ciudad Autónoma de Buenos Aires y cargarlo en el entorno de R.
2. Tome de la API de Transporte los colectivos que estén circulando en la Ciudad al momento que esté realizando la consigna. Si quiere, puede crear su propia cuenta desde [AQUÍ](#) o puede tomar el link con las credenciales que utilizamos en clase. Si lo hace de la primera forma, le recuerdo que el dataset que utilizamos es el de VehiclePositionsSimple. Recuerde que es conveniente convertir el dataset a un objeto SF antes de continuar.
3. Realice un informe en el cual estén presentes aspectos descriptivos fundamentales como un top de los barrios con más unidades operando, los barrios con mayor densidad de colectivos ajustada por área, aquellos barrios que poseen promedio de velocidad más alta y otras cuestiones que usted considere pertinente. Es necesario que haya por lo menos un gráfico o mapa (de lo que quiera) y la extensión será de máximo una carilla sin contar el o los gráficos. En el informe también se deben responder las siguientes preguntas:
  - a. ¿En qué barrio hay más unidades funcionando de la línea 166?
  - b. ¿Cuál es el Metrobús que más colectivos por metro posee en ese instante? Para esto explicite el buffer que utilizará para intersectar geometrías.
  - c. ¿Cuántos colectivos están operando en el entorno de la Línea A de Subterráneos?

**1) Descargar el dataset oficial de los barrios de la Ciudad Autónoma de Buenos Aires y cargarlo en el entorno de R.**

```
library(sf)
library(tidyverse) #dplyr
library(leaflet)
library(plotly)
library(jsonify)
library(dplyr)

#1)
#Levanto el geojson con la informacion geografica de los barrios de CABA
barrios <- st_read("barrios_Ciudad.geojson")

#Ploteo los barrios
barrios %>% leaflet() %>% addTiles() %>% addPolygons()
```



- 2) Tome de la API de Transporte los colectivos que estén circulando en la Ciudad al momento que esté realizando la consigna. Si quiere, puede crear su propia cuenta desde [AQUÍ](#) o puede tomar el link con las credenciales que utilizamos en clase. Si lo hace de la primera forma, le recuerdo que el dataset que utilizamos es el de VehiclePositionsSimple. Recuerde que es conveniente convertir el dataset a un objeto SF antes de continuar

#descargo la informacion de la API de Transporte y agregamos la columna de geometria

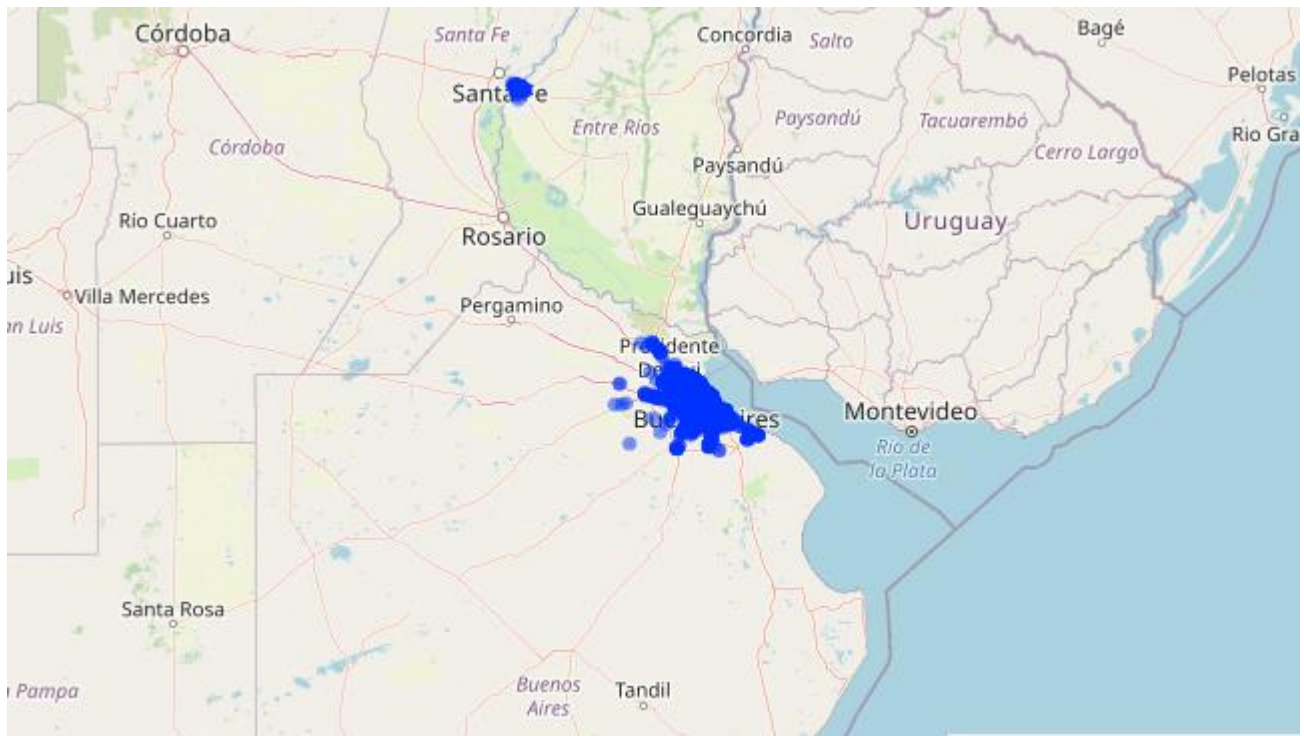
```
colectivos <-  
st_as_sf(from_json("https://apitransporte.buenosaires.gob.ar/colectivos/vehiclePositionsSimple?client_id=02bc7547f78543  
a9a217480ec619bb82&client_secret=A690E6d9390f406ea4f5246b2b0C62F1"), coords = c("longitude", "latitude"), crs =  
4326, remove = FALSE)
```

#paso la velocidad a kilometros por hora ya que viene en metros por segundos

```
colectivos <- colectivos %>% mutate(speed = speed * 3.6)
```

#Ploteo los colectivos en el mapa

```
colectivos %>% leaflet() %>% addTiles() %>% addCircles(label = ~route_short_name)
```

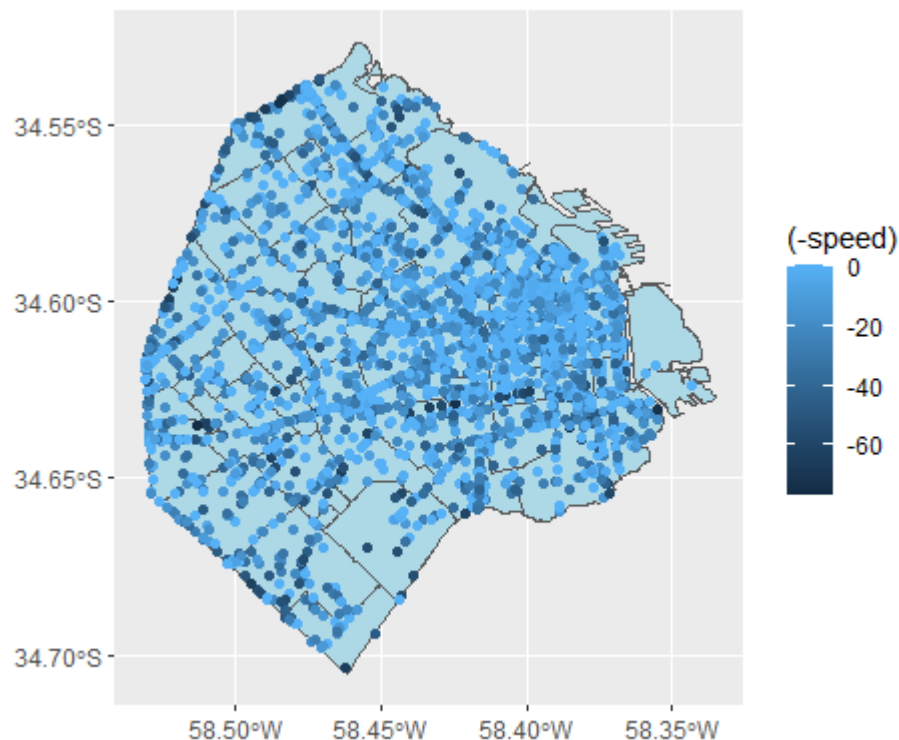


Se observa que la API de transportes no solo incluye los colectivos de Buenos Aires, si no que también brinda información de los colectivos de Paraná.

3) Realice un informe en el cual estén presentes aspectos descriptivos fundamentales como un top de los barrios con más unidades operando, los barrios con mayor densidad de colectivos ajustada por área, aquellos barrios que poseen promedio de velocidad más alta y otras cuestiones que usted considere pertinente. Es necesario que haya por lo menos un gráfico o mapa (de lo que quiera) y la extensión será de máximo una carilla sin contar el o los gráficos. En el informe también se deben responder las siguientes preguntas:

- ¿En qué barrio hay más unidades funcionando de la línea 166?
- ¿Cuál es el Metrobús que más colectivos por metro posee en ese instante? Para esto explícite el buffer que utilizará para intersectar geometrías.
- ¿Cuántos colectivos están operando en el entorno de la Línea A de Subterráneos?

A continuación, se grafican todos los colectivos que se encuentran en la Ciudad de Buenos Aires en función de su posición y velocidad (km/h):



Del gráfico se puede observar que la zona de Capital Federal más cargada de colectivos en el momento en que se consultó a la API corresponde al microcentro y sus zonas aledañas. Así mismo también se observa que en dichas zonas se presentan las velocidades más bajas en los colectivos, lo cual guarda sentido ya que una alta densidad de colectivos implica una congestión más elevada y por ende la imposibilidad de alcanzar altas velocidades.



- ❖ Para averiguar cuales son los barrios de CABA con mayor unidades operando realizo un Join entre el Dataset de los colectivos y el de los barrios:

```
#jineo los dos datasets de manera de asignarle a cada colectivo el barrio en el que se encontraba
#en el momento en que consulté a la API (descarto los colectivos que no se encontraban en CABA)
colectivos_CABA <- colectivos %>% st_join(barrios) %>% st_as_sf() %>% drop_na()
#top de los barrios con más unidades operando
colectivos_CABA_by_barrios <- colectivos_CABA %>% group_by(BARRIO) %>% summarise(densidad_bondis = n()) %>%
as.data.frame() %>% arrange(desc(densidad_bondis))
top_n(colectivos_CABA_by_barrios, 5, densidad_bondis)
```

Finalmente se obtiene el siguiente top 5 de barrios con mayores unidades operado:

	BARRIO	densidad_bondis
1	PALERMO	305
2	BALVANERA	249
3	RECOLETA	242
4	BELGRANO	170
5	CABALLITO	159

- ❖ Barrios con mayor densidad de colectivos ajustada por área.

Primero convierto el área de m2 a km2 y luego genero una nueva columna que represente la cantidad de bondis por kilometro cuadrado.

```
#los barrios con mayor densidad de colectivos ajustada por área
#asumo que el area de los barrios esta en m2 y la paso a km2
colectivos_CABA_by_barrios <- colectivos_CABA %>% group_by(BARRIO) %>% summarise(densidad_bondis = n(), AREA =
mean(AREA)/1000000) %>% as.data.frame()
colectivos_CABA_by_barrios$Bondis_x_Area = colectivos_CABA_by_barrios$densidad_bondis /
colectivos_CABA_by_barrios$AREA
top_n(colectivos_CABA_by_barrios %>% as.data.frame() %>% arrange(desc(Bondis_x_Area)) %>% drop_na, 5, Bondis_x_Area)
```

Top 5 Barrios con mayor cantidad de colectivos por kilometro cuadrado:

	BARRIO	densidad_bondis	AREA	geometry	Bondis_x_Area
1	CONSTITUCION	146	2.101995	MULTIPOINT ((-58.39031 -34....	69.45782
2	SAN NICOLAS	133	2.289008	MULTIPOINT ((-58.36813 -34....	58.10378
3	BALVANERA	249	4.342280	MULTIPOINT ((-58.4119 -34.5...	57.34314
4	MONSERRAT	119	2.198622	MULTIPOINT ((-58.36919 -34....	54.12481
5	SAN TELMO	50	1.232270	MULTIPOINT ((-58.37012 -34....	40.57552

- ❖ Barrios que poseen promedio de velocidad más alta

```
#barrios que poseen promedio de velocidad más alta
colectivos_CABA_by_barrios_vel <- colectivos_CABA %>% group_by(BARRIO) %>% summarise(vel_prom = mean(speed)) %>%
as.data.frame() %>% arrange(desc(vel_prom)) %>% drop_na()
top_n(colectivos_CABA_by_barrios_vel, 5, vel_prom)
```

Top 5 Barrios que poseen promedio de velocidad más alta:

	BARRIO	vel_prom	geometry
1	VILLA LUGANO	23.54286	MULTIPOINT ((-58.49092 -34....
2	PARQUE AVELLANEDA	21.20370	MULTIPOINT ((-58.48932 -34....
3	SAAVEDRA	20.85915	MULTIPOINT ((-58.48566 -34....
4	VILLA LURO	19.88235	MULTIPOINT ((-58.51266 -34....
5	VILLA SANTA RITA	19.16667	MULTIPOINT ((-58.4866 -34.6... )

**a. ¿En qué barrio hay más unidades funcionando de la línea 166?**

#¿En qué barrio hay más unidades funcionando de la línea 166?

```
colectivos_CABA_by_barrios_166 <- colectivos_CABA[grepl("166", colectivos_CABA$route_short_name),] %>% group_by(BARRIO)
%>% summarise(cant_166 = n()) %>% as.data.frame() %>% arrange(desc(cant_166)) %>% drop_na()
```

```
colectivos_CABA_by_barrios_166
```

```
top_n(colectivos_CABA_by_barrios_166, 1, cant_166)$BARRIO
```

	BARRIO	cant_166	geometry
1	PALERMO	8	MULTIPOINT ((-58.43619 -34....
2	VILLA CRESPO	6	MULTIPOINT ((-58.43915 -34....
3	VILLA GRAL. MITRE	3	MULTIPOINT ((-58.46826 -34....
4	VILLA LURO	2	MULTIPOINT ((-58.50977 -34....
5	VILLA SANTA RITA	2	MULTIPOINT ((-58.48078 -34....
6	FLORESTA	1	POINT (-58.48677 -34.62546)
7	LINIERS	1	POINT (-58.52355 -34.63434)
8	VELEZ SARSFIELD	1	POINT (-58.49471 -34.62943)
9	VERSALLES	1	POINT (-58.52213 -34.63382)

Por ende, el barrio en cual había mas colectivos de la línea 166 cuando consulte a la API es Palermo con un total de 8 colectivos.

**b. ¿Cuál es el Metrobús que más colectivos por metro posee en ese instante? Para esto explicité el buffer que utilizaré para intersectar geometrías.**

#levantamos el geojason con la informacion geografica de metrobuses

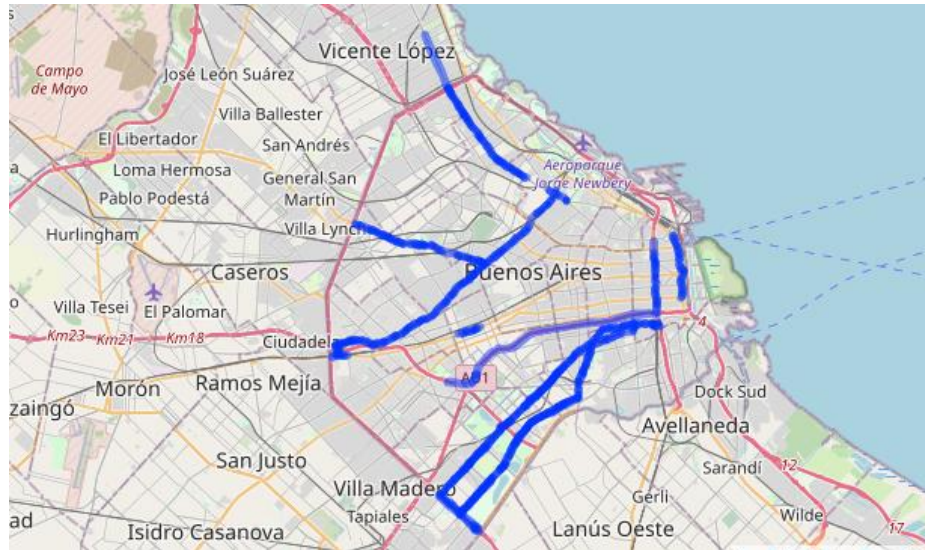
```
metrobus <- st_read("recorrido-de-metrobus.geojson")
```

#buffer. Realizo un Buffer de 10 metros sobre las líneas que describen a los metrobuses

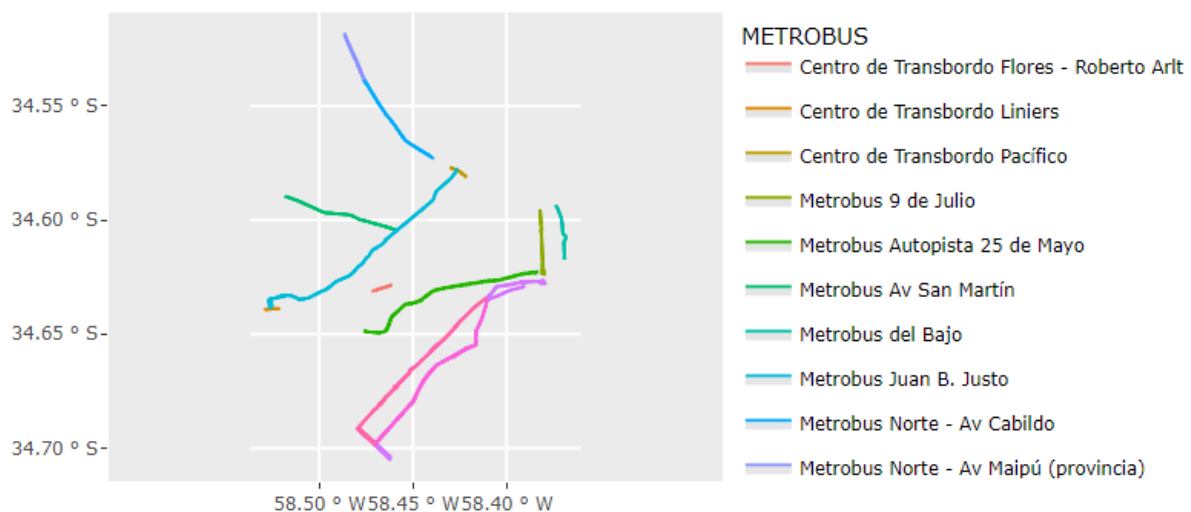
```
metrobus_buff = st_buffer(metrobus, dist = 10)
```

# Grafico los Metrobuses

```
metrobus_buff %>% leaflet() %>% addTiles() %>% addPolygons()
```



```
esquema_metrobus <- ggplot(metrobus_buff) +  
  geom_sf(aes(color = METROBUS))  
ggplotly(esquema_metrobus)
```





#juego con colectivos y grupo por metrobus

```
colectivos_en_metro_by_metro <- colectivos %>% st_join(metrobus_buff) %>% group_by(METROBUS) %>%
summarise(densidad_bondis = n()) %>% arrange(desc(densidad_bondis)) %>% as.data.frame() %>% drop_na()
```

```
colectivos_en_metro_by_metro
```

#calcula longitud de metrobus

```
long_metro <- metrobus %>% filter(COD_SENT >= 0) %>% group_by(METROBUS) %>% summarise(long_metrobus = sum(LONG))
%>% as.data.frame()
```

```
colectivos_en_metro_by_metro_2 <- merge(x=colectivos_en_metro_by_metro, y=long_metro, by="METROBUS")
```

#Creo una nueva columna con los colectivos x metro x Metrobus

```
colectivos_en_metro_by_metro_2$Colectivosxmetro = colectivos_en_metro_by_metro_2$densidad_bondis /
colectivos_en_metro_by_metro_2$long_metrobus
```

```
colectivos_en_metro_by_metro_2 <- colectivos_en_metro_by_metro_2 %>% arrange(desc(Colectivosxmetro))
```

```
dropList <- c("geometry.x", "geometry.y")
```

```
colectivos_en_metro_by_metro_2 <- colectivos_en_metro_by_metro_2[, !colnames(colectivos_en_metro_by_metro_2) %in%
dropList]
```

```
colectivos_en_metro_by_metro_2
```

	METROBUS	densidad_bondis	long_metrobus	Colectivosxmetro
1	Centro de Transbordo Pacífico	31	208.31	0.148816667
2	Metrobus Norte - Av Cabildo	140	5443.54	0.025718558
3	Centro de Transbordo Liniers	16	698.09	0.022919681
4	Metrobus del Bajo	64	2793.64	0.022909179
5	Centro de Transbordo Flores - Roberto Arlt	17	964.09	0.017633209
6	Metrobus 9 de Julio	45	3332.11	0.013504956
7	Metrobus Norte - Av Maipú (provincia)	29	2379.53	0.012187281
8	Metrobus Sur - Ramal Coronel Roca	63	8191.35	0.007691040
9	Metrobus Sur	47	6347.03	0.007405038
10	Metrobus Juan B. Justo	67	12177.80	0.005501815
11	Metrobus Av San Martín	27	5777.19	0.004673552
12	Metrobus Sur - Ramal Fernández de la Cruz	24	10144.06	0.002365917
13	Metrobus Autopista 25 de Mayo	19	11282.33	0.001684049

> |

Por lo tanto, el Metrobus con más colectivos por metro al instante de consultar la API es el "Centro de Transbordo Pacífico".

**b. ¿ Cuántos colectivos están operando en el entorno de la Línea A de Subterráneos?**

#levantamos el geojason con la informacion geografica de las lineas de subte

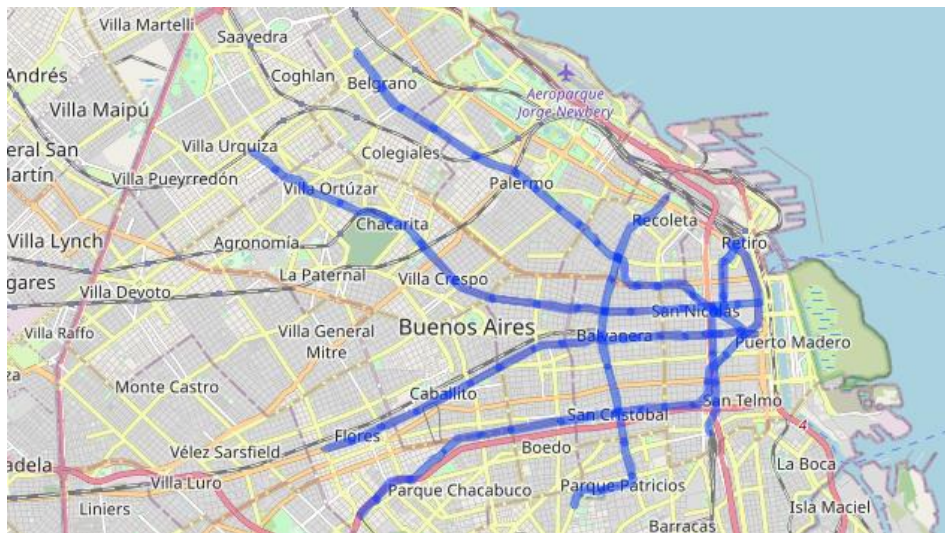
```
subtes <- st_read("subte_lineas.geojason")
```

#buffer

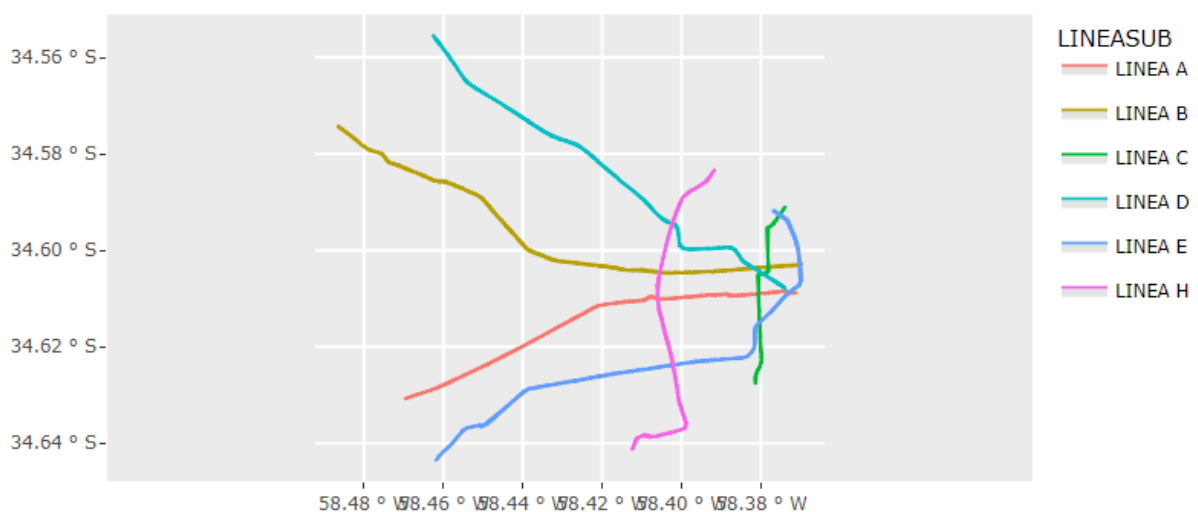
```
subtes_buff = st_buffer(subtes, dist = 10)
```

#Grafico los Metrobuses

```
subtes_buff %>% leaflet() %>% addTiles() %>% addPolygons()
```



```
esquema_subtes <- ggplot(subtes_buff) +  
  geom_sf(aes(color = LINEASUB))  
ggplotly(esquema_subtes)
```



#Me quedo solo con lo referente a la linea A de subte y joineo con los colectivos

```
colectivos_entorno_subte_A <- colectivos %>% st_join(subtes_buff) %>% filter(LINEASUB == "LINEA A") %>% group_by(LINEASUB)
%>% summarise(densidad_bondis = n()) %>% arrange(desc(densidad_bondis)) %>% as.data.frame() %>% drop_na()
```

colectivos\_entorno\_subte\_A

```
  LINEASUB densidad_bondis geometry
1  LINEA A           113 MULTIPOINT ((-58.46395 -34....
  |
```

Por lo tanto, en torno a la línea A de subte se encuentran 113 colectivos al momento de consultar la API.