

# Tarea Final

Arriola, Miquelerena, Rovetta

2023-07-07

## Introducción

Datos de trafico

## Datos

### Descripción general de los datos

Los datos utilizados son estos

- Conteo vehicular en las principales avenidas de Montevideo
- Velocidad promedio vehicular en las principales avenidas de Montevideo
- Ubicación de sensores de medición de conteo vehículos
- Semaforos

Todos provenientes de la Intendencia de Montevideo

### Observaciones

Los datos son mensuales. Van desde Enero 2021 hasta Abril 2023. Los datos de los semaforos son de septiembre del 2014

### Descripcion de variables

#### Conjunto de datos: Conteo vehicular en las principales avenidas de Montevideo

- **cod\_detector:** Entero - ID de la cámara que monitorea un carril específico para detectar vehículos.
- **id\_carril:** Entero - Número del carril monitoreado (1, 2, 3, ...).
- **fecha:** AAAA-MM-DD - Día en que se realizó la medición.
- **hora:** hh:mm:ss - Hora en que se realizó la medición.
- **dsc\_avenida:** Texto - Nombre de la avenida donde se mide el tráfico.
- **dsc\_int\_anterior:** Texto - Nombre de la vía desde donde vienen los vehículos.
- **dsc\_int\_siguiente:** Texto - Nombre de la vía hacia donde se dirigen los vehículos.
- **latitud:** Float - Latitud del lugar de medición.
- **longitud:** Float - Longitud del lugar de medición.
- **volumen:** Entero - Cantidad de vehículos detectados en el carril en los últimos 5 minutos.
- **volumen\_hora:** Entero - Cantidad de vehículos detectados en el carril en la última hora.

## Conjunto de datos: Velocidad promedio vehicular en las principales avenidas de Montevideo

- **cod\_detector**: Entero - ID de la cámara que monitorea un carril específico para detectar vehículos.
- **id\_carril**: Entero - Número del carril monitoreado (1, 2, 3, ...).
- **fecha**: AAAA-MM-DD - Día en que se realizó la medición.
- **hora**: hh:mm:ss - Hora en que se realizó la medición.
- **dsc\_avenida**: Texto - Nombre de la avenida donde se mide el tráfico.
- **dsc\_int\_anterior**: Texto - Nombre de la vía desde donde vienen los vehículos.
- **dsc\_int\_siguiete**: Texto - Nombre de la vía hacia donde se dirigen los vehículos.
- **latitud**: Float - Latitud del lugar de medición.
- **longitud**: Float - Longitud del lugar de medición.
- **velocidad\_promedio**: Entero - Promedio de las velocidades de los autos que circularon por el carril durante los últimos 5 minutos.

## Conjunto de datos: Ubicación de sensores de medición de conteo vehículos

- **dsc\_avenida**: Texto - Nombre de la avenida donde se encuentra el sensor o cámara y donde se mide el tránsito.
- **dsc\_int\_anterior**: Texto - Nombre de la vía que forma el cruce desde donde vienen los vehículos.
- **dsc\_int\_siguiete**: Texto - Nombre de la vía que forma el cruce donde está el sensor. En general, el sensor se encuentra un poco antes de esta vía. El sentido de circulación será desde el cruce con **dsc\_int\_anterior** hacia el cruce con **dsc\_int\_siguiete**.
- **latitud**: Float - Coordenada que indica la latitud de la ubicación del sensor.
- **longitud**: Float - Coordenada que indica la longitud de la ubicación del sensor.

Sobre estos datos en particular, son *100 sensores* que se van cambiando de ubicacion mes a mes

## Base de Datos

Debido a que los datos utilizados, están estrechamente relacionados y a su vez son sumamente masivos, hemos decidido utilizar una base de datos relacionales.

En la base de datos la tabla principal es **fct\_registros**

### Tabla: fct\_registros

- Cantidad de datos: 85386695
- Variables de la tabla:
  - *id\_registros* (Primary Key)
  - *id\_carril*
  - *id\_fecha* -> *d\_date(id\_fecha)* (Foreign Key)
  - *id\_hora*
  - *id\_detector* -> *d\_sensores(id\_detector)* (Foreign Key)
  - *volume*
  - *volumen\_hora*
  - *velocidad*

### Tabla: d\_sensores

- Cantidad de datos: 273
- Variables de la tabla:

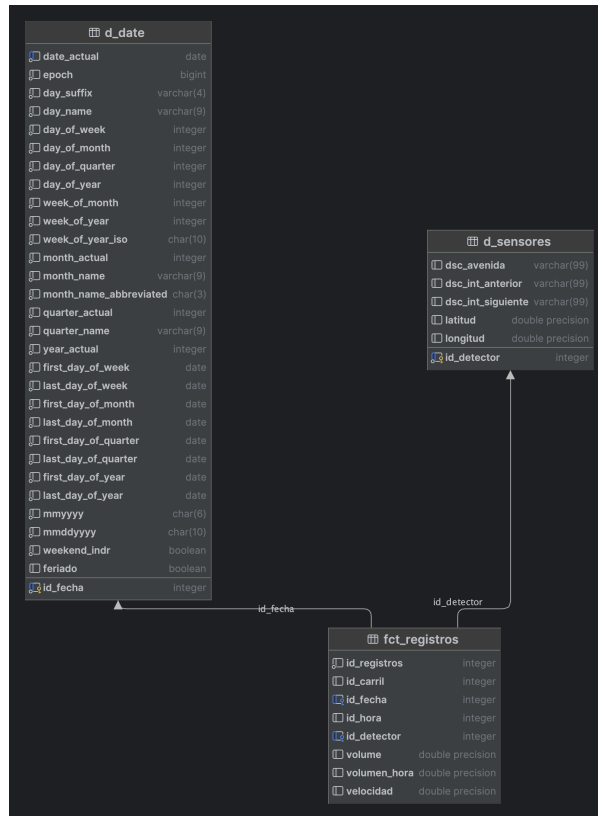


Figure 1: Diagrama de la base de datos

- *id\_detector* (Primary Key)
- *dsc\_avenida*
- *dsc\_int\_anterior*
- *dsc\_int\_siguiente*
- *latitud*
- *longitud*
- *barrio*

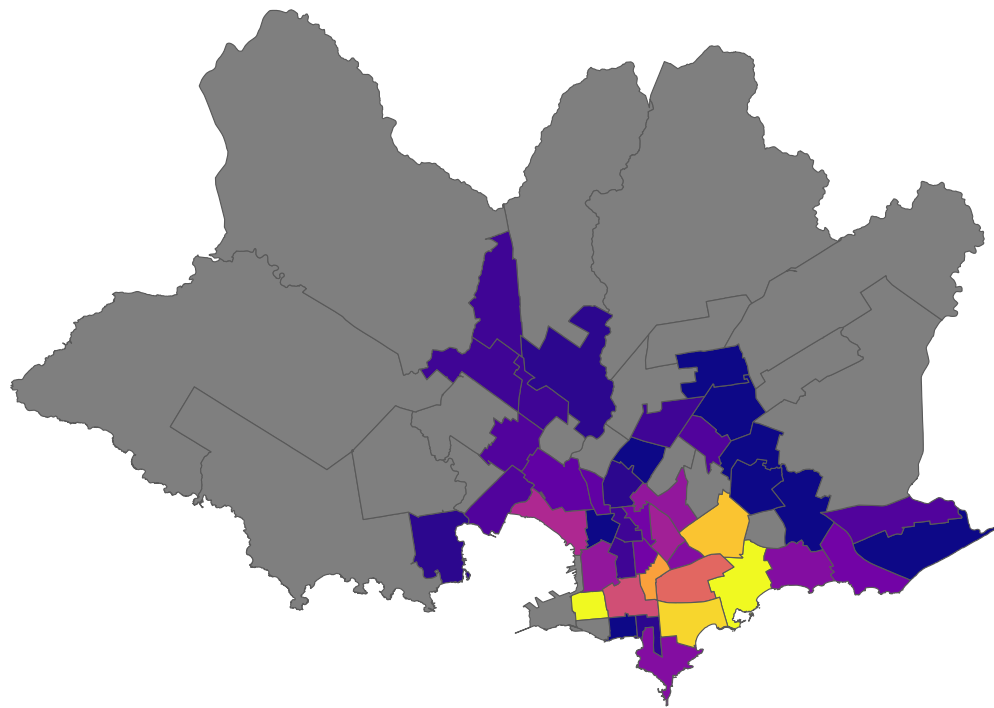
**Tabla: d\_date**

- Cantidad de datos: 3652
- Variables de la tabla:
  - *id\_fecha* (Primary Key)
  - *date\_actual*
  - *epoch*
  - *day\_suffix*
  - *day\_name*
  - *day\_of\_week*
  - *day\_of\_month*
  - *day\_of\_quarter*
  - *day\_of\_year*
  - *week\_of\_month*
  - *week\_of\_year*
  - *week\_of\_year\_iso*
  - *month\_actual*
  - *month\_name*
  - *month\_name\_abbreviated*
  - *quarter\_actual*
  - *quarter\_name*
  - *year\_actual*
  - *first\_day\_of\_week*
  - *last\_day\_of\_week*
  - *first\_day\_of\_month*
  - *last\_day\_of\_month*
  - *first\_day\_of\_quarter*
  - *last\_day\_of\_quarter*
  - *first\_day\_of\_year*
  - *last\_day\_of\_year*
  - *mmyyyy*
  - *mmddyyyy*
  - *weekend\_indr*
  - *feriado*

## Análisis exploratorio

### Mapa de los datos

Los datos de los sensores no son de todos los barrios , sino que solamente alcanzan los siguientes barrios  
 En el siguiente mapa se puede visualizar



Cantidad de sensores por barrio

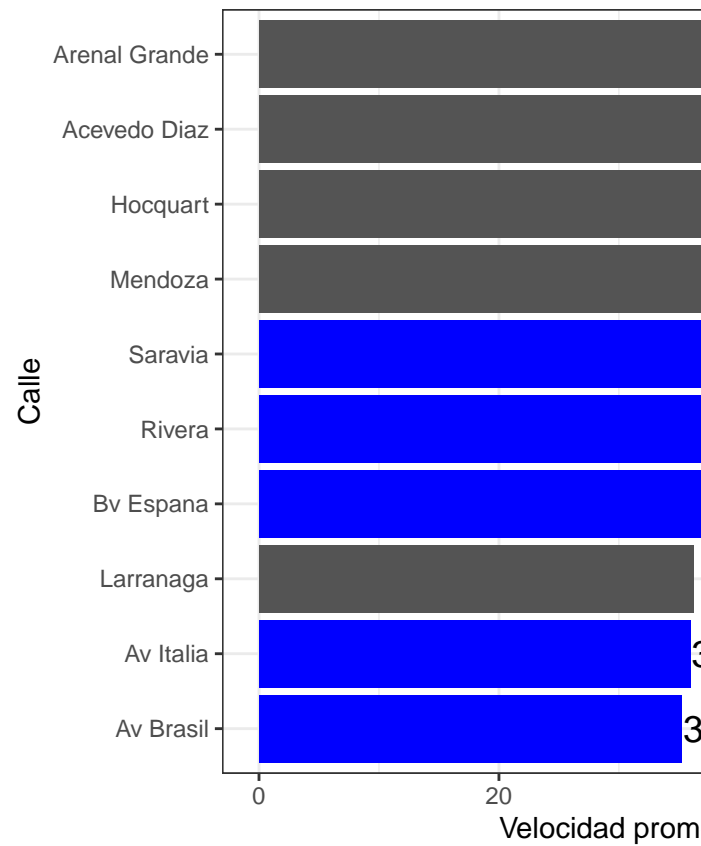
5 10 15 20

““

## Preguntas de investigación

1. ¿Existe alguna correlación entre el volumen de tráfico y la velocidad promedio en las avenidas de Montevideo?
2. ¿Cuáles son las calles con los mayores promedios de velocidad en Montevideo?
3. ¿Cómo va variando el volumen y velocidad medidos a través de la semana?

¿Cuáles son las calles con los mayores promedios de velocidad en Montevideo?  
 ¿Con que frecuencia se cometen excesos de velocidad?



Pasemos a investigar las calles con mayor promedio de velocidad.

El siguiente gráfico nos muestra que las 3 calles con mas velocidad en promedio superan el máximo de 45km/h siendo este la velocidad máxima de circulación reglamentaria. La avenida más rápida en promedio no alcanza el máximo de velocidad permitido

En promedio de velocidad circulacion los conductores son prudentes, aun asi vemos con que frecuencia se cometen excesos de velocidad.

¿Cómo va variando el volumen y velocidad medidos a traves de el tiempo?

Observemos el volumen y la velocidad promedio de todos los días de la semana

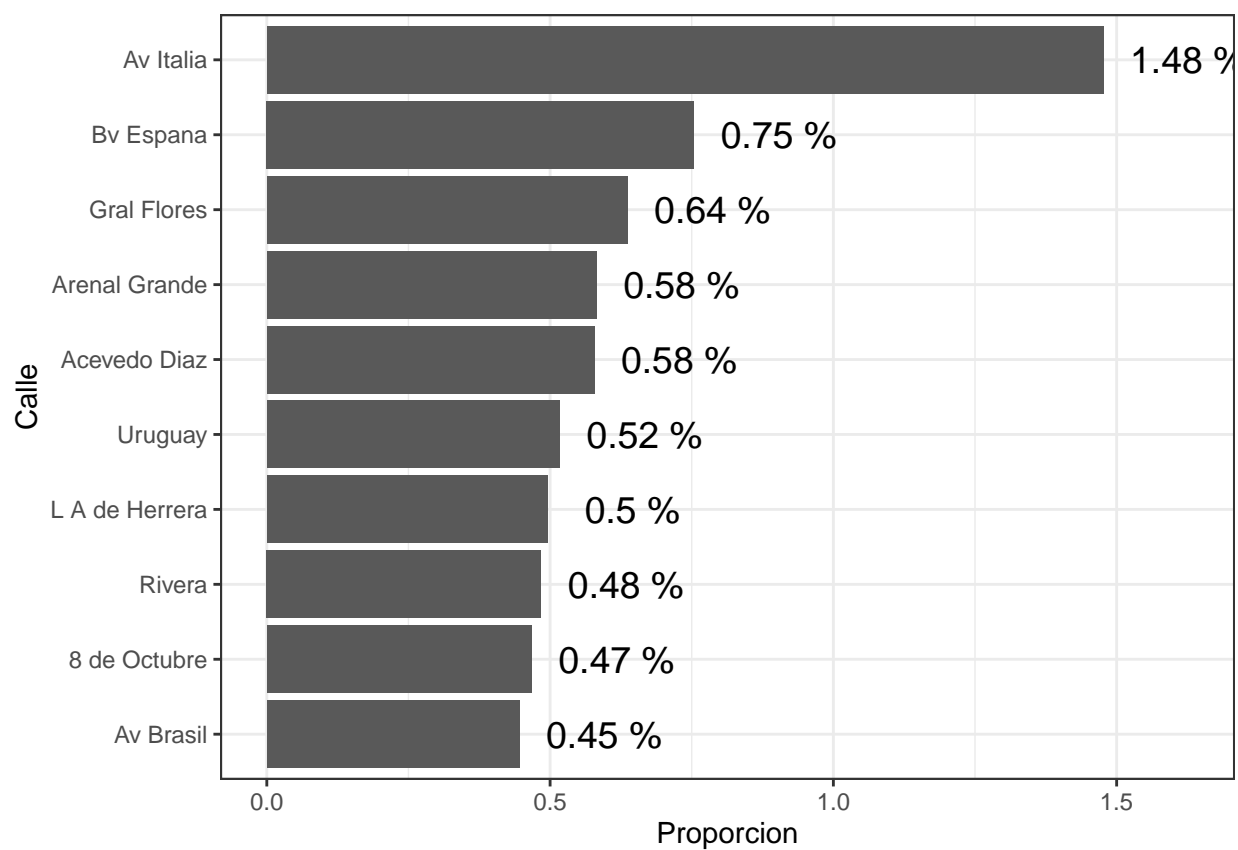
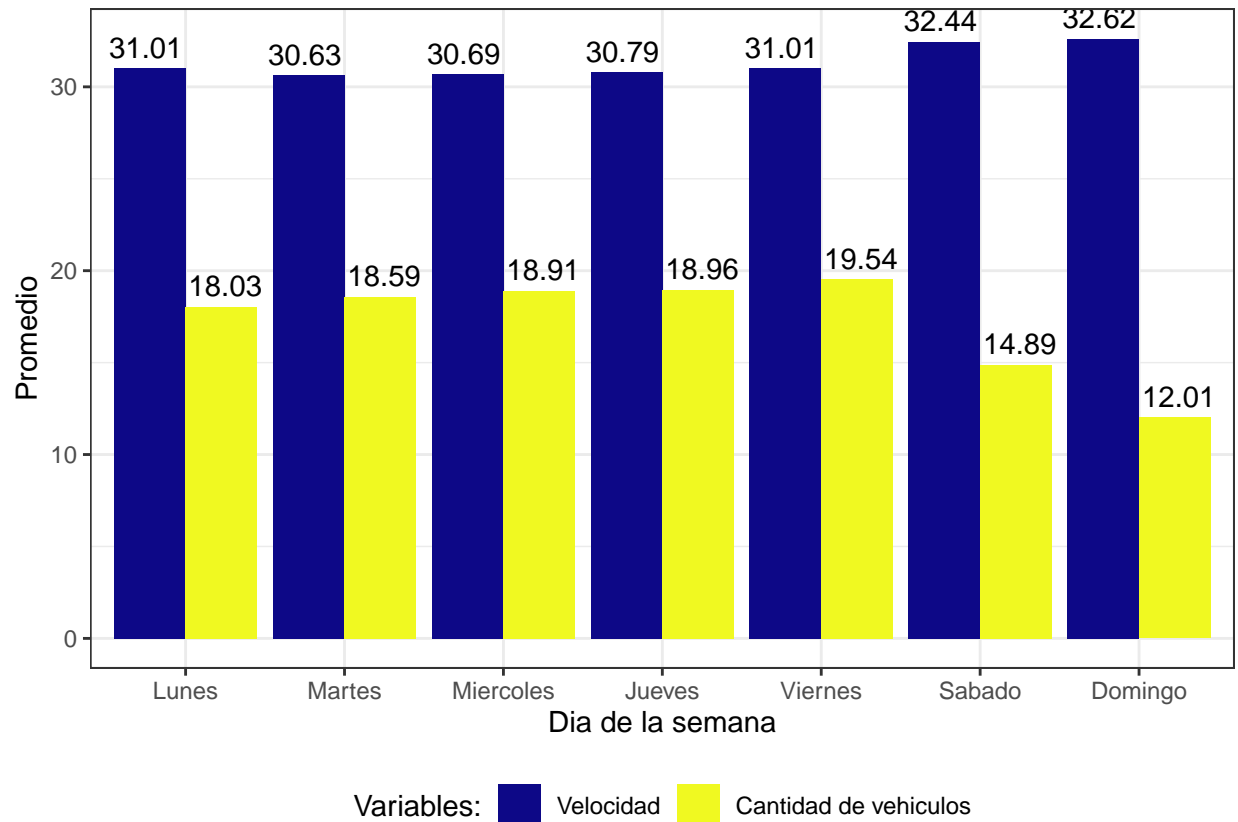


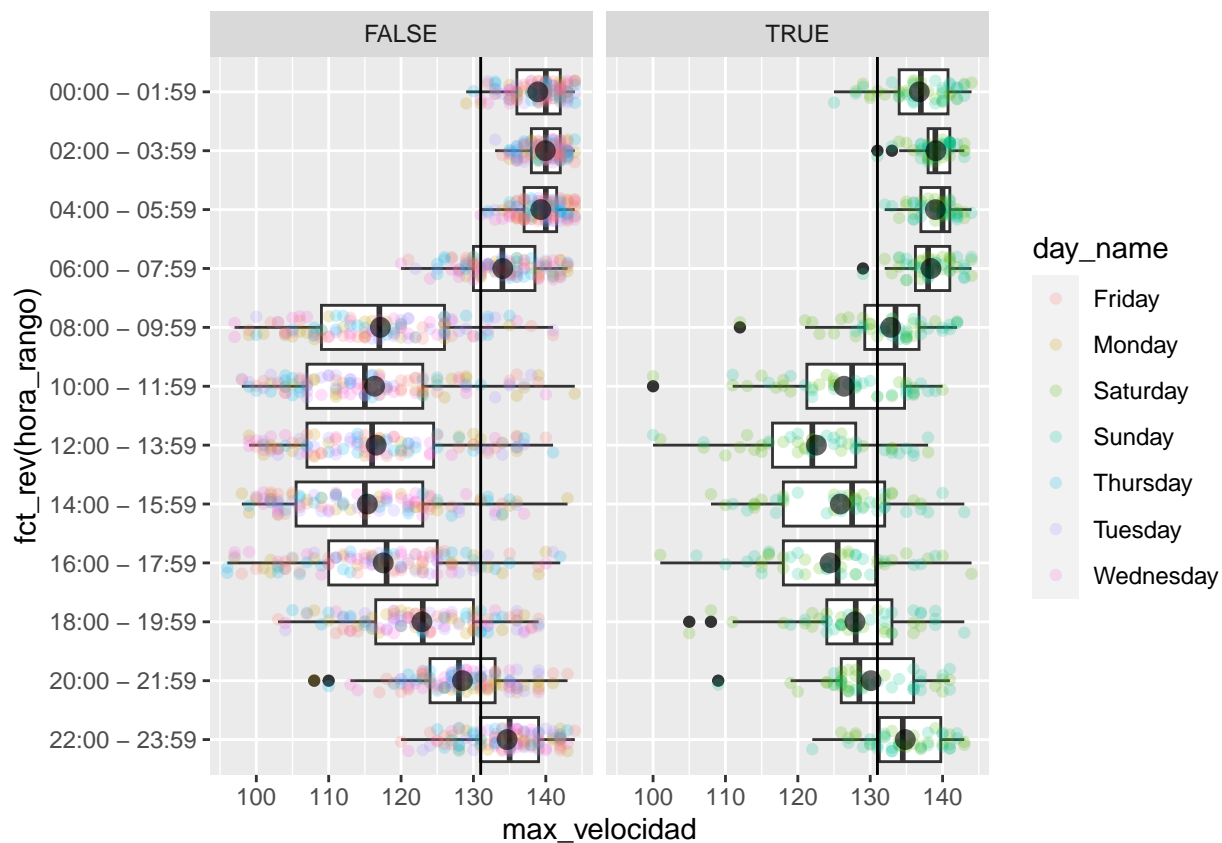
Figure 2: Es netamente nula la proporcion de excesos de velocidad, podemos decir el los conductores en general respetan los limites de velocidad



En el gráfico se puede observar: - La velocidad media se mantiene casi constante durante toda la semana, salvo una leve variación positiva los fines de semana. - El volumen medio de vehículos detectados por los radares va variando en la semana, alcanzado su pico los viernes. También es considerablemente inferior los fines de semana.

Si bien previamente obsevamos leves exesos de velocidad, dentro de esos excesos se registran velocidades maximas de mas de 100km/h, veamos esos maximos pro franja horaria segun si fue registrado en un fin de semana o





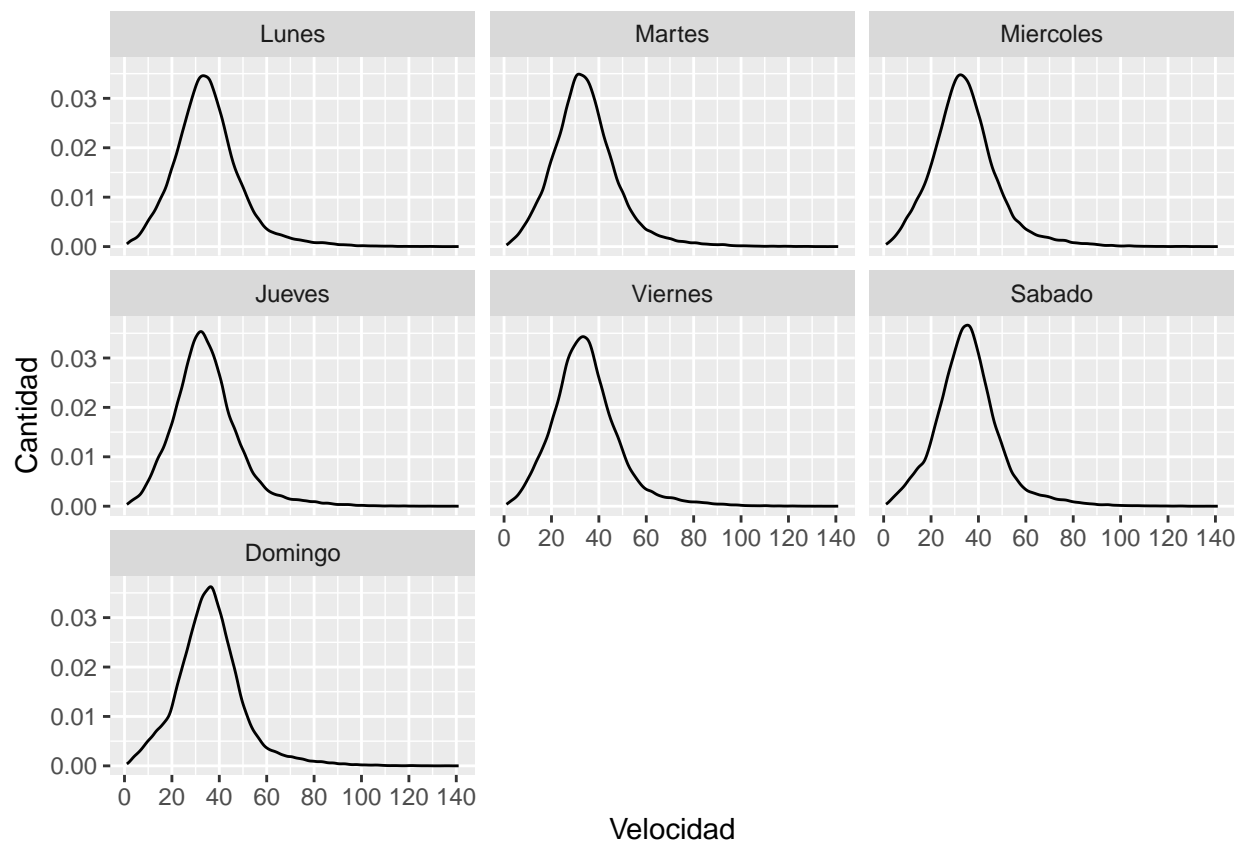
no

Entre las 00:00 y las 08:00 es donde se alcanzan picos de casi 140 km/h, luego de ese rango se reduce la tendencia por ser horario laboral para luego crecer a medida de que cae la noche. Durante el fin de semana las velocidades máximas durante el horario laboral de lunes a viernes es totalmente diferente, los conductores durante el fin de semana aprovechan la baja de volumen para circular mas rápido.

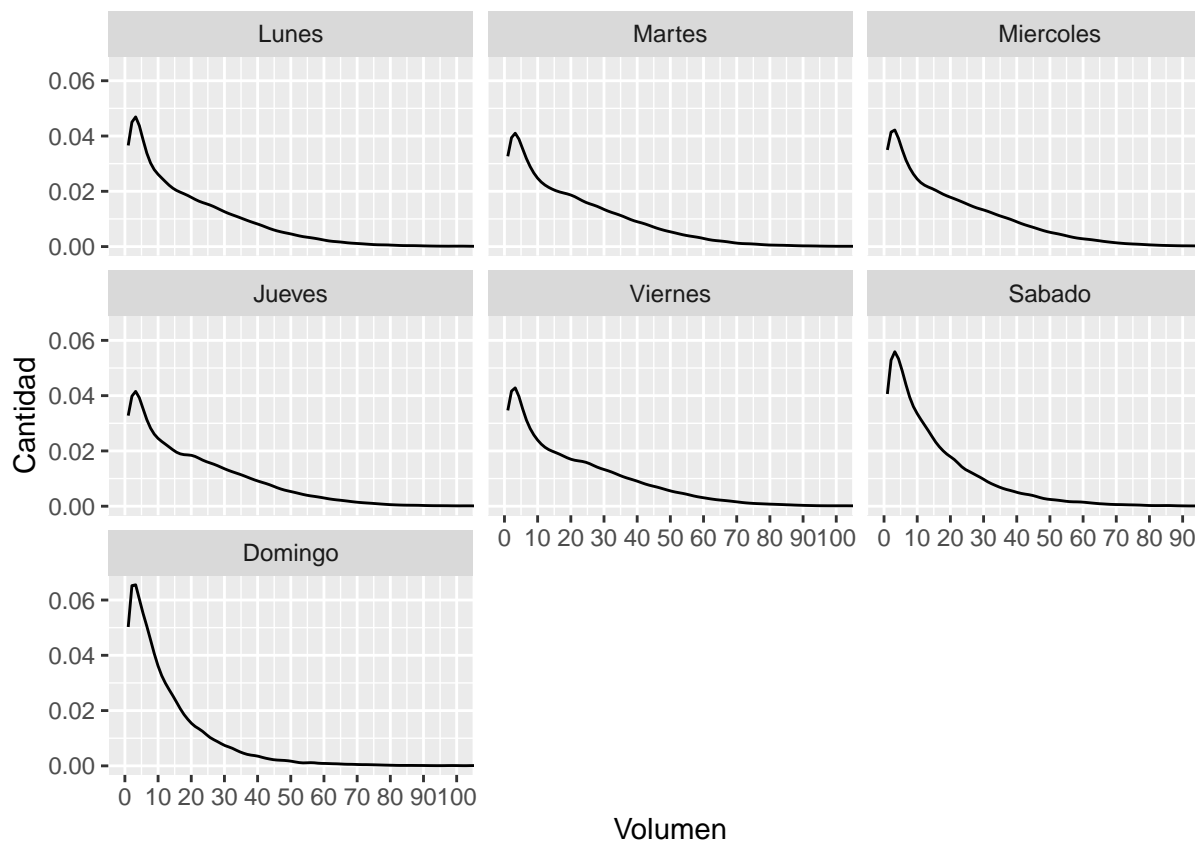
Si entramos en cada año vemos como crece la densidad de velocidad máxima registrada de lunes a viernes,



Ahora la desidad de velocidad durante los dias siempre se encuentran en un rango de  $(0,60]$



Con el volumen de trafico, hay una gran diferencia el domingo comparado con el resto de los dias, en todos los dias la mayoría de volumen se concentra en el rango  $(0,10]$  donde en dicho rango alcanza su pico de circulación



en todo Montevideo.

La verdad que no veo nada interesante, mas alla que cerrito es muy dispar al resto

## Resultados interesantes

[Presentar los resultados más relevantes obtenidos durante el análisis exploratorio.]

## Modelo estadístico

### Variables

[Describir la variable de respuesta y las variables explicativas utilizadas.]

### Elección del modelo

Intentaremos predecir como sera la velocidad en lo que queda del 2023

### Ajuste y evaluación

[Ajustar el modelo con el conjunto de entrenamiento y evaluarlo con el conjunto de prueba.]

## **Predicciones**

[Presentar las predicciones realizadas por el modelo.]

## **Interpretación de resultados**

[Interpretar los resultados obtenidos del modelo estadístico.]

## **Aplicación Shiny**

### **Descripción**

La aplicación consta de 4 partes. Un mapa con la ubicación de los semáforos, un análisis univariado, otro multivariado y el modelo. La aplicación resume todo lo hecho anteriormente. [Enlace](#)

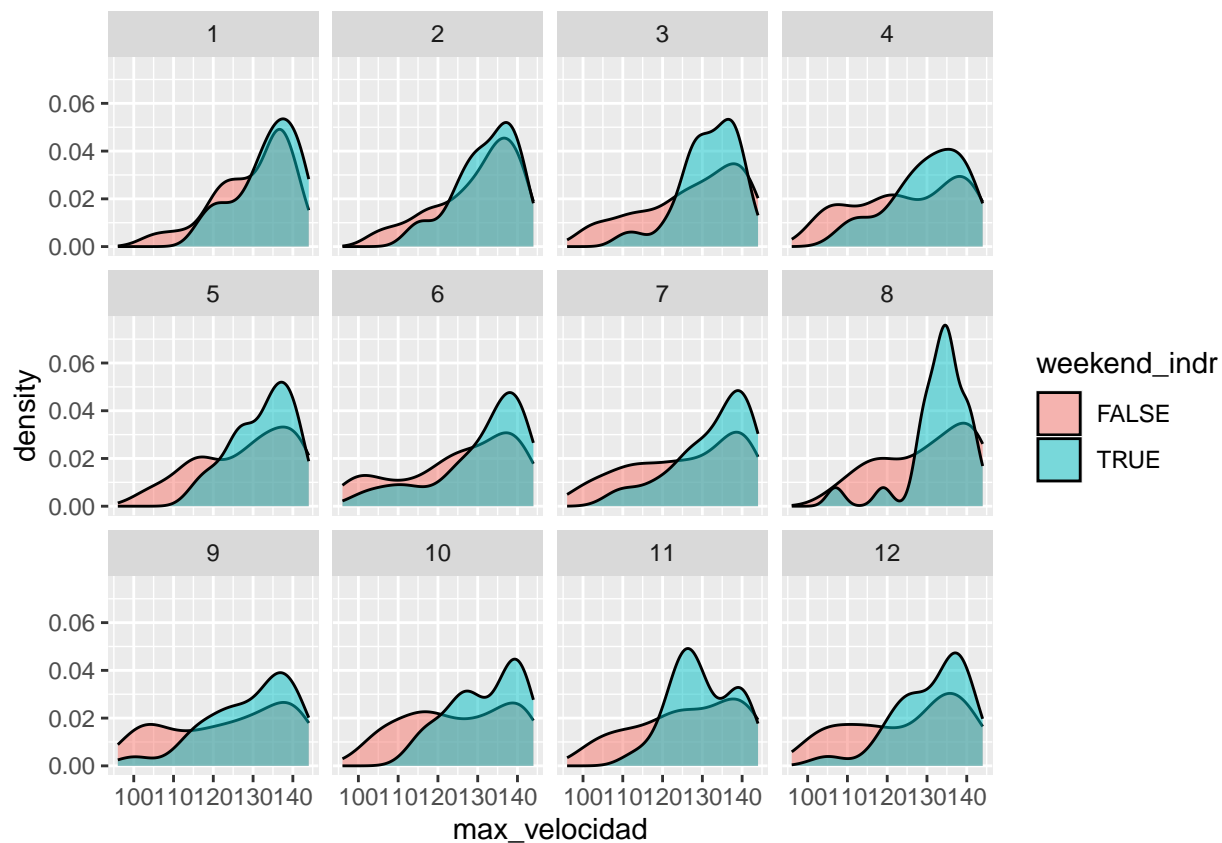
## **Comentarios finales**

### **Hallazgos principales**

[Resumir los principales hallazgos del proyecto.]

### **Posibles extensiones**

[Discutir posibles extensiones o mejoras para el proyecto.]



Se observa un pico en Agosto Dia de la nostalgia?

##	count	month_actual	year_actual
## 1	3830149	1	2022
## 2	3898585	1	2023
## 3	3714726	2	2021
## 4	3374893	2	2022
## 5	3455550	2	2023
## 6	3908628	3	2021
## 7	3746326	3	2022
## 8	3879877	3	2023
## 9	3759551	4	2021
## 10	3620380	4	2022
## 11	3930695	5	2021
## 12	2701766	5	2023
## 13	3873954	6	2021
## 14	3731149	6	2022
## 15	4078328	7	2021
## 16	3679998	7	2022
## 17	4032694	8	2021
## 18	3857623	9	2021
## 19	2501798	9	2022
## 20	3748052	10	2021
## 21	4072685	10	2022
## 22	3979504	11	2021
## 23	4009784	12	2021

Haciendo una observacion de los meses cubiertos por los datos, nos hemos dado cuenta que Agosto solo esta representado por los datos del 2021.

## Referencias

**¿Existe alguna correlación entre el volumen de tráfico y la velocidad promedio en las avenidas de Montevideo?**

El grafico no muestra una relacion aparente entre los promedios de las variables

## Muestra aleatoria de volumen y velocidad, agregados por dia de la semana

Vamos a analizar la relacion entre el volumen de trefico, la velocidad, el dia de la semana y el barrio donde se encuentra el sensor

WICKHAM, HADLEY. 2023. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'REILLY MEDIA.