

TP Estadística Descriptiva

Autores: Tomas Santucci y Fernando Silva

4 May 2022

1 Introduction

Este informe esta destinado a mostrar estadisticamente el uso de las ecobicis provenientes de la Ciudad Autonoma de Buenos Aires. Desde el mismo se podran observar datos Univariados y Bivariados para tener una mayor apreciación de la utilidad de las mismas.



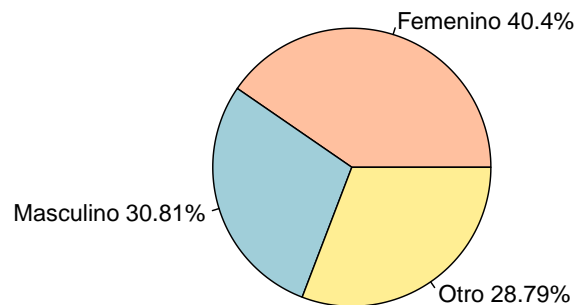
Análisis univariado

Variable género (Gráfico de Sectores)

Es una variable cualitativa cuyos valores posibles son: Masculino, Femenino y Otro

Representa el género de los usuarios que han utilizado las ecobici.

**Gráfico 1: GÉNERO DE LOS USUARIOS DE ECOBICI
CABA 2020**



Fuente: Datos sustraídos de las estaciones de las bicicletas publicas de CABA

En el Gráfico 1 se puede apreciar que la mayoría de los usuarios de las Ecobicis son Mujeres. Además hay una cantidad similar de usuarios de géneros masculino y otro.

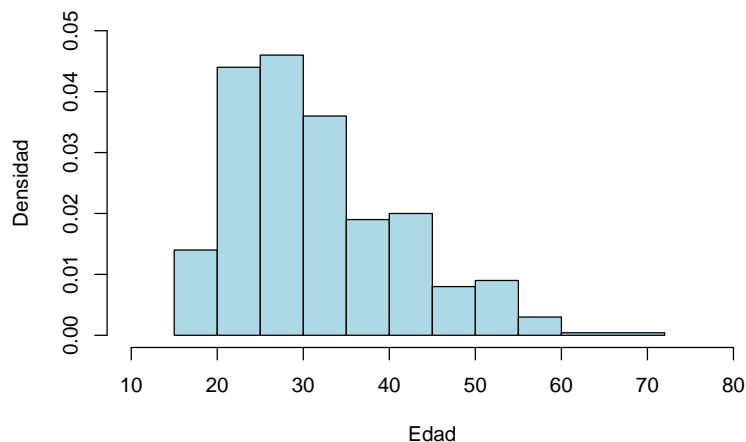
Medida de interés:

- Moda : Femenino

Variable edad (Histograma)

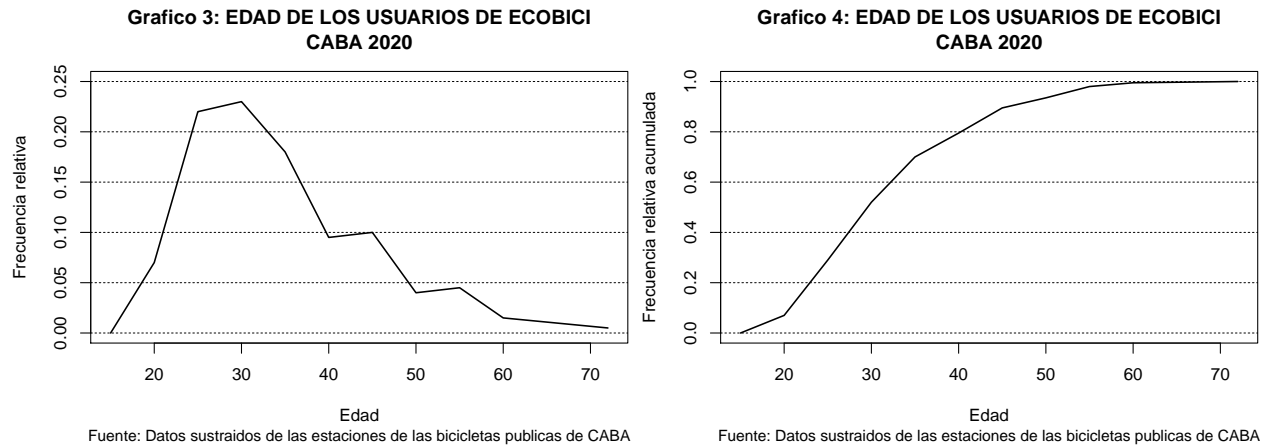
Es una variable cuantitativa continua cuyos valores están entre 16 y 72. Representa la edad de los usuarios de ecobici.

**Gráfico 2: EDAD DE LOS USUARIOS DE ECOBICI
CABA 2020**



Fuente: Datos sustraídos de las estaciones de las bicicletas publicas de CABA

Los datos resultan en un histograma unimodal asimétrico a derecha.



En el Gráfico 2 se muestra la variación del uso del servicio respecto de la edad de los usuarios. En el Grafico 3 y 4, se muestran los polígonos de Frecuencias Relativas y Absolutas de los usuarios respecto de la edad de los mismos.

En el Grafico 3, se puede ver claramente el aumento del uso de las ecobici, de manera creciente entre los usuarios cuyas edades van desde 20 a 30 años y luego va decreciendo a medida que aumenta la edad de los mismos. En otras palabras los usuarios de entre 20 a 35 años utilizan más el servicio, que aquellos usuarios mayores.

Medidas de interés:

- Mínimo: 16
- Mediana: 30
- Media o Esperanza Matematica: 31.97
- Variancia: 97.96
- Moda: 225
- Máximo: 72

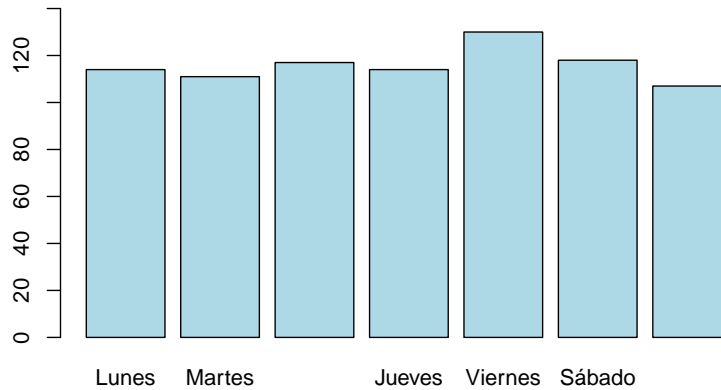
Variable días (Grafico de Barras)

Es una variable cualitativa cuyos valores posibles. Cuyos valores posibles son Lunes, Martes, Miércoles, Jueves, Viernes, Sábado y Domingo. Representa los días en que se efectuaron los viajes.

Table 1: Tabla de frecuencias

Estación	Frecuencia absoluta	Frecuencia relativa	Frecuencia absoluta acumulada	Frecuencia relativa acumulada
Lunes	114	0.1406	114	0.1406
Martes	111	0.1369	225	0.2774
Miércoles	117	0.1443	342	0.4217
Jueves	114	0.1406	456	0.5623
Viernes	130	0.1603	586	0.7226
Sábado	118	0.1455	704	0.8681
Domingo	107	0.1319	811	1

**Grafico 5: DIA DE LOS RECORRIDOS DE LAS ECOBICI
CABA 2020**



Fuente: Datos sustraídos de las estaciones de las bicicletas publicas de CABA

En el Grafico 5 se puede observar que la mayor cantidad de viajes se efectuan los dias viernes. Se puede suponer que ese día se efectuan dicha cantidad de viajes, debido a que es el ultimo día laboral de la semana y la mayoría de los usuarios obta por este servicio para poder regresar a sus casas a tiempo.

Medida de interés:

- Moda: Viernes

Variable estación de origen (Tabla de Frecuencias más Gráfico de Barras)

Es una variable cualitativa. Representa a las estaciones donde se originaron los viajes.

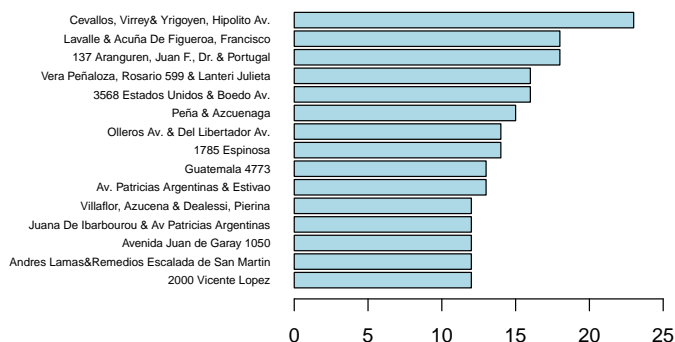
Se representan las estaciones más utilizadas, debido a la gran cantidad de estaciones de origen.

Table 2: Tabla de frecuencias

Estación	Frecuencia absoluta	Frecuencia relativa	Frecuencia absoluta acumulada	Frecuencia relativa acumulada
Cevallos, Virrey& Yrigoyen, Hipolito Av.	23	0.0284	23	0.0284
137 Aranguren, Juan F., Dr. & Portugal	18	0.0222	41	0.0506
Lavalle & Acuña De Figueroa, Francisco	18	0.0222	59	0.0727
3568 Estados Unidos & Boedo Av.	16	0.0197	75	0.0925
Vera Peñaloza, Rosario 599 & Lanteri Julieta	16	0.0197	91	0.1122
Peña & Azcuenaga	15	0.0185	106	0.1307
1785 Espinosa	14	0.0173	120	0.148
Olleros Av. & Del Libertador Av.	14	0.0173	134	0.1652
Av. Patricias Argentinas & Estivao	13	0.016	147	0.1813
Guatemala 4773	13	0.016	160	0.1973
2000 Vicente Lopez	12	0.0148	172	0.2121
Andres Lamas&Remedios	12	0.0148	184	0.2269
Escalada de San Martin				

Estación	Frecuencia absoluta	Frecuencia relativa	Frecuencia absoluta acumulada	Frecuencia relativa acumulada
Avenida Juan de Garay 1050	12	0.0148	196	0.2417
Juana De Ibarbourou & Av	12	0.0148	208	0.2565
Patricias Argentinas				
Villaflor, Azucena & Dealessi, Pierina	12	0.0148	220	0.2713

ESTACIONES DE ORIGEN MAS USADAS DE LAS ECOE CABA 2020



Fuente: Datos sustraídos de las estaciones de las bicicletas publicas de C

Se puede ver a partir de la tabla que alrededor del 27% de los viajes se originaron en sólo 15 de las 169 estaciones.

Medida de interés:

- Moda: Cevallos, Virrey & Yrigoyen, Hipolito Av.

Variable estación de Destino (Tabla de Frecuencias más Gráfico de Barras)

Es una variable cualitativa. Representa a las estaciones de destino de cada viaje.

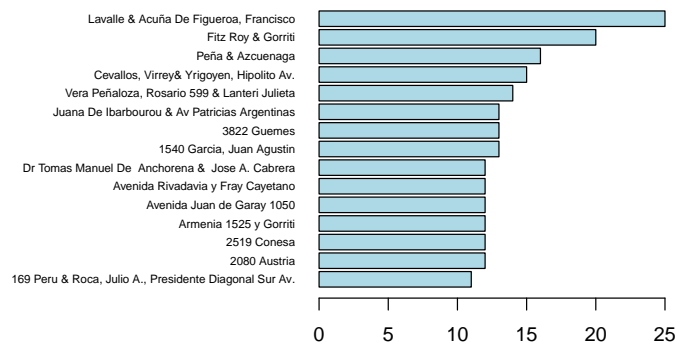
Se representan las estaciones más utilizadas, debido a la gran cantidad de estaciones de destino.

Table 3: Tabla de frecuencias

Estación	Frecuencia absoluta	Frecuencia relativa	Frecuencia absoluta acumulada	Frecuencia relativa acumulada
Lavalle & Acuña De Figueroa, Francisco	25	0.0308	25	0.0308
Fitz Roy & Gorriti	20	0.0247	45	0.0555
Peña & Azcuenaga	16	0.0197	61	0.0752
Cevallos, Virrey & Yrigoyen, Hipolito Av.	15	0.0185	76	0.0937
Vera Peñaloza, Rosario 599 & Lanteri Julieta	14	0.0173	90	0.111
1540 Garcia, Juan Agustin	13	0.016	103	0.127
3822 Guemes	13	0.016	116	0.143
Juana De Ibarbourou & Av	13	0.016	129	0.1591
Patricias Argentinas				
2080 Austria	12	0.0148	141	0.1739

Estación	Frecuencia absoluta	Frecuencia relativa	Frecuencia absoluta acumulada	Frecuencia relativa acumulada
2519 Conesa	12	0.0148	153	0.1887
Armenia 1525 y Gorriti	12	0.0148	165	0.2035
Avenida Juan de Garay 1050	12	0.0148	177	0.2182
Avenida Rivadavia y Fray Cayetano	12	0.0148	189	0.233
Dr Tomas Manuel De Anchorena & Jose A. Cabrera	12	0.0148	201	0.2478
169 Peru & Roca, Julio A., Presidente Diagonal Sur Av.	11	0.0136	212	0.2614

ESTACIONES DESTINO MAS USADAS DE LAS ECOE CABA 2020



Fuente: Datos sustraídos de las estaciones de las bicicletas publicas de CABA

Se puede ver que a partir de la tabla, alrededor de 26% de los viajes tuvieron como destino solo 15 de las 171 estaciones.

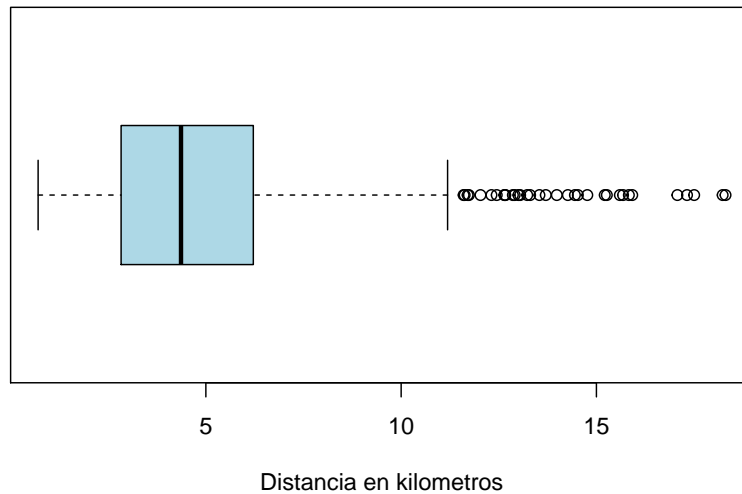
Medida de interés:

- Moda: Lavalle & Acuña De Figueroa, Francisco

Variable distancia

Es una variable cuantitativa continua. Sus valores están entre 0 y 20 km. Cualquier apreciación mayor a 20 km es considerada un error de sistema o humano y por lo tanto no es tenida en cuenta en el análisis.

DISTANCIA DE LOS RECORRIDOS DE ECOBICI CABA 2020



Fuente: Datos sustraídos de las estaciones de las bicicletas publicas de CABA

En el boxplot se puede ver que:

El mínimo recorrido fue 700 metros y el máximo fue 18.30 km

El 25% recorrió menos de 2.83 km.

El 50% recorrió menos de 4.36 km.

El 75% recorrió menos de 6.21 km.

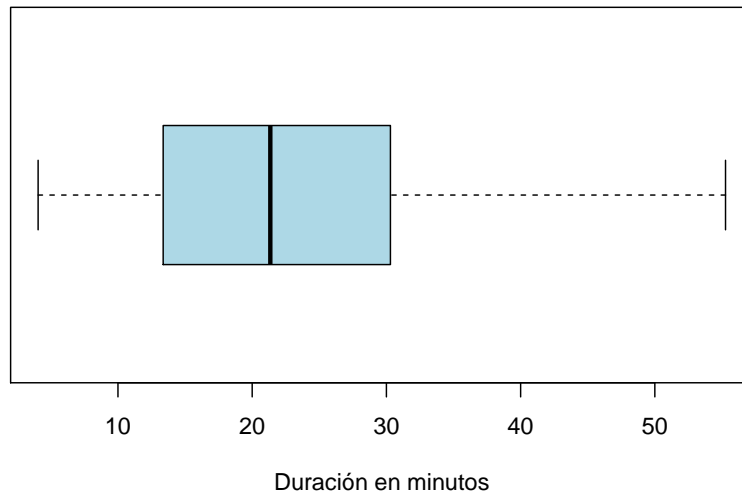
Medidas de interés:

- Mínimo: 0.70 km
- Mediana: 4.36 km
- Media o Esperanza Matematica: 4.96 km
- Variancia: 9.37 km
- Maximo: 18.30 km (outlier)

Variable duración

Es una variable cuantitativa continua. Sus valores están entre 4 min y 90 min. Cualquier apreciación mayor a 90 minutos es considerada un error de sistema o humano y por lo tanto no es tomada en cuenta en el análisis.

DURACION DE LOS RECORRIDOS DE ECOBICI CABA 2020



Fuente: Datos sustraídos de las estaciones de las bicicletas publicas de CABA

A partir del Boxplot se puede ver que:

El minimo tiempo transcurrido fue 4 min y el maximo fue 87 min

El 25% de los usuarios utilizo el servicio menos de 13 min

El 50% de los usuarios utilizo el servicio menos de 21 min

El 75% de los usuarios utilizo el servicio menos de 30 min

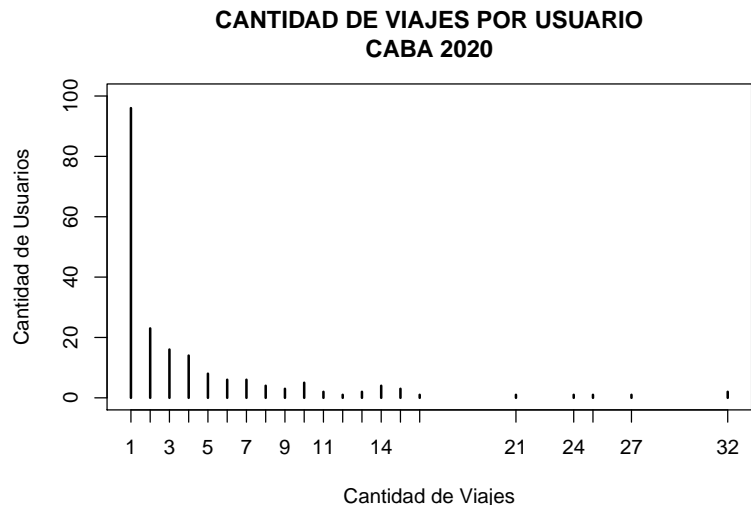
Es interesante notar que un 25% de los viajes se excedieron del tiempo límite de 30 minutos.

Medidas de interés:

- Minimo: 4.05 min
- Mediana: 21.35 min
- Media o Esperanza Matematica: 23.84 min
- Variancia: 201.10 min
- Maximo: 87.37 min (outlier)

Variable cantidad de viajes por usuario

Es una variable cualitativa, cuyos valores posibles estan entre 1 y 32 viajes.



Fuente: Datos sustraídos de las estaciones de las bicicletas publicas de CABA

A partir del gráfico de bastones se puede ver que:

96 Usuarios realizaron un solo viaje.

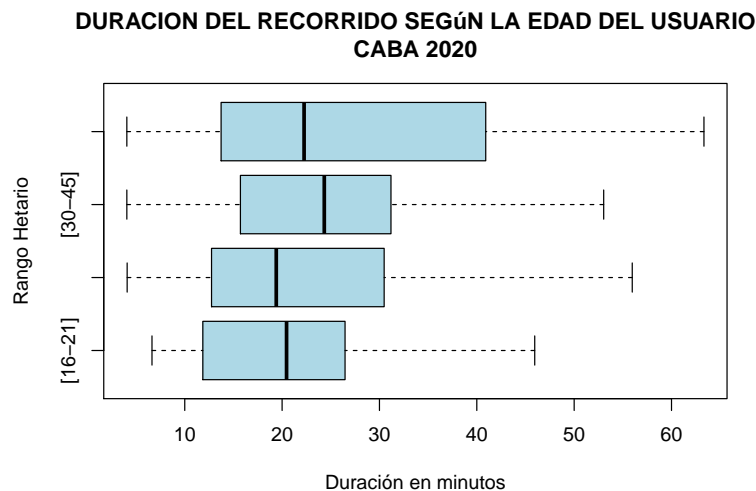
Solamente 2 usuarios realizaron 32 viajes.

Medida de interes:

- Moda: 1 viaje

Análisis bivariado

Se realiza un análisis de las variables (cuantitativas continuas) edad y distancia recorrida. Se divide la edad a partir de 4 rangos hetarios por lo que se transforma en una variable cualitativa.



Fuente: Datos sustraídos de las estaciones de las bicicletas publicas de CABA

El rango de la variable en el grupo de 45 a 72 es el más extenso y tiene el mayor máximo.

Se podría creer que una posible explicación sería que la gente mayor no acredita correctamente el fin de sus viajes.

A medida que avanzan las edades se puede ver que cada vez más gente se acerca e incluso sobrepasa el límite de 30 minutos. Siendo el grupo hetario [16,21] el que más respeta el tiempo límite. Y el grupo de [45-72] es el que más lo excede.