

Certificación Avanzada en Ciencia de Datos



Trabajo Práctico ~ Visualización de Información

Profesor Diego Ariel Aizemberg

Grupo 1

Odorico, María Del Pilar ~ Leiva, María Julia ~ Paz, Francisco

Trabajo Práctico Uso del sistema de bicicletas (CABA)

Introducción

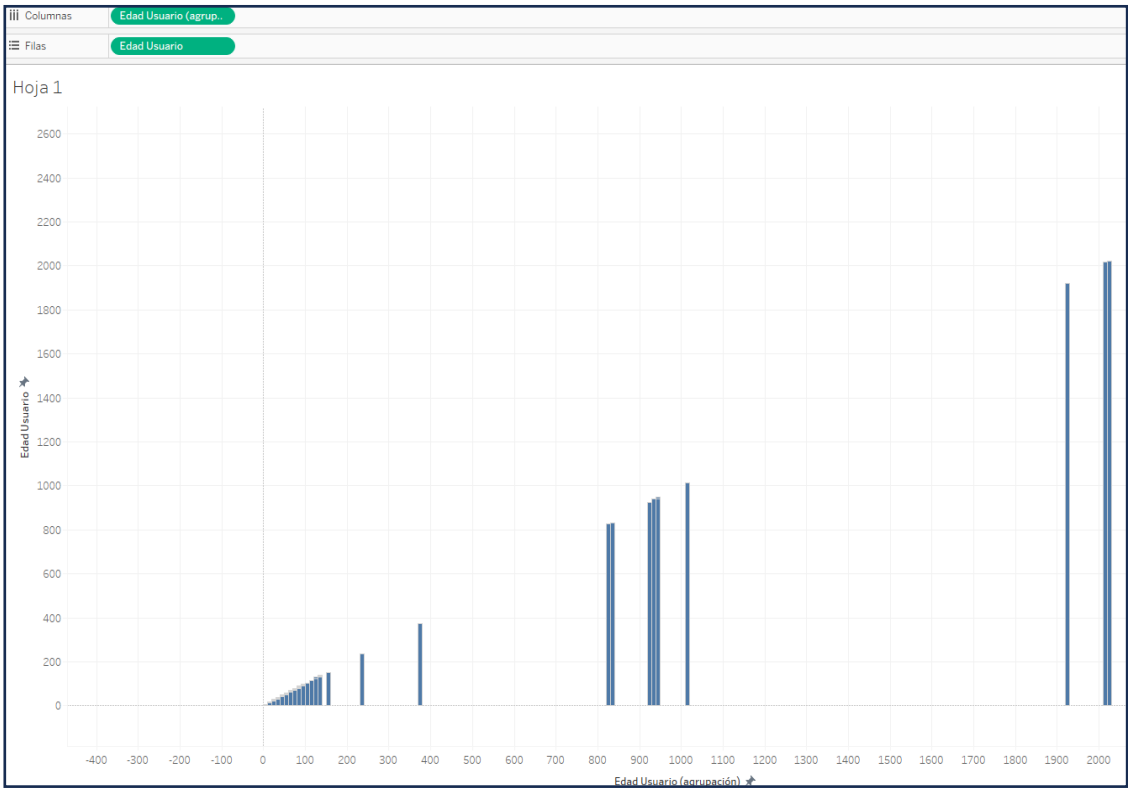
El uso de sistemas de bicicletas compartidas ha crecido significativamente en las últimas décadas como una alternativa sostenible de transporte en las grandes ciudades. La Ciudad Autónoma de Buenos Aires (CABA) no es la excepción, habiendo implementado su propio sistema de bicicletas públicas para fomentar la movilidad urbana y reducir la congestión vehicular.

Este trabajo se enfoca en analizar el uso del sistema de bicicletas de CABA, utilizando un conjunto de datos que recopila información detallada sobre los viajes realizados. Este dataset incluye variables clave como la duración de los viajes, las estaciones de origen y destino, la fecha y hora de uso, y el tipo de usuario, entre otros. A partir de este análisis, se pretende identificar patrones de uso, perfiles de usuarios, y factores que influyen en la demanda del servicio.

El objetivo es proporcionar una visión integral del funcionamiento del sistema, con el fin de ofrecer recomendaciones que puedan optimizar su operación y promover un mayor uso del mismo. Además, se busca entender cómo este servicio contribuye a la movilidad sustentable en la ciudad y qué mejoras pueden implementarse para incrementar su impacto positivo en la comunidad.

Análisis exploratorio de los datos

Dentro de lo encontrado dentro del dataset se pudieron observar gracias a un histograma varios outliers con respecto a la edad.



Columnas innecesarias que fueron eliminadas, ejemplo:

	A ^B _C modelo_bicicleta	A ^B _C género	A ^B _C Género.1
	ICONIC	OTHER	NA
	ICONIC	OTHER	NA
	ICONIC	FEMALE	NA
	ICONIC	FEMALE	NA
	ICONIC	FEMALE	NA
	ICONIC	FEMALE	NA
	ICONIC	MALE	NA
	ICONIC	OTHER	NA
	ICONIC	FEMALE	NA
	ICONIC	FEMALE	NA
	ICONIC	FEMALE	NA
	ICONIC	FEMALE	NA
	ICONIC	OTHER	NA
	ICONIC	MALE	NA
	ICONIC	MALE	NA
	ICONIC	MALE	NA
	ICONIC	MALE	NA
	ICONIC	FEMALE	NA
	ICONIC	MALE	NA
	ICONIC	MALE	NA
	ICONIC	FEMALE	NA

Valores en blanco que decidimos transformar a valores OTHER para evitar inconsistencias en los datos:

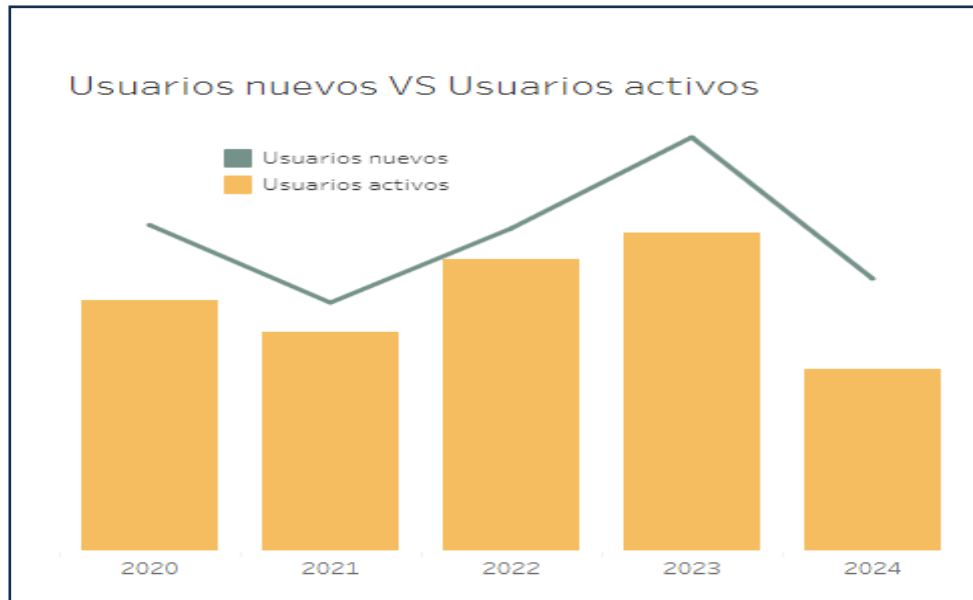
1.2	long_estacion_destino	1.2	lat_estacion_destino	A ^B _C	id_usuario	A ^B _C	modelo_bicicleta	A ^B _C	género
	-58.36882117		-34.60515159		233678				
	-58.4346883		-34.5752773		242168				
	-58.3825498		-34.5938629		208528				
	-58.4220694		-34.5982097		237648				
	-58.372763		-34.6033685		191568				
	-58.3825498		-34.5938629		208528				
	-58.4299793		-34.5829625		215098				
	-58.4111136		-34.5850763		215098				
	-58.4335573		-34.6079414		249778				
	-58.4198163		-34.5915614		197348				
	-58.4092876		-34.617509		214268				
	-58.380707		-34.626851		237408				
	-58.484297		-34.616813		199968				
	-58.3941087		-34.5936508		186808				
	-58.407741		-34.585443		238358				
	-58.4056671		-34.5825475		238358				
	-58.4016808		-34.6212681		205658				
	-58.4016808		-34.6212681		205658	BAEcobici	ICONIC		
	-58.391768		-34.597225		208528	BAEcobici	ICONIC		
	-58.46763869		-34.66036166		191568	BAEcobici	ICONIC		
	-58.436115		-34.563539		201648	BAEcobici	ICONIC		
	-58.4335573		-34.6079414		217708	BAEcobici	ICONIC		

Columnas con valores mezclados que fueron limpiados.

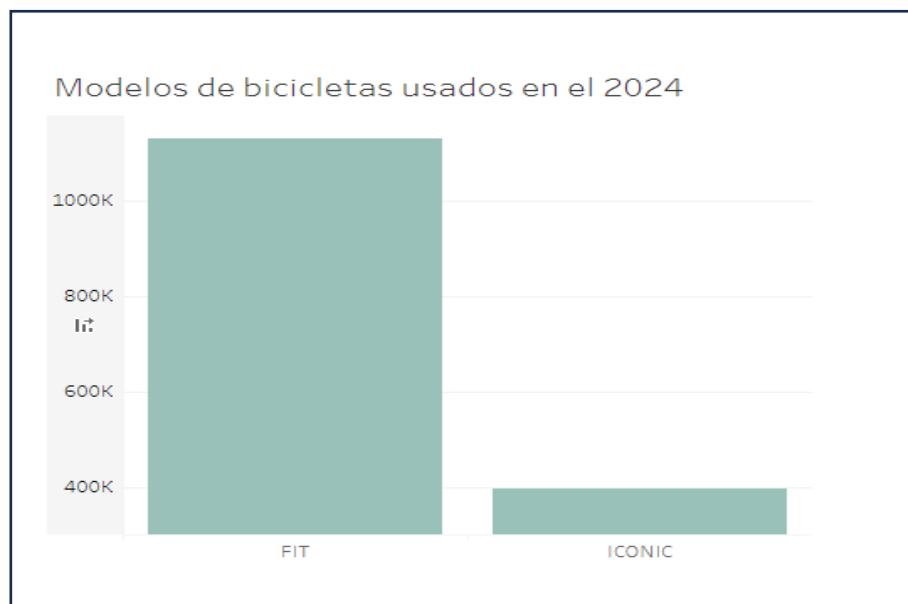
1.2	long_estacion_destino	A ^B _C	lat_estacion_destino	A
	-34.6094218	-34.6094218,-58.3893364		2
	-34.585443	-34.585443,-58.407741		6
	-34.61009	-34.61009,-58.406		7
	-34.577424	-34.577424,-58.426387		1
	-34.577424	-34.577424,-58.426387		6
	-34.577424	-34.577424,-58.426387		5
	-34.577424	-34.577424,-58.426387		4
	-34.6064101	-34.6064101,-58.4187306		6
	-34.5894269	-34.5894269,-58.4161178		6
	-34.5894269	-34.5894269,-58.4161178		6
	-34.5894269	-34.5894269,-58.4161178		7
	-34.600752	-34.600752,-58.3638723		8
	-34.5908211	-34.5908211,-58.3973698		2
	-34.5964246	-34.5964246,-58.371847		3
	-34.5964246	-34.5964246,-58.371847		6
	-34.6371232	-34.6371232,-58.4058883		4
	-34.5755148	-34.5755148,-58.4138829		7
	-34.5755148	-34.5755148,-58.4138829		7

Preguntas y Gráficos:

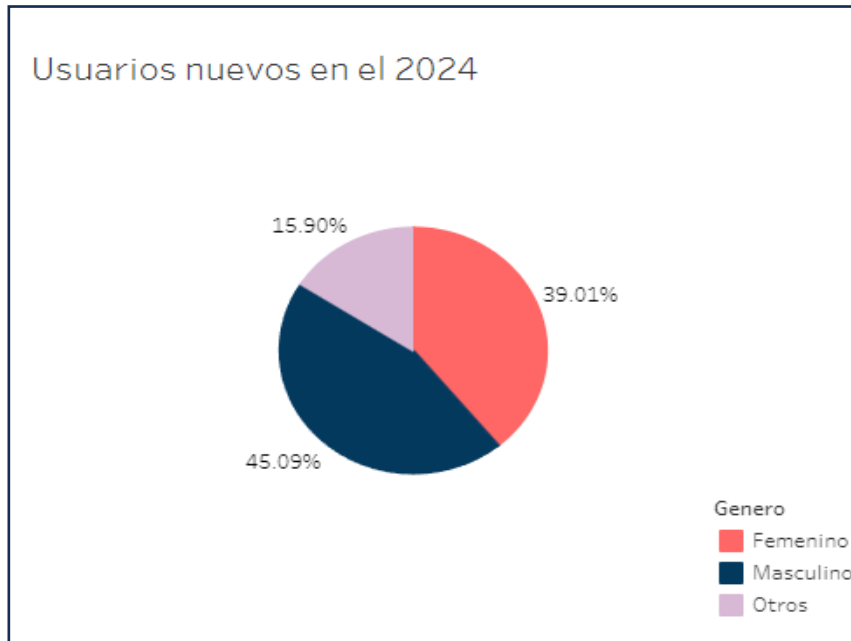
¿Cuántos usuarios nuevos y cuantos usuarios activos hay por año?



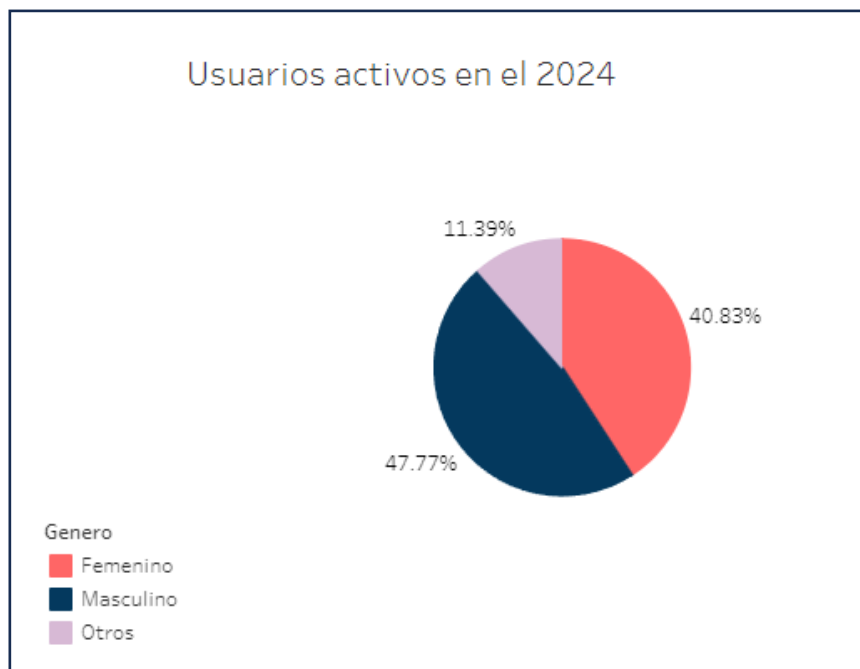
¿Al 2024, cuáles son los modelos de bicicletas que ofrece el servicio? ¿Cual de ellos es el más utilizado?



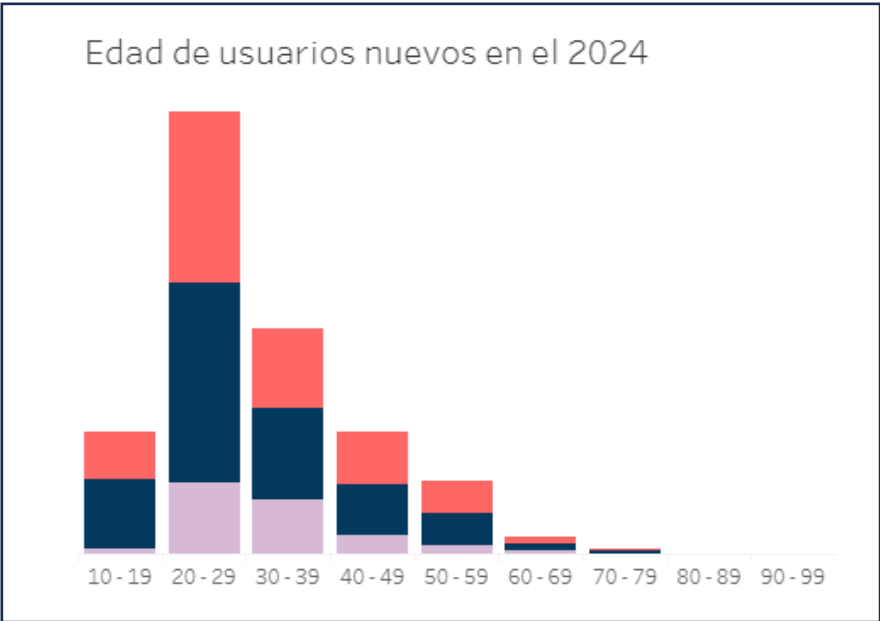
¿Cuál es la proporción de genero de los usuarios nuevos del 2024?



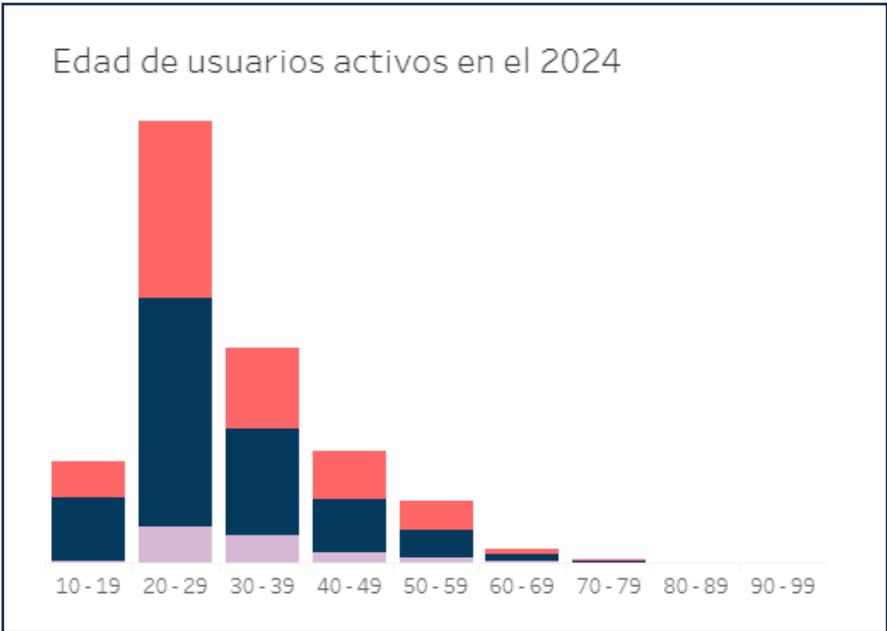
¿Cuál es la proporción de genero de los usuarios activos del 2024?



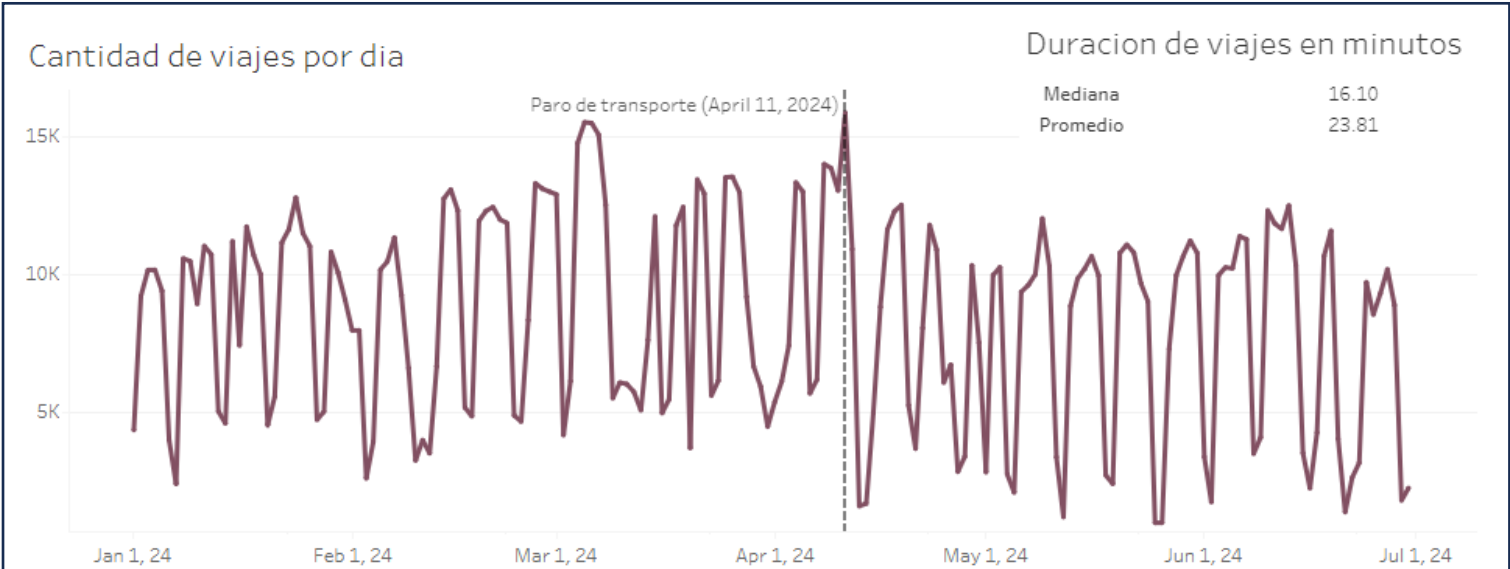
¿Cuál es la edad de los usuarios nuevos del 2024?



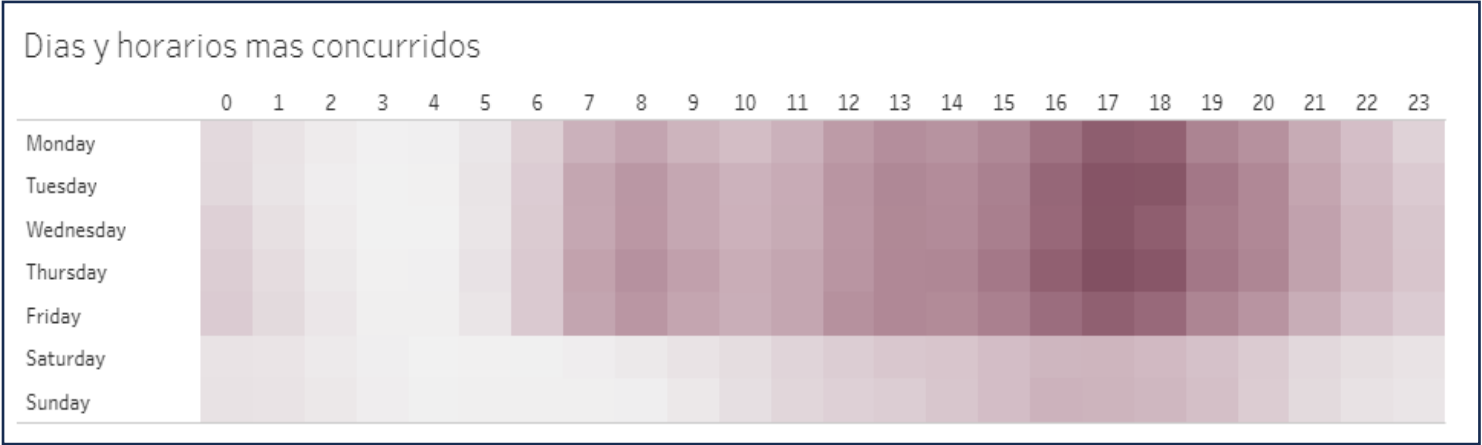
¿Cuál es la edad de los usuarios activos del 2024?



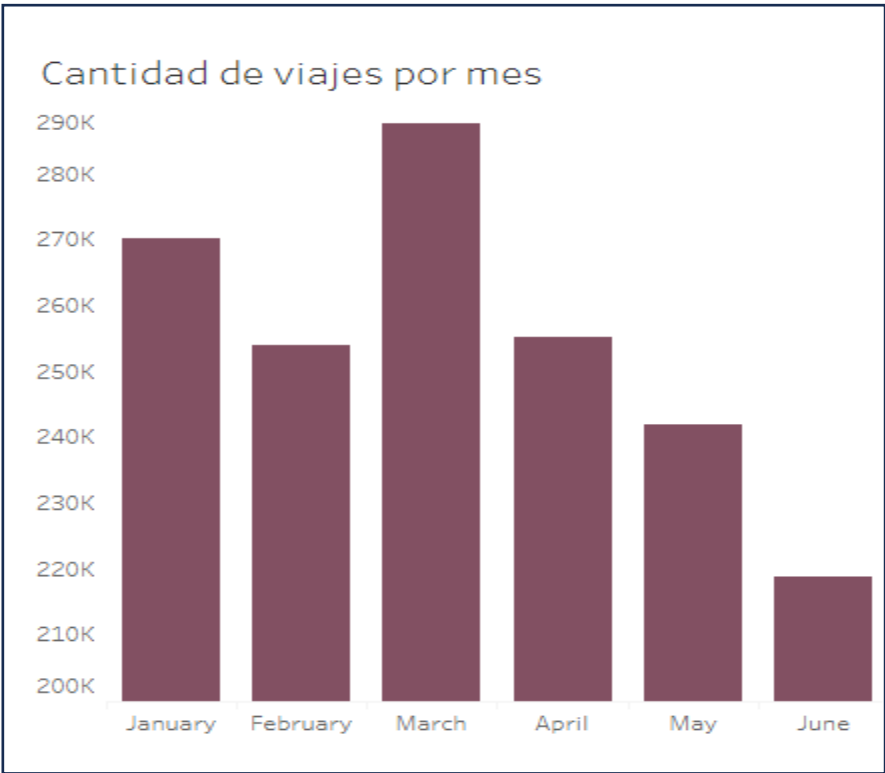
¿Cómo vario la cantidad de viajes por día en lo que va del 2024? ¿A que se debe el punto máximo de viajes el 11 de abril? ¿En general, cuanto duran los recorridos?



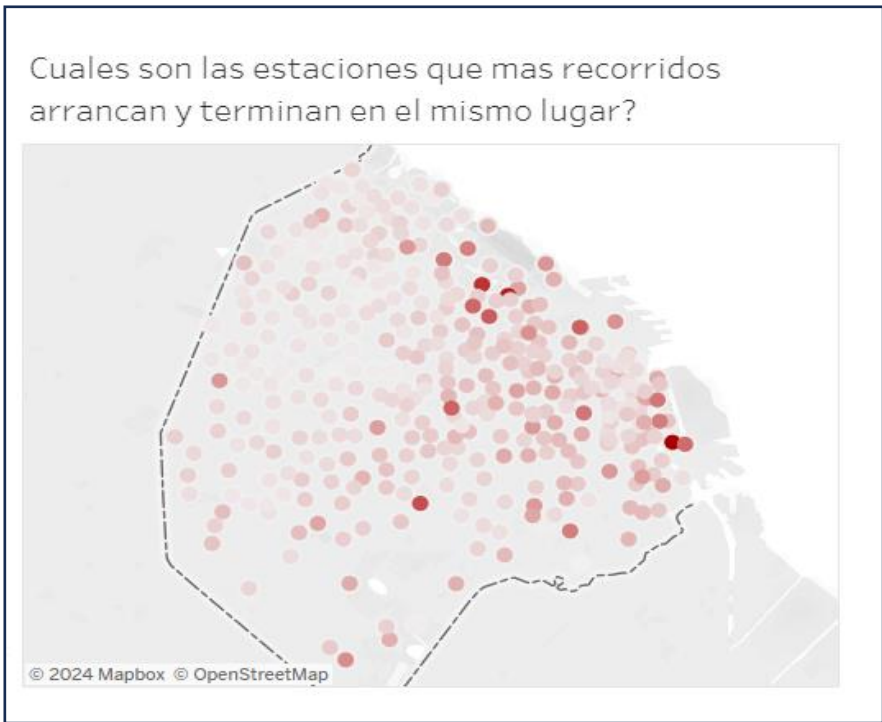
¿Cuáles son los días de la semana y los horarios en donde más se utiliza el servicio?



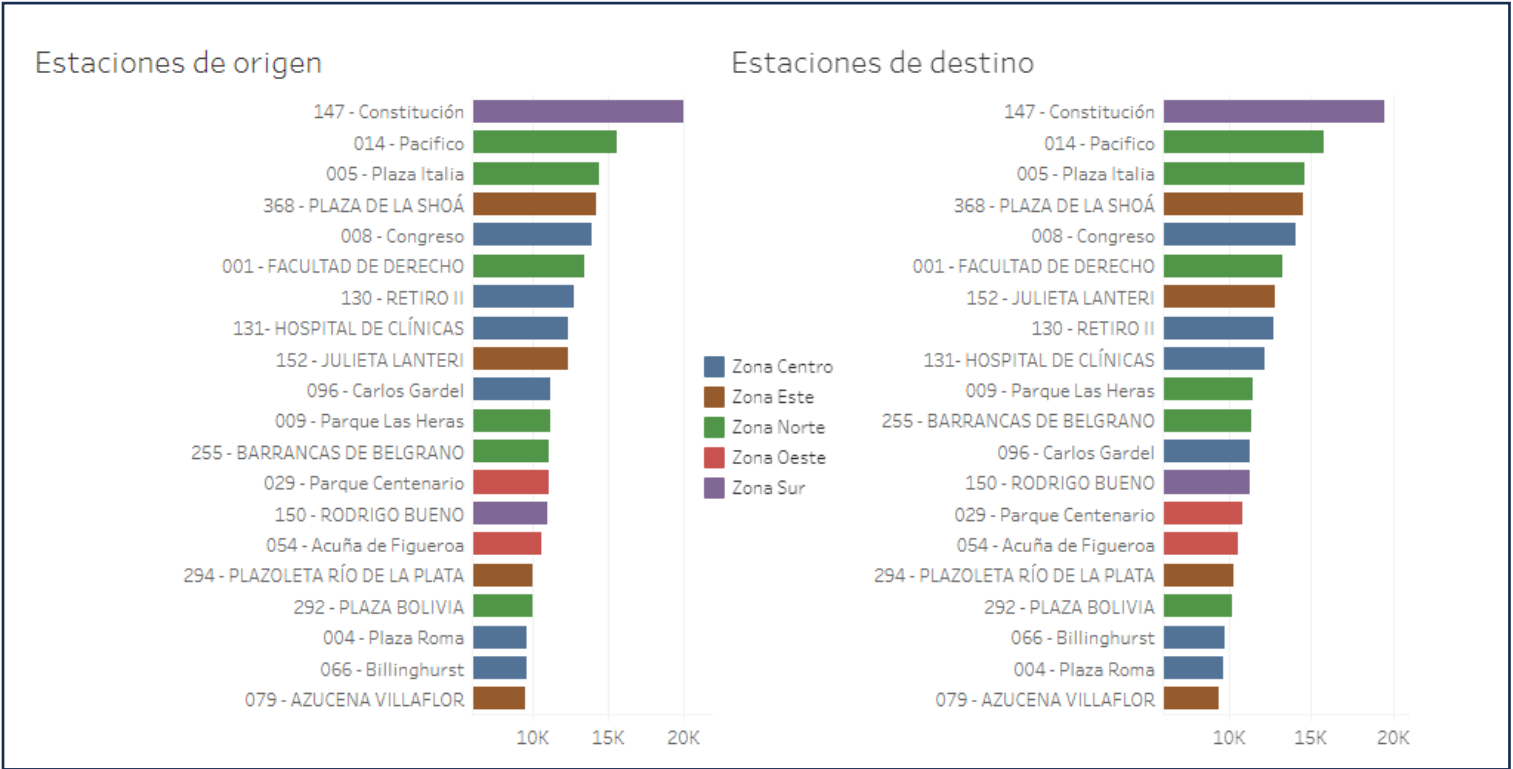
¿Cómo varía la cantidad de viajes a lo largo de los meses de lo que va del 2024?



En algunos casos, el servicio no se utiliza para llegar de un punto a otro. ¿Cuáles son las estaciones donde más recorridos inician y terminan?



¿Cuáles son las estaciones donde más recorridos inician? ¿Cuáles son las estaciones donde más recorridos Finalizan? ¿A qué zona de la ciudad pertenecen?



Tecnologías usadas:

- Power Query: EDA, Limpieza y manipulación de datos.**
- Tableau: EDA y Generación de Gráficos.**