

Sistema de transporte ECOBICI

Rudik Rompich - Alejandro Pallais

Descripción de los datos

- <https://ecobici.cdmx.gob.mx/datos-abiertos/>
- Webscraping y Concatenar
- Datos de 2010 a 2024 (5.35 GB)
- Datos de 2023 (716 MB)

Descripción de los datos (Estaciones)

	id	name	address	addressNumber	zipCode	districtCode	districtName	altitude	nearbyStations	stationType	location.lat	location.lon
0	448	448 DR. ANDRADE - ARCOS DE BELÉN	DR. ANDRADE ARCOS DE BELÉN	S/N	None	None	None	None	[448]	BIKE,TPV	19.426611	-99.144470
1	443	443 BRUNO TRAVEN-GOLONDRINAS	443 - Bruno Trave-Avenida México-Coyoacan	S/N	3340	1	Ampliación Granada	None	[432, 442, 444]	BIKE	19.359655	-99.162025
2	463	463 LAGO ANDRÓMACO-LAGO ZURICH	463 LAGO ANDRÓMACO-LAGO ZURICH	S/N	None	1	Ampliación Granada	None	[197, 242, 462]	BIKE	19.442338	-99.202492
3	416	416 RECREO-PARROQUÍA	416 - Recreo-Parroquia	S/N	3104	1	Ampliación Granada	None	[390, 391, 415]	BIKE	19.371308	-99.177607
4	437	437 MORAS-COMUNAL	437 - Moras-Comunal	S/N	3240	1	Ampliación Granada	None	[423, 436, 438]	BIKE	19.364808	-99.176022

Descripción de los datos

(Data mensual)

Genero_Usuario	Edad_Usuario	Bici	Ciclo_Estacion_Retiro	Fecha_Retiro	Hora_Retiro	Ciclo_EstacionArribo	Fecha_Arribo	Hora_Arribo	
0	M	45.0	3297082	150	31/12/2022	23:58:00	040	01/01/2023	00:06:24
1	F	27.0	8922254	214	31/12/2022	23:36:34	025	01/01/2023	00:07:08
2	M	34.0	2320075	214	31/12/2022	23:36:33	025	01/01/2023	00:07:12
3	M	58.0	2096493	137	31/12/2022	23:57:19	036	01/01/2023	00:07:48
4	M	23.0	6294433	260	31/12/2022	23:57:42	043	01/01/2023	00:09:33

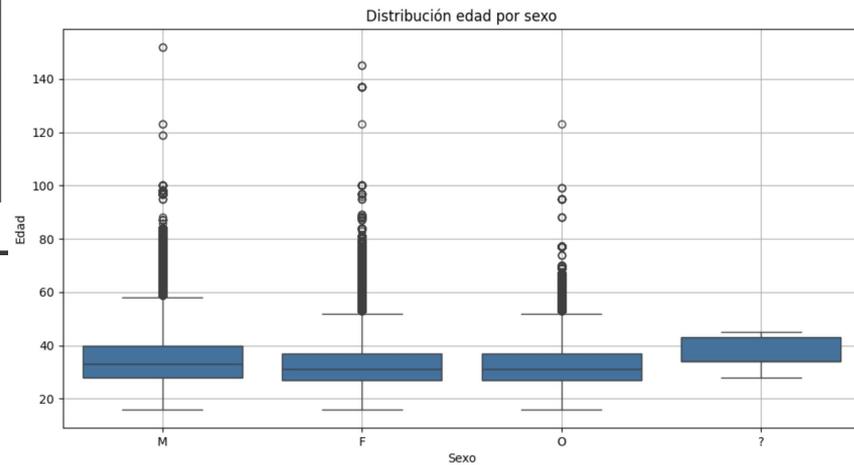
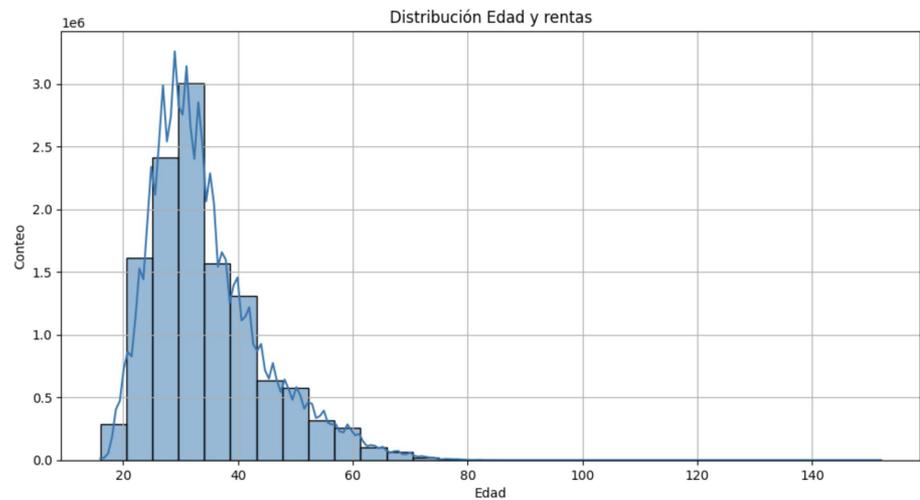
Exploración y variables importantes

- La exploración inicial reveló variables clave, incluyendo datos demográficos del usuario (**`Genero_Usuario`**, **`Edad_Usuario`**)
- Detalles de la transacción (**`Bici`**, **`Ciclo_Estacion_Retiro`**, **`Fecha_Retiro`**, **`Hora_Retiro`**, **`Ciclo_EstacionArribo`**, **`Fecha_Arribo`**, **`Hora_Arribo`**), que proporcionó información sobre patrones de uso, horas pico de alquiler y estaciones populares.

Hipótesis

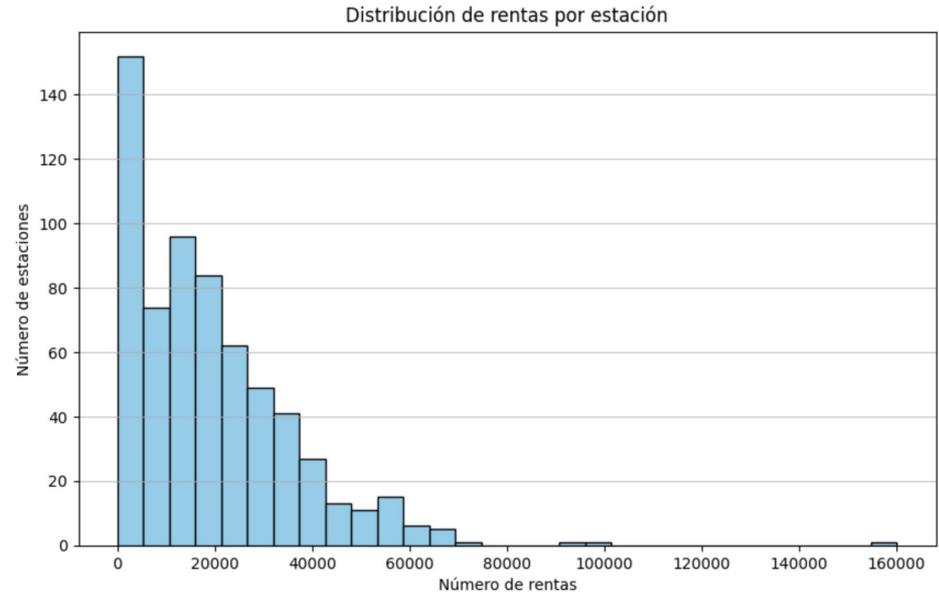
- 1. **Variación del patrón de uso:** existe una variación significativa en los patrones de alquiler de bicicletas en diferentes momentos del día y días de la semana.
- 2. **Influencia demográfica:** la demografía de los usuarios influye significativamente en el comportamiento de alquiler de bicicletas.
- 3. **Popularidad de la estación:** ciertas estaciones son más populares que otras, con patrones de uso influenciados por la ubicación, la hora del día y el día de la semana.

Distribución de edad y sexo



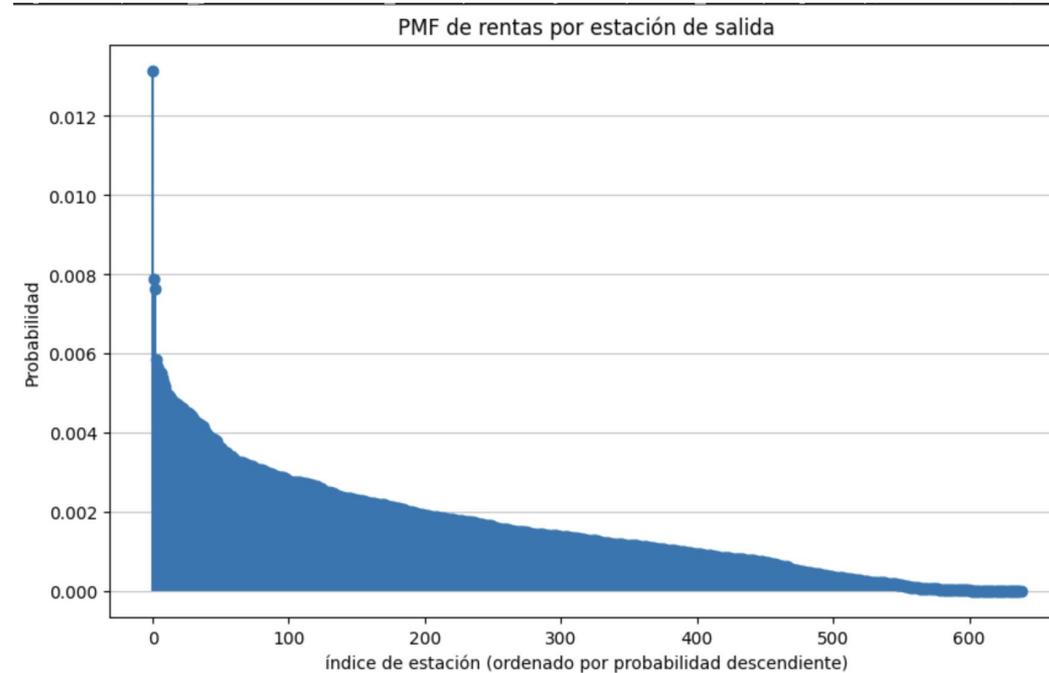
Análisis (Entropía de Shannon en rentas estación de salida)

- La entropía de Shannon de **8,05** indica un nivel moderado de diversidad e imprevisibilidad en la distribución del alquiler de bicicletas en las estaciones de salida.



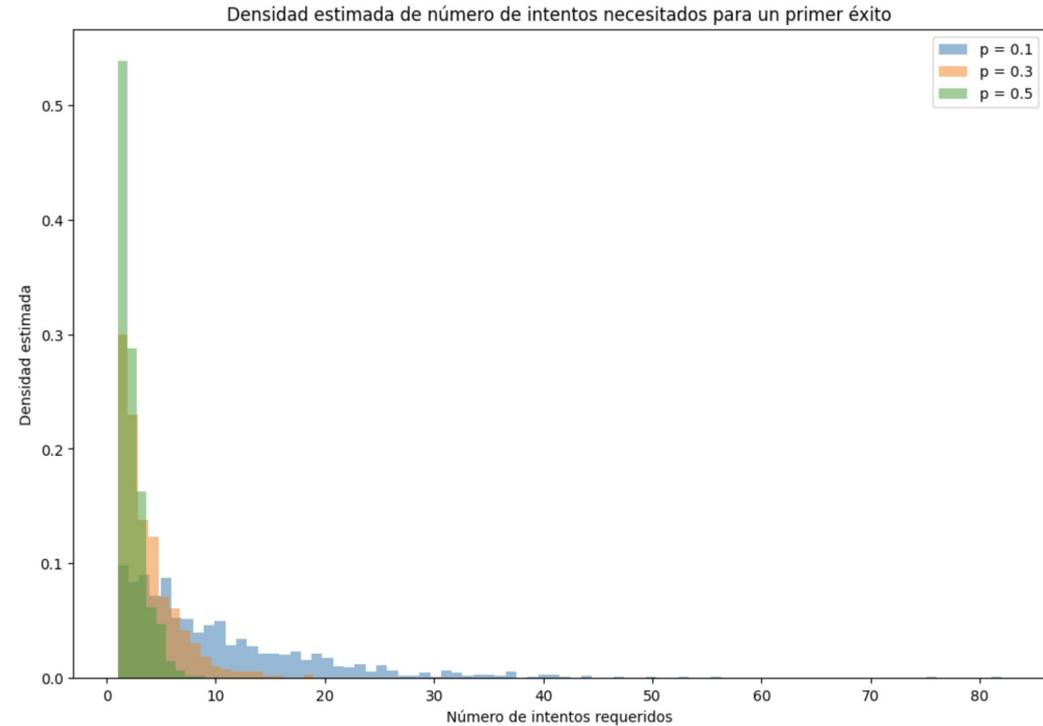
Análisis (Entropía de Shannon en rentas estación de salida)

La caída al principio indica que un pequeño número de estaciones tienen probabilidades de alquiler significativamente mayores que otras, lo que disminuye rápidamente. Esto ilustra aún más la distribución desigual de las actividades de alquiler entre las estaciones.



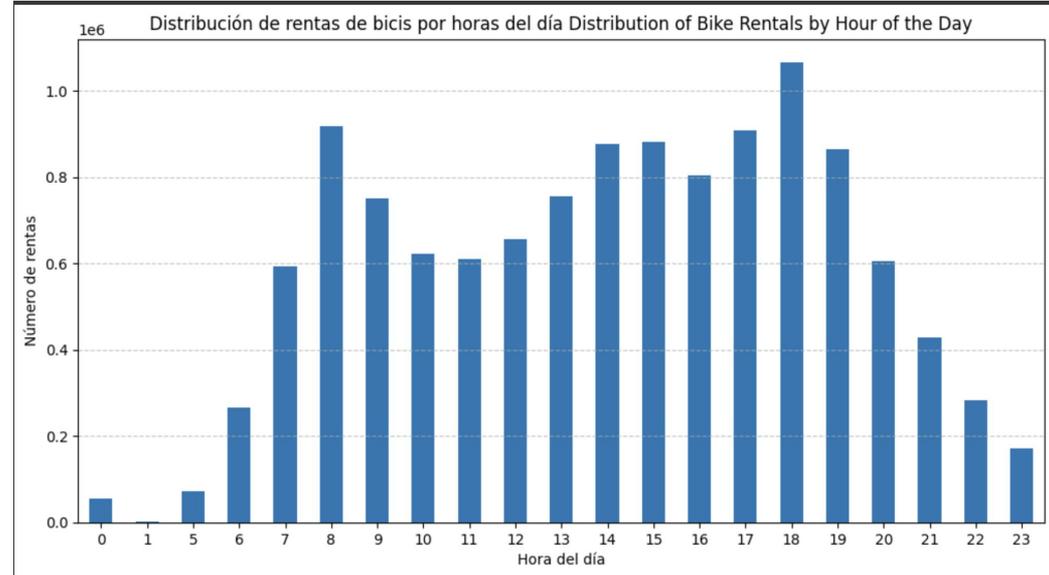
Análisis (Simulación y Estimación)

Simulamos lanzamientos de monedas para modelar el proceso de alquiler de bicicletas, estimando la distribución del número de intentos necesarios para alquilar una bicicleta con éxito. Esta simulación ayudó a comprender los patrones de demanda y optimizar la distribución de bicicletas.



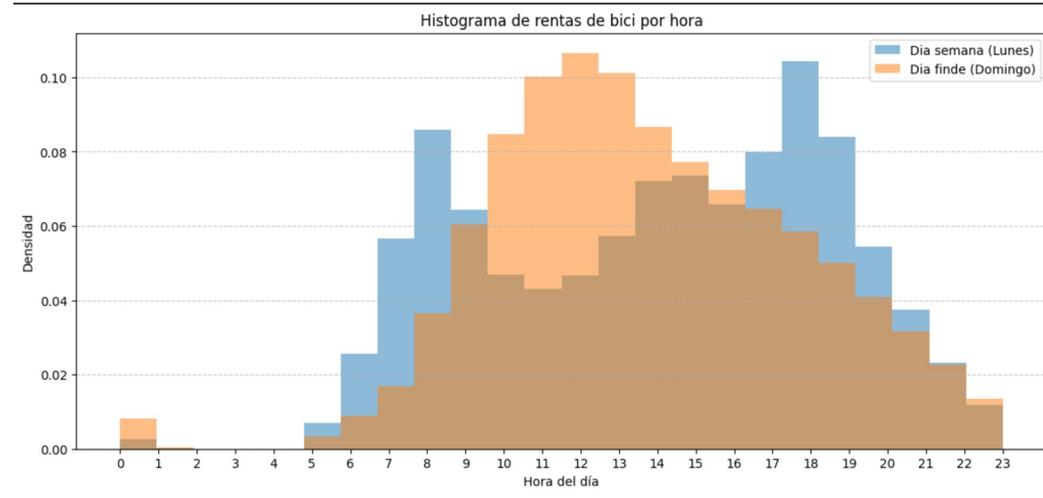
Análisis (Densidad y Distribución)

Para comparar la densidad y la distribución, analizamos la distribución del alquiler de bicicletas a lo largo del tiempo, analizando específicamente el número de alquileres por hora del día. Esto nos ayudará a comprender las horas pico de alquiler y cómo varía la distribución de los alquileres a lo largo del día.



Análisis (Test de Kolmogorov-Smirnov)

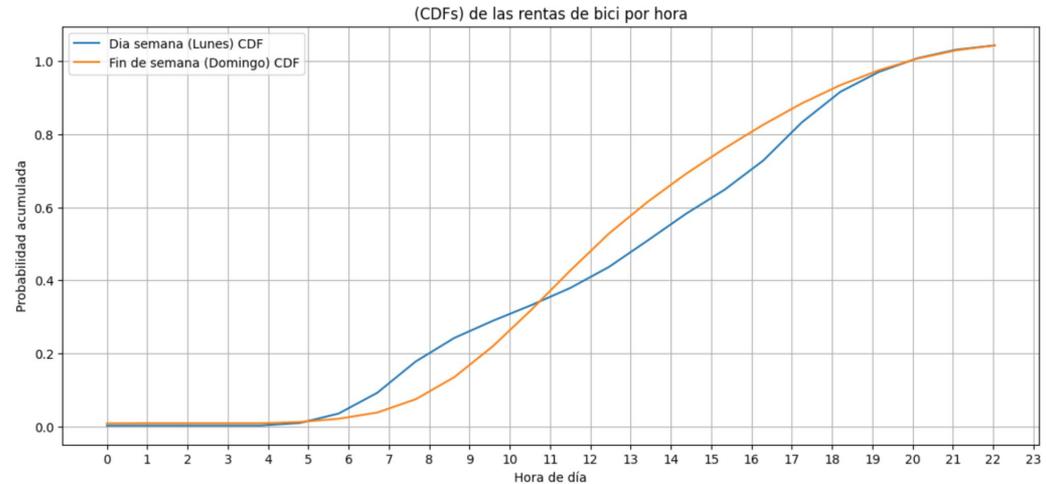
Para las pruebas estadísticas, específicamente utilizando la prueba de Kolmogorov-Smirnov (KS), podemos comparar las distribuciones de alquiler de bicicletas entre dos días diferentes de la semana para ver si hay una diferencia estadísticamente significativa en los patrones de alquiler.



Análisis (Test de Kolmogorov-Smirnov)

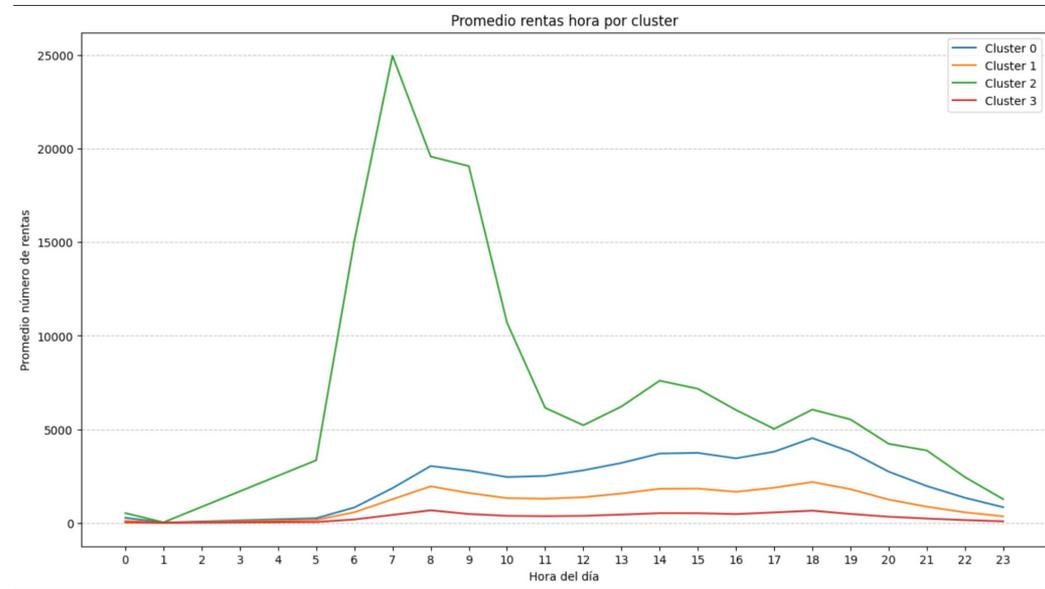
- KS 0.13
- Valor-p 0

Los resultados sugieren que la distribución del alquiler de bicicletas a lo largo del día es significativamente diferente entre un día laborable típico y un día de fin de semana, lo que refleja diferentes patrones de uso, posiblemente debido a las diferentes rutinas y actividades entre semana y fines de semana.



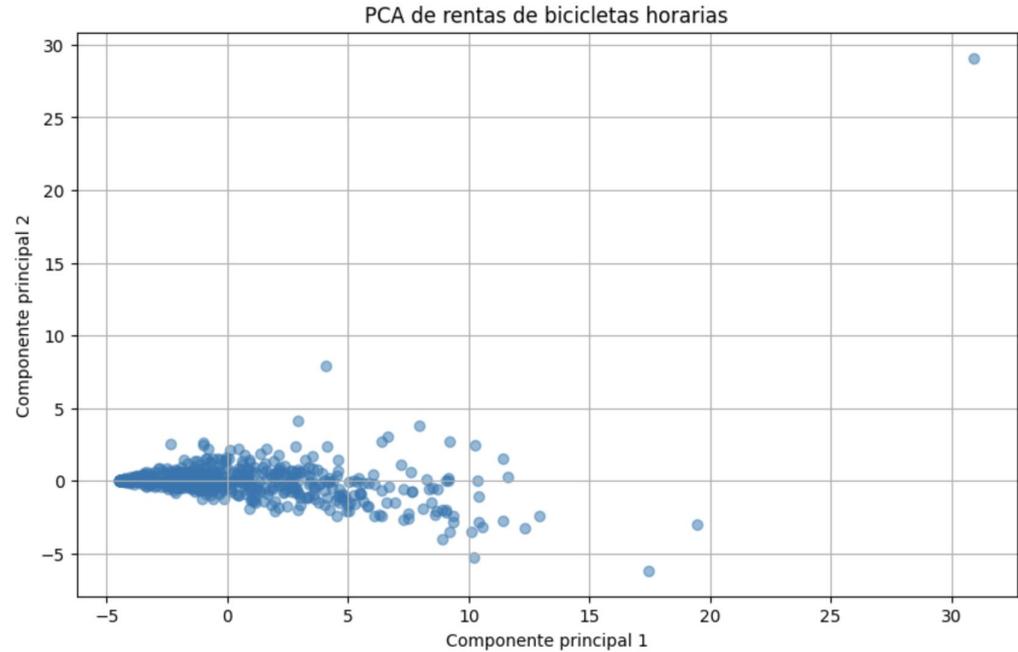
Análisis (K-means)

- Agrupamos por estaciones y horas, y su número de rentas.
- Con el método del codo encontramos 4 clusters.



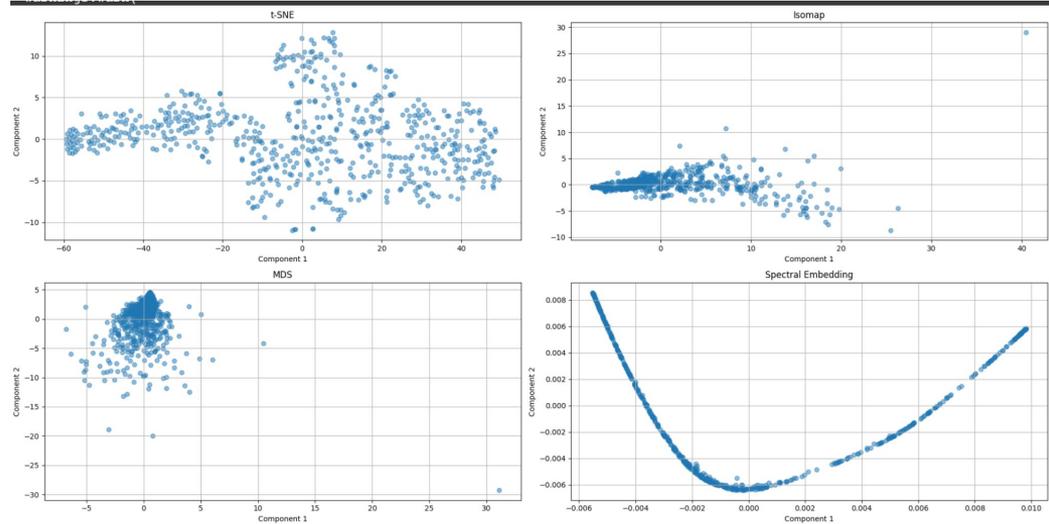
Análisis (PCA)

- Usamos misma base de datos que la anterior y agrupamos en dos componentes.



Análisis (Reducción de dimensionalidad)

- t-SNE, Isomap, MDS, Spectral embedding



Conclusiones

- Existe una variación significativa en los patrones de alquiler de bicicletas en diferentes momentos del día y días de la semana.
- La demografía de los usuarios influye significativamente en el comportamiento de alquiler de bicicletas.
- Ciertas estaciones son más populares que otras, con patrones de uso influenciados por la ubicación, la hora del día y el día de la semana.