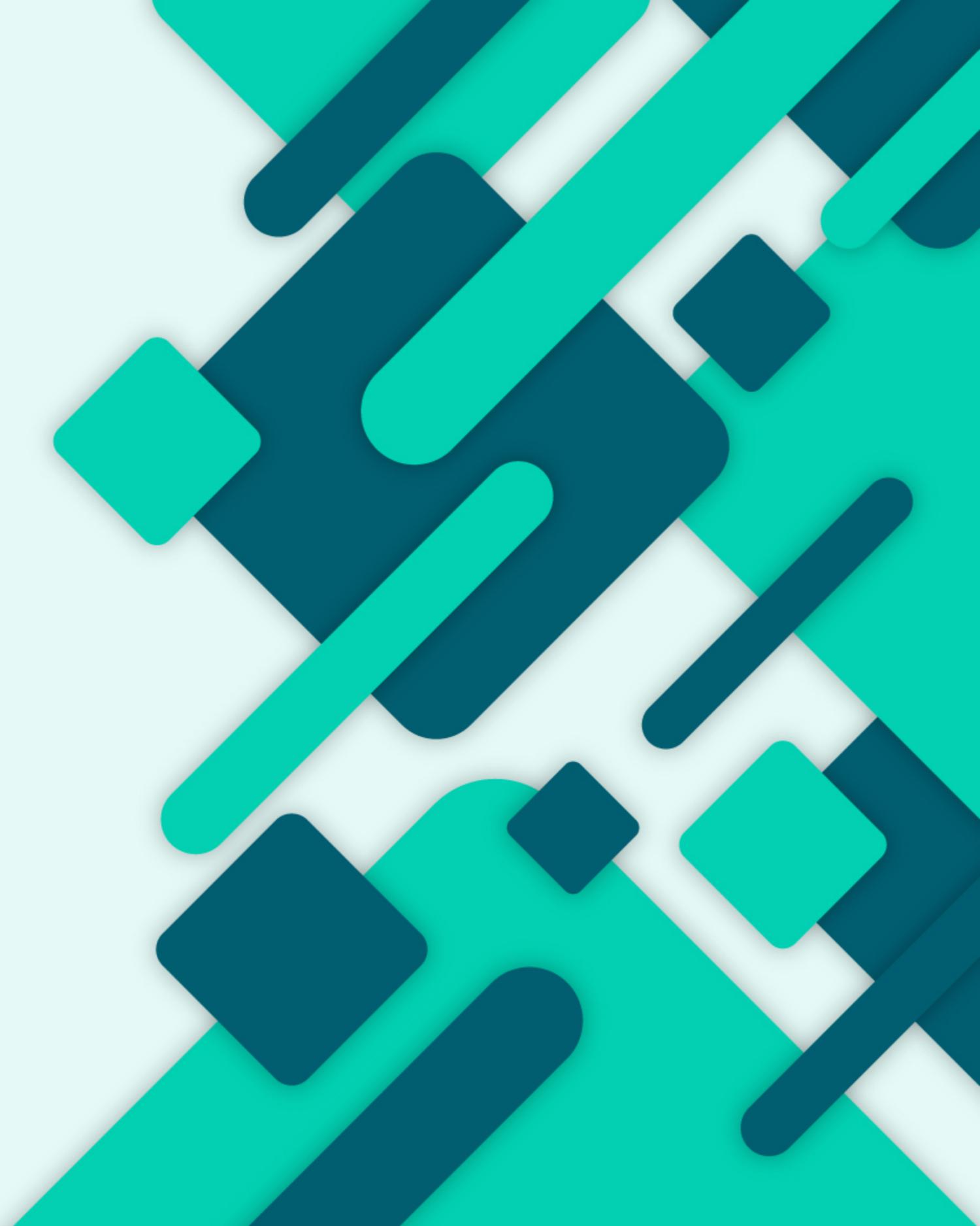


Análisis Ecobici

Majo Gil
Joshua Chicoj



ECOBICI

Datos recolectados

Podemos encontrar 2 conjuntos de datos, los obtenidos en tiempo real y los históricos. En este caso utilizaremos los datos históricos del 2023

DATOS

- Genero del usuario
- Edad del usuario
- Fecha de Retiro
- Fecha de Arribo
- Hora de Retiro
- Hora de Arribo
- Ciclo Estación Arribo
- Ciclo Estación Retiro

Ademas, tenemos las ubicaciones de las ciclo estaciones, por coordenadas y sus nombres



ANÁLISIS EXPLORATORIO

CLASIFICACIÓN DE VARIABLES

CUALITATIVAS

- Genero Usuario
- Bici
- CicloEstación Retiro
- Hora Retiro
- Fecha Retiro
- CicloEstación Arribo
- Fecha Arribo
- Hora Arribo

CUANTITATIVAS

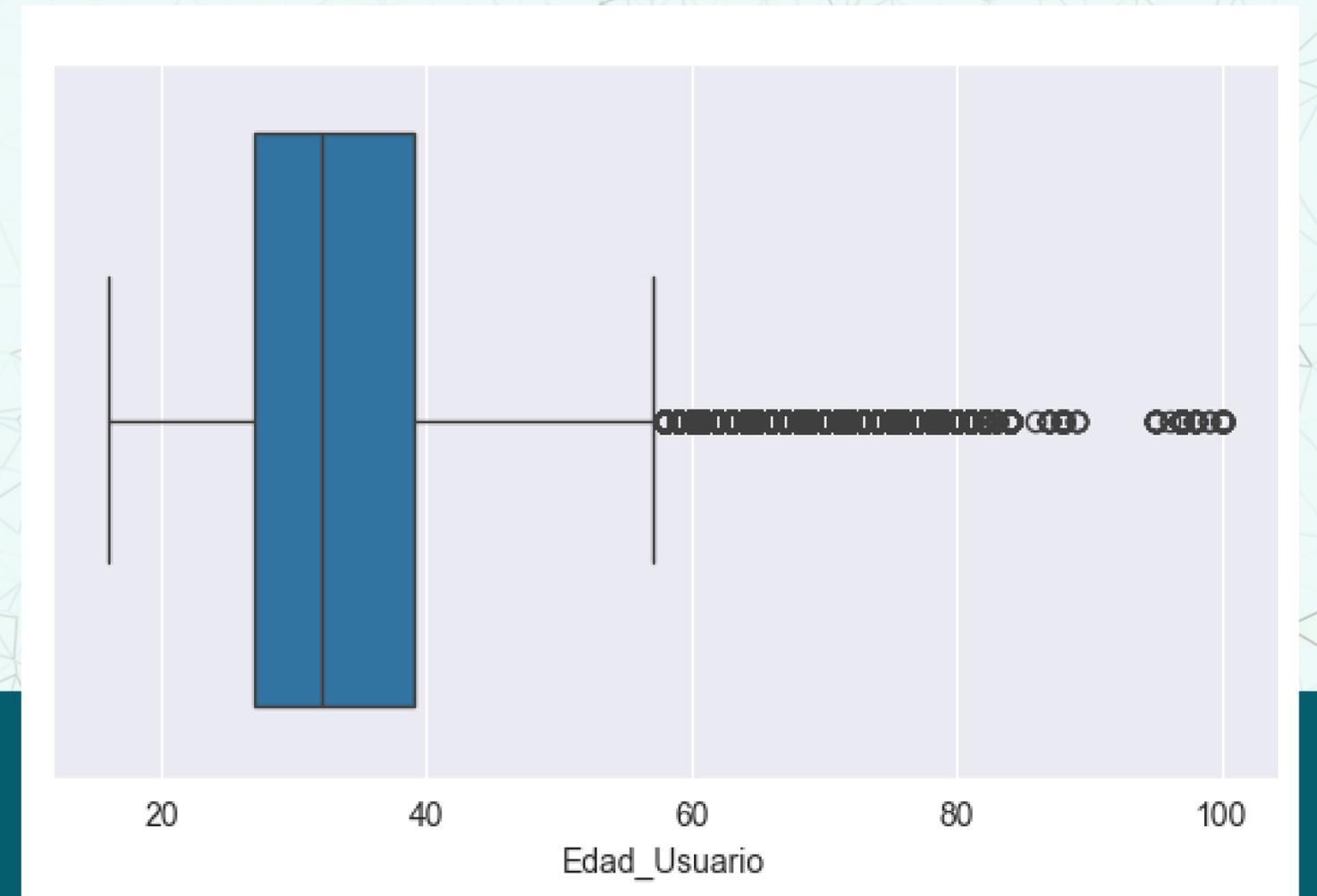
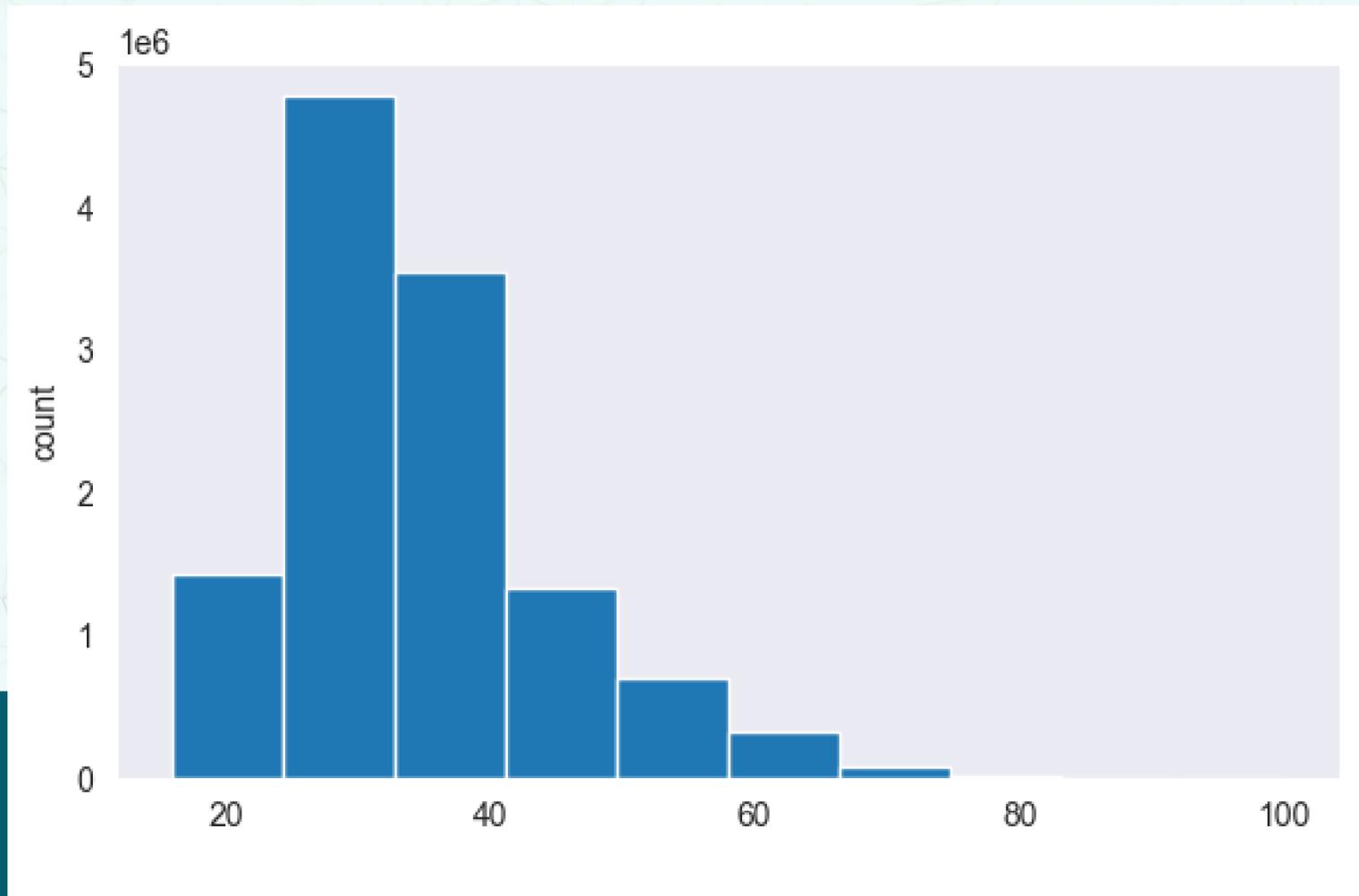
- Edad Usuario

DATOS DE LAS ESTACIONES

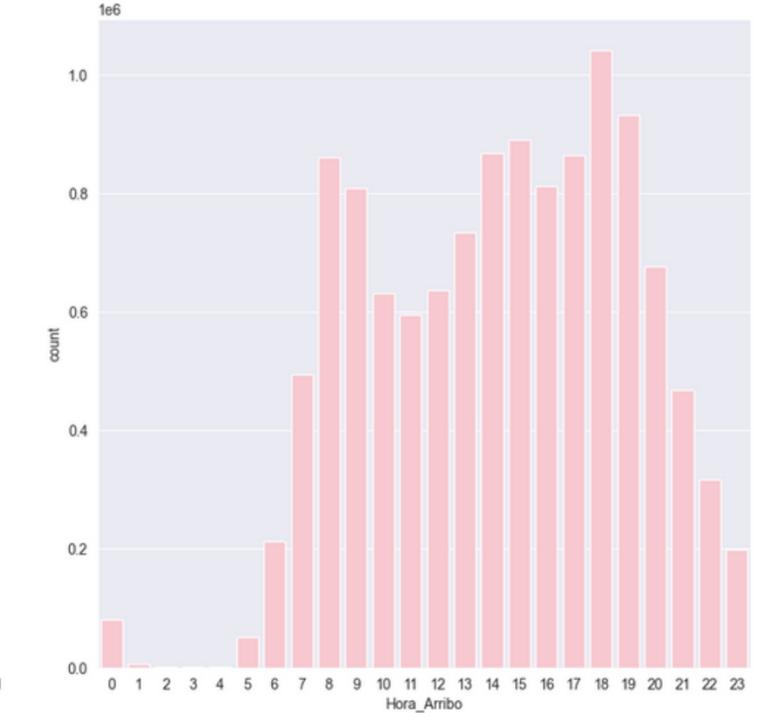
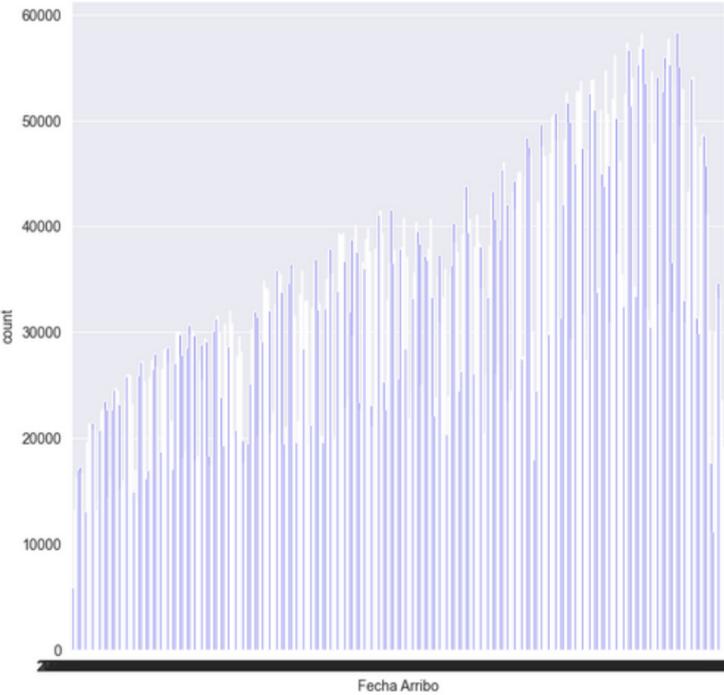
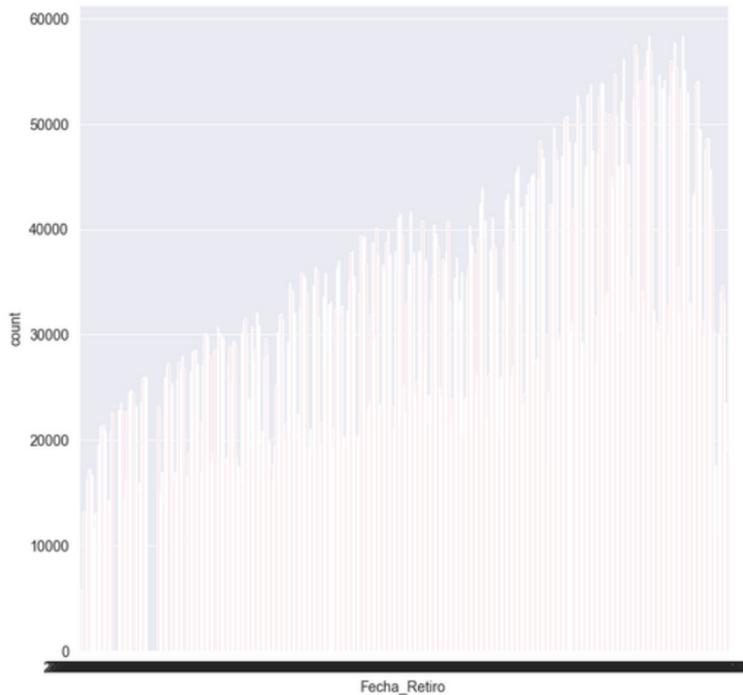
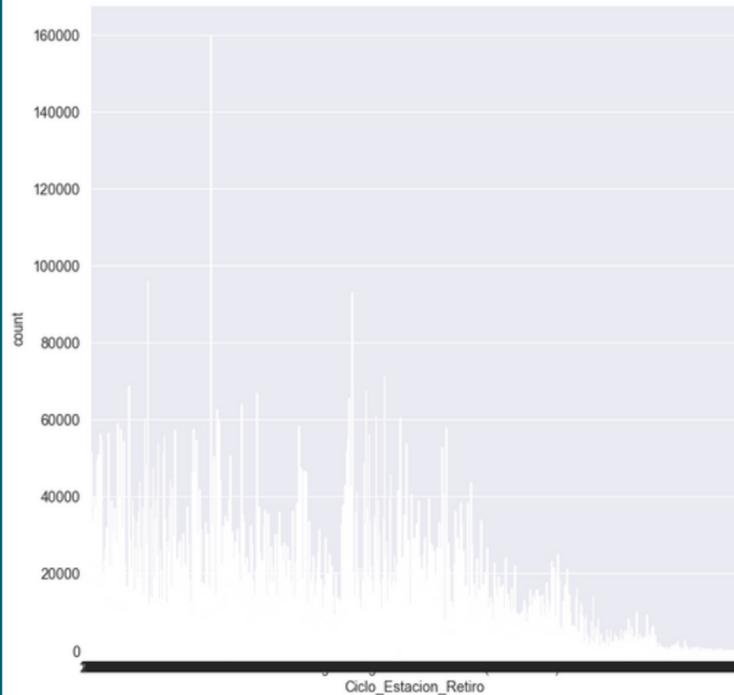
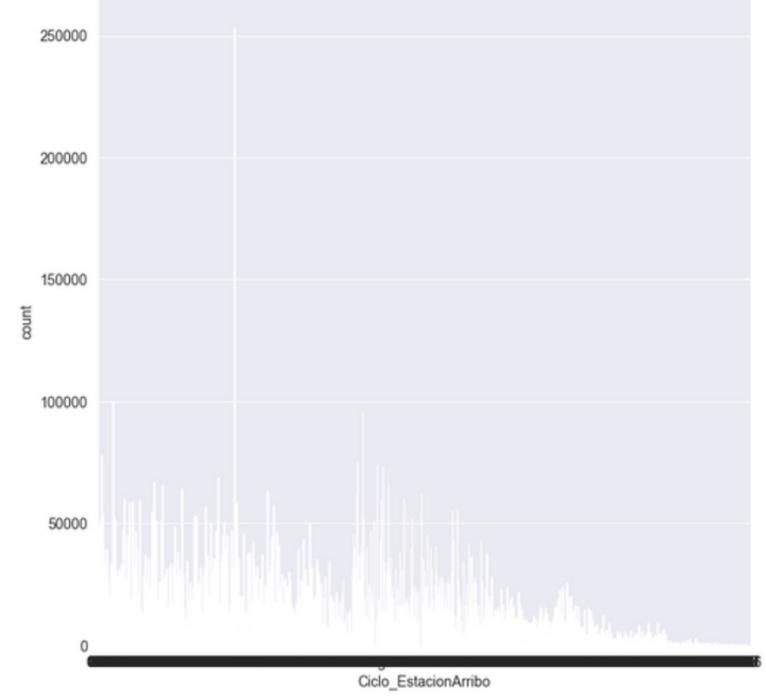
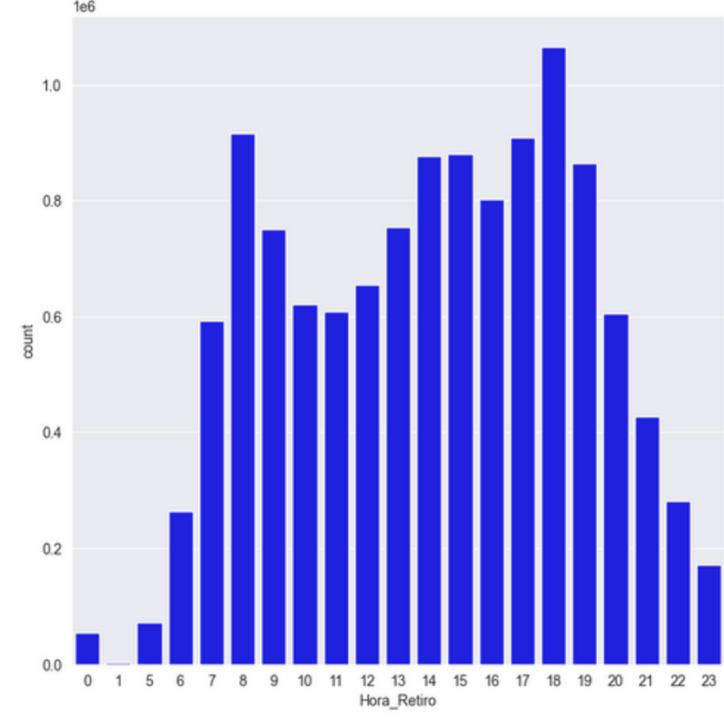
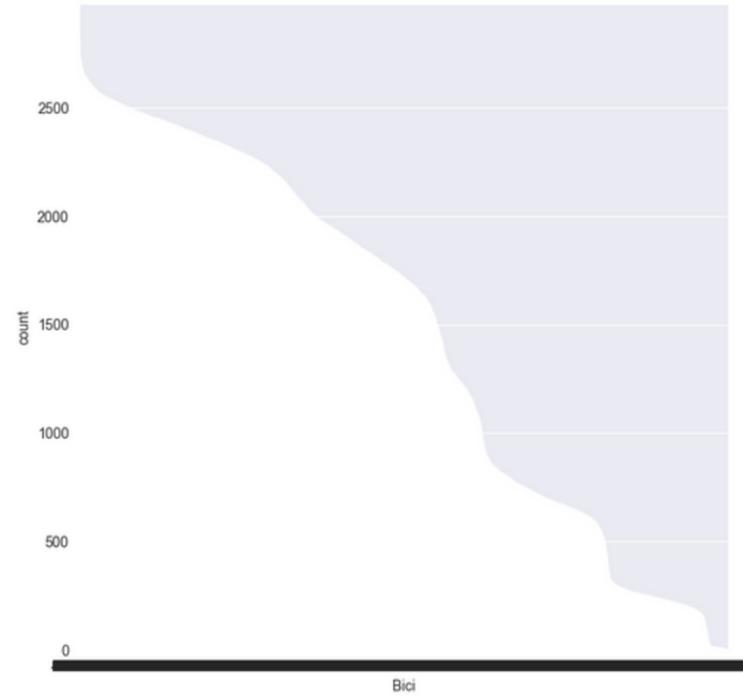
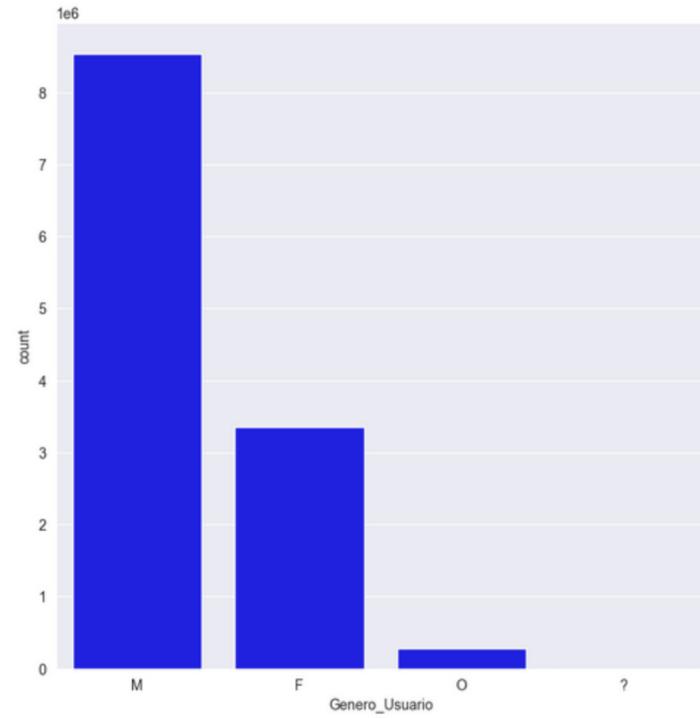
- Latitud
- Longitud
- Nombre
- Código

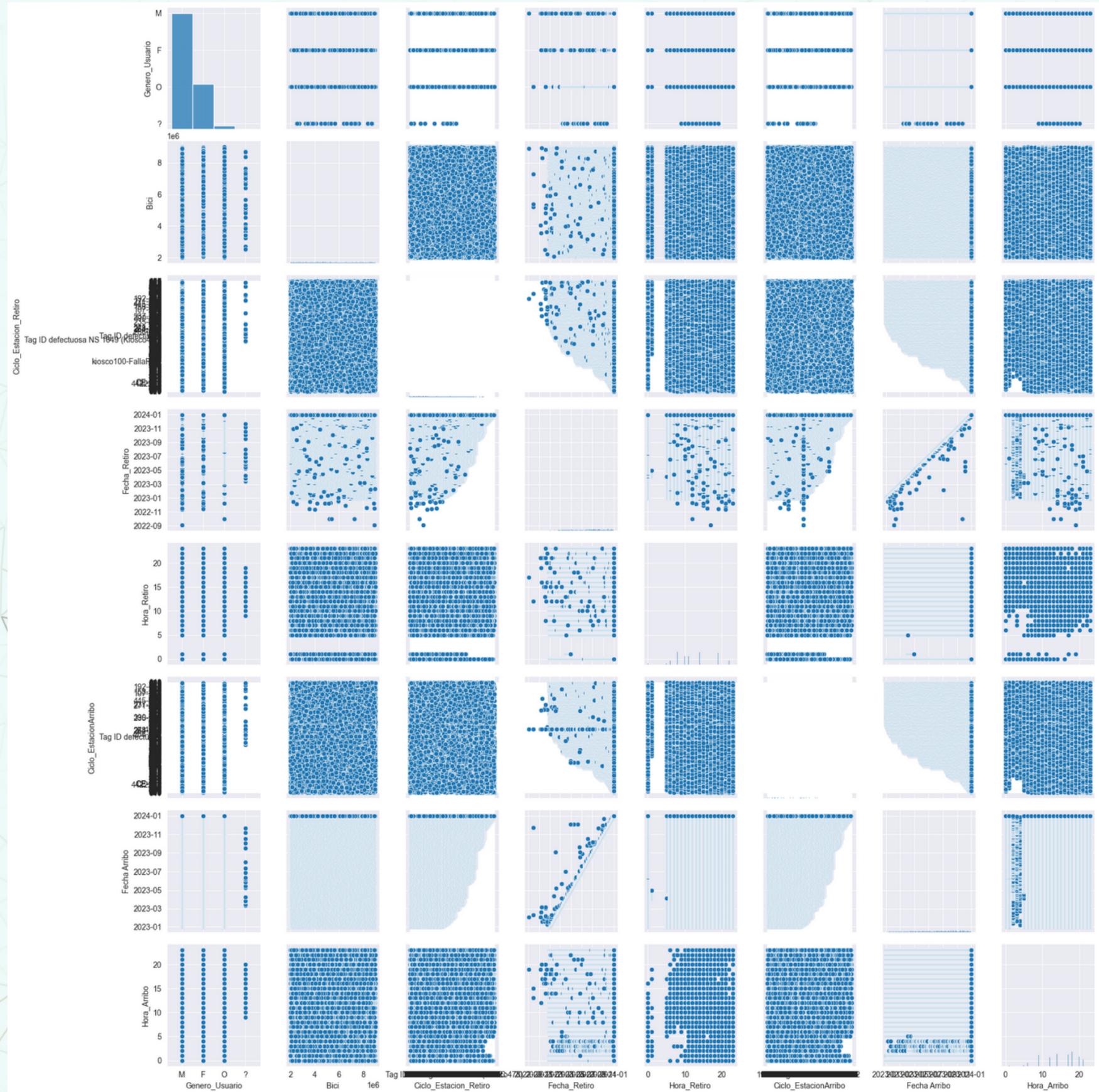
	count	unique	top	freq	mean	min	25%	50%	75%	max	std
Genero_Usuario	12177337	4	M	8542606	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Edad_Usuario	12177337.0	NaN	NaN	NaN	34.439772	16.0	27.0	32.0	39.0	100.0	9.829181
Bici	12177337.0	NaN	NaN	NaN	5471026.853...	2000776.0	3746518.0	5445259.0	7213283.0	8999708.0	2026707.513...
Ciclo_Estacion_Retiro	12177337	637	271-272	159767	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Fecha_Retiro	12177337	NaN	NaN	NaN	2023-07-26 ...	2022-09-03 ...	2023-05-06...	2023-08-08...	2023-10-23 ...	2023-12-31...	NaN
Hora_Retiro	12177337.0	NaN	NaN	NaN	14.073498	0.0	10.0	14.0	18.0	23.0	4.611318
Ciclo_EstacionArribo	12177337	635	271-272	253349	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Fecha_Arribo	12177337	NaN	NaN	NaN	2023-07-26 ...	2023-01-01 ...	2023-05-06...	2023-08-08...	2023-10-23 ...	2023-12-31...	NaN
Hora_Arribo	12177337.0	NaN	NaN	NaN	14.2652	0.0	10.0	15.0	18.0	23.0	4.651843

ANÁLISIS VARIABLES CUANTITATIVAS

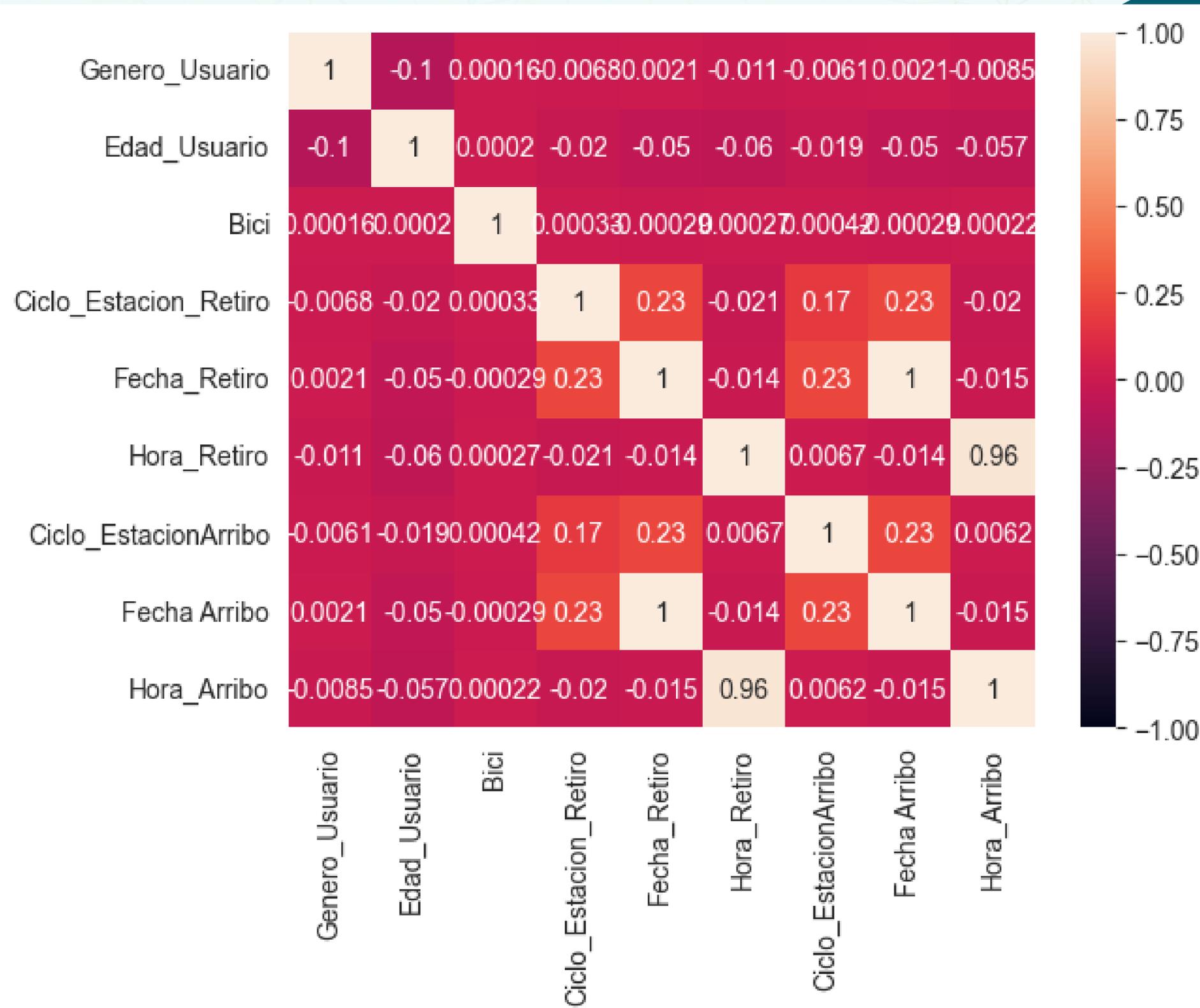


ANÁLISIS VARIABLES CUALITATIVAS





CORRELACIÓN LINEAL



ECOBICI

Preguntas

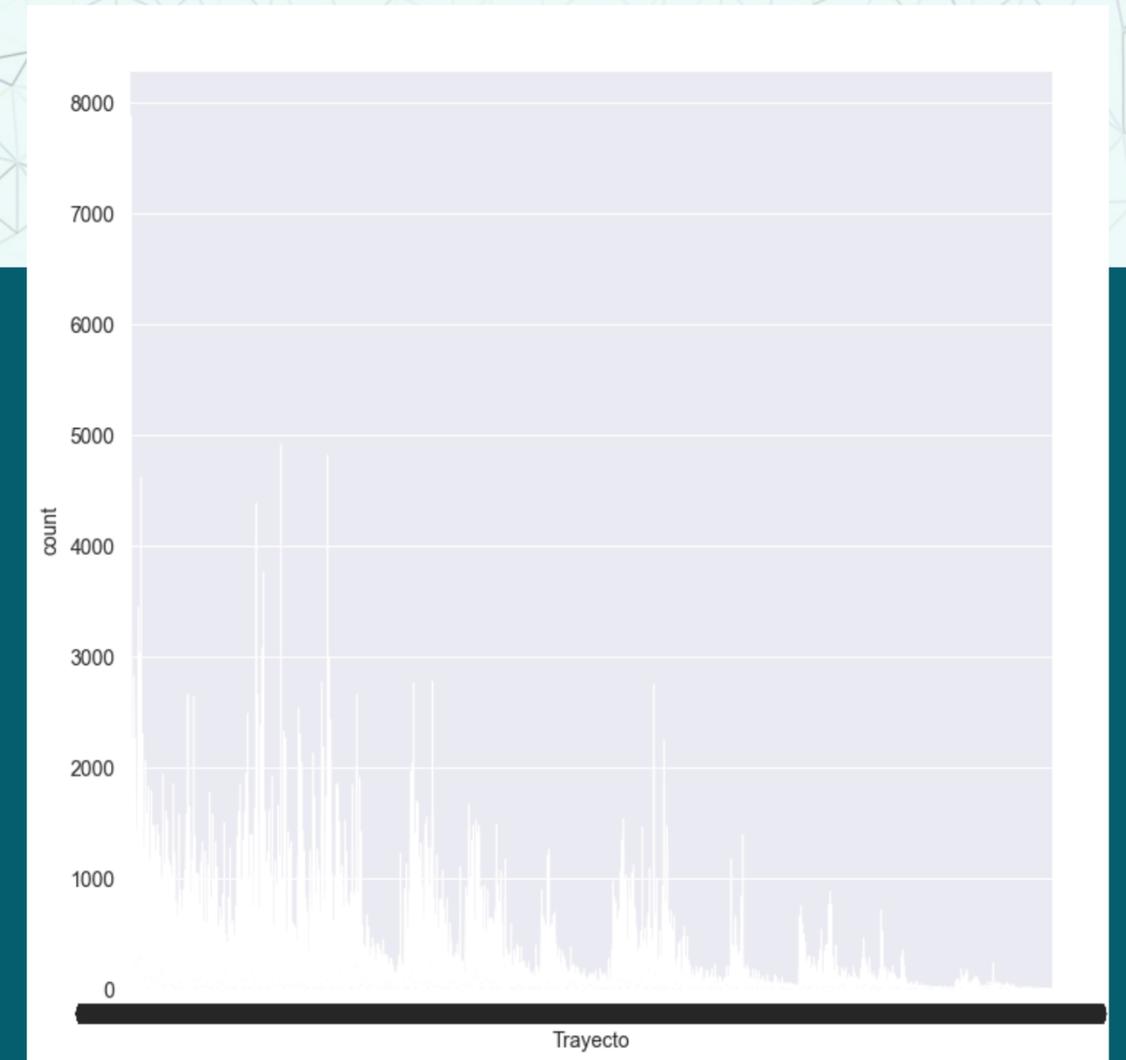
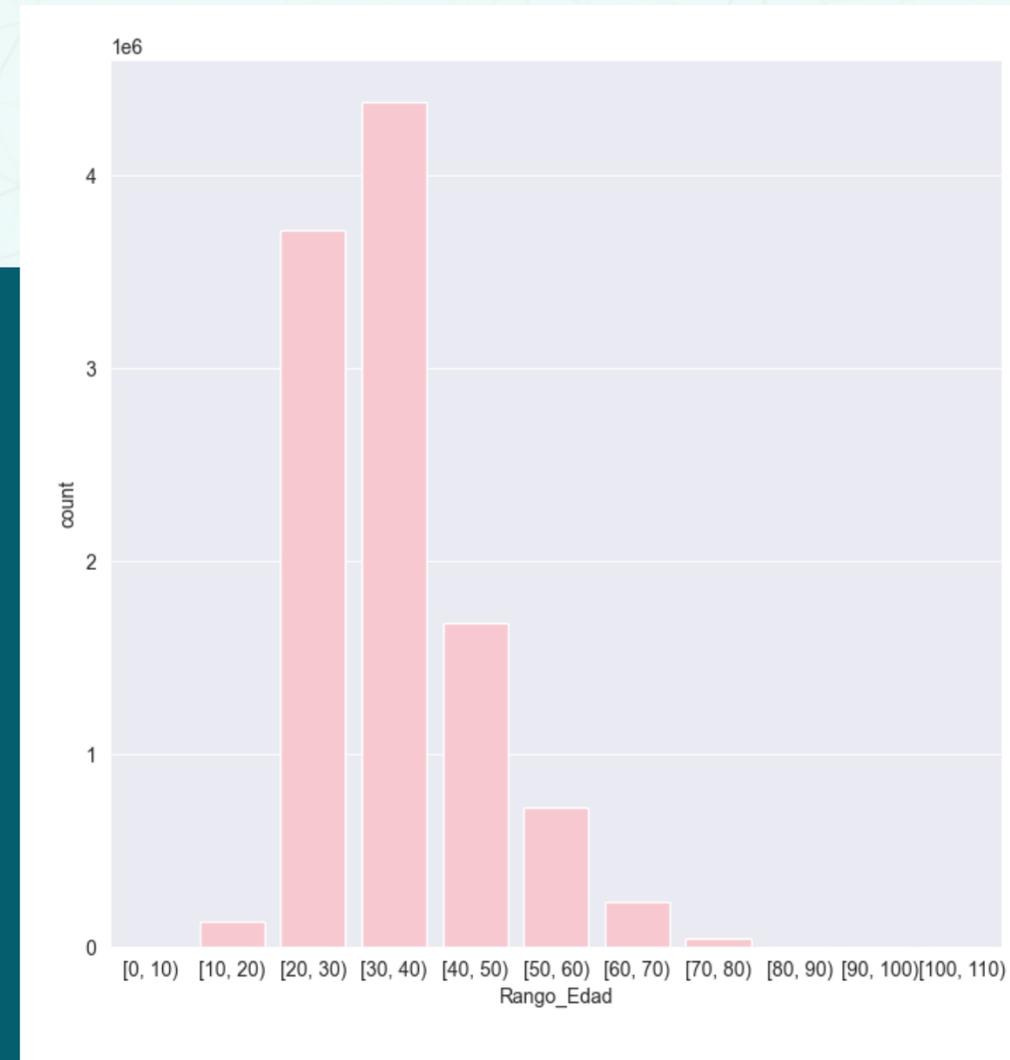
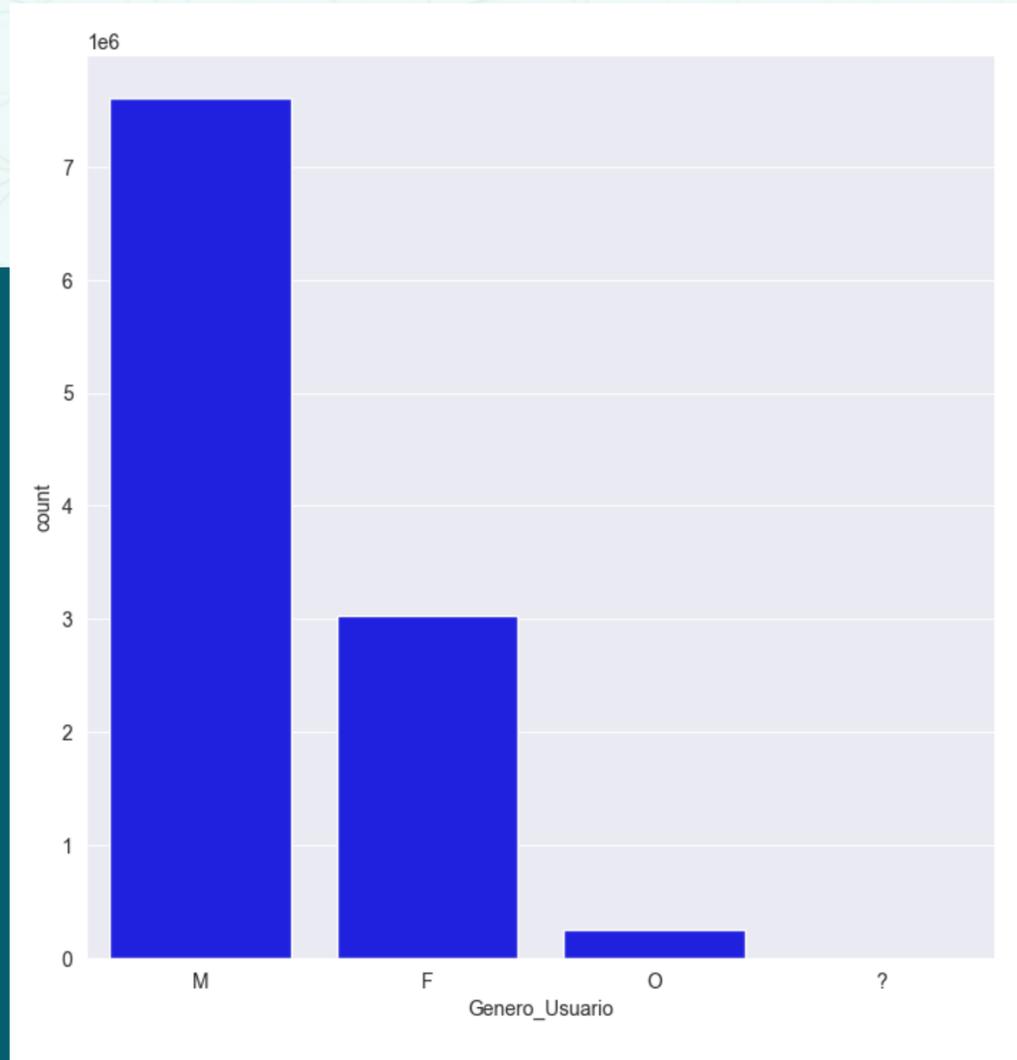
- Existe algún tipo de relación entre la edad, el tiempo y las trayectorias de las bicicletas

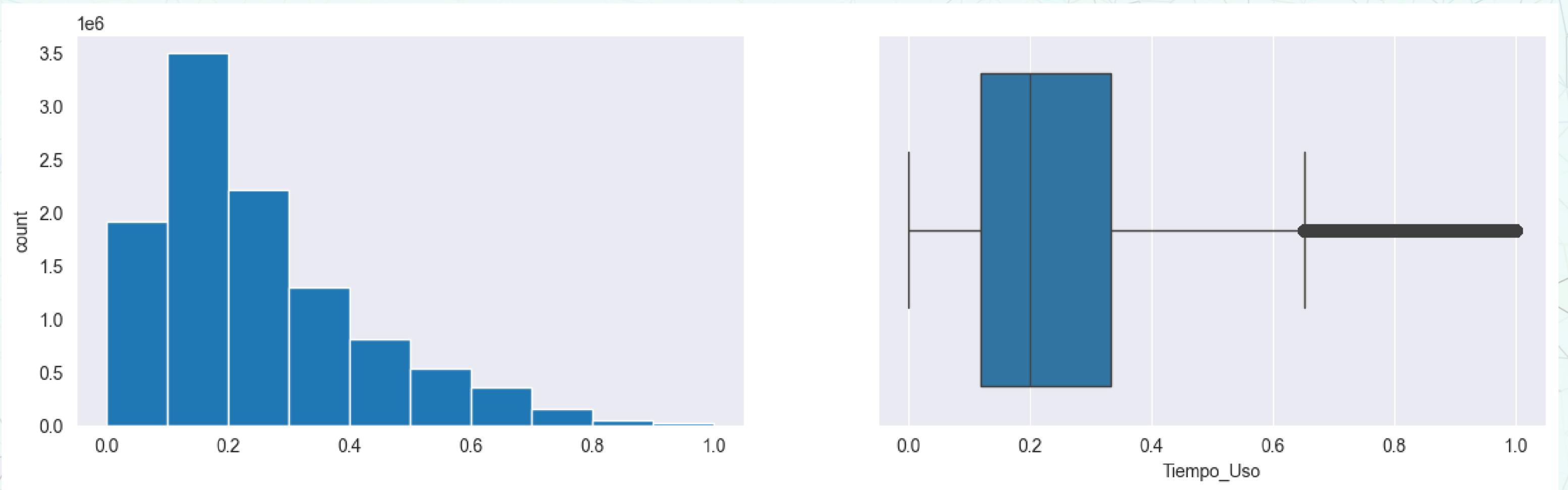
- Si el uso de las bicicletas cambia conforme la época del año
- Qué tanto afectan realmente los datos atípicos



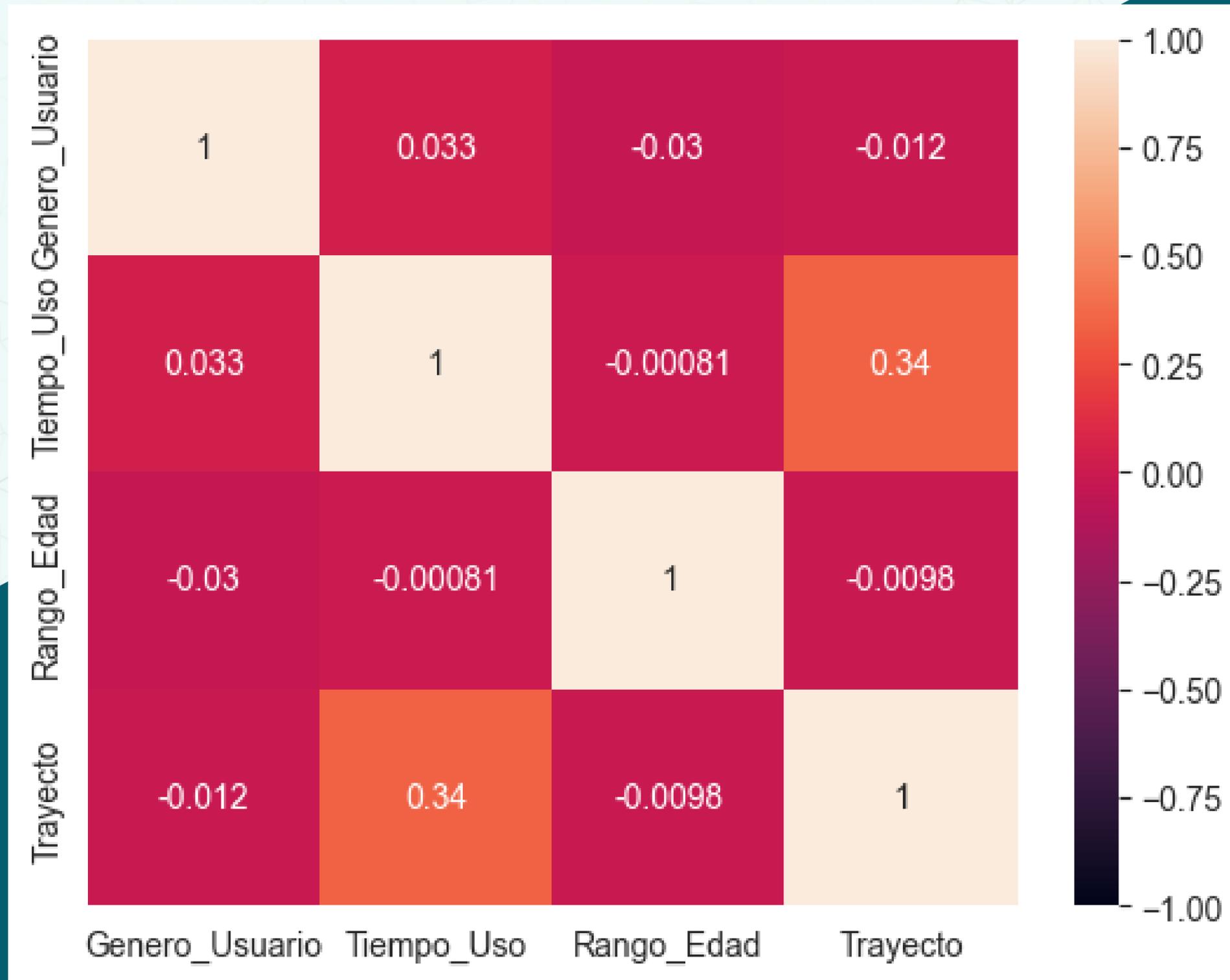
MODIFICACIÓN DE VARIABLES

Decidimos restringir la edad, y modificar los datos a modo de obtener el tiempo de uso y el trayecto utilizado

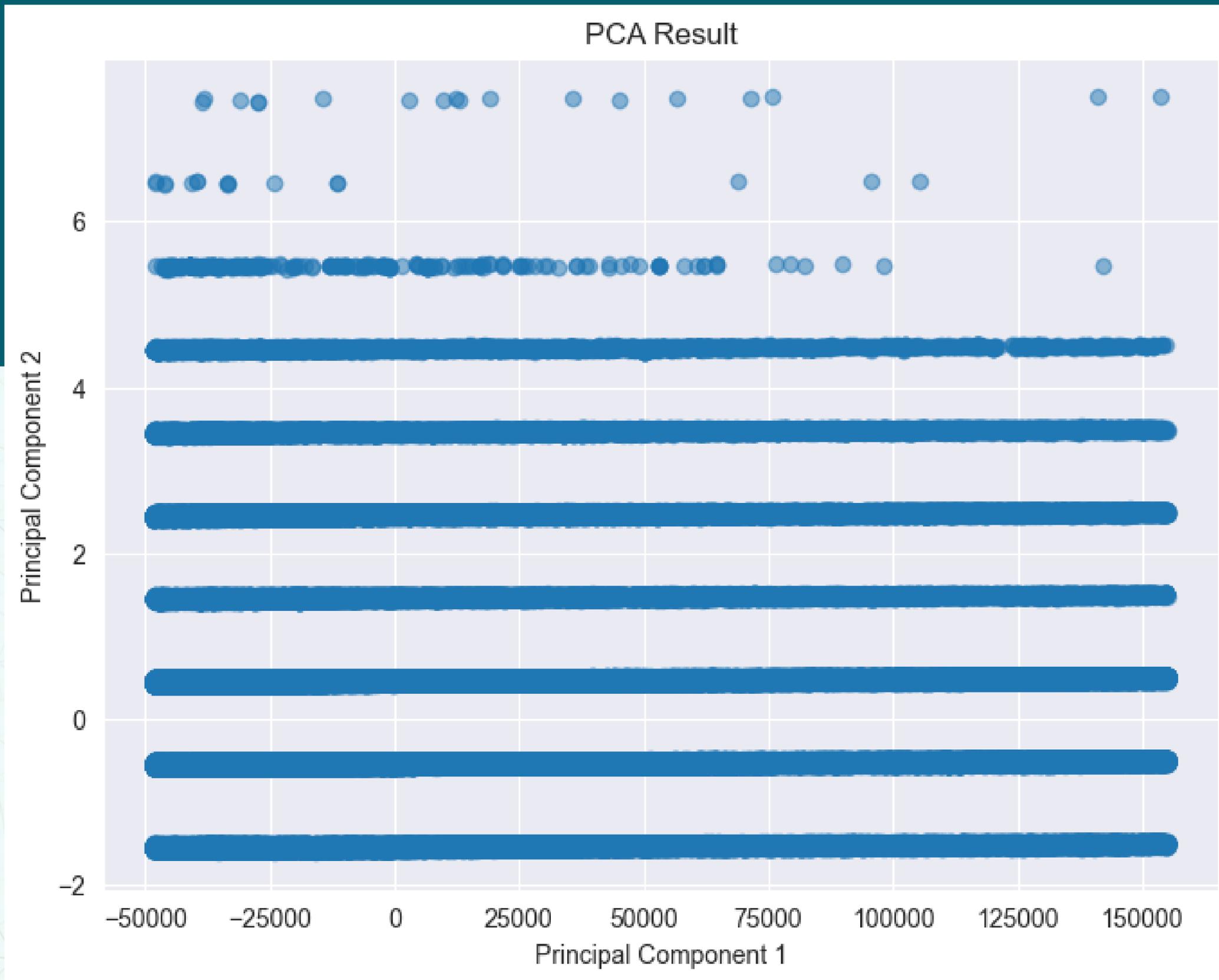




CORRELACIÓN LINEAL

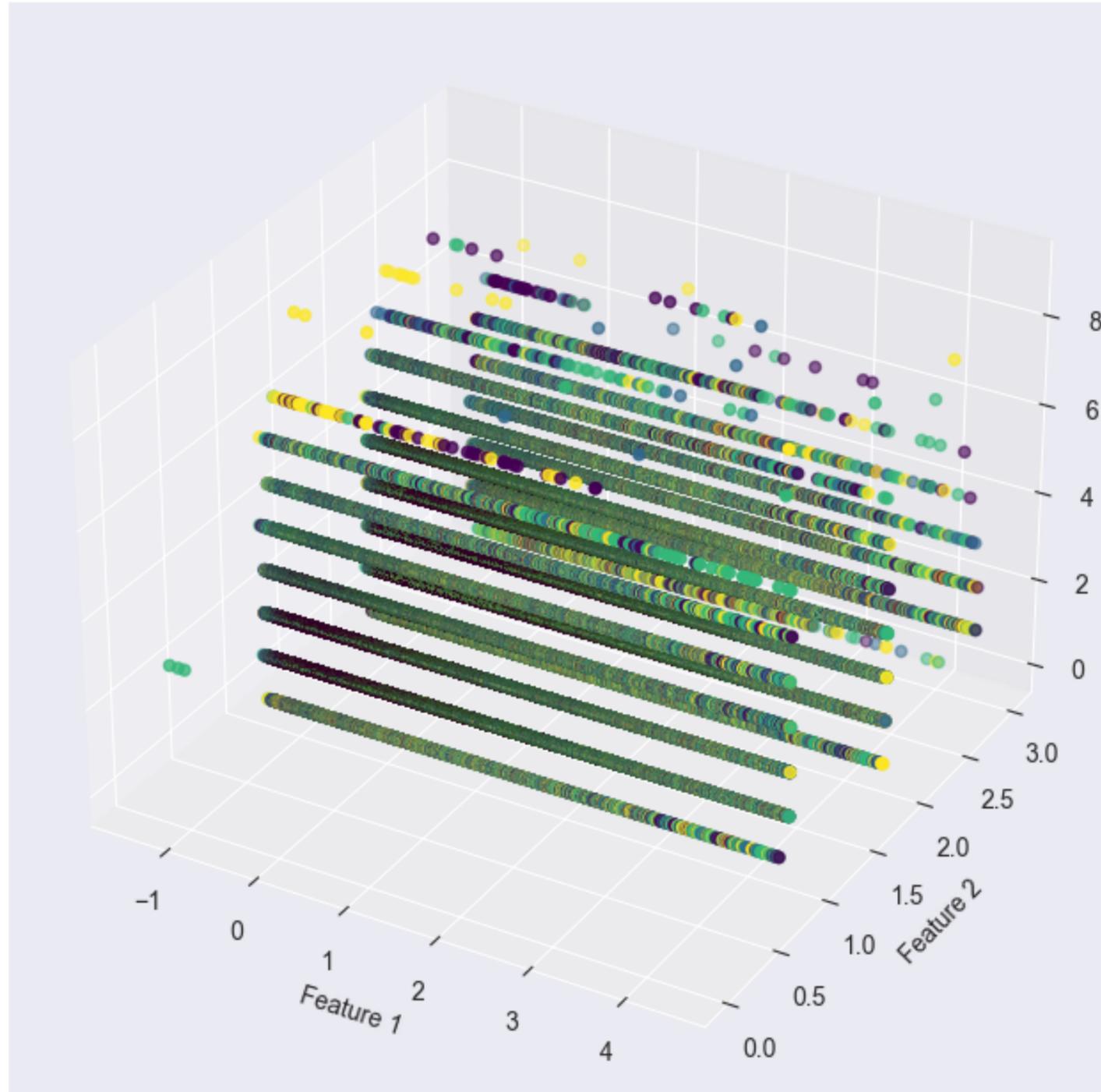


PCA

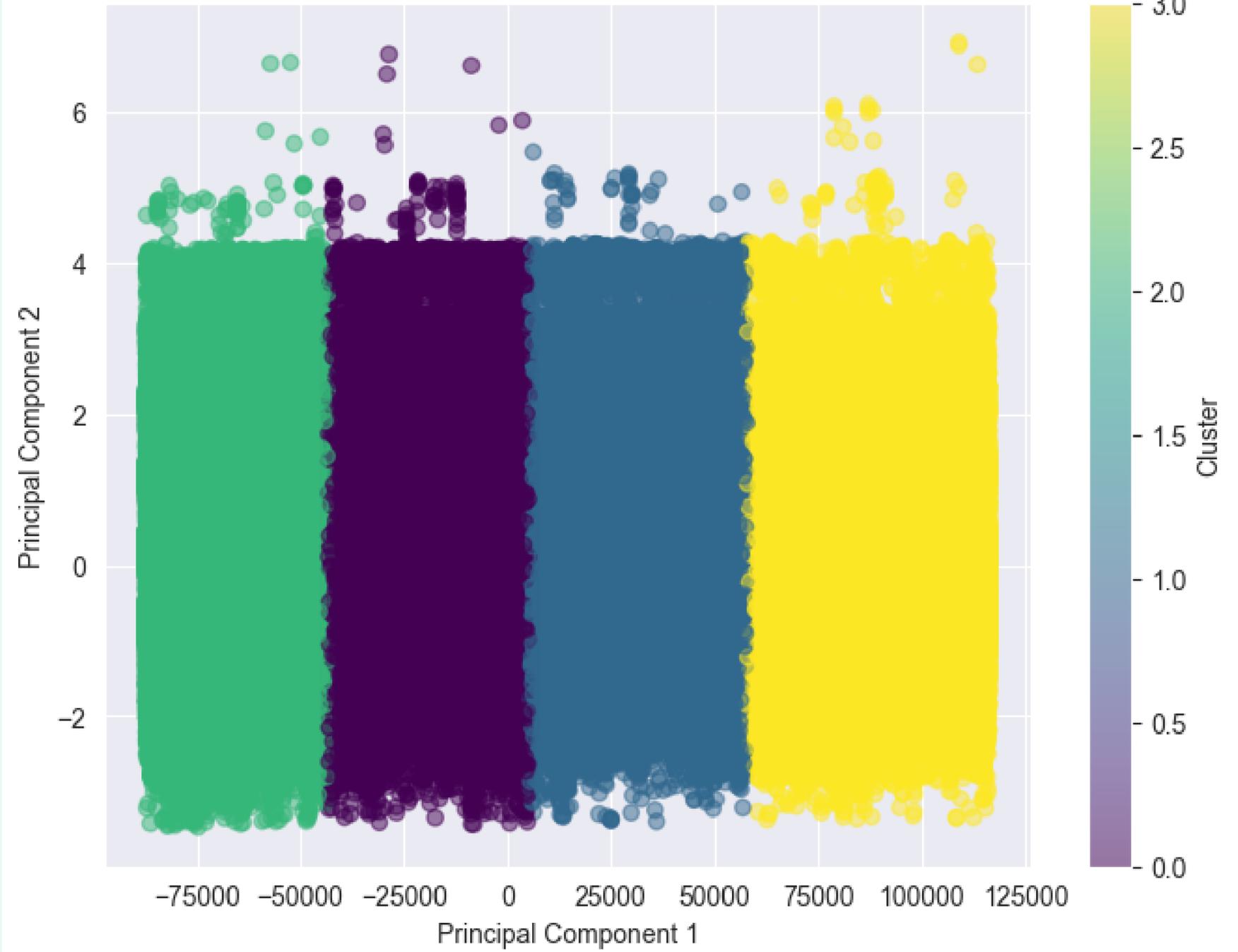


KMEANS

KMeans Clustering - 3D Visualization

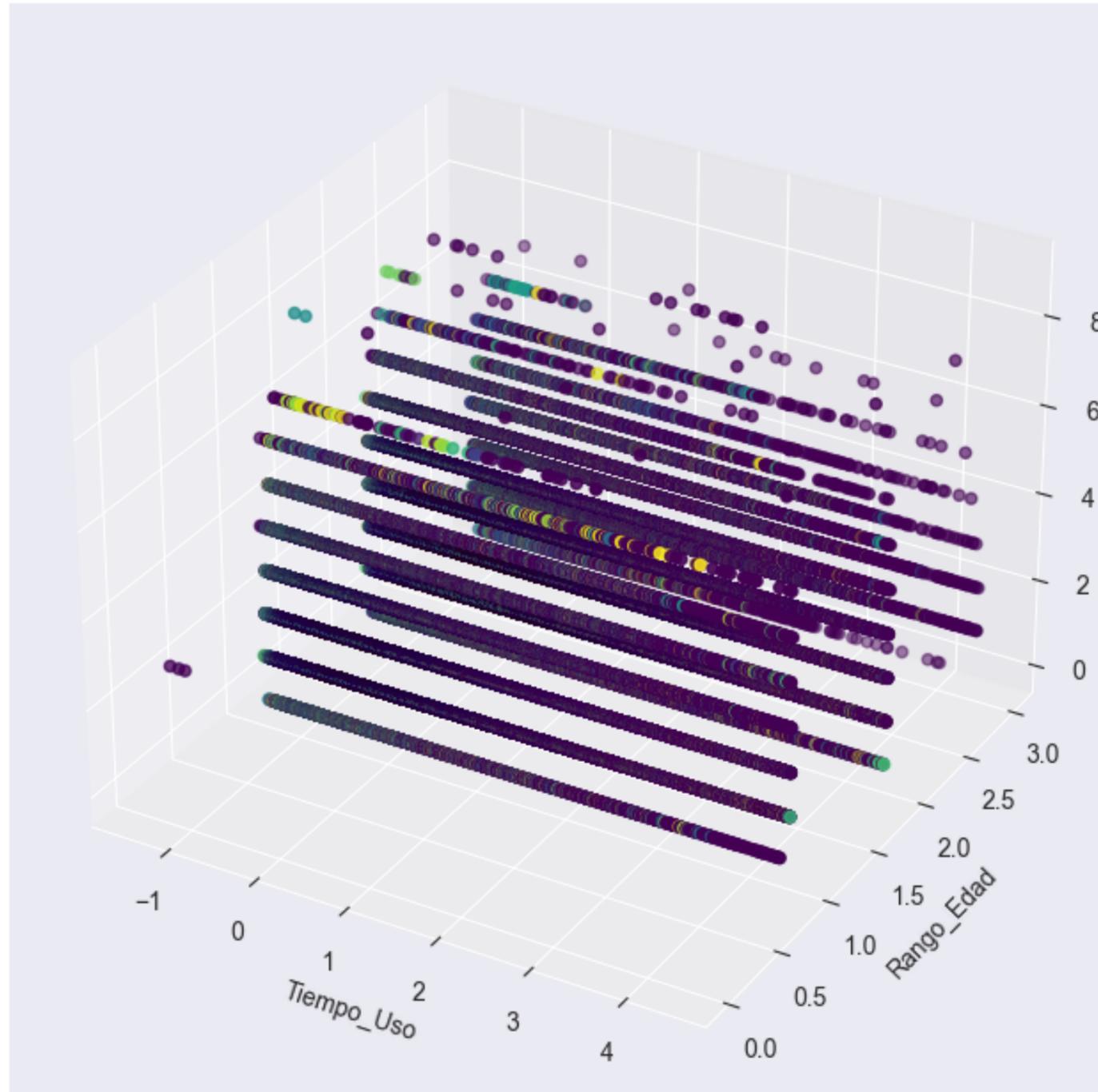


KMeans Clustering - 2D Visualization

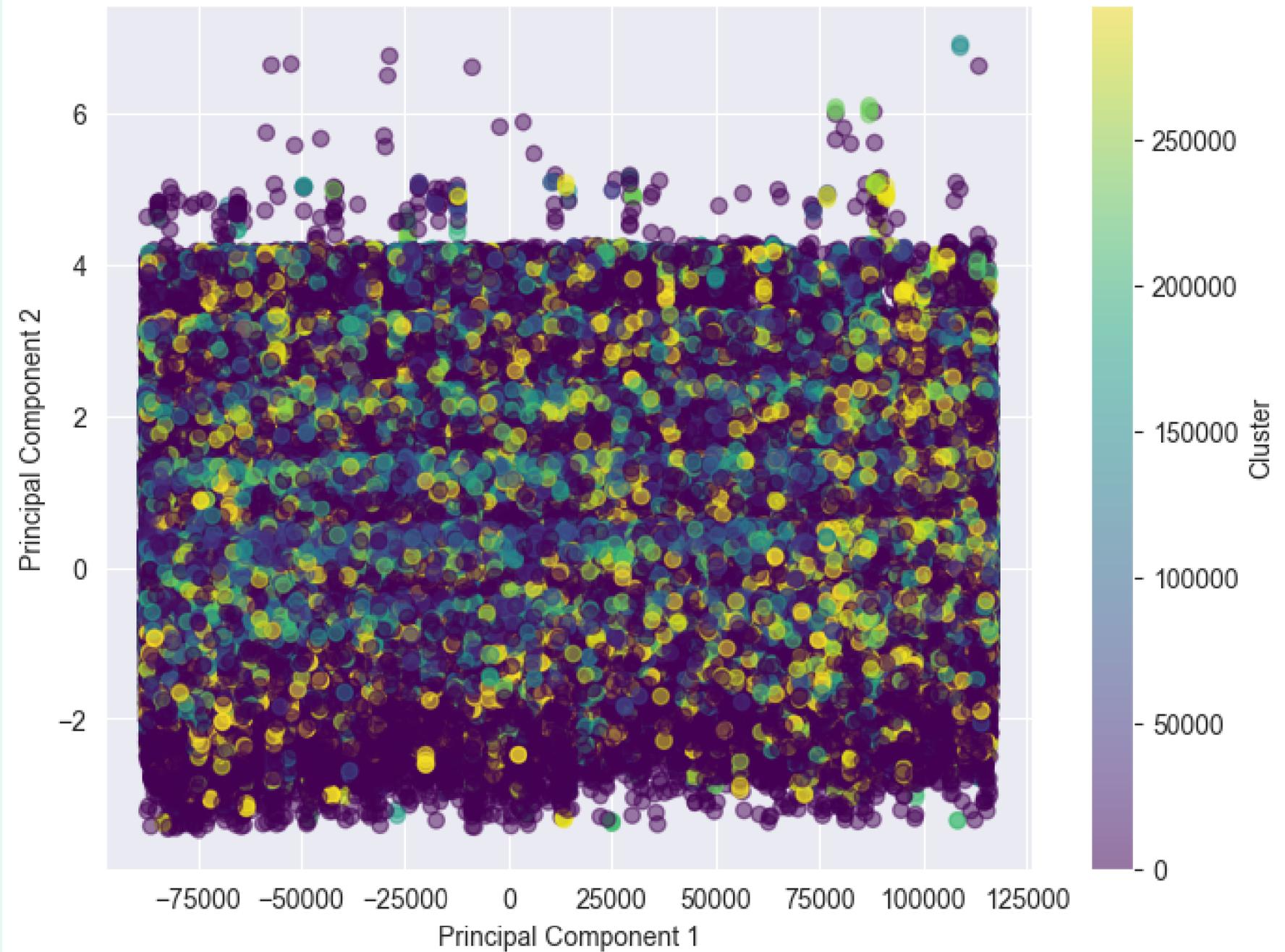


DBSCAN

DBSCAN Clustering - 3D Visualization

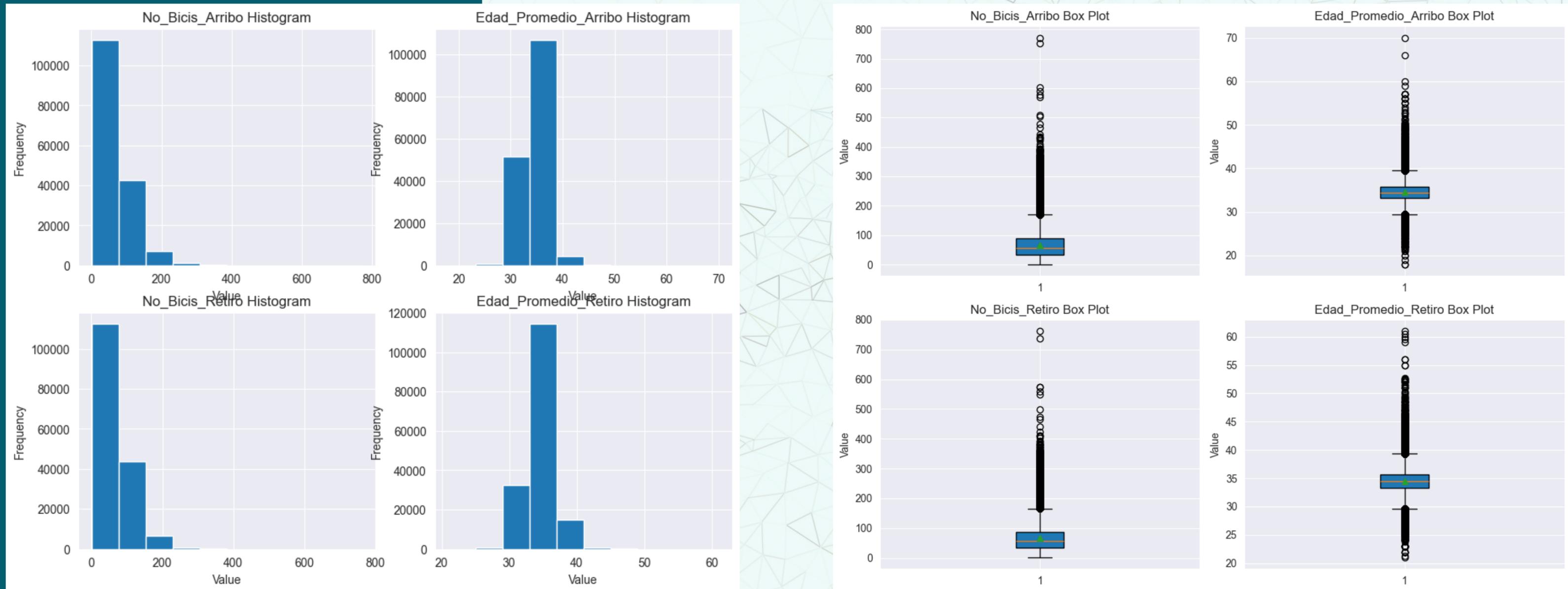


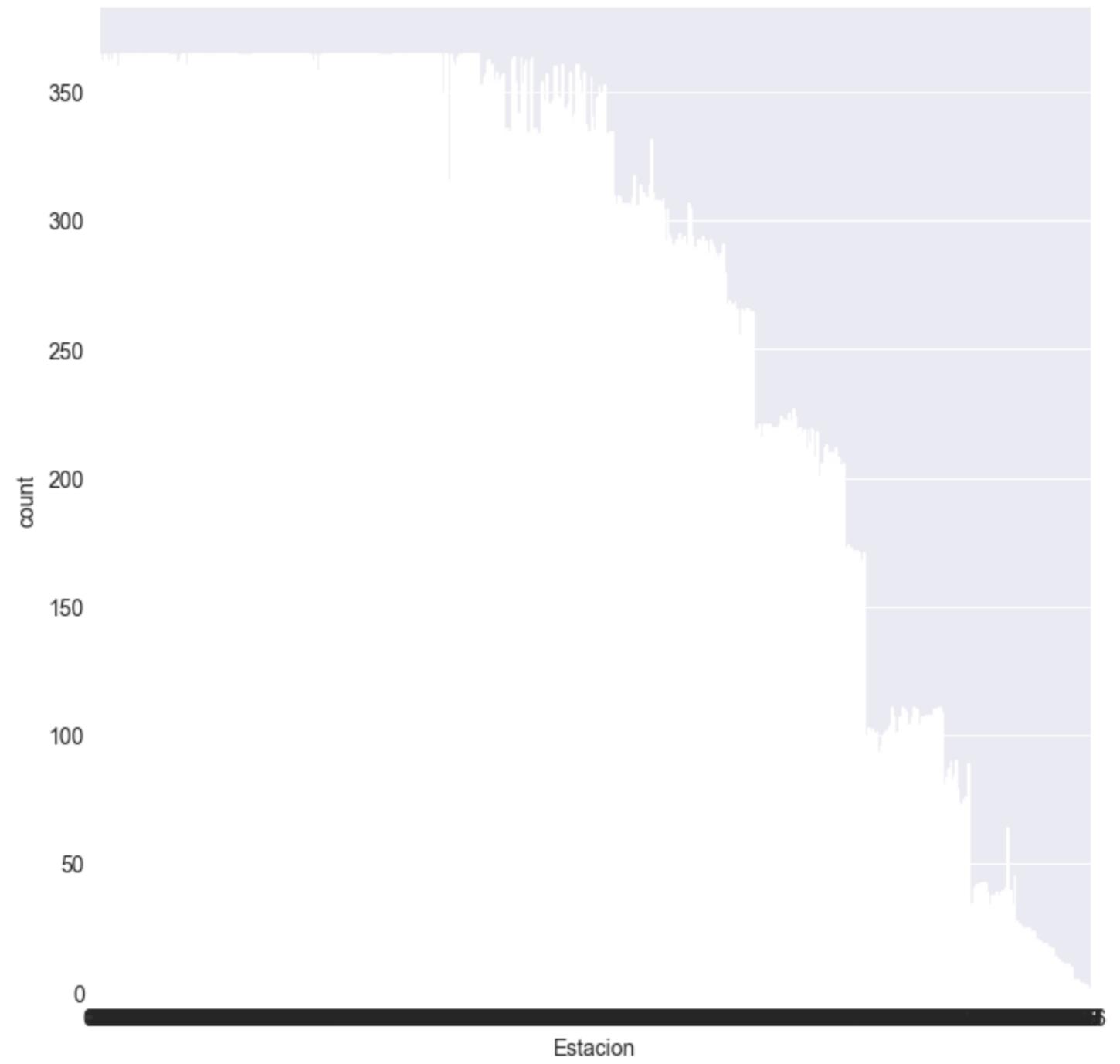
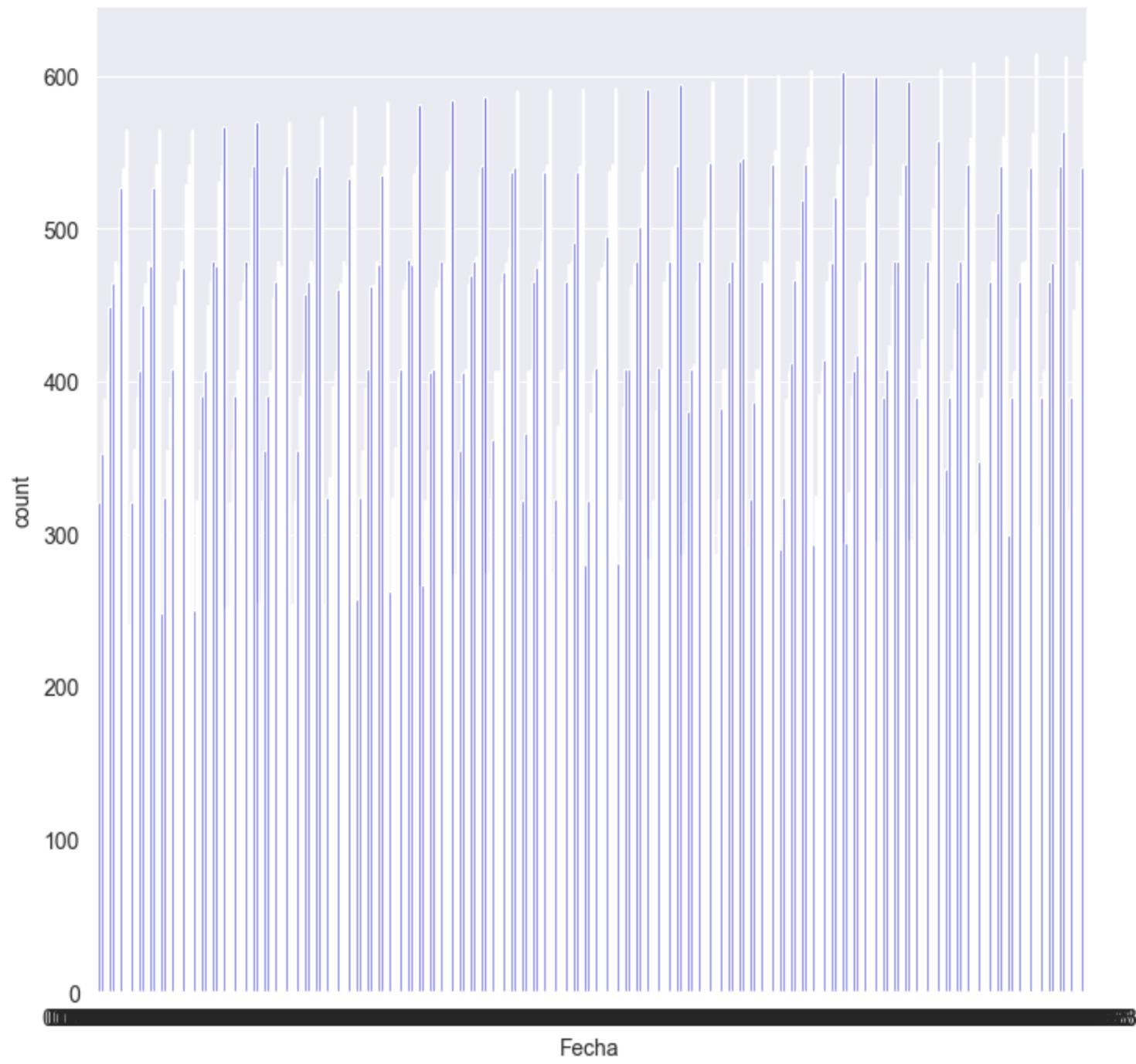
DBSCAN Clustering - 2D Visualization

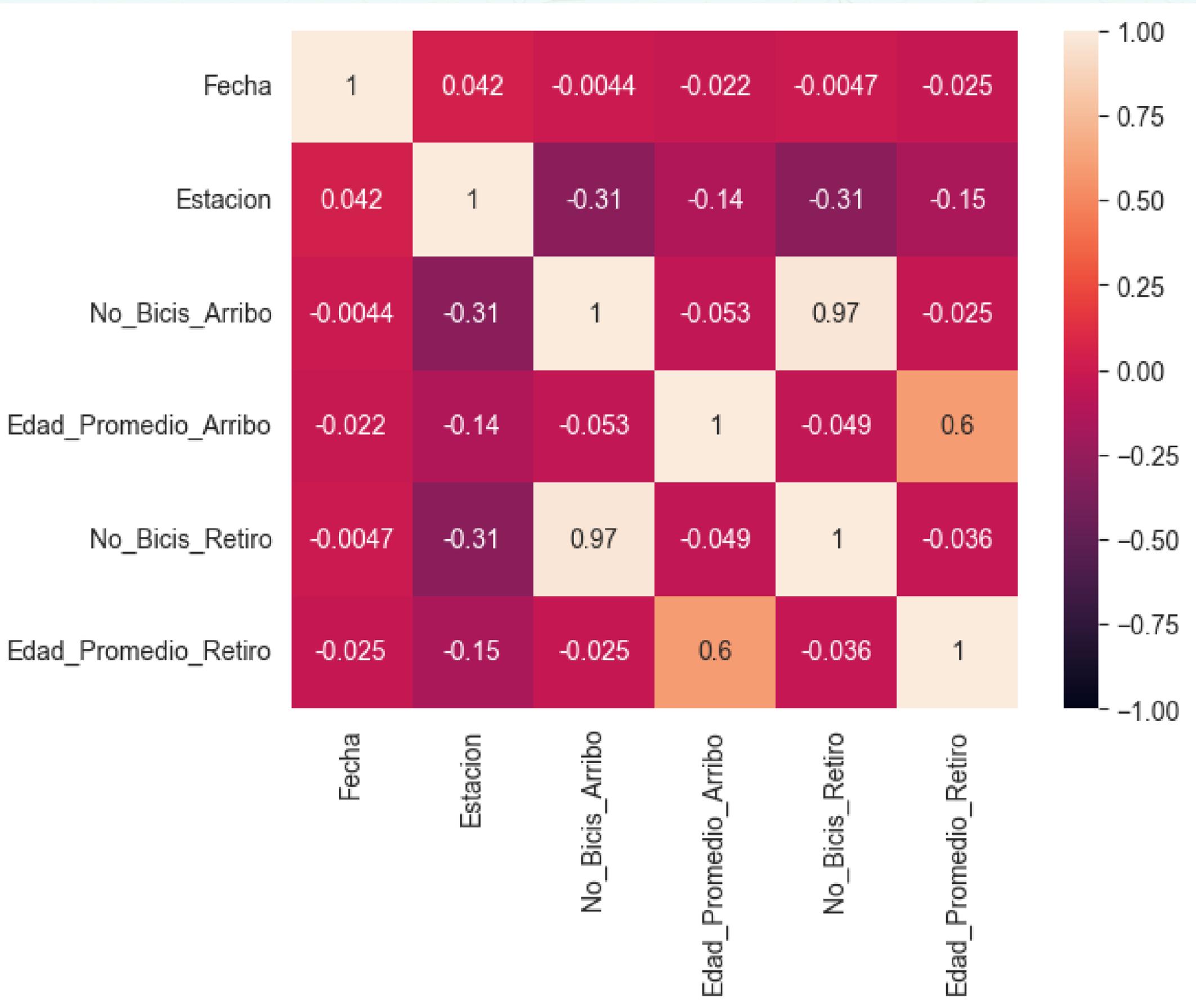


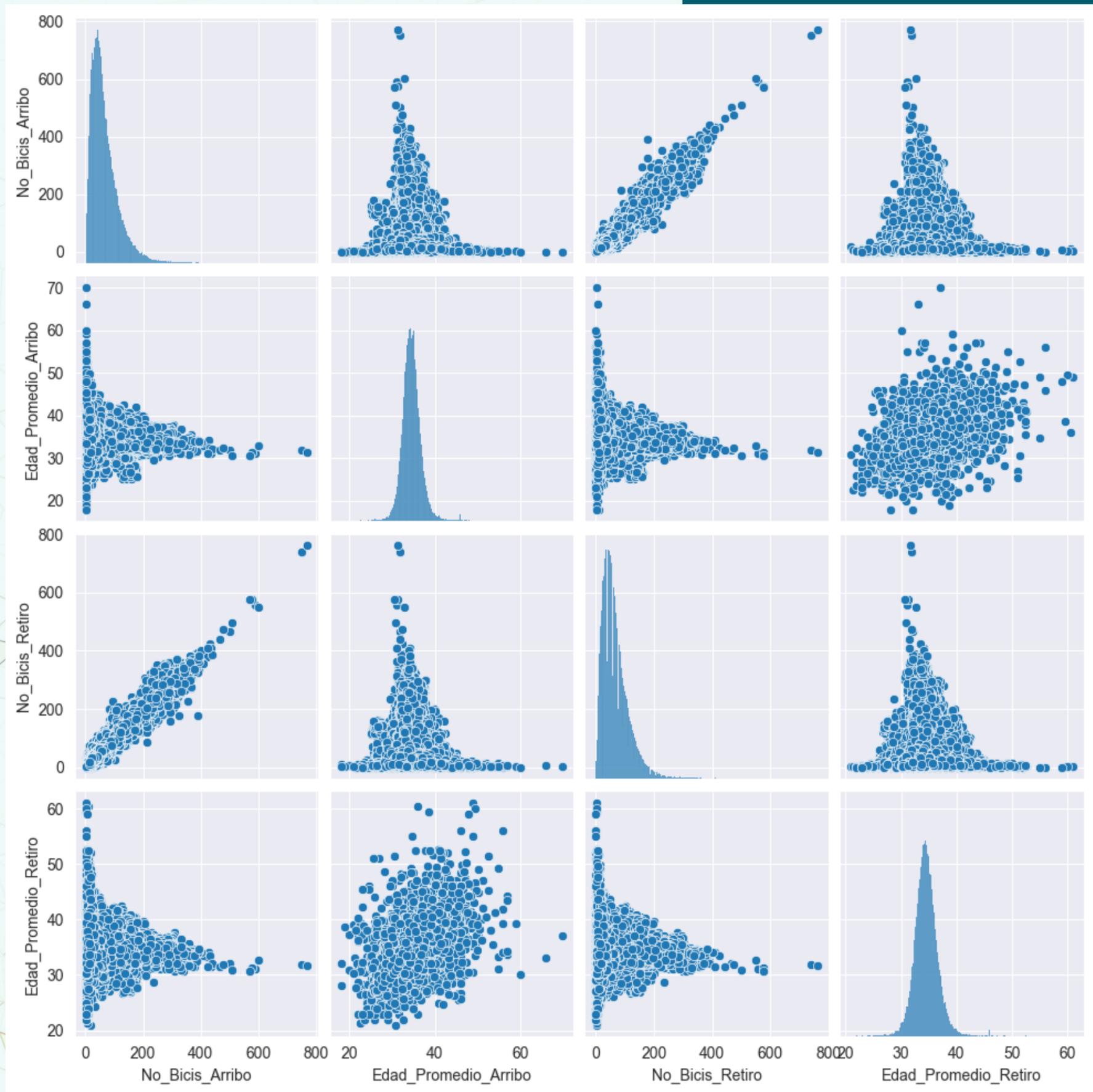
CAMBIO DE VARIABLES 2

Dado que parecía haber una relación entre los trayectos y el tiempo de uso, quisimos analizar si existía una relación entre la cantidad de bicicletas que se retiraban de cada estación y las fechas en la que lo hacían, así como si había alguna edad que predominaba en el área



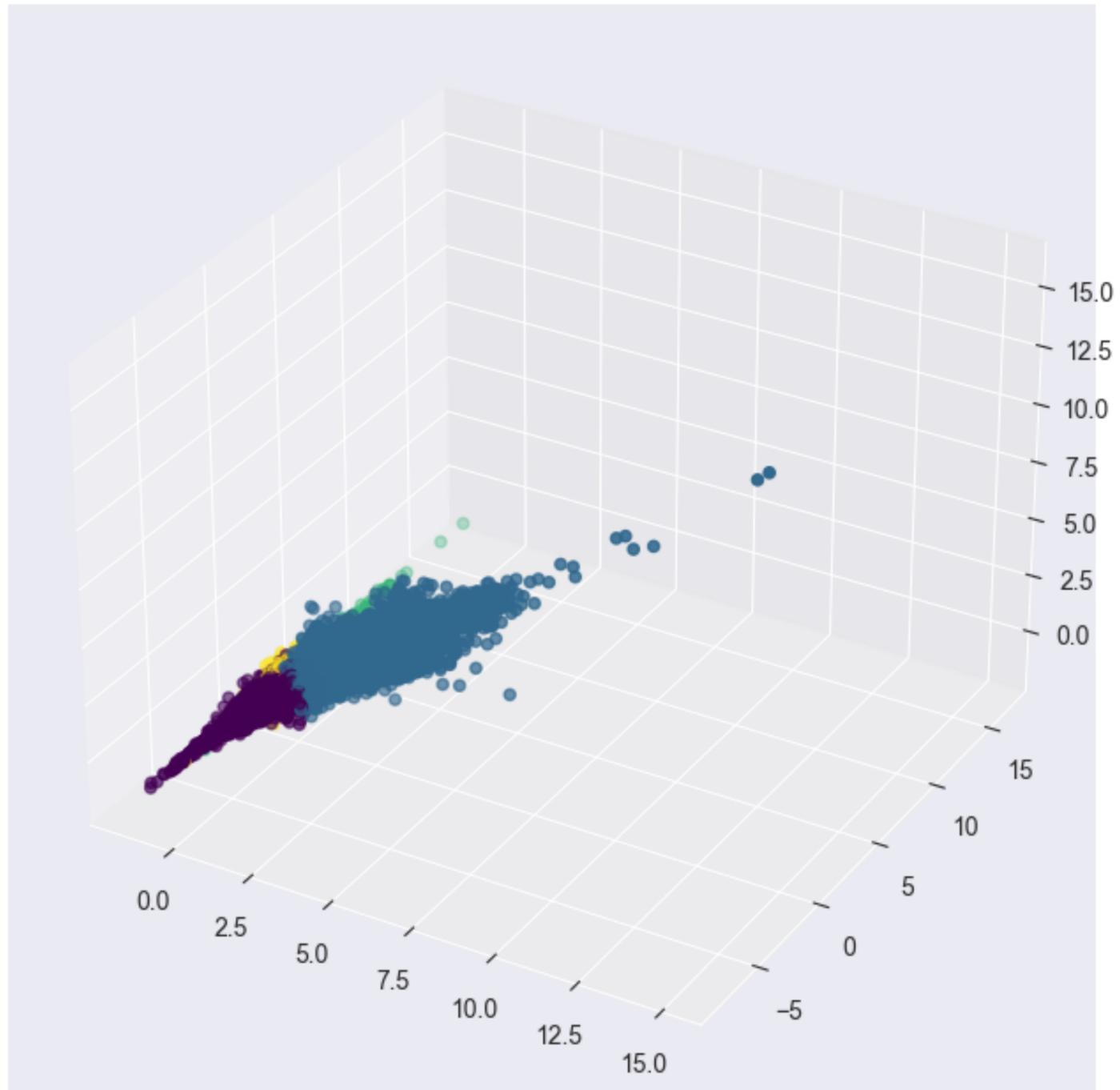




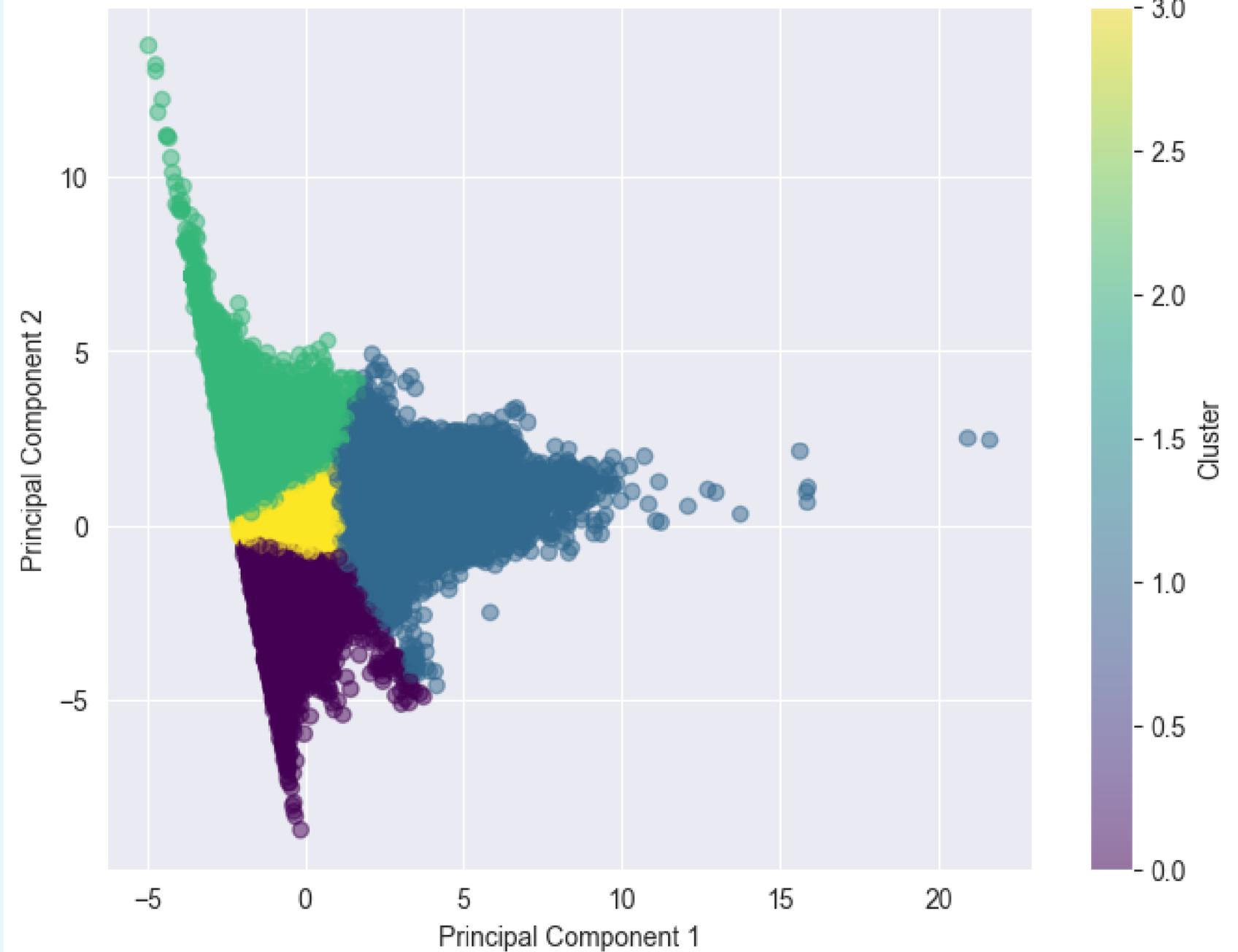


Kmeans sin Estación

KMeans Clustering - 3D Visualization

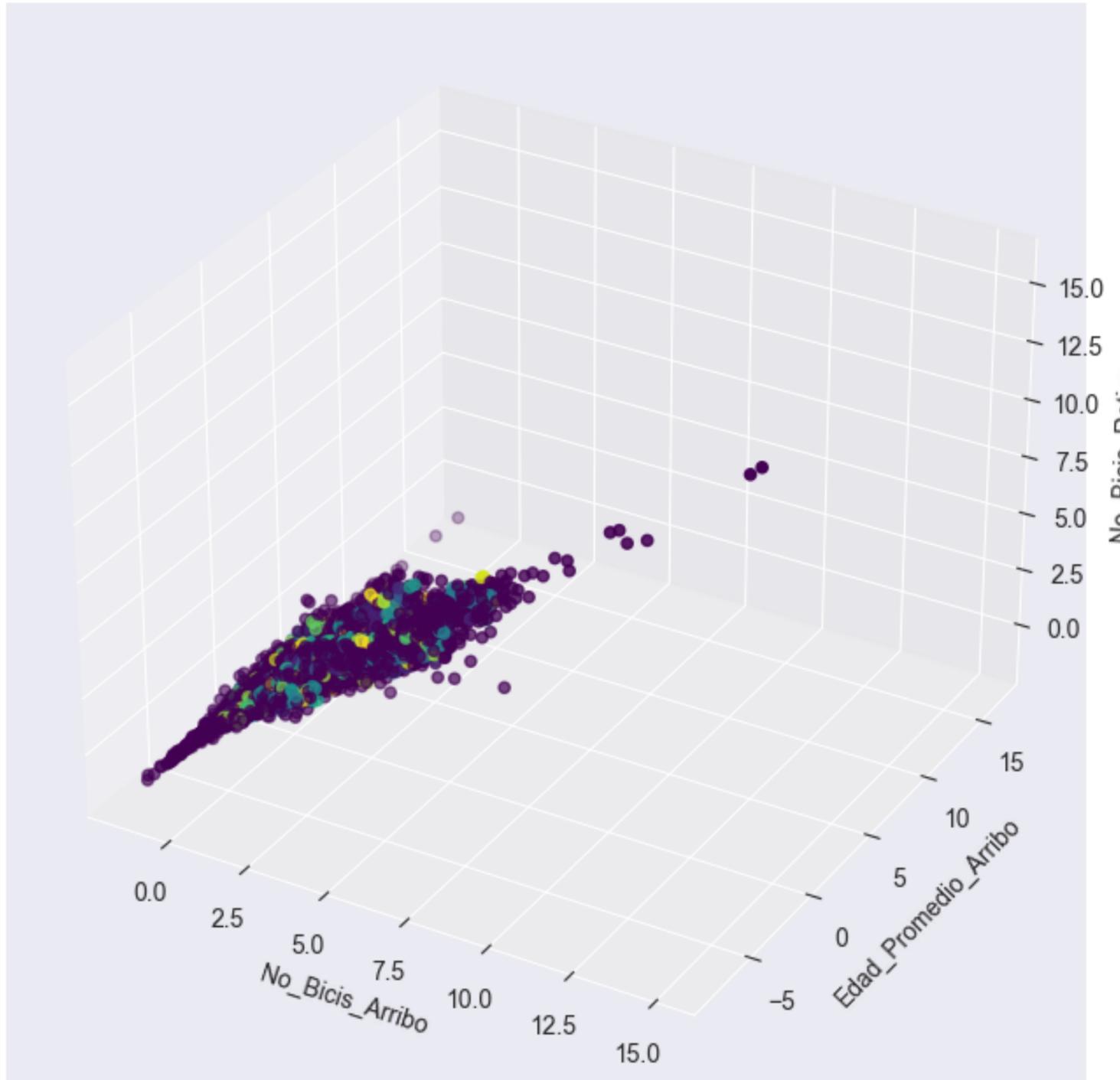


KMeans Clustering - 2D Visualization

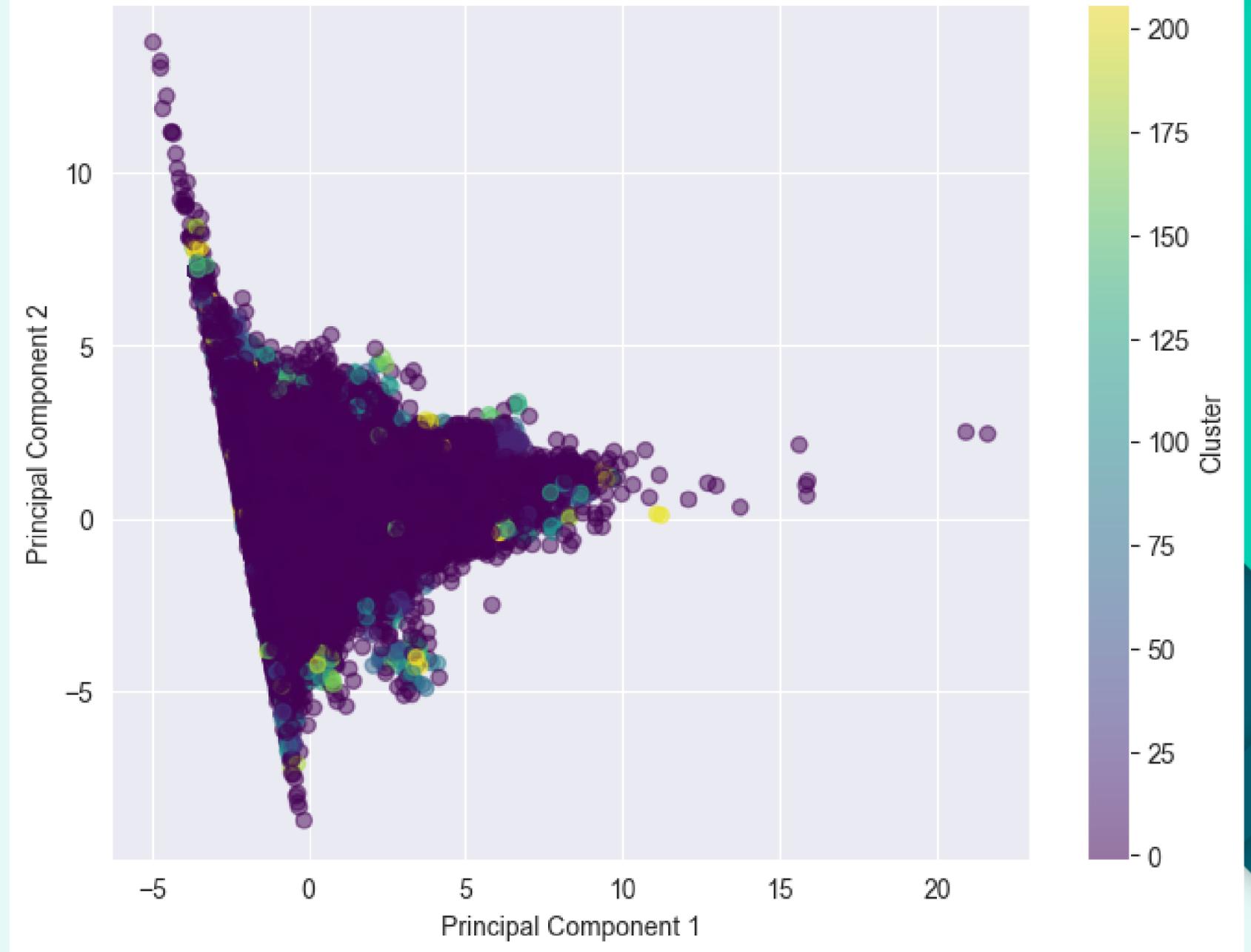


DBSCAN

DBSCAN Clustering - 3D Visualization

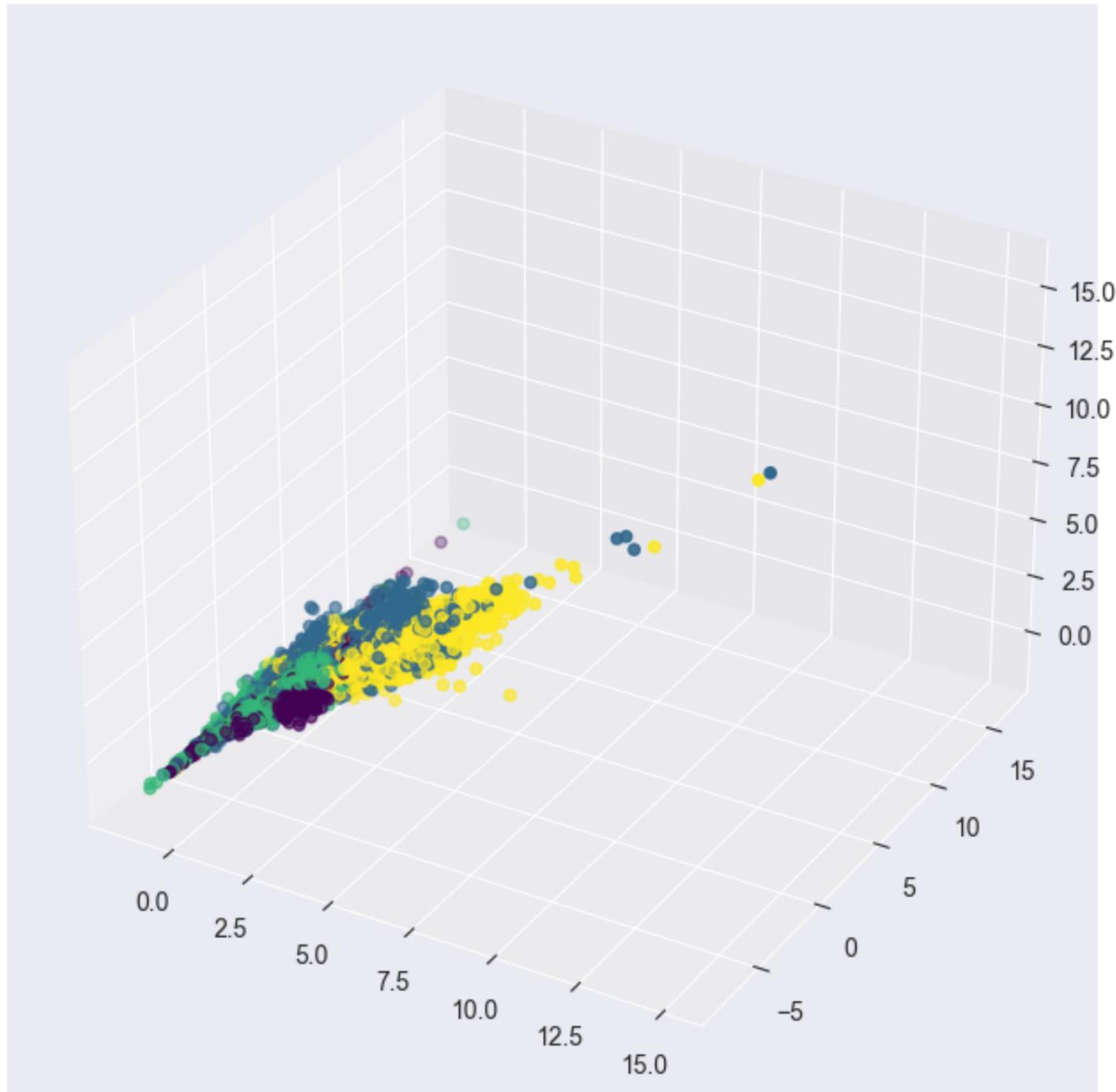


DBSCAN Clustering - 2D Visualization

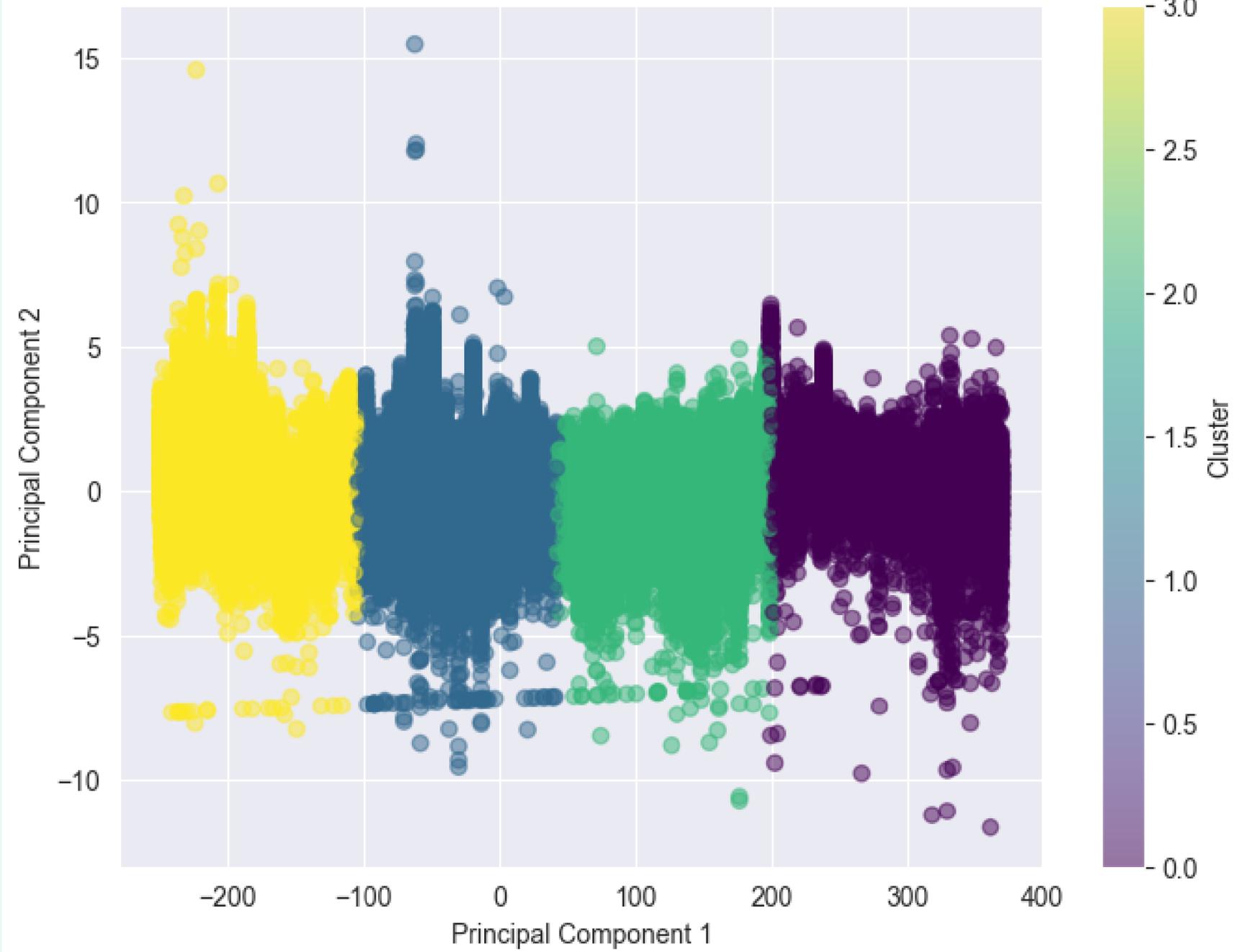


Kmeans con Estación

KMeans Clustering - 3D Visualization

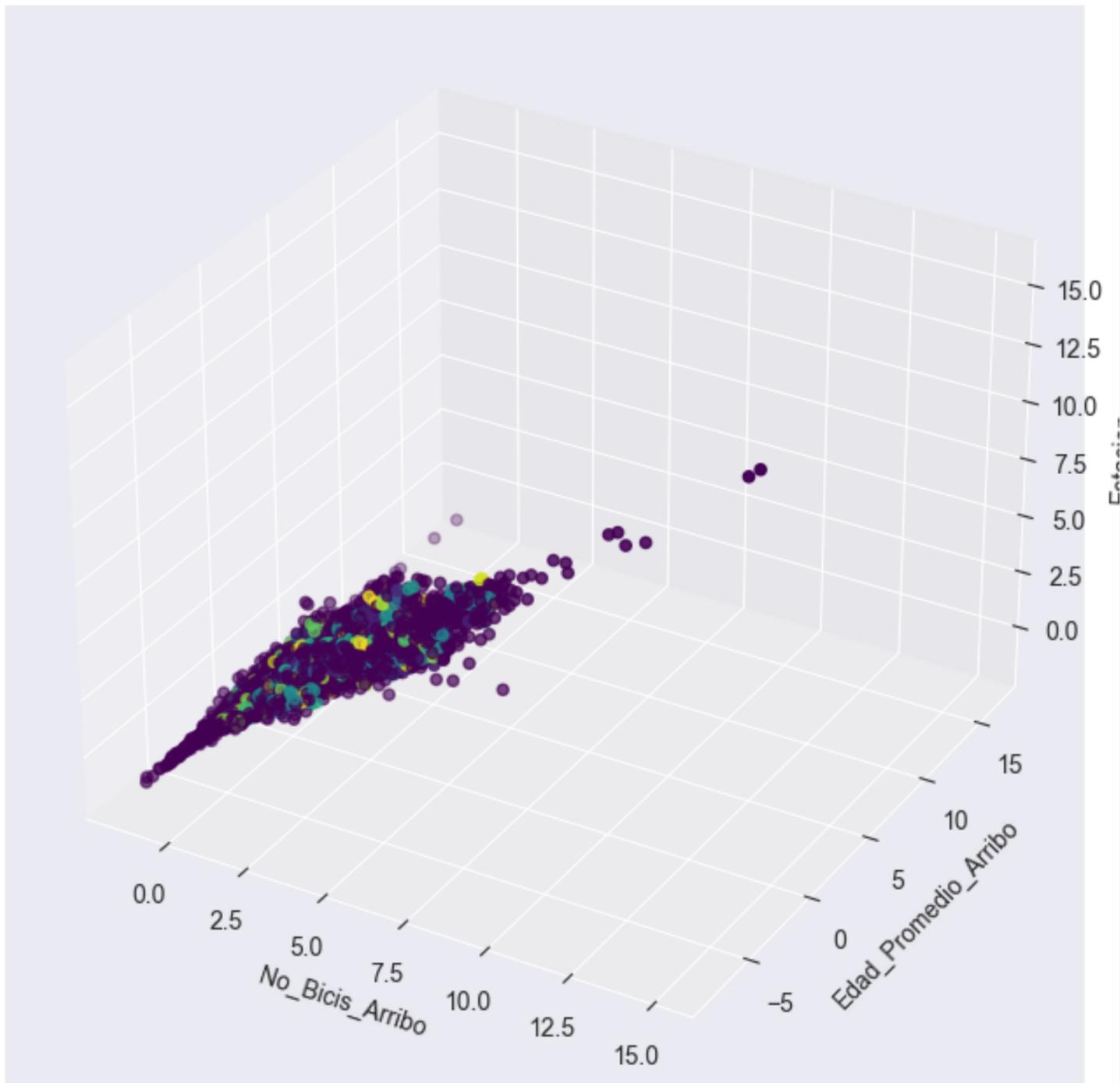


KMeans Clustering - 2D Visualization

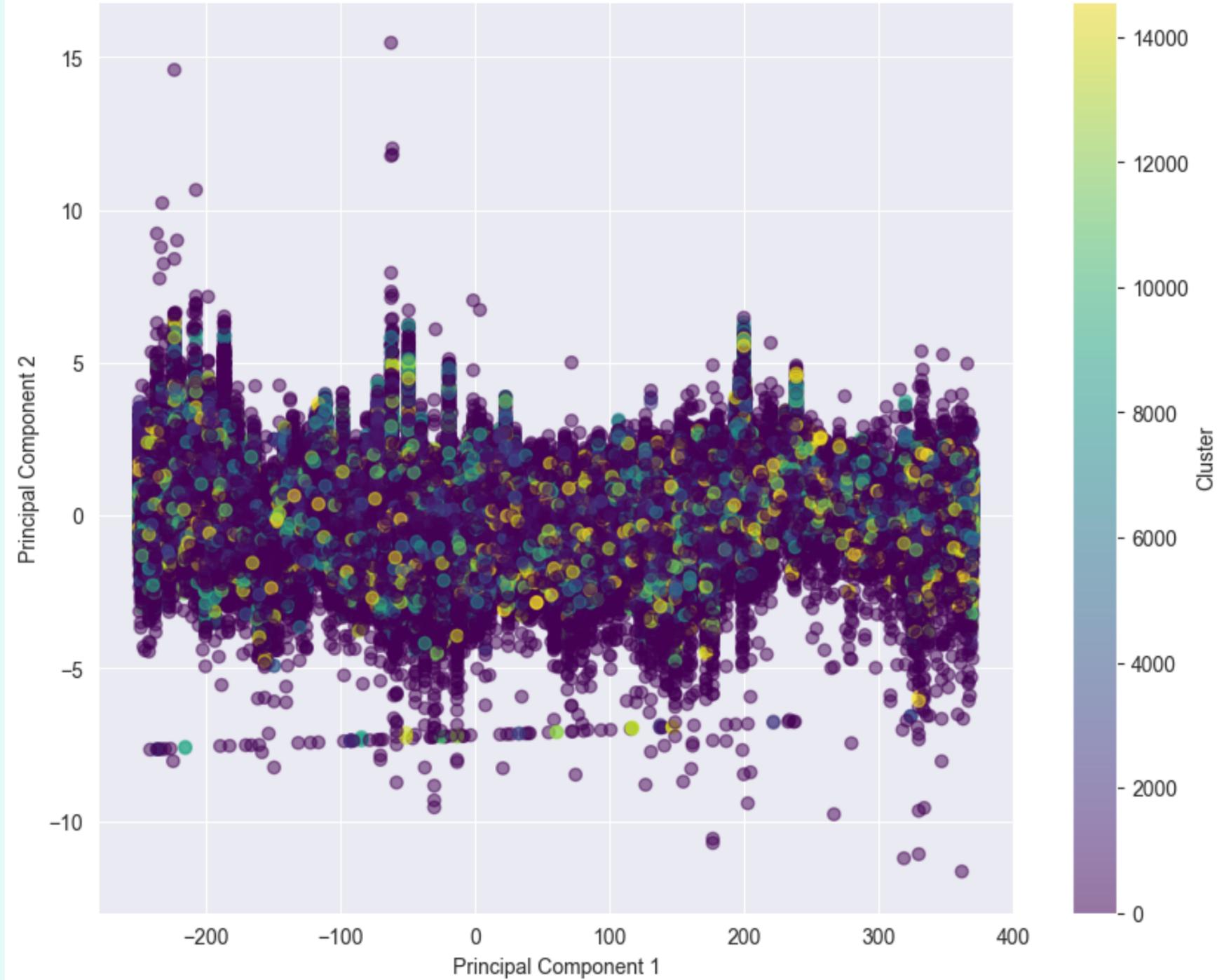


DBSCAN CON ESTACIÓN

DBSCAN Clustering - 3D Visualization with Estacion

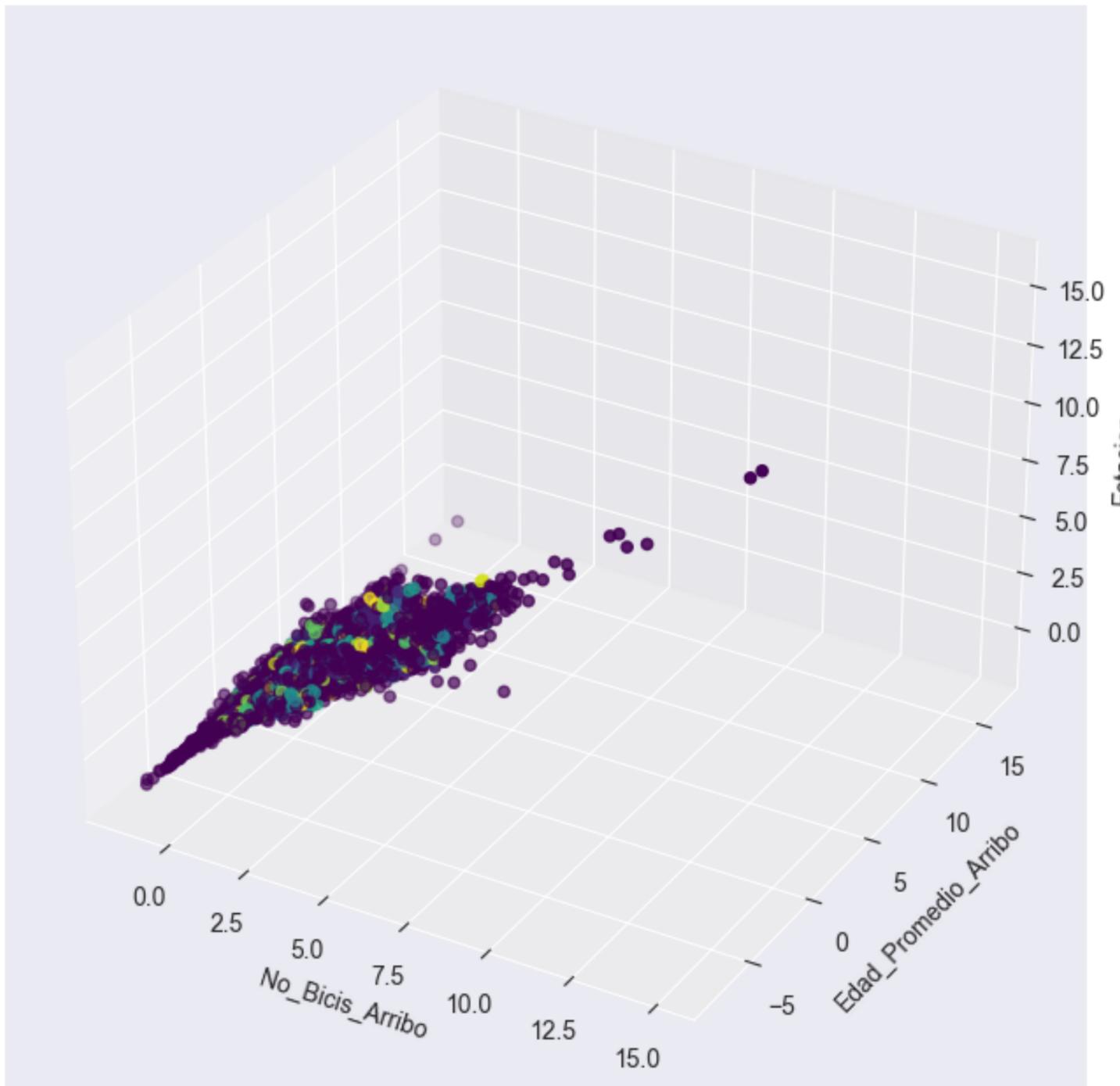


DBSCAN Clustering - 2D Visualization with Estacion

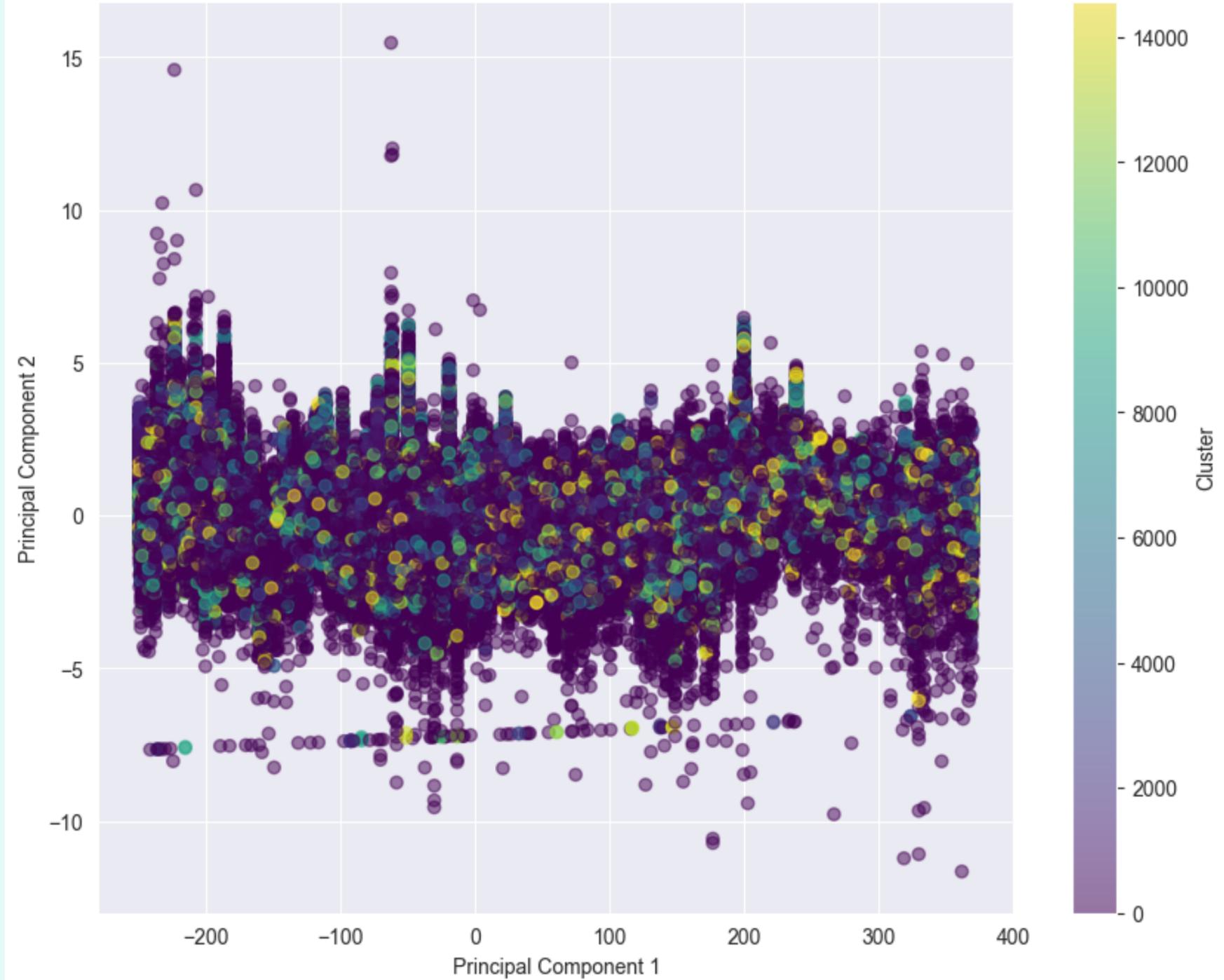


DBSCAN CON ESTACIÓN

DBSCAN Clustering - 3D Visualization with Estacion



DBSCAN Clustering - 2D Visualization with Estacion



Observaciones

- Datos atípicos
 - Edad
 - Fechas de Arribo y Recibo
 - Tiempos de viaje

Fin

