

ANÁLISIS EXPLORATORIO DE DATOS

Patrones de carga de autos eléctricos.

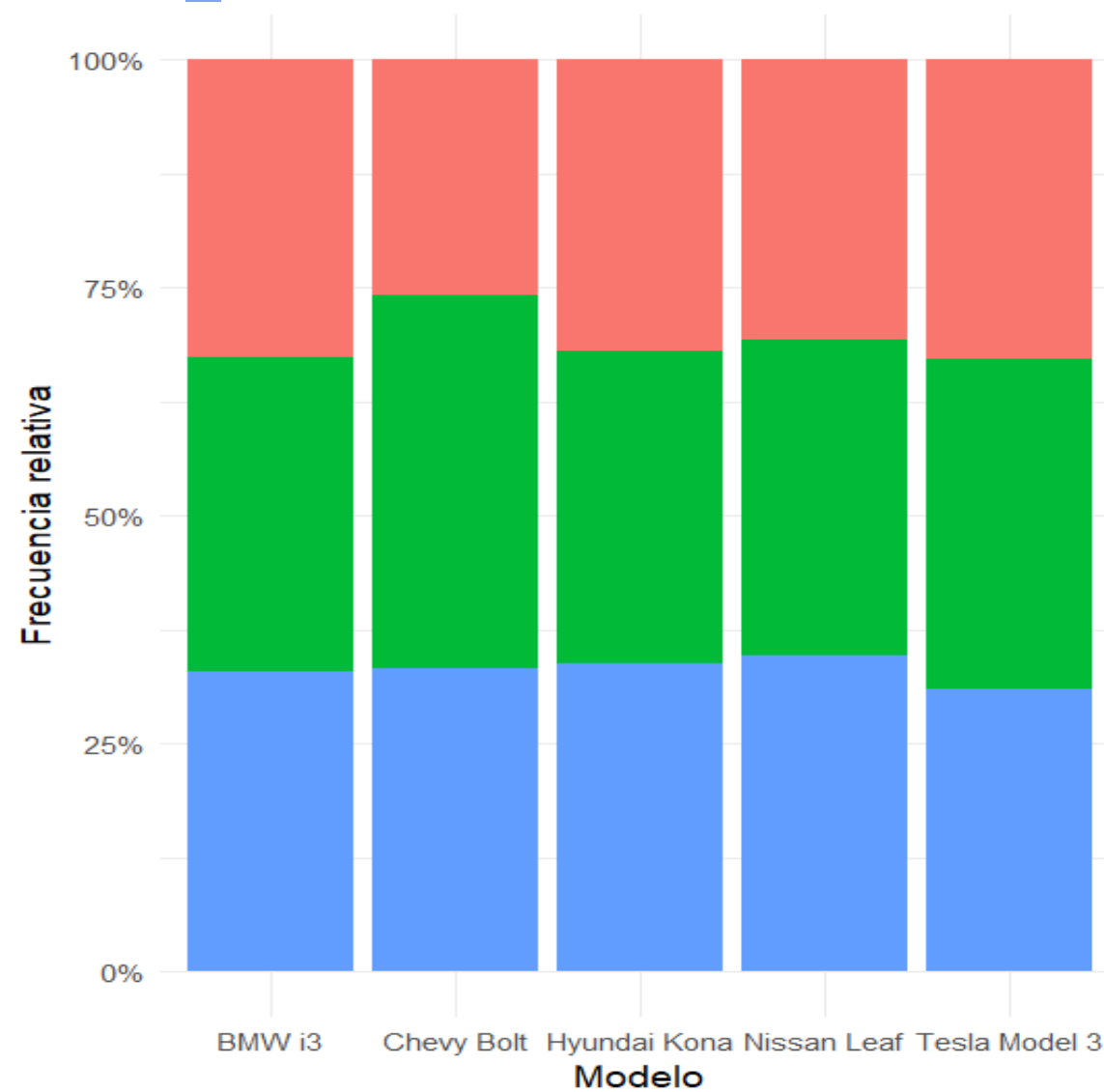
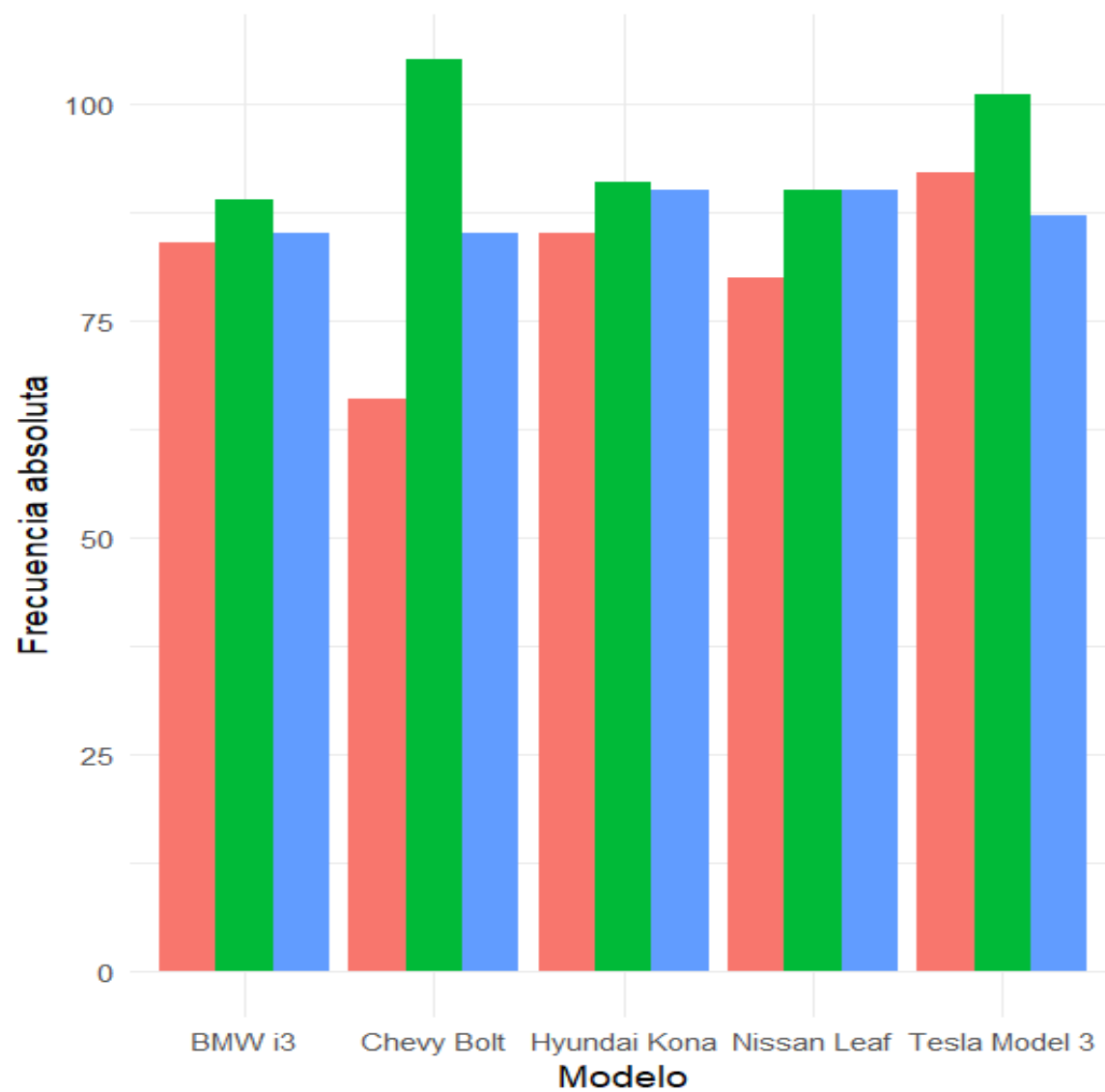
Integrantes

Matias Gonzalez Balbi. Legajo 205631

Juan Camacho. Legajo 205632

Cantidad de vehículos por tipo de usuario.

- Commuter
- Casual driver
- Long distance traveler



¿Qué nos muestra este gráfico?

En este gráfico observamos la cantidad de vehículos filtrados por tipo de usuario.

En el primer gráfico podemos observar la cantidad de vehículos por usuario de manera absoluta y en el segundo gráfico de manera relativa.

Conclusiones.

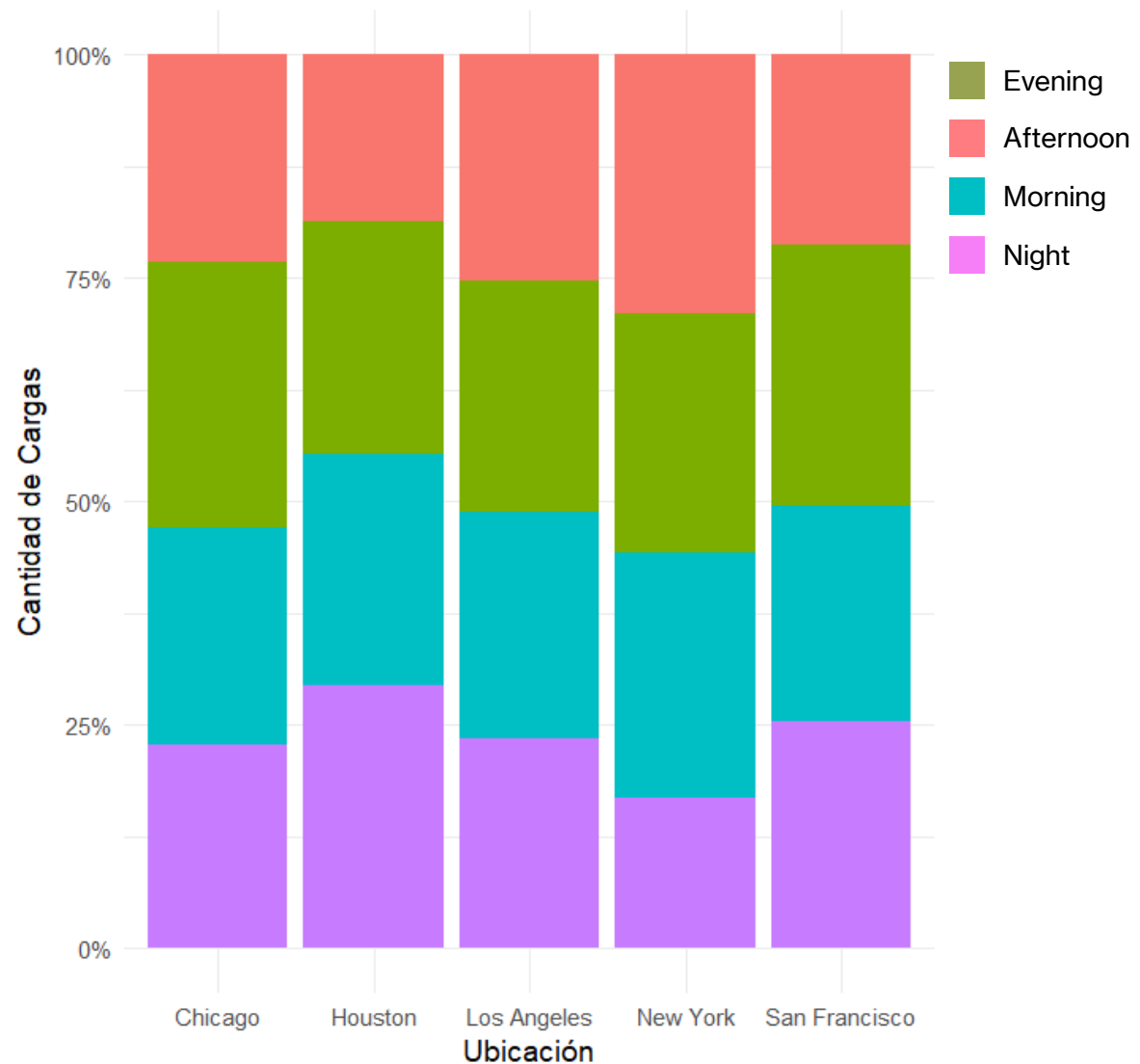
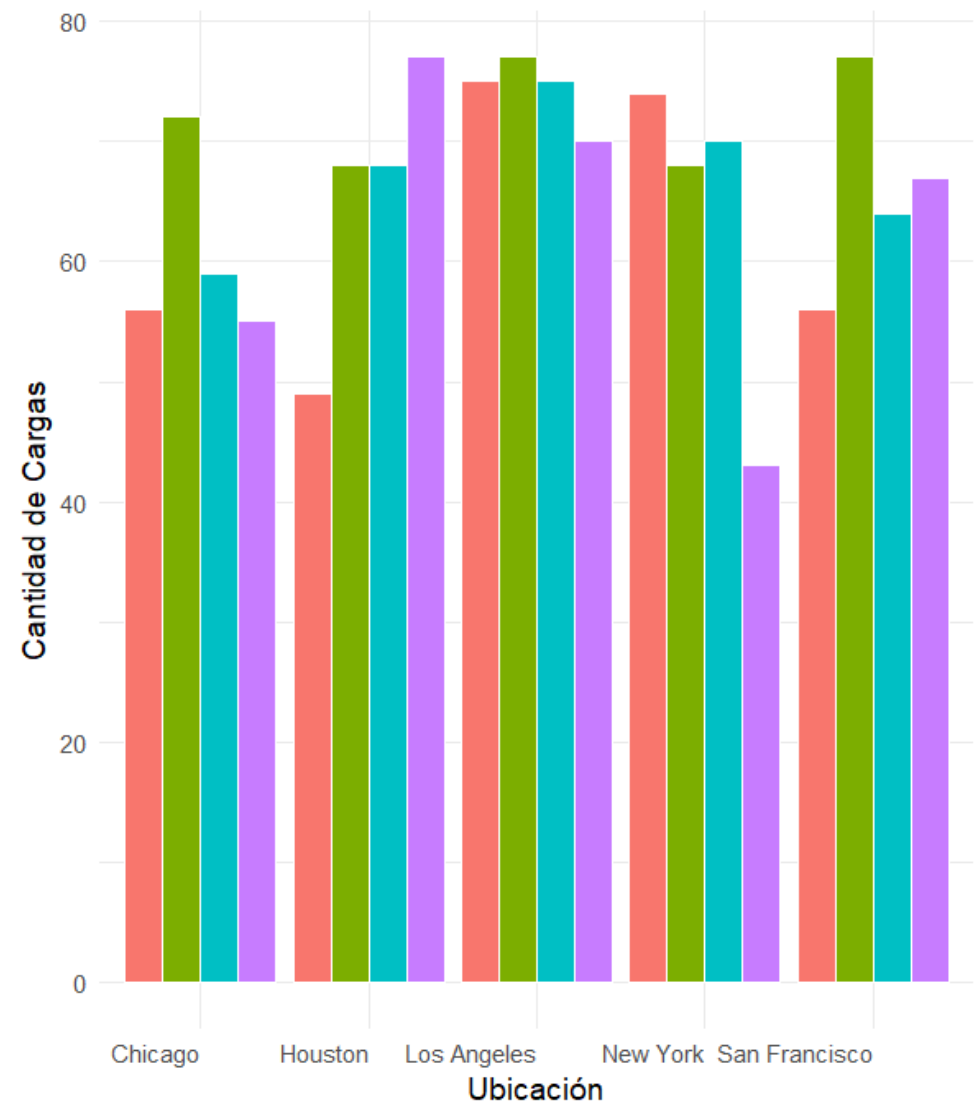
Al observar los datos podemos notar una ligera inclinación por parte de los Commuter (trabajadores diarios) por el Chevy Bolt, seguido por el Tesla Model 3.

Para el resto de los modelos no parece haber una tendencia o inclinación clara por parte de los usuarios a un tipo de vehículo u otro dependiendo de su condición.

	BMW i3	Chevy Bolt	Hyundai Kona	Nissan Leaf	Tesla Model 3
Casual Driver	84	66	85	80	92
Commuter	89	105	91	90	101
Long-Distance Traveler	85	85	90	90	87

Cargas por Ciudad

Filtrados por momento del día



¿Qué nos muestra este gráfico?

Este gráfico nos muestra las cargas que realizan los conductores por ciudad, de acuerdo con el momento del día. Nuevamente en primer gráfico se representa de manera absoluta mientras que en el de la derecha se representa de manera relativa de forma apilada.

Conclusiones.

Entendemos por estos datos que en promedio los conductores prefieren cargar sus vehículos durante la tarde noche (evening). También vemos que en New York el promedio de cargas durante la noche es significativamente menor al resto, esto puede deberse a la inseguridad en la ciudad.

	Chicago	Houston	Los Angeles	New York	San Francisco
Morning	59	68	75	70	64
Afternoon	56	49	75	74	56
Evening	72	68	77	68	77
Night	55	77	70	43	67

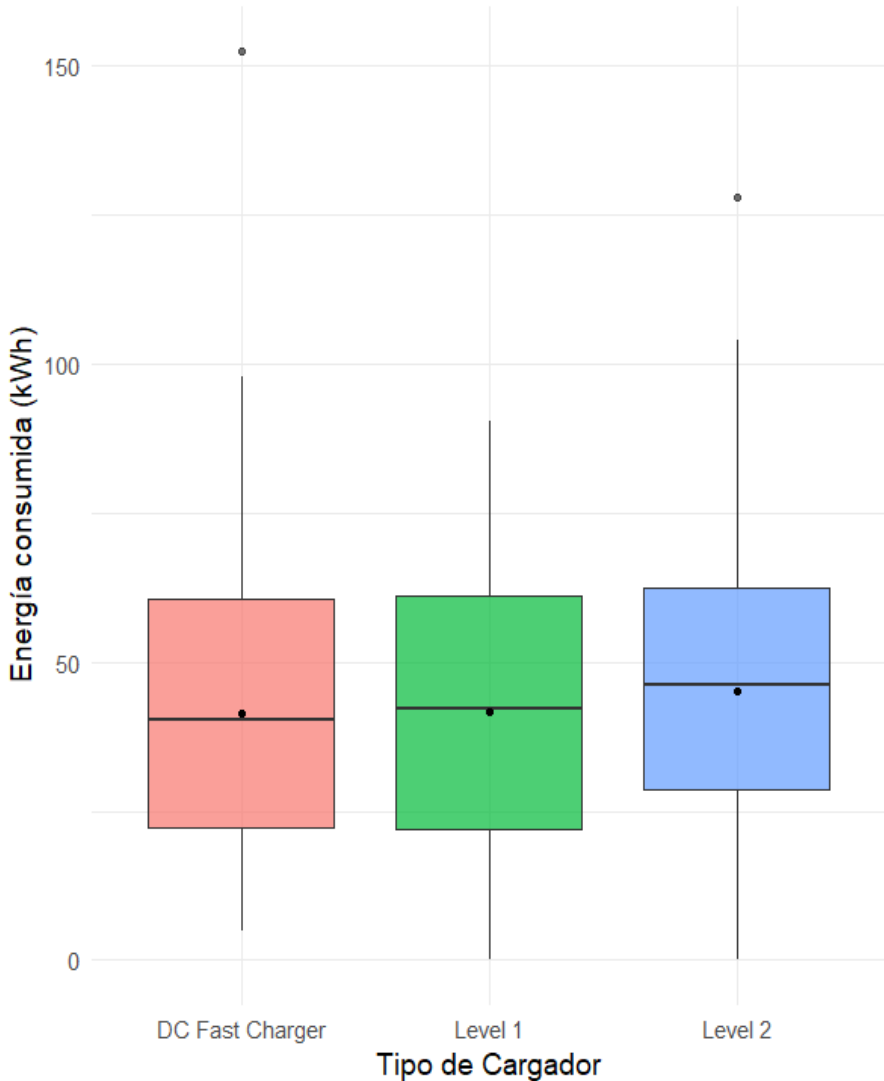
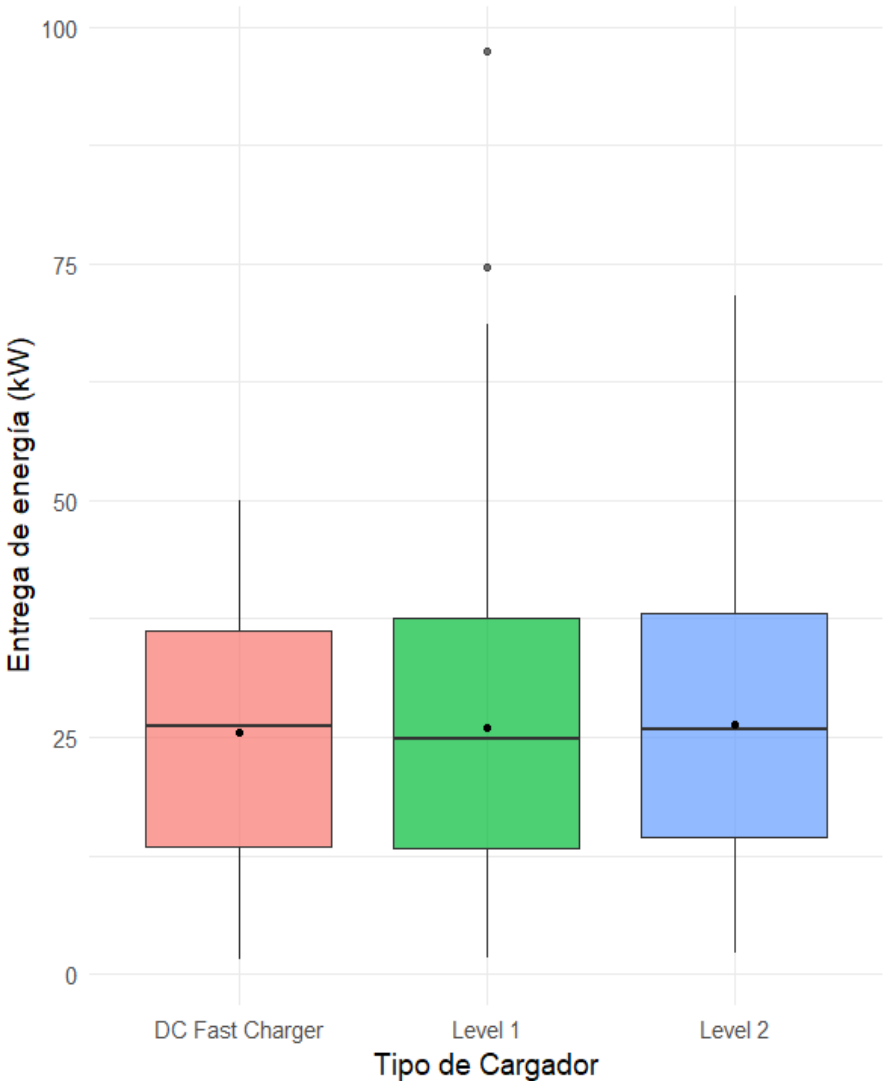
Comparación de tipo de cargador con energía consumida y velocidad de carga.

Entrega de energía
(Agrupada por tipo de cargador)

DC Fast Charger	
Media	25.6
Mediana	26.1
Moda	NA
C.V	53.1
Conteo	409

Level 1	
Media	26
Mediana	24.8
Moda	NA
C.V	56.6
Conteo	434

Level 2	
Media	26.3
Mediana	25.9
Moda	NA
C.V	52.0
Conteo	411



Energía consumida
(Agrupada por tipo de cargador)

DC Fast Charger	
Media	41.3
Mediana	40.2
Moda	NA
C.V	55.1
Conteo	407

Level 1	
Media	41.6
Mediana	42.3
Moda	NA
C.V	54.3
Conteo	437

Level 2	
Media	45.1
Mediana	46.1
Moda	NA
C.V	48.1
Conteo	410

¿Qué nos muestra este gráfico?

El primer gráfico representa la entrega de energía agrupada por tipo de cargador medida en kilovatios en promedio.

El gráfico de la derecha nos muestra la energía consumida también agrupada por tipo de cargador, medida en kilovatios hora.

Observaciones.

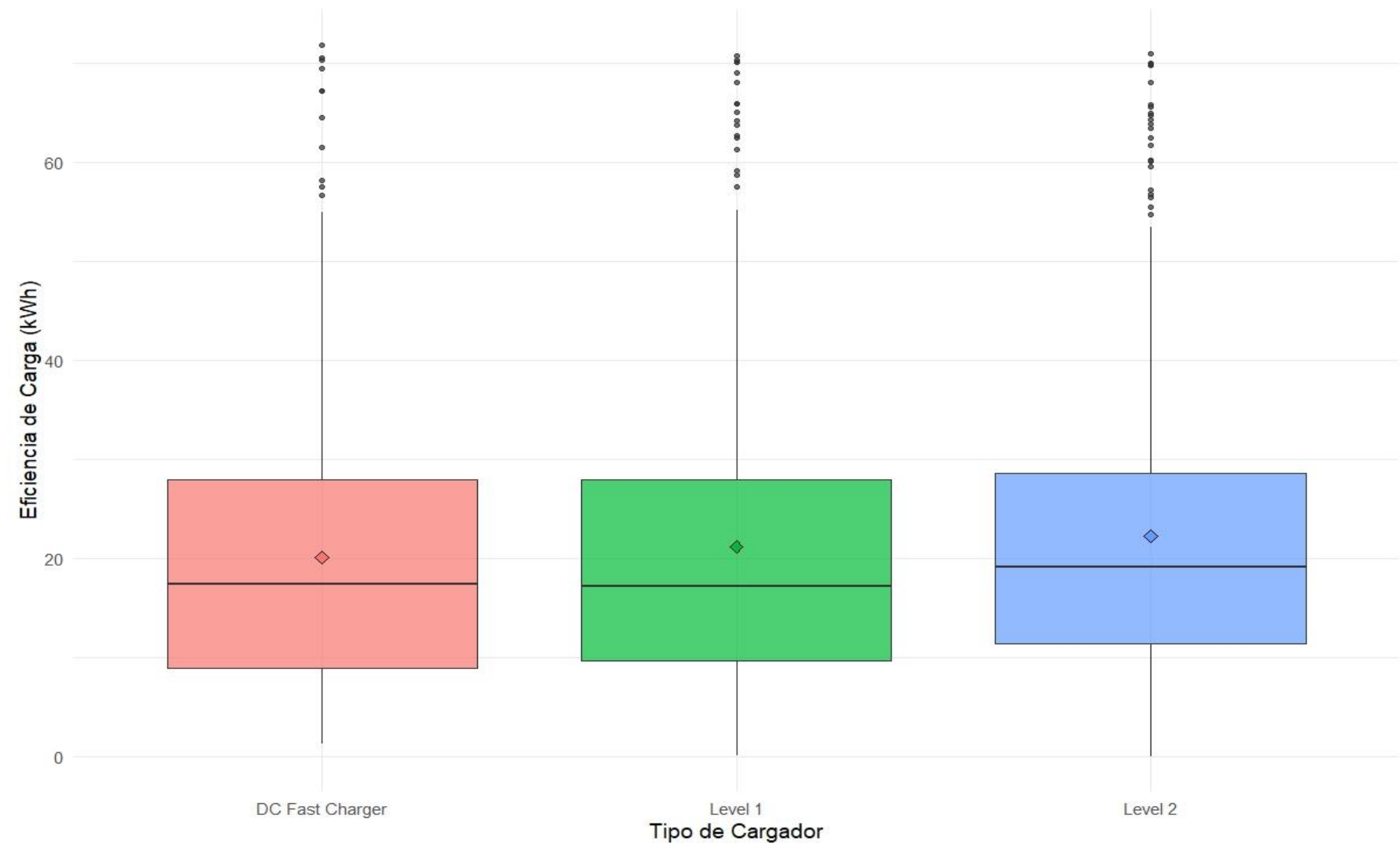
Podemos observar tanto una media como una mediana similar para ambos casos, lo que sugiere que la distribución es razonablemente simétrica.

Todos los coeficientes de variación están alrededor de 50%, lo que indica una variabilidad alta respecto a la media.

Conclusiones

A partir de estos datos podemos deducir que la energía consumida y la entrega de esta al vehículo no varía por tipo de cargador de manera significativa.

Eficiencia de carga por tipo de cargador (-5% Superior de los datos)



Eficiencia de carga (kWh)
(Agrupada por tipo de cargador)

DC Fast Charger	
Media	20.2
Mediana	17.4
Moda	NA
C.V	70.9
Conteo	386

Level 1	
Media	21.2
Mediana	17.2
Moda	NA
C.V	75.5
Conteo	413

Level 2	
Media	22.3
Mediana	19.2
Moda	NA
C.V	68.0
Conteo	392

¿Qué nos muestra este gráfico?

Para realizar este gráfico creamos la variable **Eficiencia de Carga**, realizando el siguiente calculo:

Energía consumida / Duración de la carga.

Para visualizarlo fueron agrupados por tipo de cargador y el 5% superior de los datos fue extraído debido a la presencia de outliers extremos.

Observaciones.

Se pueden observar gran cantidad de datos outliers en los tres diagramas aún habiendo quitado el 5% superior.

El promedio es levemente mayor que la mediana por lo que podemos asumir un leve sesgo positivo.

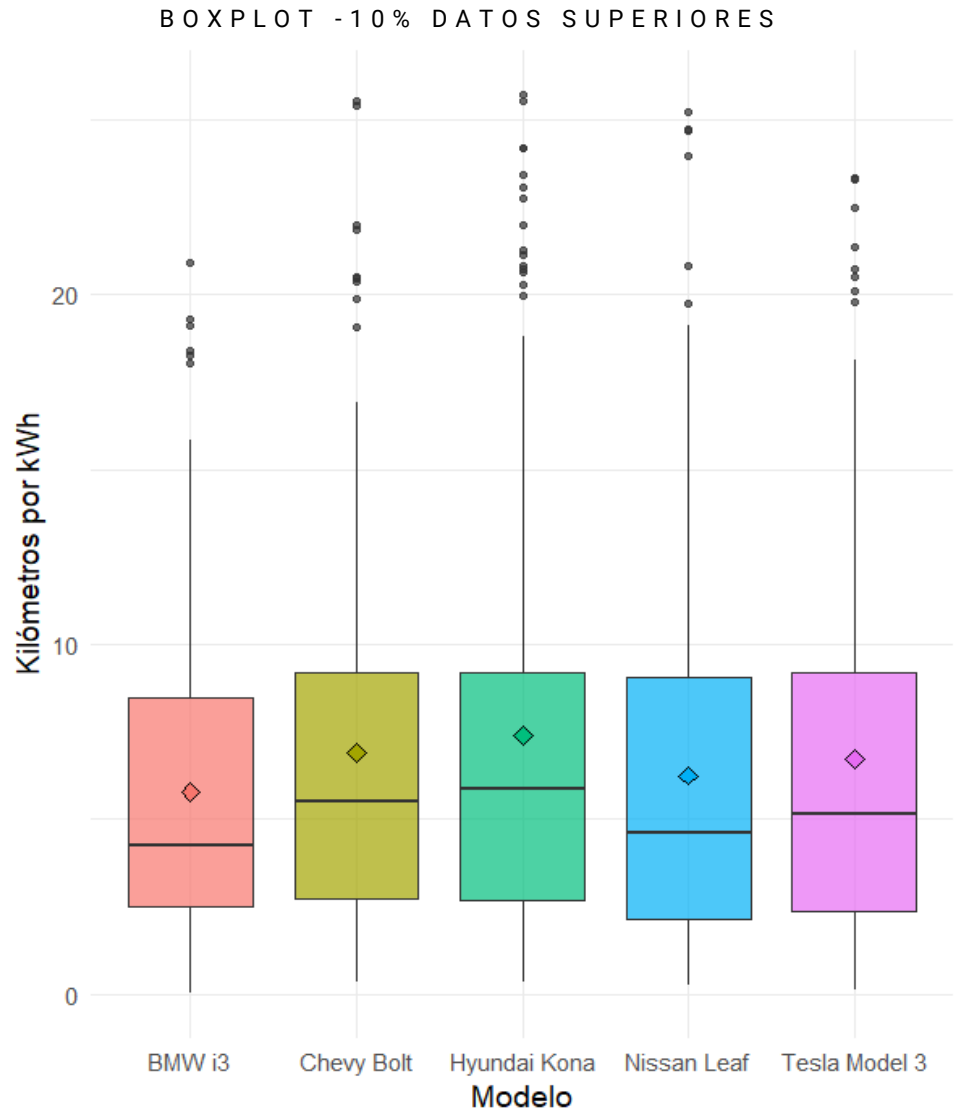
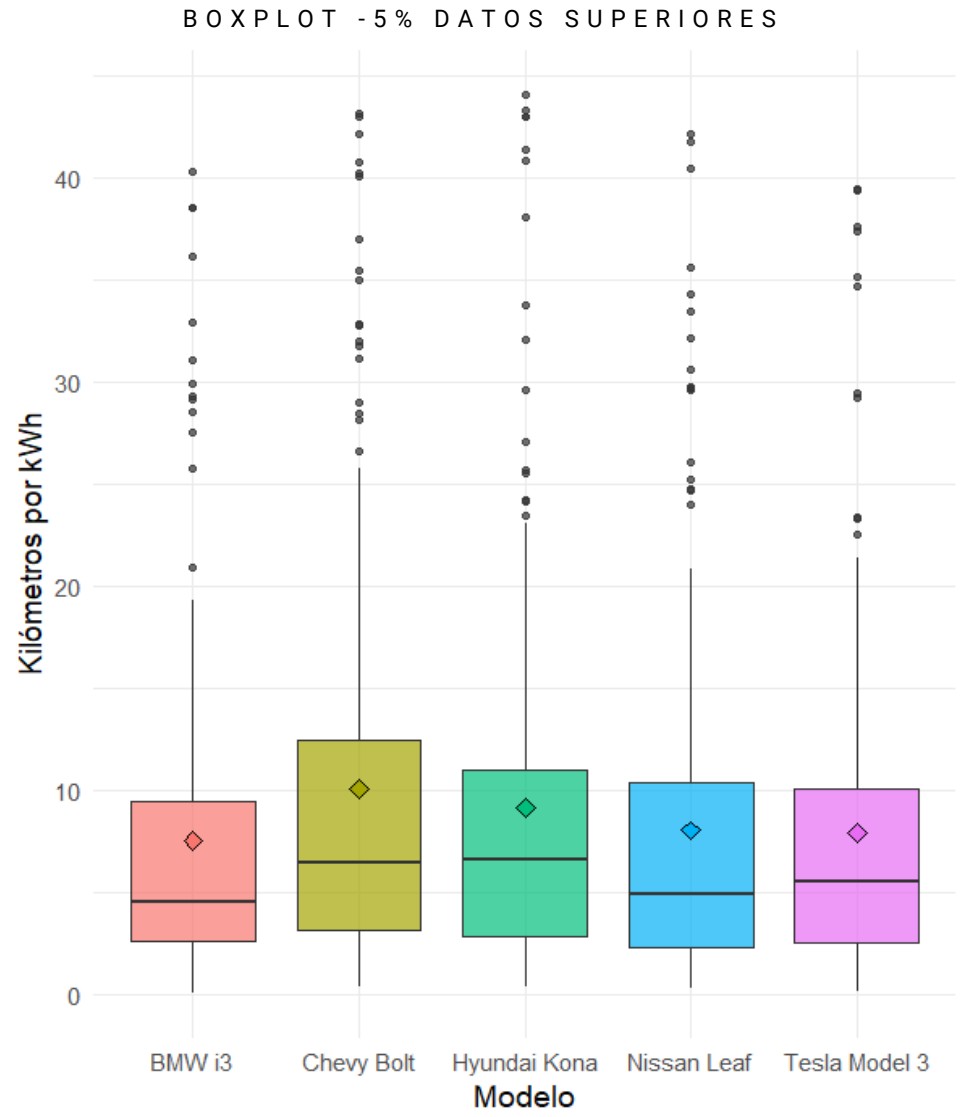
El coeficiente de variación es significativamente alto para todos los casos, cerca de un 70%, por lo que podemos decir que la media no es representativa.

Conclusiones

Mediante los datos analizados entendemos que ningún cargador es sustancialmente mas eficiente que el resto, tanto en términos de rapidez como de consumo.

Eficiencia por kWh

Agrupado por tipo de vehículo



¿Qué nos muestra este gráfico?

Para realizar este gráfico creamos la variable **Eficiencia por kWh**, realizando el siguiente calculo:
Empezamos restando **SOC End y SOC Start**, obteniendo como resultado, la energía total cargada durante la sesión.
Esta variable la utilizamos para multiplicarlo por la **capacidad de la batería**, obteniendo como resultado la variable **energía utilizada**.
Finalmente, la **distancia recorrida** la dividimos por 'batería utilizada' , para obtener la variable '**eficiencia por kWh**.'

Esta variable tiene como objetivo medir **cuántos kilómetros se recorrieron por cada kilovatio-hora de energía consumida**.

Observaciones.

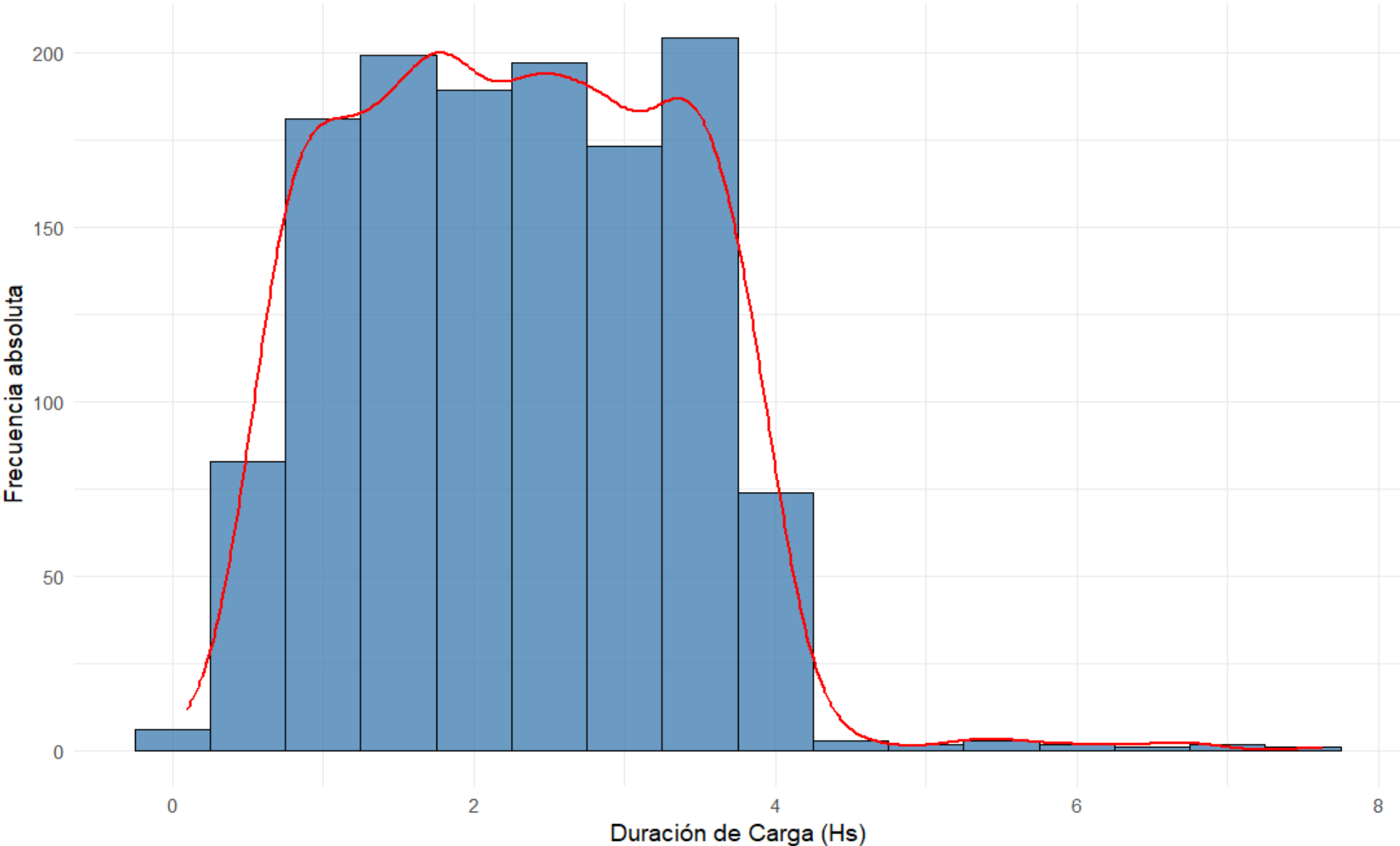
Cómo se detalla en el gráfico, se realizo una disminución del 5% y %10 de la parte superior de los datos para eliminar outliers extremos. Sin embargo, aún así notamos la presencia de valores outliers que están afectando el promedio al alza bastante por encima de la media lo que formaría una distribución con un sesgo positivo.

Conclusiones

Es difícil concluir cuál es el modelo más eficiente debido a la gran cantidad de datos outliers y la gran dispersión de los datos la cual se confirma por el coeficiente de variación y la gran diferencia entre los máximos y mínimos, sin embargo, basándonos en la mediana de los datos quitando el 10% de los datos superiores, podemos aproximar que el modelo más eficiente esta entre el Hyundai Kona y el Chevy Bolt

Modelo	promedio	mediana	moda	sd	var	cv	q1	q3	min	max	conteo
BMW i3	6.68	4.41	NA	6.4	40.9	95.8	2.58	9.03	0.0195	31	177
Chevy Bolt	8.1	5.95	NA	7.35	54	90.7	2.8	10.3	0.352	32	163
Hyundai Kona	7.75	6.16	NA	6.59	43.4	85	2.73	9.6	0.364	32	183
Nissan Leaf	6.88	4.74	NA	6.5	42.2	94.4	2.19	9.17	0.247	30.6	181
Tesla Model 3	6.95	5.19	NA	5.87	34.5	84.4	2.39	9.26	0.123	29.4	193

Distribución de la duración de carga



Duración de carga	
Media	2.27
Mediana	2.26
Moda	NA
C.V	46.75
Q1	1.4
Q3	3.11
S.D	1.06
VAR	1.13
Min	0.1
Máx	7.64

¿Qué nos muestra este gráfico?

El gráfico anterior nos muestra un histograma donde se ve reflejada la **distribución de la duración** de la carga en horas.

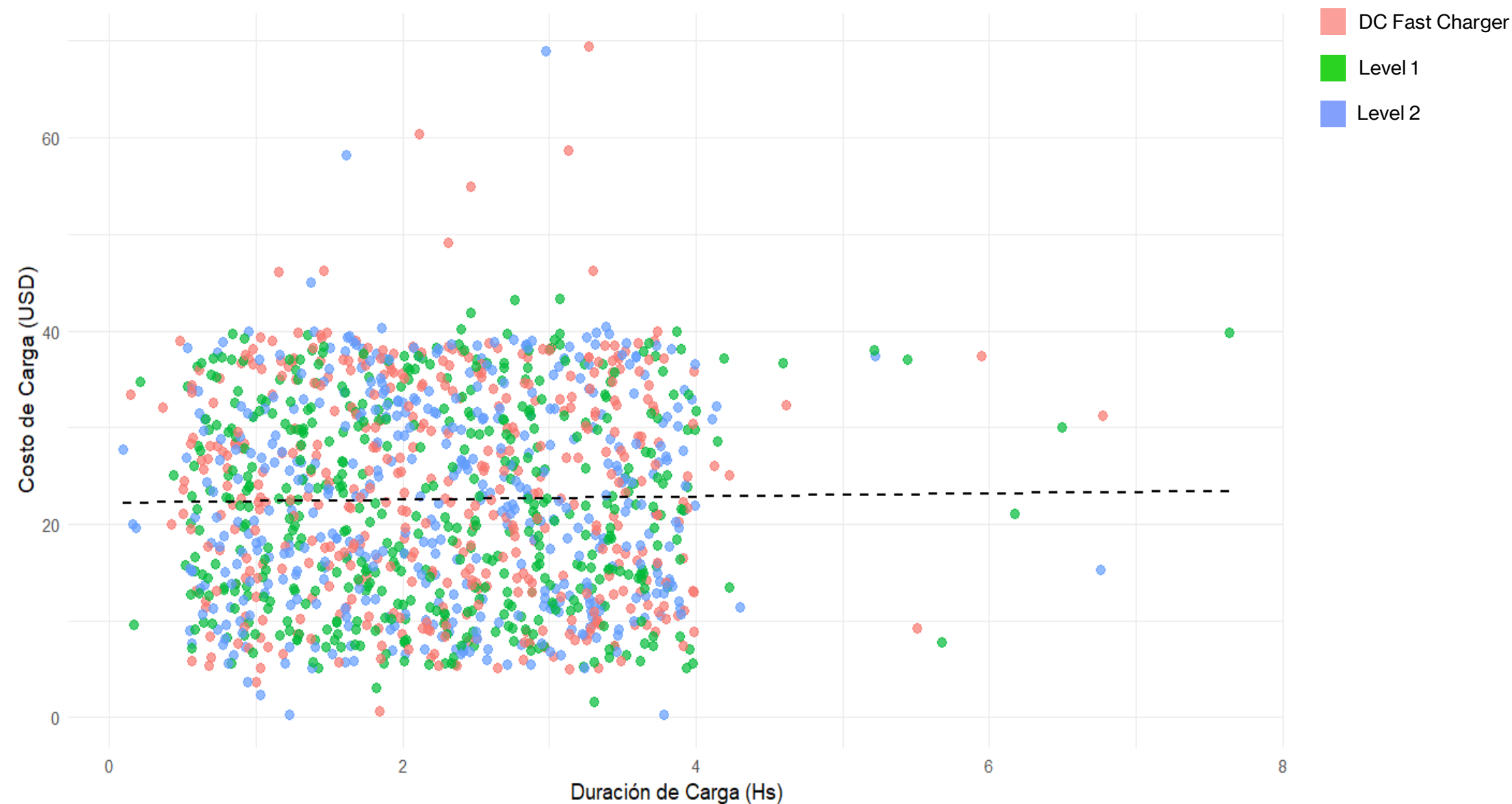
Observaciones.

Cómo podemos observar por el histograma y los datos de la media y la mediana, la estructura es bastante simétrica, sin embargo, al tener un coeficiente de variación alto encontramos gran dispersión respecto a la media. También otro factor a favor de la dispersión de los datos es el mínimo y el máximo de 0,1 y 7,64 respectivamente.

Conclusiones

Teniendo en cuenta ambos cuartiles podemos concluir que el 50% de los conductores cargan su vehículo entre 1.4 y 3.11 horas.

Relación entre la Duración de Carga y el Costo de Carga por Tipo de Estación



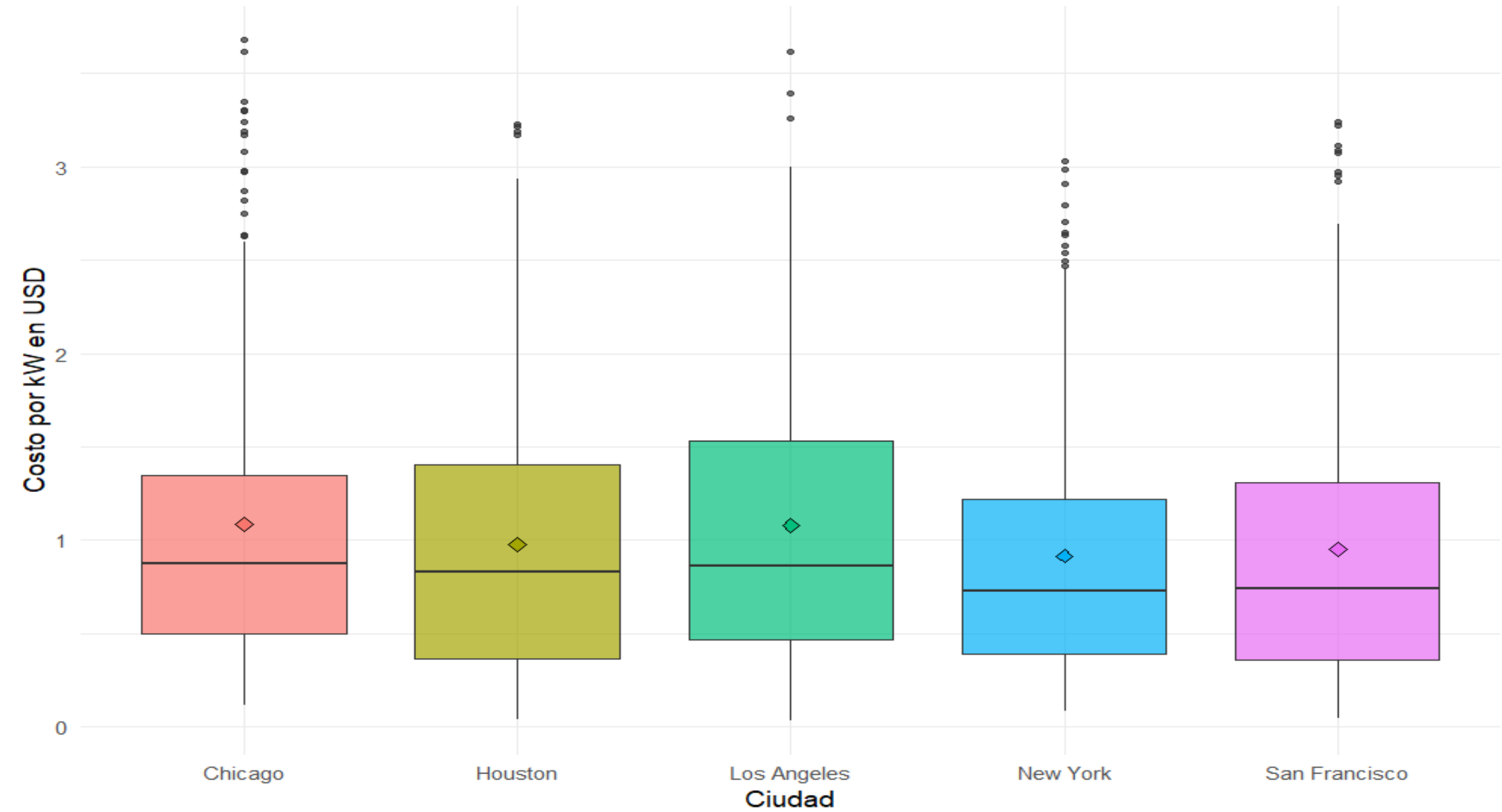
¿Qué nos muestra este gráfico?

En este gráfico observamos un diagrama de dispersión entre la duración de la carga medida en horas y el costo de la carga, filtrado por tipo de cargador.

Conclusiones

Cómo podemos observar por la línea de regresión de manera horizontal, no encontramos ninguna relación positiva o negativa entre estas variables. Tampoco parece haber una acumulación específica por tipo de cargador en determinadas zonas del gráfico. Por lo que podemos concluir que la duración de la carga no tiene implicancia en el costo de la misma ni en el tipo de cargador utilizado.

Distribución del costo por kW por Ciudad



¿Qué nos muestra este gráfico?

En este gráfico vemos reflejado el costo por kilovatio agrupado por ciudad. Esta variable fue calculada realizando el siguiente calculo:

Costo de Carga / Energía Utilizada

Para este grafico eliminamos los valores negativos y también eliminamos los valores atípicos aplicando la siguiente formula: $Q3 + 1.5 * IQR (Q3 - Q1)$

Observaciones.

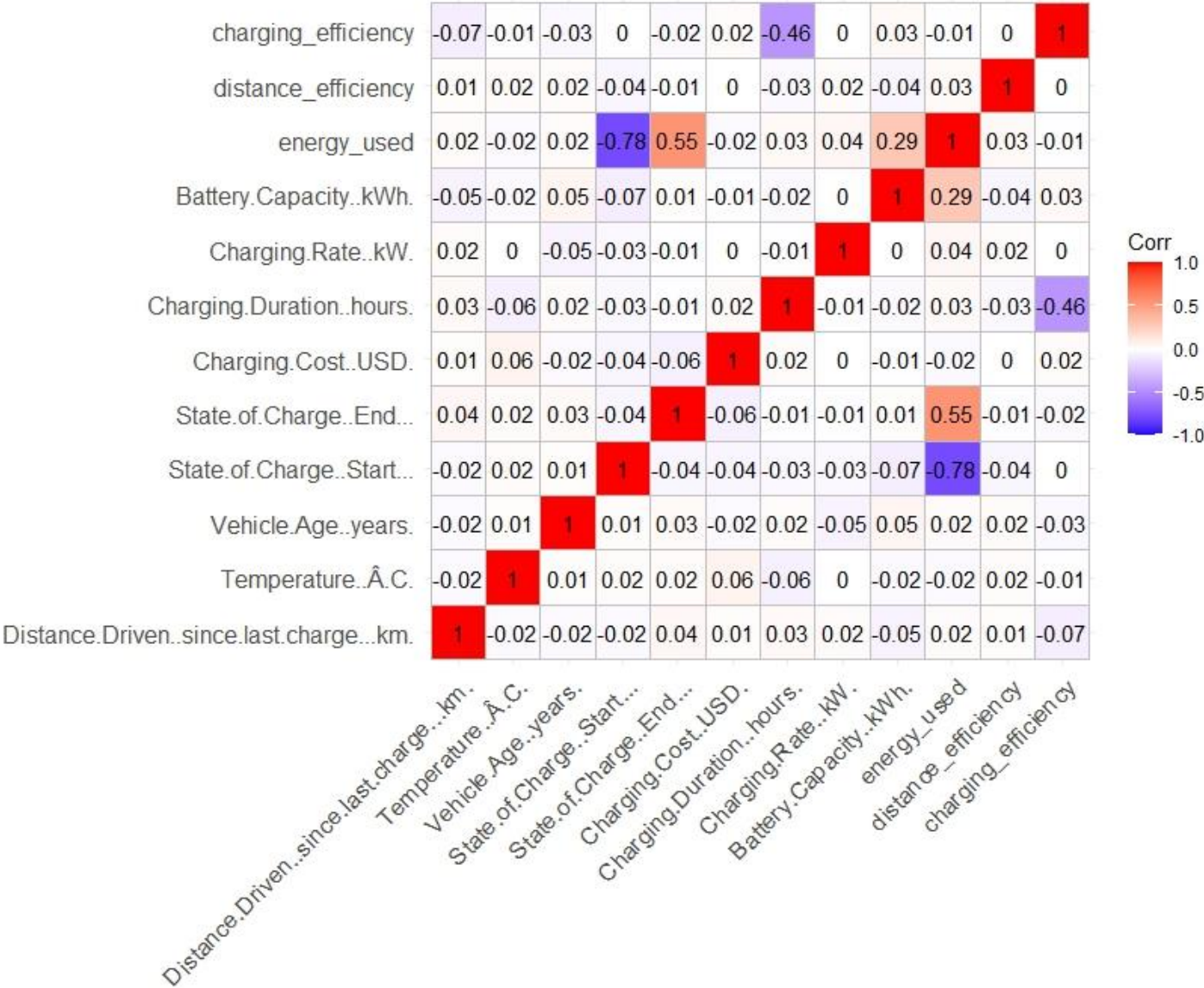
Para todos los casos la media es levemente superior al promedio por lo que podemos encontrar un sesgo positivo.

Por otro lado, el cálculo del coeficiente de variación para todos los casos ronda entre 70% y 80%, lo que nos indica una gran dispersión en los datos con respecto a la media.

Conclusiones

Basándonos en la mediana como mejor indicador debido a la gran dispersión de los datos, podemos concluir que el mejor destino en cuanto a precio para cargar energía es Nueva York, con una mediana de 0.74. Mientras que el más caro es Chicago, con una mediana de 0.86.

Matriz de correlación



Matriz de correlación

En la matriz observamos relaciones que en primera instancia parecen ser interesantes, sin embargo, estas relaciones parten de aquellas variables que fuimos creando a lo largo del análisis exploratorio de datos, cómo energía utilizada o eficiencia energética. A la hora de analizar estas relaciones podemos rápidamente entender que parten de la forma en la que son calculadas y no de una correlación real o lógica entre las variables que pueda ser explicada de manera real.

Para el resto de las variables que ya vienen incorporadas de manera predeterminada en la base de datos, podemos ver que no existe una relación fuerte o significativa dentro de la matriz.

Por ejemplo, en este diagrama de dispersión podemos ver la correlación entre energía usada y estado de carga al inicio, que no es más que el resultado de la disminución de la variable SOC Start en la manera de calcular energía utilizada:

Resto Energía = $((\text{SOC End} - \text{SOC Start}) * 100))$

Energía Usada = $(\text{Capacidad de la batería}) * (\text{Resto Energía})$

