

Prueba técnica

La tabla nos da las siguientes columnas:

'distance', 'consume', 'speed', 'temp_inside', 'temp_outside', 'specials', 'gas_type', 'AC', 'rain', 'sun', 'refill liters', 'refill gas'

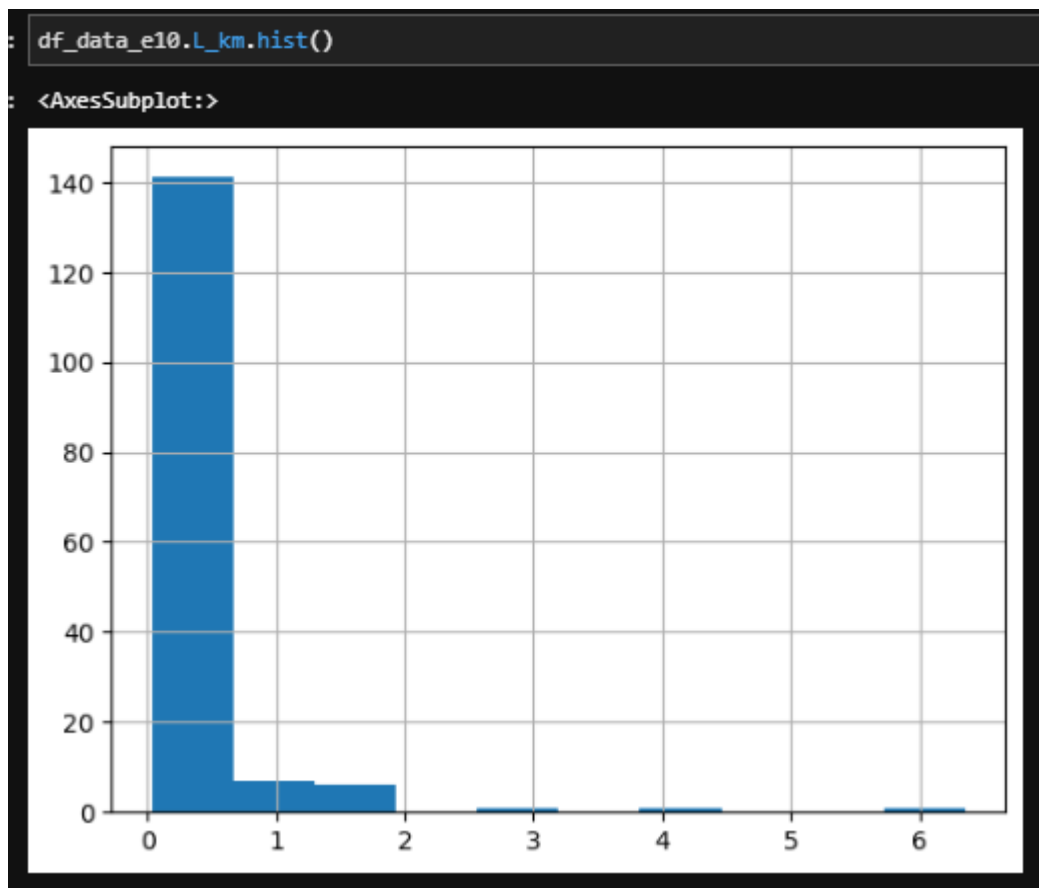
pasos realizados:

- Análisis de estatus de la tabla
- Borrado de columnas con muchos nulos ('refill liters', 'refill gas', 'specials')
- Borrado de celdas con nulos (solo fueron 12 líneas dejándonos 387 correctas)

Para el análisis de si ahí diferencia entre el uso de E10 Y SP98.

Vamos a realizar una columna donde se divida el consumo entre la distancia, para tener un parámetro de eficiencia de gas, y vamos a separar en dos tablas, una de solo E10 y otra de SP98, y las analizaremos.

Aquí graficamos E10(L/Km):



Podemos definir que tenemos valores atípicos, mayores a 2, los cuales los vamos a identificar y revisar.

Tabla de valores atípicos tanto de E10 Y SP98

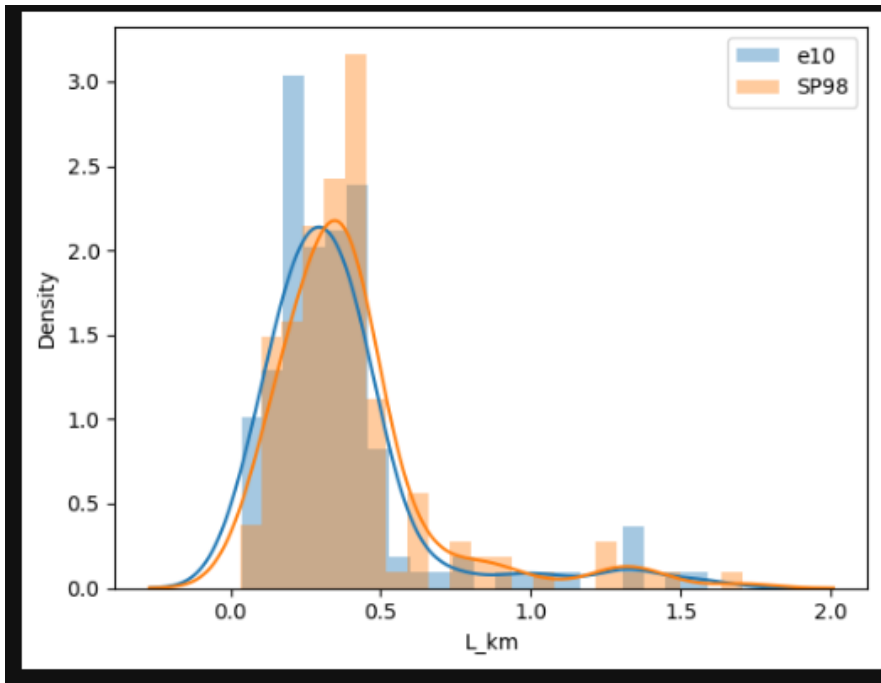
df_data_e10[df_data_e10.L_km > 2]											
	distance	consume	speed	temp_inside	temp_outside	gas_type	AC	rain	sun	L/km	L_km
147	1.7	10.8	14	21,5	10	E10	0	1	0	6.352941	6.352941
254	2.0	8.1	20	22,5	14	E10	0	0	0	4.050000	4.050000
258	2.0	6.0	22	22,5	14	E10	0	0	0	3.000000	3.000000

df_data_sp98[df_data_sp98.L_km > 2]											
	distance	consume	speed	temp_inside	temp_outside	gas_type	AC	rain	sun	L/km	
45	2.0	9.9	21	21,5	4	SP98	0	0	0	4.950000	
60	2.9	7.9	18	21,5	9	SP98	1	1	0	2.724138	
61	2.1	12.2	23	21,5	9	SP98	1	1	0	5.809524	
71	2.0	6.2	20	21,5	10	SP98	0	0	0	3.100000	
198	2.9	7.4	24	21,5	14	SP98	0	0	0	2.551724	
206	2.4	9.0	26	20	10	SP98	0	0	0	3.750000	
210	1.3	11.5	21	20	10	SP98	0	0	0	8.846154	

Podemos observar que la temperatura exterior era muy baja, y muy distinta al interior del auto, al igual que la distancia recorrida fue mínima, lo que nos podría hacer pensar que se tenía prendido el calefactor, y quizás se mantuvo quieto un rato el auto con la calefacción prendida, por ello es por lo que se consumió más gas.

Nota: Por este análisis, vamos a quitar estos valores atípicos, de las tablas de E10 Y S98, para continuar con su análisis.

Grafica de eficiencia (L/km) de E10 Y SP98



Aquí podemos ver visualmente una similitud entre el rendimiento de gas, de E10 y SP98.

Lo vamos a apoyar con análisis matemático.

Vamos a calcular un coeficiente de correlación con el valor p asociado.

```
stats.kendalltau(df_data_e10_2.L_km, df_data_sp98_2_eje.L_km)
KendalltauResult(correlation=0.06005191014138457, pvalue=0.2700033124378367)

stats.spearmanr(df_data_e10_2.L_km, df_data_sp98_2_eje.L_km)
SpearmanrResult(correlation=0.0844485775157058, pvalue=0.29773521670111514)

stats.pearsonr(df_data_e10_2.L_km, df_data_sp98_2_eje.L_km)
PearsonRResult(statistic=0.021834607428596997, pvalue=0.788095299048337)
```

NO podemos rechazar la hipótesis nula de que no hay relación y concluir que NO hay dependencia (tienden a ser similares)

Calculamos la prueba T para las medias de dos muestras independientes

```
stats.ttest_rel(df_data_e10_2.L_km, df_data_sp98_2_eje.L_km)  
Ttest_relResult(statistic=-1.2735132660849164, pvalue=0.20476733452624335)
```

No podemos rechazar la hipótesis nula y podemos concluir que las medias el consumo de gasolina/distancia (L/Km) (es decir que tienden a ser iguales)

Conclusión:

Podemos definir que la eficiencia del GAS E10, es similar al de SP98.

Nota: revisar en la carpeta de data, el código prueba, para revisar el análisis paso a paso.