

```
import math
import statistics
import scipy.stats
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import re
from sklearn import preprocessing
import seaborn as sns
from unicode import unicode
```

## Cargar Datos

```
with pd.ExcelFile("/datasets/luís-josé-molina-truyot/ISA_EQUIPOS/Dataset_Historico-CTs.xlsx") as xls:
    dfhist = pd.read_excel(xls, "Hoja1")
```

```
with pd.ExcelFile("/datasets/luís-josé-molina-truyot/ISA_EQUIPOS/Datos-CTS_Sin_Tap.xlsx") as xls:
    df_familias = pd.read_excel(xls, "Equipos")
df_familias
```

	Equipo int64 100171 - 10031587	Activo fijo float64 8005847.0 - 34003...	Fabricante object ABB ..... 31.9% 27 others ..... 67.6% Missing ..... 0.5%	Denomin.tipo obj... QDR-245 ..... 18.6% 60 others ..... 69.5% Missing ..... 11.9%	Fabr. N°-serie obj... 1HSE 88556... .. 0.1% 1253 others .... 89.7% Missing ..... 10.2%	Denominación ob... CT 230 kV-B ... 11.8% CT 230 kV-C ... 11.7% 98 others ..... 76.5%	Inic.gar.prov. dat... 1987-11-01 00:00:0...	FinGarProv dateti...
0	2008286	8039843	ABB	IMB 123	1HSE 8855622	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
1	2008280	8039837	ABB	IMB 123	1HSE 8855616	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
2	2008283	8039840	ABB	IMB 123	1HSE 8855619	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
3	2004660	8037510	ABB	IMB 123	1HSE8827803	CT 110 kV-A	NaT	NaT
4	2008236	8039793	ABB	IMB 123	1HSE 8855625	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
5	2008245	8039802	ABB	IMB 123	1HSE 8855634	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
6	2008239	8039796	ABB	IMB 123	1HSE 8855628	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
7	2008242	8039799	ABB	IMB 123	1HSE 8855631	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
8	2008198	8039755	ABB	IMB 123	1HSE 8855637	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
9	2008287	8039844	ABB	IMB 123	1HSE 8855623	CT 110 kV-B	2016-10-28 00:00:00	2019-10-28 00:00:00

## Homologar familias

```
replacements_familia = {'ABB LTDA': 'ABB', 'GEC ALSTHOM': 'GEC ALSTHOM BALTEAU', 'ALSTOM': 'ALSTHOM', 'ARTECHE'
                        'ARTECHE': 'ARTECHE', 'T
replacements_subfamilia = {'IMB 123': 'IMB123', 'AK 123/2': 'AK123/2', 'AKOF 245': 'AKOF245', 'IMB 245': 'IMB245', 'IMB 123': 'IMB123', 'IOSK 145':

df_familias['Fabricante'] = df_familias['Fabricante'].replace(replacements_familia)
df_familias['Denomin.tipo'] = df_familias['Denomin.tipo'].replace(replacements_subfamilia)
df_familias
```

	Equipo int64 100171 - 10031587	Activo fijo float64 8005847.0 - 34003...	Fabricante object ABB ..... 34.9% 18 others ..... 64.6% Missing ..... 0.5%	Denomin.tipo obj... IMB245 ..... 22.8% 51 others ..... 65.4% Missing ..... 11.9%	Fabr. N°-serie obj... 1HSE 88556... .. 0.1% 1253 others .... 89.7% Missing ..... 10.2%	Denominación ob... CT 230 kV-B ... 11.8% CT 230 kV-C ... 11.7% 98 others ..... 76.5%	Inic.gar.prov. dat... 1987-11-01 00:00:0...	FinGarProv dateti...
0	2008286	8039843	ABB	IMB123	1HSE 8855622	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
1	2008280	8039837	ABB	IMB123	1HSE 8855616	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
2	2008283	8039840	ABB	IMB123	1HSE 8855619	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
3	2004660	8037510	ABB	IMB123	1HSE8827803	CT 110 kV-A	NaT	NaT

4	2008236	8039793	ABB	IMB123	1HSE 8855625	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
5	2008245	8039802	ABB	IMB123	1HSE 8855634	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
6	2008239	8039796	ABB	IMB123	1HSE 8855628	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
7	2008242	8039799	ABB	IMB123	1HSE 8855631	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
8	2008198	8039755	ABB	IMB123	1HSE 8855637	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
9	2008287	8039844	ABB	IMB123	1HSE 8855623	CT 110 kV-B	2016-10-28 00:00:00	2019-10-28 00:00:00

Funciones

```
def clean_colfam(dataset):
    col_f=list(dataset.columns)
    lista=[]
    for columna in col_f:
        columna=re.sub(r"\.", "", columna)
        columna= unicode(columna)

        lista.append(columna)
    diccion=dict(zip(col_f,lista))
    data_nuevo=dataset.rename(diccion, axis=1)
    return data_nuevo
```

```
def clean_col(dataset):
    col_f=list(dataset.columns)
    lista=[]
    for columna in col_f:
        columna=re.sub(r"\s+", "", columna)
        columna= unicode(columna)

        lista.append(columna)
    diccion=dict(zip(col_f,lista))
    data_nuevo=dataset.rename(diccion, axis=1)
    return data_nuevo
```

```
def Normalization_tabla(df,columna):
    df[columna] = df[columna].str.strip()
    df[columna] = df[columna].str.lower()
    df[columna] = df[columna].apply(lambda x: x.replace (" ", "_"))
    df[columna] = df[columna].apply(lambda x: unicode(x))
    return df
```

```
def col_info(dataset):
    col_f=list(dataset.columns)
    lista=[]
    for columna in col_f:
        lista.append(col_f.index(columna))

    return dict(zip(lista,col_f))
```

EXTRACCIÓN DE DATOS RELEVANTES

dfhist								
	<b>Equipo</b> int64 100171 - 10031587	<b>Aviso</b> float64 45702.0 - 2000002...	<b>Creado el</b> datetim... 1999-09-29 00:00:...	<b>Hora</b> object	<b>Posición medida</b> o..	<b>Denominación</b> ob...	<b>Valor medido</b> floa...	<b>Unidad ámb.med.</b> o
10	100171	nan	2017-06-28 00:00:00	12:59:24	CT TOTAL	Inspección visual general	nan	nan
11	100171	nan	2017-06-28 00:00:00	12:59:24	CT TOTAL	Fuga de aceite	nan	nan

12	100171	nan	2017-06-28 00:00:00	12:59:24	CT TOTAL	Nivel de aceite	nan	nan
13	100171	nan	2017-05-24 00:00:00	16:48:05	CT TOTAL	Inspección visual general	nan	nan
14	100171	nan	2017-05-24 00:00:00	16:48:05	CT TOTAL	Fuga de aceite	nan	nan
15	100171	nan	2017-05-24 00:00:00	16:48:05	CT TOTAL	Nivel de aceite	nan	nan
16	100171	nan	2017-03-23 00:00:00	15:39:47	CT TOTAL	Inspección visual general	nan	nan
17	100171	nan	2017-03-23 00:00:00	15:39:47	CT TOTAL	Fuga de aceite	nan	nan
18	100171	nan	2017-03-23 00:00:00	15:39:46	CT TOTAL	Nivel de aceite	nan	nan
19	100171	365157	2015-11-04 00:00:00	11:42:02	CAPACITANCIA TOTAL	Tip Up Capacitancia...	0.001	1

# NORMALIZACIÓN DE NOMBRES VARIABLES

```
df=clean_col(dfhist)
```

```
df_fam=clean_colfam(df_familias)
```

```
df=Normalization_tabla(df, 'Denominacion')
```

SQL

Saved to variable df\_1

```
SELECT Equipo,Creadoel,Denominacion, Valormedido,"Codif.txt.cod." FROM df
```

	<div>Equipo int64 100171 - 10031587</div>	<div>Creadoel datetim... 1999-09-29 00:00:...</div>	<div>Denominacion ob... calificacion_... 11.6% inspeccion_... 10.2% 25 others ..... 78.2%</div>	<div>Valormedido float...</div>	<div>Codif.txt.cod. obj...</div>	
0	100171	2020-11-27 00:00:00	calificacion_del_eq uipo	3	None	
1	100171	2020-11-27 00:00:00	medida_resistencia _aislamiento_nucl...	21	None	
2	100171	2020-11-27 00:00:00	medida_resistencia _aislamiento_nucl...	23	None	
3	100171	2020-11-27 00:00:00	medida_resistencia _aislamiento_nucl...	28	None	
4	100171	2020-11-27 00:00:00	medida_resistencia _aislamiento_nucl...	50	None	
5	100171	2020-11-24 00:00:00	calificacion_del_eq uipo	5	None	
6	100171	2020-11-24 00:00:00	factor_de_potencia _2.5kv_ctotal	0.336	None	
7	100171	2020-11-24 00:00:00	factor_de_potencia _10kv_ctotal	0.34	None	
8	100171	2020-11-24 00:00:00	capacitancia_total	794.65	None	
9	100171	2020-10-14 00:00:00	calificacion_del_eq uipo	5	None	

SQL



Saved to variable df\_2

```
SELECT COALESCE("Valormedido", "Codif.txt.cod.") as "Valormedido" FROM df
```

	<div>Valormedido object Bueno ..... 12% Normal ..... 9.9% 4732 others ..... 78%</div>	
0	3.0	
1	21.0	

2	23.0	
3	28.0	
4	50.0	
5	5.0	
6	0.336	
7	0.34	
8	794.65	
9	5.0	

```
df_1['Valormedido']=df_2['Valormedido']
df=df_1.copy()
df=df.drop('Codif.txt.cod.',axis=1)
df
```

	<div>Equipo  int64 100171 - 10031587</div> 	<div>Creadoel  datetim... 1999-09-29 00:00:...</div> 	<div>Denominacion  ob... calificacion_... 11.6% inspeccion_... 10.2% 25 others ..... 78.2%</div>	Valormedido  object	
0	100171	2020-11-27 00:00:00	calificacion_del_eq uipo	3.0	
20	100171	2015-11-04 00:00:00	factor_de_potencia _2.5kv_ctotal	0.185	
21	100171	2015-11-04 00:00:00	factor_de_potencia _10kv_ctotal	0.184	
22	100171	2015-11-04 00:00:00	capacitancia_total	794.98	
23	100171	2015-10-28 00:00:00	medida_resistencia _aislamiento_nucl...	229.0	
24	100171	2015-10-28 00:00:00	medida_resistencia _aislamiento_nucl...	160.0	
26	100171	2015-10-28 00:00:00	medida_resistencia _aislamiento_nucl...	174.0	
19	100171	2015-11-04 00:00:00	tip_up_capacitanci a_ctotal	0.001	
27	100171	2012-09-11 00:00:00	medida_resistencia _aislamiento_nucl...	742.0	
29	100171	2012-09-11 00:00:00	medida_resistencia _aislamiento_nucl...	697.0	

Homologar denominaciones repetidas

```
df['Denominacion'].replace(['factor_de_potencia_10_kv_ctotal', 'factor_de_potencia 10kv_ctotal', 'factor_de_potencia_2.5_kv_total', 'factor_de
list(df['Denominacion'].unique())
```

```
[ 'calificacion_del_equipo',
  'medida_resistencia_aislamiento_nucleo_4',
  'medida_resistencia_aislamiento_nucleo_3',
  'medida_resistencia_aislamiento_nucleo_2',
  'medida_resistencia_aislamiento_nucleo_1',
  'factor_de_potencia_2.5kv_ctotal',
  'factor_de_potencia_10kv_ctotal',
  'capacitancia_total',
  'inspeccion_visual_general',
  'fuga_de_aceite',
  'nivel_de_aceite',
  'tip_up_capacitancia_ctotal',
  'severidad_por_termografia',
  'medida_resistencia_aislamiento_nucleo_6',
  'medida_resistencia_aislamiento_nucleo_5',
  'inspeccion_diafragma',
  'estado_de_la_porcelana',
  'estado_cajetin',
  'nivel_de_aceite_no_usar',
  'nivel_de_aceite_-_no_usar',
  'tip_up_capacitancia_total' ]
```

ETIQUETAR EQUIPOS : NUEVOS o VIEJOS

El criterio para evaluar si un equipo es nuevo o viejo es la cantidad de registros por cada valor de medida tomado.

EQUIPO VIEJO : Si tiene mas un registro en distintos años por cada valor de medida

EQUIPO NUEVO : si tiene solo un unico registro por cada valor de medida

TOMAR UNICAMENTE EL VALOR MAS RECIENTE DE CADA AÑO EN CASO DE QUE SE REPITAN MEDIDAS EN EL MISMO AÑO E INCLUSO EL MISMO MES

TABLA DATOS RECIENTES

Ya con los datos correctamente etiquetados, procedemos a crear la tabla de datos recientes, por equipo.

Aquí tomamos los dos últimos datos recientes de cada equipo y en caso de que tengan solo uno, se agrega.



SQL

Saved to variable df\_4

```
SELECT Equipo,Creadoe1 ,Denominacion, Valormedido
FROM (
  SELECT Equipo, Denominacion,Creadoe1, Valormedido,
    ROW_NUMBER() OVER (PARTITION BY Equipo, Denominacion ORDER BY Creadoe1 DESC) AS rn
  FROM df
) t
WHERE rn <= 2
ORDER BY Equipo, Denominacion, Creadoe1 DESC
```



	<div>Equipo int64 100171 - 10031587</div>	<div>Creadoe1 datetim... 2001-08-03 00:00:...</div>	<div>Denominacion ob... calificacion... 10.4% inspeccion_v... 9.1% 19 others ..... 80.6%</div>	<div>Valormedido object Bueno ..... 11.3% Normal ..... 8.8% 4358 others ... 79.9%</div>	
0	100171	2020-11-27 00:00:00	calificacion_del_eq uipo	3.0	
1	100171	2020-11-24 00:00:00	calificacion_del_eq uipo	5.0	
2	100171	2020-11-24 00:00:00	capacitancia_total	794.65	
3	100171	2015-11-04 00:00:00	capacitancia_total	794.98	
4	100171	2020-11-24 00:00:00	factor_de_potencia _10kv_ctotal	0.34	
5	100171	2015-11-04 00:00:00	factor_de_potencia _10kv_ctotal	0.184	
6	100171	2020-11-24 00:00:00	factor_de_potencia _2.5kv_ctotal	0.336	
7	100171	2015-11-04 00:00:00	factor_de_potencia _2.5kv_ctotal	0.185	
8	100171	2017-06-28 00:00:00	fuga_de_aceite	No	



9	100171	2017-05-24 00:00:00	fuga_de_aceite	No	
---	--------	---------------------	----------------	----	--

df_historicos=df_4.copy() df_historicos					
	<b>Equipo</b> int64 100171 - 10031587 	<b>Creadoel</b> datetim... 2001-08-03 00:00:00... 	<b>Denominacion</b> ob... calificacion... 10.4% inspeccion_v... 9.1% 19 others ..... 80.6%	<b>Valormedido</b> object Bueno ..... 11.3% Normal ..... 8.8% 4358 others ... 79.9%	
13	100171	2015-10-28 00:00:00	medida_resistencia _aislamiento_nucl...	174.0	
12	100171	2020-11-27 00:00:00	medida_resistencia _aislamiento_nucl...	50.0	
17	100171	2015-10-28 00:00:00	medida_resistencia _aislamiento_nucl...	160.0	
10	100171	2017-06-28 00:00:00	inspeccion_visual_g eneral	Bueno	
9	100171	2017-05-24 00:00:00	fuga_de_aceite	No	
8	100171	2017-06-28 00:00:00	fuga_de_aceite	No	
7	100171	2015-11-04 00:00:00	factor_de_potencia _2.5kv_ctotal	0.185	
6	100171	2020-11-24 00:00:00	factor_de_potencia _2.5kv_ctotal	0.336	
5	100171	2015-11-04 00:00:00	factor_de_potencia _10kv_ctotal	0.184	
4	100171	2020-11-24 00:00:00	factor_de_potencia _10kv_ctotal	0.34	

Ahora eliminaremos los registros repetidos de las variables categoricas dado que estas se evaluan por su estado actual mas no por sus historicos. Primero separaremos estos documentos de medidas por cualitativos y cuantitativos, eliminamos los repetidos de los cualitativos y por ultimo unir las tablas para así tener los datos que necesitaremos



```
documentos_interes = ['fuga_de_aceite', 'nivel_de_aceite', 'severidad_por_termografia', 'tip_up_capacitancia_ctotal', 'inspeccion_visual_genera  
df_var_cual = df_historicos[df_historicos['Denominacion'].isin(documentos_interes)].groupby(['Equipo', 'Denominacion']).apply(lambda x: x.so
```

df_var_cual					
	<b>Equipo</b> int64 100171 - 10031587 	<b>Creadoel</b> datetim... 2008-02-07 00:00:00... 	<b>Denominacion</b> ob... calificacion... 21.4% inspeccion_... 19.1% 7 others ..... 59.5%	<b>Valormedido</b> object Bueno ..... 25.8% 5.0 ..... 19.7% 110 others ..... 54.5%	
0	100171	2020-11-27 00:00:00	calificacion_del_eq uipo	3.0	
1	100171	2017-06-28 00:00:00	fuga_de_aceite	No	
2	100171	2017-06-28 00:00:00	inspeccion_visual_g eneral	Bueno	
3	100171	2017-06-28 00:00:00	nivel_de_aceite	Normal	
4	100171	2015-11-04 00:00:00	tip_up_capacitanci a_ctotal	0.001	
5	100172	2020-11-27 00:00:00	calificacion_del_eq uipo	3.0	
6	100172	2017-06-28 00:00:00	fuga_de_aceite	No	
7	100172	2017-06-28 00:00:00	inspeccion_visual_g eneral	Bueno	
8	100172	2017-06-28 00:00:00	nivel_de_aceite	Normal	
9	100172	2015-11-04 00:00:00	tip_up_capacitanci a_ctotal	0.001	

df_var_cuant = df_historicos[~df_historicos['Denominacion'].isin(documentos_interes)] df_var_cuant					
	<b>Equipo</b> int64 100171 - 10031587 	<b>Creadoel</b> datetim... 2001-08-03 00:00:00... 	<b>Denominacion</b> ob... medida_res... 14.9%	<b>Valormedido</b> object 550.0 ..... 0.6%	

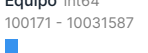
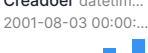
				medida_res... 14.9%	0.19 0.4%
2	100171	2020-11-24 00:00:00	capacitancia_total	794.65	
3	100171	2015-11-04 00:00:00	capacitancia_total	794.98	
4	100171	2020-11-24 00:00:00	factor_de_potencia_10kv_ctotal	0.34	
5	100171	2015-11-04 00:00:00	factor_de_potencia_10kv_ctotal	0.184	
6	100171	2020-11-24 00:00:00	factor_de_potencia_2.5kv_ctotal	0.336	
7	100171	2015-11-04 00:00:00	factor_de_potencia_2.5kv_ctotal	0.185	
12	100171	2020-11-27 00:00:00	medida_resistencia_aislamiento_nucl...	50.0	
13	100171	2015-10-28 00:00:00	medida_resistencia_aislamiento_nucl...	174.0	
14	100171	2020-11-27 00:00:00	medida_resistencia_aislamiento_nucl...	28.0	
15	100171	2015-10-28 00:00:00	medida_resistencia_aislamiento_nucl...	139.0	

df\_datos\_correctos=pd.concat([df\_var\_cual, df\_var\_cuant], axis=0)  
df\_datos\_correctos

	<b>Equipo</b> int64 100171 - 10031587 	<b>Creadoel</b> datetim... 2001-08-03 00:00:00... 	<b>Denominacion</b> ob... medida_res... 10.4% medida_res... 10.4% 19 others ..... 79.2%	<b>Valormedido</b> object Bueno ..... 7.8% 5.0 ..... 6% 4347 others ... 86.2%	
0	100171	2020-11-27 00:00:00	calificacion_del_eq uipo	3.0	
1	100171	2017-06-28 00:00:00	fuga_de_aceite	No	
2	100171	2017-06-28 00:00:00	inspeccion_visual_g eneral	Bueno	
3	100171	2017-06-28 00:00:00	nivel_de_aceite	Normal	
4	100171	2015-11-04 00:00:00	tip_up_capacitanci a_ctotal	0.001	
5	100172	2020-11-27 00:00:00	calificacion_del_eq uipo	3.0	
6	100172	2017-06-28 00:00:00	fuga_de_aceite	No	
7	100172	2017-06-28 00:00:00	inspeccion_visual_g eneral	Bueno	
8	100172	2017-06-28 00:00:00	nivel_de_aceite	Normal	
9	100172	2015-11-04 00:00:00	tip_up_capacitanci a_ctotal	0.001	

Ahora, lo que haremos es mirar cuales son los equipos que poseen dos registros por documento y los que solo tienen uno y los clasificaremos como viejo y nuevo respectivamente. Usaremos esta información para separar estos equipos en dos dataframes distintos.

counts = df\_datos\_correctos.groupby(["Equipo", "Denominacion"]).size().reset\_index(name="num\_registros")  
  
# Identificar equipos nuevos y viejos  
nuevos = counts.groupby("Equipo")["num\_registros"].max() == 1  
viejos = counts.groupby("Equipo")["num\_registros"].max() > 1  
  
# Crear una nueva columna en el DataFrame original para etiquetar los equipos como nuevos o viejos  
df\_datos\_correctos["estado\_equipo"] = "nuevo"  
df\_datos\_correctos.loc[df\_datos\_correctos["Equipo"].isin(viejos[viejos].index), "estado\_equipo"] = "viejo"  
  
# Mostrar la tabla con la nueva columna de estado de equipo  
df\_datos\_correctos

	<b>Equipo</b> int64 100171 - 10031587 	<b>Creadoel</b> datetim... 2001-08-03 00:00:00... 	<b>Denominacion</b> ob... medida_res... 10.4% medida_res... 10.4% 19 others ..... 79.2%	<b>Valormedido</b> object Bueno ..... 7.8% 5.0 ..... 6% 4347 others ... 86.2%	<b>estado_equipo</b> o... viejo ..... 84.6% nuevo ..... 15.4%	
0	100171	2020-11-27 00:00:00	calificacion_del_eq uipo	3.0	viejo	

1	100171	2017-06-28 00:00:00	fuga_de_aceite	No	viejo	
2	100171	2017-06-28 00:00:00	inspeccion_visual_g eneral	Bueno	viejo	
3	100171	2017-06-28 00:00:00	nivel_de_aceite	Normal	viejo	
4	100171	2015-11-04 00:00:00	tip_up_capacitanci a_ctotal	0.001	viejo	
5	100172	2020-11-27 00:00:00	calificacion_del_eq uipo	3.0	viejo	
6	100172	2017-06-28 00:00:00	fuga_de_aceite	No	viejo	
7	100172	2017-06-28 00:00:00	inspeccion_visual_g eneral	Bueno	viejo	
8	100172	2017-06-28 00:00:00	nivel_de_aceite	Normal	viejo	
9	100172	2015-11-04 00:00:00	tip_up_capacitanci a_ctotal	0.001	viejo	

```
# Crear dos DataFrames separados para equipos nuevos y viejos
equipos_nuevos = df_datos_correctos[df_datos_correctos["Equipo"].isin(nuevos[nuevos].index)].copy()
equipos_viejos = df_datos_correctos[df_datos_correctos["Equipo"].isin(viejos[viejos].index)].copy()

# Mostrar las tablas separadas

equipos_nuevos
```

	<b>Equipo</b> int64 108808 - 10031587	<b>Creadoel</b> datetim... 2002-09-17 00:00:...	<b>Denominacion</b> ob... calificacion... 12.4% inspeccion_v... 8.7% 16 others ..... 78.8%	<b>Valormedido</b> object 5.0 ..... 11.2% Bueno ..... 10.1% 1167 others .... 78.7%	<b>estado_equipo</b> o... nuevo ..... 100%	
557	108808	2020-10-14 00:00:00	calificacion_del_eq uipo	5.0	nuevo	
558	108808	2011-03-16 00:00:00	tip_up_capacitanci a_ctotal	0.0	nuevo	
559	108810	2020-10-14 00:00:00	calificacion_del_eq uipo	5.0	nuevo	
560	108810	2011-03-16 00:00:00	tip_up_capacitanci a_ctotal	0.0	nuevo	
593	109038	2020-10-14 00:00:00	calificacion_del_eq uipo	5.0	nuevo	
594	109038	2014-07-20 00:00:00	tip_up_capacitanci a_ctotal	0.029	nuevo	
753	112213	2023-02-25 00:00:00	calificacion_del_eq uipo	5.0	nuevo	
754	112213	2023-02-24 00:00:00	fuga_de_aceite	No	nuevo	
755	112213	2023-02-24 00:00:00	inspeccion_visual_g eneral	Bueno	nuevo	
756	112213	2023-02-24 00:00:00	nivel_de_aceite	Normal	nuevo	

Por ultimo creamos las tablas de los equipos nuevos y los equipos viejos tal que las columnas seran cada Documento de medida y los valores seran los valores de medida, indexados por sus equipos. Para el caso de los equipos nuevos, simplemente se aplica una función "pivot" para crear la nueva tabla con las características anteriormente descritas. Para el caso de los equipos viejos seguiran las columnas igual que los para los equipos nuevos pero los valores serán dados por la expresión \$(Valorfinal -valoranterior)/Valoranterior\$ esta formula nos permite ver que tanto ha sido el incremento o decrecimiento de los los valores de medida de cada documento de medida a lo largo del tiempo. A lo largo del notebook se iran agregando ciertas metricas para poder entender mejor estos valores.

Equipos nuevos

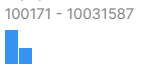
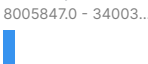
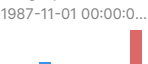
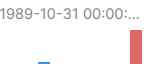
```
df_pivot_equipos_nuevos = equipos_nuevos.pivot(index="Equipo", columns="Denominacion", values="Valormedido")
df_pivot_equipos_nuevos = df_pivot_equipos_nuevos.reset_index()
df_pivot_equipos_nuevos
```

	<b>Equipo</b> int64 108808 - 10031587	<b>calificacion_del_...</b> 5.0 ..... 87.7% 8.0 ..... 7.4% 4 others ..... 4.9%	<b>capacitancia_total</b> ( ... 1479.7 ..... 0.5% 136 others ..... 38.1% Missing ..... 61.4%	<b>estado_cajetin</b> ob... 100.0 ..... 15.1% 0.0 ..... 0.8% Missing ..... 84.1%	<b>estado_de_la_po...</b> 100.0 ..... 15.3% 0.0 ..... 24.4% Missing ..... 83.8%	<b>factor_de_poten...</b> 0.155 ..... 1.4% 64 others ..... 24.4% Missing ..... 74.2%	<b>factor_de_poten...</b> 0.195 ..... 0.8% 48 others ..... 15.6% Missing ..... 83.6%	<b>fuga_de_aceite</b> o... No ..... 65.5% 2 others ..... 1.6% Missing ..... 32.9%
0	108808	5.0	705.12	nan	nan	0.158	0.158	nan



1	108810	5.0	700.93	nan	nan	0.159	0.159	nan
2	109038	5.0	839.731	nan	nan	0.201	0.23	nan
3	112213	5.0	1190.11	nan	nan	0.172	0.186	No
4	112214	5.0	1183.23	nan	nan	0.155	0.168	No
5	112215	5.0	1222.4	nan	nan	0.158	0.168	No
6	118997	5.0	1070.85	nan	100.0	0.126	0.135	No
7	118998	5.0	1058.44	nan	100.0	0.123	0.133	No
8	118999	5.0	1088.14	nan	100.0	0.151	0.157	No
9	125454	5.0	964.85	nan	nan	0.185	nan	nan

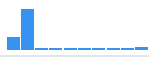
Le agregaremos la columna familia y subfamilia

df_fam								
	<b>Equipo</b> int64 100171 - 10031587 	<b>Activo fijo</b> float64 8005847.0 - 34003... 	<b>Fabricante</b> object ABB ..... 34.9% 18 others ..... 64.6% Missing ..... 0.5%	<b>Denomintipo</b> object IMB245 ..... 22.8% 51 others ..... 65.4% Missing ..... 11.9%	<b>Fabr No-serie</b> obj... 1HSE 88556... .. 0.1% 1253 others .... 89.7% Missing ..... 10.2%	<b>Denominacion</b> ob... CT 230 kV-B ... 11.8% CT 230 kV-C ... 11.7% 98 others ..... 76.5%	<b>Inicgarprov</b> dateti... 1987-11-01 00:00:0... 	<b>FinGarProv</b> dateti... 1989-10-31 00:00:0... 
0	2008286	8039843	ABB	IMB123	1HSE 8855622	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
1	2008280	8039837	ABB	IMB123	1HSE 8855616	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
2	2008283	8039840	ABB	IMB123	1HSE 8855619	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
3	2004660	8037510	ABB	IMB123	1HSE8827803	CT 110 kV-A	NaT	NaT
4	2008236	8039793	ABB	IMB123	1HSE 8855625	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
5	2008245	8039802	ABB	IMB123	1HSE 8855634	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
6	2008239	8039796	ABB	IMB123	1HSE 8855628	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
7	2008242	8039799	ABB	IMB123	1HSE 8855631	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
8	2008198	8039755	ABB	IMB123	1HSE 8855637	CT 110 kV-A	2016-10-28 00:00:00	2019-10-28 00:00:00
9	2008287	8039844	ABB	IMB123	1HSE 8855623	CT 110 kV-B	2016-10-28 00:00:00	2019-10-28 00:00:00

SQL

Saved to variable df\_4

```
SELECT e.*, f.Fabricante, f.Denomintipo
FROM df_pivot_equipos_nuevos e
LEFT JOIN df_fam f ON e.equipo = f.equipo
```

	Equipo int64 108808 - 10031587 	calificacion_del_... 5.0 ..... 87.7% 8.0 ..... 7.4% 4 others ..... 4.9%	capacitancia_total ( ... 1479.7 ..... 0.5% 136 others ..... 38.1% Missing ..... 61.4%	estado_cajetin ob... 100.0 ..... 15.1% 0.0 ..... 0.8% Missing ..... 84.1%	estado_de_la_po... 100.0 ..... 15.3% 0.0 ..... 0.8% Missing ..... 83.8%	factor_de_poten... 0.155 ..... 1.4% 64 others ..... 24.4% Missing ..... 74.2%	factor_de_poten... 0.195 ..... 0.8% 48 others ..... 15.6% Missing ..... 83.6%	fuga_de_aceite o... No ..... 65.5% 2 others ..... 1.6% Missing ..... 32.9%
0	2008286	5.0	1370.766	100.0	100.0	0.137	None	No
1	2008280	5.0	1428.1	100.0	100.0	0.129	None	No
2	2008283	5.0	1363.699	100.0	100.0	0.127	None	No
3	2004660	5.0	1245.6	None	None	0.175	None	No
4	2008236	5.0	1354.6	100.0	100.0	0.136	None	No
5	2008245	5.0	1329.0	100.0	100.0	0.138	None	No
6	2008239	5.0	1314.6	100.0	100.0	0.137	None	No

7	2008242	5.0	1309.2	100.0	100.0	0.14	None	No
8	2008198	5.0	1307.4	100.0	100.0	0.173	None	No
9	2008287	5.0	1378.751	100.0	100.0	0.142	None	No

```
col_info(df_pivot_equipos_nuevos)

{0: 'Equipo',
 1: 'calificacion_del_equipo',
 2: 'capacitancia_total',
 3: 'estado_cajetin',
 4: 'estado_de_la_porcelana',
 5: 'factor_de_potencia_10kv_ctotal',
 6: 'factor_de_potencia_2.5kv_ctotal',
 7: 'fuga_de_aceite',
 8: 'inspeccion_diafragma',
 9: 'inspeccion_visual_general',
10: 'medida_resistencia_aislamiento_nucleo_1',
11: 'medida_resistencia_aislamiento_nucleo_2',
12: 'medida_resistencia_aislamiento_nucleo_3',
13: 'medida_resistencia_aislamiento_nucleo_4',
14: 'medida_resistencia_aislamiento_nucleo_5',
15: 'medida_resistencia_aislamiento_nucleo_6',
16: 'nivel_de_aceite',
17: 'severidad_por_termografia',
18: 'tip_up_capacitancia_ctotal'}
```

Equipos viejos

equipos_viejos						
	<b>Equipo</b> int64 808584 - 808584 	<b>Creadoel</b> datetim... 2013-06-14 00:00:00 	<b>Denominacion</b> ob... medida_resi... 16.7% medida_resi... 16.7% 8 others ..... 66.7%	<b>Valormedido</b> object 5.0 ..... 16.7% Bueno ..... 16.7% 8 others ..... 66.7%	<b>estado_equipo</b> o... viejo ..... 100%	
15686	808584	2021-09-23 00:00:00	medida_resistencia_aislamiento_nucl...	27000.0	viejo	
15687	808584	2021-09-23 00:00:00	medida_resistencia_aislamiento_nucl...	37600.0	viejo	

Para los equipos viejos primero crearemos una tabla tal que tendrá como información, la cantidad porcentual de la degradación de cada documento de medida por cada equipo y esa tabla será la que será pivotada.

Primero separemos los cuantitativos y los cualitativos como hicimos anteriormente

```
documentos_interes = ['fuga_de_aceite', 'nivel_de_aceite', 'severidad_por_termografia', 'tip_up_capacitancia_ctotal', 'inspeccion_visual_genera
```

Cuantitativos

```
df_viejos_var_cuant = df_historicos[~df_historicos['Denominacion'].isin(documentos_interes)]
df_viejos_var_cuant
```

	<b>Equipo</b> int64	<b>Creadoel</b> datetim...	<b>Denominacion</b> ob...	<b>Valormedido</b> object	
15676	808584	2021-09-23 00:00:00	capacitancia_total	773.52	
15683	808584	2022-10-12 00:00:00	medida_resistencia_aislamiento_nucl...	1907.0	
15684	808584	2021-09-23 00:00:00	medida_resistencia_aislamiento_nucl...	21000.0	
15685	808584	2022-10-12 00:00:00	medida_resistencia_aislamiento_nucl...	1802.0	
15686	808584			27000.0	

15687	808584	2021-09-23 00:00:00	medida_resistencia_aislamiento_nucl...	37600.0	
-------	--------	---------------------	--	---------	--

Separaremos las medidas de resistencia de aislamiento puesto que el crietiro de decrecimiento para esta es diferente a las demas. En caso de que se presente comparacion entre valores de 1000 en adelante en ambos, se le asignará 0 a su cambio porcentual que indique que el equipo está bien y si al menos uno de los valores es menor que 100 ya se debe revisar de la manera que sea habia planteado

```
medida_resistencia=['medida_resistencia_aislamiento_nucleo_1','medida_resistencia_aislamiento_nucleo_2','medida_resistencia_aislamiento_nucl
df_viejos_var_cuant_medidas1= df_viejos_var_cuant[df_viejos_var_cuant['Denominacion'].isin(medida_resistencia)]
df_viejos_var_cuant_medidas1
```

	Equipo int64	Creadoel datetim...	Denominacion ob...	Valormedido object	
15683	808584	2022-10-12 00:00:00	medida_resistencia_aislamiento_nucl...	1907.0	
15684	808584	2021-09-23 00:00:00	medida_resistencia_aislamiento_nucl...	21000.0	
15685	808584	2022-10-12 00:00:00	medida_resistencia_aislamiento_nucl...	1802.0	
15686	808584	2021-09-23 00:00:00	medida_resistencia_aislamiento_nucl...	27000.0	
15687	808584	2021-09-23 00:00:00	medida_resistencia_aislamiento_nucl...	37600.0	

```
# Ordenar el dataframe por fecha
grupos_med =df_viejos_var_cuant_medidas1.sort_values(by='Creadoel')

# Obtener los valores más viejos para cada equipo y denominación
df_viejos = grupos_med.groupby(['Equipo', 'Denominacion']).first().reset_index()
df_viejos = df_viejos.rename(columns={'Valormedido': 'valor_inicial'})

# Obtener los valores más nuevos para cada equipo y denominación
df_nuevos = grupos_med.groupby(['Equipo', 'Denominacion']).last().reset_index()
df_nuevos = df_nuevos.rename(columns={'Valormedido': 'valor_final'})

# Unir los dataframes de valores viejos y nuevos
df_resultado = df_viejos.merge(df_nuevos, on=['Equipo', 'Denominacion'])

# Seleccionar las columnas que se desean
grupos_med = df_resultado[['Equipo', 'Denominacion', 'valor_inicial', 'valor_final']]

grupos_med
```

	Equipo int64 100171 - 10031587	Denominacion ob...	valor_inicial object	valor_final object	
0	100171	medida_resistencia_aislamiento_nucl...	174.0	50.0	
1	100171	medida_resistencia_aislamiento_nucl...	139.0	28.0	
2	100171	medida_resistencia_aislamiento_nucl...	160.0	23.0	
3	100171	medida_resistencia_aislamiento_nucl...	229.0	21.0	
4	100172	medida_resistencia_aislamiento_nucl...	403.0	81.0	
5	100172	medida_resistencia_aislamiento_nucl...	350.0	41.0	
6	100172	medida_resistencia_aislamiento_nucl...	267.0	22.0	
7	100172	medida_resistencia_aislamiento_nucl...	675.0	51.0	
8	100173	medida_resistencia_aislamiento_nucl...	1630.0	158.0	
9	100173	medida_resistencia_aislamiento_nucl...	1550.0	123.0	

```
len(grupos_med['valor_final'])
```

```
5579

grupos_med['valor_final'][0]

'50.0'

# Crear columna 'cambio_porcentual'
grupos_med=grupos_med.copy()
grupos_med['cambio_porcentual'] = np.where((grupos_med['valor_inicial'].astype(float) < 1000) | (grupos_med['valor_final'].astype(float) < 1
                                             (grupos_med['valor_final'].astype(float) - grupos_med['valor_inicial'].astype(float))/grupos_med['valor_i
```

grupos_med						
	<div>Equipo int64 100171 - 10031587</div>	<div>Denominacion ob...</div> <div>medida_res... 20.3% medida_res... 20.3% 4 others ..... 59.5%</div>	<div>valor_inicial object</div> <div>51000.0 ..... 0.9% 50000.0 ..... 0.6% 1994 others ... 98.6%</div>	<div>valor_final object</div> <div>550.0 ..... 1.4% 1100.0 ..... 0.4% 2171 others ... 98.2%</div>	<div>cambio_porcent...</div> <div>-0.999709202453...</div>	
0	100171	medida_resistencia _aislamiento_nucl...	174.0	50.0	-0.7126436782	
1	100171	medida_resistencia _aislamiento_nucl...	139.0	28.0	-0.7985611511	
2	100171	medida_resistencia _aislamiento_nucl...	160.0	23.0	-0.85625	
3	100171	medida_resistencia _aislamiento_nucl...	229.0	21.0	-0.9082969432	
4	100172	medida_resistencia _aislamiento_nucl...	403.0	81.0	-0.7990074442	
5	100172	medida_resistencia _aislamiento_nucl...	350.0	41.0	-0.8828571429	
6	100172	medida_resistencia _aislamiento_nucl...	267.0	22.0	-0.9176029963	
7	100172	medida_resistencia _aislamiento_nucl...	675.0	51.0	-0.9244444444	
8	100173	medida_resistencia _aislamiento_nucl...	1630.0	158.0	-0.9030674847	
9	100173	medida_resistencia _aislamiento_nucl...	1550.0	123.0	-0.9206451613	

SQL	Saved to variable df_18					
<pre>SELECT Equipo,Denominacion,cambio_porcentual,valor_final FROM grupos_med GROUP BY Equipo,Denominacion,cambio_porcentual,valor_final</pre>						
	<div>Equipo int64 100171 - 10031587</div>	<div>Denominacion ob...</div> <div>medida_res... 20.3% medida_res... 20.3% 4 others ..... 59.5%</div>	<div>cambio_porcent...</div> <div>-0.999709202453...</div>	<div>valor_final object</div> <div>550.0 ..... 1.4% 1100.0 ..... 0.4% 2171 others ... 98.2%</div>		
0	100171	medida_resistencia _aislamiento_nucl...	-0.7126436782	50.0		
1	100171	medida_resistencia _aislamiento_nucl...	-0.7985611511	28.0		
2	100171	medida_resistencia _aislamiento_nucl...	-0.85625	23.0		
3	100171	medida_resistencia _aislamiento_nucl...	-0.9082969432	21.0		
4	100172	medida_resistencia _aislamiento_nucl...	-0.7990074442	81.0		
5	100172	medida_resistencia _aislamiento_nucl...	-0.8828571429	41.0		
6	100172	medida_resistencia _aislamiento_nucl...	-0.9176029963	22.0		
7	100172	medida_resistencia _aislamiento_nucl...	-0.9244444444	51.0		

8	100173	medida_resistencia_aislamiento_nucl	-0.9030674847	158.0	
9	100173	medida_resistencia_aislamiento_nucl...	-0.9206451613	123.0	

Ahora, solo consideremos el resto de puntos de medida exceptuando las medidas de resistencia de aislamiento

```
medida_resistencia=['medida_resistencia_aislamiento_nucleo_1','medida_resistencia_aislamiento_nucleo_2','medida_resistencia_aislamiento_nucl
df_viejos_var_cuant_sin_med1= df_viejos_var_cuant[~df_viejos_var_cuant['Denominacion'].isin(medida_resistencia)]
df_viejos_var_cuant_sin_med1
```

	Equipo int64 100171 - 10031587	Creadoel datetim... 2008-07-25 00:00:...	Denominacion ob... capacitanci... 40.2% factor_de_p... 30.7% 2 others ..... 29.1%	Valormedido object 0.19 ..... 1.6% 0.2 ..... 1.4% 1697 others ..... 97%	
2	100171	2020-11-24 00:00:00	capacitancia_total	794.65	
3	100171	2015-11-04 00:00:00	capacitancia_total	794.98	
4	100171	2020-11-24 00:00:00	factor_de_potencia_10kv_ctotal	0.34	
5	100171	2015-11-04 00:00:00	factor_de_potencia_10kv_ctotal	0.184	
6	100171	2020-11-24 00:00:00	factor_de_potencia_2.5kv_ctotal	0.336	
7	100171	2015-11-04 00:00:00	factor_de_potencia_2.5kv_ctotal	0.185	
26	100172	2020-11-24 00:00:00	capacitancia_total	793.61	
27	100172	2015-11-04 00:00:00	capacitancia_total	791.41	
28	100172	2020-11-24 00:00:00	factor_de_potencia_10kv_ctotal	0.322	
29	100172	2015-11-04 00:00:00	factor_de_potencia_10kv_ctotal	0.146	

```
# Ordenar el dataframe por fecha
grupo1 =df_viejos_var_cuant_sin_med1.sort_values(by='Creadoel')

# Obtener los valores más viejos para cada equipo y denominación
df_viejos = grupo1.groupby(['Equipo', 'Denominacion']).first().reset_index()
df_viejos = df_viejos.rename(columns={'Valormedido': 'valor_inicial'})

# Obtener los valores más nuevos para cada equipo y denominación
df_nuevos = grupo1.groupby(['Equipo', 'Denominacion']).last().reset_index()
df_nuevos = df_nuevos.rename(columns={'Valormedido': 'valor_final'})

# Unir los dataframes de valores viejos y nuevos
df_resultado = df_viejos.merge(df_nuevos, on=['Equipo', 'Denominacion'])

# Seleccionar las columnas que se desean
grupos = df_resultado[['Equipo', 'Denominacion', 'valor_inicial', 'valor_final']]

grupos
```

	Equipo int64	Denominacion ob...	valor_inicial object	valor_final object	
354	109394	capacitancia_total	1026.09	1035.299	
355	109394	factor_de_potencia_10kv_ctotal	0.125	0.28	
356	109394	factor_de_potencia_2.5kv_ctotal	0.154	0.283	

```
grupo= grupos.copy()
grupo["cambio_porcentual"] = (grupo["valor_final"].astype(float) - grupo["valor_inicial"].astype(float))/grupo["valor_inicial"].astype(float)

grupo
```

	Equipo int64	Denominacion ob...	valor_inicial object	valor_final object	cambio_porcent...
--	--------------	--------------------	----------------------	--------------------	-------------------

354	109394	capacitancia_total	1026.09	1035.299	0.008974846261	
355	109394	factor_de_potencia_10kv_ctotal	0.125	0.28	1.24	
356	109394	factor_de_potencia_2.5kv_ctotal	0.154	0.283	0.8376623377	

SQL

Saved to variable df\_13

SELECT

Equipo,Denominacion,cambio\_porcentual,valor\_final

FROM

grupo

	Equipo int64	Denominacion ob...	cambio_porcent...	valor_final object	
1567	808584	capacitancia_total	0	773.52	

df_viejos_completo_cuanti=pd.concat([df_18,df_13], axis=0) df_viejos_completo_cuanti					
	Equipo int64 100171 - 10031587 	Denominacion ob... medida_resi... 14.7% medida_resi... 14.7% 8 others ..... 70.5%	cambio_porcent... -1.0 - inf 	valor_final object 550.0 ..... 1% 0.2 ..... 0.4% 3249 others ... 98.6%	
0	100171	medida_resistencia_aislamiento_nucl...	-0.7126436782	50.0	
2	100171	factor_de_potencia_2.5kv_ctotal	0.8162162162	0.336	
1	100171	factor_de_potencia_10kv_ctotal	0.847826087	0.34	
0	100171	capacitancia_total	-0.0004151047825	794.65	
1	100171	medida_resistencia_aislamiento_nucl...	-0.7985611511	28.0	
3	100171	medida_resistencia_aislamiento_nucl...	-0.9082969432	21.0	
2	100171	medida_resistencia_aislamiento_nucl...	-0.85625	23.0	
3	100172	capacitancia_total	0.002779848625	793.61	
4	100172	factor_de_potencia_10kv_ctotal	1.205479452	0.322	
5	100172	factor_de_potencia_2.5kv_ctotal	1.393103448	0.347	

Cualitativos

df_viejos_var_cual = df_historicos[df_historicos['Denominacion'].isin(documentos_interes)].groupby(['Equipo', 'Denominacion']).apply(lambda df_viejos_var_cual['Criterio']=df_viejos_var_cual['Valormedido'] df_viejos_var_cual=df_viejos_var_cual.drop(['Valormedido', 'Creadoel'],axis=1) df_viejos_var_cual= df_viejos_var_cual.rename(columns={'Criterio': 'cambio_porcentual'}) df_viejos_var_cual					
	Equipo int64 100171 - 10031587 	Denominacion ob... calificacion... 21.4% inspeccion_... 19.1% 9 others ..... 59.5%	cambio_porcent... Bueno ..... 25.8% 5.0 ..... 19.7% 110 others ..... 54.5%		
0	100171	calificacion_del_equip	3.0		
1	100171	fuga_de_aceite	No		
2	100171	inspeccion_visual_general	Bueno		
3	100171	nivel_de_aceite	Normal		
4	100171	tip_up_capacitancia_ctotal	0.001		

5	100172	calificacion_del_eq	3.0	
6	100172	fuga_de_aceite	No	
7	100172	inspeccion_visual_g	Bueno	
8	100172	nivel_de_aceite	Normal	
9	100172	tip_up_capacitanci	0.001	
		a_ctotal		

Aqui crearemos una tabla ya con los valores porcentuales que nos interesan y además eliminaremos los documentos de medida repetidos

Aqui simplemente se unieron las variables cualitativas con las cuantitativas

Creamos la tabla nueva con los documentos de medida como columnas y sus valores de medida correctos

Tabla pivoteada variables cuantitativas

```
df_pivot_equipos_viejos = df_viejos_completo_cuanti.pivot(index="Equipo", columns="Denominacion", values="cambio_porcentual")
df_pivot_equipos_viejos = df_pivot_equipos_viejos.reset_index()
```

df\_pivot\_equipos\_viejos

	<div>Equipo int64 100171 - 10031587</div>	<div>capacitancia_total f -0.999860045860...</div>	<div>factor_de_poten... -1.0 - 8611.440191...</div>	<div>factor_de_poten... -0.989348659003...</div>	<div>medida_resisten... -0.999 - 4759.869...</div>	<div>medida_resisten... -0.999709202453...</div>	<div>medida_resisten... -0.997916666666...</div>	<div>medida_resisten... -0.998458333333...</div>
0	100171	-0.0004151047825	0.847826087	0.8162162162	-0.7126436782	-0.7985611511	-0.85625	-0.9082969432
1	100172	0.002779848625	1.205479452	1.393103448	-0.7990074442	-0.8828571429	-0.9176029963	-0.9244444444
2	100173	-0.004502019011	1.129032258	1.137254902	-0.9030674847	-0.9206451613	-0.9055944056	-0.6911392405
3	100189	0.06263125386	0.3684210526	0.3684210526	-0.5537974684	-0.1419354839	2.005464481	3.802561366
4	100190	0.06435006435	0.7368421053	0.7368421053	-0.1854636591	-0.72109375	-0.6363636364	-0.2443396226
5	100191	0.05871330418	0.75	0.75	0	0	0	0
6	100195	0.05879458795	0.04	0.06666666667	-0.8688293371	-0.9071428571	-0.8915662651	-0.8945868946
7	100196	0.05987730061	0.272	0.252	0	0	-0.5938864629	0
8	100197	0.06470951792	0.3363636364	0.2608695652	-0.9938042131	-0.9411764706	-0.8575667656	-0.9672131148
9	100207	0.06031407035	0.26	0.2947368421	0	0	0	0

Eliminemos las columnas sobrantes y filas que contengan valores inf; esto ultimo significa la ausencia de valores de medida para documentos de medida

Tabla pivoteada variables cualitativa

```
df_pivot_equipos_viejos_cual = df_viejos_var_cual.pivot(index="Equipo", columns="Denominacion", values="cambio_porcentual")
df_pivot_equipos_viejos_cual
```

	<div>calificacion_del_... 5.0 ..... 77.3% 3.0 ..... 18.2% 2.0 ..... 4.4%</div>	<div>estado_cajetin ob... 100.0 ..... 79% 4 others ..... 13.3% Missing ..... 7.7%</div>	<div>estado_de_la_po... 100.0 ..... 95% 0.0 ..... 4.4% 99.0 ..... 0.6%</div>	<div>fuga_de_aceite o... No ..... 99.4% Missing ..... 0.6%</div>	<div>inspeccion_diafr... Bueno ..... 47.5% Malo ..... 1.7% Missing ..... 50.8%</div>	<div>inspeccion visu... Bueno ..... 100%</div>	<div>nivel_de_aceite o... Normal ..... 99.4% Missing ..... 0.6%</div>	<div>nivel_de_aceite_... Missing ..... 100%</div>
109015	3.0	100.0	100.0	No	Bueno	Bueno	Normal	nan
109041	5.0	100.0	100.0	No	Bueno	Bueno	Normal	nan
111947	5.0	99.0	100.0	No	nan	Bueno	Normal	nan
111948	5.0	99.0	100.0	No	nan	Bueno	Normal	nan
111949	5.0	99.0	100.0	No	nan	Bueno	Normal	nan


112009	3.0	99.0	100.0	No	Bueno	Bueno	Normal	nan
112010	5.0	99.0	100.0	No	Bueno	Bueno	Normal	nan
112011	5.0	99.0	100.0	No	nan	Bueno	Normal	nan
112018	5.0	100.0	100.0	No	nan	Bueno	Normal	nan
112019	3.0	100.0	100.0	No	nan	Bueno	Normal	nan

Llenaremos los datos faltantes de los cualitativos. La ausencia de datos significa que los datos son positivos

```
df_pivot_equipos_viejos_cual['estado_cajetin'].replace(np.nan, 'Bueno', inplace=True)
df_pivot_equipos_viejos_cual['estado_cajetin'].replace(100, 'Bueno', inplace=True)
df_pivot_equipos_viejos_cual['estado_cajetin'].replace(99, 'Bueno', inplace=True)
df_pivot_equipos_viejos_cual['estado_cajetin'].replace(95, 'Bueno', inplace=True)
df_pivot_equipos_viejos_cual['estado_cajetin'].replace(90, 'Bueno', inplace=True)
df_pivot_equipos_viejos_cual['estado_cajetin'].replace(0, 'Malo', inplace=True)
df_pivot_equipos_viejos_cual['estado_de_la_porcelana'].replace(100, 'Bueno', inplace=True)
df_pivot_equipos_viejos_cual['estado_de_la_porcelana'].replace(99, 'Bueno', inplace=True)
df_pivot_equipos_viejos_cual['estado_de_la_porcelana'].replace(0, 'Roto', inplace=True)
df_pivot_equipos_viejos_cual['estado_de_la_porcelana'].replace(np.nan, 'Bueno', inplace=True)
df_pivot_equipos_viejos_cual['fuga_de_aceite'].replace(np.nan, "No", inplace=True)
df_pivot_equipos_viejos_cual['inspeccion_diafragma'].replace(np.nan, "Bueno", inplace=True)
df_pivot_equipos_viejos_cual['inspeccion_visual_general'].replace(np.nan, "Bueno", inplace=True)
df_pivot_equipos_viejos_cual['nivel_de_aceite'].replace(np.nan, "Normal", inplace=True)
df_pivot_equipos_viejos_cual['severidad_por_termografia'].replace(np.nan, 5, inplace=True)
df_pivot_equipos_viejos_cual=df_pivot_equipos_viejos_cual.drop(['nivel_de_aceite_-_no_usar', 'nivel_de_aceite_no_usar'], axis=1)
```

df_pivot_equipos_viejos_cual								
	calificacion_del_...	estado_cajetin ob...	estado_de_la_po...	fuga_de_aceite o...	inspeccion_diafr...	inspeccion visu...	nivel_de_aceite o...	severidad_por_t...
	5.0 ..... 81%	Bueno ..... 85.7%	Bueno ..... 85.4%	No ..... 98%	Bueno ..... 99.7%	Bueno ..... 99.3%	Normal ..... 99.7%	5 ..... 87.8%
	3.0 ..... 13.3%	100.0 ..... 12.1%	100.0 ..... 13.8%	Mancha ..... 1.4%	Malo ..... 0.3%	Regular ..... 0.7%	No aplica ..... 0.2%	5.0 ..... 11.1%
	4 others ..... 5.7%	4 others ..... 2.2%	2 others ..... 0.7%	2 others ..... 0.6%			2 others ..... 0.2%	2 others ..... 1.1%
100171	3.0	Bueno	Bueno	No	Bueno	Bueno	Normal	5
100172	3.0	Bueno	Bueno	No	Bueno	Bueno	Normal	5
100173	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal	5
100189	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal	5
100190	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal	5
100191	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal	5
100195	3.0	Bueno	Bueno	No	Bueno	Bueno	Normal	5
100196	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal	5
100197	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal	5
100207	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal	5

```
df_pivot_equipos_viejos_cual= df_pivot_equipos_viejos_cual.reset_index()
```

df_pivot_equipos_viejos_cual								
	Equipo int64 100171 - 10031587	calificacion_del_...	estado_cajetin ob...	estado_de_la_po...	fuga_de_aceite o...	inspeccion_diafr...	inspeccion visu...	nivel_de_aceite o...
		5.0 ..... 81%	Bueno ..... 85.7%	Bueno ..... 85.4%	No ..... 98%	Bueno ..... 99.7%	Bueno ..... 99.3%	Normal ..... 99.7%
		3.0 ..... 13.3%	100.0 ..... 12.1%	100.0 ..... 13.8%	Mancha ..... 1.4%	Malo ..... 0.3%	Regular ..... 0.7%	No aplica ..... 0.2%
		4 others ..... 5.7%	4 others ..... 2.2%	2 others ..... 0.7%	2 others ..... 0.6%			2 others ..... 0.2%
0	100171	3.0	Bueno	Bueno	No	Bueno	Bueno	Normal
1	100172	3.0	Bueno	Bueno	No	Bueno	Bueno	Normal



2	100173	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal
3	100189	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal
4	100190	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal
5	100191	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal
6	100195	3.0	Bueno	Bueno	No	Bueno	Bueno	Normal
7	100196	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal
8	100197	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal
9	100207	5.0	Bueno	Bueno	No	Bueno	Bueno	Normal

```
df_pivot_equipos_viejos_cual.columns

Index(['Equipo', 'calificacion_del_equipo', 'estado_cajetin',
      'estado_de_la_porcelana', 'fuga_de_aceite', 'inspeccion_diafragma',
      'inspeccion_visual_general', 'nivel_de_aceite',
      'severidad_por_termografia', 'tip_up_capacitancia_ctotal'],
      dtype='object', name='Denominacion')
```

Tabla sin valores finales

tabla\_unida = pd.merge(df\_pivot\_equipos\_viejos, df\_pivot\_equipos\_viejos\_cual, on='Equipo')  
tabla\_unida

	Equipo int64 100171 - 10031587	capacitancia_total f -0.999860045860...	factor_de_poten... -1.0 - 8611.440191...	factor_de_poten... -0.989348659003...	medida_resisten... -0.999 - 4759.869...	medida_resisten... -0.999709202453...	medida_resisten... -0.9979166666666...	medida_resisten... -0.998458333333...
724	809222	-0.9998600459	nan	nan	0	0	0	0
746	810200	-0.9990439537	nan	nan	0	0	0	0
928	2003975	-0.99013714	0.2053571429	nan	0	0	0	0
467	132281	-0.9	-1	inf	0	-0.9997092025	0	0
829	817900	-0.8167254085	nan	nan	0	0	0	0
86	106876	-0.64	8611.440191	0.245508982	0	0	0	0
764	816300	-0.2269718902	nan	nan	10.35135135	10.11111111	11	9
740	809482	-0.209749955	nan	nan	9.310344828	8.155172414	6.810344828	4.948275862
736	809452	-0.2002713704	nan	nan	0.1428571429	0.6	0.8571428571	1.171428571
735	809451	-0.1852119678	nan	nan	-0.1804347826	0.1847826087	0.3369565217	0.5434782609

```
tabla_unida=tabla_unida.drop(['tip_up_capacitancia_ctotal', 'tip_up_capacitancia_total'], axis=1)
```


tabla\_unida

	Equipo int64 100171 - 10031587	capacitancia_total f -0.999860045860...	factor_de_poten... -1.0 - 8611.440191...	factor_de_poten... -0.989348659003...	medida_resisten... -0.999 - 4759.869...	medida_resisten... -0.999709202453...	medida_resisten... -0.9979166666666...	medida_resisten... -0.998458333333...
0	100171	-0.0004151047825	0.847826087	0.8162162162	-0.7126436782	-0.7985611511	-0.85625	-0.9082969432
1	100172	0.002779848625	1.205479452	1.393103448	-0.7990074442	-0.8828571429	-0.9176029963	-0.9244444444

2	100173	-0.004502019011	1.129032258	1.137254902	-0.9030674847	-0.9206451613	-0.9055944056	-0.6911392405
3	100189	0.06263125386	0.3684210526	0.3684210526	-0.5537974684	-0.1419354839	2.005464481	3.802561366
4	100190	0.06435006435	0.7368421053	0.7368421053	-0.1854636591	-0.72109375	-0.6363636364	-0.2443396226
5	100191	0.05871330418	0.75	0.75	0	0	0	0
6	100195	0.05879458795	0.04	0.06666666667	-0.8688293371	-0.9071428571	-0.8915662651	-0.8945868946
7	100196	0.05987730061	0.272	0.252	0	0	-0.5938864629	0
8	100197	0.06470951792	0.3363636364	0.2608695652	-0.9938042131	-0.9411764706	-0.8575667656	-0.9672131148
9	100207	0.06031407035	0.26	0.2947368421	0	0	0	0

Tabla valores finales

SQL	Saved to variable	df_15
SELECT Equipo,Denominacion,Valor_final FROM grupos		
	<div>Equipo int64 100171 - 10031587</div> 	<div>Denominacion ob... capacitanci... 40.4% factor_de_p... 30.7% 2 others ..... 28.9%</div> <div>valor_final object 0.2 ..... 1.6% 0.19 ..... 1.6% 1089 others ... 96.8%</div>
0	100171	capacitancia_total 794.65
1	100171	factor_de_potencia_10kv_ctotal 0.34
2	100171	factor_de_potencia_2.5kv_ctotal 0.336
3	100172	capacitancia_total 793.61
4	100172	factor_de_potencia_10kv_ctotal 0.322
5	100172	factor_de_potencia_2.5kv_ctotal 0.347
6	100173	capacitancia_total 793.83
7	100173	factor_de_potencia_10kv_ctotal 0.33
8	100173	factor_de_potencia_2.5kv_ctotal 0.327
9	100189	capacitancia_total 860.2

SQL	Saved to variable	df_16
SELECT Equipo,Denominacion,Valor_final FROM grupos_med		
	<div>Equipo int64 100171 - 10031587</div> 	<div>Denominacion ob... medida_res... 20.3% medida_res... 20.3% 4 others ..... 59.5%</div> <div>valor_final object 550.0 ..... 1.4% 1100.0 ..... 0.4% 2171 others .... 98.2%</div>
0	100171	medida_resistencia_aislamiento_nucl... 50.0
1	100171	medida_resistencia_aislamiento_nucl... 28.0
2	100171	medida_resistencia_aislamiento_nucl... 23.0
3	100171	medida_resistencia_aislamiento_nucl... 21.0
4	100172	medida_resistencia_aislamiento_nucl... 81.0
5	100172	medida_resistencia_aislamiento_nucl... 41.0
6	100172	medida_resistencia_aislamiento_nucl... 22.0

7	100172	medida_resistencia	51.0	
8	100173	medida_resistencia _aislamiento_nucl...	158.0	
9	100173	medida_resistencia _aislamiento_nucl...	123.0	

```
df_valores_finales_viejos=pd.concat([df_15,df_16])
```


SQL

Saved to variable df\_val\_finales

```
SELECT Equipo,Denominacion,Valor_final
FROM df_valores_finales_viejos
```

	<div>Equipo int64 100171 - 10031587</div> 	<div>Denominacion ob...</div> <div>medida_resi... 14.7% medida_resi... 14.7% 8 others ..... 70.5%</div>	<div>valor_final object</div> <div>550.0 ..... 1% 0.2 ..... 0.4% 3249 others ... 98.6%</div>	
0	100171	capacitancia_total	794.65	
1	100171	factor_de_potencia _10kv_ctotal	0.34	
2	100171	factor_de_potencia _2.5kv_ctotal	0.336	
3	100172	capacitancia_total	793.61	
4	100172	factor_de_potencia _10kv_ctotal	0.322	
5	100172	factor_de_potencia _2.5kv_ctotal	0.347	
6	100173	capacitancia_total	793.83	
7	100173	factor_de_potencia _10kv_ctotal	0.33	
8	100173	factor_de_potencia _2.5kv_ctotal	0.327	
9	100189	capacitancia_total	860.2	

```
df_equipos_viejos_val_final = df_val_finales.pivot(index="Equipo", columns="Denominacion", values="valor_final")
df_equipos_viejos_val_final=df_equipos_viejos_val_final.reset_index()
df_equipos_viejos_val_final
```

	<div>Equipo int64 100171 - 10031587</div> 	<div>capacitancia_total (</div> <div>1190.9 ..... 0.3% 822 others ..... 73.3% Missing ..... 26.5%</div>	<div>factor_de_poten...</div> <div>0.19 ..... 1.7% 219 others ..... 54% Missing ..... 44.2%</div>	<div>factor_de_poten...</div> <div>0.2 ..... 2.1% 203 others ..... 49.8% Missing ..... 48.1%</div>	<div>medida_resisten...</div> <div>550.0 ..... 1.2% 829 others ..... 97.2% Missing ..... 1.6%</div>	<div>medida_resisten...</div> <div>550.0 ..... 1.4% 820 others ..... 97% Missing ..... 1.6%</div>	<div>medida_resisten...</div> <div>550.0 ..... 1.4% 826 others ..... 95.7% Missing ..... 2.9%</div>	<div>medida_resisten...</div> <div>550.0 ..... 1.2% 793 others ..... 93.8% Missing ..... 5%</div>
0	100171	794.65	0.34	0.336	50.0	28.0	23.0	21.0
1	100172	793.61	0.322	0.347	81.0	41.0	22.0	51.0
2	100173	793.83	0.33	0.327	158.0	123.0	135.0	244.0
3	100189	860.2	0.26	0.26	705.0	665.0	550.0	4500.0
4	100190	827.0	0.33	0.33	650.0	357.0	480.0	801.0
5	100191	847.5	0.35	0.35	4190.0	7080.0	4690.0	4690.0
6	100195	860.8	0.26	0.256	93.0	65.0	36.0	37.0
7	100196	863.8	0.318	0.313	1440.0	1470.0	930.0	1220.0
8	100197	861.35	0.294	0.29	25.0	18.0	48.0	24.0
9	100207	844.01	0.252	0.246	3570.0	3180.0	3920.0	4580.0

```
col_info(df_equipos_viejos_val_final)
```

```
{0: 'Equipo',
1: 'capacitancia_total',
2: 'factor_de_potencia_10kv_ctotal',
3: 'factor_de_potencia_2.5kv_ctotal',
4: 'medida_resistencia_aislamiento_nucleo_1',
5: 'medida_resistencia_aislamiento_nucleo_2',
6: 'medida_resistencia_aislamiento_nucleo_3',
7: 'medida_resistencia_aislamiento_nucleo_4',
8: 'medida_resistencia_aislamiento_nucleo_5',
9: 'medida_resistencia_aislamiento_nucleo_6',
10: 'tip_up_capacitancia_total'}
```

Cambiar nombre de la columna final

```
# Creamos un diccionario con los nuevos nombres de las columnas
new_names = {'capacitancia_total': 'capacitancia_total_ultimo', 'factor_de_potencia_2.5kv_ctotal': 'factor_de_potencia_2.5kv_ctotal_ultimo',
'medida_resistencia_aislamiento_nucleo_2': 'medida_resistencia_aislamiento_nucleo_2_ultimo',
'medida_resistencia_aislamiento_nucleo_3': 'medida_resistencia_aislamiento_nucleo_3_ultimo',
'medida_resistencia_aislamiento_nucleo_4': 'medida_resistencia_aislamiento_nucleo_4_ultimo',
'medida_resistencia_aislamiento_nucleo_5': 'medida_resistencia_aislamiento_nucleo_5_ultimo',
'medida_resistencia_aislamiento_nucleo_6': 'medida_resistencia_aislamiento_nucleo_6_ultimo',
'tip_up_capacitancia_total': 'tip_up_capacitancia_total_ultimo'}

# Renombramos las columnas utilizando el método rename() y el diccionario
df_equipos_viejos_val_final = df_equipos_viejos_val_final.rename(columns=new_names)
df_equipos_viejos_val_final
```

	Equipo int64 100171 - 10031587	capacitancia_tot... 1190.9 ..... 0.3% 822 others ..... 73.3% Missing ..... 26.5%	factor_de_poten... 0.19 ..... 1.7% 219 others ..... 54% Missing ..... 44.2%	factor_de_poten... 0.2 ..... 2.1% 203 others ..... 49.8% Missing ..... 48.1%	medida_resisten... 550.0 ..... 1.2% 829 others ..... 97.2% Missing ..... 1.6%	medida_resisten... 550.0 ..... 1.4% 820 others ..... 97% Missing ..... 1.6%	medida_resisten... 550.0 ..... 1.4% 826 others ..... 95.7% Missing ..... 2.9%	medida_resisten... 550.0 ..... 1.2% 793 others ..... 93.8% Missing ..... 5%
0	100171	794.65	0.34	0.336	50.0	28.0	23.0	21.0
1	100172	793.61	0.322	0.347	81.0	41.0	22.0	51.0
2	100173	793.83	0.33	0.327	158.0	123.0	135.0	244.0
3	100189	860.2	0.26	0.26	705.0	665.0	550.0	4500.0
4	100190	827.0	0.33	0.33	650.0	357.0	480.0	801.0
5	100191	847.5	0.35	0.35	4190.0	7080.0	4690.0	4690.0
6	100195	860.8	0.26	0.256	93.0	65.0	36.0	37.0
7	100196	863.8	0.318	0.313	1440.0	1470.0	930.0	1220.0
8	100197	861.35	0.294	0.29	25.0	18.0	48.0	24.0
9	100207	844.01	0.252	0.246	3570.0	3180.0	3920.0	4580.0

agreguemos estos valores a la tabla con las tendencias con un join

SQL

Saved to variable df\_11

```
SELECT e.*, f.*
FROM df_equipos_viejos_val_final e
LEFT JOIN  tabla_unida f ON e.Equipo = f.Equipo
```

	Equipo int64 100171 - 10031587	capacitancia_tot... 1190.9 ..... 0.3% 822 others ..... 73.3% Missing ..... 26.5%	factor_de_poten... 0.19 ..... 1.7% 219 others ..... 54% Missing ..... 44.2%	factor_de_poten... 0.2 ..... 2.1% 203 others ..... 49.8% Missing ..... 48.1%	medida_resisten... 550.0 ..... 1.2% 829 others ..... 97.2% Missing ..... 1.6%	medida_resisten... 550.0 ..... 1.4% 820 others ..... 97% Missing ..... 1.6%	medida_resisten... 550.0 ..... 1.4% 826 others ..... 95.7% Missing ..... 2.9%	medida_resisten... 550.0 ..... 1.2% 793 others ..... 93.8% Missing ..... 5%
992	100171	794.65	0.34	0.336	50.0	28.0	23.0	21.0
0	100172	793.61	0.322	0.347	81.0	41.0	22.0	51.0
1	100173	793.83	0.33	0.327	158.0	123.0	135.0	244.0

2	100189	860.2	0.26	0.26	705.0	665.0	550.0	4500.0
3	100190	827.0	0.33	0.33	650.0	357.0	480.0	801.0
4	100191	847.5	0.35	0.35	4190.0	7080.0	4690.0	4690.0
5	100195	860.8	0.26	0.256	93.0	65.0	36.0	37.0
6	100196	863.8	0.318	0.313	1440.0	1470.0	930.0	1220.0
7	100197	861.35	0.294	0.29	25.0	18.0	48.0	24.0
8	100207	844.01	0.252	0.246	3570.0	3180.0	3920.0	4580.0

Haremos lo mismo que hicimos con la tabla de datos nuevos de agregar las familias

SQL

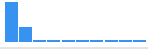
Saved to variable df\_equipos\_hist\_final

```
SELECT e.*, f.Fabricante, f.Denomintipo
FROM df_11 e
LEFT JOIN df_fam f ON e.Equipo = f.Equipo
```

	<div>Equipo int64 100171 - 10031587</div> <div></div>	<div>capacitancia_tot... 1190.9 ..... 0.3% 822 others ..... 73.3% Missing ..... 26.5%</div>	<div>factor_de_poten... 0.19 ..... 1.7% 219 others ..... 54% Missing ..... 44.2%</div>	<div>factor_de_poten... 0.2 ..... 2.1% 203 others ..... 49.8% Missing ..... 48.1%</div>	<div>medida_resisten... 550.0 ..... 1.2% 829 others ..... 97.2% Missing ..... 1.6%</div>	<div>medida_resisten... 550.0 ..... 1.4% 820 others ..... 97% Missing ..... 1.6%</div>	<div>medida_resisten... 550.0 ..... 1.4% 826 others ..... 95.7% Missing ..... 2.9%</div>	<div>medida_resisten... 550.0 ..... 1.2% 793 others ..... 93.8% Missing ..... 5%</div>
1053	100171	794.65	0.34	0.336	50.0	28.0	23.0	21.0
0	100172	793.61	0.322	0.347	81.0	41.0	22.0	51.0
1	100173	793.83	0.33	0.327	158.0	123.0	135.0	244.0
2	100189	860.2	0.26	0.26	705.0	665.0	550.0	4500.0
3	100190	827.0	0.33	0.33	650.0	357.0	480.0	801.0
4	100191	847.5	0.35	0.35	4190.0	7080.0	4690.0	4690.0
5	100195	860.8	0.26	0.256	93.0	65.0	36.0	37.0
6	100196	863.8	0.318	0.313	1440.0	1470.0	930.0	1220.0
7	100197	861.35	0.294	0.29	25.0	18.0	48.0	24.0
8	100207	844.01	0.252	0.246	3570.0	3180.0	3920.0	4580.0

Los espacios vacios en muchas de las medidas son debido a que no existen medidas en ningun año o solo poseen una medida en ese documento de medida, por lo tanto no encontré ningún porcentaje, poner el ultimo valor o asumir que es 0 ( estable) seria un error puesto que no corresponde a la información que brindan los datos y ademas seria falso decir que es estable puesto que no hay registro pasados. Algunos tambien les paso que tenian dos registros pero en un mismo año mismo me y mismo dia pero en diferentes horas y los resultados eran los mismos y el algoritmo los omite pero solo es un solo caso (hasta ahora solo se ah visto uno)


Para este caso la decisión que se puede tomar es asumir que son constantes

df_equipos_hist_final								
	<div>Equipo int64 100171 - 10031587</div> 	<div>capacitancia_tot...</div> <div>1190.9 ..... 0.3% 822 others ..... 73.3% Missing ..... 26.5%</div>	<div>factor_de_poten...</div> <div>0.19 ..... 1.7% 219 others ..... 54% Missing ..... 44.2%</div>	<div>factor_de_poten...</div> <div>0.2 ..... 2.1% 203 others ..... 49.8% Missing ..... 48.1%</div>	<div>medida_resisten...</div> <div>550.0 ..... 1.2% 829 others ..... 97.2% Missing ..... 1.6%</div>	<div>medida_resisten...</div> <div>550.0 ..... 1.4% 820 others ..... 97% Missing ..... 1.6%</div>	<div>medida_resisten...</div> <div>550.0 ..... 1.4% 826 others ..... 95.7% Missing ..... 2.9%</div>	<div>medida_resisten...</div> <div>550.0 ..... 1.2% 793 others ..... 93.8% Missing ..... 5%</div>
1053	100171	794.65	0.34	0.336	50.0	28.0	23.0	21.0
0	100172	793.61	0.322	0.347	81.0	41.0	22.0	51.0
1	100173	793.83	0.33	0.327	158.0	123.0	135.0	244.0
2	100189	860.2	0.26	0.26	705.0	665.0	550.0	4500.0
3	100190	827.0	0.33	0.33	650.0	357.0	480.0	801.0

4	100191	847.5	0.35	0.35	4190.0	7080.0	4690.0	4690.0
5	100195	860.8	0.26	0.256	93.0	65.0	36.0	37.0
6	100196	863.8	0.318	0.313	1440.0	1470.0	930.0	1220.0
7	100197	861.35	0.294	0.29	25.0	18.0	48.0	24.0
8	100207	844.01	0.252	0.246	3570.0	3180.0	3920.0	4580.0

```
df_equipos_hist_final=df_equipos_hist_final[~df_equipos_hist_final['Equipo'].isin([132382,
132383,
132384,
132385,
132386,
132387,
132388,
132389,
132390,
132391,
132392,
132393,
132394,
132395,
132396,
132397,
132398,
132399,
133167,
133168,
133169,
133181,
133182,
133183,
133184,
133185,
133186,
133187,
133188,
133189,
2003078,
2003079,
2003080,
2003081,
2003082,
2003083,
2003084,
2003085,
2003086,
2003090,
2003091,
2003092,
2003093,
2003094,
2003095,
2003096,
2003097,
2003098,
2003099,
2003100,
2003101,
2003102,
2003103,
2003104,
2003105,
2003106,
2003107,
2003108,
2003109,
2003110,
2003111,
2003112,
2003113,
2003114,
2003115,
2006252,
```

```
2006253,
2006254,
2006707,
2007781,
2008036,
2008037,
2008038,
2006181,
2006182,
2006183,
2006184,
2006185,
2006186,
2006534,
2006535,
2006536,
2006537,
2006538,
2006539,
2006540,
2006541,
2006542,
2006647,
2006648,
2006649,
2007339,
2007340,
2007341,
2007342,
2007343,
2007344,
2007516,
2007517,
2007518,
2007730,
2007731,
2007732,
2007733,
2007734,
2007751,
2007752,
2007753,
2007754,
2007755,
2007757,
2007789,
2007790,
2007791,
2008189,
2008366,
2008368,
2008369,
2008371,
2008372,
2008374,
2008375,
2008377,
2008378,
2008380,
2008381,
2008383,
2008798,
2008799,
2008800,
2009085,
2009086,
2009087,
2009088,
2012679]]]
```

df_equipos_hist_final								
	Equipo int64 100171 - 10031587	capacitancia_tot...	factor_de_poten...	factor_de_poten...	medida_resisten...	medida_resisten...	medida_resisten...	medida_resisten...
		1190.9 ..... 0.3% 766 others ..... 72% Missing ..... 27.7%	0.19 ..... 1.7% 212 others ..... 51.8% Missing ..... 46.4%	0.2 ..... 2.2% 200 others ..... 47.2% Missing ..... 50.6%	550.0 ..... 0.8% 796 others ..... 97.5% Missing ..... 1.6%	550.0 ..... 1% 787 others ..... 97.3% Missing ..... 1.6%	550.0 ..... 0.9% 791 others ..... 96.1% Missing ..... 3%	550.0 ..... 0.8% 758 others ..... 94% Missing ..... 5.2%
1053	100171	794.65	0.34	0.336	50.0	28.0	23.0	21.0

0	100172	793.61	0.322	0.347	81.0	41.0	22.0	51.0
1	100173	793.83	0.33	0.327	158.0	123.0	135.0	244.0
2	100189	860.2	0.26	0.26	705.0	665.0	550.0	4500.0
3	100190	827.0	0.33	0.33	650.0	357.0	480.0	801.0
4	100191	847.5	0.35	0.35	4190.0	7080.0	4690.0	4690.0
5	100195	860.8	0.26	0.256	93.0	65.0	36.0	37.0
6	100196	863.8	0.318	0.313	1440.0	1470.0	930.0	1220.0
7	100197	861.35	0.294	0.29	25.0	18.0	48.0	24.0
8	100207	844.01	0.252	0.246	3570.0	3180.0	3920.0	4580.0

```
df_equipos_hist_final = df_equipos_hist_final.dropna(subset=[ 'factor_de_potencia_10kv_ctotal',  
'factor_de_potencia_2.5kv_ctotal', 'capacitancia_total' ])
```

```
df_equipos_hist_final=df_equipos_hist_final.drop(['tip_up_capacitancia_total_ultimo'], axis=1)
```

df_equipos_hist_final								
	<div>Equipo int64</div> <div>100171 - 10031587</div> <div></div>	<div>capacitancia_tot...</div> <div>962.54 ..... 0.4%</div> <div>941.96 ..... 0.4%</div> <div>508 others .... 99.2%</div>	<div>factor_de_poten...</div> <div>0.19 ..... 3.6%</div> <div>0.21 ..... 2.7%</div> <div>201 others ..... 93.7%</div>	<div>factor_de_poten...</div> <div>0.2 ..... 4.6%</div> <div>0.21 ..... 2.7%</div> <div>198 others ..... 92.8%</div>	<div>medida_resisten...</div> <div>550.0 ..... 1.7%</div> <div>429 others ..... 96.2%</div> <div>Missing ..... 2.1%</div>	<div>medida_resisten...</div> <div>550.0 ..... 1.7%</div> <div>428 others ..... 96.2%</div> <div>Missing ..... 2.1%</div>	<div>medida_resisten...</div> <div>550.0 ..... 1.9%</div> <div>409 others ..... 93.7%</div> <div>Missing ..... 4.4%</div>	<div>medida_resisten...</div> <div>550.0 ..... 1.7%</div> <div>406 others ..... 91.8%</div> <div>Missing ..... 6.5%</div>
788	2005638	999.75	0.185	0.187	None	None	None	None
1056	103126	998.6	0.22	0.22	10800.0	9390.0	10100.0	11500.0
99	111949	998.19	0.161	0.17	4620.0	5080.0	4230.0	10800.0
61	106718	996.93	0.26	0.26	22100.0	19000.0	27700.0	29600.0
148	119027	996.61	0.139	0.149	75900.0	64100.0	85300.0	73200.0
59	106716	993.68	0.2	0.21	36100.0	29700.0	45400.0	31500.0
346	132270	993.11	0.226	0.229	23500.0	27000.0	22500.0	22300.0
25	103272	992.53	0.142	0.15	28400.0	23900.0	29600.0	27500.0
94	111428	988.89	0.21	0.2	1200.0	2400.0	3000.0	2400.0
787	2005637	986.9	0.19	0.19	None	None	None	None

```
df_equipos_hist_final.columns
```



```
Index(['Equipo', 'capacitancia_total_ultimo',
      'factor_de_potencia_10kv_ctotal_ultimo',
      'factor_de_potencia_2.5kv_ctotal_ultimo',
      'medida_resistencia_aislamiento_nucleo_1_ultimo',
      'medida_resistencia_aislamiento_nucleo_2_ultimo',
      'medida_resistencia_aislamiento_nucleo_3_ultimo',
      'medida_resistencia_aislamiento_nucleo_4_ultimo',
      'medida_resistencia_aislamiento_nucleo_5_ultimo',
      'medida_resistencia_aislamiento_nucleo_6_ultimo', 'Equipo_1',
      'capacitancia_total', 'factor_de_potencia_10kv_ctotal',
      'factor_de_potencia_2.5kv_ctotal',
      'medida_resistencia_aislamiento_nucleo_1',
      'medida_resistencia_aislamiento_nucleo_2',
      'medida_resistencia_aislamiento_nucleo_3',
      'medida_resistencia_aislamiento_nucleo_4',
      'medida_resistencia_aislamiento_nucleo_5',
      'medida_resistencia_aislamiento_nucleo_6', 'calificacion_del_equipo',
      'estado_cajetin', 'estado_de_la_porcelana', 'fuga_de_aceite',
      'inspeccion_diafragma', 'inspeccion_visual_general', 'nivel_de_aceite',
      'severidad_por_termografia', 'Fabricante', 'Denomintipo'],
      dtype='object')
```

```
tabla_unida['medida_resistencia_aislamiento_nucleo_5'].replace(np.nan, "No_tiene" ,inplace=True)
tabla_unida['medida_resistencia_aislamiento_nucleo_6'].replace(np.nan, "No_tiene" ,inplace=True)
tabla_unida['medida_resistencia_aislamiento_nucleo_4'].replace(np.nan, "No_tiene", inplace=True)
tabla_unida['medida_resistencia_aislamiento_nucleo_3'].replace(np.nan, "No_tiene" ,inplace=True)
```

tabla\_unida

	Equipo int64 100171 - 10031587	capacitancia_total i -0.999860045860...	factor_de_poten... -1.0 - 8611.440191...	factor_de_poten... -0.989348659003...	medida_resisten... -0.999 - 4759.869...	medida_resisten... -0.999709202453...	medida_resisten... 0.0 ..... 84.9% No_tiene ..... 2.9% 141 others ..... 12.3%	medida_resisten... 0.0 ..... 84.1% No_tiene ..... 5% 125 others ..... 11%
0	100171	-0.0004151047825	0.847826087	0.8162162162	-0.7126436782	-0.7985611511	-0.85625	-0.9082969432314411
1	100172	0.002779848625	1.205479452	1.393103448	-0.7990074442	-0.8828571429	-0.9176029962546817	-0.9244444444444444
2	100173	-0.004502019011	1.129032258	1.137254902	-0.9030674847	-0.9206451613	-0.9055944055944056	-0.6911392405063291
3	100189	0.06263125386	0.3684210526	0.3684210526	-0.5537974684	-0.1419354839	2.00546444808743167	3.8025613660618998
4	100190	0.06435006435	0.7368421053	0.7368421053	-0.1854636591	-0.72109375	-0.6363636363636364	-0.24433962264150944
5	100191	0.05871330418	0.75	0.75	0	0	0.0	0.0
6	100195	0.05879458795	0.04	0.06666666667	-0.8688293371	-0.9071428571	-0.891566265060241	-0.8945868945868946
7	100196	0.05987730061	0.272	0.252	0	0	-0.5938864628820961	0.0
8	100197	0.06470951792	0.3363636364	0.2608695652	-0.9938042131	-0.9411764706	-0.857566765578635	-0.9672131147540983
9	100207	0.06031407035	0.26	0.2947368421	0	0	0.0	0.0

Definir valores para las familias

df\_familias

	Equipo int64 100171 - 10031587	Activo fijo float64 45702.0 - 2000002...	Fabricante object GEC ALSTHOM BALTEAU	Denomin.tipo obj...	Fabr. N°-serie obj...	Denominación ob...	Inic.gar.prov. dat...	FinGarProv dateti...
268	103161	8023488	GEC ALSTHOM BALTEAU	QDR-245	54675-03	CT 230 kV-A	NaT	NaT

dfhist

	Equipo int64 100171 - 10031587	Aviso float64 45702.0 - 2000002...	Creado el datetim... 1999-09-29 00:00:...	Hora object	Posición medida o..	Denominación ob...	Valor medido floa...	Unidad ámb.med. o
--	-----------------------------------	---------------------------------------	--	-------------	---------------------	--------------------	----------------------	-------------------

0	100171	nan	2020-11-27 00:00:00	21:45:03	CT TOTAL	Calificación del equipo	3	UNIDAD
1	100171	5509002	2020-11-27 00:00:00	16:00:00	NUCLEO 4	Medida resistencia aislamiento núcle...	21	Mohmio
2	100171	5509002	2020-11-27 00:00:00	16:00:00	NUCLEO 3	Medida resistencia aislamiento núcle...	23	Mohmio
3	100171	5509002	2020-11-27 00:00:00	16:00:00	NUCLEO 2	Medida resistencia aislamiento núcle...	28	Mohmio
4	100171	5509002	2020-11-27 00:00:00	16:00:00	NUCLEO 1	Medida resistencia aislamiento núcle...	50	Mohmio
5	100171	nan	2020-11-24 00:00:00	21:18:56	CT TOTAL	Calificación del equipo	5	UNIDAD
6	100171	5509002	2020-11-24 00:00:00	16:51:27	CAPACITANCIA TOTAL	Factor de Potencia 2.5kV Ctotal	0.336	%
7	100171	5509002	2020-11-24 00:00:00	16:51:27	CAPACITANCIA TOTAL	Factor de Potencia 10kV Ctotal	0.34	%
8	100171	5509002	2020-11-24 00:00:00	16:51:27	CAPACITANCIA TOTAL	Capacitancia total	794.65	pF
9	100171	nan	2020-10-14 00:00:00	05:05:15	CT TOTAL	Calificación del equipo	5	UNIDAD

SQL

Saved to variable df\_16

```
SELECT Equipo,Denominación,"Valor teórico","Valor medido"
FROM dfhist
GROUP BY Equipo,Denominación,"Valor teórico","Valor medido"
```

	<div>Equipo int64 100171 - 10031587</div>	<div>Denominación ob... Medida resi... 11.2% Medida resi... 11.2% 33 others ..... 77.6%</div>	<div>Valor teórico float... 0.09 - 10000.0</div>	<div>Valor medido floa... -0.009 - 2830000.0</div>	
0	100171	Calificación del equipo	nan	3	
16930	100171	Medida resistencia aislamiento núcle...	nan	697	
16929	100171	Medida resistencia aislamiento núcle...	nan	229	
16928	100171	Tip Up Capacitancia...	nan	0.001	
16927	100171	Nivel de aceite	nan	nan	
16926	100171	Capacitancia total	786.96	794.65	
16925	100171	Factor de Potencia 2.5kV Ctotal	0.17	0.336	
16924	100171	Calificación del equipo	nan	5	
16923	100171	Medida resistencia aislamiento núcle...	nan	50	
11308	100171	Capacitancia total	786.96	794.88	

SQL

Saved to variable df\_17

```
SELECT e.*, f.Fabricante, f."Denomin.tipo"
FROM df_16 e
LEFT JOIN df_familias f ON e.Equipo = f.Equipo
```

	<div><div>Equipo int64</div><div>100171 - 10031587</div></div>	<div><div>Denominación ob...</div><div>Medida resi... 11.2% Medida resi... 11.2% 33 others ..... 77.6%</div></div>	<div><div>Valor teórico float...</div><div>0.09 - 10000.0</div></div>	<div><div>Valor medido floa...</div><div>-0.009 - 2830000.0</div></div>	<div><div>Fabricante object</div><div></div></div>	<div><div>Denomin.tipo obj...</div><div></div></div>	
0	100172	Calificación del equipo	nan	3	RITZ	OSKF 245.32	
1	100172	Medida resistencia aislamiento núcle...	nan	81	RITZ	OSKF 245.32	
2	100172	Factor de Potencia 2.5kV Ctotal	0.17	0.347	RITZ	OSKF 245.32	
3	100172	Capacitancia total	786.96	793.61	RITZ	OSKF 245.32	

4	100172	Inspección visual general	nan	nan	RITZ	OSKF 245.32	
5	100172	Medida resistencia aislamiento núcle...	nan	403	RITZ	OSKF 245.32	
6	100173	Inspección visual general	nan	nan	RITZ	OSKF 245.32	
7	100173	Factor de Potencia 2.5kV Ctotal	0.17	0.153	RITZ	OSKF 245.32	
8	100173	Medida resistencia aislamiento núcle...	nan	135	RITZ	OSKF 245.32	
9	100173	Medida resistencia aislamiento núcle...	nan	158	RITZ	OSKF 245.32	

```
df_17['Denominación'].unique()

array(['Calificación del equipo',
      'Medida resistencia aislamiento núcleo 1',
      'Factor de Potencia 2.5kV Ctotal', 'Capacitancia total',
      'Inspección visual general',
      'Medida resistencia aislamiento núcleo 3',
      'Medida resistencia aislamiento núcleo 4',
      'Medida resistencia aislamiento núcleo 2',
      'Factor de Potencia 10kV Ctotal', 'Nivel de aceite',
      'Fuga de aceite', 'Severidad por termografía',
      'Tip Up Capacitancia Ctotal',
      'Medida resistencia aislamiento núcleo 6',
      'Medida resistencia aislamiento núcleo 5', 'Inspección diafragma',
      'Estado de la Porcelana', 'Estado Cajetin',
      'Factor de Potencia 2 kV Ctotal', 'Tip up capacitancia Ctotal',
      'Tip Up Capacitancia CTotal', 'Capacitancia Total',
      'Factor de potencia 10 kV Ctotal', 'Tip Up Capacitancia total',
      'Factor de potencia 2.5kV Ctotal', 'Factor de Potencia 2kV Ctotal',
      'Factor de Potencia 2.5 kV total', 'Nivel de aceite NO USAR',
      'Factor de Potencia 10 kV Ctotal',
      'Factor de Potencia\x010kV Ctotal',
      'Factor de Potencia 2.5 kV Ctotal',
      'Factor de potencia 10kV Ctotal',
      'Medida resistencia aislamiento nucleo 2',
      'Nivel de aceite - NO USAR',
      'Medida resistencia aislamiento nucleo 1'], dtype=object)
```

```
# Seleccionar las filas que cumplen con la condición
filas_a_eliminar = df_17.loc[df_17['Denominación'].isin(['Inspección visual general', 'Nivel de aceite',
      'Medida resistencia aislamiento núcleo 3',
      'Tip Up Capacitancia Ctotal',
      'Fuga de aceite', 'Medida resistencia aislamiento núcleo 2',
      'Medida resistencia aislamiento núcleo 1',
      'Calificación del equipo',
      'Medida resistencia aislamiento núcleo 4',
      'Severidad por termografía',
      'Medida resistencia aislamiento núcleo 5',
      'Medida resistencia aislamiento núcleo 6', 'Inspección diafragma',
      'Estado Cajetin', 'Estado de la Porcelana',
      'Tip up capacitancia Ctotal',
      'Tip Up Capacitancia CTotal',
      'Nivel de aceite - NO USAR', 'Nivel de aceite NO USAR',
      'Tip Up Capacitancia total',
      'Medida resistencia aislamiento nucleo 2',
      'Medida resistencia aislamiento nucleo 1'])])
```

```
df_17 = df_17.drop(filas_a_eliminar.index)
```

df_17							
	<b>Equipo</b> int64 100171 - 10031587	<b>Denominación</b> ob... Capacitanc... 40.5% Factor de P... 30.4% 11 others ..... 29.1%	<b>Valor teórico</b> float... 0.09 - 1650.0	<b>Valor medido</b> floa... 0.0 - 947360.0	<b>Fabricante</b> object GEC ALSTH... _ 40% SIEMENS-T... 18.6% 13 others ..... 41.4%	<b>Denomin.tipo</b> obj... QDR-245 ..... 31.6% 40 others ..... 67.8% Missing ..... 0.6%	
1683	100171	Factor de Potencia 10kV Ctotal	0.17	0.34	RITZ	OSKF 245.32	



6069	100171	Factor de Potencia 2.5kV Ctotal	0.17	0.184	RITZ	OSKF 245.32	
6073	100171	Factor de Potencia 2.5kV Ctotal	0.17	0.189	RITZ	OSKF 245.32	
12126	100171	Factor de Potencia 2.5kV Ctotal	0.17	0.185	RITZ	OSKF 245.32	
12127	100171	Capacitancia total	786.96	794.98	RITZ	OSKF 245.32	
12131	100171	Capacitancia total	786.96	794.88	RITZ	OSKF 245.32	
18171	100171	Factor de Potencia 2.5kV Ctotal	0.17	0.336	RITZ	OSKF 245.32	
18172	100171	Capacitancia total	786.96	794.65	RITZ	OSKF 245.32	
18177	100171	Factor de Potencia 10kV Ctotal	0.17	0.186	RITZ	OSKF 245.32	
5470	100172	Factor de Potencia 10kV Ctotal	0.17	0.322	RITZ	OSKF 245.32	

Conteo de equipos por calificacion

SQL





Saved to variable df\_8

```
SELECT Equipo,Fabricante,Denomintipo,calificacion_del_equipo, COUNT(*) as cantidad
FROM df_equipos_hist_final
GROUP BY Equipo,Fabricante, Denomintipo,calificacion_del_equipo
ORDER BY Fabricante, Denomintipo
```

	<div>Equipo int64 100171 - 10031587</div> 	<div>Fabricante object GEC ALST... 50.2% ARTECHE 9.1% 8 others 40.7%</div>	<div>Denomintipo object QDR-245 39.5% 22 others 59.1% Missing 1.3%</div>	<div>calificacion_del_... 5.0 82.1% 3.0 15.6% 2.0 2.3%</div>	<div>cantidad int64 1 - 1</div> 	
10	122786	ABB	IMB245	5.0	1	
11	2011783	ABB	IMB245	5.0	1	
12	2011781	ABB	IMB245	5.0	1	
13	132279	ABB	IMB245	3.0	1	
14	132266	ABB	IMB245	5.0	1	
15	132256	ABB	IMB245	3.0	1	
16	130937	ABB	IMB245	5.0	1	
17	126905	ABB	IMB245	5.0	1	
18	132255	ABB	IMB245	5.0	1	
19	132265	ABB	IMB245	5.0	1	

SQL	Saved to variable df_7					
<pre>SELECT     Fabricante,     SUM(CASE WHEN calificacion_del_equipo = 5 THEN cantidad ELSE 0 END) AS buen_estado_5,     SUM(CASE WHEN calificacion_del_equipo = 3 THEN cantidad ELSE 0 END) AS revision_3,     SUM(CASE WHEN calificacion_del_equipo = 2 THEN cantidad ELSE 0 END) AS pre_falla_funcional_2,     SUM(CASE WHEN calificacion_del_equipo = 1 THEN cantidad ELSE 0 END) AS falla_funcional_1, FROM     df_8 GROUP BY     Fabricante;</pre>						
	<div>Fabricante object</div> <div>ABB ..... 10%</div> <div>ALSTHOM ..... 10%</div> <div>8 others ..... 80%</div>	<div>buen_estado_5 in...</div> <div>0 - 210</div> 	<div>revision_3 int64</div> <div>1 - 47</div> 	<div>pre_falla_funcio...</div> <div>0 - 7</div> 	<div>falla_funcional_1 i...</div> <div>0 - 0</div> 	
0	ABB	35	5	0	0	

1	ALSTHOM	33	3	1	0
2	ARTECHE	31	16	1	0
3	Crompton Greaves	0	1	0	0
4	GEC ALSTHOM BALTEAU	210	47	7	0
5	HAEFELY	17	1	0	0
6	HAEFELY TRENCH	26	2	2	0
7	NUOVA MAGRINI GALILEO	28	1	1	0
8	RITZ	10	3	0	0
9	SIEMENS-TRENCH	42	3	0	0

SQL	Saved to variable df_9					
<pre>SELECT   Fabricante, Denomintipo,   SUM(CASE WHEN calificacion_del_equipo = 5 THEN cantidad ELSE 0 END) AS buen_estado_5,   SUM(CASE WHEN calificacion_del_equipo = 3 THEN cantidad ELSE 0 END) AS revision_3,   SUM(CASE WHEN calificacion_del_equipo = 2 THEN cantidad ELSE 0 END) AS pre_falla_funcional_2,   SUM(CASE WHEN calificacion_del_equipo = 1 THEN cantidad ELSE 0 END) AS falla_funcional_1, FROM   df_8 GROUP BY   Fabricante, Denomintipo;</pre>						
	<div>Fabricante object</div> <div>GEC ALSTH... 20% NUOVA MAG... 20% 8 others 60%</div>	<div>Denomintipo object</div> <div>CTH 550 6.7% 22 others 86.7% Missing 6.7%</div>	<div>buen_estado_5 in...</div> <div>0 - 161</div> 	<div>revision_3 int64</div> <div>0 - 35</div> 	<div>pre_falla_funcio...</div> <div>0 - 6</div> 	<div>falla_funcional_1 i...</div> <div>0 - 0</div> 
0	ABB	None	6	0	0	0
1	ABB	IMB 245 A4	3	0	0	0
2	ABB	IMB245	26	5	0	0
3	ALSTHOM	CTH 550	16	2	1	0
4	ALSTHOM	OSFK-123	3	0	0	0
5	ALSTHOM	QDR-145	3	0	0	0
6	ALSTHOM	QDR-245	6	0	0	0
7	ALSTHOM	TDX-245	5	1	0	0
8	ARTECHE	CTG-145	2	0	1	0
9	ARTECHE	CTS 525	29	16	0	0
10	Crompton Greaves	IOSK245	0	1	0	0
11	GEC ALSTHOM BALTEAU	None	0	1	0	0
12	GEC ALSTHOM BALTEAU	CTH 550	33	9	0	0
13	GEC ALSTHOM BALTEAU	QDR-245	161	35	6	0
14	GEC ALSTHOM BALTEAU	QDR-245T	9	1	1	0
15	GEC ALSTHOM BALTEAU	TDX-245	6	1	0	0
16	GEC ALSTHOM BALTEAU	TDX245P	1	0	0	0
17	HAEFELY	TC 245	17	1	0	0

18	HAEFELY TRENCH	IOSK145	4	2	0	0	
19	HAEFELY TRENCH	IOSK245	22	0	2	0	
20	NUOVA MAGRINI GALILEO	ATH 123 B2	6	0	0	0	
21	NUOVA MAGRINI GALILEO	ATH 123 C5	3	0	0	0	
22	NUOVA MAGRINI GALILEO	ATH 245 F	11	0	1	0	
23	NUOVA MAGRINI GALILEO	ATH 245 LC5	3	0	0	0	
24	NUOVA MAGRINI GALILEO	ATH 245 LC6	3	0	0	0	
25	NUOVA MAGRINI GALILEO	TAGa123L	2	1	0	0	
26	RITZ	OSKF 245.32	10	3	0	0	
27	SIEMENS-TRENCH	IOSK 550	36	2	0	0	
28	SIEMENS-TRENCH	IOSK123	3	0	0	0	
29	SIEMENS-TRENCH	IOSK145	3	1	0	0	