

manipulacion_de_datos

June 26, 2024

0.0.1 Comenzamos cargando la libreria pandas, con esto leeremos en un primer paso los archivos “Datos Maestros VF” y “dDEC1204” este ultimo archivo es necesario utilizar el encoding “ISO-8859-1” ya que sin este no permite una carga exitosa ya que hay algun dato que pandas no puede inferir su tipo, continuando con la carga del archivo “Datos Maestros VF” solo se cargaran las columnas que se encuentran en la lista “use_col” y para la carga del archivo “dDEC1204” se le crearan sus columnas para darle un mejor formato y facilitar su lectura.

```
[1]: import pandas as pd

[4]: use_col = ['Nombre visible Agente', 'AGENTE (OFEI)', 'CENTRAL (dDEC, dSEGDES, dPRU...)', 'Tipo de central (Hidro, Termo, Filo, Menor)']
col_names = ["Central"] + [f'Hora_{i}' for i in range(1,25)]
master_data = pd.read_excel("./Prueba_Tecnica/Datos3/Datos Maestros VF.xlsx", usecols = use_col)
ddec = pd.read_csv("./Prueba_Tecnica/Datos3/dDEC1204.txt", delimiter= ",", names = col_names, encoding = 'ISO-8859-1')
```

0.0.2 Continuando con la resolucio, creamos una lista donde guardaremos los datos ‘H’ y ‘T’ ya que la utilizaremos para filtrar la columna ‘Tipo de central (Hidro, Termo, Filo, Menor)’ y tambien filtraremos la columna ‘Nombre visible Agente’, por ultimo verificamos el resultado obtenido

```
[5]: type = ['H', 'T']
filter_data = master_data[(master_data['Nombre visible Agente'] == 'EMGESA' ) & (master_data['Tipo de central (Hidro, Termo, Filo, Menor)'].isin(type))]
filter_data
```

```
[5]:
```

	Nombre visible Agente	AGENTE (OFEI)	CENTRAL (dDEC, dSEGDES, dPRU...) \
32	EMGESA	EMGESA S.A.	BETANIA
33	EMGESA	EMGESA S.A.	BETANIA
34	EMGESA	EMGESA S.A.	BETANIA
49	EMGESA	EMGESA S.A.	CTGEMG1
50	EMGESA	EMGESA S.A.	CTGEMG2
51	EMGESA	EMGESA S.A.	CTGEMG3
90	EMGESA	EMGESA S.A.	ELQUIMBO
91	EMGESA	EMGESA S.A.	ELQUIMBO
120	EMGESA	EMGESA S.A.	GUAVIO

121	EMGESA	EMGESA S.A.	GUAVIO
122	EMGESA	EMGESA S.A.	GUAVIO
123	EMGESA	EMGESA S.A.	GUAVIO
124	EMGESA	EMGESA S.A.	GUAVIO
148	EMGESA	EMGESA S.A.	PAGUA
149	EMGESA	EMGESA S.A.	PAGUA
150	EMGESA	EMGESA S.A.	PAGUA
188	EMGESA	EMGESA S.A.	PAGUA
189	EMGESA	EMGESA S.A.	PAGUA
190	EMGESA	EMGESA S.A.	PAGUA
299	EMGESA	EMGESA S.A.	ZIPAEMG2
300	EMGESA	EMGESA S.A.	ZIPAEMG3
301	EMGESA	EMGESA S.A.	ZIPAEMG4
302	EMGESA	EMGESA S.A.	ZIPAEMG5

Tipo de central (Hidro, Termo, Filo, Menor)

32	H
33	H
34	H
49	T
50	T
51	T
90	H
91	H
120	H
121	H
122	H
123	H
124	H
148	H
149	H
150	H
188	H
189	H
190	H
299	T
300	T
301	T
302	T

0.0.3 Posteriormente re realizaremos un merge entre los dos archivos y crearemos una nueva columna donde sera la suma de todas las horas de forma horizontal, por ultimo filtramos esta nueva columna para obtener los que son mayores a cero

```
[18]: data_merge = pd.merge(master_data, ddec, left_on = 'CENTRAL (dDEC, dSEGDES, dPRU...)', right_on = 'Central', how = 'inner')
col_hours = [f'Hora_{i}' for i in range(1,25)]
data_merge["Total_hours"] = data_merge[col_hours].sum(axis=1)
```

```
[21]: data_merge_filter = data_merge[data_merge["Total_hours"] > 0 ]
data_merge_filter
```

```
[21]:      Nombre visible Agente      AGENTE (OFEI) CENTRAL (dDEC, dSEGDES, dPRU...) \
0      RIO PIEDRAS  ENERGIA RIO PIEDRAS      MAGUAFRE
1      ALEJANDRÍA    PCH ALEJANDRÍA      M_ALEJANDRIA
2      EPSA          EPSA(PACIFICO)      ALBAN
3      EPSA          EPSA(PACIFICO)      ALBAN
4      EPSA          EPSA(PACIFICO)      ALBAN
..      ...
290     URRA          URRA              URRA
291     URRA          URRA              URRA
292  GENERAMOS ENERGIA  GENERAMOS ENERGIA      M_URRAO
294     VATIA          VATIA S.A.      MVENTANA_A
295     VATIA          VATIA S.A.      MVENTANA_B
```

```
      Tipo de central (Hidro, Termo, Filo, Menor)      Central  Hora_1  Hora_2 \
0      M      MAGUAFRE      7      7
1      M  M_ALEJANDRIA      8      8
2      H      ALBAN      223     223
3      H      ALBAN      223     223
4      H      ALBAN      223     223
..      ...
290     H      URRA      241     241
291     H      URRA      241     241
292     M      M_URRAO      1      1
294     M  MVENTANA_A      1      1
295     M  MVENTANA_B      2      2
```

```
      Hora_3  Hora_4  Hora_5  ...  Hora_16  Hora_17  Hora_18  Hora_19  Hora_20 \
0      7      7      7  ...      7      7      7      7      7
1      8      8      8  ...      8      8      8      8      8
2      223     223     223  ...     355     355     355     355     355
3      223     223     223  ...     355     355     355     355     355
4      223     223     223  ...     355     355     355     355     355
..      ...
290     241     241     241  ...     241     241     241     241     241
291     241     241     241  ...     241     241     241     241     241
```

292	1	1	1	...	1	1	1	1	1
294	1	1	1	...	1	1	1	1	1
295	2	2	2	...	2	2	2	2	2

	Hora_21	Hora_22	Hora_23	Hora_24	Total_hours
0	7	7	7	7	168
1	8	8	8	8	192
2	355	355	355	223	7596
3	355	355	355	223	7596
4	355	355	355	223	7596
..
290	241	241	241	241	5784
291	241	241	241	241	5784
292	1	1	1	1	24
294	1	1	1	1	24
295	2	2	2	2	48

[233 rows x 30 columns]

0.0.4 Por ultimo guardamos el archivo con el resultado final

```
[22]: data_merge_filter.to_csv("./Prueba_Tecnica/Resultados/Ejercicio2.csv",index =_
      ↪False)
```