Projet python Pole_emploi

July 16, 2020

[197]:
```
import pandas as pd

data = pd.read_excel('/home/simplon/Téléchargements/chomage_zone.xls',
                     sheet_name='txcho_ze', skiprows=5)
print(data)
```

```
                                                  ZE2010           LIBZE2010  \
0                                                   0050      Mont-de-Marsan
1                                                   0051             Alençon
2                                                   0052       Cosne-Clamecy
3                                                   0053               Mâcon
4                                                   0054    Nogent-le-Rotrou
..                                                   …                  …
300                                                 9404             Bastia
301                                                 9405  Calvi-L'Île-Rousse
302                                                 9406              Corte
303                                                 9407  Ghisonaccia-Aléria
304  Source : Insee, Estimations de taux de chômage…                    NaN

      REG         LIBREG  2003-T1  2003-T2  2003-T3  2003-T4  2004-T1  \
0     0.0  interrégional      6.0      6.1      6.0      6.3      6.4
1     0.0  interrégional      7.2      7.5      7.2      7.6      7.7
2     0.0  interrégional      7.5      7.6      7.4      7.8      8.0
3     0.0  interrégional      5.3      5.3      5.2      5.3      5.4
4     0.0  interrégional      5.9      6.3      6.4      7.0      7.3
..    …            …      …        …        …        …        …
300  94.0          Corse      9.2      8.9      8.9      9.3      9.4
301  94.0          Corse     11.7     11.4     11.2     11.4     11.2
302  94.0          Corse      8.3      7.6      8.2      8.1      8.0
303  94.0          Corse      9.6      8.8      9.0      9.4      9.6
304   NaN            NaN      NaN      NaN      NaN      NaN      NaN

     2004-T2  …  2017-T4  2018-T1  2018-T2  2018-T3  2018-T4  2019-T1  \
0        6.2  …      7.3      7.4      7.1      7.0      6.7      6.8
1        7.4  …      8.3      8.6      8.4      8.4      8.1      8.1
2        7.7  …      8.1      8.3      8.1      7.9      7.5      7.5
3        5.4  …      6.4      6.5      6.5      6.4      6.3      6.4
4        7.2  …      8.1      8.3      8.1      8.0      7.9      7.9
```

```
..       …  …      …       …       …       …       …       …
300      9.3  …      9.7     10.1      9.7      9.4      9.2      9.2
301     11.5  …     11.6     12.5     12.2     11.9     11.4     11.6
302      8.0  …      8.5      8.9      8.9      8.6      7.8      8.2
303      9.7  …     10.0     10.6     10.4     10.1      9.9     10.0
304      NaN  …      NaN      NaN      NaN      NaN      NaN      NaN

     2019-T2  2019-T3  2019-T4  2020-T1
0        6.7      6.7      6.5      6.1
1        7.9      7.9      7.8      7.4
2        7.2      7.4      7.2      6.8
3        6.3      6.3      6.1      5.9
4        7.8      8.0      7.4      7.1
..       …        …        …        …
300      8.8      8.5      8.2      7.9
301     11.0     10.8     10.5     10.1
302      7.6      7.3      7.1      7.3
303      9.4      9.4      9.0      8.6
304      NaN      NaN      NaN      NaN

[305 rows x 73 columns]
```

[199]:
```
data_paca = data.loc[data['LIBREG'] == "Provence-Alpes-Côte d'Azur", :]
```

[200]:
```
data_paca
```

[200]:
```
     ZE2010                LIBZE2010   REG                       LIBREG  \
280    9301            Digne-les-Bains  93.0  Provence-Alpes-Côte d'Azur
281    9302                  Manosque  93.0  Provence-Alpes-Côte d'Azur
282    9303                  Briançon  93.0  Provence-Alpes-Côte d'Azur
283    9304                       Gap  93.0  Provence-Alpes-Côte d'Azur
284    9305             Cannes-Antibes  93.0  Provence-Alpes-Côte d'Azur
285    9306  Menton Vallée de la Roya  93.0  Provence-Alpes-Côte d'Azur
286    9307                      Nice  93.0  Provence-Alpes-Côte d'Azur
287    9308            Aix-en-Provence  93.0  Provence-Alpes-Côte d'Azur
288    9309                     Arles  93.0  Provence-Alpes-Côte d'Azur
289    9310          Marseille-Aubagne  93.0  Provence-Alpes-Côte d'Azur
290    9311           Istres-Martigues  93.0  Provence-Alpes-Côte d'Azur
291    9312          Salon-de-Provence  93.0  Provence-Alpes-Côte d'Azur
292    9313                Draguignan  93.0  Provence-Alpes-Côte d'Azur
293    9314       Fréjus-Saint Raphaël  93.0  Provence-Alpes-Côte d'Azur
294    9315                    Toulon  93.0  Provence-Alpes-Côte d'Azur
295    9316              Cavaillon-Apt  93.0  Provence-Alpes-Côte d'Azur
296    9317                    Orange  93.0  Provence-Alpes-Côte d'Azur

     2003-T1  2003-T2  2003-T3  2003-T4  2004-T1  2004-T2  …  2017-T4  \
280      7.5      7.6      7.2      7.6      7.7      7.5  …     10.7
```

281	10.0	10.1	9.8	9.8	9.9	9.7	…	11.2
282	7.0	7.1	7.0	7.4	7.4	7.4	…	8.2
283	7.0	7.0	6.9	7.1	7.3	7.1	…	8.2
284	9.5	9.7	9.4	9.5	9.5	9.3	…	9.8
285	6.8	6.9	6.7	6.9	6.9	6.7	…	8.3
286	8.1	8.1	8.0	8.3	8.3	8.1	…	9.6
287	9.2	9.4	9.2	9.3	9.4	9.0	…	8.4
288	11.1	11.4	11.1	11.4	11.5	11.0	…	12.4
289	11.6	11.8	11.4	11.6	11.7	11.3	…	11.5
290	10.3	10.5	10.3	10.5	10.5	10.2	…	10.8
291	10.7	10.8	10.4	10.6	10.7	10.3	…	10.4
292	11.4	11.4	11.2	11.4	11.3	11.1	…	11.6
293	10.0	10.1	10.1	10.6	10.9	11.0	…	11.7
294	10.6	10.6	10.3	10.5	10.6	10.4	…	9.5
295	9.3	9.2	8.9	9.1	9.1	9.0	…	11.9
296	9.7	9.6	9.2	9.3	9.4	9.3	…	12.0

	2018-T1	2018-T2	2018-T3	2018-T4	2019-T1	2019-T2	2019-T3	2019-T4	\
280	11.0	10.7	10.6	10.4	10.3	10.0	9.8	9.4	
281	11.6	11.4	11.1	10.9	11.0	10.6	10.5	10.0	
282	8.4	8.3	8.4	8.2	8.3	8.3	8.0	7.9	
283	8.3	8.3	8.2	8.0	8.0	7.7	7.5	7.2	
284	10.1	9.9	9.7	9.5	9.5	9.1	9.1	8.7	
285	8.4	8.1	7.8	7.6	7.6	7.3	7.2	6.9	
286	9.8	9.6	9.5	9.2	9.2	8.9	8.9	8.4	
287	8.6	8.5	8.3	8.1	8.0	7.7	7.7	7.3	
288	12.5	12.5	12.3	11.7	11.8	11.4	11.4	11.0	
289	11.9	11.7	11.6	11.2	11.2	10.8	10.7	10.3	
290	11.1	10.8	10.7	10.2	10.3	10.1	9.9	9.6	
291	10.6	10.4	10.3	10.0	10.0	9.7	9.5	9.1	
292	12.0	11.8	11.5	11.2	11.2	10.9	10.8	10.2	
293	12.2	11.9	11.7	11.5	11.5	11.1	11.0	10.5	
294	9.7	9.6	9.5	9.3	9.3	8.9	8.8	8.4	
295	12.1	11.9	11.8	11.4	11.7	11.2	11.2	10.7	
296	12.2	12.1	12.0	11.7	11.6	11.3	11.2	10.5	

	2020-T1
280	9.1
281	9.6
282	7.5
283	7.0
284	8.5
285	6.7
286	8.2
287	7.0
288	10.7
289	9.9

```
290      9.2
291      8.7
292      9.8
293     10.3
294      8.1
295     10.3
296     10.1

[17 rows x 73 columns]
```

[201]:
```
data_paca["LIBZE2010"].value_counts()
```

[201]:
```
Arles                        1
Menton Vallée de la Roya     1
Cavaillon-Apt                1
Draguignan                   1
Orange                       1
Marseille-Aubagne            1
Manosque                     1
Nice                         1
Cannes-Antibes               1
Aix-en-Provence              1
Istres-Martigues             1
Digne-les-Bains              1
Gap                          1
Briançon                     1
Fréjus-Saint Raphaël         1
Toulon                       1
Salon-de-Provence            1
Name: LIBZE2010, dtype: int64
```

[203]:
```
data_t1_2020= data["2020-T1"]>7.1
```

[204]:
```
data_t1_2020
```

[204]:
```
0      False
1       True
2      False
3      False
4      False
       …
300     True
301     True
302     True
303     True
304    False
Name: 2020-T1, Length: 305, dtype: bool
```

[106]:
```
data_t1_2019= data["2019-T1"]>7.1
```

[107]:
```
data_t1_2019
```

[107]:
```
0      False
1       True
2       True
3      False
4       True
      …
300     True
301     True
302     True
303     True
304    False
Name: 2019-T1, Length: 305, dtype: bool
```

[205]:
```
data_t1_2018= data["2018-T1"]>7.1
data_t1_2018
```

[205]:
```
0       True
1       True
2       True
3      False
4       True
      …
300     True
301     True
302     True
303     True
304    False
Name: 2018-T1, Length: 305, dtype: bool
```

[214]:
```
data_7_1 = pd.DataFrame(data_t1_2018 & data_t1_2019 & data_t1_2020,␣
↪columns=['booléen'])
data_7_1['ville'] = data['LIBZE2010']
data_7_1
```

[214]:
```
     booléen               ville
0      False      Mont-de-Marsan
1       True             Alençon
2      False       Cosne-Clamecy
3      False               Mâcon
4      False    Nogent-le-Rotrou
..       …                   …
300     True             Bastia
301     True  Calvi-L'Île-Rousse
```

```
302     True               Corte
303     True  Ghisonaccia-Aléria
304    False                 NaN

[305 rows x 2 columns]
```

[215]:
```
data_7_1['booléen'].value_counts()
```

[215]:
```
True     163
False    142
Name: booléen, dtype: int64
```

[ ]:
```

```

[216]:
```
import pandas as pd

data2 = pd.read_excel ('/home/simplon/Téléchargements/pole_emploi.xlsx',␣
 ↪skiprows = 3)
data2
```

[216]:

	Code	Libellé	Pôle emploi 2018
0	1	Guadeloupe	17
1	2	Martinique	11
2	3	Guyane	8
3	4	La Réunion	17
4	6	Mayotte	3
5	11	Île-de-France	164
6	24	Centre-Val de Loire	116
7	27	Bourgogne-Franche-Comté	165
8	28	Normandie	145
9	32	Hauts-de-France	171
10	44	Grand Est	238
11	52	Pays de la Loire	104
12	53	Bretagne	91
13	75	Nouvelle-Aquitaine	251
14	76	Occitanie	262
15	84	Auvergne-Rhône-Alpes	294
16	93	Provence-Alpes-Côte d'Azur	145
17	94	Corse	31

[164]:
```

```

[217]:
```
fusion = pd.merge(data, data2, left_on ='REG', right_on ='Code')
fusion
```

[217]:

	ZE2010	LIBZE2010	REG	LIBREG	2003-T1	2003-T2	\
0	0060*	Saint-Étienne	84.0	Auvergne-Rhône-Alpes	8.0	8.2	

```
1     8201   Ambérieu-en-Bugey  84.0  Auvergne-Rhône-Alpes      5.7      5.8
2     8202     Bourg-en-Bresse  84.0  Auvergne-Rhône-Alpes      4.5      4.7
3     8203             Oyonnax  84.0  Auvergne-Rhône-Alpes      6.8      6.9
4     8204             Annonay  84.0  Auvergne-Rhône-Alpes      7.6      8.0
..     …                   …   …                     …      …      …
290   9403   Sartène-Propriano  94.0                Corse     10.7     10.7
291   9404             Bastia  94.0                Corse      9.2      8.9
292   9405  Calvi-L'Île-Rousse  94.0                Corse     11.7     11.4
293   9406              Corte  94.0                Corse      8.3      7.6
294   9407  Ghisonaccia-Aléria  94.0                Corse      9.6      8.8

     2003-T3  2003-T4  2004-T1  2004-T2  …  2018-T3  2018-T4  2019-T1  \
0        8.1      8.4      8.6      8.3  …      8.7      8.3      8.4
1        5.6      6.0      6.2      6.0  …      6.6      6.2      6.2
2        4.6      4.8      4.9      4.9  …      6.7      6.4      6.4
3        7.2      7.7      7.8      7.5  …      7.9      7.6      7.6
4        8.0      8.1      8.1      7.8  …      8.6      8.2      8.2
..       …        …        …        …    …      …        …        …
290     10.4     10.0     10.5     10.4  …     10.0     10.1     10.0
291      8.9      9.3      9.4      9.3  …      9.4      9.2      9.2
292     11.2     11.4     11.2     11.5  …     11.9     11.4     11.6
293      8.2      8.1      8.0      8.0  …      8.6      7.8      8.2
294      9.0      9.4      9.6      9.7  …     10.1      9.9     10.0

     2019-T2  2019-T3  2019-T4  2020-T1  Code               Libellé  \
0        8.1      8.3      8.0      7.7    84  Auvergne-Rhône-Alpes
1        6.1      6.3      6.0      5.8    84  Auvergne-Rhône-Alpes
2        6.3      6.5      6.2      5.9    84  Auvergne-Rhône-Alpes
3        7.4      7.4      7.3      7.2    84  Auvergne-Rhône-Alpes
4        7.9      7.9      7.7      7.3    84  Auvergne-Rhône-Alpes
..       …        …        …        …    …                    …
290      9.5      9.6      9.5      9.4    94               Corse
291      8.8      8.5      8.2      7.9    94               Corse
292     11.0     10.8     10.5     10.1    94               Corse
293      7.6      7.3      7.1      7.3    94               Corse
294      9.4      9.4      9.0      8.6    94               Corse

     Pôle emploi 2018
0                 294
1                 294
2                 294
3                 294
4                 294
..                …
290                31
291                31
292                31
```

```
293                31
294                31

[295 rows x 76 columns]
```

[220]:
```
fusion ['quantile'] = pd.qcut(fusion['Pôle emploi 2018'], q = 4, labels =␣
↪['q25', 'q50', 'q75', 'q100'])
fusion
```

[220]:
```
    ZE2010           LIBZE2010   REG                LIBREG  2003-T1  2003-T2  \
0    0060*       Saint-Étienne  84.0  Auvergne-Rhône-Alpes      8.0      8.2
1     8201   Ambérieu-en-Bugey  84.0  Auvergne-Rhône-Alpes      5.7      5.8
2     8202     Bourg-en-Bresse  84.0  Auvergne-Rhône-Alpes      4.5      4.7
3     8203             Oyonnax  84.0  Auvergne-Rhône-Alpes      6.8      6.9
4     8204             Annonay  84.0  Auvergne-Rhône-Alpes      7.6      8.0
..     …                  …   …                    …      …      …
290   9403   Sartène-Propriano  94.0               Corse     10.7     10.7
291   9404             Bastia  94.0               Corse      9.2      8.9
292   9405  Calvi-L'Île-Rousse  94.0               Corse     11.7     11.4
293   9406              Corte  94.0               Corse      8.3      7.6
294   9407  Ghisonaccia-Aléria  94.0               Corse      9.6      8.8

     2003-T3  2003-T4  2004-T1  2004-T2  …  2018-T4  2019-T1  2019-T2  \
0        8.1      8.4      8.6      8.3  …      8.3      8.4      8.1
1        5.6      6.0      6.2      6.0  …      6.2      6.2      6.1
2        4.6      4.8      4.9      4.9  …      6.4      6.4      6.3
3        7.2      7.7      7.8      7.5  …      7.6      7.6      7.4
4        8.0      8.1      8.1      7.8  …      8.2      8.2      7.9
..       …      …        …        …    …      …        …        …
290     10.4     10.0     10.5     10.4  …     10.1     10.0      9.5
291      8.9      9.3      9.4      9.3  …      9.2      9.2      8.8
292     11.2     11.4     11.2     11.5  …     11.4     11.6     11.0
293      8.2      8.1      8.0      8.0  …      7.8      8.2      7.6
294      9.0      9.4      9.6      9.7  …      9.9     10.0      9.4

     2019-T3  2019-T4  2020-T1  Code              Libellé  Pôle emploi 2018  \
0        8.3      8.0      7.7    84  Auvergne-Rhône-Alpes               294
1        6.3      6.0      5.8    84  Auvergne-Rhône-Alpes               294
2        6.5      6.2      5.9    84  Auvergne-Rhône-Alpes               294
3        7.4      7.3      7.2    84  Auvergne-Rhône-Alpes               294
4        7.9      7.7      7.3    84  Auvergne-Rhône-Alpes               294
..       …      …        …    …                    …                 …
290      9.6      9.5      9.4    94               Corse                31
291      8.5      8.2      7.9    94               Corse                31
292     10.8     10.5     10.1    94               Corse                31
293      7.3      7.1      7.3    94               Corse                31
294      9.4      9.0      8.6    94               Corse                31
```

```
     quantile
0        q100
1        q100
2        q100
3        q100
4        q100
..        …
290       q25
291       q25
292       q25
293       q25
294       q25

[295 rows x 77 columns]
```

[221]:
```
print(pd.pivot_table(fusion, values = '2020-T1', index = 'LIBREG', columns =␣
↪'quantile'))
```

```
quantile                        q25       q50       q75       q100
LIBREG
Auvergne-Rhône-Alpes            NaN       NaN       NaN  6.728571
Bourgogne-Franche-Comté         NaN  6.565217       NaN       NaN
Bretagne                   6.661111       NaN       NaN       NaN
Centre-Val-de-Loire        7.684211       NaN       NaN       NaN
Corse                      8.514286       NaN       NaN       NaN
Grand-Est                       NaN       NaN  7.264516       NaN
Hauts-de-France                 NaN  9.911538       NaN       NaN
Ile-de-France                   NaN  6.794737       NaN       NaN
Normandie                  7.150000       NaN       NaN       NaN
Nouvelle-Aquitaine              NaN       NaN  7.293548       NaN
Occitanie                       NaN       NaN       NaN  9.603333
Pays de la Loire           6.563158       NaN       NaN       NaN
Provence-Alpes-Côte d'Azur 8.864706       NaN       NaN       NaN
```

[ ]: