relatorio_data500

February 14, 2023

[ ]:
```
import pandas as pd
import numpy as np
```

[ ]:
```
data_hora = pd.read_excel('C:
 ↪\\Users\\Riallen\\Documents\\Print_de_telas\\data1\\hora_apostas.xlsx')
print(data_hora.head(10))
data_hora['Data_Hora_Aposta'] = pd.to_datetime(data_hora['Hora_Aposta'])
data_hora['Data_Hora_Aposta'] = data_hora['Data_Hora_Aposta'].dt.date
#data_hora['Hora_Aposta'] = data_hora['Hora_Aposta'].dt.strftime('%H:%M:%S')
#data_hora.rename(columns={'03:55:45': ''}, inplace=True)
#print(data_hora)
#print(len(data_hora))
#t = len(data_hora)
#for i in range(0,t):
#    print(data_hora['Hora_Aposta'][i])
#    print(type(data_hora['Hora_Aposta'][i]))
data_hora.dtypes
```

```
  Hora_Aposta
0    03:55:45
1    03:55:58
2    03:56:20
3    03:56:30
4    03:56:53
5    03:57:06
6    03:57:33
7    03:57:44
8    03:57:56
9    03:58:12
```

[ ]:
```
Hora_Aposta         object
Data_Hora_Aposta    object
dtype: object
```

[ ]:
```
data_hora["hour"] = data_hora['Hora_Aposta'].str.replace(":.+", "", regex=True).
 ↪astype("int")
```

[ ]:
```
data_hora
```

[ ]:

	Hora_Aposta	Data_Hora_Aposta	hour
0	03:55:45	2023-02-13	3
1	03:55:58	2023-02-13	3
2	03:56:20	2023-02-13	3
3	03:56:30	2023-02-13	3
4	03:56:53	2023-02-13	3
..	…	…	…
496	07:24:55	2023-02-13	7
497	07:25:10	2023-02-13	7
498	07:25:24	2023-02-13	7
499	07:25:42	2023-02-13	7
500	07:25:59	2023-02-13	7

[501 rows x 3 columns]

[ ]:
```
data_qtapost = pd.read_excel('C:
 ↪\\Users\\Riallen\\Documents\\Print_de_telas\\data1\\qt_apostadores1.xlsx')
data_qtapost.head()
t = len(data_qtapost)

for i in range(0,t):
    if data_qtapost['Qt_Apostadores'][i] // 100 == 0:
        print("Elemento: ", data_qtapost['Qt_Apostadores'][i])
        print("Indice: ", i)
```

```
Elemento:  77
Indice:  63
Elemento:  57
Indice:  89
Elemento:  50
Indice:  117
Elemento:  86
Indice:  129
Elemento:  77
Indice:  287
Elemento:  7
Indice:  417
Elemento:  4
Indice:  450
```

[ ]:
```
data_qtapost['Qt_Apostadores'][63] = 577
data_qtapost['Qt_Apostadores'][89] = 571
data_qtapost['Qt_Apostadores'][117] = 850
data_qtapost['Qt_Apostadores'][129] = 886
data_qtapost['Qt_Apostadores'][287] = 1177
data_qtapost['Qt_Apostadores'][417] = 1777
data_qtapost['Qt_Apostadores'][450] = 1771
```

[ ]:
```
t = len(data_qtapost)

for i in range(0,t):
    if data_qtapost['Qt_Apostadores'][i] // 10000 >= 1:
        print("Elemento: ", data_qtapost['Qt_Apostadores'][i])
        print("Indice: ", i)
```

```
Elemento:  14143
Indice:  301
```

[ ]:
```
data_qtapost['Qt_Apostadores'][301] = 1413
```

[ ]:
```
data_qtapost
```

[ ]:
```
     Qt_Apostadores
0               597
1               598
2               535
3               696
4               648
..              …
496            2557
497            2626
498            1942
499            1844
500            2034

[501 rows x 1 columns]
```

[ ]:
```
data_odds = pd.read_excel("C:
 ↪\\Users\\Riallen\\Documents\\Print_de_telas\\data1\\odd.xlsx")
data_odds.head()
#data_odds.describe()
#data_odds.dtypes
```

[ ]:
```
   Odds
0  3.54
1  1.19
2  2.58
3  1.00
4  2.60
```

[ ]:
```
data_odds['Odds'].describe()
```

[ ]:
```
count    501.000000
mean       9.654990
std       46.210931
min        1.000000
```

```
25%        1.350000
50%        2.080000
75%        4.140000
max      810.900000
Name: Odds, dtype: float64
```

[ ]:
```
data_geral = pd.concat([data_odds, data_qtapost, data_hora], axis = 1)
data_geral
```

[ ]:

	Odds	Qt_Apostadores	Hora_Aposta	Data_Hora_Aposta	hour
0	3.54	597	03:55:45	2023-02-13	3
1	1.19	598	03:55:58	2023-02-13	3
2	2.58	535	03:56:20	2023-02-13	3
3	1.00	696	03:56:30	2023-02-13	3
4	2.60	648	03:56:53	2023-02-13	3
..	…	…	…	… …	
496	12.47	2557	07:24:55	2023-02-13	7
497	1.32	2626	07:25:10	2023-02-13	7
498	1.22	1942	07:25:24	2023-02-13	7
499	1.08	1844	07:25:42	2023-02-13	7
500	1.65	2034	07:25:59	2023-02-13	7

[501 rows x 5 columns]

[ ]:
```
data_geral.to_excel('C:
↪\\Users\\Riallen\\Documents\\Print_de_telas\\data1\\data_geral.xlsx', index␣
↪= False)
```