

Analiza liczby zgonów

September 27, 2022

```
[1]: import pandas as pd
import numpy as np
import matplotlib as plt
import matplotlib.pyplot as plt
```

```
[2]: #Źródło danych:https://dane.gov.pl/pl/dataset/
↳1953,liczba-zgonow-zarejestrowanych-w-rejestrze-stanu-cywilnego
df = pd.read_csv(r'C:\Users\Damian\Desktop\analiza danych\zgony\Liczba_Zgonow.
↳csv') #Liczba zgonów zarejestrowanych w Rejestrze Stanu Cywilnego
df.head(5)
```

```
[2]:
```

	Kod TERYT	Nazwa województwa	Rok zgonu	Miesiąc zgonu	\
0	2	DOLNOŚLĄSKIE	2022	1	
1	2	DOLNOŚLĄSKIE	2022	2	
2	2	DOLNOŚLĄSKIE	2022	3	
3	2	DOLNOŚLĄSKIE	2022	4	
4	2	DOLNOŚLĄSKIE	2022	5	

```
Zgony kobiet Przedział wiekowy 0-64 lata \
```

0	258
1	203
2	216
3	188
4	177

```
Zgony kobiet Przedział wiekowy 65 lat i więcej \
```

0	1693
1	1440
2	1383
3	1241
4	1130

```
Zgony mężczyzn Przedział wiekowy 0-64 lata \
```

0	555
1	419
2	434
3	437

```

Zgony mężczyzn Przedział wiekowy 65 lat i więcej
0      1370
1      1191
2      1187
3      1062
4       990

```

```

[3]: df['suma_zgonow'] = df['Zgony kobiet Przedział wiekowy 0-64 lata'] + df['Zgony_
    ↪kobiet Przedział wiekowy 65 lat i więcej'] + df['Zgony mężczyzn Przedział_
    ↪wiekowy 0-64 lata'] + df['Zgony mężczyzn Przedział wiekowy 65 lat i więcej']
df

```

```

[3]:
Kod TERYT    Nazwa województwa    Rok zgonu    Miesiąc zgonu  \
0           2      DOLNOŚLĄSKIE      2022           1
1           2      DOLNOŚLĄSKIE      2022           2
2           2      DOLNOŚLĄSKIE      2022           3
3           2      DOLNOŚLĄSKIE      2022           4
4           2      DOLNOŚLĄSKIE      2022           5
...         ...                ...         ...
1339        30      WIELKOPOLSKIE      2015          12
1340        32  ZACHODNIOPOMORSKIE      2015           9
1341        32  ZACHODNIOPOMORSKIE      2015          10
1342        32  ZACHODNIOPOMORSKIE      2015          11
1343        32  ZACHODNIOPOMORSKIE      2015          12

```

```

Zgony kobiet Przedział wiekowy 0-64 lata  \
0      258
1      203
2      216
3      188
4      177
...         ...
1339      227
1340      127
1341      110
1342      141
1343      132

```

```

Zgony kobiet Przedział wiekowy 65 lat i więcej  \
0      1693
1      1440
2      1383
3      1241
4      1130
...         ...

```

1339	1029
1340	477
1341	521
1342	522
1343	541

Zgony mężczyzn Przedział wiekowy 0-64 lata \	
0	555
1	419
2	434
3	437
4	407
...	...
1339	503
1340	249
1341	284
1342	290
1343	292

Zgony mężczyzn Przedział wiekowy 65 lat i więcej		suma_zgonow
0	1370	3876
1	1191	3253
2	1187	3220
3	1062	2928
4	990	2704
...
1339	883	2642
1340	389	1242
1341	456	1371
1342	414	1367
1343	460	1425

[1344 rows x 9 columns]

```
[4]: df.describe()
```

```
[4]:
```

	Kod TERYT	Rok zgonu	Miesiąc zgonu \
count	1344.000000	1344.000000	1344.000000
mean	17.000000	2018.666667	6.500000
std	9.222976	2.055570	3.453337
min	2.000000	2015.000000	1.000000
25%	9.500000	2017.000000	3.750000
50%	17.000000	2019.000000	6.500000
75%	24.500000	2020.000000	9.250000
max	32.000000	2022.000000	12.000000

Zgony kobiet Przedział wiekowy 0-64 lata \	
--	--

count	1344.000000
mean	153.904762
std	89.568045
min	37.000000
25%	84.000000
50%	132.000000
75%	195.000000
max	530.000000

Zgony kobiet Przedział wiekowy 65 lat i więcej \

count	1344.000000
mean	920.152530
std	543.633867
min	250.000000
25%	500.500000
50%	760.000000
75%	1131.000000
max	3915.000000

Zgony mężczyzn Przedział wiekowy 0-64 lata \

count	1344.000000
mean	369.782738
std	198.011602
min	114.000000
25%	227.000000
50%	312.000000
75%	456.250000
max	1248.000000

	Zgony mężczyzn Przedział wiekowy 65 lat i więcej	suma_zgonow
count	1344.000000	1344.000000
mean	783.843750	2227.683780
std	464.444676	1281.495178
min	208.000000	673.000000
25%	439.000000	1238.250000
50%	651.000000	1847.500000
75%	954.500000	2716.500000
max	3528.000000	9193.000000

```
[5]: df.sort_values(by=['suma_zgonow'])
dane_suma = df.loc[:, ['Nazwa województwa', 'Rok zgonu', 'Miesiąc zgonu',
↪ 'suma_zgonow']]
dane_suma
```

```
[5]:
```

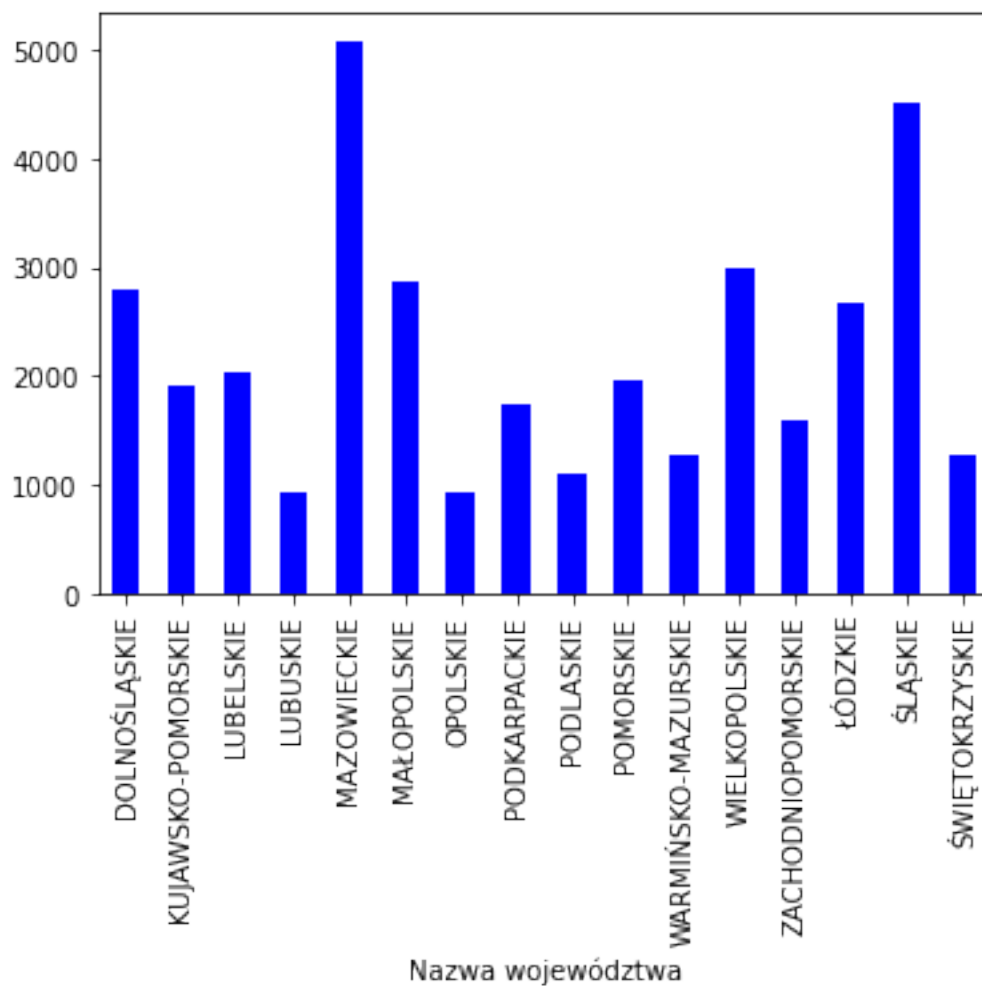
	Nazwa województwa	Rok zgonu	Miesiąc zgonu	suma_zgonow
0	DOLNOŚLĄSKIE	2022	1	3876
1	DOLNOŚLĄSKIE	2022	2	3253

2	DOLNOŚLĄSKIE	2022	3	3220
3	DOLNOŚLĄSKIE	2022	4	2928
4	DOLNOŚLĄSKIE	2022	5	2704
...
1339	WIELKOPOLSKIE	2015	12	2642
1340	ZACHODNIOPOMORSKIE	2015	9	1242
1341	ZACHODNIOPOMORSKIE	2015	10	1371
1342	ZACHODNIOPOMORSKIE	2015	11	1367
1343	ZACHODNIOPOMORSKIE	2015	12	1425

[1344 rows x 4 columns]

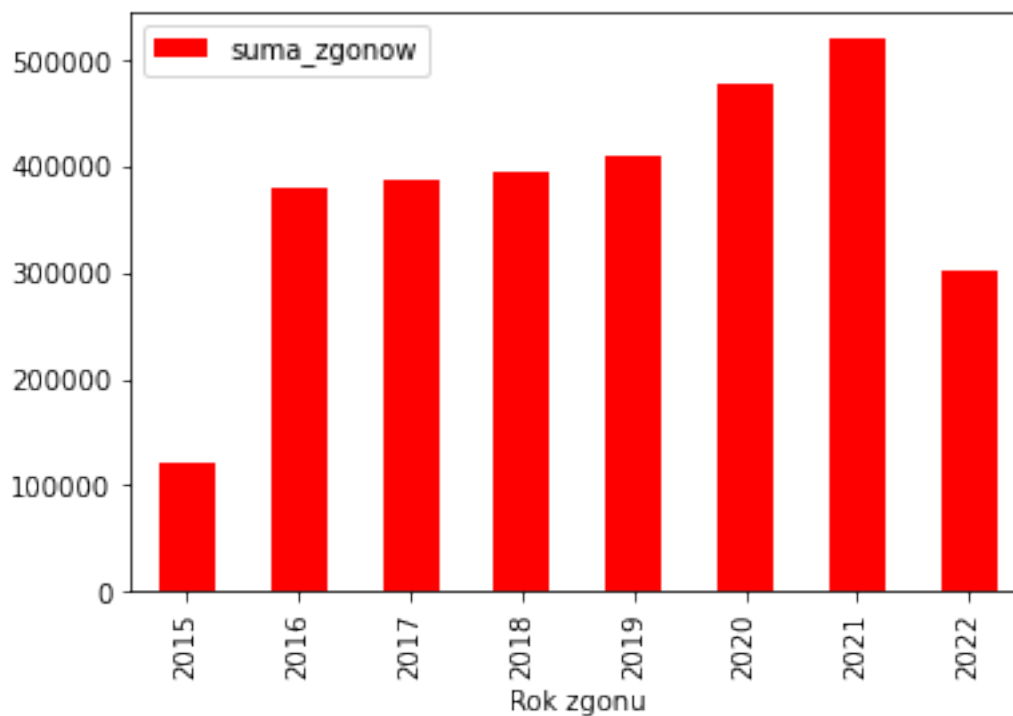
```
[6]: #Wykres średniej całkowitej liczby zgonów z podziałem na województwa
by_state = dane_suma.groupby('Nazwa województwa').suma_zgonow.mean()
by_state.plot(kind='bar', color='blue',)
```

```
[6]: <AxesSubplot: xlabel='Nazwa województwa'>
```



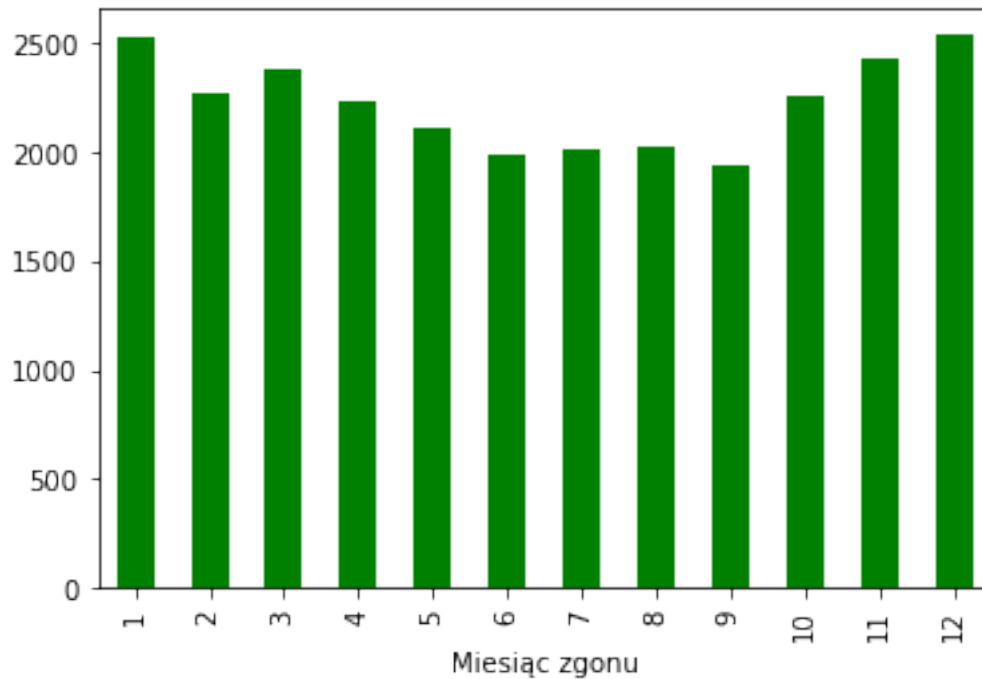
```
[7]: #suma zgonów na przestrzeni lat
dane_suma_2 = df.loc[:, ['Rok zgonu', 'suma_zgonow']]
by_year = dane_suma_2.groupby('Rok zgonu').sum()
by_year.plot(kind='bar', color='red')
```

```
[7]: <AxesSubplot:xlabel='Rok zgonu'>
```



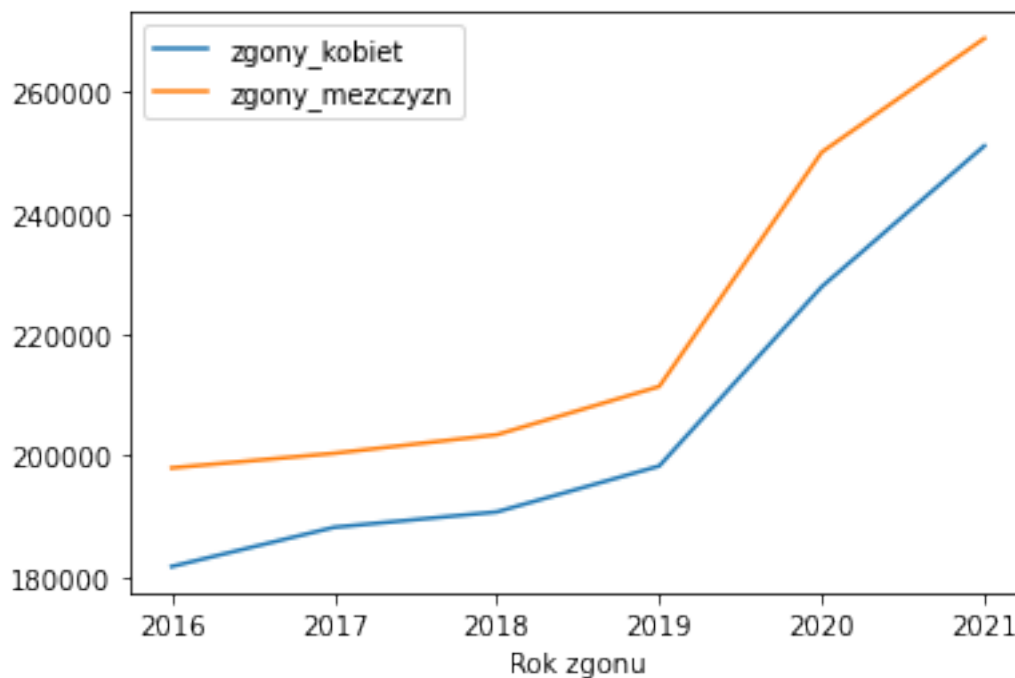
```
[8]: #Wykres średniej całkowitej liczby zgonów
by_month = dane_suma.groupby('Miesiąc zgonu').suma_zgonow.mean()
by_month.plot(kind='bar', color='green')
```

```
[8]: <AxesSubplot:xlabel='Miesiąc zgonu'>
```



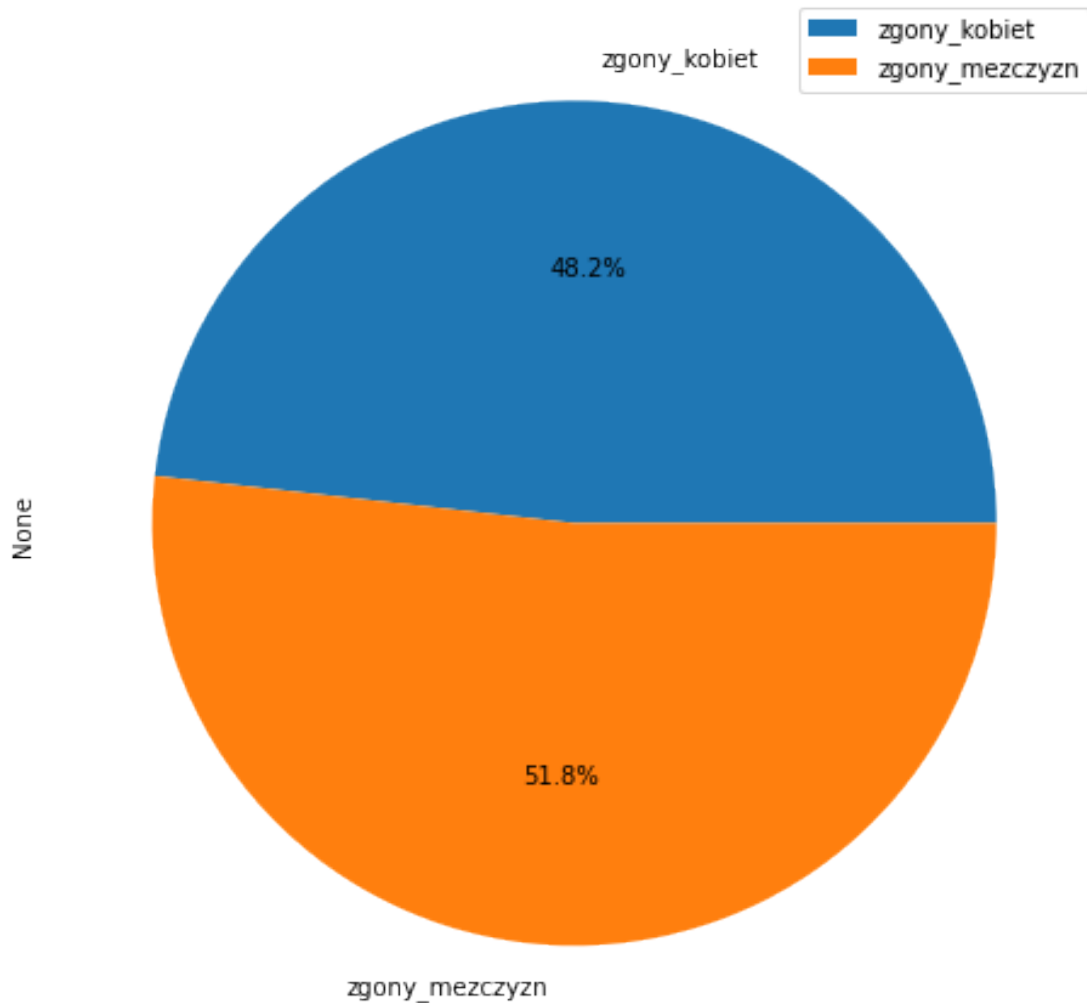
```
[9]: #Wykres całkowitej liczby zgonów na przestrzeni lat z podziałem na płeć
df['zgony_kobiet'] = df['Zgony kobiet Przedział wiekowy 0-64 lata'] + df['Zgony_
↳kobiet Przedział wiekowy 65 lat i więcej']
df['zgony_mezczyzn'] = df['Zgony mężczyzn Przedział wiekowy 0-64 lata'] +
↳df['Zgony mężczyzn Przedział wiekowy 65 lat i więcej']
dane = df.loc[:, ['Rok zgonu', 'zgony_kobiet', 'zgony_mezczyzn']]
dane.set_index('Rok zgonu')
dane.groupby('Rok zgonu').sum().drop([2015, 2022]).plot()
```

```
[9]: <AxesSubplot:xlabel='Rok zgonu'>
```



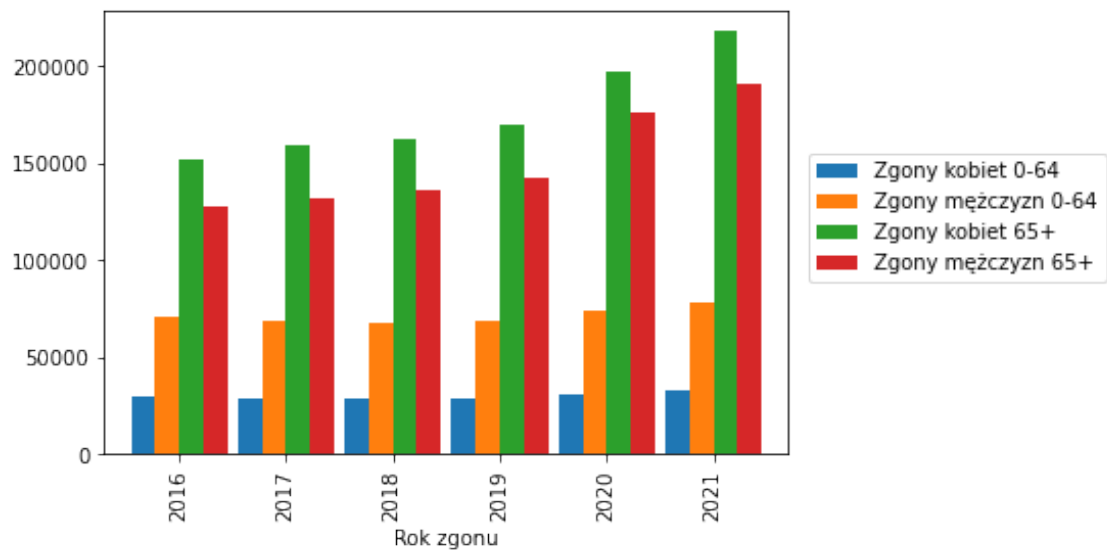
```
[10]: #Procentowy udział mężczyzn i kobiet w całkowitej liczbie zgonów  
dane_sum = dane.drop('Rok zgonu', axis=1).sum()  
dane_sum.plot.pie(autopct='%1.1f%%', legend=True, figsize=(8, 8))
```

```
[10]: <AxesSubplot:ylabel='None'>
```

```
[11]: #Wykres słupkowy liczby zgonów na przestrzeni lat z podziałem na kategorie
      ↪ wiekowe
dane_pelne = df.loc[:, ['Rok zgonu', 'Zgony kobiet Przedział wiekowy 0-64',
      ↪ lata', 'Zgony mężczyzn Przedział wiekowy 0-64 lata', 'Zgony kobiet Przedział
      ↪ wiekowy 65 lat i więcej', 'Zgony mężczyzn Przedział wiekowy 65 lat i
      ↪ więcej']].groupby('Rok zgonu').sum().drop([2015, 2022])
dane_pelne.plot.bar(width=0.9,)
plt.legend(['Zgony kobiet 0-64', 'Zgony mężczyzn 0-64', 'Zgony kobiet 65+',
      ↪ 'Zgony mężczyzn 65+'], bbox_to_anchor = (1.48, 0.7))
```

```
[11]: <matplotlib.legend.Legend at 0x24d26a2b700>
```



```
[12]: #Wykres liczby zgonów na przestrzeni lat z podziałem na kategorie wiekowe
      dane_pelne.plot()
      plt.legend(['Zgony kobiet 0-64', 'Zgony mężczyzn 0-64', 'Zgony kobiet 65+',
                  ↪ 'Zgony mężczyzn 65+'], bbox_to_anchor = (1.48, 0.7))
```

[12]: <matplotlib.legend.Legend at 0x24d26a7baf0>

