Analiza_penzioneri

August 1, 2025

[1]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

[2]:
```
umirovljenici = pd.read_excel("../../datasets/penzioneri.xlsx",␣
 ↪sheet_name="drugi")
```

[3]:
```
umirovljenici.head(10)
```

[3]:

	grad	starost	zanimanje	spol	primanja	stednja
0	Split	55.0	obrtnik	m	2500.0	10000.0
1	Zagreb	77.0	penzioner	m	1700.0	5000.0
2	Split	62.0	penzioner	m	1800.0	2000.0
3	Zagreb	NaN	obrtnik	f	3500.0	4500.0
4	Split	59.0	NaN	NaN	NaN	NaN
5	Zagreb	71.0	penzioner	m	1250.0	6500.0
6	Osijek	59.0	obrtnik	m	4000.0	12000.0
7	Pula	NaN	obrtnik	NaN	3000.0	10000.0
8	Zagreb	65.0	NaN	m	2000.0	7000.0
9	Split	52.0	obrtnik	m	5000.0	NaN

[4]:
```
umirovljenici.isna().sum()
```

[4]:
```
grad         0
starost      3
zanimanje    2
spol         2
primanja     1
stednja      3
dtype: int64
```

[5]:
```
isna_values = umirovljenici[umirovljenici.isna().any(axis=1)]
```

[6]:
```
isna_values
```

[6]:

	grad	starost	zanimanje	spol	primanja	stednja
3	Zagreb	NaN	obrtnik	f	3500.0	4500.0
4	Split	59.0	NaN	NaN	NaN	NaN

```
7     Pula      NaN   obrtnik  NaN    3000.0  10000.0
8   Zagreb     65.0       NaN    m    2000.0   7000.0
9    Split     52.0   obrtnik    m    5000.0      NaN
11    Pula      NaN   obrtnik    m    4500.0      NaN
```

[7]:
```
umirovljenici2 = umirovljenici.copy()
```

[8]:
```
umirovljenici2.dropna(how = "any")
```

[8]:
```
      grad  starost  zanimanje spol  primanja  stednja
0    Split     55.0    obrtnik    m    2500.0  10000.0
1   Zagreb     77.0  penzioner    m    1700.0   5000.0
2    Split     62.0  penzioner    m    1800.0   2000.0
5   Zagreb     71.0  penzioner    m    1250.0   6500.0
6   Osijek     59.0    obrtnik    m    4000.0  12000.0
10  Zagreb     60.0  penzioner    f    1700.0   2500.0
12  Osijek     58.0    obrtnik    f    3000.0  11000.0
13  Zagreb     83.0  penzioner    m    3500.0  14000.0
14   Split     74.0    obrtnik    f    4900.0  16000.0
15  Zagreb     92.0  penzioner    m    6000.0  20000.0
16  Osijek     62.0    obrtnik    m    5200.0   6200.0
17    Pula     69.0    obrtnik    m    4700.0   9200.0
18  Osijek     76.0    obrtnik    m    4900.0   8200.0
19  Zagreb     59.0    obrtnik    f    6300.0  25000.0
20  Zagreb     75.0  penzioner    m    8200.0  33000.0
```

[9]:
```
umirovljenici3 = umirovljenici.copy()
```

[10]:
```
umirovljenici3.head(10)
```

[10]:
```
     grad  starost  zanimanje spol  primanja  stednja
0   Split     55.0    obrtnik    m    2500.0  10000.0
1  Zagreb     77.0  penzioner    m    1700.0   5000.0
2   Split     62.0  penzioner    m    1800.0   2000.0
3  Zagreb      NaN    obrtnik    f    3500.0   4500.0
4   Split     59.0        NaN  NaN       NaN      NaN
5  Zagreb     71.0  penzioner    m    1250.0   6500.0
6  Osijek     59.0    obrtnik    m    4000.0  12000.0
7    Pula      NaN    obrtnik  NaN    3000.0  10000.0
8  Zagreb     65.0        NaN    m    2000.0   7000.0
9   Split     52.0    obrtnik    m    5000.0      NaN
```

[11]:
```
umirovljenici3.dropna(thresh=5, inplace = True)
```

[12]:
```
umirovljenici3
```

[12]:

	grad	starost	zanimanje	spol	primanja	stednja
0	Split	55.0	obrtnik	m	2500.0	10000.0
1	Zagreb	77.0	penzioner	m	1700.0	5000.0
2	Split	62.0	penzioner	m	1800.0	2000.0
3	Zagreb	NaN	obrtnik	f	3500.0	4500.0
5	Zagreb	71.0	penzioner	m	1250.0	6500.0
6	Osijek	59.0	obrtnik	m	4000.0	12000.0
8	Zagreb	65.0	NaN	m	2000.0	7000.0
9	Split	52.0	obrtnik	m	5000.0	NaN
10	Zagreb	60.0	penzioner	f	1700.0	2500.0
12	Osijek	58.0	obrtnik	f	3000.0	11000.0
13	Zagreb	83.0	penzioner	m	3500.0	14000.0
14	Split	74.0	obrtnik	f	4900.0	16000.0
15	Zagreb	92.0	penzioner	m	6000.0	20000.0
16	Osijek	62.0	obrtnik	m	5200.0	6200.0
17	Pula	69.0	obrtnik	m	4700.0	9200.0
18	Osijek	76.0	obrtnik	m	4900.0	8200.0
19	Zagreb	59.0	obrtnik	f	6300.0	25000.0
20	Zagreb	75.0	penzioner	m	8200.0	33000.0

[13]:
```
umirovljenici3a = umirovljenici.copy()
```

[14]:
```
umirovljenici3a
```

[14]:

	grad	starost	zanimanje	spol	primanja	stednja
0	Split	55.0	obrtnik	m	2500.0	10000.0
1	Zagreb	77.0	penzioner	m	1700.0	5000.0
2	Split	62.0	penzioner	m	1800.0	2000.0
3	Zagreb	NaN	obrtnik	f	3500.0	4500.0
4	Split	59.0	NaN	NaN	NaN	NaN
5	Zagreb	71.0	penzioner	m	1250.0	6500.0
6	Osijek	59.0	obrtnik	m	4000.0	12000.0
7	Pula	NaN	obrtnik	NaN	3000.0	10000.0
8	Zagreb	65.0	NaN	m	2000.0	7000.0
9	Split	52.0	obrtnik	m	5000.0	NaN
10	Zagreb	60.0	penzioner	f	1700.0	2500.0
11	Pula	NaN	obrtnik	m	4500.0	NaN
12	Osijek	58.0	obrtnik	f	3000.0	11000.0
13	Zagreb	83.0	penzioner	m	3500.0	14000.0
14	Split	74.0	obrtnik	f	4900.0	16000.0
15	Zagreb	92.0	penzioner	m	6000.0	20000.0
16	Osijek	62.0	obrtnik	m	5200.0	6200.0
17	Pula	69.0	obrtnik	m	4700.0	9200.0
18	Osijek	76.0	obrtnik	m	4900.0	8200.0
19	Zagreb	59.0	obrtnik	f	6300.0	25000.0
20	Zagreb	75.0	penzioner	m	8200.0	33000.0

[15]:
```
umirovljenici3a.dropna(thresh=2, subset=['grad', 'starost', 'zanimanje',␣
↪'stednja'])
```

[15]:

	grad	starost	zanimanje	spol	primanja	stednja
0	Split	55.0	obrtnik	m	2500.0	10000.0
1	Zagreb	77.0	penzioner	m	1700.0	5000.0
2	Split	62.0	penzioner	m	1800.0	2000.0
3	Zagreb	NaN	obrtnik	f	3500.0	4500.0
4	Split	59.0	NaN	NaN	NaN	NaN
5	Zagreb	71.0	penzioner	m	1250.0	6500.0
6	Osijek	59.0	obrtnik	m	4000.0	12000.0
7	Pula	NaN	obrtnik	NaN	3000.0	10000.0
8	Zagreb	65.0	NaN	m	2000.0	7000.0
9	Split	52.0	obrtnik	m	5000.0	NaN
10	Zagreb	60.0	penzioner	f	1700.0	2500.0
11	Pula	NaN	obrtnik	m	4500.0	NaN
12	Osijek	58.0	obrtnik	f	3000.0	11000.0
13	Zagreb	83.0	penzioner	m	3500.0	14000.0
14	Split	74.0	obrtnik	f	4900.0	16000.0
15	Zagreb	92.0	penzioner	m	6000.0	20000.0
16	Osijek	62.0	obrtnik	m	5200.0	6200.0
17	Pula	69.0	obrtnik	m	4700.0	9200.0
18	Osijek	76.0	obrtnik	m	4900.0	8200.0
19	Zagreb	59.0	obrtnik	f	6300.0	25000.0
20	Zagreb	75.0	penzioner	m	8200.0	33000.0

[16]:
```
umirovljenici4 = umirovljenici.copy()
```

[17]:
```
prosjek_godina = int(umirovljenici4["starost"].mean())
```

[18]:
```
prosjek_godina
```

[18]: 67

[19]:
```
umirovljenici4["starost"] = umirovljenici4["starost"].fillna(prosjek_godina)
```

[20]:
```
umirovljenici4
```

[20]:

	grad	starost	zanimanje	spol	primanja	stednja
0	Split	55.0	obrtnik	m	2500.0	10000.0
1	Zagreb	77.0	penzioner	m	1700.0	5000.0
2	Split	62.0	penzioner	m	1800.0	2000.0
3	Zagreb	67.0	obrtnik	f	3500.0	4500.0
4	Split	59.0	NaN	NaN	NaN	NaN
5	Zagreb	71.0	penzioner	m	1250.0	6500.0
6	Osijek	59.0	obrtnik	m	4000.0	12000.0
7	Pula	67.0	obrtnik	NaN	3000.0	10000.0
8	Zagreb	65.0	NaN	m	2000.0	7000.0

```
9    Split     52.0    obrtnik    m    5000.0      NaN
10  Zagreb     60.0  penzioner    f    1700.0   2500.0
11    Pula     67.0    obrtnik    m    4500.0      NaN
12  Osijek     58.0    obrtnik    f    3000.0  11000.0
13  Zagreb     83.0  penzioner    m    3500.0  14000.0
14   Split     74.0    obrtnik    f    4900.0  16000.0
15  Zagreb     92.0  penzioner    m    6000.0  20000.0
16  Osijek     62.0    obrtnik    m    5200.0   6200.0
17    Pula     69.0    obrtnik    m    4700.0   9200.0
18  Osijek     76.0    obrtnik    m    4900.0   8200.0
19  Zagreb     59.0    obrtnik    f    6300.0  25000.0
20  Zagreb     75.0  penzioner    m    8200.0  33000.0
```

[21]:
```
umirovljenici4["spol"].mode()[0]
```

[21]:
```
'm'
```

[22]:
```
umirovljenici4["spol"] = umirovljenici4["spol"].fillna(value = "m")
```

[23]:
```
umirovljenici4["zanimanje"] = umirovljenici4["zanimanje"].fillna(value =␣
↪umirovljenici4["zanimanje"].mode()[0])
```

[24]:
```
umirovljenici4
```

[24]:
```
      grad  starost  zanimanje spol  primanja  stednja
0    Split     55.0    obrtnik    m    2500.0  10000.0
1   Zagreb     77.0  penzioner    m    1700.0   5000.0
2    Split     62.0  penzioner    m    1800.0   2000.0
3   Zagreb     67.0    obrtnik    f    3500.0   4500.0
4    Split     59.0    obrtnik    m       NaN      NaN
5   Zagreb     71.0  penzioner    m    1250.0   6500.0
6   Osijek     59.0    obrtnik    m    4000.0  12000.0
7     Pula     67.0    obrtnik    m    3000.0  10000.0
8   Zagreb     65.0    obrtnik    m    2000.0   7000.0
9    Split     52.0    obrtnik    m    5000.0      NaN
10  Zagreb     60.0  penzioner    f    1700.0   2500.0
11    Pula     67.0    obrtnik    m    4500.0      NaN
12  Osijek     58.0    obrtnik    f    3000.0  11000.0
13  Zagreb     83.0  penzioner    m    3500.0  14000.0
14   Split     74.0    obrtnik    f    4900.0  16000.0
15  Zagreb     92.0  penzioner    m    6000.0  20000.0
16  Osijek     62.0    obrtnik    m    5200.0   6200.0
17    Pula     69.0    obrtnik    m    4700.0   9200.0
18  Osijek     76.0    obrtnik    m    4900.0   8200.0
19  Zagreb     59.0    obrtnik    f    6300.0  25000.0
20  Zagreb     75.0  penzioner    m    8200.0  33000.0
```

[25]:
```
umirovljenici4["primanja"] = umirovljenici4["primanja"].fillna(value =␣
↪umirovljenici4["primanja"].median())
```

[26]:
```
umirovljenici4["stednja"] = umirovljenici4["stednja"].fillna(value =␣
↪umirovljenici4["stednja"].median())
```

[27]:
```
umirovljenici4
```

[27]:

	grad	starost	zanimanje	spol	primanja	stednja
0	Split	55.0	obrtnik	m	2500.0	10000.0
1	Zagreb	77.0	penzioner	m	1700.0	5000.0
2	Split	62.0	penzioner	m	1800.0	2000.0
3	Zagreb	67.0	obrtnik	f	3500.0	4500.0
4	Split	59.0	obrtnik	m	3750.0	9600.0
5	Zagreb	71.0	penzioner	m	1250.0	6500.0
6	Osijek	59.0	obrtnik	m	4000.0	12000.0
7	Pula	67.0	obrtnik	m	3000.0	10000.0
8	Zagreb	65.0	obrtnik	m	2000.0	7000.0
9	Split	52.0	obrtnik	m	5000.0	9600.0
10	Zagreb	60.0	penzioner	f	1700.0	2500.0
11	Pula	67.0	obrtnik	m	4500.0	9600.0
12	Osijek	58.0	obrtnik	f	3000.0	11000.0
13	Zagreb	83.0	penzioner	m	3500.0	14000.0
14	Split	74.0	obrtnik	f	4900.0	16000.0
15	Zagreb	92.0	penzioner	m	6000.0	20000.0
16	Osijek	62.0	obrtnik	m	5200.0	6200.0
17	Pula	69.0	obrtnik	m	4700.0	9200.0
18	Osijek	76.0	obrtnik	m	4900.0	8200.0
19	Zagreb	59.0	obrtnik	f	6300.0	25000.0
20	Zagreb	75.0	penzioner	m	8200.0	33000.0