

CiscenjeUspjeh

August 1, 2025

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

```
[2]: uspjuh_df = pd.read_csv("../..../datasets/st_uspjuh.csv")
```

```
[3]: uspjuh_df.head()
```

```
[3]:
```

	Unnamed: 0	StudentID	Drzava	Smjer	Godina_Studija	\
0	0	1	Hrvatska	Informatika	1	
1	1	2	Hrvatska	Pravo	1	
2	2	3	Hrvatska	Informatika	5	
3	3	4	Hrvatska	Ekonomija	5	
4	4	5	Hrvatska	Informatika	4	

	Prosjek_Ocjena	Broj_Polozenih_Ispita	Status
0	3.83	8.0	Izvanredan
1	4.31	3.0	Izvanredan
2	2.16	3.0	Redovan
3	3.37	5.0	Redovan
4	4.51	2.0	Apsolvent

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

```
[4]:
```

	Unnamed: 0	StudentID	Godina_Studija	Prosjek_Ocjena	\
count	100.000000	100.000000	100.000000	95.000000	
mean	49.500000	50.500000	2.940000	3.392526	
std	29.011492	29.011492	1.983874	0.749001	
min	0.000000	1.000000	1.000000	1.360000	
25%	24.750000	25.750000	1.000000	2.910000	
50%	49.500000	50.500000	3.000000	3.370000	
75%	74.250000	75.250000	4.000000	3.965000	
max	99.000000	100.000000	12.000000	5.000000	

	Broj_Polozenih_Ispita
count	94.000000
mean	5.723404
std	3.105334

```

min          1.000000
25%          3.000000
50%          5.000000
75%          7.000000
max          14.000000

```

```
[5]: uspjh_df.drop(columns=['Drzava'], inplace=True)
```

```
[6]: uspjh_df.head()
```

```
[6]:
```

	Unnamed: 0	StudentID	Smjer	Godina_Studija	Prosjek_Ocjena	\
0	0	1	Informatika	1	3.83	
1	1	2	Pravo	1	4.31	
2	2	3	Informatika	5	2.16	
3	3	4	Ekonomija	5	3.37	
4	4	5	Informatika	4	4.51	

	Broj_Polozenih_Ispita	Status
0	8.0	Izvanredan
1	3.0	Izvanredan
2	3.0	Redovan
3	5.0	Redovan
4	2.0	Apsolvent

```
[7]: uspjh_df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 7 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Unnamed: 0                            100 non-null    int64
1   StudentID                             100 non-null    int64
2   Smjer                                 93 non-null     object
3   Godina_Studija                         100 non-null    int64
4   Prosjek_Ocjena                         95 non-null     float64
5   Broj_Polozenih_Ispita                 94 non-null     float64
6   Status                                96 non-null     object
dtypes: float64(2), int64(3), object(2)
memory usage: 5.6+ KB

```

```
[8]: uspjh_df.isna().sum()
```

```
[8]:
```

Unnamed: 0	0
StudentID	0
Smjer	7
Godina_Studija	0
Prosjek_Ocjena	5

```
Broj_Polozenih_Ispita    6
Status                   4
dtype: int64
```