



ACADEMIA DE
STUDII ECONOMICE



PACHETE SOFTWARE RAPORTATE AGREGATE

PROF. COORDONATOR
ANDREESCU ANCA

STUDENTE

Cazacu Irina, grupa 1081
Ciurea Claudia, grupa 1082



Introducere

Tema: Analiza numărului mediu de ore pe care un american îl doarme în medie pe zi într-un an

Somnul are o importanță crucială pentru sănătatea și bunăstarea generală a unei persoane. În cele ce urmează se prezintă o analiză a unui set de date cu privire la numărul mediu de ore pe care un American îl doarme într-o zi în funcție de sex, grupa de vârstă și tipul de zi (zi normală, weekend normal și weekend în perioada vacantei), observațiile fiind facute pe un set de date de tip panel (2003-2017).



Setul de date

Setul de date : "Time Americans Spend Slepping.csv "

Sursa : [https://www.kaggle.com/datasets/
thedevastator/how-much-sleep-do-americans-really-get](https://www.kaggle.com/datasets/thedevastator/how-much-sleep-do-americans-really-get)

Număr observații : 944

Variabile :

- *Index* - Numarul observatiei
- *Year* - Anul in care a fost preluata observatia
- *Period* - Perioada de timp căreia îi corespunde media
- *Avg hrs per day sleeping* - Valoarea medie a numarului de ore dormite
- *Standard Error* - Eroarea standard
- *Type of Days* - Tipul zilei
- *Age Group* - grupele de varsta pe care s-au realizat observatiile
- *Activity* - activitatea persoanelor pe care s-au realizat observatiile
- *Sex* - sexul persoanelor

Index	Year	Period	Avg hrs per day sleeping	Standard Error	Type of Days	Age Group	Activity	Sex
0	2003	Annual	8.57	0.018	All days	15 years and over	Sleeping	Both
1	2004	Annual	8.54	0.026	All days	15 years and over	Sleeping	Both
2	2005	Annual	8.62	0.023	All days	15 years and over	Sleeping	Both
3	2006	Annual	8.63	0.024	All days	15 years and over	Sleeping	Both
4	2007	Annual	8.57	0.024	All days	15 years and over	Sleeping	Both
5	2008	Annual	8.6	0.023	All days	15 years and over	Sleeping	Both
6	2009	Annual	8.67	0.023	All days	15 years and over	Sleeping	Both
7	2010	Annual	8.67	0.024	All days	15 years and over	Sleeping	Both
8	2011	Annual	8.71	0.026	All days	15 years and over	Sleeping	Both
9	2012	Annual	8.73	0.026	All days	15 years and over	Sleeping	Both
10	2013	Annual	8.74	0.027	All days	15 years and over	Sleeping	Both
11	2014	Annual	8.78	0.025	All days	15 years and over	Sleeping	Both
12	2015	Annual	8.83	0.027	All days	15 years and over	Sleeping	Both
13	2016	Annual	8.79	0.028	All days	15 years and over	Sleeping	Both
14	2017	Annual	8.8	0.027	All days	15 years and over	Sleeping	Both
15	2003	Annual	8.26	0.023	Nonholiday weekdays	15 years and over	Sleeping	Both
16	2004	Annual	8.21	0.034	Nonholiday weekdays	15 years and over	Sleeping	Both
17	2005	Annual	8.33	0.03	Nonholiday weekdays	15 years and over	Sleeping	Both
18	2006	Annual	8.33	0.032	Nonholiday weekdays	15 years and over	Sleeping	Both
19	2007	Annual	8.29	0.032	Nonholiday weekdays	15 years and over	Sleeping	Both
20	2008	Annual	8.3	0.032	Nonholiday weekdays	15 years and over	Sleeping	Both
21	2009	Annual	8.39	0.033	Nonholiday weekdays	15 years and over	Sleeping	Both
22	2010	Annual	8.38	0.033	Nonholiday weekdays	15 years and over	Sleeping	Both
23	2011	Annual	8.42	0.034	Nonholiday weekdays	15 years and over	Sleeping	Both



PYTHON RAPOARTE AGREGATE

Python este un limbaj de programare puternic și versatil, potrivit pentru prelucrarea analitică a datelor și crearea de rapoarte aggregate.

Există mai multe biblioteci și instrumente disponibile în Python care facilitează acest proces:

- biblioteci specializate pentru analiza datelor, cum ar fi **Pandas** pentru manipularea datelor tabulare
- biblioteci specializate pe calcule matematice și statistice precum **Numpy / Scipy**
- biblioteci specializate pentru vizualizarea grafică, precum **Matplotlib / Seaborn**
- lucrul cu tipul de date **DataFrame**, operații de filtrare, agregare și alte transformări pe tabele



Grupări și agregări

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plot
import seaborn as sns
from numpy import savetxt
```

```
→ df = pd.read_csv("Time Americans Spend Sleeping.csv")
→ df2 = df.groupby(['Year', 'Sex'], as_index=False)['Avg hrs per day sleeping'].mean()
→ df3 = df.groupby(['Age Group', 'Sex', 'Year'], as_index=False)['Avg hrs per day
sleeping'].agg(['mean', 'min', 'max'])
→ df4 = df.groupby(['Type of Days','Year'])['index'].count()
→ age_groups = ['15 years and over', '18 years and over']
→ filtered_data = df[df['Age Group'].isin(age_groups)]
→ grouped_data = filtered_data.groupby(['Year', 'Sex']).agg({'Avg hrs per day sleeping': 'mean'})
```



Grupări și agregări

- `df['Rank'] = df.groupby(['Age Group', 'Sex'])['Avg hrs per day sleeping'].transform(lambda x:
x.rank(ascending=False, method='min'))`
`ranked_data=df[['Age Group', 'Sex','Avg hrs per day sleeping','Rank']]`
`ranked_data.to_csv("Ranked_age_sex.csv")`
- `grouped_data = df.groupby(['Year', 'Type of Days']).agg({'Avg hrs per day sleeping': 'sum'})`
`total_sleep_per_year = grouped_data.groupby('Year').transform('sum')`
`grouped_data['Percentage Contribution'] = grouped_data / total_sleep_per_year * 100`
- `grouped_data = df.groupby(['Age Group', 'Activity']).agg({'Avg hrs per day sleeping': ['mean', 'std']})`
`correlation = df['Avg hrs per day sleeping'].corr(df['Year'])`



Grupări și agregări

- `f_filter = df['Sex'] == 'Women'`
- `df_filtered = df[f_filter]`
- `df_filtered_gr = df_filtered.groupby('Year')['Avg hrs per day sleeping'].mean()`
- `f_avg = df[f_filter].agg(f_MINavgHours=('Avg hrs per day sleeping', np.min),`
 `f_MAXavgHours=('Avg hrs per day sleeping', np.max))`
- `avg_hours = df[['Avg hrs per day sleeping']].values`
- `day = df['Type of Days'].unique()`
- `variable_observe = list(df.columns)[:,`



Grupări și agregări

→ **matrix = df[variable_observe].values**

matrix_sel = matrix [:315:15]

matrix_h = df["Avg hrs per day sleeping"].values

hours_list_2003_Both = matrix_h[:315:15]

→ **h_day = df["Avg hrs per day sleeping"].values**

hours_list_2017_Both = h_day[14:315:15]

→ **variable_observe = list(df2.columns)[:]**

matr1 = df2[variable_observe].values

matr2 = matr1 [::3]

matr_sort = matr2 [matr2[:, 2].argsort()]



Grafice

Barchart

Numărul de ore dormite pe fiecare an

```
plot.figure( figsize=(12, 10))
sns.barplot( x='Avg hrs per day sleeping', y='Year',
             data=df2, dodge=False)
plot.title("Ore vs Ani", fontsize=16)
plot.xlabel("Avg hrs per day sleeping")
plot.ylabel("Year")
plot.ylim(2003, 2017)
plot.show()
```

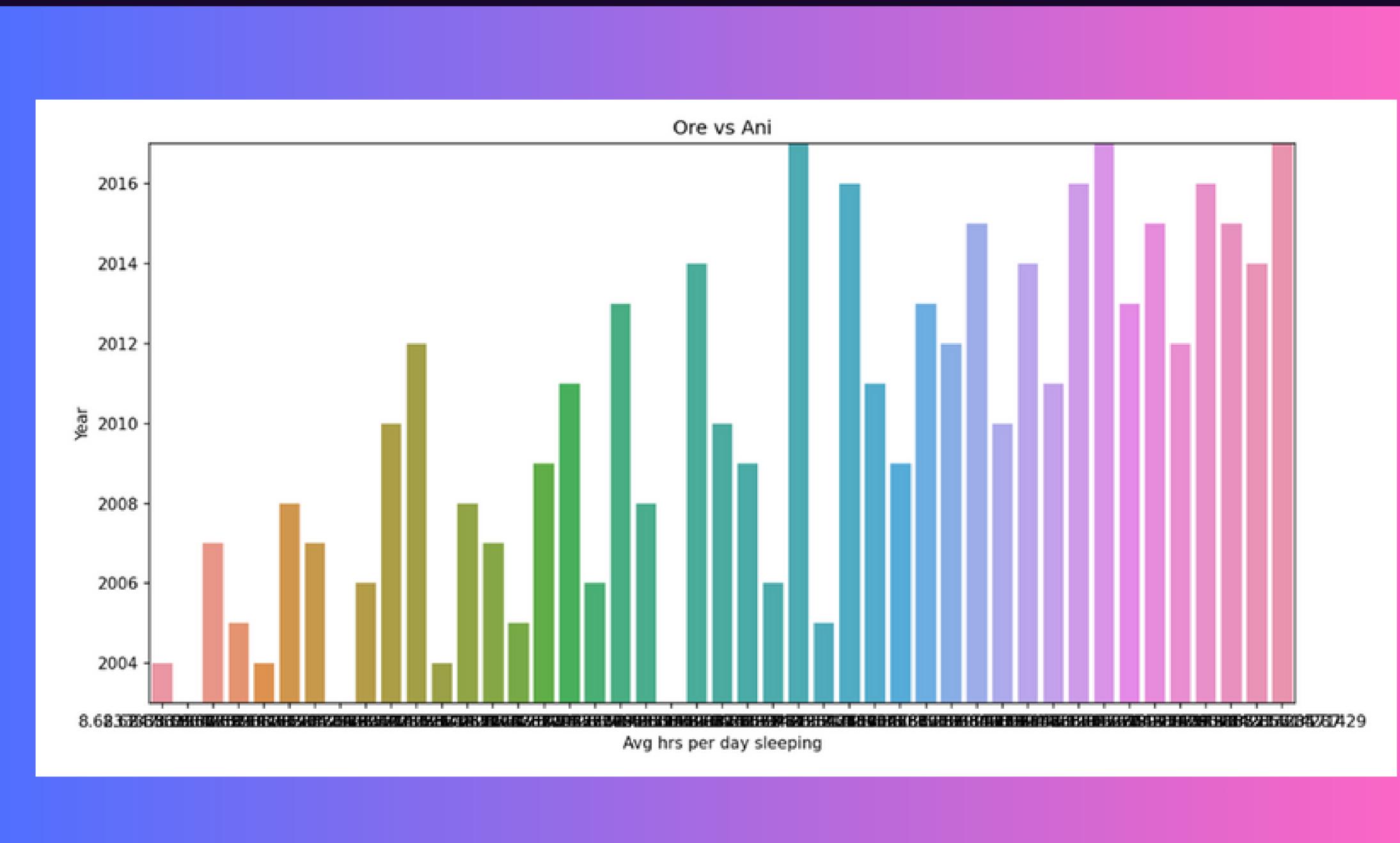
Piechart

Ponderi ore dormite pe ani si tipul zilei

```
values = grouped_data['Percentage Contribution'].values
plot.pie(values, autopct='%1.1f%%')
plot.title('Contribuția procentuală în funcție de ani si tipul zilei')
plot.show()
```

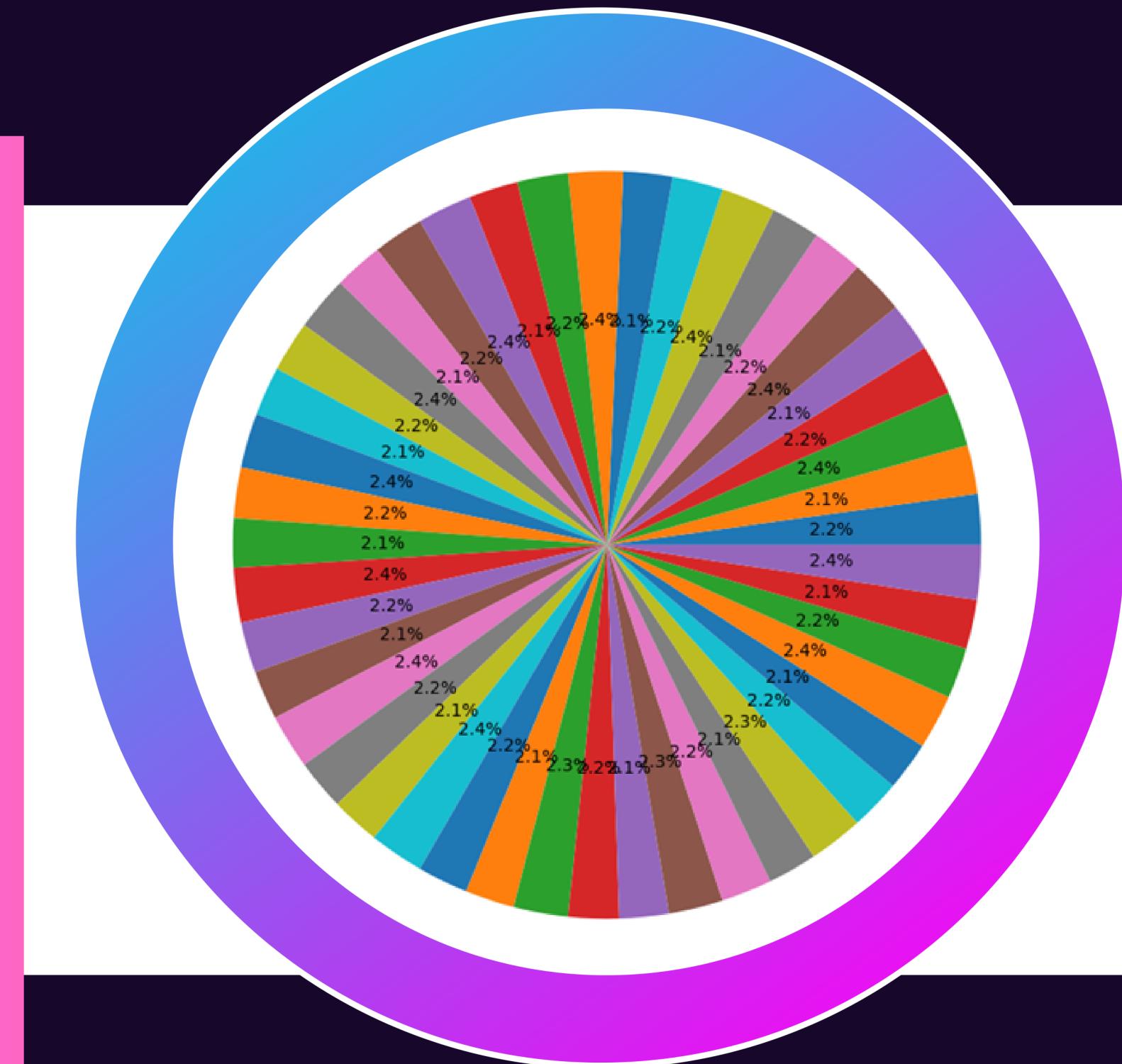


Grafice



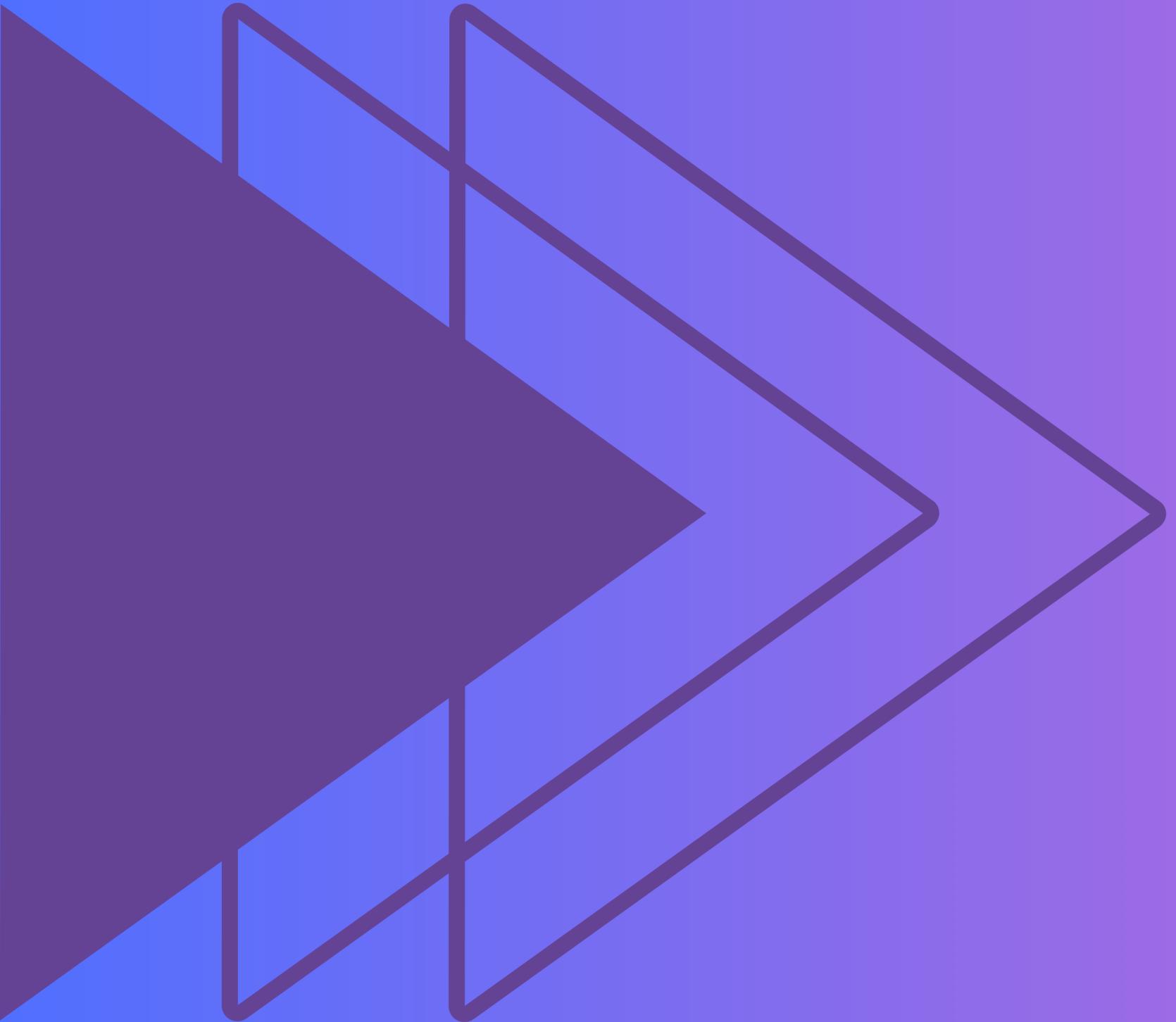
Barchart

Numărul de ore dormite pe fiecare an



Piechart

Ponderi ore dormite pe ani si tipul zilei



**MULTUMIM
PENTRU
ATENȚIE!**