

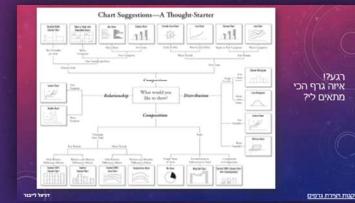
ויזואליזציה של נתונים

DATA VISUALIZATION

מה תقلס נעשה בהרצאה ?

ספריות python לרונטיזות

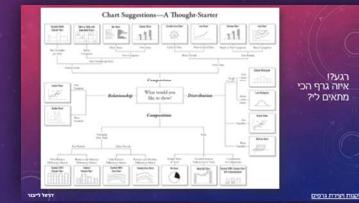
למידה ויצירת תרשימים שונים



מה תقلס נעשה בהרצאה ?

ספריות python לרונטיזות

למידה ויצירת תרשימים שונים



ספריות מרכזיות ליזואליזציה בשפת PYTHON



MATPLOTLIB

- בעicker לתרשימים בסיסים (אפשרי לעשות בה הכל אבל...)
- אין סוף אפשרויות לשחק עם התרשימים

[Examples — Matplotlib 3.7.2 documentation](#)

יתרונות :

- קלה מאוד לשימוש
- אפשר להתאים כל דבר שרוצים.

חסרונות :

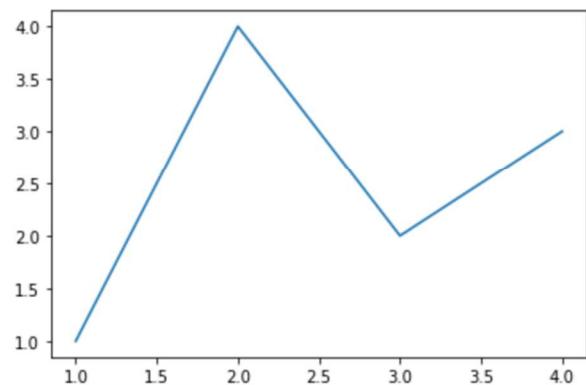
- או שהפתרונות הסופית לא מוצלחת או שכתבת הרבה קוד בשבייה (ביחס לכלים אחרים)

MATPLOTLIB

```
In [1]: import matplotlib as mpl
import matplotlib.pyplot as plt
import numpy as np

In [2]: fig, ax = plt.subplots() # Create a figure containing a single axes.
        ax.plot([1, 2, 3, 4], [1, 4, 2, 3]) # Plot some data on the axes.

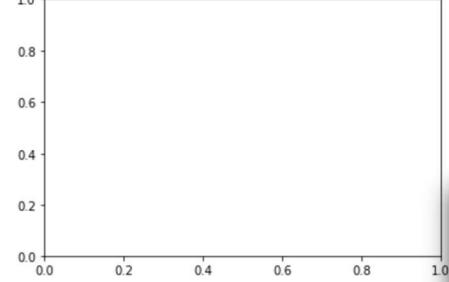
Out[2]: [<matplotlib.lines.Line2D at 0x27c40576c10>]
```



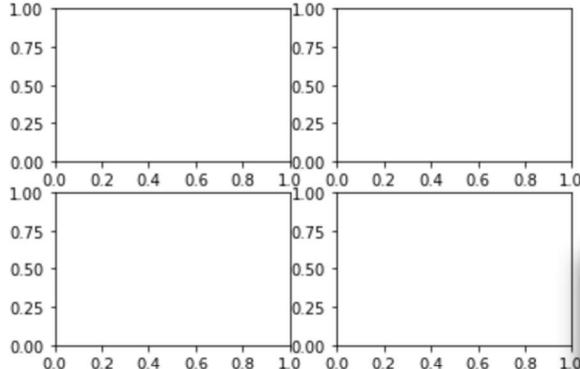
MATPLOTLIB

```
In [3]: fig = plt.figure() # an empty figure with no Axes  
fig, ax = plt.subplots() # a figure with a single Axes
```

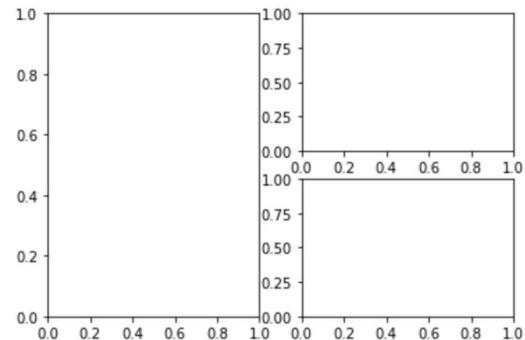
<Figure size 432x288 with 0 Axes>



```
In [4]: fig, axs = plt.subplots(2, 2) # a figure with a 2x2 grid of Axes
```



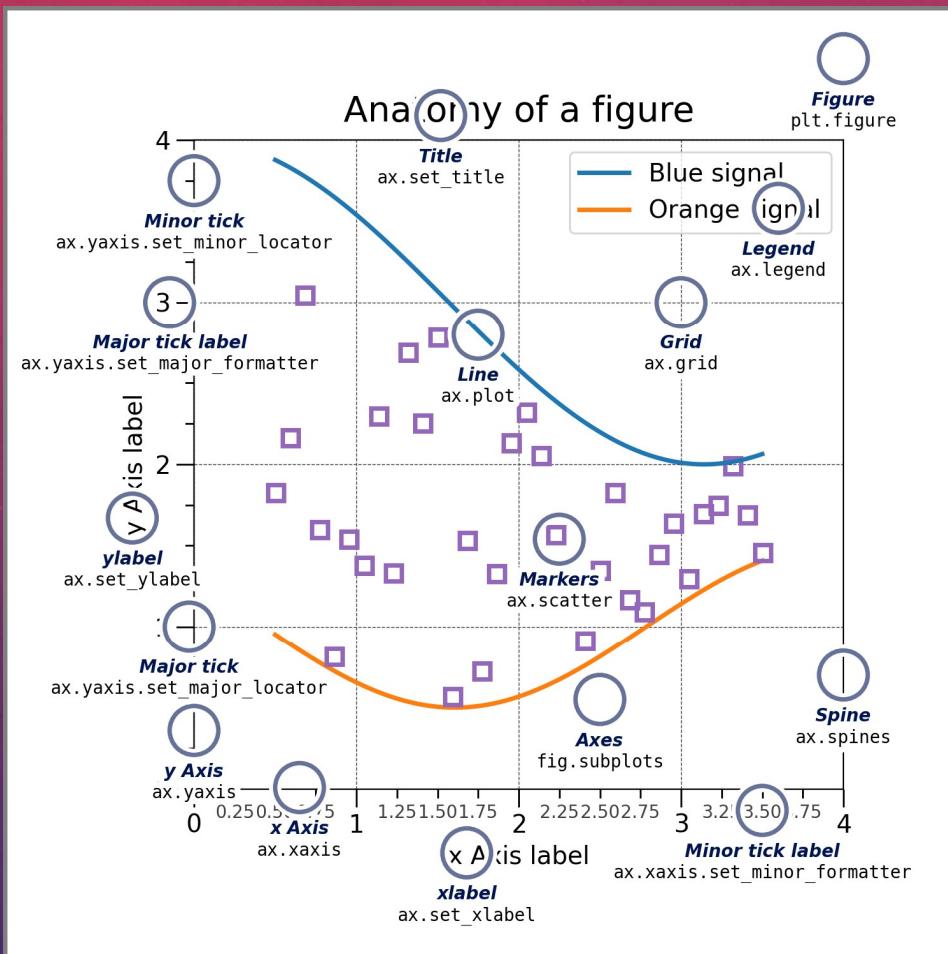
```
In [5]: # a figure with one axes on the left, and two on the right:  
fig, axs = plt.subplot_mosaic([['left', 'right_top'],  
                             ['left', 'right_bottom']])
```



MATPLOTLIB

Figure: הגורם החיצוני ביותר עבור גרף. Matplotlib יכול להכיל אובייקטים רבים של צירים.

Axes: גרפ (מערכת צירים) ספציפית (מערכת צירים עם התצוגה שעלייה)



SEABORN

- מבוסס `matplotlib`
- אופטימיזציה לטובת גרפים "יפים" בצלבים ובתוכן.
[Example gallery — seaborn 0.12.2 documentation \(pydata.org\)](#)
-

יתרונות :

- קלה מאוד לשימוש
- ה - High level של `Matplotlib`
- מותאם לשימוש על `Datasets`

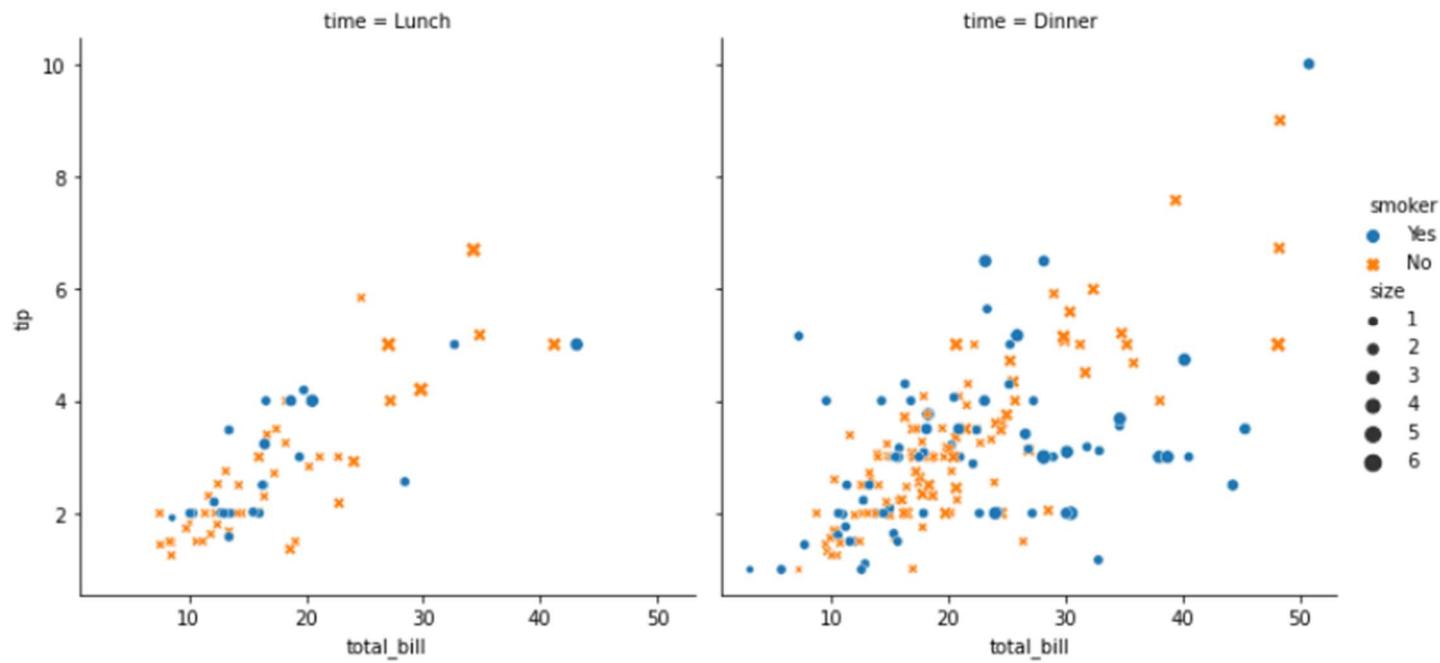
חסרונות :

- בסוף יש דברים שבשביל להתאים אותם צריך לעשות ב `matplotlib`

```
In [6]: import seaborn as sns  
tips = sns.load_dataset("tips")
```

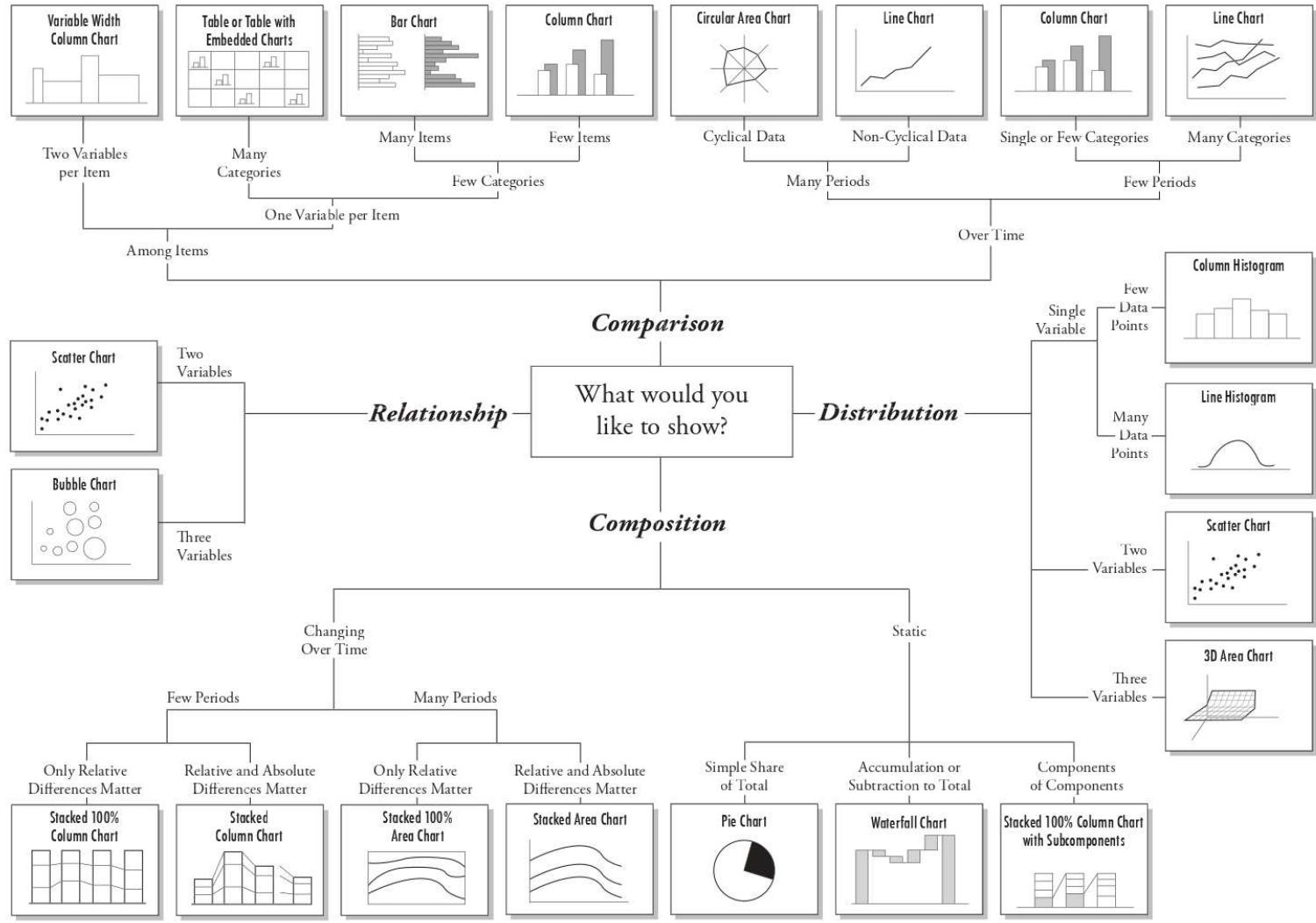
```
In [7]: # Create a visualization  
sns.relplot(  
    data=tips,  
    x="total_bill", y="tip", col="time",  
    hue="smoker", style="smoker", size="size",  
)
```

```
Out[7]: <seaborn.axisgrid.FacetGrid at 0x2517347b670>
```



SEABORN

Chart Suggestions—A Thought-Starter



רגע!
איזה גרף הכי
מתאים לי?

IRIS DATASET

Iris Setosa 1 0 0



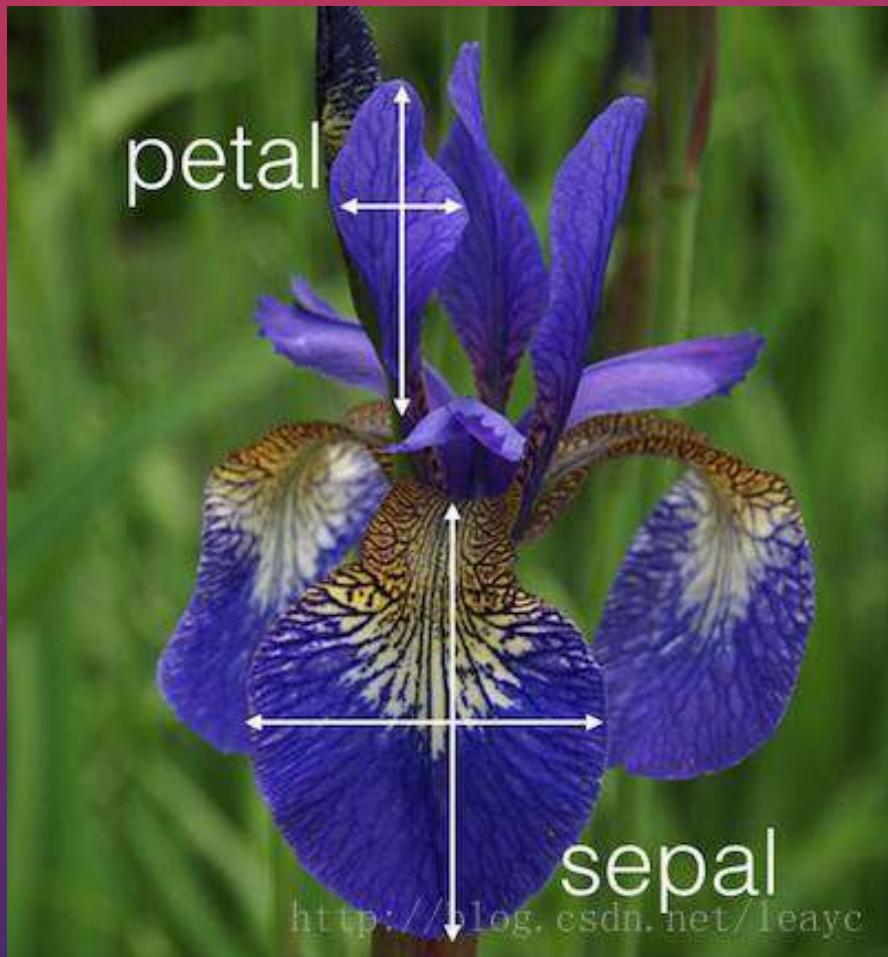
Iris Versicolor 0 1 0



Iris Virginica 0 0 1



IRIS DATASET

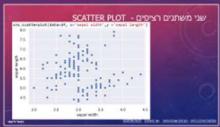


מסקנות ויצירת גרפים

הגרפים שננחת וניצור

שני משתנים

שני משתנים רציפים



משתנה רציף ומשתנה קטגוריאי



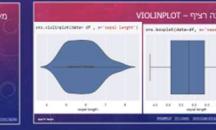
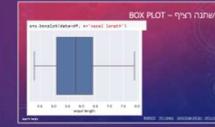
מתקנות ויצירת גרפים - בגרפים שננתח ומייצור - שני משתנים

דניאל ליבנר

משתנה יחיד

משתנה יחיד רציף

משתנה יחיד קטגוריאי



מתקנות ויצירת גרפים - בגרפים שננתח ומייצור - משתנה יחיד

דניאל ליבנר

שלושה משתנים

שני משתנים רציפים
וממשנה אחד קטגוריאי



כל המשתנים



מתקנות ויצירת גרפים - בגרפים שננתח ומייצור - שלושה משתנים

דניאל ליבנר

מסקנות ויצירת גרפים - הגרפים שננחת וניצור

משתנה יחיד

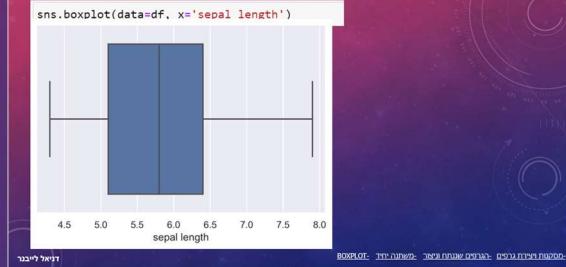
משתנה יחיד קטgoriy

COUNT PLOT – משתנה קטגורוי

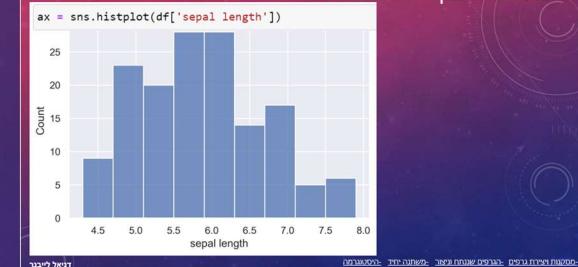


משתנה יחיד רציף

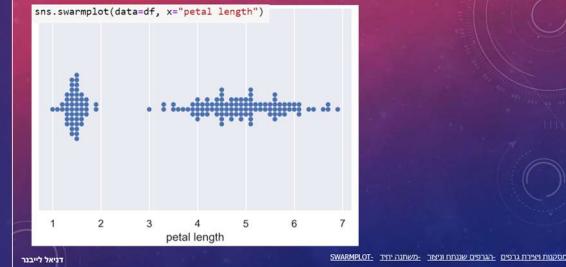
BOX PLOT – משתנה רציף



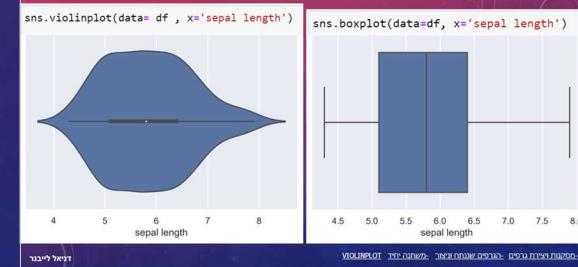
histplot – המשתנה רציף - היסטוגרמה



SWARM PLOT – משתנה רציף



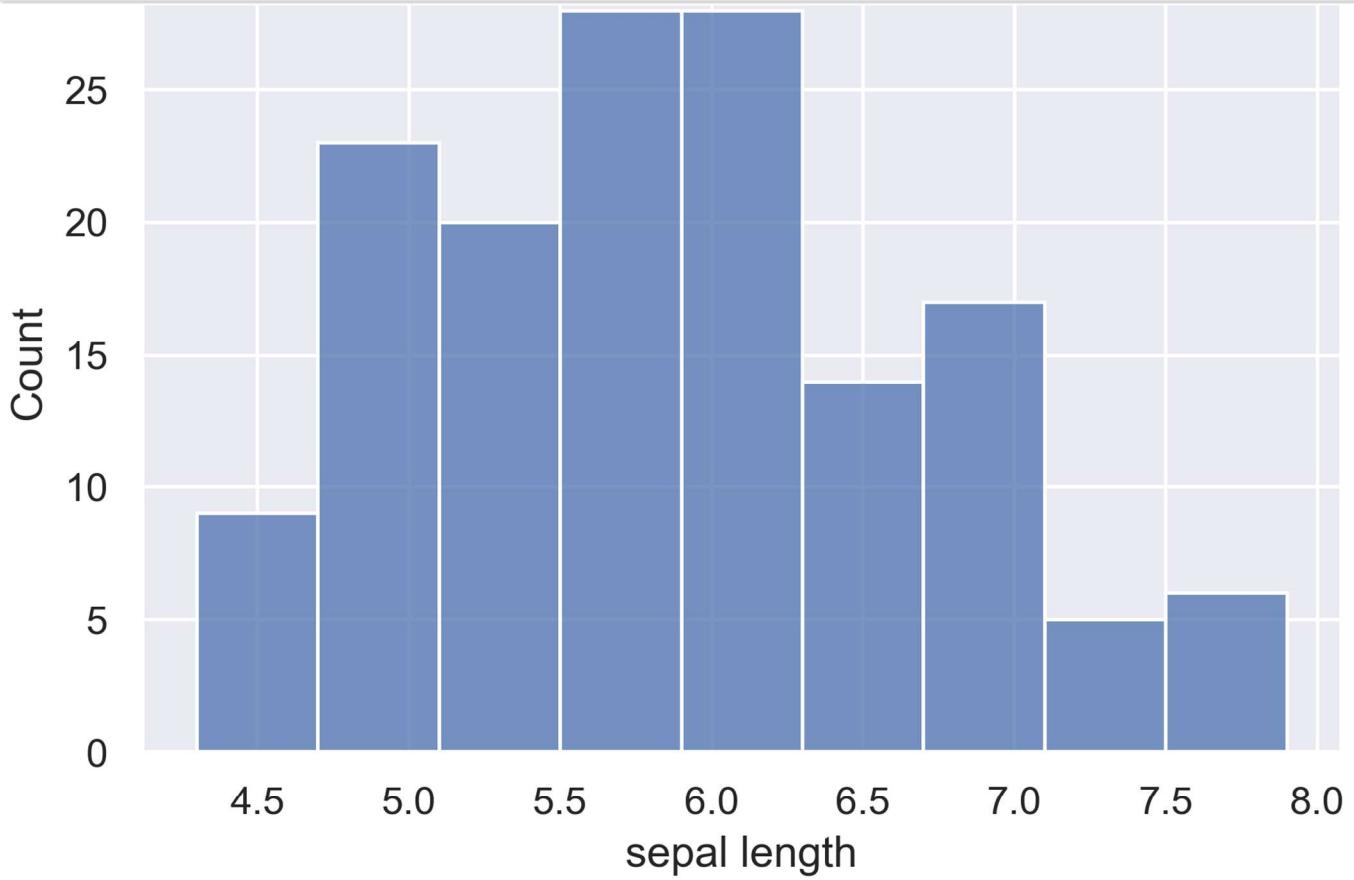
VIOLIN PLOT – המשתנה רציף - ווילינט פלט



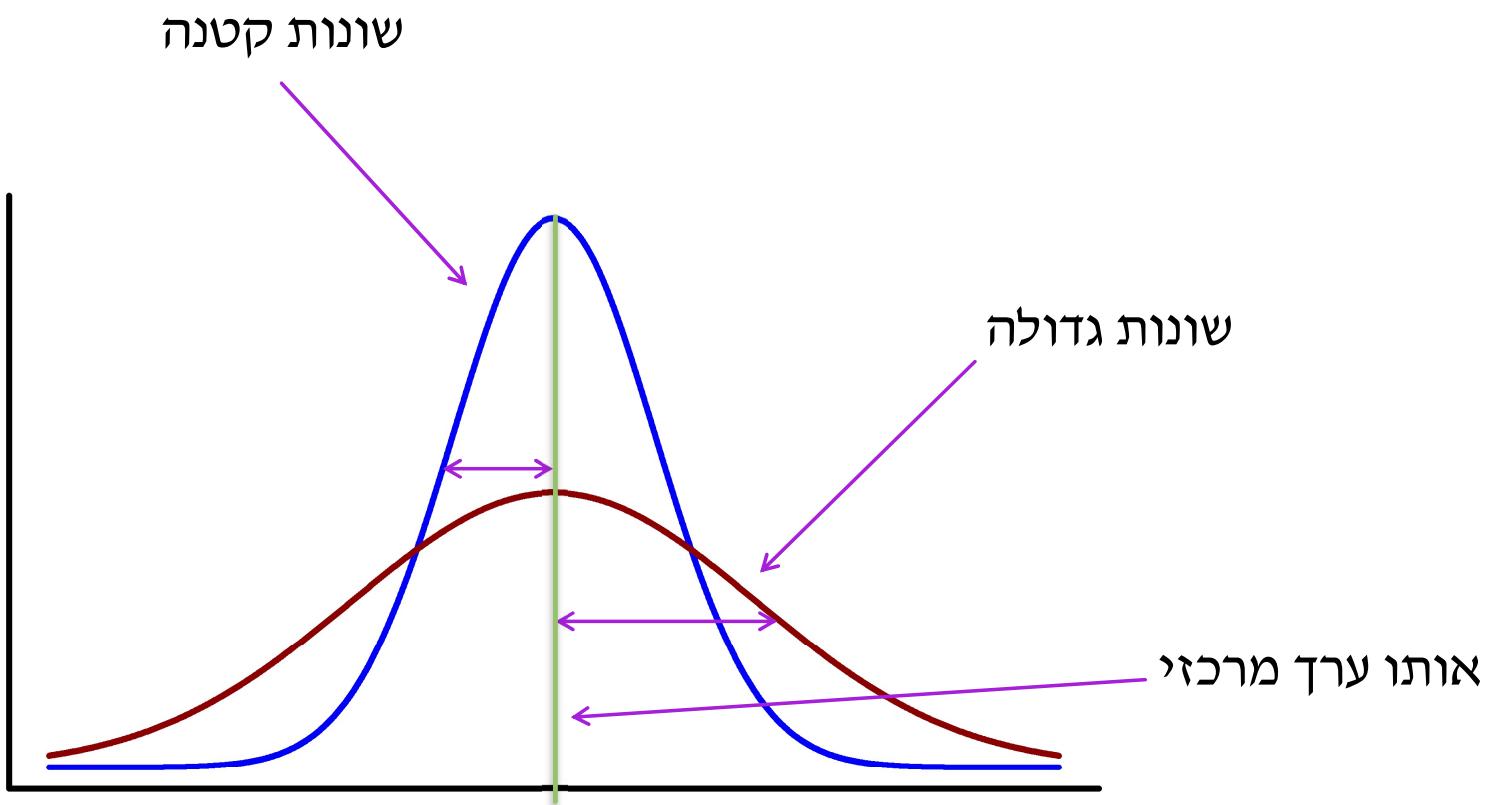
-מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ -משתנה יחיד

משתנה רציף - היסטוגרما

```
ax = sns.histplot(df['sepal length'])
```

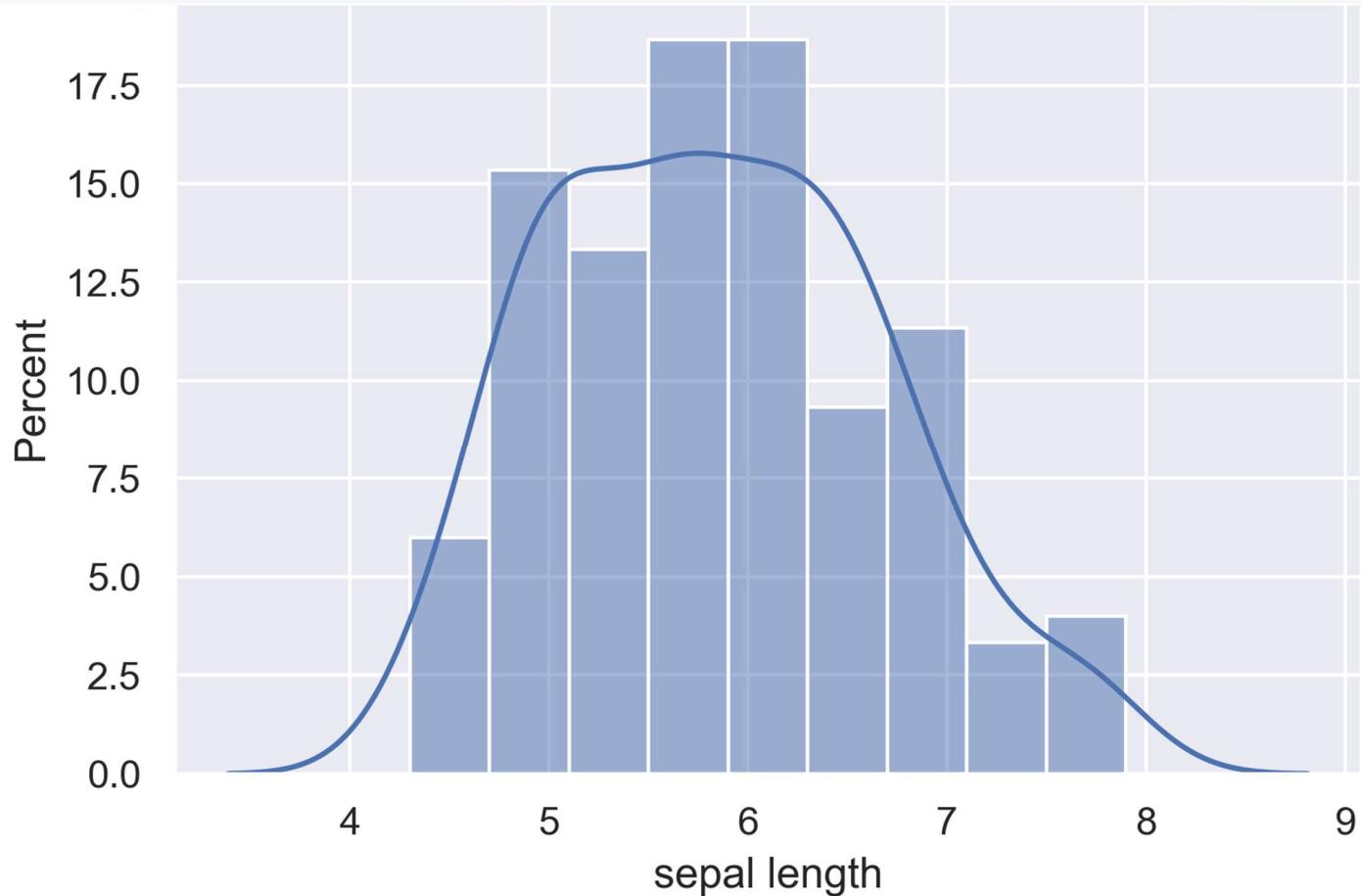


משתנה רציף - היסטוגרמה



משתנה רציף - היסטוגרמה

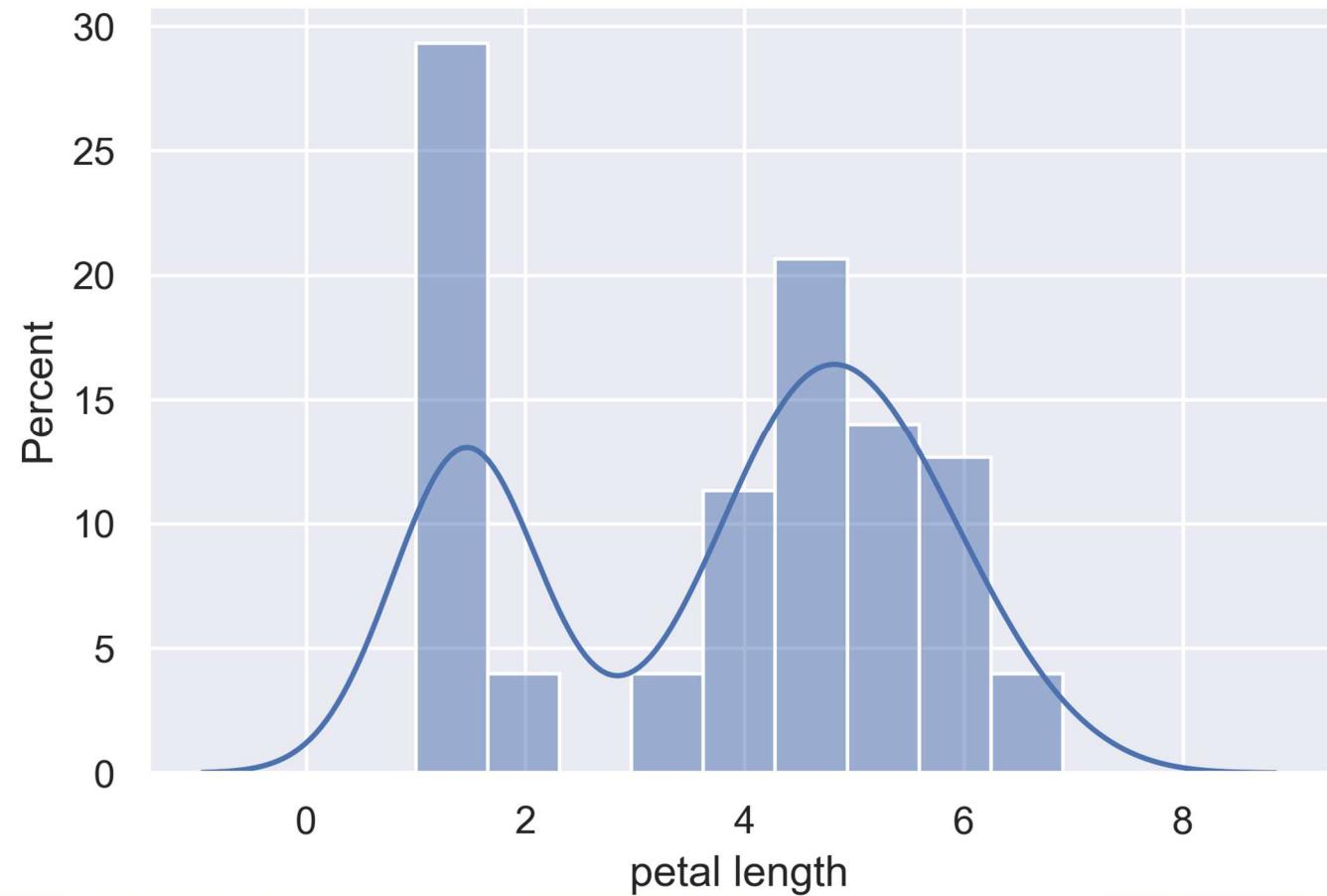
```
ax = sns.histplot(df['sepal length'], kde=True, stat='percent', kde_kws=dict(cut=3) )
```



-מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ -משתנה יחיד -ההיסטוגרמה

משתנה רציף - היסטוגרמה

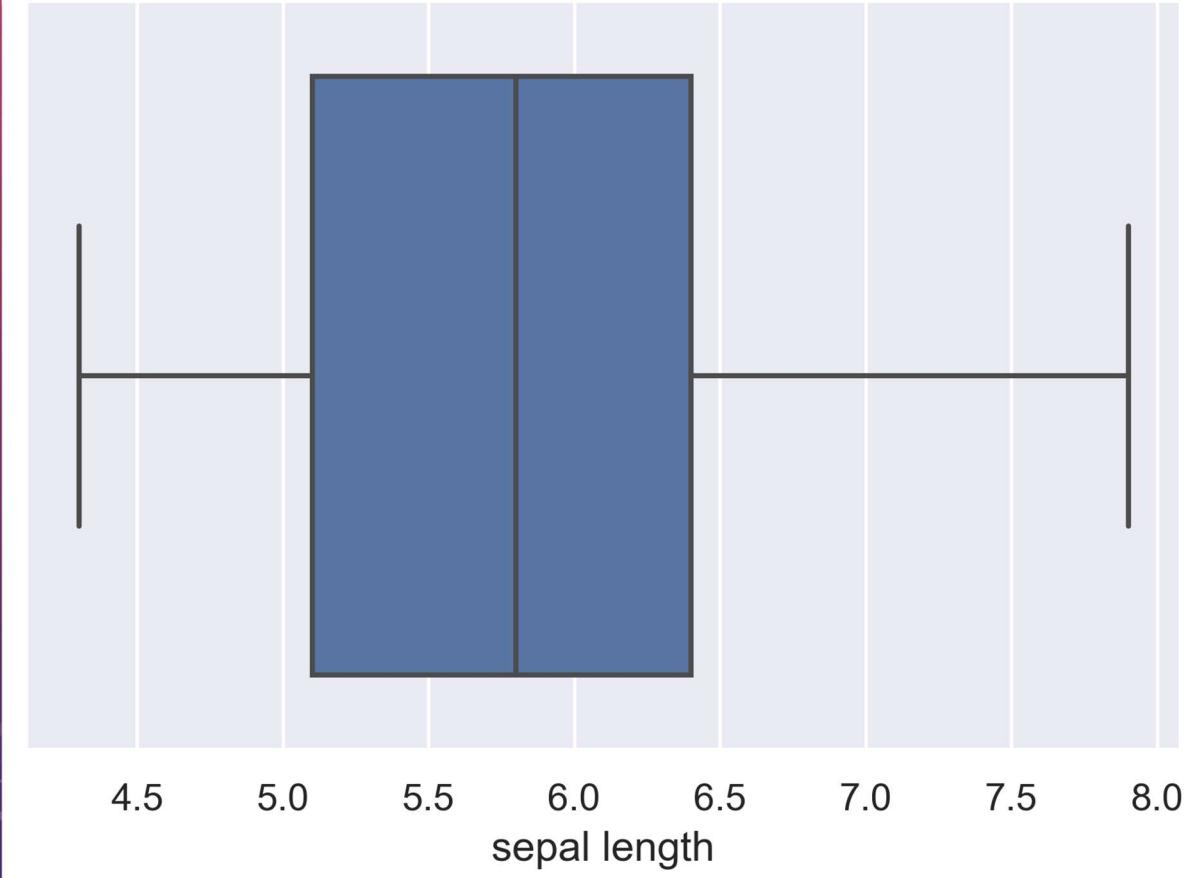
```
ax = sns.histplot(data=df, x ='petal length', kde=True, stat='percent', kde_kws=dict(cut=3) )
```



-מסקנות ויצירת גרפים -הגרפים שננתח וניצו -משתנה יחיד -ההיסטוגרמה

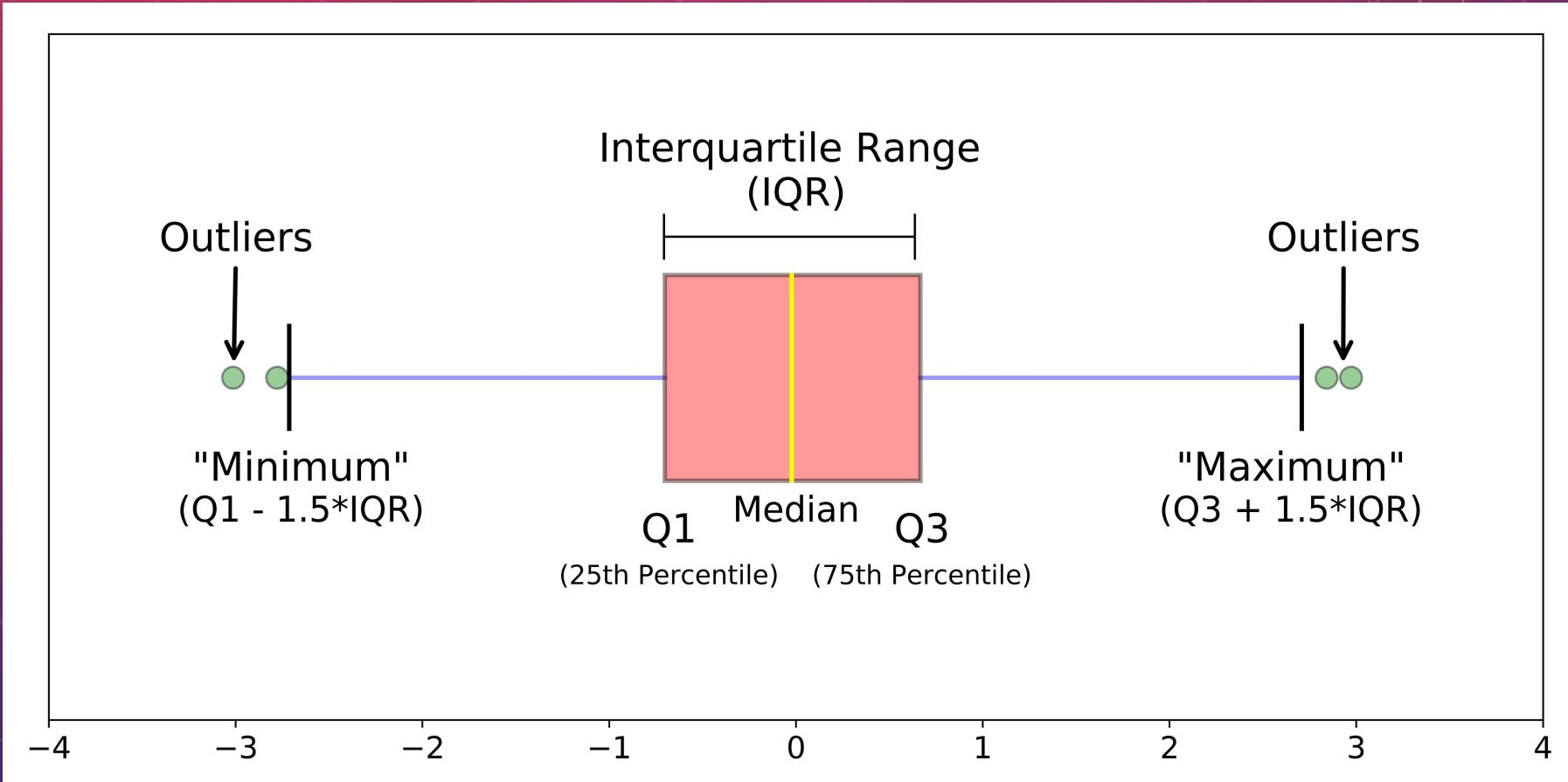
משתנה רציף - BOX PLOT

```
sns.boxplot(data=df, x='sepal length')
```

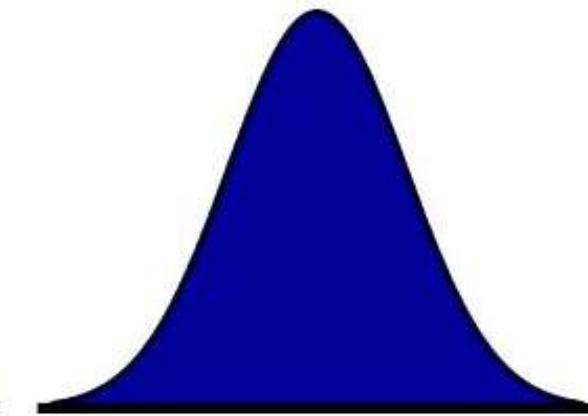
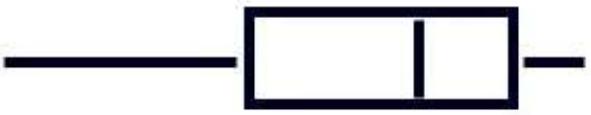
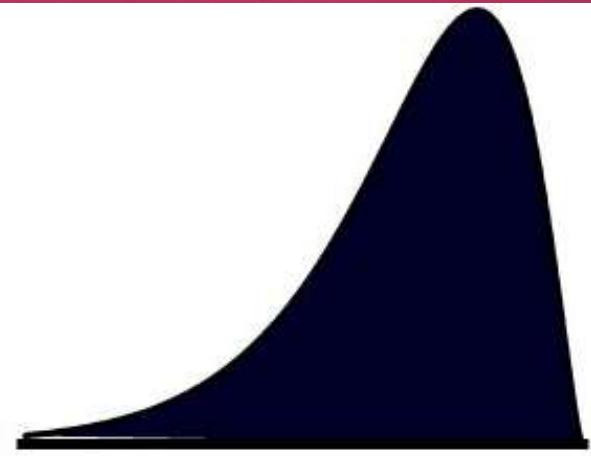


-מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ -משתנה יחיד -BOXPLOT-

משתנה רציף - BOX PLOT



-מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ -משתנה יחיד BOXPLOT-



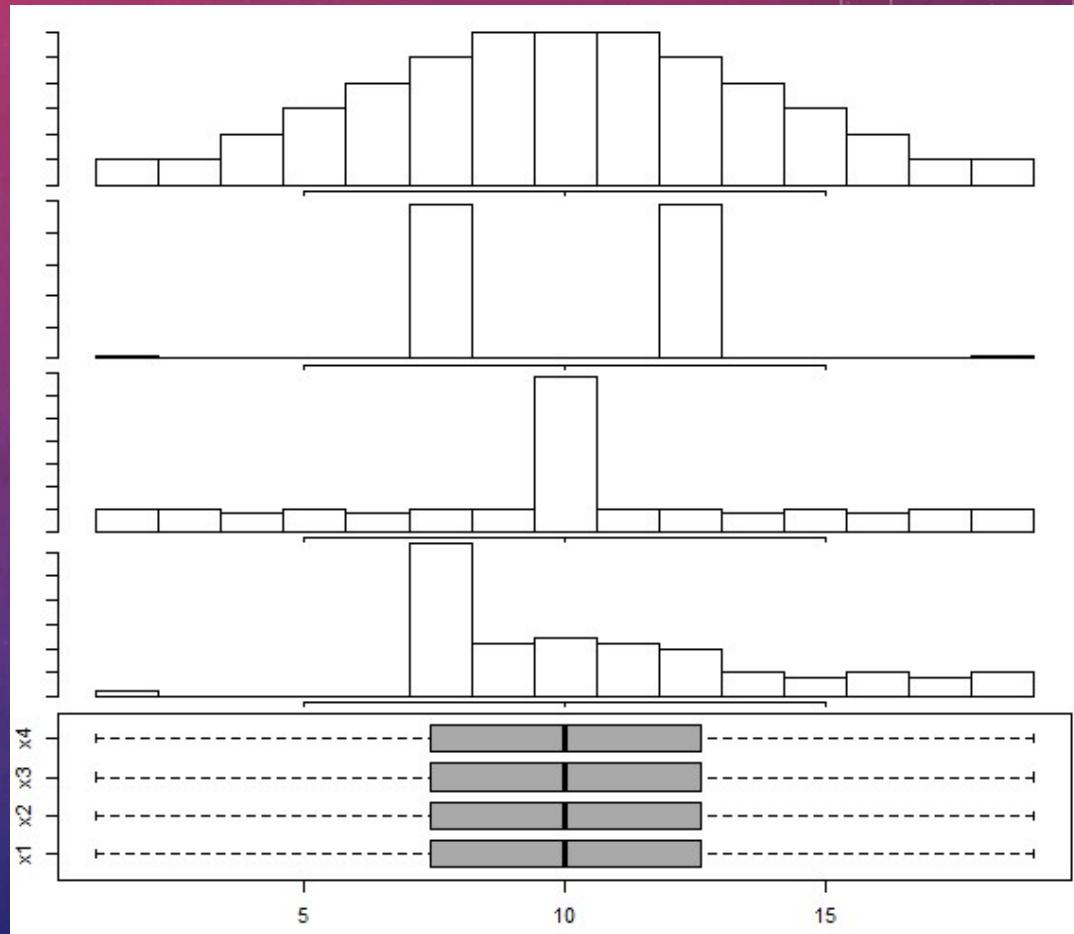
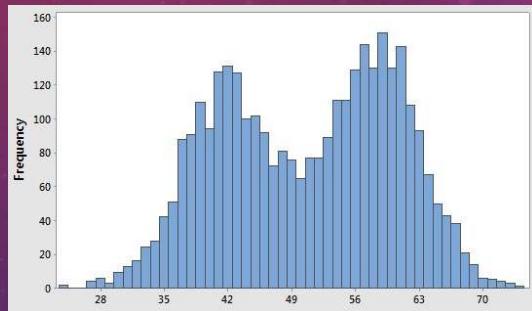
**Negatively
Skewed**

**Symmetric
(Not Skewed)**

**Positively
Skewed**

از מה הביעיות של BOX PLOT ?

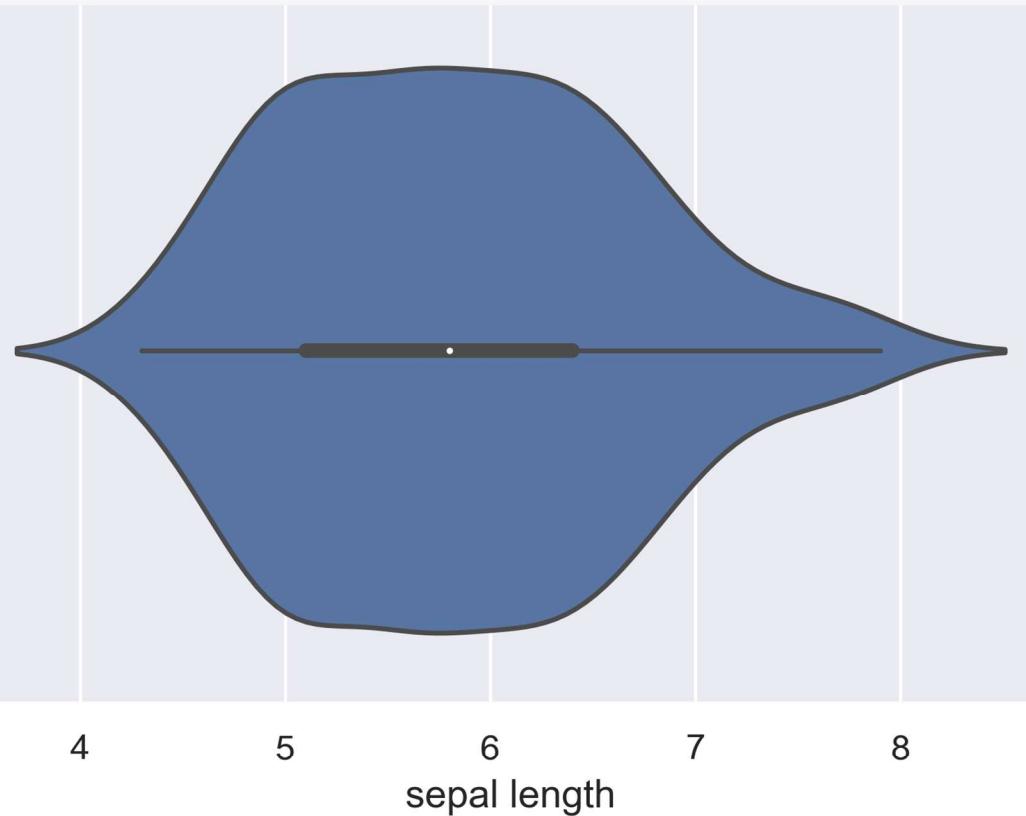
Boxplot של התפלגות נורמלית
ובו-מודלית יכולות להיות דומות!



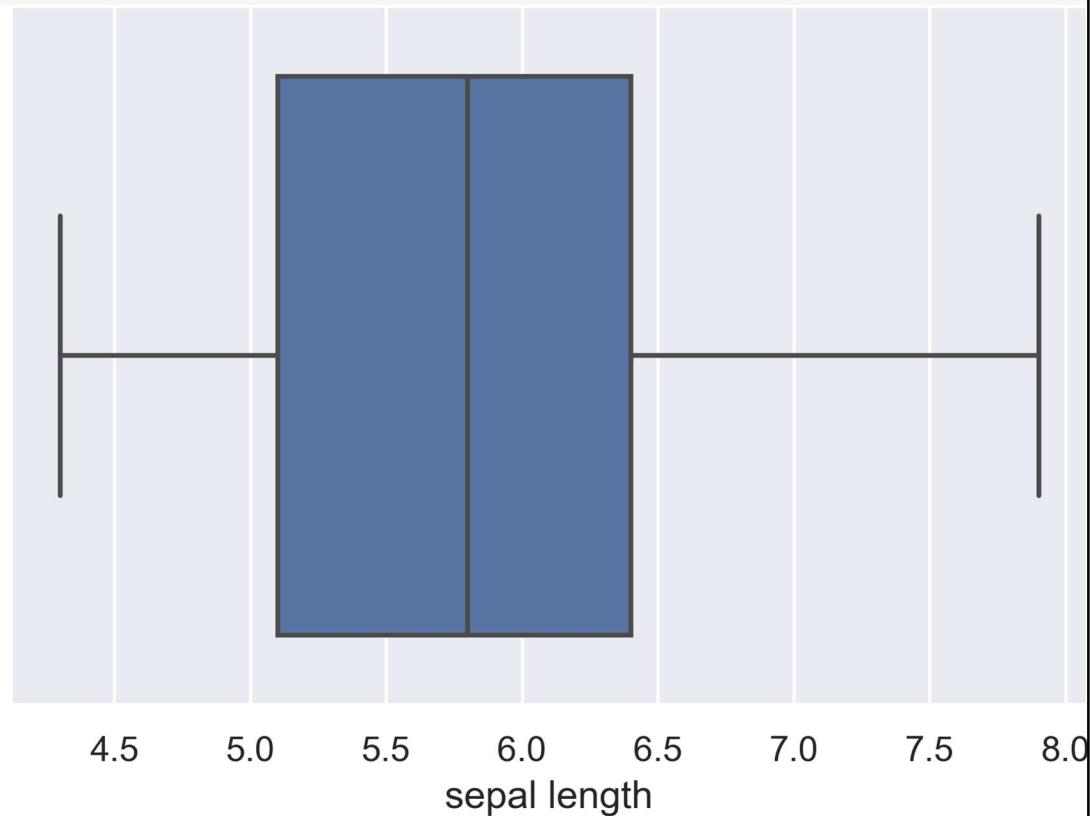
-מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ -משתנה יחיד -BOXPLOT

משתנה רציף – VIOLINPLOT

```
sns.violinplot(data= df , x='sepal length')
```

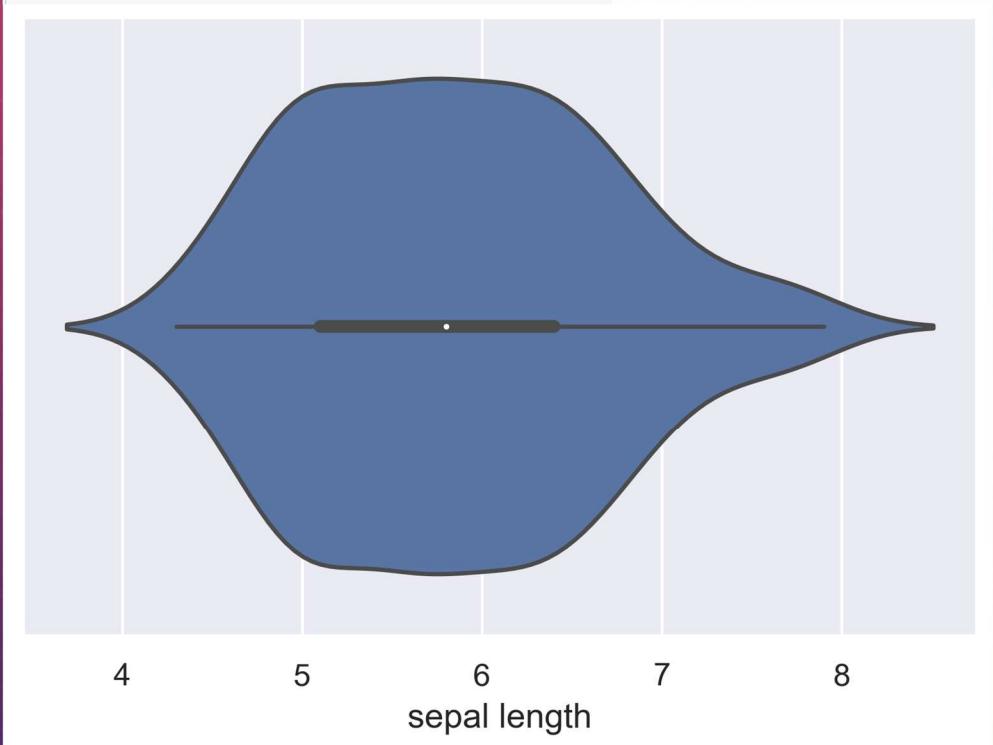


```
sns.boxplot(data=df, x='sepal length')
```

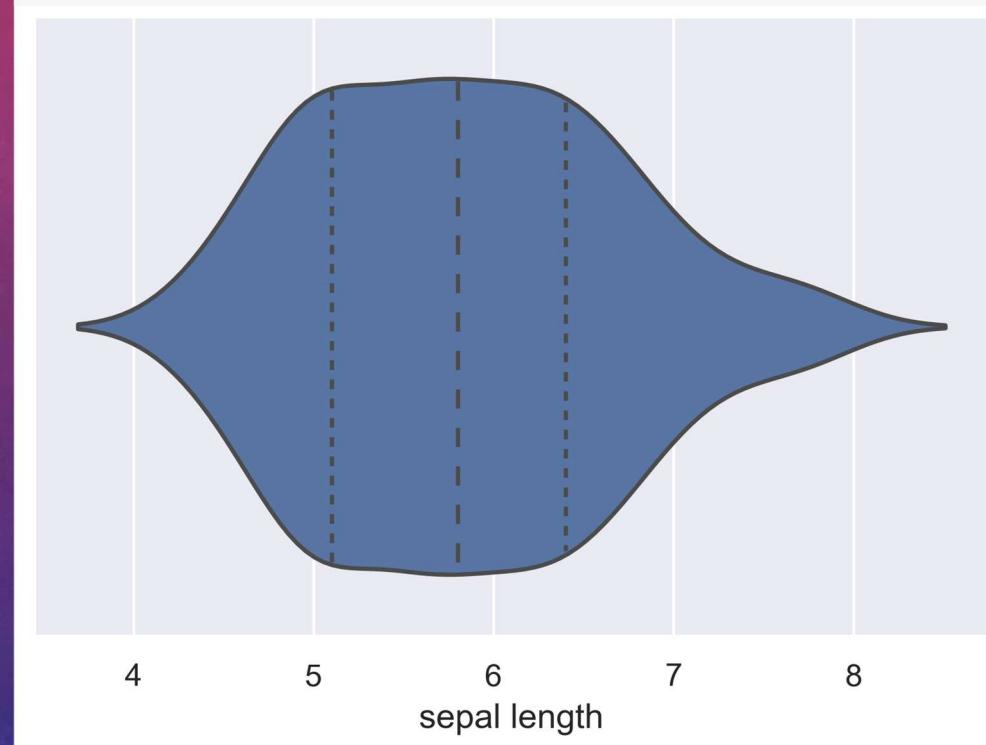


משתנה רציף – VIOLINPLOT

```
sns.violinplot(data= df , x='sepal length')
```

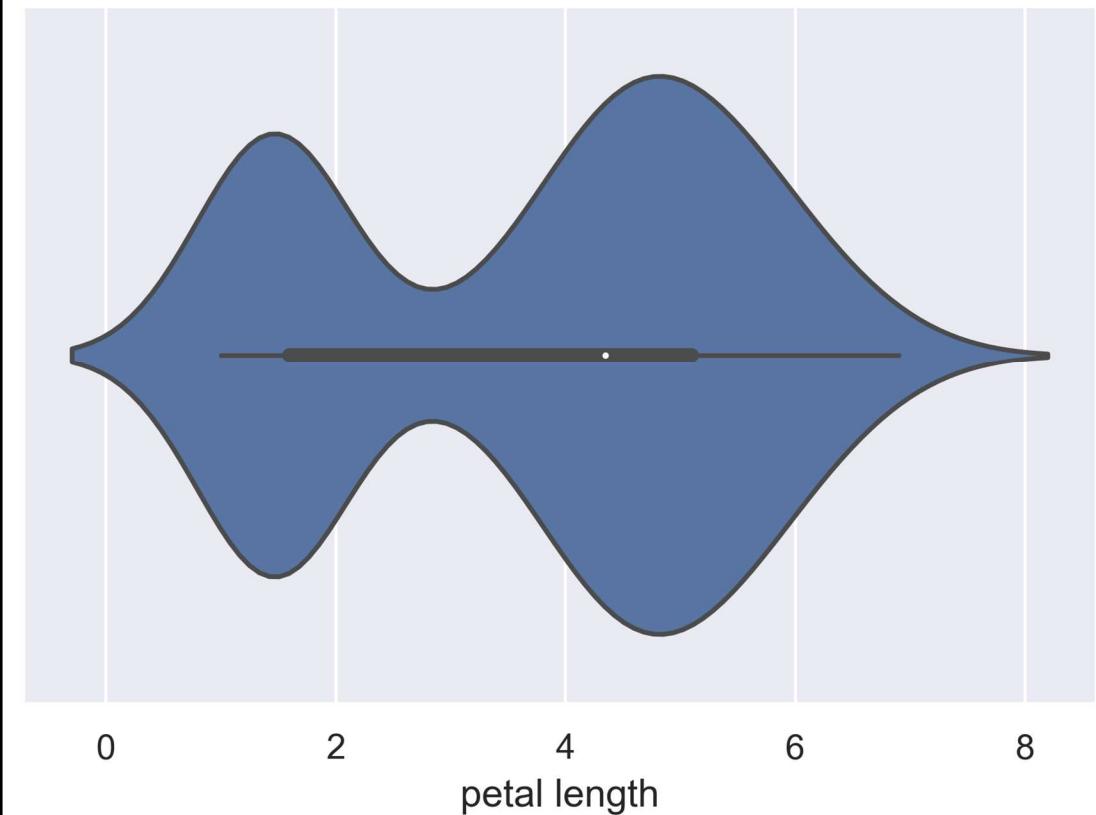


```
sns.violinplot(data= df , x='sepal length',inner="quart")
```

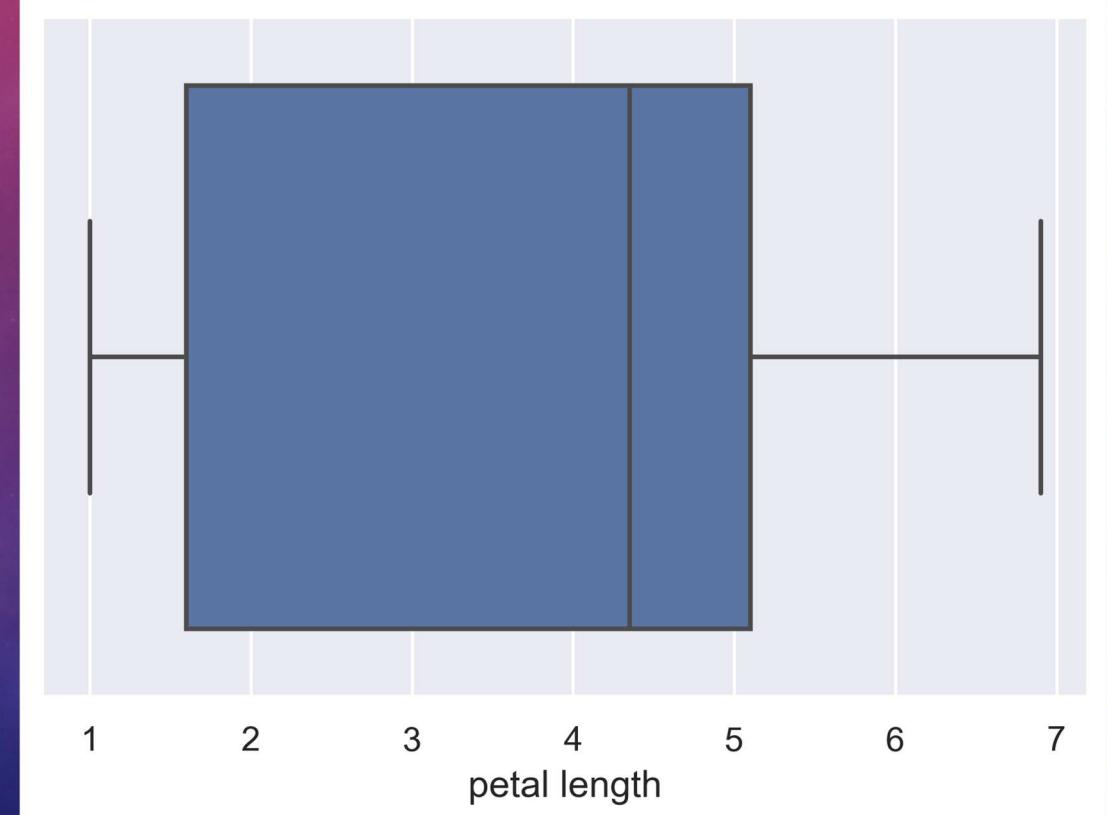


משתנה רציף – VIOLINPLOT

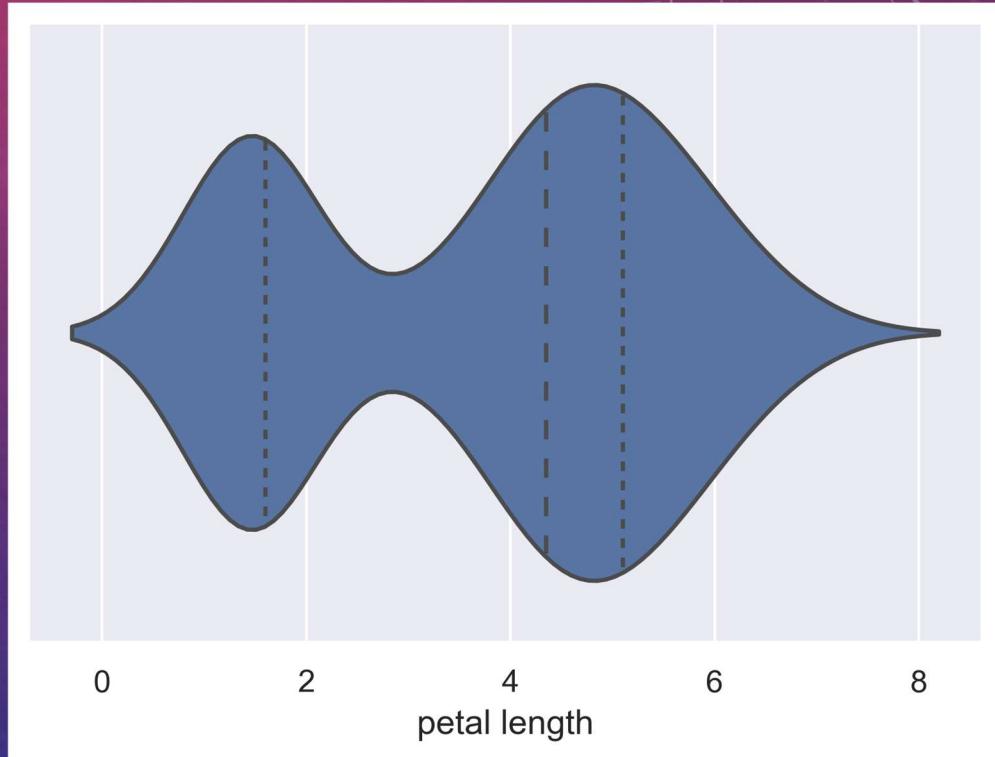
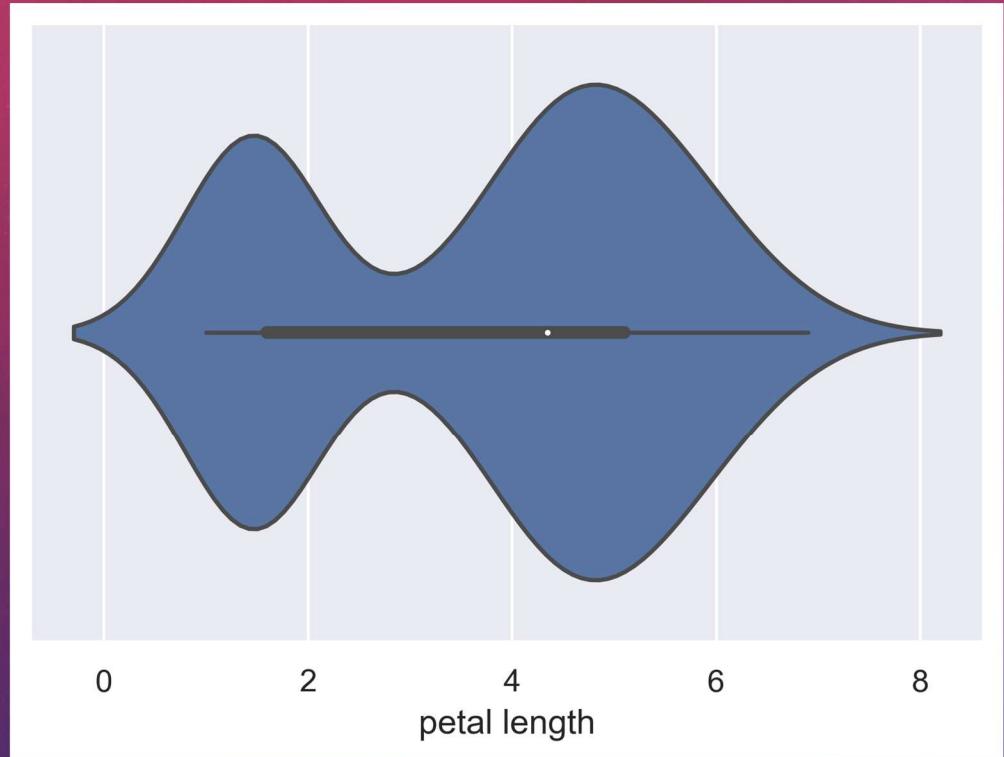
```
sns.violinplot(data= df ,x ='petal length')
```



```
sns.boxplot(data=df, x='petal length')
```



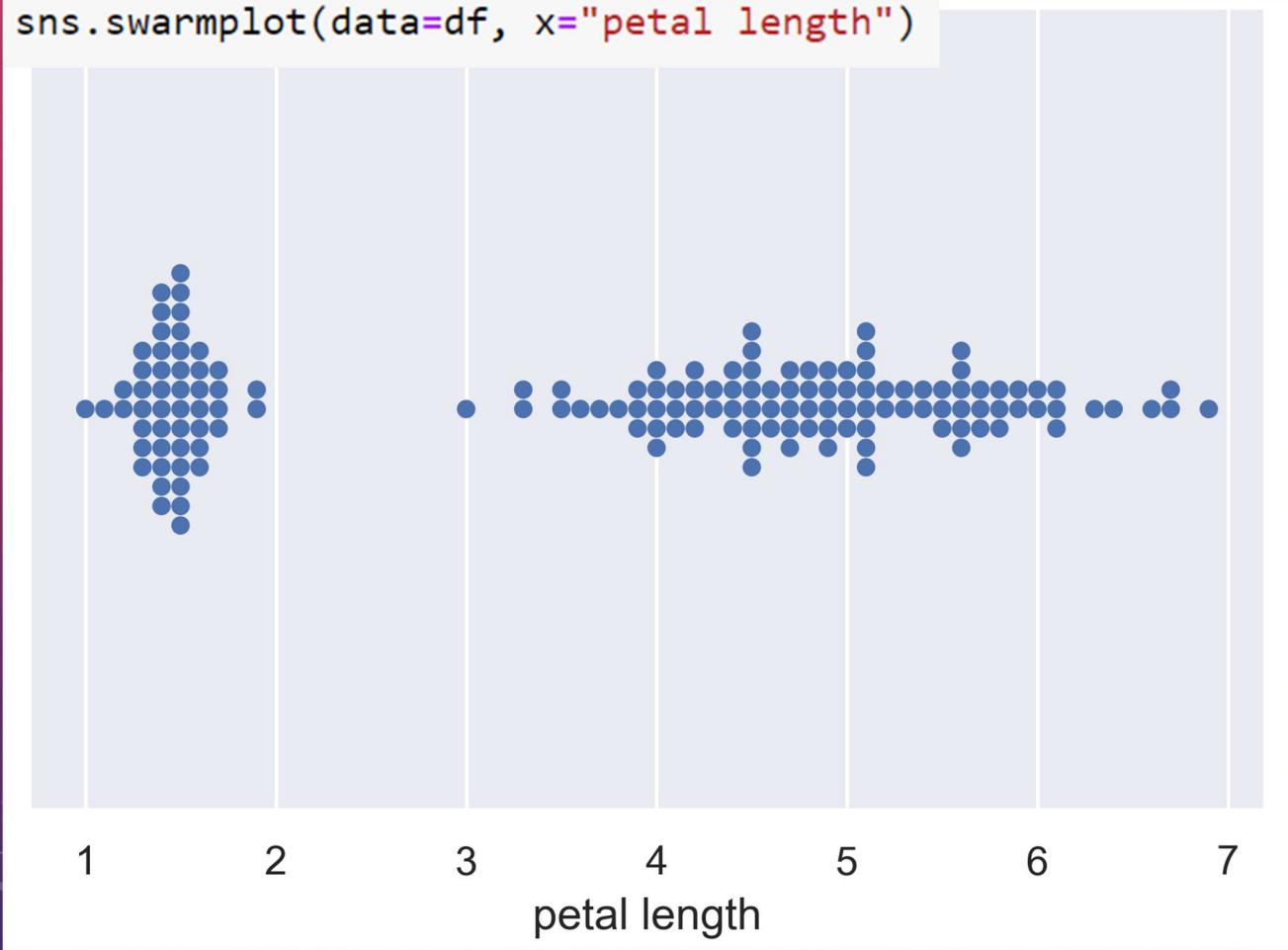
משתנה רציף – VIOLINPLOT



מסקנות ויצירת גרפים - הגרפים שננתחו וניצוור - משתנה יחיד VIOLINPLOT

משתנה רציף - SWARM PLOT

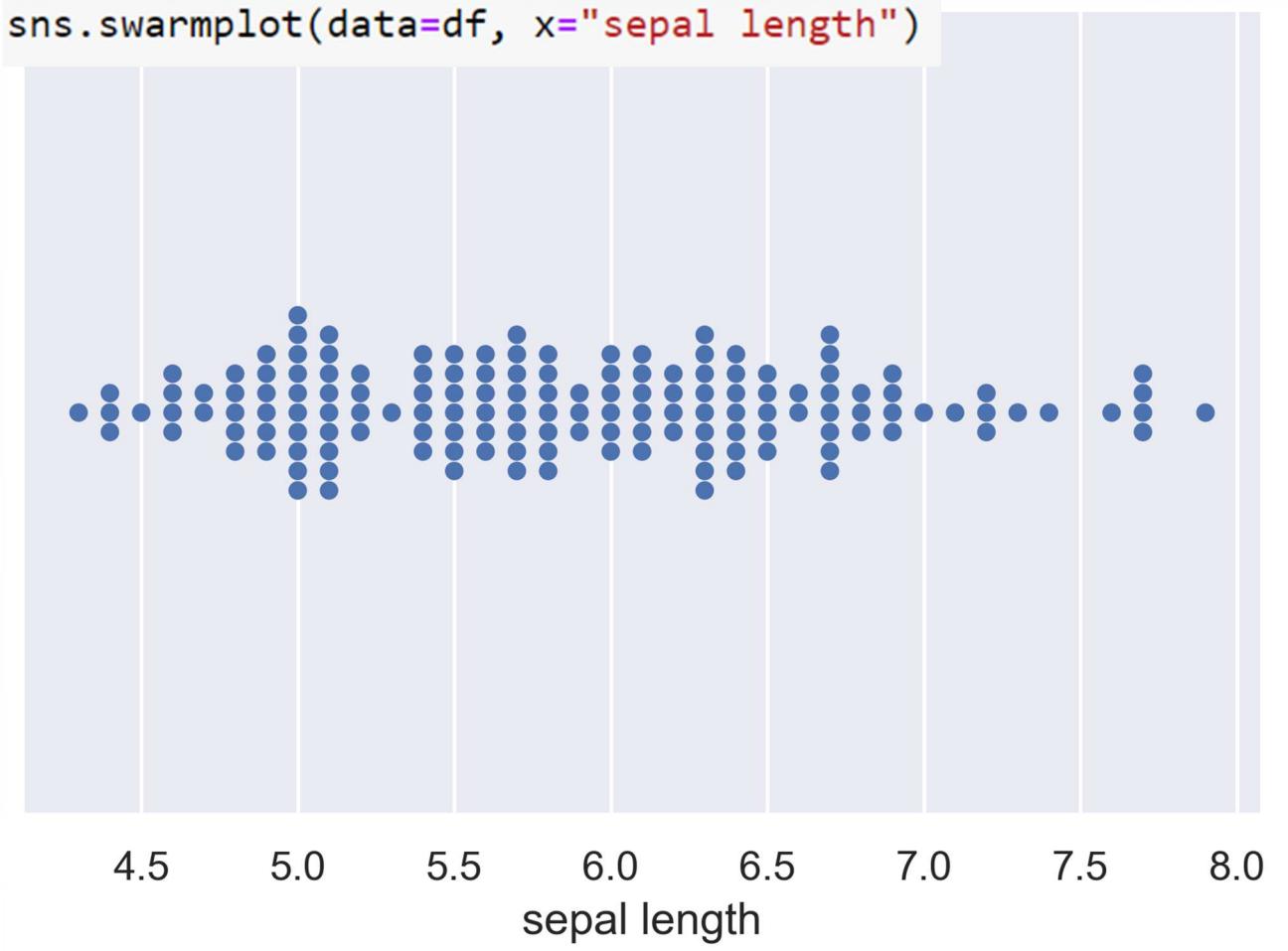
```
sns.swarmplot(data=df, x="petal length")
```



-מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ -משתנה יחיד -SWARM PLOT-

משתנה רציף – SWARM PLOT

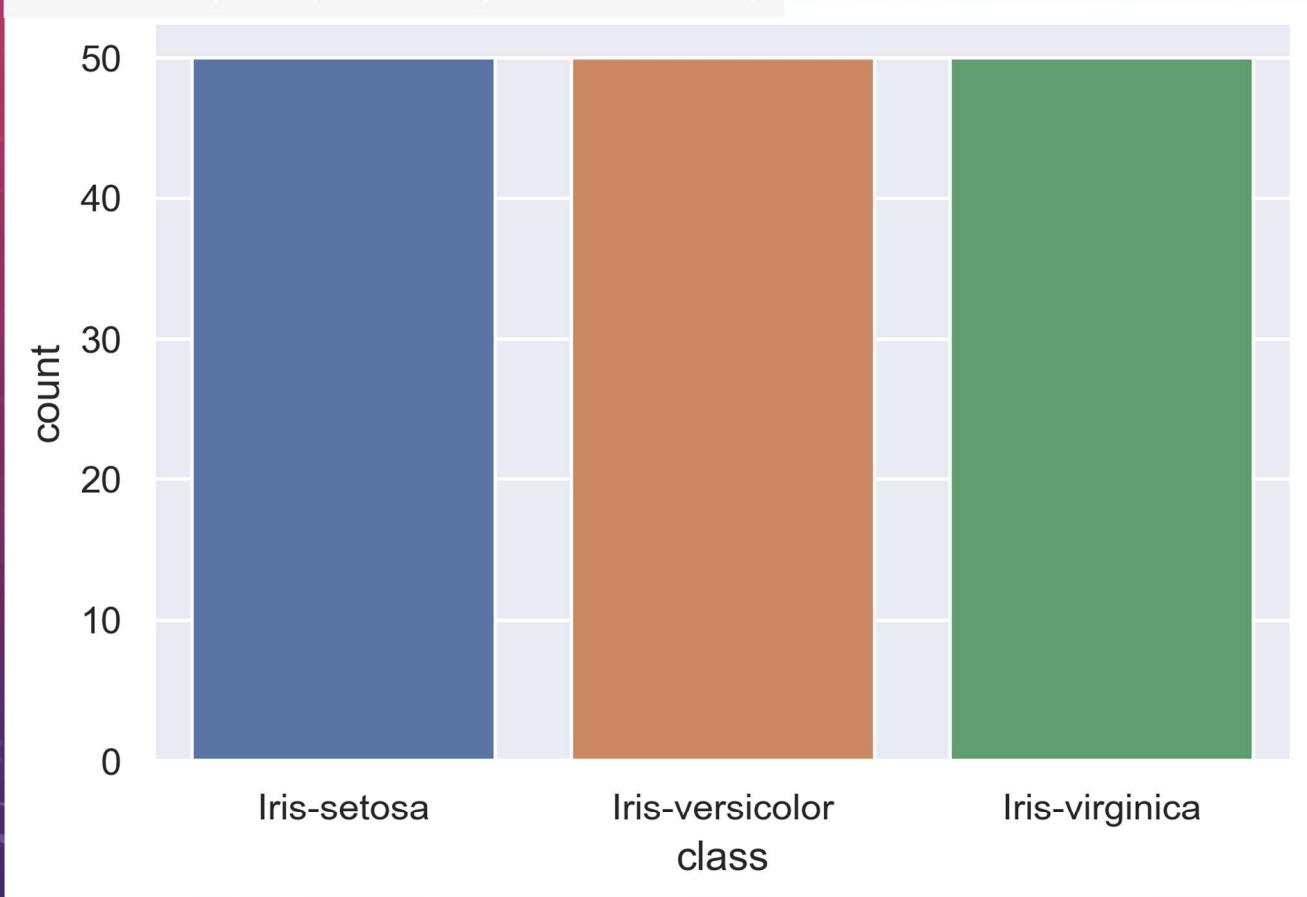
```
sns.swarmplot(data=df, x="sepal length")
```



-מסקנות ויצירת גרפים -הגרפים שננתח וניצו -משתנה יחיד -SWARM PLOT-

משתנה קטgoriy - COUNT PLOT

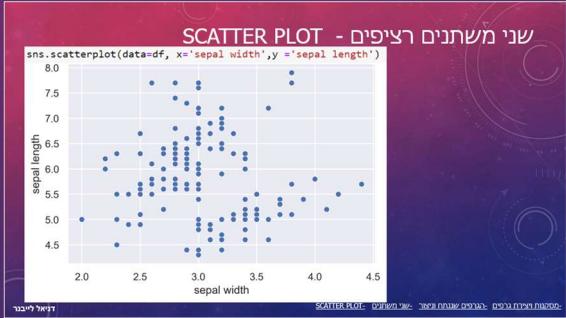
```
sns.countplot(data=df, x='class' )
```



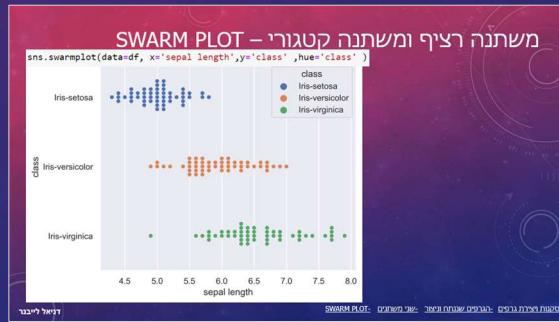
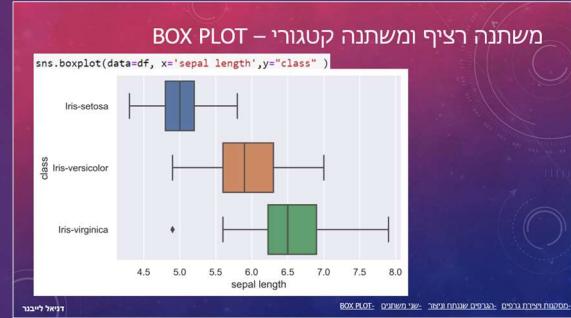
-מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ -משתנה יחיד COUNT PLOT-

שני משתנים

שני משתנים רציפים



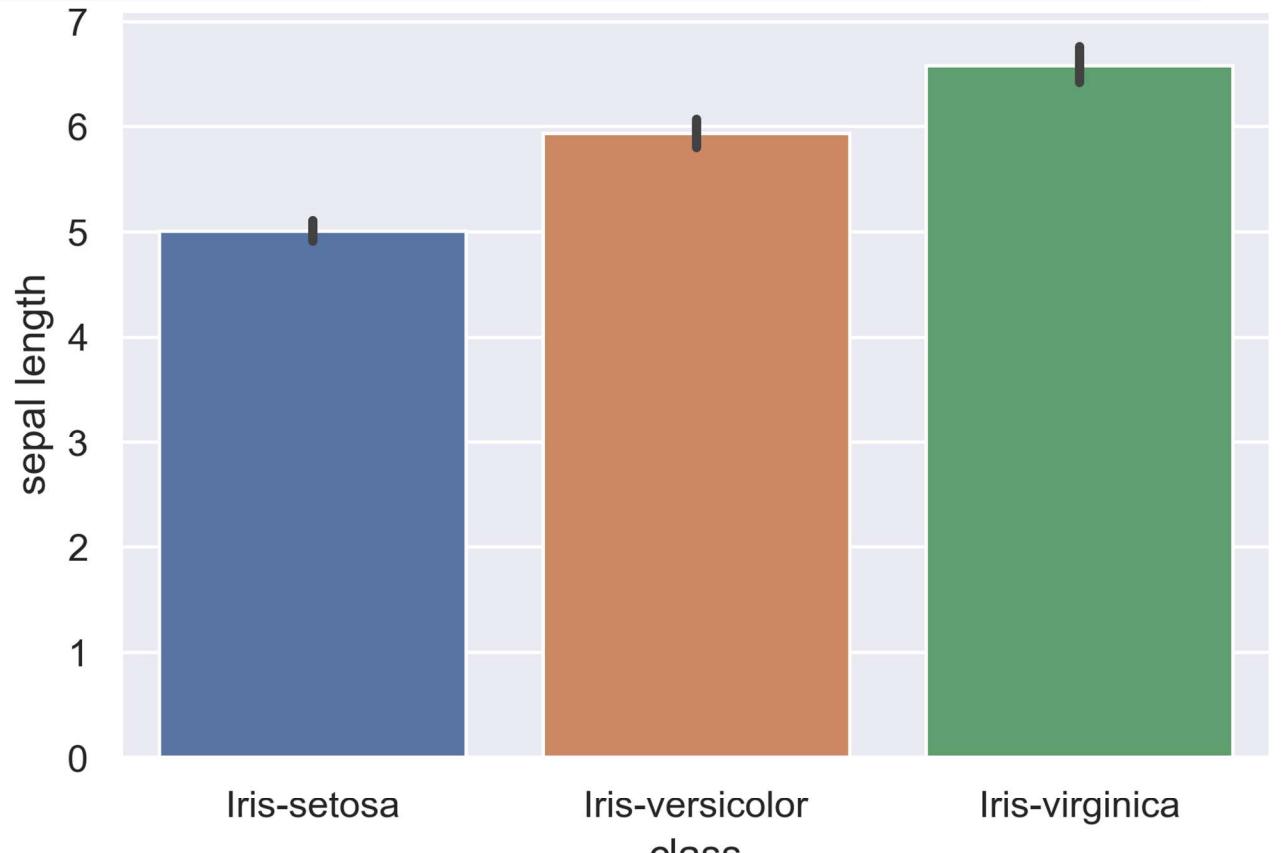
משתנה רציף ומשתנה קטgoriy



מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ - **שני משתנים**

משתנה רציף ומשתנה קטgoriy - BAR PLOT

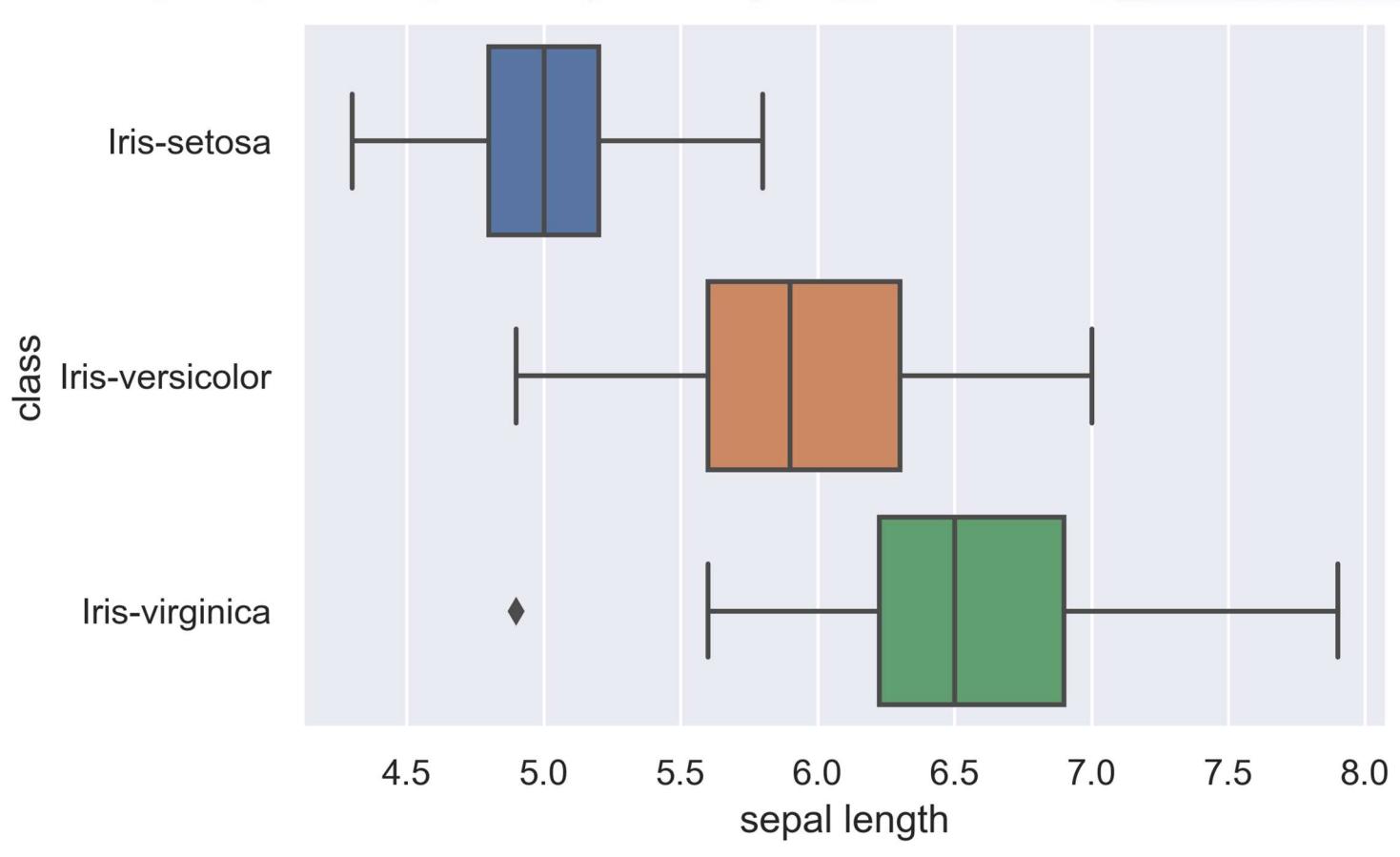
```
sns.barplot(data=df, x='class',y='sepal length' )
```



-מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ -שני משתנים BAR PLOT-

משתנה רציף ומשתנה קטgoriy - BOX PLOT

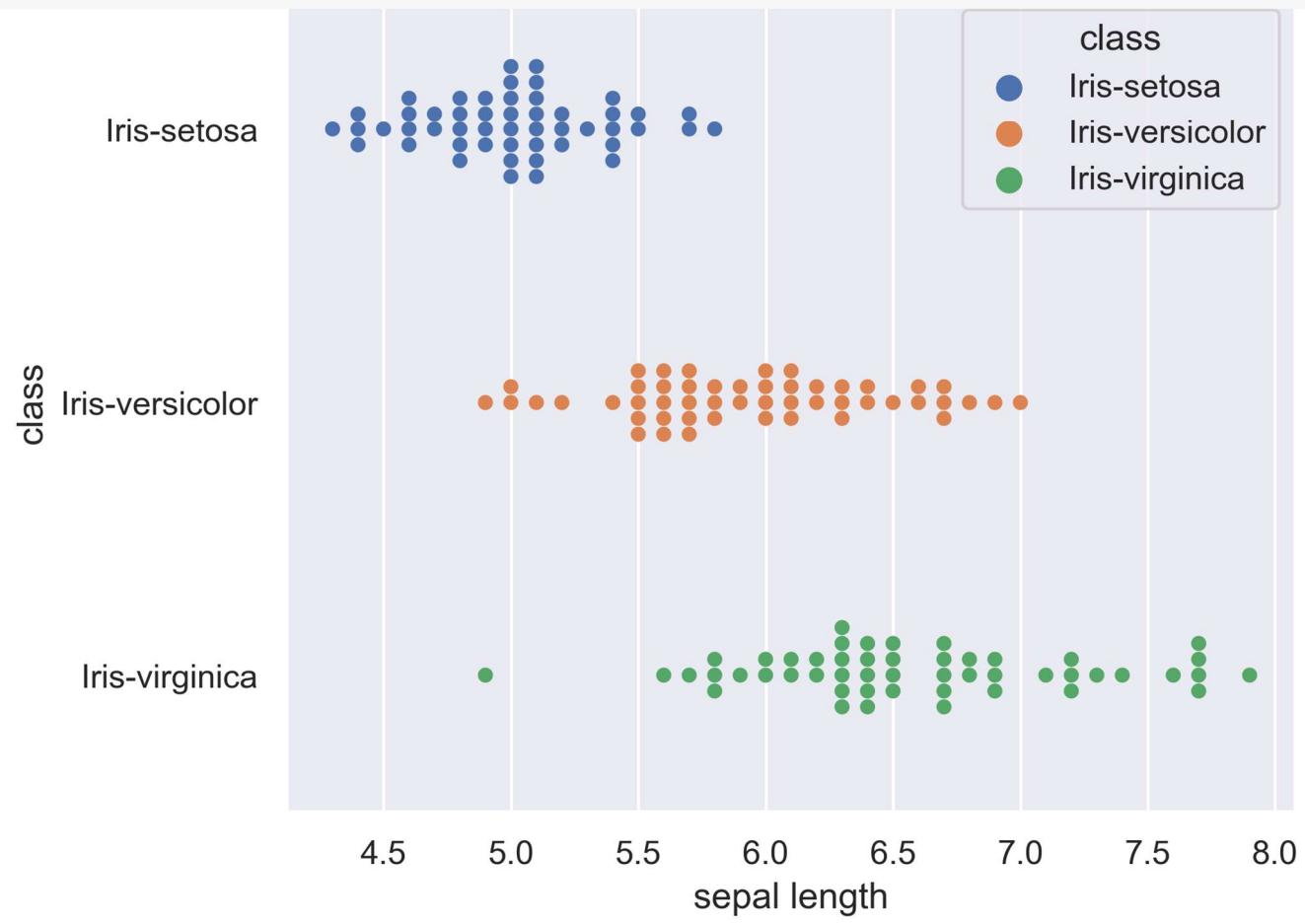
```
sns.boxplot(data=df, x='sepal length',y="class" )
```



-מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ -שני משתנים -BOX PLOT-

משתנה רציף ומינטנה קטגוריה – SWARM PLOT

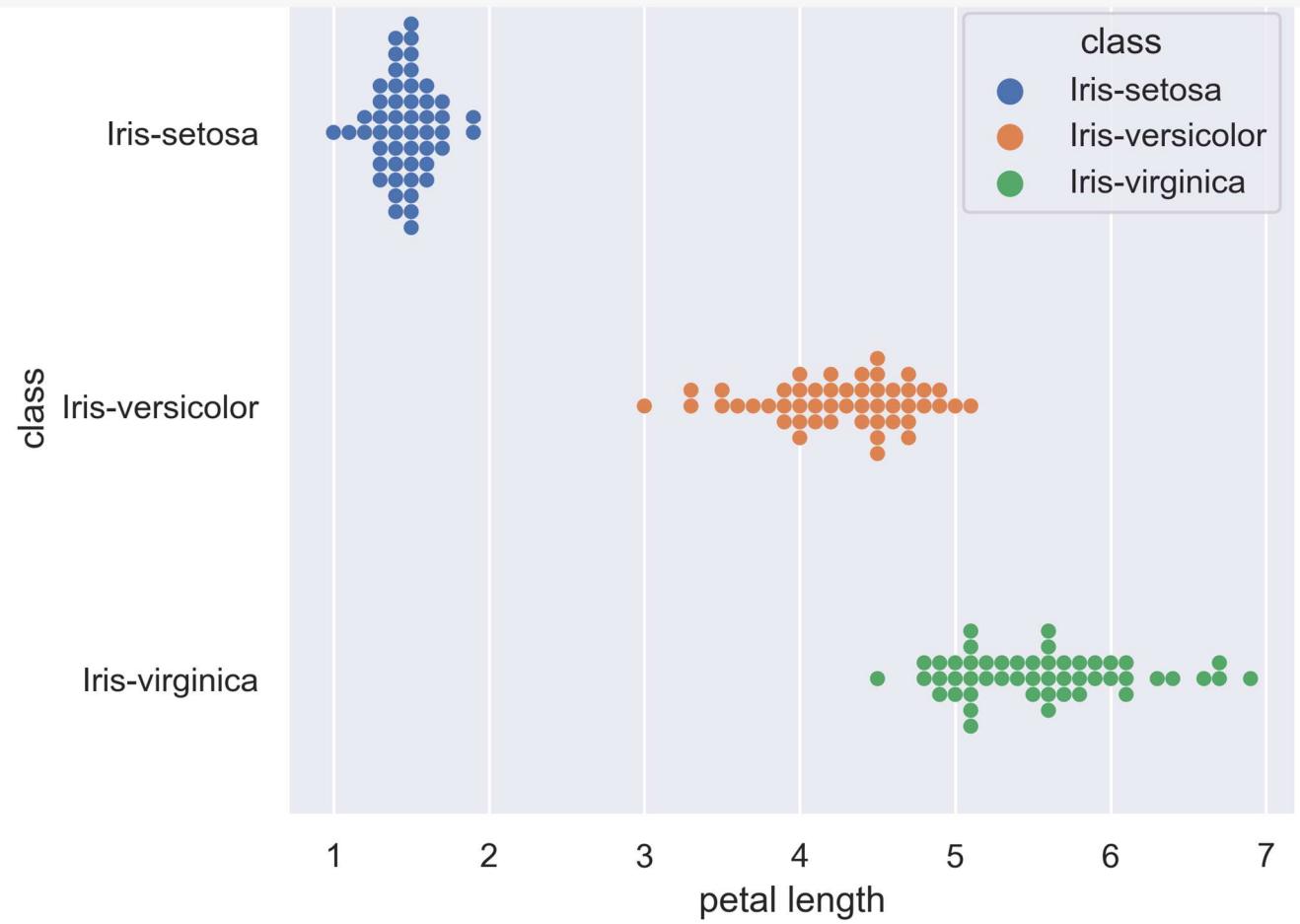
```
sns.swarmplot(data=df, x='sepal length',y='class' ,hue='class' )
```



SWARM PLOT-

משתנה רציף ומינטנה קטגוריתית – SWARM PLOT

```
sns.swarmplot(data=df, x='petal length',y='class' ,hue='class' )
```



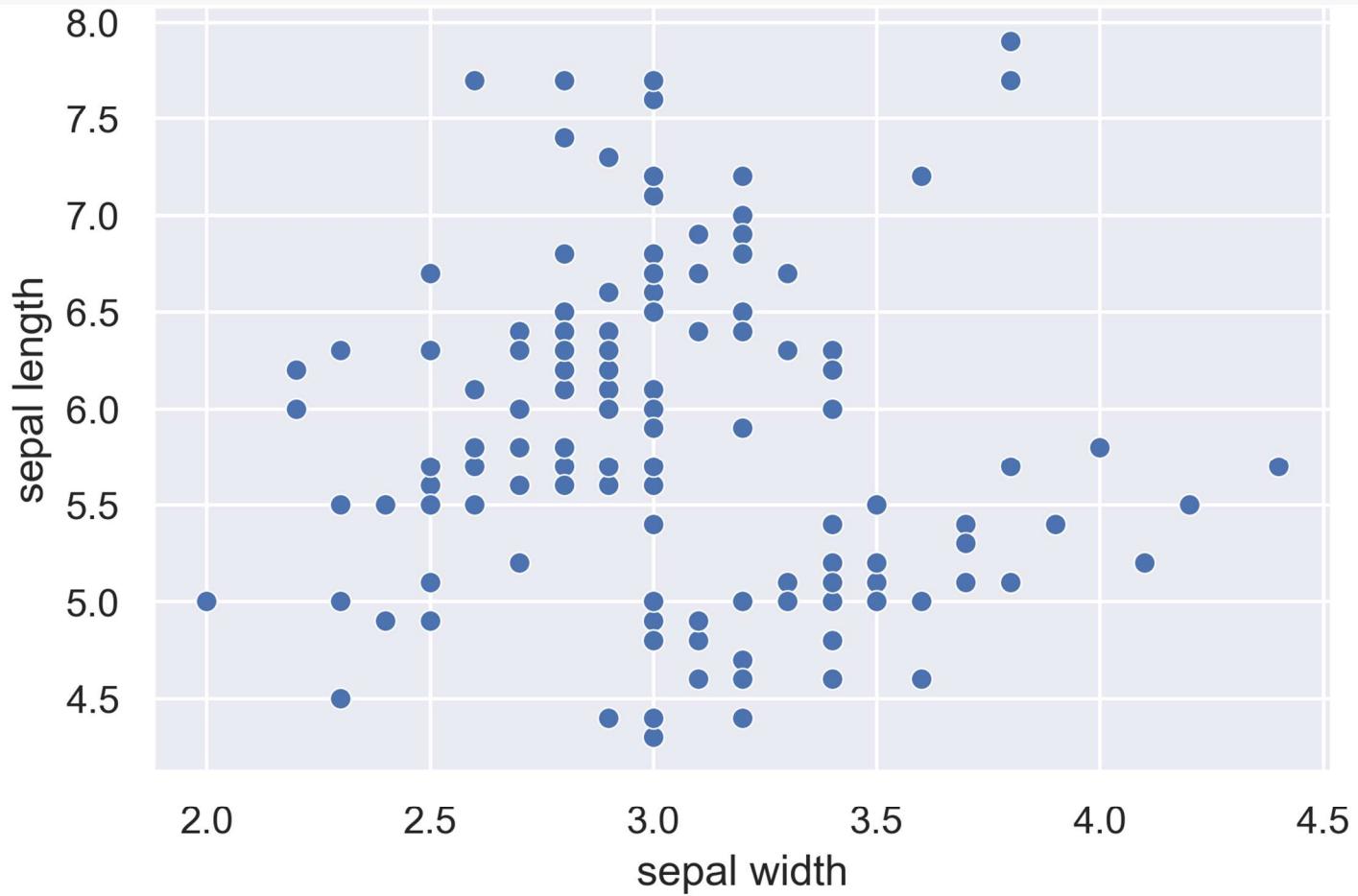
SWARM PLOT-

דניאל ליברמן

-מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ -שני משתנים

שני משתנים רציפים - SCATTER PLOT

```
sns.scatterplot(data=df, x='sepal width', y ='sepal length')
```



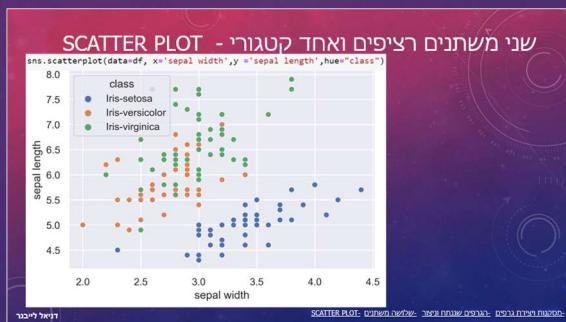
SCATTER PLOT-

דניאל ליבנר

-מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ -שני משתנים

שלושה משתנים

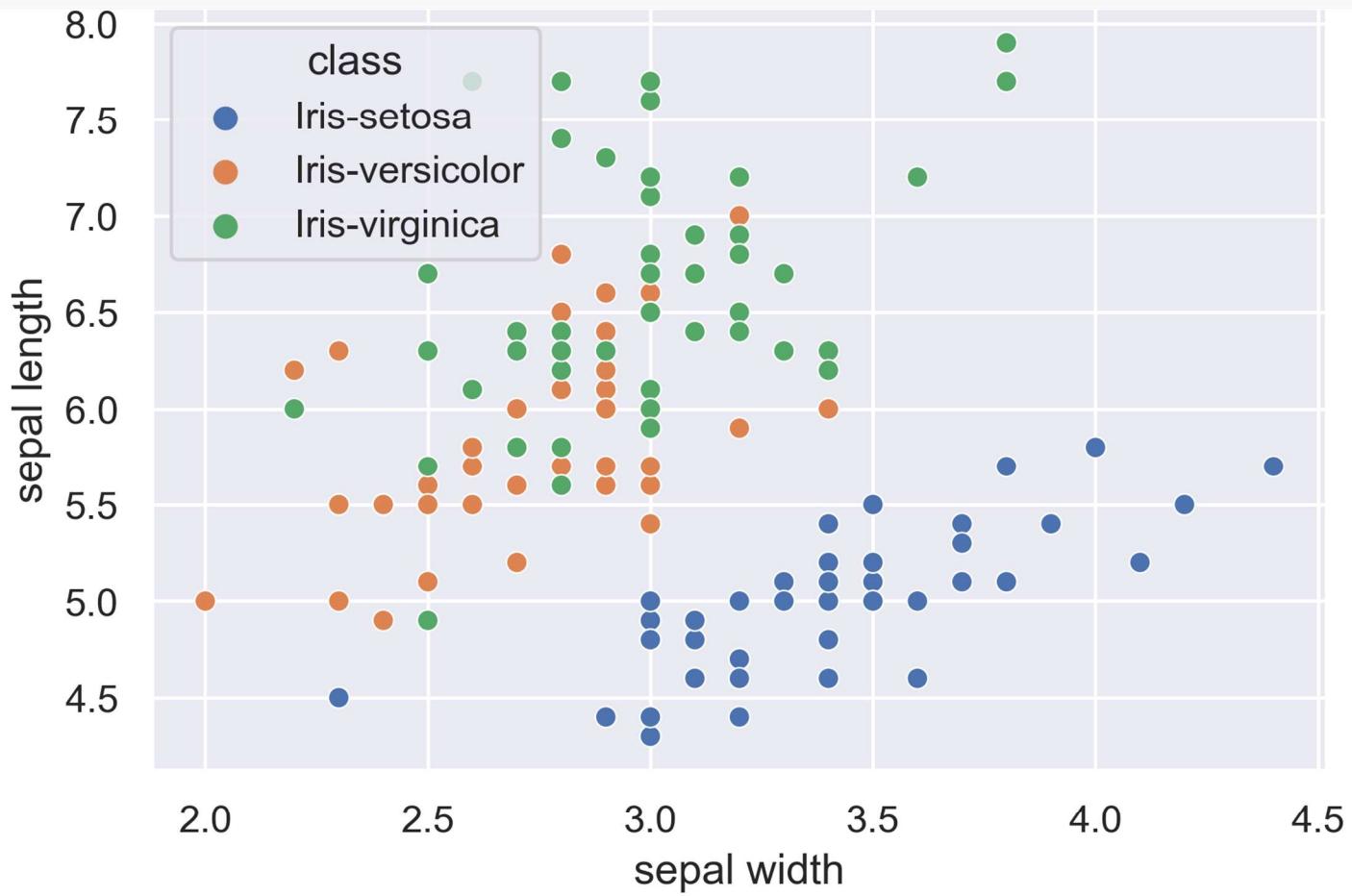
שני משתנים רציפים
ומשתנה אחד קטgoriy



מסקנות ויצירת גרפים - הגרפים שננתח וניצוץ שלושה משתנים

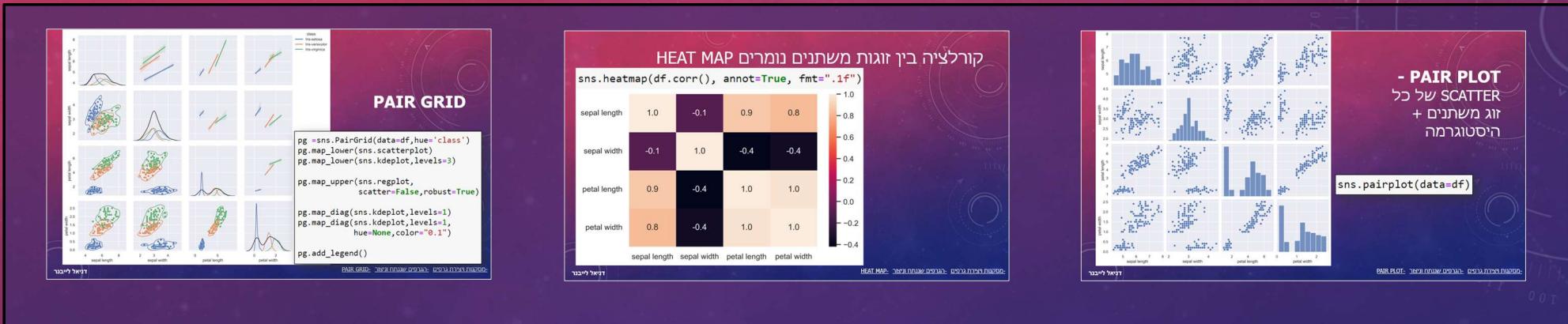
שני משתנים רציפים واحد קטgoriy - SCATTER PLOT

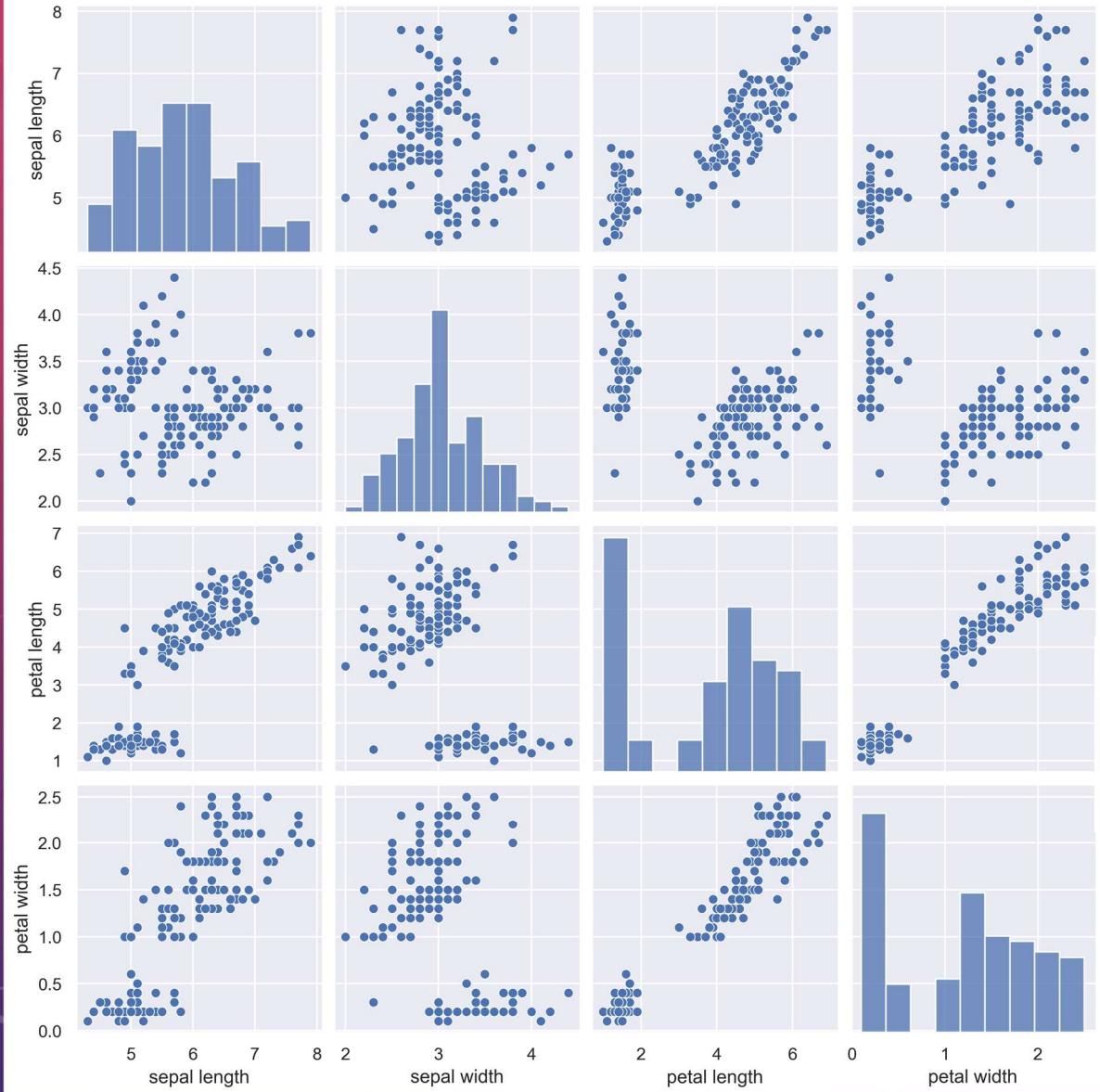
```
sns.scatterplot(data=df, x='sepal width',y = 'sepal length',hue="class")
```



מסקנות ויצירת גרפים - הגרפים שננתח וניצוץ - שלושה משתנים - SCATTER PLOT

כל המשתנים

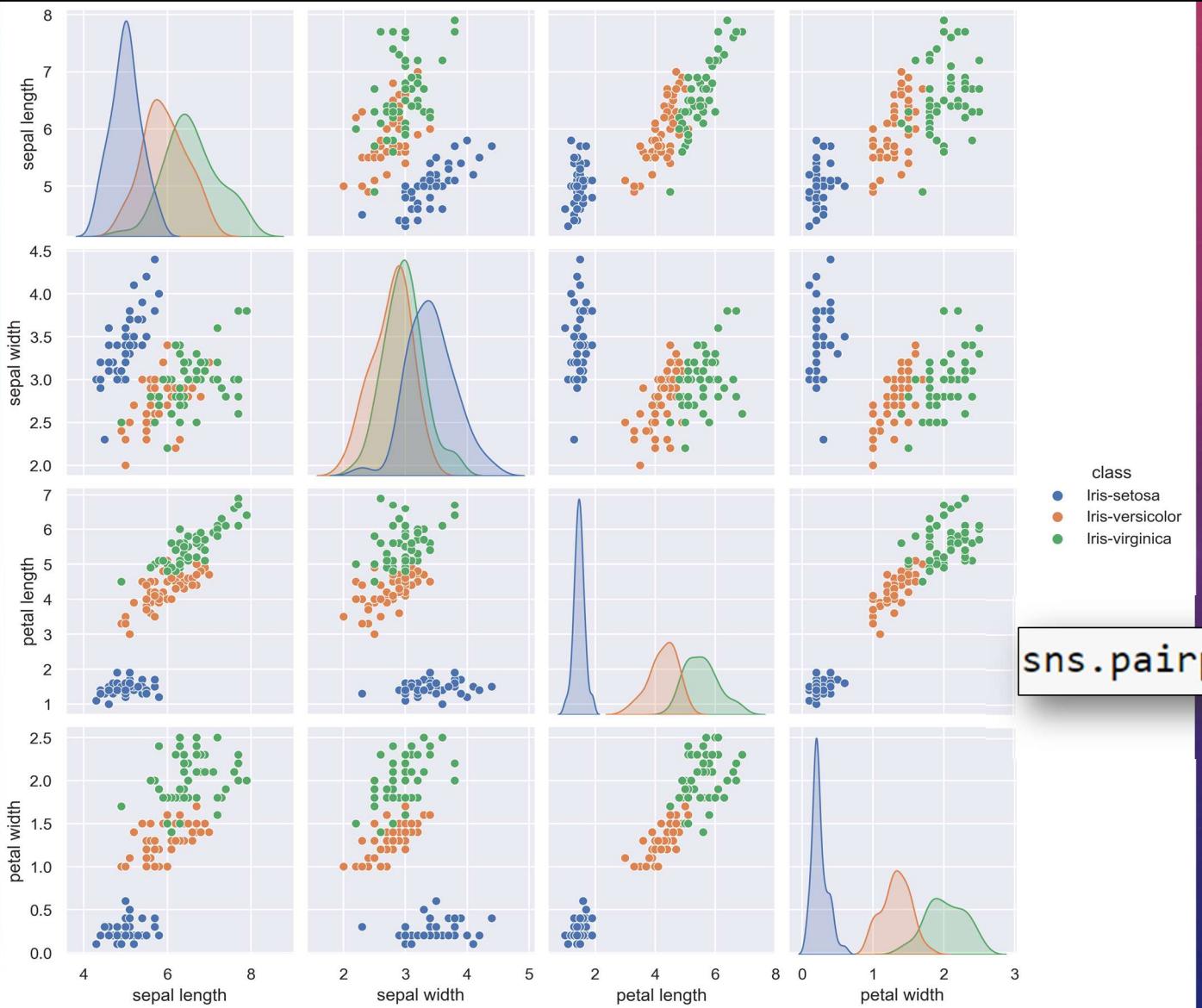




- PAIR PLOT
SCATTER של כל זוג משתנים + היסטוגרמה

```
sns.pairplot(data=df)
```

מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ -PAIR PLOT-



- PAIR PLOT
SCATTER של כל זוג משתנים

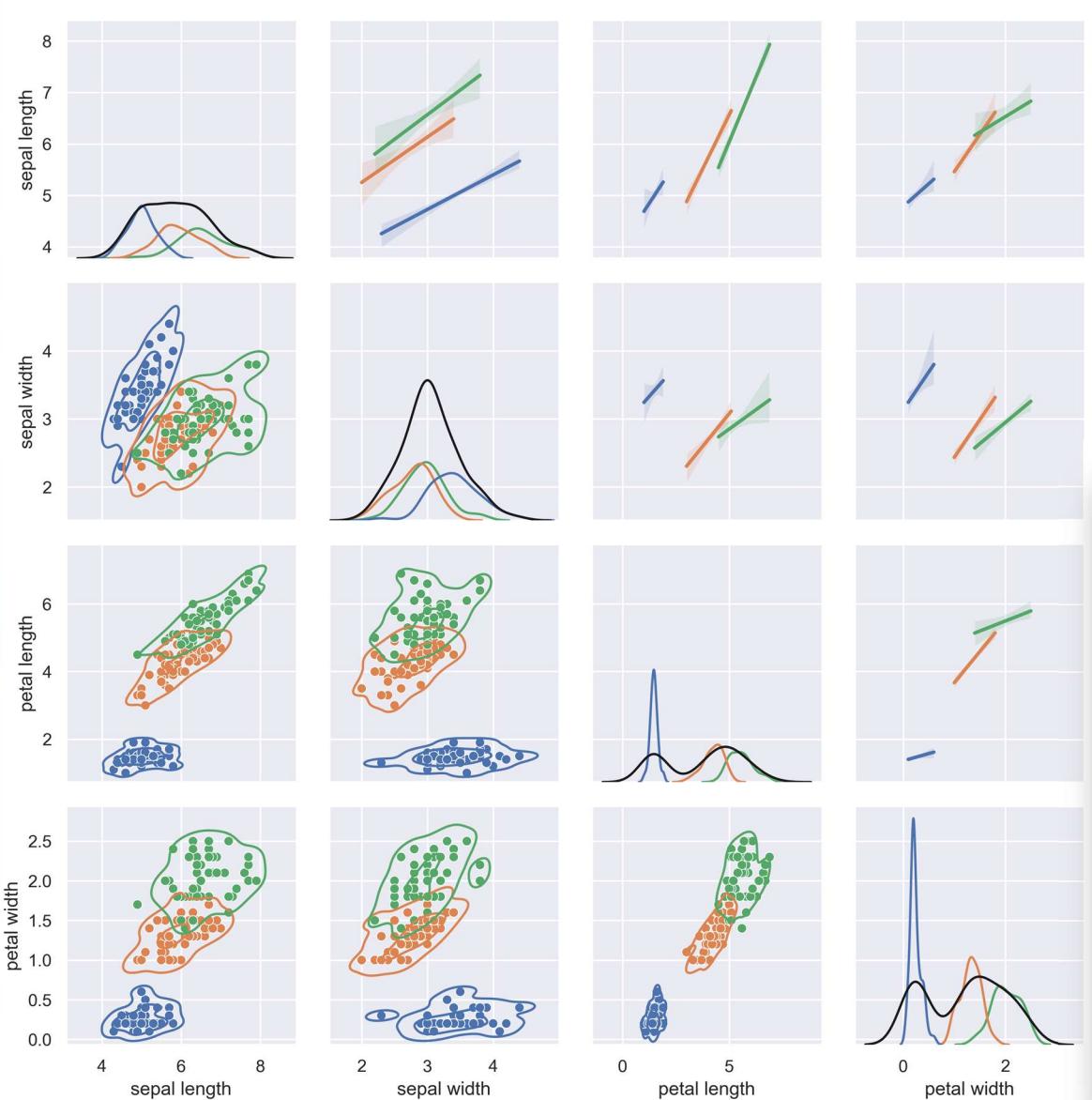
מסקנות ויצירת גרפים -הגרפים שננתח וניצוץ -PAIR PLOT-

קורלציה בין זוגות משתנים נומריים

```
sns.heatmap(df.corr(), annot=True, fmt=".1f")
```



PAIR GRID



```

pg =sns.PairGrid(data=df,hue='class')
pg.map_lower(sns.scatterplot)
pg.map_lower(sns.kdeplot,levels=3)

pg.map_upper(sns.regplot,
             scatter=False,robust=True)

pg.map_diag(sns.kdeplot,levels=1)
pg.map_diag(sns.kdeplot,levels=1,
            hue=None,color="0.1")

pg.add_legend()

```

תודה!

דניאל לייבנר

DA.LEIBNER@GMAIL.COM