

- הכרות עם שפת התכנות פייתון (PYTHON)
- הכרות עם סביבת העבודה של PYTHON
- הכרות עם התכנות הבסיסי בפייתון

➤ ספריות – הכנה ללמידת מכונה:

matplotlib✓

pandas✓

numpy✓

seaborn✓

הכל נמצא ברשת, צריך רק לדעת איפה לחפש:

אתרים ללמוד עצמי - seaborn

– <https://seaborn.pydata.org/>

– <https://reshetech.co.il/machine-learning-tutorials/simpler-better-faster-plots-with-seaborn>

- ✓ ספריה עזר להצגה ויזואלית של הנתונים – בדגש על קשרים סטטיסטיים
- ✓ מבוססת על matplotlib

ספריית seaborn

➤ **ייבוא -** `import seaborn as sns`

➤ **תיאור גרפי של נתונים מתוך קבצים בחלוקה לקטגוריות**

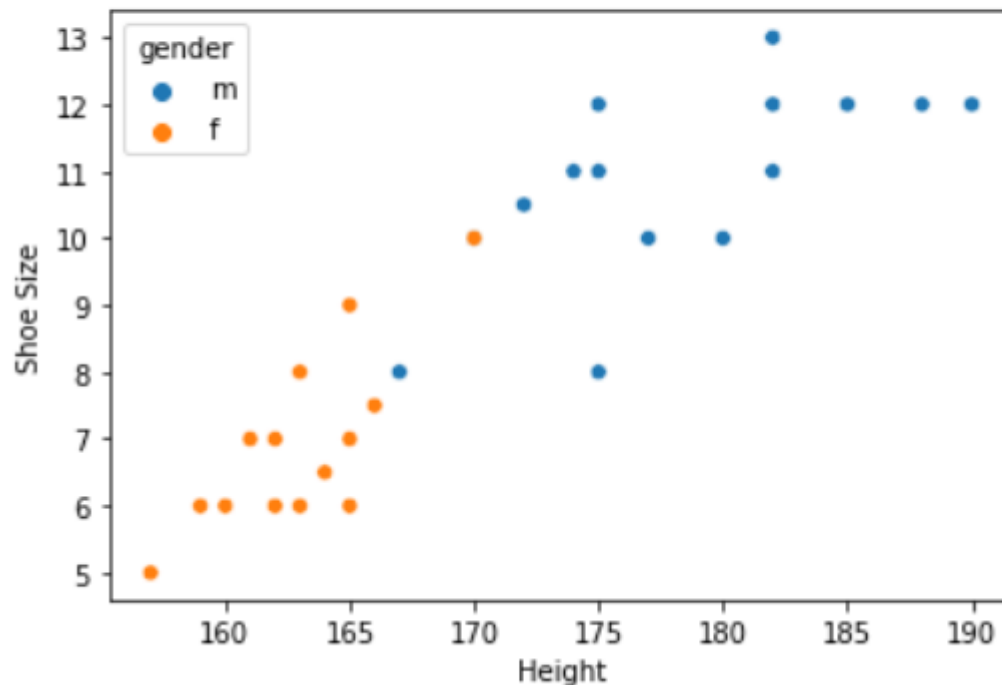
➤ **תיאור גרפי של התאמה בין כל שני מאפיינים**

➤ **גישה למאגרי נתונים – ניתוח סטטיסטי והצגתו**

תיאור גרפי של נתונים מתוך קבצים בחלוקה לקטגוריות

```
[15] sns.scatterplot(x=df['Height'],y=df['Shoe Size'],hue=df['gender'])
```

<matplotlib.axes._subplots.AxesSubplot at 0x7f24adf1c6a0>



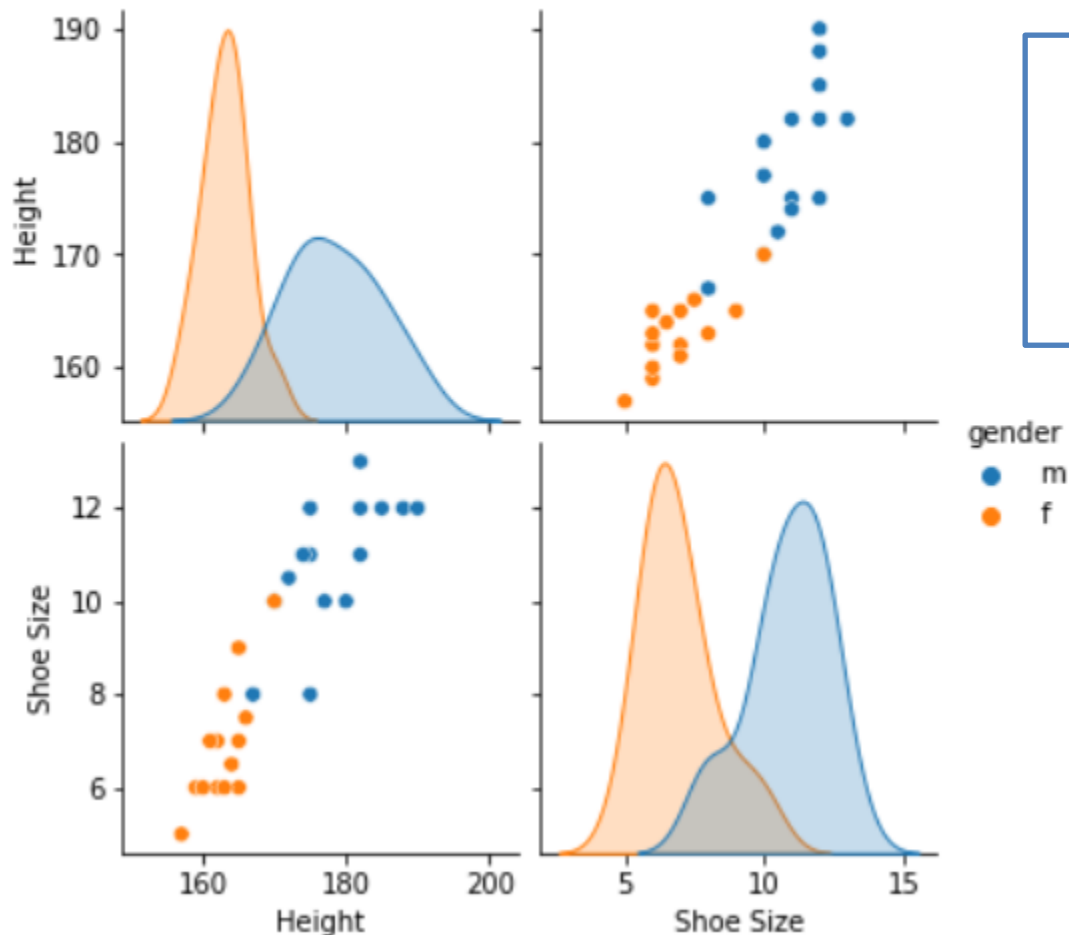
הצגת הקורלציה בין מידת
הנעלים לגובה, בנים בצבע שונה
מבנות

באמצעות גרף פיזור

תיאור גרפי של התאמה בין כל שני מאפיינים

```
sns.pairplot(data=df[['Height', 'Shoe Size', 'gender']], hue='gender')
```

```
<seaborn.axisgrid.PairGrid at 0x7f24acf1fa30>
```



הצגת הקורלציה בין מידת
הנעלים לגובה, בנים בצבע שונה
מבנות

באמצעות גרף פיזור

תרגול 1 - תצוגה גרפית של שערים לוגיים

□ יש לייצר קובץ אקסל עם הנתונים והתוצאה ולייבא את

הנתונים מהקובץ ל data frame

□ יש להציג 3 גרפים באופן הבא:

	A	B	C	D	E
1	value1	value2	OR	AND	XOR
2	0	0	0	0	0
3	1	1	1	1	0
4	1	0	1	0	1
5	0	1	1	0	1

□ גרף פיזור של שער OR

□ גרף פיזור של שער AND

□ גרף פיזור של שער XOR

□ יש לצייר בעיגול כחול ערכי false, יש לצייר בעיגול אדום ערכי

true

✓
0s



```
sns.get_dataset_names()
```

```
['anagrams',  
 'anscombe',  
 'attention',  
 'brain_networks',  
 'car_crashes',  
 'diamonds',  
 'dots',  
 'dowjones',  
 'exercise',  
 'flights',  
 'fmri',  
 'geyser',  
 'glue',  
 'healthexp',  
 'iris',  
 'mpg',  
 'penguins',  
 'planets',  
 'seaice',  
 'taxi',  
 'tips',  
 'titanic']
```

מסד הנתונים tips מכיל נתונים
אודות:
מספר הסועדים,
סכום החשבון,
היום בשבוע,
הטיפ (תשר) ששילמו בסוף
הארוחה.

מעניין אילו נתונים משפיעים
במיוחד על גובה התשר?

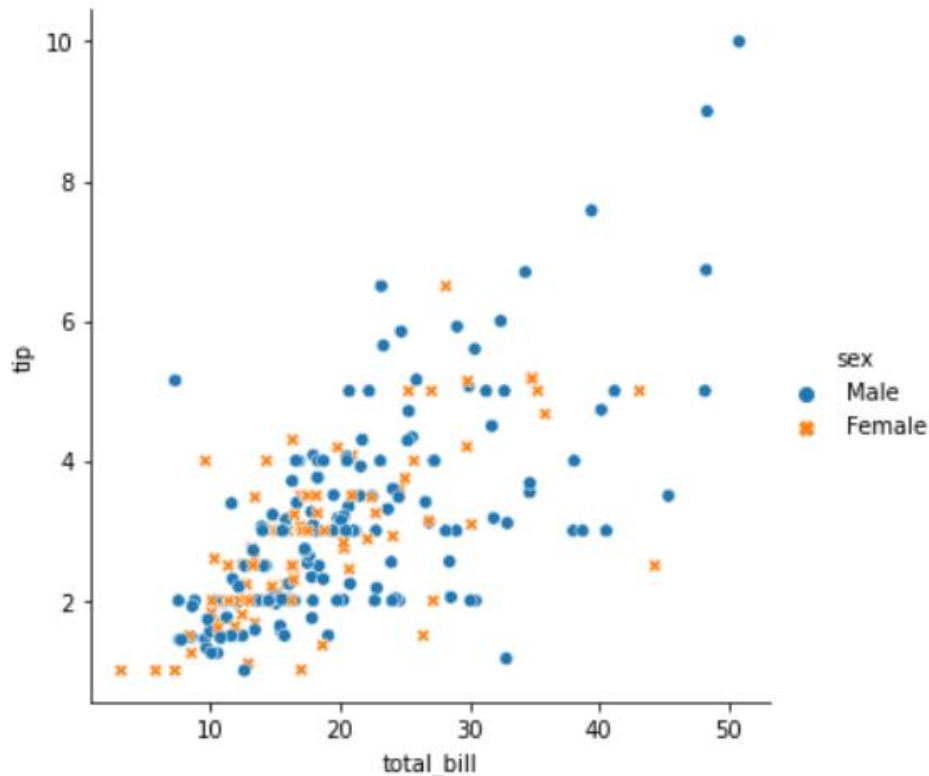
גישה למאגרי נתונים – שלב 1 מציאת קורלציות



קורלציה בין סכום החשבון לטיפ והשונויות בין בנים לבנות

```
sns.relplot(x='total_bill', y='tip', data=tdf, hue='sex', style='sex')
```

<seaborn.axisgrid.FacetGrid at 0x7f24a7c212e0>



כיצד הטיפ מושפע ממינו של האדם.

בפרמטר "hue" – קובעים את הגורם שלפיו רוצים לצבוע בצבע שונה

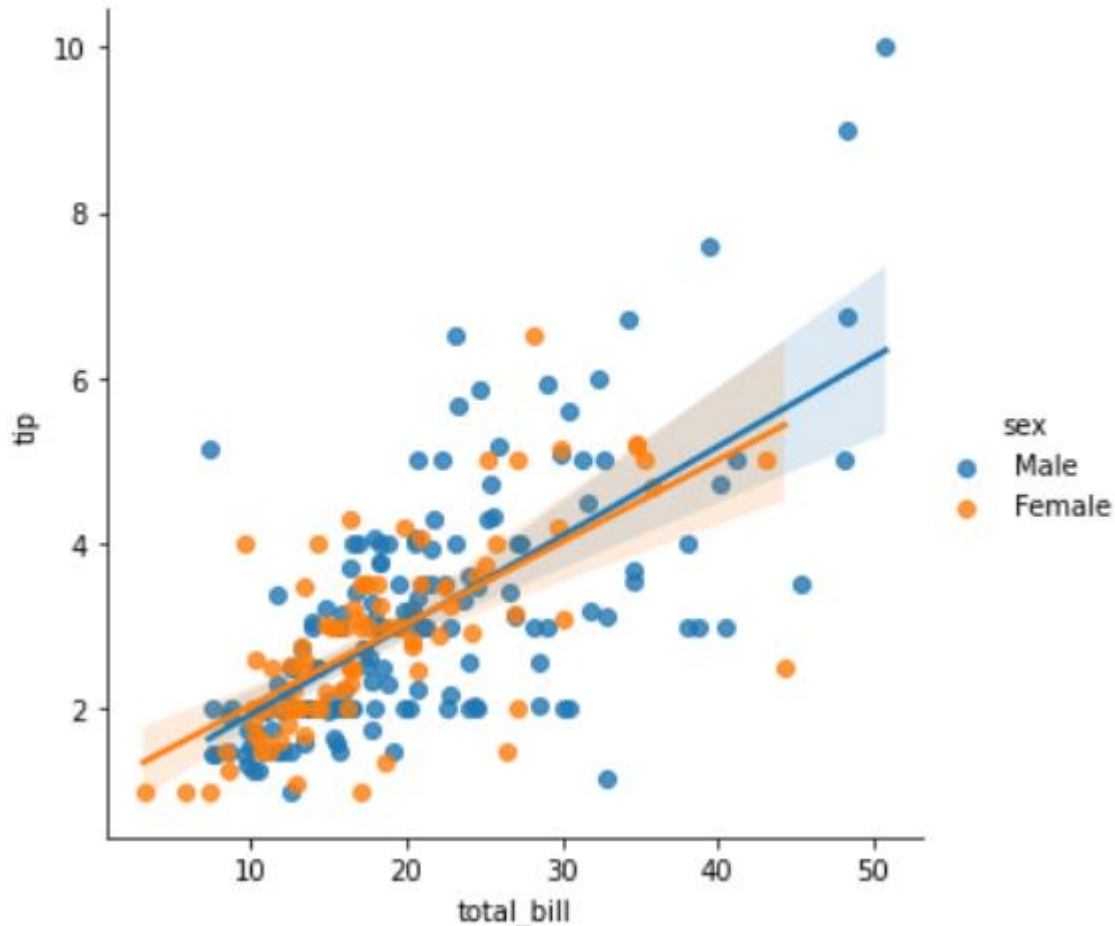
בפרמטר "style" – קובעים את הגורם שלפיו רוצים להדגיש סימון שונה

גישה למאגרי נתונים – שלב 2 הצגת קו מגמה - Implot



```
sns.lmplot(x='total_bill', y='tip', hue='sex', data=tdf)
```

<seaborn.axisgrid.FacetGrid at 0x7f24acaa05b0>



השיפוע הכחול – מייצג את קו
המגמה בין גובה החשבון לגובה
הטיפ במקרה שהסועד הוא בן

השיפוע האדום – מייצג את קו
המגמה בין גובה החשבון לגובה
הטיפ במקרה שהסועד הוא בת

גישה למאגרי נתונים – שלב 2 הצגת קו מגמה - Implot



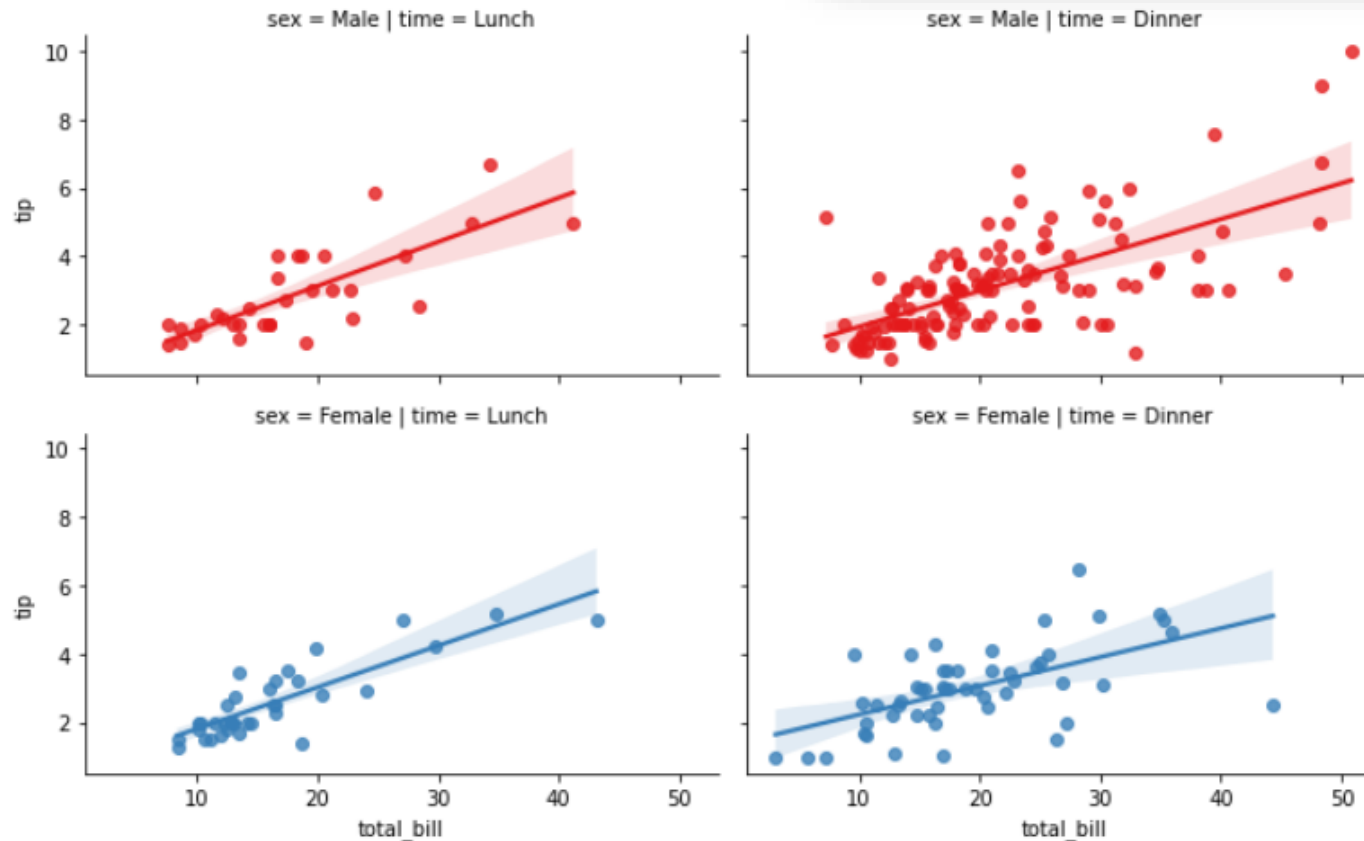
```
sns.lmplot(x='total_bill', y='tip', data=tdf,  
           row='sex', col='time',  
           hue='sex',  
           palette='Set1',  
           height=3, aspect=1.61)
```

הרחבה – בחינת קו המגמה בהתאם לשני גורמים

גורם 1 - סוג הארוחה - צהריים או ערב

גורם 2 - המגדר - בן או בת

↗ <seaborn.axisgrid.FacetGrid at 0x7f24a7cb16d0>



- מתוך מאגר הנתונים הקיים ב seaborn בחרו מאגר נתונים וייבאו את הנתונים למחברת
- העזרו בתרשים pairplot בכדי להציג את הקורלציה בין כל שני מאפיינים
- בחרו בזוג מאפיינים ובשני גורמים נוספים על פיהם תרצו לבדוק את השונות בפיזור הנקודות
- הציגו את גרף פיזור הנקודות, תוך צביעה כל גורם בצבע שונה