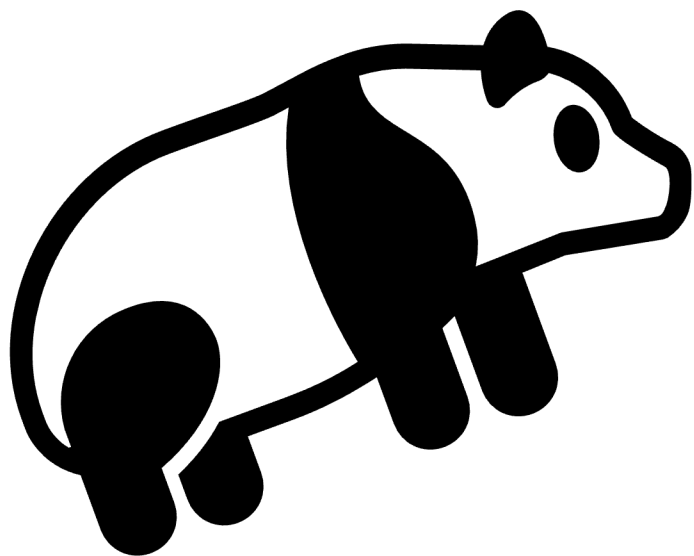


עבודה עם Pandas





על מה נעבור היום

- עבודה עם Series
- מדידת ושיפור ביצועים
- עבודה עם DataFrame
- פעולת GroupBy
- טבלאות ציר – Pivot Table
- שימוש בנתונים מקוטלגים



עבודה עם Series

• מה זה Series?

- מבנה נתונים שמשלב תכונות של List ו- Dictionary
- מורכב מ- Value ו- Index
- Index יכול להיות אובייקט, לא רק מספר
- יש סדר פנימי

עבודה עם Series

• דוגמאות

```
diseases = ['Heart Disease', 'HIV/AIDS', 'Tuberculosis', 'Liver Cancer',  
            'Respiratory infection']  
pd.Series(diseases)  
→  
0          Heart Disease  
1           HIV/AIDS  
2          Tuberculosis  
3          Liver Cancer  
4  Respiratory infection  
dtype: object
```

עבודה עם Series

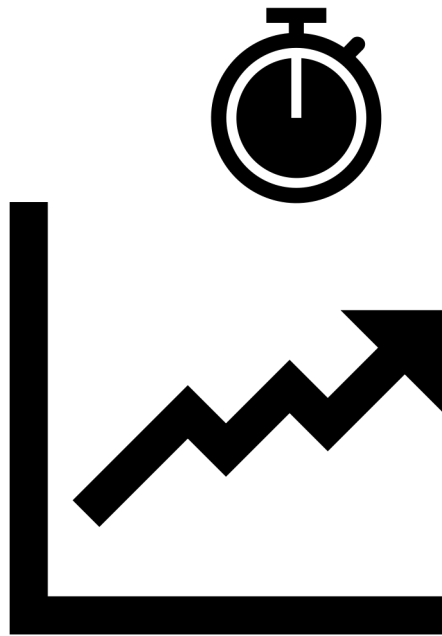
• דוגמאות

```
diseases_by_country = {'Israel' : 'Heart Disease',  
                        'Guatemala': 'HIV/AIDS',  
                        'Guinea' : 'Tuberculosis',  
                        'China' : 'Liver Cancer',  
                        'Grenada' : 'Respiratory infection'}  
  
pd.Series(diseases_by_country)  
→  
China          Liver Cancer  
Grenada        Respiratory infection  
Guatemala      HIV/AIDS  
Guinea         Tuberculosis  
Israel         Heart Disease  
dtype: object
```

מידת ושיפור ביצועים

• שימוש ב-timeit

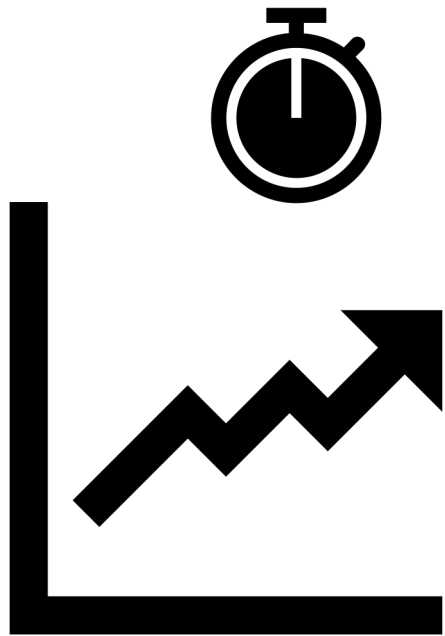
- בקודה של Jupyter
- מאפשר מידת ביצועים בצורה נוחה
- כל מה שצריך לעשות זה פקודה אחת



מידת ושיפור ביצועים

• שימוש ב-timeit

- בקודה של Jupyter
- מאפשר מידת ביצועים בצורה נוחה
- כל מה שצריך לעשות זה פקודה אחת



מתי נראה קוד רץ?!



עבודה עם Data frame

- Data frame מהווה גרסה דו-ממדית של ה-Series
- ניתן לחשוב עליו כמערך דו-ממדי
- ייבוא נתונים אל pandas הוא תהליך פשוט ומיידי

```
df = pd.read_csv('Heart_Disease_Mortality_Among_US_Adults__35__by_Territory.csv')
```




עבודה עם Data frame

- נראה כמה דוגמאות של פעולות בסיסיות על Data frame ...



עבודה עם Data frame

• מה ראינו עד עכשיו:

binary series •

Where •

Dropna •

set_index •

iloc •


Merge •



פעולת GroupBy

- מאפשר לעשות הקבצה של הנתונים ביחס לעמודה/תכונה
- מחזיר אובייקט שניתן לעבור עליו איטרטיבית
- אפשר להשתמש בפונקציית אגרגציה על התוצאה
- אפשר לעשות הקבצה באמצעות פונקציות שכתבנו

נראה דוגמאות...



טבלאות ציר – Pivot Table

- טבלאות ציר – Pivot Table – הן דרך יעילה להפעלת פונקציות בצורה אגרסיבית

נראה דוגמאות...