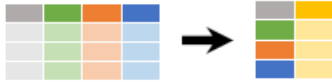


## Summarize Data

`df['w'].value_counts()`  
Count number of rows with each unique value of variable  
`len(df)`  
# of rows in DataFrame.  
`df.shape`  
Tuple of # of rows, # of columns in DataFrame.  
`df['w'].nunique()`  
# of distinct values in a column.  
`df.describe()`  
Basic descriptive and statistics for each column (or GroupBy).



### שאלות-

קרא את קובץ ה-csv הבא לתוך data frame:

<https://github.com/pythonai250824/23.03.2025/blob/main/mobile.csv>

1. צור עמודה של שם פרטי מהמילה הראשונה של שם המנוי, השתמש בפקודת `apply`

2. צור עמודה של חברת סלולר לפי תחילית מספר הטלפון:  
050 פלפון, 051 גולן, 052 סלקום, 053 מוטורלה  
(שם החברה יופיע באנגלית, לדוגמא פלפון=Peleephone)  
רמז: `replace`

3. השתמש בעמודה שיצרת בסעיף הקודם וחשב:  
a. כמה מנויים יש לכל חברת סלולר. רמז: `value counts`  
b. מהם השמות של החברות הסלולריות השונות? רמז: `unique`  
c. כמה חברות סלולריות שונות יש? רמז: `nunique`

4. \*\*רשות: צור עמודה בה יהיה כתוב סוג המשתמש:  
אם זמן השיחה גדול מזמן הגלישה יופיע ערך `talker` אחרת יופיע ערך `surfer`, פתור ב-2 דרכים:  
a. `apply lambda row`  
b. `vectorize`

5. מיון את המנויים לפי זמן שיחה מהקטן לגדול  
מיון את המנויים לפי זמן שיחה + זמן גלישה

6. הדפס את 5 הדברנים הגדולים ביותר. רמז: `nlargest`  
הדפס את 3 המנויים שגולשים הכי פחות. רמז: `nsmallest`

7. נראה שחלק מהרשומות נרשמו מספר פעמים-  
a. הדפס את הרשומות הכפולות  
b. הסר את הרשומות הכפולות

8. הוחלט להגריל 3 מנויים ולהעניק להם פרס-  
הגרל 3 מנויים כלשהם והדפס את השם שלהם ואת המספר טלפון שלהם. רמז: `sample`

את שיעורי הבית יש לשלוח ל- [pythonai250824+HWPD4@gmail.com](mailto:pythonai250824+HWPD4@gmail.com)

יש לעלות את קובץ ה- `ipynb` ל- Github

