

Information Visualization

תרגיל בית 1 – Python

תרגיל בית זה נועד לבדוק את הלימוד העצמי שלכם בתוכנת Python. כל הסעיפים עובדים עם ה-dataset של ה-cars שהתחלתם לעבוד איתו בתרגילי ביתה.

להלן רשימת המטלות שאתם צריכים לבצע:

1. כתבו תוכנית אשר ממלאת את הערכים הריקים בשדה ה-weight (המסומנים ב-*), בממוצע של הערכים האחרים בשדה זה. הדפיסו את הממוצע וסטיית התקן (הפקודה לסטיית תקן – std) של שדה ה-weight לאחר הפעולה.

עבור השאלות הבאות, אם יש שורה בטבלה שאין לה ערך בשדה שאתם משתמשים בו, מלאו קודם את השורה בממוצע כפי שנעשה בשאלה 1. בנוסף, בכל גרף הקפידו על כתיבת כותרת משמעותית לגרף, שמות לצירים ו**לשמור** כל גרף כתמונה.

2. כתבו תוכנית שתצייר באותו גרף, boxplot אחד שיראה אגרגציה של כמות דלק בכביש בין עירוני (Hwy Mpg) למספר צילינדרים (cyl). כלומר, עבור כל ערך של צילינדר (4, 6, 8, ...), יש לצייר boxplot של כמות הדלק.

3. כתבו תוכנית שתצייר, בגרף אחד שני scatterplot שונים, אחד ליד השני.

- אחד של כוח סוס יחסית ל-CITY MPG. צבעו בגרף זה צבע שונה כל מספר של צילינדרים (כלומר בצבע אחד cyl 6, בצבע אחר cyl 4, וכו'). הורידו שורות שאין להן ערכים באחד משלושת שדות אלו.
- שני של CITY MPG יחסית ל-Hwy MPG, כאשר תצבעו את הגרף ב-3 צבעים שונים: מכוניות מסוג sport, מסוג pickup, וכל השאר.

לכל גרף אמור להיות כותרת, שמות צירים ומקרא ברור. מקור מידע מורחב לגבי צביעות בגרף:

<https://www.statology.org/matplotlib-scatterplot-color-by-value/>

4. כתבו תוכנית שתשרטט מטריצה של scatterplots (תלמדו מתודת scatter_matrix – מקור אפשרי). המטריצה צריכה להראות את היחסים בין גודל מנוע, כוח סוס, צריכת דלק עירונית (city mpg), ומחיר לסוחר (dealer price).

חוץ מה שנלמד בתרגילי ביתה, יש המון מידע ברשת על Python. למשל, ניתן למצוא המון מידע על כל סוגי הדיאגרמות ע"י חיפוש: "plotting in python".

- הגשה בשלשות או זוגות בלבד
- תאריך הגשה: **16/4/2023 בחצות**.
- יש להגיש את התרגיל בצורה אלקטרונית בלבד, בקובץ ipynb. וקובץ html. לא לשכוח לכתוב שם ות.ז של המגשים.
- בקובץ עליכם לרשום ב-markdown מספר השאלה שאתם פותרים לפני כל codeblock של שאלה מסוימת.
- עליכם לוודא שבקובץ HTML יש כל תוצאות ההרצות, הגרפים וההדפסות.
- יש לתעד הקוד במידת הצורך.
- יש להגיש לתא ההגשה באתר המודל בקובץ מכוון בפורמט הבא: HW2_123456789_123456789.zip.
- ההגשה תתבצע דרך תיבת הגשה תרגיל בית 1. (לא לשכוח ללחוץ על "שלח לבדיקה"). התוכניות (script) ב-python, צריכות לעבוד, כך שירוצו כרצוי ונקבל אותם תוצאות. ניתן להניח שיש ב-working directory קובץ בשם cars.csv, אבל צריך לכתוב פקודה שתטען את הקובץ.
- יורדו כל הנקודות לתוכניות שלא ירוצו, גם אם זה יהיה רק בגלל טעות ב-syntax.

עבודה נעימה!