בס"ד

**הפרומפט לקלוד – EDA מורחב על מנת לגלות דברים לא הגיוניים בדאטה**

בצע EDA כזה, ולפני כל גרף/טבלה כתוב שורה המסבירה מה בדקת. כמובן שיש לחלק את ה-EDA בצורה מובנת ובתוספת של הערות באנגלית. כתוב אותו בצורה שיהיה קל להחיל את הפייפליין הזה גם פעמים נוספות על דאטה עם ערכים שונים. כשאני כותבת הצג את היחס הכוונה להציג באחוזים.

שים לב: אני אוסרת עליך לכתוב שנית את הקוד שכבר כתבת לי, שמחולל את הדאטהסט! כשאני כותבת "הצג את היחס", "מצא את מספר האחוזים" וכדומה איני מתכוונת שתיתן לי את התשובה לשאלה אלא שתכתוב לי קוד שכאשר מריצים אותו מתקבלת התשובה לשאלה

הצג את היחס בין בעלי הרקע הרפואי לבין כאלו שאין להם רקע רפואי.

הצג את היחס בין בעלי הרקע הפלילי לבין כאלו שאין להם רקע פלילי.

הצג את היחס בין בעלי האזרחות הזרה לבין כאלו שאין להם אזרחות זרה.

הצג את היחס הממוצע לעובדים בין הימים שיש להם בהם פעילות כלשהי לבין הימים שאין להם בהם שום פעילות.

הצג את אחוז העובדים שיש להם גם רשומות ללא פעילות.

בדוק לכמה עובדים חסרה רשומה בתאריך מסוים.

מצא את מספר הימים הזדוניים הממוצע לעובדים שסומנו על ידך כזדוניים, ואת מספר הימים הזדוניים הממוצע לעובדים שסומנו על ידך כלא זדוניים.

מצא את היחס הממוצע בין הרשומות בטבלה שבהן מספר הכניסות שווה למספר היציאות, לבין הרשומות בטבלה שבהן מספר הכניסות אינו שווה למספר היציאות.

ערוך לי פילוח לגבי הרשומות בטבלה שבהן מספר הכניסות אינו שווה למספר היציאות – מהו אחוז הרשומות הללו, מה יש ברשומה בד"כ יותר – האם יש יותר כניסות או יותר יציאות וכו'.

מצא לי מהו אחוז הרשומות שבהן כמות הדפים שהודפסו במועד חריג עולה על כמות הדפים שהודפסו.

מצא לי מהו אחוז הרשומות שבהן הפקודות החריגות עולה על כמות הפקודות הרגילות.

מצא לי מהו אחוז הרשומות שבהן הפקודות החריגות במועד חריג עולה על כמות הדפים שהודפסו.

מצא לי מהו אחוז הרשומות שבהן פקודות ההדפסה עולה על כמות הדפים שהודפסו.

מצא לי מהו אחוז הרשומות שבהן שכמות הדפים הצבעוניים עולה על כמות הדפים שהודפסו.

מצא לי מהו היחס הממוצע בין כמות ההדפסות במועד חריג לבין כמות ההדפסות לא במועד חריג באותו יום לאותו עובד, בימים שבהם היו גם הדפסות חריגות.

מצא לי מהו היחס הממוצע בין כמות הפקודות במועד חריג לבין כמות הפקודות בכלל באותו יום לאותו עובד, בימים שבהם היו גם פקודות חריגות.

מצא לי מהו היחס הממוצע בין כמות ההדפסות בצבע לבין כמות ההדפסות בשחור לבן באותו יום לאותו עובד, בימים שבהם היו גם הדפסות בצבע.

מצא את אחוז הרשומות שבהם שכמות הדפסות צבע+שחור לבן אינו משלים למספר ההדפסות בדיוק.

מצא את אחוז הרשומות שבהן כמות ההדפסות בצבע מתוך כמות ההדפסות הכוללת אינה שווה למספר שרשמת בתור אחוז ההדפסות בצבע.

מצא את אחוז הרשומת שבהן כמות פקודות ההדפסה אינה שווה לכמות הדפים שהודפסו.

מצא מהו היחס בין העובדים הקבלנים -is\_contractor לבין העובדים שלא.

מצא מהו היחס בין דרגות הסיווג השונות.

מצא בכמה רשומות avg\_request גבוה מ-max\_request