**האם אנחנו מתלוננים סתם או הכל ההתחממות הגלובלית?**

* לממשלה כבר נמאס שהאזרחים מתלוננים כל הזמן, והם מנסים למצוא סיבות ללמה זה קורה.

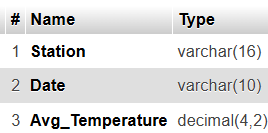
חייל מצטיין מקודקוד 3 הציע שאולי זה לא בגלל העבודה עצמה, אלא בגלל גורם חיצוני.

שלחו אותו לבדוק האם זה נכון.

* החייל המצטיין מפרוייקט קודקוד לקח שני מאגרי מידע - מאגר תלונות ומאגר טמפ'.
* שאלת המחקר: **האם יש קשר בין טמפ' גבוהה לבין כמות תלונות?**
* עזרו לחייל המצטיין לאמת או להפריך את ההשערה!
* צרו DB בשם complaints בMYSQL
* ייבאו את הDB

1. מבנה הטבלאות:  
   תלונות:  
   

טמפ



**משימות תחקור בסיסיות: (מודגש = משימות חובה)**

1. טבלת תלונות:
   1. **על איזה משרד התלוננו הכי הרבה? (שימו לב שעלולת להיות תוצאות שם משרד ריק)**
   2. **לאיזה משרד יש הכי הרבה תלונות מוצדקות?**
   3. **לאיזה משרד יש הכי הרבה תלונות לא מוצדקות?**
   4. **כמה תלונות הגישו אזרחים?**
   5. כמה תלונות לא מוצדקות הגישו אזרחים?
   6. כמה תלונות מוצדקות הגישו אזרחים?
   7. כמה תלונות הגישו עמותות?
   8. כמה תלונות מוצדקות הגישו עמותות?
   9. כמה תלונות לא מוצדקות הגישו עמותות?
   10. כמה תלונות מזכירות קורונה כחלק מהתלונה?
   11. מהי התלונה שעבר בה הכי הרבה זמן מקבלת התלונה עד לסיום הטיפול?
   12. מהי התלונה שעבר בה הכי פחות זמן מקבלת התלונה עד לסיום הטיפול?
2. טבלת טמפ:
   1. **מה הטמפ' הגבוהה ביותר שמדדה תחנת:**
      1. **ירושלים?**
      2. **תל אביב?**
      3. **חיפה?**
   2. מהי הטמפ' הנמוכה ביותר שנמדדה בין התאריכים 1.5-1.10? (קראו על STR\_TO\_DATE)
   3. מהי הטמפ' הגבוהה ביותר שנמדדה בין התאריכים 1.5-1.10? (קראו על STR\_TO\_DATE)

**שאלת המחקר:**

1. כדי לענות על שאלת המחקר צריך להגיע למצב שבו מוצגים לנו 2 פרטי מידע: כמות התלונות לתאריך מסוים, והטמפ' הממוצעת באותו תאריך.
2. אין לנו JOIN פשוט שייתן לנו את הטבלה הזאת, נדרשים כאן כמה שלבים.
3. למה?
   1. את התאריך וכמות התלונות פר תאריך אפשר לייצר מטבלת התלונות.
   2. את הטמפ' הממוצעת לפי תאריך צריך לייצר - כרגע בטבלת הטמפ' מוצגות עד 3 מדידות לאותו יום (משלוש תחנות שונות). צריך לעשות ממוצע בין שלושתן!
   3. את כל המידע הזה צריך לשים בטבלה אחת.
4. ברגע שיש טבלה אחת (אם מסתבכים - ראו שלבי פתרון בסוף המסמך), בדקו כמה תלונות מתוך סך הכל התלונות קרו כשהייתה טמפ' גבוהה מ:
   1. 20 מעלות
   2. 25 מעלות
   3. 30 מעלות.
5. חשבו על התוצאות - מה דעתכם? האם החייל המצטיין צדק?

**שלב אחרון: הנגישו את הדאטה לדורות הבאים.**

צרו תוכנת C# שתטען את הטבלה הסופית בעזרת SQL ותייצא אותה לקובץ CSV.

אקסטרה - פתחו את הקובץ באקסל (או תוכנה זהה) וצרו גרפים שמראים את הקשר שבין הטמפ' לכמות התלונות.

1. שלבים לפתרון שאלת המחקר (מי שלא צריך שינסה לבד!):
   1. בעזרת JOIN, צרו טבלה חדשה שמסכמת את הממוצע טמפ היומי לכל תאריך.
   2. יוצרים טבלה חדשה בעזרת השורה CREATE TABLE <new table name> AS לפני הSELECT, לדוגמה:  
        
      נניח שאנחנו רוצים ליצור טבלה ששומרת את המשרדים הכושלים שמספר התלונות המוצדקות נגדם הוא גדול מ50.  
        
      שאלת הSQL שעושה את זה היא (נסו לבד!)

SELECT `שם הנילון`, COUNT(\*) AS justified\_count

FROM complaints

WHERE החלטה = 'מוצדקת'

GROUP BY `שם הנילון`

HAVING COUNT(\*) > 50;

אם רוצים לשמור את זה בטבלה:

CREATE TABLE offices\_with\_many\_justified AS

SELECT `שם הנילון`, COUNT(\*) AS justified\_count

FROM complaints

WHERE החלטה = 'מוצדקת'

GROUP BY `שם הנילון`

HAVING COUNT(\*) > 50;

* 1. אחרי שיש טבלה עם ממוצע טמפ' יומי ותאריך, אפשר לעשות JOIN עם טבלת התלונות לפי תאריך. שמרו גם את התוצאה הזאת בטבלה חדשה
  2. תחקרו את הטבלה החדשה כדי למצוא כמה תלונות היו לפי טמפ'.