בס"ד

פרומפט לקלוד:

בצע בדאטהסט את השינויים הבאים:

1. את כל השמות הקבועים כמו שמות המחלקות והתפקידים כתוב באנגלית.

seniority\_years .2 אמור להיות מספר בין 0 ל-50.

3. מספר העובד אמור להיות בעל אותו מספר ספרות לכולם, כך שאם תחליט שלכל עובד יש ליצור מספר עובד בן 4 ספרות, העובד הראשון יקבל את האיידי 0001 ולא 1.

trip\_day\_number .4 עליך לעשות מסוג int.

בנוסף, שים לב שזהו רצף. מדובר במספר יום שהותו בחו"ל של העובד, כך שלא ייתכן שביום מסוים העובד היה בארץ ואילו למחרת הוא נמצא בחו"ל ביום השישי לטיולו, אלא רק ביום הראשון לטיולו. ויום למחרת היום הראשון – הערך המותר ביום זה הוא רק 2, אם העובד נמצא בטיול ועדיין לא חזר לארץ. וכך המספר עולה לפי הסדר, לפי הרצף, כל עוד העובד לא שב לארץ.

כאשר העובד בחו"ל – הוא לא יכול להיכנס למבנה ועל כן במציאות הרגילה כל הנתונים המורים על כניסתו למבנה אמורים להיות ריקים או מאופסים, כגון total\_presence\_time וכן entry\_time וכו'. אלא אם כן בזמן שהעובד היה בטיול מישהו אחר נכנס לעבודה עם התג שלו, ואז זה חשוד!

5. הוסף עמודה שבה יהיה כתוב מכמה קמפוסים שונים ביצע העובד הדפסות באותו יום (מספר אינט בין 0 ל-3).

6. מחק את העמודה print\_location\_campus ובמקומה שים עמודה printed\_fron\_other שתוכנה היא האם העובד ביצע באותו יום הדפסות מקמפוס אחד לפחות שאינו הקמפוס שלו (boolean – 0 או 1).

7. הוסף עמודה print\_campuses שבה יהיה כתוב מכמה קמפוסים שונים ביצע העובד הדפסות באותו יום (int בין 0 ל-3).

8. מחק את העמודה burn\_location\_campus ובמקומה שים עמודה שתוכנה היא האם העובד ביצע באותו יום צריבות מקמפוס אחד לפחות שאינו הקמפוס שלו (boolean – 0 או 1)

7. הוסף עמודה burn\_campuses שבה יהיה כתוב מכמה קמפוסים שונים ביצע העובד צריבות באותו יום (int בין 0 ל-3).

בצע EDA ספציפי לכל אחד מהשינויים שביצעת על מנת שאראה שביצעת אותם נכון.

---

אנחנו רוצים שהדאטה יהיה ככל היותר קרוב למציאות מבחינת פעולות חשודות של עובדים חשודים. לשם כך עליך לדעת את המידע המקדים הבא:

א. לגבי employee\_seniority\_years: עובד ותיק לרוב כבר "מוכר" לארגון – המערכות מכירות את דפוסיו. שינוי התנהגות פתאומי אצל עובד ותיק, הרבה יותר חשוד מאצל עובד חדש.

עובדים עם ותק רב שחווים: ירידת שביעות רצון, תחושת מיצוי או בעיות משמעת - נמצאים בקבוצת סיכון לפעולה נקמנית/נואשת. עובדים ותיקים בדרך כלל: צברו הרשאות גבוהות, מכירים "מעקפים" במערכת, יודעים להסתיר עקבות. לפיכך הסיכון גבוה יותר, גם אם ההתנהגות נראית רגילה יחסית.

ב. לגבי ratio\_color\_prints: אם יש הרבה עובדים שאינם מדפיסים בצבע בד"כ, אז שימוש בצבע עשוי להיות סיגנל, כאשר מסמכים צבעוניים משמשים לרוב עבור חומרי שיווק, טפסים לרשויות, דוחות גרפיים, תכניות עבודה, תרשימי ארכיטקטורה וכו'.

ג. לגבי num\_unique\_campus:

יתכן שעובד יגיע לקמפוס אחר שאינו הקמפוס שלו לצורכי פגישות וכדומה, אולם לא אמורה להיות תנועה מרובה בין הקמפוסים השונים היות שהם ממוקמים באזורים שונים ברחבי הארץ. אבל אם יש שינוי פתאומי בכמות הביקורים כאשר התפקיד לא השתנה, זה גם מחשיד.

ד. לגבי printed\_from\_other, burned\_from\_other , print \_campuses, burn\_campuses

הדפסות או צריבות מקמפוס אחר מעוררות חשד, היות שאין סיבה שעובד ידפיס/יצרוב בקמפוס שאינו שלו.

ד. לגבי is\_official\_trip - האם הנסיעה לחו"ל היא בתפקיד - -Boolean 0/1 אם עובד מתחבר מרחוק בזמן שהוא בכלל בחופשה פרטית – זה חריג. פעולה כמו העתקת קבצים רגישים/גישה למערכות בזמן זה - סיכון ברור לדליפה. איך יודעים אם הוא מתחבר מרחוק? אם הוא הדפיס/צרב קבצים.

ה. עליך לחשב risk\_travel\_indicator (Boolean 0/1) באופן הבא: if (

is\_abroad == 1 and

is\_official\_trip == 0 and

is\_hostile\_country\_trip == 1 and

(total\_printed\_pages > 0 or total\_files\_burned > 0)

):

risk\_trip\_indicator = 1

else:

risk\_trip\_indicator = 0

הסבר הבדיקה:

העובד היה בחופשה פרטית בחו"ל, ובכל זאת מישהו הדפיס/צרב דרך המשתמש שלו -זה חשד מובהק להדלפת מידע או גניבת זהות.

אפשרות א - הדלפת מידע בשיתוף פעולה עם גורם פנימי אחר

העובד משתף פעולה עם מישהו אחר מתוך הארגון, שנותן לו גישה או פועל בשמו.

זה תרחיש איום פנימי קלאסי מסוג "collusion" – אחד הקשים לגילוי.

אפשרות ב - גניבת זהות.

העובד לא יודע בכלל שמתבצעת פעולה בשמו, בזמן שהוא בחו"ל.

מישהו גנב את האישורים שלו (Credentials).

זה כבר זולג לתחום של סייבר פלילי – וחשוב מאוד שהמערכת תזהה את זה מוקדם.

ו. קח בחשבון שיש פעולות שנחשבות מאוד חשודות, כך שצפוי שרוב העובדים החשודים יהיו מתוחכמים וכמעט לא יעשו אותן, אם בכלל, על מנת לא לעורר חשד. לדוגמה: התחברות לעבודה מרחוק בזמן טיול שאינו רשמי ( is\_official\_trip הוא 0).

ז. שינויים מזמני העבודה המקובלים עלולים לעורר חשד: entered\_during\_night\_hours (כניסה במהלך שעות הלילה – בין 10 בלילה לחמש בבוקר) early\_entry\_flag(כניסה לפני השעה שש בבוקר) late\_exit\_flag (יציאה אחרונה אחרי השעה 10 בלילה) entry\_during\_weekend (כניסה בימים שישי או שבת)

ח. הוסף עמודה: האם העובד טס לארץ מוצאו. עמודה זו תהיה בוליאנית – 0 או 1 – ותהיה 1 רק כאשר is\_abroad הוא 1, ויעד הטיסה של העובד הוא ארץ המוצא שלו, וזה יהיה 1 לאורך כל ימי ה-trip שלו.

ט. לא תמיד מה שעלול לעורר חשד באופן רגיל הוא באמת חשוד ביחס לעובד עצמו. לדוגמה: יש עובד שאוהב להתחיל לעבוד בשעה מוקדמת. אצלו, הדגל של שעת כניסה מוקדמת לא יעורר חשד, היות שזו התנהגות שאופיינית לו. אבל אם הוא גם יתחיל לפתע לצאת בשעה מאוחרת (כך שיהיה שינוי בהתנהגותו הקבועה), או שמידי פעם יצא בשעה מאוחרת – זה עלול לעורר חשד ביחס לעובד עצמו. כלומר, עובד חשוד הוא עובד שיש בהתנהגותו סממנים חשודים גם ביחס להתנהגותו שלו ולא רק ביחס להתנהגות השכיחה בקרב רוב העובדים.

בנוסף, נניח שעובד טס לאיראן שנחשבת מדינה עוינת. אם זוהי ארץ מוצאו – אז זה לא מחשיד. אבל אם דפוסי הטיסה שלו השתנו, ולדוגמה במקום טיסה פעם בשנה הוא טס פעמיים בשנה – זה עשוי לעורר חשד.

כמובן שכל מה שכתבתי לך כאן זה רק דוגמאות. עליך לדעת ששינויים בהתנהגות העובד הן ביחס לתפקידו והן ביחס להתנהגותו הרגילה – עשויים לעורר חשד!

י. רמות הסיווג הן החל מ-1 – הכי פחות מסווג, ועד 4 – המסווג ביותר.