**🔐 פרויקט: ניתוח מתקפות סייבר ותוכנות זדוניות**

**🗄️ חלק 1 – עבודה מול מסד נתונים (PyMySQL + קובץ טקסט)**

* **1. קרא קובץ טקסט threats.txt המצורף לתרגיל זה, אשר מכיל יומני מתקפות (פורמט: סוג, תאריך, IP תוקף, מדינה, משך התקפה, כמות המידע שהודלף ב MB).**
* **בדוק תקינות הפורמט – וודא שכל השדות קיימים (בדיקה לפי הכמות בלבד).**
* **2. הכנס את המידע לטבלת attacks במסד הנתונים MySQL עם השדות:**
* **attack\_type**
* **date**
* **attacker\_ip**
* **country**
* **duration\_sec**
* **data\_leaked\_mb**
* **is\_investigated (ברירת מחדל False)**
* **בצע שאילתה שמחזירה את 10 כתובות ה-IP התוקפות הכי פעילות.**
* **עדכן את השדה is\_investigated ל true ל- IP הנמצאים בקובץ טקסט בשם blacklist.txt המצורף לתרגיל זה**

**📊 חלק 2 – ניתוחים וגרפים (Python + Matplotlib)**

**כל הגרפים יתבססו על הנתונים במסד הנתונים:**

* **גרף עמודות – כמות מתקפות לפי סוג (Phishing, Malware, וכו').**
* **תרשים עוגה (Pie chart) – התפלגות מתקפות לפי מדינה.**
* **Scatter plot – משך התקפה לעומת כמות המידע שהודלף.**

**📁 חלק 3 – עבודה עם קבצי טקסט**

**קבצים שמוזנים למערכת ומייצרים תובנות לצורך הזנה למסד הנתונים וניתוח נוסף:**

* **קובץ טקסט (activity\_log) המצורף לתרגל זה בנוי בפורמט קבוע: IP, timestamp, פעולה, תוצאה.**
* **זהה שורות חשודות – למשל ניסיון כניסה כושל מ-IP זר (לפי תוצאה).**
* **צור קובץ חדש suspicious\_activity.txt – כולל רק שורות חשודות.**
* **צור קובץ suspicious\_ips.txt – כתובות IP חשודות ללא כפילויות (לפי תוצאה).**

**📊 חלק 4 – עבודה עם Excel**

**עיבוד קובצי Excel והצלבת מידע עם הנתונים במסד הנתונים:**

* **תבנה קובץ אקס חדש אשרבו ישנה רשימת תוכנות זדוניות (עמודות: ארגון, תאריך גילוי, סוג, חומרה).**
* **חשב ממוצע חומרה לכל סוג – הצג כגרף עמודות.**
* **צור גיליון חדש – סיכום פר ארגון: כמות כוללת, כמות חמורות.**

**📌חלק 5 – בקשת ניסוח לקטע main:**

**כתוב קטע main מרכזי עבור הפרויקט, אשר מרכז את כל הפעולות לפי שלבי העבודה אשר נדרשו בתרגיל.**

**🧪 חלק 6 – בדיקות יחידה (Unit Tests)**

**מטרה:** להבטיח תקינות הלוגיקה בפונקציות הקריטיות של המערכת באמצעות מודול unittest של פייתון.

🔍 יש לכתוב בדיקות יחידה עבור כל הרכיבים בסעיפים הקודמים.



**🧩 הנחיות לכתיבת הקוד:**

* הקוד בפונקציית main צריך להיות **קריא, מאורגן ולכלול הדפסות מצב ביניים (print) ברורות**, כך שניתן יהיה להבין מה מתבצע בכל שלב במהלך הריצה.
* יש **להשתמש במחלקות (classes) לפי הצורך**, בעיקר לניהול חיבור למסד נתונים, טיפול בקבצים, או ייצוג ישויות מורכבות.
* יש להוסיף **הערות בקוד (comments)** במקומות מרכזיים, אשר מסבירות את הלוגיקה, ההנחות, והשלבים השונים בתהליך.
* יש להשתמש ב־**Type Hinting** עבור כל הפונקציות והפרמטרים, בהתאם לסוגי הנתונים.
* **הגשה בזמן היא חובה.** עבודות שתגענה באיחור **לא יבדקו**.
* יש לשלוח את **שמות המגישים** בהודעה פרטית בקלאסרום.  
   אם העבודה מוגשת בזוג, **רק אחד מהשותפים מגיש את העבודה**, וכולל אך ורק **קבצי פייטון, קבצי אקסל, קבצי טקסט וקובץ יצוא של בסיס הנתונים שלכם**.
* **אין להגיש קבצים מכווצים (.zip / .rar) – קבצים כאלה לא יבדקו.**