**אבן דרך 4 - "בוס סניפר"**

לפני כחודש הודלפו באפליקציית "טלגרם" **מסמכים ביטחוניים מסווגים** והוצעו למחירה למרבה במחיר. הארוע קיבל בתקשורת את השם "מכירת חיסול" וגרם למהומה בדרג הבטחוני והמדיני.

לאחר בירור מעמיק נמצא שהמסמכים הודלפו מחברת אבטחה מדינית מסווגת. אחד מעובדי החברה ביצע העברה של הקבצים דרך שרתי החברה ונכנס למספר אתרים לא חוקיים כדי לשתף את המסמכים.

עופרה רימון, מנכ"לית החברה, לא רצתה לקחת שום סיכון נוסף, וביקשה לייצר תוכנת ניטור על מחשבי העובדים בחברה!

**המטרה** - לוודא שעובדי החברה מתעסקים בדברים הרלוונטיים להם ולא מבצעים פעולות בעייתיות או אסורות.

**כאן אתן נכנסות לתמונה!**

**על המערכת**

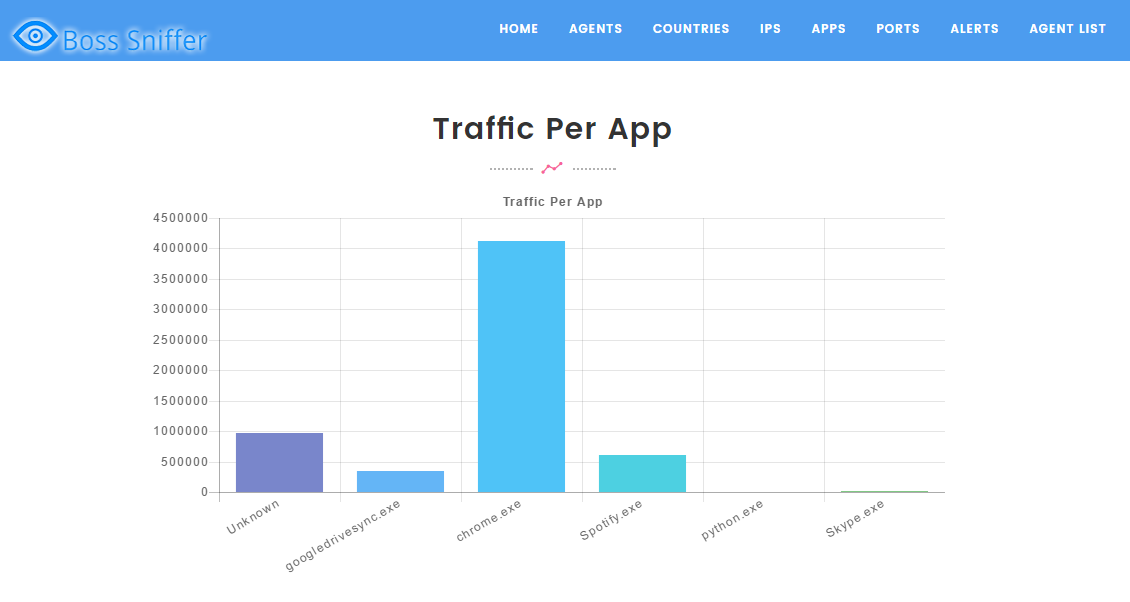
בבניית המערכת נשלב שלושה כלים אשר מאפשרים לנו לגשת לאינטרנט עם פייתון – סוקטים, סקאפי ובקשות HTTP.

אתן תשמשו כמפתחות האישיות של גברת רימון, ותבנו עבורה את המערכת המנטרת את עבודת העובדים.

**תיאור המערכת**

תוכנת המעקב "בוס סניפר" מורכבת משני חלקים:   
1) **הסוכן** – **:Agent** כלי אשר נמצא על מחשבי העובדים ומסניף בצורה שקטה, ואוסף נתונים. אחרי כל 100 פקטות שהכלי הסניף, הוא ישלח את הנתונים שבידיו למנהל, וימשיך בהסנפה.  
2) **המנהל** – **:manager** כלי אשר מקבל דיווחים מהסוכנים השונים ואוסף את המידע כולו בצורה מסכמת, עבור מנהלת החברה. הכלי ייצר דו"חות (קבצי HTML) המסכמים את המידע שהגיע מהסוכנים, ומראה אם יש עובדים שמבצעים פעולות בעייתיות ברשת.

דו"חות ה-HTML יבנו לפי טמפלייט מצורף ויראו כך:



**שלבי פיתוח המערכת**

שלב א' – פיתוח סוכן בסיסי, מנהל בסיסי והתקשורת בניהם

שלב ב' – הרחבת הפעילות של הסוכן ושל המנהל

שלב ג' – שליחת הדוחות לאינטרנט, והרחבות נוספות

**אופן העבודה והגשת הפרוייקט**

* **שימו לב! רק שלב א' יתבצע ב- "זוגות חשיבה"**

תפקידו של זוג להתייעץ, להיעזר, לחשוב ביחד ולבדוק את התוכנה יחד. כל זוג גם חייב לסכם ביחד על פרוטוקול התקשורת בין המנהל לסוכן, ולהשתמש באותו הפרוטוקול. עם זאת לכל אחת יהיה פרוייקט משלה שכולל את כל הקוד הנדרש לשלב זה.

לאחר סיום שלב א', כל אחת ממשיכה את יתר השלבים בנפרד, כאשר כמובן ש**מגישות את אבן הדרך כל אחת בנפרד**.

* **מה מגישים לאחר השבוע הראשון**

לאחר השבוע הראשון, יש להגיש דו"ח התקדמות באבן הדרך. זהו מסמך טקסט, עם תאור מילולי קצר של מצב ההתקדמות, כולל ציון השלבים שסיימנו. שימו לב- תגישו מסמך כזה בכל מקרה, המטרה היא שהמדריך ידע מה מצבכן

* **הגשת קוד** – בכל סיום שלב (מתוך השלבים א', ב' ו- ג') יש להגיש את הקוד שנכתב עד לשלב זה. שמות הקבצים יתחילו עם מספר השלב שסיימתן, לדוגמא: **step1\_**agent.py עבור סיום שלב א', 2 לשלב ב' וכך הלאה.

לכן צריך לזכור לשמור בצד גירסה בסיום כל שלב!

באופן זה תמיד תהיה לכן גירסה עובדת מוכנה שכבר הוגשה.

הגשות הקוד הן תמיד באופן פרטני, גם לאחר שלב א'.

**שלב א' (יתבצע ב"זוגות חשיבה")**

**בשלב הראשון של הפרוייקט נְפַתֵּחַ את הסוכן (הסניפר) הבסיסי ומנהל בסיסי. הסוכן ידווח למנהל על ההסנפות והמנהל רק יציג אותן - ללא ביצוע סטטיסטיקות מתקדמות.**

**1. כתיבת סוכן בסיסי**

נכתוב את הסוכן אשר אוסף נתונים על המחשב המקומי, ללא החלק שמדווח למנהל. על הסוכן להסניף בצורה אינסופית, ולאסוף תיעוד של כל החבילות שהוסנפו. עבור כל חבילה יש לשמור:

* IP איתו שוחחנו
* האם זו תעבורה נכנסת או יוצאת (בוליאני)
* פורט אצל ה-IP איתו שוחחנו (לא הפורט המקומי שלנו)
* גודל החבילה בבתים

שימו לב – אין צורך לשמור את תוכן החבילות אלא רק את נתוני המסגרת הנ"ל.

על הסוכן לבצע הסנפות בצורה של "סבבים". בכל סבב יוסנפו 100 חבילות ובסיום הסבב יתבצע עיבוד של המידע ושליחתו אל המנהל, בפרוטוקול טקסטואלי לבחירתכן. (בין סבב לסבב זה בסדר שלא תהיה הסנפה לזמן קצר). לאחר שהסבב יסתיים – יתחיל סבב חדש של הסנפה.

**דגשים:**

* על התוכנה לטפל בחבילות אשר כוללות UDP או TCP מעל IP בלבד.
* לצורך התרגיל ניתן להניח שהכתובת של המחשב עליו רץ הסוכן היא כתובת IP פרטית. (איך אפשר להחליט בפייתון אם כתובת היא חיצונית או פרטית? העזרו בגוגל)
* את הדיווח למנהל יש לבצע בפרוטוקול שתמציאו בעצמכן, שמבוסס על **UDP** (כלומר הודעה אחת בלבד ולא שיחה). נרצה להמיר את המידע ששמור במבני נתונים של פייתון (כמו רשימות ומילונים) למחרוזת על מנת לשלוח בסוקט. פעולה זאת נקראת פרסור parse. יש להשתמש [בפורמט JSON](http://docs.python-guide.org/en/latest/scenarios/json/) הפשוט והנפוץ לצורך זה, ולהמיר אליו בעזרת מודול ייעודי של פייתון. אפשר לקרוא על פורמט JSON גם במצגת המצורפת בתיקייה.
* **חלוקה לפונקציות - שימו לב!** **רלוונטי גם לסוכן וגם למנהל:** נזכור שכדי להקל על כתיבת הקוד, וכן לכתוב קוד נכון יותר, שגם קל יותר לתחזוקה- עלינו לבצע חלוקה נכונה לפונקציות, לפי העקרונות שלמדנו:
  + לפונקציה יהיה תפקיד אחד בלבד
  + נשמור שהפונקציה תשאר כללית עד כמה שניתן
  + נשמור על הפונקציה סגורה ומופרדת עד כמה שניתן
  + נשמור על הפונקציה פשוטה וקצרה

אפשר גם להתחיל את כתיבת הקוד, ולחלק אותו לפונקציות תוך כדי ההתקדמות.

* כדי להקל על תהליך הדיבאג, ניתן לעשות הדפסות שמתארות מה קורה (למשל: "מתחיל הסנפה..." "שולח את המידע..") אך כמובן לא לפחד להשתמש בדיבאג של PyCharm!
* כדי להקל על תהליך הפיתוח, הוסיפו טיפול בשגיאות (try-except) רק לאחר שהכל עובד.

**2. כתיבת מנהל בסיסי**

המנהל הוא תוכנה אשר מקבלת מידע ממספר גדול של סוכנים במקביל. המנהל הוא למעשה שרת המאזין לדיווחים שנכנסים אליו מהסוכנים השונים, וכל הזמן צובר את המידע של כולם.

בגרסה הבסיסית למנהל יהיה קובץ הגדרות בשם settings.dat (קיימת דוגמא בתיקיית התרגיל) ובו יהיו הפרטים הבאים:

* הסוכנים הפעילים – עבור כל אחד יהיה שמור ה-IP שלו ושם העובד (הפרטים קבועים מראש בקובץ אצל המנהל ולא מתעדכנים בלייב).

בגרסה הבסיסית של המנהל עליו לבצע את הדברים הבאים בלבד:

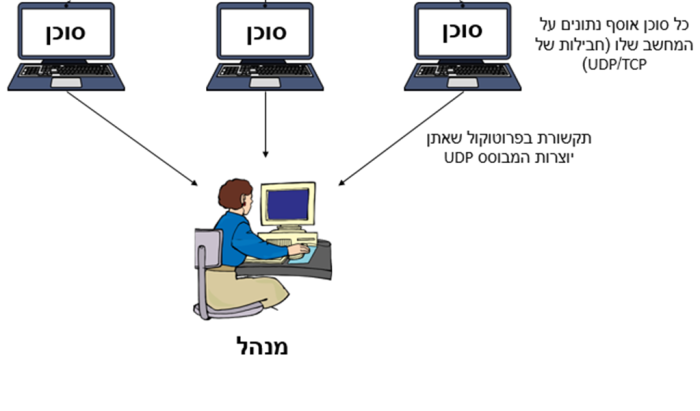
* האזנה ב-UDP לדיווחים מכל כתובות ה- IP.
* בכל פעם שמגיע דיווח יש להציג אותו (את רשימת כל הפאקטות) ובנוסף את שמו של העובד שממנו הגיע הדיווח. (במידה והכתובת מוגדרת מראש בקובץ ה- settings.dat עם שם העובד- נתייחס לשם, במידה ולא- נשתמש בכתובת עצמה)

**דגשים:**

* שימו לב, בשלב זה אין צורך לנפק דו"חות HTML.
* **גודל הבאפר-** נשים לב- הפונקציה recvfrom שמבצעת את קבלת המידע מהלקוחות (הסוכנים), מקבלת פרמטר בשם bufsize - זהו הגודל ב- bytes שהשרת יכול לקבל בבת אחת. עד עכשיו הזנו בדרך כלל 1024 כי שלחנו מעט מידע והגודל הספיק. הפעם אנחנו שולחים כמות יותר גדולה של מידע, ולכן נרצה להשתמש בבאפר גדול יותר. נזכור שכדאי להזין גודל כמספר שהוא חזקה של 2.

**מבלבל?**

אולי התרשים הבא יוכל לעזור, כך תראה התקשורת בין הסוכנים למנהל:



לסיכום, מה אנחנו רואים בתרשים? כל סוכן אוסף נתונים על המחשב שלו.  
כל סוכן מעביר את המידע על המחשב שלו למנהל (באמצעות פרוטוקול שאתן יוצרות), המנהל אוסף את הנתונים מכל הסוכנים.

נזכר שבחלק א' העבודה היא ב**- זוגות חשיבה** ועבודה משותפת. הטבלה הבאה היא הצעה לשלבי העבודה – אתן יכולות כמובן לשנות את רצף העבודה לפי בחירתכן.

**רצף משימות - הצעה:**

|  |
| --- |
| **משימות** |
| הבנת המשימה |
| יצירת סוכן ומנהל המתקשרים בניהם טכנית על בסיס UDP |
| בדיקה  (שליחת מחרוזת מהסוכן והדפסה שלה במנהל) |
| סוכן: הסנפת הפקטות והוצאת המידע הרלוונטי, בחירת מבני נתונים מתאימים |
| החלטה על פרוטוקול טקסטואלי להעברת המידע בפורמט JSON |
| סוכן: מימוש תקשורת מול המנהל ע"פ הפרוטוקול שעליו הוחלט |
| מנהל: מימוש התקשורת מול הסוכנים ע"פ הפרוטוקול שעליו הוחלט, ועל פי הסוכנים הרשומים ב- settings.dat |
| מנהל: הצגת כל דיווח שמגיע – רשימת פאקטות, העובד ממנו המידע מגיע, ועוד.. |
| בדיקות (נחשוב על מקרים שונים ונבדוק אם יש להם מענה) |
| סוכן: הוספת טיפול בשגיאות לאחר שהכל עובד, שליחת שגיאה לבדיקה בשרת |
| מנהל: הוספת טיפול בשגיאות לאחר שהכל עובד |
| בדיקות (נחשוב על מקרים שונים כולל שגיאות אפשריות ונבדוק שיש מענה) |

בסיום שלב א' שימו לב שלכל אחת יש את כל הקוד המעודכן, לפני הפיצול שממנו כל אחת עובדת בנפרד על השלבים הבאים

**שלב ב'**

**בשלב השני של הפרויקט נרחיב את פעולת ההסנפה אצל הסוכן (גילוי התוכנה של כל פאקטה, והוספת מדינה לפי IP). בצד המנהל, נאסוף סטטיסטיקות מכל המידע שנכנס אליו ונייצר פלט של דו"חות HTML אוטומטיים.**

**1.   
סוכן - הוספת שדה "תוכנה" להסנפה בסוכן**

* לפרטים אשר אנו אוספים על כל חבילה נוסיף גם את התוכנה אשר קיבלה/שלחה את החבילה.
* כדי לדעת זאת, יש להיעזר בפקודת הווינדוס netstat –nb. (שימו לב, כדי להריץ אותה אנו צריכים הרשאות אדמין. יש ללחוץ כפתור ימני על הסמל של ווינדוס ולבחור  
  Windows Power shell (ניהול) או "שורת משימה (ניהול)" כדי לפתוח cmd בהרשאות אדמין. כדי להריץ פקודת ווינדוס מתוך פיתון ולקבל את הפלט שלה, היעזרו בגוגל.
* עבור חבילות שלא מזוהה התוכנה שלהן, פשוט כתבו Unknown.

**סוכן – הוספת שדה "מדינה" להסנפה בסוכן**

* קבלת מדינה לפי IP בעזרת API :
  + כדי לקבל מדינה של כל IP יש להיעזר ב- API שמציע האתר [ip-api.com](http://ip-api.com/docs/). באתר עצמו אפשר להתנסות בשירות- להזין IP ולראות את התוצאות המתקבלות, אבל כדי להשתמש בשירות מתוך הקוד יש לגשת אליו בבקשת HTTP על פי הדקומנטציה המתוארת באתר.
  + כדי לשלוח בקשת HTTP מפייתון ישנן ספריות פייתון מוכנות לשליחת HTTP. נעזר באחת מהספריות האלו. **הסבר במצגת המצורפת לתיקייה.**
  + יש שתי אופציות לפורמטים עבור תשובה מהשירות - פורמט CSV או JSON – בחרו מה שהכי נוח לכן לפרסר**.**
  + שימו לב, אין לבצע את הגישות ל- ip-api.com בזמן ההסנפה אלא רק בין הסבבים (לפני הדיווח למנהל).
* שימו לב כי יש לבדוק לכל IP את המדינה שלו רק פעם אחת ולא שוב ושוב עבור כל חבילה (דבר שיאט מאוד את התוכנה ועשוי לחסום אתכן מהשירות החינמי בשלב מסוים). נסו לחשוב באיזה מבני נתונים של פייתון תוכלו להעזר לצורך זה.
* שימו לב כי יש לבדוק את המדינה של הכתובות החיצוניות ולא הכתובת המקומית של המחשב.

**2. מנהל - יצירת סטטיסטיקות מתקדמות**

* על המנהל לשקלל את המידע המגיע אליו מהסוכנים השונים, על מנת לנתח אותו ולהציג בדו"ח. בחרו מבני נתונים מתאימים על מנת לשמור את המידע המשוקלל.
* המנהל ייצור מהנתונים שמגיעים אליו דו"ח HTML. ישנה תזכורת על HTML במצגת המצורפת בתיקייה.
* לצורך יצירת הדוחות יש להשתמש בתבנית template.html, כאשר כל הספריות הנלוות אליה נמצאות לידה (css, fonts, js …). מומלץ לפתוח את התבנית קודם כל בדפדפן, ולאחר מכן לפתוח את הקוד ב-Notepad++ ולבצע בו חיפוש אחרי מילים שראיתם בדפדפן). עליכן לערוך את הקובץ template.html כך שיתאים לצרכים שלכן ותוכלו לבצע בו החלפות בקלות.
* יש לכן דו"ח HTML מוכן לדוגמא example.html. מומלץ לבצע השוואה בין הדו"ח המוכן לבין התבנית template.html, אפשר לעשות את זה למשל ע"י העתקת תוכן הקבצים אל אתר שמבצע diff אונליין, לדוגמא www.diffchecker.com/diff
* בכל הדוחות הערך המוצג (המספרים) הם לפי מספר בתים bytes.
* בכל פעם שיגיע דיווח נוסף מ-סוכן, המידע יתווסף אל הנתונים המשוקללים, וישמר דו"ח HTML חדש (במקום הקודם).
* יש להוסיף לקובץ ההגדרות settings.dat רשימה של כתובות Blacklisted – כתובות IP שהמנהל אוסר על העובדים להיכנס אליהן. עבור כל כתובת ישמר גם שם האתר (למשל "facebook.com:123.23.12.1"). אפשר לראות דוגמא לכך בקובץ ההגדרות לדוגמא בתיקייה.

בדו"ח ה- HTML יופיעו התראות על גישות לאתרים לא מורשים. ההתראות האלו נמצאות בסוף הדו"ח תחת כותרת Alerts.

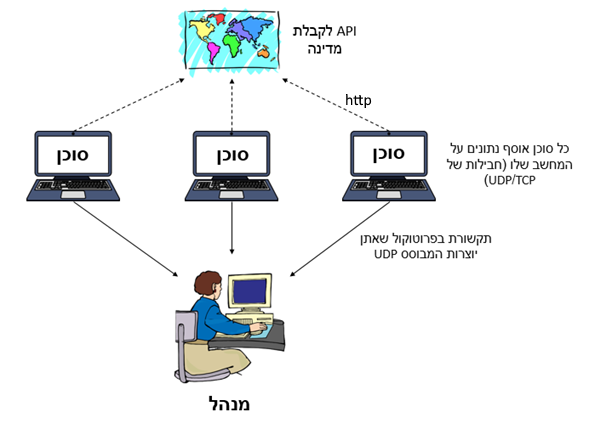
**הדו"חות אשר מפיק המנהל (תמונות בסוף המסמך):**

* תעבורה נכנסת פר משתמש.
* תעבורה יוצאת פר משתמש.
* מדינות נפוצות ביותר בקרב כל המשתמשים.
* IP נפוצים ביותר בקרב כל המשתמשים.
* תוכנות נפוצות בקרב כל המשתמשים
* פורטים נפוצים בקרב כל המשתמשים
* רשימת גישות לא מאושרת לכתובות Blacklisted (יש להציג רשימה של כל הגישות האסורות שישתמשו, עבור כל אחת יש להציג משתמש ו-IP. מספיק להציג כל IP פעם אחת בלבד).

**רשימת משימות - הצעה:**

|  |
| --- |
| סוכן: הוספת שדה תוכנה להסנפהמנהל:קבלה והצגה של שדה תוכנה  בדיקות |
| מנהל: ניתוח הדו"ח – ניתוח המידע שהגיע מהסוכנים והוספת רשימת blacklist |
| בדיקות |
| סוכן: חקירה ומימוש התממשקות מול  ip-api.com שליחת מדינה  מנהל: שימוש במידע על מדינה  בדיקות |
| מנהל: שינוי ושמירת דו"ח כHTML |
| בדיקות |

תרשים:



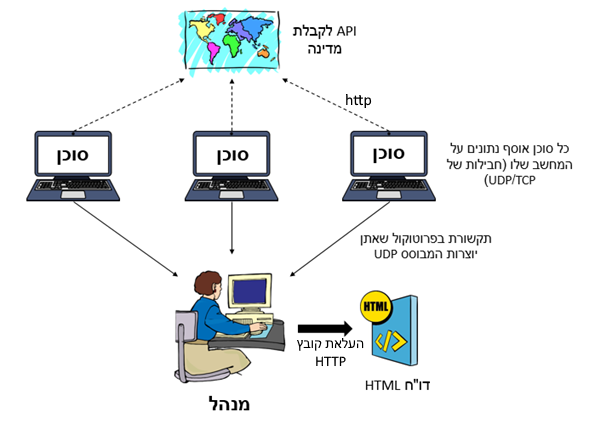
**שלב ג'**

**בשלב זה נוסיף שליחה של הדוחות לאינטרנט. בנוסף ישנן הרחבות נוספות שניתן לעשות בתור בונוס.**

**העלאה לאינטרנט**:

בתיקיית הפרוייקט מצורפת תוכנת ReportUpload אשר משתמשת בפרוטוקול מיוחד כדי להעלות קבצי HTML לשרת ה-Web של BossSniffer. השתמשו בתוכנה כדי להעלות דו"ח (קובץ (HTML לשרת, וצפו בו דרך הדפדפן לאחר העלאה, בכתובת שתתקבל. בתור שם משתמש, השתמשו בשם הפרטי שלכם + נקודה + שם המשפחה (באנגלית).

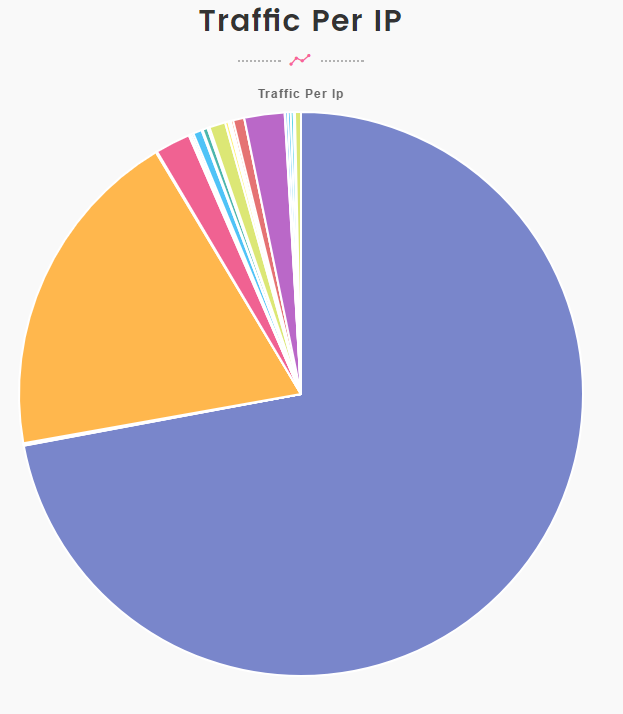
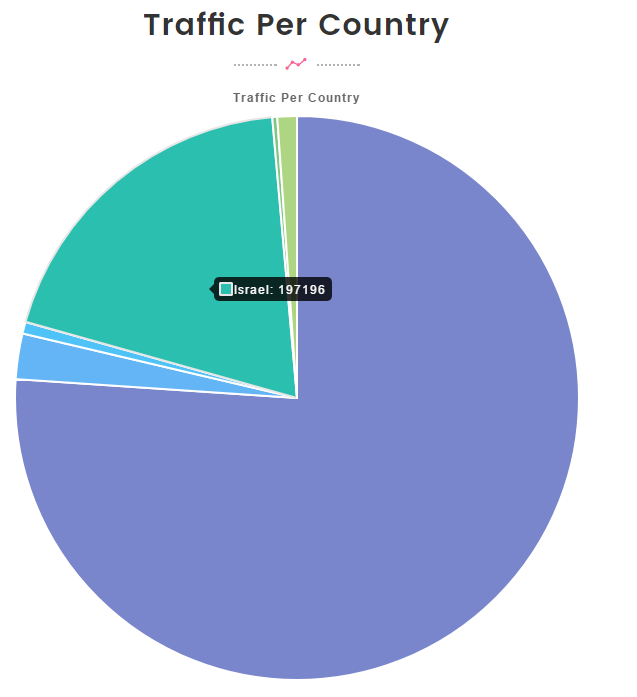
**העלאת קבצים באופן אוטומטי:**  
עליכם לחקור את התוכנה ולכתוב קוד שמעלה את הדו"חות לאתר אוטומטית בכל פעם שאתם מייצרים קובץ. חל איסור מוחלט להעלות לשרת קבצים שאינם קשורים לפרוייקט.

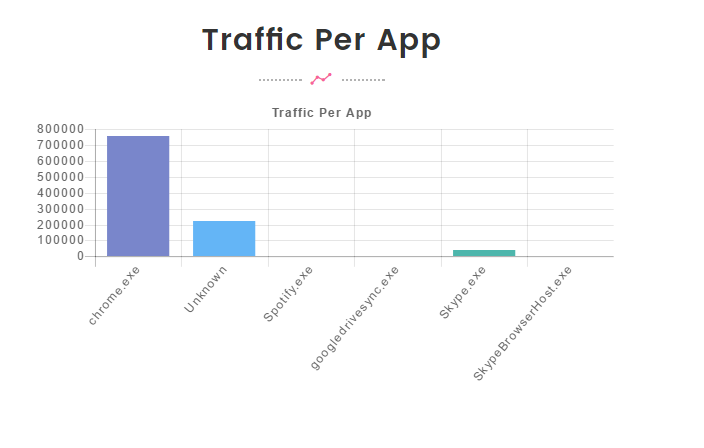
****

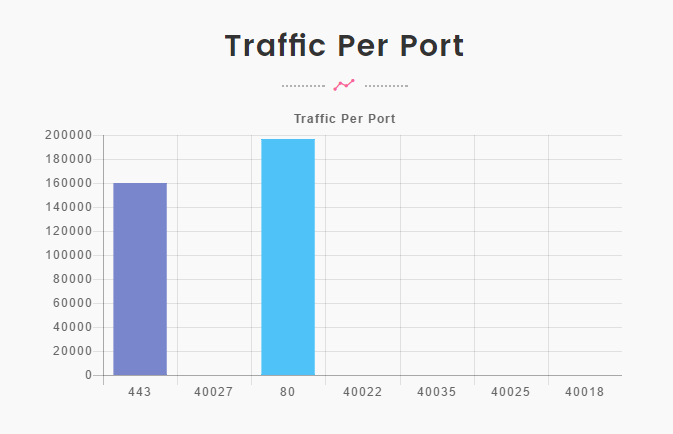
**בונוס - סטטיסטיקות מתקדמות:**

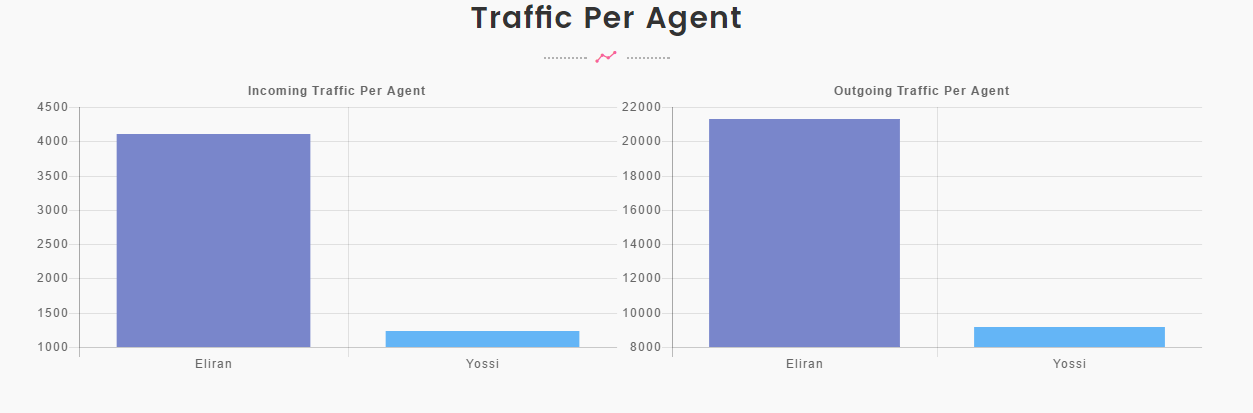
צרו עוד סטטיסטיקה נוספת שמעניינת לדעתכם. תוכלו להשתמש בנתונים הקיימים או להוסיף נתונים שיוסנפו בסוכן. תוכלו להשתמש גם בסוגים אחרים של גרפים, דוגמאות [באתר הזה.](http://tobiasahlin.com/blog/chartjs-charts-to-get-you-started/)

**איך אמורים להיראות הדו"חות:**

* IP נפוצים ביותר בקרב כל המשתמשים
* מדינות נפוצות ביותר בקרב כל המשתמשים
* תוכנות נפוצות בקרב כל המשתמשים



* פורטים נפוצים בקרב כל המשתמשים
* תעבורה נכנסת פר משתמש + תעבורה יוצאת פר משתמש



* רשימת גישות לא מאושרת לכתובות Blacklisted

