

# מציאה ואפיון של מסלולי טיול

הערה: התבקשנו להעלות PDF, ולכן לא צירפנו את הקוד שלנו (יצרנו קוד שמבצע scraping למסלולי טיול ותוכנה שמורידה את קבצי ה-gpx בהתנתן אזור גאוגרפי ומציירת אותם על מפה) או את המידע שחילצנו, אך נוכל לצרפם לפי בקשה.

## חברי הפרויקט

(1)

- (a) שם: מתן פנקס
- (b) מייל: [matan.pinkas@mail.huji.ac.il](mailto:matan.pinkas@mail.huji.ac.il)
- (c) לוגין: matan.pinkas

(2)

- (a) שם: נוי שטרנליכט
- (b) מייל: [noy.sternlicht@mail.huji.ac.il](mailto:noy.sternlicht@mail.huji.ac.il)
- (c) לוגין: noystl

(3)

- (a) שם: בר אלוני
- (b) מייל: [bar.aloni@mail.huji.ac.il](mailto:bar.aloni@mail.huji.ac.il)
- (c) לוגין: baraloni

## תיאור הבעיה

בפרויקט זה אנחנו רוצים למצוא מסלולי טיול, מוכרים ושאינם מוכרים, ולאפיין אותם. לצורך המטלה נשתמש במאגר הנתונים של openstreetmap בו מופיעים **נתיבי GPS** שתועדו על ידי משתמשים שונים. כל נתיב מתואר כאוסף נקודות שבו כל נקודה מיוצגת על ידי ערך קו גובה וערך קו אורך. לפי הגדרת הפרטיות שהגדירו משתמשים שתרמו את המידע, נוספים הפרטים: הזמן בעת המדידה, וגובה מעל פני הים.

נשתמש בנתונים אלו ובמאגר של מפות קיימות כדי להגדיר מסלול (לבחור האם להגדיר קבוצה של נתיבי GPS כמסלול טיול) ולאפיין את המסלולים שמצאנו על פי הקריטריונים הבאים: זמן, מרחק, רמת קושי, ונקודות עניין.

על מנת להוסיף מידע למסלולים שמצאנו, וכאמצעי להערכת הצלחתנו, ניעזר במידע על מסלולים מוכרים שחילצנו מאתרי טיולים. על ידי מציאת הדמיון בין מסלולים אלו למסלולים החדשים שנמצא, אנו מאמינים שנוכל לאפיין טוב יותר את המסלולים שאינם מוכרים מבחינת רמת הקושי של המסלול, זמן ממוצע לסיום המסלול, נקודות עניין וכדומה.

## מקורות מידע

(4) מסלולי GPS-

- (a) מה הדאטה: openstreetmap gps traces
- (b) תיאור: (גודל, מה מכיל...): מסלולים שאנשים הקליטו בgps, ותרמו ל-openStreetMaps. כל מסלול מורכב מנקודות המצויינות על ידי קו אורך ורוחב על הגלובוס (ובהתאם להגדרות הפרטיות של המשתמשים, גם זמן וגובה). בנוגע לגודל, ה-tar שמכיל את כל המסלולים שמשמשים על כדור הארץ העלו שוקל 21 ג'יגה:  
<https://planet.openstreetmap.org/gps>

(c) מאין השגנו: בהינתן איזור גאוגרפי, ה-API של openStreetMaps מאפשר לקבל מהאתר שלהם קובץ GPX המכיל 500 נקודות GPS שמשתמשים תרמו באותו איזור. לכן, כדי להשיג את המידע הדרוש לנו עבור איזור כלשהו, בינו תוכנה שבהינתן ריבוע התוחם את האיזור, מורידה מספר קבצי GPX כנ"ל כרצוננו.

(5) אתרי טיולים -

(a) מה הדאטה: HikingProject path data

(b) תיאור: מידע על: מסלולי טיולים ברחבי העולם הנאגרים ומשמשים להמלצות (ישנם מסלולים בודדים בישראל אך אלפי מסלולים ברחבי העולם) כל מסלול שכזה כולל מידע הכולל פירוט gps של המסלול (כולל מידע על גובה מעל פני הים), אורך מסלול, דרגת קושי ופרטי מידע נוספים (לדוגמה האם מותר לקחת כלבים), מבחינת גודל אין לנו הערכה מדויקת אך בהינתן שמסלול gps אחד שוקל בערך 30kb ויש לפי האתר 225 אלף מסלולים מדובר בסדר גודל של 6.8gb

(c) מאין השגנו: ביצענו scraping לאתר הטיולים <https://www.hikingproject.com> שמשמש בין השאר להמלצות למסלולי טיול.

(6) מפות קיימות -

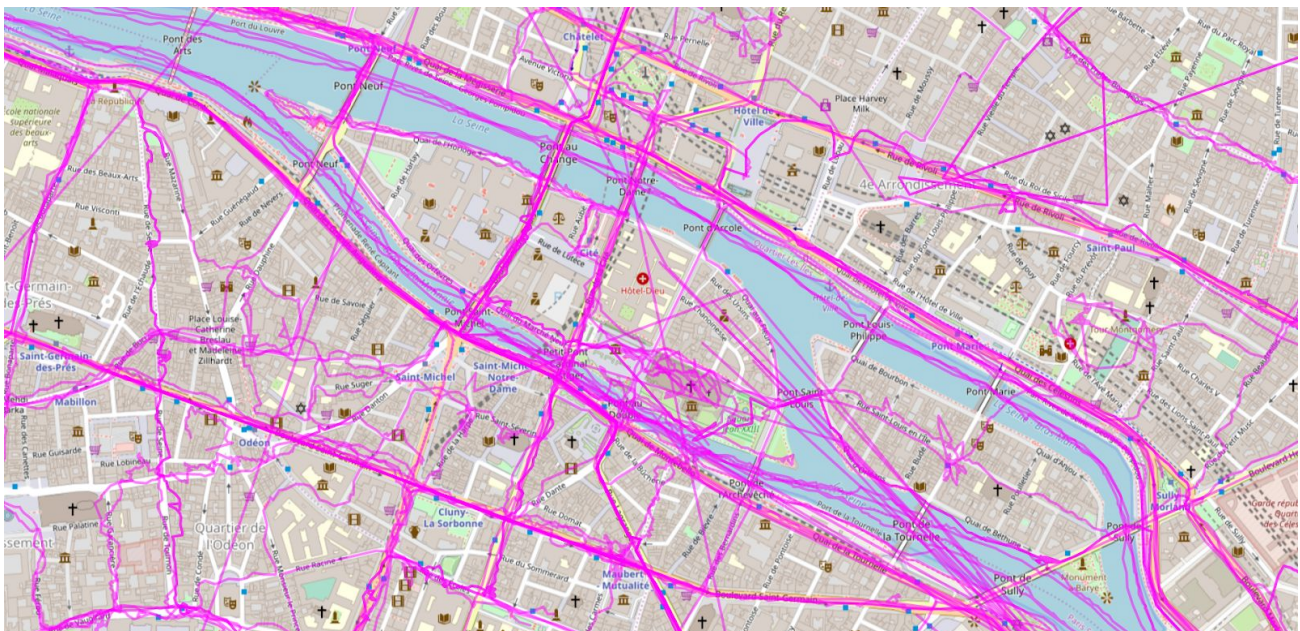
(a) מה הדאטה: מפות

(b) תיאור: המפות מכילות מידע גאוגרפי רב, כגון מוקדי עניין, מסלולים, ערים וכדומה.

(c) מאין השגנו: <https://download.geofabrik.de>

## ויזואליזציות

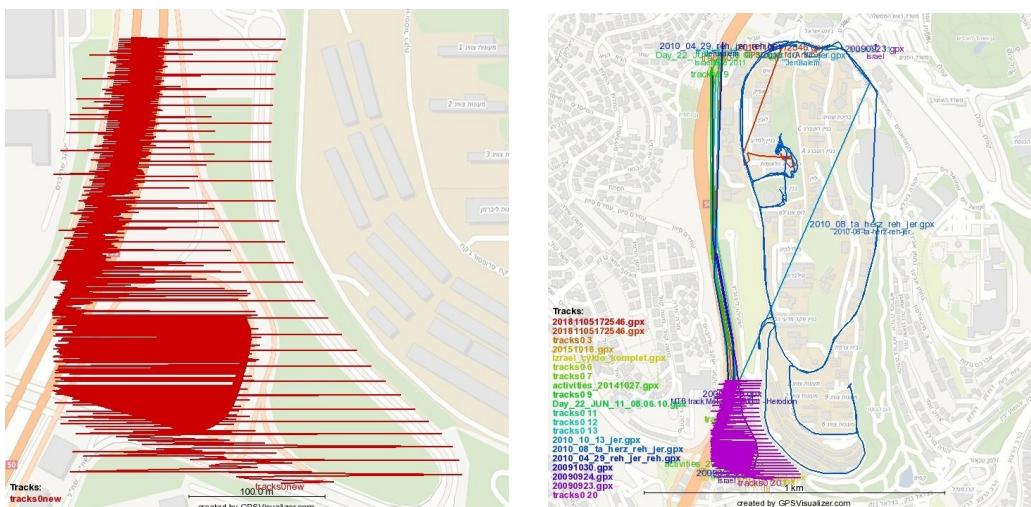
### ויזואליזציה ראשונה: רחובות פריז



על מנת ליצור את הויזואליזציה הנ"ל כתבנו סקריפט פייתון שמוריד מ-OpenStreetMaps עשרת-אלפים נקודות GPX שנתרמו על ידי משתמשים באיזור מרכז פריז. כאן, במסגרת בדיקת השפיות נצפה שהרוב המוחלט של המסלולים יתאימו לתוואי השטח (הליכה ברחובות, שהייה באזורים תיירותיים וכו') ואכן מהתבוננות בויזואליזציה נראה שכך המצב: מרבית המסלולים מתאימים לרחובות או דרכים במפה (גם המסלולים על נהר הסן תואמים את מה שנצפה ממסלולי שייט במציאות) ובכירות או במקומות תיירותיים נתן לראות "קשקושים" שמצביעים על כך שמי שהלך שם השתהה והסתובב במקום.



**הערה:** בעת עיבוד הדטה הראשוני הבחנו בהתנהגות מעניינת:



לפי מידע שמצאנו ברשת על הגדרות הפרטיות של משתמשים התורמים את המסלולים הנ"ל ([https://wiki.openstreetmap.org/wiki/Visibility\\_of\\_GPS\\_traces](https://wiki.openstreetmap.org/wiki/Visibility_of_GPS_traces)) כאשר מסלול מתוייג כ"פרטי" נקודותיו מוצגות בסדר שאינו כרונולוגי, ולכן הצגה של אוסף נקודות כזה כמסלול נראית לא הגיונית. לפיכך, בעת הצגה ויזואלית בחרנו לא להחשיב מסלולים שכאלה, אך ייתכן שבבואנו לאפיין מסלולי טיול נעשה בהם שימוש כדטה נוסף.

## ויזואליזציה שניה: שעות הטיול ביער בולון

יער בולון בפריז הוא פארק ציבורי שבשעות היום משמש את התושבים לרכיבה על אופניים, ספורט, פיקניקים וכדומה. בשעות הלילה יש בו פעילות לא חוקית ולרוב מזהירים תיירים לא להסתובב שם בחושך. עניין אותנו להבין אם באמת יש פחות ביקורים בפארק בשעות הלילה ( 20:00 - 5:00), ולכן כתבנו סקריפט פייתון שמוריד מ- OpenStreetMaps כ-4500 נקודות GPS שנתרמו על ידי משתמשים באזור הנ"ל, וצבענו מסלולים שהתרחשו בשעות היום **באדום** ומסלולים שהתרחשו בשעות הלילה **בכחול**. ניתן לראות בויזואליזציה שאכן מרבית המסלולים שתועדו בתוך הפארק נערכו במהלך היום, בשונה מאזורים ליד הפארק (שם כבר יש עדות לתזוזה בשעות הלילה).

