**האקתון דאטה 2024 – הנחיות**

**רקע:**

בעיר מנדלייביה, מתגוררים תושבים בעלי אוריינטציה מדעית ומודעות אקולוגית אשר מרבים לנסוע מנקודה לנקודה בעיר.

התושבים מעוניינים בפיתוח מערכת שתמליץ להם על נתיב מיטבי להגעה מנקודת מוצא לנקודת יעד. תושבת שתרצה להגיע מנקודת מוצא לנקודת יעד, תזין את שתי הנקודות למערכת, וזו תמליץ לה על נתיב אופטימלי בהתאם להעדפותיה (ראו הסבר בהמשך).

בעיר ישנן 96 נקודות, שכל אחת מהן יכולה לשמש כמוצא, כיעד, או כנקודת מעבר לאורך הנתיב המקשר בין המוצא ליעד. המעבר בין כל זוג נקודות יכול להתבצע באחד מהאופנים הבאים: א. בהליכה (בשני הכיוונים, ללא הגבלה), ב. ברכב פרטי (רק במידה וכיוון הנסיעה מאפשר זאת), ו-ג. בתחבורה ציבורית (רק בין נקודות בהן קיימת אפשרות כזו). כדי להגיע מנקודת מוצא לנקודת יעד ניתן לייצר נתיב שמשלב את כל האופנים המתוארים.

נסיעה בכלי תחבורה (רכב פרטי או תחבורה ציבורית) מביאה לפליטה של מזהמים (תחמוצות חנקן, תחמוצות גופרית או חומר חלקיקי), זאת בניגוד להליכה ברגל (שהיא גם בריאה יותר). ההנחה שלפנינו היא שנסיעה ברכב פרטי מביאה לפליטה של 5 יחידות מזהמים על כל 100 שניות של נסיעה, ואילו נסיעה בתחבורה ציבורית מביאה לפליטה של יחידת מזהמים אחת על כל 100 שניות. הסיבה שפליטת מזהמים בתחבורה ציבורית נמוכה יותר נובעת משני טעמים: א. פליטות הגזים מתחלקות על יותר נוסעים, ב. כל התחבורה הציבורית בעיר מבוססת על אנרגיות מתחדשות (אשר ייצורן לא מביא לפליטה של מזהמים לאוויר).

**מטרה:**

ההתניידות בין שתי נקודות (מוצא ויעד) מאופיינת במשך הזמן הנדרש, ובזיהום האוויר הנוצר. לתושבים השונים יש שיקולים אישיים שונים בבחינת החלופות הנ"ל, חלקם מוכנים להאריך את משך ההתניידות על מנת להקטין את זיהום האוויר, וחלקם פחות. מטרתנו היא פיתוח מערכת שאליה יזין התושב את הנתונים הבאים, והמערכת תחשב עבורו את המסלול המיטבי התואם את העדפותיו כפי שהן נתונות בסעיף 4 למעלה:

1. נקודת מוצא
2. נקודת יעד
3. תאריך וזמן היציאה
4. האם ברצונו של התושב למזער את פליטות המזהמים בנסיעתו
5. האם ברצונו של התושב למזער את זמן ההתניידות בין שתי הנקודות

**דאטה:**

לפניכם תיעוד של זמני הליכה ונסיעה (ברכב פרטי) כפי שנמדדו בזמנים שונים לאורך שנים, הכולל:

* תיעוד של זמן תחילת המדידה
* נקודת התחלה
* נקודת סיום
* משך הזמן שחלף בשניות (את זה יתכן שאחליף בזמן הסיום של המעבר בין הנקודות)

הערה: בכל המדידות, נקודות ההתחלה והסיום הינן נקודות סמוכות בעיר, שאין ביניהן נקודות נוספות.

בנוסף, נתון כי יש תחבורה ציבורית אשר:

* עוברת בין הנקודות הבאות ע"פ סדר רישומן למטה
* משך הנסיעה בין כל 2 נקודות סמוכות הוא בדיוק 4 דקות
* כלי תחבורה ציבורית נוסף יוצא למסלול כל 10 דקות
* מיד עם הגיעו לתחנת הסיום (Plutonium) מסתובב כלי התחבורה הציבורית ומתחיל במסלול חזרה עד לנקודה הראשונה (Beryllium)

1: Beryllium

2: Neon

3: Sulfur

4: Titanium

5: Nickel

6: Selenium

7: Zirconium

8: Palladium

9: Tellurium

10: Cerium

11: Gadolinium

12: Ytterbium

13: Osmium

14: Lead

15: Radium

16: Plutonium

**רמז:**

העיר מנדלייביה נמצאת בצפון אמריקה.

**איך תוכלו לזכות בנקודות?**

הציון על משימת ההאקתון הוא מודולרי ומשלב מספר אלמנטים. ככל שתפתחו יותר אלמנטים כך תוכלו לקבל יותר נקודות:

|  |  |  |  |
| --- | --- | --- | --- |
| אלמנט | המודל / האלמנט שיפותח | מס' נקודות מקסימלי | הערות |
| 1 | מודל חיזוי להולכי רגל בלבד | 30 |  |
| 2 | מודל חיזוי לכלי רכב פרטיים בלבד | 30 |  |
| 3 | מודל המשלב הולכי רגל + כלי רכב פרטיים | 70 | לא כולל את כל אחד מהמודלים הקודמים |
| 4 | מודל המשלב את כל אפשרויות ההתניידות (הליכה, נסיעה ברכב פרטי, נסיעה בתחב"צ) | 85 |  |
| 5 | שקלול דרישות למיזעור המזהמים | 10 | עבור מודלים 3, 4 |
| 6 | פיתוח GUI | 10 |  |

**מה צריך להגיש ומתי? ומה השלבים הבאים?**

1. את הקוד המלא בתוספת תוצאות הרצתו על סדרת מבחן יש להגיש בתום ההאקתון (קרי, ביום חמישי, 30.5.24, בשעה 21:00)
2. לאחר ההאקתון יעמדו לרשותכם חמישה ימים בהם עליכם להכין:
   1. מצגת המסכמת את תהליך העבודה, את האלגוריתם, אילו אלמנטים פיתחתם וכיצד (עד 8 שקפים)
   2. סרטון קצר (עד 3 דקות) בו אתם מתארים את תהליך העבודה ואת האלגוריתם

את שני התוצרים הללו יש להעלות לתיבת ההגשה במודל עד לתאריך 4.6 בשעה 17:30

1. צוות השיפוט יעבור על תוצרי הצוותים (קוד, מצגת וסרטון)
2. במפגש המסכם (6.6 בשעה 1800 בזום) יוכרזו הזוכים במקומות הראשונים שגם יציגו את המצגת שהכינו.
3. והשלב החשוב ביותר: שלושת הצוותים שיזכו במקומות הראשונים יזכו בפרסים כספיים.

להזכירכם – קיימת נוכחות חובה לכל המפגשים (כולל המפגש המסכם). זהו תנאי לקבלת ציון בקורס.

**בהצלחה!**

כללי :

1. מדובר בשאלת ריגרסיה
2. המטרה היא : לזהות את הנתיב האופטימלי ביותר .
3. הנתיב האופטימלי הוא : קצר ביותר או מזהם ביותר.
4. אפשר שילוב.

נתונים :

1. ישנם 96 נקודות מוצא יעד
2. מתוכן 16 נקודות הן תחנות אוטובוס
3. אוטובוס- 1 יחידות מזהמים ל 100 שניות
4. אוטובוס - משך נסיעה בין תחנה לתחנה 4 דק ( כלומר בין 2 תחנות אוטבוס מזהם 2.4)
5. אוטובוס יוצא כל 10 דקות
6. הליכה – דו כיווני
7. רכב פרטי – רק במידה וכיון הנסיעה מאפשר
8. נסיעה ציבורית – רק ב 16 תחנות, המסלול הוא הלוך חזור
9. רכב פרטי 5 יחידות מזהמים ל 100 שניות

הנחות :

1. לבחור אם עובדים בשנייה אחת או 100 שניות
2. ברכב - אם הנסעה לא מתועדת, אי אפשר לעשות את המסלול ישיר

שאלות :

1. ~~איך נדע מה הכיוון הנסיעה שמאפשר רכב?~~
2. ~~האם שהמודל מאתחל את המטריצה , הוא מאתחל את כל הקשתות~~

משימות :

1. ליצור מילון לתחנות אוטובוס
2. ~~לפצל את העמודה תאריך~~
3. ~~לאחר מכן להפוך את העמודה תאריך ל DateTime~~
4. פיציר של יום בשבוע
5. פיצר של חודש
6. פיציר של עונה (?)
7. פיציר של חגים
8. משתנה דמי של שעות , קפיצות של שעה
9. חלוקה לשל יום ל 4 קבוצות

מסקנות :

1. סך הכול יש 318 מסלולים ישירים של הולכי רגל ומכוניות שאפשריים.
2. בנוסף יש עוד 22 מסלולים שרק הולכי רגל יכולים ללכת בהם.
3. פיצול הנתונים ל-2 מודלים: להולכי רגל ולכלי רכב פרטיים.
4. בדיקה של כל צירופי המסלולים האפשריים כדי לזהות את הקשתות האפשריות ואת כיוונן.
5. יצירה pipeline המכיל את השלבים הבאים :
   1. פיצול עמודת הDate ל-2 עמודות (תאריך ושעת יציאה), ויצירת משתני דמי.
   2. יצירת מודל למידת מכונה שמקבל תאריך ושעה ובודק את כל הצירופים על מנת למצוא את משקולות הקשתות לאלגוריתם דייקסטרה.
   3. יצירה אלגוריתם דייקסטרה על מנת למצוא את המסלול המהיר ביותר מנק' המוצא של המשתמש עד לנקודת היעד.
   4. יצירת פונקציה המריצה את קובץ ה"טסט", מפעילה את כל הפונקציות ב-pipeline ומחזירה את הקובץ עם המסלולים המהירים ביותר וזמני הנסיעה.