

	<p>לפניך רשת של סטודנטים אשר נפגשים על בסיס יומימי בהרצאות השונות באוניברסיטה (כל צומת מייצגת סטודנט, קשת מקשרת בין הסטודנטים שיכולים להיפגש).</p> <p>ביום בהיר אחד, אחד הסטודנטים נדבק בוירוס.</p> <p>הניחו כי מספיק להימצא באותו חדר עם הסטודנט המודבק על מנת שהוירוס יעבור, וסטודנט מתחיל להדביק בוירוס יום אחרי שהוא נדבק בעצמו.</p> <p>א. מיהו הסטודנט אשר יביא למספר מודבקים הגדול ביותר תוך יום אחד לפי הרשת הנתונה.</p> <p>ב. האם יש סטודנט אשר לא יוכל להביא להדבקה כלל?</p>
	<p>כעת נניח כי לגרף יש כיווניות, כלומר הדבקה ע"י סטודנט א ל-ב' יכולה לקרות אך ורק אם קיימת קשת מכוונת היוצאת מסטודנט א אל סטודנט ב</p> <p>ג. כיצד תשתנה תשובתך לסעיף א?</p> <p>ד. נניח כי הסטודנט שנדבק הוא אכן הסטודנט שבחרתם בסעיף א.</p> <p>עברו יומיים והסטודנט פנה לעזרה רפואית. לתרופה נגד הוירוס לוקח יומיים כדי למנוע הדבקה עתידית, האם יש סטודנט שלא יודבק כלל?</p>

### שאלה פתוחה 3

סעיף א. סטודנט מספר 5 כיוון שלו הקישוריות הגבוהה ביותר.

סעיף ב. לא. כיוון שהגרף כולול מקושר.

סעיף ג. פוטנציאל ההדבקה הגבוה ביותר קיים אצל שני סטודנטים – מספר 4 ומספר 5.

סעיף ד. לא, כל הסטודנטים ידבקו. לפי סעיף א' הסטודנט שנדבק ופנה לעזרה רפואית הוא מספר 5. עקב כך שקיים מסלול מסטודנט 5 לכל סטודנט אחר, וקצב ההדבקה עולה על קצב ה-"חיסון" על ידי התרופה (יום מול יומיים) שגם כך הסטודנט קיבל טיפול בה רק לאחר יומיים.

### פתוחה 4

בסעיפים הבאים מתוארות תכונות שונות המאפיינות data של פרופיל ביטוי גנים במקבץ של דוגמאות. אופי הנתונים כפי שמתואר יכול להיות בעיה עבור אלגוריתם k-means. עבור כל תכונה, תאר כיצד תשפיע התכונה על האלגוריתם והצע פתרון אפשרי שיאפשר לאלגוריתם להיות עמיד לבעייתיות זו.

א. קלסטור של דוגמאות לפי פרופיל ביטוי הגנים שלהם בהנתן שכאשר מבצעים קלסטור של גנים לפי פרופיל הביטוי שלהם בדוגמאות ניתן בברור לראות כי הם מחולקים לתתי-קבוצות של גנים בעלות גודל מאוד שונה. כאשר הגנים המאפיינים כל קבוצה מראים קורלציה גבוהה ביותר בינם לבין עצמם.

השפעה: מספר הגנים הגדול מתת-קבוצה אחת יגרום להטיה חזקה של דמיון בין דוגמאות על סך גנים אלו תוך כדי "פיספוס" עושר הדוגמא.

פתרון: כיוון שגנים בכל תת-קבוצה בעלי קורלציה גבוהה, בחירה של "גנים ייצוגיים" מכל תת-קבוצה וקליסטור דוגמאות לפיהם יוכל לשמר על עושר הדוגמא תוך כדי אובדן מידע מינימלי.

ב. קלסטור של דוגמאות לפי פרופיל ביטוי גנים השונים זה מזה באופן משמעותי בטווח הביטוי

השפעה: הגנים בעלי טווח הביטוי הרחב ישיפעו באופן חזק מאוד על הקליסטור של הדוגמאות, למרות שטווח ביטוי לאו דווקא מצביע על חשיבות ביולוגית.

פתרון: נרמול של הגנים לטווח דומה (לדוגמא להתפלגות של ממוצע אפס, וסטיית תקן אחת)  
יתן משקל דומה לכל גן, אם כי, יש להזהר בהתייחסות לגנים ששונתם בין דוגמאות היא  
מינימלית ונובעת מרעש טכני.