

Project proposal

Team 7

1. Introduction

פרויקט הסיכום שלנו בקורס "מבוא לניתוח נתונים" יתמוך במחקר המתבצע בפקולטה לרפואה על ידי עמית עקיבא(סטודנט בתואר שני לסיעוד ופרמדיק במד"א) בהנחיית ד"ר אודיה כהן(ראש המחלקה לסיעוד באב"ג) וד"ר דגן שוורץ. המתנת אמבולנסים במלר"ד היא תופעה מוכרת המת"יחסת למצב שבו מטופל אשר מגיע למלר"ד עם אמבולנד לא עובר תהליך קליטה ומיון כרגיל אלא נשאר על מיטת האמבולנס מסיבות שונות הקשורות בזמניות המלר"ד למטופל. המשמעות של המתנת אמבולנסים במלר"ד היא בין היתר חוסר זמניות האמבולנסים לקראת מקרי חירום נוספים אליהם הם נדרשים. היקף התופעה נחקר רבות מצד בתי החולים ומחקרים רבים בודקים ומחפשים דרכים ושיטות לייעול תהליכים בקבלת המטופלים אך מעטים המחקרים המאפיינים את התופעה מצד שירותי המערכת רפואה דחופה. מטרת המחקר העיקרית הינה לאפיין את היקף ופרטי תופעת עיכוב המטופלים המובאים למלר"ד(מחלקה לרפואה דחופה - מיון) על ידי צוותי האמבולנס השונים של מד"א. אפיון התופעה יתבצע בין היתר על ידי "מטרות משנה" שהגדיר החוקר המוביל את המחקר.

1. לאפיין את התופעה מבחינת מיקום גיאוגרפי של הפינוי ובית החולים הקולט.

2. לתאר את הקשר בין משך ההמתנה לבין עונות השנה, ימי השבוע ושעת ההגעה למלר"ד.

3. לאפיין את המטופלים אשר מתעכבים יותר מהמוצע עד קבלת מיטה במלר"ד מבחינת גיל המטופל, דחיפות הפינוי כפי שנקבעה על ידי מד"א והעברת או אי-העברת מידע מקדים למלר"ד ומאפיינים רפואיים של המקרה.

מטרות המשנה יתמכו כחלק מפרויקט הסיכום שלנו בקורס "מבוא לניתוח נתונים". הועבר לצוות ממסד נתונים גדול על התופעה זאת על מנת להשתמש ביכולות שלמדנו בתואר בכלל ובקורס "מבוא לניתוח נתונים" בפרט כדי לסייע באפיון התופעה ובמציאת פתרון לבעיה.

2. Data

סוג המחקר: המחקר המוצע הינו מחקר שנערך בשיתוף עם מד"א. אוכלוסיית המחקר: מטופלים אשר פונו לחדר מיון באמצעות אמבולנסים של מד"א. קריטריונים להכללה: כל המטופלים אשר הופנו לחדר מיון באמצעות אמבולנסים של מד"א בשנת 2022. קריטריונים לאי-הכללה: אמבולנסים אשר פינו שני מטופלים ומעלה באותו הזמן, מטופלים אשר הוכנסו ישירות לחדר ההלם/טראומה, מטופלים אשר הגיעו לחדר מיון באמצעות מסוק, מטופלים מתחת לגיל 18, וילדות, מטופלים אשר פונו למלר"ד ללא אמבולנס של מד"א. שיטת דגימה: החובשים מחויבים למלא בכל אירוע טופס עם נתונים אלו ע"י טאבלט שנמצא ברכב האמבולנס. גודל המדגם: 436,778 תצפיות.

מספר עמודה בטבלה המפורט (מימין לשמאל)	שם המשתנה	סוג המשתנה	ערכים+הסבר
1.	דחיפות	סודר	דחוף/לא דחוף/בידוד טיפול נמרץ -מידת הדחיפות של מצבו הבריאותי של המטופל בבידוד.
2.	תאריך	רציף	תאריך בתבנית: xx/xx/xxxx -תאריך הפינוי
3.	עיר	שמי	ערים בישראל -הגדרת האזור התבצעה על בסיס הגדרות הלמ"ס. העיר בו התבצע פינוי המטופל אל הבידוד.
4.	גיל	רציף	18 ומעלה -גיל המטופל
5.	שם רכב	שמי	שם הרכב נותן מידע על המרחב אליו הוא שייך בחלוקה המרחבית של מד"א.
6.	יעד פינוי	שמי	בית החולים בו מתקבל החולה.
7.	קוד רפואי צוות	סודר	לפי מערכות מד"א. החולה מקבל אבחון ראשוני על סידת הטיפול. מחולק לקטגוריות.
8.	מועד הגעה ליעד פינוי	רציף	תאריך ושעה בתבנית: xx/xx/xxxx xxxxx -הגעת המטופל לבידוד.
9.	מועד הכרזה כפנוי	רציף	תאריך ושעה בתבנית: xx/xx/xxxx xxxxx -לאחר מסירת המטופל.

מימדי הטבלה- 9 עמודות ו436,777 שורות. מספר ערכים חסרים בטבלה היא 3 לעמודת הגיל ו4041 לעמודת העיר.

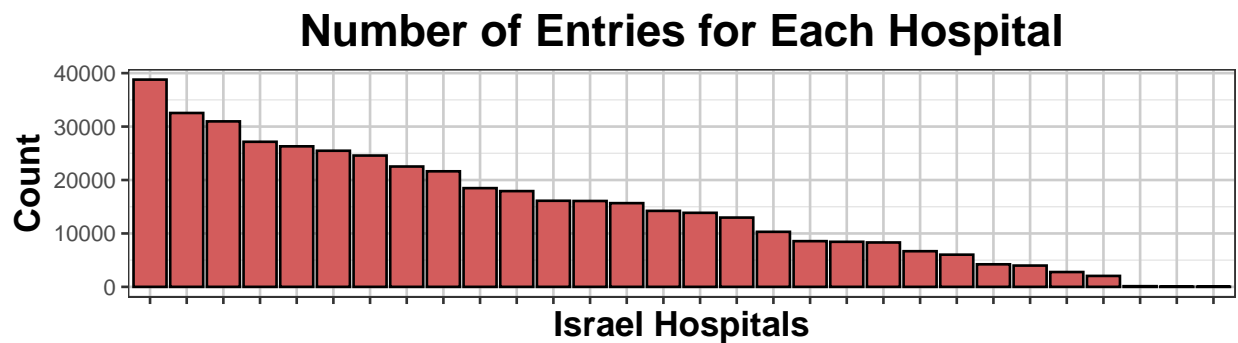
3. Preliminary results

בחלק הזה של הצעת מחקר נבצע ניתוח נתונים חקרני ראשוני כדי לוודא שהדאטא שקיבלנו מתאימה למחקר ומסגרת זמן של הקורס כפי שראינו מקודם, דאטא בעל 9 עמודות(משתנים, מאפיינים), ו 436,777 שורות. לפי דעתינו, ממבט ראשוני על הבעיה, כמות נתונים כזאת יכולה להספיק כדי להשיג תוצאות שהוצבו במחקר. אין טעם לבצע פעולת סיכום על הדאטאפריים כי רוב הערכים בה קטגוריאליים אז נתרכז בעמודת הגיל:

גיל מקסימלי בדאטאסט 221720 .
הערך הזה חריג מאוד ויכול לסמל לנו על טעות בהזנת הנתונים או באג במערכת.

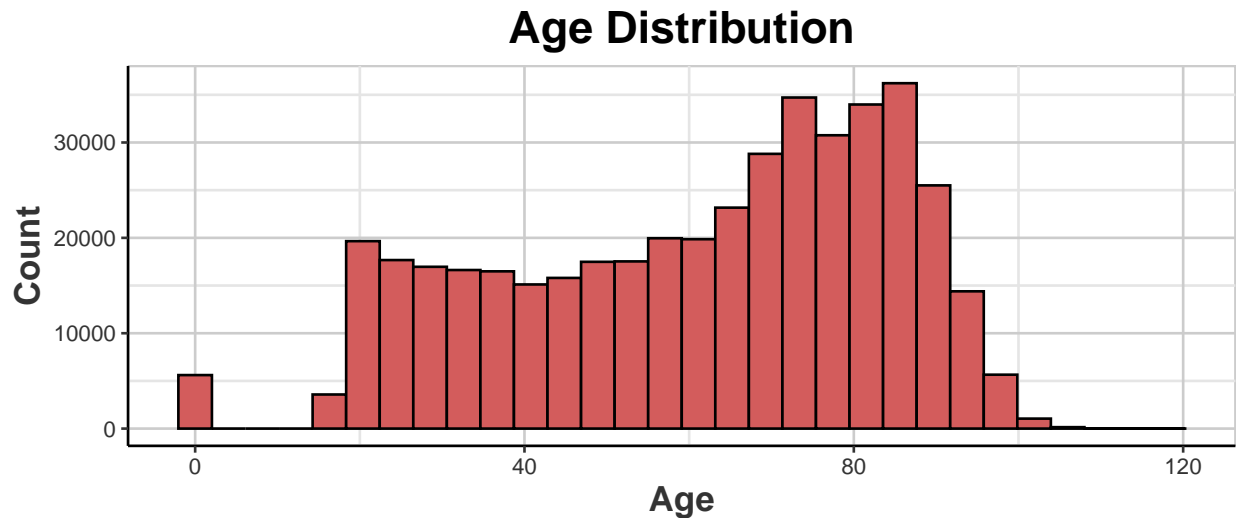
כעת נתבונן בכמה ויזאליזציות חשובות שיעזרו לנו להוכיח כשירות הנתונים למחקר

גרף של בתי חולים



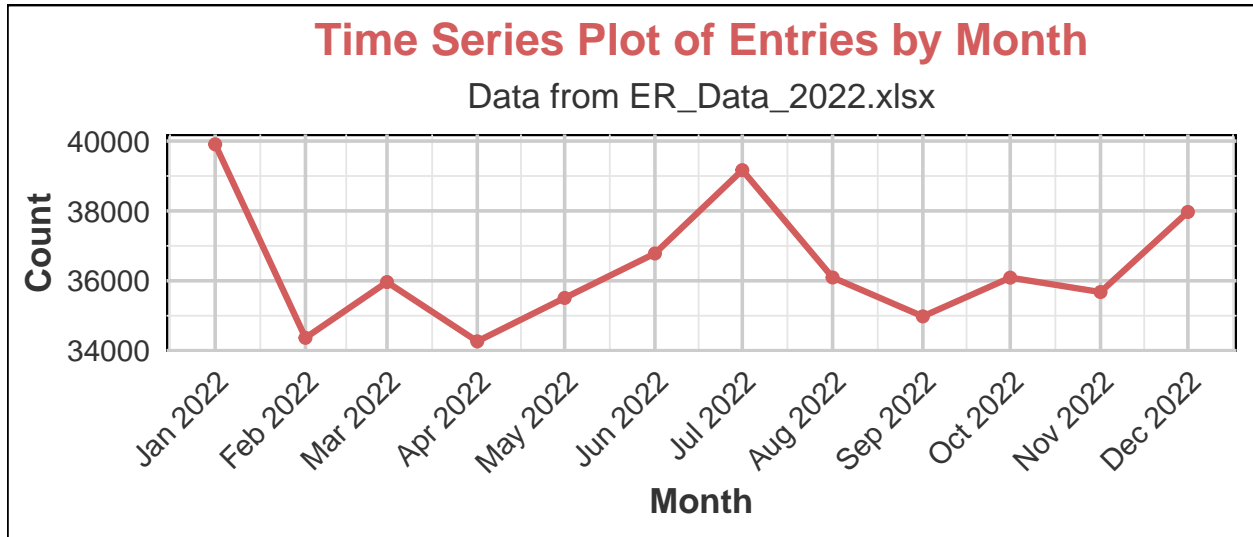
שלושת בתי החולים המובילים(איכילוב, שיבא וסורוקה) מציגים מספר רשומות גבוה במיוחד בהשוואה לאחרים. נוסף על כך, יש פה מקום להרבה סטטיסטיקות מעניינות(כמו איזה בית חולים שירת הכי הרבה מטופלי והאם זה משפיע על זמני המתנה).

גרף התפלגות הגילאים



כתוצאה מהגרף אנו יכולים לראות שקיימת התפלגות יחסית אחידה (גיל הממוצע גבוה יחסית) אך זה הגיוני בנוגע לקריאות אמבולנסים שאנשים בגיל מבוגר יותר ישתמשו בשירות הזה יותר. לכן לא קיימת הטייה כלשהי מבחנת גיל במדגם שלנו ואפשר לנתח את הנתונים בלי חשש לנתונים לא מאוזנים והטיית.

תדירות הקריאות כל חודש



מסקנות מהגרף יש שונות עונתיות ניכרת במספר הרשומות בכל חודש. ישנן פסגות בינואר ואוגוסט, עם ירידות משמעותיות בפברואר ובספטמבר. זהו עניין שכדאי לחקור ולנסות לזהות דפוסים מסויימים בין חודשי שנה. יש עלייה יציבה במספר הרשומות מאפריל עד אוגוסט. העלייה באמצע השנה יכולה להיות קשורה לגורמים עונתיים ספציפיים או לאירועים שמתרחשים במהלך תקופה זו. בהמשך המחקר, כדאי לוודא שהמגמות הנצפות אינן נובעות מטעויות בהזנת הנתונים או אי-עקביות בתהליך איסוף הנתונים.

מכסה הערים המחקר

מספר ערים שונות שהשתתפו במחקר: 2544

מספר קודי רפואי הצוות שהופיע במחקר: 130

כפי שאנו יכולים לראות המחקר מכסה הרבה מאוד ערים בארץ אז המדגם הוא גם חסר הטייה מבחנה גיאוגרפית ונתון לניתוח. כפי שניתן לראות ישנם 130 קודים רפואיים שונים שמופיעים במחקר והמטרה שלנו זה להבין אותם ווללמוד להשתמש בהם. לצורך המשימה הזאת צוות המחקר סיפק לנו קובץ עם פיענוח של קודי צוות הרפואי.

סיכום

בהתבסס על כל הממצאים הללו, ניתן לקבוע שמערך הנתונים מתאים לניתוח ולהשגת מטרות הקורס. הנתונים מייצגים בצורה אמינה את הפעילות לאורך השנה, מאפשרים זיהוי מגמות עונתיות ומספקים בסיס טוב לניתוח מעמיק יותר.

4. Data analysis plan

1. משתנים תלויים (Y) ומשתנים בלתי תלויים (X):

- משתנה תלוי (Y): זמן עיכוב בקבלת חולים לחדר המיון מהאמבולנס.
- משתנים בלתי תלויים (X): משתנים עיקריים כוללים קודי מצב רפואי, מיקום גיאוגרפי (עיר), ורמת הדחיפות של הטיפול.

2. קבוצות השוואה:

- מצבים רפואיים: השוואת מספר הביקורים בין מצבים רפואיים שונים לזיהוי המקרים הנפוצים או החמורים ביותר.
- מיקומים גיאוגרפיים: השוואת ביקורים בחדר המיון בין ערים שונות לזיהוי הבדלים אזוריים בצרכי הבריאות.

- רמות דחיפות: השוואה על בסיס רמת הדחיפות של הטיפול לזיהוי דפוסים בביקורים דחופים ולא דחופים.

3. שיטות למענה על השאלות:

- סטטיסטיקה תיאורית: לסיכום הנתונים ולמתן תובנות ראשוניות לגבי התפלגות הביקורים בחדר המיון.
- ניתוח סדרות זמן: לניתוח מגמות לאורך זמן וזיהוי דפוסים עונתיים בביקורים בחדר המיון.
- ניתוח רגרסיה: לקביעת הקשר בין המשתנים הבלתי תלויים (מצבים רפואיים, מיקומים גיאוגרפיים ורמות דחיפות) לבין המשתנה התלוי (מספר הביקורים החודשי).
- ANOVA: להשוואת הממוצעים של ביקורים בחדר המיון בין קבוצות שונות (גיל, מצבים רפואיים וכו') ובדיקת משמעות סטטיסטית.

4. תוצאות נדרשות לתמיכה בתשובה המושערת:

- מגמות משמעותיות: זיהוי מגמות משמעותיות של עלייה או ירידה בביקורים בחדר המיון לאורך זמן.
- קורלציה וסיבתיות: קביעת קורלציות משמעותיות בין המשתנים הבלתי תלויים למספר הביקורים, המציינים גורמים סיבתיים פוטנציאליים.
- הבדלים בין קבוצות: הצגת הבדלים משמעותיים סטטיסטית במספר הביקורים בין קבוצות ההשוואה (גיל, מצבים רפואיים, מיקומים, רמות דחיפות).

5. עבודה בצוות:

- חלוקת עבודה: העבודה תחולק באופן הבא להבטחת ניתוח מעמיק והתקדמות יעילה:
 - ניקוי והכנת נתונים: חבר צוות אחד יתמקד בניקוי והכנת הנתונים לניתוח.
 - ניתוח תיאורי וחקר ראשוני: חבר אחר יעסוק בחקר ראשוני של הנתונים והמחשתם לזיהוי דפוסים עיקריים ותובנות.
 - ניתוח סדרות זמן ורגרסיה: חבר שלישי יהיה אחראי על ביצוע ניתוח סדרות זמן ודגמי רגרסיה לחקר הקשרים בין המשתנים.
 - ANOVA והשוואות בין קבוצות: חבר רביעי ינהל מבחני ANOVA וניתוח השוואות בין קבוצות לקביעת הבדלים משמעותיים.
- פגישות שבועיות: כל ארבעת החברים ייפגשו אחת לשבוע להסביר את חלקם בעבודה ולהבטיח שכל הניתוחים מחוברים ומשולבים באופן קוהרנטי.

תחום הפריקט: עבור צוות בן ארבעה חברים, התחום הנוסף כולל ניתוח מעמיק יותר של פערים גיאוגרפיים בנגישות לבריאות ובחינה מפורטת של דפוסים עונתיים, שייטכן שלא יהיו אפשריים עבור צוות קטן יותר בשל עומס העבודה המוגבר והמורכבות.

Appendix

README Data