

למיזת מכונה ברפואה

ביתוח נטוונים רפואיים
והסקת מסקנות

מגייסים: דור אינברגר
אביה גمرا
אייתי בקנשטיין



מבנה

3

קשרים בין משתנים

2

הDATA ועיבוד נתונים
ראשוני

1

ספרות מחקרית

6

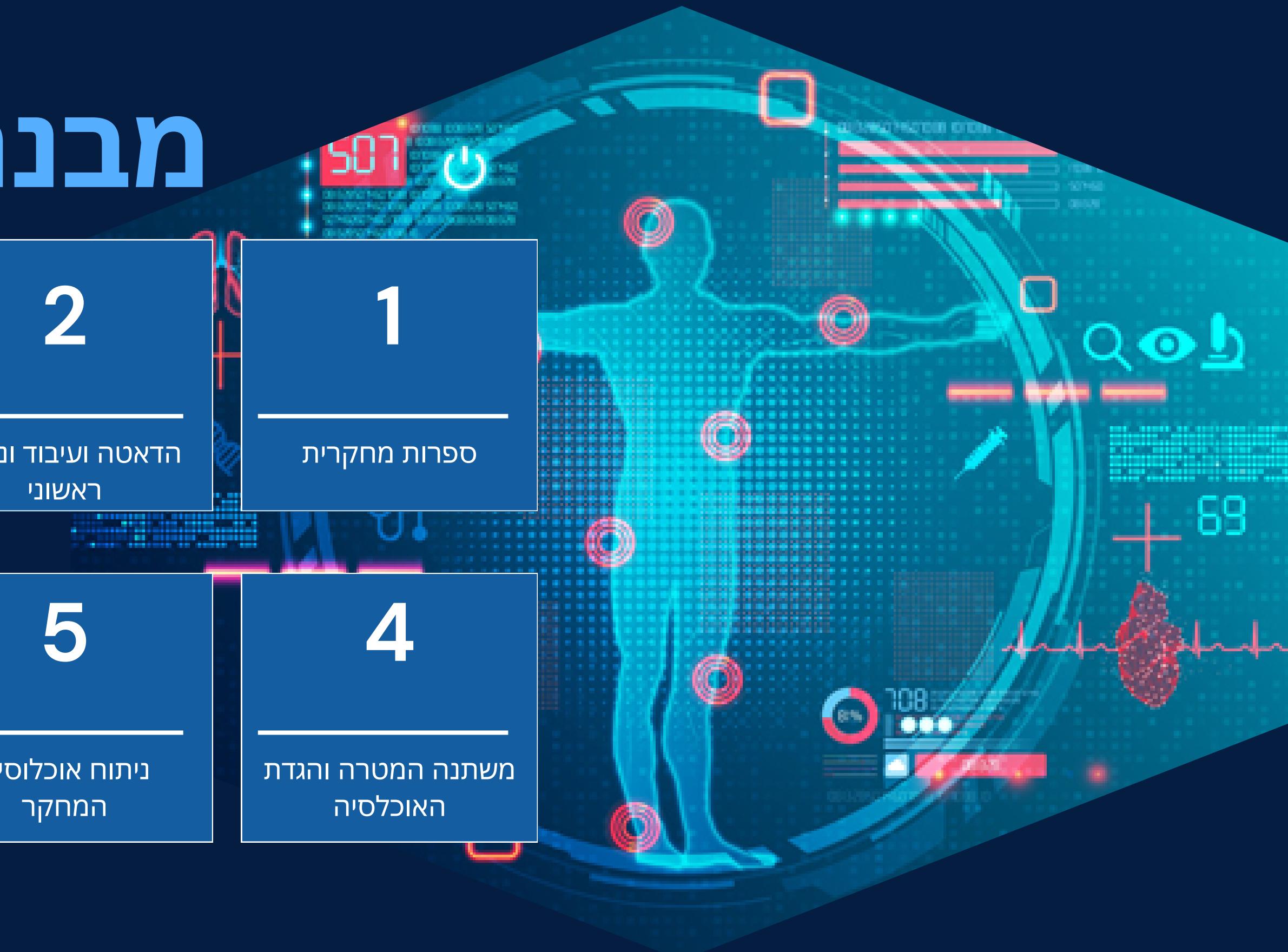
מסקנות

5

ניתוח אוכלסיות
המחקר

4

משתנה המטרה והגדרת
האוכלסיה



ממצאים עיקריים

חסיפה להחץ דם נמוך:

- חסיפה של 120 דקוט ומעלה קשורה לעלייה בסיכון למתותה וכשל איברים.

לחץ דם ממוצע (MAP):

- קשר שלילי בין MAP נמוך לבין תמותה בטיפול נמרץ.

לחץ דם סיסטולי (SBP) ודייסטולי (DBP):

- ערכים נמוכים של SBP ו-DBP קשורים לתוצאות שליליות כולל תמותה מוגברת.

מחקרים

 השפעת רכיבי לחץ דם על תחלואה ותמותה בחולי ספסיס בטיפול נמרץ

- בחינת הקשרים בין רכיבי לחץ דם (סיסטולי, דייסטולי, ממוצע ופולסי) לבין תוצאות קליניות בחולי ספסיס בטיפול נמרץ.
- מטרה: הערכת השפעתם על תמותה וכשל איברים.

השפעת רכיבי לחץ דם על תחלואה ותמותה בחולי ספסיס בטיפול נמרץ

תרומה לפרויקט

הבנת דפוסי לחץ דם קריטיים:

- **תובנות לשיפור ניתוח הקשרים בין רכיבי לחץ דם לתמותה.**

פיתוח אלגוריתמים:

- **שיפור מודלים לחיזוי תוצאות קליניות והטאמת טיפולים למניעת
תמותה.**

שיפור ניטור וניהול חולים:

- **יצירת פרוטוקולים לניטור לחץ דם ולניהול חול ספסיס.**

מסקנות:

- **ניטור דינמי:** שמירה על טווחי
לחץ דם קריטיים לשיפור תוצאות
החולים.
- **שמירה על רמות לחץ דם תקינות:**
mphיתה סיבוכים ותמותה בקרבת
חול ספסיס בטיפול נמרץ.

THE DATA



icustays

- **תועלות:** מספקת מידע על שהיית המטופלים ביחידה לטיפול נמרץ.
- **מידע:** מזהה מטופלים, מזהה אשפוז, תאריכי כניסה ויציאה מהיחידה לטיפול נמרץ.



patients

- **תועלות:** מכילה מידע דמוגרפי על המטופלים.
- **מידע:** מזהה מטופלים, תאריכי כניסה, מגדר, אתניות.



admissions

- **תועלות:** מכילה מידע על כניסה המטופלים לבית החולים.
- **מידע:** מזהה מטופלים, תאריכי כניסה ויציאה מבית החולים, סיבת אשפוז.

More Tables

lab_events

- תועלת: מכילה מידע על תוצאות בדיקות מעבדה.
- מידע: מזהה מטופלים, מזהה אשפוז, סוג בדיקות, תוצאות בדיקות.

diagnoses_icd:

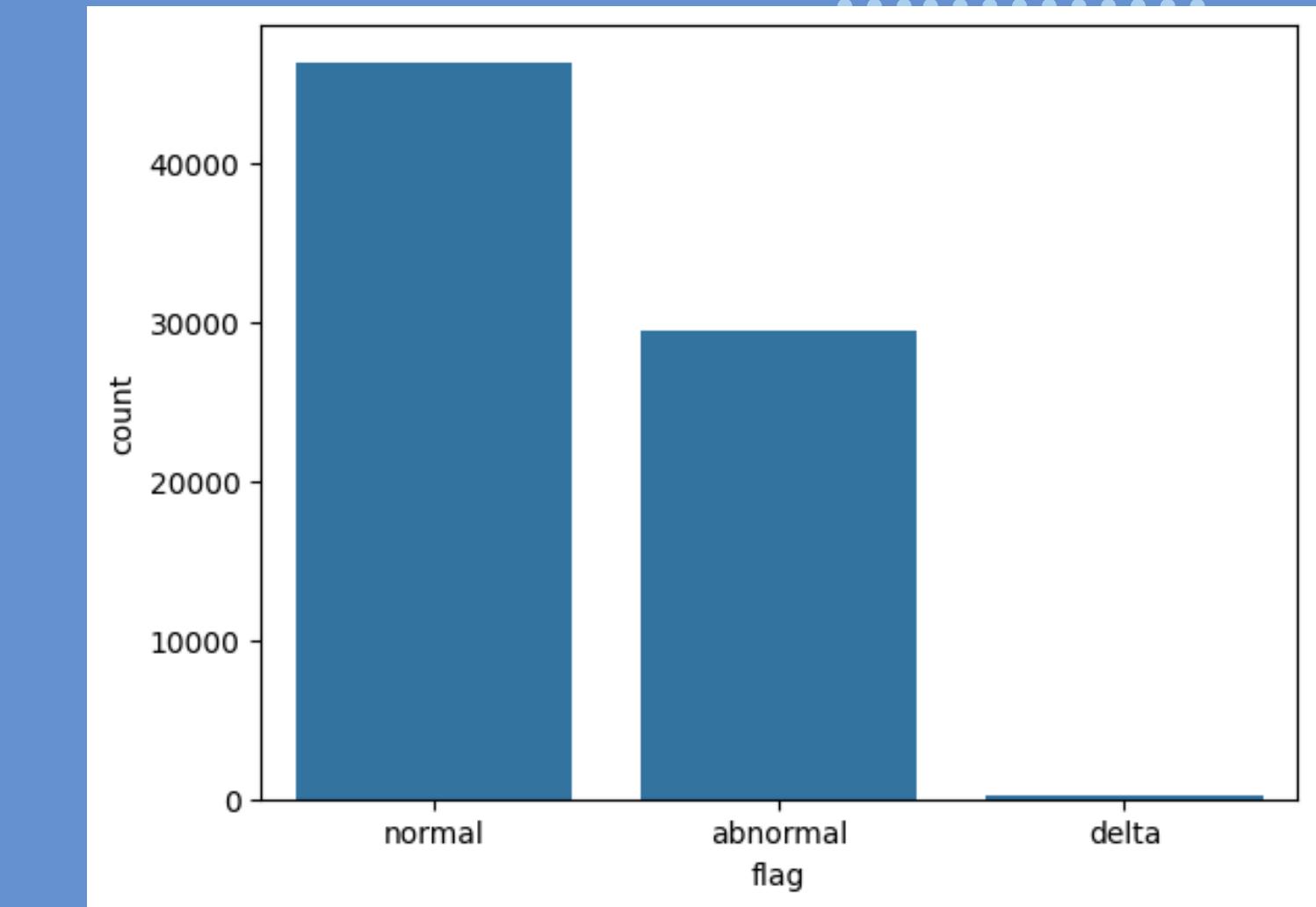
- תועלת: מכילה מידע על דיאגנוזות המטופלים.
- מידע: קוד ICD-9, תיאורי דיאגנוזות.



עיבוד ראשוני-ערכיים חסרים

בכדי לקבל נתונים ווילדים ותקינים פועלנו בכמה צעדים בכדי לצור את המשתנים בצורתם הנכונה והפייצ'ארים הבאים:

- סקירהת הערכאים החסרים בכלל הטבלאות והעמודות וביצוע אפיון לעמודות הרלוונטיות
- מילוי המילה 'normal' כאשר ערך הבדיקה הוא ריק ומהויה בדיקה תקינה ללא flag
- שימוש בהגדת הבדיקות כ"נורמליות" במילוי ערכאים חסרים ביחס **לממוצע** של אותה בדיקה המתויגת כנורמלית בכדי לנורמל את הנתונים לערכאים תקינים.
- קבלת ערכי בדיקות תקינים ומינורמליים ביחס לכל בדיקה תור יכולת לסוג את הבדיקות וליצור התפלגות מיצגית



הקוד

```
סינון ריק לערכאים נומריים #
labevents['value_numeric'] = pd.to_numeric(labevents['value'], errors='coerce')

# חישוב ממוצע של הבדיקות הנורמליות עבור כל סוג בדיקה
normal_means = labevents[labevents['flag'] == 'normal'].groupby('itemid')['value_numeric'].mean()

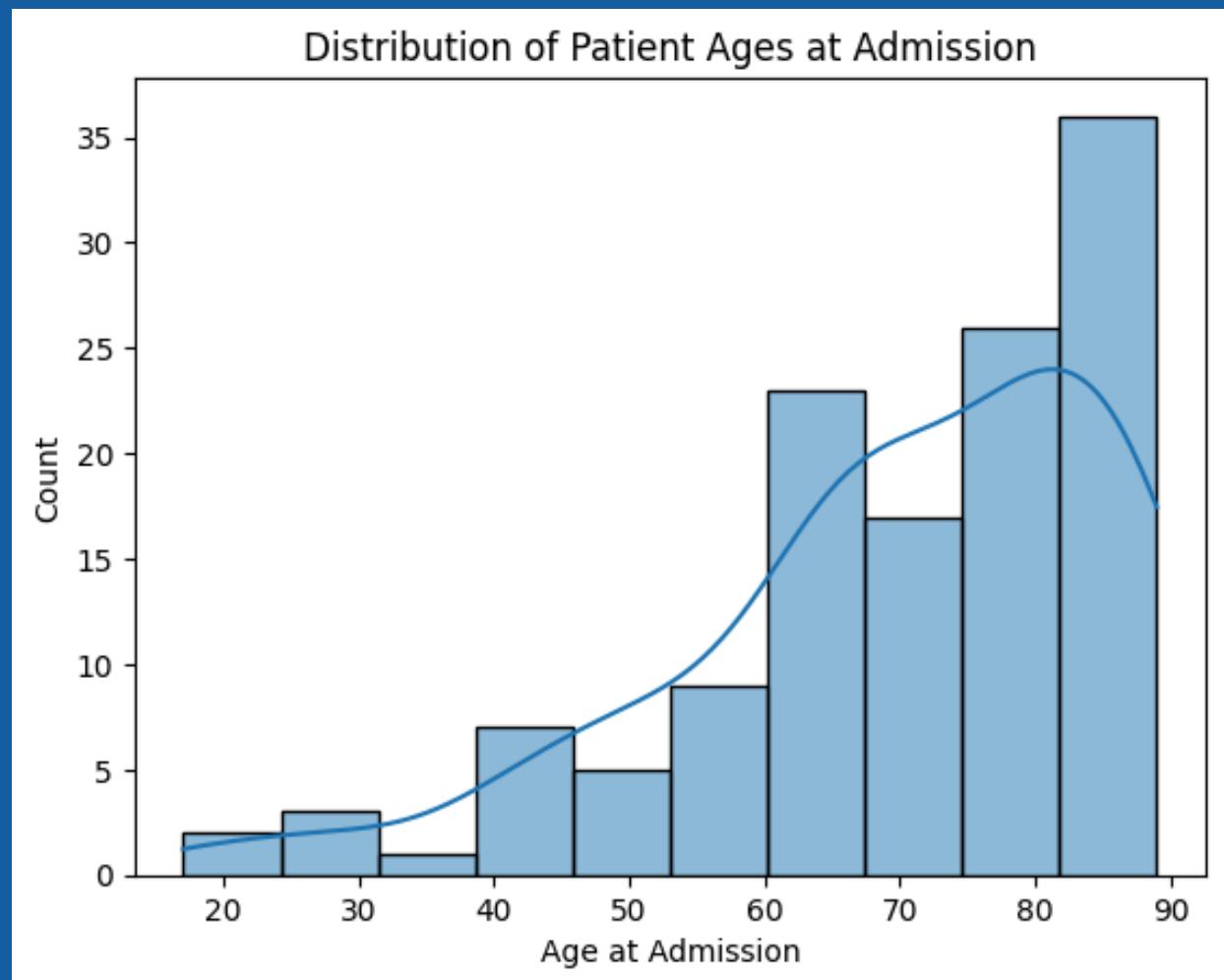
פונקציה להשלמת ערכים חסרים #
def fill_missing_values(row, normal_means):
    if pd.isnull(row['value']) and row['flag'] == 'normal':
        return normal_means.get(row['itemid'], row['value'])
    return row['value']

# כאשר הדגל הוא נורמלי משלמת הערכים החסרים בעמודות #
labevents['value'] = labevents.apply(lambda row: fill_missing_values(row, normal_means), axis=1)
```

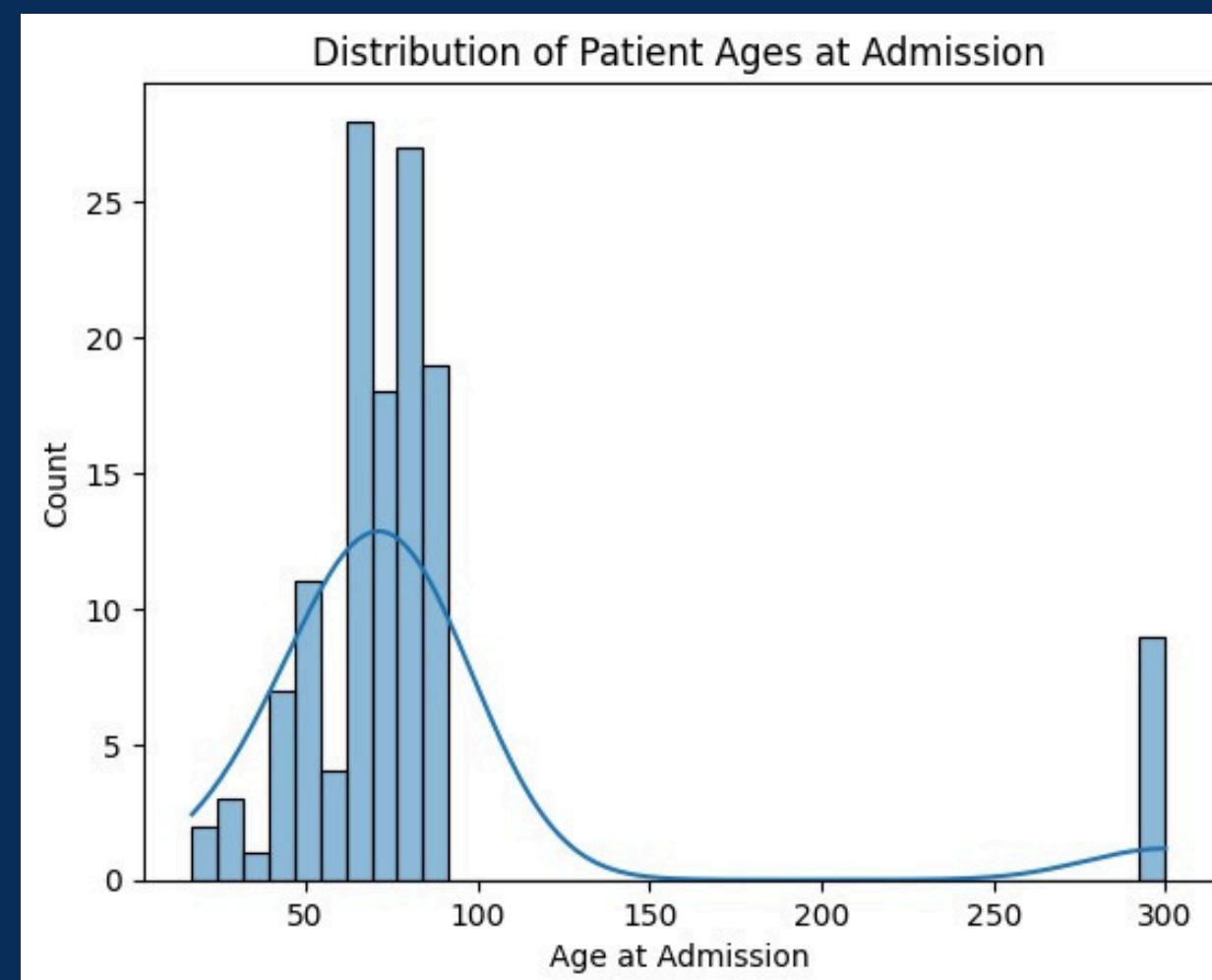
עיבוד ראשוני-חריגים

יצירת משתנה הגיל ללא חריגים (גיל "300" והמייצג מטופלים מבוגרים מאוד)

אחרי



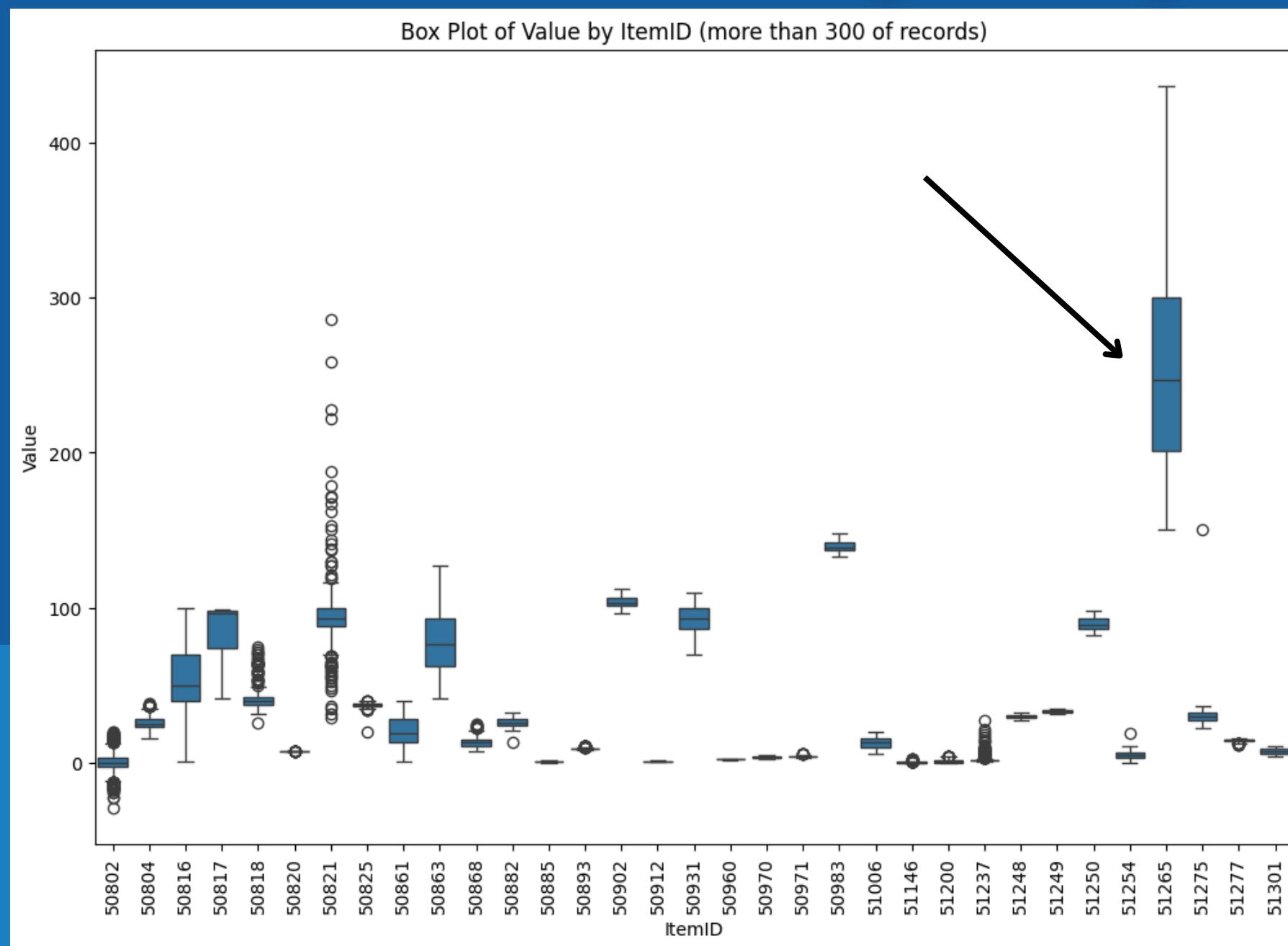
לפני



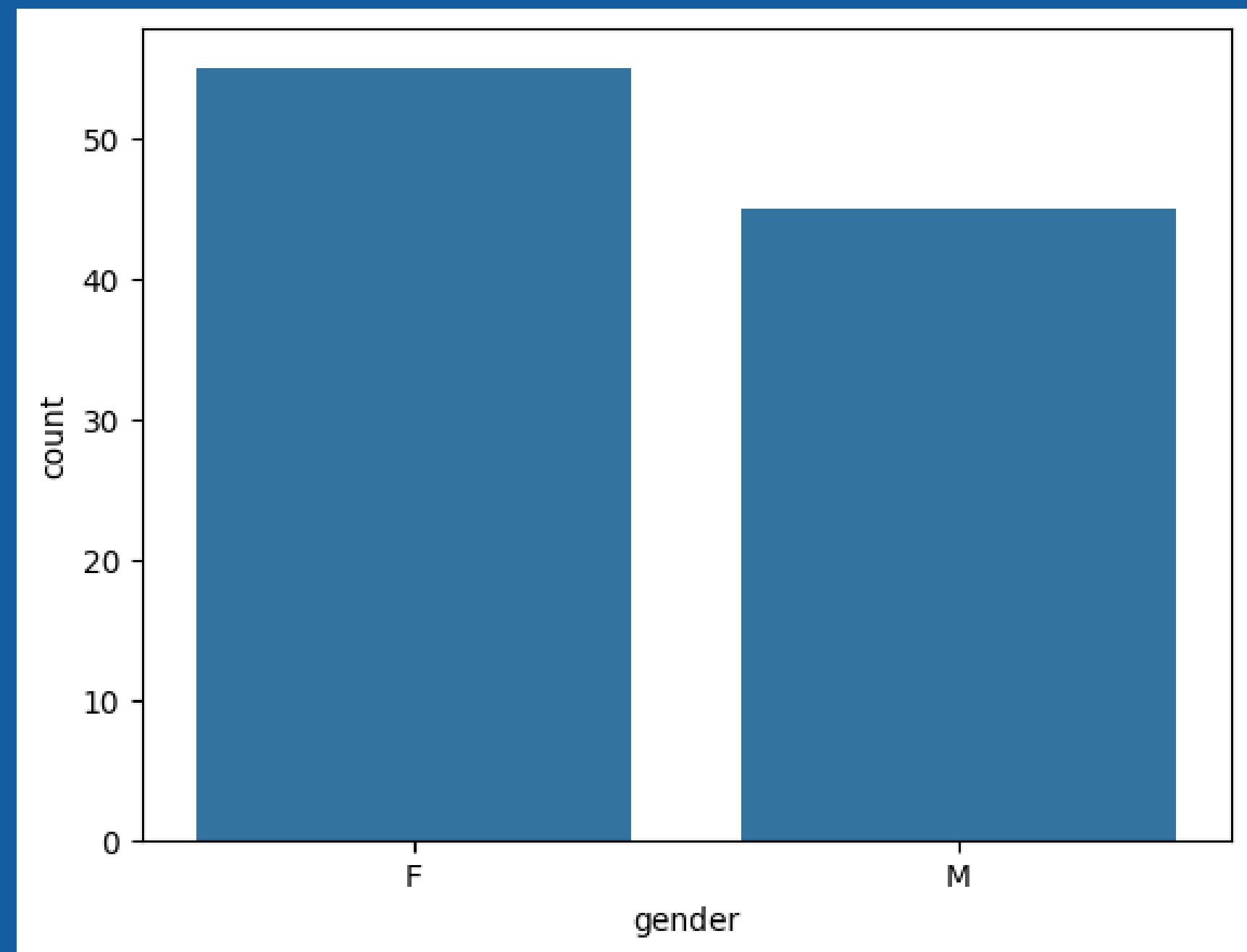
ניתוח נתונים ראשי



התפלגות ערכי הבדיקות המרכזית (TOP20) למציאת ערכי נורמה



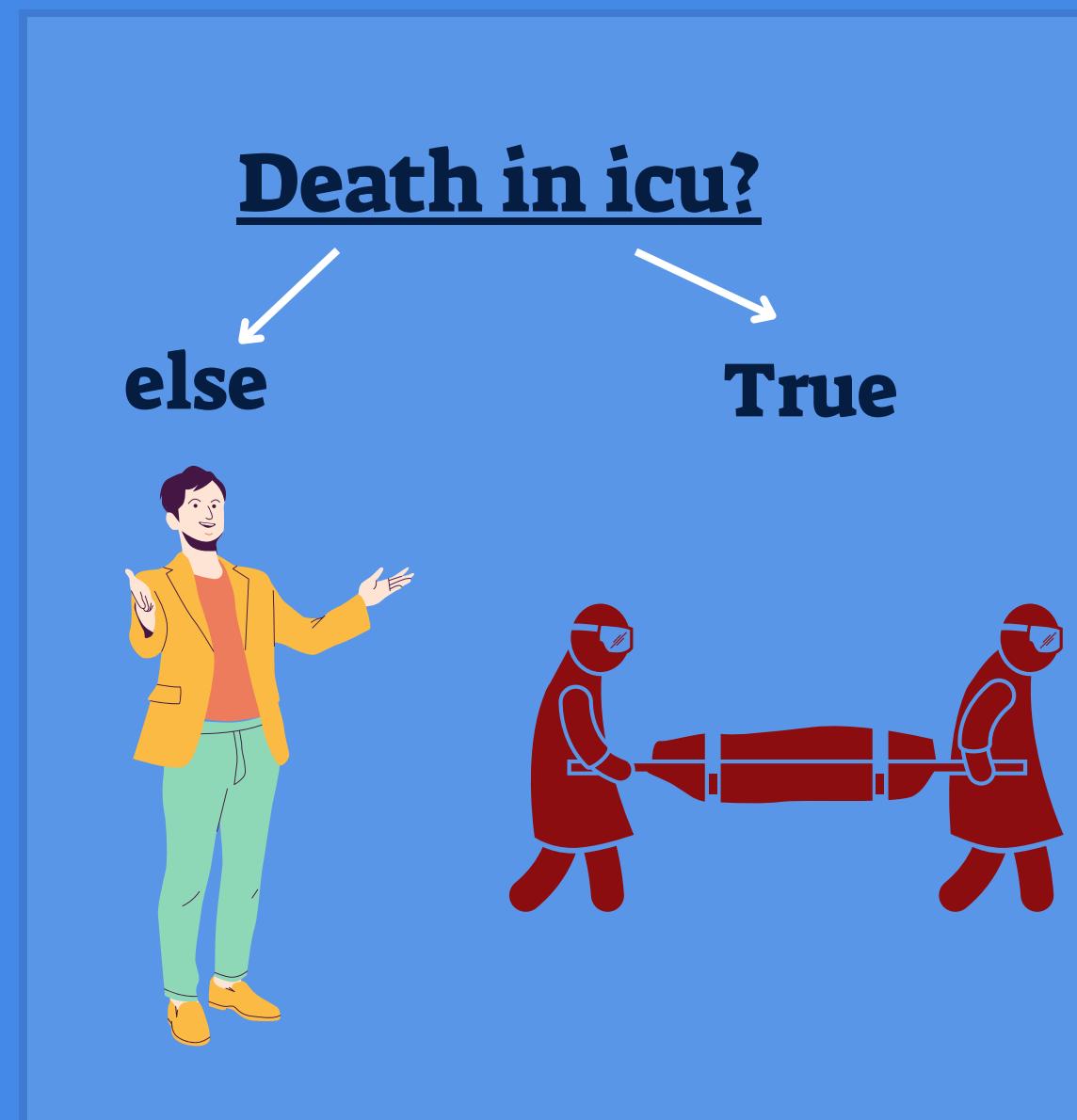
התפלגות גברים ונשים באשפוז כמשתנה דמוגרפי בסיסי



משתנה `death_in_icu`:

מוות בעת שהות בטיפול נמרץ

הקוד בודק השוואת בין זמן המוות (outtime) וזמן הבנייה (intime) לזמן הכניסה (deathtime) והיציאה (intime) לזמן המנוחה (deathtime) --> ומצביע על מטופל ממת לאחר כניסה לטיפול נמרץ.



IF המטופל מת לאחר כניסה לטיפול נמרץ --> ומצביע על **True**.

```
merged_data['deathtime'] >= merged_data['intime']
```

IF המטופל מת לפני יציאתו מהטיפול נמרץ, --> ומצביע על **True**

```
merged_data['deathtime'] <= merged_data['outtime']
```

• IF יש ערך בזמן המוות, --> ומצביע על **True**

```
merged_data['deathtime'].notna()
```

STATISTICS

בדיקות קשרים בין משתנים
ומובהקות סטטיסטיות

МОובהקות?



סקירה כל הקשרים האפשריים בין
דיאגנוזות שונות

95%



בדיקות הדיאגנוזות המרכזיות
הניידות ומציאת 2 המובילות
لتמונה בשיעור גובהה יותר
באשפוז ומובהקות סטטיסטיות

90%



בדיקות הקשר בין זמן השהות
באשפוז לבין מותם בזמן האשפוז

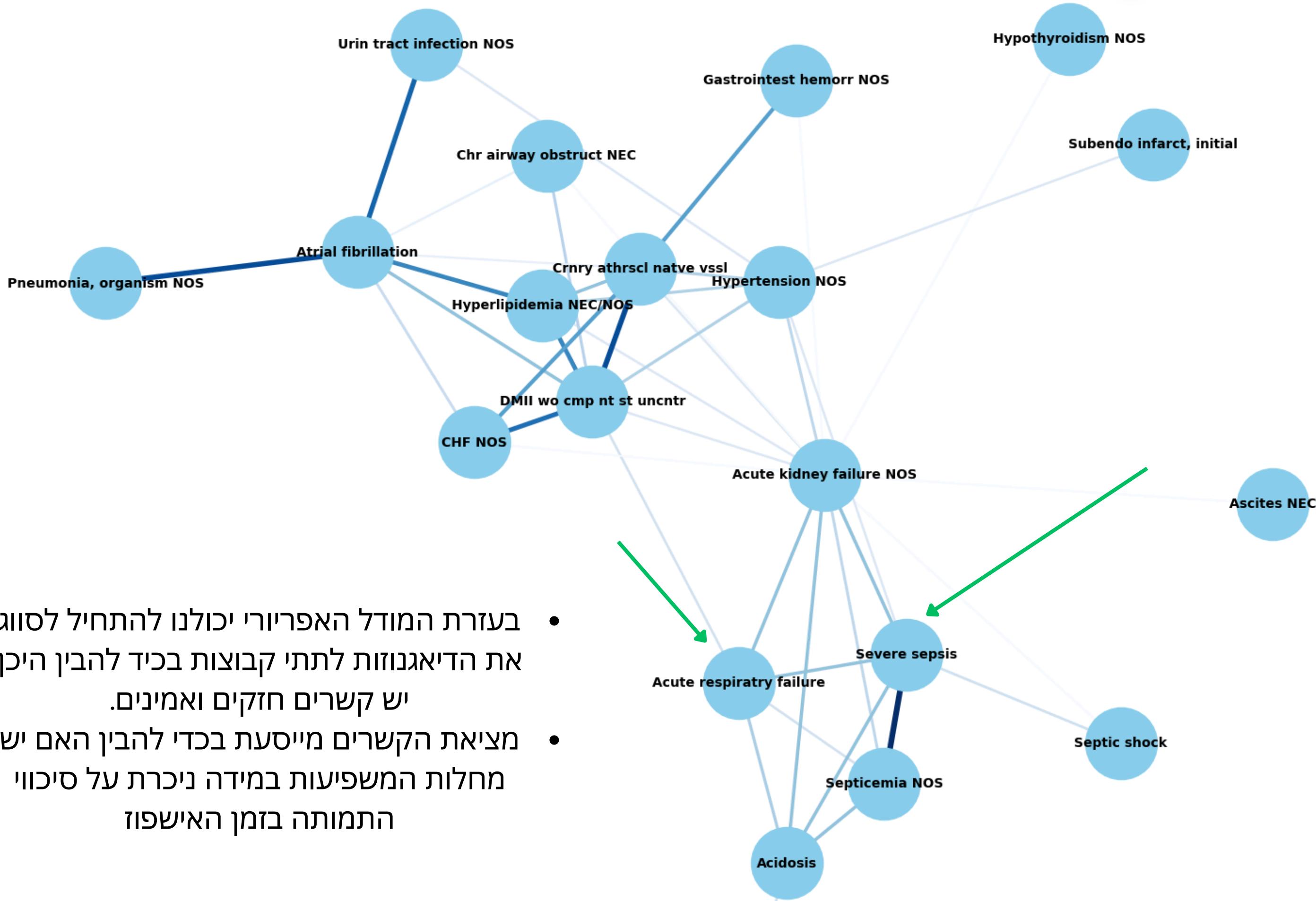
1

2

3



קשרים בין דיאגנוזות



2 מובילות

בחינת ה TOP5 של המחלות בהם חלו המטופלים שמתו במהלך האשפוז ומציאת 2 מתוכם המובהקים סטטיסטיות ביותר ב 95%-90% מובהקות

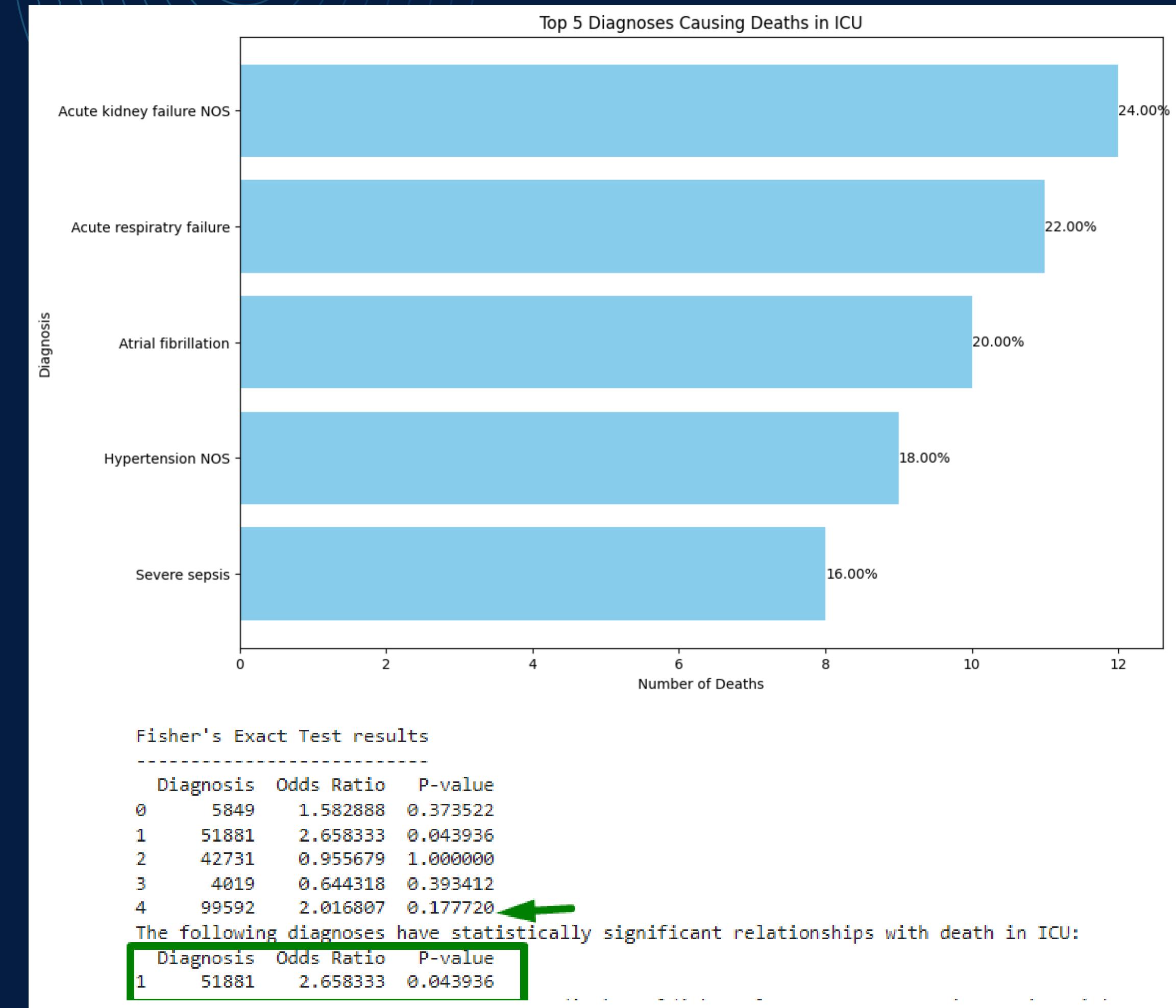
Acute respiratory failure &

"תסמונת מצוקה נשימית חריפה"

Severe sepsis 🍀

"אלח דם"

ממצאים אלו הובילו אותנו לשימוש בדיאגנוזות אלו בתור אובייקטיבת המחקר

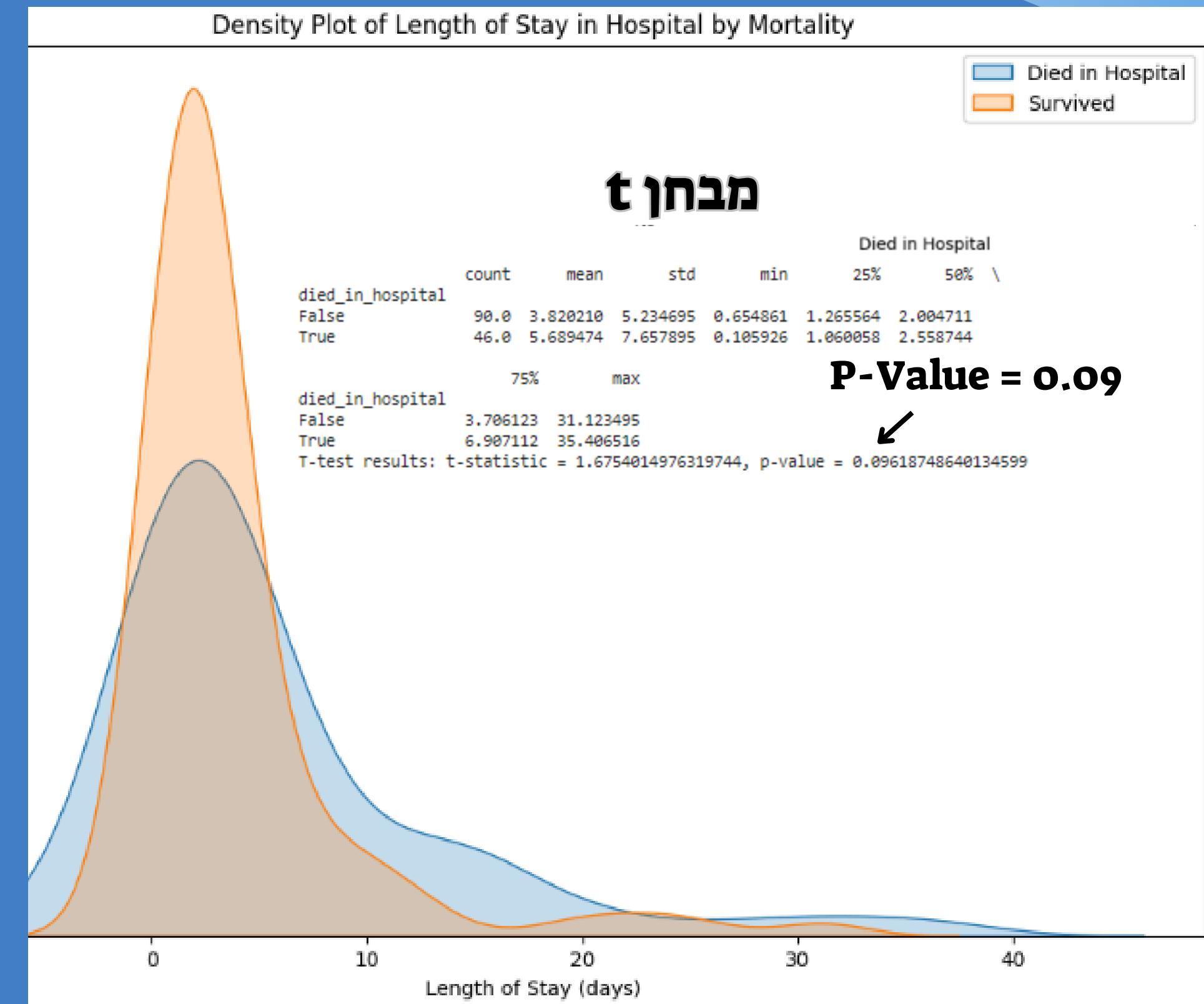


3

מאות בזמנן האשפוז

בדיקות הקשר בין זמן השהות באשפוז לבין

על פי הסתברות
בגרפים ו מבחון t
שともר בהם המסקנה
היא:
למשך השהות בטיפול
נمرץ יש השפעה
מוביקת בקרב האנשים
שנתו בטיפול נמרץ



אוכלוסית המטרה

חולים שמתו בטיפול נמרץ והיו עם המחלות

1 - תסמנת מצוקה נשימתית חריפה



2 - אלח דם Severe sepsis



אתגרים: מדגם קען המקשה על היבולת ללמידה תובנות מובהקות

מציאת משתנים דמוגרפיים

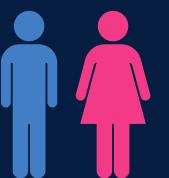


על מנת לבדוק את המשתנים הדמוגרפיים
יצרנו טבלה חדשה שמכילה מאפיינים
דמוגרפיים בקרוב לאנשים שהגיעו לטיפול נמרץ
עם הדיאגנוזות שבחרנו

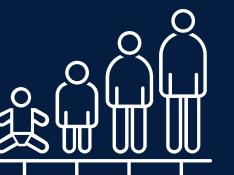
	subject_id	hadm_id	gender	age	ethnicity	insurance	icd9_code	death_in_icu
0	10019	177759	M	49	WHITE	Medicare	51881	1
1	10019	177759	M	49	WHITE	Medicare	99592	1
2	10029	132349	M	78	WHITE	Medicare	99592	0
3	10036	189483	F	89	WHITE	Medicare	99592	1
4	10040	157839	F	86	UNKNOWN/NOT SPECIFIED	Medicare	51881	0

מציאת משתנים דמוגרפיים

בשלב הבא, ניסינו לאתר מtower הubble שבחרכנו ניתנו את המשתנים הדמוגרפיים הבאים:



• מספר נשים וגברים ומתחם כמה מトー--> **Gender**



• התפלגות הגילאים ומתחם כמה מトー--> **Age**



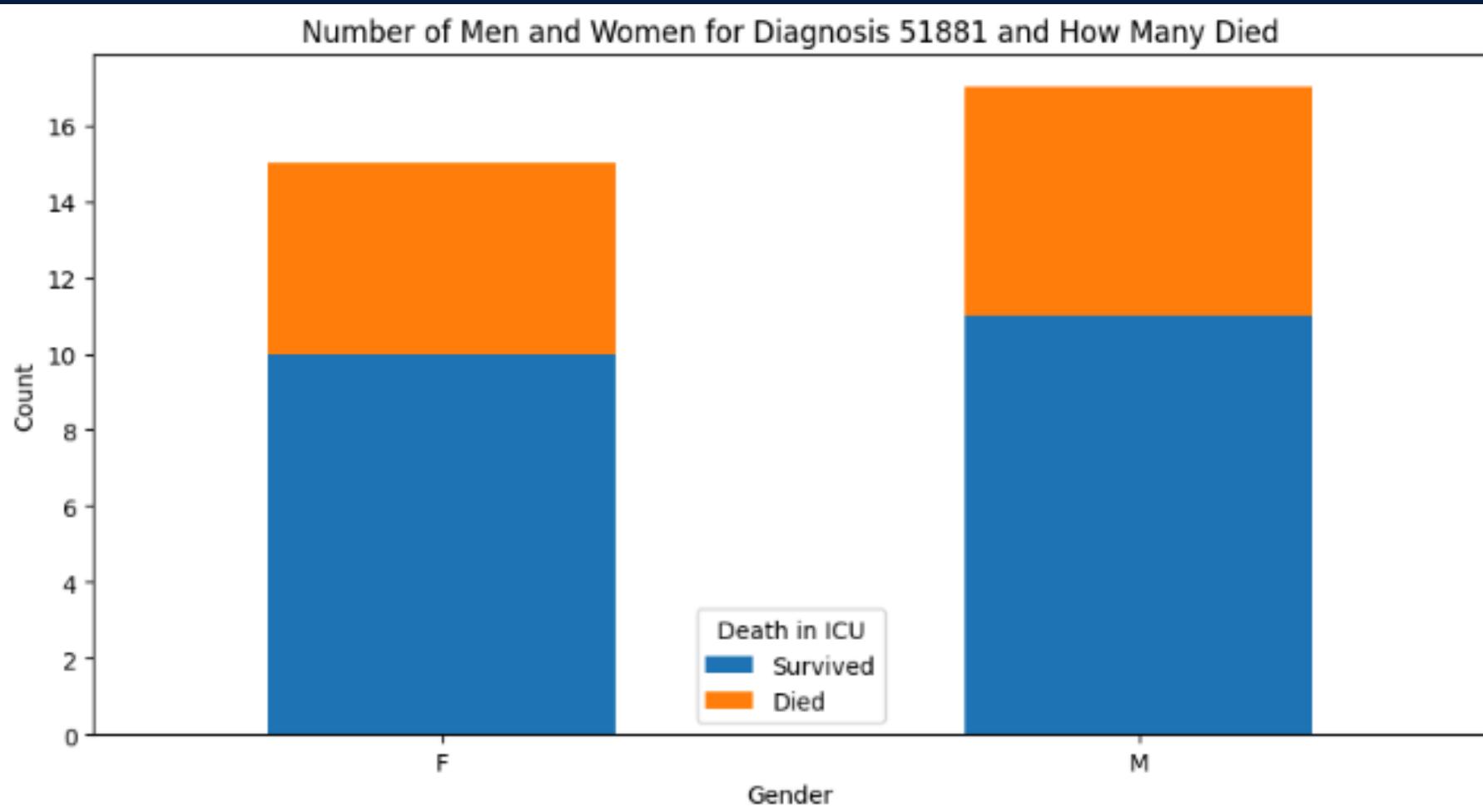
• סוגי הביטוחים (פרטי / ממשלת / סיעודי) ומתחם כמה מטו --> **Insurance**

1 - מספר נשים וגברים ומה מתו



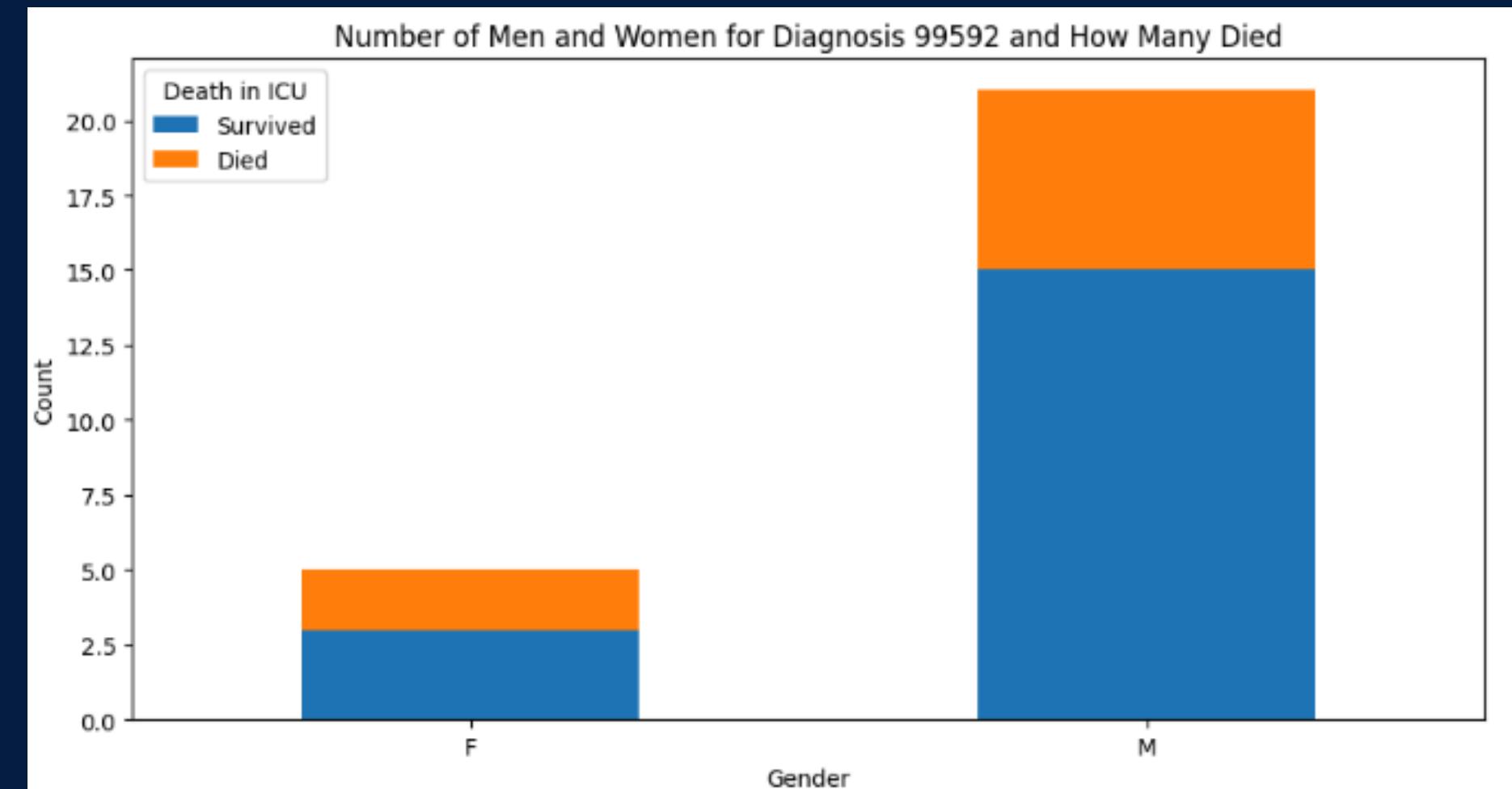
Acute respiratory failure 🦇

תסמונת מצוקה נשימית חריפה



Severe sepsis 💯

אלח דם



💡 אין שוני מרהוטי בין התפלגות גברים ונשים Acute respiratory failure

💡 נטייה חזקה יותר לגברים לחלות ב Severe sepsis

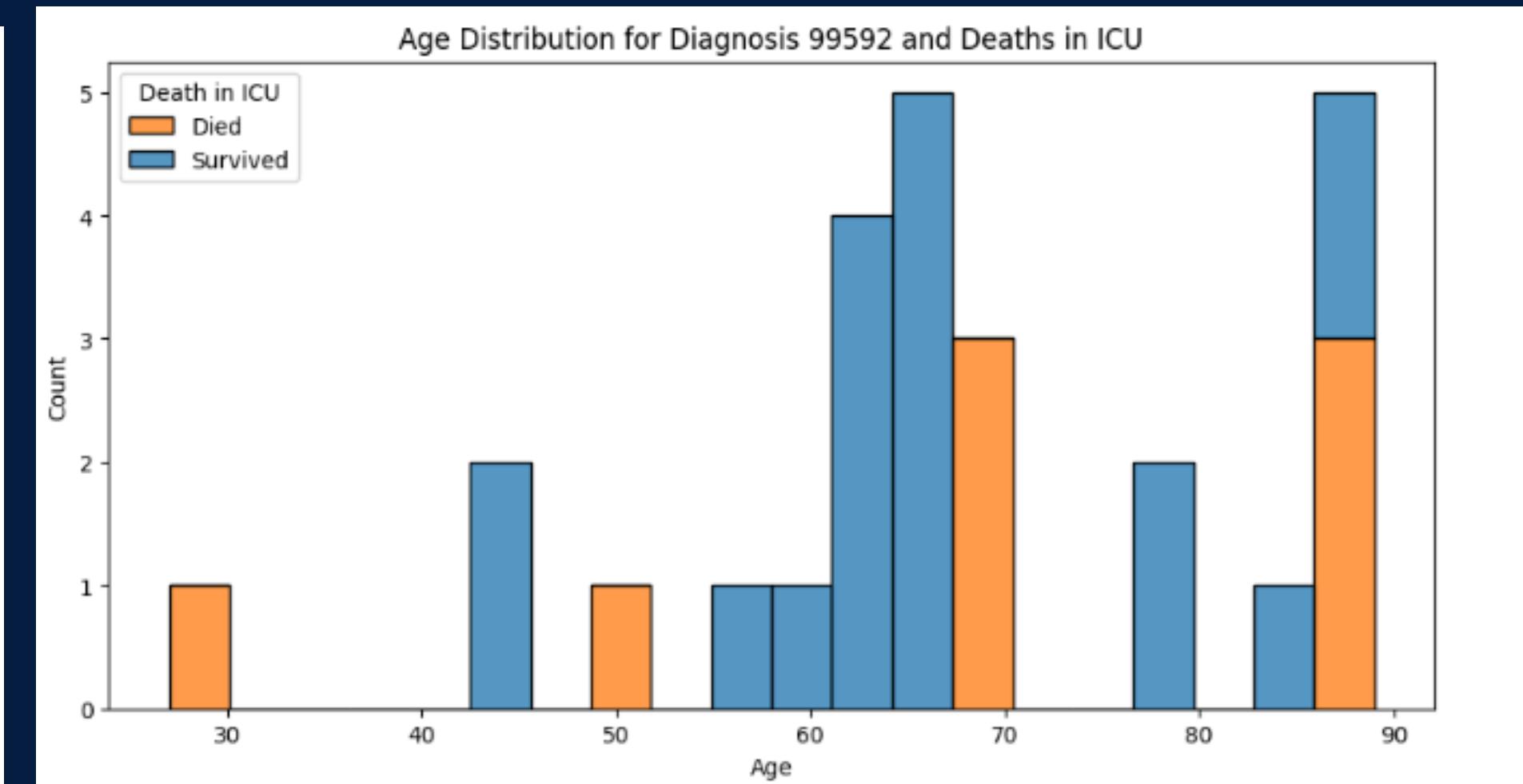
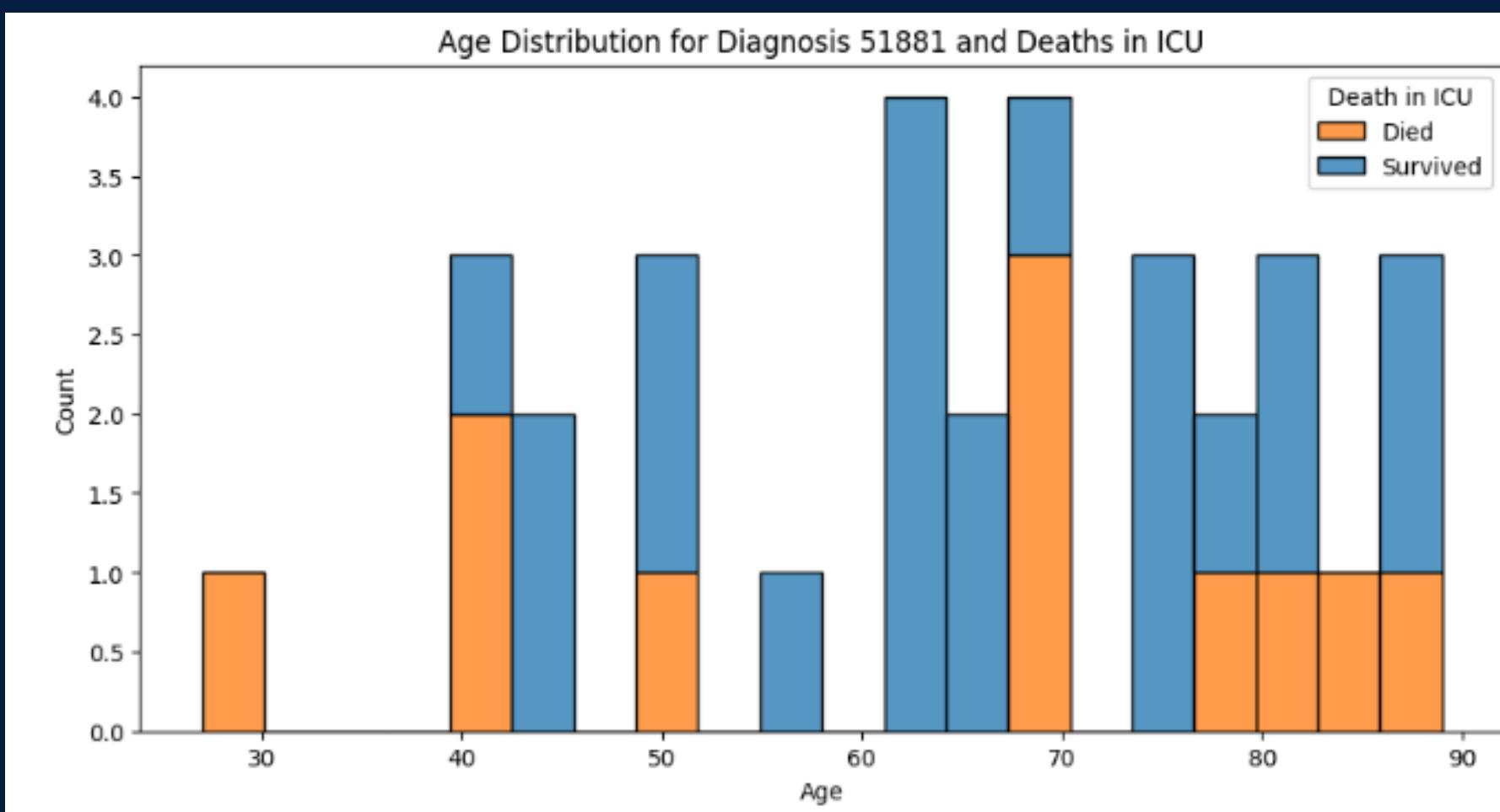
💡 אחוזי התמותה באשפוז היו יחסית זהים (יותר תמונה בקרוב נשים 6%)

2 - התפלגות האילאים ומתוכם כמה מתו



Acute respiratory failure &
תסמנות מצוקה נשימתית חריפה

Severe sepsis 🍃
אלח דם



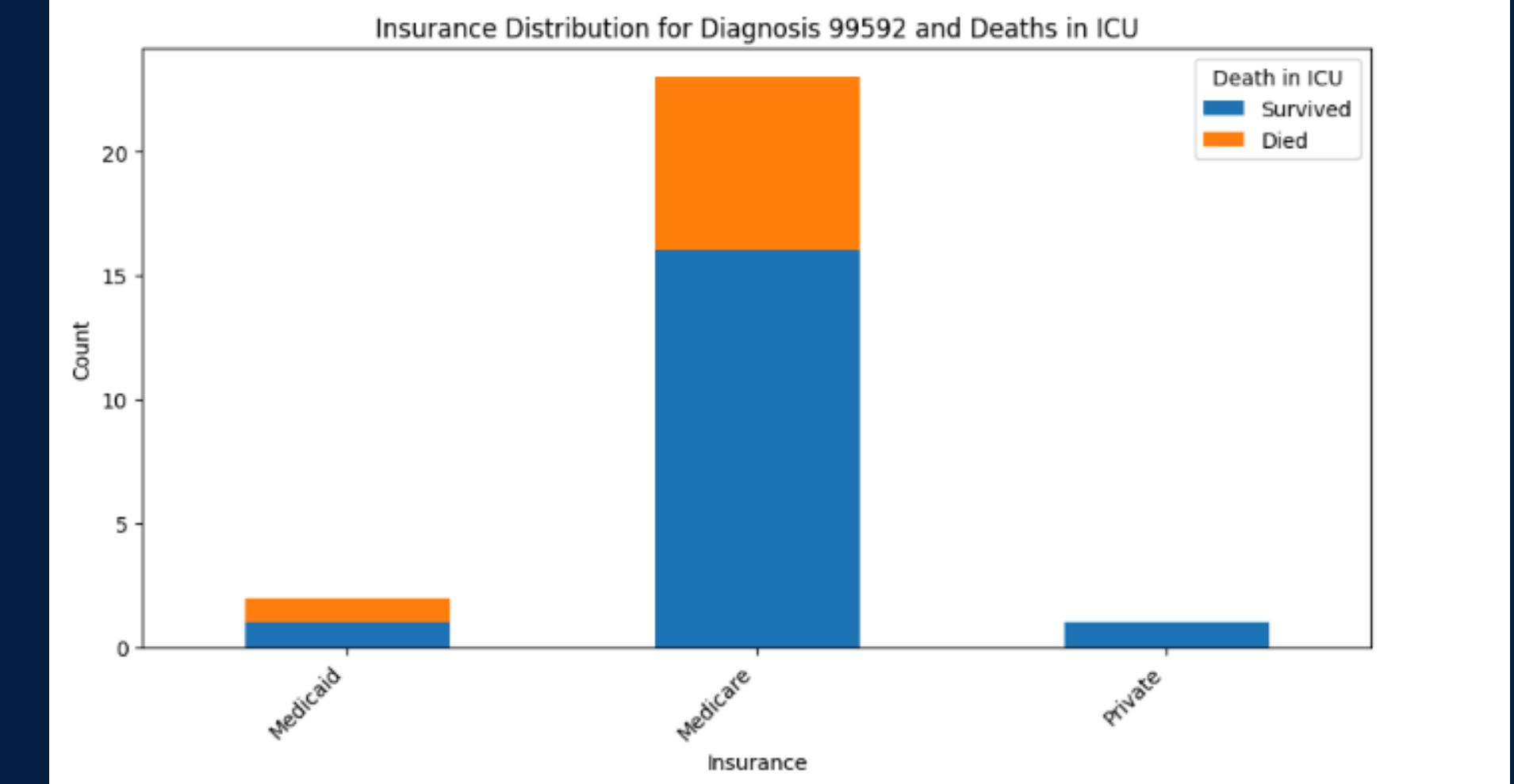
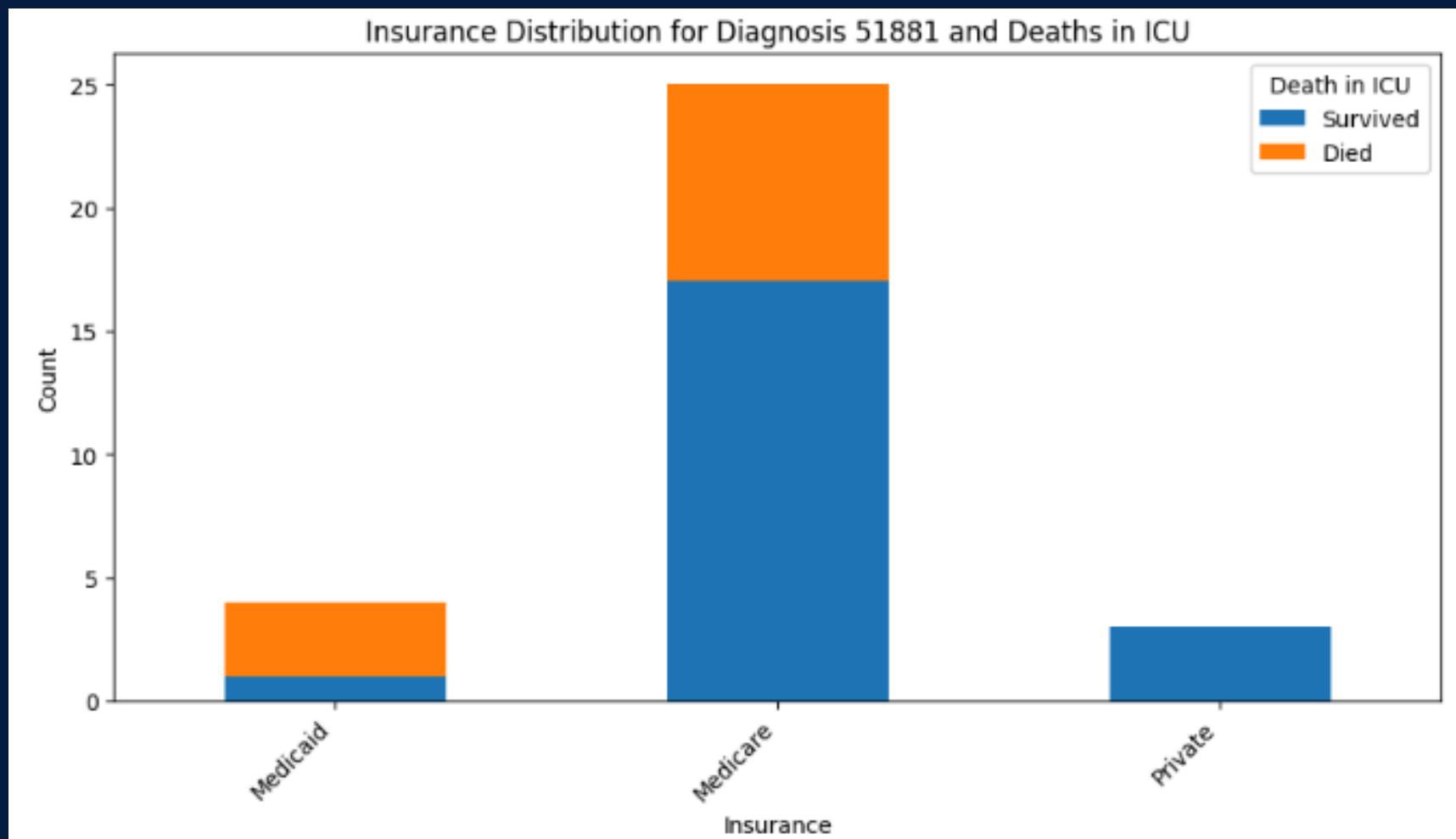
⚠️ אחוזי התמותה מעל גיל 60 הולכים ועולים עברו שני הדיאגנוזות
(אוכלוסית סיון)



3 - התפלגות ביטוח (insurance)

Acute respiratory failure &
תסמנות מצוקה נשימית חריפה

Severe sepsis 🍃
אלח דם



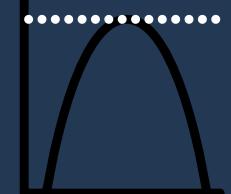
💡 אף אדם עם ביטוח פרטי לא מת בזמן האשפוז (וגם לכתיה מעת מאד חולים)

מסקנה: המשטנים הדמוגרפיים לא מהווים משטנה מרכזי בסיסי התמזה בזמן האשפוז

בדיקות מעבדה וסימניים חיצוניים



- על פי הסקירה המחקנית למדנו כי קיים קשר בין לחץ דם לדיאגנזה של אלח דם
- בנוסף, קיים קשר חזק בין רמת הקריאטינין לבין שתי הדיאגנוזות שלנו.
- יצרנו טבלת עזר חדשה שמכילה את תוצאות הבדיקות האלה עבור אנשים שמתו מרדיאגנוזות שבחרנו במהלך הטיפול נמרץ על פי ארגינזיט הבדיקות הבאות:

LAST CHANCE			
בדיקה אחרונה	ממוצע	מקסימלית	בדיקה ראשונה
נותן אינדיקציה על מצב המטופל כי <u>קרוב להחלמה או מותו</u> .	נותן תמונה מצב <u>כללית יותר</u> על <u>הבדיקה אצל כל בן אדם</u> .	נותן מידע על הערך החרגיבי יותר שנמדד עבור הבדיקה.	נותן מידע על המצב של האדם ברגע הראשו <small>ן</small> שהגיע לטיפול נמרץ.

subject_id	creatinine_first	creatinine_max	creatinine_mean	creatinine_count	creatinine_last	systolic_bp_first	systolic_bp_max	systolic_bp_mean	systolic_bp_count	systolic_bp_last	icd9_code	death_in_icu
0	40177	1.1	2.4	1.300000	11	2.4	86.0	87.0	64.727273	55	50.0	99592
1	40310	0.6	1.6	0.401170	171	0.4	141.0	152.0	110.066964	672	152.0	51881
2	42075	1.4	1.4	0.890476	21	1.0	68.0	151.0	65.421308	413	0.0	51881
3	42281	0.7	1.5	1.106522	46	0.9	85.0	134.0	92.197531	486	21.0	51881
4	43870	1.1	1.1	1.066667	3	1.1	68.0	92.0	67.846154	39	51.0	51881

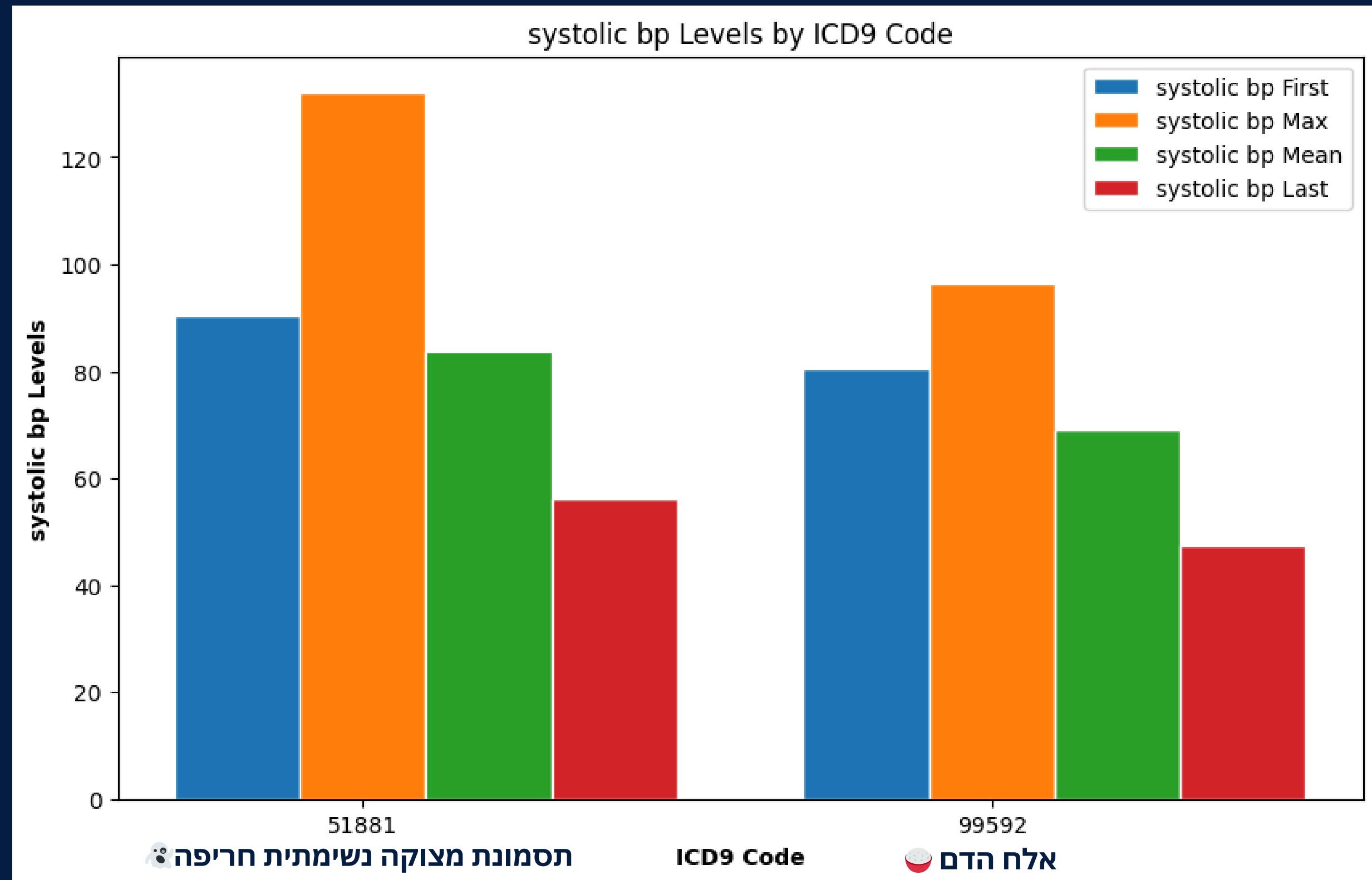
בדיקות מעבדה וסימנים חיוניים:

נתוני לחץ דם לכל אחת מהדיאגנוזות

טבלת מדדים שימושיים תקינים

Category	Systolic (mm Hg)
Dangerously low	50 or lower
Very low	60 or lower
Low	Less than 90
Normal	Less than 120

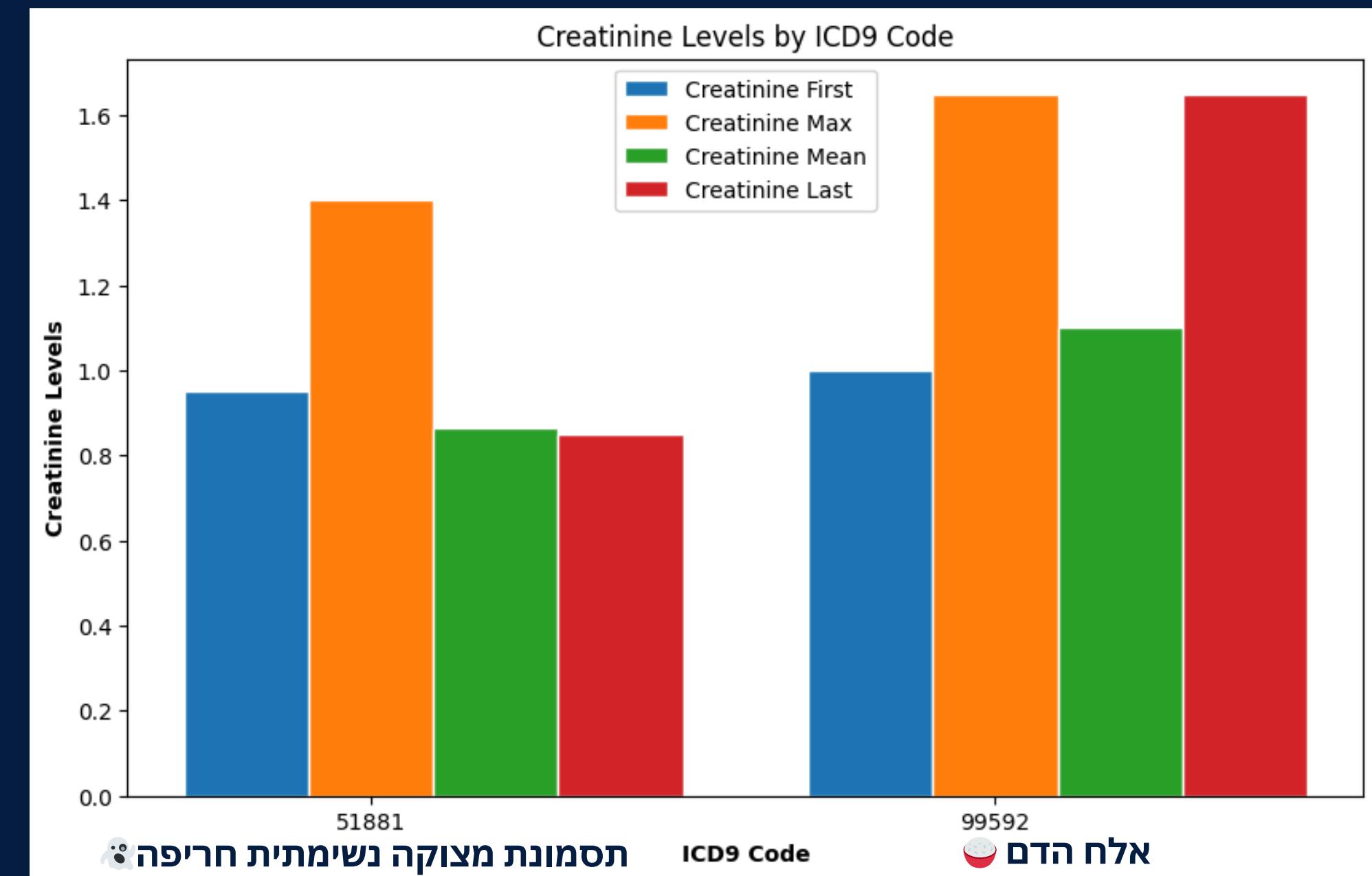
💡 התרפלגות דומה בין שני הדיאגנוזות



בדיקות מעבדה וסימנים חיוניים:

נתוני קריאטינין לבל אחת מהדיאגנוזות

ליקראטינין יש תלות בגיל המטופל ולבן הגרפ' הזה בפni עצמו לא עוזר מספיק יצרנו טבלה שמחולקת לפי גילאים את האנשים בנתונים שלנו ובודקת את ערכי הקריאטין אצלם



בדיקות קרייטאני

הטבלה שיצרנו לפי חלוקה לטווחי גילאים ורמות הקרייטניין

Age Group	Average Last Creatinine	Normal Creatinine	Difference (%)
0 18 to 29 years	7.787500	1.195	551.673640
1 30 to 39 years	0.563158	1.130	-50.163018
2 40 to 49 years	1.214887	1.065	14.073871
3 60 to 72 years	1.442222	0.935	54.248366
4 above 72 years	1.514881	0.935	62.019353

הערבים התקיניים לקרייטניין

Table 1. Normal creatinine clearance by age

Age	Normal creatinine levels
18 to 29 years	78 to 161 mL/min/BSA
30 to 39 years	72 to 154 mL/min/BSA
40 to 49 years	67 to 146 mL/min/BSA
50 to 59 years	62 to 139 mL/min/BSA
60 to 72 years	56 to 131 mL/min/BSA

מתודולוגיה: השווינו את רמת הקרייטניין "התקינה" ביחס לממוצעי הקרייטניין בבדיקה [לאחרונה](#) לממטופל.

מטרה: היא להשוות את הבדיקה הכי סטור להטאוששות/מוות במהלך האשפוז והוות אינדיקציה לסיכוי התמונותה.

מסקנות:

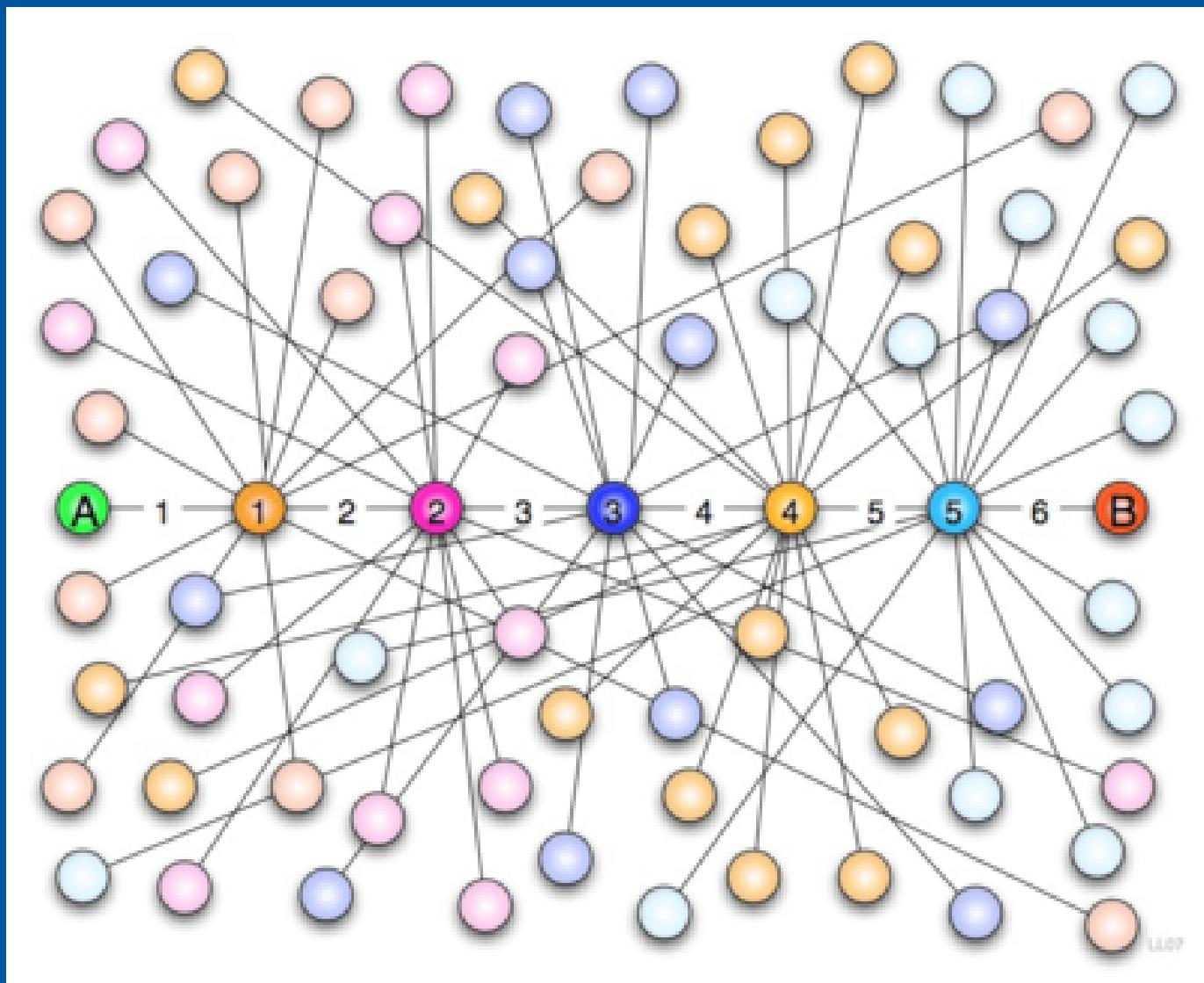
קובצת גיל 18-29 ו-60 ומעלה

 **רמות הקרייטניין גבוהות משמעותית מהנורמה, מעידות על תפקוד כלייתי לקוי חמור.**

קובצת גיל 30-39

 **רמות הקרייטניין נמוכות מהנורמה**

מציאת קשרים עמוקים בין הדיאגנוזות שבחtmp לדיאגנוזות מסוימות

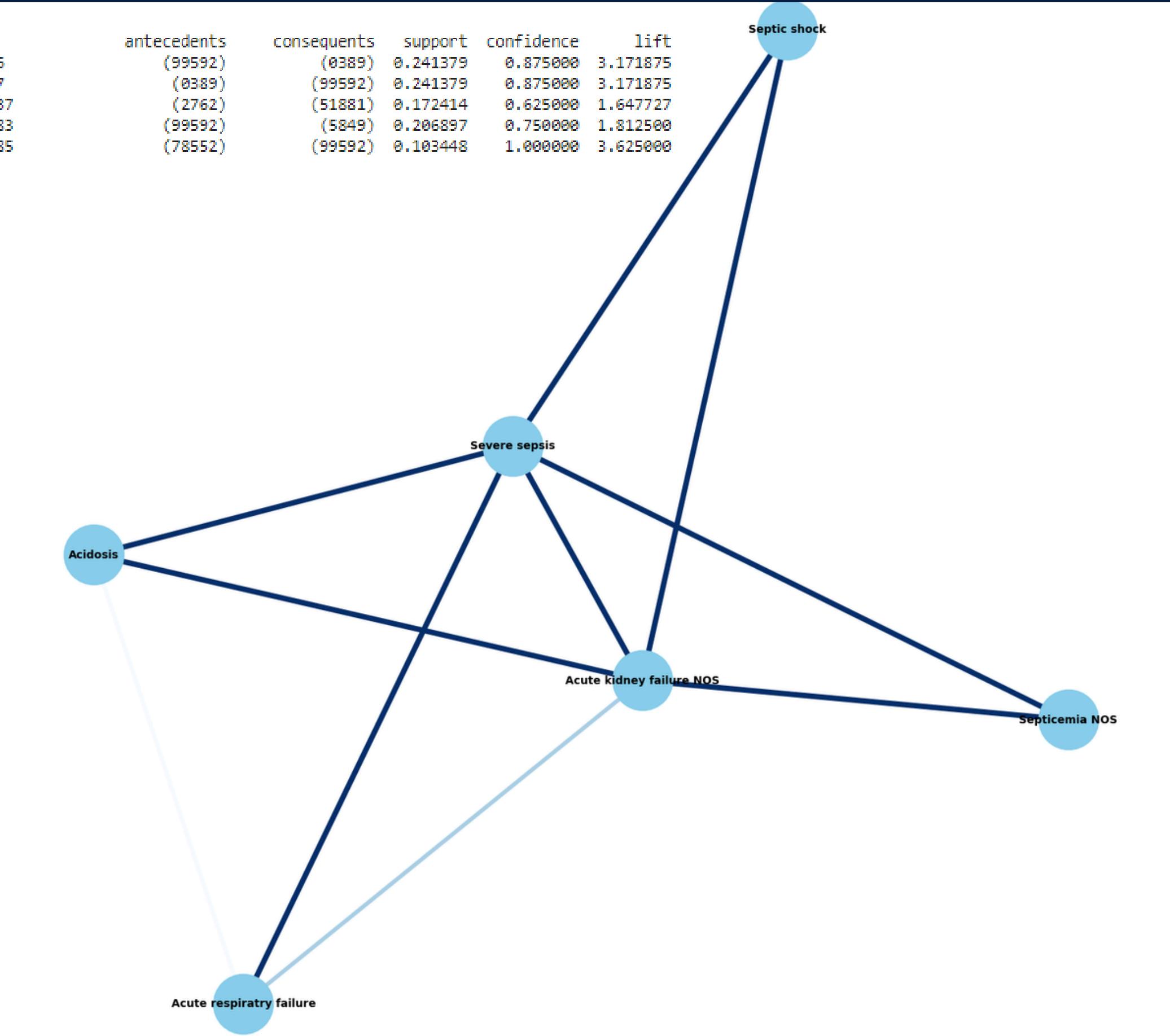


- מטור הינה שישנים דיאגנוזות הקשורות ואפילו גורמות לדיאגנוזות אחרות, רצינו למצוא תת-קבוצות של דיאגנוזות הקשורות לדיאגנוזות שבחרנו.
 - אנו מאמינים שמציאת קבוצות אicasיות תוביל לשיפור את המודל שנבנה ותגרום לנו לחתה בחשbon דיאגנוזות אחרות שיועילו לנו ואף תעוזר לנו לתת להם ציון במודלים מסוימים שנבחנו.
- בחרנו לבצע את זה בעזרת מודל שנלמד בקורס **SD-Apriori** והציגו את הקבוצות בעזרת גרף רשת

ביצענו 2 פרויקטים לחת-קבוצות:

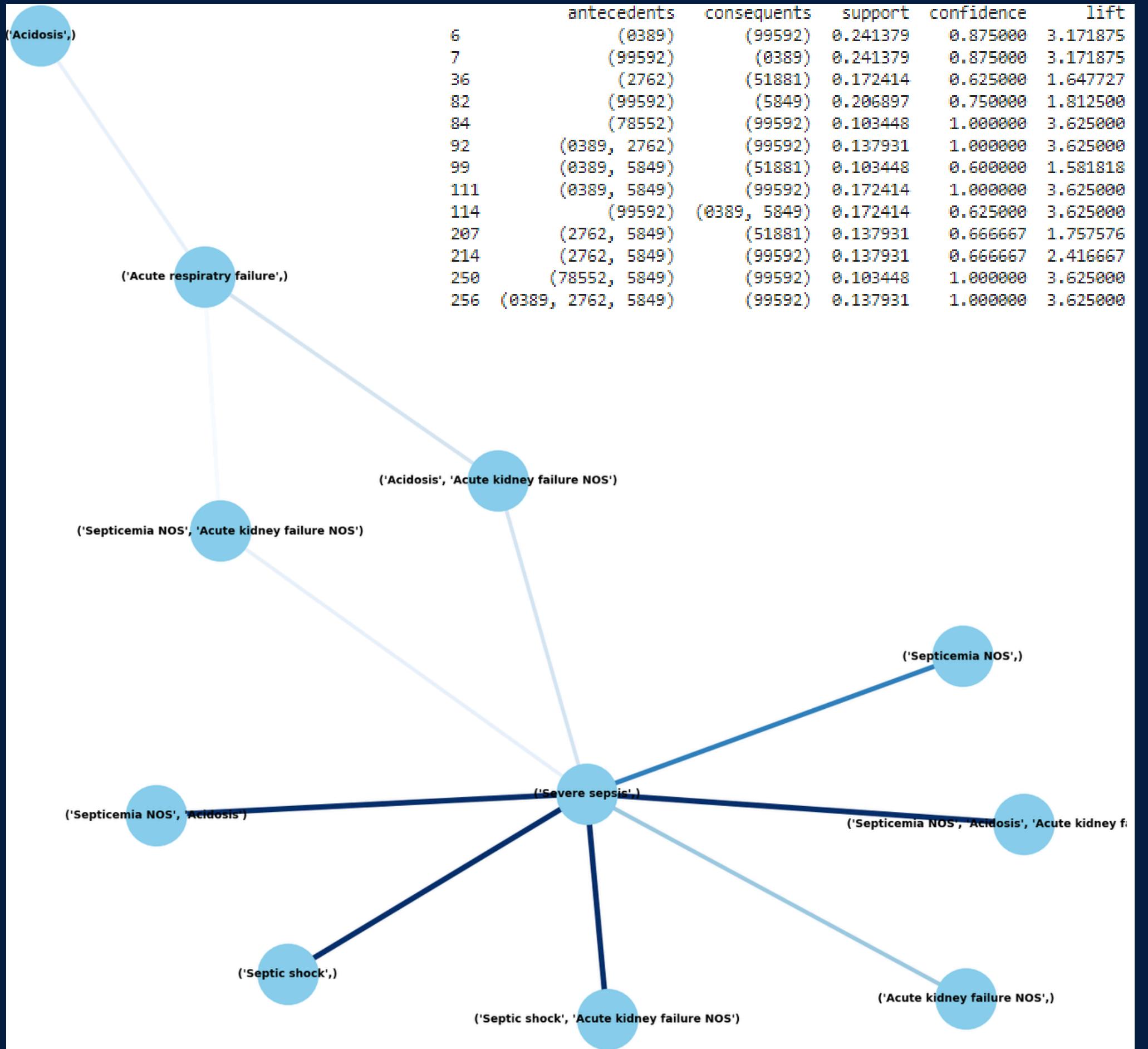
1

פירוק לחתמי קבוצות
בעזרת המודל ולקירה
של זוגות בלבד על מנת
לראות קשרים חזקים
וישירים יותר.
(דיאגנוזה אחת לשניה)



2

פירוק לחתמי קבוצות מרובות
агдоловות יותר מדיאגנוזה אחת
על מנת למצוא קשרים
מורכבים ועמוקים יותר.
**(diagnoza achet machovrat lan
diagnozot achrot)**



מסקנות: סיווע בבנויות מודל חיזוי רפואי



שיפור דיקן המודלים

- באמצעות שימוש במשתנים כמו גיל, מגדר, לחץ דם ודיאגנווזות, ניתן לשפר את הדיקן של מודלים לחיזוי תוצאות קליניות.
- שילוב משתנים דמוגרפיים וקליניים מאפשר למודל לקחת בחשבון את כל הפרמטרים המשפיעים.



זיהוי נתונים קריטיים

- באמצעות שימוש במשתנים כמו גיל, מגדר, לחץ דם ודיאגנווזות, ניתן לשפר את הדיקן של מודלים לחיזוי תוצאות קליניות.
- שילוב משתנים דמוגרפיים וקליניים מאפשר למודל לקחת בחשבון את כל הפרמטרים המשפיעים.



הכנסת נתונים בזמן אמת

- ניתוח דינמי והכנסת נתונים בזמן אמת מאפשרו חיזוי מיידי ושיפור הטיפול בזמן אמת.
- מודלים אלו יכולים לסייע לרופאים בקבלת החלטות מבוססות נתונים בזמן אמת.