

...

# למידת מכונה ברפואה

ניתוח נתונים רפואיים והסקת מסקנות



מגישים:



- עמית פאלח
- מתן שמעון
- יעדו צילמס
- קארין לב סופי

# מבנה

01

מחקר

02

דاطה ועיבוד  
ראשוני

03

ניתוח אוכלוסיית  
המחקר

04

מסקנות

## MORTALITY PREDICTION AMONG ICU INPATIENTS BASED ON MIMIC-III DATABASE RESULTS FROM THE CONDITIONAL MEDICAL GENERATIVE ADVERSARIAL NETWORK



המאמר עוסק בחולים הנמצאים ביחידות טיפול נמרץ (ICU).

### מטרת המחקר

היא להעריך ביצועי מודלים שונים לחיזוי תמותה בקרב החולים אלו. המודלים משתמשים נתונים קליניים וдמוגרפיים שנאפסו במהלך השנות של החולים בטיפול נמרץ.

### שיטת המחקר

המחקר כולל תוכנות, המורכבות מדמוגרפיה של המטופל, אבחון, סימנים חיוניים, תוצאות בדיקות, טיפול. גורמים אלו משמשים כמשתנים במודלים שהוערכו. המחקר בוחן את המודלים הבאים: SVM, II, SAPS, GAN, med-c-MLP.

### מסקנות המחקר

המודל GAN med-c השיג את הביצועים הטובים ביותר במדד PR-AUC, ROC-AUC ו-Score F1 בהשוואה ל-MLP, SVM, II, SAPS.

# THE DATA



CHARTEVENTS



DIAGNOSES



PATIENTS



ADMISSION



D\_LABITEM



CALLOUT



D\_ITEMS



ICUSTAYS

	subject_id	gender	age	ethnicity	religion	is_infected	bp_mean	pulse_mean	mortality	list_of_labels
0	10006	F	70	BLACK/AFRICAN AMERICAN	CATHOLIC	0.0	122.388889	82.000000	0	[BSA, Admit Wt, Allergy 1, Diagnosis/op, Servi...
1	10011	F	36	UNKNOWN/NOT SPECIFIED	CATHOLIC	0.0	110.829787	82.398649	1	[Abdominal Assessment, Activity, Activity Tole...
2	10013	F	87	UNKNOWN/NOT SPECIFIED	CATHOLIC	0.0	102.356322	92.146341	1	[Heart Rate, NBP Mean, Respiratory Rate, SpO2,...
3	10017	F	74	WHITE	CATHOLIC	1.0	113.236842	91.836735	0	[Abdominal Assessment, Activity, Behavior, Bow...
4	10019	M	49	WHITE	CATHOLIC	0.0	83.098039	83.333333	1	[Airway Size, Airway Type, Apnea Time Interval...

# עיבוד ראשוני - ערכים חסרים

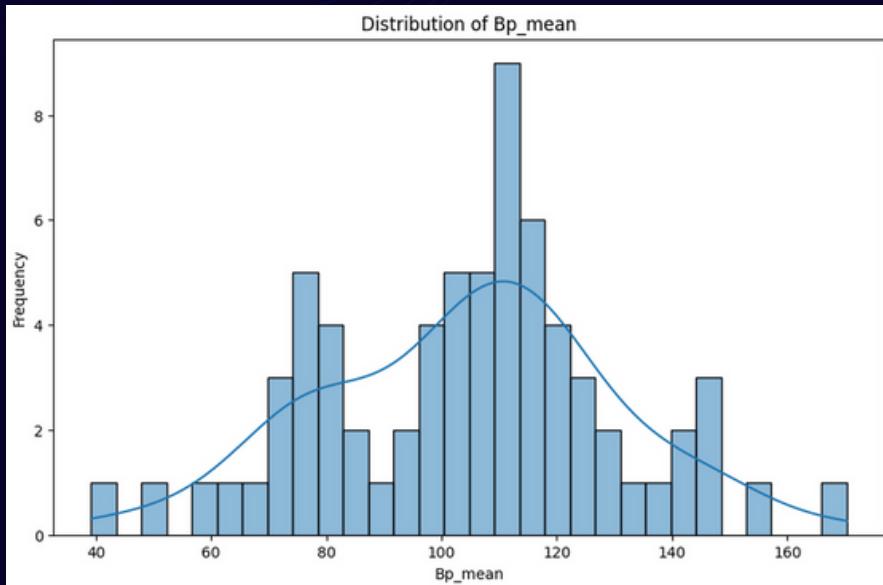


השלבים בהם פועלנו:

```
missing_bp_mean = all_patients_data['bp_mean'].isnull().sum()  
print(f'Number of missing values in bp_mean: {missing_bp_mean}')
```

```
Number of missing values in bp_mean: 29
```

```
plot_histogram_with_kde(all_patients_data, 'bp_mean')
```



```
all_patients_data['bp_mean'].median()
```

```
107.06172839506173
```

```
all_patients_data['bp_mean'].fillna(all_patients_data['bp_mean'].mean(), inplace=True)
```

- סקירהת ערכים ריקים בדאטה הרלוונטית
- בנינו פונקציה שמטרתה לבדוק את דילולות הנתונים, הפונקציה מצינה את כמות השורות המכילות יותר מ-N ערכים ריקים

עבור עמודות ערכים ריקים בודדים בחרנו למחוק רשומה (עמודות LIST OF LABELS, PULSE MEAN, RELIGION)

עמודה MEAN BP כולה מס ערכים חסרים - בחרנו לבדוק התפלגות על מנת להבין באיזה שיטה לטפל



# עיבוד ראשוני - חיריגים

...

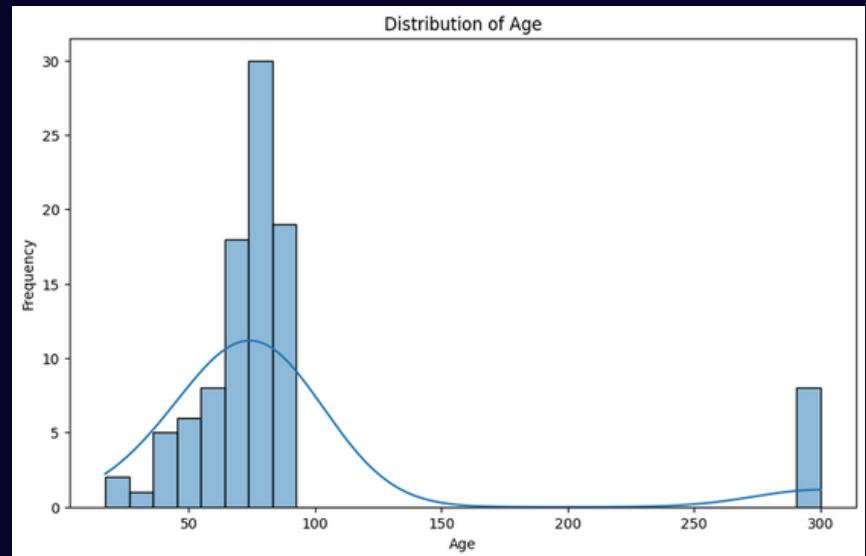
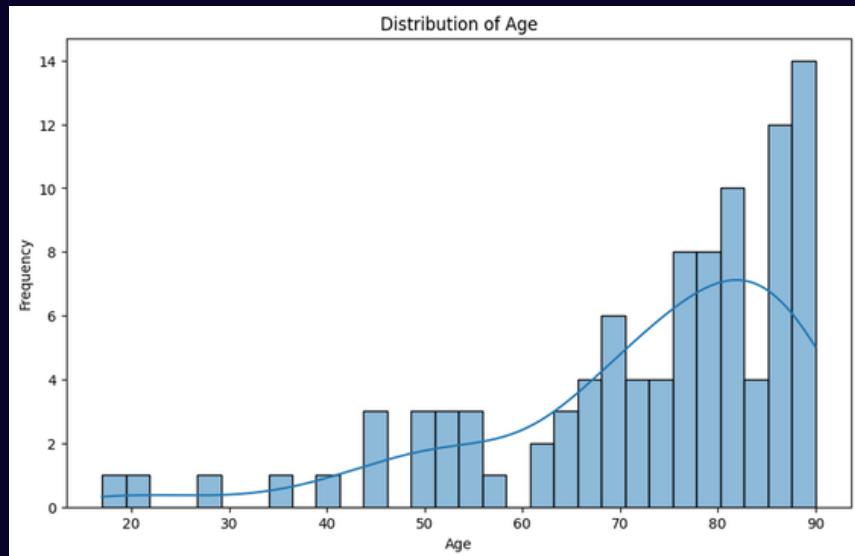
נבחן כל עמודה בנפרד לפני המשך הניתוך כדי לראות אם יש להתייחס לחריגות

אחרי הגבלת הגיל עד 100

02

לפני טיפול בחיריגים עבור  
עמודת Age

01

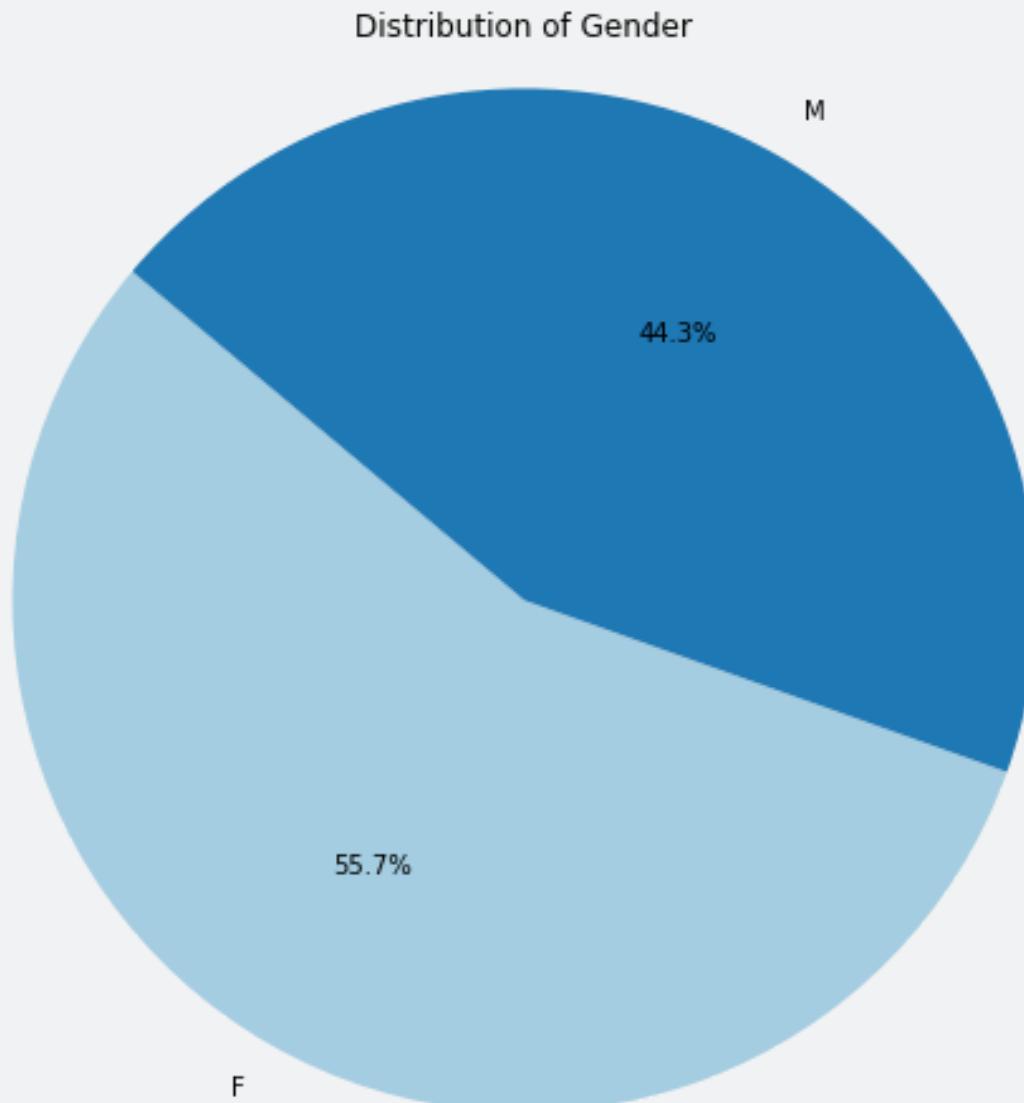


```
len(all_patients_data[all_patients_data['age'] > 100])  
8  
all_patients_data.loc[all_patients_data['age'] > 100, 'age'] = 90  
plot_histogram_with_kde(all_patients_data, 'age')
```



## משתנה המטרה

נסתכל על משתנה המטרה וצורת התפלגות שלו



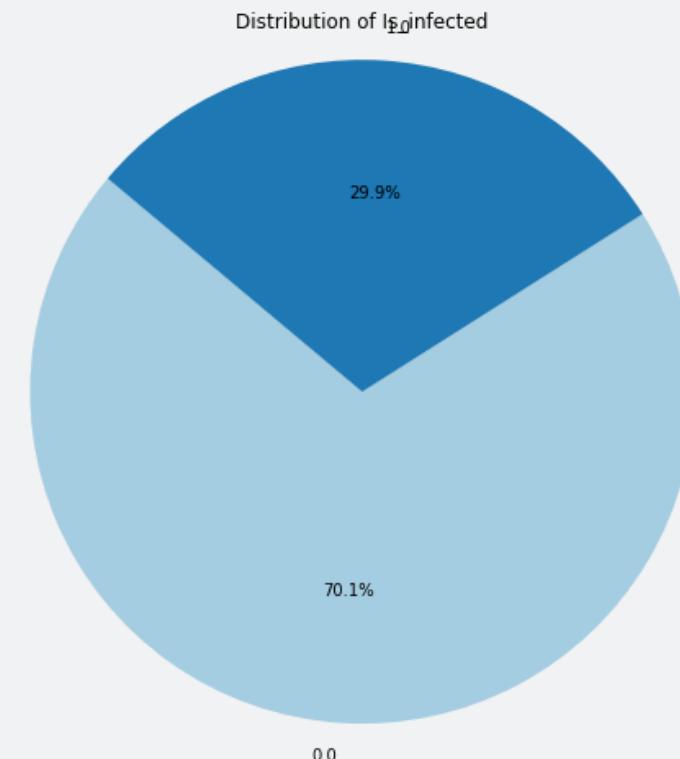
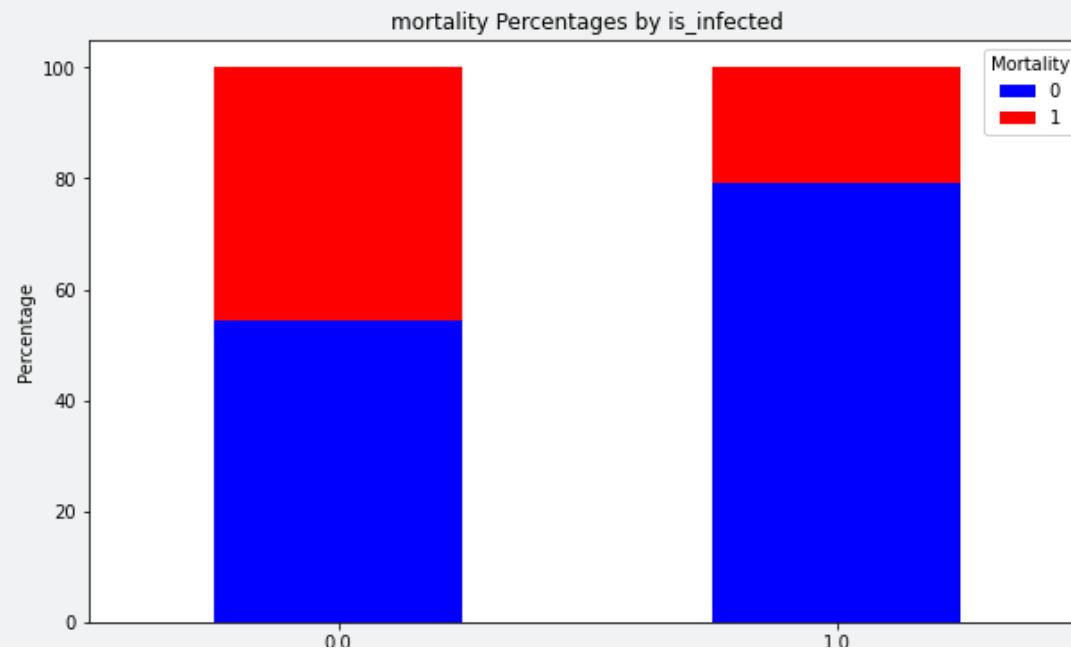
```
[35] all_patients_data.groupby('mortality')[['mortality']].count()
```

mortality	
0	60
1	37



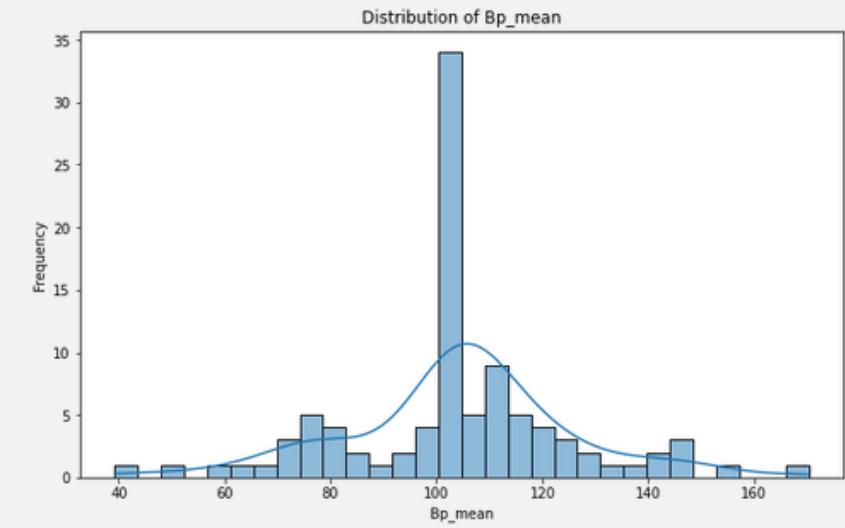
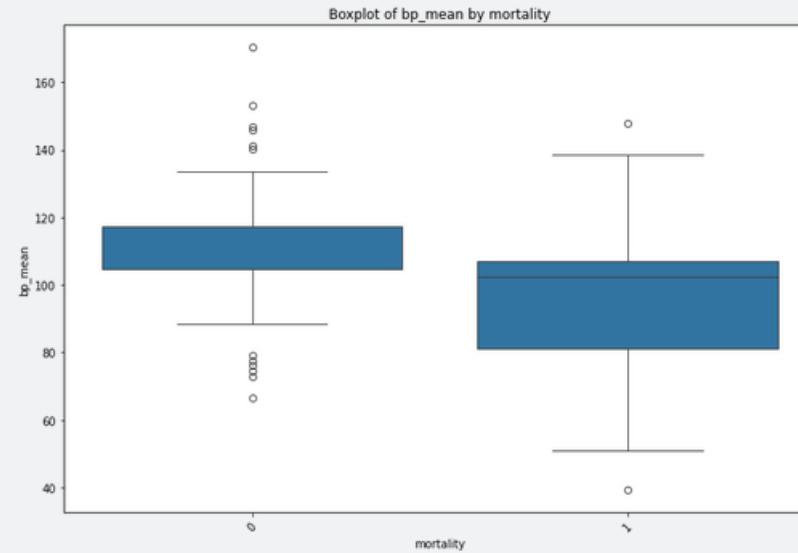
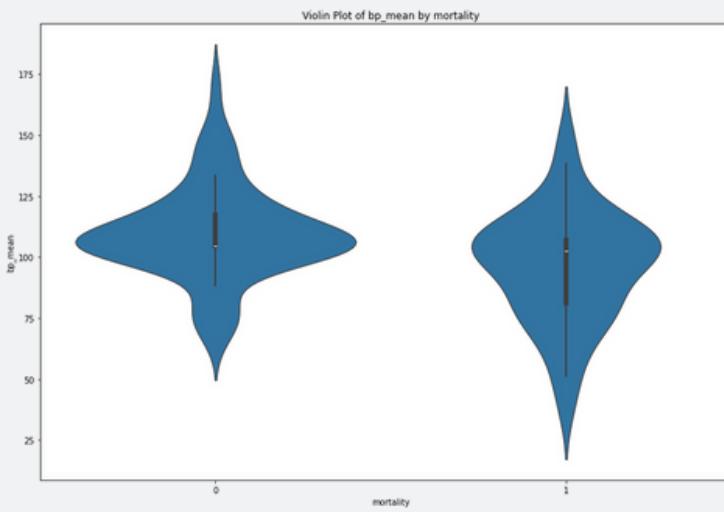
## בדיקות קשורים בין משתנה המטרה למשתנים החזויים is\_infected

מצאנו כי יש קשר מובהק בין אנשים ששורדים לבין אנשים שנדרבקו בחידוך חסין



```
[ ] chi_square_test(all_patients_data, 'mortality', 'is_infected')  
→ P-value: 0.03725999871460757  
There is a statistically significant association between the two variables.
```

## בדיקות קשרים בין משתנה המטרה למשתנים החזויים bp\_mean



```
t_test_mortality_numeric(all_patients_data, 'bp_mean')
```

P-value for bp\_mean: 0.00399135585779228

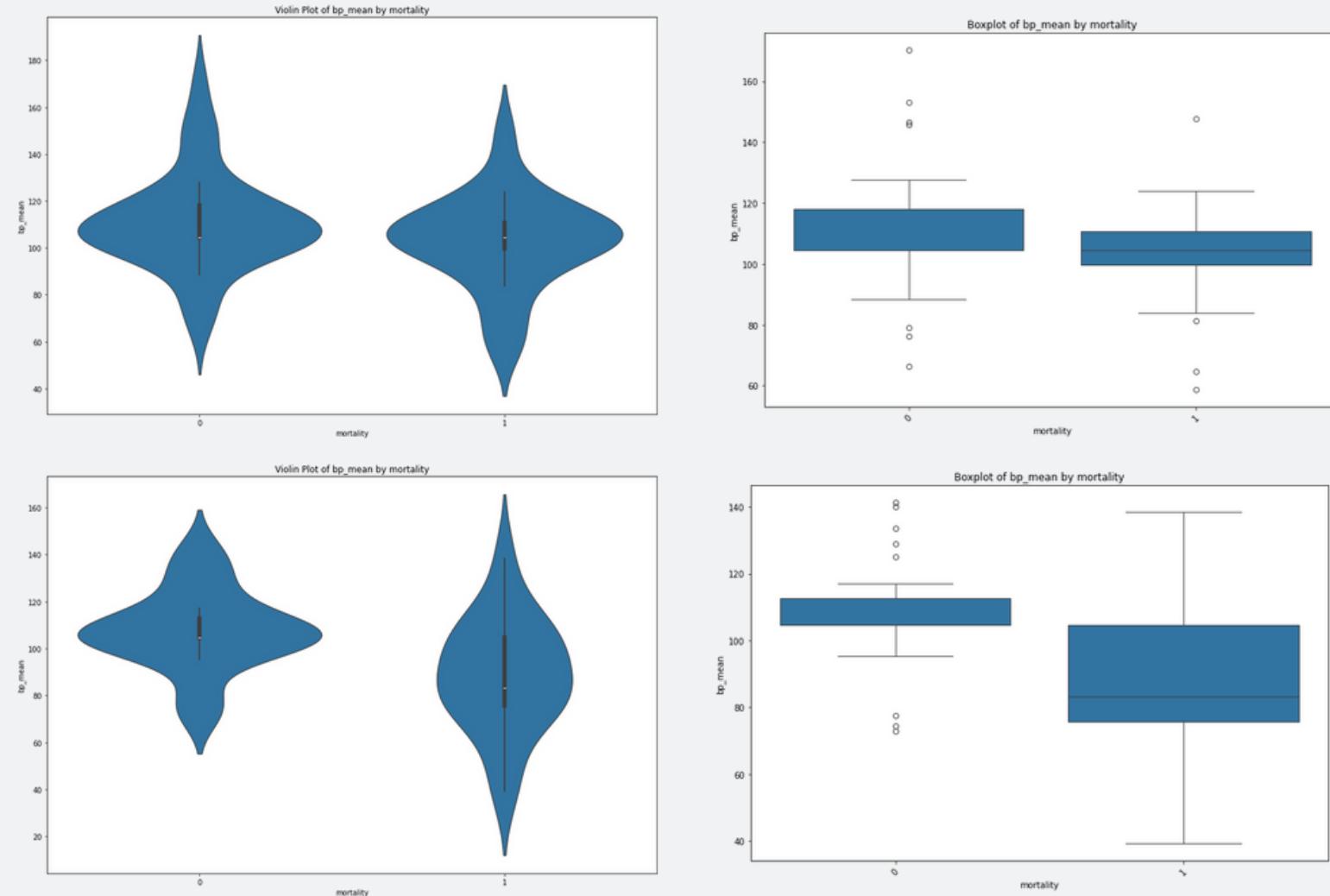
The difference in values between those who died and those who lived is statistically significant.

מיצאנו כי יש השפעה מובהקת סטטיסטית בין לחץ דם לתמותה

## בדיקות קשורים בין משתנה המטרה למשתנים החזויים

**bp\_mean** בחרנו להעמיק את הקשר ולבזק ספציפיות לפי מגדר (gender)

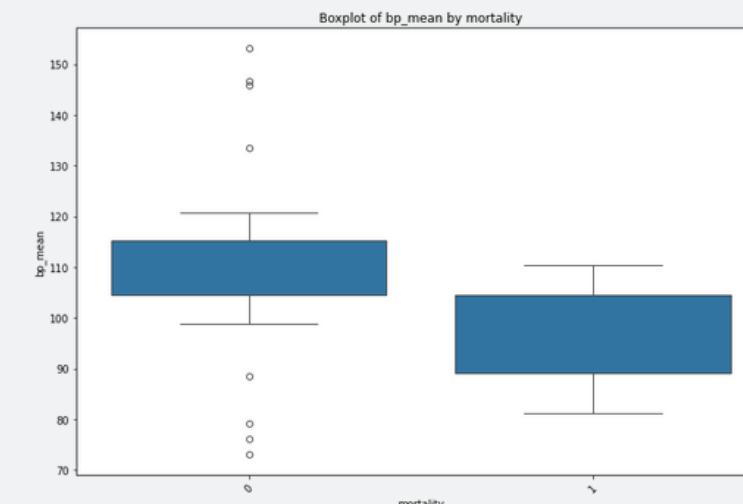
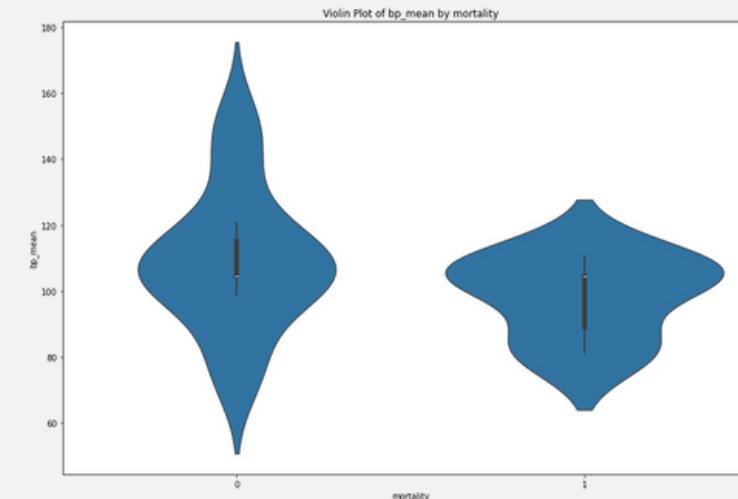
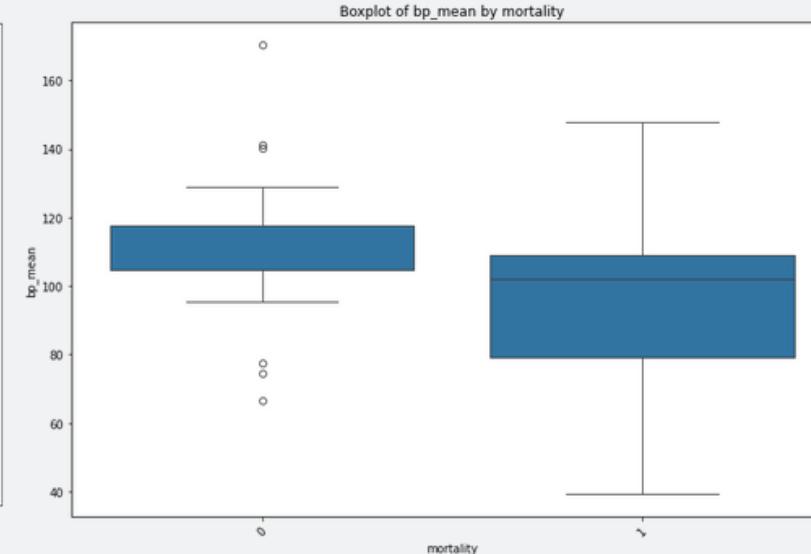
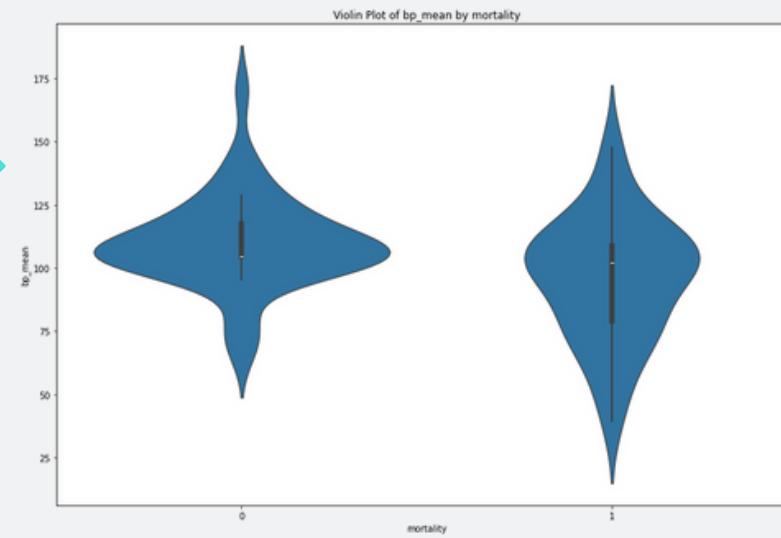
לאחר חלוקה לזכרים ונקבות, הצלחנו לחדד את מסקנה ולומר **לחץ דם יש השפעה משמעותית על התמונת אם המין של המטופל הוא זכר.** אם המטופלת היא נקבה, לחץ הדם אין השפעה משמעותית על התמונת.



## בדיקות קשורים בין משתנה המטרה למשתנים החזויים

**bp\_mean** בחרנו להעמיק את הקשר ולבזוק ספציפיות לפי חידוך(INFECTED)

♦ ניתן לראות שיש השפעה משמעותית בין לחץ דם לתמונת החולים שלא נדבקו בחידוך עמיד (מבחן T), נבדוק אם ההשפעה משתנה לפי מגדר כדי לחזק את המסקנה.



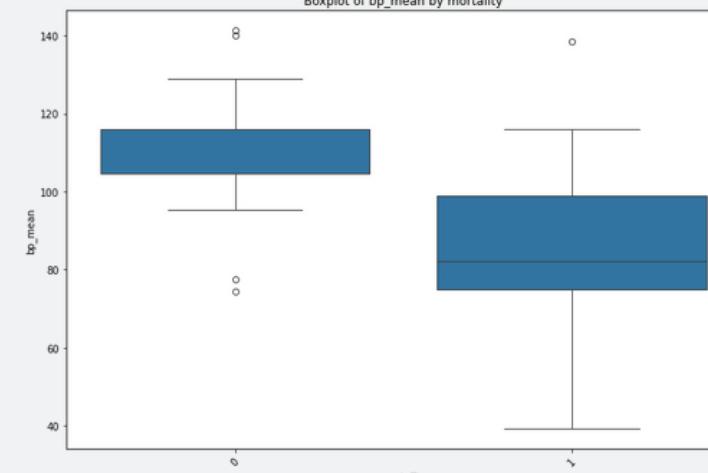
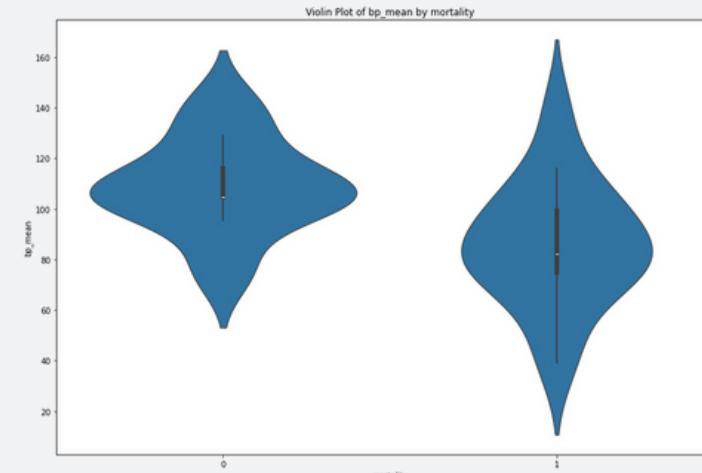
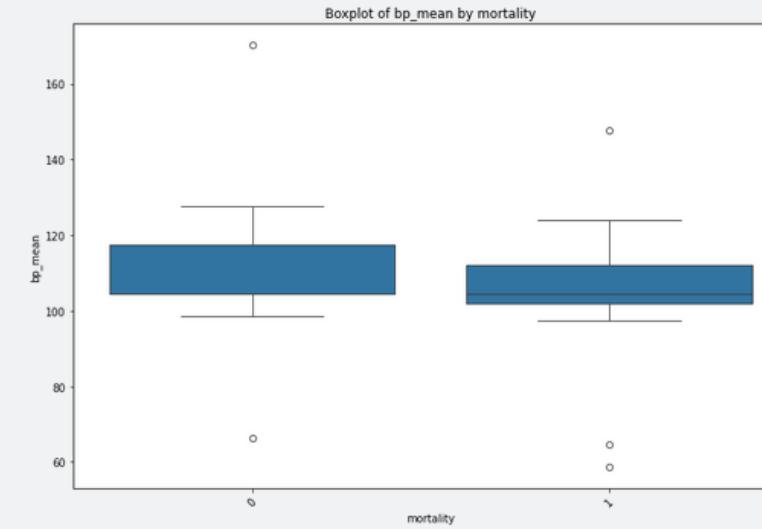
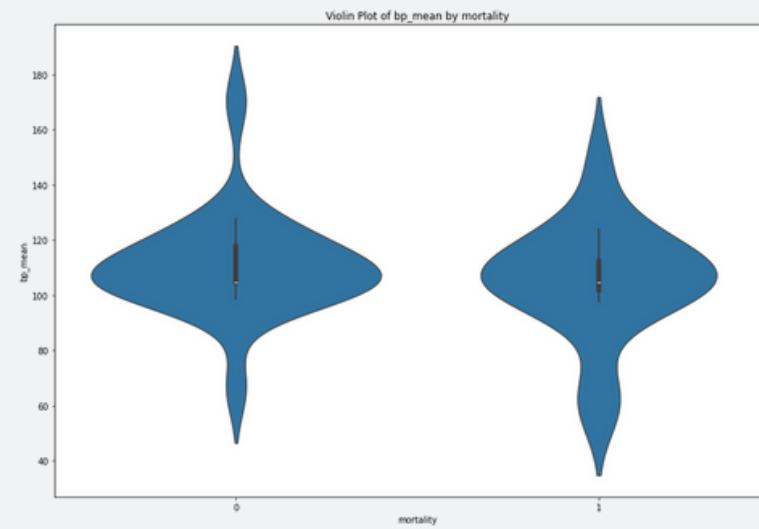
UN

IS

## בדיקת קשרים בין משתנה המטרה למשתנים החזויים

**bp\_mean** בחרנו להעמיק את הקשר ולבזק ספציפיות לפי מגדר (gender) ולפי אדם שאינו נגוע

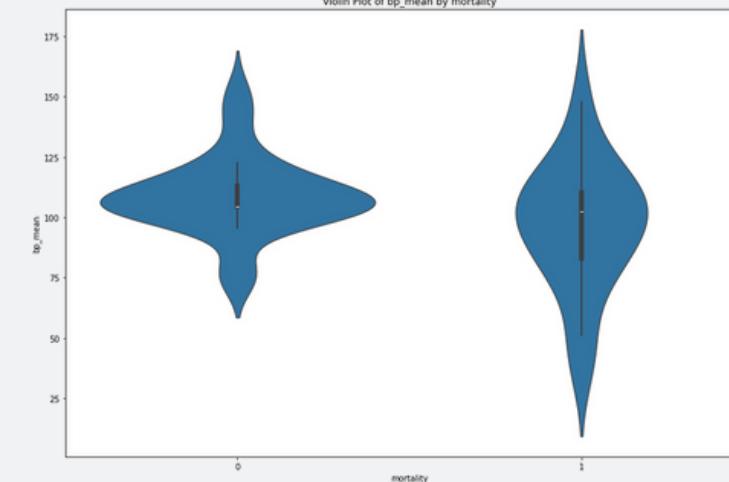
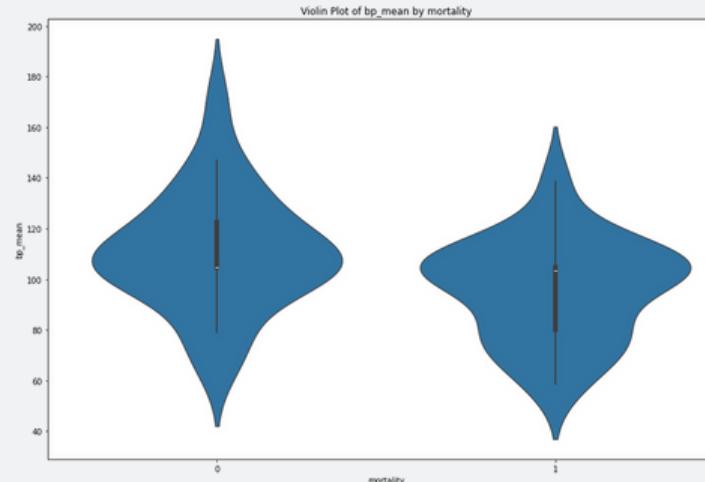
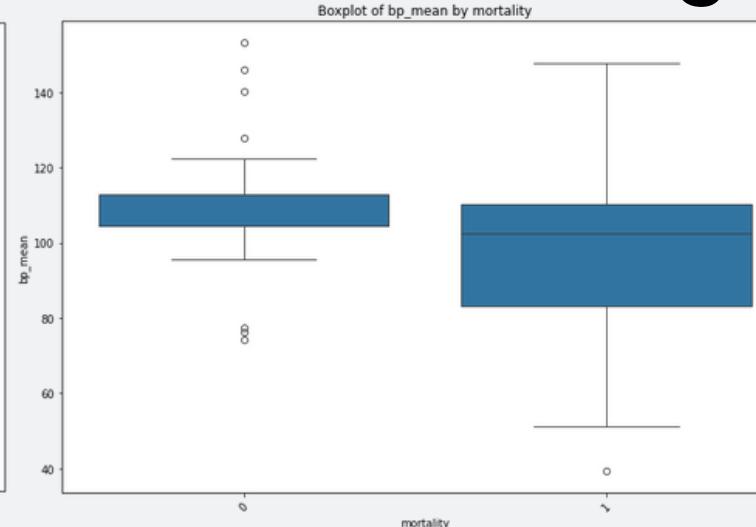
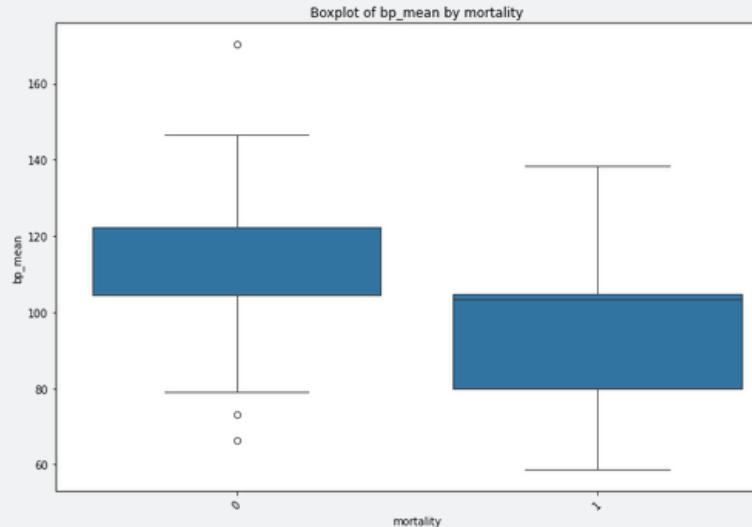
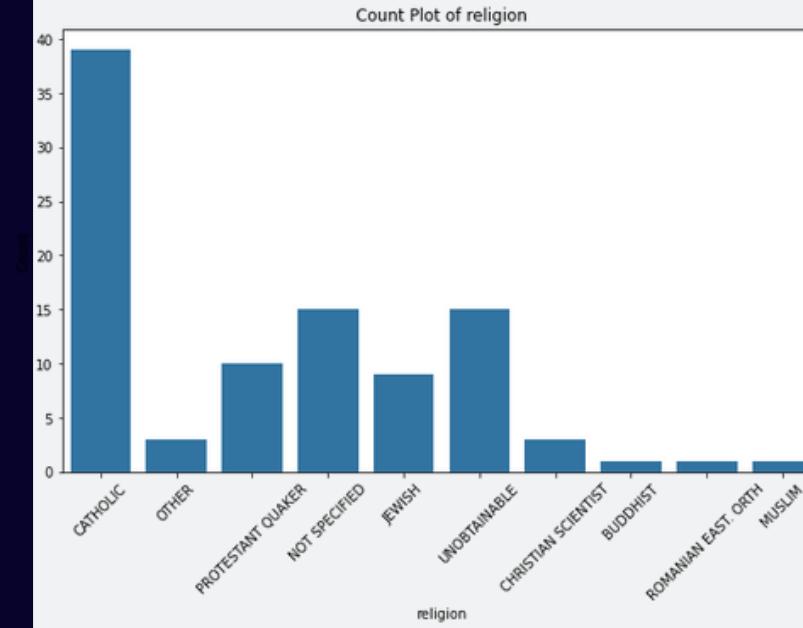
❖ ניסינו להעמיק את  
המסקנה לפי מגדר  
והראינו שאם המטופל  
הוא זכר (מין = "M")  
שאינו נגוע בנגיף חיסוני  
('0=0 infected\_is') לחץ  
הדם יש השפעה  
משמעותית על התמותה  
(חולים העומדים  
בקרייטרונים עם לחץ דם  
גבוה יותר ישרדו עם סיכוי  
גבוה יותר).



F

M

## בדיקות קשרים בין משתנה המטרה למשתנים החזויים בגלא פילוח האוכלוסייה בחרנו להתמקד בנוצרים / לא נוצרים **religion**



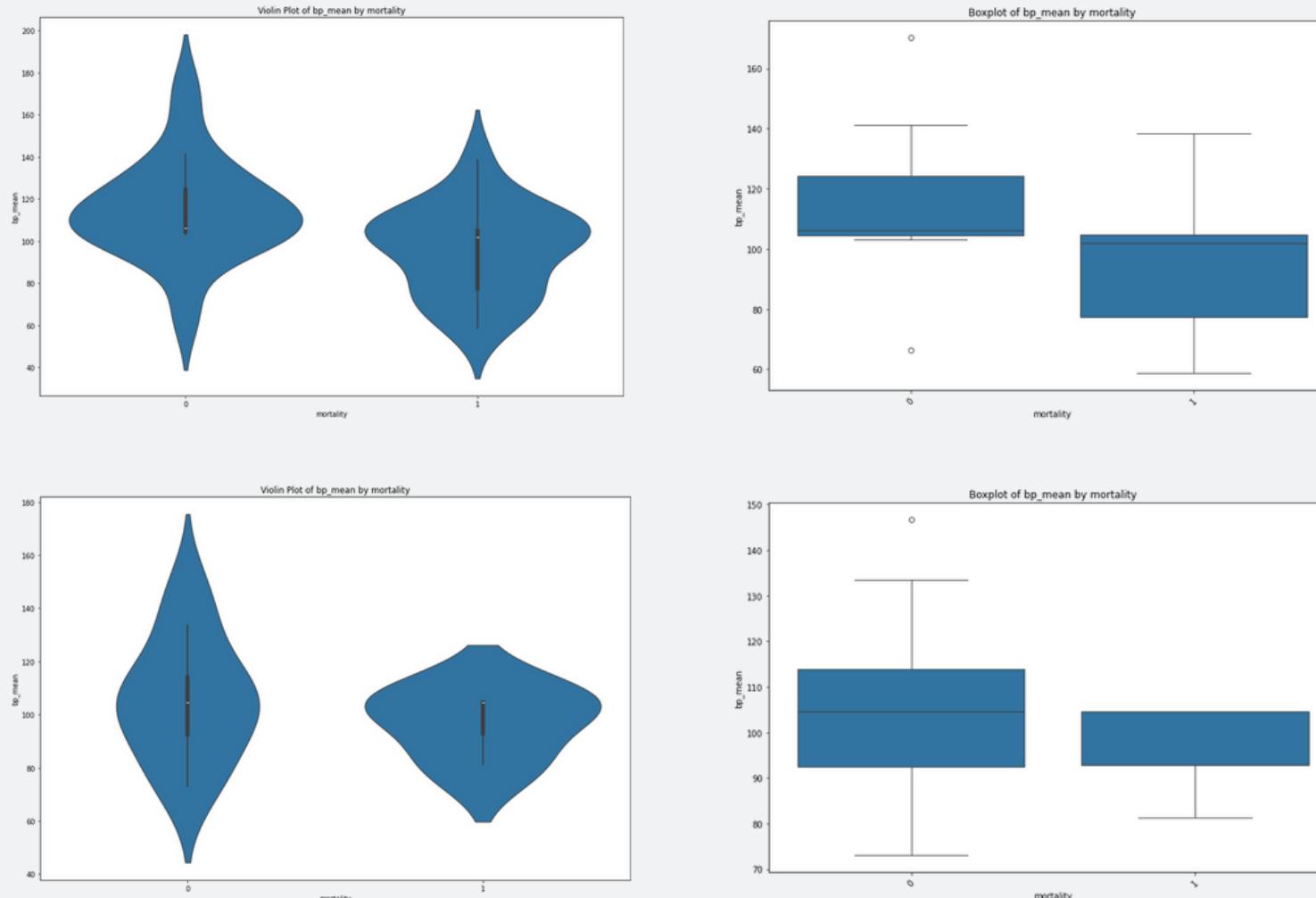
❖ ניתן לראות שם  
החוליה איננו נוצרי יש  
קשר בין לחץ הדם שלו  
لتמותה.

CH

OTH

## בדיקות קשורים בין משתנה המטרה למשתנים החזויים INFECTED religion

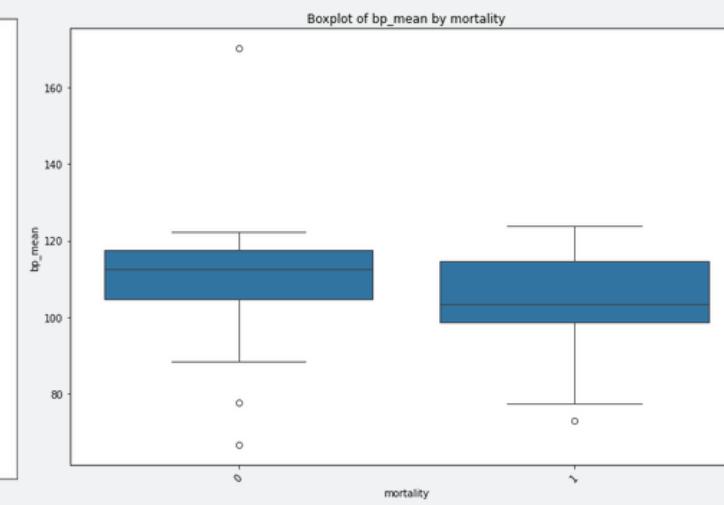
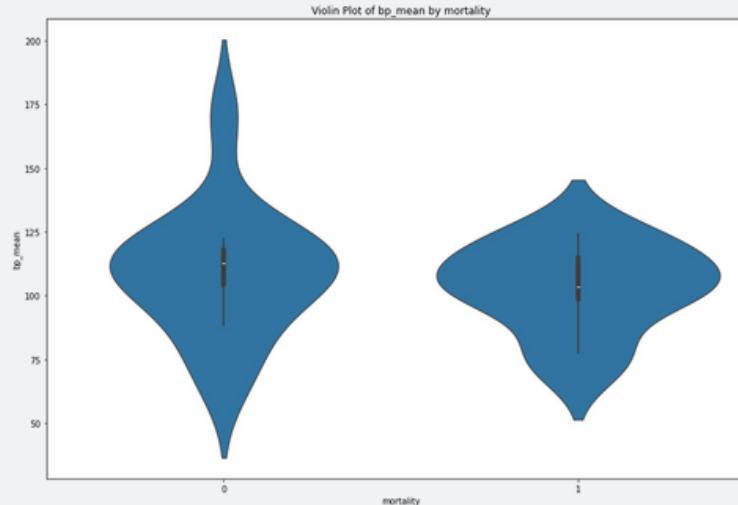
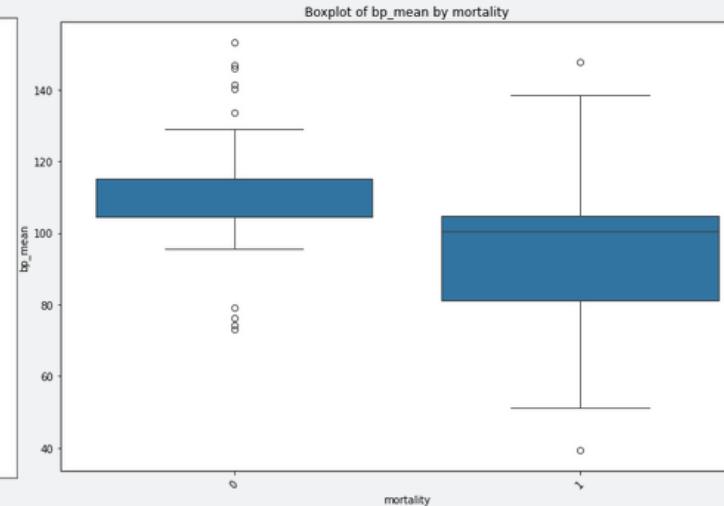
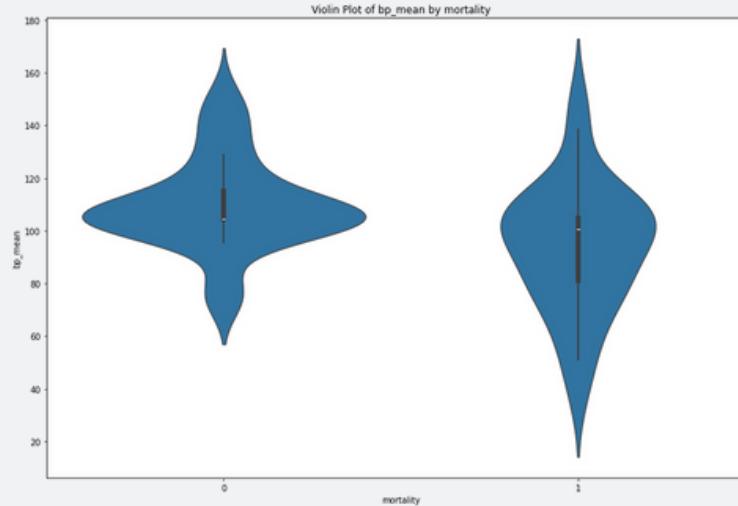
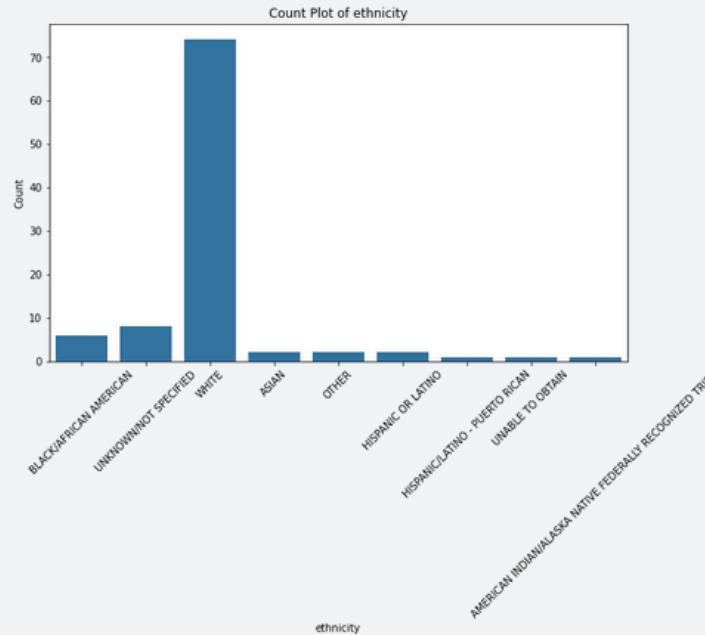
הצלחנו להוכיח שאם  
החוליה איננו נגוע בחידוק  
חיסוני והוא איננו נוצרי.  
**קיים קשר מובהק**  
סטטיסטית בין תמותה  
ולחץ דם



UN

IS

## בדיקת קשרים בין משתנה המטרה למשתנים החזויים בגלל פילוח האוכלוסייה בחרכו להתמוך לבנים/ אחר ethnicity



♦ אנו יכולים לראות שם  
מצאו האתני של החולה  
**לבן, יש קשר בין לחץ**  
**הדם שלו לבין התמותה.**

W

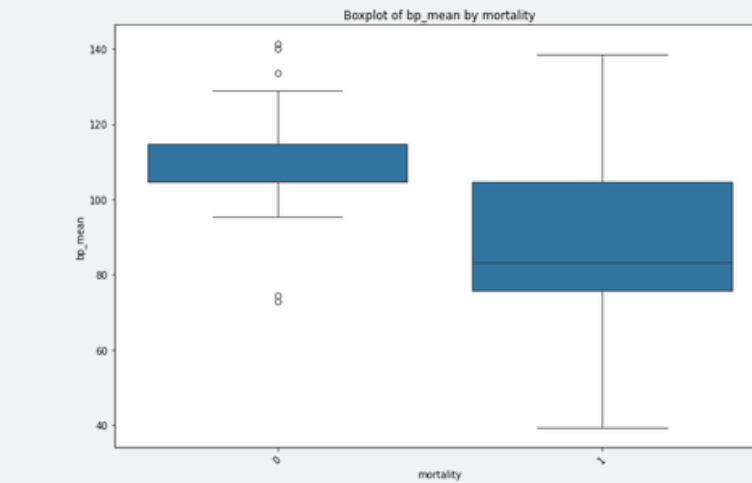
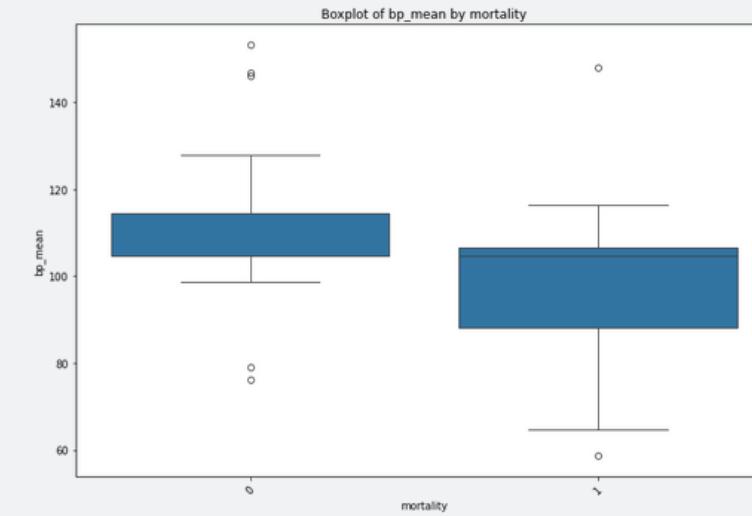
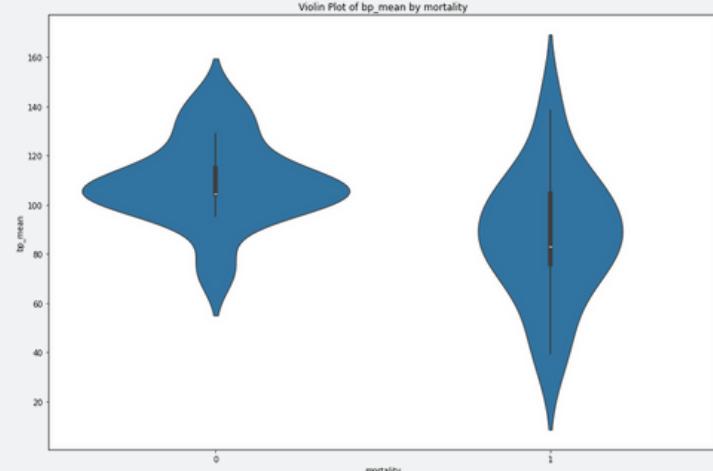
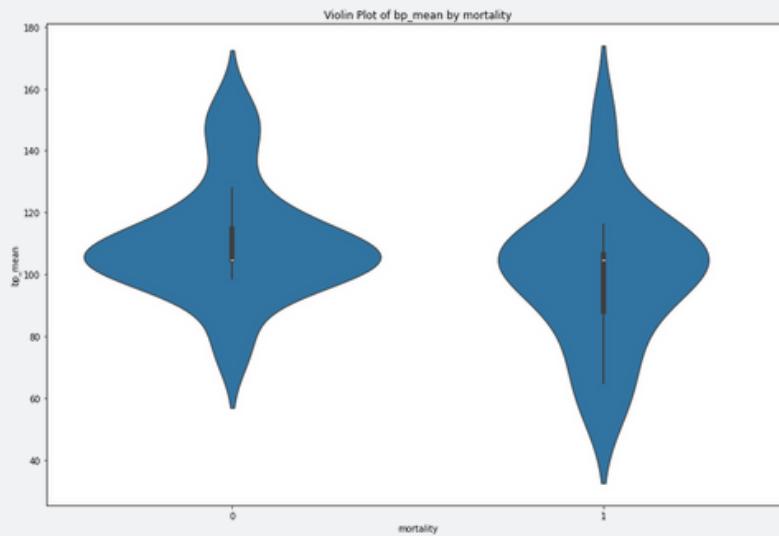
OTH

## בדיקות קשרים בין משתנה המטרה למשתנים החזויים

במה שקדם למסקנה הקודמת - אנשים לבנים בחרנו להעמיק את הקשר ולבוחן אותו מול מגדר

ניתן לראות שאם החולים הוא זכר ומוצאו האתני של החולים לבן, יש קשר בין לחץ הדם שלו לתמותה.

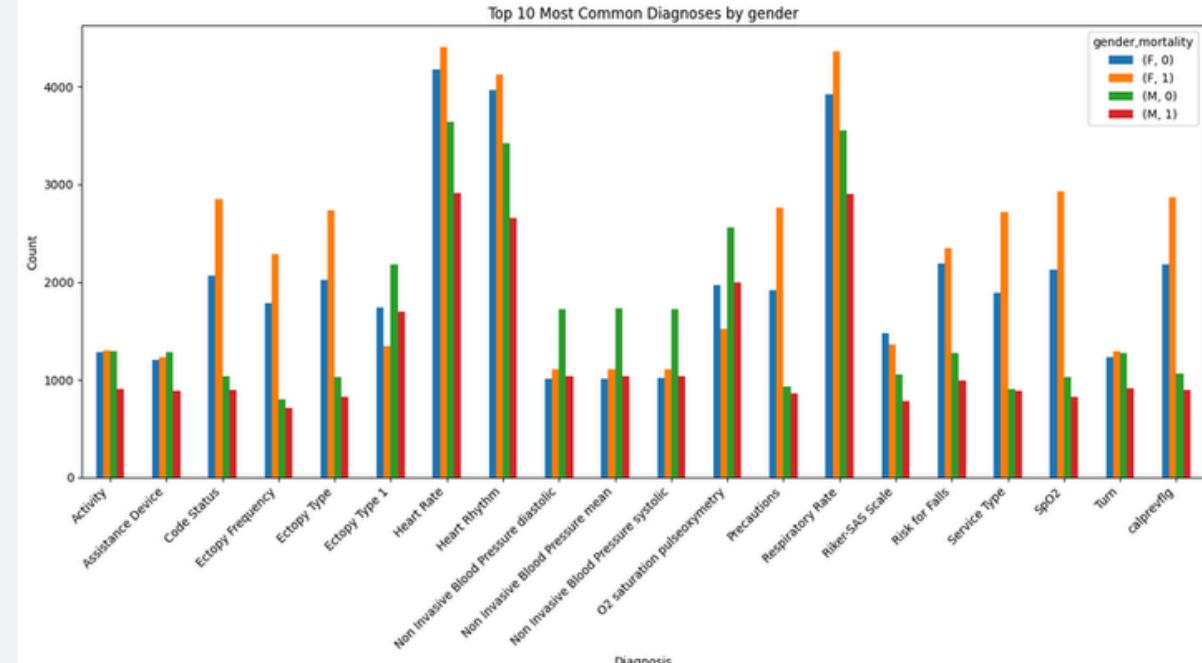
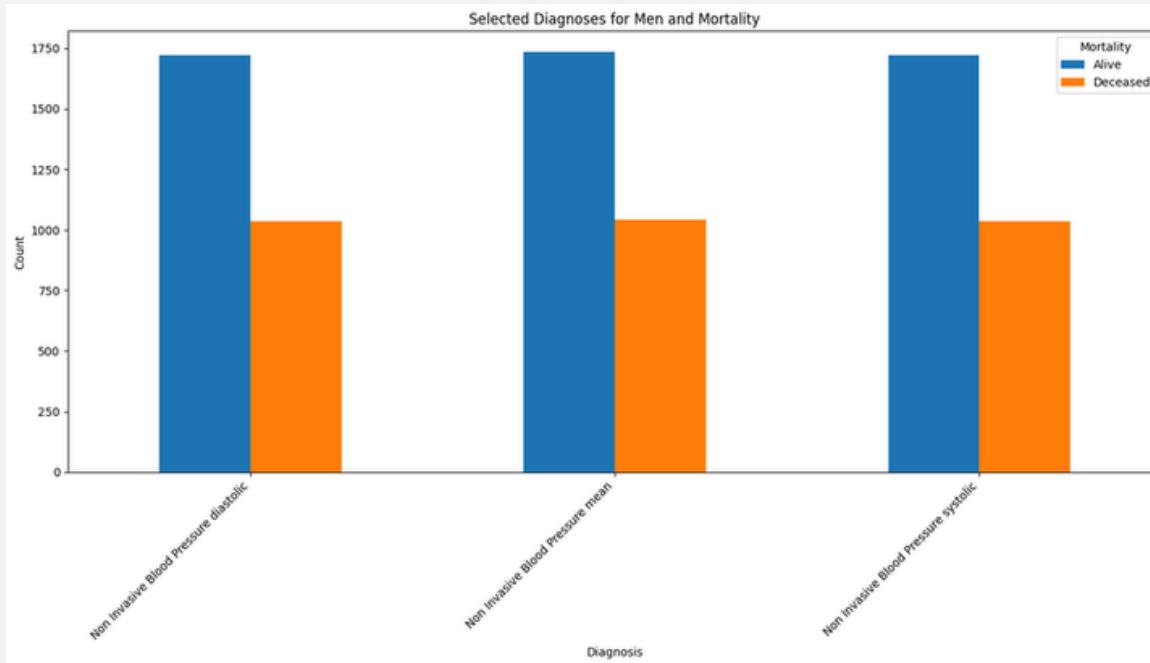
מסקנה זו מחזקת את המסכנות הקודמות כי לגברים עם מאפיינים מסוימים יש קשר מובהק סטטיסטי בין תמותה ולחץ דם



F

M

## בדיקת קשרים בין משתנה המטרה למשתנים החזויים (gender) בחרנו להסתכל על משתנה זה ולהעמיק את הקשר שלו עם משתנה גיל (list\_of\_labels)



הטונה כי לחץ הדם הממוצע השפיעה משמעותית על תמותה של חולמים גברים ביחס לטיפול נמרץ מתחזקת מההתבוננות שיש הבדל ניכר בבדיקות לחץ הדם בין הזכרים ( מגדר = "M" ) שרדן. (תמותה = 0 ) והזכרים שלא.



ראינו שיש השפעה ישירה בין תמותה למצב ההדבקה של החולים בחידק חיסוני (לפי בדיקת החי בריבוע), החולים שנדבקו בחידק נוטים לשרוד זמן רב יותר ביחס לטיפול נמרץ.

בדקנו את כל המשתנים המספריים מול תמותה וראינו שרק לחץ דם יש קשר מובהק לתמותה (לפי t-test).

כשהעמוקנו את הקשר ראיינו כמה מסקנות:

- ❖ הקשר בין לחץ דם לתמותה הוא משמעותי אם המטופל הוא זכר (לפי עמודת המגדר).
- ❖ הקשר בין לחץ דם לתמותה הוא משמעותי אם החולים אינו נגע בחידק חיסוני ומינו זכר (לפי העמודות is\_infected ו מגדר).
- ❖ הקשר בין לחץ דם ותמותה משמעותי עבור החולים שאינם נצרים ולאינם החולים בחידק חיסוני (על פי העמודות is\_infected והדת).
- ❖ הקשר בין לחץ דם ותמותה משמעותי עבור החולים גברים שהמוצא האתני שלהם לבן (על פי עמודות האתניות והמגדר).



THANK YOU

---

