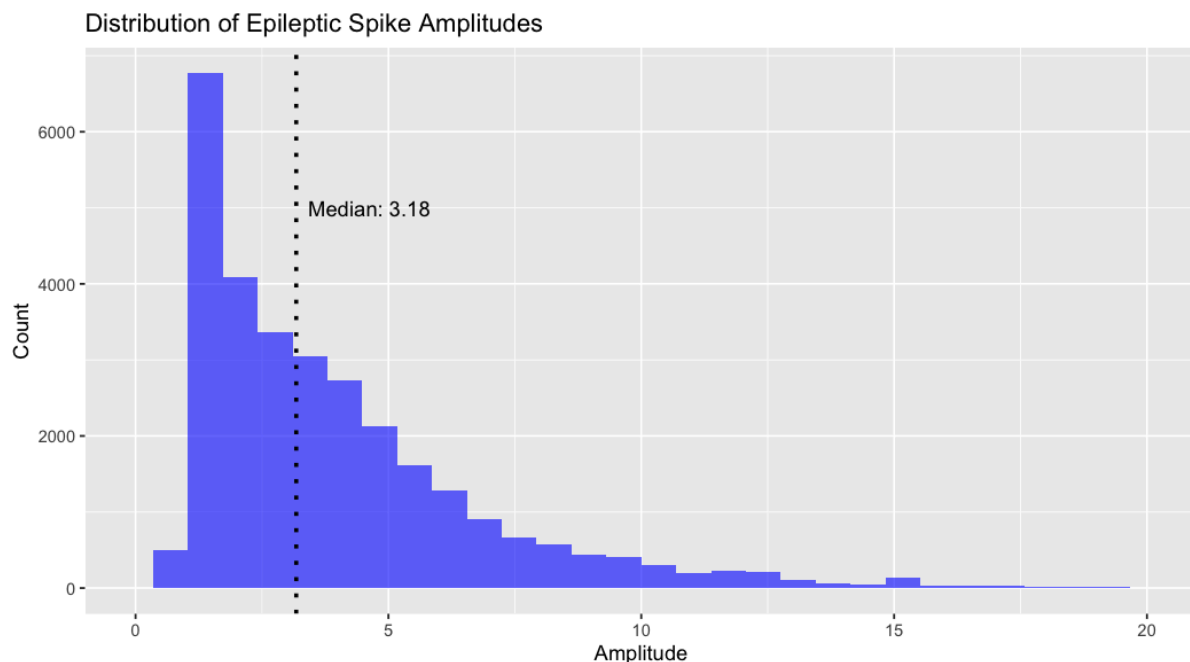


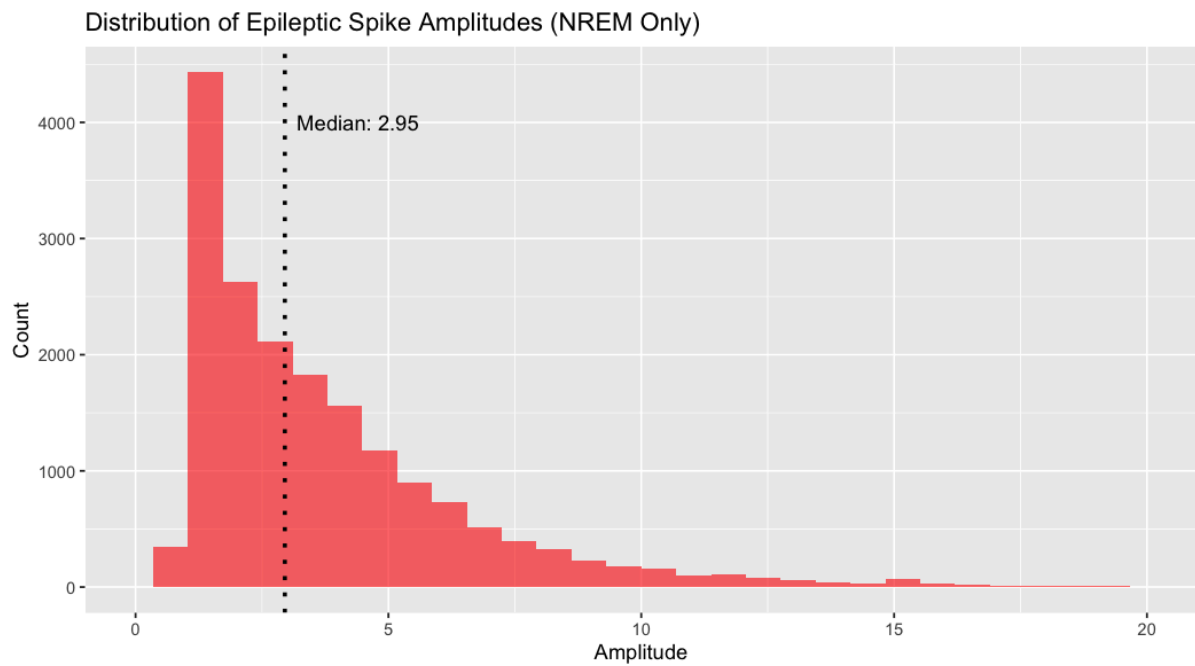
עבודה מסכמת

שלב א'

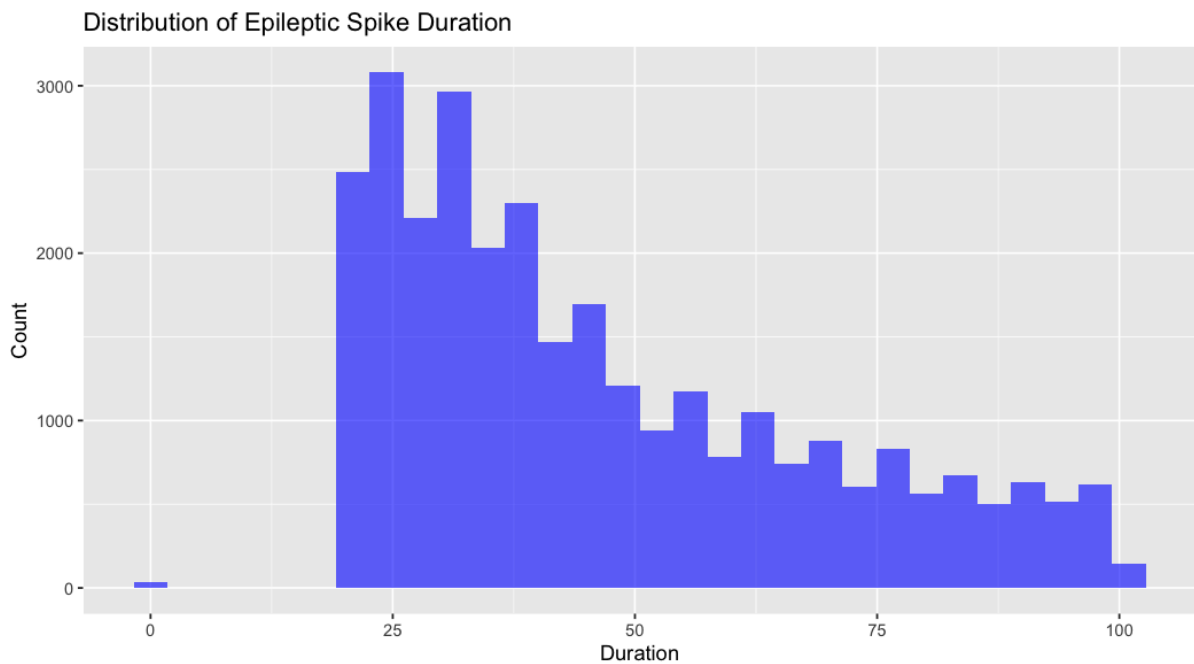
1. הדוקטורט שלי עוסק בניטור לא-פולשני של interictal spikes – אירועים אפילפטיים סאב-קליניים שמתקיימים בין התקפים ותדירותם עולה משמעותית בזמן שינה. לאחרונה, הם נחקרים גם בהקשרים של מחלות נוירולוגיות נוספות כמו אלצהיימר ושבץ. הדאטאסט שבחרתי מבוסס על נתונים שאספנו מחולי אפילפסיה ($n=6$) עם אלקטרודות מוחיות מושתלות במגוון איזורים (בעיקר באונה הטמפורלית והפרונטלית), אשר עברו גם פרדיגמת גירוי חשמלי במהלך שינה עמוקה. הנתונים כוללים מאפייני ספייקים כמו אמפליטודה, משך, מיקום מרחבי, ונפח התפשטותם באירועים קבוצתיים (כלומר ספייק שנמצא במספר ערוצים בו זמנית). בחרתי להשתמש בדאטאסט זה כדי לבחון את השפעת הגירוי על פעילות אפילפטית בזמן שינה. אין ספרות על השפעת גירוי באזור הספציפי שבחרנו (החומר הלבן של האונה הפרונטלית), ורוב הידע הקיים עוסק בהתקפים עצמם, ולא בפעילות האינטראיקטלית הקשורה לירידה קוגניטיבית ארוכת טווח. ניתוח זה יכול לספק תובנות חדשות על השפעות גירוי חשמלי בעומק המוח ולתרום להבנת מנגנוני הטיפול הקיימים באמצעות (Deep brain stimulation (DBS.
2. כדי לחוש את הדאטה בחרתי לייצא גרפים על מאפיינים שונים של הספייקים כדי לבחון התפלגות כללית, בין שלבי שינה, ובשלבים שונים של הגירוי החשמלי. להלן ניתוחים שביצעתי:
 - התפלגות האמפליטודה של הספייקים:
 - היסטוגרמה המציגה את התפלגות ערכי האמפליטודה עבור כלל הספייקים שנמדדו. הקו המקווקו מציין את חציון האמפליטודה, והציר מוגבל לערכים בין 0 ל-20 על מנת להתמקד בטווח הרלוונטי ביותר.



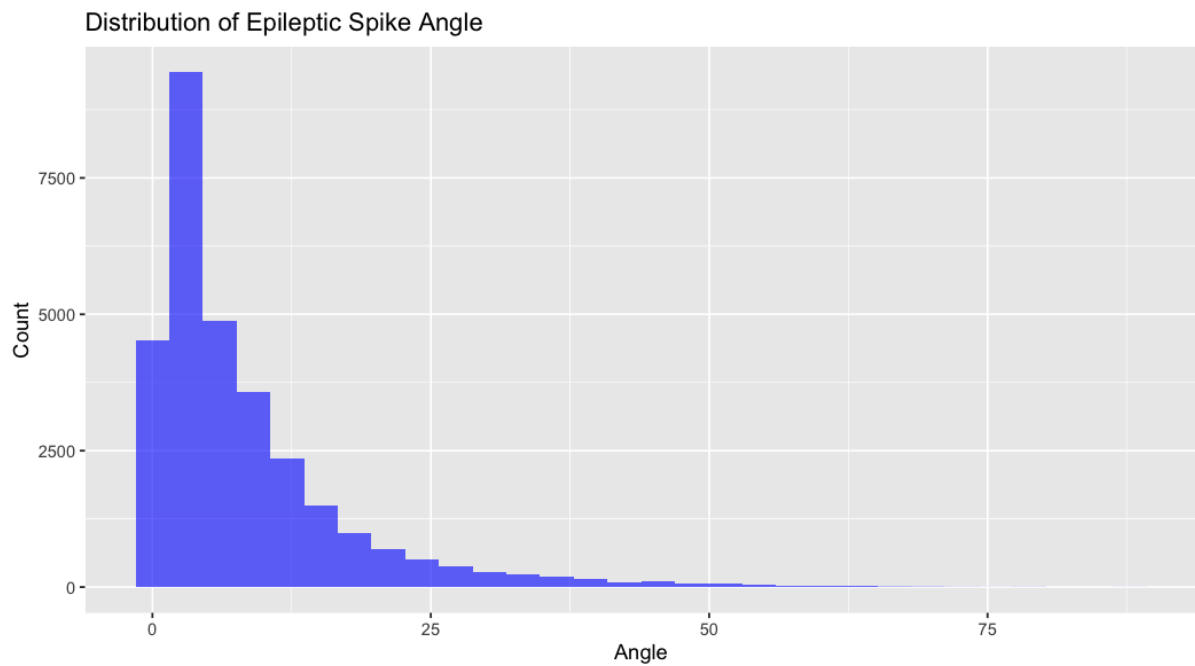
- התפלגות נוספת (אותו גרף כמו קודם), אך הפעם מוצגים רק הספייקים שהתרחשו במהלך שנת NREM. החציון מסומן שוב כדי לאפשר השוואה בין ההתפלגות הכללית לבין זו של שלב השינה הספציפי (לא נראה שיש הבדל מהותי).



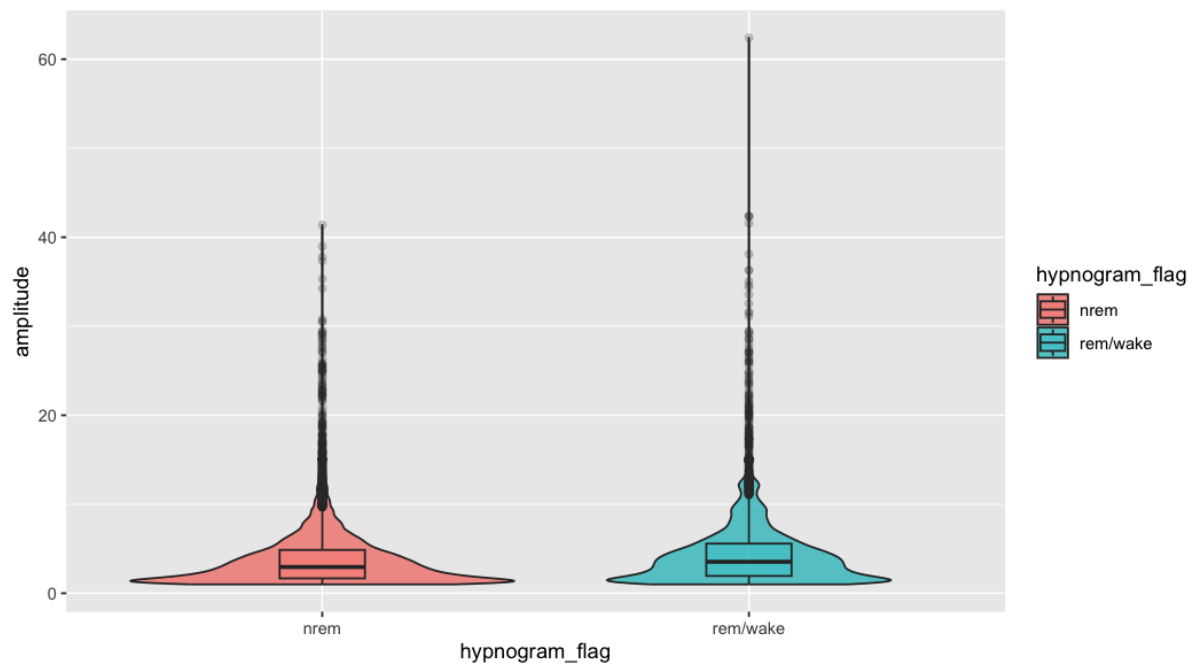
- היסטוגרמה המציגה את משך הזמן של הספייקים בכל הנבדקים. התפלגות זו מסייעת לזהות אם משכי הספייקים הם אחידים, מוטים, או מציגים כמה פיקים נפרדים. רואים שחלק מתנאי הזיהוי של הספייקים הם משך גדול מ-20.



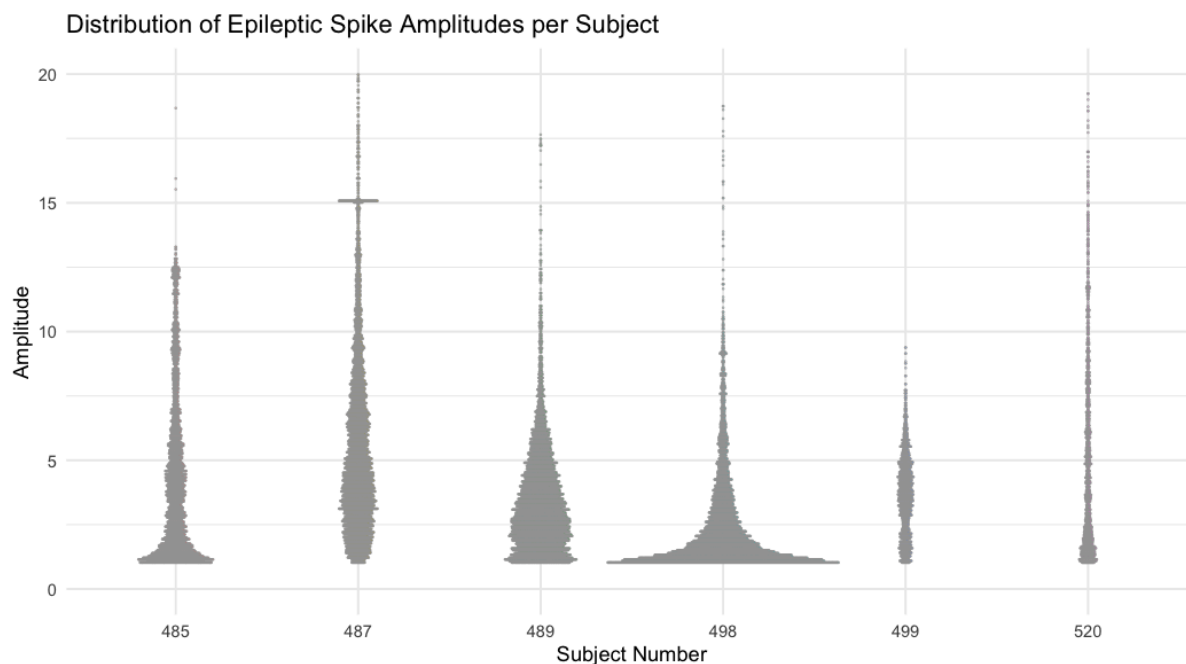
היסטוגרמה של הזוויות של הספייקים- כלומר מבחינה מורפולוגית כמה הספייק חד.



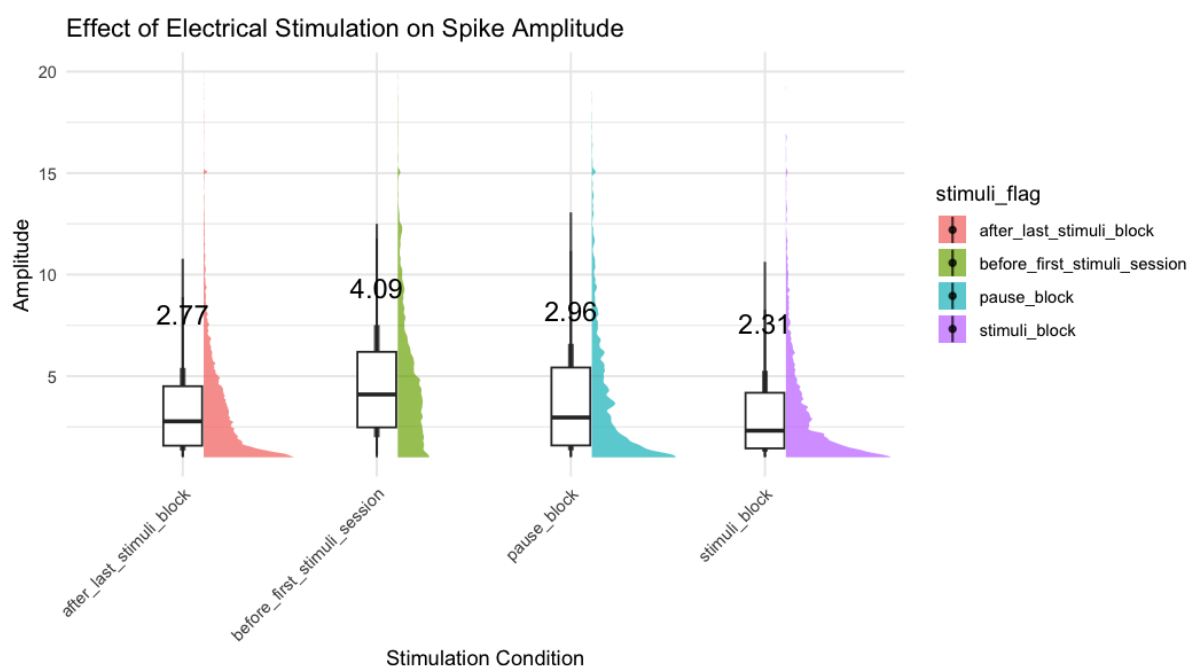
גרף Violin + Boxplot המציג את התפלגות אמפליטודת הספייקים בין שלבי השינה השונים. הוויאולין מציג את צפיפות הערכים בכל שלב שינה, בעוד שהתיבה מציגה את המגוון הסטטיסטי של הערכים (רבעונים וחציון).



גרף stat_dots המציג את התפלגות האמפליטודה עבור כל נבדק בנפרד. מטרת הגרף היא לאפשר השוואה בין נבדקים שונים ולבחון את השונות בין האינדיבידואלים במחקר. כצפוי בקרב אוכלוסיות פתולוגיות, התפלגויות שונות אחת מהשניה.

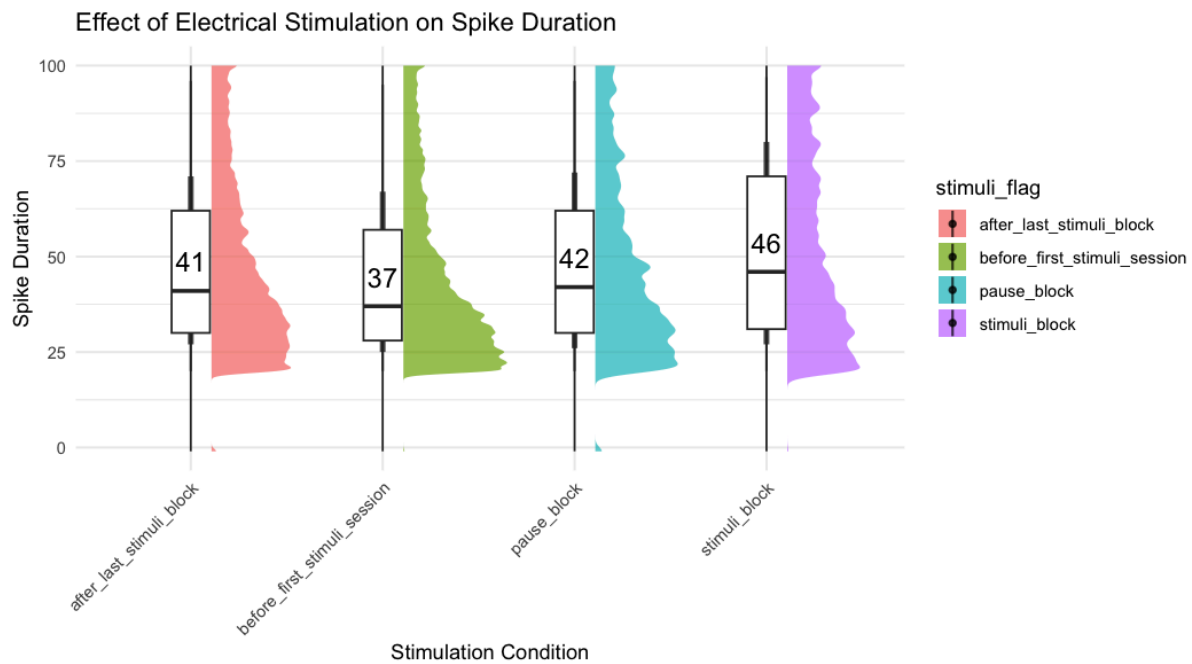


גרף Half-eye + Boxplot המשווה את התפלגות האמפליטודה בין מצבי הגירוי השונים. צורת ה-Half-eye מציגה את התפלגות הערכים, בעוד שה-Boxplot מספק מבט סטטיסטי על מגמות הנתונים. ציר Y מוגבל בין 1 ל-20, כדי להתמקד בהבדלים בטווח הרלוונטי ביותר. אפשר לראות שההתפלגות הירוקה שהיא תחילת הלילה נראת שונה מהאחרות, עם חציון גבוה יותר.

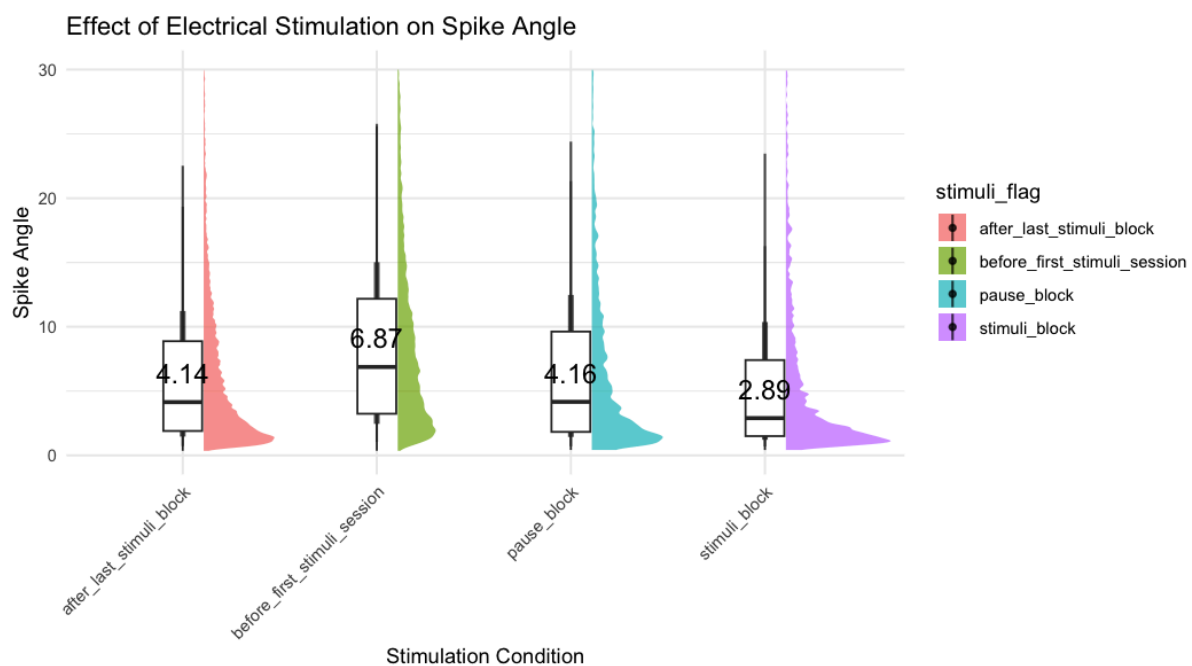


באמצעות גרפים דומים לאלה של האמפליטודה, בוצעה בדיקה של השפעת הגירוי החשמלי על המשתנים הבאים:

- משך הספייקים

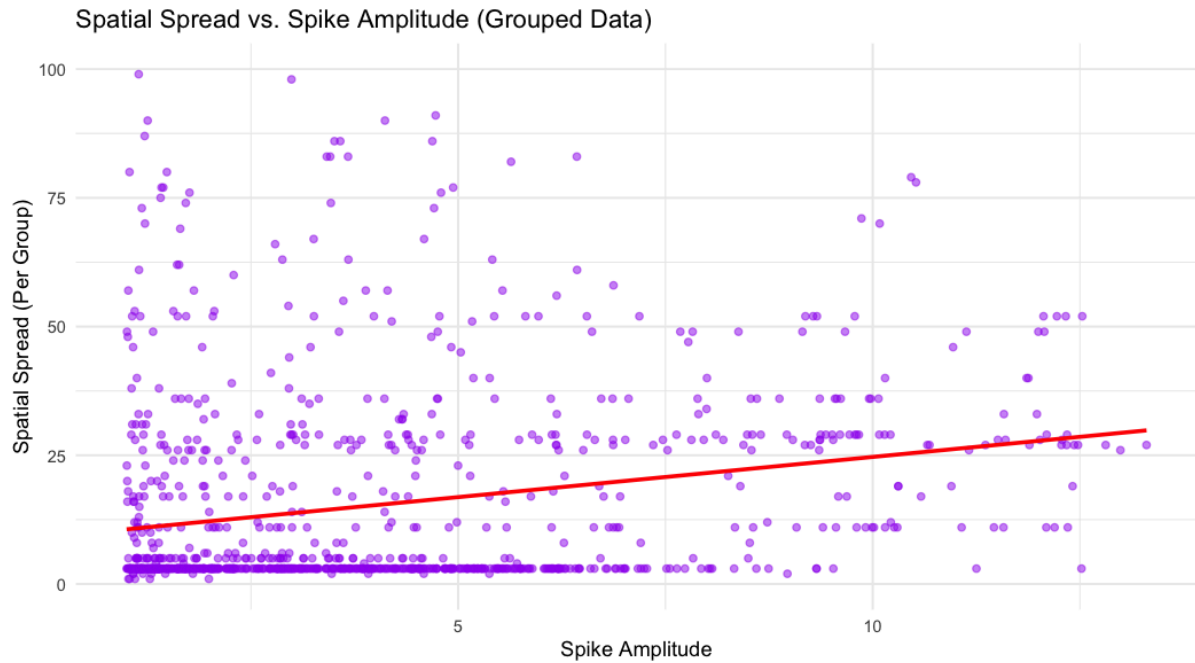


• זווית הספייקים

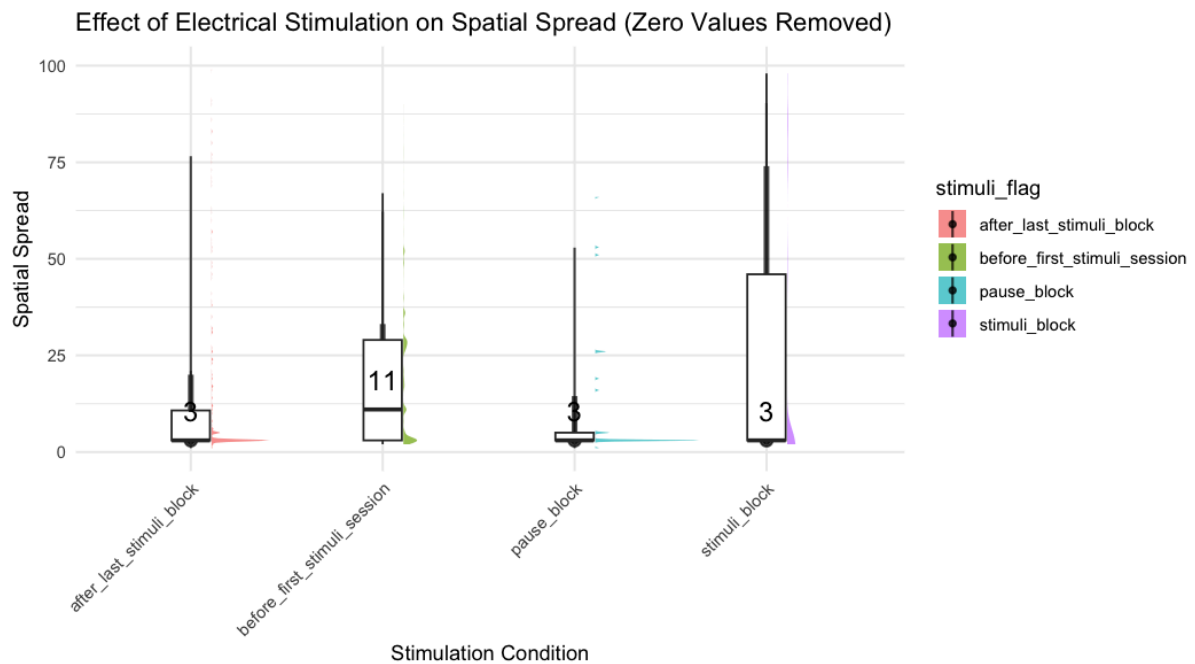


עכשיו נסתכל על ערכים של **קבוצות** ספייקים. כלומר כל הספייקים שקרו במרחק פחות מ-200 מילישניות מספייק אחר קיבלו את אותו מספר קבוצה, וככה אפשר למדוד מאפיין של התפשטות מרחבית- בכמה ערוצים האירוע הפתולוגי הופיע- ולחשב את הנפח של הצורה שהערוצים יוצרים (יש גם ספייקים שמופיעים מקומית רק בערוץ אחד ולכן ההתפשטות שלהם תהיה 0).

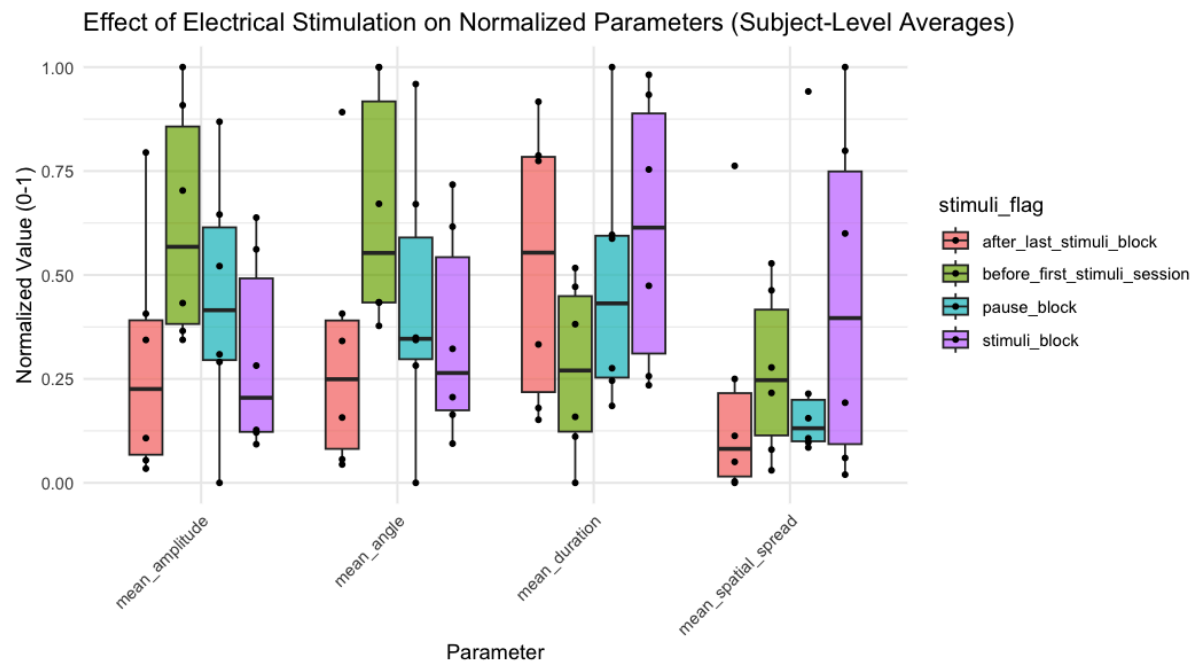
הגרף הבא מציג את הקשר בין אמפליטודת מוקד הספייק לבין ההתפשטות המרחבית של האירוע, כאשר כל נקודה מייצגת אירוע קבוצתי של ספייקים. קו הרגרסיה (באדום) מספק אינדיקציה למגמה בין המשתנים, ועוזר לבחון האם ספייקים בעוצמה גבוהה יותר נוטים להתפשט על פני אזורים נרחבים יותר במוח.



שילוב של Half-eye + Boxplot עם חציון להצגת ההשפעה של הגירוי על ההתפשטות המרחבית תוך הדגשת המגמות המרכזיות. ניתן לראות שההתפשטות ירדה משמעותית לאחר הגירוי, אבל השונות יותר גבוהה בזמן הגירוי עצמו. בנוסף כמות הדגימות ירדה משמעותית (בערך 15 אחוז).

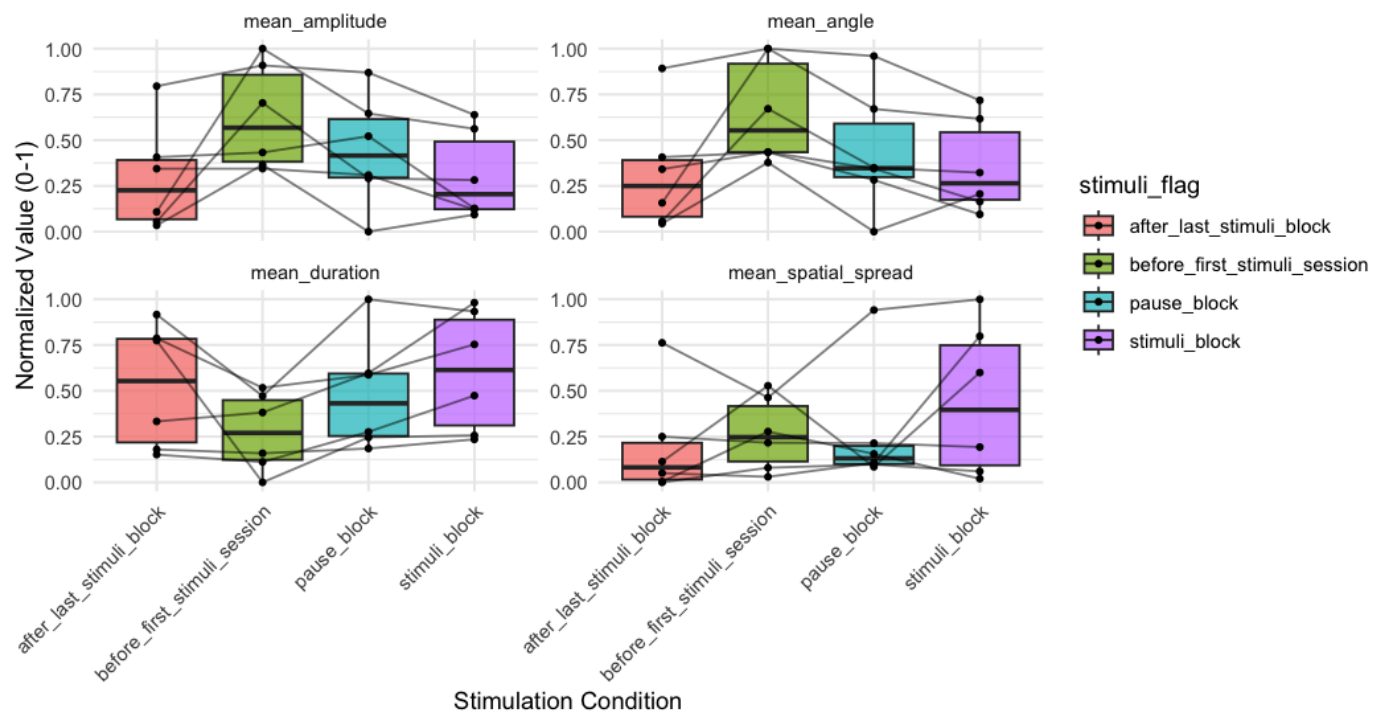


בגלל שאנחנו יודעים שהנבדקים שונים אחד מהשני, נרצה לראות אם התופעות שגילינו קורות גם ברמת כל נבדק ולא רק ברמת הממוצע הכללי. לכן יצרתי גרף Boxplot המציג את ממוצע הפרמטרים עבור כל נבדק בשני תנאי גירוי שונים. כל נקודה מייצגת את הממוצע של נבדק ספציפי. בשל הבדלים בסקאלות בין הפרמטרים, בוצע נירמול (0-1 Min-Max), המאפשר השוואה ישירה בין הפרמטרים השונים. רואים שהנתונים בירוק (לפני גירוי) הם המקסימלים באמפליטודה והזווית ביחס לתנאים האחרים, ומינימלים מבחינת משך הזמן.



גרף `facet_wrap`, שבו כל פרמטר מוצג בנפרד, עם קווים המחברים בין ערכי אותו נבדק כדי להמחיש מגמות אינדיבידואליות בהשפעת הגירוי. אפשר לראות שיש מגמות שחוזרות על עצמן בכל הנבדקים כמו ירידה באמפליטודה ובזווית לאחר ותוך כדי גירוי, ועליה במשך הזמן.

Effect of Electrical Stimulation on Normalized Parameters (Subject-Level Averages)



3. שאלת המחקר:

How do sleep stage, spike duration, spatial spread, and electrical stimulation affect "the amplitude of interictal spikes in epilepsy patients" (כיצד שלב השינה, משך הספייק, ההתפשטות המרחבית, והגירוי החשמלי משפיעים על אמפליטודת הספייקים האפילפטיים?)

שלב ב'

1. **משתנה מנובא:** אמפליטודת של ספייק אפילפטי (group_focal_amplitude) - זהו משתנה רציף המתאר את עוצמת הפעילות החשמלית של הספייקים (במידה ומדובר בקבוצת ספייקים זהו הספייק הראשון שהתרחש). נשתמש בו כרציף בניתוח רגרסיה ליניארית ונהפוך אותו למשתנה קטגורי עבור רגרסיה לוגיסטית כלומר נבדוק מה מנבא אמפליטודת גבוהה לעומת נמוכה.
משתנים מנבאים:
 - גירוי חשמלי (stimuli_flag) - פקטור שיומר ל-0 לפני תחילת גירוי ו-1 אחרי גירוי (איחוד של 3 מצבים שהופיעו בדאטה המקורי: בזמן הגירוי, בהפסקות שבין הגירוי, ולאחר הגירוי האחרון).
 - שלב שינה (hypnogram_flag) - פקטור שיומר ל-0 עבור שינה עמוקה (NREM) ו-1 עבור אחר.
 - משך הספייק (duration) - משתנה רציף המייצג את משך הזמן של הספייק במילישניות.
 - התפשטות מרחבית (group_event_spatial_spread) - משתנה רציף, המתאר את מידת ההתפשטות של הפעילות האפילפטית במוח.
2. נבצע המרה של משתנים קטגוריים לפקטורים, ננרמל את המשתנים הרציפים, ונוסיף משתנה בינארי קטגוריאלי מנובא המסמן אמפליטודת מעל החציון כ-1 ומתחת לחציון כ-0 עבור הרגרסיה הלוגיסטית. ככה הנתונים נראים כעת:

processed_data	30121 obs. of 6 variables
\$ stimuli_flag	: Factor w/ 2 levels "before_stimulation",...: 1 1 1 1 1 1 1 1 1 ...
\$ hypnogram_flag	: Factor w/ 2 levels "nrem", "rem/wake": 2 2 2 2 2 2 2 2 2 ...
\$ duration_z	: num [1:30121, 1] -0.814 -0.493 0.425 0.838 -0.906 ...
..- attr(*, "scaled:center")=	num 46.7
..- attr(*, "scaled:scale")=	num 21.8
\$ spatial_spread_z	: num [1:30121, 1] -0.416 -0.416 -0.32 -0.32 -0.416 ...
..- attr(*, "scaled:center")=	num 121
..- attr(*, "scaled:scale")=	num 292
\$ high_amplitude	: num [1:30121] 0 1 0 0 1 1 0 0 0 ...
\$ group_focal_amplitude	: num [1:30121] 1.84 3.41 2.35 2.35 9.9 ...

3. כדי להפוך את תהליך העיבוד לגמיש וניתן לשימוש חוזר, ניצור פונקציה שתאפשר לנו להפעיל את אותם שלבי עיבוד על קבצי נתונים נוספים (למשל עבור נבדקים חדשים שנאספו).
4. נשתמש בספריית recipe שמאפשרת לעבד נתונים בצורה מודולרית ומסודרת וליצור pipelines. כלומר במקום לבצע עיבוד ישירות עם scalei mutate אנחנו יוצרים "מתכון" שמעבד את הנתונים בשלבים. זה מאפשר שלבים מובנים וגמישות לשימוש חוזר. מימשת את עיבוד הנתונים שלי עם שימוש בספריה באמצעות יצירת מתכון ואז הפעלתי אותו על הדאטה.

שלב ג'

1. ביצעתי רגרסיה ליניארית מרובה כפי שהגדרתי בשלב ב' סעיף 1: רגרסיה עבור ניבוי אמפליטודת על סמך המשתנים הפקטוריאליים של שלב שינה, לפני/אחרי גירוי, משך הספייק והתפשטות מרחבית. עבור רגרסיה ליניארית השתמשנו באמפליטודת כמשתנה רציף ועבור רגרסיה לוגיסטית המרנו את הערכים ל-0 ו-1 כאשר ערכים נמוכים מתחת לחציון הם 0 וערכים גבוהים מעל החציון הם 1.

2. תוצאת הרגרסיה הליניארית:

```
Residuals:
    Min       1Q   Median       3Q      Max
-4.439 -1.791 -0.785  0.918 57.783

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    4.55290    0.03590 126.818 <2e-16 ***
stimuli_flag1 -1.29340    0.03691 -35.045 <2e-16 ***
hypnogram_flagrem/wake 0.06266    0.03526  1.777  0.0756 .
duration_z     -0.45761    0.01661 -27.558 <2e-16 ***
spatial_spread_z -0.02561    0.01655  -1.548  0.1217
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.867 on 30116 degrees of freedom
Multiple R-squared:  0.07147, Adjusted R-squared:  0.07135
F-statistic: 579.5 on 4 and 30116 DF, p-value: < 2.2e-16
```

פירוש:

כל מקדם מתאר את השינוי הצפוי באמפליטודה כתוצאה משינוי ביחידת מידה אחת של המשתנה המנבא, בהנחה שכל שאר המשתנים קבועים. משמעות ה-intercept- כאשר כל המשתנים שווים ל0 (המשתנים הקטגוריאליים בנקודת הייחוס שהיא לפני גירוי בשינה עמוקה), האמפליטודה הצפויה היא 4.55. הגירוי החשמלי מפחית בממוצע את האמפליטודה ב1.29 יחידות, מובהק מאוד והוא בעל ההשפעה הגדולה ביותר. מבחינת משך הזמן האפקט הוא מובהק וככל שהספייק ארוך יותר כך האמפליטודה נמוכה יותר. המקדם של שלב השינה יצא לא מובהק (אולי בעל טרנד מסוים), אך משמעותו היא שבשלב שינה לא עמוקים הספייקים מעט גדולים יותר. לא נמצא קשר עבור ההתפשטות המרחבית. ערך R נמוך יחסית, מה שמעיד שהמודל מסביר חלק קטן מהשונות ויש גורמים נוספים שלא נלקחו בחשבון.

תוצאת הרגרסיה הלוגיסטית:

```
Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    0.54081    0.02635  20.524 < 2e-16 ***
stimuli_flag1 -0.83711    0.02699 -31.011 < 2e-16 ***
hypnogram_flagrem/wake 0.07109    0.02554  2.783  0.00538 **
duration_z     -0.39791    0.01238 -32.139 < 2e-16 ***
spatial_spread_z -0.05897    0.01218  -4.843 1.28e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 41757 on 30120 degrees of freedom
Residual deviance: 39365 on 30116 degrees of freedom
AIC: 39375

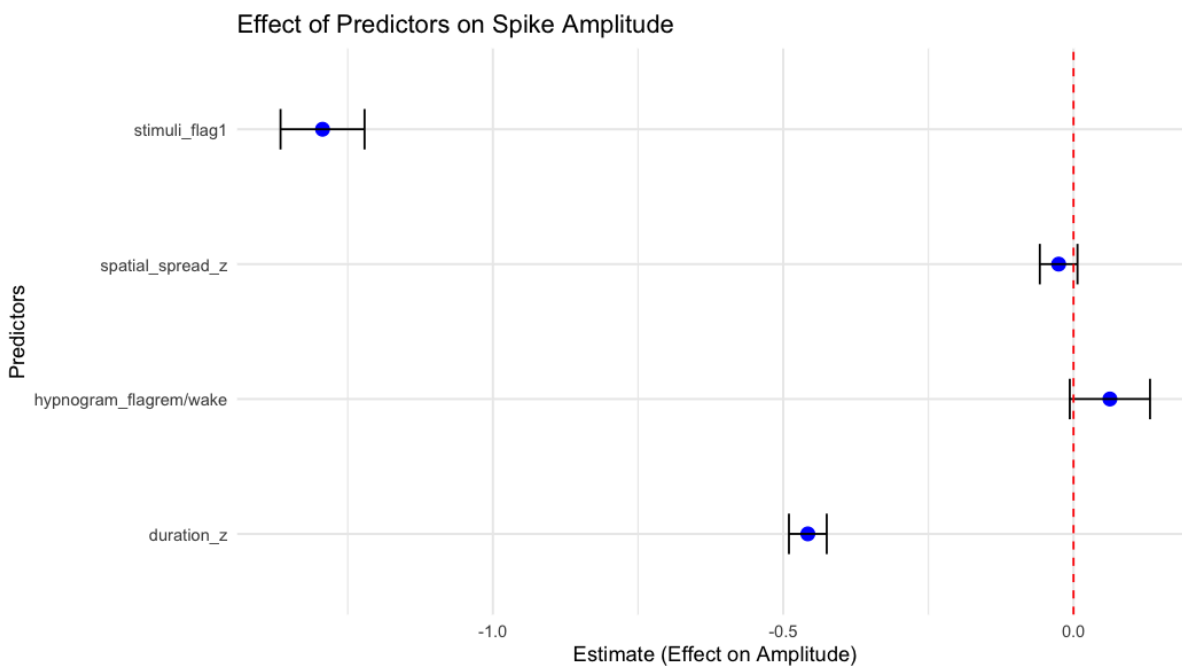
Number of Fisher Scoring iterations: 4
```

פירוש:

כל מקדם מתאר את ההשפעה של המשתנה המנבא על הסיכוי לאמפליטודה גבוהה, כאשר שאר המשתנים נשארים קבועים. אם הערך חיובי אז הסיכוי לאמפליטודה גבוהה עולה, ואם שלילי אז הסיכוי יורד.

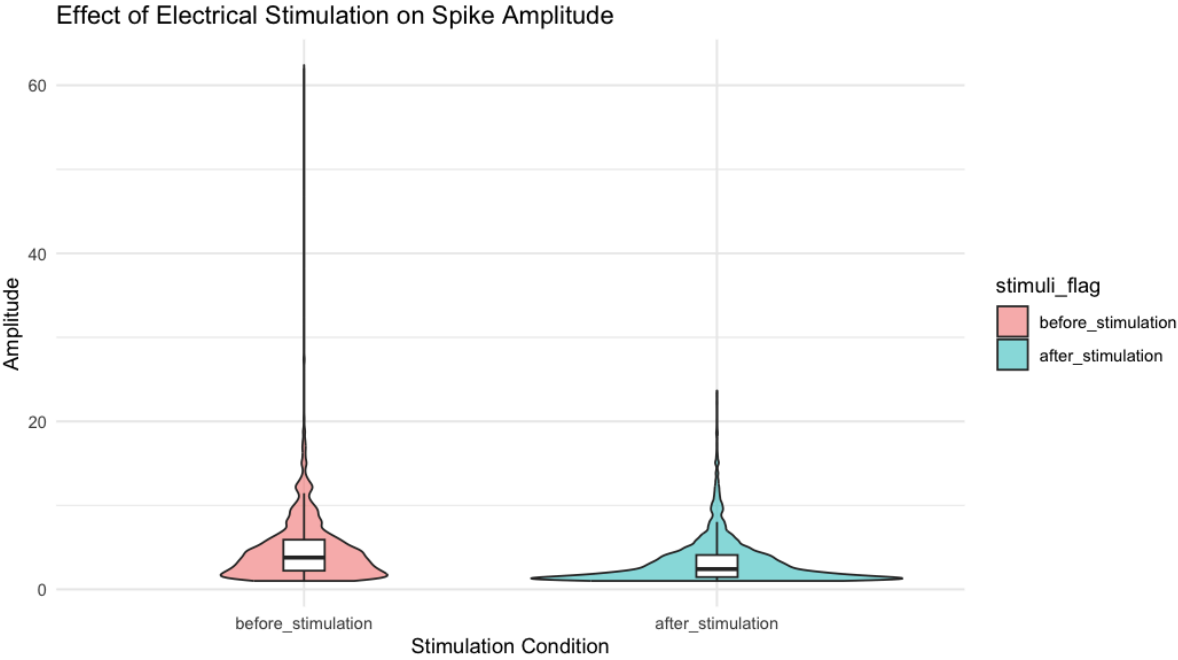
משמעות הintercept- כאשר כל המשתנים המנבאים שווים ל0 (המשתנים הקטגוריאליים בנקודת הייחוס שהיא לפני גירוי בשינה עמוקה), קיים סיכוי גבוה יותר לאמפליטודה גבוהה בגלל שהערך הוא חיובי (נראה לאחר חישוב מדויק שהסיכוי הוא 63%).
 הגירוי החשמלי מפחית משמעותית את הסיכוי לאמפליטודה גבוהה, מובהק מאוד והוא בעל ההשפעה הגדולה ביותר.
 בשונה מהמודל הליניארי, המקדם של שלב השינה יצא מובהק, ובאופן דומה משמעותו היא שבשלב שינה לא עמוקה, הספייקים מעט גדולים יותר. ההשפעה יחסית קטנה.
 מבחינת משך הזמן האפקט הוא מובהק ובעל השפעה יחסית חזקה, ככל שהספייק ארוך יותר כך גדל הסיכוי לאמפליטודה נמוכה יותר.
 ההשפעה של התפשטות מרחבית היא הקטנה ביותר והקשר הוא שככל שהיא גדולה יותר כך קטן הסיכוי לאמפליטודה גבוהה.

3. האפקטים ברגרסיה הליניארית:

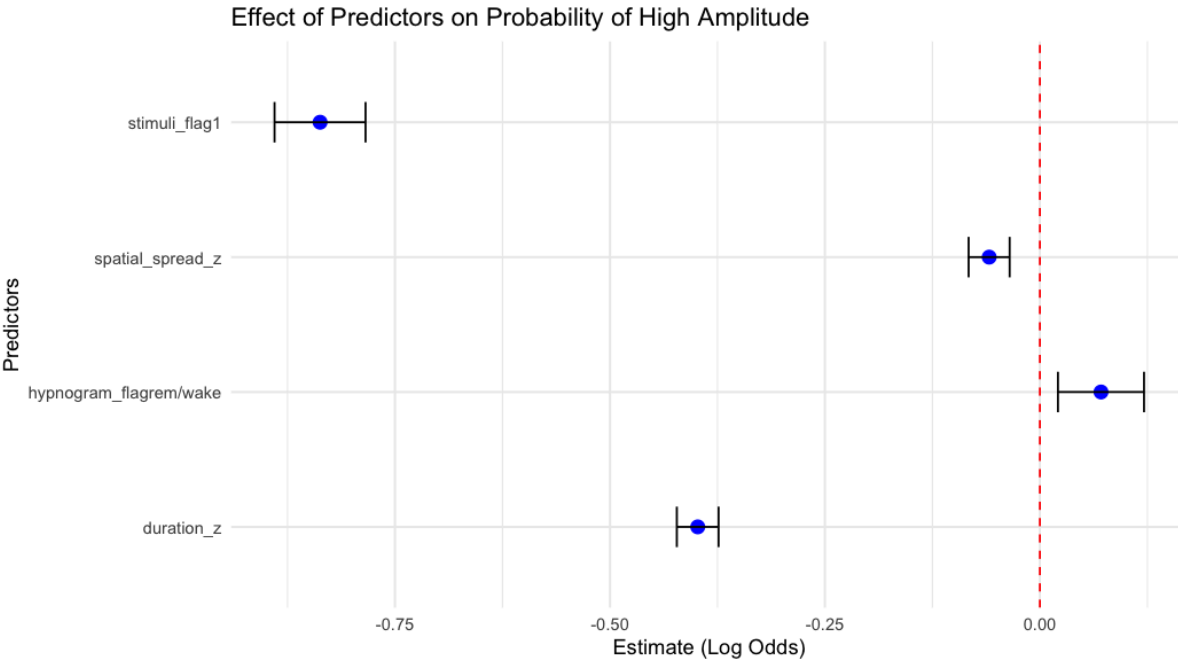


הגרף מציג את ערכי המקדמים יחד על רווח סמך שבו ניתן לראות האם הטווח מכיל את 0 או לא ולקבוע את מובהקות המקדם.
 מצאנו שמצב הגירוי (לפני או אחרי) בעל השפעה מרובה על חיזוי האמפליטודה, להלן גרף המציג את

הנתונים והתפלגותם השונה בין שני המצבים:

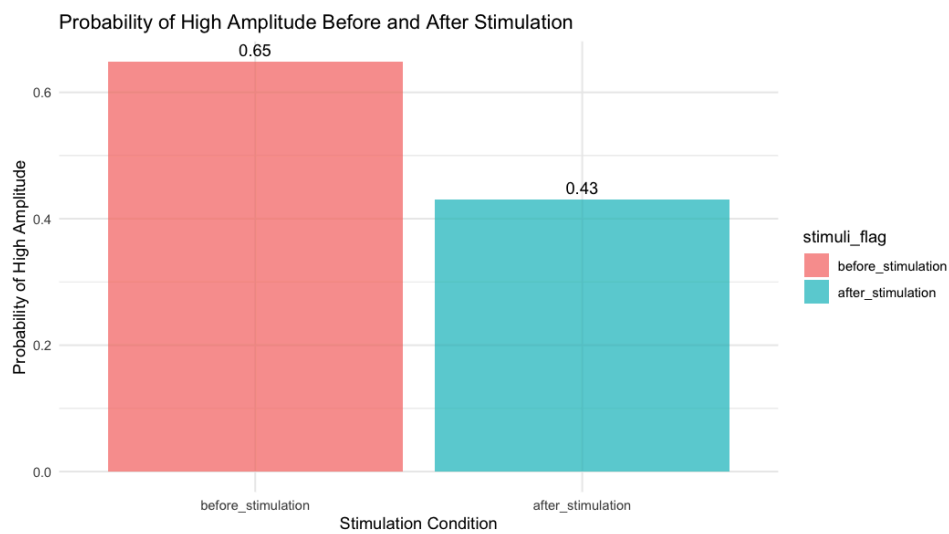


האפקטים ברגסיה הלוגיסטית (כאן כל המקדמים מובהקים):

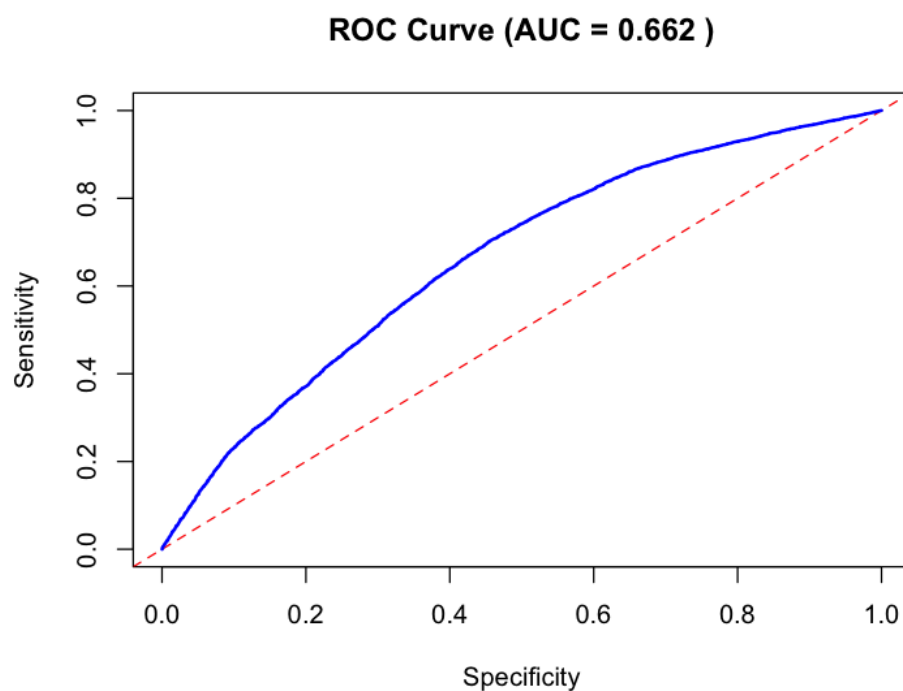


בגלל שהמרנו את המשתנה המנובא להיות בינארי, נוכל להסתכל על האפקט החזק שמצאנו שקשור למצב הגירוי (לפני/אחרי) בנתונים דרך בדיקה מה הסיכוי לאמפליטודה גבוהה בכל אחת מקבוצות

הגירוי:



4. בגרף רואים את היחס בין ספציפיות לרגישות ביחס למודל ממוצע, כאשר השטח שחושב עבור המודל מתחת לעקומה הוא 0.662.



- אסכם בכך שמצאתי פוטנציאל אמיתי בדאטה סט ואני מחכה להוסיף עוד נבדקים כדי לוודא שהאפקט של הגירוי על האירועים האפילפטיים של החולים באמת נשמר באופן חזק ויציב גם מול קבוצת קונטרול. בנוסף למדתי קצת על היתרונות של R אחרי שנים של שימוש בפייתון. אחלה פרוייקט ותודה על אחלה קורס :