**מחקר MSA - סיכום ניסוי לולאות 1#**

**1. רקע כללי**

(כאן נכתוב על מה אנחנו מנסים לבדוק, ואמל;ק על התוצאות ביחס לאפקט המרכזי)

**2. מהלך הניסוי**

(כאן נכתוב על הניסוי עצמו, הנבדקים והמאפיינים שלהם - גם מניתוח של המידע שהעבירו)

**3. עיבוד וניתוח**

להלן מתוארים שלבי הסינון והעיבוד של תוצאות הניסוי. את התוצאות הגולמיות וכל הקוד לעיבוד וניתוח של התוצאות ניתן למצוא תחת [MathThinkingLabDA\LoopDA](https://github.com/DanielVered/MathThinkingLabDA/tree/main/LoopsDA). העיבוד הראשוני נמצא ברובו בקובץ ProccessingFuncs, וניתוח התוצאות בקובץ AnalysisFuncs. אנליזה אקספלורטיבית, בעיקר לאיתור אנומליות, בוצעה בקובץ OutliersExploration.

**3.1. סינונים ראשוניים**

ראשית, הוצאנו מקובץ הנתונים את הרשומות הבאות:

* צעדי השמה של משתנה חדש (drop\_assign).
* לולאה ראשונה בכל סשן (drop\_first\_loop).

כמו כן, על מנת לבחון את אפקט ה-Priming, ביצענו את הניתוח בנפרד על השורה הראשונה בכל לולאה (שבה הוחלף או לא הוחלף סוג הלולאה, only\_first\_lines), ועל כלל השורות ביחד.

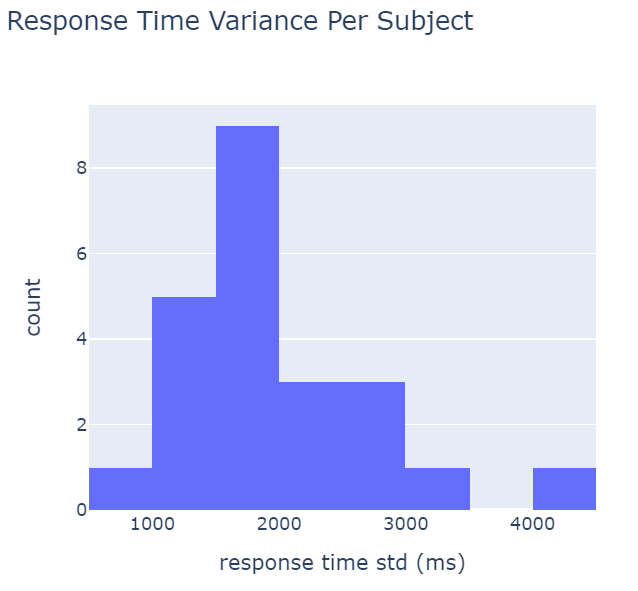
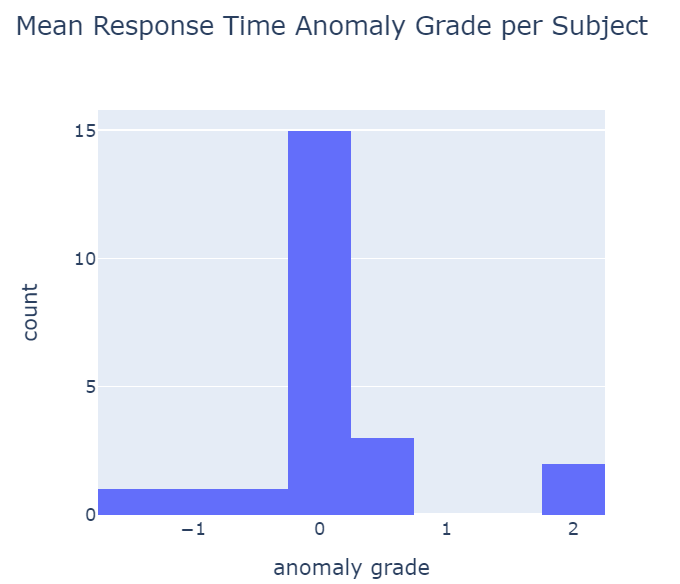
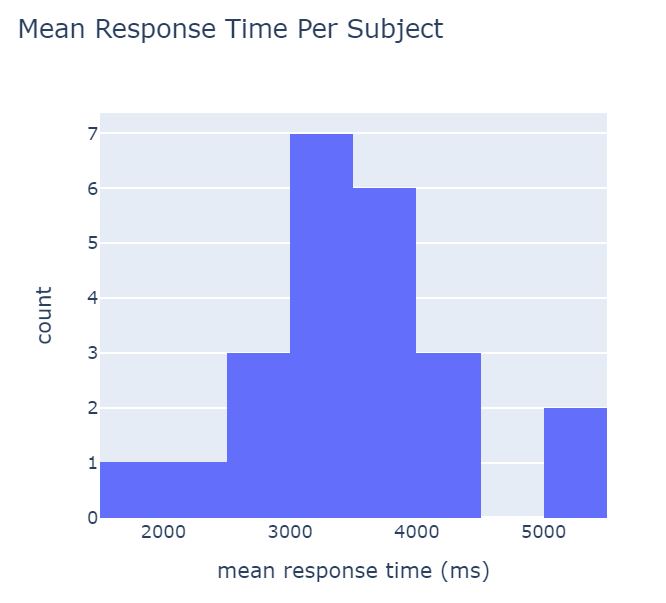
**3.2. שונות ואיתור אנומליות**

בשלב זה ביצענו מספר אנליזות במטרה לזהות חריגות שהתקבלו ברמות שונות, ולבחון את הצורך בסינון שלהם, לפני ניתוח התוצאות הסופי. שלב זה בוצע על השורות הראשונות מכל לולאה בלבד.

לאורך הניתוח, חישבנו ציון חריגה באמצעות המרחק ב-IQR מהרבעון הקרוב (). תצפיות שנמצאו בתוך הטווח הבין-רבעוני נחשבו כבעלות ציון חריגה 0.

3.2.1. נבדקים חריגים - בהיבט של זמן תגובה איטי או לחילופין אחוז הצלחה נמוך, ביחס ליתר הנבדקים.

* מבדיקת התפלגות זמני התגובה הממוצעים לכל נבדק, לא נמצאו חריגים משמעותיים, וכך גם בבחינת התפלגות שונות זמני התגובה של כל נבדק. בגרפים להלן ניתן לראות את התפלגויות אלה הן בערכים גולמיים, והן באמצעות ציוני חריגה ב-IQR.

תמונה שמכילה טקסט, צילום מסך, תרשים, גופן

התיאור נוצר באופן אוטומטי

* ככלל **אחוזי ההצלחה במטלה היו גבוהים באופן יחסי**, ובפרט מבחינת התפלגות אחוזי ההצלחה הכלליים של כל נבדק במטלה לא אותרו חריגים.

תמונה שמכילה טקסט, צילום מסך, תרשים, גופן

התיאור נוצר באופן אוטומטי

3.2.2. תכניות חריגות- בהיבט של אחוז הצלחה נמוך, ביחס ליתר התוכניות של כלל הנבדקים.

* כ-95% מכלל התכניות בוצעו באחוז הצלחה גבוה מ-85%. בסה"כ נמצאו 24 תוכניות עם אחוז הצלחה נמוך יותר, ש-22 מתוכן באחוז הצלחה גבוה מ-70%. ממצאים אלה עשויים להצביע על היותם של 24 התכניות לעיל, ובפרט 2 התוכניות האיטיות ביותר, כקיצוניות וראויות לסינון.

תמונה שמכילה טקסט, צילום מסך, תרשים, גופן

התיאור נוצר באופן אוטומטי

* בגרף הבא ניתן לראות את זמן התגובה הממוצע לתוכניות בחלוקה לאחוזי הצלחה שונים. הגרף מדגים שייתכן שקיים הבדל באפקט ה-Priming בין הקבוצות השונות. בקבוצה האמצעית, נמצא הבדל מובהק (p = 0.003), וכך גם בקבוצה הימנית (p = 0.0123).

תמונה שמכילה טקסט, צילום מסך, גופן, קו

התיאור נוצר באופן אוטומטי

* נסתייג ונציין כי כפי שמתואר בהרחבה תחת "תיאור התוצאות", נראה **שלסינון תוכניות המבוסס על אחוזי הצלחה אין השפעה משמעותית על התוצאות.**

3.2.3. צעדים חריגים - בהיבט של זמן תגובה איטי, ביחס ליתר הצעדים של אותו נבדק.

* ביחס לציון חריגה של 2.25 (IQR) אותרו כ-111 צעדים חריגים, על פני כלל הנבדקים. בהסתכלות על כלל הצעדים החריגים, נראה שקיים בהם אפקט Priming - אך הוא אינו מובהק (p = 0.187).

תמונה שמכילה טקסט, צילום מסך, גופן, תרשים

התיאור נוצר באופן אוטומטי

* עם זאת, בהסתכלות על הצעדים החריגים בחלוקה לציוני חריגה, ניתן לראות שבציונים הגבוהים **נצפה ההבדל המנוגד למצופה.**

תמונה שמכילה טקסט, צילום מסך, גופן, תרשים

התיאור נוצר באופן אוטומטי

* בבחינת מובהקות ההבדלים באמצעות מתאם פירסון, נמצא כי כלל ההבדלים אינם מובהקים   
  (p > 0.05) מלבד לצעדים בציון תקן (0.0, 1.5], בה נמצא מתאם מובהק (0.00038p = ).

תמונה שמכילה טקסט, צילום מסך, גופן, תרשים

התיאור נוצר באופן אוטומטי

* בשל הכיווניות ההפוכה באפקט בצעדים הקיצוניים, ניתן לסנן אותם מבחינת האפקט. גם כאן נסתייג ונציין כי בהתאם לכתוב תחת "תיאור התוצאות", נראה **שלסינון צעדים המבוסס על זמן תגובה חריג אין השפעה משמעותית על התוצאות.**

בסופו של דבר ביצענו את ניתוח התוצאות במקביל על **שני סטים של נתונים** - האחד ללא סינון חריגים כלל, והאחר עם סינון אגרסיבי. להלן פירוט הסינונים שבוצעו בסט הנתונים השני:

* לא אותרו ולא סוננו נבדקים חריגים.
* סוננו כלל התוכניות שאחוז ההצלחה בהן היה נמוך מ-90% (בסה"כ 146 תוכניות).
* לאחר סינון התכניות, סוננו צעדים קיצוניים בהיבט של זמן תגובה איטי, בציון חריגה של 2.25 (IQR) ומעלה (בסה"כ סוננו 75 צעדים איטיים).

בסה"כ בסט ללא הסינון נמצאו 3865 צעדים, ובסט לאחר הסינון האגרסיבי נותרו 2767 צעדים.

בקובץ ניתוח התוצאות הסופי (FinalAnalysis) ניתן לראות פירוט על התוכניות והצעדים שסוננו מכל נבדק.

**4. תיאור התוצאות**

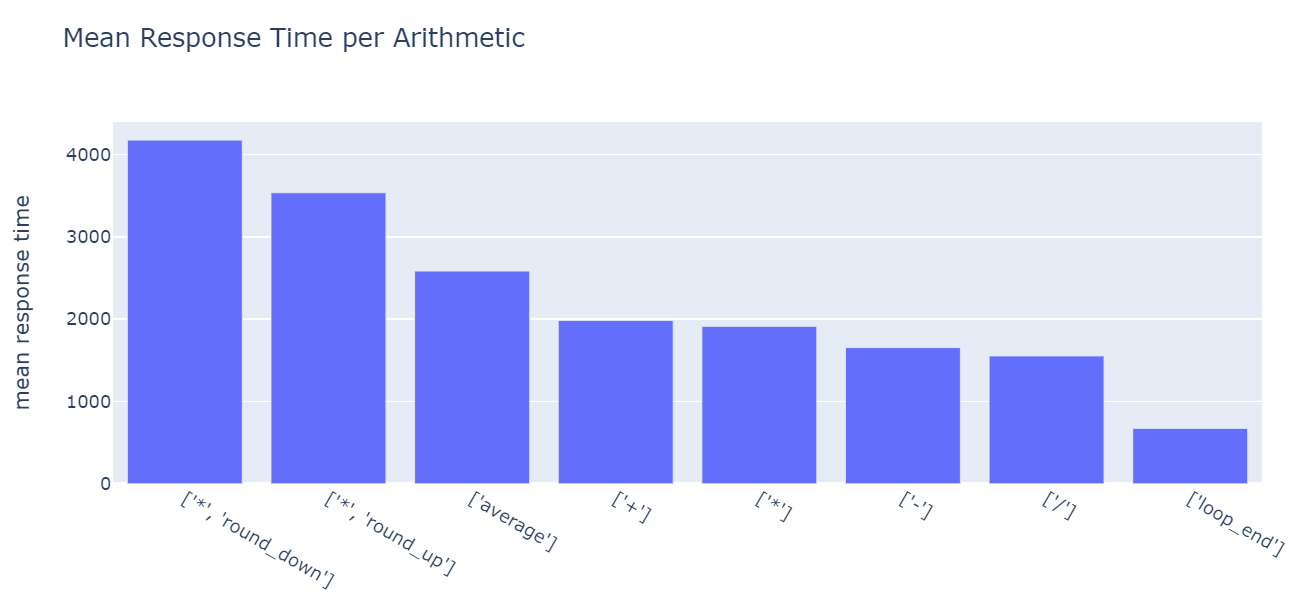
מלבד האפקט המרכזי, ביצענו כמה אנליזות כלליות יותר, וכן בחנו את השפעתם של גורמים שונים על האפקט.

**4.1. ניתוח כללי**

הניתוח הכללי התבצע על כלל הנתונים - כלומר לא רק על השורות הראשונות, וללא סינונים בכלל, מלבד להוצאת הלולאה הראשונה בכל תכנית.

4.1.1 פעולות אריתמטיות - ביצועים בהיבט זמן תגובה ממוצע ואחוז הצלחה.

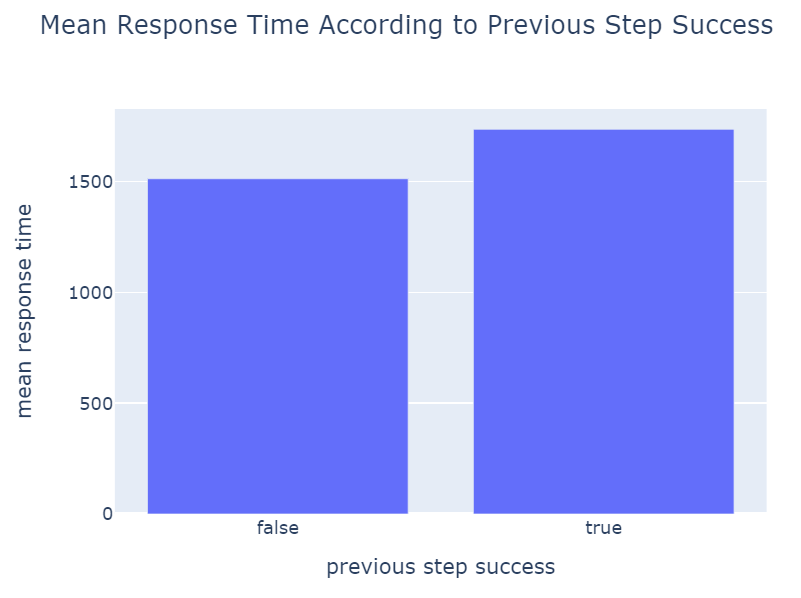
* כצפוי נרשמה שונות משמעותית בין הפעולות האריתמטיות השונות, כאשר פונקציות ערך שלם עליון / תחתון דרשו את הזמן הממושך ביותר. להלן גרף המפרט את משך הזמן הממוצע הנדרש לכל פעולה, הכולל עמודת 'loop\_end' המתייחסת לזמן הנדרש לתגובה בסיום לולאה.



* מבחינת אחוזי ההצלחה - מלבד לאותן שתי פעולות "קשות" (ערך שלם עליון / תחתון) בהן נרשמו אחוזי הצלחה של כ-86%, ביתר הפעולות נרשמו אחוזי הצלחה של כ-96%.

4.1.2 השפעת טעות על הביצועים - טעות בצעד הקודם, השפעה על זמן התגובה.

* זמן התגובה בצעדים לאחר טעות (סה"כ 672), כלומר שבצעד הקודם להן בוצעה טעות, היה **מהיר יותר ביחס ליתר הצעדים** - ובאופן מובהק (0.000219 p =).

****

4.1.2 אפקט למידה בין הסשנים - האם קיים שיפור בביצועים בהיבט של זמני תגובה ואחוז טעויות.

* נמצא כי זמן התגובה הממוצע בסשן הראשון גדול מזמן התגובה הממוצע בסשן השני. כלומר בממוצע על פני כל הנבדקים, **נראה שיפור בביצועים** (p < 0.000001).
* נרשמת ירידה גם בזמן התגובה באחוזון 90 בין הסשן הראשון לשני.

4.2**. האפקט והשפעות עליו**

כפי שצוין קודם לכן, ניתוח האפקט התבצע במקביל על שני סטים של נתונים - אחד ללא סינון (להלן all\_data) והאחר עם סינון אגרסיבי (להלן best\_data).

4.2.1. התוצאה המרכזית - לא הצלחנו לאתר אפקט Priming כלל בכל אחד מהסטים של הנתונים.

**,**

**5. מסקנות ודיון**

(כאן נכתוב מסקנות כלליות מהנתונים שהתקבלו, והצעות להמשך המחקר)