# שאלות מחקר על מדידת הזמנים: (בחינה של ממוצעים)

**הערה כללית: לאורך כל העבודה רמת המובהקות (α) הייתה 0.95**

# טבלת הסבר השמות

לאורך העבודה שמות המשתנים רשומים בקיצור ובאנגלית. בניסוי השתתפו 50 משתתפים, אך ב-3 מדדים נתקבלו תשובות רק מ-25 משתתפים – מספר התשובות ניתן בטבלה כ-N. להלן טבלת השמות המלאים:

|  |  |  |  |
| --- | --- | --- | --- |
| משתנה | יחידות מידה | שם מקוצר | n |
| מין | בינארי: 1/0 | gender | 50 |
| גיל הנבדק | מספרי: 55-73 | age | 50 |
| שפה | מקודד למספרים: 1-11 | lang | 50 |
| זמן שימוש בסלולרי ביום | שעות | hours.of.use | 50 |
| שימוש באפליקציות שירותי בריאות | בינארי: 1/0 | application | 50 |
| קלות שימוש הנתפסת | דירוג: 1-5 | easy | **25** |
| חסר! | חסר! | times.of.use | **25** |
| זמן השימוש בסלולרי ביום | דקות | minutes | **25** |
| שביעות הרצון מגודל הפונט באפליקציה הקיימת | דירוג: 1-8 | font1 | 50 |
| שביעות הרצון מגודל הפונט באפליקציה החדשה | דירוג: 1-8 | font2 | 50 |
| שביעות הרצון מכמות התפריטים באפליקצייה הקיימת | דירוג: 1-10 | n.menu1 | 50 |
| שביעות הרצון מכמות התפריטים באפליקצייה החדשה | דירוג: 1-10 | n.menu2 | 50 |
| חסר! | שניות | Login..Secs. | 50 |
| חסר! | שניות | Login..NEW..Secs. | 50 |
| זמן זימון תור לרופא משפחה באפליקציה הקיימת | שניות | Appointments1..Secs. | 50 |
| זמן זימון תור לרופא משפחה באפליקציה החדשה | שניות | Appointments1..NEW..Secs. | 50 |
| זמן זימון תור לרופא יועץ באפליקצייה הקיימת | שניות | Appointments2..Secs. | 50 |
| זמן זימון תור לרופא יועץ באפליקצייה החדשה | שניות | Appointments2..NEW..Secs. | 50 |
| זמן צפייה בתוצאות מעבדה באפליקציה הקיימת | שניות | Blood.Test..Secs. | 50 |
| זמן צפייה בתוצאות מעבדה באפליקציה החדשה | שניות | Blood.Test...NEW..Secs. | 50 |
| זמן צפייה בסיכום מידע רפואי באפליקצייה הקיימת | שניות | Medical.File..Secs. | 50 |
| זמן צפייה בסיכום מידע רפואי באפליקצייה החדשה | שניות | Medical.File...NEW..Secs. | 50 |
| זמן בקשת מרשם באפליקציה הקיימת | שניות | Mirshm..Secs. | 50 |
| זמן בקשת מרשם באפליקציה החדשה | שניות | Mirshm..NEW..Secs. | 50 |
| זמן הפקת דוח חיובים חודשי באפליקצייה הקיימת | שניות | Payments..Secs. | 50 |
| זמן הפקת דוח חיובים חודשי באפליקצייה החדשה | שניות | Payments..NEW..Secs. | 50 |
| זמן עדכון תור לרופא באפליקצייה הקיימת | שניות | Update.A..Secs. | 50 |
| זמן עדכון תור לרופא באפליקצייה החדשה | שניות | Update.A..NEW..Secs. | 50 |
| זמן ביצוע כל הפעולות באפליקציה הישנה | שניות | SUM.OLD..Secs. | 50 |
| זמן ביצוע כל הפעולות באפליקציה החדשה | שניות | SUM.NEW..Secs. | 50 |

טבלה 1 רשימת המדדים שנלקחו בניסוי, יחידות המידה, שם המדד כפי שמופיע בקוד ("שם מקוצר") ומספר המשתתפים לכל מדד (n).

## שיטת העבודה:

### בדיקת התפלגות נורמלית:

כדי לבחון את ההבדל בין הממוצעים, נדרש קודם כל לבדוק האם ההתפלגות של כל משתנה היא נורמלית. לשם כך השתמשנו במבחן שפירו-ווילקס לבדיקת התפלגות נורמלית:

H0: ערכי המדגם מתפלגים באופן נורמאלי

H1: ערכי המדגם מתפלגים באופן שאינו נורמאלי

ביצענו את המבחן הן על הנתונים הגולמיים והן על נתונים לאחר חישוב LOG

בנוסף, מצורפת ההתפלגות של הנתונים בצורה גרפית באמצעות גרף שארים (Residuals) והיסטוגרמת התפלגות.

**תוצאה:** כפי שניתן לראות בטבלה, נמצא שלכל המשתנים שנמדדו מבחן שפירו-ווילקס מובהק ולכן עלינו לדחות את השערת ה-0 ולהסיק כי ערכי המדגם מתפלגים באופן שאינו נורמאלי.

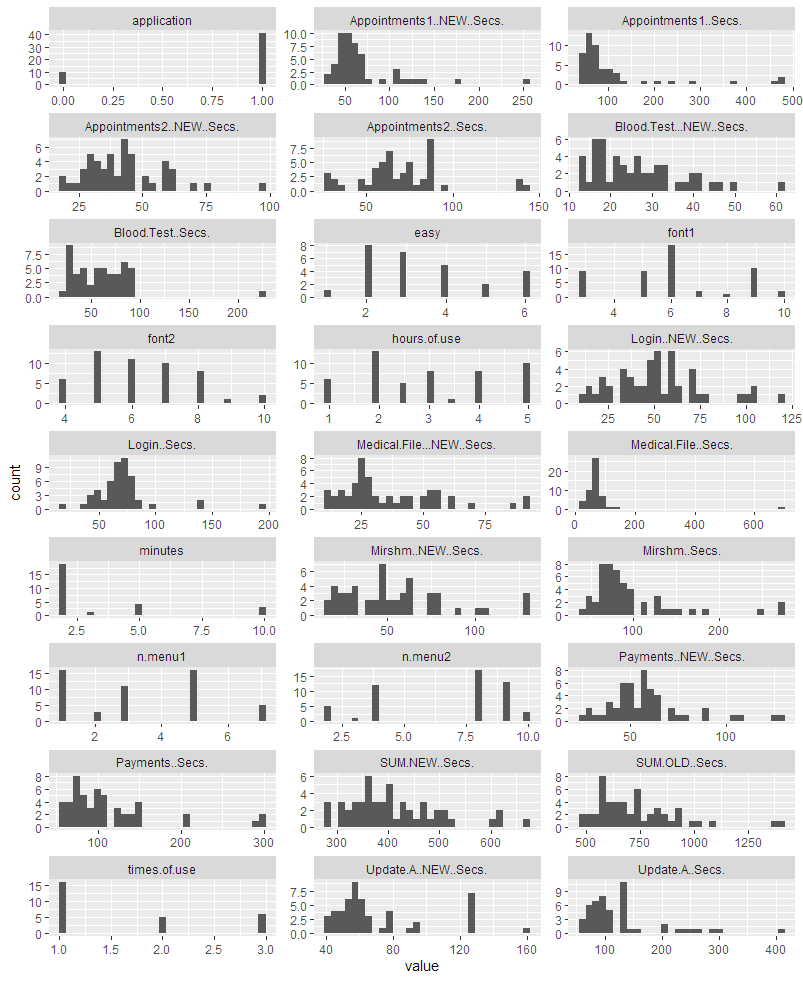
הערה: המשתנים היחידים שהתפלגו נורמלית לאחר חישוב log היו כולם מדדי זמן לאפליקצייה החדשה ולא נדרש לבחון את ההבדל ביניהם, לכן מבחינת המשך הניתוח כל המשתנים מתפלגים באופן שאינו נורמלי

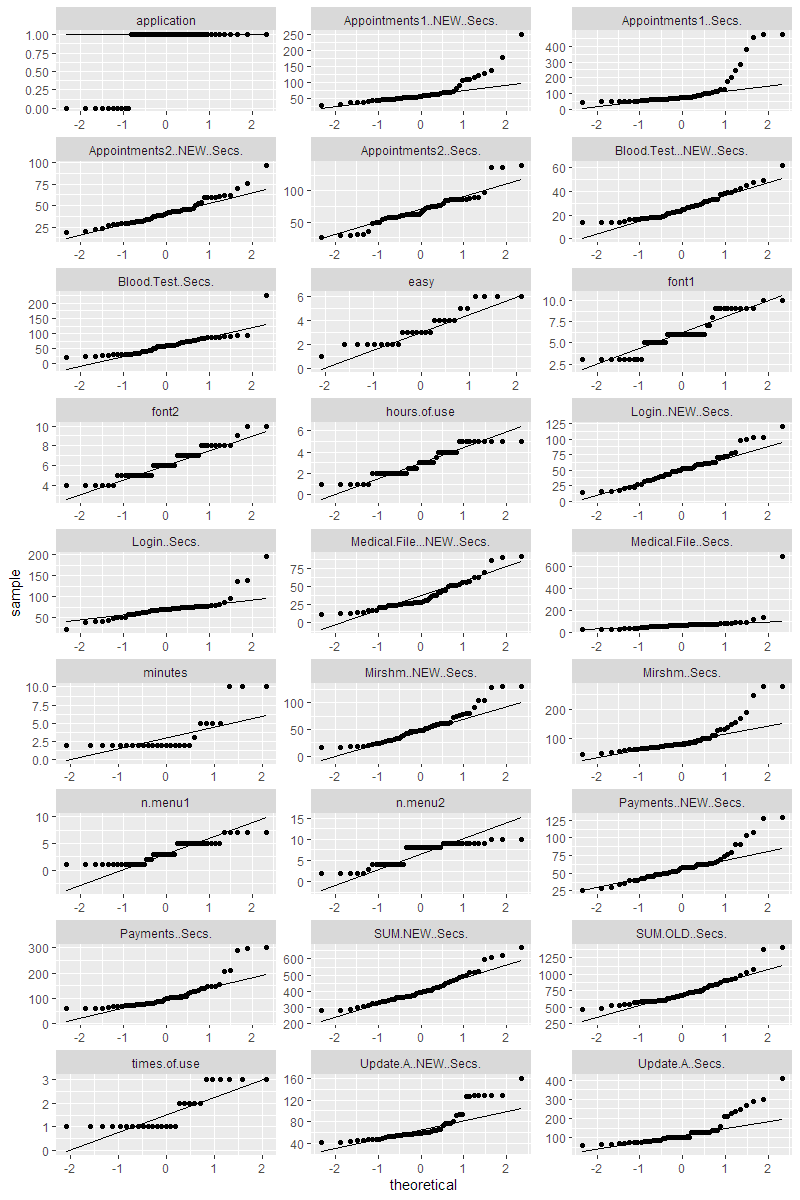
טבלה 2 בחינת התפלגות נורמלית לפי מבחן שפירו ווילקס לכל משתנה. המבחן בוצע הן לנתונים הגולמיים והן לאחר חישוב log על הנתונים. עמודת "Normal" מצביעה מתי ערך p.value נמוך מ-0.05 ולפיכך ההתפלגות אינה נורמלית.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | p.value | Normal | LOG: p.value | LOG: Normal |
| hours.of.use | 0.000539 | TRUE | 0.000211 | TRUE |
| easy | 0.009101 | TRUE | 0.031679 | TRUE |
| times.of.use | 3.56E-06 | TRUE | 3.69E-06 | TRUE |
| minutes | 1.07E-07 | TRUE | 3.66E-07 | TRUE |
| font1 | 0.000229 | TRUE | 5.36E-05 | TRUE |
| font2 | 0.005085 | TRUE | 0.012085 | TRUE |
| n.menu1 | 2.14E-05 | TRUE | 2.30E-06 | TRUE |
| n.menu2 | 1.26E-06 | TRUE | 1.35E-07 | TRUE |
| Login..Secs. | 4.08E-08 | TRUE | 8.45E-05 | TRUE |
| Login..NEW..Secs. | **0.036771** | **TRUE** | **0.060256** | **FALSE** |
| Appointments1..Secs. | 1.41E-10 | TRUE | 2.44E-06 | TRUE |
| Appointments1..NEW..Secs. | 1.73E-08 | TRUE | 0.002989 | TRUE |
| Appointments2..Secs. | 0.00109 | TRUE | 0.001907 | TRUE |
| Appointments2..NEW..Secs. | **0.003298** | **TRUE** | **0.915829** | **FALSE** |
| Blood.Test..Secs. | 1.75E-07 | TRUE | 0.014222 | TRUE |
| Blood.Test...NEW..Secs. | **0.000886** | **TRUE** | **0.145276** | **FALSE** |
| Medical.File..Secs. | 3.80E-14 | TRUE | 9.29E-07 | TRUE |
| Medical.File...NEW..Secs. | **0.000185** | **TRUE** | **0.349945** | **FALSE** |
| Mirshm..Secs. | 4.34E-08 | TRUE | 0.001519 | TRUE |
| Mirshm..NEW..Secs. | **0.00076** | **TRUE** | **0.203059** | **FALSE** |
| Payments..Secs. | 5.50E-08 | TRUE | 0.000878 | TRUE |
| Payments..NEW..Secs. | 7.07E-05 | TRUE | 0.242286 | **FALSE** |
| Update.A..Secs. | 1.47E-07 | TRUE | 0.001367 | TRUE |
| Update.A..NEW..Secs. | 2.94E-07 | TRUE | 7.48E-05 | TRUE |
| SUM.OLD..Secs. | 4.07E-05 | TRUE | 0.033301 | TRUE |
| SUM.NEW..Secs. | 0.005271 | TRUE | 0.323858 | **FALSE** |

בנוסף, התפלגות הנתונים מוצגת באופן גרפי בשני אופנים:   
היסטוגרמת התפלגות וגרף שארים (Residuals).

איור 1 התפלגות המשתנים השונים בהיסטוגרמה, ציר ה-X מציג את הערך הנמדד וציר ה-Y את כמות הדגימות בכל ערך. כל גרף מציג התפלגות של משתנה אחד.



איור 2 התפלגות המשתנים השונים לעומת התפלגות נורמלית בגרף שארים. בציר ה-X ניתנים נתוני שברונים (Quantiles) מהתפלגות נורמלית תאורטית ובציר ה-Y שברונים של המתשנה שנמדד. הקו המשורטט מראה רגרסייה ליניארית של הנתונים.

קשה להסיק מן הגרפים לגבי ההתפלגות של הנתונים ולכן נשתמש בתוצאות ממבחן ווילקס-שפירו להמשך הניתוח.

#### קוד:

השתמשנו בפקודת **shapiro.test** מתוך חבילת הבסיס של R

# NORMALITY: Shapiro test  
# From the output, the p-value > 0.05 implies that the distribution  
#of the data are not significantly different from a normal distribution.  
#In other words, we can assume the normality.  
  
Normality.Table = as.data.frame(  
 do.call(rbind, lapply(data[,numerical], function(x) shapiro.test(x)[c("p.value")] # Calculates Shapiro test for each variable, and attaches the pvalue to a new table  
 )))   
Normality.Table$Normal = Normality.Table$p.value < 0.05 # This will show you which rows have pvalue<0.05  
  
#And now the same for logs:  
logstable = log10(data[,numerical]) # Calculates the log of the data  
Logs.Normality.Table = as.data.frame(  
 do.call(rbind, lapply(logstable, function(x) shapiro.test(x)[c("p.value")]))) # Calculates Shapiro test for each variable, and attaches the pvalue to a new table  
Logs.Normality.Table$Normal = Logs.Normality.Table$p.value < 0.05 # This will show you which rows have pvalue<0.05  
  
Normality.Table.Combined = cbind(Normality.Table, Logs.Normality.Table) # Combines the regular and Log shapiro test results  
colnames(Normality.Table.Combined)[3:4] = c("LOG: p.value", "LOG: Normal") # renames the columns

### בדיקת הבדל בין ממוצעים:

היות והמשתנים אינם מתפלגים באופן נורמלי, נוכל לבחון את ההבדל בין הממוצעים רק באמצעות מבחן א-פרמטרי (שאינו מניח התפלגות נורמלית ושונויות דומות). למטרה זו נשתמש במבחן Wilcoxon rank sum test.

הערה: המשתנים היחידים שהתפלגו נורמלית לאחר חישוב log היו כולם מדדי זמן לאפליקצייה החדשה ולא נדרש לבחון את ההבדל ביניהם, לכן מבחינת המשך הניתוח כל המשתנים מתפלגים באופן שאינו נורמלי (טבלה 2)

במבחן ווילקוקסון חשוב לשים לב כי ערכי המדגם מדורגים ולכן הוא אינו בהכרח מצביע על גודל ההפרש אלא על עוצמתו.

H0: P(X>Y) = P(Y>X) = 0.5

H1: P(X>Y) ≠ P(Y>X)

או במילים:

H0: ההסתברות שמקרה אקראי אחד ממדגם X גדול ממקרה אקראי אחד ממדגם Y היא 0.5

H1: ההסתברות שמקרה אקראי אחד ממדגם X גדול ממקרה אקראי אחד ממדגם Y שונה מ-0.5

#### קוד:

# Significance testing:   
# We will use the wilcoxon Ranked Sum Test as a non-parametric test:  
colnames(data)

## [1] "ID" "gender"   
## [3] "age" "lang"   
## [5] "times.in.clalit" "hours.of.use"   
## [7] "application" "use.it"   
## [9] "easy" "times.of.use"   
## [11] "minutes" "font1"   
## [13] "font2" "n.menu1"   
## [15] "n.menu2" "Login..Secs."   
## [17] "Login..NEW..Secs." "Appointments1..Secs."   
## [19] "Appointments1..NEW..Secs." "Appointments2..Secs."   
## [21] "Appointments2..NEW..Secs." "Blood.Test..Secs."   
## [23] "Blood.Test...NEW..Secs." "Medical.File..Secs."   
## [25] "Medical.File...NEW..Secs." "Mirshm..Secs."   
## [27] "Mirshm..NEW..Secs." "Payments..Secs."   
## [29] "Payments..NEW..Secs." "Update.A..Secs."   
## [31] "Update.A..NEW..Secs." "SUM.OLD..Secs."   
## [33] "SUM.NEW..Secs."

#### Function explained ####  
# This function will work as follows:   
# Input: two variable names from your data set  
# Output: Boxplot with Wilcoxon Rank Sum Test statistic printed on it  
  
  
# Notes:   
# 1. You can ignore the warnings:   
 # No id variables; using all as measure variables  
 # Removed 2 rows containing non-finite values  
# It's just the way it handles NA lines. Which you have  
  
# 2. Box plots:   
# The middle line is the median  
#   
# The lower and upper hinges correspond to the first and third quartiles  
#   
# The whiskers extend from the hinge to the furthest value but no further than 1.5 \* IQR from the hinge   
# (where IQR is the inter-quartile range, or distance between the first and third quartiles).   
  
# Outliers: Points beyond the whisker limit  
  
 wilcox.plus.boxes <- function(var1, var2) { # Defines the function requiring two parameters  
data.for.function = data[,c(var1,var2)] # creates a dataframe from your data which is only the two parameters you chose  
colnames(data.for.function) = c(var1,var2) # Gives the relevant columns names (for later)  
  
varname = gsub("\\..\*","",var1)  
levs = c("New","Old")  
  
result = wilcox.test( # Calculates the wilcoxon test between the two variables  
 data.for.function[,1],  
 data.for.function[,2]  
 )   
  
mean.diff = mean(data.for.function[,1], na.rm = T)-mean(data.for.function[,2],na.rm = T)  
  
melted = melt(data.for.function) # Transforms the data for ggplot  
  
plot = ggplot( # This is a very long function to create the plot  
 melted, aes(x=variable,y=value,fill=variable)) + # Defines the releant data  
 geom\_boxplot()+ # Defines the plot type as Box Plot  
 theme(panel.background = element\_rect(fill="white")) + # Makes it visually clean  
 labs(x = varname,y="Time [Mins]") + # Adds X, Y axis labels  
 theme(panel.grid.major = element\_line(colour = "gray")) + # adds gray lines on the Y axis  
 theme(panel.grid.minor = element\_line(colour = "gray")) + # does the same  
 guides(fill=FALSE) + # Removes the legend (not needed for this)  
 labs( # Startes to create the complicated titles  
 title = paste("Comparison of", varname, sep = " "), # Main title: Comparison between the two variables  
   
   
   
   
 subtitle = paste( # Subtitle: gives the different statistics pasted together  
 "Result of the Wilcoxon Rank Sum Test: \ntest statistic = ",   
 result$statistic, "\n p-value = ",   
 formatC(result$p.value, format = "g", digits = 2),   
 "\n mean difference = ", round(mean.diff, digits = 1), "minutes", sep = " " )  
 ) +   
 scale\_x\_discrete(breaks = c(var1,var2), labels=c("New","Old")) +  
   
 geom\_signif(comparisons = list(c(var1, var2)), # Creates a significance star when the pvalue<0.05  
 test="wilcox.test", map\_signif\_level = c("\*"=0.05))  
   
  
plot # Saves the final plot   
  
 }  
  
names = colnames(data) # This is for my own reference to know which variables need to be tabulated  
  
# Calculations:  
# Each Q is a question from your list  
Q1 = wilcox.plus.boxes(names[33],names[32]) #SUM  
Q2 = wilcox.plus.boxes(names[19],names[18]) # Appointments1  
Q3 = wilcox.plus.boxes(names[27],names[26]) # Mirshm  
Q4 = wilcox.plus.boxes(names[29],names[28]) # Payments  
Q5 = wilcox.plus.boxes(names[31],names[30]) # Update  
Q6 = wilcox.plus.boxes(names[21],names[20]) # Appointments 2  
Q7 = wilcox.plus.boxes(names[25],names[24]) # Medical File  
Q8 = wilcox.plus.boxes(names[23],names[22]) # Blood Test

## שאלות המחקר

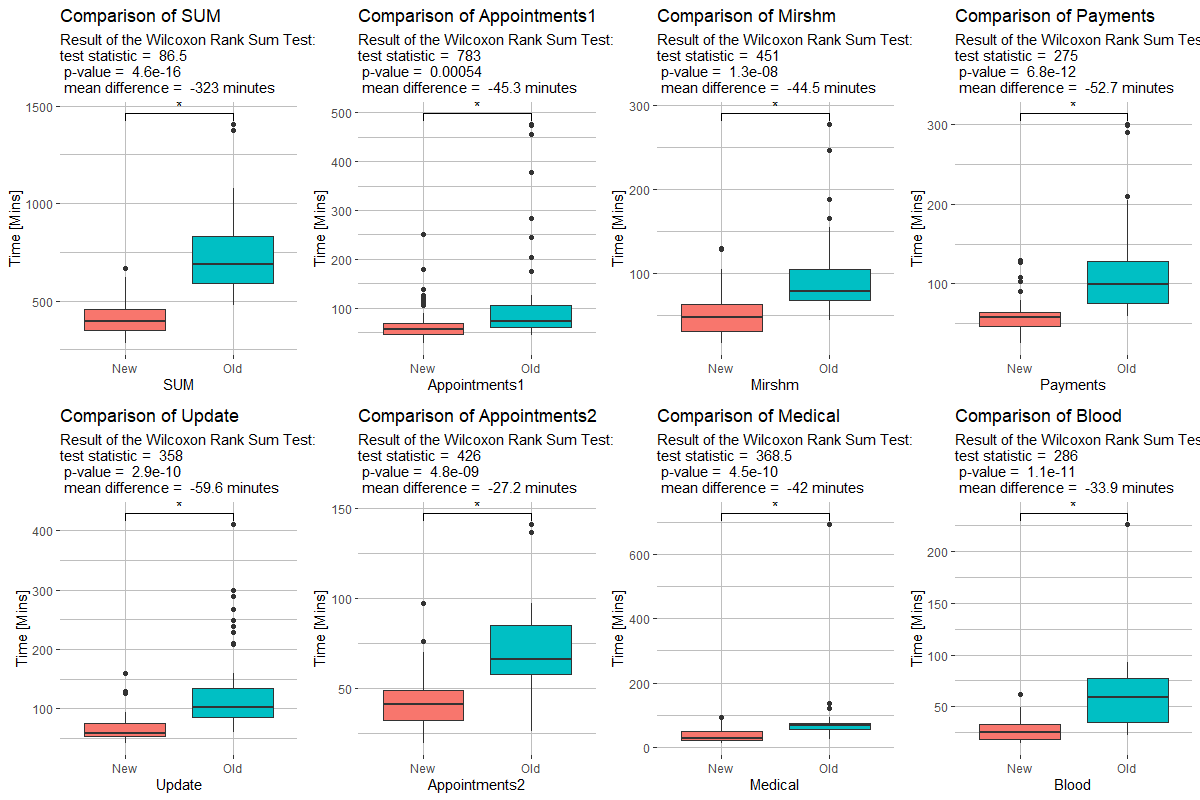
1. האם קיים הבדל בין זמן ביצוע כל המשימות בקיימת אל מול זמן ביצוע כל הפעולות בחדשה? (SUM)
2. האם קיים הבדל בין זמן זימון תור לרופא משפחה בחדשה אל מול זמן זימון תור לרופא משפחה בקיימת? (Appointment1)
3. האם קיים הבדל בין זמן בקשת מרשם בחדשה אל מול זמן בקשת מרשם בקיימת? (Mirshm)
4. האם קיים הבדל בין זמן הפקת דוח חיובים חודשי בקיימת לבין זמן הפקת דוח חיובים חודשי בחדשה? (Payments)
5. האם קיים הבדל בין זמן עדכון תור לרופא בקיימת לבין זמן עדכון תור לרופא בחדשה? (Update)
6. האם קיים הבדל בין זמן זימון תור לרופא יועץ בקיימת אל מול זמן זימון תור לרופא יועץ בחדשה? (Appointments2)
7. האם קיים הבדל בין זמן צפייה בסיכום מידע רפואי בקיימת אל מול זמן צפייה בסיכום מידע רפואי בחדשה? (Medical)
8. האם קיים הבדל בין זמן צפייה בתוצאות מעבדה בקיימת אל מול זמן צפייה בתוצאות מעבדה בחדשה? (Blood)

## תשובות

כפי שניתן לראות באיור 3, בכל המדדים שנבדקו האפליקצייה החדשה מקצרת את זמני הפעולות באופן מובהק לפי מבחן ווילקוקסון. כלומר, עלינו לדחות את השערת ה-0 ולהסיק כי, בהנתן מקרה אקראי, משך זמן ביצוע הפעולה יהיה קצר יותר באפליקציה החדשה, לכל הפעולות שנמדדו.

כאשר סוכמים את כל המדדים (SUM), זמני הפעולה באפליקציה החדשה באופן ממוצע היו קצרים ב-323 דקות, ובכל מדד באופן ספציפי הזמן היה קצר בין 27-59 דקות.

איור 3 השוואה בין ממוצע זמן ביצוע פעולות שונות באפליקציה החדשה לעומת הישנה. האפליקציה הישנה (Old) מופיעה בכחול והחדשה (new) באדום. הפעולה שנמדדה רשומה בתחתית כל גרף. ציר ה-Y מציג את זמן הפעולה. הנתונים מוצגים כקופסאות, כאשר הקו המרכזי מייצג את החציון, גבולות הקופסה מייצגים את רבעונים 25%, 75% והקווים החיצוניים (Whiskers) מייצגים את הערכים הנמצאים עד מרחק 1.5 של הטווח-הבין-רבעוני (Inter-Quantile-Range). נקודות חריגות מטווח זה מוצגות כנקודות. בכל המבחנים גודל המדגם הוא n=50. בנוסף, ההבדל בין הקבוצות נבחן באמצעות מבחן ווילקוקסון. ערך p value רשום בראש כל גרך וכאשר הערך קטן מ-0.05 מסומנת כוכבית כדי לציין דחיית השערת H0 והבדל מובהק בין הקבוצות. ההפרש בין הממוצעים ניתן גם הוא בראש כל גרף (mean difference) ומחושב ע"י החסרת הזמן הממוצע באפליקצייה הישנה מהממוצע בחדשה.



טבלה 3 סיכום ההשערות המבחנים והתוצאות לבדיקת ההבדל בין הממוצעים. told­ ו-tnew בזמן ביצוע הפעולה באפליקציה הישנה או החדשה על ידי משתתף

|  |  |  |  |
| --- | --- | --- | --- |
| מדד | השערות | מבחן | תוצאה |
| Sum | H0: P(told > tnew) = P(tnew > told) = 0.5  H1: P(told >Y) ≠ P(tnew > told) | Wilcoxon | P<0.05  ניתן לדחות את השערת H0 |
| Appointments1 | H0: P(told > tnew) = P(tnew > told) = 0.5  H1: P(told >Y) ≠ P(tnew > told) | Wilcoxon | P<0.05  ניתן לדחות את השערת H0 |
| Mirshm | H0: P(told > tnew) = P(tnew > told) = 0.5  H1: P(told >Y) ≠ P(tnew > told) | Wilcoxon | P<0.05  ניתן לדחות את השערת H0 |
| Payments | H0: P(told > tnew) = P(tnew > told) = 0.5  H1: P(told >Y) ≠ P(tnew > told) | Wilcoxon | P<0.05  ניתן לדחות את השערת H0 |
| Update | H0: P(told > tnew) = P(tnew > told) = 0.5  H1: P(told >Y) ≠ P(tnew > told) | Wilcoxon | P<0.05  ניתן לדחות את השערת H0 |
| Appointments2 | H0: P(told > tnew) = P(tnew > told) = 0.5  H1: P(told >Y) ≠ P(tnew > told) | Wilcoxon | P<0.05  ניתן לדחות את השערת H0 |
| Medical | H0: P(told > tnew) = P(tnew > told) = 0.5  H1: P(told >Y) ≠ P(tnew > told) | Wilcoxon | P<0.05  ניתן לדחות את השערת H0 |
| Blood | H0: P(told > tnew) = P(tnew > told) = 0.5  H1: P(told >Y) ≠ P(tnew > told) | Wilcoxon | P<0.05  ניתן לדחות את השערת H0 |

# שאלות מחקר על מתאם (קורלציה) מתוך השאלון:

## שיטת העבודה:

את שאלות המתאם חשוב לחלק לשני סוגים:

1. מתאם בין שני משתנים מספריים
2. מתאם בין משתנה מספרי למשתנה בינארי (כמו מין)

### שיטת העבודה למתאם בין משתנים מספריים:

1. לשרטט את הקשר בין המשתנים
2. לבחון מה כיוון הקשר (מקדם פירסון/ספירמן: r)
3. לבחון מה עוצמת הקשר (מקדם הקורלציה R2)
4. לבחון מה רמת המובהקות (p-value), או האם הקשר מובהק (p-value<0.05)

השימוש במקדם פירסון או ספירמן תלוי בסוג המשתנה. למשתני רצף (Continuous) השתמשנו במקדם פירסון ולמשתני סדר (Ordinal) במקדם ספירמן.

בחרנו בשיטה זו היות ו-r מצביע הן על עוצמת והן על כיוון הקשר בקשרים של שני משתנים.

במבחן המובהקות אנו בוחנים האם הקורלציה שונה באופן מובהק מ=0:

H0: r = 0

H1: r ≠ 0

### קוד:

#### Correlations ####  
  
# This we must break down into 3 "categories":  
# 1. Correlation between numericals   
# 2. Correlation between a binary category (like gender) and a numerical  
# 3. Correlation between a categorical and a numerical  
  
####1. Correlation between numericals ####  
  
#### Function explained ####  
# This function will work as follows:   
# Input: two variable names from your data set  
# Output: scatterplot with correlation coefficents and statistic printed on it  
#  
# This will calculate the correlation between two chosen variables and plot it  
  
  
correlation.with.plot = function(var1, var2, cor.met){ # Defines the function requiring two parameters. cor.met needs to be either "pearson" or "spearman"  
 data.for.function = data[,c(var1,var2)] # creates a dataframe from your data which is only the two parameters you chose  
 colnames(data.for.function) = c(var1,var2) # Gives the relevant columns names (for later)  
  
   
cor = cor.test(data[,var1], data[,var2], method = cor.met) # tests the correlation between the two variables  
 # This calculates all the values we need: r and p-val  
  
plot = ggplot( # This is a very long function to create the plot  
 data.for.function, aes(x=data.for.function[,1],y=data.for.function[,2])) + # Defines the relevant data  
 geom\_point() + # Defines that we will create a scatter plot  
 theme(panel.background = element\_rect(fill="white")) + # Makes it visually clean  
 labs(x = var1,y=var2) + # Adds X, Y axis labels  
 theme(panel.grid.major = element\_line(colour = "gray")) + # adds gray lines on the Y axis  
 theme(panel.grid.minor = element\_line(colour = "gray")) + # does the same  
 guides(fill=FALSE) + # Removes the legend (not needed for this)  
 labs( # Starts to create the complicated titles  
 title = paste("Comparison between",var1, "&", var2, sep = " "), # Main title: Comparison between the two variables  
 subtitle = paste( # Subtitle: gives the different statistics pasted together  
 "Correlation statistics using", cor.met, "method:  
 correlation coefficent (r) = ", round(cor$estimate, digits = 2),   
 "\n R^2 = ", round(cor$estimate^2, digits = 2),   
 "\n p-value = ",   
 formatC(cor$p.value, format = "g", digits = 2))) +  
 geom\_smooth(method="lm", se = F) # Adds a linear correlation line to the figure  
   
}  
  
# Now we can ask the questions  
names # This is again just for my reference

## [1] "ID" "gender"   
## [3] "age" "lang"   
## [5] "times.in.clalit" "hours.of.use"   
## [7] "application" "use.it"   
## [9] "easy" "times.of.use"   
## [11] "minutes" "font1"   
## [13] "font2" "n.menu1"   
## [15] "n.menu2" "Login..Secs."   
## [17] "Login..NEW..Secs." "Appointments1..Secs."   
## [19] "Appointments1..NEW..Secs." "Appointments2..Secs."   
## [21] "Appointments2..NEW..Secs." "Blood.Test..Secs."   
## [23] "Blood.Test...NEW..Secs." "Medical.File..Secs."   
## [25] "Medical.File...NEW..Secs." "Mirshm..Secs."   
## [27] "Mirshm..NEW..Secs." "Payments..Secs."   
## [29] "Payments..NEW..Secs." "Update.A..Secs."   
## [31] "Update.A..NEW..Secs." "SUM.OLD..Secs."   
## [33] "SUM.NEW..Secs."

Q10 = correlation.with.plot(names[6], names[32], cor.met = "pearson") # hours of use VS. SUM.OLD  
Q11 = correlation.with.plot(names[6], names[33], cor.met = "pearson") # hours of use VS. SUM.NEW  
Q12 = correlation.with.plot(names[9], names[32], cor.met = "spearman") # easy VS. SUM.OLD  
Q13 = correlation.with.plot(names[9], names[33], cor.met = "spearman") # easy VS. SUM.NEW  
Q14 = correlation.with.plot(names[15], names[32], cor.met = "spearman") # n.menu2 VS. SUM.OLD  
Q15 = correlation.with.plot(names[14], names[33], cor.met = "spearman") # n.menu1 VS. SUM.NEW  
Q16 = correlation.with.plot(names[3], names[11], cor.met = "pearson") # age VS. minutes

### שיטת העבודה למתאמים בין משתנה מספרי לבינארי

לשם בדיקת המתאם בין משתנים מספריים לבינאריים רצינו למעשה לבחון האם קיים הבדל בין הממוצעים. כפי שהראנו בתחילת הדו"ח, המשתנים אינם מתפלגים נורמלית ולכן השתמשנו שוב במבחן ווילקוקסון ובאותו הקוד, עם שינויים קלים.

### קוד:

# 2. Correlation between a binary category (like gender) and a numerical  
# For questions with regards to 2 categories, like gender, it makes no sense to calculate a   
# CORRELATION (as there is no 1.3 male)  
# Therefor we need to test whether there is a difference in averages like we did earlier  
  
# As we already showed no variable behaves "noramlly", we must still use the non-parametric test: Wilcoxon Ranked Sum Test  
# This function is a variation of the earlier 2 functions for this purpose.  
# Input: Two variables, the first one is BINARY and the second NUMERICAL.  
# Output: a boxplot with test statistics printed on it  
  
wilcox.categorical.data <- function(var1, var2) { # Defines the function requiring two parameters  
 data.for.function = data[,c(var1,var2)] # creates a dataframe from your data which is only the two parameters you chose  
 data.for.function[,1] = as.factor(data.for.function[,1]) # Defines the BINARY column as CATEGORICAL  
 colnames(data.for.function) = c(var1,var2) # Gives the relevant columns names (for later)  
   
 result = wilcox.test( # Calculates the wilcoxon test between the categories based on the numerical values  
 data.for.function[data.for.function==1,2],  
 data.for.function[data.for.function==0,2]  
 )   
 melted = data.for.function[complete.cases(data.for.function),] # Removes any missing data from the dataset  
 plot = ggplot( # This is a very long function to create the plot  
 melted, aes(x=melted[,1],y=melted[,2],fill=melted[,1])) + # Defines the relevant data  
 geom\_boxplot()+ # Defines the plot type as Box Plot  
 theme(panel.background = element\_rect(fill="white")) + # Makes it visually clean  
 labs(x = colnames(melted)[1],y="Time [Mins]") + # Adds X, Y axis labels  
 theme(panel.grid.major = element\_line(colour = "gray")) + # adds gray lines on the Y axis  
 theme(panel.grid.minor = element\_line(colour = "gray")) + # does the same  
 guides(fill=FALSE) + # Removes the legend (not needed for this)  
 labs( # Startes to create the complicated titles  
 title = paste("Comparison between", var1, "&", var2, sep = " "), # Main title: Comparison between the two variables  
 subtitle = paste( # Subtitle: gives the different statistics pasted together  
 "Result of the Wilcoxon Rank Sum Test: \ntest statistic = ",   
 result$statistic, "\n p-value = ",   
 formatC(result$p.value, format = "g", digits = 2))) +   
 geom\_signif(comparisons = list(c(var1, var2)), # Creates a significance star when the pvalue<0.05  
 test="wilcox.test", map\_signif\_level = c("\*"=0.05))  
   
 plot # Saves the final plot   
}  
  
  
# Now we can ask the questions:  
#2-way catergorical:  
Q9 = wilcox.categorical.data(names[7], names[18]) # application VS. Appointments 1  
Q17 = wilcox.categorical.data(names[2], names[32]) # Gender VS. SUM.OLD   
Q18 = wilcox.categorical.data(names[2], names[33]) # Gender VS. SUM.NEW

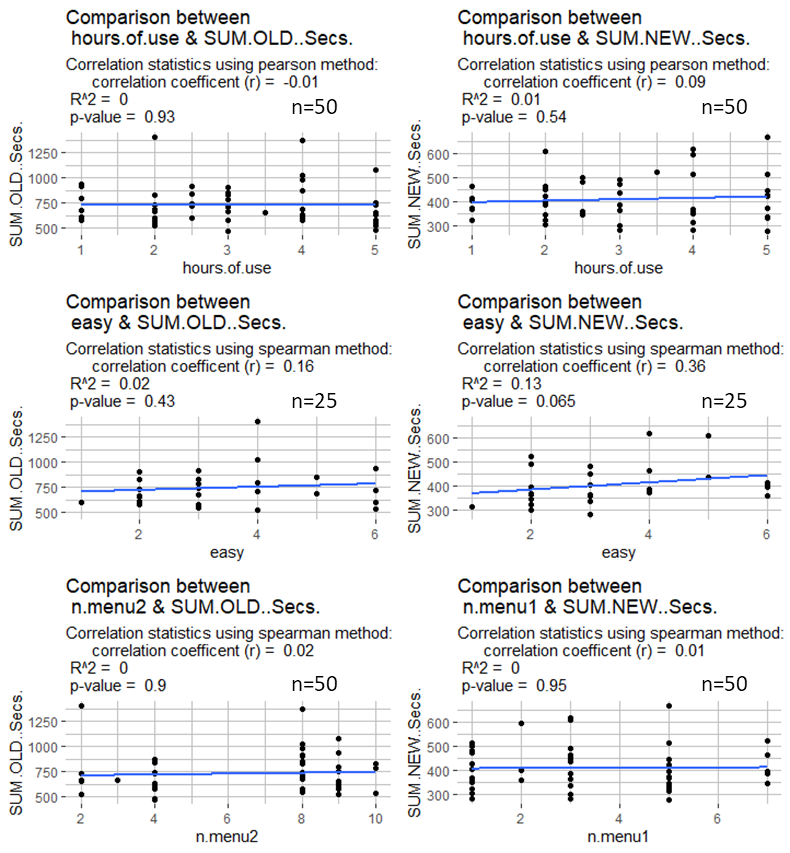
## שאלות המחקר

1. האם קיים קשר בין זמן שימוש בסלולרי ביום לבין זמן ביצוע כל המשימות באפליקציה הקיימת?
2. האם קיים קשר בין זמן שימוש בסלולרי ביום לבין זמן ביצוע כל המשימות באפליקציה החדשה?
3. האם קיים קשר בין קלות שימוש הנתפסת באפליקציה לבין זמן ביצוע כל המשימות באפליקציה הקיימת?
4. האם קיים קשר בין קלות שימוש הנתפסת באפליקציה לבין זמן ביצוע כל המשימות באפליקציה החדשה?
5. האם קיים קשר בין שביעות הרצון מכמות התפריטים באפליקציה החדשה לבין זמן ביצוע כל המשימות באפליקציה החדשה?
6. האם קיים קשר בין שביעות הרצון מכמות התפריטים באפליקציה הקיימת לבין זמן ביצוע כל המשימות באפליקציה הקיימת?
7. האם קיים קשר בין גיל הנבדק לבין זמן השימוש בסלולרי ביום?
8. האם קיים קשר בין שימוש באפליקציית שירותי בריאות (**בינארי**) לבין כמות הפעמים שהלקוח מבקר במרפאה בחודש?
9. האם קיים קשר בין מין (**בינארי**) לבין זמן ביצוע כל המשימות באפליקציה הקיימת?
10. האם קיים קשר בין מין B (**בינארי**) לבין זמן ביצוע כל המשימות באפליקציה החדשה?

## תשובות:

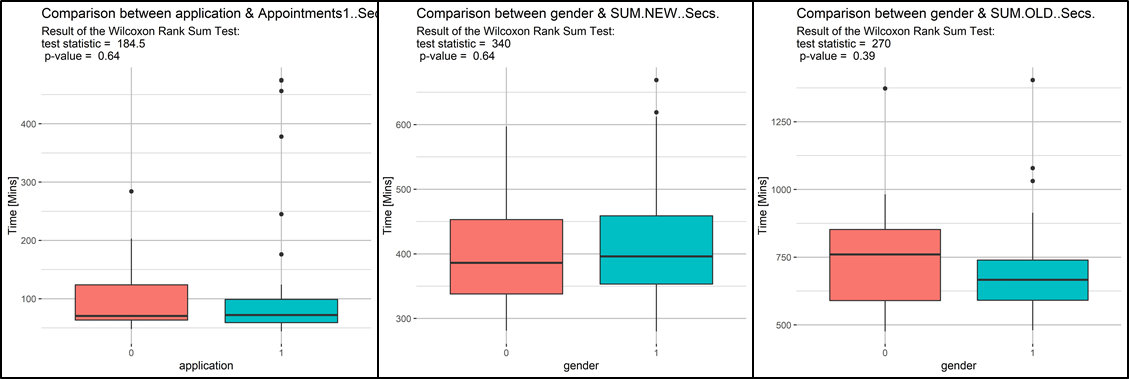
כפי שניתן לראות באיור 4, בכל הבדיקות נמצע ערך p value>0.05 ולכן עלינו לקבל את השערת ה-0 ולהסיק שלא קיים מתאם בין המשתנים. לפיכך, לא נמצא מתאם בין שימוש בסלולרי, קלות השימוש הנתפסת או שביעות הרצון מכמות התפריטים לבין זמן ביצוע הפעולות באפליקציה (הישנה והחדשה).

איור 4 קורלציות בין זמן ביצוע כל הפעולות באפליקציה החדשה והישנה לבין גורמים שונים. הנתונים לגבי האפליקציה הישנה מופיעים בצד שמאל ובאפליקציה החדשה בימין. המדד שמולו נבחנת הקורלציה רשום בתחתית כל גרף. ציר ה-Y מציג את זמן סך כל הפעולות בדקות. קו רגרסייה ליניארית מוצג בכחול. גודל המדגם מצויין באות n. בנוסף, המתאם בין הקבוצות נבחן באמצעות שיטת פירסון או ספירמן, כמפורט בראש כל גרף. ערכי r, R2 ו-pvalue מוצגים גם הם לכל מתאם.



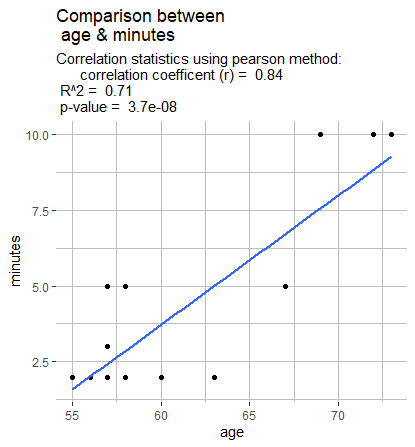
בנוסף, כפי שניתן לראות באיור 5, במבחן ווילקוקסון לא נמצא הבדל משמעותי בין הקבוצות השונות ולכן עלינו לקבל את השערת ה-0 שבהנתן מקרה אקראי, אין הבדל בין זמן ביצוע הפעולות בשתי הקבוצות. כלומר, לא נמצא קשר בין שימוש באפליקציות שירותי בריאות לכמות הפעמים שהלקוח מבקר במרפאה, או בין זמן ביצוע המשימות למין (הן באפליקציה הישנה והן בחדשה)

איור 5 השוואה בין ממוצע זמן ביצוע פעולות שונות לבין משתנים בינאריים (מין או שימוש באפליקצייה). ציר ה-Y מציג את זמן הפעולה. הנתונים מוצגים כקופסאות, כאשר הקו המרכזי מייצג את החציון, גבולות הקופסה מייצגים את רבעונים 25%, 75% והקווים החיצוניים (Whiskers) מייצגים את הערכים הנמצאים עד מרחק 1.5 של הטווח-הבין-רבעוני (Inter-Quantile-Range). נקודות חריגות מטווח זה מוצגות כנקודות. בכל המבחנים גודל המדגם הוא n=50. בנוסף, ההבדל בין הקבוצות נבחן באמצעות מבחן ווילקוקסון. ערך p value רשום בראש כל גרף.



לעומת זאת, כפי שניתן לראות באיור 4, נמצא מתאם חיובי חזק ומובהק בין גיל הנבדק לבין זמן השימוש בסלולרי ביום. חשוב להזהר בפרשנות בממצא זה כיוון שההתפלגות מוטה ומספר הדוגמאות מאוד קטן (ישנן מעט דוגמאות מעל 65 והן מושכות את כל המתאם מעלה).

איור 6 מתאם בין גיל הנבדק וזמן השימוש בסלולרי ביום. קו רגרסייה ליניארית מוצג בכחול. גודל המדגם n=25. בנוסף, המתאם בין הקבוצות נבחן באמצעות שיטת פירסון, ערכי r, R2 ו-pvalue מוצגים גם הם לכל מתאם. חשוב לציין שישנן דוגמאות חופפות ולכן מספר הנקודות נראה קטן מ-25.



כל ההשערות והתוצאות בנוגע למתאמים מפורטות בטבלה הבאה:

טבלה 4 סיכום ההשערות המבחנים והתוצאות לבדיקת מתאמים בין משתנים.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **משתנה א'** | **משתנה ב'** | **מבחן** | **השערות** | **תוצאה** |
| hours.of.use | SUM.OLD.SECS.. | Pearson | H0: r=0 H1: r≠0 | p=0.93  לא ניתן לדחות את השערת H0 |
| hours.of.use | SUM.NEW.SECS.. | Pearson | H0: r=0 H1: r≠0 | p=0.54  לא ניתן לדחות את השערת H0 |
| easy | SUM.OLD.SECS.. | Spearman | H0: r=0 H1: r≠0 | p=0.43  לא ניתן לדחות את השערת H0 |
| easy | SUM.NEW.SECS.. | Spearman | H0: r=0 H1: r≠0 | p=0.065  לא ניתן לדחות את השערת H0 |
| n.menu2 | SUM.OLD.SECS.. | Spearman | H0: r=0 H1: r≠0 | p=0.9  לא ניתן לדחות את השערת H0 |
| n.menu1 | SUM.NEW.SECS.. | Spearman | H0: r=0 H1: r≠0 | p=0.95  לא ניתן לדחות את השערת H0 |
| application | Appointments1..SECS | Wilcoxon | H0: P(A > B) = P(B > A) = 0.5  H1: P(A >B) ≠ P(B > A) | p=0.64  לא ניתן לדחות את השערת H0 |
| gender | SUM.NEW.SECS.. | Wilcoxon | H0: P(A > B) = P(B > A) = 0.5  H1: P(A >B) ≠ P(B > A) | p=0.64  לא ניתן לדחות את השערת H0 |
| gender | SUM.OLD.SECS.. | Wilcoxon | H0: P(A > B) = P(B > A) = 0.5  H1: P(A >B) ≠ P(B > A) | p=0.39  לא ניתן לדחות את השערת H0 |
| age | minutes | Pearson | H0: r=0 H1: r≠0 | **p<0.05**  **ניתן לדחות את השערת H0** |

# שאלות מחקר בעלות שלושה משתנים:

## שיטת העבודה:

על מנת להתייחס לקשר בין 3 משתנים עלינו לבצע חישובי ANOVA ל-3 משתנים, או MANOVA (Multiple Analysise of Variance.

היות והמשתנים אינם מתפלגים נורמלית, איננו יכולים לבצע מבחן MANOVA רגיל. כדי שנוכל לבצע מבחן MANOVA נשתמש בשיטה הכוללת תמורות (permutations) על הנתונים כדי לדמות מדגם גדול ונורמלי. לשם כך השתמשנו במבחן ANOVA הכולל תמורות לפי Anderson1,2.

במבחן הנ"ל, מחושב מבחן ANOVA רגיל מספר רב של פעמים, כאשר בכל פעם (כל פרמוטציה) הדוגמאות מתערבבות בין הקבוצות. באופן זה אנחנו בוחנים האם ערבוב בין הקבוצות משפיע על ההבדל ביניהן. במקרה שבו הערבוב מעלה את ערך F, ניתן לומר שהסדר האמיתי (לפני הערבוב) הוא הסדר בעל ההבדל המובהק ביותר מכל האפשרויות, ולכן ניתן להשתמש בערך ה-F להסקת מסקנות. במקרה שבו הערבוב מוריד או שאינו משנה את ערך F, ניתן להסיק שהסדר הקיים הוא אקראי ואין הבדל מובהק בין הקבוצות. באופן ריאלי, ברוב המקרים יהיה יחס של פרמוטציות שבהן ערך F עלה לעומת אלה שבהן הוא ירד. היחס הזה משמש אותנו כערך F להערכת מובהקות ההבדל בין הקבוצות. להעמקה בנושא ניתן לקרוא את המאמרים של אנדרסון מ-2001 ו-2003 (ובנימה אישית מומלץ לקרוא את הבלוג [הבא](https://www.ohbmbrainmappingblog.com/blog/a-brief-overview-of-permutation-testing-with-examples))

במבחן ה-ANOVA, השערת ה-0 היא כי לא קיים הבדל בממוצעים בין הקבוצות:

H0: μ1 = μ2

H1: μ1 ≠  μ2

כאשר אנחנו בוחנים את השוני ב-3 משתנים, אנחנו בודקים את כל הקומבינציות ולכל אחת מהן יש השערת 0 משלה.

[1] Anderson, M.J. and Robinson, J. (2001), Permutation Tests for Linear Models. Australian & New Zealand Journal of Statistics, 43: 75-88. doi:

[2] Anderson M.J. and Ter Braak C. (2003), Permutation tests for multi-factorial analysis of variance, Journal of Statistical Computation and Simulation, 73:2, 85-113, DOI: [10.1080/00949650215733](https://doi.org/10.1080/00949650215733)

### קוד:

בבדיקה זו השתמשנו בפקודת adonis מחבילת Vegan, והשתמשנו ב-1000 פרמוטציות:

#### 2-way ANOVA on non-normal data ####  
  
#Parameters needed from original data: Age [3], font 2 [13], Sum.new [33], Sum.old [32]  
data.manova = data[,c(3,13,33,32)] #makes a new dataframe with only the relevant columns  
data.manova.filtered = data.manova %>% drop\_na() #removes any lines with missing data (NAs)  
colnames(data.manova.filtered) = c("age", "font", "sum.new", "sum.old") #renames the colomns for easy writing downstream  
  
res.aov.Q1 <- adonis(sum.new ~ age \* font, data = data.manova.filtered, permutations = 1000) # Calculates the 2-way MANOVA with 1,000 permutations  
res.aov.Q2 <- adonis(sum.old ~ age \* font, data = data.manova.filtered, permutations = 1000) # Calculates the 2-way MANOVA with 1,000 permutations

## שאלות המחקר

1. האם קיים קשר בין גיל ושביעות הרצון מגודל הפונט באפליקציה החדשה לבין זמני ביצוע המשימות באפליקציה החדשה?
2. האם קיים קשר בין גיל ושביעות הרצון מגודל הפונט באפליקציה הקיימת לבין זמני ביצוע המשימות באפליקציה הקיימת?

## תשובות

1. האם קיים קשר בין גיל ושביעות הרצון מגודל הפונט באפליקציה החדשה לבין זמני ביצוע המשימות באפליקציה החדשה?

Call:

adonis(formula = sum.new ~ age \* font, data = data.manova.filtered)

Permutation: free

Number of permutations: 999

Terms added sequentially (first to last)

Df SumsOfSqs MeanSqs F.Model R2 **Pr(>F)**

age 1 0.00379 0.0037857 0.35606 0.00737 **0.544**

font 1 0.00949 0.0094898 0.89255 0.01848 **0.360**

age:font 1 0.00044 0.0004404 0.04142 0.00086 **0.894**

Residuals 47 0.49972 0.0106322 0.97329

Total 50 0.51343 1.00000

ניתן לראות מערכי ה-F כי לא קיים קשר בין המשתנים השונים. כלומר עלינו לקבל את השערת ה-0 כי הממוצעים דומים ולכן לא קיים קשר בין הגיל ושביעות הרצון מהפונט לבין זמני ביצוע המשימות באפליקציה החדשה.

1. האם קיים קשר בין גיל ושביעות הרצון מגודל הפונט באפליקציה הקיימת לבין זמני ביצוע המשימות באפליקציה הקיימת?

Call:

adonis(formula = sum.old ~ age \* font, data = data.manova.filtered,  
permutations = 1000)

Permutation: free

Number of permutations: 1000

Terms added sequentially (first to last)

Df SumsOfSqs MeanSqs F.Model R2 **Pr(>F)**

age 1 0.04447 0.044472 3.3004 0.06520 **0.08591**

font 1 0.00145 0.001448 0.1075 0.00212 **0.79121**

age:font 1 0.00285 0.002847 0.2113 0.00417 **0.66633**

Residuals 47 0.63331 0.013475 0.92850

Total 50 0.68208 1.00000

ניתן לראות מערכי ה-F כי לא קיים קשר בין המשתנים השונים. כלומר עלינו לקבל את השערת ה-0 כי הממוצעים דומים ולכן לא קיים קשר בין הגיל ושביעות הרצון מהפונט לבין זמני ביצוע המשימות באפליקציה הישנה.

**מסקנה**: לא נמצא קשר בין הגיל ושביעות הרצון מהפונט לבין זמן ביצוע המשימות באפליקציה, הישנה והחדשה.

טבלה 5 סיכום ההשערות המבחנים והתוצאות לבדיקת הבדלים בין הממוצעים של 3 משתנים

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **משתנה א'** | **משתנה ב'**  **(או משתנה ב' בהנתן ג')** | **מבחן** | **השערות** | **תוצאה** |
| SUM.NEW.SECS.. | Age | Permutation ANOVA | H0: μ1 = μ2 H1: μ1 ≠  μ2 | p=0.54  לא ניתן לדחות את השערת H0 |
| SUM.NEW.SECS.. | font | Permutation ANOVA | H0: μ1 = μ2 H1: μ1 ≠  μ2 | p=0.36  לא ניתן לדחות את השערת H0 |
| SUM.NEW.SECS.. | Age:font | Permutation ANOVA | H0: μ1 = μ2 H1: μ1 ≠  μ2 | p=0.0.89  לא ניתן לדחות את השערת H0 |
| SUM.OLD.SECS.. | Age | Permutation ANOVA | H0: μ1 = μ2 H1: μ1 ≠  μ2 | p=0.09  לא ניתן לדחות את השערת H0 |
| SUM.OLD.SECS.. | font | Permutation ANOVA | H0: μ1 = μ2 H1: μ1 ≠  μ2 | p=0.79  לא ניתן לדחות את השערת H0 |
| SUM.OLD.SECS.. | Age:font | Permutation ANOVA | H0: μ1 = μ2 H1: μ1 ≠  μ2 | p=0.67  לא ניתן לדחות את השערת H0 |

# הערה:

היו עוד 2 שאלות שביקשת לבחון, אבל הייתה בשתיהן אותה הבעיה:

1. האם קיים קשר בין שימוש באפליקציית שירותי בריאות AI וקלות השימוש הנתפסת AK לבין כמו הפעמים שהלקוח מבקר במרפאה בחודש R?

לא ניתן לבצע היות ויש הרבה ערכים חסרים (NA) ולאחר הסרתם נשאר רק משתתף אחד שלא משתמש באפליקציה:

times.in.clalit easy application

2 3 3 1

4 4 3 1

7 2 4 1

8 2 2 1

9 2 1 1

11 4 3 1

13 3 3 1

14 4 6 1

15 4 2 0

18 4 2 1

21 2 2 1

22 2 6 1

23 4 3 1

25 4 2 1

29 2 4 1

30 2 4 1

31 4 5 1

34 2 2 1

35 2 2 1

36 4 6 1

38 4 4 1

40 3 6 1

41 2 3 1

43 2 5 1

45 4 4 1

49 3 2 1

50 3 3 1

1. האם יש קשר בין כמות שעות השימוש בסלולרי AB (26) ושימוש באפליקציות השונות בסלולרי AD (28) לבין קלות השימוש באפליקציית כללית AK (35)?

לא ניתן לבצע היות ויש הרבה ערכים חסרים (NA) ולאחר הסרתם נשאר רק משתתף אחד שלא משתמש באפליקציה

data.manova.filtered

hours.of.use applications easy

2 3.0 1 3

4 3.0 1 3

7 2.0 1 4

8 1.0 1 2

9 4.0 1 1

11 2.5 1 3

13 5.0 1 3

14 2.0 1 6

15 3.0 0 2

18 3.0 1 2

21 2.0 1 2

22 1.0 1 6

23 2.5 1 3

25 3.0 1 2

29 2.0 1 4

30 1.0 1 4

31 3.0 1 5

34 2.0 1 2

35 1.0 1 2

36 2.5 1 6

38 3.0 1 4

40 2.0 1 6

41 1.0 1 3

43 2.0 1 5

45 4.0 1 4

49 3.5 1 2

50 2.0 1 3