

שאלה 1

א. עבור כל עשור נוצרה רשת וכמות המדינות שהיו מעורבות בסכוסכים בכל אחד מהעשורים היא:

72 :1950-1960

100 :1960-1970

100 :1970-1980

117 :1980-1990

131 :1990-2000

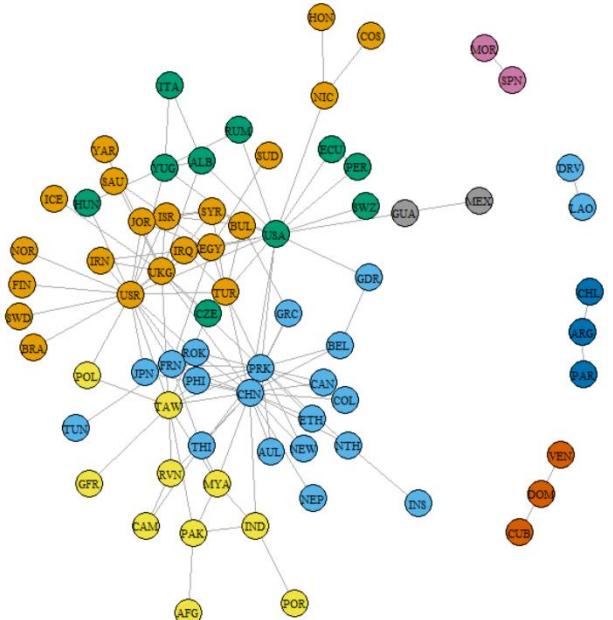
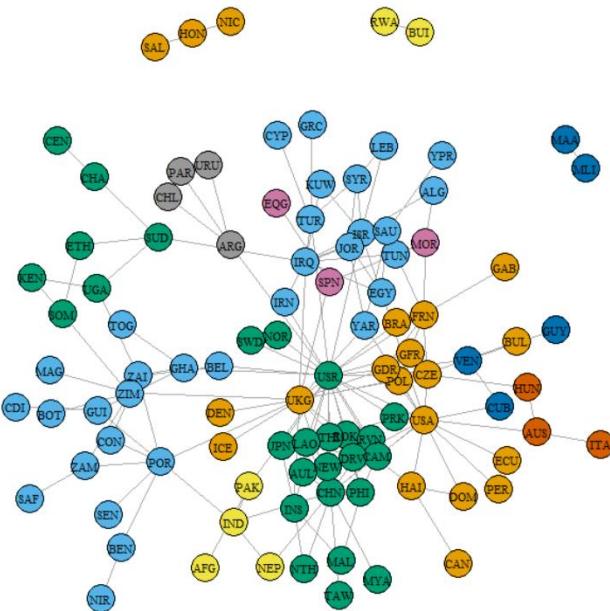
129 :2000-2010

.ב.

עבור שתי הרשותות הראשונות צציג גם את ההחזיה, עבור השאר רק את הנדרש.

רשת 1960 : 13 קבוצות, מודולריות 0.577

רשת 1950 : 10 קבוצות, מודולריות 0.472



[1] "TUN" "IRQ" "TUR" "YAR" "ISR" "JOR" "KUW" "SAU" "EGY" "SYR" "ALG" "YPR" "IRN"
[14] "CYP" "LEB" "GRC"

[1] "UKG" "USR" "EGY" "SYR" "ISR" "SAU" "IRN" "BUL" "TUR" "BRA" "ICE" "IRQ" "JOR"
[14] "NOR" "YAR" "SUD" "SWD" "FIN"

רשת 1970 : 16 קבוצות, מודולריות 0.605

[1] "ISR" "USA" "PRK" "ARG" "USR" "ECU" "IRQ" "SYR" "JOR" "SAU" "IRN" "CHL" "PAK"
[14] "URU" "AFG" "UKG" "PER" "CAN" "BOL" "JPN" "LEB" "KUW" "NOR"

רשת 1980 : 17 קבוצות, מודולריות 0.534

[1] "SAU" "EGY" "LIB" "FRN" "ISR" "YPR" "TUN" "BRA" "LEB" "OMA" "YAR" "NEW" "COM"

רשות 1990: 18 קבוצות, מודולריות 0.503

[1] "BEL" "CAN" "FRN" "GMY" "GRC" "ITA" "NTH" "SPN" "TUR" "UKG" "USA" "KUW" "RUS"
[14] "SAU" "ARM" "IRN" "IRQ" "PRK" "IND" "ZAI" "JOR" "ISR" "EGY" "BNG" "PAK" "ARG"
[27] "DOM" "POL" "SYR" "CUB" "AUL" "BAH" "OMA" "UAE" "QAT" "MOR" "HAI" "LAT" "LEB"

רשות 2000: 20 קבוצות, מודולריות 0.573

[1] "SWZ" "ISR" "EGY" "LEB"

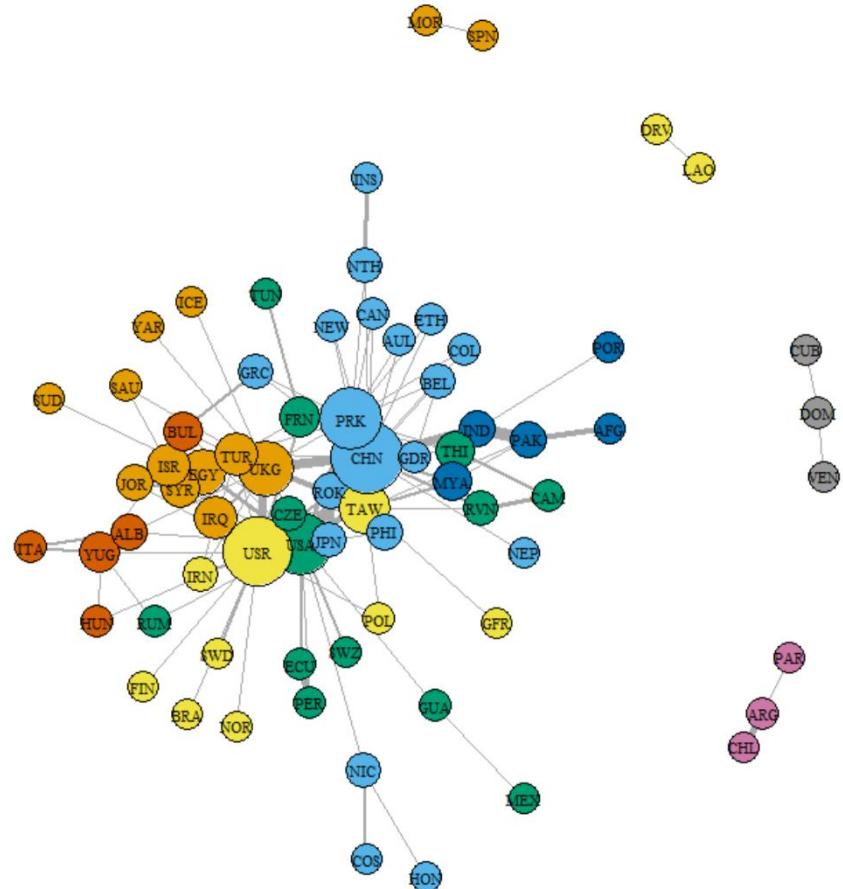
ניתן לראות ש מרבית המדינות שנמצאות עם ישראל באותה הקהילה לאוֹרְק השנים אלו הן המדינות הערביות במצרים התיכון, או מדינות קרובות גיאוגרפית. אך יש לשים לב Ci זה איננו חד-משמעי וגם לעיתים מצטרפות מדינות שאמורות להיות ביחסים טובים עם ישראל כמו אריה"ב לדוגמא ב-1970. כמות המדינות לרוב נעה בין 10-20, חוץ חרייגים (1990 לרעה, 2000 לטובה). סה"כ כמות המדינות שישראל סוגה אותה הקהילה לאוֹרְק העשורים היא 56. לבנון תמיד איתנו בקהילה החל מהעשור של 1960.

.ג.

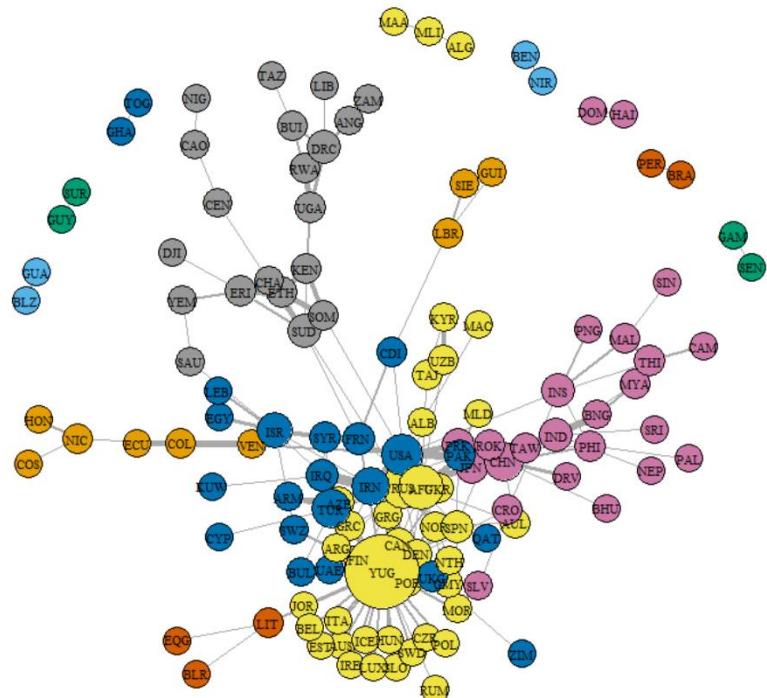
לטובת התצוגה, בניתי את הרשות כך שכן תהיה ממושקלת. המשקל מייצג מספר הסוכסוכים עם מדינה מסוימת.

צבעים לפי החלוקת לקהילות, גודל לפי הדרגה, עובי קשת לפי המשקל.

רשות 1950:



רשת 2000



.ג.

הוספתי את הקודקודים החסרים לכל רשות כדי שייו בועלות אותן קודקודים, ולאחר מכן יוכל לבצע .qaptest

QAP Test Results

Estimated p-values:

$p(f(\text{perm}) \geq f(d))$: 0.121
 $p(f(\text{perm}) \leq f(d))$: 0.928

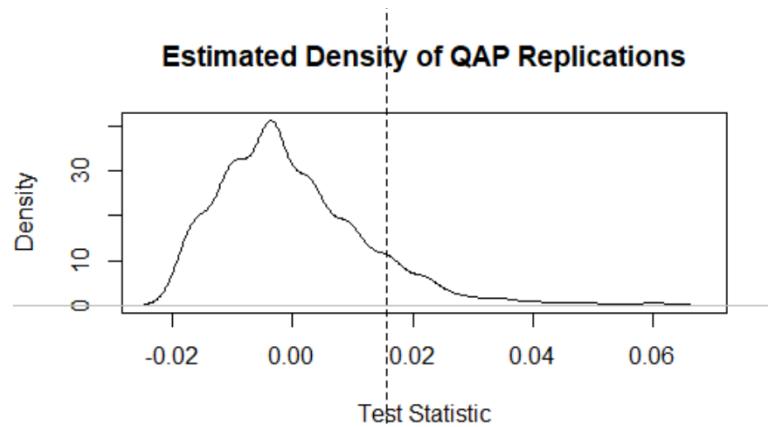
Test Diagnostics:

Test Value ($f(d)$): 0.01563618

Replications: 1000

Distribution Summary:

Min:	-0.01620681
1stQ:	-0.009838215
Med:	-0.003469617
Mean:	-5.604824e-05
3rdQ:	0.009267579
Max:	0.06021636



התתקבל קורלציה של 0.015, הבעה שהיא לא מובהקת, אך שלא באמת ניתן לקבוע כי יש קורלציה בין שתי הרשנות.

.ה.

זהו מפקד השלשות של הרשת המלאה (על פני כל השנים האפשריות):

0	1	2	3
1036188	158401	19105	3171

כדי לתמוך בטענה נרצה נריצה שיהיו קשרים יחסית מסווג 2 לעומת הקשרים האחרים. ככלומר שאכן במרבית הפעמים האויב שלו הוא האויב של חבר (משולש מסווג 2 מייצג סיטואציה זו). לעומת משולש מסווג 3 המייצג סיטואציה שבה האויב שלו גם האויב של אויבי, ומשולש מסווג 1 מייצג סיטואציה שבה האויב שלו הוא דואק לחבר.

לכן אגריל 100 רשותות רנדומליות ואבצע מבחן χ^2 לבדוק האם משולש מסווג 2 אצלנו ברשת גדול יותר בaczורה מובהקת ממה שהיינו מצפים לקבל ברשת רנדומלית שיוצרה מאותה הצפיפות ומאותו מפקד זוגות (בכדי לבצע השוואה מייצגת והוגנת).

מפקד הזוגות של הרשת שלנו הוא: 8068 קשרים הדדיים ו- 17847 קשרים ריקיים.

הנה הפלט של מבחן t:

One Sample t-test

```
data: triadNum2
t = -125.32, df = 99, p-value < 2.2e-16
alternative hypothesis: true mean is less than 19105
95 percent confidence interval:
-Inf 11083.39
sample estimates:
mean of x
10975.68
```

בוצע מבחן t חד-צדדי תחתון. השערת האפס היא שההתוחלת האמיתית (של קשרים מסוג 2) היא 19105, וההשערה האלטרנטטיבית טוענת שההתוחלת קטנה מכך (כלומר ש- μ הוא מספר גדול מדי לקשרים מסוג 2 ברשות רנדומליות דומות). קיבלנו p -value-ם נמוך מ-0.05 ולכן ההשערה נדחתה. לכן נוכל להגיד כי הרשת תומכת בטענה, ונראית כמות גדולה מהמצופה של קשרים מסוג 2 באופן מובהק.

שאלה 2

אתוייחס לרשת זו כרשות לא מכוונת מכיוון שקשת בה מייצגת חברות בין שני קודקודים (למרות שלפי ההסבר על הרשת יש הבדלה בין השולח למקבל אך איןפה רלוונטיות לטעמי לייצג את הכוון).

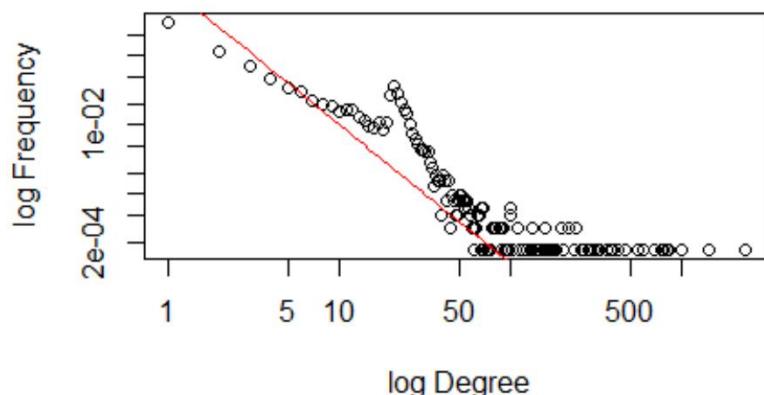
א.

ביצעתי מבחן `fit_power_law` של `igraph`, בו השערת האפס היא שהרשת מתפלגת כנטולות סקללה (החל מדרגה מינימלית 1). קיבלתי p (ערך קרוב מאוד לאפס), ולכן נדחה את ההשערה ונגיד שהרשות אינה נטולת סקללה.

אכן גраф הלוג-לוג של התפלגות הדרגה לא מתנהג כמו קו ישר שמתאר רשת נטולת סקללה.

מדד R^2 (מידת ההתאמה) יצא 0.5186.

Log-log of degree



ב.

בכדי לבחור את 316 הקודקודים שאוריך (חמיישה אחוז מתווך הקודקודים), אוחשב עבור כל קודקוד את

מדד טיב ההתאמנה R^2 שהיא מושג במידה והינו מוחקם אותו. לאחר מכן נבדוק עבור אילו קודקודים התקבלמדד טיב ההתאמנה גבוהה יחסית, ונמחק את אלו שבלעדיהם התקבלו 316 הערכים הגבוהים ביותר שלמדד טיב ההתאמנה.

לאחר שנחסיר 316 קודקודים אלו, נוכל לבדוק את מדד טיב ההתאמנה העדכני. מדד טיב ההתאמנה R^2 שהתקבל לאחר מחיקתם של 316 הקודקודים כפי שהסביר הינו 0.542, לעומת זאת שיפור של 2.31% במדד זה. יש לציין כי עדין לפני מבחן `fit_power_law` הרשות אינה מתפלגת כנטולית סקירה אך בעת היא מעת יותר קרובה לכך.

ג. הוגלו 100 רשותות רנדומליות וחושב אורך המסלול הממוצע בכל אחת ולאחר מכן מיצוע על פני כל הרשותות (L_r), ובאותו אופן חושב מקדם ההתקבצות עבור כל רשות ומיוצע על פני כל הרשותות (C_r).

חושב אורך המסלול הממוצע של רשות הפרויקט (L_{net}), וכן מקדם ההתקבצות של הרשות (C_{net}).

$$\frac{L_{net}}{L_r} = \frac{3.424}{3.656} \approx 1$$

$$\frac{C_{net}}{C_r} = \frac{0.028}{0.002} \approx 14$$

קיבלנו יחס שגובה מ-10 (ואותו סדר גודל של מסלול ממוצע), כך שהרשות עונה על הקритריון של עולם קטן. משמעות הדבר היא שברוב המקרים של צומת מסוים הם גם חברים ביניהם בלבד.

שאלה 3

.א.

הוגלה רשות מתאימה בעזרת `static.power.law.game` עם 1000 קודקודים, 124875 קשותות (0.25*(1000*999)/(2)), ומקדם חזקה 2.5. בסעיף זה עדין לא הוגלה עשר רשותות כאלה, כי לא התבקשו להפיק במהלך מדד/חישוב כלשהו.

.ב.

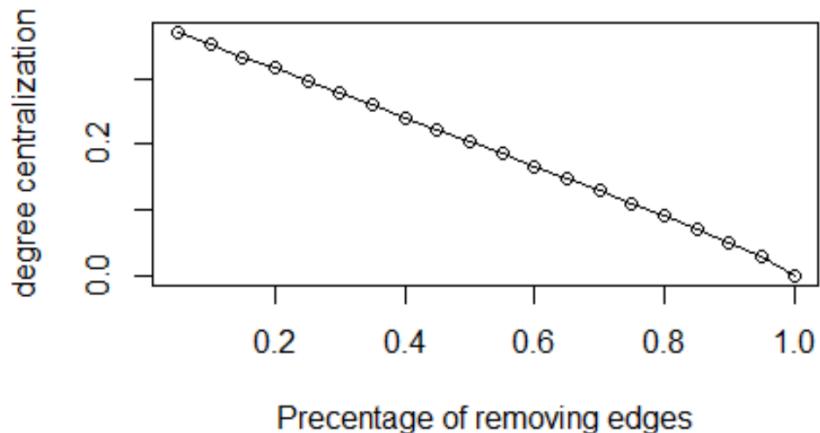
בסעיף זה הוגלה 10 רשותות לפי הדרישות של סעיף א', ובערתי בלולאה שהצעדים שלה הם הורדת האחזים הנדרשים מהקשותות. בתוך הלולאה הגדולה ישנה עוד לולאה פנימית שככל צעד שלא מורידה את הנקודות הנדרשת עבור כל אחת מהקשותות (כדי למסך על המקרים של יצירת הרשות הרנדומלית). גם בתוך הלולאה הפנימית ישנה עוד לולאה שמורידה את הנקודות הנדרשת מהרשות המסוימת עשר פעמים שונות, כלומר מגילה קשותות אחרות להורדה (כדי למסך על המקרים של בחירת הקשותות להורדה), ולכן מקבלת 10 מדדי מרכזיות שונים עבור רשות מסוימת אחת (ובbor כמוות מסוימת של קשותות להורדה) וממצעת אותן. לאחר מכן נמציע את עשר מדדי המרכזיות של כל רשות (שכאמור הם מבוססים על ממוצע עשר חזנות), ונקבל את המדד המרכזיות המתאים עבור כמוות

הקשנות הנדרשת להורדה. ככל מרבית מומחים מתאים לכמות מסוימת.

.ג.

להלן מדדי המרכזיות שחושבו כפי שהועסב בסעיף ב':

Percentage	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
Centralization	0.369	0.351	0.332	0.315	0.297	0.278	0.259	0.240	0.222	0.203	0.185	0.165	0.147	0.129	0.109	0.089	0.070	0.049	0.029	0.000



ניתן לראות שככל שמוסרים יותר קששות, כך ממד המרכזיות של הגרף יורד. שכן זה הגיוני משומש ככל שיש יותר קששות כך הרשות יותר דומה לרשות נתולות סקללה (כפי שיצרנו אותה) שמאופיינת במרכזיות גבוהה יחסית, וכך שמוסרים בה קששות, היא מאבדת את תכונה זו ונחפה לפחות לפחות מרכזיות.

שאלה 4

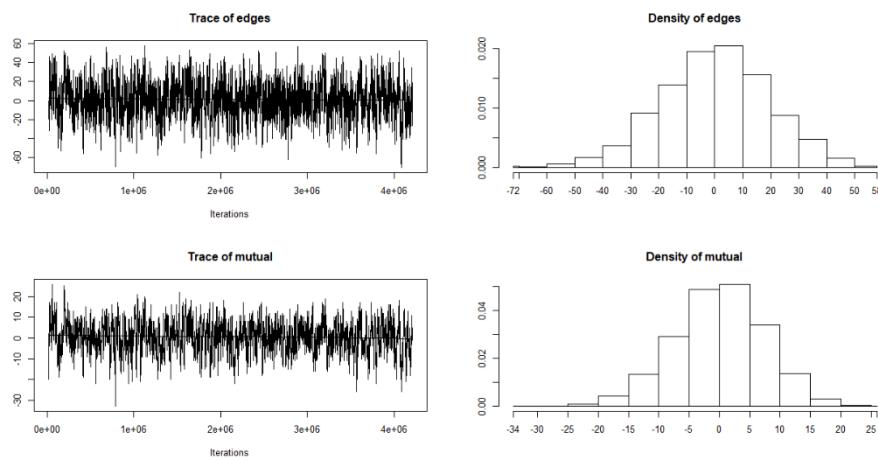
.א.

נוצרו המודלים לפי הדרישה:

```
#1
model1_fall <- ergm(fall_net~edges+mutual)
model1_spring <- ergm(spring_net~edges+mutual)
#2
model2_fall <- ergm(fall_net~edges+mutual+nodematch('Gender'))
model2_spring <- ergm(spring_net~edges+mutual+nodematch('Gender'))
#3
model3_fall <- ergm(fall_net~edges+mutual+nodematch('Gender')+absdiff('Grade'))
model3_spring <- ergm(spring_net~edges+mutual+nodematch('Gender')+absdiff('Grade'))
#4
model4_fall <- ergm(fall_net~edges+mutual+nodematch('Gender')+absdiff('Grade')+nodefactor('Community'))
model4_spring <- ergm(spring_net~edges+mutual+nodematch('Gender')+absdiff('Grade')+nodefactor('Community'))
```

.ב.

בכל הנитוחים הבאים נבדקה גם התכנסות של המודל ובדיקה האם אכן נדגם כל הטווח והיותה התכנסות סביב האפס. לראייה ה- mcmc.diagnostic של מודל 1 (שאר האבחונים מופיעים בקובץ):



: (1)

Formula: spring_net ~ edges + mutual

Iterations: 2 out of 20

Monte Carlo MLE Results:

	Estimate	Std. Error	MCMC %	z value	Pr(> z)
edges	-4.5028	0.1239	0	-36.33	<1e-04 ***
mutual	3.7070	0.3555	0	10.43	<1e-04 ***

Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ' '

Null Deviance: 8983.2 on 6480 degrees of freedom
Residual Deviance: 974.1 on 6478 degrees of freedom

AIC: 978.1 BIC: 991.7 (smaller is better.)

Formula: fall_net ~ edges + mutual

Iterations: 2 out of 20

Monte Carlo MLE Results:

	Estimate	Std. Error	MCMC %	z value	Pr(> z)
edges	-3.58759	0.07834	0	-45.80	<1e-04 ***
mutual	3.13876	0.20240	0	15.51	<1e-04 ***

Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ' '

Null Deviance: 9665 on 6972 degrees of freedom
Residual Deviance: 2244 on 6970 degrees of freedom

AIC: 2248 BIC: 2262 (smaller is better.)

ניתן לראות בשני המודלים שההדדיות יצאא מובהקת, וגם באותו סדר גודל וסימן. ככלומר קיימת נטייה להדדיות בשני הקורסים, והוא יחסית דומה. אך ניתן לראות כי עדין יש מרחק בין שני גודלים אלו של בערך 0.56 שזה 2.5 סטיות תקן של ההדדיות בראשת סטיי ו-1.6 סטיות תקן של ההדדיות בראשת אביב, כך שבהתחשב בסטיות התקן נראה שההדדיות לא זהה.

: (2)

```
Formula: spring_net ~ edges + mutual + nodematch("Gender")
Iterations: 2 out of 20
Monte Carlo MLE Results:
Estimate Std. Error MCMC % z value Pr(>|z|)
edges      -4.7181    0.1653     0 -28.537 <1e-04 ***
mutual      3.6752    0.3444     0 10.670 <1e-04 ***
nodematch.Gender 0.3814    0.1829     0  2.085  0.037 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null Deviance: 8983.2 on 6480 degrees of freedom
Residual Deviance: 969.3 on 6477 degrees of freedom
AIC: 975.3    BIC: 995.7    (smaller is better.)

Formula: fall_net ~ edges + mutual + nodematch("Gender")
Iterations: 2 out of 20
Monte Carlo MLE Results:
Estimate Std. Error MCMC % z value Pr(>|z|)
edges      -3.64361   0.09315    0 -39.116 <1e-04 ***
mutual      3.11858   0.19941    0 15.639 <1e-04 ***
nodematch.Gender 0.11205   0.10211    0  1.097  0.272
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null Deviance: 9665 on 6972 degrees of freedom
Residual Deviance: 2243 on 6969 degrees of freedom
AIC: 2249    BIC: 2270    (smaller is better.)
```

ניתן לראות ברשף הסטי המופיליה של מגדר יצא לא מובהקת, ככלומר היא לא באה לידי ביתוי (נשים מאותו המין לא בהכרח בוחרים זה את זה), ואילו ברשף האביב היא כן יצאה מובהקת עם מקדם חיובי, ככלומר אנשים מאותו המין מוטים להצביע אחד לשני יותר מאשר למשני האחר. لكن ברשף הסטי אין המופיליה וברשף האביב ישנה המופיליה מובהקת.

: (3)

```
Formula: fall_net ~ edges + mutual + nodematch("Gender") + absdiff("Grade")
Iterations: 2 out of 20
Monte Carlo MLE Results:
Estimate Std. Error MCMC % z value Pr(>|z|)
edges      -3.586702  0.098796    0 -36.304 <1e-04 ***
mutual      3.130530  0.197854    0 15.822 <1e-04 ***
nodematch.Gender 0.100795  0.102675    0  0.982  0.326
absdiff.Grade -0.006243  0.004805    0 -1.299  0.194
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null Deviance: 9665 on 6972 degrees of freedom
Residual Deviance: 2242 on 6968 degrees of freedom
AIC: 2250    BIC: 2277    (smaller is better.)
```

```
Formula: spring_net ~ edges + mutual + nodematch("Gender") + absdiff("Grade")
Iterations: 2 out of 20
Monte Carlo MLE Results:
Estimate Std. Error MCMC % z value Pr(>|z|)
edges      -4.05757  0.20349     0 -19.940 <1e-04 ***
mutual      3.47738  0.35492     0  9.798 <1e-04 ***
nodematch.Gender 0.37172  0.18649     0  1.993  0.0462 *
absdiff.Grade -0.04922  0.01092     0 -4.505 <1e-04 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null Deviance: 8983.2 on 6480 degrees of freedom
Residual Deviance: 941.3 on 6476 degrees of freedom
AIC: 949.3    BIC: 976.4    (smaller is better.)
```

ניתן לראות שברשף הסטי אין נטייה מובהקת לבחור/לא לבחור סטודנטים עם ציון דומה, ואילו ברשף האביב כן קיימת נטייה לבחור סטודנטים עם ציון דומה של. המקדם יצא מובהק ושלילי, ככלומר ככל שההפרש בין שני ציוני הסטודנטים גדול כך הסיכון קטן (ולהיפך- ככל שההפרש קטן כך הסיכון גדול), لكن ההשערה נכונה לגבי רשות האביב

: (4)

```
Formula: fall_net ~ edges + mutual + nodematch("Gender") + absdiff("Grade") +
nodefactor("Community")
```

Iterations: 2 out of 20

Monte Carlo MLE Results:

	Estimate	Std. Error	MCMC %	z value	Pr(> z)
edges	-4.153901	0.809384	0	-5.132	<1e-04 ***
mutual	3.000858	0.204103	0	14.703	<1e-04 ***
nodematch.Gender	0.110210	0.104846	0	1.051	0.2932
absdiff.Grade	0.005653	0.005528	0	1.023	0.3065
nodefactor.Community.2	0.111278	0.421130	0	0.264	0.7916
nodefactor.Community.3	0.350949	0.404280	0	0.868	0.3853
nodefactor.Community.4	0.325548	0.409124	0	0.796	0.4262
nodefactor.Community.5	0.410815	0.416833	0	0.986	0.3243
nodefactor.Community.NA	-0.903488	0.476928	0	-1.894	0.0582 .

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Null Deviance: 9665 on 6972 degrees of freedom

Residual Deviance: 2196 on 6963 degrees of freedom

AIC: 2214 BIC: 2276 (smaller is better.)

```
Formula: spring_net ~ edges + mutual + nodematch("Gender") + absdiff("Grade") +
nodefactor("Community")
```

Iterations: 2 out of 20

Monte Carlo MLE Results:

	Estimate	Std. Error	MCMC %	z value	Pr(> z)
edges	-3.69204	0.42588	0	-8.669	< 1e-04 ***
mutual	3.04289	0.36599	0	8.314	< 1e-04 ***
nodematch.Gender	0.39110	0.19099	0	2.048	0.040589 *
absdiff.Grade	-0.03846	0.01130	0	-3.403	0.000666 ***
nodefactor.Community.2	0.06547	0.21527	0	0.304	0.761039
nodefactor.Community.3	0.26250	0.22276	0	1.178	0.238622
nodefactor.Community.4	0.38631	0.33296	0	1.160	0.245954
nodefactor.Community.NA	-0.96586	0.24136	0	-4.002	< 1e-04 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Null Deviance: 8983.2 on 6480 degrees of freedom

Residual Deviance: 882.7 on 6472 degrees of freedom

AIC: 898.7 BIC: 952.9 (smaller is better.)

ניתן לראות כי בראשת הסטי השפעת השיעיות לכהילה לא יצא מובהקת עבור כל אחת מהרמות, אך לא ניתן לקבוע שיש השפעה לכך, אף קבוצה איננה שונה באופן מובהק מקבוצת הבסיס (1). בראשת האביב רק קבוצת-NA יצאת שונה משאר הקבוצות, ככלומר סטודנטים שלא התקבלה עבורם תשובה לשאלת זו (ערוך חסר) יצאו שונים משאר הקבוצות.

.ג.

כדי לענות על שאלה זו נעזר במודל 3 (שכלול בתוכו את ההדדיות, את המגדר ואת הפרש ה齧וינט), נציג את הערכים המתאימים, ונחשב את logit ההפוכה של צירוף מקדמי הבתאות והערכים המתאימים. הערכים שנציג הם: אפס להדדיות (מניח שהכוונה של השאלה הייתה מה הסיכוי שיוויזר קשר בין שניהם ולא שני קשרים), 1-less-nodematch של המגדר (אותו המין), ו-0-less-absdiff של ה齧וינט (כי לשנייהם ממוצע 90). הנחתה 'ועל' להשתמש באחד המודלים הקיימים בכך כדי לחשב הסתרות זו ולא לבנות אחד חדש, ולכן לא בנית מודל חדש עם (diff=T,'Gender', nodematch), בצד' לתמוך בסוגיות שני הגברים ולא שני אותו המין. לפי דרישות בנית המודל בסעיף א לא היה נראה שמתבקש למצוא המופiliaה לכל הרמות.

$$\text{Invlogit}(x) = \text{logit}^{-1}(x) = e^x / (1 + e^x)$$

לכן, ההסתבריות המתאימות עבור שתי הרשותות זו:

$$p_{fall} = \text{invlogit}(-3.687 + 0 * 3.13 + 1 * 0.1008 + 0 * (-0.006)) = 0.0297$$

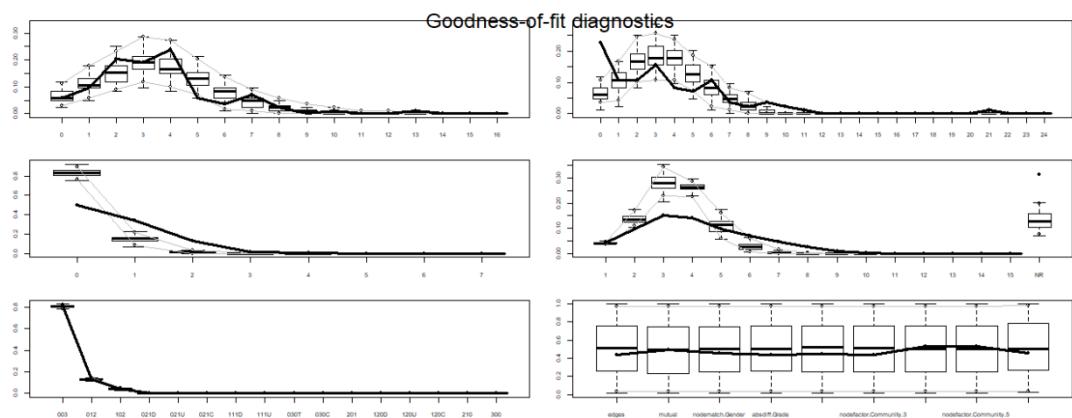
$$p_{spring} = \text{invlogit}(-4.058 + 0 * 3.477 + 1 * 0.372 + 0 * (-0.049)) = 0.0245$$

.2

נבחן את טיב ההתאמה של מודל 4 על פני פרמטרים שונים בראש הסתו:

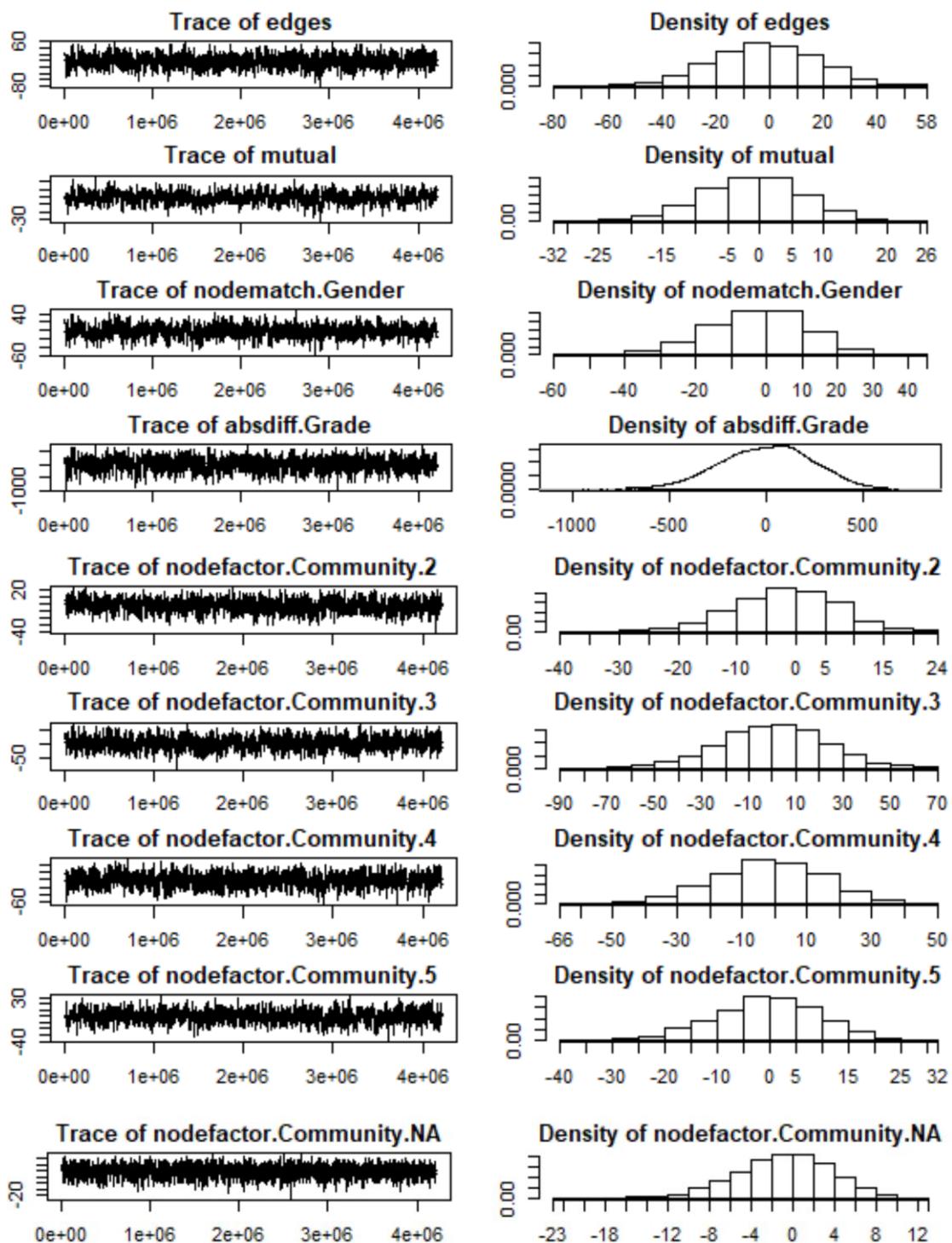
Goodness-of-fit for in-degree						Goodness-of-fit for out-degree						Goodness-of-fit for edgewise shared partner					
obs	min	mean	max	MC	p-value	obs	min	mean	max	MC	p-value	obs	min	mean	max	MC	p-value
0	5	2	5.50	12	1.00	0	19	1	4.98	12	0.00	esp0	147	208	244.00	277	0
1	8	5	9.57	15	0.78	1	9	4	10.24	18	0.78	esp1	100	18	42.58	72	0
2	17	8	14.07	21	0.46	2	9	8	13.99	26	0.10	esp2	39	0	3.86	10	0
3	16	7	16.04	26	1.00	3	13	7	15.79	23	0.56	esp3	5	0	0.25	3	0
4	20	7	14.06	25	0.12	4	7	7	14.40	26	0.08	esp4	3	0	0.02	1	0
5	5	5	11.18	19	0.02	5	6	5	11.16	19	0.06	Goodness-of-fit for minimum geodesic distance					
6	3	2	7.08	13	0.12	6	9	1	6.93	17	0.52	1	294	233	290.71	348	0.90
7	6	0	3.65	8	0.32	7	3	0	3.66	11	0.96	2	692	612	937.31	1301	0.04
8	2	0	1.66	5	1.00	8	2	0	1.71	7	0.92	3	1052	1308	1944.83	2569	0.00
9	0	0	0.75	6	1.00	9	3	0	0.76	4	0.14	4	989	1565	1838.81	2085	0.00
10	1	0	0.26	2	0.50	10	2	0	0.23	3	0.04	5	678	342	795.64	1151	0.48
11	0	0	0.12	1	1.00	11	1	0	0.10	2	0.18	6	495	30	208.68	519	0.02
12	0	0	0.03	1	1.00	12	0	0	0.05	1	1.00	7	319	0	44.45	212	0.00
13	1	0	0.02	1	0.04	21	1	0	0.00	0	0.00	8	176	0	8.76	83	0.00
14	0	0	0.01	1	1.00							9	63	0	1.94	52	0.00
												10	19	0	0.40	19	0.02
												11	7	0	0.08	4	0.00
												12	3	0	0.01	1	0.00
Goodness-of-fit for triad census						Goodness-of-fit for model statistics						Goodness-of-fit for edges					
obs	min	mean	max	MC	p-value	edges	min	mean	max	MC	p-value	edges	min	mean	max	MC	p-value
003	77355	74891	77235.52	79313	0.90	edges	294.00	233.00	290.710	348.00	0.90	edges	294.00	233.00	290.710	348.00	0.90
012	12494	10823	12687.03	14243	0.68	mutual	57.00	33.00	56.450	84.00	0.94	mutual	57.00	33.00	56.450	84.00	0.94
102	4022	2873	4025.32	5168	0.96	nodematch.Gender	156.00	127.00	153.750	197.00	0.88	nodematch.Gender	156.00	127.00	153.750	197.00	0.88
021D	419	130	184.85	253	0.00	absdiff.Grade	2151.49	1592.95	2167.714	2796.89	0.94	absdiff.Grade	2151.49	1592.95	2167.714	2796.89	0.94
021U	195	113	184.39	258	0.64	nodefactor.Community.2	53.00	29.00	52.810	78.00	0.96	nodefactor.Community.2	53.00	29.00	52.810	78.00	0.96
021C	222	255	368.40	474	0.00	nodefactor.Community.3	270.00	219.00	268.000	323.00	0.88	nodefactor.Community.3	270.00	219.00	268.000	323.00	0.88
111D	181	170	241.63	370	0.04	nodefactor.Community.4	175.00	124.00	172.900	215.00	0.90	nodefactor.Community.4	175.00	124.00	172.900	215.00	0.90
111U	238	153	241.47	334	1.00	nodefactor.Community.5	68.00	41.00	65.880	88.00	0.90	nodefactor.Community.5	68.00	41.00	65.880	88.00	0.90
030T	29	2	11.37	23	0.00	nodefactor.Community.NA	17.00	6.00	16.790	30.00	1.00	nodefactor.Community.NA	17.00	6.00	16.790	30.00	1.00
030C	1	0	3.86	10	0.24												
201	66	34	78.30	161	0.68												
120D	13	0	3.95	11	0.00												
120U	14	0	3.94	12	0.00												
120C	5	1	7.78	18	0.52												
210	21	1	5.58	13	0.00												
300	9	0	0.61	4	0.00												

מבחינת הדרגה הנכנסת וגם היוצאת, ישנה התאמה טוביה מאוד. כמעט כל הערכים ה-pvalue גבוה יחסית. מבחינת esp, התאמה איננה טוביה באף אחד מהסוגים. גם Über המרחקים הגאודזים המודל התאמה איננה טוביה במיוחד, אלא רק Über מרחוקים מסוימים ומעטים. מבחינת סוגי



המשולשים, רוב ההתאמה הינה טובה (כחזיות מכך הערכים ובעיקר עבור סוגים מסוימים הראשיונים).
כמוון שסבירות ערכיו הפרמטרים של המודל ההתאמה טובה (כפי שהיינו מצפים כי כך יוצרים את המודל). ניתן לראות באופן ויזואלי את המתואר בגרפים הבאים:

כמו כן, כשלב נדרש לבחינת טיב המודל, בוצעה בדיקה האם הערכים נדגו על כל ערכי הפרמטרים האפשריים והאם התרחשה התכנסות עבורה. ניתן לראות לפי הגרפים הבאים שאכן עבור כל הפרמטרים הנבדקים מתקיימים תנאים אלו:



הוסףתי תכונה של התחממות סטודנט. בגלל שראייתי שחלק גדול מן הסטודנטים שייכים להתחממות 'me', ולשאר התחממות יש שכיחיות נמוכות יחסית, הגדרתי את התכונה כבעל שתי רמות, כלומר האם שיש להתחממות 'me' או שיש-'Other'. הוסףתי את תכונה זו תחילת כ-nodematch כדי לבדוק האם יש המופיליה במוגמת התחממות אבל זה לא תרם למודל. لكن בדקתי האם הוספה של תכונה זו כ-nodefactor יכולה להועיל, וכן אם המודל השתפר מעט ומשתנה זה יצא מובהק. בדקתי אפשרויות community (כמו ארץ זה לא עזר למודל, אך נשארתי עם התצורה הזאת).

```
Formula: fall_net ~ edges + mutual + nodematch("Gender") + absdiff("Grade") +
nodefactor("Community") + nodefactor("Major")
```

```
Iterations: 2 out of 20
```

```
Monte Carlo MLE Results:
```

	Estimate	Std. Error	MCMC %	z value	Pr(> z)
edges	-3.84322	0.81737	0	-4.702	<1e-04 ***
mutual	2.99709	0.20634	0	14.525	<1e-04 ***
nodematch.Gender	0.11722	0.10393	0	1.128	0.2594
absdiff.Grade	0.00496	0.00553	0	0.897	0.3697
nodefactor.Community.2	0.04119	0.42624	0	0.097	0.9230
nodefactor.Community.3	0.28450	0.40959	0	0.695	0.4873
nodefactor.Community.4	0.25410	0.40963	0	0.620	0.5351
nodefactor.Community.5	0.34748	0.41588	0	0.836	0.4034
nodefactor.Community.NA	-0.90816	0.46914	0	-1.936	0.0529 .
nodefactor.Major.Other	-0.14772	0.07303	0	-2.023	0.0431 *
<hr/>					
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

```
Null Deviance: 9665 on 6972 degrees of freedom
Residual Deviance: 2192 on 6962 degrees of freedom
```

```
AIC: 2212      BIC: 2281      (smaller is better.)
```

שאלה 5

הרשות נקלטה ונחפכה ללא מכונת (לפי ההנחה בפורום).

א. נוענה על שאלה זו בעזרת שלושה מדדי מרכזיות ונבדוק מי מבין הדמויות מופעה באותו מקום גבוי ברובם.

Betweenness:

Closeness:

Degree:

Eddard Stark	6926.86414
Sansa Stark	6165.66652
Tyrion Lannister	5617.48229
Tywin Lannister	5070.39500
Joanna Lannister	4737.52448

Sansa Stark	0.0002013288
Tyrion Lannister	0.0002012882
Tywin Lannister	0.0002011668
Joanna Lannister	0.0002005616
Eddard Stark	0.0002002804

Quellon Greyjoy	12
Walder Frey	10
Oberyn Martell	10
Eddard Stark	9
Catelyn Stark	8

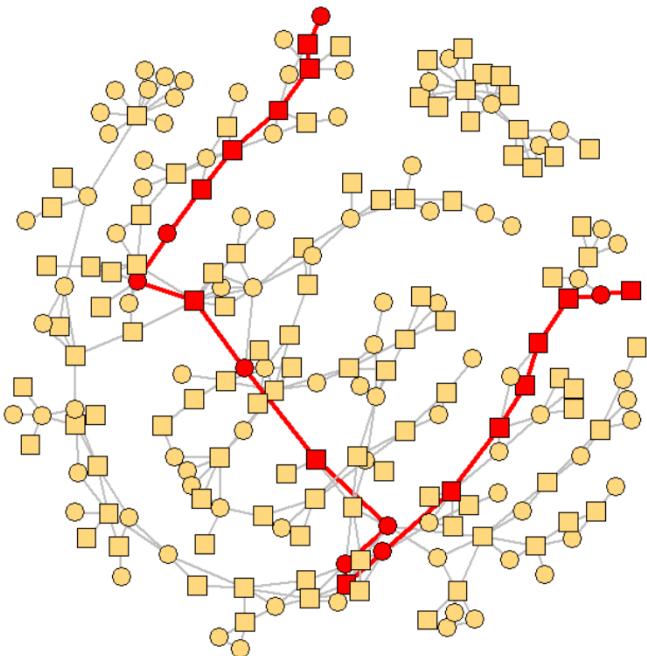
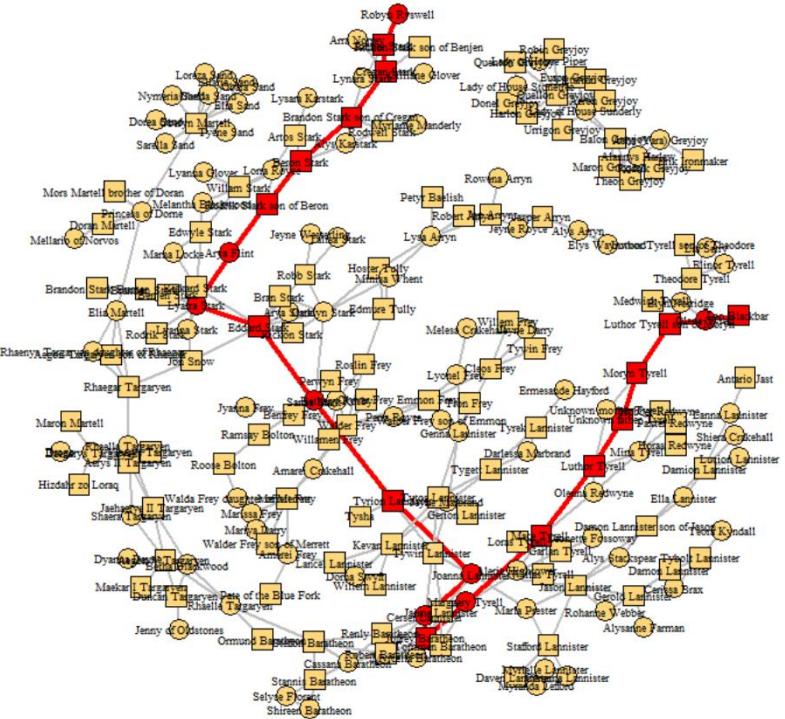
ניתן לראות כי אדוארד סטארק הוא היחיד שנמצא בין חמישת הראשוניים בכל שלושת המדדים. גם סנסה סטארק, ג'ימי לאניסטר וטיריוון דמויות מרכזיות מאוד שבממד הדרגה התבגרו קצת לאחר חמישת הראשוניים ולכן לא נכנסו לחמישייה שלן.

.ב.

קווטר הרשת הינו באורך 21 קשחות (דרך הקודקודים הבאים):

[1] Robyn Ryswell	Jonnel Stark	Cregan Stark
[4] Brandon Stark son of Cregan	Beron Stark	Rodrik Stark son of Beron
[7] Arya Flint	Lyarra Stark	Eddard Stark
[10] Sansa Stark	Tyron Lannister	Joanna Lannister
[13] Cersei Lannister	Joffrey Baratheon	Margaery Tyrell
[16] Mace Tyrell	Luthor Tyrell	Unknown father Tyrell
[19] Moryn Tyrell	Luthor Tyrell son of Moryn	Olene Tyrell
[22] Leo Blackbar		

החזית קווטר (עם ובלי שמות הקודקודים):



.ג.

זהו מפקד השלשות של הרשת (היות והיא רשת לא מכונת יש ארבעה סוגים אפשריים):

0	1	2	3
1412100	65261	790	105

נగזר 100 רשתות רנדומליות עם אותו מפקד זוגות כמו הרשת שלנו (326 הדרדים ו-21202 ריקים), ונשווה כל סוג שלשה לוקטור המתאים שיצרו הרשתות הרנדומליות בעזרת מבחן t , ונבדוק עבור איזה סוג מתקיים התוצאה החירגה ביותר.

ויג 0:

```
One Sample t-test

data: triadNum0
t = -0.7515, df = 99, p-value = 0.4541
alternative hypothesis: true mean is not equal to 1412100
95 percent confidence interval:
1411037 1412579
sample estimates:
mean of x
1411808
```

ויג 1:

```
One Sample t-test

data: triadNum1
t = 0.46035, df = 99, p-value = 0.6463
alternative hypothesis: true mean is not equal to 65261
95 percent confidence interval:
64687.43 66181.11
sample estimates:
mean of x
65434.27
```

ויג 2:

```
One Sample t-test

data: triadNum2
t = 17.571, df = 99, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 790
95 percent confidence interval:
983.7376 1033.0624
sample estimates:
mean of x
1008.4
```

ויג 3:

```
One Sample t-test

data: triadNum3
t = -343.38, df = 99, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 105
95 percent confidence interval:
4.754052 5.905948
sample estimates:
mean of x
5.33
```

ניתן לראות כי עבור שני הסוגים הראשונים, לא נדחתת השערת האפס, כלומר הם לא שונים באופן מובהק מסוגי הקשרים המתאימים ברשומות רנדומליות. בשני הקשרים האחרים ניכר השוני, בראשת שלמו יש יותר קשרים מסוג 3 ופחות קשרים מסוג 2, ושניהם אלו הבדלים מובהקים. החרג יתיר הוא קשר מסהפר 3 שאצלנו גבוהה בהרבה יותר מהסוג המתאים ברשת רנדומלית ($t = -343.38$), לעומת זאת מסוג 2 שגם חרג מאד אבל פחות ($t = 17.57$). ככל שהערך המוחלט של ה- t -גבוה יותר כך זו תוצאה חריגה יותר. ניתן להסביר זאת בכך שהקששות מיצגות קשרים משפחתיים ולכן הגיוני לצפות שאם שני קודקודים מחוברים לקודקוד מסווג (כolumbia בני משפחה שלו), הגיוני שהם גם מחוברים ביניהם וגם יש ביניהם קרבה משפחתית.

.7

הקליקה המקסימלית בגרף זה בגודל 3 ויש 105 חברים.

אריה סטארק נמצאת בклיקה אחת מהן והיא:

[[43]]

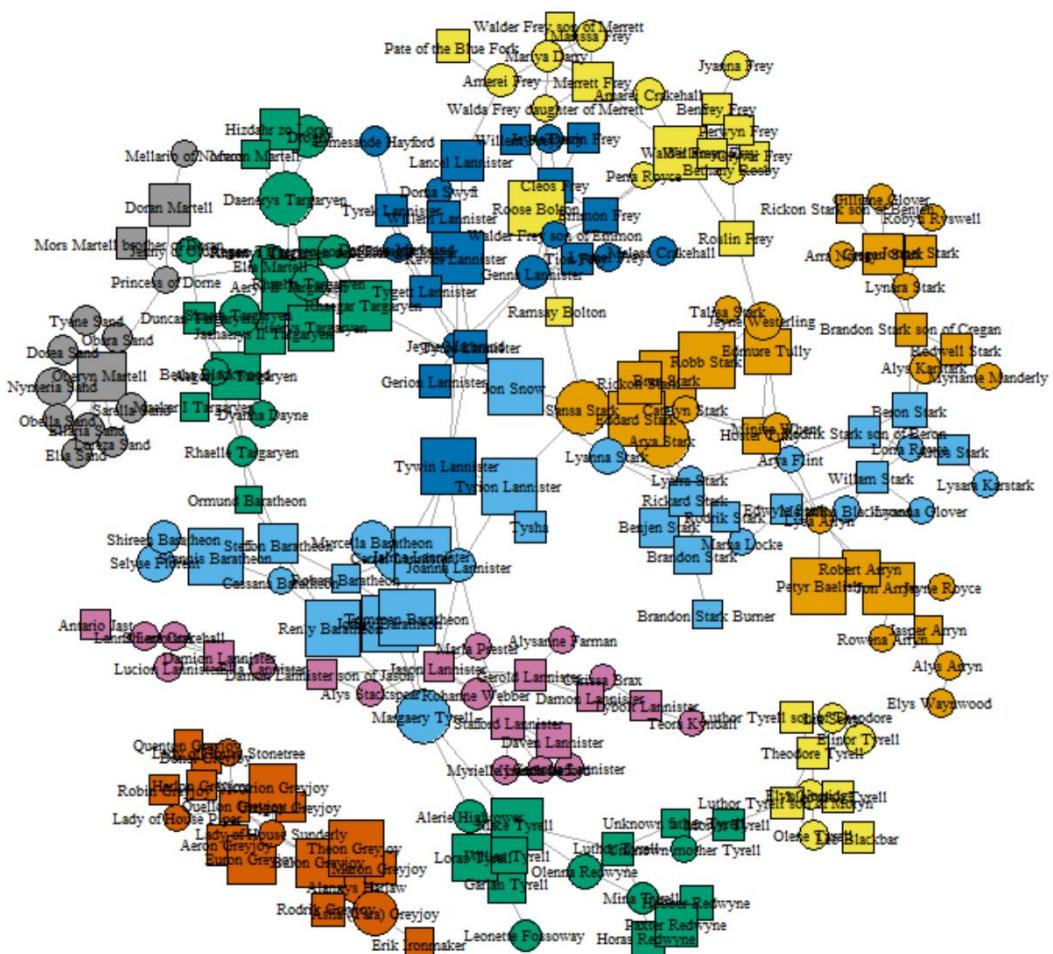
+ 3/208 vertices, named, from b9a7f90:
 [1] Arya Stark Catelyn Stark Eddard stark

הקליקה כאמור בגודל 3, והחברים בה (מלבד אריה) הם קטלין סטארק ואדוארד סטארק (שני הוריה).

.8

בחרתי לנתח לפि האופן הבא:

ביצעת חלוקה לקהילות (לפי edge.betweenness) שהפיקה את החלוקה הטובה ביותר ביותר עם מודלים של 0.836 (0.0), ואלו הם הצבעים בראשת. השתמשתי בשתי תוכנות של הגרף שזה המגדיר (הצורות בראשת-ריבוע מייצג זכר ועיגול מייצג נקבה), והפופולריות (שבאה לידי ביטוי בגודל הקודקוד- פופולריות גדולה יותר כך הקודקוד גדול יותר).



תצוגה זו עוזרת לנו להבין קודם כל את החלוקה לקהילות והאם קהילות אלו מייצגות לדוגמא את הקשר המשפחתי (שנראה שאכן מייצגות זאת יחסית). יתרה מכך, ניתן לראות האם קהילות/משפחות מסוימות מאופייניות בדמותות יותר פופולריות, והאם יש קשר אובי בין הפופולריות לבין מגדר הדמות (ב>Show Results נבדק בין תוחלות עברו בדיקת השערת זו).

לטובת המבחן יצרתי שני וקטורים, שהאחד מייצג את הפופולריות של דמויות הגברים והשני מייצג את הפופולריות של דמויות הנשים.

תוצאת המבחן:

Welch Two Sample t-test

```
data: popularityMale and popularityFemale
t = 4.58, df = 198.09, p-value = 8.21e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.1073138 0.2696046
sample estimates:
mean of x mean of y
0.3476152 0.1591559
```

מצאנו שקיים הפרש תוחלות מובהק ודוחינו את השערת האפס (שהפרש התוחלות הוא אפס).
א מייצג את הפופולריות של הגברים ו-ז מייצג את הפופולריות של הנשים. לפי הממצאים- הפופולריות של הגברים גבוהה יותר באופן מובהק מזו של הנשים. כמובן, מסיבה מסוימת (שאלוי מגיעה מתוך עלילות המלחמה הרבות), דמויות מהם גברים לרוב פופולריות יותר מדמותות שנן נשים.