**רגרסיה לינארית**

**פרויקט בניתוח סטטיסטי של מאגרי נתונים טבלאיים – חלק ב'**

[](https://www.google.co.il/imgres?imgurl=http://econedlink.org/img/lesson-images/EconEdLink-615-Inflation-Chart.jpg&imgrefurl=http://www.econedlink.org/lesson/615/What-causes-inflation&docid=FP4-y8aE2ds0aM&tbnid=ieas2P8teKejtM:&vet=10ahUKEwjSu8nJmafTAhXMA8AKHRTpAbk4ZBAzCFEoTjBO..i&w=620&h=200&bih=550&biw=865&q=inflation&ved=0ahUKEwjSu8nJmafTAhXMA8AKHRTpAbk4ZBAzCFEoTjBO&iact=mrc&uact=8)

שם הקורס: רגרסיה לינארית

שם המרצה: פרופ' ישראל פרמט, פרופ' הלל בר-גרא

בודקת הפרויקט: גב' לוי ירדן

סמסטר ב' תשע"ז

קבוצה: 9

תאריך הגשה: 29/6/17

**תוכן עניינים**

[1. תקציר מנהלים: 3](#_Toc486414945)

[1.1 המטרה: 3](#_Toc486414946)

[1.2 המשתנה המוסבר: 3](#_Toc486414947)

[1.3 המשתנים המסבירים: 3](#_Toc486414948)

[1.4 מקורות הנתונים: 3](#_Toc486414949)

[1.5 התצפיות: 3](#_Toc486414950)

[1.6 מהלך המחקר: 3](#_Toc486414951)

[1.6.1 התאמת משתנים: 3](#_Toc486414952)

[1.6.2 בחירת משתנים: 3](#_Toc486414953)

[1.6.3 בדיקת הנחות המודל: 4](#_Toc486414954)

[1.6.4 שיפור המודל: 4](#_Toc486414955)

[2. עיבוד מקדים: 5](#_Toc486414956)

[2.1 הגדרת המשתנים: 5](#_Toc486414957)

[2.2 הסבר על המשתנים: 5](#_Toc486414958)

[2.2.1 המשתנה המוסבר: 5](#_Toc486414959)

[2.2.2 המשתנים המסבירים: 5](#_Toc486414960)

[2.3 הסרה של משתנים: 6](#_Toc486414961)

[2.3.1 הסרה ע"פ מדד Pearson : 6](#_Toc486414962)

[2.3.2 הסרה ע"פ תרשימי פיזור: 6](#_Toc486414963)

[2.4 התאמת המשתנים: 7](#_Toc486414964)

[2.5 הגדרת משתני דמה: 8](#_Toc486414965)

[2.6 הוספת משתני אינטראקציה עבור משתני דמה: 9](#_Toc486414969)

[3. בניית המודל ובדיקת הנחות המודל: 10](#_Toc486414970)

[3.1 בחירת משתני המודל: 10](#_Toc486414971)

[3.2 בדיקת הנחות המודל: 10](#_Toc486414972)

[3.2.1 בדיקת הנחת הנורמליות של השגיאות: 10](#_Toc486414973)

[3.2.2 בדיקת הנחת שוויון שונויות: 11](#_Toc486414974)

[3.2.3 בדיקת הנחת הליניאריות: 12](#_Toc486414975)

[3.3 דוגמא לשימוש במודל: 12](#_Toc486414976)

[3.4 בדיקת השערות: 13](#_Toc486414977)

[3.4.1 מבחן F חלקי לבדיקה האם המודל הנבחר עדיף על המודל המלא: 13](#_Toc486414978)

[3.4.2 מבחן השערות על שונות המודל: 13](#_Toc486414980)

[3.4.3 מבחן השערות לבדיקת משתנה קטגוריאלי - עונות: 14](#_Toc486414981)

[4. שיפור המודל: 14](#_Toc486414982)

[5. מסקנות והמלצות: 16](#_Toc486414983)

[6. נספחים: 19](#_Toc486414984)

## 1. תקציר מנהלים:

## 1.1 המטרה:

בפרויקט זה מטרתנו הייתה לפתח את המודל הטוב ביותר (ע"פ הנלמד בכיתה) החוזה את האינפלציה בישראל.

## 1.2 המשתנה המוסבר:

הגדרנו את מדד המחירים לצרכן להיות המשתנה המוסבר. מדד המחירים לצרכן מפורסם מידי חודש ע"י הלשכה המרכזית לסטסטיסטיקה. מקובל לומר כי עלייה במדד המחירים לצרכן פירושה אינפלציה, וכי ירידה במדד משמעה דפלציה.

## 1.3 המשתנים המסבירים:

ישנם משתנים רבים המשפיעים על מדד המחירים לצרכן, בהיקף עבודה זו בחרנו להתחיל את פיתוח המודל עם שלושה עשר משתנים אשר שיערנו כי הם המשפיעים ביותר על האינפלציה בישראל.

* **משתנים רציפים –**תחזית אנליסטים ל-12 חודשים הבאים בישראל, ריבית בנק ישראל, שער הדולר, שער האירו, אוכלוסייה קבועה, עלייה ברוטו, שכר חודשי ממוצע, מדד המחירים לצרכן ב-OECD, תוצר לאומי גולמי, גרעון הממשלה.
* **משתנים קטגוריאליים** –חודש בשנה, כמות יחידות חבילות דואר מחו"ל ,שר האוצר.

## 1.4 מקורות הנתונים:

הנתונים נאספו מארבעה מקורות: בנק ישראל, הלמ"ס, OECD ודואר ישראל.

## 1.5 התצפיות:

המדגם הורכב מ-108 תצפיות כאשר כל תצפית היא חודש.

הנתונים מוינו על פי חודשים בין 2008 ל2016, כך שכל הנתונים רלוונטיים לסוף תקופה.

## 1.6 מהלך המחקר:

## התאמת משתנים:

* הערכנו ע"פ החומרים שהופקו בחלק א', מתאם Pearson ותרשימי הפיזור אילו משתנים כנראה נסיר מהמודל.
* איחדנו את קטגוריות 4-5 במשתנה שר האוצר ואת חודשי השנה לעונות.
* התאמנו את חבילות הדואר מחו"ל למשתנה קטגוריאלי.
* התאמנו משתני אינטראקציה {שער הדולר\*עונה}, {שער האירו\*עונה}, {מדד המחירים העולמי\*שר האוצר} ו{גרעון ממשלה\*שר האוצר}.

## 1.6.2 בחירת משתנים:

כדי לבחור את משתני המודל השתמשנו בשלושה אלגוריתמים לבדיקת מודלים (Forward, Backward, Stepwise). בחנו את תוצאות האלגוריתמים ע"פ מספר מדדים המספקים הערכה לטיב התאמה

(, , AIC, BIC). כמדדי החלטה לבחירת המודל הטוב ביותר, בחרנו במדדים BIC ו- .

האלגוריתמים Backward ו- Stepwise הניבו מודל זהה וטוב יותר מהמודל שהתקבל מאלגוריתם Forward.

המודל שהתקבל בשלב זה הינו:

j= 4.704+ 1.716Season2j+ 4.107Season3j+9.369Season4j+0.5439forcastInflation­­\_12j+ 0.9926dollarRatej + 0.7521Minister.of.Finance2 j + 0.3345Minister.of.Finance3 j -0.5684Minister.of.Finance4 j +0.23OECD\_CPI j -0.704unitMailPackets2 j  -0.8889unitMailPackets3 j -1.259unitMailPackets4 j -2.757unitMailPackets5 j -3.825unitMailPackets7 j +0.0001523GNP j -0.0000709 govermentDeficit j + 0.0002653govermentDeficit j\* Minister.of.Finance2 j+ 0.0002755govermentDeficit j\* Minister.of.Finance3 j + 0.0002467govermentDeficit j\* Minister.of.Finance4 j -0.3359dollarRate j\* Season2j -1.044dollarRate j\* Season3j -2.272 dollarRate j\* Season4j

## 1.6.3 בדיקת הנחות המודל:

בדקנו את הנחות המודל ומצאנו כי הנחת הנורמליות מתקיימת, אך הנחת הלינאריות והנחת שוויון שוניות לא מתקיימות במלואן. כדי לשפר את המודל ולטפל באי-לינאריות ובהטרו-סקדסטיות ביצענו טרנספורמציה לא לינארית של המשתנים המסבירים.

## 1.6.4 שיפור המודל:

בחלק של שיפור המודל מצאנו תחילה את המתאם של כל אחד מהמשתנים המסבירים עם המשתנה המוסבר, המשתנים בריבוע, שורש המשתנים ופונקציית ln. לאחר מכן, בחרנו בטרנספורמציית המשתנה בעל המתאם המקסימלי עבור כל משתנה. כמו כן שקלנו הוספת משתנים שהוסרו בשיטת הצעדים לאחור למתאם החדש של המשתנים המסבירים אל מול המשתנה המוסבר.

לאחר שיפור המודל ע"י טרנספורמציה לא לינארית של המשתנים המסבירים והמשתנה המוסבר ה- עלה מ-**0.9973** ל-**0.9975**. בעיית שוויון השוניות והאי-ליניאריות השתפרו מעט אך לא מספיק כדי לומר באופן חד משמעי שהנחת הליניאריות מתקיימת ושהמודל הומו-סקדסטי.

לבסוף קיבלנו את המודל:

j2=-36560+173.7Season2j+535Season3j+1436Season4j+4159forcastInflation­­\_12j-79.56forcastInflation­­\_12j+2 131.4dollarRate jּ ּּ96.53+Minister.of.Finance2 j + 35.06Minister.of.Finance3 j + 66.62Minister.of.Finance4 j + 4015log(OECD\_CPI j) -137.9unitMailPackets2 j  -168unitMailPackets3 j -236.9unitMailPackets4 j -498.6unitMailPackets5 j -694.3unitMailPackets7 j +2135log(GNP j )+ 6.275Avgwagej0.5 + 25.56Populationj0.5 -0.009236govermentDeficit j + 0.04398 govermentDeficit j\* Minister.of.Finance2 j+ 0.04295 govermentDeficit j\* Minister.of.Finance3 j + 0.0404 govermentDeficit j\* Minister.of.Finance4 j -25.62dollarRate j\* Season2j-129.76dollarRate j\* Season3j -338.1dollarRate j\* Season4j

## 2. עיבוד מקדים:

## 2.1 הגדרת המשתנים:

בחלק א' הגדרנו משתנה מוסבר אחד: Y – מדד המחירים לצרכן ו-13 משתנים מסבירים: X1 – X13. בטבלה שלהלן מוצגים משתנים אלה, תפקידם, סוג המשתנה ויחידות המידה וטווח הערכים שלו.

|  |  |  |  |
| --- | --- | --- | --- |
| שם משתנה | יחידת מידה | סוג (רציף, בידד, קטגוריאלי) | טווח ערכים |
| CPI | מדד בסיס=100 | רציף | {86.4,101.8} |
| monthInYear | חודשים | קטגוריאלי | {1,12} |
| forecastAVG\_12 | מדד בסיס=100 | רציף | {88.6,103.5} |
| interestBOI | אחוזים | רציף | {0.1,4.25} |
| dollarRate | ₪ לדולר | רציף | {3.3,4.2} |
| euroRate | ₪ ליורו | רציף | {4.0,5.7} |
| grossMigration | מספרים מוחלטים (נפש) | בדיד | {666,3900} |
| Population | אלפים (נפש) | רציף | {7290.4,8623.6} |
| AVGWage | ₪ | רציף | {7782,10189} |
| OECD\_CPI | מדד בסיס=100 | רציף | {95.8,111.5} |
| unitsMailPackets | שייכות לקטגוריה | קטגוריאלי | {1,7} |
| GNP | ₪ | רציף | {63178.17,104451.8} |
| governmentDeficit | מיליוני ₪ | רציף | {-6439.52,2374.3} |
| MinisterOfFinance | שייכות לקטגוריה | קטגוריאלי | {1,5} |

## 2.2 הסבר על המשתנים:

## 2.2.1 המשתנה המוסבר:

* CPI – מדד המחירים לצרכן: מדד המחירים לצרכן הוא מדד המהווה אחד מסוגי מדדי המחירים, המחושב בטכניקות סטטיסטיות ומשתנה מזמן לזמן, אשר מיועד לייצג את השתנות המחירים של סל מוצרים ושירותים הנצרכים על ידי קבוצת צרכנים.

## 2.2.2 המשתנים המסבירים:

* monthInYear – חודש בשנה: המדד מיצג את החודש בשנה (1-ינואר, 2-פברואר וכן הלאה).
* forecastAVG\_12– ממוצע תחזיות החזאים ל12 חודשים הקרובים למדד המחירים לצרכן.
* interestBOI – ריבית בנק ישראל: ריבית בנק ישראל היא כשמה הריבית שנקבעת על ידי הבנק המרכזי-בנק ישראל.
* dollarRate – שער הדולר:שער הדולר מהווה את היחס בין המטבע המקומי לדולר האמריקאי.
* euroRate – שער האירו: שער האירו מהווה את היחס בין המטבע המקומי לאירו.
* population – אוכלוסיה קבועה: מדד זה מייצג את ממוצע כמות התושבים לחודש נתון.
* grossMigration – עלייה ברוטו:מציין את ההגירה למדינת ישראל( ברוטו- לא כולל את האנשים שיורדים מהארץ).
* AVGWage - שכר חודשי ממוצע לשכיר במחירים שוטפים: מדד זה מציין את ממוצע השכר לשכירים בחודש נתון (נומינאלית).
* OECD\_CPI - מדד המחירים לצרכן ב-OECD:ארגון OECD הוא ארגון בינלאומי של המדינות המפותחות המקבלות את עקרונות הדמוקרטיה הליברלית והשוק החופשי. מדד המחירים לצרכן ב-OECD מהווה ממוצע משוקלל של מדד המחירים לצרכן של המדינות ששייכות לארגון.
* unitsMailPackets – כמות יחידות חבילות דואר מחו"ל:משתנה מסביר קטגוריאלי לפי טווחים- מודד את כמות היחידות של חבילות דואר מחו"ל באופן הבא: קטגוריה 1: בין מיליון למילון וחצי יחידות, קטגוריה 2: בין מיליון וחצי לשני מיליון יחידות וכן הלאה עד לקטגוריה 7: למעלה מ4 מיליון יחידות.
* GNP - תוצר לאומי גולמי:תוצר לאומי גולמי הוא מונח [כלכלי](https://he.wikipedia.org/wiki/%D7%9B%D7%9C%D7%9B%D7%9C%D7%94) המבטא את סך כל ה[מוצרים](https://he.wikipedia.org/wiki/%D7%9E%D7%95%D7%A6%D7%A8) (סחורות ושירותים) הסופיים המופקים בתקופת זמן מסוימת, נתונה במשק לאומי. הנתונים שאספנו מבנק ישראל היו לפי רבעונים, חילקנו כל רבעון לשלוש ופרסנו זאת על שלושת החודשים של הרבעון.
* governmentDeficit - גרעון הממשלה:גירעון תקציבי הוא מונח המציין כי התחייבויות או הוצאות, שנעשו בסעיפי [תקציב](https://he.wikipedia.org/wiki/%D7%AA%D7%A7%D7%A6%D7%99%D7%91) שונים עולים על הסכומים שהוקצאו להם באותו תקציב. הנתונים שאספנו מבנק ישראל היו לפי רבעונים, חילקנו כל רבעון לשלוש ופרסנו זאת על שלושת החודשים של הרבעון.
* MinisterOfFinance -שר האוצר:משתנה קטגוריאלי-קטגוריה 1:רוני בר און, קטגוריה 2:יובל שטייניץ, קטגוריה 3:יאיר לפיד, קטגוריה 4:משה כחלון וקטגוריה 5:אחר( קטגוריה זו מציינת את החודשים לאחר פיטורו של יאיר לפיד, בהם לא היה שר אוצר).

## 2.3 הסרה של משתנים:

## 2.3.1 הסרה ע"פ מדד Pearson :

כדי להחליט אילו משתנים נרצה להסיר, נרצה בתור התחלה לבדוק עבור אילו משתנים מסבירים רציפים, מקדם מתאם Pearson בינם לבין המשתנה המוסבר, קרוב לאפס.

חשוב לציין כי מדד זה אינו רלוונטי עבור המשתנים הקטגוריאליים, ולכן משתנים אלו יבדקו בהמשך לפי תרשימי הפיזור שלהם.

מקדם מתאם Pearson הינו מדד למתאם לינארי בין שני משתנים רציפים. המדד מעיד על הקשר שבין שני משתנים, כאשר ערכיו נעים בין מינוס אחד לאחד, וככל שהערך קרוב לאחד או מינוס אחד אז הקשר חזק יותר, ובהתאמה- ככל שהערך קרוב יותר ל-0 ניתן להסיק כי הקשר חלש יותר או לא קיים.

הקורלציה בין שני משתנים מקריים מוגדרת על ידי הנוסחה –

להלן תוצאות המדד:

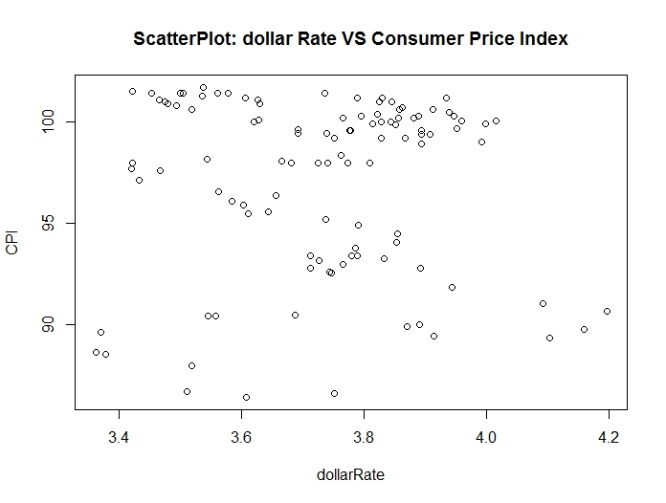
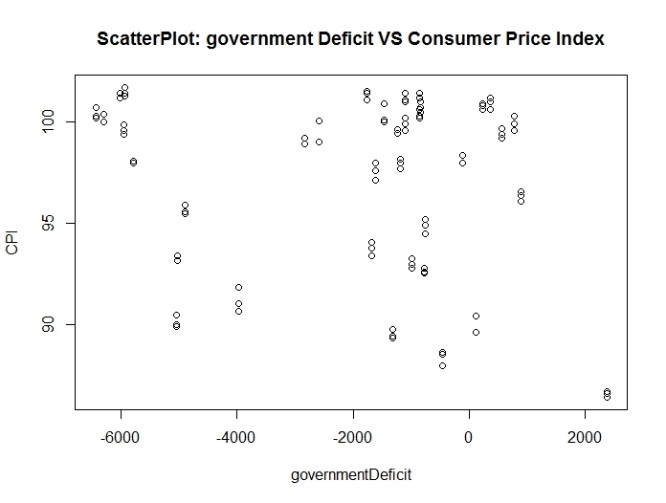
|  |  |
| --- | --- |
|  | Pearson Correlation with CPI |
| forecastInflation\_12 | 0.98495548 |
| interestBOI | -0.51294009 |
| dollarRate | -0.01850871 |
| euroRate | -0.7322672 |
| grossMigration | 0.51669982 |
| Population | 0.87347642 |
| AVGWage | 0.83900397 |
| OECD\_CPI | 0.92381172 |
| GNP | 0.88810916 |
| governmentDeficit | -0.18204891 |

לפי תוצאות המדד ניתן לראות כי עבור המשתנים שער הדולר וגרעון הממשלה ערך המדד בערך מוחלט נמוך מ0.2, ולכן נסמן אותם כמועמדים להסרה, ובשלב הבא נבחן את תרשימי הפיזור שלהם על מנת לוודא אם אכן לא קיים קשר לינארי בינם לבין המשתנה המוסבר.

## 2.3.2 הסרה ע"פ תרשימי פיזור:

מבחינה של כלל תרשימי הפיזור שהופקו בחלק א' של העבודה עולה כי ישנם שני משתנים שמידת ההסבר שלהם על המשתנה המוסבר-מדד המחירים לצרכן, עולה בספק. כפי שניתן לראות ניכר כי אין מתאם חזק בין גרעון הממשלה למדד המחירים לצרכן ועל כן נשקול להסירו. כמו כן ניכר כי המתאם בין שער הדולר למדד המחירים לצרכן אינו גבוה ועל כן נשקול להסירו. הנ"ל תואם את הבדיקה שעשינו ע"פ מתאם Pearson.

ההסקה לעיל הינה ראשונית, לא פורמאלית וויזואלית בלבד, כלומר ייתכן שלאחר בחינה סטטיסטית נבחר שלא להסיר את המשתנים- שער הדולר וגרעון ממשלה. (שאר תרשימי הפיזור מצורפים בנספחים).



## 2.4 התאמת המשתנים:

**א.** בחרנו לאחד את קטגוריית חודשי השנה – לפי עונות באופן הבא:

קטגוריה 1: חורף (ינואר-מרץ) .

קטגוריה 2: אביב (אפריל – יוני).

קטגוריה 3: קיץ (יולי- ספטמבר).

קטגוריה 4: סתיו (אוקטובר-דצמבר).

אנו סבורים שההשפעה של החודש על מדד המחירים לצרכן מקורה בעונתיות. להערכתנו השפעה החודש על מדד המחירים לצרכן מקורה בעיקר במזג האוויר, בחגים ובחופשים. לכן לדוגמא, לא נצפה הבדל בהשפעת יולי או אוגוסט על מדד המחירים לצרכן שכן שניהם חודשי קיץ ובשניהם יש חופש מהלימודים ועל כן השפעתם על המדד תהיה דומה מאוד. כמו כן ישנו חוסר התאמה בין לוח השנה העברי ללועזי ולכן החגים נחגגים בחודשים שונים על פני זמן. משיקולים הללו בחרנו לאחד את הקטגוריות חודשים לעונות השנה כמפורט לעיל.

**ב.** ייבוא חבילות דואר מחו"ל – מעבר ממשתנה רציף למשתנה קטגוריאלי.

הקטגוריות הן:

קטגוריה 1: 1 מיליון- 1.5 מיליון יחידות

קטגוריה 2: 1.5 מיליון – 2 מיליון יחידות

קטגוריה 3: 2 מיליון - 2.5 מיליון יחידות

קטגוריה 4: 2.5 מיליון – 3 מיליון יחידות

קטגוריה 5: 3 מיליון – 3.5 מיליון יחידות

קטגוריה 6: 3.5 מיליון עד 4 מיליון יחידות

קטגוריה 7: 4 מיליון יחידות ומעלה

בחרנו להתייחס למשתנה חבילות דואר מחו"ל כמשתנה קטגוריאלי ולא רציף משום שלהערכתנו עלייה של חבילת דואר בודדת משפיעה באופן זניח עד לא קיים על מדד המחירים לצרכן ,באופן מדגמי דרך התוצר הלאומי הגולמי והצריכה הפרטית. לעומת זאת עלייה של כחצי מיליון בחבילות הדואר (המשתנה הקטגוריאלי שלנו מקבל ערך שונה עבור כל שינוי של חצי מיליון חבילות דואר מחו"ל ) מצביעה על שינוי מגמתי בהרגלי הצריכה של הצרכן הישראלי, מה שלהערכתנו עשוי להשפיע על מדד המחירים לצרכן.

בחרנו לאחד את קטגוריות 4-5 במשתנה שר האוצר, הסיבה לכך היא מיעוט התצפיות של קטגוריה 5(מתחת ל2 אחוז מהמדגם-3 תצפיות), כלומר להערכתנו בגלל שמדובר במעט תצפיות אין ביכולתן להסביר את מדד המחירים לצרכן. קטגוריה 5 מציינת תקופה ללא שר האוצר, הבחירה לאחד עם קטגוריה 4 היא משום ששר מספר 4 הוא השר שכיהן מיד לאחד 3 החודשים ללא השר, באותו האופן יכולנו לאחד את הקטגוריה עם קטגוריה מס' 3-הקטגוריה לפניה הגיעו 3 החודשים ללא שר אוצר. ניתן לייחס את 3 החודשים הללו כתקופת חפיפתו של משה כחלון- קטגוריה מס' 4, כמו כן אין בידינו מידע מדויק על מה בדיוק כרגע מבחינת משרד האוצר בחודשים הללו, על כן האיחוד עם קטגוריה 4 הינו אקראי ובאותו האופן יכולנו לאחד את קטגוריות 5 ו3. חשוב לציין שהסיבה המרכזית לאיחוד הקטגוריות היא מיעוט התצפיות מקטגוריה 5.

## 2.5 הגדרת משתני דמה:

הגדרת משתני הדמה מגדירה מודל רגרסיה שונה עבור כל קבוצה- עם חותך שונה ולעתים שיפוע שונה (תלוי בהוספת משתני האינטראקציה). בדרך זו ניתן לבחון את ההשפעה השולית של כל קבוצה על המשתנה המוסבר. את המעבר ממשתנה קטגוריאלי למשתני דמה עשינו באמצעות פקודת factor (הקוד המלא מצורף בנספחים), התוכנה מגדירה k-1 משתני דמה עבור משתנה קטגוריאלי בעל k קטגוריות כפי שנלמד בכיתה ומגדירה באופן אוטומטי את הקטגוריה הראשונה כקבוצת הבסיס.

להלן הגדרות המשתנים הקטגוריאליים:

## משתנה דמה 1 – עונות- Season:

קטגוריה 1 תשמש כקבוצת הבסיס.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| סתיו | קיץ | אביב | חורף | משתנה דמה/ קטגוריה |
| 0 | 0 | 1 | 0 |  |
| 0 | 1 | 0 | 0 |  |
| 1 | 0 | 0 | 0 |  |

## משתנה דמה 2 – שר האוצר- MinisterOfFinance:

קטגוריה 1 תשמש כקבוצת הבסיס.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| משה כחלון | יאיר לפיד | יובל שטייניץ | רוני בר-און | משתנה דמי/ קטגוריה |
| 0 | 0 | 1 | 0 |  |
| 0 | 1 | 0 | 0 |  |
| 1 | 0 | 0 | 0 |  |

## משתנה דמה 3 – כמות יחידות חבילות דואר מחו"ל - unitsMailPackets:

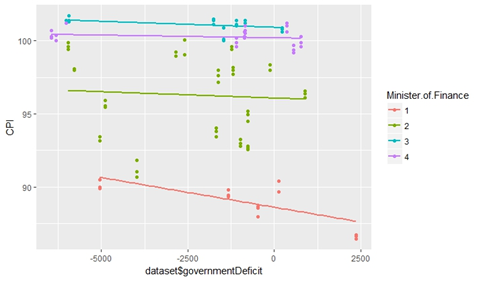
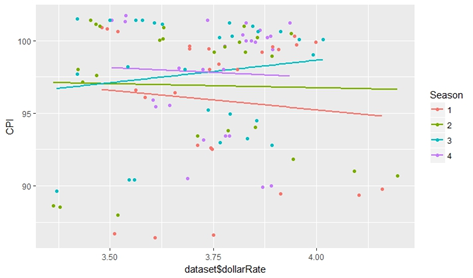
קטגוריה 1 תשמש כקבוצת הבסיס.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 4 מיליון יחידות ומעלה | 3.5 מיליון עד 4 מיליון יחידות | 3.5 מיליון עד 4 מיליון יחידות | 3 מיליון – 3.5 מיליון יחידות | 2 מיליון - 2.5 מיליון יחידות | 1.5 מיליון – 2 מיליון יחידות | 1 מיליון- 1.5 מיליון יחידות | משתנה דמי/ קטגוריה |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 |  |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 |  |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 |  |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 |  |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 |  |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 |  |

## 2.6 הוספת משתני אינטראקציה עבור משתני דמה:

בחרנו במשתני אינטראקציה {שער הדולר\*עונה} ו{שער האירו\*עונה} משום שלהערכתנו השפעתם של שער הדולר ושער האירו על מדד המחירים לצרכן משתנים מעונה לעונה. למשל בעונת הקיץ ישראלים רבים טסים לחוץ לארץ ולכן השימוש במטבע חוץ הוא רחב ובשל זאת להערכתנו השפעת שער החליפין (אירו או דולר) על מדד המחירים לצרכן בקיץ גבוהה יותר משאר עונות השנה.

בנוסף בחרנו במשתני אינטראקציה {מדד המחירים העולמי\*שר האוצר} ו{גרעון ממשלה\*שר האוצר} משום שלהערכתנו למדיניות שר האוצר קיימת השפעה על האופן שבו המשתנים הרציפים המתוארים לעיל משפיעים על מדד המחירים לצרכן בישראל. למשל עבור שר אוצר שנוקט במדיניות חוץ כלכלית רחבה נצפה שמדד המחירים בישראל יהיה דומה יחסית לזה העולמי (לעומת שר אוצר שנוקט במדיניות חוץ כלכלית צרה יותר).



ניתן לראות כי ישנם שיפועים שונים ללפחות שתי קטגוריות בגרפים לעיל (שאר הגרפים למשתני האינטראקציה שהצענו מצורפים בנספחים), כלומר-לדוגמא אם נבחן את הגרף הימני- ההשפעה השולית של שער הדולר על מדד המחירים לצרכן שונה בין עונה לעונה. בהתאם לכך עולה הצורך במשתני אינטראקציה מתאימים.

## 3. בניית המודל ובדיקת הנחות המודל:

## 3.1 בחירת משתני המודל:

על מנת לבחור את משתני המודל נבצע רגרסיה בצעדים לפי האלגוריתמים Forward Selection, Backward Elimination, Stepwise Regression. בכל איטרציה של כל אחד מהאלגוריתמים נסתמך על מבחן BIC, בחרנו במבחן זה, משום שבנוסף להתייחסות לפונקציית הנראות המדד "קונס" על פי גודל המדגם. לאחר מכן נבצע השוואה בין שלושת המודלים (המודלים שנבחרנו לפי האלגוריתמים המפורטים לעיל) לפי מדד .

הסיבה לבחירה במדד זה שהוא יכול לעלות או לרדת ככל שמספר המשתנים המסבירים עולה, בניגוד למדד שעולה ככל שמספר המשתנים המסבירים עולה.

מבחן BIC בוחן את טיב ההתאמה של המודל על סמך נראות המודל, ככל שנראות המודל עולה המדד קטן. כמו כן, המדד "קונס" את המודל על הוספת פרמטרים ועל גודל המדגם.

* ברגרסיה לפנים, נתחיל בבדיקת מודל ריק (ללא משתנים), ובכל איטרציה נוסיף למודל את המשתנה בעל ערך ה-BIC הנמוך ביותר, בתנאי שאינו גדול מערך המודל הנוכחי.
* ברגרסיה לאחור, נתחיל עם המודל המלא ובכל איטרציה נוריד את המשתנה שהסרתו תביא לערך ה-BIC הקטן ביותר, בתנאי שאינו גבוה מערך המודל הנוכחי.
* ברגרסיה דו-צדדית, נתחיל במודל המלא ונשלב את שתי השיטות שהוזכרו, כלומר בכל איטרציה נבחן הוספת משתנה או השמטת משתנה.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Forward Se-lection | Backward Elimination | Stepwise Regression |
| *BIC* | *231.79-* | -239.76 | -239.76 |
|  | 0.9957 | 0.9973 | 0.9973 |

המודלים שהתקבלו בשיטת המשולבת ובשיטת הצעדים לאחור הינם זהים, ערך ה- גבוה יותר מזה שנבחר באלגוריתם הForward ועל כן בחרנו במודל זה.

המודל הנבחר הוא:

CPIj β0+ β1Season2j+ β2Season3j+ β3Season4j +β4forcastInflation­­\_12j+ β5dollarRate j+ β6Minister.of.Finance2 j + β7Minister.of.Finance3 j + β8Minister.of.Finance4 j + β9OECD\_CPI j + β10unitMailPackets2 j  + β11unitMailPackets3 j + β12unitMailPackets4 j + β13unitMailPackets5 j + β14unitMailPackets7 j +β15GNP j + β16govermentDeficit j + β17 govermentDeficit j\* Minister.of.Finance2 j+ β18 govermentDeficit j\* Minister.of.Finance3 j + β19 govermentDeficit j\* Minister.of.Finance4 j +β20dollarRate j\* Season2j+ β21dollarRate j\* Season3j + β22dollarRate j\* Season4j +Ԑj

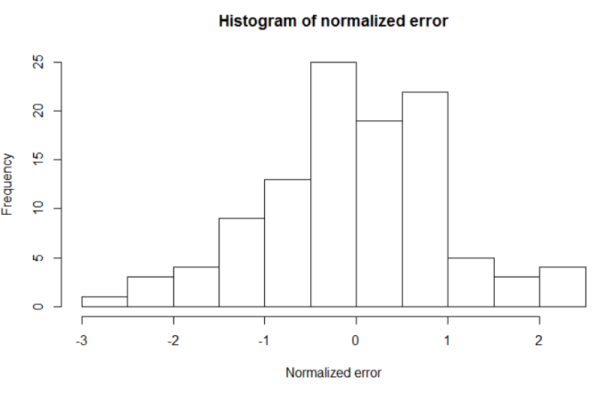
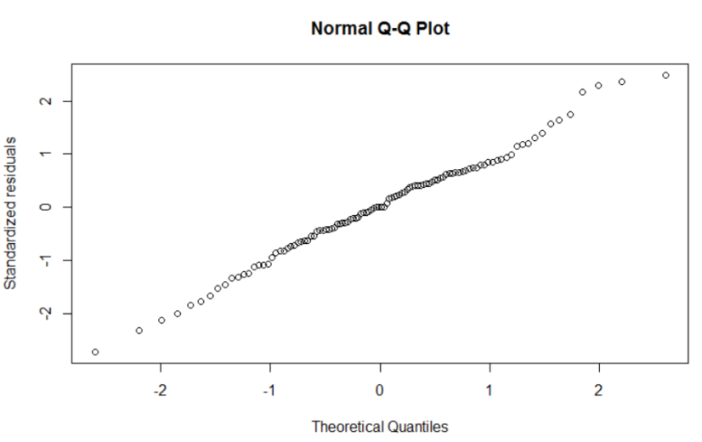
המודל הנאמד לפי ה-R:

j= 4.704+ 1.716Season2j+ 4.107Season3j+9.369Season4j+0.5439forcastInflation­­\_12j+ 0.9926dollarRatej + 0.7521Minister.of.Finance2 j + 0.3345Minister.of.Finance3 j -0.5684Minister.of.Finance4 j +0.23OECD\_CPI j -0.704unitMailPackets2 j  -0.8889unitMailPackets3 j -1.259unitMailPackets4 j -2.757unitMailPackets5 j -3.825unitMailPackets7 j +0.0001523GNP j -0.0000709 govermentDeficit j + 0.0002653govermentDeficit j\* Minister.of.Finance2 j+ 0.0002755govermentDeficit j\* Minister.of.Finance3 j + 0.0002467govermentDeficit j\* Minister.of.Finance4 j -0.3359dollarRate j\* Season2j -1.044dollarRate j\* Season3j -2.272 dollarRate j\* Season4j

## 3.2 בדיקת הנחות המודל:

## 3.2.1 בדיקת הנחת הנורמליות של השגיאות:

על מנת לבדוק הנחת הנורמליות של השגיאות נעשה שימוש בתרשים היסטוגרמה ותרשים Q-Q Plot על השגיאות המתוקננות.



מהתבוננות בתרשים ה- Q-Q Plot ניתן לראות כי רוב השגיאות ממוקמות על הציר הסימטרי בין X ל-Y, והתצפיות הקיצוניות אף הן קרובות לעקומה ודי סימטריות. בנוסף גם תרשים ההיסטוגרמה נראה סימטרי ולכן נראה כי הנחת הנורמליות מתקיימת. כדי לוודא שההנחה מתקיימת ביצענו את המבחנים הסטטיסטיים: Shapiro-Wilk ו-Kolmogorov-Smirnov .

מבחן קלמגורוב-סמירנוב:

One-sample Kolmogorov-Smirnov test

data: dataset$stan\_residuals

D = 0.053367, p-value = 0.9181

alternative hypothesis: two-sided

מבחן שפירו-ווילקס:

Shapiro-Wilk normality test

data: dataset$stan\_residuals

W = 0.99067, p-value = 0.6691

ניתן לראות כי על פי שני המבחנים ערכי הp-value שהתקבלו גדולים מ-0.05 ולכן לא נדחה את השערת האפס ונאמר כי שהשגיאות המתוקננות מתפלגות נורמלית ברמת מובהקות של 5% (השערת האפס במבחנים אלו היא שהנתונים מגיעים מהתפלגות נורמלית).

על מנת לתמוך בהחלטתנו כי הנחת הנורמליות מתקיימת נסתכל על ערכי הסטטיסטי שהתקבלו. ניתן לראות כי במבחן KS התקבל ערך נמוך של 0.053, דבר המצביע על מרחק קטן של התצפיות מההתפלגות הנורמלית. ובמבחן SW התקבל ערך גבוה של 0.99, דבר המצביע על מתאם גבוה בין הנתונים להתפלגות הנורמלית.

לכן, נחליט כי הנחת הנורמליות מתקיימת במודל.

## 3.2.2 בדיקת הנחת שוויון שונויות:

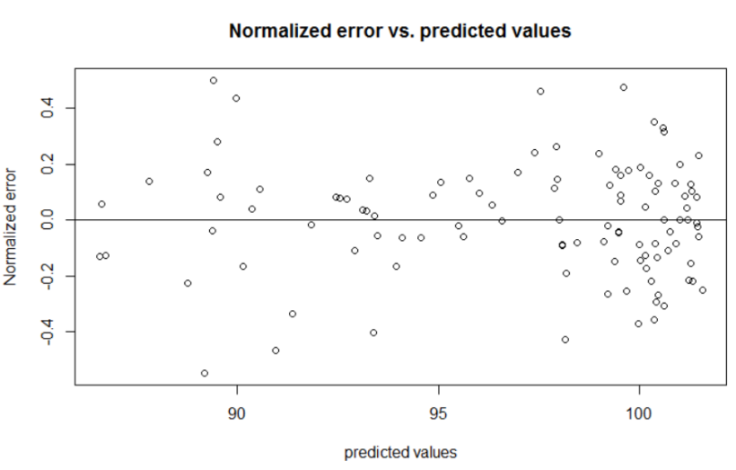
על מנת לבדוק את הנחת שוויון השונויות נעשה שימוש בתרשים פיזור השגיאות המתוקננות והאומד המוסבר.

Residuals:

Min 1Q Median 3Q Max

-0.54738 -0.12609 -0.00025 0.12834 0.49942

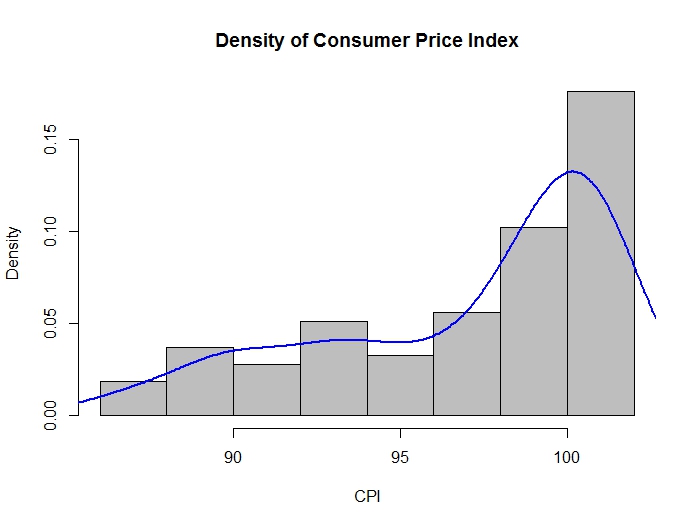
ע"פ פלט ה-R של התפלגות השגיאות נראה כי התפלגות השגיאות סימטרית סביב ה-0. החציון עומד קרוב לאפס, הערכים הקיצוניים (max ו-min) והערכים הרבעוניים סימטריים סביב האפס.



מהתבוננות בתרשים השגיאות מול האומד המוסבר נראה כי הפיזור אכן סימטרי אך לא אחיד סביב ה-0. ניתן לראות לפי התרשים כי מתקיימת הטרו-סקדסטיות במודל, כאשר שונות השגיאה מצטמצמת בין ערכים 92.5 ל-97.5 של האומד המוסבר. כלומר השונות של הטעויות איננה קבועה לכל התצפיות כי אם שונה, לכל תצפית יש את השונות שלה.

ניתן להסביר דינמיקה זו עקב שתי סיבות עיקריות:

* למדד המחירים לצרכן יש תלות סדרתית. כל תצפית תלויה בחודשים שלפנייה.
* מדד המחירים לצרכן אינו מתפלג נורמלית. קיימת אסימטריה שלילית גדולה המתבטאת בזנב שמאלי ארוך. ברוב המוחלט של המקרים, מדד המחירים לצרכן יהיה גדול מ100. ולכן לתצפיות הנ"ל תהיה השפעה מהותית על מודל הרגרסיה ולכן לא תצליח לחזות בהתאמה תצפיות חריגות (היכן שאין מספיק תצפיות) ולכן ניתן להניח שהשונות של השגיאה בכל רמה של מדד עלולה להיות שונה.



ולכן לא ניתן להסיק שהנחת שוויון השונויות מתקיימת.

## 3.2.3 בדיקת הנחת הליניאריות:

ע"פ תרשימי הפיזור שהופקו בחלק א' של הפרויקט רואים כי קיים קשר בין רוב מהמשתנים המסבירים שנבחרו במודל למשתנה המוסבר אך ככל הנראה איננו ליניארי.

בתרשים פיזור השגיאות המתוקננות והאומד המוסבר ניתן לראות כי ממוצע התצפיות לא בהכרח אפס לאורך העקומה (צד ימין) וכפי שצוין לעיל התצפיות לא מפוזרות בצורה אחידה על פני התרשים. לכן לא ניתן להסיק כי הנחת הליניאריות מתקיימת בתרשים הנ"ל.

## 3.3 דוגמא לשימוש במודל:

אנליסט בכיר מעוניין לבדוק מהי התוחלת הנחזית של מדד המחירים לצרכן בקיץ, כאשר גרעון הממשלה מאוזן (כלומר שווה לאפס), ייבוא חבילות הדואר נשאר במגמה חיובית עולה ועל כן שייך לקטגוריה 7 (ניתן לראות על פי הנתונים שכמות חבילות הדואר המיובאות מחולה עולה על פי זמן, ולכן הכלכלן מניח שאם המגמה ממשיכה גם בעתיד כמות חבילות הדואר המיובאות תמשיך ותהינה מעל 4 מיליון), התוצר הלאומי הגולמי מקסימאלי (ביחס למדגם), שר האוצר הוא עדיין משה כחלון, התחזיות לאינפלציה הן כחציון המדגם, שער הדולר הוא כממוצע המדגם ומדד המחירים העולמי מקסימאלי(ביחס למדגם).

Season0=3, govermentDeficit0=0, unitMailPackets0=7, Minister.of.Finance0=4

מחלק א' סעיף 5-

forcastInflation­­\_120=100.5710627, GNP0=104451.8, OECD\_CPI0=111.5, dollarRate0=3.733733222

j= 4.704+ 4.107Season3j+0.5439forcastInflation­­\_12j+ 0.9926dollarRatej  -0.5684Minister.of.Finance4j +0.23OECD\_CPI j -3.825unitMailPackets7 j +0.0001523GNP j -0.0000709 govermentDeficit j + 0.0002467govermentDeficit j\* Minister.of.Finance4 j -1.044dollarRate j\* Season3j

ולאחר הצבת הנתונים לעיל נקבל-

=100.4792963

## 3.4 בדיקת השערות:

## 3.4.1 מבחן F חלקי לבדיקה האם המודל הנבחר עדיף על המודל המלא:

נבצע מבחן F חלקי לבדוק האם המודל שבחרנו על פי מדדBIC ועל פי קריטריון אכן עדיף על המודל המלא, להלן ההשערות:

H0: modelbackward

H1: fullmodel

> anova(fullmudel, modelbackward)

Analysis of Variance Table

Model 1: CPI ~ Season + forecastInflation\_12 + interestBOI + dollarRate +

Season:dollarRate + euroRate:Season + euroRate + grossMigration +

population + AVGWage + OECD\_CPI + unitMailPackets + GNP +

governmentDeficit + Minister.of.Finance + Minister.of.Finance:governmentDeficit + Minister.of.Finance:OECD\_CPI

Model 2: CPI ~ Season + forecastInflation\_12 + dollarRate + OECD\_CPI +

unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

Season:dollarRate + governmentDeficit:Minister.of.Finance

Res.Df RSS Df Sum of Sq F Pr(>F)

1 74 3.6167

2 85 4.3274 -11 -0.71078 1.3221 0.2296

## ניתן לראות כי הPvalue>0.05 לכן לא ניתן לדחות את השערת האפס, נטען כי המודל שנבחר בשיטת הצעדים לאחור עדיף על המודל המלא ברמת ביטחון של 95%.

## 3.4.2 מבחן השערות על שונות המודל:

מבחן השערות על שונות המודל- האנליסט הבכיר מהסעיף הקודם רצה להוכיח שהמודל אכן מנבא את מדד המחירים לצרכן באחוז גבוה עם שונות לא גבוהה, זאת על מנת להתקבל לעבודה ב google. ראש מחלקת האנליסטים בגוגל טען ששונות המודל גבוהה מ-0.05 ושאם טעה אכן יקבל את האנליסט למשרה.

ההשערות הן:

H0:σ2<=0.05

H1: σ2>0.05

X2ST=sse/ σ2Ho=MSE\*n-p/0.05=86.522

X2CR=X285,0.95=107.5217

ניתן לראות שהסטטיסטי אינו גדול מהקריטי, לכן אין בידנו מספיק נתונים על מנת לדחות H0. לכן ראש מחלקת האנליסטים יקבל את האנליסט הבכיר למשרה.

## 3.4.3 מבחן השערות לבדיקת משתנה קטגוריאלי - עונות:

ברצוננו לבחון את חלוקת הקטגוריות שלנו, לצורך הדוגמא נבחן האם קיים הבדל מובהק בין מקדם הקיץ למקדם האביב תחת קטגוריה עונות(כלומר האם החותכים של הקיץ והאביב זהים). נחלץ את השונות המשותפת מ- R.

= 1.353165

H0: βseason-βspring=0

H1:not H0

Tst=-- (βseason-βspring)H0]\s.e(-)

=1.27179: s.e(-) =v(

לאחר הצבת האומדים מהפלט וסטיית התקן המחושבת לעיל נקבל את סטטיסטי המבחן.

Tst1.88 = Tcr= T85,0.975=+-1.66

ניתן לראות שהסטטיסטי גדול מהצד העליון של הקריטי, לכן נדחה את השערת האפס ונטען כי ברמת ביטחון של 95% החותכים של קטגוריות הקיץ והאביב אכן שונים.

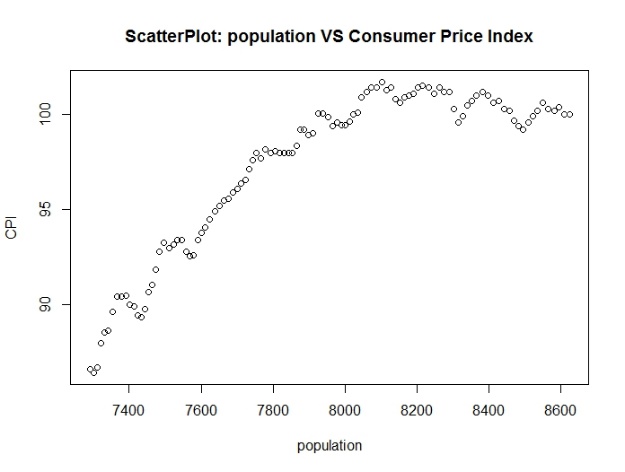
## 4. שיפור המודל:

לצורך שיפור המודל נבחן את מקדמי מתאם Pearson עבור המשתנה המוסבר על מול המשתנים הרציפים במודל שנבחר בשיטת הצעדים לאחור בכמה טרנספורמציות- ln,שורש והעלאה בריבוע.

\*הערה-לא בחנו את התהליך עבור המשתנה הרציף גרעון ממשלה מאחר והוא מקבל גם ערכים שליליים (לא מוגדרת פונקציית שורש ופונקציית ln).

נבחר את הטרנספורמציות בעלות מקדם המתאם הכי גבוה (בערך מוחלט). כמו כן נשקול הוספת משתנים שהוסרו בשיטת הצעדים לאחור באמצעות גרפים זוגיים של המשתנים המסבירים אל מול המשתנה המוסבר.

|  |  |
| --- | --- |
|  | Pearson correlation with cpi |
| forecastInflation\_12 | 0.98495548 |
| ln(forecastInflation\_12) | 0.98387195 |
| forecastInflation\_12^0.5 | 0.98445581 |
| forecastInflation\_12^2 | 0.98570342 |
| dollarRate | -0.01850871 |
| ln(dollarRate) | -0.01134585 |
| dollarRate^0.5 | -0.01490336 |
| dollarRate^2 | -0.02585915 |
| OECD\_CPI | 0.92381172 |
| ln(OECD\_CPI) | 0.92924186 |
| OECD\_CPI^0.5 | 0.92656048 |
| OECD\_CPI^2 | 0.91811672 |
| GNP | 0.88810916 |
| ln(GNP) | 0.91452938 |
| GNP^0.5 | 0.9018103 |
| GNP^2 | 0.85810542 |



ניתן לראות לדוגמא בגרף לעיל שהקשר בין האוכלוסייה למדד המחירים לצרכן אינו לינארי והפונקציה נראית כפונקציית שורש, על כן נשקול טרנספורמציית שורש. נבחן את מקדמי המתאם בשביל להחליט אם אכן לבצע את הטרנספורמציה הזו, ועוד נוספות עבור משתנים שהוסרו מהמודל.

\*הערה-נבחן רק טרנספורמציות שעולות כחשודות מהגרפים הזוגיים.

|  |  |
| --- | --- |
|  | Pearson correlation with cpi |
| AVGWage | 0.839004 |
| AVGWage^0.5 | 0.8440489 |
| Population | 0.8734764 |
| Population^0.5 | 0.8785355 |

על מנת לטפל בבעיית שיוויון השוניות ביצענו טרנספורמציית BOX-COX על המשתנה המוסבר. הלמדא המשוערת יצאה 2 (מצורף בנספחים) לכן הצגנו את ריבוע מדד המחירים לצרכן כמשתנה מוסבר.

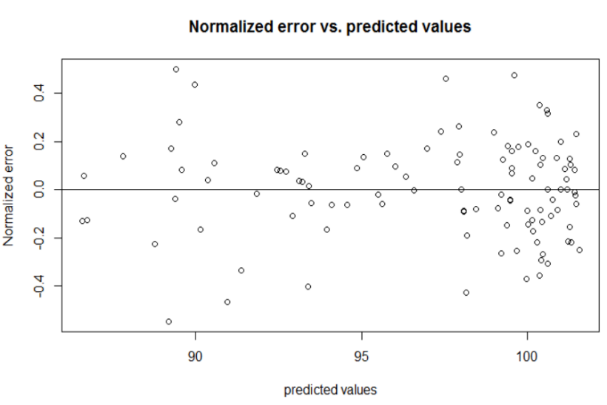
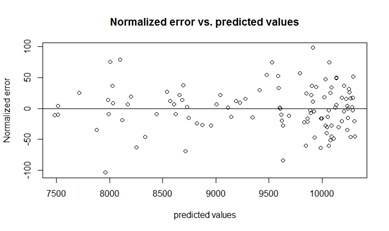
לאחר שינוי המודל באמצעות הטרנספורמציות על המשתנים הרציפים במודל שנבחר בשיטת הצעדים לאחור (מלבד הטרנספורמציה על שער הדולר שגורמת למולטי קולינאריות מלאה עם משתנה האינטרקציה שער הדולר\*עונה) והוספת המשתנים שהוסרו בטרנספורמציות שונות המודל הסופי הוא :

CPIj2=β0+ β1Season2j+ β2Season3j+ β3Season4j +β4forcastInflation­­\_12j+ β5forcastInflation­­\_12j+2 β6dollarRate j+ β7Minister.of.Finance2 j + β8Minister.of.Finance3 j + β9Minister.of.Finance4 j + β10 log(OECD\_CPI j)+ β11unitMailPackets2 j  + β12unitMailPackets3 j + β13unitMailPackets4 j + β14unitMailPackets5 j + β15unitMailPackets7 j +β16 log(GNP j )+ β17Avgwagej0.5 + β18Populationj0.5 + β19govermentDeficit j + β20 govermentDeficit j\* Minister.of.Finance2 j+ β21 govermentDeficit j\* Minister.of.Finance3 j + β22 govermentDeficit j\* Minister.of.Finance4 j +β23dollarRate j\* Season2j+ β24dollarRate j\* Season3j + β25dollarRate j\* Season4j +Ԑj

j2=-36560+173.7Season2j+535Season3j+1436Season4j+4159forcastInflation­­\_12j-79.56forcastInflation­­\_12j+2 131.4dollarRate jּ ּּ96.53+Minister.of.Finance2 j + 35.06Minister.of.Finance3 j + 66.62Minister.of.Finance4 j + 4015log(OECD\_CPI j) -137.9unitMailPackets2 j  -168unitMailPackets3 j -236.9unitMailPackets4 j -498.6unitMailPackets5 j -694.3unitMailPackets7 j +2135log(GNP j )+ 6.275Avgwagej0.5 + 25.56Populationj0.5 -0.009236govermentDeficit j + 0.04398 govermentDeficit j\* Minister.of.Finance2 j+ 0.04295 govermentDeficit j\* Minister.of.Finance3 j + 0.0404 govermentDeficit j\* Minister.of.Finance4 j -25.62dollarRate j\* Season2j-129.76dollarRate j\* Season3j -338.1dollarRate j\* Season4j

לאחר ביצוע הטרנספורמציות הנ"ל ה- עלה מ-0.9973 ל-0.9975.

בנוסף, בעיית שוויון השוניות השתפרה באופן זניח כפי שניתן לראות בגרפים(הפיזור סביב האפס מעט אקראי יותר בגרף הימני):



## 5. מסקנות והמלצות:

בפרויקט זה רצינו לבחון כיצד האינפלציה בישראל מושפעת משלושה עשר גורמים, אותם איתרנו בחלק א כגורמים שיתכן ומשפיעים על האינפלציה.

בחלק זה של הפרויקט ראשית בחנו אפשרות להסרת משתנים רציפים אשר לא נמצא קשר חזק בינם לבין המשתנה המוסבר, לפי מדד Pearson ותרשימי הפיזור. כמו כן, ביצענו התאמות למשתנים במודל , ויצרנו משתני דמה מתאימים למשתנים הקטגוריאליים ומשתני אינטראקציה מתאימים.

בחנו מספר מודלים על ידי האלגוריתמים שנלמדו בכיתה: Forward, Backward ו- Stepwise Regression. כמדדי החלטה לבחירת המודל הטוב ביותר, בחרנו במדדים BIC ו- .

האלגוריתמים Backward ו- Stepwise הניבו מודל זהה וטוב יותר מהמודל שהתקבל מאלגוריתם Forward.

המודל שהתקבל הינו:

j= 4.704+ 1.716Season2j+ 4.107Season3j+9.369Season4j+0.5439forcastInflation­­\_12j+ 0.9926dollarRatej + 0.7521Minister.of.Finance2 j + 0.3345Minister.of.Finance3 j -0.5684Minister.of.Finance4 j +0.23OECD\_CPI j -0.704unitMailPackets2 j  -0.8889unitMailPackets3 j -1.259unitMailPackets4 j -2.757unitMailPackets5 j -3.825unitMailPackets7 j +0.0001523GNP j -0.0000709 govermentDeficit j + 0.0002653govermentDeficit j\* Minister.of.Finance2 j+ 0.0002755govermentDeficit j\* Minister.of.Finance3 j + 0.0002467govermentDeficit j\* Minister.of.Finance4 j -0.3359dollarRate j\* Season2j -1.044dollarRate j\* Season3j -2.272 dollarRate j\* Season4j

בנוסף, היה עלינו לבחון את שלושת הנחות הרגרסיה הליניארית על מודל זה. נמצא כי אף על פי שהמודל עונה על הנחת הנורמליות ומתאר את הנתונים בצורה טובה (ה- עומד על 0.9973), נראה כי המודל סובל מהטרו-סקדסטיות ואינו ליניארי. הצלחנו להעלות שתי השערות בנוגע למקור ההטרו-סקדסטיות ומצאנו כי זה עלול לנבוע מ:

* למדד המחירים לצרכן יש תלות סדרתית. כל תצפית תלויה בחודשים שלפנייה.
* מדד המחירים לצרכן אינו מתפלג נורמלית. קיימת אסימטריה שלילית גדולה המתבטאת בזנב שמאלי ארוך. ברוב המוחלט של המקרים, מדד המחירים לצרכן יהיה גדול מ100. ולכן לתצפיות הנ"ל תהיה השפעה מהותית על מודל הרגרסיה ולכן לא תצליח לחזות בהתאמה תצפיות חריגות (היכן שאין מספיק תצפיות) ולכן ניתן להניח שהשונות של השגיאה בכל רמה של מדד עלולה להיות שונה.

לבסוף, בדקנו האם ביצוע טרנספורמציות על המשתנים יביא לשיפור המודל. ניסינו לפתור ללא הצלחה יתרה את האי-לינאריות ובעיית ההטרו-סקדסטיות הנ"ל על ידי טרנספורמציות שונות על משתנים המסבירים (שורש, Ln,העלאה בריבוע).

לאחר שיפור המודל ע"י טרנספורמציה לא לינארית של המשתנים המסבירים והמשתנה המוסבר ה- עלה מ-**0.9973** ל-**0.9975**. בעיית שוויון השוניות והאי-ליניאריות השתפרו מעט אך לא מספיק כדי לומר באופן חד משמעי שהנחת הליניאריות מתקיימת ושהמודל הומו-סקדסטי.

לבסוף קיבלנו את המודל:

j2=-36560+173.7Season2j+535Season3j+1436Season4j+4159forcastInflation­­\_12j-79.56forcastInflation­­\_12j+2 131.4dollarRate jּ ּּ96.53+Minister.of.Finance2 j + 35.06Minister.of.Finance3 j + 66.62Minister.of.Finance4 j + 4015log(OECD\_CPI j) -137.9unitMailPackets2 j  -168unitMailPackets3 j -236.9unitMailPackets4 j -498.6unitMailPackets5 j -694.3unitMailPackets7 j +2135log(GNP j )+ 6.275Avgwagej0.5 + 25.56Populationj0.5 -0.009236govermentDeficit j + 0.04398 govermentDeficit j\* Minister.of.Finance2 j+ 0.04295 govermentDeficit j\* Minister.of.Finance3 j + 0.0404 govermentDeficit j\* Minister.of.Finance4 j -25.62dollarRate j\* Season2j-129.76dollarRate j\* Season3j -338.1dollarRate j\* Season4j

עקב כך ובנוסף לסיבות המפורטות לעיל, אנו ממליצים:

1. כדי להתגבר על התלות סדרתית של מדד המחירים לצרכן בזמן, יש לדגום באופן אקראי ומטווחי זמן שונים.
2. הוספה של משתנים מסבירים לא טריוויאליים נוספים כגון מדד ביטחון הצרכנים, אירועים ביטחוניים, אירועים כלכליים משמעותיים, מדידה של השפעת המחאה החברתית, שינוי טעמים של הצרכנים.
3. ביצוע טרנספורמציות נוספות.
4. בחינת בעיית מולטיקולניאריות ופתרונה כפי שנלמד בכיתה.

**תוכן עניינים - נספחים**

[6. נספחים: 20](#_Toc486414984)

[6.1 מקורות הנתונים: 20](#_Toc486414985)

[6.2 נספח – התאמת משתנים: 20](#_Toc486414986)

[6.2.1 נספח – התאמת משתנים – מדד Pearson: 20](#_Toc486414987)

[6.2.2 נספח – התאמת משתנים – מדדי פיזור: 20](#_Toc486414988)

[6.2.3 נספח – התאמת משתנים - עונות: 22](#_Toc486414989)

[6.2.4 נספח – התאמת משתנים – כמות חבילות דואר מחו"ל: 22](#_Toc486414990)

[6.2.5 נספח – התאמת משתנים – שר האוצר: 22](#_Toc486414991)

[6.2.6 נספח – משתנה דמה: 22](#_Toc486414992)

[6.2.7 נספח – בחינה של משתני אינטראקציה: 22](#_Toc486414993)

[6.3 נספח – הסרת משתנים: 24](#_Toc486414994)

[6.3.1 נספח – הסרת משתנים – המודל הריק: 24](#_Toc486414995)

[6.3.2 נספח – הסרת משתנים – המודל המלא: 24](#_Toc486414996)

[6.3.3 נספח – הסרת משתנים – אלגוריתם backward: 24](#_Toc486414997)

[6.3.4 נספח – הסרת משתנים – המודל שהתקבל ע"י backward: 27](#_Toc486414998)

[6.3.5 נספח – הסרת משתנים – summary של המודל שהתקבל ע"י backward: 28](#_Toc486414999)

[6.3.6 נספח – הסרת משתנים – אלגוריתם forward: 29](#_Toc486415000)

[6.3.7 נספח – הסרת משתנים – המודל שהתקבל ע"י אלגוריתם forward: 30](#_Toc486415001)

[6.3.8 נספח – הסרת משתנים – summary של המודל שהתקבל ע"י אלגוריתם forward: 31](#_Toc486415002)

[6.3.9 נספח – הסרת משתנים –אלגוריתם Stepwise: 31](#_Toc486415003)

[6.3.10 נספח – הסרת משתנים –המודל שהתקבל ע"י אלגוריתם Stepwise: 35](#_Toc486415004)

[6.3.11 נספח – הסרת משתנים – summary של המודל שהתקבל ע"י אלגוריתם Stepwise: 36](#_Toc486415005)

[6.3.12 נספח – בדיקת הנחות המודל – פיזור השגיאות המתוקננות והאומד המוסבר: 37](#_Toc486415006)

[6.3.13 נספח – בדיקת הנחות המודל – פיזור השגיאות המתוקננות והאומד המוסבר: 37](#_Toc486415007)

[6.3.14 נספח – בדיקת הנחות המודל – היסטוגרמה: 37](#_Toc486415008)

[6.3.15 נספח – בדיקת הנחות המודל – Q-Q Plot: 37](#_Toc486415009)

[6.3.16 נספח – בדיקת הנחות המודל – Kolmogorov-Smirnov: 37](#_Toc486415010)

[6.3.17 נספח – בדיקת הנחות המודל – Shapiro-Wilk: 37](#_Toc486415011)

[6.4 נספח – בדיקת השערות : 37](#_Toc486415012)

[6.4.1 נספח – מבחן F חלקי לבדיקה האם המודל הנבחר עדיף על המודל המלא : 37](#_Toc486415013)

[6.4.2 נספח – מטריצת שוניות משותפת של האומדים למקדמי הרגרסיה בדיקת משתנה - עונות: 38](#_Toc486415014)

[6.5 נספח – שיפור המודל - טרנספורמציה לא לינארית של המשתנים המסבירים: 38](#_Toc486415015)

[6.5.1 נספח – שיפור המודל – מציאת מטריצת שוניות משותפת: 38](#_Toc486415016)

[6.5.2 נספח – שיפור המודל – טרנספורמציית בוקס קוקס: 38](#_Toc486415017)

[6.5.3 נספח – שיפור המודל –המודל שהתקבל לאחר טרנספורמציה לא לינארית: 38](#_Toc486415018)

[6.5.4 נספח – שיפור המודל – summary על המודל המשופר: 38](#_Toc486415019)

[6.5.5 נספח – שיפור המודל – פיזור השגיאות המתוקננות והאומד המוסבר: 39](#_Toc486415020)

## 6. נספחים:

## 6.1 מקורות הנתונים:

הנתונים נאספו מארבעה מקורות:

* בנק ישראל

<http://www.boi.org.il/he/DataAndStatistics/Pages/AllSubjects.aspx>

* הלמ"ס

<http://www.cbs.gov.il/reader>

* OECD

<https://data.oecd.org/>

* דואר ישראל.

(פורסם ב-ynet, קיבלנו אישור מהלל להשתמש בנתונים)

<http://www.ynet.co.il/articles/0,7340,L-4901799,00.html>

## 6.2 נספח – התאמת משתנים:

## 6.2.1 נספח – התאמת משתנים – מדד Pearson:

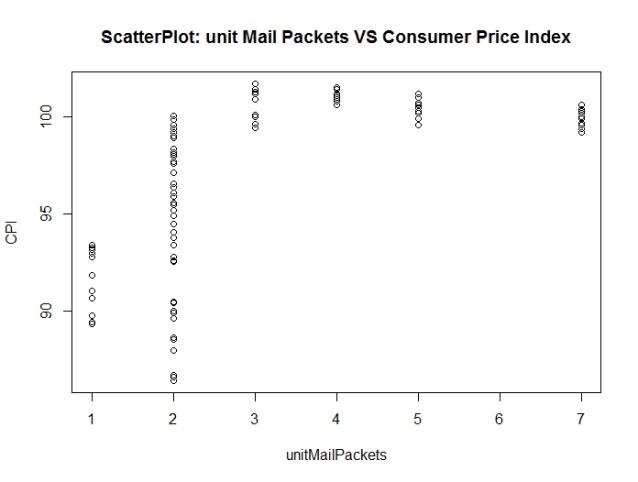
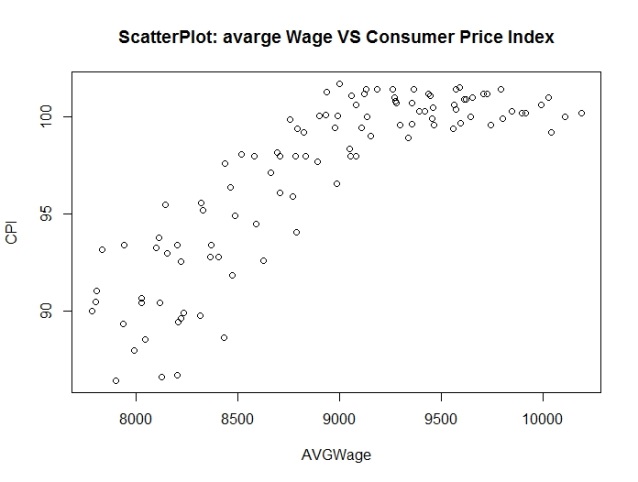
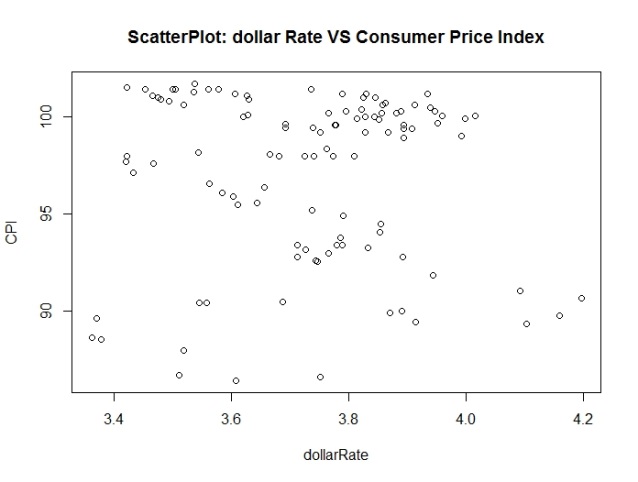
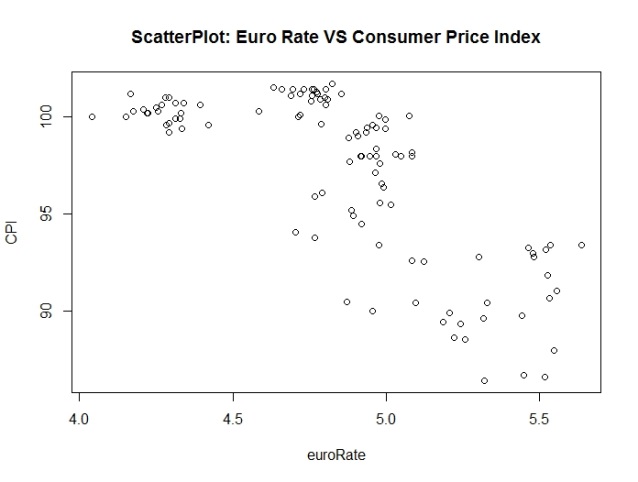
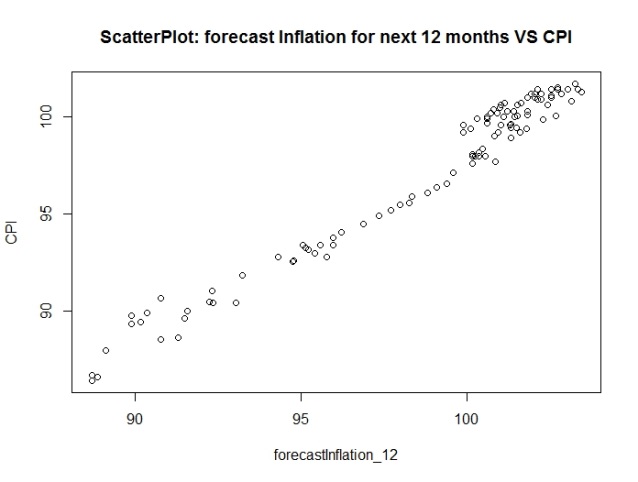
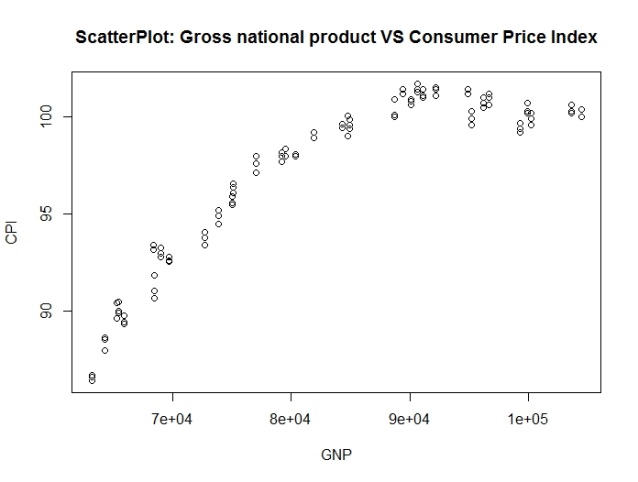
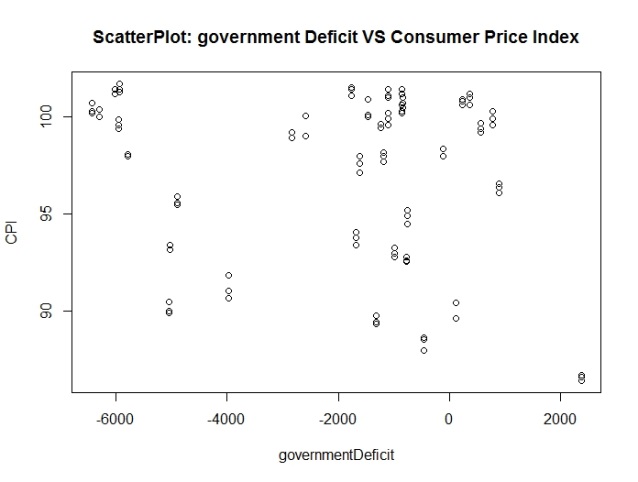
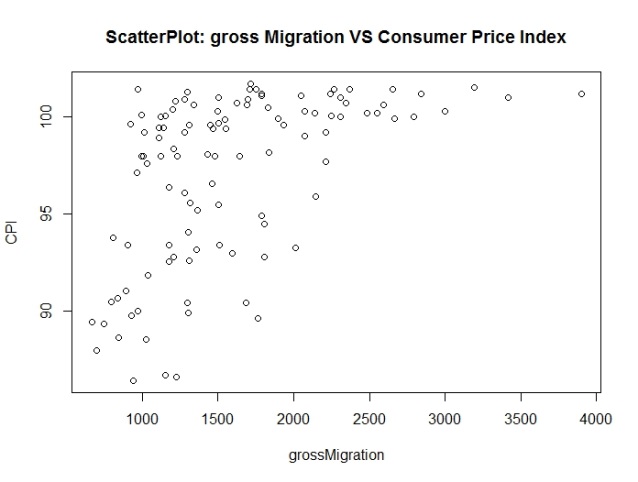
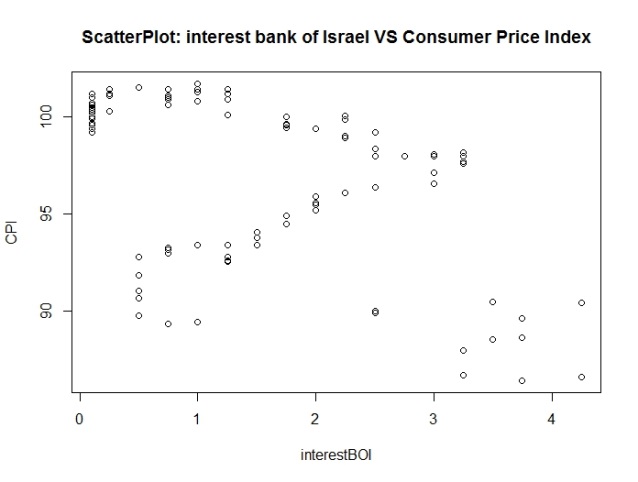
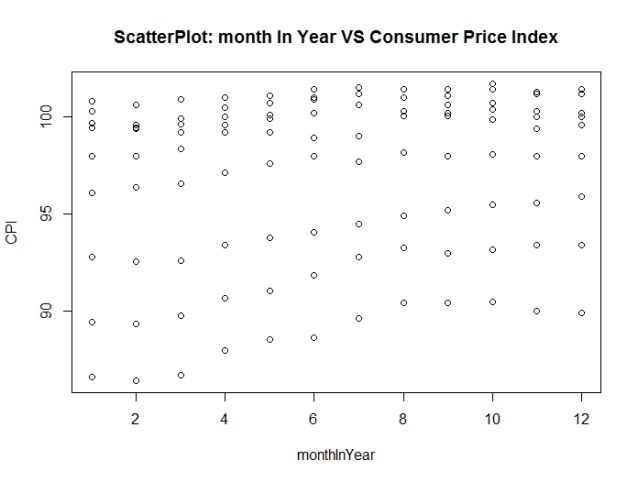
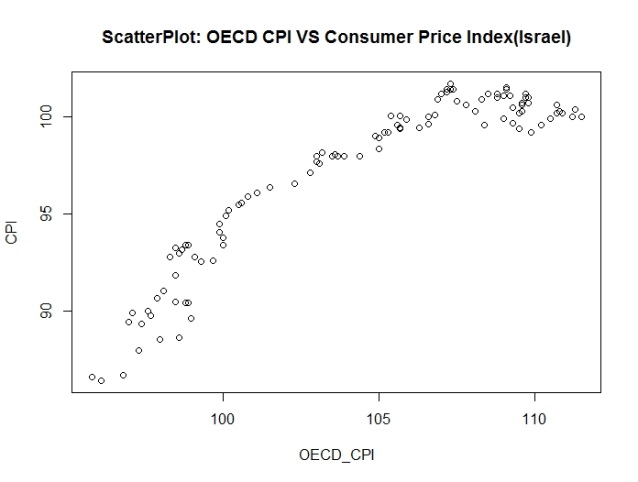
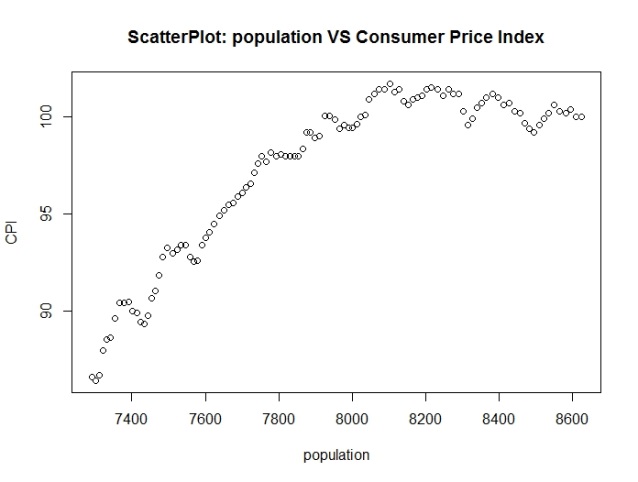
> dataset<-read.csv(file.choose(),header = T)

> attach(dataset)

> d1<-data.frame(CPI,forecastInflation\_12, interestBOI,dollarRate,euroRate, grossMigration, population,AVGWage,OECD\_CPI, GNP, governmentDeficit)

> cor(d1)

## 6.2.2 נספח – התאמת משתנים – מדדי פיזור:



## 6.2.3 נספח – התאמת משתנים - עונות:

dataset<-read.csv(file.choose(),header = T)

dataset$Season <- as.character(dataset$Season)

dataset$Season[dataset$Season == "Winter"] <- "1"

dataset$Season[dataset$Season == "Spring"] <- "2"

dataset$Season[dataset$Season == "Summer"] <- "3"

dataset$Season[dataset$Season == "Fall"] <- "4"

dataset$Season <- as.numeric(dataset$Season)

## 6.2.4 נספח – התאמת משתנים – כמות חבילות דואר מחו"ל:

dataset$unitMailPackets <- as.character(dataset$unitMailPackets)

dataset$unitMailPackets[dataset$unitMailPackets == "1-1.5"] <- "1"

dataset$unitMailPackets[dataset$unitMailPackets == "1.5-2"] <- "2"

dataset$unitMailPackets[dataset$unitMailPackets == "2-2.5"] <- "3"

dataset$unitMailPackets[dataset$unitMailPackets == "2.5-3"] <- "4"

dataset$unitMailPackets[dataset$unitMailPackets == "3-3.5"] <- "5"

dataset$unitMailPackets[dataset$unitMailPackets == "3.5-4"] <- "6"

dataset$unitMailPackets[dataset$unitMailPackets == "4+"] <- "7"

dataset$unitMailPackets <- as.numeric(dataset$unitMailPackets)

## 6.2.5 נספח – התאמת משתנים – שר האוצר:

dataset$Minister.of.Finance <- as.character(dataset$Minister.of.Finance)

dataset$Minister.of.Finance[dataset$Minister.of.Finance == "Bar-On"] <- "1"

dataset$Minister.of.Finance[dataset$Minister.of.Finance == "Steinitz"] <- "2"

dataset$Minister.of.Finance[dataset$Minister.of.Finance == "Lapid"] <- "3"

dataset$Minister.of.Finance[dataset$Minister.of.Finance == "Kachlon"] <- "4"

dataset$Minister.of.Finance[dataset$Minister.of.Finance == "nobody"] <- "4"

dataset$Minister.of.Finance <- as.numeric(dataset$Minister.of.Finance)

## 6.2.6 נספח – משתנה דמה:

dataset$Season <- factor(dataset$Season)

dataset$Minister.of.Finance <- factor(dataset$Minister.of.Finance)

dataset$unitMailPackets <- factor(dataset$unitMailPackets)

## 6.2.7 נספח – בחינה של משתני אינטראקציה:

> library(ggplot2)

> (p1 <- ggplot(data = dataset, aes(x =dataset$dollarRate , y = CPI, colour = Season)) +

+ geom\_smooth(

+ data = dataset, aes(x = dataset$dollarRate, y = CPI, fill=Season),

+ method = "lm", se = FALSE

+ ) + geom\_point())

> (p1 <- ggplot(data = dataset, aes(x = dataset$euroRate , y = CPI, colour = Season)) +

+ geom\_smooth(

+ data = dataset, aes(x = dataset$euroRate, y = CPI, fill=Season),

+ method = "lm", se = FALSE

+ ) + geom\_point())

> (p1 <- ggplot(data = dataset, aes(x = dataset$governmentDeficit , y = CPI, colour = Minister.of.Finance)) +

+ geom\_smooth(

+ data = dataset, aes(x = dataset$governmentDeficit, y = CPI, fill=Minister.of.Finance),

+ method = "lm", se = FALSE

+ ) + geom\_point())

> (p1 <- ggplot(data = dataset, aes(x = dataset$governmentDeficit , y = CPI, colour = Minister.of.Finance)) +

+ geom\_smooth(

+ data = dataset, aes(x = dataset$governmentDeficit, y = CPI, fill=Minister.of.Finance),

+ method = "lm", se = FALSE

+ ) + geom\_point())

> (p1 <- ggplot(data = dataset, aes(x = dataset$governmentDeficit , y = CPI, colour = Minister.of.Finance)) +

+ geom\_smooth(

+ data = dataset, aes(x = dataset$governmentDeficit, y = CPI, fill=Minister.of.Finance),

+ method = "lm", se = FALSE

+ ) + geom\_point())

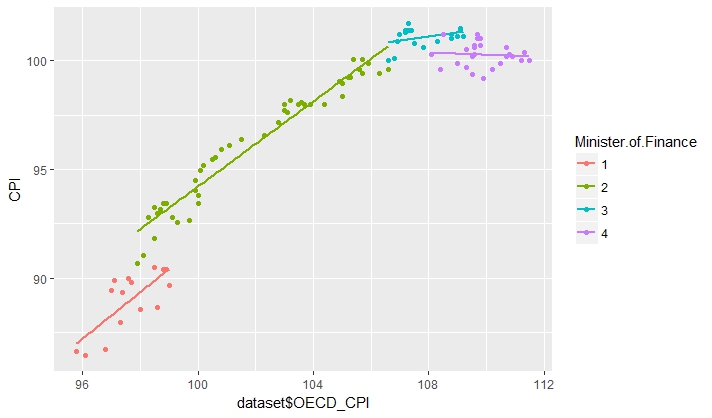
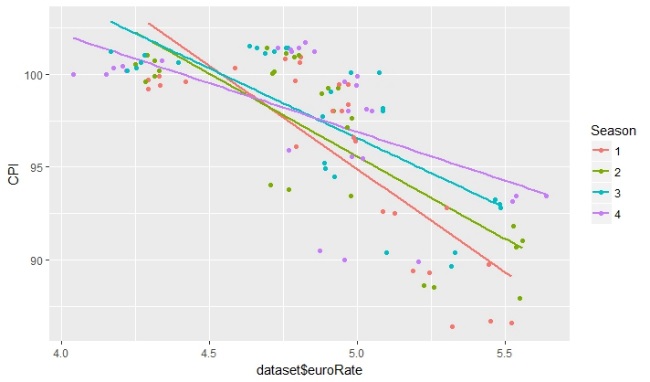
> (p1 <- ggplot(data = dataset, aes(x = dataset$OECD\_CPI , y = CPI, colour = Minister.of.Finance)) +

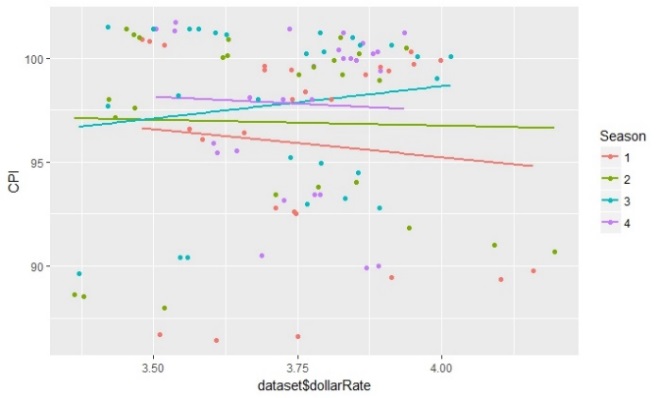
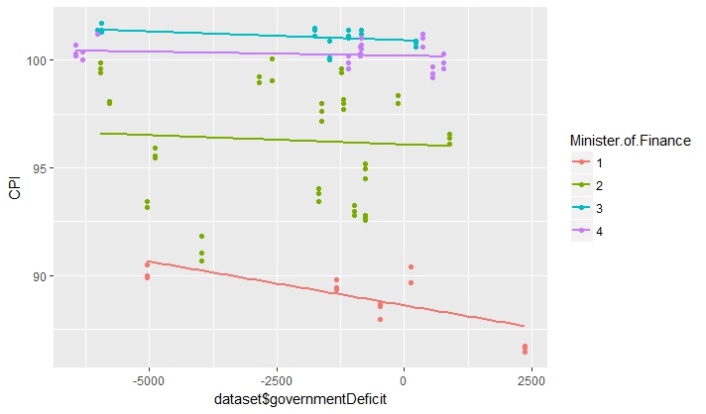
+ geom\_smooth(

+ data = dataset, aes(x = dataset$OECD\_CPI, y = CPI, fill=Minister.of.Finance),

+ method = "lm", se = FALSE

+ ) + geom\_point())





## 6.3 נספח – הסרת משתנים:

## 6.3.1 נספח – הסרת משתנים – המודל הריק:

modelnull<-lm(CPI~1)

## 6.3.2 נספח – הסרת משתנים – המודל המלא:

fullmudel<-lm(CPI~Season+forecastInflation\_12+interestBOI+dollarRate+Season:dollarRate+euroRate:Season+euroRate+grossMigration+population+AVGWage+OECD\_CPI+unitMailPackets+GNP+governmentDeficit+Minister.of.Finance+Minister.of.Finance:governmentDeficit+Minister.of.Finance:OECD\_CPI)

## 6.3.3 נספח – הסרת משתנים – אלגוריתם backward:

backBIC <- step(fullmudel, direction = "backward", test = "F", k = log(108))

Start: AIC=-207.64

CPI ~ Season + forecastInflation\_12 + interestBOI + dollarRate +

Season:dollarRate + euroRate:Season + euroRate + grossMigration +

population + AVGWage + OECD\_CPI + unitMailPackets + GNP +

governmentDeficit + Minister.of.Finance + Minister.of.Finance:governmentDeficit +

Minister.of.Finance:OECD\_CPI

Df Sum of Sq RSS AIC F value Pr(>F)

- OECD\_CPI:Minister.of.Finance 3 0.0198 3.6365 -221.09 0.1353 0.9386853

- Season:euroRate 3 0.2978 3.9145 -213.14 2.0313 0.1168310

- AVGWage 1 0.0271 3.6437 -211.51 0.5538 0.4591121

- interestBOI 1 0.0444 3.6611 -211.00 0.9094 0.3433813

- grossMigration 1 0.0876 3.7043 -209.74 1.7926 0.1847038

- population 1 0.1120 3.7287 -209.03 2.2926 0.1342549

<none> 3.6167 -207.64

- Season:dollarRate 3 0.5625 4.1792 -206.07 3.8366 0.0130438 \*

- GNP 1 0.4157 4.0323 -200.57 8.5047 0.0046861 \*\*

- unitMailPackets 5 1.2393 4.8560 -199.22 5.0716 0.0004728 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.0717 5.6884 -172.77 14.1296 2.26e-07 \*\*\*

- forecastInflation\_12 1 3.5110 7.1277 -139.05 71.8386 1.63e-12 \*\*\*

---

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

Step: AIC=-221.09

CPI ~ Season + forecastInflation\_12 + interestBOI + dollarRate +

euroRate + grossMigration + population + AVGWage + OECD\_CPI +

unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

Season:dollarRate + Season:euroRate + governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

- Season:euroRate 3 0.3404 3.9769 -225.48 2.4025 0.074034 .

- AVGWage 1 0.0309 3.6674 -224.86 0.6533 0.421438

- interestBOI 1 0.0359 3.6724 -224.71 0.7609 0.385755

- grossMigration 1 0.0773 3.7138 -223.50 1.6362 0.204695

- population 1 0.0939 3.7304 -223.02 1.9883 0.162547

- OECD\_CPI 1 0.0994 3.7359 -222.86 2.1052 0.150858

<none> 3.6365 -221.09

- Season:dollarRate 3 0.6305 4.2670 -217.87 4.4502 0.006159 \*\*

- GNP 1 0.4722 4.1087 -212.59 9.9993 0.002241 \*\*

- unitMailPackets 5 1.8784 5.5149 -199.53 7.9549 4.388e-06 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.1200 5.7565 -185.53 14.9632 9.158e-08 \*\*\*

- forecastInflation\_12 1 4.1140 7.7505 -144.05 87.1106 2.737e-14 \*\*\*

---

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

Step: AIC=-225.48

CPI ~ Season + forecastInflation\_12 + interestBOI + dollarRate +

euroRate + grossMigration + population + AVGWage + OECD\_CPI +

unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

Season:dollarRate + governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

- interestBOI 1 0.0087 3.9856 -229.92 0.1751 0.676778

- grossMigration 1 0.0472 4.0241 -228.88 0.9496 0.332749

- AVGWage 1 0.0575 4.0344 -228.61 1.1570 0.285312

- population 1 0.0724 4.0493 -228.21 1.4564 0.231062

- euroRate 1 0.1144 4.0913 -227.09 2.3022 0.133134

- OECD\_CPI 1 0.1747 4.1516 -225.52 3.5141 0.064500 .

<none> 3.9769 -225.48

- Season:dollarRate 3 0.7887 4.7656 -219.98 5.2885 0.002232 \*\*

- GNP 1 0.5087 4.4856 -217.16 10.2328 0.001978 \*\*

- unitMailPackets 5 1.8476 5.8245 -207.68 7.4333 9.050e-06 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.6746 6.6515 -183.97 17.9343 5.399e-09 \*\*\*

- forecastInflation\_12 1 6.2255 10.2024 -128.41 125.2341 < 2.2e-16 \*\*\*

---

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

Step: AIC=-229.92

CPI ~ Season + forecastInflation\_12 + dollarRate + euroRate +

grossMigration + population + AVGWage + OECD\_CPI + unitMailPackets +

GNP + governmentDeficit + Minister.of.Finance + Season:dollarRate +

governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

- grossMigration 1 0.0449 4.0305 -233.40 0.9120 0.3424235

- AVGWage 1 0.0633 4.0489 -232.90 1.2872 0.2599150

- population 1 0.0662 4.0518 -232.82 1.3462 0.2493455

- euroRate 1 0.1311 4.1167 -231.11 2.6649 0.1064631

<none> 3.9856 -229.92

- OECD\_CPI 1 0.2152 4.2008 -228.92 4.3743 0.0396183 \*

- Season:dollarRate 3 0.7955 4.7811 -224.31 5.3894 0.0019673 \*\*

- GNP 1 0.6374 4.6230 -218.58 12.9539 0.0005483 \*\*\*

- unitMailPackets 5 2.3349 6.3205 -203.53 9.4907 3.755e-07 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.6828 6.6684 -188.38 18.1746 4.140e-09 \*\*\*

- forecastInflation\_12 1 7.4821 11.4677 -120.47 152.0594 < 2.2e-16 \*\*\*

---

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

Step: AIC=-233.4

CPI ~ Season + forecastInflation\_12 + dollarRate + euroRate +

population + AVGWage + OECD\_CPI + unitMailPackets + GNP +

governmentDeficit + Minister.of.Finance + Season:dollarRate +

governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

- population 1 0.0535 4.0839 -236.65 1.0879 0.2999952

- euroRate 1 0.1068 4.1373 -235.25 2.1739 0.1442032

- AVGWage 1 0.1382 4.1687 -234.44 2.8123 0.0973544 .

<none> 4.0305 -233.40

- OECD\_CPI 1 0.2424 4.2728 -231.77 4.9309 0.0291396 \*

- Season:dollarRate 3 0.7969 4.8274 -227.96 5.4043 0.0019197 \*\*

- GNP 1 0.6710 4.7015 -221.44 13.6525 0.0003957 \*\*\*

- unitMailPackets 5 2.5989 6.6294 -203.06 10.5750 7.380e-08 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.6535 6.6840 -192.81 17.9953 4.602e-09 \*\*\*

- forecastInflation\_12 1 7.5146 11.5450 -124.42 152.8842 < 2.2e-16 \*\*\*

---

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

Step: AIC=-236.65

CPI ~ Season + forecastInflation\_12 + dollarRate + euroRate +

AVGWage + OECD\_CPI + unitMailPackets + GNP + governmentDeficit +

Minister.of.Finance + Season:dollarRate + governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

- euroRate 1 0.0752 4.1591 -239.37 1.5283 0.219852

<none> 4.0839 -236.65

- AVGWage 1 0.2102 4.2942 -235.91 4.2724 0.041854 \*

- Season:dollarRate 3 0.7877 4.8716 -231.65 5.3361 0.002068 \*\*

- OECD\_CPI 1 0.5552 4.6391 -227.57 11.2834 0.001183 \*\*

- GNP 1 1.4453 5.5292 -208.61 29.3732 5.731e-07 \*\*\*

- unitMailPackets 5 2.6089 6.6928 -206.72 10.6042 6.758e-08 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.7099 6.7938 -195.73 18.3579 3.165e-09 \*\*\*

- forecastInflation\_12 1 7.6941 11.7781 -126.94 156.3716 < 2.2e-16 \*\*\*

---

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

Step: AIC=-239.37

CPI ~ Season + forecastInflation\_12 + dollarRate + AVGWage +

OECD\_CPI + unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

Season:dollarRate + governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

- AVGWage 1 0.1683 4.3274 -239.76 3.3992 0.0687537 .

<none> 4.1591 -239.37

- Season:dollarRate 3 0.7165 4.8756 -236.25 4.8233 0.0037958 \*\*

- OECD\_CPI 1 0.6845 4.8436 -227.59 13.8236 0.0003613 \*\*\*

- GNP 1 1.3714 5.5305 -213.27 27.6965 1.070e-06 \*\*\*

- unitMailPackets 5 3.0743 7.2334 -203.01 12.4180 4.999e-09 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.8471 7.0062 -197.09 19.1672 1.470e-09 \*\*\*

- forecastInflation\_12 1 7.6216 11.7808 -131.60 153.9298 < 2.2e-16 \*\*\*

---

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

Step: AIC=-239.76

CPI ~ Season + forecastInflation\_12 + dollarRate + OECD\_CPI +

unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

Season:dollarRate + governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

<none> 4.3274 -239.76

- Season:dollarRate 3 0.7060 5.0335 -237.49 4.6226 0.004812 \*\*

- OECD\_CPI 1 0.8680 5.1954 -224.70 17.0490 8.479e-05 \*\*\*

- GNP 1 1.4486 5.7761 -213.26 28.4541 7.815e-07 \*\*\*

- unitMailPackets 5 3.1758 7.5033 -203.74 12.4759 4.362e-09 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.8815 7.2089 -198.69 18.8659 1.816e-09 \*\*\*

- forecastInflation\_12 1 7.4888 11.8162 -135.96 147.0953 < 2.2e-16 \*\*\*

---

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

## 6.3.4 נספח – הסרת משתנים – המודל שהתקבל ע"י backward:

modelbackward<-lm(CPI ~ Season + forecastInflation\_12 + dollarRate + OECD\_CPI +

+ unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

+ Season:dollarRate + governmentDeficit:Minister.of.Finance)

## 6.3.5 נספח – הסרת משתנים – summary של המודל שהתקבל ע"י backward:

summary(modelbackward)

Call:

lm(formula = CPI ~ Season + forecastInflation\_12 + dollarRate +

OECD\_CPI + unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

Season:dollarRate + governmentDeficit:Minister.of.Finance)

Residuals:

Min 1Q Median 3Q Max

-0.54738 -0.12609 -0.00025 0.12834 0.49942

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4.704e+00 4.146e+00 1.135 0.259695

Season2 1.716e+00 1.368e+00 1.255 0.212970

Season3 4.107e+00 1.566e+00 2.622 0.010348 \*

Season4 9.369e+00 2.671e+00 3.508 0.000724 \*\*\*

forecastInflation\_12 5.349e-01 4.410e-02 12.128 < 2e-16 \*\*\*

dollarRate 9.926e-01 4.193e-01 2.367 0.020185 \*

OECD\_CPI 2.300e-01 5.569e-02 4.129 8.48e-05 \*\*\*

unitMailPackets2 -7.040e-01 1.307e-01 -5.388 6.25e-07 \*\*\*

unitMailPackets3 -8.889e-01 2.428e-01 -3.661 0.000435 \*\*\*

unitMailPackets4 -1.259e+00 3.191e-01 -3.945 0.000164 \*\*\*

unitMailPackets5 -2.757e+00 4.894e-01 -5.634 2.23e-07 \*\*\*

unitMailPackets7 -3.825e+00 5.956e-01 -6.422 7.37e-09 \*\*\*

GNP 1.523e-04 2.855e-05 5.334 7.81e-07 \*\*\*

governmentDeficit -7.090e-05 5.172e-05 -1.371 0.173986

Minister.of.Finance2 7.521e-01 2.027e-01 3.711 0.000367 \*\*\*

Minister.of.Finance3 3.345e-01 2.970e-01 1.126 0.263280

Minister.of.Finance4 5.684e-01 4.184e-01 1.359 0.177864

Season2:dollarRate -3.559e-01 3.613e-01 -0.985 0.327370

Season3:dollarRate -1.044e+00 4.116e-01 -2.536 0.013038 \*

Season4:dollarRate -2.272e+00 6.761e-01 -3.360 0.001168 \*\*

governmentDeficit:Minister.of.Finance2 2.653e-04 3.735e-05 7.103 3.48e-10 \*\*\*

governmentDeficit:Minister.of.Finance3 2.755e-04 4.363e-05 6.316 1.18e-08 \*\*\*

governmentDeficit:Minister.of.Finance4 2.467e-04 3.993e-05 6.179 2.15e-08 \*\*\*

---

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

Residual standard error: 0.2256 on 85 degrees of freedom

Multiple R-squared: 0.9978, Adjusted R-squared: 0.9973

F-statistic: 1781 on 22 and 85 DF, p-value: < 2.2e-16

## 6.3.6 נספח – הסרת משתנים – אלגוריתם forward:

forwardBIC <- step(modelnull, scope = ~(Season+forecastInflation\_12+interestBOI+dollarRate+euroRate+grossMigration+population+AVGWage+OECD\_CPI+unitMailPackets+GNP+governmentDeficit+Minister.of.Finance+Minister.of.Finance:governmentDeficit+Minister.of.Finance:OECD\_CPI+dollarRate:Season+euroRate:Season), direction = "forward",k=log(108))

Start: AIC=319.85

CPI ~ 1

Df Sum of Sq RSS AIC

+ forecastInflation\_12 1 1939.25 59.69 -54.67

+ OECD\_CPI 1 1705.95 292.99 117.15

+ GNP 1 1576.64 422.30 156.63

+ Minister.of.Finance 3 1597.57 401.37 160.51

+ population 1 1525.12 473.83 169.07

+ AVGWage 1 1407.11 591.83 193.08

+ euroRate 1 1071.86 927.08 241.55

+ unitMailPackets 5 1154.83 844.11 250.16

+ grossMigration 1 533.68 1465.27 290.99

+ interestBOI 1 525.94 1473.01 291.56

<none> 1998.94 319.85

+ governmentDeficit 1 66.25 1932.70 320.89

+ dollarRate 1 0.68 1998.26 324.50

+ Season 3 65.46 1933.49 330.30

Step: AIC=-54.67

CPI ~ forecastInflation\_12

Df Sum of Sq RSS AIC

+ GNP 1 41.597 18.097 -178.885

+ OECD\_CPI 1 41.405 18.289 -177.745

+ population 1 40.699 18.995 -173.651

+ unitMailPackets 5 39.854 19.840 -150.225

+ Minister.of.Finance 3 36.318 23.376 -141.875

+ interestBOI 1 31.969 27.725 -132.810

+ AVGWage 1 28.285 31.409 -119.339

+ euroRate 1 21.116 38.577 -97.136

+ dollarRate 1 14.252 45.442 -79.449

+ grossMigration 1 8.576 51.118 -66.737

<none> 59.694 -54.669

+ governmentDeficit 1 0.518 59.175 -50.929

+ Season 3 1.086 58.608 -42.605

Step: AIC=-178.88

CPI ~ forecastInflation\_12 + GNP

Df Sum of Sq RSS AIC

+ unitMailPackets 5 7.5044 10.592 -213.32

+ interestBOI 1 3.6139 14.483 -198.26

+ dollarRate 1 2.0351 16.062 -187.09

+ euroRate 1 1.3840 16.713 -182.79

<none> 18.097 -178.88

+ OECD\_CPI 1 0.7600 17.337 -178.84

+ governmentDeficit 1 0.6462 17.451 -178.13

+ Minister.of.Finance 3 1.7301 16.367 -175.69

+ AVGWage 1 0.0664 18.030 -174.60

+ population 1 0.0025 18.094 -174.22

+ grossMigration 1 0.0009 18.096 -174.21

+ Season 3 0.6762 17.421 -168.95

Step: AIC=-213.32

CPI ~ forecastInflation\_12 + GNP + unitMailPackets

Df Sum of Sq RSS AIC

+ OECD\_CPI 1 1.13500 9.4573 -220.88

+ population 1 0.83546 9.7569 -217.51

+ dollarRate 1 0.82153 9.7708 -217.36

<none> 10.5923 -213.32

+ AVGWage 1 0.40412 10.1882 -212.84

+ euroRate 1 0.35827 10.2341 -212.35

+ interestBOI 1 0.22147 10.3709 -210.92

+ governmentDeficit 1 0.01822 10.5741 -208.82

+ grossMigration 1 0.01396 10.5784 -208.78

+ Season 3 0.67398 9.9183 -206.37

+ Minister.of.Finance 3 0.12348 10.4688 -200.54

Step: AIC=-220.88

CPI ~ forecastInflation\_12 + GNP + unitMailPackets + OECD\_CPI

Df Sum of Sq RSS AIC

+ dollarRate 1 1.15085 8.3065 -230.21

+ interestBOI 1 1.03502 8.4223 -228.71

+ euroRate 1 0.55449 8.9028 -222.72

<none> 9.4573 -220.88

+ population 1 0.37691 9.0804 -220.59

+ AVGWage 1 0.13325 9.3241 -217.73

+ grossMigration 1 0.00155 9.4558 -216.21

+ governmentDeficit 1 0.00133 9.4560 -216.21

+ Season 3 0.27383 9.1835 -210.00

+ Minister.of.Finance 3 0.15925 9.2981 -208.66

Step: AIC=-230.21

CPI ~ forecastInflation\_12 + GNP + unitMailPackets + OECD\_CPI +

dollarRate

Df Sum of Sq RSS AIC

+ euroRate 1 0.46815 7.8383 -231.79

+ interestBOI 1 0.40041 7.9061 -230.86

<none> 8.3065 -230.21

+ AVGWage 1 0.11099 8.1955 -226.98

+ grossMigration 1 0.06003 8.2464 -226.31

+ population 1 0.03248 8.2740 -225.95

+ governmentDeficit 1 0.01851 8.2880 -225.77

+ Season 3 0.57709 7.7294 -223.94

+ Minister.of.Finance 3 0.21651 8.0900 -219.01

Step: AIC=-231.79

CPI ~ forecastInflation\_12 + GNP + unitMailPackets + OECD\_CPI +

dollarRate + euroRate

Df Sum of Sq RSS AIC

<none> 7.8383 -231.79

+ interestBOI 1 0.20784 7.6305 -230.01

+ AVGWage 1 0.06955 7.7688 -228.07

+ grossMigration 1 0.04612 7.7922 -227.75

+ governmentDeficit 1 0.00725 7.8311 -227.21

+ population 1 0.00190 7.8364 -227.14

+ Season 3 0.44836 7.3900 -224.11

+ Minister.of.Finance 3 0.21502 7.6233 -220.75

## 6.3.7 נספח – הסרת משתנים – המודל שהתקבל ע"י אלגוריתם forward:

modelforward<-lm(CPI ~ forecastInflation\_12 + GNP + unitMailPackets + OECD\_CPI +

+ dollarRate + euroRate)

## 6.3.8 נספח – הסרת משתנים – summary של המודל שהתקבל ע"י אלגוריתם forward:

summary(modelforward)

Call:

lm(formula = CPI ~ forecastInflation\_12 + GNP + unitMailPackets +

OECD\_CPI + dollarRate + euroRate)

Residuals:

Min 1Q Median 3Q Max

-0.65627 -0.16516 0.00803 0.18735 1.10515

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.090e+00 3.895e+00 0.280 0.780075

forecastInflation\_12 7.000e-01 2.750e-02 25.459 < 2e-16 \*\*\*

GNP 5.363e-05 2.205e-05 2.432 0.016834 \*

unitMailPackets2 -7.195e-01 1.470e-01 -4.894 3.93e-06 \*\*\*

unitMailPackets3 -6.239e-01 2.470e-01 -2.526 0.013170 \*

unitMailPackets4 -8.194e-01 3.015e-01 -2.718 0.007778 \*\*

unitMailPackets5 -1.624e+00 3.757e-01 -4.322 3.74e-05 \*\*\*

unitMailPackets7 -2.126e+00 4.635e-01 -4.587 1.35e-05 \*\*\*

OECD\_CPI 2.185e-01 4.836e-02 4.518 1.76e-05 \*\*\*

dollarRate 8.604e-01 2.370e-01 3.630 0.000456 \*\*\*

euroRate -5.520e-01 2.293e-01 -2.407 0.017980 \*

---

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

Residual standard error: 0.2843 on 97 degrees of freedom

Multiple R-squared: 0.9961, Adjusted R-squared: 0.9957

F-statistic: 2464 on 10 and 97 DF, p-value: < 2.2e-16

## 6.3.9 נספח – הסרת משתנים –אלגוריתם Stepwise:

stepBIC <- step(fullmudel, direction = "both", k = log(108), test = "F")

Start: AIC=-207.64

CPI ~ Season + forecastInflation\_12 + interestBOI + dollarRate +

Season:dollarRate + euroRate:Season + euroRate + grossMigration +

population + AVGWage + OECD\_CPI + unitMailPackets + GNP +

governmentDeficit + Minister.of.Finance + Minister.of.Finance:governmentDeficit +

Minister.of.Finance:OECD\_CPI

Df Sum of Sq RSS AIC F value Pr(>F)

- OECD\_CPI:Minister.of.Finance 3 0.0198 3.6365 -221.09 0.1353 0.9386853

- Season:euroRate 3 0.2978 3.9145 -213.14 2.0313 0.1168310

- AVGWage 1 0.0271 3.6437 -211.51 0.5538 0.4591121

- interestBOI 1 0.0444 3.6611 -211.00 0.9094 0.3433813

- grossMigration 1 0.0876 3.7043 -209.74 1.7926 0.1847038

- population 1 0.1120 3.7287 -209.03 2.2926 0.1342549

<none> 3.6167 -207.64

- Season:dollarRate 3 0.5625 4.1792 -206.07 3.8366 0.0130438 \*

- GNP 1 0.4157 4.0323 -200.57 8.5047 0.0046861 \*\*

- unitMailPackets 5 1.2393 4.8560 -199.22 5.0716 0.0004728 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.0717 5.6884 -172.77 14.1296 2.26e-07 \*\*\*

- forecastInflation\_12 1 3.5110 7.1277 -139.05 71.8386 1.63e-12 \*\*\*

---

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

Step: AIC=-221.09

CPI ~ Season + forecastInflation\_12 + interestBOI + dollarRate +

euroRate + grossMigration + population + AVGWage + OECD\_CPI +

unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

Season:dollarRate + Season:euroRate + governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

- Season:euroRate 3 0.3404 3.9769 -225.48 2.4025 0.074034 .

- AVGWage 1 0.0309 3.6674 -224.86 0.6533 0.421438

- interestBOI 1 0.0359 3.6724 -224.71 0.7609 0.385755

- grossMigration 1 0.0773 3.7138 -223.50 1.6362 0.204695

- population 1 0.0939 3.7304 -223.02 1.9883 0.162547

- OECD\_CPI 1 0.0994 3.7359 -222.86 2.1052 0.150858

<none> 3.6365 -221.09

- Season:dollarRate 3 0.6305 4.2670 -217.87 4.4502 0.006159 \*\*

- GNP 1 0.4722 4.1087 -212.59 9.9993 0.002241 \*\*

+ OECD\_CPI:Minister.of.Finance 3 0.0198 3.6167 -207.64 0.1353 0.938685

- unitMailPackets 5 1.8784 5.5149 -199.53 7.9549 4.388e-06 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.1200 5.7565 -185.53 14.9632 9.158e-08 \*\*\*

- forecastInflation\_12 1 4.1140 7.7505 -144.05 87.1106 2.737e-14 \*\*\*

---

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

Step: AIC=-225.48

CPI ~ Season + forecastInflation\_12 + interestBOI + dollarRate +

euroRate + grossMigration + population + AVGWage + OECD\_CPI +

unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

Season:dollarRate + governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

- interestBOI 1 0.0087 3.9856 -229.92 0.1751 0.676778

- grossMigration 1 0.0472 4.0241 -228.88 0.9496 0.332749

- AVGWage 1 0.0575 4.0344 -228.61 1.1570 0.285312

- population 1 0.0724 4.0493 -228.21 1.4564 0.231062

- euroRate 1 0.1144 4.0913 -227.09 2.3022 0.133134

- OECD\_CPI 1 0.1747 4.1516 -225.52 3.5141 0.064500 .

<none> 3.9769 -225.48

+ Season:euroRate 3 0.3404 3.6365 -221.09 2.4025 0.074034 .

- Season:dollarRate 3 0.7887 4.7656 -219.98 5.2885 0.002232 \*\*

- GNP 1 0.5087 4.4856 -217.16 10.2328 0.001978 \*\*

+ OECD\_CPI:Minister.of.Finance 3 0.0624 3.9145 -213.14 0.4091 0.746902

- unitMailPackets 5 1.8476 5.8245 -207.68 7.4333 9.050e-06 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.6746 6.6515 -183.97 17.9343 5.399e-09 \*\*\*

- forecastInflation\_12 1 6.2255 10.2024 -128.41 125.2341 < 2.2e-16 \*\*\*

---

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

Step: AIC=-229.92

CPI ~ Season + forecastInflation\_12 + dollarRate + euroRate +

grossMigration + population + AVGWage + OECD\_CPI + unitMailPackets +

GNP + governmentDeficit + Minister.of.Finance + Season:dollarRate +

governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

- grossMigration 1 0.0449 4.0305 -233.40 0.9120 0.3424235

- AVGWage 1 0.0633 4.0489 -232.90 1.2872 0.2599150

- population 1 0.0662 4.0518 -232.82 1.3462 0.2493455

- euroRate 1 0.1311 4.1167 -231.11 2.6649 0.1064631

<none> 3.9856 -229.92

- OECD\_CPI 1 0.2152 4.2008 -228.92 4.3743 0.0396183 \*

+ interestBOI 1 0.0087 3.9769 -225.48 0.1751 0.6767777

+ Season:euroRate 3 0.3132 3.6724 -224.71 2.2170 0.0927441 .

- Season:dollarRate 3 0.7955 4.7811 -224.31 5.3894 0.0019673 \*\*

- GNP 1 0.6374 4.6230 -218.58 12.9539 0.0005483 \*\*\*

+ OECD\_CPI:Minister.of.Finance 3 0.0628 3.9228 -217.59 0.4163 0.7417794

- unitMailPackets 5 2.3349 6.3205 -203.53 9.4907 3.755e-07 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.6828 6.6684 -188.38 18.1746 4.140e-09 \*\*\*

- forecastInflation\_12 1 7.4821 11.4677 -120.47 152.0594 < 2.2e-16 \*\*\*

---

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

Step: AIC=-233.4

CPI ~ Season + forecastInflation\_12 + dollarRate + euroRate +

population + AVGWage + OECD\_CPI + unitMailPackets + GNP +

governmentDeficit + Minister.of.Finance + Season:dollarRate +

governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

- population 1 0.0535 4.0839 -236.65 1.0879 0.2999952

- euroRate 1 0.1068 4.1373 -235.25 2.1739 0.1442032

- AVGWage 1 0.1382 4.1687 -234.44 2.8123 0.0973544 .

<none> 4.0305 -233.40

- OECD\_CPI 1 0.2424 4.2728 -231.77 4.9309 0.0291396 \*

+ grossMigration 1 0.0449 3.9856 -229.92 0.9120 0.3424235

+ interestBOI 1 0.0064 4.0241 -228.88 0.1282 0.7212158

- Season:dollarRate 3 0.7969 4.8274 -227.96 5.4043 0.0019197 \*\*

+ Season:euroRate 3 0.2883 3.7422 -227.36 2.0285 0.1166230

- GNP 1 0.6710 4.7015 -221.44 13.6525 0.0003957 \*\*\*

+ OECD\_CPI:Minister.of.Finance 3 0.0319 3.9986 -220.21 0.2098 0.8893164

- unitMailPackets 5 2.5989 6.6294 -203.06 10.5750 7.380e-08 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.6535 6.6840 -192.81 17.9953 4.602e-09 \*\*\*

- forecastInflation\_12 1 7.5146 11.5450 -124.42 152.8842 < 2.2e-16 \*\*\*

---

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

Step: AIC=-236.65

CPI ~ Season + forecastInflation\_12 + dollarRate + euroRate +

AVGWage + OECD\_CPI + unitMailPackets + GNP + governmentDeficit +

Minister.of.Finance + Season:dollarRate + governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

- euroRate 1 0.0752 4.1591 -239.37 1.5283 0.219852

<none> 4.0839 -236.65

- AVGWage 1 0.2102 4.2942 -235.91 4.2724 0.041854 \*

+ population 1 0.0535 4.0305 -233.40 1.0879 0.299995

+ grossMigration 1 0.0321 4.0518 -232.82 0.6498 0.422518

+ interestBOI 1 0.0023 4.0816 -232.03 0.0463 0.830167

- Season:dollarRate 3 0.7877 4.8716 -231.65 5.3361 0.002068 \*\*

+ Season:euroRate 3 0.3065 3.7775 -231.03 2.1635 0.098787 .

- OECD\_CPI 1 0.5552 4.6391 -227.57 11.2834 0.001183 \*\*

+ OECD\_CPI:Minister.of.Finance 3 0.0102 4.0738 -222.88 0.0666 0.977464

- GNP 1 1.4453 5.5292 -208.61 29.3732 5.731e-07 \*\*\*

- unitMailPackets 5 2.6089 6.6928 -206.72 10.6042 6.758e-08 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.7099 6.7938 -195.73 18.3579 3.165e-09 \*\*\*

- forecastInflation\_12 1 7.6941 11.7781 -126.94 156.3716 < 2.2e-16 \*\*\*

---

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

Step: AIC=-239.37

CPI ~ Season + forecastInflation\_12 + dollarRate + AVGWage +

OECD\_CPI + unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

Season:dollarRate + governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

- AVGWage 1 0.1683 4.3274 -239.76 3.3992 0.0687537 .

<none> 4.1591 -239.37

+ euroRate 1 0.0752 4.0839 -236.65 1.5283 0.2198523

- Season:dollarRate 3 0.7165 4.8756 -236.25 4.8233 0.0037958 \*\*

+ population 1 0.0218 4.1373 -235.25 0.4378 0.5100077

+ grossMigration 1 0.0172 4.1420 -235.13 0.3444 0.5589088

+ interestBOI 1 0.0016 4.1576 -234.72 0.0318 0.8589620

- OECD\_CPI 1 0.6845 4.8436 -227.59 13.8236 0.0003613 \*\*\*

+ OECD\_CPI:Minister.of.Finance 3 0.0367 4.1224 -226.28 0.2406 0.8678118

- GNP 1 1.3714 5.5305 -213.27 27.6965 1.070e-06 \*\*\*

- unitMailPackets 5 3.0743 7.2334 -203.01 12.4180 4.999e-09 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.8471 7.0062 -197.09 19.1672 1.470e-09 \*\*\*

- forecastInflation\_12 1 7.6216 11.7808 -131.60 153.9298 < 2.2e-16 \*\*\*

---

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

Step: AIC=-239.76

CPI ~ Season + forecastInflation\_12 + dollarRate + OECD\_CPI +

unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

Season:dollarRate + governmentDeficit:Minister.of.Finance

Df Sum of Sq RSS AIC F value Pr(>F)

<none> 4.3274 -239.76

+ AVGWage 1 0.1683 4.1591 -239.37 3.3992 0.068754 .

- Season:dollarRate 3 0.7060 5.0335 -237.49 4.6226 0.004812 \*\*

+ grossMigration 1 0.0854 4.2421 -237.23 1.6909 0.197039

+ population 1 0.0766 4.2508 -237.01 1.5143 0.221923

+ euroRate 1 0.0333 4.2942 -235.91 0.6512 0.421968

+ interestBOI 1 0.0002 4.3272 -235.09 0.0042 0.948214

+ OECD\_CPI:Minister.of.Finance 3 0.0294 4.2981 -226.45 0.1867 0.905137

- OECD\_CPI 1 0.8680 5.1954 -224.70 17.0490 8.479e-05 \*\*\*

- GNP 1 1.4486 5.7761 -213.26 28.4541 7.815e-07 \*\*\*

- unitMailPackets 5 3.1758 7.5033 -203.74 12.4759 4.362e-09 \*\*\*

- governmentDeficit:Minister.of.Finance 3 2.8815 7.2089 -198.69 18.8659 1.816e-09 \*\*\*

- forecastInflation\_12 1 7.4888 11.8162 -135.96 147.0953 < 2.2e-16 \*\*\*

---

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

## 6.3.10 נספח – הסרת משתנים –המודל שהתקבל ע"י אלגוריתם Stepwise:

modelboth<-lm(CPI ~ Season + forecastInflation\_12 + dollarRate + OECD\_CPI +

+ unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

+ Season:dollarRate + governmentDeficit:Minister.of.Finance)

## 6.3.11 נספח – הסרת משתנים – summary של המודל שהתקבל ע"י אלגוריתם Stepwise:

summary(modelboth)

Call:

lm(formula = CPI ~ Season + forecastInflation\_12 + dollarRate +

OECD\_CPI + unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

Season:dollarRate + governmentDeficit:Minister.of.Finance)

Residuals:

Min 1Q Median 3Q Max

-0.54738 -0.12609 -0.00025 0.12834 0.49942

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4.704e+00 4.146e+00 1.135 0.259695

Season2 1.716e+00 1.368e+00 1.255 0.212970

Season3 4.107e+00 1.566e+00 2.622 0.010348 \*

Season4 9.369e+00 2.671e+00 3.508 0.000724 \*\*\*

forecastInflation\_12 5.349e-01 4.410e-02 12.128 < 2e-16 \*\*\*

dollarRate 9.926e-01 4.193e-01 2.367 0.020185 \*

OECD\_CPI 2.300e-01 5.569e-02 4.129 8.48e-05 \*\*\*

unitMailPackets2 -7.040e-01 1.307e-01 -5.388 6.25e-07 \*\*\*

unitMailPackets3 -8.889e-01 2.428e-01 -3.661 0.000435 \*\*\*

unitMailPackets4 -1.259e+00 3.191e-01 -3.945 0.000164 \*\*\*

unitMailPackets5 -2.757e+00 4.894e-01 -5.634 2.23e-07 \*\*\*

unitMailPackets7 -3.825e+00 5.956e-01 -6.422 7.37e-09 \*\*\*

GNP 1.523e-04 2.855e-05 5.334 7.81e-07 \*\*\*

governmentDeficit -7.090e-05 5.172e-05 -1.371 0.173986

Minister.of.Finance2 7.521e-01 2.027e-01 3.711 0.000367 \*\*\*

Minister.of.Finance3 3.345e-01 2.970e-01 1.126 0.263280

Minister.of.Finance4 5.684e-01 4.184e-01 1.359 0.177864

Season2:dollarRate -3.559e-01 3.613e-01 -0.985 0.327370

Season3:dollarRate -1.044e+00 4.116e-01 -2.536 0.013038 \*

Season4:dollarRate -2.272e+00 6.761e-01 -3.360 0.001168 \*\*

governmentDeficit:Minister.of.Finance2 2.653e-04 3.735e-05 7.103 3.48e-10 \*\*\*

governmentDeficit:Minister.of.Finance3 2.755e-04 4.363e-05 6.316 1.18e-08 \*\*\*

governmentDeficit:Minister.of.Finance4 2.467e-04 3.993e-05 6.179 2.15e-08 \*\*\*

---

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

Residual standard error: 0.2256 on 85 degrees of freedom

Multiple R-squared: 0.9978, Adjusted R-squared: 0.9973

F-statistic: 1781 on 22 and 85 DF, p-value: < 2.2e-16

## 

## 6.3.12 נספח – בדיקת הנחות המודל – פיזור השגיאות המתוקננות והאומד המוסבר:

dataset$fitted<-fitted(modelbackward) # predicted values

dataset$residuals<-residuals(modelbackward) # residuals

s.e\_res <- sqrt(var(dataset$residuals))

dataset$stan\_residuals<-(residuals(modelbackward)/s.e\_res)

## 6.3.13 נספח – בדיקת הנחות המודל – פיזור השגיאות המתוקננות והאומד המוסבר:

plot(dataset$fitted, dataset$residuals,ylab ="Normalized error", xlab="predicted values", main="Normalized error vs. predicted values")

abline(0, 0)

## 6.3.14 נספח – בדיקת הנחות המודל – היסטוגרמה:

qqnorm(dataset$stan\_residuals)

abline(a=0, b=1)

hist(dataset$stan\_residuals, xlab ="Normalized error", main="Histogram of normalized error")

## 6.3.15 נספח – בדיקת הנחות המודל – Q-Q Plot:

qqnorm(dataset$stan\_residuals, main = "Normal Q-Q Plot",

xlab = "Theoretical Quantiles", ylab = "Standardized residuals",

plot.it = TRUE, datax = FALSE)

## 6.3.16 נספח – בדיקת הנחות המודל – Kolmogorov-Smirnov:

ks.test(x= dataset$stan\_residuals,y="pnorm",alternative = "two.sided", exact = NULL)

## 6.3.17 נספח – בדיקת הנחות המודל – Shapiro-Wilk:

shapiro.test(dataset$stan\_residuals)

## 6.4 נספח – בדיקת השערות :

## 6.4.1 נספח – מבחן F חלקי לבדיקה האם המודל הנבחר עדיף על המודל המלא :

anova(fullmudel, modelbackward)

Analysis of Variance Table

Model 1: CPI ~ Season + forecastInflation\_12 + interestBOI + dollarRate +

Season:dollarRate + euroRate:Season + euroRate + grossMigration +

population + AVGWage + OECD\_CPI + unitMailPackets + GNP +

governmentDeficit + Minister.of.Finance + Minister.of.Finance:governmentDeficit +

Minister.of.Finance:OECD\_CPI

Model 2: CPI ~ Season + forecastInflation\_12 + dollarRate + OECD\_CPI +

unitMailPackets + GNP + governmentDeficit + Minister.of.Finance +

Season:dollarRate + governmentDeficit:Minister.of.Finance

Res.Df RSS Df Sum of Sq F Pr(>F)

1 74 3.6167

2 85 4.3274 -11 -0.71078 1.3221 0.2296

## 6.4.2 נספח – מטריצת שוניות משותפת של האומדים למקדמי הרגרסיה שנבחרה בשיטת הצעדים בשביל בדיקת משתנה קטגוריאלי - עונות:

vcov(modelbackward)

## 6.5 נספח – שיפור המודל - טרנספורמציה לא לינארית של המשתנים המסבירים:

## 6.5.1 נספח – שיפור המודל – מציאת מטריצת שוניות משותפת:

> d1<-data.frame(CPI,forecastInflation\_12,forcastpow,dataset$sqforcast,dataset$lnforcast,dollarRate,dollarpow,dataset$sqdollar,dataset$lndollar,OECD\_CPI,oecdcpipow,dataset$sqoecdcpi,dataset$lnoecdcpi,GNP,gnppow,dataset$sqgnp,dataset$lngnp)

> cor(d1)

## 6.5.2 נספח – שיפור המודל – טרנספורמציית בוקס קוקס:

> BoxCoxTrans(CPI)

Box-Cox Transformation

108 data points used to estimate Lambda

Input data summary:

Min. 1st Qu. Median Mean 3rd Qu. Max.

86.45 93.43 99.21 97.08 100.40 101.70

Largest/Smallest: 1.18

Sample Skewness: -0.889

Estimated Lambda: 2

## 6.5.3 נספח – שיפור המודל –המודל שהתקבל לאחר טרנספורמציה לא לינארית:

> lastmodel<-lm(cpipow~Season + poly(forecastInflation\_12, 2) + dollarRate+Season:dollarRate + dataset$lnoecdcpi + dataset$lngnp + unitMailPackets +

+ + governmentDeficit + Minister.of.Finance + Minister.of.Finance:governmentDeficit +

+ + dataset$sqavgwage + dataset$sqpopulation)

## 6.5.4 נספח – שיפור המודל – summary על המודל המשופר:

> summary(lastmodel)

Call:

lm(formula = cpipow ~ Season + poly(forecastInflation\_12, 2) +

dollarRate + Season:dollarRate + dataset$lnoecdcpi + dataset$lngnp +

unitMailPackets + +governmentDeficit + Minister.of.Finance +

Minister.of.Finance:governmentDeficit + +dataset$sqavgwage +

dataset$sqpopulation)

Residuals:

Min 1Q Median 3Q Max

-104.116 -21.120 0.549 21.631 98.333

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -3.656e+04 5.267e+03 -6.941 8.36e-10 \*\*\*

Season2 1.737e+02 2.701e+02 0.643 0.522128

Season3 5.350e+02 3.004e+02 1.781 0.078624 .

Season4 1.436e+03 4.994e+02 2.875 0.005150 \*\*

poly(forecastInflation\_12, 2)1 4.159e+03 3.702e+02 11.235 < 2e-16 \*\*\*

poly(forecastInflation\_12, 2)2 7.956e+01 9.713e+01 0.819 0.415083

dollarRate 1.314e+02 8.576e+01 1.532 0.129365

dataset$lnoecdcpi 4.015e+03 1.252e+03 3.206 0.001920 \*\*

dataset$lngnp 2.135e+03 5.304e+02 4.025 0.000126 \*\*\*

unitMailPackets2 -1.379e+02 2.506e+01 -5.504 4.15e-07 \*\*\*

unitMailPackets3 -1.680e+02 4.669e+01 -3.599 0.000546 \*\*\*

unitMailPackets4 -2.369e+02 6.774e+01 -3.498 0.000761 \*\*\*

unitMailPackets5 -4.986e+02 1.058e+02 -4.711 9.92e-06 \*\*\*

unitMailPackets7 -6.943e+02 1.353e+02 -5.131 1.90e-06 \*\*\*

governmentDeficit -9.236e-03 9.521e-03 -0.970 0.334861

Minister.of.Finance2 9.653e+01 5.316e+01 1.816 0.073054 .

Minister.of.Finance3 3.506e+01 6.110e+01 0.574 0.567674

Minister.of.Finance4 6.662e+01 8.183e+01 0.814 0.417916

dataset$sqavgwage 6.275e+00 3.794e+00 1.654 0.102002

dataset$sqpopulation 2.556e+01 5.013e+01 0.510 0.611472

Season2:dollarRate -2.562e+01 7.134e+01 -0.359 0.720412

Season3:dollarRate -1.297e+02 7.881e+01 -1.645 0.103765

Season4:dollarRate -3.381e+02 1.262e+02 -2.679 0.008923 \*\*

governmentDeficit:Minister.of.Finance2 4.398e-02 7.381e-03 5.959 6.11e-08 \*\*\*

governmentDeficit:Minister.of.Finance3 4.295e-02 7.963e-03 5.394 6.51e-07 \*\*\*

governmentDeficit:Minister.of.Finance4 4.040e-02 7.033e-03 5.744 1.52e-07 \*\*\*

---

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

Residual standard error: 41.29 on 82 degrees of freedom

Multiple R-squared: 0.9981, Adjusted R-squared: 0.9975

F-statistic: 1696 on 25 and 82 DF, p-value: < 2.2e-16

## 6.5.5 נספח – שיפור המודל – פיזור השגיאות המתוקננות והאומד המוסבר:

> dataset$fitted<-fitted(lastmodel) # predicted values

> dataset$residuals<-residuals(lastmodel) # residuals

> s.e\_res <- sqrt(var(dataset$residuals))

> dataset$stan\_residuals<-(residuals(lastmodel)/s.e\_res)

>

>

> plot(dataset$fitted, dataset$residuals,ylab ="Normalized error", xlab="predicted values", main="Normalized error vs. predicted values")

> abline(0, 0)