דו"ח קורס סדרות עתיות

**מגיש**: אייל בן ציון

**ת.ז**: 034029926

# הקדמה

היום ניתן "לעקוב" אחרי אנשים ולהסיק על אורח חייהם. אם בעבר היה צורך בתצפיות, יומנים אישים, וראיונות, היום בעזרת הטכנולוגיה ניתן לדעת איפה אדם היה ומתי. בעזרת איכוני האדם ניתן להסיק על אורח חייו, האם אדם זה הינו מכור לעבודה, האם בליין או עקר בית. בדו"ח זה אנסה לנתח את תנועת האדם מבחינת טרנד, מחזוריות וגורמים משפעים (כמו מספר תקשורות בשעה ומספר נ.צ בשעה שבהם ביקר) בעזרת רגרסיה רב משתנית (יותר ממשתנה תלוי אחד).

# עיבוד הנתונים

בסיס הנתונים הינו סדרה עתית של רשומות של תקשורת בין מכשיר סלולרי של אדם לאנטנות של הרשת הסלולרית שמסביבו. איכוני המכשיר אינם תמיד מדויקים, וקיימת סטייה באיכון המכשיר המשתנה כפונקציה בצפיפות האנטנות מסביב למכשיר (ככל שיש פחות אנטנות כך הסטייה גדולה יותר). האיכון מתבצע אך ורק על ידי הרשת הסלולרית.

כל רשומה מייצגת איזה שהיא תקשורת (אינטרנט, הודעות ואו שיחה) בין המכשיר הסלולרי לאנטנה שאליה שייך בזמן נתון. מרווחי הזמן בין רשומה לרשומה אינו קבוע, מפני שכל תקשורת תלויה ברגע שמשתמש יבצע איזה שהיא תקשורת עם האנטנה. לכן סידרה עתית זו אינה אחידה. לכן בסיס הנתונים מכיל נ.צ.ים של המיקום והזמן של התקשורת.

על מנת לנסות ולצור אחידות בסדרה העתית נקטתי במספר אפשרויות פעולה, כאשר לבסוף בחרתי את האפשרות הטובה ביותר.

# הוצאת חריגים

דבר ראשון אשר משותף לכל האפשרויות היה הוצאת חריגים. בגלל שאיכוני המכשיר הינם בעלי סטייה, וישנם רשומות שלא מייצגות את המיקום האמיתי של המכשיר היה לי חשוב להוציא רשומות אלו. הדרך שבחרתי היא על פי מדדי מהירות ותאוצה. כאשר המרחק בין רשומות עוקבות חלקי הזמן שעבר בין שתי הרשומות עוקבות היה גדול מ300 ק"מ לשעה (מהירות) ואו כאשר שני ציוני מהירויות עוקבות חלקי הזמן שעבר בין מהירויות אלו היה גדול מ9.8 מטר לשנייה בריבוע (תאוצה). כפי שניתן להבין חישוב המהירות התבסס על שתי רשומות עוקבות ואילו חישוב התאוצה התבסס על שלוש רשומות עוקבות, כל פעם הוצאתי את הרשומה האחרונה שגרמה לסטייה.

# מציאת חלון בבסיס הנתונים תקף

לאחר הוצאת החריגים חיפשתי חלון בבסיס הנתונים שבו היחס של מספר הרשומות החסרות אינו קטן ביותר. התמקדתי בשני רזולוציות שונות (של דקה ושל שעה). כאשר בכל התייחסות של גודל רזולוציה ביצעתי ברגע שהיה שינוי בדקה או שעה אז חישבתי את הממוצע/ שכיח של כל נ.צ.ים של הדקה/ שעה לפני השינוי.

יש לציין כאשר חיפשתי ברזולוציה של דקה מספר הרשומות החסרות היה גבוה מאוד, מעל שלושים אחוז. ואילו כאשר חיפשתי ברזולוציה של שעה היו רשומות חסרות בחלון שבו חיפשתי מתחת ל10 אחוז. לכן העדפתי לעבוד עם רזולוציה של שעה.

# הכנת הנתונים לרגרסיה הרב משתנית

לאחר בחירת החלון ביצעתי שני שיטות לבחירת הנ.צ. בשעה מסוימת. בשיטה הראשונה פשוט ביצעתי ממוצע לכל הנ.צ.ים אשר הופיעו בשעה מסוימת, ואליו בשיטה השנייה פשוט מצאתי את הנ.צ. השכיח בשעה מסוימת. בטבלה החדשה שמרתי את הנ.צ. הממוצע/ שכיח ואת הזמן (בשעות) מצטבר ובסדר עולה.

# ביצוע אינטרפולציה

לערכים החסרים ביצעתי אינטרפולציה בכמה שיטות לינארי.

# Fast Fourier Transform

על ידי ביצוע התמרת פורייה בדידה אנו ממירים את האות במרחב הזמן למרחב התדר. התמרת פורייה מהירה הינו אלגוריתם לביצוע התמרת פורייה בדידה, אשר בעזרת התמרת פורייה מהירה ניתן לקבל בדיוק אותו תוצאות כמו התמרת פורייה בדידה רק בזמן קצר יותר משמעותי.

 X_k =  \sum_{n=0}^{N-1} x_n e^{-{i 2\pi k \frac{n}{N}}}
\qquad
k = 0,\dots,N-1. 

x הינו וקטור האות במרחב הזמן ואילו X הינו וקטור האות במרחב התדר כאשר X הינו וקטור מורכב. על ידי ההתמרה אפשר לגלות את הפונקציות הטריגונומטרית שמרכיבות את האות.

בגלל שלא ניתן לבצע התמרת פורייה למשתנה תלוי דו ממדי, אני ביצעתי את ההתמרה לכל קורדינאטה בנפרד (פעם לקורדינאטת הרוחב ופעם לקורדינאטת האורך).

# מדוע להשתמש בהתמרת פורייה

אדם לרוב מבקר באותם מקומות לאורך תקופה מסוימת, במידה ולא שינה מקום עבודה, מקום מגורים ואו שינוי כלשהו בהרגליו. על כן כאשר מסתכלים על תנועת האדם לאורך הזמן ולמען פישוט הבעיה על ציר אחד במערכת הקורדינאטות (אורך או רוחב) ניתן לראות כי תנועת האדם בציר זה מתנהגת בדיוק כמו אות בעלת מחזוריות ואמפליטודה של כל מחזור.

על כן בעזרת התמרת פורייה אני מתכוון לגלות את מבנה האות (התנועה). כאשר אדע את התדרים שמרכיבים את האות אוכל לגלות את מבנה הפונקציה הטריגונומטרית (המורכבת ממספר תדרים המשפיעים ביותר). לאחר מכן אוכל בעזרת הפונקציה הטריגונומטרית לבטל את השפעת המחזוריות ולהפעיל מודל רגרסיה על מנת למצוא השפעת של משתנים מסבירים.

# מסנן קלמן

מסנן קלמן הוא [אלגוריתם](http://he.wikipedia.org/wiki/%D7%90%D7%9C%D7%92%D7%95%D7%A8%D7%99%D7%AA%D7%9D), המקבל כקלט סדרה של מדידות מורעשות של ערכים שונים המייצגים מצב של מערכת בזמן מסוים, ומייצר הערכה, בדרך כלל מדויקת יותר, של אותם הערכים על סמך תוצאת המסנן בעבר. המסנן עובד בצורה רקורסיבית על קלט [מורעש](http://he.wikipedia.org/wiki/%D7%A8%D7%A2%D7%A9) ומייצר הערכה [סטטיסטית](http://he.wikipedia.org/wiki/%D7%A1%D7%98%D7%98%D7%99%D7%A1%D7%98%D7%99%D7%A7%D7%94) של מצב המערכת.

משוואת האלגוריתם הינה:  \textbf{x}_{k} = \textbf{F}_{k} \textbf{x}_{k-1} + \textbf{B}_{k} \textbf{u}_{k} + \textbf{w}_{k} 

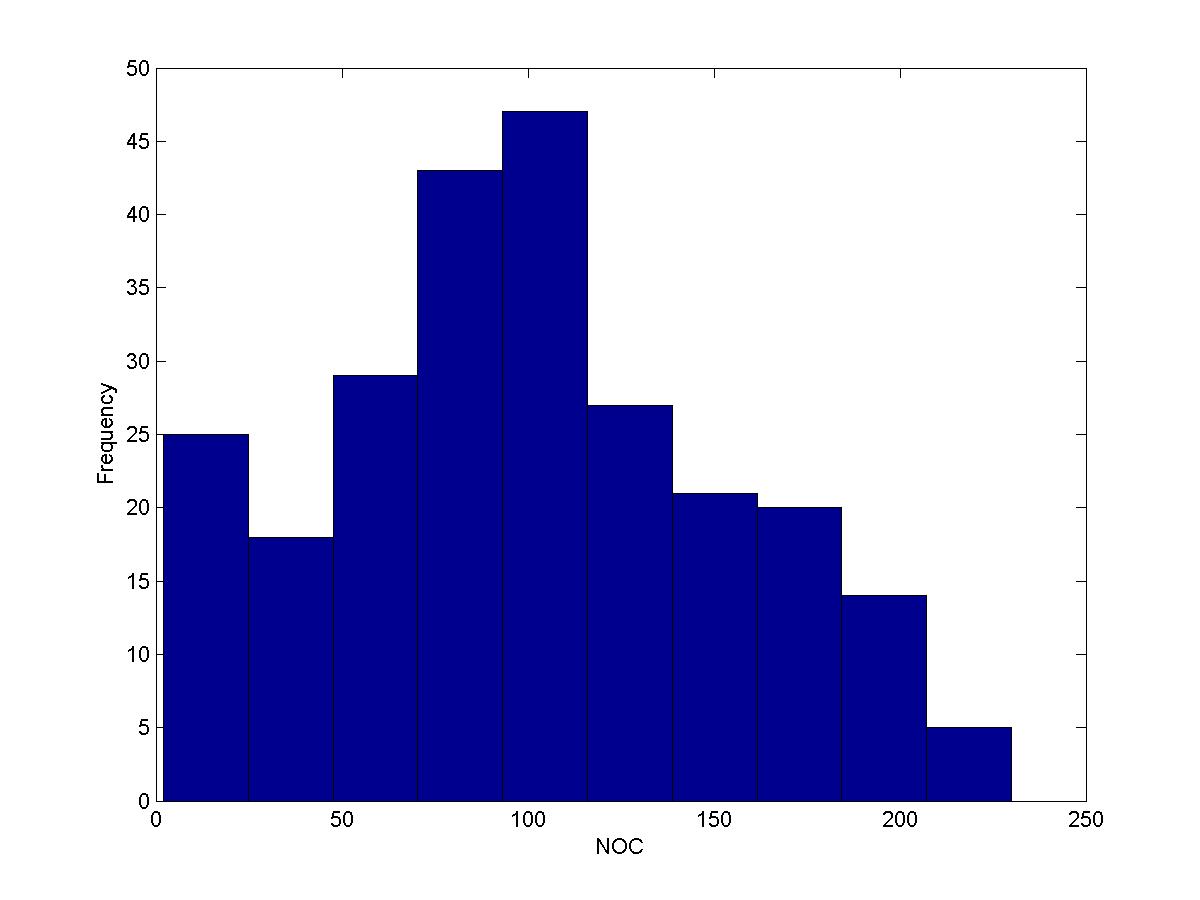
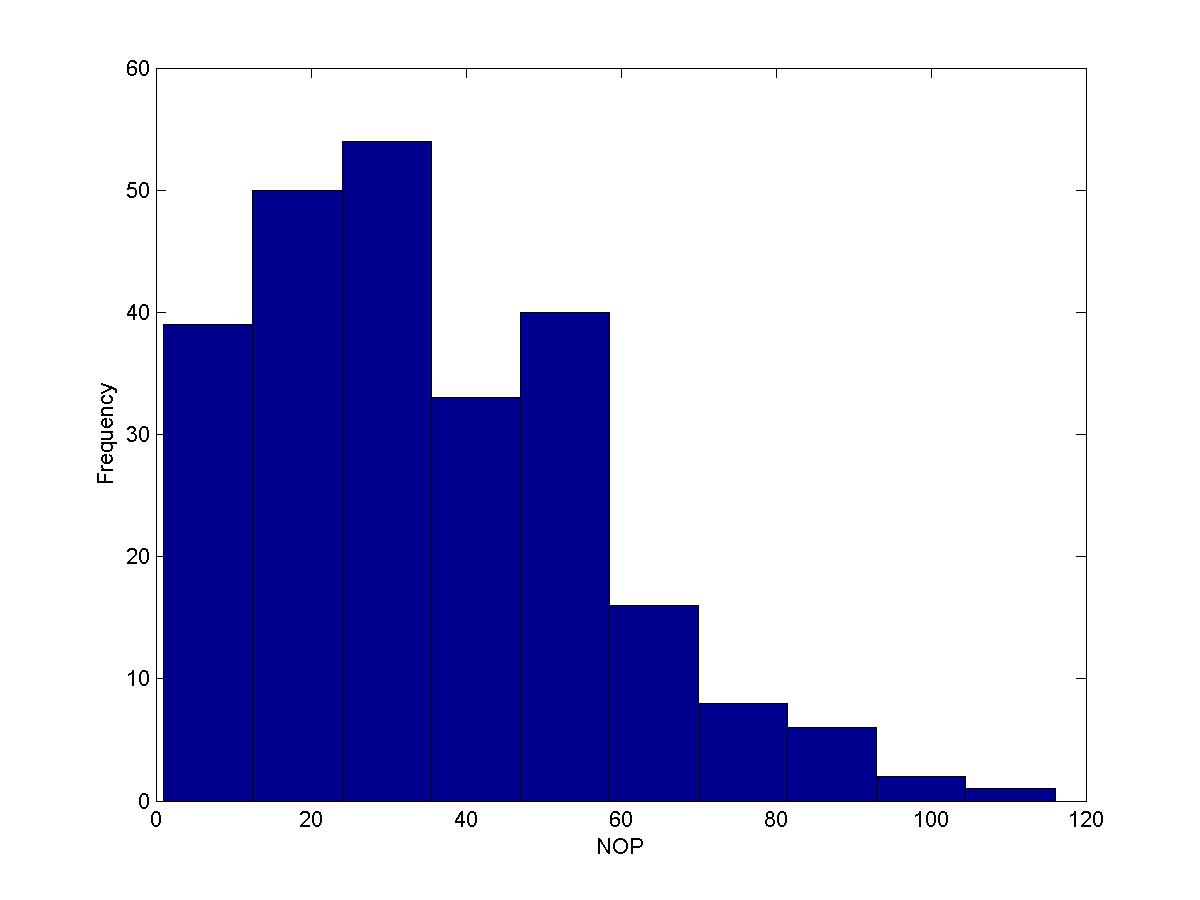
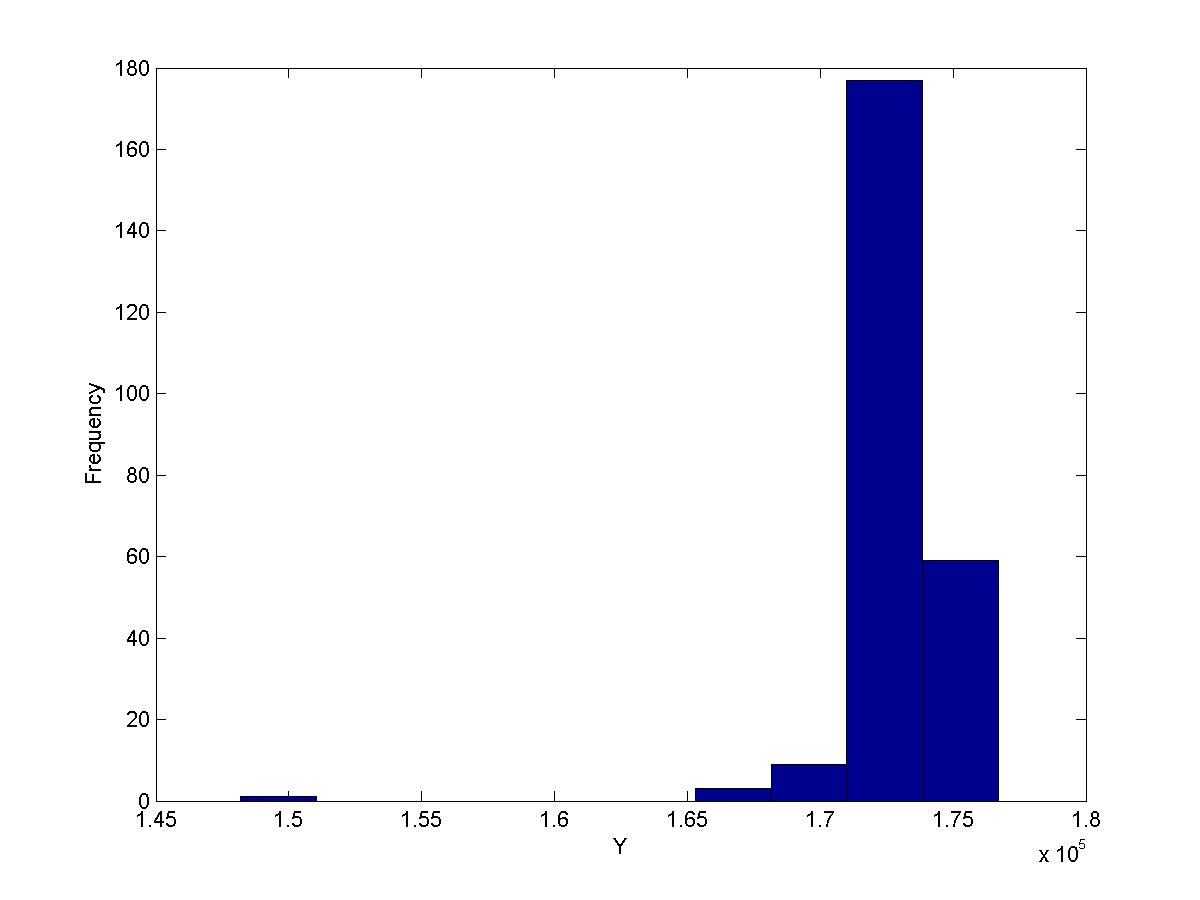
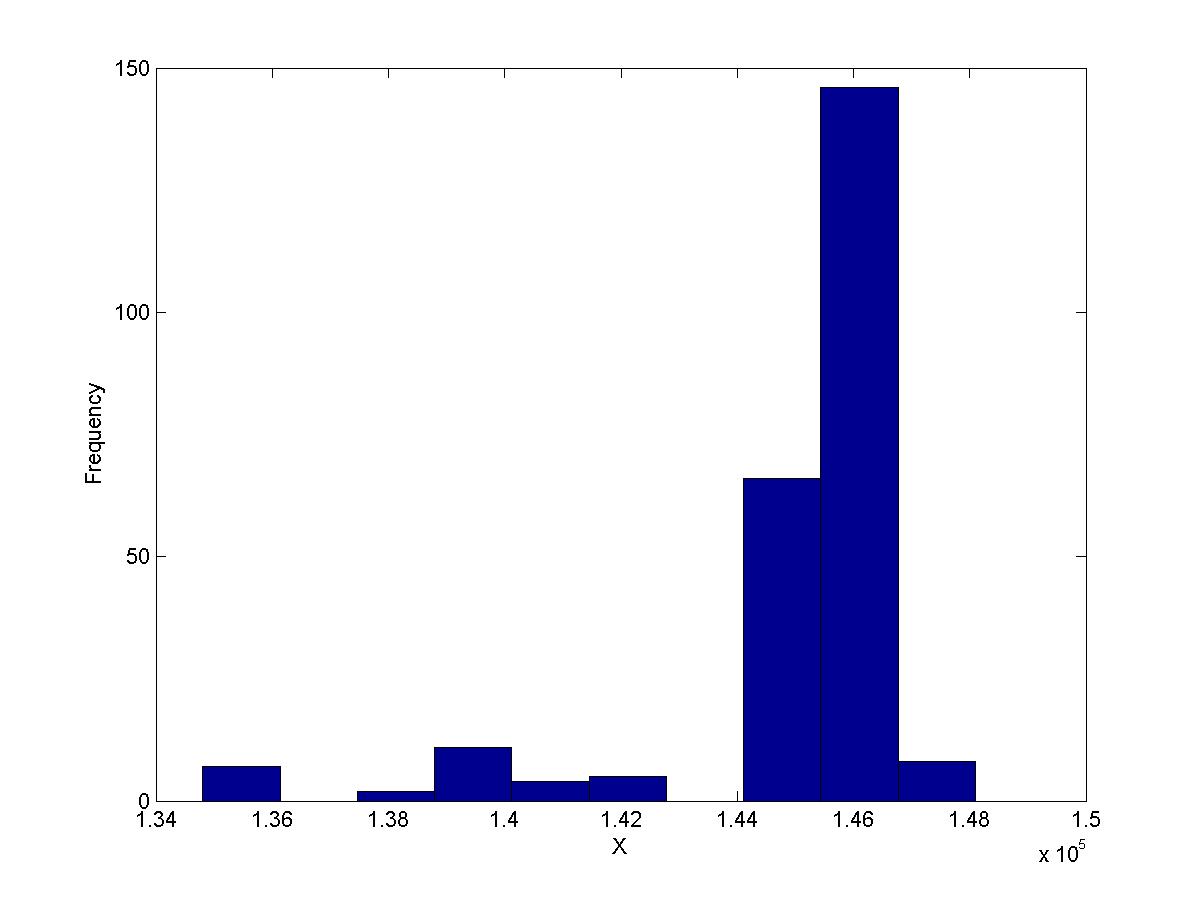
כאשר x הינו וקטור הנ.צ והמהירות, u הינה התאוצה, w הינו הרעש ואילו B וF הינה המקדמים של המשוואה. k הינו מצב נוכחי ואילו k-1 מצב קודם.

# תוצאות סטטיסטיקה תיאורית

להלן תיאור משתני הקובץ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| מספר תקשורות (NOC) | מספר מקומות (NOP) | קורדינטת Y | קורדינטת X |  |
| 103.5818 | 35.5905 | 173173.7974 | 144795.7715 | **Mean** |
| 3 | 27 | 174502 | 144571 | **Mode** |
| 101.5 | 31 | 173200 | 145687.5 | **Median** |
| 2 | 1 | 148189 | 134808 | **Min** |
| 230 | 116 | 176706 | 148095 | **Max** |
| מספר תקשורות שביצע המשתמש במהלך השעה. יש לציין כי משתנה מספר תקשורות בעצם מייצג מספר תקשורות עם האנטנות הסלולרית, כך שלא חייב להיות בהכרח שיחה (יוכל להיות הודעות, תקשורת אינטרנט ואו סתם איכון המכשיר). | מספר מקומות שבהם המשתמש ביקר במהלך השעה (על פי נ.צ השונים שהופיעו במהלך השיחות באותה שעה). יש לציין כי מספר הנ.צ השונים נלקחו גם כאשר המשתמש היה בתנועה ולא רק בעצירה של המשתמש. | נ.צ של ציר האורך | נ.צ של ציר הרוחב | תיאור |

להלן תיאור המשתנים בצורה גרפית



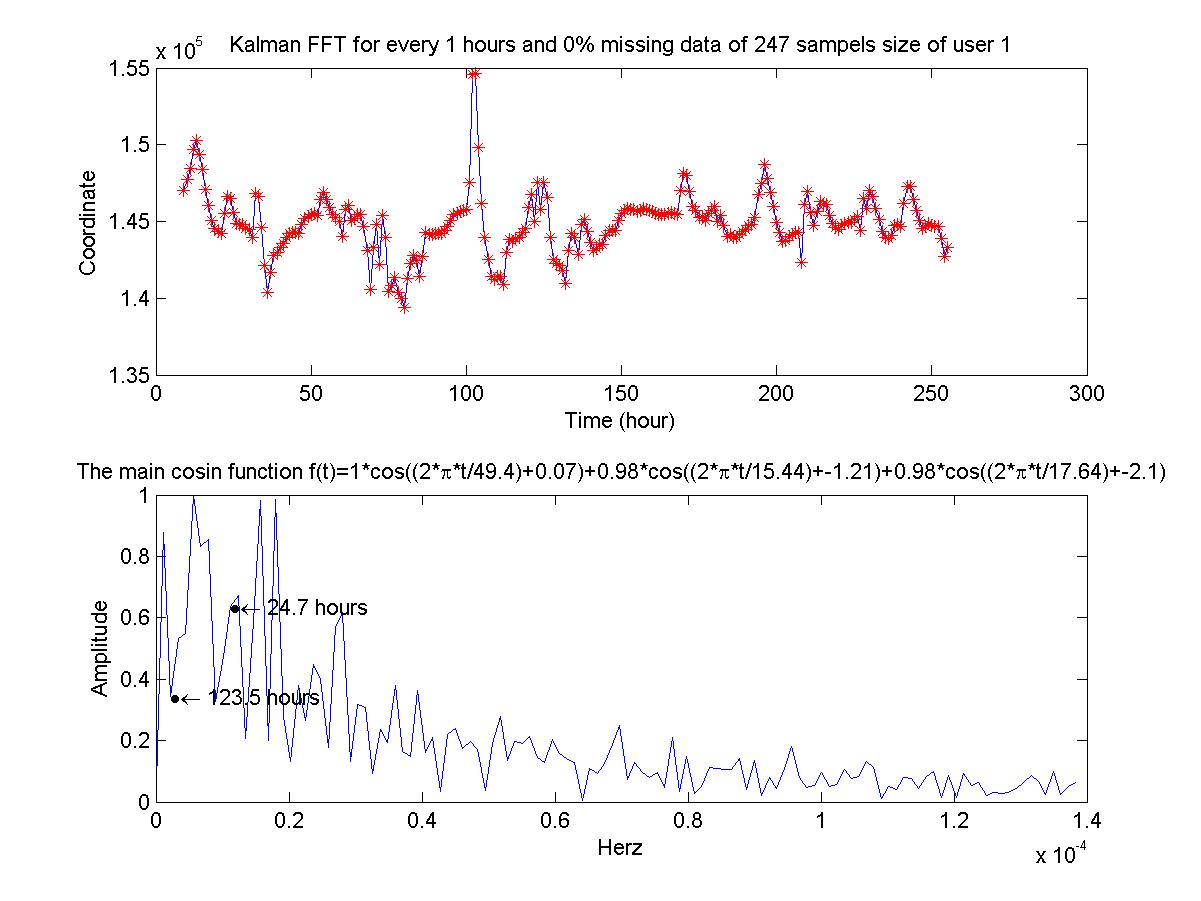
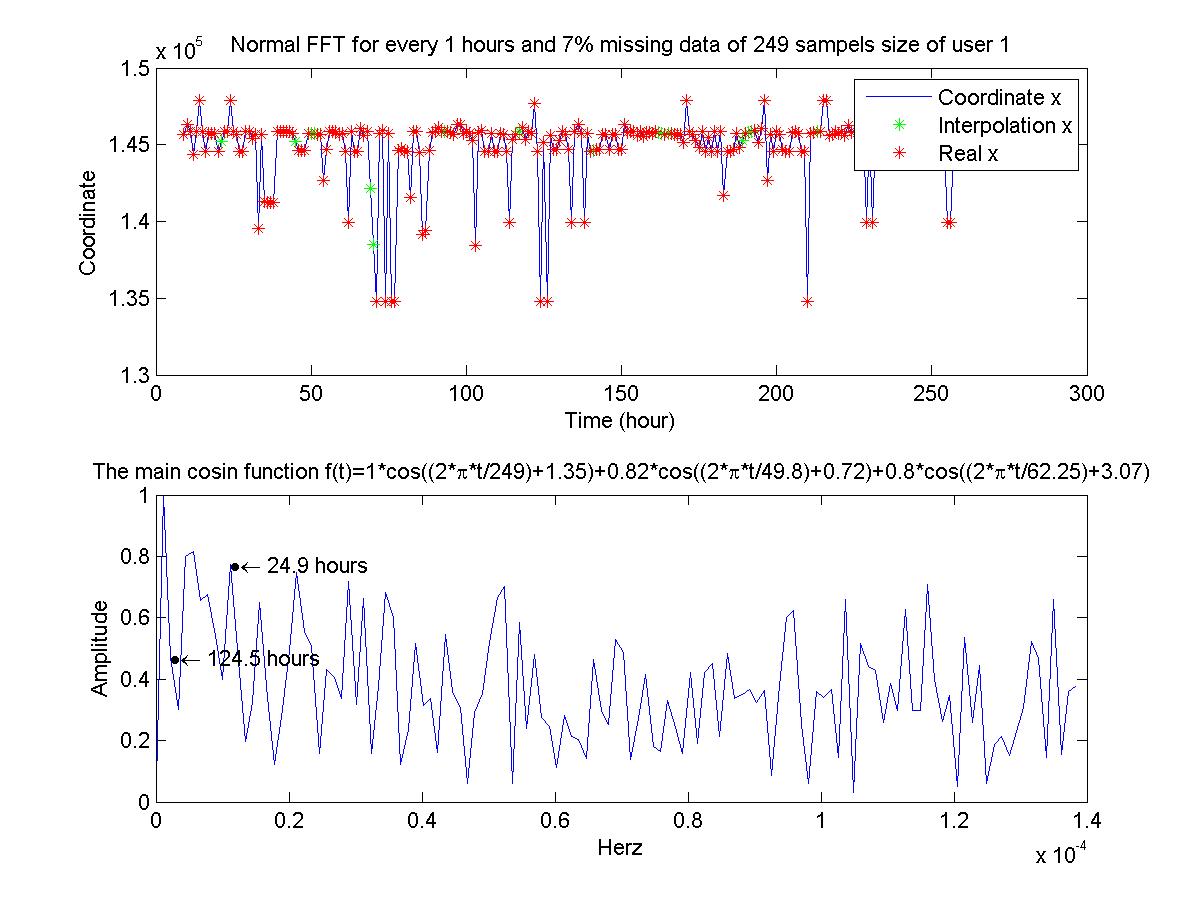
כפי שניתן לראות רק משתנים מספר תקשורות מספר מקומות מתנהגים דומה להתפלגות נורמלית (ממוצע וחציון כמעט שווים) אף הגרף דומה מזכיר התפלגות נורמלית אך עם זנב ימני.

# התמרת פורייה

הגרפים הבאים מציגים את ההתמרה לאורך ציר הרוחב (קורדינאטה x) לכל אחד משלושת המשתמשים ללא ועם מסנן קלמן. בנוסף ניתן לראות את הפונקציה הקוסינאוסית כתלות בזמן של שלושת הפיקים הגבוהים ביותר. יש לציין כי האמפליטודה מנורמלת לפי ערך האמפליטודה המקסימלית.

יש לציין שהגרפים מראים רק את התנועה בציר הרוחב כתלות בזמן, למרות שהפונקציה הטריגונומטרית נבנתה בנפרד עבור ציר הרוחב ועבור ציר האורך.

משתמש 1



ניתן לראות כי בעזרת מסנן קלמן הזנב הימני פחות רועש, זאת אומרת שיש הרבה פחות רעשים בתנועה (הרבה פחות מחזורים של רעש).

* הסיבה שבגרפים של מסנן קלמן מופיע כי אין נתונים חסרים מפני שהעיבוד נעשה לאחר האינטרפולציה
* היסה שיש פחות רשומות במסנן קלמן מפני ששני הרשומות הראשונות נאבדו לטובת חישוב המהירות והתאוצה.

# ARIMA

בדקתי שלושה מודלי ARIMA (1,1,0), (0,1,1) ו(0,1,0) לשני בסיסי הנתונים (בלי/ עם מסנן קלמן) ועם המשתנים מספר תקשורות ומספר מקומות וזמן. להלן התוצאות:

מודל (1,1,0) בלי מסנן קלמן:

עבור כל המודל NBIC= 15.894

עבור רק x NBIC= 15.957

עבור רק y NBIC= 15.83

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ARIMA Model Parameters** | | | | | | | | |
|  | | | | | Estimate | SE | t | Sig. |
| X-Model\_1 | X | No Transformation | Constant | | 83.637 | 363.098 | .230 | .818 |
| AR | Lag 1 | -.493 | .061 | -8.102 | .000 |
| Difference | | 1 |  |  |  |
| NOP | No Transformation | Numerator | Lag 0 | -.951 | 7.104 | -.134 | .894 |
| NOC | No Transformation | Numerator | Lag 0 | -.565 | 2.620 | -.216 | .829 |
| t | No Transformation | Numerator | Lag 0 | -.061 | 2.128 | -.029 | .977 |
| Y-Model\_2 | Y | No Transformation | Constant | | -223.282 | 351.203 | -.636 | .526 |
| AR | Lag 1 | -.446 | .062 | -7.152 | .000 |
| Difference | | 1 |  |  |  |
| NOP | No Transformation | Numerator | Lag 0 | 8.917 | 6.857 | 1.301 | .195 |
| NOC | No Transformation | Numerator | Lag 0 | -.586 | 2.531 | -.232 | .817 |
| t | No Transformation | Numerator | Lag 0 | -.056 | 2.061 | -.027 | .978 |

ניתן לראות כי ישנה תלות רק במצב קודם ואין תלות במשתנים מספר תקשורות, מספר מקומות וזמן. בנוסף ניתן לראות כי אין חותך למודל.

מודל (0,1,1) בלי מסנן קלמן:

עבור כל המודל NBIC= 15.64

עבור רק x NBIC= 15.72

עבור רק y NBIC= 15.559

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ARIMA Model Parameters** | | | | | | | | |
|  | | | | | Estimate | SE | t | Sig. |
| X-Model\_1 | X | No Transformation | Constant | | -3.095 | 103.586 | -.030 | .976 |
| Difference | | 1 |  |  |  |
| MA | Lag 1 | .895 | .038 | 23.580 | .000 |
| NOP | No Transformation | Numerator | Lag 0 | -1.976 | 2.426 | -.815 | .416 |
| NOC | No Transformation | Numerator | Lag 0 | .863 | .962 | .897 | .371 |
| t | No Transformation | Numerator | Lag 0 | -.176 | .364 | -.482 | .630 |
| Y-Model\_2 | Y | No Transformation | Constant | | -80.217 | 60.518 | -1.326 | .186 |
| Difference | | 1 |  |  |  |
| MA | Lag 1 | .997 | .136 | 7.352 | .000 |
| NOP | No Transformation | Numerator | Lag 0 | 2.154 | 1.754 | 1.228 | .221 |
| NOC | No Transformation | Numerator | Lag 0 | .006 | .496 | .012 | .991 |
| t | No Transformation | Numerator | Lag 0 | .091 | .119 | .763 | .446 |

ניתן לראות כי ישנה תלות בהחלקה האקספוננצילית ואין תלות במשתנים מספר תקשורות, מספר מקומות וזמן. בנוסף ניתן לראות כי אין חותך למודל.

מודל (0,1,0) בלי מסנן קלמן:

עבור כל המודל NBIC= 16.111

עבור רק x NBIC= 16.204

עבור רק y NBIC= 16.018

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ARIMA Model Parameters** | | | | | | | | |
|  | | | | | Estimate | SE | t | Sig. |
| X-Model\_1 | X | No Transformation | Constant | | 36.616 | 583.692 | .063 | .950 |
| Difference | | 1 |  |  |  |
| NOP | No Transformation | Numerator | Lag 0 | 1.736 | 10.523 | .165 | .869 |
| NOC | No Transformation | Numerator | Lag 0 | -1.173 | 4.151 | -.283 | .778 |
| t | No Transformation | Numerator | Lag 0 | .166 | 3.603 | .046 | .963 |
| Y-Model\_2 | Y | No Transformation | Constant | | -422.509 | 531.782 | -.795 | .428 |
| Difference | | 1 |  |  |  |
| NOP | No Transformation | Numerator | Lag 0 | 19.011 | 9.588 | 1.983 | .049 |
| NOC | No Transformation | Numerator | Lag 0 | -1.682 | 3.781 | -.445 | .657 |
| t | No Transformation | Numerator | Lag 0 | -.308 | 3.283 | -.094 | .925 |

ניתן לראות כי אין תלות בין המשתנים אלא רק בין מספר מקומות ומשתנה קורדינטת y.

מודל (1,1,0) עם מסנן קלמן:

עבור כל המודל NBIC= 14.103

עבור רק x NBIC= 14.16

עבור רק y NBIC= 14.045

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ARIMA Model Parameters** | | | | | | | | |
|  | | | | | Estimate | SE | t | Sig. |
| X-Model\_1 | X | No Transformation | Constant | | -199.397 | 237.452 | -.840 | .402 |
| AR | Lag 1 | .156 | .064 | 2.445 | .015 |
| Difference | | 1 |  |  |  |
| NOP | No Transformation | Numerator | Lag 0 | 1.542 | 3.583 | .430 | .667 |
| NOC | No Transformation | Numerator | Lag 0 | 1.330 | 1.528 | .871 | .385 |
| t | No Transformation | Numerator | Lag 0 | -.010 | 1.195 | -.009 | .993 |
| Y-Model\_2 | Y | No Transformation | Constant | | -315.246 | 265.959 | -1.185 | .237 |
| AR | Lag 1 | .345 | .061 | 5.687 | .000 |
| Difference | | 1 |  |  |  |
| NOP | No Transformation | Numerator | Lag 0 | 9.229 | 3.550 | 2.600 | .010 |
| NOC | No Transformation | Numerator | Lag 0 | -.253 | 1.604 | -.158 | .875 |
| t | No Transformation | Numerator | Lag 0 | .275 | 1.448 | .190 | .850 |

ניתן לראות כי ישנה תלות במצב קודם ומשתנה מספר מקומות משפיע על קורדינטת y. אין תלות במשתנים מספר תקשורות וזמן. בנוסף ניתן לראות כי אין חותך למודל.

מודל (0,1,1) עם מסנן קלמן:

עבור כל המודל NBIC= 14.095

עבור רק x NBIC= 14.15

עבור רק y NBIC= 14.04

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ARIMA Model Parameters** | | | | | | | | |
|  | | | | | Estimate | SE | t | Sig. |
| X-Model\_1 | X | No Transformation | Constant | | -188.628 | 241.056 | -.783 | .435 |
| Difference | | 1 |  |  |  |
| MA | Lag 1 | -.217 | .063 | -3.452 | .001 |
| NOP | No Transformation | Numerator | Lag 0 | 1.603 | 3.591 | .446 | .656 |
| NOC | No Transformation | Numerator | Lag 0 | 1.186 | 1.550 | .765 | .445 |
| t | No Transformation | Numerator | Lag 0 | .004 | 1.222 | .003 | .997 |
| Y-Model\_2 | Y | No Transformation | Constant | | -290.052 | 244.990 | -1.184 | .238 |
| Difference | | 1 |  |  |  |
| MA | Lag 1 | -.357 | .060 | -5.928 | .000 |
| NOP | No Transformation | Numerator | Lag 0 | 8.925 | 3.444 | 2.591 | .010 |
| NOC | No Transformation | Numerator | Lag 0 | -.437 | 1.544 | -.283 | .777 |
| t | No Transformation | Numerator | Lag 0 | .298 | 1.288 | .231 | .818 |

ניתן לראות כי ישנה תלות בהחלקה האקספוננצילית ומשתנה מספר מקומות משפיע על קורדינטת y. אין תלות במשתנים מספר תקשורות וזמן. בנוסף ניתן לראות כי אין חותך למודל.

(0,1,0) עם מסנן קלמן:

עבור כל המודל NBIC= 14.151

עבור רק x NBIC= 14.158

עבור רק y NBIC= 14.145

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ARIMA Model Parameters** | | | | | | | | |
|  | | | | | Estimate | SE | t | Sig. |
| X-Model\_1 | X | No Transformation | Constant | | -193.468 | 212.895 | -.909 | .364 |
| Difference | | 1 |  |  |  |
| NOP | No Transformation | Numerator | Lag 0 | .630 | 3.400 | .185 | .853 |
| NOC | No Transformation | Numerator | Lag 0 | 1.592 | 1.398 | 1.139 | .256 |
| t | No Transformation | Numerator | Lag 0 | -.026 | 1.022 | -.025 | .980 |
| Y-Model\_2 | Y | No Transformation | Constant | | -289.286 | 211.577 | -1.367 | .173 |
| Difference | | 1 |  |  |  |
| NOP | No Transformation | Numerator | Lag 0 | 7.213 | 3.379 | 2.134 | .034 |
| NOC | No Transformation | Numerator | Lag 0 | .188 | 1.389 | .136 | .892 |
| t | No Transformation | Numerator | Lag 0 | .257 | 1.015 | .253 | .801 |

ניתן לראות כי אין תלות בין המשתנים אלא רק בין מספר מקומות ומשתנה קורדינטת y.

מודל (0,1,1) עם מסנן קלמן הינו המודל שמתאים **לנתונים בצורה הטובה ביותר** עם NBIC= 14.095

# תוצאות רגרסיה רב משתנית

את הרגרסיה ביצעתי לכל משתמש בנפרד. בסדר הפעולות להלן:

בהתחלה בדקתי האם קיים טרנד בתנועת המשתמש. בדקתי את זה כתלות בזמן, בזמן בריבוע ובזמן בשלישית. אני מצפה לראות תלות סטציונארית או פולינומילית מדרגה שלישית מפני שתנועת המשתמש לא אמורה להיות בכיוון אחד לאורך כל זמן המדידה או בשני כיוונים לאורך כל המדידה. לדעתי התנועה אמורה להיות קבועה או לכל היותר לא להתנהג בצורה קבועה (כפי הידוע פולינום מדרגה שלישית יוכל לאפיין כל תנועה, לכן מספיק לבדוק עד דרגה שלישית).

לאחר מכן השתמשתי בפונקציות הטריגונומטרית עבור כל ציר שקיבלתי מהFFT על מנת לבדוק תנועה מחזורית. בניתי שישה משתנים קוסינוסיאלים מהפונקציה הטריגונומטרית (שלושה עבור כל ציר) ובדקתי האם הם מובהקים.

לאחר מכן את הרכיבים שיצאו מובהקים (מהטרנד והמחזוריות) הוספתי לשתי המשתנים, מספר תקשורות בשעה ומספר מקומות בשעה ובניתי מודל רגרסיה.

* לא בדקתי lag-ים מפני שלא ניתן לבצע רגרסיה רב משתנית לlag-ים (בכל אופן לא מצאתי בספרות).

להלן התוצאות:

משתמש 1:

בשני המקרים (ללא מסנן ועם מסנן) קיימת תנועה פולינומילית מסדר 3.

המודל הינו:

ואלו התוצאות:

ללא מסנן קלמן

|  |  |  |  |
| --- | --- | --- | --- |
| P- value | y | x | Beta |
| 0 | 173435.3637 | 146113.9205 | Const |
| 0.0018 | -9.6330 | -69.5848 | t |
| 0.0011 | 0.0716 | 0.6769 | t^2 |
| 0.0020 | -0.0001 | -0.0017 | t^3 |

עם מסנן קלמן

|  |  |  |  |
| --- | --- | --- | --- |
| P- Value | y | x | Beta |
| 0 | 174169.1603 | 147123.8152 | Const |
| 0.0001 | -30.5046 | -97.2178 | t |
| 0.0001 | 0.2370 | 0.8841 | t^2 |
| 0.0001 | -0.0005 | -0.0021 | t^3 |

מבחינה התנועה המחזוריות המודל עבור משתמש 1 ללא מסנן קלמן הוא:

מבחינה התנועה המחזוריות המודל עבור משתמש 1 עם מסנן קלמן הוא:

ללא מסנן קלמן

|  |  |  |  |
| --- | --- | --- | --- |
| P- value | y | x | Beta |
| 0 | 173193.7823 | 144808.2457 | Const |
| 0.0009 | 151.1490 | 707.4030 | Cosx1 |
| 0.0145 | -273.1085 | 698.1121 | Cosx2 |
| 0.0036 | -567.4046 | 704.5271 | Cosx3 |
| 0.5905 | 165.8291 | 81.4481 | Cosy1 |
| 0.2432 | -306.0054 | 167.2130 | Cosy2 |
| 0.0738 | 82.3260 | 440.4742 | Cosy3 |

כאשר המודל הנבחר ברמת מובהקות של 5% הינו

עם מסנן קלמן

|  |  |  |  |
| --- | --- | --- | --- |
| P- Value | y | x | Beta |
| 0 | 173216.6740 | 144878.6510 | Const |
| 0 | 64.5530 | 840.6060 | Cosx1 |
| 0 | 427.5240 | 787.2520 | Cosx2 |
| 1 | -8246.0910 | -5422.0100 | Cosx3 |
| 0 | 946.7490 | 391.0500 | Cosy1 |
| 0 | 605.5490 | 606.1690 | Cosy2 |
| 1 | 9243.3600 | 6920.1250 | Cosy3 |

כאשר המודל הנבחר ברמת מובהקות של 5% הינו

**בשלב הבא** בדקתי מודל אשר מתבסס על הטרנד והמחזוריות והוספתי שתי משתנים של מספר תקשורות בשעה ומספר נ.צ בשעה.

המודל ללא מסנן קלמן הינו

|  |  |  |  |
| --- | --- | --- | --- |
| P- value | y | x | Beta |
| 0 | 171310.5220 | 147441.9250 | Const |
| 0.0380 | 214.1390 | -194.7200 | t |
| 0.0400 | -2.2700 | 1.9700 | t^2 |
| 0.0390 | 0.0060 | -0.0050 | t^3 |
| 0.0690 | 3991.3110 | -2094.8230 | Cosx1 |
| 0.0040 | -305.4140 | 822.2380 | Cosx2 |
| 0.0080 | -563.2640 | 691.1720 | Cosx3 |
| 0.6100 | 1.7740 | -6.7590 | Number of places |
| 0.3800 | -2.2210 | 3.6690 | Number of connects |

המודל הסופי ללא מסנן קלמן הינו

כאשר R^2 הוא 0.140 =x 0.0520 =y

כאשר Adjusted R^2 הוא 0.1110=x 0.0200 =y

ניתן לראות כי כל שינוי יחידה בt גורם לשינוי שלילי בגודל 194.72 מטר בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי חיובי של 214.139 מטר.

ניתן לראות כי כל שינוי יחידה בt^2 גורם לשינוי חיובי בגודל 1.97 מטר בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי שלילי של 2.27 מטר.

ניתן לראות כי כל שינוי יחידה בt^3 גורם לשינוי שלילי בגודל 0.005 מטר בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי חיובי של 0.006 מטר.

ניתן לראות כי כל שינוי יחידה בCosx1 גורם לשינוי שלילי בגודל 2094.823 מטר בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי חיובי של 3991.311 מטר.

ניתן לראות כי כל שינוי יחידה בCosx2 גורם לשינוי חיובי בגודל 822.238 מטר בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי שלילי של 305.414 מטר.

המודל עם מסנן קלמן הינו

|  |  |  |  |
| --- | --- | --- | --- |
| P- Value | y | x | Beta |
| 0 | 146395.5930 | 175035.2990 | Const |
| 0 | -86.0470 | -23.1870 | t |
| 0 | 0.8000 | 0.1780 | t^2 |
| 0 | -0.0020 | 0 | t^3 |
| 0 | 829.6580 | 87.4710 | Cosx1 |
| 0 | 750.0170 | 455.9280 | Cosx2 |
| 0 | 288.6980 | 975.4860 | Cosy1 |
| 0 | 582.1690 | 846.2820 | Cosy2 |
| 0.0160 | 13.7890 | 1.2230 | Number of places |
| 0 | -1.2360 | -12.3930 | Number of connects |

המודל הסופי עם מסנן קלמן הינו

כאשר R^2 הוא 0.3580 =x 0.2540 =y

כאשר Adjusted R^2 הוא 0.3330 =x 0.2260 =y

ניתן לראות כי כל שינוי יחידה בt גורם לשינוי שלילי בגודל 23.187 מטר בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי שלילי של 86.047 מטר.

ניתן לראות כי כל שינוי יחידה בt^2 גורם לשינוי חיובי בגודל 0.178 מטר בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי חיובי של 0.8 מטר.

ניתן לראות כי כל שינוי יחידה בt^3 לא גורם לשינוי בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי שלילי של 0.002 מטר.

ניתן לראות כי כל שינוי יחידה בCosx1 גורם לשינוי חיובי בגודל 87.471 מטר בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי חיובי של 829.658 מטר.

ניתן לראות כי כל שינוי יחידה בCosx2 גורם לשינוי חיובי בגודל 455.928 מטר בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי שלילי של 750.017 מטר.

ניתן לראות כי כל שינוי יחידה בCosy1 גורם לשינוי חיובי בגודל 975.486 מטר בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי חיובי של 288.698 מטר.

ניתן לראות כי כל שינוי יחידה בCosy2 גורם לשינוי חיובי בגודל 846.282 מטר בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי חיובי של 582.169 מטר.

ניתן לראות כי כל שינוי יחידה במספר מקומות גורם לשינוי חיובי בגודל 1.223 מטר בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי חיובי של 13.789 מטר.

ניתן לראות כי כל שינוי יחידה במספר תקשורות גורם לשינוי שלילי בגודל 12.393 מטר בקורדינטת הרוחב ואליו בקורדינטת האורך שינוי שלילי של 1.236 מטר

כפי שניתן לראות כי קיימת מגמתיות בתנועת משתמש זה בנוסף (למרות שיכולנו לבדוק עוד מחזורים) ניתן לראות כי למשתמש זה עם מסנן קלמן יש מחזוריות בתנועה כל 9 שעות בערך, כל 27 שעות בערך וכל 22 שעות בערך. במחקר עתידי ניתן לבדוק עוד פיקים משמעותיים מהFFT, על מנת למצוא מחזוריות משמעותית בחיי האדם. זאת אומרת שכרגע נמצאה למשתמש זה מחזוריות בערך כל יום (24 שעות) וכל חצי יום (קרוב ל12 שעות).

בנוסף ניתן לראות כי בין הנ.צ. ים למשתנה מספר מקומות יש תלות לינארית שלילית, ואילו בין מספר תקשורות לנ.צ. ים יש תלות לינארית חיובית.

**בשלב האחרון** בדקתי מודל אשר מתבסס על הטרנד והמחזוריות והוספתי שתי משתנים של מספר תקשורות בשעה ומספר נ.צ בשעה ואת האינטראקציה ביניהם.

המודל ללא מסנן קלמן הינו

|  |  |  |  |
| --- | --- | --- | --- |
| P- value | y | x | Beta |
| 0 | 172453.107 | 146492.059 | Const |
| 0.103 | 176.962 | -163.814 | t |
| 0.104 | -1.887 | 1.651 | t^2 |
| 0.101 | 0.005 | -0.004 | t^3 |
| 0.157 | 3371.993 | -1579.965 | Cosx1 |
| 0.005 | -280.217 | 801.291 | Cosx2 |
| 0.012 | -526.332 | 660.469 | Cosx3 |
| 0.09 | -31.46 | 20.87 | Number of places |
| 0.014 | -9.515 | 9.733 | Number of connects |
| 0.025 | 0.302 | -0.251 | NOP x NOC |

המודל הסופי ללא מסנן קלמן הינו

כאשר R^2 הוא 0.152=x 0.072=y

כאשר Adjusted R^2 הוא 0.12=x 0.037=y

ניתן לראות כי על פי Adjusted R^2 המודל השתפר לעומת המודל ללא האינטראקציה.

המודל עם מסנן קלמן הינו

|  |  |  |  |
| --- | --- | --- | --- |
| P- Value | y | x | Beta |
| 0 | 176053.649 | 146547.823 | Const |
| 0 | -26.338 | -86.518 | t |
| 0 | 0.205 | 0.804 | t^2 |
| 0 | 0 | -0.002 | t^3 |
| 0 | 153.01 | 839.456 | Cosx1 |
| 0 | 451.621 | 749.373 | Cosx2 |
| 0 | 919.149 | 280.277 | Cosy1 |
| 0 | 815.916 | 577.63 | Cosy2 |
| 0.001 | -43.248 | 7.142 | Number of places |
| 0 | -22.013 | -2.674 | Number of connects |
| 0.001 | 0.409 | 0.061 | NOP x NOC |

המודל הסופי עם מסנן קלמן הינו

כאשר R^2 הוא 0.359=x 0.294=y

כאשר Adjusted R^2 הוא 0.331=x 0.265=y

ניתן לראות כי על פי Adjusted R^2 המודל פחות טוב לעומת המודל ללא האינטראקציה.

המודל הנבחר הינו המודל ללא האינטראקציה ולכן לא הסברתי כיצד הbeta –ות השונות משפיעות על השינוי במיקום.

לסיכום

כפי שניתן לראות Adjusted R^2 הינו גבוה יותר עם מסנן קלמן, זאת ניתן להסביר כי מסנן קלמן מבצע החלקה (שאולי לא תמיד מסתדרת עם ההיגיון) וכך בעצם מונע רעשים. על כן מדדי טיב התאמת המודל טובים יותר ואף ישנה השפעה של המשתנים (מספר תקשורות ומספר מקומות) במודל עם מסנן קלמן.