תרגיל 2 | IML (67329)

שם: רונאל חרדים | ת"ז:208917641

חלק I

תיאורטי:

שאלה 1

נוכיח כי $Ker(\mathbf{X}) = Ker\left(\mathbf{X}^\top \mathbf{X}\right)$**:**
יהי $u \in Ker(\mathbf{X})$ שונה מ 0 אזי מתקיים:

$$Xu = 0 \Rightarrow \mathbf{X}^\top \mathbf{X}u = 0$$

ולכן מתקיים גם כי $u \in Ker\left(\mathbf{X}^\top \mathbf{X}\right)$ ולכן מתקיים:

$$0 = \mathbf{X}^\top \mathbf{X}u = u^\top \mathbf{X}^\top \mathbf{X}u = \|\mathbf{X}u\|^2$$

מתכונות הנורמה נובע כי הנורמה שווה ל 0 אמ"מ הווקטור הוא ווקטור ה 0 ומיכוון שהגדרנו שהוא שונה מ 0 אזי הטענה מתקיימת כנדרש.

שאלה 2

נוכיח כי עבור מטריצה ריבועית מתקיים $A : Im\left(A^\top\right) = Ker(A)^\perp$**:**
יהי $b \in Im\left(A^\top\right)$ אזי קיים $x \in R^m$ כך ש $A^\top x = b$ 'יהי $w \in Ker\left(A^\top\right)$:

$$A^\top w = 0 \Rightarrow 0 = \left\langle A^\top w, x\right\rangle = \langle w, Ax\rangle = \langle w, b\rangle$$

לכן $Im(A) \subset Ker\left(A^\top\right)^\perp$.
בכיוון השני נניח כי $b \notin Im(A)$ ונראה כי $b \notin Ker\left(A^\top\right)$:
מספיק למצוא ווקטור $c \in Ker\left(A^\top\right)$ כך ש $\langle b, c\rangle \neq 0$, מההנחה ש $b \notin Im(A)$ חייב להיות ש b הוא רכיב ב $Im\left(A^\top\right)$.
יהי $c \in Im(A)^\perp$ כך ש $\langle b, c\rangle \neq 0$.

כעת מכיוון ש $c \in Im(A)^{\perp}$ אנו יודעים כי הוא אורתוגונלי לכל ווקטור בתמונה של A בייחוד $\left\langle c, AA^{\top}c \right\rangle = 0$ לכן:

$$\left\| A^{\top}c \right\|^{2} = \left\langle A^{\top}c, A^{\top}c \right\rangle = \left\langle c, AA^{\top}c \right\rangle = 0$$

לכן: $A^{\top}c = 0 \Rightarrow c \in Ker\left(A^{\top}\right)$כנדרש.

שאלה 3

מכיוון ש X איננה הפיכה למערכת המשוואות אין פתרון או שיש לה אינסוף פתרונות. בנוסף אנו יודעים כי יש פתרון יחיד אמ"מ ט נמצא בתמונה של X. לכן נובע מכך שלמערכת יש אינסוף פתרונות ומשאלה 2:

$$y \in Im(X) \iff y \in Ker(\mathbf{X})^{\perp} y \perp \iff Ker(\mathbf{X})^{\perp}$$

שאלה 4

אנו מניחים כי $X^{T}X$ איניה הפיכה. משאלה קודמת אנו יודעים כי למערכת יש אינסוף פתרונות אמ"מ $\perp$ $\mathbf{X}^{\top}\mathbf{y}$ $Ker\left(\mathbf{X}^{\top}\mathbf{X}\right)$ בנוסף הוכחנו כי $Ker(\mathbf{X}) = Ker\left(\mathbf{X}^{\top}\mathbf{X}\right)$.
לכן מספיק להראות כי $\mathbf{X}^{\top}\mathbf{y} \perp Ker\left(\mathbf{X}\right)$. לכן אם $u \in Ker(X)$ אזי $\left\langle u, \mathbf{X}^{\top}\mathbf{y} \right\rangle = \langle \mathbf{X}u, \mathbf{y} \rangle = 0$. א"כ הוכחנו כי לצורה הנורמלית חייב להיות פתרון יחודי אם $\mathbf{X}^{\top}\mathbf{X}$ הפיכה, או אינסוף פתרונות.

שאלה 5

(א)

נראה כי P סימטרית, ומתקיים $P^{T} = P$:
p היא סכום של מטריצות סימטריות ולכן היא סימטרית כנדרש.

(ב)

נוכיח כי הע"ע של P הם 0,1 וגם ווקטורי הבסיס האורתונורמלים הם ו"ע של ע"ע 0:

$$Pv_{j} = \sum_{i=1}^{k} v_{i}v_{i}^{T}v_{j} = \sum_{i=1}^{k} v_{i}\delta_{ij} = v_{j}$$

(ג)

נראה כי מתקיים לכל $v \in V$ **מתקיים** $P\mathbf{v} = \mathbf{v}$**:**

$$Pv_j = \sum_{i=1}^{k} v_i v_i^{\top} v_j = \sum_{i=1}^{k} v_i \delta_{ij} = v_j$$

(ד)

נוכיח כי $P^2 = P$:

$$P^2 = UDU^TUDU^T = UDDU^T = UDU^T = P$$

(ה)

נוכיח כי $(I-P)P = 0$**:**

$$(I-P)P = \left(UIU^{\top} - UDU^{\top}\right) UDU^{\top} = U(I-D)U^{\top}UDU^{\top} = U(I-D)DU^{\top} = U(D-D)U^{\top} = 0$$

שאלה 6

הוכחה:

ראשית נוכיח כי:

$$\left(\mathbf{X}^{\top}\mathbf{X}\right)^{-1} = UD^{-1}U^{\top}, \ \textbf{where } D = \Sigma\Sigma^{\top}$$

ולכן:

$$\mathbf{X}^{\top}\mathbf{X} = \left(U\Sigma V^{\top}\right)\left(U\Sigma V^{\top}\right)^{\top} = U\Sigma V^{\top}V\Sigma^{\top}U^{\top} = U\Sigma\Sigma^{\top}U^{\top} = UDU^{\top}$$

$$\mathbf{X}^{\top}\mathbf{X}\left(UD^{-1}U^{\top}\right) = UDU^{\top}UD^{-1}U^{\top} = UDD^{-1}U^{\top} = UU^{\top} = I$$

כעת נראה כי $\left(\mathbf{X}\mathbf{X}^{\top}\right)^{-1}\mathbf{X}^{\top} = \mathbf{X}^{\top\dagger}$**:**

$$\left(\mathbf{X}^{\top}\mathbf{X}\right)^{-1}\mathbf{X} = UD^{-1}U^{\top}U\Sigma V^{\top} = UD^{-1}\Sigma V^{\top} = U\Sigma^{\dagger\top}V^{\top} = \mathbf{X}^{\top\dagger}$$

שאלה 7

פתרון:

הדרגה של X שווה ל $span\left\{x_1, ...x_m\right\}$ מפשט פירוק SVD נובע כי הדרגה של X שווה לדרגה של X^TX.

לכן $\mathbf{x}_1,\ldots,\mathbf{x}_m span\mathbb{R}^d$ אמ"מ דרגה של X^TX שוןה ל d. מכיוון ש X^TX היא מטריצה מהצורה $d\cdot d$ היא הפיכה אמ"מ הדרגה היא d.

שאלה 8

יהי $\mathbf{X} = U\Sigma V^\top$ הפירוק של X, ונגדיר את הדרגה של X להיות r. נכתוב באופס הבא:

$$V = \left[\begin{array}{cc} V_1 & V_2 \end{array}\right], \quad U = \left[\begin{array}{cc} U_1 & U_2 \end{array}\right], \quad \Sigma = \left[\begin{array}{cc} \Sigma_1 & 0 \\ 0 & 0 \end{array}\right]$$

כאשר V_1, U_1 הן r השורות הראשונות ו V_2, U_2 שאר השורות של V, U ו Σ_1 היא מטריצה דיאגונלית עם ערכים שונים מ 0.

בהינתן w נגדיר

$$b = U^\top w, \;\; b_1 = U_1^\top w, b_2 = U_2^\top w \quad \Rightarrow \;\; b = \left[\begin{array}{c} b_1 \\ b_2 \end{array}\right]$$

U היא מטריצה אורתוגונלית ולכן איזומטרית, לכן אנו מקבלים $||b|| = ||w||$ כעת אנו יכולים לשכתב את הבעיה שלנו ולהראות שיש ל b נורמה מינימלית, מכיוון שגם V איזומטרית, נקבל:

$$\left\|y - X^\top w\right\|^2 = \left\|y - V\Sigma U^\top w\right\|^2 = \left\|V\left(V^\top y - \Sigma b\right)\right\|^2 = \left\|V^\top y - \Sigma b\right\|^2 =$$

$$\left\|\left[\begin{array}{cc} V_1 & V_2 \end{array}\right]^\top y - \left[\begin{array}{cc} \Sigma_1 & 0 \\ 0 & 0 \end{array}\right]\left[\begin{array}{c} b_1 \\ b_2 \end{array}\right]\right\|^2 = \left\|V_1^\top y - \Sigma_1 b_1\right\|^2 + \left\|V_2^\top y\right\|^2$$

לכן, בכדי למצוא פתרון מינימלי $||y - Xw||$ נפתור את

$$\Sigma_1 b_1 - V_1^\top y = 0 \Longleftrightarrow \Sigma_1 b_1 = V_1^\top y \Longleftrightarrow b_1 = \Sigma_1^{-1} V_1^\top y.$$

נמנמם את bכך:

$$b_1 = \Sigma_1^{-1} V_1^\top y \quad \text{and} \quad b_2 = 0$$

אנו יודעים ש $\widehat{\mathbf{w}} = \mathbf{X}^{\top\dagger}\mathbf{y}$ ולכן:

$$U_1^\top \hat{\mathbf{w}} = U_1^\top X^{\top\dagger}\mathbf{y} = U_1^\top U_1 \Sigma_1^{-1} V_1^\top y = \Sigma_1^{-1} V_1^\top y$$

וגם

$$U_2^\top \hat{\mathbf{w}} = U_2^\top X^{\top\dagger}\mathbf{y} = U_1^\top U_1 \Sigma_1^{-1} V_1^\top y = 0$$

לכל פתרון אחר $\bar{w}$אנו מקבלים כי התנאי הראשון חייב להתקיים אך התנאי השני לא חייב להתקיים בהכרח לכן $\|\overline{\mathbf{W}}\| \leq \|\widehat{\mathbf{w}}\|$ כנדרש.

חלק II

$LinearRegression$:

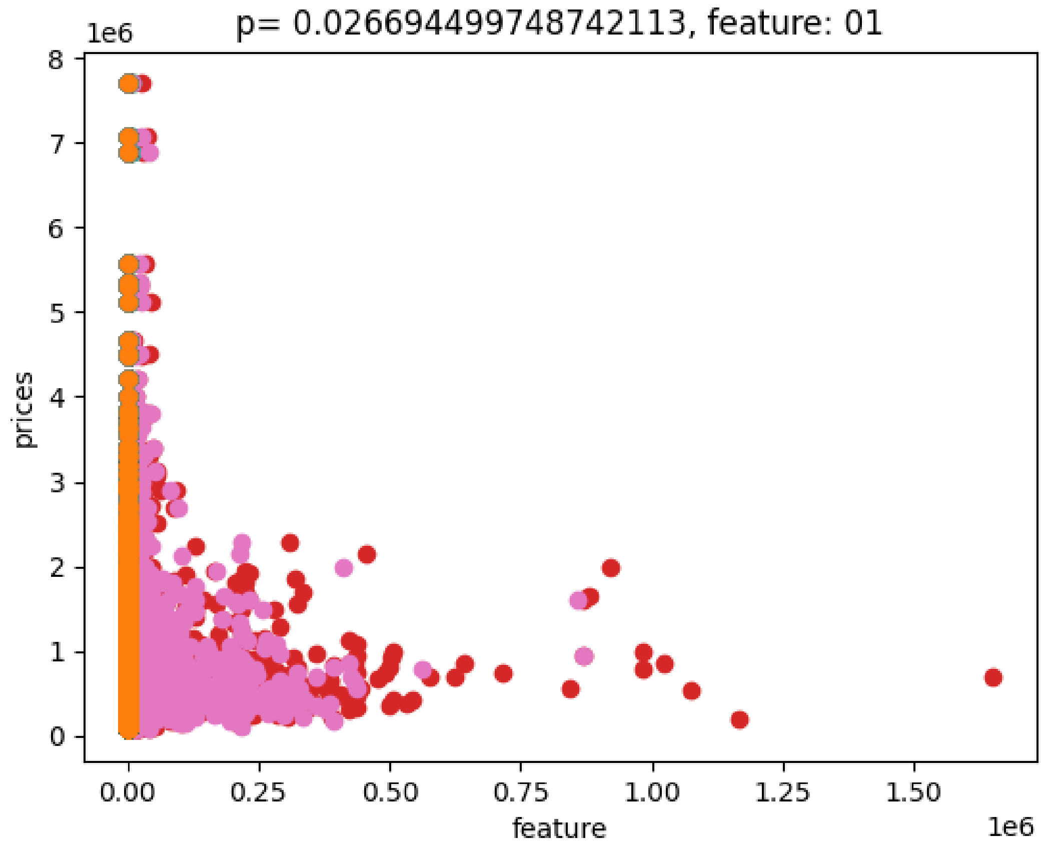
שאלה 1

התכונות שהחלטתי לא לשמור הן: $'id', date,' price',' zipcode'$ משום שהן לא מנבאות את מחיר הדירה.
תכונות קטגוריות כגון מיקוד ותאריך החלפתי למשתנה בוליאני.
בדקתי אם קיימים ערכים לא חוקיים כגון ערכי $null$ בת"ז ובתאריך, בנוסף בדקתי כי שאר הערכים שצריכים להופיע חיוביים כגון מחיר וגודל במ"ר יהיו גדולים מ 0.
לגבי שאר הערכים בעלי טווח מסויים דאגתי שהם יקיימו את התנאים וישארו בטווח.

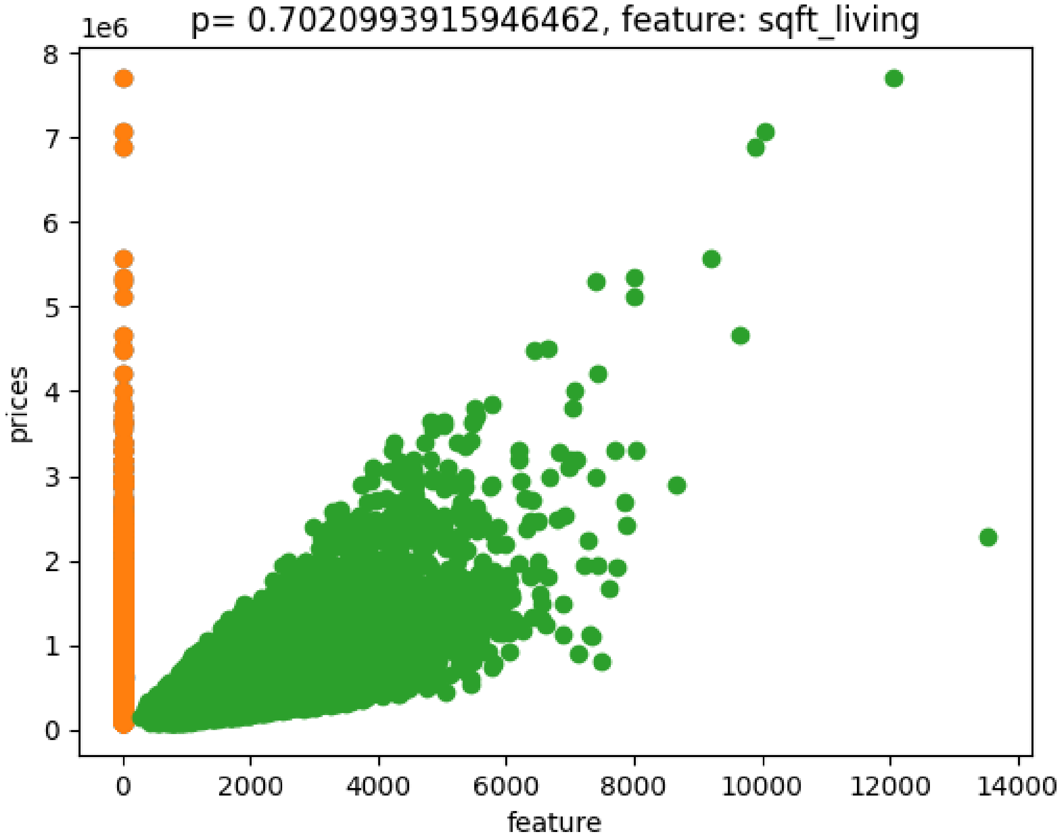
שאלה 2

נבחר שתי תכונות ונסיק את הדרוש:

הגרף הראשון הוא גרף של היום בחודש שבו נמכרה הדירה - הראשון לחודש:



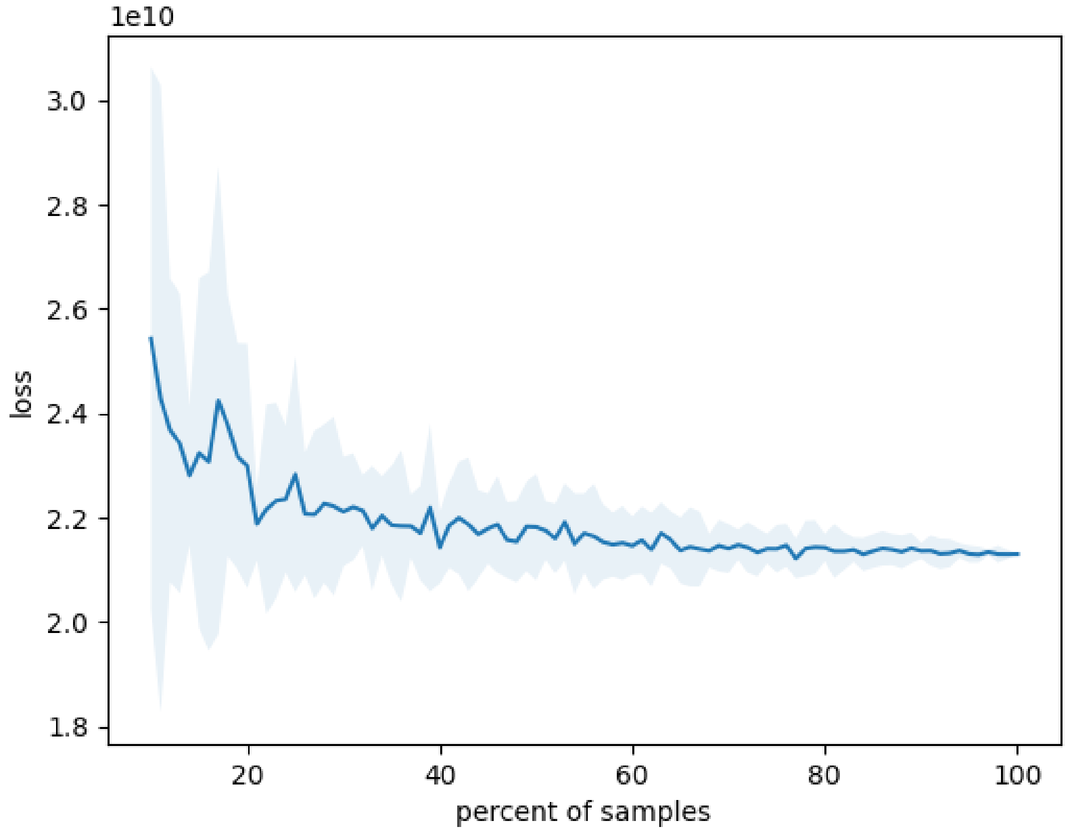
הגרף השני מייצג את גודל הבית במ"ר:



נשים לב כי הפרמטר p של התכונה הראשונה שואף ל 0 משמע שאין קורלציה בין X ל y, אך בגרף השני הפרמטר p שואף ל 1 משמע שככל עליה ב X יש עליה ב y.
בנוסף: ניתן לראות מהגרף השני כי יש קורלציה בין x, y ובגרף הראשון אין כל קורלציה בניהם.

שאלה 4

הגרף הוא הגרף הבא:



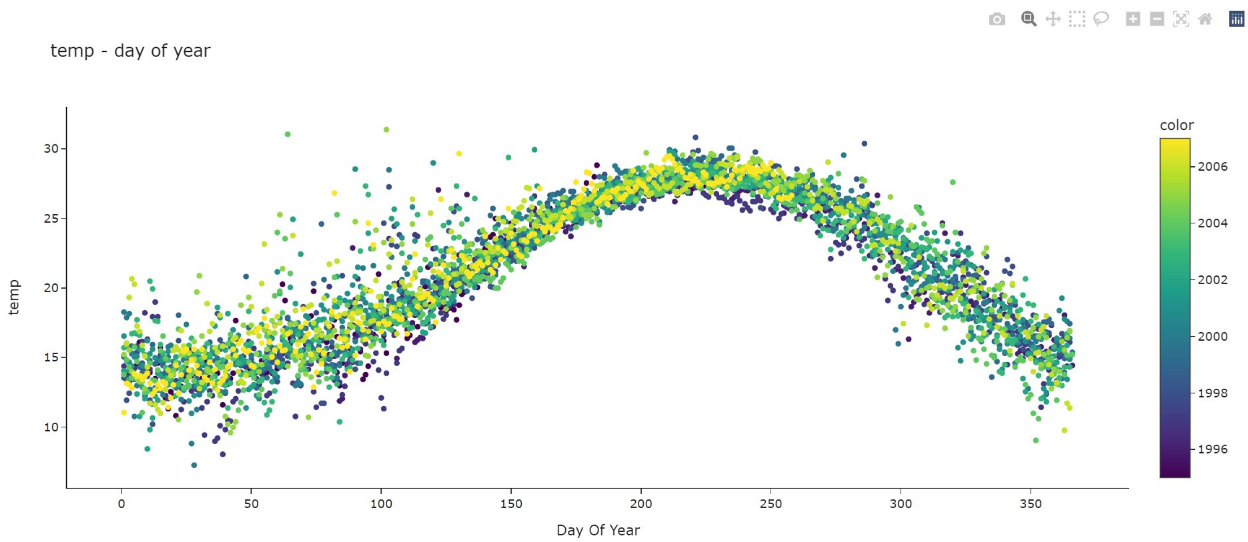
אנו רואים כי ככל שמספר הדגימות גדל ההפסד קטן ובנוסף ה std קטן גם כן ככל שמספר הדגימות עולה.

חלק III

$Polynomial\ Fitting$:

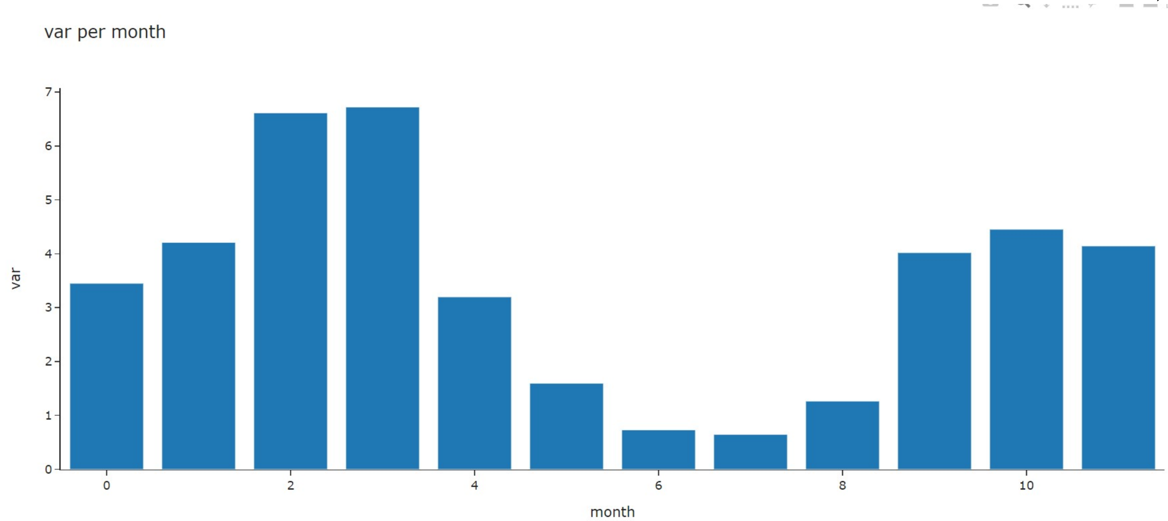
שאלה 2 :

הגרף לפי השנים הוא הגרף הבא:



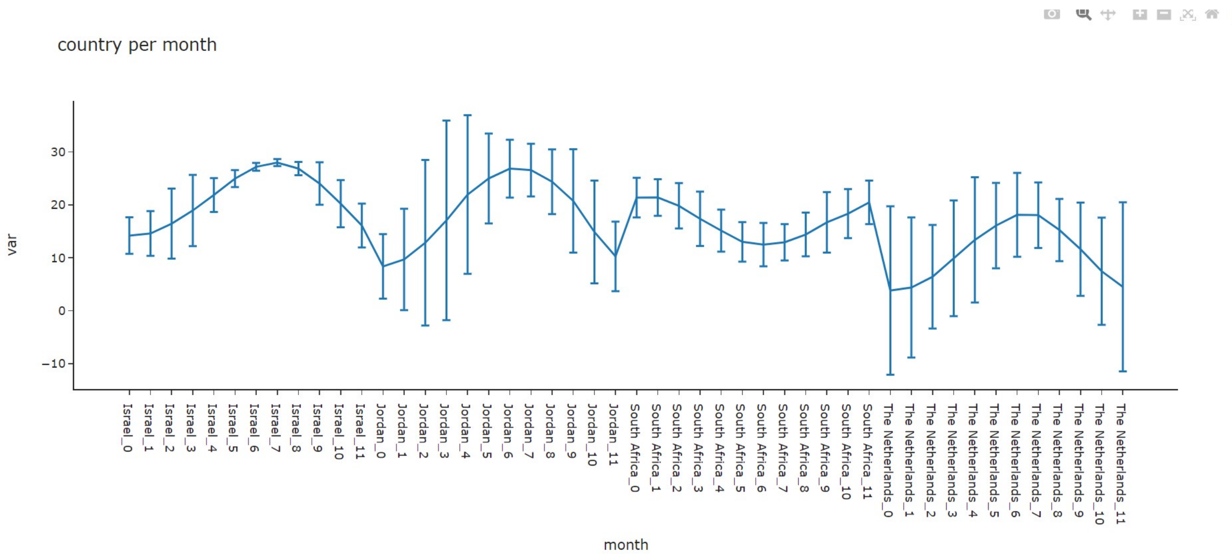
דרגת הפולינום שעשויה להתאים לגרף זה היא $k = 3$. ניתן לראות כי הפולינום לא מורכב מידיי.

הגרף המייצג את החודשים:



יש חודשים בהם הביצועים יהיו טובים יותר.
בחודשים בהם סטיית התקן גבוהה זאת אומרת שאנו רודפים אחרי הרעש, ולכן בחודשים אלו החיזוי לא יהיה מדוייק, לעומת זאת בחודשים בהם סטיית התקן נמוכה אנו נוכל לחזות טוב יותר.

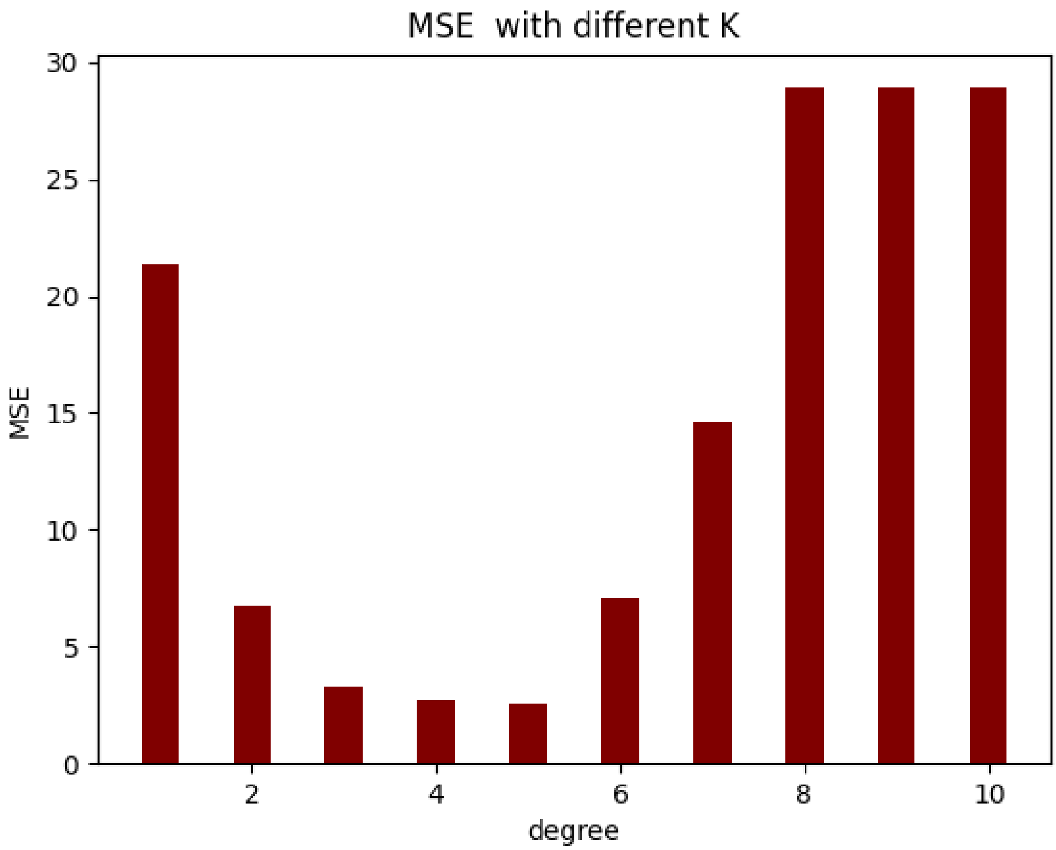
שאלה 3



אכן למעט דרום אפריקה כל המדינות חולקות דפוס דומה, המודל של ישראל עשוי לעבוד עבור הולנד וירדן.

שאלה 4

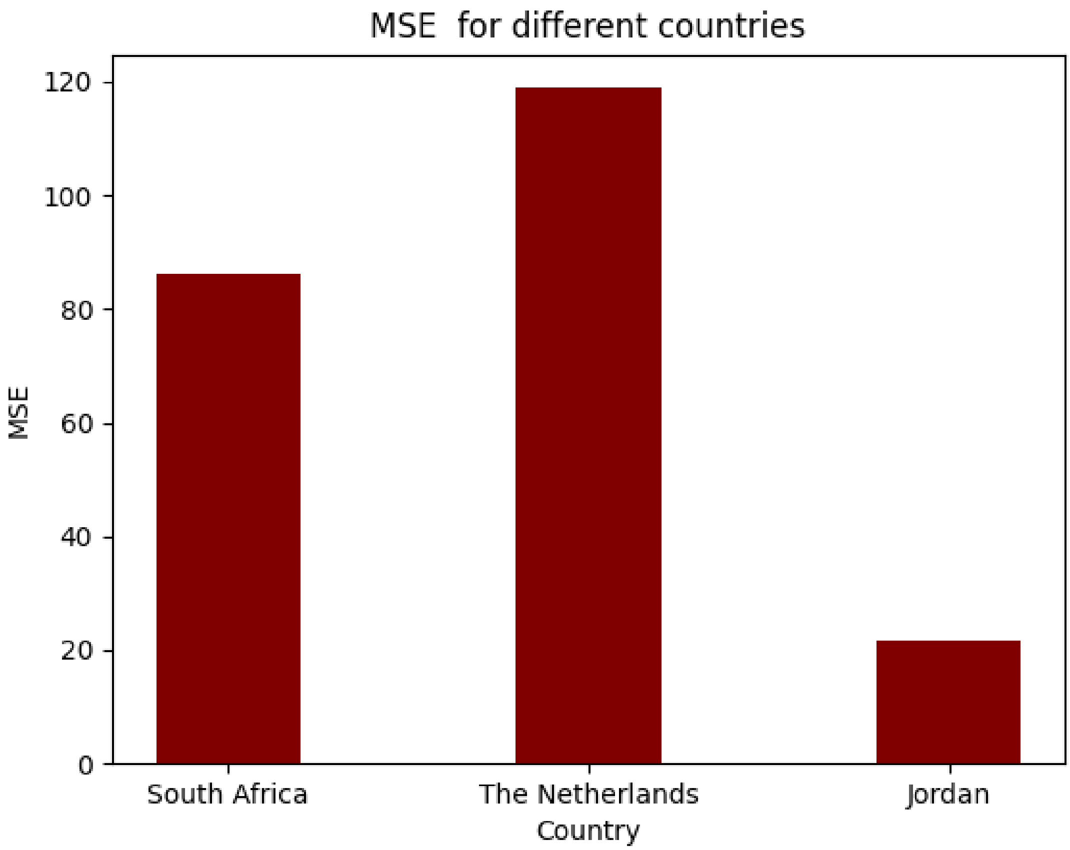
הגרף שקיבלנו הוא הגרף הבא:



ניתן לראות כי $k = 5$ היא דרגת הפולינום שתביא לנו את התחיזוי הטוב ביותר.
בנוסף מיקום העיר יוכל לתת לנו חיזוי טוב, יחד עם היום בשנה.

שאלה 5

הגרף שקיבלנו הוא:



נשים לב כי ירדן שהיא מדינה קרובה אלינו ומזג האוויר דומה אכן קיבלה $loss$ נמוך לפי חיזוי על ישראל, בעוד מדינות רחוקות קיבלו חיזוי גרוע. וזה מתקמפל עם שאלה 3 בה ראינו כי המודלים של ישראל וירדן דומים.