תרגיל 2

שם: תומר שני
ת"ז: 315315523

2.1)

1) נראה הכלה דו כיוונית, יהי $v\in\ker(X)$ כלומר $Xv=0$ אז $v\in\ker(X^TX)$ כי

$$X^TXv \underset{\text{תכונות כפל}}{=} X^T(Xv)=X^T\cdot 0=0$$

כלומר $v\in\ker(X^TX)$ כנדרש

לכן $\ker(X)\subseteq\ker(X^TX)$

בכיוון השני: יהי $0\neq v\in\ker(X^TX)$ כלומר $X^TXv=0 \;\Big/ \cdot v^T$ משמאל

$$v^TX^TXv=0$$

$$(Xv)^TXv=0$$

$$\|Xv\|^2=0 \underset{\text{תכונות נורמה}}{\Longleftrightarrow} Xv=0$$

כלומר $v\in\ker X$ כנדרש ■

2) יהי $A\in M_{n\times n}(\mathbb{R})$ אז $Im(A^t)=(\ker(A))^\perp$

יהי $v\in Im(A^t)$ כלומר קיים $u\in\mathbb{R}^d$ כך ש $A^Tu=v$

צ"ל $v\in\ker(A)^\perp$ כלומר $v\in\{x\in\mathbb{R}^d \mid \langle x,w\rangle=0\ \forall w\in\ker A\}$

יהי $w\in\ker A$ כלומר $Aw=0$ צ"ל $\langle v,w\rangle=0$

$$\langle v,w\rangle=\langle A^Tu,w\rangle=\langle u,Aw\rangle=\langle u,0\rangle=0$$

בכיוון השני:
יהי $v\in Im(A^t)^\perp$ כלומר לכל $u\in Im(A^t)$, $\langle v,u\rangle=0$ כי $u\in Im(A^t)$
יהי $w\in\mathbb{R}^n$ כך ש $A^tw=u$

$$\langle v,u\rangle=0 \Rightarrow \langle v,A^tw\rangle=0 \Rightarrow (A^tw)^tv=0$$

$$\Rightarrow w^tAv=0$$

$$\Rightarrow \langle w,Av\rangle=0$$

כלומר עבור $w=Av$ נקבל

$$\langle v,A^tw\rangle=0$$

$$\Rightarrow \langle v,A^tAv\rangle=0$$

$$\Rightarrow \langle Av,Av\rangle=0 \Rightarrow \|Av\|=0$$

כלומר $Av=0$ ולכן $v\in\ker(A)$ ולכן $Im(A^t)^\perp\subseteq\ker(A)$ ולפי תכונות ניצבים במרחב סופי נקבל

$$\ker(A)^\perp\subseteq(Im(A^t)^\perp)^\perp=Im(A^t)$$

■

(3 $y = Xw$ $X \in Mat_{n\times n}(\mathbb{R})$ לא הפיך

צ"ל $y \perp \ker(X^T)$ $\iff$ למערכת $y = Xw$ יש אינסוף פתרונות

הוכחה

$$\underset{\text{יש אינסוף פתרונות}}{y = Xw} \iff y \in \operatorname{Im} \underbrace{X}_{A^T} \underset{\text{סעיף 2}}{\iff} y \in \ker(\underbrace{X^T}_{A})^\perp$$

$$\iff y \perp \ker(X^T)$$

כנדרש ■

(4 $X^T X w = X^T y$

אם X^TX הפיכה אזי

$$X^TXw = X^Ty \Rightarrow w = (X^TX)^{-1}X^Ty$$

כלומר קיבלנו פתרון יחיד למשוואה.

אם X^TX אינה הפיכה

$$X^TXw = X^Ty$$

הוכחנו בסעיף קודם כי למשוואה יש אינסוף פתרונות אמ"מ $X^Ty \perp \ker((X^TX)^T)$

$$\Rightarrow X^Ty \perp \ker(X^TX)$$

מסעיף 1. זה שקול לכך ש

$$X^Ty \perp \ker(X)$$

ונשים לב ש $X^Ty \in \operatorname{Im}(X^T)$ מסעיף 2 נקבל כי $X^Ty \in \ker(X)^\perp$ כלומר קיבלנו

$$X^Ty \perp \ker(X)$$

ולכן המערכת אמ"מ נקבל כי למשוואה יש אינסוף פתרונות. ■

(2.2

ז) $V \subseteq \mathbb{R}^d$, $\dim(V)=k$ $v_1 \dots v_k$ בסיס א"נ של V

$$P = \sum_{i=1}^{k} v_i v_i^T = \sum_{i=1}^{k} v_i \otimes v_i$$

א) נבחין ש $\otimes$ לכל מיקום האינדקס ה ij $[v_l]_i \cdot [v_l]_j$ כלומר $[P]_{ij} = \sum_{l=1}^{k} [v_l]_i [v_l]_j$

ובאינדקס ה j,i יש לנו מיקום וריה $[v_l]_j \cdot [v_l]_i$ כלומר $[P]_{ji} = \sum_{l=1}^{k} [v_l]_j [v_l]_i$

$$= \sum_{l=1}^{k} [v_l]_i [v_l]_j = [P]_{ij}$$

חילופיות של כפל

לכן P סימטרית.

ב) יהי $\lambda \in \mathbb{R}$ ע"ע של P ויהי $u \in V$ ו"ע של λ כלומר

$$Pu = \lambda u$$

יהיו $\alpha_1 \dots \alpha_k \in \mathbb{R}$ כך ש $u = \sum_{i=1}^{k} \alpha_i v_i$ כאשר מהגדרה קיימים

$$Pu = \left(\sum_{i=1}^{k} v_i v_i^t\right) u = \sum_{i=1}^{k} v_i v_i^t u = \sum_{i=1}^{k} v_i v_i^t \sum_{l=1}^{k} \alpha_l v_l = \sum_{i=1}^{k} \sum_{l=1}^{k} \alpha_l v_i \left(v_i^t v_l\right)$$

$0 = v_i^t v_l$ $i \neq l$ כי בסיס א"נ

$$= \sum_{i=1}^{k} \alpha_i v_i = u$$

כלומר $\lambda = 1$.

כאשר יהי v_i נתון מ $v_1 \dots v_k$ נקבל כי $P v_i = \sum_{j=1}^{k} v_j \left(v_j^t v_i\right) = v_i$

כי $0 = v_j^t v_i$ $i \neq j$

ג) יהי $u \in V$ יהיו $\alpha_1 \dots \alpha_k \in \mathbb{R}$ כך ש $u = \sum_{i=1}^{k} \alpha_i v_i$

סעיף קודם

$$Pu = \sum_{i=1}^{k} \alpha_i P v_i = \sum_{i=1}^{k} \alpha_i v_i = u$$

כנדרש

ד) P סימטרית, מהמשפט הספקטרלי היא לכסינה אורתוגונלית כלומר $P = UDU^T$ כאשר U אורתוגונלית.

$$P^2 = P \cdot P = UD\underbrace{U^T U}_{I}DU^T = UD^2U^T$$

אלכסונית

מאחר שD מכיל רק את הע"ע של P על האלכסון שהם רק 1 או 0 נקבל $D^2 = D$ ולכן

$$= UDU^T = P$$

כנדרש ☑

ה)

$$(I-P)P = P - P^2 = P \cdot P = 0$$

סעיף ד

(2.3

6) נניח ש X^TX הפיכה

$$[X^TX]^{-1}X^T = [V\Sigma^T \underbrace{\cancel{U^T U}}^{\text{אורתוגונלית}} \Sigma V^T]^{-1} V\Sigma^T U^T$$

$$= [V\Sigma^T\Sigma V^T]^{-1} V\Sigma^T U^T$$

נסמן D להיות מטריצה אלכסונית מימדים $d \times d$ המתקבלת ע"י $\Sigma^T\Sigma$ וכי σ_i^2 נמצאים על האלכסון

$$= [VDV^T]^{-1} V\Sigma^T U^T$$

$$= VD^{-1}\underbrace{V^TV}_{I_d}\Sigma^T U^T$$

$$= VD^{-1}\Sigma^T U^T$$

כעת נשים לב כי הכפל של $D^{-1}\Sigma^T$ נקבל מטריצה מגודל $d \times m$ כאשר על האלכסון הראשי שלה יש את

$$[D^{-1}\Sigma^T]_{ii} = \frac{1}{\sigma_i^2}\cdot\sigma_i = \frac{1}{\sigma_i} = \Sigma^+$$

כלומר קיבלנו כי

$$\hat{W} = V\Sigma^+U^+y = X^+y$$

מוגדרת X^+ כנדרש ☒

7) X^TX הפיכה $\overset{\text{תרגיל 1}}{\iff}$ כל d הערכים הסינגולריים שונים מ-0 ב $\Sigma^T\Sigma$ $\iff$ כל d הערכים הסינגולריים של Σ שונים מ-0 $\iff$ הו"ע שמתאימים לd הערכים הסינגולריים פורשים את מרחב העמודות של X (פורשים מרחב כמימד d המשמעות d עמודות הן פורשות את R^d)

ולשון שמימד מרחב העמודות הוא הדרגה של המטריצה שזה גם הדרגה של מרחב השורות נקבל שהשורות $x_1 \ldots x_m$ פורשות את R^d.
ולשון אחר בדיוק הוא אותו תנאי עבור לומר כי X^TX הפיכה

ב) $\hat{w} = X^{+}y$

נניח ש X^TX אינו הפיך כלומר למשוואה הנורמלית יש אינסוף פתרונות

כעת נסתכל על פתרון $\bar{w}$

כלומר

$$y = X\bar{w} \iff X^Ty = X^TX\bar{w}$$

נסמן $X = U\Sigma V^T$ פירוק ה svd של X כאשר U ו V אורתוגונליות ! Σ אלכסונית

אם נסמן ב k את הדרגה של X נקבל כי

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_k & \\ 0 & & & 0 \ddots 0 \end{bmatrix}$$

כעת אם נציב את הפירוק במשוואה לעיל נקבל

$$X^Ty = X^TX\bar{w} \iff V\Sigma^TU^Ty = V\Sigma^TU^TU\Sigma V^T\bar{w}$$

$$\iff V\Sigma^TU^Ty = V\Sigma^2V^T\bar{w}$$

כפל משמאל ב V^T

$$\iff \Sigma^TU^Ty = \Sigma^2V^T\bar{w}$$

כעת נקבל

$$\Sigma^TU^Ty = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \, 0 \ddots 0 \end{bmatrix}\begin{bmatrix} - u_1 - \\ \vdots \\ - u_n - \end{bmatrix} y = \begin{bmatrix} \sigma_1\langle u_1, y\rangle \\ \vdots \\ \sigma_k\langle u_k, y\rangle \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\Sigma^2V^T\bar{w} = \begin{bmatrix} \sigma_1^2\langle V_1, \bar{w}\rangle \\ \vdots \\ \sigma_k^2\langle V_k, \bar{w}\rangle \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

כלומר לכל $i \in [k]$ מתקיים $\sigma_i^2\langle V_i, \bar{w}\rangle = \sigma_i\langle u_i, y\rangle$

המשוואה $\hat{w}$ גם הוא פתרון למשוואה ומתקיים $\hat{w} = X^{+}y$

כלומר מאותו שיקול מתקיים כי לכל $i \in [k]$

$$\sigma_i^2\langle V_i, \hat{w}\rangle = \sigma_i\langle u_i, y\rangle$$

כלומר לכל קורדינטה $i \in [k]$ מתקיים; $\sigma_i^2\langle V_i, \hat{w}\rangle = \sigma_i^2\langle V_i, \bar{w}\rangle$

$$\Rightarrow \langle V_i, \hat{w}\rangle = \langle V_i, \bar{w}\rangle$$

כעת עבור הקורדינטות ה $k+1$ עד d נקבל כי

$$\hat{w} = X^{+}y = V\Sigma^{+}U^Ty = V \cdot \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \, 0 \ddots 0 \end{bmatrix}\begin{bmatrix} \langle u_1, y\rangle \\ \vdots \\ \langle u_m, y\rangle \end{bmatrix} = V \cdot \begin{bmatrix} \sigma_1\langle u_1, y\rangle \\ \vdots \\ \sigma_k\langle u_k, y\rangle \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$\Leftarrow$ קיבלנו כי כל הקורדינטות של $\hat{w}$ הן אפסים ולכן את כל הקורדינטות של $\hat{w}$ שוות ל$\bar{w}$ ובנוסף שאר הקורדינטות של $\hat{w}$ הן אפסים כלומר אין להן תרומה כלל לסכום בחישוב הנורמה L_2 נקבל כי

$$\|\hat{w}\|_2 \leq \|\bar{w}\|_2$$

כנדרש

∎

חלק 2

2) - החלטתי להוריד את ה id, date, long, lat, zipcode. שכן הערך המספרי של כל אחד מהעמודות האלו לא גורם לשינוי הגדרה

בנוסף הורדתי את sqft_living15, sqft_lot15 שכן לפי מידע זה דירוג השכנים ולא של הבית עצמו

- על ציונים כמו view, condition, grade בדקתי שכל הערכים הם בטווח המצוין בתיאור של הנתונים

- הסרתי 5 דאטה הכולל מחיר קטן שווה 0

ובנוסף הפרדתי את עמודת המחירים כ y ! שאר הדאטה אלו המאפיינים כ X.

3) - נשים לב שהקשר בין גודל הסלון לבין המחיר של הבית גבוה מאוד ולכן הגרף נראה שהעלייה במחיר היא כמעט לינארית ביחס לעלייה בגודל הסלון

- לעומת זאת, בצורה שבה נראה הגרף של מצב הבית לא נראה שיש קשר למחיר, אפשר לראות שהמחיר די דומה עבור הציונים 3,4,5 וההתפלגות של הבתים בעלי המחירים השונים קיימת ביחס לציונים

4) נראה שאנחנו יכולים ללמוד מכך שככל שכמות הדגימות עולה ככה ה std קטן כלומר אנחנו יכולים להגיע לבטחון יותר בקונסיסטנטיות של המודל שלנו כלומר הוא פחות יושפע מהממוצע מערכי דגימות של תוצאות את הדאטהסט שלנו

לגבי הטרנד של ה loss ניתן לראות בגרף שיש קפיצות הרבה יותר גדולות בין 2 כמות דגימות עבור הדגימות הבודדות לעומת "התכנסות" של ה loss עבור הכמות דגימות הגבוהות יותר.

2) א) הייתי ממליץ על פולינום בדרגה 4 או 6 יותר לגבי שנוצר על ידי הנקודות בגרף

ב) נשים לב שבחודשים ה 7 ו 8 ה std של הטמפרטורות הוא הנמוך ביותר ולכן נסיק שרוב הנקודות בחודשים אלו קרוב לממוצע ולכן יהיה קל יותר לצפות עבור המודל שלנו.

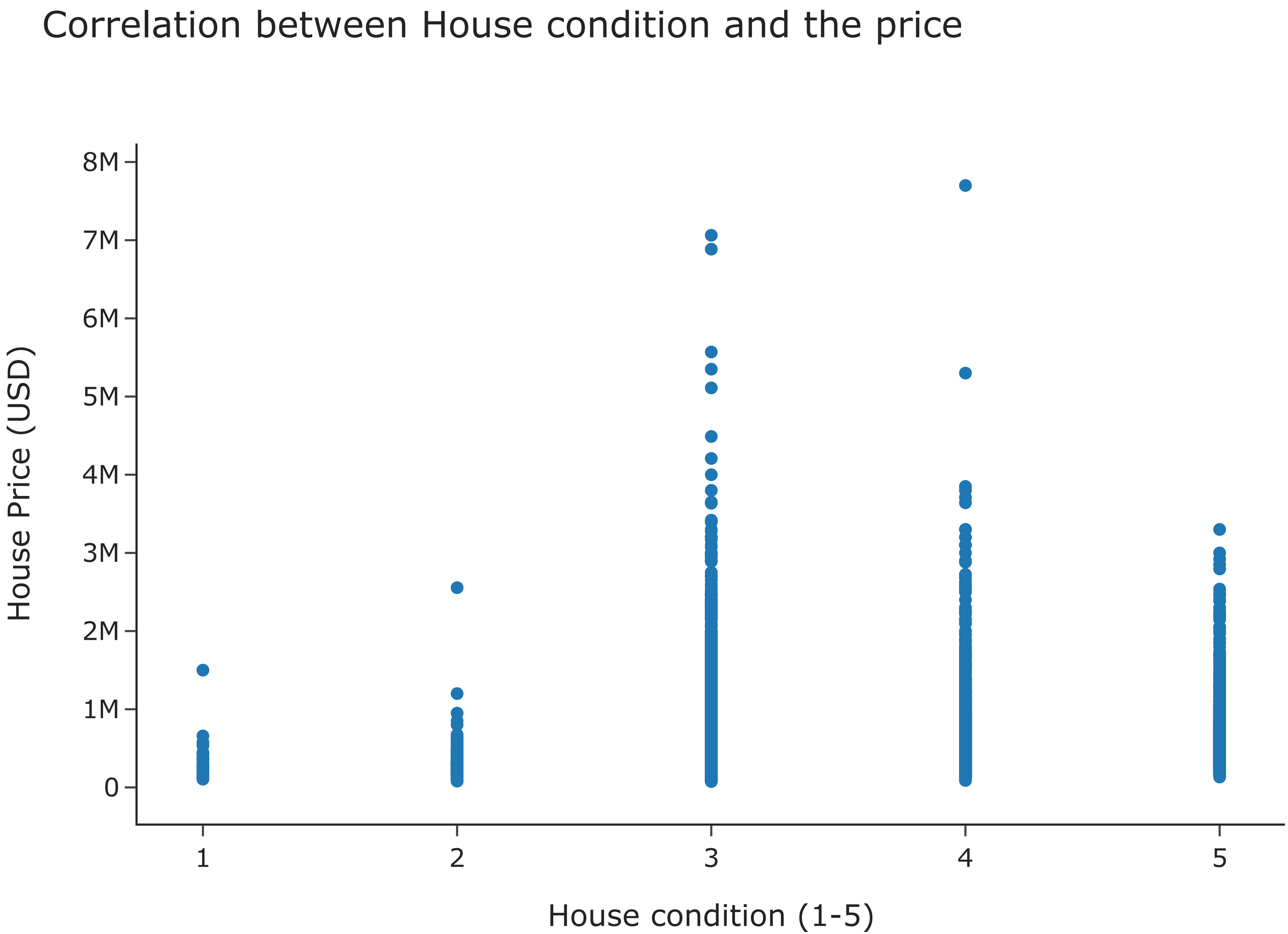
3) נשים לב שהצורה של ישראל הולנד וירדן נראית מאוד דומה אך הצורה של דרום אפריקה הפוך לישראל.

בנוסף ניתן לראות מכך לפי ישראל, ירדן שתופסת לישראל בטמפרטורה הממוצעת רוב השנה כנראה תקבל ערכים טובים יחסית לעומת זאת הולנד שמנה תופסת לישראל לא תקבל תוצאות טובות. דרום אפריקה תקבל תוצאות טובות רק במקרה קטן משנה.

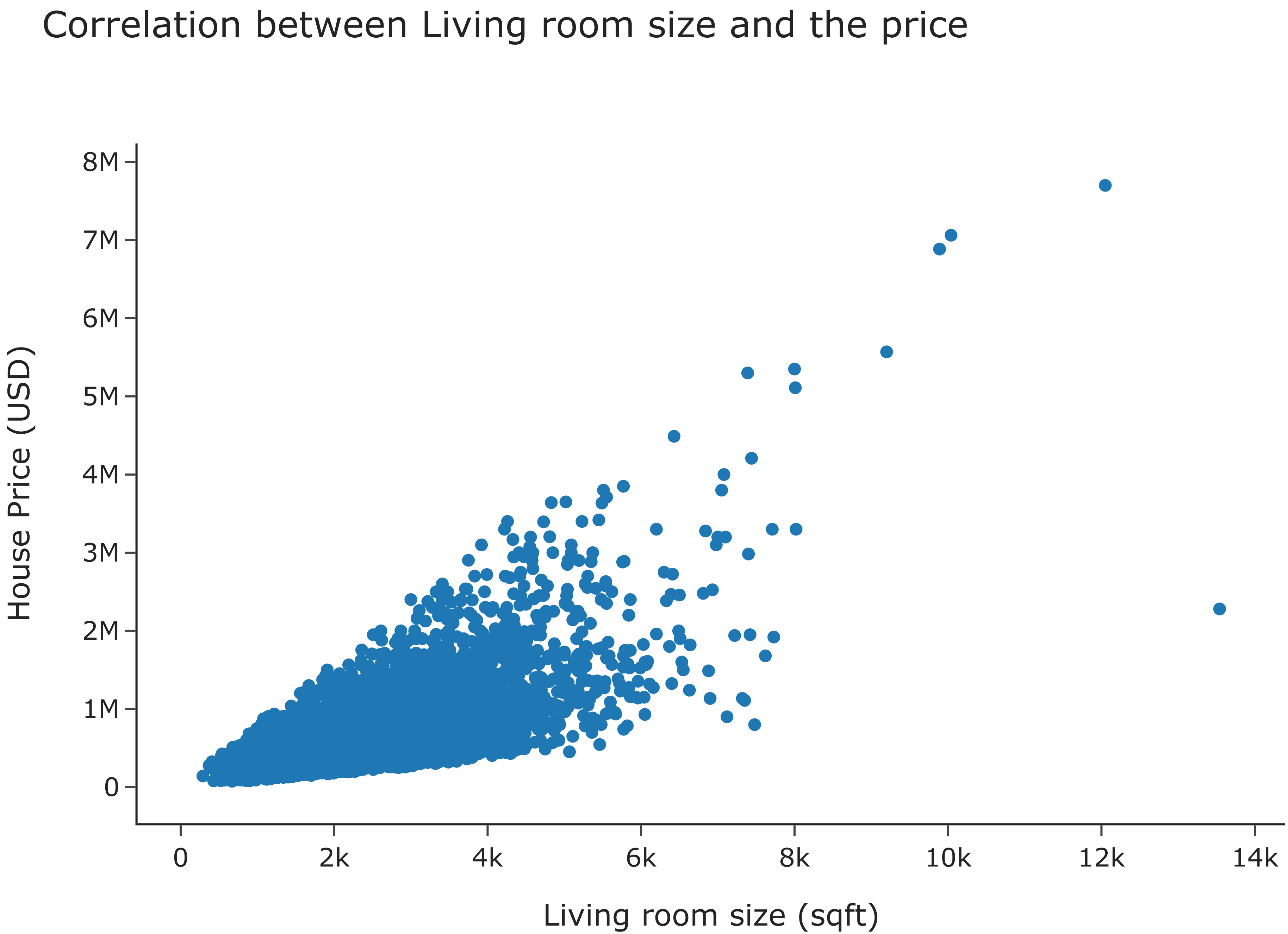
4) [21.6, 6.7, 3.58, 3.11, 3, 6.6, 14.25, 29.75, 29.75, 29.84]

בערך של k שמשיג את התוצאה הטובה ביותר הוא k=5, לדעתי אין צורך לבדוק עוד ערכים שכן נשים לב כי ככל ש k עולה ככה המודל יותר רחוק אחרי הגרף ומאבד את הדיוק התיאורטי

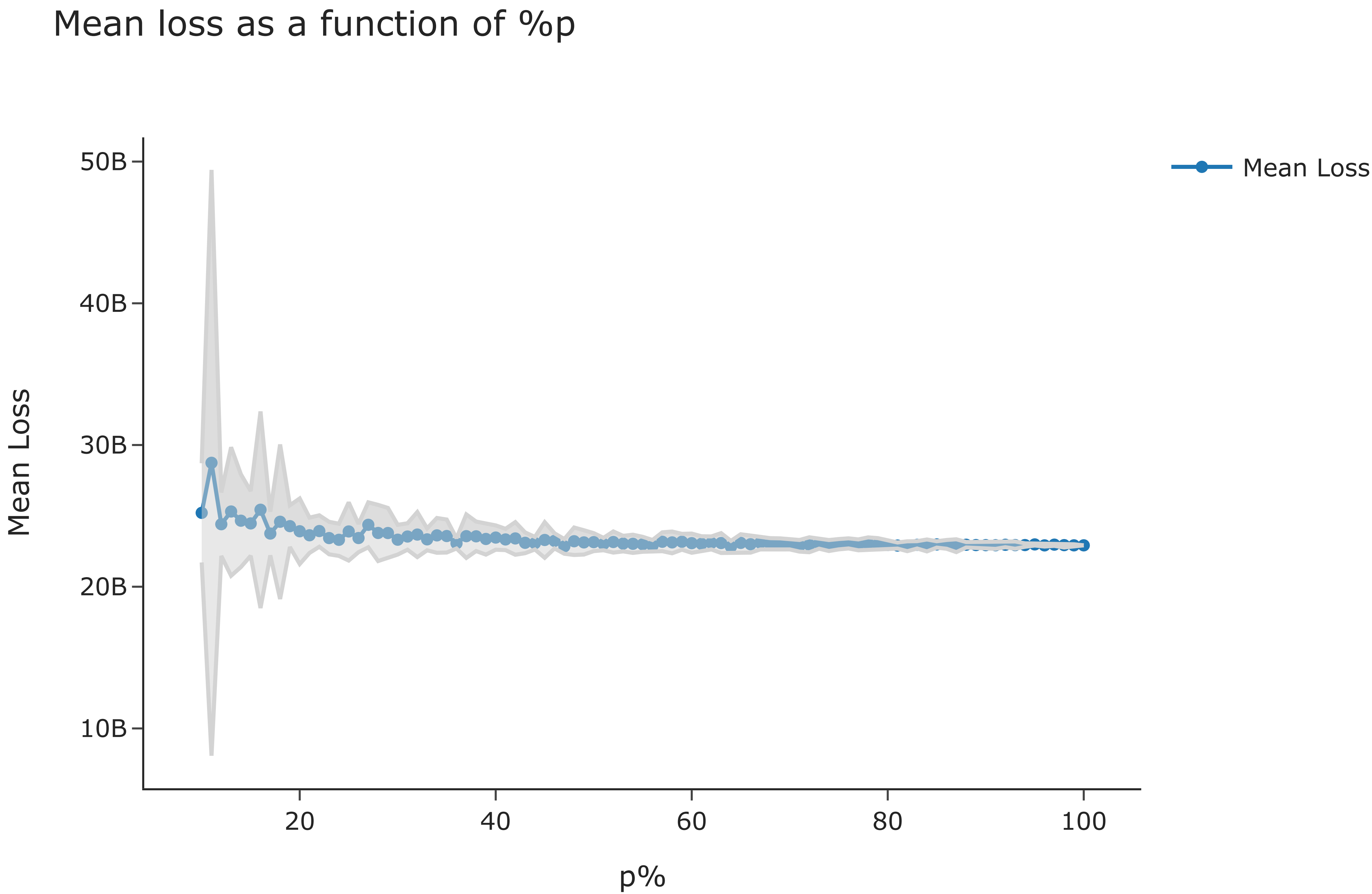
5. כמו שחשבתי בסעיף 3, קיבלנו כי הערות של המודל ביחס לירדן נמוכה ביחס להולנד ודרום אפריקה שם לא הייתה תפיסה בין המדינות האלו לישראל

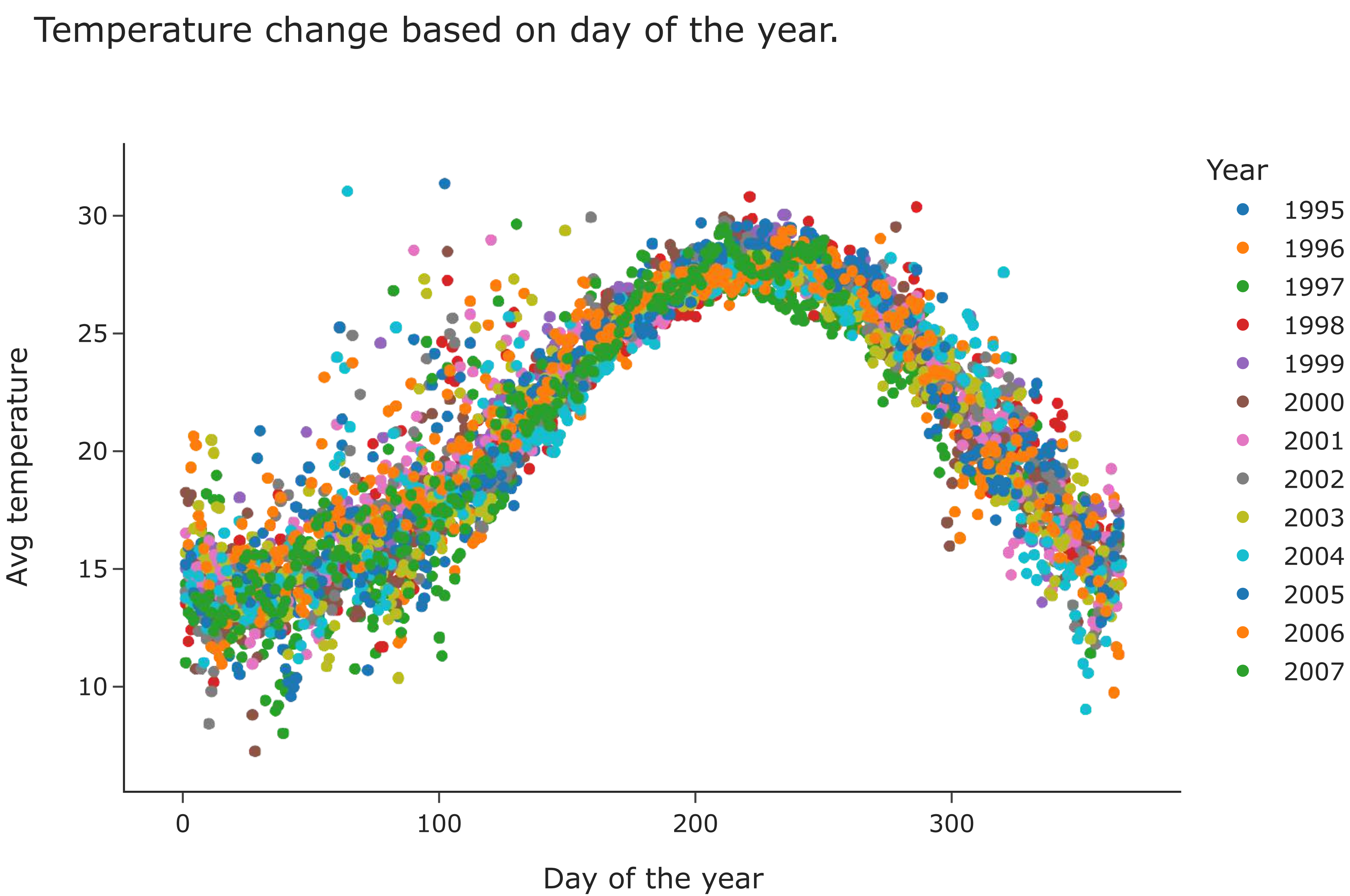


Correlation between House condition and the price
House Price (USD)
8M
7M
6M
5M
4M
3M
2M
1M
0
1
2
3
4
5
House condition (1-5)



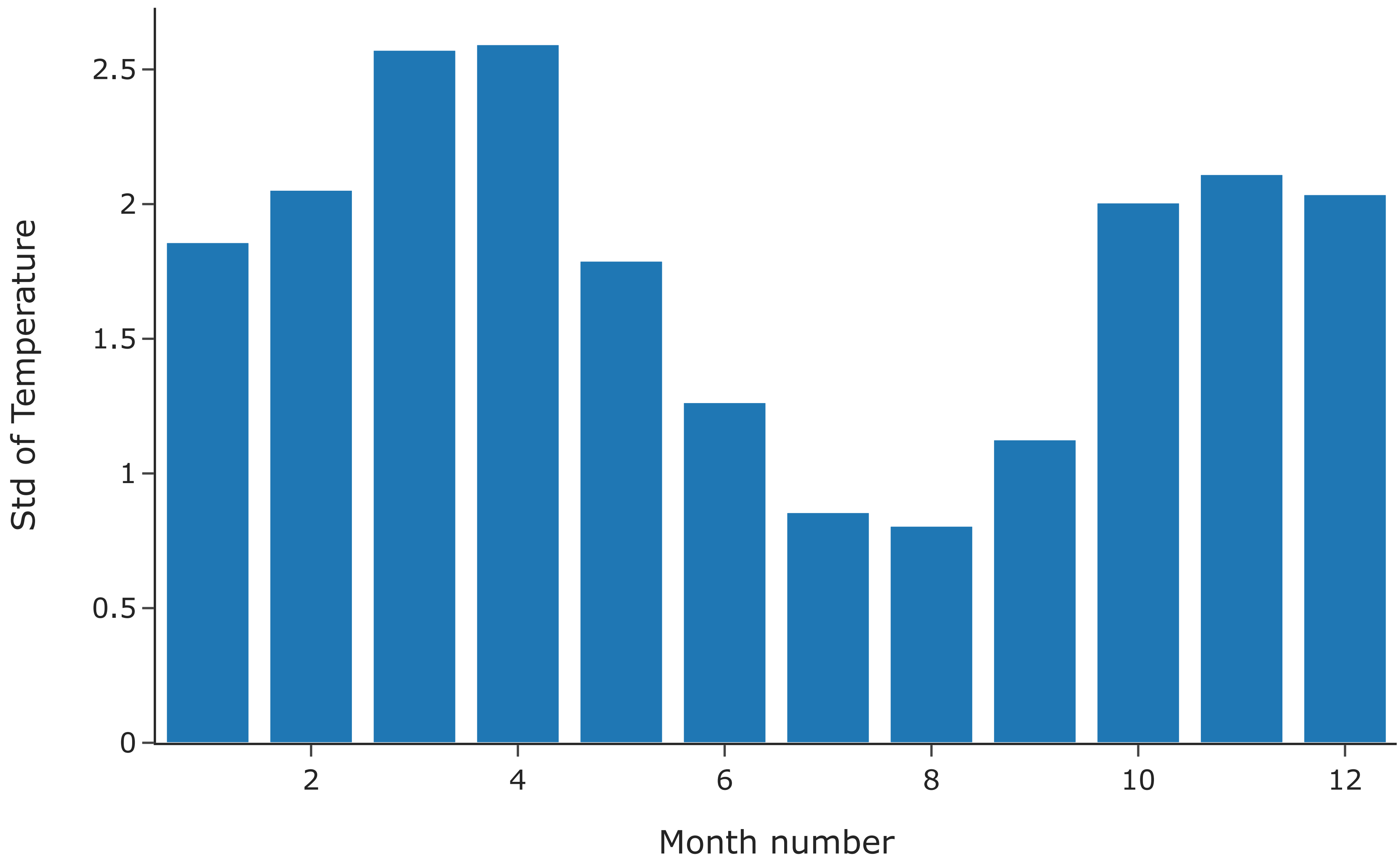
Correlation between Living room size and the price
House Price (USD)
8M
7M
6M
5M
4M
3M
2M
1M
0
0
2k
4k
6k
8k
10k
12k
14k
Living room size (sqft)

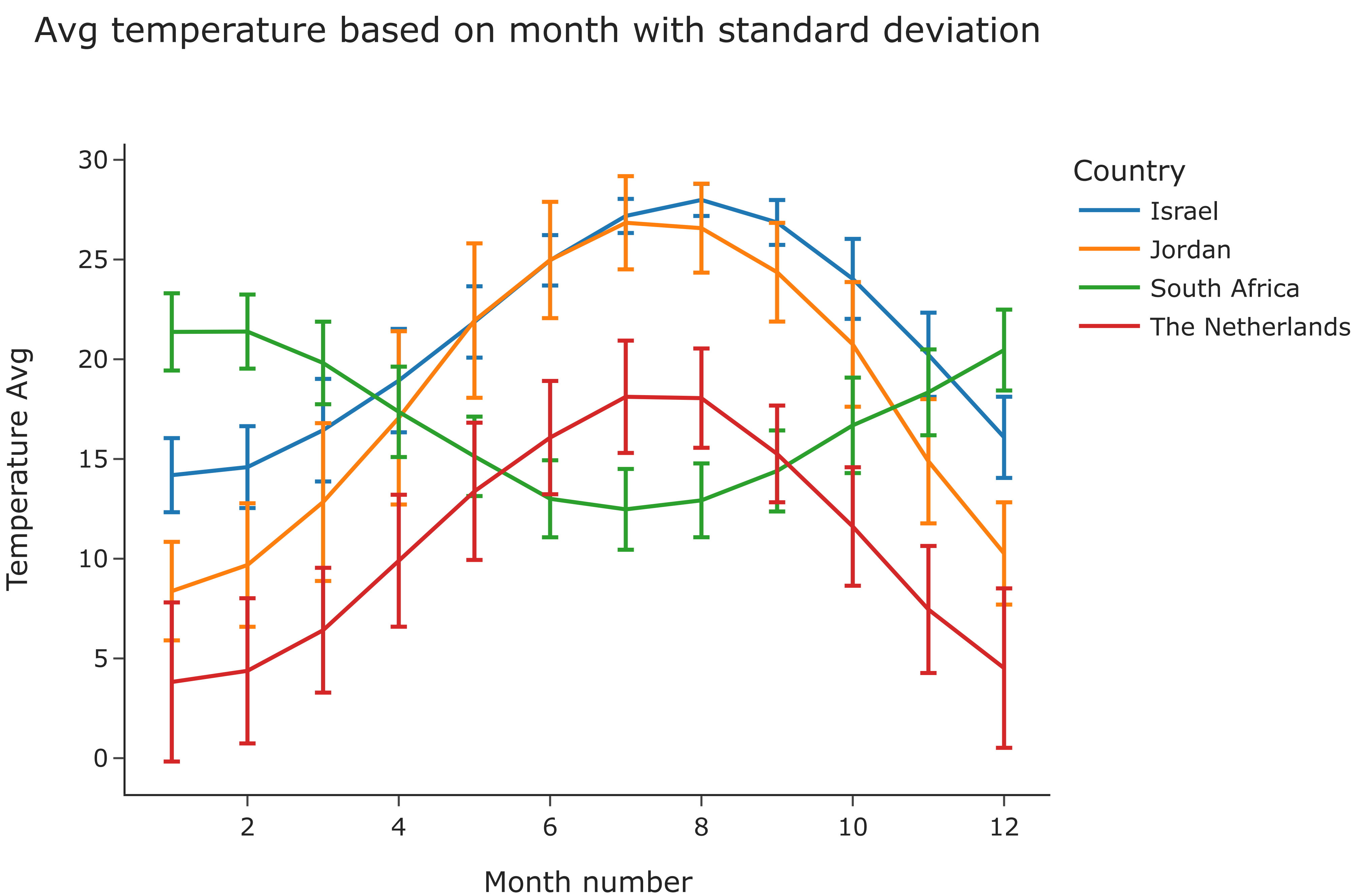

Mean loss as a function of %p
Mean Loss
50B
40B
30B
20B
10B
20
40
60
80
100
p%
Mean Loss

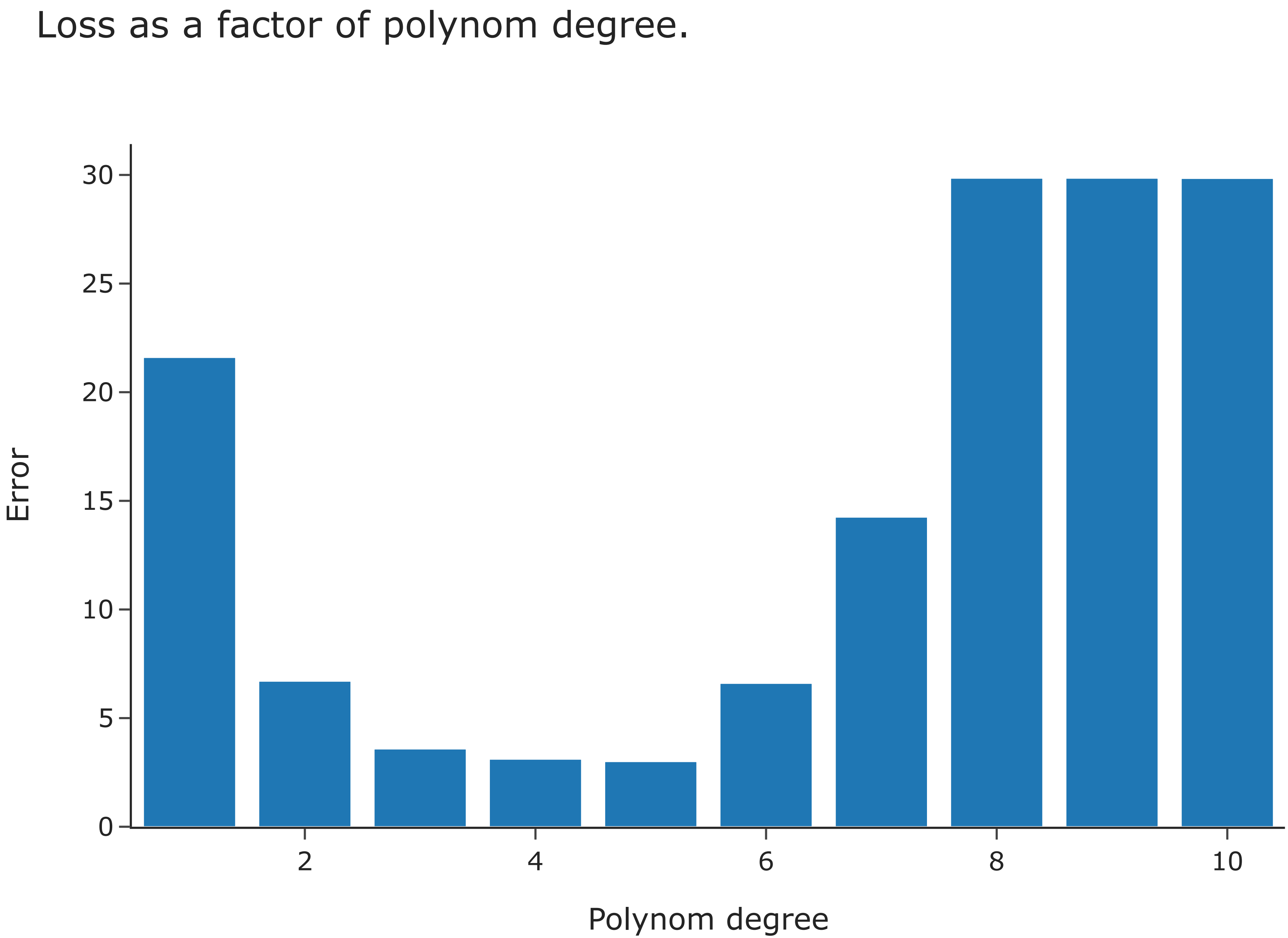


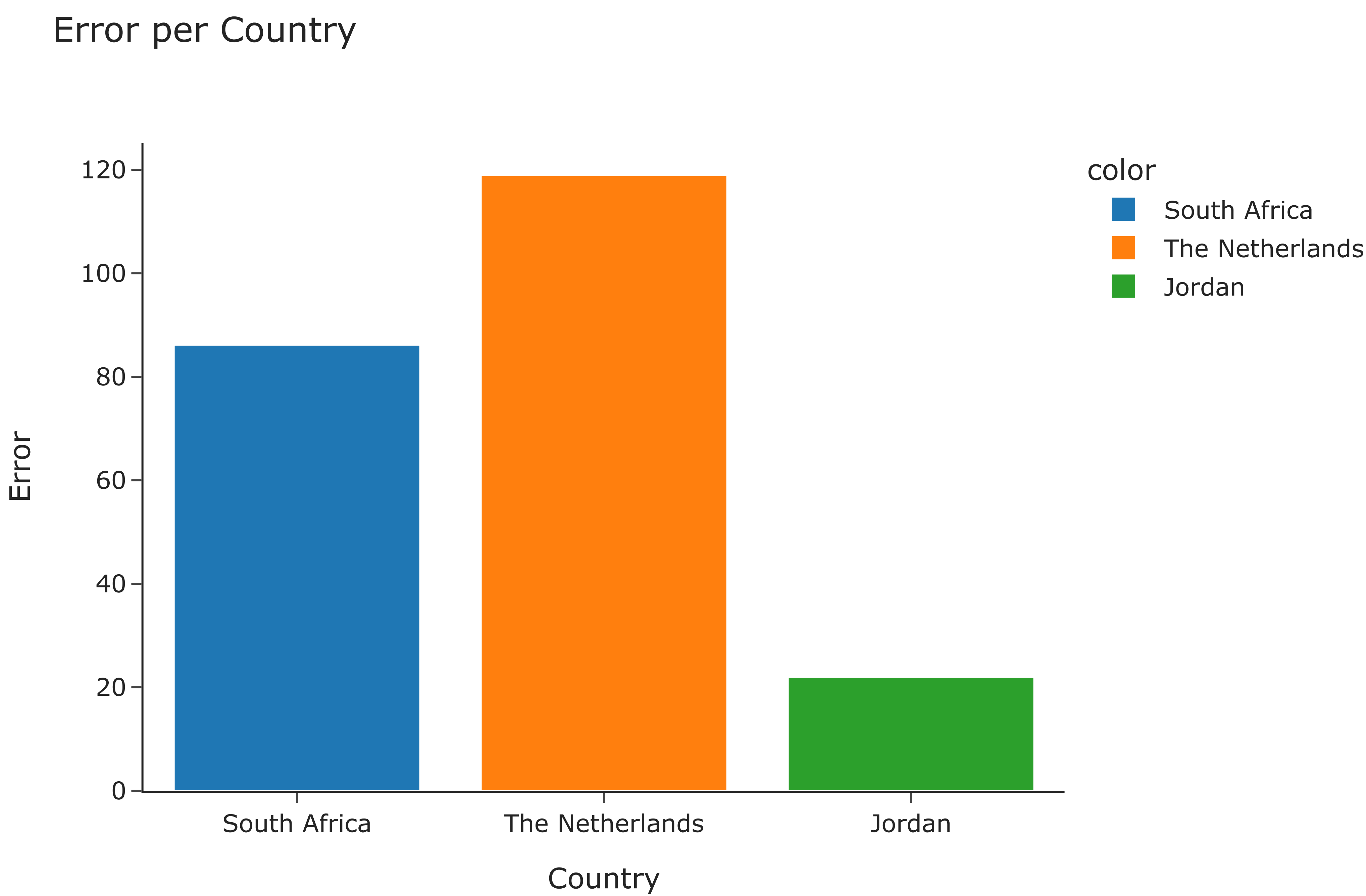
Temperature change based on day of the year.
Avg temperature
Day of the year
Year
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
10
15
20
25
30
0
100
200
300

Std of Temperature based on the Month




Avg temperature based on month with standard deviation
Country
Israel
Jordan
South Africa
The Netherlands
Temperature Avg
Month number
0
5
10
15
20
25
30
2
4
6
8
10
12


Loss as a factor of polynom degree.
Error
30
25
20
15
10
5
0
2
4
6
8
10
Polynom degree


Error per Country
color
South Africa
The Netherlands
Jordan
120
100
80
60
40
20
0
Error
South Africa
The Netherlands
Jordan
Country