איפיון אלגוריתם מחירון רכב

מטרה: יצירת טבלה ב-DB (סכמת pricelist) עם מחירי מחירון סופיים לכל תת דגם+שנתון בכל חודש (או טבלת CSV שתעלה לאחר מכן ל-DB)

האלגוריתם ירוץ כל 20 לחודש, כך שיחשב מחיר מחירון לחודש הקרוב (למשל ב-20 במאי יחושבו מחירי מחירון עבור חודש יוני)

קלט:

* קובץ log עם תאריך חישוב המחירים הסטטסטיים.

הקובץ נכתב לשרת בסיום כל ריצה המחשבת מחירים סטטסטיים ודורס את הקובץ שנכתב בחודש קודם.

יש לוודא לפני הרצת הסקריפט (חישוב מחירי מחירון סופיים) שהתאריך בקובץ גדול/שווה ל-20 לחודש.

* טבלת CSV שמייצר ה-QV עם מחיר סטטסטי לכל תת דגם+שנתון (כ-6000 רשומות) בעל כמות תצפיות מספיקה (גדול/שווה 5)

שדות:

|  |  |  |
| --- | --- | --- |
| שם השדה | תיאור | הערות |
| **SMID** | תת דגם (מחרוזת של 5 ספרות) | מפתח חד ערכי |
| **Year** | שנתון (תאריך בפורמט YYYY) |
| StatPrice | מחיר סטטסטי |  |
| N\_Rows | כמות תצפיות |  |
| Stdev | סטיית תקן לחלק למחיר סטטסטי | מס בין 0 ל-1 |

* טבלת אקסל/CSV עם פרמטרים למשפחה

|  |  |
| --- | --- |
| תיאור | שדה |
| ID משפחה | FamilyID |
| כמות ק"מ סטנדרטית בחודש לרכב במשפחה זו | FamilyKMperMonth |
| כמות ק"מ סטנדרטית בשנה | FamilyKMperYear |
| מס' חודשים שרכב סטנדרטי נמצא באותה יד | FamilyMONTHperHANDS |
| כמות שנים שרכב סטדרטי נמצא באותה יד | FamilyYEARSperHands |
| % הורדת מחיר במעבר חודש | P\_Age |
| % קנס על חריגה מק"מ סטדנרטי | P\_KM |
| % קנס על חריגה מיד סטנרטית | P\_Hands |
|  | P\_ExcesiveKmYear |
| % ירידת ערך שנתית ממוצעת לרכב במשפחה זו | AvgRedAnnual |
| % ירידת ערך שנתית מינימלית | MinRedAnnual |
| % ירידת ערך שנתית מקסימלית | MaxRedAnnual |
| % הפרש מינימלי מותר בין שתי שנים עוקבות באותו SMID המשוייך למשפחה זו | FamilyButtom |
| % הפרש מקסימלי מותר בין שתי שנים עוקבות באותו SMID המשוייך למשפחה זו | FamilyTop |
| % הפחתת מחיר לרכב בשנתו הראשונה | RedPercentFirstYear |
| % הפחתת מחיר רכב חדש על הורדת ניילונים | NewCarMonitor |

* טבלת CSV עם כל הרכבים (תת דגם+שנתון) ומידע נוסף:

|  |  |  |
| --- | --- | --- |
| שם השדה | תיאור | הערות |
| **SMID** | תת דגם (מחרוזת של 5 ספרות) | מפתח חד ערכי |
| **Year** | שנתון (תאריך בפורמט YYYY) |
| Price\_Yad2 | מחיר מחירון חודש קודם |  |
| PriceNewCar | מחיר רכב חדש (עבור השנתון המתקדם ביותר לאותו תת דגם) |  |
| infoFix |  | חודש ראשון לחישוב – לא יהיה. |
| PriceNewCar | מחיר רכב חדש (עבור השנתון המתקדם היותר) |  |

פלט:

טבלה עם שורה עבור כל אחד מ ~ 46K המחירים

מבנה הטבלה הסופית:

|  |  |  |
| --- | --- | --- |
| שם השדה | תיאור | הערות |
| **Year-Month (Pricelist)** | שנה-חודש של המחירון |  |
| **SMID** | תת דגם (מחרוזת של 5 ספרות) | מפתח חד ערכי |
| **Year** | שנתון (תאריך בפורמט YYYY) |
| Yad2\_Price | מחיר מחירון חודש נוכחי | מעוגל ל – 100 הקרוב (27,680 מעוגל ל – 27,700) |
| infoFix | קידוד המתאר את התהליך שעבר מחיר הרכב |  |
| Group | ספרה המתארת קבוצה אליה משתייך מחיר מבחינת אמינות | אפיון מסודר באחריות מאיה – סדר גודל הצמדה לID |
| Yad2\_Price\_NotRounded | מחיר מחירון חודש נוכחי – לפני עיגול ל100 הקרוב |  |
| MinPrice | מחיר מינימלי מותר לרכב |  |
| MaxPrice | מחיר מקסימלי מותר לרכב |  |
| LastPL\_Price | מחיר מחירון קודם |  |
| Mon1\_Price | מחיר שנקבע בסוף בקרה 1 |  |
| Mon2\_Gap\_StdevAvg | הפרש % סטיות תקן- בין הרכב בשנתון הבא (או הרכב אליו משווים) לרכב נוכחי |  |
| Mon2\_PercChange | % שינוי בין רכב בשנתון הבא (או הרכב אליו משווים) לרכב נוכחי |  |
| comp\_N\_Rows | כמות תצפיות עליהן מתבסס בפועל המחיר | אם מחיר הרכב השתנה לפי מחיר רכב אחר- כמות התצפיות היא הכמות של הרכב שמחירו השפיע על מחיר רכב נוכחי |
| comp\_StdevToStatPrice | % סטיית תקן עליו מתבסס בפועל המחיר |  |
| Mon2\_Price | מחיר סופי מבקרה 2 | לפני בדיקה האם המחיר עומד בגבולות המותרים לרכב |

\*המסומנים באפור זה עזר עבורנו להבין כיצד נקבע מחיר.אפשר לייצר קובץ צדדי שמכיל את כל הנתונים האלה, במקום בטבלה הסופית.

**מחירון רכב - חישוב מחיר**

**מחירון חדש**

מודל זה מתחלק לשני חלקים:

1. **מודל גדול** – מחשב את המחיר עפ"י חישוב סטטיסטי של תצפיות מתוך ההיצע.
2. **מודל קטן** – מחשב את המחיר לכל תת דגם + שנתון ש-"נפל" מהמודל הגדול. החישוב מורכב מנוסחה מתמטית יבשה. הנוסחה נעזרת בפרמטרים המחולקים היום לפי משפחה (לשקול לעדכן את הפרמטרים לפי דגם + תת דגם).

**תחילת עבודה** – סיווג תתי דגמים בין המודל הגדול לקטן

1. **בכל חודש** עושים חישוב מחיר סטטיסטי **לכל** תת דגם + שנתון (החישוב נעשה על 6 החודשים האחרונים) – יש לאגור את החישובים החודשיים
2. סיווג בין מודל גדול לקטן, תת דגם + שנתון ייכנס למודל הגדול במידה ועמד באחד מהחוקים הבאים:

* עבור אותו תת דגם + שנתון התקבל מחיר סטטיסטי החודש
* בעבר אותו תת דגם + שנתון עבר דרך המודל הגדול – דורש טבלת תתי דגמים ש"אושרו"
* אין לתת דגם + שנתון מחיר סטטיסטי אבל לתת דגם זהה בשנתון אחר יש מחיר סטטיסטי ~~(בעבר)~~ – בהווה או בעבר

**מודל גדול**

**הערה:**

נקודת המוצא של המחיר היא המחיר הסטטיסטי שחושב באותו החודש, לכל תת דגם + שנתון.

אם יש מחיר חודש קודם – הפחתה חודשית לפי משפחה

אם יש מחיר מחירון לשנתון אחר – עדכון מחיר לפי שנים (פרמטרים כמו במודל קטן)

**חישוב:**

**בקרה ראשונה**

האם המחיר **קטן** מהמחיר בחודש הקודם עד להפרש של 5%? – מתודולגיית השוואה

\*לוודא שלא עברנו מחיר רכב חדש פחות ניילונים

כן

לא

**חישוב בקרה ראשונה**

המחיר החדש קטן ממחיר חודש קודם ביותר מ 5%

**או**

המחיר החדש גדול ממחיר חודש קודם – **דגל**

משאירים את המחיר הסטטיסטי שיצא

**בקרה שנייה – תנאי התחלה**

\*נקבע תחום של מינימום ומקסימום מחיר, בסוף בקרה ראשונה, יש לבדוק שעומד בטווח.

**~~מינימום~~** ~~– מחיר חודש קודם פחות 5%.~~

**מינימום** - מחיר מקסימום -5%

**~~מקסימום~~** ~~– מחיר חודש קודם~~ **~~או~~** ~~מחיר חודש קודם + 5%~~ **~~או~~** ~~מחיר חודש קודם + ההוספה שהייתה~~

**מקסימום** – מחירון חודש קודם

בכל מקרה – לא עובר מחיר רכב חדש פחות ניילונים

[בעמוד הבא]

מחיר גבוה מדי

מחיר קטן מדי

מחיר קטן ממחיר חודש קודם ביותר מ 5% - מתקנים למחיר חודש קודם פחות 5%

מחיר גדול ממחיר חודש קודם –

האם זה חודש שלישי ברציפות שהמחיר גבוה יותר?

לא

כן

לא – 10%

כן – 4%

משאירים את מחיר חודש קודם

האם המחיר גדול ממחיר חודש קודם **בפחות** מ-5%?

מתקנים את המחיר כך שיהיה מחיר חודש קודם + 5% ממנו

(משאירים את הדגל של מחיר גבוה ממחיר חודש קודם)

משאירים את המחיר כך שיהיה המחיר הסטטיסטי של אותו החודש

(מוחקים את הדגל של מחיר גבוה מחודש קודם)

סתירה בין שנתונים

לכל זוג עם פער שאינו בטווח (מינ' מקס' מסוף בקרה 1), בודקים איזה שנתון הוא ה"בעייתי" לפי מטריצה ומתקנים בהתאם (מעגלים להפרש המינימלי המותר או המקסימלי המותר - תלוי מי השנתון הבעייתי)

\*בעת התיקון נתקן גם את כמות התצפיות עליה מתבסס מחיר הרכב ו-% סטיית תקן, לפי כמות תצפיות וסטיית תקן של הרכב שהמחיר השתנה לפיו.

**בקרה שנייה**

עוברים על כל שני שנתונים עוקבים באותו תת דגם ובודקים שהמחיר של השנתון המאוחר יותר גדול מהמחיר של השנתון המוקדם בהפרש מינמימלי ומקסימלי לפי המשפחהאליה משתייך הרכב.

\*נתחיל מהשנתון עם המחיר המבוסס ביותר (כמות תצפיות הכי גבוהה, % סטיית תקן הכי נמוך) ופעם אחת "נזוז" כלפי מטה, פעם שניה "נזוז" כלפי מעלה.

+ לבצע תיקון – אם נדרש - בכל השוואה

לא

כן

יש מחיר!

עובר לבדיקה ידנית

**בדיקה אחרונה לפני סיום**

רק מי שתוקן בבקרה שנייה - האם מחיר חודש עומד בגבולות המותרים לרכב? (מינ' מקס' מסוף בקרה 1)

**מודל קטן**

**הערה:**

המודל מבצע את החישוב שלו בכל חודש על כל תתי הדגמים + שנתונים שאינם נכללים במודל הגדול.

נקודת המוצא של המחיר בכל תת דגם + שנתון היא ה**מחיר רכב חדש** של אותו תת דגם בשנתון הזהה לשנה הנוכחית בחודש ינואר.

**חישוב:**

האם יש לרכב מחיר רכב חדש?

לא

האם יש לרכב מחיר חודש קודם?

לא

כן

עובר לבדיקה ידנית

כן

כן

לא

האם השנתון של הרכב שווה לשנה הנוכחית?

שלב א' – מחשבים את המחיר למחיר שאמור להיות בשנתון של הרכב בחודש ינואר:

לוקחים מחיר רכב חדש ומבצעים הפחתה שנתית עד לשנתון של תת הדגם:

שנה ראשונה – יורד *אחוז שנה ראשונה* (לפי המשפחה)

**+**

לכל שנה נוספת – יורד *אחוז שנה נוספת* (לפי משפחה)

מבצעים הפחתה חודשית, ממחיר רכב חדש (בחודש ינואר), עד החודש הנוכחי:

חודש ינואר – יורדים 40% מתוך *אחוז הירידה בשנה הראשונה* (לפי משפחה)

**+**

חודש פברואר – יורדים 30% מתוך *אחוז הירידה בשנה הראשונה* (לפי משפחה)

**+**

על כל חודש נוסף – יורדים 3% מתוך *אחוז הירידה בשנה הראשונה* (לפי משפחה)

כן

יש מחיר!

לא

מעגלים את המחיר למחיר חודש קודם פלוס / מינוס 5%

יש מחיר!

**בקרה אחרונה**

האם המחיר שיצא, נמוך או גבוה ממחיר חודש קודם עד 5%

שלב ב' – מבצעים הפחתה חודשית, מהמחיר שהתקבל בשלב א' (בחודש ינואר), עד החודש הנוכחי:

לכל חודש – יורד *אחוז הפחתה בחודש* (לפי משפחה) (P\_Age)

מבצעים הפחתה חודשית מחודש קודם (קבוע – P\_Age)

יש מחיר!

הערות:

לא נחשב מחירים עבור רככבים שקיבלו מחיר סטטסטי (מבוסס 55 + תצפיות) נמוך מ- 10,000K

רכבים אלה יכנסו ל"קבוצה" של רכבים שאין צורך יותר לחשב להם מחיר מחירון אף פעם..

נוסחאות

מודל גדול:

הגדרות:

SMID - תת הדגם (5 ספרות) עליו אנחנו מסתכלים

Year – השנתון של הרכב עבורו אנחנו מחשבים מחיר

N\_Rows – כמות תצפיות חודש נוכחי

STDEV\_PERC – אחוז סטיית תקן = סטיית תקן לחלק למחיר סטטסטי

StatPrice – מחיר סטטסטי שחושב לרכב

StatPrice\_Comp- מחיר סטטסטי מחושב – עבור רכבים שאין להם מחיר סטטסטי בחודש נוכחי

N\_Rows\_Comp - כמות תצפיות מחושבת– עבור רכבים שאין להם מחיר סטטסטי בחודש נוכחי ורכבים שמחירם השתנה לפי מחיר רכב אחר

STDEV\_ PERC \_Comp - אחוז סטיית תקן מחושבת – עבור רכבים שאין להם מחיר סטטסטי בחודש נוכחי ורכבים שמחירם השתנה לפי מחיר רכב אחר

Family – המשפחה אליה משתייך הרכב (ספרה בין 1 ל-9)

קבועים:

|  |  |
| --- | --- |
| תיאור | שדה |
| ID משפחה | FamilyID |
| כמות ק"מ סטנדרטית בחודש לרכב במשפחה זו | FamilyKMperMonth |
| כמות ק"מ סטנדרטית בשנה | FamilyKMperYear |
| מס' חודשים שרכב סטנדרטי נמצא באותה יד | FamilyMONTHperHANDS |
| כמות שנים שרכב סטדרטי נמצא באותה יד | FamilyYEARSperHands |
| % הורדת מחיר במעבר חודש | P\_Age |
| % קנס על חריגה מק"מ סטדנרטי | P\_KM |
| % קנס על חריגה מיד סטנרטית | P\_Hands |
|  | P\_ExcesiveKmYear |
| % ירידת ערך שנתית ממוצעת לרכב במשפחה זו | AvgRedAnnual |
| % ירידת ערך שנתית מינימלית | MinRedAnnual |
| % ירידת ערך שנתית מקסימלית | MaxRedAnnual |
| % הפרש מינימלי מותר בין שתי שנים עוקבות באותו SMID המשוייך למשפחה זו | FamilyButtom |
| % הפרש מקסימלי מותר בין שתי שנים עוקבות באותו SMID המשוייך למשפחה זו | FamilyTop |
| % הפחתת מחיר לרכב בשנתו הראשונה | RedPercentFirstYear |
| % הפחתת מחיר רכב חדש על הורדת ניילונים | NewCarMonitor |

1. קביעת מחיר סטטסטי לכל רכב(=תת דגם+שנתון):
2. אם בחודש נוכחי כמות תצפיות גדולה שווה 5 - קיים מחיר סטטסטי, אין בעיה
3. אם בחודש נוכחי כמות תצפיות קטנה מ-5:
   1. אם בעבר היה לרכב מחיר סטטסטי:

- מחיר מחירון קודם

– הקנס במשפחה על מעבר חודש (מספר שלילי)

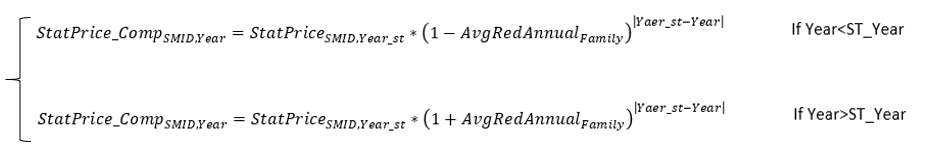
במקרה זו כמות התצפיות, ואחוז סטיית התקן יהיו הנתונים האחרונים שנראו לרכב

* 1. אם מעולם לא היה לו מחיר סטטסטי:

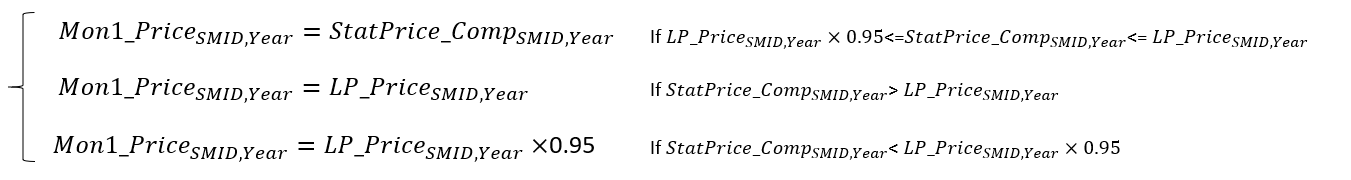
– השנתון החזק ביותר

Year – השנתון של רכב נוכחי

FamilyAvgRedAnnual - אחוז הפחתה שנתית לפי משפחה



1. בקרה ראשונה:

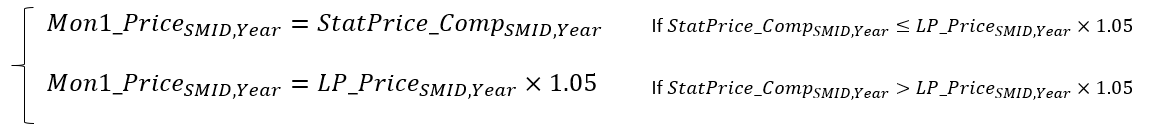


\*

\*

מקרה קצה:

אם זו פעם שלישית ברציפות בה המחיר הסטטסטי גבוה ממחיר המחירון הקודם-



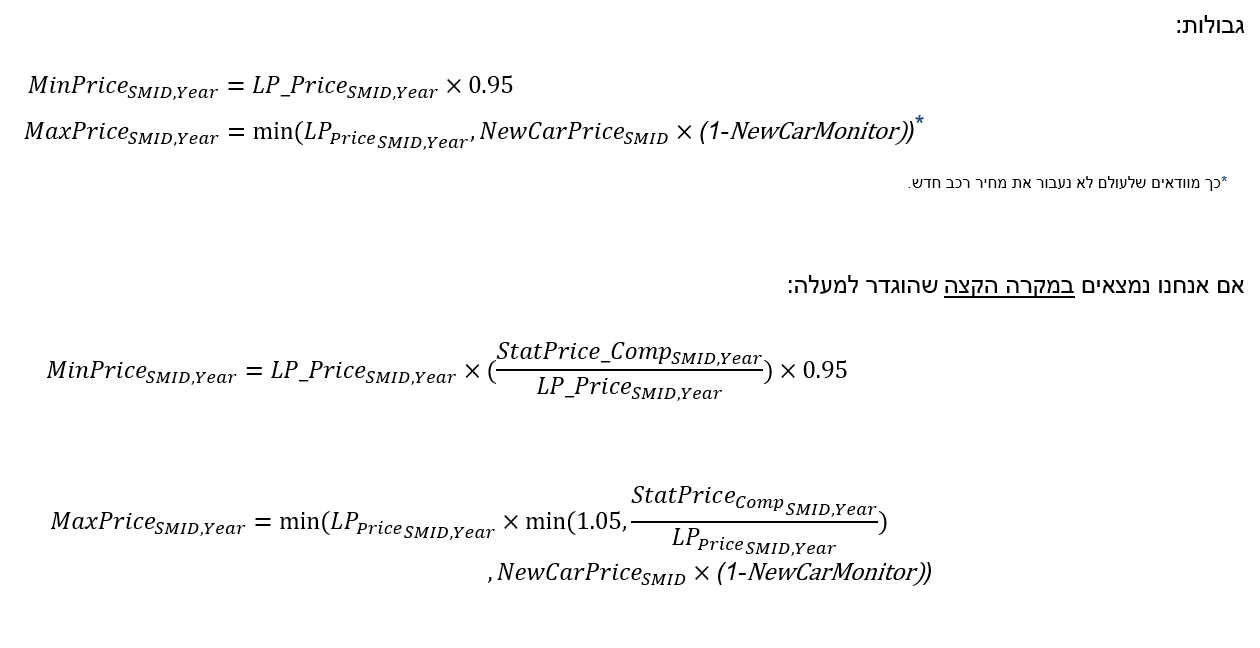
\*

במקרה הראשון הרכב לא יחשב חריג מקסימום.

במקרה השני, בו המחיר הסטטסטי דורש לתקן מעל 5% (אך האלגוריתם לא מאפשר זאת) – הרכב עדיין נחשב חריג מקסימום. אם בחודש הבא יקרה מצב כזה שוב, נמשיך בהאלאת המחיר עד לגבול של 5%.

אם יש חריגה – בכל המקרים המסומנים בכוכבית כחולה- נתקן את כמות התצפיות וסטיית התקן לכמות של חודש קודם

1. קביעת גבולות מותרים לרכב:



1. בקרה שניה:

הבדיקה היא בזוגות.

נתחיל מהשנתון החזק ביותר: הכי הרבה תצפיות, % סטיית תקן הכי נמוכה.

נזוז קדימה ואחורה בשנתונים כלומר- אם למשל לתת דגם יש שנתונים 2013 עד 2018, ו-2016 הוא החזק ביותר: נתחיל לבדוק את 2016 מול 2015, ואז 2015 מול 2014, ואז 2014 מול 2013. נבדוק את 2016 מול 2017, ואז 2017 מול 2018.

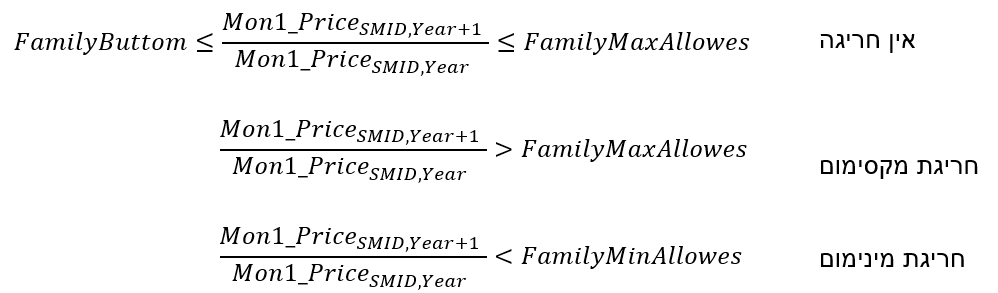
בכל בדיקה נבצע את ההשואה הבאה:

* טבלת עזר – מתוך טבלת הקבועים:

אחוז מינמלי מותר לחריגה נקרא FamilyBottom. אחוז מקסימלי - FamilyTop



איך נקבעת חריגה:



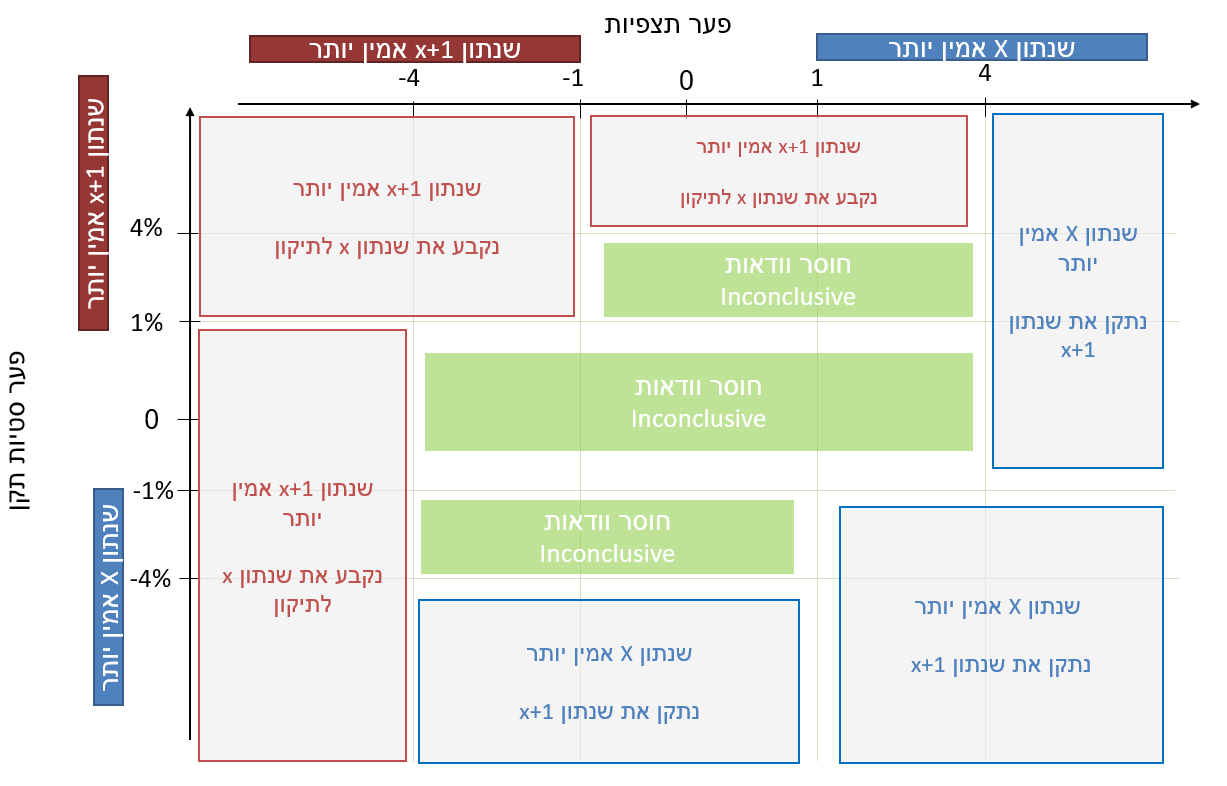
כשמתקבלת חריגה- איך לקבוע את מי לתקן:

\*כאשר לאחד מהם אין תצפיות החודש ולשני יש – נתקן את הרכב ללא התצפיות.

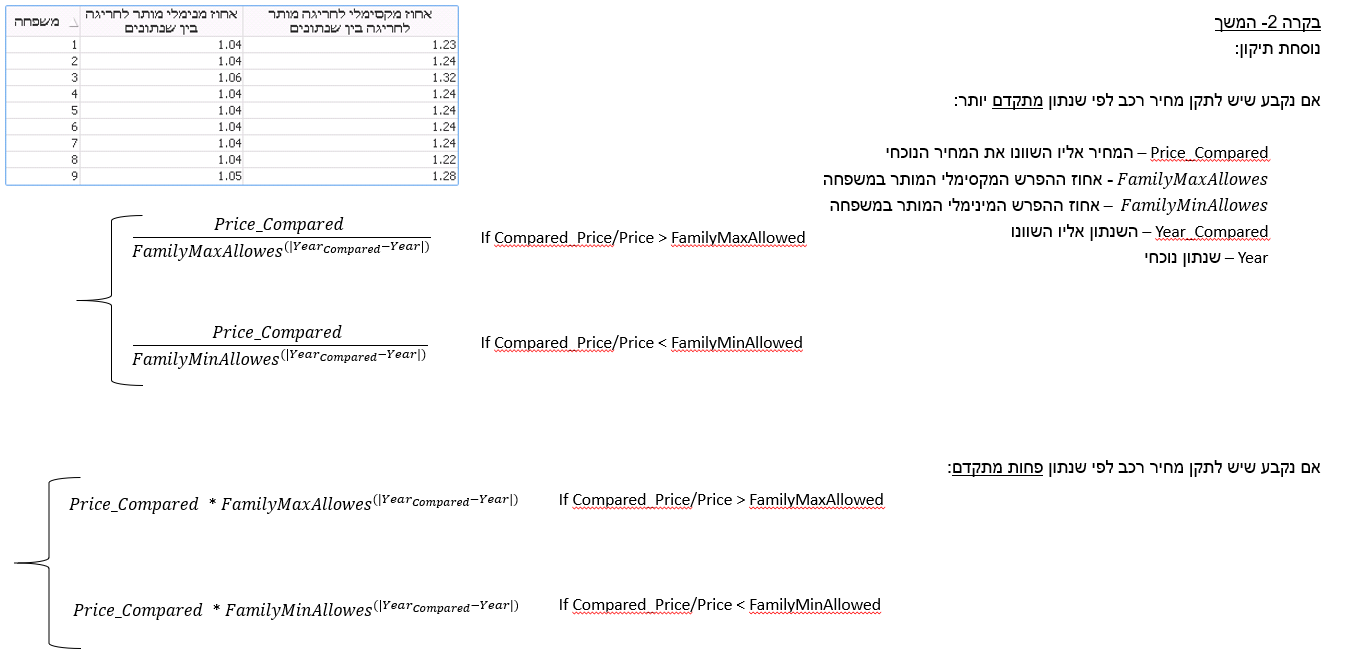
כשלשניהם יש תצפיות-

יש ליצר הפרשים בין תצפיות ובין % סטיית תקן.

נקבע את מי לתקן לפי המטריצה הבאה:



איך לתקן:



לבסוף יש לבדוק שהתיקון עומד בגבולות המותרים לרכב, וכך לקבוע את מחיר הרכב.

1. את המחיר יש לעגל ל-100 הקרוב.

מודל קטן:

