

פרויקט סיכום יחידת SQL

אנליסטים: ברק הופמן, דניאל פלטובסקי ומיכל בן יהודה.

מסד הנתונים הנחקר: AdventureWorks2016

רקע: מסד הנתונים AdventureWorks2016 מציג נתונים של חברה העוסקת במכירת אופניים ואביזרים נלווים. בין הטבלאות השונות אפשר למצוא נתונים של: מכירות, מוצרים, לקוחות, עובדים וספקים. מסד הנתונים עוקב אחר פעילות החברה בין השנים 2011-2014, מאז שנפתחה ועד שנסגרה.

מטרה: לבחון את שלוש השאלות העסקיות הבאות-

1. האם קיימת עונתיות בהכנסות וברווחים?
2. האם ישנה מגמת עלייה או ירידה בהכנסות החברה לאורך החודשים והשנים?
3. בחירת נושא אחד המשפיע על רווחיות החברה, ולאחר בחינתו- כיצד ניתן לשפר את רווחיות החברה?

חקר הנתונים:

ראשית יצרנו פאנל נתונים הנקרא Panel_Business ומורכב מאיחוד של שלוש טבלאות:

Sales.SalesOrderHeader
Sales.SalesOrderDetail
Production.Product

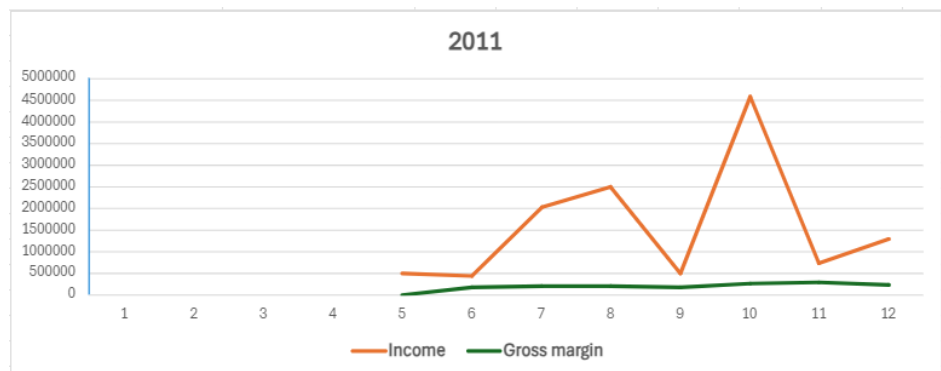
Panel_Business מציג את העמודות הבאות:

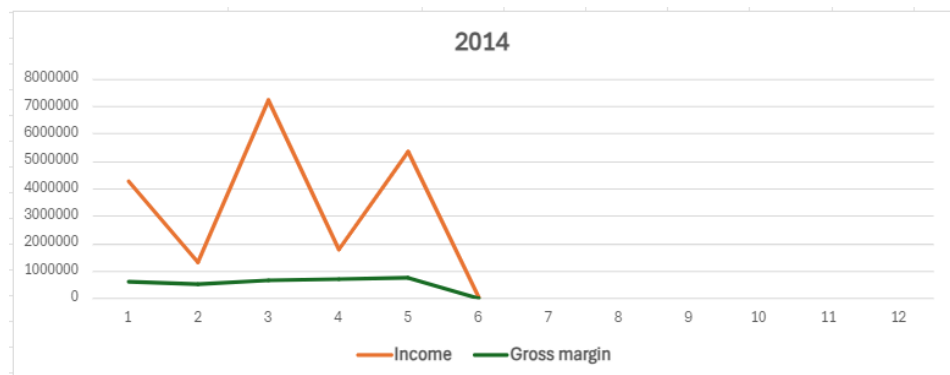
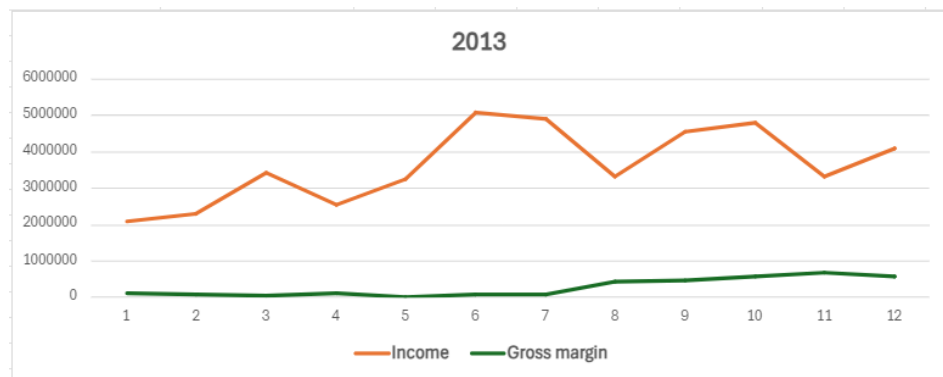
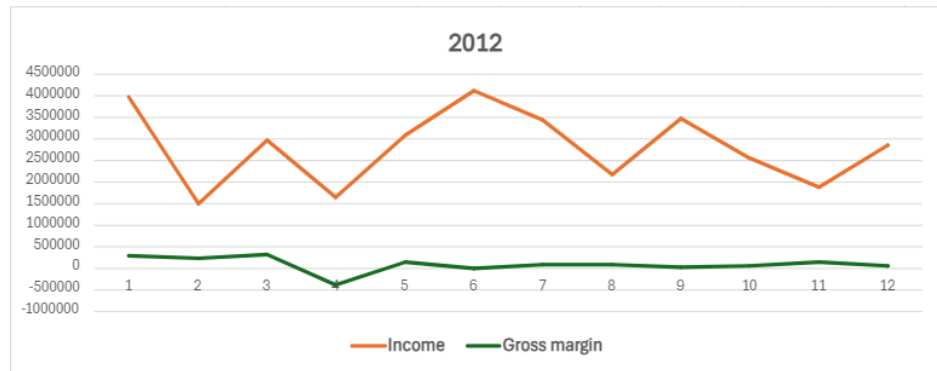
SalesOrderID	OrderDate	OrderYear	OrderMonth	subtotal	OrderQty	ProductID	UnitPrice	UnitPriceDiscount	LineTotal	Name	StandardCost	UnitGrossMargin	LineGrossMargin
43659	00:00.0	2011	5	20565.6206	1	776	2024.994	0	2024.994	Mountain-100 Black, 42	1898.0944	126.8996	126.8996
43659	00:00.0	2011	5	20565.6206	3	777	2024.994	0	6074.982	Mountain-100 Black, 44	1898.0944	126.8996	380.6988
...

הפאנל מציג עמודות מקוריות מהטבלאות שאוחדו, לצד עמודות מחושבות לצרכי המחקר. הפאנל יאפשר לנו שליפה מהירה ונוחה יותר של נתונים רלוונטיים לצורך מענה על השאלות העסקיות. את המשך חקר הנתונים ביצענו ב-MSSQL וב-EXCEL.

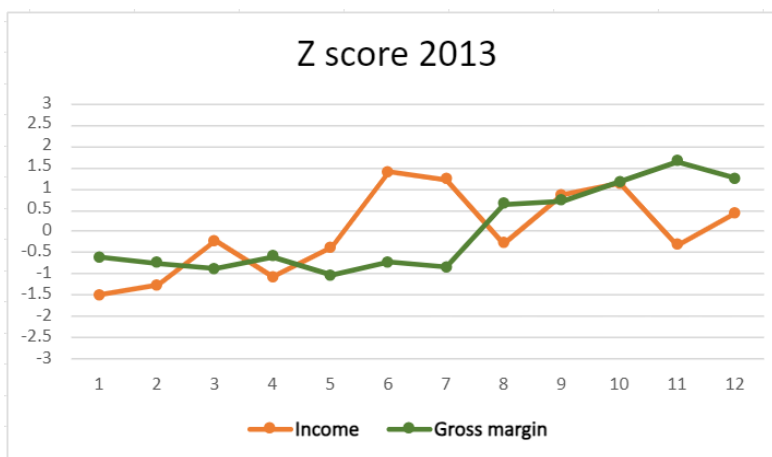
1. האם קיימת עונתיות בהכנסות וברווחים?

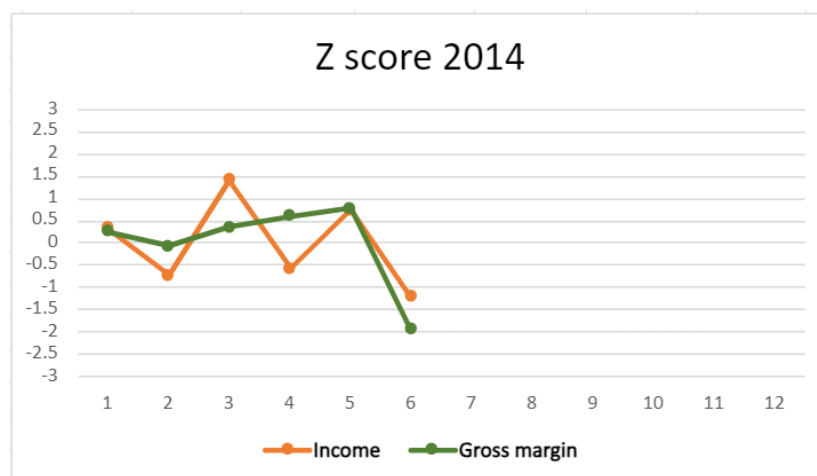
לצורך מענה על שאלה זו, בחנו את סך ההכנסות ואת סך הרווחים עבור כל שנה במטרה לזהות דפוס עונתי:





לאחר בחינה של הגרפים, ראינו שכאשר מציגים את הרווחים יחד עם ההכנסות, ניתן לראות עליות וירידות בהכנסות אך העליות והירידות של הרווחים לא ברורות מספיק. למעשה מתקבלת אשליה שהרווחים הם יציבים לאורך כמעט כל השנה, וזה בניגוד לתוצאות השאילתות שהרצנו על הרווחים שהעידו על תנודתיות גבוהה יותר. כדי לעשות השוואה בין מגמות בהכנסה לבין מגמות ברווח הגולמי, הוחלט לנרמל את הנתונים בעזרת Z-score וכך לשפר את ההבנה שלנו לגבי הנתונים האמיתיים ולסייע בקבלת החלטות מושכלות יותר:





מסקנות:

הכנסות - הבחנו בדפוס של עליה וירידה עונתית, כך שבתחילת כל עונה ישנה עליה בהכנסות ואז ירידה עד לתחילת העונה הבאה:

תחילת החורף: דצמבר-ינואר

תחילת האביב: מרץ

תחילת הקיץ: יוני

תחילת הסתיו: ספטמבר-אוקטובר

רווחים - ישנם מקרים שבהם חלה עלייה ברווחים בתחילת עונה בהתאמה לעלייה בהכנסות. עם זאת, הבחנו גם במספר חודשים שבהם נצפה זינוק בהכנסות אך מנגד הייתה ירידה ברווחים או עליה מתונה מאוד. דוגמאות שראינו: בחודשים אפריל ויוני בשנת 2012 החברה הייתה בהפסד ולא הצליחה להרוויח, דבר זה מפתיע במיוחד בהסתכלות על חודש יוני שבו נראתה עלייה חדה בהכנסות. חודשים יוני-יולי בשנת 2013 גם הם מראים תמונה של עלייה בהכנסות, אך הדבר לא משתקף ברווחים.

ממצאים אלה מביאים אותנו למסקנה שיש בעיה רווחית בחברה הגורמת להפסדים, והדבר ניכר בעיקר בחודשים שבהם ההכנסה עולה.

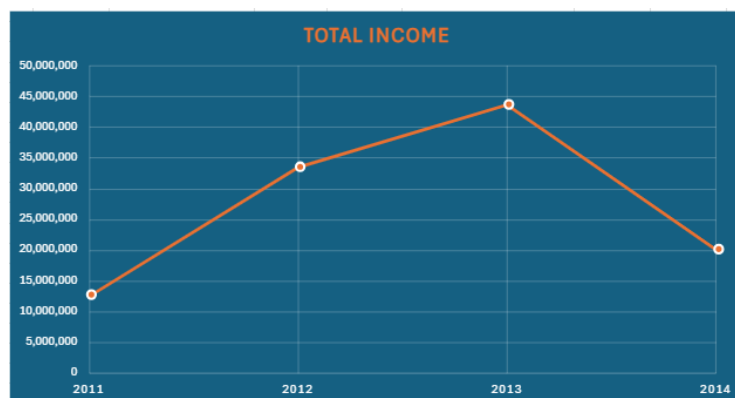
השערה שנרצה לחקור בהמשך היא: תמחור מוצרים במחיר הנמוך מהעלות שלהם.

2. האם ישנה מגמת עלייה או ירידה בהכנסות החברה לאורך החודשים והשנים?

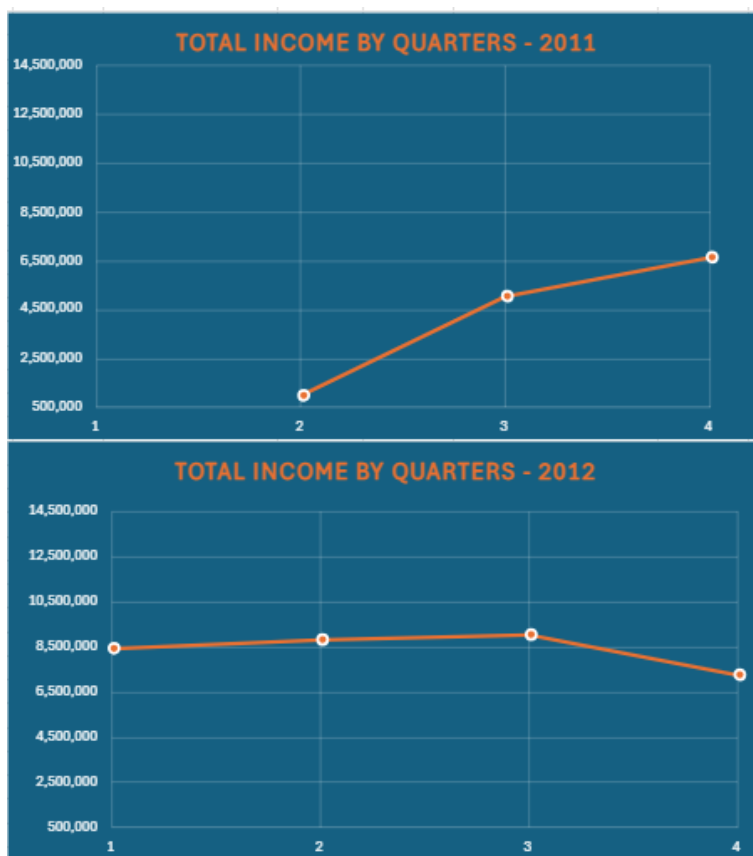
ראשית הסתכלנו על החודשים והשנים שבהן החברה הייתה פעילה:

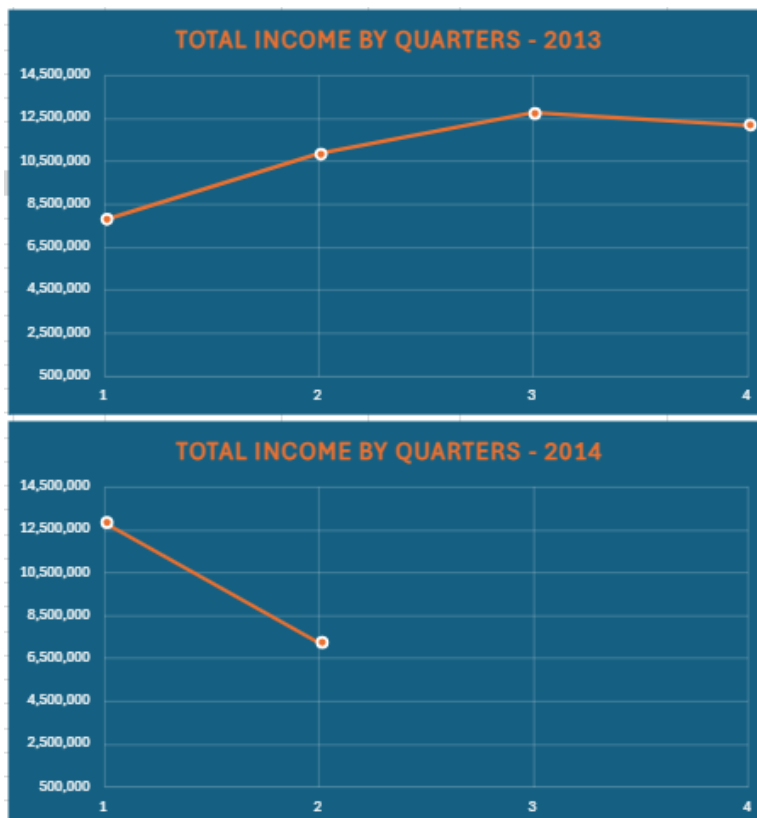
ב-2011 החברה התחילה לעבוד בחודש מאי, 2012 ו-2013 היו שנים מלאות ולבסוף ב-2014 החברה נסגרה בחודש יוני.

הגרף הבא מציג את מגמת **הכנסות** לאורך שנות פעילות החברה:



אפשר לראות שיש מגמת עלייה בהכנסות עד לשנת 2013 ואז חלה ירידה עד לסגירת העסק, אך גרף זה לא משקף את העובדה ששנים 2011 ו-2014 הן לא שנים מלאות. כדי לקבל תמונת מצב נכונה יותר של **ההכנסות** לאורך שנות פעילות החברה, החלטנו לבחון את סך ההכנסות לפי רבעונים:





מסקנות:

בהסתכלות על רבעונים, אפשר לראות שישנה עלייה בהכנסות משנה לשנה, עד לירידה ברבעון השני בשנת 2014 שבו החברה סיימה לעבוד. בנוסף, הבחנו כי בין הרבעון השני לרבעון השלישי ישנה עלייה בהכנסות בשנים 2011-2013. בהמשך לעונתיות שדנו בה בשאלה הראשונה, אנו מסיקים מכך שהמעבר מעונת האביב (הרבעון השני) לעונת הקיץ (הרבעון השלישי) הוא בעל ההכנסות הגבוהות ביותר. השנה המכניסה ביותר היא שנת 2013 ובה נצפתה עלייה בהכנסות מהמחצית הראשונה למחצית השנייה. על אף העלייה בהכנסות לאורך השנים, החברה מפסיקה את פעילותה בסוף הרבעון השני של 2014. תוצאות אלה יחד עם תוצאות השאלה הקודמת מחזקות אצלנו את ההבנה שיש גורם עסקי בחברה שמוריד את רווחיותה.

3. בחירת נושא אחד המשפיע על רווחיות החברה, ולאחר בחינתו- כיצד ניתן לשפר את רווחיות החברה?

בהמשך למסקנות מהשאלות הקודמות, הנושא אותו נרצה לבחון הוא: תמחור מוצרים, ובפרט מוצרים אשר מתומחרים פחות מעלותם. נבחן גם האם ההנחות משתלמות לחברה. השאלות העסקיות שמנחות אותנו הן:

◆ האם יש מוצרים "לא רווחיים"- כלומר שמתומחרים פחות ממחיר עלותם?

◆ האם יש מוצרים שמלכתחילה היו מתומחרים ברווח אך ההנחה גרמה להם להימכר בהפסד?

◆ כמה מוצרים "לא רווחיים" נמכרים בכל שנה?

◆ לאיזו קטגוריה שייכים רוב המוצרים ה"לא רווחיים"?

◆ איך נוכל להציג את דירוג המוצרים הכי "לא רווחיים" בחברה תוך התייחסות לכמות המכירה שלהם?

◆ האם מוצרים "לא רווחיים" נמצאים בראש רשימת המוצרים הכי נמכרים בחברה?

◆ איך הנחות משפיעות על אורך חיי הלקוח והאם הן משתלמות לחברה?

ראשית יצרנו פאנל נתונים הנקרא Panel_Products ומורכב מאיחוד של שלוש טבלאות:

Production.Product
Product.ProductSubcategoryID
Production.ProductCategory

Panel_Products מציג את העמודות הבאות:

ProductName	ProductCategoryID	CategoryName	ProductSubcategoryID	SubCategoryName
HL Road Frame - Black, 58	2	Components	14	Road Frames
HL Road Frame - Red, 58	2	Components	14	Road Frames
...

הפאנל הזה והפאנל שיצרנו בשאלה הראשונה מקושרים אחד לשני דרך עמודת ProductName. בעזרת איחוד הפאנלים נוכל לייצר טבלה חדשה שתראה לנו את הקשר בין הרווחים לבין קטגוריות המוצרים.

בעבור כל שנה, מצאנו את מספר המוצרים ה"לא רווחיים" ולכמה מוצרים מהם יש בנוסף לכך גם הנחה (ובכך ההפסד מוגדל). חישבנו את אחוז המוצרים ה"לא רווחיים" מתוך סך המוצרים שנמכרו.

Year	All sold products	Sold with loss	Sold with loss+discount	% Sold with loss
2011	12888	3064	10	23.8
2012	68579	10373	382	15.1
2013	131788	11544	1138	8.8
2014	61659	4173	287	6.8
2011-2014	274914	29154	1817	10.6

מצאנו שלאורך כל שנות פעילות החברה נמכרו כמויות גבוהות של מוצרים "לא רווחיים". חלקם נמכרו גם בהנחות משמעותיות של 20%, 35% ואפילו 40%, מה שמגדיל את ההפסד בצורה משמעותית אף יותר. בנוסף ערכנו בדיקה לכל שנה ומצאנו שאין מוצרים שמלכתחילה תומחרו ברווח אך נמכרו בהפסד בעקבות הנחה. כלומר, כל המוצרים ה"לא רווחיים" מקורם בתמחור נמוך מהעלות.

לאחר מכן, יצרנו עמודה הנקראת Total Gross Margin ומחושבת לפי Unit Gross Margin*QTY, כלומר הרווח הגולמי לפריט כפול הכמות שנמכרה מאותו פריט. בעבור מוצרים "לא רווחיים" ה- Total Gross Margin

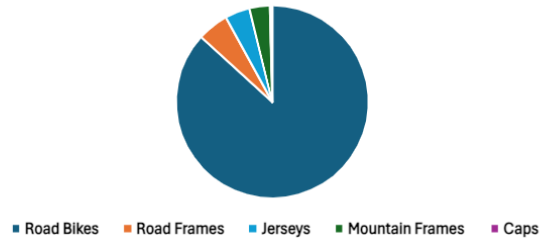
הינו מספר שלילי ומייצג הפסד גולמי.

מדד זה יאפשר לנו לשקלל את הרווח הגולמי השלילי יחד עם כמות המכירה, ובכך לקבל תמונה נכונה יותר של דירוג המוצרים ה"לא רווחיים" מאשר הסתכלות רק על הרווח הגולמי השלילי לפריט. לאחר מכן חישבנו בעבור כל שנה את סך ה-Total Gross margin לכל המוצרים ה"לא רווחיים" לפי תת-קטגוריה וקטגוריה. נתונים אלה אפשרו לנו לאפיין את המוצרים ה"לא רווחיים". ניתן לראות את כל התוצאות בטבלה, ואת חמש הקטגוריות עם הכי הרבה מוצרים "לא רווחיים" מוצגים בגרף שליד:

Unprofitable products grouped by SubCategory and their loss for 2011

CategoryName	SubCategoryName	TotalGrossMargin
Bikes	Road Bikes	-208632.0101
Components	Road Frames	-12593.2149
Clothing	Jerseys	-10007.04184
Components	Mountain Frames	-8297.9521
Clothing	Caps	-956.118364

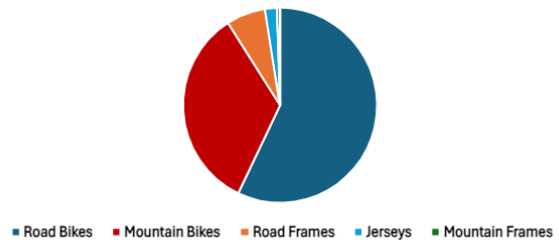
Unprofitable products - 2011



Unprofitable products grouped by SubCategory and their loss for 2012

CategoryName	SubCategoryName	TotalGrossMargin
Bikes	Road Bikes	-1224066.203
Bikes	Mountain Bikes	-725620.4608
Components	Road Frames	-139990.6654
Clothing	Jerseys	-42831.87289
Components	Mountain Frames	-10802.3357
Clothing	Caps	-3777.889284

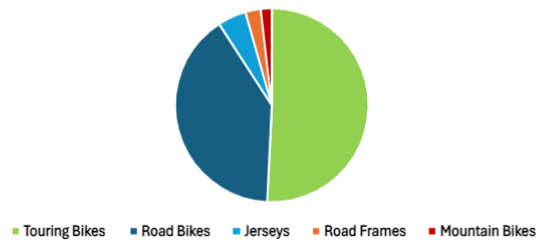
Unprofitable products - 2012



Unprofitable products grouped by SubCategory and their loss for 2013

CategoryName	SubCategoryName	TotalGrossMargin
Bikes	Touring Bikes	-1089202.634
Bikes	Road Bikes	-859645.1098
Clothing	Jerseys	-102261.4581
Components	Road Frames	-54955.49069
Bikes	Mountain Bikes	-39515.34112
Components	Touring Frames	-7127.180115
Clothing	Caps	-4716.842381
Components	Mountain Frames	-1537.72

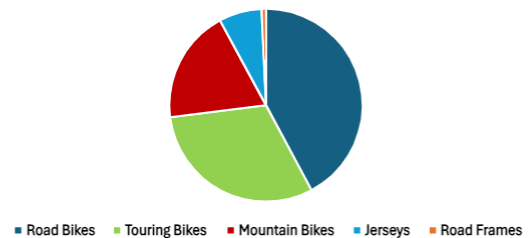
Unprofitable products - 2013



Unprofitable products grouped by SubCategory and their loss for 2014

CategoryName	SubCategoryName	TotalGrossMargin
Bikes	Road Bikes	-203093.2819
Bikes	Touring Bikes	-148392.3588
Bikes	Mountain Bikes	-91840.0962
Clothing	Jerseys	-34565.44081
Components	Road Frames	-3446.2787
Clothing	Caps	-1379.202648

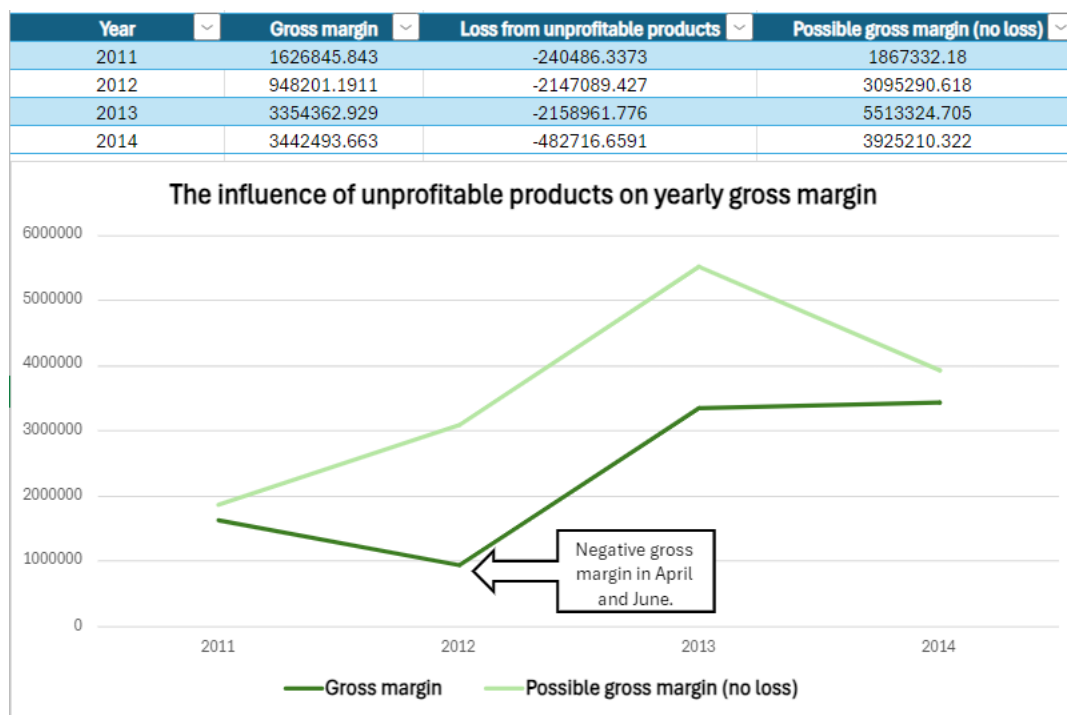
Unprofitable products - 2014



מסקנה: הקטגוריה עם הכי הרבה מוצרים "לא רווחיים" בכל שנות פעילות החברה היא "**Bikes**", ובפרט תת-קטגוריה "**Road Bikes**" אשר מדורגת במקום הראשון בשנים 2011, 2012, ו-2014, ובמקום השני בשנת 2013. שתי תת-קטגוריות נוספות של אופניים מדורגות גם הן במקומות הגבוהים לאורך השנים. מהנתונים שבחנו

עולה כי אופניים הם המוצר הכי יקר ביחס לשאר המוצרים ה"לא רווחיים", ויחד עם כמות המכירה הגבוהה שלו מתקבל הפסד משמעותי לחברה.
 היינו ממליצים לשים דגש על מוצרים מקטגוריה "Bikes" ולבדוק את התמחור שלהם. תיקון מחירם ישפיע על עליית רווחי החברה וימנע הפסד.

דבר נוסף שבדקנו לצרכי המלצות עסקיות לחברה הוא:
 מה היה יכול להיות הרווח לכל שנה, אם המוצרים ה"לא רווחיים" היו מתומחרים ללא הפסד (כלומר נמכרו במחיר עלות)?



מסקנה:

ניתן לראות את ההשפעה החיובית של תמחור נכון של מוצרים על הרווח השנתי, ומכאן ניתן להסיק שאם היה תמחור שהוא גם בעל רווח (גבוה מעל מחיר העלות) אז החברה יכלה להגדיל את רווחיה לאורך כל שנות פעילותה.

לאחר שאפינו את המוצרים ה"לא רווחיים" בחברה בעזרת קיבוץ לתתי-קטגוריות, רצינו להמשיך לחקור האם יש חפיפה בין המוצרים הכי "לא רווחיים" (על פי מדד Total gross margin) לבין המוצרים הכי נמכרים בחברה (מבחינת כמות מכירה).

את תוצאות חמש תתי-הקטגוריות הכי נמכרות עבור כל שנה ניתן לראות בגרפים הבאים:



ניתן לראות כי ישנה חפיפה בין נתונים אלה לנתונים שהתקבלו עבור דירוג המוצרים ה"לא רווחיים": תת-קטגוריה היא במקום השני (אך חשוב לזכור כי זו 2014, ובשנת 2011-2013 "מובילה ראשונה בשנים Road Bikes" אינה שנה מלאה). בנוסף מופיעות כאן קטגוריות נוספות כמו:

שדורגו גבוה גם במוצרים ה"לא רווחיים". Mountain Bikes, Road Frames, Touring Bikes. ייתכן וההנחות הגבוהות שניתנו לאופניים לאורך השנים משכו הרבה לקוחות לעסק (נבדוק זאת בהמשך), אך יחד עם זאת בגרפים האלה ניתן לראות מוצרים מקטגוריות נוספות כמו: כפפות, צמיגים וקסדות מדורגים גבוה. מדובר במוצרים שהם אביזרים נלווים לאופניים, זולים הרבה יותר מאופניים ורובם אינם נמכרים בהפסד. במסקנות העסקיות לחברה נתייחס להמלצה איך ניתן להשתמש בנתון זה כדי לשפר רווחים.

נושא אחרון אותו רצינו לחקור על מנת שנוכל לתת המלצות מבוססות נתונים על נושא התמחור הוא: **ההנחות**. הצגנו לעיל שיש הרבה מאוד מוצרים עם הנחות משמעותיות, והדבר גורם להפסד, אבל חשוב לנו להבין גם האם קיימת השפעה חיובית של ההנחות על הגדלת מכירות ושימור לקוחות.

ראשית חישבנו את הממוצע של אורך החיים של לקוחות שקיבלו הנחה לעומת לקוחות שלא קיבלו הנחה. **אורך חיי לקוח מחושב ביחידות של ימים על ידי הפרש בין תאריך ההזמנה הראשונה לבין תאריך ההזמנה האחרונה לכל לקוח.**

מספר הימים הממוצע של לקוח שקיבל הנחה היו 528, זאת לעומת 161 בלקוחות שלא קיבלו הנחה מעולם. מספר זה מראה לנו שלקוחות שקיבלו הנחה נשארים בממוצע 367 ימים יותר מלקוחות שלא קיבלו הנחה מעולם.

לאחר מכן בנינו פאנל שבדק לכל הלקוחות את אורך חייהם בחברה, כמות הכנסות, ועוד. לכל ערך שנבדק נבנתה פונקציה.

בעזרת הפאנל התחלנו בדיקות כלליות:

תחילה, בדיקה של מתאם פירסון:

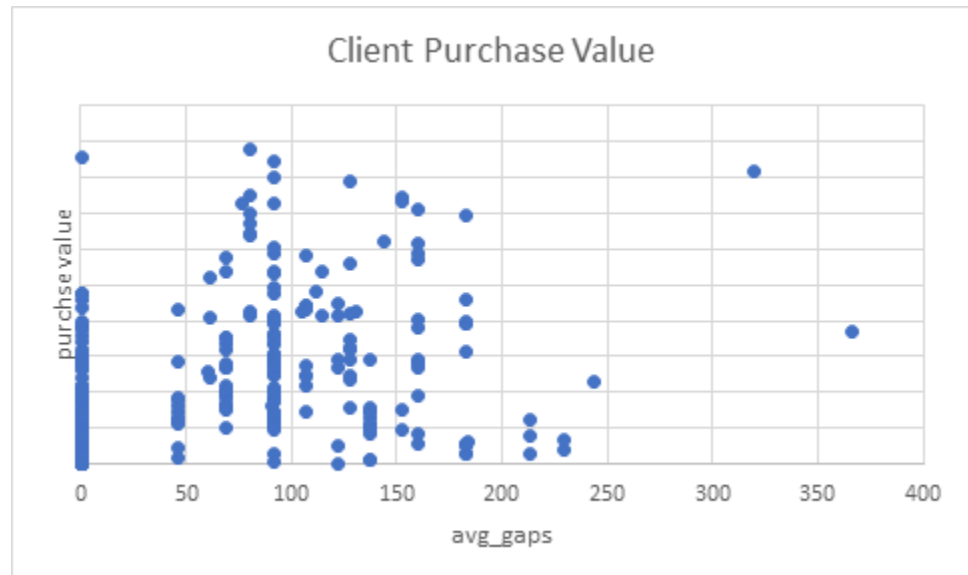
המתאם בין כמות ההכנסות לבין אורך חיי לקוח בחברה הינו 0.236344, כלומר מתאם נמוך מאוד. מה שאומר שלמרות שלקוחות נשארים יותר הם לא בהכרח מכניסים יותר כסף לחברה.

למרות כל זאת כשבודקים את ממוצע ההכנסות ללקוח של לקוחות שקיבלו הנחה, לעומת לקוחות שלא קיבלו הנחה- יש פער ניכר של 200,379 (עוגל כלפיי מטה).

מכאן מגיעה מסקנה מספר 1 בנושא ההנחות:

הנחות מגדילות חיי לקוח ואת ממוצע ההכנסות ללקוח, ולכן נמליץ לחברה לפעול להגדיל את כמות הלקוחות שמקבלים הנחות.

לאחר מכן בדקנו את כמות המרווחים בין ההנחות וכיצד זה השפיע על מדד רכישות הלקוח:

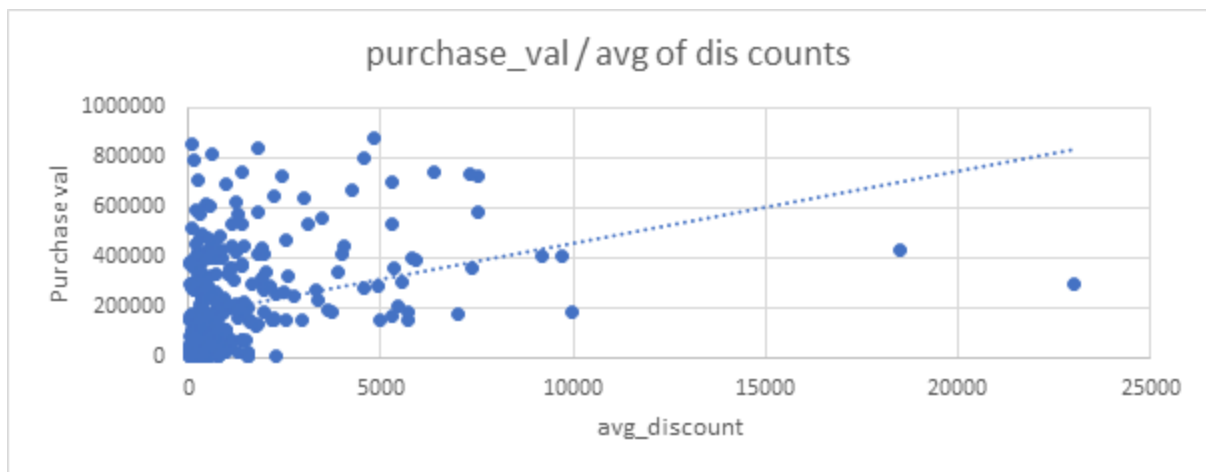


כפי שאפשר לראות הערכים גבוהים יותר בין 50-100 יום. כלומר המרווח בין נתינת ההנחות שיוביל להבאת ערך מקסימלי של רכישות הינו בין 50-100.

מכאן מגיעים למסקנה מספר 2 בנושא ההנחות:

יש לשאוף למרווח של בין 50 ל 100 יום בנתינת ההנחות ללקוח.

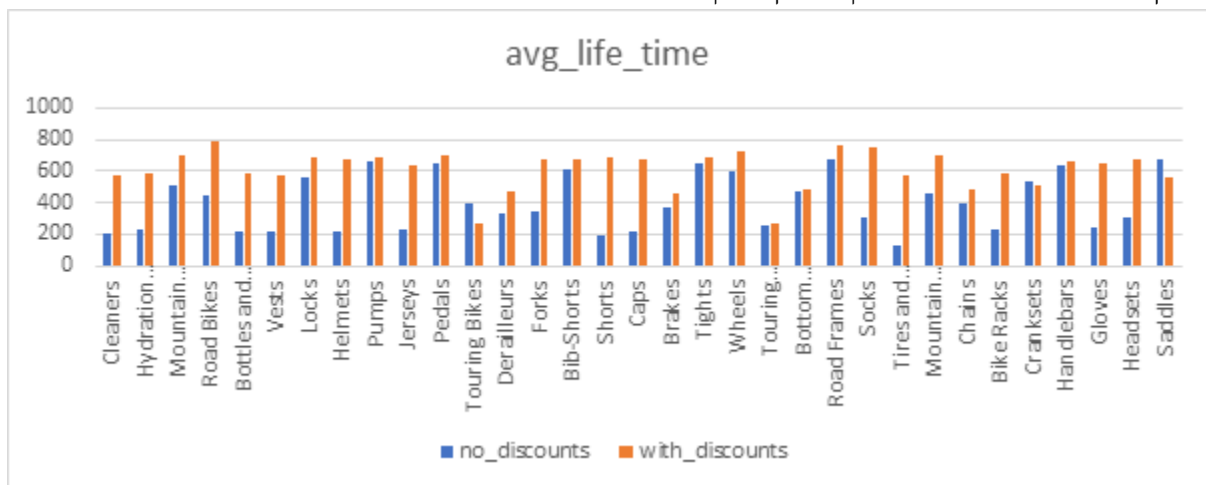
השלב הבא היה לבדוק איך כמות הכסף שחוסכים בהנחה משפיע על הכנסה מהלקוח:

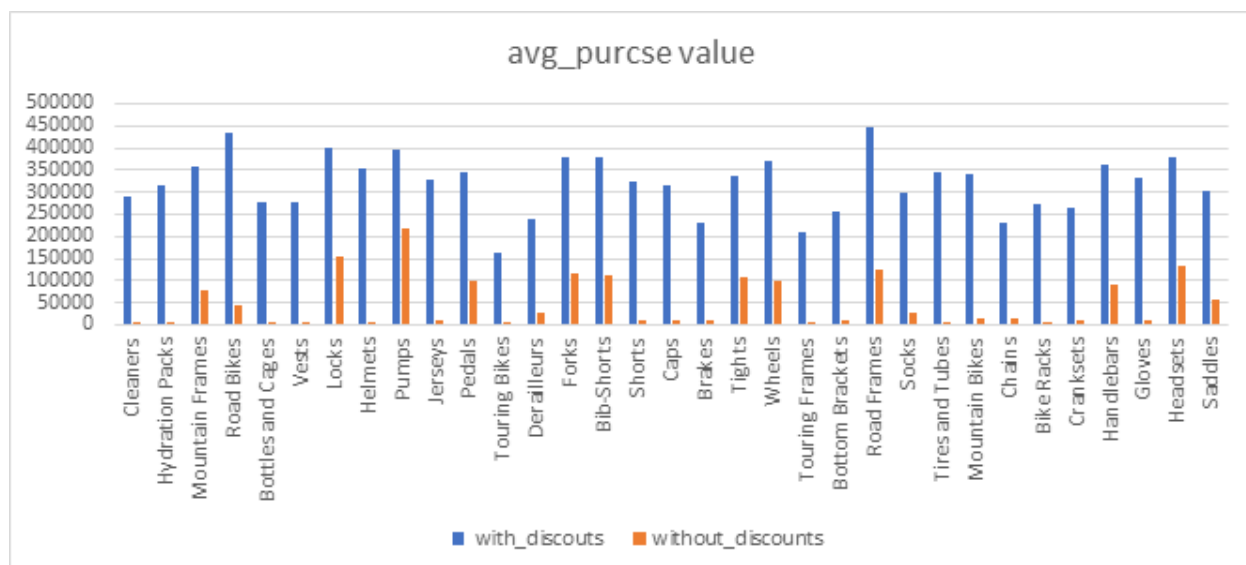


כפי שאפשר לראות מהגרף אכן יש מתאם מסוים בין כמות הכסף שהלקוח חסך בכל קניה בה הייתה הנחה בממוצע. אומנם היחס הינו חלש (מתאם פירסון של 0.32628).

מכאן מגיעים למסקנה מספר 3 בנושא ההנחות:
יש לשים דגש על עשיית ההנחה ופחות על כמות הכסף הנחסכת בהנחה.

לבסוף חישבנו את ממוצע חיי הלקוח וערך הלקוח לפי מוצרים:





בגרפים ניתן לראות איזה מוצרים בעלי הנחות מייצרים יותר לקוחות מכניסים (מבחינת חיי לקוח וגם ערך הקנייה שלו), ביחס למוצרים אחרים בעלי הנחות. כפי שאפשר לראות שני מוצרים אינם נכללים כאן: מעמדי אופניים ומגינים (Fenders), זאת מכיוון שלעולם לא ניתנה עליהם הנחה.

מסקנה מספר 4 בנושא ההנחות:

קודם כל הנתונים מראים שנתינת הנחה על המוצרים מגדילה את חיי הלקוח וגם את ערך הלקוח מבחינת כוח הקנייה שלו.

שנית, הנתונים מאפשרים לנו לקבל החלטות עסקיות על איזה מוצרים יותר כדאי לתת הנחה כדי להגדיל את המכירות ולשמר לקוחות, ועל אילו מוצרים פחות כדאי.

ניקח לדוגמה את הקטגוריה "Road Bikes" - אלה המוצרים הכי נמכרים בחברה, אך גם מדורגים הכי גבוה ברשימת המוצרים ה"לא רווחיים". בגרפים שלעיל אפשר לראות שההנחה על קטגוריה זו הגדילה משמעותית את חיי הלקוח ואת ערך הלקוח.

עם זאת ישנם עוד מוצרים, כמו ביגוד ואביזרי רכיבה, שגם הם מדורגים גבוה ברשימת "הכי נמכרים" וההנחות עליהן מעלות את חיי וערך הלקוח, אך מחירם זול בהרבה מאופניים.

אנחנו היינו ממליצים לתת פחות הנחות משמעותיות על מוצרי אופניים, ולהשתמש בכוח ההנחות דווקא במוצרים הנלווים (לפי מסקנה מספר 3 הדגש צריך להיות על עצם ההנחה ולא בהכרח על כמות הכסף שנחסכת). כך החברה תוכל להעלות רווחים על האופניים מצד אחד, ולמשוך עוד לקוחות מצד שני.

להלן מסקנות עסקיות לחברה:

1. **תיקון מחירים** - כל המוצרים אשר מתומחרים נמוך ממחיר העלות יתוקנו למחיר הגבוה ממחיר העלות, בהתאם להחלטת ההנהלה.
2. **הנחות** - החברה תמשיך להעניק הנחות למוצריה אך לא יינתנו הנחות למוצרים אם לאחר ההנחה מחיר המוצר נמוך ממחיר עלותו.
ייבחנו מקרים חריגים עבור מוצרים שלא נמכרים בכמות גבוהה או עבור חיסול מלאי סוף עונה.
נמליץ לשאוף למרווח של בין 50 ל-100 יום בנתינת ההנחות ללקוחות.
3. **קידום מוצרים שאינם אופניים** - קטגוריות כמו ביגוד ואביזרים נלווים יעמדו בראש קידומי המכירות, יקבלו הנחות ויוצגו במבצעי תחילת העונה. דבר זה יעודד את הלקוחות להמשיך לרכוש מחברה זו אופניים במקום מחברה אחרת בזכות המוצרים הנלווים במחירים משתלמים.
בנוסף, ההפסד מההנחות יקטן משמעותית משום שמדובר במוצרים שיחסית לאופניים הם זולים יותר.
המלצה עסקית נוספת בנושא זה היא להציע עסקת "חבילה" ללקוח הכוללת בכל קניית אופניים קבלת מוצר נלווה (שבהכרח יצטרך לרכיבה) בהנחה. חבילת רכישה מסוג זה תגדיל חשיפה למוצרים הנלווים, תעודד אליהם ביקוש בעתיד ותמשוך לקוחות.