**מטלה 1 – בינה עסקית**

קבוצה K

מגישים:

|  |  |
| --- | --- |
| שם | ת.ז |
| אמיתי אוירבך | 307885509 |
| רועי נדם | 311147763 |
| בן גורניצקי | 313580144 |
| אסף איסן | 313289506 |
| בר יעקבי | 203261144 |

**חלק 1**

מאגר הנתונים שבחרנו מציג עסקאות שבוצעו בחברה מסוימת. המאגר מחולק למספר קטגוריות:

0 Account 1322 non-null object מזהה לקוח

1 Division 1322 non-null object מחלקה

2 ARR 1322 non-null int64 הכנסה לשנה

3 stage 1322 non-null int64 העסקה נסגרה או לא

4 Source 483 non-null object אופן סגירת עסקה

5 Employees size 1322 non-null int64 כמות עובדים

6 Country 1322 non-null object מדינת הלקוח

7 Created Date 1322 non-null datetime64[ns]תאריך פתיחה

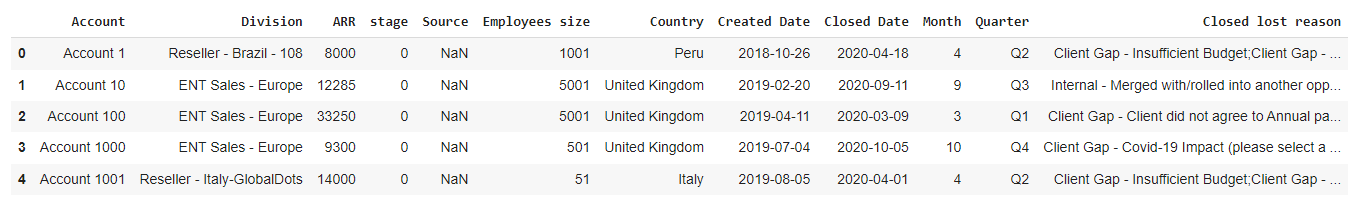
8 Closed Date 1322 non-null datetime64[ns]תאריך סגירה

9 Month 1322 non-null int64 חודש

10 Quarter 1322 non-null object רבעון

11 Closed lost reason 849 non-null objectסיבת ביטול העסקה

דגימה מתוך מאגר הנתונים:



**חלק 2**

1. שאלות המחקר:
2. **שאלת מחקר Supervised:**

האם ניתן לחזות אם העסקה תניב מעל $100,000 ARR באמצעות מדינת המקור וגודל הארגון והרבעון?

**KPI - ARR** - It is defined as the value of the contracted recurring revenue components of your term subscriptions normalized to a one-year period. ARR is the less frequently used alternative normalization method of the two common ones, ARR and MRR. It is used almost exclusively in B2B subscription businesses.

**SMART**

**S ספציפי –** מודד את ההכנסה השנתית הצפויה וכולל בתוכו רק חוזים אשר אורכם לפחות שנה.

**M מדיד –** ניתן למדידה לפי הנוסחה המוגדרת לעיל.

**A בר השגה –** ניתן לקבוע סכום שאליו אנו רוצים להגיע ולמדוד את הצלחתנו ביחס אליו, יש לקבוע את המדד בהתחשבות במדדים הנהוגים בתעשייה הרלוונטית ועבור המוצר הספציפי. במקרה זה אנו עוסקים במוצר SAAS.

**R רלוונטיות –** עבור ארגון בעל מודל עסקי של subscription ה ARR מהווה מדד חשוב עבור תחזיות לצמיחה/גרעון בשנה הקרובה.

**T זמן –** ARR כשמו כן הוא, תחום לשנה ונמדד בדרך כלל בסוף החודש.

1. **שאלת מחקר Unsupervised:**

האם ניתן לחזות את ההסתברות לסגירת עסקה בהינתן הפרמטרים - Country, ARR, Employee Size?

**KPI - Win rate** - the number of deals won out of the total number of sales opportunities.

**SMART**

**S ספציפי –** יעד זה הינו ספציפי מאחר והוא מודד את כמות העסקאות מסוימות שבוצעו.

**M מדיד –** ניתן למדוד האם שיעור העסקאות שהצליחו עלה או ירד לפי הנוסחה שהגדרנו.

**A בר השגה –** באמצעות גורמים מסוימים שיבוצעו ניתן לשיפר מדד זה.

**R רלוונטי –** שיעור העסקאות שבוצעו בהצלחה רלוונטי למדידת הצלחת החברה והגדלת ההכנסה.

**T זמן –** ניתן למדוד שיעור עסקאות מוצלחות בפרק זמן מוגדר מראש.

**חלק 3**

1. מדדי פיזור של עמודה "ARR"(הכנסה שנתית צפויה):

**count 1322**

**mean 14818**

**std 20203**

**min 76**

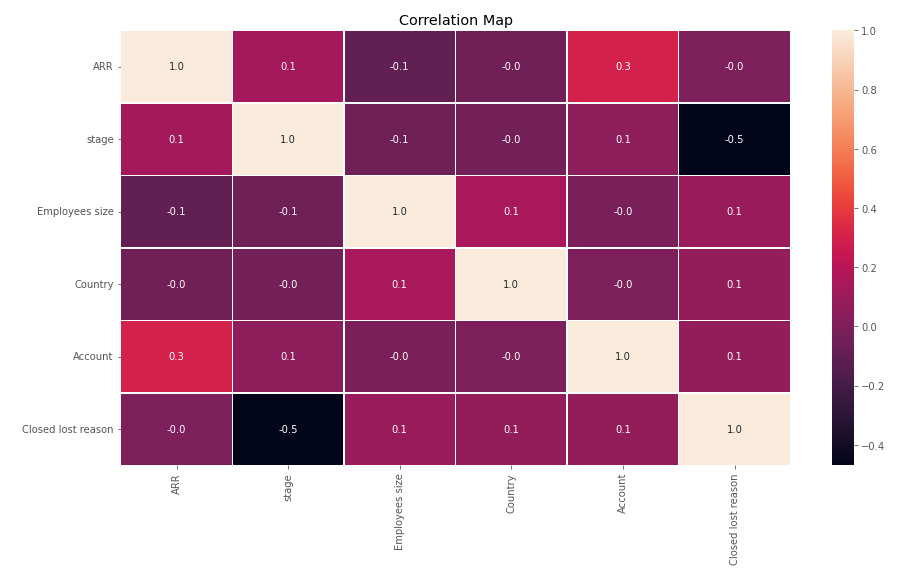
**25% 3000**

**50% 10000**

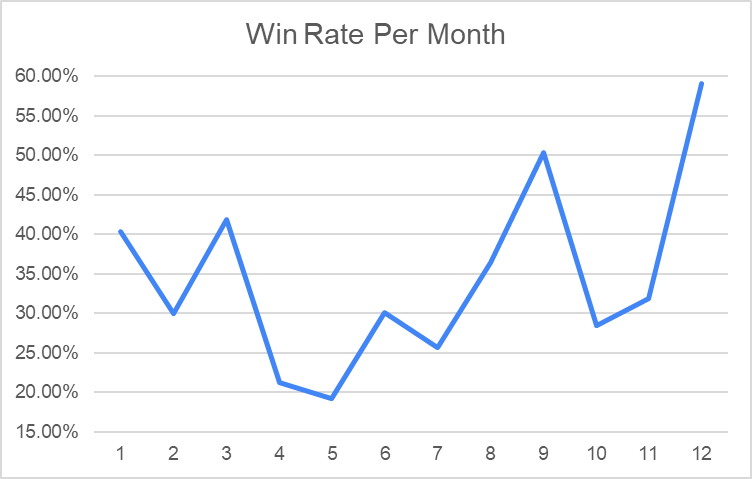
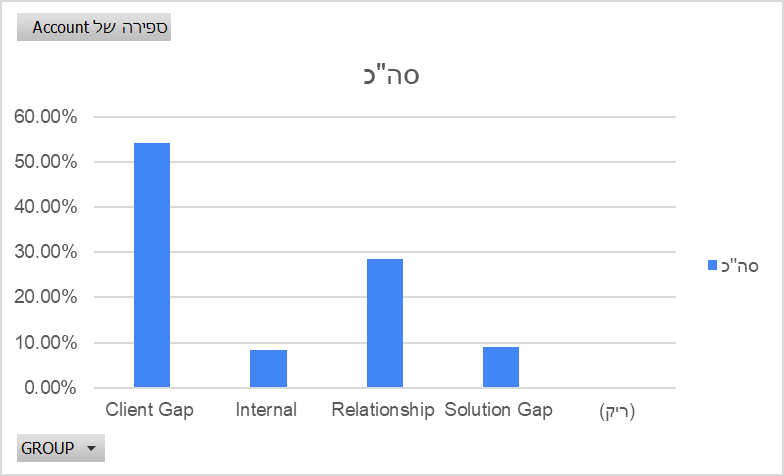
**75% 18500**

**max 290000**

**Name: ARR, dtype: float64**

1. ****קורלציה
2. מדדים סטטיסטיים נוספים:

שיעור כמות סגירת עסקאות פר חודש שיעור הסיבות בגללן לא נסגרה עסקה

****

1. מדד האנטרופיה עבור התכונות של השאלה ה- supervised:

ARR – 7.209

Month – 3.544

1. Gini-index (ARR) = 0.982

Gini-index (Month) = 0.912

Information – Gain (ARR, Month) = 1.004

1. ניתן להסיק מתחקור הנתונים כי מרבית הסיבות בגללן לא נסגרה עסקה הן עקב פערים מצד הלקוח. בנוסף, ניתן לראות לפי הגרף שהוצג כי ברבעון השני הייתה ירידה חריגה בכמות סגירת העסקאות.