



Leveraging Data Science in Market Research

מחקר שיווקי
תובנות IL



הצגה לקבוצת:

יוני 2020

נתחיל בסיפור אישי

אפשר לומר שנולדתי לתוך התעשייה של מחקרי שוק

ההורים שלי הקימו מכון מחקר בשנות ה-90...

כילד – הייתי מבלה במעבדות המחקר של אוניברסיטת חיפה

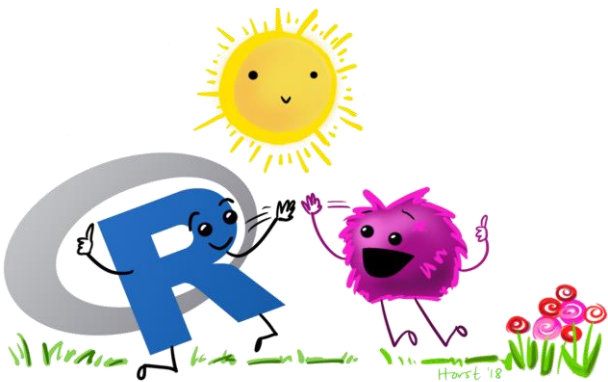
כנער – הייתי עוזר במשרד בהדפסה (או גריסה) של שאלוני Pen and Paper

כמתבגר (עתודאי) – הלכתי ללמוד מתמטיקה עם סטטיסטיקה וחקר ביצועים

(ואז תואר שני בחקר ביצועים, ואז תואר שלישי)

ובסוף יצא (כמה לא מפתיע) שאני עוסק ב-Data Science ושותף במכון מחקר

אבל – מספיק לדבר עלי. בואו נדבר עליכם, ועל Data Science (ולמה זה צריך לעניין אתכם)



מהם האתגרים המשמעותיים ביותר של תעשיית מחקרי השוק? (רשימה חלקית)

(רשימה חלקית)

שיטות

- הרבה שיטות מקבילות (כלים רבים, קשה להתמחות ולהתעדכן)

- שיעורי היענות נמוכים

לקוחות

- מצפים לתובנות אתמול בבוקר (עלול לבוא על חשבון האיכות)

- תקציבים – רוצים הרבה במעט (מובילים לפתרונות DIY לעיתים גם על חשבון איכות)

- מתקשים לנסח את הרצונות שלהם (והרצונות משתנים ללא הרף)

מהם האתגרים המשמעותיים ביותר של תעשיית מחקרי השוק? (רשימה חלקית)

תוצאות

- קשה להפוך אותם ל-Actionable insights
 - עסקים רצים מהר וקשה להדביק את הקצב (תוצאות עשויות להיות לא רלוונטיות) טכנולוגיות וחדשנות
 - לקוחות מושפעים מ-Buzz words כמו Data Science או AI (אבל לא באמת מבינים מה זה אומר)
 - להציג בידול, להביא דברים חדשים, להמציא מחדש
- ועוד ועוד...

עד סוף המפגש נענה על:
1. איך ובמה Data Science נותן מענה
2. איך אתם יכולים להפעיל Data Scientists בשביל לקבל את המענה

אז מה זה בעצם Data Science?

אין באמת תשובה, זה איפשהו בין... סטטיסטיקה, מדעי המחשב (תכנות, מחקר, ועוד...) (אין באמת תשובה, זה איפשהו בין... סטטיסטיקה, מדעי המחשב (תכנות, מחקר, ועוד...))

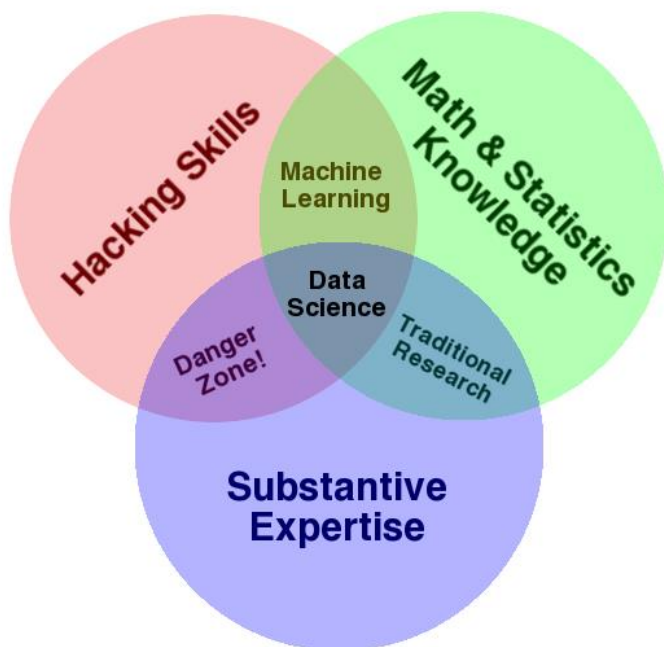
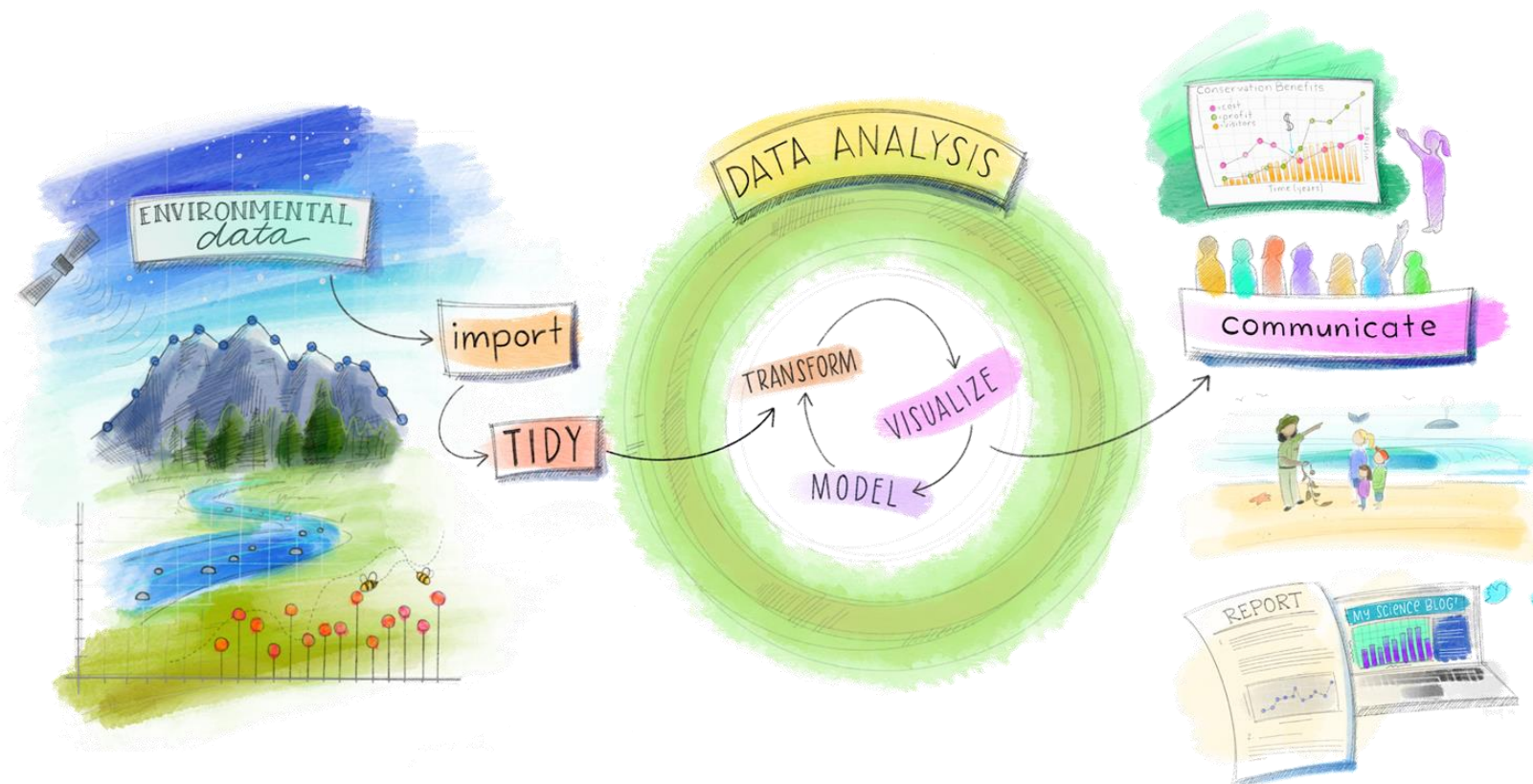


Figure 1a: Conway's Data Scientist Venn Diagram

ככה נראה "יום העבודה" של Data Scientist

אולי יראה מוכר, כי זה בעצם לא מאוד רחוק מהעבודה של חוקר שוק... (אבל עם טוויסט)



אבל באיזה כלים?

ניסוי מחשבתי. באילו כלים אתם משתמשים?

<https://app.sarid-ins.com/s3/mr-tools>

תוצאות

הנה מה שאני משתמש בו, לפי תדירות (לפעמים זה יותר מזכיר פיתוח תוכנה מאשר מחקר שוק):

Python/R <- github <- פלטפורמה לסקרים/מאגרי מידע <-
מצגות <- כונן משותף <- וורד, אקסל

מה זה אומר? שהעבודה היום-יומית שלי היא איפשהו באמצע בין מחקר לבין פיתוח

מה עושה Data Scientist? (רשימה חלקית)

מחבר נתונים מהרבה מקורות

dplyr : go wrangling



מה עושה Data Scientist? (רשימה חלקית)

מסדר את הנתונים

`dplyr::relocate()`
1.0.0
move columns around!

Default: move to FRONT
or move to
.before or .after
A SPECIFIED COLUMN!



@alison_horst

מה עושה Data Scientist? (רשימה חלקית)

מסאג' לנתונים



מה עושה Data Scientist? (רשימה חלקית)

בונה משתנים



מה עושה Data Scientist? (רשימה חלקית)

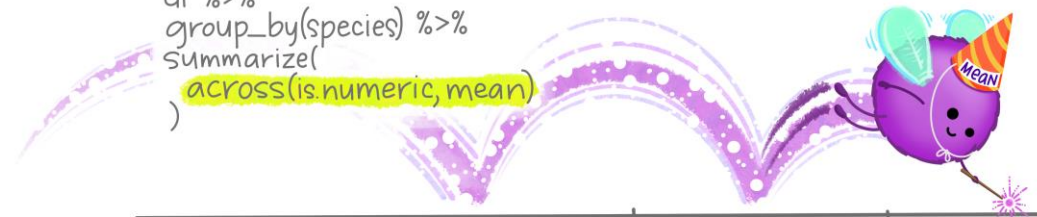
עושה טרנספורמציות

dplyr::across()
v.1.0.0

use within `mutate()` or `summarize()` to apply function(s) to a selection of columns!

EXAMPLE:

```
df %>%  
  group_by(species) %>%  
  summarize(  
    across(is.numeric, mean)  
  )
```



species	mass_g	age_yr	range_sqmi
pika	163	2.4	0.46
marmot	1509	3.0	0.87
marmot	2417	5.6	0.62

@allison_horst

מה עושה Data Scientist? (רשימה חלקית)

ויזואליזציות



מה עושה Data Scientist? (רשימה חלקית)

בונה מודלים



a standardized interface for fitting models
and making predictions.

מה עושה Data Scientist? (רשימה חלקית)

מציג הכל באמצעות.... דוחות (ועוד – מיד נדגים)



עד כה התייחסנו ל-Data Scientists

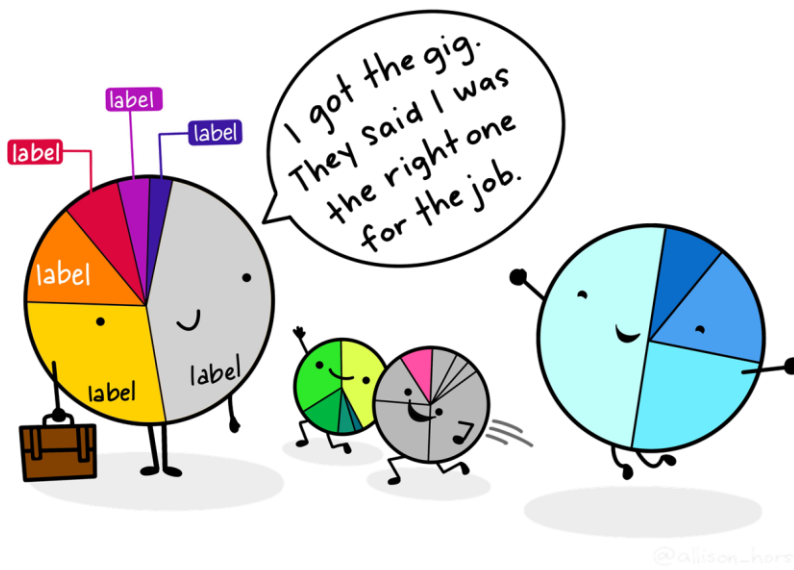
יש גם בעלי תפקידים נוספים, שלפעמים מבלבלים אותם:

Machine learning engineers •

Data analyst •

BI •

ועוד...



תפקידים שונים, עם תחומי אחריות שונים ומיומנויות שונות

מה הרקע של Data Scientist?

- רקע אקדמי
 - בדרך כלל תואר שני או שלישי
 - מקצוע אנליטי (סטטיסטיקה, מתמטיקה, מדעי המחשב, יכול להיות גם פיזיקה, או מקצועות הנדסיים)
- ניסיון נדרש
 - כדי ש-Data Scientist יצליח לשתף פעולה על פרויקט רצוי ניסיון של 5 שנים לפחות בתחום שקשור ל-Domain (או הנחיה מקצועית של חוקר בכיר יותר)
- כלים
 - כלי תכנותי-סטטיסטי (הנפוצים: R או Python, הולך ונעלם: SAS, SPSS)
 - כלים נוספים - כתלות בעסק ובאופן הפעילות

מתי לשלב Data Scientist בפרויקט?

החל מניתוח נתונים פשוט של סקרים "פשוטים"

היתרון הגדול של Data Scientists יבוא לידי ביטוי כאשר:

- מקורות נתונים מרובים שצריך לחבר ביניהם

- עבודת סידור נתונים משמעותית

- מודלים מורכבים (Clustering, Conjoint, מודלי חיזוי)

- כשרוצים להבטיח שקיפות (Transparency), ויכולת שחזור בעתיד (Reproducibility)

- כשרוצים אוטומציות (תהליכים שקורים הרבה, שיחזרו עליהם בעתיד)

- כשצריך לשלב את התובנות או המודלים במערכות Production

מה התוצרים שניתן לדרוש?

סט היכולות של Data Scientist מאפשר לו להפיק כמעט כל דבר

למה? כי כשעובדים עם כלים כמו R ו-Python כל הכלים כבר קיימים בצורת חבילות, רק צריך לדעת למצוא אותם

חשוב של-Data Scientist תהיה יכולת להבין את הכלים, המגבלות שלהם, ולהתאים אותם למצב הרלוונטי (כאן הרקע האקדמי נכנס לעניין)



ממה להיזהר?

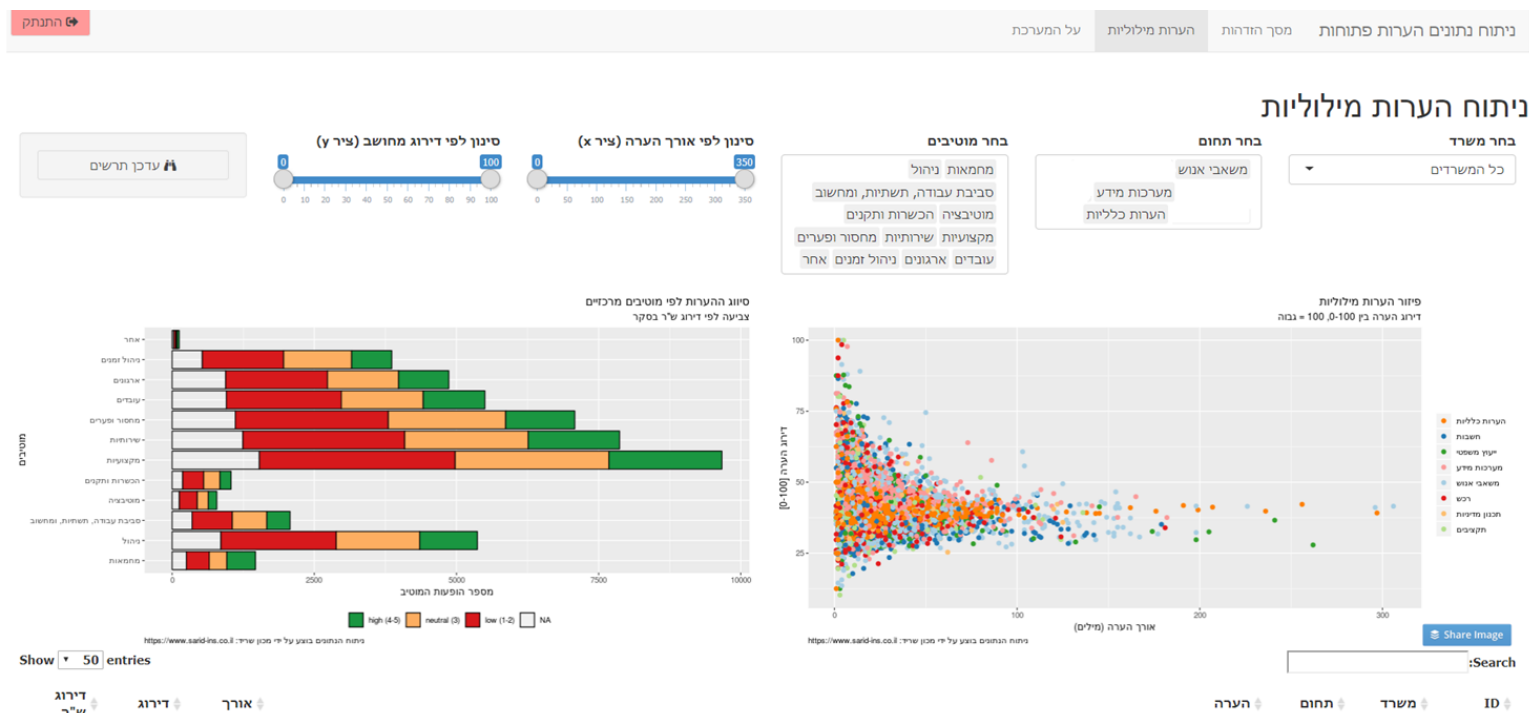
- ל-Data Scientists (במיוחד מתחילים) יש נטיה "לסבך את העניינים"
- צריך להקפיד על התאמת הרמה לקהל היעד:
- למקבלי החלטות עם קשב נמוך, ועומס קוגניטיבי - פשטות, על חשבון העמקה
- למקבלי החלטות אנליטיים, כמותיים - אפשר להעמיק למודלים ומשמעויות, להיכנס לפרטים



דוגמאות מגניבות מהשנים האחרונות...

דוגמה 1: ניתוח תוכן הערות מילוליות

- הערות מילוליות מאוד נפוצות בסקרים, וגם באפיקי תקשורת נוספים מול הלקוח (Social, Product reviews)
- "השיטה הקלאסית" במחקרי שוק היא לעבור על ההערות ולסווג אותן לקטגוריות. לוקח הרבה זמן, לא Scalable.
- בפועל – באלגוריתמים של Data Science ניתן לחשב Sentiment, לבחון קשרים, לחלץ מוטיבים אוטומטית ולתת למקבלי החלטות כלים נוספים לחקור את ההערות, מעבר ל"קובץ אקסל" או תרשים סטטי

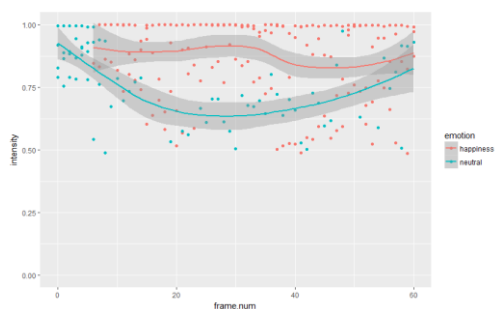


דוגמה 2: ניתוח הבעות פנים בוידאו

איך מודדים אפקטיביות של פרסומת במדיה (לפני שידור)? בשיטות הקלאסיות – צפיה ודיווח עצמי (מענה על שאלות)

כלים כיום מאפשרים לנתח הבעות פנים (יש מספר חברות שמציעות את הטכנולוגיה "כמוצר שלם")

https://sarid.shinyapps.io/MR_faceR/



Facial coding וסיווג לרגשות (שמחה, כעס, נייטרלי, פחד, גועל, ...).



הפרסומת:

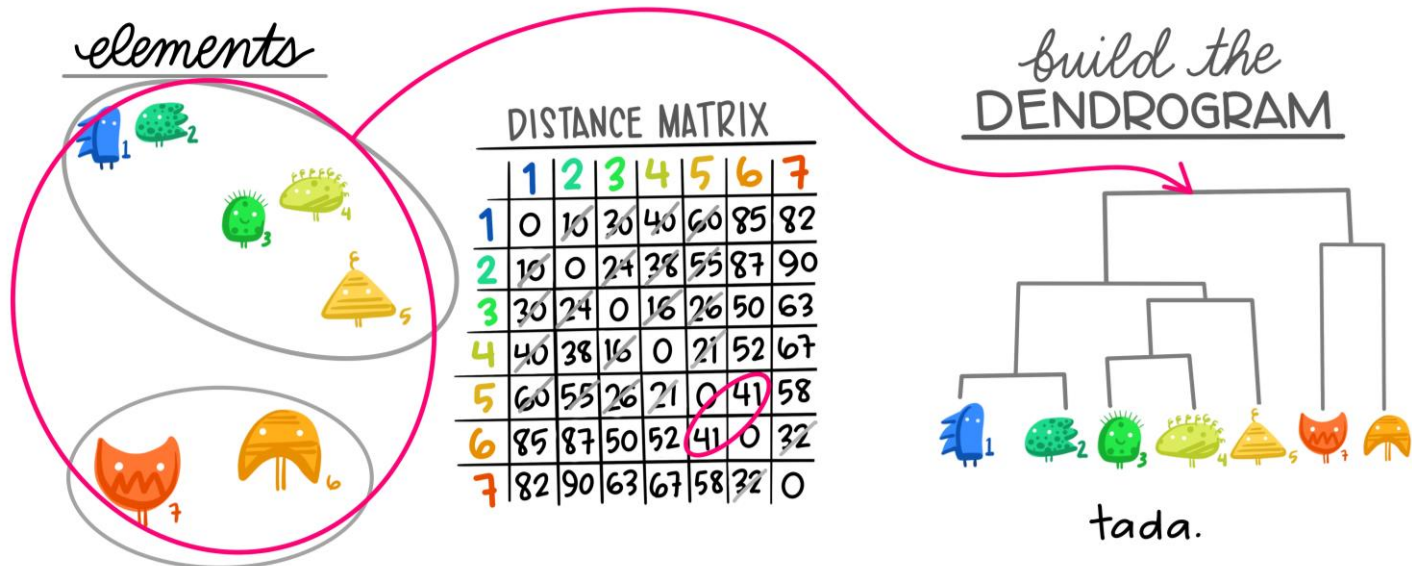


הצופים:

דוגמה 3: Clustering

סגמנטציה באמצעות Clustering – סיווג לקוחות לקבוצות על בסיס פרמטרים שנאספו (בסקרים או Database של החברה).

מאפשר למקבלי החלטות להגדיר טרמינולוגיה, להתמקד בקבוצות מסוימות, ולהכווין אסטרטגיה



דוגמה 4: אוטומציה, יצירת דוחות ב-Batch

בממשקי Point and click (אקסל, SPSS) קשה לייצר מסה של דוחות לעומת זאת...

כשכותבים את הדוח בקידוד, המאמץ דומה $1000 \approx 100 \approx 10 \approx 1$...

תוצר יכול להיות מאות מצגות, מאות מסמכים, או דשבורד

מפולחים לפי תתי-אוכלוסיות, מחלקות, או כל חלוקה ארגונית אחרת רלוונטית



דוגמה 5: ניתוח מרחבי, אופטימיזציה של מיקום מרכזי שירות

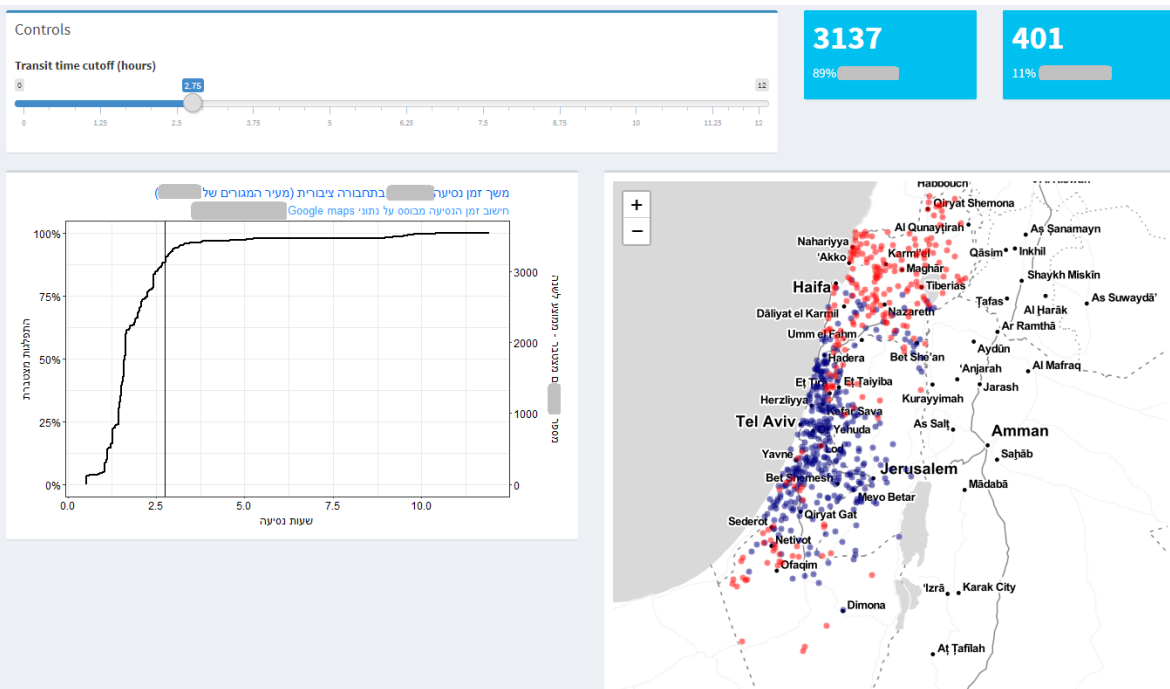
אינטגרציה של מקורות מידע פנימיים וגלויים מאפשרים לבחון פריסה מרחבית של שירותים שונים

מחד – מקום מגורים של הלקוח

מאידך – מודל לחישוב מיקום אופטימלי של מרכזי השירות (מקסום הסבירות לרכישה למול עלויות)

מרחקי נסיעה – מחושבים באמצעות API של google maps על פני אלפי תצפיות

מקבלי החלטות – מקבלים אפליקציה לבחינת Cutoff שונים, מיקום מרכזים שונים והמשמעויות שלהם על הרווחים



והרשימה ממשיכה...

Web scraping וניתוח תוכן

חיזוי מכירה

חיזוי נטישה

Time Series Analysis

ועוד...

איך מתניעים פרויקט Data Science? צ'ק ליסט

מרימים את הטלפון ומתקשרים אלי (סתתםם...)

התדריך שתיתנו ל-Data Scientist צריך לכלול:

1. מטרת הפרויקט, היכרות עם עולם התוכן של הלקוח, Context כללי
 2. מחקרים שנעשו בעבר בהקשר של הפרויקט הנוכחי, נתוני עבר (האם נדרשות השוואות? היכן נמצאים מאגרי נתונים קודמים)
 3. מידע על אופן איסוף הנתונים בפרויקט (סקר? דאטה פנימי? ריבוי מקורות? בעיות, מגבלות)
 4. ולידציות שנכנסו בפנים (לדוגמה ולידציות בסקרים על משך מענה, שאלות מלכודת)
 5. תוצרים רלוונטיים נדרשים, שאלות מחקר, לוחות זמנים
- שיחות שוטפות עם ה-Data Scientist, להיות מעורבים, לבקש ללמוד, לשאול הרבה שאלות, להכיר לו את עולם התוכן.

לקריאה נוספת

- <https://www.sarid-ins.co.il/blog> הבלוג של מכון שריד
- adisarid.github.io הבלוג הפרטי שלי (קצת יותר טכני)
- [The Rstudio blog](#)
- (אם אתם מעוניינים ללמוד R – פנו אלי, קורסים מעת לעת)



Thanks!

Adi Sarid adi@sarid-ins.co.il

@SaridResearch 

<https://www.sarid-ins.co.il>