תפקידים בעולם הדאטה ושאלות מראיונות - אבי לוסקי:  
 הרצאת "אקדמיית הפניקס" של "כולא לייק"

בהרצאה הזו נעשה סדר במגוון סוגי התפקידים בעולם הדאטה, ונבחן כמה שאלות לדוגמא.

שמי אבי ובאופן כללי אני משתדל ללמוד כל הזמן דברים חדשים, בעיקר באופן עצמאי.

בין השאר יצא לי להיות מוביל מוצר אנליטי בסטארט אפ, כשחלק ניכר מהעיסוק כלל גרפים ולמידת מכונה עם BI.

גם אני הייתי ג'וניור בדיוק כמוכם.

אם אתם סקרנים ורעבים, אנחנו בעידן בו אפשר ללמוד הכול באינטרנט, הכול נגיש וזמין,   
זו תקופה נהדרת להעשרת הידע.

חלק מהתפקידים בעולם הדאטה חופפים ופשוט נקראים אחרת בין חברות, חלק מאוד דומים.

כל עולם הדאטה הוא תהליך אחד ארוך ומתמשך, המתחיל בנקודה מסוימת וממשיך לאחרת.

חברות מאוד משקיעות בהפקת תובנות מהדאטה ההולך וגדל היום.

**סוגי תפקידים:**Data scientist

יש מהם שיותר מתעסקים בresearch, ויש שפותרים בעיה עסקית ע"י כלים של data science.

זה תפקיד שיכול להיות מאוד רחב בסטארט-אפים וחברות קטנות, אך מאוד מצומצם בחברות גדולות.

בד"כ מתעסק במודלים היותר מתקדמים, לכן מומלץ שיהיה רקע באנליזה כדי להבין טוב יותר את שלבי השרשרת ועם זה להביא תוצאות מעניינות יותר.

Data analyst ו Business analyst  
בהרבה חברות זה כמעט אותו תפקיד, משתנה לפי הגודל והתשתית איתם הן עובדות.

אם החברה שעובדת עם data base אחד, אז לרוב הם יהיו באותו צוות או שיעשו את אותה העבודה.

כדי שבכל זאת יהיה קל לזכור את ההבדל ביניהם, נתמקד בשם שלהם.

data analyst עסוק הרבה בדאטה עצמה,

וה- business עסוק בלאסוף דרישות מהלקוחות ולעשות ולידציה (תִּקּוּף) איתם.

זו בגדול החלוקה, הskill set שלהם דומה ולרוב הם ידעו אותו דבר.

הרבה פעמים business analyst בא מעולם הביזנס כuser בעצמו.

data analyst הוא figure it out kind of a guy, מקבל בעיה עם דאטה ופותר אותה.

בד"כ אם יש הפרדה בין התפקידים, פחות יהיה להם ממשק עם הלקוחות

business analyst יכול גם לעבוד באקסל, בד"כ כמעט ולא כותב קוד, בד"כ יעבוד על כמה פרויקטים במקביל.

סוג של ביזנס לצד הIT (**I**nformation **T**echnology).

יעבוד על אפיונים יותר טובים לצוות ודיבור עם יוזרים עסקיים שיסבירו את הצורך העסקי שלהם.  
לדוג', אני צריך ניתוח עזיבה של עובדים. הוא צריך להיות הגשר בין הit לבין הביזנס.

סטטיסטיקאי

יש חברות שעדיין מגייסות סטטיסטיקאים. יש חפיפה מסוימת ל data analyst ו business analyst.

הרבה פעמים הם יעבדו עם כל מיני כלים סטטיסטיים כמו Matlab, המכילים מודלים פחות נפוצים.

בחברות ענקיות כמו ביטוח, יש מחלקות סטטיסטיקה שבונות היפותזות ראשוניות ומייצרות מודלים.

משם תתחיל שרשרת שלמה של פיתוח ואנליזה עם גורמים אחרים בחברה עד הגעה לproduction.

ברוב המקרים יגייסו אדם עם רקע חזק בסטטיסטיקה.

Data/business manager   
אחרי הhype הגדול עם עולם הדאטה, הגיעו לתובנה שהכי קשה לגייס מנהלים בתחום הזה.

כמובן, הפעם מדובר בתפקיד שמצריך קצת קילומטראז' לפני שמגיעים אליו.

הייחוד של התפקיד זה הוא ההיכרות שלו מבפנים של תפקידי הדאטה/ביזנס אנליסט והפעם צריך לנהל צוות, לעמוד בתקציב וכו'. כמובן, זה ניהול מקצועי ולא hr-י.

תפקיד חשוב והכרחי להעצמת הצוות, קביעת מדיניות..

בד"כ מאוד מעורב בהחלטות ארגוניות כי הוא מתווך בין האנליסטים עצמם לארגון.

DBA

נזכיר שוב שעולם הדאטה הוא תהליך ארוך.

הDBA אחראים לתקינות מסדי הנתונים בהם אנו משתמשים.

יש היום הרבה סוגי data bases, רלציונים/לא רלציונים, big data כמו Adobe.

התשתיות הרבה פעמים מאוד מורכבות.

הmind set - קודם כל מניעת אסון, כמו נפילת גיבוי בגלל תקלה, תאימות חומרה, שדרוגים, גרסאות.

שהדאטה שלנו יהיה בעל כושר עמידות וספיגה - resilient.

אם יש אירוע לא צפוי, אפשר לעבור למסד אחר.

בד"כ מדובר בחבר'ה מאוד טכניים שבקיאים בתחום, ואנליסטים יכולים להתייעץ איתם איך לשפר עניינים שקשורים למסדי נתונים, כמו הפיכת שאילתות למורכבות יותר או שאלות הקשורות לdata base. הם פחות יהיו עם קשרי לקוח, כי זו עבודה מאחורי הקלעים.

Data engineer

תפקיד יחסית חדש.  
לרוב, נתונים צריכים לעבור שלבים רבים ושונים:   
תרגומים, הסבת פורמטים, המרות, איחוד מערכות למערכת אחת..

לכן לתפקיד זה יידרש אדם עם הרבה ניסיון פיתוחי. יכולים להשתמש בו בהמון כלים.

ברגע שמבנה הנתונים הוא מורכב אז ה data engineerהוא מאוד מרכזי.

תפקידו לבצע את הטרנספורמציה בצורה הכי יעילה וחכמה.

גם בחברות קטנות בהן יש התעסקות עם ממשקים, יהיה מהנדס מידע שאחראי על הpipe line (תהליכי עיבוד הנתונים) של הדאטה. בד"כ גם הם טכניים, בקיאים ב data warehouse(מחסן הנתונים).  
רוב הEcosystem (מערכת העבודה) תהיה big data, כנראה Adobe שהזכרנו.

בחברות גדולות מאוד זה יכול להיות גם מאות אנשים בתפקיד זה.

Data architect

תפקיד סופר קריטי בארגונים גדולים.

מטרתו בגדול היא להיות אחראי על השרשרת שהמידע עובר.   
התובנות, הפתרונות הטכנולוגיים של איזה מסד נשתמש.

יכול להיות שבתהליך אחד יהיה לנו ביג דאטה ובשלבים אחרים נעבור לnosql.

הוא רואה את כל התמונה ומחליט.

Machine learning engineer   
תפקיד פופולארי בשנים האחרונות.

ההבדל המהותי שלו הוא שבתפיסה הוא יותר engineer.

פותר יותר בעיות הנדסיות מאשר אנליטיות.

אחראי על הpipe line של מודל, לאמן, לבדוק שמגיע לתוצר/production, לבדוק שהקוד מתוחזק.

אז מה ההבדל בין data scientist לengineer?

היכולת של ה scientistלעבד מידע בכל מיני צורות.

מהנדס פחות יתעסק במה יעשו עם הדאטה או אילו תובנות יהיו ממנו.   
הוא מאחורי הקלעים יתעסק עם דאטה "מלוכלך" ויציף בעיות הקיימות בו עוד לפני שנגעו בהם.

לאיזה תפקידים אפשר להיכנס בקלות יותר בלי ניסיון?

מן הסתם, תפקידי ניהול או ארכיטקטורה פחות רלוונטיים בתור התחלה.

ברוב המקרים זה מאוד חשוב שתהיו כנים עם עצמכם - איזה סוגי בעיות יותר מעניינות אתכם?

יש כאלה שיותר הנדסיים בתפיסה, יש הרוצים לפתור בעיות טכנולוגיות,

ויש כאלה שדווקא יותר מעניינת אותם האינטראקציה עם הלקוחות, לקבל דרישה עסקית ולנסות לפתור אותה.

נכון, לפעמים בהתחלה קצת קשה להבין מה מעניין אותך.

ברוב המקרים דאטה/ביזנס אנליסט או Data scientist ייתנו אופציה לג'וניורים.

חשוב שיהיה מישהו שיעזור לכם להיקלט.

חברות קטנות זזות מהר יותר ובד"כ לא מסוגלות להשקיע בזמן חניכה והטמעה של ג'וניורים.

חברות גדולות יכולות לאפשר את זה וברוב המקרים זה משתלם להן.

האם מנסים ללמוד על טעויות העבר, או לחזות/לנבא את העתיד?  
זה תלוי במטרה של החברה. יש המנסים מהעבר ללמוד על העתיד.

לדוגמא, ההתעסקות של חברת ביטוח עם data היא אדירה ואנשים שעובדים שם עוסקים בכמויות גדולות של מידע, איתן ניתן לבנות אנליזות, credit score (דירוג אשראי), למי לתת הלוואה ולמי לא.

במקומות אחרים יתמקדו בלהבין מה היה בעבר. זה מאוד משתנה בין חברות.

מה מראיין מחפש אצל מועמד?

Diagram

Description automatically generated

*1. מוּכָנוֹת ללמוד ללא הפסקה.   
2. האומץ לתת המלצות ושיפור/שימור למקום העבודה.  
3. חשיבה זריזה ויכולת הטמעה טובה.  
4. הבנה עמוקה של טכנולוגיה דיגיטלית.  
5. היכרות של המודל העסקי.  
6. ידע רחב היקף.*

*7. היכולת לשאול את השאלות הנכונות.*

באופן כללי, בנאדם אינטראקטיבי, שמתמקד בבעיה יותר מהפתרון,  
כי זה יעזור לו לבנות את התשובות בצורה טובה ומסודרת יותר.  
ישאלו על האישיות שלכם, יבקשו הדגמה של מה שלמדתם,   
תדגישו מה שלמדתם לבד, כי זו עבודה של למידה non stop.

מה התהליך של Data analysis?

A picture containing text

Description automatically generated

ירצו להבין אם אתם רואים את כל התהליך. באופן כללי – איסוף נתונים, ניתוח שלהם ויצירת דו"ח.

יכולים להציג להם שאלת open ended ללא מיקוד, כדי לראות אם אתם יודעים לענות בצורה ממוקדת, להתחשב במה שחשוב. לא להיות picky (בררן) באזורים יחסית מינוריים, כמו סוגי פונקציות וגרסאות.

תנסו לתת דוגמא ממשהו שעשיתם/התנסיתם בו. לדוג':

"בפרויקט גמר שלי/בקורס x עשיתי פרויקט ניתוח בו קודם כל אספנו וטייבנו נתונים, ביצענו אנליזה, בנינו dashboard (גיליון הצגת נתונים)"

להשתדל להיות תמציתי ולהתמקד במה שחשוב, מקס' אם המראיין ישאל ספציפית אז תענו.

מה אתה מבין מהמושג התפלגות נורמלית?

Chart, histogram

Description automatically generated

דוגמא לשאלות בסטטיסטיקה.

יכולה להיות שאלה בסיסית כמו מה ההבדל בין ממוצע לחציון?  
במה היית משתמש במקרה כזה או אחר?

מה זה בדיקות ?A/B

Text

Description automatically generated

אם רוצים לבחון feature שעולה למוצר, אם מוסיפים נניח לחצן חיווי במקומות שונים בדף, ישחררו כמה גרסאות שונות וינסו להשוות בין קבוצה אחת לאחרת.

לכן, אחת מהן תהיה קבוצת הביקורת (A),הקבוצה הקיימת היום, ואילו השנייה (B) תהיה עם השינוי.

באיזה joins (סוגי איחודים) להשתמש בכל אחד מהמקרים?

Chart, bubble chart

Description automatically generated

שווה ללמוד מדוגמאות ולשחק עם הנתונים כדי להבין.

Data cleansing

ברוב המקרים הדאטה מגיע מלוכלך (תאים ריקים, תווים לא מפוענחים, בין 0-1..) וצריך לטייב אותו.

בריאיון ישאלו כמעט תמיד על פרויקט לדוגמא. כדאי שיהיה לכם אחד כזה, אפשר להכין מראש.

ירצו לדעת מה הייתה המטרה, האם הובלתם אותו, מה היה התפקיד שלכם, האם הייתם צריכים להתייעץ עם אנשים נוספים..

תשתדלו לא להגזים, כי קל למראיין טכני ומנוסה לצלול לפרטים קטנים.

ירצו לדעת באיזו טכנולוגיה השתמשתם, ולמה דווקא בה.

תתארו את המציאות, ותדעו את הפרטים.

המועמד לא מכיר את החברה ומה היא עוסקת, אז יכולים לשאול אותו שאלה שרלוונטית לכולם.

מה היית מנתח בתנועות האשראי שלך?

זה תרגיל מצוין כדי להבין איפה יש חורים ואיפה צריך להשתפר.

לי היא יכולה להגיד די הרבה על הבנאדם שהתמודד איתה.  
בהצלחה לכולם! 😊

**תודה רבה לאבי לוסקי ושמעון מויאל!**