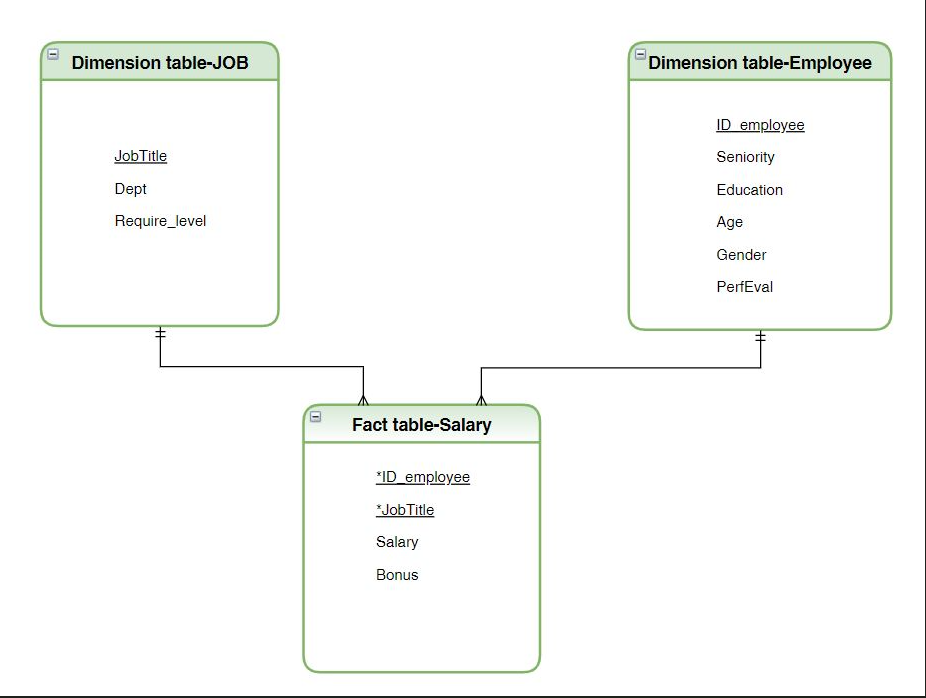
Homework 2 – Business Intelligence

**חלק 1-**נפתח פרויקט משותף בGitHub ובוצעו כלל המשימות.

**חלק 2-**

1. הסכמה שאנו בוחרים לתהליך הETL הינה: Star. יש לנו טבלת נתונים יחידה ופשוטה ולכן לא נדרשת בחירת סכמה מורכבת. סכמת star נבחרה בשל הפשטות, יכולות הפרדת הנתונים והגמישות שלה, התורמים לעיבוד נתונים יעיל, ניהול ותחזוקת המערכת.
2. Data Warehouse:



A picture containing text, screenshot, font, number

Description automatically generated

1. Use-case המתאר מדוע בחירתנו נכונה לשם קבלת החלטות למשתמש:

A picture containing diagram, sketch, drawing, line

Description automatically generated

תרשים הUSECASE שהצגנו יעיל ועוזר בקבלת החלטות המשתמש מהסיבות הבאות:

* בהירות שלבי התהליך: התרשים מתאר בבירור את השלבים העוקבים המעורבים בתהליך חיפוש העבודה. מהזנת פרטים אישיים ועד לחיפוש קטגוריית עבודה ולבסוף פתיחת בקשה לריאיון עבודה. קל לבין ולעקוב אחר זרימת התהליך. בהירות זו עוזרת למשתמש לנווט בתהליך ללא בלבול ומבטיחה חווית משתמש חלקה.
* תוצאות עבודה רלוונטיות:

השלבים "חיפוש קטגוריית משרה" ו"הצגת תוצאות עבודה" מאפשרים למשתמש לסנן ולהציג הזדמנויות עבודה ספציפיות לתחומי העניין והכישורים שלו. תכונה זו מסייעת למשתמש בקבלת החלטות על ידי הצגת אפשרויות עבודה רלוונטיות ובכך חוסכת זמן ומאמץ בסינון תפקידים לא מותאמים לבקשת המשתמש.

* מידע מפורט על התפקיד:

השלבים "פתיחת פרטי משרה רלוונטית" ו"הצגת דרישות משרה ושכר" מספקים למשתמש מידע מקיף על משרה מסוימת. על ידי גישה לתיאור המשרה המשתמש יכול לקבל החלטות מושכלות האם לממש בקשה לקביעת ריאיון עבודה למשרה מסוימת בהתאם לציפיות שלהם מהמשרה העתידית.

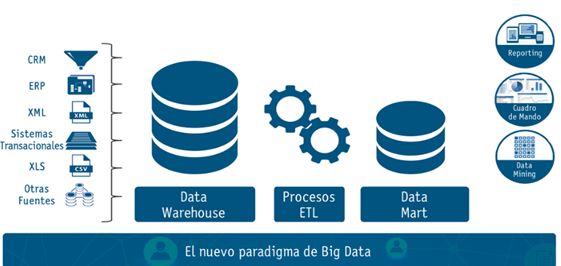
* בקשה לראיון עבודה:

שלב "פתיחת בקשה לריאיון עבודה" מאפשר למשתמש להביע את התעניינותו בעבודה מסוימת ולהתחיל את שלב הגיוס. פונקציונליות זו מסמיכה את המשתמש לנקוט בצעדים יזומים לקראת הבטחת הזדמנות עבודה ומספקת לו תחושת שליטה ומעורבות בתהליך קבלת ההחלטות.

* תפקידי משתמש:

התרשים כולל שלושה שחקנים מרכזיים: מחפש העבודה, חברת ההשמה והמעסיק. על ידי זיהוי שחקנים אלה, התרשים מכיר בבעלי העניין השונים המעורבים בתהליך חיפוש העבודה. הכרה זו עוזרת למשתמשים להבין עם מי הם מקיימים אינטראקציה ואת התפקידים שישויות אלו ממלאות, משפרת את השקיפות ומאפשרת תקשורת יעילה.

**חלק 3-**



* **חילוץ:** ניגשנו לאתר Kaggle וחיפשנו סט נתונים בהתאם לנושא שבחרנו לחקור. לאחר מציאת סט נתונים עשיר ומגוון בסוגי השדות המוכלים בו, הורדנו את סט הנתונים קובץ מסוג CSV ופתחנו אותו בעזרת תוכנת Excel.
* **טרנספורמציה:**
* **הסרת כפילויות**
* **טיפול בערכים חסרים וחריגים**
* **הבטחת שלמות הנתונים-**

המרת ערכי מגדר לערכים סטנדרטיים – זכר, נקבה.

הבטחת ערכים עקביים לגיל, שכר, ותק.

* **ביצוע חישוב נתונים-**

על סמך וותק עובד הגדרנו את רמת העובד: וותק של 1-2 שנים: junior, 3-4 medium , 5 שנים ומעלה: senior.

* **טעינה:**
* **קביעת סכימת היעד- סכימת כוכב.**

הגדרנו את הטבלאות, העמודות והקשרים על סמך הדרישות העסקיות שהגדרנו.

יצרנו תכונות מתאימות בטבלאות העובדות והממדים כדי לייעל את ביצועי השאילתות. בין התכונות הגדרנו מפתחות ראשיים, מפתחות זרים ותכונות אחרות המבוססות על דפוסי השימוש.

* **מיפוי וטעינת הנתונים-**

מיפינו את הנתונים שעברו טרנספורמציה לסכימת הכוכב וביצענו טעינה לטבלאות המתאימות.

* **בדיקת איכות הנתונים-**

אמתנו את איכות הנתונים, בדקנו האם יש חוסר עקביות, ערכים חסרים או חריגים.