# פרויקט גמר – הנדסת נתוני עתק

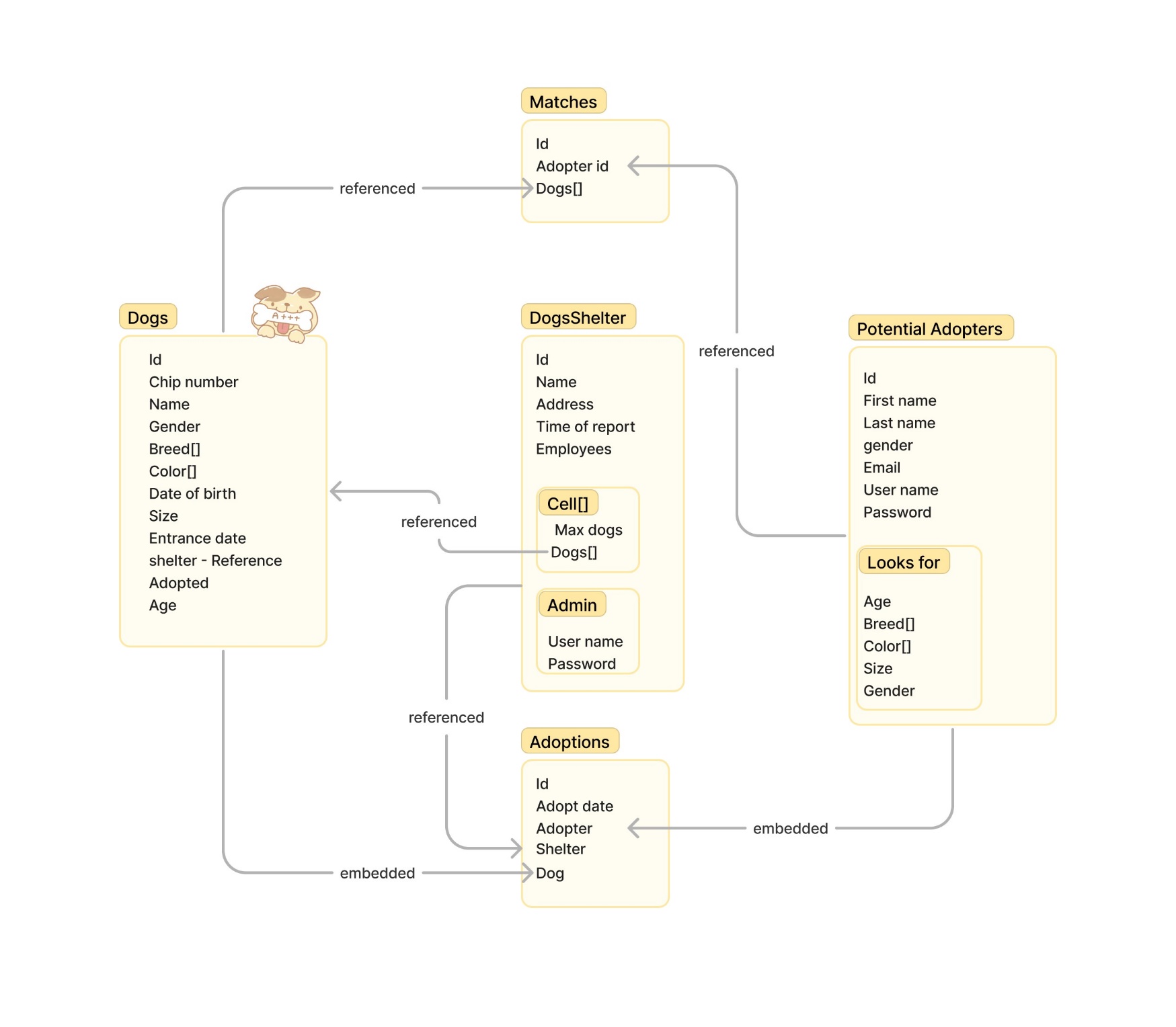
## איפיון המערכת:

הנושא שבחרנו בו הוא נושא פרויקט הגמר שלנו – הקמת מערכת "המלונה".

המערכת מספקת שירותי ניהול כלבייה ותהליכי אימוץ ופרסום.

בחרנו כמה ישויות מרכזיות מהפרויקט ובהן התמקדנו.

**\*הערה חשובה! פונקציה מס' 13 (האחרונה) הינה פונקציה אינטרקטבית בה נדרש להזין קלט ולכן כדאי להריץ אותה בנפרד ולאחר הרצת כל הקוד** 😊

**מבנה האוספים בפרויקט:**

להלן פרטי האוספים הקיימים, חלק מהאוספים נוצרים במהלך הפרויקט בפונקציות שונות.

בהמשך המסמך ליד כל אוסף רשום האם הוא מחושב.

בנוסף, בתחתית המסמך ישנה טבלה המפרטת את כלל הפונקציות כולל הסברים. לכל פונקציה משויך מספר שהוא נמצא מעל הפונקציה בקובץ ה-txt. הפונקציות החדשות שהוספנו מסומנות בצבע צהוב.

**Dog Collection:** אוסף של כלבים המיועדים לאימוץ.

|  |  |  |
| --- | --- | --- |
| **שדה** | **סוג שדה** | **הסבר** |
| Id | Int | מספר מזהה הכלב |
| Chip number | String | מס' הצ'יפ של הכלב |
| Name | String | שם הכלב |
| Gender | Char (m/f) | מין הכלב (זכר/נקבה) |
| Breed | String[] | גזע הכלב – יכול להיות בעל כמה גזעים ולכן רשימה |
| Color | String[] | צבע הכלב – יכול להיות בעל כמה צבעים ולכן רשימה |
| Date of birth | Date | תאריך לידה של הכלב |
| Size | String | גודל הכלב |
| Entrance date | Date | תאריך כניסת הכלב לכלבייה |
| shelter | DogsShelter - Reference | כלבייה בה נמצא הכלב – מופיע כ-Reference לכלבייה ע"י Id של הכלבייה |
| Adopted | Bool | שדה מחושב הנוצר כאשר כלב מאומץ. |
| Age | int | שדה מחושב שנוצר בהתחלה על מנת להפוך את התאריך לידה לגיל |

**DogsShelter Collection:** אוסף של כלביות הכולל את פרטיהן, מערך תאים אשר כל תא כולל את הכלבים בתוכו.

|  |  |  |
| --- | --- | --- |
| **שדה** | **סוג שדה** | **הסבר** |
| Id | Int | מספר מזהה של הכלבייה |
| Name | String | שם הכלבייה |
| Admin | Admin - embedded | פרטי ההתחברות של מנהל הכלבייה (פירוט בהמשך) |
| Address | String | הכתובת בה הכלבייה נמצאת |
| Cells | Cell[] – embedded | רשימה של תאים בכלבייה (פירוט בהמשך) |
| Time of report | Time | זמן להפקת דוח יומי אוטומטי |
| Employees | Int | מס' העובדים בכלבייה |

**Admin:** אובייקט מנהל כלבייה מכיל את פרטי ההתחברות של המנהל למערכת "המלונה".

|  |  |  |
| --- | --- | --- |
| **שדה** | **סוג שדה** | **הסבר** |
| User name | String | שם משתמש למערכת ניהול "המלונה" |
| Password | String | סיסמא למערכת ניהול "המלונה" |

**Cell:** אובייקט תא מכיל את מס' הכלבים המקסימלי בתא ואילו כלבים נמצאים בו ברגע נתון.

|  |  |  |
| --- | --- | --- |
| **שדה** | **סוג שדה** | **הסבר** |
| Max dogs | Int | מס' כלבים מקסימלי שהתא יכול להכיל |
| Dogs | Dog[] - Reference | רשימת הכלבים הנמצאים בתא– מופיע כ-Reference לכלב ע"י Id של הכלב |

**Potential Adopters Collection:** אוסף של מאמצים פוטנציאליים הכולל נתונים יבשים והעדפות לכלבים.

|  |  |  |
| --- | --- | --- |
| **שדה** | **סוג שדה** | **הסבר** |
| Id | Int | מספר פלאפון של המאמץ – מזהה |
| First name | String | שם פרטי המאמץ |
| Last name | String | שם משפחה המאמץ |
| gender | Char (m/f) | מין המאמץ |
| Email | String | מייל המאמץ |
| User name | String | שם משתמש לאתר האימוץ "המלונה" |
| Password | String | סיסמא לאתר האימוץ "המלונה" |
| Looks for | LooksFor - embedded | אובייקט המכיל את כל התכונות שהמאמץ מחפש בכלב (פירוט בהמשך) |

**Looks For:** אובייקט המכיל את העדפותיו של מאמץ פוטנציאלי, האובייקט קיים אך ורק בתוך מאמץ.

|  |  |  |
| --- | --- | --- |
| **שדה** | **סוג שדה** | **הסבר** |
| Age | Int | גיל הכלב |
| Breed | String[] | גזע הכלב – יכול להיות בעל כמה גזעים ולכן רשימה |
| Color | String[] | צבע הכלב – יכול להיות בעל כמה צבעים ולכן רשימה |
| Size | String | גודל הכלב |
| Gender | Char (m/f) | מין הכלב (זכר/נקבה) |

**Adoptions Collection:** אוסף מחושב המרכז את כלל האימוצים שהתבצעו במערכת.

|  |  |  |
| --- | --- | --- |
| **שדה** | **סוג שדה** | **הסבר** |
| Id | ObjectId | Id הנוצר אוטומטי ע"י מונגו |
| Adopters | PotentialAdopters - embedded | פרטי המאמץ |
| Dog | Dog - embedded | פרטי הכלב המאומץ |
| Shelter | DogsShelter - reference | אובייקט המכיל את id ושם הכלבייה ממנה אומץ הכלב |
| Adopt date | Date | תאריך בו הכלב אומץ |

**Matches Collection:** אוסף מחושב המרכז את כל הכלבים שמתאימים למאמץ פוטנציאלי עם אחוז התאמה מעל 50%, ההתאמה מתבססת על מאפייני הכלב אל מול העדפותיו של המאמץ.

|  |  |  |
| --- | --- | --- |
| **שדה** | **סוג שדה** | **הסבר** |
| Id | ObjectId | Id הנוצר אוטומטי ע"י מונגו |
| Adopter id | Potential Adopters - Reference | מאמץ פוטנציאלי |
| Dogs | Dog[] - Reference | רשימת הכלבים המתאימים למאמץ, לכל כלב שמור אחוז ההתאמה |

**הסבר הפונקציות:**

|  |  |  |
| --- | --- | --- |
| **מס' פונקציה** | **שם** | **הסבר** |
| 1 | addDogsToCells | הפונקציה עוברת על כל הכלבים ובצורה רנדומלית בוחרת כלבייה אליה יכנס הכלב, בנוסף היא עוברת על כל התאים של הכלבייה ומכניסה את הכלב לתא הראשון הפנוי בהתבסס על קיבולת התא והתכולה הקיימת, הכלב יתווסף למערך ה-dogs באובייקט Cell. |
| 2 | updateAges | הפונקציה מחשבת את גילו של כלב ומוסיפה שדה Age לכלב. |
| 3 | addLovedDogs | הפונקציה מפעילה את פונקציית updateAges על מנת לקבל את גילם המעודכן ביותר של הכלבים.  הפונקציה יוצרת את אוסף matches- כל מסמך מכיל מאמץ ומערך כלבים מתאימים.  הפונקציה עוברת על כל מערך המאמצים ועבור כל מאמץ היא בודקת את אחוז ההתאמה שלו אל מול כל כלב ע"י השוואה בין שדה LooksFor של המאמץ למאפייניהם של הכלבים, במידה ואחוז ההתאמה גדול או שווה ל50%, הכלב יתווסף למערך הכלבים של אותו מאמץ באוסף של matches. |
| 4 | פונקציית aggregate | הפונקציה מפרקת את מערך הכלבים אצל כל מאמץ באוסף matches ומחשבת לכל כלב את ממוצע אחוזי ההתאמות אצל המאמצים השונים ומחזירה את הכלב עם הממוצע הכי גבוה ואת פרטיו. |
| 5 | פונקציית aggregate | הפונקציה מכניסה לכל מאמץ את מערך הכלבים המתאימים לו הנמצא באוסף matches ובנוסף מכניסה מערך נוסף המכיל את הפרטים המלאים של הכלבים.   * הפונקציה מבצעת pipeline נוסף של אגרגציה בתוך ה-Lookup הראשון. |
| 6 | deleteDogWhoAdopted | פונקציה המקבלת מספר מזהה כלב ומוחקת אותו מכל המאמצים בהם הוא נמצא באוסף matches.   * פונקצית עזר לפונקציית adoption. |
| 7 | adoption | הפונקציה יוצרת את אוסף adoption.  הפונקציה אוספת את כל המאמצים אשר נמצאים באוסף match משמע ישנם כלבי המתאימים להם תוך התחשבות במאמצים אשר כל הכלבים שמתאימים להם כבר אומצו.  עבור מאמצים שנשארו כלבים שמתאימים להם נמיין את הכלבים לפי אחוז ההתאמה הגבוה ביותר ונבצע אימוץ ע"י יצירת מסמך חדש באוסף adoption שם יתווספו פרטי המאמץ, פרטי הכלב ופרטי הכלבייה ותאריך האימוץ. |
| 8 | genderMR – mapReduce | פונקציה המפיקה דוח המציג כמה כלבים מכל מין יש בכל כלבייה. |
| 9 | פונקציית aggregate | הפונקציה ממירה את אובייקט LooksFor של כל מאמץ למערך ומפרקת אותו וסופרת כמה פעמים מאמצים חיפשו כל מאפיין של כלב כמו גיל, צבע, גזע וכו', לבסוף מפיקה דוח המראה כמה פעמים חיפשו כל מאפיין ומה הערכים שאותם חיפשו. |
| 10 | פונקציית aggregate | הפונקציה מחזירה את כל ערכי המאפיינים של הכלבים אותם המאמצים חיפשו וממיינת לפי מספר מופעים מגדול לקטן כלומר המאפיין המחופש ביותר יופיע ראשון. |
| 11 | ShelterOccupancy – mapReduce | הפונקציה עוברת על כל הכלביות ומבצעת חישוב על אחוז התפוסה של כל כלבייה לפי סכום קיבולת התאים וסכום הכלבים הנוכחים בכלבייה. הדוח מופק מציג את קיבולת הכלבייה, מספר הכלבים הנוכחים ואחוז התפוסה. |
| 12 | CountLovedBreeds - mapReduce | הפונקציה עוברת על מערך הגזעים המבוקשים אצל כל מאמץ פוטנציאלי הנמצא בתוך אובייקט LooksFor וסופרת עבור כל גזע כמה מאמצים מעוניינים בו.  לבסוף מציגה כמה אנשים מעוניינים בכל גזע כולל פרטי ההתקשרות שלהם. |
| 13 | פונקציית שליפה ע"י קבלת קלט מהמשתמש | פונקציה המקבלת מהמשתמש מס' מזהה של כלב מתוך הרשימה הניתנה לו ומחזירה את שם הכלב, שם המאמץ ועל בסיס מה נוצרה ביניהם התאמה. |