MonkeDB

Data and Applications Project Phase 1

Team WeDontKnown

Gaurav Singh, 2020111014
Pratyay Suvarnapathaki, 2020111016
Yash Mehan, 2020111020

Our Database

Mini World: A **National Park**

Introduction

A **national park** is an area set aside by a national government for the preservation of the natural environment, and for purposes of public recreation and enjoyment or because of its historical or scientific interest. Most of the landscapes and accompanying plants and animals in a national park are kept in their natural state.
Currently, India has **104** total national parks.
This database deals with one such typical park: its **major interacting components** and **administrative tasks** have been represented systematically. Naturally, due to the scope of operations of a real national park being very expansive, it is not possible to fit every little operational nuance in this database.

Purpose

To **organise and store** the information of the various flora and fauna present in the park, staff (rangers, security personnel, vets), their personal information, visitor data, information about their tours within the park and vehicles owned by the park, with the ability to **update and retrieve** data in an efficient and convenient way.

Users

- ➔ Administrators, tech team and members of the Central and State Government.
- ➔ Various wildlife conservation bodies
- ➔ Tourists
- ➔ Universities, researchers, and National Archives

Applications

Having a proper database gives a **definite structure** to the available information regarding wildlife, park administration, and tourism, making it maintainable in the long term.
Such a database has various applications, ranging from keeping track of populations of endangered and vulnerable animal and plant populations, to serving as a reliable source of statistics for research and to suggest/implement measures to prevent species from going extinct.

Database Requirements

Entities and Attributes

1. **Visitor**
   - ➔ UniqueID (Key) : A string, at max 100 characters. Aadhar number for Indians, other equivalent ID for non-Indians.
   - ➔ Name: A string, at max 100 characters
   - ➔ Phone_Num (Multivalued) : Multiple numbers, each at max 12 digits

2. **Vehicle**
   - ➔ Lic_Number (Key) : Indian Vehicle registration ID (String with at most 11 characters)
   - ➔ Vehicle_type : A string, at max 100 characters
   - ➔ Hourly_rate : Floating point value (in rupees)

3. **Zone**
   - ➔ ZoneName (Key) : A string, at max 100 characters
   - ➔ Coordinates (Composite) (Key) : Coordinates of the rough center-point of the zone.
     - ◆ Longitude : Floating point value (degrees)
     - ◆ Latitude : Floating point value (degrees)
   - ➔ Description : A string, at max 1000 characters. May give details on zone location.

4. **Organism**
   - ➔ Sci_Nomenclature (Composite) (Key)
     - ◆ Genera : A string with max 100 characters
     - ◆ Species: A string with max 100 characters
   - ➔ Class : A string with max 100 characters
   - ➔ Common Name (Multivalued attribute) : Multiple strings with max 100 characters each.

5. **Employee (Superclass)**
   - ➔ Employee_id (Key) : Integer
   - ➔ Name : A string with max 100 characters
   - ➔ Salary : A string with max 100 characters
   - ➔ Job_type : A string with max 100 characters
   - ➔ Address : A string with max 100 characters
   - ➔ DoB : Format (dd/mm/yyyy)
   - ➔ Age (Derived from DoB) : A non-negative Integer

   Subclasses

   i. **Medic**
      - ➔ Specialization : A string with max 100 characters

   ii. **Security**
      - ➔ Weapon : A string with max 100 characters

   iii. **Ranger**
      - ➔ Ranger_grade : A string with max 100 characters

Weak Entities

6. **Tour (weak)**
   - ➔ Start_timestamp (Partial key) : UNIX Timestamp
   - ➔ End_timestamp (Partial key) : UNIX Timestamp
   - ➔ Fee (Derived from duration of the tour and the ‘Hourly_rate’ attribute of the entity ‘Vehicle’) : Floating point value (in rupees)

7. **Population (weak)**
   - ➔ Year (Partial key) : Format (yyyy)
   - ➔ Number : Integer value

Relationships

1. **PRESENT_IN**
   - ➔ Degree 2
   - ➔ Organism, Zone
   - ➔ M:N

2. **ASSIGNED_TO**
   - ➔ Degree 3
   - ➔ Zone, Security, Ranger
   - ➔ 1:M:N

3. **IS_TALLIED**
   - ➔ Degree 2
   - ➔ Organism, Population
   - ➔ 1:N

4. **CAN_TREAT**
   - ➔ Degree 2
   - ➔ Medic, Organism
   - ➔ M:N

5. **SUPERVISES** `[BONUS TASK]`
   - ➔ Degree 2
   - ➔ Employee, Employee
   - ➔ 1:N

n > 3 Relationship

6. **GOES_ON_TOUR**
   - ➔ Degree 5
   - ➔ Visitor, Zone, Ranger, Vehicle, Tour
   - ➔ 1:N:M:L:P

Functional Requirements

Examples of Modifications

1. **Insert**
    a. Insert a new visitor, named A, who submits ID B, phone C, who goes on a tour of zone D from time t=E to time t=F, is charged G rupees, with ranger H, in vehicle J.
    b. Insert a new species with scientific name 'A B', class C, common names D and E, which has a total population of X in the current year Y.
2. **Delete**
    a. From all tables delete records which visitor of UniqueID X was involved in.
3. **Update**
    a. Zone Z changes description from D1 to D2.
    b. Employee M (who happens to be a Medic) who earlier could treat only organism A1, gains expertise and now can treat A2 and A3 also.
    c. Organism A1 expands its territory and is now found in Zone Z2 as well.

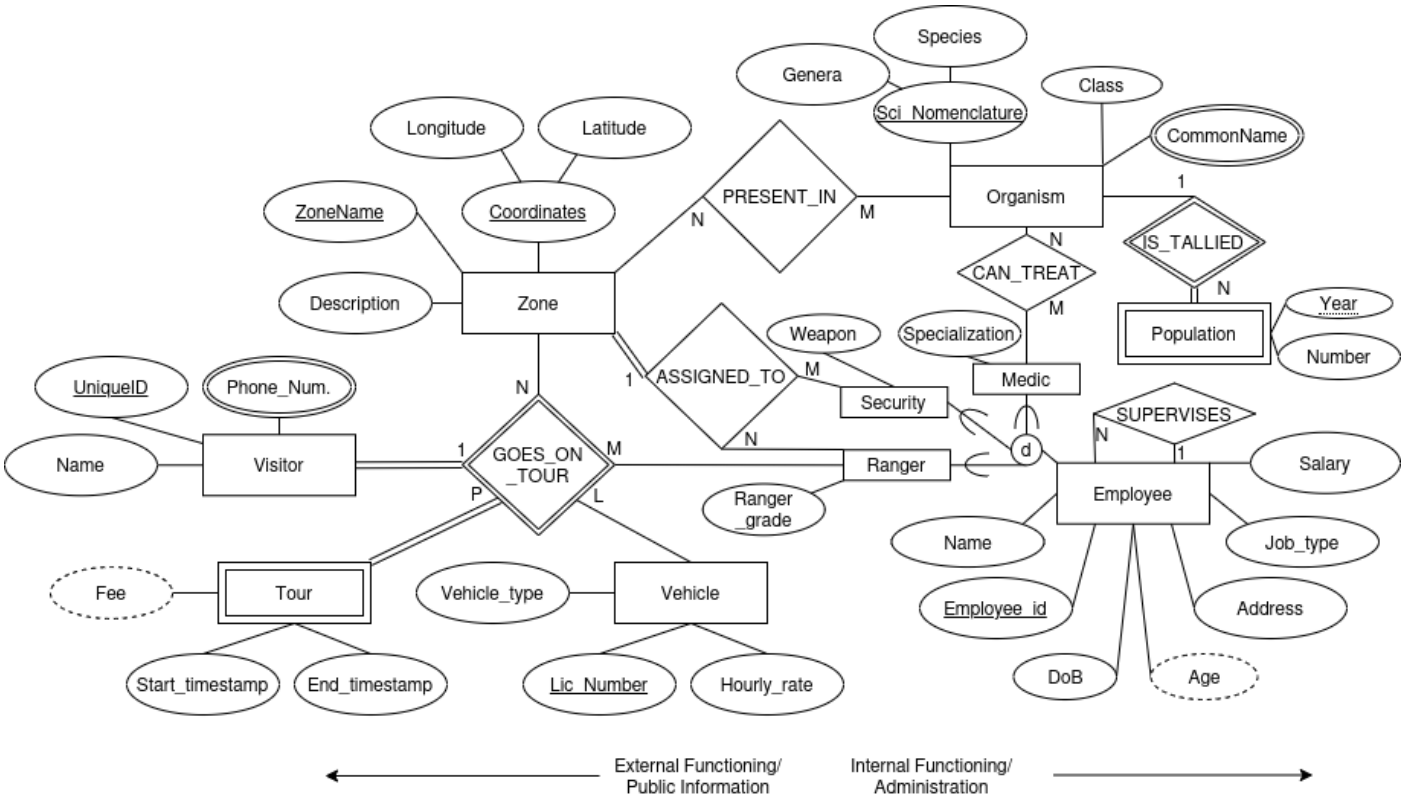
Examples of Retrievals

1. **Selection** : Retrieve data tuples for all organisms who inhabit the Z1 zone.
2. **Projection** : Names of all visitors who stayed in the park for more than 5 hours continuously in one go.
3. **Aggregate** : Average population of organism O1 across all the years' data available in the database in zone Z1.
4. **Search** : text match search such as 'Names of employees that contain the substring 'tan' with Address containing the substring 'pur'.'
5. **Analysis**

    i. Examining which zones are favoured by tourists, and at what times of the year :
    a. In the list of all tours, examine the frequency of visits to each zone per month.
    b. Obtain a list of the most favored zone(s) per month.
    c. Post this, a manual investigation may be conducted to see why certain zones are favored (factors such as natural beauty during monsoon, summer heat, etc. might play a role).

    ii. Considering a hypothetical situation where a top predator say, organism A's population variation is not in line predictions :
    a. Find which zones A is found in.
    b. Gather a list of other (potentially competing) organisms in A's zones.
    c. Analyse their population variation.
    d. Analyse number of visitors entering into A's zones.
    e. Seek inputs of medic who are qualified to treat organism A.
    f. Come up with a mitigation plan to restore population levels of A and A' preys in the respective zones.

E-R Diagram


Species
Genera
Sci_Nomenclature
Class
CommonName
Longitude
Latitude
PRESENT_IN
Organism
ZoneName
Coordinates
N
M
1
IS_TALLIED
N
CAN_TREAT
N
M
Description
Zone
Weapon
Specialization
Year
Population
Number
ASSIGNED_TO
M
Medic
UniqueID
Phone_Num.
N
1
Security
SUPERVISES
Name
Visitor
1
GOES_ON
_TOUR
M
N
Ranger
d
N
1
Salary
P
L
Ranger
_grade
Employee
Name
Job_type
Fee
Tour
Vehicle_type
Vehicle
Employee_id
Address
Start_timestamp
End_timestamp
Lic_Number
Hourly_rate
DoB
Age
External Functioning/
Public Information
Internal Functioning/
Administration