
FINAL PROJECT – DISEASE DIAGNOSIS & MEDIC RECOMMENDATION SYSTEM

ABSTRACT

In our ever-evolving world, there are those who frequently relocate to new locations, such as students like us, expats, who without giving a thought to their safety and wellbeing engage in extensive social interaction, consume unusual delicacies, follow unhealthy lifestyles which includes erratic sleep activity, unsafe sexual practices, intense smoking and drinking, and little to no physical exercise. This may result in a number of illnesses and other health issues that are difficult to treat promptly. Being away from one's own country is difficult but at times not knowing where to reach for the right medical care can even cause death. The main goal of this project is to aid expats and students who are away from home in finding and obtaining medical care as soon as possible.

DESCRIPTION

Our database comprises data for around 200-300 diseases along with its criticality and symptoms. It will assist the user in matching his/her symptoms to diseases. The database will also comprise of doctors working in Massachusetts along with their area of expertise. In-depth diagnoses can be made by the user as a result of this. And it would also help the user in locating a doctor and obtaining the necessary treatment on time. Furthermore, our database will also be able to help the user in locating a doctor upon multiple parameters such as a) Location b) Consulting Cost c) Gender d) Education e) Private/Public Consultations.

PHYSICAL MODEL

Diseases Table

CREATE TABLE Diseases (

Disease_Id INT ,

Disease_name VARCHAR(100) NOT NULL,

Signs_and_symptoms TEXT,

Diagnosis TEXT,

Treatment TEXT,

```
Disease_Category_Id INT,  
  
Doctor_Id INT,  
  
PRIMARY KEY (Disease_Id),  
  
    FOREIGN KEY (Doctor_Id)  
  
    REFERENCES Doctors (Doctor_Id),  
  
    FOREIGN KEY (Disease_Category_Id)  
  
    REFERENCES DiseaseCategory(Disease_Category_Id)  
  
);
```

Doctors Table

```
CREATE TABLE Doctors (  
  
    Doctor_Id INT NOT NULL,  
  
    License_Number TEXT,  
  
    Doctor_Name VARCHAR(150) NOT NULL,  
  
    Hospital_Id INT NOT NULL,  
  
    Gender VARCHAR(20) NOT NULL,  
  
    PRIMARY KEY (Doctor_Id)  
  
);
```

Consultation Cost Table

```
CREATE TABLE Consultation_Cost(  
  
    Id INT NOT NULL,  
  
    Cost INT NOT NULL,  
  
    Hospital_Id INT NOT NULL,  
  
    Doctor_Id INT NOT NULL,  
  
    PRIMARY KEY (Id),
```

FOREIGN KEY (Doctor_Id)

REFERENCES Doctors (Doctor_Id),

FOREIGN KEY (Hospital_Id)

REFERENCES Hospitals(Hospital_Id)

);

Disease Category Table

CREATE TABLE DiseaseCategory (

Disease_Category_Id INT NOT NULL,

Disease_Category_Name VARCHAR(150) NOT NULL UNIQUE,

PRIMARY KEY (Disease_Category_Id)

);

Hospital Affiliations Table

CREATE TABLE Hospital_Affiliations(

Hospital_Id INT NOT NULL,

Hospital_Affiliations VARCHAR(300),

PRIMARY KEY (Hospital_Id,Hospital_Affiliations),

FOREIGN KEY (Hospital_Id)

REFERENCES Hospitals (Hospital_Id)

);

Practice Specialities Table

CREATE TABLE Practice_Specialities (

Doctor_Id INT NOT NULL,

Practice_Specialities VARCHAR(300),

PRIMARY KEY (Doctor_Id , Practice_Specialities),

```
FOREIGN KEY (Doctor_Id)

REFERENCES Doctors (Doctor_Id)

);
```

Hospitals Table

```
CREATE TABLE Hospitals (

    Hospital_Id INT NOT NULL,

    Zip_Code BIGINT NOT NULL,

    PRIMARY KEY (Hospital_Id)

);
```

Pincodes State City Table

```
CREATE TABLE PINCODES_STATE_CITY (

    City VARCHAR(150) NOT NULL,

    State VARCHAR(20),

    Zip_Code INT NOT NULL,

    Hospital_Id INT NOT NULL,

    PRIMARY KEY (Hospital_Id),

    FOREIGN KEY (Hospital_Id)

        REFERENCES Hospitals (Hospital_Id)

);
```

Medicines Table

```
CREATE TABLE Medicines (  
  
    ID INT NOT NULL,  
  
    Drug_Name VARCHAR(200) NOT NULL,  
  
    Manufacturer_Name TEXT,  
  
    Price DECIMAL(10 , 2 ) NOT NULL,  
  
    Disease_Category_Id INT NOT NULL,  
  
    PRIMARY KEY (ID),  
  
    FOREIGN KEY (Disease_Category_Id)  
  
        REFERENCES DiseaseCategory (Disease_Category_Id)  
  
);
```

Degree Table

```
CREATE TABLE DEGREE (  
  
    Doctor_Id INT NOT NULL,  
  
    Doctor_Degree TEXT,  
  
    FOREIGN KEY (Doctor_Id)  
  
        REFERENCES Doctors (Doctor_Id)  
  
);
```

SCREENSHOTS

DISEASES DATA

Result Grid							Filter Rows:		Edit:		Export/Import:		Wrap Cell Content:	
	Disease_Id	Disease_name	Signs_and_symptoms	Diagnosis	Treatment	Disease_Cat	Doctor_Id							
▶	1	Acinetobacter infections	Blood infection: Fever chills vomiting c...	Culture	Supportive care	18	6							
	2	Actinomycosis	Painful abscesses	Histologic findings	Penicillin doxycycline and sulf...	18	8							
	3	Adenovirus infection	high fever that lasts 4 to 6 daysphary...	Antigen detection polymerase ...	Most infections are mild and r...	19	6							
	4	African sleeping sickness (African ...	Hemolymphatic phase: Fever lymphad...	Identification of trypanosome...	Fexinidazole by mouth or pen...	19	8							
	5	AIDS (acquired immunodeficiency ...	Opportunistic infections	Antibody test p24 antigen tes...	Treatment is typically a non-n...	18	96							
	6	Amoebiasis	Most are asymptomatic.	Microscopy	Those with symptoms require ...	19	43							
	7	Anaplasmosis	Fever leukopenia abnormally elevated ...	indirect immunofluorescence a...	Tetracycline drugs (including t...	19	23							
	8	Angiostrongyliasis	Abdominal pain nausea vomiting and w...	Lumbar puncture brain imagin...	Albendazole	18	22							
	9	Anisakiasis	Severe abdominal pain malnutrition an...	Gastroscopic examination or hl...	Albendazole	25	5							

DEGREE TABLE

7 • `SELECT * FROM DEGREE;`

Result Grid | Filter Rows: | Export: | Wrap Cell Co

	Doctor_Degree	Doctor_Id
▶	M.D	1
	M.D	2
	M.D	3
	M.D	4
	D.O.	5
	M.D	6
	M.D	7
	M.D	8
	M.D	9

PRACTICE SPECIALITIES DATA

8 • `select * from Practice_Specialities;`





Result Grid | Filter Rows: | Export: | Wrap

	Practice_Specialities	Doctor_Id
▶	Child and Adolescent Psychiatry	1
	Psychiatry	1
	Forensic Psychiatry	1
	Geriatric Psychiatry	1
	Child and Adolescent Psychiatry	2
	Psychiatry	2
	Forensic Psychiatry	2
	Forensic Psychiatry	3
	Psychiatry	3

DISEASE CATEGORY DATA

10 • `select * from DiseaseCategory;`

<








Result Grid |   Filter Rows: | Export:  | Wrap Cell Content: 

	Disease_Category_Id	Disease_Category_Name
▶	1	Child and Adolescent Psychiatry
	2	Forensic Psychiatry
	3	Medical Microbiology Pathology
	4	Anesthesiology
	5	Addiction Medicine
	6	Consultation-Liaison Psychiatry/Psychosomatic ...
	7	Adult Reconstructive Orthopaedic Surgery
	8	Adult Cardiac Anesthesiology
	9	Neurology

DOCTOR DATA

9 • `select * from Doctors;`

<



Result Grid |   Filter Rows: | Edit:    | Export/Import:  

	Doctor_Id	License_Number	Doctor_Name	Gender
▶	1	54713.0	Ward Bein	Male
	2	58952.0	Donald Condie	Transgender
	3	51517.0	Jeffrey Friedman	Male
	4	255819.0	Mihae Platt	Male
	5	288074.0	Allison Nussbaum	Transgender
	6	287676.0	Analise Peleggi	Male
	7	39144.0	Barry Sugarman	Transgender
	8	291575.0	Bella Vishnevsky	Male
	9	208243.0	Elizabeth Davis	Transgender

HOSPITAL DATA

12 • `select * from Hospitals;`

<




Result Grid |   Filter Rows: | E

	Hospital_Id	Zip_Code
▶	1	1720
	2	2141
	3	2139
	4	1752
	5	2139
	6	2139
	7	1803
	8	1805
	9	2130

PINCODES STATE CITY DATA

11 • `select * from PINCODES_STATE_CITY;`




<

Result Grid |   Filter Rows: | Export:  | Wra

	City	State	Zip_Code	Hospital_Id
▶	Acton	MA	1720	1
	Cambridge	MA	2141	2
	Cambridge	MA	2139	3
	Marlborough	MA	1752	4
	cambridge	MA	2139	5
	Cambridge	MA	2139	6
	Burlington	MA	1803	7
	Burlington	MA	1805	8
	Cambridge	MA	2130	9

HOSPITAL AFFILIATIONS




```
11 • select * from Hospital_Affiliations;
```

<			
Result Grid			
Filter Rows: <input type="text"/>			
Edit:   			
Exp			
	Id		
	Doctor_Id		
	Hospital_Affiliations		
▶	1	1	Tewksbury Hospital
	2	1	Private Practice
	3	2	Private Office
	4	2	People Care Clinic
	5	2	Vinfen Corporation
	6	3	Mount Auburn Hospital
	7	3	Brigham & Women's Hospital
	8	3	Private Practice
	9	3	Beth Israel Deaconess Medical Center

CONSULTATION COST DATA

```
14 • Select*from Consultation_Cost;
```

```
15
```

<				
Result Grid				
Filter Rows: <input type="text"/>				
Edit:   				
Exp				
	Id		Cost	
	Hospital_Id		Doctor_Id	
▶	1	100	1	1
	2	350	2	2
	3	300	3	3
	4	250	4	4
	5	250	5	5
	6	100	6	6
	7	200	7	7
	8	200	8	8
	9	100	9	9

MEDICINES DATA

```
5 • SELECT * FROM ddmrs.medicines;
```

ID	Drug_Name	Manufacturer_Name	Price	Disease_Category_Id
1	Abilify	A & Z Pharmaceuticals Inc	87.58	1
2	Abilify Maintena	A & Z Pharmaceuticals Inc	109.00	2
3	Abacavir	A & Z Pharmaceuticals Inc	117.58	3
4	Abaloparatide	A & Z Pharmaceuticals Inc	110.16	4
5	Abatacept	A & Z Pharmaceuticals Inc	110.98	5
6	Agomelatine	A & Z Pharmaceuticals Inc	121.63	6
7	AbobotulinumtoxinA	A & Z Pharmaceuticals Inc	108.63	7
8	Abrocitinib	A & Z Pharmaceuticals Inc	87.17	8
9	Acetadote	A & Z Pharmaceuticals Inc	111.84	9
10	Acetohydroxamic acid	A & Z Pharmaceuticals Inc	90.69	10

MAJOR USE-CASES AND SCREENSHOTS OF VIEWS

1. List all doctors that are not male or female working in a Massachusetts based hospital and are highly skilled in their area of expertise

```
276 • SELECT * FROM VIEW_FOR_FINDING_OTHER_GENDER_DOCTORS;
```

Doctor_Name	Gender	Degree	Practice_Specialities	Hospital_Affiliations	State
Mihae Platt	Transgender	M.D	Medical Microbiology Pathology	Inc.	MA
Mihae Platt	Transgender	M.D	Medical Microbiology Pathology	Quest Diagnostics	MA
Mihae Platt	Transgender	M.D	Molecular Genetic Pathology	Inc.	MA
Mihae Platt	Transgender	M.D	Molecular Genetic Pathology	Quest Diagnostics	MA
Mihae Platt	Transgender	M.D	Pathology	Inc.	MA
Mihae Platt	Transgender	M.D	Pathology	Quest Diagnostics	MA
Karen Brody	Transgender	M.D	Child and Adolescent Psychiatry	Lahey Clinic	MA
Karen Brody	Transgender	M.D	Psychiatry	Lahey Clinic	MA

VIEW_FOR_FINDING_OTHER_G... x

- List all doctors working in Lahey Clinic specialized in any form of Psychiatry.

122 • `SELECT * FROM VIEW_FOR_DOCTOR_WORKING_IN_LAHEY;`

Doctor_Name	Practice_Specialties	Hospital_Affiliations
Karen Brody	Child and Adolescent Psychiatry	Lahey Clinic
Karen Brody	Psychiatry	Lahey Clinic
Kenneth Lakritz	Consultation-Liaison Psychiatry/Psychosomatic ...	Lahey Clinic
Kenneth Lakritz	Geriatric Psychiatry	Lahey Clinic
Kenneth Lakritz	Psychiatry	Lahey Clinic
Stephen Pinals	Consultation-Liaison Psychiatry/Psychosomatic ...	Lahey Clinic
Stephen Pinals	Geriatric Psychiatry	Lahey Clinic
Stephen Pinals	Psychiatry	Lahey Clinic

- How many doctors are present in Boston that have educational qualification equivalent to an MD degree and Deal with Forensic Pathology?

1 `SELECT * FROM symptoms_of_actinomycosis_and_treatment_locations`

Doctor_Name	Hospital_Affiliations	City	State	Zip_Code	Signs_and_sympto...
Anisa Khaliq	Harvard University Health Services	Cambridge	MA	2138	Painful abscesses

4. What are the symptoms for Actinomycosis and where can I get it treated?

```
1 • SELECT * FROM doctors_with_md_degree_that_deal_forensic_pathology;
```

100% 19:1

Result Grid Filter Rows: Search Export:

Doctor_Name	Gender	Hospital_Affiliations	State	City	ZIP_CODE
Berry Sugarman	Transgender	Private Office	MA	Burlington	1805
Barbara Voetsch	Male	McLean Hospital	MA	Quincy	2169
Dita Mayer	Male	Lahey Clinic	MA	Cambridge	2138
Jacob Elkins	Transgender	Other	MA	Cambridge	2139
Jacob Elkins	Transgender	Mount Auburn Hospital	MA	Cambridge	2139

5. What are the medicines that I can get for Adenovirus infection and which doctor prescribes them?

```
295 • SELECT * FROM VIEW_FOR_DOCTOR_PRESCRIBED_MEDICINE_FOR_SPECIFIC_DISEASE;
```

296

297

298

299

Result Grid Filter Rows: Search Export: Wrap Cell Content:

Doctor_Name	License_Number	Disease_name	Drug_Name
Barbara Voetsch	242103.0	Adenovirus infection	Adalimumab

6. Show all the Hospitals I can book an appointment for visiting Donald Condie.

```
14 • SELECT * FROM all_appointment_locations_for_doc_donald_condie;
```

00% 63:14

Result Grid Filter Rows: Search Export:

Hospital_Affiliati...	City	State	Zip_Code	Hospital_Id
Private Office	Cambridge	MA	2139	3
People Care Clinic	Marlborough	MA	1752	4
Vinfen Corporation	cambridge	MA	2139	5

all_appointment_locations_for_doc_donald_condie 5

7. What is the price of the drug for treating Cholera type disease?

```
33 • SELECT * FROM VIEW_FOR_SEEING_COST_OF_CHOLERA_DISEASE;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

Disease_name	Disease_Category_Name	Drug_Name	Price
Cholera	Infectious Disease	Epclusa	102.93

/VIEW_FOR_SEEING_COST_OF_... x

8. Show all alternatives drugs similar to 'Adace' that can be used to treating Acinetobacter infections.

```
14 • SELECT * FROM alternative_drugs_similar_to_adace_for_treating_acinetobacter;
```

100% 78:14

Result Grid | Filter Rows: Search | Export: |

ID	Drug_Name	Manufacturer_Na...	Price
18	Adace	AbbVie Inc	86.2
146	Adace-V	Roche	75.1
147	Adace-Pil	Pfizer	120
148	Adace-A	Abbvie	120

alternative_drugs_similar_to_adace_for_treating_acinetobacter 6 Read

9. Find doctors with highest consultation who are specialized in Vascular Neurology and their Hospital.

```
74 • Select * from Doctors_VascularNeurology;
```

75

76

<

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

Doctor_Name	Hospital_Affiliations	Practice_Specialities	Cost
Matthew Tilem	Lahey Clinic	Vascular Neurology	500
Collin Culbertson	Private Practice	Vascular Neurology	400
Shruti Sonni	Lahey Clinic	Vascular Neurology	300
Shruti Sonni	Out of State Hospital	Vascular Neurology	300
Benny Kim	Lahey Clinic	Vascular Neurology	250

10. Find the female doctors treating Infectious Diseases in 'Cambridge Health Alliance'.

```
113 • Select*from female_doctors
```

	Disease_Category_Name	Doctor_name	Gender	Hospital_Affiliations
▶	Infectious Disease	Barbara Lewis	Female	Cambridge Health Alliance
	Infectious Disease	Harold Abramowitz	Female	Cambridge Health Alliance
	Infectious Disease	Kenneth Minkoff	Female	Cambridge Health Alliance
	Infectious Disease	Anne Lennon	Female	Cambridge Health Alliance

11. Find all doctors in Cambridge, having consultation cost less than \$500.

Doctor_Id	Doctor_Name	Hospital_Affiliations
▶ 1	Ward Bein	Tewksbury Hospital
1	Ward Bein	Private Practice
2	Donald Condie	Private Office
2	Donald Condie	People Care Clinic
2	Donald Condie	Vinfin Corporation
3	Jeffrey Friedman	Mount Auburn Hospital
3	Jeffrey Friedman	Brigham & Women's Hospital
3	Jeffrey Friedman	Private Practice
3	Jeffrey Friedman	Beth Israel Deaconess Medical Center
5	Allison Nussbaum	Cambridge Health Alliance - Somerville Campus
5	Allison Nussbaum	Cambridge Health Alliance - Cambridge Campus
8	Bella Vishnevsky	Lahey Hospital & Medical Center
8	Bella Vishnevsky	Part of Beth Israel Lahey Health
9	Elizabeth Davis	Cambridge Health Alliance - Somerville Campus
12	Kristin Bruning	Mental Health Clinic

12. List all Neurology drugs under \$150.

Drug_Name	Manufacturer_Name	Price
▶ Acetadote	A & Z Pharmaceuticals Inc	111.84
Actemra	Aavis Pharmaceuticals	83.11
Norvir	Aegerion Pharmaceuticals	83.77
Pegasys	Aegerion Pharmaceuticals	81.38

13. Find Doctors who provide private consultations for Dental Caries disease and their consultation cost.

```
125 • SELECT*FROM Dental_Doc;
126
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:
Doctor_Name	Disease_Name	Hospital_Affiliations	Cost
Jeffrey Friedman	Dental caries	Private Practice	300

14. Find number of doctors working in Beth Israel Deaconess Medical Center.

```
87 • select*from Beth_Israel_Deaconess
```

count(Doctor_Name)	Hospital_Affiliations
19	Beth Israel Deaconess Medical Center

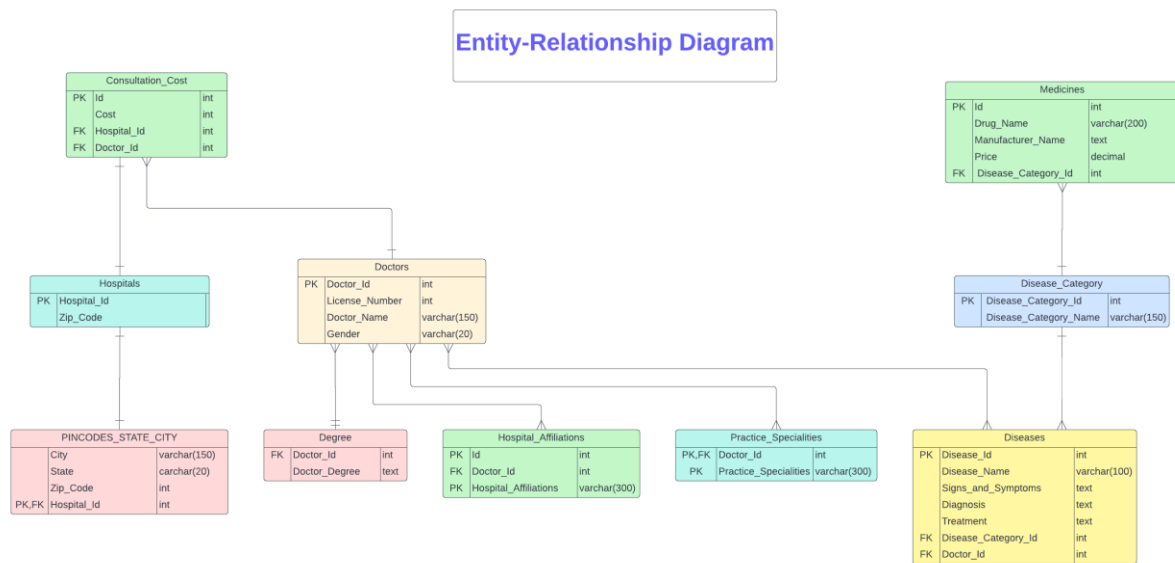
15. Which doctor who is neither a male or female, is treating AIDS disease and what is the medicine and its cost?

```
341 • SELECT * FROM VIEW_FOR_TRANS_DOCTOR_WORKING_ON_AIDS;
```

Doctor_Name	Gender	Disease_name	Drug_Name	Price
Ellen Choi	Transgender	AIDS (acquired immunodeficiency syndrome)	Adacel	86.20

VIEW_FOR_TRANS_DOCTOR_... x

E-R DIAGRAM



CODE

Please find the Python scripts and the physical schemas for our project on this GitHub repository: [Disease-Diagnosis-and-Medic-Recommendation-System](#)

STEPS FOLLOWED SO FAR

1. We developed initial Physical model for our entities as represented in the above E-R Diagram.
2. We have done the following for obtaining relevant data for our project:
 - Scraping data from Wikipedia tables
 - Scraping data from Massdoc websites
 - Scraping and parsing .CSV files
 - Data sources references: [List of Infectious diseases](#) , [List of Doctors and their Hospitals](#) , [List of Doctors v2](#) & [List of Doctors v3](#)
3. We have done the following steps to make our data obtained in step 2 relevant:
 - Data parsing in Python
 - Data cleaning in Python
 - Data reformatting in Python
4. Schema normalization was done and existing tables were split into new tables for maintaining data consistency, quicker execution, reducing data redundancy and overall database organization.
5. We rewrote the entire Python scripts for inserting data into the new normalized schemas.
6. Created views for all major use cases.

CONTRIBUTORS

1. Rohit Panicker | NUID: 002791446
2. Rohit Paul | NUID: 002908933
3. Akshatha Patil | NUID: 002657158