Table Of Contents

Contents

Table Of Contents	1
Phase I. Project Proposal	2
Group Members:	2
Project Proposal:	3
Benefits:	3
Entities:	4
Phase II. Design and Modeling	7
Executive Summary	7
Project Summary	7
Problem Description	8
Conceptual Design	9
EER diagram with all assumptions	9
(Min, Max) Notation for Relationship	10
Relational Schema	11
Relational Schema	11
Assumptions:	12
Data Format for Every Relation	12
4. Normalization	16
5. Conclusion	17
Phase III. Implementation	18
Pre-Illumination	18
Creation of Database with SQL Statements	18
7.1 Table Creation	18
7.2 A Database State	21
Query Scenario Design	30
Conclusion	37

Phase I. Project Proposal

Project Title: Immigration Database System

Group Members:

Name	Email	Role	Duties
Sheetal Mule	SSM220015@utdallas. edu	Project Manager	Technical documents, Managing tasks and resources. Follow-up on tasks and deliverables. Review ERD diagrams, database architecture and reports.
Eshan Todkar	EAT220004@utdallas.e du	Data Analyst	Analysis of data and generate reports, charts, forms.
Tanya Tiwari	TXT230007@utdallas. edu	Data Engineer	Manage Database and required tables, perform SQL queries.
Ankitha Keshavamurthy	AXK230092@utdallas. edu	Data Scientist	Review ERD and integrate it in database, Data collection, creates meaningful data from raw data.
Shrutika Pujari	SXP230057@utdallas.e du	Documentation Writer	Prepare ERD Diagram documentation and presentations.
Palak Sharma	PXS233007@utdallas.e du	Quality Analyst	Test database and its defined entities, queries to execute desired outputs. Keep track of issues.

Project Proposal:

For our project, we propose the creation of a comprehensive database for immigrants. This database will focus on entities such as immigrants, family members, country of origin, employment history, education history, visa types, citizenship status, contact information, legal documents, and port of entry records. The purpose of this database is to provide a centralized repository of information that can benefit immigrants, recruiters, researchers, government organizations, and various stakeholders to offer valuable insights into immigration trends and patterns.

- Immigrants have multiple family members accompanying them.
- Immigrants submit one or more visa application.
- Immigration Agency process multiple visa applications.
- Immigrants belongs to one country.
- Immigrant work for employer.
- Employer has multiple immigrants.
- Immigrants has several education degrees.
- Immigrant has one port of entry.
- Immigrants shares several legal documents with immigration agency.

Benefits:

- Employers can use this data for recruitment job opportunities.
- Policymakers can track how many immigrants have entered the country under which visa and analyze the impact of immigration policies on different immigrant groups.
- Researchers can mine the data to gain insights into immigration trends.
- Social service organizations can utilize the database to allocate resources and provide support to immigrant communities.
- Authorities can use the database to monitor and manage immigration-related security concerns more effectively, and it can also provide data for assessing the economic contributions of immigrants, including tax revenue, entrepreneurship, and job creation.
- Immigrants can easily access their personal records, immigration status, and legal documents, reducing the complexity of navigating immigration bureaucracy, and access to educational and employment history can help them identify opportunities for career growth and educational advancement.

Entities:

- Immigrants
- Family members
- Country of origin
- Employment history
- Education history
- Visa types
- Citizenship status
- Contact information
- Legal documents
- Port of entrys
- 1. Immigrants Table:
- immigrant id (Primary Key)
- first_name
- last name
- date of birth
- country of origin
- arrival date
- visa type
- status (e.g., resident, non-resident, citizen)
- address
- contact info

2. Visa Types Table:

- visa type id (Primary Key)
- visa category
- Description

3. Countries of Origin Table:

- country id (Primary Key)
- country name
- Continent

4. Arrival Ports Table:

- port id (Primary Key)
- port name
- state
- country

5. Family Members Table:

- family member id (Primary Key)
- immigrant id (Foreign Key referencing Immigrants Table)
- relationship
- first name
- last name
- date of birth

6. Employment History Table:

- employment id (Primary Key)
- immigrant id (Foreign Key referencing Immigrants Table)
- employer name
- job_title
- start date
- end date

7. Education History Table:

- education id (Primary Key)
- immigrant id (Foreign Key referencing Immigrants Table)
- school name
- degree
- major
- graduation date

8. Citizenship Status Table:

- citizenship status id (Primary Key)
- immigrant id (Foreign Key referencing Immigrants Table)
- status
- date granted

9. Legal Documents Table:

- document_id (Primary Key)
- immigrant_id (Foreign Key referencing Immigrants Table)
- document_type (e.g., passport, visa, green card)
- document_number
- issue_date
- expiration_date

Phase II. Design and Modeling

Executive Summary

In this project report, we delve into the logic design and modeling of our project:

- Section 1: Introduces the project details, benefits, stakeholders and problem statement.
- Section 2: ER/EER diagram, along with all attributes, underlying assumptions, derived from Section 1.
- Section 3: Relational schema, resulting from the transformation of the ER/EER diagram.
- Section 4: Documenting functional dependencies and normalizing all tables to meet the third normal form (3NF) standards. To conclude, a concise summary is offered at the end of this report.

Project Summary

Database project on immigrants in the United States aims to provide a comprehensive repository of information about immigrants and their associated data. The project involves designing a database that encompasses various entities including immigrants, family members, education and employment histories, legal documents, and more. The goal is to facilitate organized and accessible data for immigrants, policymakers, researchers, and organizations while enabling insights into immigration trends. Below are some benefits of this database system.

- Employers can use this data for recruitment job opportunities.
- Policymakers can track how many immigrants have entered the country under which visa and analyze the impact of immigration policies on different immigrant groups.
- Researchers can mine the data to gain insights into immigration trends.
- Social service organizations can utilize the database to allocate resources and provide support to immigrant communities.
- Authorities can use the database to monitor and manage immigration-related security concerns more effectively, and it can also provide data for assessing the economic contributions of immigrants, including tax revenue, entrepreneurship, and job creation.

• Immigrants can easily access their personal records, immigration status, and legal documents, reducing the complexity of navigating immigration bureaucracy, and access to educational and employment history can help them identify opportunities for career growth and educational advancement.

Project Stakeholders: Immigrants, Employers, Government Agencies, Database Managers, Visa agencies, Research institutions.

Problem Description

In this database project on immigrants in the United States, we aim to solve various challenges and issues faced by various stakeholders.

- The information of immigrants are often scattered and not well structured, making it challenging for immigrants to access data.
- Organizations hiring immigrants face difficulties in targeting required skill sets due to scattered information.
- Limited accessibility to data inhibits in-depth research on immigrant's trends.

Our project aims to generate centralized and accessible databases to resolve the above challenges. Addressing these challenges requires the development of a comprehensive, secure, and accessible database system that consolidates fragmented immigrant data, providing a single, reliable source for various stakeholders. The project aims to create a centralized platform that supports immigrants, policymakers, researchers, service providers, and authorities while ensuring data accuracy, privacy, and security.

Conceptual Design

Here is the EER diagram generated based on both our project description and real-life experiences.

EER diagram with all assumptions

CITIZENSHIP (0,1)(1,*)(0,*)WORKED AT **IMMIGRANTS** (0,*) (1,*) (1,*)**EMPLOYMENT** (1,*) BELONGS HISTORY LEGAL Date_Of_Birt DOCUMENTS HOLD Family_Member_ID COUNTRY OF (0, 1)(0,*) VISA TYPES STUDIED (1,*) (0,*)EDUCATION HISTORY (0,*)(1,*) HAS FAMILY MEMBERS ARRIVED ARRIVAL PORT

Figure 1. EER Design for Immigrants in United States Database

Page 9 of 37

(Min, Max) Notation for Relationship

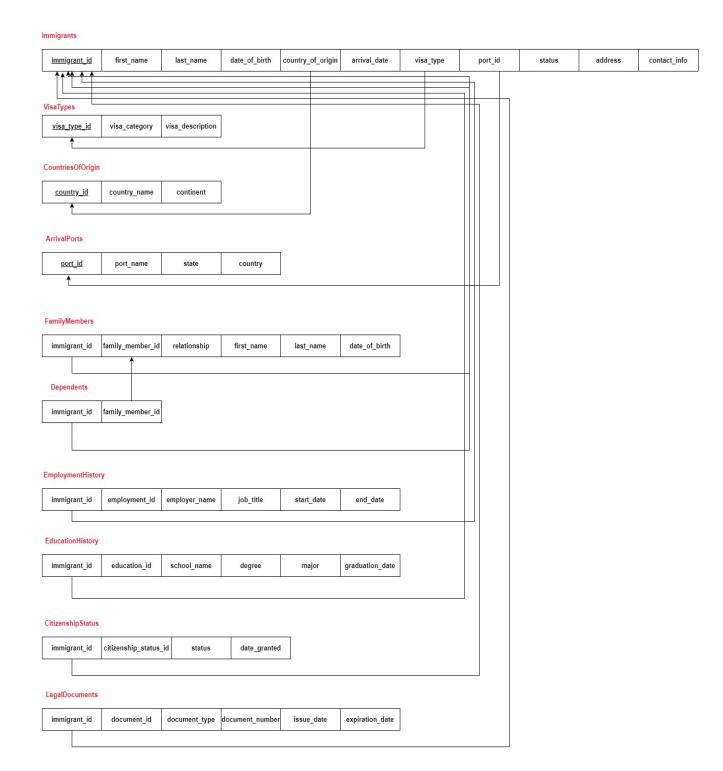
In this part we discuss the (min, max) notation for several important relationships that exist in our EER diagram. Table 1 clearly specifies how the numerical expression corresponds to the relationship between two entities.

Numerical Expression Discussion Each immigrant can be associated with at (1,M)(0,1)most one visa type. It is not mandatory for every immigrant to have a visa type, but if Immigrants Associates_with Visa_type they do, it is (0, 1). One type of visa can be associated with one immigrant or many immigrants so (1, M). (1,M)(1,1)Each immigrant belongs to one country. it is Belongs_to Immigrants Country_of_origin Each country has zero or many immigrants (0, N).

Table 1. Explanation for (Min, Max) Notation

Relational Schema

Relational Schema



Assumptions:

- Each immigrant is tied to singly port of entry.
- Each immigrant is assumed to have one citizenship status.
- The schema might assume limited set of visa types, legal documents of immigrants.
- Contact information is simplified.
- Each immigrant is tied to single country of origin.

Data Format for Every Relation

```
CREATE TABLE Immigrants (
   immigrant id INT PRIMARY KEY,
   first name VARCHAR(255) NOT NULL,
   last name VARCHAR(255) NOT NULL,
   date of birth DATE NOT NULL,
   country of origin VARCHAR(255),
   arrival date DATE,
   port id VARCHAR(255) NOT NULL,
   visa type VARCHAR(255) NOT NULL,
   status VARCHAR(255) NOT NULL,
   address VARCHAR(255),
   contact info VARCHAR(255)
   CONSTRAINT FK visa id FOREIGN KEY (visa type)
   REFERENCES VisaTypes (visa type id),
   CONSTRAINT FK port id FOREIGN KEY (port id)
   REFERENCES ArrivalPorts (port id),
   CONSTRAINT FK country id FOREIGN KEY (country of origin)
   REFERENCES CountriesOfOrigin (country id)
);
```

```
CREATE TABLE VisaTypes (
   visa type id INT PRIMARY KEY,
   visa category VARCHAR(255) NOT NULL,
   visa description VARCHAR(255)
);
CREATE TABLE CountriesOfOrigin (
   country_id INT PRIMARY KEY,
   country name VARCHAR(255) NOT NULL,
   continent VARCHAR(255)
);
CREATE TABLE ArrivalPorts (
   port id INT PRIMARY KEY,
   port_name VARCHAR(255) NOT NULL,
   state VARCHAR(255),
   country VARCHAR(255),
   arrival date DATE
);
CREATE TABLE FamilyMembers (
   family member id INT PRIMARY KEY,
   immigrant id INT,
   relationship VARCHAR(255) NOT NULL,
   first name VARCHAR(255),
   last name VARCHAR(255),
   date of birth DATE
   CONSTRAINT FK immigrant id FOREIGN KEY (immigrant id)
   REFERENCES Immigrants (immigrant id),
);
```

```
CREATE TABLE Dependents (
   family member id INT PRIMARY KEY,
   immigrant id INT,
   CONSTRAINT FK dependent id FOREIGN KEY (family member id)
   REFERENCES FamilyMembers (family member id),
   CONSTRAINT FK immigrant id FOREIGN KEY (immigrant id)
   REFERENCES Immigrants (immigrant id),
);
CREATE TABLE EmploymentHistory (
   employment id INT PRIMARY KEY,
   immigrant id INT,
   employer name VARCHAR(255) NOT NULL,
   job title VARCHAR(255) NOT NULL,
   start date DATE NOT NULL,
   end date DATE NOT NULL,
   CONSTRAINT FK immigrant id1 FOREIGN KEY (immigrant id)
   REFERENCES Immigrants (immigrant id),
);
CREATE TABLE EducationHistory (
   education id INT PRIMARY KEY,
   immigrant id INT,
   school name VARCHAR(255) NOT NULL,
   degree VARCHAR(255) NOT NULL,
   major VARCHAR(255) NOT NULL,
   graduation date VARCHAR(255) NOT NULL
   CONSTRAINT FK_immigrant id2 FOREIGN KEY (immigrant id)
   REFERENCES Immigrants (immigrant id),
);
```

```
CREATE TABLE CitizenshipStatus (
   citizenship status id INT PRIMARY KEY,
   immigrant id INT,
   CONSTRAINT FK immigrant id3 FOREIGN KEY (immigrant id)
   REFERENCES Immigrants (immigrant id),
   status VARCHAR(255) NOT NULL,
   date granted DATE NOT NULL
);
CREATE TABLE LegalDocuments (
 document id INT PRIMARY KEY,
 immigrant id INT,
 document_type VARCHAR(255) NOT NULL,
 document number VARCHAR(255) NOT NULL,
 issue date DATE NOT NULL,
 expiration_date DATE NOT NULL
 CONSTRAINT FK immigrant id4 FOREIGN KEY (immigrant id)
 REFERENCES Immigrants (immigrant id),
);
```

4. Normalization

In this part, we apply the principles of normalization to ensure all the tables conform to 3NF. To do this, we document all functional dependencies, and indicate how the normalization is performed.

Step 1: Strong Entity

Immigrants Table: [immigrant_id (PK), first_name, last_name, date_of_birth, country of origin, arrival date, visa type, status, address, contact info]

Step 2: Weak Entity

FamilyMembers Table: [family_member_id (PK), relationship, first_name, last_name, date of birth]

Dependents: [family member id (FK), immigrant id (FK)]

EmploymentHistory Table: [employment_id (PK), immigrant_id (FK), employer_name, job_title, start_date, end_date]

EducationHistory Table: [education_id (PK), immigrant_id (FK), school_name, degree, major, graduation_date]

CitizenshipStatus Table: [citizenship_status_id (PK), immigrant_id (FK), status date_granted]

LegalDocuments Table: [document_id (PK), immigrant_id (FK), document_type, document_number, issue_date, expiration_date]

ContactInformation Table: [contact_id (PK), immigrant_id (FK), address, email, phone_number]

VisaTypes Table: [visa_type_id (PK), immigrant_id (FK), visa_category, visa_description]

CountriesOfOrigin Table: [country_id (PK), immigrant_id (FK), country_name, continent]

ArrivalPorts Table: [port_id (PK), immigrant_id (FK), port_name, state, country]

In this 3NF representation, each table has a clear primary key, and the other attributes are functionally dependent on the primary keys. This form reduces redundancy and ensures data consistency and integrity within the database.

5. Conclusion

In this part, we discuss and design the relational schema of the immigrants in the United States Database. Our EER diagram and its associated relational schema show the conceptual and logical designs of the system. We also define data types and formats for each attribute in the relation. The next step is to implement this database, different SQL queries to retrieve information, perform data analysis. In the future, we may change some designs due to practical difficulties and other requirements.

Phase III. Implementation

Pre-Illumination

This report outlines the implementation phase of the database project, focusing on the creation of the database, table setup, data population, SQL queries. Our project utilizes the MySQL database management system. Part 1 is the creation of the database, including tables, all other structures as well as constraints, data type and format, and Part 2 is the query scenario design and implementation.

Creation of Database with SQL Statements

7.1 Table Creation

First, we created 10 tables using the following SQL statement:

• Immigrants:

```
CREATE TABLE Immigrants (
   immigrant id INT PRIMARY KEY,
   first name VARCHAR (255) NOT NULL,
   last name VARCHAR (255) NOT NULL,
   date of birth DATE NOT NULL,
   country of origin VARCHAR (255),
   arrival date DATE,
   visa type INT,
   status VARCHAR(255) NOT NULL,
   contact info INT,
      CONSTRAINT FK visa id FOREIGN KEY (visa type)
   REFERENCES VisaTypes (visa type id),
   CONSTRAINT FK port id FOREIGN KEY (port id)
   REFERENCES ArrivalPorts (port id),
   CONSTRAINT FK country id FOREIGN KEY (country of origin)
   REFERENCES CountriesOfOrigin (country id),
      CONSTRAINT FK contact id FOREIGN KEY (contact info)
   REFERENCES contactinformation (contact id)
) ;
```

Page 18 of 37

```
VisaTypes:
```

```
CREATE TABLE VisaTypes (
   visa_type_id INT PRIMARY KEY,
   visa_category VARCHAR(255) NOT NULL,
   visa_description VARCHAR(255)
);
```

• CountriesOfOrigin:

```
CREATE TABLE CountriesOfOrigin (
   country_id INT PRIMARY KEY,
   country_name VARCHAR(255) NOT NULL,
   continent VARCHAR(255)
);
```

ContactInformation:

```
CREATE TABLE ContactInformation (
   contact_id INT PRIMARY KEY,
   address VARCHAR(255) NOT NULL,
   email VARCHAR(255) NOT NULL,
   phone_number VARCHAR(255) NOT NULL
);
```

EducationHistory:

```
CREATE TABLE EducationHistory (
   education_id INT PRIMARY KEY,
   immigrant_id INT,
   CONSTRAINT FK_immigrant_id2 FOREIGN KEY (immigrant_id)
   REFERENCES Immigrants (immigrant_id),
   school_name VARCHAR(255) NOT NULL,
   degree VARCHAR(255) NOT NULL,
   major VARCHAR(255) NOT NULL,
   graduation_date VARCHAR(255) NOT NULL
);
```

EmploymentHistory:

```
CREATE TABLE EmploymentHistory (
   employment id INT PRIMARY KEY,
   immigrant id INT,
   CONSTRAINT FK immigrant id1 FOREIGN KEY (immigrant id)
   REFERENCES Immigrants (immigrant id),
   employer name VARCHAR(255) NOT NULL,
   job title VARCHAR(255) NOT NULL,
   start date DATE NOT NULL,
   end date DATE NOT NULL
   FamilyMembers:
CREATE TABLE FamilyMembers (
   family member id INT PRIMARY KEY,
   immigrant_id INT,
   CONSTRAINT FK immigrant id FOREIGN KEY (immigrant id)
   REFERENCES Immigrants (immigrant id),
   relationship VARCHAR(255) NOT NULL,
   first name VARCHAR(255),
   last name VARCHAR(255),
   date of birth DATE
ArrivalPorts:
CREATE TABLE ArrivalPorts (
   port id INT PRIMARY KEY,
   port name VARCHAR(255) NOT NULL,
   state VARCHAR(255)
LegalDocuments:
CREATE TABLE LegalDocuments (
 document id INT PRIMARY KEY,
 immigrant id INT,
 CONSTRAINT FK immigrant id4 FOREIGN KEY (immigrant id)
 REFERENCES Immigrants (immigrant id),
```

Page 20 of 37

document_type VARCHAR(255) NOT NULL, document number VARCHAR(255) NOT NULL,

issue_date DATE NOT NULL,
expiration date DATE NOT NULL

);

7.2 A Database State

To ensure the database is populated for testing and development purposes, sample data was inserted into each table. The following records were added to each table, maintaining data consistency and validity.

INSERTION OF TABLE IMMIGRANTS

```
INSERT INTO Immigrants (immigrant id, first name, last name,
date of birth, country of origin, arrival date, visa type, status,
contact info, port id)
VALUES
   (123123, 'John', 'Doe', '1990-05-15', 'Canada', '2022-01-01', 1,
'Active', 1,1),
   (123124, 'Jane', 'Smith', '1985-08-20', 'India', '2021-12-15', 2,
'Inactive', 2, 2),
   (123125, 'Carlos', 'Garcia', '1992-03-10', 'Mexico', '2022-02-28', 3,
'Active', 3, 3),
   (123126, 'Aisha', 'Khan', '1988-12-05', 'Pakistan', '2022-03-15', 4,
'Active', 4, 4),
   (123127, 'Alex', 'Wong', '1995-06-30', 'China', '2021-11-10', 5,
'Inactive', 5, 5),
   (123128, 'Elena', 'Martinez', '1993-09-25', 'Spain', '2022-04-20', 6,
'Active', 6, 6),
   (123129, 'Ahmed', 'Ali', '1987-04-12', 'Egypt', '2021-10-05', 7,
'Inactive', 7, 7),
   (123130, 'Yuki', 'Tanaka', '1994-02-18', 'Japan', '2022-05-12', 8,
'Active', 8, 8),
   (123133, 'Anna', 'Novak', '1991-07-08', 'Russia', '2021-09-01', 9,
'Active', 9, 9),
   (123134, 'Miguel', 'Lopez', '1986-11-03', 'Mexico', '2022-06-08', 10,
'Inactive', 10, 10);
```

Table 1. shows the states for Immigrants database schemas.

	immigrant_id	first_name	last_name	date_of_birth	country_of_origin	arrival_date	visa_type	status	contact_info	port_id
•	123123	John	Doe	1990-05-15	Canada	2022-01-01	1	Active	1	1
	123124	Jane	Smith	1985-08-20	India	2021-12-15	2	Inactive	2	2
	123125	Carlos	Garcia	1992-03-10	Mexico	2022-02-28	3	Active	3	3
	123126	Aisha	Khan	1988-12-05	Pakistan	2022-03-15	4	Active	4	4
	123127	Alex	Wong	1995-06-30	China	2021-11-10	5	Inactive	5	5
	123128	Elena	Martinez	1993-09-25	Spain	2022-04-20	6	Active	6	6
	123129	Ahmed	Ali	1987-04-12	Egypt	2021-10-05	7	Inactive	7	7
	123130	Yuki	Tanaka	1994-02-18	Japan	2022-05-12	8	Active	8	8
	123133	Anna	Novak	1991-07-08	Russia	2021-09-01	9	Active	9	9
	123134	Miguel	Lopez	1986-11-03	Mexico	2022-06-08	10	Inactive	10	10

INSERTION OF TABLE VISATYPES

INSERT INTO VisaTypes (visa_type_id, visa_category, visa_description)
VALUES

- (1, 'B-1', 'Business Visitor Business-related activities, such as conferences and negotiations'),
- (2, 'B-2', 'Tourist/Visitor for Pleasure Tourism, vacation, or visits with friends and relatives'),
- (3,'F-1', 'Student Academic studies at a U.S. educational institution'),
- (4, 'H-1B', 'Specialty Occupation Worker Employment in a specialty occupation requiring a higher education degree'),
- (5,'J-1', 'Exchange Visitor Educational and cultural exchange programs, including students and researchers'),
- (6,'L-1', 'Intracompany Transferee Intra-company transfer of employees within multinational companies'),
- (7,'0-1', 'Extraordinary Ability or Achievement Individuals with extraordinary ability or achievement in their field'),
- (8,'K-1', 'Fiancé(e) of U.S. Citizen Fiancé(e) coming to the U.S. to get married'),
- (9,'U', 'Crime Victim Victims of certain crimes willing to assist law enforcement'),
- (10, 'Green Card', 'Lawful Permanent Resident Permanent residency in the U.S.');

visa_type_id visa_category visa_description 1 B-1 Business Visitor - Business-related activities, suc... 2 B-2 Tourist/Visitor for Pleasure - Tourism, vacation, ... 3 F-1 Student - Academic studies at a U.S. education... 4 H-1B Specialty Occupation Worker - Employment in a ... 5 J-1 Exchange Visitor - Educational and cultural exch... 6 L-1 Intracompany Transferee - Intra-company tran... 7 Extraordinary Ability or Achievement - Individua... 0-1 8 K-1 Fiancé(e) of U.S. Citizen - Fiancé(e) coming to t... 9 Crime Victim - Victims of certain crimes willing to ... 10 Green Card Lawful Permanent Resident - Permanent reside...

Table 2. shows the states for VisaTypes database schemas.

INSERTION OF TABLE COUNTRIESOFORIGIN

INSERT INTO CountriesOfOrigin (country id, country name, continent)

```
VALUES
(1,'USA', 'North America'),
(2,'Canada', 'North America'),
```

```
(3,'France', 'Europe'),
(4,'Germany', 'Europe'),
(5,'Japan', 'Asia'),
(6,'Brazil', 'South America'),
(7,'Australia', 'Oceania'),
(8,'South Africa', 'Africa'),
(9,'India', 'Asia'),
(10,'Mexico', 'North America'),
(11,'South Korea', 'Asia'),
(12,'Italy', 'Europe'),
(13,'Russia', 'Europe');
```

Table 3. shows the states for CountriesOfOrigin database schemas.

country_id	country_name	continent
1	USA	North America
2	Canada	North America
3	France	Europe
4	Germany	Europe
5	Japan	Asia
6	Brazil	South America
7	Australia	Oceania
8	South Africa	Africa
9	India	Asia
10	Mexico	North America
11	South Korea	Asia
12	Italy	Europe
13	Russia	Europe

INSERTION OF TABLE CITIZENSHIPSTATUS

Table 4. shows the states for citizenshipstatus database schemas.

citizenship_status_id	immigrant_id	status	date_granted
1	123123	Granted	2023-01-15
2	123124	Pending	2021-02-20
3	123125	Granted	2020-03-05
4	123126	Denied	2014-04-10
5	123127	Granted	2017-05-18
6	123128	Pending	2023-06-25
7	123129	Granted	2020-07-30
8	123130	Denied	2019-08-12
9	123133	Granted	2017-09-03
10	123134	Pending	2005-10-22

INSERTION OF TABLE CONTACTINFORMATION

```
INSERT INTO ContactInformation (contact id, address, email,
phone number)
VALUES
 (1,'123 Elm St, Springfield, IL 62704', 'john.doe@email.com',
'+1234567890'),
 (2,'456 Oak Ave, Lexington, KY 40502', 'jane.smith@email.com',
'+0987654321'),
 (3,'789 Pine Dr, Boulder, CO 80301', 'sam.jackson@email.com',
'+1122334455'),
 (4, '321 Maple Ln, Raleigh, NC 27601', 'emily.white@email.com',
'+9876543210'),
 (5,'654 Birch Blvd, Portland, OR 97201', 'michael.green@email.com',
'+5678901234'),
 (6, '876 Cedar Ct, Austin, TX 78701', 'olivia.brown@email.com',
'+1234509876'),
 (7,'109 Pine St, Tampa, FL 33602', 'daniel.jones@email.com',
'+6789012345'),
 (8,'543 Oak Rd, Atlanta, GA 30301', 'sophia.smith@email.com',
'+3456789012'),
 (9,'987 Maple Ave, Denver, CO 80202', 'logan.miller@email.com',
'+2109876543'),
 (10,'210 Elm Blvd, Seattle, WA 98101', 'amelia.davis@email.com',
'+8765432109');
```

Table 5. shows the states for ContactInformation database schemas.

	contact_id	address	email	phone_number
١	1	123 Elm St, Springfield, IL 62704	john.doe@email.com	+1234567890
	2	456 Oak Ave, Lexington, KY 40502	jane.smith@email.com	+0987654321
	3	789 Pine Dr, Boulder, CO 80301	sam.jackson@email.com	+1122334455
	4	321 Maple Ln, Raleigh, NC 27601	emily.white@email.com	+9876543210
	5	654 Birch Blvd, Portland, OR 97201	michael.green@email.com	+5678901234
	6	876 Cedar Ct, Austin, TX 78701	olivia.brown@email.com	+1234509876
	7	109 Pine St, Tampa, FL 33602	daniel.jones@email.com	+6789012345
	8	543 Oak Rd, Atlanta, GA 30301	sophia.smith@email.com	+3456789012
	9	987 Maple Ave, Denver, CO 80202	logan.miller@email.com	+2109876543
	10	210 Elm Blvd, Seattle, WA 98101	amelia.davis@email.com	+8765432109

INSERTION OF TABLE EDUCATIONHISTORY

INSERT INTO EducationHistory (education_id, immigrant_id, school_name,
degree, major, graduation_date)

VALUES

- (1, 123123, 'Harvard University', 'Bachelor of Science', 'Computer Science', '2020-05-15'),
- (2, 123124, 'University of Oxford', 'Bachelor of Arts', 'English Literature', '2019-12-20'),
- (3, 123125, 'INSEAD', 'Master of Business Administration', 'Business Administration', '2021-06-30'),
- (4, 123126, 'Massachusetts Institute of Technology (MIT)', 'Bachelor of Engineering', 'Mechanical Engineering', '2018-08-10'),
- (5, 123127, 'University of Tokyo', 'Bachelor of Science', 'Biomedical Sciences', '2022-04-25'),
- (6, 123128, 'Stanford University', 'Master of Science', 'Data Science', '2020-10-05'),
- (7, 123129, 'University of Cambridge', 'Master of Arts', 'Psychology', '2019-06-15'),
- (8, 123130, 'London School of Economics and Political Science (LSE)', 'Bachelor of Business Administration', 'Marketing', '2021-02-28'),
- (9, 123133, 'ETH Zurich', 'Master of Engineering', 'Civil Engineering', '2017-07-12'),
- (10, 123134, 'Sorbonne University', 'Bachelor of Arts', 'History', '2022-01-08');

Table 6. shows the states for EducationHistory database schemas.

education_id	immigrant_id	school_name	degree	major	graduation_date
1	123123	Harvard University	Bachelor of Science	Computer Science	2020-05-15
2	123124	University of Oxford	Bachelor of Arts	English Literature	2019-12-20
3	123125	INSEAD	Master of Business Administration	Business Administration	2021-06-30
4	123126	Massachusetts Institute of Technology (MIT)	Bachelor of Engineering	Mechanical Engineering	2018-08-10
5	123127	University of Tokyo	Bachelor of Science	Biomedical Sciences	2022-04-25
6	123128	Stanford University	Master of Science	Data Science	2020-10-05
7	123129	University of Cambridge	Master of Arts	Psychology	2019-06-15
8	123130	London School of Economics and Political Scienc	Bachelor of Business Administration	Marketing	2021-02-28
9	123133	ETH Zurich	Master of Engineering	Civil Engineering	2017-07-12
10	123134	Sorbonne University	Bachelor of Arts	History	2022-01-08

INSERTION OF TABLE EMPLOYMENTHISTORY

INSERT INTO EmploymentHistory (employment id, immigrant id, employer name, job title, start date, end date) VALUES (1, 123123, 'Tech Corp', 'Software Engineer', '2018-05-15', '2021-08-20'), (2, 123124, 'Global Finance Ltd', 'Financial Analyst', '2017-12-10', '2020-06-30'), (3, 123125, 'Marketing Solutions Inc', 'Marketing Manager', '2019-03-01', '2022-01-15'), (4, 123126, 'Manufacturing Innovations', 'Mechanical Engineer', '2016-08-25', '2019-11-30'), (5, 123127, 'BioHealth Research Labs', 'Biomedical Scientist', '2020-02-15', '2023-05-10'), (6, 123128, 'Data Analytics Co', 'Data Scientist', '2019-07-01', '2022-09-30'), (7, 123129, 'MindWorks Psychology Clinic', 'Clinical Psychologist', '2018-06-15', '2021-12-31'), (8, 123130, 'Digital Marketing Pro', 'Marketing Specialist', '2020-03-01', '2022-11-30'), (9, 123133, 'City Engineering Services', 'Civil Engineer', '2015-09-12', '2018-12-20'), (10, 123134, 'Historical Archives Foundation', 'Archivist', '2021-02-01', '2023-04-15');

Table 7. shows the states for EmploymentHistory database schemas.

employment_id	immigrant_id	employer_name	job_title	start_date	end_date
1	123123	Tech Corp	Software Engineer	2018-05-15	2021-08-20
2	123124	Global Finance Ltd	Financial Analyst	2017-12-10	2020-06-30
3	123125	Marketing Solutions Inc	Marketing Manager	2019-03-01	2022-01-15
4	123126	Manufacturing Innovations	Mechanical Engineer	2016-08-25	2019-11-30
5	123127	BioHealth Research Labs	Biomedical Scientist	2020-02-15	2023-05-10
6	123128	Data Analytics Co	Data Scientist	2019-07-01	2022-09-30
7	123129	MindWorks Psychology Clinic	Clinical Psychologist	2018-06-15	2021-12-31
8	123130	Digital Marketing Pro	Marketing Specialist	2020-03-01	2022-11-30
9	123133	City Engineering Services	Civil Engineer	2015-09-12	2018-12-20
10	123134	Historical Archives Foundation	Archivist	2021-02-01	2023-04-15

INSERTION OF TABLE FAMILYMEMBERS

INSERT INTO FamilyMembers (family_member_id, immigrant_id, relationship,
first name, last name, date of birth)

```
VALUES
```

```
(1, 123123, 'Spouse', 'Mary', 'Doe', '1985-03-12'),
(2, 123123, 'Child', 'Emma', 'Doe', '2010-08-25'),
(3, 123124, 'Parent', 'Robert', 'Smith', '1950-11-18'),
(4, 123125, 'Sibling', 'Lisa', 'Garcia', '1990-06-03'),
(5, 123126, 'Spouse', 'David', 'Khan', '1982-09-20'),
(6, 123127, 'Child', 'Sophie', 'Wong', '2015-04-15'),
(7, 123128, 'Parent', 'Emily', 'Martinez', '1968-07-28'),
(8, 123129, 'Sibling', 'Alex', 'Ali', '1988-12-10'),
(9, 123130, 'Spouse', 'Oliver', 'Tanaka', '1980-05-03'),
(10, 123133, 'Child', 'Ava', 'Novak', '2000-02-14');
```

Table 8. shows the states for FamilyMembers database schemas.

family_member_id	immigrant_id	relationship	first_name	last_name	date_of_birth
1	123123	Spouse	Mary	Doe	1985-03-12
2	123123	Child	Emma	Doe	2010-08-25
3	123124	Parent	Robert	Smith	1950-11-18
4	123125	Sibling	Lisa	Garcia	1990-06-03
5	123126	Spouse	David	Khan	1982-09-20
6	123127	Child	Sophie	Wong	2015-04-15
7	123128	Parent	Emily	Martinez	1968-07-28
8	123129	Sibling	Alex	Ali	1988-12-10
9	123130	Spouse	Oliver	Tanaka	1980-05-03
10	123133	Child	Ava	Novak	2000-02-14

INSERTION OF TABLE ARRIVALPORTS

```
INSERT INTO ArrivalPorts (port_id, port_name, state)
VALUES

(1, 'Port of Los Angeles', 'California'),
 (2, 'Port of Long Beach', 'California'),
 (3, 'Port of New York and New Jersey', 'New York'),
 (4, 'Port of Savannah', 'Georgia'),
 (5, 'Port of Houston', 'Texas'),
 (6, 'Port of Seattle', 'Washington'),
 (7, 'Port of Charleston', 'South Carolina'),
 (8, 'Port of Miami', 'Florida'),
 (9, 'Port of San Francisco', 'California'),
 (10, 'Port of Boston', 'Massachusetts');
```

Table 9. shows the states for ArrivalPorts database schemas.

port_id	port_name	state
1	Port of Los Angeles	California
2	Port of Long Beach	California
3	Port of New York and New Jersey	New York
4	Port of Savannah	Georgia
5	Port of Houston	Texas
6	Port of Seattle	Washington
7	Port of Charleston	South Carolina
8	Port of Miami	Florida
9	Port of San Francisco	California
10	Port of Boston	Massachusetts

INSERTION OF TABLE LEGALDOCUMENTS

INSERT INTO LegalDocuments (document_id, immigrant_id, document_type, document_number, issue_date, expiration_date)
VALUES
 (1, 123123, 'Passport', 'AB123456', '2019-05-15', '2024-05-14'),
 (2, 123123, 'Visa', 'X987654', '2020-01-10', '2022-01-09'),
 (3, 123123, 'Green Card', 'GC456789', '2018-06-30', '2028-06-29'),
 (4, 123124, 'Driver License', 'DL789012', '2017-08-10', '2023-08-09'),
 (5, 123125, 'Work Permit', 'WP345678', '2021-04-25', '2023-04-24'),
 (6, 123126, 'ID Card', 'ID123456', '2019-10-05', '2024-10-04'),
 (7, 123127, 'Residence Permit', 'RP987654', '2018-06-15',
'2023-06-14'),
 (8, 123128, 'Social Security Card', 'SSC456789', '2020-02-28',
'2030-02-27'),

```
(9, 123130, 'Visa', 'V123456', '2017-07-12', '2022-07-11'), (10, 123133, 'ID Card', 'ID789012', '2022-01-08', '2027-01-07'), (11, 123125, 'ID Card', 'ID199456', '2021-10-05', '2024-10-04'), (12, 123127, 'ID Card', 'ID156456', '2020-10-05', '2024-10-04');
```

Table 10. shows the states for LegalDocuments database schemas.

document_id	immigrant_id	document_type	document_number	issue_date	expiration_date
1	123123	Passport	AB123456	2019-05-15	2024-05-14
2	123123	Visa	X987654	2020-01-10	2022-01-09
3	123123	Green Card	GC456789	2018-06-30	2028-06-29
4	123124	Driver License	DL789012	2017-08-10	2023-08-09
5	123125	Work Permit	WP345678	2021-04-25	2023-04-24
6	123126	ID Card	ID123456	2019-10-05	2024-10-04
7	123127	Residence Permit	RP987654	2018-06-15	2023-06-14
8	123128	Social Security Card	SSC456789	2020-02-28	2030-02-27
9	123130	Visa	V123456	2017-07-12	2022-07-11
10	123133	ID Card	ID789012	2022-01-08	2027-01-07
11	123125	ID Card	ID 199456	2021-10-05	2024-10-04
12	123127	ID Card	ID156456	2020-10-05	2024-10-04

Till now we have finished the process of creating tables and database states.

Query Scenario Design

Query 01: First name, last name, and number of legal documents for immigrants who have more than one document. Query result should be grouped by Immigrant ID.

```
SELECT Immigrants.first_name, Immigrants.last_name,
COUNT(LegalDocuments.document_id) AS num_documents
FROM Immigrants
LEFT JOIN LegalDocuments ON Immigrants.immigrant_id =
LegalDocuments.immigrant_id
GROUP BY Immigrants.immigrant_id
HAVING num documents > 1;
```

Result of Query 01:

first_name	last_name	num_documents
John	Doe	3
Carlos	Garcia	2
Alex	Wong	2

Query 02: First name, last name, and degree information of immigrants who have advanced degrees.

```
SELECT Immigrants.immigrant_id, Immigrants.first_name,
Immigrants.last_name, EducationHistory.degree
FROM Immigrants
JOIN EducationHistory ON Immigrants.immigrant_id =
EducationHistory.immigrant_id
WHERE EducationHistory.degree IN ('Master of Arts', 'Master of Engineering');
```

Result of Query 02:

immigrant_id	first_name	last_name	degree
123129	Ahmed	Ali	Master of Arts
123133	Anna	Novak	Master of Engineering

Query 03: First name, last name, job title, country of origin, and US citizenship status for immigrants who are engineers.

```
SELECT i.first_name,
i.last_name,
em.job_title,
i.country_of_origin,
cs.status AS us_citizenship_status
FROM immigrants AS i
JOIN employmenthistory AS em ON i.immigrant_id = em.immigrant_id
JOIN citizenshipstatus AS cs ON i.immigrant_id = cs.immigrant_id
WHERE em.job title LIKE '%engineer%';
```

Result of Query 03:

first_name	last_name	job_title	country_of_origin	us_citizenship_status
John Doe		Software Engineer	Canada	Granted
Aisha	Khan	Mechanical Engineer	Pakistan	Denied
Anna	Novak	Civil Engineer	Russia	Granted

Query 04: The average age of immigrants based on their visa category.

```
SELECT
```

```
vt.visa_category,
   ROUND(AVG(YEAR(CURDATE()) - YEAR(i.date_of_birth))) AS avg_age
FROM visatypes AS vt
JOIN immigrants AS i ON vt.visa_type_id = i.visa_type
GROUP BY vt.visa category;
```

Result of Query 04:

visa_category	avg_age
B-1	33
B-2	38
F-1	31
H-1B	35
J-1	28
L-1	30
0-1	36
K-1	29
U	32

Green Card 37

Query 05: Top 5 immigrants with the most family members.

Result of Query 05:

immigrant_id	first_name	last_name	family_member_count
123123	John	Doe	2
123124	Jane	Smith	1
123125	Carlos	Garcia	1
123126	Aisha	Khan	1
123127	Alex	Wong	1

Query 06: The number of immigrants arriving through each port.

```
SELECT p.port_name, COUNT(i.immigrant_id) AS num_immigrants
FROM immigrants i
JOIN arrivalports p ON i.port_id = p.port_id
GROUP BY p.port_name
ORDER BY num immigrants DESC;
```

Result of Query 06:

port_name	num_immigrants
Port of Los Angeles	1
Port of Long Beach	1
Port of New York and New Jersey	1
Port of Savannah	1
Port of Houston	1
Port of Seattle	1
Port of Charleston	1
Port of Miami	1
Port of San Francisco	1

Port of Boston	1

Query 07: The duration of immigrants' previous employment history.

```
CREATE VIEW Emp_History_Duration AS
SELECT employment_id, immigrant_id, job_title,
DATEDIFF(end_date, start_date) / 365.25 AS employment_duration_years
FROM employmenthistory;
SELECT * FROM Emp History Duration;
```

Result of Query 07:

employment_id	immigrant_id	job_title	employment_duration_years
1	123123	Software Engineer	3.2663
2	123124	Financial Analyst	2.5544
3	123125	Marketing Manager	2.8775
4	123126	Mechanical Engineer	3.2635
5	123127	Biomedical Scientist	3.2307
6	123128	Data Scientist	3.2498
7	123129	Clinical Psychologist	3.5455
8	123130	Marketing Specialist	2.7488
9	123133	Civil Engineer	3.2717
10	123134	Archivist	2.1985

Query 08: Filtering out visa types that are economically beneficial to the US.

```
SELECT * FROM visatypes;
CREATE VIEW Filtered_VisaTypes AS
SELECT visa_category, visa_description, visa_type_id,
COUNT(visa_category) AS distinct_count
FROM visatypes
WHERE visa_category IN ('F-1', 'J-1', 'B-2')
GROUP BY visa_category, visa_description, visa_type_id;
SELECT * FROM Filtered VisaTypes;
```

Result of Query 08:

visa_category	visa_description	visa_type_id	distinct_count
B-2	Tourist/Visitor for Pleasure - Tourism, vacation, or visits with friends and relatives	2	1
F-1	Student - Academic studies at a U.S. educational institution	3	1
J-1	Exchange Visitor - Educational and cultural exchange programs, including	5	1

students and researchers

Query 09: Create a view table of immigrants with required columns.

```
SELECT * FROM immigrants;
CREATE VIEW US_Immigrant_1 AS
SELECT immigrant_id, visa_type, country_of_origin
FROM immigrants;
SELECT * FROM US_Immigrant_1;
```

Result of Query 09:

immigrant_id	visa_type	country_of_origin
123123	1	Canada
123124	2	India
123125	3	Mexico
123126	4	Pakistan
123127	5	China
123128	6	Spain
123129	7	Egypt
123130	8	Japan
123133	9	Russia
123134	10	Mexico



Query 10: Join 3 tables.

```
SELECT
```

```
i.immigrant_id AS immigrant_id_immigrants,
i.visa_type AS visa_type_immigrants,
i.country_of_origin,
e.employment_id AS emp_id,
e.job_title,
e.employment_duration_years AS emp_duration,
v.visa_category,
v.visa_type_id AS visa_type_id_visatypes,
v.visa_description
FROM US_Immigrant_1 i
INNER JOIN Emp_History_Duration e ON i.immigrant_id = e.immigrant_id
INNER JOIN Filtered VisaTypes v ON i.visa type = v.visa type id;
```

Result of Query 10:

immigrant_id	visa_type	country_of_orig	emp_id	job_title	emp_duration	visa_category	visa_type_id_visatypes	visa_description
123124	2	India	2	Financial Analyst	2.5544	B-2	2	Tourist/Visitor for Pleasure - Tourism, vacation,
123125	3	Mexico	3	Marketing Manager	2.8775	F-1	3	Student - Academic studies at a U.S. education
123127	5	China	5	Biomedical Scientist	3.2307	J-1	5	Exchange Visitor - Educational and cultural exch

Query 11: Immigrants with more than one family member accompanying them.

```
SELECT i.first_name, i.last_name, COUNT(f.family_member_id) AS
family_members_count
FROM Immigrants i
JOIN FamilyMembers f ON i.immigrant_id = f.immigrant_id
GROUP BY i.immigrant_id
HAVING COUNT(f.family member id) > 1;
```

Result of Query 11:

first_name	last_name	family_members_count
John	Doe	2

Query 12: Calculating the percentage of immigrants whose visa status is active.

```
SELECT status, (COUNT(*) / (SELECT COUNT(*) FROM Immigrants)) * 100 AS
active_visa_percentage
FROM Immigrants
WHERE status = 'Active';
```

Result of Query 12:

status active_visa_percentage

Active 60.0000

Query 13: Immigrants with the longest immigration history.

```
SELECT i.first_name, i.last_name, MAX(DATEDIFF(e.end_date,
e.start_date)) AS max_employment_duration
FROM Immigrants i
JOIN EmploymentHistory e ON i.immigrant_id = e.immigrant_id
GROUP BY i.immigrant_id
ORDER BY max_employment_duration DESC
LIMIT 5;
```

Result of Query 13:

first_name	last_name	max_employment_duration
Ahmed	Ali	1295
Anna	Novak	1195
John	Doe	1193
Aisha	Khan	1192
Elena	Martinez	1187

Query 14: Top 3 countries of origin with the most immigrants and their counts.

```
SELECT c.country_name, COUNT(*) AS immigrant_count
FROM Immigrants i
JOIN CountriesOfOrigin c ON i.country_of_origin = c.country_name
GROUP BY c.country_name
ORDER BY immigrant_count DESC
LIMIT 3;
```

Result of Query 14:

country_name	immigrant_count				
Mexico	2				
Canada	1				
Japan	1				

Query 15: Immigrants with the highest number of legal documents.

```
SELECT i.first_name, i.last_name, COUNT(l.document_id) AS
document_count
FROM Immigrants i
JOIN LegalDocuments l ON i.immigrant_id = l.immigrant_id
GROUP BY i.immigrant_id
ORDER BY document_count DESC
LIMIT 8;
```

Result of Query 15:

first_name	last_name	document_count			
John	Doe	3			
Carlos	Garcia	2			
Alex	Wong	2			
Jane	Smith	1			
Aisha	Khan	1			
Elena	Martinez	1			
Yuki	Tanaka	1			
Anna	Novak	1			

Query 16: Count of the number of distinct immigrant last names.

```
SELECT COUNT (DISTINCT last_name)
FROM immigrants;

Result of Query 16:
COUNT(DISTINCT last_name)
10
```

Query 17: Information of immigrants and their education history.

```
SELECT *
FROM immigrants
INNER JOIN educationhistory ON immigrants.immigrant_id =
educationhistory.immigrant id;
```

Result of Query 17:

immigrant_id	first_name	last_name	date_of_birth	country_of_origin	arrival_date	visa_type	status	contact_info	port_id	education_id	immigrant_id	school_name	degree	major	graduation_dat
123123	John	Doe	1990-05-15	Canada	2022-01-01	1	Active	1	1	1	123123	Harvard University	Bachelor of Science	Computer Science	2020-05-15
123124	Jane	Smith	1985-08-20	India	2021-12-15	2	Inactive	2	2	2	123124	University of Oxford	Bachelor of Arts	English Literature	2019-12-20
123125	Carlos	Garcia	1992-03-10	Mexico	2022-02-28	3	Active	3	3	3	123125	INSEAD	Master of Business Administration	Business Administration	2021-06-30
123126	Aisha	Khan	1988-12-05	Pakistan	2022-03-15	4	Active	4	4	4	123126	Massachusetts Institute of Technology (MIT)	Bachelor of Engineering	Mechanical Engineering	2018-08-10
123127	Alex	Wong	1995-06-30	China	2021-11-10	5	Inactive	5	5	5	123127	University of Tokyo	Bachelor of Science	Biomedical Sciences	2022-04-25
123128	Elena	Martinez	1993-09-25	Spain	2022-04-20	6	Active	6	6	6	123128	Stanford University	Master of Science	Data Science	2020-10-05
123129	Ahmed	Ali	1987-04-12	Egypt	2021-10-05	7	Inactive	7	7	7	123129	University of Cambridge	Master of Arts	Psychology	2019-06-15
123130	Yuki	Tanaka	1994-02-18	Japan	2022-05-12	8	Active	8	8	8	123130	London School of Economics and Political Scienc	Bachelor of Business Administration	Marketing	2021-02-28
123133	Anna	Novak	1991-07-08	Russia	2021-09-01	9	Active	9	9	9	123133	ETH Zurich	Master of Engineering	Civil Engineering	2017-07-12
123134	Miguel	Lopez	1986-11-03	Mexico	2022-06-08	10	Inactive	10	10	10	123134	Sorbonne University	Bachelor of Arts	History	2022-01-08

Conclusion

In this report we describe the implementation process. This includes steps for setting up the database, populating it with data, and, if applicable, developing the software component. Challenges faced during implementation are outlined along with their resolutions.