

CPSC 304 Project Cover Page

Milestone #: 4

Date: November 30, 2023

Group Number: 69

Name	Student Number	CS Alias (Userid)	Preferred E-mail Address
Allison Jiao	63646590	y4j5h	alliOn@student.ubc.ca
Manjot Singh	78562972	b3r7o	manjjott@student.ubc.ca
William Sun	74695198	p9k3b	williamsunedu1617@gmail.com

By typing our names and student numbers in the above table, we certify that the work in the attached assignment was performed solely by those whose names and student IDs are included above. (In the case of Project Milestone 0, the main purpose of this page is for you to let us know your e-mail address, and then let us assign you to a TA for your project supervisor.)

In addition, we indicate that we are fully aware of the rules and consequences of plagiarism, as set forth by the Department of Computer Science and the University of British Columbia

Group 69 Milestone 4 PDF

a. A short description of the final project, and what it accomplished.

Our project is a Canadian National census database, the database contains census facts on both individuals and households. It is a comprehensive database for local and national governments to access vital information for budget allocation in different areas. The final project involved the creation of a well-structured Canadian National Census database. This database was designed to effectively capture, store, and analyze detailed demographic and occupational information of Canadian citizens and residents.

Key accomplishments of the project include:

- Our database organization facilitated easier data entry, retrieval, and management.
- The project ensured high data integrity and accuracy. This reduced data redundancy and inconsistencies, leading to more reliable census information.
- The database design offered scalability and flexibility, allowing for future expansions and updates as required by changing census parameters or additional data requirements.

Overall, the project played a crucial role in modernizing and enhancing the efficiency of Canada's national census process, which happens once every 5 years. Our database should record accurate and up-to-date population data.

b. A description of how your final schema differed from the schema you turned in.

i. If the final schema differed, explain why. Note that turning in a final schema that's different from what you planned is fine, we just want to know what changed and why.

We figured that we did not need too many attributes thus we cut down 8 attributes that we initially had. Cutting down the number of attributes mainly served the purpose of making implementation easier for us.

SIN → Individual Name, Gender, Age, Income, **Ethnicity, Birthplace**, Level Name, Status ID, Address, Postal Code, Skill_ID

SIN, Authorization ID → Occupation ID

Occupation ID → Occupation Name, **Occupation Average Income**

Level Name → **Education Average Income**

Status ID → Status Name, **Average Age of Marriage**

Address, Postal Code → Number of Members, ED Name, City Name, Province Name, Language Name

ED Name → MP, MP Party

City Name, Province Name → Population

The following highlighted attributes in yellow were taken out. Our rationale for excluding each attribute was mostly arbitrary. However information such as “occupation average income”, “education average income”, “average age of marriage”, “MP”, “MP party” and “population” are mostly expendable. “Ethnicity” and “birthplace” are not commonly recorded on the Canadian Census. Throughout our implementation, we also added the table Individual Skills to our project, hence we have the new attribute Skill_ID.

c. A copy of the schema and screenshots that show what data is present in each relation after the SQL script from item #2 is run.

i. Do not use DDL statements for this deliverable. List the relational schemas with the primary key attributes underlined and foreign keys bolded. For example, Student(sid, name, dob).

ii. You can complete this deliverable with screenshots from SQL Plus or create a representation of your relational instances through a spreadsheet program. No matter what you use, be sure to clearly label which relation a given instance refers to.

- **Individual** (Individual Name, Gender, Age, SIN, Income, **Address**, **Postal Code**, **Occupation ID**, **Education ID**, **Status ID**)
- **Household** (Address, Postal Code, Number of Members, **ED Name**, **City Name**, **Province Name**, **Language Name**) (ED Name, City Name, Province Name cannot be null)
- **Occupation** (Occupation ID, Occupation Name)
- **Education Level** (Education ID, Level Name)
- **Marital Status** (Status ID, Status Name)
- **Electoral District** (ED Name)
- **City** (City Name, Province Name)
- **Language** (Language Name, **Address**, **Postal Code**)
- **Employed** (Employer, Job Title, Occupation ID, Occupation Name)
- **Unemployed** (Seeking Employment, Occupation ID, Occupation Name)
- **Student** (Institution Name, Level of Study, Occupation ID, Occupation Name)
- **Retired** (Age at Retirement, Occupation ID, Occupation Name)
- **Individual Skills** (Skill_ID, **SIN**)

```
SQL> SELECT table_name from user_tables;
```

```
TABLE_NAME
```

```
INDIVIDUAL
HOUSEHOLD
OCCUPATION
EDUCATION_LEVEL
MARITAL_STATUS
ELECTORAL_DISTRICT
CITY
LANGUAGE
EMPLOYED
UNEMPLOYED
STUDENT
```

```
TABLE_NAME
```

```
RETIRED
SKILLS
INDIVIDUALSKILLS
```

Fig. 1 Screenshot of all the Tables

```
SQL> DESCRIBE Individual;
```

Name	Null?	Type
Individual_Name		VARCHAR2(20)
Gender		VARCHAR2(20)
Age		NUMBER(38)
SIN	NOT NULL	NUMBER(38)
Income		NUMBER(38)
Address		VARCHAR2(20)
Postal_Code		VARCHAR2(6)
Occupation_ID		NUMBER(38)
Education_ID		NUMBER(38)
Status_ID		NUMBER(38)

Fig. 2 Screenshot of Individual

```
SQL> DESCRIBE Household;
```

Name	Null?	Type
Address	NOT NULL	VARCHAR2(20)
Postal_Code	NOT NULL	VARCHAR2(6)
Number_of_Members		NUMBER(38)
ED_Name		VARCHAR2(20)
City_Name		VARCHAR2(20)
Province_Name		VARCHAR2(20)
Language_Name		VARCHAR2(20)

Fig. 3 Screenshot of Household

```
SQL> DESCRIBE Occupation;
Name                                     Null?      Type
-----
Occupation_ID                           NOT NULL   NUMBER(38)
Occupation_Name                                    VARCHAR2(30)
```

Fig. 4 Screenshot of Occupation

```
SQL> DESCRIBE EDUCATION_LEVEL;
Name                                     Null?      Type
-----
Education_ID                             NOT NULL   NUMBER(38)
Level_Name                                         VARCHAR2(20)
```

Fig. 5 Screenshot of Education Level

```
SQL> describe marital_status;
Name                                     Null?      Type
-----
Status_ID                                NOT NULL   NUMBER(38)
Status_Name                                        VARCHAR2(20)
```

Fig. 6 Screenshot of Marital Status

```
SQL> describe electoral_district;
Name                                     Null?      Type
-----
ED_Name                                  NOT NULL   VARCHAR2(20)
```

Fig. 7 Screenshot of Electoral district

```
SQL> describe city;
Name                                     Null?      Type
-----
City_Name                                NOT NULL   VARCHAR2(20)
Province_Name                            NOT NULL   VARCHAR2(20)
```

Fig. 8 Screenshot of City

```
SQL> describe language;
Name                                     Null?      Type
-----
Language_Name                            NOT NULL   VARCHAR2(20)
Address                                  NOT NULL   VARCHAR2(20)
Postal_Code                              NOT NULL   VARCHAR2(6)
```

Fig. 9 Screenshot of Language

```
SQL> describe employed;
```

Name	Null?	Type
Employer		VARCHAR2(20)
Job_Title		VARCHAR2(20)
Occupation_ID	NOT NULL	NUMBER(38)
Occupation_Name		VARCHAR2(20)

Fig. 10 Screenshot of Employed

```
SQL> describe unemployed;
```

Name	Null?	Type
Seeking_Employment		NUMBER(1)
Occupation_ID	NOT NULL	NUMBER(38)
Occupation_Name		VARCHAR2(20)

Fig. 11 Screenshot of Unemployed

```
SQL> describe student;
```

Name	Null?	Type
Institution_Name		VARCHAR2(30)
Level_of_Study		VARCHAR2(20)
Occupation_ID	NOT NULL	NUMBER(38)
Occupation_Name		VARCHAR2(20)

Fig. 12 Screenshot of Student

```
SQL> describe retired;
```

Name	Null?	Type
Age_at_Retirement		NUMBER(38)
Occupation_ID	NOT NULL	NUMBER(38)
Occupation_Name		VARCHAR2(20)

Fig. 13 Screenshot of Retired

```
SQL> describe skills;
```

Name	Null?	Type
Skill_ID	NOT NULL	NUMBER(38)
Skill_Name		VARCHAR2(50)

Fig. 14 Screenshot of Skills

```
SQL> describe individualskills;
```

Name	Null?	Type
SIN	NOT NULL	NUMBER(38)
Skill_ID	NOT NULL	NUMBER(38)

Fig. 15 Screenshot of Individual skills

```
SQL> select * from individual;
```

Individual_Name	Gender	Age	SIN	Income
Address	Postal	Occupation_ID	Education_ID	Status_ID
Martinez, John	M	45	987615431	85000
#123 ExampleStreet	V6T1Z4	1	1	1
Wong, Ashley	F	38	987615432	150000
#123 SQLStreet	V6E1M7	2	2	2
Reynolds, Olivia	F	29	987615433	70000
#111 NoSQLStreet	H2Z1B2	3	3	3
Individual_Name	Gender	Age	SIN	Income
Address	Postal	Occupation_ID	Education_ID	Status_ID
Patel, Neha	F	35	987615435	100000
#123 DataStreet	M4C1A1	5	5	5

Fig. 16 Screenshot of Individual Selection

```
SQL> select * from household;
```

Address	Postal	Number_of_Members	ED_Name
City_Name	Province_Name	Language_Name	
#123 ExampleStreet Edmonton	T5J0R4 AB	2	Edmonton Centre English
#123 SQLStreet Vancouver	V6E1M7 BC	3	Vancouver Granville Chinese
#111 NoSQLStreet Montreal	H2Z1B2 QC	1	Saint-Laurent English
Address	Postal	Number_of_Members	ED_Name
City_Name	Province_Name	Language_Name	
#123 DataStreet Toronto	M4C1A1 ON	1	Toronto-Danforth English
#123 ExampleStreet	V6T1Z4		

Fig. 17 Screenshot of Household Selection

```
SQL> select * from occupation;
```

Occupation_ID	Occupation_Name
1	Civil Engineer
2	Family Physician
3	Teacher
4	Software Engineer
5	Data Scientist

Fig. 18 Screenshot of Occupation Selection

```
SQL> select * from education_level;
```

Education_ID	Level_Name
1	Bachelors
2	MD
3	Bachelor
4	Bachelors
5	PhD

Fig. 19 Screenshot of education level Selection

```
SQL> select * from marital_status;
```

Status_ID	Status_Name
1	Single
2	Married
3	Divorced
4	Widowed
5	Separated

Fig. 20 Screenshot of marital status Selection

```
SQL> select * from electoral_district;
```

ED_Name
Edmonton Centre
Saint-Laurent
Toronto-Danforth
Toronto-St
Vancouver Granville

Fig. 21 Screenshot of electoral district Selection

```
SQL> select * from city;
```

City_Name	Province_Name
Edmonton	AB
Montreal	QC
Toronto	ON
Vancouver	BC

Fig. 22 Screenshot of city Selection

```
SQL> select * from employed;
```

Employer	Job_Title	Occupation_ID	Occupation_Name
Pacific Land Group	Civil Engineer	1	Civil Engineer
Google	Software Engineer	4	Software Engineer
SAP	Data Scientist	5	Data Scientist

Fig. 23 Screenshot of employed Selection

```
SQL> select * from student;
```

Institution_Name	Level_of_Study	Occupation_ID
UBC Civil Engineer	Bachelors	1
UofT Family Physician	MD	2
McGill Teacher	Bachelor	3
Stanford Software Engineer	Bachelors	4
UofT Data Scientist	PhD	5

Fig. 24 Screenshot of student Selection

```
SQL> select * from retired;
```

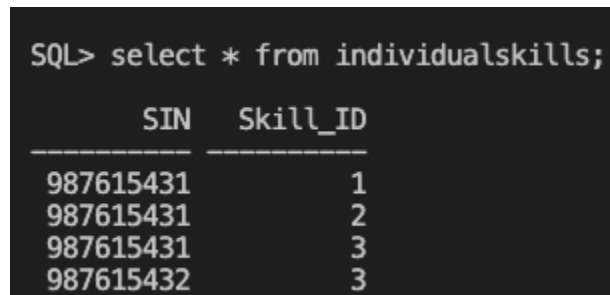
Age_at_Retirement	Occupation_ID	Occupation_Name
65	1	Civil Engineer
67	2	Family Physician
70	3	Teacher
66	4	Software Engineer
66	5	Data Scientist

Fig. 25 Screenshot of retired Selection

```
SQL> select * from skills;
```

Skill_ID	Skill_Name
1	Programming
2	Data Analysis
3	Project Management

Fig. 26 Screenshot of skills Selection



```
SQL> select * from individualskills;
```

SIN	Skill_ID
987615431	1
987615431	2
987615431	3
987615432	3

Fig. 27 Screenshot of individual skills Selection

d. A list of all SQL queries used and where it can be found in the code (i.e., file name and line number(s)). For SQL query requirements, check the rubric listed on Canvas for **Milestone 4**.

Our queries can be found in **Node_Project/queries.sql**

Our list:

-- INSERT

-- "insert an individual"

INSERT INTO

Individual (

"Individual_Name",

"Gender",

"Age",

"SIN",

"Income",

"Address",

"Postal_Code",

"Occupation_ID",

"Education_ID",

"Status_ID"

)

VALUES

(

:individualName,

:gender,

:age,

:sin,

:income,

:address,

```
:postalCode,  
:occupationId,  
:educationId,  
:statusId  
);
```

-- DELETE

-- "delete a household, deletes the individual (ON DELETE CASCADE)"

```
DELETE FROM  
    Household  
WHERE  
    "Address" = :address  
    AND "Postal_Code" = :postalCode;
```

-- UPDATE

-- "update individual address, postal code, also updates in the household using sin"

```
UPDATE  
    Individual  
SET  
    "Address" = :newAddress,  
    "Postal_Code" = :newPostalCode  
WHERE  
    "SIN" = :sin;
```

-- SELECTION

-- "select all the individuals having particular gender and income > than given income "

```
SELECT  
    *  
FROM  
    Individual  
WHERE  
    "Gender" = :gender  
    AND "Income" > :income;
```

-- PROJECTION

-- "selects name,age and address of the individual"

```
SELECT  
    "Address",  
    "Postal_Code",  
    "Individual_Name"
```

```
FROM
    Individual;
```

-- JOIN

-- "Utilizes the join of Individual and household to filter out individuals in any given city"

```
SELECT
    I."Individual_Name",
    I."Address",
    I."Postal_Code",
    H."City_Name"
FROM
    Individual I
    JOIN Household H ON I."Postal_Code" = H."Postal_Code"
WHERE
    H."City_Name" = :cityName;
```

-- AGGREGATION WITH GROUP BY

-- Get Average income by education

```
SELECT
    E."Level_Name",
    AVG(I."Income") AS "Average_Income"
FROM
    Individual I
    JOIN Education_Level E ON I."Education_ID" = E."Education_ID"
GROUP BY
    E."Level_Name";
```

-- AGGREGATION WITH HAVING

-- Get Average Income by Occupation

```
SELECT
    O."Occupation_Name",
    AVG(I."Income") AS "Average_Income"
FROM
    Individual I
    JOIN Occupation O ON I."Occupation_ID" = O."Occupation_ID"
GROUP BY
    O."Occupation_Name"
HAVING
    AVG(I."Income") > :givenIncome;
```

-- NESTED AGGREGATION BY GROUP BY

-- Get average of average income of city, only city where average income is greater than given income

```
SELECT
    City_Avg."City_Name",
    AVG(City_Avg."Average_Income") AS "City_Avg_Income"
FROM
    (
        SELECT
            H."City_Name",
            I."Individual_Name",
            AVG(I."Income") AS "Average_Income"
        FROM
            Individual I
            JOIN Household H ON I."Postal_Code" = H."Postal_Code"
        GROUP BY
            H."City_Name",
            I."Individual_Name"
    ) City_Avg
GROUP BY
    City_Avg."City_Name"
HAVING
    AVG(City_Avg."Average_Income") > :givenIncome;
```

-- DIVISION

-- Find all individuals having every skill

```
SELECT
    I."Individual_Name"
FROM
    Individual I
WHERE
    NOT EXISTS (
        SELECT
            S."Skill_ID"
        FROM
            Skills S
        WHERE
            NOT EXISTS (
                SELECT
                    ISk."SIN"
                FROM
                    IndividualSkills ISk
                WHERE
```

```

        ISk."SIN" = I."SIN"
        AND ISk."Skill_ID" = S."Skill_ID"
    )
);

```

e. Screenshots demonstrating the functionality of each query using the GUI. We want to see a before/during/after progression of events. For example, the before screenshot would be what data is in the table before you run the query, the during screenshot(s) is how the query is triggered using the GUI, and the after screenshot is what data is in your table afterwards. Please label each set of screenshots with the name of the query it is meant to address (e.g., “Insert Operation”)

i. You need only to include screenshots for the required queries – if you implemented more than what was required, screenshots are not needed for those extra queries.

INSERT:

Before:

Insert Individual (Insert)

Individual name:

Test Name

Gender:

F

Age:

21

SIN:

999999999

Income:

55000

Address:

123 testStreet

Postal Code:

V6T1Z3

Occupation ID:

6

Education ID:

6

Status ID:

6

Insert Individual

Individual inserted successfully!

AFTER:

```
SQL> SELECT * FROM Individual;
```

Individual_Name	Gender		Age	SIN	Income
Address	Postal	Occupation_ID	Education_ID	Status_ID	
Martinez, John #123 ExampleStreet	M T5J0R4	1	45 1	987615431 1	85000
Wong, Ashley #123 SQLStreet	F V6E1M7	2	38 2	987615432 2	150000
Reynolds, Olivia #111 NoSQLStreet	F H2Z1B2	3	29 3	987615433 3	70000
Individual_Name	Gender		Age	SIN	Income
Address	Postal	Occupation_ID	Education_ID	Status_ID	
Patel, Neha #123 DataStreet	F M4C1A1	5	35 5	987615435 5	100000
Test Name 123 testStreet	F V6T1Z3	6	21 6	999999999 6	55000

DELETE:

Before:

```
SQL> SELECT * FROM Household;
```

Address	Postal	Number_of_Members	ED_Name
City_Name	Province_Name		Language_Name
#123 ExampleStreet Edmonton	T5J0R4 AB	2	Edmonton Centre English
#123 SQLStreet Vancouver	V6E1M7 BC	3	Vancouver Granville Chinese
#111 NoSQLStreet Montreal	H2Z1B2 QC	1	Saint-Laurent English
Address	Postal	Number_of_Members	ED_Name
City_Name	Province_Name		Language_Name
#123 DataStreet Toronto	M4C1A1 ON	1	Toronto-Danforth English
123 testStreet	V6T1Z3		

After:

Delete Household (Delete)

Address:

Postal Code:

Delete Household

Household deleted successfully

```
SQL> SELECT * FROM Household;
```

Address	Postal	Number_of_Members	ED_Name
City_Name	Province_Name	Language_Name	
#123 ExampleStreet Edmonton	T5J0R4 AB	2	Edmonton Centre English
#123 SQLStreet Vancouver	V6E1M7 BC	3	Vancouver Granville Chinese
#111 NoSQLStreet Montreal	H2Z1B2 QC	1	Saint-Laurent English
Address	Postal	Number_of_Members	ED_Name
City_Name	Province_Name	Language_Name	
#123 DataStreet Toronto	M4C1A1 ON	1	Toronto-Danforth English

Update:

Before:

```
SQL> SELECT * FROM Individual;
```

Individual_Name	Gender	Age		SIN	Income
Address	Postal	Occupation_ID	Education_ID	Status_ID	
Martinez, John #123 ExampleStreet	M T5J0R4	1	45 1	987615431 1	85000
Wong, Ashley #123 SQLStreet	F V6E1M7	2	38 2	987615432 2	150000
Reynolds, Olivia #111 NoSQLStreet	F H2Z1B2	3	29 3	987615433 3	70000
Individual_Name	Gender	Age		SIN	Income
Address	Postal	Occupation_ID	Education_ID	Status_ID	
Patel, Neha #123 DataStreet	F M4C1A1	5	35 5	987615435 5	100000

After

Update Address or Postal Code of Individual (Update)

SIN:

987615431

New Address

New Postal Code

Updated Postal Code:

V6T1Z4

Update Household

Address or Postal Code updated

SQL> SELECT * FROM Individual;

Individual_Name	Gender	Age	SIN	Income
Address	Postal	Occupation_ID	Education_ID	Status_ID
Martinez, John	M	45	987615431	85000
#123 ExampleStreet	V6T1Z4	1	1	1
Wong, Ashley	F	38	987615432	150000
#123 SQLStreet	V6E1M7	2	2	2
Reynolds, Olivia	F	29	987615433	70000
#111 NoSQLStreet	H2Z1B2	3	3	3
Individual_Name	Gender	Age	SIN	Income
Address	Postal	Occupation_ID	Education_ID	Status_ID
Accounts				
Reynolds, Neha	F	35	987615435	100000
#123 DataStreet	M4C1A1	5	5	5

Selection:

Search Individual by Gender and Income (Selection)

Gender: Female

Minimum Income:

50000

Search

Results

Name	Gender	Age	SIN	Income	Address	Postal Code	Occupation ID	Education ID	Status ID
Wong, Ashley	F	38	987615432	150000	#123 SQLStreet	V6E1M7	2	2	2
Reynolds, Olivia	F	29	987615433	70000	#111 NoSQLStreet	H2Z1B2	3	3	3
Patel, Neha	F	35	987615435	100000	#123 DataStreet	M4C1A1	5	5	5

Projection:

Individuals' Addresses and Names (Projection)

Load Individuals

Name	Address	Postal Code
Martinez, John	#123 ExampleStreet	V6T1Z4
Wong, Ashley	#123 SQLStreet	V6E1M7
Reynolds, Olivia	#111 NoSQLStreet	H2Z1B2
Patel, Neha	#123 DataStreet	M4C1A1

Join:
Before:

```
SQL> SELECT * FROM Household;
```

Address	Postal	Number_of_Members	ED_Name
City_Name	Province_Name	Language_Name	
#123 ExampleStreet Edmonton	T5J0R4 AB	2	Edmonton Centre English
#123 SQLStreet Vancouver	V6E1M7 BC	3	Vancouver Granville Chinese
#111 NoSQLStreet Montreal	H2Z1B2 QC	1	Saint-Laurent English
Address	Postal	Number_of_Members	ED_Name
City_Name	Province_Name	Language_Name	
#123 DataStreet Toronto	M4C1A1 ON	1	Toronto-Danforth English
#123 ExampleStreet	V6T1Z4		

After:

```
SQL> SELECT * FROM City;
```

City_Name	Province_Name
Edmonton	AB
Montreal	QC
Toronto	ON
Vancouver	BC

City Name:

Vancouver

Search

Results

Name	Address	Postal Code	City Name
Wong, Ashley	#123 SQLStreet	V6E1M7	Vancouver

Aggregation:

Average Income by Education (Aggregation)

Load Individuals

Level Name	Average Income
MD	150000
Bachelors	85000
Bachelor	70000
PhD	100000

Aggregation with having

Occupation having greater average income (Aggregation with Having)

Income

Search

Results

Occupation	Average Income
Civil Engineer	85000
Data Scientist	100000
Teacher	70000
Family Physician	150000

Nested aggregation by group by

City with higher average income (Nested Aggregation)

Income

Search

Results

City	Average Income
Toronto	100000
Montreal	70000
Vancouver	150000

Division:

Individuals' with Every Skill (Division)

Load

Name
Martinez, John

```
SQL> SELECT * FROM INDIVIDUALSKILLS;

      SIN      Skill_ID
-----
987615431         1
987615431         2
987615431         3
987615432         3
```