1   Preliminaries

2   Before starting on this assignment, please be sure to read the General Instructions that are on Piazza (under Resources->Resources->General Resources), and also make sure you are able to log in to the class PostgreSQL server. You'll get help on this in your Lab Section, not the Lectures, so *be sure to attend Lab Sections*.

3   Goal

The goal of the first assignment is to create a PostgreSQL data schema with 6 tables. That is all that is required in this assignment. In your Lab Sections, you may be given information about how to load data into a table and issue simple SQL queries, because that's fun, but loading data and issuing queries is **not** required in this assignment. (That will show up in the Lab2 assignment.)

4   Lab1 Description

4.1 Create PostgreSQL Schema Lab1

As we note on the general instructions, you will create a Lab1 schema to set apart the database tables created in this lab from tables you will create in future labs, as well as from tables (and other objects) in the default (public) schema. Note that the meaning of schema here is specific to PostgreSQL, and distinct from the general meaning of schema. See here for more details on PostgreSQL schemas. You create the Lab1 schema using the following command:

```
CREATE SCHEMA Lab1;
```

Now that you have created the schema, you want to make Lab1 be the default schema when you use psql. If you do not set Lab1 as the default schema, then you will have to qualify your table names with the schema name (e.g., by writing Lab1.Customers, rather than just Customers). To set the default schema, you modify your search path as follows. (For more details, see here.)

```
ALTER ROLE username SET SEARCH_PATH to Lab1;
```

You will need to log out and log back in to the server for this default schema change to take effect.

4.2 Tables

The IRS is the Internal Revenue Service, which is responsible for collecting U.S. government taxes. You will create tables in schema Lab1 for tables TaxReturns, Payments, Individuals, Businesses, IRSagents and Delinquents as described below. Datatypes for the attributes in these "IRS" tables are described in the next section.

TaxReturns(<u>taxpayerID, taxYear</u>, kind, dateFiled, income, taxOwed)

Payments(<u>taxpayerID, datePaid</u>, amountPaid)

Individuals(<u>taxpayerID</u>, name, address, spouseID, lastDateFiled, lastDatePaid, totalTaxOwed)

Businesses(<u>taxpayerID</u>, name, address, businessType, lastDateFiled, lastDatePaid, totalTaxOwed)

IRSagents(<u>IRSagentID</u>, taxpayerID, jobLevel, active)

Delinquents(<u>taxpayerID</u>, IRSagentID, collectionAgencyID)

Important: To receive full credit, you must use the attribute names as given, and the attributes must be in the order given. Also, the datatypes must match the specifications given in the next section.   Follow directions; <u>do not</u> do more than you're asked to do in this assignment.

The underlined attribute (or attributes) identifies the primary key of the table. The tables of the schema record IRS information. *Tax returns* are submitted by tax payers who may be *individuals* (kind is 'I') or *businesses* (kind is 'B'). Tax payers may make *payments*. Slightly different information is kept about individuals and businesses. A business may have businessType 'C' (corporation'), 'P' (partnership) or 'L" (limited liability company). *IRS agents* are employees of the IRS. *Delinquents* are tax payers who have owed money for a long time; the IRS may assign an IRS agent and a collection agency to try to get payment from a delinquent.

In this assignment, you'll just have to create tables with the correct table names, attributes, datatypes and primary keys. (<u>Don't forget the primary keys!</u>)

4.2.1 Datatypes

Sometimes an attribute (such as taxpayerID) appears in more than one table. Such attributes might not have the same datatype in some schemas, but in our schema, they do.

- For taxpayerID, collectAgencyID, taxYear, spouseID and jobLevel attributes, use *integer*.
- For kind, use *character* with fixed length 1.
- For IRSagendaID, use *character* with fixed length 6.
- For name and address attributes, use *character* of variable length, with maximum length 30.
- income, taxOwed, amountPaid and totalTaxOwed should be *decimal,* with at most 9 digits to the left of the decimal point and 2 decimal digits after it.
- The dateFiled, datePaid and lastDateFiled attributes should be of type *date*.
- For the active attribute, which indicates whether an IRS agent is currently working, use *boolean*.

You must write a CREATE TABLE statement for each of the six tables in Section 3.2. Write the statements in the order that the tables are listed above. Save the statements in the file create.sql You do not need any enumerated types in this assignment, just the datatypes mentioned above.

5 Testing

While you're working on your solution, it is a good idea to drop all objects from the database every time you run the create.sql script, so you can start fresh. Dropping each object in a schema may be tedious, and sometimes there may be a particular order in which objects must be dropped. The following command, which you should put at the top of your create.sql, will drop your Lab1 schema (and all the objects within it), and then create the (empty) schema again:

DROP SCHEMA Lab1 CASCADE;
CREATE SCHEMA Lab1;

Before you submit your Lab1 solution, login to your database via psql and execute your create.sql script. As you'll learn in Lab Sections, the command to execute a script is: \i <filename>. Verify that every table has been created by using the command: \d

Also, verify that the attributes of each table are in the correct order, and that each attribute is assigned its correct datatype using the following command: \d <table>.

6 Submitting

1. Save your script as create.sql   You may add informative comments to your scripts if you want. Put any other information for the Graders in a separate README file that you may submit.

2. Zip the file(s) to a single file with name Lab1_XXXXXXX.zip where XXXXXXX is your 7-digit student ID, for example, if a student's ID is 1234567, then the file that this student submits for Lab1 should be named Lab1_1234567.zip

   If you have a README file (which is <u>not</u> required), you can use the Unix command:

   zip Lab1_1234567 create.sql README

   If you don't have a README file, to create the zip file you can use the Unix command:

   zip Lab1_1234567 create.sql

   (Of course, you use **your own student ID**, not 1234567.)

3. Submit the zip file on Canvas under Assignment Lab1. Please be sure that you have access to Canvas for CMPS 180.

   If you are working on the UNIX timeshare and your zip file is located there, you will need to copy your file to your computer so that you can upload it to Canvas through your browser. For that, you will need an FTP (File Transfer Protocol) client to securely transfer files from the UNIX timeshare. A widely used secure FTP client is Filezilla. Installation instructions are found in the site of <u>FileZilla</u> (make sure you install the distribution suitable for your operating system). After opening the Filezilla client, you will need to set the host field to unix.ucsc.edu, the username to your CruzId and the password to your Blue password, while the port number should be set to 22

(the default port for remote login). By clicking the Quickconnect button, if your credentials are correct, you will connect and be able to see the contents of your remote Unix folder at the right pane (under the title "Remote site"), while the left pane (under the title "Local site") will display the contents of your local file system. With the mouse, you can drag the file from the Unix folder and drop it to the desired location at your computer. This will transfer the file to your local machine, without erasing it from its original remote location. As noted above, Filezilla is only one of several options for an FTP client. If you are finding it difficult to install the necessary tools and successfully do file transfers, you should <u>promptly</u> ask for help in the lab sections (do not postpone it for the day of the deadline). The computers at the Lab also have pre-installed SSH and FTP clients (Putty and PSFTP).

Other approaches to copy files includes using SCP (Secure Copy) and using Cut-and-Paste, where you copy the contents of the file from the unix system, and then paste contents into a file on your computer. Cut-and-Paste may work with for small files, but it's a hack that would not work well for large files.

The CMPS 180 Teaching Assistants, Avani and Akhil, will discuss approaches with you during Lab Sections. Be sure that you know how to handle this,

4. Lab1 is due by 11:59pm on Wednesday, Jan 24. **Late submissions will not be accepted (Canvas won't take them), and there will be no make-up Lab assignments.** Check to make sure that your submission is on Canvas, and that you've submitted the correct file. You will receive **no credit** if you accidentally submitting the wrong file, even if you show that you completed the correct file on time.