CLiC Microscopic Data Organization Database (Group 68)

Stone Chen, Yunwen Ji, Wenquan Li, Yuki Zhang

February 4, 2019

1 Requirement Analysis

1.1 Introduction

1.1.1 Purpose

The database revolves around the CLiC raw data in building its connections with three most relevant components: the lab member, the experiment, and the analysis results. This architecture would benefit all users (most likely the lab members) in efficiently retrieving pertinent information, such as the individual responsible for a dataset, the experimental procedures, the file PATHs, etc.

1.1.2 Scope and Special Design Choice

We limit the sphere of our interest to only the biomedical imaging component of the lab, as there is huge need for safe and efficient data management, particularly for this branch of the lab. Note that this database is not responsible for storing multidimensional arrays such as videos, images, and analysis related information; instead, those information are backed up in the lab cloud drive, of which the respective PATHs are saved as attributes to corresponding entities.

1.1.3 Terminology

CLiC Microscopy: Convex Lense induced Confinement Microscopy

1.1.4 Resources

https://pubs.acs.org/doi/10.1021/ac101041s. Please also see the appendix for a brief overview of this method.

1.2 Database Description

1.2.1 Entities and Attributes

- **Lab Members**: Those who are currently working or have worked in the lab, including research assistants and researchers. They both share the attribute name, contact information, and status. Name is key attribute as there is very low chance of duplication. Status can have value of active/inactive, indicating whether the individual still works here.
- **Research Assistant**: A research assistant (undergraduates) is a subclass of lab members, it has no additional attribute but unique relationships.
- **Researcher**: A researcher is a subclass of lab members with a special attribute named level, which indicates the degree/job level in the lab, for example, PhD. It also has unique relationship compared to research assistant
- **CLiC Raw Data**: The raw data is generated using a microscope installed with CLiC in the format of TIFF time series. This multidimensional array will not be saved directly in the database but rather backed up in the cloud drive, where the unique PATHs are captured by our database. This entity has an

artificial ID as key as well as an additional attribute date. Furthermore, since there are different types of data generated, which will subsequently be studied differently, this raw data contains two subclasses depending on the type of data.

- **Nanoparticle Data**: This subclass of raw data represent the data used in the studies of nanoparticle. It has one unique relationship compared to the other subclass.

- **DNA Unwinding Data**: This subclass of raw data represent the data used in the studies of DNA unwinding. It has one unique relationship compared to the other subclass.

- **Binding Analysis Result**: This entity represents the result analyzed by research assistant based on DNA unwinding data. These results are different from nanoparticle analysis results in the procedures performed as well as the information saved, which includes binding count, diffusion coefficient etc. Those results are not saved directly as attributes but backed up in the cloud drive, of which the PATH is stored as an attribute of this entity. In addition, this entity contains two other attribute: date and analysis type. Date is just the date those results are generated, while analysis type can have values of 'rough','decent' or 'finalized', which represents the degree of confidence in the accuracy of a particular result. We decide to save this as an attribute because this can help users decide whether or not they would like to get the PATH and check this result. For example, if a result is of analysis type 'rough', the user may not consider it valuable to its purposes. Finally, the key of this entity is an artificial ID.

- **Nanoparticle Analysis Result**: This entity represents the result analyzed by research assistant based on nanoparticle data. These results are different from DNA binding analysis results in the procedures performed as well as the information saved, which includes mean square displacement, number of particles etc. Those results are not saved directly as attributes but backed up in the cloud drive, of which the PATH is stored as an attribute of this entity. In addition, this entity contains two other attribute: date and MSD (Mean Square Displacement). Date is just the date those results are generated, while MSD can assume certain numerical values. We decided to save MSD as an attribute because again it is an important analysis result that a researcher would like to know before taking a closer look at the whole dataset. For instance, if a researcher decides to review all results with MSD larger than k, then this setup makes it convenient to do so. Finally, the key of this entity is an artificial ID.

- **CLiC Dashboard**: The entity represent the quality control result of CLiC raw data, which is performed before all analysis and presented in format of a dashboard. It includes general data information and specific quality information that includes the focus, SNR (signal-to-Noise Ratio) and contaminant level. All those information will be backed up in cloud drive, where the corresponding PATH is saved as an attribute of this entity. In addition, since quality information are the strict requirements to deem a raw data as valuable for analysis, we decided to save them as attributes so that users can quickly obtain these information before proceeding to analysis. Last but not least, this entity also has an attribute named date to store the date these results are generated. Finally, the key of this entity is an artificial ID.

- **Protocol**: This is the outline of experimental procedures, assays and safety information etc. Researchers formulate protocols before performing experiments so that they ensure they have made comprehensive planning. Each researcher has their own ways of naming their protocols, including typically their initials and some keywords regarding their experiment, which is always unique. Thus we use those protocol IDs as key of protocol. Since those protocols are usually saved and backed up somewhere, an attribute of their PATHs is also saved.

- **Microscope**: This entity represents the microscopes, which are used to perform experiments and generate raw data. There are currently two sets of microscopes in the lab and a third one is underway. Each device has a unique ID that starts with 'Nikon' and a number, for example 'Nikon 1'. Therefore, those unique microscope IDs are used as the key of this entity. Furthermore, those microscopes can be on temporary maintenance sometimes so an attribute named status is used; it assumes values of either 'active' or 'inactive'.

- **Experiment**: This weak entity represents an experiment that follows a certain protocol with a particular set of microscope. Often times, researchers replicate experiments using the same protocol to collect more data or verify their results. To represent this in ER model, experiment is designed to be a weak entity that is dependent on a protocol; it has a partial key named 'try number' that indicates the replicate number. For example, if try number is equal to 3, then this particular experiment is the third replicate of a certain protocol. Obviously this try number is not unique unless it is tailored to a specific protocol that has a unique protocol ID.

1.2.2 Relationships

- **is a (from lab members)**: *Lab members* have two different subclasses: *researchers* and *research assistants*, who conduct distinctive tasks in the lab.

- **supervise**: A *researcher* supervises a *research assistant*. This is a many-to-many relationship because a researcher can supervise more than one research assistant and one research assistant can be supervised by more than one researcher. There is a participation constraint on the research assistant, which requires that every research assistant must be supervised by at least one researcher.

- **formulate**: A *researcher* formulates a *protocol*. This is a many-to-one relationship because a researcher can formulate many protocols but a protocol can only be formulated by one researcher. There is also a participation constraint on the protocol because every protocol must be formulated by at least one researcher, indicating that every protocol must be formulated by exactly one researcher incorporated with the key constraint.

- **generate**: A *researcher* generates a set of *CliC raw data* through an *experiment*. This is a many-one-many relationship because an experiment can only be conducted by one researcher but a researcher can conduct many experiments. In addition, a researcher can generate many sets of data and a set of data can be generated by multiple researchers.

- **using**: The *experiment* uses the *protocol*. This is a one-to-many relationship since one experiment can only apply one protocol while one protocol can be utilized by replicate experiments. There is a participation constraint on the experiment because every experiment must be conducted according to at least one protocol, indicating that every experiment uses exactly one protocol incorporated with the key constraint.

- **with**: An *experiment* is conducted with a *microscope*. This is a one-to-many relationship because an experiment can only use one microscope while one microscope can be used by plenty of experiments. There is also a participation constraint on the experiment because every experiment must be conducted with at least one microscope, indicating that every experiment must use exactly one microscope incorporated with the key constraint.

- **is a (Clic Raw Data)**: The CliC new data has two subclasses: nanoparticle data and DNA unwinding data.

- **Assess**: The *research assistant* assesses the *CliC new data* to generate the *CliC dashboard*. This is a many-one-many relationship since one research assistant can assess more than one set of CliC raw data to get more than one CliC dashboard. One set of CliC raw data or one dashboard can only be assessed by one research assistant and one dashboard can only be generated from one set of raw data. There is a participation constraint on the CliC dashboard because every dashboard must be generated by at least one research assistant and at least one set of CliC raw data.

- **Analyze (DNA unwinding data or nanoparticle data)**: The *research assistant* analyzes the *DNA unwinding data (nanoparticle data)* to generate *binding analysis results (nanoparticle analysis results)*. This is a many-one-many relationship since one research assistant can analyze more than one set of DNA unwinding data to generate more than one set of binding analysis results. One set of DNA unwinding data or binding analysis results can only be analyzed or obtained by one research assistant. One set of binding analysis results can only be produced by one set of DNA unwinding data. There is a participation

constraint on the binding analysis results since every set of binding analysis results must be achieved by at least one research assistant and at least one set of DNA unwinding data.

2 Functional Analysis

2.1 Overview

2.1.1 Why do we need to store the CLiC data and its derivatives?

The lab members perform their research utilizing this microscopy technique extensively, which means that a huge amount of data is generated every day. Those data play a critical role in testing hypothesis and supporting conclusions in the day-to-day work of every researcher. Therefore, the lab has already set up cloud and local drives for backup purposes as well as a set of strict rules to follow in storing those data.

2.1.2 Why do need this database?

Despite those data are safely stored, researchers in the lab still complains about the difficulty to retrieve those data and their relevant information efficiently. For example, a researcher could not recall whether a dataset has been analyzed or not and spent a whole day looking for it on the cloud. He ended up finding nothing so he asked a research assistant to analyze it again which took him another whole day. This example can be even worse if that data has indeed been analyzed but just no where to find. Therefore, if there can be a magic button that one can simply press to find and present all they want in front of them, that will be exactly what we long to have. In practice, we can have something very similar which only requires more than one simple press — the database that store the links between CLiC data and all its relevant aspects!

2.2 Application Description

The central component of this database is the CLiC raw data and the application seeks to link all relevant components, which includes three major components: the lab members, the experiments, and the analysis results. Currently, those that need to be backed up on the hard drive are consisted of research protocol, CLiC raw data, CLiC dashboard, and CLiC data analysis results (in order of when they are generated). As soon as those are backed up on the cloud, their corresponding component in the database is expected to be updated accordingly. In this way, if one needs to retrieve relevant information on one dataset, all relationship that stems from the raw data entity will offer immediate feedback. Potentially, we would like the database to have the following functions:

2.2.1 Experimental details? Is it legit?

Researchers and their PI usually review the experimental details before looking at the results to ensure the data is valuable to answer some questions. Given a particular dataset, this database can report on where the protocol is, which microscope was used as well as the number of replicates (try number) etc.

2.2.2 Data analyzed? I need it NOW!!

The main intersection between the work of a researcher and a researcher assistant is one ultimate question — Is this dataset analyzed? This database will help them interact better and more efficiently since this question can be answered by checking whether there is a analysis result entity with date specified linked to the particular set of raw data. Researchers will be able to know whether their assistants are idling, while research assistants can have a clearer idea of the upcoming work to do.

2.2.3 This dataset is of such poor quality!! Prove it.

From time to time, some dataset generated can be of poor quality and luckily we have quality control measures in place. This assessment can be performed by any research assistant and backed up, however more importantly, those who will potentially be assigned this analysis work need to be aware of it so that they do not waste their

time. Rest assure, our database can capture this information through the relationship between raw data and CLiC Dashboard. Some of the hardline quality control parameters are saved directly as attributes, so that in most cases, with a quick reference of those values, a dataset can be trashed immediately if not qualified.

2.2.4 Data analyzed!! Great, where is the one I want?

The greatest pain often comes from actually finding the particular subsets of data out of the pools of dataset. For example, find PATHs of all DNA unwinding analysis result in after Sept, 2018 that are decent counts. This is such a tedious task that every researcher suffers from in the lab. Nevertheless, with this database, all the PATHs can be output immediately with a line of query and potentially we can implement this as a user-friendly function (the magic button).

2.2.5 Who is responsible??

This is a general question asked for everything in the lab. Conveniently, our database captures the person responsible for every task, or in other way around, what a person is responsible for in his work. This even includes the supervision responsibility between researcher and research assistant.

3 Relations

Entities

labMember (Name, contactInformation, Status)
Researcher (Name, contactInformation, Status, Level)
researchAssistant (Name, contactInformation, Status)

CLiCDashboard (dID, Date, PATH, Focus, SNR, Contaminants, crID, Name) (crID ref CLiC Raw Data(crID), Name ref researchAssistant (Name))

nanoparticleAnalysisResults (narID, Date, PATH, MSD, Name, crID) (crID ref nanoparticleData (crID), Name ref researchAssistant (Name))

bindingAnalysisResults (barID, Date, PATH, nnalysisType, Name, crID) (crID ref DNAUnwindingData (crID), Name ref researchAssistant (Name))

CLiCRawData (crID, Date, PATH)
nanoparticleData (crID)
DNAUnwindingData (crID)
Protocol (ProtocalID, PATH, Name) (Name ref Researcher (Name))
Microscope (MicroscopeID, Status)

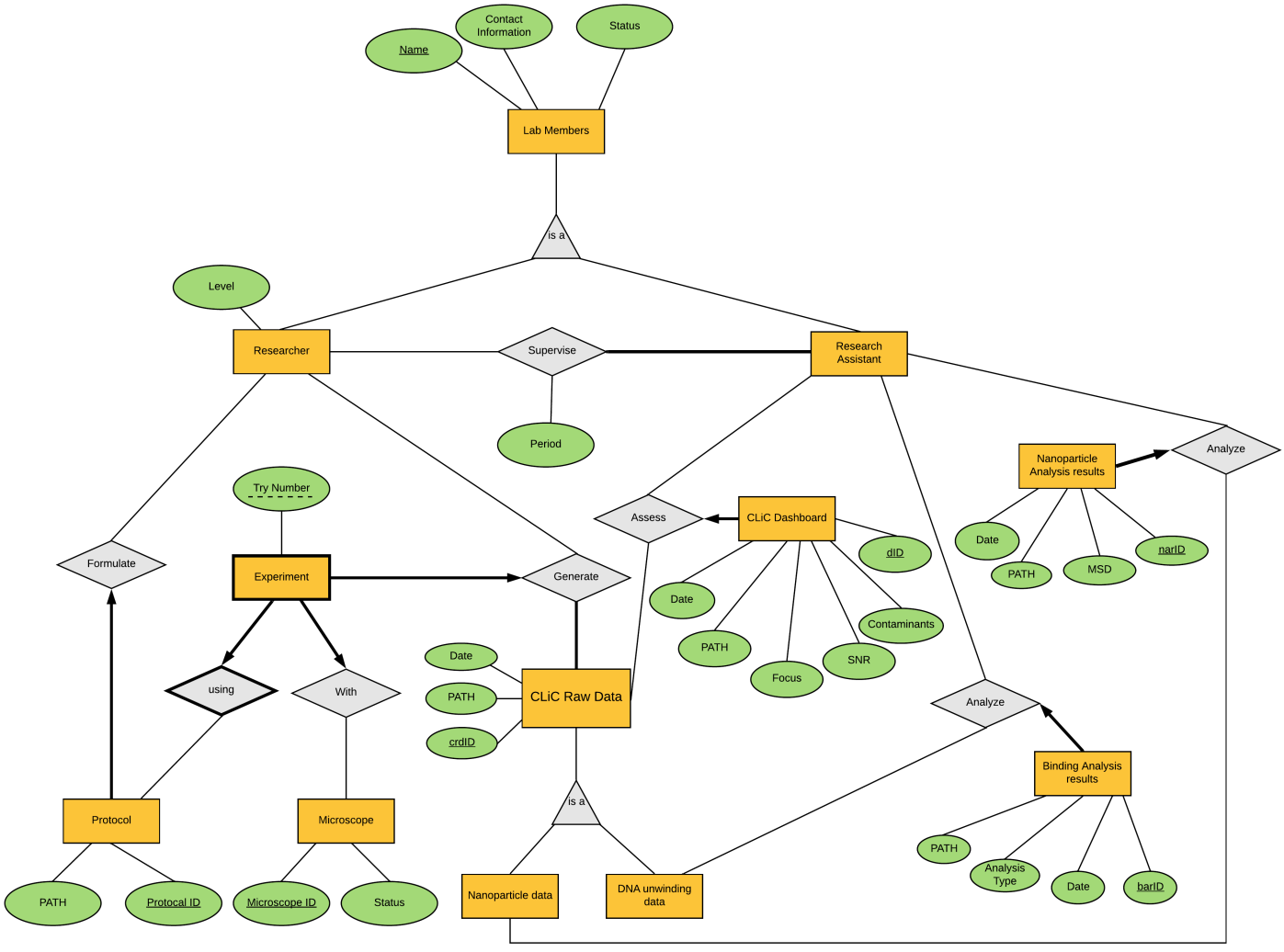
Weak Entity

Experiment(protocolID, tryNumber, microscopeID, crdID, Name) (microscopeID ref Microscope(microscopeID), protocolID ref Protocol(protocolID), crdID ref CliCRawData(crdID), Name ref Researcher(Name))

Relationship

Supervise (Period, rName, raName) (rName ref Researcher (Name), raName ref researchAssistant (Name))

We did not feel there was a way to combine any relations. For example, if we did not use a relation table for the ER relationship **Supervise**, but save this as an attribute in the **researchAssistant** table, we would not be able store past records. This sitation could occur if a research assistant had worked with multiple researcher.


Contact Information
Status
Name
Lab Members
is a
Level
Researcher
Supervise
Research Assistant
Period
Analyze
Nanoparticle Analysis results
Try Number
Assess
CLiC Dashboard
Formulate
Experiment
Generate
dID
Date
PATH
MSD
narID
Contaminants
Date
PATH
SNR
Focus
using
With
Date
PATH
CLiC Raw Data
crdID
Analyze
Binding Analysis results
Protocol
Microscope
is a
PATH
Analysis Type
Date
barID
PATH
Protocal ID
Microscope ID
Status
Nanoparticle data
DNA unwinding data

CLiC Microscopic Data Organization Database P2 (Group 68)

Stone Chen, Yunwen Ji, Wenquan Li, Yuki Zhang

March 2, 2019

1 Modified Relations

The feedback we received is as follows:

Sub-entities in ISA hierarchies should explicitly reference the super-entity in the relational translation.

Entities

LabMember (Name, contactInformation, Status)
Researcher (Name, contactInformation, Status, Level) (Name ref LabMember(Name))
ResearchAssistant (Name, contactInformation, Status) (Name ref LabMember(Name))

CLiCDashboard (dID, Date, PATH, Focus, SNR, Contaminants, crID, Name) (crID ref CLiC Raw Data(crID), Name ref researchAssistant (Name))

nanoparticleAnalysisResults (narID, Date, PATH, AvgRadius, Name, crID) (crID ref nanoparticleData (crID), Name ref researchAssistant (Name))

bindingAnalysisResults (barID, Date, PATH, analysisType, Name, crID) (crID ref DNAUnwindingData (crID), Name ref researchAssistant (Name))

CLiCRawData (crID, Date, PATH)
nanoparticleData (crID)(crID ref CLiCRawData(crID))
DNAUnwindingData (crID)(crID ref CLiCRawData(crID))
Protocol (ProtocalID, PATH, Name) (Name ref Researcher (Name))
Microscope (MicroscopeID, Status)

Weak Entity

Experiment(protocolID, tryNumber, microscopeID, crdID, Name) (microscopeID ref Microscope(microscopeID), protocolID ref Protocol(protocolID), crdID ref CliCRawData(crdID), Name ref Researcher(Name))

Relationship

Supervise (rName, raName, StartDate, EndDate) (rName ref Researcher (Name), raName ref researchAssistant (Name))

2 Database Construction

Please refer to ...sql for more information. The tables created after running all the setup commands are demonstrated in Fig.1.

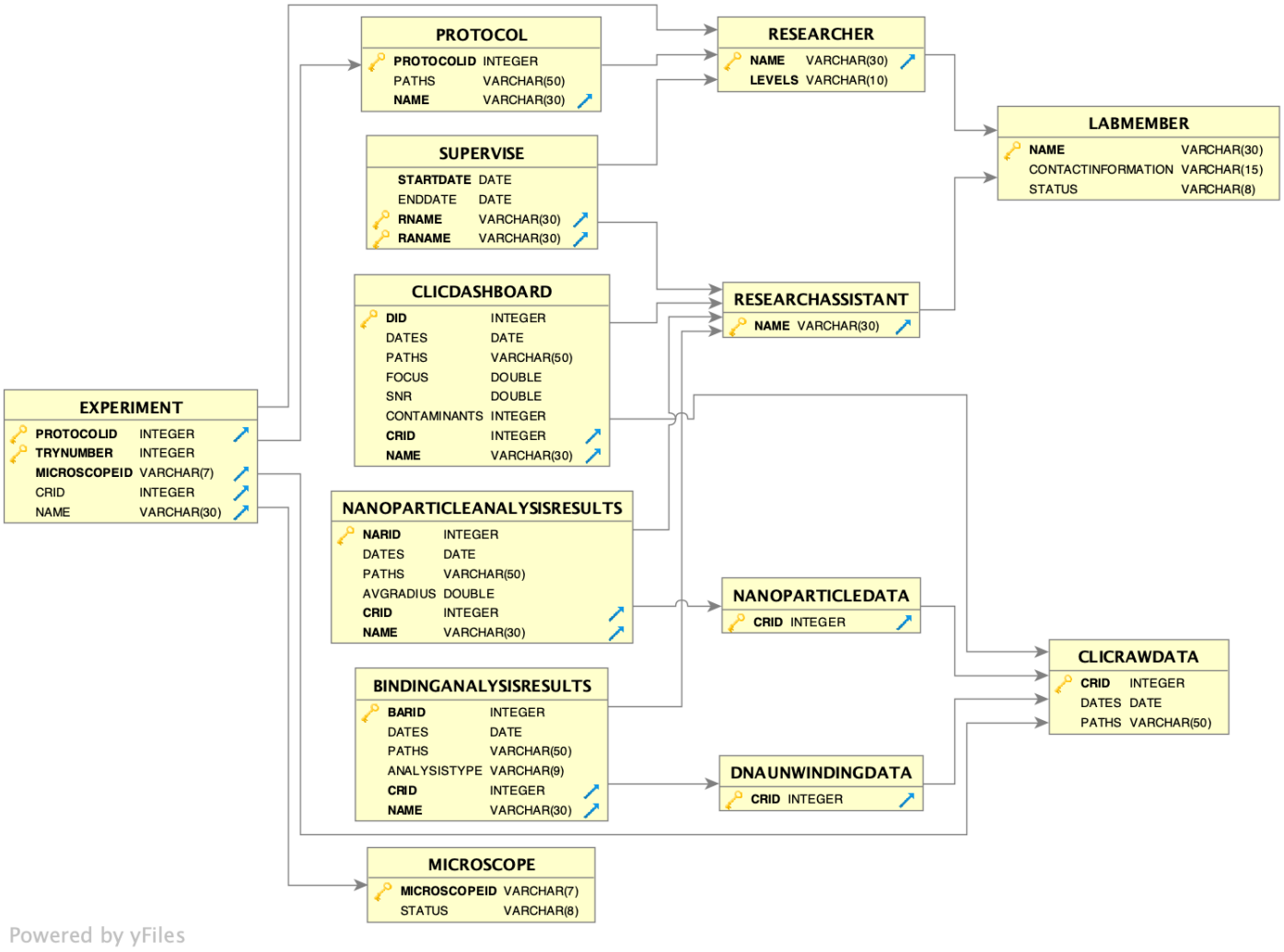


Figure 1: Database framework after running all create table commands. This representation is generated with DBVisualizer. Bold Text = "Not Null"; Key sign = "Primary Key"; Blue Arrow = "Foreign Key"

3 Manual Insert

Please refer to *q3.sql* for details on the insert commands. We inserted multiple triplet data into the *labmember* relation as well as its sub-classes *researchassistant* and *researcher*. In particular, *name* and *status* contain real data, while *contactinformation* is randomly generated. The newly inserted information in each of the three tables is shown, respectively:

```
[db2 => SELECT * FROM labmember
[db2 (cont.) => ;

NAME                          CONTACTINFORMATION STATUS
----------------------------- ------------------ --------
Wendy Ji                      (514)-561-5587     active
Stone Chen                    (514)-531-3322     active
Zhi Zhang                     (514)-557-3232     active
Zach Friedenberger            (514)-324-1237     active
Gracia Gu                     (514)-545-5123     active
Zhi Yue Zhang                 (514)-512-5567     active
Cindy Shaheen                 (514)-532-8521     active
Francis Stabile               (514)-274-4887     active
Daniel Berard                 (514)-244-9817     active
Kim Metera                    (514)-274-4921     active
Radin Tahvildari              (514)-284-1237     active
Romain Berti                  (514)-224-9387     active
Sabrina Leslie                (514)-224-2837     active
Shane Scott                   (514)-274-9287     inactive
Haoran Liao                   (514)-271-1927     inactive
Yash Patel                    (514)-292-1927     inactive

  16 record(s) selected.

[db2 => SELECT * FROM researcher;

NAME                          LEVELS
----------------------------- ----------
Zhi Yue Zhang                 master
Cindy Shaheen                 PhD
Francis Stabile               PhD
Daniel Berard                 associate
Kim Metera                    associate
Radin Tahvildari              associate
Romain Berti                  associate
Sabrina Leslie                PI
Shane Scott                   PhD

  9 record(s) selected.

[db2 => SELECT * FROM researchassistant
[db2 (cont.) => ;

NAME
----------------------------
Gracia Gu
Haoran Liao
Stone Chen
Wendy Ji
Yash Patel
Zach Friedenberger
Zhi Zhang

  7 record(s) selected.
```

4 Auto Data Generation

Please refer to *autoDataGen.py* and *autoDataGen.sql* for details on the automatically generated *Insert* commands. We used python to automatically generate 200 *Insert* commands for each table. Their connections are randomly constructed but still stick to the rules specified for each relation. Some inserted information are shown below:

```
db2 => SELECT * FROM protocol FETCH FIRST 10 ROWS ONLY

PROTOCOLID  PATHS                                              NAME
----------- -------------------------------------------------- --------------------------------
          2 C:\Protocol\2.pdf                                  Kim Metera
          3 C:\Protocol\3.pdf                                  Cindy Shaheen
          4 C:\Protocol\4.pdf                                  Radin Tahvildari
          5 C:\Protocol\5.pdf                                  Radin Tahvildari
          6 C:\Protocol\6.pdf                                  Francis Stabile
          7 C:\Protocol\7.pdf                                  Radin Tahvildari
          8 C:\Protocol\8.pdf                                  Daniel Berard
          9 C:\Protocol\9.pdf                                  Francis Stabile
         10 C:\Protocol\10.pdf                                 Cindy Shaheen
         11 C:\Protocol\11.pdf                                 Daniel Berard

  10 record(s) selected.


db2 => SELECT * FROM experiment FETCH FIRST 10 ROWS ONLY

PROTOCOLID  TRYNUMBER   MICROSCOPEID CRID        NAME
----------- ----------- ------------ ----------- --------------------------------
          2           1 Nikon01              183 Daniel Berard
          2           2 Nikon01               11 Daniel Berard
          2           3 Nikon01               33 Daniel Berard
          2           4 Nikon01               37 Daniel Berard
          3           1 Nikon01               33 Cindy Shaheen
          3           2 Nikon02              175 Cindy Shaheen
          4           1 Nikon02               35 Sabrina Leslie
          4           2 Nikon02               93 Sabrina Leslie
          4           3 Nikon02              105 Sabrina Leslie
          4           4 Nikon02              174 Sabrina Leslie

  10 record(s) selected.


db2 => SELECT * FROM BINDINGANALYSISRESULTS FETCH FIRST 10 ROWS ONLY

BARID       DATES      PATHS                                              ANALYSISTYPE CRID        NAME
----------- ---------- -------------------------------------------------- ------------ ----------- --------------------------------
          1 12/16/2018 C:\BindingAnalysisResults\1                        rough                154 Stone Chen
          2 05/22/2018 C:\BindingAnalysisResults\2                        rough                 57 Zach Friedenberger
          3 11/01/2018 C:\BindingAnalysisResults\3                        rough                 24 Stone Chen
          4 10/21/2018 C:\BindingAnalysisResults\4                        accurate             167 Wendy Ji
          5 06/01/2018 C:\BindingAnalysisResults\5                        accurate             128 Stone Chen
          6 01/12/2019 C:\BindingAnalysisResults\6                        rough                 95 Wendy Ji
          7 08/05/2018 C:\BindingAnalysisResults\7                        decent               185 Stone Chen
          8 09/15/2018 C:\BindingAnalysisResults\8                        accurate             156 Stone Chen
          9 09/22/2018 C:\BindingAnalysisResults\9                        rough                124 Zhi Zhang
         10 12/18/2018 C:\BindingAnalysisResults\10                       accurate             178 Zhi Zhang

  10 record(s) selected.


db2 => SELECT * FROM CLICDASHBOARD FETCH FIRST 10 ROWS ONLY

DID         DATES      PATHS                                              FOCUS                     SNR                       CONTAMINANTS CRID        NAME
----------- ---------- -------------------------------------------------- ------------------------- ------------------------- ------------ ----------- --------------------------------
          1 06/27/2018 C:\CLiCDashBoard\1                                  +1.50000000000000E+000    +5.47000000000000E+000          2539         175 Zhi Zhang
          2 02/23/2019 C:\CLiCDashBoard\2                                  +1.76000000000000E+000    +1.21000000000000E+000           600         186 Stone Chen
          3 05/07/2018 C:\CLiCDashBoard\3                                  +4.82000000000000E+000    +9.82000000000000E+000           145          43 Stone Chen
          4 10/04/2018 C:\CLiCDashBoard\4                                  +3.06000000000000E+000    +6.63000000000000E+000          2646         127 Zach Friedenberger
          5 08/22/2018 C:\CLiCDashBoard\5                                  +2.90000000000000E+000    +4.62000000000000E+000          1957         128 Zhi Zhang
          6 07/03/2018 C:\CLiCDashBoard\6                                  +2.95000000000000E+000    +9.03000000000000E+000           636          80 Stone Chen
          7 09/05/2018 C:\CLiCDashBoard\7                                  +4.71000000000000E+000    +1.85000000000000E+000          1845         192 Stone Chen
          8 10/06/2018 C:\CLiCDashBoard\8                                  +1.12000000000000E+000    +8.89000000000000E+000          3427         153 Zach Friedenberger
          9 07/24/2018 C:\CLiCDashBoard\9                                  +4.01000000000000E+000    +1.40000000000000E+000           773           1 Wendy Ji
         10 08/28/2018 C:\CLiCDashBoard\10                                 +3.09000000000000E+000    +5.41000000000000E+000          1288         115 Stone Chen

  10 record(s) selected.


db2 => SELECT * FROM CLICDASHBOARD FETCH FIRST 10 ROWS ONLY

DID         DATES      PATHS                                              FOCUS                     SNR                       CONTAMINANTS CRID        NAME
----------- ---------- -------------------------------------------------- ------------------------- ------------------------- ------------ ----------- --------------------------------
          1 06/27/2018 C:\CLiCDashBoard\1                                  +1.50000000000000E+000    +5.47000000000000E+000          2539         175 Zhi Zhang
          2 02/23/2019 C:\CLiCDashBoard\2                                  +1.76000000000000E+000    +1.21000000000000E+000           600         186 Stone Chen
          3 05/07/2018 C:\CLiCDashBoard\3                                  +4.82000000000000E+000    +9.82000000000000E+000           145          43 Stone Chen
          4 10/04/2018 C:\CLiCDashBoard\4                                  +3.06000000000000E+000    +6.63000000000000E+000          2646         127 Zach Friedenberger
          5 08/22/2018 C:\CLiCDashBoard\5                                  +2.90000000000000E+000    +4.62000000000000E+000          1957         128 Zhi Zhang
          6 07/03/2018 C:\CLiCDashBoard\6                                  +2.95000000000000E+000    +9.03000000000000E+000           636          80 Stone Chen
          7 09/05/2018 C:\CLiCDashBoard\7                                  +4.71000000000000E+000    +1.85000000000000E+000          1845         192 Stone Chen
          8 10/06/2018 C:\CLiCDashBoard\8                                  +1.12000000000000E+000    +8.89000000000000E+000          3427         153 Zach Friedenberger
          9 07/24/2018 C:\CLiCDashBoard\9                                  +4.01000000000000E+000    +1.40000000000000E+000           773           1 Wendy Ji
         10 08/28/2018 C:\CLiCDashBoard\10                                 +3.09000000000000E+000    +5.41000000000000E+000          1288         115 Stone Chen

  10 record(s) selected.


db2 => SELECT * FROM CLICRAWDATA FETCH FIRST 10 ROWS ONLY

CRID        DATES      PATHS
----------- ---------- --------------------------------------------------
          1 07/20/2018 C:\CliCRawData\1.mat
          2 11/15/2018 C:\CliCRawData\2.mat
          3 07/17/2018 C:\CliCRawData\3.mat
          4 07/11/2018 C:\CliCRawData\4.mat
          5 05/07/2018 C:\CliCRawData\5.mat
          6 10/20/2018 C:\CliCRawData\6.mat
          7 10/02/2018 C:\CliCRawData\7.mat
          8 05/14/2018 C:\CliCRawData\8.mat
          9 06/11/2018 C:\CliCRawData\9.mat
         10 07/13/2018 C:\CliCRawData\10.mat

  10 record(s) selected.


db2 => SELECT * FROM MICROSCOPE FETCH FIRST 10 ROWS ONLY

MICROSCOPEID STATUS
------------ --------
Nikon01      active
Nikon02      active
Nikon03      inactive

  3 record(s) selected.


db2 => SELECT * FROM NANOPARTICLEANALYSISRESULTS FETCH FIRST 10 ROWS ONLY

NARID       DATES      PATHS                                              AVGRADIUS                 CRID        NAME
----------- ---------- -------------------------------------------------- ------------------------- ----------- --------------------------------
          0 02/18/2019 C:\NanoparticleAnalysisResults\1                   +4.88400000000000E+001            85 Stone Chen
          1 12/26/2018 C:\NanoparticleAnalysisResults\2                   +3.70800000000000E+001             6 Wendy Ji
          2 08/08/2018 C:\NanoparticleAnalysisResults\3                   +4.17500000000000E+001           102 Wendy Ji
          3 08/20/2018 C:\NanoparticleAnalysisResults\4                   +4.21600000000000E+001           193 Stone Chen
          4 08/09/2018 C:\NanoparticleAnalysisResults\5                   +3.45600000000000E+001           152 Zhi Zhang
          5 08/11/2018 C:\NanoparticleAnalysisResults\6                   +3.82200000000000E+001            60 Wendy Ji
          6 08/07/2018 C:\NanoparticleAnalysisResults\7                   +3.64000000000000E+001           189 Zach Friedenberger
          7 08/06/2018 C:\NanoparticleAnalysisResults\8                   +2.87700000000000E+001            29 Stone Chen
          8 05/20/2018 C:\NanoparticleAnalysisResults\9                   +2.25200000000000E+001            64 Stone Chen
          9 12/28/2018 C:\NanoparticleAnalysisResults\10                  +4.97600000000000E+001           133 Zach Friedenberger

  10 record(s) selected.
```

```
db2 => SELECT * FROM NANOPARTICLEDATA FETCH FIRST 10 ROWS ONLY

CRID
-----------
          1
          2
          3
          4
          5
          6
          7
          8
          9
         10

  10 record(s) selected.


db2 => SELECT * FROM SUPERVISE FETCH FIRST 10 ROWS ONLY

STARTDATE  ENDDATE    RNAME                           RANAME
---------- ---------- ------------------------------- -------------------------------
05/01/2018 -          Radin Tahvildari                Wendy Ji
05/01/2018 -          Radin Tahvildari                Zach Friedenberger
02/01/2017 06/01/2018 Radin Tahvildari                Haoran Liao
02/01/2017 06/01/2018 Radin Tahvildari                Yash Patel
05/01/2018 -          Cindy Shaheen                   Stone Chen
05/01/2018 -          Cindy Shaheen                   Zhi Zhang
05/01/2018 -          Cindy Shaheen                   Gracia Gu

  7 record(s) selected.


db2 => SELECT * FROM RESEARCHASSISTANT FETCH FIRST 10 ROWS ONLY

NAME
-------------------------------
Gracia Gu
Haoran Liao
Stone Chen
Wendy Ji
Yash Patel
Zach Friedenberger
Zhi Zhang

  7 record(s) selected.


db2 => SELECT * FROM RESEARCHER FETCH FIRST 10 ROWS ONLY

NAME                            LEVELS
------------------------------- ----------
Zhi Yue Zhang                   master
Cindy Shaheen                   PhD
Francis Stabile                 PhD
Daniel Berard                   associate
Kim Metera                      associate
Radin Tahvildari                associate
Romain Berti                    associate
Sabrina Leslie                  PI
Shane Scott                     PhD

  9 record(s) selected.
```

5 Analytical Query

5.1 Digging around!!

List the paths of all data analysis results (including binding analysis and nano-particle analysis) analyzed by an active researcher named "Wendy Ji":

```
1 SELECT PATHS
2 FROM NANOPARTICLEANALYSISRESULTS
3 WHERE NAME = 'Wendy Ji'
4 UNION
5 SELECT PATHS
6 FROM BINDINGANALYSISRESULTS
7 WHERE NAME = 'Wendy Ji';
```

7:25 [144] INS

Log | 1: NANOPARTICLEANALYSISRESULTS [93] ×

*	PATHS
22	C:\BindingAnalysisResults\179
23	C:\BindingAnalysisResults\18
24	C:\BindingAnalysisResults\182
25	C:\BindingAnalysisResults\185
26	C:\BindingAnalysisResults\186
27	C:\BindingAnalysisResults\187
28	C:\BindingAnalysisResults\19
29	C:\BindingAnalysisResults\199
30	C:\BindingAnalysisResults\23
31	C:\BindingAnalysisResults\25
32	C:\BindingAnalysisResults\27
33	C:\BindingAnalysisResults\28
34	C:\BindingAnalysisResults\4
35	C:\BindingAnalysisResults\44
36	C:\BindingAnalysisResults\47
37	C:\BindingAnalysisResults\50
38	C:\BindingAnalysisResults\6
39	C:\BindingAnalysisResults\77
40	C:\BindingAnalysisResults\93
41	C:\BindingAnalysisResults\94
42	C:\BindingAnalysisResults\95
43	C:\NanoparticleAnalysisResults\101
44	C:\NanoparticleAnalysisResults\105
45	C:\NanoparticleAnalysisResults\114
46	C:\NanoparticleAnalysisResults\118

Format: Text

Time	Status	Command	Exec	Fetch	Rows	Message	SQL/Command
22:25:53	STARTED					Executing for: 'cs421g' [DB2 LUW], Schema: CS421G68	
22:25:53	SUCCESS	SELECT	0.025	0.018	93	Result set fetched	SELECT PATHS ...
22:25:54	FINISHED		0.025	0.018	93	Success: 1	

5.2 Who has been slacking?

Rank active research assistants (name) by the number of CLiC raw data they have analyzed during January and February of 2019:

```
9
10 SELECT C.NAME, COUNT(UNIQUE(CRID)) AS NUMBEROFDATA
11 FROM CLICDASHBOARD C, LABMEMBER L
12 WHERE C.DATES >= '2019-01-01' AND C.DATES <= '2019-02-28' AND C.NAME = L.NAME AND L.STATUS='active'
13 GROUP BY C.NAME
14 ORDER BY NUMBEROFDATA;
15
```

7:25 [144] INS

Log | 1: CLICDASHBOARD [4]

*	NAME	NUMBEROFDATA
1	Zhi Zhang	4
2	Zach Friedenberger	8
3	Stone Chen	12
4	Wendy Ji	12

Time	Status	Command	Exec	Fetch	Rows	Message	SQL/Command
22:28:51	STARTED					Executing for: 'cs421g' [DB2 LUW], Schema: CS421G68	
22:28:51	SUCCESS	SELECT	0.029	0.018	4	Result set fetched	SELECT C.NAME, COUNT(UNIQUE(CRID)) AS NUMBEROFDATA...
22:28:51	FINISHED		0.029	0.018	4	Success: 1	

5.3 Who is worst at experiments?

Rank active researchers (name) by the number of raw data they produced that is associated with a dashboard containing focus > 3 (this needs to be normalized by the total number of experiments they have done):

```
20 SELECT M.NAME, M.NUMEBRDATA, N.NUMEBRDATA, ROUND(CAST(M.NUMEBRDATA AS DOUBLE)/N.NUMEBRDATA,3)
21 FROM
22 (SELECT Y.NAME, COUNT(X.CRID) AS NUMEBRDATA
23 FROM
24 (SELECT UNIQUE(D.CRID) AS CRID FROM CLICDASHBOARD D, CLICRAWDATA R
25 WHERE D.CRID = R.CRID AND D.FOCUS > 3) X, LABMEMBER Y, EXPERIMENT Z
26 WHERE Y.STATUS = 'active' AND Z.CRID = X.CRID AND Y.NAME = Z.NAME
27 GROUP BY Y.NAME
28 ORDER BY NUMEBRDATA)M,
29 (SELECT Y.NAME, COUNT(X.CRID) AS NUMEBRDATA
30 FROM
31 (SELECT UNIQUE(D.CRID) AS CRID FROM CLICDASHBOARD D, CLICRAWDATA R
32 WHERE D.CRID = R.CRID ) X, LABMEMBER Y, EXPERIMENT Z
33 WHERE Y.STATUS = 'active' AND Z.CRID = X.CRID AND Y.NAME = Z.NAME
34 GROUP BY Y.NAME
35 ORDER BY NUMEBRDATA)N
36 WHERE M.NAME=N.NAME
37
```

22:45 [520] INS

Log | 1: CLICDASHBOARD [6]

*	NAME	NUMEBRDATA	NUMEBRDATA	4
1	Kim Metera	18	28	0.643
2	Sabrina Leslie	22	36	0.611
3	Francis Stabile	24	36	0.667
4	Radin Tahvildari	24	34	0.706
5	Daniel Berard	28	35	0.8
6	Cindy Shaheen	30	38	0.789

Time	Status	Command	Exec	Fetch	Rows	Message	SQL/Command
13:23:35	STARTED					Executing for: 'cs421g' [DB2 LUW], Schema: CS421G68	
13:23:35	SUCCESS	SELECT	0.019	0.002	6	Result set fetched	SELECT M.NAME, M.NUMEBRDATA, N.NUMEBRDATA, ROUND(CAST(M.NUMEBRDAT...
13:23:35	FINISHED		0.019	0.002	6	Success: 1	

5.4 Get me all my good data!! Meeting with PI tomorrow!!

List the paths of all DNA unwinding data whose binding analysis results have analysis type "Decent" or "Accurate":

```
23
24
25 SELECT PATHS
26 FROM DNAUNWINDINGDATA D, BINDINGANALYSISRESULTS B
27 WHERE D.CRID=B.CRID AND (B.ANALYSISTYPE IN ('decent','accurate'));
28
29
30
```

34:1 [798] INS

Log | 1: DNAUNWINDINGDATA [140]

*	PATHS
1	C:\BindingAnalysisResults\4
2	C:\BindingAnalysisResults\5
3	C:\BindingAnalysisResults\7
4	C:\BindingAnalysisResults\8
5	C:\BindingAnalysisResults\10
6	C:\BindingAnalysisResults\11
7	C:\BindingAnalysisResults\12
8	C:\BindingAnalysisResults\13
9	C:\BindingAnalysisResults\14
10	C:\BindingAnalysisResults\15
11	C:\BindingAnalysisResults\16
12	C:\BindingAnalysisResults\17
13	C:\BindingAnalysisResults\18
14	C:\BindingAnalysisResults\20
15	C:\BindingAnalysisResults\21
16	C:\BindingAnalysisResults\22
17	C:\BindingAnalysisResults\24
18	C:\BindingAnalysisResults\25
19	C:\BindingAnalysisResults\26
20	C:\BindingAnalysisResults\29
21	C:\BindingAnalysisResults\31
22	C:\BindingAnalysisResults\32
23	C:\BindingAnalysisResults\33
24	C:\BindingAnalysisResults\36

Time	Status	Command	Exec	Fetch	Rows	Message	SQL/Command
22:30:51	STARTED					Executing for: 'cs421g' [DB2 LUW], Schema: CS421G68	
22:30:51	SUCCESS	SELECT	0.024	0.020	154	Result set fetched	SELECT PATHS ...
22:30:51	FINISHED		0.024	0.020	154	Success: 1	

5.5 Getting ready to publish!!

List the paths of all protocols that satisfy:

- they are used in generating nano-particle data with nano-particle analysis results, which produce average radius higher than 40;
- they are used in the experiments for more than 2 times (try number in experiment is bigger than or equal to 2)

```
29
30 SELECT DISTINCT P.PATHS
31 FROM PROTOCOL P, EXPERIMENT E, NANOPARTICLEANALYSISRESULTS N
32 WHERE N.CRID = E.CRID AND N.AVGRADIUS > 40 AND E.PROTOCOLID=P.PROTOCOLID AND E.TRYNUMBER >= 2;
33
34
35
```

27:67 [791] INS

Log | 1: PROTOCOL [50]

*	PATHS
1	C:\Protocol\102.pdf
2	C:\Protocol\104.pdf
3	C:\Protocol\105.pdf
4	C:\Protocol\111.pdf
5	C:\Protocol\119.pdf
6	C:\Protocol\125.pdf
7	C:\Protocol\128.pdf
8	C:\Protocol\130.pdf
9	C:\Protocol\137.pdf
10	C:\Protocol\142.pdf
11	C:\Protocol\143.pdf
12	C:\Protocol\145.pdf
13	C:\Protocol\15.pdf
14	C:\Protocol\152.pdf
15	C:\Protocol\154.pdf
16	C:\Protocol\169.pdf
17	C:\Protocol\17.pdf
18	C:\Protocol\18.pdf
19	C:\Protocol\181.pdf
20	C:\Protocol\188.pdf
21	C:\Protocol\190.pdf
22	C:\Protocol\193.pdf
23	C:\Protocol\196.pdf
24	C:\Protocol\197.pdf
25	C:\Protocol\2.pdf

Format: <Select a Cell>

Time	Status	Command	Exec	Fetch	Rows	Message	SQL/Command
22:31:44	STARTED					Executing for: 'cs421g' [DB2 LUW], Schema: CS421G68	
22:31:45	SUCCESS	SELECT	0.048	0.018	50	Result set fetched	SELECT DISTINCT P.PATHS...
22:31:45	FINISHED		0.048	0.018	50	Success: 1	

5.6 The BIG Picture!

Researchers supervising undergraduates need to keep track of the progress of their assistants. For every active assistant, supervisors would like to see the average metrics regarding data dashboard in the last year. In addition, they also want to know how many datasets assistants have finished analyzing in order to compare the performance.

```
120 SELECT RNAME, RANAME, AVGFOCUS, AVGSNR, AVGCONT, NBCRID
121 FROM
122 (SELECT *
123 FROM SUPERVISE S
124 INNER JOIN
125 (SELECT NAME, AVG(FOCUS) AS AVGFOCUS, AVG(SNR) AS AVGSNR, AVG(CONTAMINANTS) AS AVGCONT, COUNT(DISTINCT CRID) AS NBCRID
126 FROM CLICDASHBOARD
127 WHERE YEAR(DATES)=YEAR(CURRENT DATE) < 2
128 GROUP BY NAME)
129 ON NAME=S.RANAME) A, LABMEMBER L
130 WHERE A.RANAME=L.NAME AND L.STATUS='active'
131 ORDER BY RNAME, RANAME;
132
```

119:1 [2917] INS

Log | 1: SUPERVISE [4]

*	RNAME	RANAME	AVGFOCUS	AVGSNR	AVGCONT	NBCRID
1	Cindy Shaheen	Stone Chen	3.16142857142857 15	5.633214285714286	2272	46
2	Cindy Shaheen	Zhi Zhang	2.5334090909090907	5.752272727272728	2468	35
3	Radin Tahvildari	Wendy Ji	3.0553061224489797	5.5916326530612235	2526	46
4	Radin Tahvildari	Zach Friedenberger	3.108653846153845	5.386923076923078	2506	46

Time	Status	Command	Exec	Fetch	Rows	Message	SQL/Command
23:03:54	STARTED					Executing for: 'cs421g' [DB2 LUW], Schema: CS421G68	
23:03:54	SUCCESS	SELECT	0.028	0.024	4	Result set fetched	SELECT RNAME, RANAME, AVGFOCUS, AVGSNR, AVGCONT, NBCRID...
23:03:54	FINISHED		0.028	0.024	4	Success: 1	

6 Data Modification

6.1 Research assistant has graduated...

Find the research assistants named 'Stone Chen' and 'Wendy Ji' and modify their status to 'inactive'. For those research assistants who just left, update their end dates of supervision to "2019-05-01":

- Table LABMEMBER:

Previous Data:

```
112 SELECT * FROM LABMEMBER;
113
114
115
```

108:29 [2957] INS

Log | 1: LABMEMBER [16]

*	NAME	CONTACTINFORMATION	STATUS
1	Wendy Ji	(514)-561-5587	active
2	Stone Chen	(514)-531-3322	active
3	Zhi Zhang	(514)-557-3232	active
4	Zach Friedenberger	(514)-324-1237	active
5	Gracia Gu	(514)-545-5123	active
6	Zhi Yue Zhang	(514)-512-5567	active
7	Cindy Shaheen	(514)-532-8521	active
8	Francis Stabile	(514)-274-4887	active
9	Daniel Berard	(514)-244-9817	active
10	Kim Metera	(514)-274-4921	active
11	Radin Tahvildari	(514)-284-1237	active
12	Romain Berti	(514)-224-9387	active
13	Sabrina Leslie	(514)-224-2837	active
14	Shane Scott	(514)-274-9287	inactive
15	Haoran Liao	(514)-271-1927	inactive
16	Yash Patel	(514)-292-1927	inactive

Data after modification:

```
38 UPDATE LABMEMBER
39 SET STATUS = CASE
40 WHEN NAME IN ('Stone Chen','Wendy Ji') THEN 'inactive'
41 ELSE STATUS
42 END;
43
44 SELECT * FROM LABMEMBER;
45
```

35:1 [977] INS

Log | 1: LABMEMBER [16]

*	NAME	CONTACTINFORMATION	STATUS
1	Wendy Ji	(514)-561-5587	inactive
2	Stone Chen	(514)-531-3322	inactive
3	Zhi Zhang	(514)-557-3232	active
4	Zach Friedenberger	(514)-324-1237	active
5	Gracia Gu	(514)-545-5123	active
6	Zhi Yue Zhang	(514)-512-5567	active
7	Cindy Shaheen	(514)-532-8521	active
8	Francis Stabile	(514)-274-4887	active
9	Daniel Berard	(514)-244-9817	active
10	Kim Metera	(514)-274-4921	active
11	Radin Tahvildari	(514)-284-1237	active
12	Romain Berti	(514)-224-9387	active
13	Sabrina Leslie	(514)-224-2837	active

- Table SUPERVISE:

Previous Data:

```
61 SELECT * FROM SUPERVISE;
62
63
```

54:19 [1373] INS

Log | 1: SUPERVISE [7]

*	STARTDATE	ENDDATE	RNAME	RANAME
1	2018-05-01	(null)	Radin Tahvildari	Wendy Ji
2	2018-05-01	(null)	Radin Tahvildari	Zach Friedenberger
3	2017-02-01	2018-06-01	Radin Tahvildari	Haoran Liao
4	2017-02-01	2018-06-01	Radin Tahvildari	Yash Patel
5	2018-05-01	(null)	Cindy Shaheen	Stone Chen
6	2018-05-01	(null)	Cindy Shaheen	Zhi Zhang
7	2018-05-01	(null)	Cindy Shaheen	Gracia Gu

Data after modification:

```
46 UPDATE BINDINGANALYSISRESULTS
47 SET ANALYSISTYPE = 'decent'
48 WHERE NAME='Zhi Zhang' AND ANALYSISTYPE = 'rough';
49
50 UPDATE SUPERVISE S
51 SET ENDDATE = '2019-05-01'
52 WHERE EXISTS
53 (SELECT L.NAME
54 FROM LABMEMBER L
55 WHERE L.STATUS = 'inactive' AND L.NAME=S.RANAME AND S.ENDDATE IS NULL);
56
57 SELECT * FROM SUPERVISE;
```

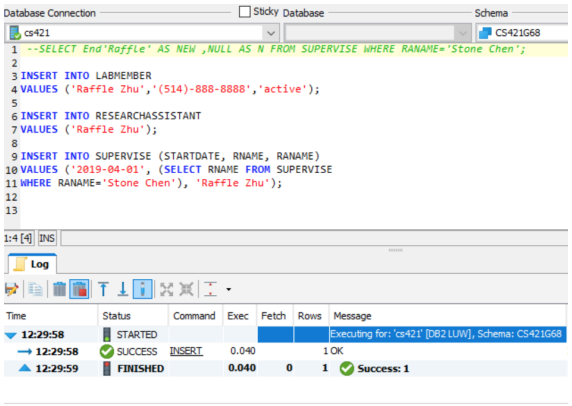
42:5 [1128] INS

Log | 1: SUPERVISE [7]

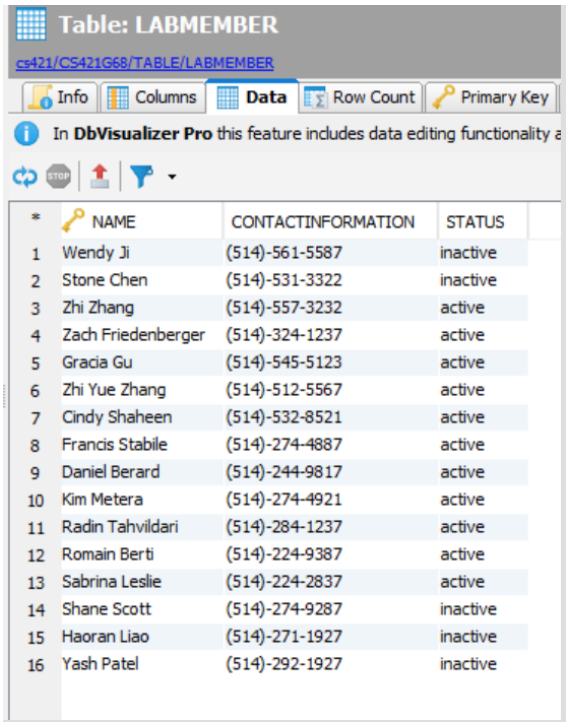
*	STARTDATE	ENDDATE	RNAME	RANAME
1	2018-05-01	2019-05-01	Radin Tahvildari	Wendy Ji
2	2018-05-01	(null)	Radin Tahvildari	Zach Friedenberger
3	2017-02-01	2018-06-01	Radin Tahvildari	Haoran Liao
4	2017-02-01	2018-06-01	Radin Tahvildari	Yash Patel
5	2018-05-01	2019-05-01	Cindy Shaheen	Stone Chen
6	2018-05-01	(null)	Cindy Shaheen	Zhi Zhang
7	2018-05-01	(null)	Cindy Shaheen	Gracia Gu

6.2 Newbie to the lab!!

Fortunately, someone is replacing those who left. Insert into research assistant and lab member "Raffle Zhu" with phone number (514)-888-8888 with 'active' status and supervision start date "2019-04-01" by "Zhi Yue Zhang":

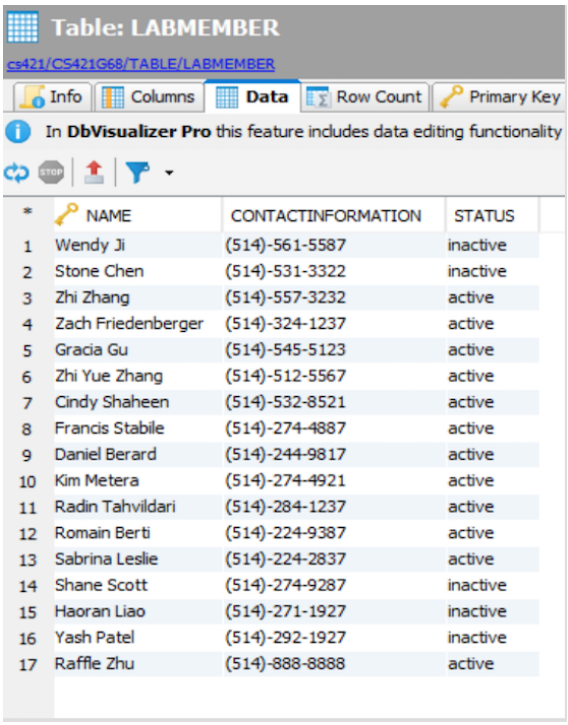


Previous Data:



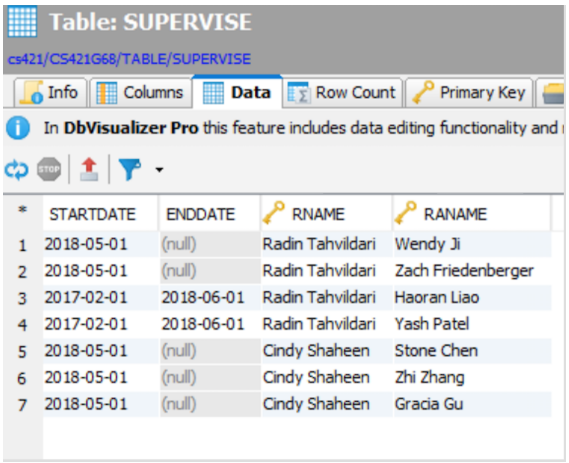
*	NAME	CONTACTINFORMATION	STATUS
1	Wendy Ji	(514)-561-5587	inactive
2	Stone Chen	(514)-531-3322	inactive
3	Zhi Zhang	(514)-557-3232	active
4	Zach Friedenberger	(514)-324-1237	active
5	Gracia Gu	(514)-545-5123	active
6	Zhi Yue Zhang	(514)-512-5567	active
7	Cindy Shaheen	(514)-532-8521	active
8	Francis Stabile	(514)-274-4887	active
9	Daniel Berard	(514)-244-9817	active
10	Kim Metera	(514)-274-4921	active
11	Radin Tahvildari	(514)-284-1237	active
12	Romain Berti	(514)-224-9387	active
13	Sabrina Leslie	(514)-224-2837	active
14	Shane Scott	(514)-274-9287	inactive
15	Haoran Liao	(514)-271-1927	inactive
16	Yash Patel	(514)-292-1927	inactive

Data after modification:



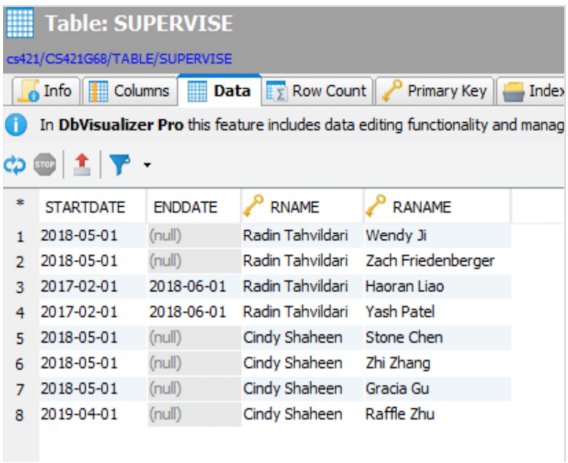
*	NAME	CONTACTINFORMATION	STATUS
1	Wendy Ji	(514)-561-5587	inactive
2	Stone Chen	(514)-531-3322	inactive
3	Zhi Zhang	(514)-557-3232	active
4	Zach Friedenberger	(514)-324-1237	active
5	Gracia Gu	(514)-545-5123	active
6	Zhi Yue Zhang	(514)-512-5567	active
7	Cindy Shaheen	(514)-532-8521	active
8	Francis Stabile	(514)-274-4887	active
9	Daniel Berard	(514)-244-9817	active
10	Kim Metera	(514)-274-4921	active
11	Radin Tahvildari	(514)-284-1237	active
12	Romain Berti	(514)-224-9387	active
13	Sabrina Leslie	(514)-224-2837	active
14	Shane Scott	(514)-274-9287	inactive
15	Haoran Liao	(514)-271-1927	inactive
16	Yash Patel	(514)-292-1927	inactive
17	Raffle Zhu	(514)-888-8888	active

Previous Data:



*	STARTDATE	ENDDATE	RNAME	RANAME
1	2018-05-01	(null)	Radin Tahvildari	Wendy Ji
2	2018-05-01	(null)	Radin Tahvildari	Zach Friedenberger
3	2017-02-01	2018-06-01	Radin Tahvildari	Haoran Liao
4	2017-02-01	2018-06-01	Radin Tahvildari	Yash Patel
5	2018-05-01	(null)	Cindy Shaheen	Stone Chen
6	2018-05-01	(null)	Cindy Shaheen	Zhi Zhang
7	2018-05-01	(null)	Cindy Shaheen	Gracia Gu

Data after modification:



*	STARTDATE	ENDDATE	RNAME	RANAME
1	2018-05-01	(null)	Radin Tahvildari	Wendy Ji
2	2018-05-01	(null)	Radin Tahvildari	Zach Friedenberger
3	2017-02-01	2018-06-01	Radin Tahvildari	Haoran Liao
4	2017-02-01	2018-06-01	Radin Tahvildari	Yash Patel
5	2018-05-01	(null)	Cindy Shaheen	Stone Chen
6	2018-05-01	(null)	Cindy Shaheen	Zhi Zhang
7	2018-05-01	(null)	Cindy Shaheen	Gracia Gu
8	2019-04-01	(null)	Cindy Shaheen	Raffle Zhu

Previous Data:

Table: RESEARCHASSISTANT

cs421/CS421G68/TABLE/RESEARCHASSISTANT

Info | Columns | Data | Row Cou

In **DbVisualizer Pro** this feature includes data

*	NAME
1	Gracia Gu
2	Haoran Liao
3	Stone Chen
4	Wendy Ji
5	Yash Patel
6	Zach Friedenberger
7	Zhi Zhang

Data after modification:

Table: RESEARCHASSISTANT

cs421/CS421G68/TABLE/RESEARCHASSISTANT

Info | Columns | Data | Row Coun

In **DbVisualizer Pro** this feature includes data

*	NAME
1	Gracia Gu
2	Haoran Liao
3	Raffle Zhu
4	Stone Chen
5	Wendy Ji
6	Yash Patel
7	Zach Friedenberger
8	Zhi Zhang

6.3 Correct a mistake!!

Find all DNA binding analysis results done by "Zhi Zhang" in 2019 with analysis type "rough". Correct the types of those analysis to "decent":

```
72 DELETE FROM CLICDASHBOARD
73 WHERE FOCUS < 2 OR (YEAR(CURRENT DATE)-YEAR(DATES)) > 1;
74
75
76 SELECT * FROM CLICDASHBOARD;
```

66:31 [1579] INS

Log

Time	Status	Command	Exec	Fetch	Rows	Message	SQL/Command
12:56:59	STARTED					Executing for: 'cs421g' [DB2 LUW], Schema: CS421G68	
12:56:59	SUCCESS	DELETE	0.012		57	OK	DELETE FROM CLICDASHBOARD...
12:56:59	FINISHED		0.012	0	57	Success: 1	

Previous Data:

```
76 SELECT * FROM CLICDASHBOARD;
```

78:20 [1915] INS

Log | 1: CLICDASHBOARD [201]

*	DID	DATES	PATHS	FOCUS	SNR	CONTAMINANTS	CRID	NAME
189	189	2018-08-28	C:\CLiCDashBoard\189	4.64	4.09	2900	188	Zhi Zhang
190	190	2018-06-07	C:\CLiCDashBoard\190	4.39	2.99	459	67	Stone Chen
191	191	2019-01-29	C:\CLiCDashBoard\191	2.36	4.97	2307	27	Wendy Ji
192	192	2019-02-26	C:\CLiCDashBoard\192	2.84	7.62	2127	189	Stone Chen
193	193	2018-12-20	C:\CLiCDashBoard\193	3.09	5.85	4671	47	Stone Chen
194	194	2018-07-05	C:\CLiCDashBoard\194	1.73	6.64	4948	105	Zhi Zhang
195	195	2018-10-22	C:\CLiCDashBoard\195	3.84	8.71	2359	194	Wendy Ji
196	196	2018-07-14	C:\CLiCDashBoard\196	4.8	5.92	4848	50	Stone Chen
197	197	2018-07-06	C:\CLiCDashBoard\197	4.75	8.12	4942	87	Stone Chen
198	198	2019-01-16	C:\CLiCDashBoard\198	1.07	7.27	2758	170	Zhi Zhang
199	199	2019-01-12	C:\CLiCDashBoard\199	3.87	8.73	3283	114	Zhi Zhang
200	200	2018-07-21	C:\CLiCDashBoard\200	2.56	4.98	3778	159	Zach Friedenberger
201	202	2016-03-03	C:\CLiCDashBoard\202	4.0	3.14	1000	115	Stone Chen

Data after modify:

```
72 DELETE FROM CLICDASHBOARD
73 WHERE FOCUS < 2 OR (YEAR(CURRENT DATE)-YEAR(DATES)) > 1;
74
75
76 SELECT * FROM CLICDASHBOARD;
```

66:43 [1591] INS

Log | 1: CLICDASHBOARD [144]

*	DID	DATES	PATHS	FOCUS	SNR	CONTAMINANTS	CRID	NAME
132	185	2018-05-26	C:\CLiCDashBoard\185	2.39	8.36	2333	125	Wendy Ji
133	187	2019-01-31	C:\CLiCDashBoard\187	2.95	3.4	3044	197	Zach Friedenberger
134	188	2019-02-09	C:\CLiCDashBoard\188	3.41	8.12	4003	153	Wendy Ji
135	189	2018-08-28	C:\CLiCDashBoard\189	4.64	4.09	2900	188	Zhi Zhang
136	190	2018-06-07	C:\CLiCDashBoard\190	4.39	2.99	459	67	Stone Chen
137	191	2019-01-29	C:\CLiCDashBoard\191	2.36	4.97	2307	27	Wendy Ji
138	192	2019-02-26	C:\CLiCDashBoard\192	2.84	7.62	2127	189	Stone Chen
139	193	2018-12-20	C:\CLiCDashBoard\193	3.09	5.85	4671	47	Stone Chen
140	195	2018-10-22	C:\CLiCDashBoard\195	3.84	8.71	2359	194	Wendy Ji
141	196	2018-07-14	C:\CLiCDashBoard\196	4.8	5.92	4848	50	Stone Chen
142	197	2018-07-06	C:\CLiCDashBoard\197	4.75	8.12	4942	87	Stone Chen
143	199	2019-01-12	C:\CLiCDashBoard\199	3.87	8.73	3283	114	Zhi Zhang
144	200	2018-07-21	C:\CLiCDashBoard\200	2.56	4.98	3778	159	Zach Friedenberger

6.4 Clean up some trash data!!

We delete those useless data (satisfy both of the following conditions):

- Data with serious errors (the focus is less than 2);

- Data generated long ago (has been generated for more than 1 years)

Previous Data:

```
70
71
72
73
74 SELECT * FROM CLICDASHBOARD;
75
76
77
78
79
80
81
```

74:29 [1923] INS

Log | 1: CLICDASHBOARD [202]

*	DID	DATES	PATHS	FOCUS	SNR	CONTAMINANTS	CRID	NAME
183	183	2019-01-08	C:\CLiCDashBoard\183	3.09	9.14	1912	81	Wendy Ji
184	184	2019-02-10	C:\CLiCDashBoard\184	1.43	8.61	1225	63	Wendy Ji
185	185	2018-05-26	C:\CLiCDashBoard\185	2.39	8.36	2333	125	Wendy Ji
186	186	2018-06-03	C:\CLiCDashBoard\186	1.71	2.22	1105	63	Zhi Zhang
187	187	2019-01-31	C:\CLiCDashBoard\187	2.95	3.4	3044	197	Zach Friedenberger
188	188	2019-02-09	C:\CLiCDashBoard\188	3.41	8.12	4003	153	Wendy Ji
189	189	2018-08-28	C:\CLiCDashBoard\189	4.64	4.09	2900	188	Zhi Zhang
190	190	2018-06-07	C:\CLiCDashBoard\190	4.39	2.99	459	67	Stone Chen
191	191	2019-01-29	C:\CLiCDashBoard\191	2.36	4.97	2307	27	Wendy Ji
192	192	2019-02-26	C:\CLiCDashBoard\192	2.84	7.62	2127	189	Stone Chen
193	193	2018-12-20	C:\CLiCDashBoard\193	3.09	5.85	4671	47	Stone Chen
194	194	2018-07-05	C:\CLiCDashBoard\194	1.73	6.64	4948	105	Zhi Zhang
195	195	2018-10-22	C:\CLiCDashBoard\195	3.84	8.71	2359	194	Wendy Ji
196	196	2018-07-14	C:\CLiCDashBoard\196	4.8	5.92	4848	50	Stone Chen
197	197	2018-07-06	C:\CLiCDashBoard\197	4.75	8.12	4942	87	Stone Chen
198	198	2019-01-16	C:\CLiCDashBoard\198	1.07	7.27	2758	170	Zhi Zhang
199	199	2019-01-12	C:\CLiCDashBoard\199	3.87	8.73	3283	114	Zhi Zhang
200	200	2018-07-21	C:\CLiCDashBoard\200	2.56	4.98	3778	159	Zach Friedenberger
201	201	2016-03-03	C:\CLiCDashBoard\201	1.0	0.67	5000	115	Stone Chen
202	202	2016-03-03	C:\CLiCDashBoard\202	4.0	3.14	1000	115	Stone Chen

Format: <Select a Cell>

Data after modification:

```
64
65
66
67
68 DELETE FROM CLICDASHBOARD
69 WHERE CONTAMINANTS/FOCUS > 4500 AND (YEAR(CURRENT DATE)-YEAR(DATES)) > 2;
70
71
72 SELECT * FROM CLICDASHBOARD;
73
74
75
```

73:1 [1922] INS

Log | 1: CLICDASHBOARD [201]

*	DID	DATES	PATHS	FOCUS	SNR	CONTAMINANTS	CRID	NAME
182	182	2018-12-16	C:\CLiCDashBoard\182	1.84	3.76	822	141	Zhi Zhang
183	183	2019-01-08	C:\CLiCDashBoard\183	3.09	9.14	1912	81	Wendy Ji
184	184	2019-02-10	C:\CLiCDashBoard\184	1.43	8.61	1225	63	Wendy Ji
185	185	2018-05-26	C:\CLiCDashBoard\185	2.39	8.36	2333	125	Wendy Ji
186	186	2018-06-03	C:\CLiCDashBoard\186	1.71	2.22	1105	63	Zhi Zhang
187	187	2019-01-31	C:\CLiCDashBoard\187	2.95	3.4	3044	197	Zach Friedenberger
188	188	2019-02-09	C:\CLiCDashBoard\188	3.41	8.12	4003	153	Wendy Ji
189	189	2018-08-28	C:\CLiCDashBoard\189	4.64	4.09	2900	188	Zhi Zhang
190	190	2018-06-07	C:\CLiCDashBoard\190	4.39	2.99	459	67	Stone Chen
191	191	2019-01-29	C:\CLiCDashBoard\191	2.36	4.97	2307	27	Wendy Ji
192	192	2019-02-26	C:\CLiCDashBoard\192	2.84	7.62	2127	189	Stone Chen
193	193	2018-12-20	C:\CLiCDashBoard\193	3.09	5.85	4671	47	Stone Chen
194	194	2018-07-05	C:\CLiCDashBoard\194	1.73	6.64	4948	105	Zhi Zhang
195	195	2018-10-22	C:\CLiCDashBoard\195	3.84	8.71	2359	194	Wendy Ji
196	196	2018-07-14	C:\CLiCDashBoard\196	4.8	5.92	4848	50	Stone Chen
197	197	2018-07-06	C:\CLiCDashBoard\197	4.75	8.12	4942	87	Stone Chen
198	198	2019-01-16	C:\CLiCDashBoard\198	1.07	7.27	2758	170	Zhi Zhang
199	199	2019-01-12	C:\CLiCDashBoard\199	3.87	8.73	3283	114	Zhi Zhang
200	200	2018-07-21	C:\CLiCDashBoard\200	2.56	4.98	3778	159	Zach Friedenberger
201	202	2016-03-03	C:\CLiCDashBoard\202	4.0	3.14	1000	115	Stone Chen

Format: <Select a Cell>

7 Views

7.1 You are FIRED!!

Of course this does not happen and this is just a simulation. Rank all active research assistants who started in 2018 (check this with supervision start date) by the number of data they have analyzed (including both types of data):

```
73
74 DROP VIEW SOMENAME;
75
76 CREATE VIEW SOMENAME (NAME) AS
77 SELECT B.NAME
78 FROM BINDINGANALYSISRESULTS B, SUPERVISE S, LABMEMBER L
79 WHERE B.NAME = S.RANAME AND YEAR(S.STARTDATE) >= 2018 AND B.NAME = L.NAME AND L.STATUS = 'active'
80 UNION ALL
81 SELECT N.NAME
82 FROM NANOPARTICLEANALYSISRESULTS N, SUPERVISE S, LABMEMBER L
83 WHERE N.NAME = S.RANAME AND YEAR(S.STARTDATE) >= 2018 AND N.NAME = L.NAME AND L.STATUS = 'active';
84
85 SELECT * FROM SOMENAME;
86 SELECT NAME,COUNT(*) AS NUMBERDA
87 FROM SOMENAME
88 GROUP BY NAME
89 ORDER BY NUMBERDA;
90
91
92
93
94
95
96
97
```

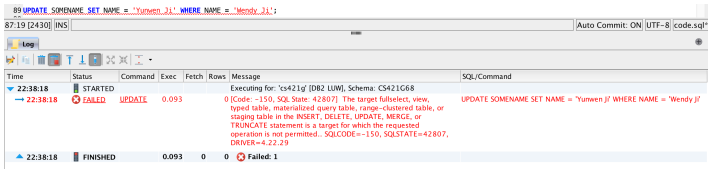
86:33 [2385] INS

Log | 1: SOMENAME [4]

*	NAME	NUMBERDA
1	Wendy Ji	93
2	Zach Friedenberger	93
3	Zhi Zhang	97
4	Stone Chen	117

Time	Status	Command	Exec	Fetch	Rows	Message	SQL/Command
22:33:19	STARTED					Executing for: 'cs421g' [DB2 LUW], Schema: CS421G68	
22:33:20	SUCCESS	DROP	0.026		0	OK. No rows were affected	DROP VIEW SOMENAME
22:33:20	SUCCESS	CREATE	0.026		0	OK. No rows were affected	CREATE VIEW SOMENAME (NAME) AS...
22:33:20	SUCCESS	SELECT	0.044	0.022	400	Result set fetched	SELECT * FROM SOMENAME
22:33:20	SUCCESS	SELECT	0.124	0.020	4	Result set fetched	SELECT NAME,COUNT(*) AS NUMBERDA...
22:33:20	FINISHED		0.220	0.042	404	Success: 4	

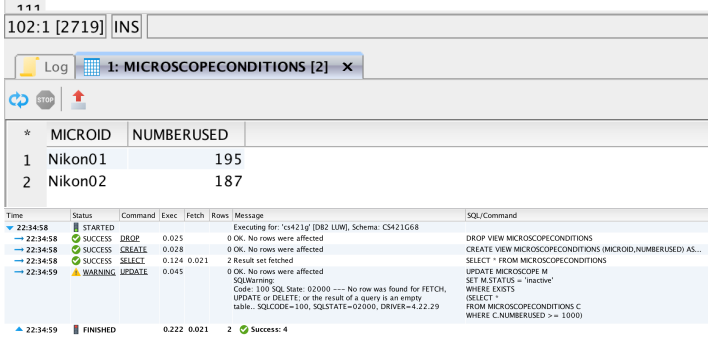
The View SOMENAME is not updatable since the attribute list of the view does not include the primary key, which makes the update unable to locate the entry accurately.



7.2 Time to check the equipment...

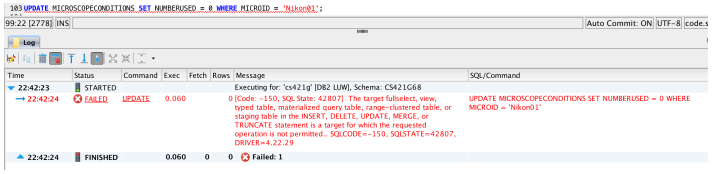
Laboratory equipment needs to be checked and maintained frequently to ensure the smooth running of the experiment and the accuracy of its result, so we will count the number of times a microscope has been used. One will check and maintain this equipment when the number of uses exceeds 1000 and reset the number of uses to 0 after maintenance.

```
DROP VIEW MICROSCOPECONDITIONS;

CREATE VIEW MICROSCOPECONDITIONS (MICROID,NUMBERUSED) AS
SELECT MICROSCOPEID,COUNT(*) AS NUMBERUSED
FROM
(SELECT E.MICROSCOPEID
FROM EXPERIMENT E, MICROSCOPE M
WHERE  E.MICROSCOPEID = M.MICROSCOPEID AND M.STATUS = 'active')T
GROUP BY MICROSCOPEID
;

SELECT * FROM MICROSCOPECONDITIONS;


UPDATE MICROSCOPE M
SET M.STATUS = 'inactive'
WHERE EXISTS
(SELECT *
FROM MICROSCOPECONDITIONS C
WHERE C.NUMBERUSED >= 1000);
```



*	MICROID	NUMBERUSED
1	Nikon01	195
2	Nikon02	187

The View MICROSCOPECONDITIONS is not updatable it contains the GROUP BY clause:



7.3

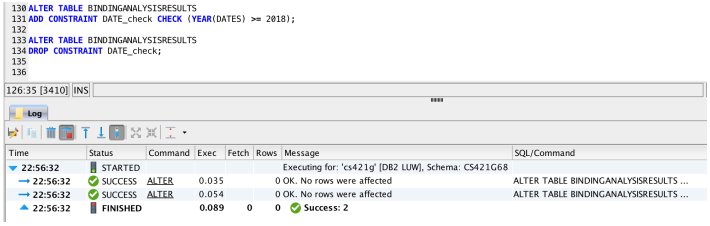
The View is updatable when it satisfies all the following conditions:

- The view is defined based on only one table and includes its PRIMARY KEY upon creation.
- The view should not have any field made out of aggregate and arithmetic functions.
- The view must not have any GROUP BY, HAVING, SUBQUERIES or DISTINCT clause in its definition.
- If the view you want to update is based upon another view, the later should be updatable.
- Any of the selected output fields (of the view) must not use constants, strings or value expressions.
- (DB2) The view is deletable and not view only.
- (DB2) UNION ALL must have matching types for each field.

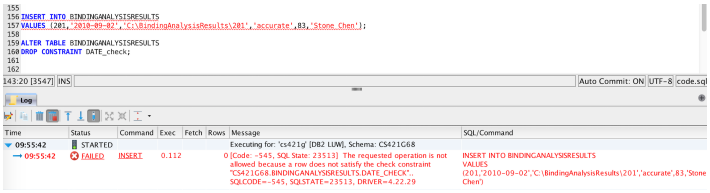
8 CHECK Constraint

8.1 Date check

Add the check constraint to the table BINDINGANALYSISRESULTS, to make sure that all the data stored in the database are after Jan, 2018.

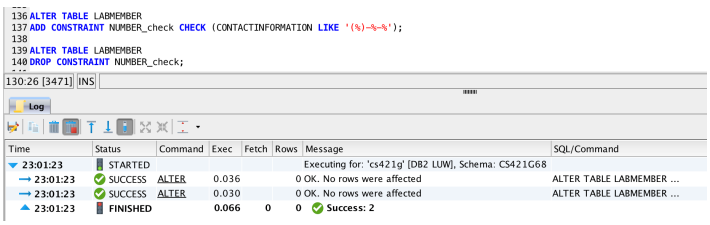


We tried to insert a tuple whose date is earlier than 2018, leading to failure of the sql execution.

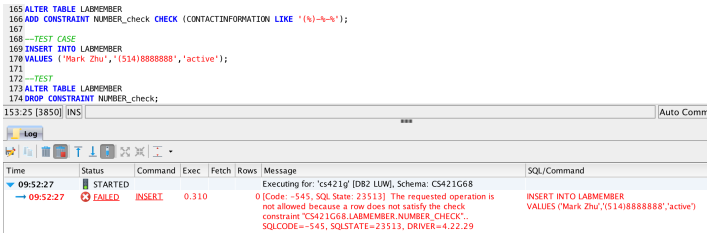


8.2 Number check

Add the check constraint to the contact information (the phone number) in the table LABMEMBER to make sure the phone number is valid.



We tried to insert a tuple which violated our phone number constraint, leading to an failure of the sql command.



9 Creative Points

We have implemented or performed the followings to our database:

- **Automatic Data Generation**: We wrote a python script to automatically construct SQL insert statements with randomly generated data. Please refer to autoDataGen.py and autoDataGen.sql.
- **Some Real Data**: We incorporated some real data in our database, including lab members and their corresponding status, as well as microscope information. In terms of other automatically generated data, they also follow the format of real data.
- **Complex Analytical Query**: All queries in section 5 are real life questions one may ask in the lab and some of them involve certain complexities.

CLiC Microscopic Data Organization Database P3(Group 68)

Stone Chen, Yunwen Ji, Wenquan Li, Yuki Zhang

March 21, 2019

1 Stored Procedure

1.1 Dubious Results

We created a stored procedure to if the average contaminants generated from a data source is above an acceptable threshold. If above, the procedure would change the analysis result type conducted on this data from 'decent' to 'rough' since the data quality was questionable.

```
cs421=> SELECT * from bindinganalysisresults where crid=1;
 barid |   dates    |             paths             | analysistype | crid |        name
-------+------------+-------------------------------+--------------+------+--------------------
   183 | 2018-12-14 | C:\BindingAnalysisResults\183 | decent       |    1 | Zach Friedenberger
   102 | 2018-10-05 | C:\BindingAnalysisResults\102 | decent       |    1 | Wendy Ji
(2 rows)

cs421=> SELECT * FROM CLICDASHBOARD WHERE CRID=1;
 did |   dates    |        paths          | focus | snr  | contaminants | crid |        name
-----+------------+-----------------------+-------+------+--------------+------+--------------------
   9 | 2018-07-24 | C:\CLiCDashBoard\9    |  4.01 |  1.4 |          773 |    1 | Wendy Ji
  31 | 2019-01-14 | C:\CLiCDashBoard\31   |  4.74 | 2.16 |         3849 |    1 | Zach Friedenberger
 147 | 2018-08-18 | C:\CLiCDashBoard\147  |  4.39 | 5.03 |         2231 |    1 | Zach Friedenberger
(3 rows)

cs421=> create or replace function downgrade (target integer, threshold numeric) returns void as $$
cs421$> declare
cs421$>     c refcursor;
cs421$>     c_rec record;
cs421$>     num integer;
cs421$>     average numeric;
cs421$> begin
cs421$>     num := 0;
cs421$>     average := 0;
cs421$>     open c for select * from clicdashboard where crid=target;
cs421$>     loop
cs421$>         fetch c into c_rec;
cs421$>         exit when not found;
cs421$>         if c_rec.contaminants is not null then
cs421$>     num = num + 1;
cs421$>     average = (average*(num-1)+c_rec.contaminants)/num;
cs421$>         end if;
cs421$>     end loop;
cs421$>     close c;
cs421$>     if average > threshold and threshold > 0 then
cs421$>         update bindinganalysisresults set analysistype='rough' where crid=target and analysistype='decent';
cs421$>     end if;
cs421$> end;
cs421$> $$ language plpgsql;
CREATE FUNCTION
cs421=> select downgrade(1,2000);
 downgrade
-----------


(1 row)

cs421=> select * from bindinganalysisresults where crid = 1;
 barid |   dates    |             paths             | analysistype | crid |        name
-------+------------+-------------------------------+--------------+------+--------------------
   183 | 2018-12-14 | C:\BindingAnalysisResults\183 | rough        |    1 | Zach Friedenberger
   102 | 2018-10-05 | C:\BindingAnalysisResults\102 | rough        |    1 | Wendy Ji
(2 rows)
```

As we can see, the created function changed the analysistype of the two results related to the problematic data source with crid=1 because the average contaminant was above 2000.

2 GUI Based Data Retriever

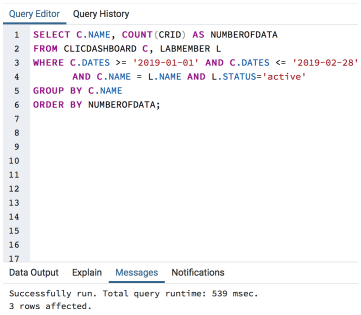
We chose to do a demo on Mar 27, 2019 1:50-2:00 P.M. Our application was designed to allow database user to conveniently retrieve information on the CLiC data and its relevant metrics. In particular, the application is GUI based, consisted of panels with quick actions, custom queries as well as a CLiC data filter. All errors are handled at the back-end and error messages received from the database are displayed on the GUI. For most of the user input, especially in the CLiC data filter, input validity is assessed real time with feedback displayed.

3 Index

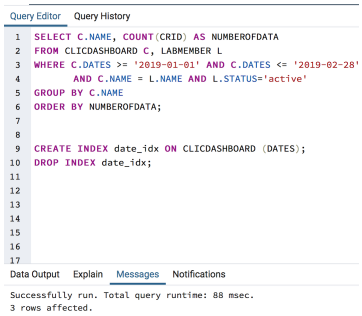
3.1 Chronological is Logical

We created index on DATES of CLICDASHBOARD.

Previous Data:



After:



The results without the indexes were returned in 539 msec, while the results with the indexes were took 88 msec.

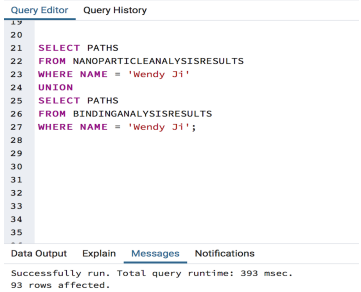
It is useful if we want to find the dates in a certain range and it should shorten the time required for the queries.

This is a considerable difference, because entries can help quickly locate data without having to search every row in a database table every time a database table is accessed, so our query is speeding up.

3.2 We Value Personal Achievements

We create index on NAME of NANOPARTICLEANALYSISRESULTS and BINDINGANALYSISRESULTS.

Previous Data:



After:

```
Query Editor  Query History
18
19
20
21  SELECT PATHS
22  FROM NANOPARTICLEANALYSISRESULTS
23  WHERE NAME = 'Wendy Ji'
24  UNION
25  SELECT PATHS
26  FROM BINDINGANALYSISRESULTS
27  WHERE NAME = 'Wendy Ji';
28
29  CREATE INDEX info_idx ON NANOPARTICLEANALYSISRESULTS (NAME);
30  CREATE INDEX info1_idx ON BINDINGANALYSISRESULTS (NAME);
31  DROP INDEX info_idx;
32  DROP INDEX info1_idx;
33
34
Data Output  Explain  Messages  Notifications
Successfully run. Total query runtime: 159 msec.
93 rows affected.
```

The results without the indexes were 393 msec, while the results with the indexes were took 159 msec.

It is useful because we want to find the paths of a certain name which is similar to locate data, so indexes can also help us to reduce the time in having to search every row in tables. Therefore our query also speed up.

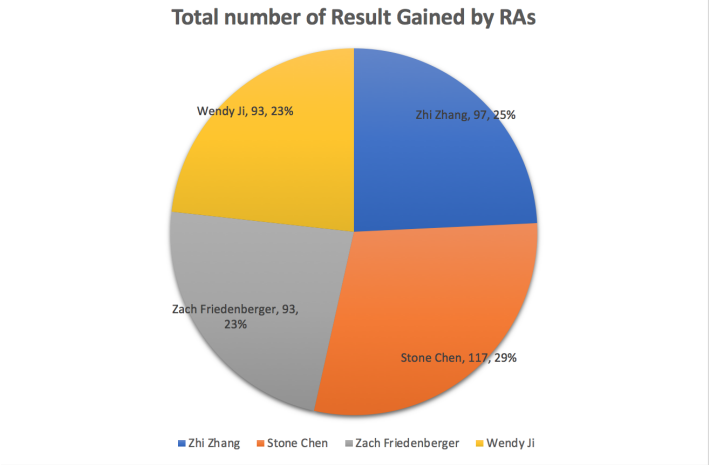
4 Visualization

4.1 Who is the hardest worker?

Rank the total number of results analyzed by the four research assistant who joined the lab at the same time, where results were counted from both nanoparticle analysis and binding analysis:

```
cs421=> \COPY (SELECT NAME, COUNT(*) FROM (SELECT NAME FROM NANOPARTICLEANALYSISRESULTS UNION ALL SELECT NAME FROM BINDINGANALYSISRESULTS)T GROUP BY NAME)TO results.csv WITH CSV
COPY 4
```

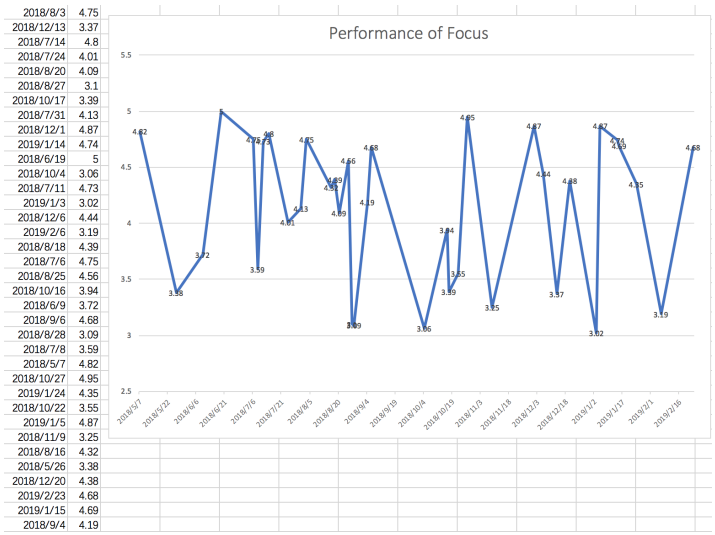
Zhi Zhang	97					
Stone Chen	117					
Zach Friedenberger	93					
Wendy Ji	93					



4.2 Focus Measurement as Experiment Quality Metric

Get the focus measurements of Cindy Shaheen's experiment to find out whether the focus in her experiment is stable or not, indicating the quality of her experiments:

```
cs421=> \COPY (SELECT E.NAME, C.DATES, C.FOCUS FROM EXPERIMENT E, CLICDASHBOARD C WHERE C.CRID = E.CRID AND C.FOCUS > 3 AND
E.NAME = 'Cindy Shaheen' GROUP BY E.NAME, C.FOCUS,C.DATES)TO results.csv WITH CSV
```



2018/8/3	4.75
2018/12/13	3.37
2018/7/14	4.8
2018/7/24	4.01
2018/8/20	4.09
2018/8/27	3.1
2018/10/17	3.39
2018/7/31	4.13
2018/12/1	4.87
2019/1/14	4.74
2018/6/19	5
2018/10/4	3.06
2018/7/11	4.73
2019/1/3	3.02
2018/12/6	4.44
2019/2/6	3.19
2018/8/18	4.39
2018/7/6	4.75
2018/8/25	4.56
2018/10/16	3.94
2018/6/9	3.72
2018/9/6	4.68
2018/8/28	3.09
2018/7/8	3.59
2018/5/7	4.82
2018/10/27	4.95
2019/1/24	4.35
2018/10/22	3.55
2019/1/5	4.87
2018/11/9	3.25
2018/8/16	4.32
2018/5/26	3.38
2018/12/20	4.38
2019/2/23	4.68
2019/1/15	4.69
2018/9/4	4.19

5 Creative Points

5.1 Sophisticated GUI

We created a graphical user interface for our database application. Refer to Section 3.

5.2 No Ghost Member

We created a trigger to update the end date of a research assistant when his status goes from active to inactive. This will help professors when a student leaves because they only need to change his status as a 'labmember' and an automatic end date will be generated on the 'supervise' table.

```
cs421=> select * from labmember where name='Zhi Zhang';
   name     | contactinformation | status
------------+--------------------+--------
 Zhi Zhang | (514)-557-3232     | active
(1 row)

cs421=> select * from supervise where raname='Zhi Zhang';
 startdate  | enddate |     rname      |  raname
------------+---------+----------------+-----------
 2018-05-01 |         | Cindy Shaheen | Zhi Zhang
(1 row)
```

```
cs421=> CREATE OR REPLACE FUNCTION f_leavingMember()
cs421->   RETURNS trigger AS
cs421-> $BODY$
cs421$> BEGIN
cs421$>  IF NEW.status='inactive' AND OLD.status='active' THEN
cs421$>  UPDATE supervise SET enddate=NOW() WHERE raname=OLD.name;
cs421$>  END IF;
cs421$>  RETURN NEW;
cs421$> END;
cs421$> $BODY$ language plpgsql;
CREATE FUNCTION
cs421=>
cs421=> CREATE TRIGGER leavingMember
cs421->   BEFORE UPDATE
cs421->   ON labmember
cs421->   FOR EACH ROW
cs421->   EXECUTE PROCEDURE f_leavingMember();
CREATE TRIGGER
cs421=> update labmember set status='inactive' where name='Zhi Zhang';
UPDATE 1
cs421=> select * from labmember where name='Zhi Zhang';
   name    | contactinformation |  status
-----------+--------------------+----------
 Zhi Zhang | (514)-557-3232     | inactive
(1 row)

cs421=> select * from supervise where raname='Zhi Zhang';
 startdate  |  enddate   |     rname     |  raname
------------+------------+---------------+-----------
 2018-05-01 | 2019-03-22 | Cindy Shaheen | Zhi Zhang
(1 row)
```

As we can see, the created trigger automatically changed the enddate of the research assistant when his status as labmember is changed to inactive from active.