INST 733 SG01, Spring 2016, Dr. Vidat Diker

Natasha Hurwitz (Peer Coach, Nibret Data)

Final Project

Grants Management Database

**Introduction**

The federal government has several agencies that issue research grants. This database design project focuses on federal research grants issued to institutions supporting military nurse researchers.

Grants can have lifecycles that span months, years and even decades. It can be challenging to keep track of all of the important information related to a grant. A funded grant project has start and end dates, a budget, a principal investigator (PI), key personnel, performance sites and a grantee organization. Each grantee organization has grants officers (GOs) with contact information, as well as its own address. Each key personnel, including the PI, has a rank, name, address, phone number and email address as well as a duty station. Each duty station has a location. Each budget has costs that need to be tracked. Each project may have multiple changes to the grant agreement between the grantor (in this case the TriService Nursing Research Program or TSNRP) and the grantee, or recipient of the grant award. These changes can include additional funding, changes in performance site or key personnel, extensions in time as well as scientific changes to the project. All of the information in a project can change throughout the course of the study and it is important to have up-to-date records.

While it is important to keep track of information related to individual grants, it is also useful to aggregate information to meet reporting requirements. Congress oversees the program and in years prior, our office would receive a congressional inquiry requiring all staff to completely focus their attention on collecting the information needed to respond to the query. A database was created so that requests for information could be quickly and efficiently answered, but the database manager is off-site and is paid an hourly rate for any database updates. If Congress makes a request for information that is not part of the views in our current database, staff must try pulling the information by manually dealing with multiple reports or pulling information from individual records. In some cases, information would be pulled from paper files. While this pulling information manually successfully answers the requests, it is time consuming and open to errors. A database, that could be maintained on site by salaried employees, would provide a more cost effective solution.

This project is the beginning of that database that would replace the existing database. If TSNRP could use server space at their home institution, Uniformed Services University of the Health Sciences (USUHS), we could house our own database maintained by our own staff and reduce costs for updates.

**Logistic Design Considerations**

Database Users:

The TSNRP Administrative Staff includes grants managers, a program manager, the Executive Director, and also the executive assistant. These staff members would be the primary users of the database. Grants managers ought to be able to enter data into the system that can be viewed and understood by other staff members. The program manager might want to quickly retrieve final budget numbers for pending awards, while the Executive Director might want to know the total amount of funding that was granted over a certain fiscal year. In the long run, some of the reports (views) from the database could be available on our website to provide other researchers with the opportunity to learn more about our programs.

Logistic Design:

In order to keep track of actions on different grants, the database needs information on the tracking numbers, the amount and timing of the award, the people who are involved in the projects and the percent of their effort they are devoting to the study, the grantee organization who is managing the project and the locations and contact information for everyone and every place involved. The Entity Relationship Diagram, Figure 1, shows the information that is tracked by the database and which tables the information is stored in.

The Project Proposal drove the logical design of the database. The database was designed to include tables that would answer many of the proposed questions. In some cases, which are noted above, it was not possible to include all of the information originally proposed.

Entities not Included:

At this stage, there are certain entities that will not be included, but should be considered later.  For example, when an application is received, the cover packet tells us which research priorities are met by the project.  This is important to the collection of studies we have funded, but not to reporting the life of the specific project.  I will not include the publications and presentations that are the result of the research.  Research subjects, human or animal, will also not be included.  Denied requests for amendments will also not be logged.  Although many of these records would be necessary, this database is focusing on the idea of reporting grant amendments to get a snapshot of one grant.

Figure 1

**Physical Design Considerations**

Physical Design

From the ERD, the database was reverse engineered into a schema titled grants\_management\_db2. After the tables were created, sample data was entered, select queries were created and tested and then saved as views in the database. Views were created to meet the requirements of the project and to display useful information.

Sample Data:

Sample data could not be obtained from the TSNRP. Instead, sample data was created using knowledge of the information that is used by TSNRP portfolio of projects, grantee institutions, locations, funding of projects and grant agreement modifications. In some cases, sample data was created using mockaroo.com. In other cases, the data was manually entered into Excel with headings. The headings were then deleted, the file was saved as a .csv file and it was then imported into the tables in MySQL. In other cases, the data was manually entered directly into MySQL, particularly when small changes needed to be made.

In most cases, the sample data was easily imported and test queries worked well. In other cases, data needed to be updated so that the queries would be more interesting. For example, key personnel were assigned to one of five duty stations instead of all 15 so that the queries would demonstrate that multiple people were at each duty station. Grant Modifications were entered for only 5 projects so that queries could demonstrate multiple modifications in projects.

Create, Read, Update, Delete

Sample Data was importing into the database using the MySQL “Create Table” function and importing .csv files that were either created with Mockaroo.com or in Excel files. In some cases, data entered directly in MySQL after the table was created. Some tables were dropped (deleted) and re-created and in some cases records were dropped and re-entered or updated.

Questions and Queries:

The following questions were initially proposed in the Project Proposal. Most questions have been addressed by the project. In cases where the questions were not addressed, notes have been made, here:

1. What amendments have been approved for this project? (see view a\_grant\_amendments)
2. When was the application funded and when were amendments made? (see view a2\_Dates\_Funded\_Mods)
3. Who approved the amendments? (This question was outside of the scope this project. A new table would need to be created listing internal (TSNRP) staff members and this project focused on only the grantee institutions and people working on the projects, extramurally.)
4. Is the grantee current with progress reports? (The decision was made not to include a progress reports table at this time.)
5. How much funding did the Navy, Army and Air Force each receive in 2015? (see view a3\_funding\_by\_service)
6. Which branch had the most awards in 2015? (see view a4\_funding\_by\_service)
7. Which branch received the greatest amount of money and is it different from the branch that received the most awards? (see view a5\_service\_branch\_totals\_and\_counts)
8. How many funded projects does a certain Principal Investigator have? How many have they had and what is the total funding? (see view a6\_pi\_funding\_totals\_and\_counts)
9. Which PIs and which key personnel are working on projects with similar titles? (Answering this question would have required tables that linked keywords to other similar key words. It is almost a separated project to create such a database. At one time, the NIH was keeping a database of keywords called CRISP, but the functions of CRISP were replaced by a larger project. I would need to gather more information to answer this question.)
10. Have PI\_X and Key Personnel\_Y ever worked together on a TSNRP project?  If so, what project? (see view a7\_pis\_with\_kp)

Additional questions the database answers:

1. Which grantee organizations received awards in each fiscal year? (see view 1, grantee\_organization\_with\_grants)
2. How many key personnel are at the different duty stations? (see view 2, key\_personnel\_at\_duty\_stations)
3. Which key personnel are working on which projects? (see view 3, key\_personnel\_by\_grant\_number)
4. If I grant received additional funding, how much total funding did the grant receive? (see view 4, original\_plus\_additional)
5. What is the PI’s phone number and where is the PI stationed? (see view 5, pi\_phone\_Number\_Duty\_Station)
6. Who is the PI for a given project and what is the project title? (see view 6, project\_with\_pis)
7. Does a given PI have an award with an approaching end date and what is the project number? (see view 7, subquery\_pi\_with\_end\_date)
8. How much total funding did TSNRP obligate in a given fiscal year? (see view 8, total\_funding\_by\_year)

Select queries and views meet requirements

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E |
| a\_view a\_grant\_amendments | X |  |  | X |  |
| a2\_Dates\_Funded\_Mods | X |  |  | X |  |
| a3\_Funding\_By\_Service | X | X | X | X |  |
| a4\_Funding\_By\_Service\_2015 | X | X | X | X |  |
| a5\_Service\_Branch\_Totals\_And\_Counts | X | X | X | X |  |
| a6\_PI\_Funding\_Totals\_and\_Counts | X | X | X | X |  |
| a7\_PIs\_with\_KP | X | X |  | X |  |
| 1. grantee\_organization\_with\_grants | X | X |  | X |  |
| 1. key\_personnel\_at\_duty\_station | X | X | X | X |  |
| 1. key\_personnel\_by\_grant\_number | X | X |  | X |  |
| 1. original\_plus\_additional | X | X | X |  |  |
| 1. pi\_phone\_number\_duty\_station | X | X |  | X |  |
| 1. projects\_with\_pis | X | X |  |  |  |
| 1. subquery\_pi\_with\_end\_date | X |  | X |  | X |
| 1. total\_funding\_by\_year | X | X | X | X |  |

Queries and Descriptions:

1. grantee\_organization\_with\_grants - This query lists the TSNRP Grant Number, Grantee Organization and Fiscal Year in order of Grantee Organization Name. The purpose is to show which grantee organizations are receiving funding give a picture of how many grants they are receiving each fiscal year. The TSNRP funds 30-40 projects each year so this view gives a picture of where funding is going and which project goes with which grantee.
2. key\_personnel\_at\_duty\_station - This query will count the number of key personnel at each duty. This provides a way to see which duty stations have the most personnel working on TSNRP grant projects. It is ordered by duty station name.
3. key\_personnel\_by\_grant\_number - This query lists the TSNRP Grant Number, and the Key Personnel associated with the grant number, ordered by grant number. The query can be used to show who is working on each project. It could easily be updated to show which projects each key personnel is working on, which would also be an interesting view.
4. original\_plus\_additional – This query reports the amount of the original budget for the project in one column and then the total amount that has been funded (original budget + all additional funds that were awarded). This is useful to see how munch total funding a project has received over the course of the grant.
5. pi\_phone\_number\_duty\_station – This query lists the PI's Rank, First and Last Name in one field, then phone number, duty station name and duty station city. It is often useful when contacting a PI to know where they are stationed so you can consider the time zone. Knowing the duty station can also give you an idea of what their work schedule might be like at that location.
6. projects\_with\_pis - This query reports the TSNRP Grant Number, Project Title and PI Last Name. It is a useful list to have so that if you know any piece of this information, you can quickly know the rest. Often, a grants manager thinks of grants in terms of PI or TSNRP Grant Number, but when someone calls, they might be talking about the project title.
7. subquery\_pi\_with\_end\_date - This query reports the TSNRP Grant Number, PI last name, and Award end date. The Executive Director asks for this information, frequently, during meetings.
8. total\_funding\_by\_year – This query lists the total funding obligated by the TSNRP by Fiscal Year. This information is included in our annual report.

**Lessons Learned**

As I was adding data to the database, I was importing the data in the incorrect way. Dr. Diker worked with me by phone to get me back on track so that I could import it properly. I learned that it is much easier to do things properly the first time than to undo mistakes.

I also learned to make sure when creating the sample data that there are several records associated with one other record so that when queries aggregate the data, there results demonstrate that the query is working.

While completing the project, I became focused on practicing writing queries and started to log them and create them as views. I was reviewing the requirements for the project as I was going along and making great progress. Then I began writing this report and realized that I had not answered the questions that were in my original proposal. I learned that I should look back at the original proposal at each stage during the course.

I also learned how to pull data from the same table using different aliases for the joined tables. For example, in query a7\_PIs\_with\_KP, both the PI and the KP linked to the Personnel table. By using two aliases, p1 and p2 for the personnel table, Dr. Diker showed me how to correctly get both pieces of information into the table. After consulting with Dr. Diker for a solution, I also learned how to have the query not duplicate records where the PI, who is also key personnel, would be the same.

I learned that while I needed a primary key for the Grant\_Modifications table, it was also necessary to add another field called mod\_number to provide an order for the modifications and a way to display them all in one table, with different fields for modification 1, 2, 3, etc. I was trying to follow a subquery model from page 271 of the Fehily text (2008), but I was not able to get it working, properly. I developed the query a different way with a different output and answered the question, but I might work on trying to figure this out. I learned that sometimes, it is better to accept what you can do and save other goals for the future.

**Future Improvements and Extensions**

There are many tables that could be added to the database to provide more information for example:

* progress reports
* a list of internal staff members who take actions related to the grants that would like to the modifications of the awards
* links to databases of keywords in project titles that would allow easy searches for projects

Also, the more knowledge and skills would be needed to create reports that would display the originally proposed grant snapshot with links to the data being displayed. Perhaps we will learn more about this in the next course.

**Summary**

I was able to create a database that incorporates many of the aspects of grants management. I was also excited to learn how to use MySQL this semester and to learn the theory behind database creation. I thoroughly enjoyed the course and the guidance of the instructor. It was fun and easy to learn the concepts from the course modules and videos. I feel like I have come a long way since the beginning of this semester and I am looking forward to using this database in the next course.

**References:**

Diker, V. (2016) Personal Communication, University of Maryland, iSchool.

Diker, V. (2016) Classroom Modules, University of Maryland, iSchool.

Fehily, C. (2008). Visual Quickstart Guide, SQL, Third Edition, Berkeley, CA: Peachpit Press.