Jon Chu, John Vu
CSE 140 Homework 9

Optimizing Employee Review and Wage Systems through a Scheduling System
By Jon Chu and John Vu

Summary of Research questions and Results

Overarching goal: Other than creating and viewing a schedule, what else can knowing how often a person works and who tell you?

1. What is the best way to determine who should do a staff review on who at work?
    - We are trying to compute how often a person works with any of his/her coworkers and use that to determine who is best to evaluate who for a staff review. This is important to determine because part of working in a workplace is to be evaluated how you work in that environment.
    - Our results showed that people who work a person enough over a quarter are deemed qualified to do a review on that person.
2. How much is spent on employee salaries per quarter?
    - We are computing and comparing the total amount spent on employee wages per quarter. An employer could use this information to see how much is spent each quarter and see whether they are spending too little or too much. Also, in connection to the reviews, management can see if the reviews an employee receives justifies how much they are currently getting paid.
    - We found out that different amounts were spent on employee salaries over the past three quarters.

Motivation and Background

At where John works, management implemented a staff review process Fall quarter and they are doing the same this quarter. The current staff review process is as follows:

- Each consultant does a review on themselves
- Three Leads does a review for each consultant
- Each consultant meets with the supervisor for a one-one-one discussion. The supervisor aggregates the scores and talks to the consultant about the results.

However, the system has come under some scrutiny because it is not comprehensive enough and the results themselves aren't very indicative of how each consultant is doing. The problem is twofold. First, it isn't very beneficial to review yourself because there is a natural tendency to not critique yourself as hard as others will. This is analogous to reading your own paper and not finding the mistakes that would otherwise be found if someone else were to revise and edit it. Second, the Leads do not work outside along with the consultants very often as they usually work in the office with the supervisor. These two problems results in reviews that do not have much meaning behind them.

In an effort to find a better way to make the staff reviews benefit the consultants, John suggested having consultants review each other. However, we do not want all consultant reviewing each other because not everyone works the same amount of hours. Say, if a person only works with another person one hour a week, the time they spend with each other would not warrant a reliable review. Choosing who would review who would not be random, but instead be chosen in a systematic way; it would be based on how often a person works with another person. By knowing who to suggest to review another person, it address the two problems stated above and adds a new layer of reliability to the results.

Following the previous task, we would have data for both employee schedules and their wage rates. We decided that as an employer, it would be useful to implement both datasets to determine how much is spent on employee wages over a period of time. The information would be conveyed as a graph plotting the overall amount spent on employee wages for each quarter. From this graph, one could determine both spending trends and absolute spending amounts. This information would be useful to factor in with other budgetary concerns when determining how to adjust employee wages for future quarters.

Dataset

Each quarter, consultants sign up for thirty-minute blocks to work, where each shift usually has at least one person working and can have up to two people working for each shift. The format of the schedule view is sectioned off in weeks and show the times, the days of the week, and who works each shift. Since we wanted our data to consist of who works with whom, John had to ask the the web developer to give him an option to export the schedule into a CSV file, so that we can parse and do our analysis on. Access to the scheduler is NetID authenticated, so only people who work at Learning Technologies and Catalyst can access it. The whole dataset consists of 4 files, three of which have all the shifts for the last three quarters (Spring 2012, Fall 2012, and Winter 2013), and the last one being a CSV file with the wage information. Since someone outside the workplace cannot download the files directly, we have uploaded the contents our four files at the following website: http://students.washington.edu/jtvu/final_project/.
You can go to the site, and download the two files (the Schedules folder and the wages.csv) and run our program in the same directory as the two files.

Methodology

The program is given both CSV files containing data on employee work schedules and data on employee wages. The first output should be a recommendation of the best people each employee should review. The second output should be a graph plotting overall spending on employee wages per quarter. To accomplish this, one first needs to parse the CSV file and then group all employees working the same time block together. We then clean the data of text that is non-representative of employee names, e.g. ‘XXX’ or ‘-’.

Using this clean data, we should be able to count how many times an employee shares a time block with another employee. The program would then recommend people who each employee should review based on how many hours they work with them. We established a threshold

where if person X works N amount of hours that is over the threshold, with person Y, person X is qualified to do a review on the person Y, and vice versa. The threshold is based on the overall average number of hours an employee works a week and the overall average number of people an employee works with. The program will output a list of the best people who each employee should review.

To produce the graph, we first add up the number of time blocks taken by an employee over a quarter and then multiply those blocks by their hourly wage. We do this same process for every employee over the quarter, and then repeat this process for other quarters. Finally, we plot the total amount spent on employee wages per quarter for multiple quarters.

Results

Our program demonstrates an efficient method for optimizing the employee evaluation system. By simply providing schedule data as an input one can quickly inform each employee who they should review on the basis of how often they work with each other. Our algorithm recommends only employees who have worked over an established threshold of hours with a given employee. We found that the program recommends on average three other employees for an employee to review. Surprisingly there were also two employees who did not spend enough time with other employees for the program to recommend a strong evaluator. These employees should instead review the employee they have worked the most hours with. The program can significantly improve workplace performance by facilitating the production of more comprehensive feedback on employee strengths and weaknesses.

One can also utilize the program to visualize trends in overall spending on employee wages per quarter. Surprisingly, our results indicate that spending was not as consistent over the last three quarters as expected. These findings suggest that employers may have significant flexibility in adjusting employee wage spending over time. By combining this understanding with the improved employee evaluation system, employers can make better informed decisions about how to adjust employee wages according to performance.

Reproducing your results

1. Go http://students.washington.edu/jtvu/final_project/.
2. Download the two files: Schedules and wages.csv. Do not download the files inside Schedules individually.
3. Store the two files you just downloaded in the same directory as our python program.
4. Open up the command line shell or interpreter, navigate to the directory and run the following command: python master.py Schedules/Winter\ 2013.csv wages.csv > output.txt
    - The command takes three additional arguments and is in the following format:
    - python [python file] [file you want to the review analysis on in the Schedules folder] [Schedules folder] [wages file] > output.txt
5. The output.txt file should be created in the same directory and show the results as stated

above, along with an image file of the graph showing how much was spent on salaries.

Collaboration
No one helped us with our project.

Reflection
Jon- I learned that it is much easier to start with a problem that is interesting to you and then find a dataset to fit the problem. Before I had started this assignment, I wish that we had a bit more specifics on the type and scale of dataset we could use. For future students, I would recommend choosing a problem that is interesting, but straightforward, as it is very easy to get sucked into a problem that goes beyond the scope of the course.

John - I learned that a lot of effort and time goes into thinking of a research project. It isn't as easy as just choosing some topic and then starting to create your algorithm for your analysis, but there are factors such as the type of dataset you want to use and whether that dataset fits in the scope of your project. For future students, I would suggest starting early because a lot of time can be spent on figuring out what you want to do. You will go through different datasets and not all of them will be of interest to you.