

Salary Budget Simulator

Xi He

University of Wisconsin-Madison

1 Introduction

Every year, City of Madison spends its budget on Purchase Services, Benefits, Salaries, and many other services. Among these spending, salary spending takes up a great percentage of government's yearly spending. Therefore, making an accurate prediction for salary spending will help City of Madison to make a better prediction for yearly total budget.

The data used in this project is Monthly Position Track from City of Madison, which provided by Financial Department of Madison. This data-set contains employees' salary information of all Madison government departments from 2017 February to 2019 February. The goal of this project is to build a budget simulator for all agency in Madison Department to generate and predict their future yearly employee salary budget.

2 Data-set

The data-set that used in this project contains 6 columns (variables) and 72408 rows (samples) for all agency in City of Madison.

- Location: the department that individual employee is belong to
- Position Description: job title
- Current Salary: The current yearly salary (in dollars) of individual employee
- Filled FTE Percent: The ratio units are FTE units or equivalent employees working full-time
- YearMonth: The time when a sample is recorded

| Location | Position Description | Current Salary | Filled FTE Percent | YearMonth |
|-------------|----------------------|----------------|--------------------|-----------|
| ATTORNEY | ASST CITY ATTY | 143627 | 1 | 2017/03 |
| ATTORNEY | LITIGATION ASST 1 | 67231 | 1 | 2017/03 |
| ATTORNEY | CLERK-TYP 2 | 48896 | 1 | 2017/03 |
| ATTORNEY | ASST CITY ATTY | 144932 | 1 | 2017/03 |
| ATTORNEY | SECRETARY 3-LEGAL | 59614 | 1 | 2017/03 |
| ... | ... | ... | ... | ... |
| CD A HOUSIN | INFORMATION CLERK | 44338 | 1 | 2019/02 |
| CD A HOUSIN | CUSTODIAL WKR 2 | 45930 | 1 | 2019/02 |
| SEWER | S/D MAINT TECH 2 | 62043 | 1 | 2019/02 |
| PARKING | PKG MAINT WKR 1 | 48722 | 1 | 2019/02 |

Fig. 1. Cleaning Employee Salary data-set

3 Data Analysis

In order to build a simulator, we need to find out the possible factors that will affect the salary of employees and ignore insignificant factors. Several analyses have been conducted. We use Engineer Department as an example to find out the significant variables.

Outline:

- Section 3.1: We explore whether individual employee's job position will affect his/her salary.
- Section 3.2: We explore whether employees' turnover will be affected by several employees' salary level
- Section 3.3: We explore the relationship between the students' salary growth probability and the job position that they have.
- Section 3.4: We explore the relationship between the students' salary growth Rates and the job position that they have.
- Section 3.5: Several calculated rates are shown here.
- Section 3.6: The conclusion of data analysis.

3.1 Average Salary Growth Rates and Job Position

Figure 2 shows the relationship between average current salary and job positions that employees have. The bar with '/' inside represents the job title(ENGR CITY) with the highest average salaries. Employees who have position title ENGR CITY can earn about 160 thousands dollars per year. The bar with '-' represents the job title(Custodial WKR 2) with the lowest average salaries. Employees who have position title ENGR CITY can earn about 40 thousands dollars per year. Therefore, the differences in positions have great impact on employees' salaries. Therefore, the individual employee's job title should be considered in salary budget simulator. Each position's average current yearly salary will be used as beginning salary for new employees.

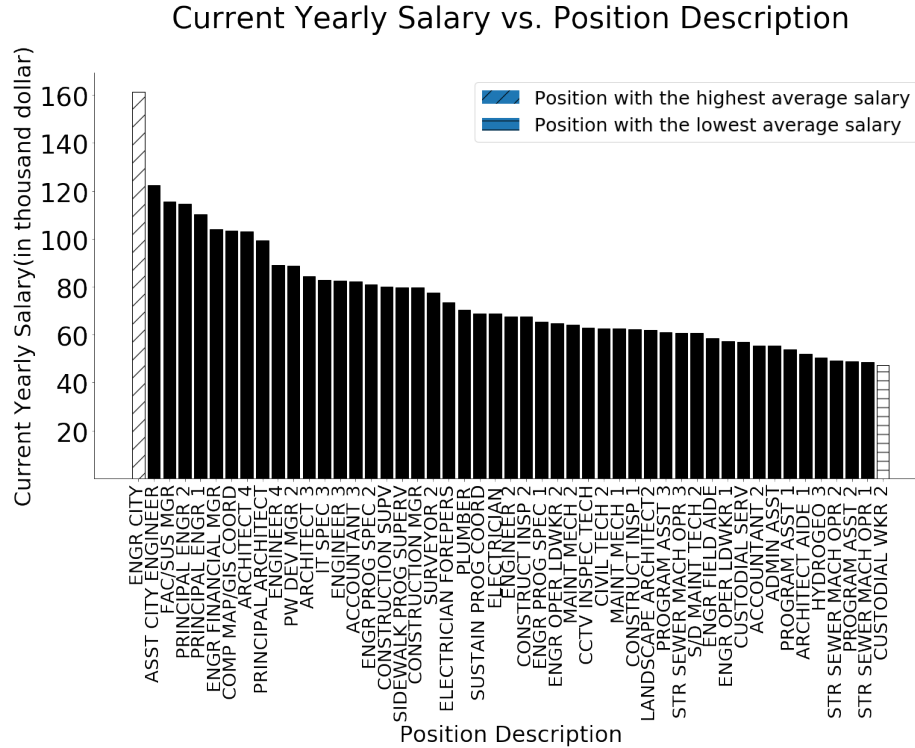


Fig. 2. Average Yearly Salary by Position Description

3.2 Employee Turnover and Salary

This section explore whether employee turnover (number of retire people + number of hire people) would be affected by different salary levels.

Figure 3 shows Madison engineer department's employee turnover through time. In order to avoid the number of employees' influence, this analysis uses 20,40,60,80,100 salary percentile as the separate point. As graph shows, there is no big differences in employee turnover between groups with different salary levels. Moreover, 2017/07, 2017/12, and 2018/12 have high turnover which means that Engineer Department tends to hire more people and people tends to quit during these three months.

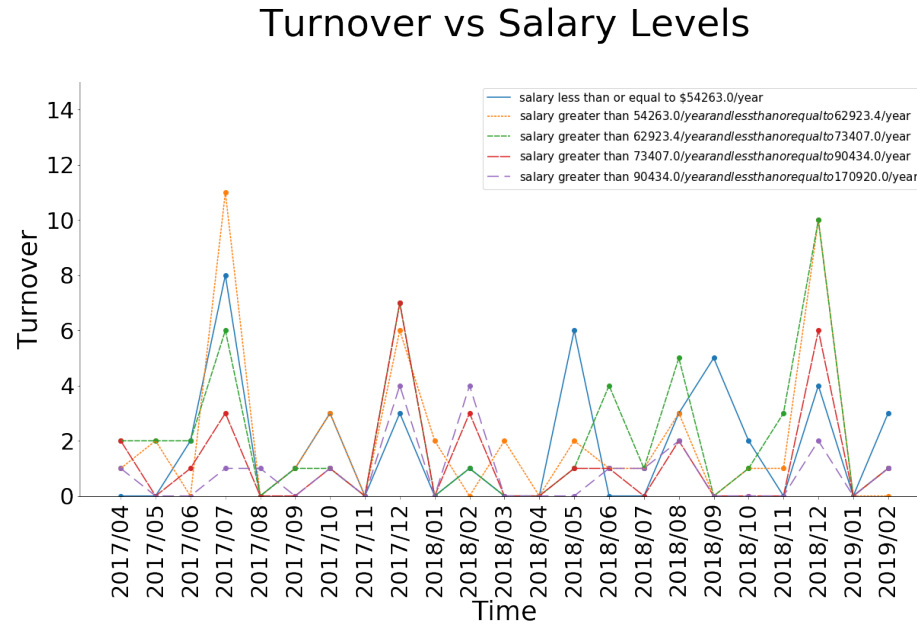


Fig. 3. Employee Flow Chart

3.3 Salary Growth Probability and Position Description

Monthly Position Track from City of Madison shows that employees' wages grow, every once in a while. This section discuss whether different positions' salary growth probability will affect employees' salaries. In this project, each job's average Number of Salary Increased per Year represents salary growth probability. For instance, if number of salary increased per year is 1, it means that this job position's salary is likely to increase 1 time per year.

Figure 4 shows the relationship between number of salary increased per year and position description. The highest number of salary increased per year is 3 and the lowest number of salary increased per year is 0.25. Year by year, this small difference can create a significant difference on different job positions' salaries.

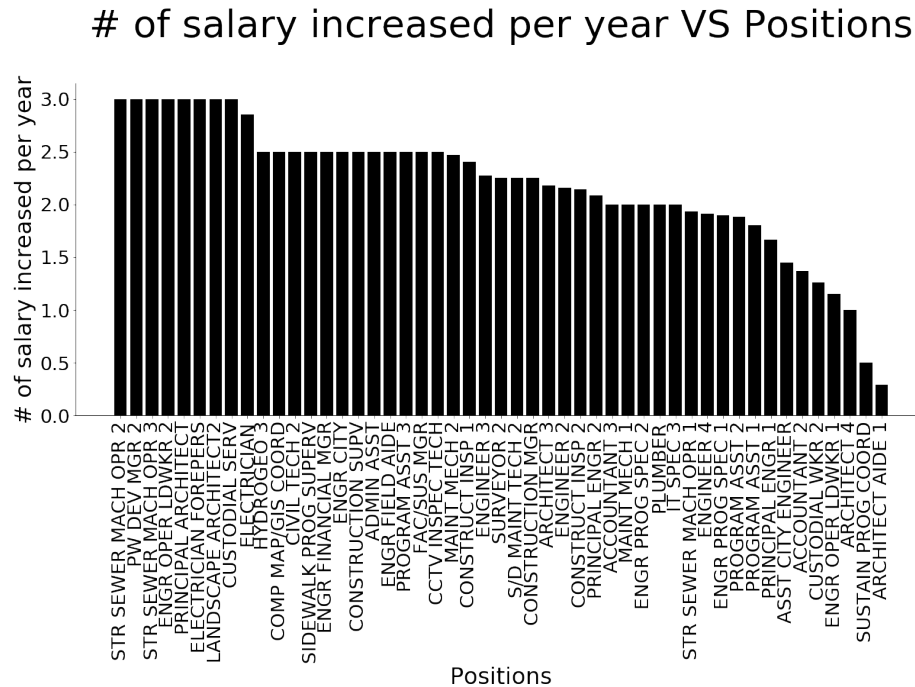


Fig. 4. Number of Salary Increased per Year by Position Description

3.4 Salary Growth Rates and Position Description

This section discuss whether different positions' salary growth probability will affect employees' salary growth rates.

Figure 5 shows the relationship between number of salary growth rates per year and position description. The highest salary growth rates is 0.05 for ENGR OPER LDWKR 1. The lowest number of salary increased per year is negative ENGR OPER LDWKR 2. Containing and calculating salary growth rates based on different job title would improve the accuracy of prediction.

Figure 5 shows that the salary growth rates of ENGR OPER LDWKR 2 is negative, which may cause by limited data. This position only have one employee and the salary growth rates for this employee is negative. Therefore, the average growth rates based on all position is assigned to ENGR OPER LDWR 2.

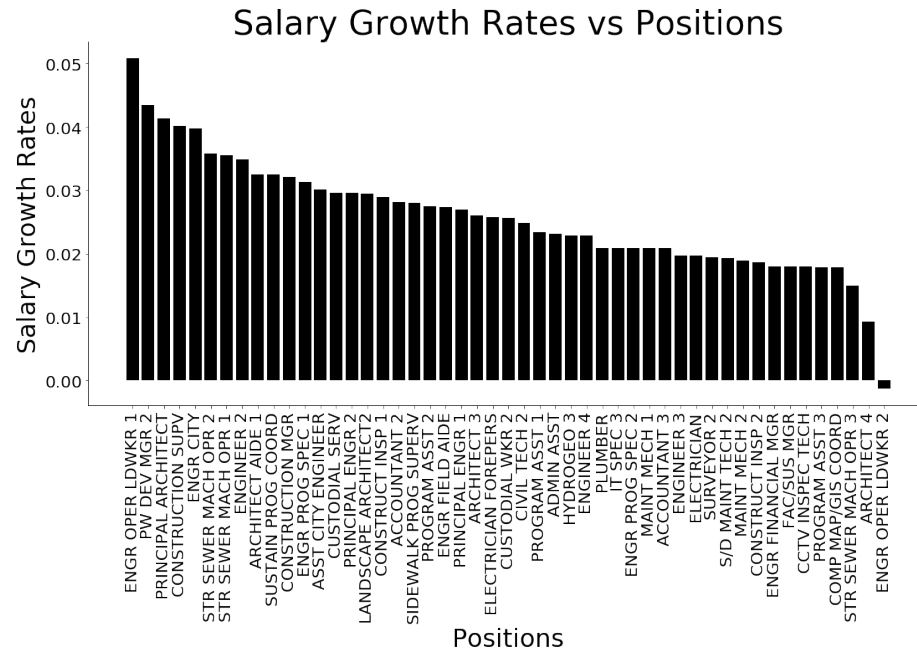


Fig. 5. Salary Group Rates based on different positions

3.5 Hire Rates and Retire Rates

In order to build a simulator, it is necessary to contain hire rates and retire rates in the budget simulator.

Using retire rates and hire rates based on employees' job position in the simulator would improve the accuracy of the prediction. However, because Monthly Salary Tracker only contains 2 years data, it is impossible to calculate retire and hire rates for positions who only have few available data. Therefore, this project uses overall average rates to represent the hire rates and retire rates of positions which have less than 25 available data.

Table 1 shows the overall average hire rates and retire rates for Engineer Department.

Table 1.

| | |
|------------------------------|--------|
| Overall Average Hire Rates | 0.0103 |
| Overall Average Retire Rates | 0.0111 |

3.6 Data Analysis Conclusion

According to analysis analysis, significant variables will be selected and insignificant variables will be ignore.

Budget Simulator Variables:

- Position Description: employees' job titles
- Beginning Salary: each position's average current yearly salary
- Hire Rates: For positions with less than 25 available data, overall average hire rates will be used. For position with more than 25 available data, overall average hire rates will be used.
- Retire Rates: For positions with less than 25 available data, overall average retire rates will be used. For position with more than 25 available data, overall average retire rates will be used.
- Number of Salary Increased per Year: average number of times that different positions' salary will increase per year
- Salary Growth Rates: The growth rates that different positions' salary will increase per time

Insignificant Variables:

- Turnover based on different amount of salary

4 Budget Simulator

4.1 How to Generate a Budget Simulator

Figure 6 shows the fundamental process of running a salary budget simulator. This project uses 2017 as the base year, so data from 2017 are used to generate simulations. For every employees, the simulator will do three steps. First, Python will generate a random number from 0 to 1. If this random number is below the hire rates, the simulator will hire a new employee based on the job position of the current employee in the process, and the starting salary is based on the average salary for this position in the previous year. Second, Python will generate another random variable, if this number is less than the retire rates, the current employee will retire. The hire rates and retire rates will either on depend the position of employees or using the overall average rates. Third, the salary of individual employee for following year will be predicted by using the mathematical formula provided in the figure. Then, the simulator will repeat these three steps for 1000 times.

Upper limits for salaries are created to avoid salaries grow forever. Upper limits are established based on intuition. For positions which earn less than 50000 dollars per year, their upper limit is 100000 dollars. For positions which earn less than 100000 dollars and more than 50000 dollars, their upper limit is 150000 dollars. For positions which earn more than 100000 dollars, their upper limit is 250000 dollars.

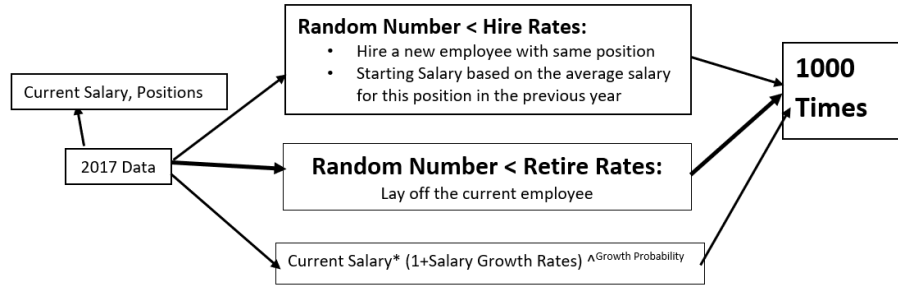


Fig. 6. Budget Simulator

4.2 Results

This Budget Simulator uses 2017 as the base year. Then, the employee budget from 2018 - 2028 is predicted by the simulator. This simulator can make predictions for all department. This paper reports the results of four departments-Engineer, Fiance, Police, and Attorney Department for 2020.

Engineer Department The mission of the City of Madison Engineering Division is to provide a multi faceted combination of Public Works services to the citizens and customers of Madison in a fair and consistent manner that allows for and encourages public input.

Figure 7 shows the salary budget distribution of 1000 simulations for 2020. The black lines in the figure represents the 95 percent confidence interval and mean of 1000 simulations. 95 percent of 1000 simulation is between about 8.63 million dollars and 8.91 million dollars. Based on the average of these 1000 simulation, the salary budget for 2020 will be about 8.77 million dollars.

Figure 8 represents Cumulative Distribution Function of 1000 simulations for 2020. Cumulative Distribution Function is probability that a variable is less than or equal to certain value. Fifty percent of result shows that Engineer Department will spend less than or equal to 8.77 million dollar in paying their employees' salaries.

Salary Simulation for 2020(1000 simulations)

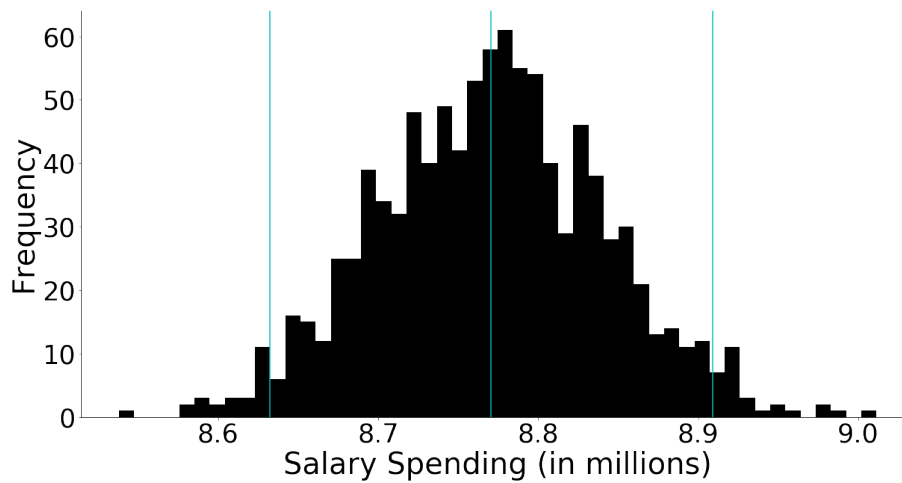


Fig. 7. 1000 Salary Simulations Distribution for 2020

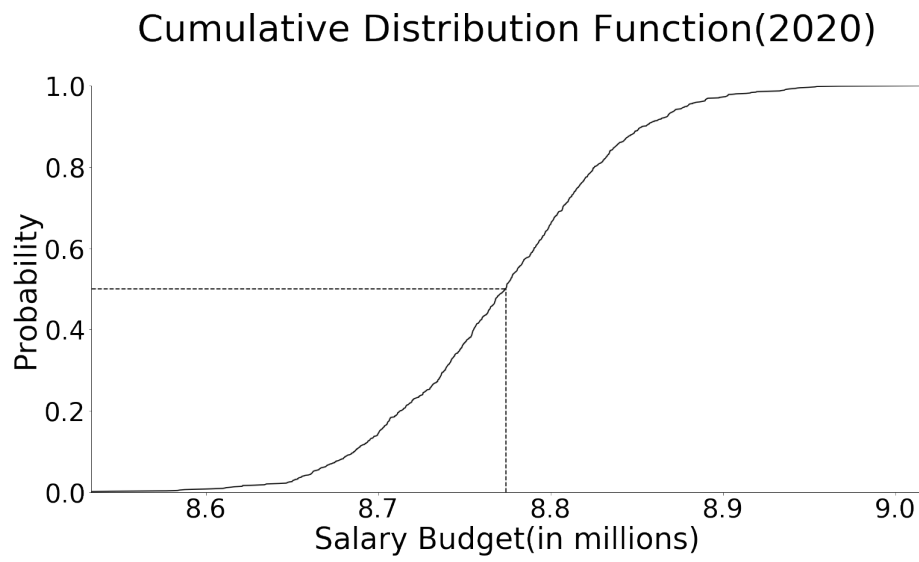


Fig. 8. Cumulative Distribution Function for 2020

Finance Department The mission in the Finance Department is to enhance the financial health of the City and provide stewardship of City resources through financial information, advice and support to the public, employees, City agencies and policymakers.

Figure 9 shows the salary budget distribution of 1000 simulations for 2020. 95 percent of 1000 simulation is between about 3.34 million dollars and 3.5 million dollars. Based on the average of these 1000 simulation, the salary budget for 2020 will be about 3.41 million dollars.

Figure 10 represents Cumulative Distribution Function of 1000 simulations for 2020. Fifty percent of result shows that Finance Department will spend less than or equal to 3.41 million dollar in paying their employees' salaries.

Salary Simulation for 2020(1000 simulations)

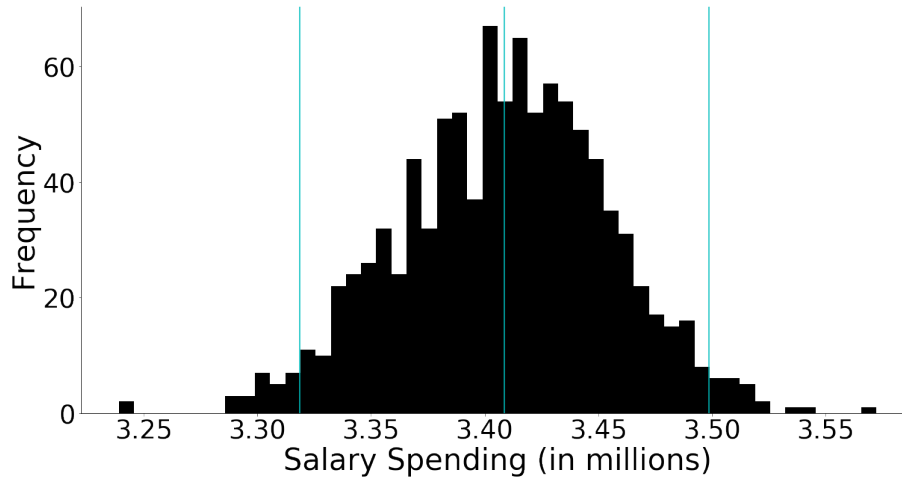


Fig. 9. 1000 Salary Simulations Distribution for 2020

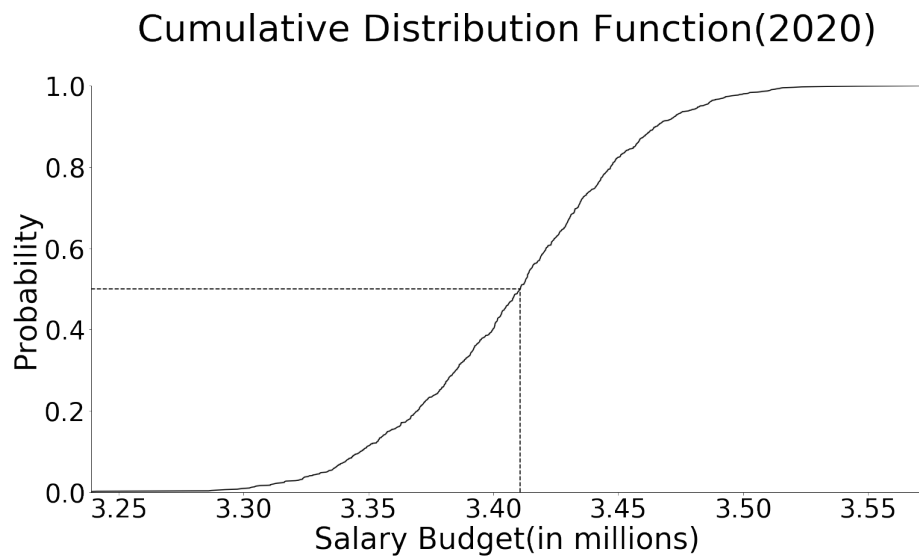


Fig. 10. Cumulative Distribution Function for 2020

Police Department Police Department is committed to providing high quality police services that are accessible to all members of the community.

Figure 11 shows the salary budget distribution of 1000 simulations for 2020. 95 percent of 1000 simulation is between about 49.2 million dollars and 49.7 million dollars. Based on the average of these 1000 simulation, the salary budget for 2020 will be about 49.44 million dollars.

Figure 12 represents Cumulative Distribution Function of 1000 simulations for 2020. Fifty percent of result shows that Police Department will spend less than or equal to 49.44 million dollar in paying their employees' salaries.

Salary Simulation for 2020(1000 simulations)

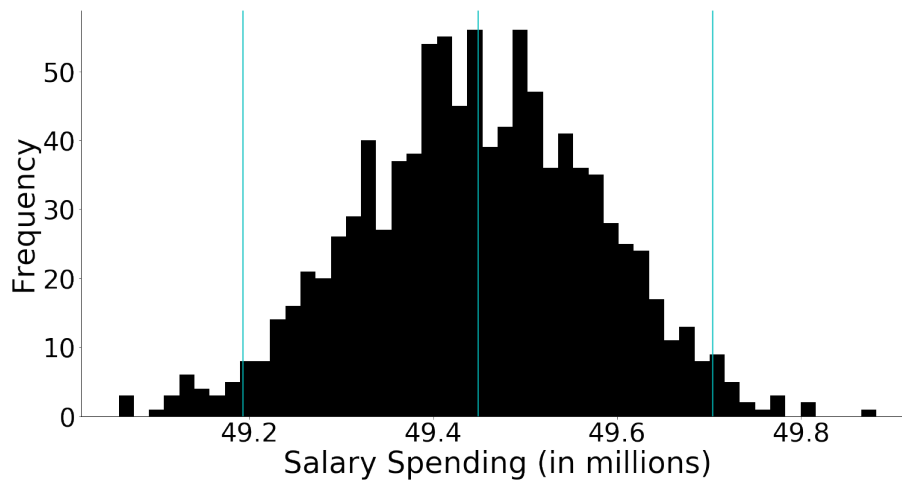


Fig. 11. 1000 Salary Simulations Distribution for 2020

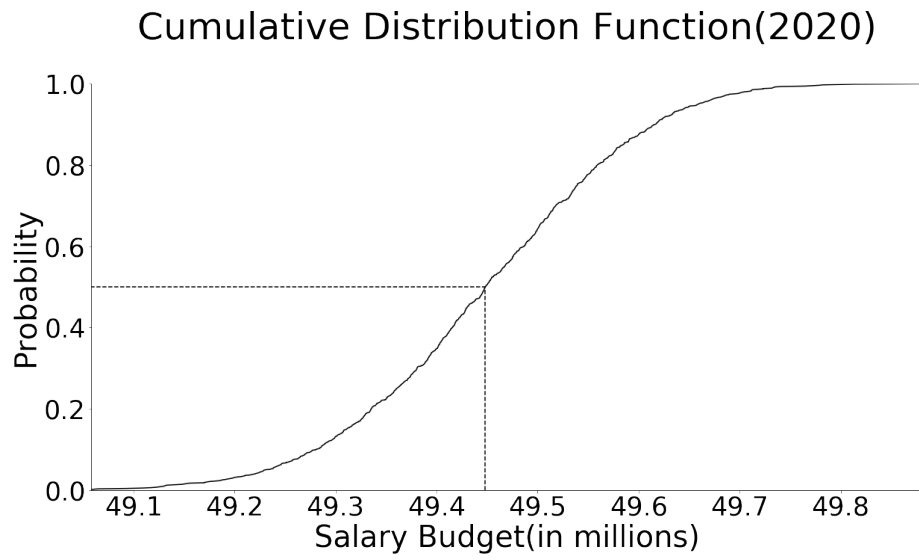


Fig. 12. Cumulative Distribution Function for 2020

Attorney Department The Office of the City Attorney will provide professional legal services and representation of the highest quality to the City of Madison government.

Figure 13 shows the salary budget distribution of 1000 simulations for 2020. 95 percent of 1000 simulation is between about 2.81 million dollars and 2.92 million dollars. Based on the average of these 1000 simulation, the salary budget for 2020 will be about 2.86 million dollars.

Figure 14 represents Cumulative Distribution Function of 1000 simulations for 2020. Fifty percent of result shows that Attorney Department will spend less than or equal to 2.86 million dollar in paying their employees' salaries.

Salary Simulation for 2020(1000 simulations)

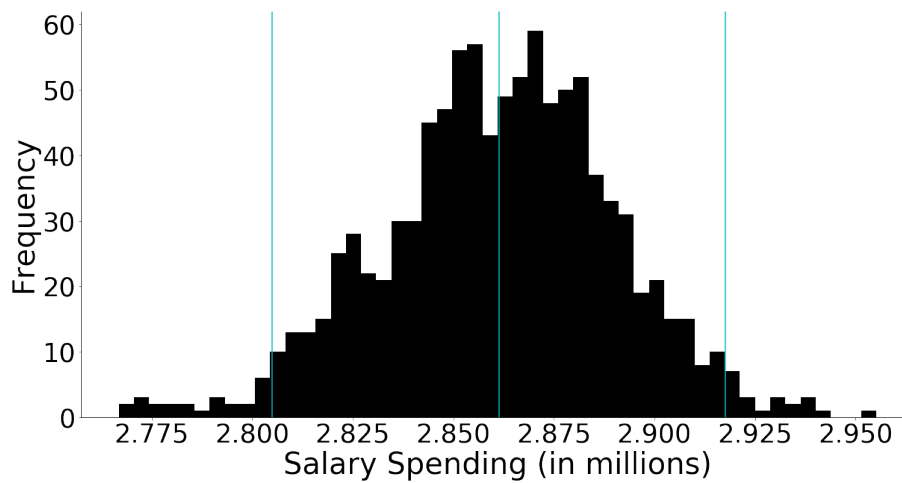


Fig. 13. 1000 Salary Simulations Distribution for 2020

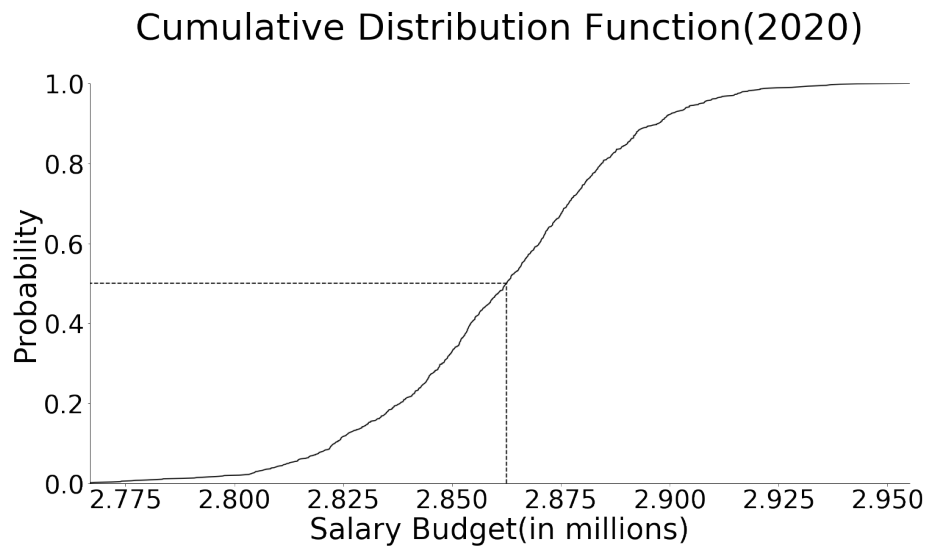


Fig. 14. Cumulative Distribution Function for 2020

5 Conclusion and Suggestions

Position Description, Hire Rates, Retire Rates, Number of Salary Increased per Year, and Salary Growth Rates are important variables for the Budget Simulator. From the result of 1000 simulations, the salary budget simulator predicts that the salary budget will be about 8.77 million dollars for Engineer Department, about 3.41 million dollars for Finance Department, about 49.44 million dollars for Police Department, and about 2.86 million dollars for Attorney Department in 2020. The budget for those department would continue to increase in next several years.

6 Limitations and Future Directions

6.1 Limitations

Monthly Position Track only contains 24 months data from 2017 March to 2019 February. This short period may not accurately make a prediction for a long period. Moreover, because the limitation of data, many mathematical algorithms cannot be calculated. For example, the hire rates and retire rates based on different job positions cannot be calculated because some positions may not have employee turnover.

Some factors related to employees' information may also affect the result of the budget simulation. The information about individual employee, such as working experience and age, is not given because of the employees' privacy. Therefore, this salary budget simulator may not be perfect.

The main analytical method of this project is using plots to tell information. This project does not contain many statistical analytical approaches. Adding some statistical methods may enhance the accuracy of this salary budget simulator.

6.2 Future Directions

- To analyze whether Filled FTE Percent is related to employees' salary