

Who Earns More – *Comparing City Employee Payroll Data Between New York City and Chicago*

Jiyuan Lu
New York University
New York, USA
jl11046@nyu.edu

Abstract—

By comparing city employee payroll data between NYC and Chicago, we can gain insights as to which city has higher salary overall, what kinds of city employee jobs tend to be better paid, and how each city's budget is being devoted overtime. By comparing NYC payroll data over the years, we can observe trends whether a city department is hiring more employees, and whether the city employees are getting better paid. Finally, we give recommendations for perspective city employees on which city is more suitable for them to pursue their careers.

Keywords—Salary, City Employee, NYC, Chicago

I. INTRODUCTION

Current City of Chicago Employee Names, Salaries, and Position Titles is a list of all current City of Chicago employees with full names, departments, positions, and annual salaries for the most recent year. *Citywide Payroll Data* contains payroll data for New York City government employees range from 2014 to the most recent year. *NYC Jobs* contains current job postings available on the City of New York's official jobs site.

By comparing the first two datasets, we can gain insights such as which city has higher salary for city employees, what kinds of jobs tend to be better paid, and how each city's budget is being devoted overtime. We also dived deep into the police department to investigate how many employees they hire, how the budget is spent on, and what is the constitution of the police department in each city. We then give recommendations for perspective city employees on which city is more suitable, in terms of salary, for them to pursue their careers.

By analyzing the last two datasets, we can observe the trends in salary and city job openings in New York City. This helps job seekers in New York City understand the trend in the job market so that they can better prepare themselves and make decisions accordingly.

II. MOTIVATION

The purpose of this analytic is to help both the city government and the job seekers better understand the current salary levels and the salary and job openings trends in the city employee job market. City government and agencies will have a better idea on how the city's financial resources are allocated

and how much of the city's budget is being devoted to, compared to other cities, each city department. Therefore, a more balanced budget distribution could be made for the next year. Job seekers will get to know which city tends to have more openings and better salary for their dream jobs. They can also observe the current salary and job opening trend for each city job and make better decisions.

III. RELATED WORK

#1 Mayors and money: Fiscal policy in New York and Chicago

The politics of large cities in the United States is most often explained in terms of powerful individual personalities. New York and Chicago are but two of a larger group of cities whose primary political relationships developed during the Depression and were reinforced during subsequent years. Like an extraordinary number of America's major cities, New York and Chicago experienced fiscal crises in the 1930s. After the short-term causes of their crises were resolved, however, the fiscal policy-making structures that emerged in the two cities differed substantially. They offer a means for identifying the political interactions that are crucial for understanding why some cities experienced fiscal problems during the 1970s while others remained fiscally stable. And these political interactions have continued to affect the fiscal conditions of American cities through the 1990s.

This paper compares the fiscal structure between major American cities, using New York and Chicago as examples. This gives insight as to explain how the budget is distributed among city departments each year and for what reason.

#2 Wage Determination in the Public Sector

There is a growing body of evidence that government employment is attractive in terms of both wages and job security. A recent U.S. Bureau of Labor Statistics survey found that clerical, data processing, and manual workers employed by municipalities in eleven large urban areas were substantially better paid than their counterparts in private industry. In most cases, federal employees in the same cities were also paid more than comparable private sector workers. Fringe benefits in the public sector are also as good or better than those in the private sector, according to a national survey of U.S. municipalities. Furthermore, job hiring and tenure practices provide considerable security to public workers: in 1971, 57 percent of

nonfarm private employees worked a full year, whereas in the public sector, the proportion was 77 percent.³ Attractive wages and salaries, steady demand for public services, and tenure practices all combine to produce low rates of employee turnover-19 percent in state and local government and 22 percent in the federal service in 1970, compared to 58 percent in private manufacturing.

Because these rather surprising findings conflict with popular notions about government pay, it is appropriate to examine the process of wage determination in the public sector and the outcomes of this process for different occupational groups in government employment. Governmental wages and salaries affect the respective government budget and, therefore, the citizens' tax burden; they influence the relative attractiveness of employment in the public and private sectors; and they are an important factor in the continuing debate over the size and role of government in American society.

This paper analyzes the process of wage determination in the public sector and the outcomes of this process for different occupational groups in government employment. This helps explain the decision-making process of distributing government budget and the discrepancies between New York and Chicago.

IV. DATASETS

A. Current City of Chicago Employee Names, Salaries, and Position Titles

Description:

List of all current City of Chicago employees with full names, departments, positions, annual salaries, and approximate annual salary with furlough days/salary reductions.

Schema before cleaning:

Field Name	Field Type	Description
Name	Text	Name of employee
Job Titles	Text	Title of employee
Department	Text	Department where employee worked
Full or Part-Time	Text	Whether the employee was employed full- (F) or part-time (P)
Salary or Hourly	Text	Whether the employee was paid on an hourly basis or salary basis
Typical Hours	Number	The typical amount of work for hourly employees.
Annual Salary	Number	Annual salary rates
Hourly Rate	Number	Hourly salary rates

Schema after cleaning:

Field Name	Field Type	Description
Department	String	Department the employee works for
Title	String	Title of the employee
Salary	Integer	Annual salary of the employee

Where the data was found:

Chicago Data Portal.

How the data was collected:

Directly download on Dumbo using curl.

Frequency of collection:

Static. Download once.

Size of the data:

2.19 MB.

B. New York City Citywide Payroll Data (Fiscal Year)

Description:

Data is collected because of public interest in how the City's budget is being spent on salary and overtime pay for all municipal employees. Data is input into the City's Personnel Management System ("PMS") by the respective user Agencies. Each record represents the following statistics for every city employee: Agency, Last Name, First Name, Middle Initial, Agency Start Date, Work Location Borough, Job Title Description, Leave Status as of the close of the FY (June 30th), Base Salary, Pay Basis, Regular Hours Paid, Regular Gross Paid, Overtime Hours worked, Total Overtime Paid, and Total Other Compensation (i.e. lump sum and/or retro payments).

Schema before cleaning:

Field Name	Field Type	Description
Fiscal Year	Number	
Payroll Number	Number	
Agency Name	Text	The payroll agency that the employee works for
Last Name	Text	Last name of employee

First Name	Text	First name of employee
Mid Init	Text	Middle initial of employee
Agency Start Date	Date & Time	Date which employee began working for their current agency
Work Location Borough	Text	Borough of employee's primary work location
Title Description	Text	Civil service title description of the employee
Leave Status	Text	Status of employee as of the close of the relevant fiscal year: Active, Ceased, or On Leave
Base Salary	Number	
Pay Basis	Text	List whether the employee is paid on an hourly, per diem or annual basis
Regular Hours	Number	Number of regular hours employee worked in the fiscal year
Regular Gross Paid	Number	The amount paid to the employee for base salary during the fiscal year
OT Hours	Number	Overtime hours worked by employee in the fiscal year
Total OT Paid	Number	Total overtime pay paid to the employee in the fiscal year
Total Other Pay	Number	Includes any compensation in addition to gross salary and overtime pay

Schema after cleaning:

Field Name	Field Type	Description
Department	String	Department the employee works for

Title	String	Title of the employee
Salary	Integer	Annual salary of the employee

Where the data was found:

NYC OpenData.

How the data was collected:

Directly download on Dumbo using curl.

Frequency of collection:

Static. Download once.

Size of the data:

444 MB.

C. NYC Jobs

Description:

This dataset contains current job postings available on the City of New York's official jobs site. Internal postings available to city employees and external postings available to the general public are included.

Schema before cleaning:

Field Name	Field Type	Description
Job ID	Text	
Agency	Text	Name of the City agency where the vacancy exists
Posting Type	Text	Internal or External job posting
# Of Positions	Text	Number of vacancies to be filled
Business Title	Text	
Civil Service Title	Text	
Title Classification	Text	
Title Code No	Text	Civil Service title Code
Level	Text	Civil Service title level
Job Category	Text	
Full-Time/Part-Time indicator	Text	

Career Level	Text	
Salary Range From	Text	
Salary Range To	Text	
Salary Frequency	Text	
Work Location	Text	Agency Location
Division/Work Unit	Text	Department/Division within the hiring agency
Job Description	Text	
Minimum Qual Requirements	Text	Minimum qualifications required for position
Preferred Skills	Text	
Additional Information	Text	Additional information provided by the hiring agency
To Apply	Text	Instructions on how to apply for this position
Hours/Shift	Text	Working hours and shift information
Work Location 1	Text	Specific work location for this opening
Recruitment Contact	Text	Recruitment contact information
Residency Requirement	Text	Residency requirements for this position
Posting Date	Date & Time	
Post Until	Text	Date Posting will be removed
Posting Updated	Date & Time	Last Modification Date
Process Date	Date & Time	Dataset created date

Schema after cleaning:

Field Name	Field Type	Description
Year	Date	Posting year of the job opening
Department	String	Department the employee works for
Title	String	Title of the employee
Salary	Integer	Annual salary of the employee

Where the data was found:

NYC OpenData.

How the data was collected:

Directly download on Dumbo using curl.

Frequency of collection:

Static. Download once.

Size of the data:

13.1 MB.

V. DESCRIPTION OF ANALYTIC

Spark and Spark SQL are the main tools used in the backend of the application. Hive tables are also used to enable the access to the dataset on DUMBO HDFS for Tableau.

First, Datasets are cleaned, profiled and combined using core Spark. Next, the transformed datasets are analyzed using Spark SQL. Finally, the analytical results are stored into Hive tables and visualized using Tableau.

Here are the analytic results and their interpretations:

1. In 2019, the budget in NYC for full time city employees, full time police department employees, and full-time police officers is 3.76, 3.06, and 1.85 times the budget for those of Chicago.

The first two ratios are in correspondence with the fact that NYC (8.4M) has approximately 3.1 times population than Chicago (2.7M). This means approximately the same amount of budget is spent on city employees and police staffs per capita in both cities. However, the last ratio shows that NYC spends less budget on police officers per capita (only 60% of that in Chicago). This shows the demand for police officers in NYC per capita is much less than that in Chicago, which in turn may indicates that NYC is safer and has lower crime rates than Chicago. In 2016, the violent crimes rates is 903.8 in Chicago, but only 585.8 in NYC.; the property crimes rates is 2946.2 in Chicago, but only 1518.7 in NYC.

2. In 2019, among 22705 full time city employees in Chicago, 60.9% belong to the police department, among which 82.7% are police officers; among 104525 full time city employees in NYC, 48.5% belong to the police department, among which 48.7% are police officers.

Chicago is putting most of its financial resources in its police department, but this proportion is less than a half in New York. Within the police department, Chicago is primarily hiring police officers, while this proportion is also below a half in New York. This shows that the primary responsibility of the police department in Chicago may be to maintain order and control crimes, while the responsibility of the police department in NYC is more varied, e.g., park rangers.

3. In 2019, for Chicago, 60.2% of the city budget is spent on the police department, of which 79.2% is spent on police officers; for NYC, 48.8% of the city budget is spent on the police department, of which 48.0% is spent on police officers.

In both cities, police and fire department employees make up for majority of the city employees. And most of the city budget is spent on these two departments. This shows that the major functionality for both cities is to maintain social order.

4. In 2019: The average salary for Chicago full time city employees is around 89.9K, with a standard deviation of about 20.7K, while the average salary for NYC full time city employees is around 73.4K, with a standard deviation of about 29.4K. The average salary for Chicago police staffs is around 88.8K, with a standard deviation of around 18.0K, while the average salary for NYC police staffs is around 74.1K, with a standard deviation of around 29.0K. The average salary for Chicago police officers is around 85.1K, with a standard deviation of about 11.8K, while the average salary for NYC police officers is around 73.1K, with a standard deviation of about 19.1K.

Overall, the average salary for full-time city employees in Chicago is higher than those in NYC, and the salary is also less variant in Chicago. The same conclusion can also be made for the police department and for police officers. There is a tradeoff between salary and job opportunities for job seekers: Working in Chicago gets you better paid, but NYC offers more job opportunities. Also, higher salary for Chicago police officers may indicate that they potentially suffer from more risks than their counterparts in NYC due to higher crime rates.

5. City departments that provide better salary in Chicago than NYC: Budget & Management, Buildings, Business Affairs, City Clerk, Cultural Affairs, Fire, Health, Housing, Human Relations, Law, Police, Sanitation, Transportation. City departments that provide better salary in NYC than Chicago: Board of Election, City Council, Emergency, Finance, Mayor's Office, Treasure, and Water Management.

6. The highest paid city department in Chicago is the Buildings department. The highest paid city department in NYC is the Treasure department. The lowest paid city department is the Board of Election department in both cities.

7. The highest paid city job in Chicago is the superintendent of police, at an annual rate of around 260K, and the lowest paid city job in Chicago is the aldermanic aide, at an annual rate of around 14K. The highest in NYC is the chief actuary, at an annual rate of around 302K, and the lowest in NYC is the temporary clerk, at an annual rate of around 28K.

8. For NYC: The budget, average salary, and number of city jobs all increased each year. 29.2% in total budget, 16.3% in average salary, and 11.1% in the number of city jobs. However, the average salary of new city jobs does not vary much across the years.

Although there are more job openings and better salaries for full-time city jobs each year in NYC, the pay for new city employees stays approximately the same over the years. For job seekers, don't get frustrated by the low starting salary rate. You might get a pay rise soon!

9. The population of New York City is approximately three times that of Chicago. This is reflected both in the budget spent on city employees and police staffs. However, NYC spends 40% less budget on police officers and has 30% fewer police officers per capita, and NYC gives 14% less salary for police officers than Chicago.

The evidence shows that the Chicago government puts more financial resources on maintaining social order and controlling crimes, and this corresponds to the fact that Chicago has a higher crime rate than NYC.

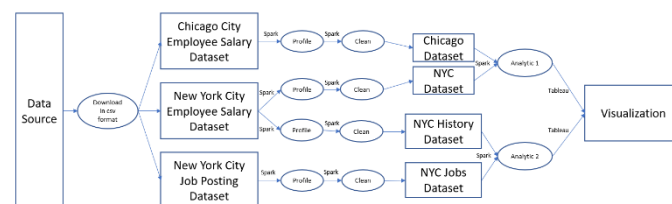
10. The police department and the fire department are the most important agencies in both cities. Most of the city budget is spent on the two departments. Majority of the city employees also belong to these two departments.

Actions can be made according to the analytical findings:

1. The NYC government could spend more budget on police officers and cut down the number of unessential employees in the police department.
2. The Chicago government could re-evaluate whether it is necessary to spend most of its budget on the police department.
3. A notification email could be sent to the job seekers who subscribe to the NYC Jobs dataset when there is a new job posting that matches their search criteria.

VI. APPLICATION DESIGN

Design Diagram:



The first step is to import the three relevant data sources on NYU DUMBO, the Hadoop cluster where we store our datasets and run the analytics. The command line tool curl is used for the data digest. All three datasets are downloaded once (statically) and stored in HDFS.

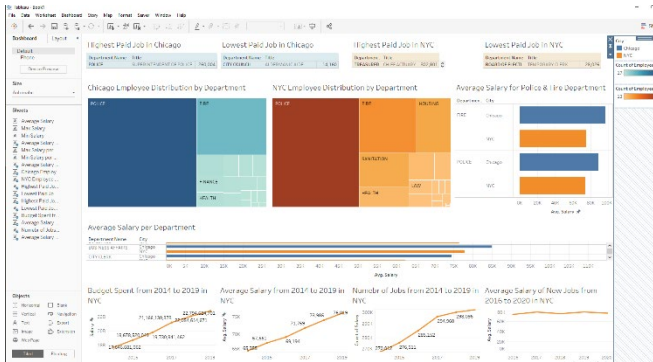
The second step is to clean the datasets. For the first analytic, which compares city employee salary between Chicago and NYC, three columns are extracted from the raw datasets: Department, Title, and Salary. Also, we picked only the 20 city departments that are present in both cities. For the second analytic, which analyzes the trend for city employee salary and job openings, four columns are extracted from the raw datasets: Department, Title, Salary, and Year.

The third step is to profile the raw and the cleaned datasets. This helps us better understand the datasets by providing information such as how many records there are in the raw vs. the cleaned datasets, what the extreme salary values are, and various statistics of the salary column.

The fourth step is to run the two analytics: One for the comparison between NYC and Chicago salaries for year 2019, and another for the trends of NYC salaries and job openings from 2014 to 2019.

The last step is to visualize the insights gained from the two analytics by creating an interactive Tableau dashboard. Apart from looking at the default plots, users can click on the plots and further explore the datasets.

Visualization/UI Rendering:



A Tableau dashboard is used for the visualization. The dashboard contains figures and plots that users can interact with. A user can hover his/her mouse over the plot to see more details about the plot, or click on the sidebar to go to the corresponding plot and design his/her own visualizations based on the data.

The dashboard shows comparisons between Chicago and NYC as well as comparison within NYC across different years. Different colors are used to distinguish NYC Chicago data and NYC data, with Chicago data in blue and NYC data in orange.

At the top of the dashboard, we can see several quantities. They represent the maximal salary, minimal salary, and average salary of the city employees in 2019 for each city.

Beneath the quantities, there are two tree maps showing the distribution of the city employees with respect to city departments for each city.

To the right of the tree maps, there are four bar charts comparing the average salary for police and fire department employees between the two cities.

Going further down, there are multiple bar charts comparing the average salaries for each department between the two cities.

Finally, at the bottom of the dashboard, we can see several line plots which represent the trend for city budget, average salary for city employees, number of city jobs, average salary of new city jobs for NYC from 2014 to 2019.

VII. ACTUATION OR REMEDIATION

For job seekers who wish to find a city job in NYC, the good news is that the city is spending more budget on paying city employees and more city job openings are available each year. However, the average salary for new jobs remains relatively the same over the years.

For job seekers who wish to pursue a job at the police or fire department, choosing to work in Chicago may result in higher and less variant pay. However, they must also be aware that Chicago has more crime activities than NYC and higher salary may indicate higher risk.

The NYC government could spend more budget on police officers and cut down the number of unessential employees in the police department.

The Chicago government could re-evaluate whether it is necessary to spend most of its budget on the police department.

A notification email could be sent to the job seekers who subscribe to the NYC Jobs dataset when there is a new job posting that matches their search criteria.

VIII. ANALYSIS

Experimental setup:

Both the frontend and the backend of this analytic are run on NYU DUMBO. Data used in the analytic are also stored in HDFS on the same Hadoop cluster.

Spark and Spark SQL are the main tools used in the backend of the application. Hive tables are also used to enable the access to the dataset on DUMBO HDFS for Tableau.

Tableau is used for the frontend of the application. In order to interact with data stored on DUMBO, first we enable SSH port forwarding, then we start Tableau Desktop on our laptop and choose SparkSQL to connect to the Thrift server on DUMBO, and finally we choose the schema and tables that we own and build a dashboard based on the data.

Problems encountered in the analytic:

1. Although there are many city departments in common between Chicago and NYC, these departments are named differently. For example, "Fire" in the Chicago dataset but "Fire Department" in the NYC dataset. Therefore, the department names must be standardized before comparison.

2. When running Spark applications in cluster mode, there are few error messages. So first run the application on a subset of the data using the client mode to get more detailed error messages. After the bug is fixed, the application can be deployed on the cluster.

3. Tableau cannot identify .csv files in HDFS on DUMBO. In addition to saving files in .csv format, also create the corresponding Hive tables that Tableau can access.

Lessons learned:

1. Although there are many similar datasets potentially suitable for comparison studies available, different schemas between them can pose a huge problem to programmers. Complex logic might be needed in order to reformat the datasets. However, field knowledge and regular expression matching can save us some work.

2. When developing a Spark application, it is a good idea to first interactively explore the datasets using Spark shell, then compile and test the program on a small portion of the datasets using client mode, and finally deploy the application using cluster mode once you believe it is bug-free.
3. Tableau does not recognize datasets stored in pure text files, but it can access Hive tables stored on the cluster. So store the dataframes also as Hive tables in your Spark program to give Tableau access to your data.

Limitation of the application:

1. The datasets used in the application are not truly “big data”. The datasets are limited in two aspects: a. There are only two cities used for comparison, and b. The history of the records used for analyzing the trends are very short. Therefore, the analytic results might be biased.
2. The application only runs on NYU DUMBO. If a user wants to interact with the analytic, he or she must have access to the Hadoop cluster. This is not very convenient for users outside NYU.
3. The ideas in the Actuation or Remediation part is proposed but not implemented. Users could benefit from the email notification system once it is implemented.

Recommendations for users of the analytic:

1. For job seekers who wish to pursue a city job, especially as a police officer, Chicago provides better salary but higher risks, while NYC provides more job opportunities but less salary.
2. For the NYC government. They should consider spending more budget on police officers and cut down the number of unessential employees in the police department.
3. For the Chicago government. They should re-evaluate whether it is worthwhile to spend most of the city’s budget on its police department. And they should also investigate whether there is a more cost-effective way to control crimes than increasing the number of police officers.

IX. CONCLUSION

In this paper, we revealed the current salary levels for city employees in Chicago and New York City, as well as the salary and job openings trends in New York City in recent years. And we also provided the corresponding visualizations to help readers better understand the numbers and the trends. We noticed that Chicago provides better salaries for city employees but NYC provides more city job opportunities. We also found that the police department and the fire department consume most of the city’s budget. Therefore, finding out a cost-effective

organization for both departments is essential to both cities’ financial resource distributions. Finally, we observed that although there are more job openings and better salaries for full-time city jobs each year in NYC, the pay for new city employees stays approximately the same over the years. And we would like to encourage job seekers not to get frustrated. You might get a pay rise soon!

X. FUTURE WORK

Future work could contain the following aspects:

1. Incorporate more datasets into the analytic. Payroll data from cities other than Chicago and NYC can be used for comparison. And this would give us a more comprehensive view of the city employee salary levels in American cities.
2. If we had data over a longer duration, instead of just from 2014 to 2019, we could build predictive models that forecast the salary or job opening trends.
3. As described in the Actuation or Remediation section, an email notification mechanism could be set up to remind job seekers whenever there are updates on job postings that match their search criteria.

ACKNOWLEDGMENT

Firstly, I appreciate the supports from NYU HPC staffs. They maintained the cluster and made it possible for me to store large data files and run heavy jobs on the cluster. Secondly, I would like to thank NYC Big Data Platform and Chicago Data Portal for generously sharing their data sources to the public. Thirdly, I want to thank Tableau for providing me with the free trial license for the visualization. Finally, I want to thank Professor McIntosh for giving me valuable advice and helping me solve technical problems throughout the research for this project.

REFERENCE

1. Ester R. Fuchs. Mayors and money: Fiscal policy in New York and Chicago, University of Chicago Press, Chicago, IL, May 1992.
2. Walter Fogel & David Lewin. Wage Determination in the Public Sector, Industrial and Labor Relations Review, February 1974.
3. [Current Employee Names, Salaries, and Position Titles. Chicago Data Portal.](#)
4. [Citywide Payroll Data \(Fiscal Year\). NYC OpenData.](#)
5. [NYC Jobs. NYC OpenData.](#)