

A CUSTOM PROJECT ON

CITY PAYROLL ANALYSIS

Project by:

Kripanshu Bhargava – kxb162030

Rahul Sengupta – rxs161630

Table of Contents

1.Introduction

2.Target Audience

3.Description of Data Sources

4.Data Integration

5.Data Product Results

6.Justification

7.Summary

1.Introduction

The United States is a land of opportunities. There are many jobs available for people, but the main concern is to relocate to the different city. Los Angeles and New York are one of the most populated cities with the most number of job opportunities. It is not easy to select a profession in these cities. There are several factors which come into play for relocation to these cities. But few of them are the Basic Pay, Annual Salary, Average Hour pay, Overtime etc. We provide a tool to visualize these factors along with other factors such as job position, department etc. This tool uses semantic web technologies to give the meaning to the data, doing this we not only visualize the data within LA and New York city but also compares these two cities.

2.Target Audience

Our project extracts information from two datasets and displays lucient information on the different types of factors affecting the relocation to the cities. There are several sources to collect the data, but all of them are distributed over the internet, and if are available then it's in a form which is hard for users to understand. Now, even if they understand the datasets then also it's hard to visualize the data as there are millions of rows and columns.

Keeping all these complexities in mind, we made a tool which targets all the employees or job seekers who are willing to move to LA city or New York.

If the audience has this tool in their hand then it becomes easy for them to visualize the million items dataset containing departments, pay, hour rate etc. with simple graphs.

3.Description of data sources

For this project, we are using two large datasets

1. LA Payroll Data which has the following attributes:

1. Year
2. Department Title
3. Job Class Title
4. Pay Grade
5. Employment Type
6. Project Annual Salaries
7. Q1 Payments
8. Q2 Payments
9. Q3 Earnings
10. Q4 Payments
11. Earnings over regular pay
12. % Over Regular
13. Total Earnings
14. Base Pay
15. Permanent Bonus Pay
16. Longevity Bonus Pay
17. Temporary Bonus Pay

18. Overtime
19. Lump Sum Pay
20. Other Pay and Benefits
21. Average City Health Costs
22. Average City Dental Costs
23. Average City Basic Life
24. Total Average Benefit Cost
25. MOU
26. Payroll Department
27. Payroll Division
28. Other Pay
29. MOU Title
30. FMS Department
31. Job Class
32. Benefits Plan
33. Employee ID
34. Job Class Link Description
35. Job Class Link

2. NYC Employee Payroll Data which has the following attributes:

- a) Fiscal Year
- b) Agency Name
- c) Last Name
- d) First Name
- e) Mid Init
- f) Agency Start Date
- g) Work Location Borough
- h) Title Description
- i) Leave Status as of June 30
- j) Base Salary
- k) Pay Basis
- l) Regular Hours
- m) Regular Gross Paid
- n) OT Hours
- o) Total OT Paid
- p) Total Other Pay

	Fiscal Year	Agency Name	Last Name	First Name	Mid Init	Agency Start Date	Work Location	Job Title	Description	Leave Status	Base Salary	Pay Basis	Regular Hours	Regular Gross	POT Hours
1	2017	DISTRICT ATTORNEY	XXX	XXX	L	02/10/1992	MANHATTAN	SUPERVISING RACKETS INVESTIGATOR		CEASED	\$89,780.00	per Annum	1,235.43	\$78,923.55	160
2	2017	FIRE DEPARTMENT	OSTROVE	GREGORY	M	04/18/2016	QUEENS	EMERGENCY MEDICAL SPECIALIST-EMT		ACTIVE	\$36,155.00	per Annum	2,126.37	\$36,194.21	196.25
3	2017	FIRE DEPARTMENT	OSULLIVAN	BRIAN	J	09/14/2003	MANHATTAN	LIEUTENANT		ACTIVE	\$90,304.00	per Annum	2,085.72	\$92,746.94	419.5
4	2017	FIRE DEPARTMENT	OSULLIVAN	STEPHEN	J	01/21/1990	BROOKLYN	LIEUTENANT		ACTIVE	\$90,304.00	per Annum	2,085.72	\$92,173.90	351.25
5	2017	FIRE DEPARTMENT	O'SULLIVAN	BRIAN		09/08/1991	MANHATTAN	LIEUTENANT		ACTIVE	\$106,175.00	per Annum	2,085.72	\$109,232.32	906.35
6	2017	FIRE DEPARTMENT	O'SULLIVAN	CHRISTOPHER	M	02/28/1994	BROOKLYN	LIEUTENANT		ACTIVE	\$90,304.00	per Annum	2,085.72	\$91,972.90	189.98
7	2017	FIRE DEPARTMENT	O'SULLIVAN	DANIEL	J	01/26/1997	MANHATTAN	CAPTAIN		ACTIVE	\$121,875.00	per Annum	2,085.72	\$125,384.34	365.25
8	2017	FIRE DEPARTMENT	O'SULLIVAN	DERMOTT	M	07/07/1999	BRONX	FIREFIGHTER		ACTIVE	\$85,292.00	per Annum	2,085.72	\$87,147.45	419.25
9	2017	FIRE DEPARTMENT	O'SULLIVAN	JAMES	J	04/21/2014	QUEENS	EMERGENCY MEDICAL SPECIALIST-EMT		ACTIVE	\$42,610.00	per Annum	2,044.87	\$39,934.06	88.75
10	2017	FIRE DEPARTMENT	O'SULLIVAN	JOHN	J	11/14/1999	BROOKLYN	FIREFIGHTER		ACTIVE	\$85,292.00	per Annum	2,085.72	\$87,147.45	291.8
11	2017	FIRE DEPARTMENT	O'SULLIVAN	MICHAEL	P	08/08/1999	BROOKLYN	FIREFIGHTER		ACTIVE	\$85,292.00	per Annum	2,085.72	\$87,147.45	825.87
12	2017	FIRE DEPARTMENT	O'SULLIVAN	MICHAEL	E	05/04/2003	QUEENS	FIRE MARSHAL		ACTIVE	\$95,527.00	per Annum	2,085.72	\$97,605.15	255.5
13	2017	FIRE DEPARTMENT	O'SULLIVAN	PATRICK	J	12/12/2016	BROOKLYN	FIREFIGHTER		ACTIVE	\$43,904.00	per Annum	1,154.28	\$23,575.86	77.98
14	2017	FIRE DEPARTMENT	O'SULLIVAN	RICHARD		02/02/1998	QUEENS	SUPERVISOR CARPENTER		ACTIVE	\$370.16	per Day	1.825	\$99,943.20	845.75

Figure 1: NYC Employee Payroll Data

	Year	Department	Tr	Job Class Title	Pay Grade	Employment	Projected Annual	Hourly or Event	R	Q1 Payments	Q2 Payments	Q3 Earnings	Q4 Earnings	Earnings over R	% Over R	Total Earnings	Base Pay	
1	2015	Zoo		Zoo Veterinarian III	3	Full Time	\$131,836.32	\$63.14		\$30,262.88	\$35,358.40	\$0.00		\$1,536.48	2.4%	\$65,621.28	\$64,084.6	
2	2014	Zoo		Zoo Veterinarian III	3	Full Time	\$131,547.13	\$63.00		\$28,350.74	\$33,444.04	\$28,666.32	\$35,280.84	\$9,358.82	8.04%	\$125,741.94	\$116,383.1	
3	2013	Zoo		Zoo Veterinarian II	2	Full Time	\$115,194.96	\$55.17		\$24,652.80	\$28,761.60	\$25,478.40	\$30,895.20	\$0.00	0%	\$109,788.00	\$109,788.0	
4	2012	Zoo		Zoo Veterinarian II		Full Time	\$97,906.00	\$46.89		\$11,254.00	\$0.00	\$0.00	\$0.00	\$-86,652.00	-89%	\$11,254.00		
5	2015	Zoo		Zoo Veterinarian II	2	Part Time	\$106,198.29	\$50.86		\$406.89	\$0.00	\$0.00	\$0.00	\$10.89	2.75%	\$406.89	\$396.0	
6	2015	Zoo		Zoo Veterinarian II	2	Part Time	\$106,198.29	\$50.86		\$5,162.42	\$5,543.88	\$0.00	\$0.00	\$286.55	2.75%	\$10,706.30	\$10,419.7	
7	2015	Zoo		Zoo Veterinarian II	2	Part Time	\$103,356.00	\$49.50		\$8,712.00	\$15,840.00	\$0.00	\$0.00	\$0.00	0%	\$24,552.00	\$24,552.0	
8	2015	Zoo		Zoo Veterinarian II	2	Part Time	\$103,356.00	\$49.50		\$22,468.05	\$5,865.75	\$0.00	\$0.00	\$0.00	0%	\$28,333.80	\$28,333.8	
9	2015	Zoo		Zoo Veterinarian II	2	Full Time	\$118,368.72	\$56.69		\$27,210.28	\$31,746.40	\$0.00	\$0.00	\$485.48	0.83%	\$58,956.68	\$58,471.2	
10	2014	Zoo		Zoo Veterinarian II	2	Part Time	\$106,198.29	\$50.86		\$0.00	\$0.00	\$0.00	\$406.89	\$10.89	2.75%	\$406.89	\$396.0	
11	2014	Zoo		Zoo Veterinarian II	2	Part Time	\$103,356.00	\$49.50		\$0.00	\$0.00	\$0.00	\$19,483.20	\$0.00	0%	\$19,483.20	\$19,483.2	
12	2014	Zoo		Zoo Veterinarian II	2	Part Time	\$106,198.29	\$50.86		\$4,475.79	\$5,925.35	\$4,882.68	\$5,289.57	\$550.64	2.75%	\$20,573.39	\$20,022.7	
13	2014	Zoo		Zoo Veterinarian II	2	Full Time	\$109,118.88	\$52.26		\$25,084.80	\$29,265.60	\$25,084.80	\$25,154.48	\$8,431.28	8.77%	\$104,589.68	\$96,158.4	
14	2014	Zoo		Zoo Veterinarian II	2	Full Time	\$118,362.72	\$56.69		\$26,481.60	\$30,895.20	\$26,724.37	\$31,744.79	\$974.21	0.85%	\$115,845.96	\$114,871.7	
15	2013	Zoo		Zoo Veterinarian II	2	Part Time	\$101,581.20	\$48.65		\$1,167.60	\$0.00	\$0.00	\$0.00	\$0.00	0%	\$1,167.60	\$1,167.6	
16	2013	Zoo		Zoo Veterinarian II	2	Part Time	\$106,198.29	\$50.86		\$4,398.95	\$5,248.74	\$4,868.72	\$5,518.44	\$536.25	2.75%	\$20,034.85	\$19,498.6	
Totals										21,479,627,266.55	\$38,992	\$4,804,936,189.22	\$5,409,309,320.82	\$3,683,797,005.57	\$4,214,232,641.03	\$83,503,892.00		\$9,614,736,687.2

Figure 2: LA City Payroll Data

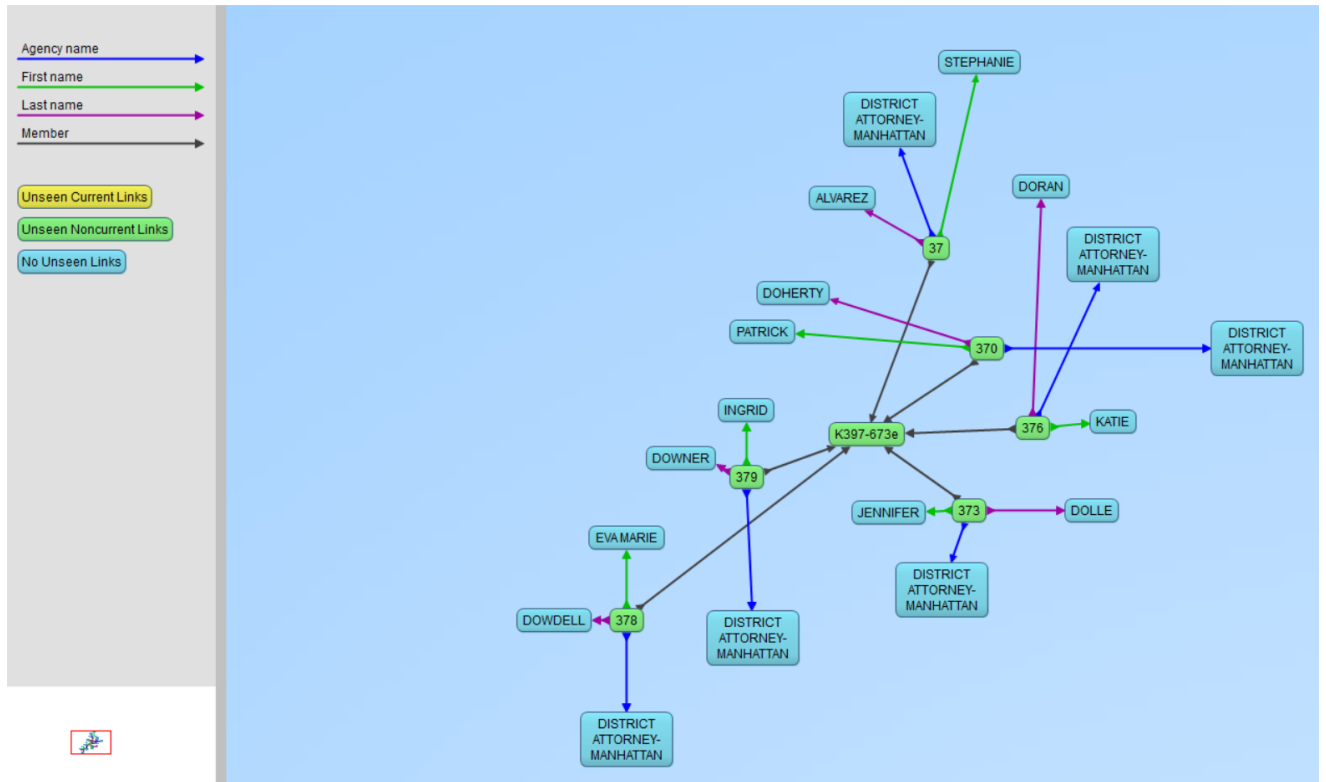


Figure 3: NYC Employee Payroll Ontology

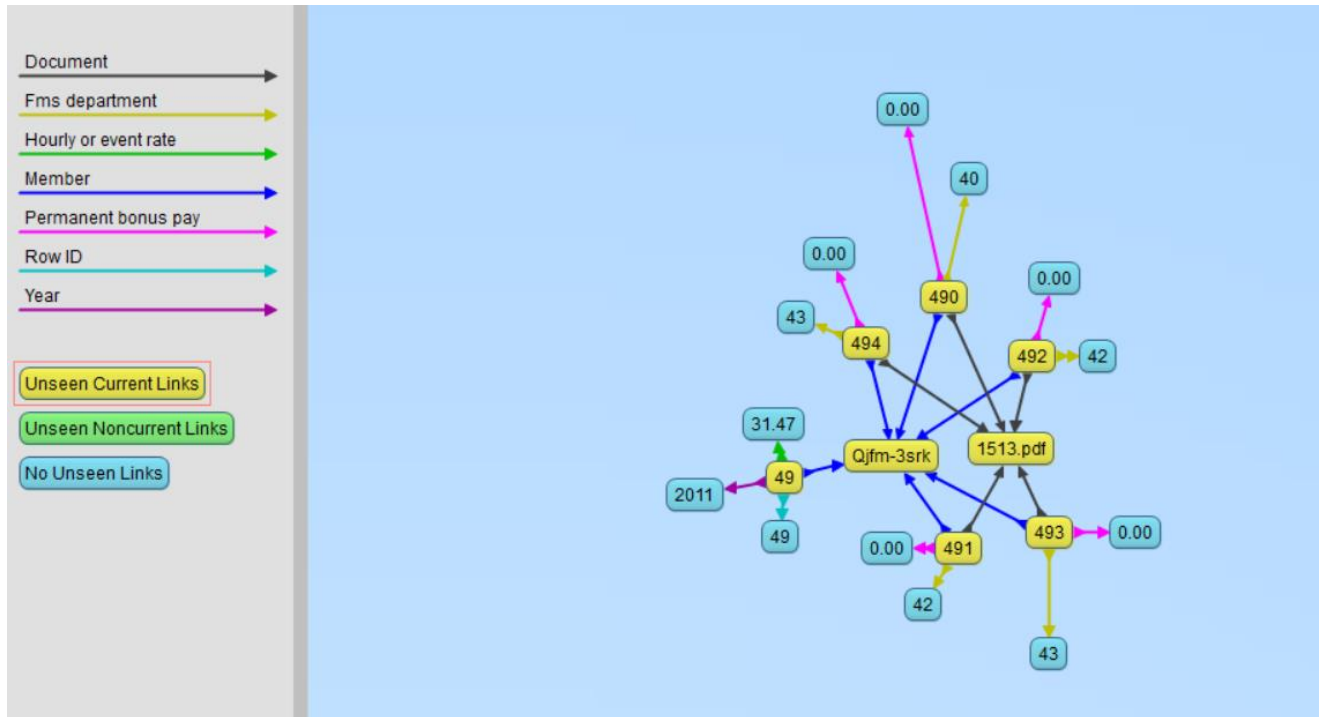


Figure 4: LA City Payroll Ontology

4.Data Integration

The two datasets LA City Payroll and NYC Employee Payroll are hosted in a local Fuseki Server where we look upon the query patterns. The graphs of the different datasets are properly analyzed to present interesting inferences. The query patterns are extracted and saved locally which is then used by our web page to visualize interesting results using D3. The visualization part happens simultaneously and there is no lag between the Fuseki server and the web page which is hosted.

Some sample queries we have used:

1. Query: Compare Base Salaries of LA City and NYC

```
prefix g1:<http://localhost:3030/payrolls/data/nyc>
prefix g2:<http://localhost:3030/payrolls/data/lacity>
prefix p1:<https://data.cityofnewyork.us/resource/k397-673e/>
prefix p2:<https://controllerdata.lacity.org/resource/qjfm-3srk/>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
SELECT ?g ?c (avg(xsd:decimal(?salary)) AS ?average)
WHERE
{
  GRAPH ?g{{
    ?s p1:agency_name ?c .
    ?s p1:base_salary ?salary .
  }}UNION{
```

```

        ?s p2:department_title ?c .
        ?s p2:base_pay ?salary
    }}
}

```

GROUP BY ?c ?g

2. Query: Compare Hourly Rate vs Departments of LA City

```

prefix g1:<http://localhost:3030/payrolls/data/nyc>
prefix g2:<http://localhost:3030/payrolls/data/lacity>
prefix p1:<https://data.cityofnewyork.us/resource/k397-673e/>
prefix p2:<https://controllerdata.lacity.org/resource/qjfm-3srk/>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
SELECT ?deptTitle ?jobClass (avg(xsd:decimal(?hrate)) AS ?avgSal)
WHERE {
    GRAPH g2:{
        ?s p2:department_title ?deptTitle .
        ?s p2:job_class_title ?jobClass .
        ?s p2:hourly_or_event_rate ?hrate
    }
}
GROUP BY ?jobClass ?deptTitle

```

3. Query: Compare Average Total Earnings vs Departments of LA City

```

prefix g1:<http://localhost:3030/payrolls/data/nyc>
prefix g2:<http://localhost:3030/payrolls/data/lacity>
prefix p1:<https://data.cityofnewyork.us/resource/k397-673e/>
prefix p2:<https://controllerdata.lacity.org/resource/qjfm-3srk/>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
SELECT ?deptTitle ?jobClass (avg(xsd:decimal(?tEarnings)) AS ?avgSal)
WHERE {
    GRAPH g2:{
        ?s p2:department_title ?deptTitle .
        ?s p2:job_class_title ?jobClass .
        ?s p2:actual_earnings_example ?tEarnings
    }
}
GROUP BY ?jobClass ?deptTitle

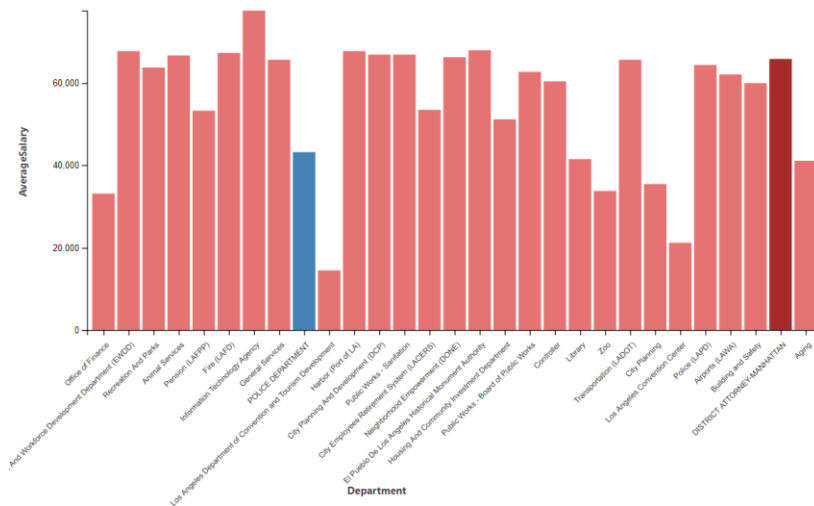
```

5.Data Product Results

The results are shown on our webpage, which is shown to the end users, who in our case are professional job seekers and existing employees looking to relocate or switch. The different patterns obtained through the dataset are shown in the form of visualizations, where they can infer the results. The visualizations of the above three sample queries are shown below:

Please select a query to be run

--choose a query--



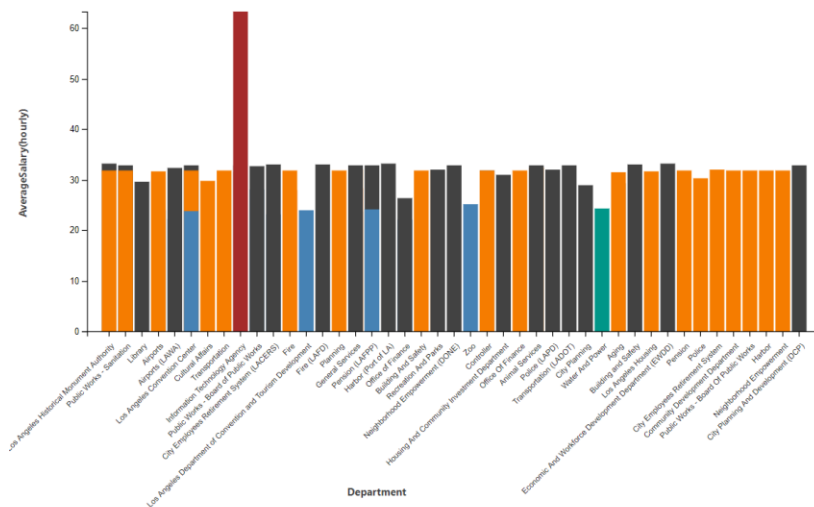
This barchart shows the **average base salary** vs **departments of Los Angeles and NYC**

Department: DISTRICT ATTORNEY-MANHATTAN has base salary: 65840.969012096774193548387091 of city: nyc

Figure 5: Compare Base Salaries of LA City and NYC

Please select a query to be run

--choose a query--

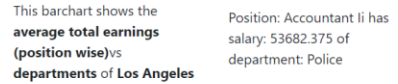


This barchart shows the **average hourly salary (position wise)** vs **departments of Los Angeles**

Position: 311 Director has salary: 63.256666666666666666666666666666 of department: Information Technology Agency

Figure 6: Compare Hourly Rate vs Departments of LA City

--choose a query-- ▼



Using Semantic Web technologies like Fuseki for hosting RDF files, HTML and JavaScript/d3.js for visualization, we have derived useful information from the datasets, and made a tool to visualize a data in a simple and efficient way.