Subproject 1

Data Mining

Daniel Lin

Norman Lim

Manveer Bhangu

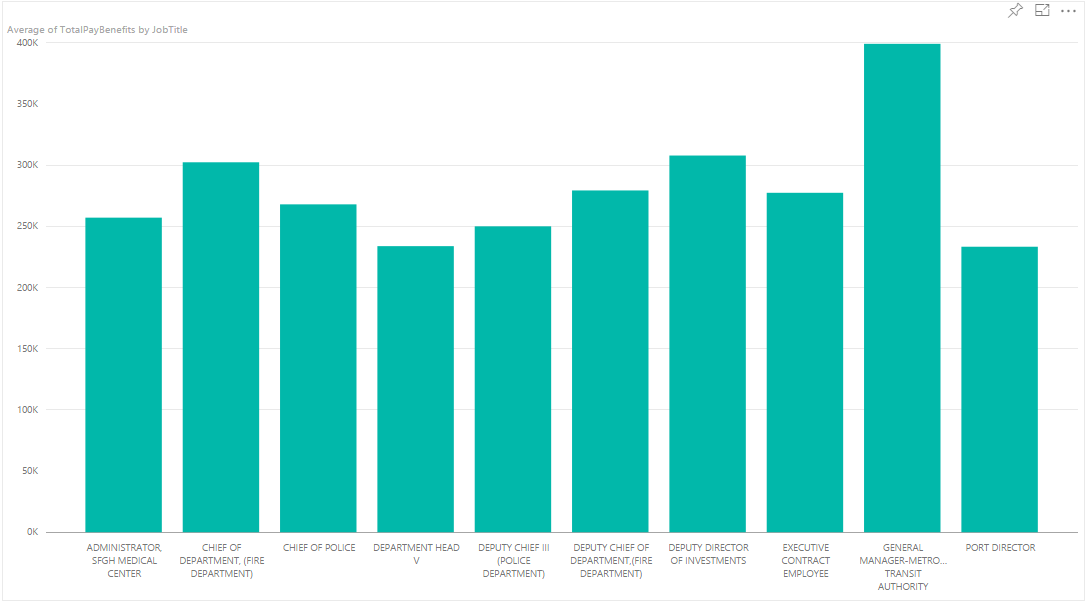
SF Salaries

|  |
| --- |
| BI Insights |
| Are public servants for essential services working more overtime?  Public employees that is essential to the smooth operation of a city occupies a significant portion of the overtime pay. With the transit operator and firefighter to be the top two, followed by police officer.    We can look at the instance (count) of actual overtime pay between these professions.   |  |  | | --- | --- | | JobTitle | TRANSIT OPERATOR | |  |  | | **Count of OvertimePay** |  | | 9424 |  |  |  |  | | --- | --- | | JobTitle | Police Officer | |  |  | | **Count of OvertimePay** |  | | 1476 |  |   Police officer and Transit operator being some of the most common professions to be paid to work overtime. In comparison, a common profession like a junior clerk have far less occurrences of a paid overtime.   |  |  | | --- | --- | | JobTitle | JUNIOR CLERK | |  |  | | **Count of OvertimePay** |  | | 755 |  | |

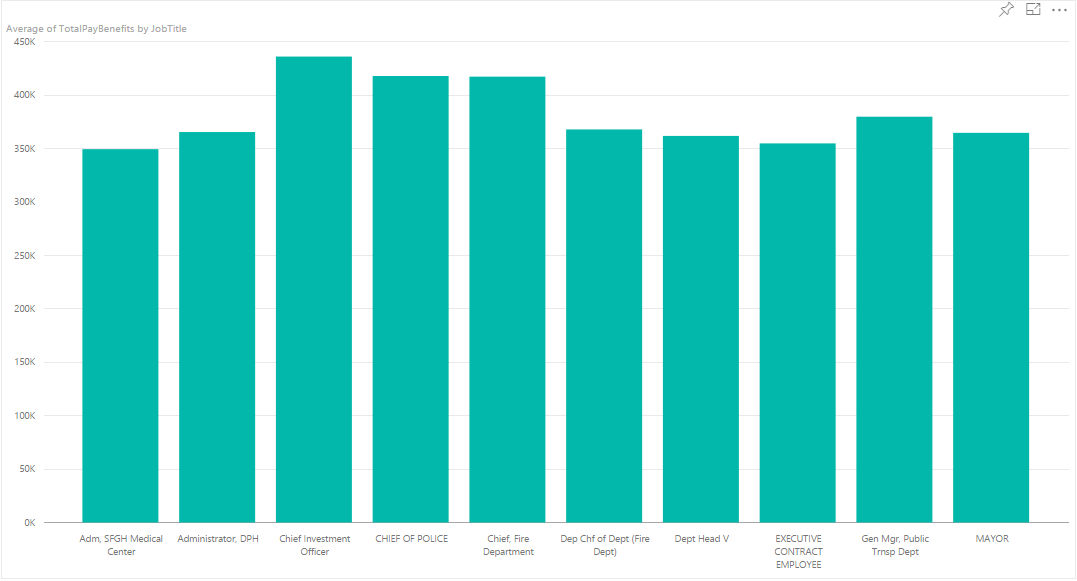
|  |
| --- |
| Excel Insights |
| How has pay change over the years? We want to find if the apparent increase of the salary is significant.   |  |  |  |  | | --- | --- | --- | --- | | **Row Labels** | **Average of BasePay** | **Average of OvertimePay** | **Average of OtherPay** | | 2011 | 63595.95652 | 4531.065429 | 3617.081926 | | 2012 | 65436.40686 | 5023.417824 | 3653.437583 | | 2013 | 69630.03022 | 5281.64198 | 3819.969007 | | 2014 | 66564.42192 | 5401.993737 | 3505.421251 |   Using t-Test (unequal variance): Two Sample for Means, we examined total salary from 2011 and 2014 to see if there are any significant difference over 4 years. We found that there is a significant probable difference between the year 2011 and 2014 in terms of salary.   |  |  |  | | --- | --- | --- | | t-Test: Two-Sample Assuming Unequal Variances | |  | |  |  |  | |  | *Variable 1* | *Variable 2* | | Mean | 71678.18572 | 75335.15144 | | Variance | 2249389889 | 2668368509 | | Observations | 33962 | 35712 | | Hypothesized Mean Difference | 0 |  | | df | 69586 |  | | t Stat | -9.740602164 |  | | P(T<=t) one-tail | 1.04569E-22 |  | | t Critical one-tail | 1.644875525 |  | | P(T<=t) two-tail | 2.09139E-22 |  | | t Critical two-tail | 1.959998076 |  | |

***Top 10 paying job in 2011 compared to 2014***

**Top paying jobs in 2011**



**Top paying jobs in 2014**



**Overall pay increase**

From the two graph we can see that top paying jobs didn’t really change a whole lot but one interesting thing we noticed from this data is that there was an average of 50K increase in the top jobs from 2011 to 2014. This made us curious so from this test, we decided to test if lower paying jobs saw any type of large increase over 3 years or not.

**Lower paying jobs test**

Hypothesis:

Since we know there was a significant increase in pay overall in the top paying jobs we believe that there will also be an increase in lower paying jobs from 2011 to 2014. It may just be a slight increase but we think there will be one.

Data:

For our data we randomly choose 10 jobs between the pay of $25,000 and $32,000 in 2011 and compared them to 2014. We choose a double bar graph to help represent our findings.

2011:

Accountant Intern 29031

Assistant Recreation Director 29442

Beautician 31301

Clerk 28968

Crafts Instructor 31213

Home Health Aide 28066

Museum Guard 29352

Museum Preparator 25896

Recreation Specialist 31016

Sheriff’s Cadet 29312

2014:

Accountant Intern 32406

Assistant Recreation Director 19801

Beautician 70411

Clerk 55279

Crafts Instructor 53383

Home Health Aide 57535

Museum Guard 42266

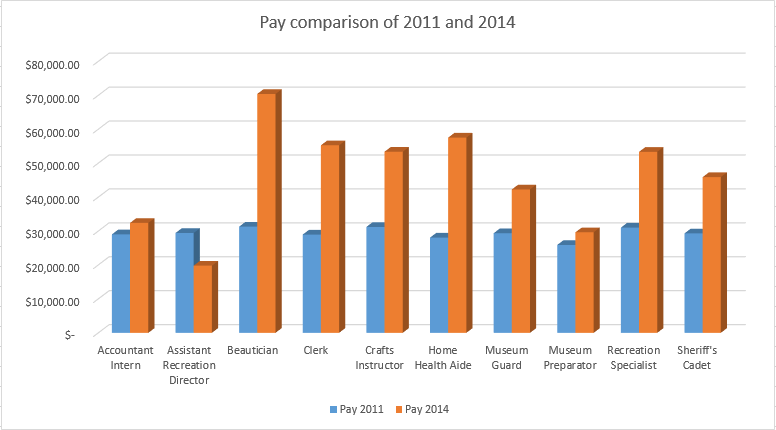
Museum Preparator 29636

Recreation Specialist 53355

Sheriff's Cadet 45913

Conclusion:

From the graph below we can see that aside from Assistant Recreation Director, the pay increase for some of the jobs went up drastically such as a Beautician who saw an increase of 30K from 2011 to 2014. Overall our hypothesis was correct and we can conclude that most jobs, even lower paying jobs, saw an increase in pay.



***GOOGLE SOME SHIT ABOUT WHY ALL JOBS SAW SUCH A SOLID INCREASE, EX. HOUSING MARKET, COST OF LIVING ECT. AND PUT IT HERE***

Top paying information expanded

Snip tool

Jobtitle investment, only guy yearly, if you are highly paid, then you might have the word invrestment in your job title.

SELECT DISTINCT

JOBTITLE,

AVG(TOTALPAYBENEFITS) AS TPB

FROM SALARIES

WHERE

YEAR = 2014

GROUP BY JOBTITLE

ORDER BY TPB DESC

SELECT

EMPLOYEENAME,

JOBTITLE,

YEAR,

AVG(TOTALPAYBENEFITS) AS TPB

FROM SALARIES

WHERE

--YEAR = 2014

JOBTITLE LIKE '%Investment%'

GROUP BY JOBTITLE, EMPLOYEENAME, YEAR

ORDER BY TPB DESC

An hourly rate of $12.25 equates to a weekly pay of $490, monthly pay of $2,123, and an annual salary of $25,480.

RANGE WE CHOSE TO USE 50% UP AND BELOW

12740

38220

Narrow until we got around 10 job titles close to the average.

2011

Accountant Intern 29031

Assistant Recreation Director 29442

Beautician 31301

Clerk 28968

Crafts Instructor 31213

Home Health Aide 28066

Museum Guard 29352

Museum Preparator 25896

Recreation Specialist 31016

Sheriff’s Cadet 29312

2014

Accountant Intern 32406

Assistant Recreation Director 7801

Beautician 70411

Clerk 55279

Crafts Instructor 53383

Home Health Aide 57535

Museum Guard 42266

Museum Preparator 29636

Recreation Specialist 53355

Sheriff's Cadet 45913

Overtime shows who worked the hardest