Subproject 1

Data Mining

Daniel Lin

Norman Lim

Manveer Bhangu

SF Salaries

|  |
| --- |
| BI Insights |
| 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. |

|  |
| --- |
| 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 |  | |