Michael D and Gabrielle K

Professor Jeneralczuk

STAT 501: Methods of Applied Statistics

University of Massachusetts Amherst

11 May 2017

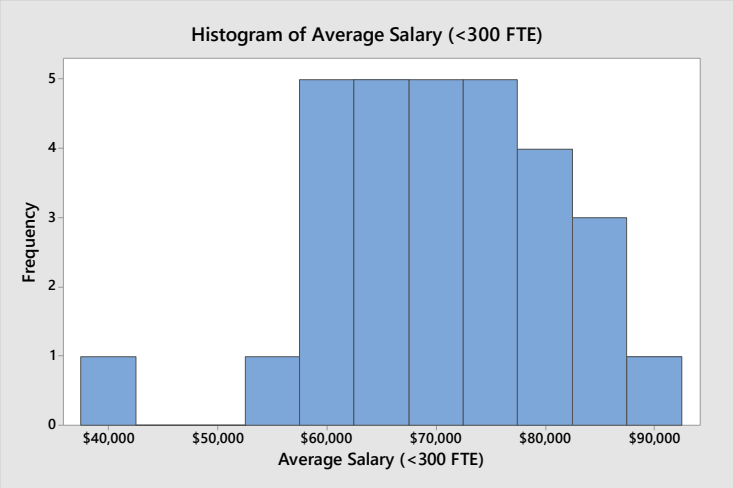
Analysis of the Differences Between Average Massachusetts Teacher Salaries by District

Introduction

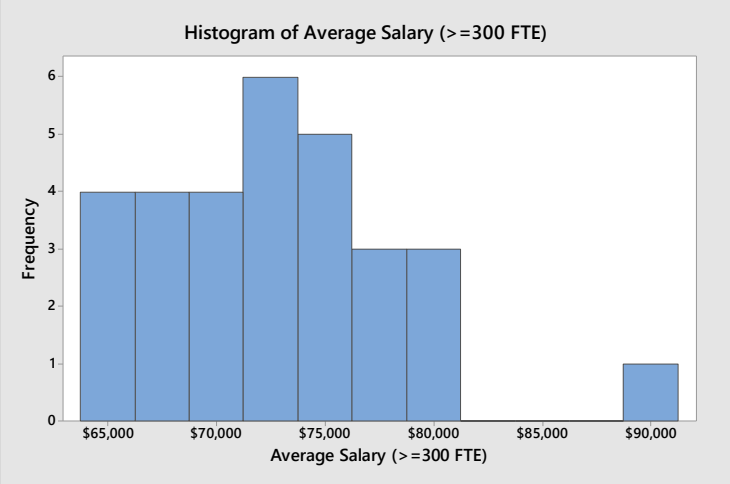
The salaries of elementary and secondary school teachers in different locations are far from uniform. For example, in Massachusetts, any given city may have an average annual teacher salary that is about $20,000 higher than another given city. There is a plethora of direct and indirect causes of this difference. In choosing where to live in Massachusetts, it may be useful for someone who has or plans to have children to consider how much money teachers get paid in different cities in case that might affect the quality of their children's education. They may, for instance, choose to live in a larger city if the teachers are paid more. This same information may also be helpful for elementary and secondary school teachers in choosing where to work. The goal of this study is to determine whether there is a difference between the mean of all average teacher salaries of cities that employ a large amount of full-time equivalent (FTE) teachers and the mean of all average teacher salaries, by city, in Massachusetts.

Data

The data examined in this study are the 2014-15 averages of elementary and secondary school teacher salaries, listed by city/district, in Massachusetts. The data was obtained as an online report by the Massachusetts Department of Education (DoE). One data unit represents the overall average of annual salaries of all teachers in a given city/district. The averages were computed by the DoE by dividing the total teaching salaries in each city/district by its number of FTE teachers. Available variables include the district salary totals, average salaries, and number of FTE teachers. The response variable is the average salary, and the key experimental variable is the number of FTE teachers. We took a random sample of 30 averages of districts with less than 300 FTE teachers and a sample of 30 averages of districts with 300 or more FTE teachers (a random number generator program was used to choose the samples).



The histogram of average salaries of cities/districts that have less than 300 teachers is normally distributed with an outlier salary of about $40,000.



The histogram of average salaries of cities/districts that have more than 300 teachers is normally distributed with an outlier salary of about $90,000.

Descriptive Statistics: Average Salary (<300 FTE)

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Average Salary (<300 FTE)	30	0	69982	1885	10323	40903	62305	70153	77770	87633

Descriptive Statistics: Average Salary (>=300 FTE)

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Average Salary (>=300 FTE)	30	0	72898	1035	5671	64515	68314	73279	76205	88790

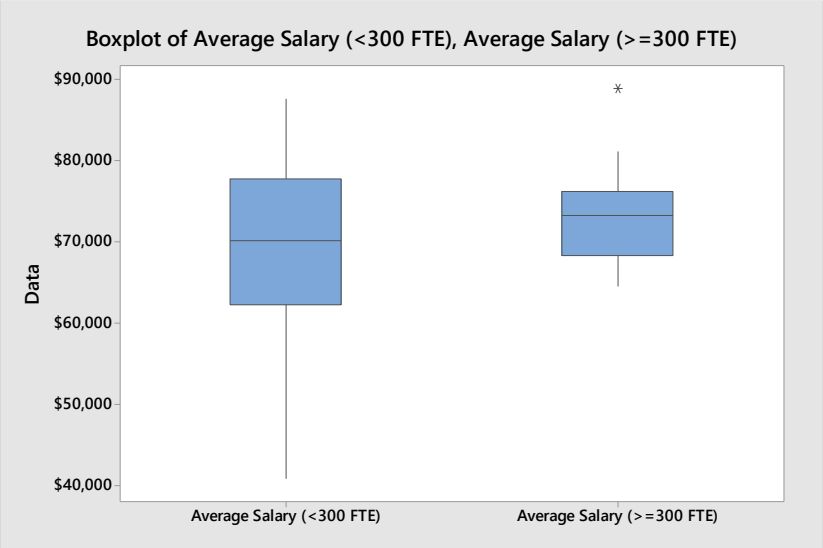
One-Sample T: Average Salary (<300 FTE)

```
Variable                      N   Mean  StDev  SE Mean      95% CI
Average Salary (<300 FTE)  30  69982  10323     1885  (66128, 73837)
```

One-Sample T: Average Salary (>=300 FTE)

```
Variable                       N   Mean  StDev  SE Mean      95% CI
Average Salary (>=300 FTE)  30  72898   5671     1035  (70780, 75015)
```

The descriptive statistics and 95% confidence intervals for the true (population) means. The descriptive statistics of the two samples show about a $3,000 higher mean average salary for districts with at least 300 teachers ($72,898) than that of those with less than 300 teachers ($69,982). The confidence intervals for the population means are close in value.



The boxplot of the first sample appears to have a wide range of salaries. The second sample has an outlying salary of about $90,000. The means of both samples appear to be close.

```
Difference = μ (Average Salary (>=300 FTE)) - μ (Average Salary (<300 FTE))
Estimate for difference:  2915
95% CI for difference:  (-1416, 7246)
T-Test of difference = 0 (vs ≠): T-Value = 1.36  P-Value = 0.182  DF = 45
```

The confidence interval for the difference in the true (population) means (-$1,416, $7,246). Notice 0 is contained in the interval.

```
t-observed=1.36 alpha=0.05 df=45 t-alpha=2.014
```

The test statistic t for the difference in means. Notice the value of t-observed (1.36) is less than that of t-alpha (2.014).

```
p-value=0.182 alpha=0.05
```

The p-value for the difference in means (0.182). Notice its value is less than that of alpha (0.05).

The same histograms as before, except with the outliers, one at about $40,000 per year and the other at about $90,000 per year, have been removed.

Descriptive Statistics: Average Salary (<300 FTE)

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Average Salary (<300 FTE)	29	1	70985	1652	8895	57322	62697	70226	77876	87633

Descriptive Statistics: Average Salary (>=300 FTE)

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Average Salary (>=30 FTE)	29	1	72350	909	4896	64515	68049	73237	76174	81211

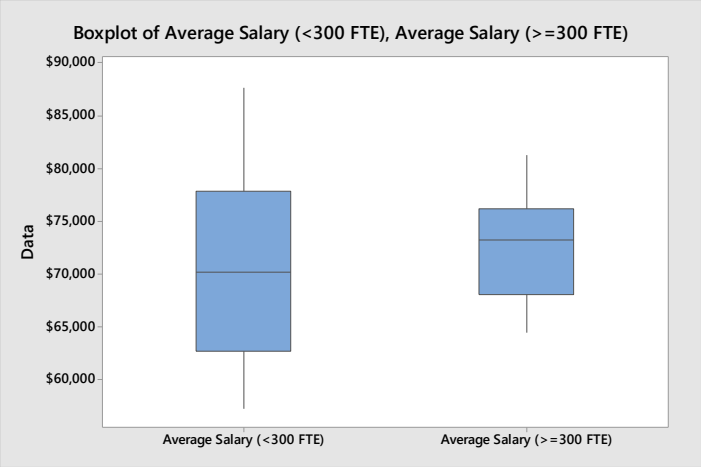
One-Sample T: Average Salary (<300 FTE)

Variable	N	Mean	StDev	SE Mean	95% CI
Average Salary (<300 FTE)	29	70985	8895	1652	(67602, 74369)

One-Sample T: Average Salary (>=300 FTE)

Variable	N	Mean	StDev	SE Mean	95% CI
Average Salary (>=300 FTE)	29	72350	4896	909	(70487, 74212)

The descriptive statistics and 95% confidence intervals for the true (population) mean for each sample, after outliers were removed. The descriptive statistics of the two samples show about a $1,000 higher mean average salary for districts with at least 300 teachers ($72,350) than that of those with less than 300 teachers ($70,985). The confidence intervals for the population means are still close in value.



Boxplot of the samples with the outliers removed. The means of both samples still remain close.

```
Difference = μ (Average Salary (>=300 FTE)) - μ (Average Salary (<300 FTE))
Estimate for difference:  1365
95% CI for difference: (-2438, 5167)
T-Test of difference = 0 (vs ≠): T-Value = 0.72 P-Value = 0.473 DF = 43
```

The 95% confidence interval for the true means, after outliers were removed. Notice that 0 is contained within the interval.

```
t-observed=0.72 alpha=0.05 df=43 t-alpha=2.016
```

The t-test run without outliers. Notice that the value of t-observed (0.72) is less than that of t-alpha (2.016).

```
p-value=0.473 alpha=0.05
```

The p-value (0.473) after the outliers were removed. Notice its value is less than that of alpha (0.05).

Analysis

Confidence intervals, t-tests, and p-values were computed for the population means of each sample and also for the difference in population means between each sample. In our analysis of these methods, it is safe to assume the samples are independent because they were taken from different cities. The sample is random because a random number generator was used to select the values that make up each sample (the numbers generated corresponded to the row numbers of the data in Microsoft Excel). Each of the samples are sufficiently large. The samples were distributed normally, as the histograms show, aside from two outliers which were removed such that the data could be tested again. When calculating the confidence intervals for the difference of two means, it is assumed that the variances are not equal because their standard deviations differ.

Zero is included in both of the 95% confidence intervals for the difference in means. Therefore, the two means are close in value and may even be equal. All of the t-observed values are lesser in value than the t-alpha value. All of the p-values computed are lesser in value than that of alpha. All of these results indicate the same conclusion: the data does not give sufficient evidence to claim that cities/districts with more than 300 teachers have a higher mean salary than

those with less than 300 teachers. It is important to note that the data without outliers gives the same result.

Conclusions

This study found no significant evidence that indicates cities/districts with more than 300 teachers have a higher mean salary than those with less than 300 teachers, but it did not prove that such a difference does not exist. Future studies might examine data of individual teacher salaries between cities rather than averages of all salaries between cities. This way, sample sizes will be larger and it may, thus, be possible to find a difference in means amongst cities with less than, say, 1000 teachers and those with more than 1000, rather than 300. It may also be helpful to incorporate the standard of living factor, as a higher standard of living along with a higher annual salary for a teacher may not indicate that such teachers are making a higher net income. If data were collected for the purpose of further research on this matter, it would be helpful to note where teachers work as well as where they live (for, they could be commuters), the standard of living for where they live/work, and to differentiate between public school salaries and those of private schools or other types of schools.