

Can Money Buy Brains?

A statistical analysis of Washington State teacher salaries and SAT scores

Ming Fong, Halle Hwang, Rae Gerking, Ethan Dang

January 24, 2020

1 Abstract

It's often said that *money can't buy happiness*. Yet in today's economic society, money seems to remain the driving factor in American lives, especially in the education system. While the United States remains the wealthiest nation in the world, it ranks 27th in healthcare and education, trailing behind a host of countries including Iceland, Finland, and the Netherlands. Furthermore, there has been a long debate in the American education system concerning teachers' salaries and whether or not they are paid enough. To this day, many educators feel that they have been denied competitive, professional pay for too long and now, with 2018 being the biggest year of teacher strikes in this generation, the effects could have an impact on the performance of American students. In this retrospective study, we will observe these issues by researching and analyzing average teacher salaries throughout public high schools in Washington state and directly compare them with average student test scores from the respective schools. Our goal in this study is to determine whether or not there is a direct association between teacher salaries and student performance.

2 Data Collection

We collected our data using online government datasets that documented average SAT scores¹ by school district for the 2018-2019 school year, as well as the salary of every public-school teacher². There are 208 data points for the state of Washington where both a district's mean SAT score and secondary teacher salaries were reported. Possible sources of bias include the fact that we only have salary data for public schools, some school districts did not report mean SAT scores and some teacher salaries may have been unreported. In order to draw the most accurate conclusion, we used a large dataset to counteract these biases. The original datasets can be found in the appendix.

¹SAT data: https://github.com/evilpegasus/APStats/blob/master/WA%20_District_SAT_2019.xlsx

²Salary data: https://github.com/evilpegasus/APStats/blob/master/WA_School_Personnel_2018-2019.xlsx

2.1 Data Processing

With the raw salary data, all rows not containing a secondary teacher salary were filtered out and removed. Using a Python script, the secondary teacher salaries were read into a Pandas Dataframe, and the average of each district was found. To combine the datasets, the district names on the SAT dataset were renamed to match the naming convention of the salary spreadsheet. The renamed SAT dataset was then directly read into a Pandas Dataframe. The two Dataframes were merged by matching district names. Because not every district in the SAT dataset was listed on the salary dataset and not every salary district was listed on the SAT dataset, we only included districts with data from both sets.

3 Analysis

3.1 Unit 1

All statistical descriptions for both data sets were calculated with a computer. For the mean SAT data (see Figure 1), the distribution is roughly symmetric and unimodal. It is approximately normal-shaped. The median is at 1046.5 points and the mean is approximately 1043.995 points. The standard deviation is about 90.661 points. The minimum and maximum values are Taholah's low mean score of 803 points and Mercer Island's high mean score of 1316 points, which is an outlier (see Figure 2). The range is 513 points. The first quartile is 982 points and the third quartile is 1103.75 and the $1.5 * IQR$ is 182.625 points. This gives an upper fence of 1286.375 and a lower fence of 799.375. This confirms that Mercer Island's mean score of 1316 is an outlier and Taholah's mean of 803 points is not.

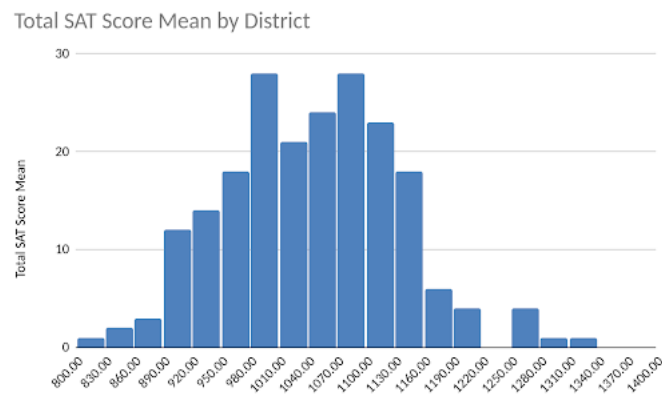


Figure 1: Histogram

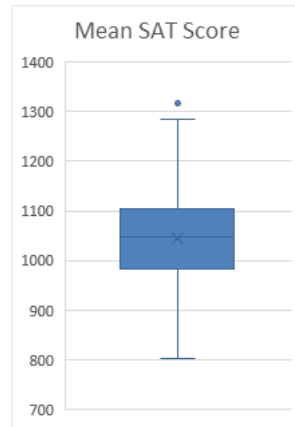


Figure 2: Box and Whiskers Plot

For the average teacher salaries (see Figure 3), the data is multimodal and roughly symmetric and normal shaped. The mean is approximately \$80,040.33 and the median is \$80,158.9058. The standard deviation is about \$9,241.6781. There are two outliers, \$47,796.83333 (Coulee-Hartline) and \$107,466.3848 (Everett). The range is about \$59,669.5515. The first quartile is \$73,999.66 and the third quartile is \$85,988.53 and the $1.5 * \text{IQR}$ is \$17,983.3. The fences are \$56,016.36 and \$103,971.8 which confirms that Coulee-Hartline and Everett are indeed outliers (see Figure 4).

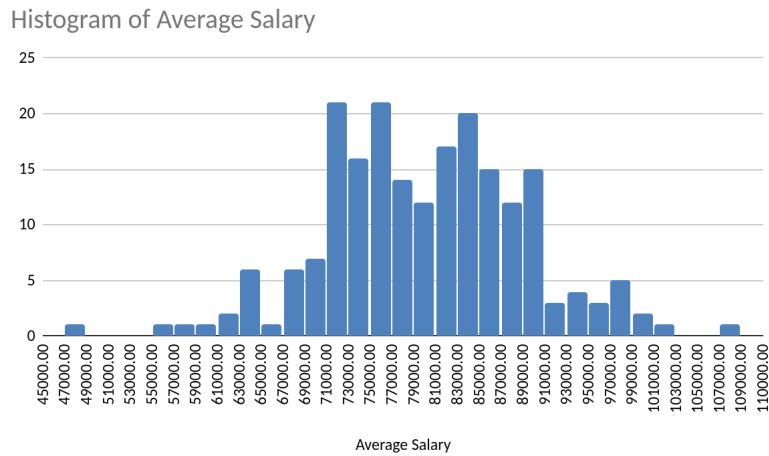


Figure 3: Histogram

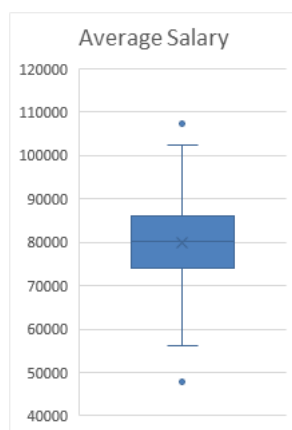


Figure 4: Box and Whiskers Plot

3.2 Unit 2

Our scatter plot showed some positive correlation between teacher salary and mean SAT scores (see Figure 5). After plotting the residuals, we found the R^2 value to be $R^2=0.09423098779302452$. This R^2 value is small, but not negligible, and indicates that the linear relationship is weak. The R^2 value tells us that the model accounts for approximately 9.42% of the variance in average SAT scores by average teacher salaries. Some districts with interesting outliers in teacher salaries include Coulee-Hartline (\$47,796.83) and Everett (\$107,466.3848), the first district being a low outlier and the second being a high outlier. The district of Mercer Island (1316) is a high outlier in terms of average SAT score. Some lurking variables that might have adverse effects on our data and analysis include demographics, urban vs rural areas, and school district funding. For instance, in urban areas, there are more opportunities to enroll in SAT prep programs or get private tutoring than in rural areas. Likewise, the demographics within each school affects the population's abilities to further their SAT prep (through programs or tutoring) outside of school. School district funding positively correlates to SAT scores

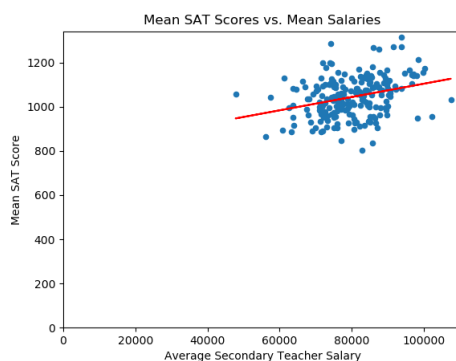


Figure 5: Scatterplot and Linear Regression Line

Linear Regression Equation

$$\hat{y} = 0.00301137x + 802.96409135$$

For each dollar increase in average secondary teacher salary, the mean SAT score of the district is predicted to increase by 0.00301137 points.

Our regression model predicts that a district paying teachers an average of \$0 will have a mean SAT score of 802.96409135. This situation is not realistic as no teacher would work for free. Additionally, a prediction at the y-intercept significantly extrapolates beyond our dataset: the lowest mean salary reported was \$47,796.83.

The R^2 value and the regression equation coefficients were both computed with the Scikit-learn software library for Python.

3.3 Unit 4

The Renton School District has 91 secondary teachers with a 6-figure salary (≥ 100000) out of a total of 295 secondary teachers. Given a simple random sample of 295 teachers from the set of all reported secondary teachers, what is the chance that more than 91 teachers are selected with six-figure salaries? There are 7234 secondary teachers in Washington State with six-figure salaries and 25511 total reported secondary teachers.

We will use a binomial distribution to model this. We first check the conditions:

$$np = 82.610 \geq 10$$

$$nq = 212.4 \geq 10$$

$$n = 295 < 0.10 * 25511$$

We can reasonably assume that most randomly selected salaries will be independent of each other. Using the binomial cumulative distribution function, we can find the probability that a random sample has equal to less than 91 teachers with six-figure salaries. The complement of that output will be our answer.

$$\sum_{i=0}^x \binom{n}{i} (p)^i (1-p)^{n-i} = 0.8447$$
$$1 - 0.8447 = 0.1553$$

Therefore, the probability that a randomly selected sample of 295 teachers has more than 91 teachers with six-figure is 15.53%. This means that the Renton School District has a relatively high concentration of high-earning secondary teachers.

4 Conclusion

This study focused on if teacher salaries and student performance have a direct association and the possible factors involved. From this study, we can conclude that there was a positive association between teacher salaries and student performance in SAT scores

during the 2018-2019 school year. In context, if a district pays their secondary teachers more, their average SAT scores are likely to be higher.

The lurking variables that we theorized were demographics, urban vs rural areas, and school district funding, but it would be impractical to attempt to lessen the influence of these factors in our model. Similarly, another external factor is the variation in the cost of living between districts, which may have facilitated the differences in the demographics and teacher salaries of each school. For further research in the future, we could gather and use data from other states in the US similar to the data used in this study. We could also further consider the lurking factors stated above. Ultimately, our study exposed the flaws of the American education system and the unfair advantages wealth may give to students.

5 Appendix

WA Sat Scores: <https://www.k12.wa.us/student-success/support-programs/dual-credit-programs/exam-based-dual-credit>

WA Salary Data: <https://www.k12.wa.us/safs-database-files>

Repository of raw data, processed data, and analysis scripts: <https://github.com/evilpegasus/apstats>

The following pages contain the processed and combined data that was used in the analysis.

School District	Total Mean SAT Score	Average Salary
Aberdeen	1050	83186.2
Adna	1029	77052.75
Anacortes	1149	97004.68182
Arlington	1140	98069.94444
Asotin-Anatone	1035	72936.66667
Auburn	947	98183.28165
Bainbridge	1267	85921.72727
Battle Ground	1129	80003.29358
Bellevue	1273	91583.06452
Bellingham	1169	81914.19713
Bethel	1046	76995.17841
Blaine	1135	85511.80769
Bremerton	1024	89494.55118
Brewster	890	69024.88
Bridgeport	932	71198.59259
Burlington-Edison	994	87945.53012
Camas	1182	85689.535
Cape Flattery	1028	74721.05
Cascade	1057	69241.22222
Cashmere	1069	82226.97674
Castle Rock	999	71649.9
Central Kitsap	1145	96549.60949
Central Valley	1142	85352.76271
Centralia	1018	84166.94203
Chehalis	1091	86177.83333
Cheney	1092	71898.16197
Chewelah	1045	75127
Chimacum	1101	72636.72
Clarkston	1085	85806.6129
Cle Elum-Roslyn	1081	63626.92308
Clover Park	1054	88222.20721
Colfax	1080	64535.61111
College Place	1037	68112.13333
Colton	1044	74406.44444
Columbia (Wal)	1057	81437.63636
Colville	1006	77341.01923
Coulee-Hartline	1057	47796.83333
Coupeville	1004	89897.74074
Curlew	1043	57367.36364
Cusick	895	60756.41667
Davenport	962	68025.375
Dayton	1035	67641.38889
Deer Park	1056	77768.56061

East Valley (Spo)	1008	82436.41667
East Valley (Yak)	1066	81084.90141
Eastmont	1056	82602.39286
Eatonville	1047	76782.2037
Edmonds	1129	89397.0744
Ellensburg	1107	74097.85333
Elma	1088	87594.15385
Enumclaw	1075	88423.42391
Ephrata	1009	80652.15254
Everett	1031	107466.3848
Evergreen (Clark)	1076	83793.30694
Federal Way	934	84720.11148
Ferndale	1117	90355.06422
Fife	1097	89722.66265
Finley	954	85126.30769
Franklin Pierce	914	83493.28736
Freeman	1115	66346.06897
Goldendale	1130	71497.41667
Grand Coulee Dam	951	63743.13043
Grandview	903	75837.16667
Granger	938	84947.10526
Granite Falls	1079	92116.77551
Highland	925	80937.14286
Highline	947	86321.36364
Hockinson	1091	77484.36
Hoquiam	1067	84651.43902
Inchelium	846	77006.09091
Issaquah	1270	93727.46506
Kalama	1102	75589.16667
Kelso	1016	72694.62385
Kennewick	1054	85505.72881
Kent	1007	84460.24194
Kettle Falls	1116	74731.42857
Kiona-Benton City	931	84956.9
Kittitas	968	71547.86364
La Conner	1070	90610
Lacenter	1033	76775.51282
Lake Chelan	964	80185.97826
Lake Stevens	1139	97158.78495
Lake Washington	1259	87408.32385
Lakewood	1060	90627.64815
Liberty	1076	84189.92857
Longview	1078	78267.67424

Lyle	866	56195.5
Lynden	1111	82627.10769
Mabton	885	71505.95
Manson	923	75870.1
Mary Walker	975	75033.47368
Marysville	954	102265.7524
Mead	1109	88774.76423
Medical Lake	1120	76985.32653
Mercer Island	1316	93696.79508
Meridian	1116	80670.40541
Methow Valley	1128	61129.66667
Monroe	1084	89931.52941
Montesano	1045	90575.79412
Moses Lake	961	87724.80769
Mossyrock	956	86010.8
Mount Adams	990	76754.64286
Mount Baker	1076	88024.38298
Mt Vernon	1033	82926.31953
Mukilteo	1154	99805.66667
Naches Valley	1000	71116.7
Napavine	1020	70947.04348
Naselle-Grays R.	1008	76456.88235
Newport	1101	74007.51515
Nine Mile Falls	1085	75538.63636
Nooksack Valley	1019	80131.83333
North Beach	915	63963.1875
North Franklin	929	81265.1875
North Kitsap	1143	88934.8
North Mason	1022	83878.5
North Thurston	1084	81185.42297
Northshore	1214	98325.44651
Oak Harbor	1113	90054.10853
Oakesdale	1048	71763.57143
Ocean Beach	1070	70176.96296
Ocosta	1110	85015.27778
Okanogan	966	74351.03704
Olympia	1196	74368.97984
Omak	1003	75693.83333
Onalaska	970	72044.72
Orcas Island	1285	74052.35714
Oroville	929	68657.94118
Orting	1057	82200.90566
Othello	1004	76333.51163

Pasco	987	77127.07634
Pateros	927	79064.30769
Peninsula	1135	82913.5
Pomeroy	1007	71862
Port Angeles	1048	82678.05814
Port Townsend	1155	76128.05556
Prescott	891	71967
Prosser	1014	77702.65714
Pullman	1200	73997.04478
Puyallup	1093	91953.35714
Quilcene	952	72115.09091
Quillayute Valley	981	71301.7037
Quincy	945	80546.32203
Rainier	968	79170.33333
Raymond	993	78827.25
Reardan	994	73648.83333
Renton	999	86721.72881
Richland	1145	83950.09477
Ridgefield	1146	77876.08333
Ritzville	985	78109.07143
Riverside	1036	75479.22581
Riverview	1167	89699.13253
Rochester	1101	73636.82143
Royal	958	73285.65789
San Juan Island	1128	83634.15385
Seattle	1103	84796.65109
Sedro-Woolley	1104	86852.28829
Selah	1058	76075.33721
Selkirk	1005	63658.2
Sequim	1136	82465.89394
Shelton	1035	78774.64103
Shoreline	1167	95997.62136
Snohomish	1174	100287.8651
Snoqualmie Valley	1151	84148.99281
Soap Lake	910	72980.16667
South Bend	1014	78757.14286
South Kitsap	985	89551.08673
South Whidbey	1083	93654.05
Spokane	995	89246.97577
Stanwood-Camano	1104	96648.09375
Steilacoom Hist.	1126	84311.07463
Stevenson-Carson	1061	77063.81818
Sultan	1055	90180.73214

Sumner	1090	87388.9122
Sunnyside	913	82104.39189
Tacoma	964	89634.7971
Taholah	803	82888.4
Tahoma	1150	94550.24571
Tenino	1078	70665.48387
Toledo	999	84824.12
Tonasket	989	67519.80645
Toppenish	905	74984.91579
Toutle Lake	1091	74942.38889
Trout Lake	1176	72777.88889
Tukwila	905	87008.21622
Tumwater	1140	83924.30872
University Place	1155	87628.6055
Vancouver	1001	79194.02925
Vashon Island	1197	71903.14894
Wahkiakum	997	62502.58824
Wahluke	939	73278.31667
Walla Walla	1126	74381.31783
Wapato	896	80603.025
Warden	977	72615.22222
Washougal	1081	81542.03448
Waterville	888	63312.5
Wellpinit	834	85841.41176
Wenatchee	1073	79671.39583
West Valley (Spo)	1090	67000.46392
West Valley (Yak)	1016	84908.36538
White Pass	950	75081.16667
White River	1028	83172.96667
White Salmon	1095	70157.05882
Wilbur	986	71092.66667
Winlock	957	75497.6087
Wishram	903	70032.6
Woodland	1020	76084.64912
Yakima	927	86729.04278
Yelm	981	75847.46212