EDUC 643 Assignment 03 Key

1. Descriptive statistics

1.1 Summarize the two categorical variables in your data set (*level* and *locale*). **(1 point)**

In the National Education Resource Database on Schools (NERD$), there are two categorical variables: *level*, which defines the grade-band level of schooling for each school and *locale,* which measures the Census-based location of schools attended by a majority of its students: Urban, Rural, Suburban or Town. As we show in Table 1, in our data, out of a total of 1,193 schools, roughly half are elementary schools with another quarter divided between middle- and high-school levels, each. A very few (6 schools) are at the level of early education. The plurality of schools (34.3 percent) are in urban settings, with others in town (28.2 percent), suburban (22.7 percent), and rural (14.8 percent) locales.

**Table 1**

Descriptive statistics of the grade-band levels and geographic locales of Oregon public schools, 2018-19.

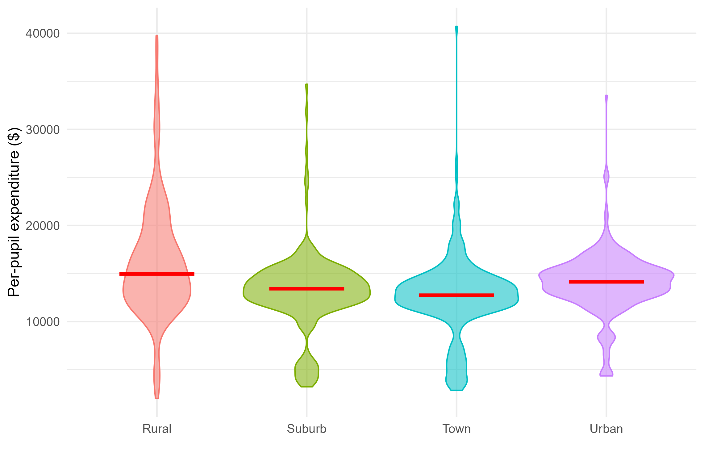
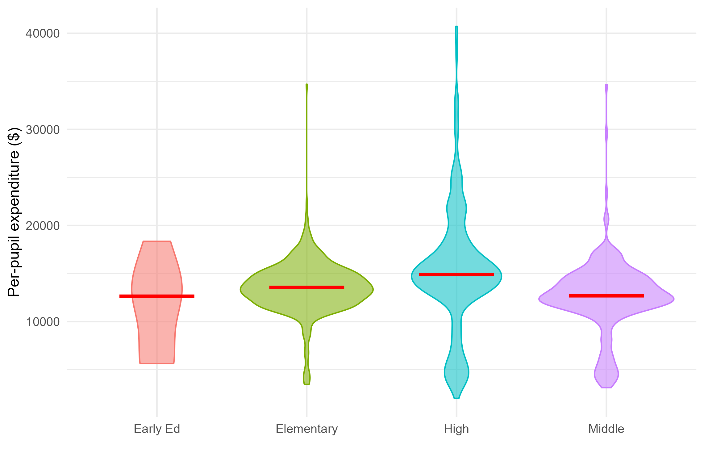
|  |  | N | % |
| --- | --- | --- | --- |
| Grade-band level | Early Ed | 6 | 0.5 |
|  | Elementary | 585 | 49.0 |
|  | High | 307 | 25.7 |
|  | Middle | 295 | 24.7 |
| Geographic locale | Rural | 177 | 14.8 |
|  | Suburb | 271 | 22.7 |
|  | Town | 336 | 28.2 |
|  | Urban | 409 | 34.3 |
| Total |  | 1193 | 100.0 |

1.2 Create one figure describing how per-pupil expenditures differs across different levels of schooling and another figure showing how per-pupil expenditures differ by schooling locale. Interpret these figures. **(2 points)**

In Figure 1, we highlight that there is meaningful variation in per-pupil expenditure both within- and between-grade level (Panel A) and locale (Panel B). In particular, high schools seem to have both higher median expenditure as well as a higher absolute range and a larger standard deviation ($6,331). The mean expenditure at the high-school level is $15,202, compared to only $12,518 at the middle-school level. In Panel B, we demonstrate that rural schools spend more ($16,219) per-pupil, on average, than other geographic locations, which all spend roughly the same median amount, though again there is substantial variability within each location, with standard deviations in PPE between $3,000 and $6,000.

**Figure 1**

*Violin plot of per-pupil expenditure by grade-band level and geographic locale of Oregon public schools, 2018-19*

**

(a) Grade-band level (b) Geographic locale

2. Do schools that serve different educational grade-band levels spend differing amounts on their students?

2.1 State your null hypothesis regarding the above research question. **(1 point)**

**In the population of Oregon public schools, there is no difference in the average per-pupil expenditure for early education, elementary, middle and high schools**

2.2 Test your null hypothesis using the statistical test that most directly and efficiently answers the above research question. Interpret the results of this test. **(2 points)**

We test our null-hypothesis that the mean per-pupil expenditure across all schooling levels is equivalent (H0: *μearly = μelem = μmiddle = μhigh*) via an Analysis of Variance (ANOVA). We calculate an *F*-statistic (*df* = 3, 1189) of 22.1, which exceeds our alpha-threshold of 0.05 and therefore we reject our null hypothesis and conclude that mean per-pupil expenditure across these groups of schools are different, on average in the population.

2.3 State the magnitude by which schools at each of the different grade-bands differ from each other in their per-pupil spending. Select (and justify) the most sensible group to serve as the comparison type of school. **(2 points)**

**In order to test for differences in per-pupil expenditure across each schooling level and interpret the magnitude of these differences, we employ an Ordinary Least Squares estimation strategy. Early education is a poor candidate to serve as a reference category because there are only six schools in this category. As a result, all of our estimates will be imprecise because we will be comparing them to a reference group whose value we can estimate only imprecisely. As a result, we use elementary schools—our largest grouping of schools—as our reference category. We present our results in Table 2. We note that early education settings spend less than elementary schools (-$1959), but we are unable to reject the null that the difference is statistically equivalent to zero. However, we do find that high schools spend more ($1669) per student than elementary schools, whereas middle schools spend less (-$1015).**

**Table 2**

**Ordinary Least Squares estimates of the relationship between schooling level and per-pupil expenditure in Oregon public schools, 2018-19**

|  | (1) |
| --- | --- |
| (Intercept) | 13532.67\*\*\* |
|  | (110.58) |
| Early Education | -1959.04 |
|  | (2219.01) |
| High School | 1669.35\*\*\* |
|  | (378.42) |
| Middle School | -1014.89\*\*\* |
|  | (243.41) |
| Num.Obs. | 1193 |
| R2 | 0.053 |
| + p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001. Cells report coefficients and heteroscedastic-robust standard errors in parentheses. Each observation is one school. | |

2.4 High schools are generally understood to have greater resource needs than other K-12 schools due to fundings items such as science laboratories, extra-curriculars and specialized classes. Test whether high schools spend more per-pupil than all other schools and interpret your results. **(1 point)**

**We examine the question of whether high schools spend more per student than other schooling levels in two ways. First, we compare high schools to each of the other three levels of schooling separately, then we test whether high schools differ in their spending patterns from all schools that are *not* high schools. We present our results in Table 3. We find that, in fact, high schools spend more per-student than elementary schools ($1,669) and much more than middle schools ($2,684). On average, as we show in Model 2, high schools spend over $2,020 more per student than all other schools in Oregon.**

**Table 3**

**Ordinary Least Squares estimates of high-school per-pupil expenditure in Oregon public schools compared to other levels of schooling, 2018-19**

|  | (1) | (2) |
| --- | --- | --- |
| (Intercept) | 15202.02\*\*\* | 13181.49\*\*\* |
|  | (361.90) | (104.63) |
| Early Education | -3628.40 |  |
|  | (2245.60) |  |
| Elementary School | -1669.35\*\*\* |  |
|  | (378.42) |  |
| Middle School | -2684.24\*\*\* |  |
|  | (421.89) |  |
| High School |  | 2020.53\*\*\* |
|  |  | (376.72) |
| Num.Obs. | 1193 | 1193 |
| R2 | 0.053 | 0.043 |
| + p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001. Cells report coefficients and heteroscedastic-robust standard errors in parentheses. Each observation is one school. | | |

3. Do schools that educate a larger proportion of students with financial need spend more per student on their education, after accounting for other school-level variation?

3.1 Review the variables you have at your disposal and select a set of substantively sensible (and non multi-collinear) **continuous** and **categorical** covariates that might explain schools’ per-pupil expenditure and help clarify the relationship between *ppe* and *frpl*. **(1 point)**

**As before, we note in Table 4 that while some of our candidate covariates are highly correlated with each other, they are all only modestly correlated with either our outcome or our primary question predictor. As a result, we are not particularly concerned about problems of multi-collinearity. However, for parsimony, we plan to use the following substantively relevant set of covariates: schooling grade-band level, total enrollment, geographic locale, average socio-economic status of families residing in the district, unemployment rates in the district and four-year degree holding rates in the district.**

**Table 4**

Pearson correlation coefficients of the characteristics of Oregon public schools and the communities in which they are located, 2018-19.

|  | ppe | enroll | sesavgall | baplusavgall | unempavgall | snapavgall | frpl | inc50avgall |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ppe | 1 | . | . | . | . | . | . | . |
| enroll | -.078 | 1 | . | . | . | . | . | . |
| sesavgall | -.053 | .137 | 1 | . | . | . | . | . |
| baplusavgall | .025 | .222 | .694 | 1 | . | . | . | . |
| unempavgall | -.016 | -.050 | -.793 | -.457 | 1 | . | . | . |
| snapavgall | .017 | -.081 | -.957 | -.587 | .754 | 1 | . | . |
| frpl | .184 | -.129 | -.529 | -.419 | .352 | .515 | 1 | . |
| inc50avgall | -.132 | .193 | .887 | .599 | -.579 | -.804 | -.475 | 1 |

3.2 Present a **table** and **plot** characterizing the regression-adjusted relationship between *frpl* and *ppe*. Substantively interpret these results as you would for an academic journal article. **(3 points)**

The proportion of students receiving free- and reduced-price lunch (FRPL) in a school is meaningfully related to the amount the school spends on its students. In particular, after accounting for other school and community characteristics, we estimate that a school that enrolls 10 percentage points more students receiving FRPL could be expected to spend $333.77 more per-student on their education. In Model 1 in Table 5, we show the simple bivariate relationship, which is somewhat more modest ($263.43). Once we adjust for schooling level, locale, enrollment, SES status, unemployment rates and college-degree holding rates, the relationship is even stronger. As we show in Figure 2, one of the largest determinants of total spending is the level of schooling. High schools consistently spend more than elementary and (especially) middle schools, but across all three levels, schools enrolling more students from low-income families spend more, on average.

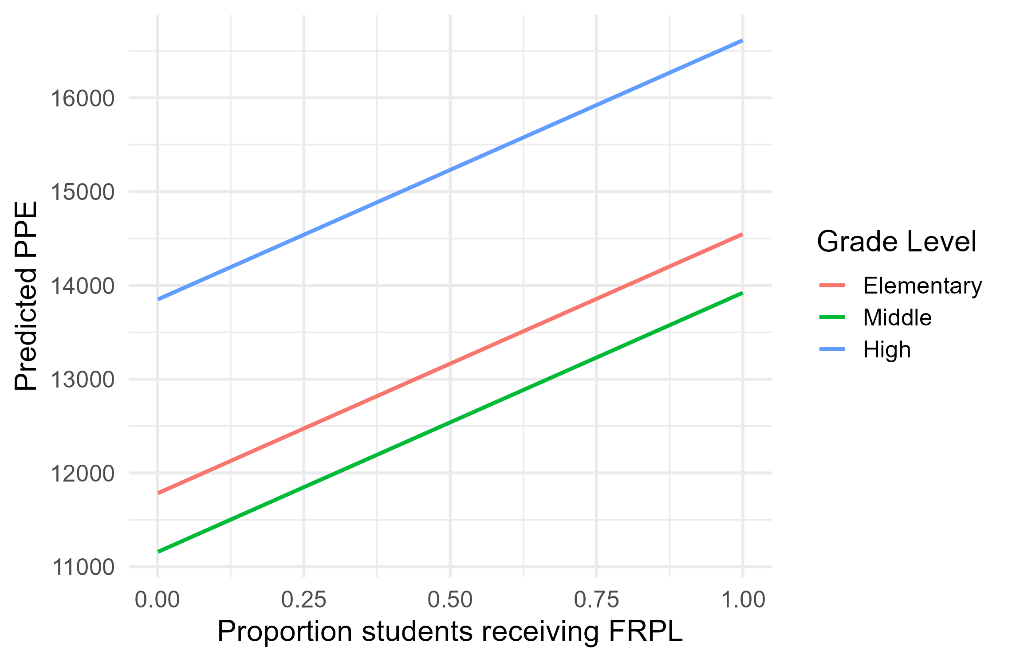
**Table 5**

**Ordinary Least Squares estimates of the relationship between free- and reduced-price lunch enrollment rates and per-pupil expenditure in Oregon public schools, 2018-19**

|  | (1) | (2) |
| --- | --- | --- |
| (Intercept) | 12226.90\*\*\* | 14525.34\*\*\* |
|  | (271.68) | (1399.65) |
| Prop. receiving FRPL (0/1) | 2634.33\*\*\* | 3337.70\*\*\* |
|  | (427.86) | (543.93) |
| **Covariates?** | **No** | **Yes** |
| Num.Obs. | 1193 | 1193 |
| R2 | 0.034 | 0.189 |
| *Notes*: + p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001. Cells report coefficients and heteroscedastic-robust standard errors in parentheses. Each observation is one school. Covariates in Model 2 are grade-band level of school, total enrollment, geographic locale, avg. SES status, unemployment rate, and four-year college degree holding rate. | | |

**Figure 2**

*Relationship between rate of school-level free-and reduced-price lunch (FRPL) receipt and predicted per-pupil expenditure (PPE) at different levels of schooling, adjusting for other school and community factors.*



*Notes:* estimates adjust for total enrollment, geographic locale, avg. SES status, unemployment rate, and four-year college degree holding rate