EDUC 643 Assignment 01 Key

1. Descriptive statistics

1.1. Summarize the dataset. Specifically, create a table to show the means and standard deviations of the continuous variables (i.e., exclude the geographical identifiers and the *level* and *locale* categorical variables). Write 2-3 sentences to report and interpret these statistics. (1 point)

**Table 1**

***Summary statistics for per-pupil expenditures and school and community characteristics for Oregon public schools, 2018-19.***

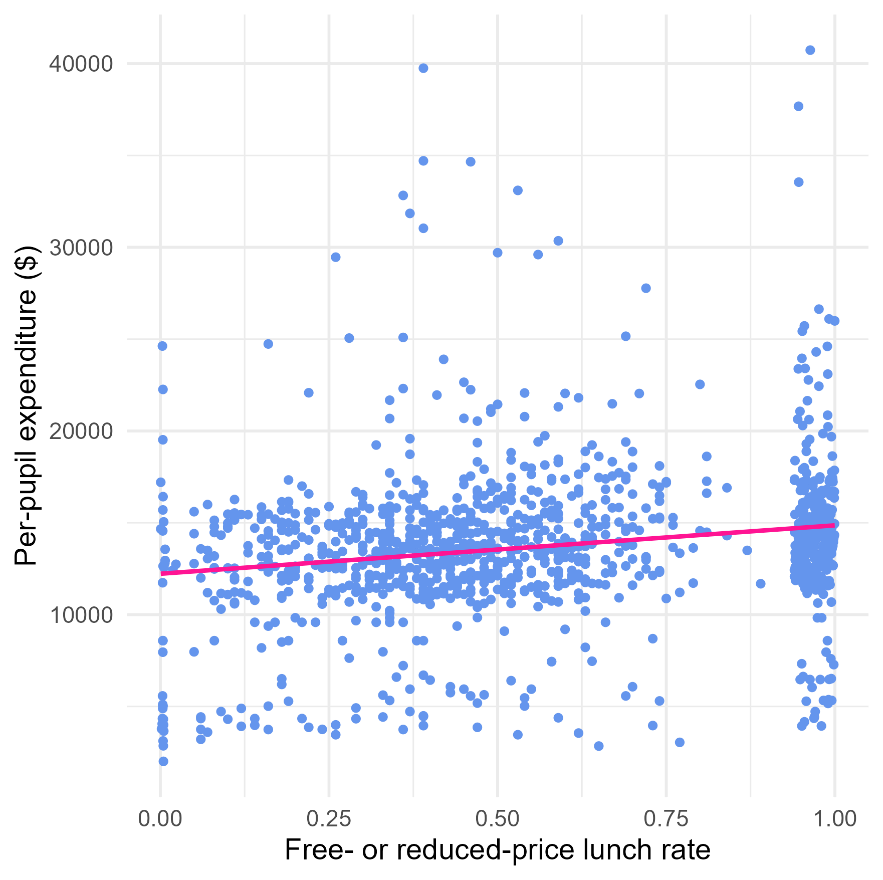
|  | Mean | SD | Min | Median | Max |
| --- | --- | --- | --- | --- | --- |
| Per-pupil expenditure ($) | 13701.44 | 4273.52 | 2011.41 | 13586.98 | 40734.41 |
| Enrollment | 467.28 | 355.60 | 20.70 | 405.30 | 2866.09 |
| SES index for all families | 0.16 | 0.56 | -1.29 | 0.07 | 1.83 |
| Median family income (log $) | 10.87 | 0.20 | 10.26 | 10.83 | 11.65 |
| BA+ rate | 0.27 | 0.13 | 0.08 | 0.24 | 0.67 |
| Unemployment rate | 0.07 | 0.01 | 0.02 | 0.07 | 0.11 |
| SNAP receipt rate | 0.14 | 0.05 | 0.02 | 0.14 | 0.28 |
| Free- or reduced-price lunch rate | 0.56 | 0.30 | 0.00 | 0.51 | 1.00 |
| Source: National Education Resource Database on Schools (NERD$) | | | | | |

There are 1,193 school-level observations in the data; none of the variables have any missing values. As we document in Table 1, with respect to our variables of interest, the mean per-pupil expenditure was $13,701, with a standard deviation of $4,274, suggesting substantial variability in expenditures. The school-level mean of the proportion of students receiving free- or reduced-price lunch (FRPL) was 0.56, with a standard deviation of 0.30, again suggesting a high degree of variability around the mean FRPL rate.

1.2 Create a plot (make sure to label the x and y axis) to visualize the relationship between the variables *ppe* and *frpl*. Include a line of best fit on this display. (1 points)

**Figure 1**

*Relationship between school-level per-pupil expenditure in U.S. dollars and the proportion of students receiving free- or reduced-price lunch for Oregon public schools 2018-19.*



1.3 Write 1-2 sentences to interpret this visualized relationship, relying on the five features of bivariate relationships we introduced in class. (1 point)

In Figure 1, on average, it appears that there is a positive relationship between the proportion of a school’s student body who receive free- or reduced-price lunch, and the school’s per-pupil expenditures. This relationship is roughly linear and fairly weak in both magnitude and strength with seemingly substantial variation away from the line of best fit throughout the distribution. There are a few potentially outlying values of schools that spend considerably more per-student than others.

#### 2. Research question

2.1. Write a formal linear model that describes the simple, bivariate relationship between school-level per-pupil expenditure and the school-level average receipt of free- or reduced-price lunch and interpret each of the terms in this model. (1 point)

We postulate the following linear model to characterize the relationship between school-level per pupil expenditure (*PPE*) as a function of the average receipt of free- or reduced-price lunch (*FRPL*) at the school:

where *PPE* for the *i*th school is a function of an intercept (*β0*), which will represent the value of *PPE* when *FRPL* is equal to 0. Our parameter of interest is *β1*, which represents the slope of the relationship between *FRPL* and *PPE*. Finally, *εi* is an idiosyncratic school-level error term, which represents the variance in PPE unexplained by our model.

2.2. State your null hypothesis about the relationship between per-pupil expenditure and the proportion of students receiving free- or reduced-price lunch. (1 point)

There is no relationship between per-pupil expenditure and the proportion of students receiving free- or reduced-price lunch in school, on average in the population of Oregon public schools.

2.3 Formally test your hypothesis using an Ordinary Least Squares estimation strategy. Report your results in a formatted table. (2 points)

**Table 2**

***Ordinary Least Squares estimates of the relationship between school-level per-pupil expenditures and free- and reduced price lunch rates for Oregon public schools, 2018-19***

|  | Model 1 |
| --- | --- |
| (Intercept) | 12226.90\*\*\* |
|  | (258.38) |
| Free- or reduced-price lunch (0-1) | 2634.33\*\*\* |
|  | (407.23) |
| Num.Obs. | 1193 |
| R2 | 0.03 |
| RMSE | 4198.60 |
| *Notes*: + *p* < 0.1, \* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001. Cells report coefficients and standard errors in parentheses. | |

2.4. Interpret the results of your test in 1-2 sentences. (2 points)

We estimate that an Oregon school in which all students (100 percent) receive free- or reduced-price would spend, on average, $2,634.33 more per-student than a school in which no students (0 percent) receive free- or reduced-price lunch. At an alpha-threshold of 0.05, we reject the null and conclude that, on average in the population of Oregon public schools, there is a relationship between FRPL rates and per-pupil expenditures.

2.5. Assess the quality of your model fit using the *R2* statistic. (1 point)

The *R2* for our model is 0.03, which means that our model explains 3 percent of the overall variation in per-pupil expenditure. This seems like a relatively small proportion of the overall variance and suggests that many other factors contribute to variation in *PPE*.

2.6. Select an alpha-threshold and describe the corresponding confidence interval for your estimates of the relationship between *frpl* and *ppe*. (1 point)

With an alpha-threshold of 0.05, our corresponding 95 percent confidence interval for the relationship between *frpl* and *ppe* is $1,835.36 - $3,433.30. This implies that were we to repeat the exercise of measuring the proportion of students receiving FRPL and school PPE—as well as the relationship between the two—for the population of Oregon schools 100 times, that in 95 cases we would expect our estimate of this relationship to fall between $1,835 and 3,433.

2.7. Imagine you are an analyst working for Colt Gill, Director of the Oregon Department of Education. You have been tasked with describing to him whether schools with greater levels of student financial need receive more money. Write a short paragraph reporting the results of your analysis, while introducing appropriate caveats and nuances as needed. In particular, you should think about how you want to introduce the ideas of relationship magnitude, model fit, omitted variables, and correlation for a lay audience. (1 point)

Using publicly available data from the National Education Resource Database on Schools, we found that public schools in Oregon who educate more children from low-income families spend *more* per-child than schools educating fewer children from low-income families. In particular, we estimate that schools have 10 percentage points more free- and reduced-price lunch recipients spend, on average, $263 more per child per year. Given that the median school in Oregon enrolls around 405 students, this represents an annual budgetary difference of more than $106,700 for school educating 10 percentage points more FRPL-receiving students. While this seems like a meaningful difference, there is little concrete evidence suggesting just how much more compensatory funds would be required to promote a more equitable schooling system for students living in poverty. Further, there is substantial variation in per-pupil expenditure that is not explained by our model—some schools with very few low family-income students spend close to $25,000 per-student, per-year while some schools that educate nearly exclusively low family-income students spend less than $5,000 per-student, per-year. It may be valuable to investigate further the role that school size, school location, school grade-level, community resources, and local governance play in determining school funding. While there is some evidence of socio-economic-based equity funding in our analysis, if the goal of the Oregon Department of Education is to systematically provide more resources to schools that educate more students living in poverty, our analysis also shows that there is substantial scope to strengthen and increase the magnitude of this relationship.