School District Analysis Project

Hemanth Pothineni

Northern Illinois University

OMIS 645 : Applied Business Analytics (SAS)

Kishen Iyengar



# **Abstract**

A school district report based on a SAS project using multiple analyses, such are simple linear regression, factor analysis, and correlation analysis. Specifically, this report presents recommendations based on statistical test reports. These reviews are based on research questions I have done on school datasets. Collectively, these unique questions provide a definitive analysis of the school district's dataset.

# **Summary**

This research analyzes teacher retention, teacher evaluation, attendance, and student enrolment using a school district dataset and several statistical approaches. I responded to a study query concerning the association between academic achievement in mathematics. Chemistry. The primary determinants influencing teacher retention were identified using factor analysis. The investigation indicated three major factors: total teacher employment and male teacher percentage, average teacher experience and teacher evaluation rate, and teacher whiteness proportion and teacher attendance rate.

The association between teacher evaluation rates and teacher attendance rates was investigated using simple linear regression while adjusting for other factors. The analysis revealed a substantial positive association between these factors. Correlation analysis was performed to investigate how student enrollment, attendance, and movement rates linked with academic achievement in ELA, mathematics, and science. The findings revealed a strong association between student attendance and academic achievement, but no significant correlation between student enrolment and transfer rates.

Overall, this effort sheds light on the factors that drive teacher retention, the link between teacher evaluations and attendance, and the association between student attendance and academic success in the school district dataset.

# **Factor Analysis**

What are the principle factors that effect the teacher’s retention rate?

Factor analysis is used to solve research questions according to the school district data set. The purpose of this study is to examine how different factors interact with each other. An attempt is then made to explain the factors found in terms of their common fundamental dimensions. That is, find all principal components in the dataset and return a new dataset. An introduction to the current school district data set. After the data is successfully imported into SAS, the dataset is read into files and folders on the server. After selecting the dataset, a spreadsheet will appear on the screen. Selecting a dataset expands the tasks and utilities on the left side of the screen. Go to Tasks for this step, edit Multivariate Analysis again, and select Factor Analysis. The new tab that opens in the SAS screen for this data selects the school district data set and takes analysis variables. In Analysis Variables, factor analysis only uses numerical variables because it can only perform numerical analyzes and is designed to use only the variables of interest. We used total school days, total teacher FTE, % teachers-white, % teachers-male, % teachers. – Women, average teacher experience, teacher attendance rate, and teacher evaluation rate as variables. Open the Options tab, set the rotation method as oblique varimax rotation, select standard and additional plots under Plots, check the path graph, and run the report.

Table

Description automatically generated

Examine the eigenvalues table to see which components are greater than 1. And because I obtained three eigenvalues larger than one, I set the number of factors to three in the settings tab and ran the report. Examine the path diagram; if the two factors connect to the same variable, the factor analysis is incorrect, and my %teachers-female & Total Number of School Days have comparable components. So I deleted those tables and ran the report again, and now I have the exact route diagram that I need.

Factor 1 : Total Teacher FTE, %teacher male  
Factor 2 : Avg. Teaching Experience, Teacher Evaluation Rate

Factor 3: % teacher white, Teacher Attendance Rate

# **Simple Linear Regression**

Is there a substantial association between teacher evaluation rate and teacher attendance rate when other variables for relevant aspects are controlled for?

According to my research question, I choose simple linear regression to infer the results of this research question. So, under Tasks & Utilities, select Tasks, Expand Linear Model, and select Linear Regression. Select the school district data set as the data. Linear regression requires a dependent variable, a continuous variable, and a classification variable. A dependent variable cannot have more than one missing variable. If you get this error, you will not be able to run your report.

To do this, we used the teacher evaluation rate as the dependent variable, the major turnover rate within 6 years as the categorical variable, and the numeric variable total teacher FTE, mean. Teaching experience, teacher retention, and teacher attendance as numerical variables for this question. On the Model tab, under Intersect, select Edit, add the extracted numeric data set, and click OK to run the report. Check the analysis of variance table which shows that the p-value must be <0.001. Otherwise the simple linear regression analysis is wrong and the numerical analysis needs to be changed. See chart for clear results of research questions.

# **Correlation Analysis**

How do student enrollment, attendance, and mobility rates correlate with academic achievement in ELA, math, and science?

According to the question I need to choose the correlation analysis because we need to do the statistical methods to correlate the required variables with the academic achievements in the subjects given in the dataset. Select the tasks and utilities, now choose the tasks and elaborate the statistics, and select the correlation analysis.

In the data tab select the school district dataset as the data and for the analysis variables I took student enrollment, student attendance rate, and student mobility rates and I correlate them with the %ELA proficiency, %math proficiency, and %science proficiency because of the direct question we can select the variables like this and these correlate with each other.

In the options tab for the statistics option select the selected statistics and check the display p-values option for that and click on the run after running the report we got the result as the Pearson correlation coefficients table in that we can observe the correlations between the taken analysis variables and correlation variables.

# **Conclusion**

I conclude that According to factor analysis, the three primary factors influencing teacher retention are: 1) teacher experience and teacher evaluation rate; 2) teacher knowledge rate and teacher attendance rate; and 3) teacher knowledge rate and teacher attendance rate. After controlling for the primary 6-year turnover rate and other pertinent numerical factors, a simple linear regression analysis revealed a strong correlation between teacher evaluation rates and teacher attendance rates. Correlation analyses discovered relatively favorable relationships between school enrollment, attendance, and mobility, as well as academic achievement in ELA, mathematics, and science. This shows that higher rates of enrollment, attendance, and mobility are related to higher levels of academic success in these courses.

# **Appendix**

**Factor Analysis**

A screenshot of a computer

Description automatically generated with medium confidence

Diagram

Description automatically generated Diagram

Description automatically generated

**Final Run Initial Run**

**Simple Linear Regression**

**Table

Description automatically generated**

**Table

Description automatically generated**

**Correlation Analysis**

**Table

Description automatically generated**