INVESTIGATING THE RELATIONSHIP BETWEEN AP TEST SCORES AND AP COURSE GRADES THROUGH CORRELATION AND MACHINE LEARNING MODELS

DIANE NICOLL

DSE 200X

ABSTRACT

A dataset of 605 AP students from one high school from the 2018-2019 academic year was created for this investigation from the schools' Student Information System. The research questions were: "Are AP test scores a reflection of subject mastery or are they a reflection of test taking skills and overall academic performance?", and "Can a regression analysis model predict AP test scores given AP course grades or are SAT scores and GPA needed as well?". The methods used in this investigation included correlation analysis, scatterplots, and machine learning regression analysis modeling. This investigation provided evidence that AP test scores reflect a combination of subject mastery, test taking skills, and overall academic performance, based on correlation analysis and scatterplot visualization. A regression analysis model was able to predict AP test scores with slightly better prediction accuracy when provided AP course grades, SAT scores, and GPA as compared to only being provided AP course grades as the input, providing additional support for AP test scores being a reflection of subject mastery, test taking skills, and overall academic performance.

MOTIVATION

The motivation to investigate the relationship between AP test scores and AP course grades comes from a curiosity to understand whether AP test scores reflect subject matter content knowledge (likely to be correlated with AP course grades) or an ability to take standardized tests (likely to be correlated with SAT scores). Another possibility considered was that AP test scores may be related to overall education and knowledge, so transcript weighted academic GPA also was considered.

It is important to understand if AP tests truly reflect subject matter knowledge as many colleges and universities are no longer accepting passing scores on AP tests (a score of 3 or higher out of 5) as college credit for that course or subject matter, while some colleges are only accepting a high passing score (a score of 4 or 5). Insights gained from this investigation would be valuable to high schools deciding to offer additional AP courses, students and families deciding whether to take AP courses, and colleges and universities deciding whether or not to accept passing AP test scores for college credit.

DATASET

The dataset analyzed for this project contains data for students who enrolled in an AP course in the 2018-2019 academic year at one high school. This year was selected as AP test scores have not yet been released for the 2019-2020 academic year. I created the dataset from the high school Student Information System (SIS) where I work. All data is publicly available on an aggregate level through the school district, Department of Education, and College Board websites. All personal identifiers have been removed from the dataset for privacy. The SIS is a non-SQL database, but accepts its own type of queries. I queried the 2018-2019 SIS database for all students who met the following criteria:

- Took an AP course that academic year
- Took the associated AP test

I queried the SIS to acquire each students' GPA and SAT scores (where applicable) and merged the datasets in Excel. I saved this file as a CSV to import into Jupyter Notebook. The final dataset contains the following variables:

- Student ID (re-assigned to a generic number for privacy), student name (re-assigned to a generic name for privacy), grade, gender, AP course taken, AP course grade for each trimester (letter grade and numeric equivalent), an average of the numeric course grades, AP test score, transcript academic weighted GPA, transcript total weighted GPA, and SAT score.

The original dataset contained 716 samples; after dropping rows with null values (all missing SAT scores), the final dataset contained 605 samples.

DATA PREPARATION AND CLEANING

To prepare the data for analysis, a dataset containing student information, AP course grades, and AP test scores was merged with another dataset containing GPA and SAT scores.

Additional columns were created for numeric-based AP course grades to accommodate the regression analysis that was planned (i.e. a letter grade of an “A” was set to a 5.0 numeric-based grade as AP course grades are on a weighted scale out of 5.0). These numeric equivalent trimester AP course grades were used to create an average course grade for each student (average across all three trimesters of the course).

Once the dataset was created and prepared, null values were dropped to finalize the cleaning process.

RESEARCH QUESTIONS

Are AP test scores a reflection of subject mastery, or are they a reflection of test taking skills and overall academic performance?

Can a regression analysis model predict AP test scores given AP course grades, or are SAT scores and GPA needed as well?

METHODS

A correlation analysis was run to determine the relationship (strength and direction of the correlation) between the variables. This is appropriate as the data of interest is numeric and this analysis is essential to answer the first research question.

Scatterplots were conducted to visualize the relationships between average AP course grade and AP test score, transcript academic weighted GPA and AP test score, and SAT score and AP test score. These visualizations are appropriate to further understand the relationships between these variables of interest and answer the first research question.

Note: Transcript academic weighted GPA was the GPA used as this investigation was focused on academic measures; therefore, transcript total weighted GPA was not used.

METHODS (CONT.)

Regression models were built, trained, and tested on the dataset to determine if a machine learning regression model could learn and predict AP test scores (the target value) based on two different sets of inputs:

- In the first regression model, the inputs were the trimester numeric AP course grades and average AP course grade
- In the second regression model, the inputs were the trimester numeric AP course grades, average AP course grade, transcript academic weighted GPA, and SAT score

Both regression models used two different modeling techniques to analyze whether a simple or more complex regressor model provided better prediction accuracy (as measured by root mean squared error):

- linear regression model
- Decision Tree regression model

Regression models are appropriate for this dataset as they take numeric inputs and have numeric targets, which the dataset contained. This is a supervised task, and the dataset contains all targets. This analysis is essential to answer the second research question.

FINDINGS

Pearson correlation:

	Average_Course_Grade	AP_Test_Score	Transcript_Academic_ Weighted_GPA	Transcript_Total_ Weighted_GPA	SAT_Score
Average_Course_Grade	1	0.63458	0.701911	0.699116	0.520574
AP_Test_Score	0.63458	1	0.560577	0.559055	0.561424
Transcript_Academic_ Weighted_GPA	0.701911	0.560577	1	0.999288	0.751918
Transcript_Total_ Weighted_GPA	0.699116	0.559055	0.999288	1	0.751575
SAT_Score	0.520574	0.561424	0.751918	0.751575	1

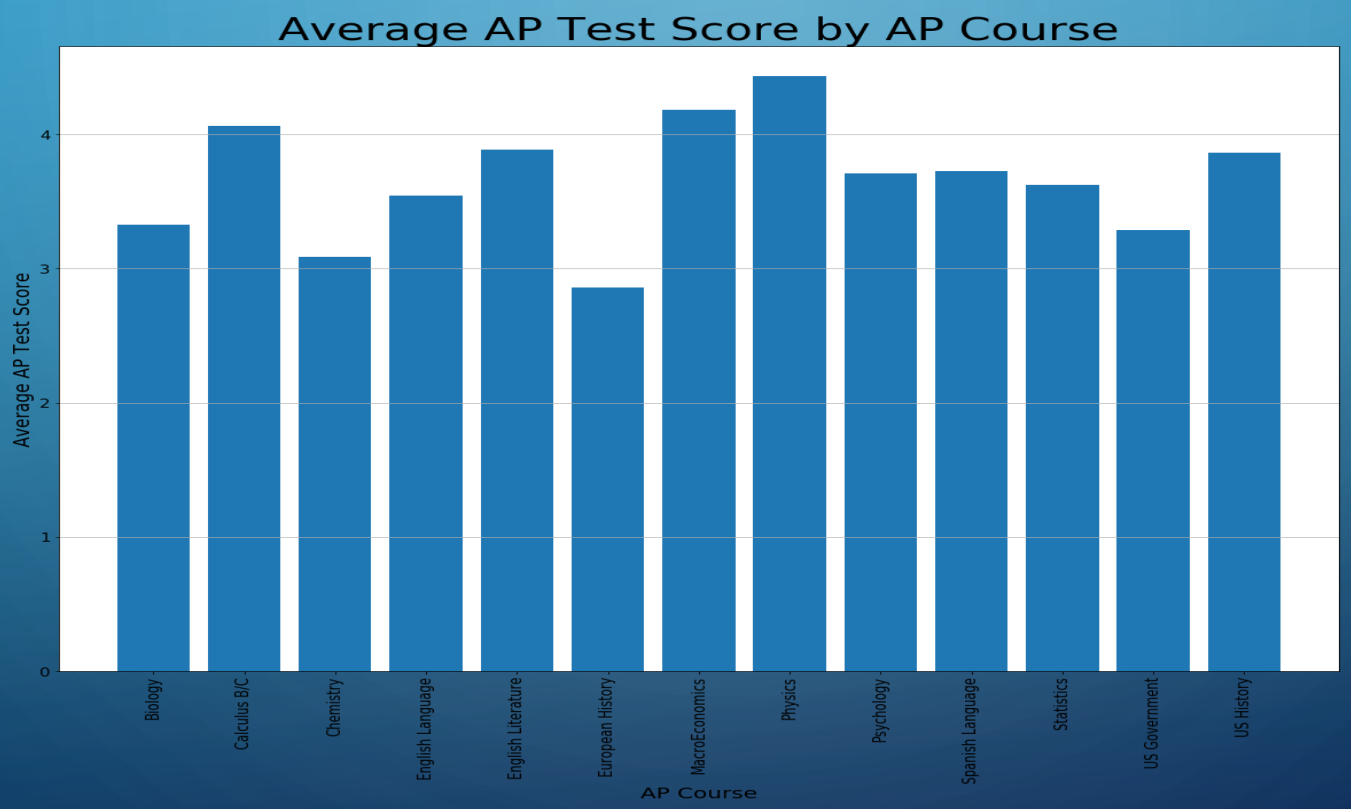
The data indicate there is a positive, moderately strong correlation between:

- Average AP course grade and AP test score ($r = 0.63458$)
- Transcript academic and total weighted GPAs and AP test score ($r = 0.560577$ and $r = 0.559055$, respectively)
- SAT score and AP test score ($r = 0.561424$)

Although the correlation coefficient for average AP course grade and AP test score is slightly higher than seen for GPA and AP test score or SAT score and AP test score, it appears that AP test scores reflect a combination of subject mastery, test taking skill, and overall academic performance as the difference between correlation coefficients was quite small.

Thus, the data indicate that there are positive relationships between average AP course grade, weighted GPA, SAT score and AP test score; as one variable increases, the other variable increases (i.e. as average AP course grade increases, AP test score increases).

FINDINGS (CONT.)



FINDINGS (CONT.)

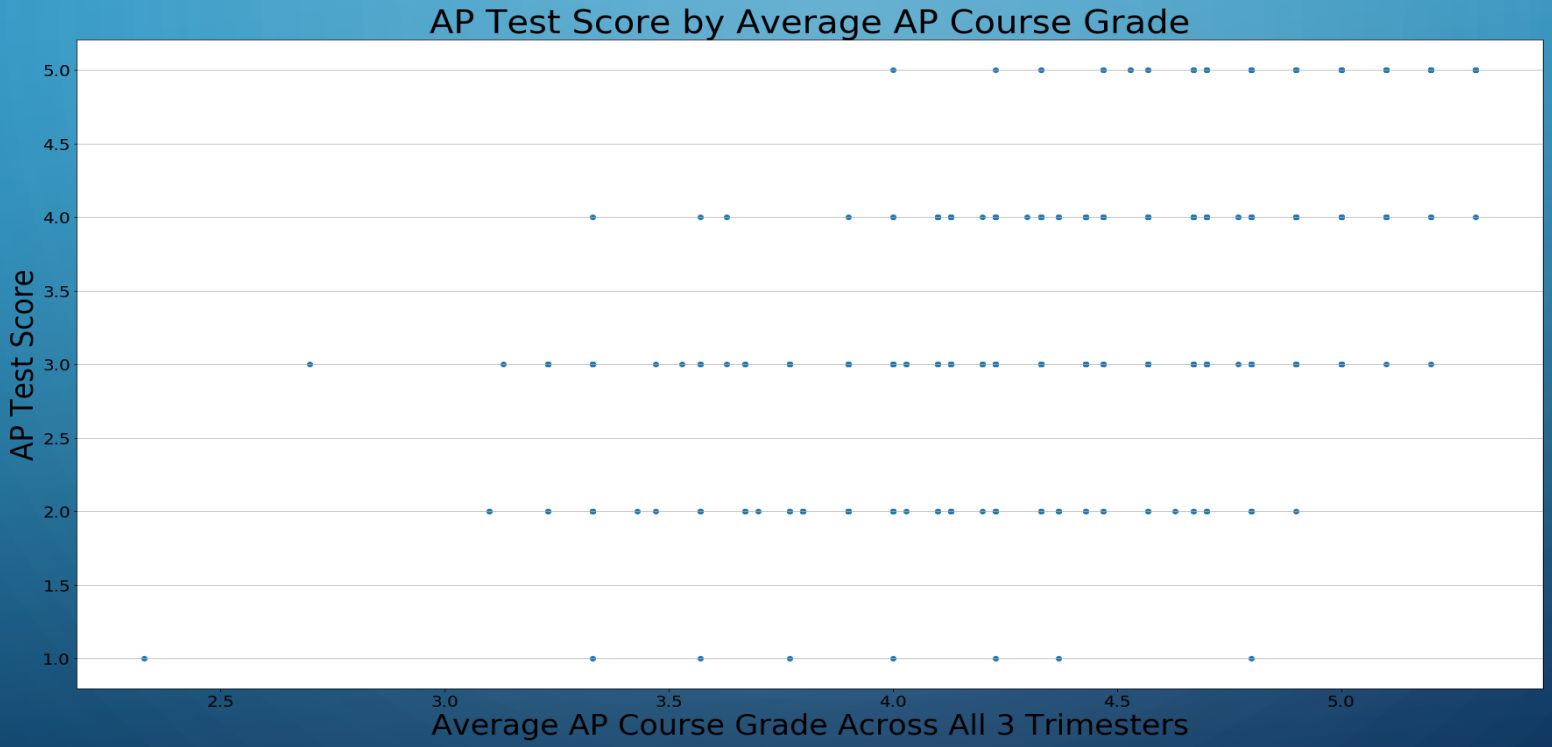
Histogram of average AP test score by AP course:

- The AP courses where students earned the highest average AP test scores were senior-level (12th grade) courses: Physics, MacroEconomics, and Calculus B/C

- The AP course where students earned the lowest average AP test score (and the only course with average AP test scores below "3" i.e. not passing) was European History, a 10th grade course

- These findings may indicate AP test scores reflect test taking skills, as 12th grade students could have taken AP courses and tests for three years, in addition to likely having taken the SAT. Alternatively, these findings could reflect differences in student understanding of course material, differences in AP tests for each specific subject, differences in teacher/course preparation for the AP test, or other factors not measured in this investigation. Further research into trends and patterns of differences between average AP test score by AP course would be needed to clarify.

FINDINGS (CONT.)

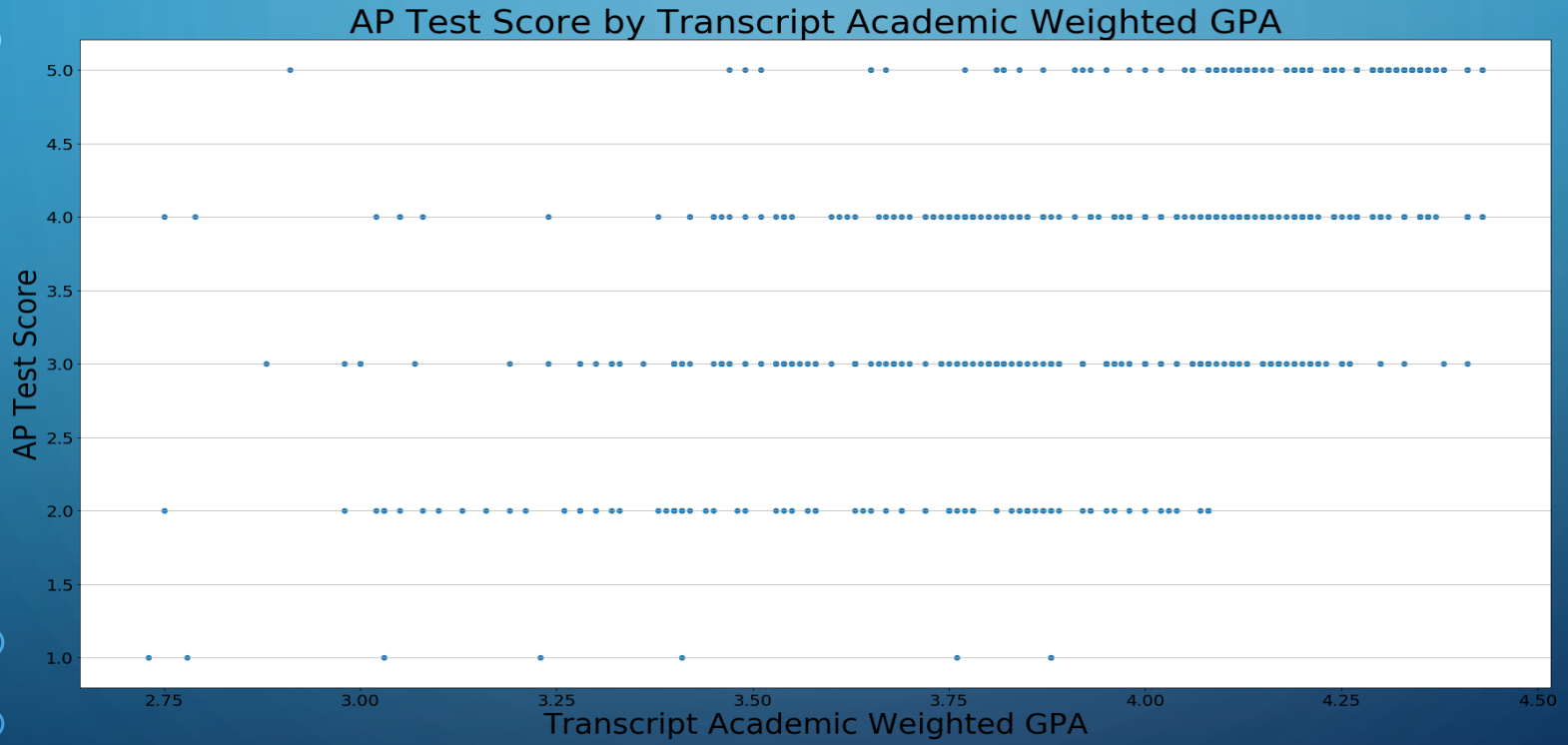


FINDINGS (CONT.)

Scatterplot of AP test score by average AP course grade:

- Students earning the highest scores on the AP test (score of 4 or 5) earned higher average AP course grades (i.e. students earning a 5 on the AP test earned average grades of "B" or better (4.0 out of 5.0 on a weighted scale))
- Students earning passing scores on the AP test (score of 3+) are associated with higher average AP course grades as compared to students earning failing scores (score of 1 or 2)
- Students earning a 1 on the AP test showed the lowest high-end average AP course grades ("A-" or 4.7 on a weighted scale), but they did not earn the lowest overall average AP course grades
- These findings indicate a relationship between AP test scores and AP course grades, and thus support AP test scores being a reflection of subject mastery

FINDINGS (CONT.)

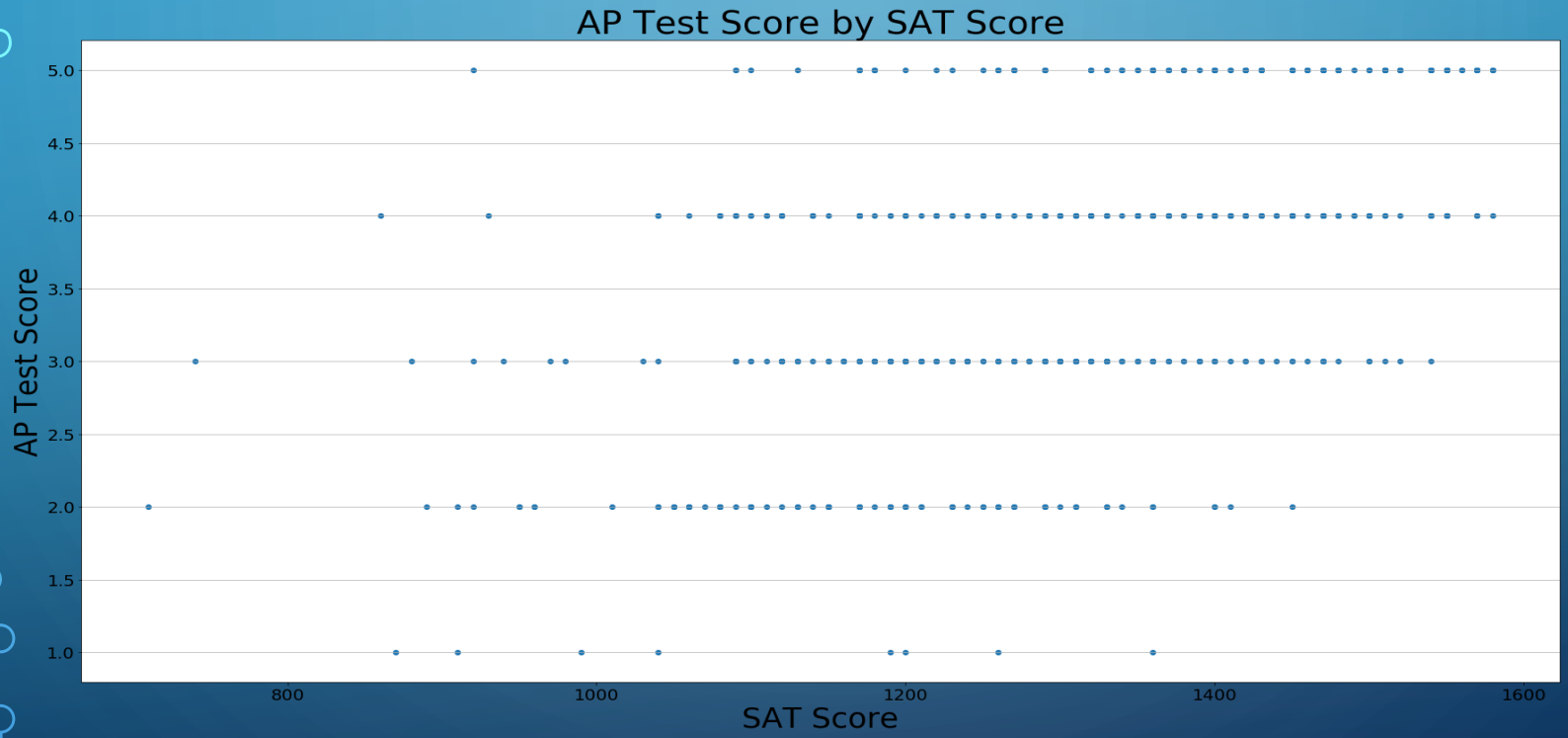


FINDINGS (CONT.)

Scatterplot of AP test score by transcript academic weighted GPA:

- Focusing on the densest plot ranges for each AP test score, higher AP test scores are overall associated with higher transcript academic weighted GPAs:
  - AP score of 5 is associated with the most density of GPA values between ~3.75-4.40
  - AP score of 4 is associated with the most density of GPA values between ~3.40-4.40
  - AP score of 3 is associated with the most density of GPA values between ~3.25-4.30
  - AP score of 2 is associated with the most density of GPA values between ~3.00-4.10
  - AP score of 1 is associated with the most density of GPA values between ~2.70-3.85
- These data indicate there is a relationship between AP test scores and GPA, and thus support AP test scores being a reflection of overall academic performance

FINDINGS (CONT.)



FINDINGS (CONT.)

Scatterplot of AP test score by SAT score:

- Students earning the highest scores on the AP test (score of 4 or 5) earned higher SAT scores as compared to students earning lower AP test scores
- Differences between AP test score groups are most defined for high-end SAT scores (i.e. students who earned a 4 or 5 on the AP test earned SAT scores as high as ~1580, whereas students who earned a 2 or 3 on the AP test earned SAT scores as high as ~1420 and 1500, respectively)
- SAT scores overlapped in the middle and lower-end ranges for all AP test scores
- These findings indicate a relationship between AP test scores and SAT scores, and thus support AP test scores being a reflection of test taking skill

FINDINGS (CONT.)

Machine learning models: regression analysis

The first regression analysis model tested the relationship between AP course grades and AP test score

- Inputs: Trimester 1 AP course grade, Trimester 2 AP course grade, Trimester 3 AP course grade, average AP course grade
-Target: AP test score

Results:

Linear regressor model RMSE = 0.755

Decision Tree regressor model RMSE = 0.805

The regression analysis model was able to learn and predict AP test scores from AP course grades with better accuracy using the linear regressor model compared to the Decision Tree regressor model

FINDINGS (CONT.)

The second regression analysis model tested the relationship between AP course grades, GPA, and SAT scores and AP test score

- Inputs: Trimester 1 AP course grade, Trimester 2 AP course grade, Trimester 3 AP course grade, average AP course grade, transcript academic weighted GPA, SAT score

-Target: AP test score

Results:

Linear regressor model RMSE = 0.727

Decision Tree regressor model RMSE = 1.033

The regression analysis model was able to learn and predict AP test scores from AP course grades, GPA, and SAT scores with better accuracy using the linear regressor model compared to the Decision Tree regressor model. This second model (with additional inputs of GPA and SAT score), using the linear regressor model, provided the lowest RMSE, and thus best prediction accuracy overall.

These findings indicate AP test scores reflect a combination of subject mastery, test taking skill, and overall academic performance.

LIMITATIONS

The limitations of this investigation include a relatively small sample size (n = 605) taken from one time period (2018-2019 academic year) from one high school. Therefore, these findings may not be representative of, or generalizable to, all high school students across the globe.

Additional research into the relationship between AP test scores and AP course grades using larger sample sizes, longer time periods, and samples from various locations would be beneficial to understand if these results can be replicated in other states and countries.

CONCLUSIONS

Research question: Are AP test scores a reflection of subject mastery, or are they a reflection of test taking skills and overall academic performance?

This investigation provided evidence that AP test scores reflect a combination of subject mastery, test taking skills, and overall academic performance through correlation analysis, scatterplot visualization, and machine learning regression analysis models. These findings provide initial support for continued acceptance of passing AP test scores by colleges and universities for course or subject credit, as they indicate a positive, moderately strong correlation between average AP course grade and AP test score.

High schools could benefit from the insight that AP test scores were positively correlated with three markers of educational success—course grades, GPA, and SAT scores—to continue to offer current AP courses and possibly consider additional AP courses as needed.

Students and families can use this insight to continue to enroll in AP courses as an opportunity to challenge themselves academically and potentially receive college credit from a passing AP test score.

CONCLUSIONS (CONT.)

Research question: Can a regression analysis model predict AP test scores given AP course grades, or are SAT scores and GPA needed as well?

A regression analysis model was able to predict AP test scores with slightly better prediction accuracy when provided AP course grades, SAT scores, and GPA as compared to only being provided AP course grades as the input.

Better prediction accuracy was found using a linear regressor model as compared to a Decision Tree regressor model, suggesting this data is linear in nature, or does not require a more complex model for prediction accuracy.

ACKNOWLEDGEMENTS AND REFERENCES

The database used for this investigation was created through Aeries Student Information System. I queried all data from the SIS. I did not receive any feedback on this project.

References:

Aeries Student Information System was used to gather data for this project