Decoding Reading Performance: Findings from the PISA Dataset

Meixian Wu

What is PISA and Why Does It Matter?

- **PISA Dataset**: Assesses 15-year-old students' performance globally in reading, math, and science. Includes data on demographics, parental background, school characteristics, and behaviors.
- **Global Importance**: It provides critical data on educational systems' effectiveness, allowing countries to benchmark and improve student outcomes.
- **Objective**: Identify key factors influencing reading performance to uncover what drives student success.
- **Focus**: Provide actionable insights for educators and policymakers to enhance literacy and academic outcomes.

How We Prepared the Data

Handling Missing Values:

- For numeric columns (e.g., `int64`, `float64`), missing values were filled using the median.
- For categorical variable `raceeth`, missing values were filled using the most frequent value (mode).

**Transforming Categorical Variables**:

- Applied **one-hot encoding** to the `raceeth` column to convert it into numerical format, creating binary columns for each category.

Final Dataset:

- After cleaning, all missing values were handled, and categorical variables were transformed. The dataset was ready for further analysis with no missing values remaining.
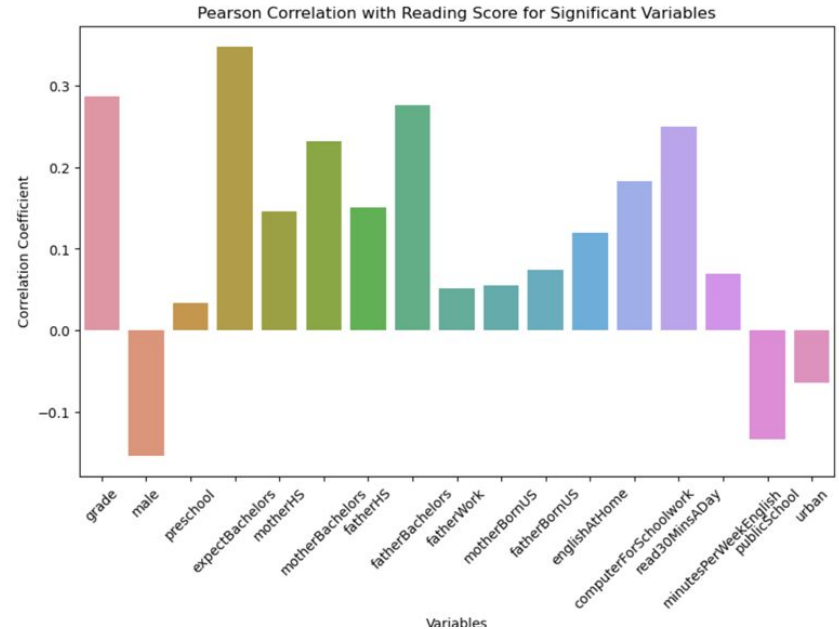
Correlation matrix & bar chart

- Computed the correlation matrix and created a **heatmap** to visualize relationships between variables.

Correlation Matrix Heatmap
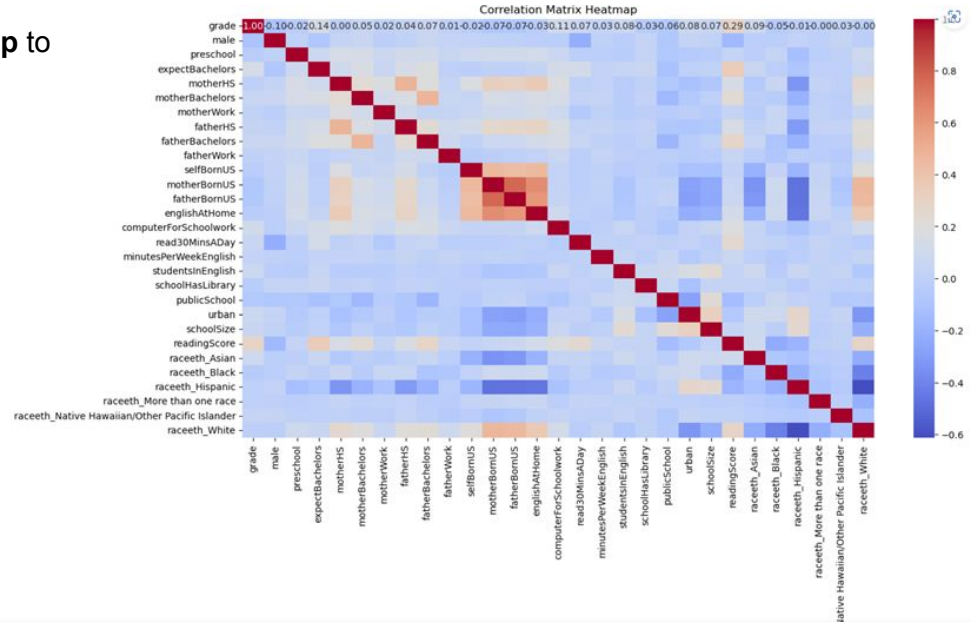
Correlation matrix & bar chart

- Conducted **Pearson correlation analysis** to identify significant predictors of reading scores.

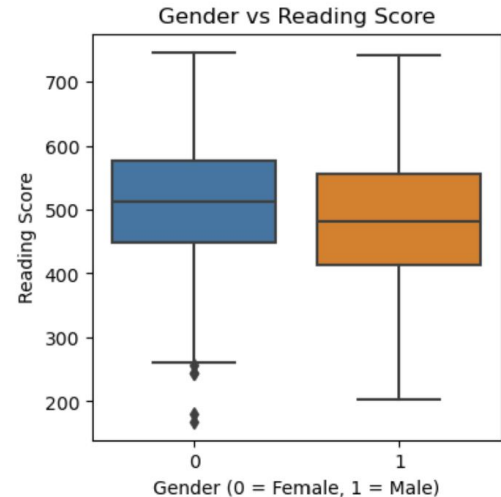Pearson Correlation with Reading Score for Significant Variables

Gender and Reading Performance: What Did We Find?

T-Test Findings: Gender & Reading Scores

Significant Results:

- **T-statistic**: -9.43
- **P-value**: 7.33e-21 (highly significant)
- **Conclusion**: Female students significantly outperform male students in reading scores
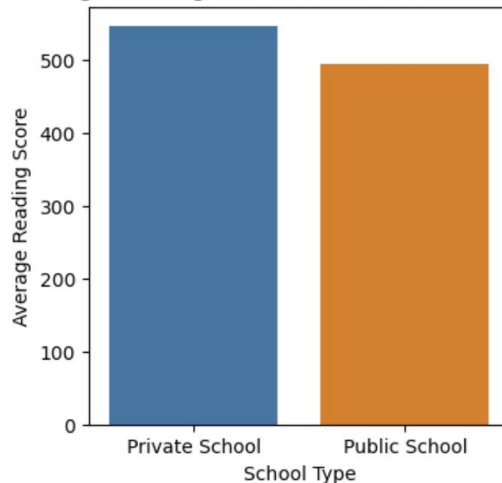

Gender vs Reading Score

Does School Type Matter? Public vs Private Schools

ANOVA Findings: School Type & Reading Scores

Significant Results:

- **F-statistic**: 66.59
- **P-value**: 4.56e-16 (highly significant)
- **Conclusion**: Private school students outperform public school students in reading scores


Average Reading Scores: Public vs Private Schools

What Factors Predict Reading Performance?

Regression Model Output:

- **Coef**: Positive → predictor increases, dependent variable increases.
- **Std Err**: Smaller = more precise estimate.
- **T-value**: |t| > 2 → significant effect on the outcome.
- **P-value**:
  - **p < 0.05** → significant impact.
  - **p > 0.05** → no significant impact.
- **95% CI**:
  - Includes 0 → not significant.
  - Excludes 0 → significant.

Regression Coefficients
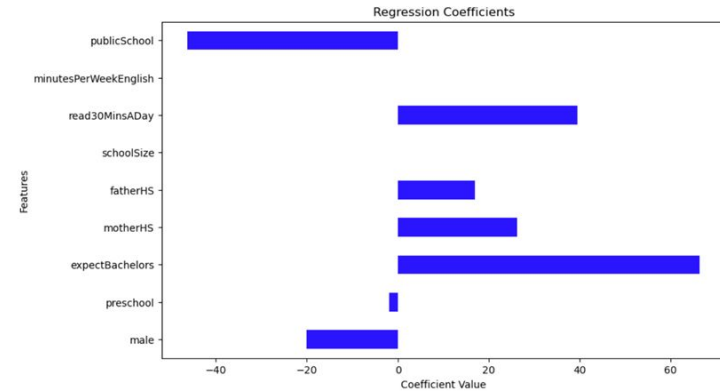
	coef	std err	t	P>|t|	[0.025	0.975]
const	429.7240	8.631	49.788	0.000	412.802	446.646
male	-20.0841	2.905	-6.914	0.000	-25.779	-14.389
preschool	-1.9164	3.208	-0.597	0.550	-8.205	4.373
expectBachelors	66.5094	3.557	18.699	0.000	59.536	73.483
motherHS	26.2140	5.024	5.218	0.000	16.364	36.064
fatherHS	16.8967	4.786	3.530	0.000	7.513	26.280
schoolSize	0.0075	0.002	4.364	0.000	0.004	0.011
read30MinsADay	39.5284	3.213	12.302	0.000	33.229	45.828

Key Findings from the Analysis

Gender:

- Male students score significantly lower in reading.
- **Significance**: Aligns with research on female outperformance in literacy.

Public vs. Private School:

- Private school students outperform public school students.
- **Significance**: Supports research but raises concerns about resource inequality.

Parental Education:

- Higher parental education, especially maternal, strongly boosts reading scores.
- **Significance**: Confirms the role of parental background in academic success.

Study Habits:

- More time spent reading leads to better performance.
- **Significance**: Reinforces the value of consistent reading habits.

- **Alignment**:Findings such as gender differences, parental education, and study habits strongly align with existing educational research.

- **Challenge**: The significant impact of private schooling over public schooling may challenge assumptions about equal educational opportunities, suggesting a need for more focus on public school resources.

Moving Forward: What Can Be Done?

- **Actionable Steps**:
  - Implement reading programs targeting boys.
  - Promote parental engagement through educational workshops.
  - Invest in public school improvements, focusing on teacher training and resource allocation.
- **Call to Action**:
  - Educational policymakers and school administrators should adopt these strategies to improve reading performance across different student demographics.