

TECHNICAL REPORTS: Student Performance Analysis Using Microsoft Excel

Outline

This technical report provides a structured breakdown of the student performance analysis project carried out using Microsoft Excel.

It covers:

1. Introduction
2. Story of Data
3. Data Splitting and Preprocessing
4. Pre-Analysis
5. In-Analysis
6. Post-Analysis and Insights
7. Data Visualizations & Charts
8. Recommendations and Observations
9. Conclusion
10. References & Appendices

Introduction

Objective of the Project

The goal of this project is to analyze the key factors influencing student performance using a dataset of 6,607 students. The analysis focuses on variables such as study hours, attendance, parental involvement, teacher quality, resource access, and family income to understand how they relate to exam outcomes.

Problem Being Addressed

Student performance is shaped by both academic and socio-economic factors. Schools and parents need to understand which variables contribute the most to higher scores so they can design better interventions. This analysis aims to uncover the strongest performance drivers.

Key Datasets and Methodologies

Dataset:

- 6,607 students records
- Variables include study habits, attendance, income level, disabilities, school type, peer influence, and exam scores.

Excel Techniques Used:

- Pivot Tables
- Pivot Charts
- Slicers
- Cards
- Data Cleaning and Label Standardization

Story of Data

Data Source

The dataset originates from an education performance simulation used for academic analytics, downloaded from *Kaggle*.

Data Collection Process

Data is assumed to be collected from school records, learning assessments, and student/parent surveys.

Data Structure

Each row represents a student. Main columns include:

- Hours Studied
- Attendance
- Parental Involvement
- Teacher Quality
- Access_to_Resources
- Family Income
- Peer Influence
- Exam Score
- Gender
- School Type

Important Features & Their Significance

- **Hours Studied:** Core academic effort indicator
- **Attendance:** Measures consistency and class engagement

- **Parental Involvement:** Links to motivation and emotional support
- **Teacher Quality:** Direct influence on instruction
- **Family Income:** Socio-economic advantage and resource access

Data Limitations or Biases

Peer influence and motivation rely on subjective reports
Income categories lack exact value ranges
No geographic segmentation
Learning disabilities listed only as “Yes” or “No”

Data Splitting and Preprocessing

Data Cleaning

- Removed duplicates
- Standardized labels across categories
- Checked numeric fields for valid ranges

Handling Missing Values

Minor blank fields were filled using Excel’s filter and mostly using average of each values.
No major missing values in key metrics.

Data Transformations

- Categorized attendance into High, Medium, and Low
- Aggregated total and average metrics using Pivot Tables
- Grouped income into Low, Medium, High

Data Splitting

Data split using

Dependent Variable: Exam_Score

Independent Variables: Study hours, attendance, involvement level, income, peer influence, motivation etc.

Industry Context

This project is within the **education analytics sector**, focusing on performance modeling.

Stakeholders

- School administrators
- Teachers and counselors

- Education policy makers
- Parents

Value to the Industry

The analysis supports decision-making in:

- Improving student outcomes
- Prioritizing resources
- Designing targeted intervention programs

Pre-Analysis

Key Early Trends Observed

- Public school students make up most of the dataset.
- Majority of students come from medium and low-income households.
- Most students report positive peer influence.
- High teacher quality appears in fewer records than medium quality.

Potential Correlations

- Longer study hours seem to link with better exam performance.
- Higher attendance appears to contribute to higher scores.
- Parental involvement shows a positive relationship with performance.

Initial Insights

Several variables show promise as strong predictors of better exam outcomes, especially study hours and attendance.

In-Analysis

Unconfirmed Insights

Initial pivot analysis suggested:

- Students with greater study hours consistently achieved higher scores.
- Those with high attendance tended to perform better.
- Students from high-income households performed better on average.

These correlations were further validated in post-analysis.

Recommendations Based on Early Findings

- Encourage longer structured study hours.
- Improve attendance monitoring systems.
- Increase parental engagement support programs.

Excel Techniques Used

- Pivot Tables for statistical aggregation
- Pivot Charts for pattern visualization
- Slicers for dynamic analysis
- Conditional formatting to highlight performance extremes

Post-Analysis and Insights

Key Findings

- Average exam score: 67.2
- Total students: 6,607

Top influencers of performance include:

1. Study Hours
2. Attendance
3. Parental Involvement
4. Resource Access
5. Teacher Quality
6. Family Income

Comparison with Initial Findings

Most assumptions were validated, particularly those relating to attendance, study hours, and involvement. Income differences also played a significant role.

Data Visualizations & Charts

The dashboard includes:

- Column charts (Teacher Quality, Study Hours)
- Bar charts (Family Income, Peer Influence)
- Doughnut charts (Attendance, Disabilities)
- KPI cards for total students and average scores
- Slicers (Gender, Motivation, School Type, Resource Access)

Explanation of Visuals

- Long study hours had a significantly larger total study time (241,631 hours).
- Positive peer influence dominates the dataset.
- Distance from home shows an effect on attendance and total study hours.

Actionable Insights

- Introduce structured study programs targeting low and medium performers.
- Provide support to low-income groups through learning resources.
- Strengthen teacher quality through targeted training programs.
- Encourage parental involvement initiatives, especially for low-engagement households.

Business / Academic Implications

- Schools can refine student engagement strategies.
- Policymakers can focus interventions on attendance and resource allocation.
- Parents can support consistent study habits and involvement.

Unexpected Outcomes

- Medium teacher quality had the highest record count, indicating imbalance in teacher distribution.
- Some students with high resource access still performed inconsistently.

Conclusion

This analysis delivers a comprehensive understanding of the academic and socio-economic factors influencing student exam performance.

Study hours, attendance, parental engagement, and resource conditions are key drivers.

Limitations

- No psychological or geographic segmentation
- Self-reported categories may include bias
- Analysis limited to one academic year

Future Work

- Predictive modeling of performance
- Longitudinal performance tracking
- Comparison between public and private institutions

References & Appendices

References:

- Kaggle Dataset
- Excel Pivot Table Documentation
- Vephla University

Appendices:

- Dashboard Screenshots
- Pivot Table Outputs
- Dataset Sample View