Week 1 Task: Data Strategy and Planning

Internship: Virtual Power BI Data Insights Internship

Prepared by: Mohammad Saniya

Date: June 14, 2025

1. Objective

The objective of this task is to define a detailed data strategy and roadmap for analyzing educational data using Power BI. This plan focuses on identifying business objectives related to student performance, discovering reliable data sources, and laying out a step-by-step process for data collection, preparation, modeling, and visualization aligned with Power BI principles. The strategy is designed to help academic institutions derive insights that support informed decision-making.

2. Business Objectives

- Monitor student academic performance trends across semesters
- Track and compare attendance and its correlation with marks
- Measure student engagement through extracurricular activity participation
- Identify at-risk students early for intervention
- Enable faculty and administrators to make data-driven decisions

3. Public Data Sources

The following public data sources are considered suitable for building the educational data model:

- Kaggle Datasets (e.g., Student Performance Data, School Data)
- UCI Machine Learning Repository (Student Alcohol Consumption Dataset)
- Government portals such as data.gov.in for school/college statistics
- World Bank Open Education Data
- National Sample Survey Reports (NSSO)

These sources provide a good mix of academic, demographic, and institutional datasets for creating a holistic model.

4. Data Strategy Roadmap

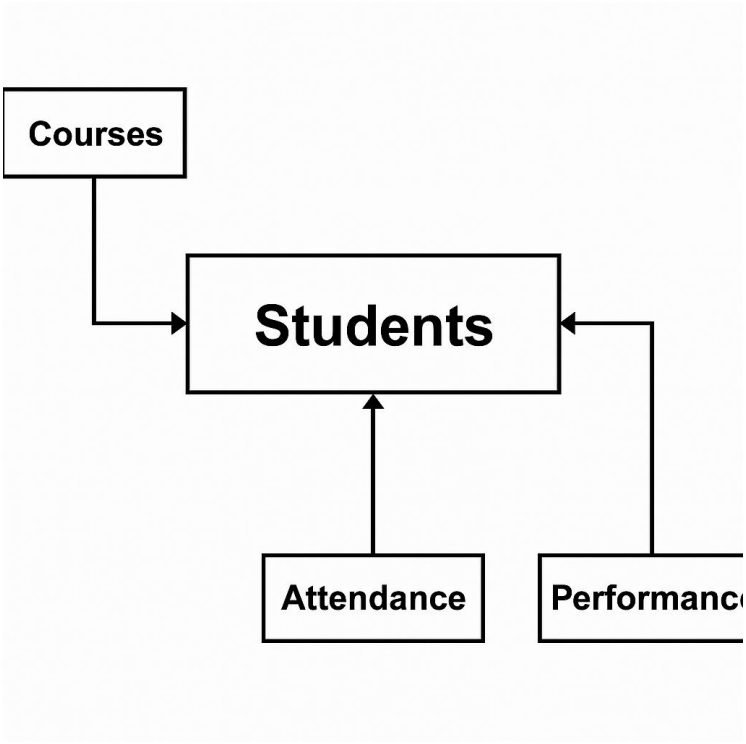
The roadmap includes the following steps:

1. Data Collection: Extract relevant datasets from identified public sources.

2. Data Cleaning: Remove missing values, normalize grading formats, unify student IDs.
3. Data Transformation: Calculate average scores, attendance percentage, activity scores.
4. Data Integration: Merge performance, attendance, and engagement tables by student ID.
5. Modeling in Power BI: Use Power BI to build relationships between tables and define calculated fields using DAX.
6. Visualization: Create dynamic dashboards with cards, charts, and slicers for better interpretation.

5. Conceptual Data Model

The model includes three key tables:
- Students (student_id, name, gender, age)
- Performance (student_id, subject, marks)
- Attendance (student_id, date, status)
- Engagement (student_id, activity_type, participation_level)
All tables are connected through `student_id`. Relationships are one-to-many from Students to other tables.



6. Tools and Techniques

- **Power Query:** For transforming and cleaning data.
- **DAX (Data Analysis Expressions):** For creating measures like average marks, attendance rate.
- Data View: For relationship creation among tables.

- **Visuals:** Bar charts, line charts, pie charts, slicers, and card visuals.
- **Filters and Interactivity:** To enable dynamic slicing of data by class, gender, etc.

7. Mock Dashboard Layout (Descriptive)

- **KPI Cards:** Average score, Attendance %, Participation Rate.
- **Line Chart:** Performance trends across semesters.
- **Matrix Table:** Subject-wise marks comparison.
- **Pie Chart:** Distribution of extracurricular activities.
- **Slicers:** Gender, Class, Activity Type.

8. Conclusion

This Week 1 task successfully outlines a well-structured plan for building a Power BI-based education analytics system. The business objectives are clear, the data sources are identified, and the roadmap aligns with Power BI's methodology. The inclusion of a mock data schema and dashboard layout provides a realistic preview of the project. This groundwork ensures the internship project can proceed smoothly into implementation in the following weeks.