

## Overview:

In schools, students misbehave and skip classes (sometimes for valid excused reasons but not always. These events are recorded as student absences, misbehavior and misbehavior consequences (or resolutions).

Student Absences, Behavior & Resolution data is split by the five school districts (AL, NE, NW, SE & SW) with the data being stored in the following:

District	dbms	Database
NE	MySQL	school_sor_ne
NW	PostgreSQL	school_sor_nw
SE	SQL Server	school_sor_se
SW	Oracle	school_sor_sw
AL	MySQL	school_sor_al
DW	SQL Server	school_dw

The project deliverables:

- Create Schools\_DW database using schools\_dw.sql for SQL Server
- load these data sources into the School DW using Talend
- create data visualizations and dashboards answering a list of questions below using Microsoft PowerBI and Tableau

## Data Integration:

- Load data from the data sources provided into your Schools DWs
  - Talend Enterprise Data Integration
  - DW database: School\_DW – SQL Server
- Separate “reject” rows, i.e. rows that have invalid foreign keys (FKs) in regards to dimensional data provided in data sources. Also put a rejection reason in the Rejects table using column such as: DW\_Reject\_Code (write the specific reason for rejection) – int not null
- Each table needs to record:
  - DI\_Job\_ID (ETL process ID) - **nvarchar (20)** NULL
  - DI\_Create\_Date (Create Date of row) - **datetime** NOT NULL **DEFAULT** (**getdate()**)

### BI Data Visualizations:

- Create data visualizations, tabular reports and dashboards to answer question below using Tableau and PowerBI
- Educators and parents are interested on getting BI reports on student absences, misbehavior and behavior resolutions listing number of students, number of instances, types of instances, types of resolution and location of instances.

### Types of analysis:

- Absence
  - Reason type, reason, gender, language, grade level, class affiliation
  - amount of absences (#, periods, minutes), # of tardy
- Behavior
  - Event type, event, role, event location, school, school type, district, grade level, city, gender
  - Number of occurrences
- Behavior Resolution
  - Resolution(s), event type, event, role, event location, school, school type, district, grade level, city, gender
  - Number of occurrences

### Deliverables:

#### Part 1:

- ETL jobs – job & workflow diagrams (PDF) – generate Talend documentation
- Table rows counts including breakdown of reject rows by reject reason
- Database backup of Schools DW (localhost db, user: info7370& password info7370)

#### Part 2:

- Dashboard(s) – Tableau & Power BI
  - Screenshots
  - Source files for each BI tool zipped

#### Grading:

- Review (Zoom) with Data Integration & BI
  - Need to demonstrate each job works with TAs