

Classroom Music Dataset: Setup & Timeline Guide

STAT 109: Introductory Biostatistics

Classroom Music Dataset: Setup & Timeline Guide



This document is currently in draft status and is not yet finalized.

Last Updated: January 3, 2026

Goal

Use one song per class as a shared, low-effort data-generating object for teaching statistics.
Data are collected at the song × class-day level (not individual level).

BEFORE THE SEMESTER (30–45 minutes total)

1. Decide the scope

- **Target:** ~20–40 songs total for the semester
- **Budget:** \$50–\$100 (buy only songs you actually use)

2. Choose acquisition & playback

- Buy songs (Amazon MP3, AutoRip CDs, or 7digital)
- **Player:** Windows 11 Media Player, VLC, or MusicBee
- Create one local folder:
 - STAT109_Class_Songs_Spring2026

3. Create the Google Sheet

Set up columns (keep it simple at first):

Identification - Date - Class number - Song title - Artist

Song-level (looked up) - Year released - Genre - Song length (sec)

Class-level (show of hands) - students present - familiar - liked - energizing

(Optional) - Notes

You can always add columns later.

DAY 1 (10–15 minutes total)

1. Frame the activity (1 minute)

Say something like:

“We’re going to use music as a running data source this semester.”

2. Seed the dataset

- Enter 2 songs you like into the sheet
- Briefly point out:
 - rows = observations

- columns = variables

3. Collect student suggestions (5 minutes)

- Give students index cards
- Ask them to write:
 - song title
 - artist
- Collect cards (no tech, no pressure)

4. Introduce variable thinking (optional)

Ask: “What could we measure about these songs or our reactions?”

BETWEEN CLASSES (10–20 minutes per week early on)

1. Curate

- Select a subset of student-suggested songs
- You do not need to include every suggestion

2. Buy & download

- Purchase tracks
- Add to your local playlist

3. (Optional) Pre-fill song-level variables

Year, genre, length can be looked up in advance or assigned to students later

EACH CLASS DAY (2–4 minutes)

1. Play the song

- Full song or ~30–60 seconds
- Start after class officially begins

2. Collect class-level data (show of hands)

Ask, count, and record:

- “How many people are here today?”
- “Who’s heard this song before?”
- “Who liked it?”
- “Who found it energizing?”

3. Enter counts live

You or a student types numbers into the sheet

EARLY SEMESTER: TYPES OF VARIABLES DAY

In-class activity (10 minutes)

As a class, decide:

- Which variables are categorical vs quantitative
- Which are ordinal vs nominal

Vote on:

- 1–2 response variables to keep all semester

This gives students ownership and reinforces concepts immediately.

MID–LATE SEMESTER USES (no extra setup)

Use the same dataset for:

- bar charts (genre)
- histograms (year)
- proportions & rates
- comparisons across days
- confounding (attendance, time of day)
- missing data
- simple regression (with caveats)

MAINTENANCE (VERY LIGHT)

- Buy songs only as needed
- Skip days when time is tight
- Let the dataset be a little messy — that's a feature

Why this works

- Minimal time cost per class
- No individual data
- No ads or streaming issues
- Grows naturally with the course
- Feels human, not artificial