Files:

* Eval Data: course\_evaluations\_long\_20001\_to\_20243.csv
* SIS Data Course Instructors Data: SIS Course Instructors.xlsx
* Teaching Credit Database (TCDB) Data: TCDB\_CBS\_Courses.xlsx

GOAL:

* One record per course per instructor.
  + Identify duplicates and remove duplicates
  + Compare enrollment totals to the SIS dataset. If it seems the sum of the enrollment in the evaluations data is close to the total enrollment in the SIS dataset, create a weighted average of the BIE fields. If that doesn’t seem to be the case, look at the rating fields in the TCDB dataset, choose the one that is closest to that. If that doesn’t seem to be the case, choose the ones with the highest enrollment.
* Fill UNIs for as many people as possible and make sure UNIs are accurate for those that already exist
  + Can use the SIS course data
* Make sure class sections in the eval data match those that are pulled from SIS. Identify those that do not.
  + [eval data].course\_number = right([SIS Data].Course\_Identifier,5)
* Compare the eval data to the teaching credit database data. Identify discrepancies in BIE #s
  + bie\_course = RatingCourse
  + bie\_professor = RatingInstructor
* Fill in missing course names from TCDB data or SIS Data