

Description of Data

DSO 570 Final Project (Spring 2018)

There are six raw excel files provided from USC Marshall:

1. **Marshall_Course_Enrollment_1516_1617.xlsx**: Last two years of schedules at Marshall, the times, departments, course numbers, max # of students allowed, and # of students enrolled.
2. **Cancelled_Courses_1516_1617.xlsx**: List of cancelled classes in last two years (important as at the time of initial scheduling, these needed to be accommodated).
3. **Marshall_Room_Capacity_Chart.xlsx**: Rooms within Marshall and max capacity for each. Certain classes may have been scheduled in rooms outside of Marshall, and the capacity of those rooms are not available in this data set.
4. **Summary_Special_Session_Codes_1516_1617.xlsx**: Explanation of session codes, which is a field in the first file that potentially links a course with a specialized program.
5. **Department_Allocations_20171.xlsx**: For spring 2017, the initial slots that each department was given to make their own initial allocations. (Current scheduling work flow is not centralized, but each department receives a set of slots based on last year's allocation and allocate them first. Then there's another round of centralized allocation to fill in remaining slots.)
6. **Student_Course_Selection_1516_1617.xlsx**: Anonymized student level data on what courses each Marshall student chose during these two years. This data can be used to determine courses that students often take together and therefore should not be in conflict with one another. One can also refer to Marshall course sequence for each program for additional information.

Below is a description of each column of each file.

Marshall_Course_Enrollment_1516_1617.xlsx:

1. **Course** : Alpha numeric course name.
2. **Course Prefix** : Alphabetical part of course name.
3. **Course Suffix** : Numerical part of course name.
4. **Department** : Department that owns this course.
5. **First Begin Time** : Time of day that the class begins (First for first meeting time).
6. **First Days** : Day of the week that the class meets.
7. **First End Time** : Time of day that the class ends.
8. **First Instructor** : Name of instructor.
9. **First Instructor UID** : UID= university ID
10. **First Room** : Room that the class meets.
11. **Link** : See "Min Units" below.
12. **Max Units** : Usually the same as min units below.
13. **Min Units** : Courses with a value of "0" are discussion sections or labs associated with a core course. You can see the relationships between a lecture and discussion/lab by matching the "Link" field.
14. **Mode** : Method in which course is delivered. C=lecture, L=lab, D=discussion, CD=combined lecture/discussion
15. **Reg Count** : Number enrolled.
16. **Seats** : Number of seats offered in section. Not always same as classroom size.
17. **Second Begin Time** : Second meeting time for the class (only a handful of classes have multiple meeting times).

18. **Second Days** : Day of week for the second meeting time.
19. **Second End Time** :
20. **Second Instructor** : Certain courses have multiple instructors, sometimes one teaches half and the other the other half.
21. **Second Instructor UID** : UID= university ID
22. **Second Room** : Room for second meeting time.
23. **Section** :
24. **Session** : Determine date range and fee structure for course. "Regular" session codes: 001 (University rate full term used for core and electives--used mostly by UG), 033: Graduate elective, full-term, 411: graduate elective first half, 415 graduate elective second half. See Summary_Special_Session_Codes_1516_1617.xlsx for explanation.
25. **Term** : Terms are identified beginning with the calendar year and then assigned a 1: spring, 2: summer, 3: fall. So, 20153= fall 2015
26. **Title** : Descriptive name of course

Cancelled Courses 1516 1617.xlsx: Same as the enrollment file.

Marshall Room Capacity Chart.xlsx:

1. **Room** : Code defining rooms in various buildings
2. **Size** : Seating capacity of the room

Summary Special Session Codes 1516 1617.xlsx:

1. **Term** : Terms are identified beginning with the calendar year and then assigned a 1: spring, 2: summer, 3: fall. So, 20153= fall 2015
2. **Session** : Code used in the enrollment files under "Session."
3. **Purpose** : Description of the session code for the specific term. (Multiple session codes may be used for essentially the same purpose. There may be nuanced differences that are not important for this project.)

Department Allocations 20171.xlsx:

1. **Room**: Identifier for the room.
2. **Capacity**: capacity of the room.
3. **Days/Start Time/End Time**: Time slot that is assigned.
4. **Department**: Department that this time slot is assigned to. The departments first use these slots to make initial allocations, then the remaining courses are put into the remaining slots in a semi-centralized iterative process.

Student Course Selection 1516 1617.xlsx:

1. **Randomized Unique Identifier**: A scrambled identifier for the student (each maps to a particular USC ID.)
2. **Major/Class/Owner/Objective**: Information about the student's degree of study, year in school, major, etc.
3. **Term**: The academic term. 20153 is Fall 2015, 20161 is Spring 2016, 20162 is Summer 2017, and so on.
4. **Course**: Identifier for the course.
5. **Title**: Title for the course.
6. **Section**: Section enrolled in.
7. **Instructor**: Instructor for the course.

8. **# Students:** # of students enrolled at the time the student enrolls. (The first student who enrolls in the section gets number 1, the second student number 2, until the class fills.)
9. **Units:** # of academic units.
10. **Enroll:** The way student enrolled. Usually is L. Not important for this project.
11. **Withdraw:** Whether the student withdrew later.