

Description of the Project-Team Formation Problem Data

1. Title: Project-Team Formation Problem (PTFP) Data
2. Source Information:
 - a) Owners:
Fatih Cavdur, Asli Sebatli, Merve Kose-Kucuk, Cagla Rodoplu
Uludağ University, College of Engineering, Department of Industrial Engineering
 - b) Contact:
Fatih Cavdur: fatihcavdur@uludag.edu.tr
Asli Sebatli: aslisebatli@gmail.com
Merve Kose-Kucuk: mervekose@uludag.edu.tr
 - c) Date: January, 2018
3. First Usage:
[Cavdur, F., Sebatli, A., Kose-Kucuk, M., Rodoplu, C. 2018. A two-phase binary-goal programming-based approach for optimal project-team formation. Journal of the Operational Research Society. (article in press) <https://doi.org/10.1080/01605682.2018.1457480>]
4. Relevant Information:
PTFP Data includes three real-life datasets and each dataset consists of three files as (i) Adviser Preferences (ii) Student Preferences and (iii) Student Qualifications.
 - a) Adviser Preferences data files (i.e., “ds1_ap.csv”) contain the preferences of academic advisers in a matrix structure. The rows of a matrix represent the academic adviser index whereas the columns show the project index. Smaller numbers indicate more preferred projects. Missing cells (N/A) show that the academic adviser does not indicate her/his preference for the project.
 - b) Student Preferences data files (i.e., “ds1_sp.csv”) contain the preferences of students in a matrix structure. The rows of the matrix represent the student index whereas columns show the project index. Smaller numbers indicate more preferred projects. Missing cells (N/A) show that the student does not indicate her/his preference for the project.
 - c) Student Qualifications data files (i.e., “ds1_sq.csv”) contain five different (two hierarchical and three categorical) qualifications of the students each in a vector structure. First qualification is represented by the Grade Point Average (GPA) of the students and it has a numerical value between zero and four. Second qualification is the language skills of the students represented by their language class grades with possible values between FF (worst grade) and AA (best grade). The other three qualifications are qualitative which are set to either YES or NO and they represent general software, computer programming and database system management skills, respectively.
5. Number of instances:

	# of advisers	# of students	# of projects
Dataset1	11	78	23
Dataset2	11	50	18
Dataset3	12	70	25