hstrick2_Final

Hannah Strickling

2022-11-23

Problem Set Up

What factors affect the success of groups? My three factors are GPA, attendance, and background. The GPA will range on a scale of 0.0 - 4.0. The higher the GPA, the better chance that student has at success. Next will be attendance. We will assume that this is an end of course project for a class that had 30 class meetings. Attendance will be marked on a scale 0 - 30. The number will correspond the the number of days the students missed. The closer to 0, the better chance the student has at success. The final factor is student background. Some students might be taking this class for a major course requirement whereas others are taking it to fulfill an elective requirement. We will use binary coding for this. 0 will represent no subject background whereas 1 represent course background. We can assume that our 1 (students with course background) will score higher.

When considering this data, GPA and class attendance are the two most important factors. Based on my personal experience, I have found that if I go to class and actively participate, I am able succeed regardless of my background. Typically a student who actively participates in class tends to care about the course and their grade. A hardworking student tends to be associated with a higher GPA. Course background matters, however not as much as these other two factors. Just because the subject is your background does not mean you are willing to work hard. To collect this data, we will randomly generate 12 numbers for each factor in the range described above.

```
round(runif(12, min = 0, max = 4), 2)
```

```
##  [1] 0.46 2.95 3.86 0.17 1.93 1.03 2.64 2.06 0.64 2.91 1.13 3.97
```

Every time I run this code, the random numbers will change so the results for the random GPAs are as follows: 2.81, 0.96, 2.3, 1.01, 2.36, 2.23, 3.55, 3.56, 3.9, 2.07, 0.18, 2.54

```
round(runif(12, min = 0, max = 30), 0)
```

```
##  [1] 15 28 20 17  8 27 20 26 28 29 25 18
```

Every time I run this code, the random numbers will change so the results for the random days of class missed are as follows: 26, 30, 10, 7, 8, 3, 18, 23, 9, 25, 16, 16

```
round(runif(12, min = 0, max = 1), 0)
```

```
##  [1] 0 1 1 1 1 0 1 1 0 1 1 1
```

Every time I run this code, the random numbers will change so the results for the random class backgrounds are as follows: 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 0, 0

Putting this random data together will now give us the following students

x1: GPA = 2.81, Missed Days = 26, Class Background = Yes (1) x2: GPA = 0.96, Missed Days = 30, Class Background = No (0) x3: GPA = 2.30, Missed Days = 10, Class Background = Yes (1) x4: GPA = 1.01, Missed Days = 7, Class Background = Yes (1) x5: GPA = 2.36, Missed Days = 8, Class Background = Yes (1) x6: GPA = 2.33, Missed Days = 3, Class Background = Yes (1) x7: GPA = 3.55, Missed Days = 18, Class Background = Yes (1) x8: GPA = 3.56, Missed Days = 23, Class Background = Yes (1) x9: GPA = 3.90, Missed Days = 9, Class Background = No (0) x10: GPA = 2.07, Missed Days = 25, Class Background = Yes (1) x11: GPA = 0.81, Missed Days = 16, Class Background = No (0) x12: GPA = 2.54, Missed Days = 16, Class Background = No (0)

In order to work with scales that are consistent with the ranking importance described above, we will convert missed days to a 0-4 scale. To do this, we will take number of days missed x 4, then divide by 30. GPA and Missed days can impact the grade by up to 4 "points" whereas the class background only 1 point. The factors that are most important will have the most impact on the grade.

x1: GPA = 2.81, Missed Days = 3.46, Class Background = Yes (1) x2: GPA = 0.96, Missed Days = 4, Class Background = No (0) x3: GPA = 2.30, Missed Days = 1.33, Class Background = Yes (1) x4: GPA = 1.01, Missed Days = 0.93, Class Background = Yes (1) x5: GPA = 2.36, Missed Days = 1.06, Class Background = Yes (1) x6: GPA = 2.33, Missed Days = 0.4, Class Background = Yes (1) x7: GPA = 3.55, Missed Days = 2.4, Class Background = Yes (1) x8: GPA = 3.56, Missed Days = 3.06, Class Background = Yes (1) x9: GPA = 3.90, Missed Days = 1.2, Class Background = No (0) x10: GPA = 2.07, Missed Days = 2.5, Class Background = Yes (1) x11: GPA = 0.81, Missed Days = 2.13, Class Background = No (0) x12: GPA = 2.54, Missed Days = 2.13, Class Background = No (0)

Putting this together we will get

Student x1 2.81 - 3.46 + 1 = 0.35. The 0.35 represents the student's score. Here, the higher the score, the more potential the student has to score well on a class assignment. Likewise, a lower score would represent a student who would not score as well. Following this logic, the score for each student is as follows.

From this, we can calculate the class average to be 0.9666667. This means that is we have 1 group with 12 students, the highest "score" is 0.9666667. Therefore, the groups should be distributed in a way that each group averages 0.966667. To do this, we can look at the variance of each student's individual grade with respect to the class grade. Our goal is to minimize the variance of all four groups. Below is the individual student variance calculations.

```
Mean <- 0.966667


x1 <- 0.35 - Mean
x2 <- -3.04 - Mean
x3 <- 1.97 - Mean
x4 <- 1.08 - Mean
x5 <- 2.3 - Mean
x6 <- 2.93 - Mean
x7 <- 2.15 - Mean
x8 <- 1.5 - Mean
x9 <- 2.7 - Mean
x10 <- 0.57 - Mean
x11 <- -1.32 - Mean
x12 <- 0.41 - Mean

x1
```

```
## [1] -0.616667
```

```
x2
```

```
## [1] -4.006667
```

```
x3
```

```
## [1] 1.003333
```

```
x4
```

```
## [1] 0.113333
```

```
x5
```

```
## [1] 1.333333
```

```
x6
```

```
## [1] 1.963333
```

```
x7
```

```
## [1] 1.183333
```

```
x8
```

```
## [1] 0.533333
```

```
x9
```

```
## [1] 1.733333
```

```
x10
```

```
## [1] -0.396667
```

```
x11
```

```
## [1] -2.286667
```

```
x12
```

```
## [1] -0.556667
```

For the students above, if the variance is negative, it shows that the student is below average. Positive numbers are above average and 0 is average. To set this up similar to a problem we have seen in class, we can now think of the variances as the "cost" of each student. Our goal is to minimize the variance or cost. Since we are trying to have our variance be as close to 0 as possible, we will need to consider the distance the variance is from 0. This means that we will take the absolute value of our variances. Referring back to student one, we found that taking the students score - class average = 0.62 or a cost of 0.62

Objective: (Minimize) 0.62 S1G1 + 0.62 S1G2 + 0.62 S1G3 + 0.62S1G4 + 4.01 S2G1 + 4.01 S2G2 + 4.01 S2G3 + 4.01 S2G4 + 1 S3G1 + 1 S3G2 + 1 S3G3 + 1 S3G4 + 0.11 S4G1 + 0.11 S4G2 + 0.11 S4G3 + 0.11 S4G4 + 1.33 S5G1 + 1.33 S5G2 + 1.33 S5G3 + 1.33 S5G4 + 1.96 S6G1 + 1.96 S6G2 + 1.96 S6G3 + 1.96 S6G4 + 1.18 S7G1 + 1.18 S7G2 + 1.18 S7G3 + 1.18 S7G4 + 0.53 S8G1 + 0.53 S8G2 + 0.53 S8G3 + 0.53 S8G4 + 1.73 S9G1 + 1.73 S9G2 + 1.73 S9G3 + 1.73 S9G3 + 0.4 S10G1 + 0.4 S10G2 + 0.4 S10G3 + 0.4 S10G4 + 2.29 S11G1 + 2.29 S11G2 + 2.29 S11G3 + 2.29 S11G4 + 0.56 S12G1 + 0.56 S12G2 + 0.56 S12G3 + 0.56 S12G4

Where S stands for the student number and G refers to the group number. Ex: S1G1 is Student 1 Group 1

Subject to the following constraints:

1. The total number of students in each group must be 3. Therefore

(Group One 3 Student Constraint) S1G1 + S2G1 + S3G1 + S4G1 + S5G1 + S6G1 + S7G1 + S8G1 + S9G1 + S10G1 + S11G1 + S12G1 = 3 Students

(Group Two 3 Student Constraint) S1G2 + S2G2 + S3G2 + S4G2 + S5G2 + S6G2 + S7G2 + S8G2 + S9G2 + S10G2 + S11G2 + S12G2 = 3 Students

(Group Three 3 Student Constraint) S1G3 + S2G3 + S3G3 + S4G3 + S5G3 + S6G3 + S7G3 + S8G3 + S9G3 + S10G3 + S11G3 + S12G3 = 3 Students

2. Next, each student may only appear in one group. Therefore

S1G1 + S1G2 + S1G3 + S1G4 = 1 S2G1 + S2G2 + S2G3 + S2G4 = 1 S3G1 + S3G2 + S3G3 + S3G4 = 1 S4G1 + S4G2 + S4G3 + S4G4 = 1 S5G1 + S5G2 + S5G3 + S5G4 = 1 S6G1 + S6G2 + S6G3 + S6G4 = 1 S7G1 + S7G2 + S7G3 + S7G4 = 1 S8G1 + S8G2 + S8G3 + S8G4 = 1 S9G1 + S9G2 + S9G3 + S9G4 = 1 S10G1 + S10G2 + S10G3 + S10G4 = 1 S11G1 + S11G2 + S11G3 + S11G4 = 1 S12G1 + S12G2 + S12G3 + S12G4 = 1

3. We also want to make sure that each student is used. Therefore

(S1G1 + S1G2 + S1G3 + S1G4) + (S2G1 + S2G2 + S2G3 + S2G4) + (S3G1 + S3G2 + S3G3 + S3G4) + (S4G1 + S4G2 + S4G3 + S4G4) + (S5G1 + S5G2 + S5G3 + S5G4) + (S6G1 + S6G2 + S6G3 + S6G4) + (S7G1 + S7G2 + S7G3 + S7G4) + (S8G1 + S8G2 + S8G3 + S8G4) + (S9G1 + S9G2 + S9G3 + S9G4) + (S10G1 + S10G2 + S10G3 + S10G4) + (S11G1 + S11G2 + S11G3 + S11G4) + (S12G1 + S12G2 + S12G3 + S12G4) = 12

4. Our next constraint is that the output must be binary. This will turn our problem into a BIP problem where our output will be 1 or zero where,

1 = decision is yes, the student will be placed in the group 0 = decision is no, the student does not belong to the group

5. Finally, we want each group to be as equal as possible. This means that each group's variance should be equal to each other and as close to zero as possible.

$$(S1G1 + S2G1 + S3G1 + S4G1 + S5G1 + S6G1 + S7G1 + S8G1 + S9G1 + S10G1 + S11G1 + S12G1) - (S1G2 + S2G2 + S3G2 + S4G2 + S5G2 + S6G2 + S7G2 + S8G2 + S9G2 + S10G2 + S11G2 + S12G2) = 0$$

$$(S1G1 + S2G1 + S3G1 + S4G1 + S5G1 + S6G1 + S7G1 + S8G1 + S9G1 + S10G1 + S11G1 + S12G1) - (S1G3 + S2G3 + S3G3 + S4G3 + S5G3 + S6G3 + S7G3 + S8G3 + S9G3 + S10G3 + S11G3 + S12G3) = 0$$

$$(S1G1 + S2G1 + S3G1 + S4G1 + S5G1 + S6G1 + S7G1 + S8G1 + S9G1 + S10G1 + S11G1 + S12G1) - (S1G4 + S2G4 + S3G4 + S4G4 + S5G4 + S6G4 + S7G4 + S8G4 + S9G4 + S10G4 + S11G4 + S12G4) = 0$$

$$(S1G2 + S2G2 + S3G2 + S4G2 + S5G2 + S6G2 + S7G2 + S8G2 + S9G2 + S10G2 + S11G2 + S12G2) - (S1G3 + S2G3 + S3G3 + S4G3 + S5G3 + S6G3 + S7G3 + S8G3 + S9G3 + S10G3 + S11G3 + S12G3) = 0$$

$$(S1G2 + S2G2 + S3G2 + S4G2 + S5G2 + S6G2 + S7G2 + S8G2 + S9G2 + S10G2 + S11G2 + S12G2) - (S1G4 + S2G4 + S3G4 + S4G4 + S5G4 + S6G4 + S7G4 + S8G4 + S9G4 + S10G4 + S11G4 + S12G4) = 0$$

$$(S1G3 + S2G3 + S3G3 + S4G3 + S5G3 + S6G3 + S7G3 + S8G3 + S9G3 + S10G3 + S11G3 + S12G3) - (S1G4 + S2G4 + S3G4 + S4G4 + S5G4 + S6G4 + S7G4 + S8G4 + S9G4 + S10G4 + S11G4 + S12G4) = 0$$

We can now begin the problem formulation step

```
lprec <- make.lp(0, 48)
set.objfn(lprec, c(.62,.62,.62,.62,4.01,4.01,4.01,4.01,1,1,1,1,.11,.11,.11,.11,1.33,1.33,1.33,1.33,1.96
lp.control(lprec, sense = 'min', all.bin=TRUE)
```

```
## $anti.degen
## [1] "fixedvars" "stalling"
##
## $basis.crash
## [1] "none"
##
## $bb.depthlimit
## [1] -50
##
## $bb.floorfirst
## [1] "automatic"
##
## $bb.rule
## [1] "pseudononint" "greedy"       "dynamic"      "rcostfixing"
##
## $break.at.first
## [1] FALSE
##
## $break.at.value
## [1] -1e+30
##
## $epsilon
##       epsb       epsd      epsel      epsint epsperturb   epspivot
##      1e-10      1e-09      1e-12       1e-07      1e-05      2e-07
##
## $improve
## [1] "dualfeas" "thetagap"
```

```
##
## $infinite
## [1] 1e+30
##
## $maxpivot
## [1] 250
##
## $mip.gap
## absolute relative
##    1e-11    1e-11
##
## $negrange
## [1] -1e+06
##
## $obj.in.basis
## [1] TRUE
##
## $pivoting
## [1] "devex"    "adaptive"
##
## $presolve
## [1] "none"
##
## $scalelimit
## [1] 5
##
## $scaling
## [1] "geometric"   "equilibrate" "integers"
##
## $sense
## [1] "minimize"
##
## $simplextype
## [1] "dual"   "primal"
##
## $timeout
## [1] 0
##
## $verbose
## [1] "neutral"
```

```
add.constraint(lprec, c(1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0
```

```
add.constraint(lprec, c(0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0
```

```
add.constraint(lprec, c(0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0
```

```
add.constraint(lprec, c(0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1
```

```
add.constraint(lprec, c(1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
```

```
add.constraint(lprec, c(0,0,0,0,1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
```

```
add.constraint(lprec, c(0,0,0,0,0,0,0,0,1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
```

```
add.constraint(lprec, c(0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
add.constraint(lprec, c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
add.constraint(lprec, c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
add.constraint(lprec, c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0
add.constraint(lprec, c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,0,0,0,0,0,0,0,0
add.constraint(lprec, c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,0,0,0,0
add.constraint(lprec, c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1
add.constraint(lprec, c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
add.constraint(lprec, c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0

add.constraint(lprec, c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1

add.constraint(lprec, c(1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,(
add.constraint(lprec, c(1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,(
add.constraint(lprec, c(1,0,0,-1,1,0,0,-1,1,0,0,-1,1,0,0,-1,1,0,0,-1,1,0,0,-1,1,0,0,-1,1,0,0,-1,1,0,0,-1,1,0,0,-1,
add.constraint(lprec, c(0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,(
add.constraint(lprec, c(0,1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,0,1,0,-1,
add.constraint(lprec, c(0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,0,0,1,-1,

solve(lprec)

## [1] 0

get.objective(lprec)

## [1] 15.72

get.variables(lprec)

##  [1] 0 0 1 0 1 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0
## [39] 0 1 1 0 0 0 0 0 1 0
```

Conclusions:

From this, we will see that our output has formed 4 groups, with three students in each group. Each student is used with no repeats. The groups will be formed as follows:

Group 1: Students 2,6,11 x2: GPA = 0.96, Missed Days = 30, Class Background = No (0) x6: GPA = 2.33, Missed Days = 3, Class Background = Yes (1) x11: GPA = 0.81, Missed Days = 16, Class Background = No (0)

Group 1 Averages GPA:1.36 Missed Days:16 Class Background: 0.3

Group 2: Students 5,7,9 x5: GPA = 2.36, Missed Days = 8, Class Background = Yes (1) x7: GPA = 3.55, Missed Days = 18, Class Background = Yes (1) x9: GPA = 3.90, Missed Days = 9, Class Background = No (0)

Group 2 Averages GPA:3.27 Missed Days: 12 Class Background: 0.67

Group 3: Students 1,3,12 x1: GPA = 2.81, Missed Days = 26, Class Background = Yes (1) x12: GPA = 2.54, Missed Days = 16, Class Background = No (0) x3: GPA = 2.30, Missed Days = 10, Class Background = Yes (1)

Group 3 Averages GPA: 2.55 Missed Days: 17 Class Background: 0.67

Group 4: Students 4,8,10 x4: GPA = 1.01, Missed Days = 7, Class Background = Yes (1) x8: GPA = 3.56, Missed Days = 23, Class Background = Yes (1) x10: GPA = 2.07, Missed Days = 25, Class Background = Yes (1)

Group 4 Averages GPA: 2.21 Missed Days: 18 Class Background: 1