ASSIGNING ENGINEERING STUDENTS TO DTC PROJECT TEAMS

Phase 2: Nov 28, 2022

Group: Jessie Lanin, Victor Agaba, Emily Wei

Question 1:

Mathematical Formulation

- Sets
  - S : Set of students (for every student i)
  - P : Set of team projects (for every project p)
  - $M = \{1, 2, 3\}$: Set of major combinations (for every major combination c)
    1 = CIV-MECH-MADE
    2 = IE-ECE-ESAM
    3 = BME-CHEME-MSE
  - G : Set of genders (for every gender k)

- Parameters
  - n : Number of students per project
  - r_{ip} : Rating per student i of project p
  - a_i : GPA per student i
  - a_l and a_u : Lower and upper bounds on average student GPA
  - m_{ic} : 1 if student i is in major group c, 0 otherwise
  - m_{l2} : minimum number of students in CS major combination
  - g_{ik} : 1 if student i is gender k, 0 otherwise
  - μ : Maximum proportion of students with the same major per group
  - H_l: Lower bound on total happiness of the group

- Decision variables
  - s_{ip} : Binary variable = 1 if student i is assigned to project p, 0 otherwise

Derivation of the objective function:

The Gini index is a measure of the diversity of a group in relation to a specified quality. If p_k is the proportion of students of gender k, then

$$p_k = \frac{\sum_{i \in S} g_{ik} s_{ip}}{\sum_{i \in S} s_{ip}} = \frac{1}{n} \sum_{i \in S} g_{ik} s_{ip}$$

$$Gini\ index\ \equiv 1 - \sum_{k \in G} {p_k}^2 = 1 - \frac{1}{n^2} \sum_{k \in G} \left(\sum_{i \in S} g_{ik} s_{ip} \right)^2$$

This quantity considered per project is at a maximum when $\sum_{k\in G}\left(\sum_{i\in S} g_{ik}s_{ip}\right)^2$ is at a minimum, thereby simplifying the objective function. The same method can be applied to question 2, which considers major proportions instead of gender.

Linear Program :

Maximize the gender diversity, measured by the total *Gini* index across the project teams

$$\min \sum_{p\in \mathrm{P}}\sum_{k\in G}\left(\sum_{i\in S} g_{ik}s_{ip}\right)^2$$

s.t.

Happiness total minimum per project

$$\sum_{i\in \mathrm{S}}\sum_{p\in P} r_{ip}s_{ip} \geq H_l$$

Number of students per project

$$\sum_{i\in S} s_{ip} = n \quad \forall p \in P$$

Number of projects per student, and each student must be assigned

$$\sum_{p\in P} s_{ip} = 1 \quad \forall i \in S$$

Average GPA lower bound per project

$$\sum_{i\in S} a_i s_{ip} \geq a_l\left(\sum_{i\in S} s_{ip}\right) \quad \forall p \in P$$

Average GPA upper bound per project

$$\sum_{i\in S} a_i s_{ip} \leq a_u\left(\sum_{i\in S} s_{ip}\right) \quad \forall p \in P$$

Minimum CS combination students per project

$$\sum_{i\in S} m_{i2}s_{ip} \geq m_{l2} \quad \forall p \in P$$

Maximum major composition ratio

$$\sum_{i\in S} m_{ic}s_{ip} \leq \mu\left(\sum_{i\in S} s_{ip}\right) \quad \forall p \in P, c \in M$$

Binary decision variable

$$s_{ip} = \{0, 1\} \quad \forall i \in S, p \in P$$

Solution Report

We tested the data with 40 students and 5 projects, assigning a set number of 8 students per project. Our results for the optimal solution (Gender Proportion = 176, where the ratio is roughly 5:3 amongst all the teams) are summarized below:

$$Av\ Gini\ index = \frac{1}{\#Proj}\left(\#Proj - \frac{Gender\ Proportion}{n^2}\right) = 0.45$$

Projects	1	2	3	4	5
Assigned student IDs	11, 12, 16, 20, 25, 26, 28, 32	6, 10, 15, 17, 27, 34, 35, 36	1, 2, 3, 4, 5, 13, 21, 37	7, 14, 18, 19, 23, 24, 29, 40	8, 9, 22, 30, 31, 33, 38, 39
Average GPA	3.165	3.095	3.041	3.631	3.123
Major composition	CIV: 1 IE: 6 BME: 1	CIV: 1 IE: 4 BME: 3	CIV: 1 IE: 3 BME: 4	CIV: 3 IE: 4 BME: 1	CIV: 4 IE: 2 BME: 2
Gender diversity	M: 6 F: 2	M: 5 F: 3	M: 5 F: 3	M: 5 F: 3	M: 5 F: 3

<u>**Question 2:**</u>

Mathematical Formulation

- Sets
  - S : Set of students (for every student i)
  - P : Set of team projects (for every project p)
  - $M = \{1, 2, 3\}$: Set of major combinations (for every major combination c)
    1 = CIV-MECH-MADE
    2 = IE-ECE-ESAM
    3 = BME-CHEME-MSE
  - G : Set of genders (for every gender k)

- Parameters
  - n : Number of students per project
  - r_{ip} : Rating per student i of project p
  - a_i : GPA per student i
  - a_l and a_u : Lower and upper bounds on average student GPA

- m_{ic} : 1 if student i is in major group c, 0 otherwise
- m_{l2} : minimum number of students in CS major combination
- g_{ik} : 1 if student i is gender k, 0 otherwise
- γ : Maximum proportion of students of the same gender per group
- H_l: Lower bound on total happiness of the group

- Decision variables
  - s_{ip} : Binary variable $= 1$ if student i is assigned to project p, 0 otherwise

Linear Program :

Using a similar derivation as in question 1, the objective function with respect to major proportion can be modified from a maximization of $1 - \frac{1}{n^2}\sum_{c\in M}\left(\sum_{i\in S} m_{ic}s_{ip}\right)^2$ to a minimization of $\sum_{p\in \mathrm{P}}\sum_{m\in M}\left(\sum_{i\in S} m_{ic}s_{ip}\right)^2$.

To maximize the major composition, the new objective is

$$\min \sum_{p\in \mathrm{P}}\sum_{m\in M}\left(\sum_{i\in S} m_{ic}s_{ip}\right)^2$$

s.t.

Happiness total minimum per project

$$\sum_{i\in \mathrm{S}}\sum_{p\in P} r_{ip}s_{ip} \geq H_l$$

Number of students per project

$$\sum_{i\in S} s_{ip} = n \quad \forall p \in P$$

Number of projects per student, and each student must be assigned

$$\sum_{p\in P} s_{ip} = 1 \quad \forall i \in S$$

Average GPA lower bound per project

$$\sum_{i\in S} a_i s_{ip} \geq a_l\left(\sum_{i\in S} s_{ip}\right) \quad \forall p \in P$$

Average GPA upper bound per project

$$\sum_{i\in S} a_i s_{ip} \leq a_u\left(\sum_{i\in S} s_{ip}\right) \quad \forall p \in P$$

Minimum CS combination students per project

$$\sum_{i \in S} m_{i2} s_{ip} \geq m_{l2} \quad \forall p \in P$$

Maximum gender composition ratio

$$\sum_{i \in S} g_{ik} s_{ip} \leq \gamma \left(\sum_{i \in S} s_{ip} \right) \quad \forall p \in P, k \in G$$

Binary decision variable

$$s_{ip} = \{0, 1\} \quad \forall i \in S, p \in P$$

Solution Report

We tested the data with 40 students and 5 projects, assigning a set number of 8 students per project. Our results for the optimal solution (Major Proportion = 118, where the ratio is roughly 2:4:2 amongst all the teams) are summarized below:

$$Av\ Gini\ index = \frac{1}{\#Proj} \left(\#Proj - \frac{Major\ Proportion}{n^2} \right) = 0.63$$

Projects	1	2	3	4	5
Assigned student IDs	2, 4, 14, 20, 25, 26, 28, 32	1, 6, 12, 15, 16, 22, 35, 39	3, 5, 10, 13, 17, 21, 34, 38	7, 18, 23, 24, 27, 29, 31, 40	8, 9, 11, 19, 30, 33, 36, 37
Average GPA	3.193	3.193	3.071	3.47	3.129
Major composition	CIV: 2 IE: 4 BME: 2	CIV: 2 IE: 3 BME: 3	CIV: 2 IE: 4 BME: 2	CIV: 3 IE: 3 BME: 2	CIV: 2 IE: 4 BME: 2
Gender diversity	M: 5 F: 3	M: 5 F: 3	M: 5 F: 3	M: 5 F: 3	M: 6 F: 2

Question 3:

Mathematical Formulation

- Sets
  - S : Set of students (for every student i)
  - P : Set of team projects (for every project p)
  - $M = \{1, 2, 3\}$: Set of major combinations (for every major combination c)

    $1 =$ CIV-MECH-MADE

    $2 =$ IE-ECE-ESAM

    $3 =$ BME-CHEME-MSE
  - G : Set of genders (for every gender k)

- Parameters
  - n : Number of students per project
  - r_{ip} : Rating per student i of project p
  - a_i : GPA per student i
  - a_l and a_u : Lower and upper bounds on average student GPA
  - m_{ic} : 1 if student i is in major group c, 0 otherwise
  - m_{l2} : minimum number of students in CS major combination
  - g_{ik} : 1 if student i is gender k, 0 otherwise
  - μ : Maximum proportion of students with the same major per group
  - γ : Maximum proportion of students of the same gender per group
  - H_l: Lower bound on total happiness of the group

- Decision variables
  - s_{ip} : Binary variable $= 1$ if student i is assigned to project p, 0 otherwise

Linear Program :

Minimize total conflict of interest

$$\min \sum_{i \in S} \sum_{j \in S} \sum_{p \in P} \left| r_{ip} - r_{jp} \right| \cdot s_{ip} s_{jp}$$

s.t.

Happiness total minimum per project

$$\sum_{i \in S} \sum_{p \in P} r_{ip} s_{ip} \geq H_l$$

Number of students per project

$$\sum_{i \in S} s_{ip} = n \quad \forall p \in P$$

Number of projects per student, and each student must be assigned

$$\sum_{p \in P} s_{ip} = 1 \quad \forall i \in S$$

Average GPA lower bound per project

$$\sum_{i \in S} a_i s_{ip} \geq a_l \left(\sum_{i \in S} s_{ip} \right) \quad \forall p \in P$$

Average GPA upper bound per project

$$\sum_{i \in S} a_i s_{ip} \leq a_u \left(\sum_{i \in S} s_{ip} \right) \quad \forall p \in P$$

Minimum CS combination students per project

$$\sum_{i \in S} m_{i2} s_{ip} \geq m_{l2} \quad \forall p \in P$$

Maximum major composition ratio

$$\sum_{i \in S} m_{ic} s_{ip} \leq \mu \left(\sum_{i \in S} s_{ip} \right) \quad \forall p \in P, c \in M$$

Maximum gender composition ratio

$$\sum_{i \in S} g_{ik} s_{ip} \leq \gamma \left(\sum_{i \in S} s_{ip} \right) \quad \forall p \in P, k \in G$$

Binary decision variable

$$s_{ip} = \{0, 1\} \quad \forall i \in S, p \in P$$

Solution Report

We tested the data with 40 students and 5 projects, assigning a set number of 8 students per project. Our results for the optimal solution (Total Conflict of Interest = 338) are summarized below:

Projects	1	2	3	4	5
Assigned student IDs	10, 12, 15, 21, 22, 23, 32, 38	6, 7, 13, 29, 33, 35, 36, 39	2, 3, 5, 11, 16, 17, 20, 27	4, 14, 18, 19, 24, 25, 28, 34	1, 8, 9, 26, 30, 31, 37, 40
Average GPA	3.24	3.02	3.29	3.36	3.15

Major composition	CIV: 2 IE: 5 BME: 1	CIV: 1 IE: 4 BME: 3	CIV: 1 IE: 4 BME: 3	CIV: 4 IE: 2 BME: 2	CIV: 2 IE: 4 BME: 2
Gender diversity	M: 6 F: 2	M: 6 F: 2	M: 5 F: 3	M: 4 F: 4	M: 6 F: 2

Combined Model and Sensitivity Analysis

We generated a combined model by considering the average major and gender proportions per team from questions 1 and 2, then using those results as constraints while minimizing the total conflict of interest.

The idea behind this combined model is to find a "sweet spot" that maintains a fairly diverse set of teams while minimizing their total conflict of interest at a given level of total happiness. We have included considerations for conflict of interest to increase the sense of inclusion in each team, because this would also be a contributor to individual student happiness as opposed to the overall.

The results for major and gender proportions are as follows:

Question 1 Proportions per Project						
Gender	P1	P2	P3	P4	P5	Av Prop'n
M	6/8	5/8	5/8	5/8	5/8	0.65
F	2/8	3/8	3/8	3/8	3/8	0.35
Major						
CIV	1/8	1/8	1/8	3/8	4/8	0.25
IE	6/8	4/8	3/8	4/8	2/8	0.475
BME	1/8	3/8	4/8	1/8	2/8	0.275

Question 2 Proportions per Project						
Gender	P1	P2	P3	P4	P5	Av Prop'n
M	5/8	5/8	5/8	5/8	6/8	0.65
F	3/8	3/8	3/8	3/8	2/8	0.35
Major						
CIV	2/8	2/8	2/8	2/8	2/8	0.25
IE	4/8	3/8	4/8	4/8	4/8	0.475
BME	2/8	3/8	2/8	2/8	2/8	0.275

Av Q1 & Q2	
Gender	Av Prop'n
M	0.65
F	0.35

Major	
CIV	0.25
IE	0.475
BME	0.275

Mathematical Formulation

Revisions to Question 3 for the combined model are highlighted in red.

- Sets
  - S : Set of students (for every student i)
  - P : Set of team projects (for every project p)
  - $M = \{1, 2, 3\}$: Set of major combinations (for every major combination c)
    - $1 =$ CIV-MECH-MADE
    - $2 =$ IE-ECE-ESAM
    - $3 =$ BME-CHEME-MSE
  - G : Set of genders (for every gender k)

- Parameters
  - n : Number of students per project
  - r_{ip} : Rating per student i of project p
  - a_i : GPA per student i
  - a_l and a_u : Lower and upper bounds on average student GPA
  - m_{ic} : 1 if student i is in major group c, 0 otherwise
  - m_{l2} : minimum number of students in CS major combination
  - g_{ik} : 1 if student i is gender k, 0 otherwise
  - μ_c : Average proportion of students per group with major c
  - $t_\mu = 0.1$: Tolerance on average major proportion per project
  - γ_k : Average proportion of students per group of gender k
  - $t_\gamma = 0.1$: Tolerance on average gender proportion per project
  - H_l: Lower bound on total happiness of the group

- Decision variables
  - s_{ip} : Binary variable $= 1$ if student i is assigned to project p, 0 otherwise

Linear Program :

Minimize total conflict of interest

$$\min \sum_{i \in S} \sum_{j \in S} \sum_{p \in P} |r_{ip} - r_{jp}| \cdot s_{ip} s_{jp}$$

s.t.

Happiness total minimum per project

$$\sum_{i \in S} \sum_{p \in P} r_{ip} s_{ip} \geq H_l$$

Number of students per project

$$\sum_{i \in S} s_{ip} = n \quad \forall p \in P$$

Number of projects per student, and each student must be assigned

$$\sum_{p \in P} s_{ip} = 1 \quad \forall i \in S$$

Average GPA bounds per project

$$\sum_{i \in S} a_i s_{ip} \geq a_l \left(\sum_{i \in S} s_{ip} \right) \quad \forall p \in P$$

$$\sum_{i \in S} a_i s_{ip} \leq a_u \left(\sum_{i \in S} s_{ip} \right) \quad \forall p \in P$$

Minimum CS combination students per project

$$\sum_{i \in S} m_{i2} s_{ip} \geq m_{l2} \quad \forall p \in P$$

Bounds on major composition ratio per project

$$\sum_{i \in S} m_{ic} s_{ip} \geq (\mu_c - t_\mu) \sum_{i \in S} s_{ip} \quad \forall p \in P, c \in M$$

$$\sum_{i \in S} m_{ic} s_{ip} \leq (\mu_c + t_\mu) \sum_{i \in S} s_{ip} \quad \forall p \in P, c \in M$$

Bounds on gender composition ratio per project

$$\sum_{i \in S} g_{ik} s_{ip} \leq (\gamma_k - t_\gamma) \sum_{i \in S} s_{ip} \quad \forall p \in P, k \in G$$

$$\sum_{i \in S} g_{ik} s_{ip} \leq (\gamma_k + t_\gamma) \sum_{i \in S} s_{ip} \quad \forall p \in P, k \in G$$

Binary decision variable

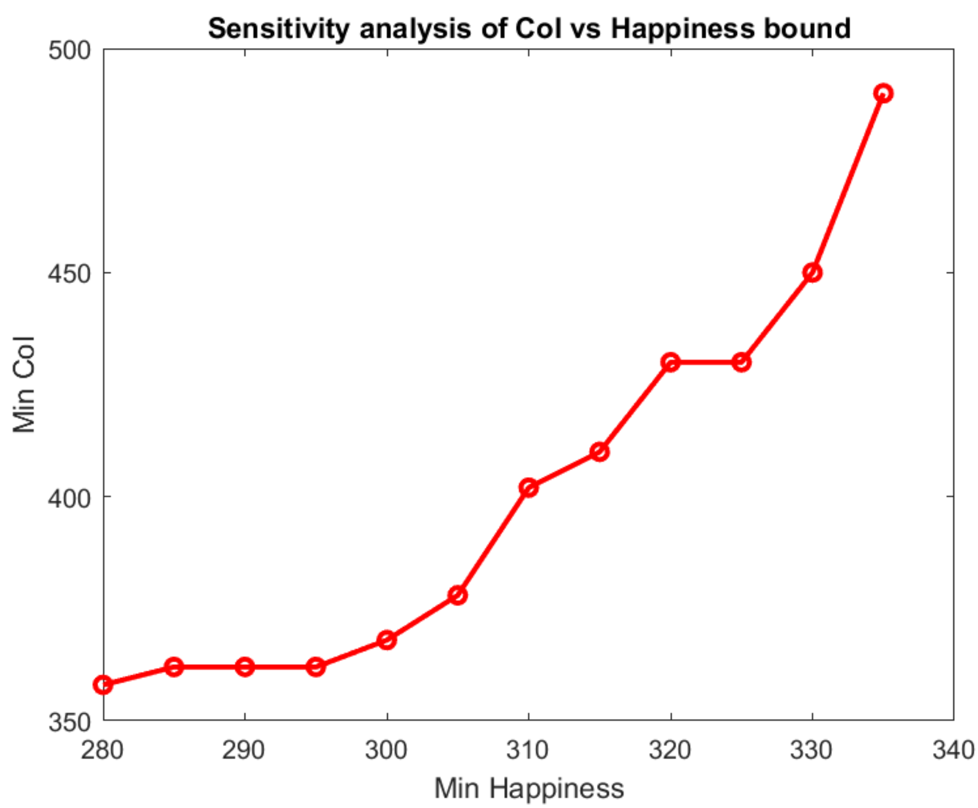
$$s_{ip} = \{0, 1\} \quad \forall i \in S, p \in P$$

Sensitivity Analysis

The model was run in AMPL for different values of H_l to investigate the effect of perturbating minimum student happiness on the total conflict of interest. The following results were obtained and synthesized into Figure 1:

Minimum Happiness	Optimal Conflict of Interest	Number of Iterations	Number of Branches
335	490	1749	0
330	450	66905	668
325	430	313637	2594
320	430	770948	7744
315	410	1736777	12431
310	402	3283768	26192
305	378	4775047	27719
300	368	4959250	35002
295	362	8420055	48761
290	362	17911591	99154
285	362	19932412	100392

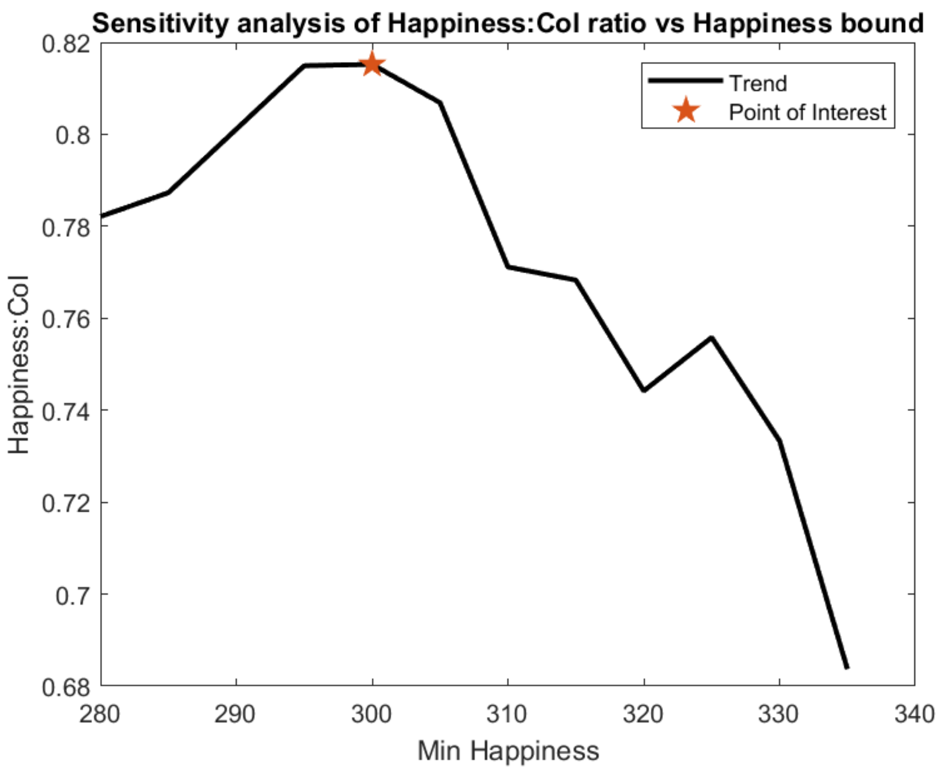
Figure 1: Results of the sensitivity analysis performed on the Happiness bound



We can observe a quasi-exponential trend from the figure, which shows that the objective function becomes increasingly sensitive to the happiness bound as it approaches an infeasible system.

It also suggests that there is a point at which the opportunity cost of increasing happiness becomes greater than its intended benefit, where the conflict of interest increases at a greater rate. A graph of $\frac{Happiness}{Conflict\ of\ Interest}$ for different values of the happiness bound verifies this.

Figure 2: A graph of (Happiness/Conflict of Interest) against Happiness



From this graph, we can deduce that the Happiness/CoI ratio is maximized when Happiness is at a relative maximum and Conflict of Interest is at a relative minimum. This occurs when $H_l = 300$. Using this value, the combined optimal solution is as follows:

Projects	1	2	3	4	5
Assigned student IDs	10, 12, 15, 18, 21, 22, 28, 38	6, 7, 13, 29, 34, 35, 36, 39	2, 3, 5, 11, 17, 20, 27, 32	4, 14, 16, 19, 23, 24, 25, 33	1, 8, 9, 26, 30, 31, 37, 40
Average GPA	3.20	3.16	3.44	3.11	3.15
Major composition	CIV: 2 IE: 4 BME: 2	CIV: 2 IE: 4 BME: 2	CIV: 2 IE: 3 BME: 3	CIV: 2 IE: 4 BME: 2	CIV: 2 IE: 4 BME: 2

Gender diversity	M: 6 F: 2	M: 6 F: 2	M: 5 F: 3	M: 4 F: 4	M: 6 F: 2

Realistic Application

Company X needs to send workers to their different branches all over the world. Each worker has a certain productivity level (rated on a scale from 1 to 5), a set revenue they make for the company, and a rating for each location (rated on a scale from 1 to 5). Each location has a quota for the amount of revenue they must generate and a certain amount they must pay their workers. Each location has a required minimum productivity and rating average for all its workers, and each location must have at least 3 workers. The company wants to maximize its total profit across all its branches.

Mathematical Formulation

- Sets
  - W: Set of workers (for every student i)
  - L : Set of locations (for every location j)

- Parameters
  - p_i : Productivity level per worker i
  - p_l : Minimum average productivity level at each location
  - m_i : Revenue generated by worker i
  - q_j : Minimum quota per location j
  - r_{ij} : Rating per worker i for location j
  - r_l : Minimum average rating at each location
  - a_j : Amount each worker must be paid at location j
  - n : Minimum number of workers at each location

- Decision variables
  - x_{ij} : Binary variable $= 1$ if worker i is assigned to location j, 0 otherwise

Linear Program:

Maximize total profit across all the locations

$$\max \sum_{i \in \mathrm{W}} \sum_{j \in L} (x_{ij} m_i - a_j)$$

s.t.

Minimum average productivity per location

$$\sum_{i \in W} p_i x_{ij} \geq p_l \left(\sum_{i \in W} x_{ij} \right) \quad \forall j \in L$$

Minimum average rating per location

$$\sum_{i \in W} r_{ij} x_{ij} \geq r_l \left(\sum_{i \in W} x_{ij} \right) \quad \forall j \in L$$

Minimum required revenue quota at each location

$$\sum_{i \in W} m_i x_{ij} \geq q_j \quad \forall j \in L$$

Each location must have a minimum number of workers

$$\sum_{i \in W} x_{ij} \geq m \quad \forall j \in L$$

Binary decision variable

$$x_{ij} = \{0,1\} \quad \forall i \in W, j \in L$$