**Freelance assignment**

The goal of this project is to develop an optimization model to match freelancing jobs with the best freelancers to do the job. We have two datasets derived from this API: https://developers.freelancer.com/docs You can check the data sets here: https://drive.google.com/drive/folders/1YsStz\_Pz0OEaz4CoZ\_YAD9m97bXeqVSJ?usp=sharing The first data set “Active Projects” lists all the active/open projects related to “Data Science” that have not been assigned to a freelancer yet (the employer posted the job and is still looking/interviewing/negotiating to find the best freelancer). The second data set “All Profiles” lists all the freelancers who have “Data science skills” along with their profile information. Both data sets were obtained from by searching the API for a number of query strings related to Data Science (queries like “Python”, “machine learning” , “statistical analysis” and so on). We want to develop a theoretical matching model that assigns each active project to a number of appropriate freelancers that are best fit to do the job. For example, the model should take in the active projects, sift through all freelancer profiles, and then matches each project with 5 to 10 freelancers that are best to do the job. When developing the model, you will have to make a business case as to why you developed the model in this certain way. This is where the real work is. You will have to think of things like: matching text queries in job and freelancer profiles, not assigning a freelancer to many jobs so it’s fair to everyone and the freelancer is not overwhelmed, assigning freelancers who have experience in big projects to jobs with bigger budgets, assigning new employers in platform to highly rated freelancers so as to make sure they have a good experience, excluding non-active freelancers from the matching. These are just random ideas that I’m throwing but you should do more research here. You can also consult the academic literature on two-sided matching and how best to do it. Check out assignment linear program and Shapley cooperative games as well. The final deliverable of this project is a proposal document in Word. The proposal should give a brief into on the problem at hand, explain the business case as outlined above, and formally write the optimization model with all sets, parameters, objective function, and constraints in Latex (check the word doc attached to see how you should write the optimization model). Also attached is full description.

**Problem 1**

Basic Box Company is considering 5 new box designs of different sizes to package 4 upcoming lines of computer monitors. The following table shows the wasted space that each box would have if used to package each monitor. Missing values indicate a box that cannot be used for a particular monitor.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Box | Monitor | | | |
| 1 | 2 | 3 | 4 |
| 1 | 5 | ­– | 10 | – |
| 2 | 20 | – | – | 25 |
| 3 | 40 | – | 40 | 30 |
| 4 | – | 10 | 70 | – |
| 5 | – | 40 | 80 | – |

Basic wants to choose the smallest number of box designs needed to pack all products and to decide which box design to use for each monitor, to minimize waste. Formulate this problem as a BIP model and provide the optimal solution. You do not need to upload your code.

Hint : Minimizing the number of box designs has a high priority compared to minimizing waste; you can treat the cost of each box design as a big M, e.g., 10,000, in the objective.

**Optimization Model:**

We can formulate the above problem as a BIP model by defining the following:

Sets:

Parameters:

The values of w can be obtained from the table above.

Variables:

Note: we create the variable Xij only if it has a valid Wij value from the table above.

Formulation:

s.t.

The objective function (1) ensures that we minimize the total wasted space from the box designs chosen for each monitor in addition to minimizing cost of box designs to be manufactured. In the objective function, we used the Big M method and applied it to the cost of establishing each box design because it’s a higher priority than minimizing the wasted space. Constraint (2) ensures that each monitor is assigned exactly one box design. Constraint (3) ensures that we cannot assign a box design to pack a monitor unless we decide to manufacture this box design in the first place. Constraint (4) is the standard binary definition of the Xij variables where they can only take the value 0 or 1. Constraint (5) is the standard binary definition of the Yi variables where they can only take the value 0 or 1.

When I coded this in Gurobi-Python, I had to slightly change the formulation to make it work with Python.

Graphical user interface, text

Description automatically generated

Specifically, I had to hard-code constraints to make Xij = 0 to any pair (i,j) that has a missing value in the wasted space table defined in the question (red rectangle).

I get the following output:

Text, letter

Description automatically generated

This means that we get the following optimal solution:

* Basic Box company will use box design 3 and 4 only for their monitors
* Monitor 1, 3, and 4 will use box design 3
* Monitor 2 will use box design 4
* The total wasted space 40+40+30+10 = 120

**Problem 2**

Top-T shirt company imprints T-shirts with cartoons and celebrity photographs. For each of their 4 pending contracts, the following table shows the number of days of production required, the earliest day the order can begin, and the day the order is due.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 |
| Production | 10 | 3 | 16 | 8 |
| Earliest | 0 | 20 | 1 | 12 |
| Due Date | 12 | 30 | 20 | 21 |

The company wants to design an optimal schedule assuming that contracts can be processed in any sequence, but that production cannot be interrupted once a job has started.

1. Formulate an MIP model to minimize the mean completion time of all contracts. Provide the optimal solution, but you do not need to upload your code.

**Optimization Model:**

We can formulate the above problem as a MIP model by defining the following:

Sets:

Parameters:

Variables:

Formulation:

s.t.

The objective function (1) is to minimize the mean completion time of all contracts. Constraint (2) ensures the completion time of each contract is greater than or equal to its processing time plus the earliest date in which production can start for that contract. Constraint (3) and (4) are disjunctive constraints which enforce that either contract j is processed before contract k or contract k is processed before contract j for any pair of contracts. We use the big M method in these two constraints to enable the disjunctive feature. Constraint (5) is the non-negativity variable definition of Cj. Constraint (6) is the binary 0,1 variable definition of Yjk.

I solved this problem by using Gurobi-Python and I got the following results:

Text

Description automatically generated

We can summarize our results and provide the optimal schedule in the following table:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Contracts | 1 | 2 | 3 | 4 |
| Start Day | 0 | 21 | 24 | 13 |
| Completion Day | 10 | 24 | 40 | 21 |
| Mean Optimal Completion Time | | | 23.75 days | |

We can see that we will process contract 1 first then contract 4 then contract 2 then contract 3.

1. How would you change the model if the objective is to minimize the maximum lateness of all contracts? Write your answer, but you do not need to solve the model.

**Optimization Model:**

If we consider maximum lateness as our objective to minimize, first we have to define it.

So,

is the lateness for each job. It’s the maximum of 0 and the difference between its completion time and due date.

is the maximum value when considering all jobs.

Now, we can define our new objective function as:

We will still have the same sets, variables, and parameters as defined in part (a).

As for the constraints, we will also have all constraints from (2) to (6) but we will add one more constraint as follows:

Constraint (7) ensures that we have the proper definition of maximum lateness embedded in the MIP. It will either be 0 which means the contract will not be late past its due date or if it’s late then we will have some positive value indicating how many days late past the due date.