DASC 3203: Optimization Methods in Data Science
Course Project

Deadline: May 10

1 Overview and Purpose

The purpose of this assignment is to solve a realistic problem by applying optimization. In order to be successful on this assignment, each student must identify a problem that can be solved using optimization, formulate an optimization model to solve the problem, gather data to populate the optimization model, and then use optimization solvers in order to perform analysis of the problem.

2 Identify a Problem

Select an optimization problem that will be the focus of your project. The following list contains some ideas that may help in identifying a topic for your project:

A. Choose a problem related to your DASC concentration. For example:

   If you are pursuing the geospatial data analytics, operations analytics, or supply chain analytics concentration, you might seek to locate a collection of facilities (you choose which type) to provide service to geographically dispersed areas and/or population.

   If you are pursuing the business data analytics or accounting analytics concentration, you might seek to optimize the allocation of funds in an investment portfolio.

B. Choose an established problem and/or application in the operations research literature. For recent applications, I suggest taking a look at some of the papers recently published in the *INFORMS Journal on Applied Analytics*. You could also consider looking in the other INFORMS Journals if you'd like.

C. Choose an idea based on your own interests or hobbies. For some ideas, take a look at some recent projects that were completed by students at the Ohio State University (see https://u.osu.edu/davanloo/teaching/).

If your idea is based on someone else's work, you must clearly cite it. In addition, your project must go beyond simply reproducing what someone else's work has already done. I have provided the following guidelines to clarify:

- You may (with citation) use someone else's work as the motivation for your project.
- You may use (with citation) an optimization model that is published in the peer-reviewed academic literature; otherwise, you must formulate your own optimization model.
- You may not (under any circumstances) use someone else's implementation of an optimization model. You are required to implement the model yourself.
- You may not choose an optimization model that is covered in a course you are taking or have taken.

Whatever you choose, I strongly recommend each of you to discuss your project topic with me before investing a significant amount of time in completing it.

3 Define the Problem

Before modeling your problem, you should define and scope your problem. That is, you should explain the motivation and/or application for your problem and make the idea of the problem as clear as possible without math.

4 Formulate an Optimization Model

Formulate a mathematical optimization model for the problem you described in Section 3. The model should be stated in general terms (i.e., no input data specified) so that any problem instance could be solved using your formulation. Clearly explain the idea of the model, and make sure you are also able to answer questions about what each component of the model does.

It is understood that you may have to make some assumptions and/or simplifications along the way to ensure you can populate and solve the model. Be sure to state and justify any assumptions and/or simplifications you have made.

5 Gather and Synthesize the Data

How will you populate the parameters of your model? Be sure to cite the sources of your data, explain what information is contained in the data, and explain in detail how the data is used to populate the parameters of your model. Within reason, it is expected that students will use high-quality data sources that are of a reasonable scale. If reasonable for your problem, you should provide visualizations to assist in understanding the data.

6 Solve and Analyze the Model

Implement and solve the model(s) using an appropriate solver. Communicate and interpret the results within the context of the original problem. For example, (i) are the results sensible and (ii) does the optimized solution improve simple benchmark solutions that you could have evaluated without optimization or by using a simpler optimization model?. Are there generalizable insights about the structure of an optimal solution (e.g., where facilities tend to be located)? How would the results change in response to a change in the parameters? You many need to solve a number of instances of the problem to answer this question thoroughly.

If reasonable for your problem, you should provide visualizations of model output to support your analysis.

7 Deliverables

A. During the final exam period, each student will deliver a 10-minute presentation to communicate the results of the project. The presentation must include:

- The background and motivating information required for the audience to understand the problem and its importance.

- The formulation of the optimization model.
- A summary of the data used to populate and solve the model.
- A summary and discussion of the results obtained from solving the model.

In preparing for the presentation, students must adhere to the following rules:

1. Presentations are expected to last at least 9 minutes but no more than 11 minutes.
2. To the extent possible, slides should be created using the assertion-evidence style (https://www.assertion-evidence.com/).
3. The first slide must include the following information:
    a. Project title
    b. Student full name
    c. Date of the presentation

4. Each subsequent slide must include the slide number in the footer.
5. The final slide should include 3-4 key takeaways from the presentation.

To facilitate a high level of quality for these presentations, each student will submit draft presentation slides on the following dates:

April 12	Draft Problem Definition, Model Formulation, Summary of Data
May 3	Draft of Model Solution and Analysis

B. Each student must submit all data and share a link to the code used to build and solve the model. Each student should also submit a final copy of the slides used for the presentation. These materials must be uploaded to Blackboard before the project deadline.

8 Rubric

Each student's project will be graded according to the following rubric:

	Unsatisfactory (0 – 4)	Needs Improvement (5 – 6)	Adequate (7 – 8)	Excellent (9 – 10)	Grade
Define/scope the problem		Both of {motivation, problem idea} are unclear	One of {motivation, problem idea} is unclear or not compelling	Both of {motivation, problem idea} are clear and compelling	
Model formulation		Numerous errors with correctness, not stated in general terms	Few errors, stated in general terms	Correct, stated in general terms	
Gather and synthesize the data		Lacking in quantity/quality or not relevant to the problem	Relevant data but lacking in quantity, errors in citation	Thorough and relevant to the problem, appropriately cited	
Solve and analyze the model		One instance solved successfully, no significant insights or observations drawn	Fewer instances solved, provides basic observations about solutions	Instances solved correctly, provides significant insights into the specific problem	
Presentation content		Factual errors, inaccurate or incomplete descriptions, conclusions not justified by analysis	Conclusions justified by supporting analysis, but descriptions contain some inaccuracies	Strong conclusions justified by supporting analysis, all descriptions accurate	
Quality of slides		Problems with organization, assertions, and/or evidence that detract from communicating the main ideas	A few problems with organization, assertions, and/or evidence, but effectively communicates the main ideas	Assertions and evidence are strong, follows a logical organization	
Delivery of presentation		Problems with multiple items in the "excellent" list	Problems with some of the items in the "excellent" list, but overall delivery still mostly effective	Confident/ enthusiastic/ knowledgeable, organized and effective with transitions, well-rehearsed but extemporaneous	
Creativity		Directly applies well-known models	Attempts to customize ideas/ models/ datasets/ experiments/ etc. based on the needs of the specific project	Successfully customizes ideas/ models/ datasets/ experiments/ etc. based on the needs of the specific project	