

SAS Optimization Challenge: Case Study

Title: **HELP YOUR OPTIMIZATION PROFESSOR!!!!**

TIER 3 (of 5)

Tier 3: Modeling Details

The model that you will build in **Tier 3** builds off the model you recently completed in **Tier 2**. Use the SAS program you submitted for **Tier 2** as the starting point for **Tier 3**.



The **Tier 3** objective is to force certain questions into the exam, regardless of their difficulty.

Your professor always includes the following eight (8) questions in his final exam, and he wants to make sure they're included in this semester's final exam:

ITEM_01.01.01
ITEM_02.02.02
ITEM_02.02.07
ITEM_03.01.02
ITEM_03.03.10
ITEM_04.01.07
ITEM_04.02.10
ITEM_05.02.07

Your objective in **Tier 3** is to choose the fifty (50) questions that produce the most difficult exam possible, while ensuring that each *topic objective* contains the correct number of questions, ensuring that no more than one question in each frenemy group is assigned to the exam, and **ensuring the eight (8) questions above are on the final exam.**

Use the `create data` statement to generate one (1) output SAS data set containing only the fifty (50) questions that your model chose for the final exam. Include the following columns in the output data set: **Question**, **Frenemy_Group**, **Topic**, **Objective**, **Force_Status**, and **Correct**.

The **Force_Status** column is explained below in the **Tier 3: Getting Started** section. Questions that were not chosen should not appear in your output data set.

Additionally, report the average percentage correct (i.e., take the average of the **Correct** column) across all fifty (50) questions chosen by your model. This value can be calculated by any means necessary (e.g., in SAS, Excel, calculator, etc.).

To do this in SAS, copy and paste the following SQL procedure code at the end of your program, and replace *<insert output data set name>* with the name of your output data set that you created in the OPTMODEL procedure.

```
proc sql;
  select avg(Correct) as Avg_Correct
  from <insert output data set name>;
quit;
```

Tier 3: Getting Started

Since **Tier 3** builds off the solution developed in **Tier 2**, copy and paste your **Tier 2** solution (i.e., DATA step and datalines code creating the two data sets and OPTMODEL code) into a new programming editor to begin.

In-between the DATA step/datalines code and your OPTMODEL code, copy and paste the following DATA step:

```
data work.question_bank;
set work.question_bank;
  if Question in('ITEM_01.01.01',
                 'ITEM_02.02.02',
                 'ITEM_02.02.07',
                 'ITEM_03.01.02',
                 'ITEM_03.03.10',
                 'ITEM_04.01.07',
                 'ITEM_04.02.10',
                 'ITEM_05.02.07')
  then force_status = 'force';
  else force_status = ' ';
run;
```

This code adds a column called **Force_Status** to the question bank data set. Questions with a **Force_Status** = 'force' must appear in the final exam. Questions with **Force_Status** set to missing are not required to appear in the final exam.

Tier 3: Hints

The hints below are provided to help overcome relatively nuanced programming syntax and model formulation barriers that may arise in [Tier 3](#).

- I. The overwhelming majority of parameters used to build optimization models are numeric, which are declared using the `num` statement inside of the OPTMODEL procedure. However, the **Force_Status** column is character, so you'll need to declare it using `str`.

(e.g., `str force_status{QUESTIONS};`)

- II. Since these eight (8) questions do not need to be determined by the optimization solver, consider using the `fix` statement within a `for{}` loop to force these questions into the final exam.