Instructions:

We will practice modeling non-linear optimization problems in this lab.

Whenever required, we will model problems by loading information from files in this lab.

Please follow the instructions given below:

- Please use different notebooks for solving different problems.
- The notebook name for Exercise 1 should be `YOURROLLNUMBER_IE507_Lab8_Ex1.ipynb`.
- Similarly, the notebook name for Exercise 2 should be `YOURROLLNUMBER_IE507_Lab8_Ex2.ipynb`, etc.
- Please ask your doubts so that TAs can clarify.

For more details on pyomo, please consult https://pyomo.readthedocs.io/en/stable/index.html.

There are only 2 exercises in this lab. Try to solve all problems on your own. If you have difficulties, ask the Instructor or TAs.

Only the questions marked **[R]** need to be answered in the notebook. You can either print the answers using `print` command in your code or you can write the text in a separate text tab. To add text in your notebook, click `+Text`. Some questions require you to provide proper explanations; for such questions, write proper explanations in a text tab.

After completing this lab's exercises, click File → Download .ipynb and save your files to your local laptop/desktop. Create a folder with name `YOURROLLNUMBER_IE507_Lab8` and copy your `.ipynb` files to the folder. Also copy the `.csv` files to the folder. Some questions require the appropriate files to be included in folder. Please include all related files required to execute your code in the folder. Then zip the folder to create `YOURROLLNUMBER_IE507_Lab8.zip`. Then upload only the `.zip` file to Moodle.

The deadline for today's lab submission is **tomorrow, 11 59 PM Indian Standard Time (IST)**.

Exercise 1: Solving a non-linear optimization problem with non-linear constraints [15 marks]
Consider the following optimization problem **(OP)**:

$$
\begin{aligned}
\min\ & f(x_1,x_2,x_3,x_4,x_5) = (x_1-1)^2 + (x_1-x_2)^2 + (x_2-x_3)^2 + (x_3-x_4)^4 + (x_4-x_5)^4, \\
\text{s.t. } & x_1 + x_2^2 + x_3^3 - 2 - 3\sqrt{2} = 0, \\
& x_2 - x_3^2 + x_4 + 2 - 2\sqrt{2} = 0, \\
& x_1 x_5 - 2 = 0, \\
& -5 \le x_1 \le 5, \\
& 0 \le x_3 \le 3.
\end{aligned} \tag{OP}
$$

1. Write a `pyomo` model to solve the optimization problem (OP).
2. Use a starting point $(x_1,x_2,x_3,x_4,x_5) = (2,2,2,2,2)$ and use `ipopt` solver to solve optimization problem (OP).
3. **[R]** Report the solver status, time taken to solve and solver termination condition.
4. **[R]** Report the optimal objective function value and optimal values of the optimization variables.
5. **[R]** Report if the constraints are active or inactive at the optimal solution.
6. **[R]** Solve the model using the following choices of starting points: $(x_1,x_2,x_3,x_4,x_5) = (2^i,2^i,2^i,2^i,2^i)$ for $i = 2,4,8,10,12$. Tabulate the optimal solution, optimal values of optimization variables and time taken to solve, for each i. Comment on your observations.

Exercise 2: Fitting data using a linear model. [25 Marks]
μtube video hosting website allows users to upload 10 minute long videos that they have created. Since the service is free for general users, μtube introduces some ads into user videos to sustain its business. Every ad has a skip option and can be skipped without watching by a viewer. To motivate users to upload many videos, μtube supports a revenue model where the users who upload videos are paid in a currency called μmoney, which the users can use for some online purchases. The μmoney is deposited daily in the uploader's account based on the following two factors:

- average time spent per view by μtube viewers on a video
- average number of ads watched without skipping, per view

The following data is made available for a few videos that you have uploaded in μtube. The data corresponds to October 11, 2022.

Video i	μmoney R_i	Views V_i	Total view-time B_i	Total unskipped ads A_i
1	175	945	2800	23
2	450	832	3900	40
3	1050	2007	24900	78
4	900	800	36450	64
5	19	780	366	4
6	3550	5005	38500	150

Table 1: μmoney for different videos

You wish to build a model which can forecast the potential revenue. You start with an assumption that a linear model of the form $R = a_0 + a_1T + a_2D$ would be useful, where R denotes the daily revenue, T denotes the daily average view-time per view, D denotes the daily average number of unskipped ads per view. You believe that this model would be reasonably accurate and easy to use. Here a_1 and a_2 would indicate how the revenue varies with average view-time and average number of unskipped ads, respectively, while a_0 would reflect a base revenue for your videos. You would like to select the *best* linear model in some sense. If you knew the three parameters a_0, a_1 and a_2, the six observations in the table would each provide a forecast of the revenue as follows:

$$\widetilde{R}_i = a_0 + a_1T_i + a_2D_i \quad i = 1, 2, \ldots, 6.$$

However, since a_0, a_1 and a_2 cannot, in general, be chosen so that the actual revenues R_i are exactly equal to the forecast revenues $\widetilde{R}_i$ for all observations, you would like to minimize the sum of $|Q_i|^p$ obtained over the videos, where $Q_i = R_i - \widetilde{R}_i$ for some $p \geq 1$.
In this lab, we will solve this problem for 6 data points in Table 1.

1. **[R]** Fix $p = 1$. Write a **linear** optimization problem using the objective provided in the description. You may assume that $a_0 \geq 0$ and a_1, a_2 do not have bounds on them. (**Hint:** Recall that $|x| \geq x$ and $|x| \geq -x$ for any real number x.)
2. Create an appropriate `.csv` file from Table 1, which you will use during the construction of the objective function and the constraints for your linear program.
3. Construct `pyomo` model for the linear program you have formulated.
4. **[R]** Solve the linear program using `cbc` solver, report the solver status, solver termination condition, optimal solution values for the decision variables, and the optimal objective function value. Comment on your observations.

5. [**R**] Now, consider $p > 1$. Write a general **nonlinear** optimization problem for $p > 1$. You may assume that $a_0 \geq 0$ and a_1, a_2 do not have bounds on them.
6. Construct `pyomo` model for the non-linear optimization model you have formulated.
7. [**R**] Solve the respective non-linear optimization problems for $p = 1.5, 2$ using `ipopt` solver, and for each p value report the solver status, solver termination conditions, time taken by solver, optimal solution values for the decision variables, and the optimal objective function value. Comment on your observations.
8. [**R**] For each value of $p \in \{1, 1.5, 2\}$, prepare different plots depicting the actual revenue R_i and the forecast value $\widetilde{R}_i$ obtained using your optimization models for the data provided. Based on the plots, comment with suitable justifications if the forecast model that you have assumed would be good fit for the given data.
9. [**R**] Based on your results and plots, suggest using appropriate reasons which value of p is good for your problem among the three different choices you tried.
10. IMPORTANT: Do not forget to upload the `.csv` file with your submission.