

40.014 Engineering Systems Architecture

Term 5, 2020

Homework 1

Due date: April 5, 2020, 11:59 pm

Note. You must complete the homework on your own and submit it individually on eDimension (in a **single PDF file**). Put your name at the beginning of the file. Please use complete English sentences to explain the results.

Problem description. You are working in collaboration with the organizers of the “Tour de France” (an annual men’s multiple-stage bicycle race primarily held in France) to design the route for the 2021 edition. With the aid of Google maps, you designed 1,153 different routes (or alternatives) that must be evaluated with respect to three criteria: (1) the racing fatigue (to be minimized), which can lead to injuries in such a long race; (2) the number of expected crashes (to be minimized) during the race; (3) the disruption caused to the regular traffic (to be minimized).

All data are contained in the file *Tour_de_France.csv*. The columns contain the values of the racing fatigue (“Fatigue”), expected crashes (“Crashes”), and disruption caused to the regular traffic (“Disruptions”), respectively. Each row represents a different route, or alternative. Note that the values of the criteria have different ranges of variability.

Task 1. Using R and ggplot2, produce three 2D scatter plots to visualize: (1) Fatigue (x-axis) vs Crashes (y-axis), (2) Fatigue (x-axis) vs Disruptions (y-axis), (3) Crashes (x-axis) vs Disruptions (y-axis). Save and attach the three scatter plots.

Task 2. Describe with your own words the relation between the three objectives. In other words, what do the 2D scatter plots show?

Task 3. Using R and ggplot2, produce one 2D scatter plot to visualize: Fatigue (x-axis) vs Crashes (y-axis) vs Disruptions (color of the dots). Save and attach the scatter plot.

Task 4. The plot generated at Task 3 should provide a clear illustration of the trade-off between the three objectives. Such illustration can also help us identify groups, or clusters, of solutions that are dominated. Can you identify any group? Explain your reasoning.

Task 5. Use the *ParetoSorting_adv.R* function (provided in the folder *Homework 1*) to identify non-dominated (or Pareto-efficient) and dominated alternatives. How many alternatives are non-dominated? How many alternatives are dominated?

Task 6. Produce a parallel coordinate plot to visualize non-dominated and dominated alternatives. Order the axis in the following way: Fatigue, Crashes, Disruptions. Use colors to differentiate dominated and non-dominated solutions. Use the scaling option "uniminmax". Save and attach the parallel coordinate plot.

Task 7. Determine the coordinates of Ideal Objective (Utopia) and Nadir vectors.

Task 8. What is the 'best' alternative according to the Utopia point method (report the index of the best alternative across all solutions)? What is the corresponding minimum distance? (Hint: remember that the Utopia Point method works best with normalized data.)

Task 9. Using R and ggplot2, produce one 2D scatter plot to visualize: Fatigue (x-axis) vs Crashes (y-axis) vs Disruptions (size of the dots). For the sake of simplicity, use the normalized data produced at Task 8, and work only with the Pareto-efficient solutions. Use different colors to identify the Utopia point, the selected solution, and the remaining Pareto-efficient solutions. Save and attach the scatter plot.

Task 10. Produce a radar chart to visualize the 'best' alternative found in Task 8 (use the normalized data). The radar chart should include all objectives, namely Fatigue, Crashes, and Disruptions. Save and attach the radar chart.