SMDE Second assignment

# Executive summary

We are going to simulate the behavior of different groups of runners in Boston Marathon, based on real data taken from 2017 episode.

In particular we are going to simulate a subgroup of all the runners, men from 30 to 40 years old and women from 25 to 40 years old.

This simulation is going to be useful for planning all different resources for next version of the same event, such as Water Spots, WC, Solid meals and so on.

# System description, introduction

A marathon in a competition that consist in running 42.195 Km distance with a maximum time of 6 hours.

At the beginning line runners are divided in groups depending on their performance category, for example those runners who can run in 2:30 minutes/kilometer, 3, 3:30 and so on. The ones who perform the better start at the beginning.

On different spots alongside the marathon the organization service is going to provide different resources to runners in order to support the competition. This resources are multiple and can vary according to the organization budget and preferences. Such resources can be: Water, Solid meals like fruits, Glucose gels, WC, Medical spots, etc.

In our case we are going to do certain assumptions and reductions that can be found in the different hypotheses on the following sections.

On the other hand to have a successful simulation for a better analysis it is important to be based on previous statistical data from previous version of the same competition.

# Problem description

In our model, the main goal is to find the right amount of resources in order to avoid collisions and waiting queues in the running line.

At the beginning we have found that having a few amount of resources in Water spots, less that 50 per sport, led to issues and long queues for runners.

We have set this number to 100 resources of water per spot.

Since WC and Solid is less frequently use for runners, and our distribution function was set on that sense, we haven’t found an issue with the resources in that sense. Having for example 8 WC on the WC spots is enough. Also having 100 resources in the Solid spot and with 1 spot only is enough for the whole run.

## Systemic Structural, Systemic Data and Simplifying Hypotheses

Complete the problem description with the hypotheses. Usually they are going to be presented with an identifier.

SH\_01 Since we cannot predict the weather with exact precision we are not taking into account the weather as a factor

SH\_02 We are not going to measure or simulate injured people. We assume that anyone is injured or abandon the marathon

SH\_03 We are going to reduce the amount of water, wc and solid spots in order to simplify DOE after and not to have so many factors.

SH\_04 Where there is a 2 colliding points, for example water and solid or water and bathroom, we are going to distribute the runners between the 2 points. There is not going to be runners that go to both resources.

SS\_01 The marathon cannot last more than 6 hours. The system will be consider finished after 6 hours.

SS\_02 The total distance of the marathon is 42.195 Km

SS\_03 Time performance for each runner is going to be taken every 5 km.

SS\_04 We are going to analyze 2 groups, Male 30-40 years old and Women of 25-40 years old.

SD\_01 We are going to use a linear regression model for predict the time that takes to each runner to cross each segment

SD\_02 Elite runners are going to be calculated as a percentage based on Boston Marathon 2017 on each group

SD\_03 Elite runners are going to take less to cross each performance measurement based on the same percentage defined in SD\_02

SD\_04 There is going to be 3 water points at Km 10, 21 and 30, and it can deliver 100 unit of water at time

SD\_05 There is going to be a fruit and glucose point at Half marathon and 35km and on each point there is going to be capacity of 500 units at time.

SD\_06 There is going to be 2 bathroom points at Km 25 and 30.

# Model specification

The model entities, operations and processes that defines the behavior of the model. We can use DEVS, Petri Nets or SDL here. Since we are not going to use those formal languages, define a flow diagram to simplify the definition of the model. If using GPSS you can use the GPSS icons of the language.

# Codification

Check file marathon\_boston.gps

## Data.

The data used along the model was:

* boston\_marathon\_men\_3040.csv and boston\_marathon\_men\_3040.xlsx which is the dataset with the marathon of 2017 for that specific group
* InputMenExperiment.txt which is the input file with the different experiments for the Factorial analysis

# Definition of the experimental framework

Explain the DOE to be used. Detail the process to execute the replications.

1. Define a DOE to explore with what parametrization of the factors the answer obtains the best value (define what means best, i.e. maximize or minimize the value).
2. Detect and analyze the interactions.

# Model validation

Propose some methods to perform a validation of your model. The validation will be done in the third assignment, hence here we will only propose the methods.

# Results /Conclusions

Be as concise as possible, describe the main conclusions of the analysis done.