

# Travelling salesman problem

Miguel Ángel Juárez Garzón

## Libraries and data load

```
library(GA)
```

```
data("eurodist", package = "datasets")  
D <- as.matrix(eurodist)
```

## Exercise 2: Fitness function definition

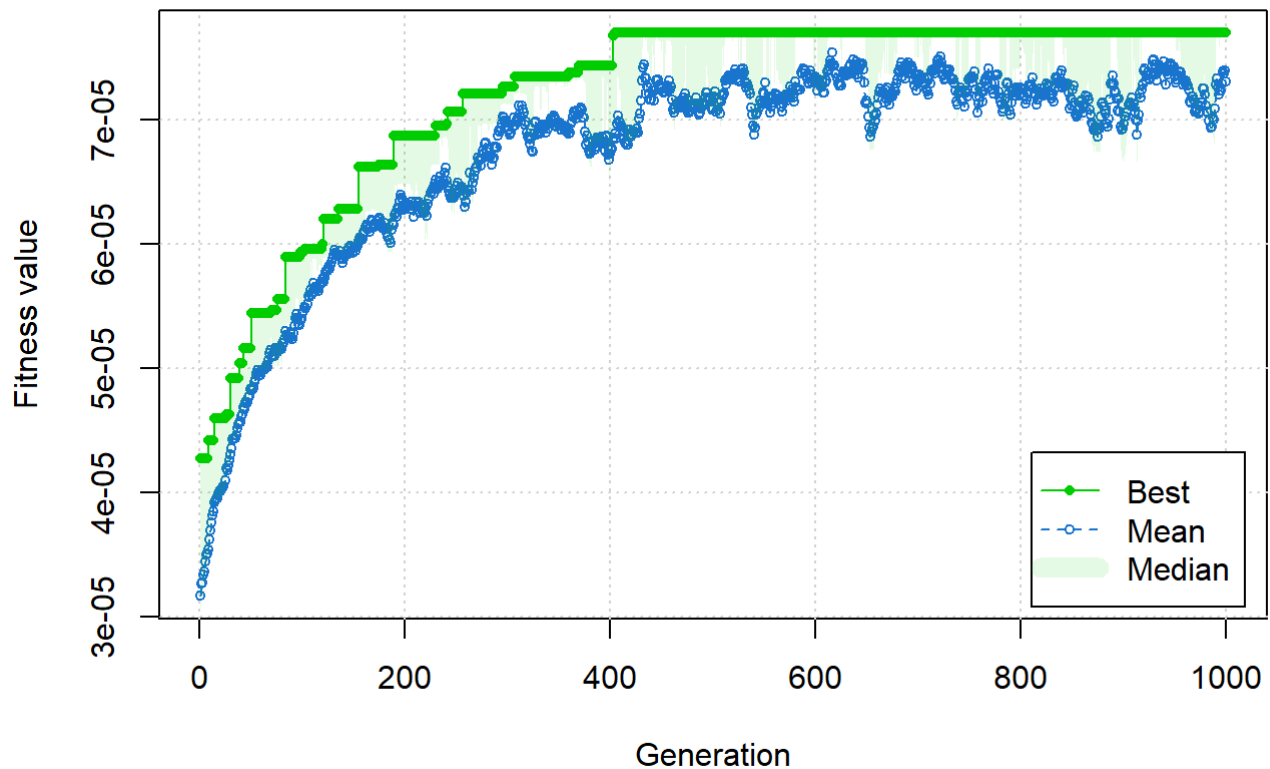
```
# Given a route calculate the total distance  
dist_ruta <- function(tour, distMatrix) {  
  tour <- c(tour, tour[1])  
  ruta <- embed(tour, 2)[, 2:1]  
  sum(distMatrix[ruta])  
}  
# Fitness function  
tpsFitness <- function(tour, ...) 1/dist_ruta(tour, ...)
```

## Exercise 3: Optimal route via GA algorithm

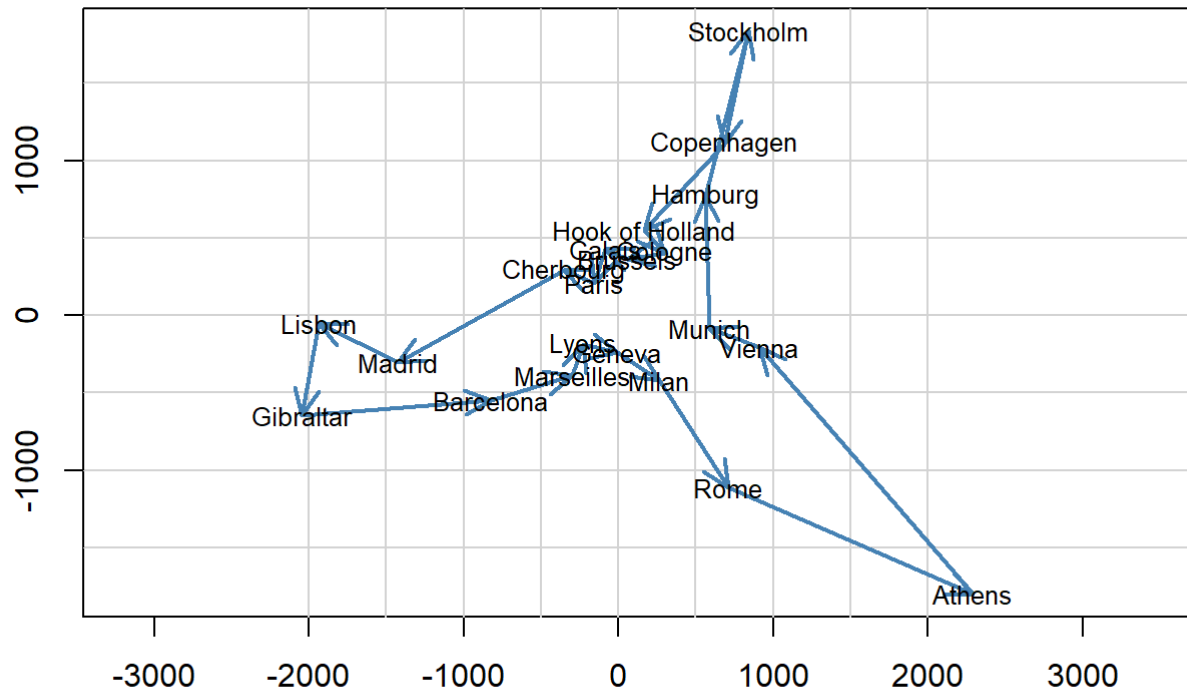
```
# GA Algorithm  
GA <- ga(type = "permutation", fitness = tpsFitness, distMatrix = D, lower = 1, upper  
= attr(eurodist, "Size"), popSize = 100, maxiter = 1000, pmutation = 0.1, pcrossover  
= 0, monitor = NULL)
```

## Plotting

```
plot(GA)
```



```
mds <- cmdscale(eurodist)
x <- mds[, 1]
y <- -mds[, 2]
plot(x, y, type = "n", asp = 1, xlab = "", ylab = "")
abline(h = pretty(range(x), 10), v = pretty(range(y), 10),
col = "light gray")
tour <- GA@solution[1, ]
tour <- c(tour, tour[1])
n <- length(tour)
arrows(x[tour[-n]], y[tour[-n]], x[tour[-1]], y[tour[-1]],
length = 0.15, angle = 25, col = "steelblue", lwd = 2)
text(x, y, labels(eurodist), cex=0.8)
```



If we plot the fitness, we can observe how it converges around generation 400 approximately. There are small fluctuations in the average fitness between generations.

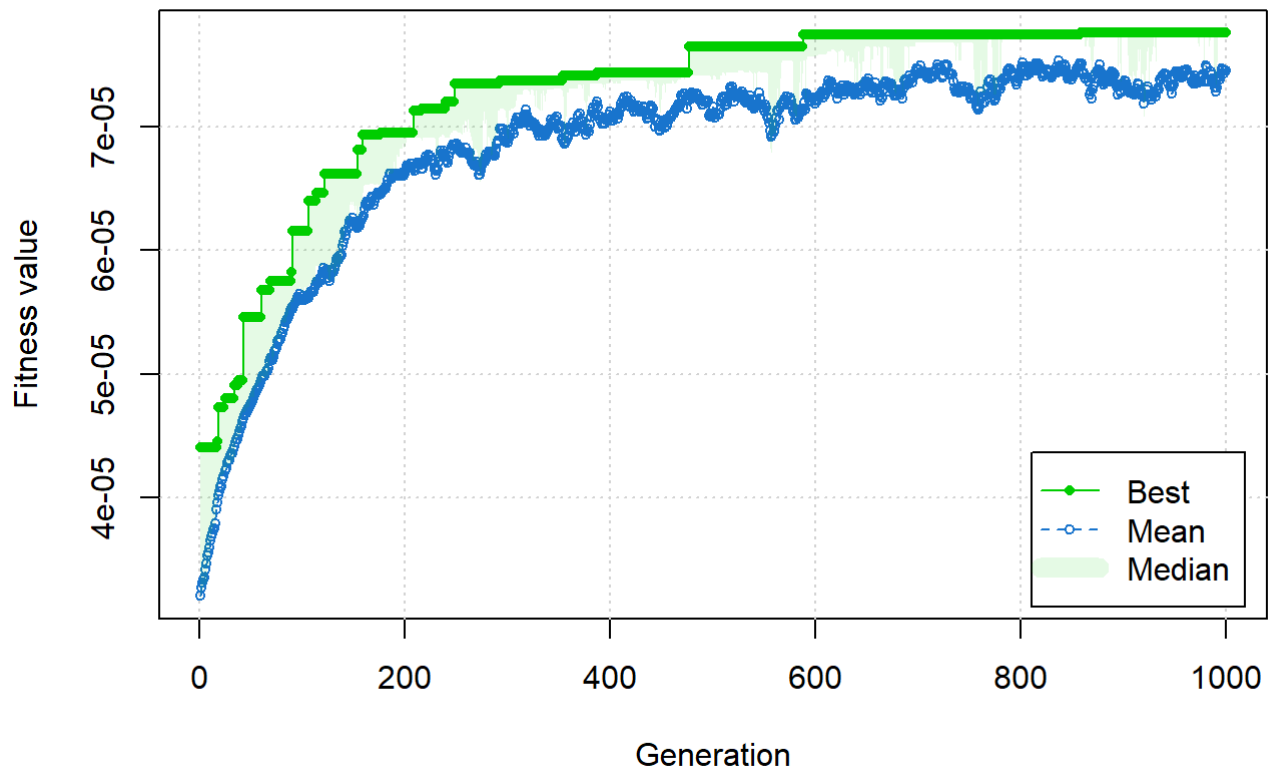
Furthermore, the route with the highest fitness is a route that traverses the cities in a counterclockwise direction, forming a circular-like path

## Exercise 4: Effect of parameters on GA variation

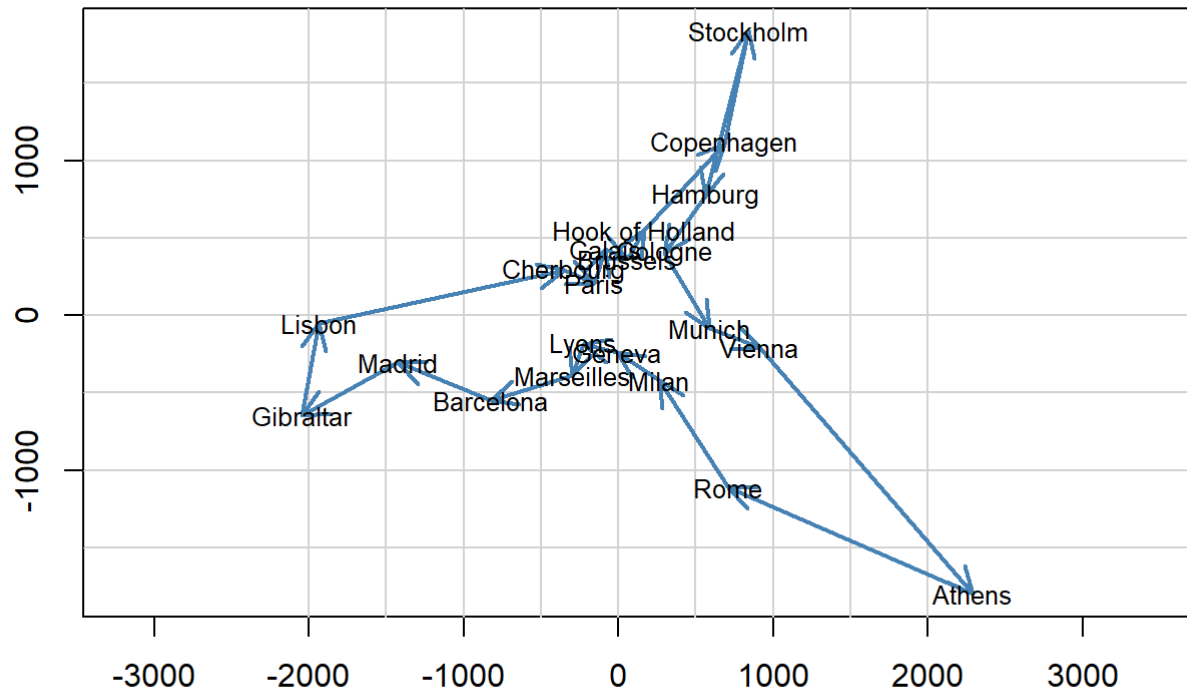
### Sample size duplication

```
GA <- ga(type = "permutation", fitness = tpsFitness, distMatrix = D, lower = 1, upper = attr(eurodist, "Size"), popSize = 200, maxiter = 1000, pmutation = 0.1, pcrossover = 0, monitor = NULL)
```

```
plot(GA)
```



```
mds <- cmdscale(eurodist)
x <- mds[, 1]
y <- -mds[, 2]
plot(x, y, type = "n", asp = 1, xlab = "", ylab = "")
abline(h = pretty(range(x), 10), v = pretty(range(y), 10),
col = "light gray")
tour <- GA@solution[1, ]
tour <- c(tour, tour[1])
n <- length(tour)
arrows(x[tour[-n]], y[tour[-n]], x[tour[-1]], y[tour[-1]],
length = 0.15, angle = 25, col = "steelblue", lwd = 2)
text(x, y, labels(eurodist), cex=0.8)
```



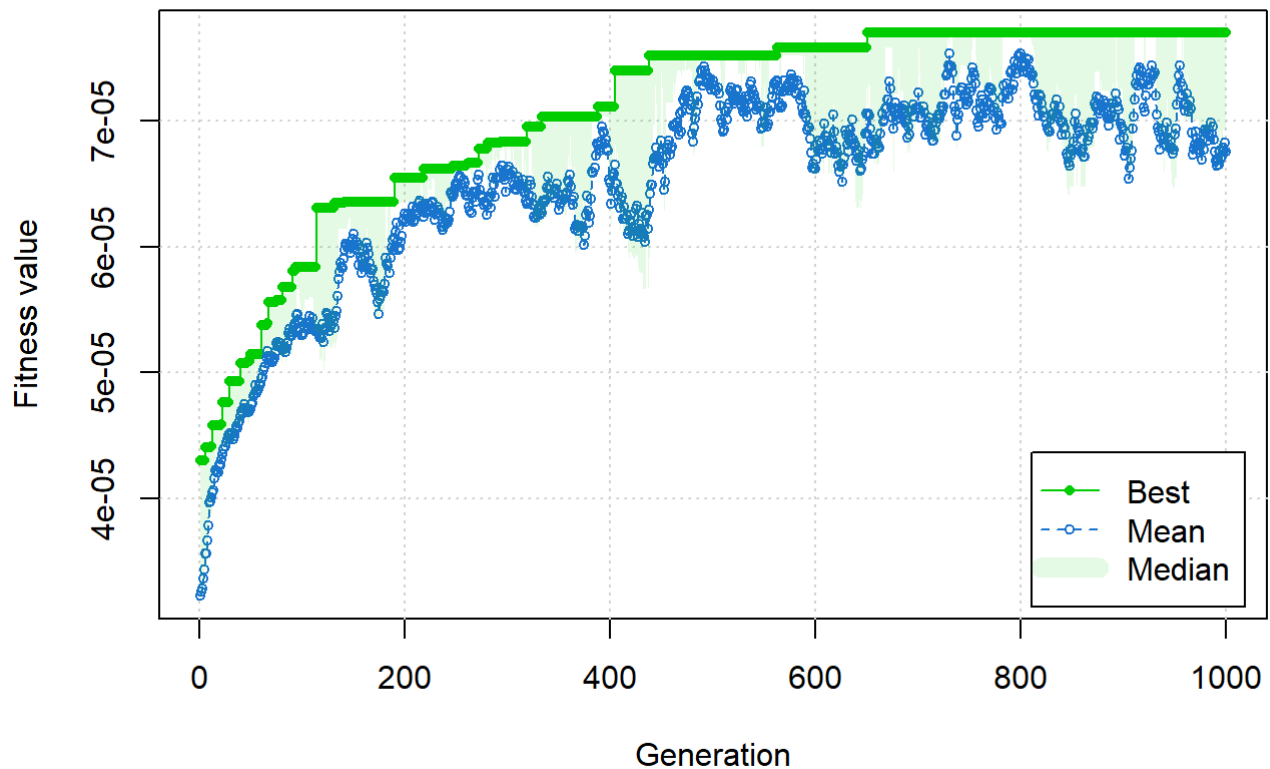
In this case, we can observe that the highest fitness value settles in the same point as with a smaller population. However, the fluctuations in the average fitness are smaller compared to a smaller population.

Regarding the route, we can see that in this case, it is the same route as in the previous case but in the reverse direction.

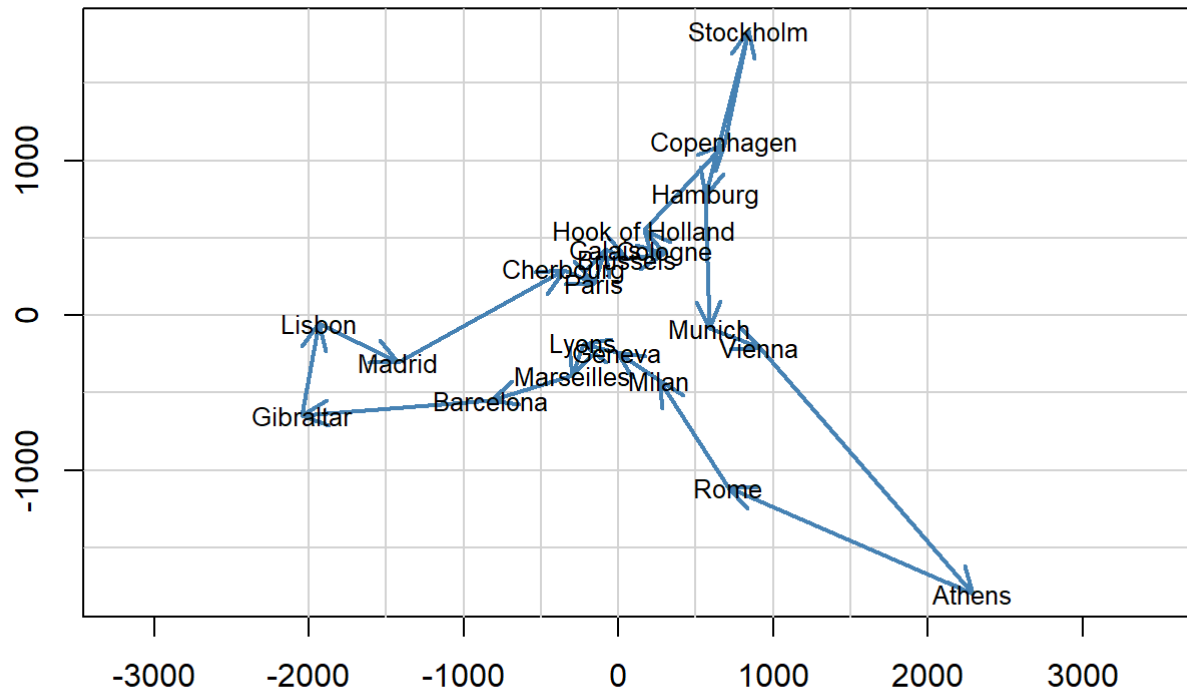
### Reduction of sample size by half

```
GA <- ga(type = "permutation", fitness = tpsFitness, distMatrix = D, lower = 1, upper = attr(eurodist, "Size"), popSize = 50, maxiter = 1000, pmutation = 0.1, pcrossover = 0, monitor = NULL)
```

```
plot(GA)
```



```
mds <- cmdscale(eurodist)
x <- mds[, 1]
y <- -mds[, 2]
plot(x, y, type = "n", asp = 1, xlab = "", ylab = "")
abline(h = pretty(range(x), 10), v = pretty(range(y), 10),
col = "light gray")
tour <- GA@solution[1, ]
tour <- c(tour, tour[1])
n <- length(tour)
arrows(x[tour[-n]], y[tour[-n]], x[tour[-1]], y[tour[-1]],
length = 0.15, angle = 25, col = "steelblue", lwd = 2)
text(x, y, labels(eurodist), cex=0.8)
```



We can see that the best fitness value again converges to a similar point as in the cases with a larger population. However, we can see that the average fitness in this case is lower and has higher fluctuations compared to the cases with a larger population.

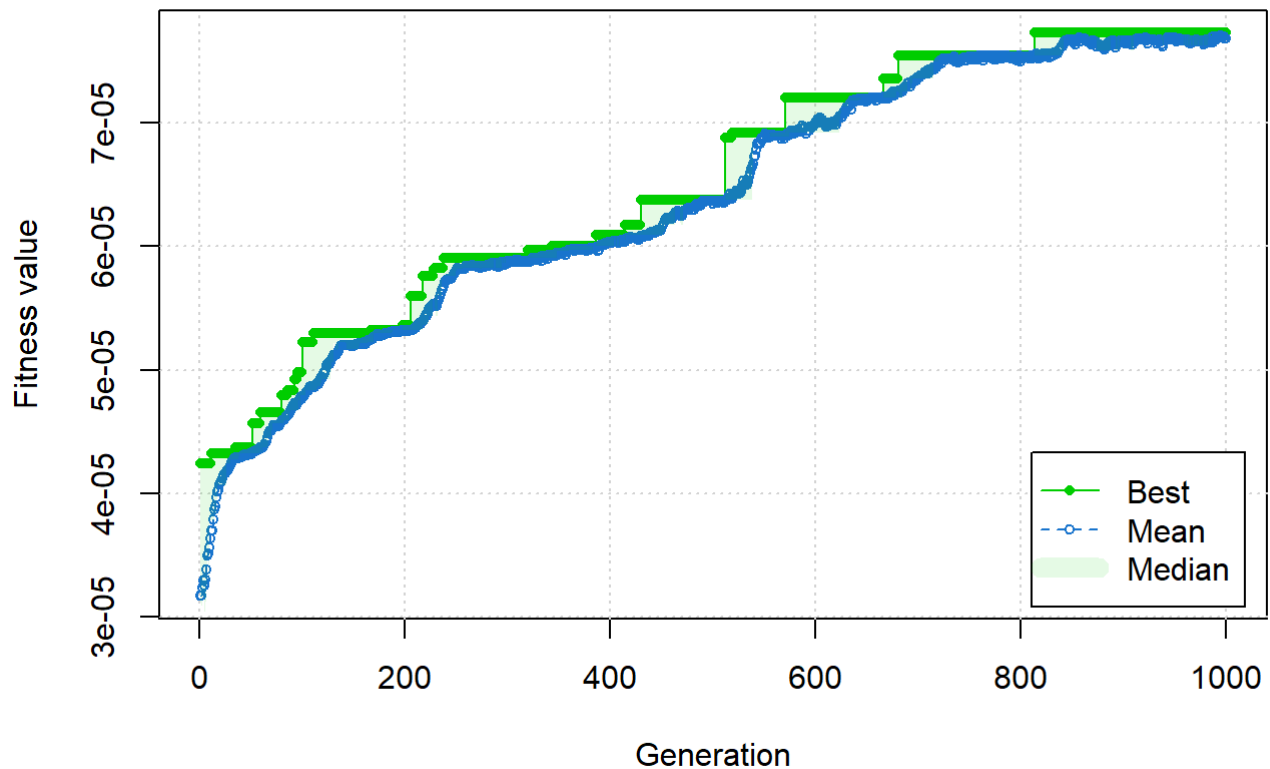
The followed route is similar to previous cases

## Mutation rate modification

### Disminution by one order of magnitude

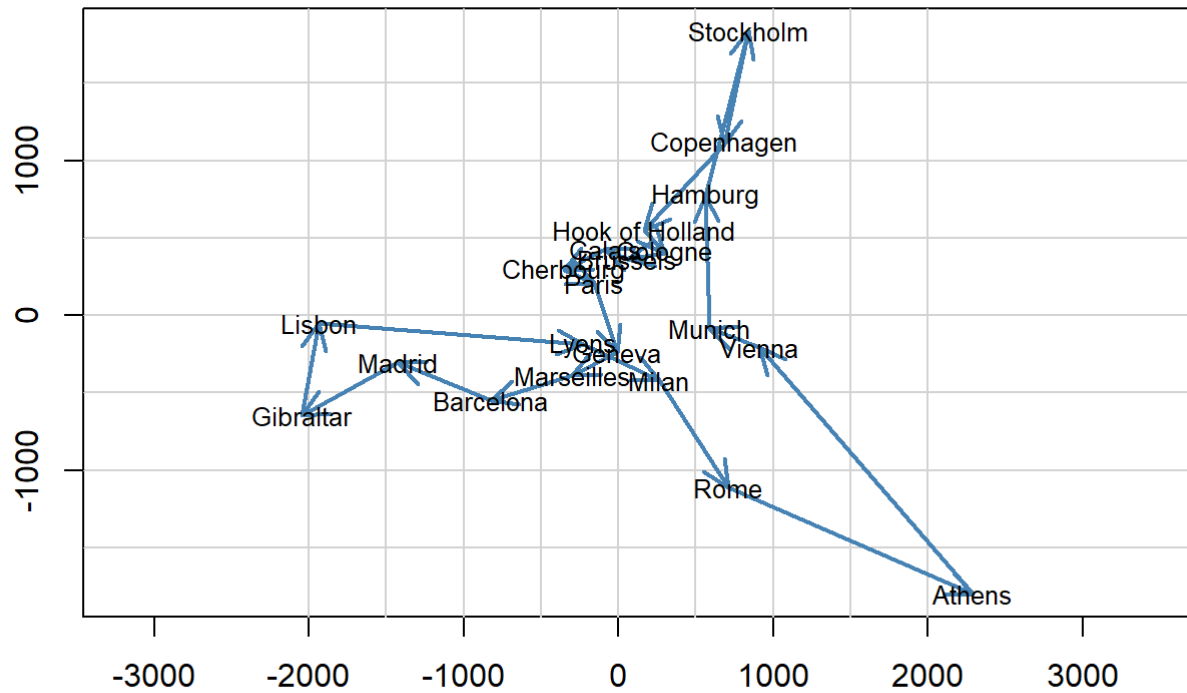
```
GA <- ga(type = "permutation", fitness = tpsFitness, distMatrix = D, lower = 1, upper = attr(eurodist, "Size"), popSize = 100, maxiter = 1000, pmutation = 0.01, pcrossover = 0, monitor = NULL)
```

```
plot(GA)
```



```
mds <- cmdscale(eurodist)
x <- mds[, 1]
y <- -mds[, 2]
plot(x, y, type = "n", asp = 1, xlab = "", ylab = "")
abline(h = pretty(range(x), 10), v = pretty(range(y), 10),
col = "light gray")
tour <- GA@solution[1, ]
tour <- c(tour, tour[1])
n <- length(tour)
arrows(x[tour[-n]], y[tour[-n]], x[tour[-1]], y[tour[-1]],
length = 0.15, angle = 25, col = "steelblue", lwd = 2)
text(x, y, labels(eurodist), cex=0.8)
```





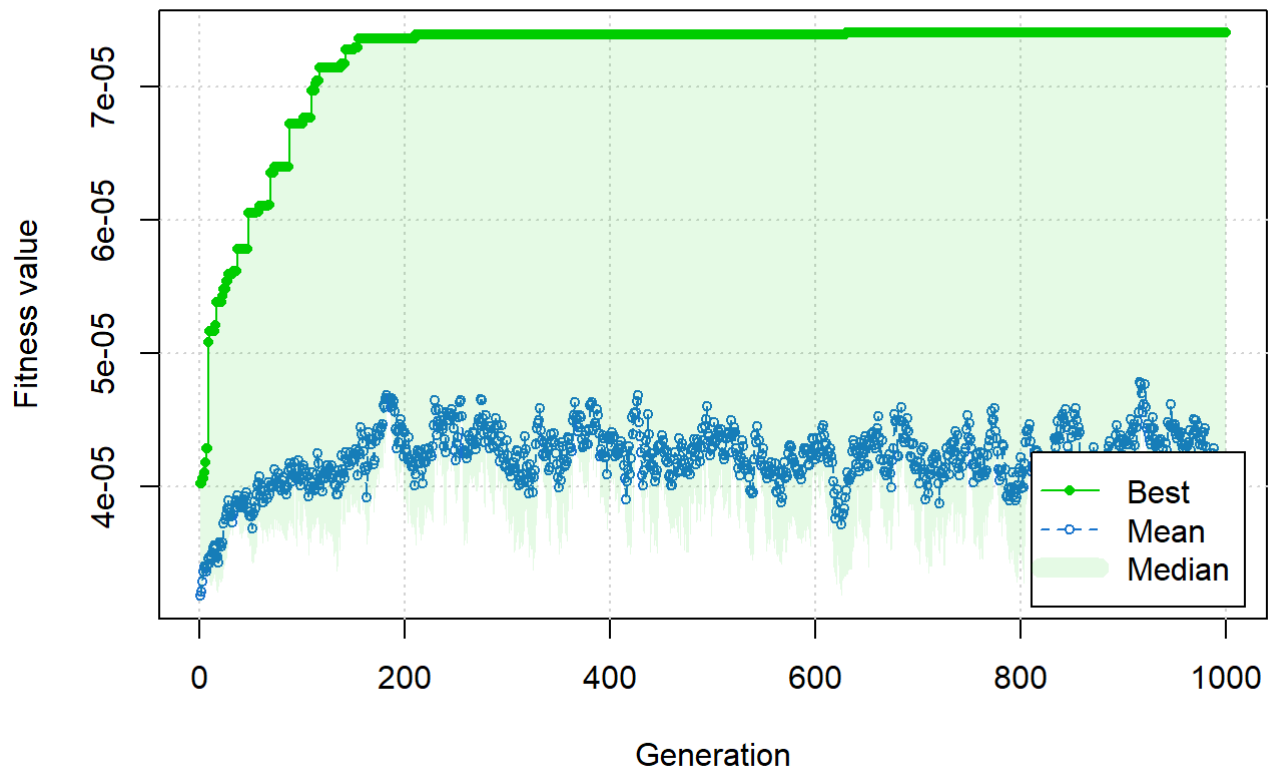
In this case, it is clearly visible that the fitness evolves almost linearly. This is reasonable because by eliminating the mutation rate, we are removing the main source of variation and potential improvement between generations. Nevertheless, it still converges to a similar fitness point as before.

Regarding the route, we can see that in this case, the circular structure is somewhat lost, and there are crossings between different steps of the route.

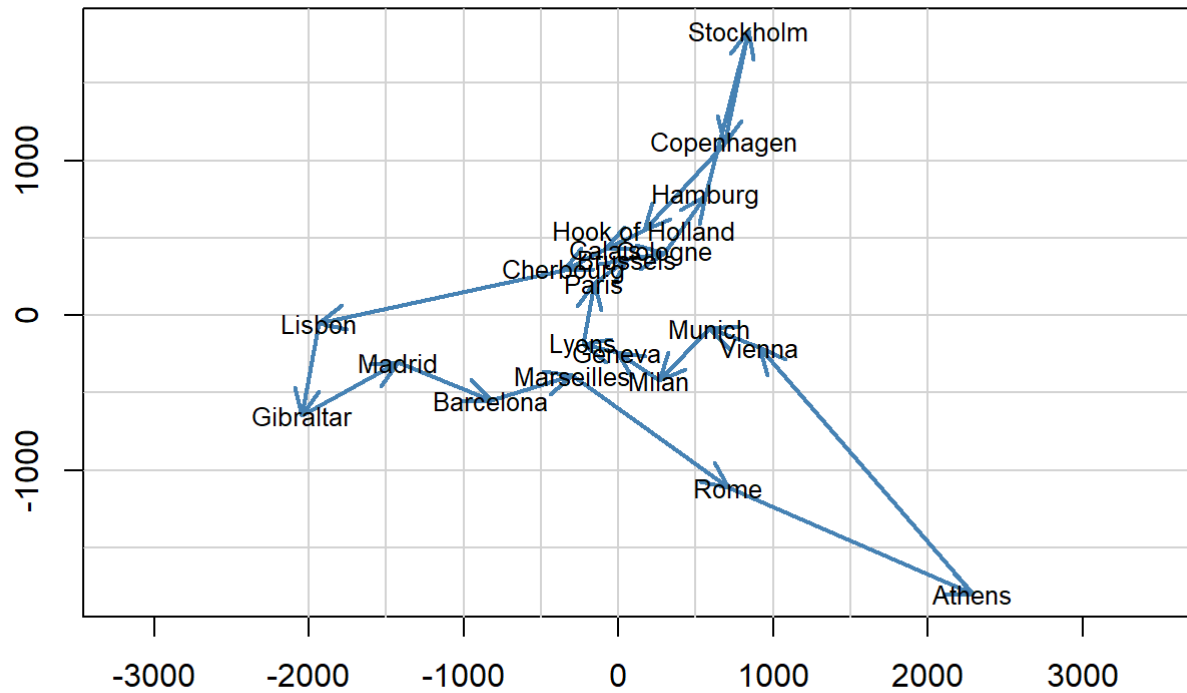
### Increase of one order of magnitude

```
GA <- ga(type = "permutation", fitness = tpsFitness, distMatrix = D, lower = 1, upper = attr(eurodist, "Size"), popSize = 100, maxiter = 1000, pmutation = 1, pcrossover = 0, monitor = NULL)
```

```
plot(GA)
```



```
mds <- cmdscale(eurodist)
x <- mds[, 1]
y <- -mds[, 2]
plot(x, y, type = "n", asp = 1, xlab = "", ylab = "")
abline(h = pretty(range(x), 10), v = pretty(range(y), 10),
col = "light gray")
tour <- GA@solution[1, ]
tour <- c(tour, tour[1])
n <- length(tour)
arrows(x[tour[-n]], y[tour[-n]], x[tour[-1]], y[tour[-1]],
length = 0.15, angle = 25, col = "steelblue", lwd = 2)
text(x, y, labels(eurodist), cex=0.8)
```



In this case, the fitness evolves very differently from other cases, with very low averages and a large difference between the averages and the best fitness value. This is reasonable because by increasing the mutation rate significantly, we are introducing a high level of variation within each generation, leading to more defective variants. Nevertheless, the maximum fitness value still converges to a similar point as before.

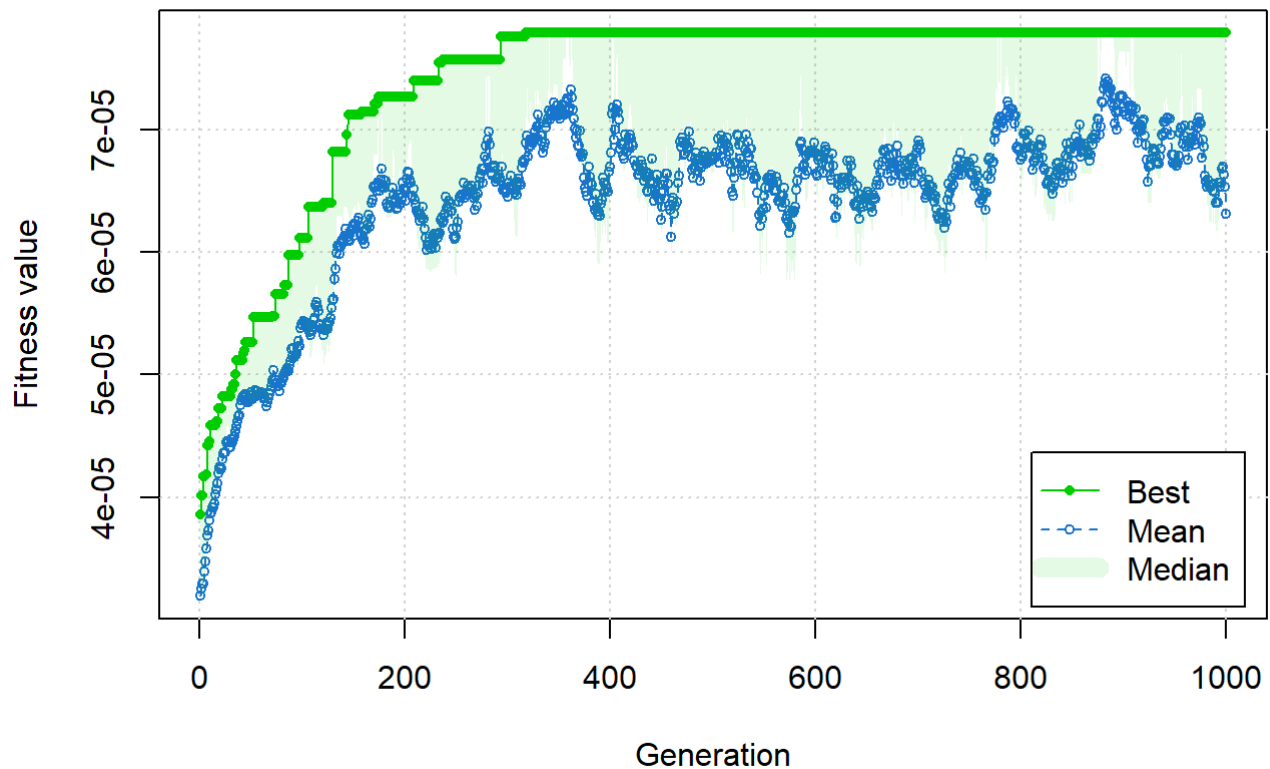
Regarding the route, we can see that in this case, the circular structure with counterclockwise steps in the route is maintained.

## Crossover

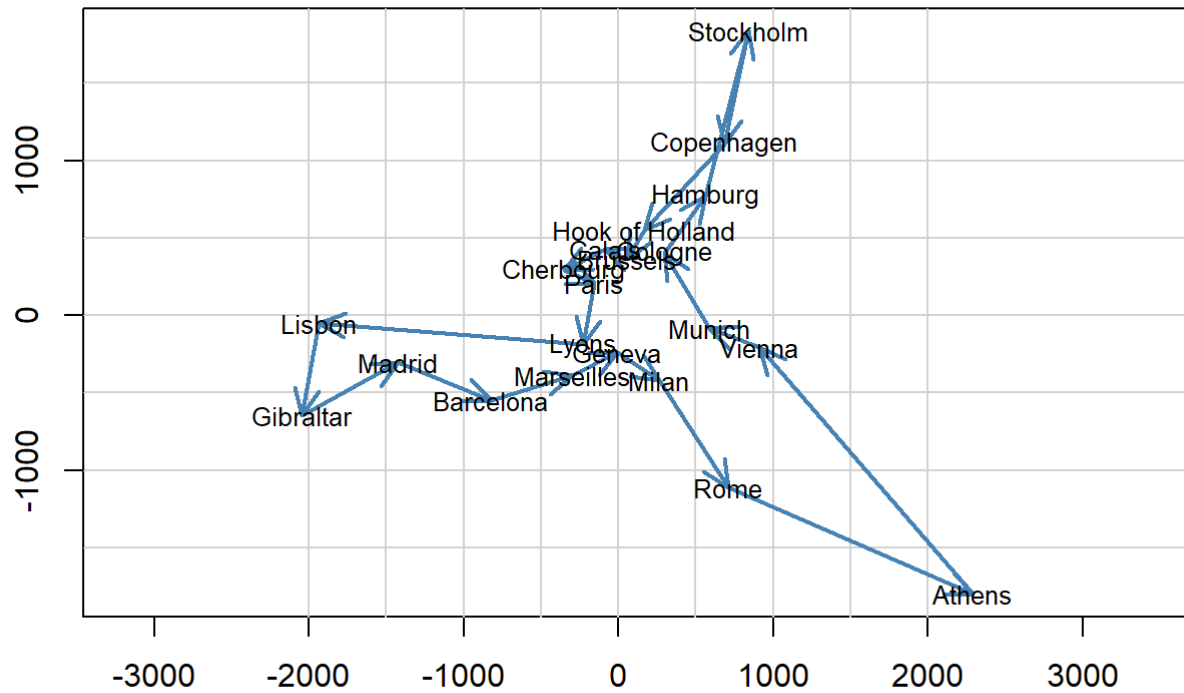
### Low crossover probability

```
GA <- ga(type = "permutation", fitness = tpsFitness, distMatrix = D, lower = 1, upper = attr(eurodist, "Size"), popSize = 100, maxiter = 1000, pmutation = 0.1, pcrossover = 0.1, monitor = NULL)
```

```
plot(GA)
```



```
mds <- cmdscale(eurodist)
x <- mds[, 1]
y <- -mds[, 2]
plot(x, y, type = "n", asp = 1, xlab = "", ylab = "")
abline(h = pretty(range(x), 10), v = pretty(range(y), 10),
col = "light gray")
tour <- GA@solution[1, ]
tour <- c(tour, tour[1])
n <- length(tour)
arrows(x[tour[-n]], y[tour[-n]], x[tour[-1]], y[tour[-1]],
length = 0.15, angle = 25, col = "steelblue", lwd = 2)
text(x, y, labels(eurodist), cex=0.8)
```



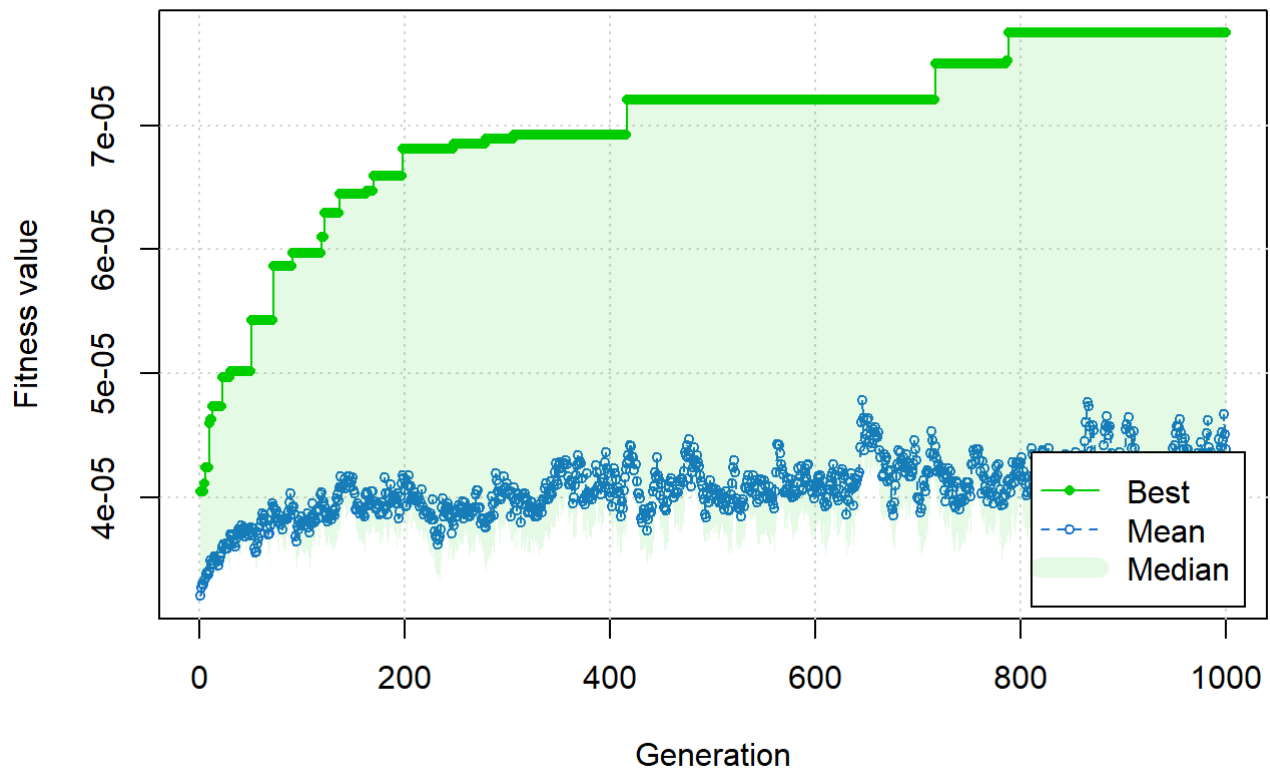
By introducing crossovers, we are incorporating a new source of variability, although it may be mild in this case. Nevertheless, we can see this variability in the differences in means between different generations and the slightly lower average fitness values.

Regarding the route, the circular structure with counterclockwise direction is maintained.

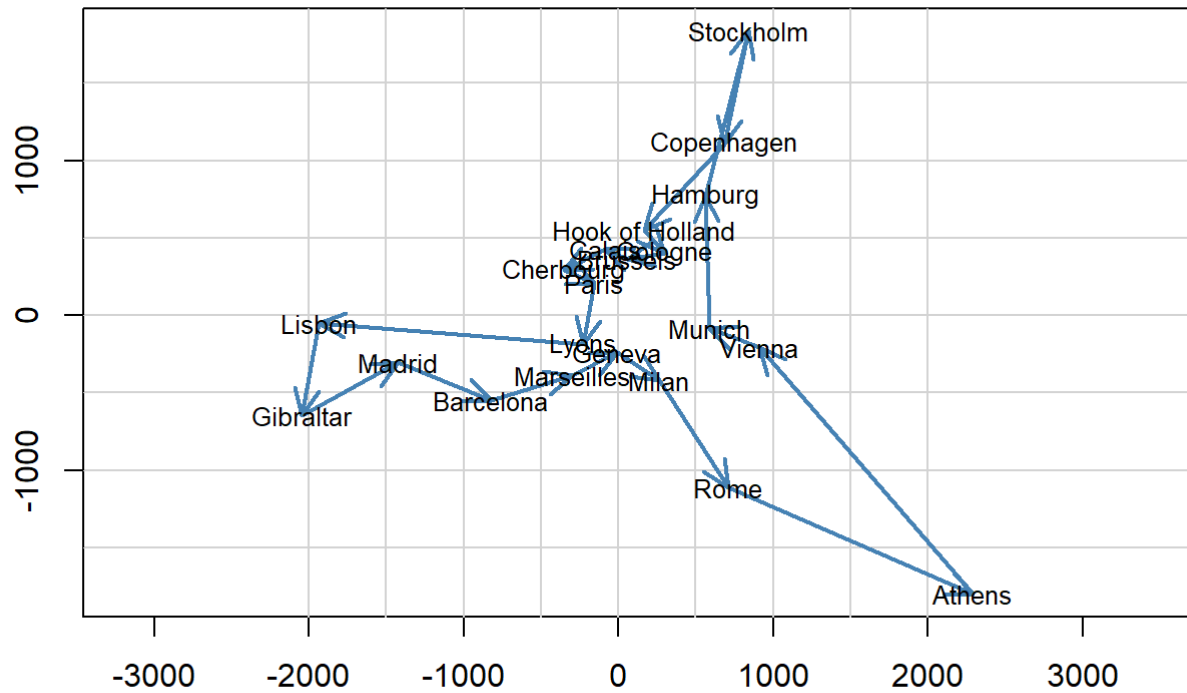
### High crossover probability

```
GA <- ga(type = "permutation", fitness = tpsFitness, distMatrix = D, lower = 1, upper = attr(eurodist, "Size"), popSize = 100, maxiter = 1000, pmutation = 0.1, pcrossover = 0.8, monitor = NULL)
```

```
plot(GA)
```



```
mds <- cmdscale(eurodist)
x <- mds[, 1]
y <- -mds[, 2]
plot(x, y, type = "n", asp = 1, xlab = "", ylab = "")
abline(h = pretty(range(x), 10), v = pretty(range(y), 10),
col = "light gray")
tour <- GA@solution[1, ]
tour <- c(tour, tour[1])
n <- length(tour)
arrows(x[tour[-n]], y[tour[-n]], x[tour[-1]], y[tour[-1]],
length = 0.15, angle = 25, col = "steelblue", lwd = 2)
text(x, y, labels(eurodist), cex=0.8)
```



Similar to the previous case, the crossover probability leads to increased variability. As we can see in this case, if the probability value is very high, the variability within and between generations is also very high. It is worth noting the variability within each generation, as seen in the low levels of the means.

The route with the highest fitness again maintains a circular path, this time in a counterclockwise direction.

## Conclusions

1. The best routes tend to converge regardless of the mean and median values obtained, and regardless of parameters such as mutation or crossover rate. The only thing that tends to vary is the direction of rotation.
2. Mutation and crossover probabilities are the main sources of variability between and within generations.
3. Higher mutation and crossover rates lead to greater variability, reaching a point where this variability greatly reduces the average fitness of the generation.
4. The evolution of different generations follows a linear pattern when we delete the probabilities of mutation and crossover.