Tools for Analytics

Project assignment week 41

In this assignment, we consider a simple **crowd shipping system**. The users of the system (the customers) submits a request at a certain time. The request is for transport from her pickup location to her drop off location.

A number of drivers are available to service the customers. A driver is a private person that needs to perform a trip from his start location to his end location. The driver's departure time is fixed. The driver is willing to take up to three passengers with him on the trip, but the agreement must be made before his departure. In fact, he is willing to perform a detour to do this. For simplicity, we will not limit the allowed length of the detour (even though in practice, one would do that).

In terms of geography, all locations are given as (x , y)-coordinates in a square of size 100 x 100. All distances are Euclidean. So the distance between point 1 (x1,y1) = (6,7) and point 2 (x2, y2) = (10,15) is $\sqrt{(x2-x1)^2+(y2-y1)^2}$ = $\sqrt{(10-6)^2+(15-7)^2} = \sqrt{4^2+8^2} = \sqrt{80} = 8.94$. Feel free to round the distances to two decimals using =round(number;2) if you prefer. We assume that one distance unit corresponds to one minute, so points 1 and 2 are 8.94 minutes apart.

Time is given as minutes past the start of the system, e.g. minutes past 8am.

Data:

Below you see an example of a data file (this particular file is Data9.xlsx).

The number of drivers is given in cell C3. For each of these drivers, the columns A-G contain information about the driver's departure time, start location (x,y)-coordinate, end location (x,y)-coordinate, and the length of the initial trip calculated as the Euclidian distance using the formula above. For example, driver 3 departs from location (90 , 50) at time 180 and drives to location (40 , 10), and this trip takes 64,03 minutes.

Columns J-O provide information regarding the customers. Column K states the time that the customer submits his request. You can assume that they are listed in the order in which they occur. Columns L-M gives the pickup location and columns N-O is the drop off location. For example, customer 1 submits her request to the system at time 10. She needs to go from location (54 , 21) to location (25 , 19).

The Data folder contains 9 data files of varying size.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Number of drivers		3												
2	Drivers:									Customers					
3	ID	Departure time	Start location		End Location		Initial trip duration			ID	Time of request	Pickup location		Drop Location	
4			X	Y	X	Y						X	Y	X	Y
5	1	120	10	10	100	100	127,28			1	10	54	21	25	19
6	2	150	25	90	25	10	80			2	19	44	99	36	75
7	3	180	90	50	40	10	64,03			3	29	86	36	51	50
8										4	55	45	51	74	47
9										5	67	43	76	36	94
10										6	78	59	95	52	80
11										7	98	84	35	88	38
12										8	122	3	63	42	49
13										9	154	77	70	61	55
14										10	192	95	86	75	60
15															

Figure 1 Example of data

System:

When a customer submits a request to the system, the system must determine if the request can be fulfilled (i.e. the customer is accepted) or if the customer has to be rejected. This happens as follows:

> **Rule 1:** The system considers the drivers that have not yet departed and are not yet filled to their capacity (remember, each driver can transport up to 3 customers). If no such driver exists, the customer is rejected. If there are multiple possible drivers, the customer is assigned to the driver whose total detour including service of this customer is smallest.

Assigning a customer to a driver and calculating the detour:

In the following, it is explained how a customer is assigned to a route. Consider driver 2 in the example in Figure 1. He needs to go from (25 , 90) to (25 , 10), a trip that takes 80 minutes.

At time 19, customer 2, who needs to go from (44 , 99) to (36 , 75) is assigned to driver 2. Consequently, driver 2 has the route shown in the left hand side of figure 2. The driver starts in the blue dot (25 , 90), drives along the black line, transports customer 2 along the green line, and plans to go along the dashed black line to his end location (25 , 10). The total length of this route is 21.02 + 25.30 + + 65.92 = 112.24. This results in a total detour of 112.24-80 = 32.24. At this point the solid black and the solid green (transport of customer 2) parts of the route are fixed (cannot be changed), and the dashed black part is expected (completing his tour).

In this illustration, we ignore what happens to customer 3.

At time 55, customer 4 requests transport from (45 , 51) to (74 , 47). Her trip is illustrated by the dashed orange line in the right hand side of figure 2. We need to determine what the total detour will be if customer 4 is assigned to driver 4. The rule is as follows:

> **Rule 2:** New customer services are always assigned to the end of the trip, immediately before the driver goes to his End location.

(from a planning perspective, this is not a good rule, but it is the rule that we will use)

If we apply that rule, assigning customer 4 to driver 2 will result in the route shown to the right in figure 2, first servicing customer 2 (green), then customer 4 (dashed orange), and finally making it to his end location. Consequently, the total length for driver 2 would be 21.02 + 25.30 + 25.63 + 29.27 + 61.4 = 162,62. Hence, if customer 4 is assigned to driver 2, his detour will become 162.62 – 80 = 82.62.

Whether customer 4 is indeed assigned to driver 2 depends on the detours of the other drivers as explained in rule 1.

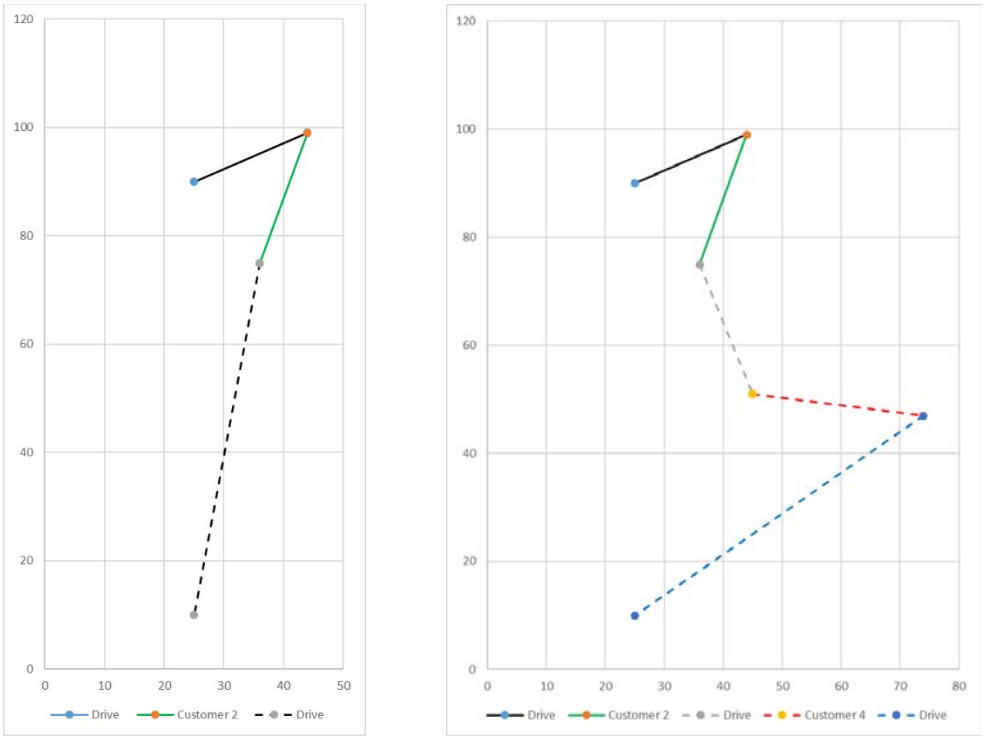

Figure 2 Assignment and calculation of detour

Output:

By applying rules 1 and 2 to each customer in turn, we obtain the plan shown in figure 3.

Cell D1 gives the number of rejected customers, and they are listed in row 2. In the example, only one customer is rejected (number 10).

Each driver has a block of four rows (5-9, 10-14, 15-19, etc). The output data is now illustrated for driver 2:

Cell A10 states the driver ID. Cell B11 is the number of customers transported by the driver and cell B13 is the total length of his route. Each column starting at D, corresponds to one action as follows: Driver start location (25 , 90) (column D), Pickup customer 2 at (44 , 99) (column E), drop off customer 2 at (36 , 75), …, and finally driver end location (25,10) (column K because this driver services 3 customers).

	A	B	C	D	E	F	G	H	I	J	K
1	Number of rejected customers			1							
2	Customers	10									
3											
4	Drivers										
5	1	Num	Who	Driver	3	3	4	4	7	7	Driver
6		3	Action	Start	Pickup	Drop	Pickup	Drop	Pickup	Drop	End
7		Length	X	10	86	51	45	74	84	88	100
8		237,1489	Y	10	36	50	51	47	35	38	100
9											
10	2	Num	Who	Driver	2	2	5	5	6	6	Driver
11		3	Action	Start	Pickup	Drop	Pickup	Drop	Pickup	Drop	End
12		Length	X	25	44	36	43	36	59	52	25
13		187,3076	Y	90	99	75	76	94	95	80	10
14											
15	3	Num	Who	Driver	1	1	8	8	9	9	Driver
16		3	Action	Start	Pickup	Drop	Pickup	Drop	Pickup	Drop	End
17		Length	X	90	54	25	3	42	77	61	40
18		278,334	Y	50	21	19	63	49	70	55	10
19											

Figure 3 Output

Your task:

Make a number of procedures that can be used by the crowd shipping system to assign customers to drivers and determine the rejected customers by applying rules 1 and 2.

Your code should be able to open the data file (one file at a time or all files by looping through them - that is up to you).

Your code should create output with the format shown in figure 3. It must be presented either in a separate sheet of the file with the code (named in a proper manner) or in a new file with proper name.