Question 1

In order to reproduce the following results, we made the assumption that the radius of the earth r is equal to 3958.8 [miles].

We can see on Table 1 the set of optimal cities in order to minimize the price and the distribution of the satisfaction of the demand.

City	Number of clients
Buffalo	19
Houston	15
Minneapolis	9
Oakland	8
Orlando	9
Paradise CDP	20
Winston-Salem	20

Table 1: Optimal cities and how many clients they have

With this set of optimal cities, the (optimal) objective value is $z^* = 4.47007 \times 10^8$ [$].

Question 2

We find the same optimal solution with the second formulation ($z^* = 4.47007 \times 10^8$ [$]) than with the first one. Indeed, feasible solutions of the integer problem regarding constraint (3) ($x_{ij} \leq y_i \ \forall i \in \{1, ..., M\}, j \in \{1, ..., N\}$) will also be feasible solutions regarding constraint (6) ($\sum_{j \in M} x_{ij} \leq |M| y_i \ \forall i \in \{1, ..., N\}$) and reciprocally :

$$\begin{aligned} y_i = 0 \quad &: \quad 0 \leq x_{ij} \leq 0 \iff 0 \leq \sum_{j \in M} x_{ij} \leq 0 \qquad (x_{ij} \geq 0 \text{ by constraint (4)}) \\ y_i = 1 \quad &: \quad 0 \leq x_{ij} \leq 1 \iff 0 \leq \sum_{j \in M} x_{ij} \leq |M| \quad (x_{ij} \leq 1 \text{ by constraint (4)}) \end{aligned} \tag{1}$$

However, we can observe significant differences in the details of the solver for the 2 formulations.

	First formulation (Q1)	Second formulation (Q2)
Relaxation : iterations [#]	189	35
Relaxation : time [s]	0.01	0.00
Relaxation : solution [$]	4.47×10^8	2.34×10^8
Integer : iterations [#]	189	10812
Integer : time [s]	0.08	0.57
Optimal solution [$]	$z^* = 4.47 \times 10^8$	

Table 2: Comparaison between both formulations

The LP relaxation provides a bound on the optimal value of the IP problem and we can see in Table 2 that the optimal value of the relaxation is lower for the second formulation. Differences are due to a better formulation in the first case than in the second one. Indeed, the 2 formulations of the IP problem are equivalent :

$$\begin{aligned} &\min z & &= \quad \max z \\ &\quad x \in P_1 & &\qquad x \in P_2 \\ &\quad y \in \{0,1\}^n & &\qquad y \in \{0,1\}^n \end{aligned} \tag{2}$$

But the feasible set of the LP relaxation of the 2 formulations are not the same and thus, the optimal value can be different :

$$\begin{aligned} &\min z & &\geq \quad \max z \\ &\quad x \in P_1 & &\qquad x \in P_2 \\ &\quad y \in [0,1]^n & &\qquad y \in [0,1]^n \end{aligned} \tag{3}$$

Where P_1 is the feasible set under constraints (2), (3), (4) and P_2 is the feasible set under constraints (2), (4), (6).

This comes from the fact that the first formulation is a better formulation than the second. Indeed we can prove that P_1 is a subset of P_2 (with x the vector containing all variables) :

$$\begin{aligned} x \in P_1 &\iff x_{ij} \leq y_j \quad (\forall i \in \{1, ..., M\}, j \in \{1, ..., N\}) \\ &\implies \sum_{i=1}^{M} x_{ij} \leq \sum_{i=1}^{M} y_j = M y_j \quad (\forall j \in \{1, ..., N\}) \\ &\iff x \in P_2 \end{aligned} \tag{4}$$

However, P_2 is not a subset of P_1. We define x as following : let $y_j = 0.5$ for one $j \in \{1, ..., N\}$, and let $x_{ij} = 1$ for all $i \leq 10$ and $x_{ij} = 0$ for all $i > 10$. Hence, recalling that $M = 100$ (from the statement) :

$$\begin{aligned} x_{ij} > y_j \quad (\forall i \leq 10) &\implies x \notin P_1 \\ 10 = \sum_{i=0}^{M} x_{ij} \leq M y_j = 50 &\implies x \in P_2 \end{aligned} \tag{5}$$

Moreover, in this case, the first formulation is not only a better formulation, it is the *ideal* formulation. In other words, it is the convex hull of the IP feasible set and thus the LP relaxation provides the exact optimal value of the IP problem as we can see in Table 2.