

Directions: This is a closed book/notes test. Complete the following problems by hand. Calculators and computers are not allowed.

1. (12 pts) Consider two data objects with attribute values listed in the table below. Compute the similarity/distance between P and Q using the measures listed in the table below.

P	0	1	0	0	0	1	0	0	0	0
Q	1	0	0	0	0	1	0	0	0	1

measure	value
City Block Distance	3
Euclidean Distance	$\sqrt{3}$
Supremum Distance	1
Cosine Similarity	$\frac{1}{\sqrt{2 \cdot 3}} = \frac{1}{\sqrt{6}}$
Simple Matching Coefficient	$\frac{7}{10} = 0.7$
Jaccard Coefficient	$\frac{1}{4} = 0.25$

2. (12 pts) Compute the indicated statistics (if possible) of each attribute in the table below.

price	gender	local time	age
\$4.00	F	01-01-01 02:00 pm	child
\$2.00	M	01-01-01 04:00 pm	teenager
\$2.00	M	01-01-01 11:00 am	senior
\$3.00	F	01-01-01 01:00 pm	adult
\$1.00	M	01-01-01 01:00 pm	baby

b c t a s

1 2 2 3 4 11 13 13 14 16

statistics	price	gender	local time	age
median	\$2.0	X	01-01-01 01:00pm	teenager
mode	\$2.0	M	01-01-01 01:00pm	baby, child, teenager, adult, senior
mean	\$2.4	X	01-01-01 01:24pm	X

$$\begin{array}{r}
 2.4 \\
 5 \overline{) 12} \\
 \underline{10} \\
 20
 \end{array}
 \quad
 \begin{array}{r}
 26 \\
 + 30 \\
 \hline
 56 \\
 + 11 \\
 \hline
 67
 \end{array}
 \quad
 \begin{array}{r}
 13.4 \\
 5 \overline{) 67} \\
 \underline{5} \\
 17 \\
 \underline{15} \\
 20
 \end{array}
 \quad
 0.4 \cdot 60 = 24$$

3. (10 pts) The heights (in inches) of ten people are list below. Construct a box plot for this data set.

40, 55, 56, 58, 60, 60, 62, 63, 65, 80

