## 第二讲 （属性、统计度量、可视化、相似性）

1、离散属性总是具有有限个值。（）

A．对

B．错

2、“数值之间具有有意义的秩”的属性是什么类型的属性（）

A.标称属性

B.二元属性

C.序数属性

D.数值属性

3、属性“成绩”的取值有A+ A A- B+ B B- 等，则“成绩”这个属性属于哪一种属性（）

A.标称属性

B.二元属性

C.序数属性

D.数值属性

4、以下属性,属于离散属性有（）.

A.标称属性

B.二元属性

C.序数属性

D.数值属性

5、假设数据是倾斜的（非对称），在度量数据的中心趋势时，应该选择哪种度量方式？（）

A.均值

B.中位数

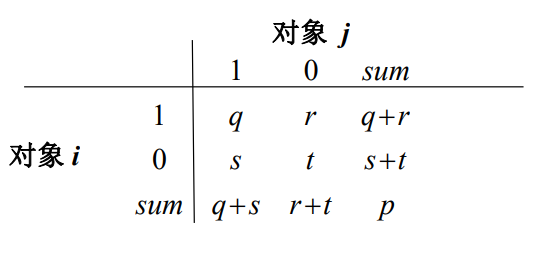
C.四分位数

D.众数

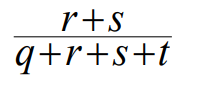
6、为了避免因度量单位的不同造成的影响，数据应该标准化()。

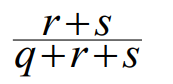
A．对

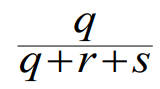
B．错

7、

如图给定两个二元数据，则Jaccard系数的公式为（）

A. 

B. 

C. 

8、求向量类型的数据之间的相似度时，应该选择余弦度量。（）

A．对

B．错

9、下面哪个不属于数据的属性类型：()

A.标称

B.序数

C.区间

D.相异

10、考虑值集{12,24,33,2,4,55,68,26}，其四分位数极差是：()

A.31

B.24

C.55

D.3

11、下列哪个不是专门用于可视化时间空间数据的技术： ()

A.等高线图

B.饼图

C.曲面图

D.矢量场图

12、离群点可以是合法的数据对象或者值。（）

A．对

B．错

13、五数概括指的是平均数、四分位数Q1和Q3，最小观测值和最大观测值（）

A．对

B．错

14、假设所分析的数据包括属性age，它在数据元组中的值（以递增序）为13，15，16，16，19，20，20，21，22，22，25，25，25，25，30，33，33，35，35，35，35，36，40，45，46，52，70。该数据的均值是\_\_\_，中位数是\_\_\_（）。

A.30 25

B.33 20

C.25 30

D.30 22

15、假设所分析的数据包括属性age，它在数据元组中的值（以递增序）为13，15，16，16，19，20，20，21，22，22，25，25，25，25，30，33，33，35，35，35，35，36，40，45，46，52，70。该数据的众数是（）。

A.25

B.35

C.25和35

D.以上选项都不是

16、假设所分析的数据包括属性age，它在数据元组中的值（以递增序）为13，15，16，16，19，20，20，21，22，22，25，25，25，25，30，33，33，35，35，35，35，36，40，45，46，52，70。该数据的中列数是（）。

A.35

B.40

C.41.5

D.45

17、假设所分析的数据包括属性age，它在数据元组中的值（以递增序）为13，15，16，16，19，20，20，21，22，22，25，25，25，25，30，33，33，35，35，35，35，36，40，45，46，52，70。（粗略地）找出该数据的第一个四分位数Q1是\_\_\_，第三个四分位数Q3是\_\_\_()。

A.15 52

B.16 35

C.20 52

D.20 35