

نام و نام خانوادگي: وزارت علوم و تحقیقات و فناوری عنوان امتحان: داده کاوی شماره دانشجويي: تاريخ امتحان: 29/1/97

نام استاد: ملكوتي مدت امتحان: 90 دقیقه

1. چهار نوع اصلی صفات اشیاء داده ای در داده کاوی و زیر مجموعه های آن ها را با ذکر مثال در مورد هر یک بیان کنید.
2. مجموعه داده زیر مفروض است:

5،2،10،6،2،8،8،1،2،5،7،3،3،4

* میانگین، میانه و مد داده ها را بدست آورید.
* نمودار جعبه ای(Box Plot) داده ها را رسم کنید.
* آیا داده ها حاوی داده پرت هستند؟ چرا؟

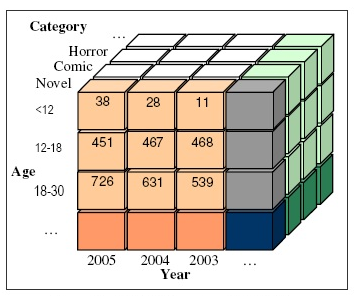
1. فرض کنید سه مشتری هر کدام یک رکورد خرید به شکل زیر در پایگاه داده ما دارند. با فرض اینکه صفات نوع جنس و رنگ، صفات اسمی و صفت سایز صفت ترتیبی حاوی مقادیر کوچک، متوسط و بزرگ باشد، ماتریس عدم تشابه بین آن ها را محاسبه کنید:

|  |  |  |  |
| --- | --- | --- | --- |
| سایز | رنگ | نوع جنس | شماره مشتری |
| متوسط | سفید | پیراهن | 1 |
| متوسط | سفید | شلوار | 2 |
| بزرگ | سبز | پیراهن | 3 |

1. چهار وظیفه اصلی مرحله پیش پردازش داده ها را با توضیح مختصر بیان کنید.
2. داده های زیر را با روش دلخواه و با ذکر نام روش نرمالسازی کنید:

200،300،400،600،1000

1. پایایی در انبار داده به چه معناست؟
2. سه مدل انبار داده سازمان، بازار داده و انبار داده مجازی را با ذکر تفاوت بین آنها توضیح دهید.
3. مکعب داده زیر را که برای نمایش اطلاعات مربوط به تعداد کتاب های به امانت برده شده در یک کتابخانه بر حسب نوع کتاب، سن امانت گیرنده و زمان می باشد را در نظر بگیرید:



* ویژگی های بعد (dimention) و معیار (measure) در این مکعب کدامند؟
* مدل مفهومی برای این انبار داده را با استفاده از شمای ستاره ای ترسیم کنید.
* یک نمونه عمل drill down در این مکعب مثال بزنید.
* یک نمونه عمل dice در مکعب مثال بزنید.

موفق باشید.