پروژه درس مبانی داده کاوی

فاز اول

علی جعفری (9622762431)

سوال اول:

با توجه به کد و مشاهده دیتاست دیوار، جدول خواسته شده به صورت زیر پر شده است (معنی \* این است که بازه مقادیر ستون موردنظر، به علت طولانی بودن، در پایین جدول آورده شده است):

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ردیف | نام ویژگی | نوع | بازه مقادیر | Min | Max | Mean | Mode | Median | مقادیر پرت |
| 1 | archive\_by\_user | Symmetric Binary | False/True | - | - | - | True | - |  |
| 2 | brand | Nominal | \* | - | - | - | پراید صندوق دار | - |  |
| 3 | cat1 | Nominal | \* | - | - | - | for-the-home | - |  |
| 4 | cat2 | Nominal | \* | - | - | - | furniture-and-home-decore | - |  |
| 5 | cat3 | Nominal | \* | - | - | - | light | - |  |
| 6 | city | Nominal | \* | - | - | - | Tehran | - |  |
| 7 | created\_at | Nominal | - | - | - | - | Saturday 12PM | - |  |
| 8 | desc | Nominal | - | - | - | - | کاملا سالم | - |  |
| 9 | id | Nominal | \* | - | - | - | - | - |  |
| 10 | image\_count | Numeric | \* | 0 | 16 | 1.78 | 0, 1 | 2 |  |
| 11 | mileage | Numeric | \* | 0 | 500,000 | 100862.29 | 200000 | 75,000 |  |
| 12 | platform | Symmetric Binary | mobile/ web | - | - | - | mobile | - |  |
| 13 | price | Numeric | \* | “\*” | 400,000,000 | 3,398,461.52 | “\*” | 150,000 |  |
| 14 | title | Nominal | \* | - | - | - | بوفه | - |  |
| 15 | type | Nominal | \* | - | - | - | women | - |  |
| 16 | year | Numeric | \* | 1366 | 1395 | 1387.64 | 1393 | 1389 |  |