In [1]:

```
import pandas as pd
import numpy as np
import warnings
warnings.filterwarnings("ignore")
```

In [2]:

```
data = pd.read_csv("BX-Book-Ratings.csv", encoding= 'unicode_escape',sep=';')
```

In [3]:

```
data.head()
```

Out[3]:

	User-ID	ISBN	Book-Rating
0	276725	034545104X	0
1	276726	0155061224	5
2	276727	0446520802	0
3	276729	052165615X	3
4	276729	0521795028	6

In [4]:

```
data.shape
```

Out[4]:

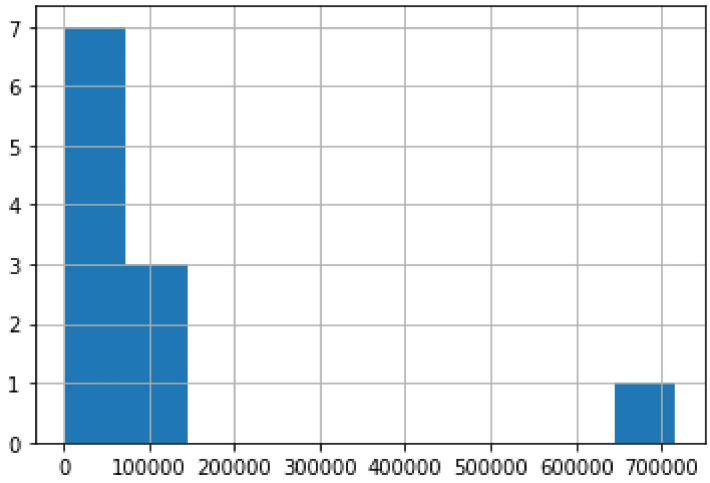
```
(1149780, 3)
```

In [5]:

```
data["Book-Rating"].value_counts().hist()
```

Out[5]:

```
<AxesSubplot:>
```



In [6]:

```
data["Book-Rating"].unique()
```

Out[6]:

```
array([ 0,  5,  3,  6,  8,  7, 10,  9,  4,  1,  2], dtype=int64)
```

In [9]:

```
#!pip3 install surprise
#!pip3 install scikit-surprise
```

In [10]:

```
from surprise import Dataset, SVD, accuracy, Reader
from surprise.model_selection import train_test_split
```

In [11]:

```
# mention the range of rating
reader = Reader(rating_scale=(0,7))
```

In [12]:

```
data = Dataset.load_from_df(data,reader)
```

In [13]:

```
trainset,testset = train_test_split(data,test_size=0.3,random_state=1)
```

In [14]:

```
svd = SVD(n_factors=200)
```

In [15]:

```
svd.fit(trainset)
```

Out[15]:

```
<surprise.prediction_algorithms.matrix_factorization.SVD at 0x1fda23d19d0>
```

In [16]:

```
predictions = svd.test(testset)
accuracy.rmse(predictions)
```

```
RMSE: 3.4648
```

Out[16]:

```
3.464783011120837
```

In [17]:

```
svd.predict(276725,'0155061224')
```

Out[17]:

```
Prediction(uid=276725, iid='0155061224', r_ui=None, est=2.864944992638342, d
etails={'was_impossible': False})
```

In [ ]:

```

```