In [8]:
```
from completejourney_py import get_data

cj_data = get_data()

import pandas as pd

transactions = cj_data['transactions']
transactions.head()
```

Out[8]:

	household_id	store_id	basket_id	product_id	quantity	sales_value	retail_disc	coupon_disc	coupon_match_disc	week	transaction_timestamp
0	900	330	31198570044	1095275	1	0.50	0.00	0.0	0.0	1	2017-01-01 11:53:26
1	900	330	31198570047	9878513	1	0.99	0.10	0.0	0.0	1	2017-01-01 12:10:28
2	1228	406	31198655051	1041453	1	1.43	0.15	0.0	0.0	1	2017-01-01 12:26:30
3	906	319	31198705046	1020156	1	1.50	0.29	0.0	0.0	1	2017-01-01 12:30:27
4	906	319	31198705046	1053875	2	2.78	0.80	0.0	0.0	1	2017-01-01 12:30:27

In [10]:
```
transactions['total_disc'] = transactions['retail_disc'] + transactions['coupon_disc'] + transactions['coupon_match_disc']
```

In [12]:
```
transactions.head()
```

Out[12]:

	household_id	store_id	basket_id	product_id	quantity	sales_value	retail_disc	coupon_disc	coupon_match_disc	week	transaction_timestamp	total_disc
0	900	330	31198570044	1095275	1	0.50	0.00	0.0	0.0	1	2017-01-01 11:53:26	0.00
1	900	330	31198570047	9878513	1	0.99	0.10	0.0	0.0	1	2017-01-01 12:10:28	0.10
2	1228	406	31198655051	1041453	1	1.43	0.15	0.0	0.0	1	2017-01-01 12:26:30	0.15
3	906	319	31198705046	1020156	1	1.50	0.29	0.0	0.0	1	2017-01-01 12:30:27	0.29
4	906	319	31198705046	1053875	2	2.78	0.80	0.0	0.0	1	2017-01-01 12:30:27	0.80

In [39]:
```
def DiscRating(transactions):
    if (transactions['sales_value'] == 0):
        return 'none'
    else:
        total_disc = transactions['total_disc']/transactions['sales_value']*100
        if total_disc == 0:
            return 'none'
        elif total_disc > 0 and total_disc < 25:
            return 'low'
        elif total_disc >= 25 and total_disc <75:
            return 'medium'
        elif total_disc >= 75:
            return 'high'
        else:
            return 'other'

transactions['disc_rating'] = transactions.apply(DiscRating, axis=1)
```

In [41]:
```
transactions.head()
```

Out[41]:

	household_id	store_id	basket_id	product_id	quantity	sales_value	retail_disc	coupon_disc	coupon_match_disc	week	transaction_timestamp	total_disc	disc_ratii
0	900	330	31198570044	1095275	1	0.50	0.00	0.0	0.0	1	2017-01-01 11:53:26	0.00	no
1	900	330	31198570047	9878513	1	0.99	0.10	0.0	0.0	1	2017-01-01 12:10:28	0.10	lo
2	1228	406	31198655051	1041453	1	1.43	0.15	0.0	0.0	1	2017-01-01 12:26:30	0.15	lo
3	906	319	31198705046	1020156	1	1.50	0.29	0.0	0.0	1	2017-01-01 12:30:27	0.29	lo
4	906	319	31198705046	1053875	2	2.78	0.80	0.0	0.0	1	2017-01-01 12:30:27	0.80	mediu

In [43]:
```
transactions.tail()
```

Out[43]:

	household_id	store_id	basket_id	product_id	quantity	sales_value	retail_disc	coupon_disc	coupon_match_disc	week	transaction_timestamp	total_disc	d
1469302	679	447	41453103606	14025548	1	0.79	0.20	0.0	0.0	53	2018-01-01 03:50:03	0.20	
1469303	2070	311	41453083334	909894	1	1.73	0.17	0.0	0.0	53	2018-01-01 04:01:20	0.17	
1469304	2070	311	41453083334	933067	2	5.00	2.98	0.0	0.0	53	2018-01-01 04:01:20	2.98	
1469305	2070	311	41453083334	1029743	1	2.60	0.29	0.0	0.0	53	2018-01-01 04:01:20	0.29	
1469306	2070	311	41453083334	1061220	1	1.19	0.13	0.0	0.0	53	2018-01-01 04:01:20	0.13	

In [49]:
```
transactions['disc_rating'].value_counts()
```

Out[49]:
```
disc_rating
none      728244
medium    333714
low       280599
high      126750
Name: count, dtype: int64
```

In [ ]:
```

```