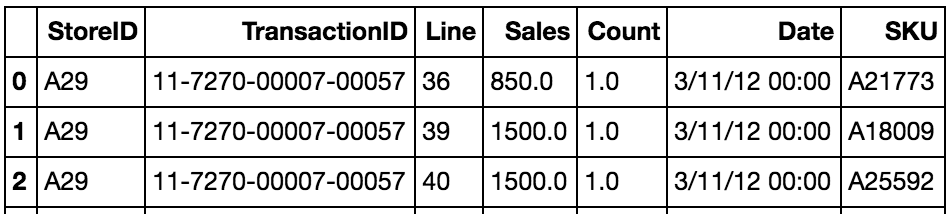
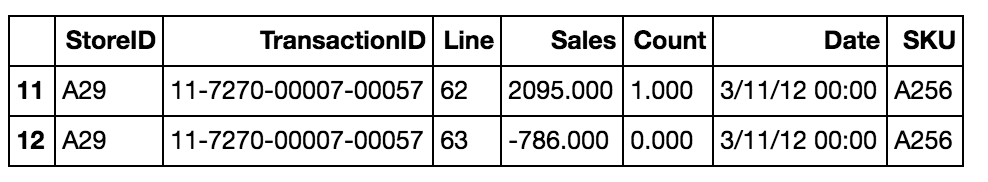
**All Scenarios**

**There are total 1M rows and 7 features in the dataset**

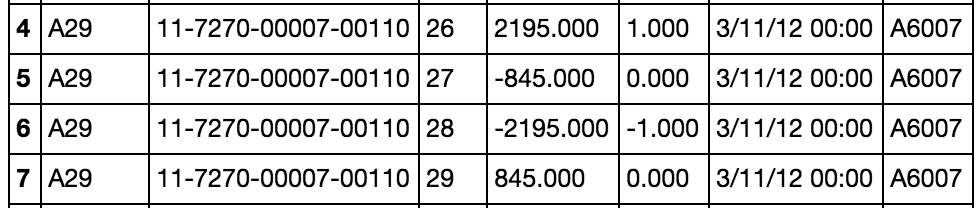
****

**scenario 1**: when count=0, sales is negative, then 'it is a discount'

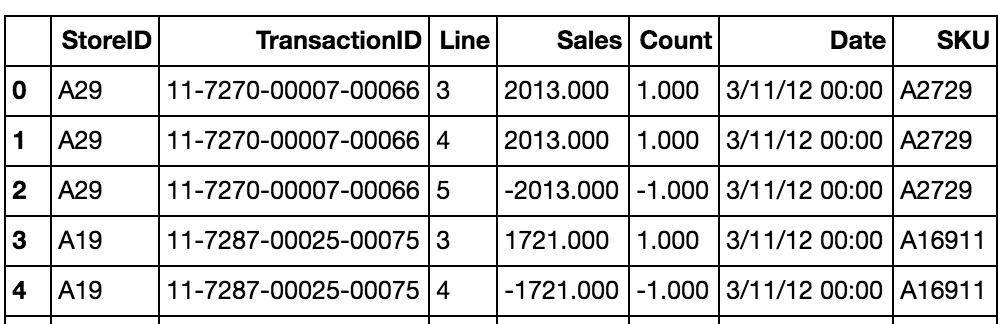


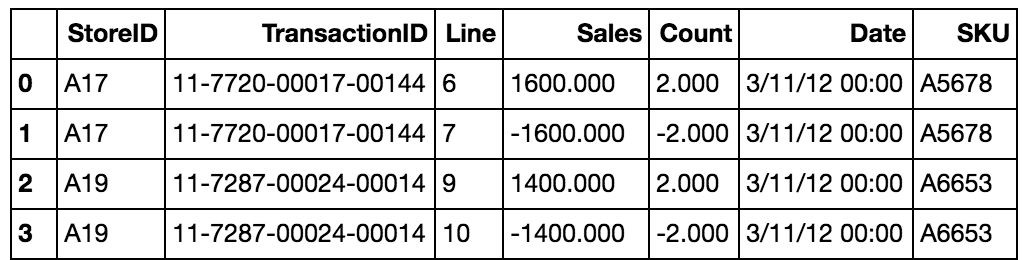
**scenario 2**: when count=0, sales is positive, then 'it is a return'

The positive sales amount is basically the credit back to the store for the discount extended to the customer earlier.



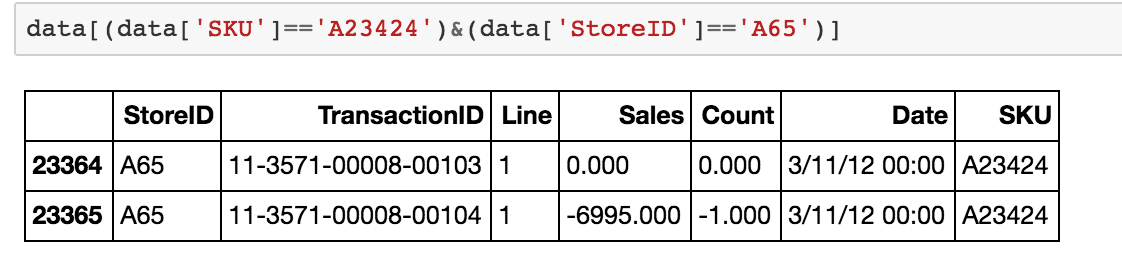
**scenario 3**: when count=-1 or -2, basically a negative number, then 'it could be a promotion strategy like buy 1 get 1 free, or a return', in this analysis, I treated it as a return.

****

****

**scenario 4:** no **corresponding** positive sales when count<0 for SKU in a specfic store, I found count is negative meaing is a return, but no corresponding sales associated with it. **Meaning that it returned sth bought earlier from other stores or that is not on the record.** Or when item is returned, I can’t find a prior mathcing sales record for it.

**Number of SKUs that exhibit such behavior across all the stores, there might be duplicate SKUs in this number**: 1020

****

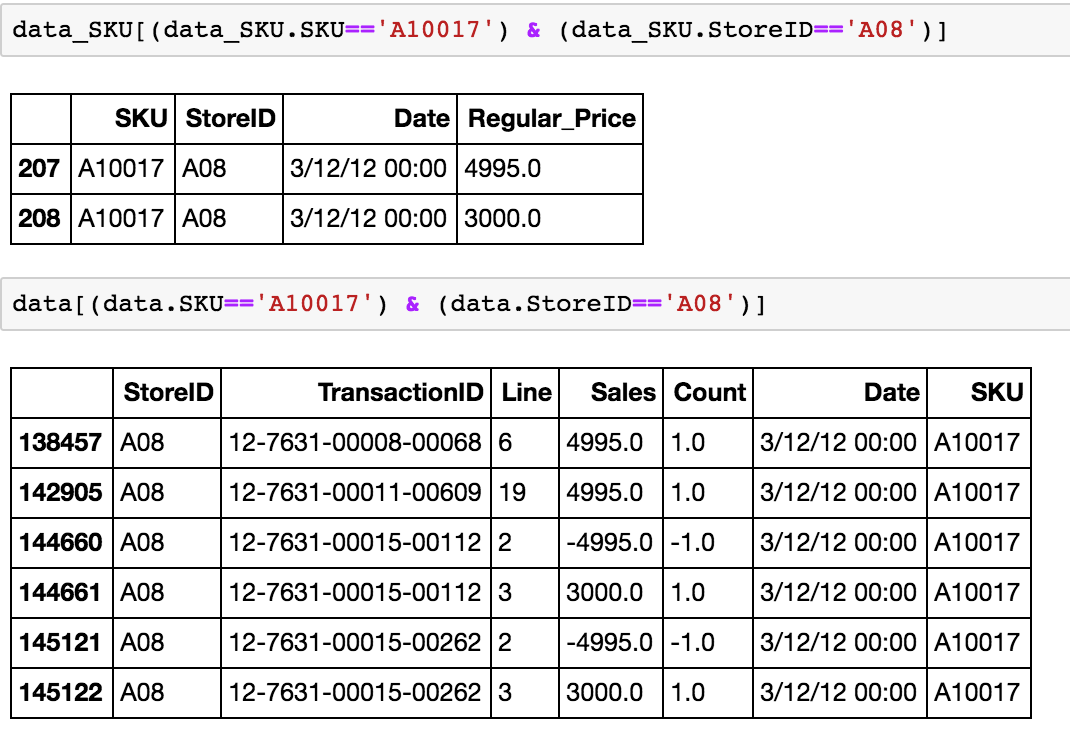
**Next scenarios will be based on SKU level per store, per day**

**(Total count of row numbers of SKU on the per day, per store level that exhibit such behaviors: 200K)**

**scenario 5:** 2 distinct regular price

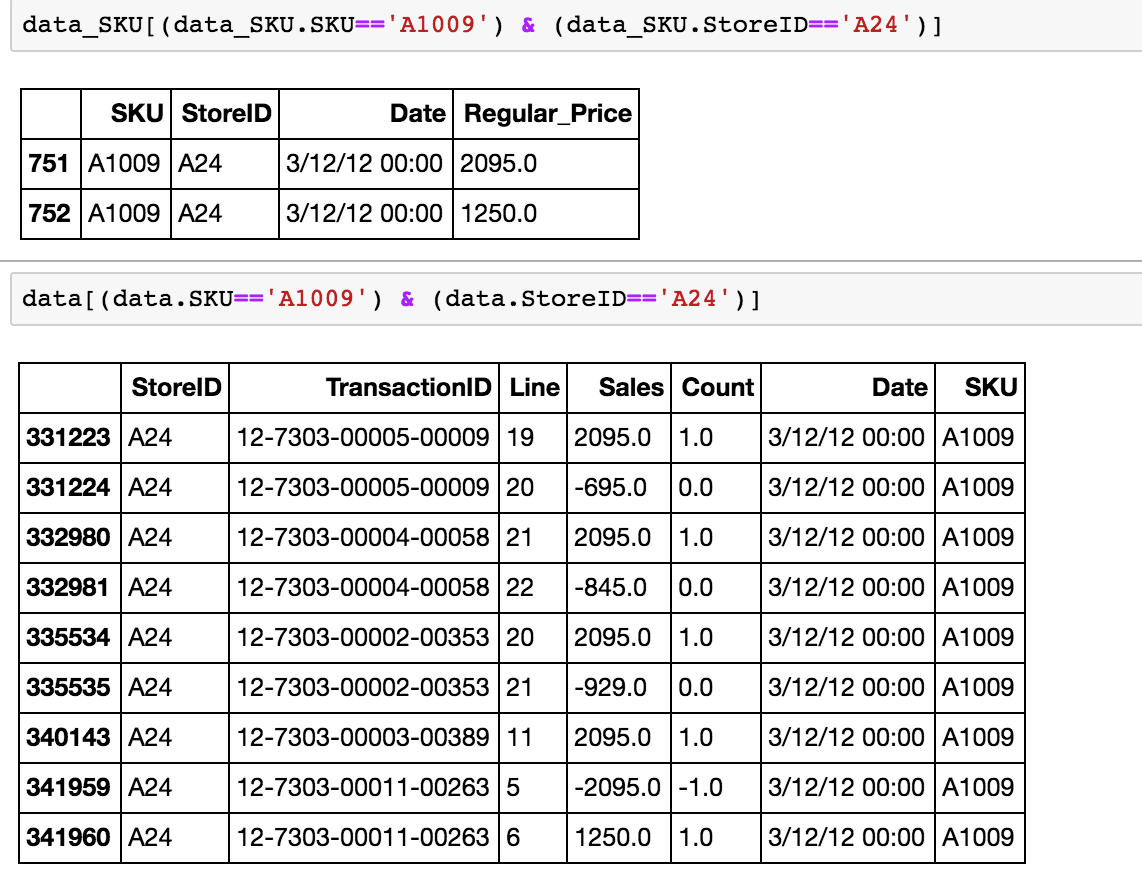
Number of cases on SKU level per store, per day: 1270

a. Promotion can start anytime

****

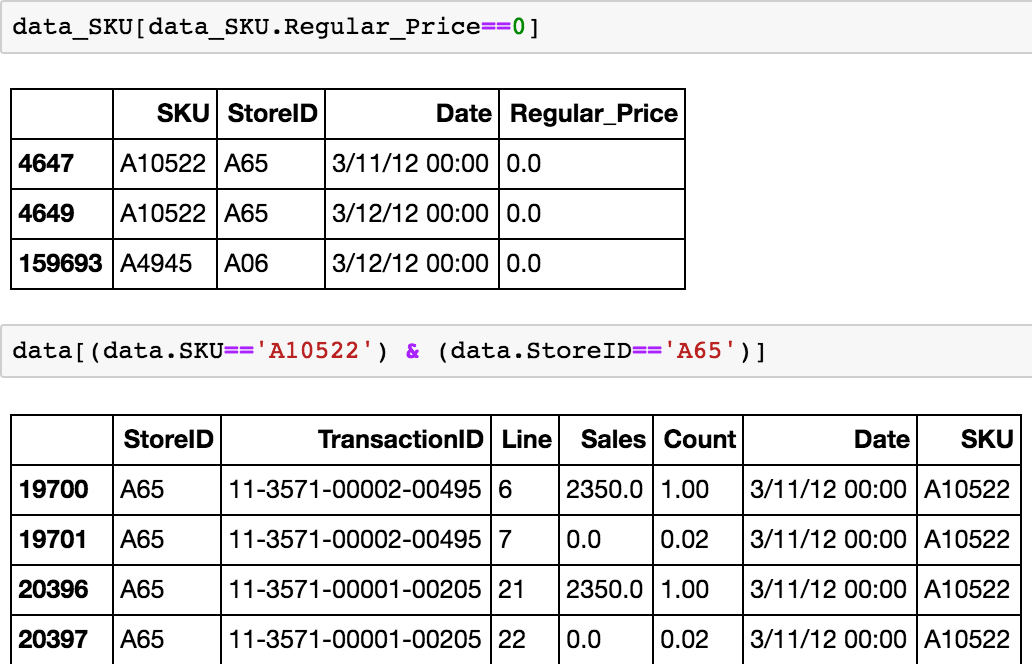
**Explanation:** a store's promotion can start anytime during a day. It could happen that the transaction captures sale of an item before and after a promotion starts.

b. Different discount value and regular price equals to the price after discount

****

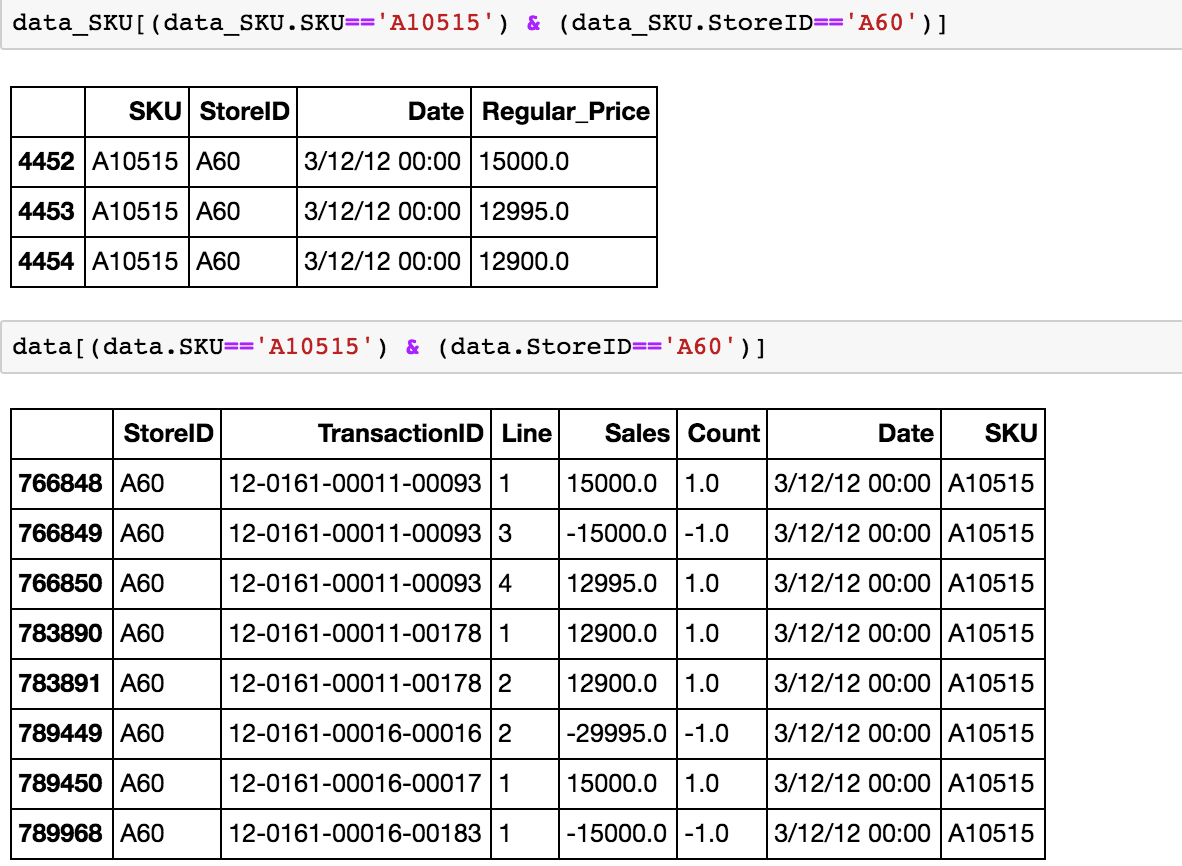
c. Sales=0 and Count is a small number

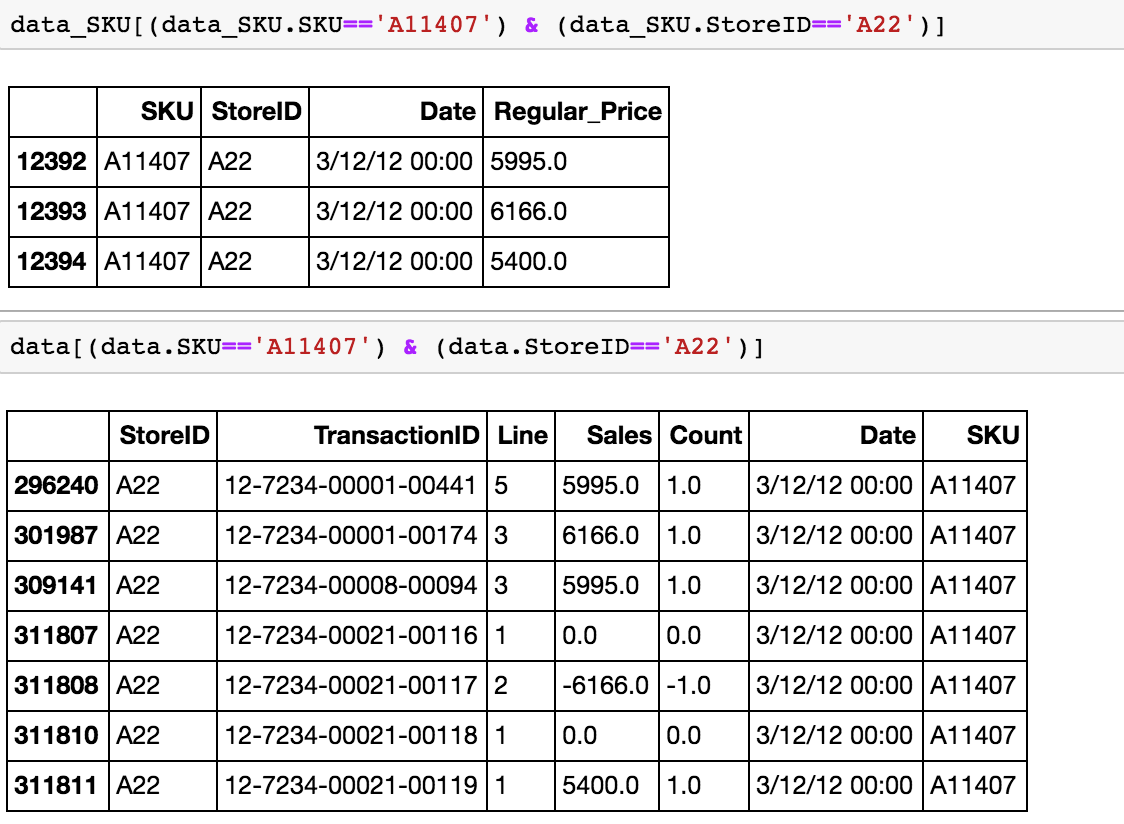
Number of cases on SKU level per store, per day: 3

****

**scenario 6:** 3 distinct regular price on SKU level per store, per day

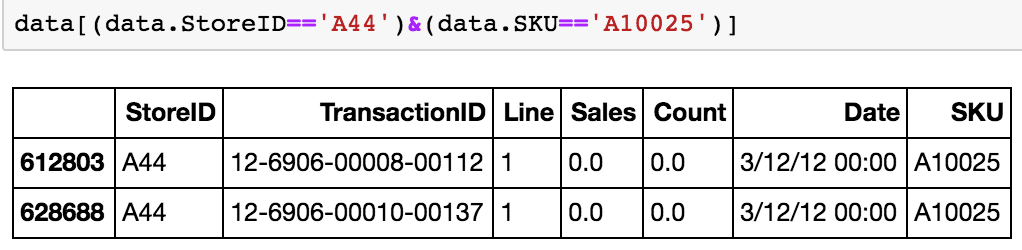
Number of cases on SKU level per store, per day: 30



****

**scenario 7:** There are only obersvations when Sales=0 & Count=0 on SKU level per store, per day

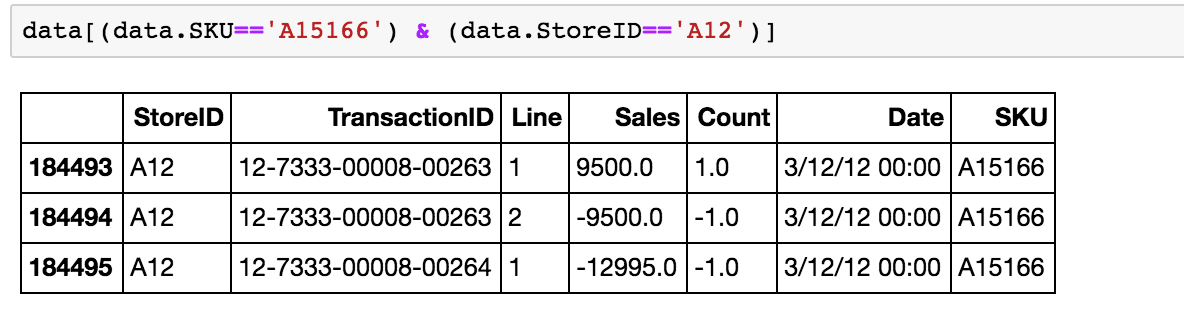
Number of cases on SKU level per store, per day: 477

****

**scenario 8:** Unit Price after sales is bigger than regular price on SKU level per store, per day

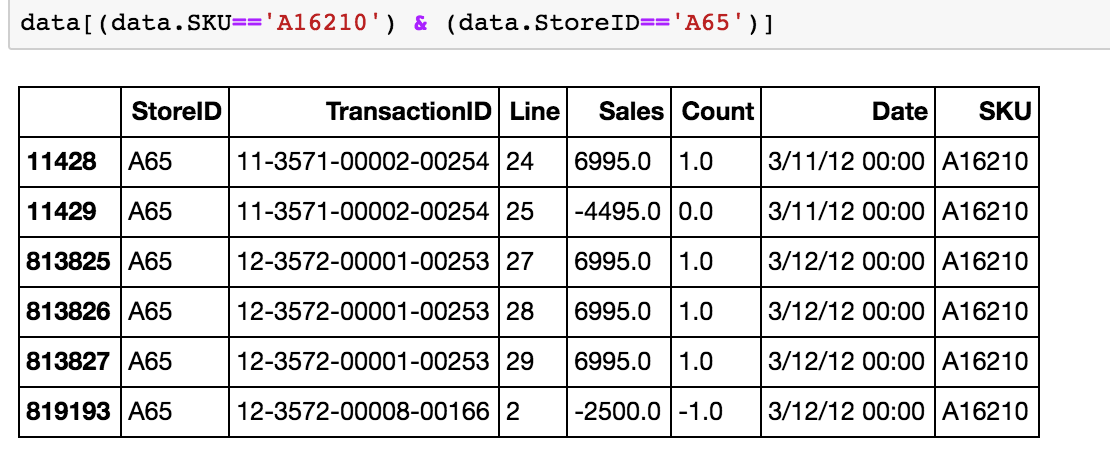
Number of cases on SKU level per store, per day: 22

a.

****

**Explanation:** It only have returned history, but no purchased history. It could be the reason that the item was bought outside timeperiod of dataset, so it was not recorded in this dataset.

b.

****

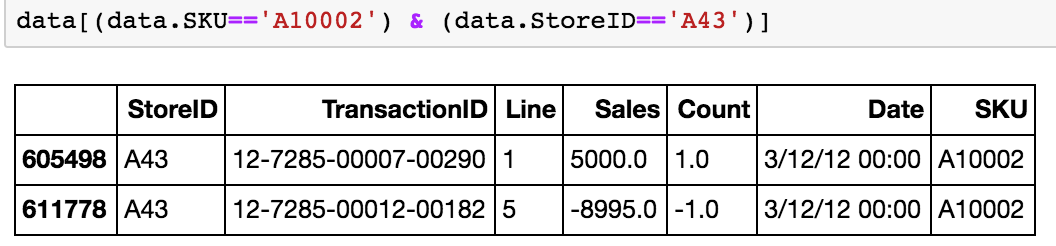
**Explanation:** Returned price is lower than the regular price. (Maybe date is wrong?)

**scenario 9:** Discount rate is missing value

Number of cases on SKU level per store, per day: 1041

a. Total Sales<0, Total Count=0

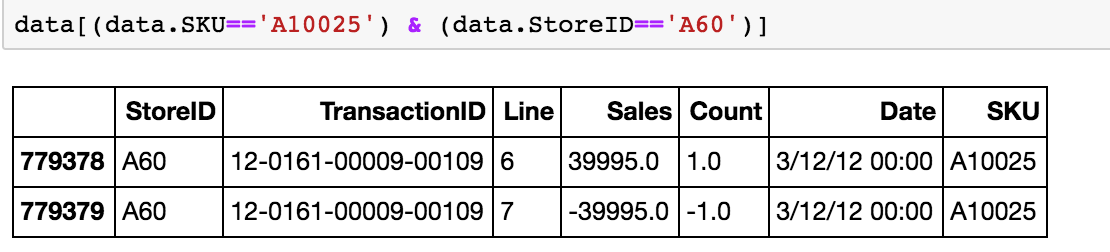
Number of cases on SKU level per store, per day: 51



**Explanation:** Returned price is higher than the regular price and total count equalts to 0.

b. Total Sales=0, Total Count=0

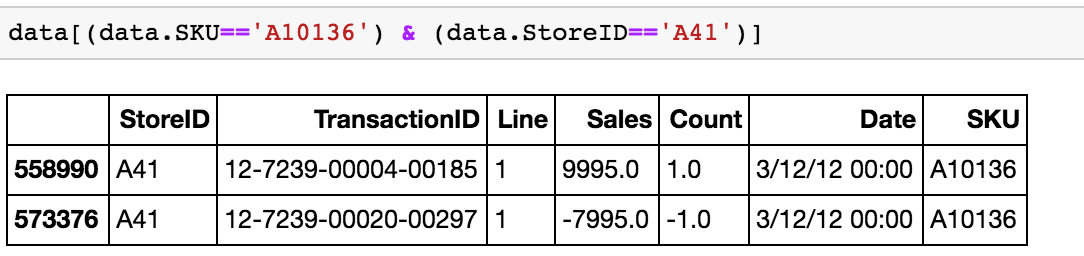
Number of cases on SKU level per store, per day: 957



**Explanation:** Sales cancel each other out.

c. Total Sales>0, Total Count=0

Number of cases on SKU level per store, per day: 33

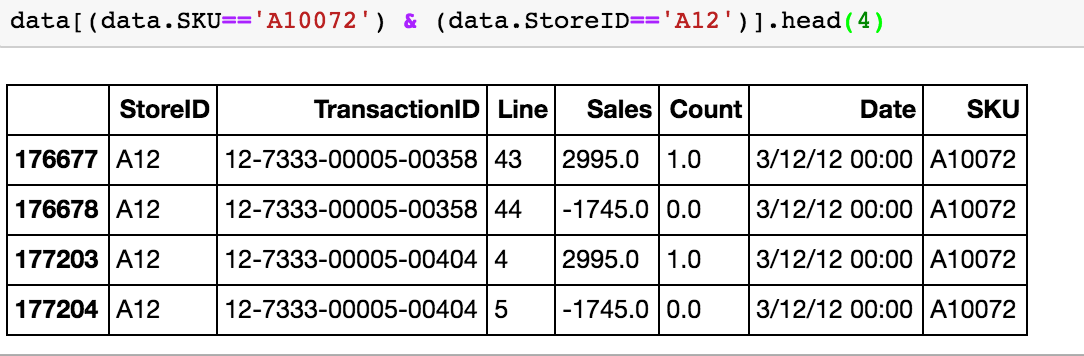


**Explanation:** Returned price is lower than the regular price and total count equals to 0.

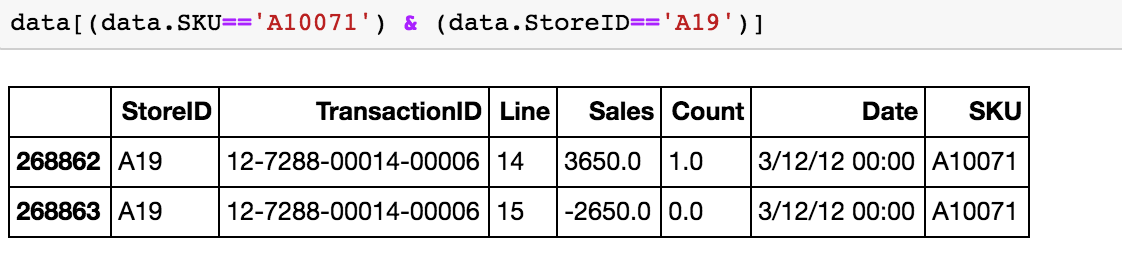
**scenario 10:** Discount rate is higher than 30%

Number of cases on SKU level per store, per day: 11831

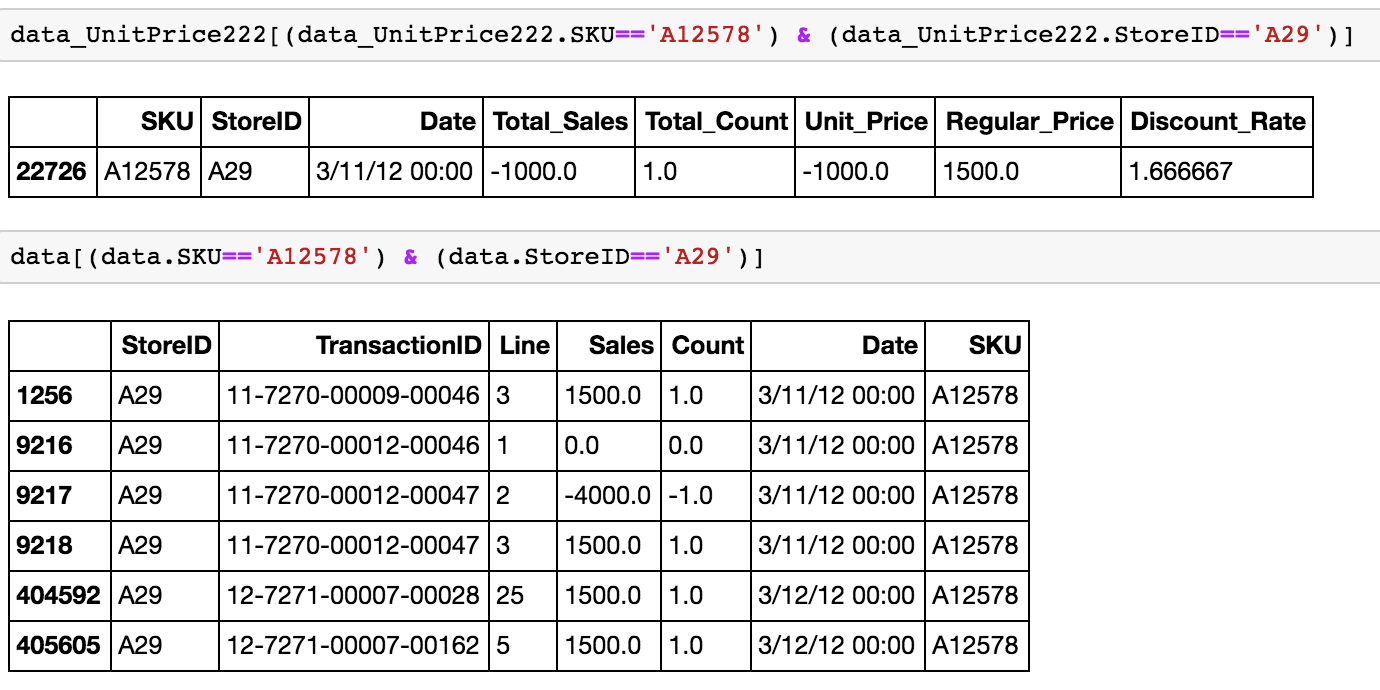
a. Discount Rate is 58%

****

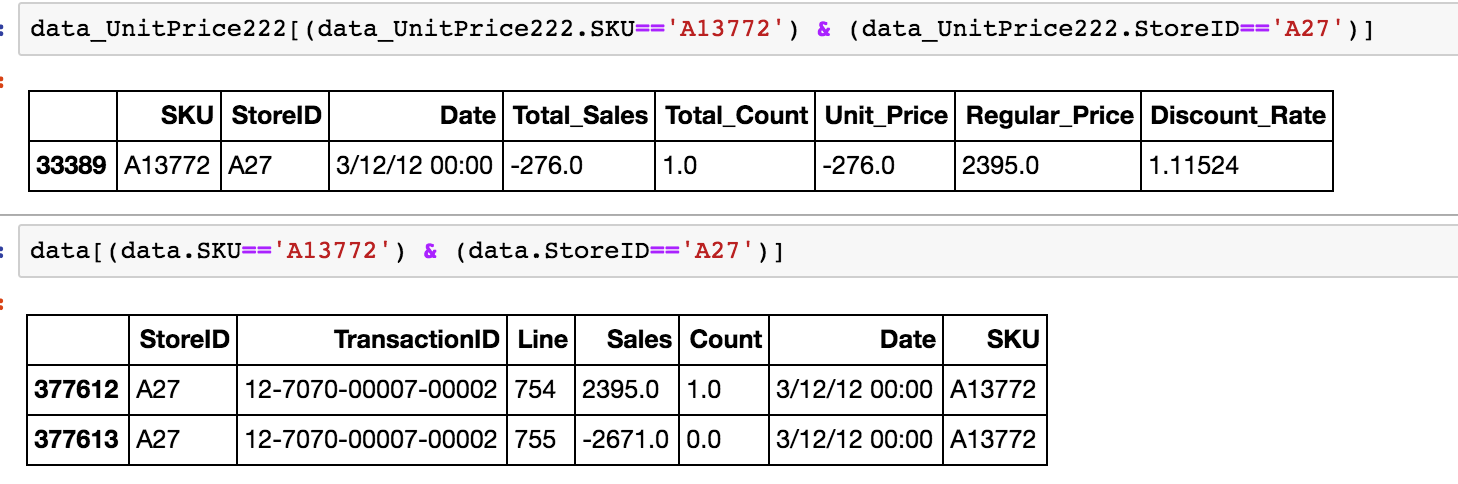
b. Discount Rate is 72%



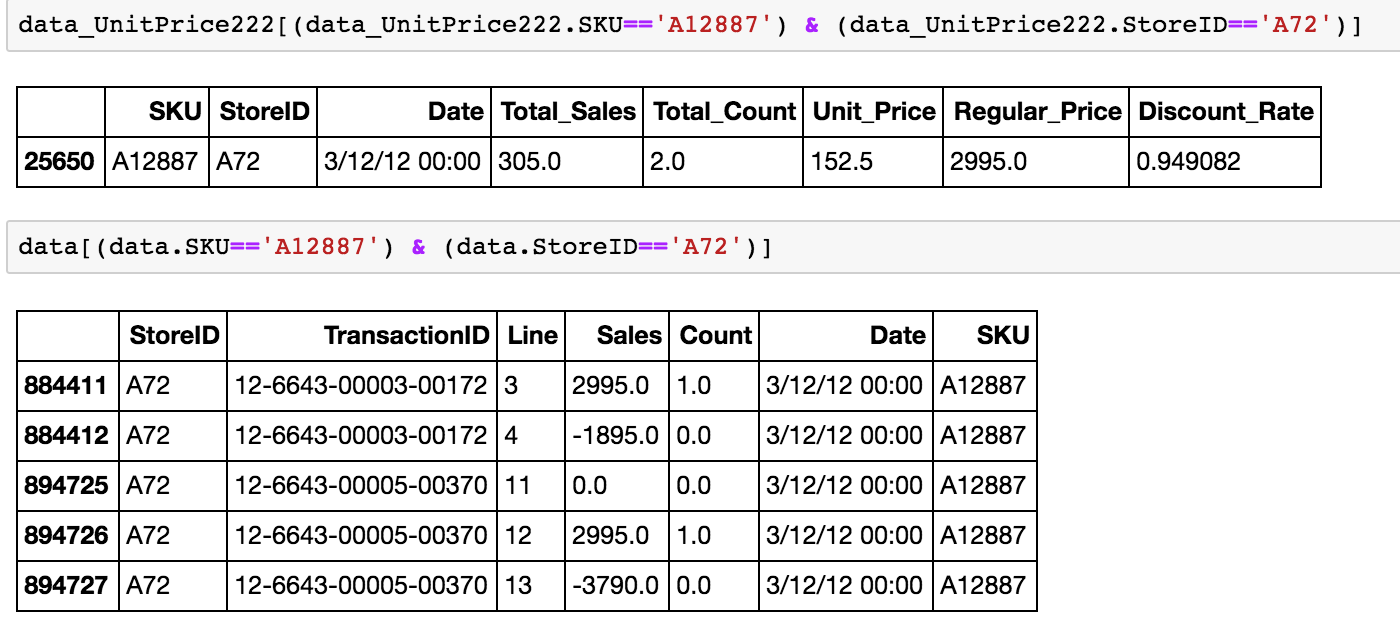
c. Returned price is higher than the regular price



d. Discount is larger than the regular price



e. Different discount value and discount is even larger than regular price



f. No credit back to the store when returning the discounted items

