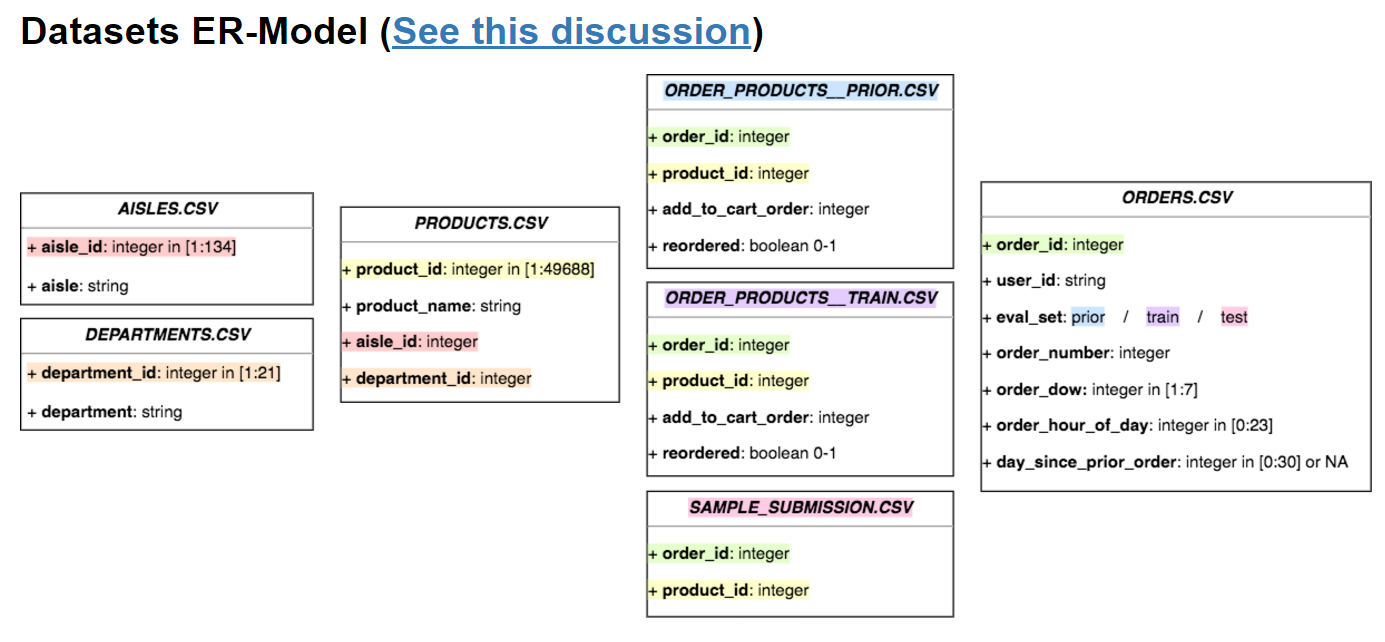
1. So there are 206,209 customers in total. Out of which, the last purchase of 131,209 customers are given as train set and we need to predict for the rest 75,000 customers.
2. Eval\_set tells us as to which of the three datasets (prior- 32434489, train- 1384617 or test- 75000) the given row goes to.
3. 
4. Orders.csv(3421083)+order\_products\_train.csv(1384617)+goods.csv=order\_details.csv(1384617)---on “order\_id”
5. Order\_details.csv+(concat)+merge(merge(op\_prior.csv(32434489),goods), orders)=order\_details.csv(33819114)
6. test\_orders = orders[orders.eval\_set == 'test']
7. Total products: 49688