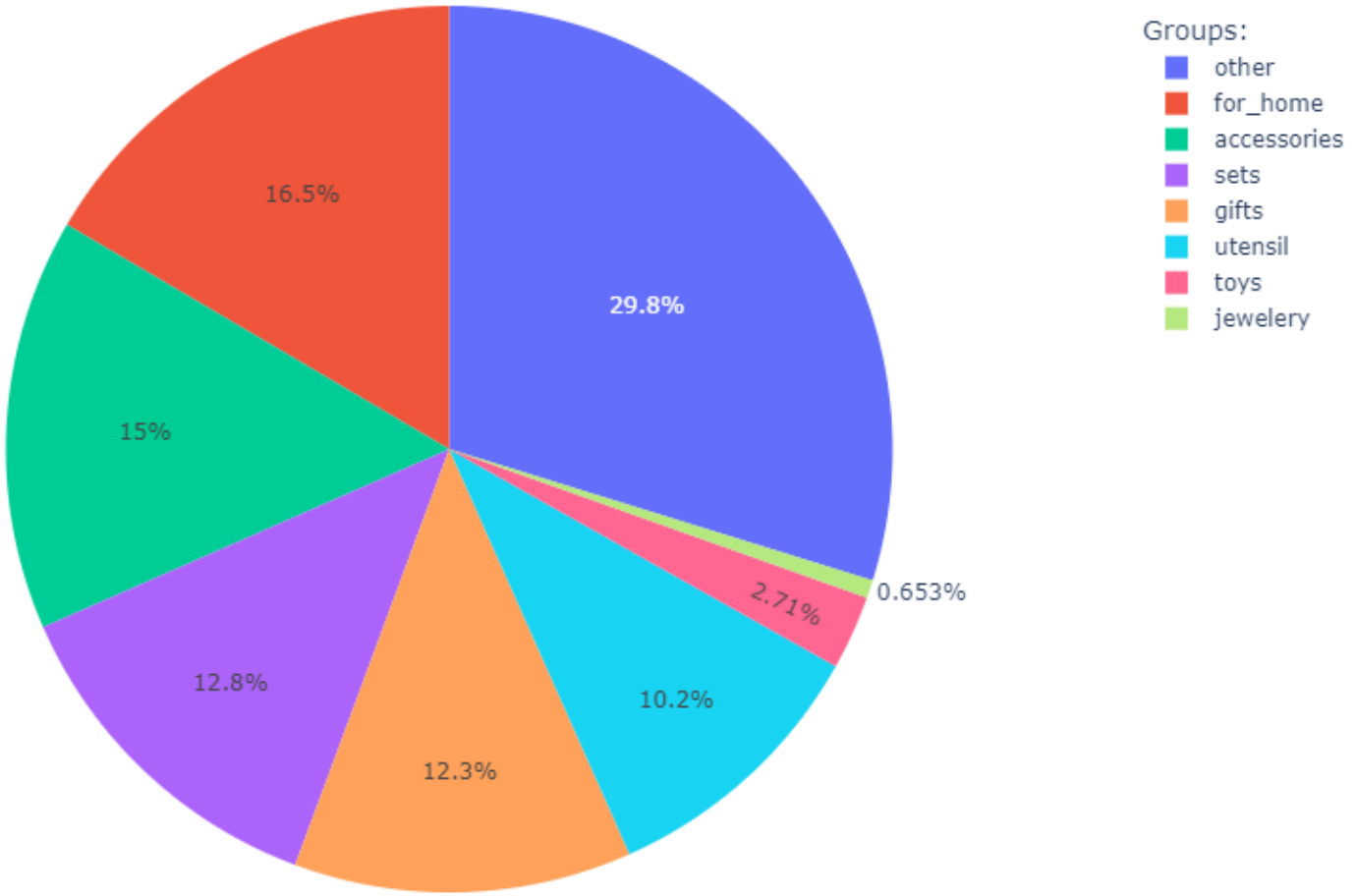
The results of the e-commerce project

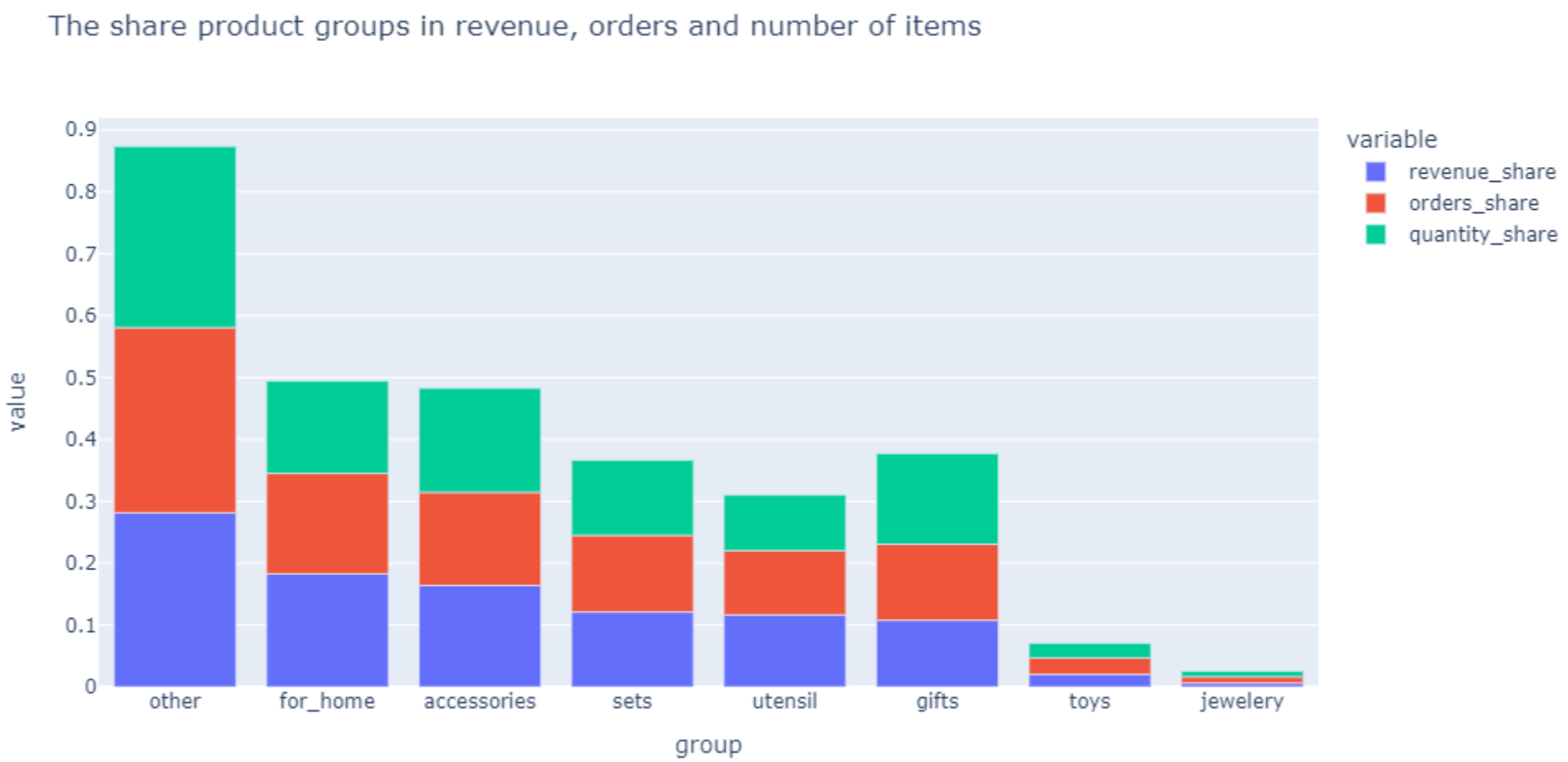
Ekaterina Dudakova 06/2022

Table of Contents

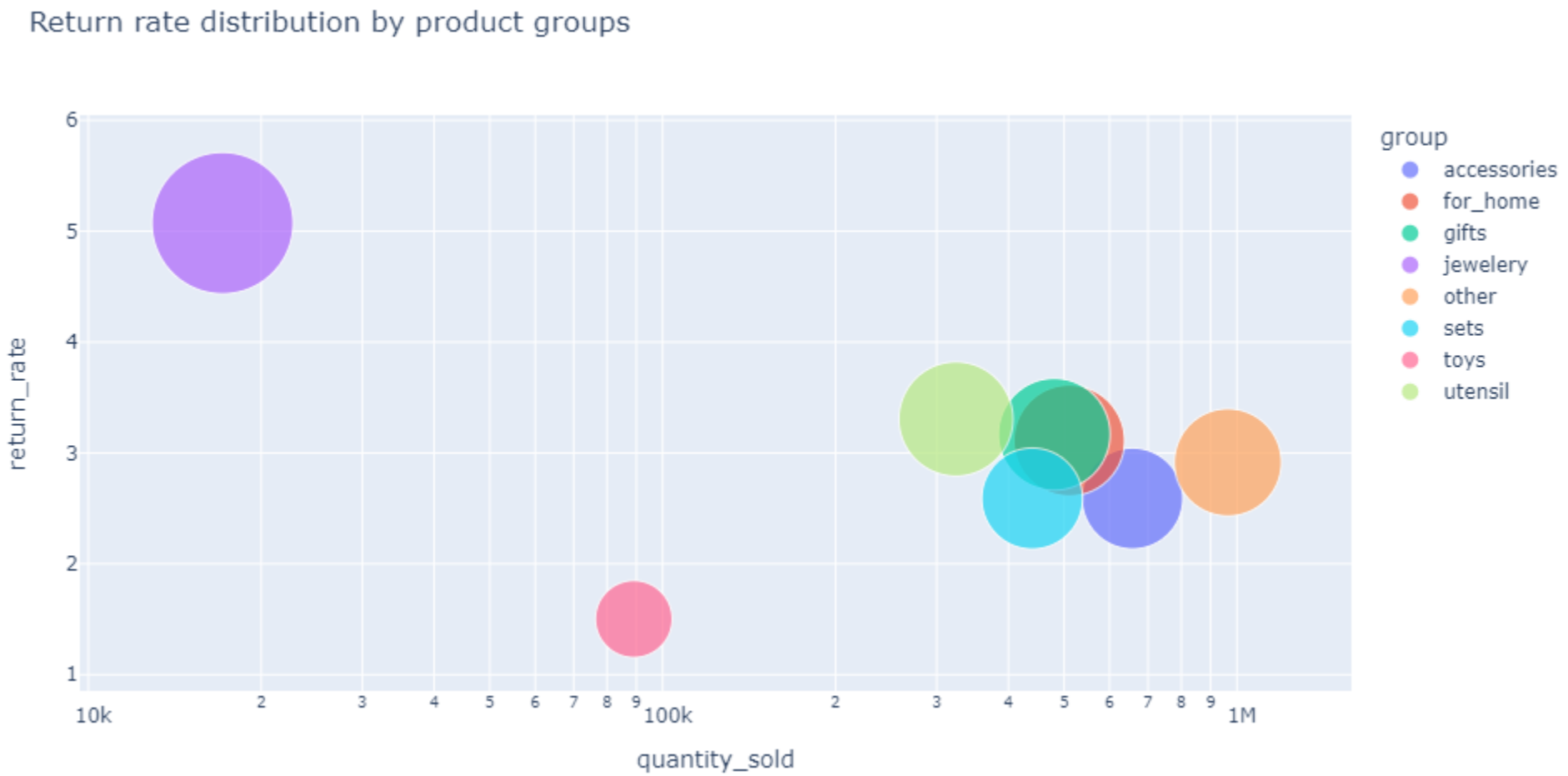
The proportions of the various product groups



- We had almost 4000 unique product descriptions, based on them grouped products by keywords which we determined by the lemmas of the words contained in the product description.
- The result – 8 categories of products.
- The biggest share in "other' group - it's logic, because we didn't manage separate around 30% of range our store - unfortunately, it's the biggest group.

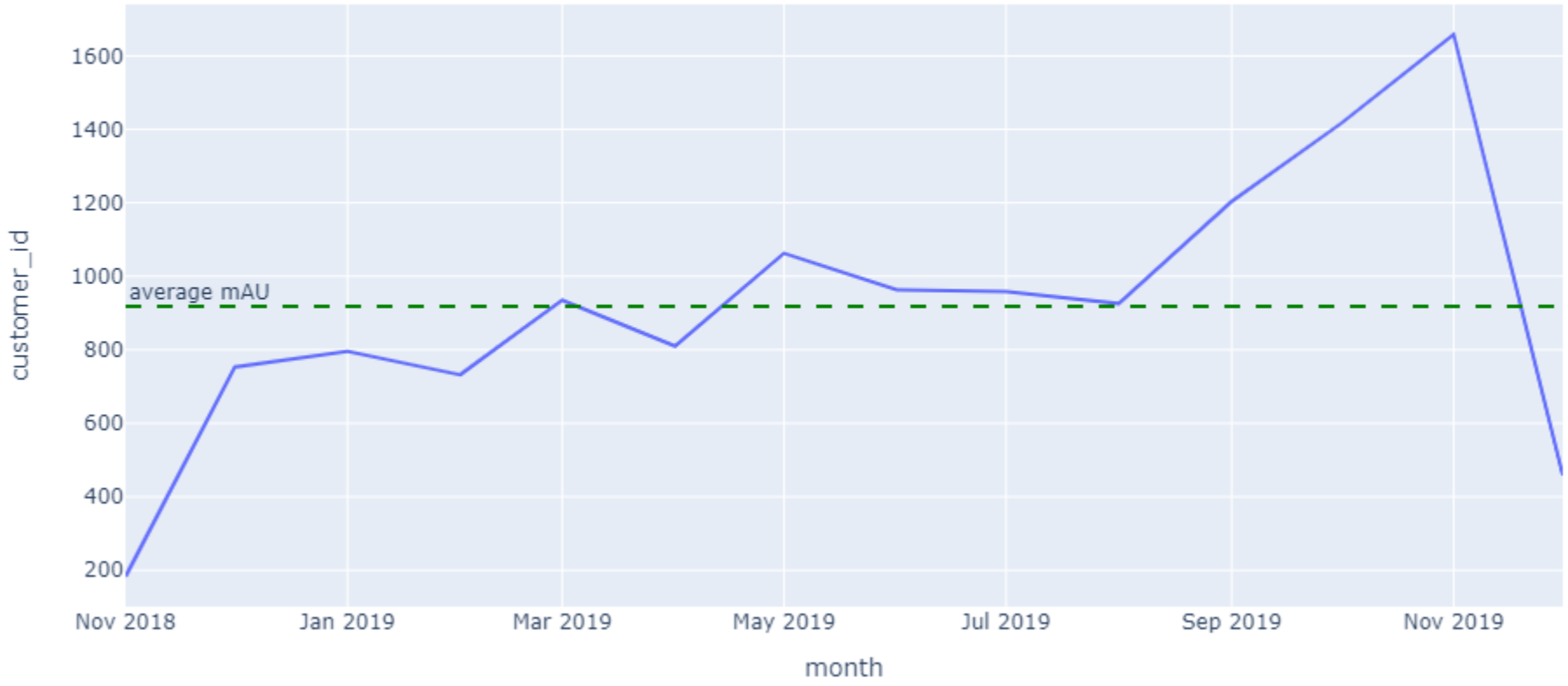


- In general, our groups have equal shares in revenue, number of orders and quantity, but
- For "utensil" and "for home" groups share in revenue bigger then share in orders (people ordered them less times but they brought more revenue), because of higher mean price per item,
- For "jewelry", "toys" and "gifts" it's opposite.

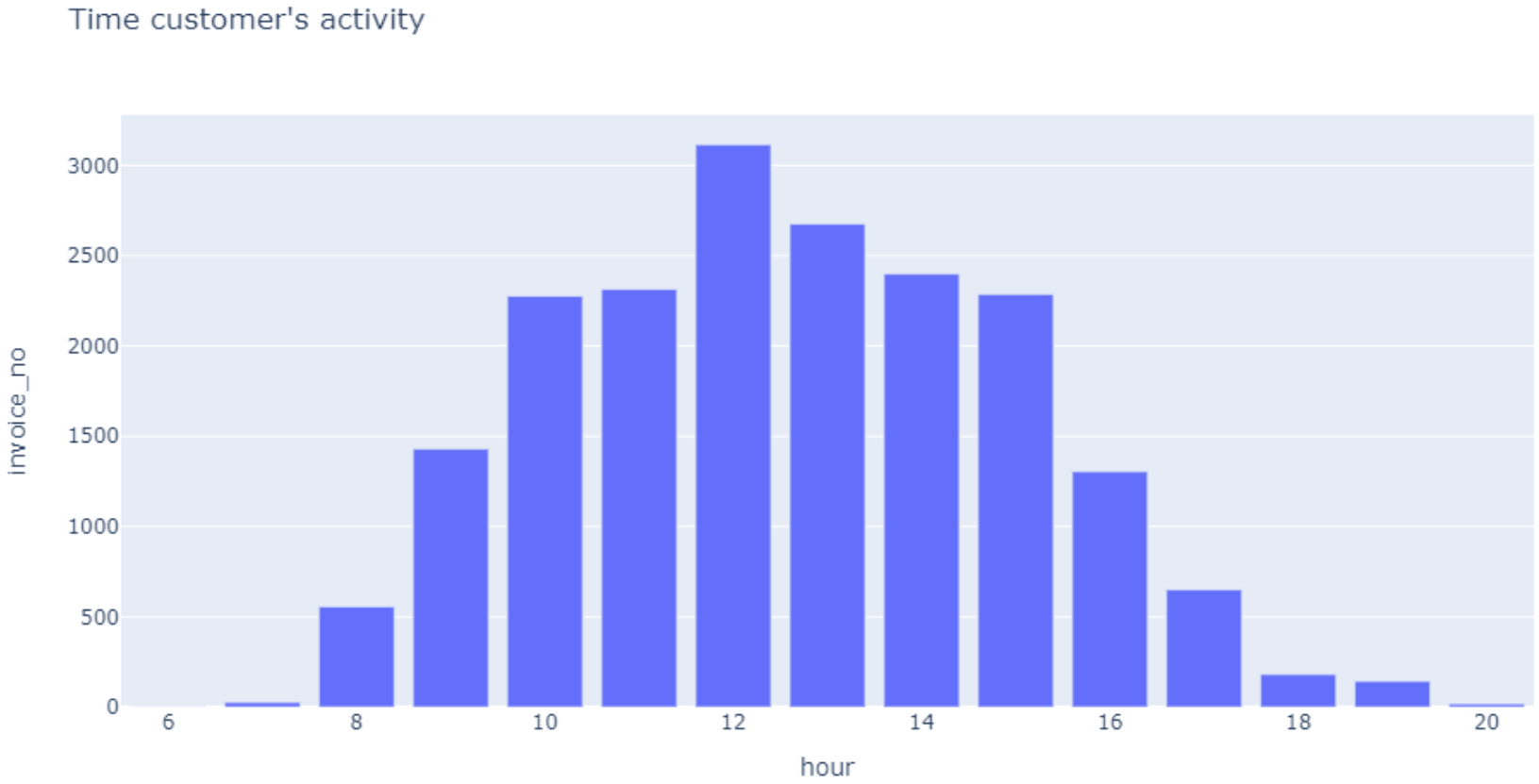


- For almost all groups return rate even lower - from 1.5% to 3.3%,
- But for "jewelry" it's higher than average - 5%, even though this group has the lowest number of products sold,
- For "utensil", 'gifts" and “for home” it's also a bit bigger then mean.

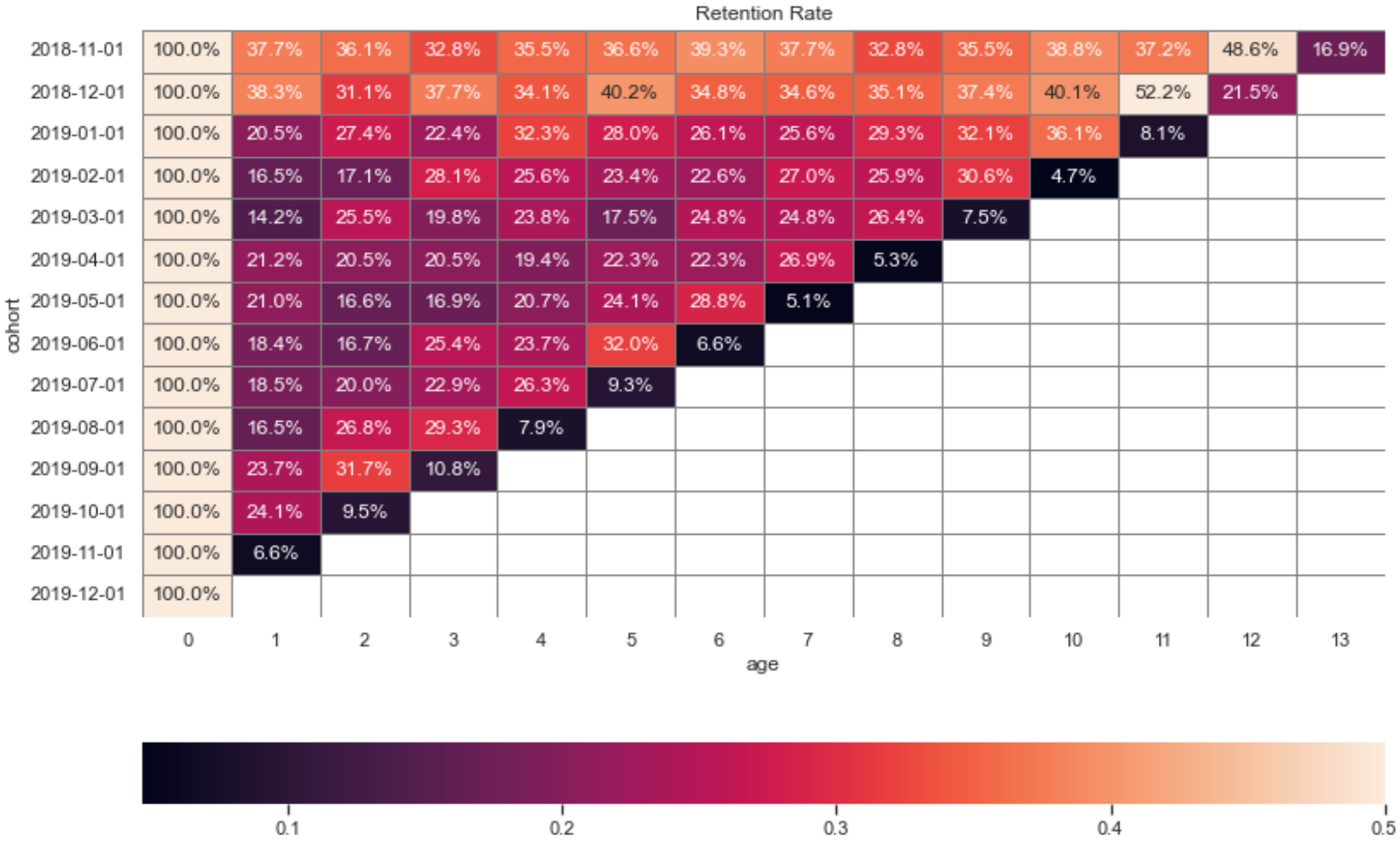
MAU dynamics.



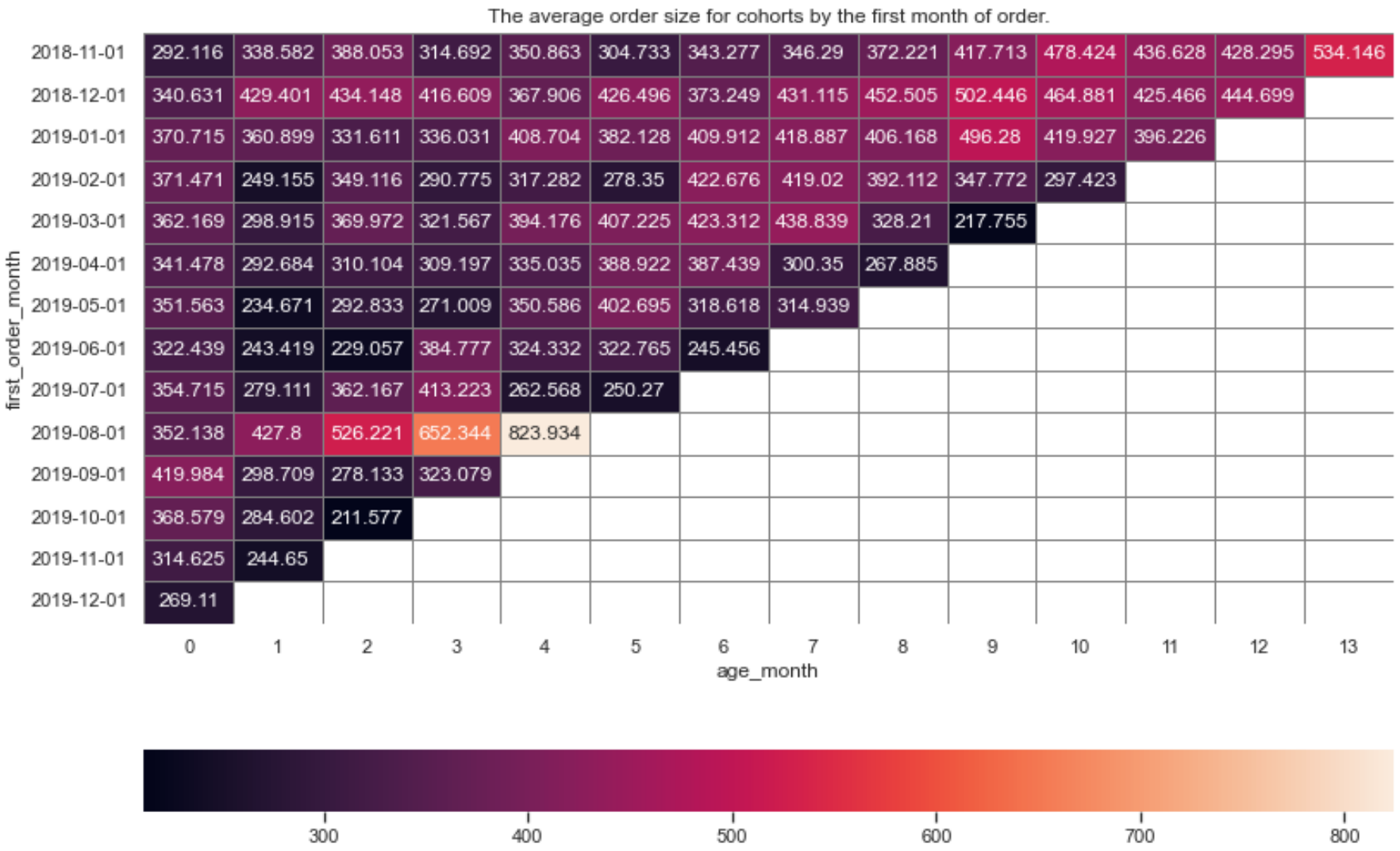
- The average number of unique customers: per day is 53, per month is 1049.
- In May, the number of customers reaches the average, and from August the growth begins until November (when people start buying goods for Christmas and New Year), after which the activity drops below the average.



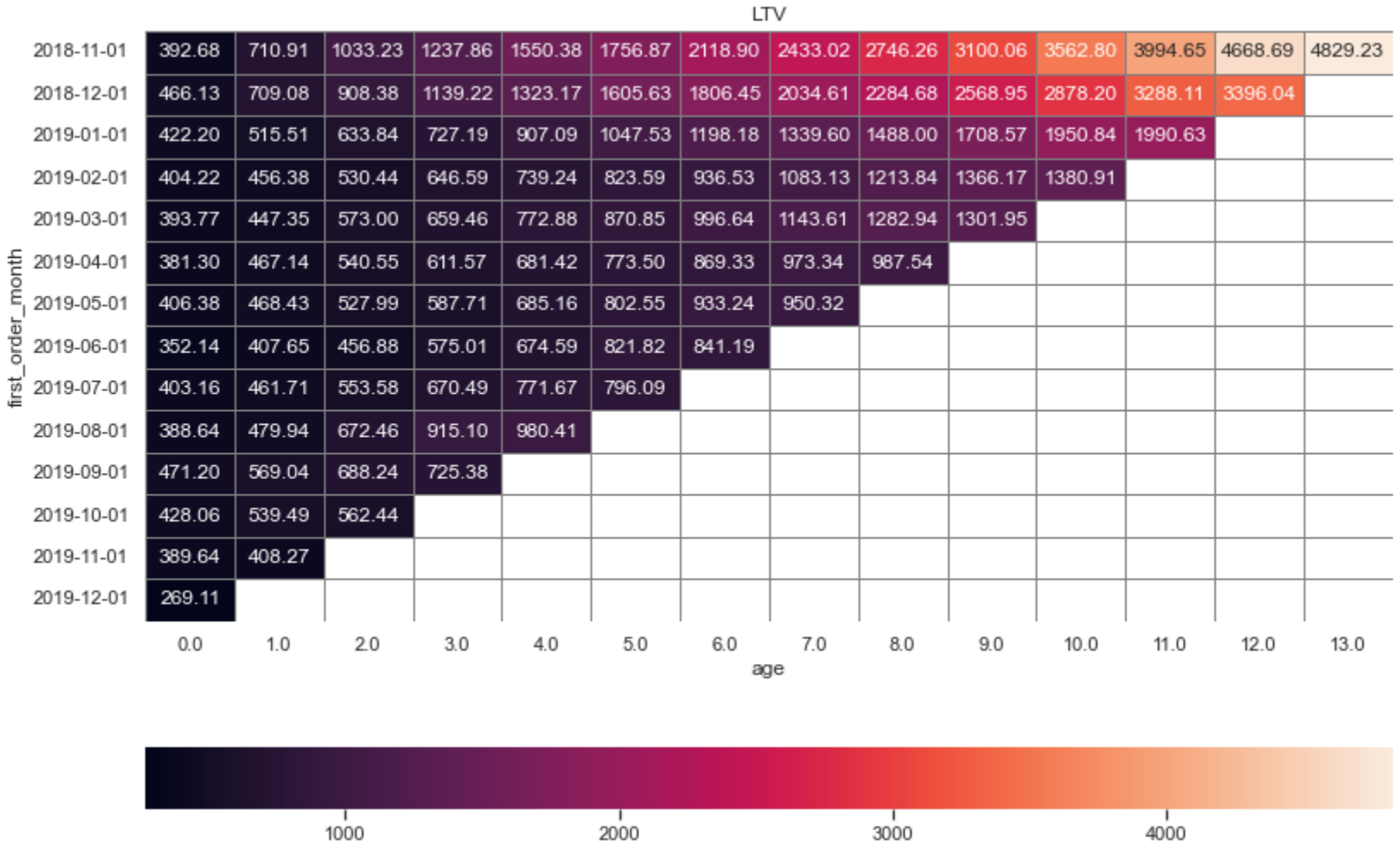
- There is normal distribution with peak at 12-13, the lunch time.



- Retention heatmap kindly remind MAU plot: we can see that December has the lowest retention for all cohorts, November's retention is the highest for all cohorts, some increases in January, March and May for all cohorts.
- In general, the first cohort is the most profitable compared to all other cohorts - we should pay attention for those customers.
- The retention rate is high but the overall trend of rate decrease.



- The average order size is 379,
- Overall order size is close to average for all cohorts,
- In the analysis of cohorts, the most interesting is August cohort - the only one cohort where average increases every month.



- The overall average LTV is 185.
- For the first three months, LTV for all cohorts grows approximately the same except first two cohorts - they are leading every month (November 2018 and December 2018), and August 2019 and January 2019 cohorts - their LTV grows not so fast as for first two cohorts, but faster then for rest.

Recommendations

- For future - it's better to use machine learning to divide resent product range by group (latent semantic analysis or word-embeddings, for example) and in accordance with them, mark each new product that will be added. Marked product range can be analyzed more effective.
- The store has to pay more attention "utensil" and "for home" categories, they look promising for expanding the range and increasing sales; and to "jewelry" - this category has lowest number of orders and highest return rate, this fact should be studied separately - may be it's better to stop sale jewelry.
- We know that our customers more active at afternoon - we can use it for advertisement to show them more propositions to buy.
- -In general, the first retention cohort is the most profitable compared to all other cohorts - we should pay attention for those customers. Also, August cohort by month of first order - it's the one cohort where average purchase size increases every month. From the point of view of LTV, August cohort is also worth paying attention, but the most interesting leaders - November 2018 and December 2018.