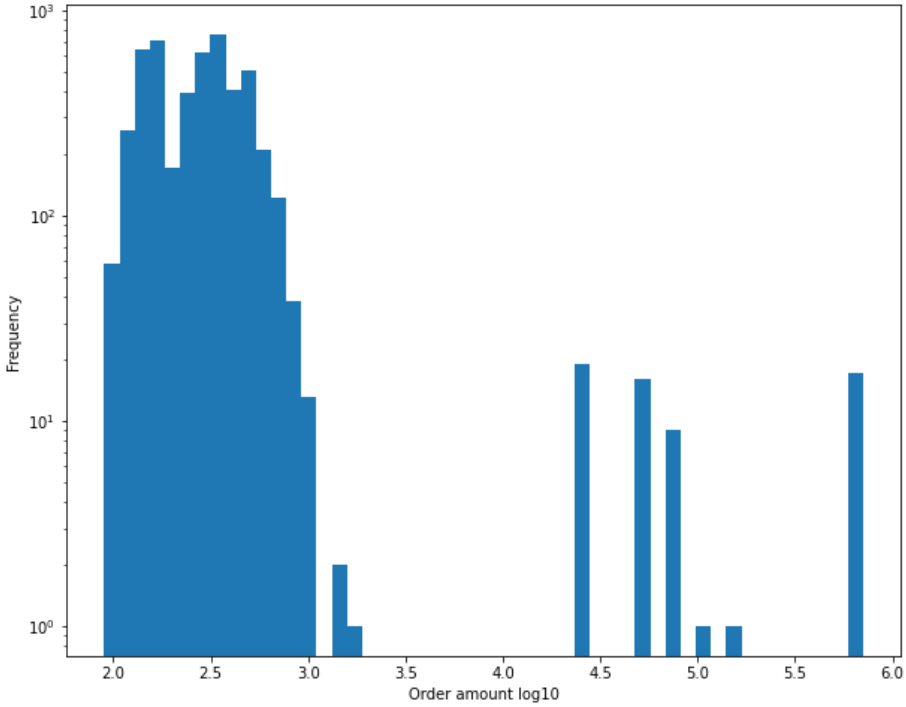
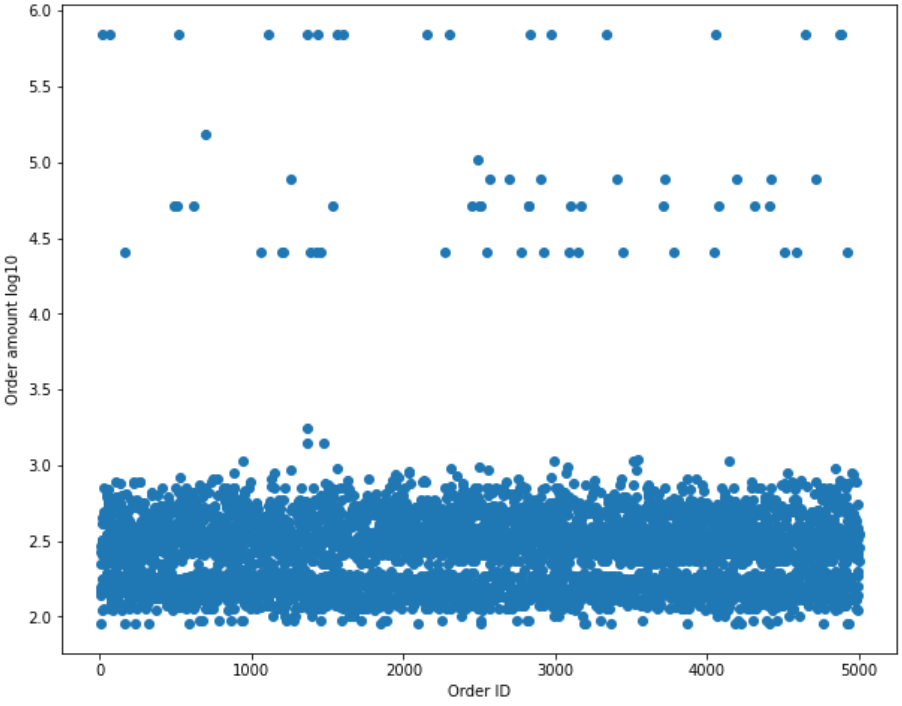
We plot the order amount as histogram and scatter plot to examine the data.





As the dataset is shown to be heavily skewed, using median for AOV is better than using mean.
Median order amount for entire dataset is 284.0

Median total items for entire dataset is 2.0

For further information, we can see that the order amount can roughly be split into 3 types: regular, large, and very large.
Median order amount for regular orders is 284.0
Median total items for regular orders is 2.0
The regular orders represent most of the orders, which consist of a one or a few pairs of sneakers that cost several hundred dollars.

Median order amount for large orders is 51450.0
Median total items for large orders is 2.0
The large orders represent the sales of expensive sneakers, which consist of a one or a few pairs of sneakers that cost tens of thousands of dollars.

Median order amount for very large orders 704000.0
Median total items for very large orders 2000.0
We can further examine that this type of very large order is a regular bulk order from the same store by the same user.

	order_id	shop_id	user_id	order_amount	total_items	payment_method	created_at
15	16	42	607	704000	2000	credit_card	2017-03-07 4:00:00
60	61	42	607	704000	2000	credit_card	2017-03-04 4:00:00
520	521	42	607	704000	2000	credit_card	2017-03-02 4:00:00
1104	1105	42	607	704000	2000	credit_card	2017-03-24 4:00:00
1362	1363	42	607	704000	2000	credit_card	2017-03-15 4:00:00

Conclusions:

A. Think about what could be going wrong with our calculation. Think about a better way to evaluate this data.

   The data is skewed by some outliers that have high order amounts, using mean for the average will get values that are too big.

B. What metric would you report for this dataset?

   Median is a better metric.

C. What is its value?

   $284.0