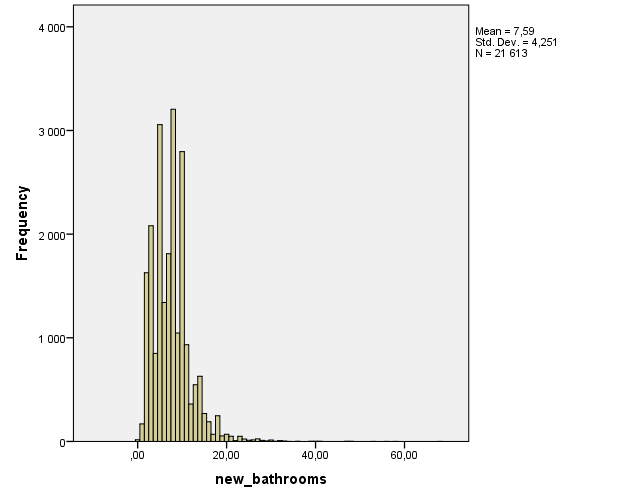
**Bathrooms**

This variable represent the number of bathroom per house.

The number of bathroom is between and 68. It’s kinda strange

Below there is a histogram of bathroom per house.



We can see the majority of the houses has between 2 and 10 bathroom (17813/21613=82%)

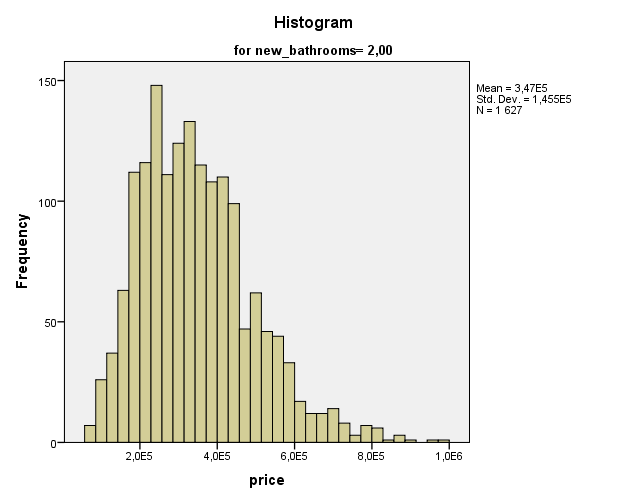
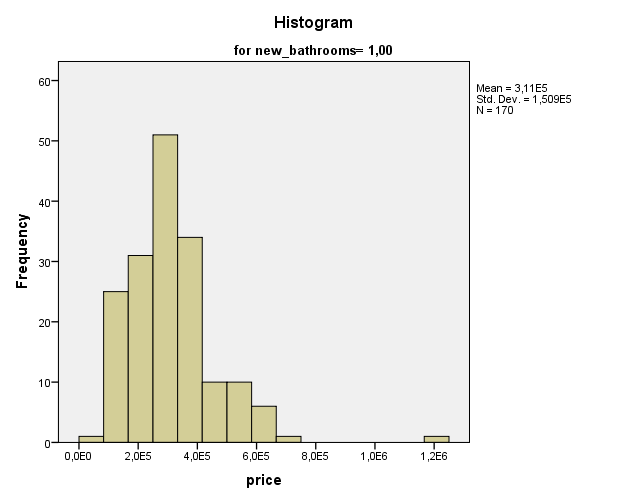
**Basic statistics**

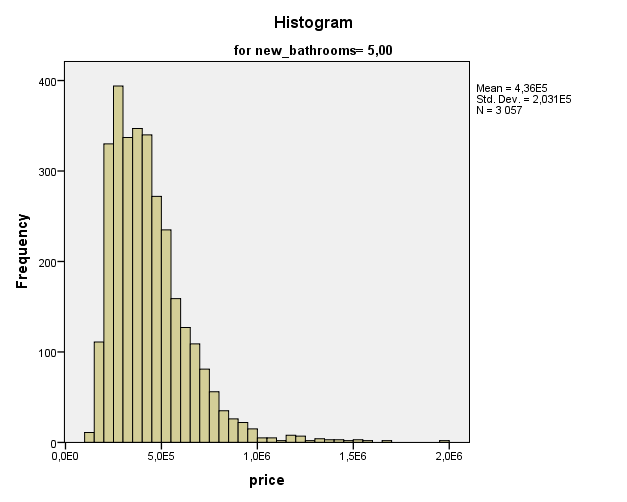
There are the statistics for houses price ($) acording to the number of bathroom. Because there are a lot of different number of bathroom per house, i’ll just present the statistics for some category. Maybe we should group the value (ie 0-5,6-10,11-15 …).

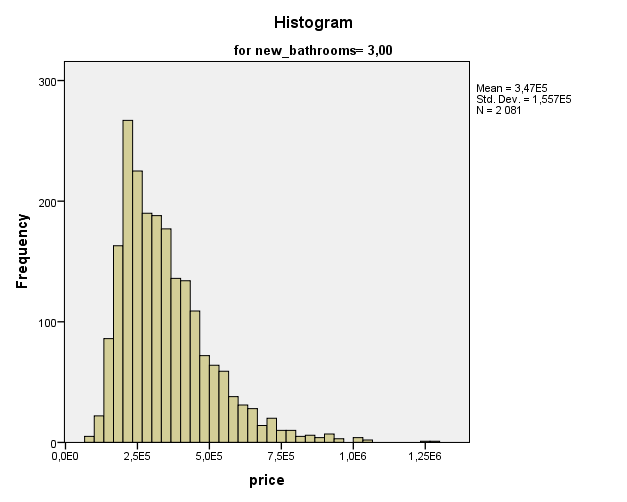
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 5 | 8 | 10 | 20 |
| Mean | 3,1\*10^5 | 3,5\*10^5 | 3,5\*10^5 | 4,3\*10^5 | 5,2\*10^5 | 5,9\*10^5 | 1,3\*10^6 |
| Median | 2,9\*10^5 | 3,3\*10^5 | 3,2\*10^5 | 4\*10^5 | 4,7\*10^5 | 5,4\*10^5 | 1\*10^6 |
| Variance | 2,2\*10^10 | 2,1\*10^10 | 2,4\*10^10 | 4,1\*10^10 | 6\*10^10 | 9\*10^10 | 4,3\*10^10 |
| Std Deviation | 1,5\*10^5 | 1,5\*10^5 | 1,6\*10^5 | 2\*10^5 | 2,5\*10^5 | 3\*10^5 | 6,6\*10^5 |
| Minimum | 8\*10^4 | 7,8\*10^4 | 8,2\*10^4 | 1\*10^5 | 1,5\*10^5 | 1,3\*10^5 | 2,7 \*10^5 |
| Maximum | 1,2\*10^6 | 1\*10^6 | 1,3\*10^6 | 2\*10^6 | 3\*10^6 | 3,6\*10^6 | 3\*10^6 |
| Range | 1,7\*10^5 | 9,2\*10^5 | 1,2\*10^6 | 1,9\*10^6 | 2,9\*10^6 | 3,5\*10^6 | 2,8\*10^6 |

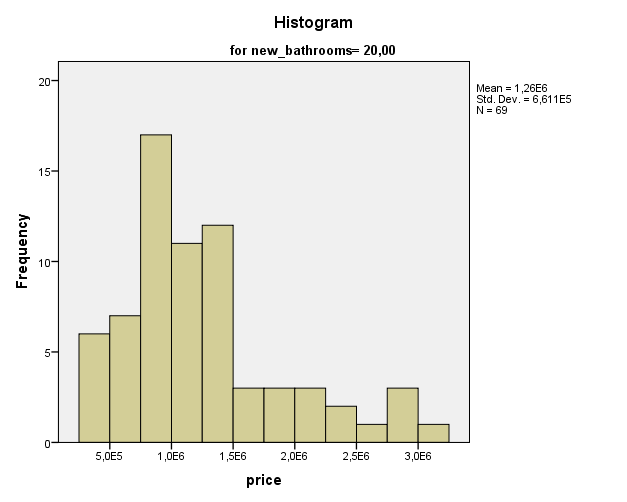
It seems to have a relation between price and the number of bathrooms, more there are bathrooms, more expensive is the house.

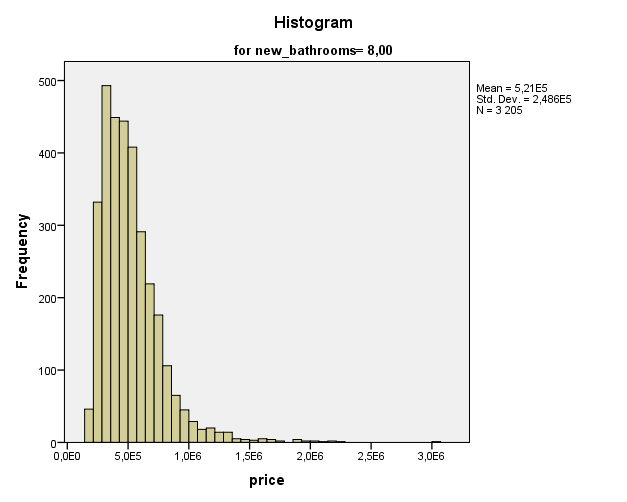
**Distribution of prices**





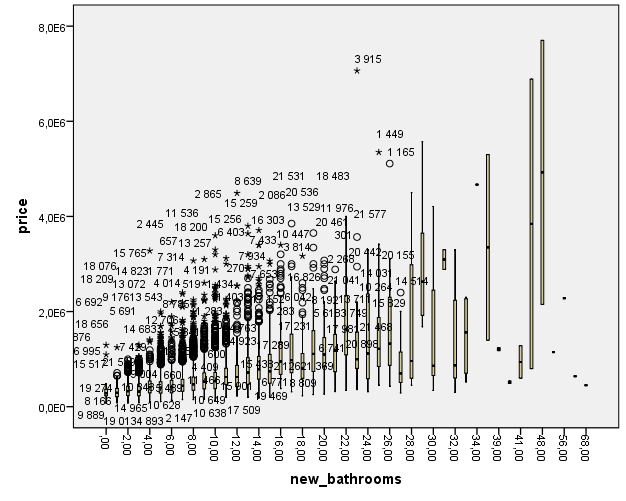






As see before, the distribution of the price seem to be higher when the number of bathroom increases.

**Outliers**



With this graphic we can see there is not below outliers. It’s probably because an house can’t be sold below 0.

There are a lot of outliers in absolute value but in percentage it’s low. In fact this graph is quite difficult to read because of the number of different category. We probably should group the data, we have to talk about it with the teacher.

Below, the 5 highest outliers for houses price by number of bathrooms

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | 1 | 2 | 3 | 4 | 5 |
| 1 | ID | 8446 | 9889 | 8166 | 19274 | 9564 |
| Price | 1,2\*10^6 | 7,1\*10^5 | 6,6\*10^5 | 6,5\*10^5 | 6,5\*10^5 |
| 2 | ID | 13029 | 7320 | 7197 | 1629 | 19337 |
| Price | 1\*10^6 | 9,7\*10^5 | 9\*10^5 | 8,8\*10^5 | 8 ,7\*10^5 |
| 3 | ID | 9548 | 4560 | 21549 | 18686 | 7477 |
| Price | 1,3\*10^6 | 1,3\*10^6 | 1\*10^6 | 1\*10^6 | 1\*10^6 |
| 5 | ID | 6692 | 9176 | 18076 | 15765 | 10983 |
| Price | 2\*10^6 | 2\*10^6 | 1,7\*10^6 | 1,7\*10^6 | 1,6\*10^6 |
| 8 | ID | 657 | 13072 | 19685 | 7701 | 541 |
| Price | 3\*10^6 | 2,3\*10^6 | 2,2\*10^6 | 2,2\*10^6 | 2,1\*10^6 |
| 10 | ID | 2865 | 18200 | 11536 | 4191 | 11226 |
| Price | 3,6\*10^6 | 3,3\*10^6 | 3,2\*10^6 | 2,9\*10^6 | 2,6\*10^6 |
| 20 | ID | 301 | 20442 | 21041 | 2268 | 10264 |
| Price | 3\*10^6 | 3\*10^6 | 2,9\*10^6 | 2,8\*10^6 | 2,7\*10^6 |