1. host\_reponse\_rate: The outliers are close to zero, but it is meaningful because a low reponse rate will impact the price negatively.

2. host\_identity\_verified: no outlier

3. host\_total\_listings\_count: there are extremely high counts, which may result from real estate companies, needs to be removed

4. city: it is encoded with number, needs to be hot\_coded ?

5. is\_location\_exact: no outlier

6. property\_type: one hot code

7. room type: one hot code

8. accommodates: check the correlation with other features and decide later

9. bathrooms, bedrooms, beds and bed\_type are corellated, keep them.

10. amentities: keep

11. price: highly skewed, remove the highest outliers.

12. minimun nights: remove highest outlier

13. number\_of\_reviews: keep, because this number reflects many facts of the pricing and it should be kept

14. review\_scores\_rating: check correlation to decide

15. instant\_bookable: keep

16. cancellation\_polity: one hot code

17. reviews\_per\_month: select one between this and the number of reviews if not important.