**Open Source SW Contribution**

**1. combination of various data scaling and categorical features encoding methods.**

9 combinations of each encoding, scaling methods.

One\_hot encode - standard scale

- min-max scale

- robust scale

Label encode - standard scale

- min-max scale

- robust scale

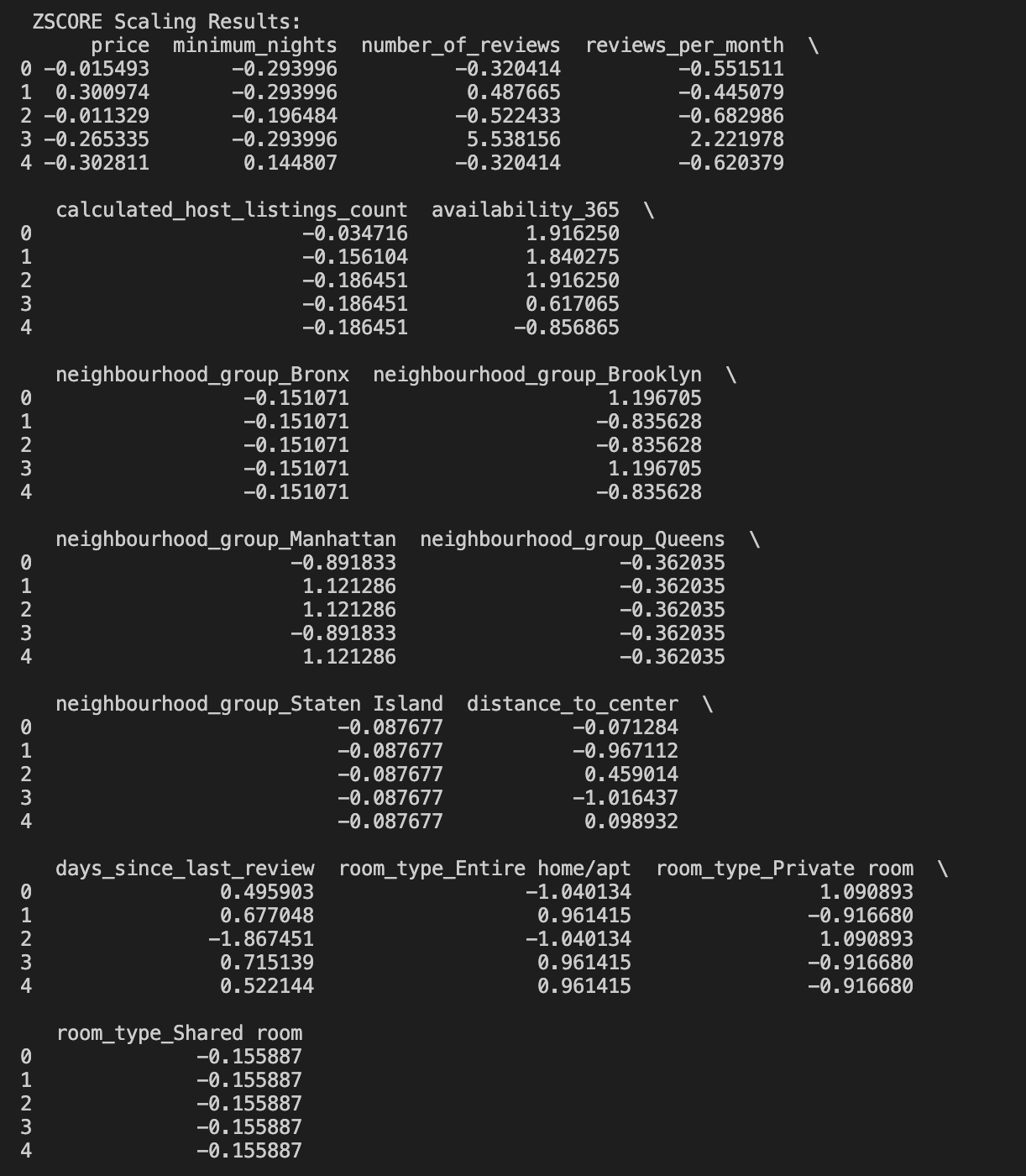
Ordinal encode - standard scale

- min-max scale

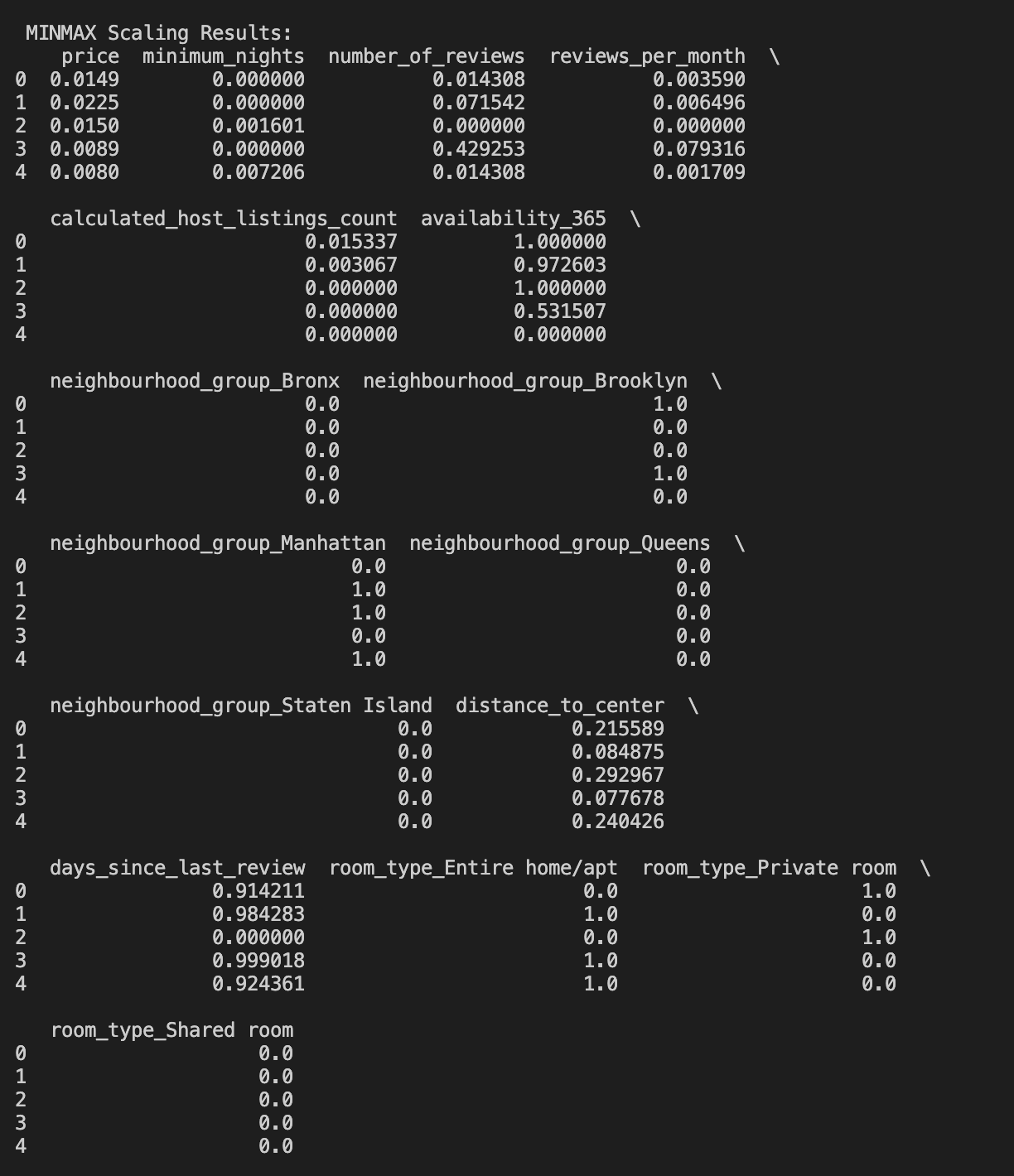
- robust scale

There are results of each method combinations below.

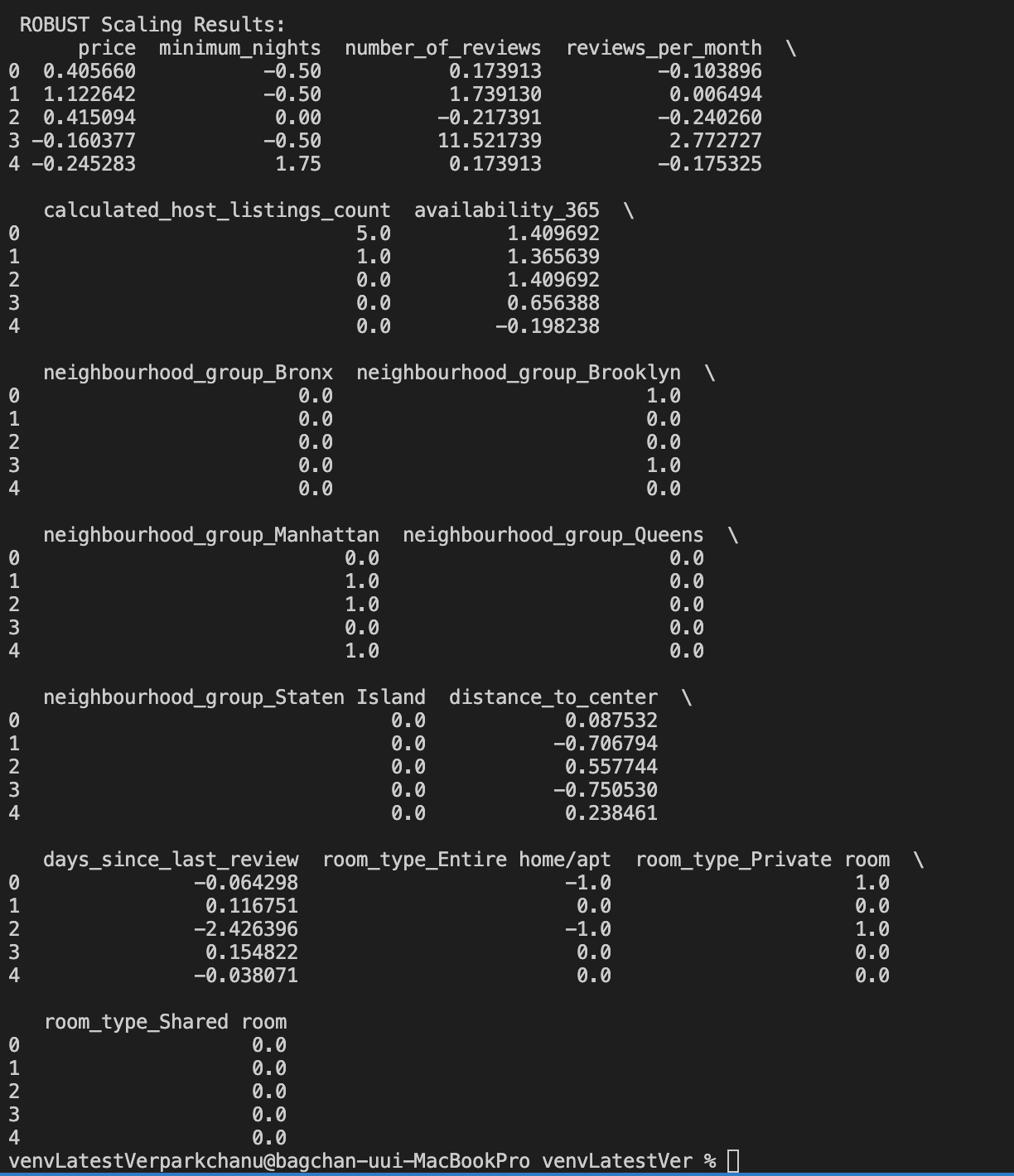
One-hot encode & Standard(Z-score) scale:



One-hot encode & Min-max scale:

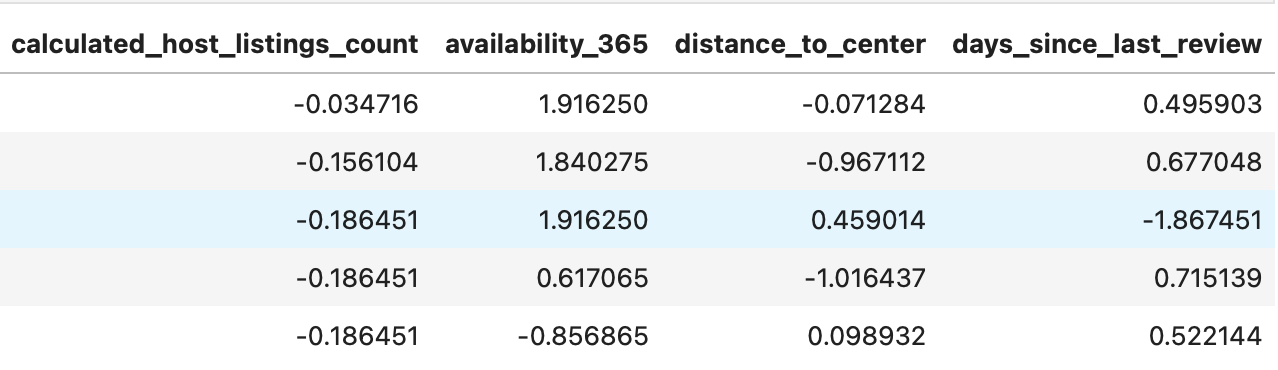


One-hot encode & Robust scale:

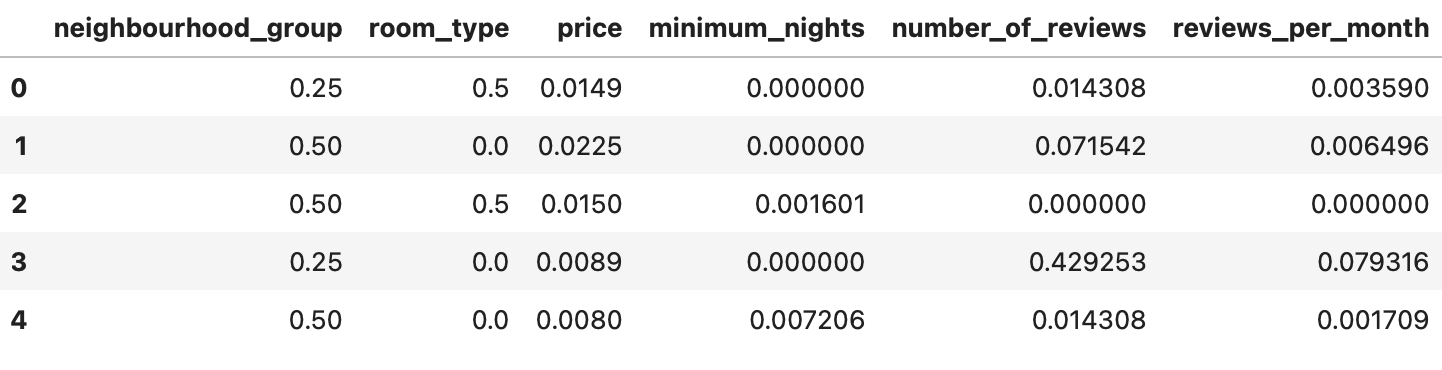


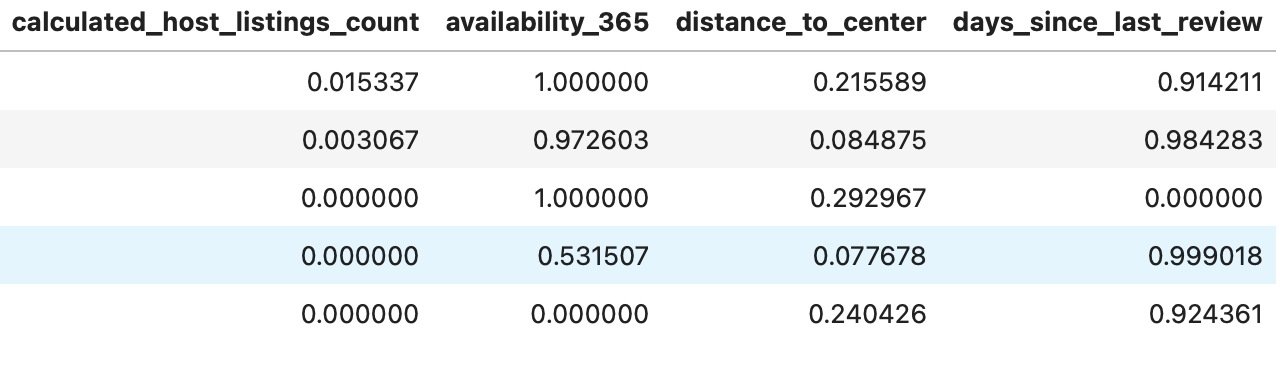
Label encode - standard scale:





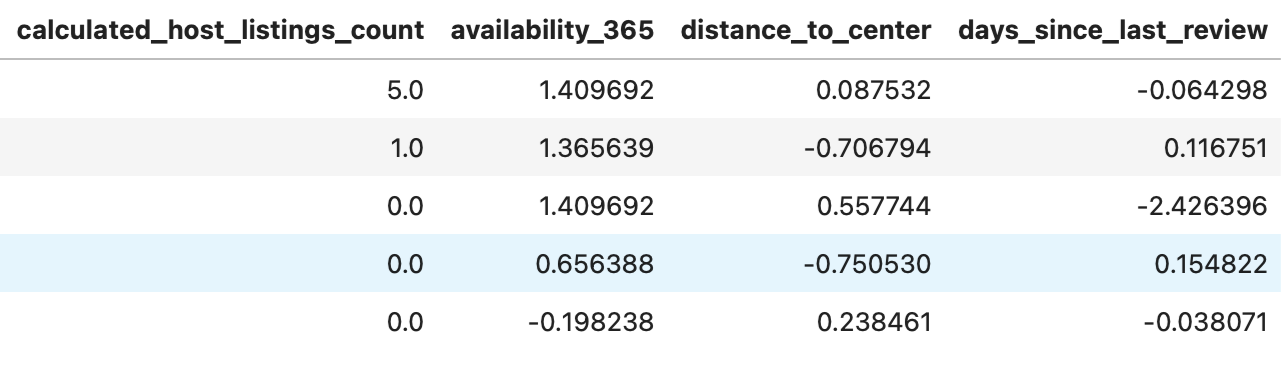
Label encode – Min-max scale:



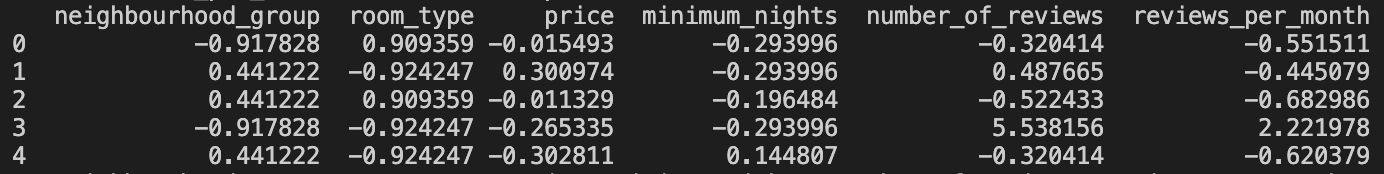


Label encode – Robust scale:





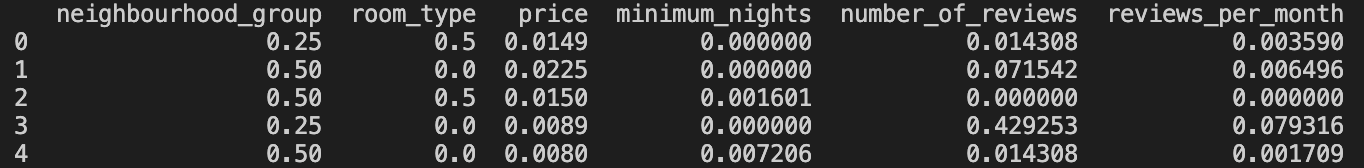
Ordinal encode - standard scale:

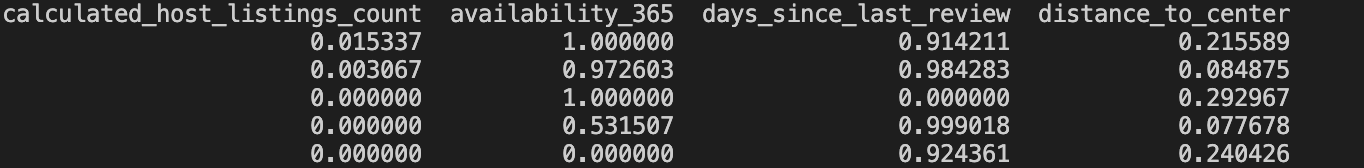


텍스트, 스크린샷, 폰트, 블랙이(가) 표시된 사진

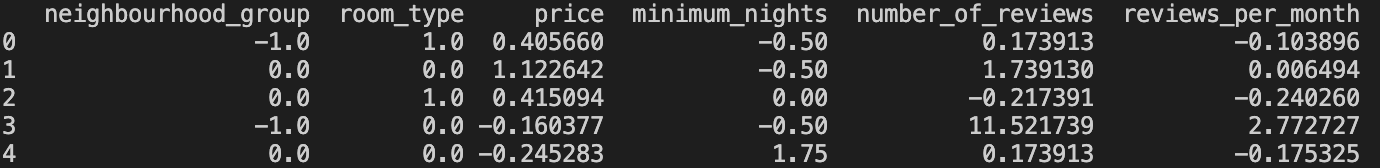
AI가 생성한 콘텐츠는 부정확할 수 있습니다.

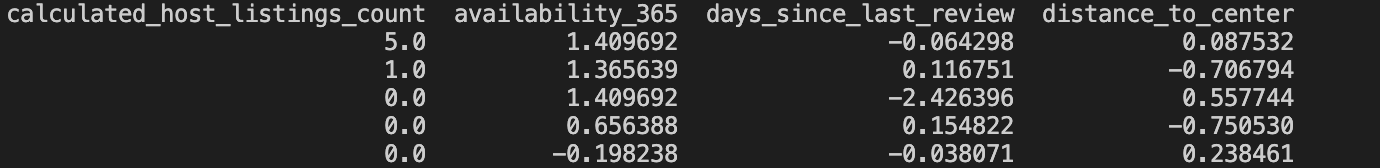
Ordinal encode – Min-max scale:





Ordinal encode – Robust scale:





From the above results, we can see that the label encoding and the ordinary encoding methods produce the same results. This seems to be because both methods map categorical data into integers in advance and start encoding. Therefore, we considered the label encoding and the Ordinal encoding to be the same, and compared the results of the one-hot encoding with a total of six results of the label encoding.

Therefore, we considered the label encoding and the Ordinal encoding to be the same, and compared the results of the one-hot encoding with a total of six results of the label encoding.

(클러스터링, 클래시피케이션 하려면 standard scailing 써야돼서 사실 선택지가 없음. Encode 방식만 비교하면 될듯. 이거 나중에 보강해야함 ….)