| | | Student ID: | NetID: |
|---|-------|-------------|--------|
| ſ | t) W | | |
| ı | 1 1 | | |
| ı | `++ m | | |

AE 04: Build better data (I)

Discussion questions

1. Identify at least four ways to represent the arrival_date column that would potentially be useful for predicting avg_price_per_room.

- 2. Arrange these data preprocessing/feature engineering steps in the correct order for a KNN model:
 - Center and scale numeric predictors
 - Remove highly correlated predictors
 - Convert arrival_date to indicators for day of week, month, year
 - Remove arrival_date
 - Indicators for holidays
 - Remove zero-variance predictors
 - Convert categorical predictors to binary indicators