**Documentation: Initial Correlation Analysis**

**Key Finding 1: No strong linear correlations with target variable**

* The correlation analysis revealed no significant linear relationships between hotel\_cluster (our target variable) and any other features in the dataset.
* This suggests that linear models alone might not be sufficient to capture the relationships in this data.

**Key Finding 2: Strong correlations between features**

* Strong negative correlation (-0.767) between posa\_continent and site\_name
  + This indicates that specific point-of-sale sites are strongly associated with particular continents
  + This relationship makes intuitive sense as users often use region-specific Expedia sites

**Key Finding 3: Moderate correlation between room and adult counts**

* Moderate positive correlation (0.521) between srch\_rm\_cnt (number of rooms) and srch\_adults\_cnt (number of adults)
  + This is expected as larger groups typically require more rooms

**Implication for modeling:** Since the target variable doesn't show strong linear correlations with individual features, we should explore:

1. Non-linear relationships
2. Feature interactions
3. More complex modeling approaches beyond simple linear models