ADBMS Project Review

Dataset selected:

Boston AirBNB Open Dataset.

This dataset has 3 files including every details about area wise apartment/house listings in Boston.

These files are: calendar.csv, listings.csv, revies.csv

Following are the patterns and analysis I want perform on this dataset.

- User comments analysis, Latest reviews about apartments -Filtering
- Number of reviews for every apartment- summarization
- Bloom filters- Finding good or bad comment
- For every neighborhood - minimum and maximum rents for apartments
- Number of listings per area- summarization patterns
- Number of transit options for a given area- Bloom filters.
- Listing and review join
- Binning pattern- 2 bedroom, 3 bedroom separations
- Hierarchy pattern(Organization) – Listing→ reviews.
- Range for ratings
- Top 5 listings per area- filtering pattern