Normalization

We have to tables to start with: Listings and Reviews

**1NF**

Listings

Id, listing\_url, scrape\_id, last\_scraped, source, name, description, neighborhood\_overview,picture\_url, host\_id, host\_url, host\_name, host\_location, host\_about, host\_response\_time, host\_response\_rate, host\_acceptance\_rate, host\_is\_superhost, host\_thumbnail\_url, host\_picture\_url, host\_neighbourhood, host\_listings\_count, host\_verifications, host\_has\_profile\_pic, host\_identity\_verified, neighbourhood, neighbourhood\_cleansed, neighbourhood\_group\_cleansed, latitude, longitude, property\_type, room\_type, accommodates, bathrooms\_text, beds, amenities, price, minimum\_nights, maximum\_nights, has\_availability, availability\_30, availability\_60, availability\_90, availability\_365, number\_of\_reviews, number\_of\_reviews\_ltm, number\_of\_reviews\_l30d, first\_review, last\_review, review\_scores\_rating, review\_scores\_accuracy, review\_scores\_cleanliness, review\_scores\_checkin, review\_scores\_communication, review\_scores\_location, review\_scores\_value, reviews\_per\_month

Reviews

listing\_id, review\_id, date, reviewer\_id, reviewer\_name comments

**2NF:**

host\_id ->

host\_url, host\_name, host\_location, host\_about, host\_response\_time,

host\_response\_rate, host\_acceptance\_rate, host\_is\_superhost,

host\_thumbnail\_url,host\_picture\_url, host\_neighbourhood, host\_listings\_count,

host\_verifications, host\_has\_profile\_pic, host\_identity\_verified

id ->

listing\_url, scrape\_id, last\_scraped, source, name, description, neighborhood\_overview,

picture\_url, latitude, longitude, property\_type, room\_type, accommodates, bathrooms\_text,

beds, amenities, price, minimum\_nights, maximum\_nights, has\_availability, availability\_30,

availability\_60, availability\_90, availability\_365, number\_of\_reviews, number\_of\_reviews\_ltm,

number\_of\_reviews\_l30d, reviews\_per\_month, first\_review, last\_review, review\_scores\_rating,

review\_scores\_accuracy, review\_scores\_cleanliness, review\_scores\_checkin,

review\_scores\_communication, review\_scores\_location, review\_scores\_value, latitude,

longitude, neighbourhood, neighbourhood\_cleansed, neighbourhood\_group\_cleansed

review\_id ->

listing\_id, review\_id, date, reviewer\_id, reviewer\_name comments

**3NF:**

host\_id ->

host\_url, host\_name, host\_location, host\_about, host\_response\_time,

host\_response\_rate, host\_acceptance\_rate, host\_is\_superhost,

host\_thumbnail\_url,host\_picture\_url, host\_neighbourhood, host\_listings\_count,

host\_verifications, host\_has\_profile\_pic, host\_identity\_verified

id ->

listing\_url, scrape\_id, last\_scraped, source, name, description, neighborhood\_overview,

picture\_url, latitude, longitude, property\_type, room\_type, accommodates, bathrooms\_text,

beds, amenities, price, minimum\_nights, maximum\_nights, has\_availability, availability\_30,

availability\_60, availability\_90, availability\_365, number\_of\_reviews, number\_of\_reviews\_ltm,

number\_of\_reviews\_l30d, reviews\_per\_month

latitude + longitude ->

neighbourhood, neighbourhood\_cleansed, neighbourhood\_group\_cleansed

id ->

first\_review, last\_review, review\_scores\_rating, review\_scores\_accuracy,

review\_scores\_cleanliness, review\_scores\_checkin, review\_scores\_communication,

review\_scores\_location, review\_scores\_value

review\_id ->

listing\_id, review\_id, date, reviewer\_id, reviewer\_name comments