**Final Columns**

'id'

'scrape\_id'

'last\_scraped'

'host\_id'

'host\_name',

'host\_is\_superhost'

'host\_total\_listings\_count'

'street'

'neighbourhood\_cleansed'

'city'

'state'

'zipcode'

'smart\_location'

'country\_code'

'country'

'latitude'

'longitude'

'property\_type'

'room\_type'

'accommodates'

'bathrooms'

'bedrooms'

'beds'

'bed\_type'

'amenities'

'price'

'weekly\_price'

'monthly\_price'

'security\_deposit'

'cleaning\_fee'

'guests\_included'

'extra\_people'

'minimum\_nights'

'maximum\_nights'

'has\_availability'

'availability\_30'

'availability\_60'

'availability\_90'

'availability\_365'

'number\_of\_reviews'

'review\_scores\_rating'

'review\_scores\_accuracy'

'review\_scores\_cleanliness'

'review\_scores\_checkin'

'review\_scores\_communication'

'review\_scores\_location'

'review\_scores\_value'

'reviews\_per\_month'

**Dropped these columns**

"name"

"summary"

"neighborhood\_overview"

"transit"

"access"

"interaction"

'host\_since'

'host\_location'

'host\_response\_time'

'host\_response\_rate'

'host\_acceptance\_rate'

"host\_listings\_count"

"host\_neighbourhood"

"neighbourhood"

"neighbourhood\_group\_cleansed"

“market"

"is\_location\_exact"

"calendar\_updated"

"first\_review", "last\_review", "square\_feet"

**MISSING VALUES**

**Changed Null (missing) values to “missing”:**

#host\_name

#city

**Changed Null (missing) values to “f”:**

#host\_is\_superhost

**Changed Null (missing) values to** "0" (zero)

host\_total\_listings\_count

bathrooms

bedrooms'

'beds'

square\_feet

price',

weekly\_price'

monthly\_price'

'security\_deposit'

'cleaning\_fee'

'guests\_included'

'extra\_people'

number\_of\_reviews

review\_scores\_rating',

'review\_scores\_accuracy'

'review\_scores\_cleanliness',

'review\_scores\_checkin'

'review\_scores\_communication',

'review\_scores\_location',

'review\_scores\_value'

'reviews\_per\_month'

**Changed Null (missing) values to**  "Victoria"

“State”

**Changed Null (missing) values to**  0000" - string

Zipcode

**CORRECTING DATA TYPES**

**Corrected to be floats (decimal numbers)**

bathrooms

bedrooms

beds

reviews\_per\_month

**Converting currency from string to float (decimal)**

'price'

weekly\_price

monthly\_price

security\_deposit

cleaning\_fee

extra\_people

**Corrected to be integers (whole numbers)**

review\_scores\_rating object

review\_scores\_accuracy object

review\_scores\_cleanliness object

review\_scores\_checkin object

review\_scores\_communication object

review\_scores\_location object

review\_scores\_value object

**CHECKING FOR STRANGE VALUES**

**host\_is\_superhost** OK

['f' 't']

**neighbourhood\_cleansed** OK  
'Banyule' 'Bayside' 'Boroondara' 'Brimbank' 'Cardinia' 'Casey' 'Darebin'

'Frankston' 'Glen Eira' 'Greater Dandenong' 'Hobsons Bay' 'Hume'

'Kingston' 'Knox' 'Manningham' 'Maribyrnong' 'Maroondah' 'Melbourne'

'Melton' 'Monash' 'Moonee Valley' 'Moreland' 'Nillumbik' 'Port Phillip'

'Stonnington' 'Whitehorse' 'Whittlesea' 'Wyndham' 'Yarra' 'Yarra Ranges']

**City** Messy – not fixed – do we need to fix?

**smart\_location** Messy – not fixed – do we need to fix?

**country\_code** OK

['AU']

**country** OK

['Australia']

**has\_availability** OK

['t']