

Feature name	Source	Value type	Scale	Data type	Unique count	Measure unit	Conversion factor	Min	Max	Outlier treatment	Null	Notes
listing_id	calendar_summery	Categorical	Nominal	Integer	22552			2015	29867350			
date	calendar_summery	Date	Time-series	Date	367							
aviliable	calendar_summery	Categorical	Nominal	Boolean	2							
price	calendar_summery	Numeric	Interval	Integer	684	dolar		9	8600		NA	
id	listing	Categorical	Nominal	Integer	22552			2015	29867350			
name	listing	Categorical	Nominal	Varchar	5997							
host_id	listing	Categorical	Nominal	Integer	22552			2217	2.25E+08			
host_name	listing	Categorical	Nominal	Varchar	21873							
neighbourhood_group	listing	Categorical	Nominal	Varchar	12							
neighbourhood	listing	Categorical	Nominal	Varchar	136							
latitude	listing	Numeric	Interval	Float	22552			52.345803	52.65167			
longitude	listing	Numeric	Interval	Float	22552			13.103557	13.757642			
room_type	listing	Categorical	Nominal	Varchar	3							
price	listing	Numeric	Interval	Integer	295			0	9000			
minimum_nights	listing	Numeric	Interval	Integer	102			1	5000			
number_of_reviews	listing	Numeric	Interval	Integer	306			1	498			
last_review	listing	Date	Time-series	Date	1312							
reviews_per_month	listing	Numeric	Interval	Float	768			0.01	36.67			
calculated_host_listings_count	listing	Numeric	Interval	Integer	23			1	45			
availability_365	listing	Numeric	Interval	Integer	366			0	365			
id	listing_summery	Categorical	Nominal	Integer	22552			2015	29867350			
listing_url	listing_summery	Categorical	Nominal	Varchar	22552							
scrape_id	listing_summery	Categorical	Nominal	Integer	1			2.02E+13	2.02E+13			
last_scraped	listing_summery	Date	Time-series	Date	2							
name	listing_summery	Categorical	Nominal	Varchar	21874						NA	59 missing values
summary	listing_summery	Categorical	Nominal	Text	21041						NA	963 missing values
space	listing_summery	Categorical	Nominal	Text	13598						NA	8532 missing values
description	listing_summery	Categorical	Nominal	Text	21997						NA	203 missing values
experiences_offered	listing_summery	Categorical	Nominal	Varchar	1							
neighborhood_overview	listing_summery	Categorical	Nominal	Text	10781						NA	11012 missing values
notes	listing_summery	Categorical	Nominal	Text	7215						NA	15337 missing values
transit	listing_summery	Categorical	Nominal	Text	12308						NA	9516 missing values
access	listing_summery	Categorical	Nominal	Text	9946						NA	11715 missing values
interaction	listing_summery	Categorical	Nominal	Text	9584						NA	12146 missing values
house_rules	listing_summery	Categorical	Nominal	Text	10350						NA	11103 missing values
thumbnail_url	listing_summery										NA	All values are missing
medium_url	listing_summery										NA	All values are missing
picture_url	listing_summery	Categorical	Nominal	Text	22465							
xl_picture_url	listing_summery										NA	All values are missing
host_id	listing_summery	Categorical	Nominal	Integer	19180			2217	2.25E+08			
host_url	listing_summery	Categorical	Nominal	Text	19180							
host_name	listing_summery	Categorical	Nominal	Varchar	5997						NA	26 missing values
host_since	listing_summery	Date	Time-series	Date	2914						NA	26 missing values
host_location	listing_summery	Categorical	Nominal	Text	1036						NA	116 missing values
host_about	listing_summery	Categorical	Nominal	Text	8879						NA	11363 missing values
host_response_time	listing_summery	Categorical	Nominal	Text	4						NA	12894 missing values
host_response_rate	listing_summery	Numeric	Ordinal	Float	64						NA	12895 missing values
host_acceptance_rate	listing_summery										NA	All values are missing
host_is_superhost	listing_summery	Categorical	Nominal	Boolean	2						NA	26 missing values
host_thumbnail_url	listing_summery	Categorical	Nominal	Text	19100						NA	26 missing values
host_picture_url	listing_summery	Categorical	Nominal	Text	19100						NA	26 missing values

host_neighbourhood	listing_summery	Categorical	Nominal	Varchar	181						NA	5094 missing values
host_listings_count	listing_summery	Numeric	Interval	Integer	22526			0	1676		NA	26 missing values
host_total_listings_count	listing_summery	Numeric	Interval	Integer	22526			0	1676		NA	26 missing values
host_verifications	listing_summery	Categorical	Nominal	Text	301							will have to split it
host_has_profile_pic	listing_summery	Categorical	Nominal	Boolean	2						NA	26 missing values
host_identity_verified	listing_summery	Categorical	Nominal	Boolean	2						NA	26 missing values
street	listing_summery	Categorical	Nominal	Text	86							
neighbourhood	listing_summery	Categorical	Nominal	Varchar	91						NA	1131 missing values
neighbourhood_cleansed	listing_summery	Categorical	Nominal	Varchar	136							
neighbourhood_group_cleansed	listing_summery	Categorical	Nominal	Varchar	12							
city	listing_summery	Categorical	Nominal	Varchar	60						NA	5 missing values
state	listing_summery	Categorical	Nominal	Varchar	19						NA	84 missing values
zipcode	listing_summery	Categorical	Nominal	Varchar	215						NA	656 missing values
market	listing_summery	Categorical	Nominal	Varchar	6						NA	63 missing values
smart_location	listing_summery	Categorical	Nominal	Text	61							
country_code	listing_summery	Categorical	Nominal	Varchar	1							
country	listing_summery	Categorical	Nominal	Varchar	1							
latitude	listing_summery	Categorical	Nominal	Float	22526			52.345803	52.65167			
longitude	listing_summery	Categorical	Nominal	Float	22526			13.103557	13.757642			
is_location_exact	listing_summery	Categorical	Nominal	Boolean	2							
property_type	listing_summery	Categorical	Nominal	Varchar	33							
room_type	listing_summery	Categorical	Nominal	Varchar	3							
accommodates	listing_summery	Numeric	Interval	Integer	22526			1	16			
bathrooms	listing_summery	Numeric	Interval	Float	22526			0	8.5		NA	32 missing values
bedrooms	listing_summery	Numeric	Interval	Integer	22526			0	12		NA	18 missing values
beds	listing_summery	Numeric	Interval	Integer	22526			0	22		NA	40 missing values
bed_type	listing_summery	Categorical	Nominal	Varchar	5							
amenities	listing_summery	Categorical	Nominal	Text	19985							
square_feet	listing_summery	Numeric	Interval	Float	22526			0	4639		NA	22106 missing values
price	listing_summery	Numeric	Interval	Integer	295	dolar		0	9000			
weekly_price	listing_summery	Numeric	Interval	Integer	387	dolar		65	3650		NA	18871 missing values
monthly_price	listing_summery	Numeric	Interval	Integer	426	dolar		265	13900		NA	19893 missing values
security_deposit	listing_summery	Numeric	Interval	Integer	142	dolar		0	4280		NA	9361 missing values
cleaning_fee	listing_summery	Numeric	Interval	Integer	112	dolar		0	2000		NA	7146 missing values
guests_included	listing_summery	Numeric	Interval	Integer	22526			1	16			
extra_people	listing_summery	Numeric	Interval	Integer	65	dolar		0	250			
minimum_nights	listing_summery	Numeric	Interval	Integer	22526			1	5000			
maximum_nights	listing_summery	Numeric	Interval	Integer	22526			1	1.10E+10			
calendar_updated	listing_summery	Categorical	Nominal	Varchar	75							
has_availability	listing_summery	Categorical	Nominal	Boolean	1							All values are true
availability_30	listing_summery	Numeric	Interval	Integer	31			0	30			
availability_60	listing_summery	Numeric	Interval	Integer	61			0	60			
availability_90	listing_summery	Numeric	Interval	Integer	91			0	90			
availability_365	listing_summery	Numeric	Interval	Integer	366			0	365			
calendar_last_scraped	listing_summery	Date	Time-series	DateTime	2							
number_of_reviews	listing_summery	Numeric	Interval	Integer	22552			0	498			
first_review	listing_summery	Date	Time-series	Date	2228						NA	3914 missing values
last_review	listing_summery	Date	Time-series	Date	1312						NA	3908 missing values
review_scores_rating	listing_summery	Numeric	Interval	Integer	46			20	100		NA	4389 missing values
review_scores_accuracy	listing_summery	Categorical	Ordinal	Integer	9			2	10		NA	4414 missing values
review_scores_cleanliness	listing_summery	Categorical	Ordinal	Integer	9			2	10		NA	4411 missing values
review_scores_checkin	listing_summery	Categorical	Ordinal	Integer	8			2	10		NA	4432 missing values

review_scores_communication	listing_summery	Categorical	Ordinal	Integer	9			2	10		NA	4418 missing values
review_scores_location	listing_summery	Categorical	Ordinal	Integer	8			2	10		NA	4431 missing values
review_scores_value	listing_summery	Categorical	Ordinal	Integer	9			2	10		NA	4435 missing values
requires_license	listing_summery	Categorical	Nominal	Boolean	2							
license	listing_summery	Categorical	Nominal	Varchar	1387						NA	20914 missing values
jurisdiction_names	listing_summery											All values are missing
instant_bookable	listing_summery	Categorical	Nominal	Boolean	2							
is_business_travel_ready	listing_summery	Categorical	Nominal	Boolean	1							All values are false
cancellation_policy	listing_summery	Categorical	Nominal	Varchar	5							
require_guest_profile_picture	listing_summery	Categorical	Nominal	Boolean	2							
require_guest_phone_verification	listing_summery	Categorical	Nominal	Boolean	2							
calculated_host_listings_count	listing_summery	Numeric	Interval	Integer	23			1	45			
reviews_per_month	listing_summery	Numeric	Interval	Integer	768			0	36.67		NA	3914 missing values
neighbourhood_group	neighbourhoods	Categorical	Nominal	Varchar	12						NA	1 missing value
neighbourhood	neighbourhoods	Categorical	Nominal	Varchar	138						NA	1 missing value
listing_id	reviews	Categorical	Nominal	Integer	18638			2015	29867350			
date	reviews	Date	Time-series	Date	2833							
listing_id	reviews summary	Categorical	Nominal	Integer	18638			2015	29867350			
id	reviews summary	Categorical	Nominal	Integer	401963			4.28E+03	3.46E+08			
date	reviews summary	Date	Time-series	Date	2833							
reviewer_id	reviews summary	Categorical	Nominal	Integer	370096			5	2.24E+08			
reviewer_name	reviews summary	Categorical	Nominal	Varchar	52345							
comments	reviews summary	Categorical	Nominal	Text	388778						NA	496 missing values