Daduosu原始数据描述

By 李昕达

edit #1 2021-05-01：描述CN、US 各自detail、order、price

CN

包含 detail   order   price   三张数据表

表描述

表名称	表描述
detail	房源描述
order	房源订购记录
price	房源标价

detail表结构

表 detail

```
CREATE TABLE `detail` (
    `id` bigint(20) UNSIGNED NOT NULL AUTO_INCREMENT,
    `listingId` varchar(80) CHARACTER SET utf8 COLLATE utf8_bin NULL,
    `title` varchar(511) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `Lat` decimal(16,10) NULL,
    `Lng` decimal(16,10) NULL,
    `chinaTitleDetails` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin
NULL,
    `kickers` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `ogImage` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `previewTags` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `propertyType` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `reviewCount` int(10) UNSIGNED NULL,
    `reviewScore` int(10) UNSIGNED NULL,
    `reviewTagSummary` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin
NULL,
    `hostAbout` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `hostBadges` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `hostId` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `hostIntroTags` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `hostName` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `amenity_jian_shen_fang` tinyint(1) NULL DEFAULT 0 COMMENT '健身房',
    `amenity_you_xian_dian_shi` tinyint(1) NULL DEFAULT 0 COMMENT '有线电视',
    `amenity_dian_shi` tinyint(1) NULL DEFAULT 0 COMMENT '电视',
               ##########  共计36个amenity    ###########
    `previewTags_you_zhi_fang_yuan` tinyint(1) NULL DEFAULT 0 COMMENT '优质房源',
    `previewTags_ke_cun_xing_li` tinyint(1) NULL DEFAULT 0 COMMENT '可存行李',
    `previewTags_ri_ben_te_hui` tinyint(1) NULL DEFAULT 0 COMMENT '日本特惠',
```

```
########## 共计15个previewTags ###########
    `chinaTitleDetails_BED` varchar(20) CHARACTER SET utf8 NULL COMMENT '_BED',
    `chinaTitleDetails_GUESTS` varchar(20) CHARACTER SET utf8 NULL COMMENT
'_GUESTS',
    `chinaTitleDetails_BATH` varchar(20) CHARACTER SET utf8 NULL COMMENT
'_BATH',
    `chinaTitleDetails_ROOM` varchar(20) CHARACTER SET utf8 NULL COMMENT
'_ROOM',
    `reviewSummary_wei_zhi_bian_li` decimal(5,2) NULL COMMENT '位置便利',
    `reviewSummary_ru_zhu_bian_jie` decimal(5,2) NULL COMMENT '入住便捷',
    `reviewSummary_ru_shi_miao_shu` decimal(5,2) NULL COMMENT '如实描述',
    `reviewSummary_gan_jing_wei_sheng` decimal(5,2) NULL COMMENT '干净卫生',
    `reviewSummary_gou_tong_shun_chang` decimal(5,2) NULL COMMENT '沟通顺畅',
    `reviewSummary_gao_xing_jia_bi` decimal(5,2) NULL COMMENT '高性价比'
)
```

字段名	字段样例	描述
listingId	44600698	房源ID
title	【青竹居.6室】植物园旁清雅青竹小院包场 地址：南山街道 公园路168号查看地图	房源名称
Lat&Lng	10.3296100000&123.8844500000	经纬度
kickers	宿务城·公寓型住宅里的独立房间	房源位置与房型的综合描述
ogImage	https://z1.muscache.cn/im/pictures/1814643/2f8cbe52_original.jpg?aki_policy=x_large	房源主图链接
propertyType	整套公寓	房源类型
reviewCount	47	评价数量
reviewScore	83	评价分数（满分100）
hostAbout	I am 27 years old & workin........	房东描述
hostId	644122	房东Id
hostIntroTags	["125 条评价", "已验证身份", "超赞房东"]	房东的标签
hostName	Arthur	房东名字
健身房等36种amenity	1	amenity
优质房源等15种 previewTags	1	房源标签
chinaTitleDetails_BED& _GUESTS&_BATH&_ROOM	4张床	床、可住人数、卫生间、卧室描述
位置便捷等6种 reviewSummary	4.7	6种评价评分项（满分5）

detail表内容

表内容样例见附件 **detail.xlsx**

order表结构

```
CREATE TABLE `order` (
    `id` bigint(11) UNSIGNED NOT NULL AUTO_INCREMENT,
    `house_id` varchar(30) CHARACTER SET utf8 COLLATE utf8_general_ci NULL,
    `fetch_date` date NULL,
    `order_date` date NULL,
    `repeat_flag` varchar(60) CHARACTER SET utf8 COLLATE utf8_general_ci NULL,
)
```

字段名	字段样例	描述
house_id	44600698	房源ID
fetch_date	2021-04-28	订购房源的日期
order_date	2021-05-27	房源被订购的日期

指房源44600698在2021-04-28被订购要在2021-05-27入住

order表内容

表内容样例见附件 **order.xlsx**

price表结构

```
CREATE TABLE `price` (
    `id` bigint(11) UNSIGNED NOT NULL AUTO_INCREMENT,
    `house_id` varchar(30) CHARACTER SET utf8 COLLATE utf8_general_ci NULL,
    `fetch_date` date NULL,
    `order_date` date NULL,
    `price` int(11) NULL,
    `repeat_flag` varchar(60) CHARACTER SET utf8 COLLATE utf8_general_ci NULL,
)
```

字段名	字段样例	描述
house_id	978531	房源ID
fetch_date	2021-04-28	订购房源的日期
order_date	2021-06-27	房源被订购的日期
price	1033	价格

指房源978531在2021-04-28被房东修改2021-06-27当日的单日价格为￥1033

price表内容

表内容样例见附件 **price.xlsx**

US

与CN相比，US在detail_us表上有部分字段差异，在price_us和order_us两张表上无任何不同

美国（及周边地区）房源共796851套

表描述

表名称	表描述
detail_us	房源描述
order_us	房源订购记录
price_us	房源标价

detail_us表结构

表 detail_us

```
CREATE TABLE `detail_us` (
    `id` bigint(20) UNSIGNED NOT NULL AUTO_INCREMENT,
    `listingId` varchar(80) CHARACTER SET utf8 COLLATE utf8_bin NULL,
    `title` varchar(511) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `description` text CHARACTER SET utf8 COLLATE utf8_bin NULL,
    `Lat` decimal(16,10)  NULL,
    `Lng` decimal(16,10)  NULL,
    `titleDetails` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `reviewCount` int(10) UNSIGNED NULL,
    `reviewScore` decimal(10,2)  NULL,
    `ogImage` varchar(255) CHARACTER SET utf8 COLLATE utf8_bin NULL,
    `hostAbout` text CHARACTER SET utf8 COLLATE utf8_bin NULL,
    `hostBadges` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `hostId` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `hostIntroTags` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `hostName` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NULL,
    `titleDetails_BED` varchar(20)  NULL COMMENT '_BED',
    `titleDetails_GUESTS` varchar(20)  COMMENT '_GUESTS',
    `titleDetails_BATH` varchar(20)  COMMENT '_BATH',
    `titleDetails_ROOM` varchar(20)  COMMENT '_ROOM',
    `reviewSummary_Accuracy` decimal(4,2)  NULL,
    `reviewSummary_Check-in` decimal(4,2)  NULL,
    `reviewSummary_Cleanliness` decimal(4,2)  NULL,
    `reviewSummary_Communication` decimal(4,2)  NULL,
    `reviewSummary_Location` decimal(4,2)  NULL,
    `reviewSummary_Value` decimal(4,2)  NULL,
    `amenity` varchar(1024) CHARACTER SET utf8 COLLATE utf8_bin NULL,
) ENGINE=InnoDB
```

字段名	字段样例	描述
listingId	4098003	房源ID
title	Ocean Front Condo Minutes from Coronado	房源名称
description	Beautiful oceanfront living in this gorgeous 2 bedroom, 1 ba...	房源描述
Lat&Lng	32.5684100000&-117.1321700000	经纬度
reviewCount	1	评价数量
reviewScore	5.00	评价分数（满分5）
ogImage	https://z1.muscache.cn/im/pictures/51512097/fe3c21d2_original.jpg?aki_policy=x_large	房源主图
hostAbout	We are beach lovers and have always been so.	房东描述
hostBadges	SUPER_HOST	SUPER_HOST
hostId	21255757	房东Id
hostIntroTags	["9 Reviews", "Identity verified"]	房东的标签
hostName	Sabrina	房东名字
titleDetails_BED&_GUESTS&_BATH&_ROOM	2 beds	床、可住人数、卫生间、卧室描述
Accuary等6种 reviewSummary	4.90	6种评价评分项（满分5）

detail_us表内容

表内容样例见附件 **detail_us.xlsx**

order_us & price_us

order_us / price_us 与 order / price 结构完全一致

表内容样例见order_us.xlsx 与 price_us.xlsx