

广州市旅游局关于2018年广州民宿名录信息服务项目

根据工作安排，2018年，我局将开展广州民宿基础数据普查工作，切实掌握我市民宿的基础设施情况。我处拟通过购买服务，委托一家专业机构，运用网络爬虫技术，获取我市民宿基本信息。现邀请你单位根据服务内容，就2018年广州民宿名录信息服务项目进行报价。

一、服务内容

1、按要求对数据采集进行布局与规划设计（包括编码对标、数据爬取规则与配置等）。

2、运用网络爬虫技术对以下热门民宿预订网站进行数据爬取：爱彼迎(airbnb)、途家网、蚂蚁短租、木鸟短租、小猪短租。

3、对民宿进行地图定位，标注民宿准确位置信息（地址信息），匹配民宿名称及区域统计。民宿数据信息包含但不限于以下内容：民宿名称、区域分布（具体到镇、街道）、房间数、床位数、可住人数、类型（公寓、乡村民宿、别墅、独栋房等）、注册时间信息、住宿单价等。

4、对全部民宿数据进行查重去重，统一格式,清洗数据，保留有效数据。

5、按统计区划代码和城乡划分代码归类，提取归集民宿信息，制成Excel表格。