

## 酒店价格分布模型

### 一、模型概述

本模型旨在对 X 市酒店的价格分布情况进行系统分析，通过将价格划分为不同区间，即 100（含）~300、300（含）~500、500（含）~700、1000（含）以上，直观呈现各区间内酒店数量或占比情况，为游客选择合适价位的酒店、酒店经营者掌握市场价格定位以及旅游管理部门了解行业价格态势提供清晰的数据参考，助力优化 X 市酒店市场的价格体系与资源配置。

### 二、数据来源与整理

**在线旅游平台数据：**像携程、去哪儿、飞猪等主流在线旅游平台汇聚了大量 X 市酒店的价格信息，涵盖不同房型、不同时间段的标价情况，这些数据实时更新且样本量大，能全面反映各酒店在市场中的定价策略，是分析酒店价格分布的关键数据源。通过平台提供的数据接口或者网络爬虫技术（遵循平台规则及法律法规）收集相关数据，可获取各酒店各房型的基础价格信息。

**酒店预订系统数据：**各酒店自身的预订系统中记录了详细的房间价格设定，包括针对不同会员等级、不同预订渠道、不同季节等因素下的差异化价格情况，这些数据能更精准地体现酒店实际运营中的价格策略，与在线旅游平台数据相互印证、补充，确保价格数据的准确性和完整性，为后续分析提供可靠依据。

**市场调研数据：**由旅游管理部门、酒店行业协会或专业调研机构定期开展的针对 X 市酒店价格的实地调研，通过电话咨询、实地走访等方式收集部分酒店的价格信息，尤其是一些未完全接入在线旅游平台或者价格变动较为特殊的酒店情况，进一步完善价格数据覆盖范围，使分析结果更贴合 X 市酒店市场的实际情况。

**数据整理流程：**首先，将从上述多渠道收集到的数据统一汇总到大数据存储平台。接着，运用数据清洗工具和人工核查相结合的方法，去除重复、错误以及逻辑不清晰的数据记录，比如剔除在线旅游平台因系统故障出现的异常价格数据、纠正酒店预订系统中录入错误的价格信息等。然后，按照酒店名称对数据进行分类，

针对每一家酒店，提取各房型的标准价格（取一定时间段内的平均价格或者常规标价等具有代表性的价格数值），再依据设定的价格区间（100（含）~300、300（含）~500、500（含）~700、1000（含）以上）对价格数据进行分类整理，统计各区间内包含的酒店数量或各酒店房型在各区间的占比情况，为后续的可视化展示做好准备。

### 三、核心算法与分析逻辑

**价格区间归类统计：**借助数据库管理系统（如 MySQL、Oracle 等）或数据分析软件（如 Python 的数据分析库 pandas 等），依据整理好的酒店价格数据以及设定的价格区间范围，运用条件判断语句（如 SQL 语句中的 CASE WHEN 语句，或 pandas 中的 where 方法等），对每个酒店的各房型价格进行判断，将其归入对应的价格区间类别中，然后运用计数函数（如 SQL 语句中的 COUNT 函数，或 pandas 中的 size 方法）统计各价格区间内的酒店数量或者各酒店房型在各区间的占比情况，形成价格分布的数据集。

**数据汇总与整合（可选，用于多维度分析）：**除了单纯统计各价格区间的酒店数量或占比，还可进一步按照酒店星级、地理位置（如市区、景区周边等）、酒店类型（商务型、度假型等）等维度进行分类汇总统计，分析不同特征酒店在各价格区间的分布差异，更深入地挖掘 X 市酒店价格分布的内在规律，为不同需求的分析提供更细致的数据支持。

### 四、模型输出与可视化

**柱状图展示：**构建一个可视化的网页界面或电子看板，生成以价格区间为横坐标，酒店数量（或占比情况，根据具体展示需求选择）为纵坐标的柱状图。不同颜色的柱子代表不同的价格区间，柱子的高低直观反映该价格区间内酒店的数量多少（或占比大小），通过柱状图可以清晰对比各价格区间酒店的分布差异，便于快速查看 X 市酒店价格的整体分布态势。

**饼图展示（可选，用于体现占比关系）：**同时，也可以制作饼图来展示各价格区

间酒店数量（或房型数量等）在总体中的占比情况，将整个圆按照各价格区间的占比划分不同扇形区域，每个扇形区域标注对应的价格区间及占比数值，一目了然地呈现各价格区间在整体中所占的比重关系，辅助分析不同价位酒店在市场中的份额结构。

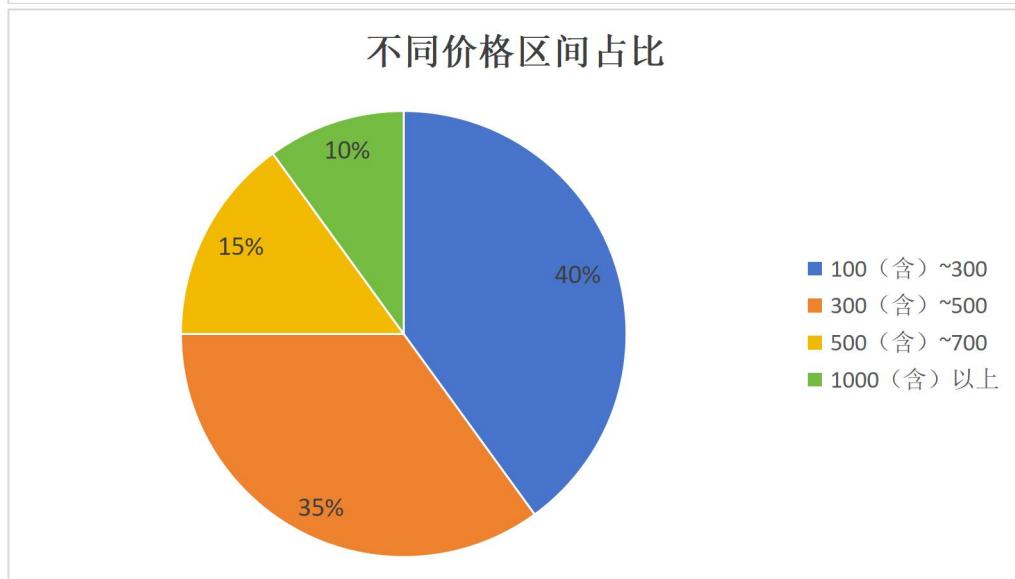
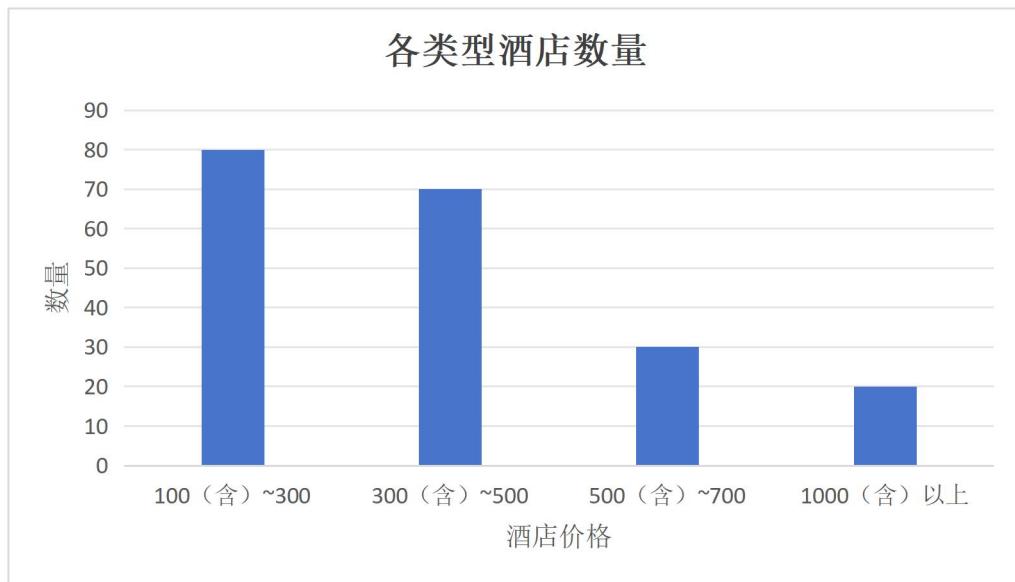
**交互功能实现：**借助前端可视化开发技术（如 JavaScript 结合 Echarts 可视化库等），为上述可视化图表添加交互功能。当用户点击某一价格区间对应的柱状图柱子或饼图扇形区域时，在界面上弹出信息框，展示该价格区间内详细的酒店列表（包含酒店名称、星级、大致地理位置等关键信息），方便用户进一步了解处于该价格区间的具体酒店情况，满足不同用户深入查看详情的需求。

## 五、模拟数据示例与可视化展示

假设 X 市有部分酒店经过数据整理和统计后，得到以下模拟的价格分布相关数据（以酒店数量为例）：

价格区间（元）	酒店数量（家）	占比
100（含）~300	80	40%
300（含）~500	70	35%
500（含）~700	30	15%
1000（含）以上	20	10%

对应的可视化展示如下：



(此处应插入一个简单的柱状图示例，横坐标为上述 4 个价格区间，纵坐标为酒店数量，柱子高低体现各区间酒店数量差异，鼠标点击柱子有信息框弹出展示该区间详细酒店列表示例图；若有饼图展示，也插入一个饼图示例，按照各区间占比划分扇形，标注区间及占比数值，鼠标点击有交互效果示例图)

## 六、应用场景与价值

**游客出行决策辅助：**游客在计划前往 X 市旅游并预订酒店时，通过查看该模型展示的酒店价格分布情况，能快速了解 X 市酒店价格的大致范围以及各价位区间内的酒店数量情况，结合自身预算和对住宿品质的期望，更精准地筛选出符合需求的酒店，提高预订效率，避免因价格信息不透明而盲目选择酒店。

**酒店定价策略参考:** 酒店经营者可以依据模型呈现的同区域、同类型酒店的价格分布情况，分析自身酒店在市场价格体系中的位置，判断定价是否合理。若发现所处价格区间竞争过于激烈，可考虑调整价格策略、优化产品服务或推出差异化的套餐等，提升酒店在市场中的竞争力，吸引更多目标客户群体。

**旅游市场监管与规划:** X 市旅游管理部门借助该模型，全面掌握全市酒店行业的价格结构与分布动态，对于价格虚高、低价恶性竞争等不合理现象及时进行监管和调控，同时，依据长期的价格数据变化趋势，合理规划酒店建设布局、引导投资方向，促进 X 市酒店市场的健康、有序、均衡发展。