

Problem Y: Understanding Used Sailboat Prices

像许多奢侈品一样，帆船的价值随着它们的老化和市场条件的变化而变化。所附的 "2023_MCM_Problem_Y_Boats.xlsx" 文件包括2020年12月在欧洲、加勒比海和美国销售的约3500艘36至56英尺长的帆船广告数据。一位船艇爱好者向COMAP提供了这些数据。像大多数现实世界的数据集一样，它可能有数据缺失或其他问题，需要在分析前进行一些数据清理。

Excel文件包括两个标签，一个是单体帆船，一个是双体帆船。在每个选项卡中，各栏都标有品牌、型号、长度（英尺）、地理区域、国家/地区/州、上市价格（美元）和年份（制造）。

对于一个给定的品牌、型号和年份，除了所提供的Excel文件外，还有许多其他来源可以提供关于特定帆船特征的详细描述。你可以用你选择的任何其他数据来补充所提供的数据集；但是，你必须在你的建模中包括 "2023_MCM_Problem_Y_Boats.xlsx" 中的数据。请确保充分识别和记录所使用的任何补充数据的来源。

帆船经常通过经纪人出售。为了更好地了解帆船市场，中国香港（特区）的一位帆船经纪人委托贵公司团队编写一份关于二手帆船定价的报告。

该经纪人希望你能

- 建立一个数学模型，解释所提供的电子表格中每艘帆船的上市价格。包括任何你认为有用的预测因素。你可以利用其他资料来了解特定帆船的其他特征（如船宽、吃水、排水量、索具、帆面积、船体材料、发动机小时数、睡眠容量、净空、电子设备等）以及按年份和地区划分的经济数据。识别和描述所有使用的数据来源。包括讨论你对每一种帆船的价格估计的精确度。
- 用你的模型来解释区域对上市价格的影响（如果有的话）。讨论任何区域性的影响是否在所有帆船变体中都是一致的。讨论任何区域效应的实际意义和统计意义。
- 讨论你对给定的地理区域的建模如何在香港（特区）市场上发挥作用。从提供的电子表格中选择一个信息量大的帆船子集，分为单体船和双体船。从香港（特区）市场找到该子集的可比上市价格数据。如果有的话，模拟香港（特区）对你子集中的每艘帆船价格的区域影响是什么。对双体船和单体船的影响是否相同？
- 识别并讨论你的团队从数据中得出的任何其他有趣的、有信息量的推论或结论。
- 为香港（特区）的帆船经纪人准备一份一到两页的报告。包括一些精心选择的图形，以帮助经纪人理解你的结论。

注意：MCM比赛有25页的限制。你提交的所有内容都计入25页的限制（摘要，目录，报告，给经纪人的一到两页的报告，参考文献列表，和任何附录）。你必须为你的想法、数据、图像和报告中使用的任何其他材料注明来源。

词汇表

船宽：船的最宽处的宽度。

经纪人：为卖家和/或买家服务的专家，作为出售或购买财产的代表。在这个问题中，财产是一艘帆船。

双体船：一种多船体的水上交通工具，具有两个大小相同的平行船体。

排水量：一艘船所排出的水的体积的重量。

吃水：船只在不触底的情况下漂浮所需的最小水深。

发动机小时数：船只的发动机从新开始运行的小时数。

净空：在船舱内可站立的高度。

船体：一艘船或其他船只的主体或外壳，包括底部、侧面和甲板。

船体材料：船体由哪些材料制成。使用的材料包括玻璃纤维、钢、木材和复合材料。

挂牌价：卖方要求的价格。船只可能以不同的价格出售。

制造：帆船的制造商。

单体帆船：只有一个船体的帆船，通常以一个沉重的龙骨（中心叶片）为中心。

索具：由绳索、电缆和滑轮组成的系统，用于支持和控制船的风帆、转向和其他系统。

帆面积：船只完全升起时的帆的总表面积。

型号：识别某一特定型号帆船的名称。例如，"Sun Odyssey 54 DS"。