

## 这个冬天去哪玩?

——出境旅游线路分析



目录

CONTENTS

01/ 相关背景

02 数据介绍&处理

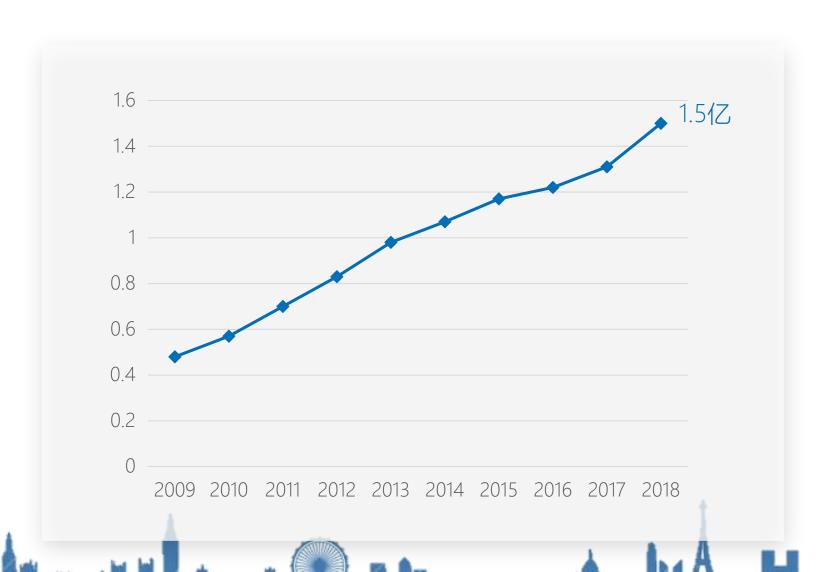
03/ 描述分析

04 模型构建&分析

05 结论&展望



## ☆ 中国公民出境游人次



#### 相关背景





















## 相关背景



旅行社竞争激烈

如何让自己的产品脱颖而出?



#### 相关背景







总有一些关于出国旅行的提问



其中有一位提问者是学霸

今年刚拿了很多很多奖学金

寒假想带着家人出国旅行

有什么线路推荐呢?

霸霸请稍等

渣渣这就去调查清楚!







数据爬取



**2000**+ 条

日本 欧洲 美洲

冬季出行

上海出发

变量名统一

缺失值处理

哑变量处理

## ☆ 变量表

变量类型		变量名	详细说明
因变量		线路价格	单位:元
自变量		产品标签	文本型变量
	产品性质	旅行社	0 1变量,是否为"牛人专线"
		满意度	百分比
		累积旅行人数	单位:人
	· 行程安排 ·	日均游览经典景点数	单位: 个/天
		日均游览景点数	单位: 个/天
		日均餐饮数	单位:次/天
		出行天数	单位:天
		游玩地区	哑变量处理
		节假日状况	0 1变量,是否为节假日线路

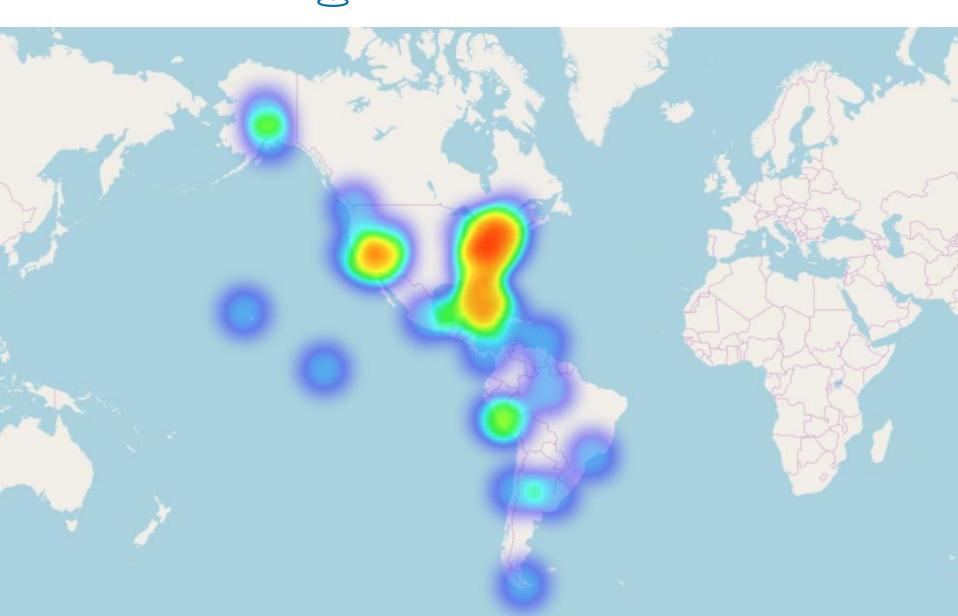


描述分析



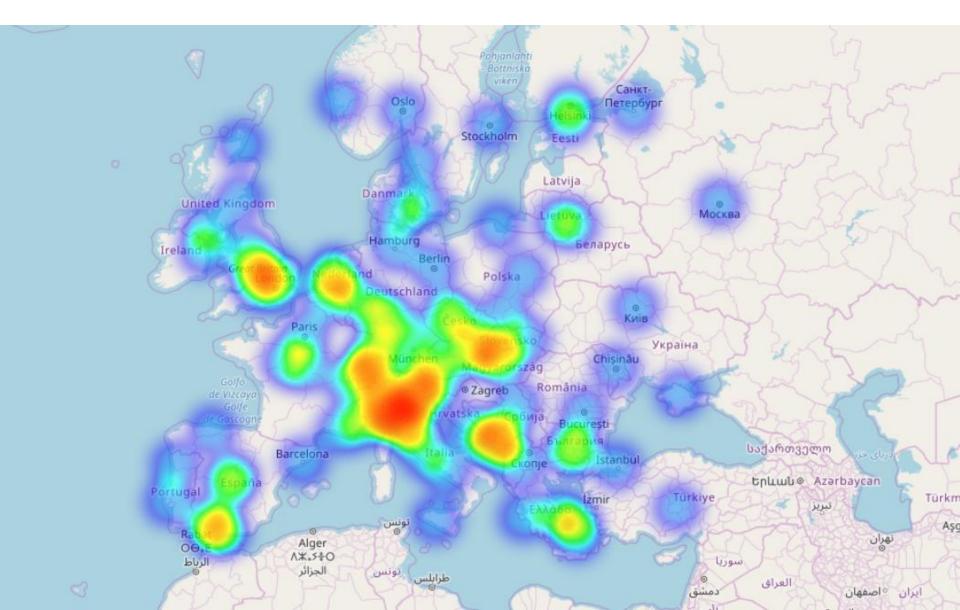


## 美洲 热门游玩地区分布



## 0

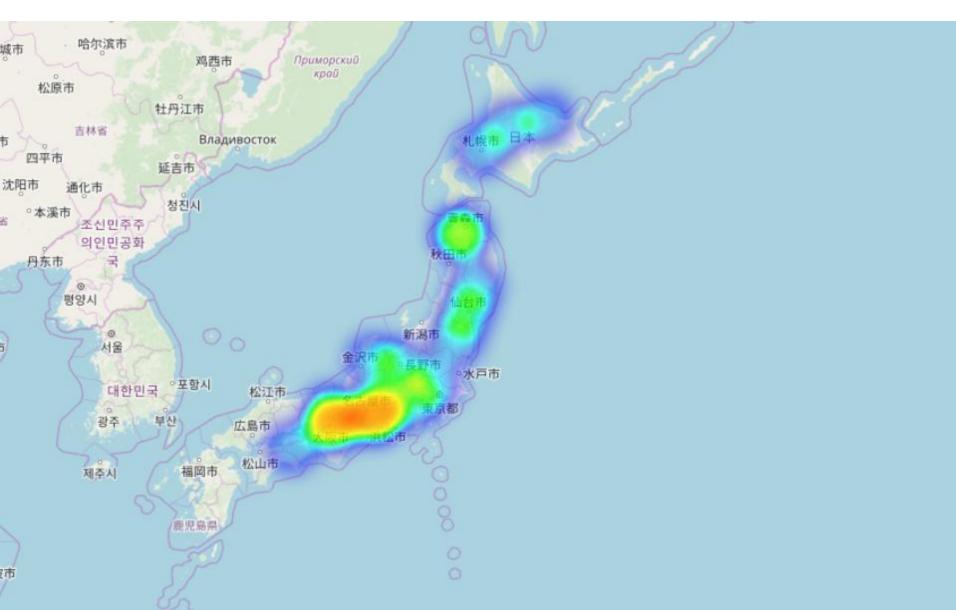
## 欧洲 热门游玩地区分布



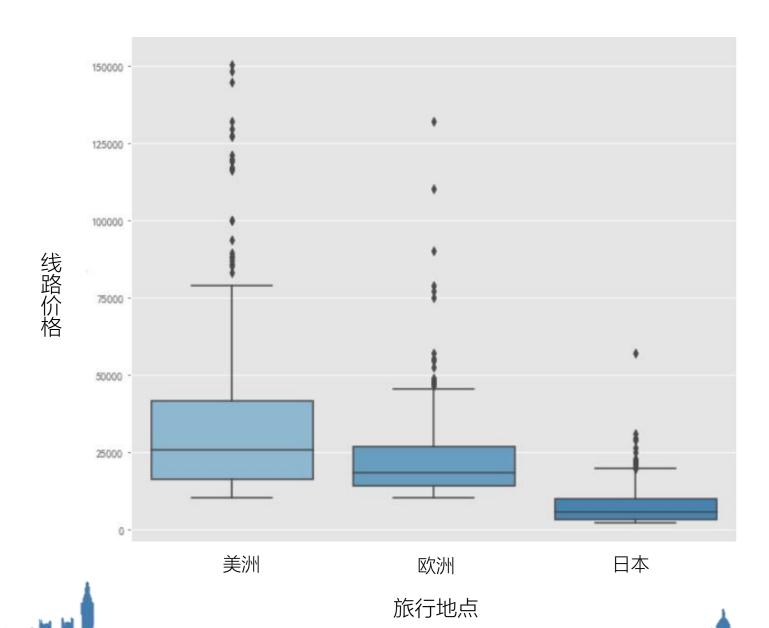
### 描述分析



## 日本 热门游玩地区分布



## 描述分析



那就决定去美洲吧

反正奖学金太多了

总是花不完







## 线路特色标签

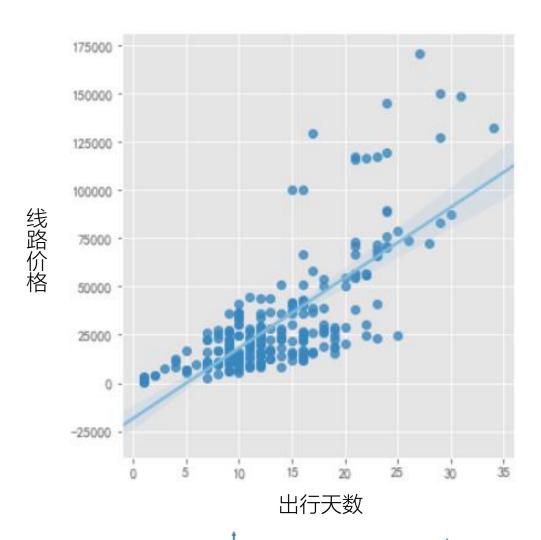




## ■ 线路参观景点

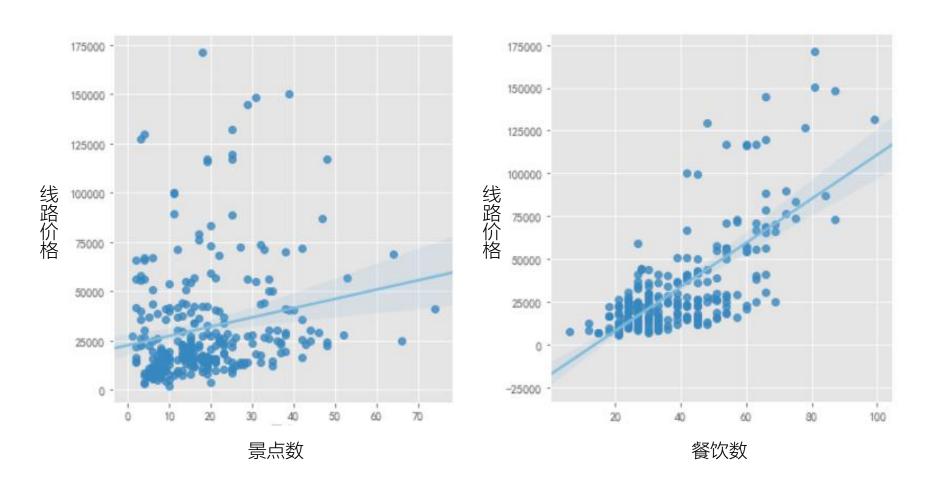


## ► 出行天数对价格的影响

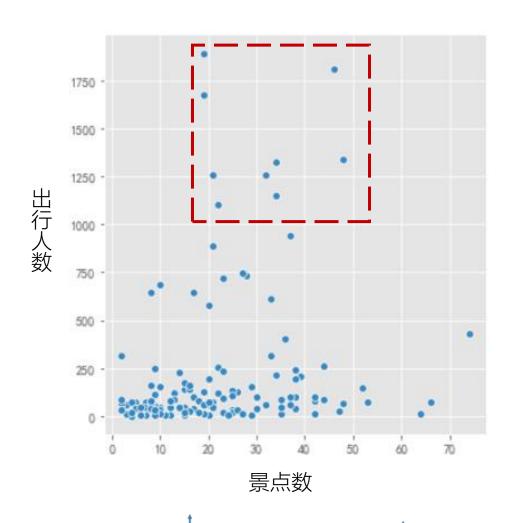




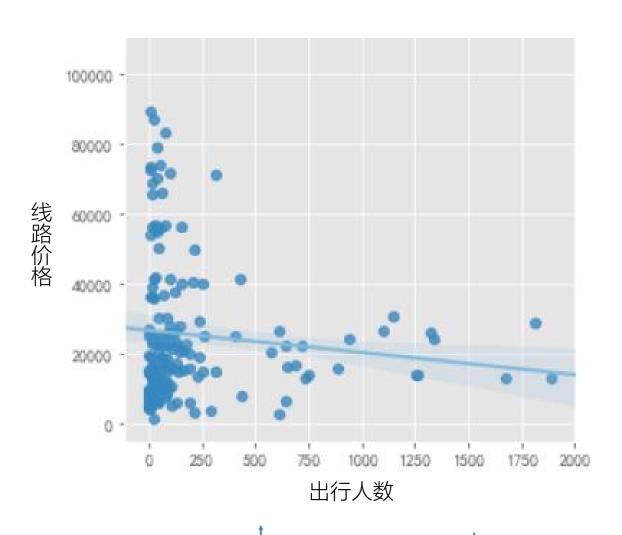
## 景点与餐饮数对价格的影响



## ■ 景点数与出行人数的关系

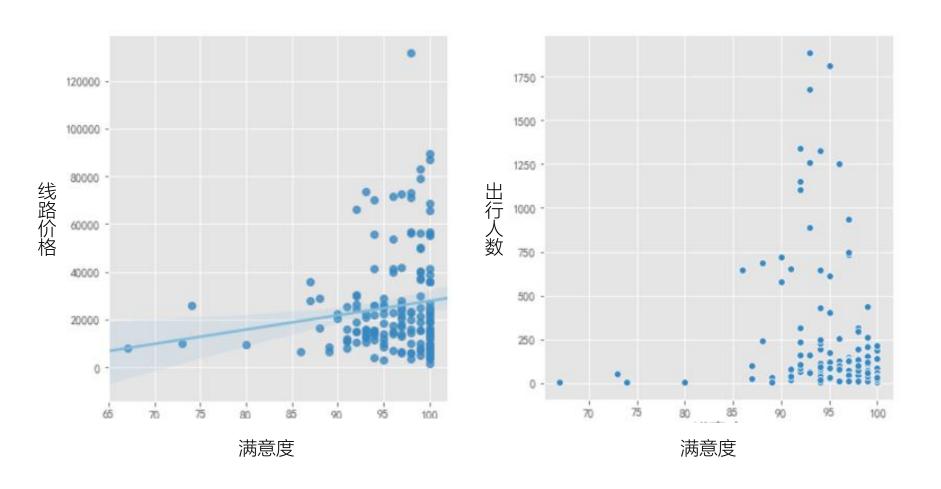


## 出行人数与线路价格关系





## **が** 满意度与价格/出行人数关系





✔ 线性回归模型

△ 回归树模型

以美洲为例

#### ✔ 线性回归模型



价格的取值范围较大,分布严重右偏



筛除变量



让系数有说服力



## 模型构建&分析

## → 线性回归模型

变量		回归系数	p值	显著性	基准
截距项		10.42	<2e-16	***	-
是否节假日		0.17	<0.01	**	-
日均餐饮数		-0.05	0.03	*	-
出行天数		0.48	<2e-16	***	-
	凯恩县 (美国)	-1.15	<0.001	***	
	圣莫尼卡 (美国)	-1.17	<0.001	***	
游玩地区	马丘比丘 (秘鲁)	0.52	0.03	*	布伦特伍德湾 (加拿大)
	利马 (秘鲁)	0.48	0.05	·	
	• • •		• • •		

Adjusted  $R^2=0.80$ 



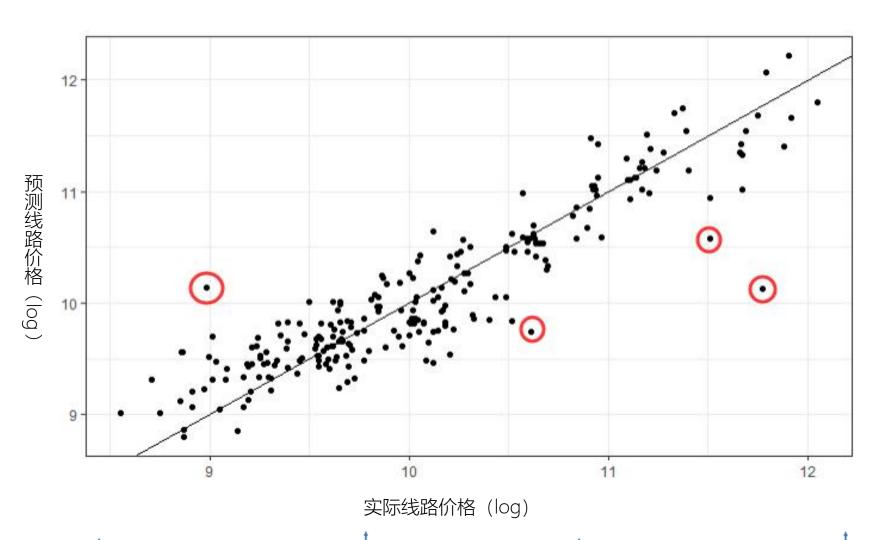
## 

变量		回归系数	p值	显著性	基准
截距项		10.42	<2e-16	***	-
是否节假日		0.17	<0.01	**	-
日均餐饮数		-0.05	0.03	*	-
出行天数		0.48	<2e-16	***	-
	凯恩县 (美国)	-1.15	<0.001	***	布伦特伍德湾 (加拿大)
	圣莫尼卡 (美国)	-1.17	<0.001	***	
游玩地区	马丘比丘 (秘鲁)	0.52	0.03	*	
	利马 (秘鲁)	0.48	0.05	·	
	• • •		• • •		

Adjusted  $R^2 = 0.80$ 



## ★ 线性回归模型预测情况



#### ★◆ 线性回归模型离群点分析



实际线路价格

#### **-----** 线性回归模型离群点分析

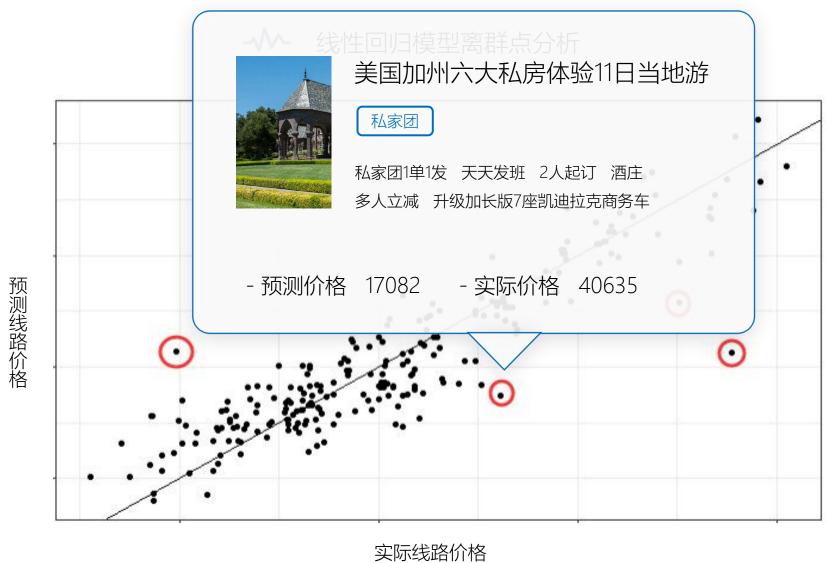


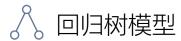
实际线路价格

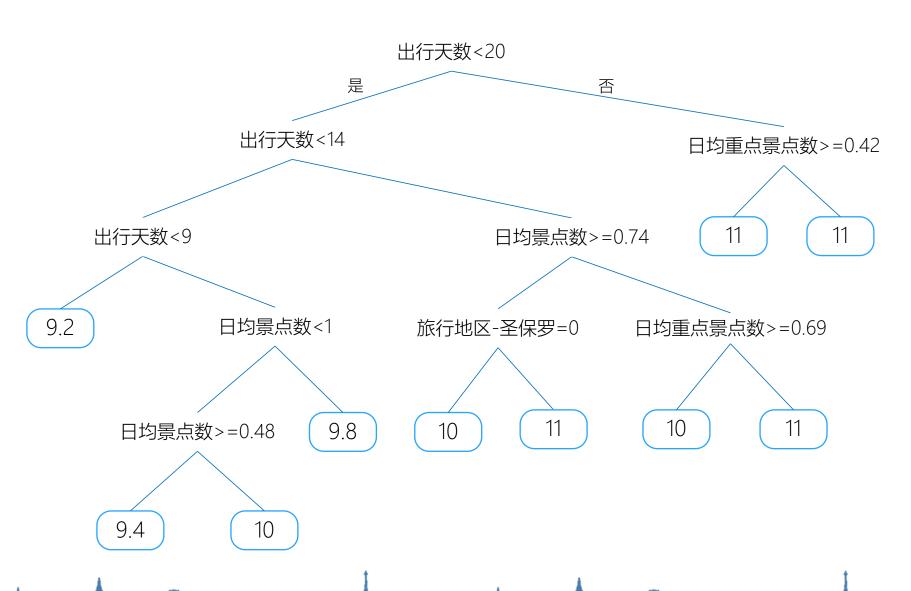
#### **-----** 线性回归模型离群点分析

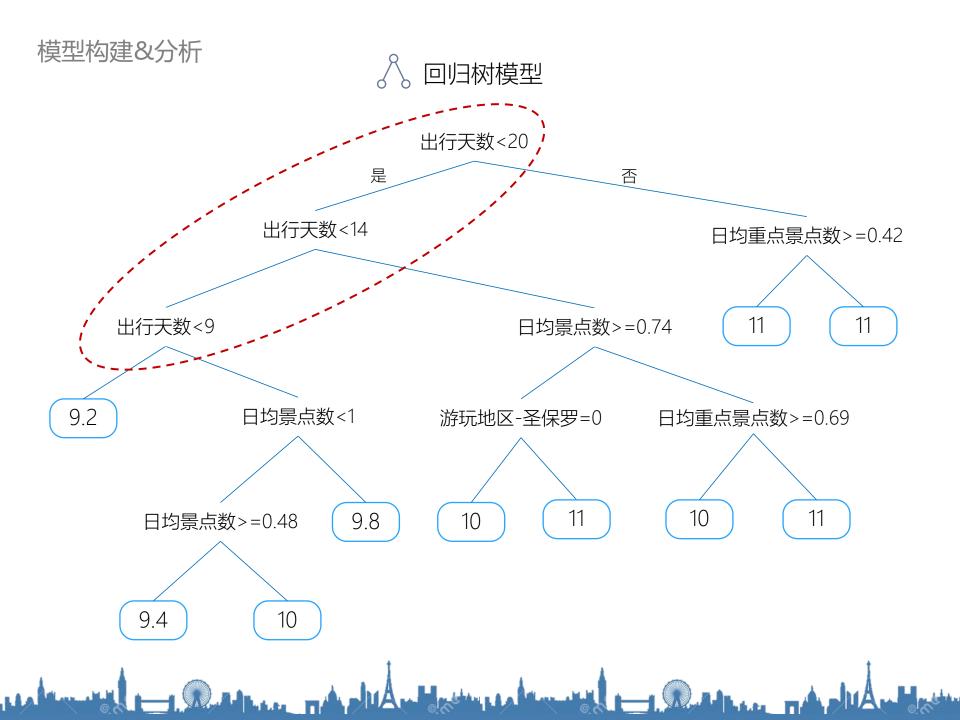


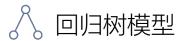
实际线路价格

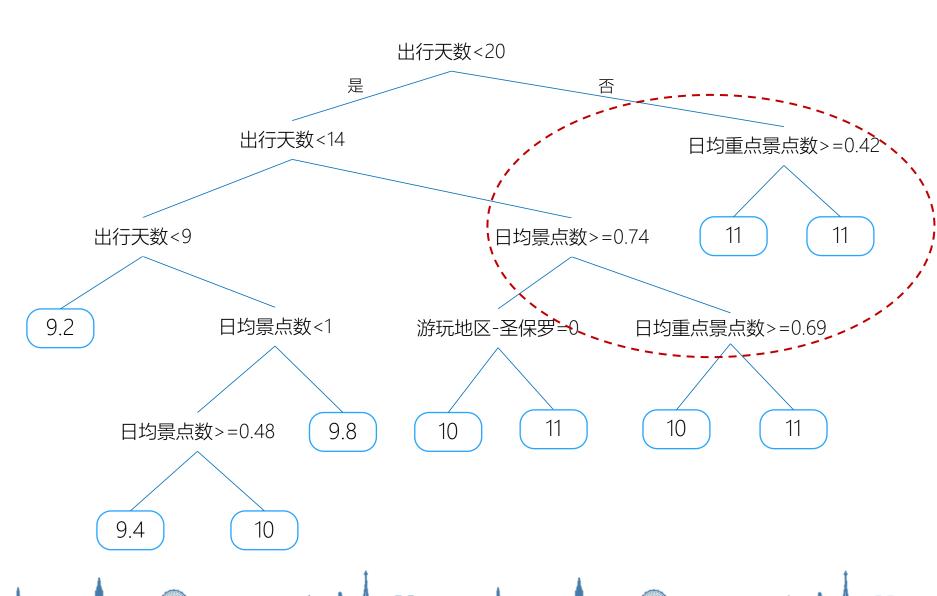


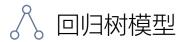


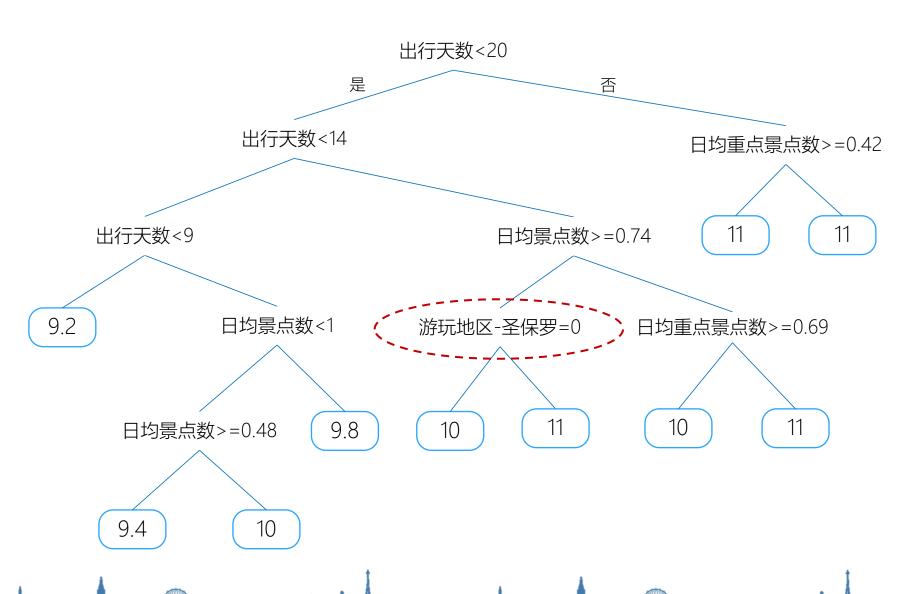








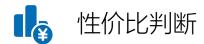


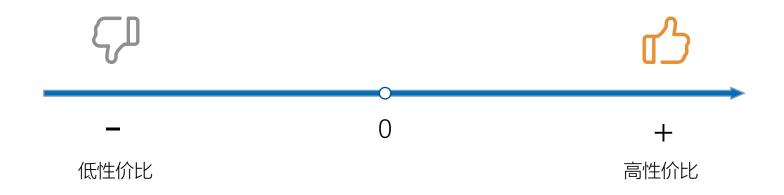




结论&展望











#### 墨西哥-古巴12日游

春节

口碑行程

玛雅文明日月金字塔 奇琴伊察 哈瓦那 海明威故居 梅里达含马车 坎昆五星 邂逅巴拉德罗 打卡网红景点 酒店升级

- 自变量

春节 古巴-巴拉德罗 12天 13个景点 33次用餐 牛人专线 128人参与过 ....

- 预测价格 32116

- 实际价格 22983

- 性价比参数 0.40





#### 结论&展望



#### 巴西-阿根廷16-17日游

南美必发团

口碑产品

北上出发 保证拼房 拒签无忧 含司导 2人wifi 4/5星酒店 2国看瀑布 中转地住1晚 17天含乌拉圭

- 自变量

非节假日 阿根廷-布宜诺斯艾利斯 16天 22个景点 42次用餐 253人参与过 ....

- 预测价格 39363

- 实际价格 39999

- 性价比参数 -0.02



#### 结论&展望



#### 美国东西海岸-夏威夷14-15日游

#### 春节

全国联运 本土无购物无自费 司导全含 拉斯夜游 圣地亚哥 纽约第五大道 更放心更乐游

- 自变量

春节 美国-可可尼诺县 14天 23景点 39次用餐 718人参与过 ....

- 预测价格 19090

- 实际价格 22253

- 性价比参数 -0.14



- 日本出境代价总体比欧洲,美洲低很多
- 若去欧美旅行,我们模型的预测可以帮助选择性价比较高的线路
- 请尽量选择牛人专线噢~





# 激 谢谢大家