

Customer Analysis of the Hotel Industry

Group 1

Leland Ball Carlos Doria Laura Pomerene



Table of Contents

- Overview
- Data Quality Assessment
- Cleansing Methodology & Final Fields
- Initial Analysis & Visualizations
- Models
 - Linear Models as Measures for Customer Priorities
 - Predicting NPS with GP_Tier
 - Statistical Score Improvement Model
- Summary of Techniques and Interpretation
- Actionable Insight

Overview

Data Set Description

Project Scope

Business Questions

- What do our best customers look like?
- What are our top performing hotels? How would we define that? In revenue or net promoter?
- Is there a survey metric correlated to the guest's likelihood to recommend?
- Does membership in gold/platinum/diamond rewards programs affect the overall hotel ratings?
- Which hotels are improving or up and coming? Are any declining?
- Do owned or franchised hotels perform better?

Data Quality Assessment

- Initial data included hotels around the globe
- Subset of United States
- All 3 relationship types are measurable
- 1 outlier from a hotel stay in the previous year

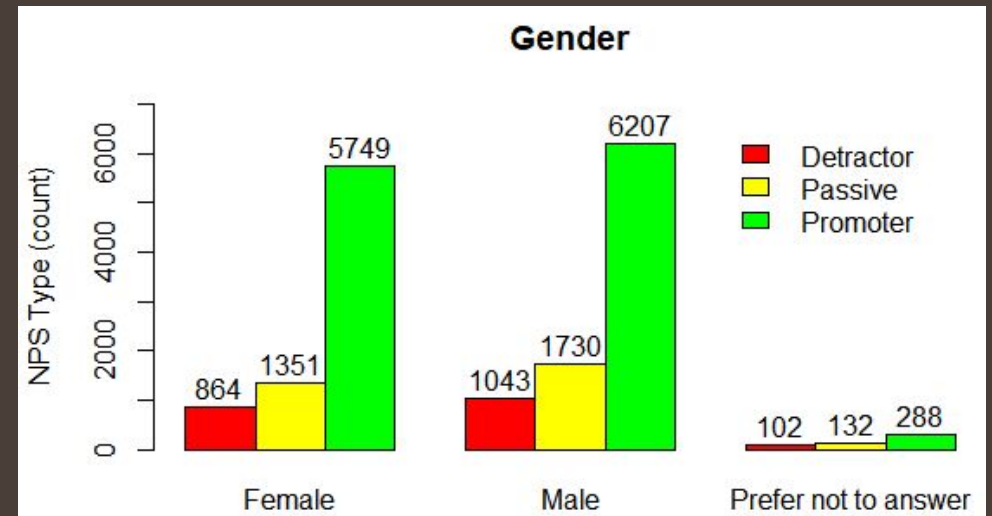
Cleansing Methodology and Final Fields

- Data load function updates format for each variable
- GP_Tier munged
- Binary fields for models
- Hotel's state, using lat/lon to state-name library
- Final count: 22 fields from data set plus 6 supporting fields

Initial Analysis & Visualizations

Net Promoter Score Customer Profile

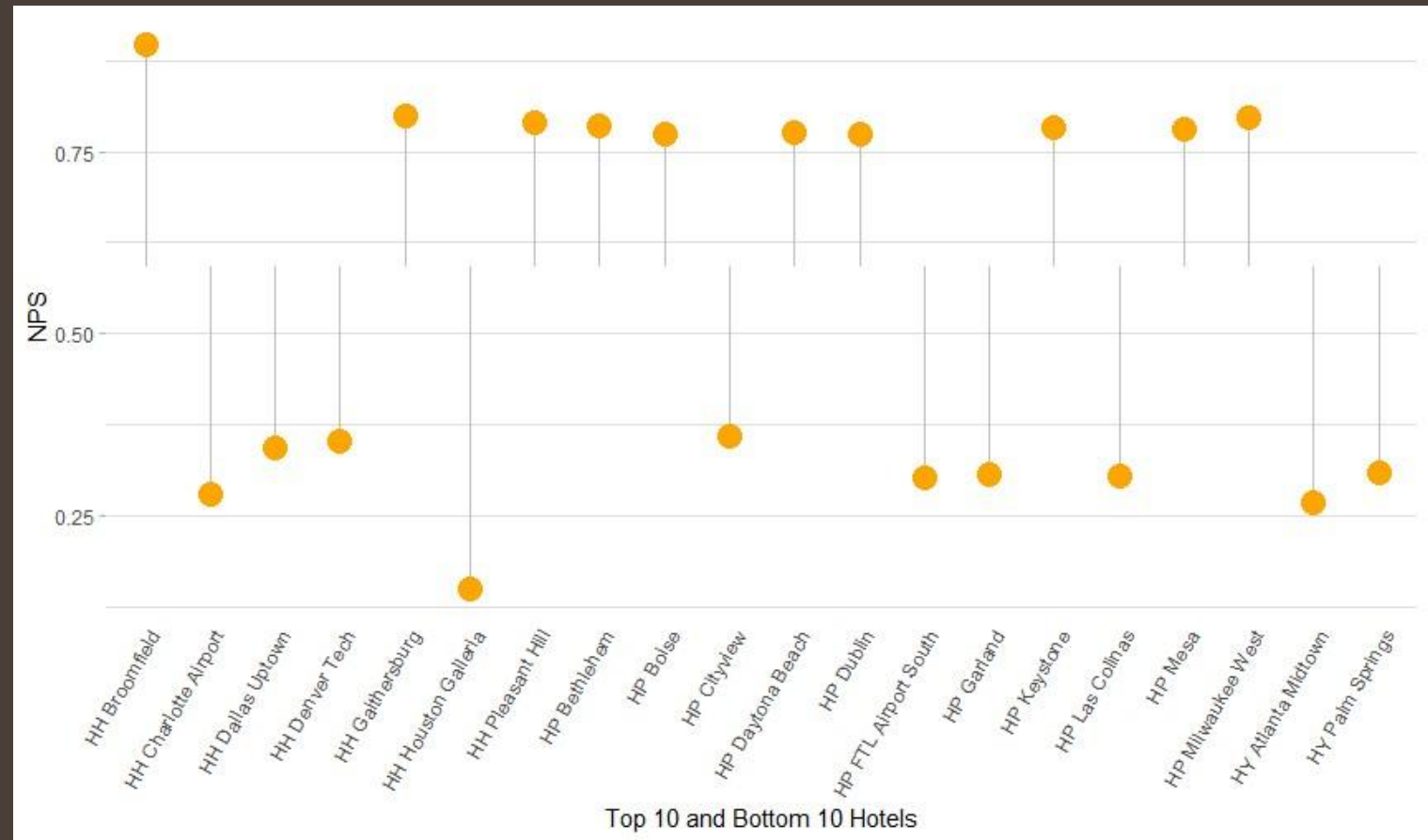
- Median age range 46-55
- 52% male
- 70% travels for business
- 91% of the time without children
- Females are more likely to be promoters



Gender_H	RECORD_CT	NPS_SCORE
Female	7964	61.33852
Male	8980	57.50557
Prefer not to answer	522	35.63218

Initial Analysis & Visualizations

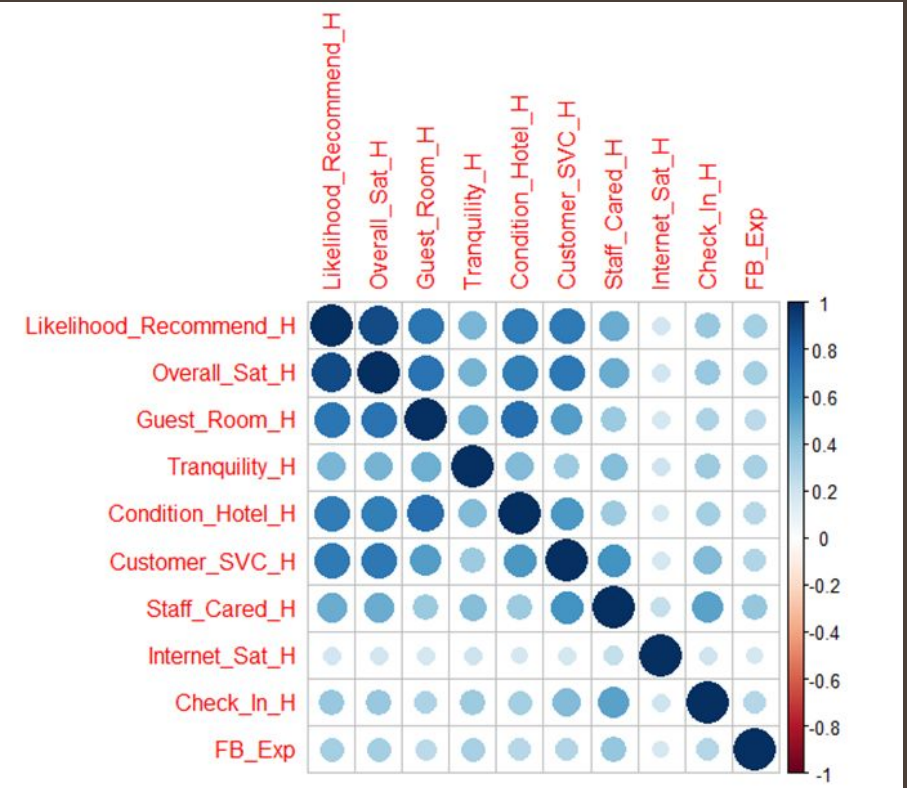
Top 10 and Bottom 10 Hotels
Highest NPS 89.7 vs. lowest 32.5



Linear Models as Measures for Customer Priorities

What drives a customer's likelihood to recommend a hotel?

- Correlation Matrix
- Linear regression model
 - Using highest correlated variables
 - Adjusted R Square 0.82
- 3 key metrics drive NPS
- Internet Satisfaction
not significant



Survey Category	Relationship	Significance	R-Squared
Guest Room Satisfaction	directly	***	0.672
Customer Service	directly	***	0.589
Hotel Condition	directly	**	0.566
Internet Satisfaction	directly	**	0.136

Predicting NPS by GP Tier with Multiple Models

Does a customer's Tier make them more likely to recommend a hotel?

If Net Present Value corresponded with NPS this could mean:

- Tiers could be a form of "carrot" to entice higher scores
- Amenities given to customers in a tier could convert passives to promoters, and detractors into passives

Linear models and Naive Bayes (NB) classifiers were attempted, but established no clear links between the two

While a NB model managed to predict successfully 84% of the time, it was only ever classifying customers as promoters

GP Tier	Record ct	Mean NPS
diamond	647	0.62
platinum	2088	0.60
gold	9274	0.58
none	5446	0.59

		Predicted		
		Correct: 84.3%	Detractor	Passive
Actual	Detractor	0	0	689
	Passive	0	0	1070
	Promoter	0	0	4063

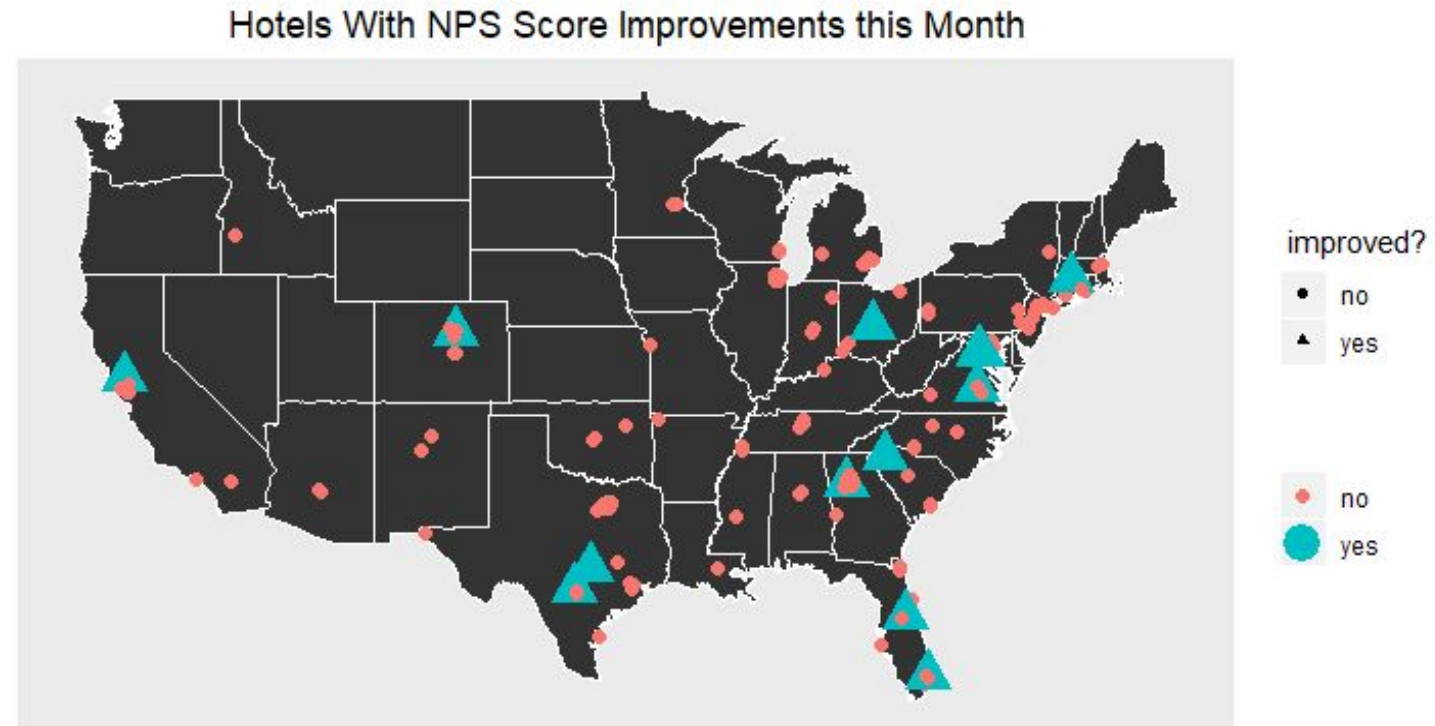
Conclusion: tiers are not correlated with NPS

Statistical Score Improvement Model

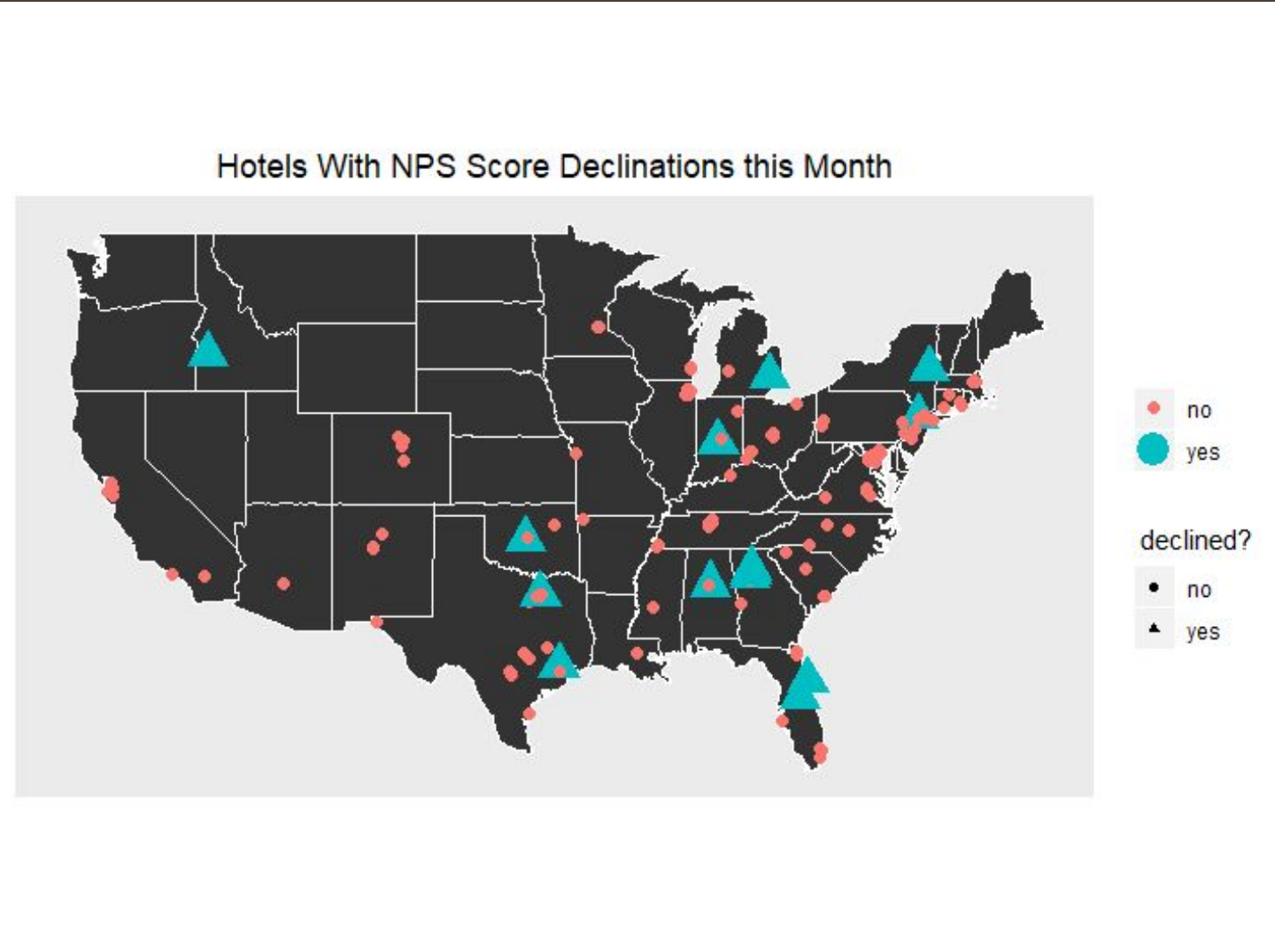
Which hotels have improved over the course of this timeframe?

- NPS averages were calculated per hotel, per day
 - Varying number of responses and timeframes
- Data was grouped into “before” and “after” groups
- Then a t-test was conducted to test whether the variation in the two groups was truly different, implying a substantial increase or decrease
- This information identifies hotels that need attention, and hotels that can be learned from

Statistical Score Improvement Model



Statistical Score Improvement Model



Summary of Techniques & Interpretation

Models and Techniques

- Measures for Customer Priorities
 - Multiple linear models
- Net Promoter Score with GP_Tier
 - Linear models
 - Naive Bayes
- Statistical Score Measurement Model
 - Student's t-tests

Actionable Insight

- Focus on hotels with declining NPS
- Focus on metrics that have proven a strong correlation to drive NPS
- Re-evaluate GP-Tier program
 - Align tier rewards with metrics identified as being impactful
 - Do not use tier membership to drive customer satisfaction
- Garner best practices from top hotels
- Revisit survey instrument
 - Tailor future questions to mesh well with questions identified as being significant



THANKS!