

Theoretically Stats

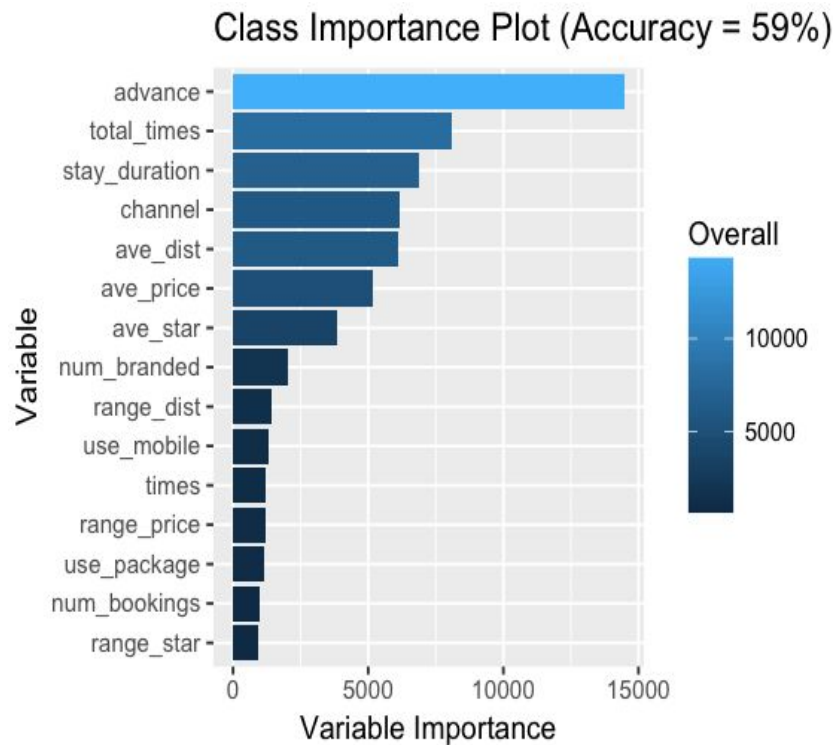
Azka Javaid, Andrew Kim, Brendan Seto, Jason Seto

Class Descriptions

- Data grouped by user's trips
 - User ID, search terms
 - Reduced size by 20%!
- Useful for Advertisements
 - Families: more expensive, more competition

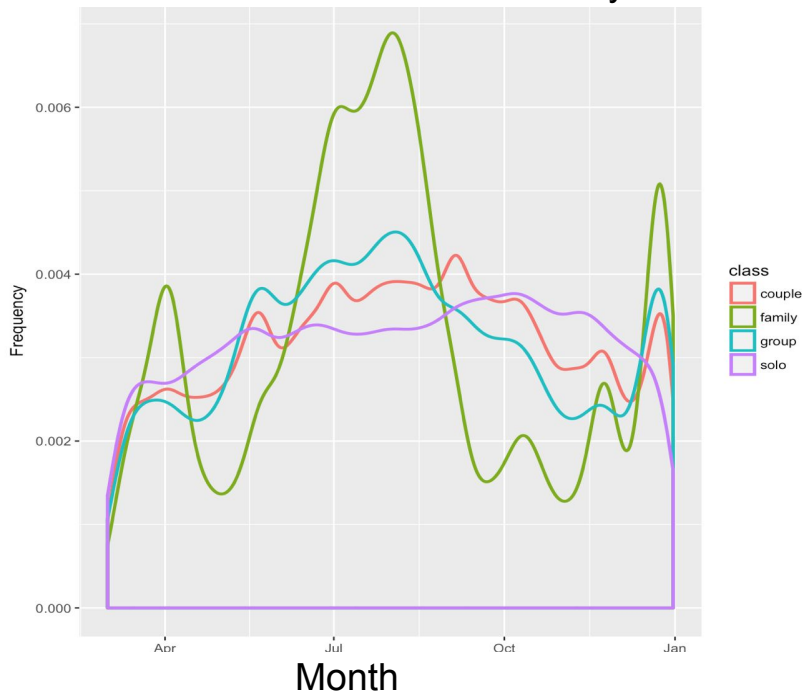
Coefficients of Multinomial Logistic Regression (Relative Odds to Couple)

	totaltimes	stayduration	ave_dist	ace_price
family	0.0177	0.0343	0.04050	0.0729
group	0.0111	0.0359	-0.00863	-0.0000
solo	0.0175	0.0357	-0.02850	-0.1070

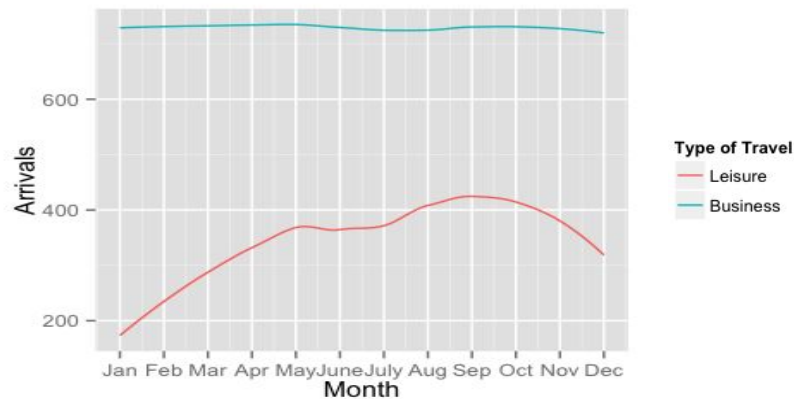


Trends Over Time

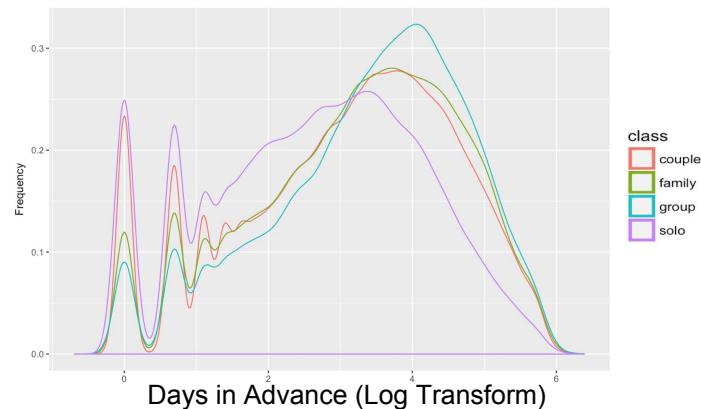
Distribution of Check-in Times by Class



Arrivals to UK Customs (External)

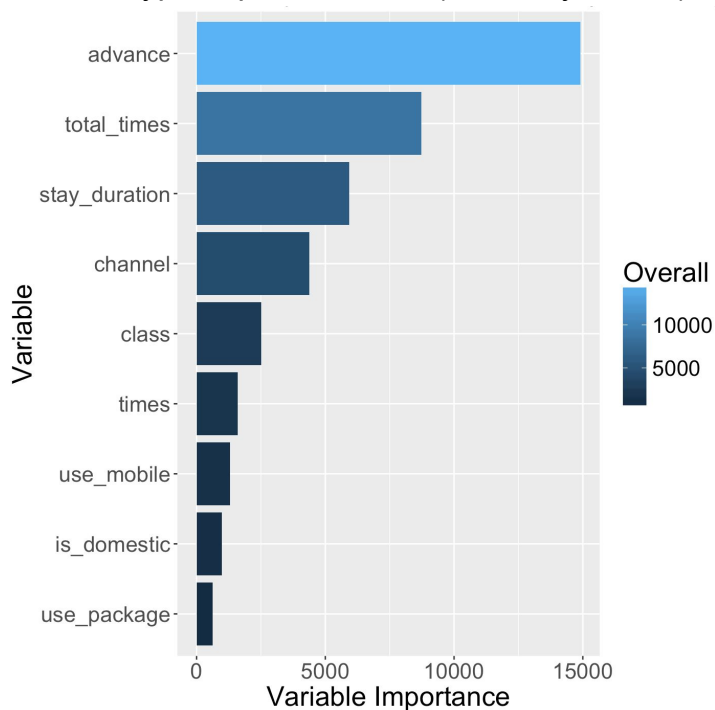


How Far in Advance People Book



Hotel Clustering

Hotel Type Importance Plot (Accuracy = 68%)



- 4 Different Types
 - Cheap/Expensive, Popular/Not
- Useful for Search Algorithm
 - Solo: Cheap foremost importance, unpopular also preferred

Coefficients of Multinomial Logistic Regression (Relative Odds to Cheap and Popular)

	advance	total_times	stay_duration	family	group	solo
Cheap,Unpopular	0.000016	-0.000673	0.01480	-0.0885	-0.1100	0.157
Expensive,Popular	0.000795	0.008900	0.01540	0.1000	-0.0352	-0.235
Expensive,Unpopular	0.000827	0.012600	0.00709	0.1280	-0.0132	-0.234