

How customers' approach on booking through Expedia

◆ Presented by:

◆ Wanyu Bian Haoyang Li Wenbin Zhou Xiwen Yang

Specified the U.S customer approach :

- Mobile device effect
- Seasonal effect
- Location and destination effect

Mobile device effect are highly influenced by region

```
Call:
glm(formula = is_booking ~ site_name.f1, family = binomial, data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.432  -0.432  -0.432  -0.432   2.239

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -2.324770    0.001641 -1416.412 < 2e-16 ***
site_name.f1EXPEDIA.CA -0.096012    0.004687  -20.486 < 2e-16 ***
site_name.f1EXPEDIA.DE -0.065298    0.008497   -7.685 1.53e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

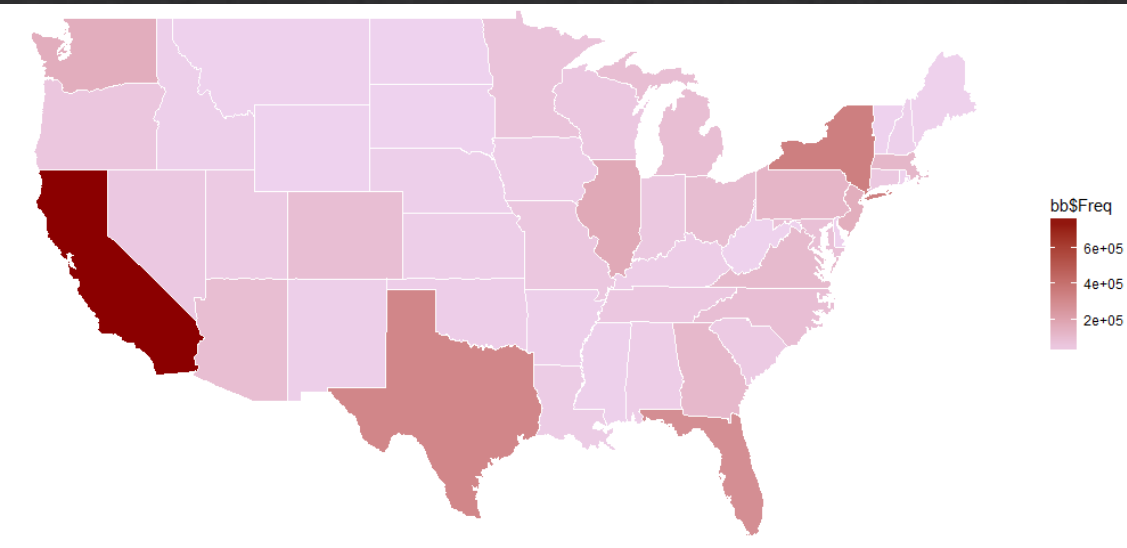
    Null deviance: 3248206  on 5453631  degrees of freedom
Residual deviance: 3247738  on 5453629  degrees of freedom
AIC: 3247744

Number of Fisher Scoring iterations: 5
```

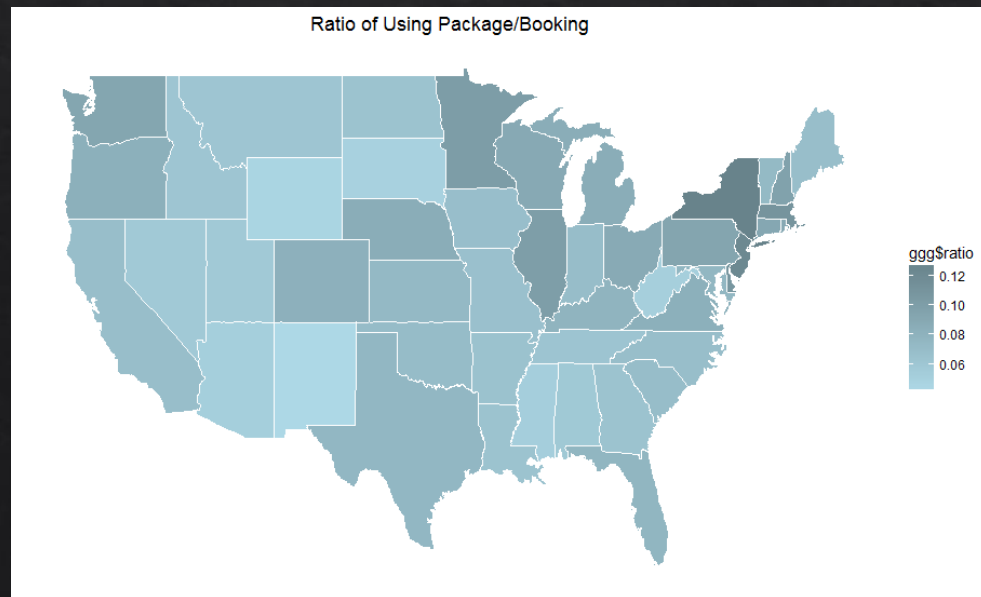


1. United States has the highest probability to book.





The graph shows the frequency that people using the app on mobile device. California has the highest frequency, along with the Washington state, Illinois, New York and Texas.



Another interest thing we found was the ratio on people who book package on mobile device in each state.

Seasonal effect based on when people start browsing.

```
glm(formula = is_booking ~ Month, family = binomial, data = data)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-0.4444	-0.4400	-0.4281	-0.4192	2.2338

Coefficients:

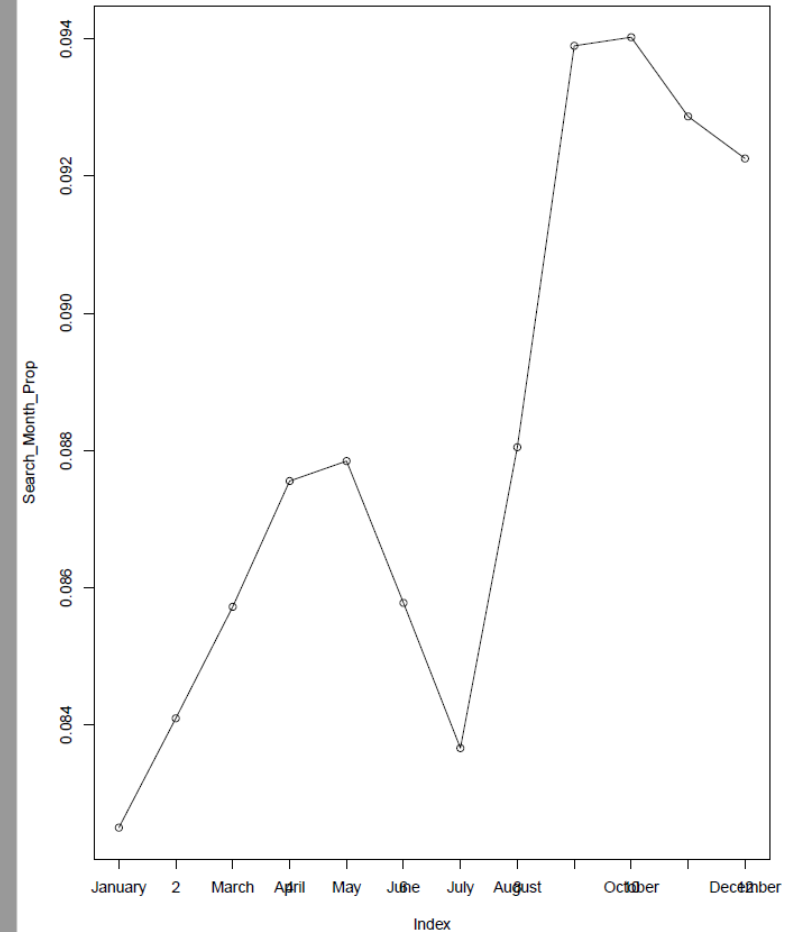
	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-2.408854	0.005907	-407.781	< 2e-16	***
Month02	0.020883	0.008152	2.562	0.0104	*
Month03	0.041808	0.007821	5.346	9.01e-08	***
Month04	0.064986	0.007950	8.175	2.96e-16	***
Month05	0.068628	0.007808	8.789	< 2e-16	***
Month06	0.042535	0.007656	5.556	2.76e-08	***
Month07	0.015223	0.007589	2.006	0.0448	*
Month08	0.071142	0.007629	9.325	< 2e-16	***
Month09	0.141940	0.007881	18.011	< 2e-16	***
Month10	0.143412	0.007953	18.033	< 2e-16	***
Month11	0.129779	0.008073	16.075	< 2e-16	***
Month12	0.122477	0.008186	14.963	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

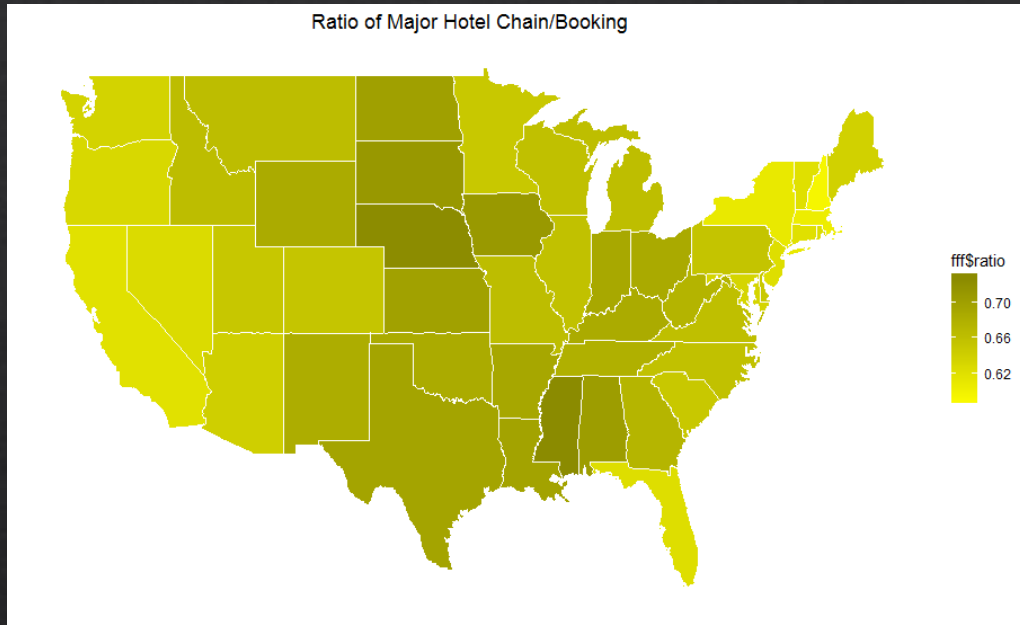
(Dispersion parameter for binomial family taken to be 1)

Null deviance: 3248206 on 5453631 degrees of freedom
Residual deviance: 3247214 on 5453620 degrees of freedom
AIC: 3247238

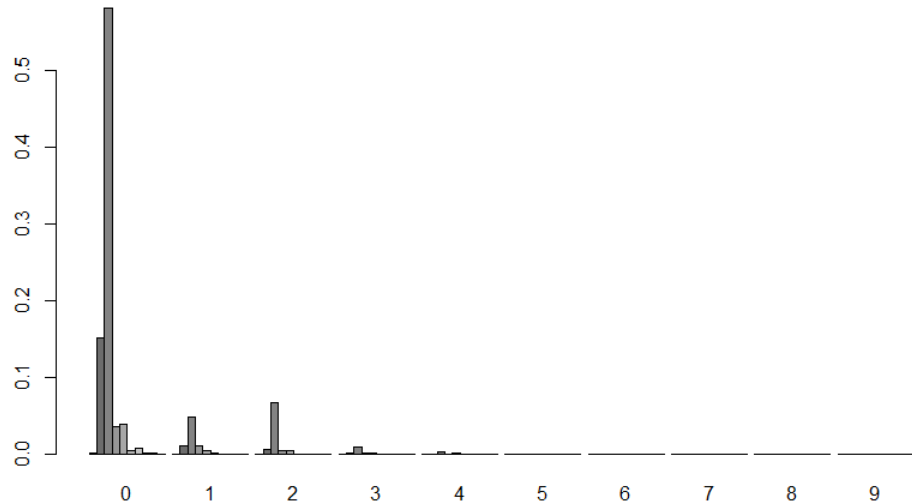
Number of Fisher Scoring iterations: 5



Location and destination effect



→ We find out people who booked hotels in inner region more specific by brands but people who booked near coastal region not really sensitive on brands.



→ Most of people who booked through Expedia are more likely being party of two. And they might reach agreement on searching a typical band they both knew.