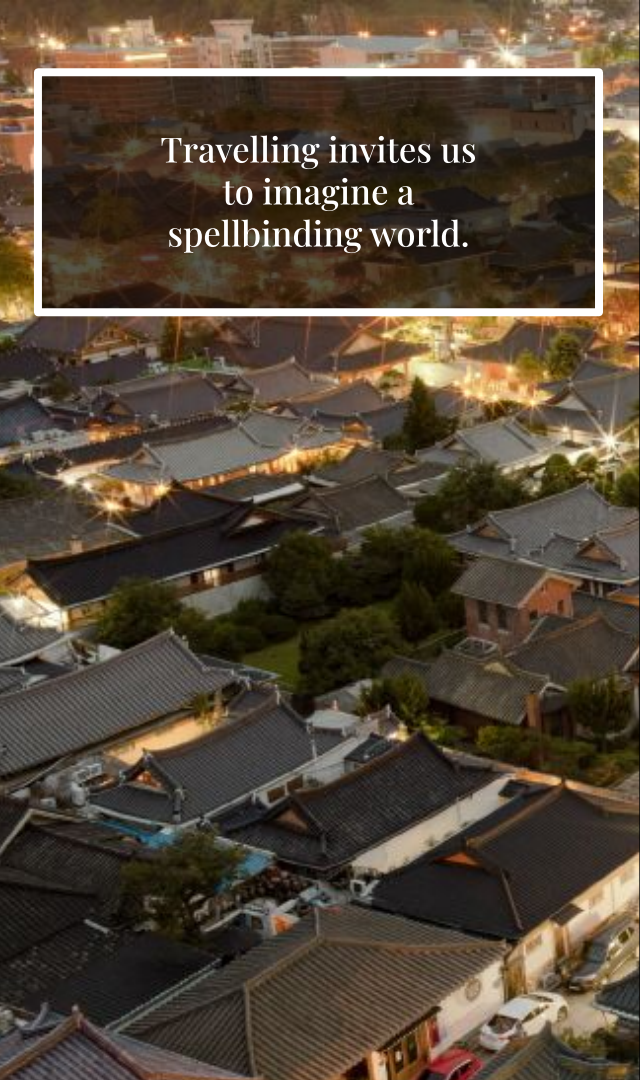


You can now enjoy **Data-Driven Travel**


Viewing the world from the Instagram perspective

CS564
Team 8: Final

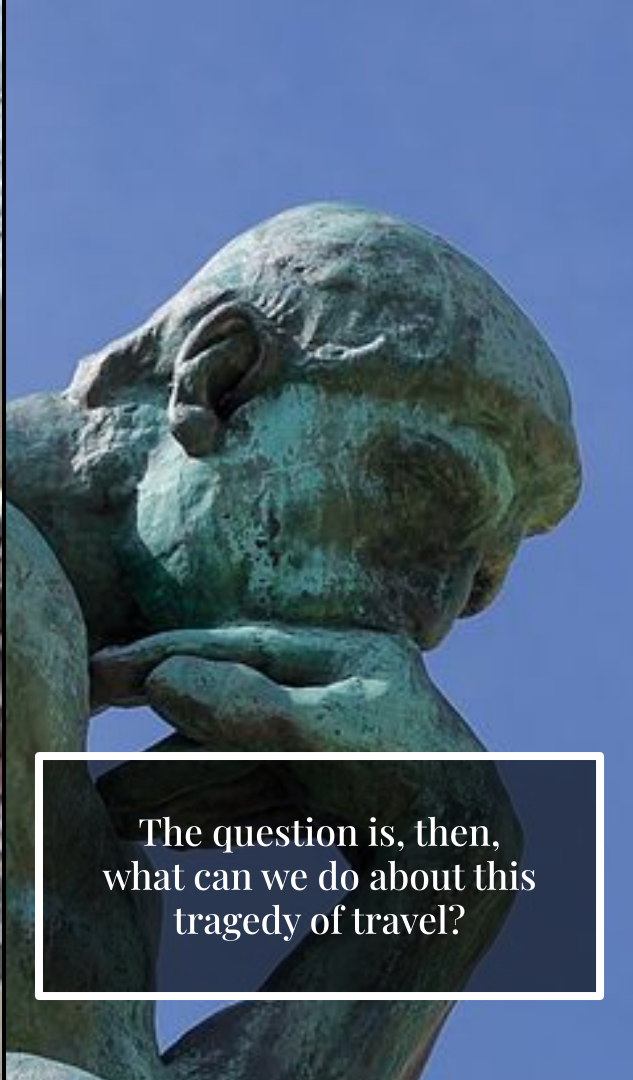




Travelling invites us
to imagine a
spellbinding world.

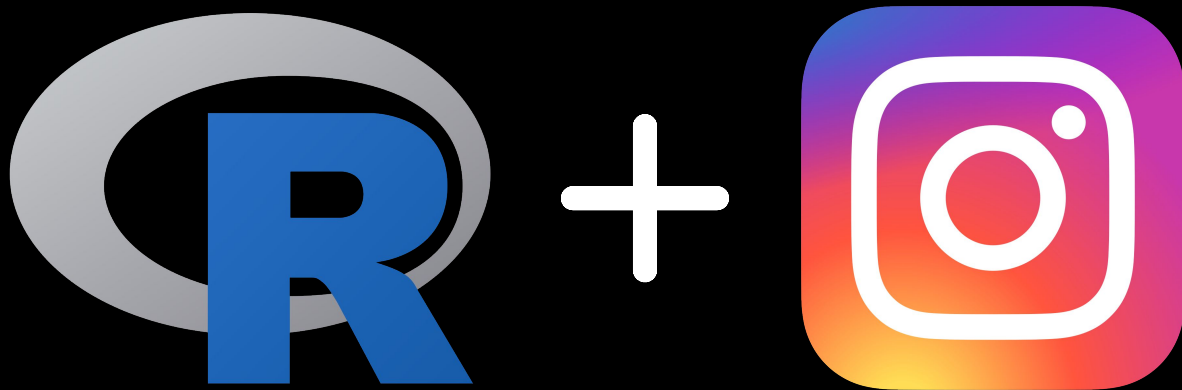


All too often, we are greeted
with unpleasant surprises.



The question is, then,
what can we do about this
tragedy of travel?

We proposed one hope:



Thanks For The Attention!



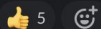
Meeyoung Cha 30 days ago

How would you know which hashtag is about locations vs topics. Is there an easy way? Also how would you scope down on topics? which tests will be used? (edited)



Jeongyeon Kim 30 days ago

How are you guys going to filter the themes? Are they manual inputs or extracted from the existing hashtags? I mean, not all the noun keywords are appropriate as the themes, so was wondering if there's any criterion to be classified as themes



20160843_Peerapon 30 days ago

how will you find those certain keys to capture (theme) the correlation? Or will you choose some specific keys?



Eunji Woo 30 days ago

Are you going to exclude postings for advertising?
And how would you choose the destination>



Seung Eon Lee 30 days ago

Actually, the travel means a path of sight points. How about path recommendation including the popular place introduced by your work?



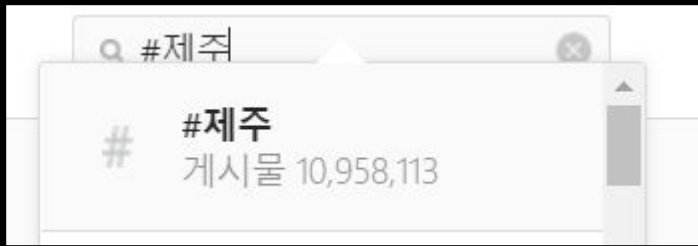
Jeeyoung Choi 30 days ago

There will be many fake hashtags that not really related to the post, how would you refine data

Some improvements we made:

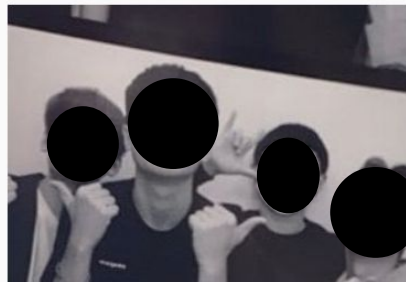
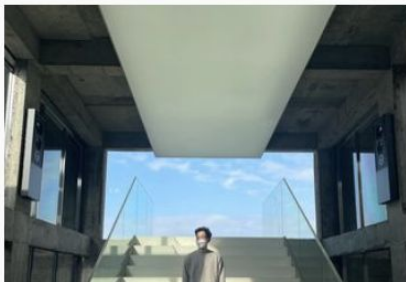
- The use of TFIDF to extract important hashtags
- We were careful about the effect of fake hashtags when performing analysis
- Inspired by the path suggestion, we performed time-dependent analysis

The Data That We Got

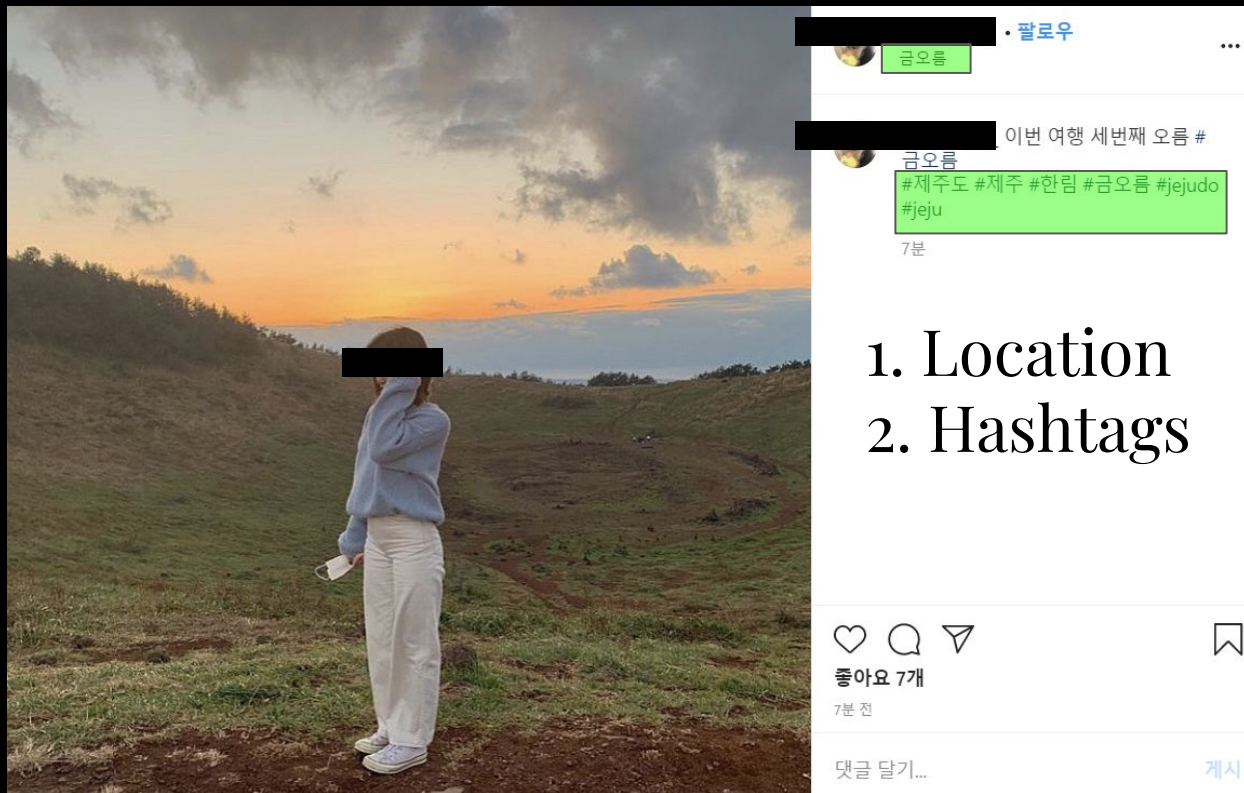


The Data That We Got

최근 사진



The Data That We Got



Our Data: Statistics



934

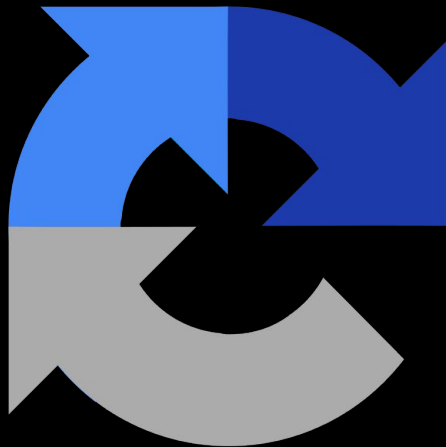
Unique Posts



8082

Hashtags

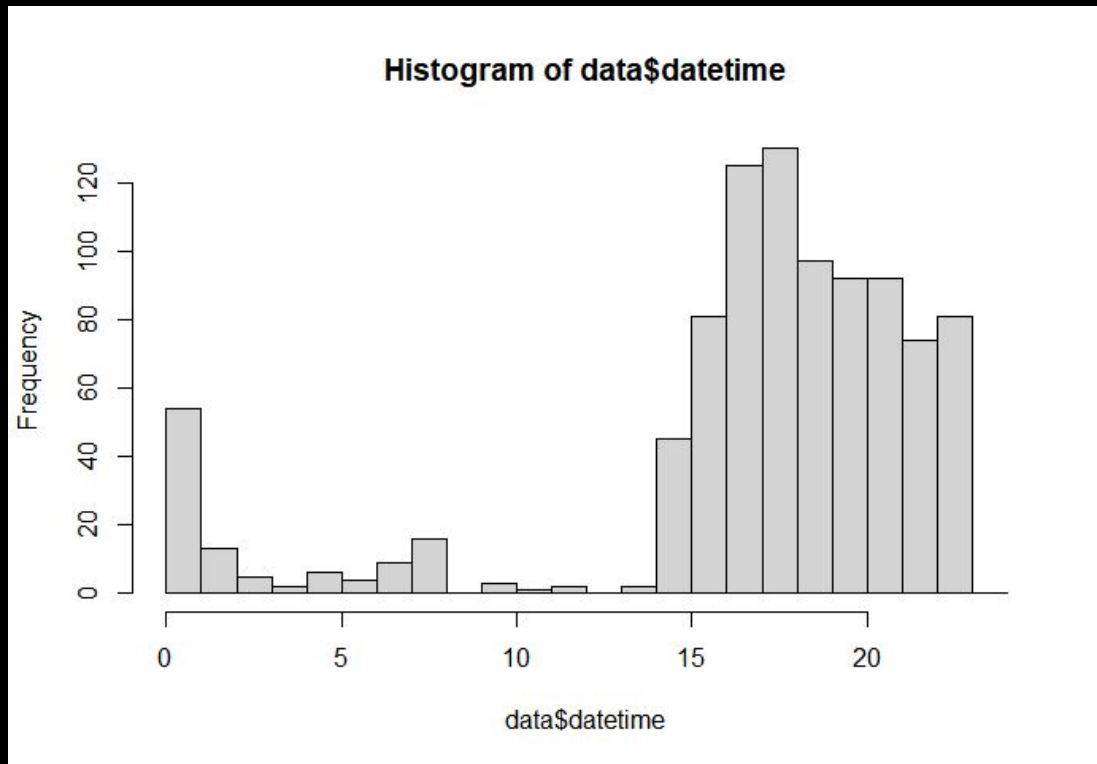
Our Data: Caveats



reCAPTCHA



Our Data: Caveats



Our data spans about a day.



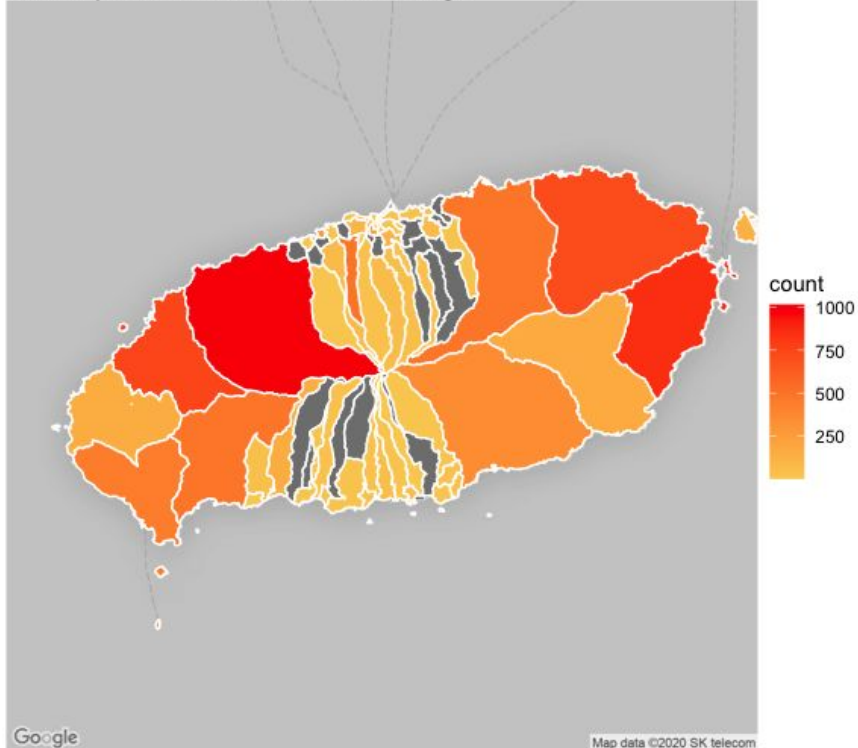
Data does *not* span full day.



Only picks up Instagram trends

Hot Places In Jeju

heatmap based on number of hashtags



```
> chisq.test(mo)
```

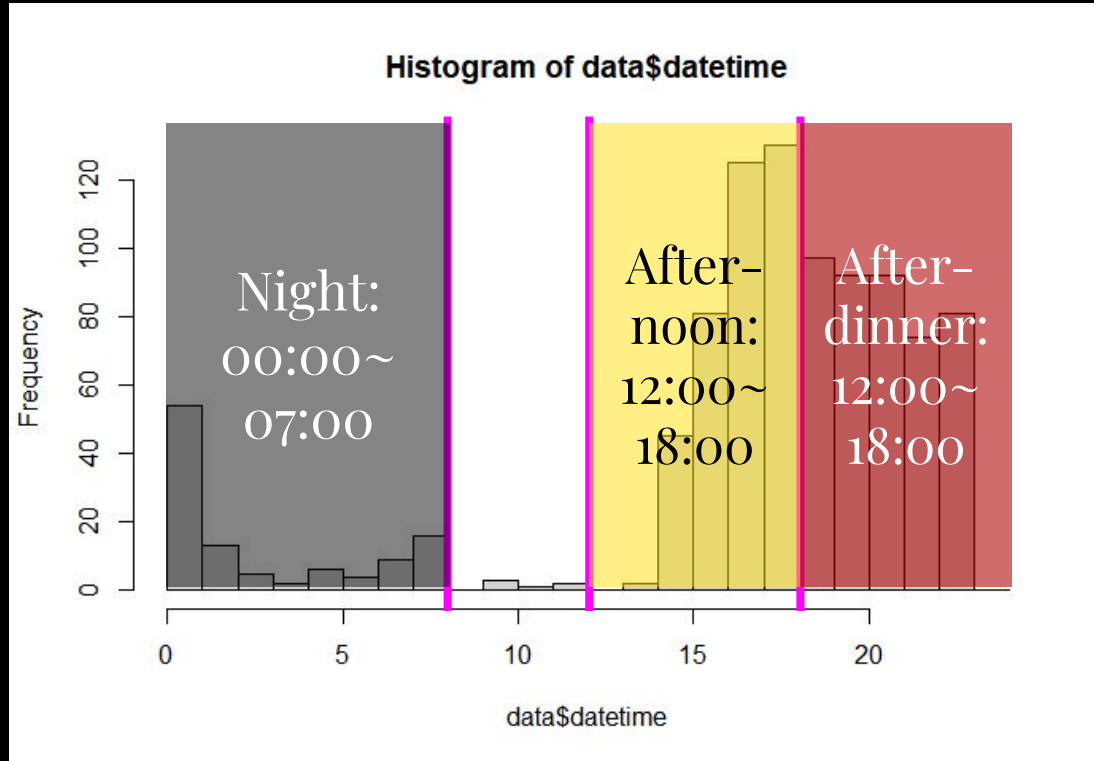
Chi-squared test for given probabilities

```
data: mo
```

```
X-squared = 4985.9, df = 64, p-value < 2.2e-16
```

```
> |
```

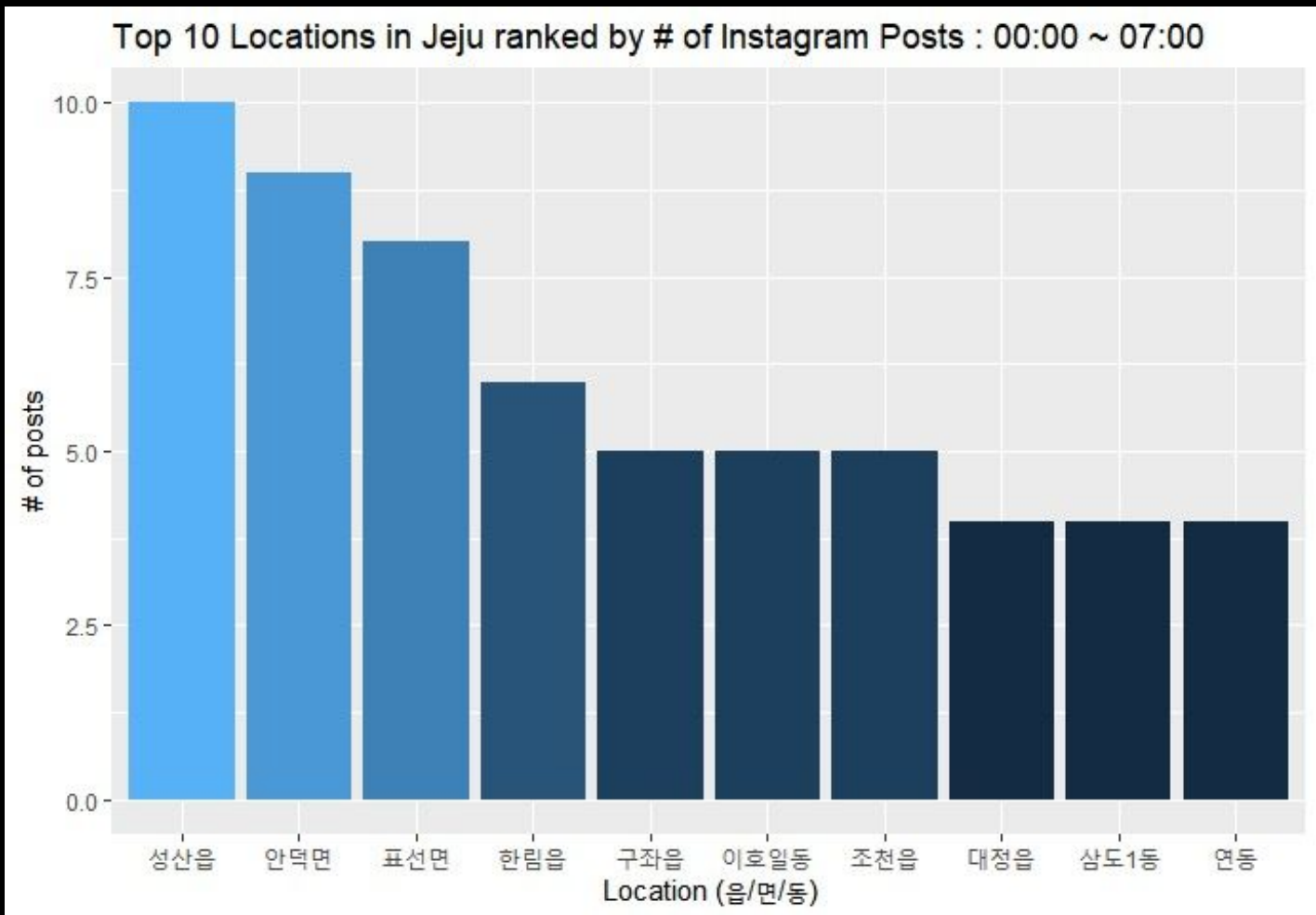
When Places are Hot



When Places Are Hot

Top 3 in
Night:

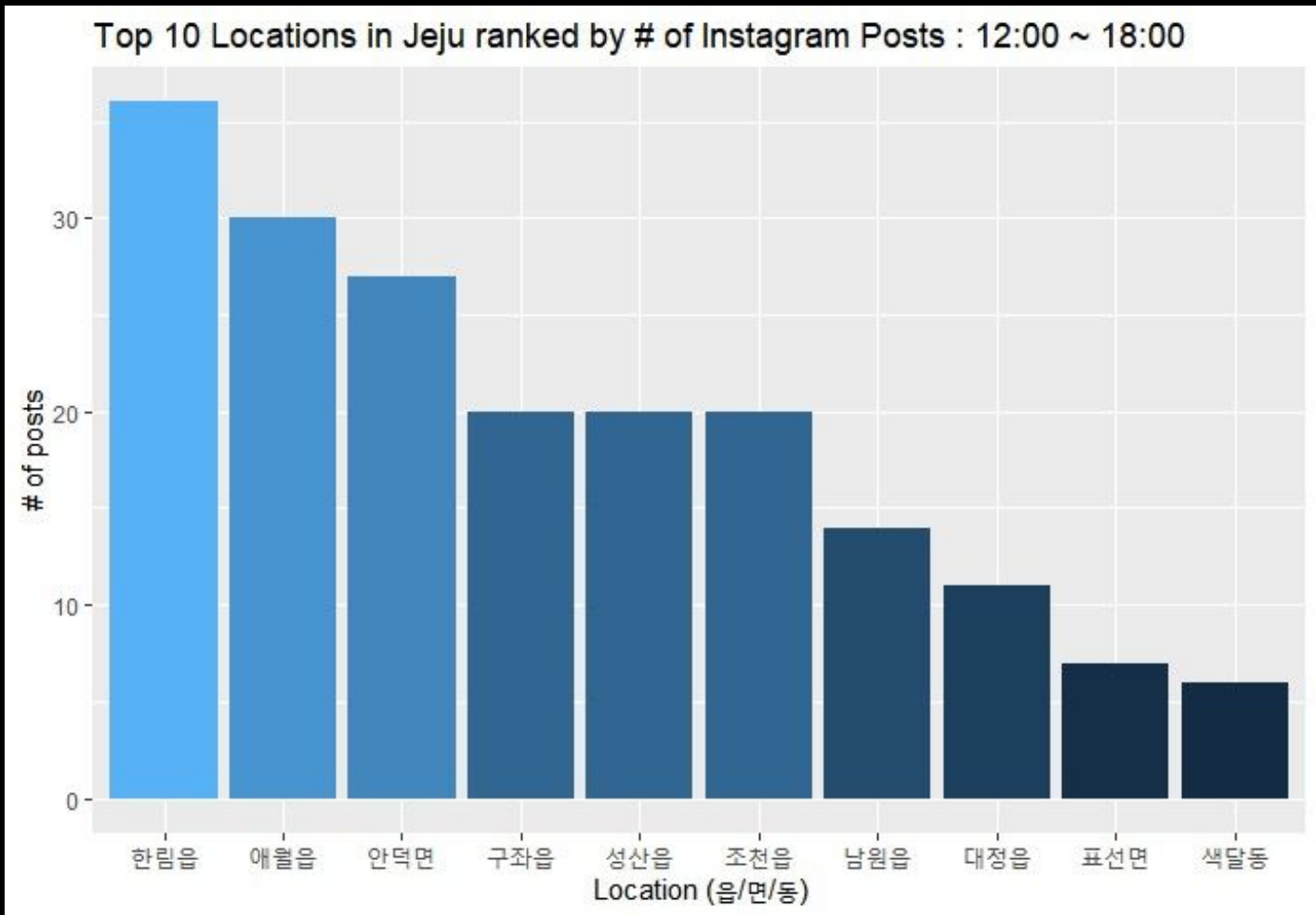
- 성산
- 안덕
- 표선



When Places Are Hot

Top 3 in
Afternoon:

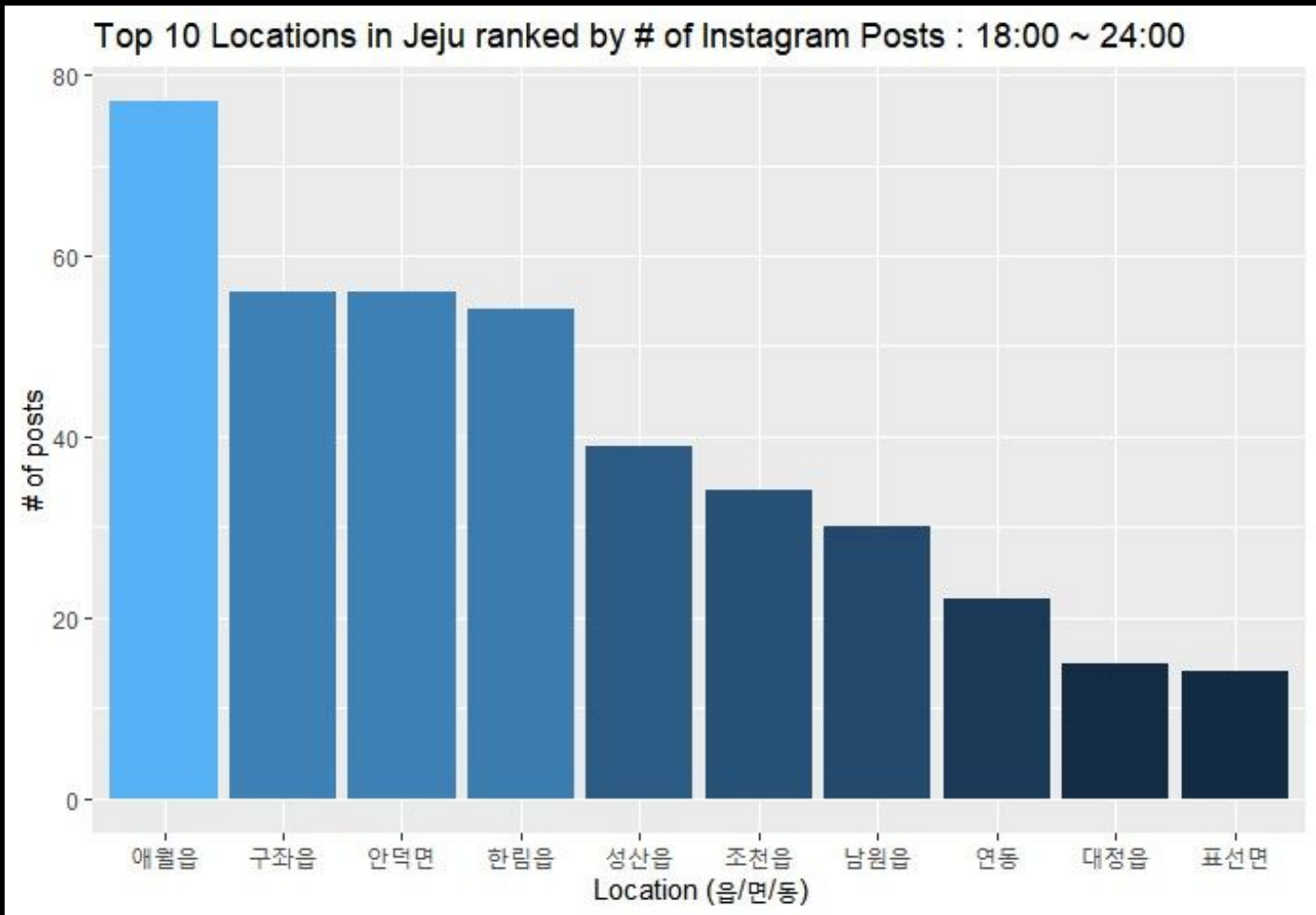
- 한림
- 애월
- 안덕



When Places Are Hot

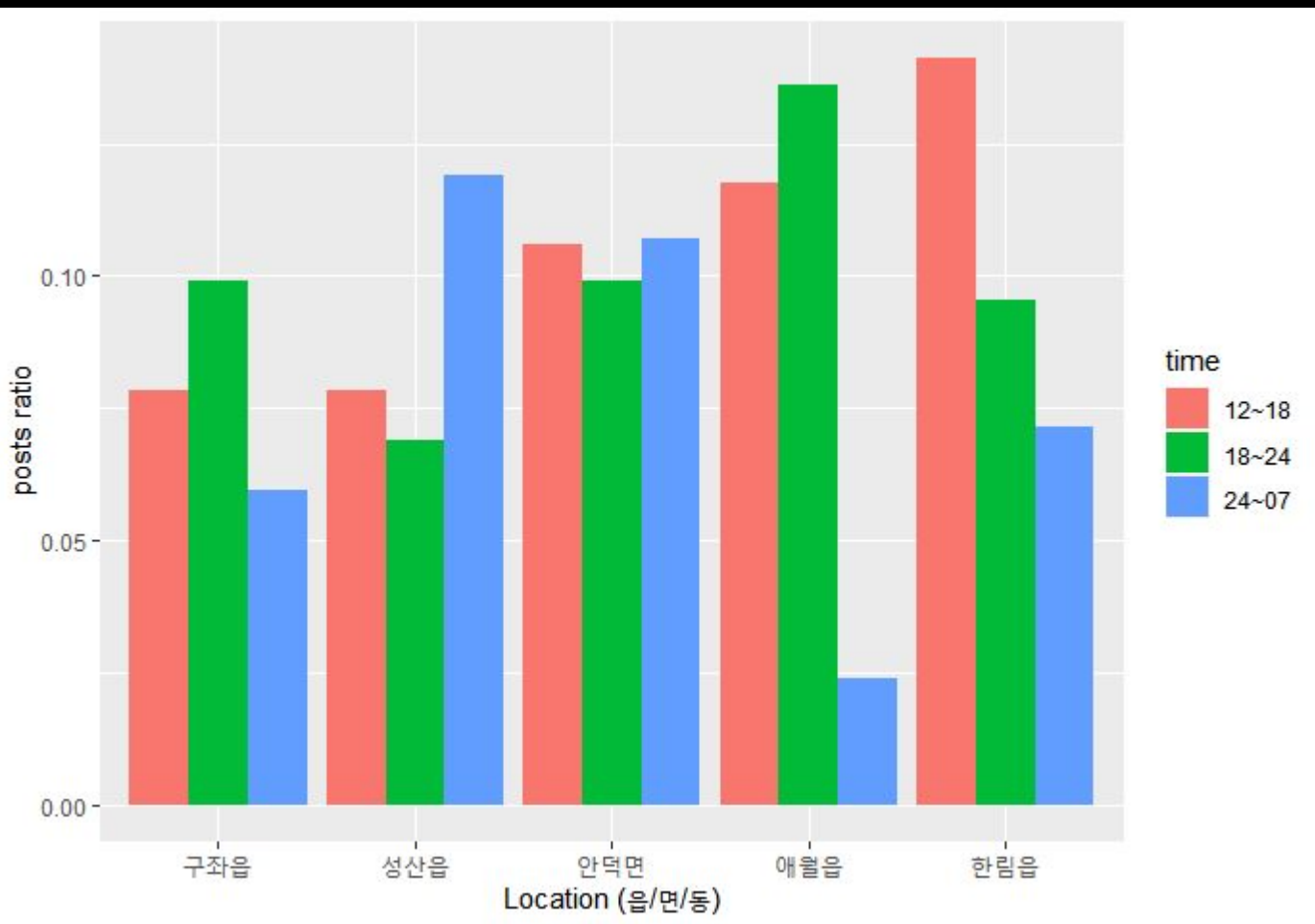
Top 3 in
After-dinner:

- 애월
- 구좌
- 안덕



When Places Are Hot

Time-hotness
In hot regions

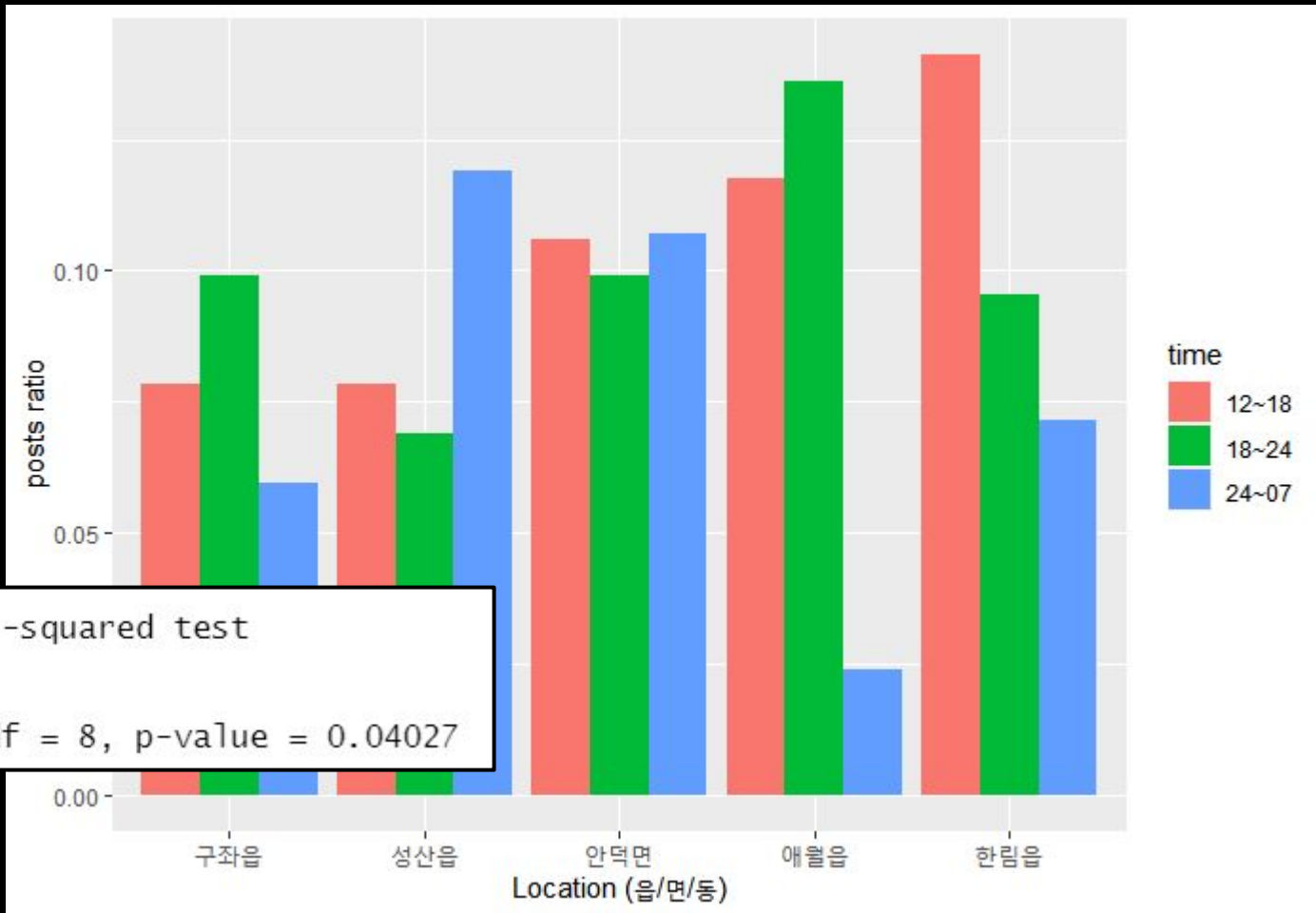


When Places Are Hot

Time-hotness
In hot regions

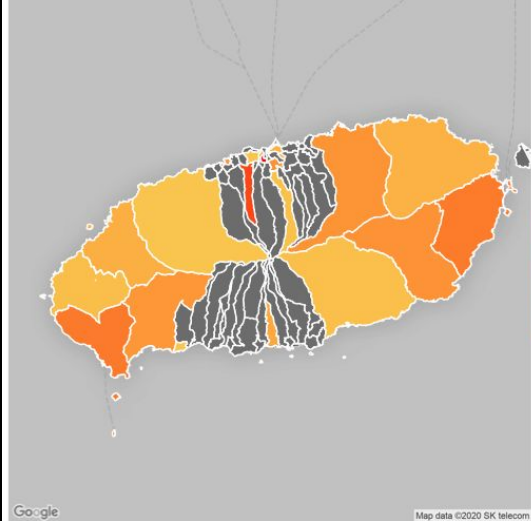
Pearson's Chi-squared test

```
data: as.table(t)
X-squared = 16.151, df = 8, p-value = 0.04027
```



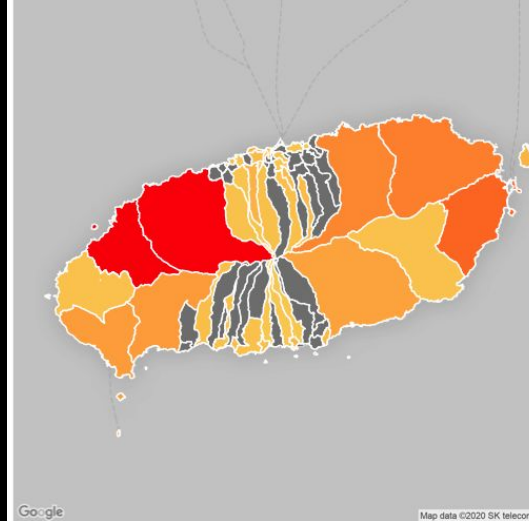
When Places are Hot

heatmap based on num of hashtags: 0AM~7AM



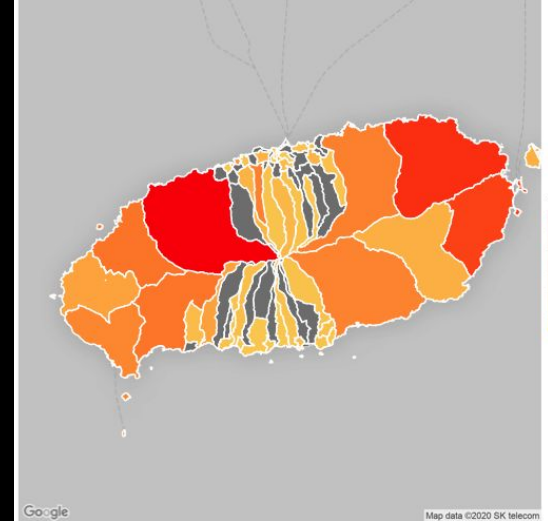
Night
(00-07)

heatmap based on num of hashtags: 12PM~6PM



Afternoon
(12-18)

heatmap based on num of hashtags: 6PM~12AM



After-dinner
(18-24)

Why Places are Hot



Step 1.
Get #s from region

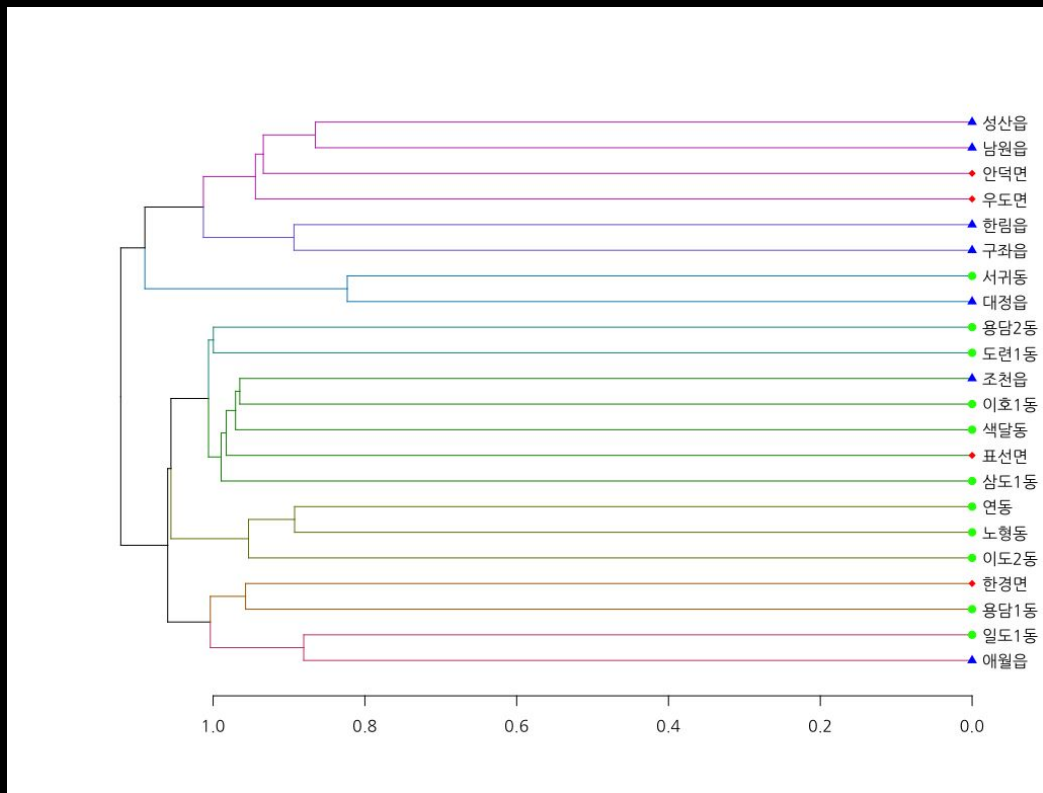


Step 2.
Split hashtags

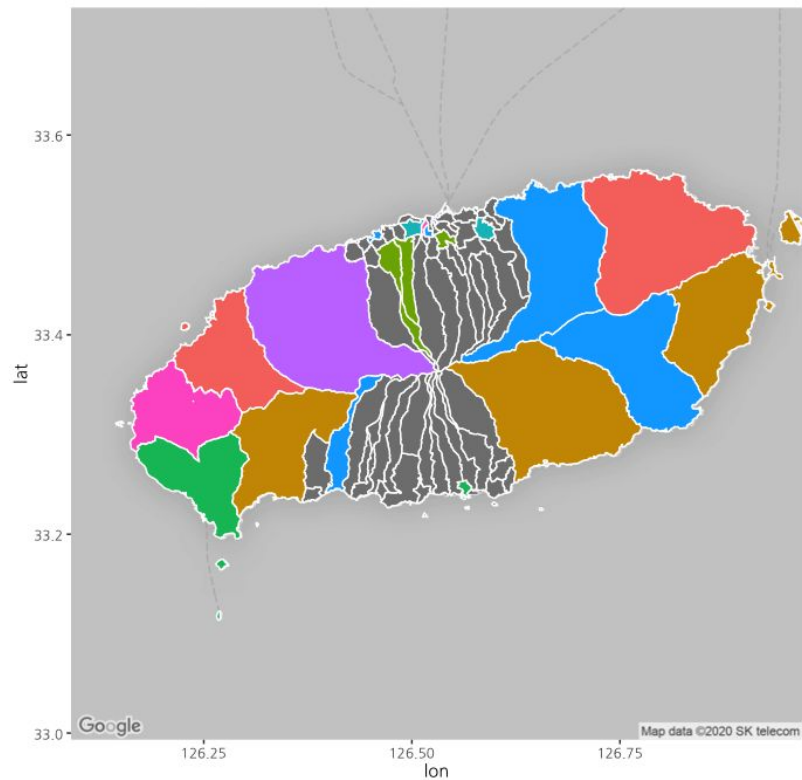
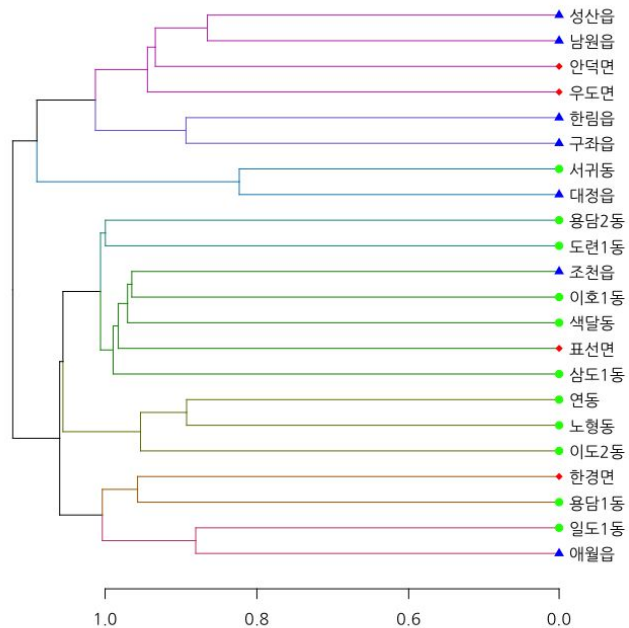
제주	0.01
여행	0.03
도	0.02
애월	1.10
맛집	0.50
바다	0.12
풍경	0.13
오션	0.44
뷰	0.12

Step 3.
Get TFIDF vector

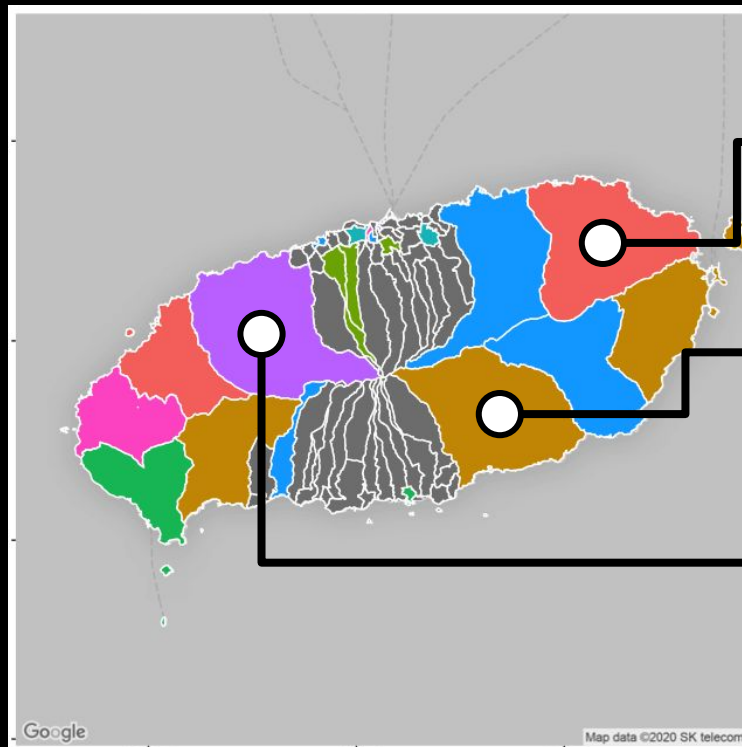
Why Places are Hot



Why Places are Hot



Why Places are Hot

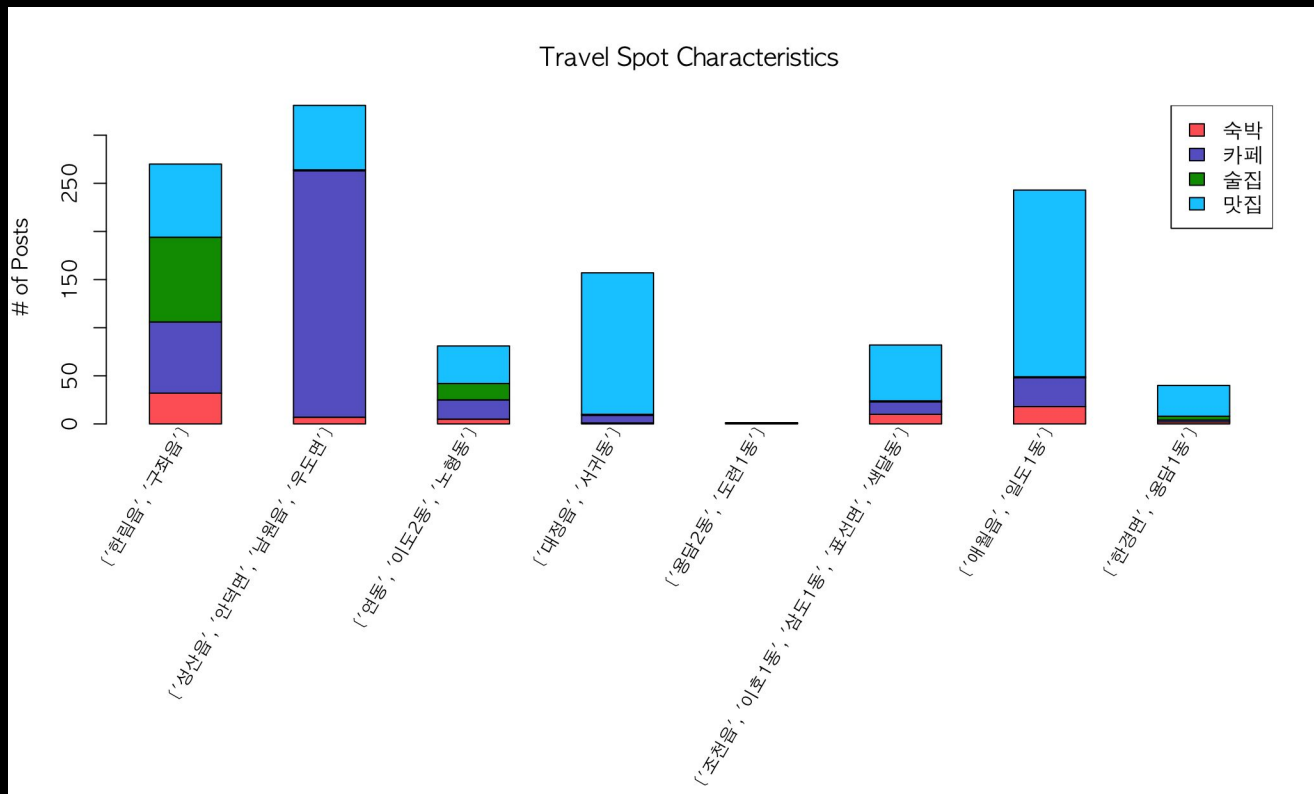


● #협재술집 #제주도술집 #바다

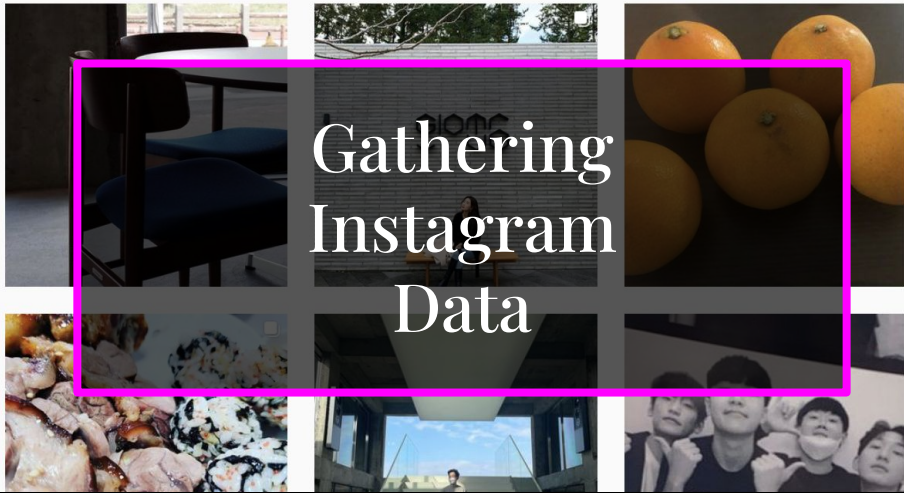
● #서귀포카페 #빛의병커

● #애월맛집 #애월밥집추천

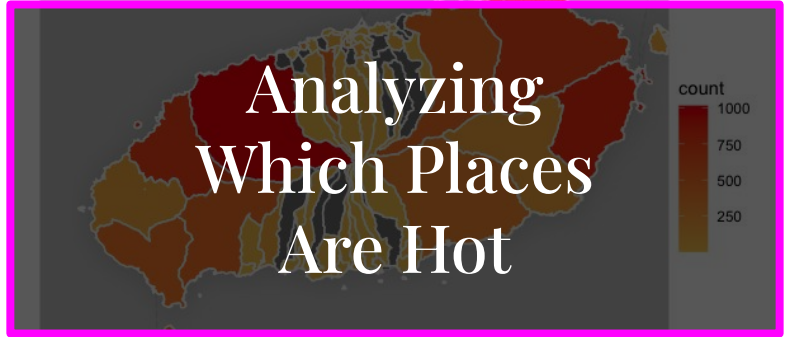
Why Places are Hot



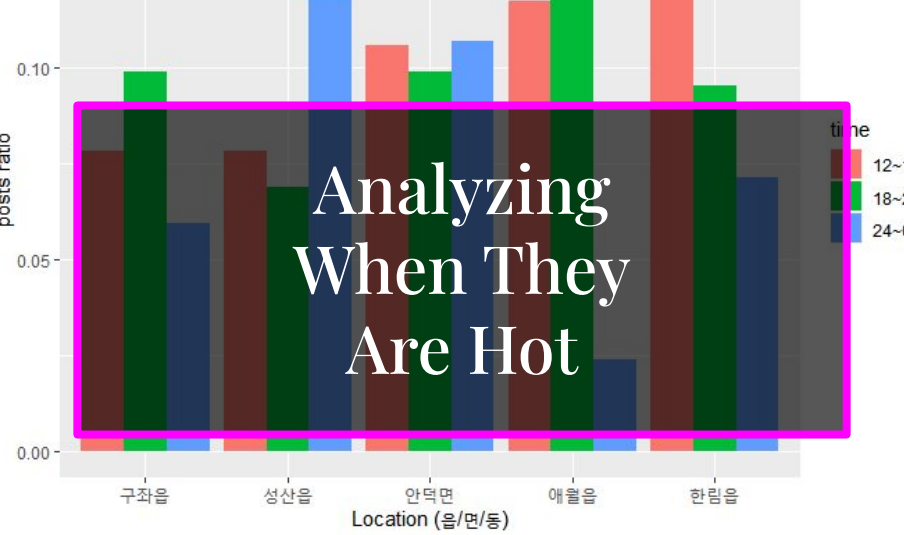
[illegible]



Gathering Instagram Data



Analyzing Which Places Are Hot



Analyzing When They Are Hot



Thanks for joining us.

We welcome any questions.

CS564
Team 8

