

Using Location Posts as a Proxy for Nature Reserve visits

Roman Guerin, Marinus van den Oever

LIACS, Leiden University, Niels Bohrweg 1, Leiden, The Netherlands
romanguerin@gmail.com, marinusvandenoever@gmail.com

1 Introduction

On the 23rd of March 2020, the Dutch government announced the ‘intelligent lockdown’ in the Netherlands. To stop the spread of Covid-19, citizens were advised to work from home. Since work is one of the most important sources of physical activity [1], and commentators have expressed concerns over the impact of these measures on fitness [2], which is essential for humans as it reduces the risk of many health problems [3]. It is possible however, that in parts of the society, the imposed limitations of normal daily activities, combined with an increase in flexibility and free time, may have led to compensatory behaviour, for instance by an increase in visits to nature reserves. Natuurmonumenten, a Dutch nature conservation association, reported an increase in the number of website visitors, downloads and memberships [4]. It is suggested that this is related to COVID-19 crisis, but little more information is provided on the distributions and patterns of the increased interest in natural reserves. We hypothesize that Instagram locations posts can be used to give an insight to visits into nature, since 47.1% of the Dutch population has an Instagram account [5]. Instagram is a social media platform where users can share photos and videos with the world. Posts can be found by searching for an account (user), tag (hashtag) and place (location). The platform allows users to add a location of where the photo/video was taken. The post can then be found under the specified location tag; the posts are currently publicly available by default unless users have set their profile to private. The location tag has existed since the beginning of Instagram, making it possible to compare time spans with each other. Collecting this data may also give insight into the number of visits to specific locations. Since Instagram did not secure their API, anyone can retrieve data without an access token.

2 Method

To test this proxy, a selection was made of 18 nature reserves (figure 1) in South Holland-based on their popularity on Google Maps. The locations used are forest areas, small lakes and dunes, concentrated around Leiden and the rest of South Holland. These locations were then looked up on Instagram, and their associated IDs were entered in a JSON file. Some locations that can be found under different names were combined. For example, Tiengemeten is also known as Natuureiland Tiengemeten. The metadata (caption, image-link, description, time-stamp, location) of posts between 1 January 2019 and 31 December 2020 were retrieved using a Python command-line application [6]. It scrapes posts, and metadata from Instagram’s explore location page.

The application can be used as follows:

```
python scraper.py -max "2020/12/31" --min  
"2019/01/01" --location "Tiengemeten"
```

As this is an automated process, we have manually counted some of the data to validate that it is indeed complete. The script saves the data into a JSON lines file. We have written a custom script that extracts Post ID, date and location as relevant to this study. The data has been converted to CSV format in order to be analysed in Microsoft Excel. Duplicates were removed by detecting ID’s that occurred twice in the dataset.



Figure 1: Map of the selected nature reserves

3 Results

Figure 2 shows the total number of posts from all locations per week for 2019 and 2020. It shows an average increase of 41,9% in 2020 in comparison with the previous year. The line of 2019 is more consistent compared to 2020, which generally runs higher from April to December. The most significant difference between both years is a steep increase in March to May. In both years there is a peak in the warmer months from May to November. Besides, we highlighted three periods of stricter lockdown measures in 2020. From June 1st to August 18th lockdown measures have decreased due to less corona cases. The number of infections increased from August 18th and a second lockdown was announced on October 13th [7].

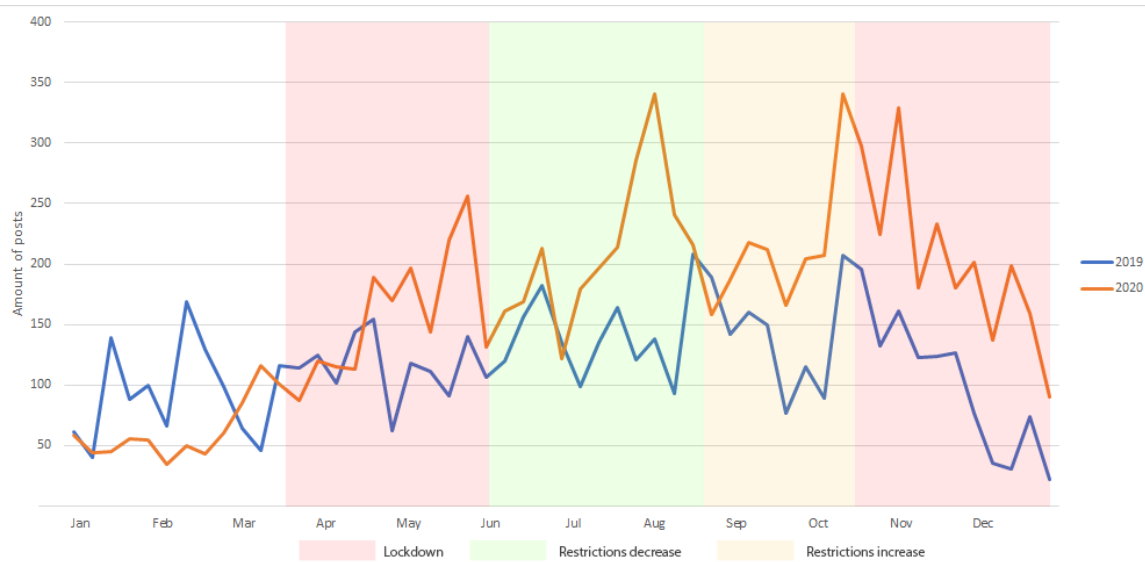


Figure 2: Weekly count of all location posts, years 2019 and 2020. Red indicates the lockdown, green indicates the restrictions decrease and orange when the restrictions increased.

The same pattern can be seen in figures 3 and 4 for individual nature reserves. Some areas which were rarely visited in the spring of 2019, like Landgoed De Horsten or Kagerplassen, showed a sudden increase after April 2020. Most nature reserves were relatively poorly visited in late winter and early spring in both 2019 and 2020. Meijndel is an apparent exception, but this is caused by an absence of data from January and February 2020. For figure 3 & 4 a monthly count instead of a weekly count was chosen to make the graph more readable.

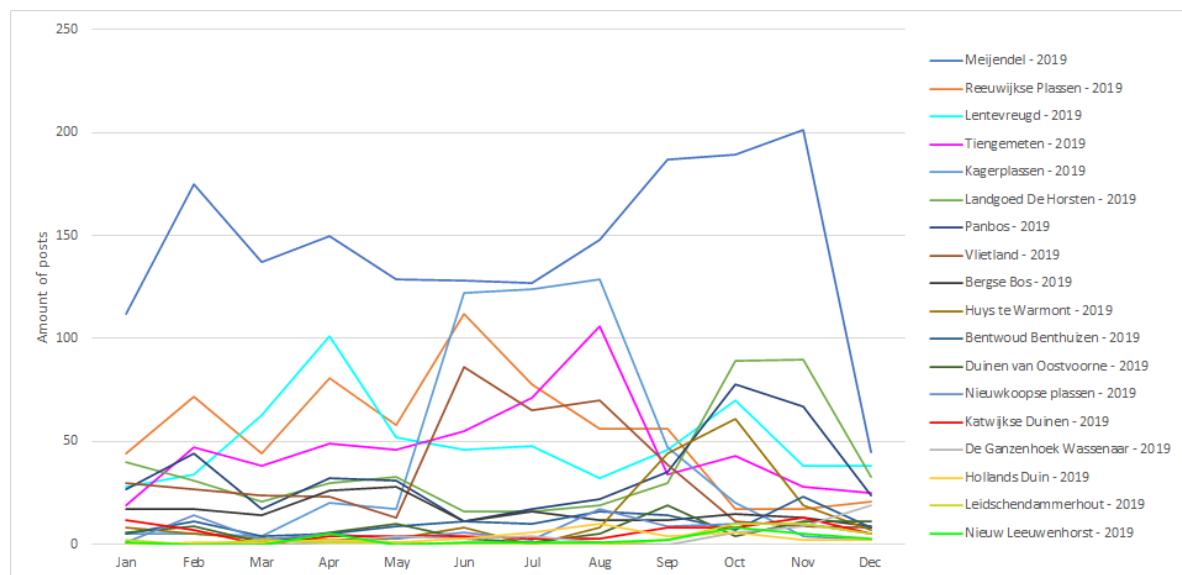


Figure 3: Monthly number of posts for each location from 1 January 2019, to 31 December 2019. Each individual line represents a specific location.

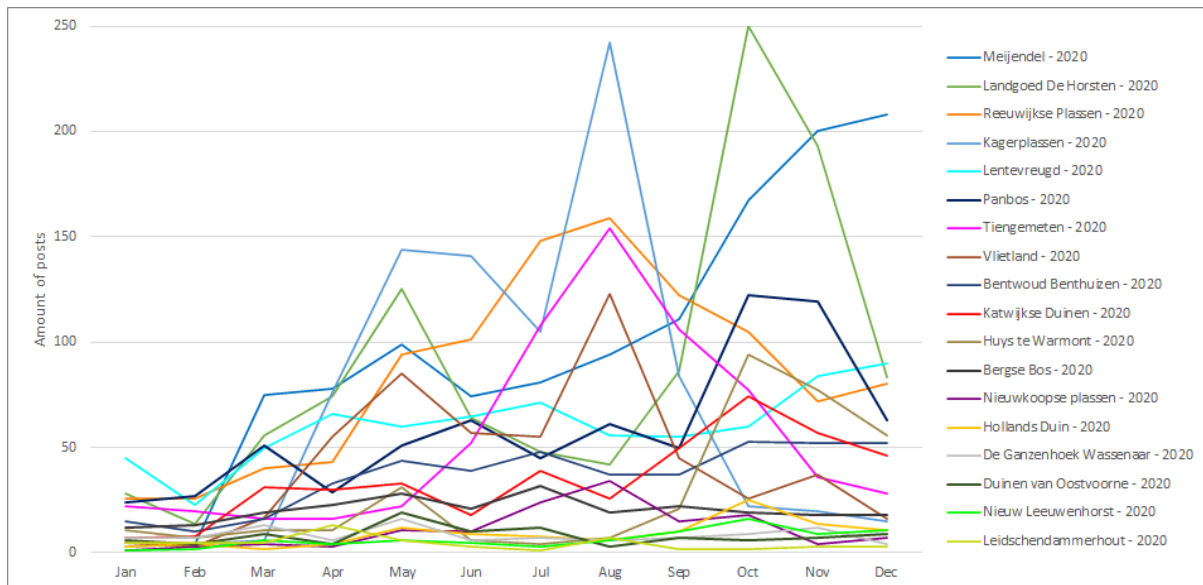


Figure 4: Monthly number of posts for each location from 1 January 2020, to 31 December 2020. Each individual line represents a specific location.

In both years, people tend to visit nature more on the weekends, as shown in Figure 5. These visits also demonstrated the smallest increases, of 46% on Saturdays and 13% on Sundays. In general, larger increases were observed during the weekdays with a range of 45-63% an average of 54%.

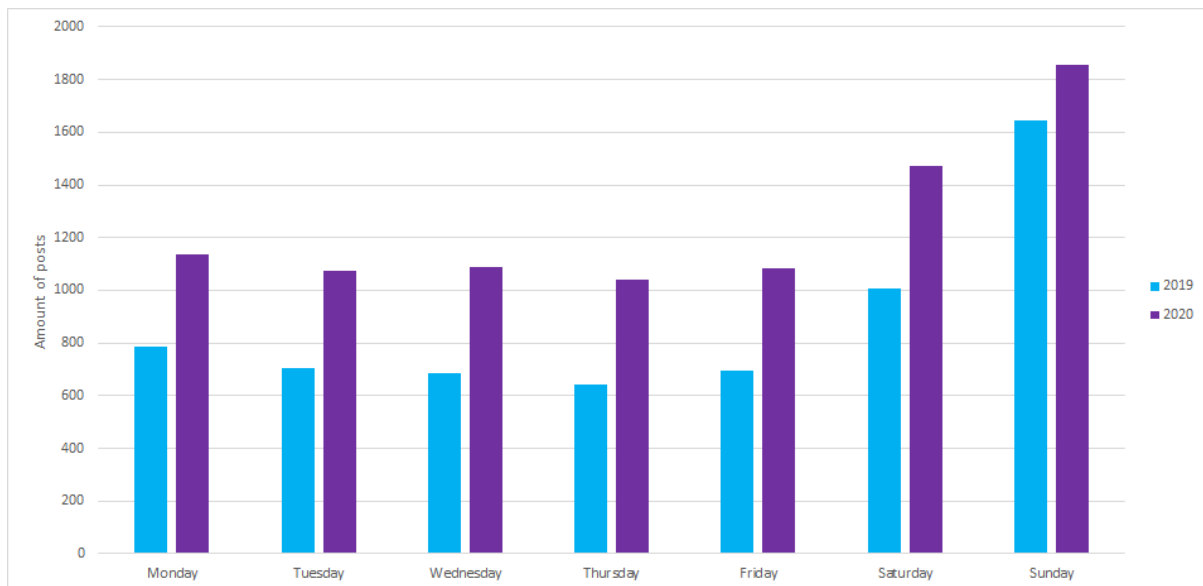


Figure 5: Number of visits per weekday, years 2019 and 2020.

The number of unique Instagram users have been counted to determine if the increase in posts is not just due to the same users posting more photos/videos. The number of unique users in 2019 was 3371 and in 2020 was 4452, which is an increase of 35%.

4 Discussion

The outcomes of this study are in line with the current reports of an increase in visits to nature reserves. The results show an average increase of 41,9% in the number of posts in 2020. We highlighted in figure 2, three different periods of measures imposed by the government due to Covid-19 in 2020 [7]. There seem to be some massive peaks in nature visits in May, August and October 2020, whereas visits in 2019 were somewhat more stable. The largest increase occurred from March to May, probably due to the announcement of the lockdown around that time. In both years there is an increase in the month's May to November, indicating that people tend to visit nature more often in the warmer months. However, in August another massive peak is observed, possibly because people stayed in the Netherlands during the holidays. At the beginning of November, right after lockdown measures have tightened, a strong peak is noticed which was not seen in the previous year.

Some nature reserves, like the forest park De Horsten in autumn, which is normally mostly visited during autumn, were now suddenly being visited in spring. However, the data still reveals in which seasons certain areas are most visited. For example, Reeuwijkse plassen, Kagerplassen, Tiengemeten and Vlietland are all locations with lakes, which explains the larger peak in the summer months.

Visits during the weekdays tended to increase more on weekdays than during the weekends (55% vs 30% on average). Perhaps working from home offers the possibility for a better organization of time schedules. Sunday, which has always been relatively popular, rose slightly by 13%, which could reflect a limit to the capacity of nature reserves to accommodate visitors on popular days. In absolute terms, people still tend to visit nature most during the weekends. One reason for this is that parents work during the week and their children go to school. Another is that people working at home will tend to largely maintain their regular weekday working hours.

Our analysis was based in Instagram posts, and the increase from 2019 to 2020 could therefore also (partly) be due to changes in the number of Instagram users. From December 2019 to December 2020 Instagram users increased from 6.322.000 to 8.130.000 (+28%), which represents 47.1% of the total population in the Netherlands [5][8]. This general growth should be taken into account for the interpretation of the 41.9% increase that we observed in our analysis. Nevertheless, the increased popularity of Instagram does not fully explain the growth of nature reserve posts, and it seems fair to attribute this mostly to larger numbers of visitors.

This study is limited to publicly available posts, as private posts are not shown under the location tag. Therefore, the number of posts is not representative of the actual number of posts. Also, one post is not equal to one visit since a single person can post multiple photos of the same visit. This is partly solved by the possibility to combine multiple photos under a single post. Both functionalities did not change over the years 2019 and 2020, however, so these limitations had little influence on the data's validity.

We have noticed that some users use a hashtag to indicate the location of a photo. This is used in combination with the whole area's location tag, for example, #haagsebos and The Hague. This method could be used to gather more posts, but it also requires more effort due to various spellings of hashtag and location. Another thing we noticed is that some IDs of nature reserves changed over time, which makes it hard to compare different years. An example of this is the location Meijendel; the posts of 2019 are placed under a different ID than those of 2020. This caused a gap in our results; for some reason, there were no posts with either one of the Meijendel ID's in January 2020, and virtually none in February, as see figure 4. At some locations, there were no posts before 2019; therefore, these have been excluded from the analysis as these cannot be compared to 2020.

Even though scraping is an automated process, acquiring the data takes a considerable amount of time. The various names of each location have to be looked up. After that, it must be checked whether there were posts in both 2019 and 2020. Sometimes specific IDs could not be found with the Instagram search function, which meant that Google had to be used. For all these reasons, the selection is limited to Zuid Holland, a province that we are familiar with. Zuid Holland is a densely populated area, so our sample represents a large proportion of Dutch citizens.

In agreement with our results, Natuurmonumenten cites mobility reports from Google which reveal that the number of visits to nature reserves increased by 30% in 2020 [3][4]. This data is not readily available for further exploration, however. Google has a dataset intended to help redress the impact of COVID-19, wherein data shows how visits to places, like grocery stores and parks, are changing in each region. Similar to us, it shows the mobility trend of parks such as national parks, public beaches, marinas, dog parks and public gardens [9]. It does show an increase comparing it with the baseline of 3 January to 6 February 2020. However, the number of visitors also depends on the season, therefore this does not provide information on the influence of the COVID-19 crisis. Our season-based analysis shows clear changes compared to the previous year. Since nature visits are highly dependent on weather circumstances, further exploration of other years could be more reliable than comparison to the single year 2019. Nonetheless, the differences between the two years and the apparent associations with lock down periods, support our view that our results largely reflect the impact of the COVID-19 measures.

In conclusion, Instagram location posts seem to be a reasonable indicator for monitoring crowd behaviour. Furthermore, it confirms the notion that people tend to go more to nature reserves during the pandemic. This may have compensated for some of the loss of activity that was imposed by forced home work and public transport restrictions.

References

1. Rowe, Geoff, et al. "Can We Make Time for Physical Activity? Simulating Effects of Daily Physical Activity on Mortality." *Epidemiology Research International*, Hindawi, 12 July 2012.
2. B. Nussbaumer-Streit, V. Mayr, et al. "Physical Activity during COVID-19 Induced Lockdown: Recommendations." *Journal of Occupational Medicine and Toxicology*, BioMed Central, 12 Aug. 2020.
3. Lee, I-Min, et al. "Effect of Physical Inactivity on Major Non-Communicable Diseases Worldwide: an Analysis of Burden of Disease and Life Expectancy." *Lancet (London, England)*, U.S. National Library of Medicine, 21 July 2012.
4. Koek, Marjolein "Massaal Op Zoek Naar Natuur Tijdens Coronacrisis." *Natuurmonumenten*, 1 Dec. 2020.
5. "Instagram Users in Netherlands - December 2020." *NapoleonCat*, Dec. 2020.
6. Singh, Asmit, and Divyanshu Sharma. "Asmitks/Insta-GraphQL-Scraper." *GitHub*, 27 Dec. 2020.
7. Ministerie van Algemene Zaken. "Dutch Measures against Coronavirus: Basic Rules for Everyone." *Coronavirus COVID-19 | Government.nl*, Ministerie Van Algemene Zaken, 27 Jan. 2021.
8. Department, Published by Statista Research, and Dec 10. "Netherlands: Instagram Users 2018-2020." *Statista*, 10 Dec. 2020.
9. "Mobility Changes Netherlands." *Google COVID-19 Community Mobility Report*, Google LCC, 22 Jan. 2021.