

# Determining an optimal time window for roaming data for tourism statistics

Martijn Tennekes [m.tennekes@cbs.nl](mailto:m.tennekes@cbs.nl)  
 May Offermans [mpw.offerfans@cbs.nl](mailto:mpw.offerfans@cbs.nl)  
 Nico Heerschap [n.heerschap@cbs.nl](mailto:n.heerschap@cbs.nl)

Statistics Netherlands

## Introduction

- Collaboration between Statistics Netherlands, Vodafone and Mezuro.
- Tourism statistics require data about the number of days foreign tourists stay in the Netherlands.
- Currently, a new IMSI is created every day for roaming CDR data, so foreign tourists cannot be traced for longer than one day.
- Suppose this restriction is dropped, what is the preferred time window to renew IMSI numbers? In other words, how long should foreign devices be traced in order to support tourism statistics?

## Data

Aggregated CDR roaming data from 1 to 31 May 2014.

Table 1: **consecutive stays**, i.e. devices that have generated events every day during the stay.

First date	Last date	Number of roaming devices
------------	-----------	---------------------------

Table 2: **non-consecutive stays**, i.e. devices that have generated events on the first and last day of the stay, but not necessarily in between.

First date	Last date	Country	Number of roaming devices
------------	-----------	---------	---------------------------

## Results & Discussion

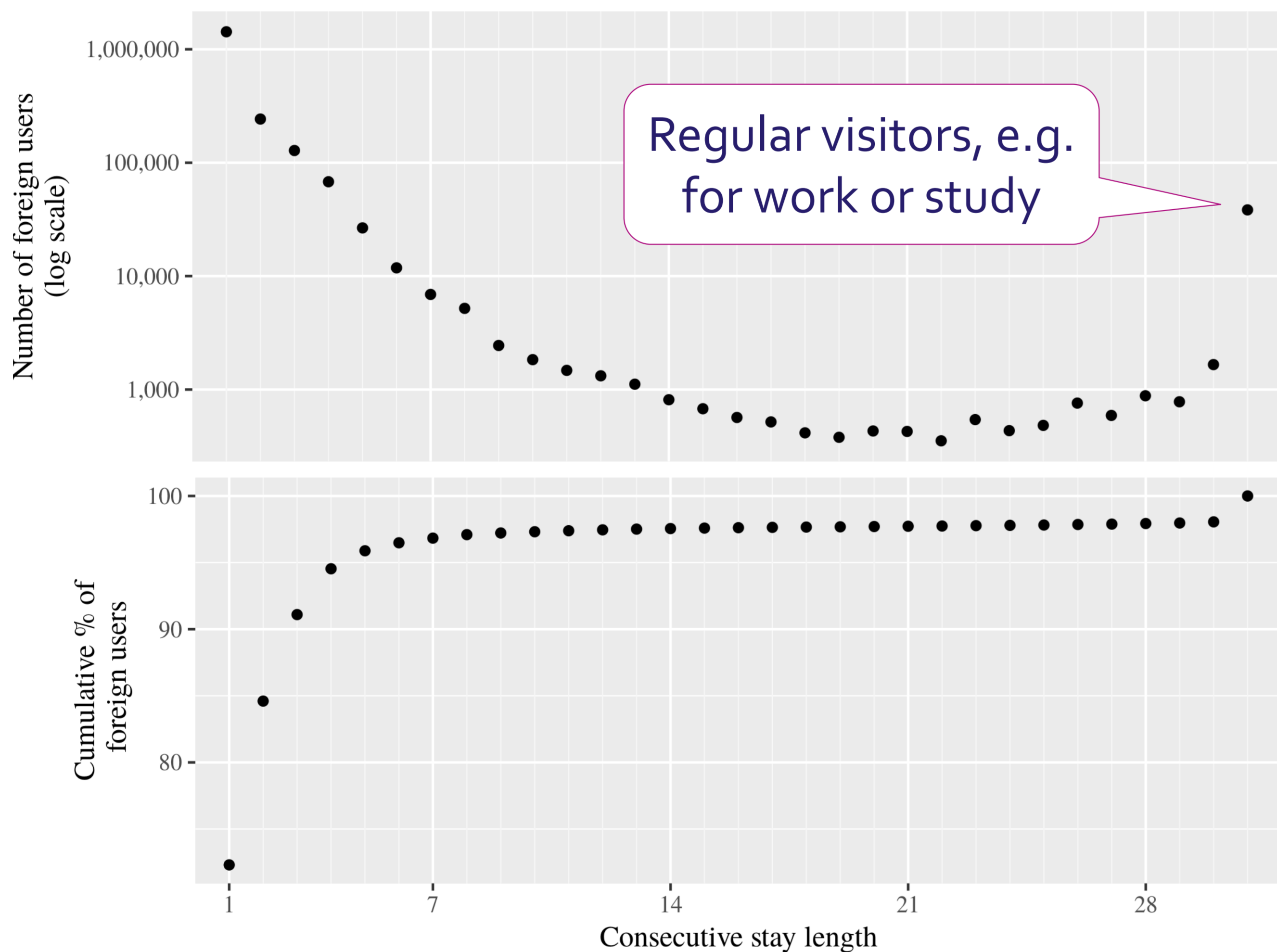


Fig. 1. Number of foreign users per consecutive stay length. Absolute numbers are shown above, cumulative below.

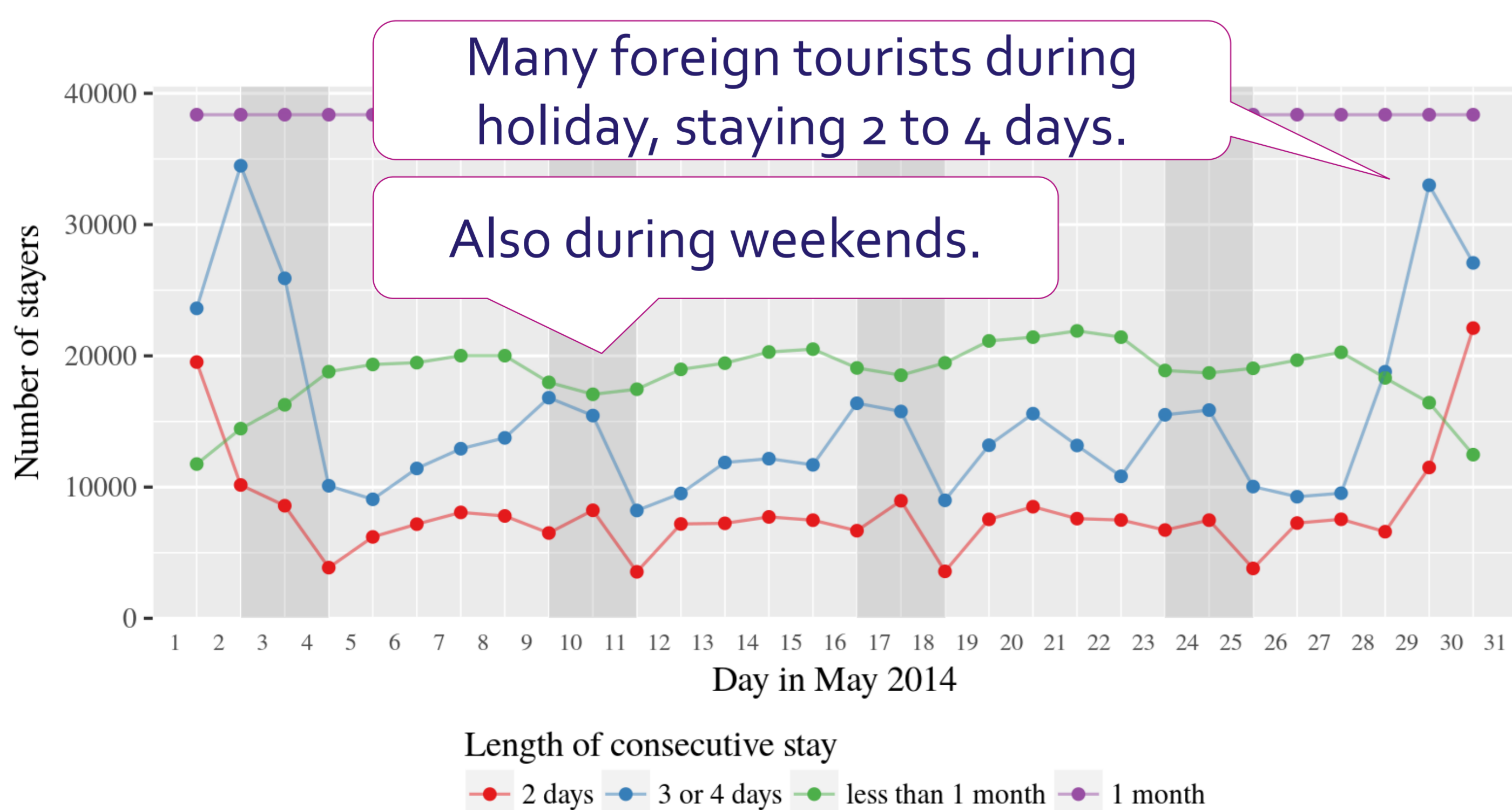


Fig. 2. Consecutive period of foreign users.

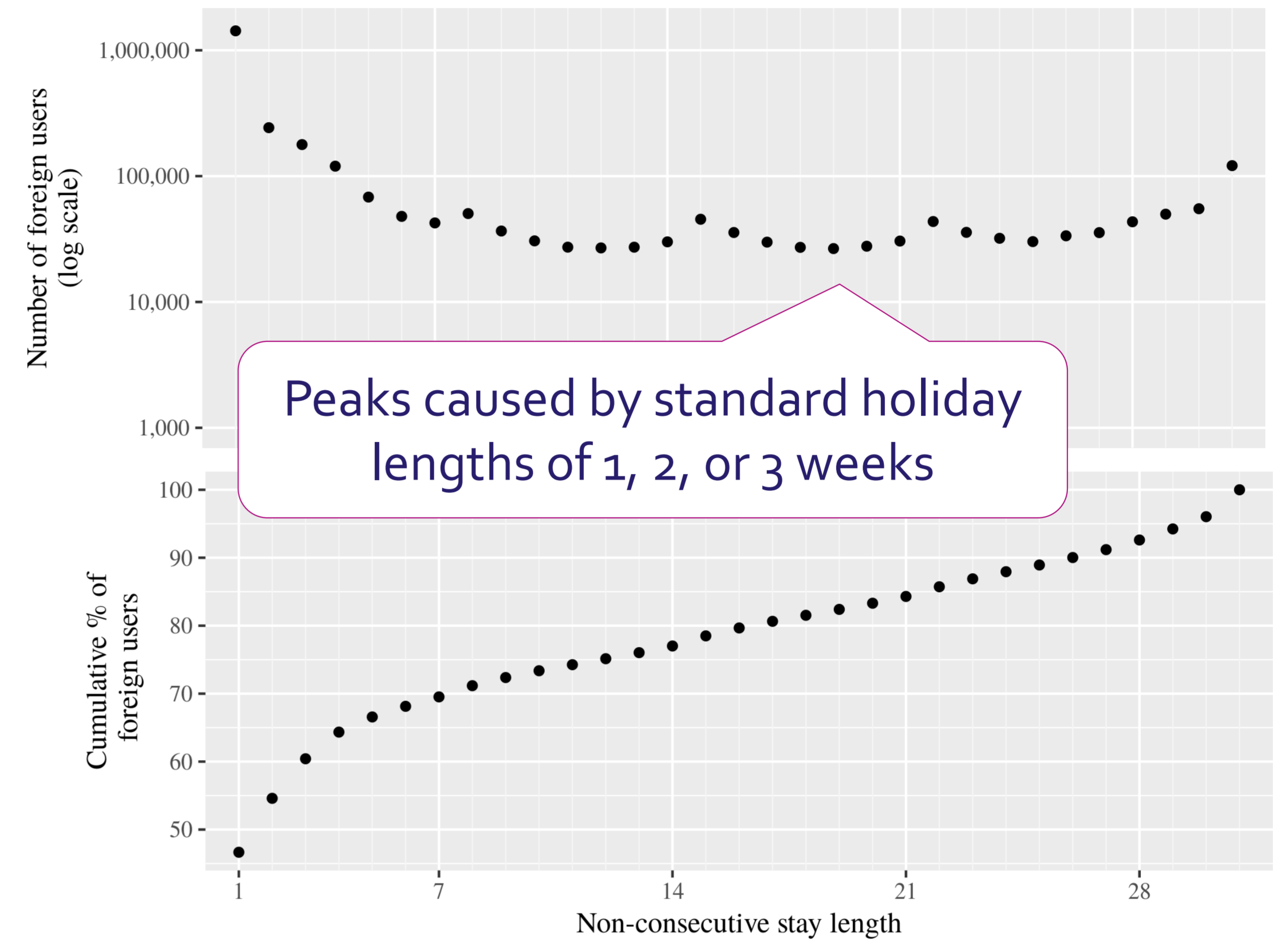


Fig. 3. Number of foreign users per consecutive stay length. Absolute numbers are shown above, cumulative below.

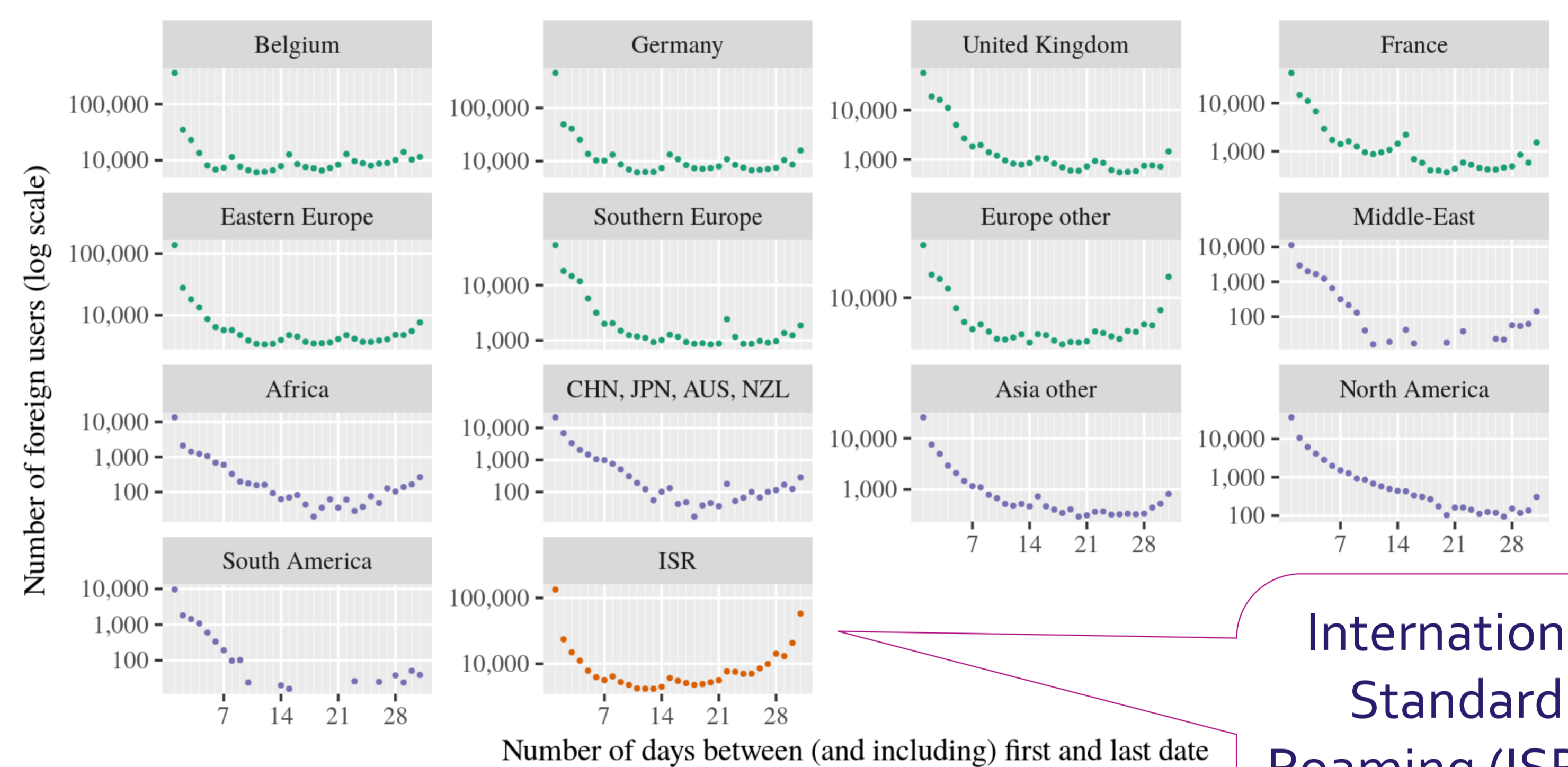


Fig. 4. Consecutive period of foreign users.

Many people from Belgium and Germany.

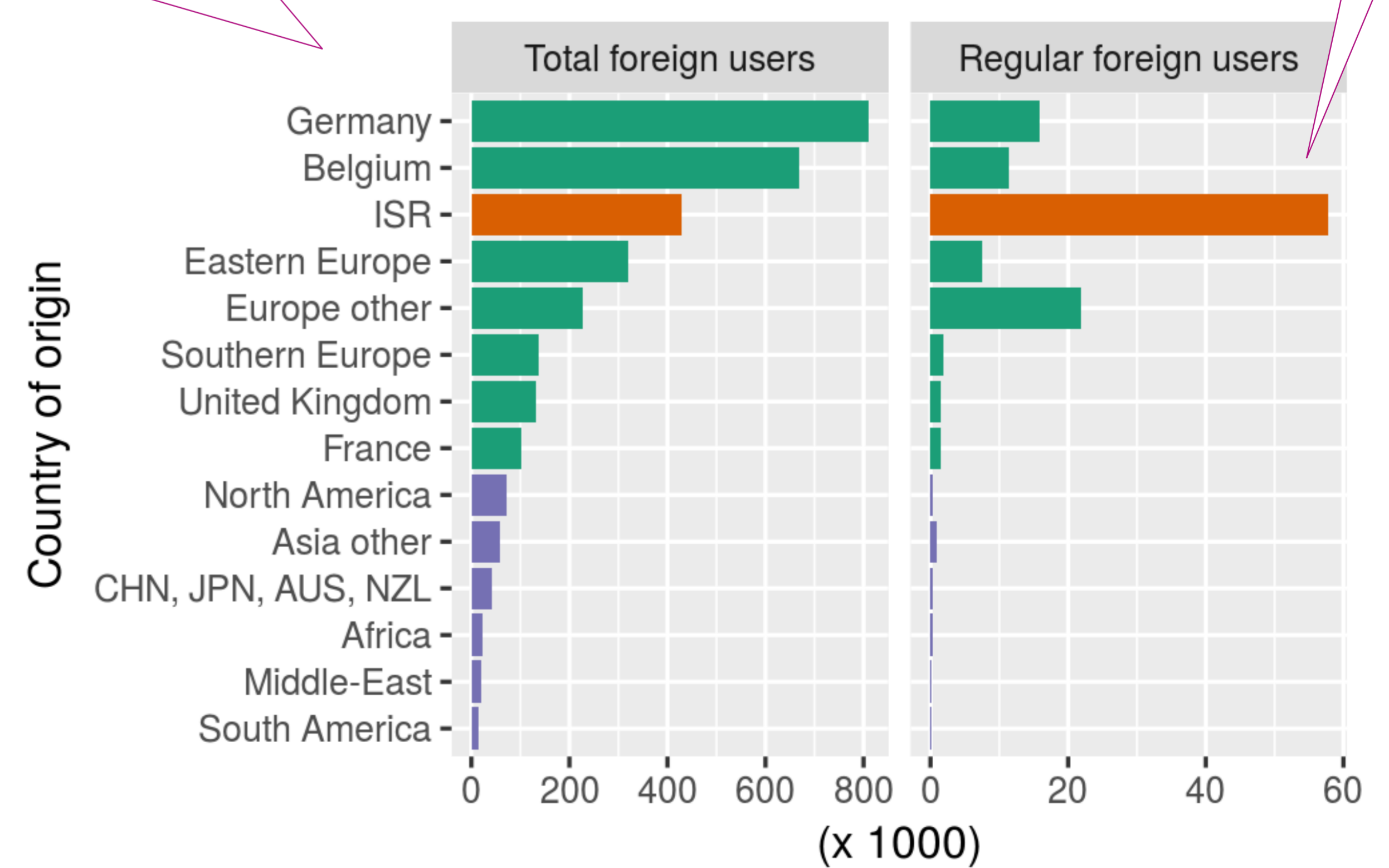


Fig. 5. Consecutive period of foreign users.

## Conclusions

- Time window of 15 days is sufficient to cover 98% of consecutive stays and 80% of non-consecutive stays. Longer is advisable due to overlap.
- However, one month is preferable for tourism statistics in order to compare with current statistics.

