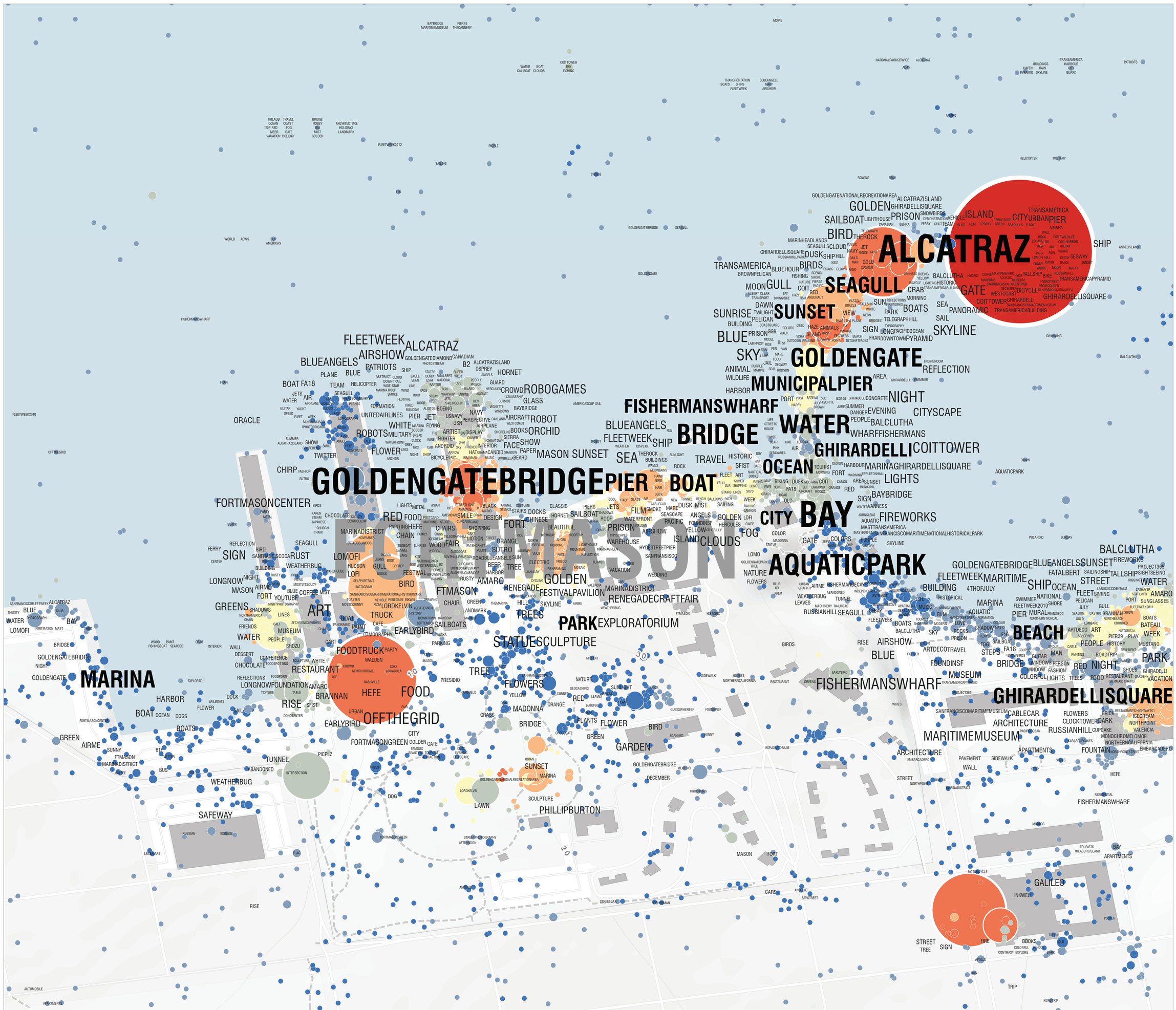
Spatial Pattern-Analysis of Flickr-Photo Distribution

Spatiotemporal Tag Cloud Fort Mason Center, San Francisco



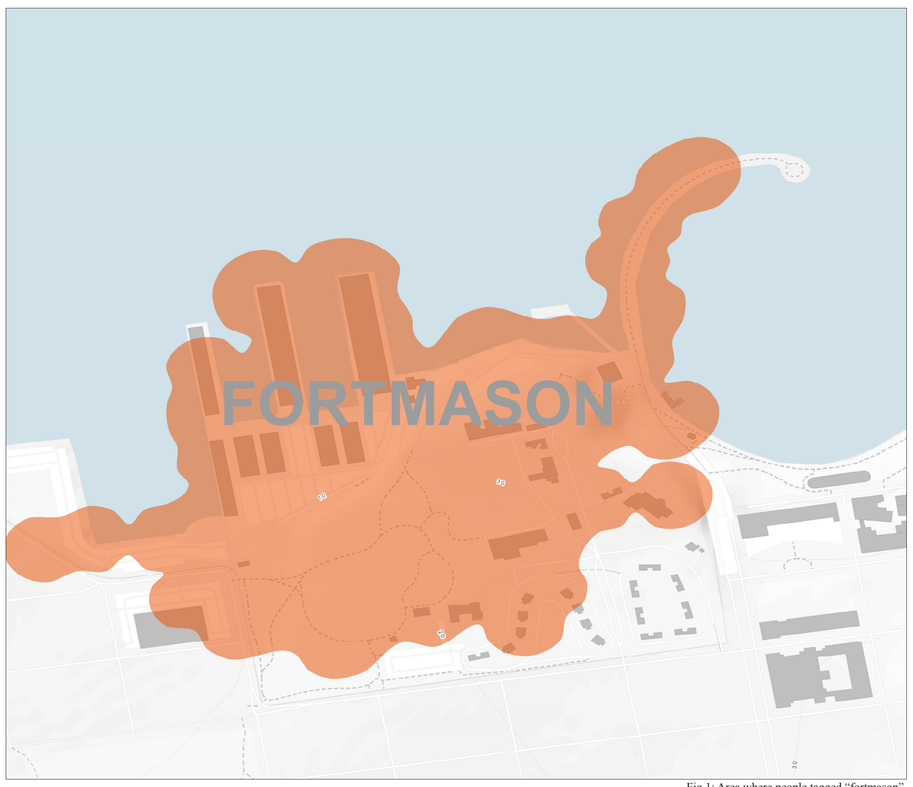
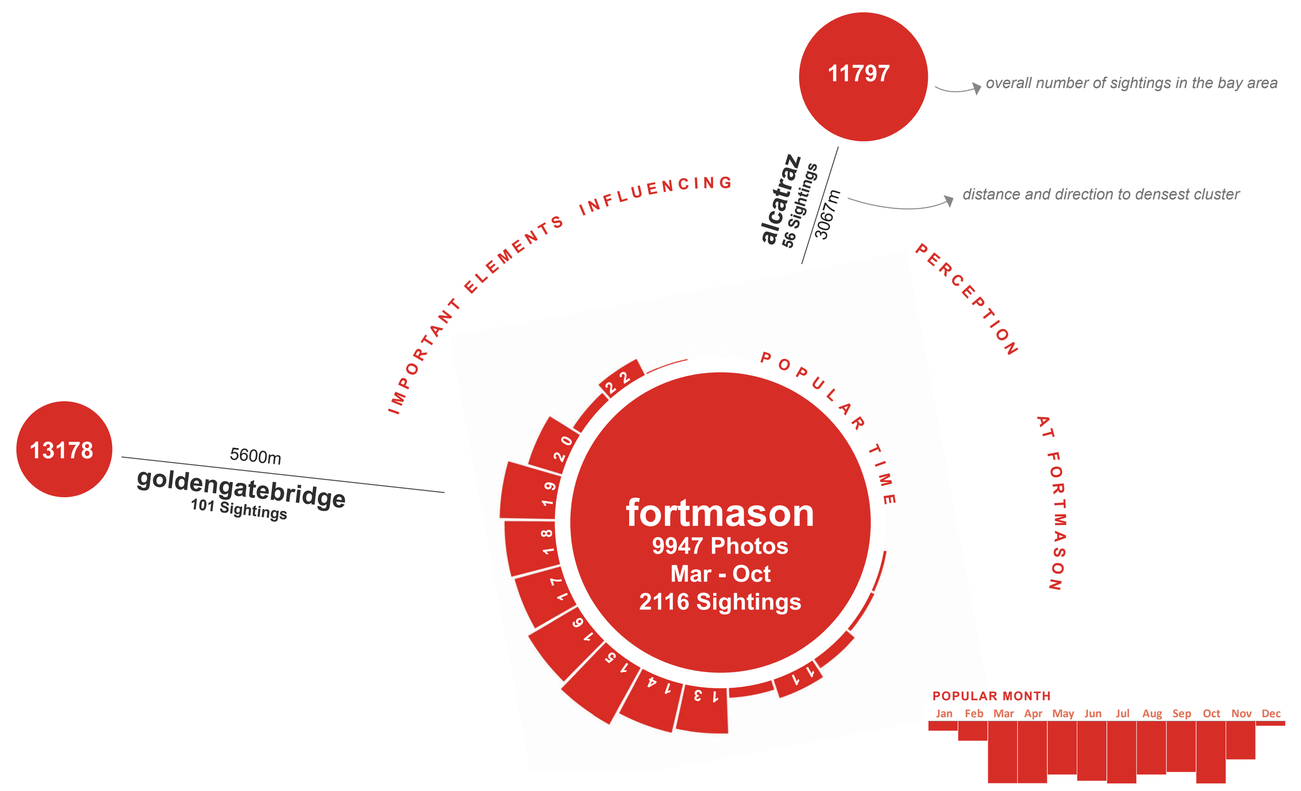


Fig.1: Area where people tagged "fortmason".



Abstract

The goal of this analysis was to evaluate the spatial pattern of geotagged images from Flickr, which is a widely used Photo-Sharing-Community. The result map can be described best as a **statistically weighted map of what is influencing people's perception at certain areas**. In addition to other data, this data may be helpful in many planning processes where it is important to focus on the identification of peoples with their environment and the uniqueness (or missing uniqueness) of each part of the landscape - especially on a bigger scale, where field work and data acquisition is not always possible.

The map consists of two parts. **First**, the spatial pattern of photos is shown as dots of different size and color. The size indicates first level clustering, which shows different locations of interest. The second level of clustering was calculated using the Gettis-Ord GI-Star statistic, which evaluates the data (number of photos taken at certain locations) by comparing the local mean to the global mean and then determining whether the difference between them is statistically significant. The results are displayed in different color variations: red for hot spots where significant more pictures were taken compared to the overall area of investigation and blue for cool spots, which means given the overall area of investigation, these areas do not get as much attention.
Second, the tags for each photo were evaluated to label certain areas. The font size was determined using a formula, which includes the number of occurrences of each tag in a certain area. In addition, the cluster of each tag where the most occurrences appear are written in bold. This way, tags which were only used by a minority of photographers appear small on the map, whereas statistically often used tags appear bigger. The placement for each tag is calculated by determining the arithmetic center of each tag cluster, but in areas where tags accumulate the placement can shift.

It is important to keep two things in mind:
1. This map shows what is statistically influencing the perception. This means that labels may occur at different places compared to where the actual "object" is located (i.e. the Golden Gate Bridge influences people's perception at many places in the Bay Area but exists only once). Usually this means, the cluster of each tag where the most occurrences appear (written in bold) is also the location where the object itself is located.
2. The analysis is based on the data of a limited group of people (the photographers) and a limited period of time (1/1/2006 to 10/01/2011). This means in this particular analysis, 4964 unique photographers took 129.397 photos during this period, which are georeferenced in this area. Ultimately, the data shows only what was or is important to this limited group of people and therefore should be treated with caution and carefull interpretation for use in planning purposes.

Legend

Scale 1:1500

Distribution of Images

Second Level Clustering:
Color
Hot-Spot-Analysis

Z-Score Range	
< -0.80	Significant Cold Spots
-0.80 - -0.40	
-0.39 - 0.0	
0.1 - 2.0	
2.1 - 4.0	Significant Hot Spots
4.1 - 6.0	
> 6.0	

First Level Clustering:
Size
Number of Photos

1
5
10
20
> 50

Distribution of Tags

Font Size:	Number of occurrences:
fortmason	1
fortmason	10
fortmason	50
fortmason	100
fortmas	> 200

Font Weight/ Color:

Light-Grey Color/ Shown in Background for most used Tags: fortmason

Bold font weight for the densest area for each Tag used: **fortmason**

Normal font weight and dark-grey Font Color for all other values: fortmason

Please note: because of accumulation of tags in certain areas, only a fraction of available tags can be shown on the map above.

2012/11/14, Alexander Dunkel