

Venice through the Lens of Instagram: A Visual Narrative of Tourism in Venice

Luca Rossi¹, Eric Boscaro², Andrea Torsello²

1. Aston University, United Kingdom

2. Università Ca' Foscari Venezia, Italy



The 8th International Workshop on Location and the Web, 24 April 2018, Lyon, France



- The tourism industry has boasted virtually uninterrupted growth over time
- International tourist arrivals have increased from 25 million globally in 1950 to 278 million in 1980, 674 million in 2000, and **1,235 million in 2016**

Source: <https://www.e-unwto.org/doi/pdf/10.18111/9789284419029>



- Tourists make an increasing use of photo-sharing social media like Instagram and Flickr to share their experiences online
- Geotagged data provides a rich source of information to study tourism consumption

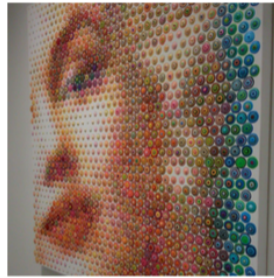


- The city of Venice (Italy) provides an interesting case study, being one of the most popular destinations in one of the world most visited countries

Source: <http://blog.euromonitor.com/2016/01/top-100-city-destinations-ranking-2016.html>

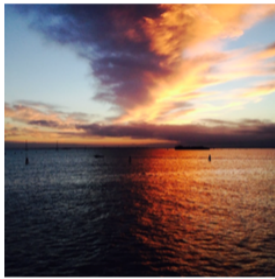
Dataset

- We retrieve **90,000** geotagged Instagram photos taken in Venice from **Jan 2014 to Dec 2015**
- We group these images into **6 categories**



Dataset

- We retrieve **90,000** geotagged Instagram photos taken in Venice from **Jan 2014 to Dec 2015**
- We group these images into **6 categories**

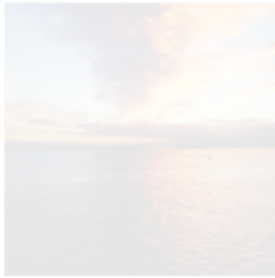


Lagoon



Dataset

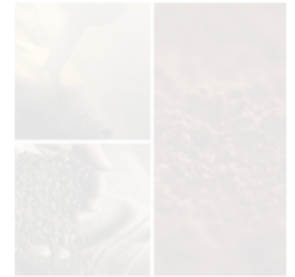
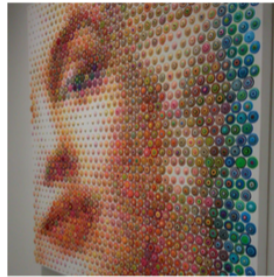
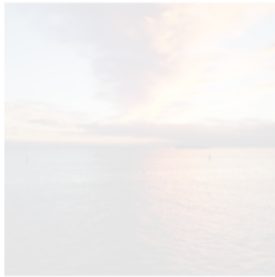
- We retrieve **90,000** geotagged Instagram photos taken in Venice from **Jan 2014 to Dec 2015**
- We group these images into **6 categories**



Townscape

Dataset

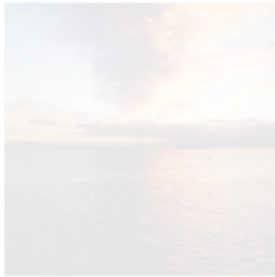
- We retrieve **90,000** geotagged Instagram photos taken in Venice from **Jan 2014 to Dec 2015**
- We group these images into **6 categories**



Art

Dataset

- We retrieve **90,000** geotagged Instagram photos taken in Venice from **Jan 2014 to Dec 2015**
- We group these images into **6 categories**

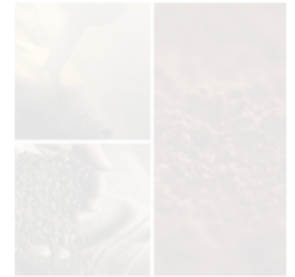
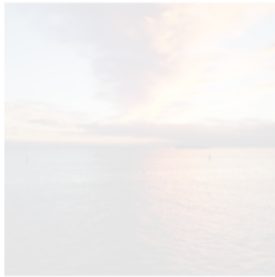


Folklore



Dataset

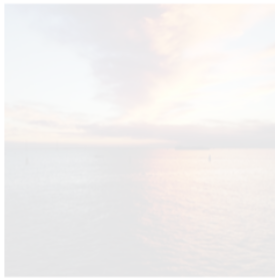
- We retrieve **90,000** geotagged Instagram photos taken in Venice from **Jan 2014 to Dec 2015**
- We group these images into **6 categories**



Food

Dataset

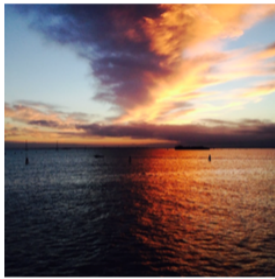
- We retrieve **90,000** geotagged Instagram photos taken in Venice from **Jan 2014 to Dec 2015**
- We group these images into **6 categories**



Other

Dataset

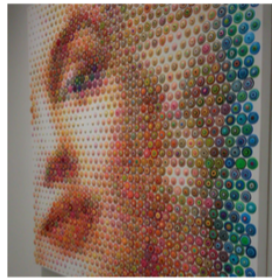
- We retrieve **90,000** geotagged Instagram photos taken in Venice from **Jan 2014 to Dec 2015**
- We group these images into **6 categories**



Lagoon



Townscape



Art



Folklore



Food



Other

Image classification framework

- We create a training set of 600 manually annotated images, 100 per class
- With this training set, we classify the remaining images using a combination of **SIFT** features, **BOW** representations and **SVM** classifiers



Image classification framework

- We create a training set of 600 manually annotated images, 100 per class
- With this training set, we classify the remaining images using a combination of **SIFT** features, **BOW** representations and **SVM** classifiers

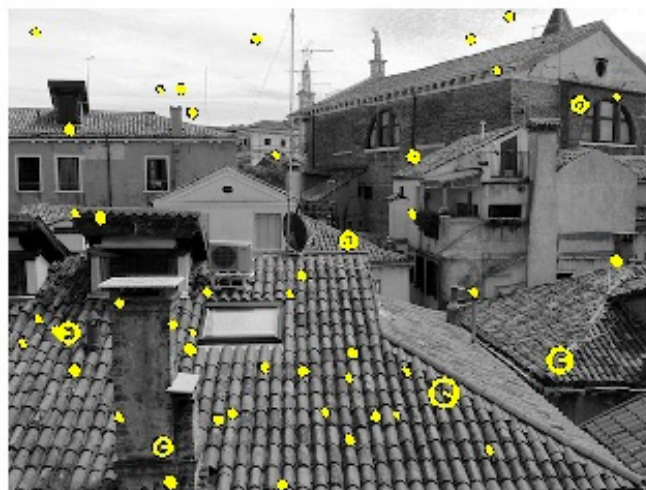


Image classification framework

- We create a training set of 600 manually annotated images, 100 per class
- With this training set, we classify the remaining images using a combination of **SIFT** features, **BOW** representations and **SVM** classifiers

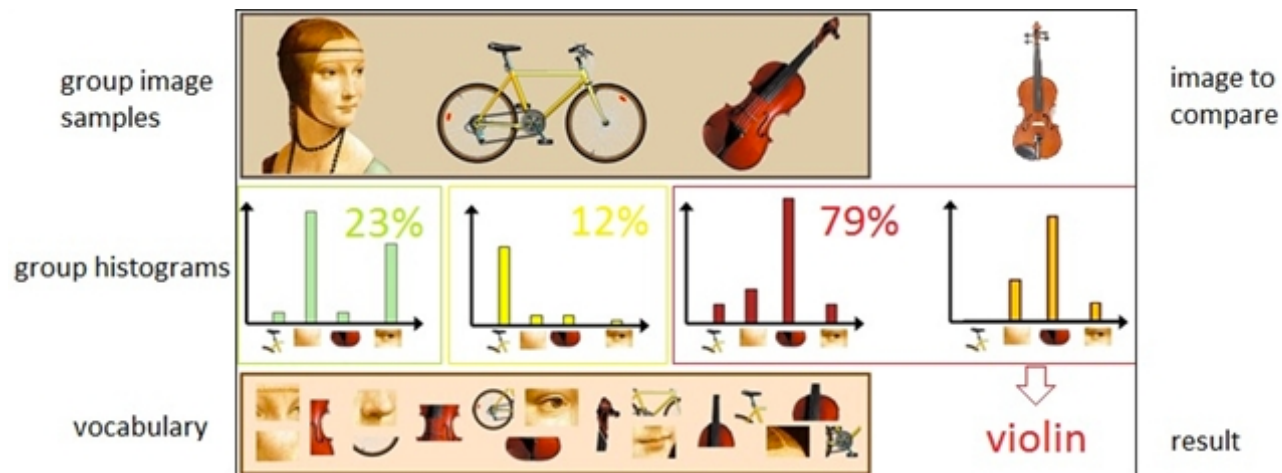


Image classification framework

- We create a training set of 600 manually annotated images, 100 per class
- With this training set, we classify the remaining images using a combination of **SIFT** features, **BOW** representations and **SVM** classifiers
- We perform 5-fold cross validation to compute the average classification accuracy of the classifier (68%)

1st classifier: confusion matrix

- Low misclassification rate for every class except *Other*
 - Over 50% misclassification rate for this class!

Lagoon	73	11	2	6	0	8
Townscape	11	72	2	8	0	7
Art	1	12	71	4	1	11
Folklore	0	10	11	64	7	8
Food	0	2	5	2	86	5
Varies	3	15	12	16	12	42
	Lagoon	Townscape	Art	Folklore	Food	Varies

1st classifier: confusion matrix

- We can reduce this issue by artificially increasing the probability of assigning an image to *Other*
 - Multiply prob of assigning to *Other* by **M**

Lagoon	73	11	2	6	0	8
Townscape	11	72	2	8	0	7
Art	1	12	71	4	1	11
Folklore	0	10	11	64	7	8
Food	0	2	5	2	86	5
Varies	3	15	12	16	12	42
	Lagoon	Townscape	Art	Folklore	Food	Varies

2nd classifier: confusion matrix

- **M** is optimised through 5-fold cross-validation on the training set
- Optimal value is found to be **M=2**

Lagoon	68	5	1	4	2	20
Townscape	9	68	1	7	0	15
Art	0	7	65	0	0	28
Folklore	0	8	8	55	5	24
Food	0	0	4	0	77	19
Varies	0	9	4	2	6	79
Lagoon						
Townscape						
Art						
Folklore						
Food						
Varies						

Misclassified photos: example



Other
(wrong)



Townscape
(wrong)

Misclassified photos: example

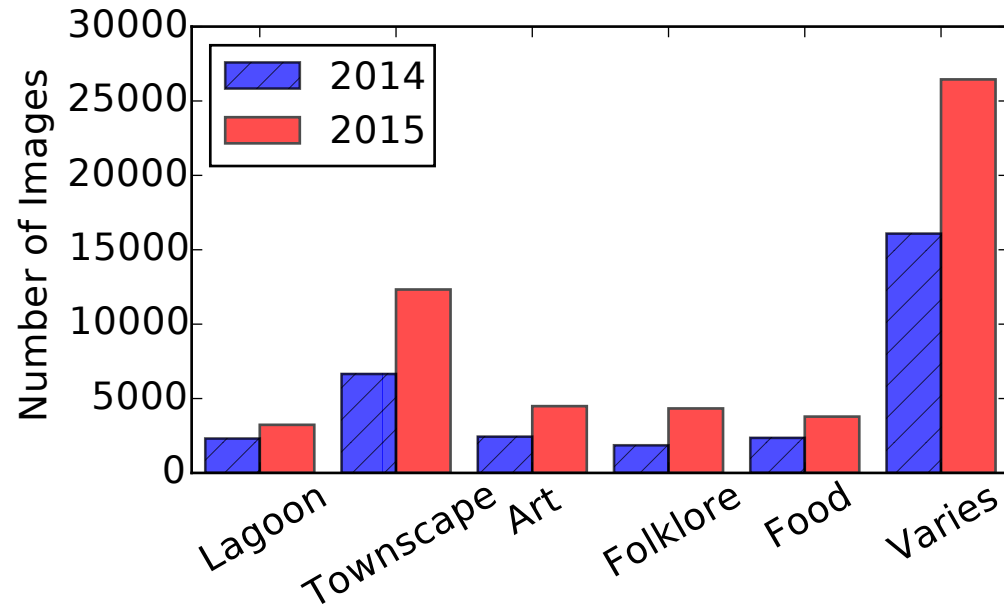


Townscape
(correct)

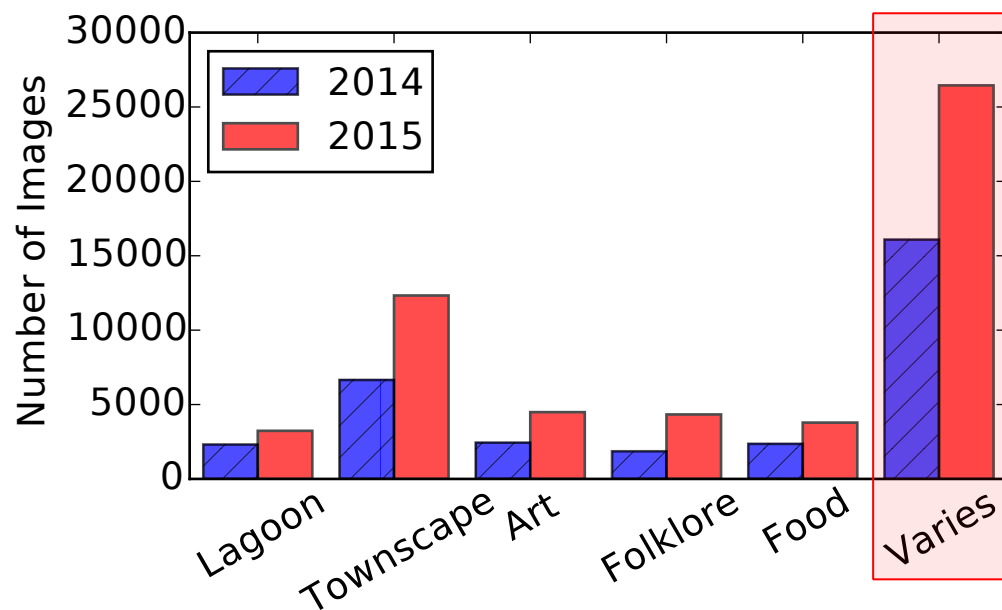


Folklore
(correct)

Categories distribution

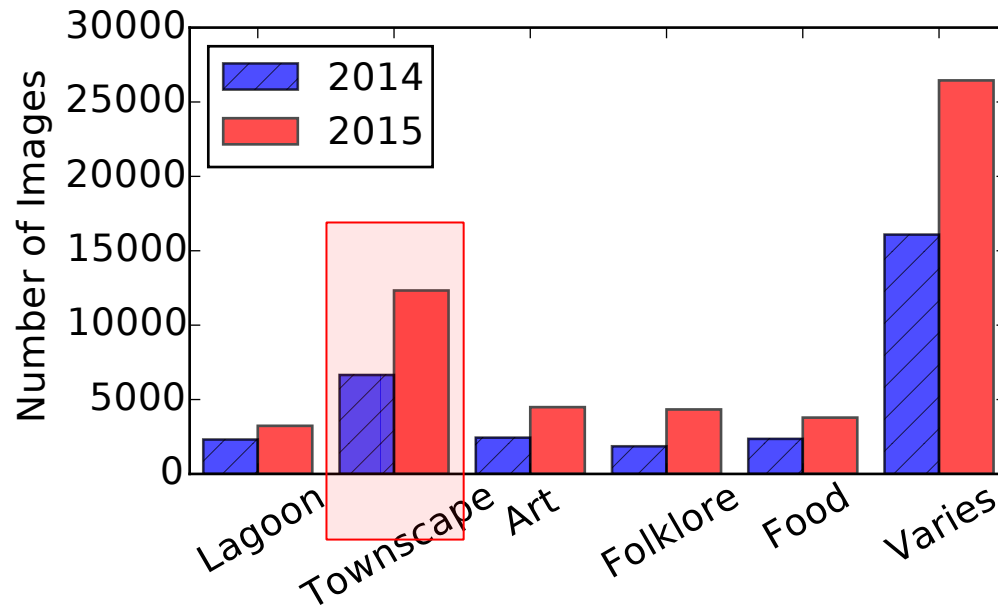


Categories distribution



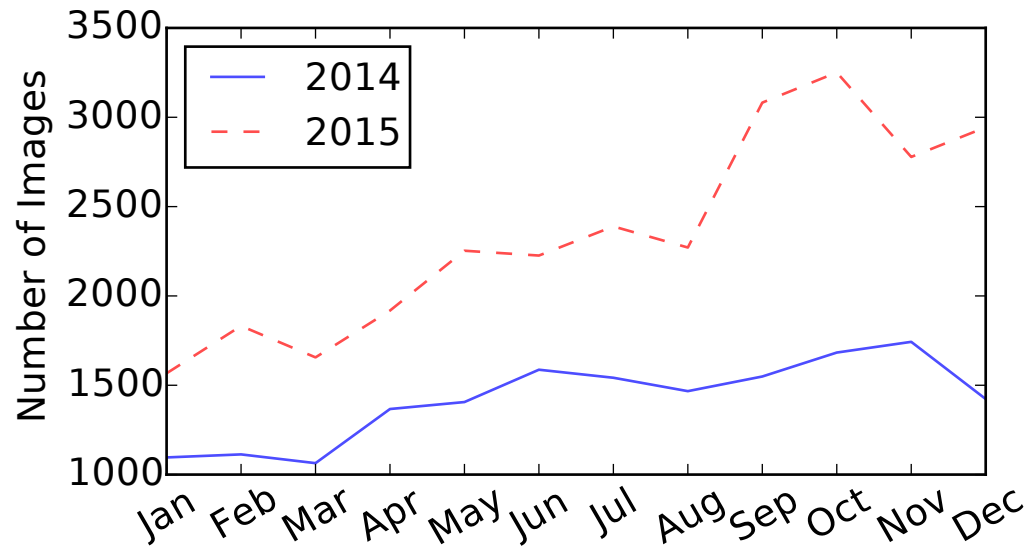
In both years, about 50% of the images are assigned to *Other*, for a total of 44k out of 90k images

Categories distribution



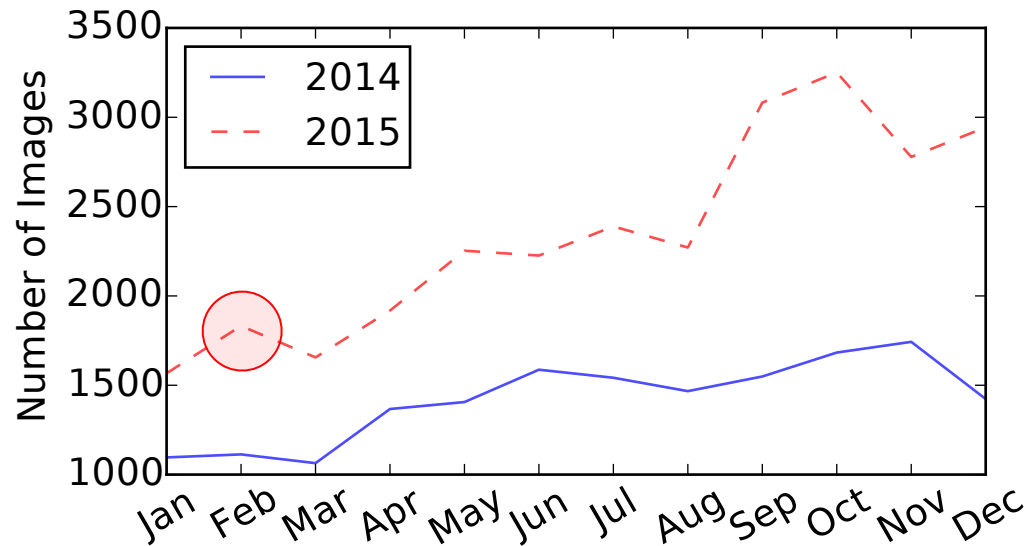
25% of the photos are in the *Townscape* category, which comprises architectural elements such as bridges, churches, squares, highlighting the rich architectural heritage of Venice

Number of photos taken over time



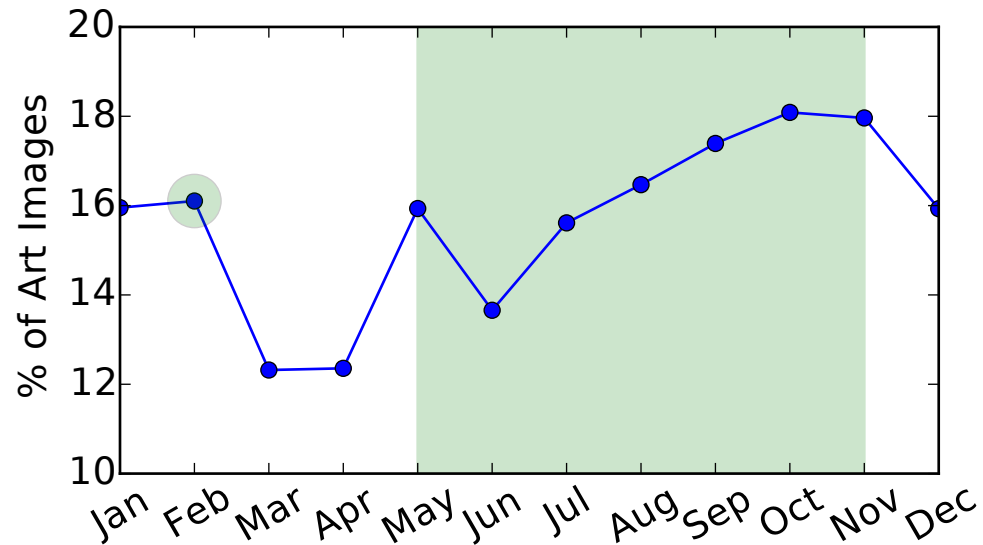
The trend is in line with the worldwide growth in Instagram active users, which have more than doubled from the beginning of 2014 to the end of 2015

Number of photos taken over time



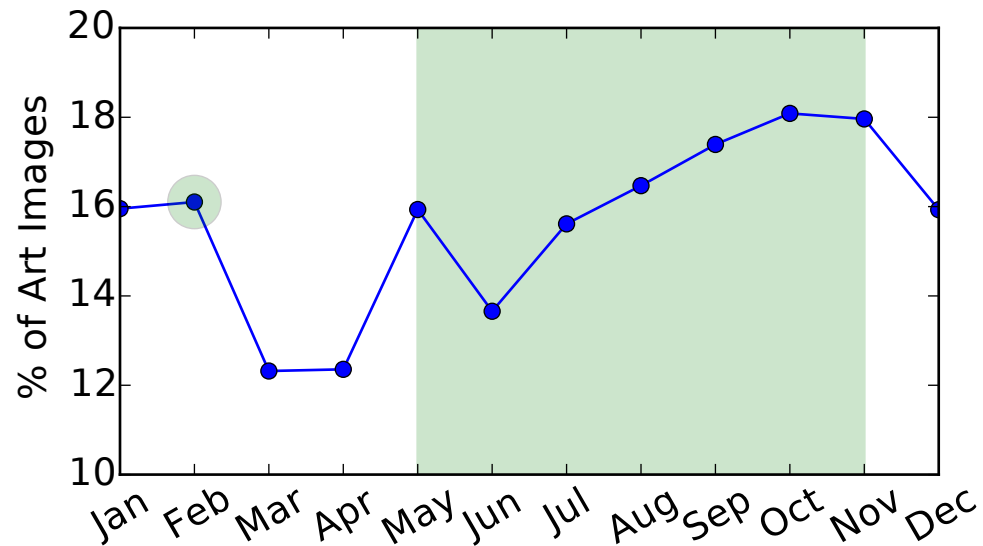
Peak corresponding to Carnival 2015

Frequency of *Art* photos in 2015

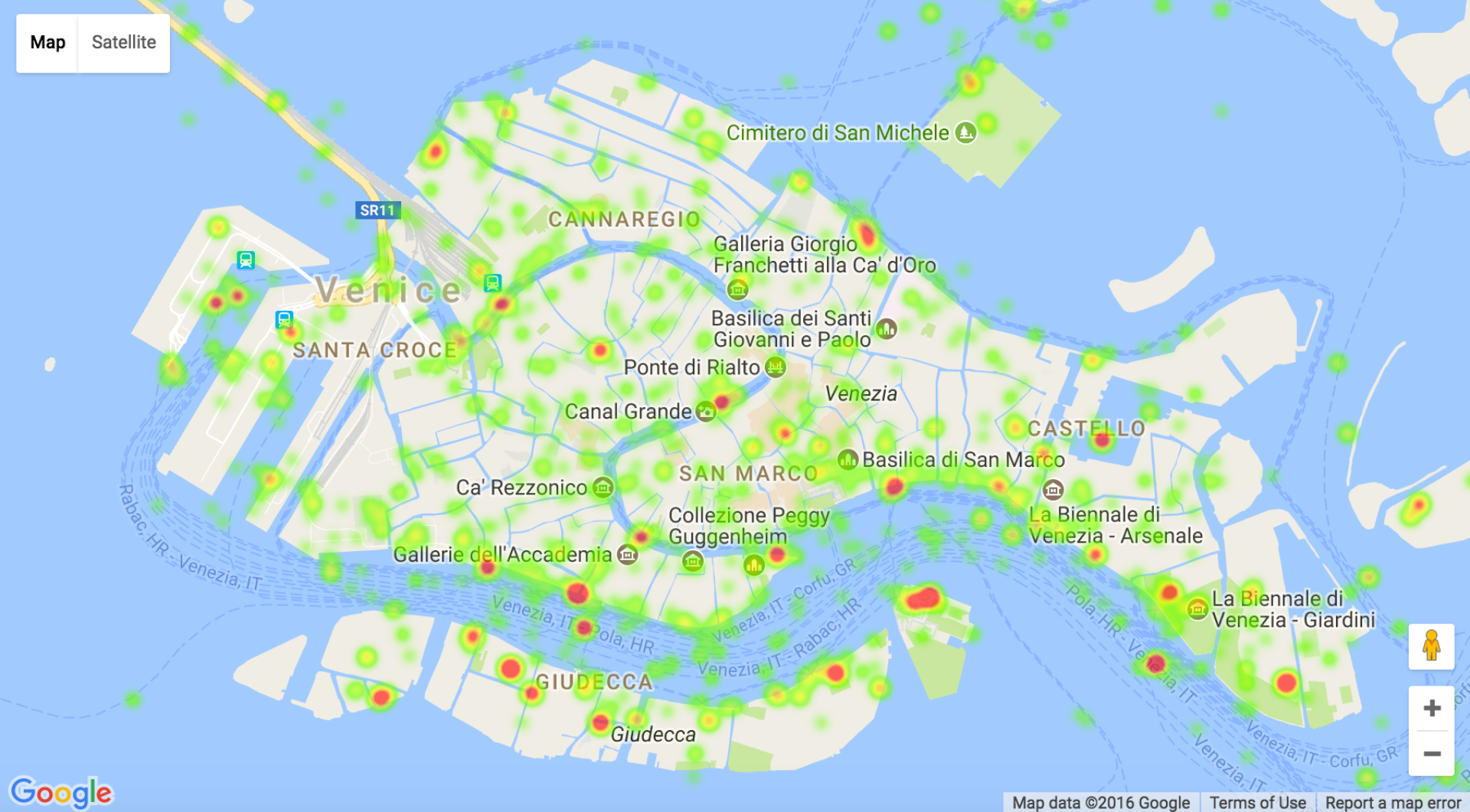


The shaded area shows the increase in the **frequency** of *Art* photos during the 56th Art Biennale

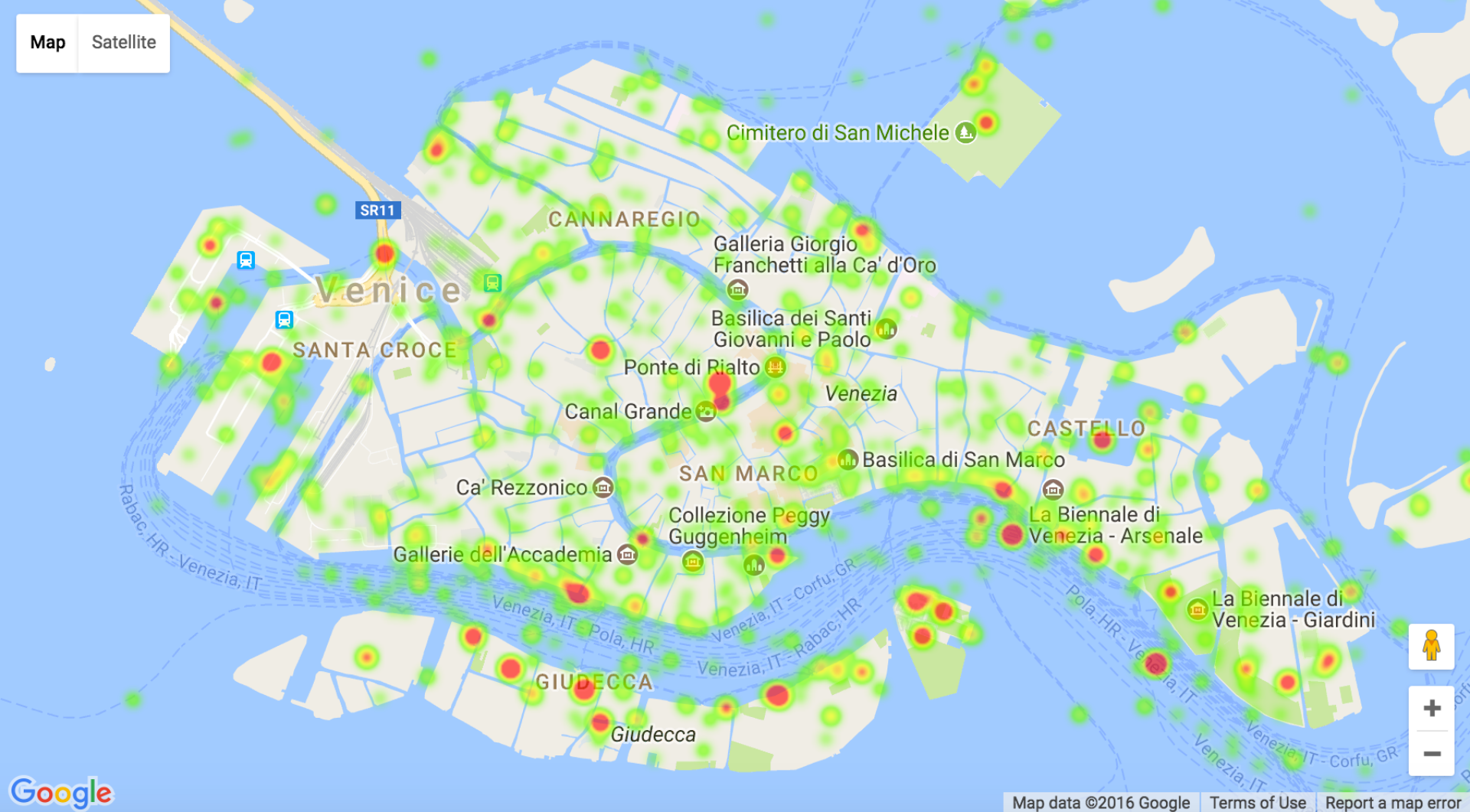
Frequency of *Art* photos in 2015



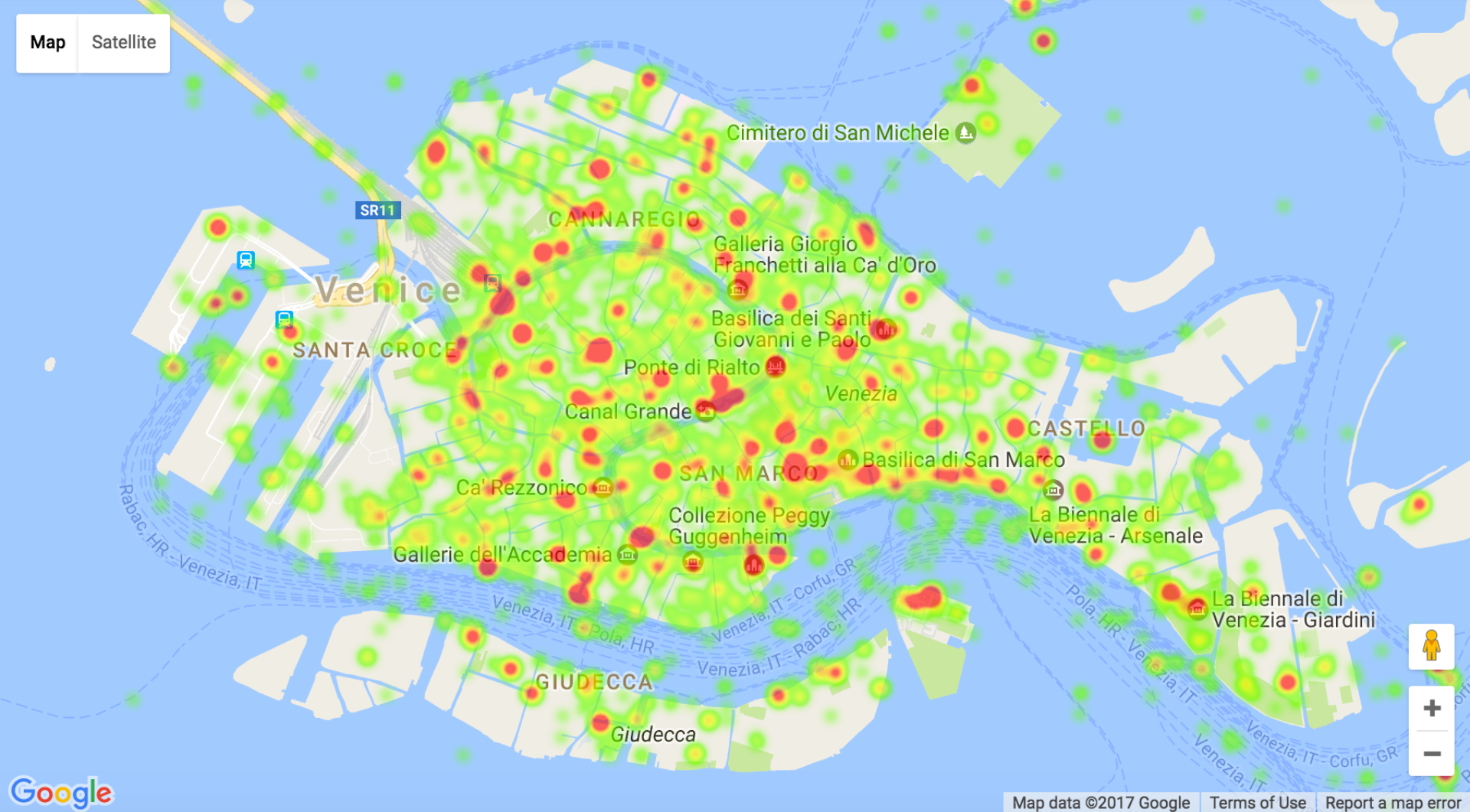
Note also the slight increase during the Carnival period.
This may be due to the increased number of cultural events organised in museums and galleries during that period



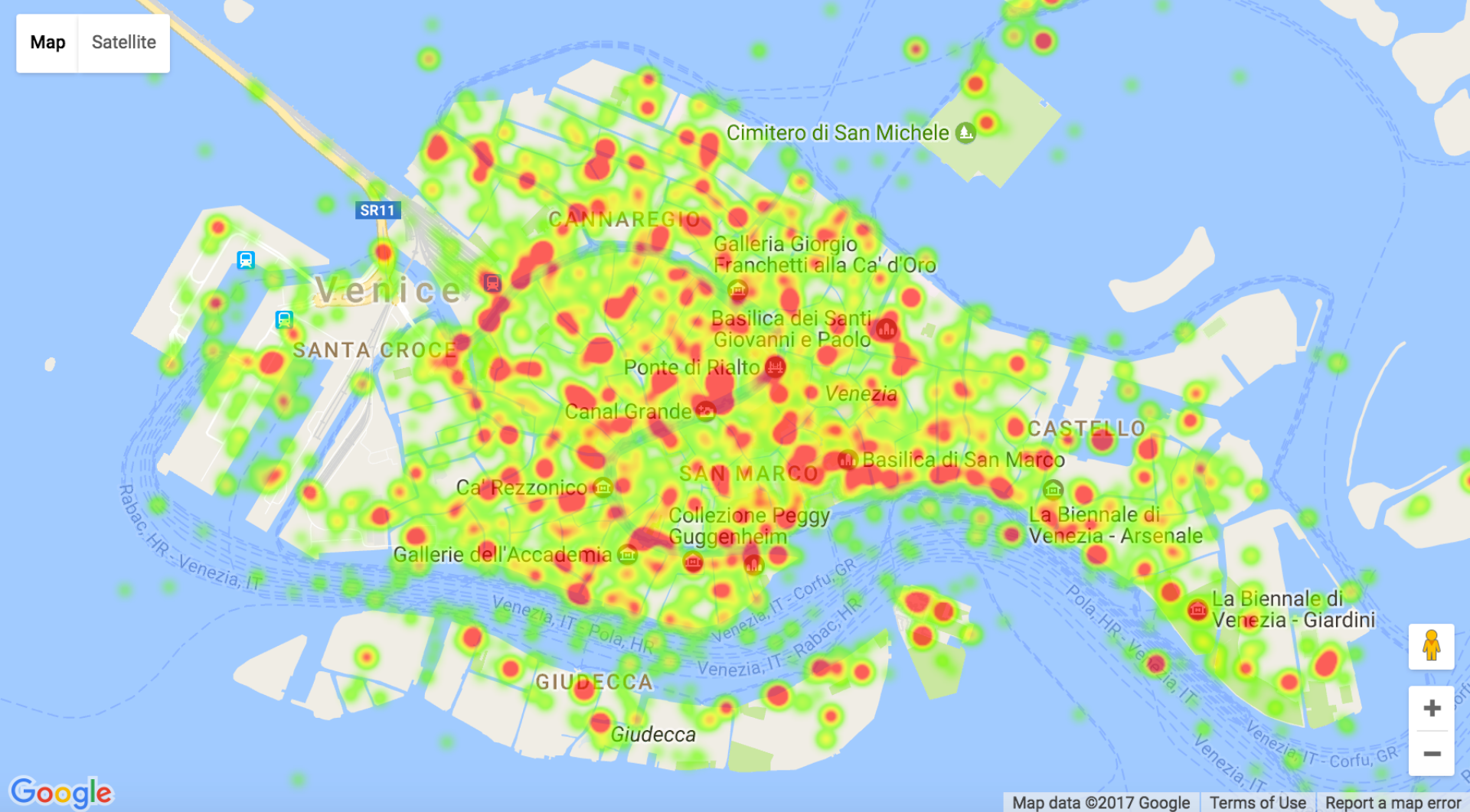
Heatmap: Lagoon 2014



Heatmap: Lagoon 2015



Heatmap: Townscape 2014



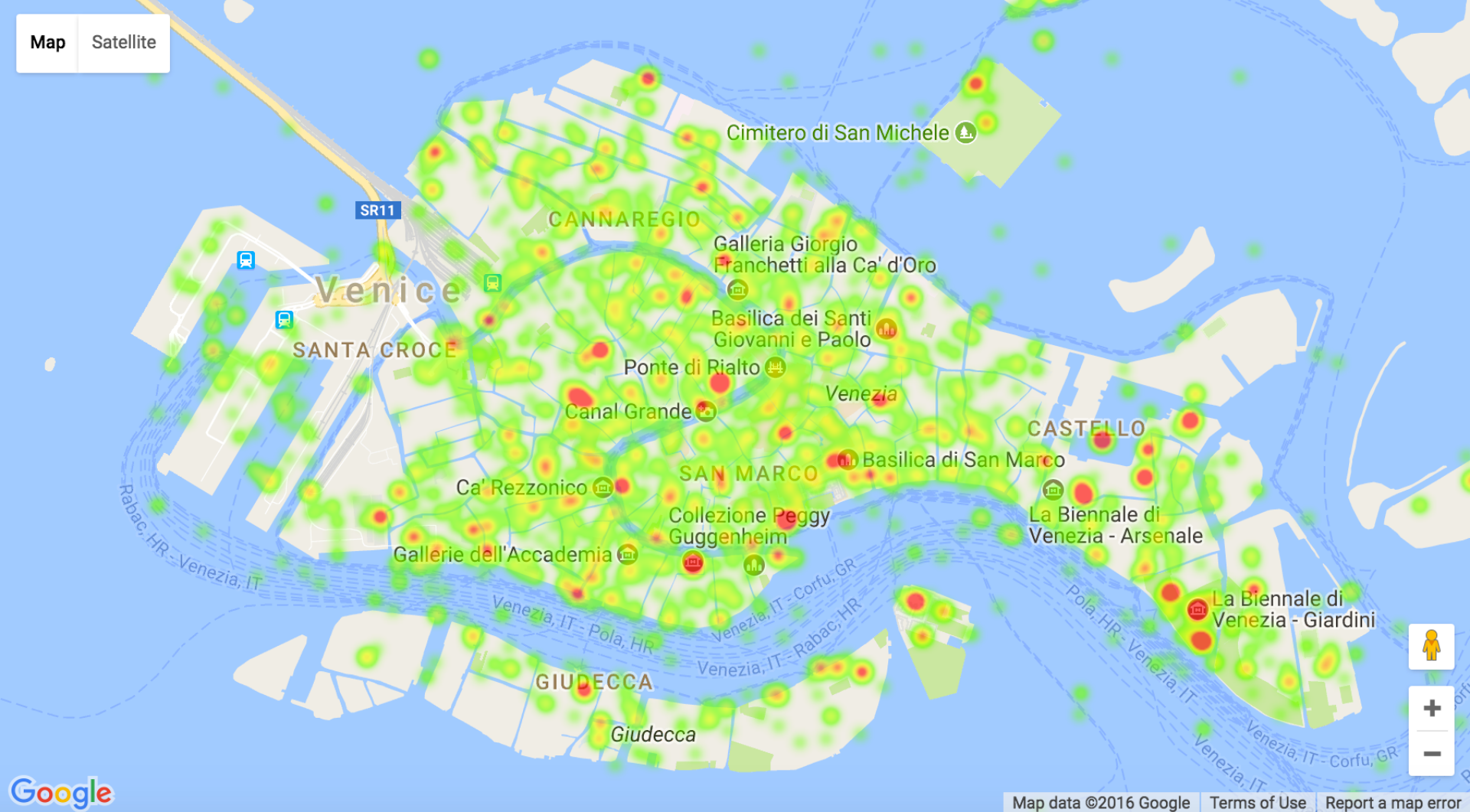
Heatmap: Townscape 2015



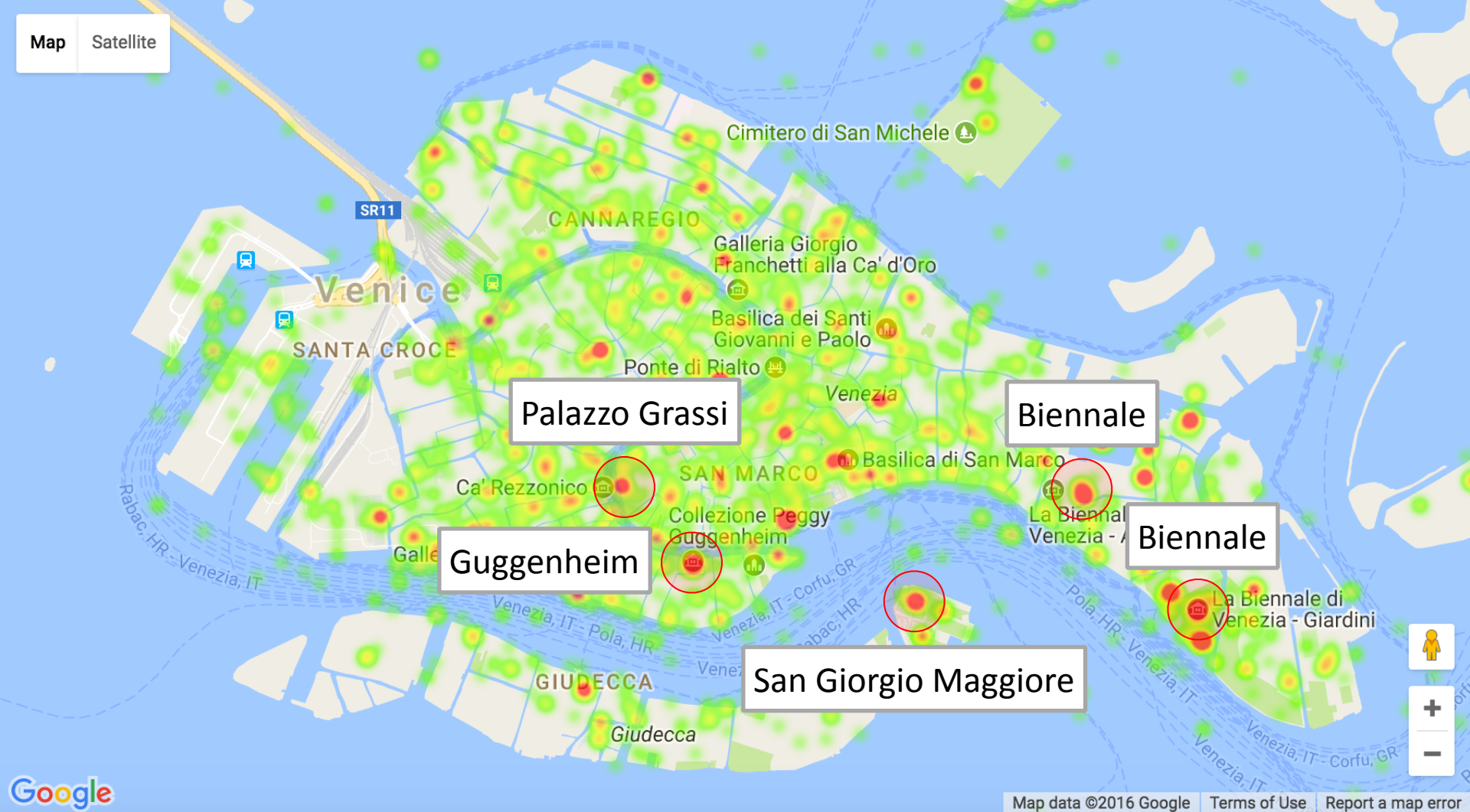
Heatmap: Art 2014



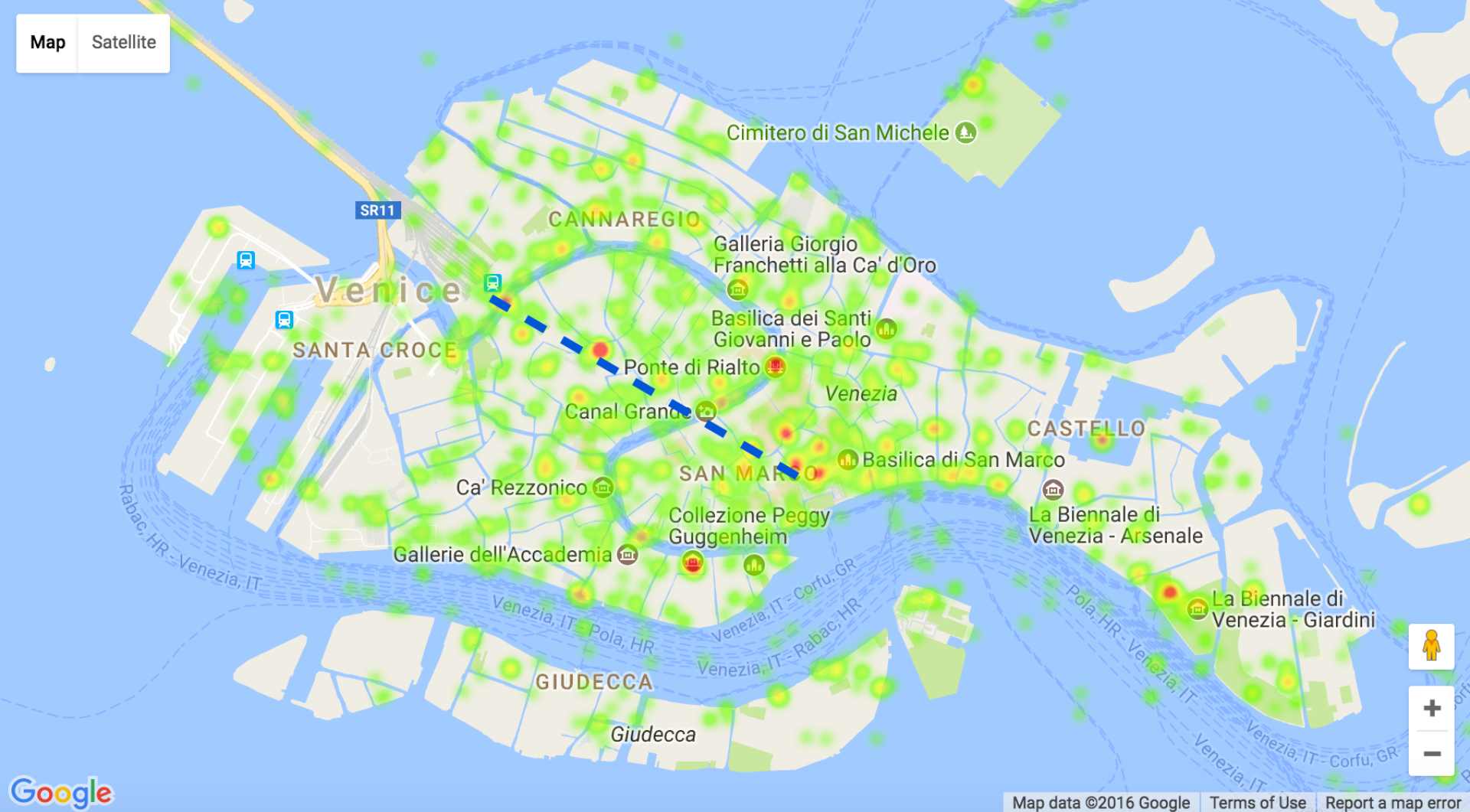
Heatmap: Art 2014



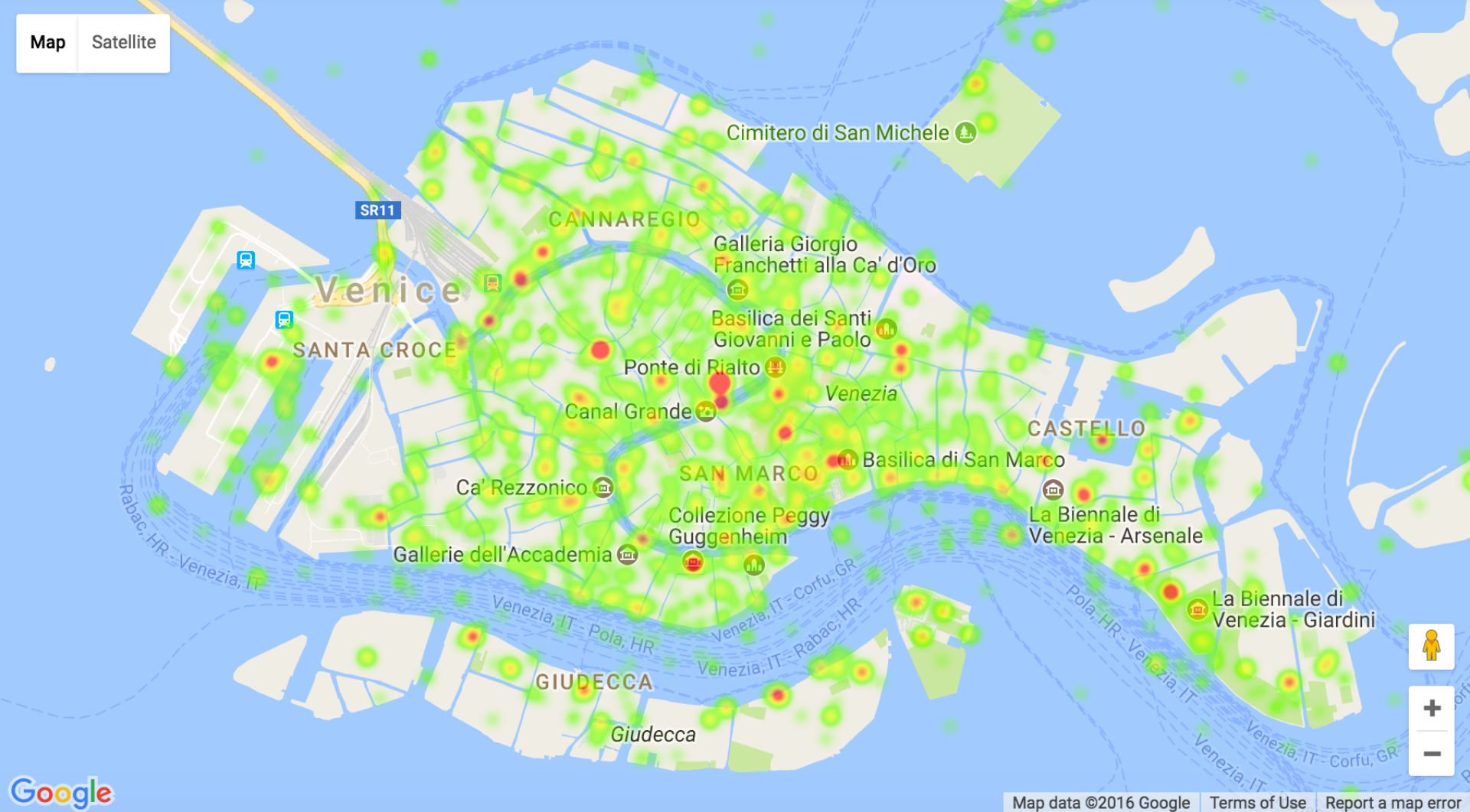
Heatmap: Art 2015



Heatmap: Art 2015



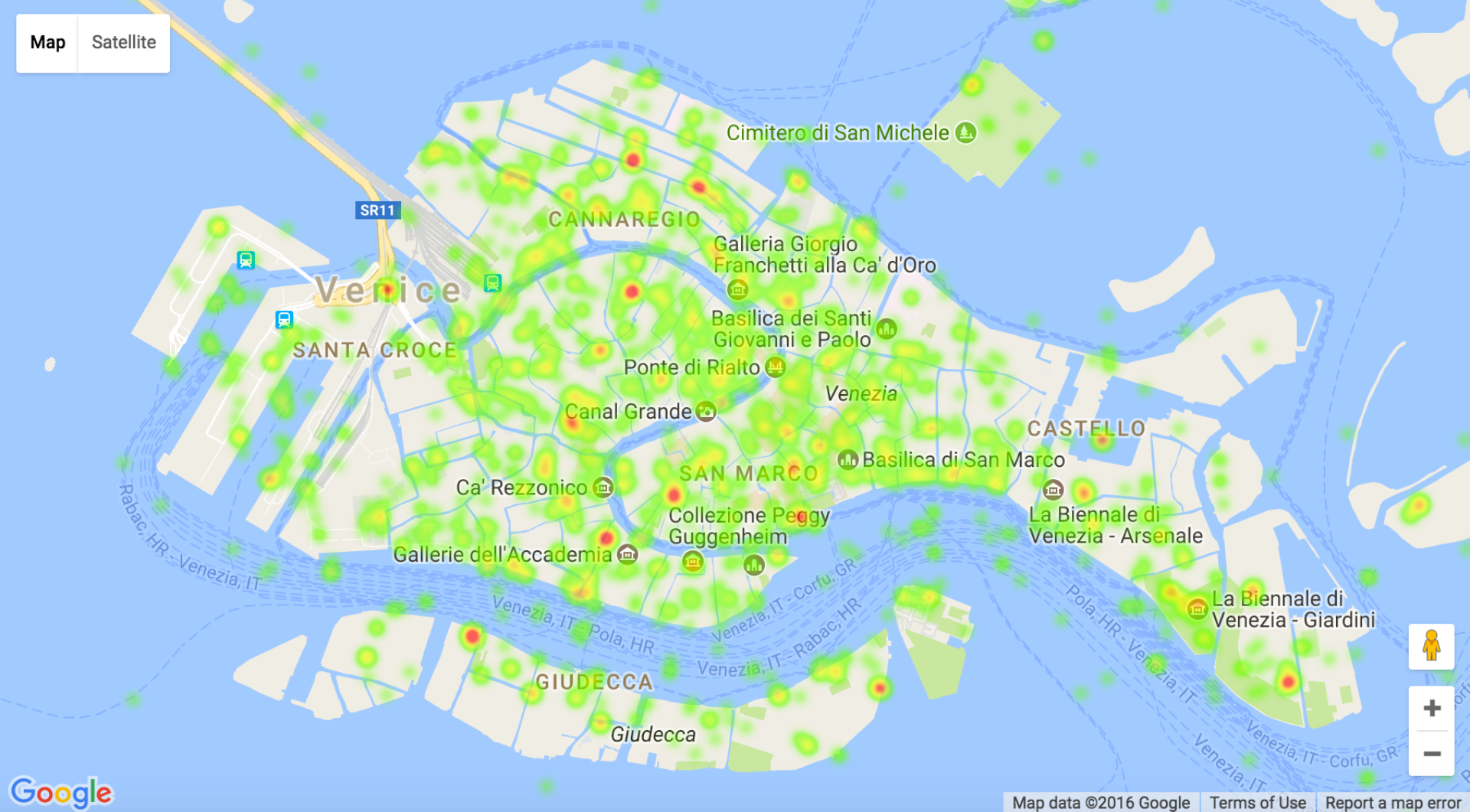
Heatmap: Folklore 2014



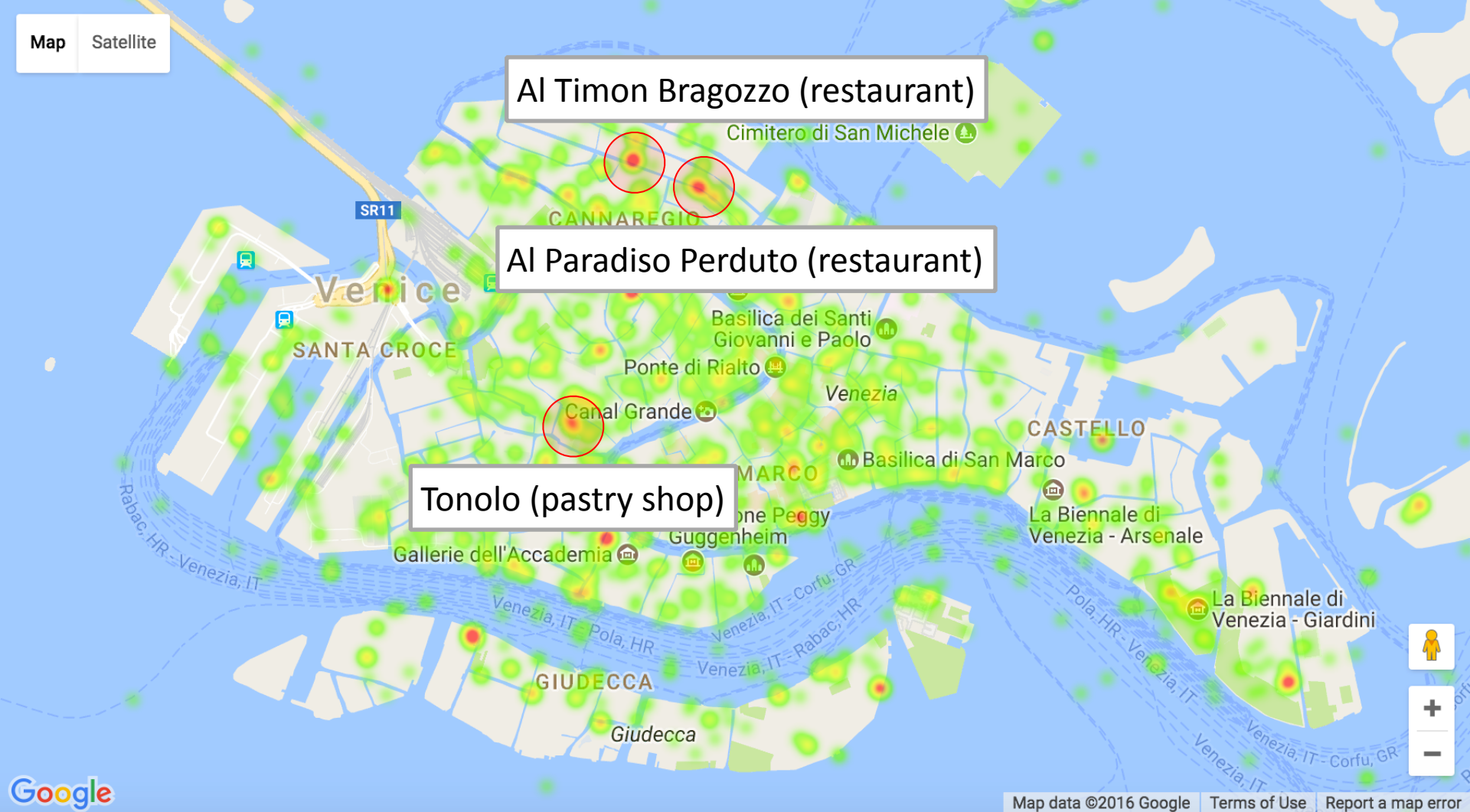
Heatmap: Folklore 2015



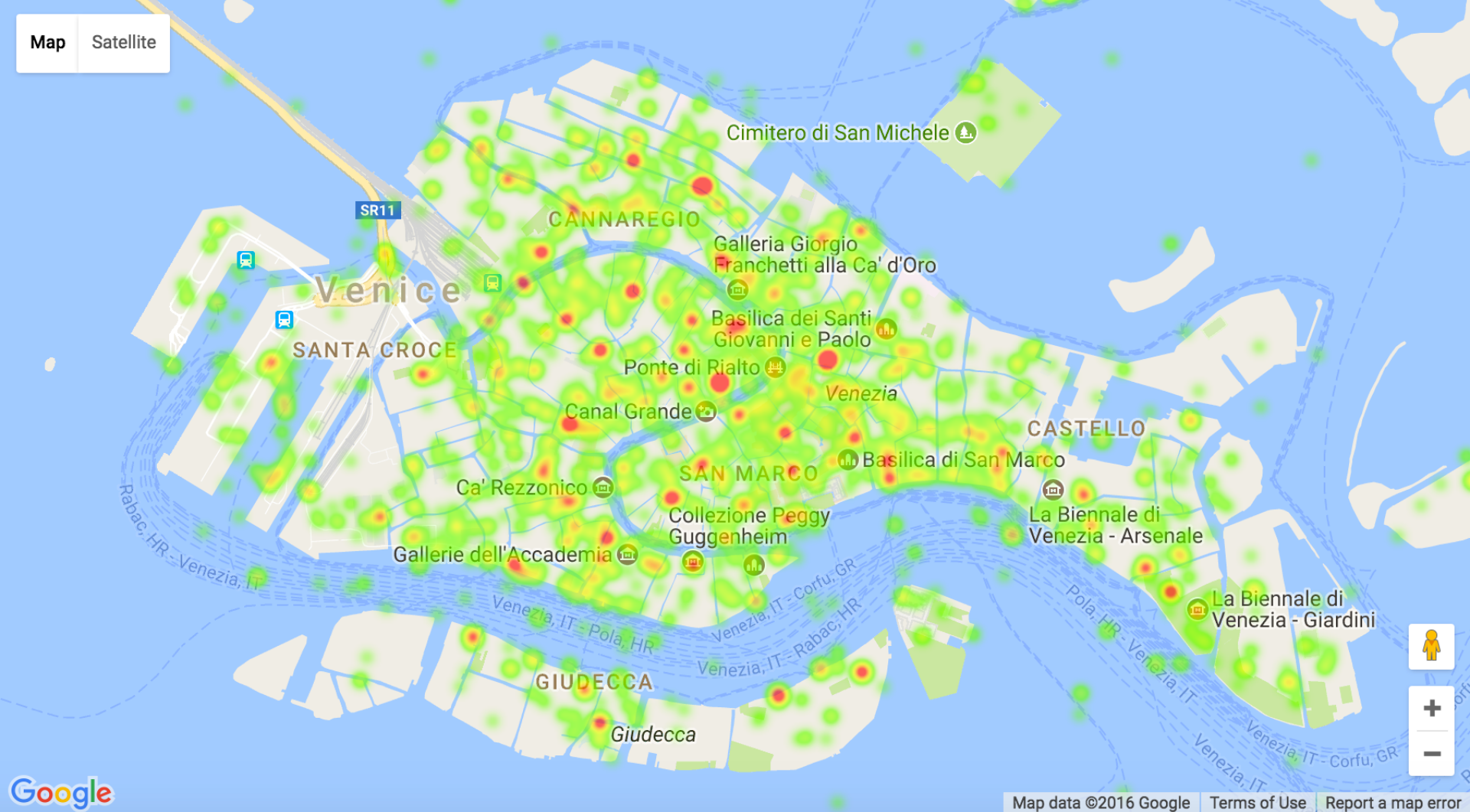
Heatmap: Folklore 2015



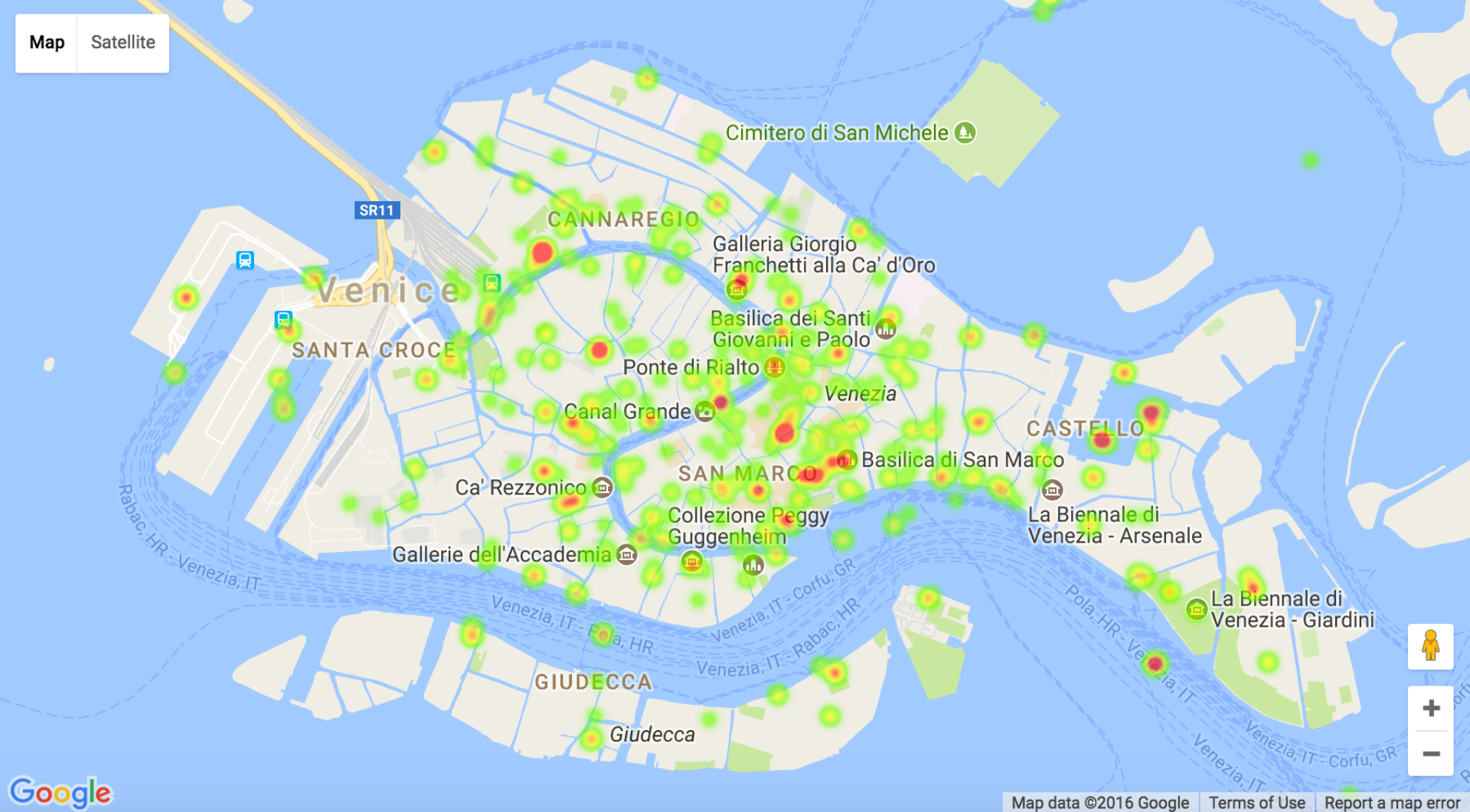
Heatmap: Food 2014



Heatmap: Food 2014



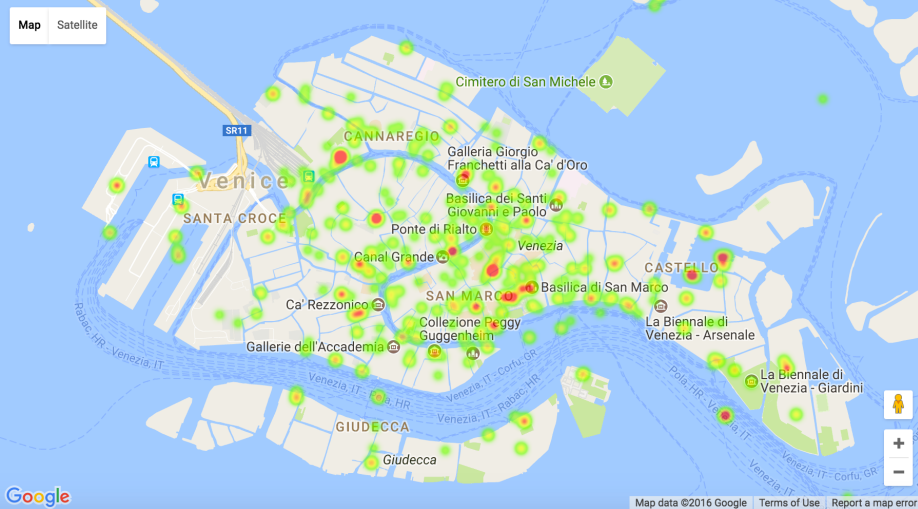
Heatmap: Food 2015



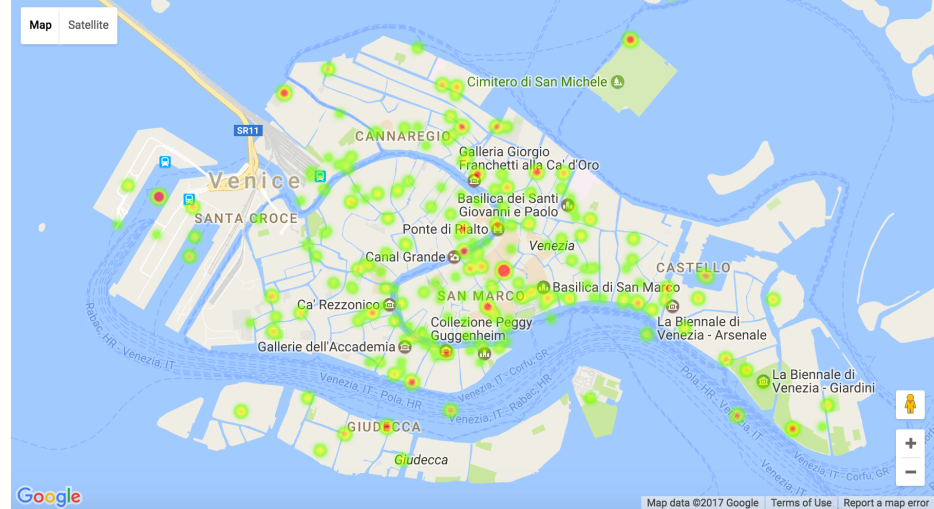
Heatmap: February 2015 (during Carnival)



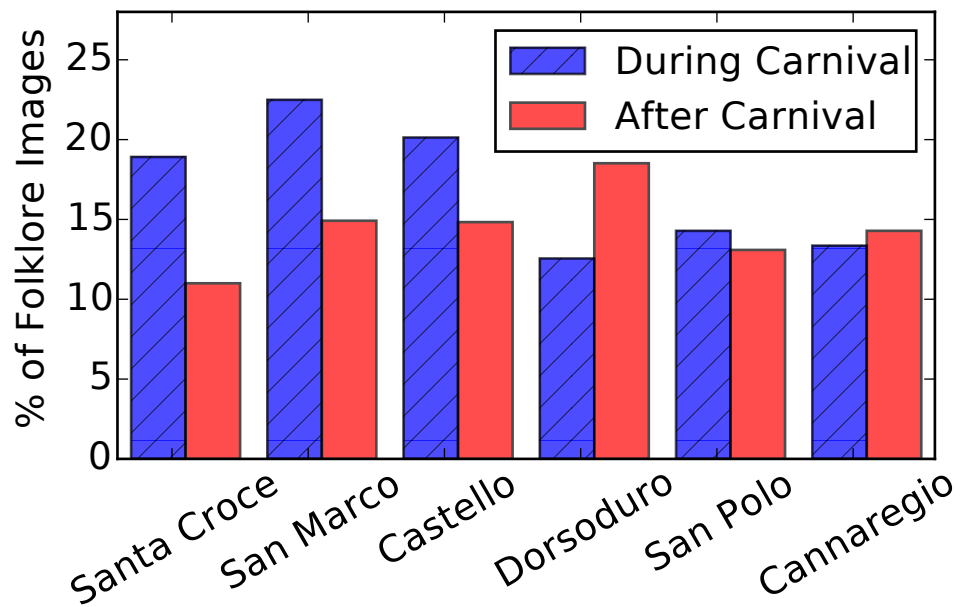
Heatmap: March 2015 (after Carnival)



During Carnival



After Carnival

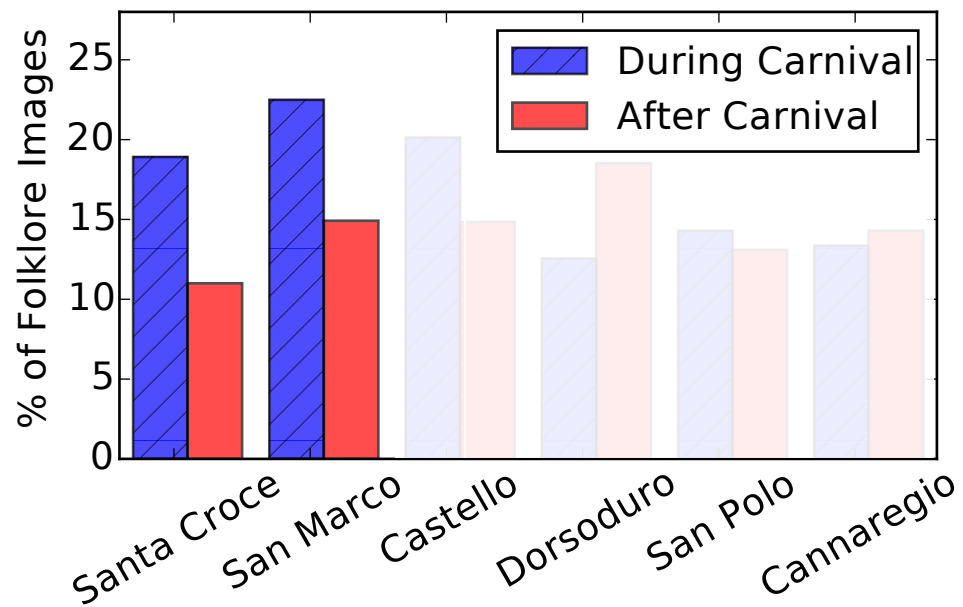




During Carnival



After Carnival





1. Study limited to users of Instagram
2. Image classification not state-of-the-art
3. Large number of photos classified as *Other*
4. Textual information (e.g, #hashtags) discarded

Conclusion and future work

- We explored tourism consumption through the lens of Instagram
- The analysis of 90k photos over two years highlights the presence of touristic hotspots
- The signal is influenced by external events and can reveal preferred touristic routes during such events

Conclusion and future work

- Potential areas of applications:
 - Urban planning
 - Marketing and advertising campaigns
 - Personalised tourist guide by linking city representation to user preferences, as determined by his/her shared photos
- Future work will investigate text to associate sentiment to places and will use CNN to improve image classification

Questions?

<https://cs.aston.ac.uk/~rossil/>

l.rossi@aston.ac.uk

 blextar