SPATIOTEMPORAL DYNAMICS OF THE REGIONS OF INTEREST OBTAINED FROM GEOTAGGED PHOTOS

Dissertation presented to the Federal University of Viçosa, as part of the requirements of the Post-Graduate Program in Computer Science, to obtain the title of Magister Scientiae.

Laura Pérez Vera

Giovanni V. Comarela

Jugurta Lisboa-Filho

UFV
Universidade Federal de Viçosa

DEPARTAMENTO DE INFORMÁTICA

Tourism and politics

Tourism is much more than “a place to stay during holidays”. The political impact on a region can carry a significant cultural and psychic/symbolic impact in the tourism of that region. This relation can be shown when studying the dynamics in the tourism related to a political event.

Volunteered Geographic Information (VGI) and Spatial footprints

- VGI brings a new notion of infrastructure to collect, synthesize, verify and redistribute geographic data through geolocation technology, mobile devices and geographic databases.
- Spatial footprints are a special type of VGI on social media platforms.
- People record and note their spatial footprints when they travel and experience the world.

Regions of Interest (ROI)

- Spatial footprints directly show specific phenomena, scenes or status of reality, which is a great opportunity to discover valuable geographic information and notice changes in reality over time.

Objective

Analyze the relations between the ROIs of a geographical area and the political events in a country, taking as a data source the spatial footprints of the users that visited that area. This analysis has the purpose to make a contribution to the tourism in the area, providing insights to the local government.

Specific objectives

1. Obtaining spatial footprints of the geographical area of Havana;
2. Obtention of ROIs in the geographic area Havana by clustering methods and algorithms to define the geometry of the clusters;
3. Identify, through analysis of the dynamics of ROIs over the years, events that had an impact on tourism.

Scientific Articles

1. “What about Cuba? Political tourism or tourism policies?” presented in the ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems 2020;
2. “Analysis of the international tourism in Havana through their ROIs dynamics.” article to be submitted.

ARTICLE I: What about Cuba? Political tourism or tourism policies?

4th ACM SIGSPATIAL International Workshop on Geospatial Humanities

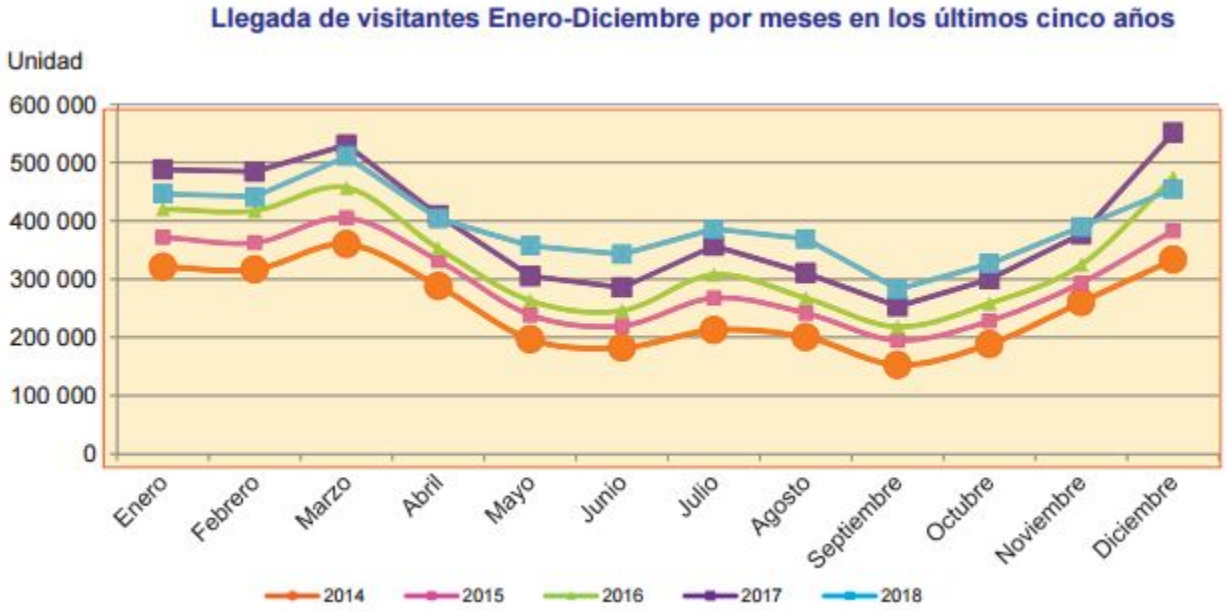
Universidade Federal de Viçosa

acm SIGSPATIAL

What about Cuba?

1. Political history.
2. Tourism in Cuba is one of the main sources of revenue for the island.
3. Changes in their economy in the last years

Statistics tourism in Cuba



ONEI
OFICINA NACIONAL
DE ESTADÍSTICA E INFORMACIÓN
República de Cuba

Given the political changes, are there changes in tourism?

Spatial footprint

Spatial footprint is a special kind of people-generated VGI on social media platforms. By posting stories or sharing information on social media, people can geo-refer posts by attaching their locations .

Cheng, Z.; Caverlee, J.; Lee, K.; Sui, D.Z. Exploring millions of footprints in location sharing services.312ICWSM2011,2011, 81–88

Region Of Interest (ROI)

A ROI may contain several co-located geographic features, such as restaurants on a pedestrian street or several nearby landmarks. The ROI can also include areas that do not have important landmarks, but simply provide panoramic views (e.g.: areas in Paris that provide a good view of the Eiffel Tower).

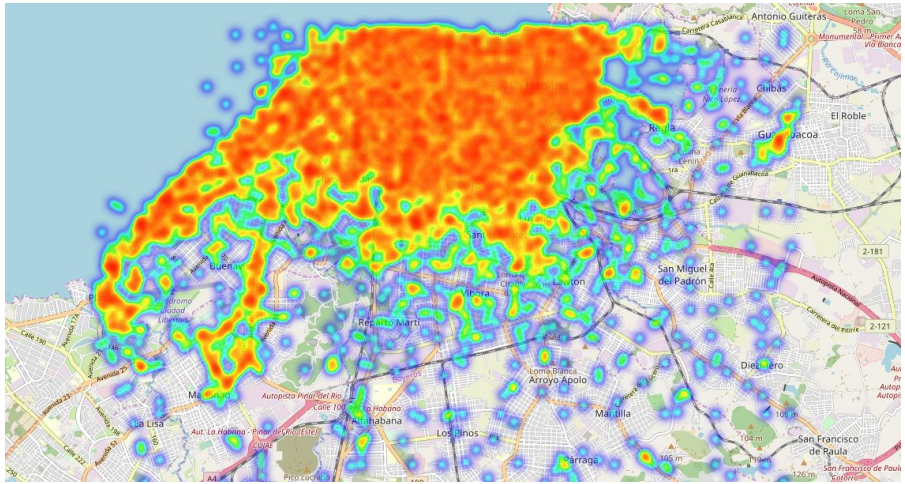
Elias, B. Extracting landmarks with data mining methods. International Conference on Spatial Information Theory. Springer, 2003, pp. 375–389

Political event

The plane of the president of the United States on the approach maneuver to the José Martí International Airport in Havana. Photo by: Yander Zamora

Data collection

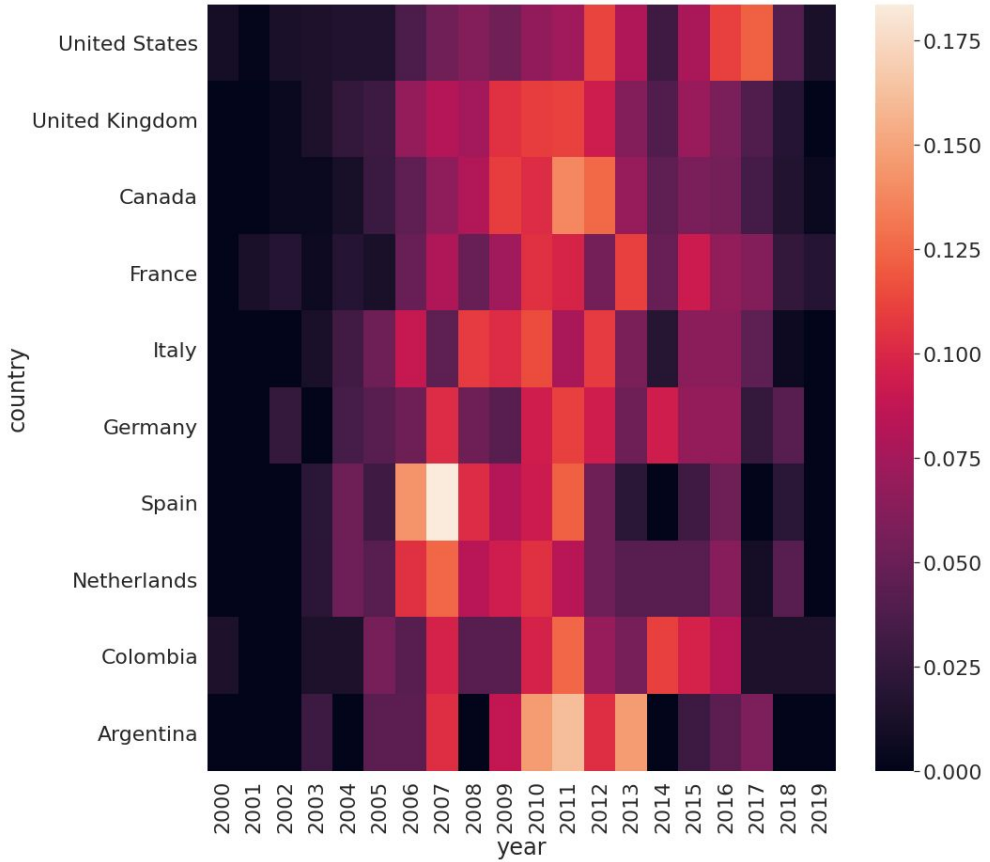
The data was collected from the geographic area of Havana using the coordinates belonging to the bounding box defined by latitude in the interval [23.0499, 23.1470]and longitude in the interval [-82.4551, -82.3025]. In total, **136,970 records** within the bounding box of Havana were retrieved, with a total of **7323 users**.



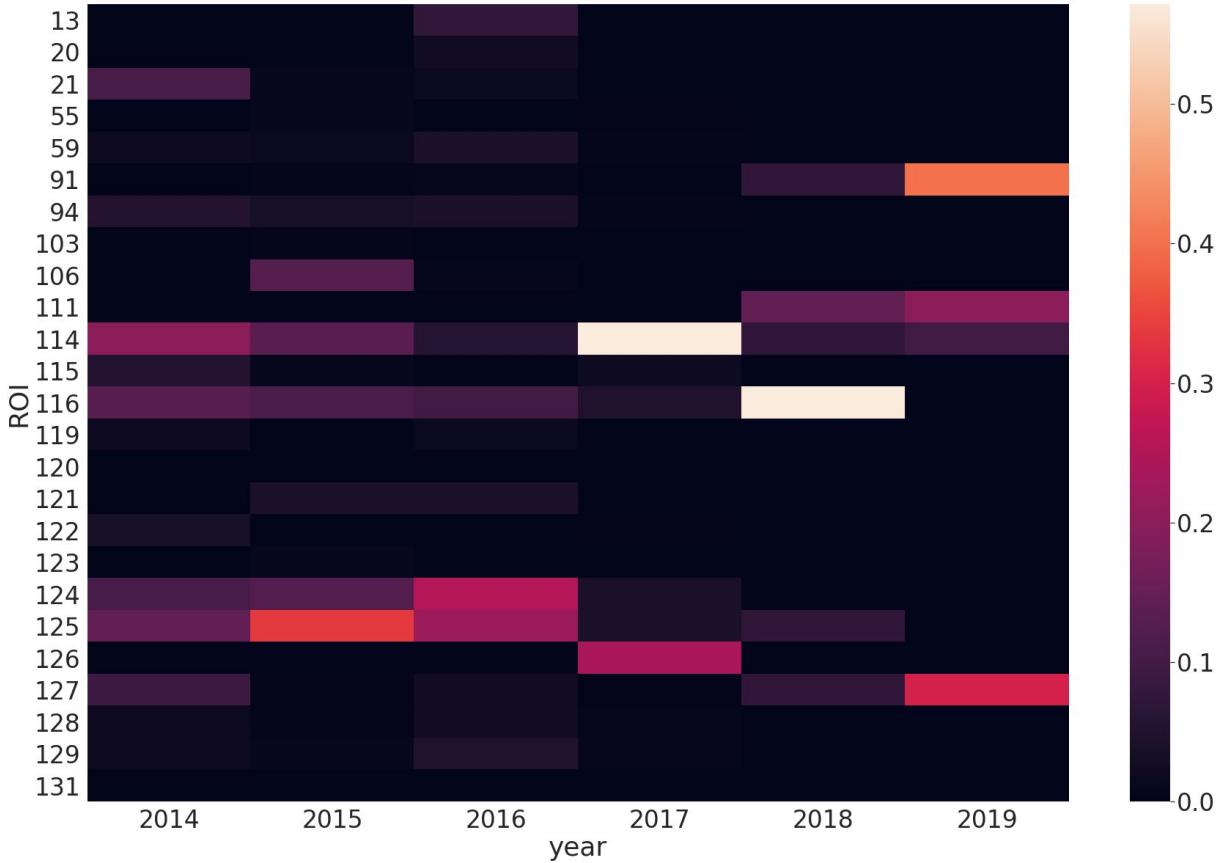
Clustering

- **HDBSCAN** clustering algorithm. A total of 136 ROIs, with different shapes and sizes.
- ***α*-shape** algorithm

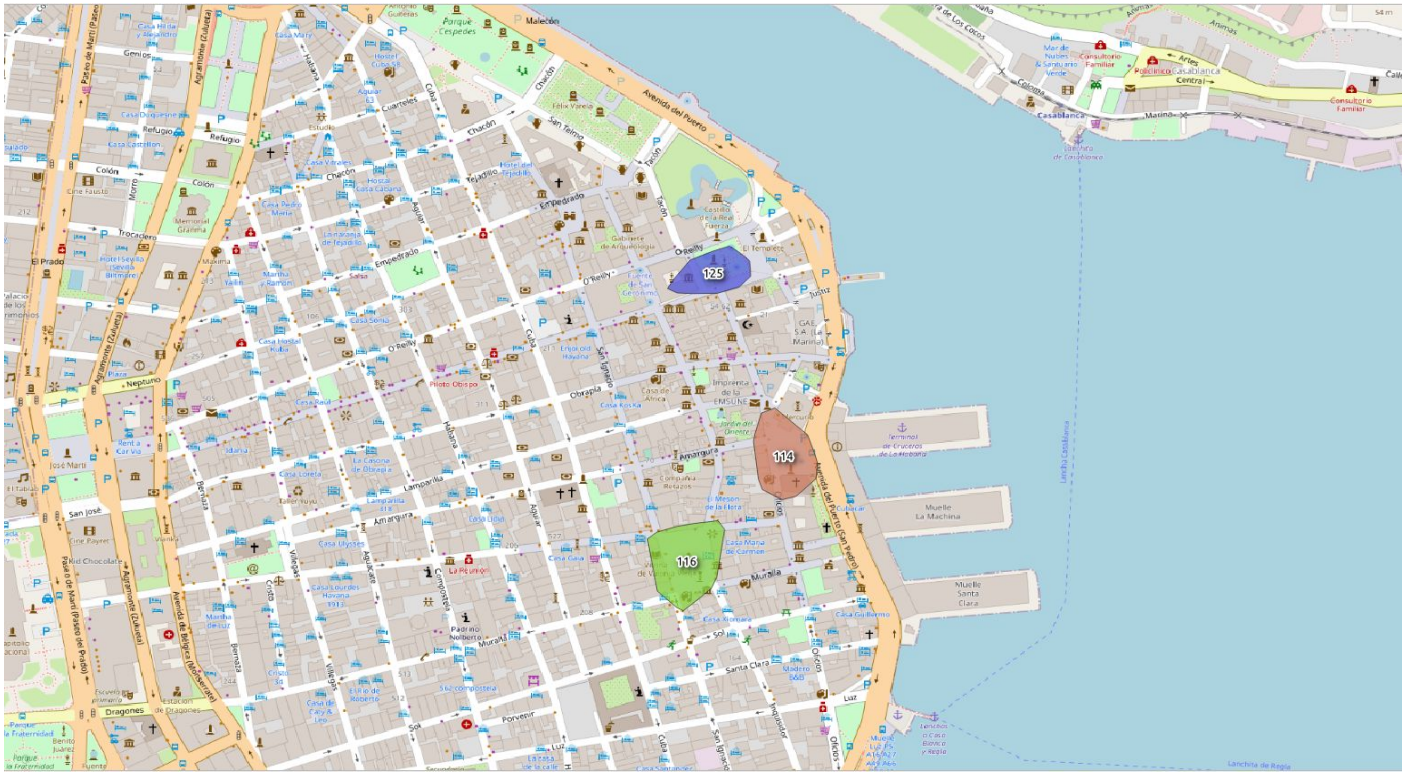
Temporal Analysis



Temporal Analysis



ROIs



ROIs mostly visited by tourists from the United States: "Plaza San Francisco de Asís" (114), "Plaza Vieja" (116) and "Plaza de Armas" (125).

Discussions

- Changes in the United States’ policies with regard to Cuba had a significant impact on the flow of US residents visiting Cuba.
- Most affected area was nearby the port of Havana.

Future work

- Analyze the tourism dynamics for all the ROIs in order to understand how these constructions and renovations affected the tourists’ interest.
- Analyze in a more wide range the ROIs that were visited by the American tourists in a way to look for political interests.

ARTICLE II: Analysis of the international tourism in Havana through their ROIs dynamics.

UFV
Universidade Federal de Viçosa

Objective

Analyze the spatial distribution and temporal dynamics of the ROIs obtained from geotagged photos.

Research Question

Is there a connection between the dynamics of the visits to the country and the political events that happened in the country? If so, how this connection affects the interest on the main ROIs on a temporal basis.

Related Works

A three-layer framework for extract the AOI from geotagged photos and understanding their spatiotemporal dynamics is presented by Hu et al. (Hu et al.,2015), retrieved data is used from Flickr through their public API (Application Programming Interface).

Related Works

Kuo et al (Kuo et al.,2018) proposed a framework to obtain the ROIs, using as data source the geotagged photos from Flickr. The attractive footprints discovery is conducted to eliminate noises and select valid footprints with a local maximum, using a voting value equation.

Related Works

The study of Li et al. (Li et al.,2013) conducted an exploratory analysis to identify tourist attractions as hot spots in Flickr usage patterns, using partial least squares regression.

Data collection and Pre-processing

- Data extraction process the Flickr's public API was used providing a wide range of information of each photo that has been uploaded.
- **flickr.photos.search** method was used, overcoming the limitation of return 4,000 results for any given search query with the development of a method.

Clustering

- The data clustering process in this study are **mainly based on the first stages of the Spatial Overlap method (SO) described by Kuo et al.** (Kuo et al.,2018).
- The process of obtaining those spatial footprints which represent **attractive footprints**, in order to eliminate some noise and to identify those spatial footprints that are going to form each ROI.
- The **pattern discovery** method for each attractive footprint.

Attractive Footprints Discovery

- Conducted under the assumption that people will gather in areas that they are interested in, for each spatial footprint within each subset is calculated the voting value.

$$v_{p_i} = \sum_{i=1}^{j} w_{ij}, w_{ij} = e^{-\frac{||i-j||}{2\sigma^2}}$$

- A threshold needs to be set for finding those spatial footprints that received special attention on each subset.

$$Q = (p * (n+1))^{th} Term$$

Pattern Discovery

- To discover the pattern for each attractive footprint that is going to be clustered, the pattern discovery process is specified. The pattern discovery process is based on the one proposed by Kuo et al. (Kuo et al.,2018).

Pattern Discovery

```
def distance_radius_subset(subset):
        RADIANT_TO_KM_CONSTANT = 6367
        radius_km = 50/1e3                    # 50 meters of radius
        radius_radian = radius_km / RADIANT_TO_KM_CONSTANT
        lat_long=[np.radians(list(zip(x.long, x.lat))) for x in subset]
        A = radius_neighbors_graph(lat_long, radius_radian,mode='distance',
            metric='haversine')
        return A
def pattern_yearly(sp):
    X = [0]*12
    for n in sp.dict_owner.values():
        m = month of n
        X[m-1] += 1
    return X
def pattern_monthly(sp):
    X = [0]*y                                 #y:number of years of being analyzed
    for n in sp.dict_owner.values():
        m = year of n
        X[m] += 1
    return X
```

Regions of Interest

- For generating the polygons objects from the points inside each cluster the **alpha-shape** algorithm is used.
- In order to cleaning some of the polygons that spatially was giving no information, the function **roi** was implemented.

Regions of Interest

```
def roi(subset):
        for cluster in subset.clusters:
            points=[x.geom for x in cluster.sp]
            alpha_shape = alphashape(points)
            if alpha_shape.geom_type == 'Polygon':
                patches.append(alpha_shape)
        if len(patches)>=1:
            polygons=[]
            polygons.append(patches[0])
            for pol1 in range(1,len(patches)):
                flag=False
                for pol2 in range(0,len(polygons)):
                    if patches[pol1].intersects(polygons[pol2]):
                        por=(patches[pol1].intersection(polygons[pol2]).area/polygons[pol2].area)*100
                        if por>30:
                            flag=True
                            break
                if not flag:
                    polygons.append(patches[pol1])

            return polygons
```

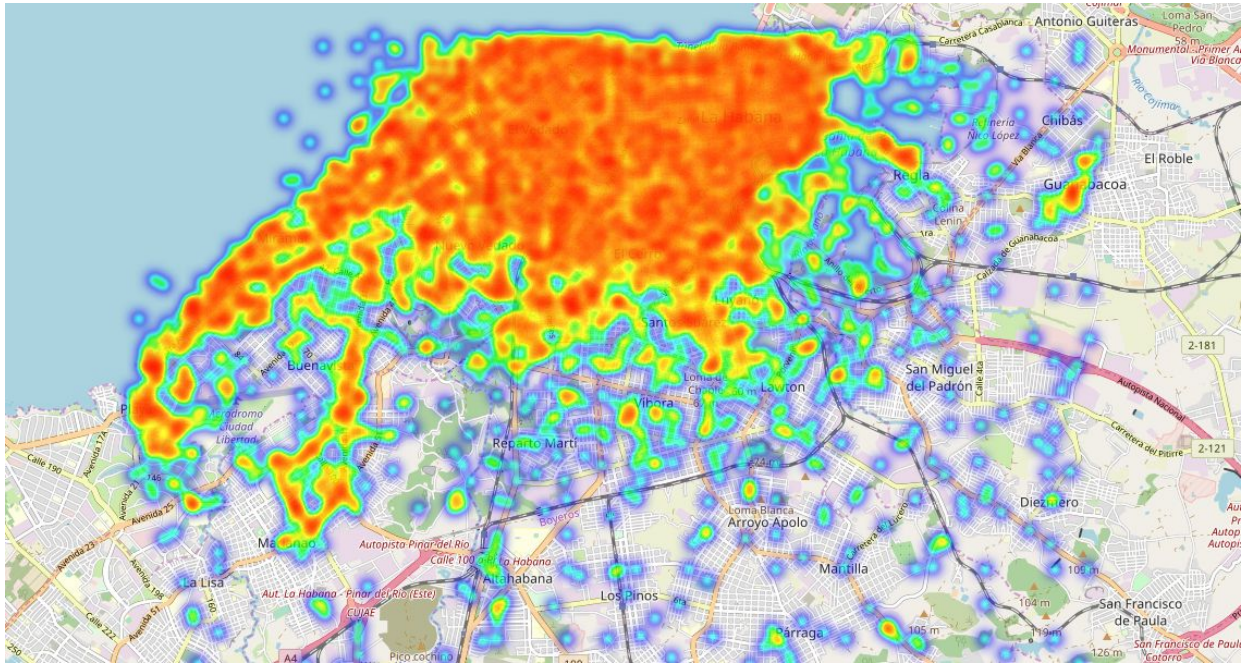
Dynamics of the ROI

- Find the uniqueness of each polygon in order to be able to identify them inside each subset
- For this paper two polygons are going to be considered as equals if the Hausdorff distance is less than 0.003.

Results

- The data was collected from the geographic area of Havana using the coordinates belonging to the bounding box: [-82.4551, 23.0499, -82.3025, 23.1470]
- In total 136,970 records within the bounding box of Havana were retrieved, with a total of 7323 users, using the flickr.photos.search method.

Results

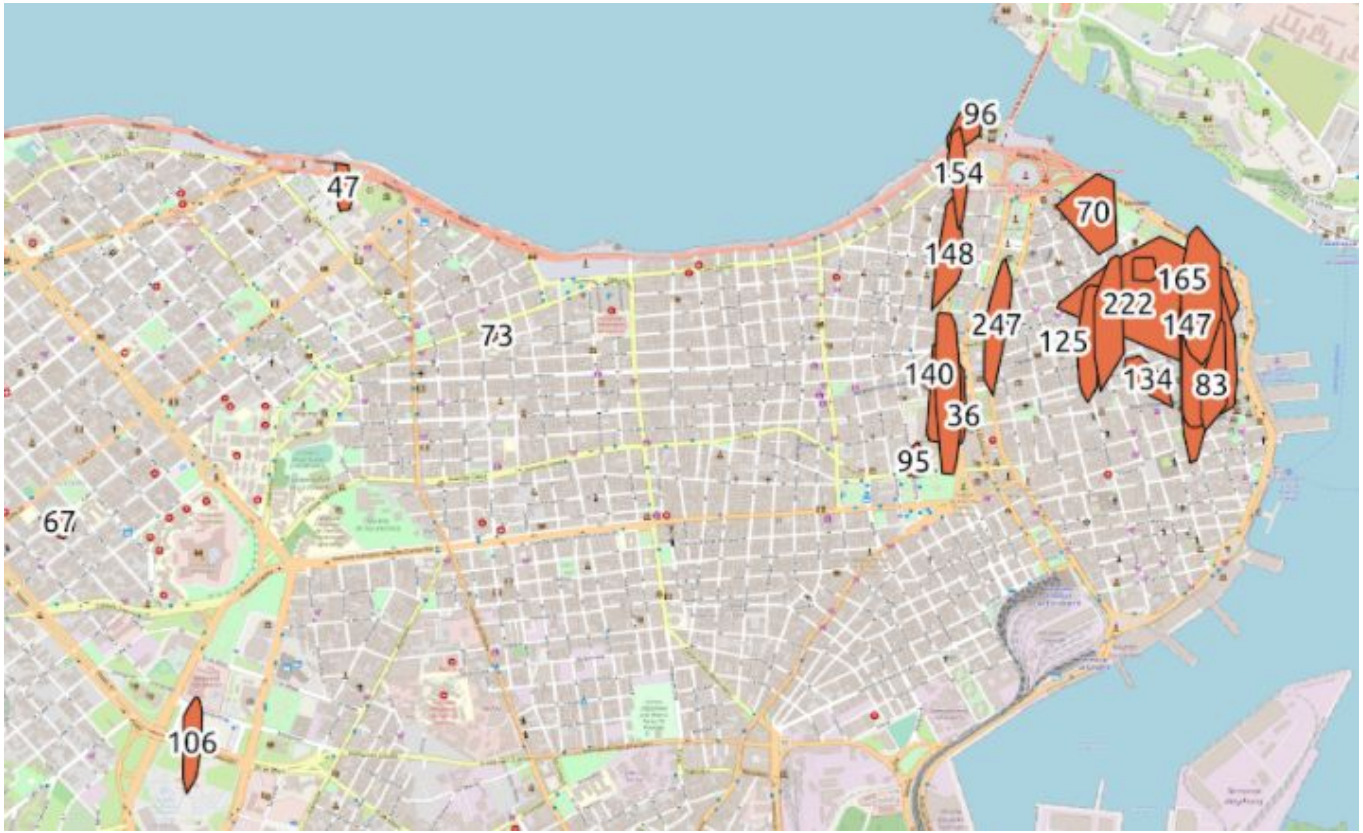


Heatmap of the photos collected from Flickr

Results

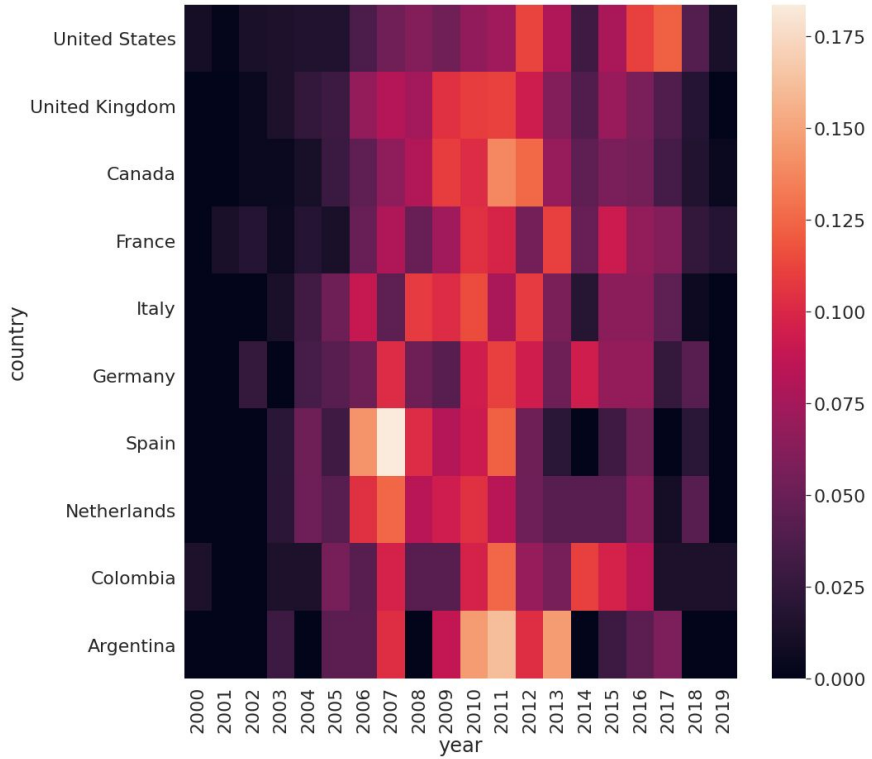
Result of the spatiotemporal method per year

Results



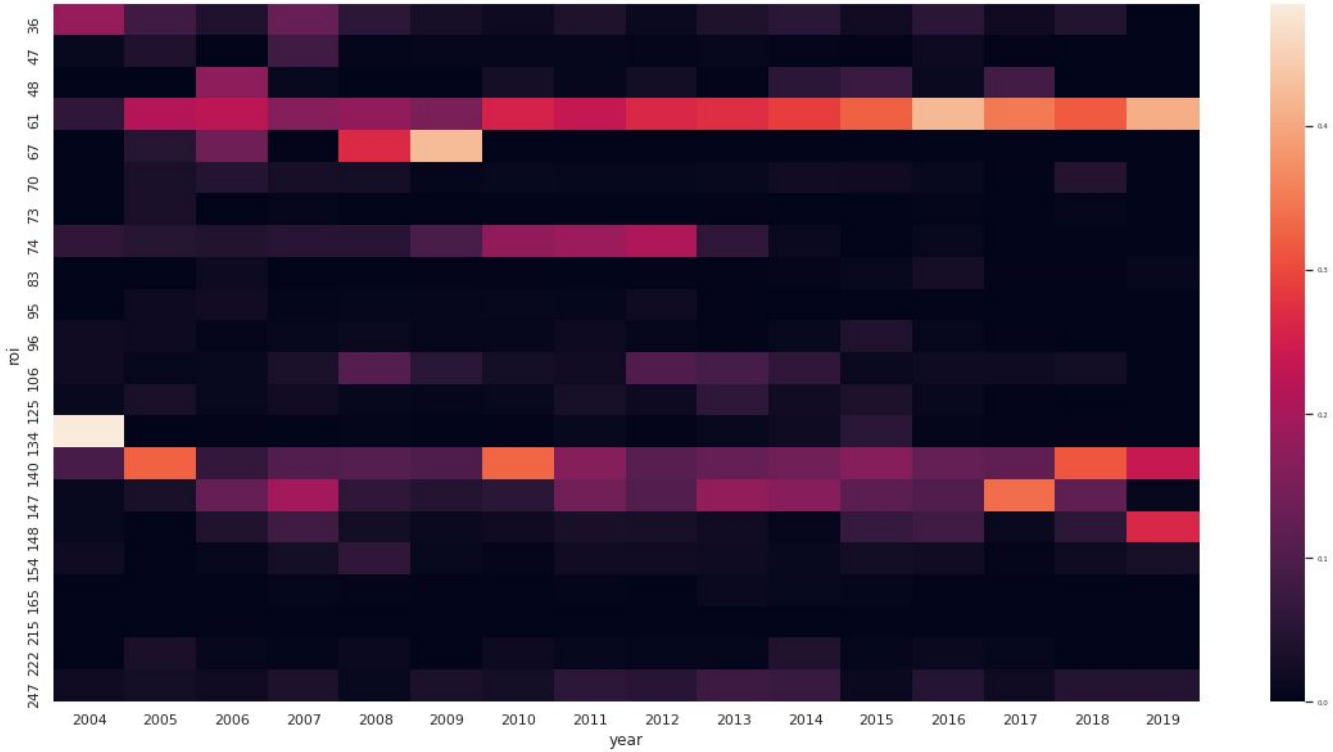
ROIs obtained after apply similarity section

Temporal Analysis



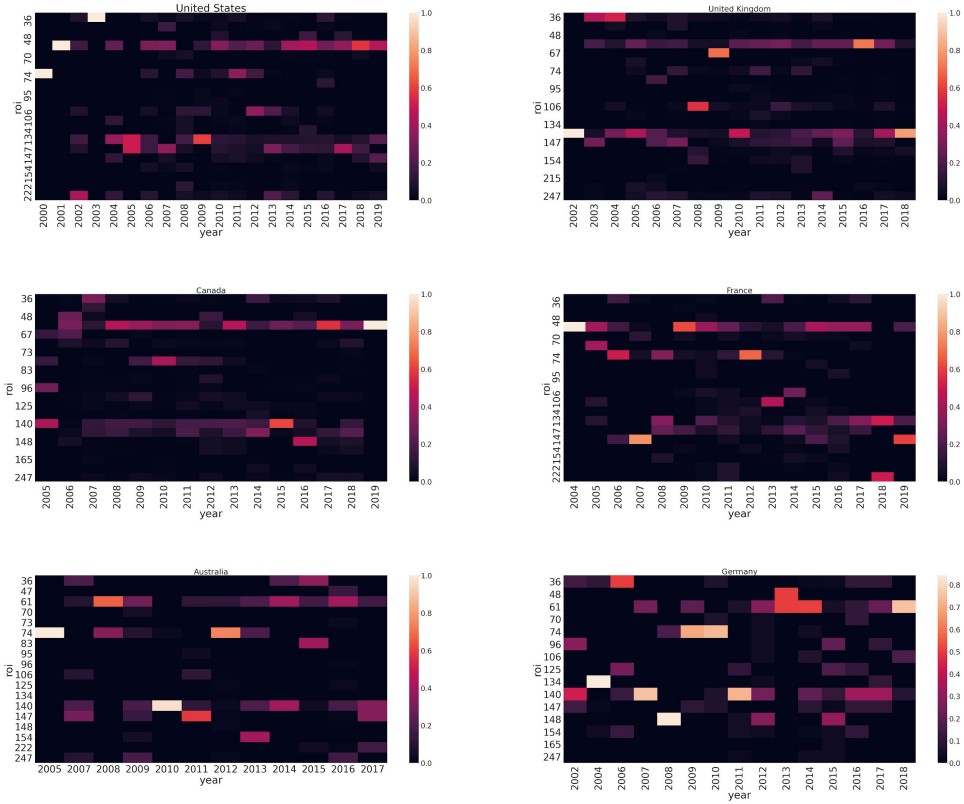
Temporal distribution of visits to Havana per country of origin.

Temporal Analysis



Temporal distribution of visits to Havana per ROIs. Values are normalized according to the sum of each row.

Temporal Analysis



Temporal distribution of visits to Havana per country of origin and per ROIs. Values are normalized according to the sum of each row.

Conclusions

- The changes in the United States policies regarding Cuba had a significant impact on the flow of the American tourists visiting Havana.
- Canada's policies towards the island had an effect in the flow of Canadians visiting the city concluding not only that these policies had an effect in the Canadian tourism but that the measures applied by the government of Cuba were effective regarding to increase the flow of Canadians visitors.

Conclusions

- The visit of the president of France in 2015 to Cuba had an impact and increased the visits in 2016 where agreements were signed between the two countries.
- There is a relation between the political events of the main foreign countries that visited the island and the tourism in Cuba.
- The internal policies of the country of the approval of the activity of rental houses in Cuba, had an effect on existing businesses of that type in the city of Havana.

General Conclusions

- The spatiotemporal polygons of the region of Havana obtained suggest that the series of measures that were applied on the island to increase the tourism, had an effect in the Canadian tourism to Havana, also the rapprochement between the Cuban government and Canadian had an effect in the tourism in the city.

General Conclusions

- There is a relation between the political events of the main foreign countries that visited the island and the tourism in Cuba, suggesting also that the internal policies of the country of the approval of the activity of rental houses in Cuba, had an effect on existing businesses of that type in the city of Havana.

Future Work

- As future work it is suggested to be analyzed in greater depth the ROIs that were visited by the American tourists in a way to look for political interests. Also these ROIs can be analyzed in order to understand how the new constructions and renovations affected the tourists' interest.

Thanks for the attention...

questions?

laura.vera@ufv.br

UFV
Universidade Federal de Viçosa

DEPARTAMENTO DE INFORMÁTICA