Developing Dorling Cartogram Algorithm for Yelp Data

Members:
Weijia Jin
Jialiang Zhang
Motivation

- Lack of development of Dorling Cartogram algorithm in the community
- No current implementation of Dorling Cartogram on Yelp Dataset
- Dataset contains geographical attributes which is suited for visualization by cartograms
- Visualizing the large set of data on business and reviews can generate useful insights
Background Survey

- Area cartograms: their use and creation, Dorling, D. (1996)
  - Cartograms are maps which are distorted by size, shape or distance
  - Size could be made proportional to any measurable feature
Dorling Cartogram

- Uses
  - ability to shock
  - different areas can be shaded with different colours to show variation over space
  - observe trends and patterns
  - compare across different geographic regions or measures
Dorling Cartogram algorithm

For each region

Calculate the radius of a circle so that its area is proportional to population*

While the forces calculated below are not negligible

For each region (the order of calculation has no effect)

For each region which overlaps with the region

Record a force away from the overlap in proportion to it

For each region which originally neighboured the region

Record a force towards it proportional to distance away

If the forces of repulsion are greater than attraction

Scale the forces to less than the distance of the closest circle

Combine the two aggregate forces for each circle

For each region

Apply the forces recorded to be acting on each circle to its centroid
Yelp data format

```json
{
    'type': 'business',
    'business_id': (encrypted business id),
    'name': (business name),
    'neighborhoods': [(hood names)],
    'full_address': (localized address),
    'city': (city),
    'state': (state),
    'latitude': latitude,
    'longitude': longitude,
    'stars': (star rating, rounded to half-stars),
    'review_count': review count,
    'categories': [(localized category names)]
    'open': True / False (corresponds to closed, not business hours),
    'hours': {
        (day_of_week): {
            'open': (HH:MM),
            'close': (HH:MM)
        },
        ...
    },
    'attributes': {
        (attribute_name): (attribute_value),
        ...
    }
}
```
Existing Yelp Visualizations

- Maps
- Word clouds
Las Vegas Zip Code Map

Source: Las Vegas Real Estate Authority
Demo 1: Dorling Cartogram

Dorling Cartogram of Yelp Dataset

Based on Protovis
Dorling Cartogram Algorithm

```javascript
var collisionConstraint = pv.Constraint.collision(function(d) d.radius + 1),
positionConstraint = pv.Constraint.position(function(d) d.position);
linkConstraint = pv.Force.spring(100).links(borders);
```
Demo 2: Pseudo-Dorling Cartogram

Pseudo-Dorling Cartogram of Yelp Dataset

Based on Mike Bostock’s D3 Implementation
Pseudo-Dorling Cartogram Algorithm

- It lacks links between adjacent features.
- Instead of trying to preserve connectedness, it tries to preserve locality.
- Each circle stays as close as possible to its origin without overlapping.

```javascript
function tick(e) {
    node.each(gravity(e.alpha * .1))
    .each(collide(.5));
}
function gravity(k) { .... }
function collide(k) { .... }
```

*Based on Mike Bostock’s D3 Implementation*
<table>
<thead>
<tr>
<th>Deadline</th>
<th>Deliverable</th>
<th>Member</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/17</td>
<td>Filtering and aggregating Yelp data</td>
<td>Weijia</td>
</tr>
<tr>
<td></td>
<td>Adapting base algorithm for our data</td>
<td>Jialiang</td>
</tr>
<tr>
<td>11/24</td>
<td>Experimenting with different algorithm constraints</td>
<td>Jialiang</td>
</tr>
<tr>
<td></td>
<td>Experimenting with different measures and coordinates</td>
<td>Weijia</td>
</tr>
<tr>
<td>12/1</td>
<td>Creating interactive visualization in D3</td>
<td>Jialiang, Weijia</td>
</tr>
<tr>
<td>12/3</td>
<td>Final poster presentation</td>
<td>Jialiang, Weijia</td>
</tr>
</tbody>
</table>

   http://www.dannydorling.org/books/visualisation/Homepage.html

3. Noncontiguous cartogram of USA:  
   http://show.mappingworlds.com/usa/

4. Cartogram Central:  
   http://www.ncgia.ucsb.edu/projects/Cartogram_Central/

5. Noncontiguous cartograms in OpenLayers and Polymaps:  
Questions and Feedbacks

Thank you!