Methods for creating interactive map to accompany "Clayton County: on the road to fitness" story

By Julianne Wyrick

1. Initial data analysis
   a. First, I downloaded the County Health Rankings Georgia Dataset
   b. Then, I hid columns such that county and obesity were easier to compare.
   c. I sorted counties by obesity prevalence from highest to lowest, and noticed Clayton County was in the top 12 most obese counties.
   d. I found a list of the 10 counties governed by the Atlanta Regional Commission (group of counties in the metro Atlanta area).
   e. I found that seven out of the other nine counties in the area around Atlanta governed by the Atlanta Regional Commission have obesity rates below 30%. (Cherokee – 28%, Clayton – 36%, Cobb – 23%, DeKalb – 26%, Douglas -35%, Fayette -25%, Fulton -23, Gwinnett- 25%, Henry -29%, Rockdale -33%.
   f. I decided to map obesity prevalence in Georgia, emphasizing these 10 counties using a chloropleth (shaded) map.
   g. I investigated data in County Health Rankings Georgia dataset for potential factors involved in obesity rate for these counties, specifically median household income, physical inactivity, access to exercise opportunities and limited access to healthy foods. Clayton County differed from the surrounding counties in many of these factors, so I decided that I would include this information in tool tips on the map.
2. Data cleaning
   a. I created a new Excel file and copied columns the from County Health Rankings Georgia dataset corresponding to county, median household income, physical inactivity, access to exercise opportunities and limited access to healthy foods in order to get rid of the extra data.
   b. I checked to make sure data was copied properly.
3. Importing data for map into TileMill
   a. I started a New Project in TileMill, checking the option for "Default data."
   b. I dound nation-based cartographic boundary file (shapefile) with county boundaries from U.S. Census data (http://www.census.gov/geo/maps-data/data/tiger-cart-boundary.html) and downloaded it.
   c. I used QuantamGIS to edit the shapefile so it only included county boundaries for Georgia as opposed to the entire nation. First, I downloaded QGIS, Gdal Complete, GSL Framework, and Matplotlib (http://www.qgis.org/en/site/forusers/download.html). Then, I opened the shapefile in QGIS by using the "Add Vector Layer" button. Next, I right-clicked the new layer and chose "Open attribute table." Then, I filtered the table for the Georgia county boundaries by filtering for "13" (the state number for Georgia) in the "State" column. Then, I selected the filtered data and right-clicked on the layer name in the left column of QGIS and chose "Save selection as." I saved the shapefile with a new name (ex: "Georgia Counties").

d. I merged the County Health Rankings data with the shapefile.
   i. To do this, I created a "key," a field or column in both data and shapefile sharing common values. I opened the Excel file of my data and found that it had a FIPS column, which includes a number identifying the county, such as "13001" for Appling County. The number represents the state (Georgia = 13) and the county number (Appling = 001). I opened the shapefile in QGIS by using the "Add Vector Layer" button. By right clicking on the layer I added and choosing "Open attribute table," I saw that the shapefile had the state number in a "State" column and the county number in a "County" column. I returned to Excel to create a column that would correspond to the "County" column in the shapefile, so that I would have a "key" for merging the data and shapefile. I used the RIGHT formula in Excel to extract the last three characters (the county number) from the FIPS number and place them into a new column, which I labeled "County" like the corresponding column in my shapefile.
   ii. I downloaded LibreOffice and opened the Excel data file in this program. I saved the Excel data file as a .dbf file, choosing UTF-8 format.
   iii. I used the "Add vector layer" button to open the .dbf file in QGIS.
   iv. I opened the shapefile's properties by double-clicking on the layer name. I went to the "Joins" tab and clicked the "+" button to add a join. I selected .dbf data layer for the "Join layer" and "County" for the key field to be joined on the .dbf data layer. I selected "County" for target field to be joined on the shapefile.
   v. I right-clicked the shapefile layer to save it with a new name. I saved it in ESRI Shapefile format.
e. I imported the new shapefile with merged boundaries and data in TileMill as new layer.

4. Styling Map (used http://colorbrewer2.org/ to find complimentary colors)
   a. I used the following styling for the #countries layer.
      i. #countries {
         line-color:#226688;
         line-width:0.5;
         polygon-opacity:1;
         polygon-fill:#36b666;
         }
   b. I used the following styling for the #GeorgiaCounties layer.
      i. #GeorgiaCounties {
         line-color: #0a2a02;
         line-width:1;
         polygon-opacity:1;
         polygon-fill:#ae8;
         }

c. I created a color gradient to show varying levels of obesity prevalence in counties.
   i. I chose a gradient using http://colorbrewer2.org/.
   ii. I determined I would represent 4 classes of the variable (obesity prevalence as measured in a percentage):
      1. Class 1: obesity prevalence >20 or obesity prevalence < 26
      2. Class 2: obesity prevalence >= 26 or obesity prevalence <31
      3. Class 3: obesity prevalence >= 31 or obesity prevalence <36
      4. Class 4: obesity prevalence >= 36
   iii. I used this code under the initial #GeorgiaCounties code above to tell TileMill the line color and the different shadings to color the counties as the variable (obesity prevalence) varies from low to high:
      @line: #226688;
      @shade0: #fee5d9;
      @shade1: #fcae91;
      @shade2: #fb6a4a;
      @shade3: #cb181d;
   iv. I used this code under the code above to apply the color gradients to the #Georgia Counties layer:

```
#Georgia Counties {
 line-color: @shade0;
line-width: 0;
polygon-fill: @shade0;
polygon-opacity:1;

[Workshee_4 > 20][Workshee_4 < 26] { polygon-fill: @shade0; }
[Workshee_4 >= 26][Workshee_4 < 31] { polygon-fill: @shade1; }
[Workshee_4 >= 31][Workshee_4 < 36] { polygon-fill: @shade2; }
[Workshee_4 >= 36] { polygon-fill: @shade3; }
[zoom > 10] { line-width: 1; }
}
```

d. I added tooltips to the map show obesity prevalence data when a user hovers over each county and median household income, physical inactivity, access to exercise opportunities and limited access to healthy foods when the user clicks on a county. I also added a legend to the map.
   i. Code for Teaser:
      {{{NAME}}}
      Obesity Prevalence: {{{Workshee_4}}}%
   ii. Code for Full
      {{{NAME}}}
      Obesity Prevalence: {{{Workshee_8}}}%

Median Household Income:
${{{Workshee_15}}}
Residents with Access to Exercise Opportunities:
{{{Workshee_12}}}%
Physically inactive residents:
{{{Workshee_18}}}%
Residents with limited access to healthy food:
{{{Workshee_6}}}%

iii. Code for legend

```
<div class='my-legend'>
<div class='legend-title'>Obesity in Georgia
Counties</div>
<div class='legend-scale'>
  <ul class='legend-labels'>
<li><spanstyle='background:#fee5d9;'></span>20
% to 25%</li>
<li><span
style='background:#fcae91;'></span>26% to 30%</li>
<li><span
style='background:#fb6a4a;'></span>31% to
35%</li>
 <li><span
style='background:#cb181d;'></span>36% and
greater</li>
</ul>
</div>
<div class='legend-source'>Source: <a
href="http://www.countyhealthrankings.org/app/ge
orgia/2014/downloads">2014 County Health
Rankings</a></div>
</div>

<style type='text/css'>
  .my-legend .legend-title {
    text-align: left;
    margin-bottom: 5px;
    font-weight: bold;
    font-size: 90%;
    }
  .my-legend .legend-scale ul {
    margin: 0;
    margin-bottom: 5px;
    padding: 0;
    float: left;
    list-style: none;
    }
```

```
.my-legend .legend-scale ul li {
  font-size: 80%;
  list-style: none;
  margin-left: 0;
line-height: 18px;
 margin-bottom: 2px;
  }
.my-legend ul.legend-labels li span {
  display: block;
  float: left;
  height: 16px;
  width: 30px;
  margin-right: 5px;
  margin-left: 0;
  border: 1px solid #999;
  }
.my-legend .legend-source {
  font-size: 70%;
  color: #999;
  clear: both;
  }
.my-legend a {
  color: #777;
  }
</style>
```

5. I took a screenshot of the map and uploaded it to this site http://www.color-blindness.com/coblis-color-blindness-simulator/ to check that it will be able to be easily viewed someone who is colorblind.
6. I uploaded the map to my MapBox account, restricting the zoom to Z7-Z9.
7. I copied the iframe code into the html for my website, adding add "scrolling=no" after "height" and adjusting height and width to fit properly on the webpage.