1. The output on my end shows:

> g1966@proj4string

CRS arguments:

+proj=utm +zone=12 +datum=NAD83 +units=m +no\_defs

In this case, Zone stands for an area denoted by the UTM projection, which divides the Earth into sixty zones. A projected coordinate system is basically defining the earth’s surface on a flat, tow-dimensional surface, and it has constant length, angles and areas across the two dimensions. In this system, locations are defined by x and y coordinates on a grid, denoting their relative horizontal and vertical distance from the center of that grid. Based on the info above, we can see that the spatial projection is meant for a specific area in North America, under the definition of The North American Datum of 1983.

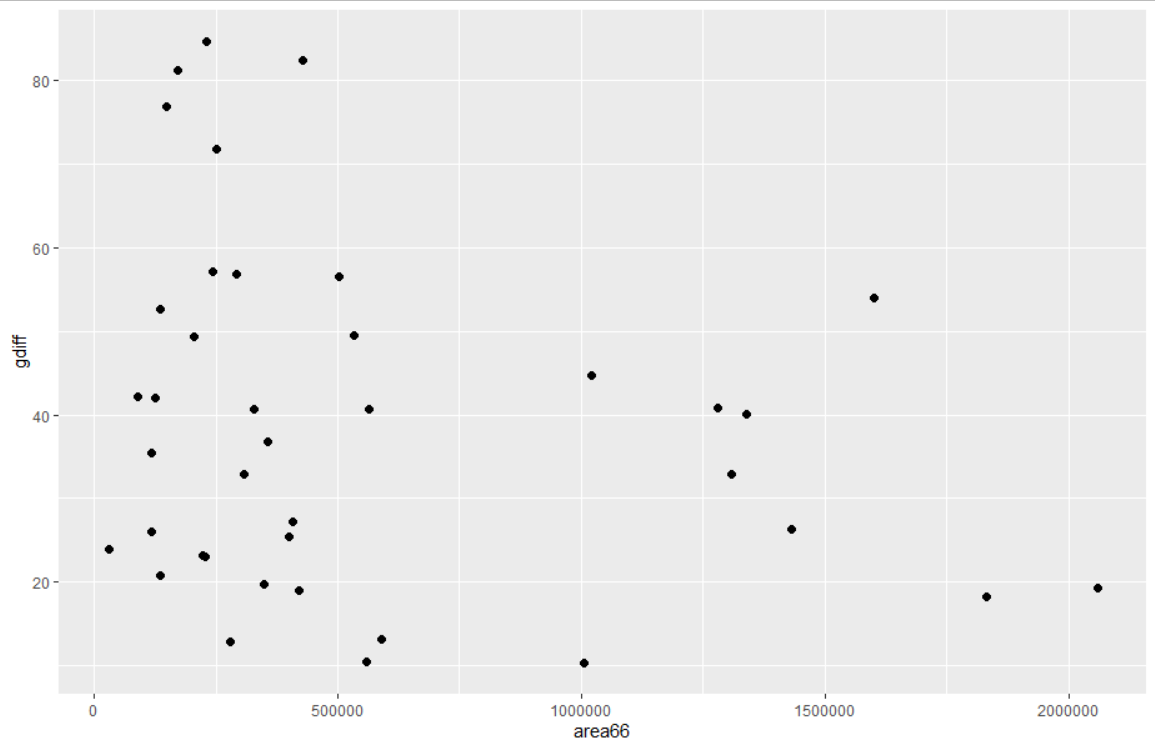
2. The background imagery is very green, meaning that the mountain range we are looking at has a lot vegetation. Most of the glaciers defined by the shapefiles are mostly surrounded by the white color from background, basically within the snow line. Most of the glaciers are located in the top right half, and there’s only small part that is close to the towns on the bottom left.

3. Given the graph is a bit hard to read, I think that there are some differences in glacial recession across the space we are investigating. Namely, at the top left corner, almost half of the glacier disappeared, but for the middle/lower area, most of the decrease of glacier seems to be not as severe.

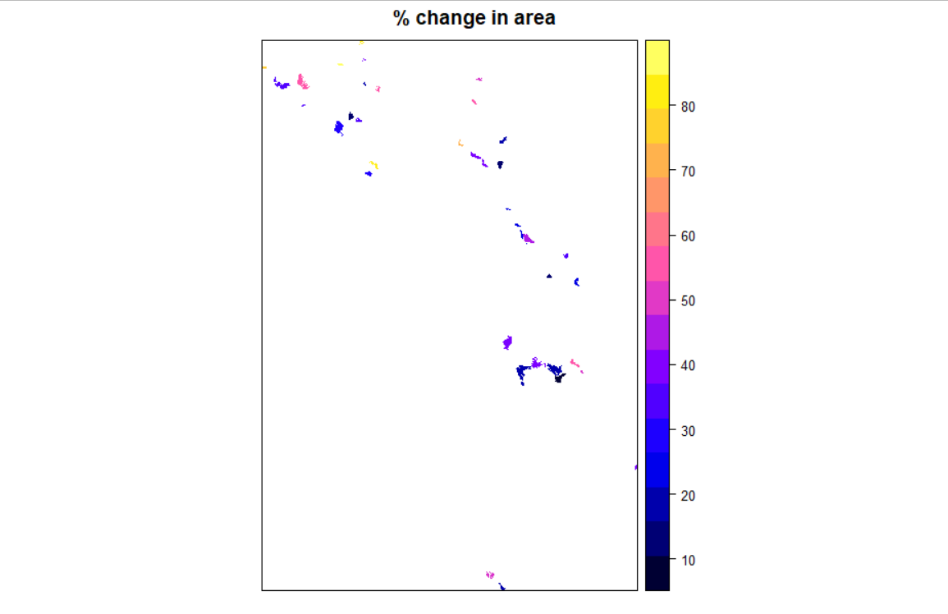
4. The top two small patch did not lose much from 1966 to 2015, and the most recession takes place on the center of the image. In the center portion, the top left glacier lost approximately 30% on the northern end; for the central part, recession took place around the glacier equilibrium line (those areas are still “white”, but most likely converted from accumulation zone to ablation zone), and overall seems like a 20% lost; the tiny strip to the right experienced more recession, and around 60% of the coverage is gone.

5. The new object is a data frame created by performing the full join operation on two prepared data frames, each contains the glaciers’ name and their area at the given time. Because it is a full join, all the data are remained, and because earlier on we made sure the 39 names are identical in 1966 and 2015 data, we should have all columns and rows perfectly aligned. The new data frame has 39 rows and 3 columns, which are the name, area in 1966, and area in 2015 respectively. If we did not conform the name columns and performed the full join, we will have 41 rows, where 2 of the rows will have 2 NA value under 1966 area and 2 rows having 2 NA values under 2015 area.

6. A join of two or more tables provides a means of gathering and manipulating data in a single table. Left join keeps everything in the “left” dataset, while adding the “right” data set with matching indices and leaving unavailable value NA; the right join is basically a reverse of left join, keeping right and map data on the left to it; full join retains all the data from two sets, and try to join the, by indices, leaving the ones unavailable to join as NA; inner join only keep the ones that match the given indices and left those not perfectly matching out of the new data set.

7. 

From the simple graph shown above, intuitively there are no correlations between the area of glaciers in 1966 and the percent recession of the glacier. So, I would claim that the percent glacial loss does not depend on the original glacier size in 1966. But to draw a more statistically convincing conclusion, I would first calculate the correlation coefficient between the variables, and if that results is not convincing enough, I would try to fit linear/log/quadratic…etc models to the plot to show that they cannot explain the variability of the data.

8. From the graph it east to see that the percent change in area varies in locations, but similar to my previous observation, the glaciers to the north seems to have more changes.

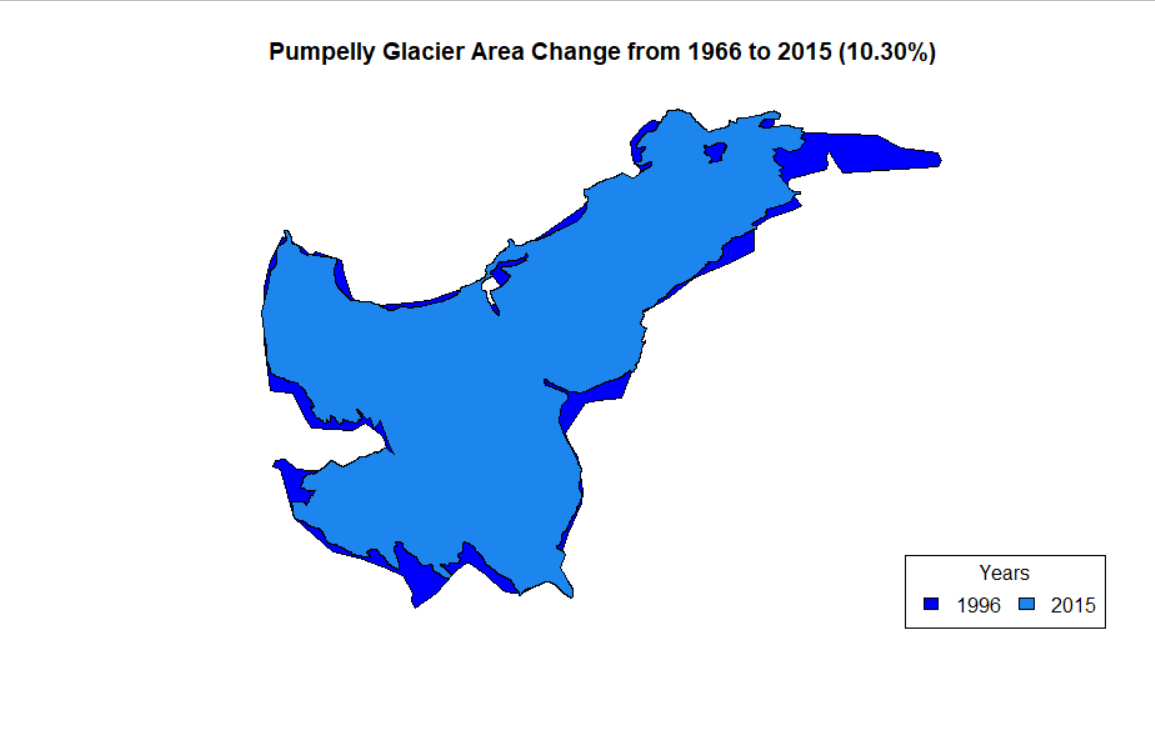
* Mean percent change: 39.02311%
* SD of percent change: 20.71525%
* Largest percent loss: Boulder Glacier, 84.72%
* Smallest percent loss: Pumpelly Glacier, 10.30%
* Largest area in 1966:

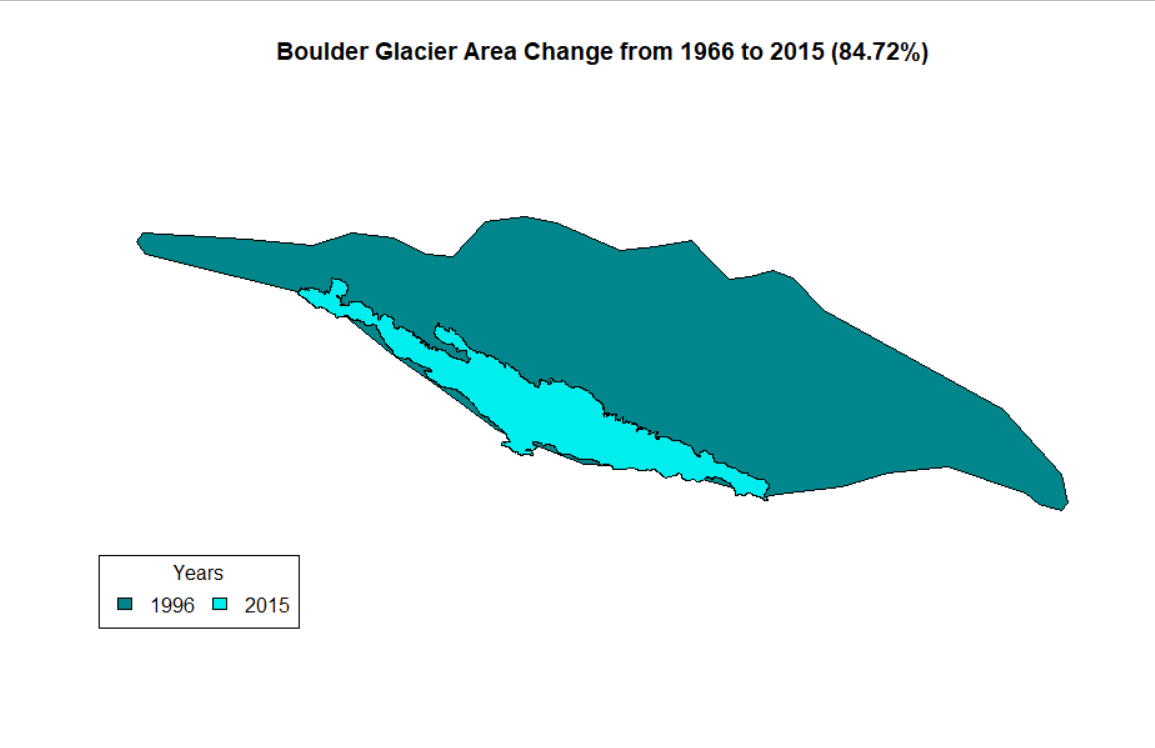
Harrison Glacier, 2059377, 19.32235%

* Smallest area in 1966:

Gem Glacier, 29140, 23.96504%

From the descriptive data above, it is easy to see that the change in glacial area is quite volatile in the Glacier National Park. The standard deviation of percent change is 20% with a mean around 40%. The largest change took place in Pumpelly Glacier where there are only around 15% of the glacier left, meanwhile, at Pumpelly Glacier, only 10% has disappeared. The largest area and the smallest area of glaciers all experienced some loss, but nothing drastic. There are no clear correlations between the original glaciers’ area in 1966 and the percent changes they experience till 2015.

9. 



10.

https://github.com/guozhaosengzs/ENVDS/blob/master/activity6/activity6\_script.R