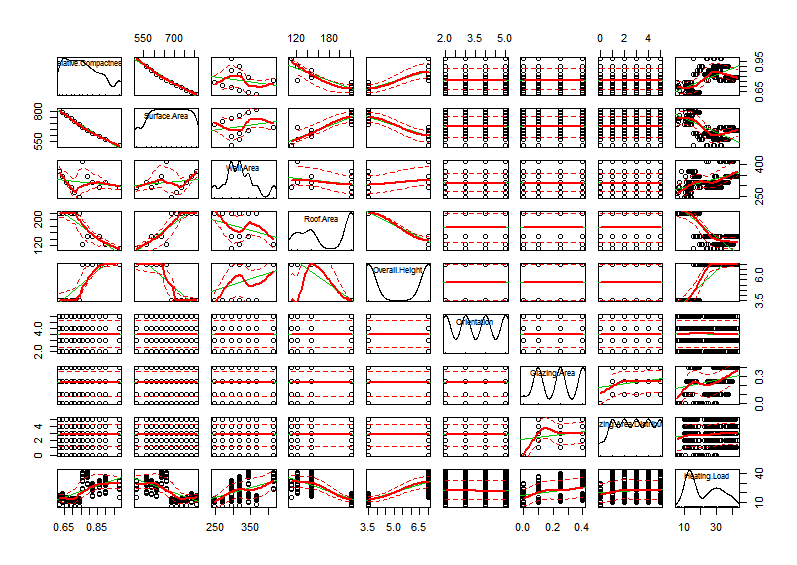
# UW Data Science – Methods of Data Analysis Dave Wine 8430191

## Summary

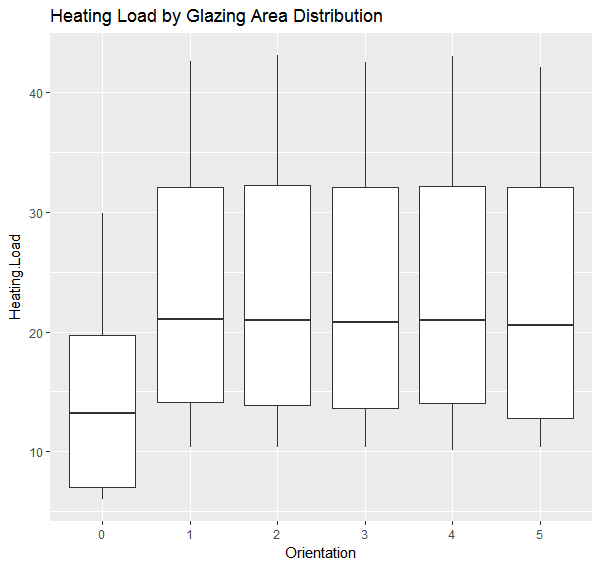
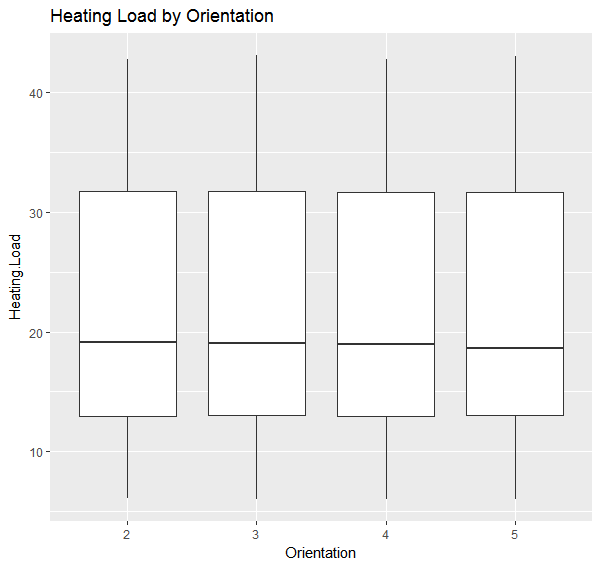
## Analysis

## Observations - Heating Load

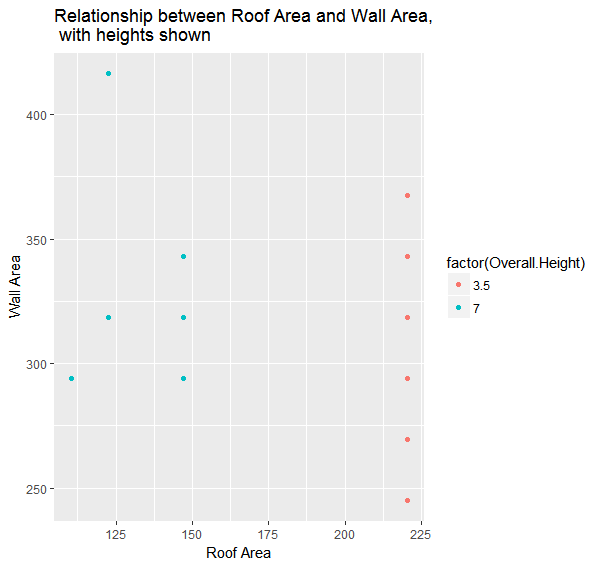
* Looked at overall correlations:



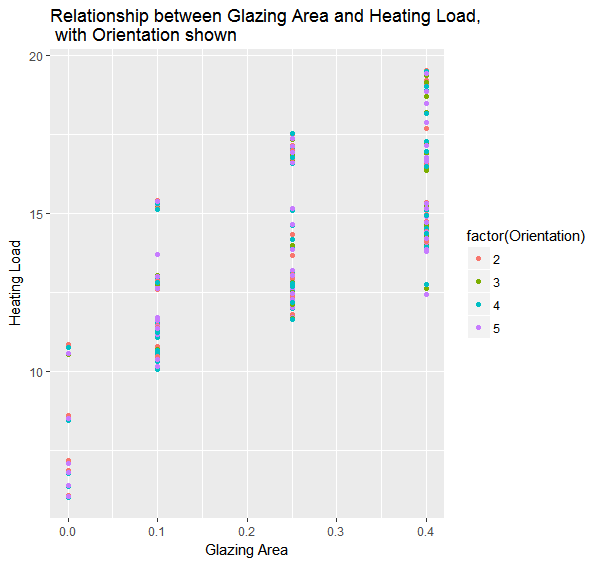
* Heating Load is obviously bimodal.
* Heating Load does not appear to depend strongly on Orientation or Glazing Area Distribution.
* Notice that Overall Height and Roof Area are inversely coupled to Wall Area.



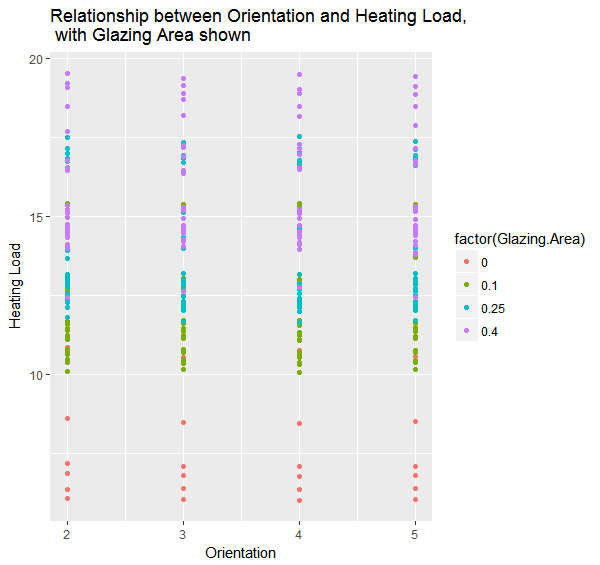
* Affected slightly by Glazing Area but it’s not the only factor.
* Some positive correlation with Overall Height, and negative correlation with Roof Area. Both are not surprising – tall buildings have more overall area, and tall thin buildings (low Roof Area) will not hold heat as well.
* Apparently, all short buildings have the same roof area:



* So if they are all the same height and have the same roof area (confirmed by summary statistics) then Wall Area must be the non-glazed area of the walls. The buildings are all the same size, but have different amounts of glazing. So plot heating load vs glazed area for short buildings only:

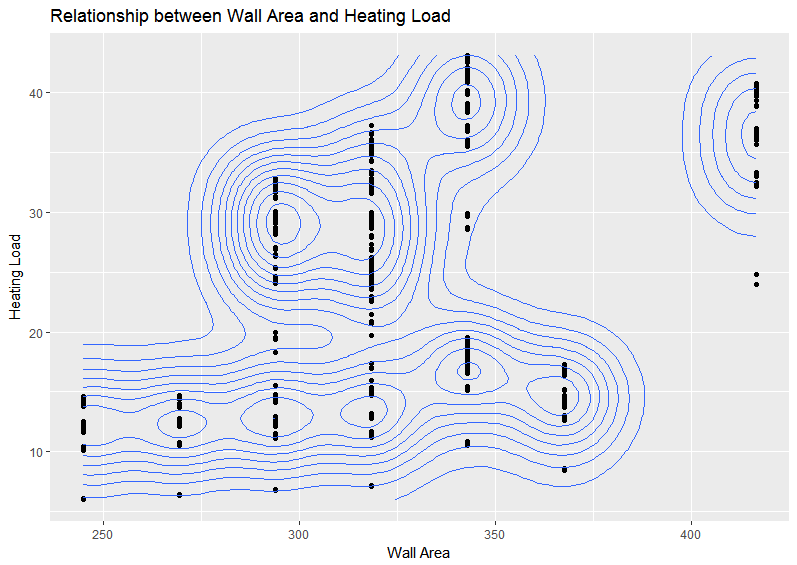


Hmm, try reversing the factors:

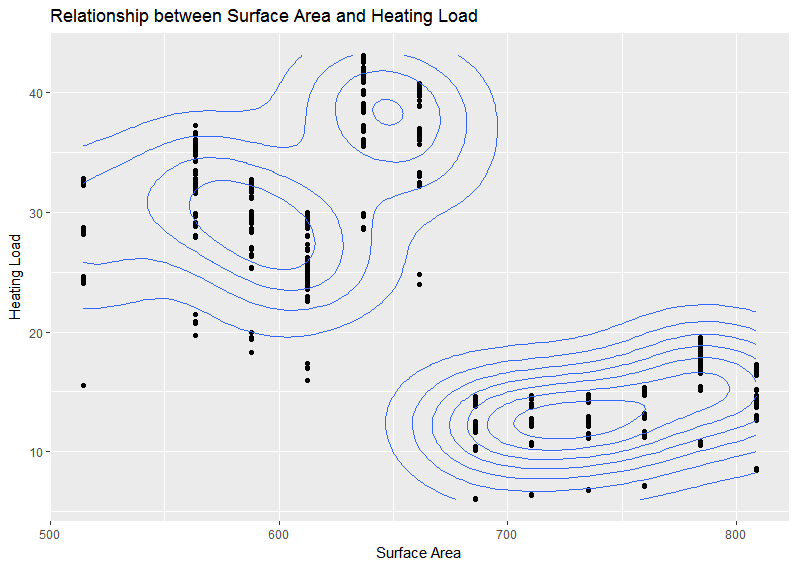


Ah ha! Now we can tell the difference between warehouses (no windows), old office buildings (some windows) and new office buildings (lots of windows) except t

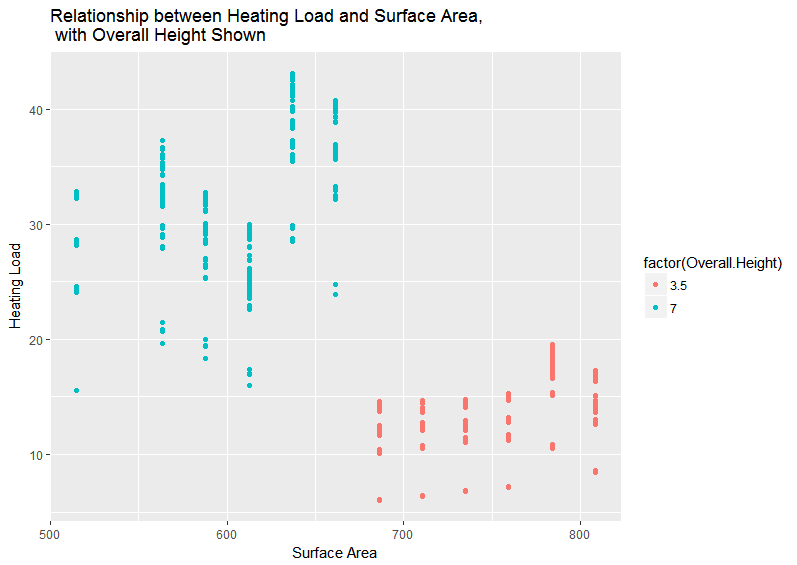
* Another unsurprising relationship is that relative compactness is inversely correlated to surface area
* Lots of confounding factors in Wall Area vs Heating Load



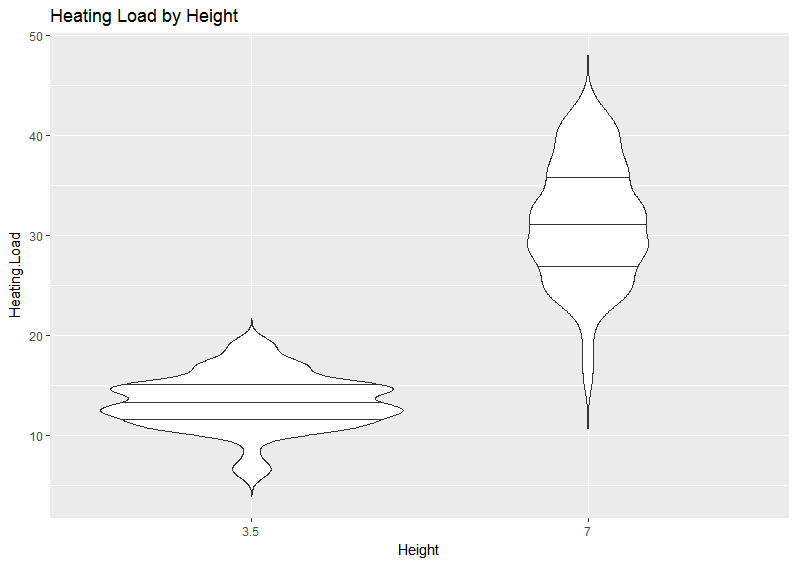
* The sum of (wall area + roof area) should be roughly equal to the surface area, so look at HL vs SA:



* So height is clearly important. Another way to look at the same data:

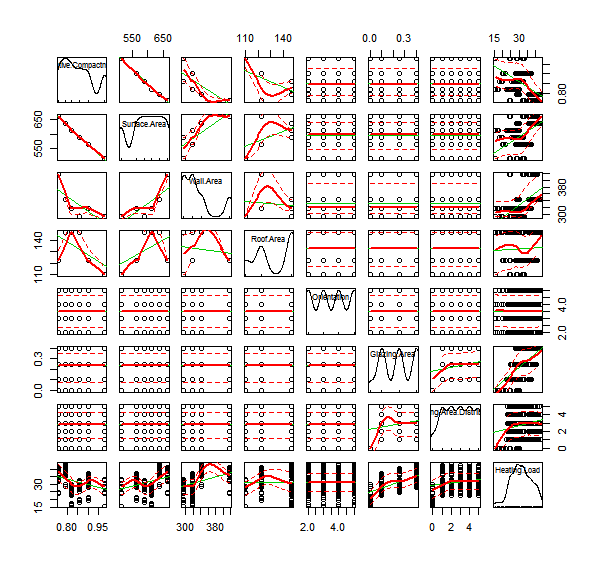


Tall buildings have in general lower surface area than shorter buildings. Then why does it take more energy to cool them? A violin plot shows the differences more clearly:

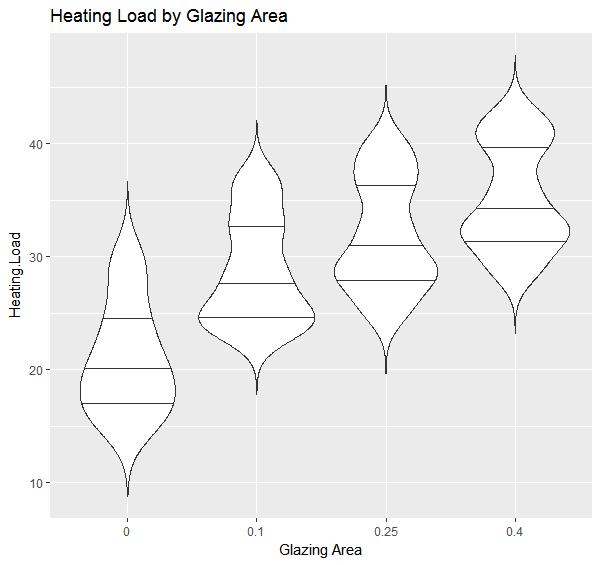


* It explains the difference in the Surface Area clusters and to a first order, the bimodality. The different shapes indicate that there may be differing factors at work between tall and short buildings. One obvious thing is that taller buildings generally will have a larger aspect ratio, changing the conduction and convection coefficients. Let’s compare

Look at Tall Buildings



Glazing area is the only obvious quasi-linear correlation:



This violin plot illustrates that there are probably a couple more factors inside, but it’s hard to se

The same is true for Cooling Load:

