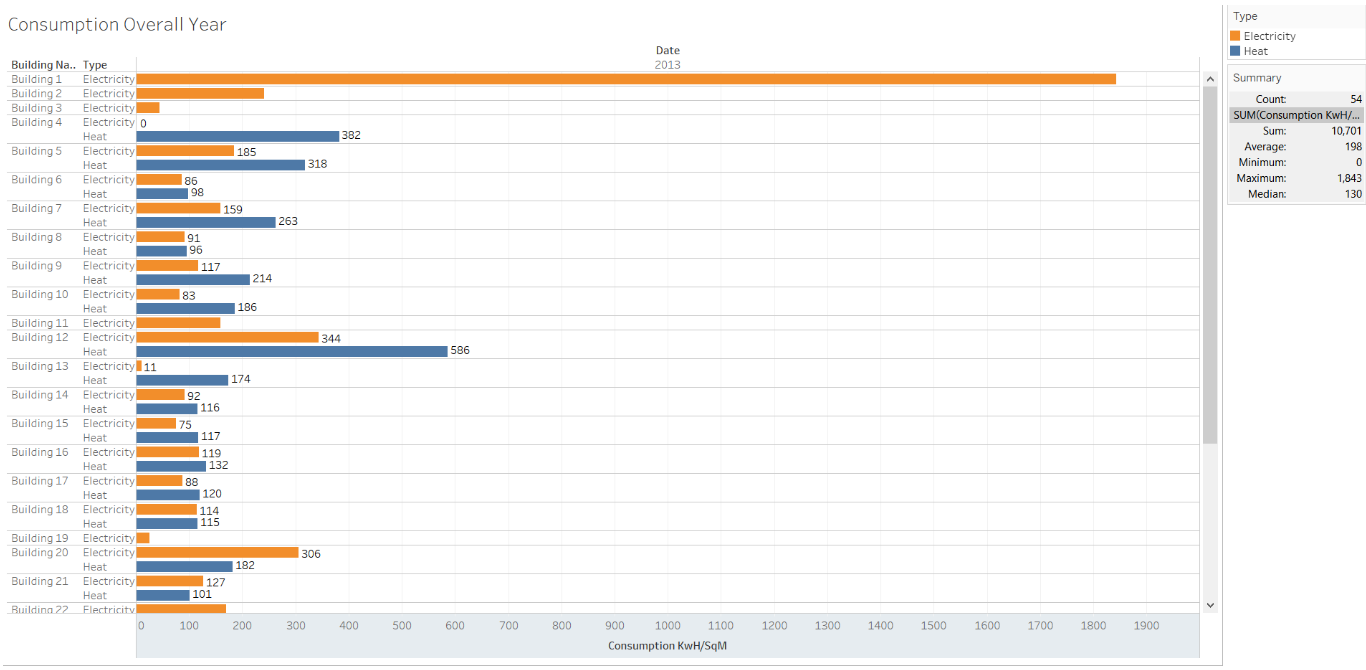
Part 3: Visualization and Dashboards

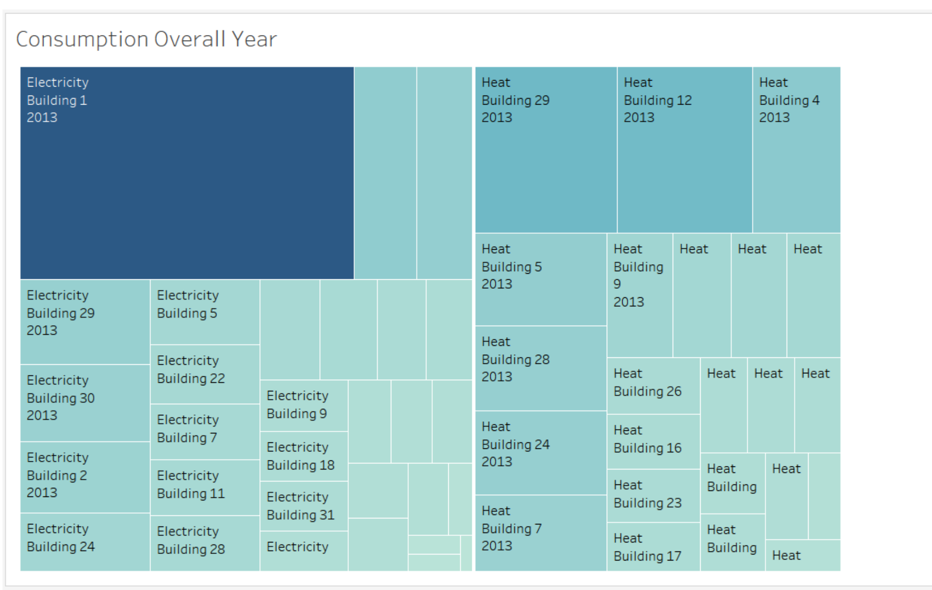
1) Overall consumption of each building with each type of electricity



Caption

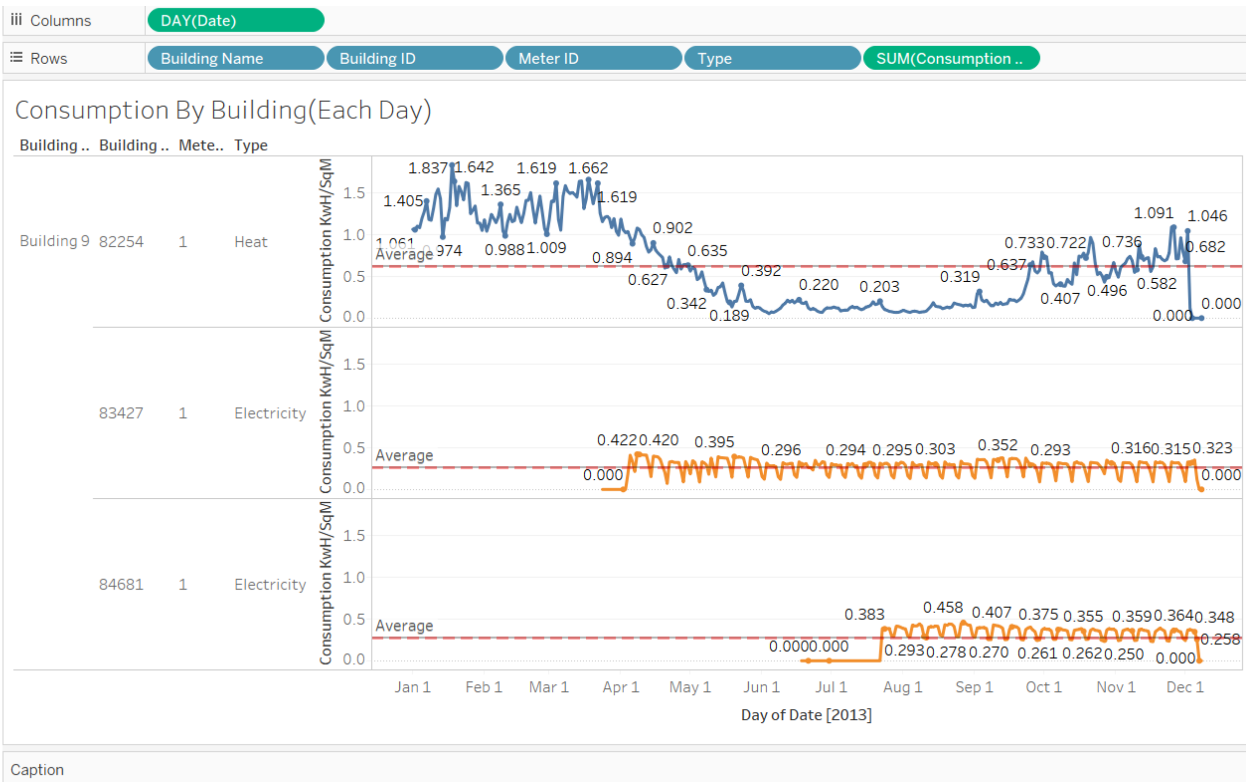
Sum of Consumption KwH/SqM for each Type broken down by Date Year vs. Building Name. Color shows details about Type.

We can see that Building 1 has consumed maximum energy in the year 2013 for Electricity. Consumption is measured with respect to KwH/ Sq.m. Same can be observed with heat map as shown below



2) Consumption of each building throughout the year

The trend of sum of Consumption KwH/SqM for Date Day broken down by Building Name, Building ID, Meter ID and Type. Color shows details about Type. The data is filtered on Action (Building Name,YEAR(Date),Type), which keeps 2 members. The view is filtered on Building Name and Type.

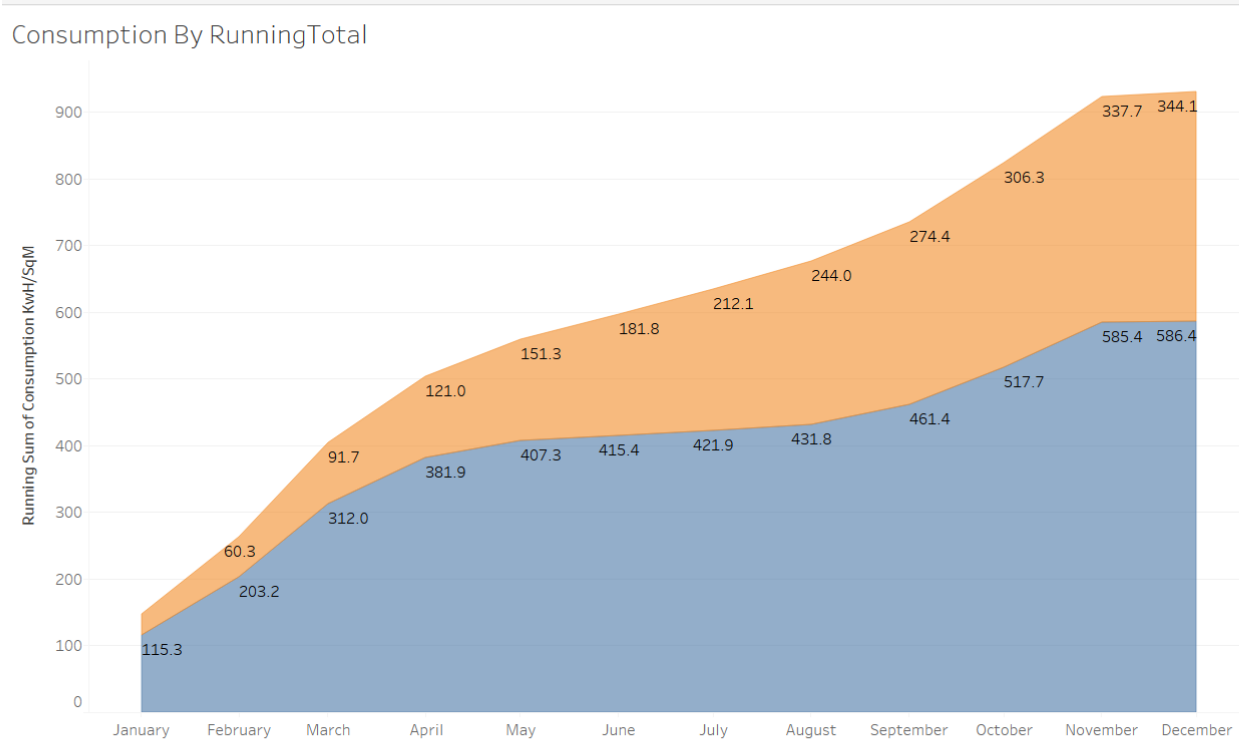


From this graph, we can see analysis of each Building and see there trend on consuming energy. It is very useful to see if any building has changed the electricity plan and if any line was aborted due to some reason or not. In this case, we can see that Building 9 which has 3 meters, has been using meter 1 for heat consistently. However, electricity was not supplied explicitly to the building until April 1, 2013 (which could probably be the official year to initialize the plan)

3) Consumption by Running Total

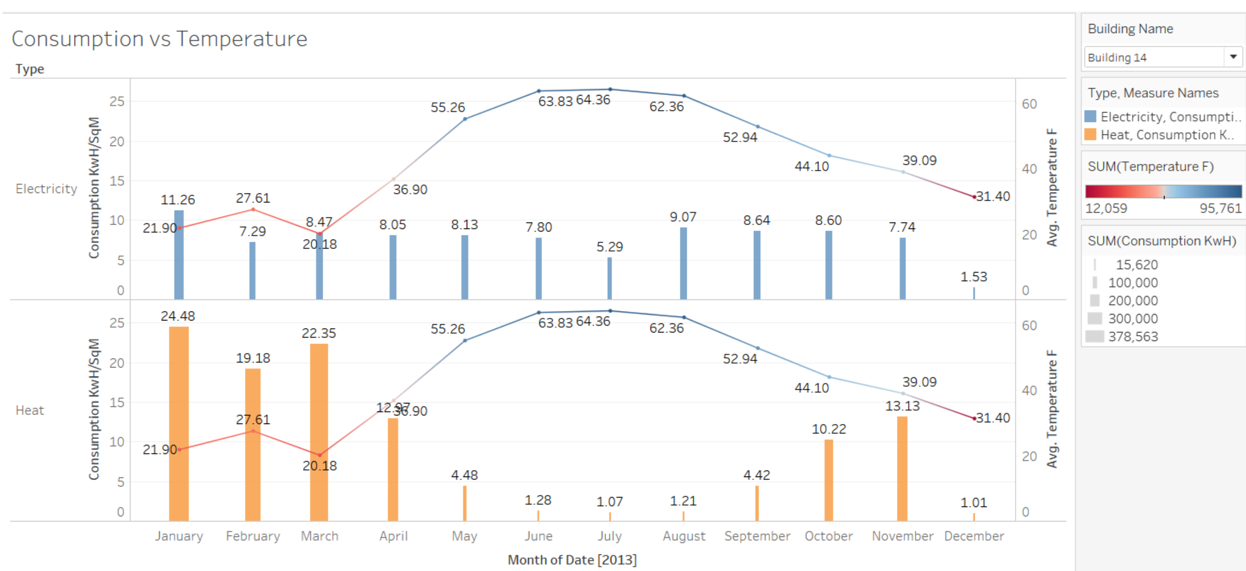
The plot of Running Sum of Consumption KwH/SqM for Date Month. Color shows details about Type. The data is filtered on Building Name, which keeps Building 12.

We can analyze the running consumption total for each building using the Area graph and distinguishing with electricity type we can see the which type of electricity has been consumed the most



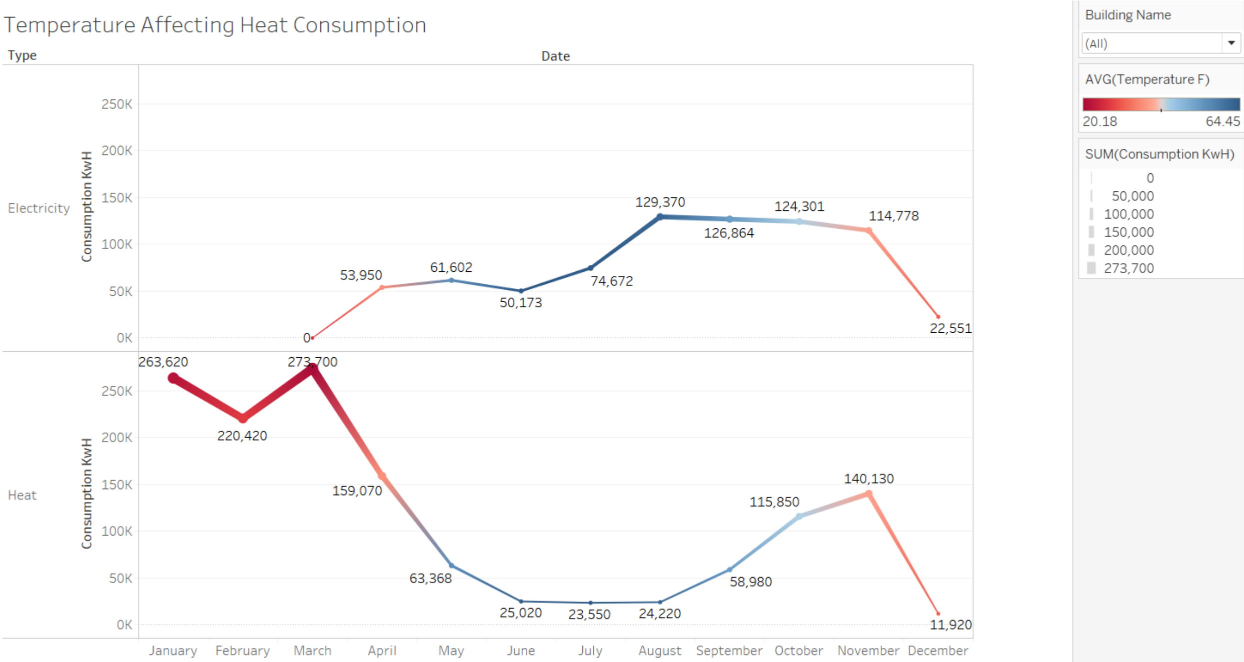
4) Does temperature affect the consumption of Building?

It can be very useful to see the major factor which is causing the consumption of the electricity. We can pot temperature as a measure and see the effect on consumption for a particular Building



It can be clearly seen that Heat has huge impact based on the temperature. There is a rise in 1st quarter for heat consumption which makes sense due to cold temperature. It is interesting to see that electricity consumption however has no significant effect due to Weather conditions

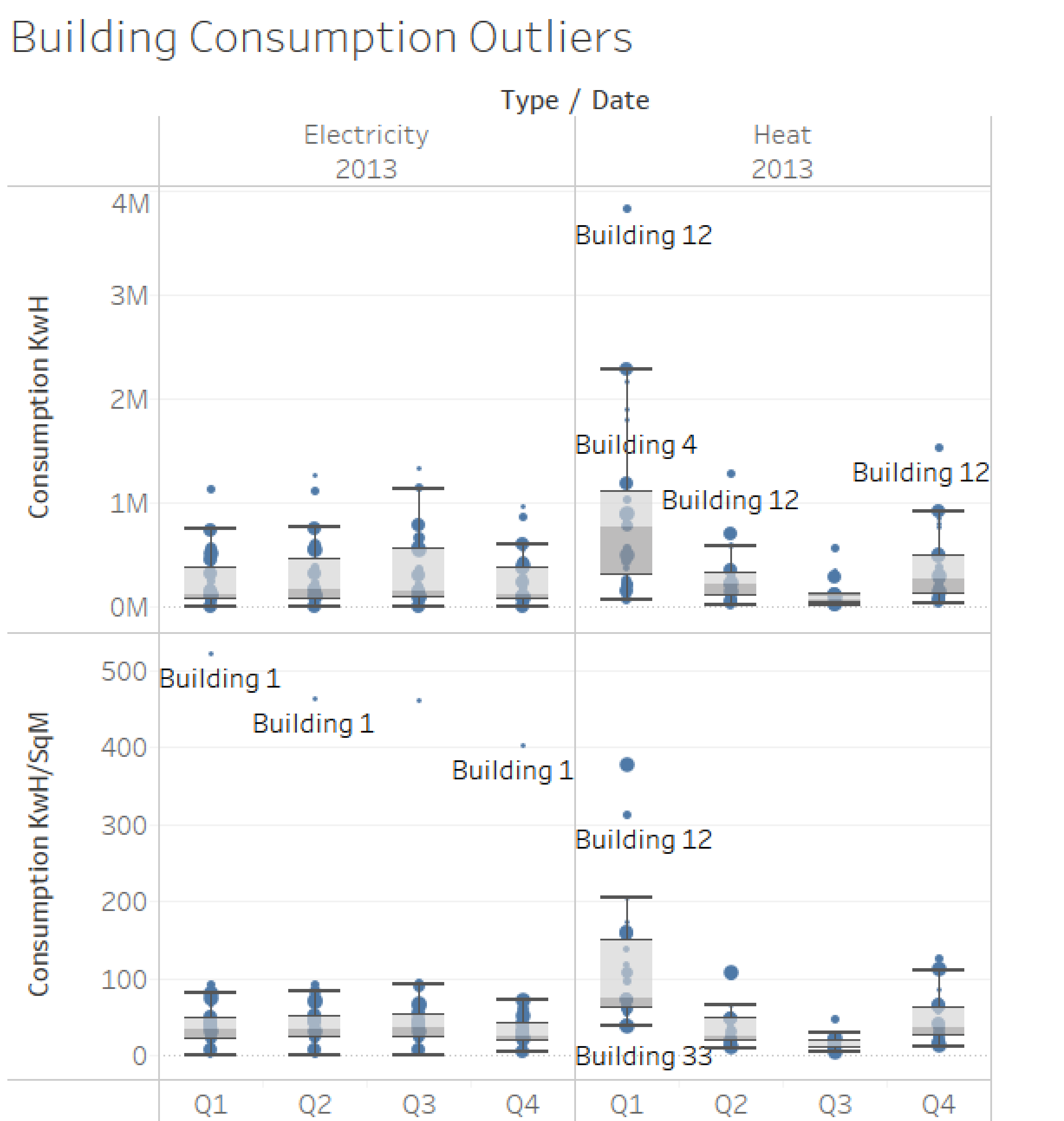
The same can be observed with a different graph as shown below



5) Does any Building consume more energy than other's does?

It can be very interesting to see how other building's are contributing to the consumption of energy. Are they any major contributors or not.

A boxplot representation for this case could be ideal. A plot can be seen below

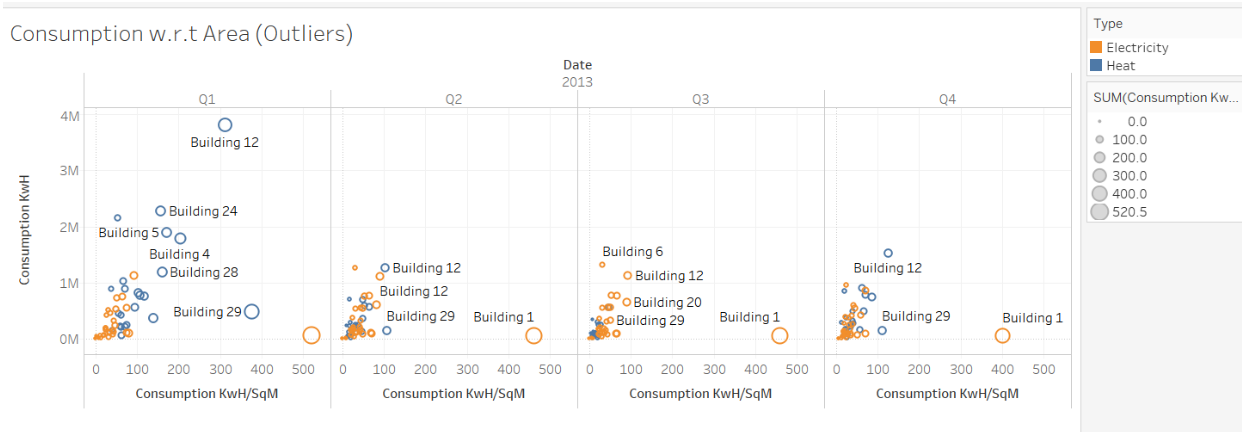


There are two very distinct observations noted here

1) If we consider the first half of the graph, where we have plotted Consumption in KwH for all the Buildings, it is clear that there are not major outliers in terms of Electricity. However, on the other hand Building 12 has been using the heat consistently higher. Looking at the address of Building 12, we can only see that it is located very close to coastal area and can have significant impact during the winters.

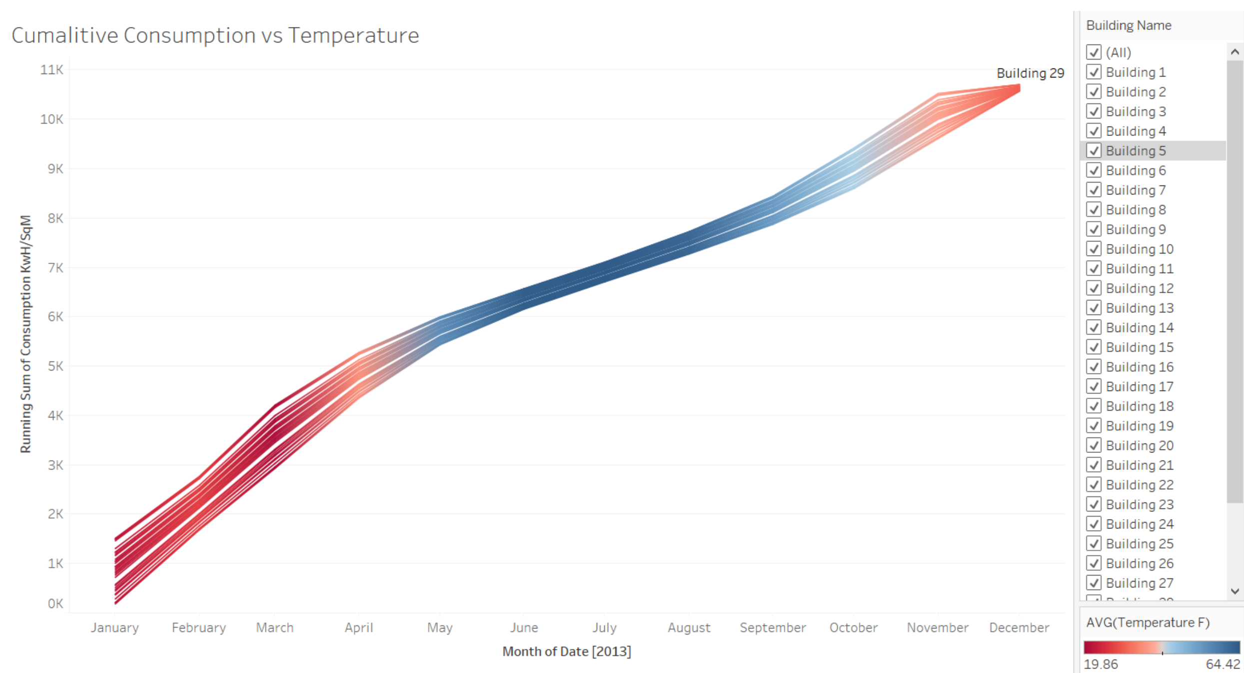
2) On the other hand, if we consider consumption per square foot area, data says something else. Building 1 is consuming lot more energy than its counterpart considering the size of the Building. Building 1 has an area of 110 sq.m. but the consumption of the building is 57,254 Kwh during Q1 of 2013 as compared to other buildings which has significant area coverage. A reason for this behavior can be due to the fact that Building 1 is solely running on electricity and has no direct supply of heat.

Other way to represent the same information in a more organized manner can be as shown below



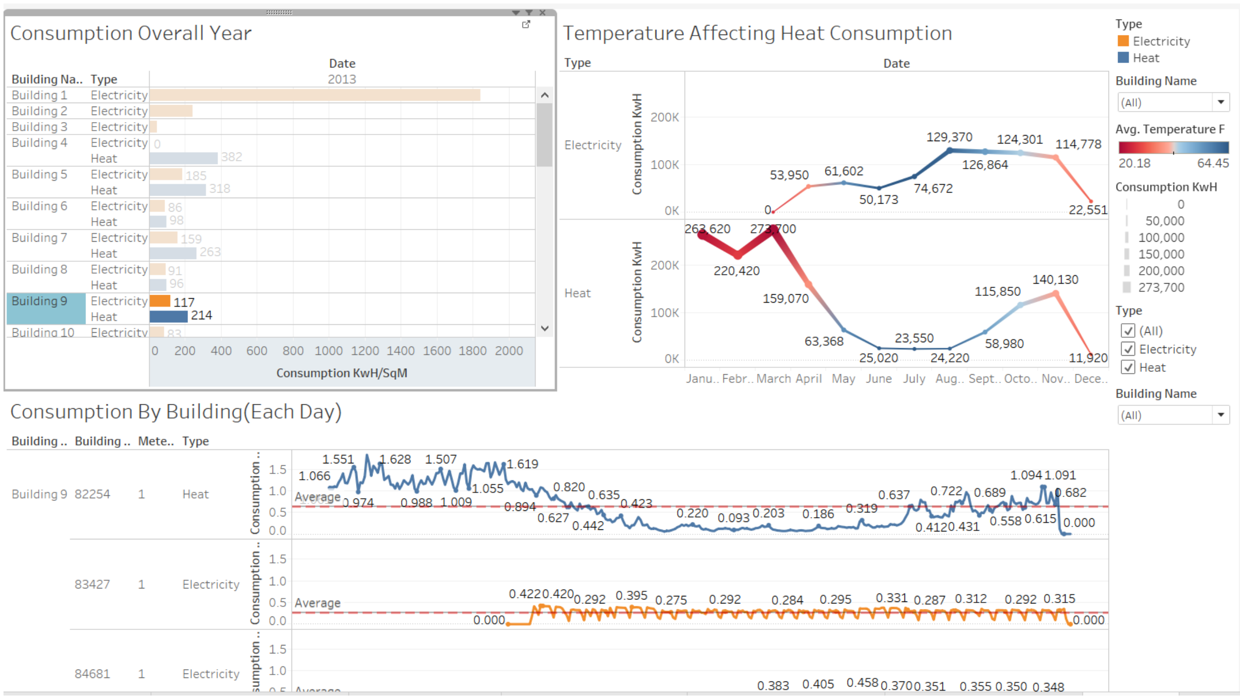
6) Cumulative consumption of Buildings w.r.t to temperature

We can see the running sum of consumption in Kwh/sq.m. vs temperature for all the Buildings.



7) Dashboard

Here we can see overall analysis of the building's in one graph. Filter is implemented on each graph to see the overall effect of Finland Energy consumption



R-Serve integration

We have used R-server as an integration component to see the results of R modelling in visual representation

Predictive Analysis

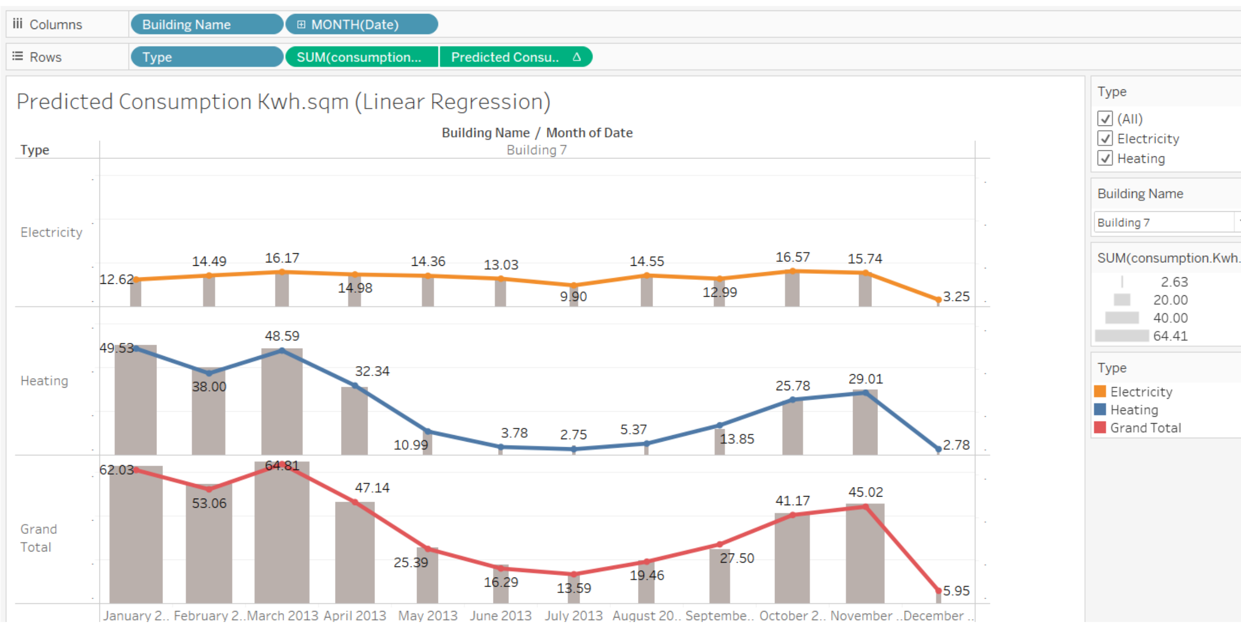
We have used Linear regression to predict the consumption of the energy for the Building. R-script calculated field for Linear regression can be seen below

Predicted Consumption

Results are computed along Table (across).

```
SCRIPT_REAL("
fit <- lm(.arg1 ~ .arg2 + .arg3 + .arg4 + .arg5 + .arg6 + .arg7 + .arg8 + .arg9)
fit$fitted",
SUM([consumption.Kwh.sqm]),
SUM([Weekday]),
SUM([Holiday]),
SUM([Base Hr Usage]),
SUM([Temperature F]),
SUM([Dew PointF]),
SUM([Wind SpeedMPH]),
SUM([Wind Dir Degrees]),
SUM([Base Hour Flag]))
```

A plot showing the actual consumption and the predicted consumption is as shown below



The trends of sum of consumption.Kwh.sqm and Predicted Consumption for Date Month broken down by Building Name vs. Type. Bar shows the Actual consumption and the lines shows the Predicted consumption per sqm. Error rate is very low and adjusted R square for linear regression is pretty good (0.8912).

Classification Analysis

We can similarly leverage the R functionality to predict the Base Hour class classification using Logistic regression. R-script code written for Logistic regression is as shown below

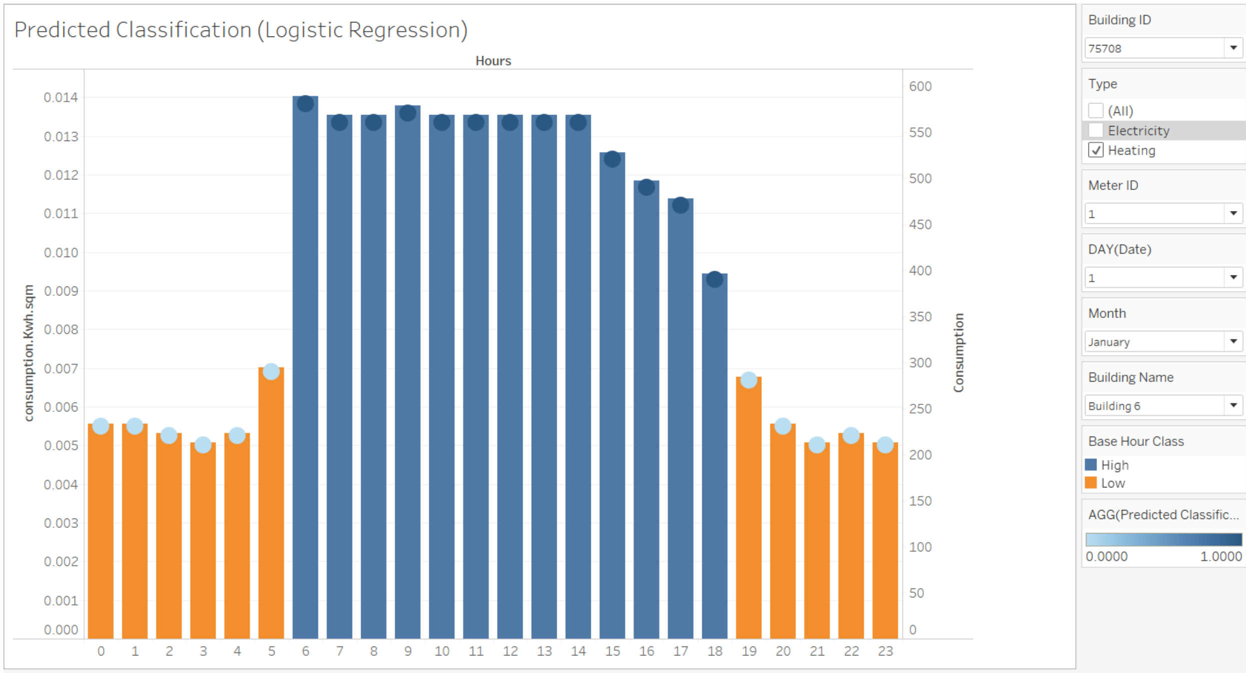
Predicted Classification

Results are computed along Table (across).

```
SCRIPT_REAL('mydata <- data.frame(
Base_Hour_Class=.arg1,
Hours=.arg2,
consumption.Kwh.sqm=.arg3,
Weekday=.arg4,
Base_Hour_Flag=.arg5,
TemperatureF=.arg6,
Dew_PointF=.arg7,
Wind_SpeedMPH=.arg8,
WindDirDegrees=.arg9,
base_hr_usage=.arg10,
Consumption=.arg11
);

logit <- glm(Base_Hour_Class ~ Hours + consumption.Kwh.sqm + Weekday + Base_Hour_Flag
prob <- predict(logit, newdata = mydata, type = "response")',
AVG([Base Hour Class_Numeric]),AVG([Hours]),AVG([consumption.Kwh.sqm]),AVG([Weekday]),
AVG([Base Hour Flag]),AVG([Temperature F]),AVG([Dew PointF]),AVG([Wind SpeedMPH]),
AVG([Wind Dir Degrees]),AVG([Base Hr Usage]),AVG([Consumption]))
```

Classifying the actual base hour class and the predicted base hour class is as shown below. We have filtered down the results of the actual base hour class to a grain level of Hour for a particular Building for a particular day as shown in the filter section to the right of the plot. Bar chart shows the consumption with Blue indicating actual High and Orange indicating actual Low as highlighted in the legend section.

Predicted classification is symbolized with circles on top of the bar. Since Logistic returns a numeric value in the range of 0 to 1, it is shown as a measure. As we can see the predicted and actual is very accurate, with Logistic regression being trained to an accuracy of 95%.



Clustering.

We can cluster our building records using K-means algorithm. A cluster of 3 is selected as an optimum cluster with high within sum of squarer error.

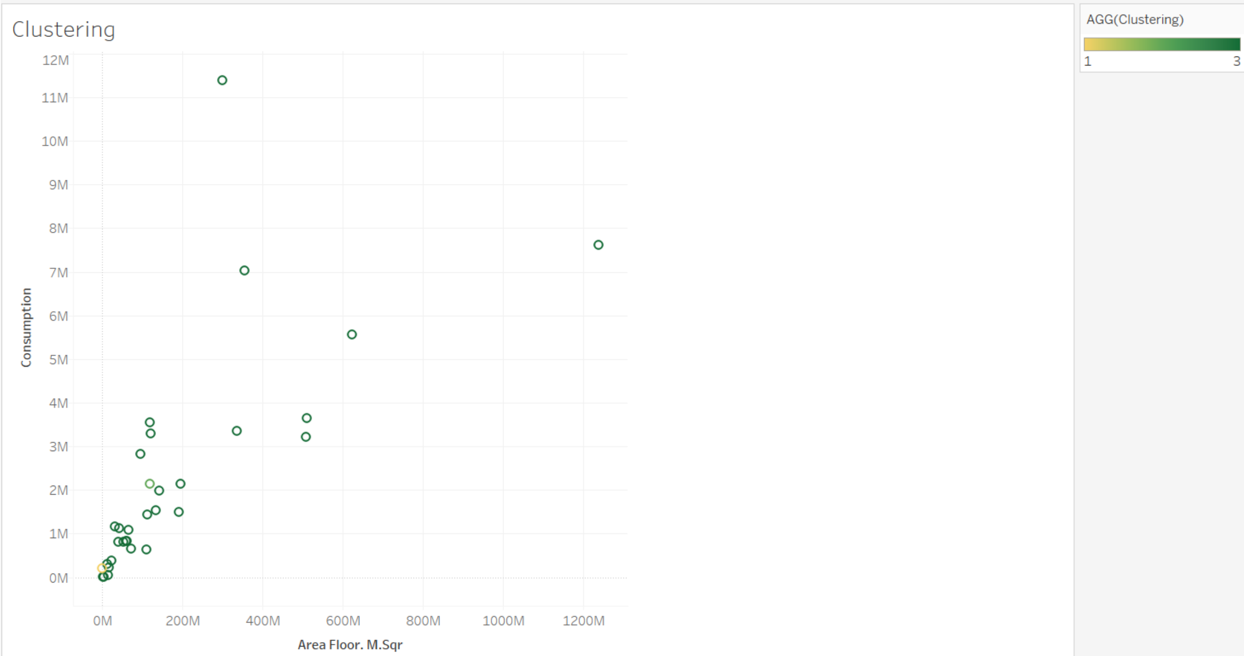
Code for writing k-means clustering with R-server calculated field is as shown below

Clustering

Results are computed along Table (across).

```
SCRIPT_INT("
    ## Sets the seed

    ##set.seed( .arg16[1] )

    ## Scale the variable

    h <- scale(.arg1,center = TRUE, scale = TRUE)
    c <- scale( .arg2,center = TRUE, scale = TRUE)
    wkday <- scale( .arg3,center = TRUE, scale = TRUE)
    baseflag <- scale( .arg4,center = TRUE, scale = TRUE)
    holiday <- scale( .arg5 ,center = TRUE, scale = TRUE)
    tmp <- scale( .arg6,center = TRUE, scale = TRUE)
    df <- scale( .arg7,center = TRUE, scale = TRUE)
    hmd <- scale( .arg8,center = TRUE, scale = TRUE)
    sealevel <- scale( .arg9 ,center = TRUE, scale = TRUE)
    wsph <- scale( .arg10,center = TRUE, scale = TRUE)
    windir <- scale( .arg11,center = TRUE, scale = TRUE)
    usage <- scale( .arg12,center = TRUE, scale = TRUE)
    area <- scale( .arg13,center = TRUE, scale = TRUE)
    consump <- scale( .arg14,center = TRUE, scale = TRUE)
    dat <- cbind(h,c,wkday,baseflag,tmp,df,hmd,sealevel,wsph,windir,usage,area,consump

    ##num <- .arg15[1]

    ## Creates the clusters

    kmeans(dat, 3)$cluster
",

AVG( [Hours] ), AVG( [consumption.Kwh.sqm] ), AVG( [Weekday] ),
AVG( [Base Hour Flag] ), AVG( [Holiday] ),
AVG( [Temperature F] ),AVG( [Dew PointF] ),
AVG( [Humidity] ),AVG( [Sea Level PressureIn] ),AVG( [Wind SpeedMPH] ),
AVG( [Wind Dir Degrees] ),AVG( [Base Hr Usage] ),
AVG( [Area Floor. M.Sqr] ),AVG( [Consumption] )
)
```

Clustering is as shown below in terms of Area.Floor and Consumption, the two main driving factors for clustering



We can see the buildings are placed in different clusters based on area floor clearly.