INFO 7390

**Assignment 3 Report**

Processor: Sri Krishnamurthy

Group 1

Chang Yu, Songzhe Zhang, You Li

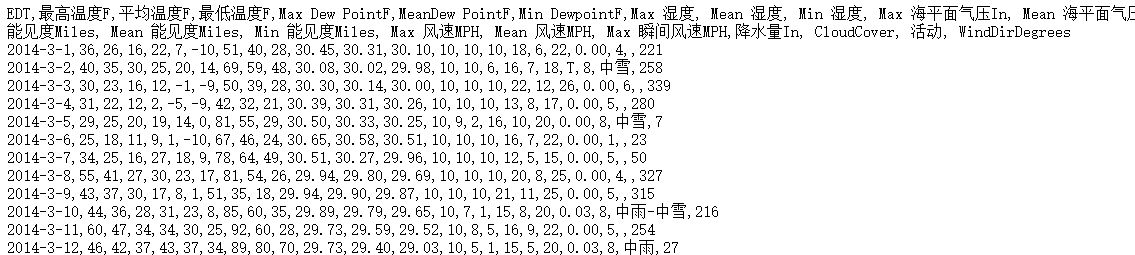
# **1.Pre-process**

1.1 Transform the Structure of the data

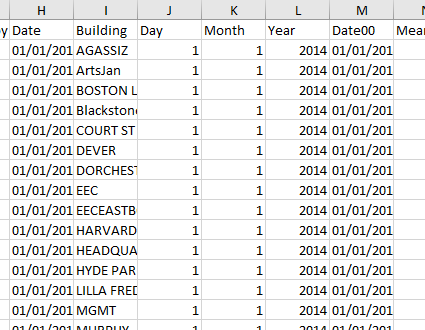
. 1.x Combing Columns

Firstly, we import the data from: <https://www.wunderground.com/>

In the type of csv format;

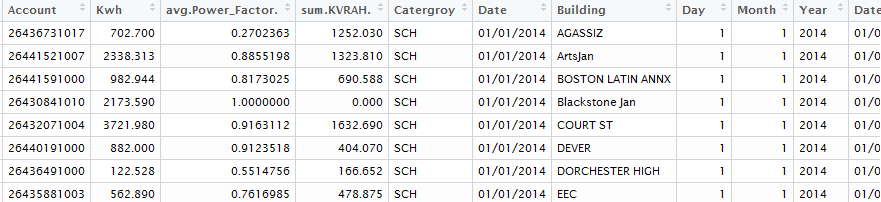


Then we transform the date data in to the same type of that in Boston Energy and combine them together in R-sql





After we collect weather info from web source, we using sqldf() function combine different functions together.



1.y Clean NA and Zero Value Data

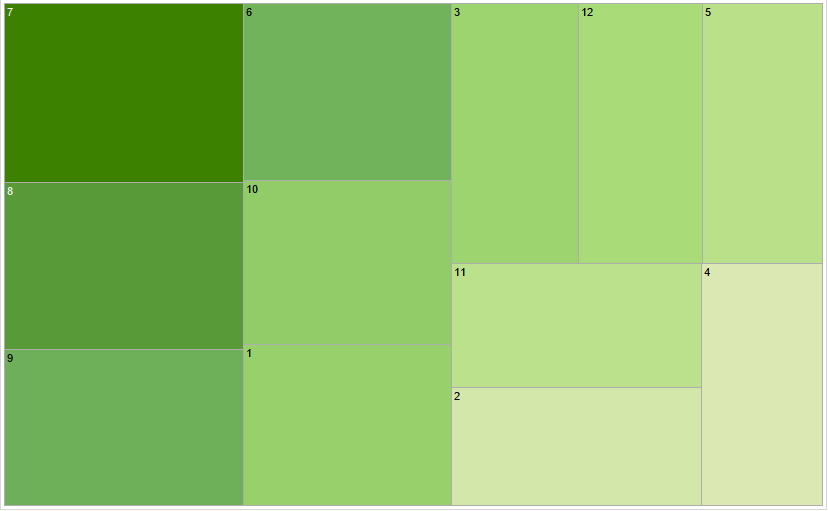


# **2.Visualization Analytic**

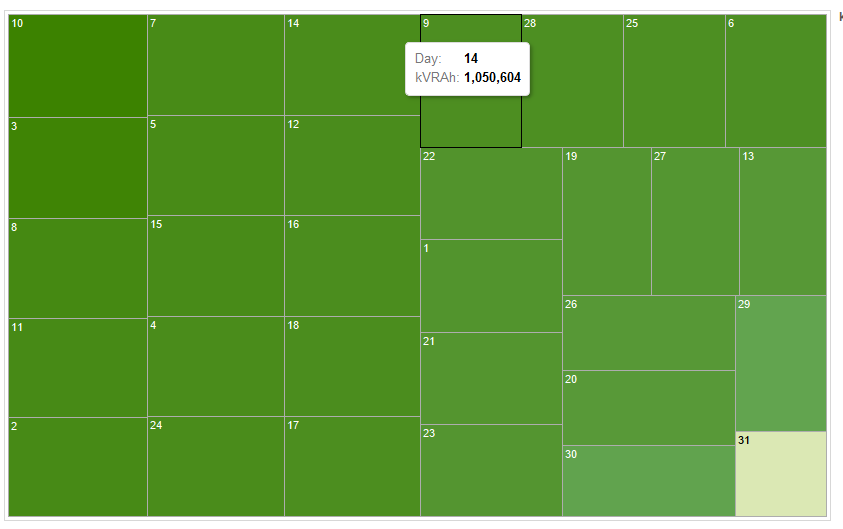
2.1 Marco Analysis

Total Energy Consumption – Month/Day

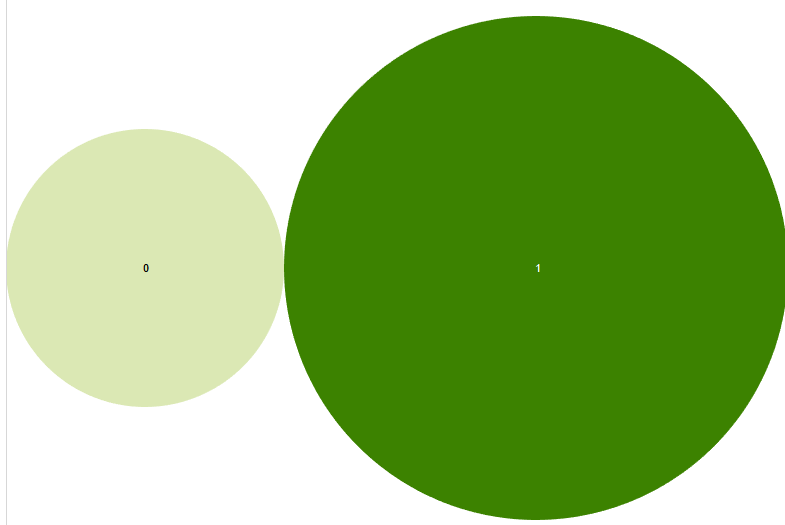
In the graph, it’s very clear that the macro energy consumption changes month by month. This will help us understand the relationship between different month and their energy consumptions.



The graph below displays the range of daily usage with day’s change in a month.

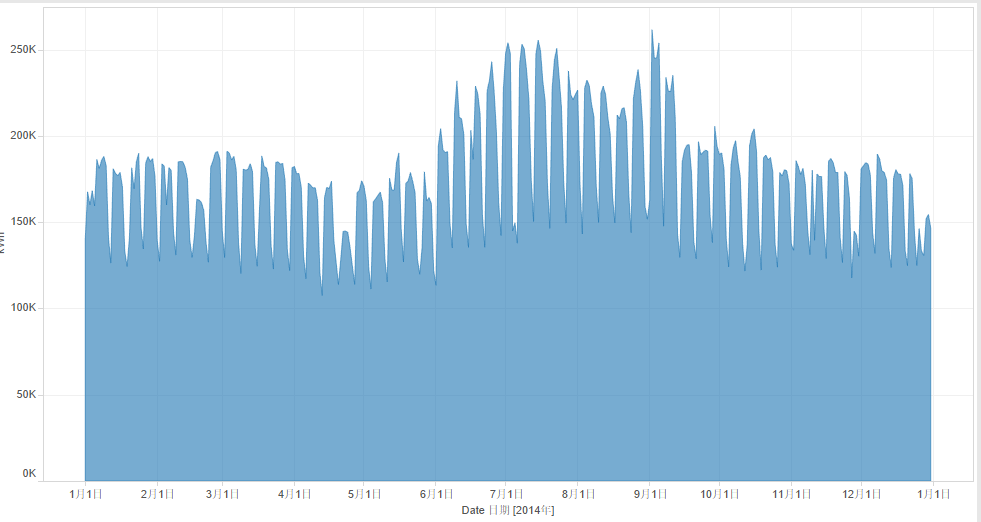


The Marco Energy Consumption with working day and no-working day

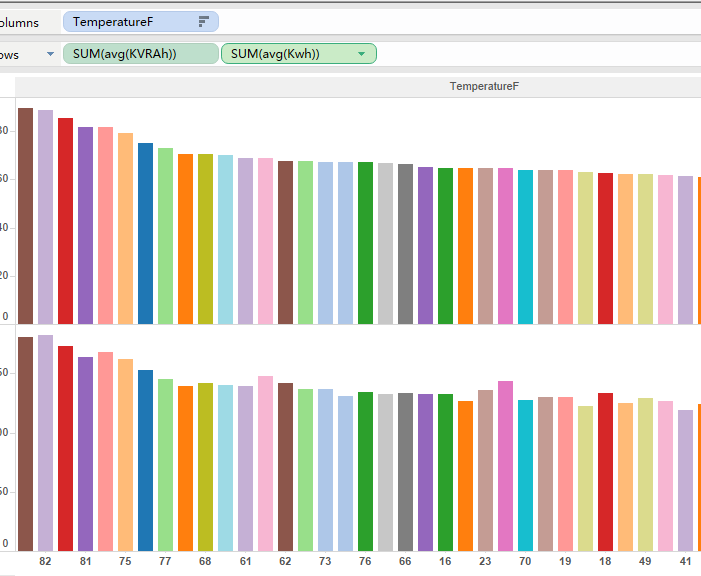


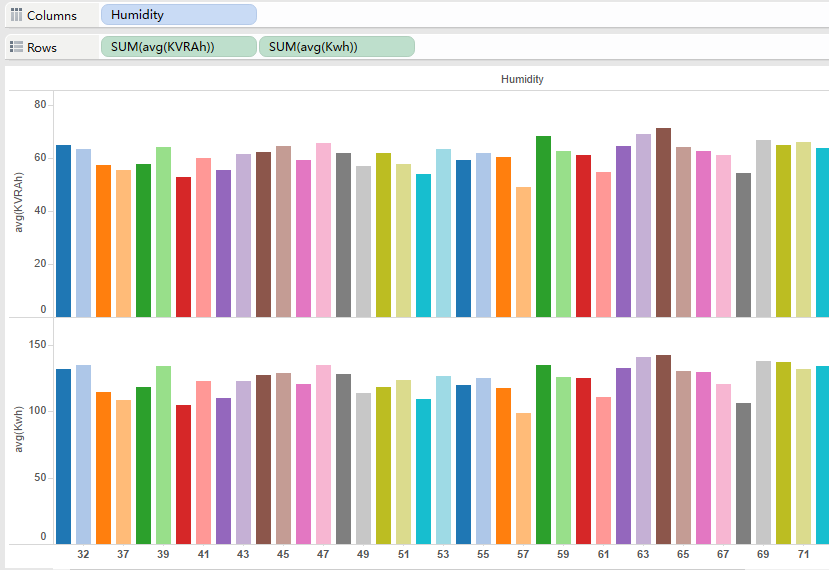
Energy Consumption – Total Value Through the Year

According to the graph we can easily distinguish the changes from month to month in one year, and changes from day by day in each month. The next step for our analysis is to get the whole trend of energy consumptions in the whole year on the unit of date.



Marco Energy Consumption – Temperature/Humidity

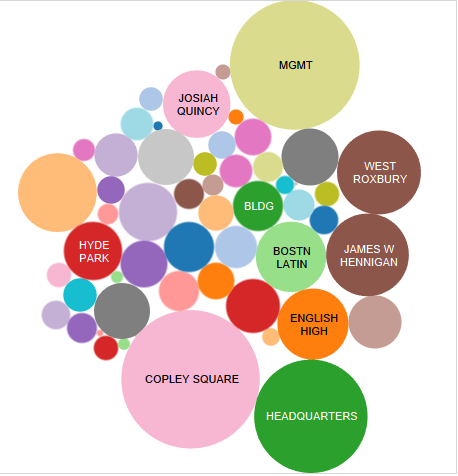




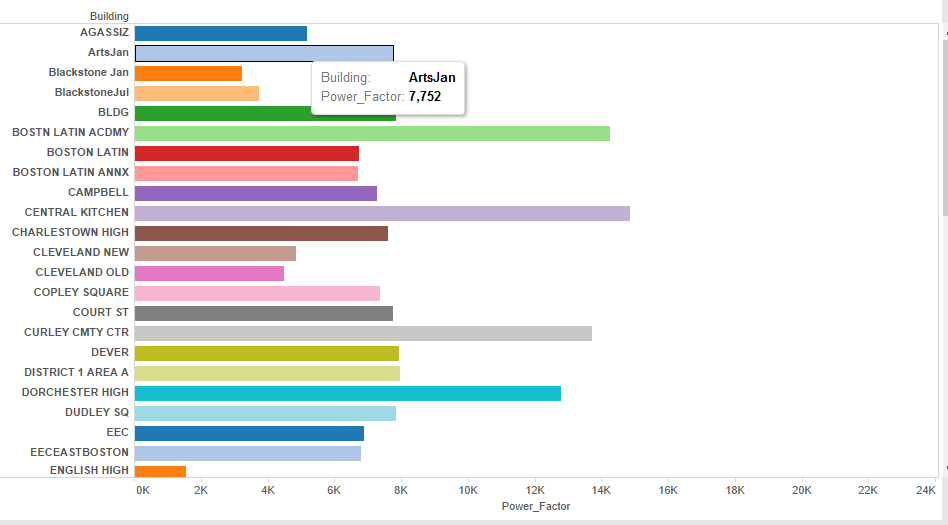
Micro-Analysis

After analysis the Marco trend of energy consumption, we now focus more on micro- analysis and their energy consumptions related to temperature, humidity and so on.

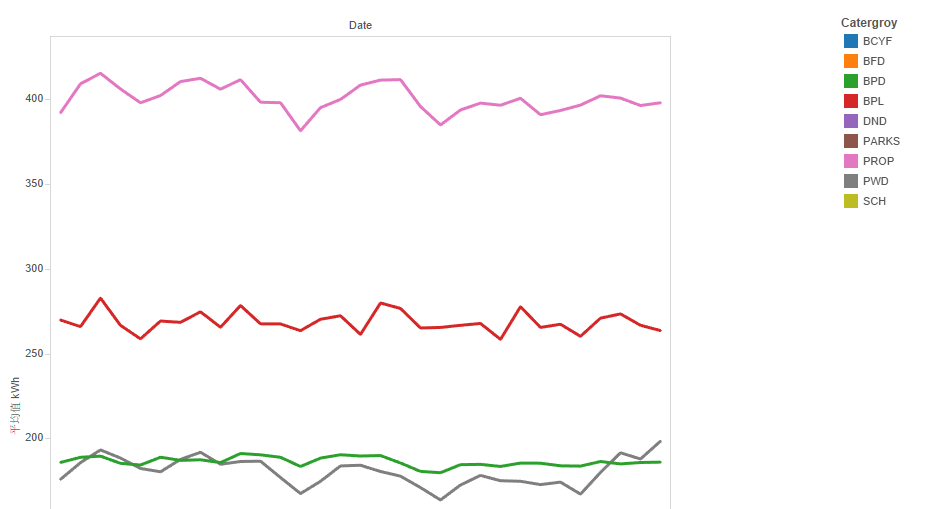
First diagram shows the energy consumption differs area by area.



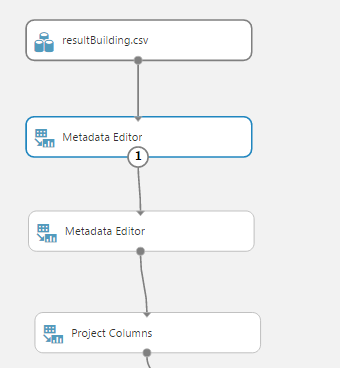
The box-graph also offer us a clear visual for building’s power-factors.



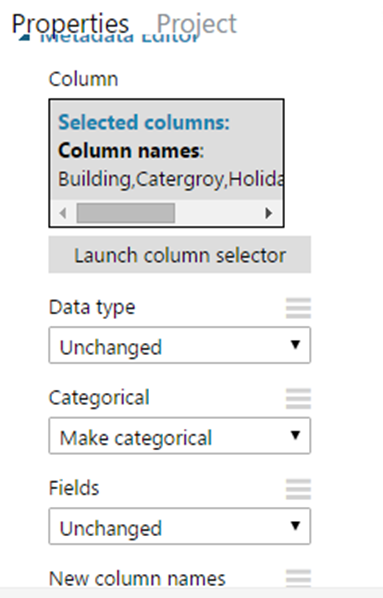
There is also a diagram display the use of electricity of each day, though difference exit in each building, while compare to themselves, the graph seems to be stable.



# **3.Model Building**

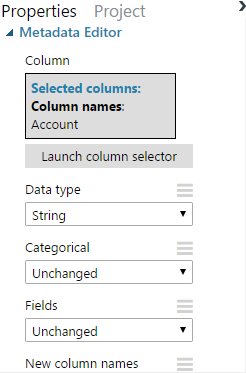


3.1 Transform and pre-clean the data for analysis



3.2 Define category variables

For this exercise, the column that should be cast as categorical are: Building, Category, Holiday, working Day, Events. To cast these columns, drag in the metadata editor. Specify the columns to be cast, then change the “Categorical” parameter to “Make categorical”



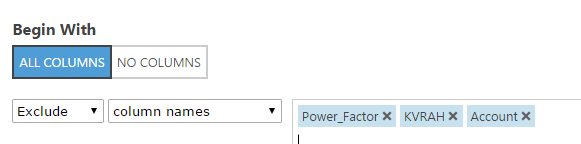
For no-categorical values. We define the value of Account in type of String and stay in its original type.

3.3 Choosing the Variables for Training

We have already define the type of different variables. Now, we should continue by identifying columns that add little-to-no value for predictive modeling. These columns will be dropped.

The first, most obvious candidate to be dropped is Account. All information provided by Account can be replaced by the Buildings. Therefore, the keys could have been completely needlessness and may add false correlations or noise to our model.

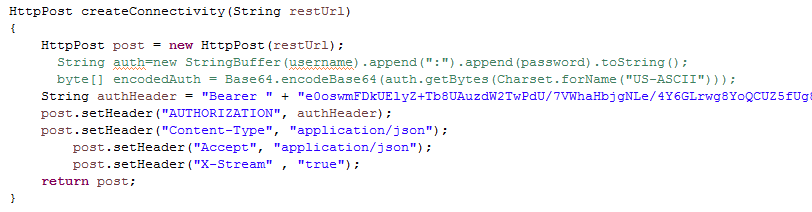
The second candidate for removal is the Power\_Factor column and KVRAH. Normally, there can be generated together with Kwh, their strong relationship may have bad effect for our prediction.



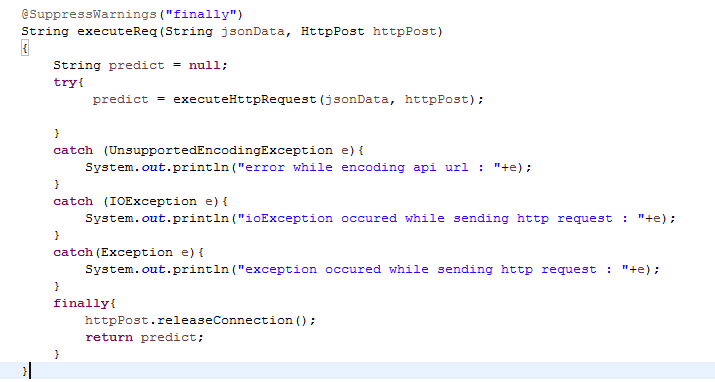
3.4 Split the Data and Training the model

# **4. Web Building**

2.1 Set the header for building Connections

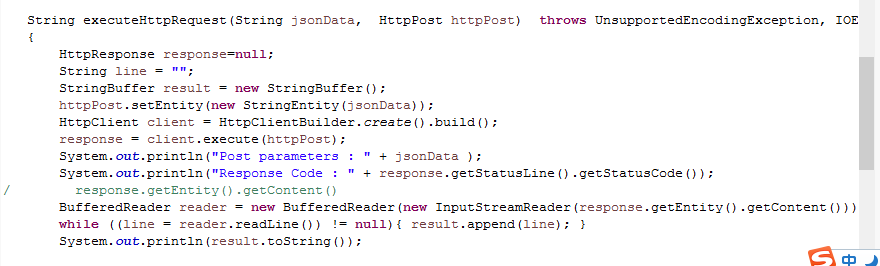


Firstly, we set the post’s header by setting parameters and connect the author Header to our API. The format of our data are setting in JSON.



We get our final result in this format.

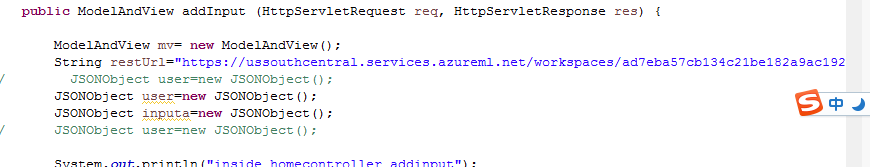
老李你自己编吧，我编不下去了

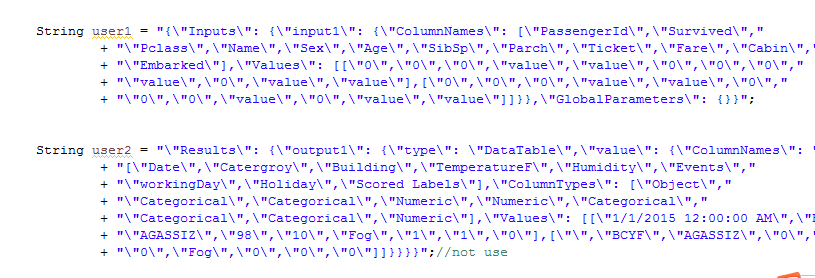


Executing process.

4.2 Executing process.

In the Home Controller, we firstly direct get the data from front page and transform it into JSON format.





Since we have already define the Httppost method, what we need to do next is to send the data into Azure through our API key. Then print the value of its output.

