

ENERGY CONSUMPTION OF BOSTON CITY



**BY: SHREYA JAIN
MUSHTAQ RIZVI
SHRUTI MANDKE**

Problem Statement



- Forecast energy consumption based on historical and consumption data
- Compare the energy consumption of different school, library, hospital
- Validate the Boston model, compare the results of simulation to actual energy consumption data
- Predict future power consumption
- Predict the demand forecast of each region

End Solution



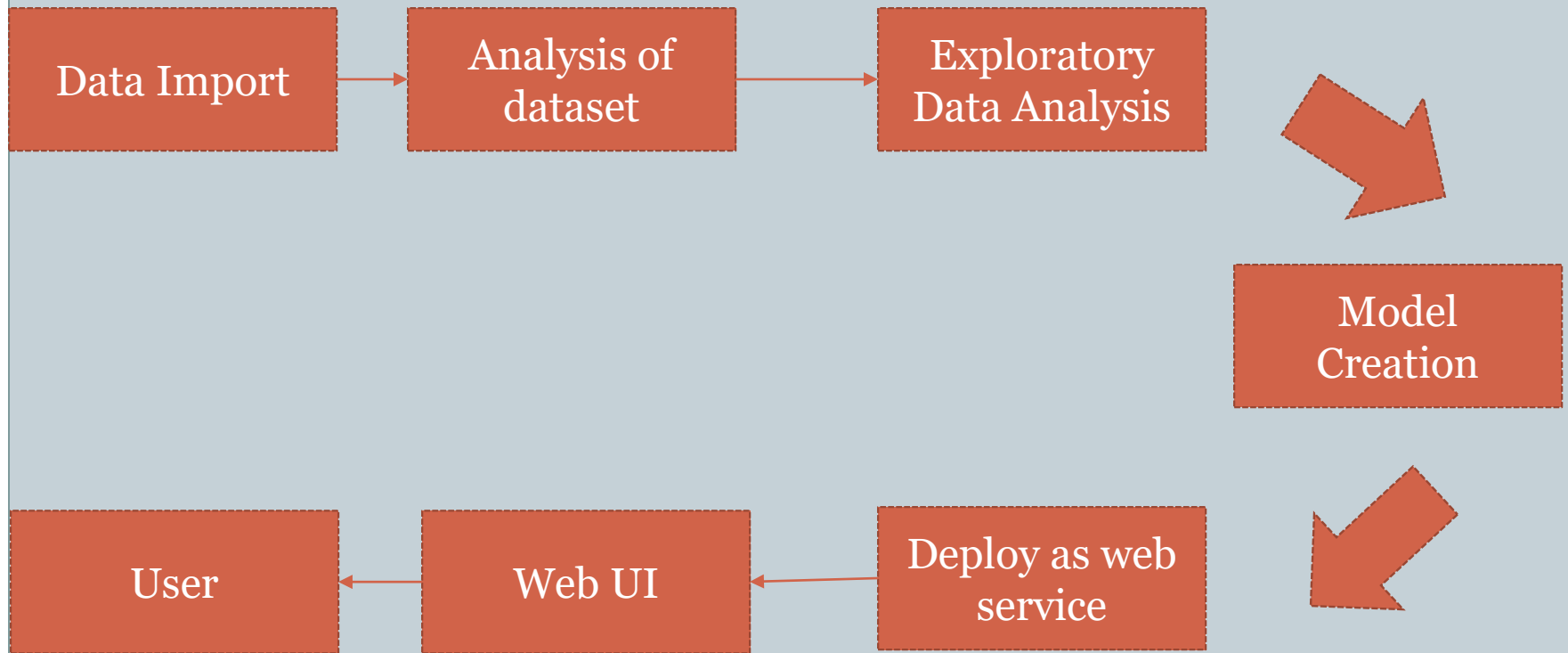
- We'll do this using the related dataset so the models become more and more accurate
- With this model, the city will have a map to help target and reach proper goals
- Lower costs, reduce greenhouse gas emissions, and make Boston's energy system more resilient.

Things we plan to do:



- Exploratory Analysis – R/Tableau
- Data Cleansing – R/Wrangler
- Partitioning of data - Azure
- Modeling and predicting – Azure/R
- Evaluate the model – Azure/R
- Deploy on the web – Azure
- Setup Dashboards path
- UI Design– Bootstrap/JavaScript

Flow Diagram



Dataset



A	B	C	D	E	F	G	H	I	J	K	L
	Account	Date	month	day	year	Time	KWh	powerfact	KVARh	hour	minute
11	2.85E+10	1/1/2014	1	1	2014	X0.05	5.454	0.963654	1.512	0	5
1.1	2.85E+10	1/1/2014	1	1	2014	X0.10	5.166	0.956996	1.566	0	10
1.2	2.85E+10	1/1/2014	1	1	2014	X0.15	5.094	0.956797	1.548	0	15
1.3	2.85E+10	1/1/2014	1	1	2014	X0.20	5.364	0.960791	1.548	0	20
1.4	2.85E+10	1/1/2014	1	1	2014	X0.25	5.148	0.957642	1.548	0	25
1.5	2.85E+10	1/1/2014	1	1	2014	X0.30	5.076	0.957452	1.53	0	30
1.6	2.85E+10	1/1/2014	1	1	2014	X0.35	5.112	0.957081	1.548	0	35
1.7	2.85E+10	1/1/2014	1	1	2014	X0.40	5.274	0.959522	1.548	0	40
1.8	2.85E+10	1/1/2014	1	1	2014	X0.45	5.346	0.974058	1.242	0	45
1.9	2.85E+10	1/1/2014	1	1	2014	X0.50	5.724	0.96141	1.638	0	50
1.1	2.85E+10	1/1/2014	1	1	2014	X0.55	5.814	0.982071	1.116	0	55
1.11	2.85E+10	1/1/2014	1	1	2014	X1.00	5.796	0.976543	1.278	1	0
1.12	2.85E+10	1/1/2014	1	1	2014	X1.05	5.634	0.981534	1.098	1	5
1.13	2.85E+10	1/1/2014	1	1	2014	X1.10	5.13	0.957363	1.548	1	10
1.14	2.85E+10	1/1/2014	1	1	2014	X1.15	5.328	0.96029	1.548	1	15
1.15	2.85E+10	1/1/2014	1	1	2014	X1.20	5.058	0.956219	1.548	1	20
1.16	2.85E+10	1/1/2014	1	1	2014	X1.25	4.95	0.954418	1.548	1	25
1.17	2.85E+10	1/1/2014	1	1	2014	X1.30	5.112	0.956142	1.566	1	30
1.18	2.85E+10	1/1/2014	1	1	2014	X1.35	5.148	0.957642	1.548	1	35
1.19	2.85E+10	1/1/2014	1	1	2014	X1.40	4.968	0.954726	1.548	1	40
1.2	2.85E+10	1/1/2014	1	1	2014	X1.45	5.094	0.953939	1.602	1	45
1.21	2.85E+10	1/1/2014	1	1	2014	X1.50	5.67	0.980581	1.134	1	50

R Code for the Analysis of Data



```
###Dudley BPD

setwd("C:/Users/Mushtaq/Downloads/ADS/COB Interval data 2014/CY2014 COB Interval data 1")
BPDDudley<-read.csv("COB-BPD.DUDLEY SQ.2014.csv",header=T)
summary(is.na(BPDDudley))
BPDDudleyTime<-BPDDudley[,c(5:292)]
BPDDudleyTranspose<-t(BPDDudleyTime)
head(BPDDudleyTranspose)
a<-0

for(i in seq(from=1,to=ncol(BPDDudleyTranspose),by=3)){
  b<-i
  c<-i+2
  a<-rbind(a,BPDDudleyTranspose[,b:c])
}
a=a[-1,]
write.csv(a,file="mushtaq31.csv")
xyz<-read.csv("mushtaq31.csv",header=FALSE)
head(xyz)
xyz=xyz[-1,]
colnames(xyz)<-c("Time", "KWh", "powerfactor", "KVARh")
head(xyz)

library(stringr)
out<-str_split_fixed(xyz$Time,".",2)
nrow(xyz)
nrow(out)
head(xyz)
xyz<-cbind(xyz,out)
xyz<-xyz[,-5]
head(xyz)
colnames(xyz)<-c("Time", "KWh", "powerfactor", "KVARh", "time")
```



```
#Now back to main file
BPDDudleyAccount<-BPDDudley[,c(1:2)]
head(BPDDudleyAccount)

#ExpandRows
newDat<-BPDDudleyAccount[c(10),]

a<-0
for(i in seq(from=1,to=nrow(BPDDudleyAccount),by=3))
{
  b<-i
  a<-rbind(a,expandRows(BPDDudleyAccount[c(b),],288,count.is.col = FALSE))
}
head(a)
a=a[-1,]

#Splitting date
datetxt<-a[,c(2)]
datetxt<-as.character(datetxt)
datetxt<-as.Date(datetxt,"%m/%d/%Y")

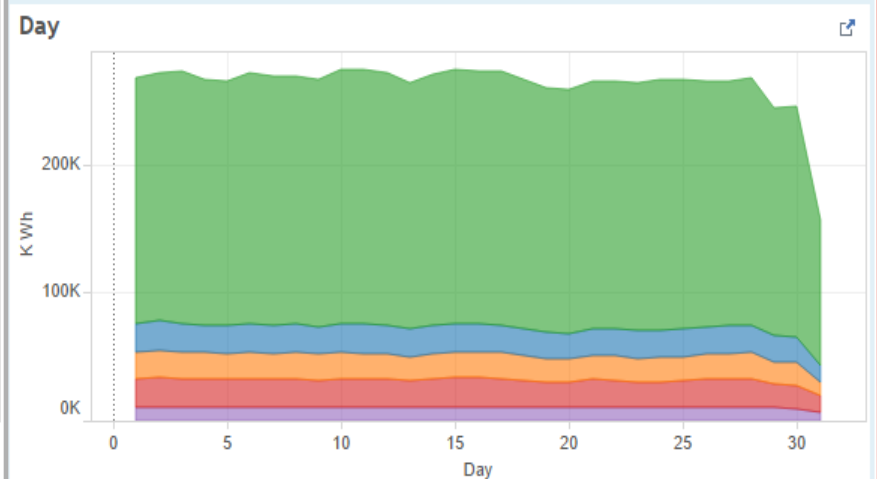
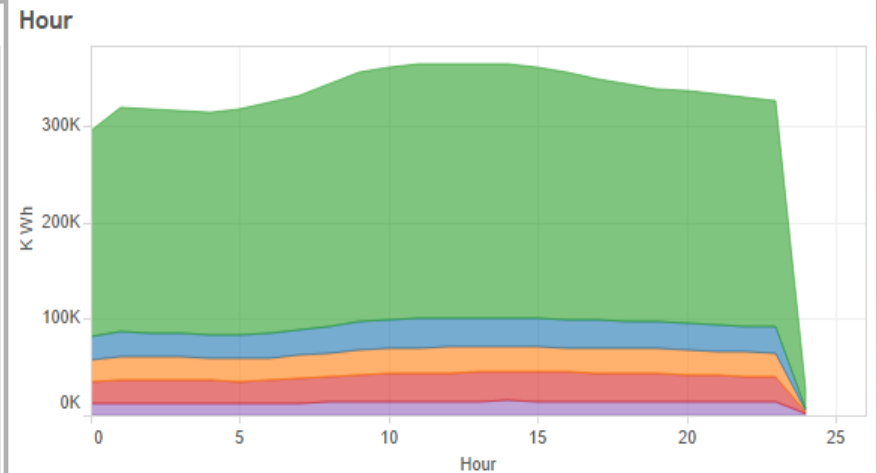
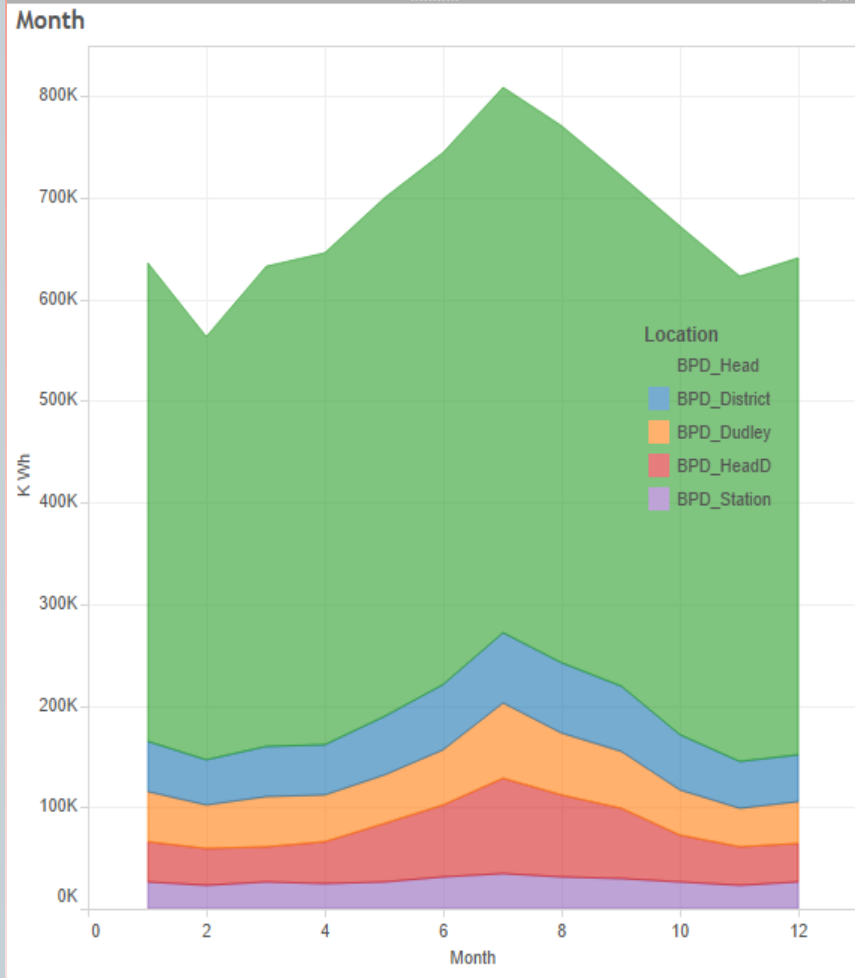
class(datetxt)
df <- data.frame(month = as.numeric(format(datetxt, format = "%m")),
                 day = as.numeric(format(datetxt, format = "%d")),
                 year = as.numeric(format(datetxt, format = "%Y")))

head(df)

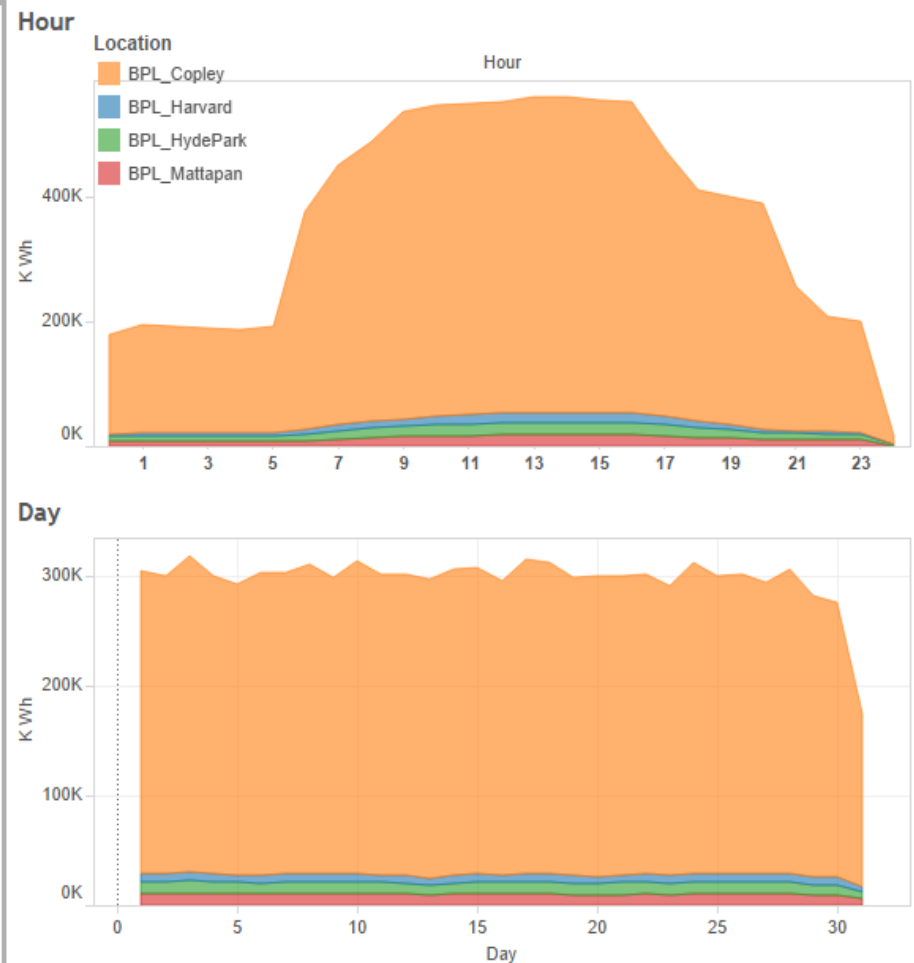
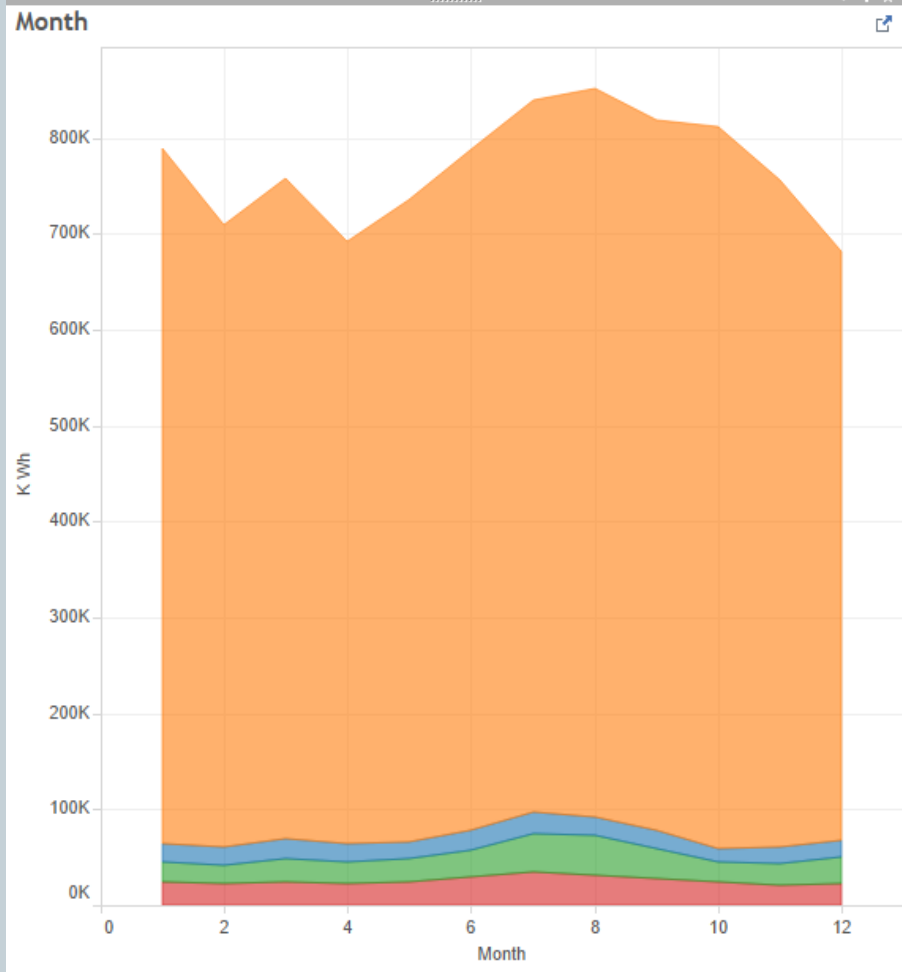
#Column binding
nrow(a)
nrow(xyz)
finalBPDDudley<-cbind(a,df)
finalBPDDudley<-cbind(finalBPDDudley,xyz)
head(final)
write.csv(finalBPDDudley,file="finalBPDDudley.csv")
```

##

Exploratory Data Analysis – Power Consumption for Police Department



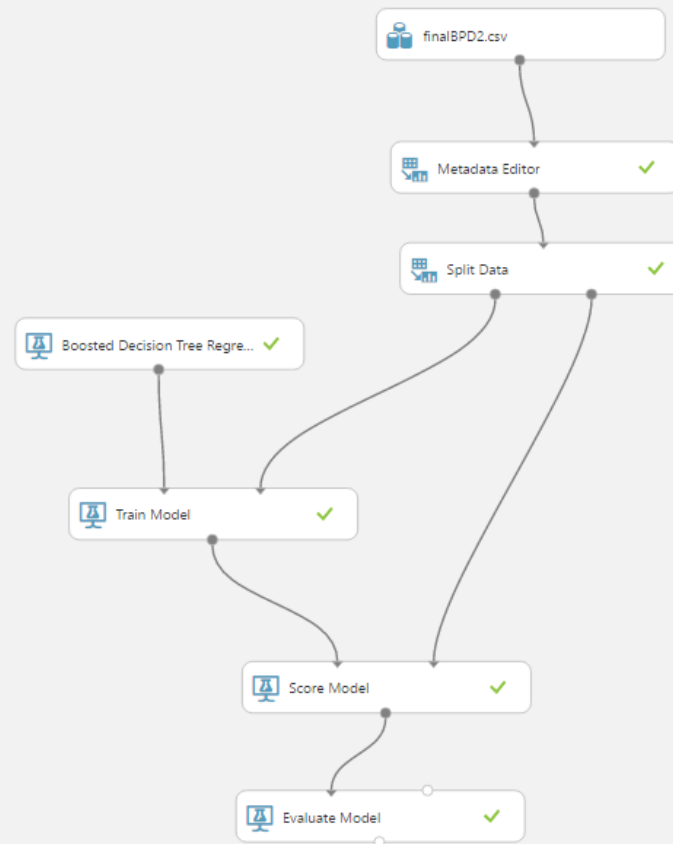
Exploratory Data Analysis – Power Consumption for Libraries



Boosted Decision Tree Regression for Boston Police Department at several places

BPDfinal

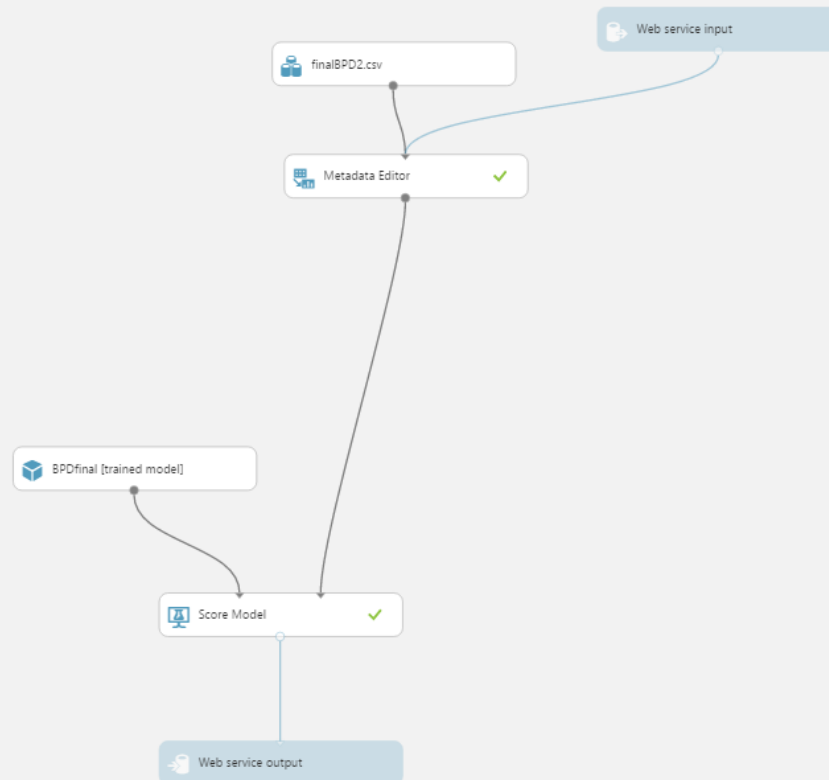
Finished running ✓



Deployed Model



BPDfinal [Predictive Exp.]



Predictive Experiment of Boosted Decision Tree Regression



The screenshot displays the Microsoft Azure Machine Learning Studio interface. The browser address bar shows the URL: <https://studio.azureml.net/Home/ViewWorkspaceCached/6531413850d9471e8ea0f903219d82e2#Workspaces/WebServiceGroups/WebServiceGroup/2ca02684abb14ea89c60b9eb30e4eb02/dashboard>. The page title is "bpdfinal [predictive exp.]".

The interface includes a sidebar with icons for Dashboard, Configuration, General, Published experiment, View snapshot, View latest, Description, No description provided for this web service, API key, and Default Endpoint.

The main content area shows the "API HELP PAGE" with a table of endpoints:

API HELP PAGE	TEST	APPS	LAST UPDATED
REQUEST/RESPONSE	Test	Excel 2013 or later Excel 2010 or earlier workbook	4/17/2016 8:43:02 PM
BATCH EXECUTION		Excel 2013 or later workbook	4/17/2016 8:43:02 PM

Below the table, a message indicates the test result: "BPDFinal [Predictive Exp.] test returned [0,0,26429921005,1/1/2015 12:00:00 AM,1,1,2014,5,null,0.8,12,1,5,BPD_HEAD,697.479431152344]...". A green checkmark indicates a successful result. The result details are as follows:

Result: ("Results":{"output1":{"type":"table","value":{"ColumnNames":["Column 0","X","Account","Date","month","day","year","Time","KWh","powerfactor","KVARh","hour","minute","location","Scored Labels"],"ColumnTypes":["Int32","Double","Int64","DateTime","Int32","Int32","Int32","String","String","String","String","Int32","Int32","String","Double"],"Values":[["0","0","26429921005","1/1/2015 12:00:00 AM","1","1","2014","5","null","0.8","12","1","5","BPD_HEAD","697.479431152344"]]]})

The bottom of the interface shows a taskbar with various application icons and the system clock displaying 9:20 PM.

Evaluation Results of Boosted Decision Tree

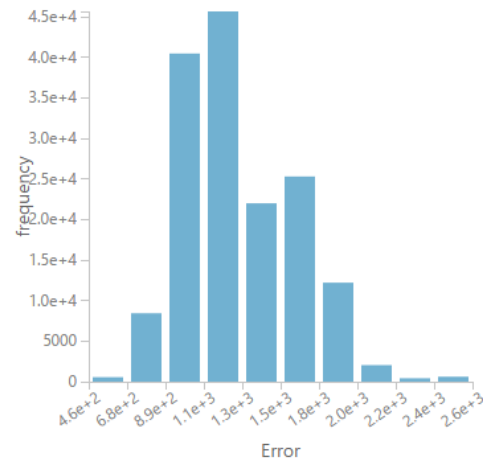


BPDFinal > Evaluate Model > Evaluation results

Metrics

Mean Absolute Error	1306.434861
Root Mean Squared Error	1344.536808
Relative Absolute Error	9.642065
Relative Squared Error	69.345879
Coefficient of Determination	-68.345879

Error Histogram



Link for Tableau Public



- https://public.tableau.com/views/BPD-Boston-Day/Day?:embed=y&:display_count=yes&:showTabs=y
- https://public.tableau.com/views/BPL-Boston-dashboard/Dashboard1?:embed=y&:display_count=yes&:showTabs=y