Zeppelin

Capstone/Visualizat...

default

FINISHED

```
%pyspark
import pandas as pd
from pyspark.sql.types import StringType
from pyspark import SQLContext
```

Took 40 sec. Last updated by anonymous at April 14 2017, 8:45:23 PM.

FINISHED

```
%pyspark
### Import Libraries
import pandas as pd
from pyspark.sql.types import StringType
from pyspark import SQLContext
import numpy as np  # arrays and numerical processing
from sklearn.cluster import KMeans  # cluster analysis by partitioning
from sklearn.metrics import silhouette_score as silhouette_score
from __future__ import division, print_function
import plotly.plotly as py
import plotly.graph_objs as go
from datetime import datetime
from pyspark.sql import functions as F
from pyspark.sql.functions import col,udf, unix_timestamp
from pyspark.sql.types import DateType
from pyspark.sql.functions import from_unixtime
from pyspark.sql.functions import col, unix_timestamp, round
```

Took 44 sec. Last updated by anonymous at April 14 2017, 8:45:32 PM.

FINISHED

```
%pyspark
from pyspark.sql.types import *
custschema = StructType([
StructField("status", StringType(), True),
StructField("avgMeasuredTime", StringType(), True),
StructField("avgSpeed", StringType(), True),
StructField("extID", StringType(), True),
StructField("medianMeasuredTime", StringType(), True),
StructField("TIMESTAMP", StringType(), True),
StructField("vehicleCount", StringType(), True),
StructField("_id", StringType(), True),
StructField("REPORT_ID", StringType(), True)
])
```

Took 0 sec. Last updated by anonymous at April 14 2017, 8:45:41 PM.

FINISHED

```
%pyspark
import timeit
start = timeit.timeit()
road_df1 = sqlContext.read.format("com.databricks.spark.csv").option("header", "true").load("
```

Capstone/visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Capstone/visualizat...

default

```
/capstone/capstone2/*.csv", schema=custschema)
road_df3 = sqlContext.read.format("com.databricks.spark.csv").option("header", "true").load("
/capstone/capstone3/*.csv")
end = timeit.timeit()
print(end - start)
```

```
-0.002839237975422293
```

Took 24 sec. Last updated by anonymous at April 14 2017, 8:46:09 PM.

FINISHED

```
%pyspark
meta_df = sqlContext.read.format("com.databricks.spark.csv").option("header", "true").load("/
/meta/trafficMetaData.csv")
```

Took 1 sec. Last updated by anonymous at April 14 2017, 8:46:14 PM.

FINISHED

```
%pyspark
import timeit
start = timeit.timeit()
road_df1 = road_df1.unionAll(road_df2)
road_df1.count()
end = timeit.timeit()
print(end - start)
```

```
-0.008102782972855493
```

Took 13 sec. Last updated by anonymous at April 14 2017, 8:46:30 PM.

FINISHED

```
%pyspark
import timeit
start = timeit.timeit()
road_df1 = road_df1.unionAll(road_df3)
road_df1.count()
end = timeit.timeit()
print(end - start)
```

```
-0.006811239989474416
```

Took 16 sec. Last updated by anonymous at April 14 2017, 8:46:50 PM.

FINISHED

```
%pyspark

meta_df = meta_df.drop('extID')
meta_df = meta_df.drop('_id')
metadata_columns = list(meta_df.columns)
merged_df = road_df1.join(meta_df, (road_df1.REPORT_ID == meta_df.REPORT_ID) ).drop(meta_df.RI
merged_df.printSchema()
type(merged_df)
```

```
|-- POINT_2_NAME: string (nullable = true)
|-- POINT_2_LNG: string (nullable = true)
|-- POINT_2_STREET: string (nullable = true)
|-- NDT_IN_KMH: string (nullable = true)
|-- POINT_2_POSTAL_CODE: string (nullable = true)
```

Capstone/visualization

Capstone/Visualization | Zeppelin

Capstone/Visualizat...

default

```
 |-- POINT_2_STREET_NUMBER: string (nullable = true)
 |-- ORGANISATION: string (nullable = true)
 |-- POINT_1_LAT: string (nullable = true)
 |-- POINT_2_LAT: string (nullable = true)
 |-- POINT_1_POSTAL_CODE: string (nullable = true)
 |-- POINT_2_STREET_NUMBER: string (nullable = true)
 |-- POINT_2_CITY: string (nullable = true)
 |-- ROAD_TYPE: string (nullable = true)
 |-- POINT_1_LNG: string (nullable = true)
 |-- POINT_1_COUNTRY: string (nullable = true)
 |-- DISTANCE_IN_METERS: string (nullable = true)
 |-- REPORT_NAME: string (nullable = true)
```

Took 0 sec. Last updated by anonymous at April 14 2017, 8:46:55 PM.

READY

FINISHED

```
%pyspark
import timeit
start = timeit.timeit()
road_merged_df = merged_df.select([c for c in merged_df.columns if c not in{'DURATION_IN_SEC'
,'POINT_2_NAME'}])
#road_merged_df.printSchema()
rd_df = road_merged_df.select([c for c in road_merged_df.columns if c not in
{'DURATION_IN_SEC','DISTANCE_IN_METERS','POINT_2_NAME'}])
format = "yyyy-MM-dd'T'HH:mm:ss"
rd_df = rd_df.select('avgMeasuredTime','avgSpeed','TIMESTAMP','vehicleCount','REPORT_ID','POI
,'POINT_2_LAT','POINT_1_LNG',  from_unixtime(unix_timestamp('TIMESTAMP',format)).cast("timest
rd_data_selected = rd_df.select([c for c in rd_df.columns if c not in
{'TIMESTAMP','REPORT_ID','avgMeasuredTime'}])
#rd_data_sel = rd_data_selected.select('avgSpeed','POINT_2_LNG','POINT_1_LAT','POINT_2_LAT','
).alias('hour'),  #.dayofyear('date').alias('dayofy') )
rd_data_sel = rd_data_selected.select('avgSpeed','vehicleCount','POINT_1_LAT','POINT_1_LNG',
F.hour('date').alias('hour'),F.dayofyear('date').alias('dayofy'),F.minute('date').alias('min'
end = timeit.timeit()
print(end - start)
```

```
-0.002913238975452259
```

Took 2 sec. Last updated by anonymous at April 14 2017, 8:47:03 PM.

FINISHED

```
%pyspark
rd_data_sel.show()
```

```
|      79|          19|56.21399075104399|10.145073305557275|56.21740644105506|10.145073305557
7275|  12|    44| 25|2014-02-13 12:25:...|
|      85|          18|56.21399075104399|10.145073305557275|56.21740644105506|10.14507330555
7275|  12|    44| 30|2014-02-13 12:30:...|
|      87|          17|56.21399075104399|10.145073305557275|56.21740644105506|10.14507330555
7275|  12|    44| 35|2014-02-13 12:35:...|
|      88|          23|56.21399075104399|10.145073305557275|56.21740644105506|10.14507330555
7275|  12|    44| 40|2014-02-13 12:40:...|
|      88|          27|56.21399075104399|10.145073305557275|56.21740644105506|10.14507330555
7275|  12|    44| 45|2014-02-13 12:45:...|
|      88|          14|56.21399075104399|10.145073305557275|56.21740644105506|10.14507330555
```

Capstone/Visualizat...

Capstone/Visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Capstone/Visualizat default

```
|      76|          19|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  12|    44| 55|2014-02-13 12:55:...|
|      76|          24|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  13|    44|  0|2014-02-13 13:00:...|
|      79|          21|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  13|    44|  5|2014-02-13 13:05:...|
```

Took 2 sec. Last updated by anonymous at April 14 2017, 8:47:10 PM.

FINISHED

```
%pyspark
rd_data_selected2 = rd_df.select([c for c in rd_df.columns if c not in
{'TIMESTAMP','REPORT_ID','avgMeasuredTime'}])
#rd_data_selected2.show()
from_pattern = 'MMM d, yyyy h:mm:ss aa'
to_pattern = 'yyyy-MM-dd'
df2 = rd_data_sel.withColumn('date', from_unixtime(unix_timestamp(rd_data_selected2['date'],
df2.show()
```

```
7275|  12|    44| 25|2014-02-13|
|      85|          18|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  12|    44| 30|2014-02-13|
|      87|          17|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  12|    44| 35|2014-02-13|

|      88|          23|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  12|    44| 40|2014-02-13|
|      88|          27|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  12|    44| 45|2014-02-13|
|      88|          14|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  12|    44| 50|2014-02-13|
|      76|          19|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  12|    44| 55|2014-02-13|
|      76|          24|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  13|    44|  0|2014-02-13|
|      79|          21|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  13|    44|  5|2014-02-13|
```

Took 0 sec. Last updated by anonymous at April 14 2017, 8:47:15 PM.

FINISHED

```
%pyspark
df2.show()
```

```
|      79|          19|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  12|    44| 25|2014-02-13|
|      85|          18|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  12|    44| 30|2014-02-13|
|      87|          17|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  12|    44| 35|2014-02-13|
|      88|          23|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  12|    44| 40|2014-02-13|
|      88|          27|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
7275|  12|    44| 45|2014-02-13|
|      88|          14|56.21399075104399|10.145073305557275|56.217406441055O6|10.14507330555
```

Capstone/Visualization

Capstone/Visualization

Zeppelin

Capstone/Visualizat

default

Took 0 sec. Last updated by anonymous at April 14 2017, 8:47:23 PM.

FINISHED

```
%pyspark
df3 = df2.where((col("POINT_1_LAT") ==  56.21399075104399) & (col("POINT_1_LNG") == 10.145073
 (col("POINT_2_LAT" )== 56.21740644105506) &(col("POINT_2_LNG") == 10.10702759027481))
df3 =  df3.select([c for c in df3.columns if c not in
{'POINT_1_LAT','POINT_1_LNG','POINT_2_LAT','POINT_2_LNG'}])
df3.show()
```

```
+--------+------------+----+------+---+----------+
|avgSpeed|vehicleCount|hour|dayofy|min|      date|
+--------+------------+----+------+---+----------+
|      93|          15|  11|    44| 30|2014-02-13|
|      90|          20|  11|    44| 35|2014-02-13|
|      84|          20|  11|    44| 40|2014-02-13|
|      88|          25|  11|    44| 45|2014-02-13|
|      90|          23|  11|    44| 50|2014-02-13|
|      88|          22|  11|    44| 55|2014-02-13|
|      87|          21|  12|    44|  0|2014-02-13|
|      84|          19|  12|    44|  5|2014-02-13|
|      88|          19|  12|    44| 10|2014-02-13|
|      90|          10|  12|    44| 15|2014-02-13|
|      80|          16|  12|    44| 20|2014-02-13|
|      79|          19|  12|    44| 25|2014-02-13|
|      85|          18|  12|    44| 30|2014-02-13|
|      87|          17|  12|    44| 35|2014-02-13|
```

Took 1 sec. Last updated by anonymous at April 14 2017, 8:47:28 PM.

FINISHED

```
%pyspark
df3.count()
df3.repartition(1).write.csv("/Users/jyothi/Downloads/latdata")
```

Took 46 sec. Last updated by anonymous at April 14 2017, 8:48:54 PM.

FINISHED

```
%pyspark
#df2.repartition(1).write.csv("/Users/jyothi/Downloads/datedata")

df2.registerTempTable("datedata")
```

Took 0 sec. Last updated by anonymous at April 14 2017, 8:48:59 PM.

ERROR

```
%sql
select  avg(avgSpeed),  date , POINT_2_LNG, POINT_1_LAT, POINT_2_LAT,  POINT_1_LNG  from date
POINT_1_LAT, POINT_2_LAT,  POINT_1_LNG,   date
```

Capstone/Visualization

Capstone/Visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Capstone/Visualizat...

```
org.apache.spark.sql.AnalysisException: cannot resolve '`POINT_2_LNG`' given input columns: [_hour, date, min, dayofy, POINT_1_LAT, avg
d, POINT_1_LNG, vehicleCount, POINT_2_LAT]; line 1 pos 107
set zeppelin.spark.sql.stacktrace = true to see full stacktrace
```

Took 0 sec. Last updated by anonymous at April 14 2017, 8:49:15 PM.

default

FINISHED

```
%pyspark
rd_data_sel.show()
```

```
+--------+------------+-----------------+-----------------+-----------------+-------------
----+----+------+---+--------------------+
|avgSpeed|vehicleCount|      POINT_1_LAT|      POINT_1_LNG|      POINT_2_LAT|       POINT_1
_LNG|hour|dayofy|min|                date|
+--------+------------+-----------------+-----------------+-----------------+-------------
----+----+------+---+--------------------+
|      93|          15|56.21399075104399|10.145073305557275|56.21740644105506|10.14507330555
7275|  11|    44| 30|2014-02-13 11:30:...|
|      90|          20|56.21399075104399|10.145073305557275|56.21740644105506|10.14507330555
7275|  11|    44| 35|2014-02-13 11:35:...|
|      84|          20|56.21399075104399|10.145073305557275|56.21740644105506|10.14507330555
7275|  11|    44| 40|2014-02-13 11:40:...|
|      88|          25|56.21399075104399|10.145073305557275|56.21740644105506|10.14507330555
7275|  11|    44| 45|2014-02-13 11:45:...|
|      90|          23|56.21399075104399|10.145073305557275|56.21740644105506|10.14507330555
7275|  11|    44| 50|2014-02-13 11:50:...|
|      88|          22|56.21399075104399|10.145073305557275|56.21740644105506|10.14507330555
7275|  11|    44| 55|2014-02-13 11:55:...|
```

Took 0 sec. Last updated by anonymous at April 14 2017, 8:49:36 PM.

FINISHED

```
%pyspark
rd_data_sel.registerTempTable("minutedata")
```

Took 0 sec. Last updated by anonymous at April 14 2017, 8:49:43 PM.

FINISHED

```
%sql
select min , avg(avgSpeed)  from minutedata group by min
```

settings

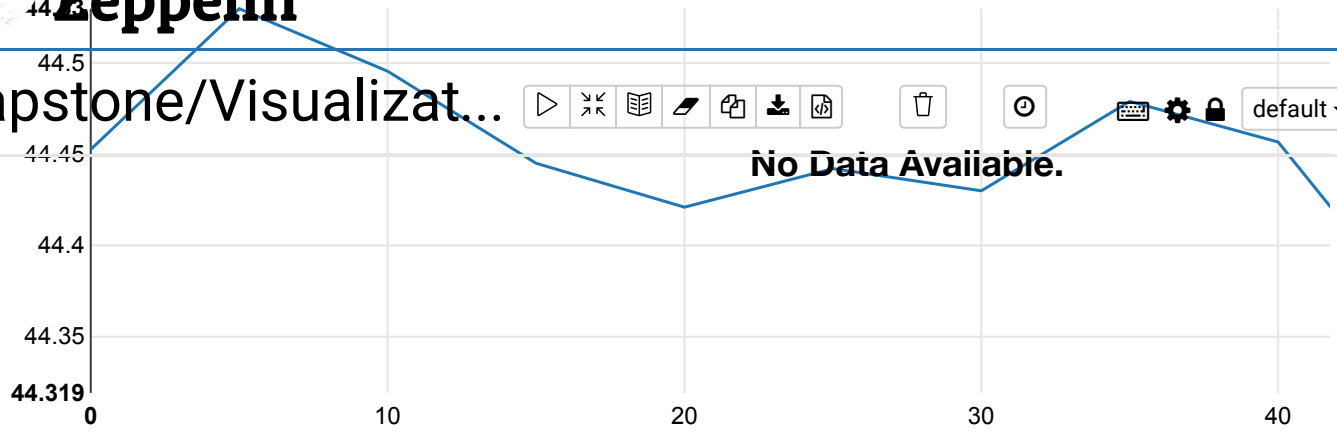
Capstone/Visualization

Capstone/Visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Capstone/Visualizat...

default

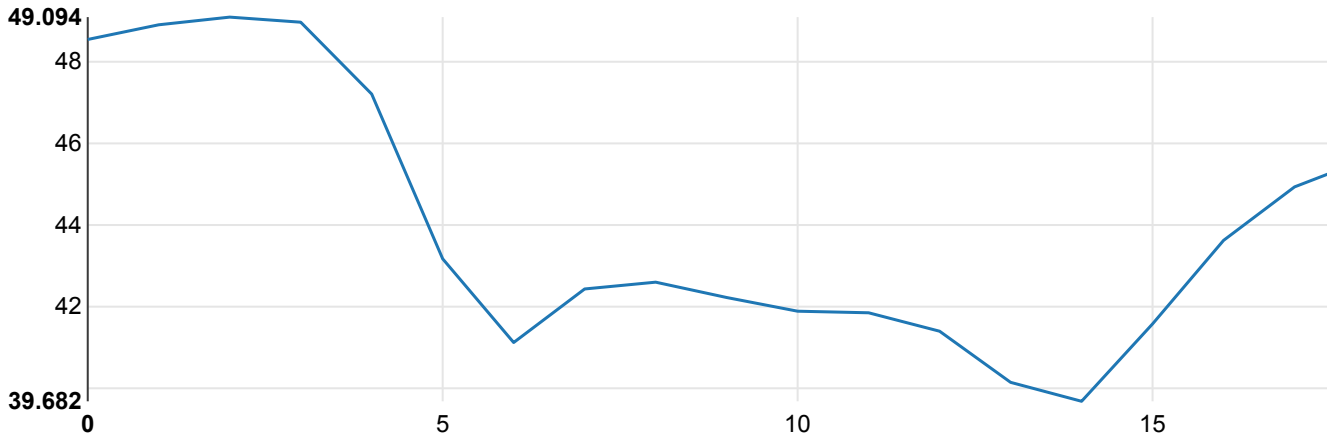


Took 1 min 13 sec. Last updated by anonymous at April 14 2017, 8:51:00 PM.

FINISHED

```
%sql
select hour  as time  ,avg(avgSpeed) as avg_speed   from minutedata group by hour
```

settings



Took 1 min 5 sec. Last updated by anonymous at April 14 2017, 8:52:11 PM.

FINISHED

```
%sql
select dayofy  as time   ,avg(avgSpeed) as avg_speed   from minutedata group by dayofy
```

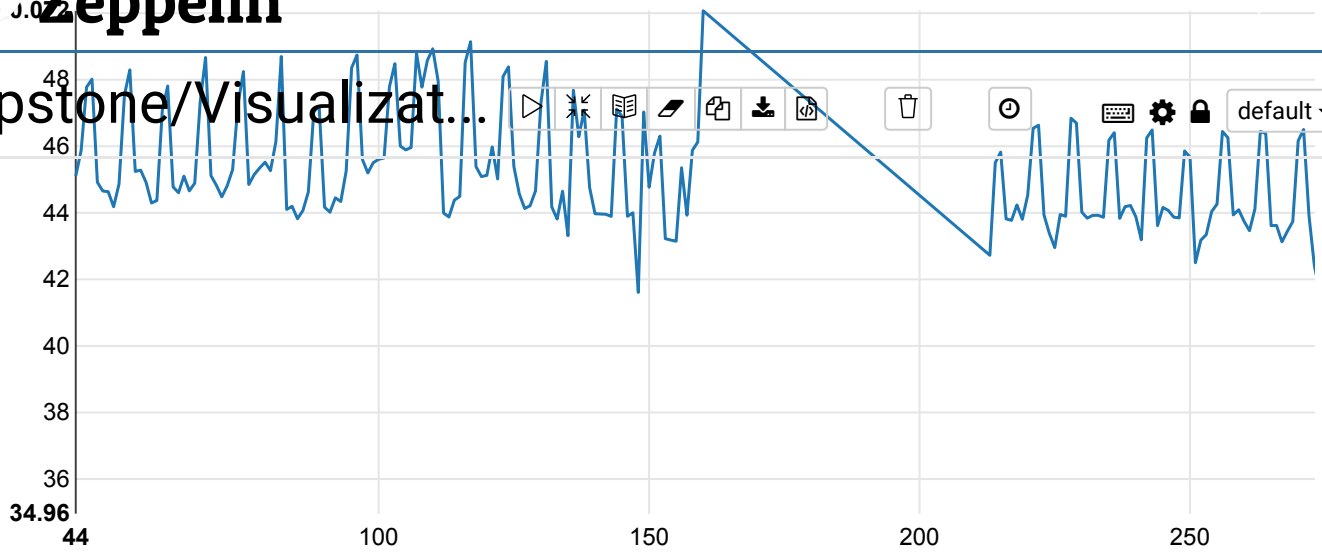
settings

Capstone/Visualization

Capstone/Visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Capstone/Visualizat...



Took 57 sec. Last updated by anonymous at April 14 2017, 3:48:54 AM.

FINISHED

```
%pyspark
import seaborn as sns
import matplotlib
matplotlib.use('Agg')
import matplotlib.pyplot as plt
from io import StringIO
```

Took 0 sec. Last updated by anonymous at April 14 2017, 3:48:59 AM.

FINISHED

```
%pyspark
py_minute_df = sqlContext.sql("select vehicleCount , avg(avgSpeed) as avgSpeed , hour from mi
vehicleCount" )
```

Took 0 sec. Last updated by anonymous at April 14 2017, 3:49:02 AM.

FINISHED

```
%pyspark
#Pandas Data frame after calculating hour , minute data
pd_minute_df = py_minute_df.toPandas()
pd_minute_df.head()
```

	vehicleCount	avgSpeed	hour
0	19	45.586107	6
1	40	60.528859	13
2	0	31.217195	11
3	40	70.050955	10
4	38	79.096386	16

Took 1 min 0 sec. Last updated by anonymous at April 14 2017, 3:50:05 AM.

Capstone/Visualization

Capstone/Visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Capstone/Visualizat... default

```
%pyspark
import matplotlib
matplotlib.use('Agg')
import matplotlib.pyplot as plt
value = "avgSpeed"
x = "vehicleCount"
grouping = ["hour"]


heatmap_data =pd_minute_df.pivot_table(values=value, index=x, columns=grouping)
heatmap_data = heatmap_data[0:100]
a4_dims = (len(heatmap_data.columns),50)
fig, ax = plt.subplots(figsize=a4_dims)
ax.set_title("Avg Speed")
sns.heatmap(heatmap_data, ax=ax, annot=True, fmt=".02f")
```

```
<matplotlib.axes._subplots.AxesSubplot object at 0x12395e7b8>
```

Took 5 sec. Last updated by anonymous at April 14 2017, 3:50:15 AM.

FINISHED

```
%pyspark

def show(p):
  img = StringIO()
  p.savefig(img, format='svg')
  img.seek(0)
  print( "%html " + img.read())
```

Took 0 sec. Last updated by anonymous at April 14 2017, 3:50:20 AM.

FINISHED

```
%pyspark
show(plt)
```

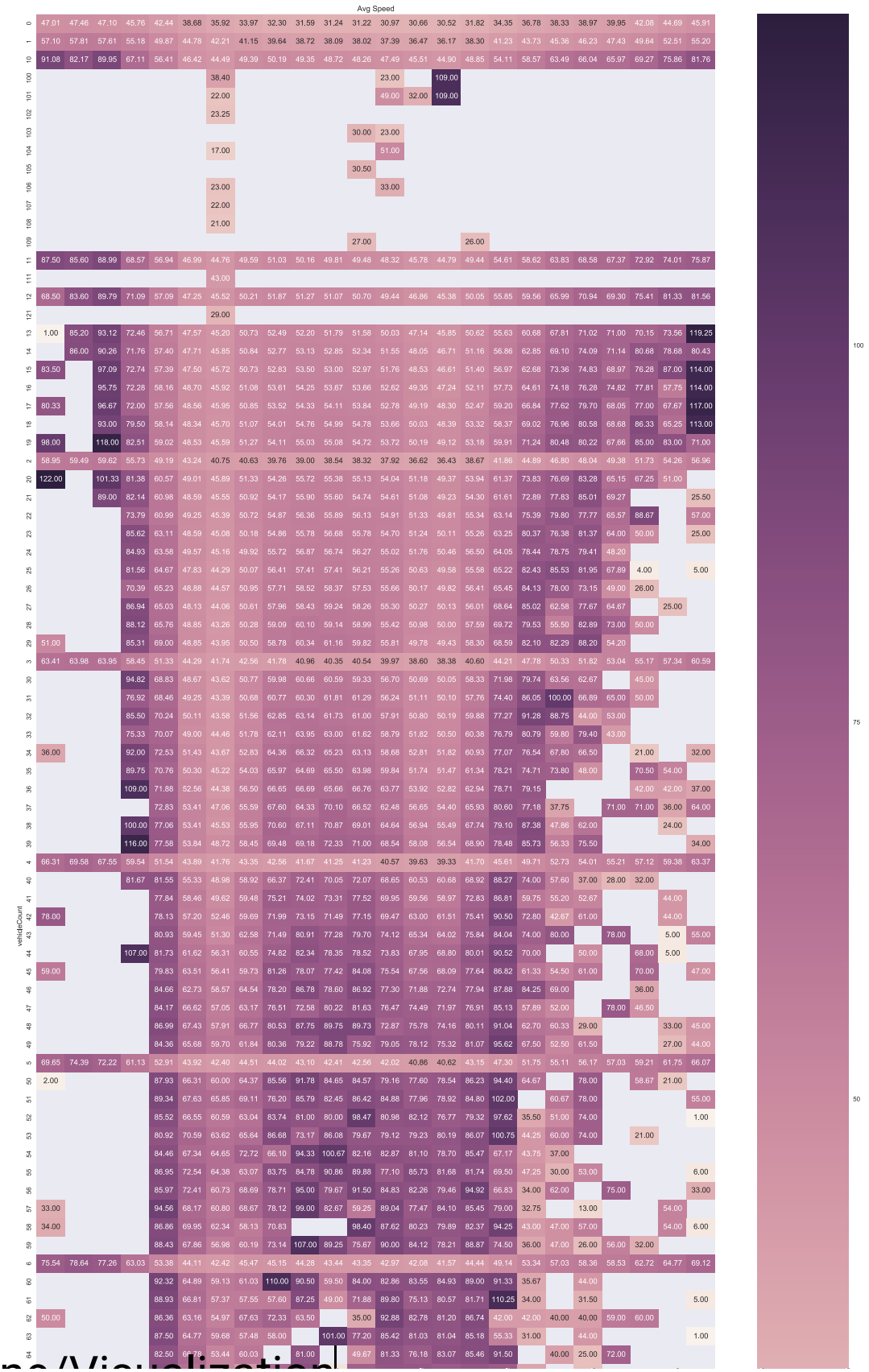
Capstone/Visualization

Capstone/Visualization | Untitled | Untitled | Untitled | Untitled | Untitled | Untitled | Untitled | Untitled | Ur

Zeppelin

Capstone/Visualizat...

default

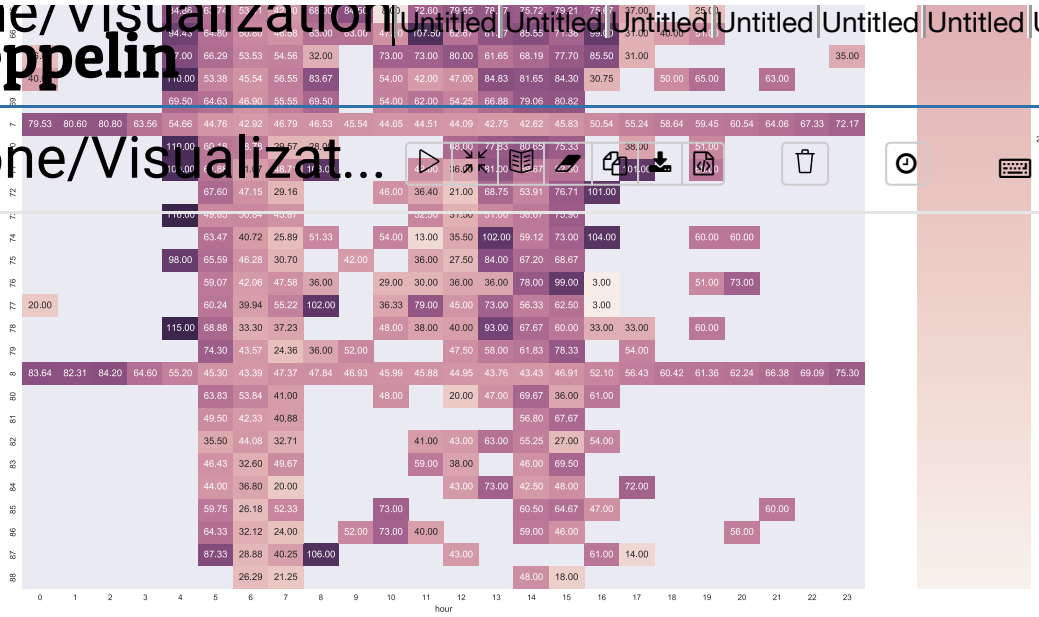


Capstone/Visualization

Capstone/Visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Capstone/Visualizat...



default

Took 7 sec. Last updated by anonymous at April 14 2017, 3:50:35 AM.

FINISHED

```
%pyspark
import matplotlib.pyplot as plt

py_minute_df = sqlContext.sql("select avg(vehicleCount) as vehicleCount  ,avg(avgSpeed) as av
 group by hour" )
pd_df = py_minute_df.toPandas()
ax = pd_df[['vehicleCount','avgSpeed']].plot(kind='bar', title ="V comp", figsize=(15, 10), l
ax.set_xlabel("Hour", fontsize=12)
ax.set_ylabel("V", fontsize=12)
show(plt)
```

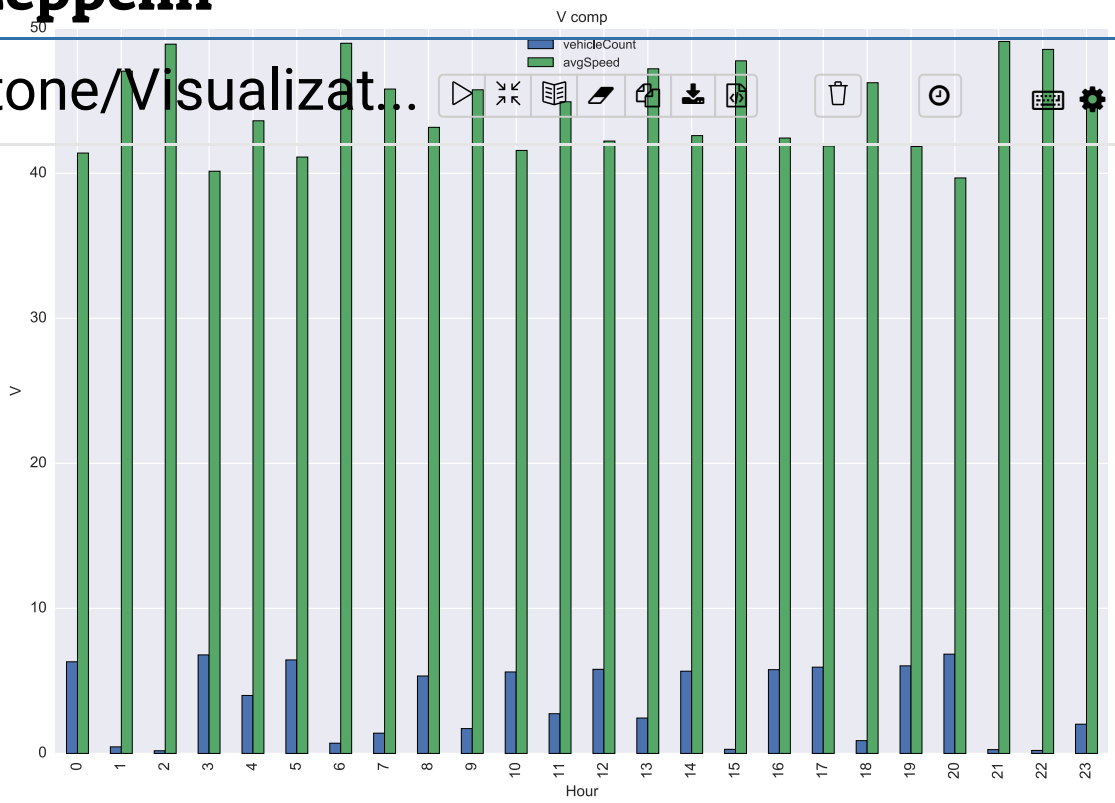
Capstone/Visualization

Capstone/Visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Capstone/Visualizat...

default



Took 1 min 1 sec. Last updated by anonymous at April 14 2017, 3:51:43 AM.

FINISHED

```
%r
library(readr)
library(dplyr)
library(data.table)
library(sqldf)
library(ggplot2)
library(ggmap)
library(maps)
data1 <- read_csv("/Users/jyothi/Downloads/rddata/data1.csv")
print(data1)
```

```
Source: local data frame [214,859 x 5]
```

Took 4 sec. Last updated by anonymous at April 14 2017, 8:54:05 PM.

FINISHED

```
%r
mydata <- data1
wss <- (nrow(mydata)-1)*sum(apply(mydata,2,var))
for (i in 2:15) wss[i] <- sum(kmeans(mydata,centers=i)$withinss)
plot(1:15, wss, type="b", xlab="Number of Clusters",
     ylab="Within groups sum of squares")
```

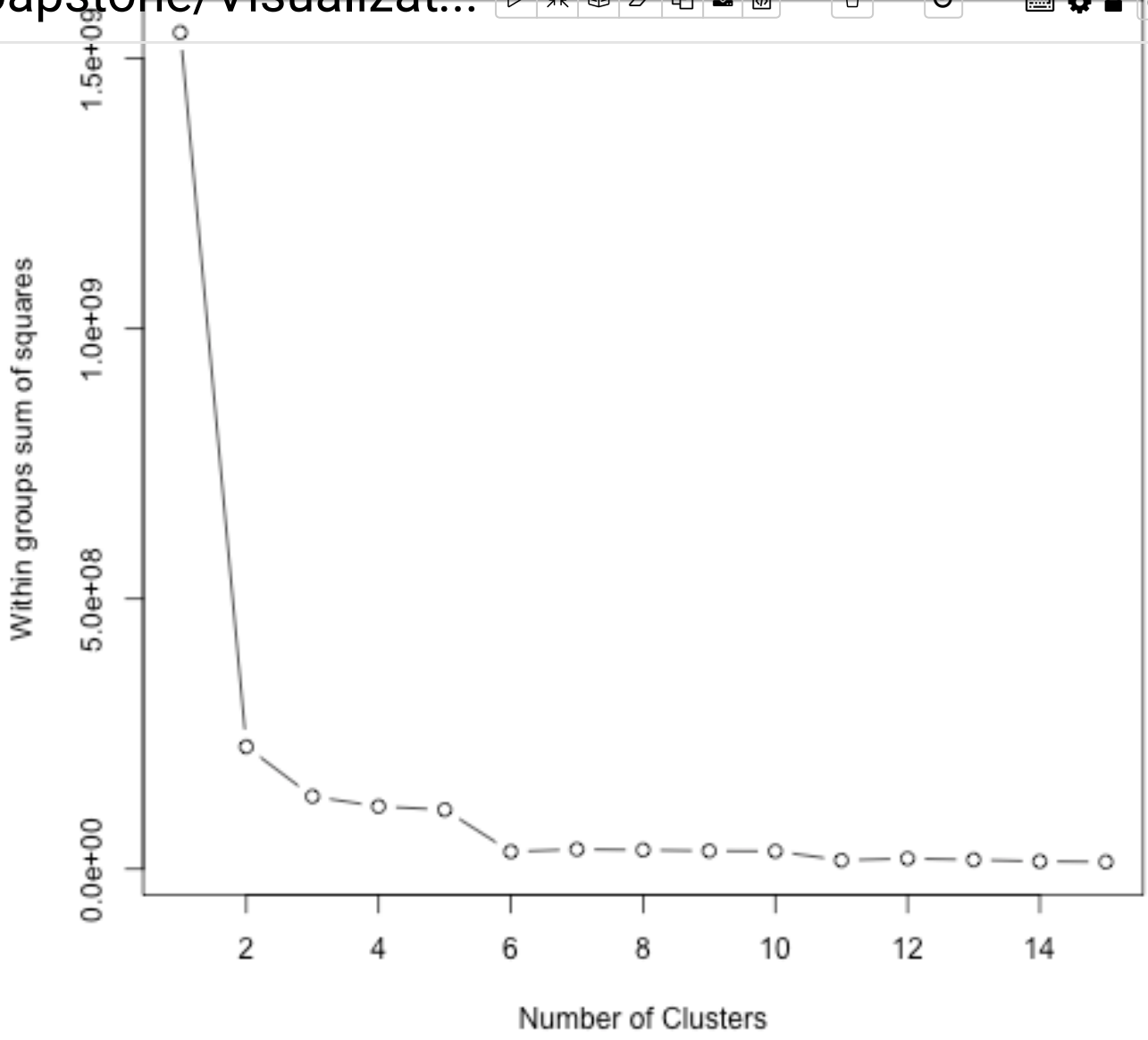
Capstone/Visualization

Capstone/Visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Capstone/Visualizat...

default



Took 5 sec. Last updated by anonymous at April 14 2017, 3:53:20 AM.

FINISHED

```
%r
tbl1<-subset(data1, data1$hr_grp == 1)
#tbl1<-subset(tbl, data1$mean > 50)
result4<-kmeans(tbl1,3)
tbl1$cluster <- result4$cluster
#View(tbl1)
tbl1$cluster <- as.factor(tbl1$cluster)
mapgilbert <- get_map(location = c(lon = mean(tbl1$long), lat = mean(tbl1$lat)), zoom = 14, m
=6)
# plotting the map with some points on it
ggmap(mapgilbert) + geom_point(data = tbl1, aes(x = as.numeric(long), y = as.numeric(lat),col
size = 3, shape = 20) + guides(fill=FALSE, alpha=FALSE, size=FALSE)
```

Capstone/Visualization

Capstone/Visualization | Untitled | Untitled | Untitled | Untitled | Untitled | Untitled | Untitled | Untitled | Ur

Zeppelin

```
ror: scale.b... c(1, 2, 4) is not TRUE
```

Capstone/Visualizat...

default

Took 4 sec. Last updated by anonymous at April 14 2017, 3:57:57 AM.

FINISHED

Took 0 sec. Last updated by anonymous at April 14 2017, 3:58:16 AM.

FINISHED

```
%r
tbl1<-subset(data1, data1$hr_grp == 2)
#tbl1<-subset(tbl, data1$mean > 50)
result4<-kmeans(tbl1,3)
```

Capstone/Visualization

Capstone/Visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Capstone/Visualizat...

default

```
tbl1$cluster <- as.factor(tbl1$cluster)
mapgilbert <- get_map(location = c(lon = mean(tbl1$long), lat = mean(tbl1$lat)), zoom = 13, ma
2)
# plotting the map with some points on it
ggmap(mapgilbert) + geom_point(data = tbl1, aes(x = as.numeric(long), y = as.numeric(lat),colo
size = 3, shape = 20) + guides(fill=FALSE, alpha=FALSE, size=FALSE)
```

Took 7 sec. Last updated by anonymous at April 14 2017, 3:57:10 AM.

READY

```
%pyspark
py_date_df = sqlContext.sql("select vehicleCount , avg(avgSpeed) as avgSpeed , hour from minu
vehicleCount" )
```

Capstone/Visualization

Capstone/ Visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Capstone/Visualizat...

FINISHED

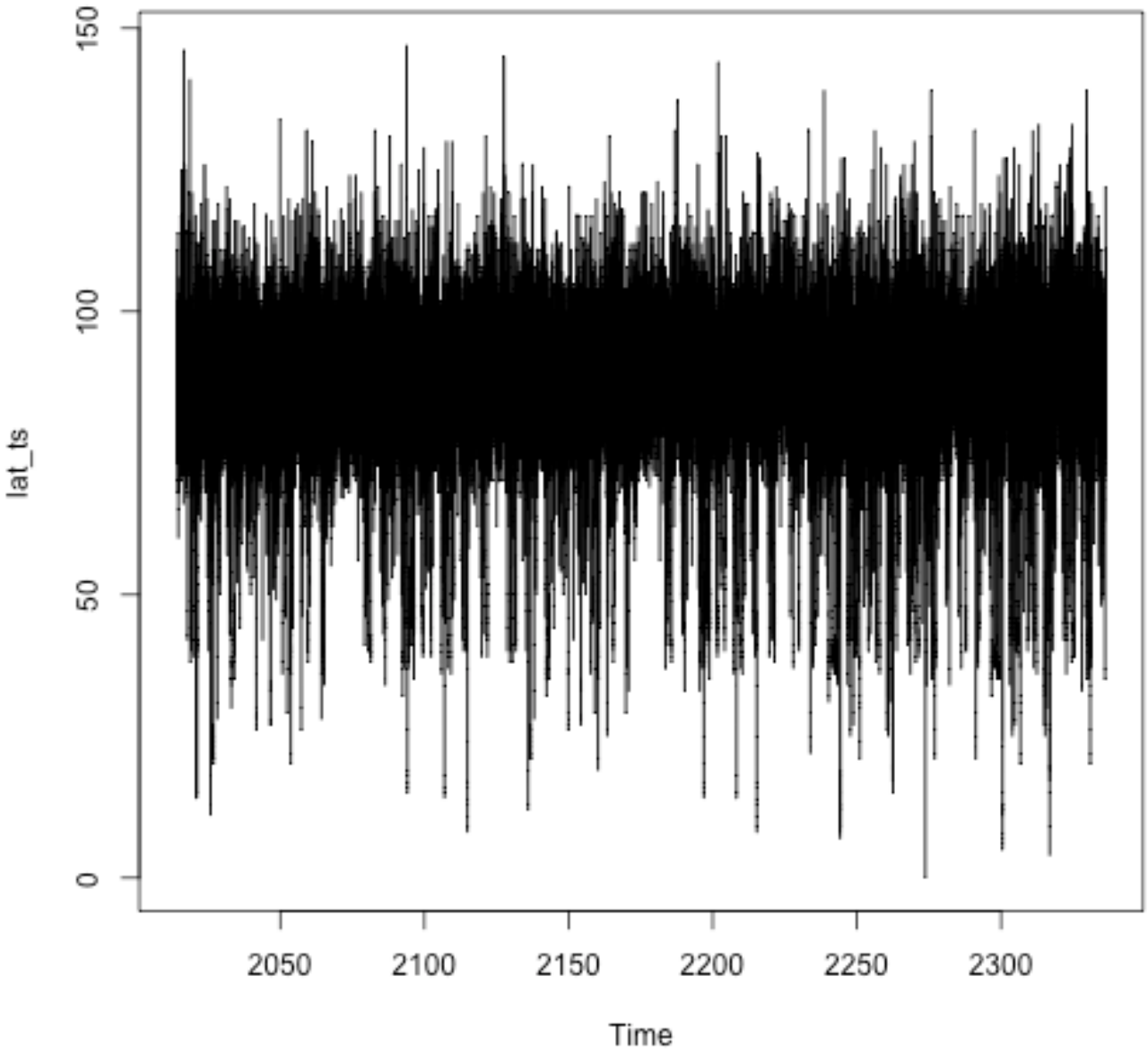
```
%r
data_ts = read.csv("C:/Users/jyothi/Downloads/latdata/latdata.csv")
```

default

Took 1 sec. Last updated by anonymous at April 14 2017, 8:56:06 PM.

FINISHED

```
%r
lat_ts <- ts(data_ts$avgspeed,start=c(2014,2),frequency=24*60/05)
plot(lat_ts)
acf(lat_ts)
pacf(lat_ts)
```



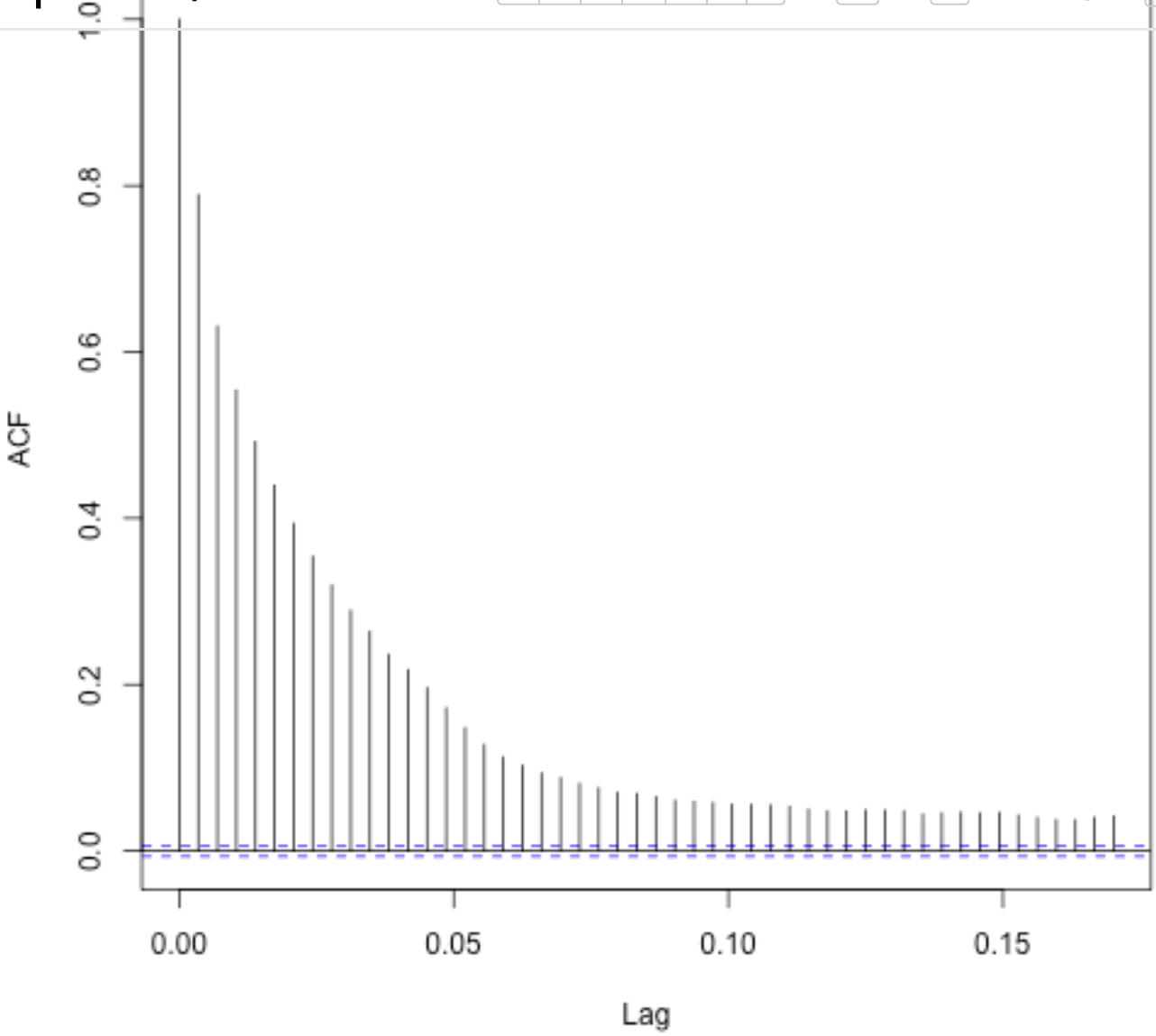
Capstone/Visualization

Capstone/Visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Series lat_ts

Capstone/Visualizat... default



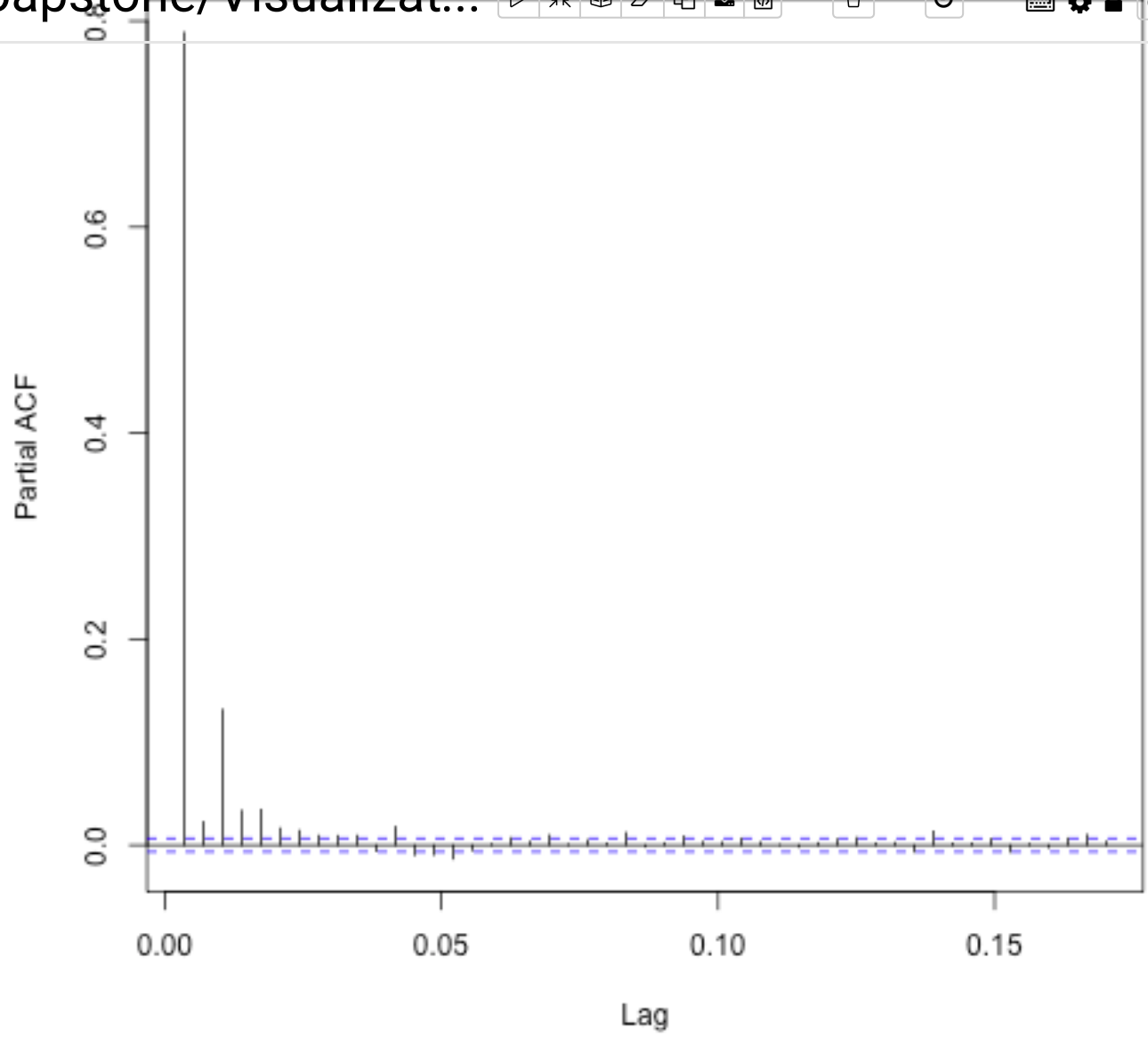
Capstone/Visualization

Capstone/Visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Series lat_ts

Capstone/Visualizat... default




```
Error in eval(expr, envir, enclos): could not find function "sarima"
```

```
Error in eval(expr, envir, enclos): could not find function "sarima"
```

```
Error in eval(expr, envir, enclos): could not find function "sarima"
```

Took 2 sec. Last updated by anonymous at April 14 2017, 8:56:12 PM. (outdated)

FINISHED

```
%r
library(astsa)
```

Capstone/Visualization

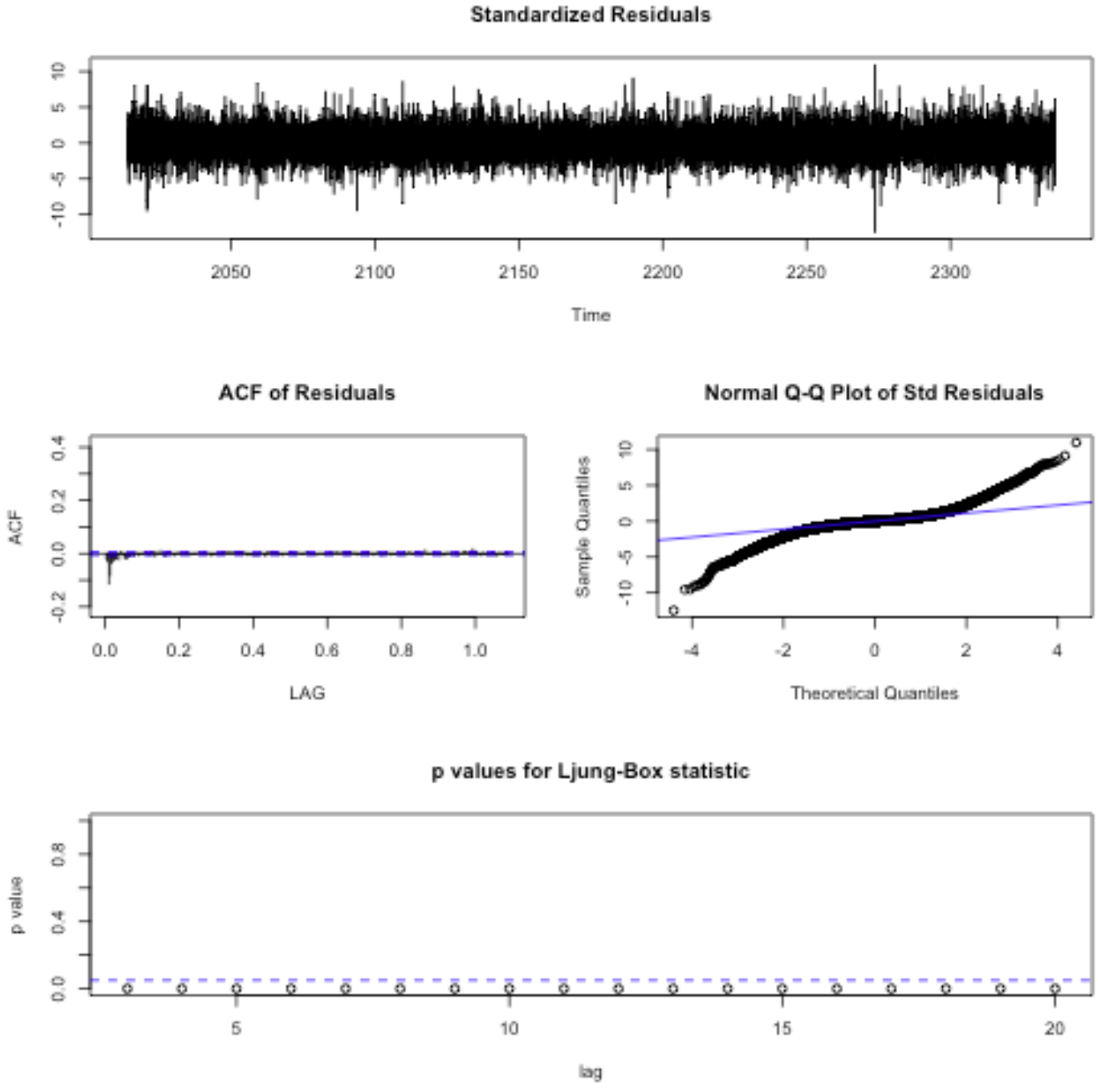
Capstone/Visualization Untitled Untitled Untitled Untitled Untitled Untitled Untitled Untitled Ur

Zeppelin

Capstone/Visualizat... default

```
sarima(latt , 2,0,0)
```

initial value 2.039744 iter 2 value 2.009463 iter 3 value 2.008967 iter 4 value 2.008962 iter 4 value 2.008962 iter4 value 2.008962 final value 2.008962 converged initial value 2.008956 iter 1 value 2.008956 final value 2.008956 converged



$fit

Call: stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S), xreg = constant, optim.control = list(trace = trc, REPORT = 1, reltol = tol))

Coefficients: ar1 ar2 constant -0.1522 -0.2112 -0.0001 s.e. 0.0032 0.0032 0.0179

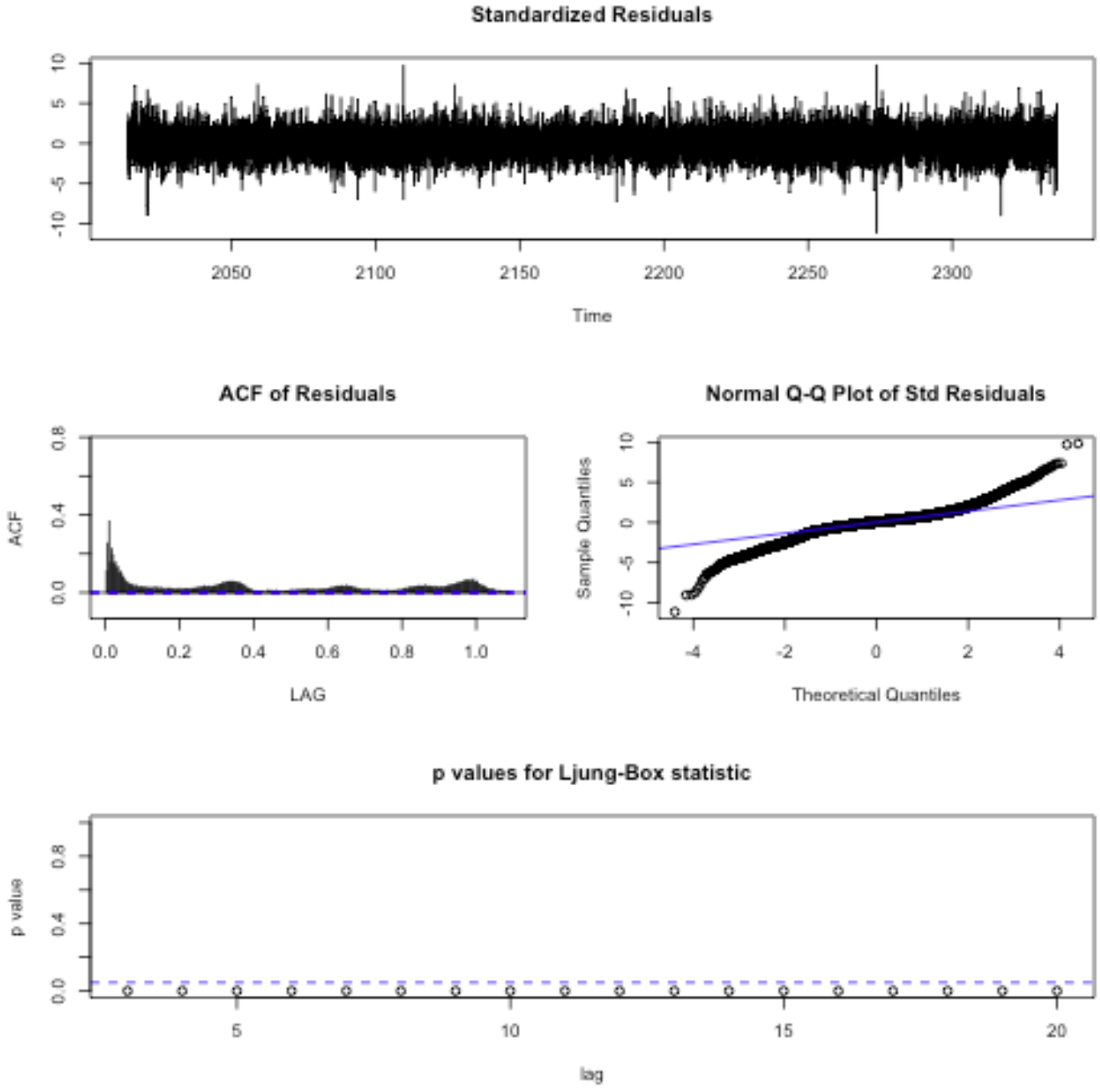
$sigma^2$ estimated as 55.58: log likelihood = -318211.5, aic = 636430.9

$AIC [1] 5.017976

Capstone/Visualization

$aic [1] 5.047997
$BIC [1] 4.078281

initial value 2.470901 iter 2 value 2.156621 iter 3 value 2.131212 iter 4 value 2.096544 iter 5 value 2.082398 iter 6 value 2.080082 iter 7 value 2.079883 iter 8 value 2.079873 iter 9 value 2.079873 iter 9 value 2.079873 final value 2.079873 converged initial value 2.079874 iter 1 value 2.079874 final value 2.079874 converged



$fit

Call: stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S), xreg = xmean, include.mean = FALSE, optim.control = list(trace = trc, REPORT = 1, reltol = tol))

Coefficients: ma1 ma2 xmean 0.7979 0.3588 87.1514 s.e. 0.0030 0.0026 0.0566

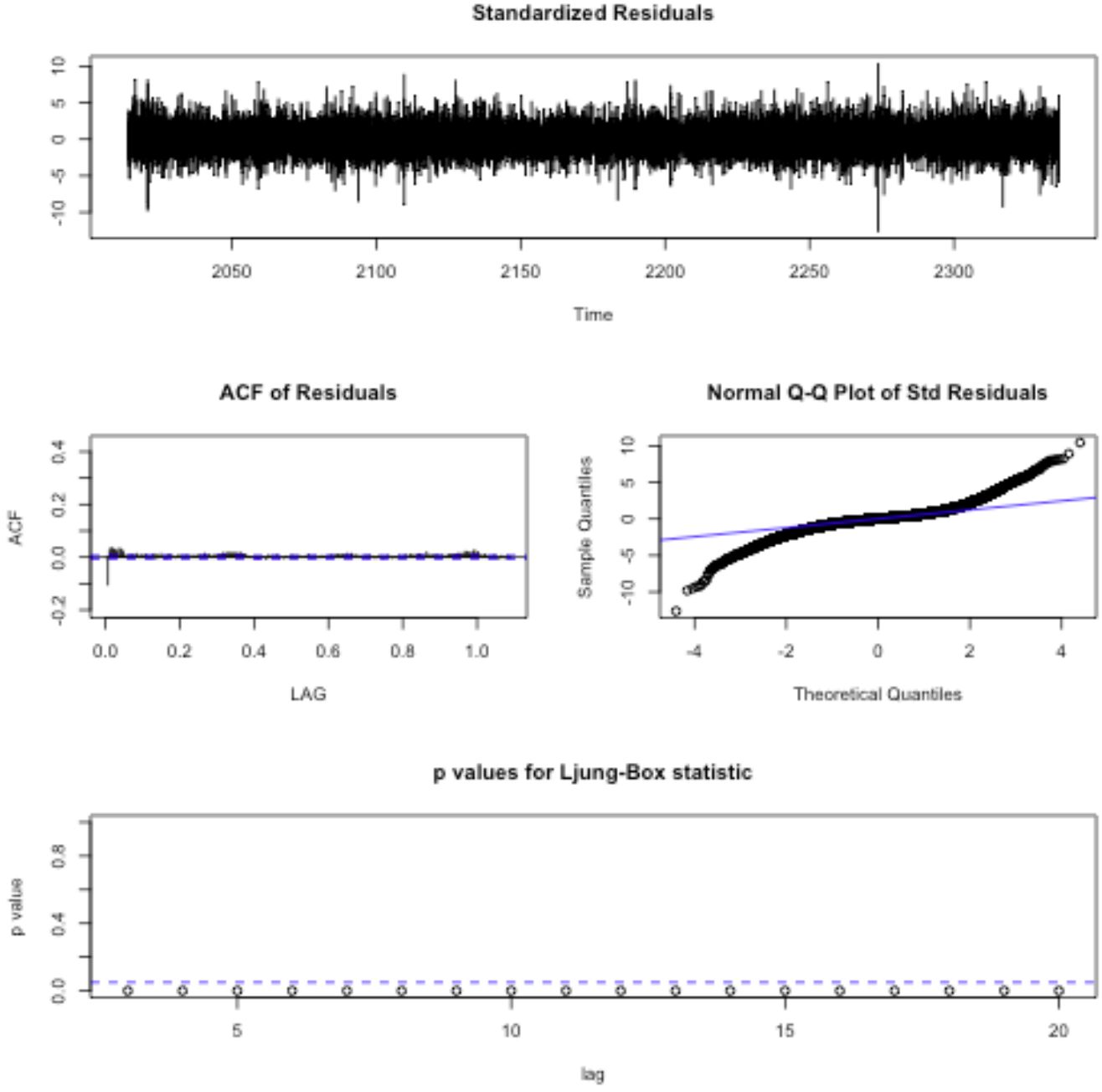
$sigma^2$ estimated as 64.05: log likelihood = -324798.3, aic = 649604.6

$AIC [1] 5.159806

$aic [1] 5.159828

$bic [1] 4.160141

initial value 2.470910 iter 2 value 2.336910 iter 3 value 2.026516 iter 4 value 1.995159 iter 5 value 1.983960 iter 6 value 1.983718 iter 7 value 1.983718 iter 7 value 1.983718 iter 7 value 1.983718 final value 1.983718 converged initial value 1.983714 iter 1 value 1.983714 final value 1.983714 converged



$fit

Call: stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S), xreg = xmean, include.mean = FALSE, optim.control = list(trace = trc, REPORT = 1, reltol = tol))

Coefficients: ar1 ar2 xmean 0.7711 0.0225 87.1514 s.e. 0.0033 0.0033 0.1156

sigma^2 estimated as 52.85: log likelihood = -315871.7, aic = 631751.3

$AIC [1] 4.967482

$AIC [1] 4.967504

$BIC [1] 3.967787

Zeppelin

Capstone/Visualizat...

Took 15 sec. Last updated by anonymous at April 14 2017, 9:05:05 PM.

default

READY