Data

FINISHED ▷ ⌗ 📖 ⚙

```pyspark
%pyspark
import csv
import json
import pandas as pd
import datetime
```

Took 1 sec. Last updated by anonymous at March 27 2017, 8:49:31 PM. (outdated)

FINISHED ▷ ⌗ 📖 ⚙

```pyspark
%pyspark
li=[]
with open("/home/scarface/Desktop/sem-3/capstone/sravan/raw_weather_data_aarhus/wspdm.txt", "r") as 
    for each in f:
        js = json.loads(each)
        li.append(js)
alLi = {}
for a in li:
    for x in a:
        ac=datetime.datetime.strptime(x, "%Y-%m-%dT%H:%M:%S")
        alLi[ac]=a[x]
lis=[]
for k,v in sorted(alLi.items()):
    lis.append([str(k),v])
pd.DataFrame(lis).to_csv("/home/scarface/Desktop/sem-3/capstone/sravan/raw_weather_data_aarhus/wspdm
```

Took 1 sec. Last updated by anonymous at March 27 2017, 8:49:34 PM. (outdated)

FINISHED ▷ ⌗ 📖 ⚙

```pyspark
%pyspark
import csv
import numpy
import matplotlib.pyplot as plt
```

Took 0 sec. Last updated by anonymous at March 27 2017, 8:49:47 PM. (outdated)

FINISHED ▷ ⌗ 📖 ⚙

```pyspark
%pyspark
f1 = open("/home/scarface/Desktop/sem-3/capstone/sravan/raw_weather_data_aarhus/hum.csv", "rt")
f2 = open("/home/scarface/Desktop/sem-3/capstone/sravan/raw_weather_data_aarhus/tempm.csv", "rt")
f3 = open("/home/scarface/Desktop/sem-3/capstone/sravan/raw_weather_data_aarhus/dewptm.csv", "rt")
f4 = open("/home/scarface/Desktop/sem-3/capstone/sravan/raw_weather_data_aarhus/pressurem.csv", "rt
f5 = open("/home/scarface/Desktop/sem-3/capstone/sravan/raw_weather_data_aarhus/vism.csv", "rt")
f6 = open("/home/scarface/Desktop/sem-3/capstone/sravan/raw_weather_data_aarhus/wdird.csv", "rt")
f7 = open("/home/scarface/Desktop/sem-3/capstone/sravan/raw_weather_data_aarhus/wspdm.csv", "rt")
```

Took 0 sec. Last updated by anonymous at March 27 2017, 8:49:50 PM. (outdated)

Zeppelin   Notebook

Assign9_sravan

FINISHED ● anonymous

default ▾

```
hum = csv.reader(f1)
tem = csv.reader(f2)
dew = csv.reader(f3)
pre = csv.reader(f4)
vis = csv.reader(f5)
wdi = csv.reader(f6)
wsp = csv.reader(f7)
```

Took 0 sec. Last updated by anonymous at March 27 2017, 8:49:56 PM. (outdated)

FINISHED

```
%pyspark
humLi=[]
temLi=[]
dewLi=[]
preLi=[]
visLi=[]
wdiLi=[]
wspLi=[]
dateLi=[]
#43
```

Took 0 sec. Last updated by anonymous at March 27 2017, 8:50:05 PM. (outdated)
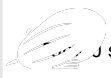
FINISHED

```
%pyspark
for h in hum:
    if h[2]=="":
        humLi.append(0)
    else:
        humLi.append(float(h[2]))
    dateLi.append(h[1])
for t in tem:
    if t[2]=="":
        temLi.append(0)
    else:
        temLi.append(float(t[2]))
```

Took 0 sec. Last updated by anonymous at March 27 2017, 8:50:11 PM. (outdated)

FINISHED

```
%pyspark
for d in dew:
    if d[2]=="":
        dewLi.append(0)
    else:
        dewLi.append(float(d[2]))
for p in pre:
    if p[2]=="":
        preLi.append(0)
    else:
        preLi.append(float(p[2]))
```

Zeppelin Notebook

0 sec. Last updated by anonymous at March 27 2017, 8:50:17 PM. (outdated)

Assign9_sravan

FINISHED

```
%pyspark
for v in vis:
    if v[2]=="":
        visLi.append(0)
    else:
        visLi.append(float(v[2]))
for wd in wdi:
    if wd[2]=="":
        wdiLi.append(0)
    else:
        wdiLi.append(float(wd[2]))
for ws in wsp:
    if ws[2]=="":
        wspLi.append(0)
    else:
        wspLi.append(float(ws[2]))
#79
```

default

Took 0 sec. Last updated by anonymous at March 27 2017, 8:50:22 PM. (outdated)

FINISHED

```
%pyspark
print (len(dewLi))
corHT=numpy.corrcoef(humLi,temLi)
corTD=numpy.corrcoef(temLi,dewLi)
corDP=numpy.corrcoef(dewLi,preLi)
corPV=numpy.corrcoef(preLi,visLi)
corVW=numpy.corrcoef(wdiLi,wspLi)

print (corHT)
print (corTD)
print (corDP)
print (corPV)
print (corVW)
```

```
8222
[[ 1.         -0.50706437]
 [-0.50706437  1.        ]]
[[ 1.          0.71304343]
 [ 0.71304343  1.        ]]
[[ 1.          0.02042398]
 [ 0.02042398  1.        ]]
[[ 1.         -0.00833772]
 [-0.00833772  1.        ]]
[[ 1.          0.03616407]
 [ 0.03616407  1.        ]]
```

Took 0 sec. Last updated by anonymous at March 27 2017, 8:55:24 PM.

FINISHED

```
%pyspark
plt.plot(humLi)
```

Zeppelin Notebook

Took 0 sec. Last updated by anonymous at March 27 2017, 8:50:34 PM. (outdated)

Assign9_sravan

● anonymou

FINISHED

```
%pyspark
plt.plot(wdiLi)
plt.plot(wspLi)
plt.show()
```

default

Took 0 sec. Last updated by anonymous at March 27 2017, 8:50:37 PM. (outdated)

FINISHED

```
%pyspark
plt.plot(preLi)
plt.plot(visLi)
plt.show()
```

Took 0 sec. Last updated by anonymous at March 27 2017, 8:51:32 PM. (outdated)

FINISHED

```
%pyspark
plt.plot(dewLi)
plt.plot(preLi)
plt.show()
```

Took 0 sec. Last updated by anonymous at March 27 2017, 8:52:08 PM. (outdated)

READY