ZONE 3532

```
import pandas as pd
zone_3532_flow_df = pd.read_csv("/home/datascience/Downloads/cleaning1/3532/flow.tsv",
                na_values=['-'], sep='\t', names = ['f_d1','f_d2','f_d3','f_d4','f_d5','f_d6'])
zone_3532_speed_df = pd.read_csv("/home/datascience/Downloads/cleaning1/3532/speed.tsv",
                 na_values=['-'], sep='\t', names = ['s_d1','s_d2','s_d3','s_d4','s_d5','s_d6'])
zone_3532_occupancy_df =
pd.read_csv("/home/datascience/Downloads/cleaning1/3532/occupancy.tsv",
                 na_values=['-'], sep='\t', names = ['o_d1','o_d2','o_d3','o_d4','o_d5','o_d6'])
zone_3532_d1_vector = pd.concat([zone_3532_flow_df['f_d1'], zone_3532_speed_df['s_d1'],
zone_3532_occupancy_df['o_d1']], axis=1)
def isNaN(x):
   return (x == x) == False
is_3532_f_d1_Nan = ~isNaN(zone_3532_d1_vector['f_d1'])
zone_3532_d1_vector = zone_3532_d1_vector[is_3532_f_d1_Nan]
is_3532_s_d1_Nan = ~isNaN(zone_3532_d1_vector['s_d1'])
zone_3532_d1_vector = zone_3532_d1_vector[is_3532_s_d1_Nan]
is_3532_o_d1_Nan = ~isNaN(zone_3532_d1_vector['o_d1'])
zone_3532_d1_vector = zone_3532_d1_vector[is_3532_o_d1_Nan]
zone_3532_d2_vector = pd.concat([zone_3532_flow_df['f_d2'], zone_3532_speed_df['s_d2'],
zone_3532_occupancy_df['o_d2']], axis=1)
is_3532_f_d2_Nan = ~isNaN(zone_3532_d2_vector['f_d2'])
zone_3532_d2_vector = zone_3532_d2_vector[is_3532_f_d2_Nan]
is_3532_s_d2_Nan = ~isNaN(zone_3532_d2_vector['s_d2'])
zone_3532_d2_vector = zone_3532_d2_vector[is_3532_s_d2_Nan]
is_3532_o_d2_Nan = ~isNaN(zone_3532_d2_vector['o_d2'])
zone_3532_d2_vector = zone_3532_d2_vector[is_3532_o_d2_Nan]
zone_3532_d3_vector = pd.concat([zone_3532_flow_df['f_d3'], zone_3532_speed_df['s_d3'],
zone_3532_occupancy_df['o_d3']], axis=1)
is_3532_f_d3_Nan = ~isNaN(zone_3532_d3_vector['f_d3'])
zone_3532_d3_vector = zone_3532_d3_vector[is_3532_f_d3_Nan]
is_3532_s_d3_Nan = ~isNaN(zone_3532_d3_vector['s_d3'])
zone_3532_d3_vector = zone_3532_d3_vector[is_3532_s_d3_Nan]
is_3532_o_d3_Nan = ~isNaN(zone_3532_d3_vector['o_d3'])
zone_3532_d3_vector = zone_3532_d3_vector[is_3532_o_d3_Nan]
zone_3532_d4_vector = pd.concat([zone_3532_flow_df['f_d4'], zone_3532_speed_df['s_d4'],
zone_3532_occupancy_df['o_d4']], axis=1)
is_3532_f_d4_Nan = ~isNaN(zone_3532_d4_vector['f_d4'])
zone_3532_d4_vector = zone_3532_d4_vector[is_3532_f_d4_Nan]
is_3532_s_d4_Nan = ~isNaN(zone_3532_d4_vector['s_d4'])
zone_3532_d4_vector = zone_3532_d4_vector[is_3532_s_d4_Nan]
is_3532_o_d4_Nan = ~isNaN(zone_3532_d4_vector['o_d4'])
zone_3532_d4_vector = zone_3532_d4_vector[is_3532_o_d4_Nan]
```

```
zone_3532_d5_vector = pd.concat([zone_3532_flow_df['f_d5'], zone_3532_speed_df['s_d5'],
zone_3532_occupancy_df['o_d5']], axis=1)
is_3532_f_d5_Nan = ~isNaN(zone_3532_d5_vector['f_d5'])
zone_3532_d5_vector = zone_3532_d5_vector[is_3532_f_d5_Nan]
is_3532_s_d5_Nan = ~isNaN(zone_3532_d5_vector['s_d5'])
zone_3532_d5_vector = zone_3532_d5_vector[is_3532_s_d5_Nan]
is_3532_o_d5_Nan = ~isNaN(zone_3532_d5_vector['o_d5'])
zone_3532_d5_vector = zone_3532_d5_vector[is_3532_o_d5_Nan]
zone_3532_d6_vector = pd.concat([zone_3532_flow_df['f_d6'], zone_3532_speed_df['s_d6'],
zone_3532_occupancy_df['o_d6']], axis=1)
is_3532_f_d6_Nan = ~isNaN(zone_3532_d6_vector['f_d6'])
zone_3532_d6_vector = zone_3532_d6_vector[is_3532_f_d6_Nan]
is_3532_s_d6_Nan = ~isNaN(zone_3532_d6_vector['s_d6'])
zone_3532_d6_vector = zone_3532_d6_vector[is_3532_s_d6_Nan]
is_3532_o_d6_Nan = ~isNaN(zone_3532_d6_vector['o_d6'])
zone_3532_d6_vector = zone_3532_d6_vector[is_3532_o_d6_Nan]
zone_3532_d1_vector.columns = ['flow','speed','occupancy']
zone_3532_d2_vector.columns = ['flow','speed','occupancy']
zone_3532_d3_vector.columns = ['flow','speed','occupancy']
zone_3532_d4_vector.columns = ['flow','speed','occupancy']
zone_3532_d5_vector.columns = ['flow','speed','occupancy']
zone_3532_d6_vector.columns = ['flow','speed','occupancy']
zone_3532_vector = zone_3532_d1_vector
zone_3532_vector = zone_3532_vector.append(zone_3532_d2_vector)
zone_3532_vector = zone_3532_vector.append(zone_3532_d3_vector)
zone_3532_vector = zone_3532_vector.append(zone_3532_d4_vector)
zone_3532_vector = zone_3532_vector.append(zone_3532_d5_vector)
zone_3532_vector = zone_3532_vector.append(zone_3532_d6_vector)
from scipy.stats import multivariate_normal
import numpy as np
mean = np.mean(zone_3532_vector)
cov = zone_3532_vector.cov()
zone_3532_vector['probability'] = multivariate_normal.pdf(zone_3532_vector, mean, cov)
zone_3532_vector['flow'] = zone_3532_vector['flow'].astype(int)
zone_3532_vector['speed'] = zone_3532_vector['speed'].astype(int)
zone_3532_vector['occupancy'] = zone_3532_vector['occupancy'].astype(int)
is_occ_neg = zone_3532_vector['occupancy'] <0
zone_3532_vector_is_neg_occ = zone_3532_vector[is_occ_neg]
is_speed_neg = zone_3532_vector['speed'] <0
zone_3532_vector_is_neg_speed = zone_3532_vector[is_speed_neg]
is_flow_neg = zone_3532_vector['flow'] <0
zone_3532_vector_is_neg_flow = zone_3532_vector[is_flow_neg]
```

```
zone_3532_vector_is_neg =
zone_3532_vector_is_neg_flow.append(zone_3532_vector_is_neg_speed.append(zone_3532_vector_is
_neg_occ))
zone_3532_vector_is_neg = zone_3532_vector_is_neg.drop_duplicates()
zone_3532_vector_is_neg['probability'] = 0
zone_3532_vector = zone_3532_vector.append(zone_3532_vector_is_neg)
zone_3532_vector['index'] = zone_3532_vector.index
zone_3532_vector= zone_3532_vector.drop_duplicates(['index'], take_last = True)
zone_3532_vector.set_index = zone_3532_vector['index']
zone_3532_vector = zone_3532_vector.drop('index', axis =1)
zone_3532_vector = zone_3532_vector.sort_index()
zone_3532_sorted_vector = zone_3532_vector.sort(['probability'])
zone_3532_sorted_vector = zone_3532_sorted_vector.reset_index(drop=True)
is_100th_row = zone_3532_sorted_vector.index % 100 == 0
zone_3532_sorted_vector_is_100th = zone_3532_sorted_vector[is_100th_row]
zone_3532_sorted_vector_is_100th['probability'] =
zone_3532_sorted_vector_is_100th['probability'].apply('{:.8f}'.format)
zone_3532_sorted_vector_is_100th.to_csv("/home/datascience/Downloads/cleaning1/3532/3532.txt",
sep='\t', index=False)
```