In [25]:
```
import numpy as np
import pandas as pd
d = pd.read_csv(r"Behavior of the urban traffic of the city of Sao Paulo in Brazil.csv")
d.head()
```

Out[25]:

	Hour (Coded)	Immobilized bus	Broken Truck	Vehicle excess	Accident victim	Running over	Fire vehicles	Occurrence involving freight	Incident involving dangerous freight	Lack of electricity	Fire	Point of flooding	Manifestations	Defect in the network of trolleybuses	Tre c th roa
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	
2	3	0	0	0	0	0	0	0	0	0	0	0	0	0	
3	4	0	0	0	0	0	0	0	0	0	0	0	0	0	
4	5	0	0	0	0	0	0	0	0	0	0	0	0	0	

In [26]:
```
#Encoding
from sklearn.preprocessing import OneHotEncoder
o = OneHotEncoder()
d['Hour (Coded)']=pd.cut(x=d['Hour (Coded)'], bins=[0,5,10,15,20,27])
d["Hour (Coded)"]=d["Hour (Coded)"].astype("category")
d["Hour (Coded)"] = d["Hour (Coded)"].cat.codes
day = ["morning","noon","afternoon","evening","night"]
q=pd.DataFrame(o.fit_transform(d[["Hour (Coded)"]]).toarray(),columns=day)
d = d.join(q)
d.head(100)
```

Out[26]:

	Hour (Coded)	Immobilized bus	Broken Truck	Vehicle excess	Accident victim	Running over	Fire vehicles	Occurrence involving freight	Incident involving dangerous freight	Lack of electricity	...	Defect in the network of trolleybuses	Tree on the road	Semaphore off	Intermitten Semaphor
0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
1	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
2	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
3	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
4	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
95	2	1	2	0	0	0	0	1	0	0	...	0	0	1	
96	3	0	3	0	1	0	0	0	0	0	...	0	0	0	
97	3	0	1	0	1	0	0	0	0	1	...	0	0	0	
98	3	0	1	0	0	0	0	0	0	0	...	1	0	0	
99	3	0	1	0	0	0	0	0	0	0	...	0	1	0	

100 rows × 23 columns

In [30]:
```
from sklearn.preprocessing import normalize
names=d.columns
s = normalize(d)
d = pd.DataFrame(s,columns=names)
x = d.drop(["Slowness in traffic (%)","Hour (Coded)"],axis=1)
y = d.iloc[:,17:18]
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.2)
```

In [31]:
```
#from sklearn.model_selection import train_test_split
#x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.2)
from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import PolynomialFeatures
l = LinearRegression()
p = PolynomialFeatures()
p=p.fit(x_train,y_train)
l=l.fit(x_train,y_train)
y_pred = l.predict(x_test)
```

In [32]:
```
from sklearn.metrics import mean_squared_error,mean_absolute_error,median_absolute_error,r2_score
mean = mean_squared_error(y_test,y_pred)
mean_ab = mean_absolute_error(y_test,y_pred)
med_ab = median_absolute_error(y_test,y_pred)
r = r2_score(y_test,y_pred)
print(mean)
print(mean_ab)
print(med_ab)
print(r)
```

```
0.002456313441402896
0.027018715865154534
0.008551448868987066
0.08634448378917459
```

In [ ]:
```
```