MAT 3772 Project 2

November 15, 2019

0.1 MAT 3772 Project 2

0.1.1 ABSTRACT

Some people claim in poor countries that as the population of a given area increase the average income tends to decrease.

0.1.2 Introduction

Our goal in this analysis is to determine if the relationship between income and population density is linear.We will breakdown our analysis as follows.

****TABLE OF CONTENTS****

1.) Procedure:

We will use a **one variable linear model of the form Y = aX + b** where **X** represent the **population density** and **Y** the **income per Capita**

2.) Variables

Y: represent the **population density**

X: the **income per Capita**

[12]:
```
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
```

3) Determining if there is a linear relationship the income and population density

[13]:
```
# Represent the outcome or population density or output
X = np.array([168,117,122,171,172,142,188,267,252,325])

# Represent the income per Capita or input vector
Y = np.array([290,101,1740,139,71,76,458,284,578,45])
```

[14]:
```
linearModel = LinearRegression()
reg = linearModel.fit(X.reshape(-1,1),Y)
```

[15]:
```
reg.coef_
```

[15]:
```
array([-2.07524704])
```

[16]:
```
reg.intercept_
```

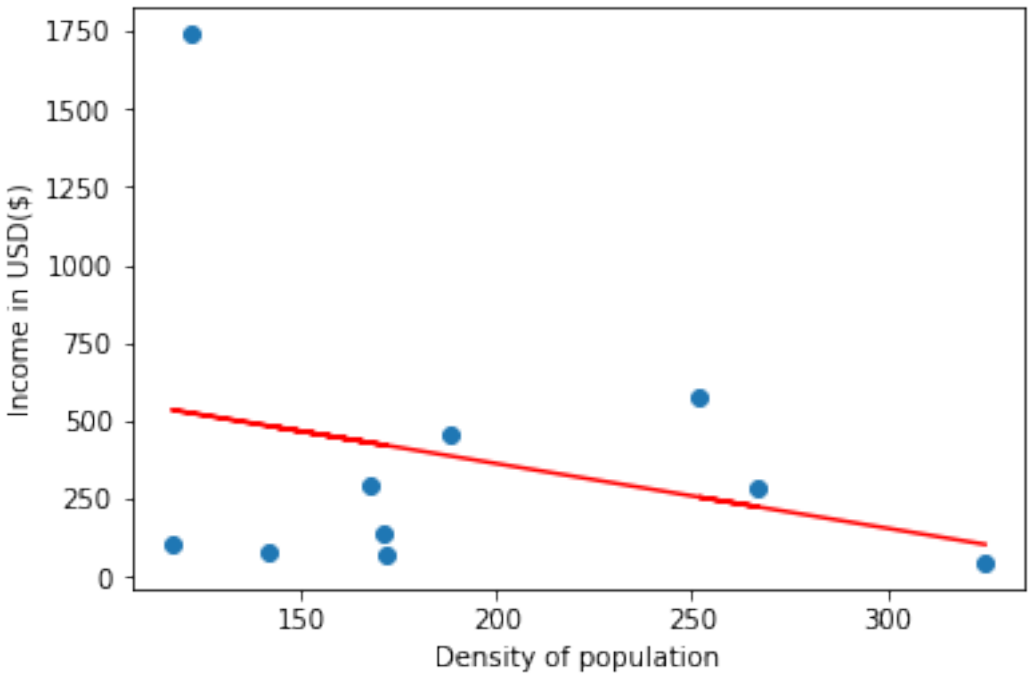
[16]:
```
777.4775298280897
```

The model is Y = -0.0365X + 206.227

[17]:
```
plt.scatter(X,Y)
plt.plot(X,reg.predict(X.reshape(-1,1)),color='red')
plt.ylabel('Income in USD($)')
plt.xlabel('Density of population')
```

[17]:
```
Text(0.5,0,'Density of population')
```

[17]:



The variations between the predicted

4) Determining the R^2 value

[18]:
```
reg.score(X.reshape(-1,1),Y)
```

[18]: 0.07587371620591976

This model represents about **7.58% of the variations** of income and population density per square mile

5) By decreasing the population of Nepal for example how will it affects the income predicted by the model?

Nepal has 290 people living per square mile and has a $168 income. Decreasing 290 by 25% gives 217.5 people. Lets compute the predicted income

according to the model of 217 people per square mile.

Computing the predicted value can be done using predict function of the regression like the following

[19]:
```
reg.predict(np.array([217]).reshape(-1,1))
```

[19]: array([327.1489229])

The income almost doubles according to the model.

0.1.3 Conclusion

In this analysis, there was no a linear relationship between the data because R^2 value was close to 0(0.075). Therefore this linear model fails to predict effectively income for a given population density. That is why small changes in the income, for example for population density of Nepal which was 290, result a big change in the output.

[0]: