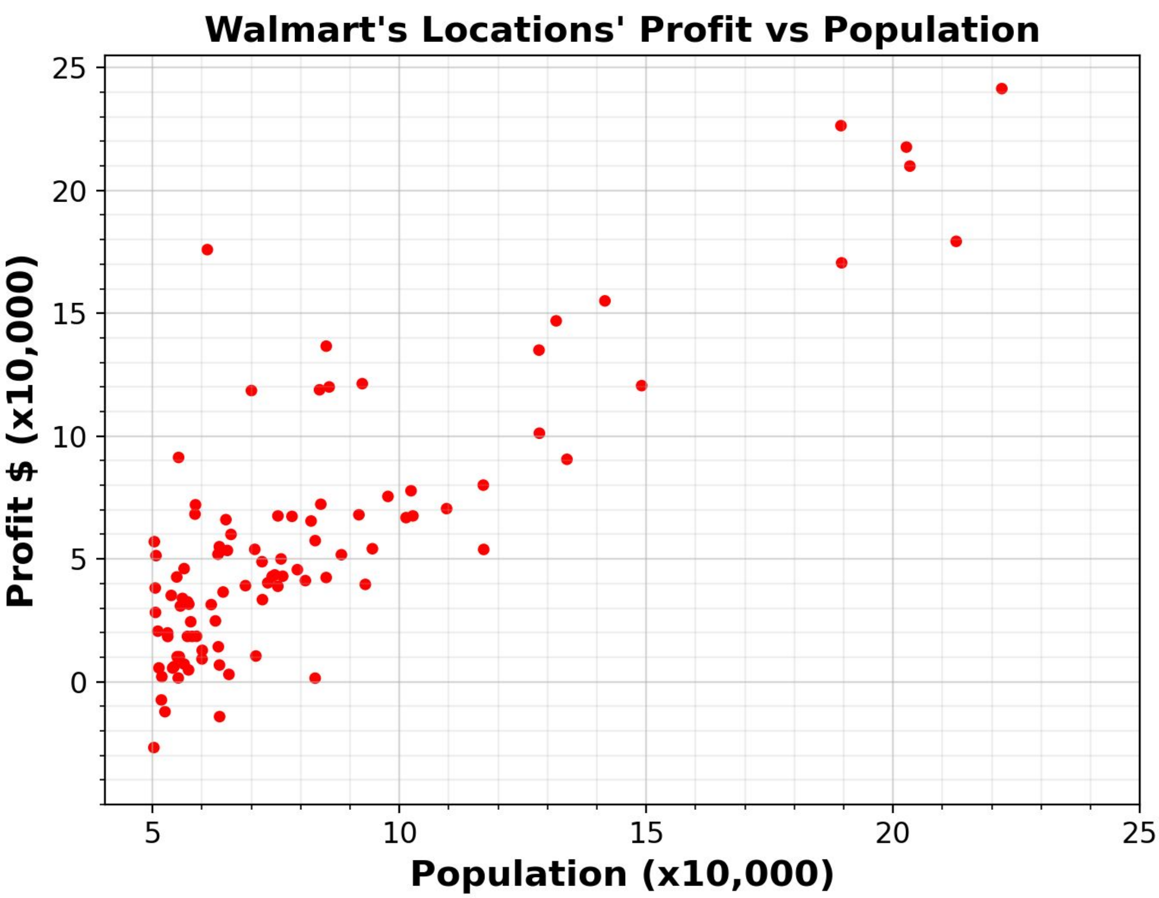
Daniel Lacayo 32247970
Michael Zink
ECE 241 FA 20
4 December 2020

Project 3

Problem 1.

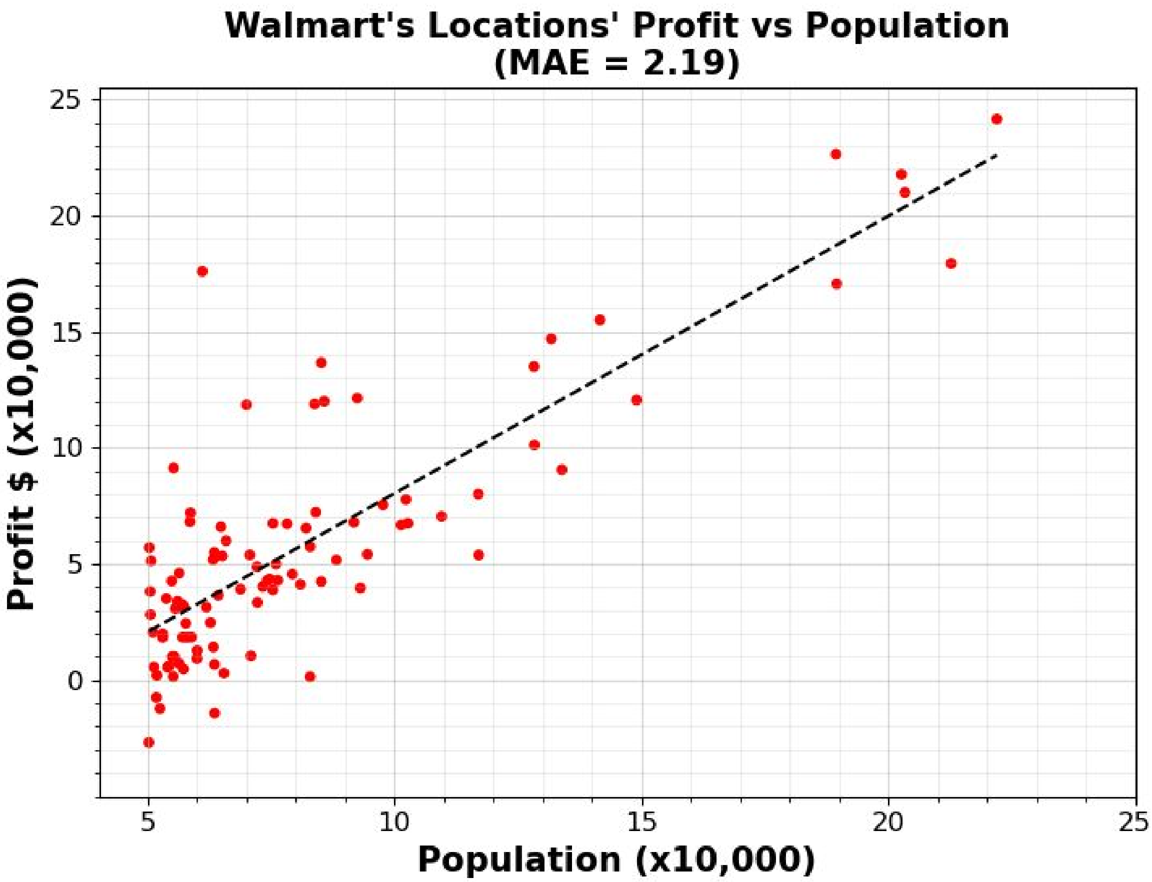
a) Plot



b) See method: plotData_with_model()

```
# model profit as a function of population
self.predict = np.poly1d(np.polyfit(self.population, self.profit, 1))
# self.predict(x) => y
```

c) Plot with model



d) This line shows a positive linear relationship between population and profit at Walmart stores. The mean absolute error is 2.19, signifying the correlation between population and profit is strong and reliable.

e)

City	Population (x10,000)	Profit $ (x10,000)
A	7.8	5.41
B	4.4	1.35
C	4.7	1.71
D	6.12	3.41

E	8.55	6.30
F	6.7	4.10
G	9.8	7.80
H	7.01	4.47

In descending order of profit: G > E > A > H > F > D > C > B

f) The mean absolute value of the model is 2.19, meaning this is a rather accurate model.

Output

```
/Users/daniellacayo/.conda/envs/Project3/bin/python
/Users/daniellacayo/PycharmProjects/Project3/Walmart.py
Future stores in descending order of profit:
Store G expects $7.80 (x10,000) of profit
Store E expects $6.30 (x10,000) of profit
Store A expects $5.41 (x10,000) of profit
Store H expects $4.47 (x10,000) of profit
Store F expects $4.10 (x10,000) of profit
Store D expects $3.41 (x10,000) of profit
Store C expects $1.71 (x10,000) of profit
Store B expects $1.35 (x10,000) of profit

Process finished with exit code 0
```

Sources:
https://socialstatisticsfun.wordpress.com/2012/12/18/deriving-alpha-and-beta/

Project 3, Problem 2

Daniel Lacayo
32247970

a)

$$f(x) = \sum_{i=1}^{m} (wx_i + \beta - y_i)^2$$

partial derivative with respect to β:

$$\frac{\partial f(x)}{\partial \beta} = \sum_{i=1}^{m} 2(\hat{w}x_i + \hat{\beta} - y_i)(0+1-0)$$

partial derivative with respect to w:

$$\frac{\partial f(x)}{\partial w} = \sum_{i=1}^{m} 2(\hat{w}x_i + \hat{\beta} - y_i)(x_i + 0 - 0)$$

set derivative to 0 $\longrightarrow$

$$0 = 2\sum_{i=1}^{m}(\hat{w}x_i + \hat{\beta} - y_i) \qquad\qquad 0 = \sum_{i=1}^{m} 2(\hat{w}x_i + \hat{\beta} - y_i)\, x_i$$

divide by 2, distribute sigma $\longrightarrow$

$$0 = \sum_{i=1}^{m} \hat{w}x_i + \hat{\beta}m - \sum_{i=1}^{m} y_i \qquad\qquad 0 = \sum_{i=1}^{m} \hat{w}x_i^2 + \sum_{i=1}^{m} \hat{\beta}x_i - \sum_{i=1}^{m} y_i x_i \quad (2)$$

$$\hat{\beta}m = \sum_{i=1}^{m} y_i - \sum_{i=1}^{m} \hat{w}x_i$$

$$\hat{\beta} = m^{-1}\sum_{i=1}^{m} y_i - m^{-1}\sum_{i=1}^{m} \hat{w}x_i \quad (1)$$

For the linear model $\sum_{i=1}^{m} \varepsilon_i^2 = \sum_{i=1}^{m} (wx_i + \beta - y_i)^2$ where $f(x) = \sum_{i=1}^{m} \varepsilon_i^2$

and if $y_i = wx_i + \beta + \varepsilon_i$, then the linear model that minimizes training error is: $\boxed{\hat{y}_i = \hat{w}x_i + \hat{\beta}}$

b) simplify in terms of variance & covariance

Goal:

$$\frac{COV_{xy}}{S_x^2} = \frac{\cancel{(N-1)^{-1}}\sum(x_i - \bar{x})(y_i - \bar{y})}{\cancel{(N-1)^{-1}}\sum(x_i - \bar{x})^2}$$

(for my sanity $\sum = \sum_{i=1}^{m}$)

Starting with eq. (2) and subs. eq. (1):

$$0 = \sum \hat{w}x_i^2 + \left(m^{-1}\sum y_i - m^{-1}\sum \hat{w}x_i\right)\sum x_i - \sum y_i x_i$$

$$0 = \sum \hat{w}x_i^2 + m^{-1}\sum y_i x_i - m^{-1}\sum \hat{w}x_i^2 - \sum y_i x_i$$

$$\sum \hat{w}x_i^2 - m^{-1}\sum \hat{w}x_i^2 = \sum y_i x_i - m^{-1}\sum y_i x_i$$

$$\hat{w}\sum x_i^2 - m^{-1}\sum x_i^2 = \sum y_i x_i - m^{-1}\sum y_i x_i$$

$$\hat{w} = \frac{\sum y_i x_i - m^{-1}\sum y_i x_i}{\sum x_i^2 - m^{-1}\sum x_i^2}$$

$$\left(m^{-1}\sum y_i x_i = N\bar{y}\bar{x}, \quad m^{-1}\sum x_i^2 = N\bar{x}^2\right)$$

$$\hat{w} = \frac{\sum y_i x_i - N\bar{y}\bar{x}}{\sum x_i^2 - N\bar{x}^2} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

Finally we have:

$$\boxed{\hat{w} = \frac{\sum_{i=1}^{m}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{m}(x_i - \bar{x})^2}}$$