

HWK 2 - Math 536

Elijah Amirianfar

2024-02-17

Download the dataset hw2.csv from Titanium. In this dataset you will find financial marketing data on 268 start-up companies working in the tech sector. The variables within this dataset are spend and revenue. Spend contains the amount of capital resources that were spent initially on over the first year of each company. Revenue quantifies the amount of revenue generated during the first year of the company. You have been hired by a new tech start-up who is interested in advertising. Please analyze the data and provide relevant output that addresses each of your employers asks. If you need any clarity on the data feel free to ask Dr. Nichols.

```
titanium <-  
  read.csv("~/My Drive/04. CSU FULLERTON 2023-2025/2. Spring 2024/MATH 536/Data Sets/hw2.csv")  
  
Spend = titanium$Spend # vector of the "spend" column in our Data Set  
Revenue = titanium$Revenue # vector of the "revenue" column in our data set
```

1. What is the relationship between advertising spend and revenue within the first year of a start-up?

```
model_titanium = lm(Revenue~Spend)  
# This gives us a linear regression model between Revenue and Spend  
model_titanium
```

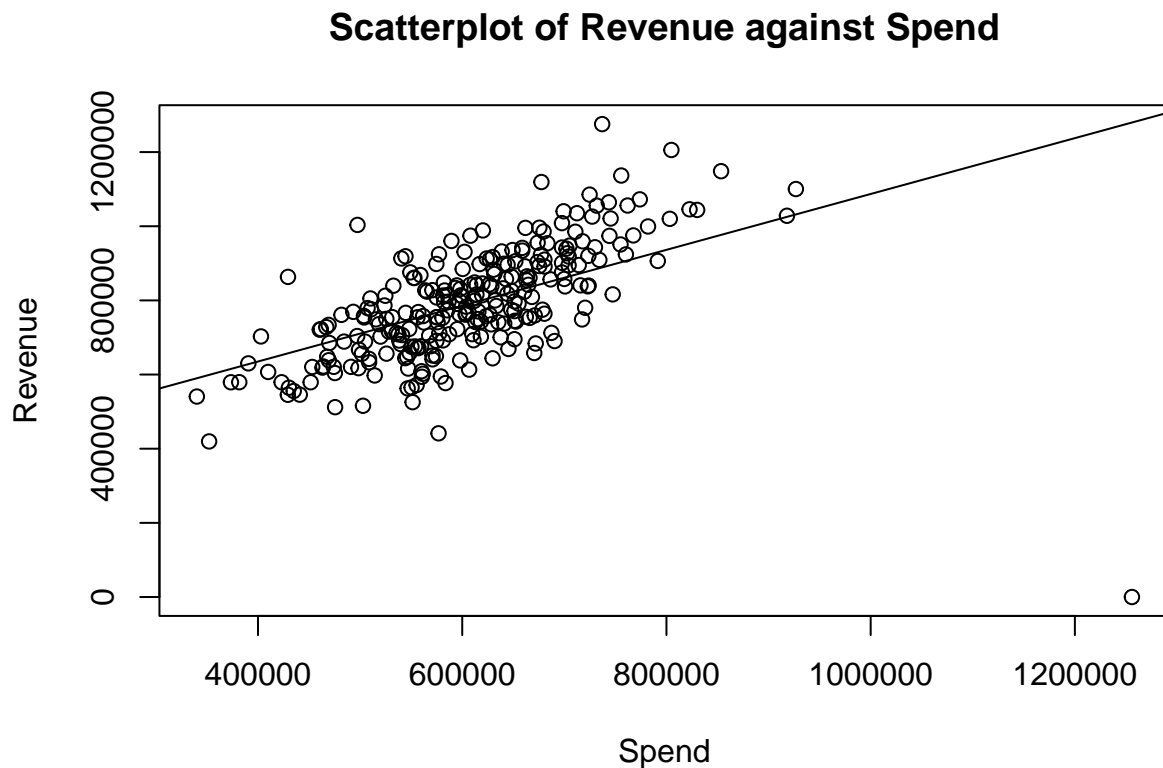
```
##  
## Call:  
## lm(formula = Revenue ~ Spend)  
##  
## Coefficients:  
## (Intercept)      Spend  
##  3.339e+05    7.531e-01
```

```
summary(model_titanium)
```

```
##  
## Call:  
## lm(formula = Revenue ~ Spend)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1279716   -70754    3568    63913   386477
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3.339e+05  4.642e+04  7.193 6.49e-12 ***
## Spend       7.531e-01  7.540e-02  9.988 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 128200 on 265 degrees of freedom
## Multiple R-squared:  0.2735, Adjusted R-squared:  0.2708
## F-statistic: 99.77 on 1 and 265 DF,  p-value: < 2.2e-16

# Gives us a summary of our model.
plot(Spend,Revenue, xlab="Spend", ylab="Revenue", main="Scatterplot of Revenue against Spend")
abline(lm(Revenue~Spend, data=titanium))
```



As we can see in our linear regression model, the relationship between the amount of money spent vs the amount of money in revenue is linear. In simpler terms, the more money we spend on titanium leads us to a higher revenue.

However, our model does not make sense 100% of the time because when we do not spend any money, our revenue is around \$339,000.

2. Currently the company is debating between spending \$500,000 and \$700,000 on advertising, please provide guidance.

```
newdata_500 = data.frame(Spend = 500000)
# Used to predict our revenue when 500k is spent
newdata_700 = data.frame(Spend = 700000)
# Used to predict our revenue when 700k is spent

predict(model_titanium,newdata_500,interval="predict")
```

```
##          fit          lwr          upr
## 1 710461.1 457082.9 963839.2
```

```
# Predicts our confidence interval for how much revenue we make when spending 500k
predict(model_titanium,newdata_700,interval="predict")
```

```
##          fit          lwr          upr
## 1 861078 607819 1114337
```

```
# Predicts our confidence interval for how much revenue we make when spending 700k
```

In the above data, we can see that if the company spends \$500,000 on Titanium, our revenue is between \$457,082.90 and \$963,839.20. However, if the company spends \$700,000 on Titanium, our revenue is between \$607,819 and \$1,114,337.

If we apply these intervals to our linear regression model above, we obtain an approximate revenue of \$710,461 when \$500k is spent on titanium (which gives us a profit of around \$210k), and an approximate revenue of \$861,078 when \$700k is spent on titanium (which gives us a profit of around \$161k).

After looking at the profits, and taking into account our linear regression model, it is better to spend \$500k on titanium versus spending \$700k since the profit is higher after spending \$500k.