**Devon Morgan**

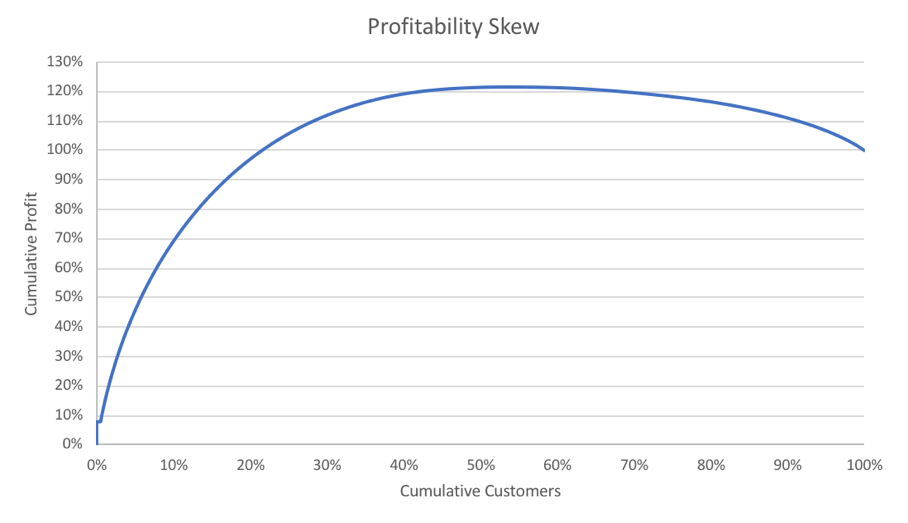
MKTK 3050 Pilgrim Bank Case Assignment

**Part 1:** Heterogeneity in Profitability

1. Need a minimum of about 21% of customers to realize 100% of current profit.
2. About 47% of customers are unprofitable.

122% of total profit could realize by targeting profitable customers only.

1. Age and income could explain variation in profitability across customers. Since most teenagers will likely have a small account balance, each time they transact with a bank it will lose money. On the other hand, adults are likely a profitable group since they likely have a larger account balance and the interest revenue offsets their transaction costs.



**Part 2:** Effect of Online/Offline – Introductory Analysis

1. Online average profitability = $116.67

Offline average profitability = $110.79

Standard error of online average profitability = $1.63

Standard error of offline average profitability = $4.57

95% confidence interval of online = $3.31

95% confidence interval of offline = -$15.07



Standard error reflects uncertainty in the data due to random error while confidence interval shows that the true value of the percentage is likely between the upper and lower CI.

From my visual, I believe there is a significant difference between the two groups because the average profit of online customers is higher with more certainty than the offline customers.

1. T-Value = -1.25389

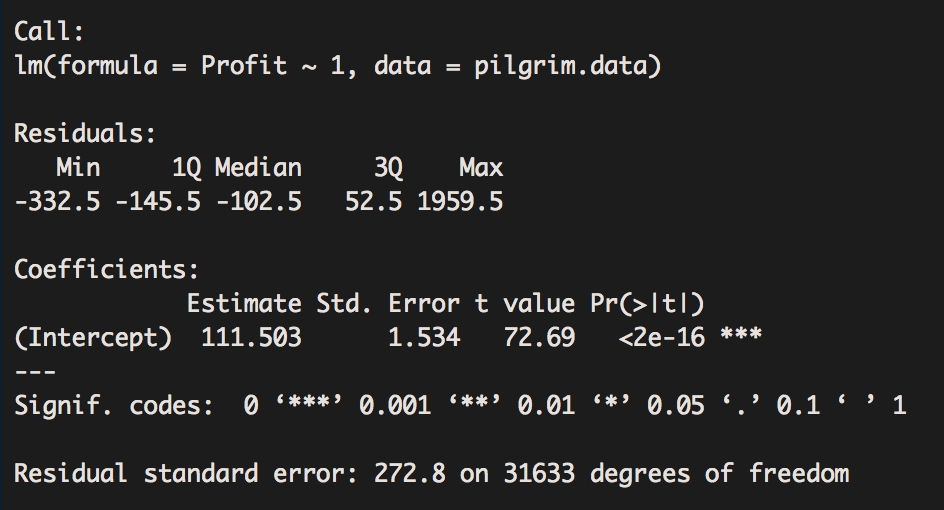
Degrees of freedom = 31632

P-value = 0.2099

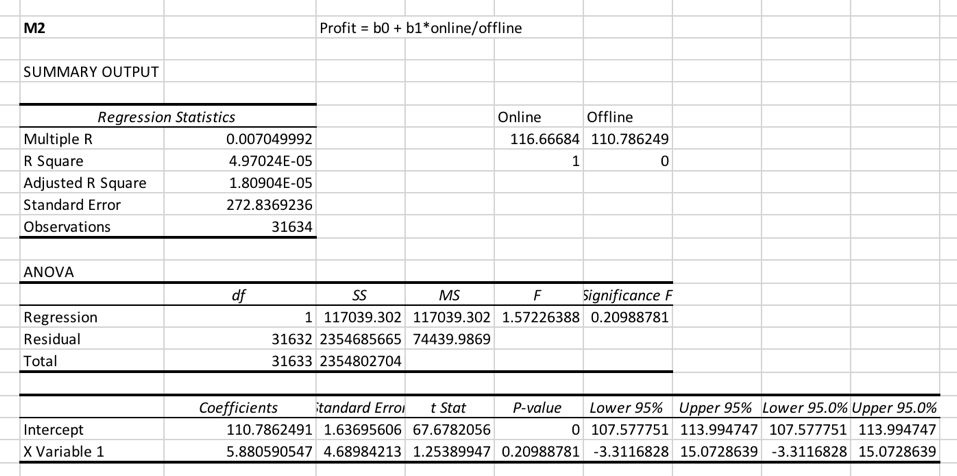
A T-test is used to determine if two populations means are equal. This test shows that the offline mean is smaller than the online mean.

**Part 3:** Simple Regression

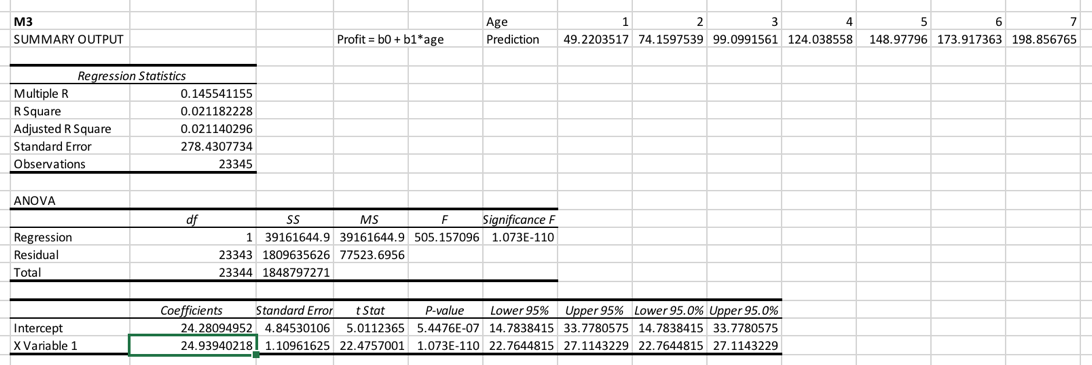
1. **Model 1:** The parameter estimate of the intercept, $111.50, is interpreted as the predicted average profit for offline customers.



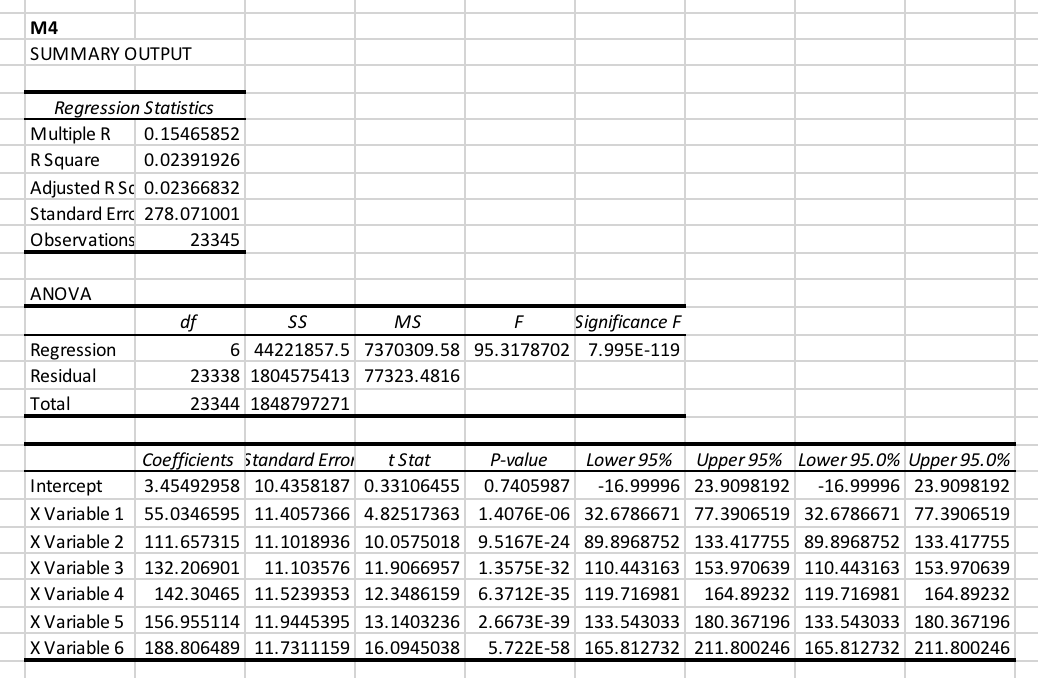
1. **Model 2:** This model relates to the t-test above because they both yield the same result considering that it shows how coefficients of the intercept are the same. The intercept is the predicted value when all variables equal 0. The coefficients for online/offline mean that there is $5.88 difference in profit between online and offline customers.



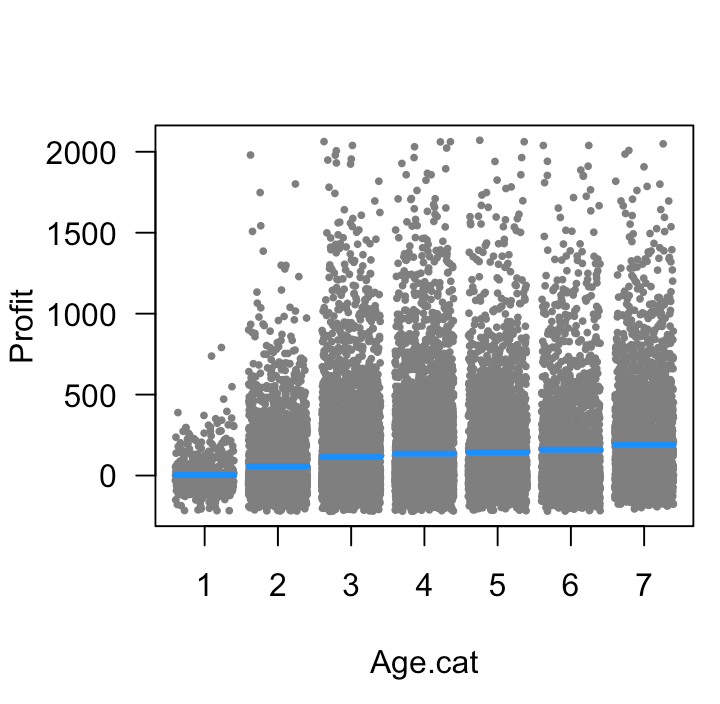
1. **Model 3:** The coefficient of age is the effect of age, meaning that as each age bracket increases by 1, profit increases by $25. The intercept is the predicted value for people in the age range of 0, which does not exist. This model shows that age is a good predictor of profit.



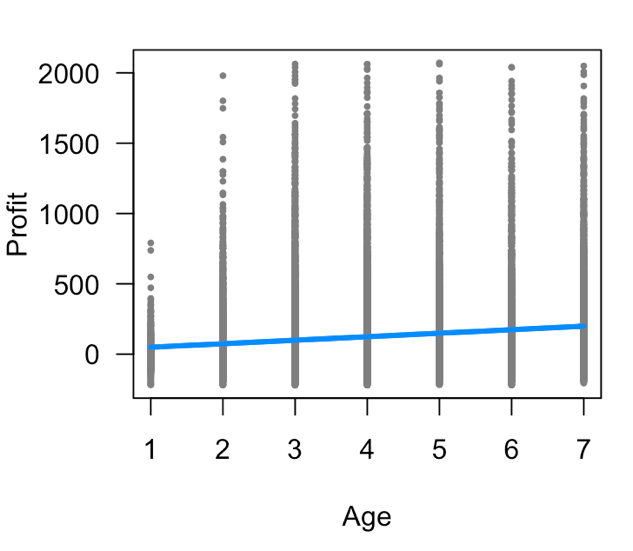
1. **Model 4:** The intercept is predicting value for age bracket 1. The coefficients show how much different each age bracket is from the reference category.



1. There is more flexibility in the categorical model since it allows for more parameters. On the other hand, the continuous model is simpler because it restricts the number of parameters to two. As a result, the categorical model reveals more variation in the profit.



Categorical Graph



Continuous Graph

**Part 4: Multiple Regression**

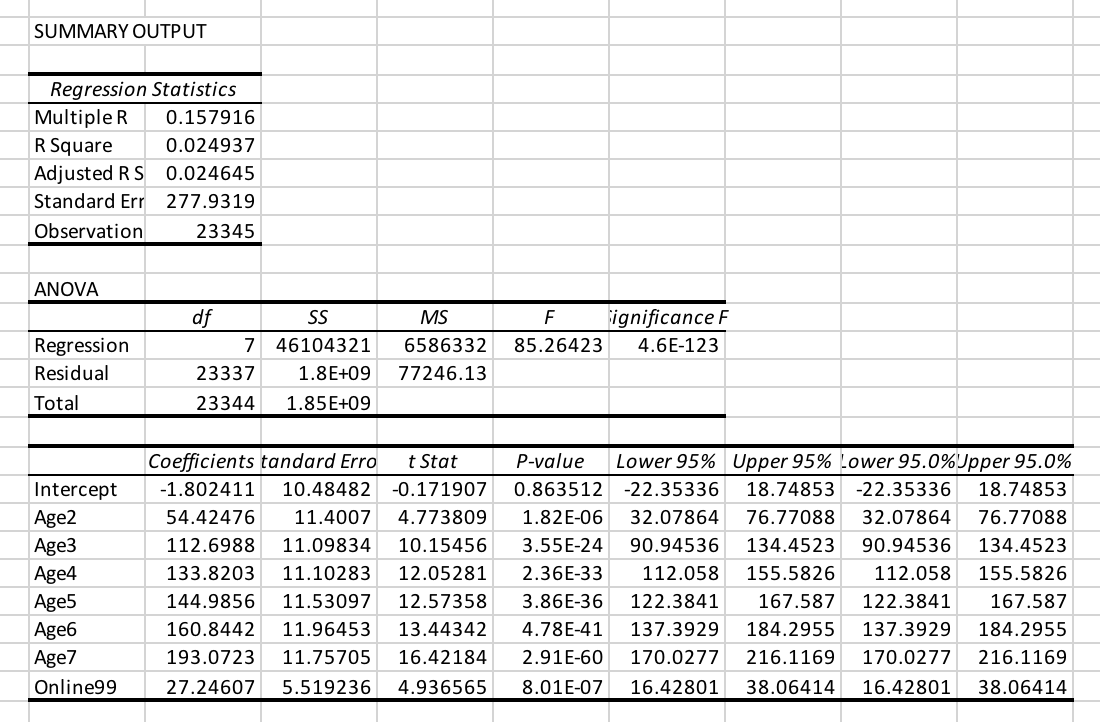
1. Averages

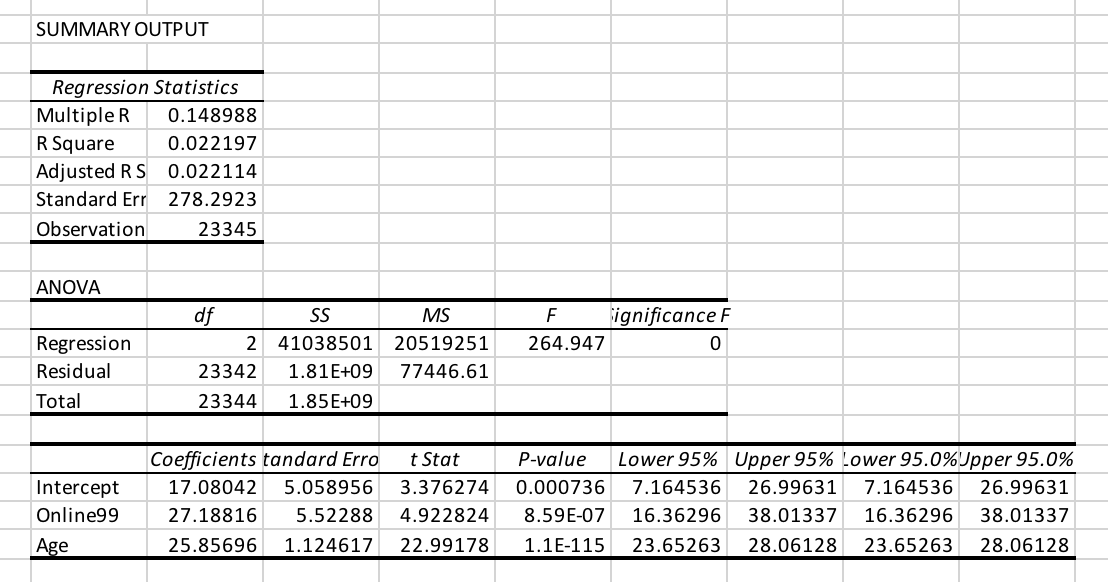
|  |  |  |
| --- | --- | --- |
|  | **Offline** | **Online** |
| Age | 4.1505 | 3.3366 |
| Income | 5.3854 | 5.9533 |
| Tenure | 10.372 | 8.6538 |
| Geographic Distribution | 1203.125 | 1203.633 |

The average age and tenure differ the most between groups.

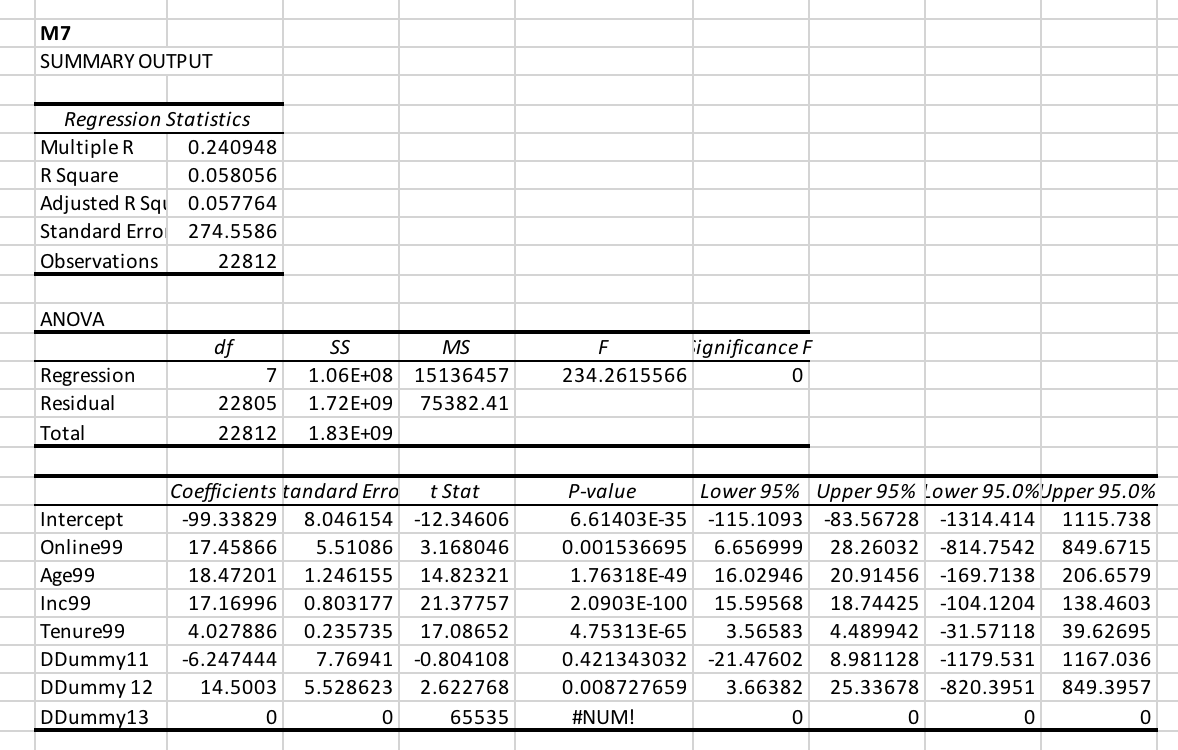
This raises issues for the t-test analysis and Model 2 in part 3 because it assumes equal variances, but they are not equal. they do not have equal variances.

1. **Model 5 & Model 6:** In Model 5, coefficients for age are interpreted as. For Model 6, one interprets the age coefficient as. I believe Model 5 is better because it adjusts for the age buckets that appear to have an impact on profit.





1. **Model 7:** This model shows which predictors have the most significance on profit.

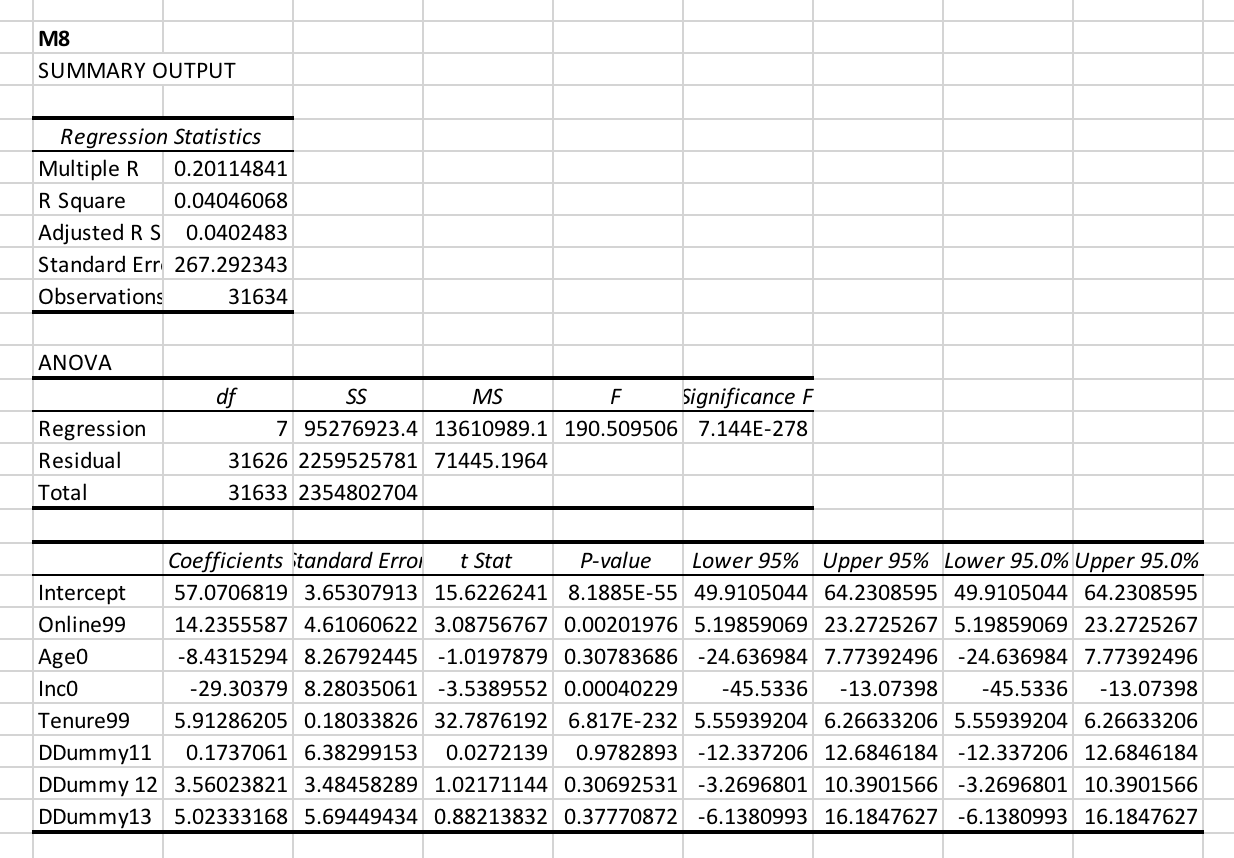


1. Customer 1: $200.95

Customer 2: $69.67

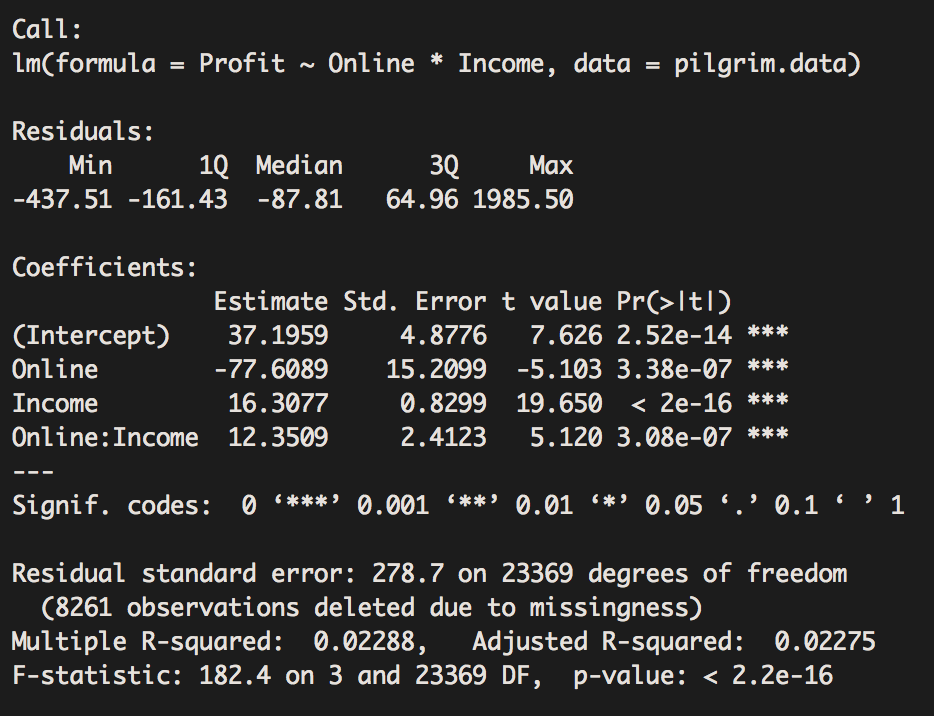
**Part 5: Missing Data**

1. Missing data is a problem because it can reduce the power of a model, lead to biased estimates, and cause invalid conclusions.
2. **Model 8:**



**Part 6: Interaction Modeling**

1. **Model 9:**

****

**Part 7: Reflection and Recommendations**