FaceLyft: Using Multiple Linear Regression Models to Predict Total Interactions to Effectively Position Lyft Driver Base

By: Shivani Karingula

Variables for Multivariate Linear Regression:

Predictors (Independent Variables):

$X_{1=}$ Lifetime Post Consumers

$X_{2=}$ Lifetime Engaged Users

$X_{3=}$ Lifetime Post Total Impressions

Target (Dependent Variable):

Total Interactions

Regression Results

OLS Regression Results

Dep. Variable:	Total Interactions	R-squared:	0.963
Model:	OLS	Adj. R-squared:	0.963
Method:	Least Squares	F-statistic:	4345.
Date:	Tue, 05 Nov 2019	Prob (F-statistic):	0.00
Time:	15:17:58	Log-Likelihood:	-2852.8
No. Observations:	500	AIC:	5714.
Df Residuals:	496	BIC:	5731.
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	9.4309	4.474	2.108	0.036	0.641	18.221
Lifetime Post Consumers	-1.3738	0.015	-91.267	0.000	-1.403	-1.344
Lifetime Engaged Users	1.4126	0.014	102.603	0.000	1.386	1.440
Lifetime Post Total Impressions	-1.568e-06	4.65e-05	-0.034	0.973	-9.3e-05	8.99e-05

Omnibus:	201.197	Durbin-Watson:	1.703
Prob(Omnibus):	0.000	Jarque-Bera (JB):	5915.368
Skew:	1.113	Prob(JB):	0.00
Kurtosis:	19.703	Cond. No.	1.13e+05

Regression Equation

$$\text{Y-Pred} = -1.3738X_1 + 1.4126\,X_2 + -1.568e^{-06}X_3$$

Analysis:

- This shows us that ‘Lifetime Post Consumers’ is negatively correlated to the prediction of Total Interactions while ‘Lifetime Engaged Consumers’ is positively correlated to the prediction of Total Interactions. For a given amount of Lifetime Engaged Consumers and Lifetime Post Total Impressions, an increase of 1 Lifetime Post Consumer is associated with a decrease in Total Predictions by 1.3738. Similarly, for a given amount of Lifetime Post Consumers and Lifetime Post Total Impressions, an increase of 1 Lifetime Engaged Consumer is associated with an increase in Total Predictions by 1.4126.

Model Accuracy/Fixes

Because the coefficient of the third predictor, Lifetime Post Total Impressions, is so close to 0, we can assume that there is not enough strength to validate this relationship, meaning that there is likely no correlation between Lifetime Post Total Impressions and Total Interactions. Another indicator that this is a poor predictor is it's high p-value of of 0.973 whereas the first two predictors have a p-value of 0.00.

The model may be more accurate using just the Lifetime Post Consumers and Lifetime Engaged Variables to predict the number of Total Interactions to deploy the supply of Lyft drivers more effectively.