hw07 Part I

Bohan Yin

11/14/2019

Contents

Estimate a basic (single variable) linear regression model of the relationship between the importance of the video and feelings towards Donald Trump. 1

Estimate a linear regression model of attitudes towards Donald Trump given the variables you have available. 3

Coefficient plot 4

Evaluation 4

Estimate a basic (single variable) linear regression model of the relationship between the importance of the video and feelings towards Donald Trump.

The basic linear regression model of the relationship between the importance of the video and feelings towards Donald Trump is expressed as:

- Attitudes = beta0 + beta1video

Here is the result of the regression model. As we can see, the p.value is zero, meaning that the video variable is very significant in influncing feelings towards Trump in this basic model.

```
## # A tibble: 2 x 5
##   term        estimate std.error statistic p.value
##   <chr>          <dbl>     <dbl>     <dbl>   <dbl>
## 1 (Intercept)     71.2     0.779      91.4       0
## 2 video          -16.1     0.295     -54.6       0
```

Here is a visualized result. From Figure 1 we can see that as importance of the video drops, the feelings towards Trump rises.

Figure 1: The relationship between the importance of the video and feelings towards Donald Trump

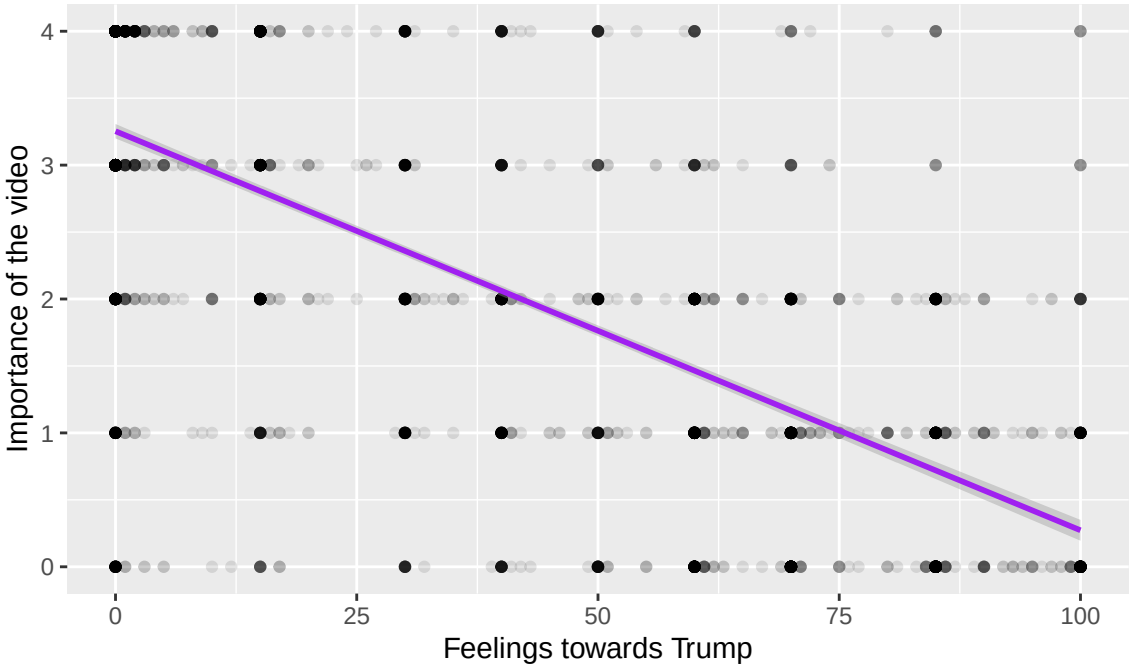
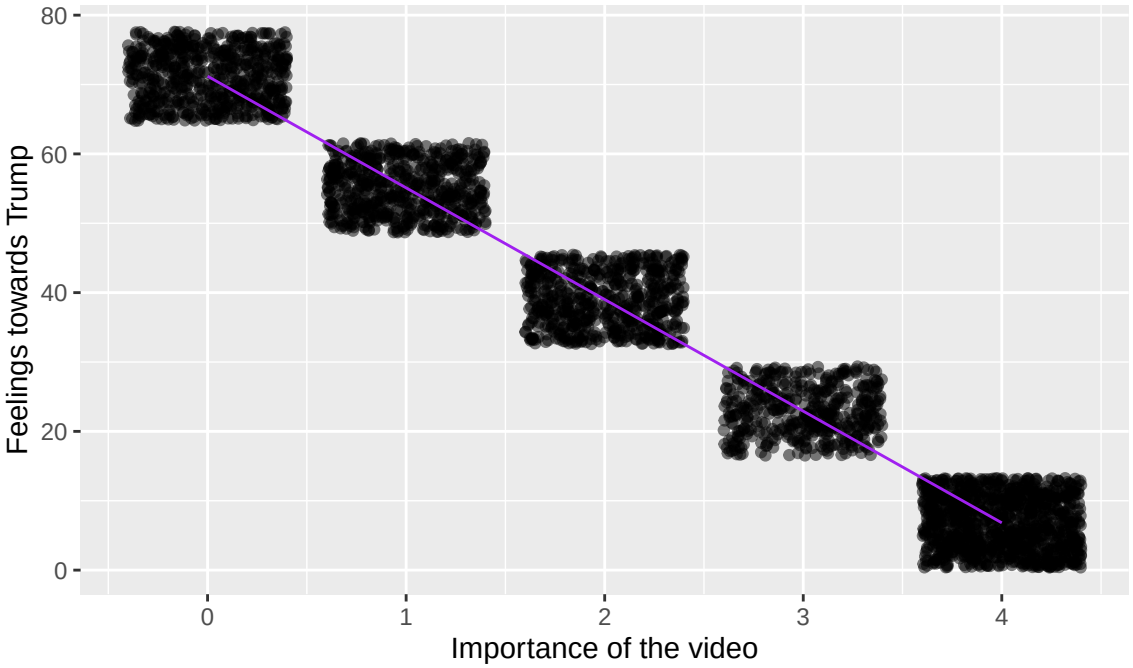


Figure 2: Relationship between the importance of the video and feelings towards Donald Trump, using predicted values



If we want to see the predictive power of this model, Figure 2 offers a more intuitive visualization. From Figure 2, we can see that the feelings and importance of the video are clustered in different rectangles. Each rectangle represents different set of combinations of the range of feelings and range of video importance. For example, the second rectangle starting from the left shows that people who rate importance of the video as 1 feel warmth toward Trump ranging from 50 to 60. In addition, the fifth rectangel starting from the left

seems the darkest, meaning that this cluster has the most people who think video is very important and don't like Trump.

Estimate a linear regression model of attitudes towards Donald Trump given the variables you have available.

The linear regression model of attitudes towards Donald Trump given the variables available is written as:

- Attitudes = beta0 + beta1video + beta2female + beta3pi + beta4age + beta5educ + beta6video*pid

Notice that I added interaction term: video*pid, as there might be some joint effects of both attitude toward video and party identification. The table below shows the result of the model:

Table 1: Regression Results Importance of the Video and Feelings towards Trump

	Dependent variable:
	Feelings towards Trump
Importance of the video	-7.468^{***}
	(0.491)
Female	-0.851
	(0.773)
Party Identification(pid)	8.761^{***}
	(0.357)
Age	0.124^{***}
	(0.022)
Education	-1.304^{***}
	(0.170)
Joint effect of video and pid	-0.992^{***}
	(0.136)
Constant	39.656^{***}
	(2.522)
Observations	3,231
R^2	0.612
Adjusted R^2	0.612
Residual Std. Error	21.739 (df = 3224)
F Statistic	849.054^{***} (df = 6; 3224)
Note:	*p<0.1; **p<0.05; ***p<0.01

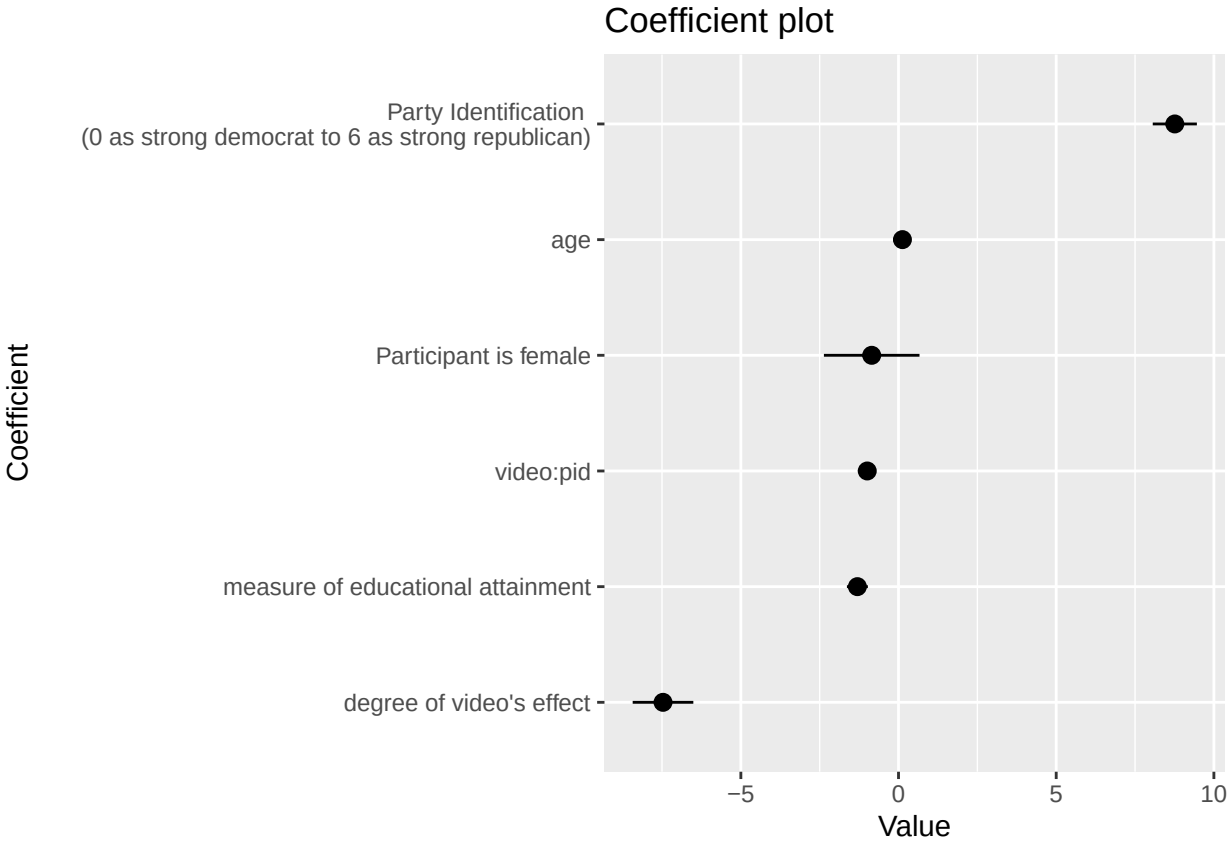
From the table, we can see that besides female variable, the p-value for all other variables are smaller than 0.05, meaning they are significant in impacting the attitude towards Trump, and we will ignore the effect of female as it is not significant. Specifically, when we talk about participants' attitudes toward Trump in terms of feeling thermometer rating:

- For every unit increase in video response, the attitude drops roughly 7.5 on the rating.

- For every unit increase in party identification, or becomeing more republican, the attitude rises roughly 8.8 on the rating.
- For every unit increase in age, the attitude slightly rises 0.1 on the rating.
- For every unit increase in education, the attitude drops roughly 1.3 on the rating.

Coefficient plot

If we want to check the result in an intuitive manner, here is the coefficient plot that contains coefficients of variables on affecting the attitude:



Same as the analysis stated previously, this plot offers a more intuitive result. As we can see, the increase in party identification positively affects the attitude toward Trump, meaning that the more republican you are, the more favor you have on Trump; the increase in degree of video's effect negatively affects the attitude toward Trump, meaning that the more important you think the video is, the less favor you have on Trump.

Evaluation

We can also do a simple evaluation of this model. As we can see from Table 2 below, the R-squared is 0.6124219, which means that the model explains 61.24% variations of the dependent variable (in this case, the attitude towards Trump).

Table 2: Regression Results

Variable	Coefficient	Std. Error	T-Statistic	P Value
(Intercept)	39.6562	2.5222	15.7226	0.0000
video	-7.4683	0.4906	-15.2224	0.0000

Variable	Coefficient	Std. Error	T-Statistic	P Value
female	-0.8505	0.7732	-1.1001	0.2714
pid	8.7609	0.3569	24.5466	0.0000
age	0.1241	0.0221	5.6190	0.0000
educ	-1.3044	0.1697	-7.6880	0.0000
video:pid	-0.9920	0.1361	-7.2892	0.0000

Here is the residual plot for the model, which only has symmetrical distribution when predicted attitude is around 50. And residuals do not really clustered around the lower single digits of the y-axis. This means tht this model still has a huge room of improvement.