Real Statistical Analysis From Real Men
Part I
What type of Dickheads make more money?

Pendacan

June 10, 2018

1 Introduction

The age old question: What makes a person rich?

What makes Sean earn more money than any other audio engineer? What makes other Dickhead researchers earn over $\$15,000$ a year, but I, also in the research field, earn less than $\$9,000$ a year? Is it because I'm gay[1]? Let's find out!

Skipping literature review because nobody cares, I will begin with **Section 2.1:"Data Transformation"** which describes the transformations I did to the variables from the TDS Census data[2]. Then in **Section 2.2:"How To Read Stats"** I will explain how you can interpret the numbers in regression models. Finally in **Section 3:"Analysis"** I will have many a stats for you.

2 Data

For this analysis I used the second wave of TDS Census data collected between 6/4/2018 and 6/7/2018, which includes more Facebook users than in the initial wave. Filthy, I know, but it is a sacrifice we must make in order to get a more general picture of the entire TDS community.

The data set contains 725 entries, with the obviously joke responses already deleted by Lago-m-orph. There were 21 questions to the survey, 3 of which were exclusive to female or tranny respondents. I deleted those questions because they don't matter. I also deleted some variables that are irrelevant to the current question at hand, or will serve as the dependent variable in other investigations.

Entries with ANY missing data in the relevant fields are removed. This leaves us with 523 entries for analysis. Your petty concern for privacy is getting in the way of stats. If you want to be relevant, answer every question.

2.1 Data Transformation

I leave **Age** and **Income** in their original state as **continuous variables** because they are more useful for our purposes serving as exact numbers instead of being sliced into ordinal categories. We want to know how to get more money, not how to get into an income range.

Most of our variables are **categorical**, which means their values are "apples and oranges" instead of "2 apples and 15 apples." I transform them into **dummy variables**. This means their values are cut into strict "1"s and "0"s. This is simple to do for questions with one "yes," and slightly less simple to do for questions that are multiple-choice. For multiple-choice

[1] I'm not actually gay.

[2] Lago-m-orph. 2018

questions I need to extract a new dummy variable for every answer option denoting "yes" to that particular option and "no" to all other ones. This is important for later.

2.2 How To Read Stats

If the standard "effect of x on y" plot is a monogamous money shot, Multiple Regression Model is a bukake. We pit many "x"s against one another for the ultimate prize - statistical significance.

In each regression table you will see many numbers. These numbers are coefficients. The bigger a coefficient is the bigger that independent variable's effect is on the dependent variable. However, just like a big cock that can't put out, the coefficients don't matter unless they have *s next to them. The more *s they have, the more statistical significant they are. Rule of thumb: Look for numbers with stars next to them. Any statistician will stab my guts for dare uttering the previous sentence, but it's good enough for our purposes.

Our regression tables will contain several models, each of them containing a different combination of variables we want to investigate.

In our regression table you will also see three types of variables. They are easy to differentiate:

- **Continuous Variables** are placed on their own
- **Single Choice Dummy Variables** are also placed on their own.
- **Multiple-Choice Dummy Variables** are placed in a group with other options for the same question. These options are compared to the [baseline option] indicated in brackets

Treat **Continuous Variable** entries as "the higher the X, the higher/lower the Y."

Treat **Single Choice Dummy Variable** entries as "People who say 'Yes' on X have higher/lower Y than those who say 'No."

Treat **Multiple-Choice Dummy Variable** entries as "People who are X have higher/lower Y than people who are [baseline]." These variables require a [baseline] comparison because they each are divided into a set of dummy variables that are mutually exclusive to one another and together make up of 100% of the observation (e.g. The % of people who hail from Reddit + % of people who hail from Facebook + % of people who hail from elsewhere = 100%), and including all the options would give the model nothing to compare these variables to. Thus, we throw in n-1 options for each question into the regression while constraining the [baseline] to be NULL or 0.

3 Analysis

We will begin by running a hierarchical linear regression, as shown in Table 1.

It is linear because our dependent variable **income** is continuous. This will be the case for every regression table from now on in this paper, so I won't mention this again.

It is hierarchical because we are adding in models sequentially so we can test different combinations of independent variables see their effects when applied together v.s. on their own. This will be the case for every regression table from now on, so I won't mention this again.

	Model1	Model2	Model3	Model4	Model5
(Intercept)	33274.034**	−39759.233**	32311.462**	−28650.336*	−32503.809*
Control Variables [v.s. from elsewhere]					
com.reddit	1705.795	4253.200	−1590.278	1846.532	1818.296
com.facebook	−10210.782	−10922.200	−14813.347	−13599.423	−13609.060
want.baby	−898.628	−1090.508	−5251.063	−3699.223	−3528.240
[v.s. poli centrist]					
poli.auth	2338.275	−117.502	2530.841	392.445	−215.533
poli.lib	6146.766	6247.832	6365.361	6364.214	6292.016
[v.s. economic centrist]					
eco.left	4581.873	3365.508	3663.659	3003.418	3024.785
eco.right	9311.478	8347.133	8870.548	8234.133	8050.311
[v.s. unmanly]					
piss.through	15653.948*	11378.946	17491.416*	13177.664	13286.048
piss.over	17935.214*	13410.015	20228.851**	15525.424*	15727.998*
vote.trump	10094.187	10796.985*	9243.785	10168.760*	10182.240*
Main Effects					
Age		2654.863***		2229.819***	2364.666***
married			25800.614***	15646.196**	35053.450
Age:married					−621.372
AIC	12752.817	12703.518	12726.537	12695.640	12697.190
BIC	12803.932	12758.893	12781.911	12755.274	12761.084
Log Likelihood	−6364.409	−6338.759	−6350.268	−6333.820	−6333.595
Deviance	1137423864687.668	1031155634432.352	1077552567806.370	1011861592511.972	1010992276687.664
Num. obs.	523	523	523	523	523

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 1: Who make more money?

3.1 Age & Marriage

Let's look at the boring variables first: **Age** & **Marriage**. Model 2 and Model 3 show that a Dickhead's age and marital status are both positively correlated with income. Even when they are placed together in Model 4, which means they are controlling for one another, we still see stars in both variables, albeit with a smaller coefficients. This means that age and

marriage status both eat up a bit of each other's effects.

In Model 5 I introduce an interaction term between the **Age** and **Marital Status**. **"Age:married"** is a dummy variable that shows up as "Yes" only when **Married** is "Yes" and **Age** is high; You can read it as "old, married people." Having an interaction term in the mix changes the meaning of **Age** and **Marriage**. Now **Age** means "The effect of age on income when not married." and **Marriage** means "The effect of being married on income when age is low." Model 5 demonstrates that **Marriage** is a free-rider to **Age**'s glory.[3][4] Not only is its effects not significant for when age is low, the interaction term "Age:married" itself is not significant either, meaning that even with high age, being married does not correlate with income.

This is caused by the close correlation between age and being married, as shown in Figure 1.[5]

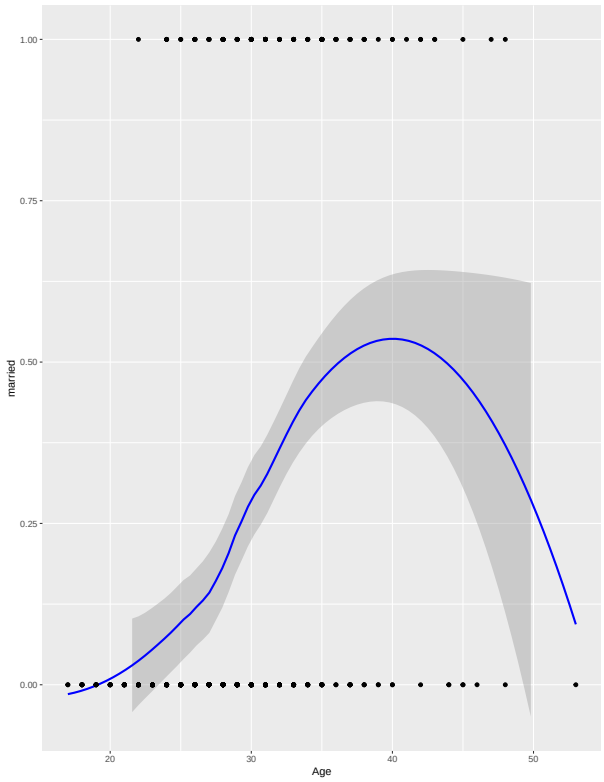

Figure 1: Correlation Between Age And Marital Status

[3] Richard Williams. 2015. "Interaction effects between continuous variables." Course Handout, p1-2

[4] Richard Williams. 2015. "Interaction effects and group comparisons." Course Handout, p13

[5] There are few data points on the right end of the plot, so we look at the general direction of the graph before that part.

3.2 Pissing Style

If you look closely at Table 1, you will see that **pissing over one's waist band** and **pissing through one's zipper hole** both yield more income than **pissing sitting down** or being a woman (same thing really). This effect gets amplified when marriage comes into play, disappears when age comes into play, but when both age and marriage are present, the effect of **Pissing Over** reappears .

What the hell is going on? In Figure 2 and Figure 3, we see that both pissing styles have a cubic correlation with Age, and an almost linear correlation with marital status.[6] **Pissing Over** is able to retain its relevance only because it is more positively correlated with age at the right end of the age spectrum compared to **Pissing Through**. Very disappointing.

[6] The hint of a cubic figure likely is just another dominating contribution of Age. Age is making everyone its cuck.

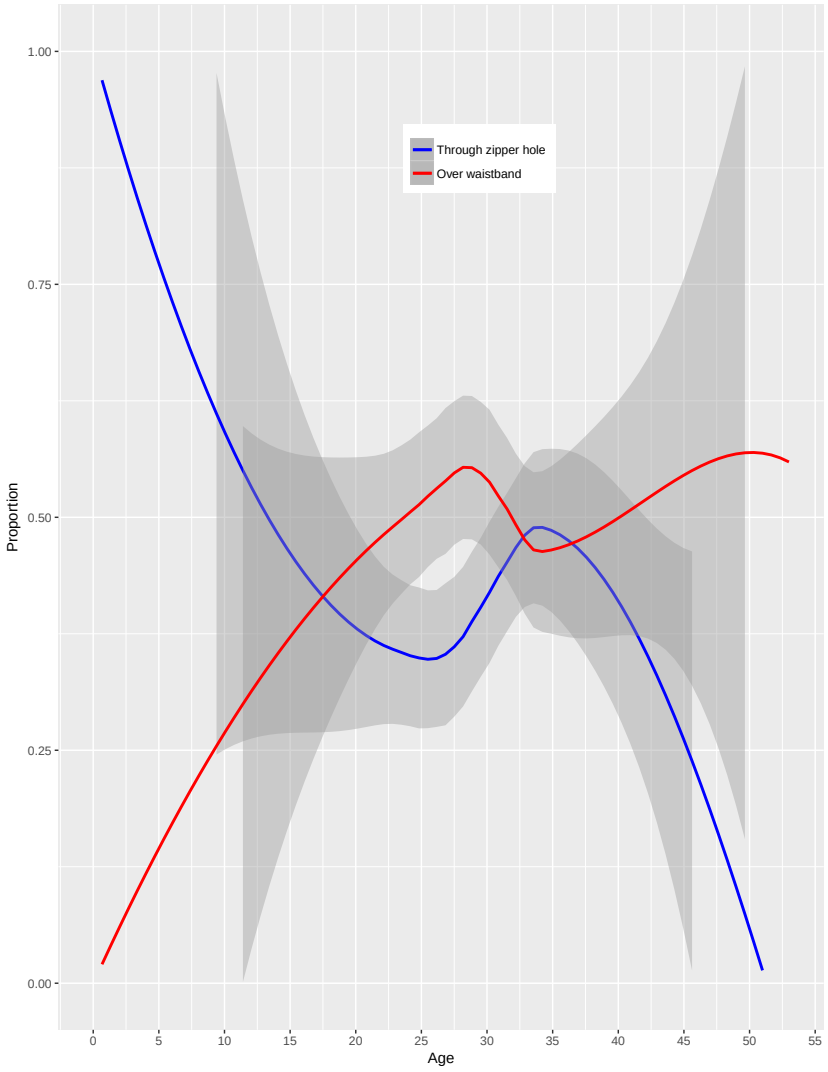

Figure 2: Correlation Between Age And Pissing Styles

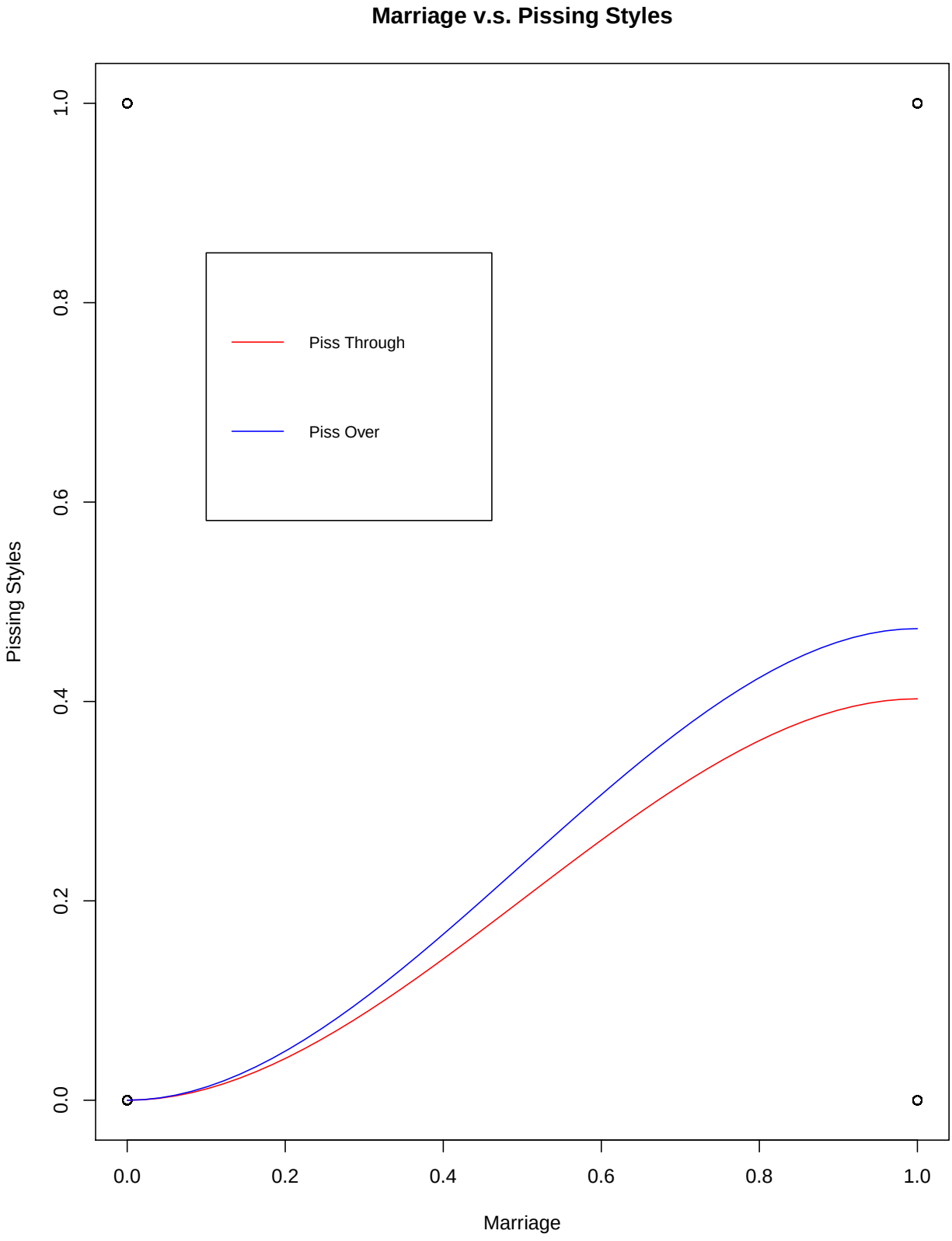

Figure 3: Correlation Between Marital Status And Pissing Styles

But we all know that because of the patriarchy and its toxic definition of masculinity, gay men and women make less money than men. Some feminists even claim a 30% difference between how much a man makes and how much a woman makes, and Dickheads can agree that the only way men can afford hookers is by paying them less than 100% of the money a man makes.

Proof.

$$\int_{Last\ month's\ pay\ day}^{This\ month's\ pay\ day} men\ wage\,dt \ > \int_{Last\ month's\ pay\ day}^{This\ month's\ pay\ day} women\ wage\,dt$$

$$\int_{Last\ month's\ pay\ day}^{This\ month's\ pay\ day} men\ wage\,dt = \sum_{i=0}^{k} \int_{Last\ month's\ pay\ day}^{This\ month's\ pay\ day} man\ wage\,dt \qquad \text{[By Definition]}$$

$$Man's\ money = \int_{Last\ month's\ pay\ day}^{This\ month's\ pay\ day} man\ wage\,dt \qquad \text{[By definition]}$$

$$Man's\ expenses = \int_{Last\ month's\ pay\ day}^{This\ month's\ pay\ day} Man's\ bills\,dt \qquad \text{[By definition]}$$

$$All\ my\ expenses \leq All\ my\ money \qquad \text{[By "No Debt" Theorem]}$$

$$\text{Money for hookers} + \text{Food money} \leq All\ my\ expenses \qquad \text{[By "Don't Starve" Theor}$$

$$\text{Money for hookers} \leq All\ my\ expenses - \text{Food money} \qquad \text{[By Additive Property of Inequalities]}$$

$$\text{Woman} = \text{Hooker} \qquad \text{[By Definition]}$$

$$\text{Money for hookers} = \int_{Last\ month's\ pay\ day}^{This\ month's\ pay\ day} hooker\ wage\,dt \qquad \text{[By Definition]}$$

$$\int_{Last\ month's\ pay\ day}^{This\ month's\ pay\ day} hooker\ wage\,dt = \int_{Last\ month's\ pay\ day}^{This\ month's\ pay\ day} woman\ wage\,dt \qquad \text{[By above]}$$

$$\int_{Last\ month's\ pay\ day}^{This\ month's\ pay\ day} women\ wage\,dt = \sum_{i=0}^{m} \int_{Last\ month's\ pay\ day}^{This\ month's\ pay\ day} woman\ wage\,dt \qquad \text{[By Definition]}$$

$$\int_{Last\ month's\ pay\ day}^{This\ month's\ pay\ day} men\ wage\,dt \ > \int_{Last\ month's\ pay\ day}^{This\ month's\ pay\ day} women\ wage\,dt \qquad \text{by [Transitive Property of Inequalities]}$$

$\therefore$ By the principle of deduction, the claim holds for all k where k = number of men in society and m = number of women in society

■

And there you have it. A mathematical proof that women make less money than men do. Go feminism!

Anyway, what really matters is whether there is an advantage to each manly pissing style. After deleting all the sitter-pissers and women entries, we shall investigate, whether, when one is a small child, would the choice of **Pissing Over** or **Pissing Through** make a difference in one's life-long financial prospects, independent from the effects of age??????????????????????????????????????

	Model1	Model2	Model3	Model4	Model5
(Intercept)	−29474.251*	−29474.251*	−28693.854	−28693.854	−31702.196
Control Variables					
[v.s. from elsewhere]					
com.reddit	2759.486	2759.486	2789.303	2789.303	2821.651
com.facebook	−14666.432	−14666.432	−14463.571	−14463.571	−14482.729
want.baby	−3637.207	−3637.207	−3524.350	−3524.350	−3551.034
[v.s. poli centrist]					
poli.auth	5094.579	5094.579	5291.510	5291.510	5049.948
poli.lib	6509.351	6509.351	6613.392	6613.392	6616.502
[v.s. economic centrist]					
eco.left	1490.378	1490.378	1292.877	1292.877	1173.935
eco.right	8313.672	8313.672	8113.412	8113.412	7957.206
vote.trump	11121.200*	11121.200*	11225.505*	11225.505*	11147.924*
Main Effects					
Age	2871.842***	2871.842***	2869.330***	2869.330***	2977.535***
piss.through			−1926.453	−1926.453	5880.973
Interaction					
Age:piss.through					−269.126
AIC	11556.419	11556.419	11558.214	11558.214	11560.105
BIC	11602.216	11602.216	11608.174	11608.174	11614.229
Log Likelihood	−5767.210	−5767.210	−5767.107	−5767.107	−5767.053
Deviance	977658133225.291	977658133225.291	977236346135.568	977236346135.568	977012225699.890
Num. obs.	475	475	475	475	475

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 2: What kind of pissers make more money?

No.

3.3 Voting Trump

After learning that it is futile to look down at our crotch, it is time we change our perspective and look upwards, at our glorious leader. President **Donald Trump**. Will the Donald save this paper from being completely uninformative?

	Model1	Model2	Model3	Model4	Model5	Model6
(Intercept)	33792.835**	−38765.966**	33274.034**	−39759.233**	−17678.095	10465.909
Control Variables [v.s. from elsewhere]						
com.reddit	2265.925	4836.797	1705.795	4253.200	6464.135	9085.256
com.facebook	−9425.390	−10078.183	−10210.782	−10922.200	−9255.682	−5985.848
want.baby	−164.715	−304.651	−898.628	−1090.508	−422.741	214.961
[v.s. poli centrist]						
poli.auth	3882.835	1548.697	2338.275	−117.502	−1710.246	−100084.086*
poli.lib	6462.753	6585.080	6146.766	6247.832	8136.338*	−29275.701
[v.s. economic centrist]						
eco.left	3464.043	2177.622	4581.873	3365.508	2515.681	1672.772
eco.right	11176.834*	10347.374*	9311.478	8347.133	7041.631	6909.075
[v.s. unmanly]						
piss.through	15758.670*	11516.586	15653.948*	11378.946	12360.396	11629.227
piss.over	17689.642*	13174.622	17935.214*	13410.015	13246.739	12127.055
Main Effects						
Age		2638.920***		2654.863***	1793.270***	748.755
vote.trump			10094.187	10796.985*	−102555.853***	−106661.451***
Interactions						
Age:vote.trump					3934.866***	4080.635***
poli.auth:Age						3336.298*
poli.lib:Age						1293.277
AIC	12754.744	12706.467	12752.817	12703.518	12684.931	12682.519
BIC	12801.599	12757.582	12803.932	12758.893	12744.565	12750.672
Log Likelihood	−6366.372	−6341.234	−6364.409	−6338.759	−6328.465	−6325.259
Deviance	1145996068627.122	1040959222417.694	1137423864687.668	1031155634432.352	991352839828.441	979273701970.575
Num. obs.	523	523	523	523	523	523

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 4: MAGA MAKE MONEY?

The answer is yes.

Take a look at Model 3 and Model 4. **Voting for Trump** has no statistically significant correlation with income unless **Age** comes into play. How the fuck does that work?

We will know if we look at Model 5. The interaction shows that being old and **voting for Trump** is a winning combination. Furthermore, it shows that there is no real correlation between the effects of **Voting for Trump** and the effects of **Age**; This checks out, as shown in Figure 4.

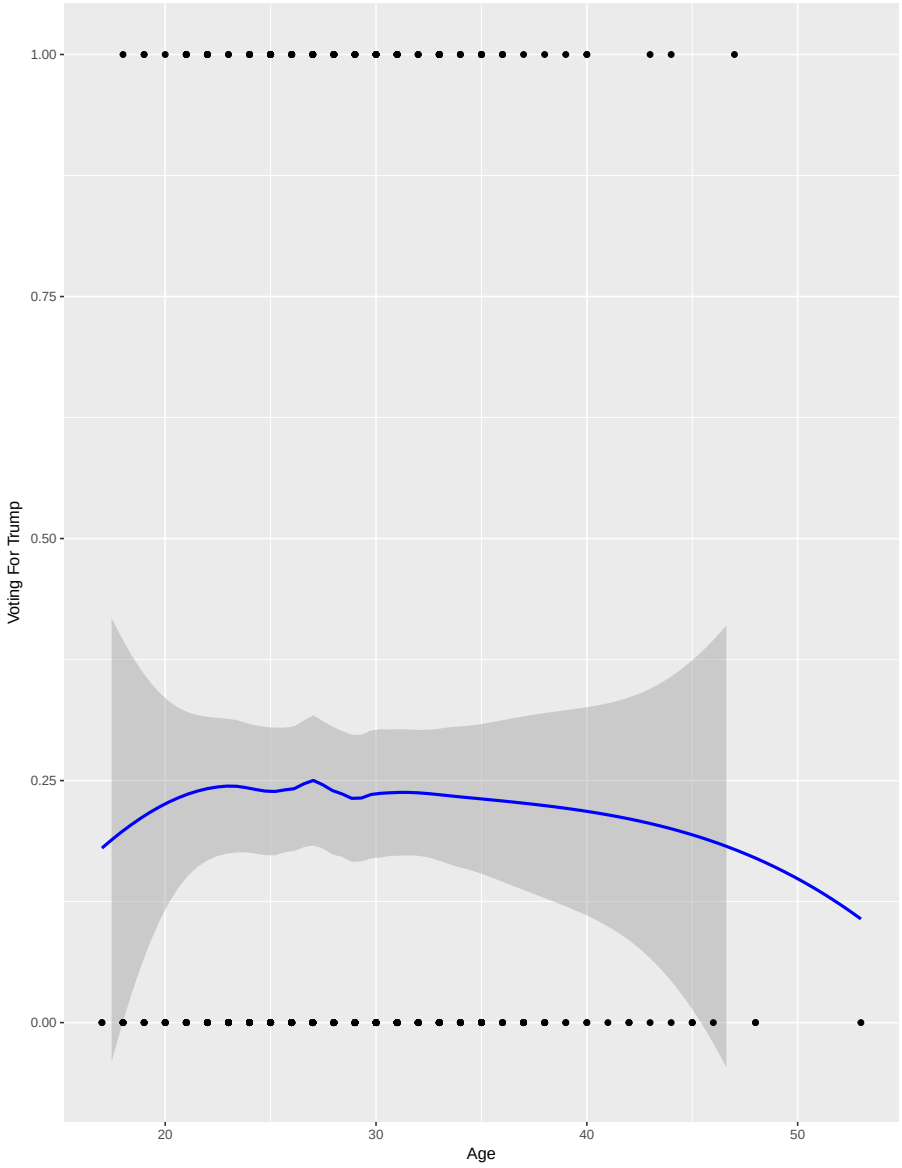

Figure 4: Correlation Between Age and Voting For Trump

But what's the deal with that giant negative number in Model 5? With an interaction term in the same model, the **Voting For Trump** coefficient is saying: "Being a young trump voters correlates with making an ass load of jack shit." Another note-worthy thing is that introducing the interaction term diminishes the effect of **Age**, meaning that a part of Age's correlation with income comes from the power of **Voting For Trump**.

Because we have three directions, there are three ways to spin this result. Here are our options:

y = **Voting For Trump**	People who vote for Trump were either young people trying to make more money or old people who already make tons.
y = **Age**	If you are poor and you vote for Trump, you are going to stay young forever; If you are rich and you vote for Trump, you are going to have accelerated aging.
y = **Income**	If you are young, voting for Trump will make you poor; If you are old, voting for Trump will make you rich.

Take your pick. Since the purpose of this paper is how to make more money, I'm gonna go with the third option. That way, when I get older, all I have to do is vote for Trump to get rich. Fucking $800 per month man.....

But wait, what about Model 6? Why is that in there? Well you see. Model 5 clutches yet another scrotum pearl close to its bosom, for it signals that being **Political Liberals** correlates with making more money. We can't have that. We can't end on a note that implies Asterios wins. Asterios needs to lose. So I constructed from this stinking pile of garbage a new model that will turn your entire world around. Behold...

Going down the list, compared to Model 5, by adding in the interaction terms between **Political Leanings** and **Age**, Model 6 demonstrates that:

1. The positive effects of being a **Liberal** disappears
2. Being a **young Authoritarian** correlates with having less income
3. The effect of **Age** disappears
4. The effect of **Voting for Trump** at low age remains the same
5. The effect of being **old** and **Voting For Trump** remains the same
6. Being an **old Authoritarian** correlates with having more income

And so we come to it finally. It is not the age of the trump voter that matters. It is ones political leanings that dominates one's destiny. Either that, or one's financial successes destines one to become a **Political Authoritarian**. This checks out as shown in Figure 5.

Conclusion: You need to **vote for Trump when you are** an **old Authoritarian** in order to **get rich**.

Everything else gets you fucked. Don't vote for Trump when you're young. Don't vote for Trump when you're a young liberal. Don't be a young liberal. Don't be a liberal, period.[7]

Get fucked Asterios.

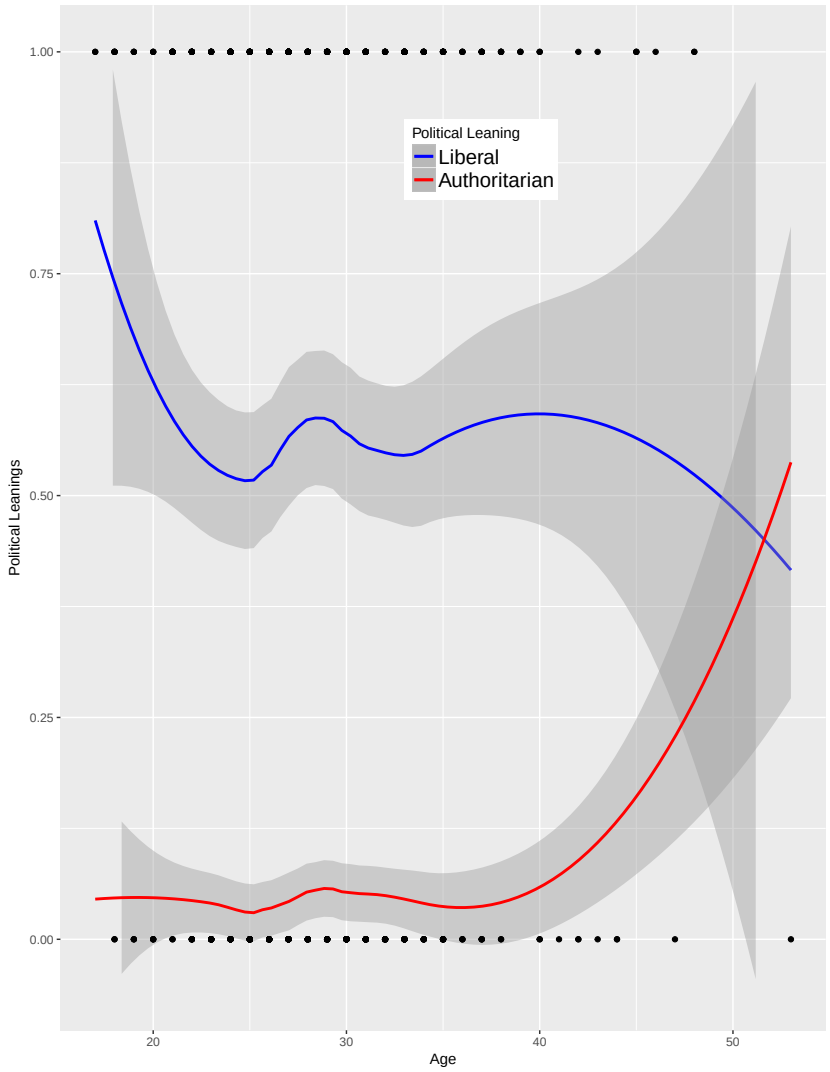

Figure 5: Correlation Between Age and Political Leanings

[7] Not actually supported by data. Model 6 actually is saying that it's more financially advantageous to be an **Authoritarian** when you are **old**, and more advantageous to be a **Political Centrist** if you are **young**. You can't infer from it anything about the effects of being liberal. But if I say that I don't get to attack Asterios. So there.