Testing the Effect of Marital Status On Weight

Zachary Safir, PUBHLTH 490ST, Spring 2018

March 7, 2018

My Story and Its Importance:

For this project I wanted to explore the relationship between a person's weight, and their marital status. It is important for everyone to know, as all of us here will eventually either be in relationship of some sort, or have never had one at all. As such this applies to everyone.

The Data I Used

I used data from the National Health Interview Survey 2015, which is a massive set of data that holds the responses of many adults who took a health survey. I used the variables AWEIGHTP, R_MARITL, and SEX. Representing, in the order listed, weight (in pounds), marital status, and sex. I also created in this analysis, a new variable, Average Weight, which represents the average weights for each marital status.

Cleaning of the Data:

In cleaning the data, I got rid of all unnecessary, and NA variables by simply making subsets where I only included people with a weight less than 300. I used the number 300, as that was the max weight recorded. For anyone who refused, information was unavailable, not ascertained, or simply did not know, they were recorded with a weight ranging from 996-999. As this was not applicable, and would only skew my results unnecessarily, I simply chose not to include them. From there I cleaned it further by filtering out anyone who did not give a proper response to the marital question. I chose to simply focus on the known statuses here to prevent making an over complicated analysis. In total I created created three data sets, one including people of both sexes, and two others split up by gender.

Everyones Weight vs Marital Status

Marital Status	Average Weight	Count	SD
Divorced	177.00	1280	38.99
Living With Partner	177.31	572	39.55
Married-in household	177.62	3888	39.14
Married-not in household	173.52	137	37.08
Never Married	174.30	2036	40.67
Seperated	178.18	266	40.19
Widowed	162.85	857	37.06

Womens Weight vs Marital Status

Marital Status	Average Weight	Count	SD
Divorced	164.27	766	36.91
Living With Partner	160.93	275	36.49
Married-in household	159.82	1988	34.14
Married-not in household	162.69	74	35.60
Never Married	161.90	1031	39.19
Seperated	169.28	166	38.08
Widowed	155.24	661	33.78

Mens Weight vs Marital Status

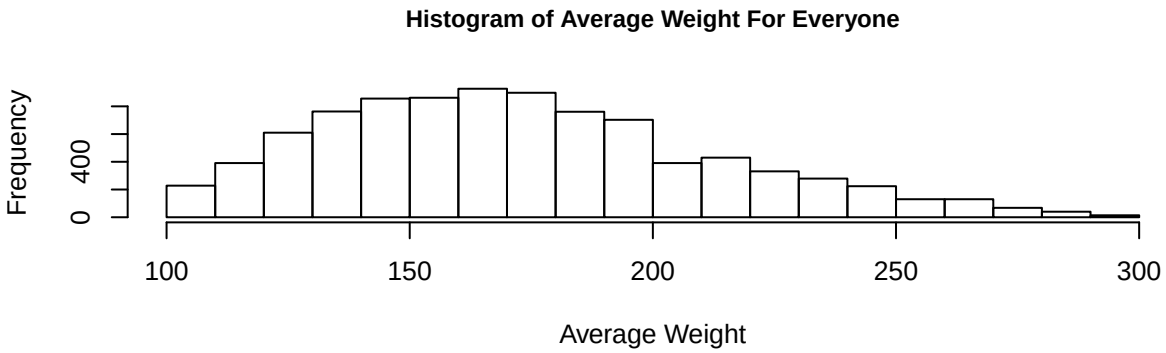
Marital Status	Average Weight	Count	SD
Divorced	195.97	514	34.01
Living With Partner	192.47	297	36.12
Married-in household	196.24	1900	35.17
Married-not in household	186.24	63	34.92
Never Married	187.02	1005	38.18
Seperated	192.96	100	39.41
Widowed	188.55	196	36.14

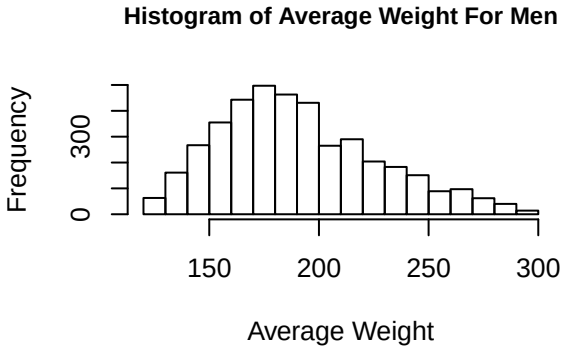
Description of My Key Summary Statistics

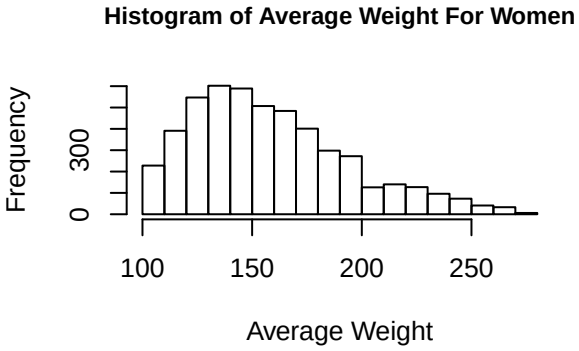
Starting from the first table, which has summary statistics for both sexes, we see that the marital status with the highest overall average is separated. The status with the lowest average is widowed, which seems to be significantly less than the others, being 10 pounds lower than the second lowest status. Looking to the counts, it seems the marriage status with the most people is married in household, with the lowest being married but not in household. The standard deviations for each seems to fluctuate between 37 and 40. Now onto the differences between both sexes. As would make sense, the first immediately obvious fact is that the average weight for men is higher than women. Looking to where each group has the highest and lowest average weight, we again see differences. For women we see the highest average weight when the person is currently separated. For men however, the highest weight seems to be when they are married and living in the same household. For the lowest weight we see for women, that they at their lowest when widowed, and for men it when they are married, but not living in the same household. We can also see some big differences in count. For women, there is a huge number of widowed people. Men on the other hand, don't even make up 20 percent of the total widowed people recorded in this survey. Everything else seems to split up almost evenly with some bigger differences here and there, but nothing nearly as big as the gap in widowed status. And finally, for the standard deviation, it's hard to see if there is any big differences for each, as they both seem to both have similar ranges and both don't vary too much.

```
##
##  Shapiro-Wilk normality test
##
## data:  m.weight$AWEIGHTP
## W = 0.9691, p-value < 2.2e-16
```

```
##
##  Shapiro-Wilk normality test
##
## data:  f.weight$AWEIGHTP
## W = 0.95369, p-value < 2.2e-16
```







Testing Normality

As you can see, in this section I first ran the Shapiro Wilks normality test on the men and women data sets I created. The p value for both tests is highly significant indicating a normal distribution, but as these are both big data sets, the test becomes less reliable, and was completely unable to deal with the data set including the weights of both sexes as it contained too many variables. As the histograms show, while the shapes for both seem to be somewhat normal, they appear to be skewed to the right. Of course, in a set as big as this we don't need to worry so much about normality as per the central limit theorem. As such, in the next part I still used the parametric ANOVA test.

Anova Table for Everyones Weght vs Marital Status

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	93602015.488	1	59954.951461	0.0000000
R_MARITL	9082.149	1	5.817394	0.0158882
Residuals	14103932.825	9034	NA	NA

Anova Table for Womens Weght vs Marital Status

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	39228665.144	1	30262.285368	0.0000000
R_MARITL	6404.811	1	4.940882	0.0262736
Residuals	6428296.743	4959	NA	NA

Anova Table for Mens Weght vs Marital Status

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	56085783.81	1	43055.18239	0e+00
R_MARITL	37824.68	1	29.03674	1e-07
Residuals	5305688.76	4073	NA	NA

My ANOVA Tables

The ANOVA tests shown above reveals some interesting findings. All three p values for the ANOVA tables are below .05, indicating a significant difference in one or more of the levels for each data set. The p value for the mens data set though, seems to be much more significant then then the womens. The p value for everyone's weight is, of course a number in between both.

	diff	lwr	upr	p.adj
Living With Partner-Divorced	-3.4955456	-11.238666	4.2475746	0.8371551
Married-in household-Divorced	0.2751792	-5.006544	5.5569023	0.9999989
Married-not in household-Divorced	-9.7288308	-23.909654	4.4519920	0.3996212
Never Married-Divorced	-8.9450355	-14.705794	-3.1842772	0.0000975
Seperated-Divorced	-3.0069261	-14.617882	8.6040301	0.9882491
Widowed-Divorced	-7.4210077	-16.339366	1.4973510	0.1764827
Married-in household-Living With Partner	3.7707248	-2.857929	10.3993790	0.6309022
Married-not in household-Living With Partner	-6.2332852	-20.968895	8.5023245	0.8752544
Never Married-Living With Partner	-5.4494899	-12.465813	1.5668334	0.2484288
Seperated-Living With Partner	0.4886195	-11.793756	12.7709955	0.9999998
Widowed-Living With Partner	-3.9254621	-13.701943	5.8510188	0.9002269
Married-not in household-Married-in household	-10.0040100	-23.608380	3.6003599	0.3126346
Never Married-Married-in household	-9.2202147	-13.363822	-5.0766074	0.0000000
Seperated-Married-in household	-3.2821053	-14.181529	7.6173190	0.9743994
Widowed-Married-in household	-7.6961869	-15.666147	0.2737735	0.0664629
Never Married-Married-not in household	0.7837953	-13.013618	14.5812088	0.9999981
Seperated-Married-not in household	6.7219048	-10.366003	23.8098123	0.9088216
Widowed-Married-not in household	2.3078231	-13.077865	17.6935115	0.9994342
Seperated-Never Married	5.9381095	-5.201333	17.0775522	0.7000914
Widowed-Never Married	1.5240278	-6.771152	9.8192075	0.9982045
Widowed-Seperated	-4.4140816	-17.469271	8.6411078	0.9547067

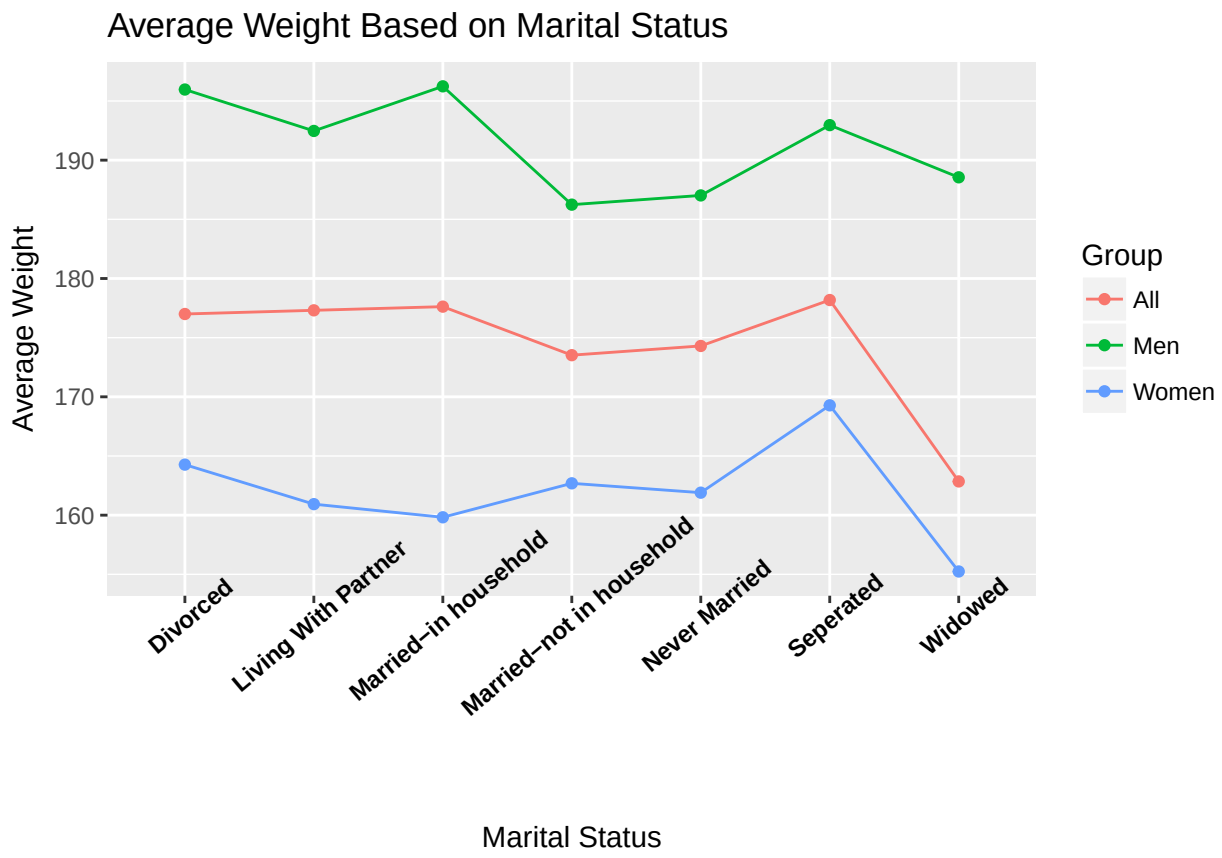
Tukey Test For Mens Weight vs Marital Status

The tukey test above tests the differences between the levels in the data you feed it. For the men data set we see that there are multiple levels that have a significant difference between them. The biggest difference for men seems to exist between married but not in same household, and married in same household. And the smallest difference seems to be between being married in the same household,and being divorced.

	diff	lwr	upr	p.adj
Living With Partner-Divorced	-3.3396060	-10.7862114	4.1069995	0.8414554
Married-in household-Divorced	-4.4582561	-8.9630311	0.0465189	0.0545436
Married-not in household-Divorced	-1.5849622	-14.4800146	11.3100901	0.9998217
Never Married-Divorced	-2.3740544	-7.4269932	2.6788843	0.8096548
Seperated-Divorced	5.0029570	-4.0659143	14.0718283	0.6646801
Widowed-Divorced	-9.0381454	-14.6616957	-3.4145951	0.0000450
Married-in household-Living With Partner	-1.1186501	-7.9338947	5.6965946	0.9990528
Married-not in household-Living With Partner	1.7546437	-12.1175203	15.6268078	0.9997893
Never Married-Living With Partner	0.9655515	-6.2237983	8.1549014	0.9997014
Seperated-Living With Partner	8.3425630	-2.0689205	18.7540464	0.2145016
Widowed-Living With Partner	-5.6985394	-13.2997814	1.9027026	0.2894575
Married-not in household-Married-in household	2.8732938	-9.6677588	15.4143464	0.9939083
Never Married-Married-in household	2.0842016	-1.9812401	6.1496434	0.7376704
Seperated-Married-in household	9.4612131	0.9031728	18.0192533	0.0192242
Widowed-Married-in household	-4.5798893	-9.3359323	0.1761536	0.0679255
Never Married-Married-not in household	-0.7890922	-13.5373153	11.9591309	0.9999969
Seperated-Married-not in household	6.5879192	-8.2184790	21.3943175	0.8464267
Widowed-Married-not in household	-7.4531831	-20.4381483	5.5317821	0.6208192
Seperated-Never Married	7.3770114	-1.4818398	16.2358627	0.1758021
Widowed-Never Married	-6.6640909	-11.9422664	-1.3859154	0.0037384
Widowed-Seperated	-14.0411024	-23.2373719	-4.8448329	0.0001390

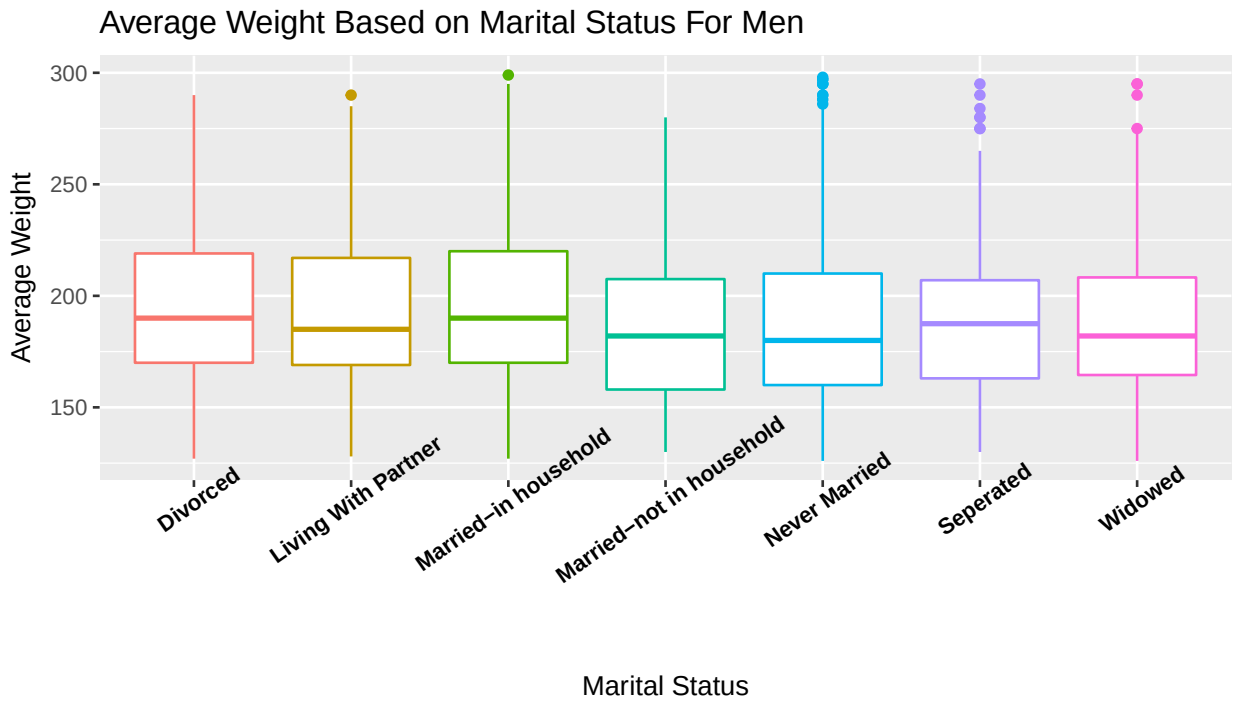
Tukey Test For Womens Weight vs Marital Status

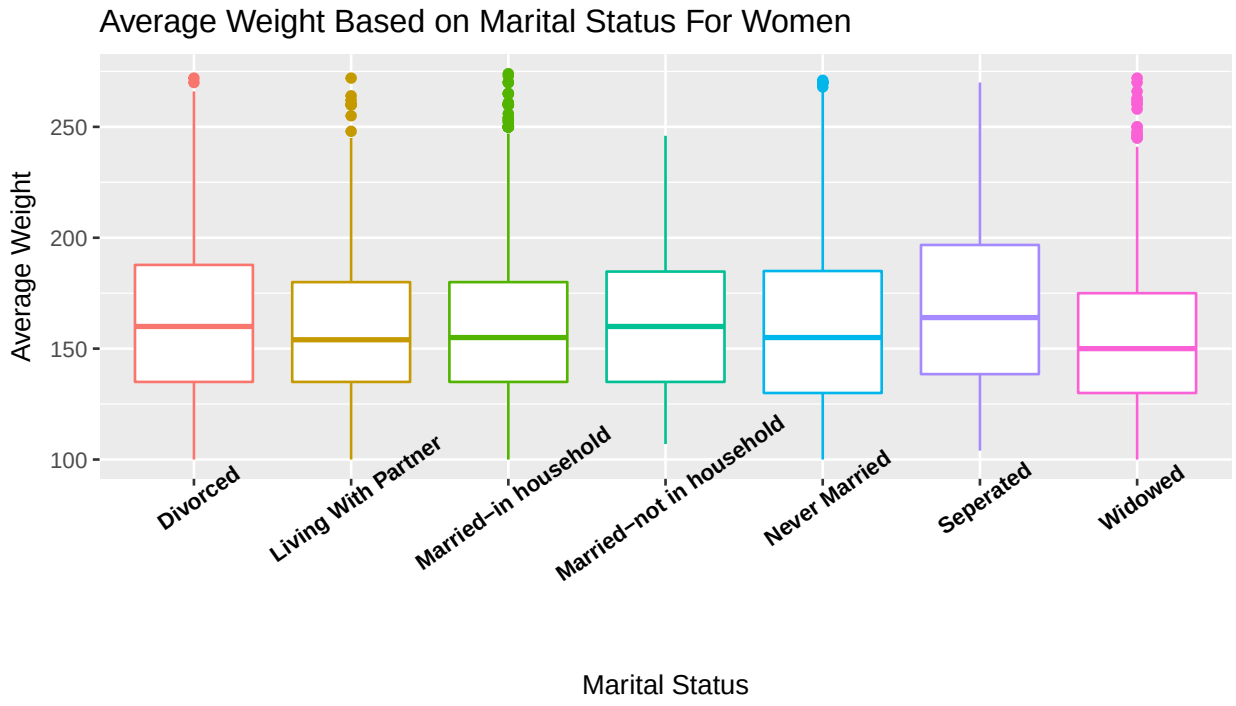
For the women data set we see similar results, there are many levels here that are significantly different. The biggest difference for women however, exists between widowed and seperated. The smallest difference here seems to exists between never married, and married but not in same household.



First Graphic

To start with I created a simple line graph to clearly display the different means for each data set. It's clear from looking at this graphic that weight does vary based on what marital status the respondent currently is in. We can also clearly see above the marital groups with the highest average weight for each sex, and how women with a widowed status are in the majority, as the data set with everyone weights also drops significantly to reflect the women's average weight there. This graphic does however fail to convey the distribution of everyone's weights.


Average Weight Based on Marital Status For Men
Average Weight
300
250
200
150
Divorced
Living With Partner
Married−in household
Married−not in household
Never Married
Seperated
Widowed
Marital Status


Average Weight Based on Marital Status For Women
Average Weight
250
200
150
100
Divorced
Living With Partner
Married−in household
Married−not in household
Never Married
Seperated
Widowed
Marital Status

Second Graphics

To properly finish this off, I created two bar plots, one for each gender, to clearly show the distribution of weight in each marital status. The graphs above indeed reveal some interesting information that would have been otherwise left unknown. Looking at the boxes for both graphics, I noticed that for men, the distribution in divorced, living with partner, married in household, and never married all seem to be slightly right skewed. The groups in the women data set though, seem to be more evenly distributed, with only separated showing some more significant skewing. What does this mean? When the data is right skewed it usually indicates that the mean of the data is not in the middle. In this case, that means there are people with abnormally high weights skewing these distribtuons. This is more commonly seen in men, according to this data at least.

Conclusions

I set out to find if marital status had any effect on a person's weight, and after doing everything here I feel like I have come to a solid conclusion. If, like I did, you were to ignore all factors outside of gender and marital status, I can soundly say there is a noticeable, significant difference in weights for people of different marital statuses. For both men and women, it seems apparent that your weight can be affected by your current marital status. Now, like I have mentioned many times before in this write up, I am fully aware that the reality is that there are way more factors at play than what I accounted for here. Until I take into complete account other such factors as age, and ethnicity I can't say for sure if marital status is what causes the different weights found in my analysis. However, I can definitively conclude that the differences exist. The p values from the ANOVA test and the turkey tests values for difference, are all statistical proof for this conclusion. In future investigations it will be interesting to investigate how exactly we can break this down based on the other factors I mentioned, and hopefully become able to find a proper cause for the different weights found in my data. For now, this is everything that can be said about the work I have done here.