Name
Professor
Course
Date

Sleep vs. Coffee

Introduction

When considering possible choices for comparable variables, I was struck by the amount of coffee people tend to drink on a regular basis. I have also become curious as to whether men or women tend to drink more coffee each day. Finally, I was also curious about the cause of that coffee consumption. Because of this I also want to know if the numbers of hours of sleep that a person gets per night has an effect on their coffee consumption (Gangwisch et al. 2007).

Thus, I initially want to compare the hours of sleep each gender gets per night to their average coffee consumption per day. This data will all be discrete as each person will be asked to round their sleep to the nearest hour, and coffee consumption to the whole cup. In this way, the independent variable will be the average hours of sleep per night, and the dependent variable will be the average cups of coffee consumed per day (Salazar-Martinez et al. 2004).

Hypothesis

Research Question:

Do the hours of sleep a person gets per night affect the amount of coffee they drinks per day?

The following null and alternative hypotheses are consistent for both groups (men and women):

H_0: There is no correlation between the hours of sleep and a person's average coffee consumption per night.

H_1: There is a correlation between the hours of sleep and a person's average coffee consumption per night.

Based on personal experience, I think that there will be a strong correlation between the hours of sleep a person gets per night and their coffee consumption per day. That being said, I would also think that women consume more cups of coffee per day than men do, while at the same time men most likely get less sleep per night than women do.

Methods

I distributed a survey to 30 of my classmates so that they might ask their parents how much sleep they get per night as well as how much coffee they drink per day. This survey consisted of only two questions:

1. On average How many hours of sleep do you get per night?
2. On average how many cups of coffee do you drink per day?

Returned to me was data on 30 men and 30 women.

Although it is extremely difficult to not produce a biased data sample, the best method to use is Simple Random Sampling. However, as that is not possible, considering that would consist of a gaining a sample of every person within the entire state, country, or world, this study will be limited to the population of parents of students attending this school. Within this limited range a random sampling of students was taken by handing out my survey to the first 30 students who would take it as I passed them in the hall way.

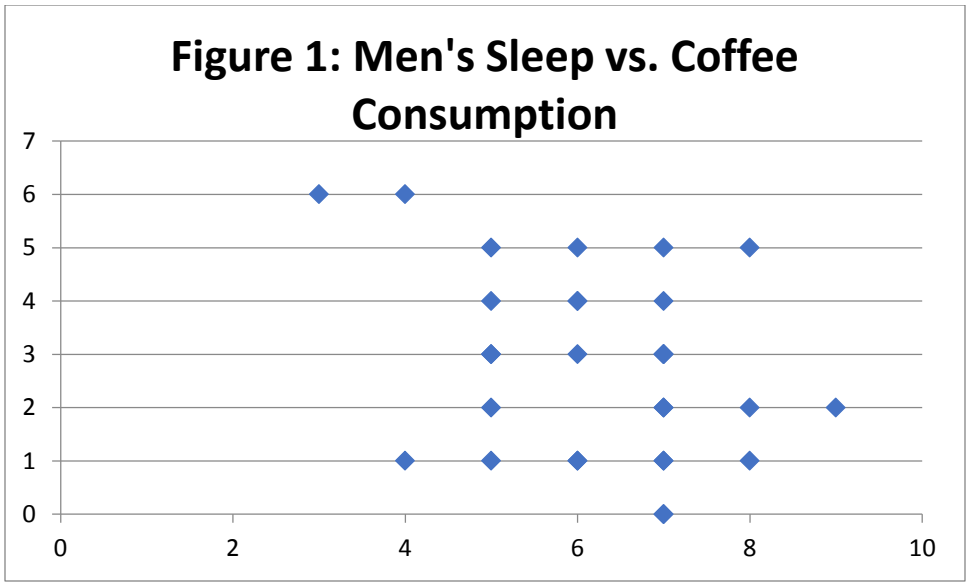
Data

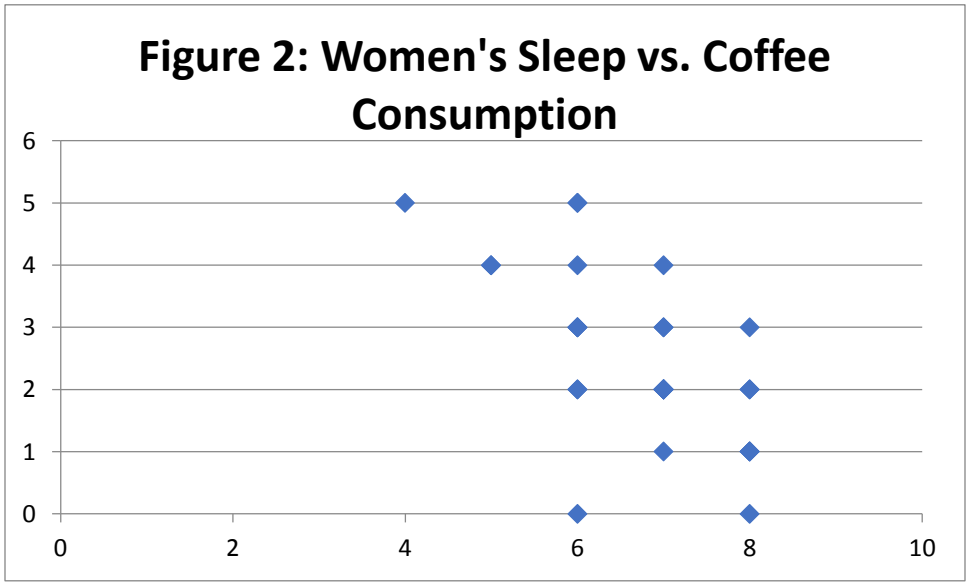
The compiled data for every man's average hours of sleep and average cups of coffee consumed per day is shown in Table 1 below. The same data is shown for every woman who was part of this study is shown below in Table 2.

Table 1: Men		
subject #	hours of sleep	coffee consumption
1	6	1
2	5	5
3	7	3
4	7	4
5	8	2
6	6	1
7	4	1
8	7	5
9	7	0
10	7	1
11	6	3
12	8	5
13	5	1
14	5	2
15	5	3
16	5	3
17	7	3
18	6	4
19	7	2
20	7	3
21	7	0
22	9	2
23	4	6
24	3	6
25	6	4
26	6	5
27	7	2
28	5	4
29	8	1
30	7	1

Table 2: Women		
subject #	hours of sleep	coffee consumption
1	8	3
2	4	5
3	6	5
4	7	2
5	8	2
6	8	1
7	7	4
8	8	0
9	6	0
10	6	2
11	7	2
12	7	1
13	6	5
14	6	4
15	6	3
16	5	4
17	7	2
18	8	1
19	7	3
20	6	2
21	7	2
22	7	2
23	5	4
24	6	3
25	6	3
26	7	2
27	7	2
28	7	3
29	8	2
30	8	1

Figure 1 and 2 show the previous data displayed in their respective scatter plots.





Results

Men

Mean:

$$\bar{x} = \frac{\sum x}{n}$$

$$\bar{x} = \frac{6+5+7+7+8+6+4+7+7+7+6+8+5+5+5+5+7+6+7+7+7+9+4+3+6+6+7+5+8+7}{30}$$

$$\bar{x} = 6.23$$

Square Mean:

$$\overline{x^2} = \frac{\sum x^2}{n}$$

$$\overline{x^2} = \frac{6^2 + 5^2 + 7^2 + 7^2 + 8^2 + 6^2 + 4^2 + 7^2 + 7^2 + 7^2 + 6^2 + 8^2 + 5^2 + 5^2 + 5^2 + 5^2 + 7^2 + 6^2 + 7^2 + 7^2 + 7^2 + 9^2 + 4^2 + 3^2 + 6^2 + 6^2 + 7^2 + 5^2 + 8^2 + 7^2}{30}$$

$$\overline{x^2} = 40.63$$

Standard Deviation:

$$S_x = \sqrt{\overline{x^2} - \bar{x}^2}$$

$$S_x = \sqrt{40.63 - 6.23^2}$$

$$S_x = 1.3338$$

Mean:

$$\bar{y} = \frac{\sum y}{n}$$

$$\bar{y} = \frac{1 + 5 + 3 + 4 + 2 + 1 + 1 + 5 + 0 + 1 + 3 + 5 + 1 + 2 + 3 + 3 + 3 + 4 + 2 + 3 + 0 + 2 + 6 + 6 + 4 + 5 + 2 + 4 + 1 + 1}{30}$$

$$\bar{y} = 2.77$$

Square Mean:

$$\overline{y^2} = \frac{\sum y^2}{n}$$

$$\bar{y} = \frac{1 + 25 + 9 + 16 + 4 + 1 + 1 + 25 + 0 + 1 + 9 + 25 + 1 + 4 + 9 + 9 + 9 + 16 + 4 + 9 + 0 + 4 + 36 + 36 + 16 + 25 + 4 + 16 + 1 + 1}{30}$$

$$\bar{y} = 10.57$$

Standard Deviation:

$$S_y = \sqrt{\overline{y^2} - \bar{y}^2}$$

$$S_y = \sqrt{10.57 - (2.77)^2}$$

$$S_y = 1.7065$$

Covariance:

$$\overline{xy} = \frac{\sum xy}{n}$$

$$\overline{xy} = \frac{6 + 25 + 21 + 28 + 16 + 6 + 4 + 35 + 0 + 7 + 18 + 40 + 5 + 10 + 15 + 15 + 21 + 24 + 14 + 21 + 0 + 18 + 24 + 18 + 24 + 30 + 14 + 20 + 8 + 7}{30}$$

$$\overline{xy} = 16.47$$

$$S_{xy} = \overline{xy} - \bar{x} \cdot \bar{y}$$

$$S_{xy} = 16.47 + 6.23 \cdot 2.77$$
$$S_{xy} = -0.7789$$

Coefficient of Correlation:

$$r = \frac{S_{xy}}{S_x \cdot S_y}$$
$$r = \frac{-0.7789}{1.3338 \cdot 1.7065}$$
$$r = -0.3422$$

Slope:

$$m = \frac{s_{xy}}{{s_x}^2}$$
$$m = \frac{-0.7789}{1.3338^2}$$
$$m = -0.4379$$

Constant:

$$c = \bar{y} - m\bar{x}$$
$$c = 2.7667 - (-0.4379)6.2333$$
$$c = 5.4959$$

Women

Mean:

$$\bar{x} = \frac{\sum x}{n}$$

$$\bar{x} = \frac{8+4+6+7+8+8+7+8+6+6+7+7+6+6+6+5+7+8+7+6+7+7+5+6+6+7+7+7+8+8}{30}$$

$$\bar{x} = 6.70$$

Square Mean:

$$\overline{x^2} = \frac{\sum x^2}{n}$$

$$\overline{x^2} = \frac{64^2+16^2+36^2+49^2+64^2+64^2+49^2+64^2+36^2+36^2+49^2+49^2+36^2+36^2+36^2+25^2+49^2+64^2+49^2+36^2+49^2+49^2+25^2+36^2+36^2+49^2+49^2+49^2+64^2+64^2}{30}$$

$$\overline{x^2} = 45.90$$

Standard Deviation:

$$S_x = \sqrt{\overline{x^2} - \bar{x}^2}$$

$$S_x = \sqrt{45.90 - 6.70^2}$$

$$S_x = 1.0050$$

Mean:

$$\bar{y} = \frac{\sum y}{n}$$

$$\bar{y} = \frac{3+5+5+2+2+1+4+0+0+2+2+1+5+4+3+4+2+1+3+2+2+2+4+3+3+2+2+3+2+1}{30}$$

$$\bar{y} = 2.50$$

Square Mean:

$$\overline{y^2} = \frac{\sum y^2}{n}$$

$$\bar{y} = \frac{9+25+25+4+4+1+16+0+0+4+4+1+25+16+9+16+4+1+9+4+4+4+16+9+9+4+4+9+4+1}{30}$$

$$\bar{y} = 8.03$$

Standard Deviation:

$$S_y = \sqrt{\overline{y^2} - \bar{y}^2}$$

$$S_y = \sqrt{8.03 - (2.50)^2}$$

$$S_y = 1.3354$$

Covariance:

$$\overline{xy} = \frac{\sum xy}{n}$$

$$\overline{xy} = \frac{24+20+30+14+16+8+28+0+0+12+14+7+30+24+18+20+14+8+21+12+14+14+20+18+18+14+14+21+16+8}{30}$$

$$\overline{xy} = 15.90$$

$$S_{xy} = \overline{xy} - \bar{x} \cdot \bar{y}$$

$$S_{xy} = 15.90 + 6.70 \cdot 8.03$$

$$S_{xy} = -0.8500$$

Coefficient of Correlation:

$$r = \frac{S_{xy}}{S_x \cdot S_y}$$

$$r = \frac{-0.8500}{1.0050 \cdot 1.3354}$$

$$r = -0.6333$$

Slope:

$$m = \frac{s_{xy}}{{s_x}^2}$$

$$m = \frac{-0.8500}{1.0050^2}$$

$$m = -0.8416$$

Constant:

$$c = \bar{y} - m\bar{x}$$

$$c = 2.50 - (-0.8500)6.70$$
$$c = 8.1386$$

The three following Tables (3, 4, 5) confirm the previous calculations shown above.

Table 3: Men					
	independent variable		dependent variable		
subject #	hours of sleep	sleep^2	coffee consumption	coffee^2	sleep*coffee
1	6	36	1	1	6
2	5	25	5	25	25
3	7	49	3	9	21
4	7	49	4	16	28
5	8	64	2	4	16
6	6	36	1	1	6
7	4	16	1	1	4
8	7	49	5	25	35
9	7	49	0	0	0
10	7	49	1	1	7
11	6	36	3	9	18
12	8	64	5	25	40
13	5	25	1	1	5
14	5	25	2	4	10
15	5	25	3	9	15
16	5	25	3	9	15
17	7	49	3	9	21
18	6	36	4	16	24
19	7	49	2	4	14
20	7	49	3	9	21
21	7	49	0	0	0
22	9	81	2	4	18
23	4	16	6	36	24
24	3	9	6	36	18
25	6	36	4	16	24
26	6	36	5	25	30
27	7	49	2	4	14
28	5	25	4	16	20
29	8	64	1	1	8
30	7	49	1	1	7

Table 4: Women			
	independent variable	dependent variable	

subject #	hours of sleep	sleep^2	coffee consumption	coffee^2	sleep*coffee
1	8	64	3	9	24
2	4	16	5	25	20
3	6	36	5	25	30
4	7	49	2	4	14
5	8	64	2	4	16
6	8	64	1	1	8
7	7	49	4	16	28
8	8	64	0	0	0
9	6	36	0	0	0
10	6	36	2	4	12
11	7	49	2	4	14
12	7	49	1	1	7
13	6	36	5	25	30
14	6	36	4	16	24
15	6	36	3	9	18
16	5	25	4	16	20
17	7	49	2	4	14
18	8	64	1	1	8
19	7	49	3	9	21
20	6	36	2	4	12
21	7	49	2	4	14
22	7	49	2	4	14
23	5	25	4	16	20
24	6	36	3	9	18
25	6	36	3	9	18
26	7	49	2	4	14
27	7	49	2	4	14
28	7	49	3	9	21
29	8	64	2	4	16
30	8	64	1	1	8

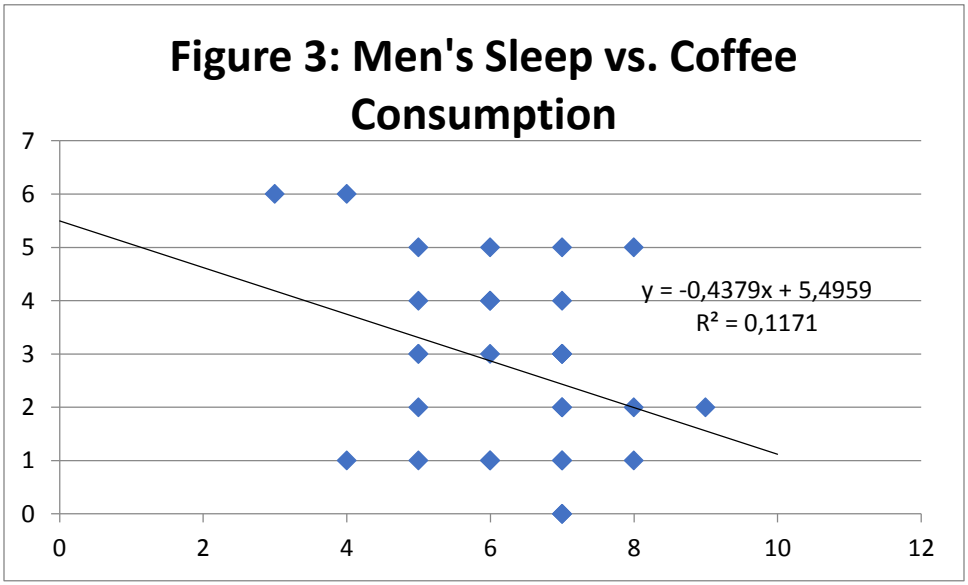
Table 5: Calculations

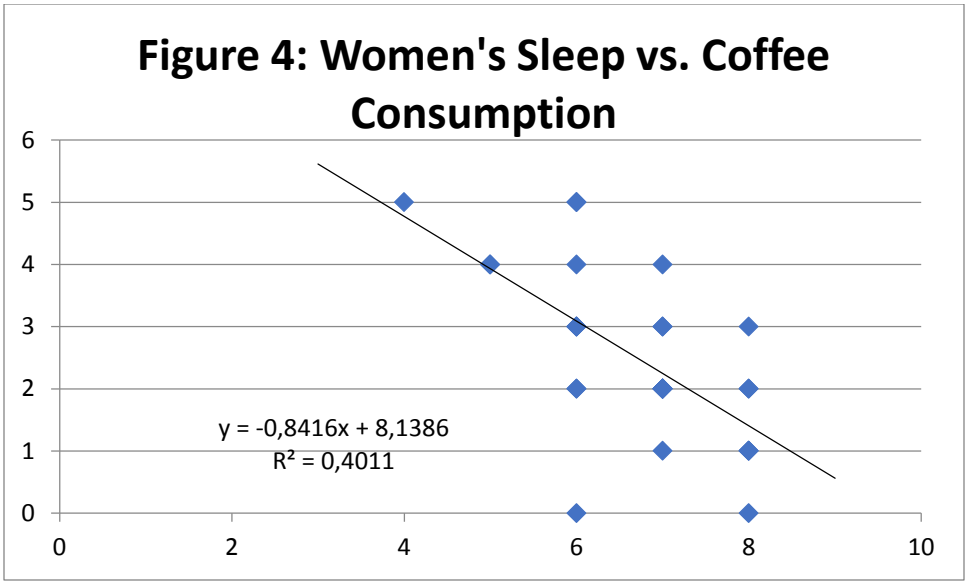
Men			Women		
Sleep	**Coffee**		**Sleep**	**Coffee**	
x-bar	**y-bar**	**xy-bar**	**x-bar**	**x-bar**	**xy-bar**
6.23333333	2.76666667	16.46666667	6.7	2.5	15.9
(x-bar)^2	**(y-bar)^2**		**(x-bar)^2**	**(y-bar)^2**	
40.6333333	10.5666667		45.9	8.033333333	
sd	**sd**	**covariance**	**sd**	**sd**	**covariance**
1.33374993	1.70652343	-0.77888889	1.004987562	1.335415042	-0.85

n
30
r
-0.34220696
m
-0.43785134
c
5.49594004

n
30
r
-0.63334738
m
-0.84158416
c
8.138613861

Figures 3 and 4 below show the scatter plots with trend lines for men and women respectively. Also shown on the graphs are their respective trend line equations as well as their R-squared values. All of these equations and values are consistent with those found by the previous equations.





Analysis and Conclusion

It can be seen from the previous data as well as from Figures 3 and 4, that there is a weak correlation for both genders between their hours of sleep per night and their cups of coffee consumed per day. For men the Correlation Coefficient $r = -0.3422$ while for women $r = -0.6333$ which means that for both genders, there is a noticeable negative correlation between these two variables. It can also be seen that there is a stronger relationship between women's sleep and coffee consumption than in men.

The negative r and m values are indicative of the negative correlation which implies that as people get more sleep, they drink fewer cups of coffee. Thus, there is enough evidence to reject the null hypothesis and conclude that there is a correlation (though weak) between the number of hours a person sleeps per night and their coffee consumption throughout the next day.

I do think that my method of analysis is effective though rather constrained by resources to only students and their parents within my own personal access, which may make the data unintentionally biased. It would be nice to expand this analysis to a larger study group, possibly including the entire school or even going so far as to find out these correlations for the entire country (Lopez-Garcia et al. 2006). More advanced methods would have to be employed of course but previous studies like this have been conducted (Horne 2018) though the comparison of men to women in this context, I have not seen before.

Personally this process has shown me how difficult it can be to come up with a viable project idea as I had to change my topic three times before I discovered something I was both interested in and able to easily accomplish. It was also challenging to word my title as well as my survey questions so that they were both easily understood and answerable within a short time frame. This has given me a new found respect for mathematicians and statisticians as I have gotten a small glimpse of the work that goes into a study like this.

Works Cited

Gangwisch, J.E., Heymsfield, S.B., Boden-Albala, B., Buijs, R.M., Kreier, F., Pickering, T.G., Rundle, A.G., Zammit, G.K., and Malaspina, D., "Sleep Duration as a Risk Factor for Diabetes Incidence in a Large US Sample." *Sleep Duration and Diabetes* v. 30, 12. (2007) pp. 1667-1673.

Horne, J., "Do Women Need More Sleep Thank Men?" National Sleep Foundation. (2018). https://sleepfoundation.org/sleep-news/do-women-need-more-sleep-men

Lopez-Garcia, E., van Dam, R. M., Willett, W. C., Rimm, E. B., Manson, J. E., Stampfer, M. J., Rexrode, K. M., and Hu, F. B., "Coffee Consumption and Coronary Heart Disease in Men and Women: A Prospective Cohort Study." *Circulation Journal of the American Heart Association* (2006) p. 10.

Salazar-Martinez, E., Willett, W. C., Ascherio, A., Manson, J. E., Leitzmann, M. F., Stampfer, M. J., and Hu, F. B., "Coffee Consumption and Risk for Type 2 Diabeties mellitus." *Annals of Internal Medicine*. (2004) p. 9.