Student Survey

Myranda Swartzwelter

1/25/2022

Student Survey Covariance

Covariance is the averaged sum of combined deviations. A positive covariance indicates that as one variable deviates from the mean, the other variable also deviates from the mean in the same direction. A negative covariance indicates that as one variable deviates from the mean, the other variable deviates from the mean in the other direction.The magnitude of difference is dependent on the scale of the data being measured.

Table 1: Covariance of Student Survey variables

	TimeReading	TimeTV	Happiness	Gender
TimeReading	3.0545455	-20.3636364	-10.350091	-0.0818182
TimeTV	-20.3636364	174.0909091	114.377273	0.0454545
Happiness	-10.3500909	114.3772727	185.451422	1.1166364
Gender	-0.0818182	0.0454545	1.116636	0.2727273

These results indicate that:

- TimeReading is inversely related to TimeTV, Happiness and Gender and vice versa.
- Time TV is positively related to Happiness and Gender and vice versa.
- Happiness is positively related to Gender and vice versa.

Examining Variables

Table 2: Example of Student Survey Variables

TimeReading	TimeTV	Happiness	Gender
1	90	86.20	1
2	95	88.70	0
2	85	70.17	0
2	80	61.31	1
3	75	89.52	1
4	70	60.50	1

Looking at the variables in the student survey table, we find that all of the variables are using different scales of measurement. For example, it appears that TimeReading is likely measured in hours whereas TimeTV is likely measured in minutes. Although if corrected for this by converting TimeReading to minutes

or TimeTV to hours, it would not change the sign of the covariance (so the relationship between variables remains positive or inverse), it would change the magnitude of the covariance between variables

Table 3: Covariance of Student Survey variables

	TimeReading	TimeReading_min	TimeTV	Happiness	Gender
TimeReading	3.0545455	183.272727	-20.3636364	-10.350091	-0.0818182
TimeReading_min	183.2727273	10996.363636	-1221.8181818	-621.005455	-4.9090909
TimeTV	-20.3636364	-1221.818182	174.0909091	114.377273	0.0454545
Happiness	-10.3500909	-621.005455	114.3772727	185.451422	1.1166364
Gender	-0.0818182	-4.909091	0.0454545	1.116636	0.2727273

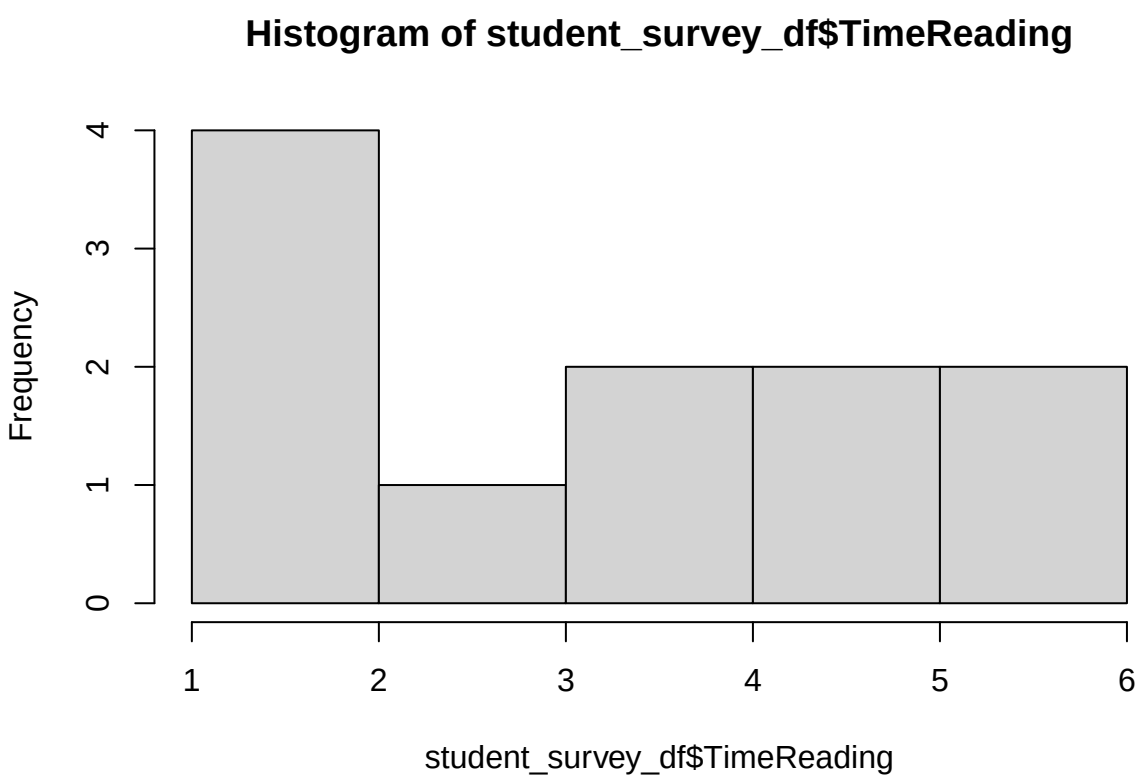
Happiness is likely on a scale from 1-100 but we have no concept of a scale for this metric. It is unable to be changed into a minute or hour measure like the previous variables were.

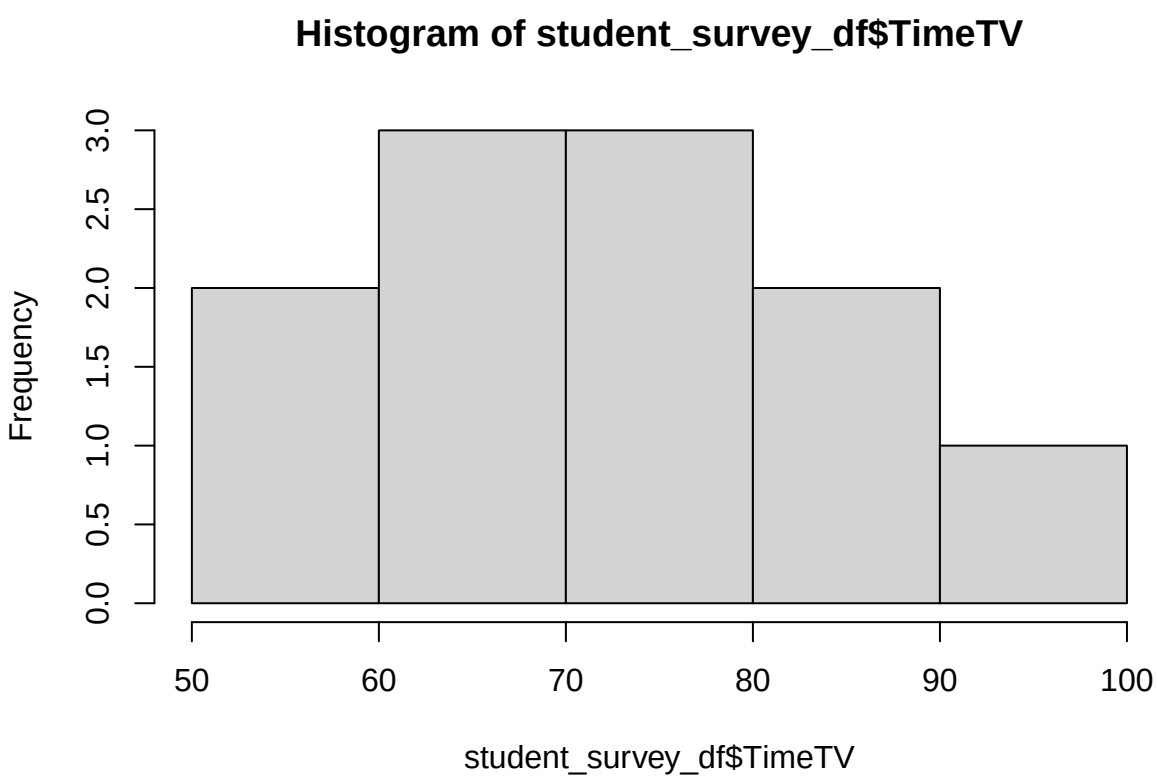
Gender is an indicator function, indicating a student is on gender if they are labeled as '1' and the other gender if labeled as '0'.

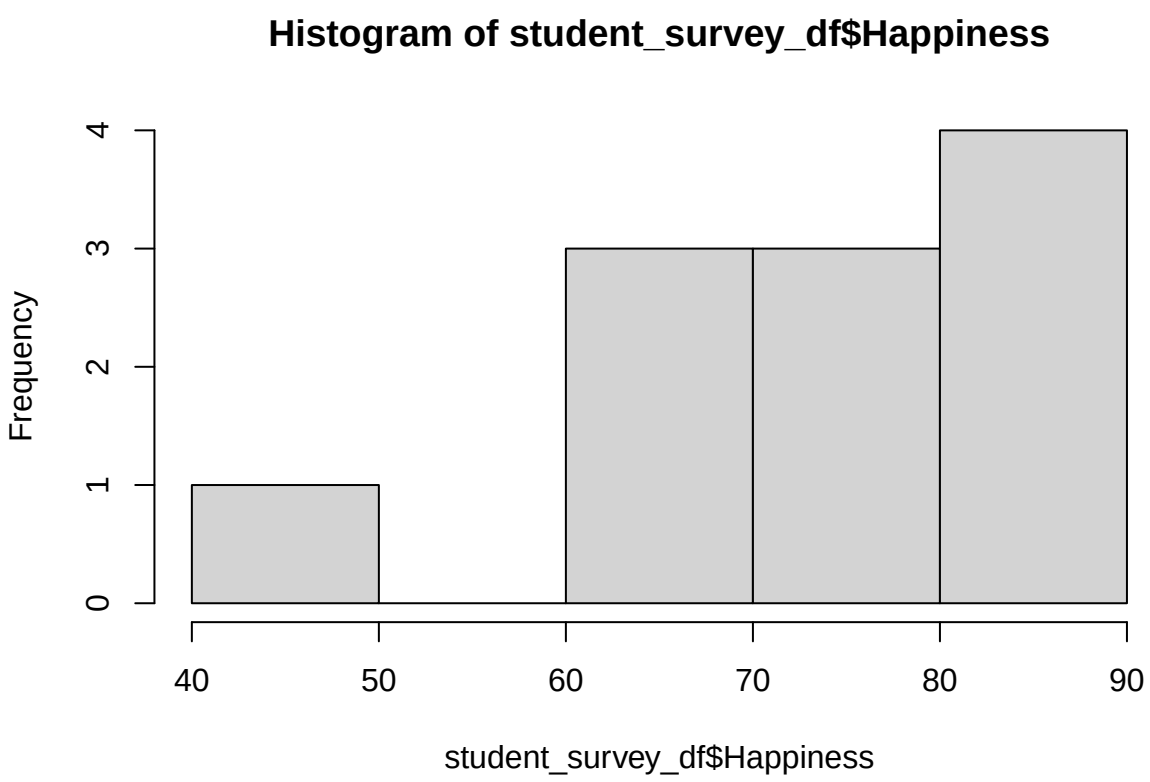
If you swap the gender indicators (i.e. the 1's become 0's and the 0's become 1's), you also change the direction of the relationship between Gender and the other variables, but maintain the magnitude. See table for examples.

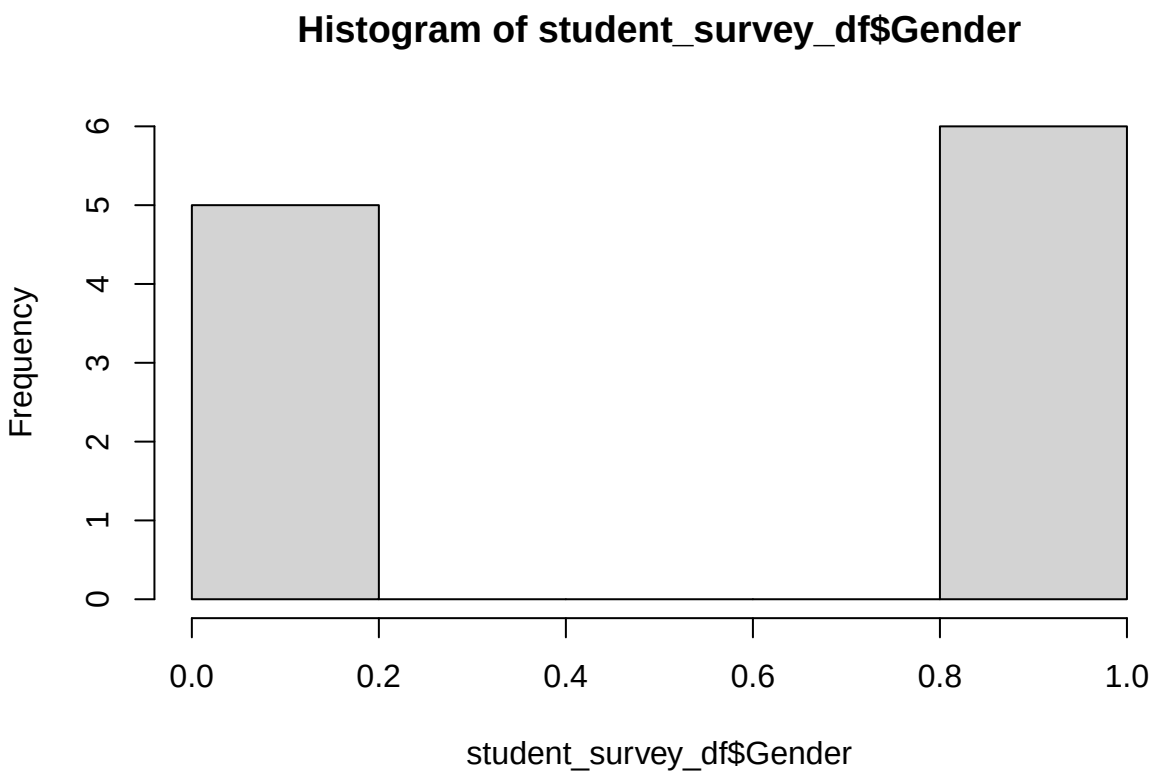
Table 4: Covariance of Student Survey variables

	TimeReading	TimeTV	Happiness	Gender	Gender_Swap
TimeReading	3.0545455	-20.3636364	-10.350091	-0.0818182	0.0818182
TimeTV	-20.3636364	174.0909091	114.377273	0.0454545	-0.0454545
Happiness	-10.3500909	114.3772727	185.451422	1.1166364	-1.1166364
Gender	-0.0818182	0.0454545	1.116636	0.2727273	-0.2727273
Gender_Swap	0.0818182	-0.0454545	-1.116636	-0.2727273	0.2727273


Histogram of student_survey_df$TimeReading
Frequency
0
1
2
3
4
1
2
3
4
5
6
student_survey_df$TimeReading


Histogram of student_survey_df$TimeTV
Frequency
0.0
0.5
1.0
1.5
2.0
2.5
3.0
50
60
70
80
90
100
student_survey_df$TimeTV


Histogram of student_survey_df$Happiness
Frequency
0
1
2
3
4
40
50
60
70
80
90
student_survey_df$Happiness


Histogram of student_survey_df$Gender
Frequency
0
1
2
3
4
5
6
0.0
0.2
0.4
0.6
0.8
1.0
student_survey_df$Gender