Design and analysis of experiments

Daniel Kartin - 20163822
Design and analysis of experiments 8 - Correlation

May 31, 2018

I'll be making these assignments with R instead of matlab, so the syntax and plots will look different. I really feel like the course is too tool dependent, and R really should be taught in tandem with matlab.

1

I load in the csv file, and put it into a data frame called data. And then correlating the pairs using Pearson's method, as seen in Table 1.

```
1 data <- read.csv('anscombe.csv')
```

Table 1

	X1 & Y1	X2 & Y2	X3 & Y3	X4 & Y4
Pearson's r	0.81642	0.81623	0.81628	0.81652
P-value	0.00217	0.00217	0.00217	0.00216

And the correlational plot is as follows:

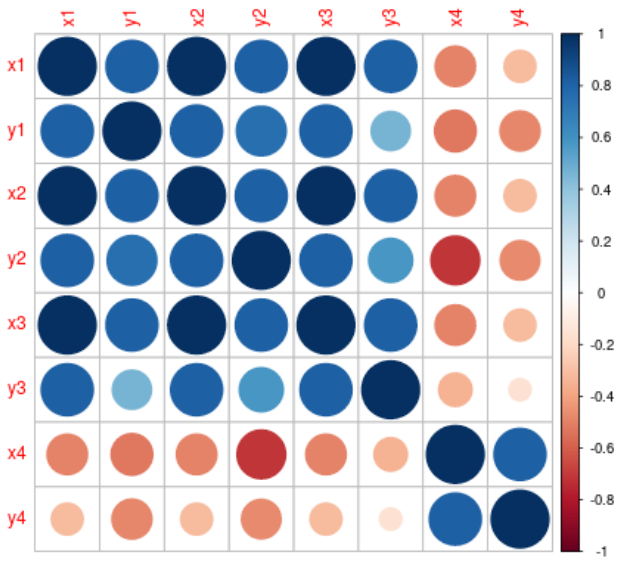

Figure 1

2

The level of measurement of the the happy data is interval.

Table 2

	Cash & Happy
Pearson's r	-0.2945288
P-value	2.038e-7

The correlation plot is seen in Figure 2.

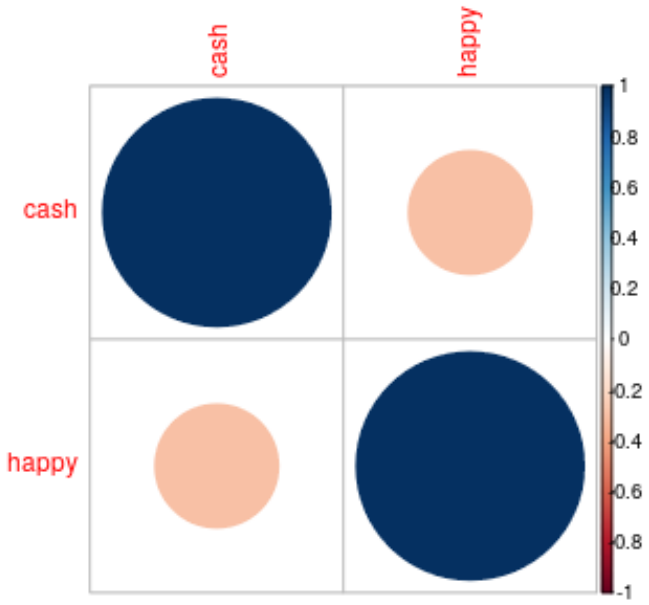

Figure 2

3

H_0 = Average weekly gaming hours are the same for male and female Medialogy students.
H_1 = Average weekly gaming hours are not the same for male and female Medialogy students.

The two sample independent t test returns a p-value of 0.004901, which is below the the confidence level of 0.05, hence the H_0 can be rejected.

To calculate the effect size I made the Cohen's d function in R, it looks as follows:

```
cohens_d <- function(x, y) {
    lenX <- length(x) - 1
    lenY <- length(y) - 1
    meanDiff <- abs(mean(x) - mean(y))
    pssd <- lenX * var(x) + lenY * var(y)
    pssd <- pssd/(lenX + lenY)
    pssd <- sqrt(pssd)
```

```
8       print("Cohen's D")
9       print(cd <- meanDiff/pssd)
10  }
```

It outputs a d value of 0.7572345, which according to the rule of thumb, is close to being a big effect.