worksheet4

Sunny Lee

2/10/2021

1)

a) A corr(X, Y) value of -.79 shows that X and Y are closely inversely correspondent/dependent.

b) Since the correlation coefficient is close to zero, we find that X and Y are not very dependent in this case.

c) corr(X, X) should be 1 since X should be perfectly depenedent with X.

2) a, b)

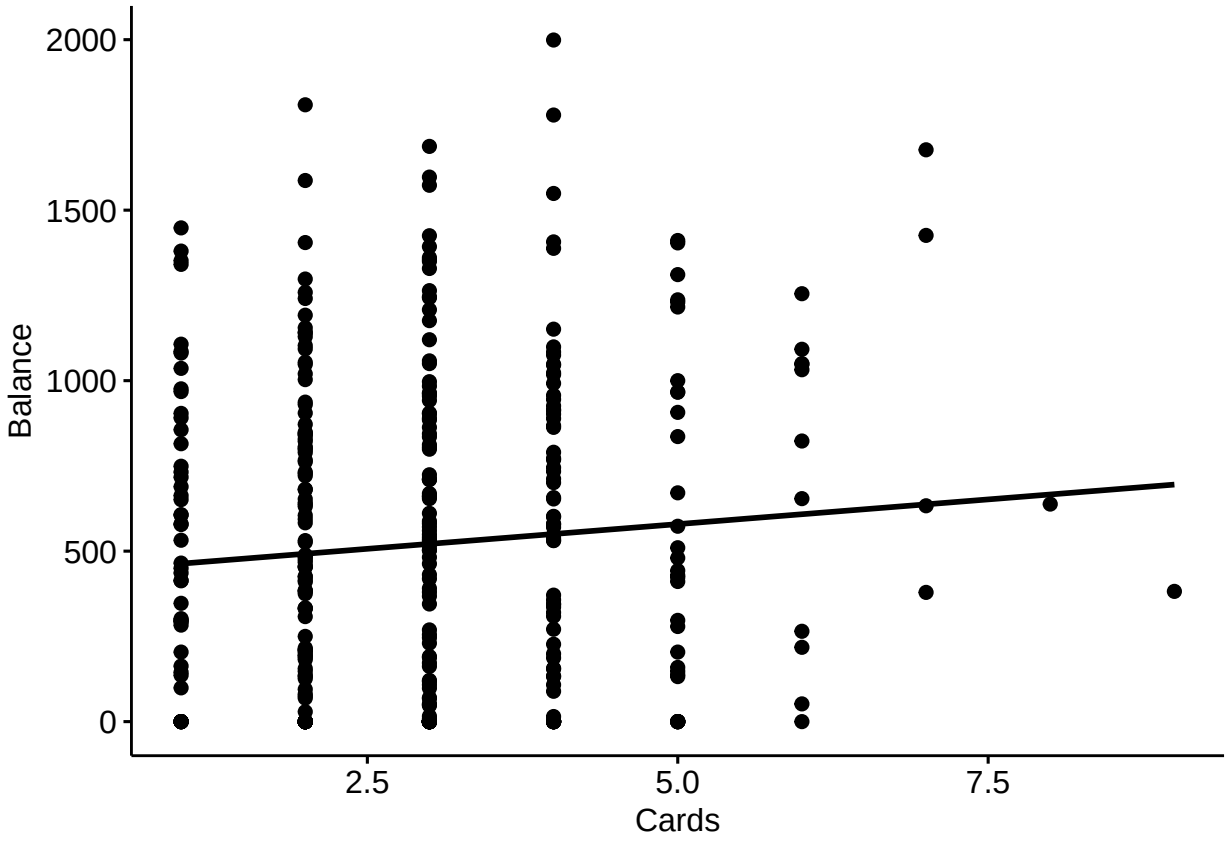
```
library("ggpubr")
```

```
## Loading required package: ggplot2
```

```
credit <- read.csv("Credit.csv")

ggscatter(credit, x = "Cards", y = "Balance", add = "reg.line")
```

```
## `geom_smooth()` using formula 'y ~ x'
```

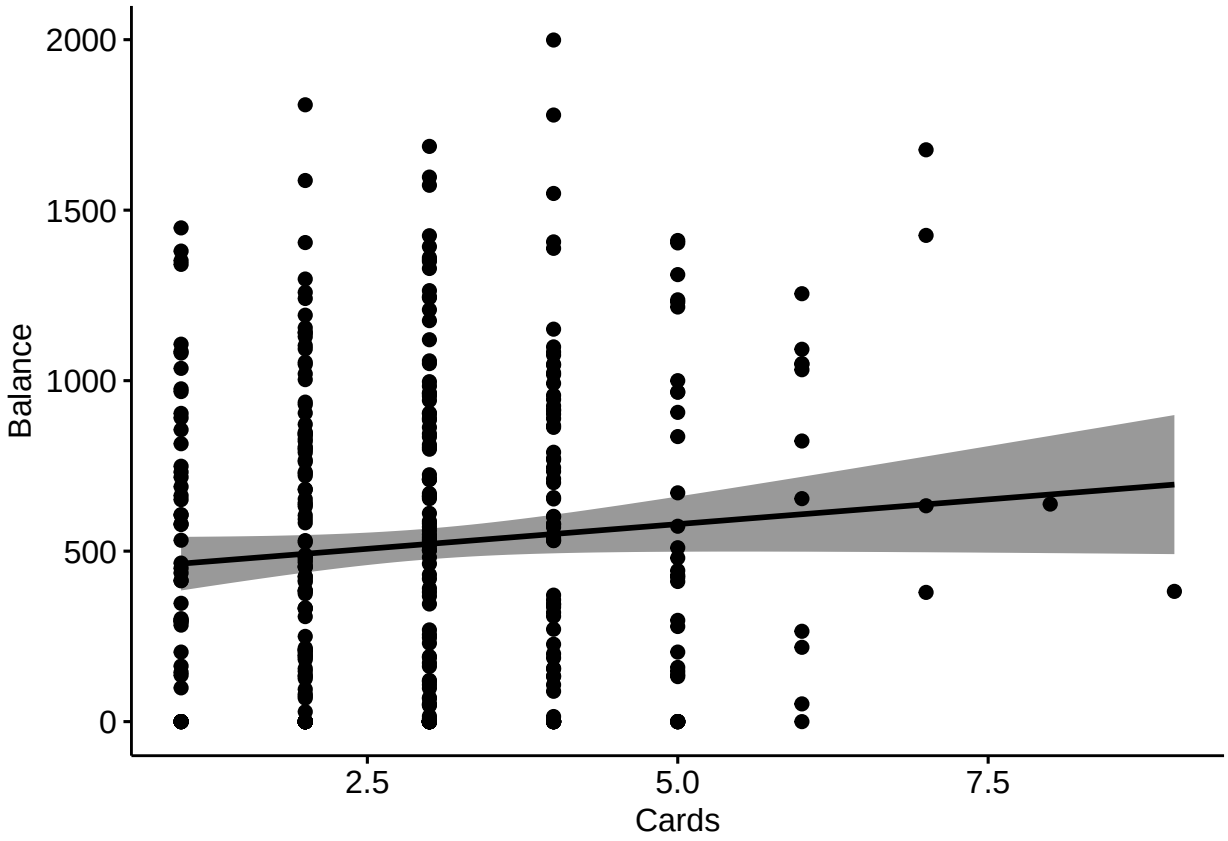


c) From the above graph, we find that Balance does seem to be correlated with Cards, and there seems to be a positive correlation.

d)

```
ggscatter(credit, x="Cards", y="Balance", add="reg.line", conf.int=TRUE)
```

```
## `geom_smooth()` using formula 'y ~ x'
```

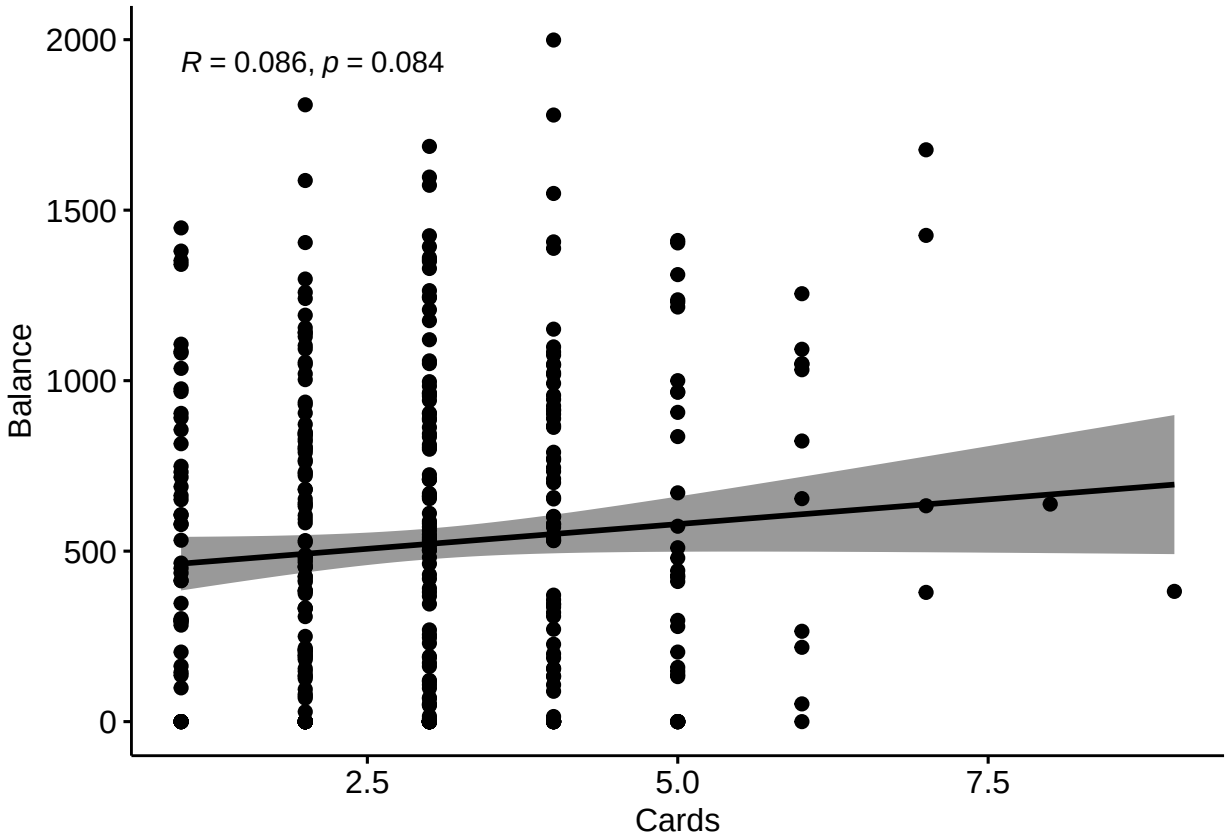


e) The shaded region does look like it could capture a line with the $\hat{\beta}_1 = 0$. One such line could be $\hat{\beta}_0 = 500$, which would be a straight line only at $Balance = 500$ which looks like it could be contained in the shaded region.

f)

```
ggscatter(credit, x="Cards", y="Balance", add="reg.line", conf.int=TRUE, cor.coef=TRUE)
```

```
## `geom_smooth()` using formula 'y ~ x'
```



g) From what we have seen above, we can conclude that Cards is not a good predictor for Balance, as the p value is very low, and the shaded region does seem to contain a line where $\hat{\beta}_1 = 0$