InClassScript3

Qingqing Yang

03/02/2022

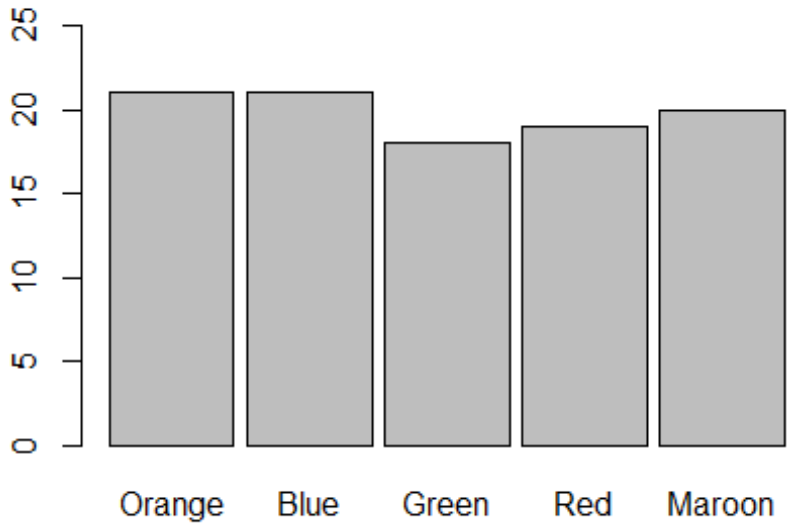
Question 1

Create data frame Q1_df

```
Q1_df <- data.frame(
  categories=c("Orange","Blue","Green","Red","Maroon"),
  freq = c(21,21,18,19,20)
)
```

Plot the data

```
barplot(Q1_df$freq,
        names.arg = Q1_df$categories,
        ylim=c(0,25),
        space=0.1
)
```



Question 2

Create data frame Q2_df

```
library(knitr)
Q2_df <- data.frame(
  Participant = 1:15,
  Introspection = c(51,65,74,74,75,76,87,86,51,85,52,92,97,70,86),
  Optimism = c(55,80,71,66,74,68,51,67,74,67,81,83,87,74,84)
)
kable(Q2_df)
```

Participant	Introspection	Optimism
1	51	55
2	65	80
3	74	71
4	74	66
5	75	74
6	76	68
7	87	51
8	86	67
9	51	74
10	85	67
11	52	81
12	92	83
13	97	87
14	70	74
15	86	84

Calculate Pearson's r

```
cor(Q2_df$Introspection, Q2_df$Optimism,method = "pearson")
```

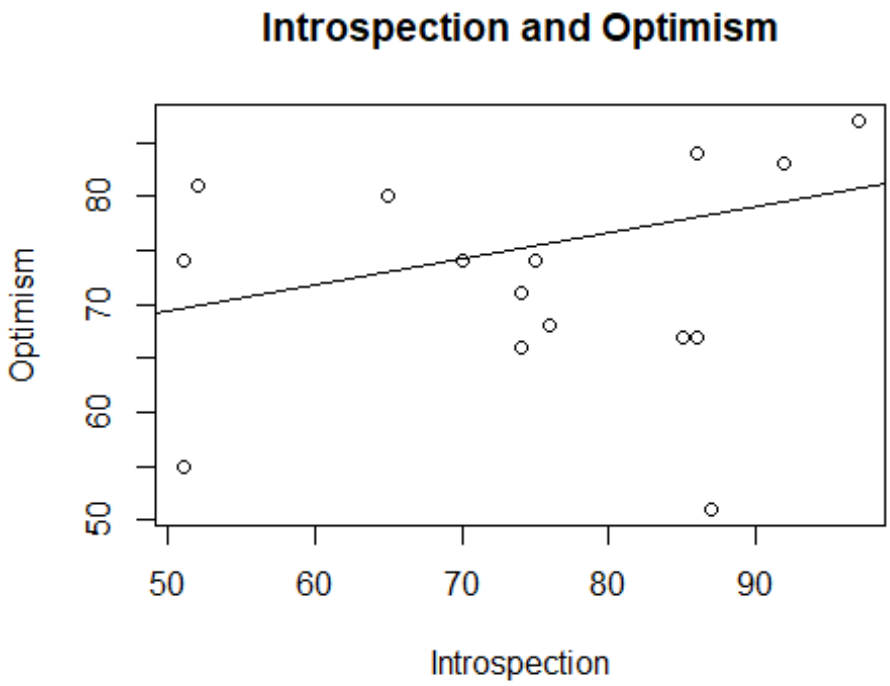
```
## [1] 0.1683006
```

The pearson's r equals 0.168.

Plot the data and trend line

```
plot(Q2_df$Introspection, Q2_df$Optimism,
     main = "Introspection and Optimism",
     xlab = "Introspection",
     ylab = "Optimism"
)
```

```
lm_Q2=lm(Q2_df$Introspection ~ Q2_df$Optimism)
abline(lm_Q2)
```




```
lm_Q2
```

```
##
## Call:
## lm(formula = Q2_df$Introspection ~ Q2_df$Optimism)
##
## Coefficients:
##    (Intercept)  Q2_df$Optimism
##        57.2557          0.2423
```

Therefore, the equation should be: $Optimism = 0.2423 * Introspection + 57.2557$

Question 3

Get the data frame for Q3

```
Q3_df <- data.frame(
  Participant = 1:12,
  Visual_Imagery_Rating= c(44,38,40,43,47,31,42,42,48,28,40,41),
  Vivid_Dreaming = NaN
)
kable(Q3_df)
```

Participant	Visual_Imagery_Rating	Vivid_Dreaming
1	44	NaN
2	38	NaN
3	40	NaN
4	43	NaN
5	47	NaN
6	31	NaN
7	42	NaN
8	42	NaN
9	48	NaN
10	28	NaN
11	40	NaN
12	41	NaN

Temporary set the vivid dreaming value to be NaN

Calculate the y by $y = bx + a$

```
slope = 0.004
intercept= 6.023
for (i in 1:12){
Q3_df$Vivid_Dreaming[i]=
  slope*Q3_df$Visual_Imagery_Rating[i]+intercept
}
kable(Q3_df)
```

Participant	Visual_Imagery_Rating	Vivid_Dreaming
1	44	6.199
2	38	6.175
3	40	6.183
4	43	6.195
5	47	6.211
6	31	6.147
7	42	6.191
8	42	6.191
9	48	6.215
10	28	6.135
11	40	6.183
12	41	6.187

Question 4

Data frame Q4_df

```
Q4_df <- data.frame(
  x = c(16,12,15,13,11),
  y_predicted= c(56.86,146.38,79.2415,124,168.76),
  residuals= c(-21.86,-66.38,30.76,26,31.24),
  y = NaN
)
kable(Q4_df)
```

x	y_predicted	residuals	y
16	56.8600	-21.86	NaN
12	146.3800	-66.38	NaN
15	79.2415	30.76	NaN
13	124.0000	26.00	NaN
11	168.7600	31.24	NaN

Find y by $y_i = \hat{y}_i + residual_i$

```
for (i in 1:5){
  Q4_df$y[i]=Q4_df$y_predicted[i]+Q4_df$residuals[i]
}
kable(Q4_df)
```

x	y_predicted	residuals	y
16	56.8600	-21.86	35.0000
12	146.3800	-66.38	80.0000
15	79.2415	30.76	110.0015
13	124.0000	26.00	150.0000
11	168.7600	31.24	200.0000

Plot x and y with range on x and y

```
plot(Q4_df$x, Q4_df$y,
     main = "x and y",
     xlab = "x",
     ylab = "y",
    xlim = c(0,25),
    ylim = c(0,250),
)
```

```r
# The h= and v= forms draw horizontal and vertical lines at the specified coo
rdinates
abline(v=12)
```