quiz6

YANING JIN

Simulate

```
set.seed(123)
hours <- 9:17

employee_1 <- rnorm(length(hours), mean=8, sd=2)

employee_2 <- employee_1 + rnorm(length(hours), mean=0, sd=1)

employee_1[1] <- 0
employee_2[1] <- 0

data <- data.frame(hour=hours, employee_1=round(employee_1), employee_2=round(employee_2))

print(data)
```

```
  hour employee_1 employee_2
1    9          0          0
2   10          8          9
3   11         11         11
4   12          8          9
5   13          8          8
6   14         11         11
7   15          9         11
8   16          5          6
9   17          7          5
```

```
cor.test(data$employee_1, data$employee_2)
```

```
    Pearson's product-moment correlation

data:  data$employee_1 and data$employee_2
t = 8.2144, df = 7, p-value = 7.696e-05
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.7821584 0.9900903
sample estimates:
      cor
0.9518455
```

```
t.test(data$employee_1, data$employee_2, paired = TRUE)
```

```
    Paired t-test

data:  data$employee_1 and data$employee_2
t = -0.89443, df = 8, p-value = 0.3972
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 -1.1927303  0.5260637
sample estimates:
mean difference
     -0.3333333
```

```
normality_test_employee_1 <- shapiro.test(data$employee_1)
normality_test_employee_2 <- shapiro.test(data$employee_2)
print(normality_test_employee_1)
```

```
    Shapiro-Wilk normality test

data:  data$employee_1
W = 0.86171, p-value = 0.1001
```

```
print(normality_test_employee_2)
```

```
    Shapiro-Wilk normality test

data:  data$employee_2
W = 0.85625, p-value = 0.08726
```

```r
morning_hours <- data$hour < 12
afternoon_hours <- data$hour >= 12

morning_hours <- data$hour < 12
afternoon_hours <- data$hour >= 12

t_test_morning_afternoon_employee_1 <- t.test(data$employee_1[morning_hours], data$employe
t_test_morning_afternoon_employee_2 <- t.test(data$employee_2[morning_hours], data$employe
print(t_test_morning_afternoon_employee_1)
```

```
    Welch Two Sample t-test

data:  data$employee_1[morning_hours] and data$employee_1[afternoon_hours]
t = -0.49266, df = 2.2516, p-value = 0.6662
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -14.76949  11.43616
sample estimates:
mean of x mean of y
 6.333333  8.000000
```

```r
print(t_test_morning_afternoon_employee_2)
```

```
    Welch Two Sample t-test

data:  data$employee_2[morning_hours] and data$employee_2[afternoon_hours]
t = -0.47161, df = 2.3738, p-value = 0.6771
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -14.79308  11.45975
sample estimates:
```

```
mean of x mean of y
 6.666667  8.333333
```

```r
outliers_employee_1 <- boxplot.stats(data$employee_1)$out
outliers_employee_2 <- boxplot.stats(data$employee_2)$out
print(outliers_employee_1)
```

```
[1] 0
```

```r
print(outliers_employee_2)
```

```
numeric(0)
```

```r
lm_fit <- lm(employee_2 ~ employee_1, data = data)
summary(lm_fit)
```

```
Call:
lm(formula = employee_2 ~ employee_1, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-2.3202 -0.4384 -0.1133  0.6502  1.6207

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.1133     1.0139   0.112    0.914
employee_1    1.0296     0.1253   8.214  7.7e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.191 on 7 degrees of freedom
Multiple R-squared:  0.906, Adjusted R-squared:  0.8926
F-statistic: 67.48 on 1 and 7 DF,  p-value: 7.696e-05
```

Explore

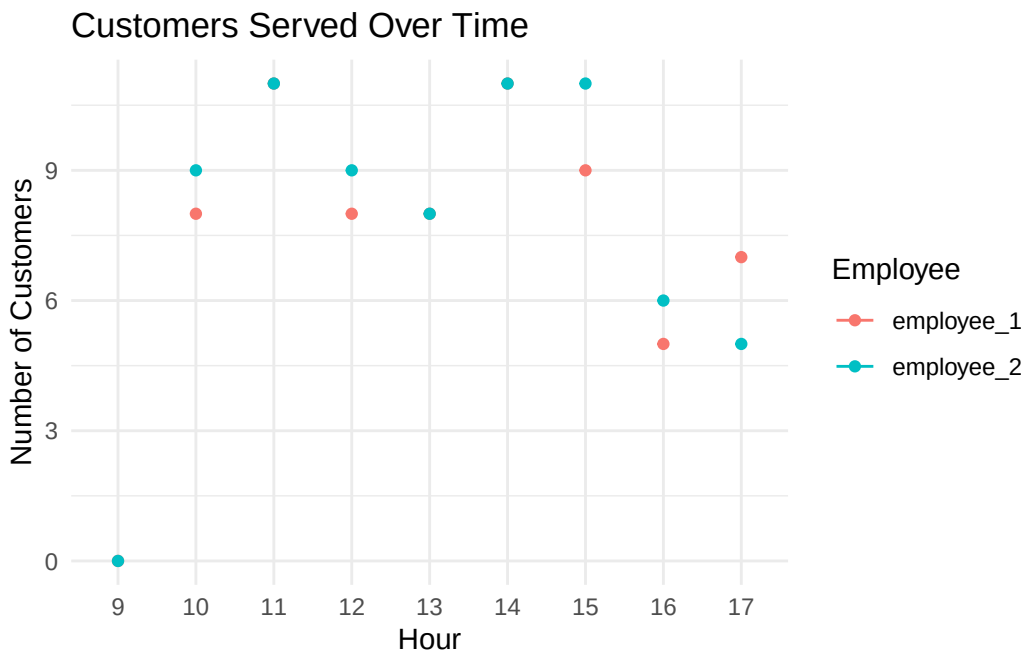
```
library(ggplot2)

data$hour <- factor(data$hour)

melted_data <- reshape2::melt(data, id.vars = 'hour', variable.name = 'employee')

ggplot(melted_data, aes(x = hour, y = value, color = employee)) +
  geom_line() +
  geom_point() +
  labs(title = "Customers Served Over Time",
       x = "Hour",
       y = "Number of Customers",
       color = "Employee") +
  theme_minimal()
```

```
`geom_line()`: Each group consists of only one observation.
i Do you need to adjust the group aesthetic?
```



Testthat

```
library("testthat")

german_cities <- c("Berlin", "Hamburg", "Munich", "Cologne", "Frankfurt", "Rostock")

library("testthat")
test_that("german_cities contains exactly the specified cities", {
  expected_cities <- c("Berlin", "Hamburg", "Munich", "Cologne", "Frankfurt", "Rostock")
  expect_equal(sort(german_cities), sort(expected_cities))
})
```

```
Test passed
```

```
test_that("german_cities contains exactly the specified cities", {
  expected_cities <- c("Berlin", "Hamburg", "Munich", "Cologne", "Frankfurt", "Rostock")
  expect_equal(sort(german_cities), sort(expected_cities))
})
```

```
Test passed
```

```
test_that("german_cities contains no duplicate entries", {
  expect_equal(length(german_cities), length(unique(german_cities)))
})
```

```
Test passed
```

```
test_that("german_cities is a character vector", {
  expect_true(is.character(german_cities))
})
```

```
Test passed
```