31_question_7

Sudarshan Budhathoki

2024-05-31

```r
# Load the airquality dataset
data(airquality)
```

```r
# a) Perform goodness-of-fit test on Temp variable to check if it follows normal distribution
shapiro.test(airquality$Temp)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  airquality$Temp
## W = 0.97617, p-value = 0.009319
```

```r
# Interpretation: The Shapiro-Wilk test is used to test the normality of the "Temp" variable.
# The null hypothesis is that the data is normally distributed. The p-value will indicate if
#the data significantly deviates from a normal distribution.

# b) Perform goodness-of-fit test on Temp variable by month variable to
#check if the variable of mpg are equal or not on am variable categories
#leveneTest(Temp ~ Month, data = airquality)

# Interpretation: The Levene's test is used to test the equality of variances
#for the "Temp" variable across the different months.
# It checks if the variances of "Temp" are equal across the month categories.

# c) The independent sample test that can be used to compare "Temp" variable by
#"Month" variable categories could be the one-way ANOVA test.
# This test can determine if there are statistically significant differences in
#the mean "Temp" values across the different months.

# d) Perform the one-way ANOVA test for the data
temp_anova <- aov(Temp ~ Month, data = airquality)
summary(temp_anova)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## Month         1   2413  2413.0   32.52 6.03e-08 ***
## Residuals   151  11205    74.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Interpretation: The one-way ANOVA test results provide information on whether
#there are statistically significant differences in the mean "Temp" values across
#the different months.
# The ANOVA results include the F-statistic, p-value, and other relevant statistics
#for interpretation.
```