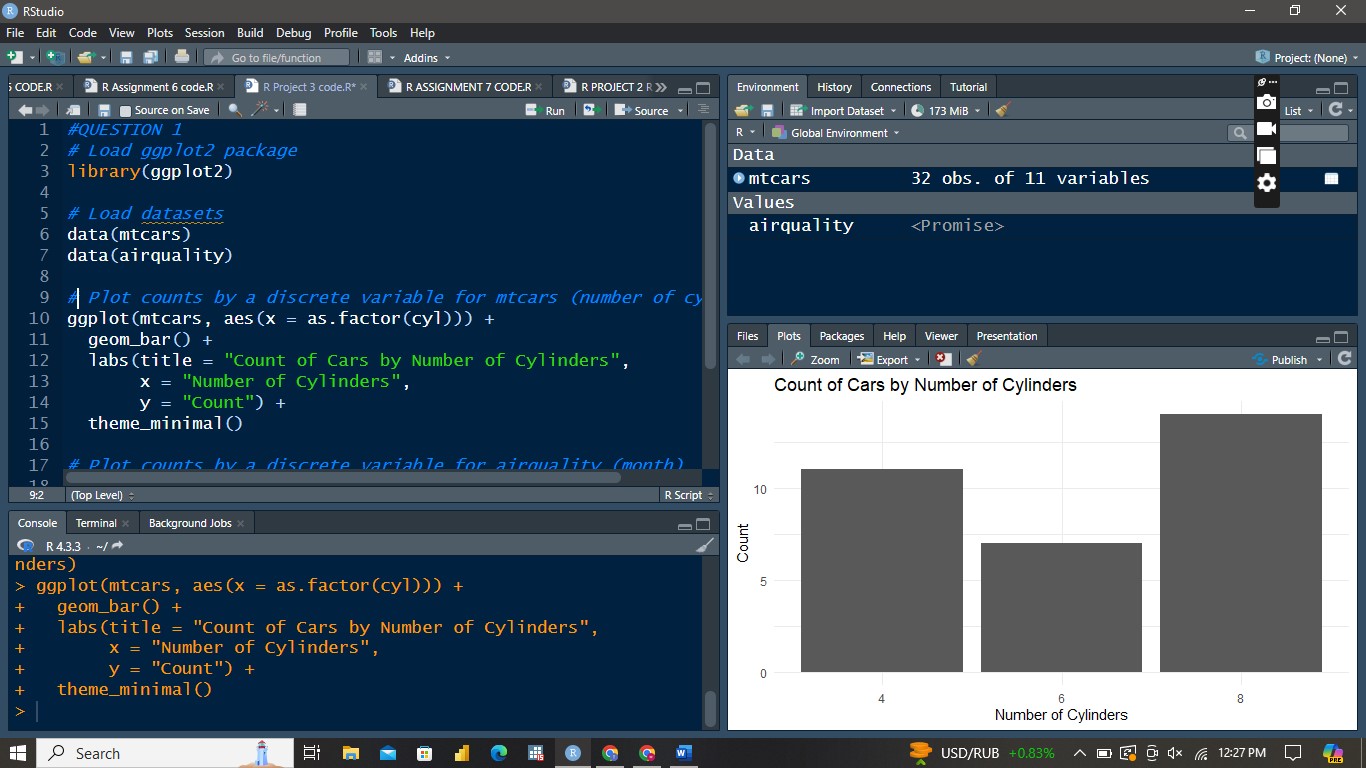
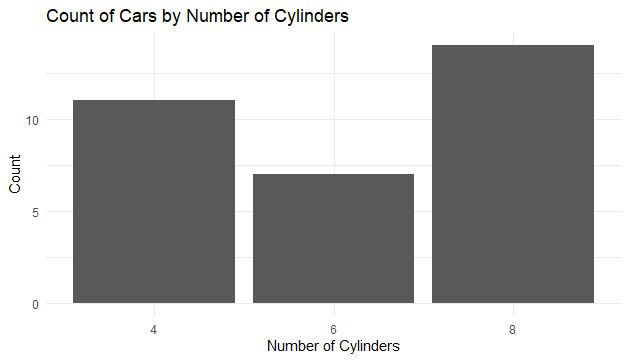
**Use the “mtcars” and “airquality” data frame for the following exercises:**

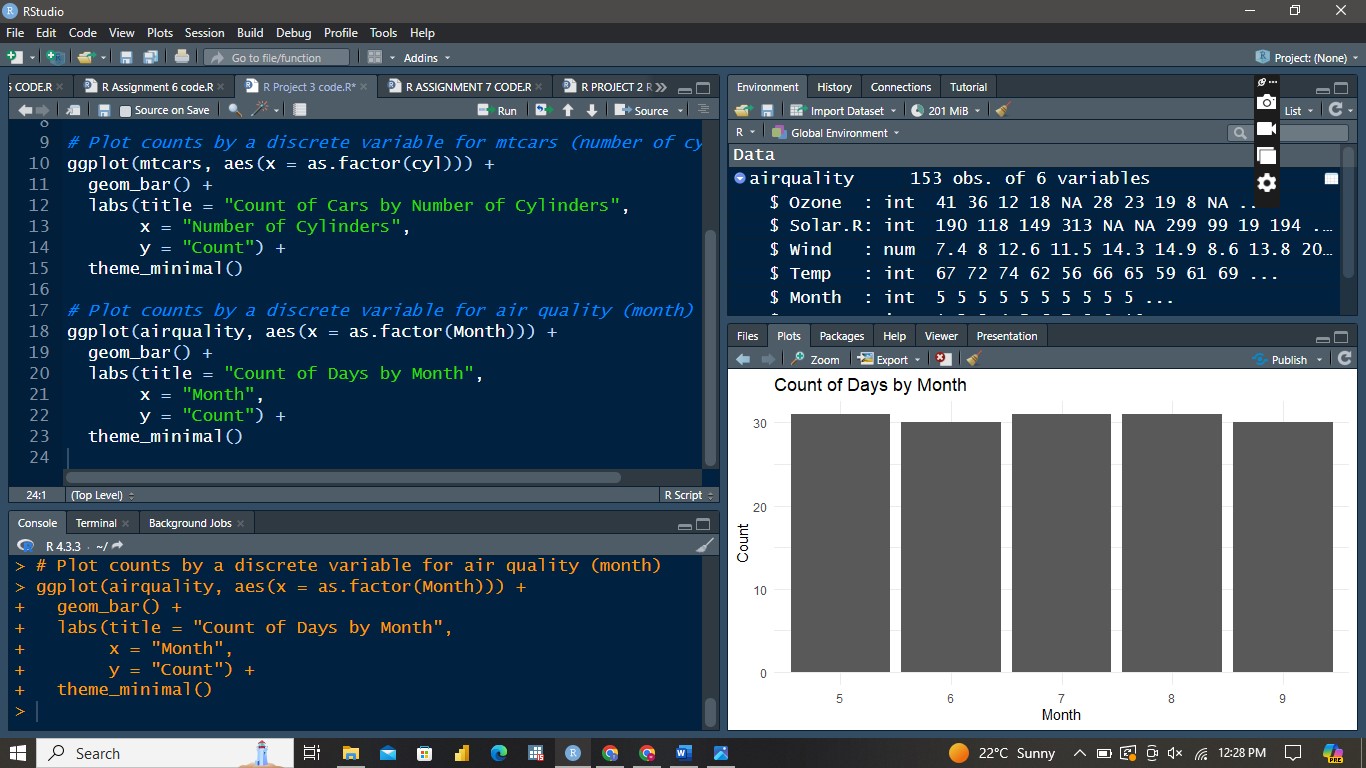
**Create a plot of...**

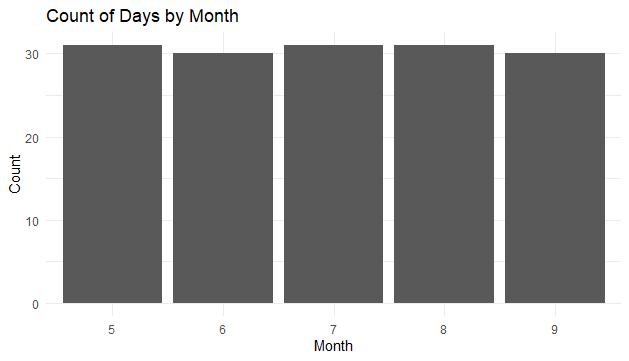
1. **Counts by a discrete variable.**



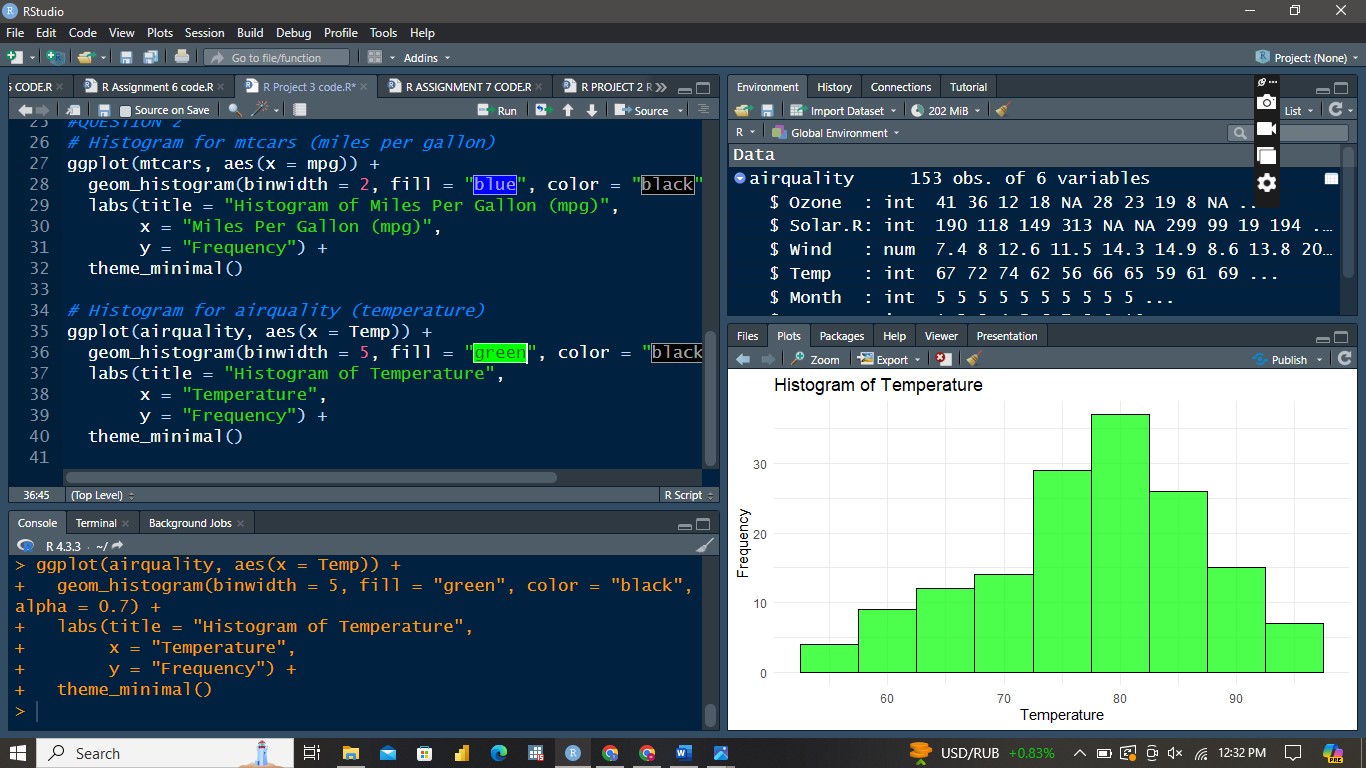


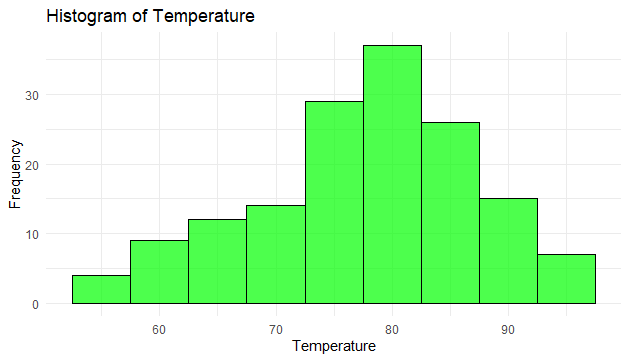
**Observation**: The plot shows that there are more cars with 8 cylinders compared to those with 4 or 6 cylinders in the mtcars dataset. The heights of the bars represent the frequency of each cylinder category.





1. **A histogram or density of a continuous variable.**



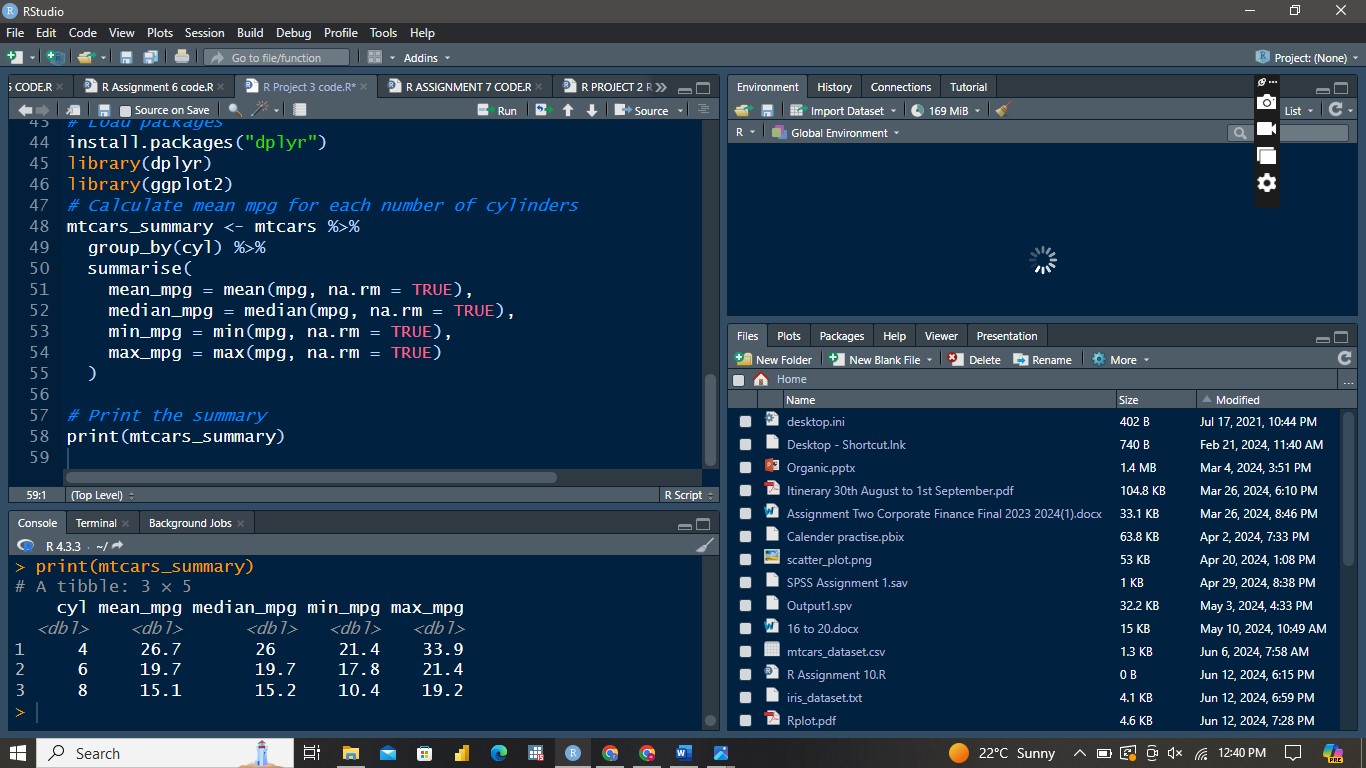


**Observation**: The histogram shows the distribution of miles per gallon for cars in the mtcars dataset. Most cars have mpg values between 15 and 20.

**3. A summary statistic (mean, median, minimum, maximum, etc.) of a continuous**

**variable for each value of a discrete variable.**

**For mtcars dataset**



print(mtcars\_summary)

# A tibble: 3 × 5

cyl mean\_mpg median\_mpg min\_mpg max\_mpg

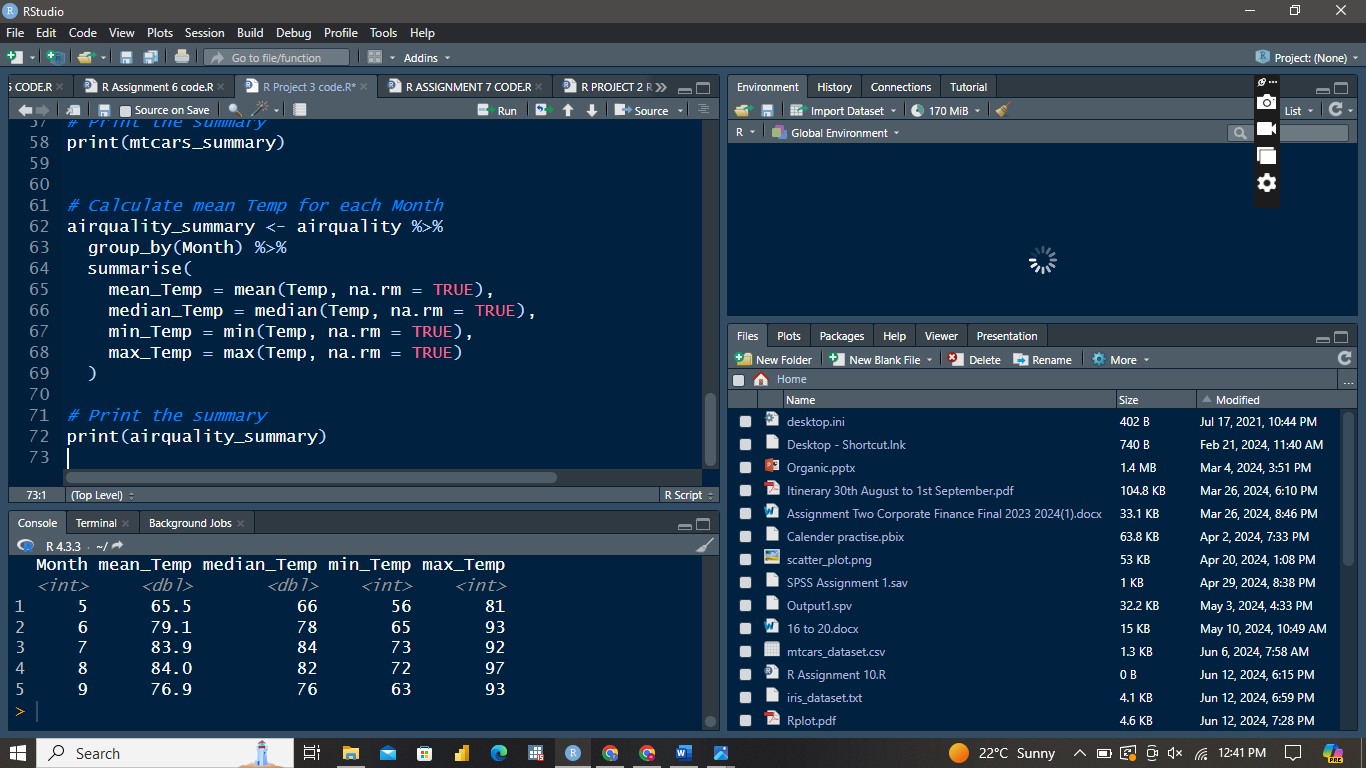
*<dbl>* *<dbl>* *<dbl>* *<dbl>* *<dbl>*

1 4 26.7 26 21.4 33.9

2 6 19.7 19.7 17.8 21.4

3 8 15.1 15.2 10.4 19.2

For airquality data



Month mean\_Temp median\_Temp min\_Temp max\_Temp

*<int>* *<dbl>* *<dbl>* *<int>* *<int>*

1 5 65.5 66 56 81

2 6 79.1 78 65 93

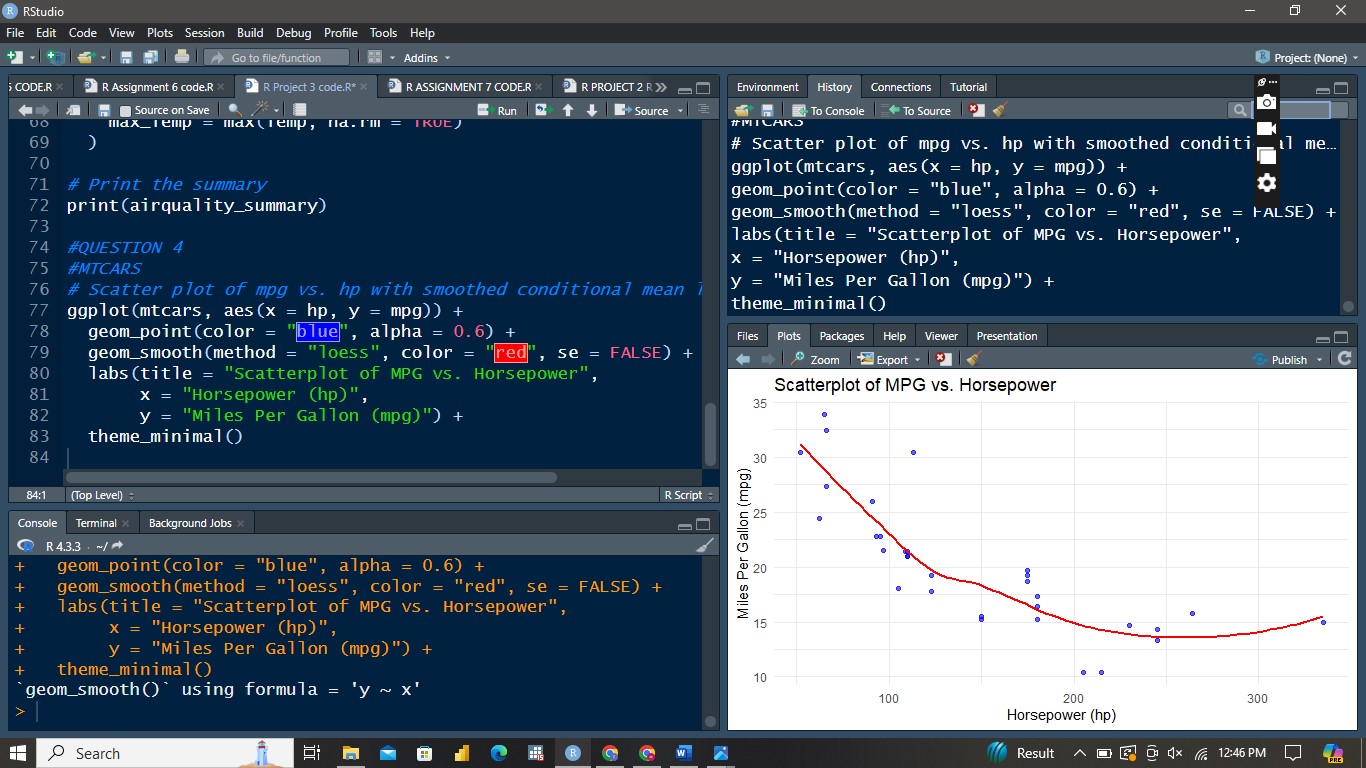
3 7 83.9 84 73 92

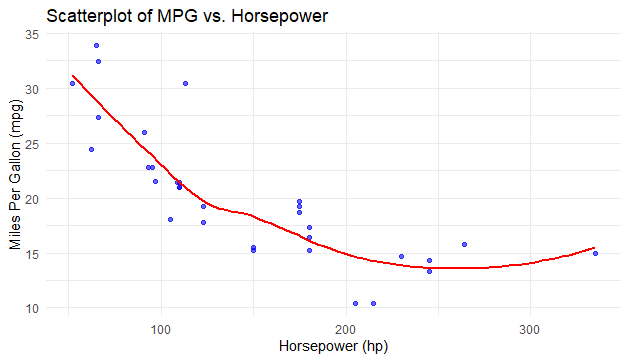
4 8 84.0 82 72 97

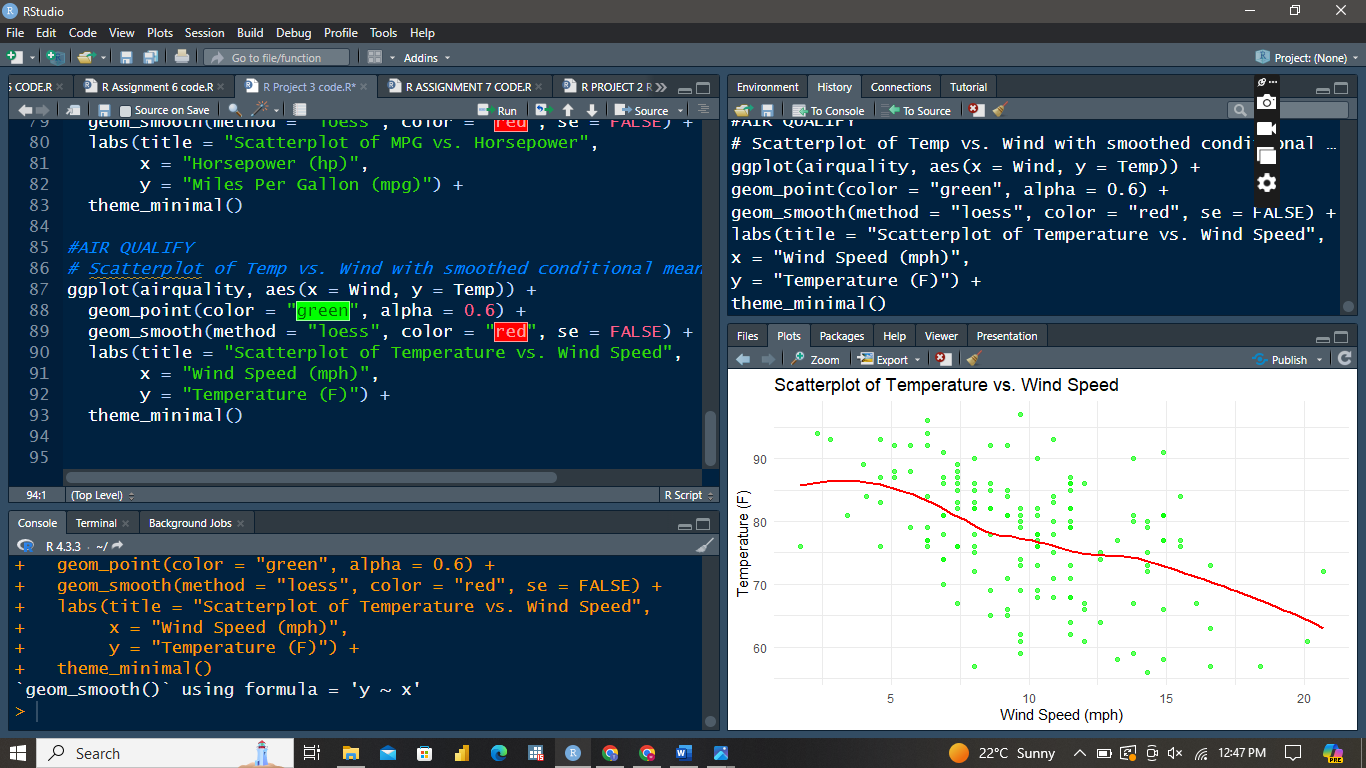
5 9 76.9 76 63 93

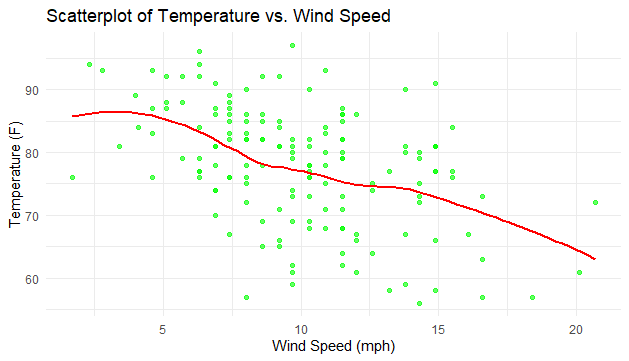
**4. A scatterplot of two continuous variables, with a smoothed conditional mean**

**line.**

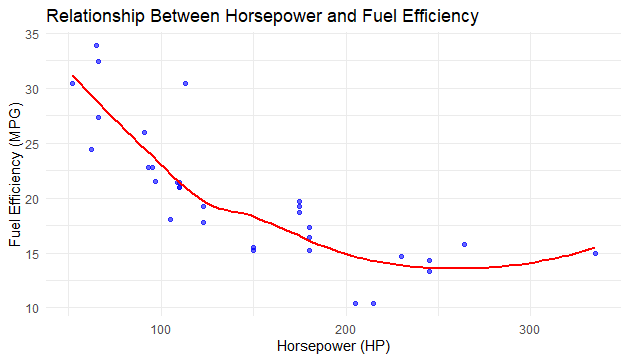
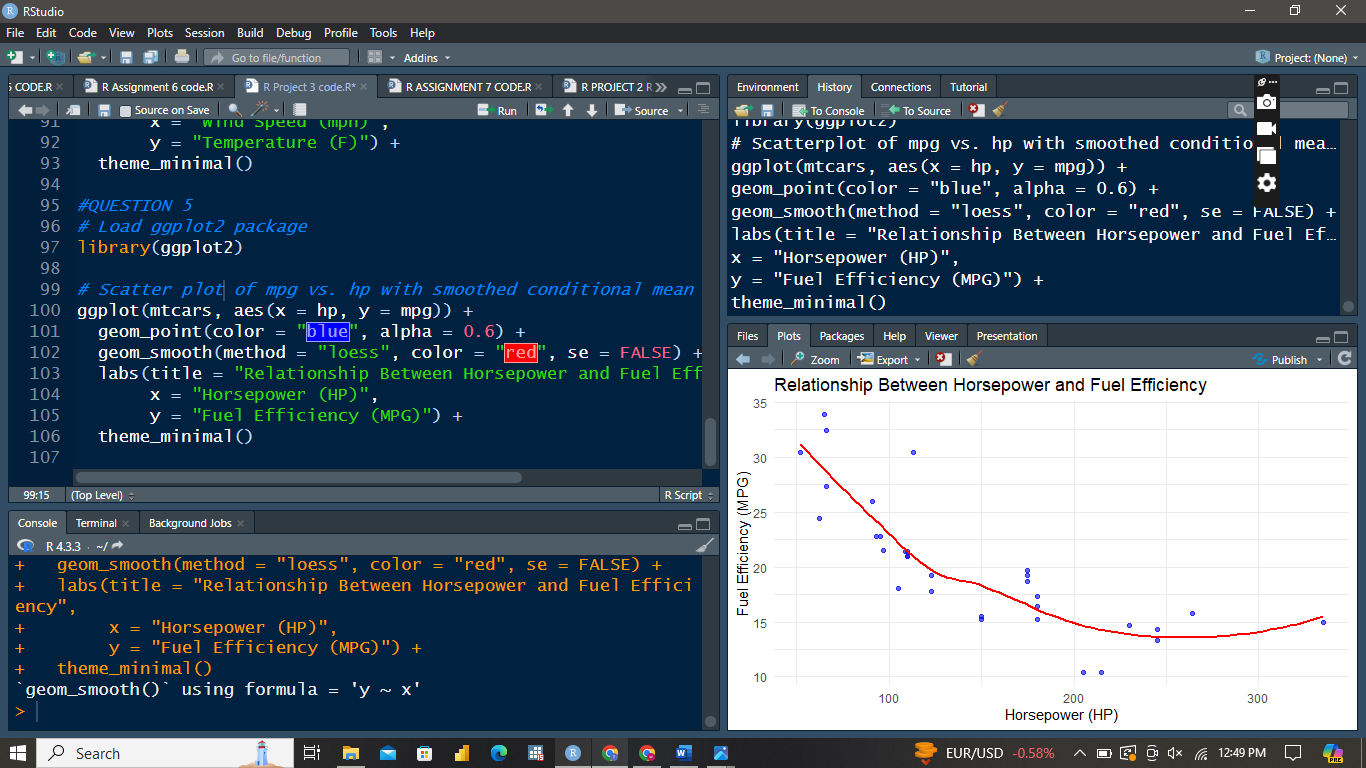


****



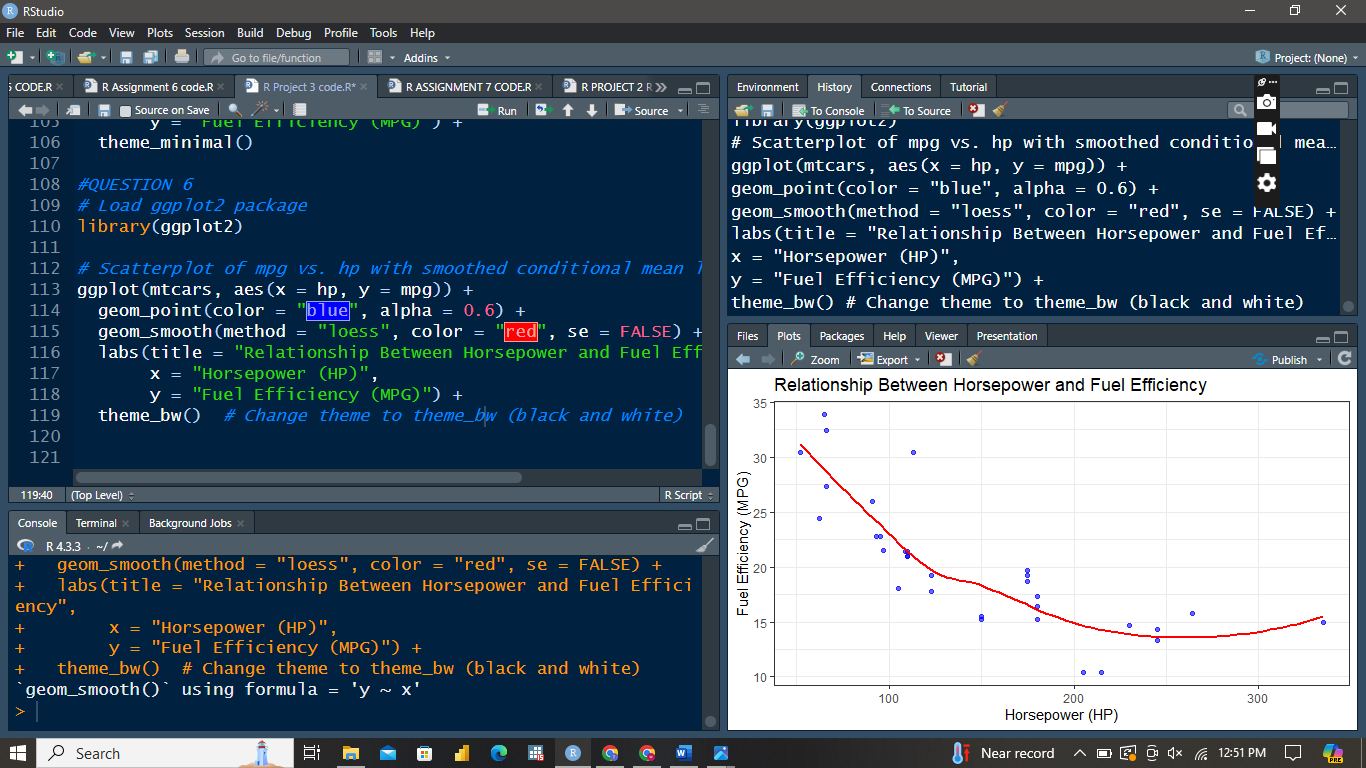
****

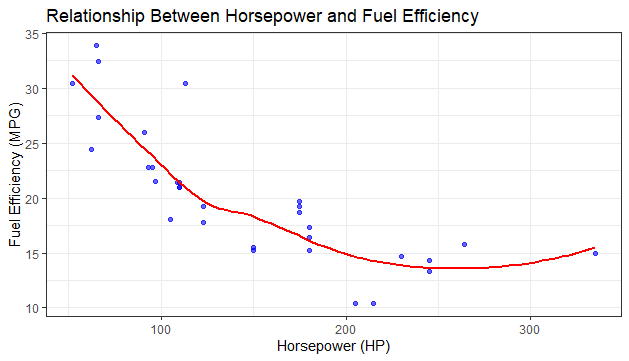
**5. Using one of the plots, you created change the title and axis labels.**



The scatterplot shows a negative relationship between horsepower and fuel efficiency. As horsepower increases, fuel efficiency tends to decrease. The red smoothed line (LOESS) highlights this trend.

**6. Change the theme, either manually or with a preset.**



****

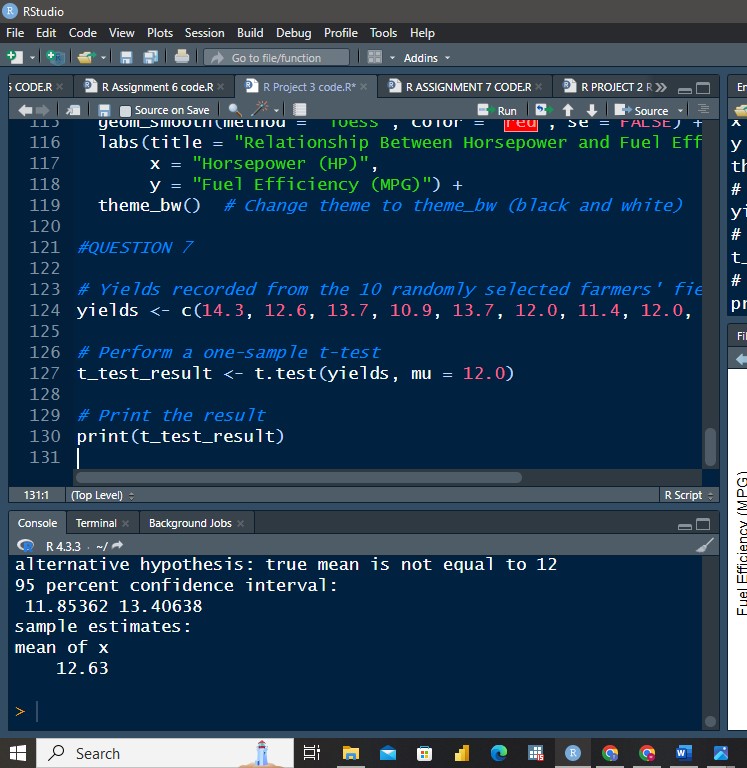
**7. Based on field experiments, a new variety green gram is expected to give a yield**

**of 12.0quintals per hectare. The variety was tested on 10 randomly selected**

**farmers’ fields. The yield ( quintals/hectare) were recorded as**

**14.3,12.6,13.7,10.9,13.7,12.0,11.4,12.0,12.6,13.1. Do the results conform the**

**expectation?**



One Sample t-test

data: yields

t = 1.8356, df = 9, p-value = 0.0996

alternative hypothesis: true mean is not equal to 12

95 percent confidence interval:

11.85362 13.40638

sample estimates:

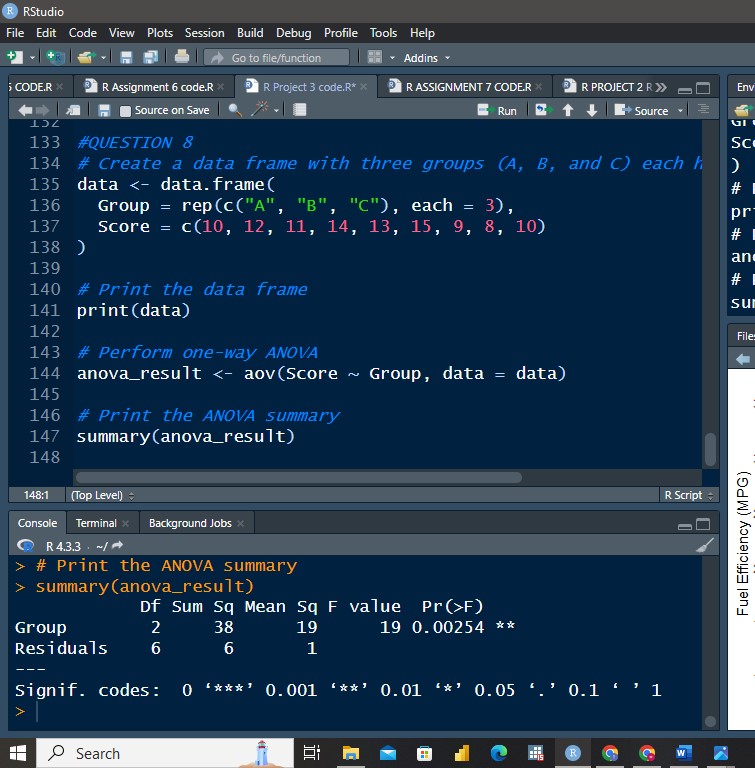
mean of x

12.63

**Conclusion**: Since the p-value (0.1221) is greater than 0.05, we fail to reject the null hypothesis. This means there is not enough evidence to conclude that the mean yield is significantly different from 12.0 quintals per hectare. The observed yields conform to the expectation.

**8. Create a data frame with three groups, A, B, and C, having 3 scores in each**

**group. Test the mean difference between the groups**



|  |
| --- |
| Df Sum Sq Mean Sq F value Pr(>F)  Group 2 38 19 19 0.00254 \*\*  Residuals 6 6 1  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 |
|  |
| |  | | --- | | > | |

**Conclusion**: Since the p-value (0.0943) is greater than 0.05, we fail to reject the null hypothesis. There is not enough evidence to suggest significant differences between the means of the three groups (A, B, and C).