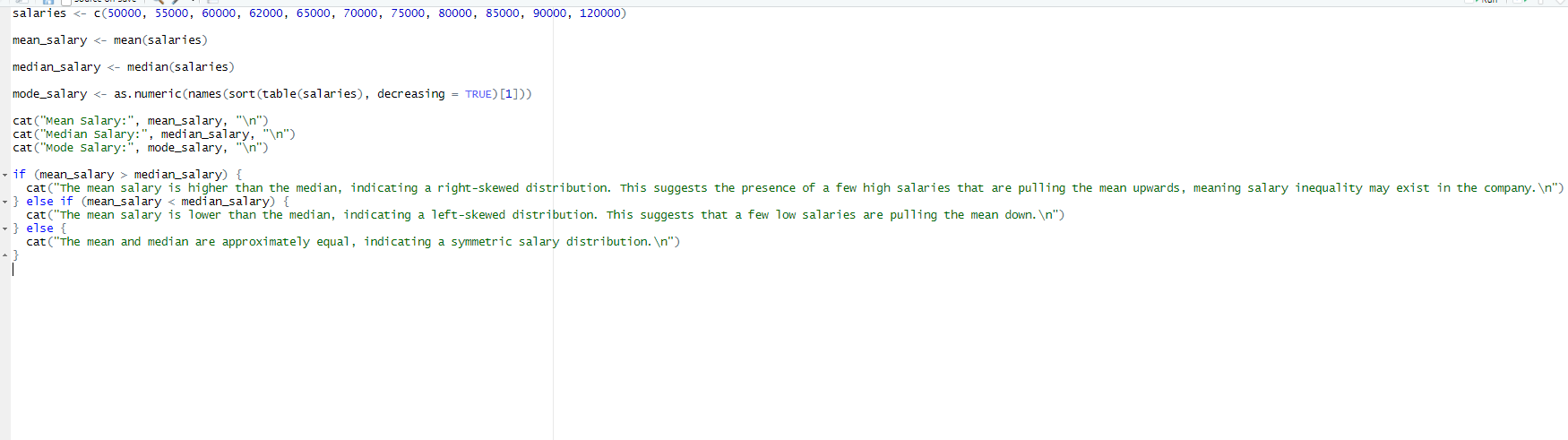
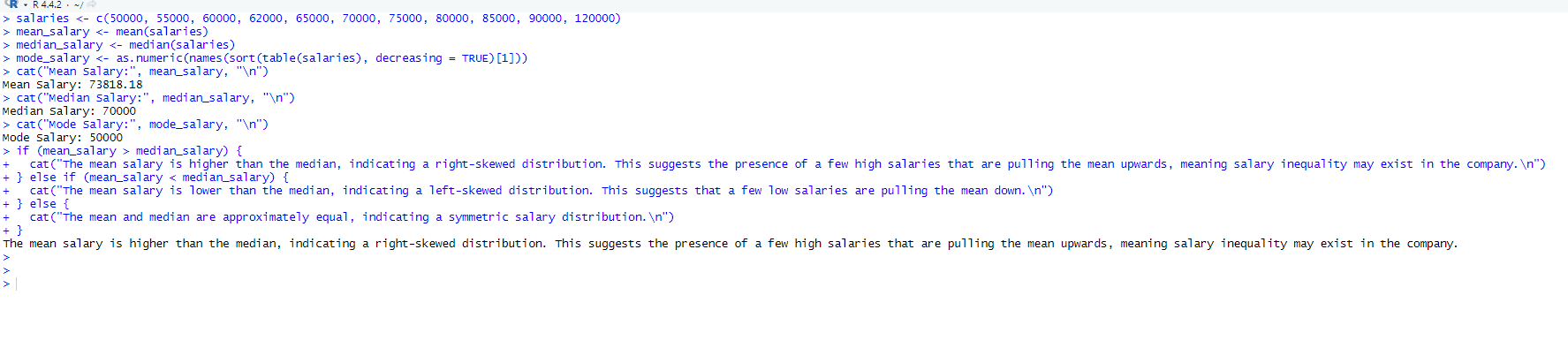
1.A company is analysing the salaries of its employees. Apply your understanding of Univariate EDA to calculate and compare the mean, median, and mode of the salary distribution. Discuss the implications of any disparities between these measures. Write R code to calculate the mean, median, and mode of the salary distribution and interpret the results in the context of employee salaries.

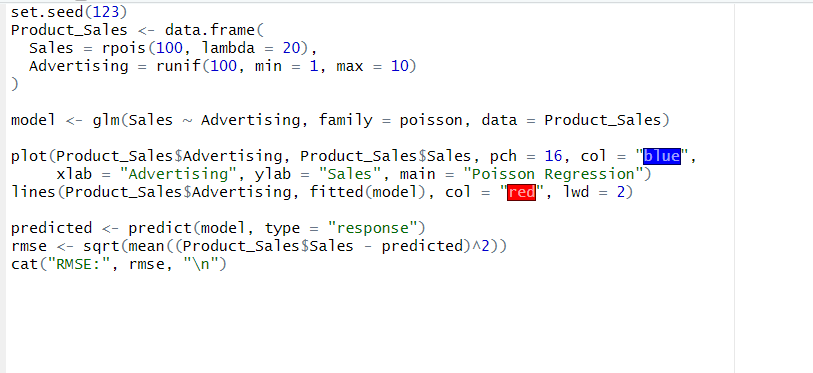
R-CODE:



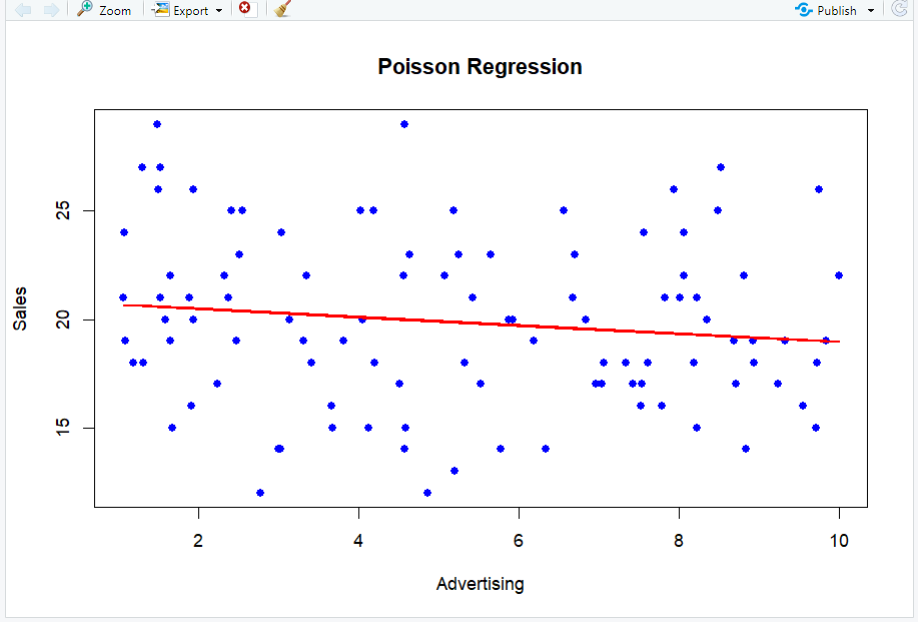
OUTPUT:



2. Create a dataset named “Product\_Sales”. Write R code to perform a poisson regression analysis, create a scatter plot with the regression line, and evaluate the accuracy of the model using appropriate metrics.

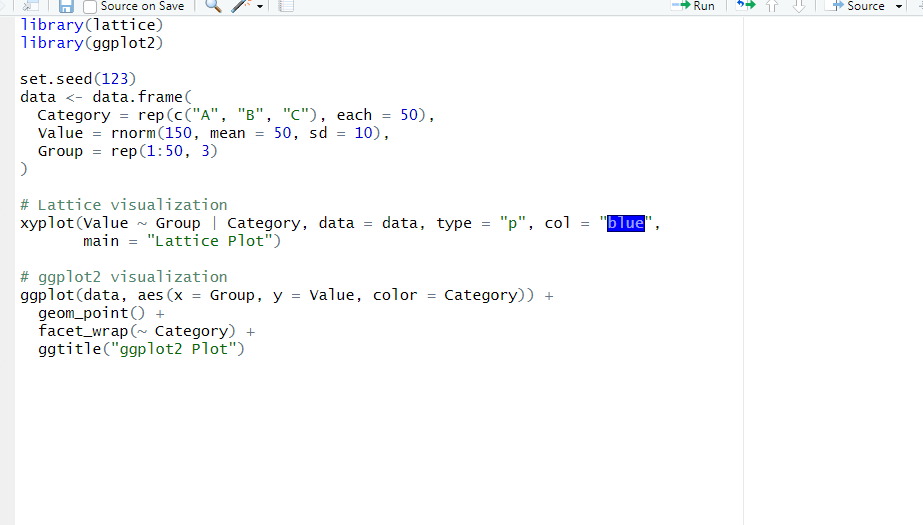


OUTPUT:

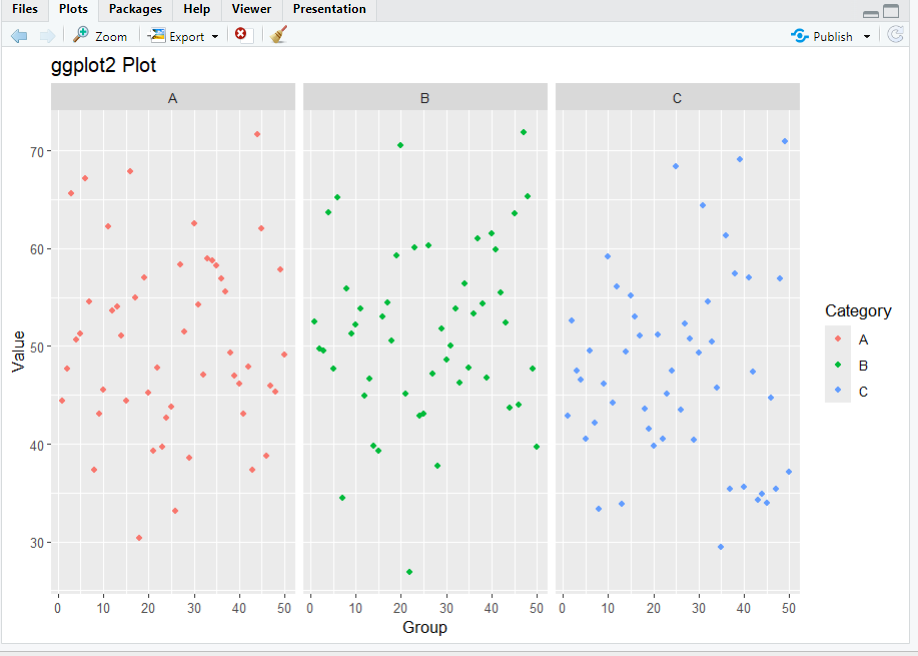


3. In a research project, you are required to create visualizations for multiple variables and subsets of the data. Compare and contrast the use of lattice graphs and ggplot2 for producing advanced graphics. Discuss the strengths and weaknesses of each approach. Develop visualizations using both lattice graphs and ggplot2, and evaluate the effectiveness of each for representing complex relationships in the data.

R-CODE:

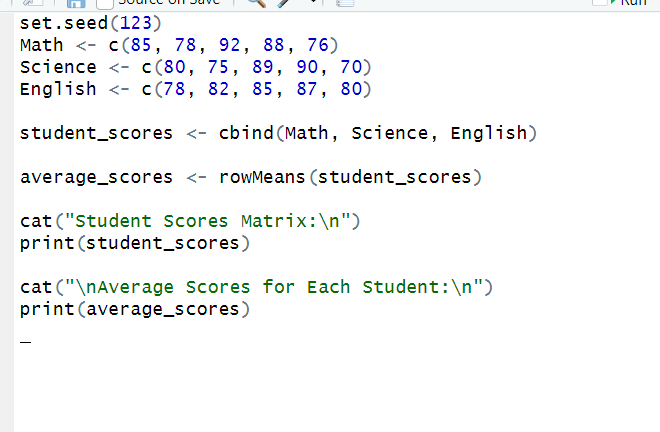


OUTPUT:

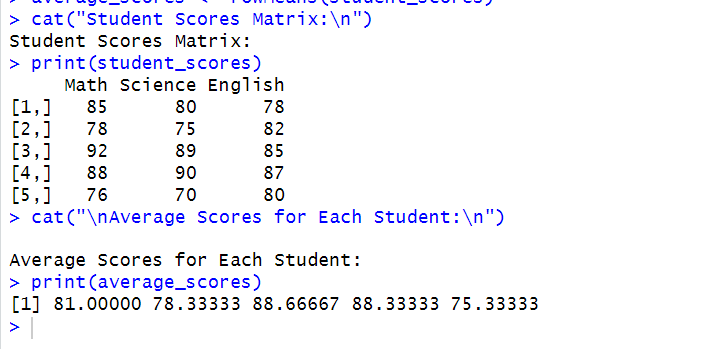


4. Imagine you are working on a project where you need to analyse the performance of students in multiple subjects. Your dataset is currently in the form of individual vectors for each subject. Apply your knowledge of matrices in R to combine these vectors into a matrix named "student scores." Also, calculate and display the average score for each student. Use R code to create the "student scores" matrix and calculate the average score for each student.

R-CODE:

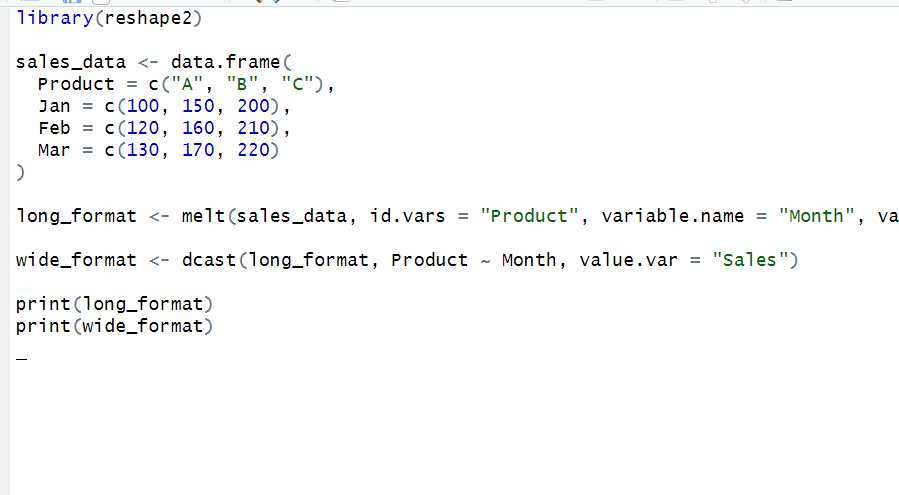


OUTPUT:

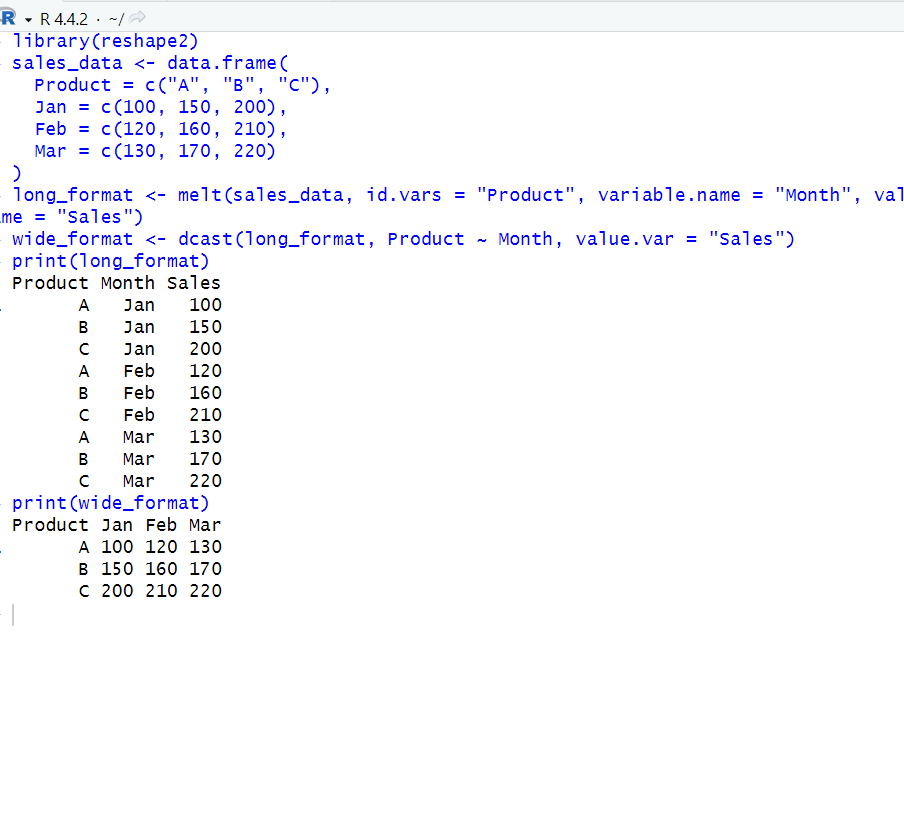


5. You have a wide-format dataset with information about sales for different products and months. Apply the concepts of melting and casting in R to transform the data into a long-format, where each row represents a unique combination of product, month, and sales. Provide R code to perform the melting and casting operations on the dataset, and explain the advantages of the resulting long-format for analysis.

R-CODE:

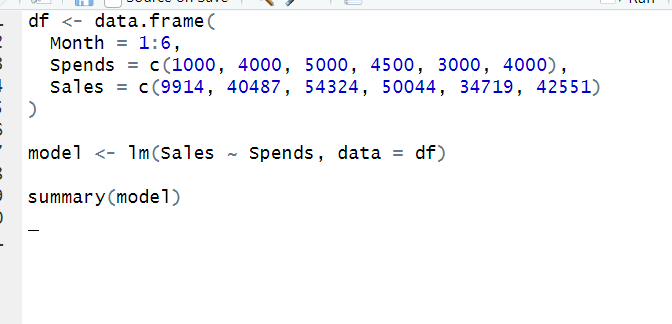


OUTPUT:

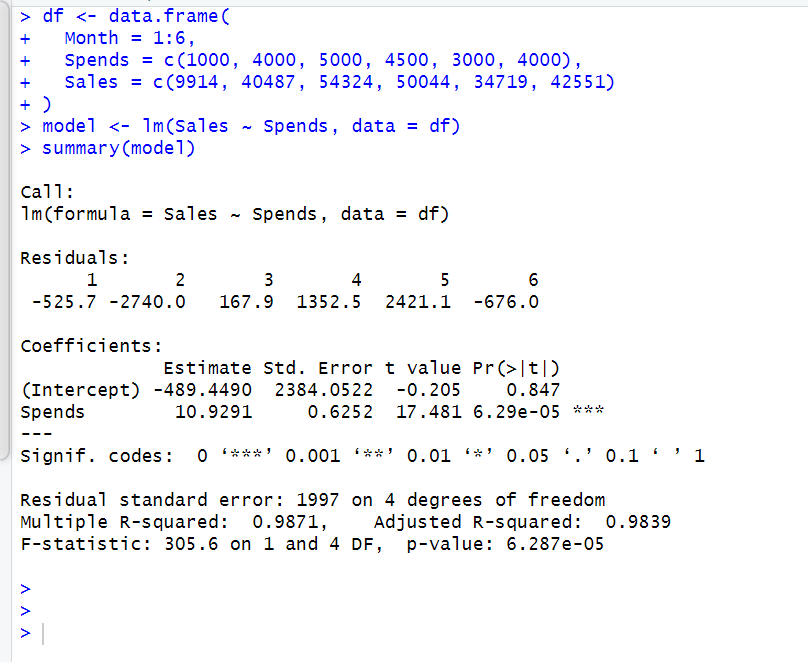


6. a.Create a data frame based on below table. Month 1 2 3 4 5 6 Spends 1000 4000 5000 4500 3000 4000 Sales 9914 40487 54324 50044 34719 42551 b. Create a regression model for that data frame table to show the amount of sales (Sales) based on the how much the company spends (Spends) in advertising.

R-CODE:

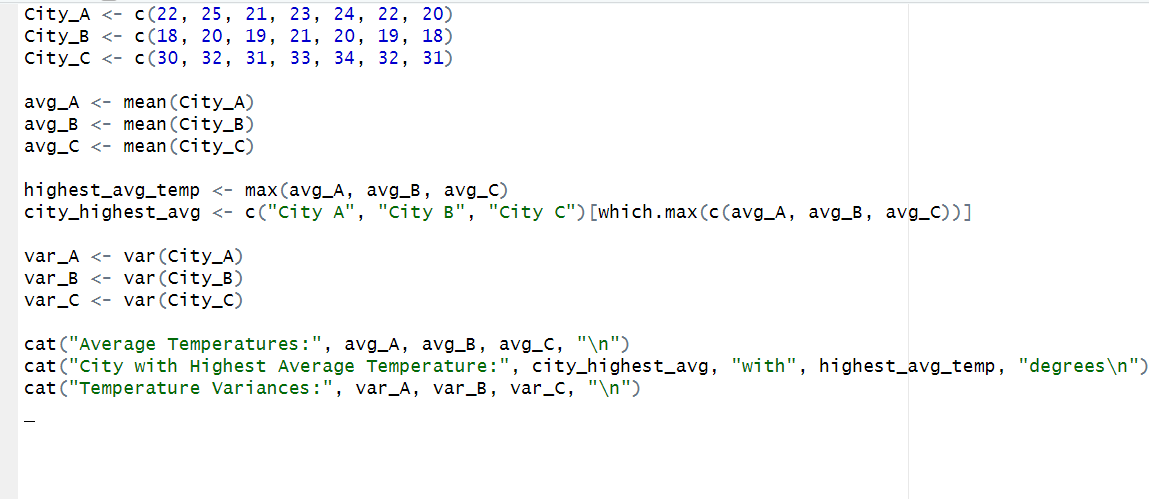


OUTPUT:

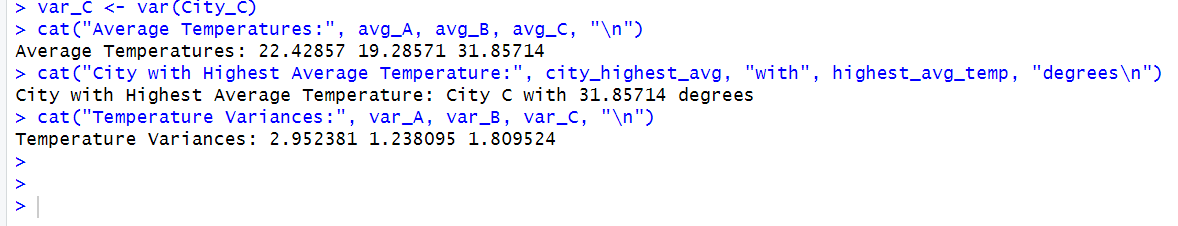


7. Consider an R script designed to analyze temperature data collected from three different cities over a week. The temperature data (in degrees Celsius) are stored in three atomic vectors as follows: - City A: c (22, 25, 21, 23, 24, 22, 20) - City B: c (18, 20, 19, 21, 20, 19, 18) - City C: c (30, 32, 31,33, 34, 32, 31) Using this data, perform the following tasks: i. Calculate the average temperature for each city over the week. ii. Identify the city with the highest average temperature and report its value. iii. Calculate the variance in temperature for each city to assess temperature fluctuation throughout the week.

R-CODE:

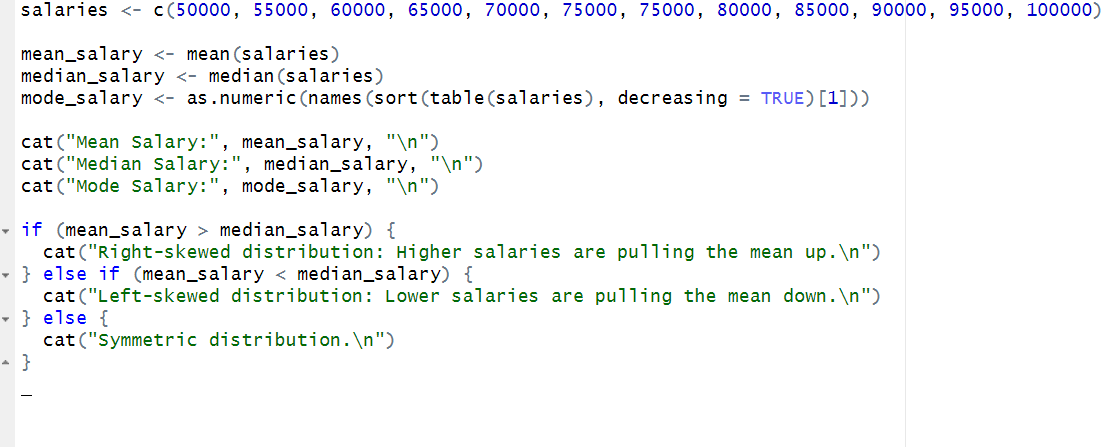


OUTPUT:

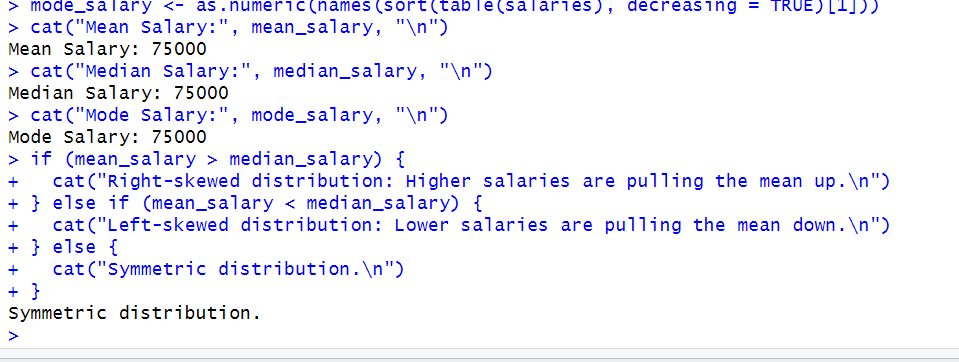


8. A company is analysing the salaries of its employees. Apply your understanding of Univariate EDA to calculate and compare the mean, median, and mode of the salary distribution. Discuss the implications of any disparities between these measures. Write R code to calculate the mean, median, and mode of the salary distribution and interpret the results in the context of employee salaries.

r-code:

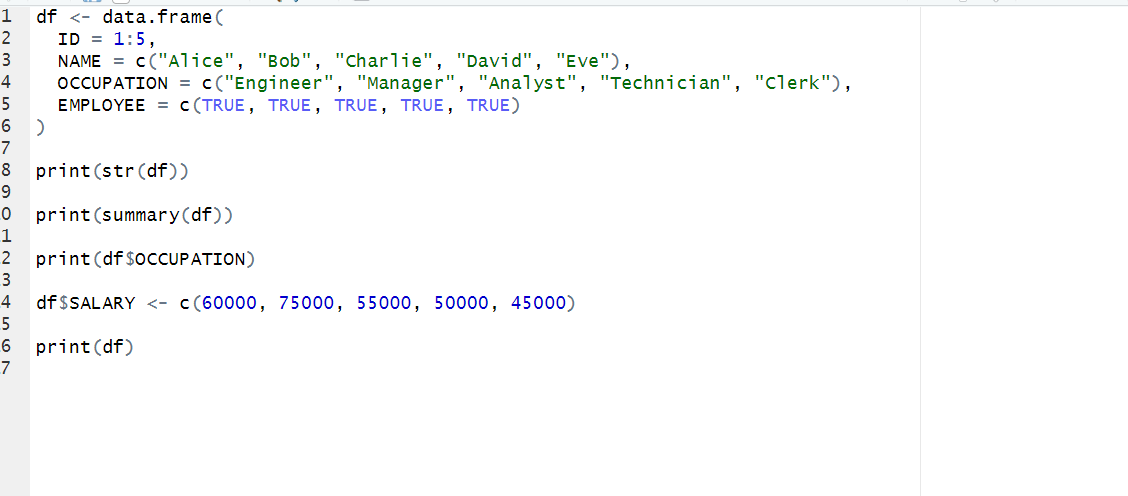


Output:

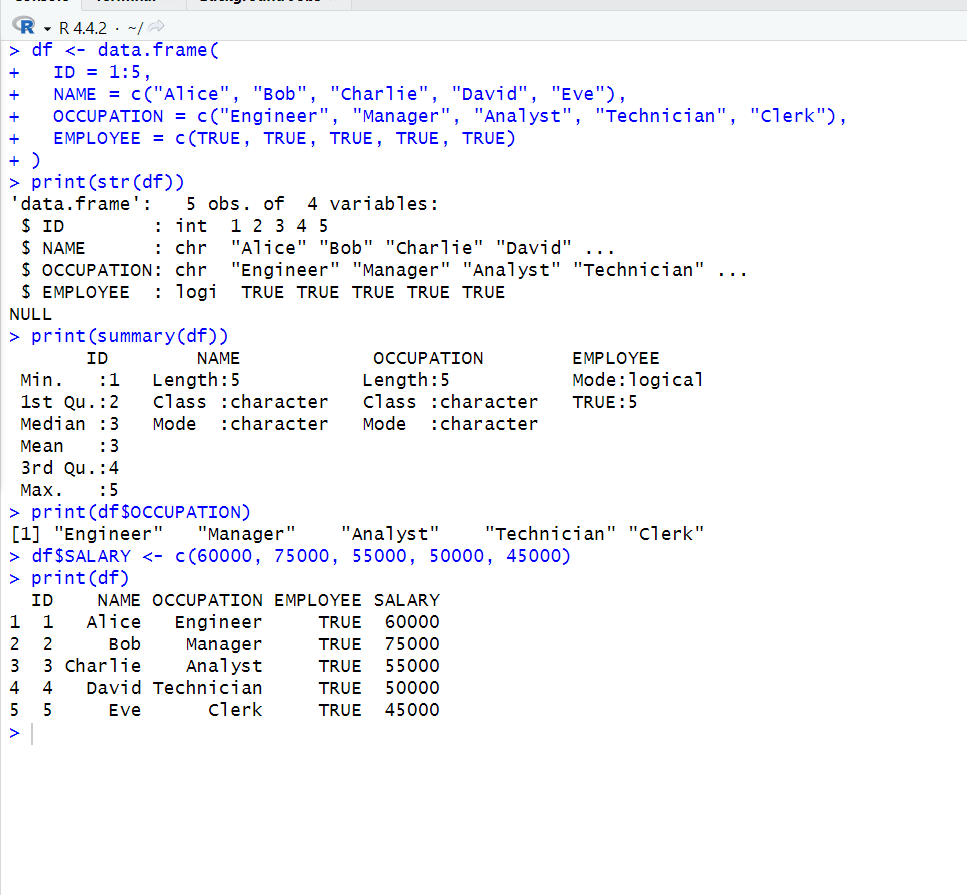


9. Create a data frame and print the structure of the data frame in R. a. Create a data frame of 5 sonar company Employee details: ID, NAME, OCCUPATION, EMPLOYEE b. Apply summary() to find the summary of the data in the data frame and display the results. c. Extract data(OCCUPATION) from the data frame. d. Expand data frame including SALARY.

r-code:



Output:



10. Consider the following data present. Create this file using Windows Notepad.

Save the file as input.csv using the save As All files(\*.\*) option in Notepad.

i.Use appropriate R commands to read the input.csv file.

ii. Analyze the CSV File and compute the following.

a. Get the maximum salary

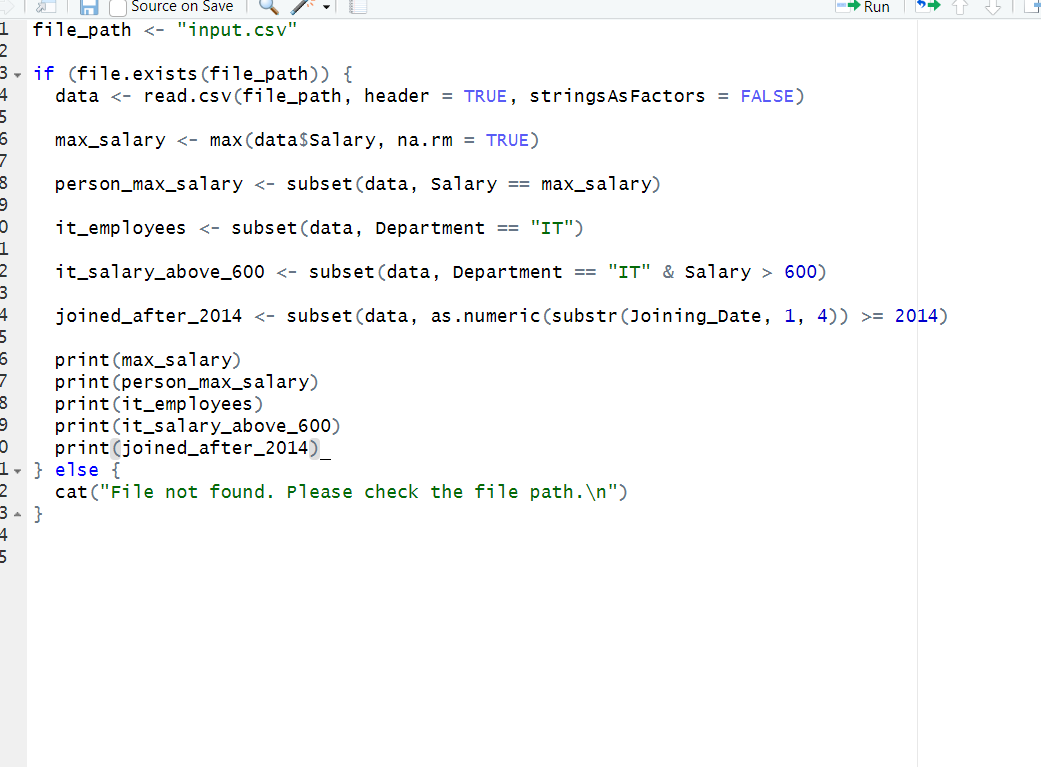
b. Get the details of the person with max salary

c. Get all the people working in IT department

d. Get the persons in IT department whose salary is greater than 600

e. Get the people who joined on or after 2014

r-code:



Output:

