Question (Share R script)

1. You have been provided with a dataset containing information about cancer patients, including their age at diagnosis, tumor size, and treatment history. Your task is to perform the following analyses using this dataset:

   a) Read the dataset from the file into R.

   b) Check the normality of the "Age_at_diagnosis" and "Tumour_size" columns using appropriate statistical tests.

   c) Conduct a t-test to compare the mean age at diagnosis between patients who received chemotherapy and those who did not.

   d) Perform a Wilcoxon signed-rank test to determine if there is a significant difference in tumor size before and after radiotherapy.

   e) Fit a linear regression model to predict tumor size based on age at diagnosis.

   f) Plot the scatter plot of tumor size against age at diagnosis and add the regression line to the plot.

   g) Generate a summary table showing the coefficients and p-values of the regression model.

   h) Check the correlation between tumor size and lymph node status using Pearson correlation.

   i) Perform a chi-square test to investigate the association between chemotherapy and radiotherapy treatment.

   j) Visualize the distribution of survival time for cancer patients with positive and negative lymph nodes