ADVERTISING

GROUP MEMBERS:
ATENA JAFARI PARSA
AVA ARABI
PEYMAN KHODABANDEHLOUEI

INTRODUCTION

The effect of three of the most common (and classic) types of advertisements:

- TV Ads
- Radio Ads
- Newspaper Ads

on companies' sales

DATASET

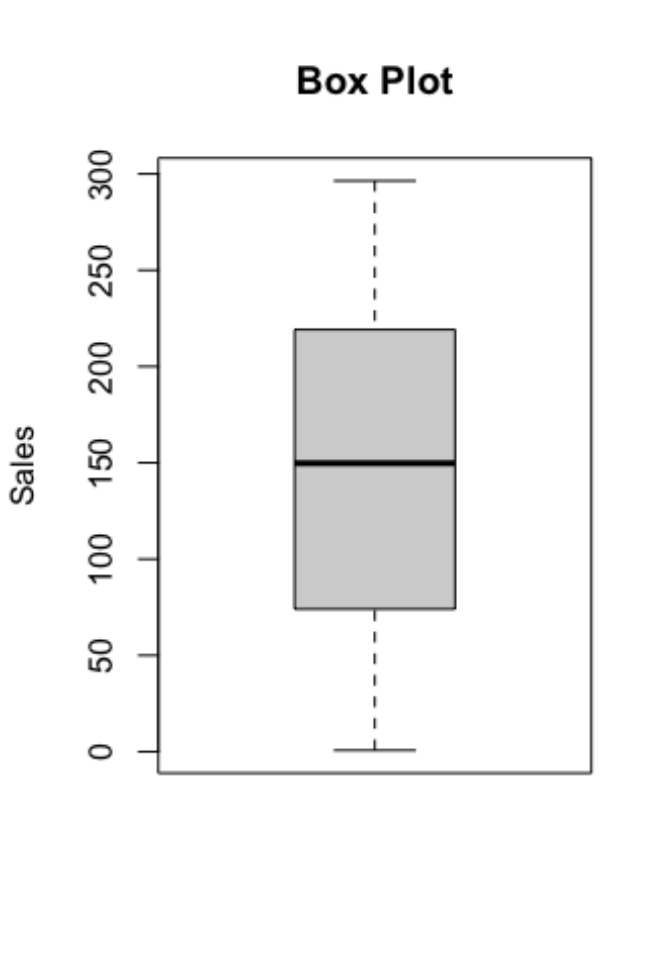
- The dataset chosen is Advertising Dataset of Sales with respect to TV, radio and newspaper advertisements budgets Advertising Dataset | Kaggle .
- The sales are in thousands of units and the budget is in thousands of dollars.

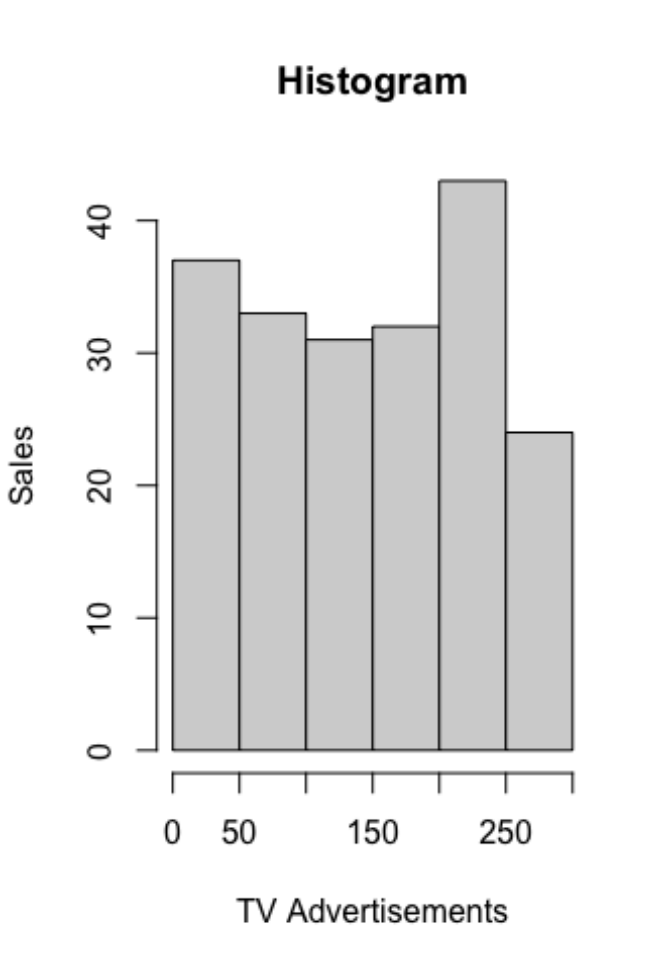
a several data points of each variable

```
> str(advertising_data)
'data.frame':   200 obs. of  4 variables:
 $ TV       : num  230.1 44.5 17.2 151.5 180.8 ...
 $ Radio    : num  37.8 39.3 45.9 41.3 10.8 48.9 32.8 19.6 2.1 2.6 ...
 $ Newspaper: num  69.2 45.1 69.3 58.5 58.4 75 23.5 11.6 1 21.2 ...
 $ Sales    : num  22.1 10.4 12 16.5 17.9 7.2 11.8 13.2 4.8 15.6 ...
```

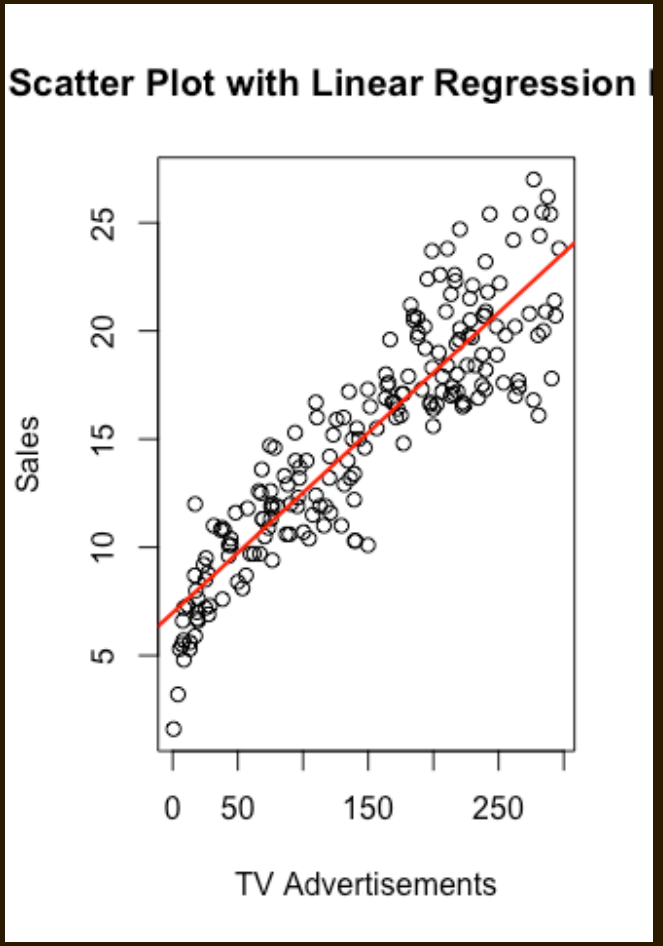
The data retrieved from:
https://www.kaggle.com/datasets/ashydv/advertising-dataset

Data Analysis

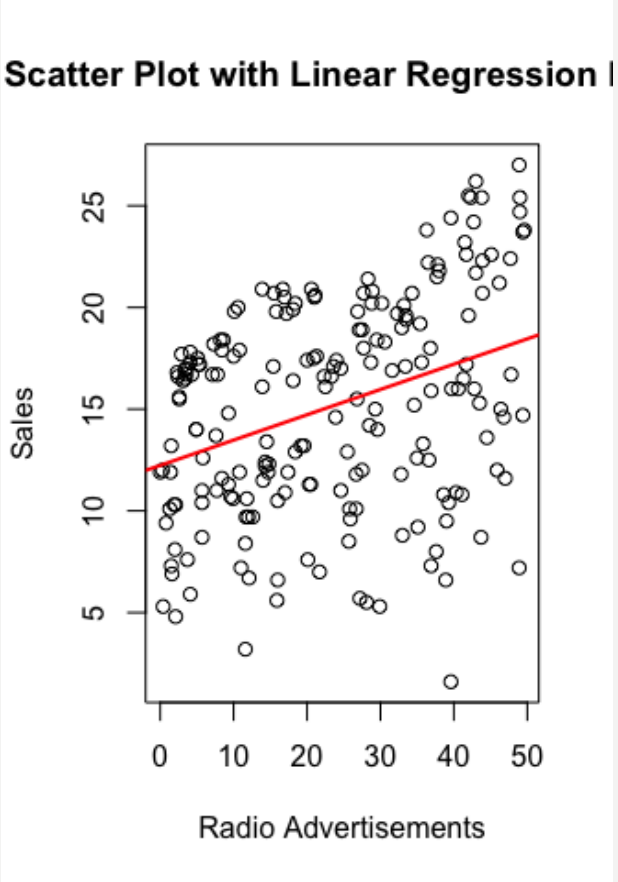


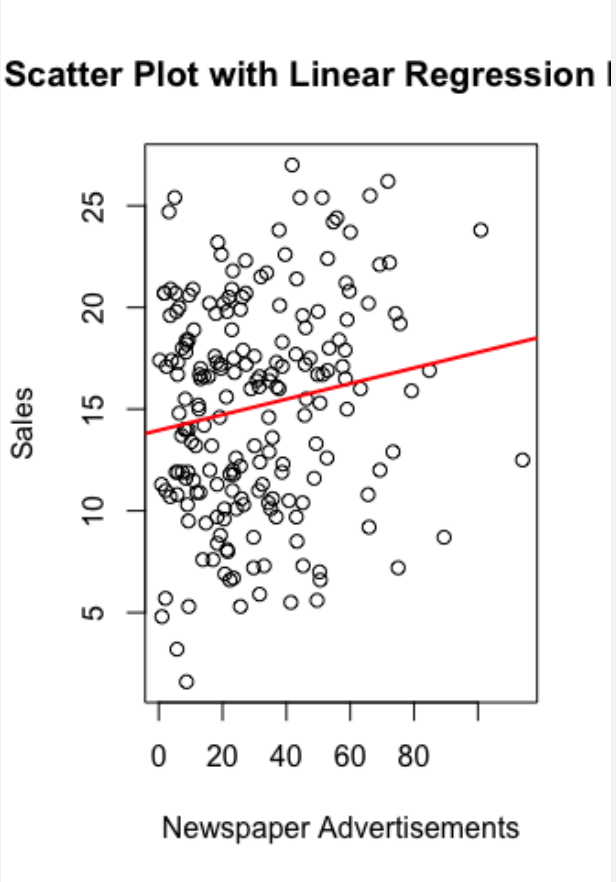


TV ADVERTISEMENTS BUDGET VS SALES



Strong Positive Linear Relationship



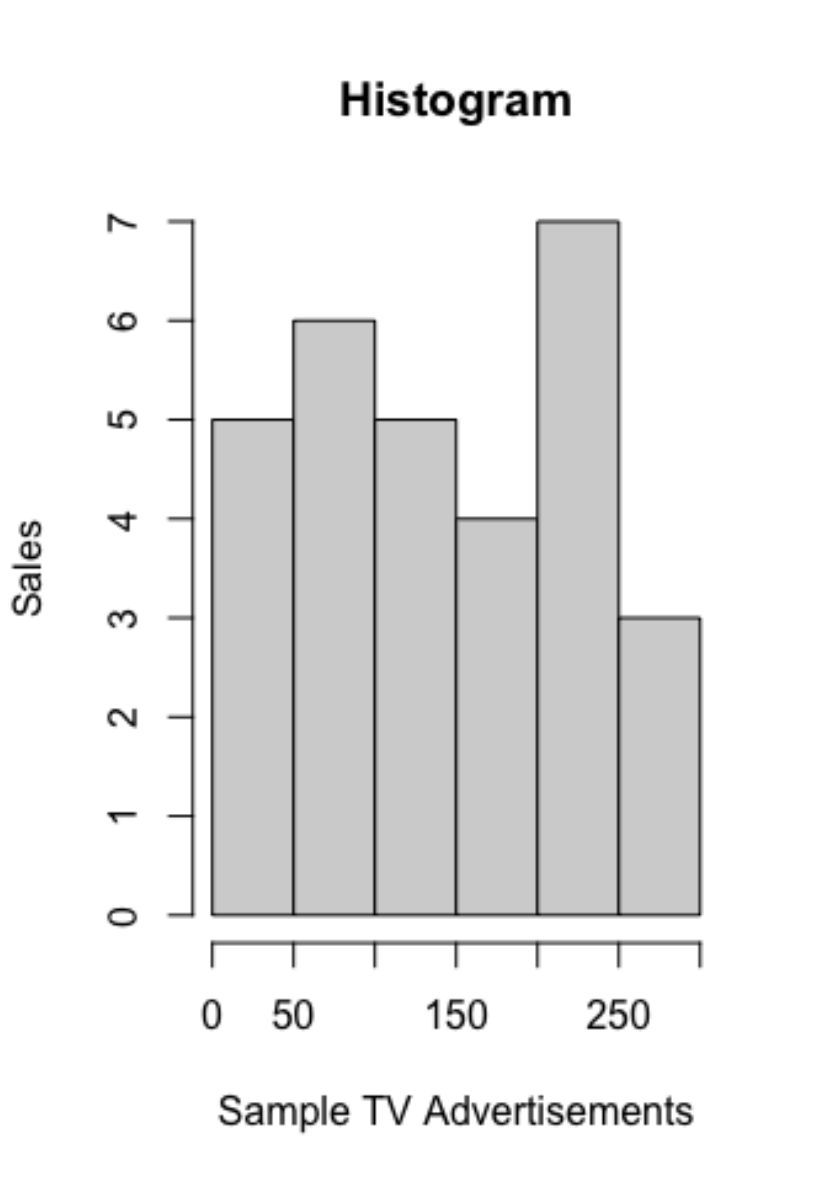


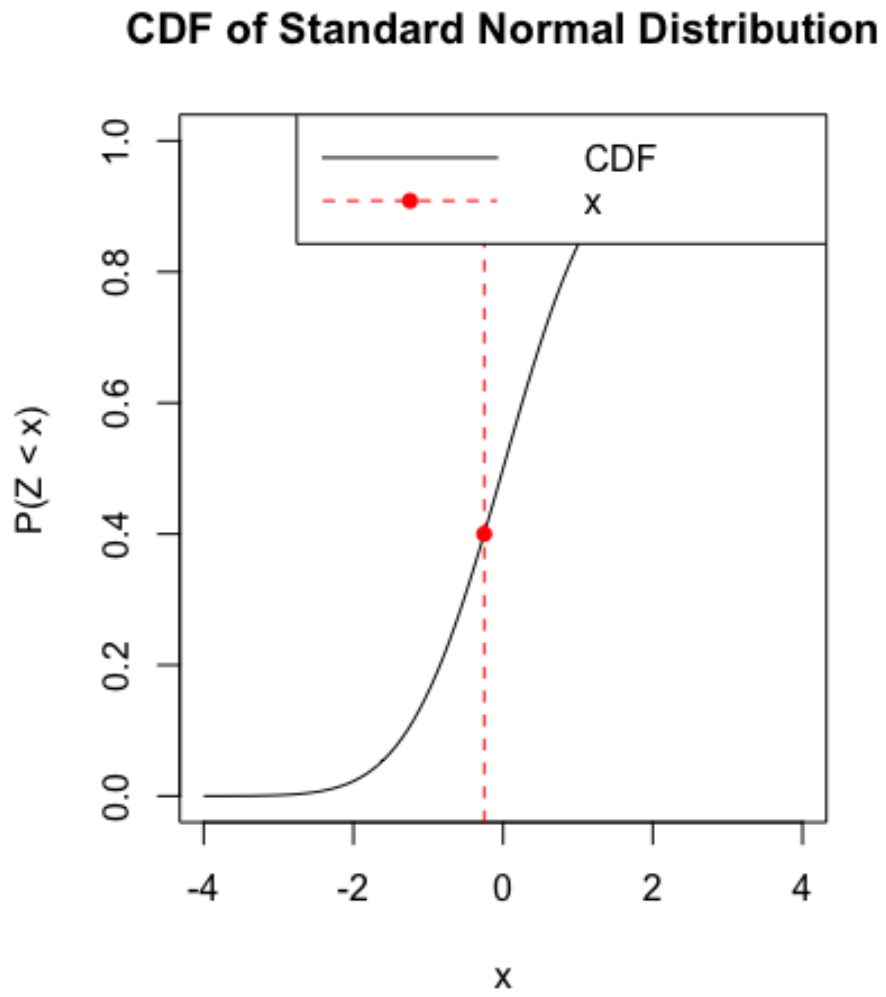
RADIO AND NEWSPAPER ADVERTISEMENTS BUDGET VS SALES

No clear relationship

SAMPLE STATISTICS

- A random sample of size 30
- In both the sample and the population: The highest sales' corresponding TV ads budget of $200K-$250K budget of TV advertisements
- The lowest sales correspond to TV ads budget larger than $250K.
- The other bins relatively show a pattern similar to the population histogram as well.





NORMALITY

The Shapiro-Wilk test proved normal distribution for sales population.

Find x as the value of sales for which the z-value of the data points below it (less than x) would make up 0.4 of the data.

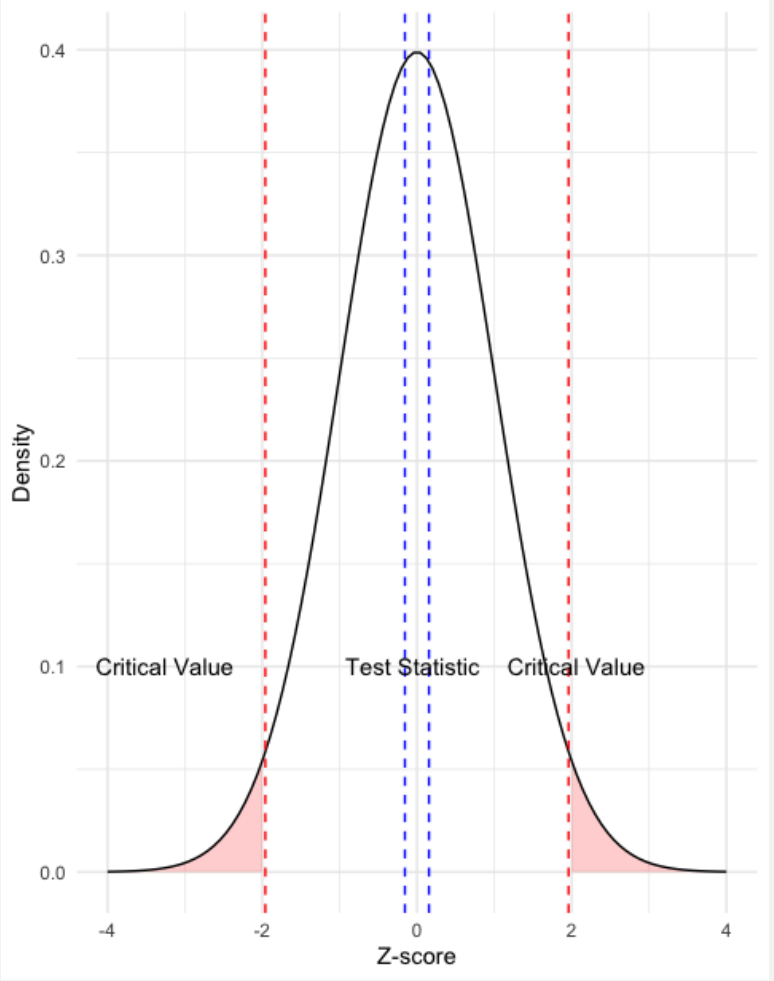
$P(Z<x) = 0.4$

HYPOTHESIS TEST 3 CONFIDENCE INTERVALS

- Claim: The population mean of sales is not equal to 15.13. (alpha = 0.05, n = 30)
- H0: mu = 15.13
- Ha: mu != 15.13
- n = 30 -> CLT -> normally distributed
- qnorm() and pnorm() functions

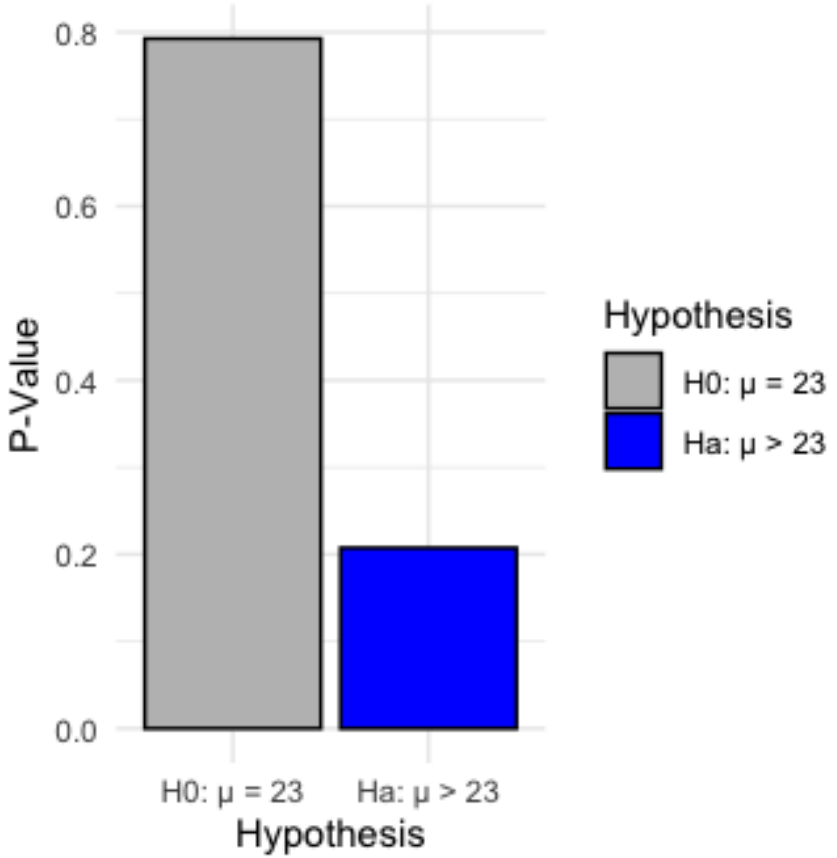
p-value: 0.8764363

Conclusion: Not enough evidence to reject H0



HYPOTHESIS TEST 2

- Claim: The population mean of Radio is greater than 23. (alpha = 0.05, n = 20)
- H0: mu = 23
- Ha: mu > 23
- t.test() function
- P-value = 0.2074528

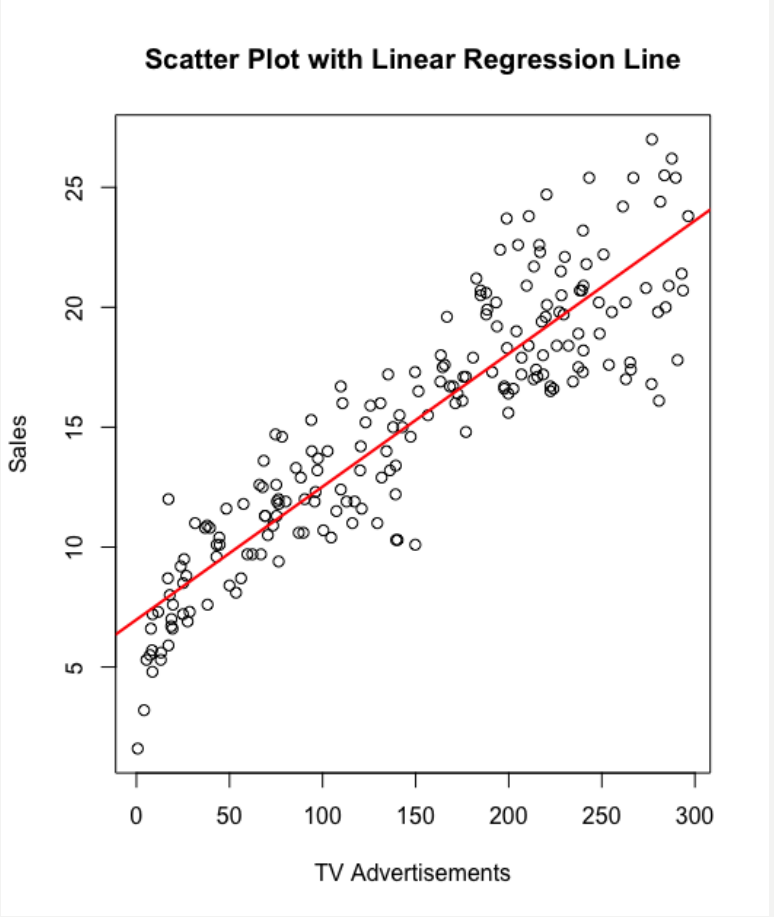


DETECTING THE DISTRIBUTION

- Goodness of fit tests
- The Chi-squared test to see whether the TV Budget population is normally distributed. (alpha = 0.05)
- H0: The distribution is normal.
- Ha: The distribution is not normal.
- Chisq.test() function
- P-value < 2.2e-16 < alpha
- Reject H0 in favor of Ha
- TV Budget population is not normally distributed.

LINEAR REGRESSION

- TV: $230.1K, Radio: $37.8K, Newspaper: $69.2K. The predicted sale?
- Fit the model.
- predict()
- Residual error?
- Observed – predicted = | 22.1 – 21.2 | = 0.9



ANOVA

- Claim: There is a significant difference in the mean values of the three advertising methods.(alpha = 0.05)
- H0: There is no significant difference in the means.
- Ha: There is a significant difference in the means.
- library(“dplyr”)
- aov() and summary()
- P-value < alpha
- Conclusion: H0 is rejected in favor of Ha.
- There is a significant difference in the mean values of the three advertising methods.

APPLICATIONS OF NONPARAMETRIC TESTS

- Single Sample Sign Test: Is there a significant difference in the sample median from the hypothesized value? (sample sales)
- H0: There is not a significant difference in the sample median from the hypothesized value.
- Ha: There is a significant difference in the sample median from the hypothesized value.
- P-value < alpha
- Conclusion: Reject H0 in favor of Ha.
- There is a significant difference in the sample median from the hypothesized value

THANK YOU FOR LISTENING!