

301033 Introduction to Data Science

Week 3 Lab Exercises

1. Multiple Linear Regression: Use Advertising data set

- Find the Covariance and Correlation Matrix of Sales, TV, Radio and Newspaper.
- Construct the multiple linear regression model and find the least square estimates of the model parameters.
- Test the significance of the parameters and find the resulting model to model Sales in terms of advertising modes, TV, Radio and Newspaper
- Assess the overall accuracy of the model.
- Calculate the predicted values and residuals
- Plot the residuals against the predicted values
- Plot the histogram of the residuals
- Comment on the residual plots
- Use the multivariate model for prediction

2. Non Linear Models: Use Advertising data set

- Add the Interaction Term $TV \times Radio$ and test the significance of the interaction term
- Give the resulting model after considering this interaction term.
- Construct the Polynomial Regression Model of order 3 and test the model significance
- Give the resulting selected model

3. If Time Permits repeat above analysis for Auto Data Set and upload via Turnitin Link located in vUWS Learning Material page.