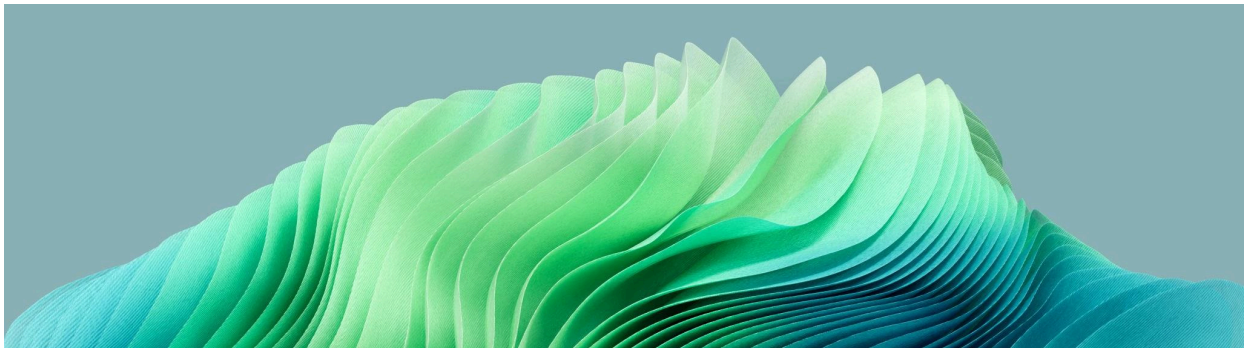

Chapter 3 : Linear Regression

1.Describe the null hypotheses to which the p-values givenin Table3.4 correspond. Explain what conclusions you can draw based on these p-values.Your explanation should be phrased in terms of sales,TV, radio,and newspaper, rather than in terms of the coefficients of the linear model.

	Coefficient	Std. error	t-statistic	p-value
Intercept	2.939	0.3119	9.42	< 0.0001
TV	0.046	0.0014	32.81	< 0.0001
radio	0.189	0.0086	21.89	< 0.0001
newspaper	−0.001	0.0059	−0.18	0.8599

TABLE 3.4. *For the* Advertising *data, least squares coefficient estimates of the multiple linear regression of number of units sold on TV, radio, and newspaper advertising budgets.*

As we see above in table 3.4, the p values for all the predictors except newspaper is <0.05 making the values obtained by them , (Coefficient, Std.Error , t-stat) statistically significant, Whereas for newspaper , p value>0.05 , meaning there is insufficient evidence to conclude that a relationship between the predictor 'newspaper' and the response variable sales exists. A p value dictates that how much of the result one obtained from an experiment was by chance or random error , this difference arises as the sample and population datasets are different , so there can be more than 1 sample for a population and all the different samples can predict different outcomes. Hence having a lesser p value(<0.05) means that our null hypothesis of the predictor and response variable having no relationship among them (beta=0){coefficients in this case} is false and there is in fact a relationship among them which is not by chance.

2. Carefully explain the differences between the KNN classifier and KNN regression methods.

The major difference lies between what kind of **problem is being solved or Our Goal** , if the response variable is qualitative then we use a classifier , if it is quantitative then we use regression. | Both the methods are used for prediction and both are Non Parametric as well
A lower K value corresponds to a Model that has low bias but high variance due to it being highly dependent on just one dataset whereas
A high K value gives a smoother fit with high bias but low variance as now the predicted value is dependent on more than one data point , so even if one changes the other data points stabilizes it.
Hence a Propper K value is chosen via Bias Variance tradeOff comparing Test MSE values for regression and error/accuracy test for Classification of different models with different K values

	KNN Classifier	**KNN Regression**
Output Produced	**Qualitative/ Categorical**	**Quantitative / numerical**
How is K used?	The nearest K data points are chosen and a majority rule is applied by default , if the differing variables are the same, then a tie breaker rule is applied.	The average of the closest K datapoints are taken and the predicted value is that average of the nearest K values.