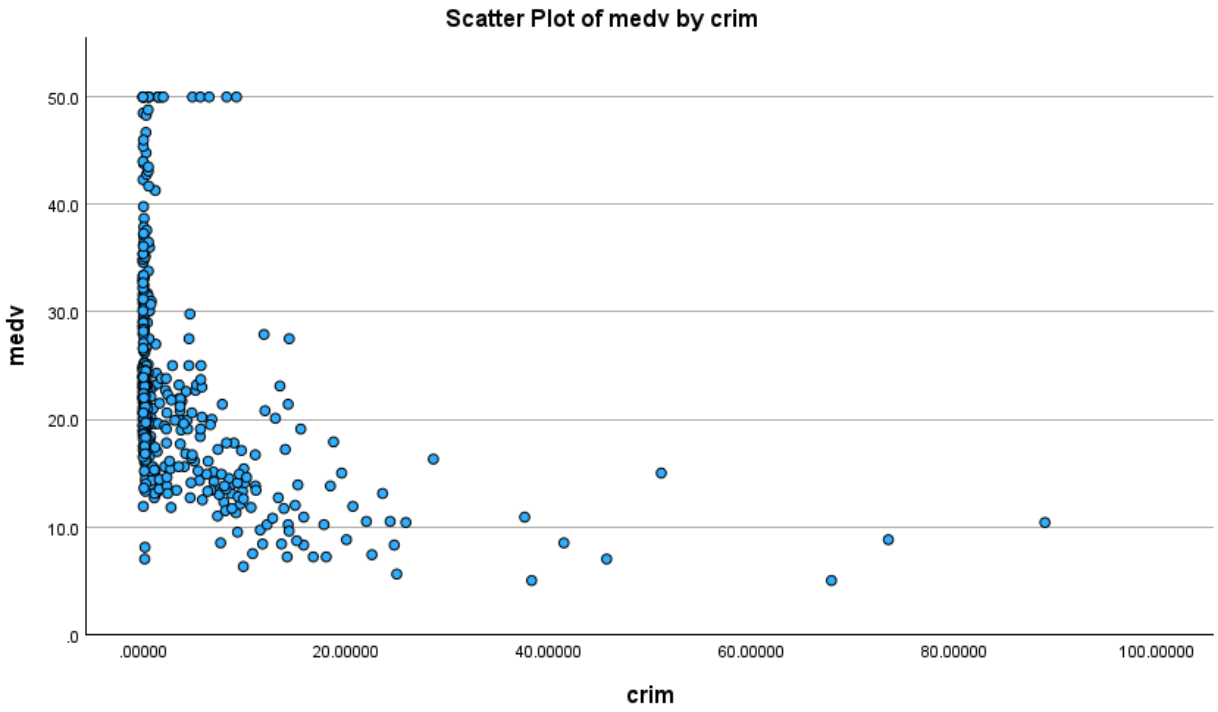
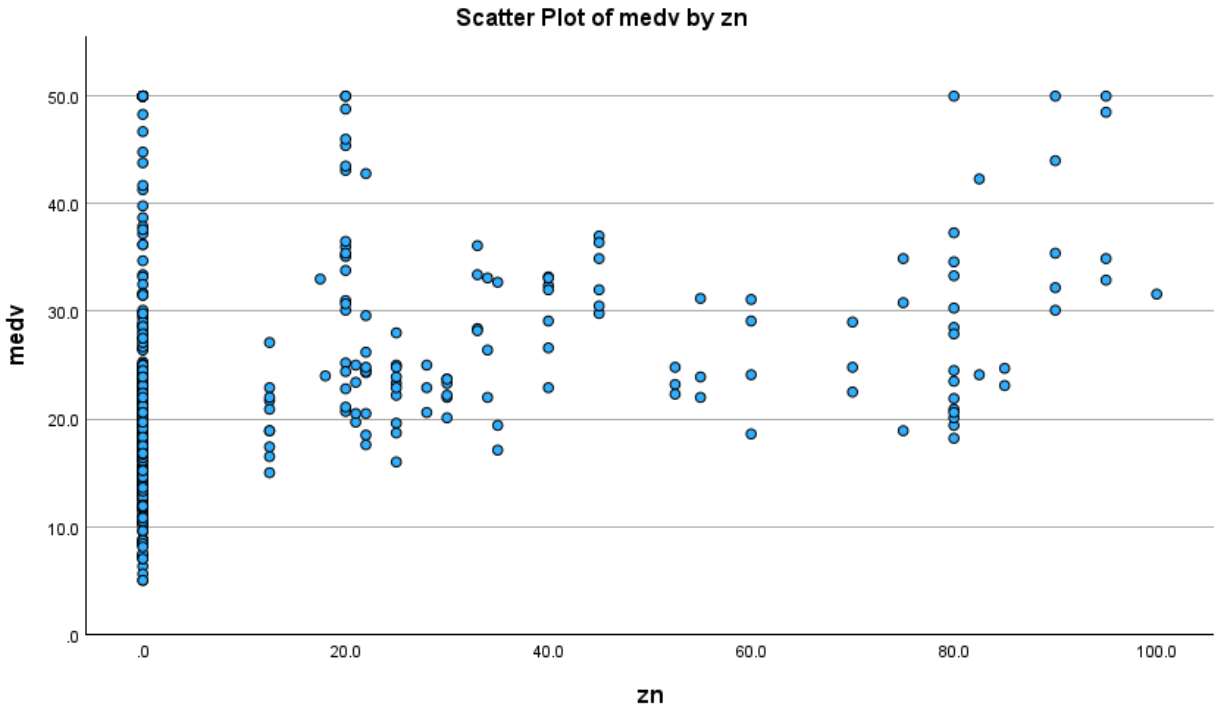
GGraph

[DataSet4]



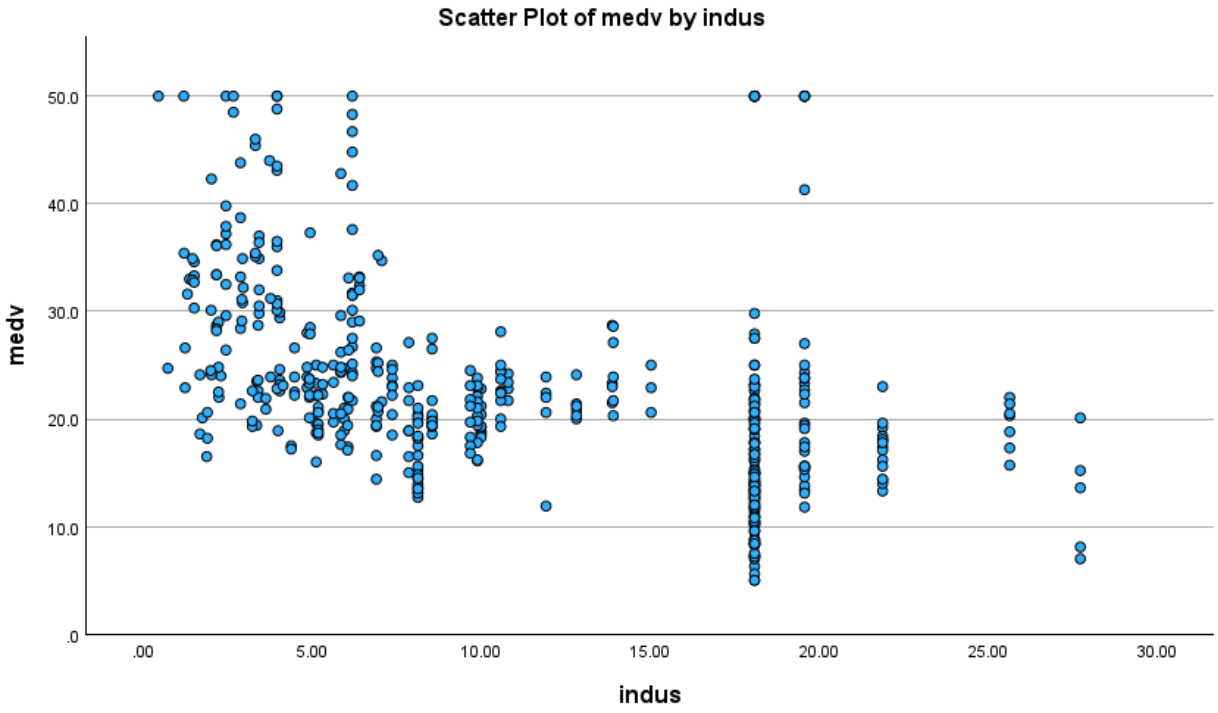
This indicates a moderate positive correlation between the independent variable crim (crime rate) and the dependent variable medv.

GGraph



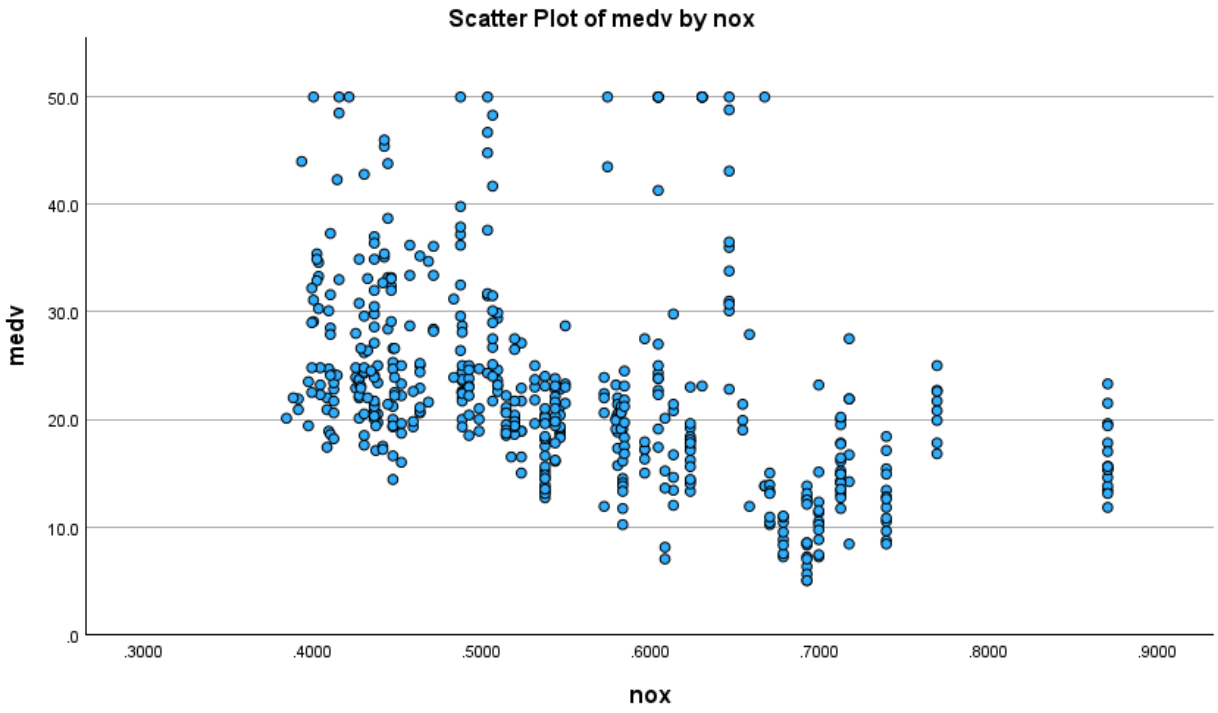
This indicates a moderate positive correlation between the independent variable zn (proportion of residential land zoned for lots over 25,000 sq. ft.) and the dependent variable medv.

GGraph



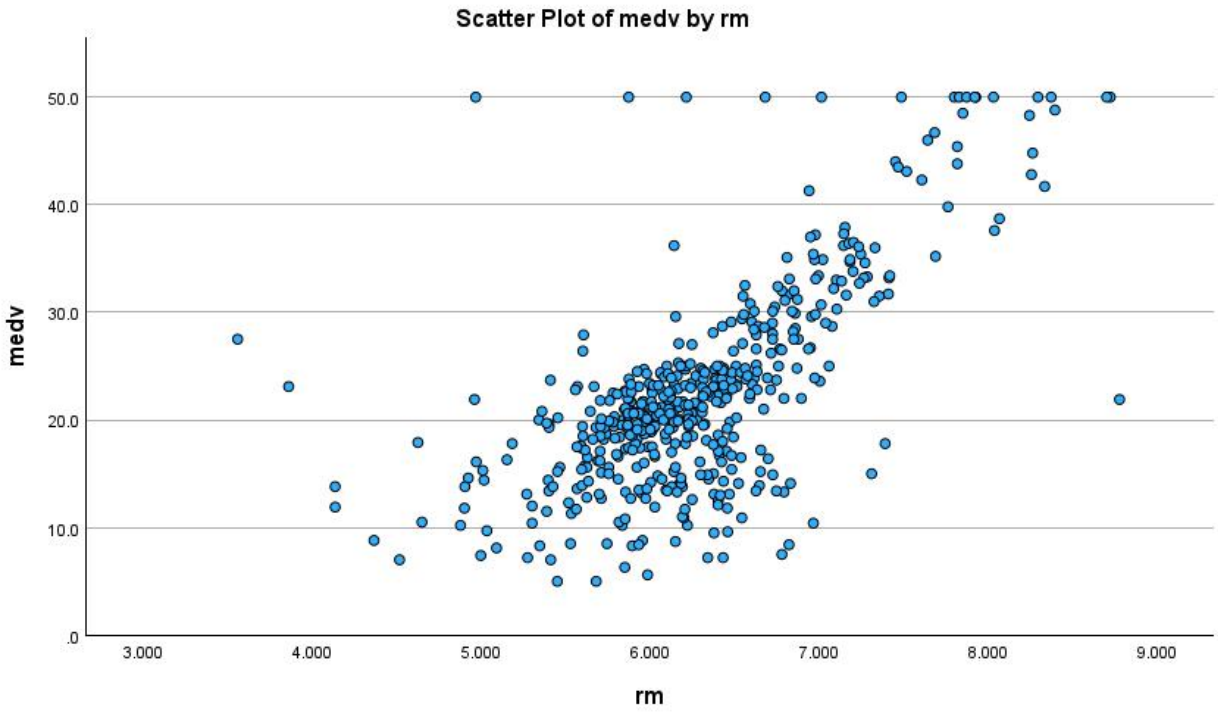
This indicates a moderate positive correlation between the independent variable indus (proportion of non-retail business acres per town) and the dependent variable medv (median home value).

GGraph



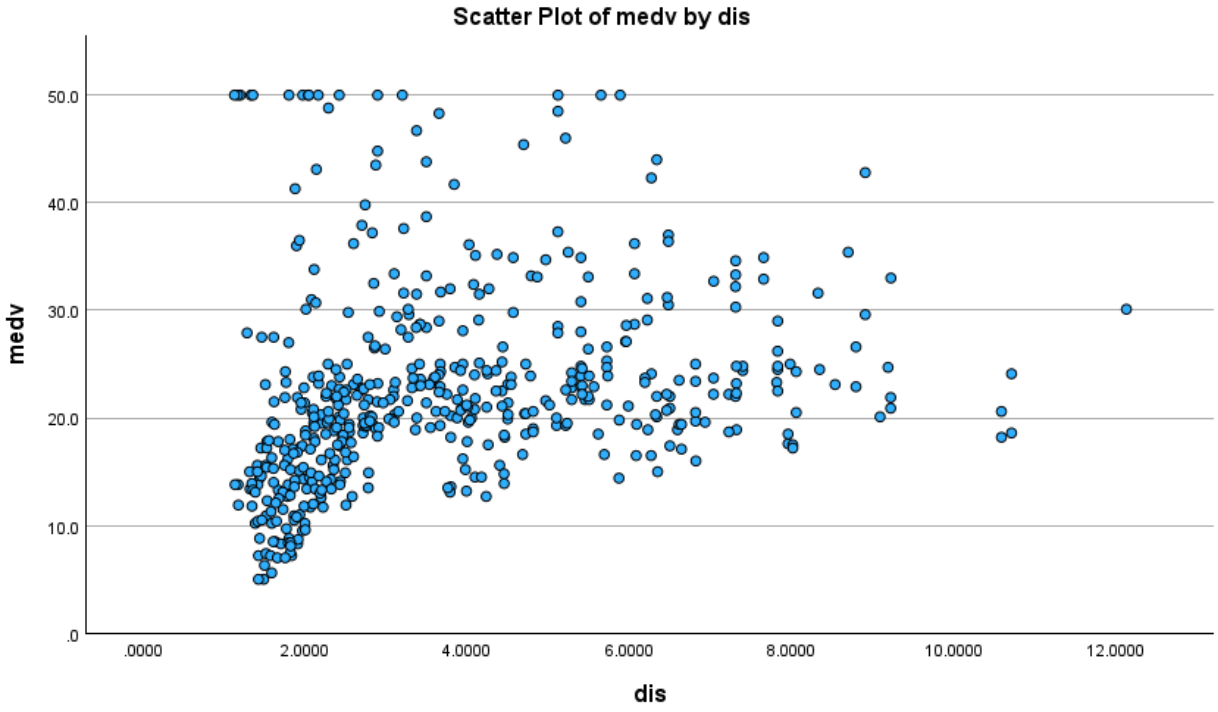
This indicates moderate positive correlation between the independent variable nox (nitric oxides concentration) and the dependent variable medv (median home value).

GGraph



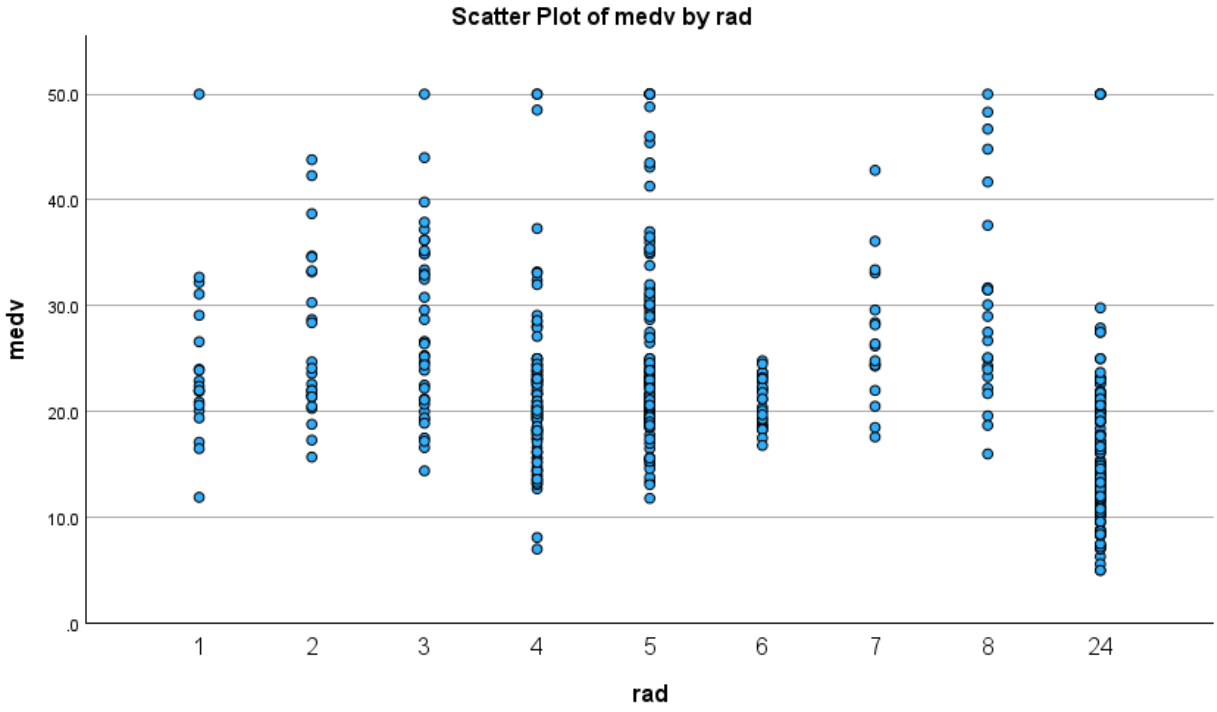
This indicates a strong positive correlation between the independent variable rm (average number of rooms per dwelling) and the dependent variable medv (median home value).

GGraph



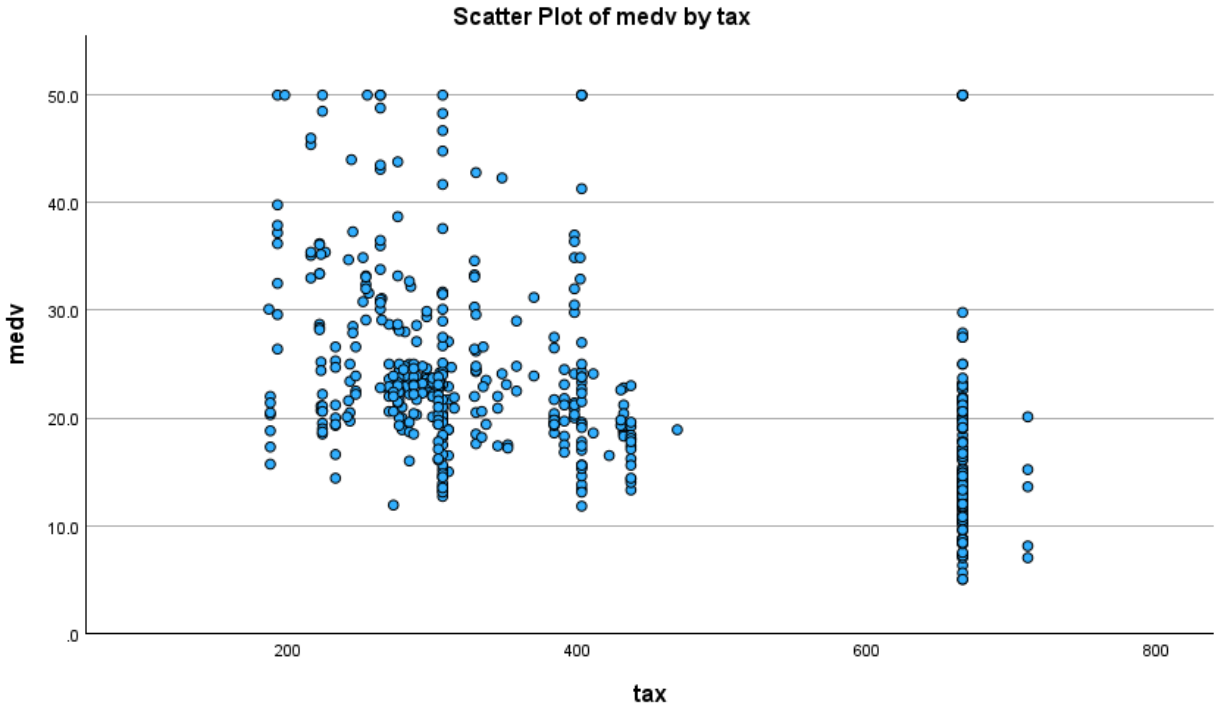
This indicates a weak positive correlation between the independent variable dis (weighted distances to five Boston employment centers) and the dependent variable medv (median home value).

GGraph



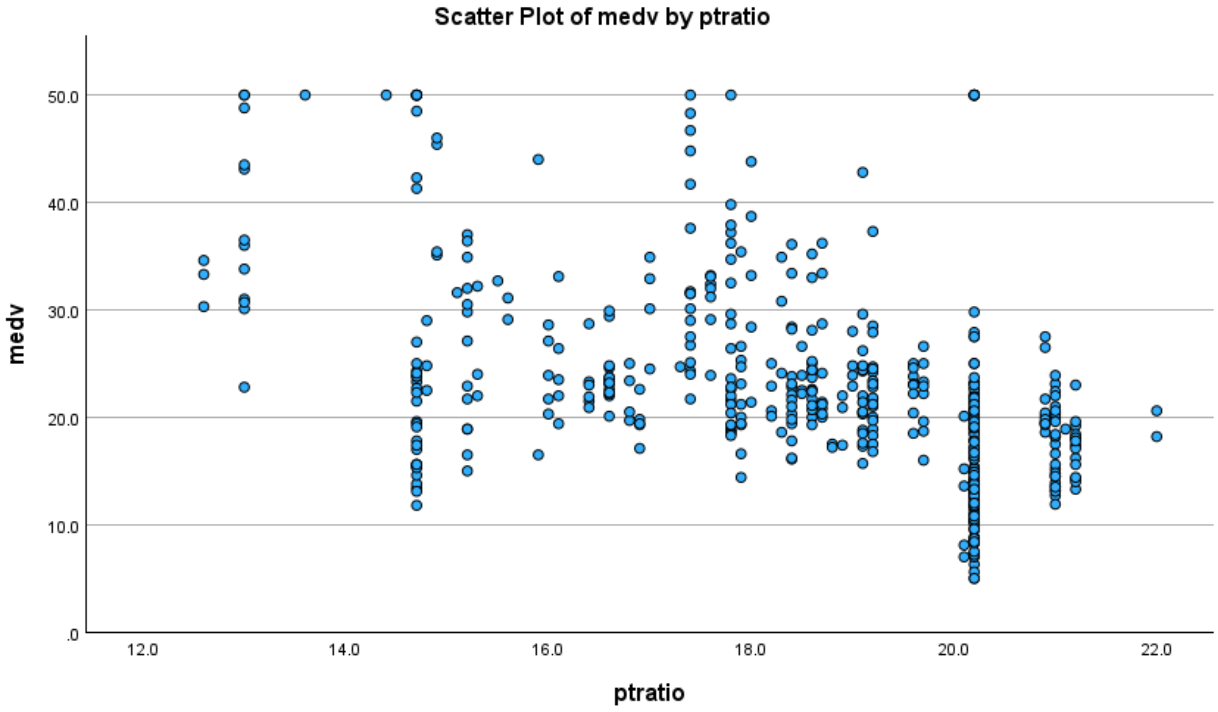
This indicates a moderate positive correlation between the independent variable rad (index of accessibility to radial highways) and the dependent variable medv (median home value).

GGraph



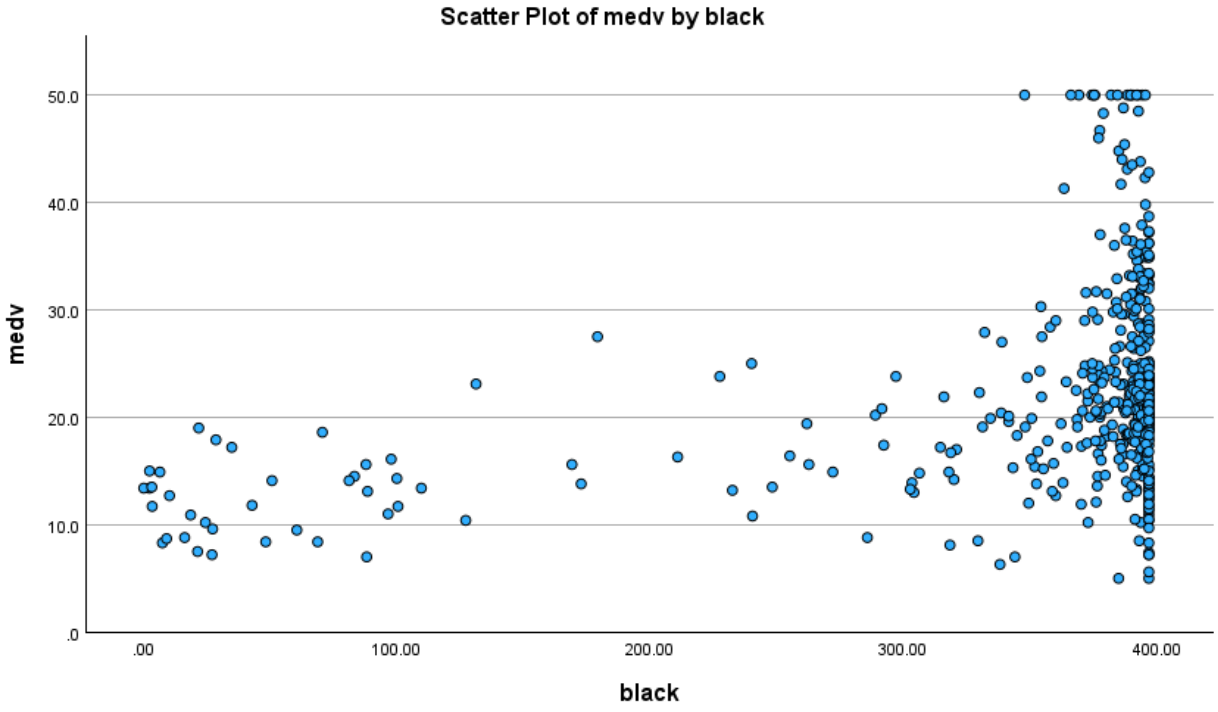
This indicates a moderate positive correlation between the independent variable tax and the dependent variable medv (median home value).

GGraph



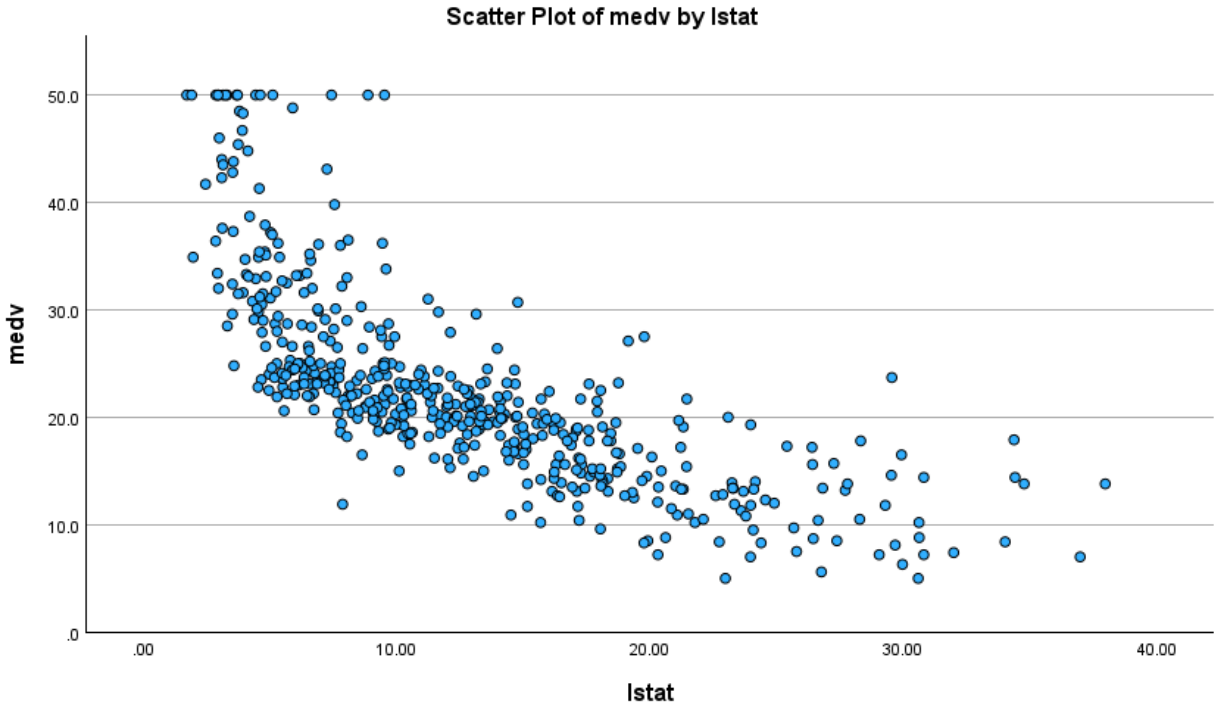
This indicates a moderate positive correlation between the independent variable ptratio (pupil teacher ratio by town) and the dependent variable medv (median home value).

GGraph



This indicates a weak to moderate positive correlation between the independent variable black (proportion of Black residents by town) and the dependent variable medv (median home value).

GGraph



This indicates a strong positive correlation between the independent variable lstat (percentage of lower-status residents) and the dependent variable medv (median home value).

Regression

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	lstat, chas, black, ptratio, zn, crim, rm, indus, age, rad, dis, nox, tax[b]	.	Enter

a. Dependent Variable: medv

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.861[a]	.741	.734	4.7453	.741	108.077	13	492	<.001

a. Predictors: (Constant), lstat, chas, black, ptratio, zn, crim, rm, indus, age, rad, dis, nox, tax

1. R (Multiple Correlation Coefficient) = .861

• The R value of 0.861 indicates a strong positive correlation between the observed values of the dependent variable and the values predicted by the model.

2. R Square (Coefficient of Determination) = .741
• The R Square value of 0.741 means that approximately 74.1% of the variability in the dependent variable can be explained by the independent variables included in the model.

3. Adjusted R Square = .734
• The Adjusted R Square value of 0.734 is slightly lower than the R Square value, which is expected when more variables are included in the model. It suggests that even when
adjusting for the number of predictors, about 73.4% of the variance in the dependent variable is still explained by the model.

4. Standard Error of the Estimate = 4.7453
• The Standard Error of the Estimate (4.7453) measures the average distance that the observed values fall from the regression line. In this context, a standard error of 4.7453
suggests that, on average, the predicted values deviate from the actual values by about 4.75units.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	31637.511	13	2433.655	108.077	<.001[b]
	Residual	11078.785	492	22.518		
	Total	42716.295	505			

a. Dependent Variable: medv

b. Predictors: (Constant), lstat, chas, black, ptratio, zn, crim, rm, indus, age, rad, dis, nox, tax

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	36.459	5.103		7.144	<.001	26.432	46.487
	crim	-.108	.033	-.101	-3.287	.001	-.173	-.043
	zn	.046	.014	.118	3.382	<.001	.019	.073
	indus	.021	.061	.015	.334	.738	-.100	.141
	chas	2.687	.862	.074	3.118	.002	.994	4.380
	nox	-17.767	3.820	-.224	-4.651	<.001	-25.272	-10.262
	rm	3.810	.418	.291	9.116	<.001	2.989	4.631
	age	.001	.013	.002	.052	.958	-.025	.027
	dis	-1.476	.199	-.338	-7.398	<.001	-1.867	-1.084
	rad	.306	.066	.290	4.613	<.001	.176	.436
	tax	-.012	.004	-.226	-3.280	.001	-.020	-.005
	ptratio	-.953	.131	-.224	-7.283	<.001	-1.210	-.696
	black	.009	.003	.092	3.467	<.001	.004	.015
	lstat	-.525	.051	-.407	-10.347	<.001	-.624	-.425

a. Dependent Variable: medv

To create a regression equation using only the independent variables with p-values less than 0.05 (indicating statistical significance), we'll include only those variables. Based on the coefficients
table.

Since the variables indus and age have p-values greater than 0.05, indicating that they are not statistically significant predictors in the model, they have been removed from the regression
equation.

Equation

medv =
36.459−0.108(crim)+0.046(zn)+2.687(chas)−17.767(nox)+3.810(rm)−1.476(dis)+0.306(rad)−0.012(tax)−0.953(ptratio)+0.009(black)
−0.525(lstat)