MULTIPLE REGRESSION

Some of the figures in this presentation are taken from "An Introduction to Statistical Learning, with applications in R" (Springer, 2013) with permission from the authors: G. James, D. Witten, T. Hastie and R. Tibshirani

- What is it?
- What can it do? (use cases)
- How does it work?
- JMP Mechanics
- Interpret results (statistically)
- Interpret results (application)
- How to apply the results
- How to understand the managerial implications

WHAT IS IT?

Multiple linear regression is an approach for predicting a quantitative response variable (Y) with a multiple predictor variables (X).

Y	X	Objective	Summary of Fit Measures	Statistical Significance Measure	Operational Significance	Mgt Insights
Continuous	Continuous or Categorical	Explanatory	RSquare Adj, Root Mean Square Error	Prob > F (p-value)	Mean and Individual Confidence Limits	Profiler; Variable Importance

WHAT CAN IT DO? (USE CASES)

- Is there a relationship between advertising budget and sales?
- Is there a relationship between sales and price?
- Explaining customer spending based on demographic information and historical buying patterns
- Developing pricing strategies based on product mix and consumer characteristics
- Establishing housing prices

HOW DOES IT WORK?

Multiple Regression

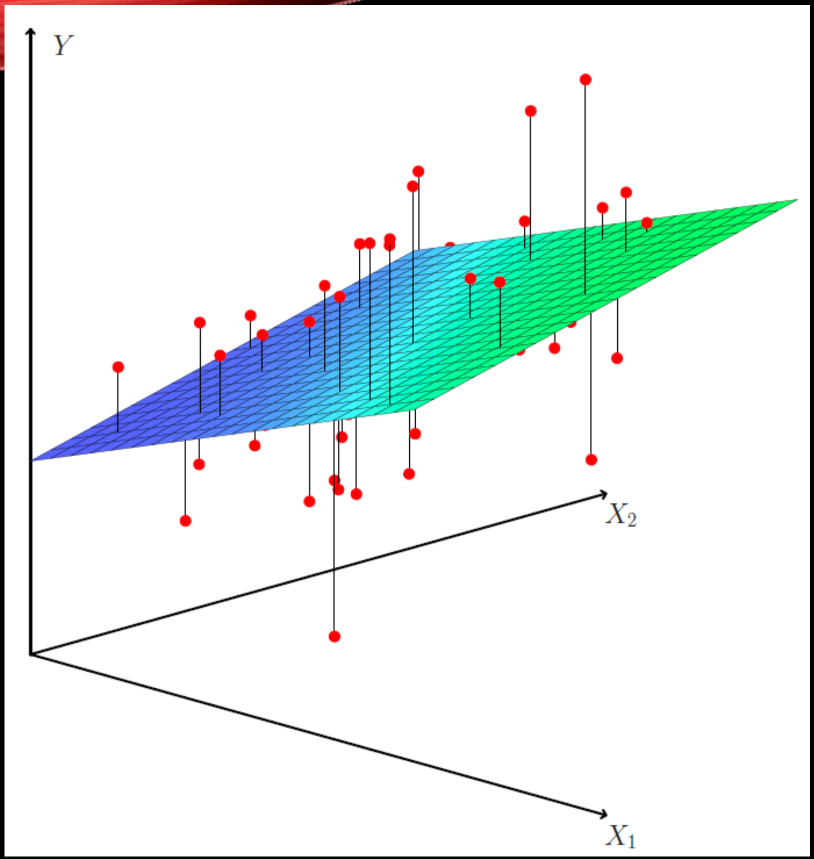


Figure 3.4, **An Introduction to Statistical Learning** by James, et al, (Springer 2013)

Least Squares Regression:

Fit to minimize the sum of squared residuals
$e_1^2 + e_2^2 + \ldots + e_n^2$
where $e = y - \hat{y}$

Assumptions:

When fitting a least squares line, we generally require

Linearity. The data should show a linear trend. If there is a nonlinear trend (e.g. left panel of Figure 7.13), an advanced regression method from another book or later course should be applied.

Nearly normal residuals. Generally the residuals must be nearly normal. When this condition is found to be unreasonable, it is usually because of outliers or concerns about influential points, which we will discuss in greater depth in Section 7.3. An example of non-normal residuals is shown in the second panel of Figure 7.13.

Constant variability. The variability of points around the least squares line remains roughly constant. An example of non-constant variability is shown in the third panel of Figure 7.13.

Be cautious about applying regression to data collected sequentially in what is called a **time series**. Such data may have an underlying structure that should be considered in a model and analysis. There are other instances where correlations within the data are important. This topic will be further discussed in Chapter 8.

Chapter 7, Introduction to Linear Regression, OpenIntro.org

OpenIntro Statistics

Assumptions

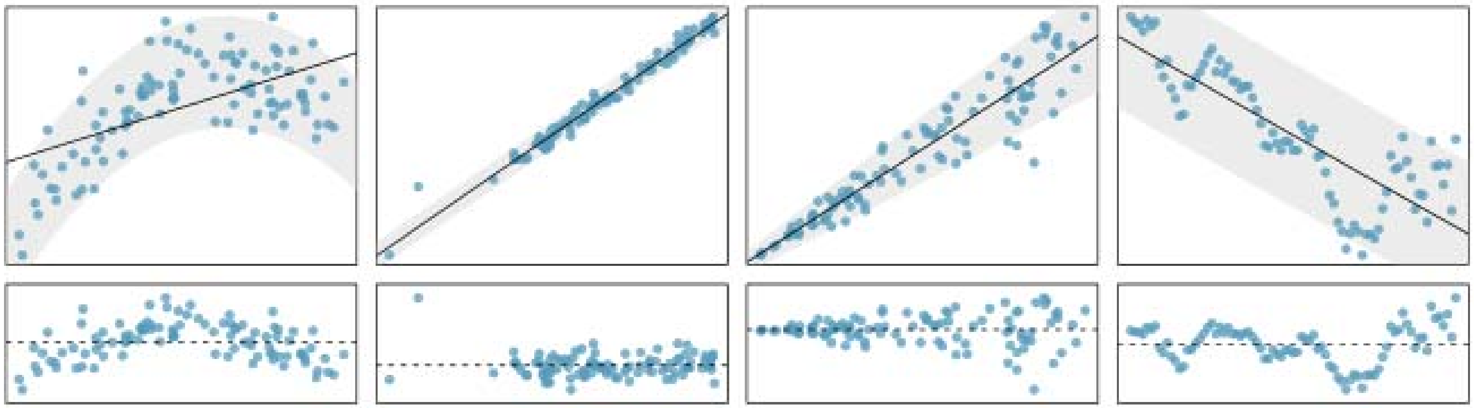


Figure 7.13: Four examples showing when the methods in this chapter are insufficient to apply to the data. In the left panel, a straight line does not fit the data. In the second panel, there are outliers; two points on the left are relatively distant from the rest of the data, and one of these points is very far away from the line. In the third panel, the variability of the data around the line increases with larger values of x. In the last panel, a time series data set is shown, where successive observations are highly correlated.

Chapter 7, Introduction to Linear Regression, OpenIntro.org

[OpenIntro Statistics]

JMP MECHANICS

jmp.com/learn rev 02/2014

Multiple Linear Regression

Multiple linear regression is used to model the relationship between a continuous response variable and continuous or categorical explanatory variables.

Multiple Linear Regression Using Fit Model

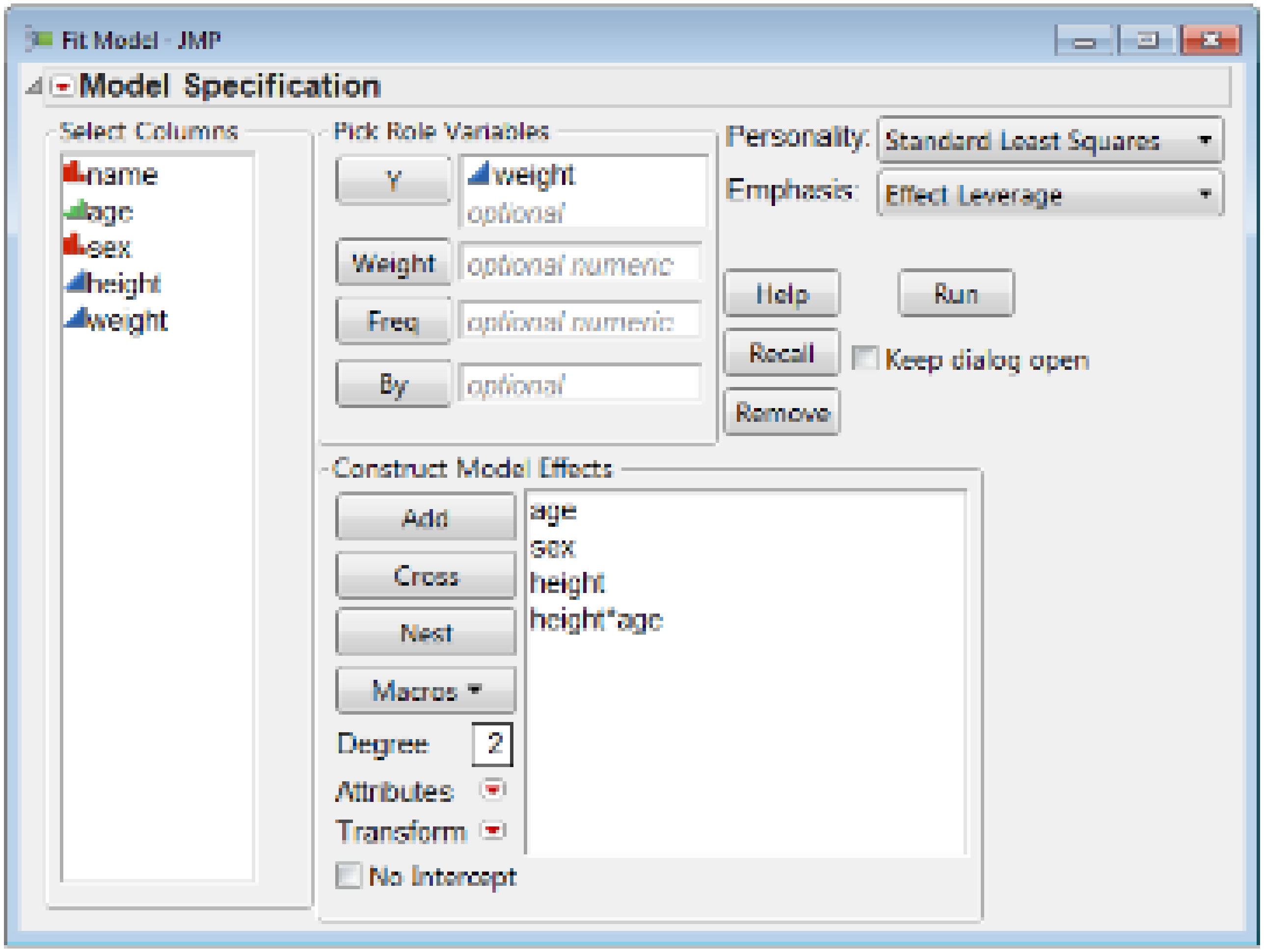
1. From an open JMP® data table, select **Analyze > Fit Model**.
2. Click on a continuous variable from **Select Columns**, and click **Y** (continuous variables have blue triangles).
3. Choose explanatory variables from **Select Columns**, and click **Add**.
4. Click **Run Model**.

By default, JMP will provide the following results:

- Actual by Predicted Plot.
- Summary of Fit table.
- Analysis of Variance table.
- Parameter Estimates table, and more (not shown).

JMP also provides Leverage Plots for each explanatory variable in the model, and for nominal and ordinal variables, the least squares means tables.

Example: Big Class.jmp (Help > Sample Data)



Tips:

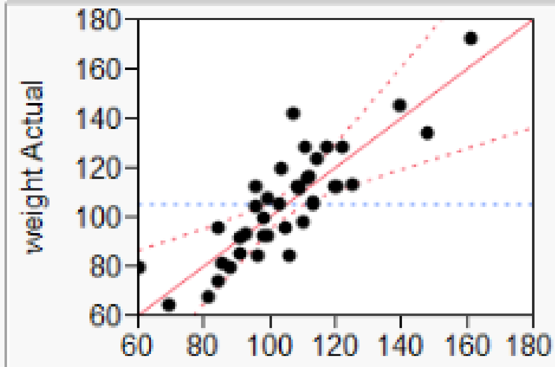
- To add interactions terms: In the **Fit Model Specification** window, select the variables under **Select Columns** and click **Cross**. The term **age*height** in the first figure is a two-way interaction. Higher-order terms can also be added.
- To save the prediction formula, predicted values, residuals or other values to the data table, click on the **top red triangle**, select **Save Columns.**

  JMP will create new columns in the data table.
- To view **indicator parameterization** (using 0, 1 coding), select **Estimates > Indicator Parameterization Estimates** from the **top red triangle**.

Response weight

Whole Model

Actual by Predicted Plot



Summary of Fit

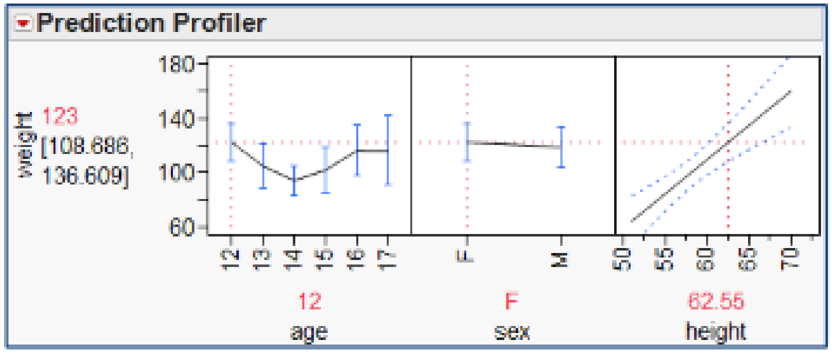
RSquare	0.738499
RSquare Adj	0.622276
Root Mean Square Error	13.64511
Mean of Response	105
Observations (or Sum Wgts)	40

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	12	14196.898	1183.07	6.3542
Error	27	5027.102	186.19	**Prob > F**
C. Total	39	19224.000		<.0001*

- To view the effect of an explanatory variable on the predicted response, click on the **top red triangle**, select **Factor Profiling** and choose **Profiler**.

  In the **Prediction Profiler**, click and drag the vertical red line for a variable to change the level or value. The predicted mean response and CI are displayed.



Note: For more details on regression analysis, see the book ***Fitting Linear Models*** (under **Help > Books**) or search for “regression” in the JMP Help.

HOUSING PRICES EXAMPLE

Dataset: HousingPrices.jmp

A real estate company that manages properties around a ski resort in the United States wishes to improve its method for pricing homes. Sample data is obtained on a number of measures, including size of the home and property, location, age of the house, and a strength-of-market indicator.

The Data HousingPrices.jmp

The data set contains information on about 45 residential properties near a popular North American ski resort sold during a recent 12-month period, and is a representative sample of the full set of properties sold during that time period. The variables in the data set are:

Price: Selling price of the property (in thousands of dollars)

Beds: Number of bedrooms in the house

Baths: Number of bathrooms in the house

Square Feet: Size of the house in square feet

Miles to Resort: Miles from the property to the downtown resort area

Miles to Base: Miles from the property to the base of the ski resort's facing mountain

Acres: Lot size in number of acres

Cars: Number of cars that will fit into the garage

Years Old: Age of the house at the time it was listed in years

DoM: Number of days the house was on the market before it was sold

http://www.jmp.com/en_us/academic/case-study-library.html

Multiple Regression Modeling Process

State modeling objective

Data gathering and preparation

Data exploration — Explore univariate | bivariate | multivariate relationships for data quality and to inform understanding

**

*

Model — Fit Model

Check assumptions and conditions: independence, equal variance, normal — Residual v. predicted; actual v. predicted: studentized residuals; normal plot;

Check high influence points, outliers and multicollinearity — Cook's D; VIF

State and test hypotheses about multiple regression coefficients — Analysis of variance; parameter estimate; Model reduction (stepwise; manual)

Interpret regression statistics - whole model results: adj R^2, overall p-value; RMSE

Interpret signs of regression coefficients

Assess usefulness of model (s_e, prediction intervals; profiler; variable importance)

Apply to stated model objective

* Transform y | Re-form model (quadratic, interaction terms,)

** Use stepwise as needed | Remove influential point | re-fit model

Multiple Regression Modeling Process

State modeling objective

Data gathering and preparation

Data exploration — Explore univariate | bivariate | multivariate relationships for data quality and to inform understanding

Model — Fit Model

Check assumptions and conditions: independence, equal variance, normal — Residual v. predicted; actual v. predicted: studentized residuals; normal plot;

Check high influence points, outliers and multicollinearity — Cook's D; VIF

* Transform y | Re-form model (quadratic, interaction terms,)

** Use stepwise as needed | Remove influential point | re-fit model

State and test hypotheses about multiple regression coefficients

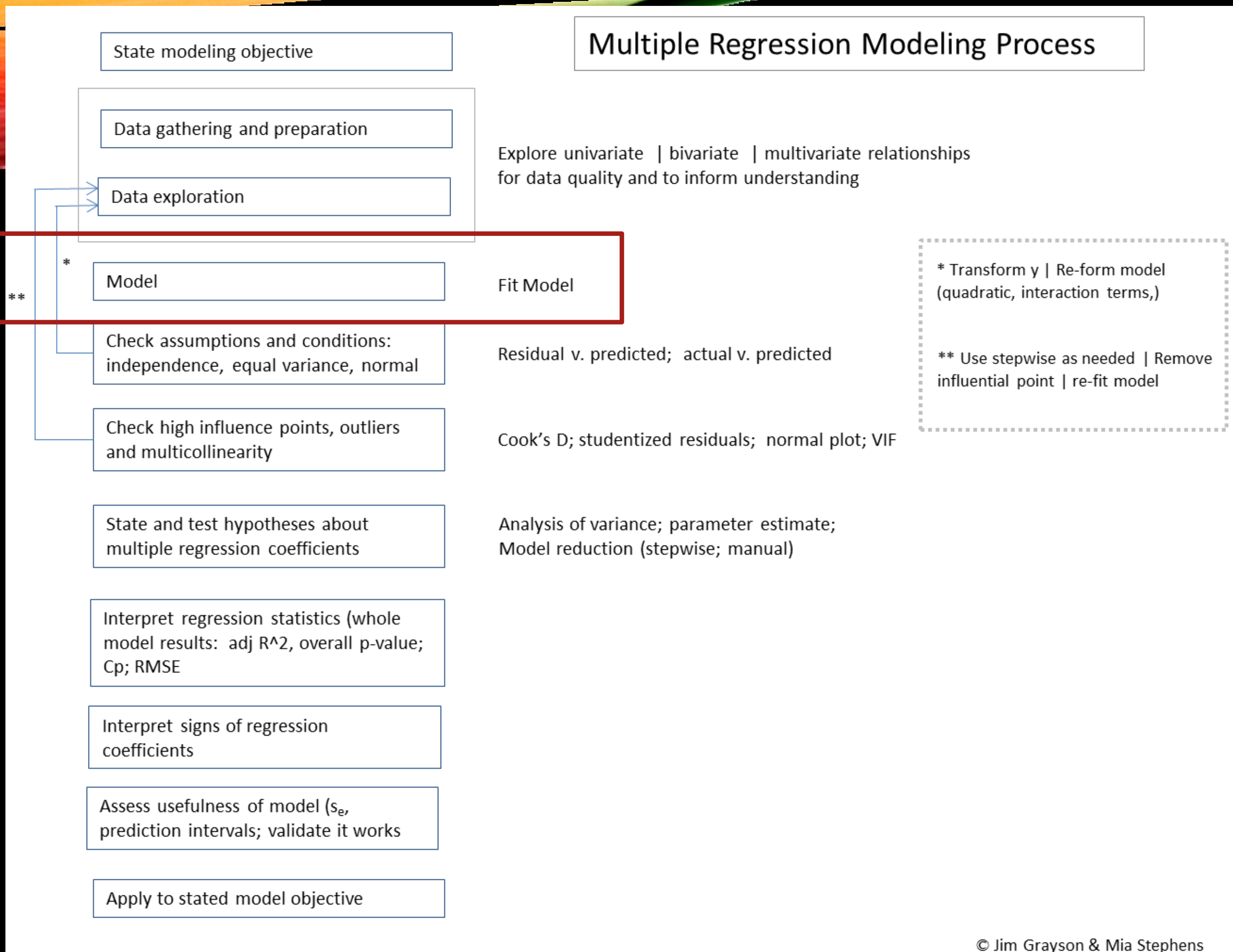
Analysis of variance; parameter estimate;
Model reduction (stepwise; manual)

Interpret regression statistics - whole model results: adj R^2, overall p-value; RMSE

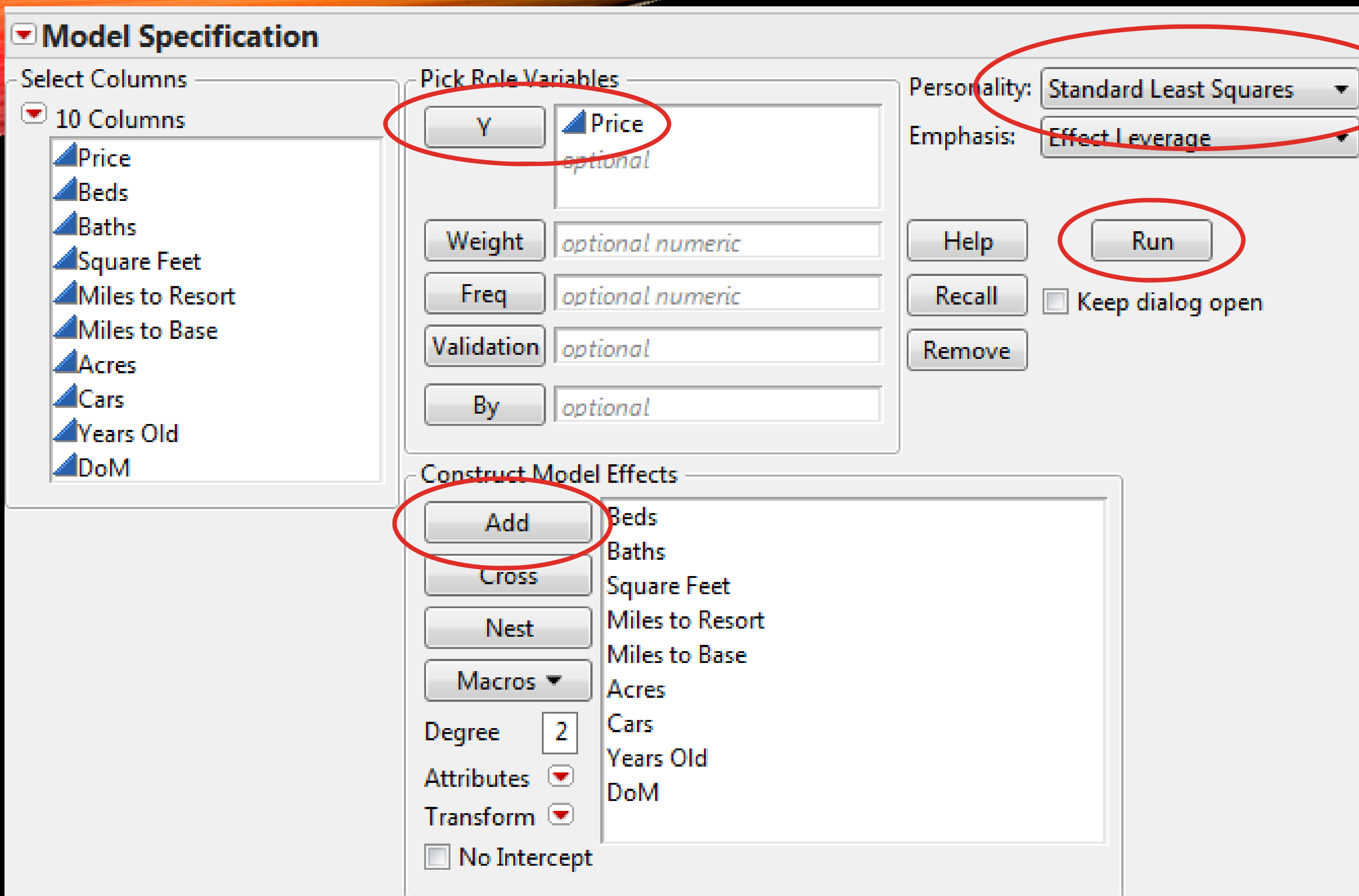
Interpret signs of regression coefficients

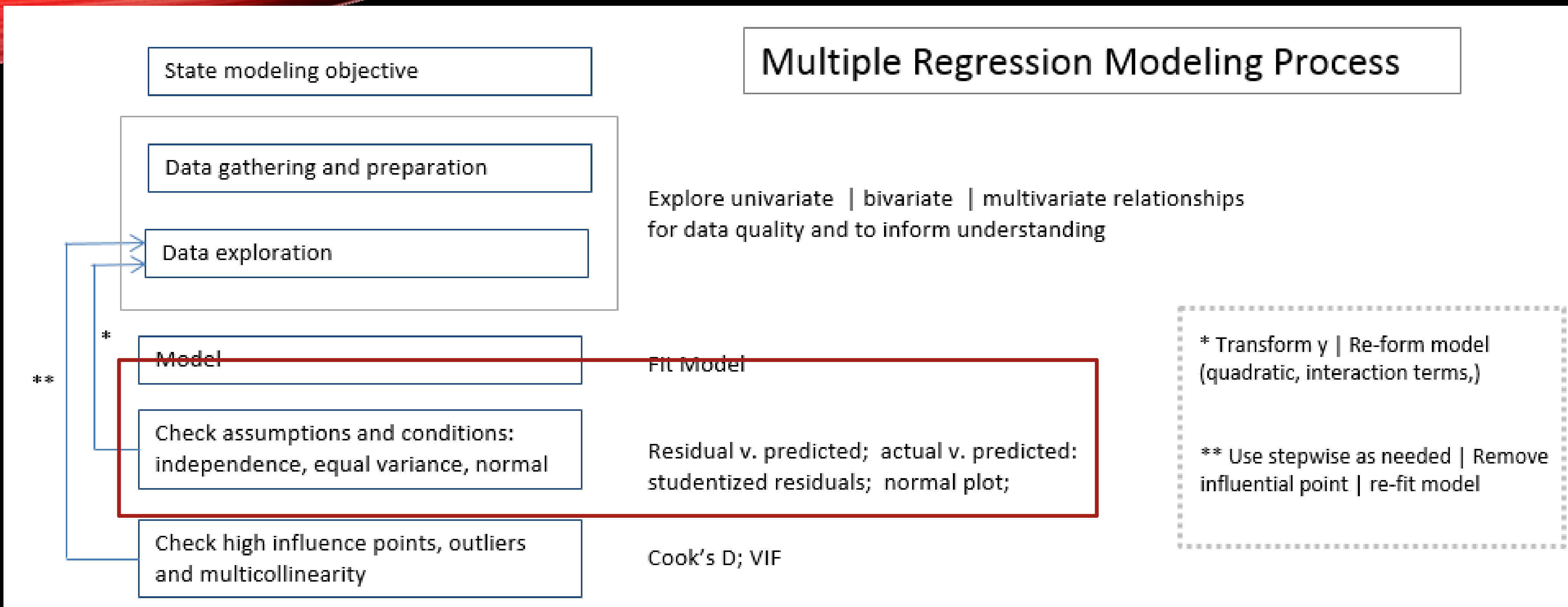
Assess usefulness of model (s_e, prediction intervals; profiler; variable importance)

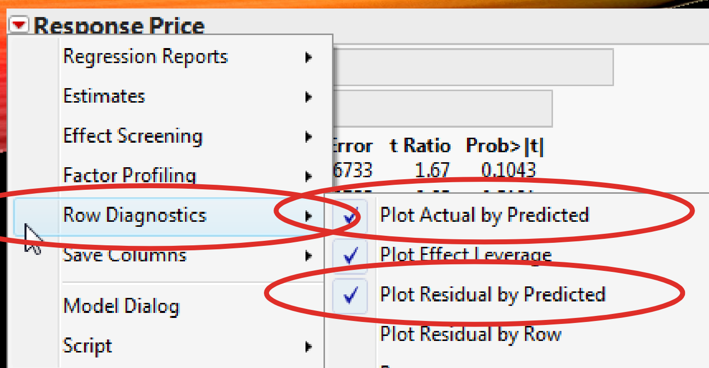
Apply to stated model objective


Multiple Regression Modeling Process
State modeling objective
Data gathering and preparation
Data exploration
Explore univariate | bivariate | multivariate relationships for data quality and to inform understanding
*
**
Model
Fit Model
* Transform y | Re-form model (quadratic, interaction terms,)
** Use stepwise as needed | Remove influential point | re-fit model
Check assumptions and conditions: independence, equal variance, normal
Residual v. predicted; actual v. predicted
Check high influence points, outliers and multicollinearity
Cook's D; studentized residuals; normal plot; VIF
State and test hypotheses about multiple regression coefficients
Analysis of variance; parameter estimate; Model reduction (stepwise; manual)
Interpret regression statistics (whole model results: adj R^2, overall p-value; Cp; RMSE
Interpret signs of regression coefficients
Assess usefulness of model (s_e, prediction intervals; validate it works
Apply to stated model objective
© Jim Grayson & Mia Stephens

Analyze > Fit Model




Multiple Regression Modeling Process
State modeling objective
Data gathering and preparation
Data exploration
Explore univariate | bivariate | multivariate relationships
for data quality and to inform understanding
*
**
Model
Fit Model
Check assumptions and conditions:
independence, equal variance, normal
Residual v. predicted; actual v. predicted:
studentized residuals; normal plot;
Check high influence points, outliers
and multicollinearity
Cook's D; VIF
* Transform y | Re-form model
(quadratic, interaction terms,)
** Use stepwise as needed | Remove
influential point | re-fit model

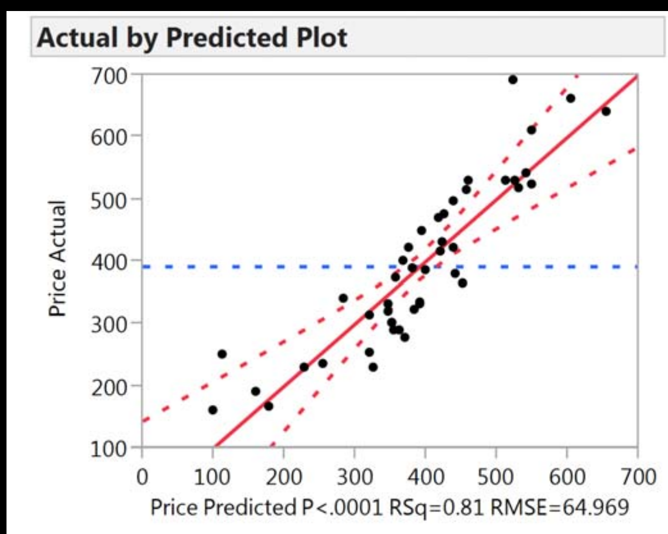


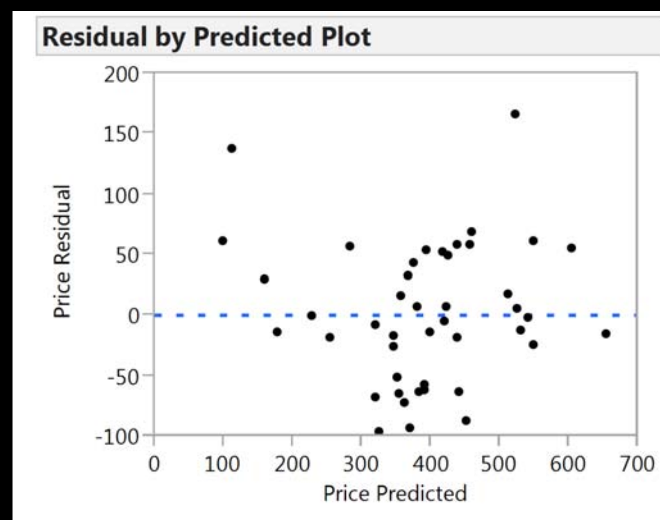
LIEN

L = linear
I = independent
E = equal variance
N = normally distributed errors

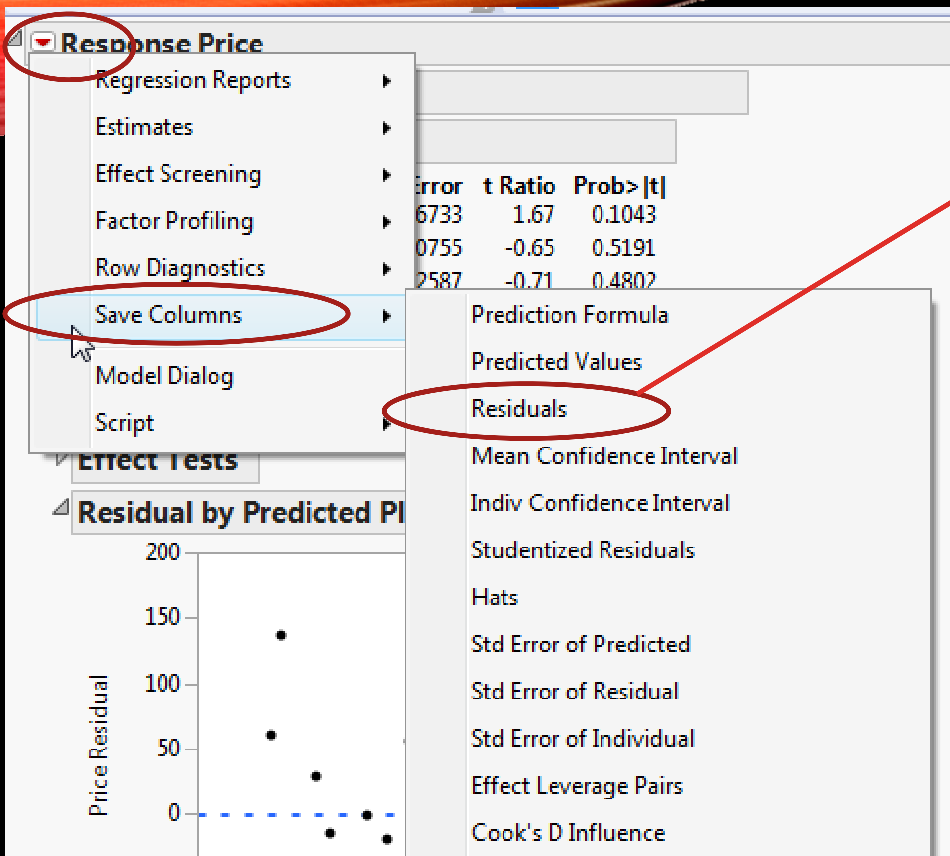
We use 'residual by predicted' to look for patterns and want to see values scattered randomly about zero and

'actual by predicted' to look for "leverage" for the whole model – points on the extremes and outside the confidence intervals are more likely to have larger contributions to the sum of squares





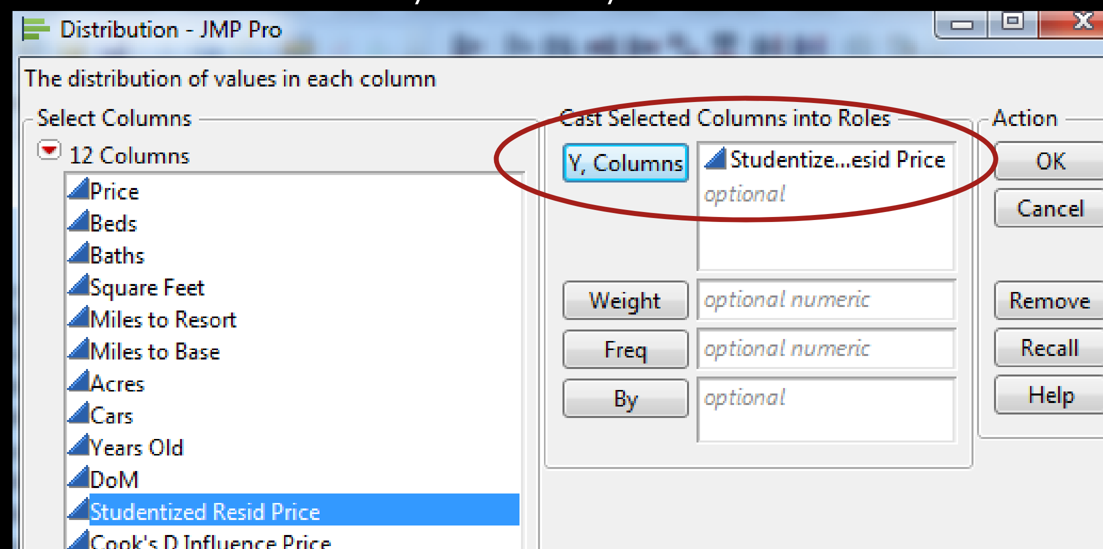
To obtain "studentized residuals"

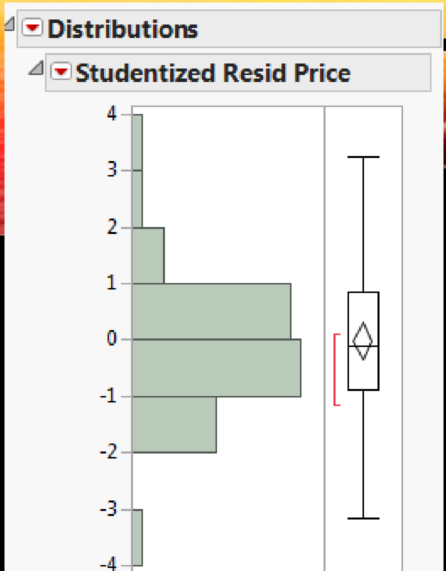


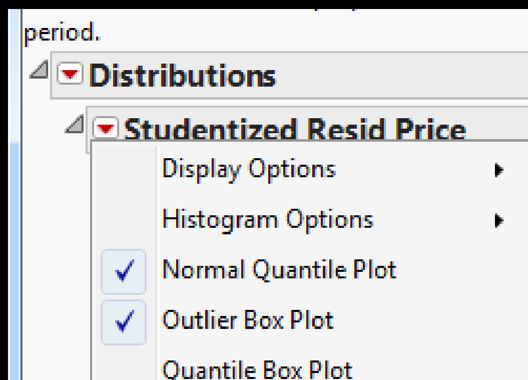
DoM	Studentized Resid Price	Cook's D Influence...
127	-0.294263733	0.0009972419
98	0.5363823513	0.0053877939
105	-0.09731404	0.0000874884
103	0.7680221538	0.0213680011
39	0.9461784813	0.0152031926
403	3.2456067936	0.6645816365

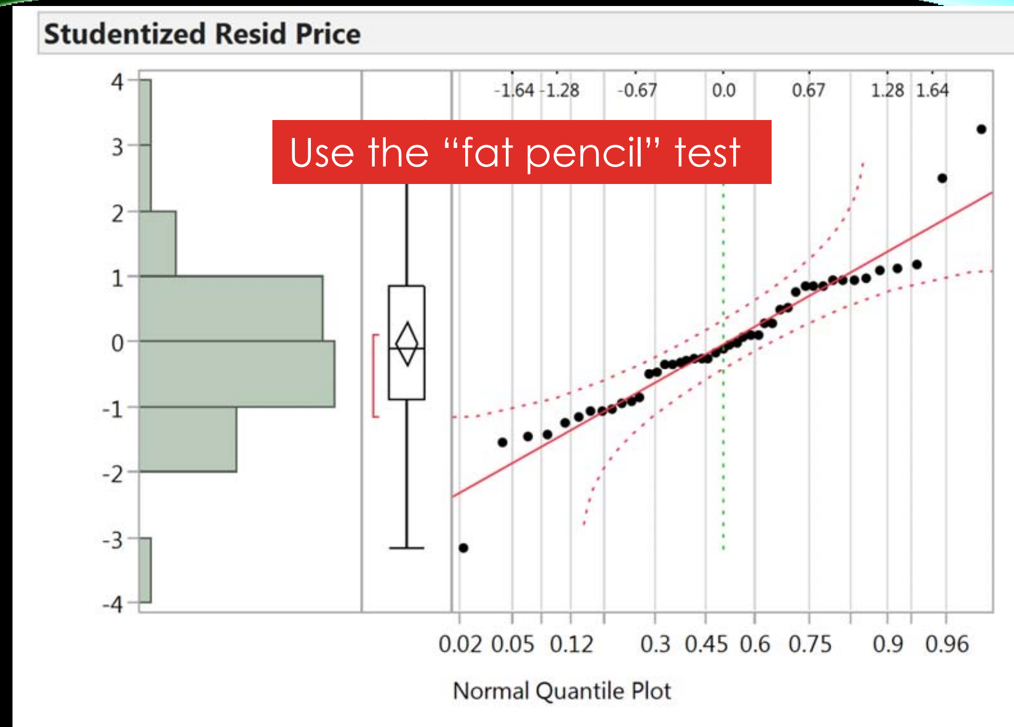
Studentized residuals are saved to the data table

We look at these by first Analyze > Distribution









LIEN

L = linear
I = independent
E = equal variance
N = normally distributed errors

We use studentized residuals to check for normally distributed errors

Dealing with Non Constant Variance - Transformations

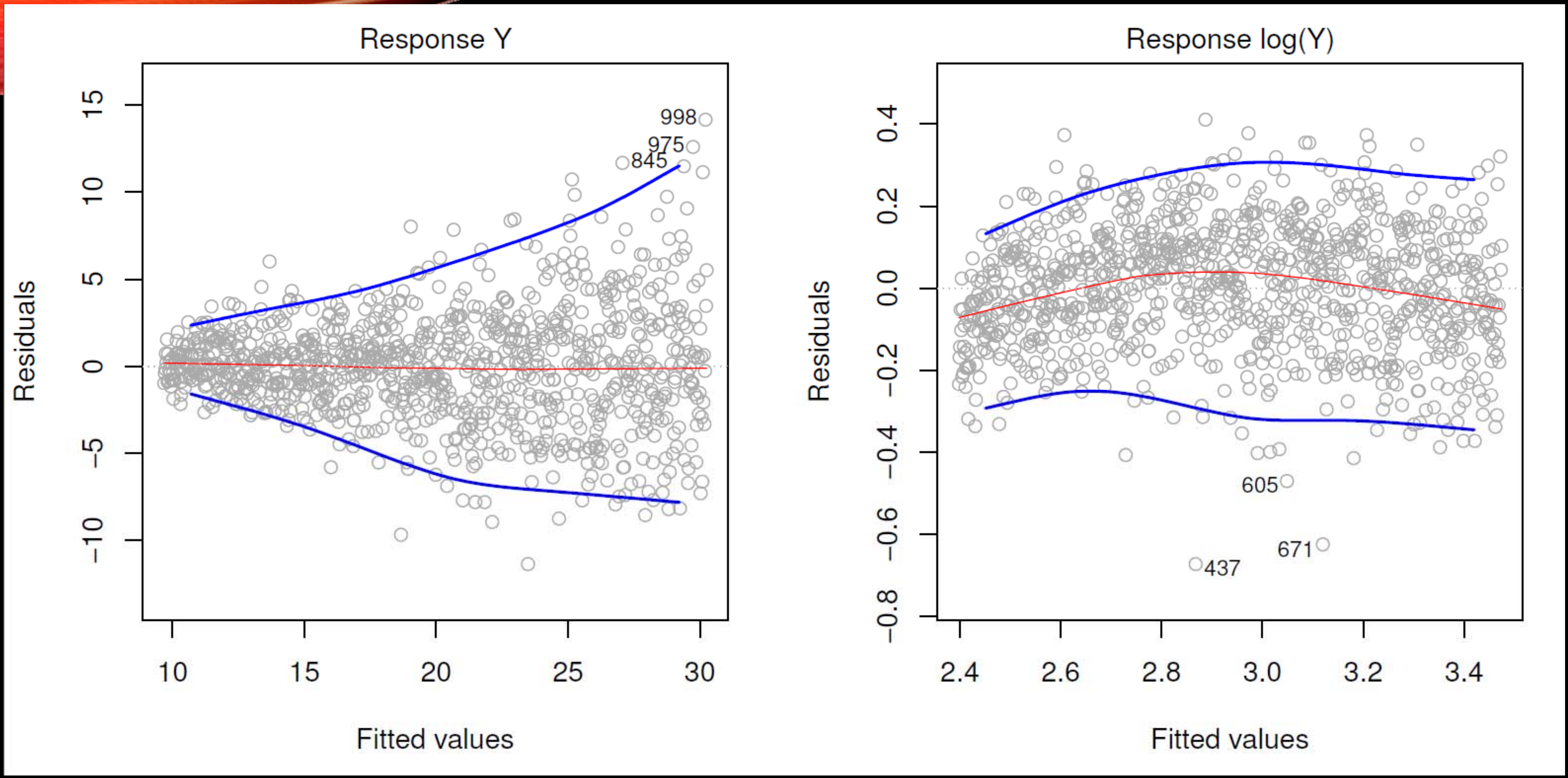
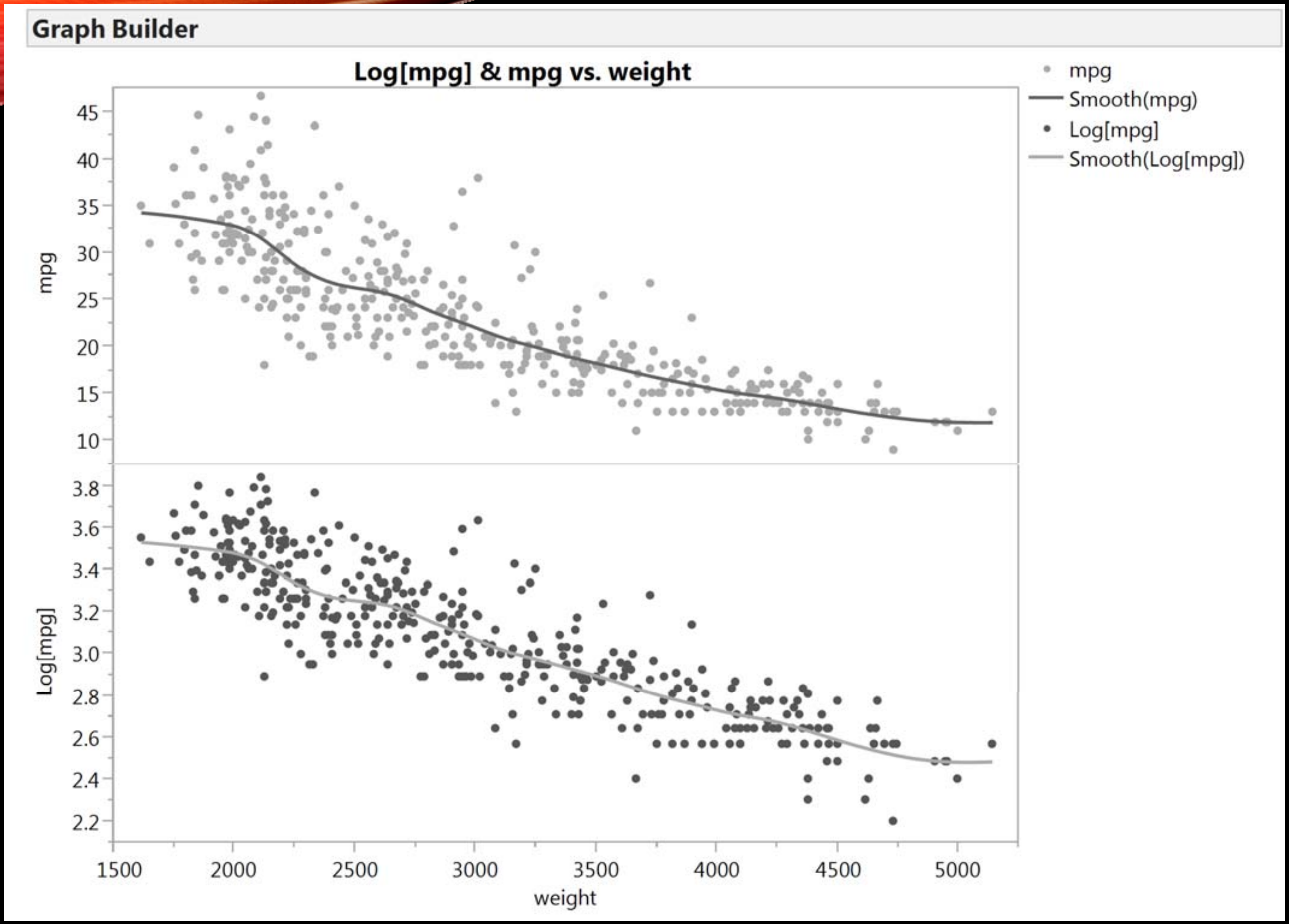


Figure 3.11 ISLR

Dealing with Non Constant Variance – Auto Example



Multiple Regression Modeling Process

State modeling objective

Data gathering and preparation

Data exploration — Explore univariate | bivariate | multivariate relationships for data quality and to inform understanding

Model — Fit Model

Check assumptions and conditions: independence, equal variance, normal — Residual v. predicted; actual v. predicted: studentized residuals; normal plot;

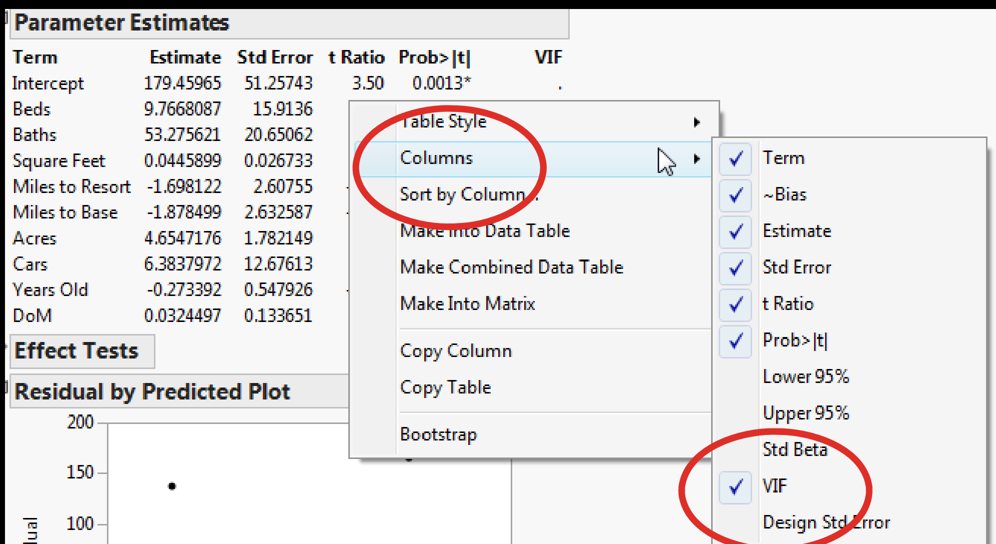
Check high influence points, outliers and multicollinearity — Cook's D; VIF

* Transform y | Re-form model (quadratic, interaction terms,)

** Use stepwise as needed | Remove influential point | re-fit model

Collinearity: when a predictor can be predicted well from other predictors

We can use the Variance Inflation Factor (VIF) to help identify collinearity in the model's predictors

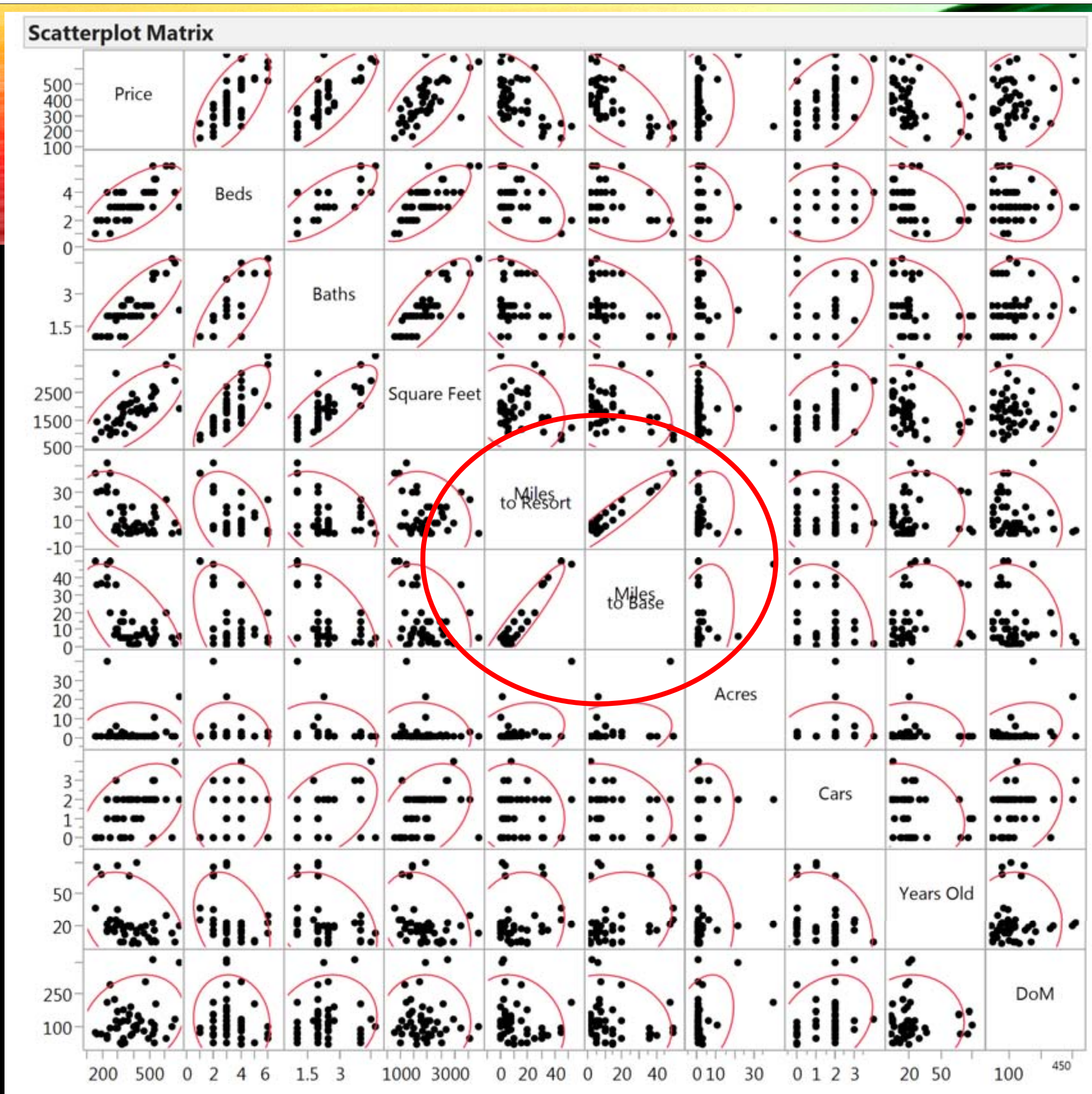


Parameter Estimates

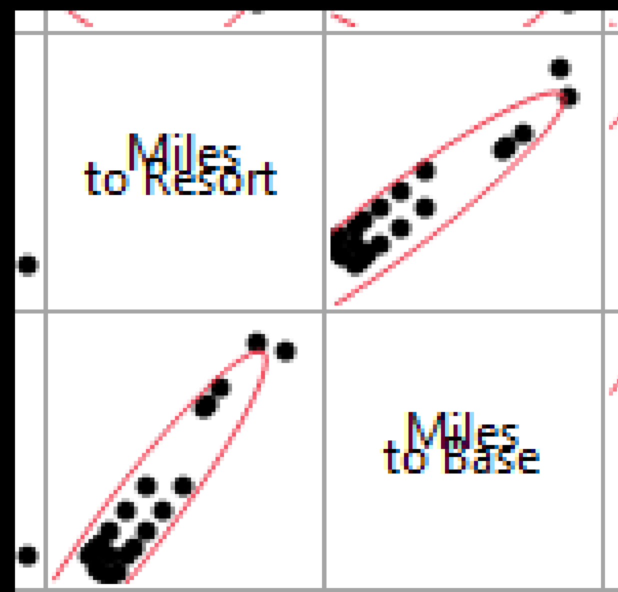
Term	Estimate	Std Error	t Ratio	Prob>|t|	VIF
Intercept	179.45965	51.25743	3.50	0.0013*	.
Beds	9.7668087	15.9136	0.61	0.5434	3.5544519
Baths	53.275621	20.65062	2.58	0.0142*	4.3822145
Square Feet	0.0445899	0.026733	1.67	0.1043	3.4757465
Miles to Resort	-1.698122	2.60755	-0.65	0.5191	13.822325
Miles to Base	-1.878499	2.632587	-0.71	0.4802	14.510754
Acres	4.6547176	1.782149	2.61	0.0132*	1.5212784
Cars	6.3837972	12.67613	0.50	0.6177	1.7883544
Years Old	-0.273392	0.547926	-0.50	0.6209	1.2901804
DoM	0.0324497	0.133651	0.24	0.8096	1.5627487

Rule of thumb:

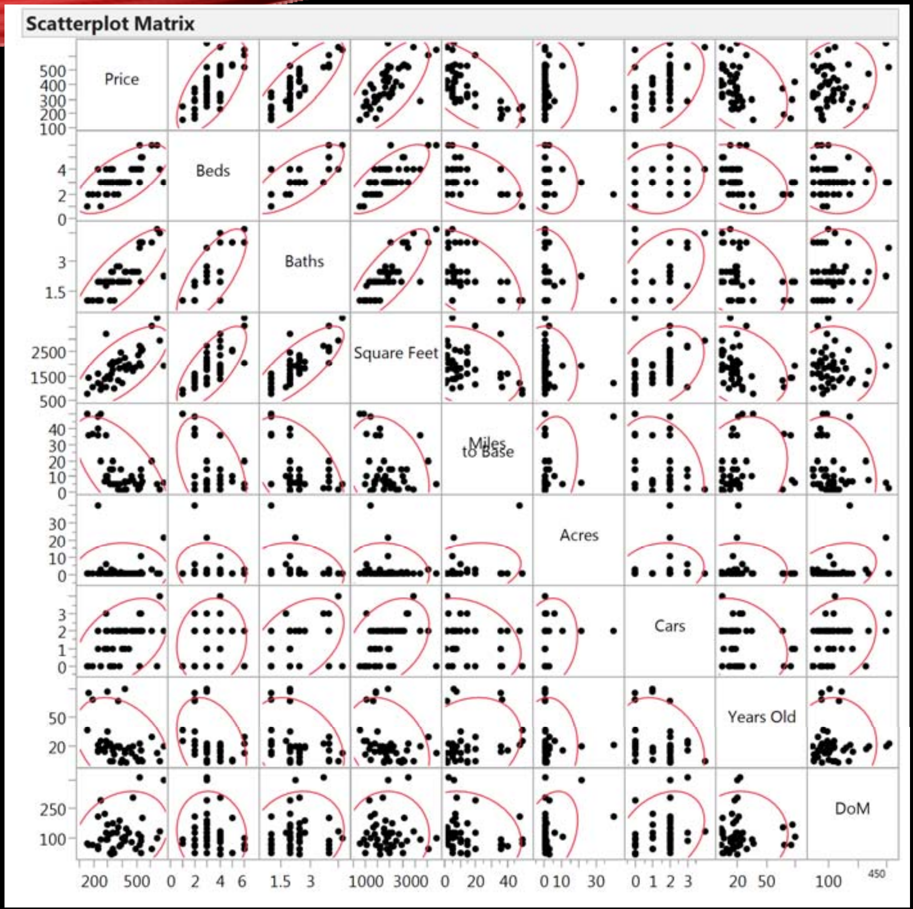
VIFs > 10 should be investigated



When we examine the scatterplot matrix we see a strong association between Miles to Resort and Miles to Base



We remove one of these predictors – say, Miles to Resort, and re-examine the VIFs and scatterplot matrix



Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|	VIF
Intercept	181.44396	50.75587	3.57	0.0010*	.
Beds	11.370069	15.59576	0.73	0.4707	3.4693829
Baths	50.724242	20.11275	2.52	0.0162*	4.224488
Square Feet	0.0430186	0.026411	1.63	0.1121	3.4474345
Miles to Base	-3.493138	0.877935	-3.98	0.0003*	1.6400357
Acres	4.3625859	1.710918	2.55	0.0152*	1.4248944
Cars	5.4694553	12.49696	0.44	0.6642	1.766415
Years Old	-0.192613	0.529416	-0.36	0.7181	1.2240612
DoM	0.0592831	0.12612	0.47	0.6412	1.414216

Leverage
Points that fall horizontally away from the center of the cloud tend to pull harder on the line, so we call them points with **high leverage**.

Points that fall horizontally far from the line are points of high leverage; these points can strongly influence the slope of the least squares line. If one of these high leverage points does appear to actually invoke its influence on the slope of the line – as in cases (3), (4), and (5) of Example 7.23 – then we call it an **influential point**. Usually we can say a point is influential if, had we fitted the line without it, the influential point would have been unusually far from the least squares line.

It is tempting to remove outliers. Don't do this without a very good reason. Models that ignore exceptional (and interesting) cases often perform poorly. For instance, if a financial firm ignored the largest market swings – the "outliers" – they would soon go bankrupt by making poorly thought-out investments.

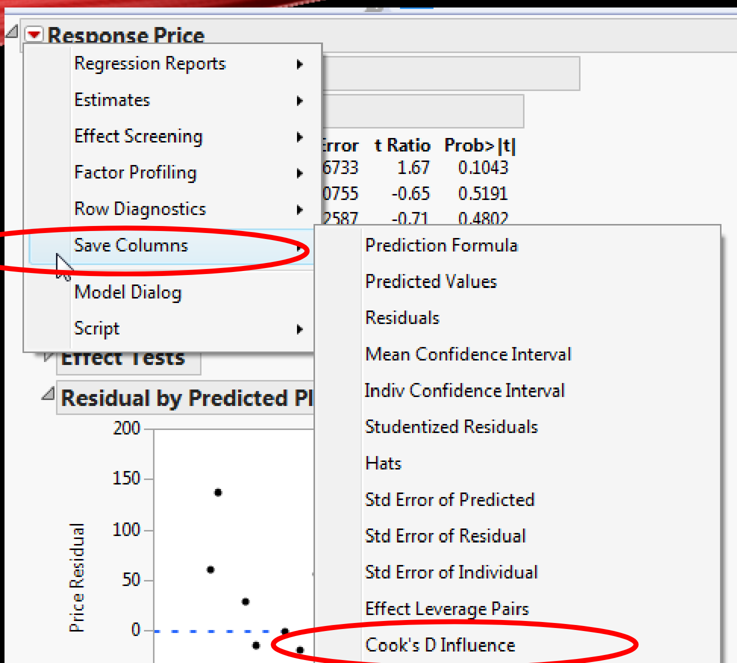
Caution: Don't ignore outliers when fitting a final model
If there are outliers in the data, they should not be removed or ignored without a good reason. Whatever final model is fit to the data would not be very helpful if it ignores the most exceptional cases.

Caution: Outliers for a categorical predictor with two levels
Be cautious about using a categorical predictor when one of the levels has very few observations. When this happens, those few observations become influential points.

Chapter 7, Introduction to Linear Regression, OpenIntro.org
OpenIntro Statistics

Leverage:

We can use the Cook's D measure to help identify undue influence

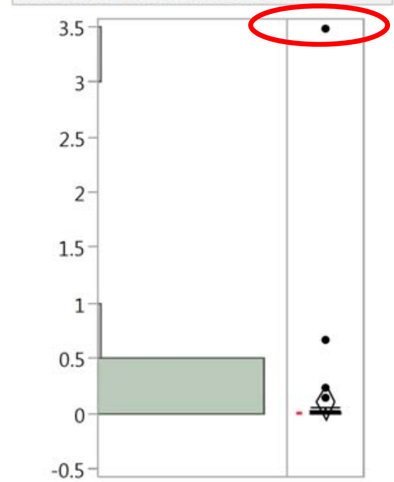


DoM	Studentized Resid Price	Cook's D Influence...
127	-0.294263713	0.0009972419
98	0.5363823513	0.0053877939
105	-0.0973140	0.0000874884
103	0.768022153	0.0213680011
39	0.946178481	0.0152031926
403	3.2456067936	0.6645816365

"Rule of Thumb":

Cook's D < 1

Cook's D Influence Price



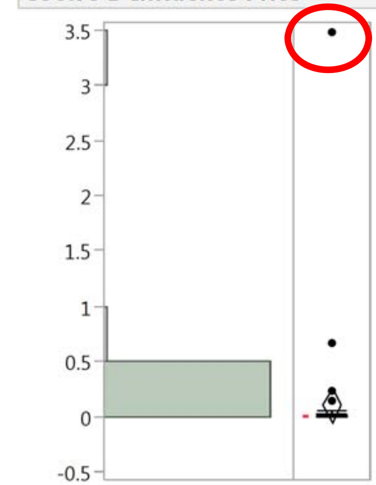
Quantiles

100.0%	maximum	3.48077
99.5%		3.48077
97.5%		3.05834
90.0%		0.09425
75.0%	quartile	0.02756
50.0%	median	0.01107
25.0%	quartile	0.00193
10.0%		0.00015
2.5%		6.91e-6
0.5%		1.38e-6
0.0%	minimum	1.38e-6

Summary Statistics

Mean	0.1134296
Std Dev	0.5238164
Std Err Mean	0.0780859
Upper 95% Mea	0.2708015
Lower 95% Mean	-0.043942
N	45

Cook's D Influence Price



Quantiles

100.0%	maximum	3.48077
99.5%		3.48077
97.5%		3.05834
90.0%		0.09425
75.0%	quartile	0.02756
50.0%	median	0.01107
25.0%	quartile	0.00193
10.0%		0.00015
2.5%		6.91e-6
0.5%		1.38e-6
0.0%	minimum	1.38e-6

Summary Statistics

Mean	0.1134296
Std Dev	0.5238164
Std Err Mean	0.0780859
Upper 95% Mea	0.2708015
Lower 95% Mean	-0.043942
N	45

	Price	Beds	Baths	Feet	Resort	Base	Acres	Cars	Years Old	DoM
1	330	3	2	1771	15	20	0.23	2	4	127
2	400	3	2	1213	5	1	0.17	1	5	98
3	416	3	2.5	1884	2	7	0.18	2	16	105
4	420	3	2	1922	1	6	0.29	1	80	103
5	496	4	2.5	1858	0	5	0.52	2	9	39
6	690	3	2.25	1948	1	6	21.91	2	20	403
7	230	2	1	1200	52	48	40	2	21	211
8	448	3	2	2002	2	7	0.18	1	18	126

We "select" the data point at the top (about 3.48) and that is associated with row 7.

When we examine row 7 we see the house is a ranch (40 acres) and therefore should be excluded from our analysis. We select "rows > hide and exclude" and rerun the analysis. Cook's D is now in an acceptable range.

INTERPRET RESULTS (STATISTICALLY VALID)

Relationship between response and predictors?

State and test hypotheses about multiple regression coefficients	Analysis of variance; parameter estimate; Model reduction (stepwise; manual)

Interpret regression statistics - whole model results: adj R^2, overall p-value; RMSE

Interpret signs of regression coefficients

Assess usefulness of model (s_e, prediction intervals; profiler; variable importance)

Apply to stated model objective

Summary of Fit

RSquare	0.809043
RSquare Adj	0.759939
Root Mean Square Error	64.96897
Mean of Response	391.1911
Observations (or Sum Wgts)	45

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	9	625914.08	69546.0	16.4763
Error	35	147733.84	4221.0	**Prob > F**
C. Total	44	773647.92		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	179.45965	51.25743	3.50	0.0013*
Beds	9.7668087	15.9136	0.61	0.5434
Baths	53.275621	20.65062	2.58	0.0142*
Square Feet	0.0445899	0.026733	1.67	0.1043
Miles to Resort	-1.698122	2.60755	-0.65	0.5191
Miles to Base	-1.878499	2.632587	-0.71	0.4802
Acres	4.6547176	1.782149	2.61	0.0132*
Cars	6.3837972	12.67613	0.50	0.6177
Years Old	-0.273392	0.547926	-0.50	0.6209
DoM	0.0324497	0.133651	0.24	0.8096

Fit

Hypotheses
(Model Significance)

Details
(Variable Significance)

Summary of Fit

RSquare	0.858737
RSquare Adj	0.826448
Root Mean Square Error	54.91143
Mean of Response	394.8545
Observations (or Sum Wgts)	44

Model Fit: RSquare Adj about 83%; RMSE = 54.9

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	8	641540.54	80192.6	26.5955
Error	35	105534.29	3015.3	Prob > F
C. Total	43	747074.83		<.0001*

Null Hypothesis $b_1 = b_2 = b_3 = \ldots = b_n = 0$
Alternate Hypothesis: At least one $b \neq 0$

Decision: Reject Ho and Accept Ha (at least one $b \neq 0$) with pvalue < .0001

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	173.94428	43.29045	4.02	0.0003*
Beds	9.1754331	13.30058	0.69	0.4948
Baths	58.081524	17.24473	3.37	0.0019*
Square Feet	0.0376742	0.022546	1.67	0.1036
Miles to Base	-2.852459	0.766609	-3.72	0.0007*
Acres	12.449926	2.570654	4.84	<.0001*
Cars	6.1533383	10.64938	0.58	0.5671
Years Old	-0.205754	0.451095	-0.46	0.6511
DoM	-0.031255	0.110042	-0.28	0.7781

Given that all variables are in the model we notice that Baths, Miles to Base, Acres are significant.

We will undertake an iterative process of eliminating one variable at a time, re-examining the p-values and then eliminating the next variable with a p-value > 0.10

Summary of Fit

RSquare	0.854004
RSquare Adj	0.83903
Root Mean Square Error	52.88352
Mean of Response	394.8545
Observations (or Sum Wgts)	44

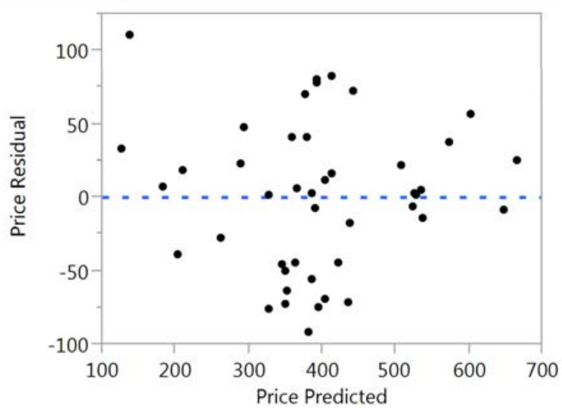
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	638004.83	159501	57.0326
Error	39	109070.00	2797	Prob > F
C. Total	43	747074.83		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|	VIF
Intercept	175.46632	30.1538	5.82	<.0001*	.
Baths	63.61719	14.57999	4.36	<.0001*	3.1587539
Square Feet	0.0472992	0.019337	2.45	0.0191*	2.6773845
Miles to Base	-3.025002	0.694498	-4.36	<.0001*	1.3192172
Acres	12.462947	2.237932	5.57	<.0001*	1.0187721

Residual by Predicted Plot



The results of our iterative process:

Notice that p-values are less than 0.02, Rsquare Adj is about 84%, VIFs are low and residual by predicted looks better that before.

INTERPRET RESULTS (USEFULNESS)

State and test hypotheses about multiple regression coefficients

Analysis of variance; parameter estimate;
Model reduction (stepwise; manual)

Interpret regression statistics - whole model results: adj R^2, overall p-value; RMSE

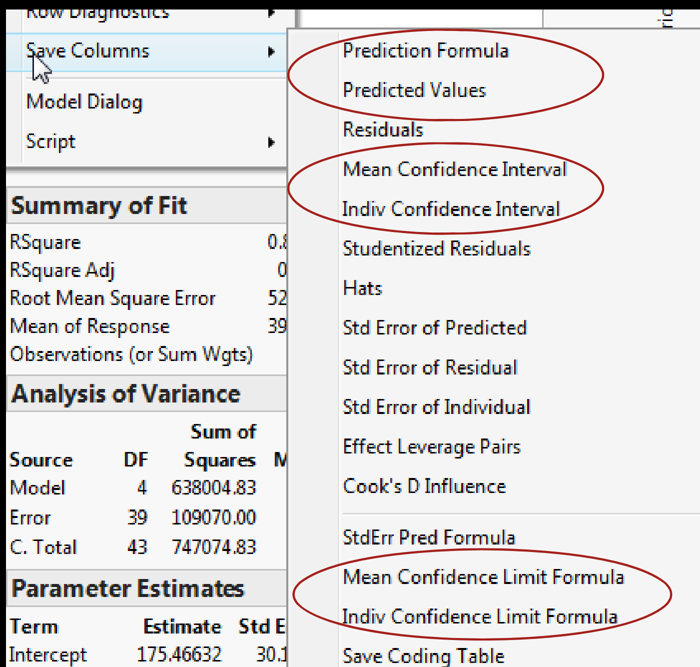
Interpret signs of regression coefficients

Assess usefulness of model (s_e, prediction intervals; profiler; variable importance)

Apply to stated model objective

Summary of Fit	
RSquare	0.854004
RSquare Adj	0.83903
Root Mean Square Error	52.88352
Mean of Response	394.8545
Observations (or Sum Wgts)	44

The Rsquare / Rsquare Adj gives us a measure of the variability in the data explained with our model. Larger values are preferred.



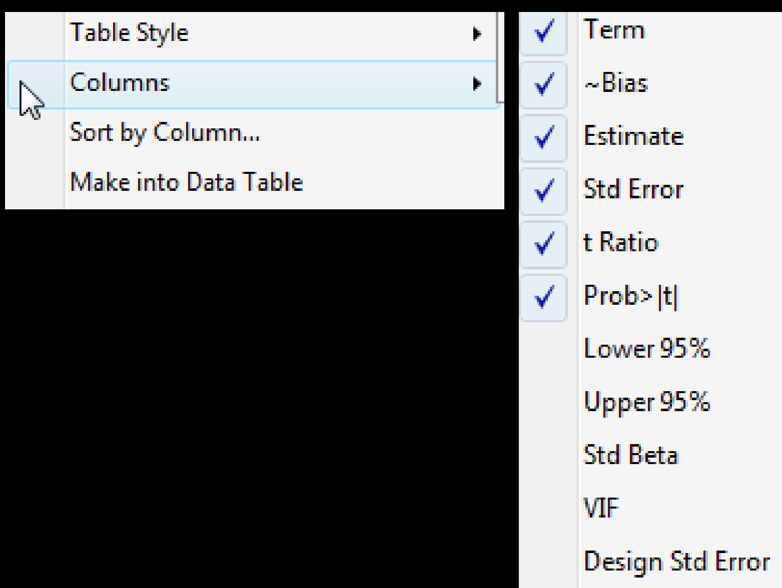
The Root Mean Square Error tells us the standard deviation of the error (ε) – a measure of variability of the model.

These "Save Columns" are added to the Data Table columns

Pred Formula Price	Predicted Price	Lower 95% Mean Price	Upper 95% Mean Price	Lower 95% Indiv Price	Upper 95% Indiv Price
328.83406706	328.83406706	309.797172	347.87096213	220.1862594	437.48187472
359.16836166	359.16836166	327.26202573	391.07469758	247.54418413	470.79253918
404.6893553	404.6893553	385.47911534	423.89959526	296.01104074	513.36766986
379.07405722	379.07405722	354.15523976	403.99287467	269.24286766	488.90524677

Parameter (Coefficients) Estimation and Variability

Select the “parameter” table and right click >
Select Columns >
Select Lower 95% and Upper 95%



Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|	Lower 95%	Upper 95%
Intercept	175.46632	30.1538	5.82	<.0001*	114.47449	236.45814
Baths	63.61719	14.57999	4.36	<.0001*	34.126385	93.107996
Square Feet	0.0472992	0.019337	2.45	0.0191*	0.0081874	0.086411
Miles to Base	-3.025002	0.694498	-4.36	<.0001*	-4.429757	-1.620247
Acres	12.462947	2.237932	5.57	<.0001*	7.9363021	16.989592

The Lower 95% and Upper 95% are confidence intervals around the coefficients.

HOW TO APPLY THE RESULTS

Prediction

The Prediction Formula:

175.466319174148 + 63.6171900658549 * :Baths +
0.0472992279684479 * :Square Feet +
-3.02500213727099 * :Miles to Base + 12.4629468282909 * :Acres

For a 2 Bath, 3000 SF, 2 Miles to Base, 1 Acres

Pred Formula Price	Predicted Price	Lower 95% Mean Price	Upper 95% Mean Price	Lower 95% Indiv Price	Upper 95% Indiv Price
451.01132576	•	391.76403471	510.25861682	328.73220149	573.29045004

Predicted Price +/- Confidence Intervals for the Mean and Individual Prediction Intervals

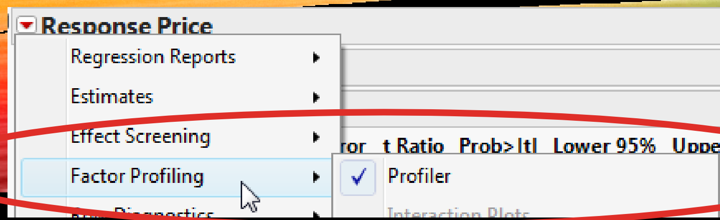
MANAGERIAL IMPLICATIONS

Insights: Profiler; Variable Importance; Confidence and Prediction Intervals

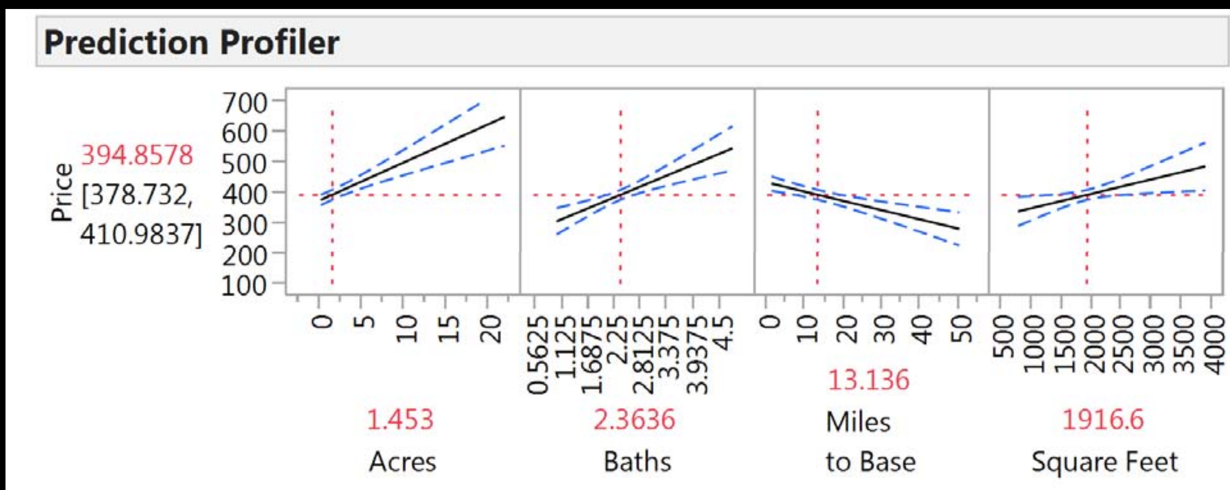
Confidence Intervals and Prediction Intervals

Confidence Intervals and Prediction Intervals can be examined to determine the usefulness of the prediction – is the model useful taking into account the variability in the intervals

Lower 95% Mean Price	Upper 95% Mean Price	Lower 95% Indiv Price	Upper 95% Indiv Price
391.76403471	510.25861682	328.73220149	573.29045004



Profiler



The Profiler allows an interactive interface to vary the value of predictor(s) and observe the impact on the response (Price) and the impact of changing one predictor on the other predictors. The profiler can help build understanding and intuition about the relationship of the predictors to explaining the response variable.

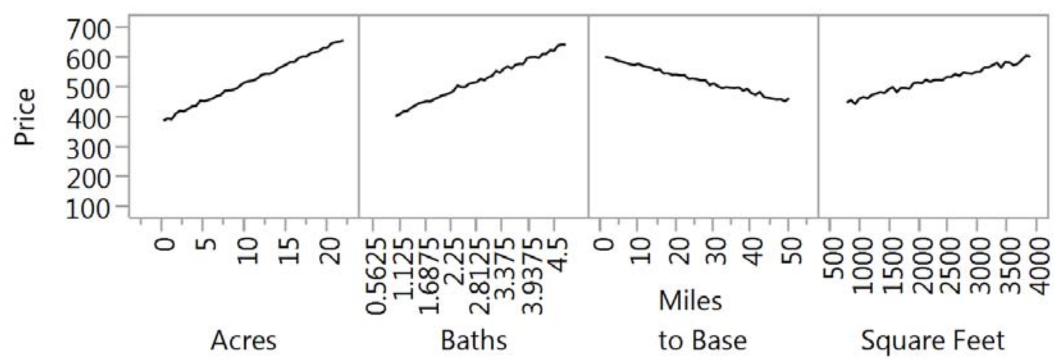
Variable Importance

Variable Importance: Independent Uniform Inputs

Summary Report

Column	Main Effect	Total Effect
Acres	0.21	0.211
Baths	0.157	0.157
Miles to Base	0.068	0.068
Square Feet	0.065	0.065

Marginal Model Plots



This tool shows the relative importance of the predictor variables, first in relative impact – that is, which variables with a +/- change in value will have the largest impact on the response (think of this as a sensitivity analysis) and then secondly, the marginal model plots provide insight into the main effect of the variables