Summary of Data

The CONTENTS Procedure

Data Set Name	WORK.NBA_SD_2021_2022	**Observations**	188
Member Type	DATA	**Variables**	33
Engine	V9	**Indexes**	0
Created	22/06/2022 13:20:11	**Observation Length**	520
Last Modified	22/06/2022 13:20:11	**Deleted Observations**	0
Protection		**Compressed**	NO
Data Set Type		**Sorted**	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information	
Data Set Page Size	131072
Number of Data Set Pages	1
First Data Page	1
Max Obs per Page	251
Obs in First Data Page	188
Number of Data Set Repairs	0
Filename	/saswork/SAS_workD5060000D388_odaws04-usw2.oda.sas.com/SAS_work1BE80000D388_odaws04-usw2.oda.sas.com/nba_sd_2021_2022.sas7bdat
Release Created	9.0401M6
Host Created	Linux
Inode Number	1946902
Access Permission	rw-r--r--
Owner Name	u58912630
File Size	256KB
File Size (bytes)	262144

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat
6	Age	Num	8	BEST12.	BEST32.
11	Assists	Num	8	BEST12.	BEST32.
18	Blocks	Num	8	BEST12.	BEST32.
13	Defensive_Rebounds	Num	8	BEST12.	BEST32.
8	Field_Goal_Attempts	Num	8	BEST12.	BEST32.
7	Field_Goals	Num	8	BEST12.	BEST32.
17	Freethrow_Attempts	Num	8	BEST12.	BEST32.
16	Freethrows	Num	8	BEST12.	BEST32.
4	Games	Num	8	BEST12.	BEST32.
5	Minutes_Played	Num	8	BEST12.	BEST32.
12	Offensive_Rebounds	Num	8	BEST12.	BEST32.
1	Player_name	Char	78	$78.	$78.
9	Points	Num	8	BEST12.	BEST32.
3	Position	Char	2	$2.	$2.
19	Salary	Char	12	$12.	$12.
15	Steals	Num	8	BEST12.	BEST32.

Summary of Data

The CONTENTS Procedure

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
2	Team	Char	3	$3.	$3.
14	Total_Rebounds	Num	8	BEST12.	BEST32.
10	Turnover	Num	8	BEST12.	BEST32.
20	Year	Num	8		
23	astg	Num	8	6.3	
24	fgg	Num	8	6.3	
32	fgper	Num	8		
28	ftg	Num	8	6.3	
33	ftper	Num	8		
30	log_salary	Num	8	6.3	
22	mpg	Num	8	6.3	
31	player_name_year	Char	200		
25	ptsg	Num	8	6.3	
27	rbg	Num	8	6.3	
21	salary_num	Num	8	COMMA10.	
29	stlg	Num	8	6.3	
26	tovg	Num	8	6.3	

Data contains 188 observations of NBA point guards and shooting guards from the 2020-2021 regular season and 2021-2022 regular season, includes game statistics and salaries.

Correlation of Data

The CORR Procedure

13 With Variables:	mpg ptsg astg tovg rbg fgper ftg ftper stlg Games Total_Rebounds Freethrows Steals
2 Variables:	salary_num log_salary

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
mpg	188	26.29499	6.61103	4943	10.87500	37.87692
ptsg	188	12.91179	6.40657	2427	3.20313	31.98413
astg	188	3.61240	2.26247	679.13086	0.43333	11.73846
tovg	188	1.59582	0.96413	300.01334	0.26667	4.80000
rbg	188	3.43695	1.54137	646.14714	0.87097	11.53846
fgper	188	0.44026	0.04070	82.76917	0.35317	0.61628
ftg	188	1.89736	1.49681	356.70287	0.06667	7.68254
ftper	188	0.81012	0.08533	152.30271	0.44444	1.00000
stlg	188	0.90389	0.36180	169.93224	0.20000	2.02941
Games	188	66.22340	6.12989	12450	56.00000	81.00000
Total_Rebounds	188	228.19681	104.90397	42901	54.00000	750.00000
Freethrows	188	124.97340	97.02827	23495	4.00000	500.00000
Steals	188	59.96809	24.90837	11274	12.00000	138.00000
salary_num	188	9745967	9978694	1832241719	925258	45780966
log_salary	188	15.62204	0.98124	2937	13.73783	17.63938

Pearson Correlation Coefficients, N = 188
Prob > |r| under H0: Rho=0

	salary_num	log_salary
mpg	0.60511 <.0001	0.66581 <.0001
ptsg	0.61268 <.0001	0.62041 <.0001
astg	0.57608 <.0001	0.55388 <.0001
tovg	0.57800 <.0001	0.55898 <.0001
rbg	0.54598 <.0001	0.51185 <.0001
fgper	0.12211 0.0950	0.02496 0.7339
ftg	0.53438 <.0001	0.50671 <.0001
ftper	0.09115 0.2135	0.13058 0.0741
stlg	0.44439 <.0001	0.47239 <.0001
Games	0.03039 0.6788	0.06117 0.4044
Total_Rebounds	0.53025 <.0001	0.50136 <.0001
Freethrows	0.53605 <.0001	0.50899 <.0001
Steals	0.42495 <.0001	0.46026 <.0001

Correlation of Data

The CORR Procedure

Correlation of salary and log transform of salary between key variables to see strength of linear relationship.

The strongest correlations results are from the following variables minutes per game (mpg), points per game (ptsg), assists per game (astg), and turnovers per game (tovg) which will be used to construct a linear model.

Model 1 Salary

The REG Procedure
Model: MODEL1
Dependent Variable: salary_num

Number of Observations Read	188
Number of Observations Used	188

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	8.186051E15	2.046513E15	35.89	<.0001
Error	183	1.043435E16	5.701831E13		
Corrected Total	187	1.86204E16			

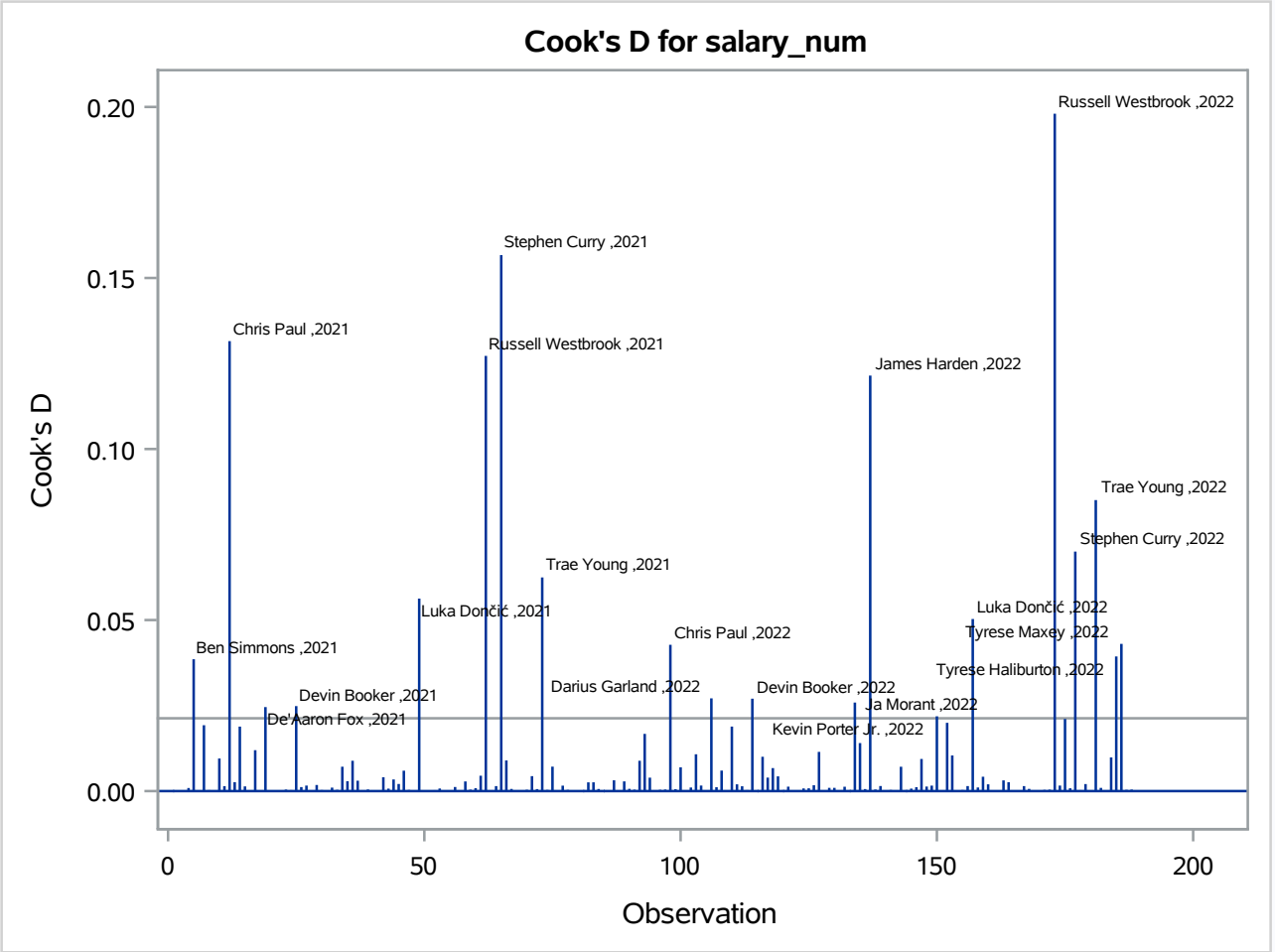
Root MSE	7551047	R-Square	0.4396
Dependent Mean	9745967	Adj R-Sq	0.4274
Coeff Var	77.47869		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-7952042	2714147	-2.93	0.0038
mpg	1	295498	164111	1.80	0.0734
ptsg	1	560741	209373	2.68	0.0081
astg	1	1557482	493716	3.15	0.0019
tovg	1	-1841399	1568285	-1.17	0.2419

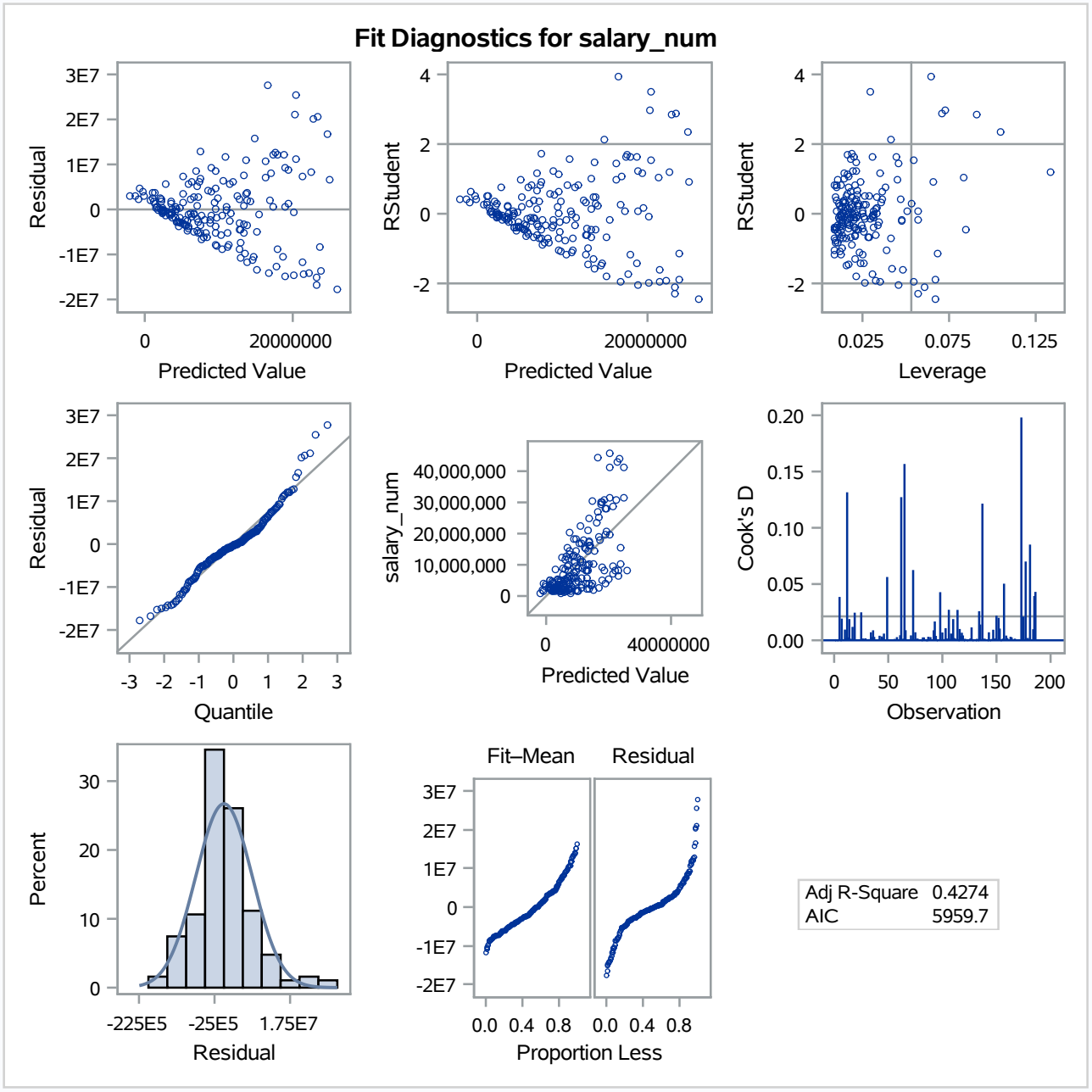
Model 1 Salary

The REG Procedure
Model: MODEL1
Dependent Variable: salary_num



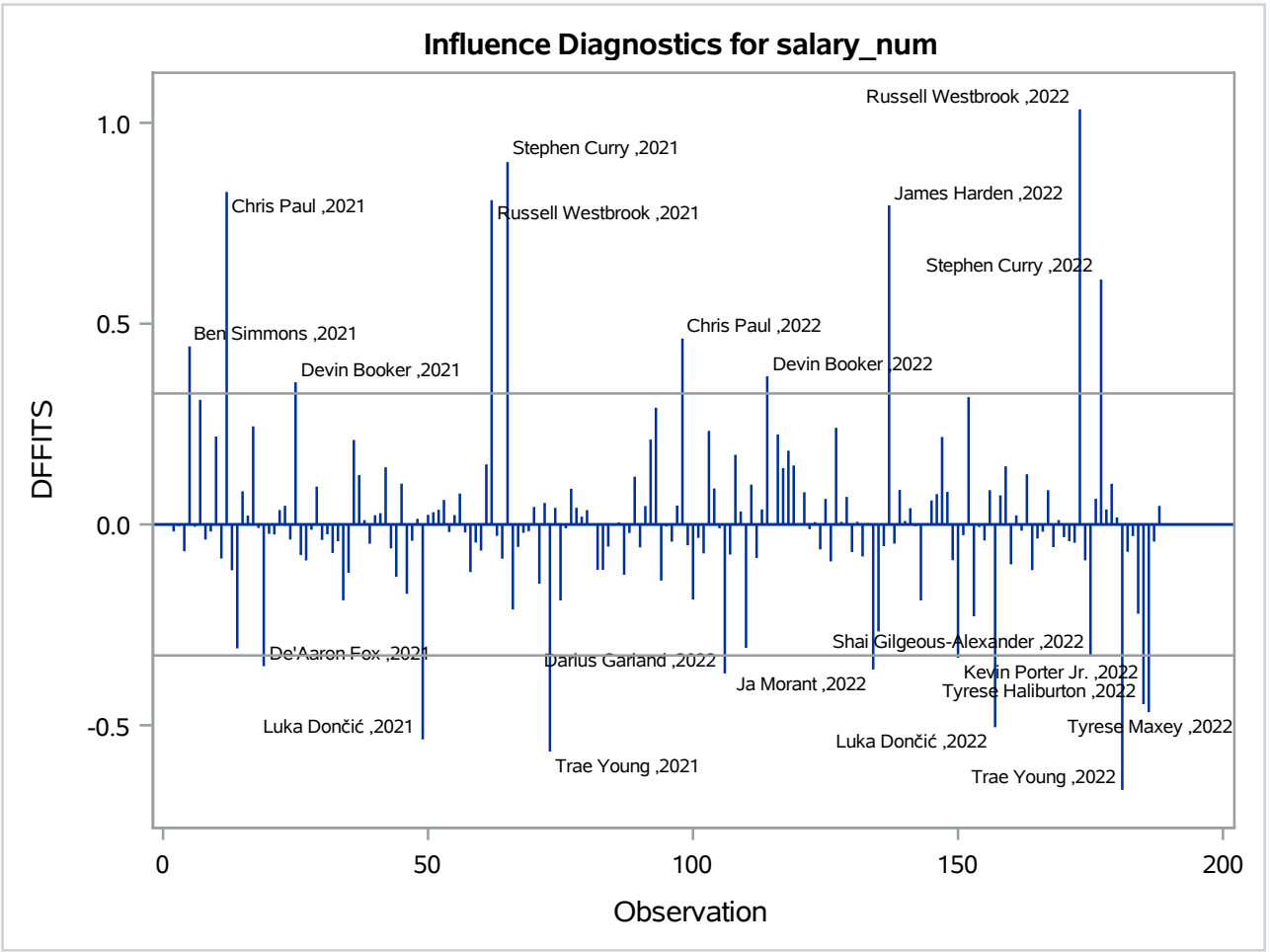
Model 1 Salary

The REG Procedure
Model: MODEL1
Dependent Variable: salary_num



Model 1 Salary

The REG Procedure
Model: MODEL1
Dependent Variable: salary_num



Model 1 Salary

The REG Procedure
Model: MODEL1
Dependent Variable: salary_num

The first constructed model using dependent variable salary and explanatory variables minutes per game (mpg), points per game (ptsg), assists per game (astg), and turnovers per game (tovg).

From the model analysis the explantory variables that are not statistically significant are turnovers per game (tovg), and minutes per game (mpg). The adjusted R-squared of this model is 0.4274 (Which measures the goodness-of-fit of a linear model) and the akaike information criterion (AIC) is 5959.7 (Which estimates therelative amount of information lost by a model).

Model 1 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary

Number of Observations Read	188
Number of Observations Used	188

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	85.64212	21.41053	41.50	<.0001
Error	183	94.40612	0.51588		
Corrected Total	187	180.04825			

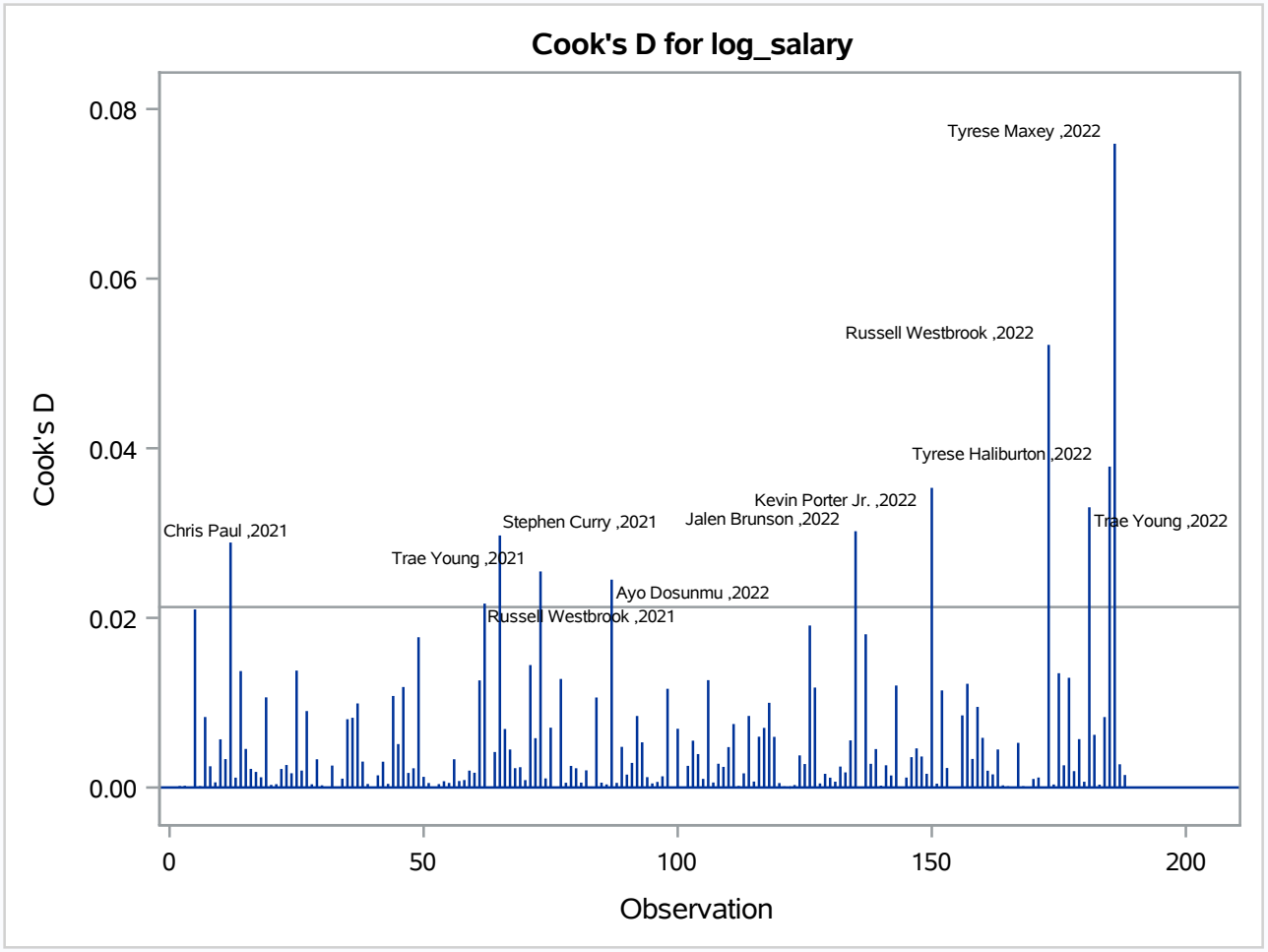
Root MSE	0.71825	R-Square	0.4757
Dependent Mean	15.62204	Adj R-Sq	0.4642
Coeff Var	4.59766		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	13.37487	0.25817	51.81	<.0001
mpg	1	0.06259	0.01561	4.01	<.0001
ptsg	1	0.04156	0.01992	2.09	0.0383
astg	1	0.12604	0.04696	2.68	0.0079
tovg	1	-0.24475	0.14917	-1.64	0.1026

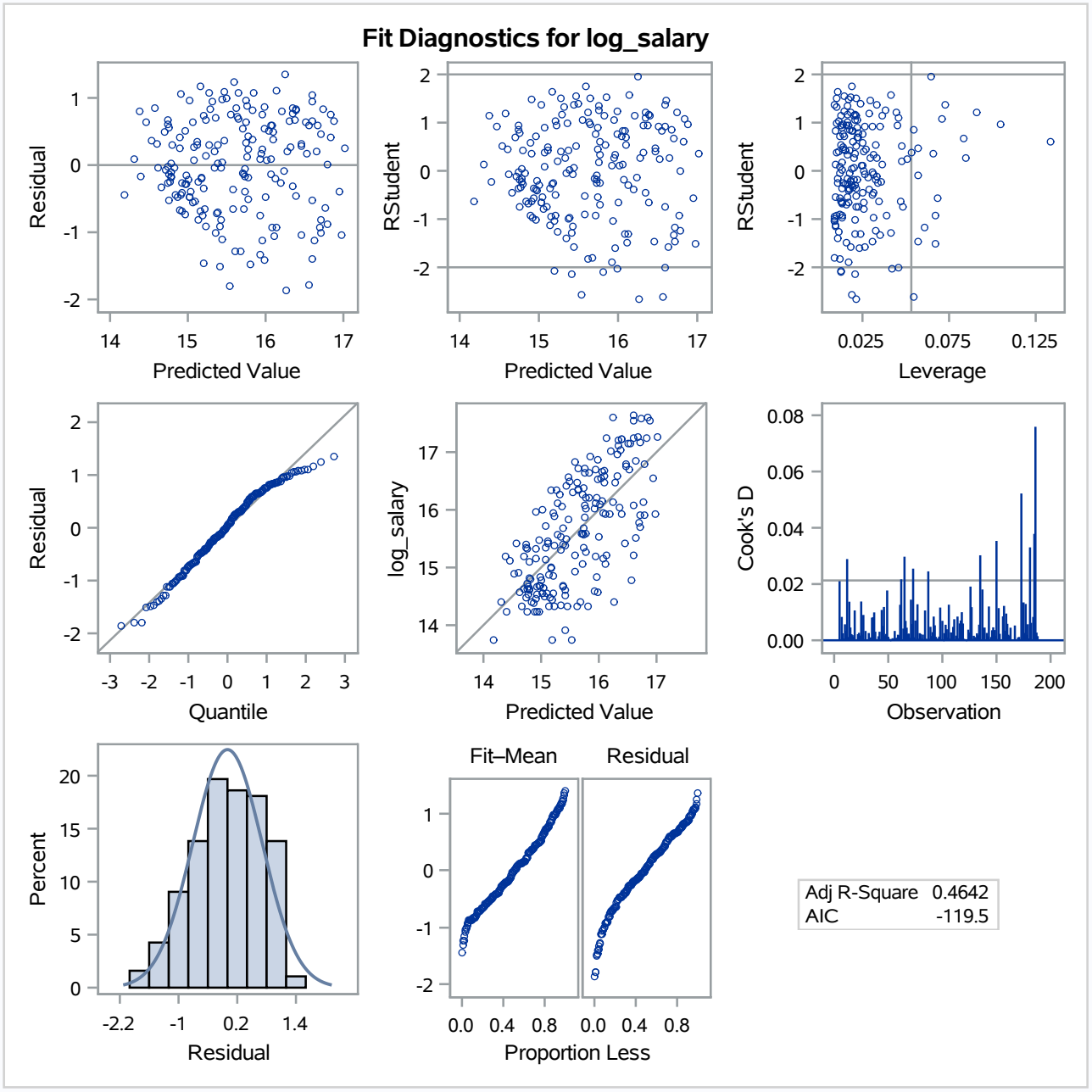
Model 1 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary



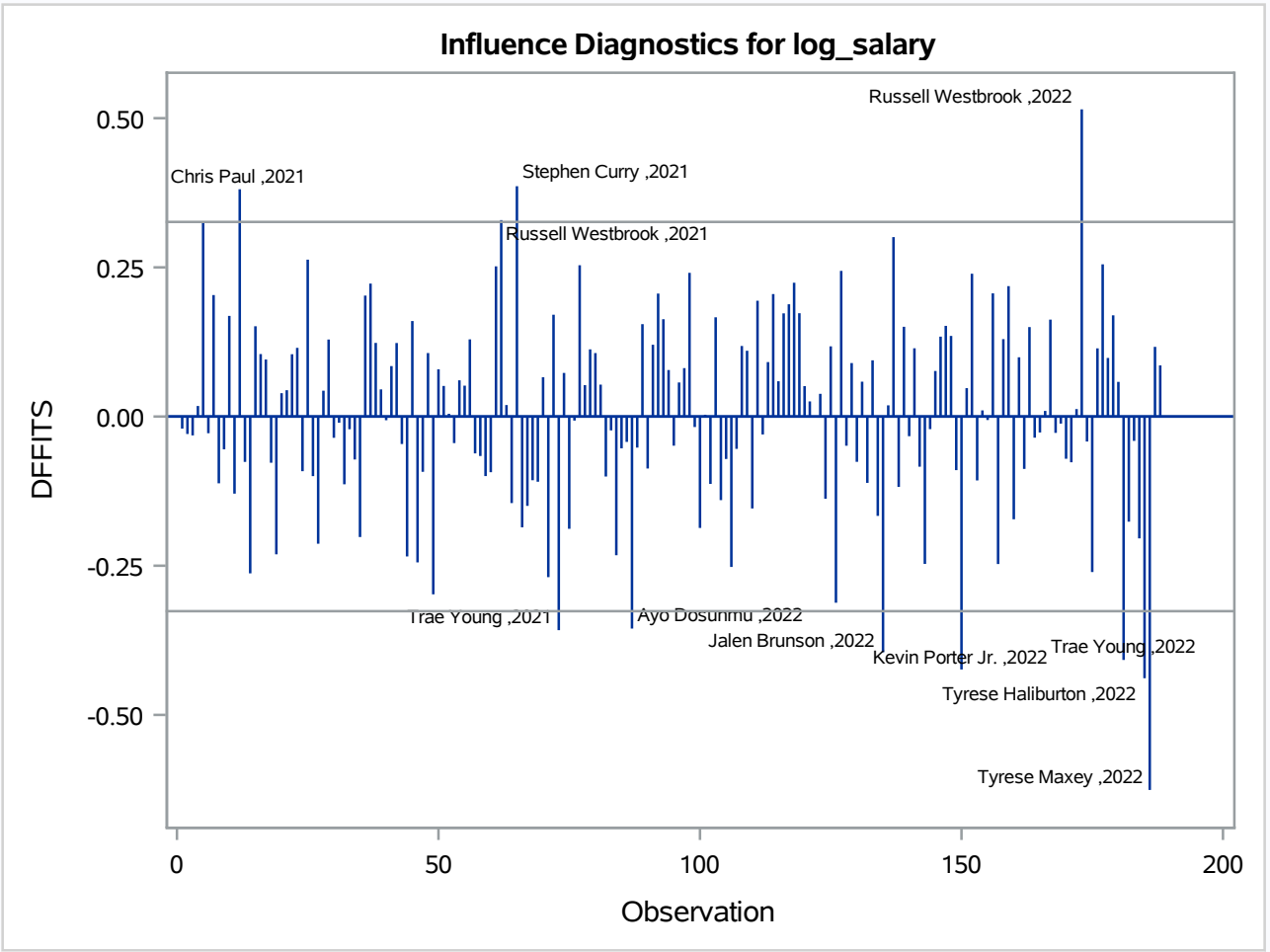
Model 1 Log Transform Salary

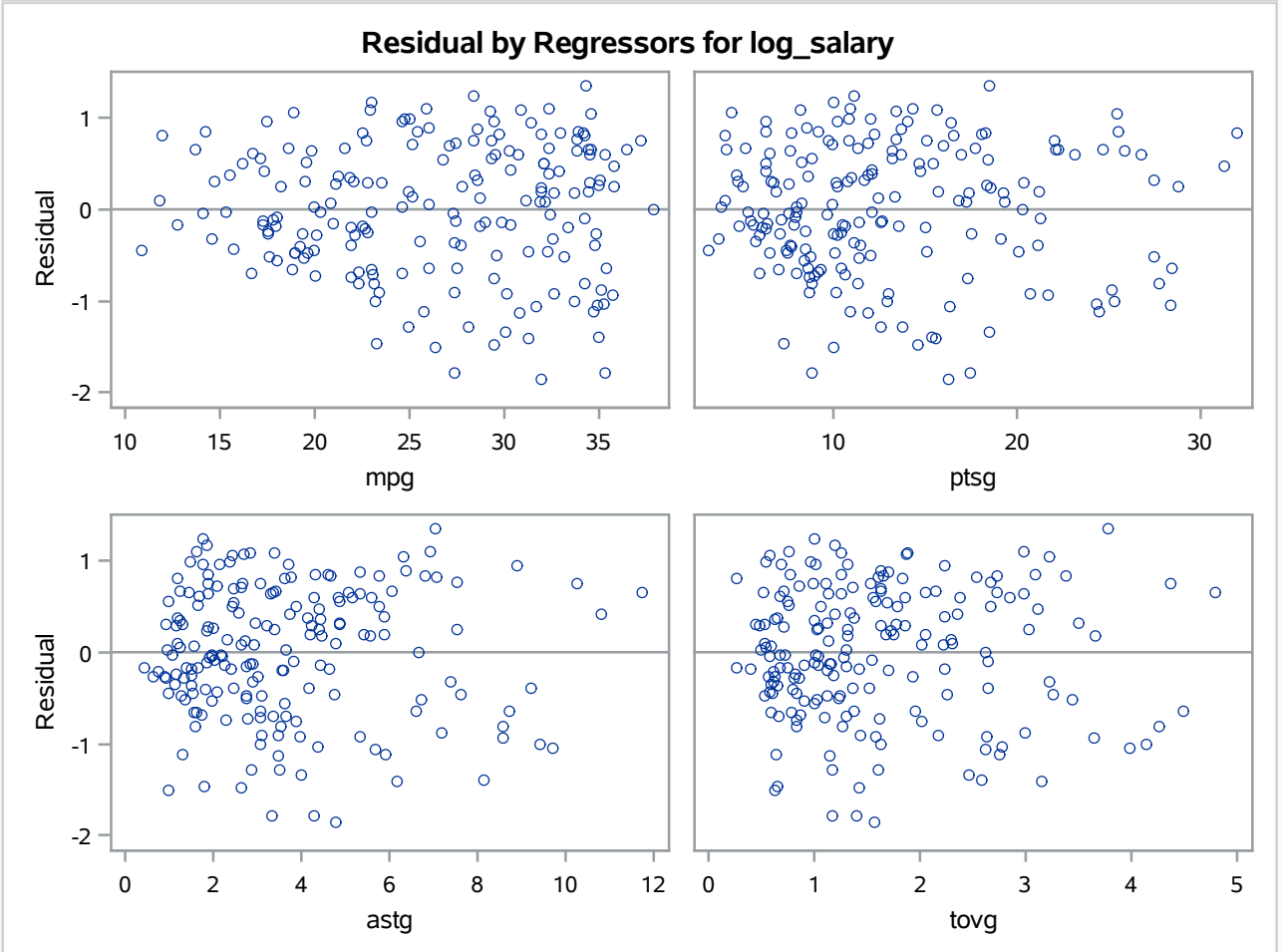
The REG Procedure
Model: MODEL1
Dependent Variable: log_salary



Model 1 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary





Model 1 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary

The first constructed using dependent variable log transformed salary and explanatory variables minutes per game (mpg), points per game (ptsg), assists per game (astg), and turnovers per game (tovg).

From the model analysis the only explanatory variable that is not statistically significant is turnovers per game (tovg).The adjusted R-squared of this model is 0.4642 and the akaike information criterion (AIC) is -119.5,which is an improvement on both measurements when using a log transformation on the dependent variable salary.

Model 2 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary

Number of Observations Read	186
Number of Observations Used	186

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	86.48138	21.62034	44.02	<.0001
Error	181	88.90756	0.49120		
Corrected Total	185	175.38894			

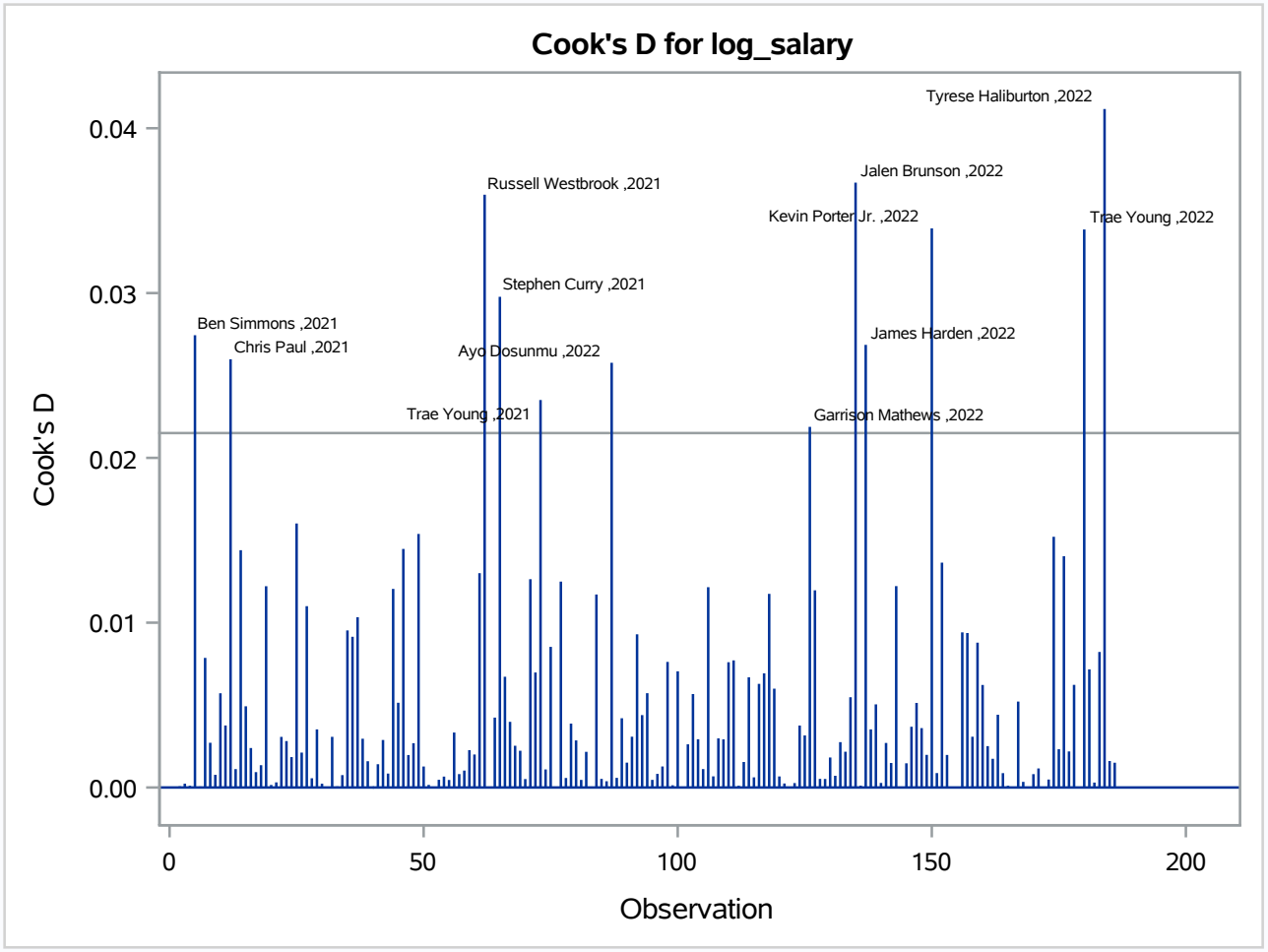
Root MSE	0.70086	R-Square	0.4931
Dependent Mean	15.61595	Adj R-Sq	0.4819
Coeff Var	4.48809		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	13.32076	0.25389	52.47	<.0001
mpg	1	0.06504	0.01534	4.24	<.0001
ptsg	1	0.05202	0.01975	2.63	0.0092
astg	1	0.14790	0.04630	3.19	0.0017
tovg	1	-0.38382	0.15164	-2.53	0.0122

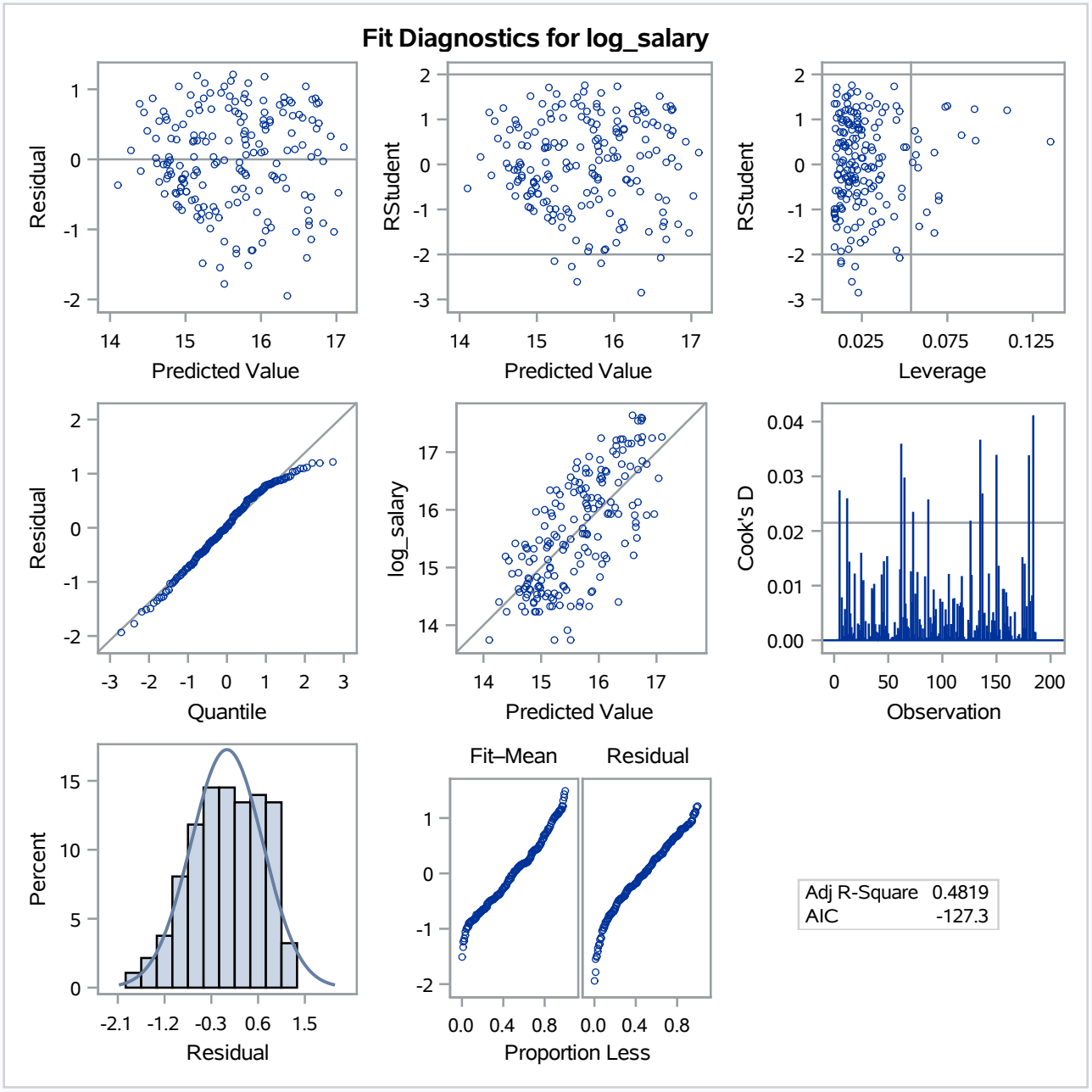
Model 2 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary



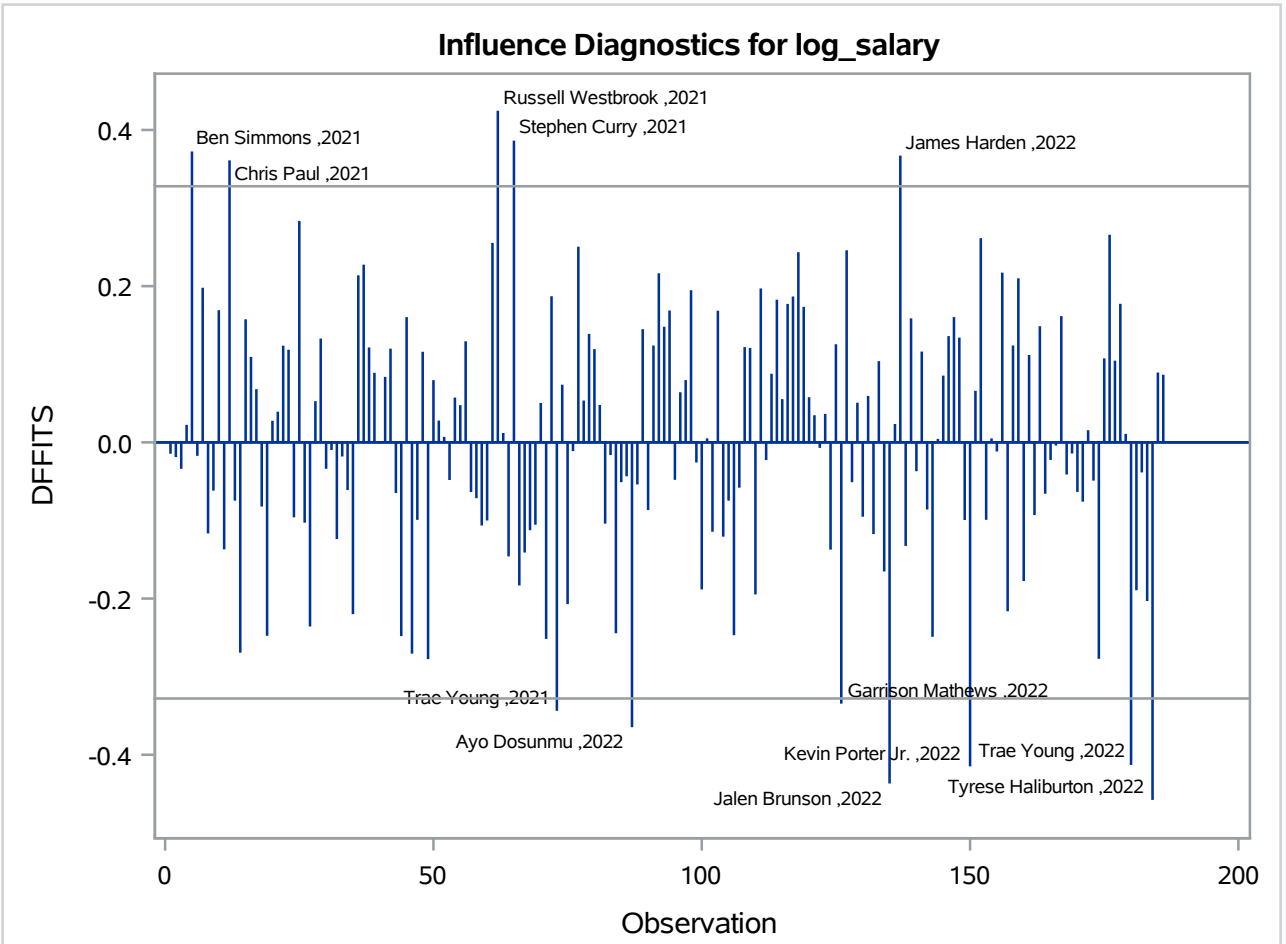
Model 2 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary



Model 2 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary



Model 2 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary

The second constructed using dependent variable log transformed salary and explanatory variables minutes per game (mpg), points per game (ptsg), assists per game (astg), and turnovers per game (tovg).

From the model analysis after the removal of two outliers all of the explanatory variables are statically significant.The adjusted R-squared of this model is 0.4819 and the akaike information criterion (AIC) is -127.3,which is an improvement on both measurements after the removal of outliers.

Model 3 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary

Number of Observations Read	183
Number of Observations Used	183

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	87.80366	21.95091	47.50	<.0001
Error	178	82.25685	0.46212		
Corrected Total	182	170.06051			

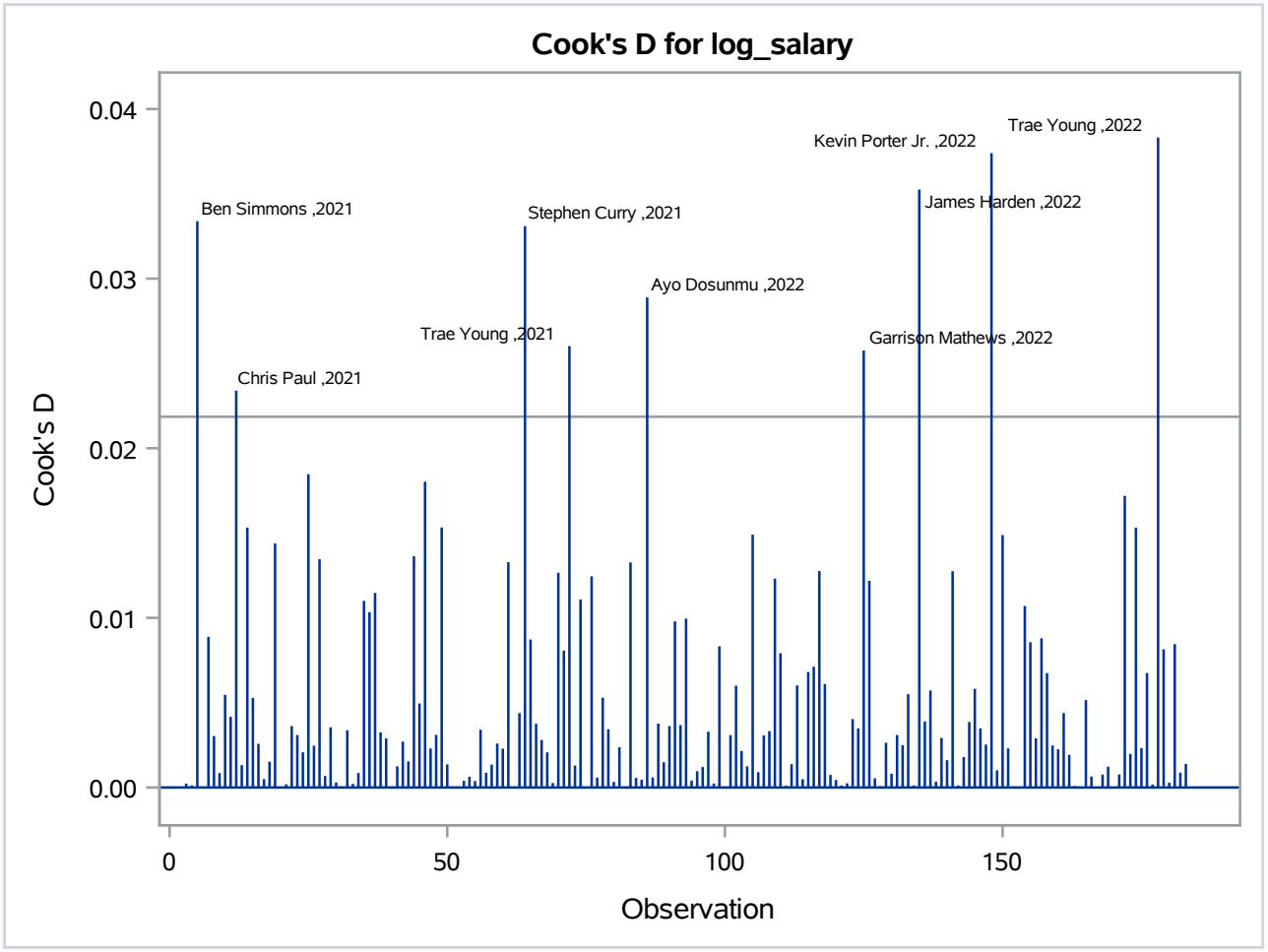
Root MSE	0.67979	R-Square	0.5163
Dependent Mean	15.61430	Adj R-Sq	0.5054
Coeff Var	4.35365		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	13.23279	0.24787	53.39	<.0001
mpg	1	0.07056	0.01497	4.71	<.0001
ptsg	1	0.05422	0.01945	2.79	0.0059
astg	1	0.16699	0.04554	3.67	0.0003
tovg	1	-0.47161	0.15011	-3.14	0.0020

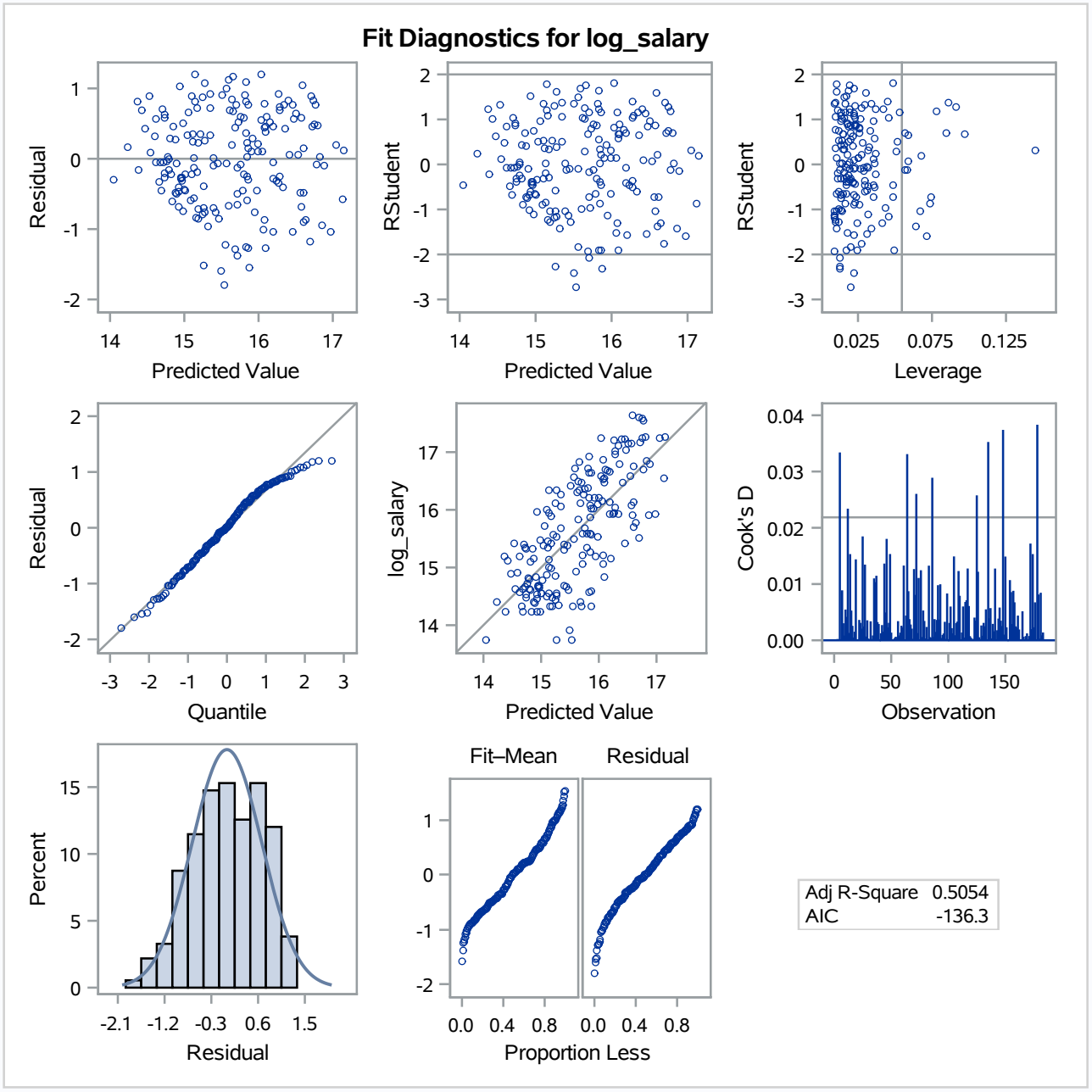
Model 3 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary



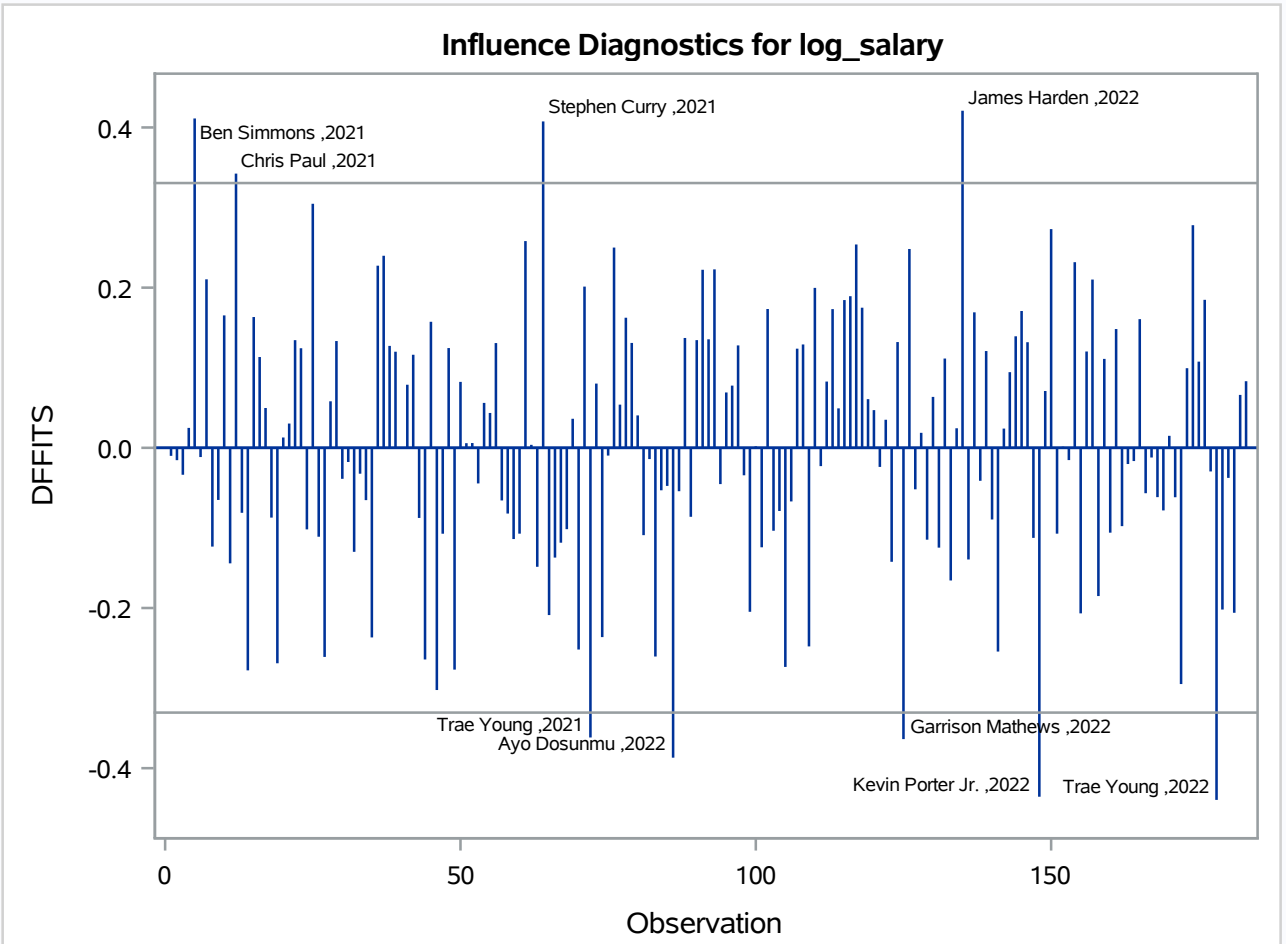
Model 3 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary



Model 3 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary



Model 3 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary

The third constructed model using dependent variable log transformed salary and explanatory variables minutes per game (mpg), points per game (ptsg), assists per game (astg), and turnovers per game (tovg).

From the model analysis after the removal of three more outliers all of the explanatory variables are statically significant.The adjusted R-squared of this model is 0.5054 and the akaike information criterion (AIC) is -136.3,which is an improvement on both measurements after the removal of outliers.

Model 4 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary

Number of Observations Read	180
Number of Observations Used	180

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	86.17218	21.54304	47.87	<.0001
Error	175	78.76170	0.45007		
Corrected Total	179	164.93388			

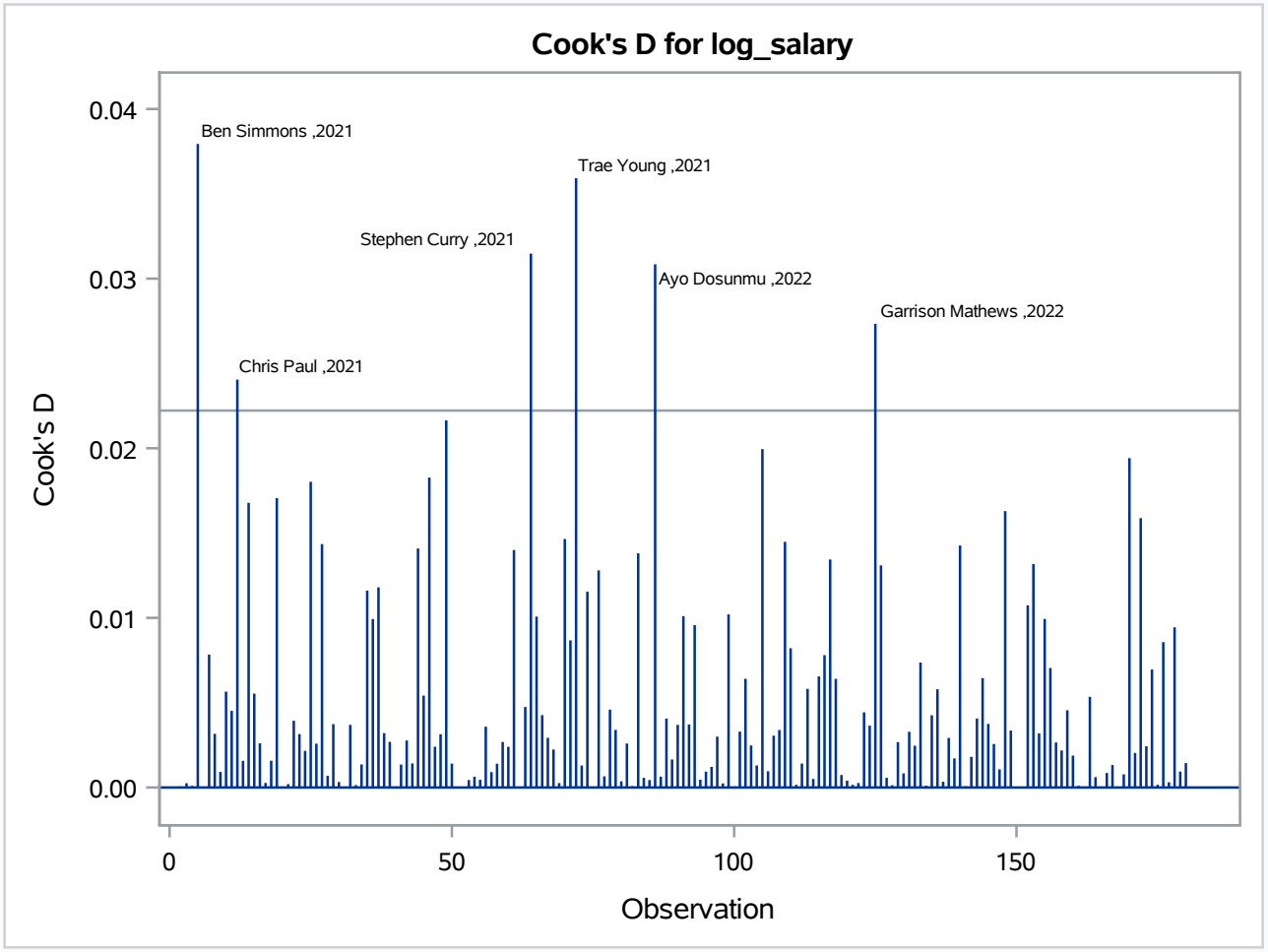
Root MSE	0.67087	R-Square	0.5225
Dependent Mean	15.60730	Adj R-Sq	0.5115
Coeff Var	4.29844		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	13.25437	0.24535	54.02	<.0001
mpg	1	0.06823	0.01493	4.57	<.0001
ptsg	1	0.05545	0.01977	2.80	0.0056
astg	1	0.16781	0.04546	3.69	0.0003
tovg	1	-0.45313	0.15277	-2.97	0.0034

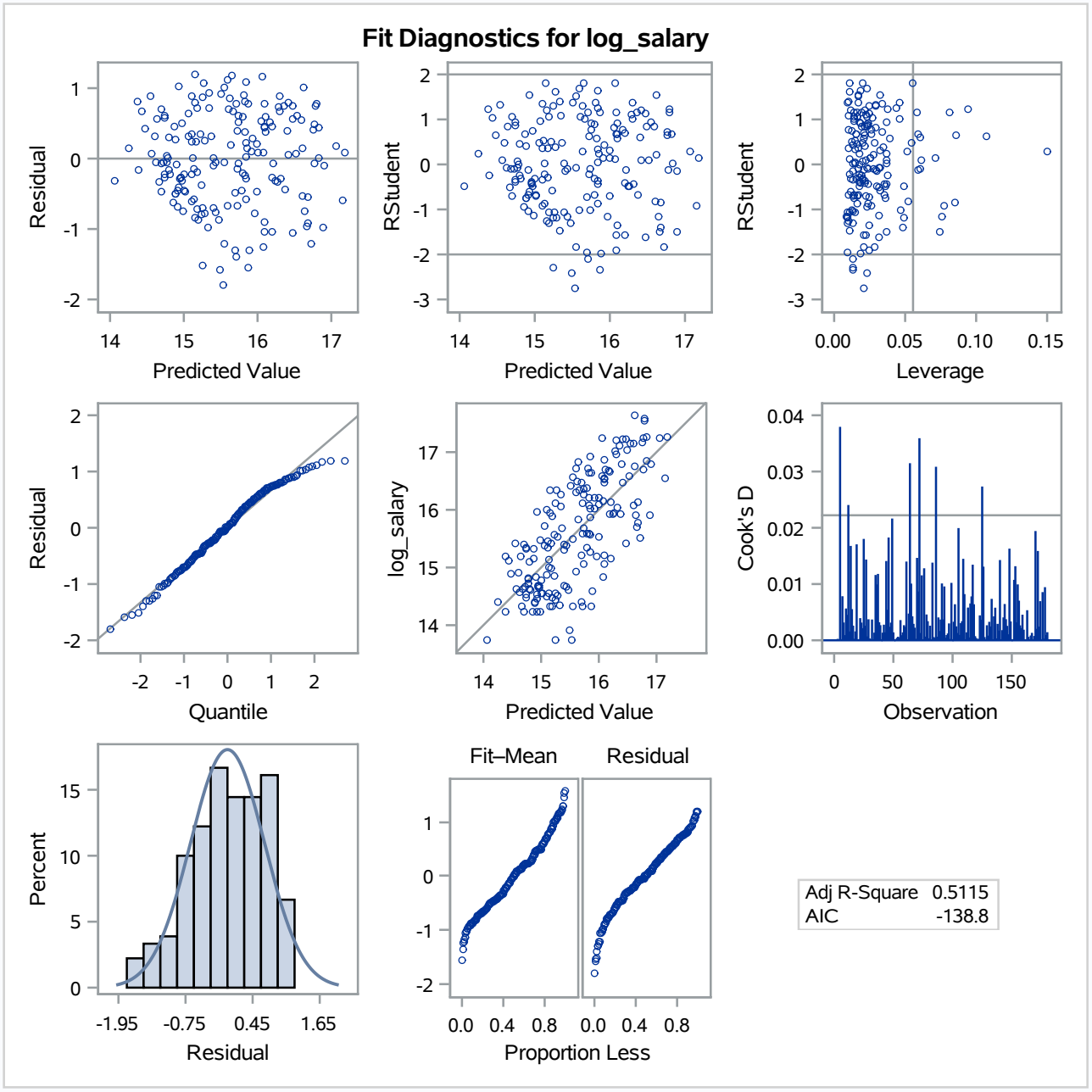
Model 4 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary



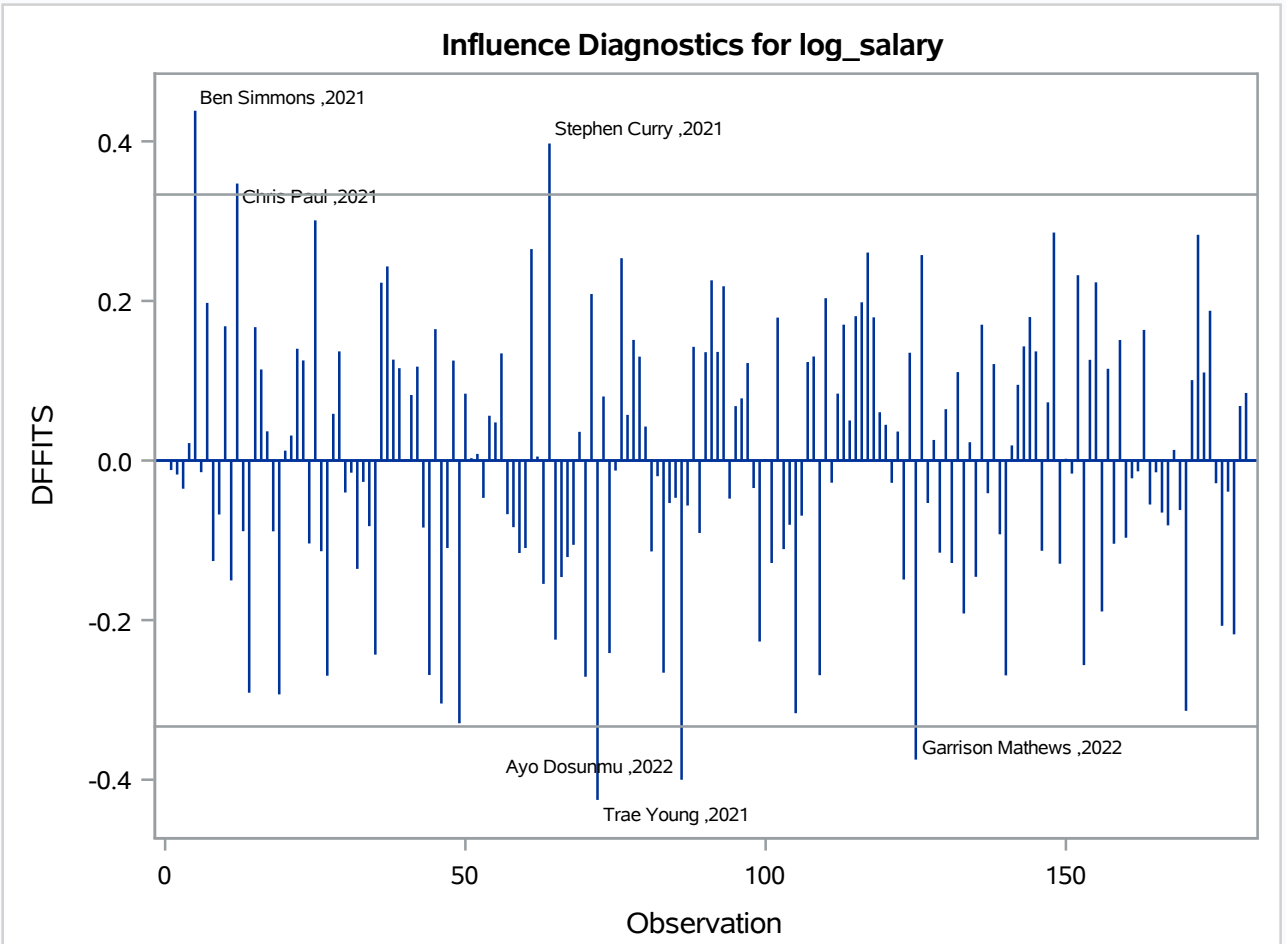
Model 4 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary



Model 4 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary



Model 4 Log Transform Salary

The REG Procedure
Model: MODEL1
Dependent Variable: log_salary

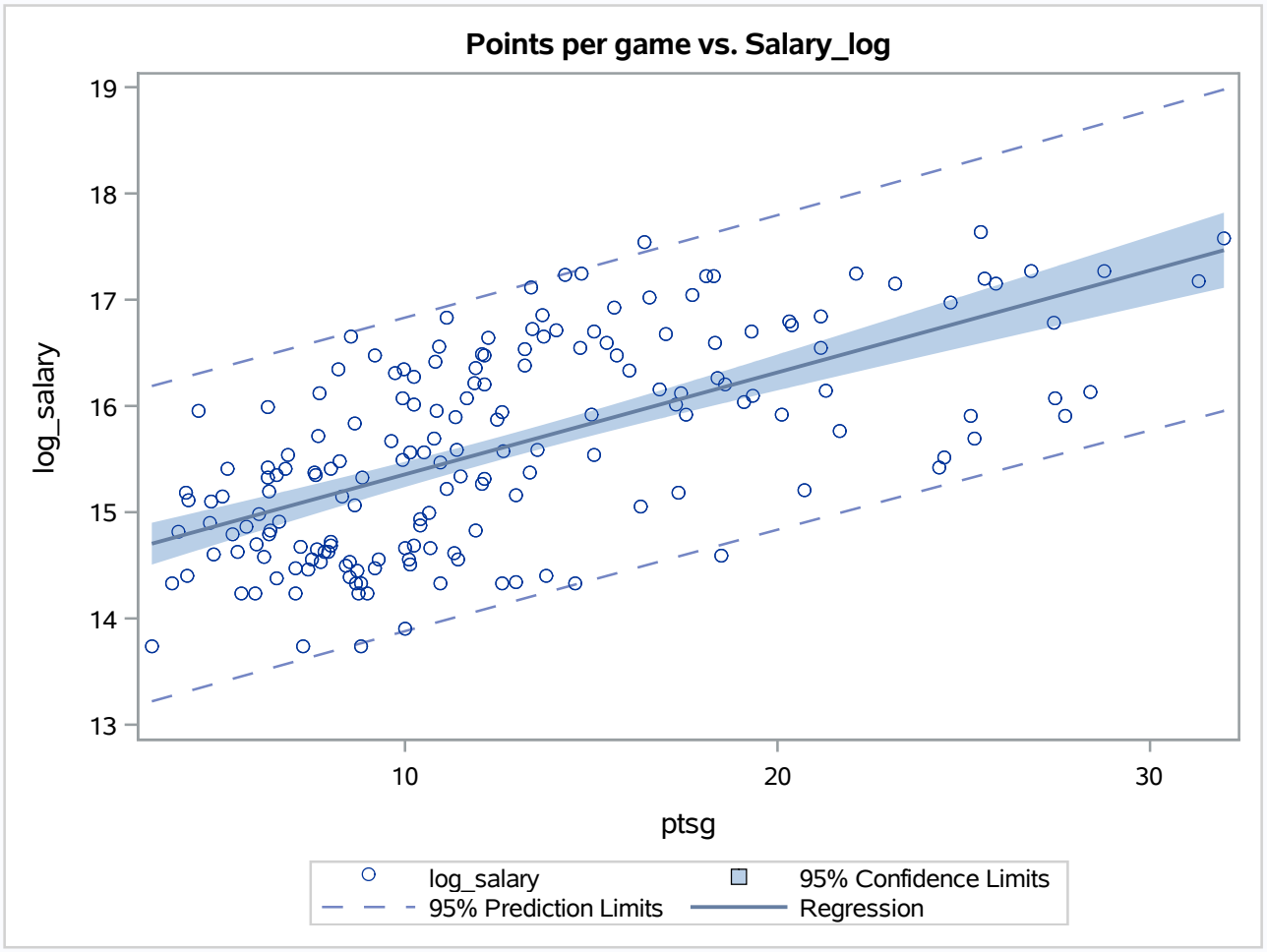
The fourth constructed model using dependent variable log transformed salary and explanatory variables minutes per game (mpg), points per game (ptsg), assists per game (astg), and turnovers per game (tovg).

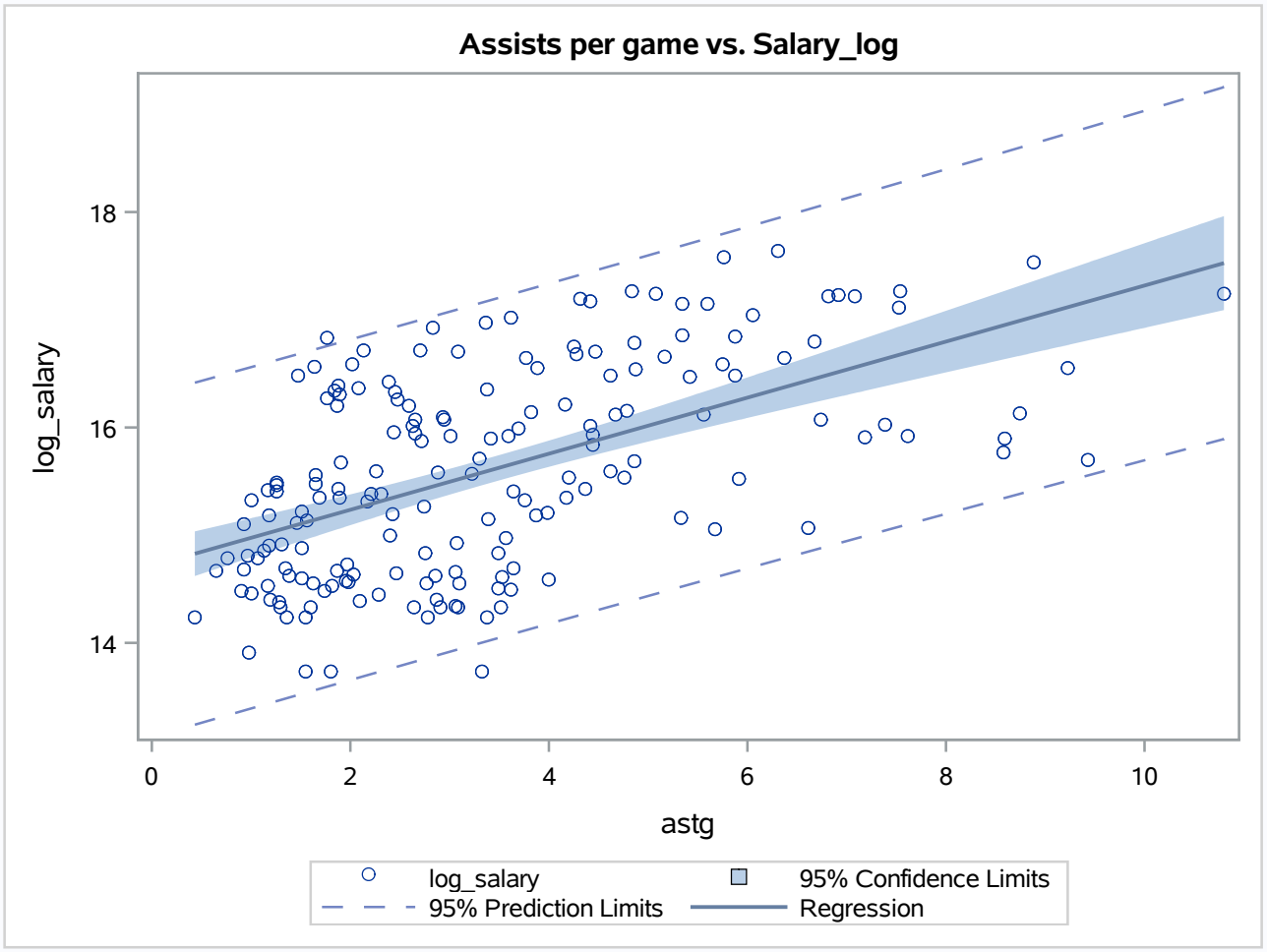
From the model analysis after the removal of three more outliers all of the explanatory variables are statically significant.The adjusted R-squared of this model is 0.5115 and the akaike information criterion (AIC) is -138.8,which is an improvement on both measurements after the removal of outliers.

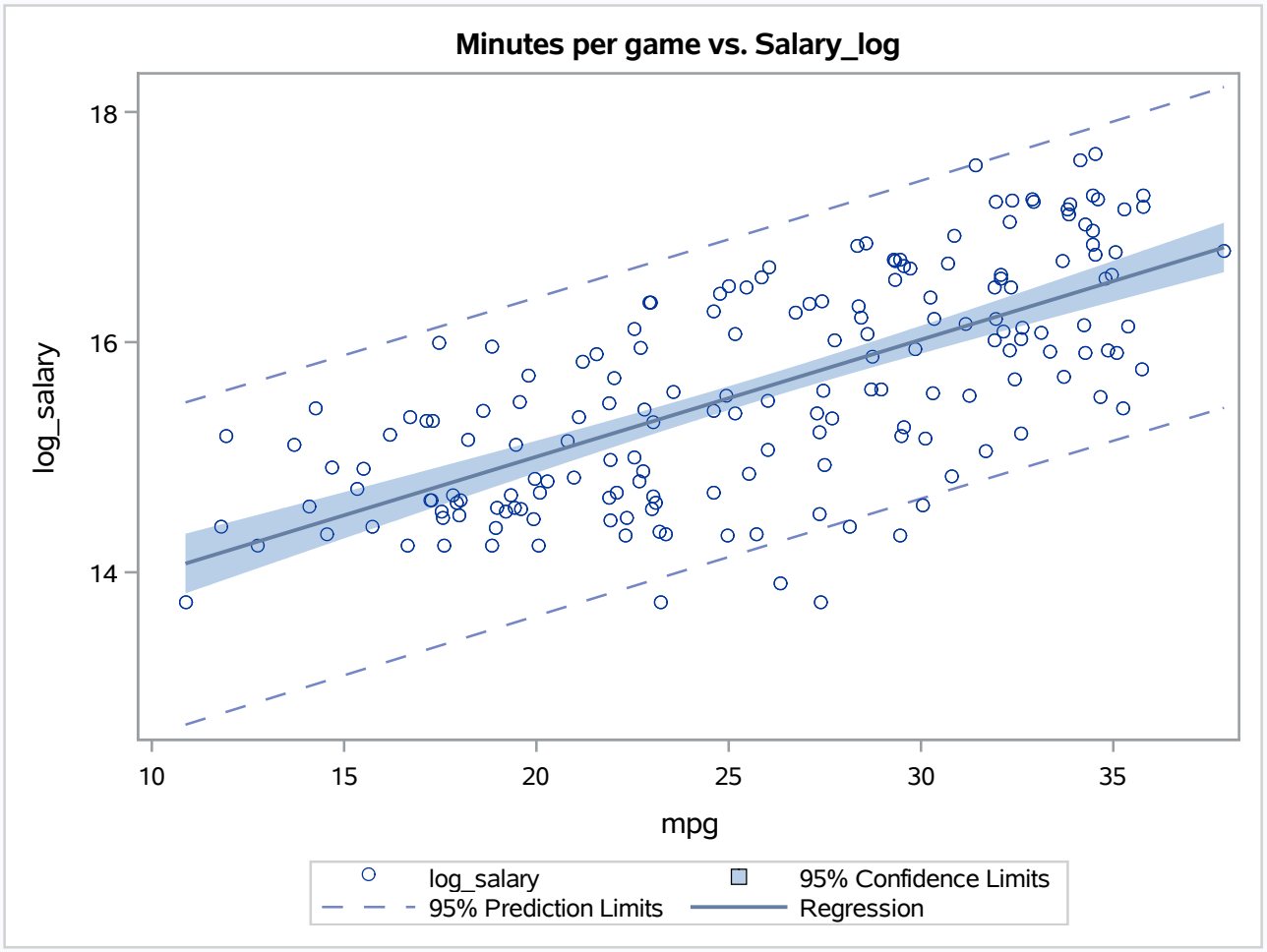
Magic Johnson Prediciton

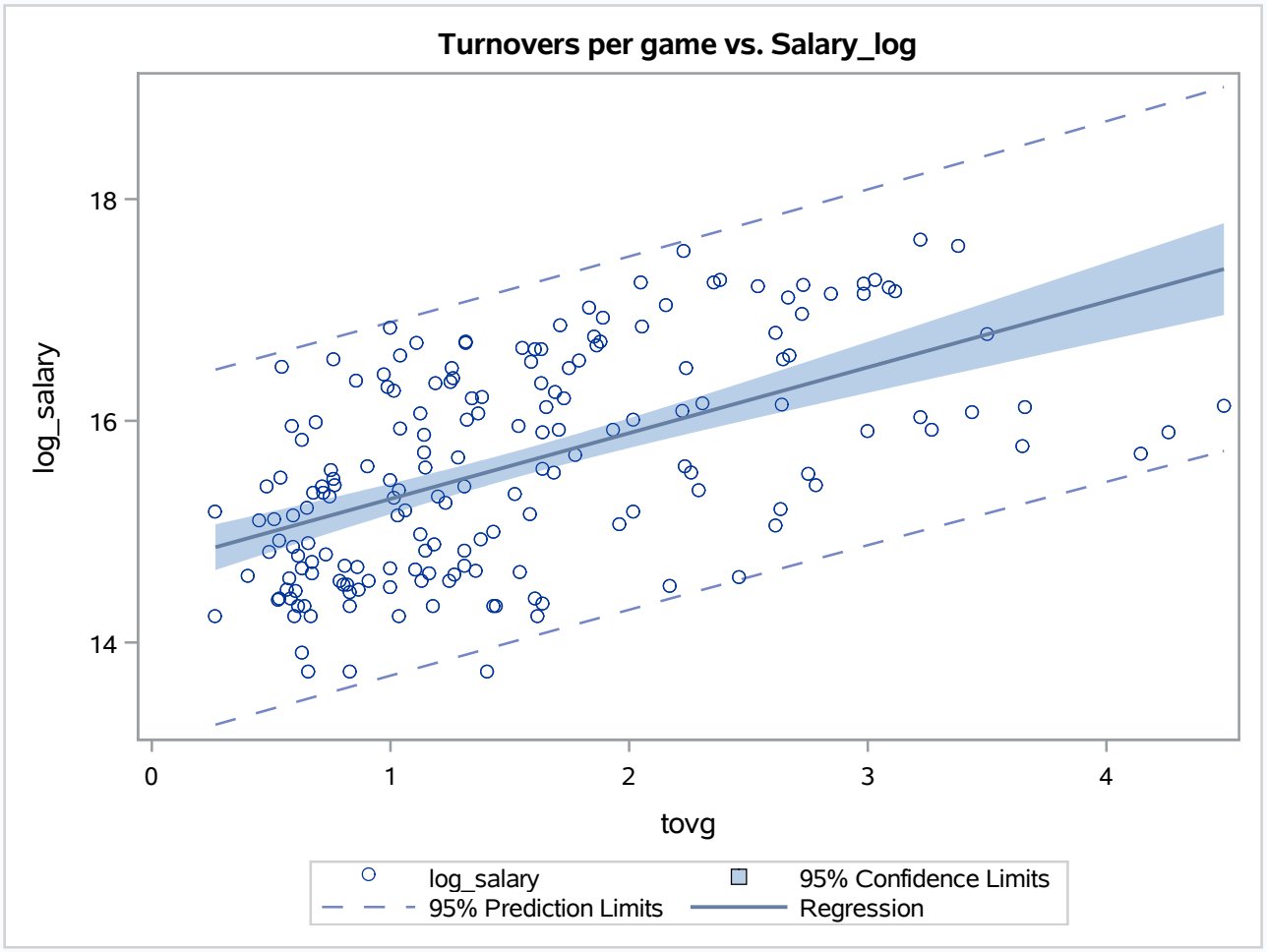
Obs	intercept	pts	ast	to	mp	salary
1	13.2544	19.5	11.2	3.9	36.7	23,021,774

Using the selected fourth model to make a prediction on the salary if Earvin (Magic) Johnson was to play in the NBA today. In his 13 seasons in the NBA he recorded a average of 36.7 minutes per game (mpg), 19.5 points per game (ptsg), 11.2 assists per game (astg), and 3.9 turnovers per game (tovg). The model predicts that he would earn $23,021,774 (USD).


Points per game vs. Salary_log
log_salary
19
18
17
16
15
14
13
10
20
30
ptsg
log_salary
95% Confidence Limits
95% Prediction Limits
Regression


Assists per game vs. Salary_log
log_salary
18
16
14
0
2
4
6
8
10
astg
log_salary
95% Confidence Limits
95% Prediction Limits
Regression


Minutes per game vs. Salary_log
log_salary
18
16
14
10
15
20
25
30
35
mpg
log_salary
95% Confidence Limits
95% Prediction Limits
Regression



Plots of variables chosen for model.

Removed Outliers

Player_name	Age	salary_num	mpg	ptsg	astg	tovg	Year
Russell Westbrook	33	44,211,146	34.333	18.474	7.051	3.782	2022
James Harden	32	43,848,000	37.215	22.031	10.262	4.369	2022
Russell Westbrook	32	41,358,814	36.446	22.231	11.738	4.800	2021
Trae Young	23	8,326,471	34.895	28.355	9.697	3.987	2022
Tyrese Haliburton	21	4,023,600	35.000	15.338	8.156	2.584	2022
Tyrese Maxey	21	2,602,920	35.333	17.480	4.280	1.173	2022
Kevin Porter Jr.	21	2,130,240	31.262	15.557	6.164	3.148	2022
Jalen Brunson	25	1,802,057	31.949	16.266	4.772	1.570	2022