Using SPSS for Multiple Regression

UDP 520 Lab 8

Lin Lin

December 6th , 2007

UNIVERSITY OF WASHINGTON

Research Question

- What factors are associated with BMI?
- Predict BMI.

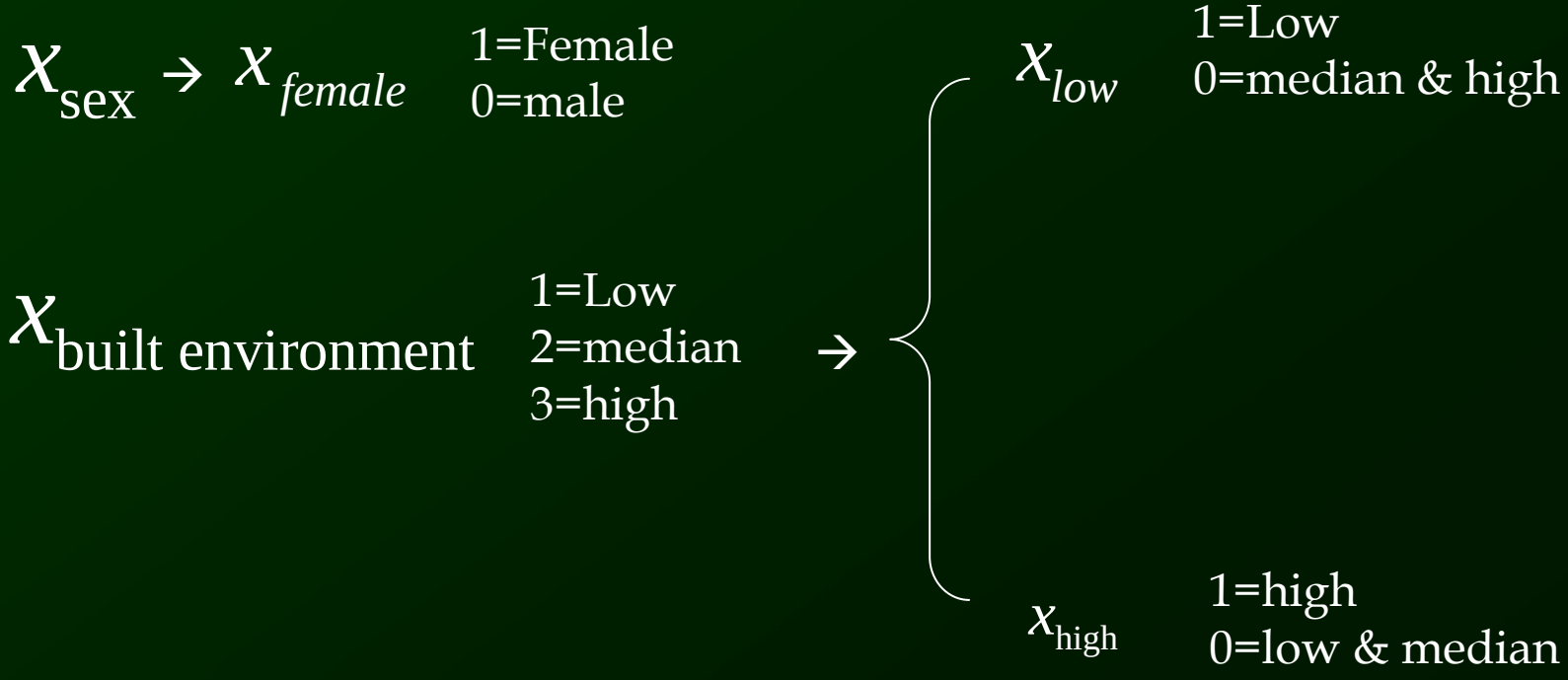
OLS Equation

- Multiple regression

$$Y_{\text{BMI}} = \beta_0 + \beta_1 x_{\text{calorie}} + \beta_2 x_{\text{exercise}} + \beta_3 x_{\text{sex}} + \beta_4 x_{\text{income}} + \beta_5 x_{\text{education}} + \beta_6 x_{\text{built environment}} + \varepsilon$$

Including Nominal Independent in OLS Model

- To include nominal independents in the OLS model in SPSS, nominal variables with n categories have been recoded to n-1 dummy variables

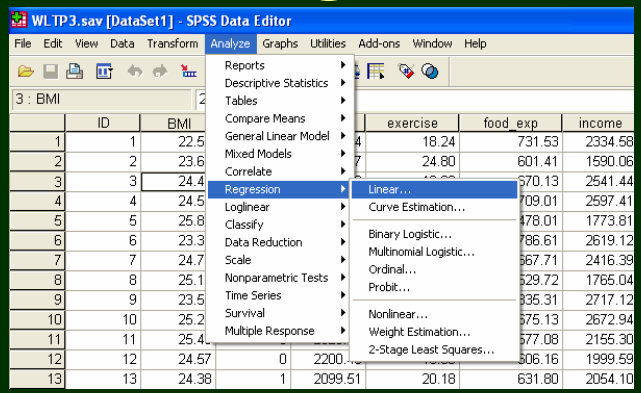


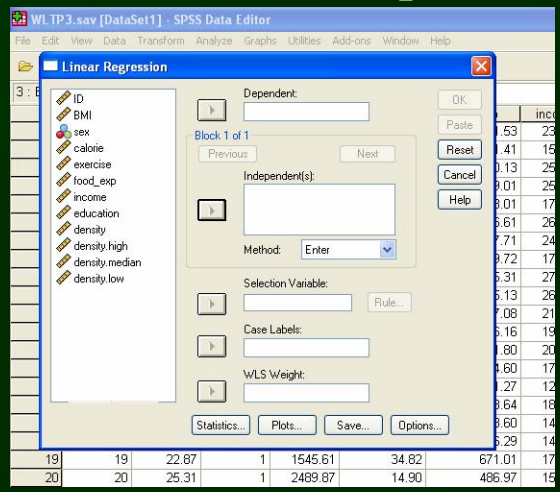
$$Y_{\text{BMI}} = \beta_0 + \beta_1 x_{\text{calorie}} + \beta_2 x_{\text{exercise}} + \beta_3 x_{\text{female}} + \beta_4 x_{\text{income}} + \beta_5 x_{\text{education}} + \beta_6 x_{\text{high}} + \beta_7 x_{\text{low}} + \varepsilon$$

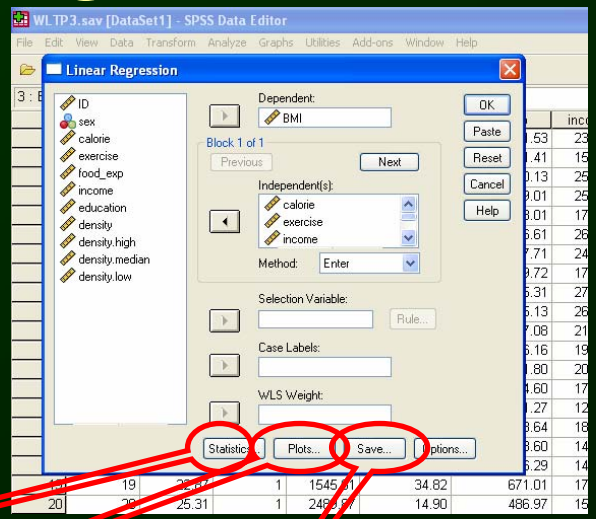
Median was excluded from the regression model.

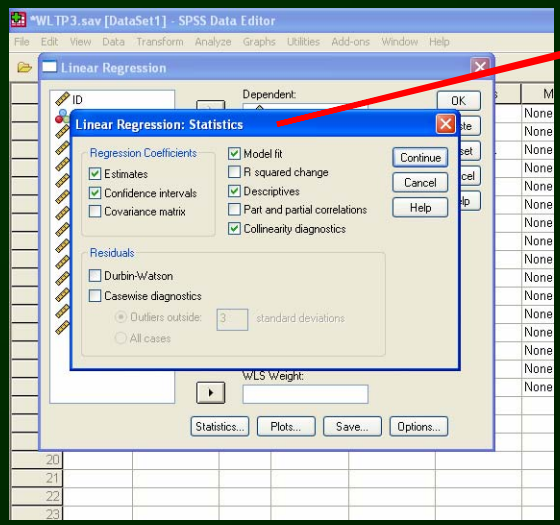
Median was used as a “reference” category.

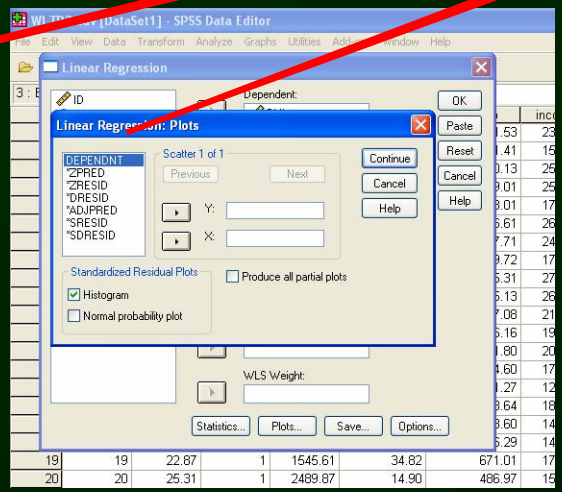
Using SPSS for Multiple Regression

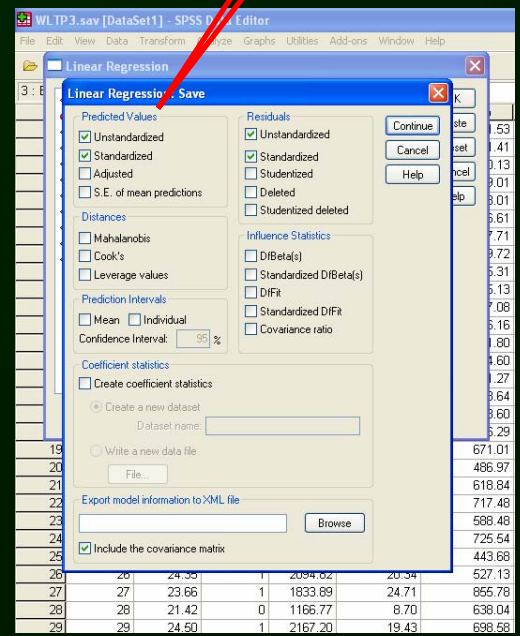












SPSS Output Tables

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	20.778	.212		97.779	.000	20.361	21.195		
	calorie	.002	.000	.754	39.106	.000	.002	.002	.962	1.039
	exercise	-.027	.003	-.162	-8.393	.000	-.034	-.021	.959	1.043
	sex	.028	.049	.011	.580	.562	-.067	.124	.997	1.003
	income	8.39E-005	.000	.033	1.751	.080	.000	.000	.992	1.008
	education	-.001	.006	-.003	-.176	.860	-.014	.011	.991	1.009
	density.high	-.176	.061	-.064	-2.894	.004	-.295	-.057	.734	1.363
	density.low	-.083	.059	-.031	-1.401	.161	-.199	.033	.733	1.364

a. Dependent Variable: BMI

- High: controlling for other variables constant, people who live in high density area have 0.176 lower BMI than people who live in median density area.

- Low: controlling for other variables constant, people who live in low density area have 0.083 lower BMI than people who live in median density area.

Controlling for other variables constant, females have 0.028 higher BMI than males.

Descriptive Statistics

	Mean	Std. Deviation	N
BMI	24.0674	1.28663	1000
calorie	2017.7167	513.71981	1000
exercise	21.7947	7.66196	1000
sex	.51	.500	1000
income	2005.1981	509.49088	1000
education	19.95	3.820	1000
density.high	.3200	.46671	1000
density.low	.3600	.48024	1000

Correlations

		BMI	calorie	exercise	sex	income	education	density.high	density.low
Pearson Correlation	BMI	1.000	.784	-.310	.018	.033	.011	-.043	-.005
	calorie	.784	1.000	-.193	.004	-.009	.004	.012	-.011
	exercise	-.310	-.193	1.000	-.011	-.030	-.046	.021	-.007
	sex	.018	.004	-.011	1.000	.038	.024	.006	-.028
	income	.033	-.009	-.030	.038	1.000	.069	-.028	.012
	education	.011	.004	-.046	.024	.069	1.000	-.010	-.033
	density.high	-.043	.012	.021	.006	-.028	-.010	1.000	-.514
	density.low	-.005	-.011	-.007	-.028	.012	-.033	-.514	1.000
Sig. (1-tailed)	BMI	.	.000	.000	.287	.148	.361	.086	.432
	calorie	.000	.	.000	.444	.391	.451	.355	.358
	exercise	.000	.000	.	.362	.175	.072	.258	.417
	sex	.287	.444	.362	.	.113	.222	.425	.185
	income	.148	.391	.175	.113	.	.014	.186	.353
	education	.361	.451	.072	.222	.014	.	.382	.151
	density.high	.086	.355	.258	.425	.186	.382	.	.000
	density.low	.432	.358	.417	.185	.353	.151	.000	.
N	BMI	1000	1000	1000	1000	1000	1000	1000	1000
	calorie	1000	1000	1000	1000	1000	1000	1000	1000
	exercise	1000	1000	1000	1000	1000	1000	1000	1000
	sex	1000	1000	1000	1000	1000	1000	1000	1000
	income	1000	1000	1000	1000	1000	1000	1000	1000
	education	1000	1000	1000	1000	1000	1000	1000	1000
	density.high	1000	1000	1000	1000	1000	1000	1000	1000
	density.low	1000	1000	1000	1000	1000	1000	1000	1000

Variables Entered/Removed(b)

Model	Variables Entered	Variables Removed	Method
1	density.low, exercise, sex, income, education, calorie, density.high(a)	.	Enter

a All requested variables entered.
b Dependent Variable: BMI

Model Summary(b)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.803(a)	.646	.643	.76873

a Predictors: (Constant), density.low, exercise, sex, income, education, calorie, density.high
b Dependent Variable: BMI

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1067.546	7	152.507	258.069	.000(a)
	Residual	586.225	992	.591		
	Total	1653.771	999			

a Predictors: (Constant), density.low, exercise, sex, income, education, calorie, density.high
b Dependent Variable: BMI

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	20.778	.212		97.779	.000	20.361	21.195		
	calorie	.002	.000	.754	39.106	.000	.002	.002	.962	1.039
	exercise	-.027	.003	-.162	-8.393	.000	-.034	-.021	.959	1.043
	sex	.028	.049	.011	.580	.562	-.067	.124	.997	1.003
	income	8.39E-005	.000	.033	1.751	.080	.000	.000	.992	1.008
	education	-.001	.006	-.003	-.176	.860	-.014	.011	.991	1.009
	density.high	-.176	.061	-.064	-2.894	.004	-.295	-.057	.734	1.363
	density.low	-.083	.059	-.031	-1.401	.161	-.199	.033	.733	1.364

a Dependent Variable: BMI

Collinearity Diagnostics(a)

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions							
				(Constant)	calorie	exercise	sex	income	education	density.high	density.low
1	1	6.036	1.000	.00	.00	.00	.01	.00	.00	.00	.01
	2	1.000	2.456	.00	.00	.00	.00	.00	.00	.26	.22
	3	.465	3.604	.00	.00	.00	.92	.00	.00	.05	.04
	4	.278	4.657	.00	.01	.02	.06	.01	.00	.66	.71
	5	.110	7.411	.00	.10	.70	.00	.02	.01	.00	.00
	6	.060	10.041	.00	.41	.04	.00	.56	.00	.00	.00
	7	.041	12.192	.01	.21	.05	.00	.25	.56	.00	.00
	8	.010	24.198	.99	.27	.19	.00	.15	.43	.02	.02

a Dependent Variable: BMI

Residuals Statistics(a)

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	21.8204	26.9458	24.0674	1.03374	1000
Residual	-3.29477	4.89907	.00000	.76604	1000
Std. Predicted Value	-2.174	2.784	.000	1.000	1000
Std. Residual	-4.286	6.373	.000	.996	1000

a Dependent Variable: BMI