Multinomial Logistic Regression
(categorical and continuous predictors)

Dr Daniel Boduszek
University of Huddersfield
d.boduszek@hud.ac.uk

University of
HUDDERSFIELD

Presentation Outline

- Introduction
- SPSS procedure
- Interpretation of SPSS output
- Presenting results

Intro to Multinomial Logistic Regression (MLR)

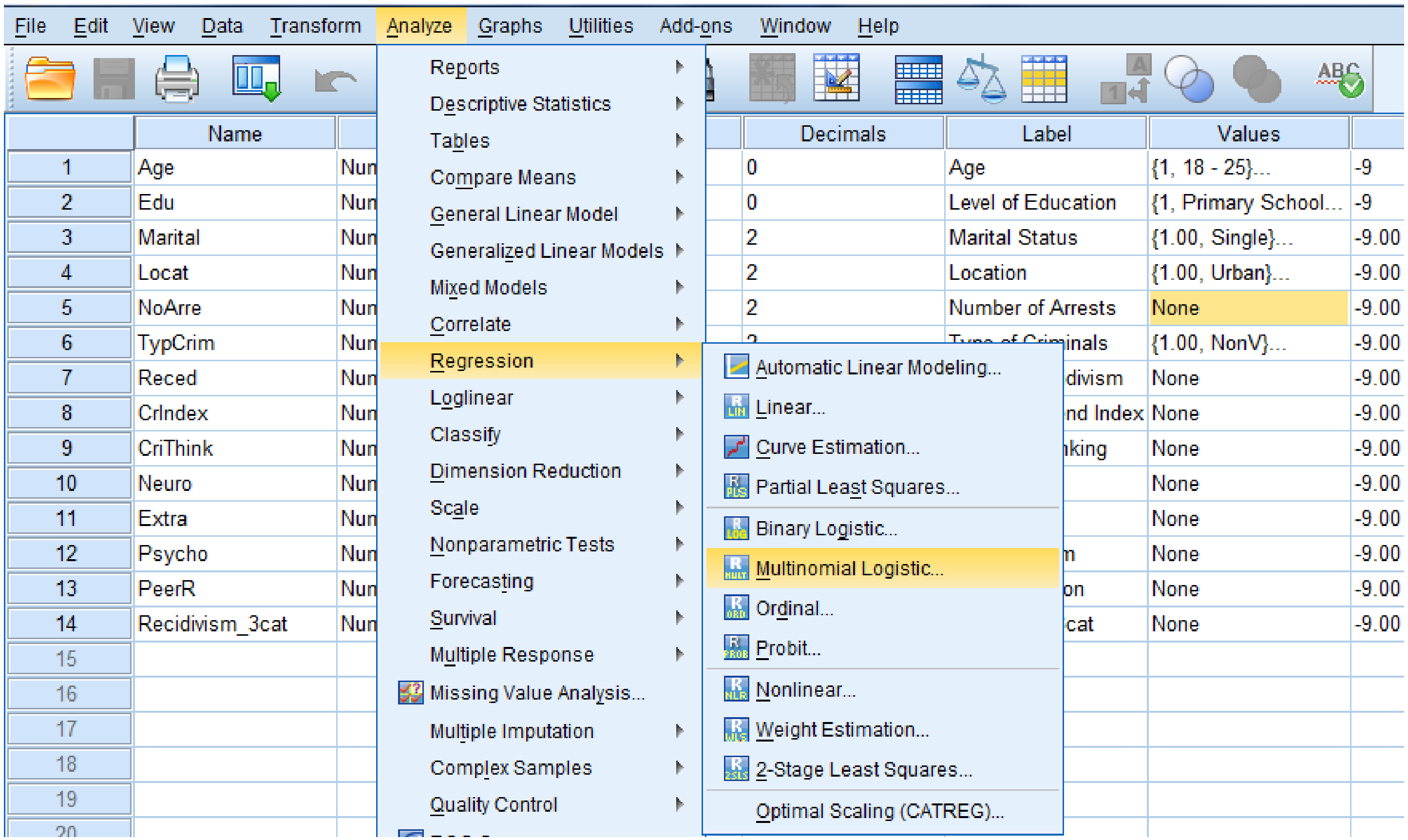
- MLR is an extension of binary logistic regression
- MLR is appropriate when the outcome variable is categorical with more than two categories and the predictors are of any type: nominal, ordinal and / or continuous
- Multinomial logistic regression does not require the use of a coding strategy (i.e. dummy coding) for including categorical predictors in the model. Categorical predictor variables can be included directly as factors in the multinomial logistic regression dialog menu box.

Research Question

- **Variables** in the model
  - DV = Recidivism (3 categories)
    - 1 - first time in prison (reference category)
    - 2 - second time in prison
    - 3 - third time (or more) in prison
  - IVs Categorical
    - Type of criminal offence (0 = non-violent; 1 = violent (ref))
    - Location of offences (1 = urban; 2 = rural; 3 = outside the country (ref))
  - IVs Continuous
    - Criminal Friends Index
    - Criminal Thinking Style
- **The main interest of current analysis is to focus on the relationship between criminal thinking & criminal friends and recidivism while controlling for type and location of offences.**

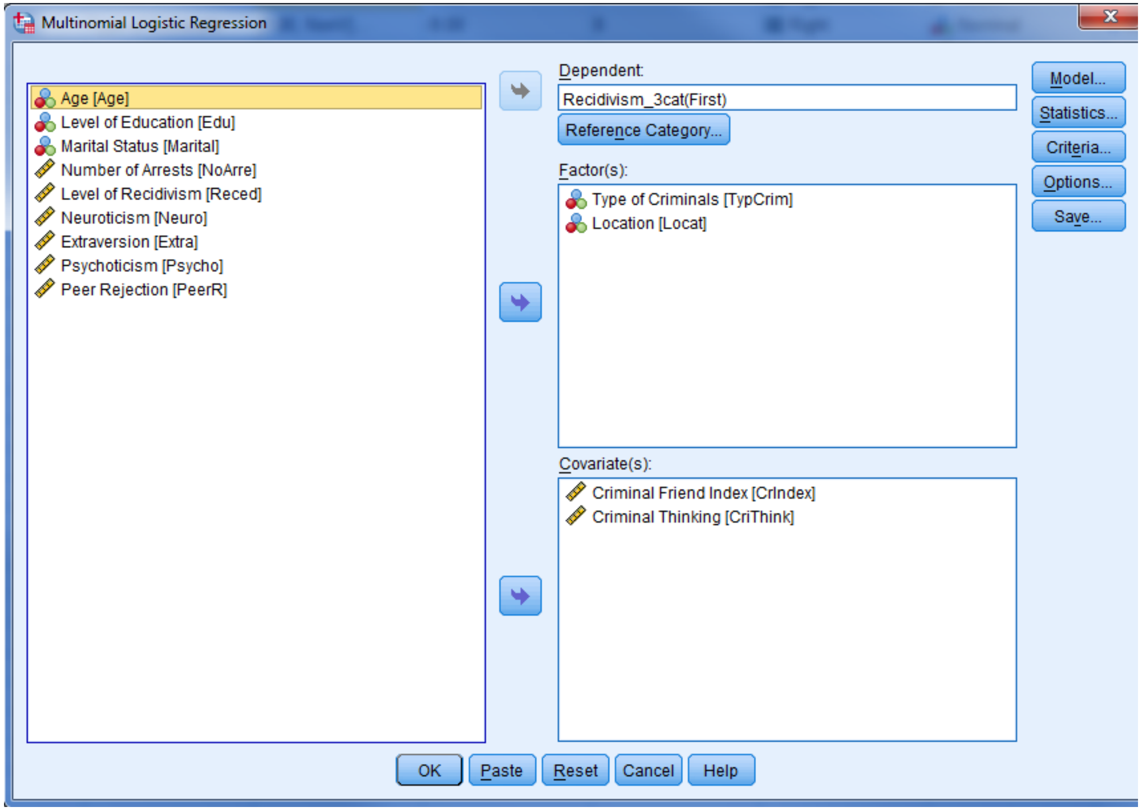
SPSS procedure for MLR

- From the menu at the top of the screen click **Analyze**, then select **Regression**, then **Multinominal Logistic**

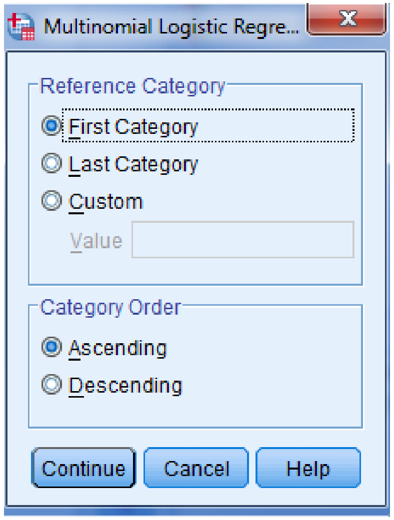


SPSS procedure for MLR

- Choose you categorical DV (Recidivism_3cat) and move it into the **Dependent** box
- Move the categorical IVs (Type of Criminals and Location) into **Factor(s)** box and continuous IVs (Criminal Friend Index and Criminal Thinking) into **Covariate(s)** box.



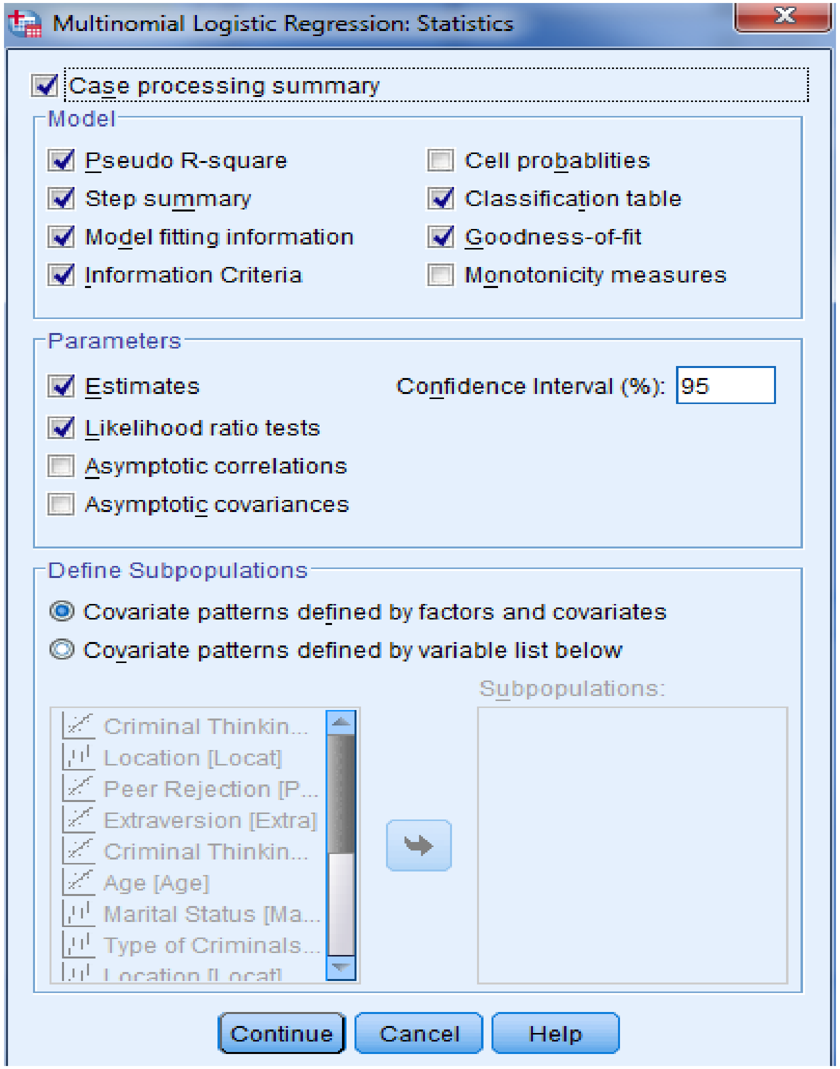
Click on the **Reference Category...** button and select **First Category**



SPSS procedure for MLR

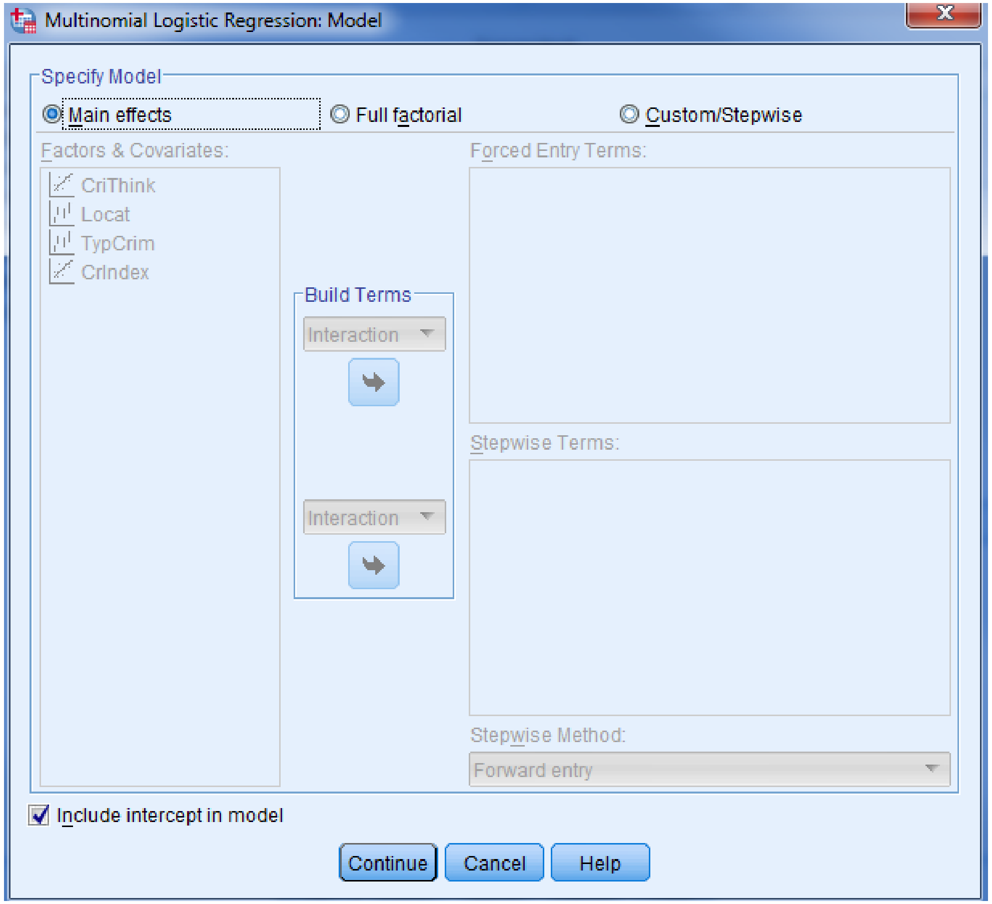
- Click on the **Statistics** button and select the following

- Click **Continue**



SPSS procedure for MLR

- Click on the **Model** button.
- In the **Specify Model** section, click on **Main effects**.
- Click on **Continue**.
- And **OK**



Interpretation of the SPSS output

- The **Case Processing Summary** table simply shows how many cases or observations were in each category of the outcome variable (as well as their percentages) and categorical predictors.
- It also shows if there was any missing data.

Case Processing Summary

		N	Marginal Percentage
Recidivism 3cat	1.00	131	42.0%
	2.00	104	33.3%
	3.00	77	24.7%
Type of Criminals	.00	179	57.4%
	NonV	133	42.6%
Location	Urban	158	50.6%
	Rural	118	37.8%
	3.00	36	11.5%
Valid		312	100.0%
Missing		0	
Total		312	
Subpopulation		287[a]	

a. The dependent variable has only one value observed in 282 (98.3%) subpopulations.

Interpretation of the SPSS output

Model Fitting Information

Model	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	668.421	675.907	664.421			
Final	598.490	643.406	574.490	89.931	10	.000

- The Model Fitting Information table shows various indices for assessing the null model and the final model which includes all the predictors and the intercept (sometimes called the full model)

- Both the AIC and the BIC are information theory based model fit statistics. Lower values of indicate better model fit and both can be below zero (i.e. larger negative values indicate better fit)

- Here, we see model fit is significantχ^2 (10) = 89.93, p < .001, which indicates our full model predicts significantly better, or more accurately, than the null model

- To be clear, you want the *p*-value to be *less than* your established cutoff (generally 0.05) to indicate good fit

Interpretation of the SPSS output

- The Goodness-of-Fit table provides further evidence of good fit for our model. Again, both the Pearson and Deviance statistics are chi-square based methods and subject to inflation with large samples.

- Here, we interpret lack of significance as indicating good fit. To be clear, you want the *p*-value to be *greater than* your established cutoff (generally 0.05) to indicate good fit.

- The Pseudo R-Square table displays three metrics which have been developed to provide a number familiar to those who have used traditional, standard multiple regression. They are treated as measures of effect size, similar to how R^2 is treated in standard multiple regression. However, these metrics do not represent the amount of variance in the outcome variable accounted for by the predictor variables. Higher values indicate better fit.

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	611.745	562	.072
Deviance	567.559	562	.427

Pseudo R-Square

Cox and Snell	.250
Nagelkerke	.283
McFadden	.134

Interpretation of the SPSS output

- The statistics in the Likelihood Ratio Tests table are the same types as those reported for the null and full models above in the Model Fitting Information table. Here however, each element of the model is being compared to the full model in such a way as to allow the research to determine if each element should be included in the full model. In other words, does each element (predictor) contributed meaningfully to the full effect.

Likelihood Ratio Tests

Effect	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC of Reduced Model	BIC of Reduced Model	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	598.490	643.406	574.490[a]	.000	0	.
CrIndex	614.475	651.906	594.475	19.985	2	.000
CriThink	608.373	645.803	588.373	13.883	2	.001
TypCrim	612.176	649.606	592.176	17.685	2	.000
Locat	607.161	637.105	591.161	16.671	4	.002

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Interpretation of the SPSS output

- **Table**

Parameter Estimates

Recidivism 3cat[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp (B)	
								Lower Bound	Upper Bound
2.00	Intercept	-3.407	.739	21.270	1	.000			
	CrIndex	.023	.014	2.697	1	.101	1.023	.996	1.052
	CriThink	.066	.020	10.689	1	.001	1.068	1.027	1.111
	[TypCrim=.00]	-.165	.294	.313	1	.576	.848	.477	1.510
	[TypCrim=1.00]	0[b]	.	.	0	.	.	.	.
	[Locat=1.00]	1.053	.483	4.745	1	.029	2.866	1.111	7.390
	[Locat=2.00]	1.278	.509	6.313	1	.012	3.590	1.325	9.729
	[Locat=3.00]	0[b]	.	.	0	.	.	.	.
3.00	Intercept	-3.969	.873	20.660	1	.000			
	CrIndex	.065	.016	17.344	1	.000	1.067	1.035	1.100
	CriThink	.068	.024	7.770	1	.005	1.070	1.020	1.123
	[TypCrim=.00]	-1.313	.342	14.743	1	.000	.269	.138	.526
	[TypCrim=1.00]	0[b]	.	.	0	.	.	.	.
	[Locat=1.00]	.649	.571	1.292	1	.256	1.914	.625	5.861
	[Locat=2.00]	1.663	.591	7.911	1	.005	5.273	1.655	16.797
	[Locat=3.00]	0[b]	.	.	0	.	.	.	.

a. The reference category is: 1.00.

b. This parameter is set to zero because it is redundant.

Interpretation of the SPSS output

- The Parameter Estimates table, shows the logistic coefficient (B) for each predictor variable for each alternative category of the outcome variable. Alternative category meaning, not the reference category. The logistic coefficient is the expected amount of change in the logit for each one unit change in the predictor. The logit is what is being predicted; it is the odds of membership in the category of the outcome variable which has been specified (here the first value: 1 was specified, rather than the alternative values 2 or 3). The closer a logistic coefficient is to zero, the less influence the predictor has in predicting the logit. The table also displays the standard error, Wald statistic, *df*, Sig. (*p*-value); as well as the Exp(B) and confidence interval for the Exp(B). The Wald test (and associated *p*-value) is used to evaluate whether or not the logistic coefficient is different than zero. The Exp(B) is the odds ratio associated with each predictor. We expect predictors which increase the logit to display Exp(B) greater than 1.0, those predictors which do not have an effect on the logit will display an Exp(B) of 1.0 and predictors which decease the logit will have Exp(B) values less than 1.0. As an example, we can see that a one unit change in x3 does not significantly change the odds of being classified in the first category of the outcome variable relative to the second or third categories of the outcome variable, while controlling for the influence of the other predictors.

Interpretation of the SPSS output

The Classification Table (above) shows how well our full model correctly classifies cases. A perfect model would show only values on the diagonal--correctly classifying all cases. Adding across the rows represents the number of cases in each category in the actual data and adding down the columns represents the number of cases in each category as classified by the full model. The key piece of information is the overall percentage in the lower right corner which shows our model (with all predictors & the constant) is 99.2% accurate; which is excellent.

Classification

Observed	Predicted			
	1.00	2.00	3.00	Percent Correct
1.00	95	24	12	72.5%
2.00	48	35	21	33.7%
3.00	18	24	35	45.5%
Overall Percentage	51.6%	26.6%	21.8%	52.9%

Presenting the results from MLR

- ## Table 1

	2nd Incarceration (n=104)		3rd (or more) Incarceration (n=77)	
Variable	OR (95% CI)	SE	OR (95% CI)	SE
Criminal Friends	1.02 (1.00/1.05)	.01	1.07(1.04/1.10)***	.02
Criminal Thinking	1.07(1.03/1.11)***	.02	1.07(1.02/1.12)**	.02
Type of offence				
Non-violent	.85 (.48/1.51)	.29	.27(.14/.53)***	.34
Violent	1		1	
Location of offence				
Urban	2.87 (1.11/7.39)***	.48	1.91(.63/5.86)	.57
Rural	3.59 (1.32/9.73)***	.51	5.27(1.66/16.80)**	.59
Outside the country	1		1	

Note. Reference group: 1st Incarceration (n=131). OR = Odds Ratio. SE = Standard Error. 95% CI = Confidence Interval. * p<.05; ** p<.01; *** p<.001

Presenting the results from MLR

- A Multinomial Logistic Regression was used to analyse predictors for an unordered group classification, such as prisoners who were incarcerated for the first time, prisoners who were incarcerated for the second time, and prisoners who were incarcerated for the third time (or more). The reference category for the outcome variable was 'first incarcerated prisoners'; each of the other two categories was compared to this reference group. The main interest of current analysis was focused on the relationship between criminal thinking & criminal friends and recidivism (3 categories) while controlling for type and location of offences.
- The first column in Table 1 has the outcome of "second incarceration" compared to "first incarceration" (reference category). The results suggest that criminal friends have no significant effect on the recidivism. However, higher levels of criminal thinking style (OR = 1.07) significantly increase the probability of recidivism. In relation to the location of offences participants from urban (OR = 2.87) and rural areas (OR = 3.59), compared to offences committed outside of the country, are more likely to be incarcerated more than once. Type of offence was not a significant predictor of recidivism.
- The second column in Table 1 has the outcome of "third (or more) incarceration" compared to "first incarceration" (reference category). Statistical analysis shows that those participants who reported higher level of criminal thinking (OR = 1.07) and associations with criminal friends (OR = 1.07) were significantly more likely to report recidivism. Non-violent offences (OR = .27), compared to violent offences decrease the probability of the recidivism. In relation to the location of offences participants from rural areas (OR = 5.27), compared to offences committed outside of the country, are over five times more likely to be incarcerated on 3 or more occasions.