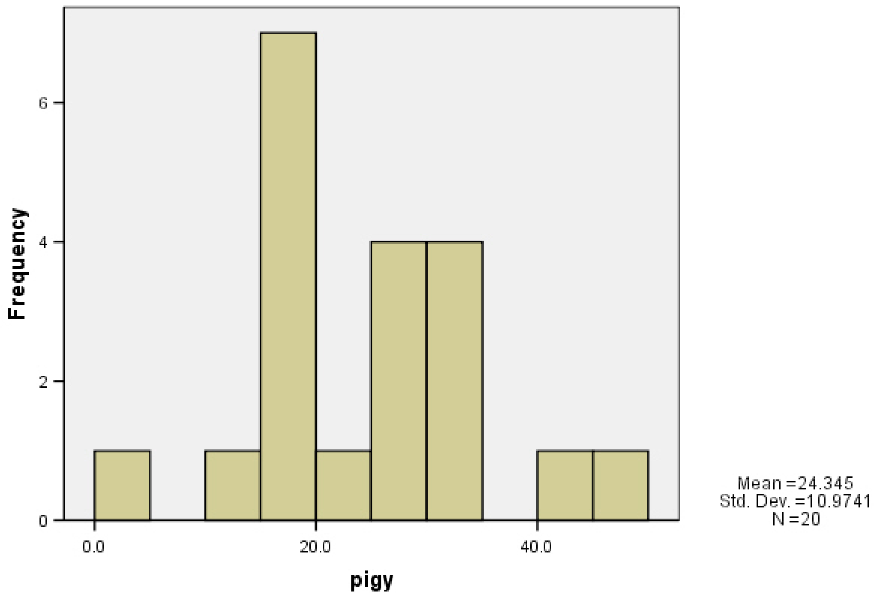
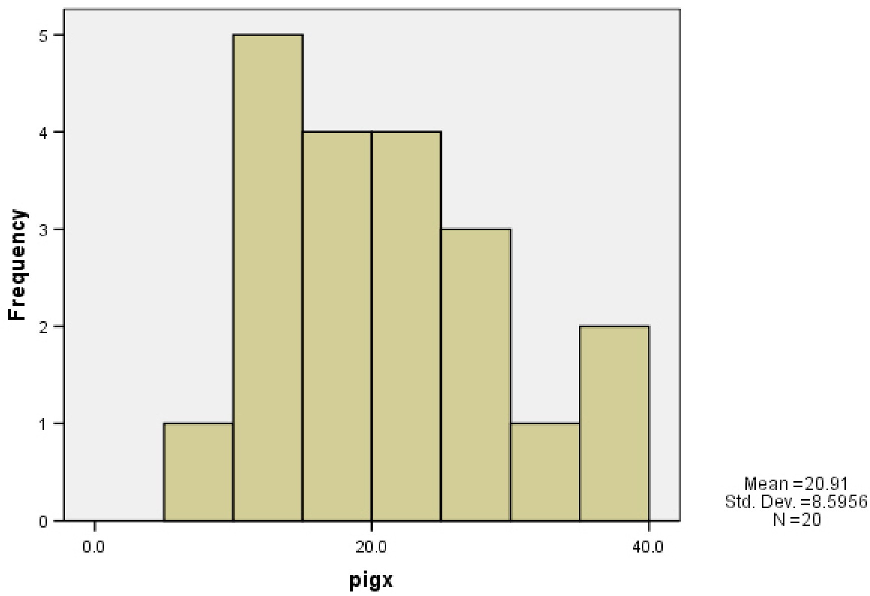
Chapter 3: Datasets

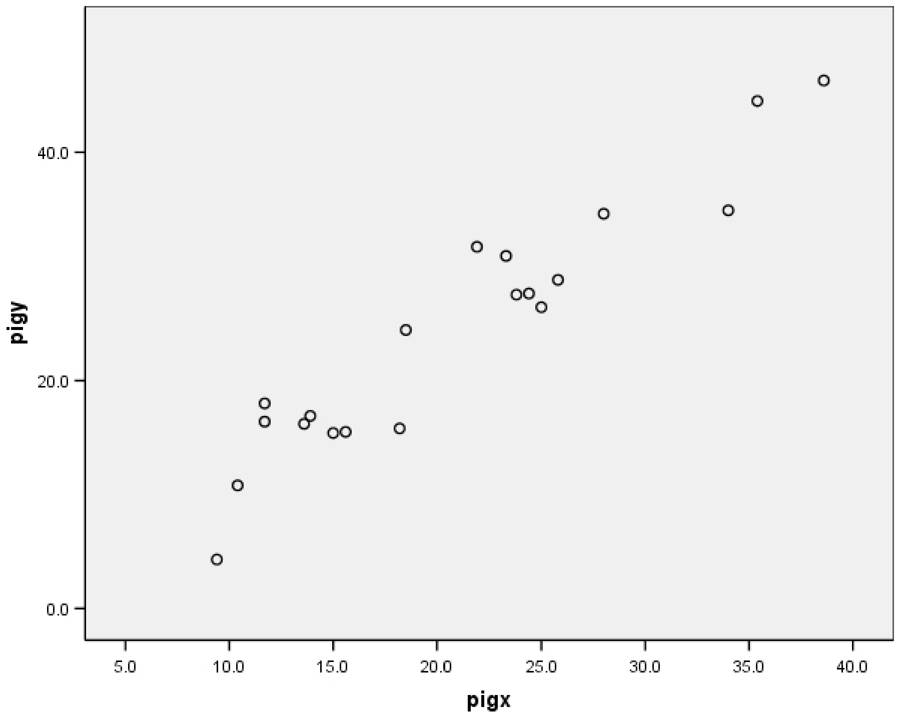
Dataset used: **pigx** & **pigy** from www.luchsinger-mathematics.ch/outputpig.txt

Statistics

		pigx	pigy
N	Valid	20	20
	Missing	0	0
Mean		20.910	24.345
Median		20.200	25.400
Variance		73.885	120.431
Minimum		9.4	4.3
Maximum		38.6	46.3







Correlations

		pigx	pigy
pigx	Pearson Correlation	1	.951**
	Sig. (2-tailed)		.000
	N	20	20
pigy	Pearson Correlation	.951**	1
	Sig. (2-tailed)	.000	
	N	20	20

**. Correlation is significant at the 0.01 level

Chapter 7: Datasets

Dataset used: **znz1**, **znz2** & **znz3** from www.luchsinger-mathematics.ch/Data.txt

Descriptives

			Statistic	Std. Error
znz1	Mean		-2.70	.784
	95% Confidence Interval for Mean	Lower Bound	-4.28	
		Upper Bound	-1.13	
	5% Trimmed Mean		-2.68	
	Median		-3.60	
	Variance		31.312	
	Std. Deviation		5.596	
	Minimum		-16	
	Maximum		9	
	Range		24	
	Interquartile Range		9	
	Skewness		-.037	.333
	Kurtosis		-.573	.656

Statistics

znz1

N	Valid	51
	Missing	0

Descriptives

			Statistic	Std. Error
znz2	Mean		-9.694	1.1904
	95% Confidence Interval for Mean	Lower Bound	-12.125	
		Upper Bound	-7.262	
	5% Trimmed Mean		-9.600	
	Median		-10.500	
	Variance		43.929	
	Std. Deviation		6.6279	
	Minimum		-23.1	
	Maximum		.9	
	Range		24.0	
	Interquartile Range		10.1	
	Skewness		-.095	.421
	Kurtosis		-.832	.821

Statistics

znz2

N	Valid	31
	Missing	0

Descriptives

			Statistic	Std. Error
znz3	Mean		-11.402	1.1497
	95% Confidence Interval for Mean	Lower Bound	-13.726	
		Upper Bound	-9.079	
	5% Trimmed Mean		-11.267	
	Median		-10.600	
	Variance		54.198	
	Std. Deviation		7.3619	
	Minimum		-30.7	
	Maximum		2.7	
	Range		33.4	
	Interquartile Range		10.1	
	Skewness		-.302	.369
	Kurtosis		-.103	.724

Statistics

znz3

N	Valid	41
	Missing	0

Chapter 7: Tests

7.5.3: $x_1, \ldots, x_n$ **from** $\mathcal{N}(\mu, \sigma^2)$ **with** σ^2 **unknown: is** μ **equal to some** μ_0 **(for example equal to 0)? [1 SAMPLE T-TEST]**

SPSS: Analyze > Compare Means > One-Sample T Test

SPSS gives us the following output with data from Data.txt: t=-3.451, df=50, Sig. (2-tailed): 0.001 (should be 0.00114 if more precise), Mean Difference=-2.704, 95 % CI: [-4.278, -1.130].

Dataset used: **znz1**

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
pressure	51	-2.70	5.596	.784

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
pressure	-3.451	50	.001	-2.704	-4.28	-1.13

Chapter 7: Tests

7.5.4: $x_1, \ldots, x_m$ **from** $\mathcal{N}(\mu_1, \sigma^2)$, $y_1, \ldots, y_n$ **from a** $\mathcal{N}(\mu_2, \sigma^2)$, **independent and variance is equal. Is** $\mu_1 = \mu_2$**? [2 SAMPLE T-TEST]**

SPSS: Analyze > Independent Samples > T Test

SPSS gives us the following output with data from Data.txt; see SPSS-Applications

---

Dataset used: **znz2** & **znz3**

Group Statistics

	stdudy	N	Mean	Std. Deviation	Std. Error Mean
pressure	drug	31	-9.694	6.6279	1.1904
	placebo	41	-11.402	7.3619	1.1497

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
pressure	Equal variances assumed	.134	.715	1.017	70	.312	1.7089	1.6796	-1.6409	5.0587
	Equal variances not assumed			1.033	67.816	.305	1.7089	1.6550	-1.5937	5.0115

7.6 An example of a non-parametrical test: χ^2-Test for independence in contingency tables

SPSS: Analyze > Descriptive Statistics > Crosstables

Dataset used: n = 100 people tested in Maths; 46 male, 54 female; 12 male fail, 22 female fail.

sex * math Crosstabulation

Count

		math		Total
		failed	succeed	
sex	male	12	34	46
	female	22	32	54
Total		34	66	100

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.377[b]	1	.123		
Continuity Correction[a]	1.769	1	.184		
Likelihood Ratio	2.405	1	.121		
Fisher's Exact Test				.142	.091
Linear-by-Linear Association	2.353	1	.125		
N of Valid Cases	100				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 15. 64.

7.6 An example of a non-parametrical test: χ^2-Test for independence in contingency tables

SPSS: Analyze > Descriptive Statistics > Crosstables

Dataset used: 326 Criminals; 166 African American, 160 White; 17 of 166 African American where sentenced to death, 19 of 160 White where sentenced to death.

race * sentence Crosstabulation

Count

		sentence		
		sentenced to death	not sentenced to death	Total
race	african american	17	149	166
	white	19	141	160
Total		36	290	326

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.221[b]	1	.638		
Continuity Correction[a]	.086	1	.769		
Likelihood Ratio	.221	1	.638		
Fisher's Exact Test				.725	.384
Linear-by-Linear Association	.221	1	.638		
N of Valid Cases	326				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 17.67.

Chapter 7: Tests

7.7.2 1-Way-ANOVA

SPSS: Analyze > Compare Means > One-Way-ANOVA

SPSS gives us the following output with data from Data.txt:

Dataset used: **znz5**

ONEWAY deskriptive Statistiken

growth

	N	Mittelwert	Standardabweichung	Standardfehler	95%-Konfidenzintervall für den Mittelwert		Minimum	Maximum
					Untergrenze	Obergrenze		
1	6	40.800	5.4904	2.2414	35.038	46.562	33.3	47.8
2	4	39.325	5.7384	2.8692	30.194	48.456	35.4	47.6
3	5	35.480	5.2604	2.3525	28.948	42.012	29.6	42.8
4	6	41.450	3.1085	1.2691	38.188	44.712	38.5	45.5
Gesamt	21	39.438	5.1103	1.1152	37.112	41.764	29.6	47.8

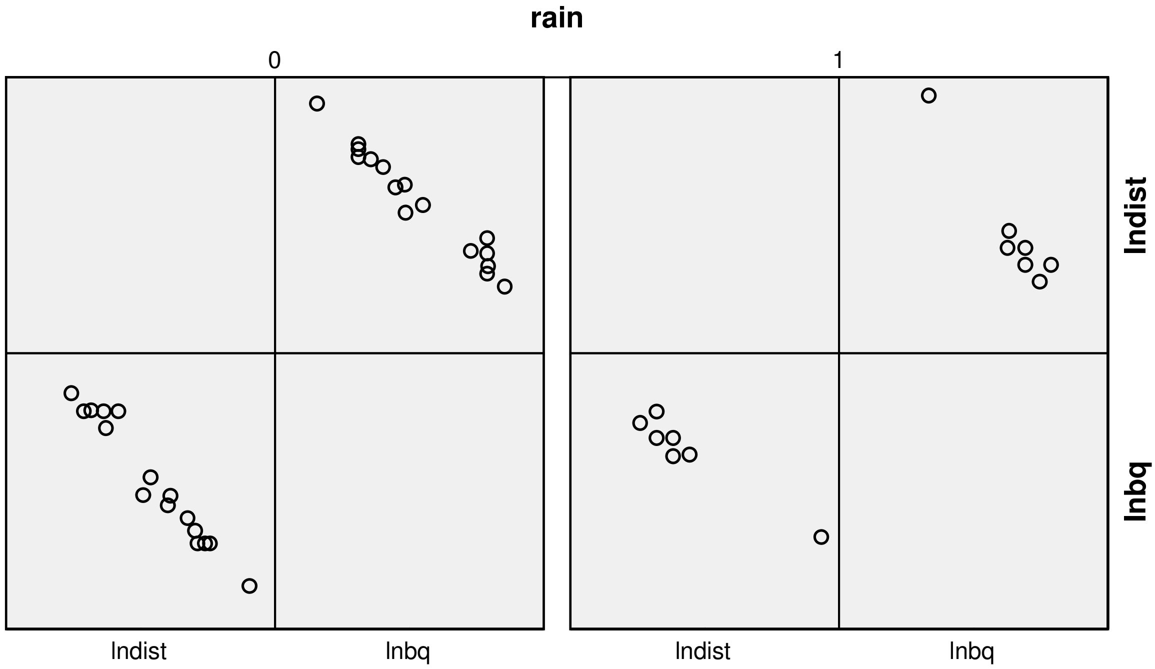
ONEWAY ANOVA

growth

	Quadratsumme	df	Mittel der Quadrate	F	Signifikanz
Zwischen den Gruppen	113.799	3	37.933	1.579	.231
Innerhalb der Gruppen	408.510	17	24.030		
Gesamt	522.310	20			

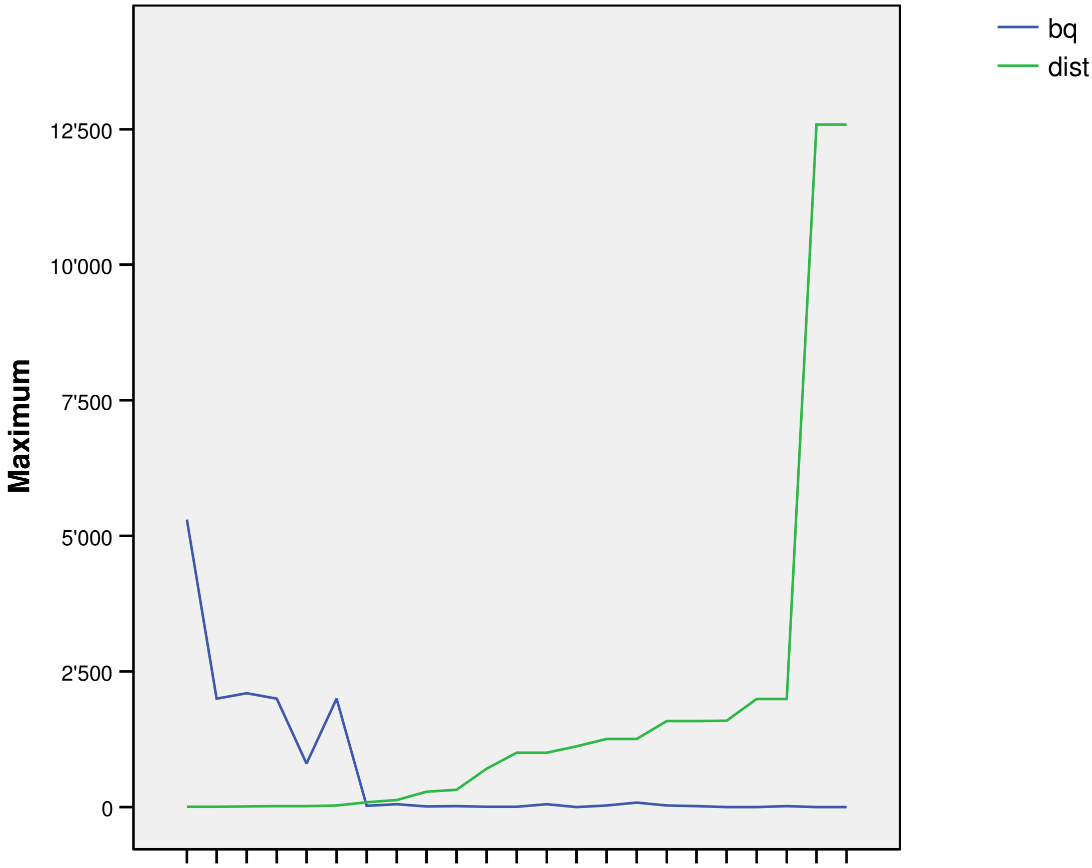
Chapter 8: Dataset

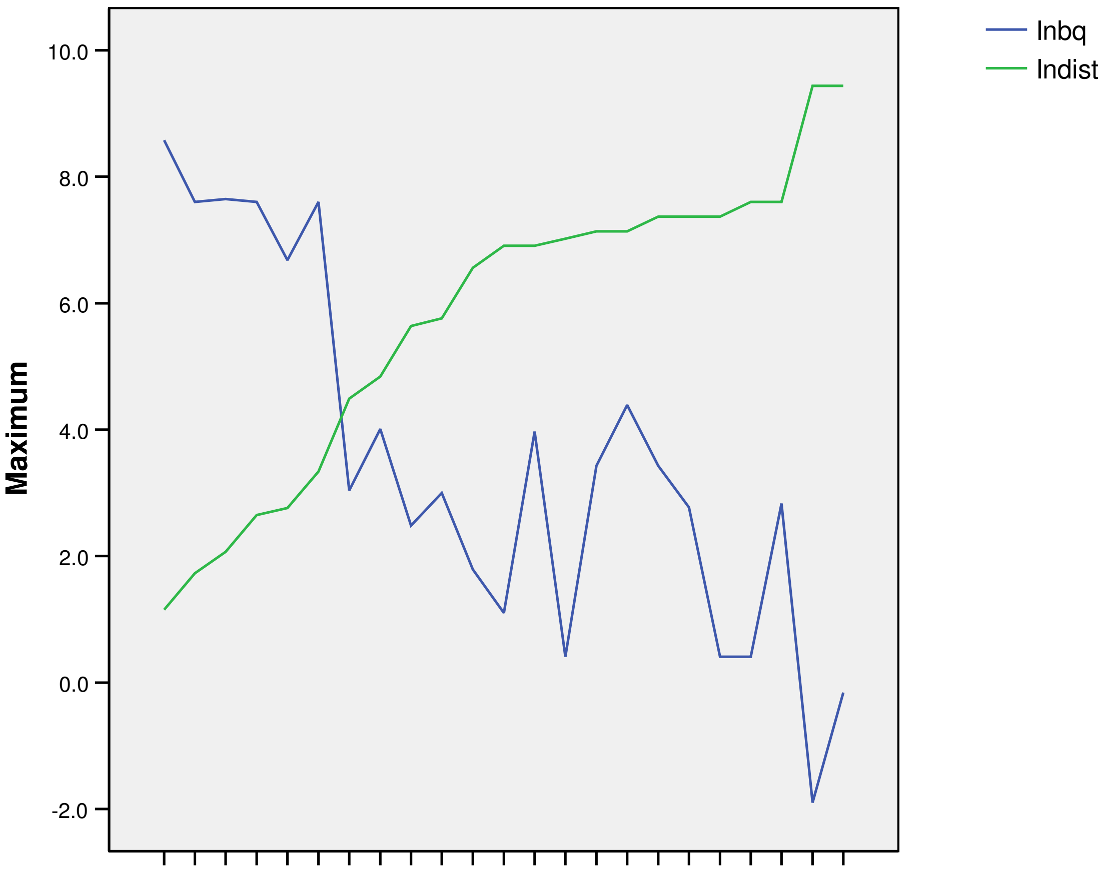
Dataset used: **znz4** from www.luchsinger-mathematics.ch/Data.txt



Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
lnbq	23	-1.9	8.6	3.527	2.9171
Valid N (listwise)	23				


bq
dist
Maximum
12'500
10'000
7'500
5'000
2'500
0


lnbq
lndist
Maximum
10.0
8.0
6.0
4.0
2.0
0.0
-2.0

Chapter 8: Analysis

8.1.1 Try to explain (or predict) radioactivity with distance alone

SPSS: Analyze > Regression > Linear

SPSS gives us the following output with data from Data.txt:

Dataset used: **znz4**

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.904[a]	.816	.808	1.2794

a. Predictors: (Constant), lndist

b. Dependent Variable: lnbq

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	152.832	1	152.832	93.373	.000[a]
	Residual	34.373	21	1.637		
	Total	187.205	22			

a. Predictors: (Constant), lndist

b. Dependent Variable: lnbq

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	9.802	.702		13.962	.000
	lndist	-1.091	.113	-.904	-9.663	.000

a. Dependent Variable: lnbq

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	-.496	8.547	3.527	2.6357	23
Residual	-1.8639	2.3770	.0000	1.2500	23
Std. Predicted Value	-1.526	1.905	.000	1.000	23
Std. Residual	-1.457	1.858	.000	.977	23

a. Dependent Variable: lnbq

Chapter 8: Analysis

8.1.2 Try to explain (or predict) radioactivity with rain alone

SPSS: Analyze > Regression > Linear

SPSS gives us the following output with data from Data.txt:

Dataset used: **znz4**

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.133[a]	.018	-.029	2.9590

a. Predictors: (Constant), rain

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3.333	1	3.333	.381	.544[a]
	Residual	183.872	21	8.756		
	Total	187.205	22			

a. Predictors: (Constant), rain

b. Dependent Variable: lnbq

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.779	.740		5.108	.000
	rain	-.827	1.341	-.133	-.617	.544

a. Dependent Variable: lnbq

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	2.951	3.779	3.527	.3892	23
Residual	-5.6788	4.8012	.0000	2.8910	23
Std. Predicted Value	-1.479	.647	.000	1.000	23
Std. Residual	-1.919	1.623	.000	.977	23

a. Dependent Variable: lnbq

8.1.3 Try to explain (or predict) radioactivity with distance *and* rain

SPSS: Analyze > Regression > Linear

SPSS gives us the following output with data from Data.txt:

Dataset used: **znz4**

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.979[a]	.959	.955	.6173

a. Predictors: (Constant), rain, lndist

b. Dependent Variable: lnbq

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	179.585	2	89.792	235.659	.000[a]
	Residual	7.621	20	.381		
	Total	187.205	22			

a. Predictors: (Constant), rain, lndist

b. Dependent Variable: lnbq

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	10.522	.349		30.110	.000
	lndist	-1.360	.063	-1.127	-21.507	.000
	rain	2.721	.325	.439	8.379	.000

a. Dependent Variable: lnbq

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	-2.317	8.958	3.527	2.8571	23
Residual	-1.3755	1.6204	.0000	.5885	23
Std. Predicted Value	-2.045	1.901	.000	1.000	23
Std. Residual	-2.228	2.625	.000	.953	23

a. Dependent Variable: lnbq

Chapter 8: Analysis

Variance reduction

SPSS: Graphs > Boxplot

SPSS gives us the following output with data from Data.txt:

Dataset used: **znz4**