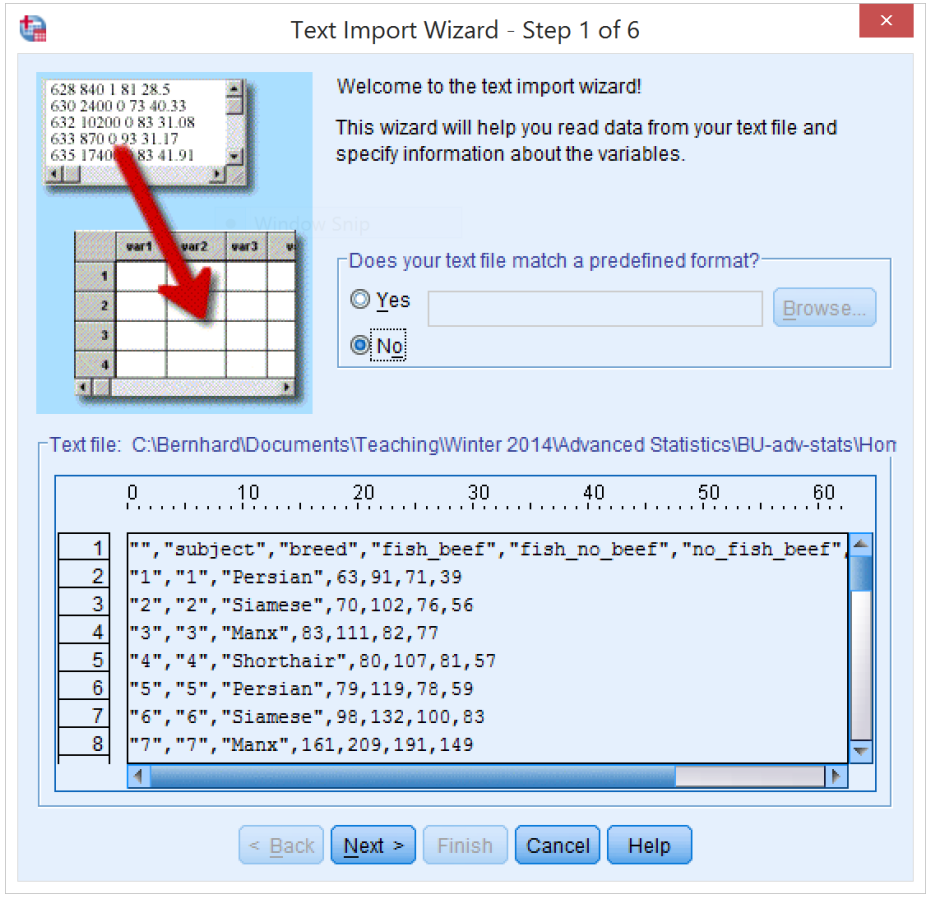
Homework 4
How to do it in SPSS

Open the data file

- Get the data file (from Github or myBU) and put it in a directory where you can find it later).
- Open the data file (Homework4_data_spss.csv) by selecting “Read Text Data…” from the File menu.
- This will open the Text Import Wizard
- (For the assignment, an SPSS .sav file is provided, so you can skip this step)

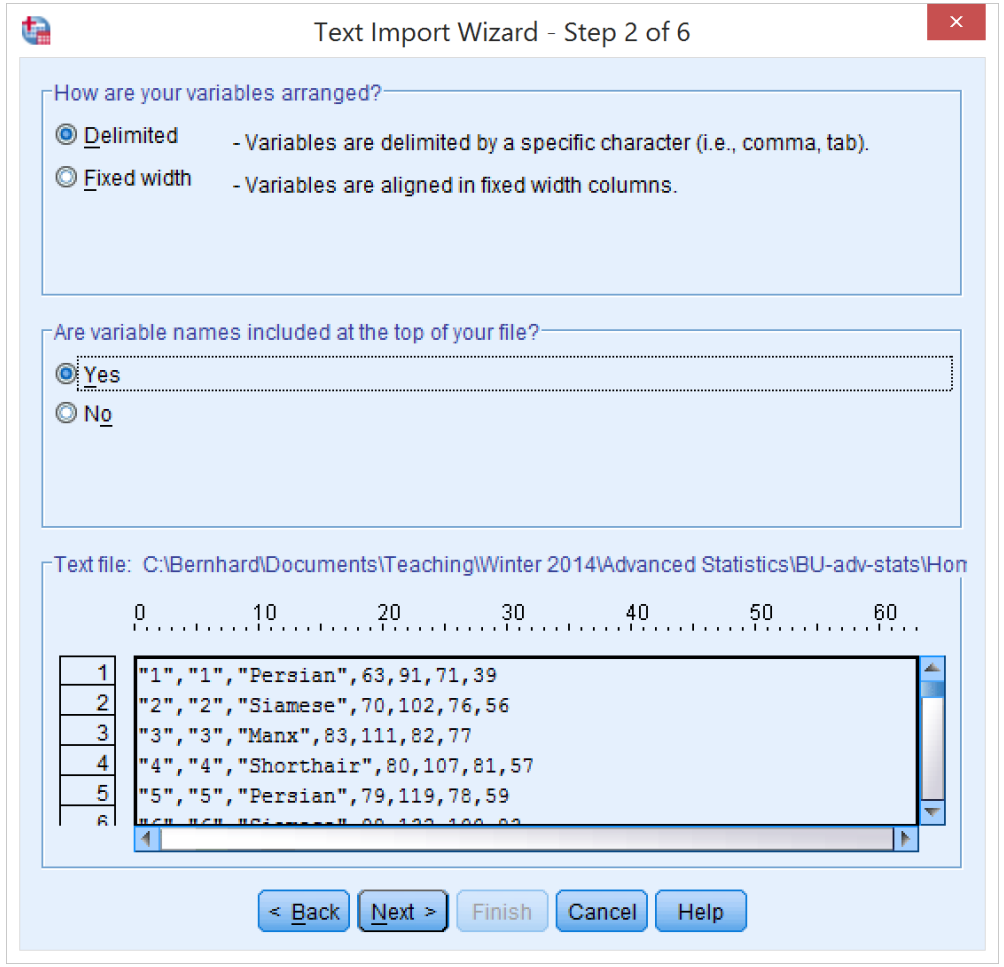
Text import wizard

- Press Next



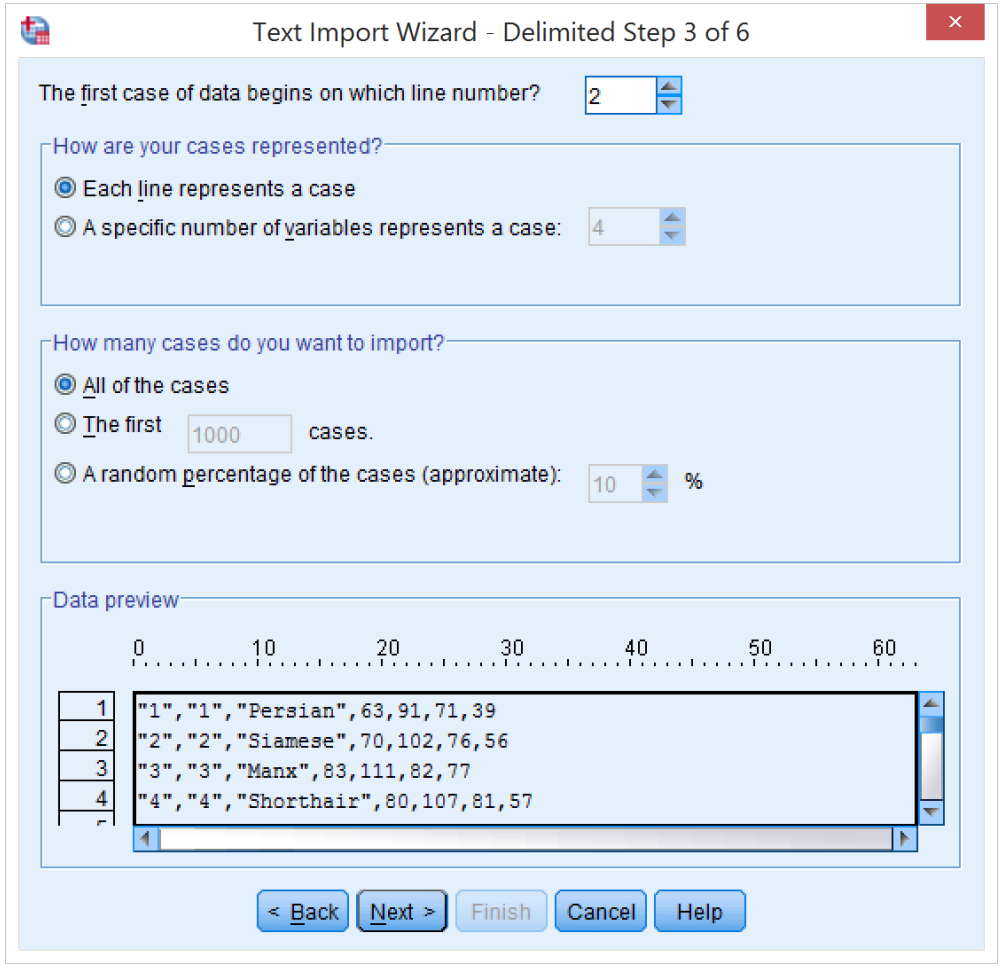
Text import wizard

- Select “Delimited” (since the data are comma delimited)
- Select “Yes”, since the first line contains column names
- Press Next



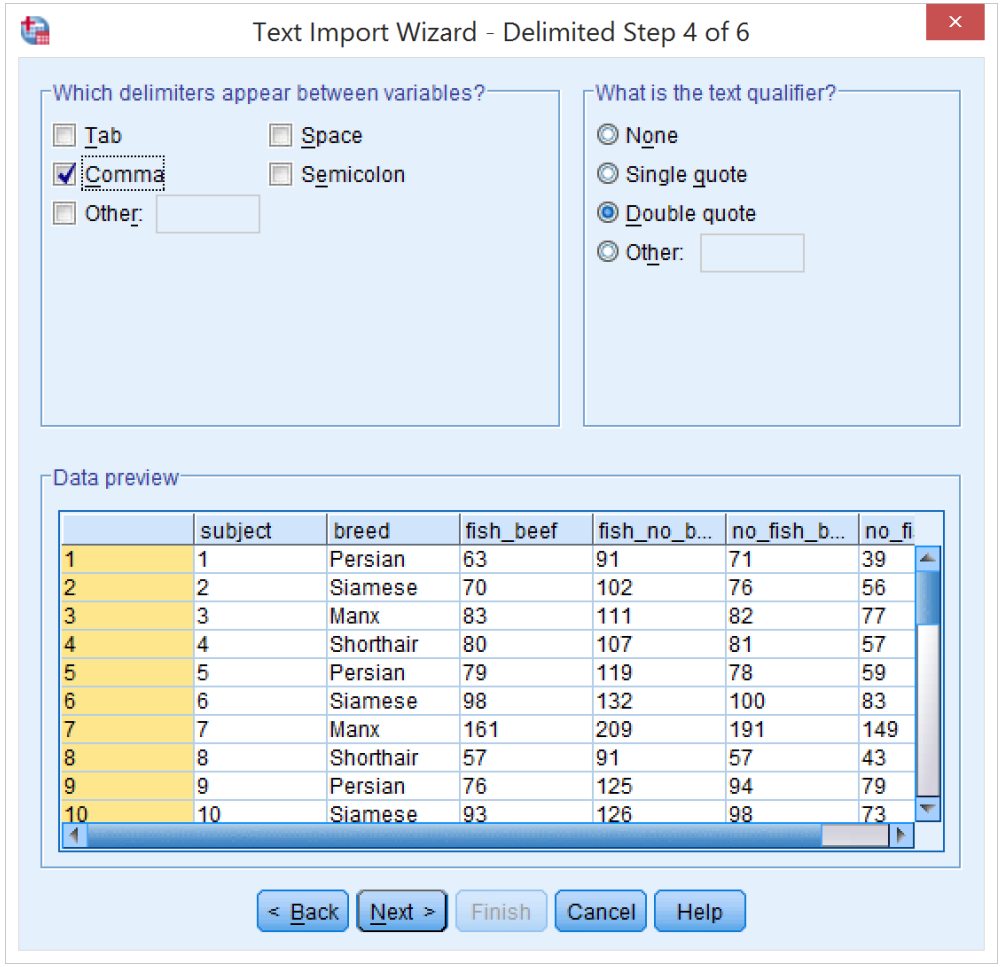
Text import wizard

- These values should be correct, but double-check:
  - First case of data begins on line 2
  - Each line represents a case
  - Import all of the cases
- Press Next



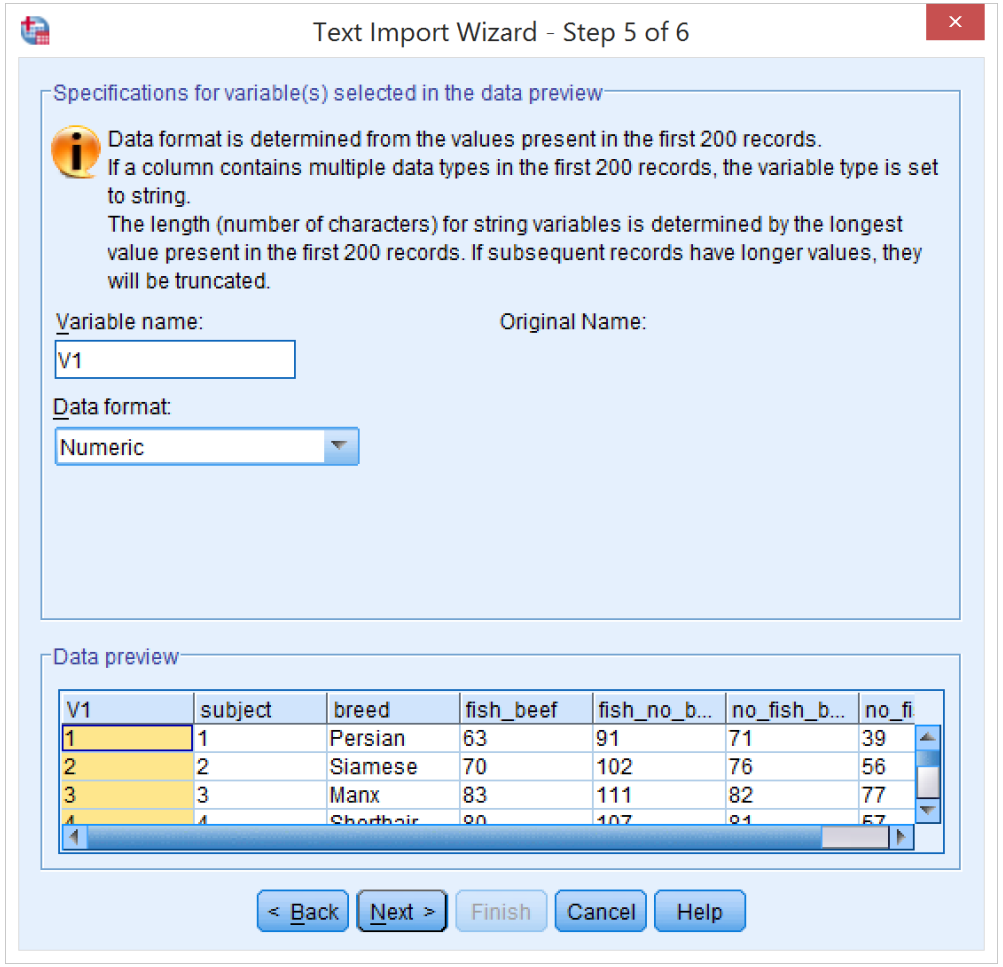
Text import wizard

- These values should be correct, but double-check:
  - Delimiters that appear between variables are commas (and nothing else)
  - Text qualifier is Double quote
- Press Next



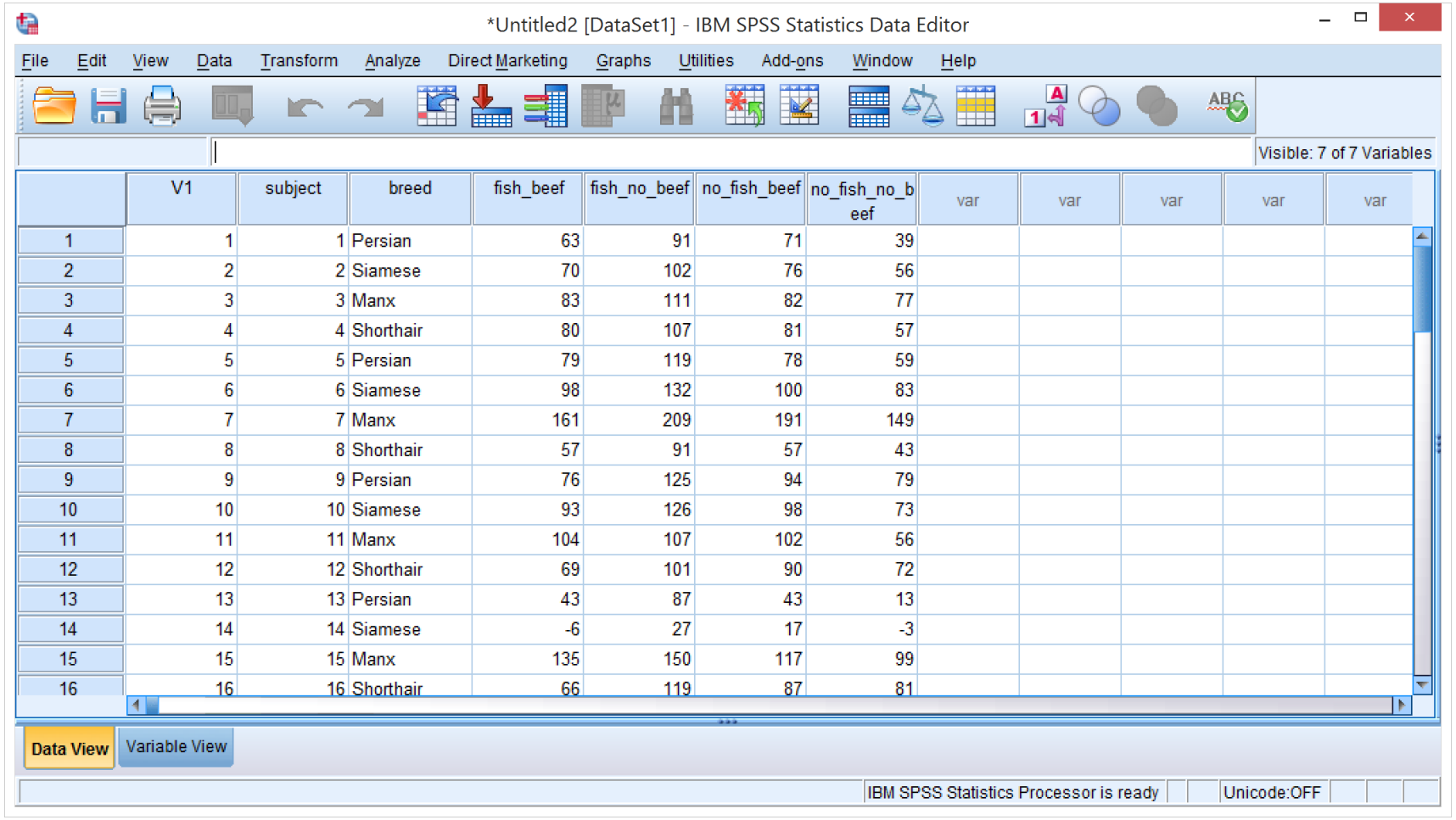
Text import wizard

- SPSS may complain about invalid variable names, but it fixes them just fine.
- Press Next, then Finish.



Data View

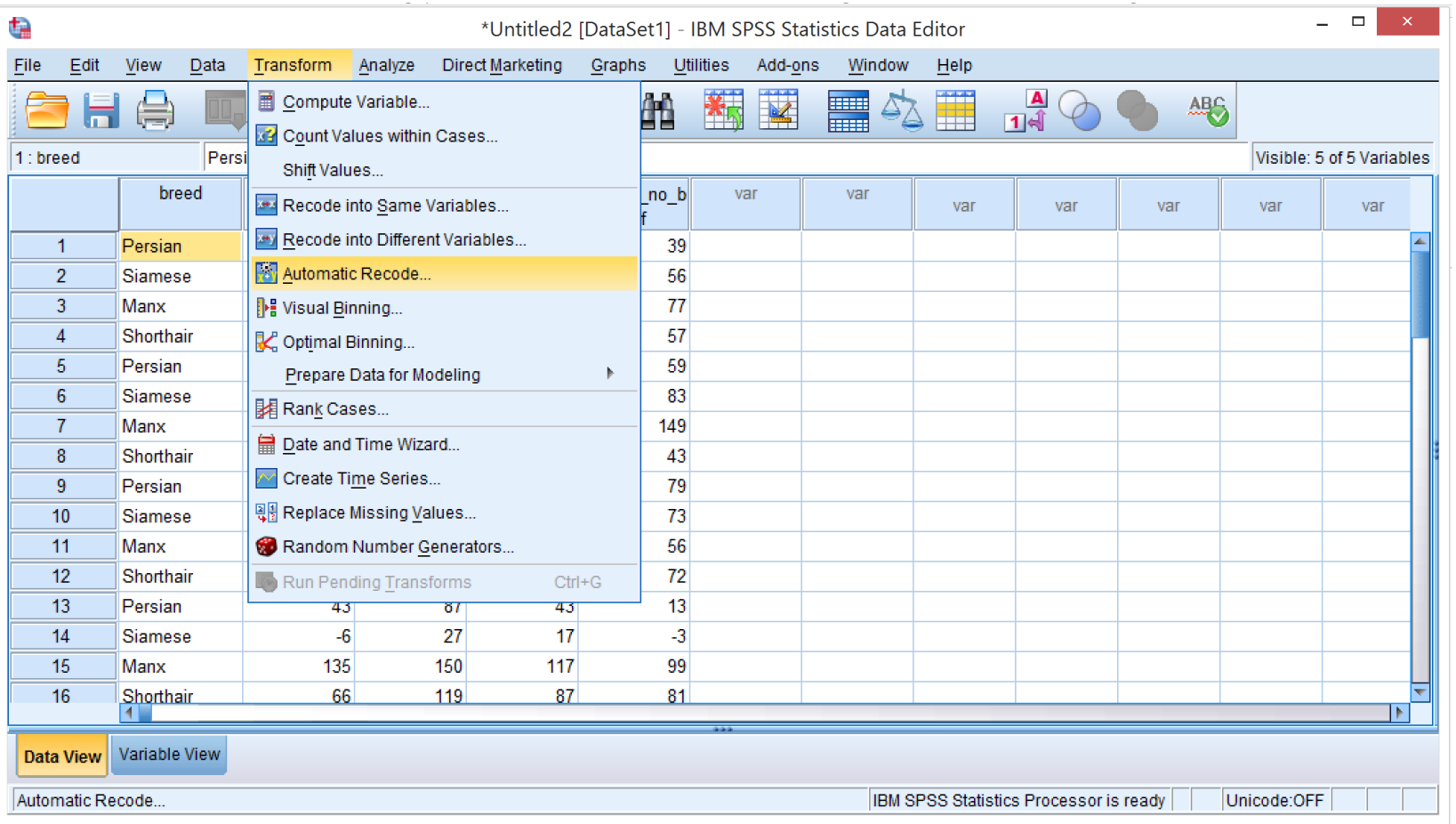
- Your data should now be imported like this.
- You can clear the “V1” and “subject” variables if you want (but they don’t hurt either)



	V1	subject	breed	fish_beef	fish_no_beef	no_fish_beef	no_fish_no_beef	var	var	var	var	var
1	1	1	Persian	63	91	71	39					
2	2	2	Siamese	70	102	76	56					
3	3	3	Manx	83	111	82	77					
4	4	4	Shorthair	80	107	81	57					
5	5	5	Persian	79	119	78	59					
6	6	6	Siamese	98	132	100	83					
7	7	7	Manx	161	209	191	149					
8	8	8	Shorthair	57	91	57	43					
9	9	9	Persian	76	125	94	79					
10	10	10	Siamese	93	126	98	73					
11	11	11	Manx	104	107	102	56					
12	12	12	Shorthair	69	101	90	72					
13	13	13	Persian	43	87	43	13					
14	14	14	Siamese	-6	27	17	-3					
15	15	15	Manx	135	150	117	99					
16	16	16	Shorthair	66	119	87	81					

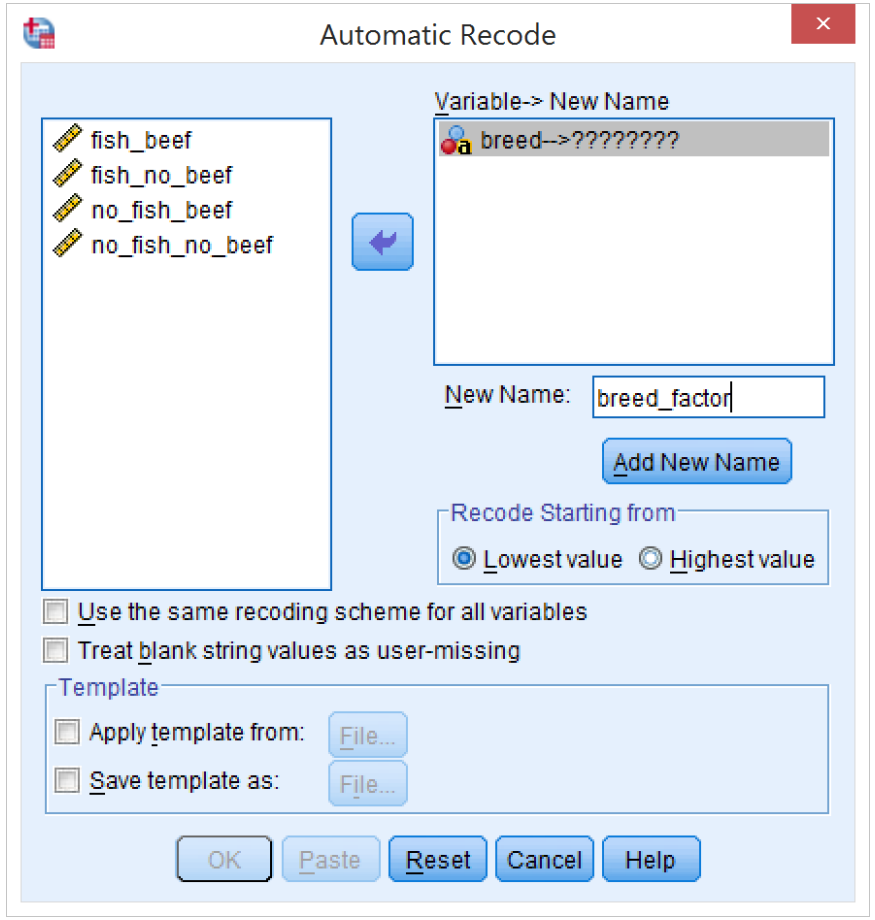
Recode breed into a discrete factor

- From the Transform menu select “Automatic Recode…”
- (If you are working from the SPSS .sav file for the assignment, you can skip this step).

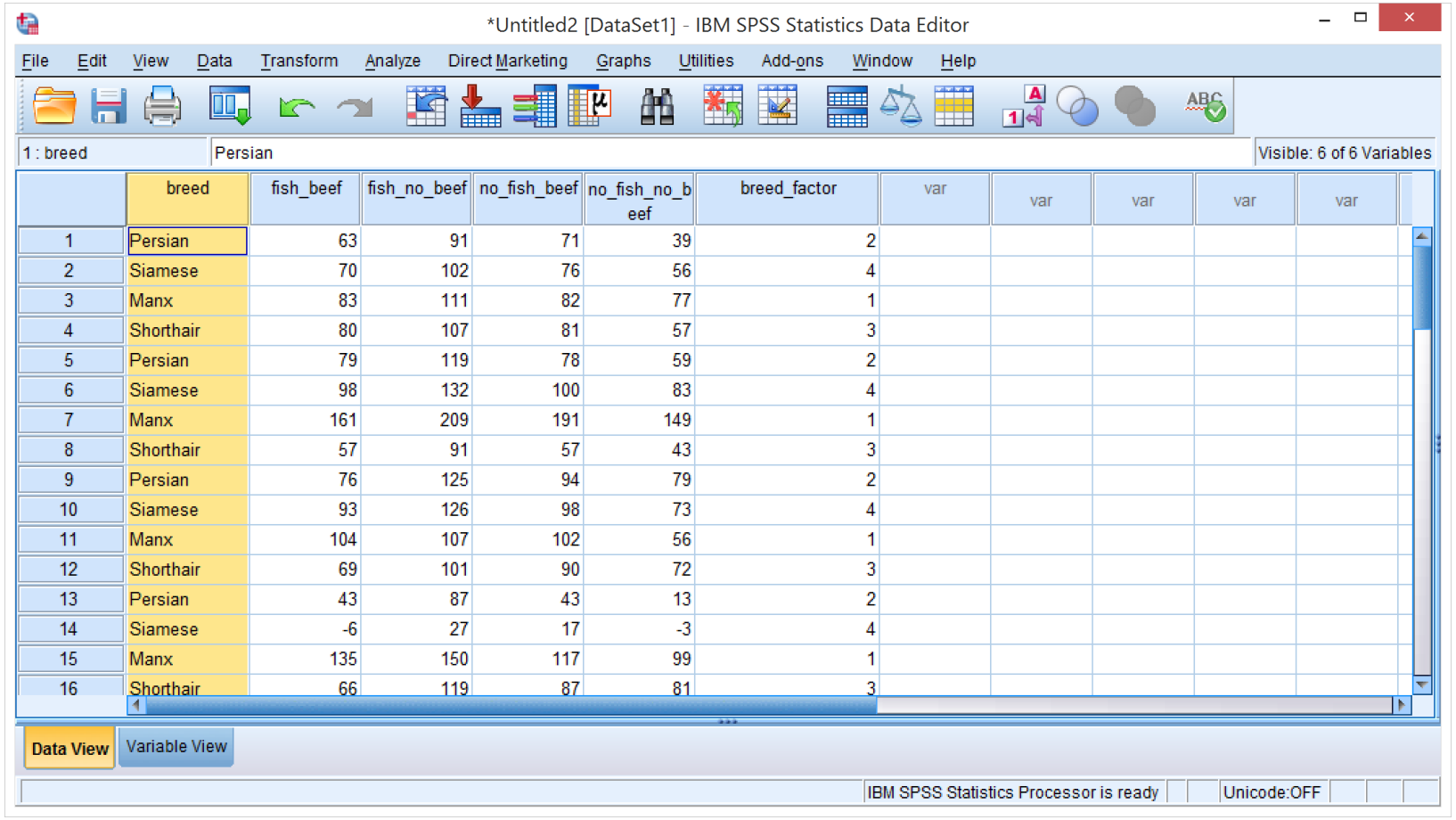


Recode breed into a discrete factor

- Select “breed” from the list on the left and click the arrow to move it to the right
- Enter a new name for the recoded variable
- Recode starting from “Lowest value” is fine for our purposes
- This will assign a different number to each breed, starting alphabetically (i.e. with Manx).
- Click “Add New Name” and then OK



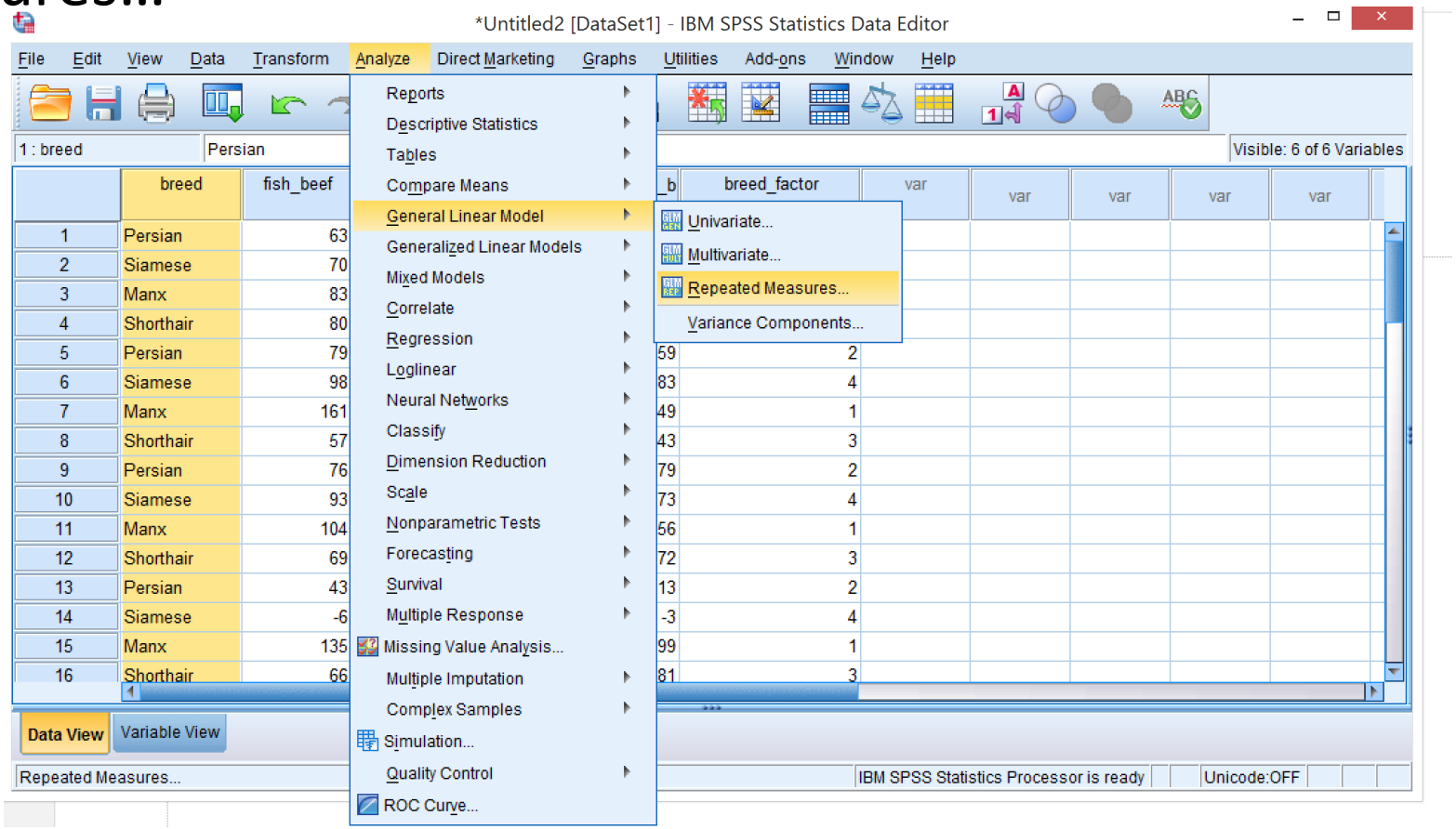
You should now have a new variable "breed_factor"



	breed	fish_beef	fish_no_beef	no_fish_beef	no_fish_no_beef	breed_factor	var	var	var	var	var
1	Persian	63	91	71	39	2					
2	Siamese	70	102	76	56	4					
3	Manx	83	111	82	77	1					
4	Shorthair	80	107	81	57	3					
5	Persian	79	119	78	59	2					
6	Siamese	98	132	100	83	4					
7	Manx	161	209	191	149	1					
8	Shorthair	57	91	57	43	3					
9	Persian	76	125	94	79	2					
10	Siamese	93	126	98	73	4					
11	Manx	104	107	102	56	1					
12	Shorthair	69	101	90	72	3					
13	Persian	43	87	43	13	2					
14	Siamese	-6	27	17	-3	4					
15	Manx	135	150	117	99	1					
16	Shorthair	66	119	87	81	3					

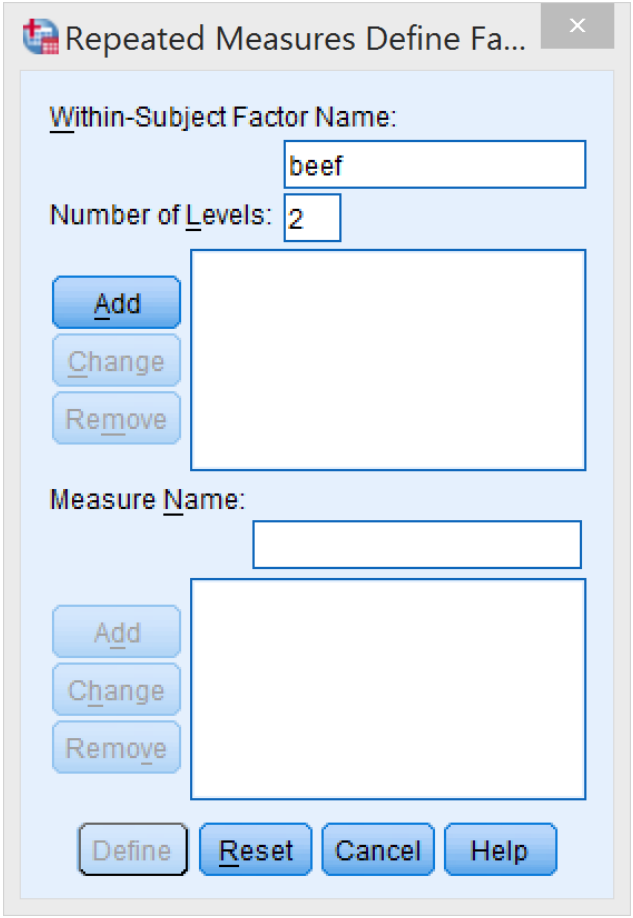
Run the analysis

- From the Analyze menu, select General Linear Model and Repeated Measures…



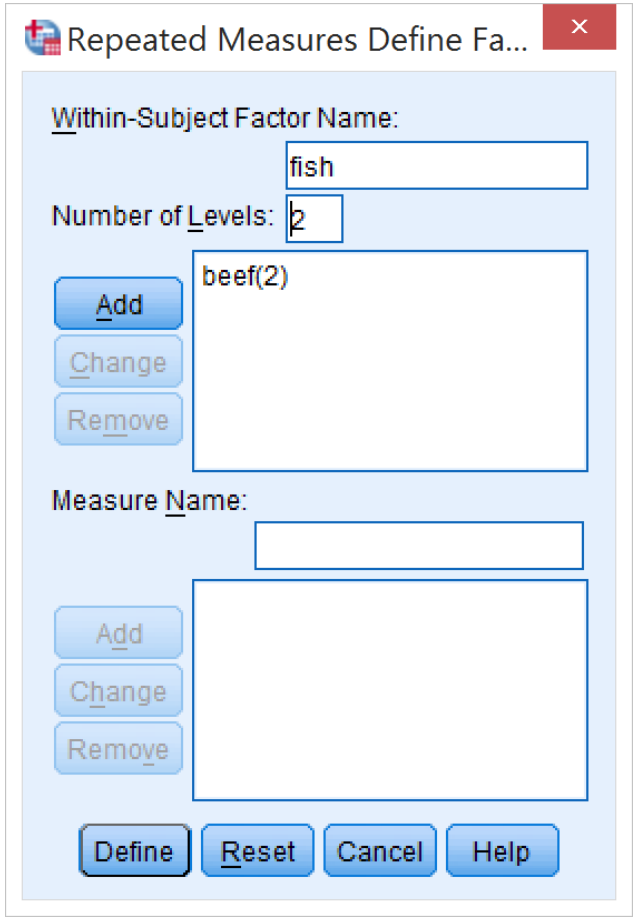
Run the analysis

- First, you have to re-code the data (i.e. split the columns back up into the two within-subject factors)
- Add the name for the first factor (“beef”) and the number of levels (2)
- Click “Add”



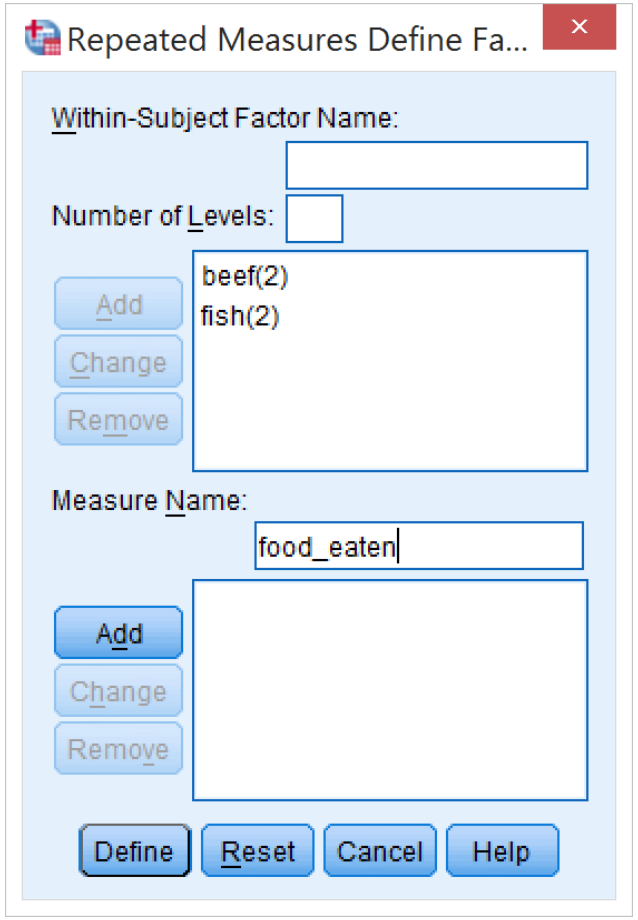
Run the analysis

- Add the name for the second factor (“fish”) and the number of levels (2)
- Click “Add”



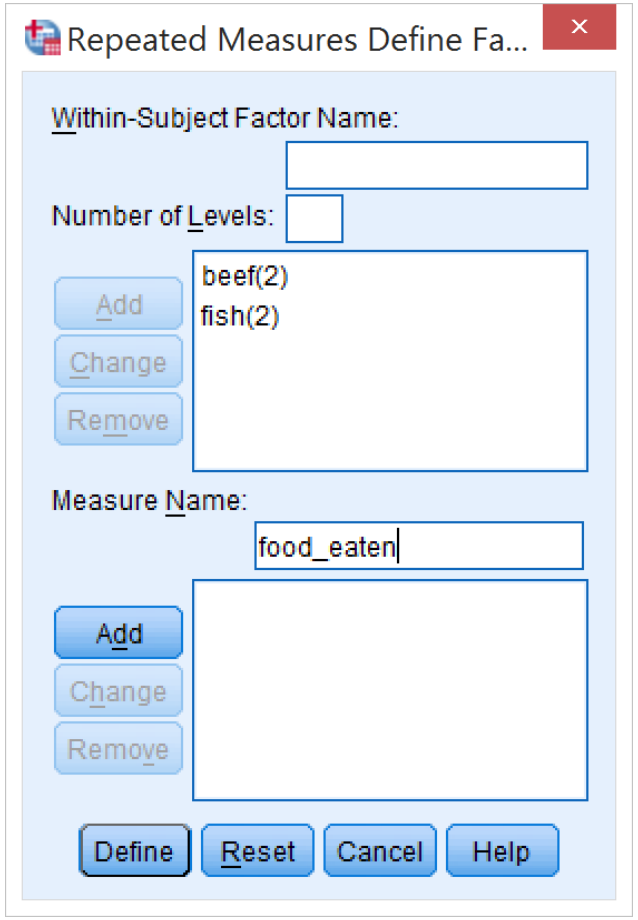
Run the analysis

- Add a name for the dependent measure (e.g. “food_eaten”)
- Click “Add” and “Define”.



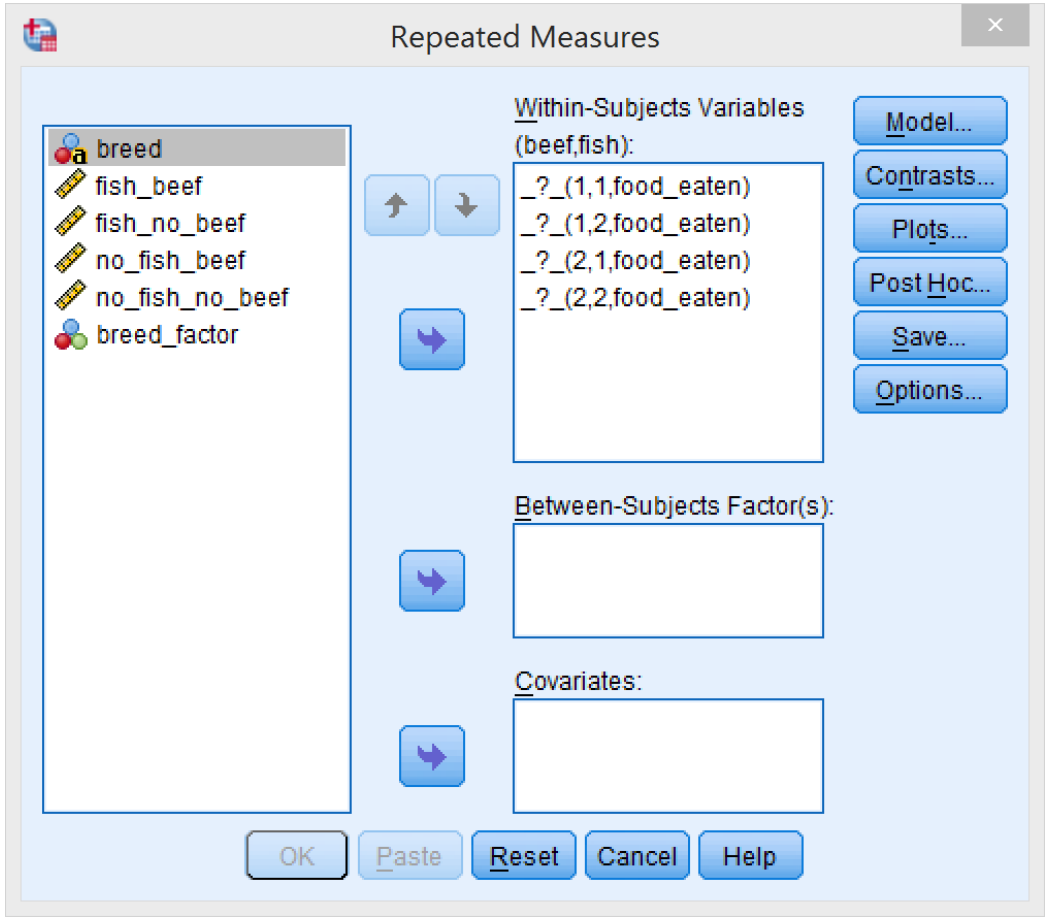
Run the analysis

- Add a name for the dependent measure (e.g. “food_eaten”)
- Click “Add” and “Define”.



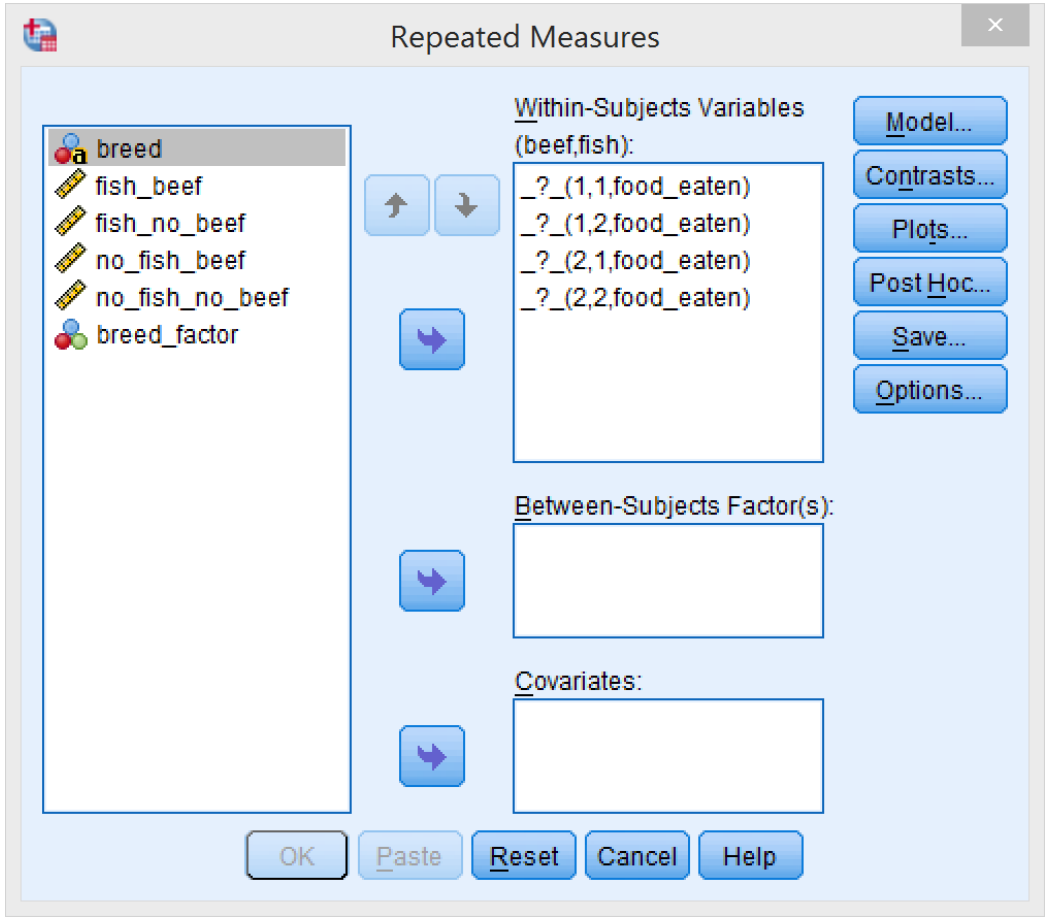
Assign within-subjects variables

- Click on each column name, then on the corresponding level of the within subjects variables and then on the arrow to assign it.
- You need to decide which level corresponds to which numbers
  - e.g. 1,1 could correspond to beef, fish and 2,2 could correspond to no fish, no beef, etc.



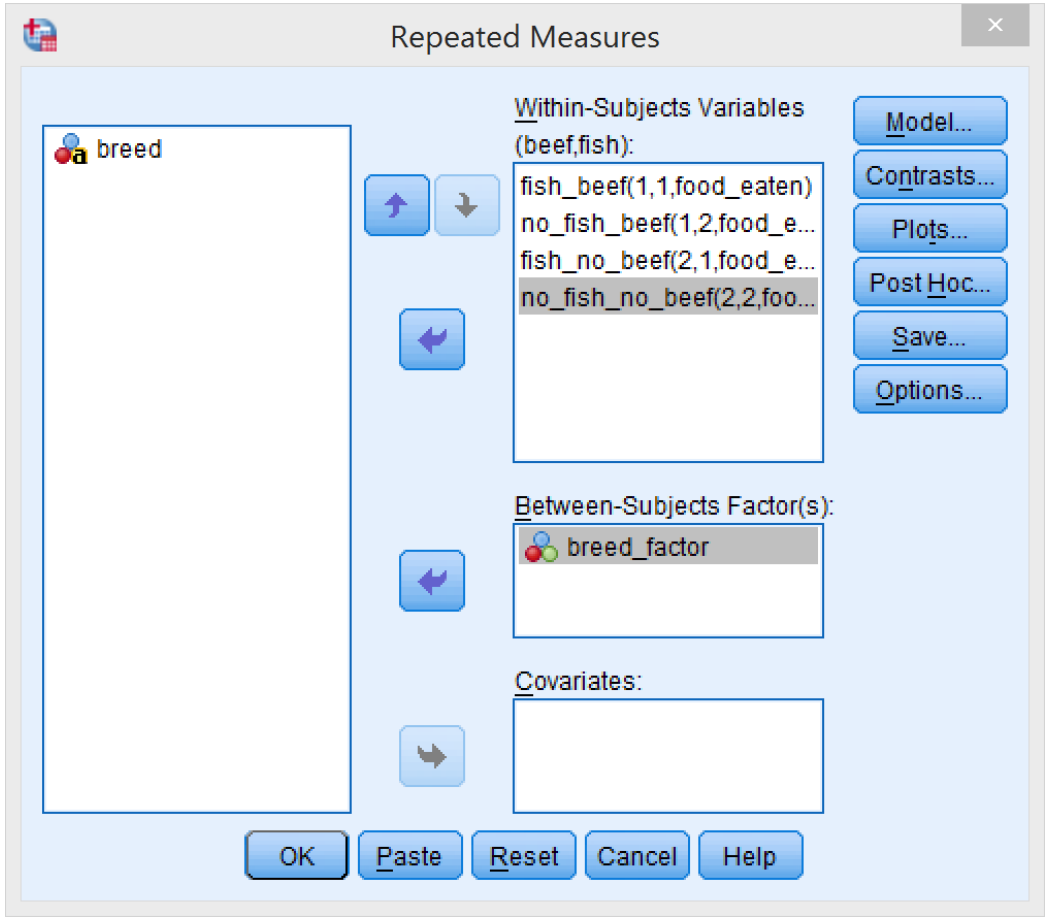
Assign within-subjects variables

- Click each corresponding level of the within subjects variables, then on the column name on the left, and then on the arrow to assign it.
- You need to decide which level corresponds to which numbers
  - e.g. 1,1 could correspond to beef, fish and 2,2 could correspond to no fish, no beef, 1,2 could correspond to beef, no fish, and 2,1 could correspond to no beef, fish
- If you get this wrong, you'll be very confused later if there's an interaction!



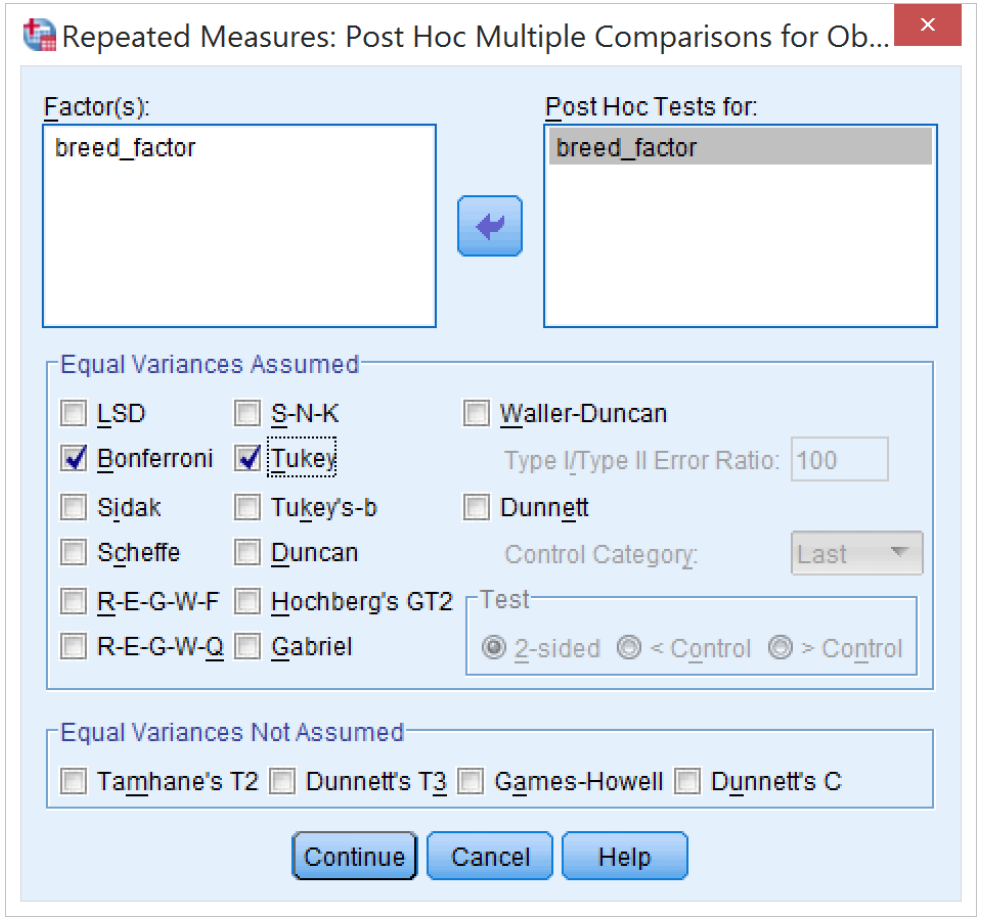
Assign between-subjects factor

- Assign the recoded factor (breed_factor) to the Between-Subjects Factor(s) field.



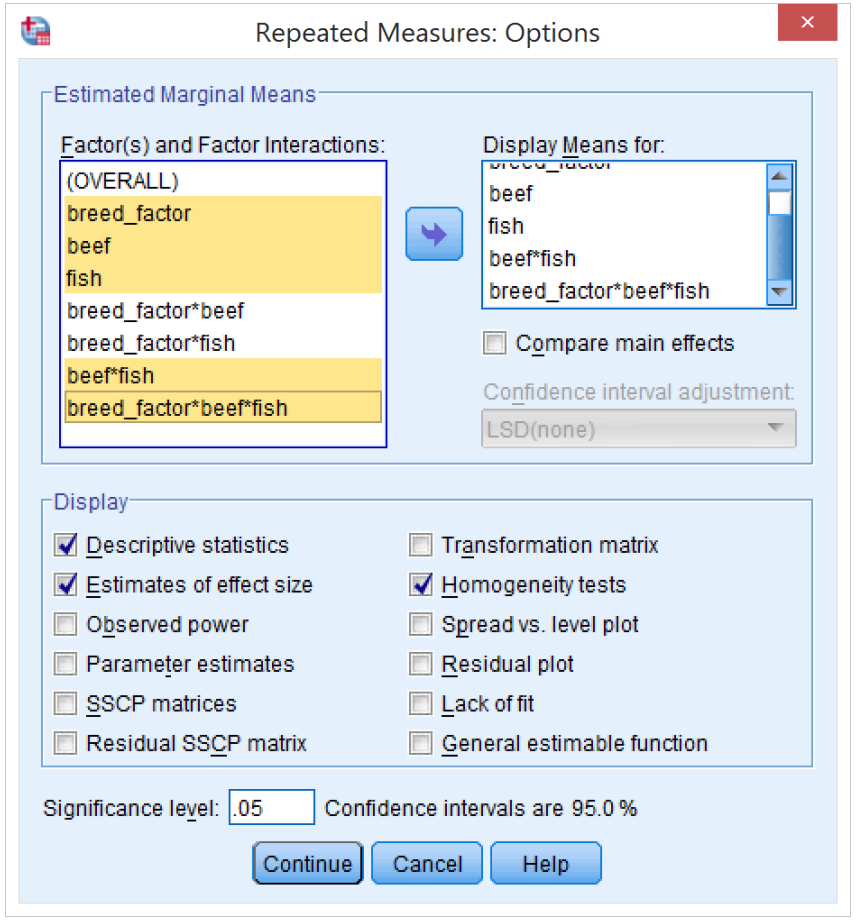
Post-hoc comparisons

- Click “Post Hoc…”
- Note that only factors with 3 or more levels are shown in the post hoc window
- Select “breed_factor” and move it over to the “Post hoc tests for” field
- Select any tests you want, e.g. Bonferroni and Tukey
- Click “Continue”



Options

- Click “Options…”
- Get marginal means for all main effects and the three way interaction (the two-way interactions are less interesting)
- Get descriptive statistics and estimates of effect size
- Get homogeneity tests
- Click “Continue”
- Click “OK” to run the analysis.



Output

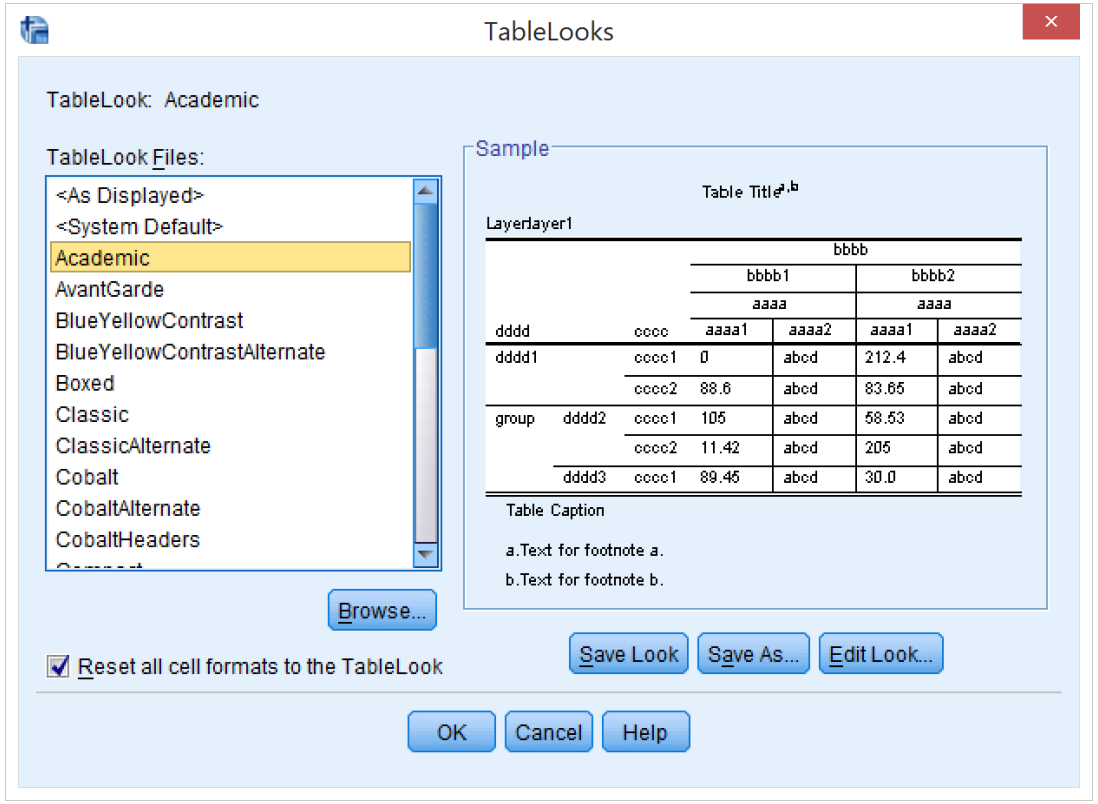
- Here you have (almost) everything you need for your homework report
- First, get the descriptive statistics table
- You can edit this table by double-clicking on it

Descriptive Statistics

	breed_factor	Mean	Std. Deviation	N
fish_beef	Manx	108.30	35.097	10
	Persian	75.40	20.961	10
	Shorthair	62.60	17.122	10
	Siamese	69.00	42.045	10
	Total	78.83	34.341	40
no_fish_beef	Manx	116.10	40.154	10
	Persian	85.70	24.432	10
	Shorthair	78.30	20.876	10
	Siamese	79.30	34.916	10
	Total	89.85	33.698	40
fish_no_beef	Manx	134.90	41.235	10
	Persian	106.30	20.056	10
	Shorthair	98.00	21.833	10
	Siamese	102.00	40.642	10
	Total	110.30	34.527	40
no_fish_no_beef	Manx	88.30	36.363	10
	Persian	59.90	24.076	10
	Shorthair	56.40	22.127	10
	Siamese	64.10	38.740	10
	Total	67.18	32.533	40

Output

- After double-clicking the table, right-click it and select TableLooks…
- Select “Academic”
- This will remove all vertical lines
- According to APA style, tables should not have vertical lines.



Output

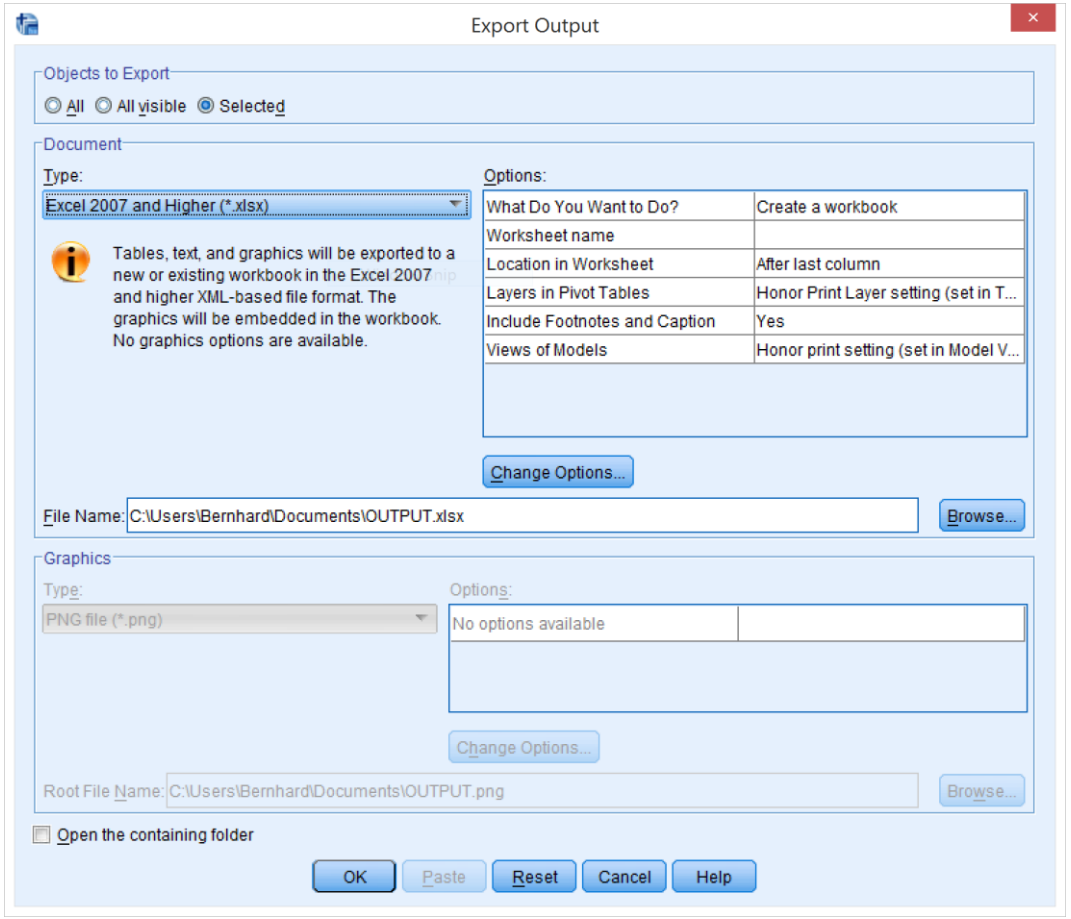
- Still in the table editing mode, double-click on the condition labels to edit them into something nicer-looking

Descriptive Statistics

	breed_factor	Mean	Std. Deviation	N
fish_beef	Manx	108.30	35.097	10
	Persian	75.40	20.961	10
	Shorthair	62.60	17.122	10
	Siamese	69.00	42.045	10
	Total	78.83	34.341	40
no_fish_beef	Manx	116.10	40.154	10
	Persian	85.70	24.432	10
	Shorthair	78.30	20.876	10
	Siamese	79.30	34.916	10
	Total	89.85	33.698	40
fish_no_beef	Manx	134.90	41.235	10
	Persian	106.30	20.056	10
	Shorthair	98.00	21.833	10
	Siamese	102.00	40.642	10
	Total	110.30	34.527	40
no_fish_no_beef	Manx	88.30	36.363	10
	Persian	59.90	24.076	10
	Shorthair	56.40	22.127	10
	Siamese	64.10	38.740	10
	Total	67.18	32.533	40

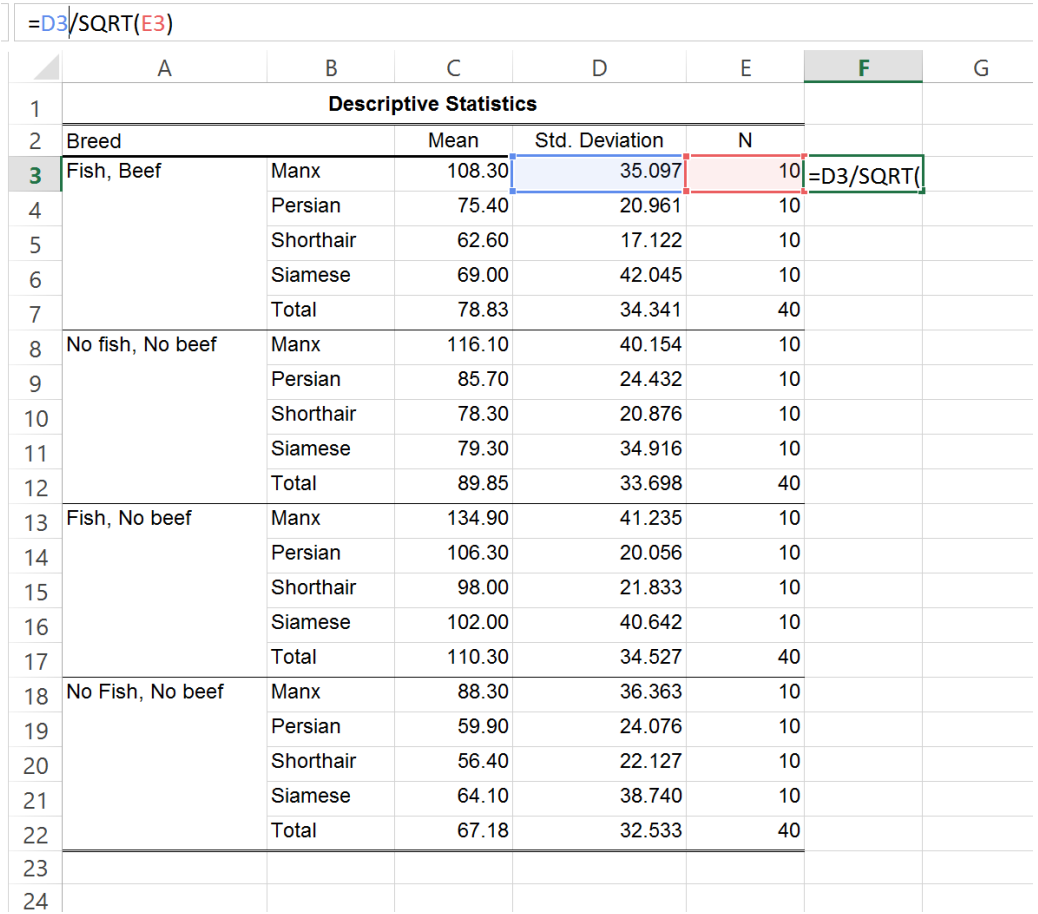
Output

- Now it's time to get this table in Excel (and then in Word). Exit the table edit mode by clicking outside it, then right-click it and select "Export…"
- Select a file name and path that you can find again later on…
- Click OK.
- Now open the file in Excel



Prepare table in Excel

- You need to add a column for the Standard error of the mean.
- Remember the formula? It's standard deviation / square root of the sample size
- In a new column in Excel, enter that formula (in this case, it should be =D3/SQRT(E3)
- Press ENTER, then double-click the little square at the lower-right corner of the cell



	A	B	C	D	E	F	G
1	Descriptive Statistics						
2	Breed		Mean	Std. Deviation	N		
3	Fish, Beef	Manx	108.30	35.097	10	=D3/SQRT(	
4		Persian	75.40	20.961	10		
5		Shorthair	62.60	17.122	10		
6		Siamese	69.00	42.045	10		
7		Total	78.83	34.341	40		
8	No fish, No beef	Manx	116.10	40.154	10		
9		Persian	85.70	24.432	10		
10		Shorthair	78.30	20.876	10		
11		Siamese	79.30	34.916	10		
12		Total	89.85	33.698	40		
13	Fish, No beef	Manx	134.90	41.235	10		
14		Persian	106.30	20.056	10		
15		Shorthair	98.00	21.833	10		
16		Siamese	102.00	40.642	10		
17		Total	110.30	34.527	40		
18	No Fish, No beef	Manx	88.30	36.363	10		
19		Persian	59.90	24.076	10		
20		Shorthair	56.40	22.127	10		
21		Siamese	64.10	38.740	10		
22		Total	67.18	32.533	40		
23							
24							

Prepare table in Excel

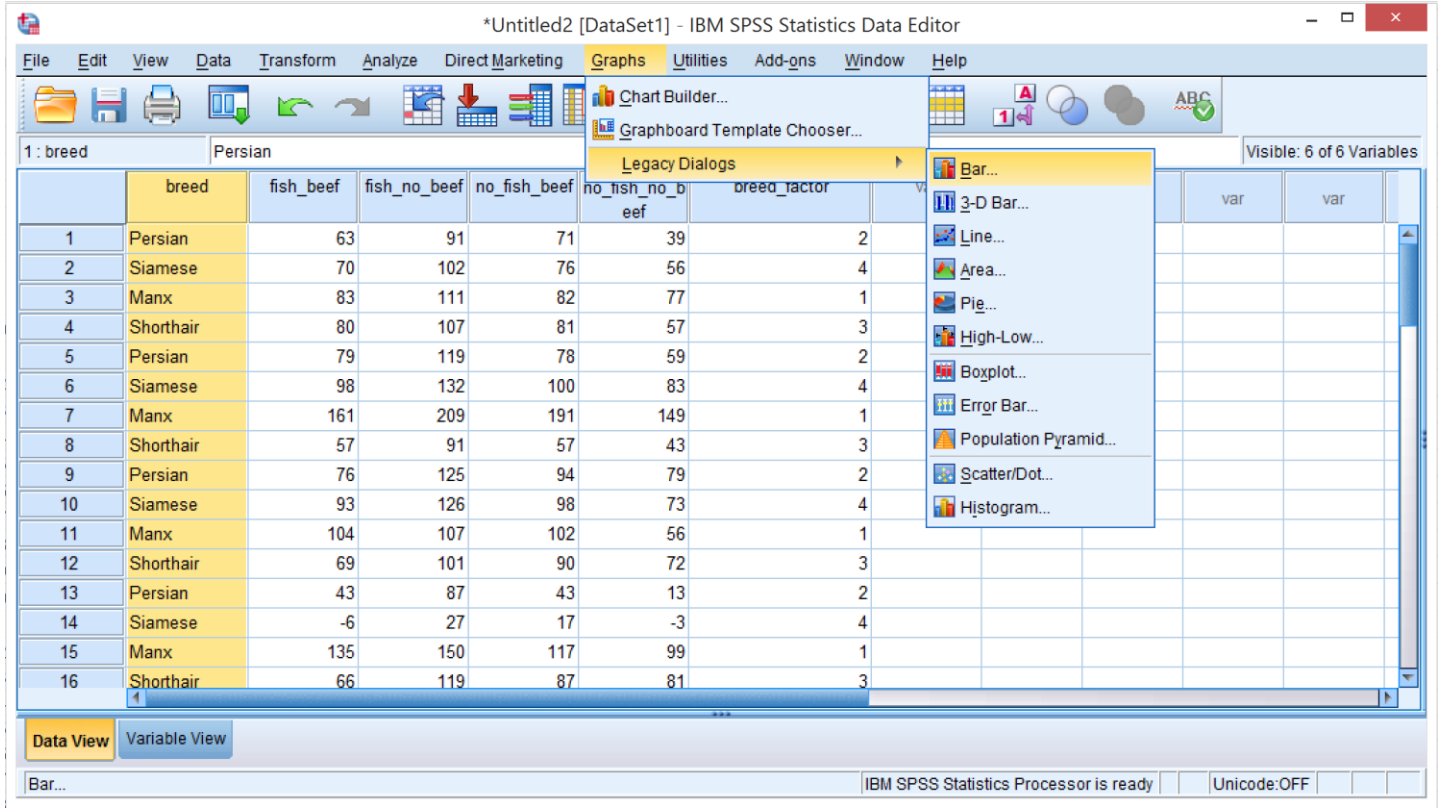
- Now just fix the formatting (Right-click + Format cells…)
  - Set Number format to Number with 2 decimal places
  - Add the horizontal lines (Format cells → Border)
  - Align the “Breed” label by unclicking “Merge and Center”
- When you’re done, copy and paste this table to Word.
- Don’t forget to add a caption!

Descriptive Statistics

Breed		Mean	Std. Deviation	N	SE
Fish, Beef	Manx	108.30	35.097	10	11.0986
	Persian	75.40	20.961	10	6.628558
	Shorthair	62.60	17.122	10	5.414384
	Siamese	69.00	42.045	10	13.29578
	Total	78.83	34.341	40	5.429724
No fish, No beef	Manx	116.10	40.154	10	12.69773
	Persian	85.70	24.432	10	7.72593
	Shorthair	78.30	20.876	10	6.601431
	Siamese	79.30	34.916	10	11.04139
	Total	89.85	33.698	40	5.328149
Fish, No beef	Manx	134.90	41.235	10	13.03964
	Persian	106.30	20.056	10	6.342187
	Shorthair	98.00	21.833	10	6.904105
	Siamese	102.00	40.642	10	12.85215
	Total	110.30	34.527	40	5.459196
No Fish, No beef	Manx	88.30	36.363	10	11.49884
	Persian	59.90	24.076	10	7.613511
	Shorthair	56.40	22.127	10	6.997142
	Siamese	64.10	38.740	10	12.25058
	Total	67.18	32.533	40	5.143939

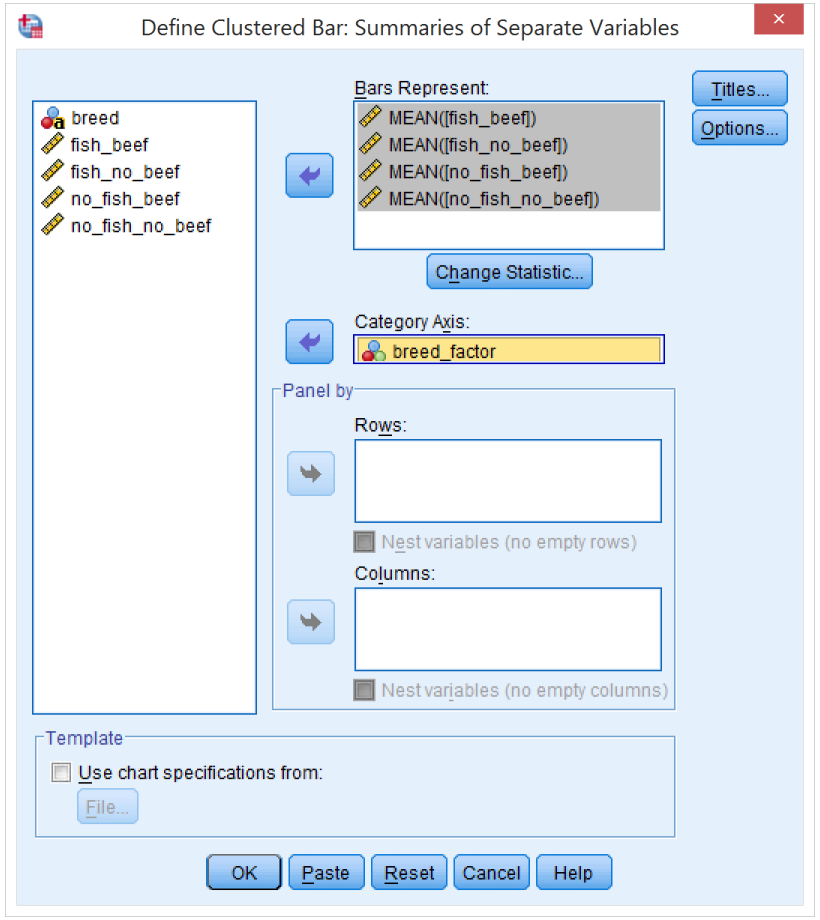
Make plots

- In SPSS, go to Graphs → Legacy Dialogs → Bar…



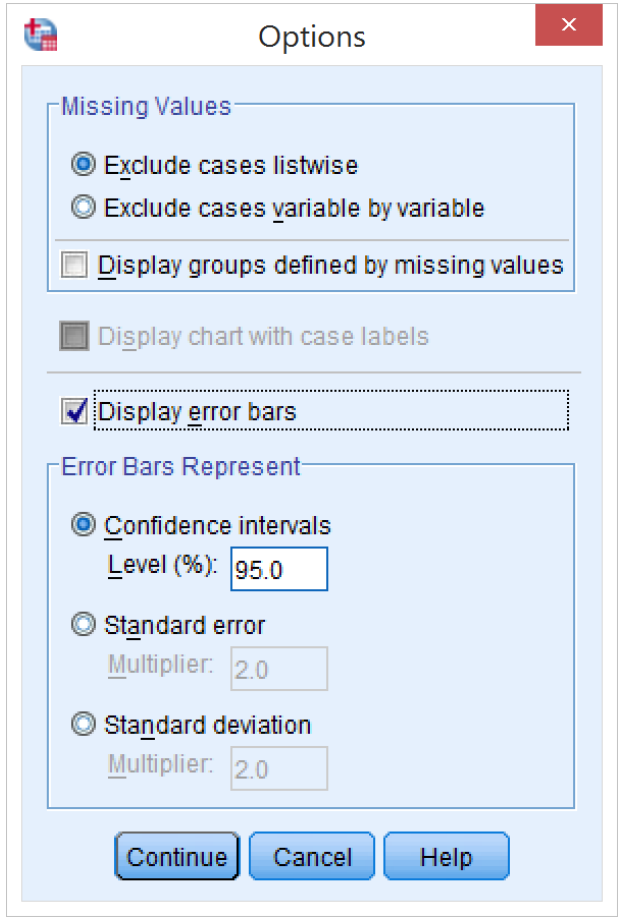
Make plots

- Assign the four different beef/fish columns to “Bars represent”
  - The default statistic (“MEAN”) is exactly what we want
- Assign the between-subject factor (“breed_factor”) to “Category Axis”



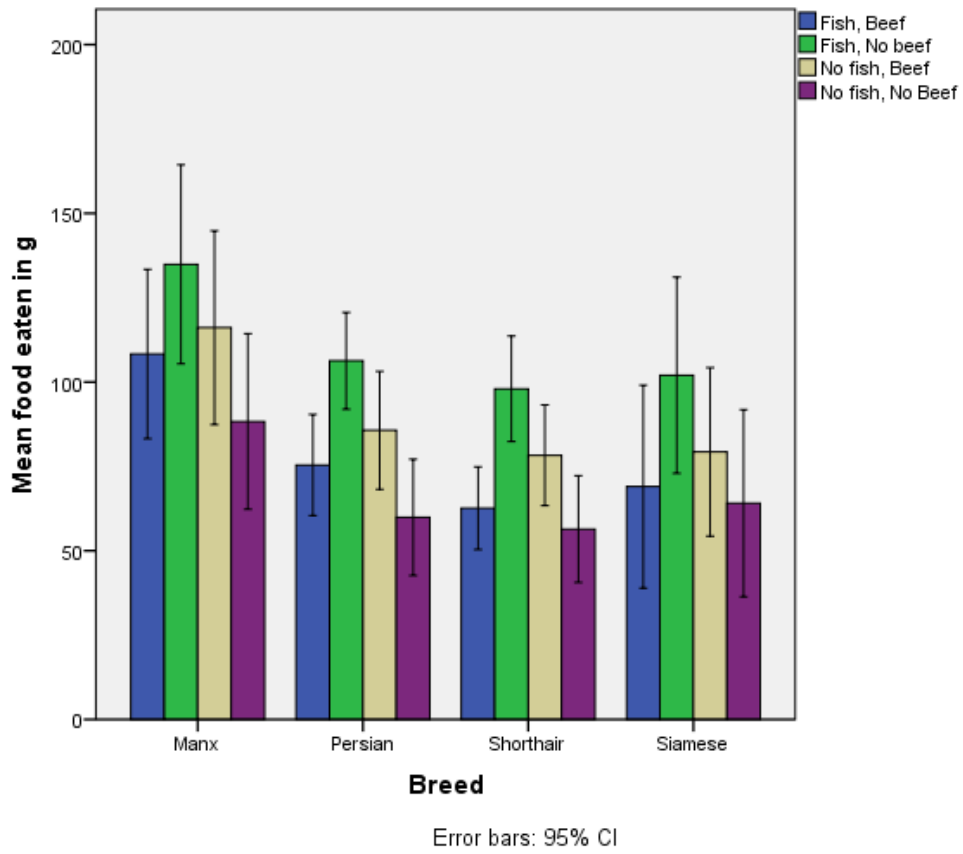
Add error bars

- Click “Options…”
- Check “Display error bars”
- The default setting (95% CI) is what we want.
- Click “Continue”



Final touches

- Double-click the resulting graph to open it in the chart editor
- Double-click axis and legend labels to edit them and make them look nice
- There we go!



The hypothesis tests

- Do NOT copy and paste any SPSS output directly
- Look at the tables and pick out the correct numbers
- First, for the within subjects effects, look at Mauchly's test of sphericity
- In this case, there is no Sig value, since neither beef nor fish have more than 2 levels
- If one of the tests is significant, check the Greenhouse-Geisser Epsilon
  - If it's over .75, use the Huynh-Feldt Epsilon to correct for sphericity
  - If it's under .75, use the G-G Epsilon

Mauchly's Test of Sphericity[a]

Measure: food_eaten

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
beef	1.000	.000	0	.	1.000	1.000	1.000
fish	1.000	.000	0	.	1.000	1.000	1.000
beef * fish	1.000	.000	0	.	1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + breed_factor
Within Subjects Design: beef + fish + beef * fish

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

The hypothesis tests

- The results of the F-tests on the Within-Subjects factors are in the “Tests of Within-Subjects Effects” table
- In the “Sig.” column, check which p-values are below .05
  - You have to report those effects in detail
- The correct error term is the first one listed *below* the predictor
- If you need to make a sphericity correction, pick the correct row (see last slide), otherwise use the values from the row “Sphericiy assumed”

The hypothesis tests

- Note down:
  - df numerator (df effect)
  - df denominator (df error)
  - F-value
  - p-value
  - Partial Eta squared
- Then report them like this in the text:
- There was a significant effect of beef, $F(1, 36) = 8.63$, $\eta_P^2 = 0.01$, $p < .01$.
- Do this for every significant factor.

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
beef	Sphericity Assumed	774.400	1	774.400	8.627	.006	.193
	Greenhouse-Geisser	774.400	1.000	774.400	8.627	.006	.193
	Huynh-Feldt	774.400	1.000	774.400	8.627	.006	.193
	Lower-bound	774.400	1.000	774.400	8.627	.006	.193
beef * breed_factor	Sphericity Assumed	541.950	3	180.650	2.012	.130	.144
	Greenhouse-Geisser	541.950	3.000	180.650	2.012	.130	.144
	Huynh-Feldt	541.950	3.000	180.650	2.012	.130	.144
	Lower-bound	541.950	3.000	180.650	2.012	.130	.144
Error(beef)	Sphericity Assumed	3231.650	36	89.768			
	Greenhouse-Geisser	3231.650	36.000	89.768			
	Huynh-Feldt	3231.650	36.000	89.768			
	Lower-bound	3231.650	36.000	89.768			

The hypothesis tests

- Note down:
  - df numerator (df effect)
  - df denominator (df error)
  - F-value
  - p-value
  - Partial Eta squared
- Then report them like this in the text:
- There was a significant effect of beef, $F(1, 36) = 8.63$, $\eta_P^2 = 0.01$, $p < .01$.
- Do this for every significant factor.

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
beef	Sphericity Assumed	774.400	1	774.400	8.627	.006	.193
	Greenhouse-Geisser	774.400	1.000	774.400	8.627	.006	.193
	Huynh-Feldt	774.400	1.000	774.400	8.627	.006	.193
	Lower-bound	774.400	1.000	774.400	8.627	.006	.193
beef * breed_factor	Sphericity Assumed	541.950	3	180.650	2.012	.130	.144
	Greenhouse-Geisser	541.950	3.000	180.650	2.012	.130	.144
	Huynh-Feldt	541.950	3.000	180.650	2.012	.130	.144
	Lower-bound	541.950	3.000	180.650	2.012	.130	.144
Error(beef)	Sphericity Assumed	3231.650	36	89.768			
	Greenhouse-Geisser	3231.650	36.000	89.768			
	Huynh-Feldt	3231.650	36.000	89.768			
	Lower-bound	3231.650	36.000	89.768			

The hypothesis tests

- For those factors with only two levels, you can interpret and report the marginal means directly
- There was a significant effect of beef, $F(1, 36) = 8.63$, $\eta_P^2 = 0.01$, $p < .01$. When there was beef present in the food, cats ate more food on average (88.7 g) than when there was no beef in the food (84.3 g).

3. beef

Measure: food_eaten

beef	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	84.338	4.800	74.603	94.072
2	88.738	4.930	78.739	98.736

The hypothesis tests

- In the presence of significant interactions, you do not have to interpret the main effects and marginal means in this way (as they are modulated by the interaction).
- But you do have to interpret the interaction
- The fish and beef main effects were modulated by a significant two-way interaction, $F(1, 36) = 571.82$ $\eta_P^2 = 0.17$, $p < .01$. This interaction showed that, when no beef was in the food, the cats ate more when there was fish in the food (mean eaten = 110 g) than when there was no fish (mean eaten = 67 g). When the food contained beef, cats ate less food when there was also fish in the food (mean eaten = 79 g) than when there was only beef (mean eaten = 90 g).

5. beef * fish

Measure: food_eaten

beef	fish	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	78.825	4.830	69.030	88.620
	2	89.850	4.914	79.883	99.817
2	1	110.300	5.142	99.871	120.729
	2	67.175	4.932	57.172	77.178

Assumption tests

- Mauchly's test (see previous slides)
- Levene's test
  - See table in the output
  - Note that this is split by within-subject conditions
  - Report if one or more of the tests are significant

Levene's Test of Equality of Error Variances[a]

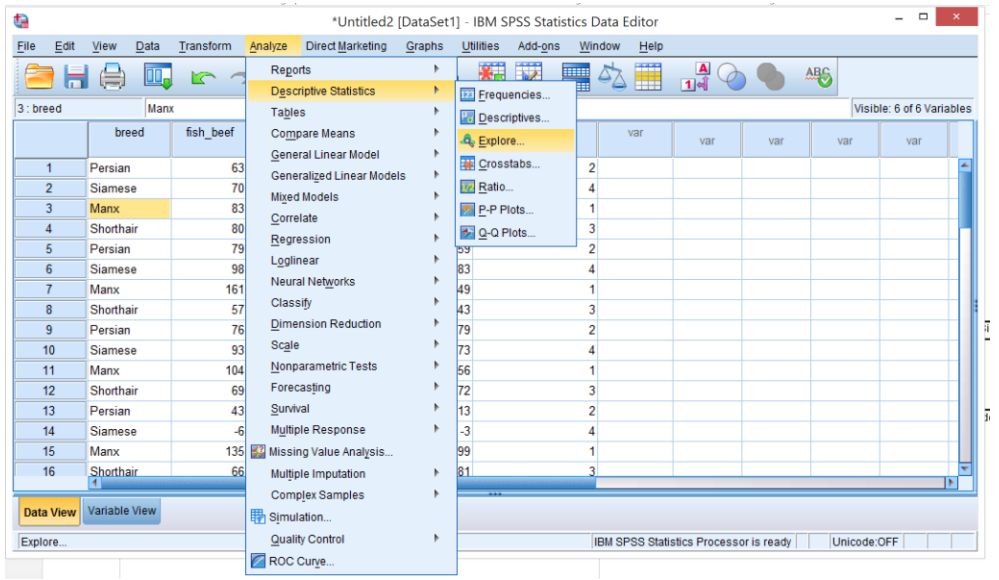
	F	df1	df2	Sig.
fish_beef	2.787	3	36	.055
no_fish_beef	1.196	3	36	.325
fish_no_beef	2.526	3	36	.073
no_fish_no_beef	1.211	3	36	.320

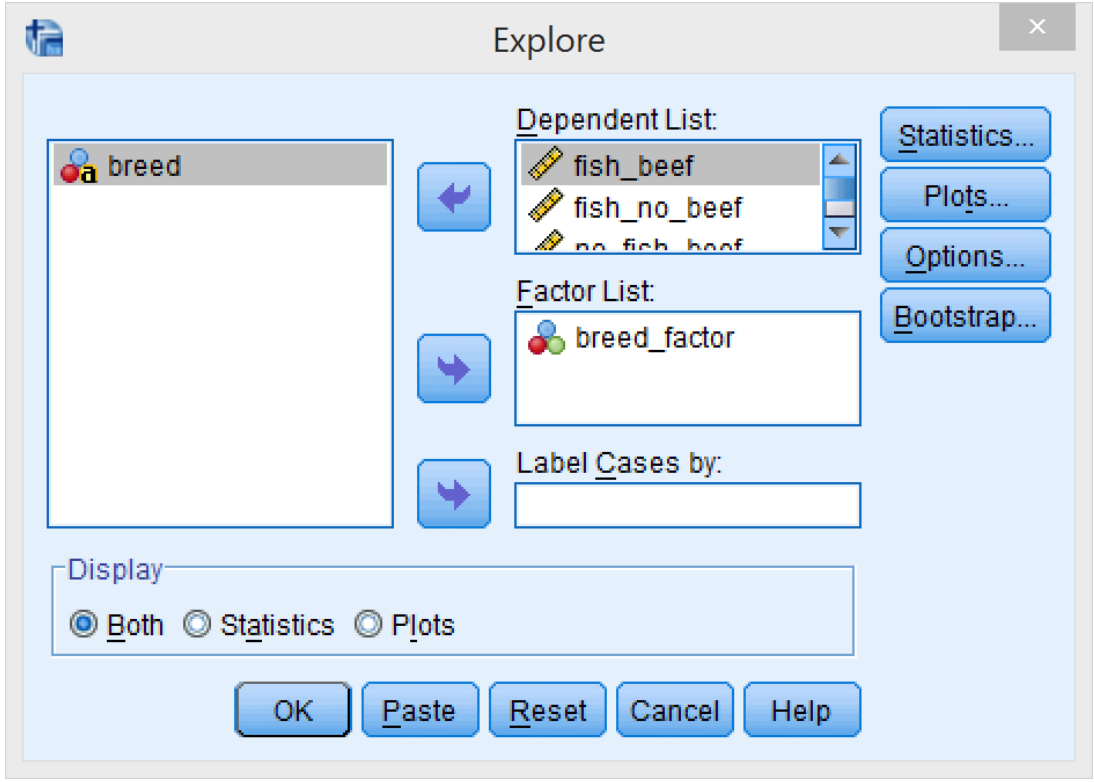
Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + breed_factor
Within Subjects Design: beef + fish + beef * fish

Assumption tests

- Normality
  - Surprisingly involved!
- Go to Analyze → Descriptive Statistics → Explore…
- Move all within-subjects columns to “Dependent List”
- Move the between-subjects factor to “factor list”
- Click on “Options…”





Assumption tests

- Check “Normality plots with tests”
- Click “Continue”
- Click “OK”
- Check the resulting table for p-values below .05
- If there are any, report them.

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	breed_factor	Statistic	df	Sig.	Statistic	df	Sig.
fish_beef	Manx	.151	10	.200*	.977	10	.947
	Persian	.154	10	.200*	.963	10	.816
	Shorthair	.208	10	.200*	.946	10	.622
	Siamese	.209	10	.200*	.935	10	.495
fish_no_beef	Manx	.149	10	.200*	.970	10	.892
	Persian	.254	10	.068	.874	10	.111
	Shorthair	.240	10	.107	.924	10	.391
	Siamese	.200	10	.200*	.958	10	.766
no_fish_beef	Manx	.191	10	.200*	.980	10	.964
	Persian	.138	10	.200*	.984	10	.982
	Shorthair	.126	10	.200*	.982	10	.976
	Siamese	.220	10	.185	.925	10	.402
no_fish_no_beef	Manx	.158	10	.200*	.966	10	.852
	Persian	.186	10	.200*	.914	10	.312
	Shorthair	.172	10	.200*	.888	10	.160
	Siamese	.186	10	.200*	.964	10	.831

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Post-hoc tests

- Look at the “Multiple Comparisons” table
- Choose the test with the highest power (Tukey’s in this case)
- Check Sig column for < .05
- Report those effects along with the means

Multiple Comparisons

Measure: food_eaten

	(I) breed_factor	(J) breed_factor	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Manx	Persian	30.08	13.597	.139	-6.55	66.70
		Shorthair	38.08*	13.597	.039	1.45	74.70
		Siamese	33.30	13.597	.086	-3.32	69.92
	Persian	Manx	-30.08	13.597	.139	-66.70	6.55
		Shorthair	8.00	13.597	.935	-28.62	44.62
		Siamese	3.22	13.597	.995	-33.40	39.85
	Shorthair	Manx	-38.08*	13.597	.039	-74.70	-1.45
		Persian	-8.00	13.597	.935	-44.62	28.62
		Siamese	-4.78	13.597	.985	-41.40	31.85
	Siamese	Manx	-33.30	13.597	.086	-69.92	3.32
		Persian	-3.22	13.597	.995	-39.85	33.40
		Shorthair	4.78	13.597	.985	-31.85	41.40
Bonferroni	Manx	Persian	30.08	13.597	.200	-7.89	68.04
		Shorthair	38.08*	13.597	.049	.11	76.04
		Siamese	33.30	13.597	.116	-4.66	71.26
	Persian	Manx	-30.08	13.597	.200	-68.04	7.89
		Shorthair	8.00	13.597	1.000	-29.96	45.96
		Siamese	3.22	13.597	1.000	-34.74	41.19
	Shorthair	Manx	-38.08*	13.597	.049	-76.04	-.11
		Persian	-8.00	13.597	1.000	-45.96	29.96
		Siamese	-4.78	13.597	1.000	-42.74	33.19
	Siamese	Manx	-33.30	13.597	.116	-71.26	4.66
		Persian	-3.22	13.597	1.000	-41.19	34.74
		Shorthair	4.78	13.597	1.000	-33.19	42.74

Based on observed means.
The error term is Mean Square(Error) = 924.457.

*. The mean difference is significant at the .05 level.

Writing things up

- Refer to the “Homework 4 worked.docx” file for guidance.