Additional Multivariate Exercise 13

Data: The expanded dataset on canine mandible measurements is described in detail in the exercise. Compared to the previously analyzed prehistoric dogs dataset, two species have been omitted (Chinese wolf and dingo), individual samples have been added for the species present (ranging from 10 and 20 samples), and the sex of the animal is available for 4 of the species. Finally, the mandible measurements have been expanded from 6 to 9, with variables $x_4 - x_8$ of the present, expanded dataset apparently also included in the aggregated data.

Descriptive statistics: With 77 samples distributed on 5 species and 2 gender groups, and 9 measurements per sample, many options exist for graphical representations. The most obvious ones are perhaps a matrix scatterplot with symbols corresponding to species (perhaps also sexes), and for each of the variables boxplots and intervals plots across species. To keep this solution reasonably brief, we do not include any of these plots, but present a table of descriptive statistics across periods and variables (without accounting for sex, but this was requested for the first question of the exercise).

Statistic	Species	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
mean	cuon	133	10.7	24.1	23.6	21.5	8.49	29.0	37.7	6.61
	golden jackal	111	8.2	18.6	17.0	18.2	6.82	30.4	33.4	4.81
	Indian wolf	157	11.6	26.2	24.7	24.7	9.34	40.2	44.8	7.41
	modern dog	126	9.7	21.4	21.1	19.4	7.68	32.1	36.6	5.87
	prehist. dog	123	10.3	20.0	22.9	19.3	8.19	32.8	35.9	6.17
one-way ANOVA: F		71.1	54.2	28.5	53.8	110	52.0	110	58.0	69.5
stand. dev.	cuon	6.4	0.61	1.68	1.58	1.01	0.34	1.50	1.93	0.48
	golden jackal	3.9	0.50	1.14	0.97	0.83	0.48	1.14	1.18	0.31
	Indian wolf	12.5	0.90	3.87	1.90	1.14	0.67	2.04	3.04	0.68
	modern dog	8.5	0.85	2.31	1.63	0.89	0.48	1.44	2.73	0.42
	prehist. dog	8.4	0.77	1.94	2.85	0.95	0.75	2.10	1.85	0.48
Levene's equal var. test: P		.047	.27	.003	.036	.77	.42	.13	.025	.18

All F-statistics for testing equality of means are very strongly significant (e.g., $F_{.999}(4, 72) = 5.18$), whereas only the variance test for x_3 is strongly significant (with x_1, x_4, x_8 weakly significant). The normal plots for the (standardized) residuals in each of the univariate analyses look reasonably good; a few outlying points are seen, and the residual distribution for x_2 seems somewhat right-skewed, but there are no glaring deviations from normality. With the fairly large sample size, the normality assumption looks tenable. Some of the plots show increasing variance with the means, and the equal variances assumption is probably the biggest concern for the multivariate analysis. It could be suggested to try a Box-Cox analysis for individual variables to utilize any improvements by transformation. It would seem most natural to use the same transformation for all variables, even if the Box-Cox analysis will typically suggest different values for the power λ. The Box-Cox results suggest that an across-the-board log transformation could work well. We will, however, for this solution keep the analysis on original scale. The pairwise correlations between variables (estimated within species) are almost all positive, but of visibly larger magnitude within some species than others; the weakest correlations are for the golden jackal, and all other species have some pairwise correlations above 0.8 (details not shown).

Multivariate tests: The MANOVA test and Box M-test for equality of mean vectors and variance matrices across periods are, not unsurprisingly, clearly significant, as seen in the listings below.

```
. manova x1-x9=Species
                       Number of obs =         77

                       W = Wilks' lambda      L = Lawley-Hotelling trace
                       P = Pillai's trace     R = Roy's largest root

             Source | Statistic        df    F(df1,     df2) =   F   Prob>F
          -----------+------------------------------------------------------
            Species |W   0.0022         4     36.0    241.6    27.67 0.0000 a
                    |P   2.5892               36.0    268.0    13.66 0.0000 a
                    |L  25.1290               36.0    250.0    43.63 0.0000 a
                    |R  16.3476                9.0     67.0   121.70 0.0000 u
                    |-------------------------------------------------------
           Residual |                  72
          -----------+------------------------------------------------------
              Total |                  76
          ------------------------------------------------------------------
                       e = exact, a = approximate, u = upper bound on F

. mvtest covariances x1-x9, by(Species)
Test of equality of covariance matrices across 5 samples

      Modified LR chi2 =  364.9761
     Box F(180,6213.6) =      1.40      Prob > F =  0.0004
         Box chi2(180) =    263.21   Prob > chi2 =  0.0001
```

These overall results are perhaps not too useful. The question specifically asks for comparisons between the prehistoric dog and the other species. For each variable individually, these would be two-sample t-tests, and in the multivariate setup we can use Hotelling's T^2-statistic. Significance levels may be adjusted for 4 comparisons, if desired. For the multivariate tests, it makes no difference for the conclusions in our situation (all tests are strongly significant), as seen from the (condensed) listings below.

```
. forval i=1(1)4 {
  2.   hotelling x1-x9 if Species==`i'|Species==5, by(Species)
  3.   }
--------------------------------------------------------------------------------------------------------------
-> Species = Cuon
-> Species = Prehistoric dog

2-group Hotelling's T-squared = 1083.885
F test statistic: ((27-9-1)/(27-2)(9)) x 1083.885 = 81.89353

H0: Vectors of means are equal for the two groups
              F(9,17) =   81.8935
       Prob > F(9,17) =    0.0000
--------------------------------------------------------------------------------------------------------------
-> Species = Golden jackal
-> Species = Prehistoric dog

2-group Hotelling's T-squared = 296.44117
F test statistic: ((30-9-1)/(30-2)(9)) x 296.44117 = 23.527077

H0: Vectors of means are equal for the two groups
              F(9,20) =   23.5271
       Prob > F(9,20) =    0.0000
```

```
--------------------------------------------------------------------------------------------------------
-> Species = Indian wolf
-> Species = Prehistoric dog

2-group Hotelling's T-squared = 243.2363
F test statistic: ((24-9-1)/(24-2)(9)) x 243.2363 = 17.198526

H0: Vectors of means are equal for the two groups
              F(9,14) =   17.1985
       Prob > F(9,14) =    0.0000
--------------------------------------------------------------------------------------------------------
-> Species = Modern dog
-> Species = Prehistoric dog

2-group Hotelling's T-squared = 119.61352
F test statistic: ((26-9-1)/(26-2)(9)) x 119.61352 = 8.8602607

H0: Vectors of means are equal for the two groups
              F(9,16) =    8.8603
       Prob > F(9,16) =    0.0001
```

Finally, we inspect the error correlation matrix. All correlations are positive, ranging from 0.25 for (x_3, x_7) to 0.80 for (x_1, x_8). The correlations with x_3 and x_4 tend to be smaller in magnitude, whereas the largest correlations are spread across several of the other variables.

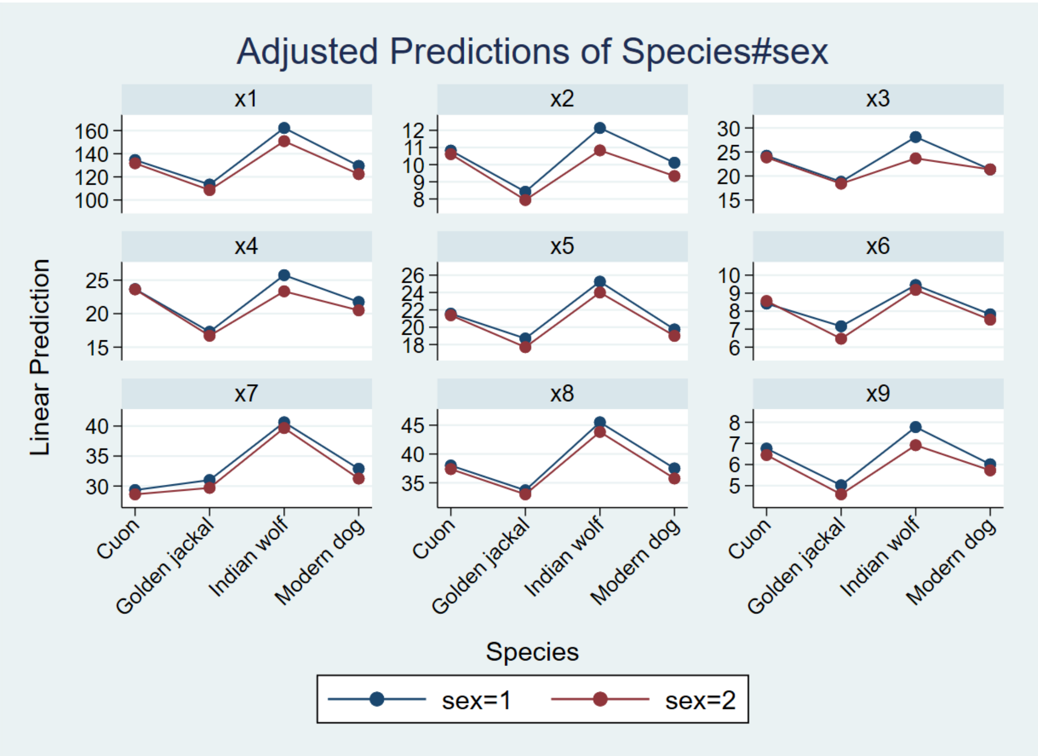
Multivariate analyses involving sex: All prehistoric dogs have the sex recorded as unknown, while the values of other species are either male or female. This means that if species and sex are included in a model together, there will be a complete confounding/collinearity between the effect of prehistoric dogs and unknown sex categories (only one of them will be estimated). This is typically undesired because it confuses the interpretations, and it it is clear that a meaningful comparison of prehistoric dogs to other species will only be possible when some decision has been made about their sexes. In a simple way, one could set the sexes to either all female or all male, and view this as a sensitivity analysis. If the conclusions from comparisons between prehistoric dogs and the other species are not strongly affected, one could have some confidence that they are valid generally. Another option could be to look for obvious size-related clusters in the measurements from prehistoric dogs, and interpret the clusters to reflect sex and insert according values. This is an imputation approach, and more rigorous imputation methods exist (beyond our scope).

For our purposes here, we will exclude the prehistoric dogs from models involving sex. We start by showing the results of the two-way MANOVA.

```
 manova x1-x9=Species##sex if Species<5

                      Number of obs =         67

                      W = Wilks' lambda      L = Lawley-Hotelling trace
                      P = Pillai's trace     R = Roy's largest root

             Source | Statistic        df    F(df1,     df2) =   F   Prob>F
         -----------+--------------------------------------------------
              Model |W   0.0011         7     63.0    293.3    10.85 0.0000 a
                    |P   2.7665               63.0    399.0     4.14 0.0000 a
                    |L  36.3812               63.0    345.0    28.46 0.0000 a
```

```
             |R   24.8665               9.0      57.0    157.49 0.0000 u
             |-----------------------------------------------------------
    Residual |                   59
-------------+-----------------------------------------------------------
     Species |W    0.0020         3    27.0     149.6     40.69 0.0000 a
             |P    2.3270              27.0     159.0     20.36 0.0000 a
             |L   34.5979              27.0     149.0     63.64 0.0000 a
             |R   24.3179               9.0      53.0    143.21 0.0000 u
             |-----------------------------------------------------------
         sex |W    0.5903         1     9.0      51.0      3.93 0.0008 e
             |P    0.4097               9.0      51.0      3.93 0.0008 e
             |L    0.6942               9.0      51.0      3.93 0.0008 e
             |R    0.6942               9.0      51.0      3.93 0.0008 e
             |-----------------------------------------------------------
 Species#sex |W    0.5082         3    27.0     149.6      1.45 0.0866 a
             |P    0.5849              27.0     159.0      1.43 0.0932 a
             |L    0.7937              27.0     149.0      1.46 0.0809 a
             |R    0.4983               9.0      53.0      2.93 0.0067 u
             |-----------------------------------------------------------
    Residual |                   59
-------------+-----------------------------------------------------------
       Total |                   66
-------------------------------------------------------------------------
                 e = exact, a = approximate, u = upper bound on F
```

The interaction tests are hovering around the usual significance level, and one would want to explore this further. The usual next step is interaction plots, and one may try something like the following (it is not so easy to get Stata to display a nice panel of interaction plots...):



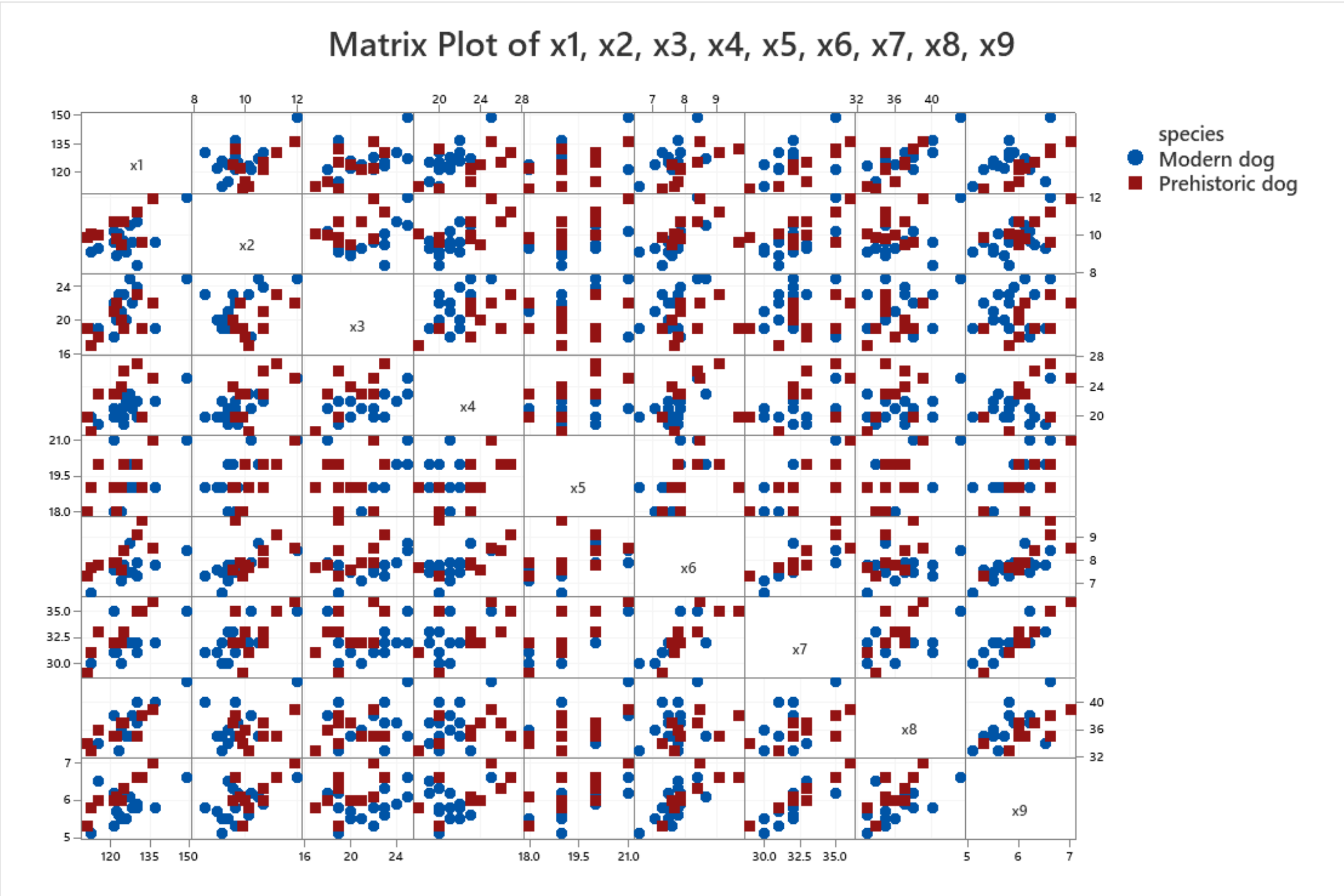
It is probably still not quite obvious where any important interaction effects reside. Some different ways to continue from here (see do-file for commands to implement these):

a) look at the univariate analyses (or the variable-specific estimates), to possibly identify those variables that show signs of an interaction,

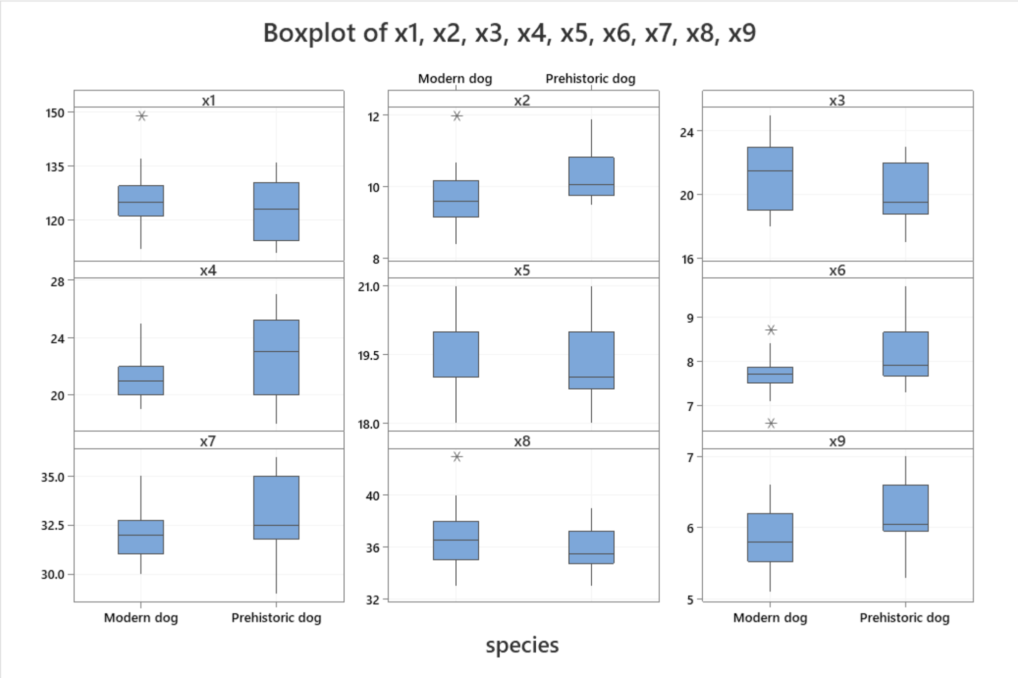
b) look at each species separately and test for effects of sex, the interest then being whether differences among the species in those results can be seen.

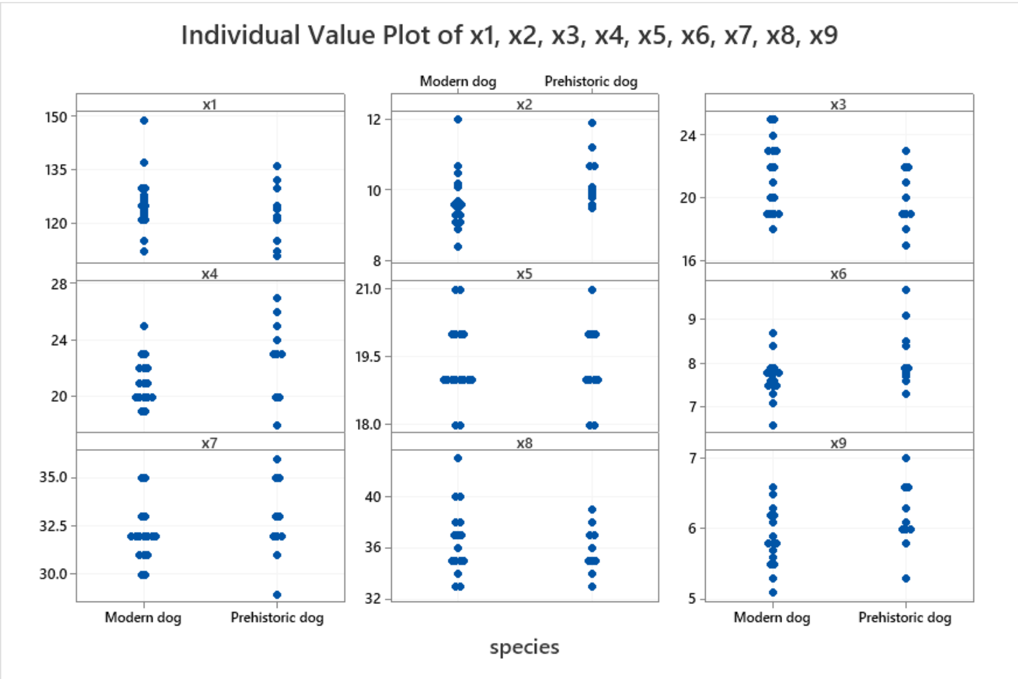
In an additive model, the multivariate test for sex gives $F = 3.23$, corresponding to $P = 0.003$ in an $F(9, 54)$-distribution. This strong significance clearly borrows strength across the 4 species, because the species-specific multivariate tests for sex have much weaker P-values ($P = 0.25, 0.058, 0.040, 0.40$ for the 4 species). The parameter estimates show that females consistently have smaller mandible dimensions than males, with different magnitudes (and P-values) across the 9 variables.

Graphical comparisons of prehistoric and modern dogs: Several types of plots are possible. In particular, when combining with dimension-reduction techniques or distance-based techniques there are many ways to represent differences between the two species. Here we will however focus on simple graphical displays, not preceded by analysis. A scatterplot matrix with different symbols for the two species is an option, this is easily done in Minitab.



The plot does get difficult to look at with so many variables, and although it is useful to be able to see the variables against each other one will also want to look at the distributions for each variable within the two species. Boxplots or individual value plots will do this. Both were constructed in Minitab by menus for "One Y, With Groups", with multiple variables graphed and the choice of separate panels within the same graph in the Multiple Graphs submenu.





Looking at these graphs, there does not appear to be a simple division between the species based on a cut-off for a single variable; all distributions show overlap.