

## Effect of exercise type and diet on pulse rates

- 18 subjects in study to assess effects of exercise type and diet on pulse rate.
- Each subject randomly assigned to one of three types of exercise:
  - racquetball,
  - aerobic stair climbing,
  - weights.
- Each subject also classified as a meat-eater or a vegetarian (with 3 meat-eaters and 3 vegetarians doing each type of exercise.)
- Response variable: pulse rate.
- So far, two-factor ANOVA with 3 replicates per combo of exercise- type and diet.
- But: **3 responses** for each subject: pulse rate at low, moderate, high intensity in their exercise. Makes it **repeated measures**.
- 3 factors:
  - diet and exercise type “between subjects”,
  - intensity-of-exercise “within subjects”.
- MANOVA to assess within-subjects factors and their interactions with all else
- ordinary ANOVA to assess between-subjects factors & interactions with each other. (Based on response averaged over the repeated measures.)

## Data:

stairs	112	166	215	meat
stairs	111	166	225	meat
stairs	89	132	189	meat
stairs	95	134	186	veg
stairs	86	109	150	veg
stairs	89	119	177	veg
racquetball	105	177	241	meat
racquetball	85	117	186	meat
racquetball	97	137	185	meat
racquetball	93	151	197	veg
racquetball	77	122	178	veg
racquetball	78	119	173	veg
weights	81	134	205	meat
weights	88	133	180	meat
weights	88	157	224	meat
weights	58	99	131	veg
weights	85	132	186	veg
weights	78	110	164	veg

SAS code:

```
options ls=70;
```

```
data ex;  
  infile "exercise.dat";  
  input exercise $ pulse1 pulse2 pulse3  
    diet $;
```

```
proc means;  
  class exercise diet;  
  var pulse1--pulse3;
```

```
proc glm data=ex;  
  class exercise diet;  
  model pulse1 pulse2 pulse3 =  
    exercise diet / noint;  
  repeated intensity;
```

Would normally do "exercise|diet" on model line,  
but interaction turns out non-significant, so  
removed from here.

## Proc means output:

exercise	diet	NObs	Variable	N	Mean	Std Dev
-----						
racquetb	meat	3	pulse1	3	95.6666667	10.0664459
			pulse2	3	143.6666667	30.5505046
			pulse3	3	204.0000000	32.0468407
	veg	3	pulse1	3	82.6666667	8.9628864
			pulse2	3	130.6666667	17.6729549
			pulse3	3	182.6666667	12.6622799
stairs	meat	3	pulse1	3	104.0000000	13.0000000
			pulse2	3	154.6666667	19.6299092
			pulse3	3	209.6666667	18.5831465
	veg	3	pulse1	3	90.0000000	4.5825757
			pulse2	3	120.6666667	12.5830574
			pulse3	3	171.0000000	18.7349940
weights	meat	3	pulse1	3	85.6666667	4.0414519
			pulse2	3	141.3333333	13.5769412
			pulse3	3	203.0000000	22.0680765
	veg	3	pulse1	3	73.6666667	14.0118997
			pulse2	3	113.6666667	16.8027775
			pulse3	3	160.3333333	27.6827263
-----						

- Must be intensity-of-exercise effect
- Might be diet effect
- Probably not a type-of-exercise effect.

## Test for intensity-of-exercise effect:

MANOVA Test Criteria and Exact F Statistics  
for the Hypothesis of no intensity Effect  
H = Type III SSCP Matrix for intensity  
E = Error SSCP Matrix

S=1      M=0      N=5.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.01778139	359.05	2	13	<.0001
Pillai's Trace	0.98221861	359.05	2	13	<.0001
Hotelling-Lawley Trace	55.23856418	359.05	2	13	<.0001
Roy's Greatest Root	55.23856418	359.05	2	13	<.0001

Strongly significant as expected.

## Intensity-of-exercise by exercise type interaction:

MANOVA Test Criteria and F Approximations for  
the Hypothesis of no intensity\*exercise Effect  
H = Type III SSCP Matrix for intensity\*exercise  
E = Error SSCP Matrix

S=2      M=-0.5      N=5.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.88850421	0.40	4	26	0.8098
Pillai's Trace	0.11235730	0.42	4	28	0.7952
Hotelling-Lawley Trace	0.12451746	0.40	4	14.623	0.8079
Roy's Greatest Root	0.11617098	0.81	2	14	0.4633

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

NOTE: F Statistic for Wilks' Lambda is exact.

Nothing here. Whatever effect of exercise type is same at all intensity levels.

## Intensity-of-exercise by diet interaction:

MANOVA Test Criteria and Exact F Statistics for  
the Hypothesis of no intensity\*diet Effect

H = Type III SSCP Matrix for intensity\*diet

E = Error SSCP Matrix

S=1      M=0      N=5.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.64025576	3.65	2	13	0.0551
Pillai's Trace	0.35974424	3.65	2	13	0.0551
Hotelling-Lawley Trace	0.56187584	3.65	2	13	0.0551
Roy's Greatest Root	0.56187584	3.65	2	13	0.0551

Marginally significant. The effect of exercise intensity in pulse rates is different for each diet type. Investigate later.

## Effect of diet and exercise type on pulse rate (averaged over intensity-of-exercise):

The GLM Procedure  
Repeated Measures Analysis of Variance  
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value
exercise	2	1524.77778	762.38889	1.03
diet	1	7800.01852	7800.01852	10.51
Error	14	10390.70370	742.19312	

Source	Pr > F
exercise	0.3835
diet	0.0059
Error	

- Exercise type has no effect, but diet *does* have an effect on pulse rates overall.
- Diet by intensity interaction significant, so effect of diet on pulse rate differs by intensity as well.
- Look back at PROC MEANS output (over):



exercise	diet	NObs	Variable	N	Mean	Std Dev
racquetb	meat	3	pulse1	3	95.6666667	10.0664459
			pulse2	3	143.6666667	30.5505046
			pulse3	3	204.0000000	32.0468407
	veg	3	pulse1	3	82.6666667	8.9628864
			pulse2	3	130.6666667	17.6729549
			pulse3	3	182.6666667	12.6622799
stairs	meat	3	pulse1	3	104.0000000	13.0000000
			pulse2	3	154.6666667	19.6299092
			pulse3	3	209.6666667	18.5831465
	veg	3	pulse1	3	90.0000000	4.5825757
			pulse2	3	120.6666667	12.5830574
			pulse3	3	171.0000000	18.7349940
weights	meat	3	pulse1	3	85.6666667	4.0414519
			pulse2	3	141.3333333	13.5769412
			pulse3	3	203.0000000	22.0680765
	veg	3	pulse1	3	73.6666667	14.0118997
			pulse2	3	113.6666667	16.8027775
			pulse3	3	160.3333333	27.6827263

- Meat pulse rates a little higher than corresponding veg pulse rates eg:

weights	meat	3	pulse1	3	85.6666667	4.0414519
			pulse2	3	141.3333333	13.5769412
			pulse3	3	203.0000000	22.0680765
	veg	3	pulse1	3	73.6666667	14.0118997
			pulse2	3	113.6666667	16.8027775
pulse3			3	160.3333333	27.6827263	

- significant diet main effect
- Difference in pulse rates between meat and veg *increases* as intensity increases: for weights, means differ by 12, 28, 43 for three intensities. This pattern consistent across exercise types.
  - significant diet by intensity interaction.

Looking at means is one way to assess how groups differ; discriminant analysis is another.

First make a variable with all combos of exercise and diet:

```
data ex;  
  infile "exercise.dat";  
  input exercise $ pulse1 pulse2 pulse3 diet $;  
  gp=cat(diet,"-",exercise);  
  one=pulse1/pulse1;
```

(variable "one" explained later)

Then see what it is about pulse1 pulse2 pulse3 that explains differences among groups:

```
proc discrim can out=fred;  
  class gp;  
  var pulse1--pulse3;
```

Output (edited):

# Generalized Squared Distance to gp

		meat	meat	meat	veg	veg
From gp		-racquetb	-stairs	-weights	-racquetb	-stairs
meat		0	1.16623	3.41757	3.41347	2.93105
-racquetb						
meat	-stairs	1.16623	0	7.38962	6.45912	3.17538
meat		3.41757	7.38962	0	1.23812	9.55256
-weights						
veg		3.41347	6.45912	1.23812	0	6.12220
-racquetb						
veg	-stairs	2.93105	3.17538	9.55256	6.12220	0
veg		7.33211	11.27939	3.86302	1.06346	7.79109
-weights						

		veg
From gp		-weights
meat		7.33211
-racquetb		
meat	-stairs	11.27939
meat		3.86302
-weights		
veg		1.06346
-racquetb		
veg	-stairs	7.79109
veg		0
-weights		

Where generalized squared distance is large, groups are very different (in pulse rate profile) – eg. meat-stairs and veg-weights. Most similar groups: meat-racquetball & stairs; veg-racquetball & weights; veg-racquetball & meat-weights (!)

Canonical variables are combinations of original variables that explain differences among groups. How many do we need?

Test of H0: The canonical correlations in the current row and all that follow are zero

	Likelihood Ratio	Approximate F Value	Num DF	Den DF	Pr > F
1	0.14954639	1.85	15	28.007	0.0778
2	0.47530543	1.24	8	22	0.3236
3	0.89044735	0.49	3	12	0.6943

Only 1 near significance.

Raw Canonical Coefficients

Variable	Can1	Can2	Can3
pulse1	0.1965980326	-.0565263460	0.0251750782
pulse2	-.0834670872	0.0092799901	-.1315596602
pulse3	0.0109568453	0.0533423277	0.0875031123

Can1 mostly pulse1 with a little (negative) pulse2 (pulse3 near 0). *But* pulse2 and pulse3 more variable than pulse1, so might be more influential than they look:

Pooled Within-Class Standardized Canonical Coefficients

Variable	Can1	Can2	Can3
pulse1	1.940143588	-0.557834818	0.248442296
pulse2	-1.618486870	0.179945685	-2.551036460
pulse3	0.250534283	1.219701610	2.000806707

Pulse2 more influential than appears.

## How do groups differ on canonical variables?

Class Means on Canonical Variables

gp		Can1	Can2	Can3
meat	-racquetb	0.759973767	0.519620228	0.281658046
meat	-stairs	1.542241535	0.452920427	-0.459854929
meat	-weights	-1.022206867	1.009888050	0.249376692
veg	-racquetb	-0.944474555	-0.004146803	-0.202075451
veg	-stairs	1.204085362	-1.133800397	0.277268748
veg	-weights	-1.539619242	-0.844481506	-0.146373106

Meat values generally more positive than veg ones on Can1 (comparing like exercise). Pattern for exercise probably affected by variability in data.

Next part of discriminant analysis: see how well data were classified into correct groups.

# Number of Observations and Percent Classified into gp

From gp	meat -racquetb	meat -stairs	meat -weights	veg -racquetb	veg -stairs	veg -weights	Total
meat - racquetb	1 33.33	0 0.00	1 33.33	0 0.00	1 33.33	0 0.00	3 100.00
meat -stairs	1 33.33	2 66.67	0 0.00	0 0.00	0 0.00	0 0.00	3 100.00
meat -weights	0 0.00	0 0.00	2 66.67	1 33.33	0 0.00	0 0.00	3 100.00
veg - racquetb	0 0.00	0 0.00	0 0.00	3 100.00	0 0.00	0 0.00	3 100.00
veg -stairs	0 0.00	0 0.00	0 0.00	0 0.00	3 100.00	0 0.00	3 100.00
veg -weights	0 0.00	0 0.00	0 0.00	2 66.67	0 0.00	1 33.33	3 100.00
Total	2 11.11	2 11.11	3 16.67	6 33.33	4 22.22	1 5.56	18 100.00

11.11

## Error Count Estimates for gp

	meat -racquetb	meat -stairs	meat -weights	veg -racquetb	veg -stairs	veg -weights
Rate	0.6667	0.3333	0.3333	0.0000	0.0000	0.6667
Priors	0.1667	0.1667	0.1667	0.1667	0.1667	0.1667

Total  
Rate  
Priors  
0.3333

Overall 33% of 18 observations misclassified – meat-racquetball and veg-weights hardest to get right.

Was classification clear-cut or a close call? Look at output data set (2 parts here):

Obs		gp	pulse1	pulse2	pulse3	Can1		_INTO_
1	meat	-stairs	112	166	215	2.22750	meat	-stairs
2	meat	-stairs	111	166	225	2.14047	meat	-stairs
3	meat	-stairs	89	132	189	0.25875	meat	-racquetb
4	veg	-stairs	95	134	186	1.23853	veg	-stairs
5	veg	-stairs	86	109	150	1.16138	veg	-stairs
6	veg	-stairs	89	119	177	1.21234	veg	-stairs
7	meat	-racquetb	105	177	241	0.21806	meat	-weights
8	meat	-racquetb	85	117	186	0.69149	meat	-racquetb
9	meat	-racquetb	97	137	185	1.37037	veg	-stairs
10	veg	-racquetb	93	151	197	-0.45308	veg	-racquetb
11	veg	-racquetb	77	122	178	-1.38628	veg	-racquetb
12	veg	-racquetb	78	119	173	-0.99407	veg	-racquetb
13	meat	-weights	81	134	205	-1.30566	meat	-weights
14	meat	-weights	88	133	180	-0.11993	veg	-racquetb
15	meat	-weights	88	157	224	-1.64104	meat	-weights
16	veg	-weights	58	99	131	-3.71687	veg	-weights
17	veg	-weights	85	132	186	-0.56051	veg	-racquetb
18	veg	-weights	78	110	164	-0.34147	veg	-racquetb

Some groups very consistent on Can1 (eg. veg-stairs), some not.



Now look at “posterior probabilities” of group membership given pulse rates:

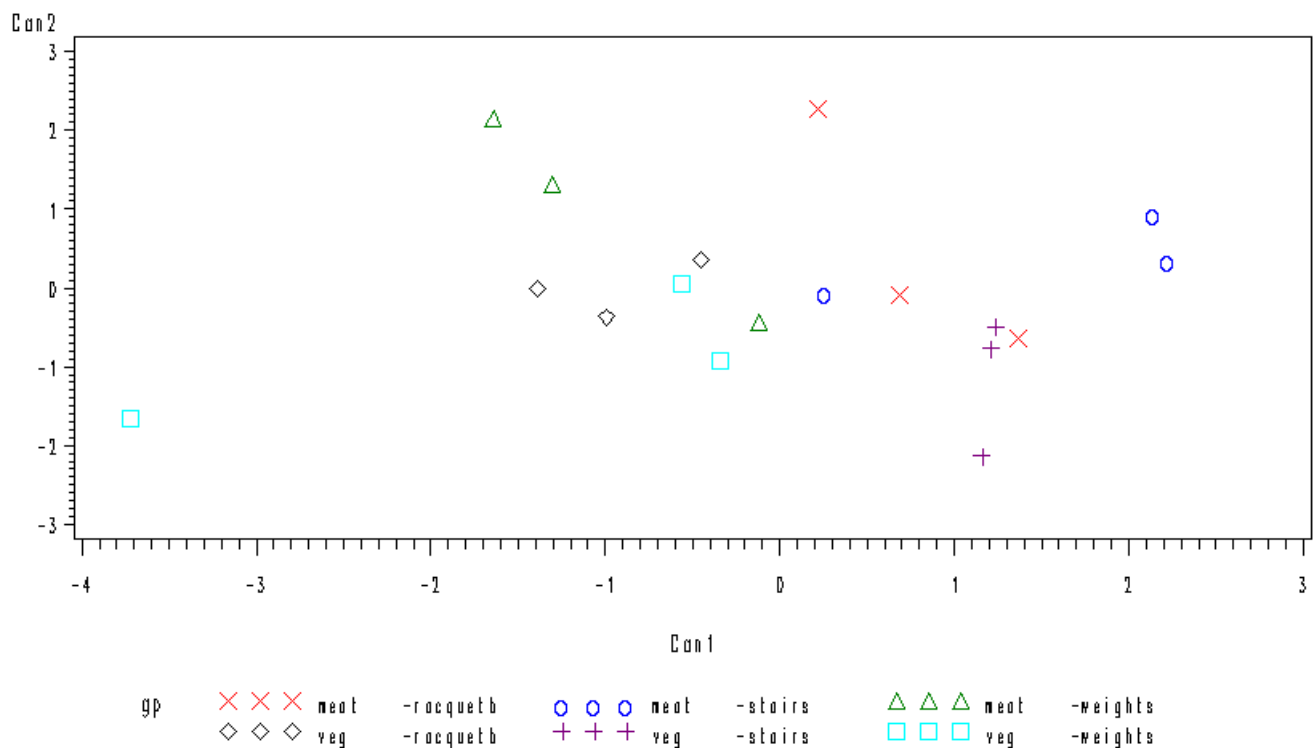
Obs	Group	Pulse Rate	meat		meat		veg		veg	
			stairs	racquetball	stairs	racquetball	stairs	racquetball	stairs	racquetball
1	meat	-stairs	0.13770	0.77947	0.00180	0.00467	0.07608	0.00028		
2	meat	-stairs	0.25337	0.68596	0.00503	0.00485	0.05059	0.00020		
3	meat	-stairs	0.34811	0.13063	0.11847	0.19098	0.15494	0.05687		
4	veg	-stairs	0.25831	0.28436	0.01226	0.04113	0.39391	0.01003		
5	veg	-stairs	0.03924	0.04843	0.00096	0.01759	0.87601	0.01778		
6	veg	-stairs	0.24437	0.12467	0.01025	0.02856	0.58268	0.00948		
7	meat	-racquetb	0.32225	0.20603	0.37332	0.09132	0.00322	0.00386		
8	meat	-racquetb	0.48368	0.07753	0.06972	0.06051	0.29235	0.01620		
9	meat	-racquetb	0.20082	0.34074	0.00714	0.03384	0.40909	0.00837		
10	veg	-racquetb	0.12599	0.09311	0.19375	0.43418	0.02208	0.13090		
11	veg	-racquetb	0.04244	0.00423	0.27034	0.37524	0.00911	0.29863		
12	veg	-racquetb	0.06944	0.00977	0.18520	0.38139	0.03172	0.32249		
13	meat	-weights	0.06891	0.00348	0.70775	0.17292	0.00171	0.04523		
14	meat	-weights	0.16749	0.09678	0.09419	0.34370	0.12551	0.17232		
15	meat	-weights	0.02769	0.00245	0.80732	0.14137	0.00015	0.02102		
16	veg	-weights	0.00003	0.00000	0.00630	0.07856	0.00004	0.91506		
17	veg	-weights	0.14777	0.03807	0.22633	0.37720	0.04105	0.16958		
18	veg	-weights	0.13810	0.02505	0.08622	0.27658	0.21368	0.26036		

Obs 1 clearcut meat-stairs, obs 2 less so, obs 3 misclassified as meat-racquetball. Obs 4 correct but very close call. Obs 14 totally wrong!

Plotting 1<sup>st</sup> 2 canonical variables gives nice 2-dimensional picture of data & how well groups are separated:

```
goptions reset=all;
symbol1 c=red v=x;
symbol2 c=blue v=o;
symbol3 c=green v=triangle;
symbol4 c=black v=diamond;
symbol5 c=purple v=plus;
symbol6 c=cyan v=square;
```

```
proc gplot;
  plot can2*can1=gp;
```



Groups rather mixed up. Can2 sometimes helps (racquetball) but not significant.

Actually exercise type not significant, so really only need discriminant analysis by diet:

```
proc discrim data=ex can out=ginger;
  class diet;
  var pulse1--pulse3;

proc print;
  var diet pulse1 pulse2 pulse3 can1 meat veg _into_;

proc gplot;
  plot one*can1=diet;
```

output:

Test of H0: The canonical correlations in the  
current row and all that follow are zero

	Likelihood Ratio	Approximate F Value	Num DF	Den DF	Pr > F
1	0.56370159	3.61	3	14	0.0404

Only one canonical variable (2 groups), which is significant. (With 2 groups, same test as MANOVA.)

What does canonical variable depend on?

Raw Canonical Coefficients

Variable	Can1
pulse1	0.0199203412
pulse2	-.0305445631
pulse3	0.0631215151

Mainly pulse3 now. (Different.)

## How are groups distinguished on Can1?

### Class Means on Canonical Variables

diet	Can1
meat	0.8294513579
veg	-.8294513579

Meat-eaters +, veg-eaters -.

How well can observations be classified back into groups?

### Number of Observations and Percent Classified into diet

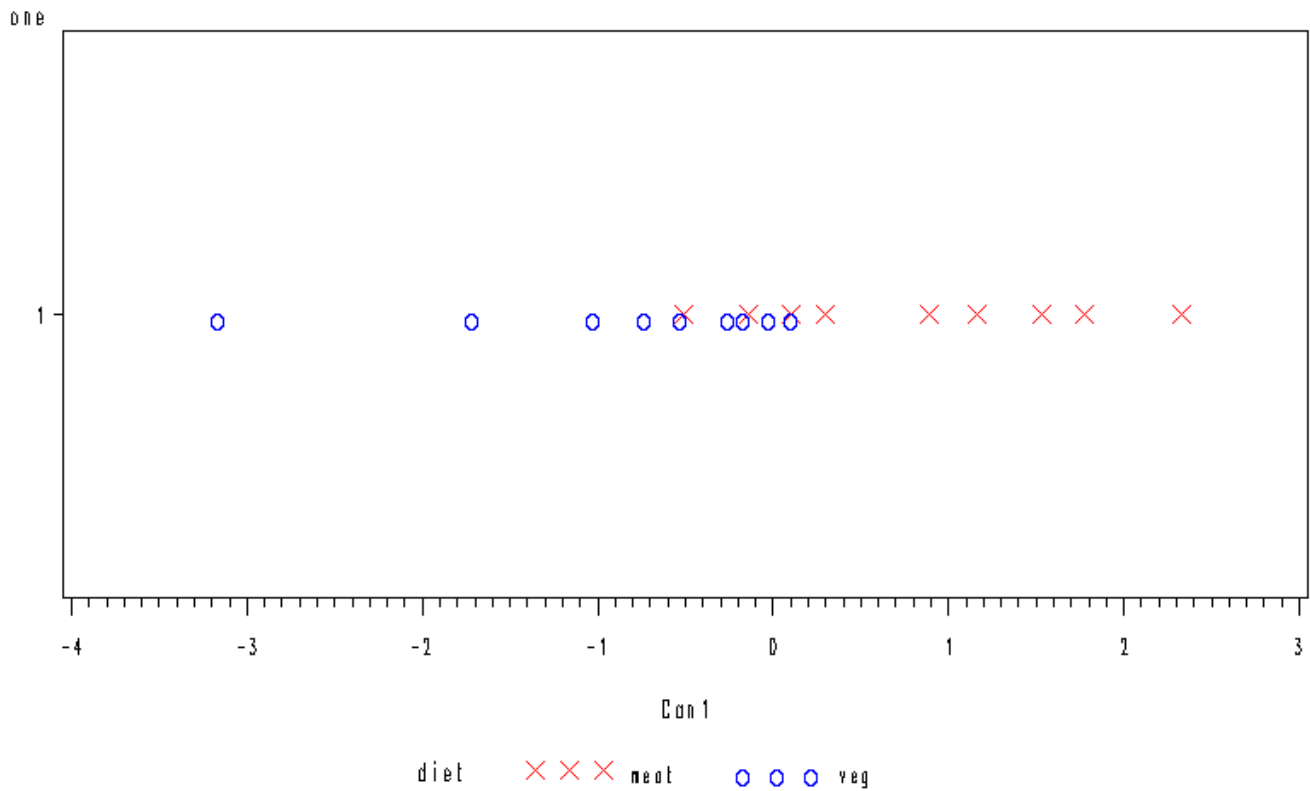
From diet	meat	veg	Total
meat	7 77.78	2 22.22	9 100.00
veg	1 11.11	8 88.89	9 100.00
Total	8 44.44	10 55.56	18 100.00
Priors	0.5	0.5	

### Error Count Estimates for diet

	meat	veg	Total
Rate	0.2222	0.1111	0.1667
Priors	0.5000	0.5000	

About as well as before, though simpler to see.

On a graph:



Now see purpose of "one". Shows distinction difficult in middle but easy at the ends.

## Output data set:

Obs	diet	pulse1	pulse2	pulse3	Can1	meat	veg	_INTO_
1	meat	112	166	215	1.16811	0.87411	0.12589	meat
2	meat	111	166	225	1.77940	0.95035	0.04965	meat
3	meat	89	132	189	0.10730	0.54438	0.45562	meat
4	veg	95	134	186	-0.02363	0.49020	0.50980	veg
5	veg	86	109	150	-1.71168	0.05523	0.94477	veg
6	veg	89	119	177	-0.25308	0.39656	0.60344	veg
7	meat	105	177	241	2.33384	0.97960	0.02040	meat
8	meat	85	117	186	0.29642	0.62051	0.37949	meat
9	meat	97	137	185	-0.13855	0.44279	0.55721	veg
10	veg	93	151	197	0.11160	0.54615	0.45385	meat
11	veg	77	122	178	-0.52064	0.29657	0.70343	veg
12	veg	78	119	173	-0.72469	0.23109	0.76891	veg
13	meat	81	134	205	0.89679	0.81573	0.18427	meat
14	meat	88	133	180	-0.51126	0.29982	0.70018	veg
15	meat	88	157	224	1.53302	0.92711	0.07289	meat
16	veg	58	99	131	-3.16331	0.00523	0.99477	veg
17	veg	85	132	186	-0.16175	0.43332	0.56668	veg
18	veg	78	110	164	-1.01788	0.15596	0.84404	veg

- Sometimes classification clear (obs 1, 2, 5);
- sometimes very unclear (obs 3, 4, 9)
- when pulse3 high, "meat",
- when low, "veg",
- when middling, hard to tell
- but look at obs 8 and 10