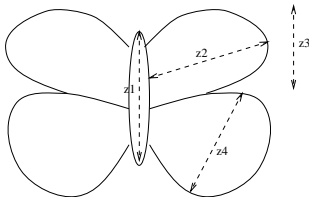


# Les papillons : les données

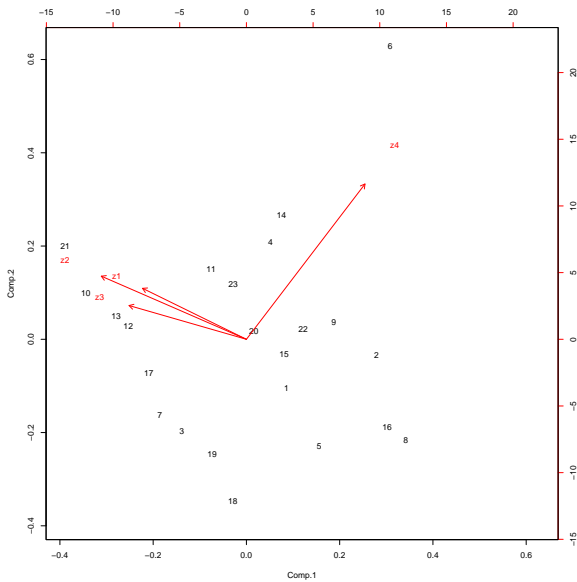
- Collecte de 23 papillons par les enfants d'une classe
- 4 mesures en mm
- Objectif : retrouver les différentes espèces de papillon uniquement avec ces 4 mesures
- Proposition : projeter sur le 1er plan de l'ACP



	$z_1$	$z_2$	$z_3$	$z_4$
1	22	35	24	19
2	24	31	21	22
3	27	36	25	15
4	27	36	24	23
5	21	33	23	18
6	26	35	23	32
7	27	37	26	15
8	22	30	19	20
9	25	33	22	22
10	30	41	28	17
11	24	39	27	21
12	29	39	27	17
13	29	40	27	17
14	28	36	23	24
15	22	36	24	20
16	23	30	20	20
17	28	38	26	16
18	25	34	23	14
19	26	35	24	15
20	23	37	25	20
21	31	42	29	18
22	26	34	22	21
23	24	38	26	21

# Les papillons : analyse en composantes principales

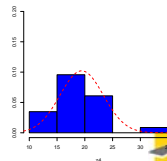
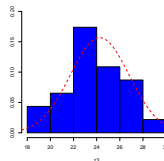
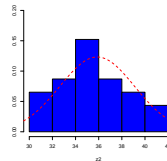
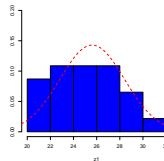
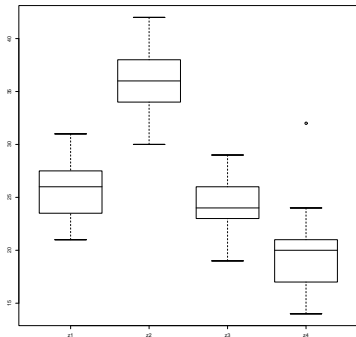
Numéro	1	2	3	4
Valeur propre	2.66	0.90	0.42	0.0147
Pourcentage de variance	66.52	22.53	10.58	0.3685
Pourcentage cumulé	66.52	89.05	99.63	100.0000



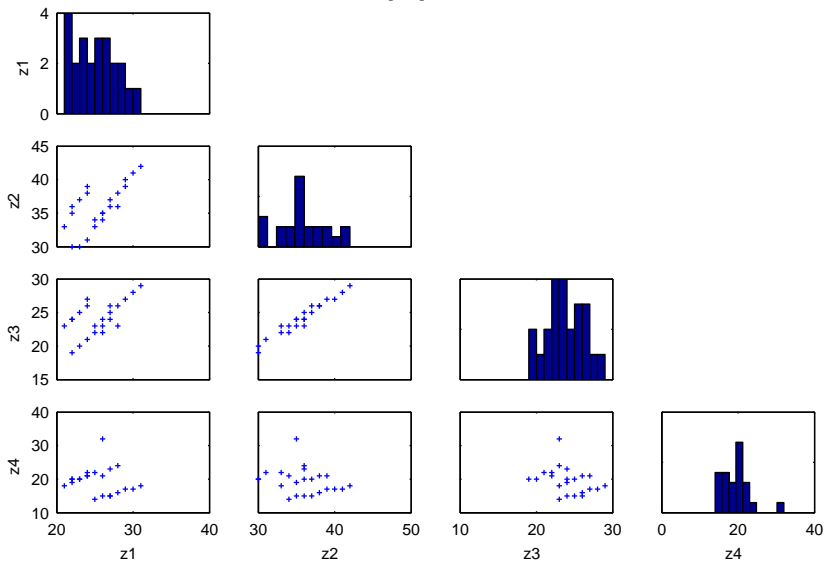
# Les papillons : description élémentaire

Var.	Mean	Var.	S - D	Min	Max
z1	26	7.8	2.8	21	31
z2	36	10.5	3.2	30	42
z3	24	6.5	2.5	19	29
z4	19	15.3	3.9	14	32

var	z1	z2	z3	z4
z1	1.00			
z2	0.69	1.00		
z3	0.64	0.98	1.00	-0.34
z4	-0.18	-0.23	-0.34	1.00

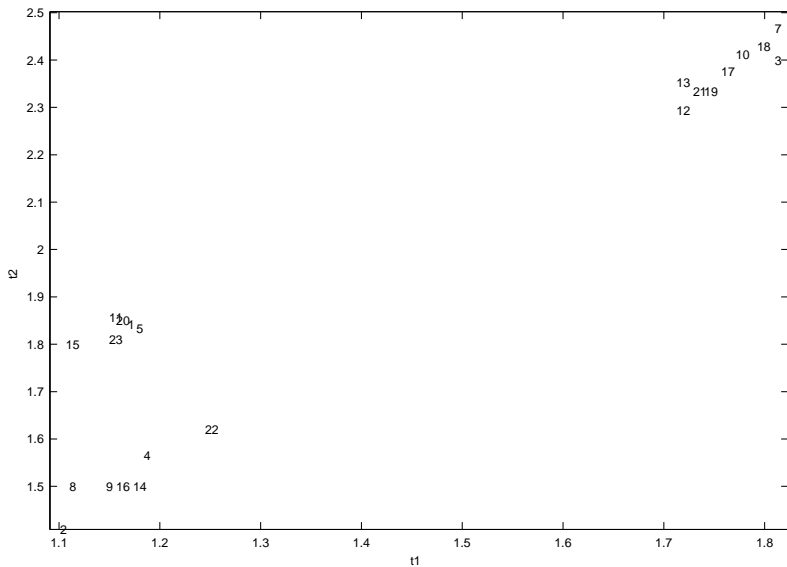


# Les papillons



- Papillon 6 atypique pour la variable  $z_4$  : on l'enlève
- Présence de l'effet taille (direction générale d'allongement) : on remplace les variables initiales par
  - $t_1 = z_1/z_4$
  - $t_2 = z_2/z_4$
  - $t_3 = z_3/z_4$
- $z_2$  et  $z_3$  variables très corrélées : on peut supprimer l'une des 2
- On peut alors visualiser les 2 variables restantes : graphique de dispersion

# Les papillons + creqan + creqan + creqan + delind



# relation Principal Component Analysis: Les papillons + creqan + creqan + creqan + delind (individ

