An abstract graphic consisting of several thin, black, overlapping lines that form various geometric shapes and polygons, primarily located in the upper-left and central portions of the page.

CLASE 6: BASES DE DATOS Y ANÁLISIS I

OBJETIVOS

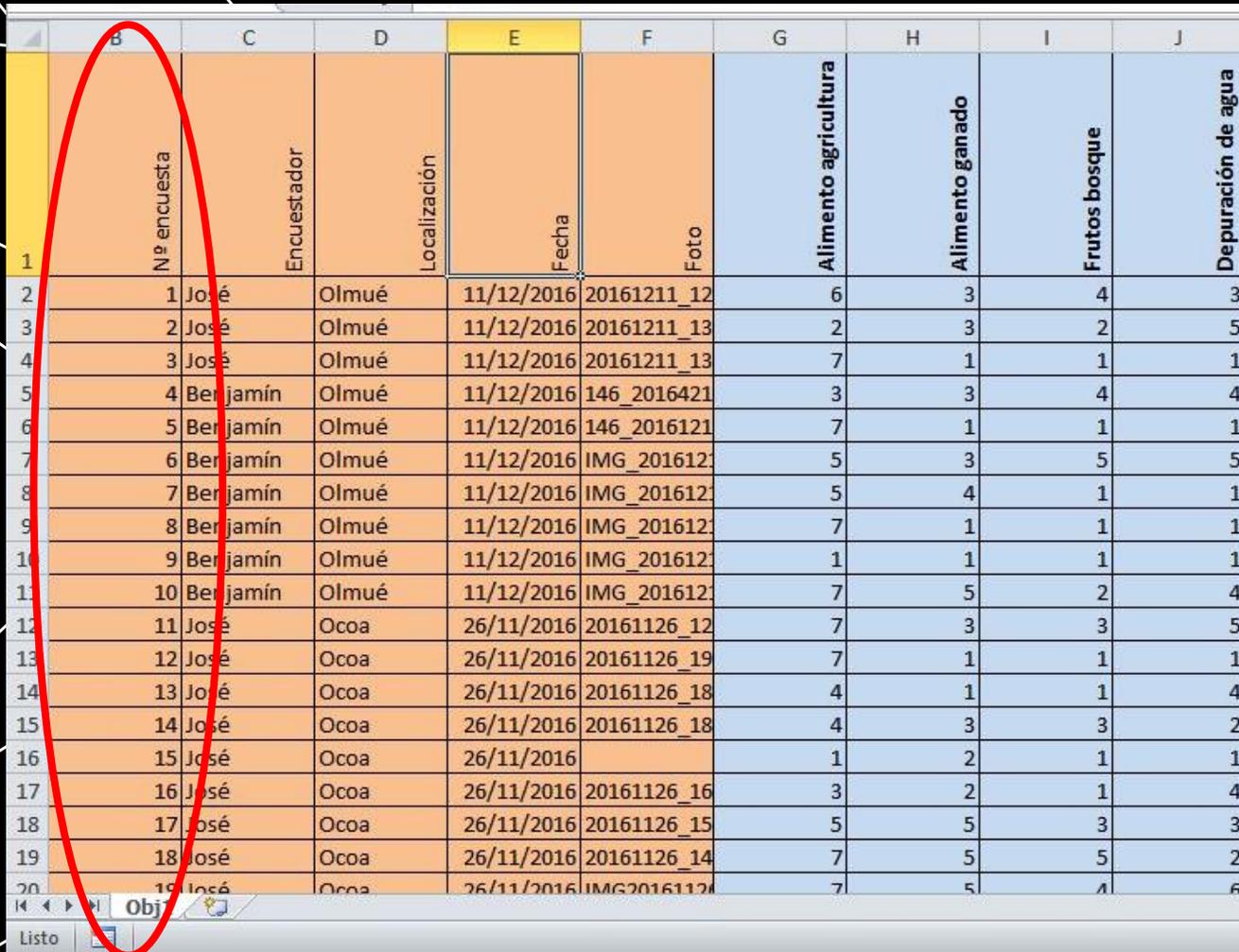
Ordenar los datos obtenidos y análisis descriptivos

CONSTRUCCIÓN DE BASES DE DATOS

Una parte esencial para la valoración de los servicios es poder llevar la encuesta y/o entrevista realizada a una planilla de datos que luego se pueda analizar estadísticamente. En general las variables se colocan en columnas y cada encuesta o entrevista como fila.

	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Nº encuesta	Encuestador	Localización	Fecha	Foto	Alimento agricultura	Alimento ganado	Frutos bosque	Depuración de agua	Control calidad aire	Belleza paisajística	Conocimiento local	Puntaje dominancia
2	1	José	Olmué	11/12/2016	20161211_12	6	3	4	3	4	4	3	3,857
3	2	José	Olmué	11/12/2016	20161211_13	2	3	2	5	1	1	2	2,286
4	3	José	Olmué	11/12/2016	20161211_13	7	1	1	1	7	1	1	2,714
5	4	Benjamín	Olmué	11/12/2016	146_2016421	3	3	4	4	4	1	1	2,857
6	5	Benjamín	Olmué	11/12/2016	146_2016121	7	1	1	1	7	1	4	3,143
7	6	Benjamín	Olmué	11/12/2016	IMG_201612	5	3	5	5	4	2	2	3,714
8	7	Benjamín	Olmué	11/12/2016	IMG_201612	5	4	1	1	5	2	1	2,714
9	8	Benjamín	Olmué	11/12/2016	IMG_201612	7	1	1	1	1	1	1	1,857
10	9	Benjamín	Olmué	11/12/2016	IMG_201612	1	1	1	1	1	1	1	1,000
11	10	Benjamín	Olmué	11/12/2016	IMG_201612	7	5	2	4	7	2	3	4,286
12	11	José	Ocoa	26/11/2016	20161126_12	7	3	3	5	7	5	6	5,143
13	12	José	Ocoa	26/11/2016	20161126_19	7	1	1	1	6	2	2	2,857
14	13	José	Ocoa	26/11/2016	20161126_18	4	1	1	4	7	1	1	2,714
15	14	José	Ocoa	26/11/2016	20161126_18	4	3	3	2	4	5	3	3,429
16	15	José	Ocoa	26/11/2016		1	2	1	1	3	4	2	2,000
17	16	José	Ocoa	26/11/2016	20161126_16	3	2	1	4	4	1	1	2,286
18	17	José	Ocoa	26/11/2016	20161126_15	5	5	3	3	5	3	3	3,857
19	18	José	Ocoa	26/11/2016	20161126_14	7	5	5	2	6	6	3	4,857
20	19	José	Ocoa	26/11/2016	IMG20161126	7	5	4	6	6	5	4	5,286

SELECCIÓN ÚNICA/MÚLTIPLE



	B	C	D	E	F	G	H	I	J
1	Nº encuesta	Encuestador	Localización	Fecha	Foto	Alimento agricultura	Alimento ganado	Frutos bosque	Depuración de agua
2	1	José	Olmué	11/12/2016	20161211_12	6	3	4	3
3	2	José	Olmué	11/12/2016	20161211_13	2	3	2	5
4	3	José	Olmué	11/12/2016	20161211_13	7	1	1	1
5	4	Berjamín	Olmué	11/12/2016	146_2016421	3	3	4	4
6	5	Berjamín	Olmué	11/12/2016	146_2016121	7	1	1	1
7	6	Berjamín	Olmué	11/12/2016	IMG_201612	5	3	5	5
8	7	Berjamín	Olmué	11/12/2016	IMG_201612	5	4	1	1
9	8	Berjamín	Olmué	11/12/2016	IMG_201612	7	1	1	1
10	9	Berjamín	Olmué	11/12/2016	IMG_201612	1	1	1	1
11	10	Berjamín	Olmué	11/12/2016	IMG_201612	7	5	2	4
12	11	José	Ocoa	26/11/2016	20161126_12	7	3	3	5
13	12	José	Ocoa	26/11/2016	20161126_19	7	1	1	1
14	13	José	Ocoa	26/11/2016	20161126_18	4	1	1	4
15	14	José	Ocoa	26/11/2016	20161126_18	4	3	3	2
16	15	José	Ocoa	26/11/2016		1	2	1	1
17	16	José	Ocoa	26/11/2016	20161126_16	3	2	1	4
18	17	José	Ocoa	26/11/2016	20161126_15	5	5	3	3
19	18	José	Ocoa	26/11/2016	20161126_14	7	5	5	2
20	19	José	Ocoa	26/11/2016	IMG2016112	7	5	4	6

»Primero es conveniente colocar el localizador o número de encuesta para evitar confusiones. Para ello, es necesario un trabajo previo de etiquetado de las encuestas.

RESULTADOS CUALITATIVOS

		Proteger humedal	Control Acceso	Educacion tenencia
1	¿Qué medidas cree que son las más adecuadas para disminuir el efecto de la libre circulación de animales domésticos?			
2	PROTEGER MAS EL HUMEDAL, CONTROLAR ACCESO A HUMEDAL		1	1
3	EDUCACION A PERSONAS EN TENENCIA RESPONSABLE			1
4	CORRAL MUNICIPAL PARA ANIMALES CALLEJEROS			
5	QUE PERSONAS BOTEN ANIMALES ???			
6	SEGURIDAD EN EL HUMEDAL, PERSONAL ENCARGADO DE CUIDAR EL LUGAR			
7	CONTROL EN EL RECINTO DEL HUMEDAL			1
8	MATAR PERROS Y GATOS SIN DUEÑO PARA DISMINUIR POBLACION			
9	TENENCIA RESPONSABLE			
10	CONTROL DE POBLACION DE PERROS VAGOS, TENENCIA RESPONSABLE			
11	TENENCIA RESPONSABLE (DUEÑOS HACERSE CARGO), CAMPAÑAS MUNICIPALES PARA ESTERILIZACION			
12	TENDRIAN QUE MANDAR A UN CAMION PARA QUE RECOJAN A LOS PERROS DE LA CALLE			
13	TENERLOS ENCERRADOS EN SU CASA			
14	NA			
15	HACERSE RESPONSABLE DE CADA ANIMAL, QUE NO BOTEN PERROS			
16	NA			
17	NA			
18	CONTROLAR ACCESO A ANIMALES			1
19	ESTERILIZACION GRATUITA, CONTROL A PERSONAS QUE CUIDAN ANIMALES			
20	PROHIBICION DE INGRESO AL HUMEDAL			1

ESCALAS DE PUNTUACIÓN

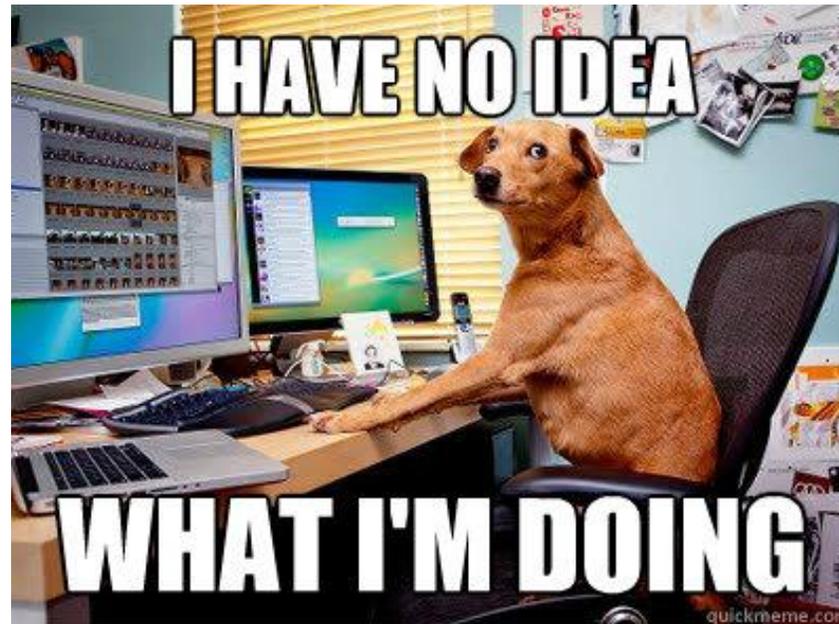
Belleza escénica	Subsistencia	Familiaridad	Valor cultural/patrimonial	Valor ecológico	Sentido de pertenencia
0	1	0	0	1	0
0	0	1	0	0	1
1	0	0	0	1	0
0	0	1	0	1	0
1	1	0	0	0	0
0	1	0	0	1	1
1	1	0	0	0	0
1	0	1	0	0	0
1	0	0	0	1	0
0	0	0	1	1	0
0	1	0	0	1	0
0	1	0	0	1	0
1	0	1	0	0	0

Variables dicotómicas:

Aquellas variables que sean de elección, por ejemplo, preguntas de si o no, pueden categorizarse como 1 o 0. Esto ayuda para calcular estadísticas básicas.

ANÁLISIS ESTADÍSTICOS

- El análisis estadístico es la parte más compleja del proceso debido a la gran cantidad de metodologías y test estadísticos existentes.
- Veremos las formas de análisis más comunes y útiles y los criterios para elegir un método u otro.
- Disponer de unos objetivos claros y una pregunta que queramos responder sea clara.



ANÁLISIS ESTADÍSTICOS

Depuración de la base de datos: Eliminar observaciones sin datos, corregir errores, etc.

Estadísticos descriptivos, cálculo de índices.

Análisis estadísticos

ANÁLISIS DESCRIPTIVOS

- Una primera mirada a los datos
- No desestimar, muchas veces las principales conclusiones vienen de esta fase.
- Para datos cuantitativos (continuos)
 - Media (datos cuantitativos con distribución normal), Mediana (más recomendable la mayor parte de las veces).
 - Varianza, error estándar.
 - Coeficiente de variación (permite comparar variación entre distintas variables)
 - Rango intercuartil
 - Coeficientes de confianza
 - Distribuciones
- Para datos cualitativos (o discretos)
 - Moda
 - Frecuencias
 - Frecuencias relativas

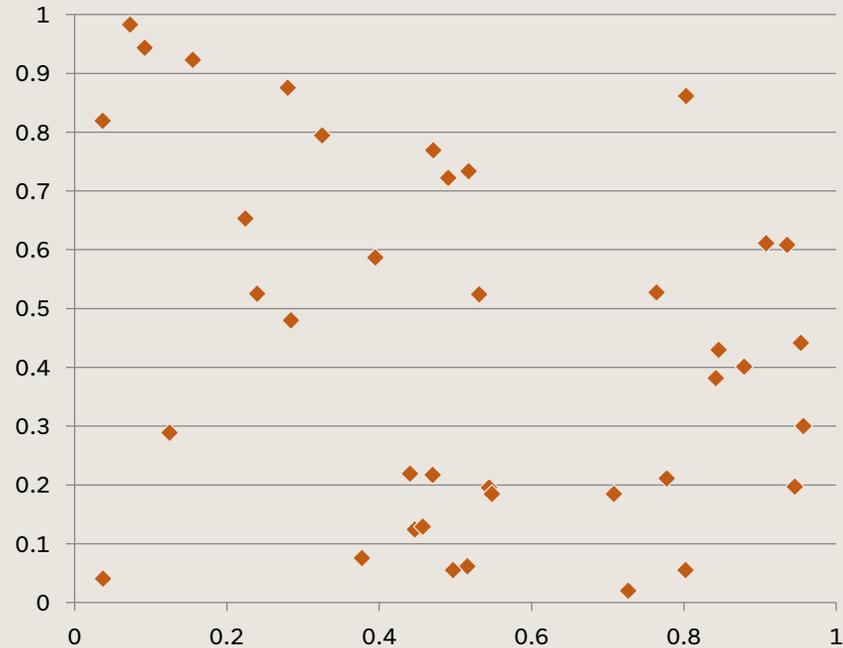
ANÁLISIS ESTADÍSTICOS

1. Tengo una **serie de variables**, por ejemplo, **puntuaciones** dados por cada encuestado a una lista de servicios y quiero relacionarlas con variables socio-culturales o socio-económicas. Mi estudio busca ser **exploratorio**, quiero una imagen comprensible de las respuestas.
2. Tengo una **serie de variables**, por ejemplo, **puntuaciones** dados por cada encuestado a una lista de servicios y quiero relacionarlas con variables socio-culturales o socio-económicas. Tengo una **hipótesis concreta** en la que quiero relacionar la valoración de varios servicios ecosistémicos **con una sola variable cualitativa**, por ejemplo, ver si al aumentar el nivel de estudio se valoran más los servicios culturales.
3. Tengo una **serie de variables**, por ejemplo, **puntuaciones** dados por cada encuestado a una lista de servicios y quiero relacionarlas con variables socio-culturales o socio-económicas. Tengo una **hipótesis concreta** en la que quiero relacionar la valoración de varios servicios ecosistémicos **con una sola variable cuantitativa**, por ejemplo, ver si al aumentar número de años de residencia en la comuna cambia la valoración de los servicios de provisión.

ANÁLISIS ESTADÍSTICOS

4. Tengo una serie de variables, por ejemplo, **puntuaciones** dados por cada encuestado a una lista de servicios y quiero relacionarlas **con variables cuantitativas socio-culturales o socio-económicas**. Tengo una **hipótesis concreta** en la que quiero relacionar la valoración de servicios ecosistémicos con varias variables de distintas características, por ejemplo, ver si al aumentar el nivel de estudio se valoran más los servicios culturales.
5. Tengo una serie de variables **dicotómicas**. La mayor parte de mi base de datos son **ceros o variables de texto y dicotómicas** (si/no, 1/0)
6. **Quiero reducir variables o crear un modelo simple**. La estructura de la encuesta se basa en una serie de preguntas agrupadas por la variable que miden. Por ejemplo, una batería de preguntas que buscan la relación de las personas con el entorno, otra batería con preguntas para averiguar el grado de conocimiento de la naturaleza del lugar.
7. **Quiero agrupar variables o valores de una variable**. Es común tener muchos valores de profesiones. Es conveniente agrupar estos valores en base a la distinta valoración de servicios.

1. DATOS DE PUNTUACIONES, ESTUDIO EXPLORATORIO



Análisis de Componentes Principales (ACP o PCA)

Análisis multivariante de ordenación. Busca relaciones entre variables respuesta.

Ejemplo: que servicios tienen puntuaciones similares

Da una imagen comprensible de los datos

Permite formular hipótesis, no responderlas

1. DATOS DE PUNTUACIONES, ESTUDIO EXPLORATORIO

Análisis de Componentes Principales (ACP o PCA)

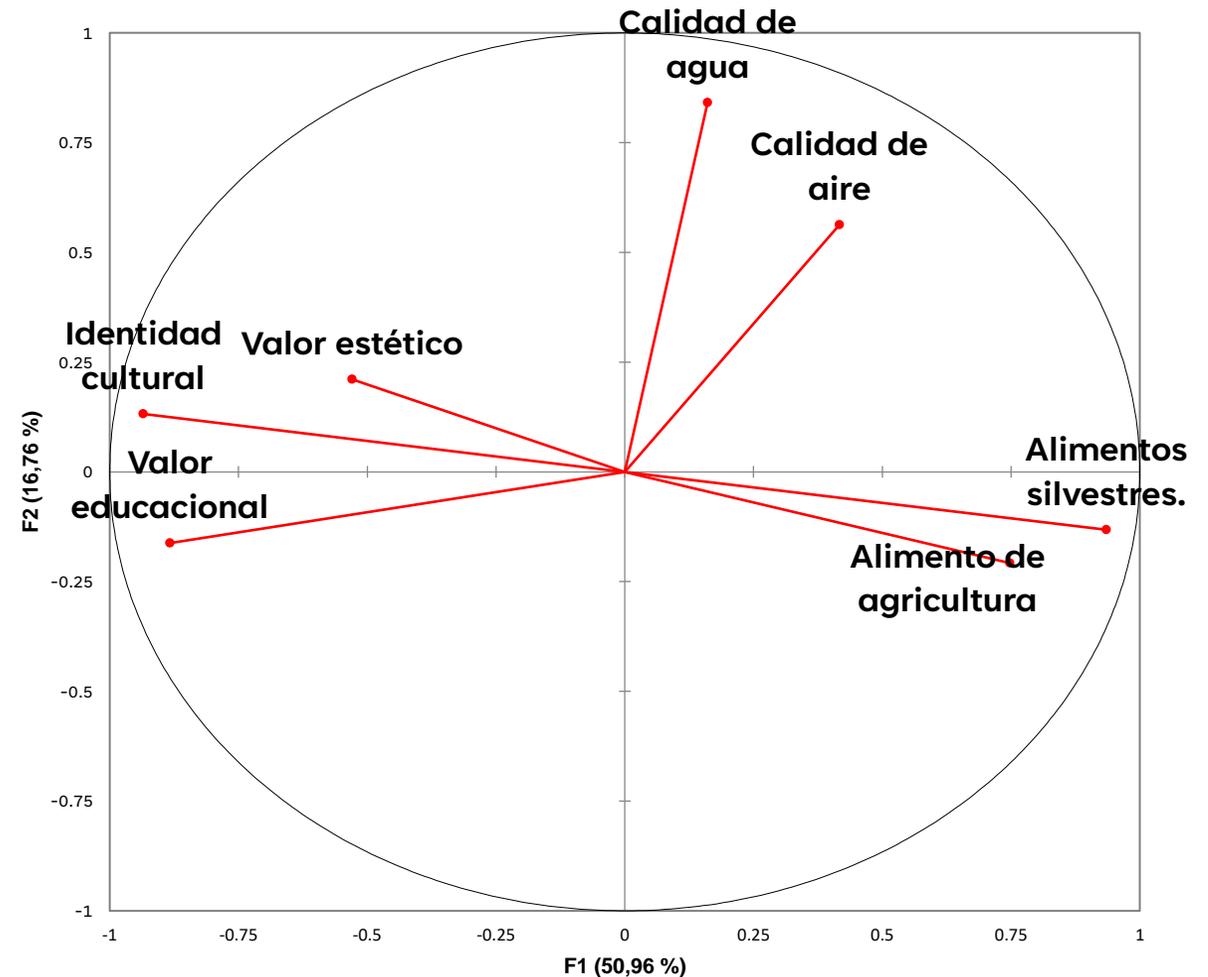
- Análisis de Componentes Principales (ACP o PCA) revisamos:
 - Los vectores propios y la variabilidad recogida por cada eje
 - Nos indican la cantidad de variabilidad de los datos recogida por los ejes.
 - Cuanto mayor es el vector propio mayor cantidad de variabilidad explica.
 - Es habitual tener la duda de cuántos ejes debemos considerar para otros análisis o para representar en un gráfico. Se deben tomar los ejes con valor propio mayor que 1 (criterio de Kaiser).

	F1	F2	F3	F4	F5	F6
Eigenvalue	3,567	1,173	0,835	0,776	0,444	0,204
Variability (%)	50,964	16,756	11,932	11,091	6,342	2,914
Cumulative %	50,964	67,720	79,652	90,744	97,086	100,000

1. DATOS DE PUNTUACIONES, ESTUDIO EXPLORATORIO

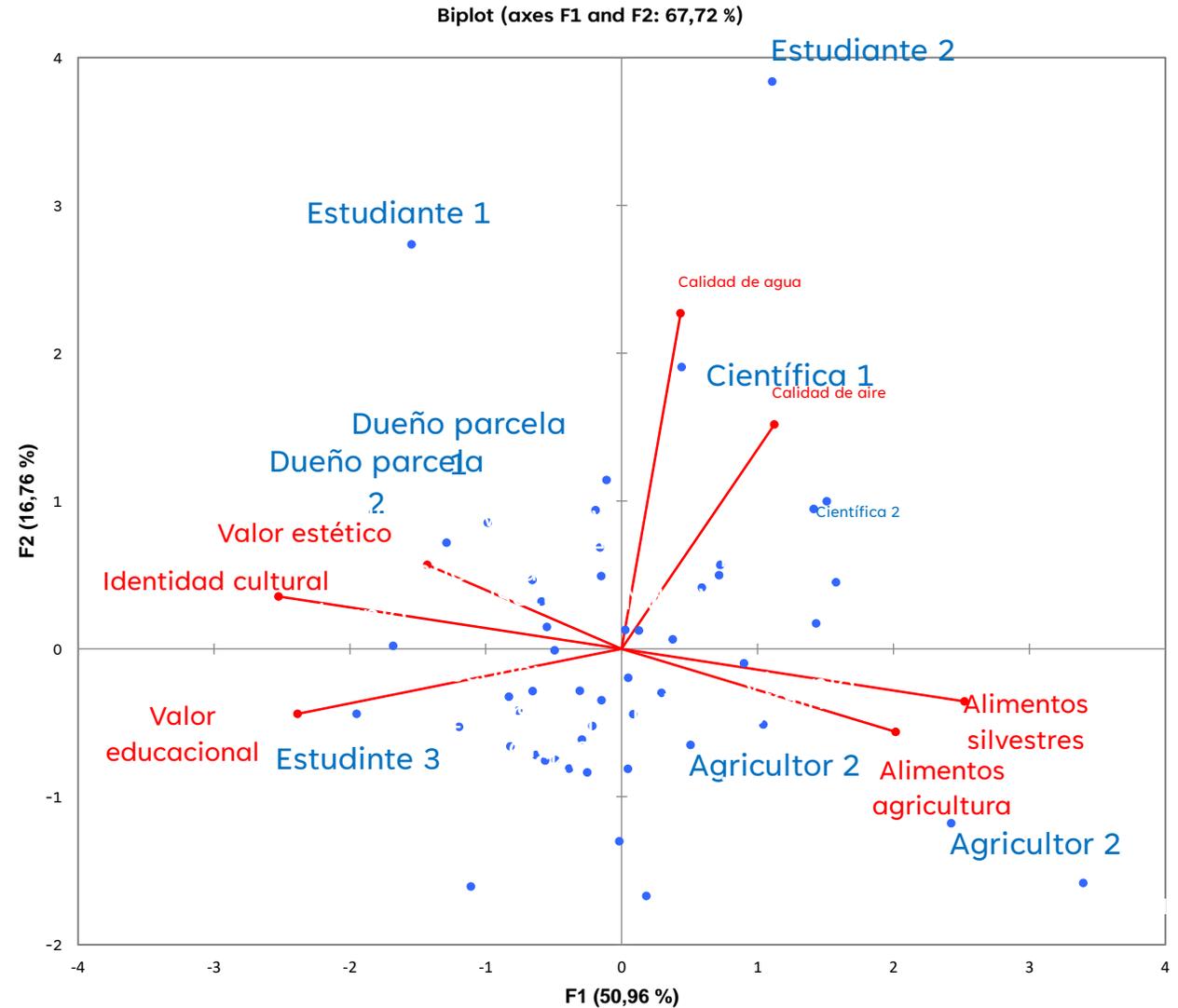
	F1	F2
Calidad del agua	0,161	0,841
Valor estético	-0,529	0,211
Calidad del aire	0,417	0,563
Alimentos de agricultura	0,748	-0,208
Valor educacional	-0,883	-0,162
Alimentos silvestres	0,935	-0,132
Identidad cultural	-0,935	0,132

Variables (axes F1 and F2: 67,72 %)



1. DATOS DE PUNTUACIONES, ESTUDIO EXPLORATORIO

Es posible también obtener centroides de las observaciones. Obtendré un punto central de las observaciones con la misma característica. Por ejemplo entre los estudiantes



2. DATOS DE PUNTUACIONES, UNA VARIABLE CUALITATIVA, HIPÓTESIS PREVIA

ANOVA o Kruskal Wallis

Pest control	Environmental affine	A		1,96	12,4896	0,0059
	Olive Grove affine	A		2,01		
	Urban affine	A		2,07		
	Cattle affine	B		2,60		
Cultural identity	Olive Grove affine	A		2,54	16,8900	0,0007
	Urban affine	A		2,73		
	Environmental affine	B		2,32		
	Cattle affine	B		2,33		

- Permite ver diferencias significativas entre la puntuación de varios servicios basados en una variable cualitativa (por ejemplo: pertenencia a asociación medioambiental). $P < 0,05 \rightarrow$ diferencias significativas
- Interesante acompañarlo de estadísticos descriptivos (media, mediana) o gráficos (histograma de frecuencias o boxplot).
- Kruskal Wallis es apropiado para escalas de Likert, no requiere normalidad en los datos.
- Utilizar comparaciones múltiples a pares

3. DATOS DE PUNTUACIONES, UNA VARIABLE CUANTITATIVA

Correlación o regresión

- Si me interesa solamente relacionar dos variables y observar si al aumentar una, crece otra, decrece o se mantiene →
Correlación
- Me interesa crear una formula que prediga como al aumentar una variable, aumenta la otra →
Regresiones, modelos lineales
 - Requiere una hipótesis

