

Análisis científico y métodos estadísticos

Erika Kanía K.
Ing. Agr. Dr.



1

Profesores

✓ **Erika Kanía K. Ing. Agr. Dr.**
 Directora académica. Diplomado Análisis Estadístico para Estudios Agropecuarios
 Profesor Asistente. Facultad de Ciencias Agronómicas, Universidad de Chile.
 Profesora Responsable "Análisis científico y métodos estadísticos" (Obligatorio / Pregrado)
 Profesora Responsable "Taller estrategias de análisis de datos multivariados" (Electivo / Pregrado)
 Profesora Colaboradora "Estadística Aplicada a los Recursos Naturales Renovables II" (Electivo / Postgrado)
 Doctorado en Cs. Silvopropagandarias, Universidad de Chile.



2

Unidad	Contenidos
Estadística descriptiva y análisis exploratorio de datos	<ul style="list-style-type: none"> - Población y muestra - Tipos de variables <ul style="list-style-type: none"> - Cualitativas: nominales u ordinales - Cuantitativas: discretas o continuas - Categorización de variables cuantitativas - Manejo de datos - Recursos gráficos <ul style="list-style-type: none"> - Tablas (Tablas de frecuencia, Tablas de contingencia) - Gráficos (Gráfico sectores, Gráfico barra, Gráfico barra apilada, Gráfico de puntos, Gráfico dispersión, Histograma, Gráfico de cajas, etc. - Estadística descriptiva <ul style="list-style-type: none"> - Medidas de posición: Moda, mediana, media, cuartiles y percentiles - Medidas de dispersión: Varianza, Desviación estándar, Coeficiente de Variación
Estadística inferencial	<ul style="list-style-type: none"> - Estimación de parámetros y contrastes de hipótesis - Conceptos y definiciones: Nivel de significancia, Valor P

3

Unidad	Contenidos
Regresión y Correlación	<ul style="list-style-type: none"> - Análisis de regresión lineal simple - Análisis de correlación lineal
Introducción al Diseño de Experimentos	<ul style="list-style-type: none"> - Elementos del Diseño de experimentos. Conceptos y notaciones. Unidad experimental Unidad de observación; Variable respuesta; Tratamientos; Factores; Niveles; Estructura factorial de tratamientos; Repetición; Submuestreo; Error Experimental; Aleatorización; Bloques; - Análisis de la Varianza - Verificación de los supuestos del ANDEVA (Alternativas al no cumplirse los supuestos) - Pruebas de Comparaciones múltiples - Diseño Completamente Aleatorizado (DCA) - Diseño en Bloques Completos al azar (DBCA) - DCA o DBCA con estructura factorial de tratamientos.



4

¿Qué es la biometría?

Es la disciplina que se encarga de obtener información a partir de datos biológicos

¿Cómo?

Mediante gráficos, medidas numéricas de resumen, estudios comparativos, estudios predictivos, etc.

5

¿ Por qué están ustedes acá?

6

¿ Por qué están ustedes aca?

Para aprender a transformar datos en información

↓

Desarrollar habilidades en el análisis de datos experimentales

7

¿ Por qué están ustedes aca?

Para aprender a transformar datos en información

↓

Desarrollar habilidades en el análisis de datos experimentales

Ser buenos consumidores de la información que recibimos

↓

Saber interpretar lo que escuchamos o leemos y poder decidir criteriosamente sobre la confiabilidad de la información

8



9

Pedir ayuda a un estadístico una vez que el experimento ya fue realizado puede no ser más que pedirle efectuar una autopsia: posiblemente, lo único que pueda hacer sea decir de qué murió el experimento

Ronald Fisher

10

Caso	Zona	N-E	P-E	K-E	Ch-E	B-E	Na-E	Mn-E	D-E	N-B	P-B	K-B	Ch-B	B-B	Na-B	Mn-B
1	Norte	6.50	0.90	4.00	1.50	15.00	0.50	26.00	9.56	36.10	5.10	21.90	3.10	90.00	1.40	144.00
2	Norte	6.90	0.90	5.10	1.10	14.00	0.40	41.00	8.89	33.50	5.60	22.30	1.80	69.00	1.30	174.00
3	Norte	9.00	2.10	4.60	1.10	14.00	0.50	19.00	6.76	34.30	5.30	21.20	0.80	49.00	0.60	70.00
4	Norte	8.50	1.30	4.20	0.90	10.00	0.10	33.00	8.70	34.20	4.70	22.40	0.50	86.00	0.60	158.00
5	Norte	7.80	1.40	5.40	1.20	13.00	0.40	55.00	7.42	32.30	4.30	22.20	1.10	89.00	1.40	188.00
6	Norte	7.90	1.30	5.00	1.00	12.00	0.20	59.00	8.17	32.90	4.90	22.40	1.10	68.00	0.50	235.00
7	Norte	7.60	1.20	5.10	0.80	10.00	0.10	25.00	7.41	32.40	4.40	21.10	1.30	36.00	0.40	137.00
8	Norte	7.40	1.00	5.90	0.90	11.00	0.10	26.00	7.89	33.30	4.40	21.40	1.60	43.00	0.40	154.00
9	Norte	7.40	1.10	5.90	0.70	15.00	0.20	1.00	8.40	39.00	5.60	25.80	1.70	75.00	0.50	210.00
10	Centro	8.10	1.30	4.90	0.70	11.00	0.20	40.00	11.17	37.50	5.30	19.40	1.00	29.00	0.10	158.00
11	Centro	6.70	0.90	5.10	0.50	10.00	0.20	50.00	11.00	39.60	6.00	23.10	0.90	30.00	0.10	285.00
12	Centro	8.10	1.40	6.10	0.60	11.00	0.30	29.00	10.92	31.10	3.80	17.40	1.30	30.00	0.10	123.00
13	Centro	7.30	1.00	5.50	0.60	11.00	0.20	24.00	11.67	37.80	4.80	19.80	1.40	36.00	0.10	147.00
14	Centro	7.20	1.20	6.30	0.50	10.00	0.20	28.00	15.50	43.00	7.30	26.00	1.10	22.00	0.10	182.00
15	Centro	7.30	1.20	6.10	0.90	9.00	0.20	26.00	15.00	33.60	6.90	26.30	1.00	22.00	0.10	184.00
16	Centro	8.00	1.30	5.90	0.80	8.00	0.20	42.00	16.00	42.90	7.80	25.60	1.00	23.00	0.40	161.00
17	Centro	7.60	0.90	6.60	0.50	10.00	0.20	33.00	13.08	33.60	5.00	18.10	0.60	60.00	0.10	380.00
18	Centro	7.00	1.10	6.40	0.50	11.00	0.20	29.00	11.00	41.70	7.20	21.90	0.70	81.00	0.30	431.00

11

Caso	Zona	N-E	P-E	K-E	Ch-E	B-E	Na-E	Mn-E	D-E	N-B	P-B	K-B	Ch-B	B-B	Na-B	Mn-B
1	Norte	6.50	0.90	4.00	1.50	15.00	0.50	26.00	9.56	36.10	5.10	21.90	3.10	90.00	1.40	144.00
2	Norte	6.90	0.90	5.10	1.10	14.00	0.40	41.00	8.89	33.50	5.60	22.30	1.80	69.00	1.30	174.00
3	Norte	9.00	2.10	4.60	1.10	14.00	0.50	19.00	6.76	34.30	5.30	21.20	0.80	49.00	0.60	70.00
4	Norte	8.50	1.30	4.20	0.90	10.00	0.10	33.00	8.70	34.20	4.70	22.40	0.50	86.00	0.60	158.00
5	Norte	7.80	1.40	5.40	1.20	13.00	0.40	55.00	7.42	32.30	4.30	22.20	1.10	89.00	1.40	188.00
6	Norte	7.90	1.30	5.00	1.00	12.00	0.20	59.00	8.17	32.90	4.90	22.40	1.10	68.00	0.50	235.00
7	Norte	7.60	1.20	5.10	0.80	10.00	0.10	25.00	7.41	32.40	4.40	21.10	1.30	36.00	0.40	137.00
8	Norte	7.40	1.00	5.90	0.90	11.00	0.10	26.00	7.89	33.30	4.40	21.40	1.60	43.00	0.40	154.00
9	Norte	7.40	1.10	5.90	0.70	15.00	0.20	1.00	8.40	39.00	5.60	25.80	1.70	75.00	0.50	210.00
10	Centro	8.10	1.30	4.90	0.70	11.00	0.20	40.00	11.17	37.50	5.30	19.40	1.00	29.00	0.10	158.00
11	Centro	6.70	0.90	5.10	0.50	10.00	0.20	50.00	11.00	39.60	6.00	23.10	0.90	30.00	0.10	285.00
12	Centro	8.10	1.40	6.10	0.60	11.00	0.30	29.00	10.92	31.10	3.80	17.40	1.30	30.00	0.10	123.00
13	Centro	7.30	1.00	5.50	0.60	11.00	0.20	24.00	11.67	37.80	4.80	19.80	1.40	36.00	0.10	147.00
14	Centro	7.20	1.20	6.30	0.50	10.00	0.20	28.00	15.50	43.00	7.30	26.00	1.10	22.00	0.10	182.00
15	Centro	7.30	1.20	6.10	0.90	9.00	0.20	26.00	15.00	33.60	6.90	26.30	1.00	22.00	0.10	184.00
16	Centro	8.00	1.30	5.90	0.80	8.00	0.20	42.00	16.00	42.90	7.80	25.60	1.00	23.00	0.40	161.00
17	Centro	7.60	0.90	6.60	0.50	10.00	0.20	33.00	13.08	33.60	5.00	18.10	0.60	60.00	0.10	380.00
18	Centro	7.00	1.10	6.40	0.50	11.00	0.20	29.00	11.00	41.70	7.20	21.90	0.70	81.00	0.30	431.00

Generar una base de datos es muy costoso...

12

Caso	Zona	N-E	P-E	K-E	Cl-E	B-E	Na-E	Mn-E	D-E	N-B	P-B	K-B	Cl-B	B-B	Na-B	Mn-B
1	Norte	6.50	0.90	4.00	1.50	15.00	0.50	26.00	9.56	36.10	5.10	21.90	3.10	90.00	1.40	144.00
2	Norte	6.90	0.90	5.10	1.10	14.00	0.40	41.00	8.89	33.50	5.60	22.30	1.80	69.00	1.30	174.00
3	Norte	9.00	2.10	4.60	1.10	14.00	0.50	19.00	6.76	34.30	5.30	21.20	0.80	49.00	0.60	70.00
4	Norte	8.50	1.30	4.20	0.90	10.00	0.10	33.00	8.70	34.20	4.70	22.40	0.50	86.00	0.60	158.00
5	Norte	7.80	1.00	4.00	0.90	10.00	0.10	33.00	7.42	32.30	4.30	22.20	1.10	89.00	1.40	188.00
6	Norte	7.80	1.00	4.00	0.90	10.00	0.10	33.00	7.42	32.30	4.30	22.20	1.10	89.00	1.40	188.00
7	Norte	7.80	1.00	4.00	0.90	10.00	0.10	33.00	7.42	32.30	4.30	22.20	1.10	89.00	1.40	188.00
8	Norte	7.80	1.00	4.00	0.90	10.00	0.10	33.00	7.42	32.30	4.30	22.20	1.10	89.00	1.40	188.00
9	Norte	7.80	1.00	4.00	0.90	10.00	0.10	33.00	7.42	32.30	4.30	22.20	1.10	89.00	1.40	188.00
10	Centro	8.70	1.00	4.00	0.90	10.00	0.10	33.00	11.17	37.50	5.30	19.40	1.00	29.00	0.10	158.00
11	Centro	8.70	1.00	4.00	0.90	10.00	0.10	33.00	11.17	37.50	5.30	19.40	1.00	29.00	0.10	158.00
12	Centro	8.70	1.00	4.00	0.90	10.00	0.10	33.00	11.17	37.50	5.30	19.40	1.00	29.00	0.10	158.00
13	Centro	7.30	1.00	5.50	0.60	11.00	0.20	24.00	10.92	31.10	3.80	17.40	1.30	30.00	0.10	123.00
14	Centro	7.30	1.00	5.50	0.60	11.00	0.20	24.00	10.92	31.10	3.80	17.40	1.30	30.00	0.10	123.00
15	Centro	7.20	1.20	6.30	0.50	10.00	0.20	28.00	15.50	43.00	7.30	26.00	1.10	22.00	0.10	182.00
16	Centro	7.30	1.20	6.10	0.90	9.00	0.20	26.00	15.00	33.60	6.90	26.30	1.00	22.00	0.10	184.00
17	Centro	8.00	1.30	5.90	0.80	8.00	0.20	42.00	16.00	42.90	7.80	25.60	1.00	23.00	0.40	161.00
18	Centro	7.60	0.90	6.60	0.50	10.00	0.20	33.00	13.08	33.60	5.00	18.10	0.60	60.00	0.10	380.00
18	Centro	7.00	1.10	6.40	0.50	11.00	0.20	29.00	11.00	41.70	7.20	21.90	0.70	81.00	0.30	431.00

Generar una base de datos es muy costoso...

Datos ≠ Información

13

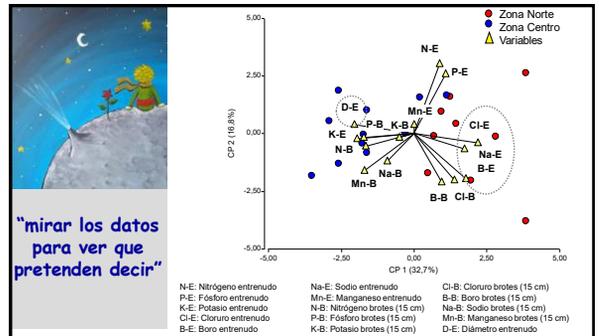
Ci-E	B-E	Na-E	Mn-E	D-E	N-B	P-B	K-B	Cl-B	B-B	Na-B	Mn-B
1.50	15.00	0.50	26.00	9.56	36.10	5.10	21.90	3.10	90.00	1.40	144.00
1.10	14.00	0.40	41.00	8.89	33.50	5.60	22.30	1.80	69.00	1.30	174.00
1.10	14.00	0.50	19.00	6.76	34.30	5.30	21.20	0.80	49.00	0.60	70.00
0.90	10.00	0.10	33.00	8.70	34.20	4.70	22.40	0.50	86.00	0.60	158.00
1.20	13.00	0.40	55.00	7.42	32.30	4.30	22.20	1.10	89.00	1.40	188.00
1.00	12.00	0.20	59.00	8.17	32.90	4.90	22.40	1.10	68.00	0.50	235.00
0.80	10.00	0.10	25.00	7.41	32.40	4.40	21.10	1.30	36.00	0.40	137.00
0.90	11.00	0.10	26.00	7.89	33.30	4.40	21.40	1.60	43.00	0.40	154.00
0.70	15.00	0.20	1.00	8.40	39.00	5.60	25.80	1.70	75.00	0.50	210.00
0.70	11.00	0.20	40.00	11.17	37.50	5.30	19.40	1.00	29.00	0.10	158.00
0.50	10.00	0.20	50.00	11.00	39.60	6.00	23.10	0.90	30.00	0.10	285.00
0.60	11.00	0.30	29.00	10.92	31.10	3.80	17.40	1.30	30.00	0.10	123.00
0.50	10.00	0.20	28.00	15.50	43.00	7.30	26.00	1.10	22.00	0.10	182.00
0.80	8.00	0.20	42.00	16.00	42.90	7.80	25.60	1.00	23.00	0.40	161.00
0.50	10.00	0.20	33.00	13.08	33.60	5.00	18.10	0.60	60.00	0.10	380.00
0.50	11.00	0.20	29.00	11.00	41.70	7.20	21.90	0.70	81.00	0.30	431.00

14

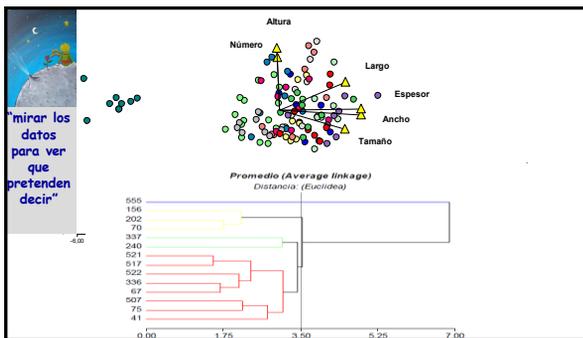
Ci-E	B-E	Na-E	Mn-E	D-E	N-B	P-B	K-B	Cl-B	B-B	Na-B	Mn-B
1.50	15.00	0.50	26.00	9.56	36.10	5.10	21.90	3.10	90.00	1.40	144.00
1.10	14.00	0.40	41.00	8.89	33.50	5.60	22.30	1.80	69.00	1.30	174.00
1.10	14.00	0.50	19.00	6.76	34.30	5.30	21.20	0.80	49.00	0.60	70.00
0.90	10.00	0.10	33.00	8.70	34.20	4.70	22.40	0.50	86.00	0.60	158.00
1.20	13.00	0.40	55.00	7.42	32.30	4.30	22.20	1.10	89.00	1.40	188.00
1.00	12.00	0.20	59.00	8.17	32.90	4.90	22.40	1.10	68.00	0.50	235.00
0.80	10.00	0.10	25.00	7.41	32.40	4.40	21.10	1.30	36.00	0.40	137.00
0.90	11.00	0.10	26.00	7.89	33.30	4.40	21.40	1.60	43.00	0.40	154.00
0.70	15.00	0.20	1.00	8.40	39.00	5.60	25.80	1.70	75.00	0.50	210.00
0.70	11.00	0.20	40.00	11.17	37.50	5.30	19.40	1.00	29.00	0.10	158.00
0.50	10.00	0.20	50.00	11.00	39.60	6.00	23.10	0.90	30.00	0.10	285.00
0.60	11.00	0.30	29.00	10.92	31.10	3.80	17.40	1.30	30.00	0.10	123.00
0.50	10.00	0.20	24.00	11.67	37.80	4.80	19.80	1.40	36.00	0.10	147.00
0.50	10.00	0.20	28.00	15.50	43.00	7.30	26.00	1.10	22.00	0.10	182.00
0.90	9.00	0.20	26.00	15.00	33.60	6.90	26.30	1.00	22.00	0.10	184.00
0.80	8.00	0.20	42.00	16.00	42.90	7.80	25.60	1.00	23.00	0.40	161.00
0.50	10.00	0.20	33.00	13.08	33.60	5.00	18.10	0.60	60.00	0.10	380.00
0.50	11.00	0.20	29.00	11.00	41.70	7.20	21.90	0.70	81.00	0.30	431.00

"mirar los datos para ver que pretenden decir"

15

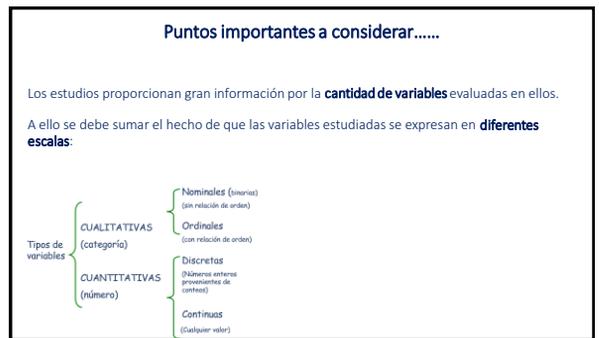


16



"mirar los datos para ver que pretenden decir"

17



18

Puntos importantes a considerar.....

Las estadísticas proporcionan gran información por la **cantidad de variables** evaluadas en ellos.

A ello se debe sumar el hecho de que las variables estudiadas se expresan en **diferentes escalas**:

Tipos de variables	{	CUALITATIVAS (categoría)	Nominales (Inerías) (sin relación de orden)
		ORDINALES (con relación de orden)	Ordinales (con relación de orden)
	{	CUANTITATIVAS (número)	Discretas (Números enteros provenientes de conteos)
		CONTINUAS (Cualquier valor)	Continuas (Cualquier valor)



19

Puntos importantes a considerar.....

Las estadísticas proporcionan gran información por la **cantidad de variables** evaluadas en ellos.

A ello se debe sumar el hecho de que las variables estudiadas se expresan en **diferentes escalas**:

Tipos de variables	{	CUALITATIVAS (categoría)	Nominales (Inerías) (sin relación de orden)
		ORDINALES (con relación de orden)	Ordinales (con relación de orden)
	{	CUANTITATIVAS (número)	Discretas (Números enteros provenientes de conteos)
		CONTINUAS (Cualquier valor)	Continuas (Cualquier valor)



ESTRATEGIAS DE ANÁLISIS DE DATOS

☐ Según la naturaleza de los datos y el objetivo de la investigación

20

- ✓ Estudios observacionales v/s Estudios experimentales
- ✓ Estadística descriptiva v/s Estadística Inferencial
- ✓ Estadística univariada v/s Estadística multivariada
- ✓ Tipos de variables
- ✓ Historia de los métodos estadísticos

21

Estudios observacionales v/s Estudios experimentales

- ✓ ESTUDIOS OBSERVACIONALES
Sobre un proceso existente se observan o miden (registra información) una o más variables.
- ✓ ESTUDIOS EXPERIMENTALES
En general, son estudios en que se realiza una selección de unidades experimentales, se aplican tratamientos y se estudian las respuestas algún tiempo después.

22

Estadística descriptiva v/s Estadística Inferencial

ESTADÍSTICA
 Es la ciencia que recoge, organiza, presenta, analiza e interpreta datos con el fin de facilitar la toma de decisiones.

23

Estadística descriptiva v/s Estadística Inferencial

ESTADÍSTICA
 Es la ciencia que recoge, organiza, presenta, analiza e interpreta datos con el fin de facilitar la toma de decisiones.

- ✓ **ESTADÍSTICA DESCRIPTIVA** - Consiste en un conjunto de procedimientos para organizar, resumir y presentar datos de manera informativa.

24

Estadística descriptiva v/s Estadística Inferencial

ESTADÍSTICA

Es la ciencia que recoge, organiza, presenta, analiza e interpreta datos con el fin de facilitar la toma de decisiones.

- ✓ **ESTADÍSTICA DESCRIPTIVA** - Consiste en un conjunto de procedimientos para organizar, resumir y presentar datos de manera informativa.
- ✓ **ESTADÍSTICA INFERENCIAL** - Corresponde a la obtención de conclusiones acerca de las características de una población a partir de una muestra de ésta.

25

Estadística univariada v/s Estadística multivariada

26

Estadística univariada v/s Estadística multivariada

ESTADÍSTICA UNIVARIADA

Observaciones (Caso)	Variable Peso fruto
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
...	
n	

INFORMACIÓN:

1 variable

↓

Métodos Estadísticos Univariados

27

Estadística univariada v/s Estadística multivariada

ESTADÍSTICA MULTIVARIADA

Observaciones (Caso)	Variables					
	x1	x2	x3	x4	...	p
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
...						
n						

INFORMACIÓN:

> 1 variable

↓

Métodos Estadísticos Multivariados

28

Estadística univariada v/s Estadística multivariada

ESTADÍSTICA MULTIVARIADA

Observaciones (Caso)	Variables					
	x1	x2	x3	x4	...	p
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
...						
n						

Métodos Estadísticos Multivariados

La matriz o tabla de datos:

a) **p** columnas, donde cada una de ellas corresponde a una variable

b) **n** filas, las cuales corresponden a cada unidad de observación.

La intersección de una fila y una columna da lugar al **dato**.

29

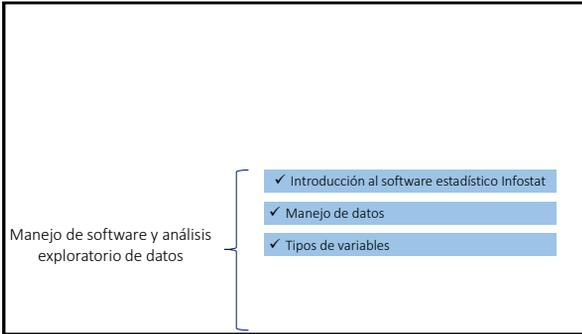
Estadística univariada v/s Estadística multivariada

ESTADÍSTICA MULTIVARIADA

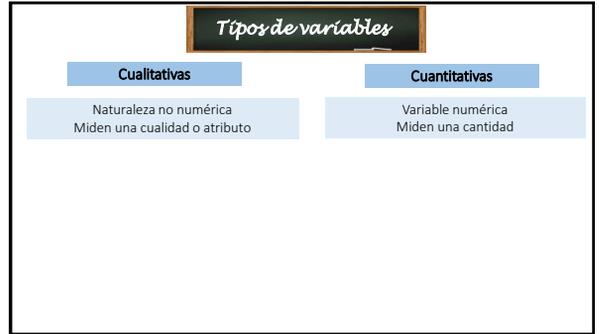
Legend: Zona Norte (red circle), Zona Centro (blue circle), Variables (yellow triangle)

Promedio (Average linkage) (Distancia (Euclídea))

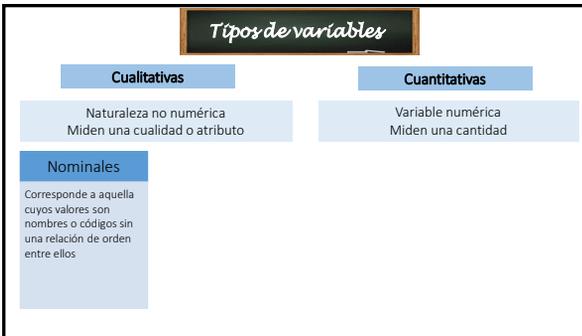
30



31



32



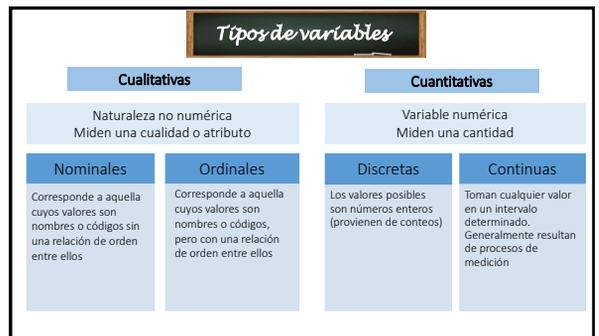
33



34



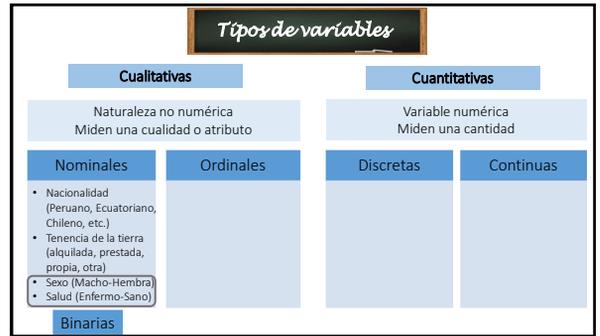
35



36



37



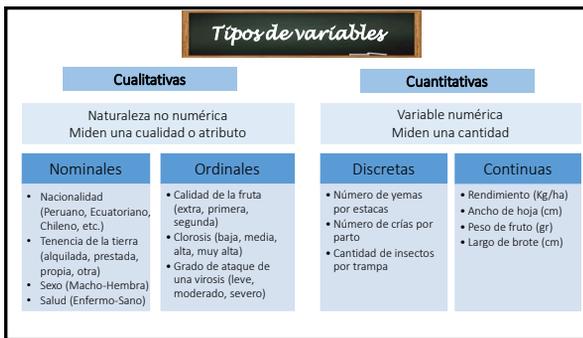
38



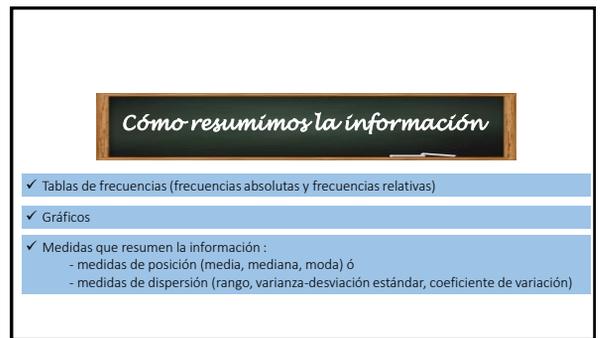
39



40



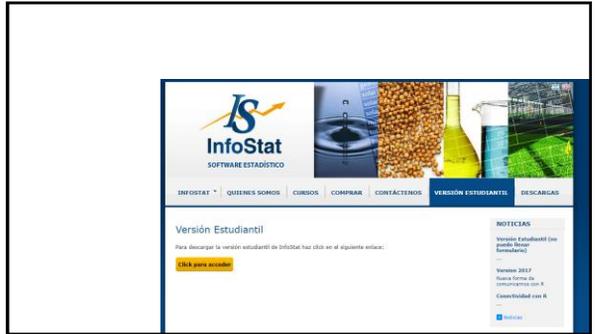
41



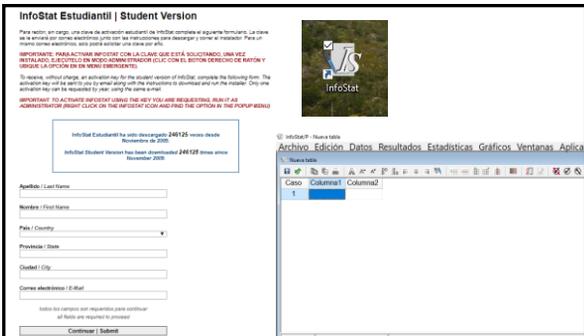
42



43



44



45



46



47