

**El Sistema que utilizo
con mis alumnos y
proyectos para
convertir datos en
resultados de impacto
en una sola tarde**



Cómo conseguir presentar tus resultados
con seguridad y confianza ahorrando
mucho tiempo y dolores de cabeza

Las claves para convertirte en un profesional imparable

- Los 3 ingredientes de la Ciencia de Datos
- Cómo aprender y dominar la estadística aunque no seas de números
- Un Sistema para tratar tus datos paso a paso con seguridad y confianza
- El software de datos que mejor se adapta para tu caso



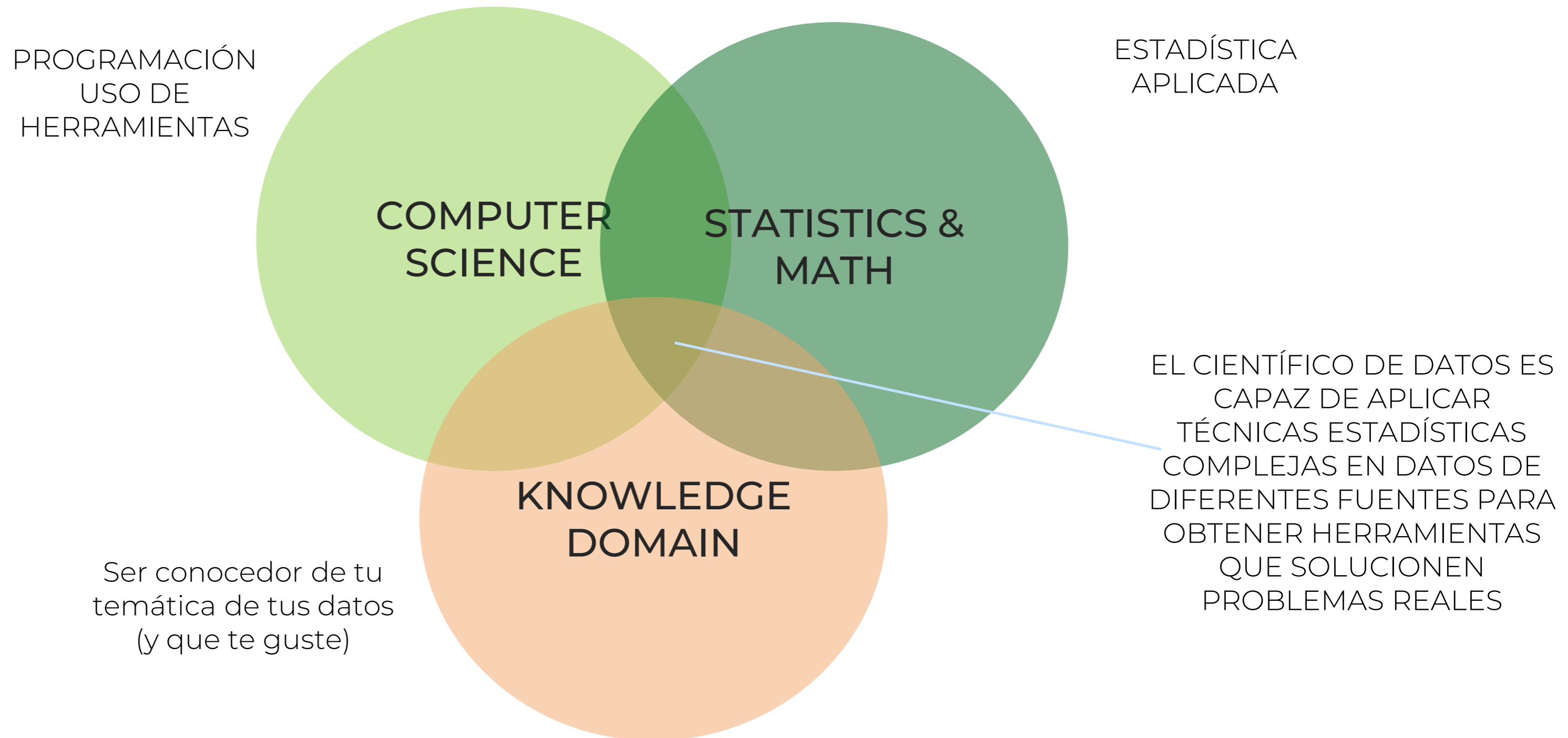
Los Tres Ingredientes de la Ciencia de Datos

... y las etapas de todo
proyecto con datos



Los tres ingredientes de la Ciencia de Datos

Qué conocimiento necesitas para ser científico de datos

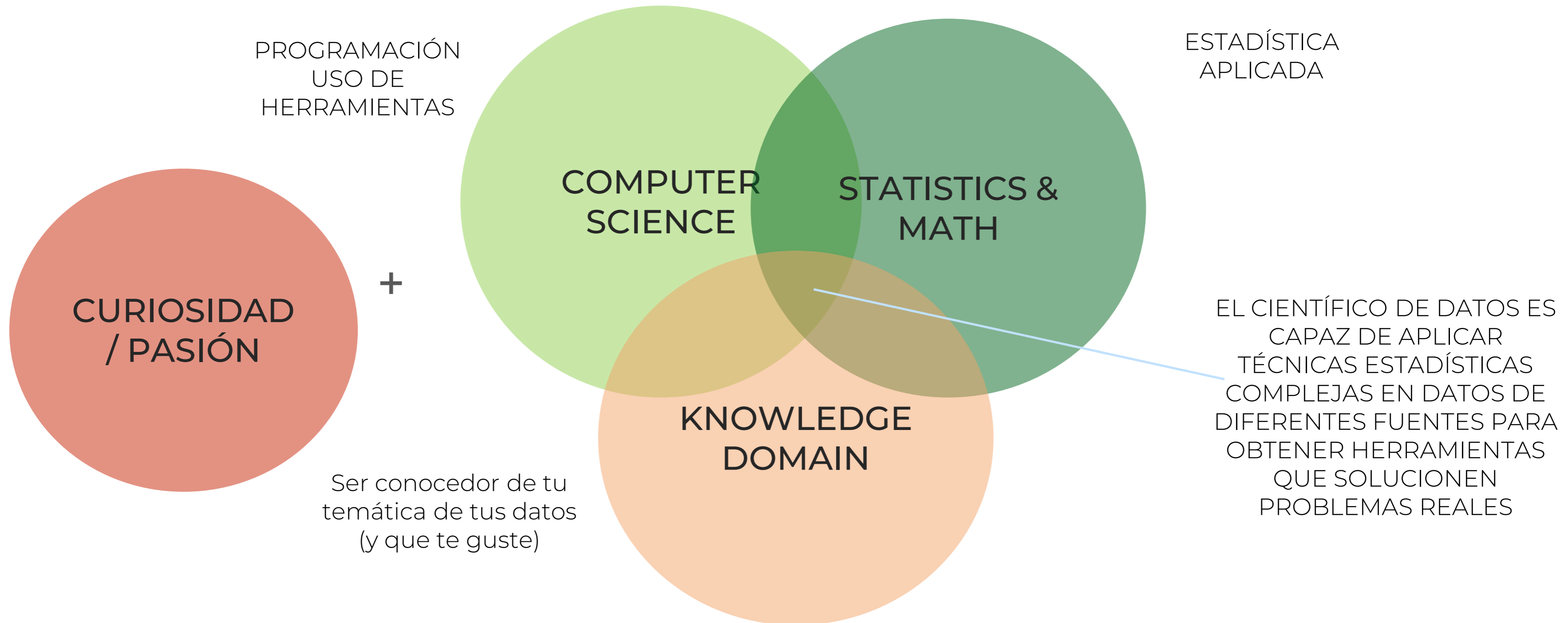


La curiosidad es
fundamental



Los tres ingredientes de la Ciencia de Datos

Qué conocimiento necesitas para ser científico de datos





El Análisis de Datos es un
oficio. Una habilidad

El papel de la Estadística

Observa la Realidad



Gráficos y Resultados



Resultados



Experto
Tiene una pregunta

MONTH	NO CANDIDATES	CAPACITY	% ASSISTANCE	NO CANDIDATES	CAPACITY	% ASSISTANCE	NO CANDIDATES	CAPACITY	% ASSISTANCE	WIFI	BENEFIT
1	20	28	71%	14	18	78%	8	12	66%	14.82	148.12
2	24	28	86%	16	18	89%	8	12	67%	156.82	156.82
3	18	28	64%	10	18	56%	8	12	67%	106.85	106.85
4	19	27	69%	11	18	61%	8	12	67%	107.82	107.82
5	20	26	77%	12	24	50%	8	12	67%	122.85	122.85
6	20	28	71%	9	18	50%	11	12	92%	105.75	111.75
7	25	28	89%	16	18	100%	4	12	33%	100.04	100.04
8	18	26	69%	8	18	50%	10	12	83%	104.14	104.14
9	21	28	75%	11	18	61%	10	12	83%	93.11	93.11
10	21	28	75%	16	24	67%	7	12	58%	81.65	81.65
11	18	22	82%	10	18	56%	8	12	67%	75.17	75.17
12	15	20	75%	9	18	50%	10	12	83%	75.65	75.65
13	17	24	71%	8	18	44%	8	12	67%	67.65	67.65
14	9	17	53%	5	18	28%	4	12	33%	61.65	61.65
15	18	24	75%	10	18	56%	8	12	67%	47.65	47.65
16	17	32	53%	11	24	46%	0	8	0%	54.15	54.15
17	9	17	53%	7	18	39%	7	12	58%	54.17	54.17
18	14	24	58%	11	18	61%	1	8	13%	51.62	51.62
19	11	28	39%	8	18	44%	8	12	67%	50.08	50.08
20	7	17	41%	7	18	39%	7	12	58%	45.87	45.87
21	7	18	39%	3	18	17%	4	8	50%	48.48	48.48
22	4	28	14%	8	24	33%	2	8	25%	46.22	46.22
23	6	17	35%	7	18	39%	7	12	58%	47.16	47.16
24	7	18	39%	1	18	6%	5	8	63%	31.39	31.39
25	4	4	100%	0	0	0%	4	8	50%	21.68	21.68
26	1	4	25%	0	0	0%	1	4	25%	53.52	53.52
27	5	17	29%	1	18	6%	7	12	58%	31.55	31.55
28	1	4	25%	0	0	0%	1	4	25%	40.78	40.78
29	5	12	42%	1	8	13%	1	8	13%	61.71	61.71
30	2	17	12%	1	18	6%	1	4	25%	48.76	48.76

Tabla de Datos

Estadística



Respuestas
Conclusiones

Cómo puedes aprender y
dominar la estadística
aunque no seas de números

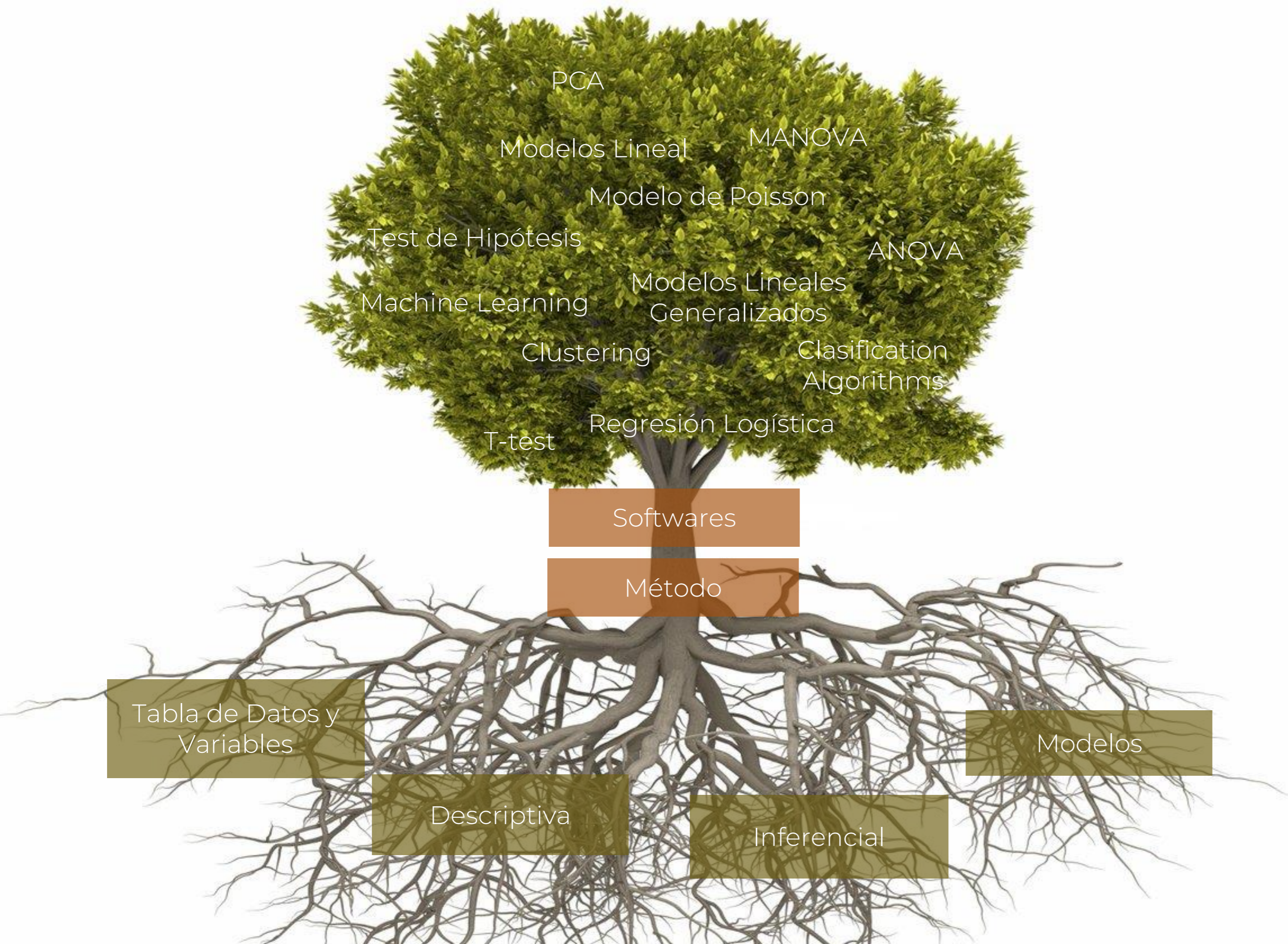


¿Qué técnica aplico en los distintos casos?



Sólo necesitas un Elemento ...





TÉCNICAS DE
ANÁLISIS DE
DATOS

SISTEMA

BASES
ESTADÍSTICAS



TÉCNICAS DE
ANÁLISIS DE
DATOS

SISTEMA

BASES
ESTADÍSTICAS

Fisher's Iris Data

Largo de sépalo ⇅	Ancho de sépalo ⇅	Largo de pétalo ⇅	Ancho de pétalo ⇅	Especies ⇅
5.0	2.0	3.5	1.0	<i>I. versicolor</i>
6.0	2.2	4.0	1.0	<i>I. versicolor</i>
6.0	2.2	5.0	1.5	<i>I. virginica</i>
6.2	2.2	4.5	1.5	<i>I. versicolor</i>
4.5	2.3	1.3	0.3	<i>I. setosa</i>
5.0	2.3	3.3	1.0	<i>I. versicolor</i>
5.5	2.3	4.0	1.3	<i>I. versicolor</i>
6.3	2.3	4.4	1.3	<i>I. versicolor</i>
4.9	2.4	3.3	1.0	<i>I. versicolor</i>
5.5	2.4	3.7	1.0	<i>I. versicolor</i>
5.5	2.4	3.8	1.1	<i>I. versicolor</i>
4.9	2.5	4.5	1.7	<i>I. virginica</i>
5.1	2.5	3.0	1.1	<i>I. versicolor</i>
5.5	2.5	4.0	1.3	<i>I. versicolor</i>
5.6	2.5	3.9	1.1	<i>I. versicolor</i>
5.7	2.5	5.0	2.0	<i>I. virginica</i>
6.3	2.5	4.9	1.5	<i>I. versicolor</i>
6.3	2.5	5.0	1.9	<i>I. virginica</i>
6.7	2.5	5.8	1.8	<i>I. virginica</i>
5.5	2.6	4.4	1.2	<i>I. versicolor</i>
5.7	2.6	3.5	1.0	<i>I. versicolor</i>
5.8	2.6	4.0	1.2	<i>I. versicolor</i>
6.1	2.6	5.6	1.4	<i>I. virginica</i>

Estructura

Header

Variables

Observaciones

Largo de sépalo *	Ancho de sépalo *	Largo de pétalo *	Ancho de pétalo *	Especies
2.0	3.5	1.0	1.0	I. versicolor
2.2	4.0	1.0	1.0	I. versicolor
2.2	5.0	1.5	1.5	I. virginica
2.2	4.5	1.5	1.5	I. versicolor
2.3	1.3	0.3	0.3	I. setosa
2.3	3.3	1.0	1.0	I. versicolor
2.3	4.0	1.3	1.3	I. versicolor
2.3	4.4	1.3	1.3	I. versicolor
2.4	3.3	1.0	1.0	I. versicolor
2.4	3.7	1.0	1.0	I. versicolor
2.4	3.8	1.1	1.1	I. versicolor
2.5	4.5	1.7	1.7	I. virginica
2.5	3.0	1.1	1.1	I. versicolor
2.5	4.0	1.3	1.3	I. versicolor
2.5	3.9	1.1	1.1	I. versicolor
2.5	5.0	2.0	2.0	I. virginica
2.5	4.9	1.5	1.5	I. versicolor
2.5	5.0	1.9	1.9	I. virginica
2.5	5.8	1.8	1.8	I. virginica
2.6	4.4	1.2	1.2	I. versicolor
2.6	3.5	1.0	1.0	I. versicolor
2.6	4.0	1.2	1.2	I. versicolor



BASES
ESTADÍSTICAS

Tabla de Datos y
Variables

Descriptiva

Inferencial

Modelos

Características
medibles de los
individuos de
una muestra

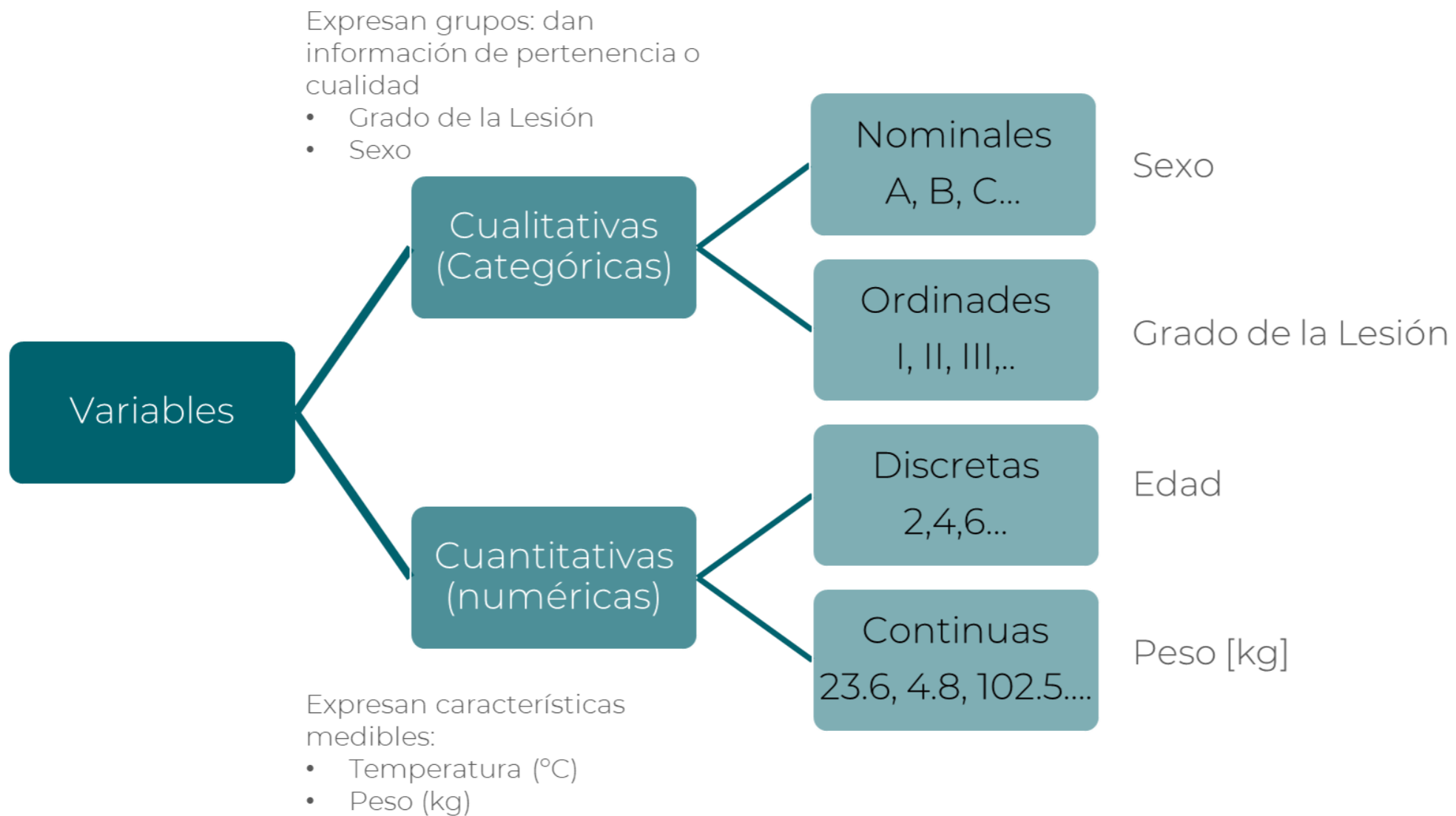


Tabla de Datos y
Variables

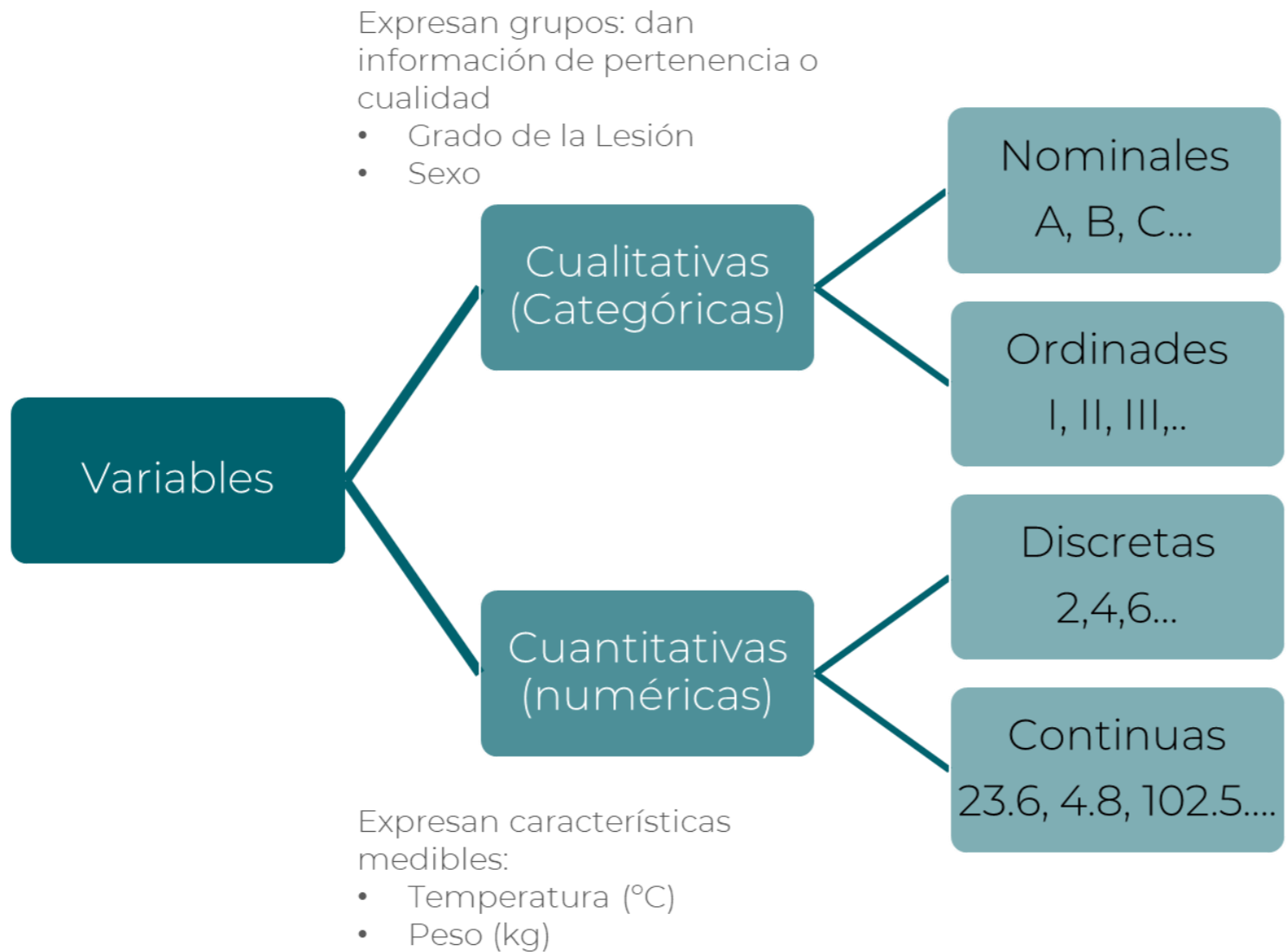
Descriptiva

Inferencial

Modelos

BASES
ESTADÍSTICAS

Características medibles de los individuos de una muestra



Fisher's Iris Data				
Largo de sépalo ⇅	Ancho de sépalo ⇅	Largo de pétalo ⇅	Ancho de pétalo ⇅	Especies ⇅
5.0	2.0	3.5	1.0	<i>I. versicolor</i>
6.0	2.2	4.0	1.0	<i>I. versicolor</i>
6.0	2.2	5.0	1.5	<i>I. virginica</i>
6.2	2.2	4.5	1.5	<i>I. versicolor</i>
4.5	2.3	1.3	0.3	<i>I. setosa</i>
5.0	2.3	3.3	1.0	<i>I. versicolor</i>
5.5	2.3	4.0	1.3	<i>I. versicolor</i>
6.3	2.3	4.4	1.3	<i>I. versicolor</i>
4.9	2.4	3.3	1.0	<i>I. versicolor</i>
5.5	2.4	3.7	1.0	<i>I. versicolor</i>
5.5	2.4	3.8	1.1	<i>I. versicolor</i>
4.9	2.5	4.5	1.7	<i>I. virginica</i>
5.1	2.5	3.0	1.1	<i>I. versicolor</i>
5.5	2.5	4.0	1.3	<i>I. versicolor</i>
5.6	2.5	3.9	1.1	<i>I. versicolor</i>
5.7	2.5	5.0	2.0	<i>I. virginica</i>
6.3	2.5	4.9	1.5	<i>I. versicolor</i>
6.3	2.5	5.0	1.9	<i>I. virginica</i>
6.7	2.5	5.8	1.8	<i>I. virginica</i>
5.5	2.6	4.4	1.2	<i>I. versicolor</i>
5.7	2.6	3.5	1.0	<i>I. versicolor</i>
5.8	2.6	4.0	1.2	<i>I. versicolor</i>
6.1	2.6	5.6	1.4	<i>I. virginica</i>

NUM. CONTINUA
(decimales)



Tabla de Datos y
Variables

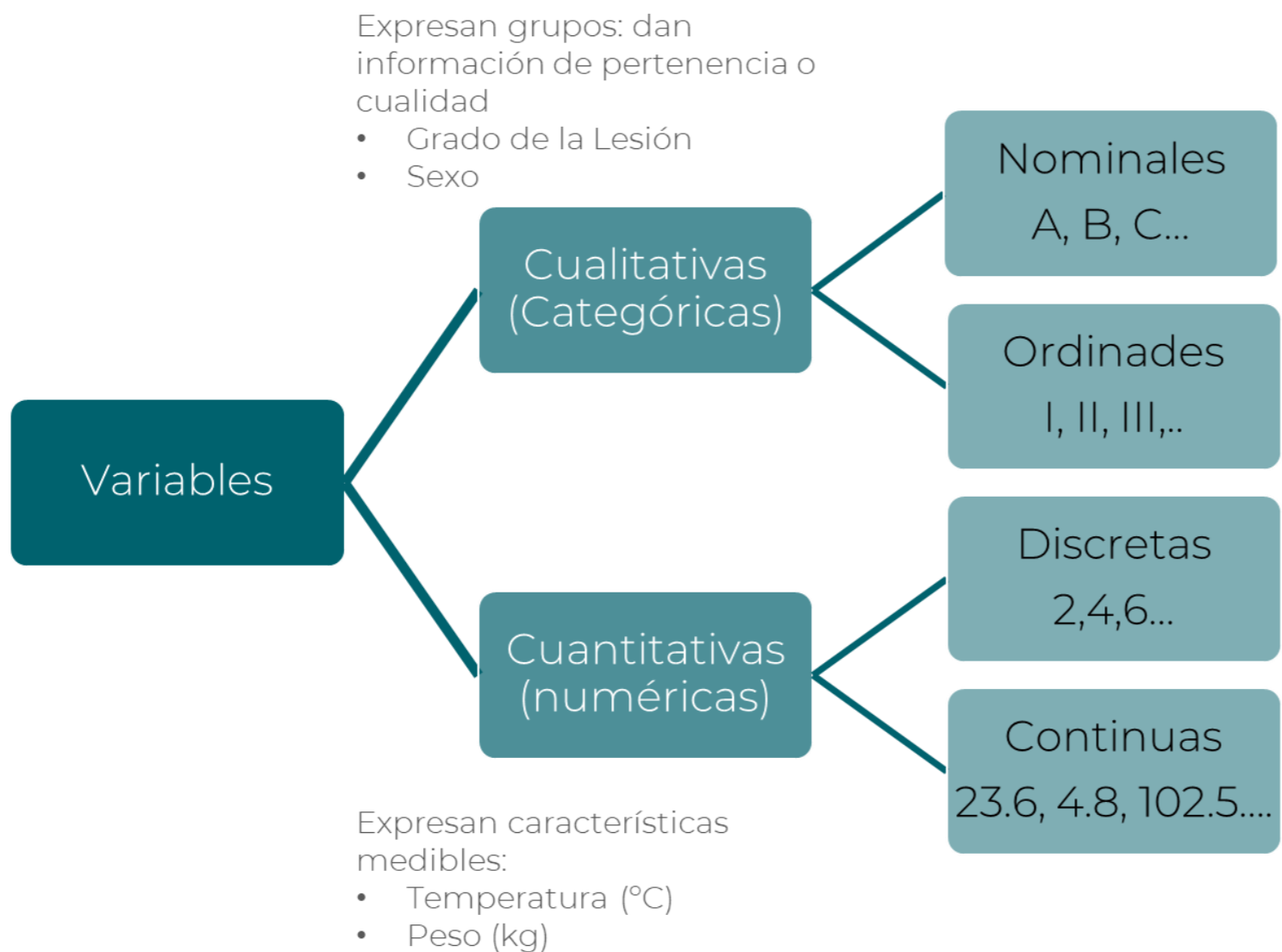
Descriptiva

Inferencial

Modelos

BASES
ESTADÍSTICAS

Características medibles de los individuos de una muestra



NUM. DISCRETA
(no decimales)

9.23 °C

NUM. CONTINUA
(decimales)

Tabla de Datos y
Variables

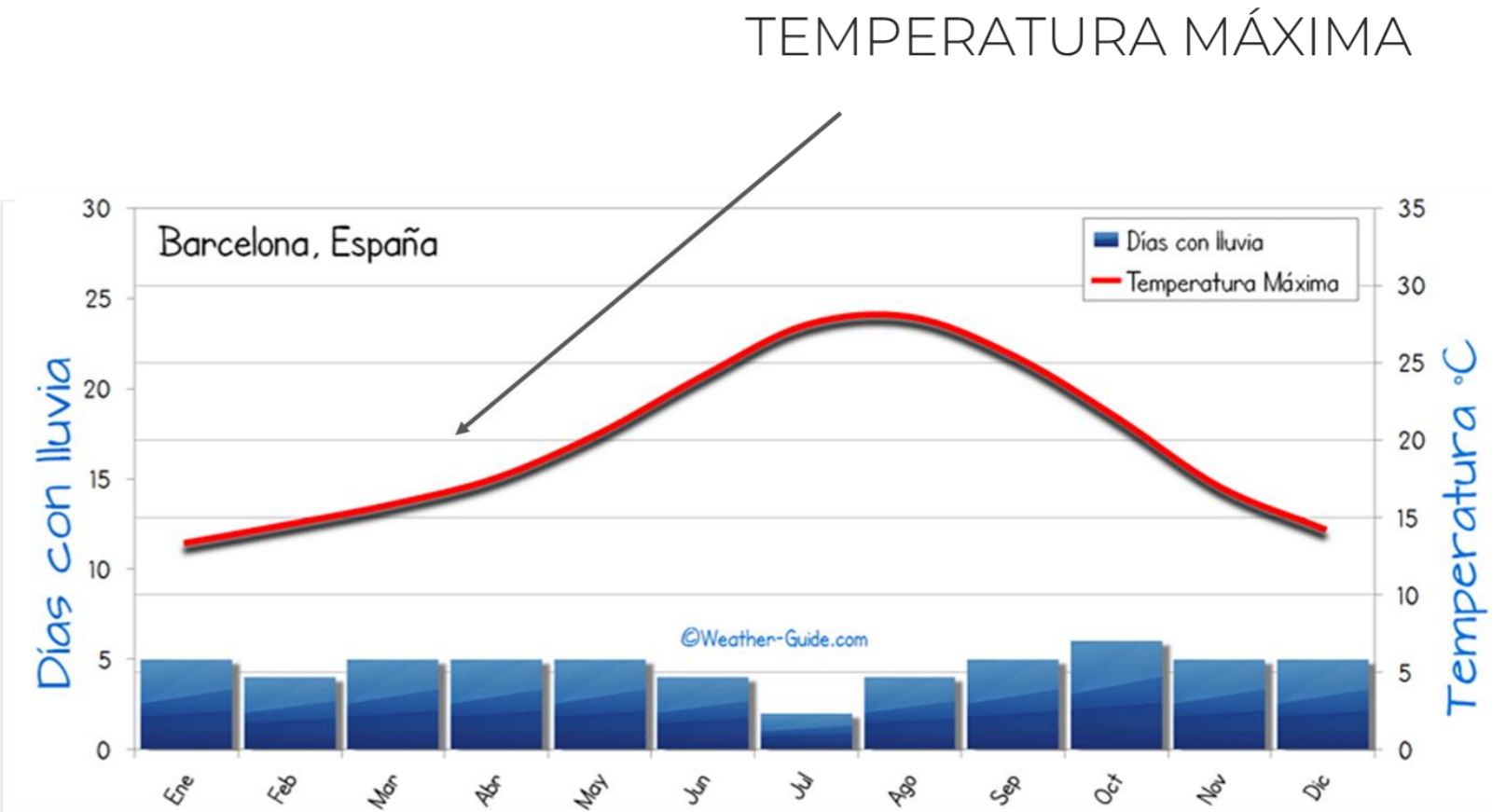
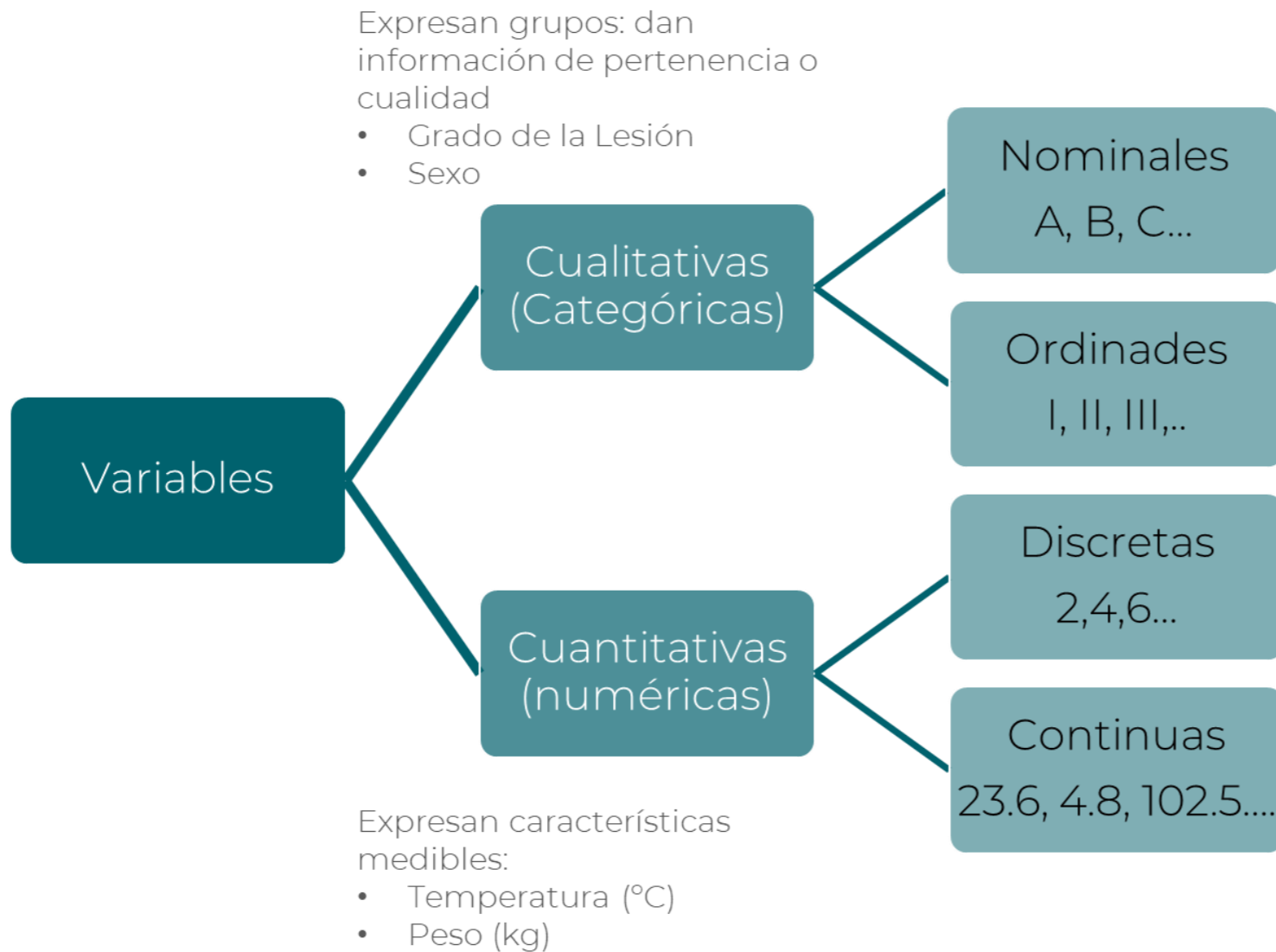
Descriptiva

Inferencial

Modelos

**BASES
ESTADÍSTICAS**

Características medibles de los individuos de una muestra



NUM. DISCRETA
(contar)

NÚMEROS DE
DÍAS CON
LLUVIA

Tabla de Datos y
Variables

Descriptiva

Inferencial

Modelos

BASES
ESTADÍSTICAS

Características medibles de los individuos de una muestra

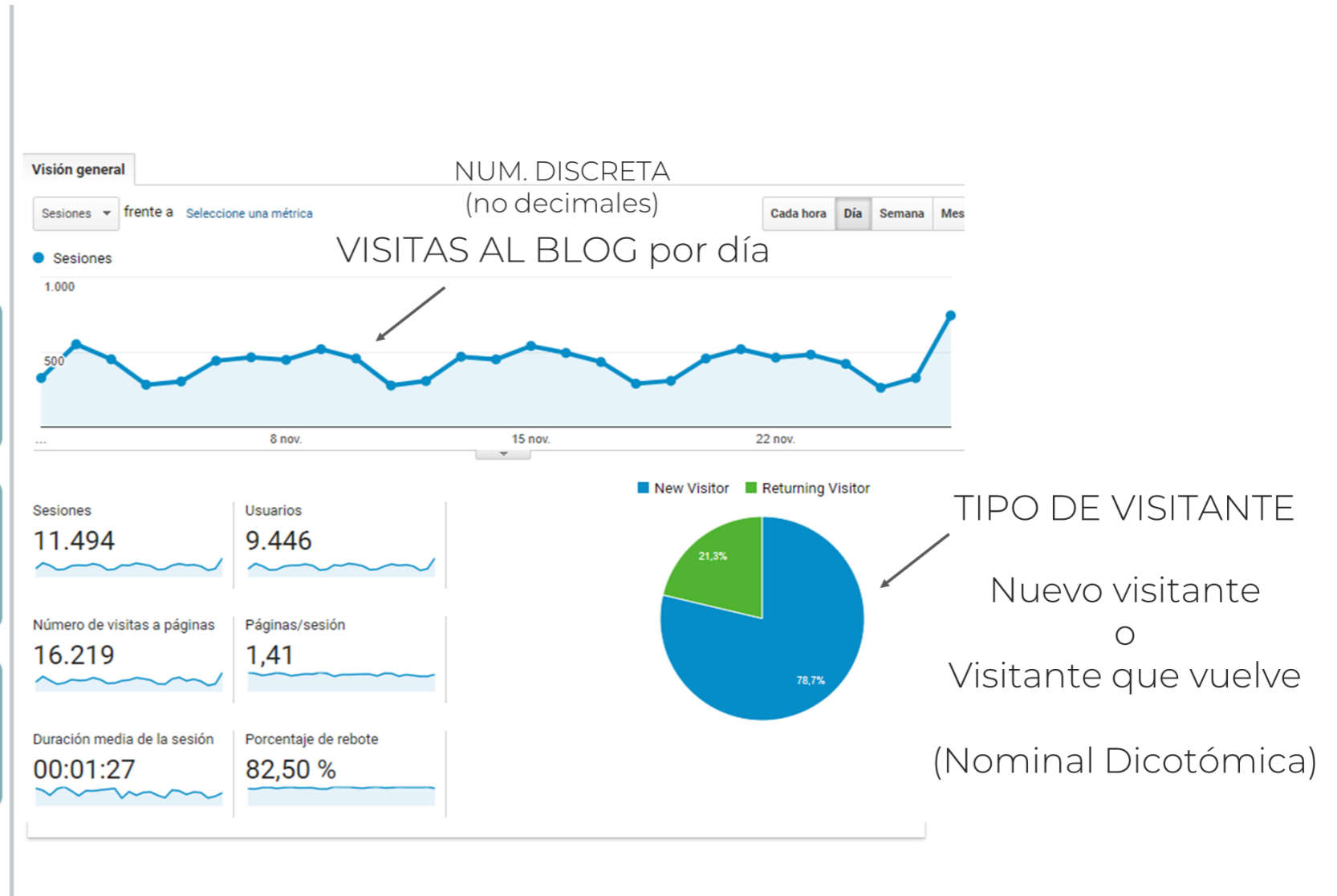
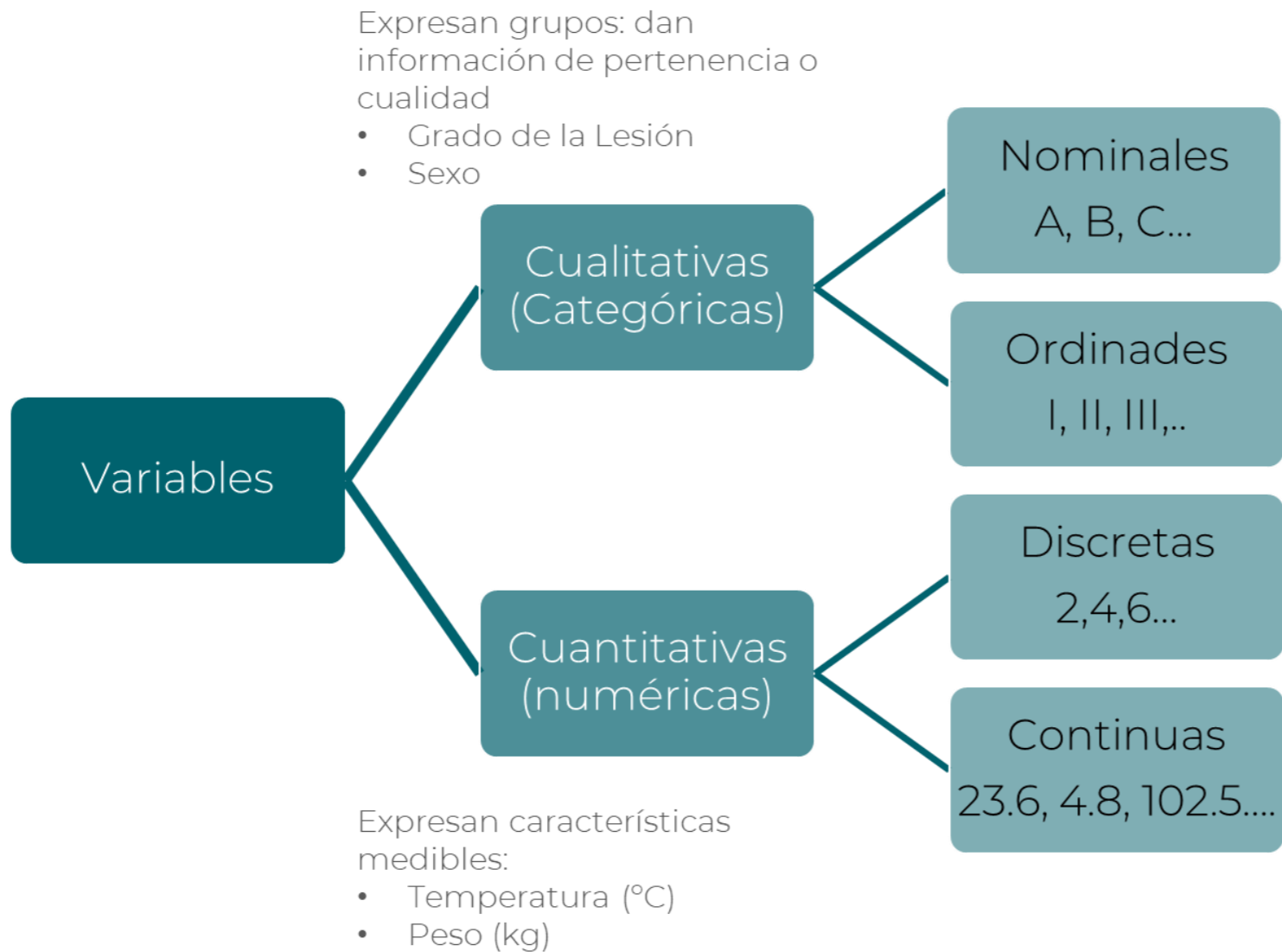


Tabla de Datos y Variables

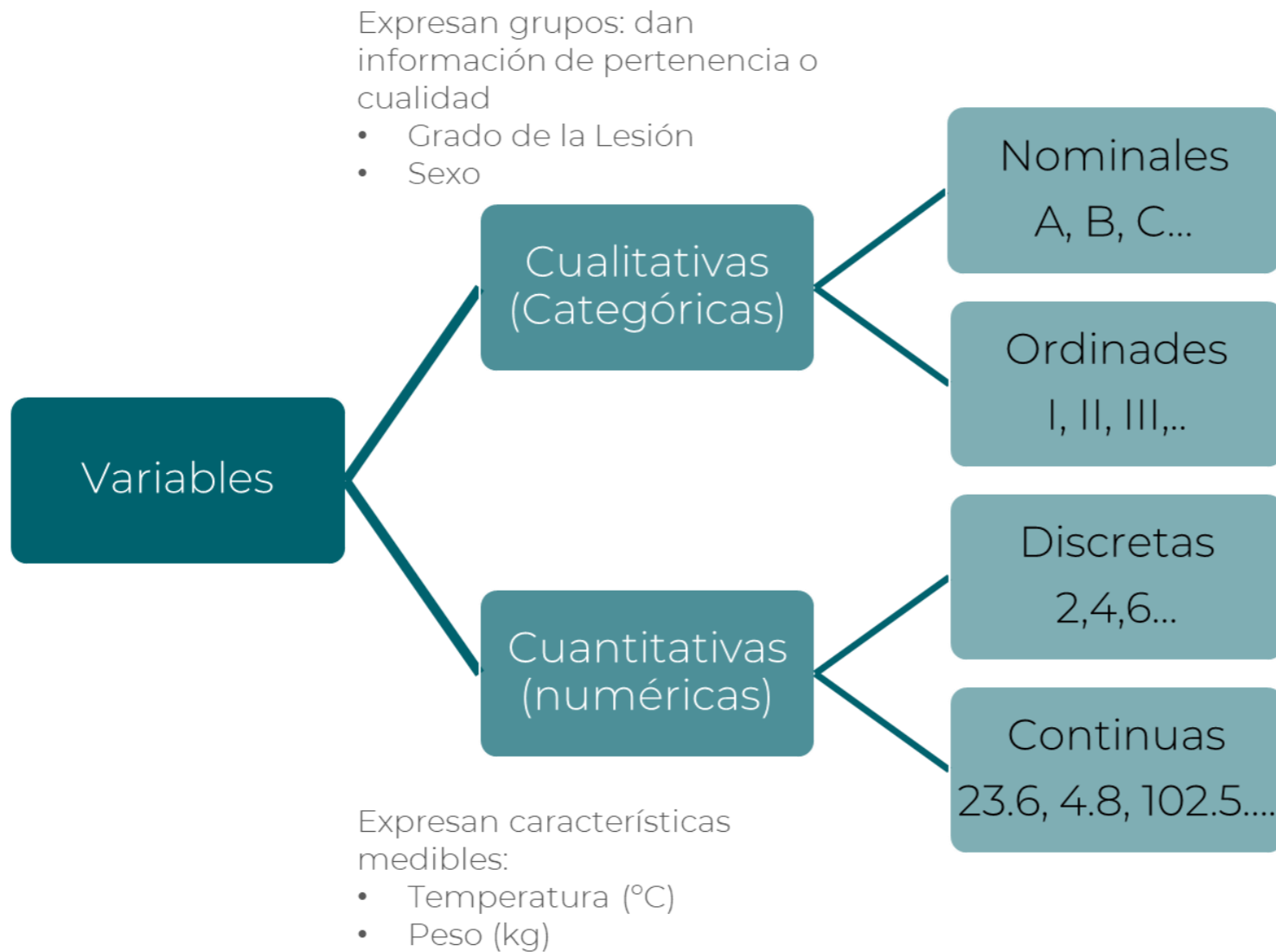
Descriptiva

Inferencial

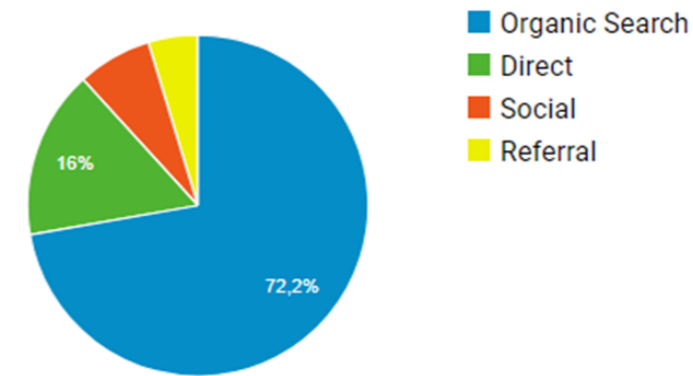
Modelos

BASES ESTADÍSTICAS

Características medibles de los individuos de una muestra



Canales principales

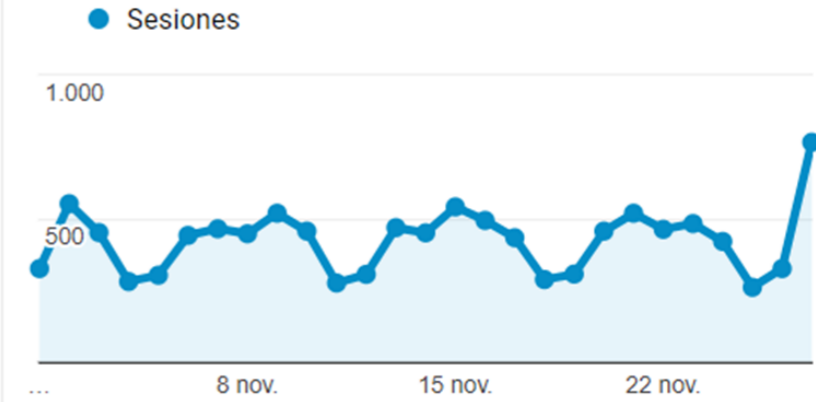


TIPO DE CANAL

Google
Directo
Redes Sociales
Link externo

(Nominal Politómica)

Sesiones



TIEMPO
Cualitativa Ordinal

Tabla de Datos y Variables

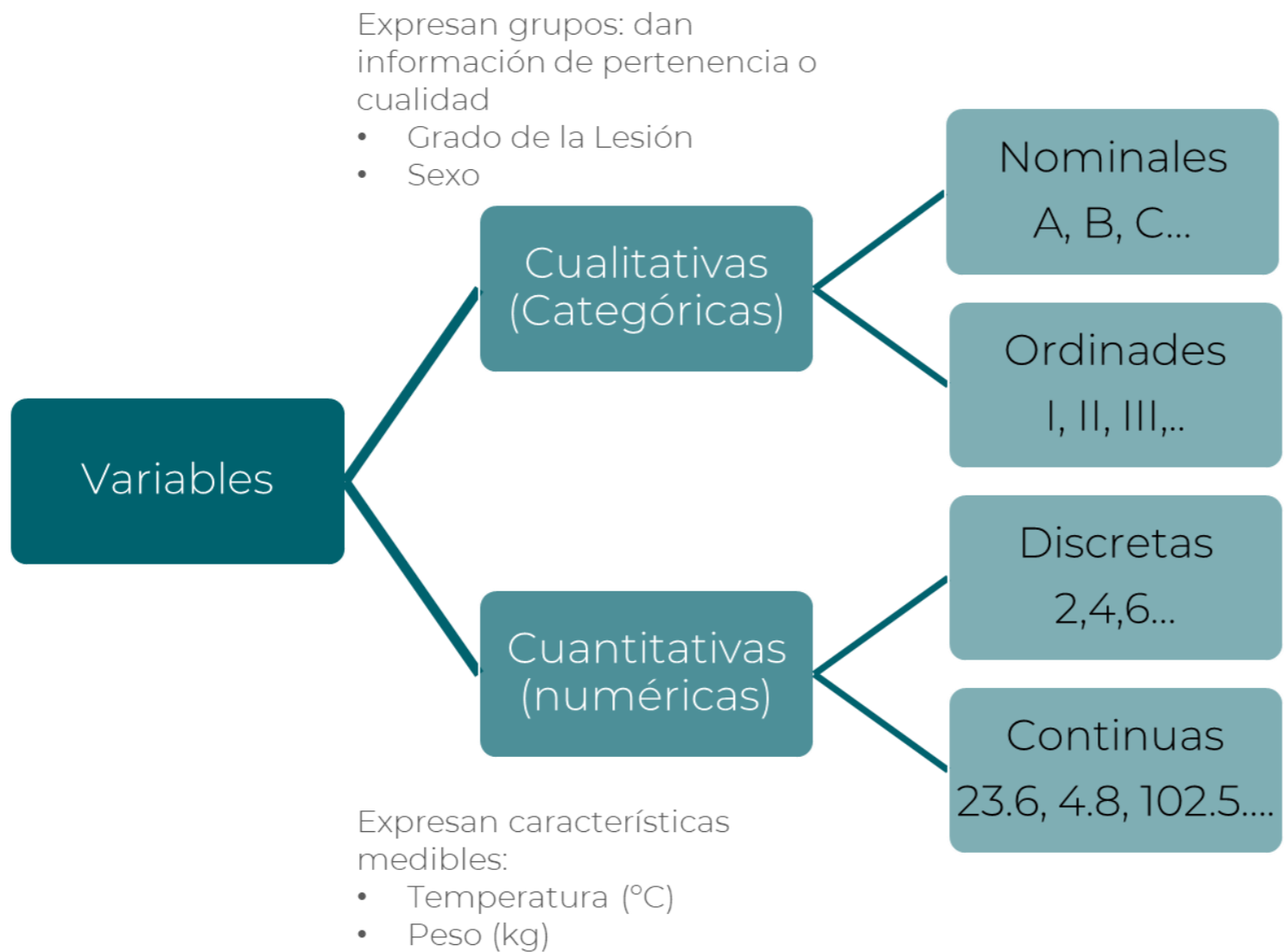
Descriptiva

Inferencial

Modelos

BASES
ESTADÍSTICAS

Características medibles de los individuos de una muestra



Fisher's Iris Data				
Largo de sépalo ↕	Ancho de sépalo ↕	Largo de pétalo ↕	Ancho de pétalo ↕	Especies ↕
5.0	2.0	3.5	1.0	<i>I. versicolor</i>
6.0	2.2	4.0	1.0	<i>I. versicolor</i>
6.0	2.2	5.0	1.5	<i>I. virginica</i>
6.2	2.2	4.5	1.5	<i>I. versicolor</i>
4.5	2.3	1.3	0.3	<i>I. setosa</i>
5.0	2.3	3.3	1.0	<i>I. versicolor</i>
5.5	2.3	4.0	1.3	<i>I. versicolor</i>
6.3	2.3	4.4	1.3	<i>I. versicolor</i>
4.9	2.4	3.3	1.0	<i>I. versicolor</i>
5.5	2.4	3.7	1.0	<i>I. versicolor</i>
5.5	2.4	3.8	1.1	<i>I. versicolor</i>
4.9	2.5	4.5	1.7	<i>I. virginica</i>
5.1	2.5	3.0	1.1	<i>I. versicolor</i>
5.5	2.5	4.0	1.3	<i>I. versicolor</i>
5.6	2.5	3.9	1.1	<i>I. versicolor</i>
5.7	2.5	5.0	2.0	<i>I. virginica</i>
6.3	2.5	4.9	1.5	<i>I. versicolor</i>
6.3	2.5	5.0	1.9	<i>I. virginica</i>
6.7	2.5	5.8	1.8	<i>I. virginica</i>
5.5	2.6	4.4	1.2	<i>I. versicolor</i>
5.7	2.6	3.5	1.0	<i>I. versicolor</i>
5.8	2.6	4.0	1.2	<i>I. versicolor</i>
6.1	2.6	5.6	1.4	<i>I. virginica</i>

TIPO DE ESPECIE

Versicolor
Virginica
Setosa

(Cualitativa Nominal)



Tabla de Datos y
Variables

Descriptiva

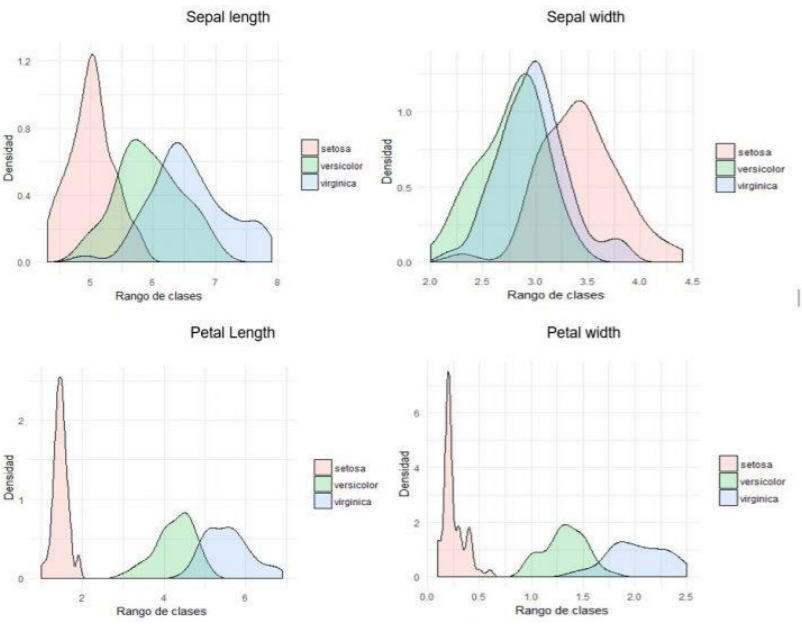
Inferencial

Modelos

BASES
ESTADÍSTICAS

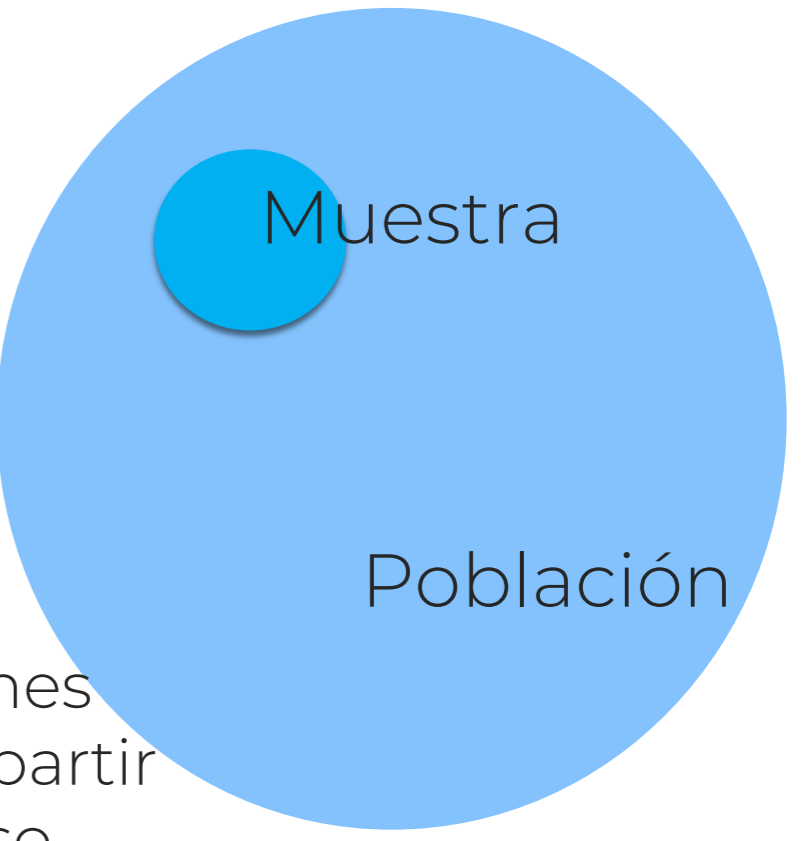
DESCRIPTIVA

INFERENCIAL



Fisher's Iris Data

Largo de sépalo ↕	Ancho de sépalo ↕	Largo de pétalo ↕	Ancho de pétalo ↕	Especies ↕
5.0	2.0	3.5	1.0	<i>I. versicolor</i>
6.0	2.2	4.0	1.0	<i>I. versicolor</i>
6.0	2.2	5.0	1.5	<i>I. virginica</i>
6.2	2.2	4.5	1.5	<i>I. versicolor</i>
4.5	2.3	1.3	0.3	<i>I. setosa</i>
5.0	2.3	3.3	1.0	<i>I. versicolor</i>
5.5	2.3	4.0	1.3	<i>I. versicolor</i>
6.3	2.3	4.4	1.3	<i>I. versicolor</i>
4.9	2.4	3.3	1.0	<i>I. versicolor</i>
5.5	2.4	3.7	1.0	<i>I. versicolor</i>
5.5	2.4	3.8	1.1	<i>I. versicolor</i>
4.9	2.5	4.5	1.7	<i>I. virginica</i>
5.1	2.5	3.0	1.1	<i>I. versicolor</i>
5.5	2.5	4.0	1.3	<i>I. versicolor</i>
5.6	2.5	3.9	1.1	<i>I. versicolor</i>
5.7	2.5	5.0	2.0	<i>I. virginica</i>
6.3	2.5	4.9	1.5	<i>I. versicolor</i>
6.3	2.5	5.0	1.9	<i>I. virginica</i>
6.7	2.5	5.8	1.8	<i>I. virginica</i>
5.5	2.6	4.4	1.2	<i>I. versicolor</i>
5.7	2.6	3.5	1.0	<i>I. versicolor</i>
5.8	2.6	4.0	1.2	<i>I. versicolor</i>
6.1	2.6	5.6	1.4	<i>I. virginica</i>



Conclusiones
generales a partir
de un caso
particular

Modelizar Datos

Dibuja la tabla de
datos con gráficos y
calcula sus
características

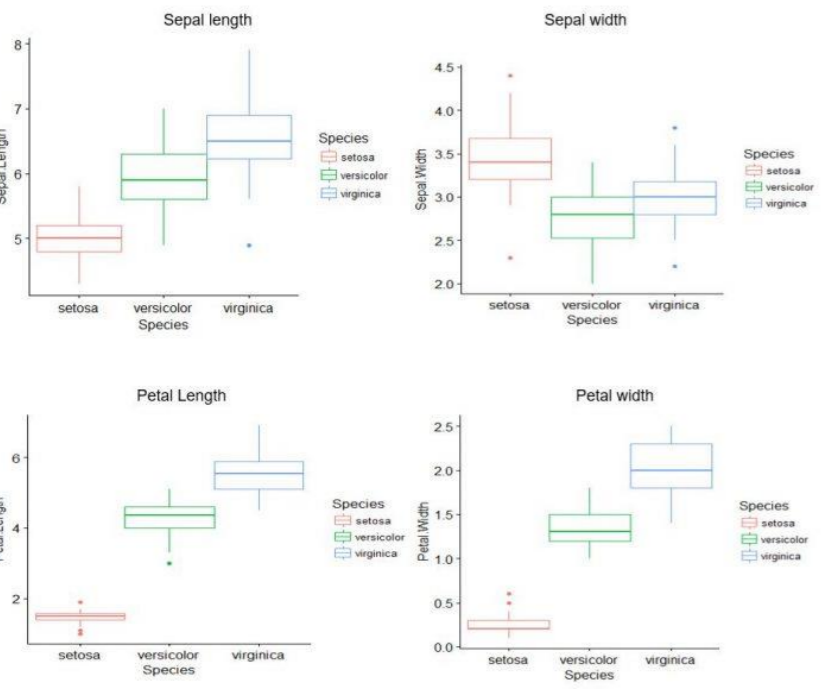


Tabla de Datos y
Variables

Descriptiva

Inferencial

Modelos

BASES
ESTADÍSTICAS

coggle
made for free at coggle.it

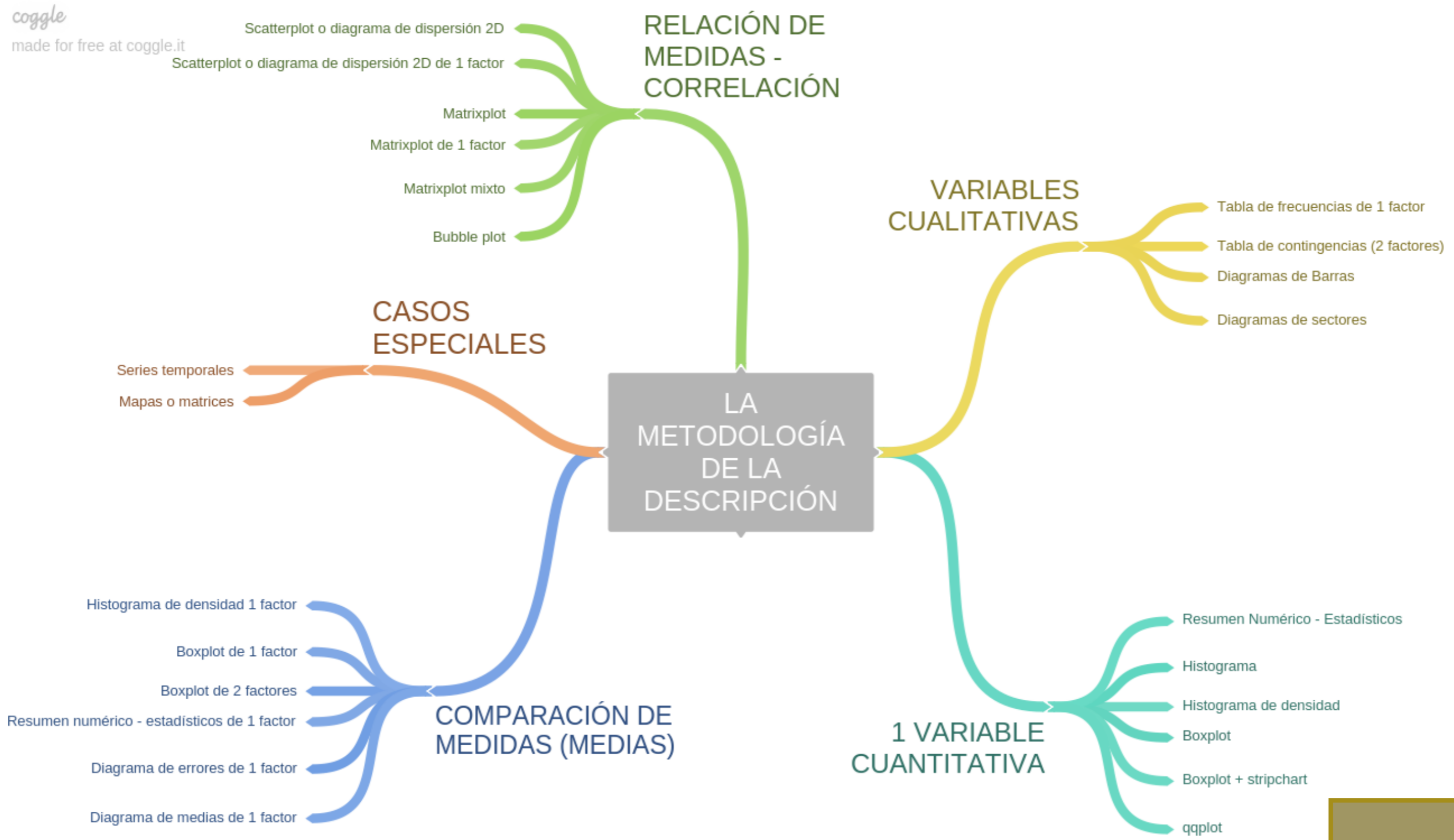


Tabla de Datos y Variables

Descriptiva

Inferencial

Modelos

BASES
ESTADÍSTICAS

Estadística descriptiva

Podemos clasificar los conceptos en función de qué variables describimos o dibujamos:

Variables cuantitativas:

- 5. Tabla de frecuencias (relativas, absolutas)
- 6. Valor central (media o mediana))
- 7. Dispersión (desviación estándar, rango intercuartílico)
- 8. Medidas de posición
- 9. Distribución (frecuencia)
- 10. Correlación y coeficiente de correlación
- 11. Scatterplot y regresión

Variables cualitativas

- 12. Tablas de contingencias (relativas, absolutas, condicionadas)
- 13. Concepto de proporciones
- 14. Distribución
- 15. Diagrama de barras y sectores

Estadística inferencial

En la **estadística inferencial** es vital entender:

- 16. Muestra y población
- 17. Qué son las distribuciones de variables
- 18. Qué es y para qué sirve la distribución de densidad de probabilidad y las probabilidades
- 19. El poder de la distribución normal
 - o Media y desviación estándar
 - o Intervalos de confianza
- 20. El poder de la distribución t
 - o T-student
- 22. El contraste de hipótesis y el p-valor
 - o Qué es una hipótesis
 - o Hipótesis alternativa
 - o Hipótesis nula
 - o P-valor
 - o Test estadístico
- 23. La esencia de los modelos predictivos y sus dos súper poderes
 - o Coeficientes
 - o Residuos o error

conceptosclaros.com/estadistica-aplicada

Tabla de Datos y Variables

Descriptiva

Inferencial

Modelos

BASES
ESTADÍSTICAS

¿Qué técnica aplico en los distintos casos?

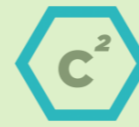




TÉCNICAS DE
ANÁLISIS DE
DATOS

SISTEMA

BASES
ESTADÍSTICAS



Modelos que nos
permitan predecir
acontecimientos,
patrones

HERRAMIENTAS

**NIVEL III – Predictivo y
reconocimiento de
patrones**

Buscamos responder a
preguntas ...

RESPUESTAS

NIVEL II – Inferencial

- Comparativo
- Relacional
- Causa-efecto

Buscamos indicios, información

TENDENCIAS / PATRONES

NIVEL I - Descriptivo

El mapa de Técnicas



- I- Comparativas
- II- Relacionales
- III- Causales
- IV- Reconocimiento de Patrones
- V- Predictivas



El mapa de Técnicas



- I- Comparativas
- II- Relacionales
- III- Causales
- IV- Reconocimiento de Patrones
- V- Predictivas



I- Comparativas

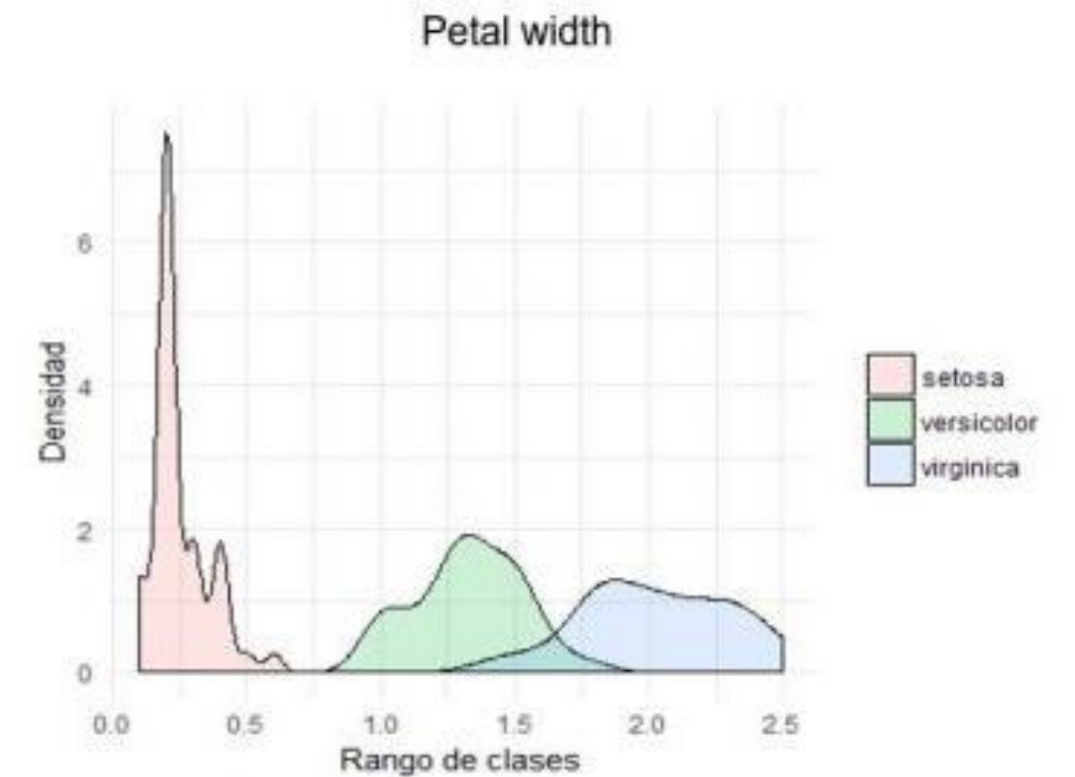
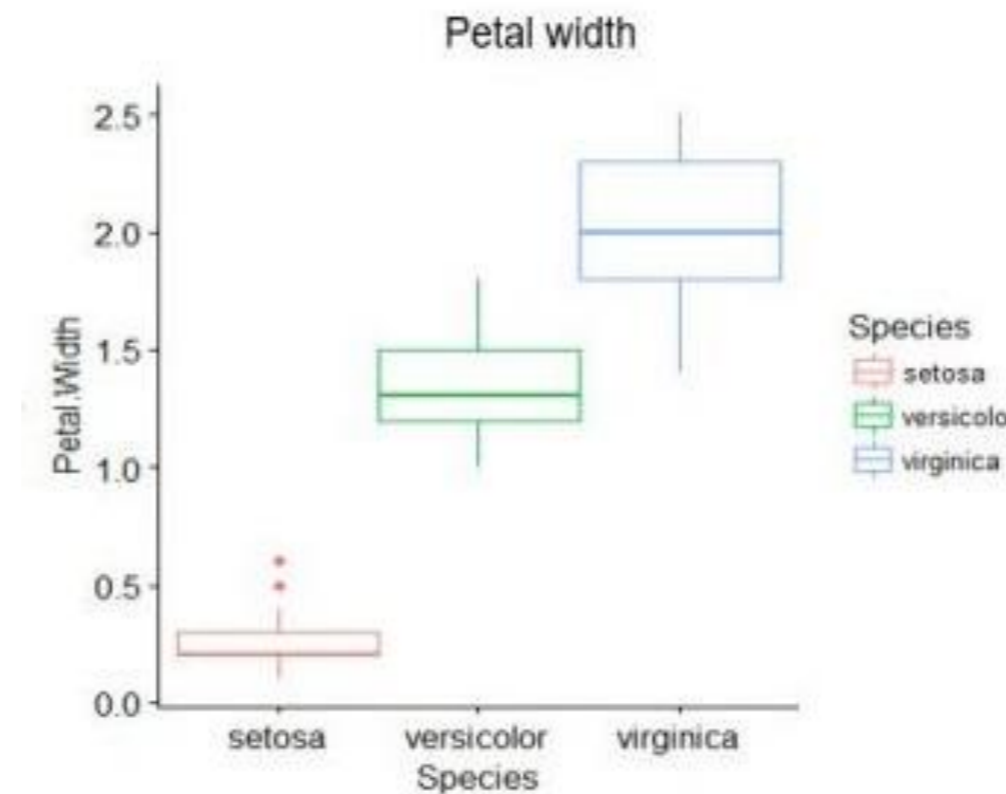
1. Cuantitativa vs Grupos
2. Cualitativa vs Grupos

Comparar
Medias

T-test
ANOVA
MANOVA
etc...

Ancho de pétalo ↕	Especies ↕
1.0	<i>I. versicolor</i>
1.0	<i>I. versicolor</i>
1.5	<i>I. virginica</i>
1.5	<i>I. versicolor</i>
0.3	<i>I. setosa</i>
1.0	<i>I. versicolor</i>
1.3	<i>I. versicolor</i>
1.3	<i>I. versicolor</i>
1.0	<i>I. versicolor</i>
1.0	<i>I. versicolor</i>
1.1	<i>I. versicolor</i>
1.7	<i>I. virginica</i>
1.1	<i>I. versicolor</i>
1.3	<i>I. versicolor</i>
1.1	<i>I. versicolor</i>
2.0	<i>I. virginica</i>
1.5	<i>I. versicolor</i>
1.9	<i>I. virginica</i>
1.8	<i>I. virginica</i>
1.2	<i>I. versicolor</i>
1.0	<i>I. versicolor</i>
1.2	<i>I. versicolor</i>
1.4	<i>I. virginica</i>

El mapa de Técnicas

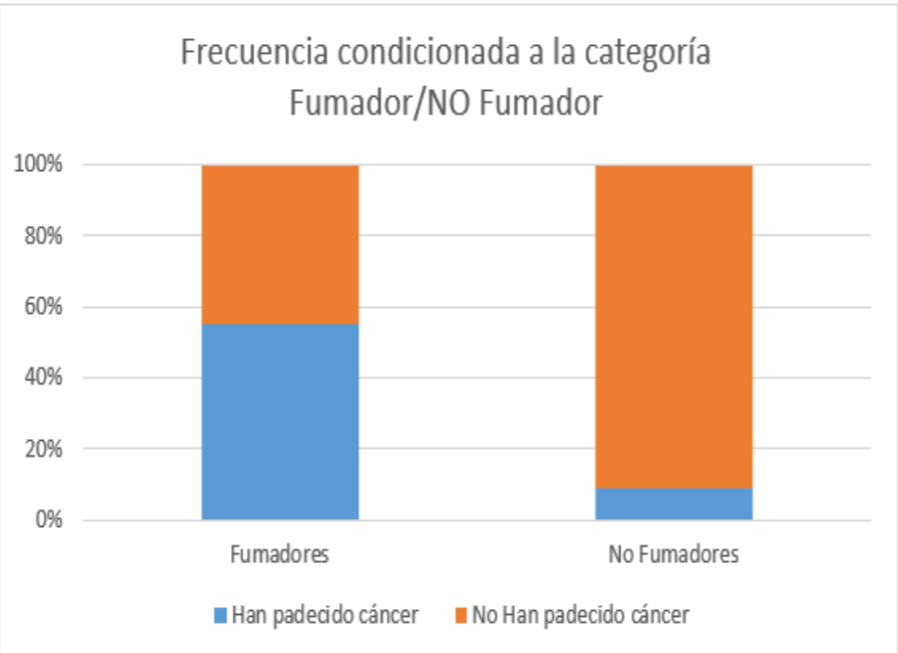
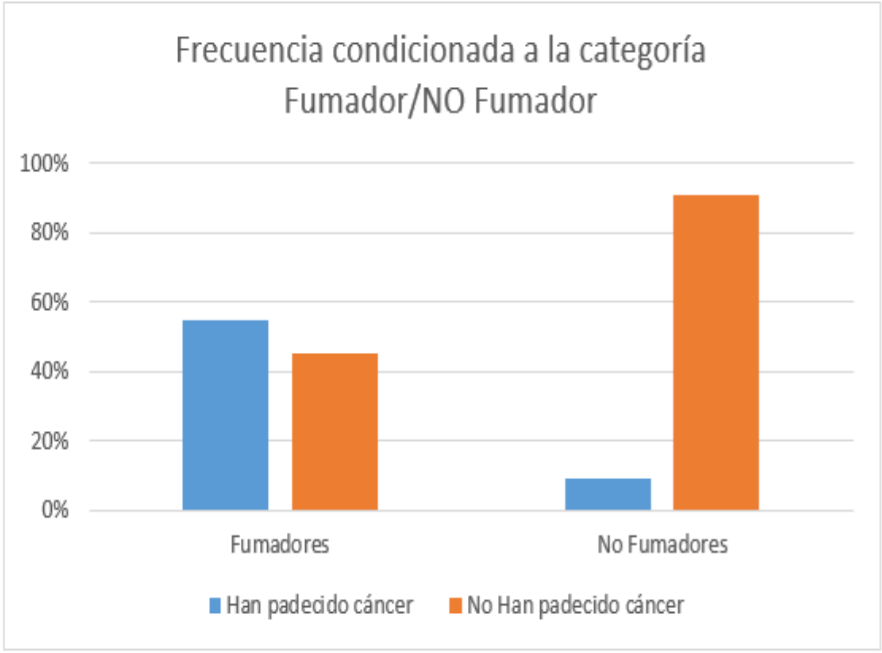
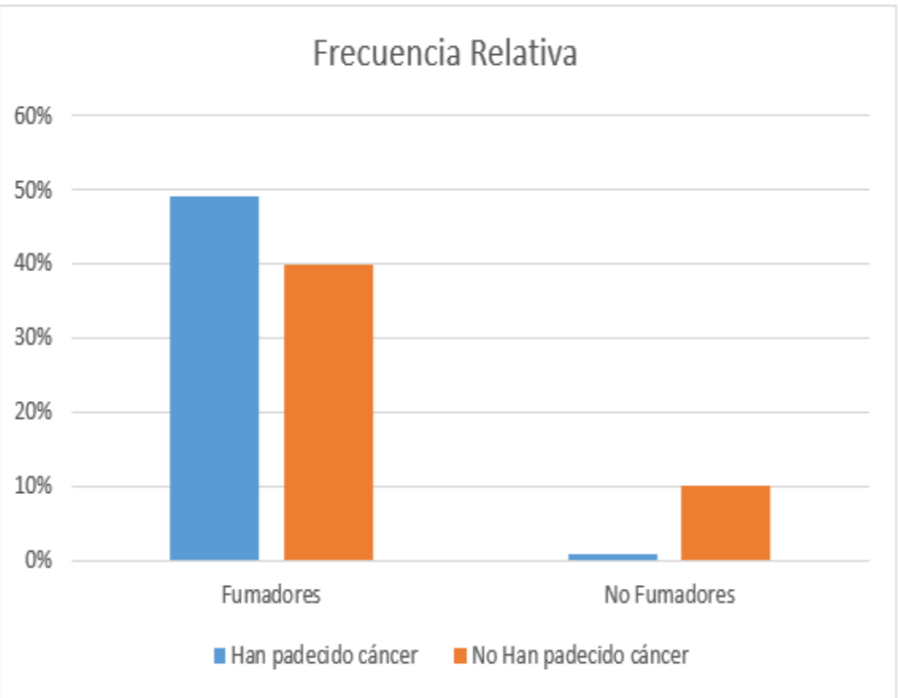
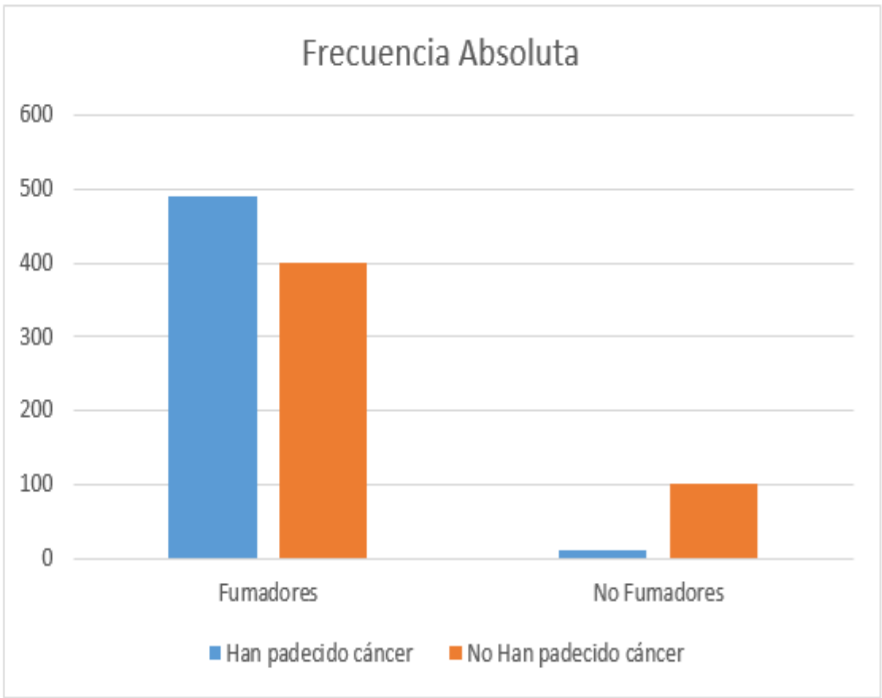


Normalmente la comparación de grupos se utiliza para ver la influencia de variables categóricas con una numérica

I- Comparativas

- 1. Cuantitativa vs Grupos
- 2. Cualitativa vs Grupos

El mapa de Técnicas



FRECUENCIAS ABSOLUTAS

	Fumadores	No Fumadores	Total
Han padecido cáncer	490	10	500
No Han padecido cáncer	400	100	500
Total	890	110	1000

FRECUENCIAS RELATIVAS

	Fumadores	No Fumadores	Total
Han padecido cáncer	49%	1%	50%
No Han padecido cáncer	40%	10%	50%
Total	89%	11%	100%

FRECUENCIA CONDICIONADA

	Fumadores	No Fumadores
Han padecido cáncer	55%	9%
No Han padecido cáncer	45%	91%
Total	100%	100%

La Tabla de contingencia se utiliza para contar las observaciones / individuos de diferentes grupos

El mapa de Técnicas



- I- Comparativas
- **II- Relacionales**
- III- Causales
- IV- Reconocimiento de Patrones
- V- Predictivas



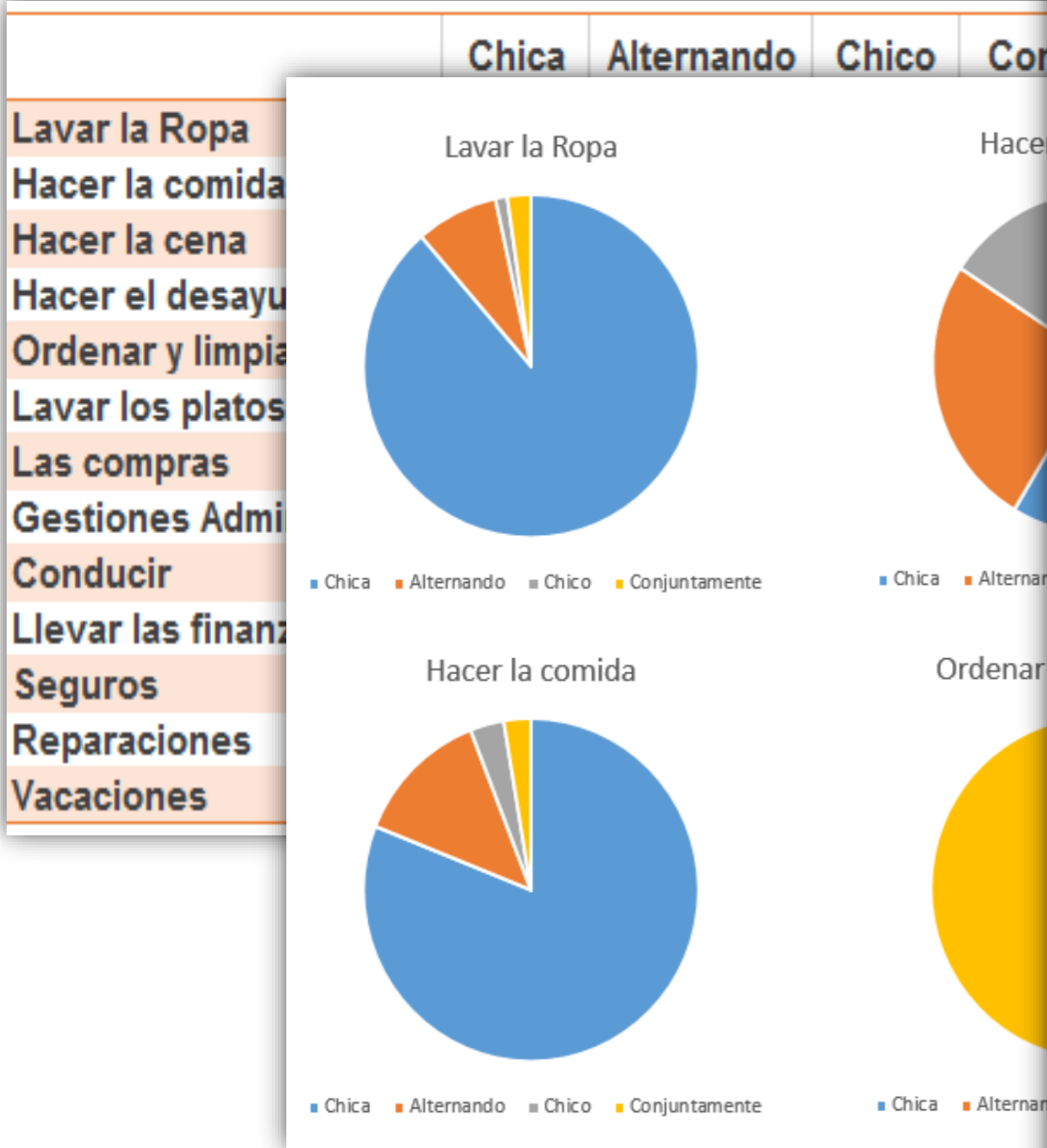
II- Relacionales

1. Cualitativa vs Cualitativa

2. Cuantitativa vs Cuanti.

...

El mapa de Técnicas



Asociar
proporciones

Chi Cuadrado
Residuos de
Pearson
Contribuciones

...

Asociar significa ver si
hay influencia entre los
grupos de 2 variables
categóricas

II- Relacionales

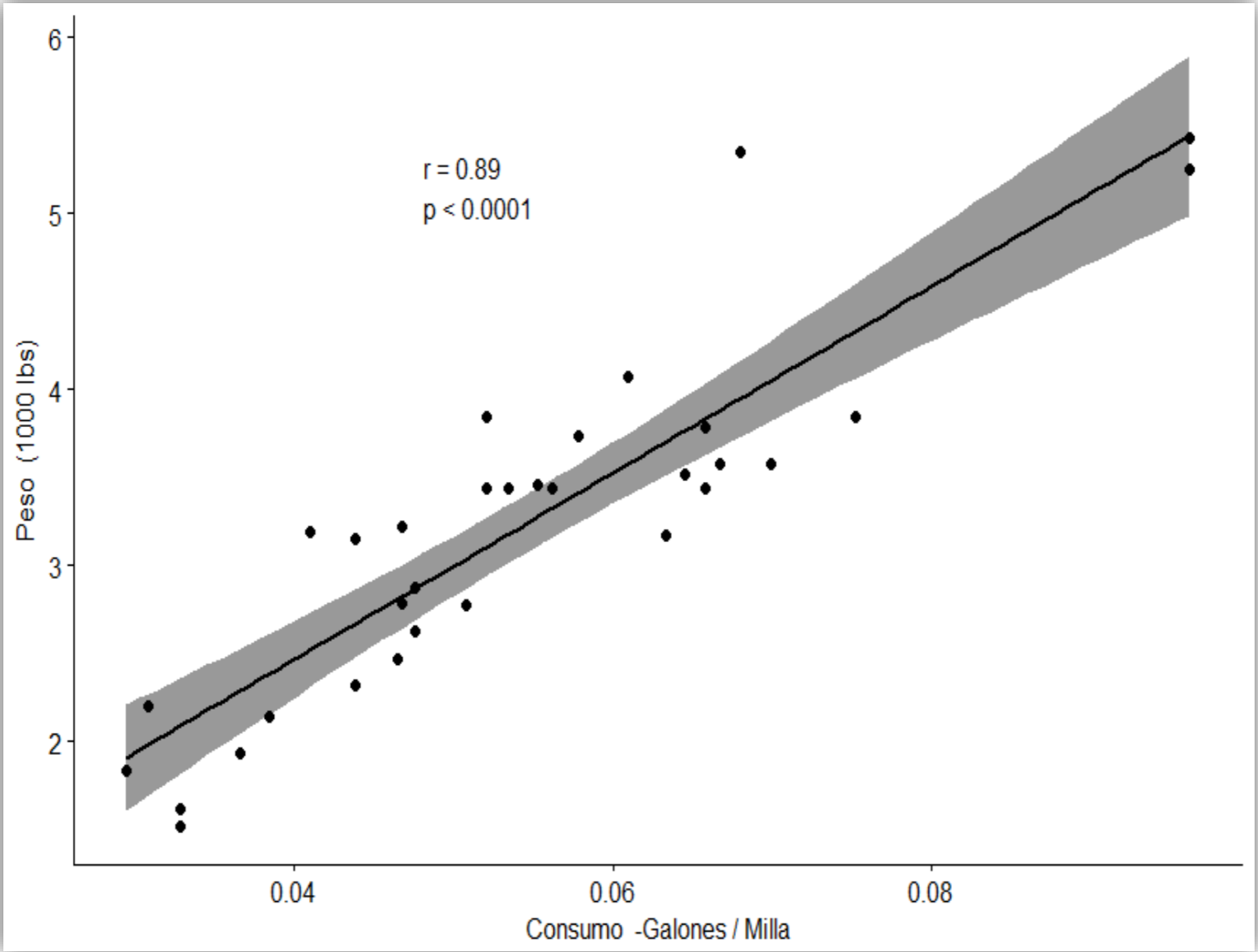
- 1. Cualitativa vs Cualitativa
- 2. Cuantitativa vs Cuanti.

...

El mapa de Técnicas



consumo	peso
0.04761905	2.620
0.04761905	2.875
0.04385965	2.320
0.04672897	3.215
0.05347594	3.440
0.05524862	3.460
0.06993007	3.570
0.04098361	3.190
0.04385965	3.150
0.05208333	3.440
0.05617978	3.440
0.06097561	4.070
0.05780347	3.730
0.06578947	3.780
0.09615385	5.250
0.09615385	5.424
0.06802721	5.345
0.03086420	2.200
0.03280474	1.615



Asociar medidas

Análisis de
Correlación
Pearson,
Spearman,
Kendall

...

Asociar variables
cuantitativas es encontrar
numéricamente cómo de
relacionadas están 2
variables

El mapa de Técnicas

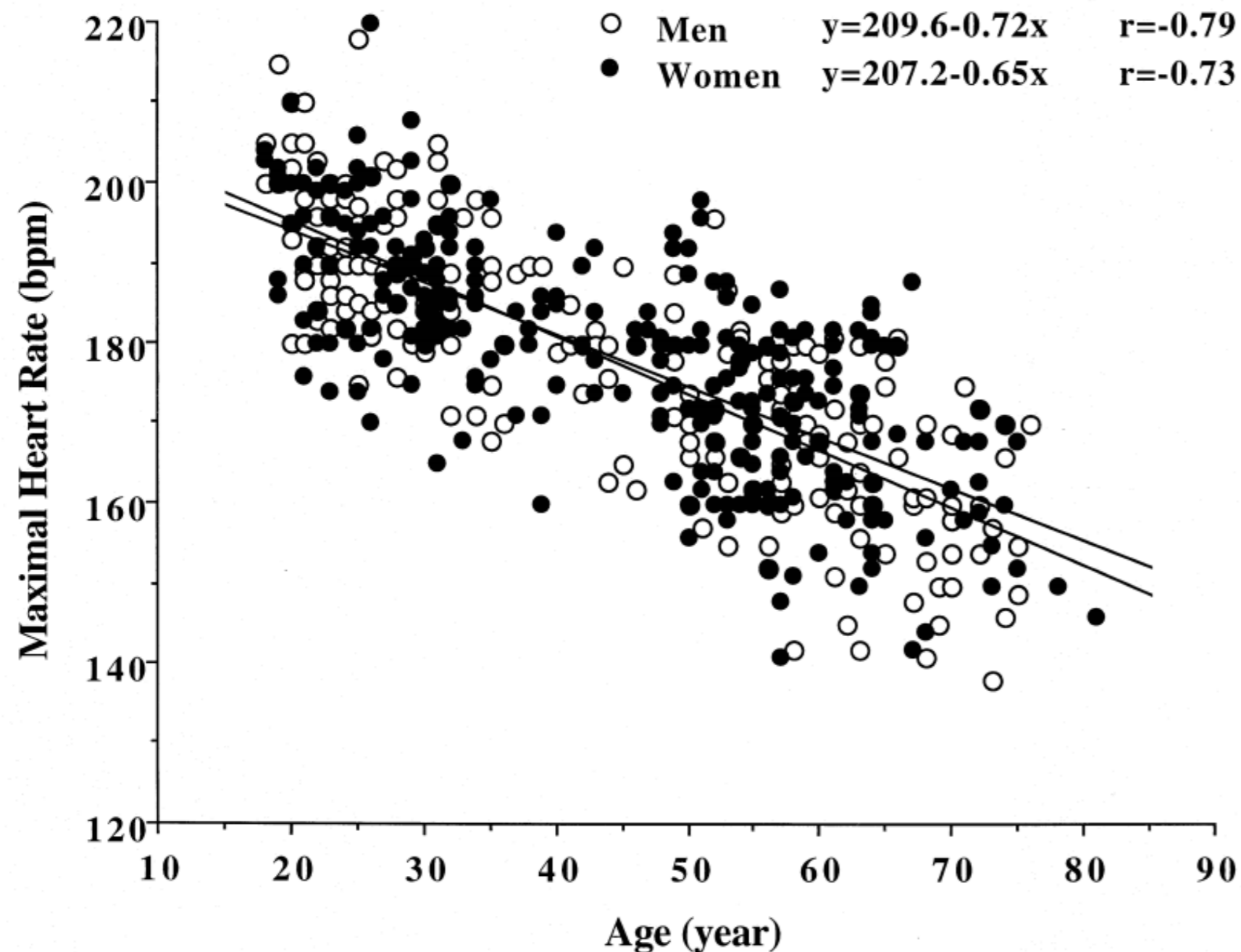


- I- Comparativas
- II- Relacionales
- **III- Causales (modelos estadísticos)**
- IV- Reconocimiento de Patrones
- V- Predictivas



III- Causales (modelos Estadísticos)

El mapa de Técnicas



Un modelo es una función matemática

Y nos sirva para:

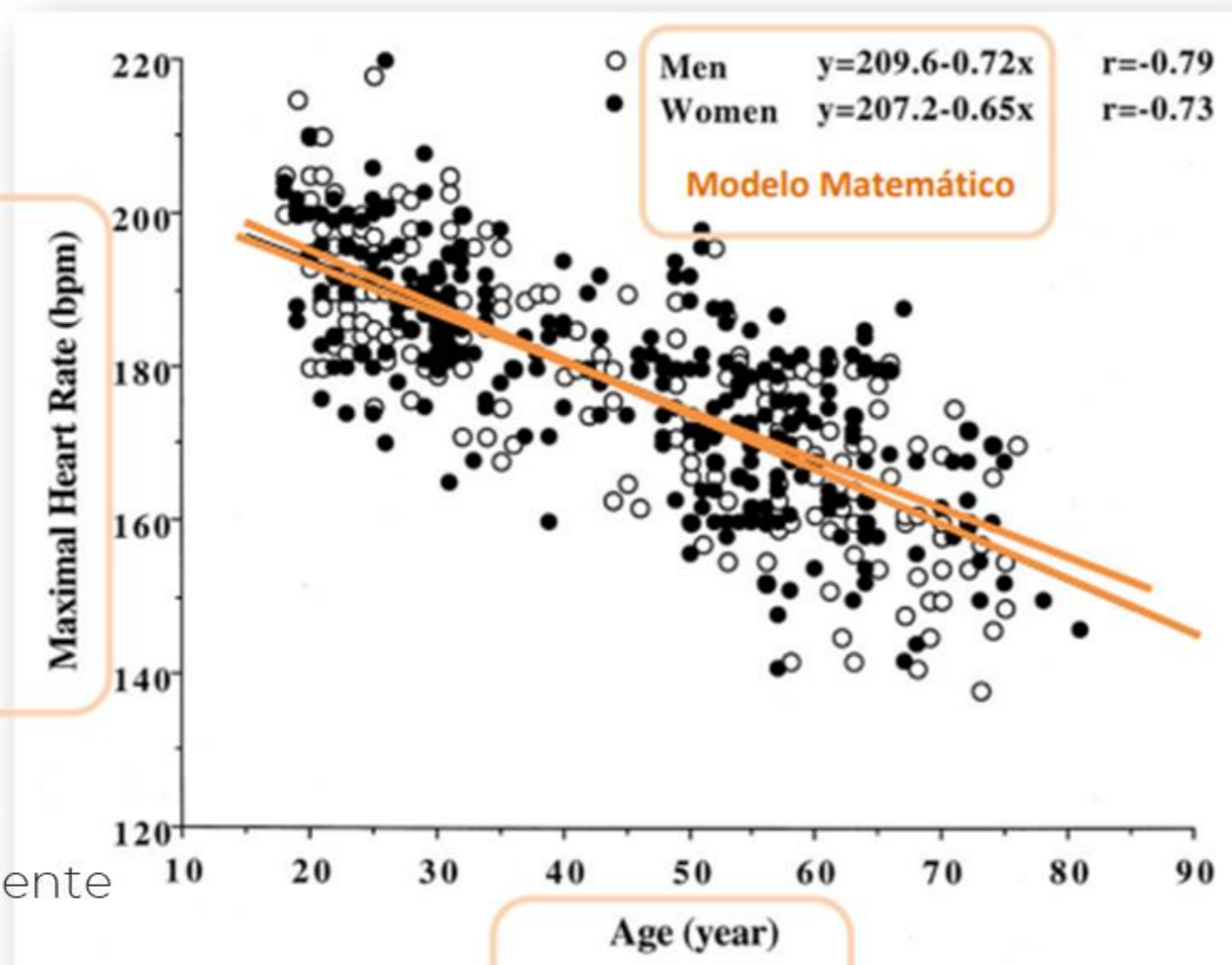
- Predecir
- Explicar relaciones

Modelos de
Predicción

Regresión lineal
Logísticos
Lineales
Generalizados

III- Causales (modelos Estadísticos)

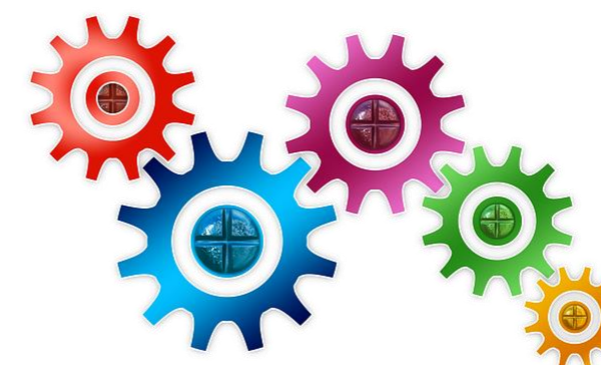
El mapa de Técnicas



Variable
Respuesta
O dependiente

X
Variable de Entrada

Variable
Respuesta
O independiente



X - EDAD
**Variable de
Entrada**

○ Men $y=209.6-0.72x$
● Women $y=207.2-0.65x$
Modelo Matemático

Y - FREC.
CARDÍACA MÁX
**Variable de
Salida**

Variable Salida – Y	Frecuencia Cardíaca Máxima
Variable Entrada – X	Edad
Función matemática	○ Men $y=209.6-0.72x$ ● Women $y=207.2-0.65x$

El mapa de Técnicas



- I- Comparativas
- II- Relacionales
- III- Causales (modelos estadísticos)

- IV- Reconocimiento de Patrones
- V- Predictivas

MACHINE
LEARNING



IV y V- Machine Learning

El mapa de Técnicas



V- Predictivas

IV- Reconocimiento de Patrones

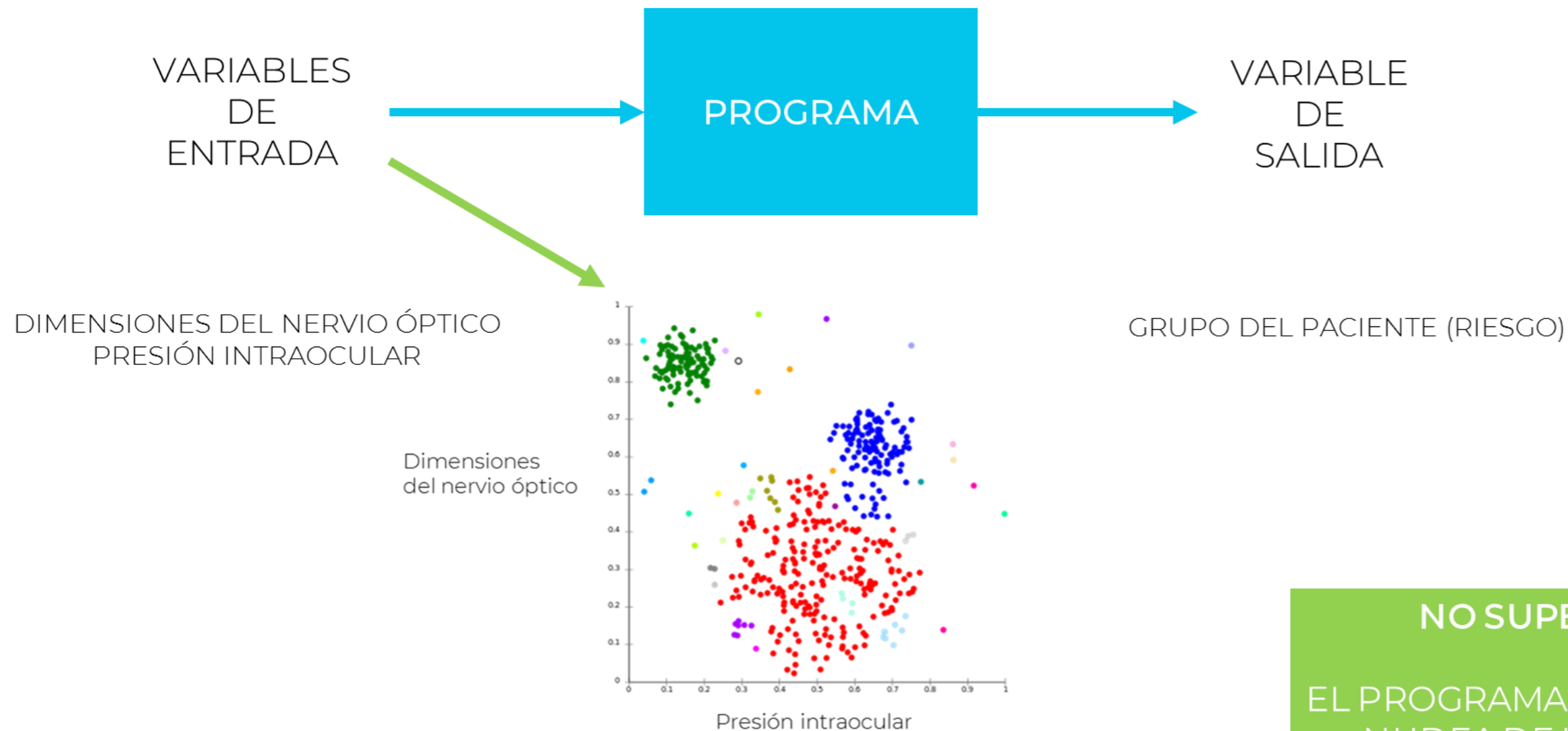


Se pueden dividir en:

- Supervisadas – reconocimiento de patrones
- No supervisadas - predictivas



- Reconocimiento de Patrones - NO Supervisadas



Se pueden dividir en:

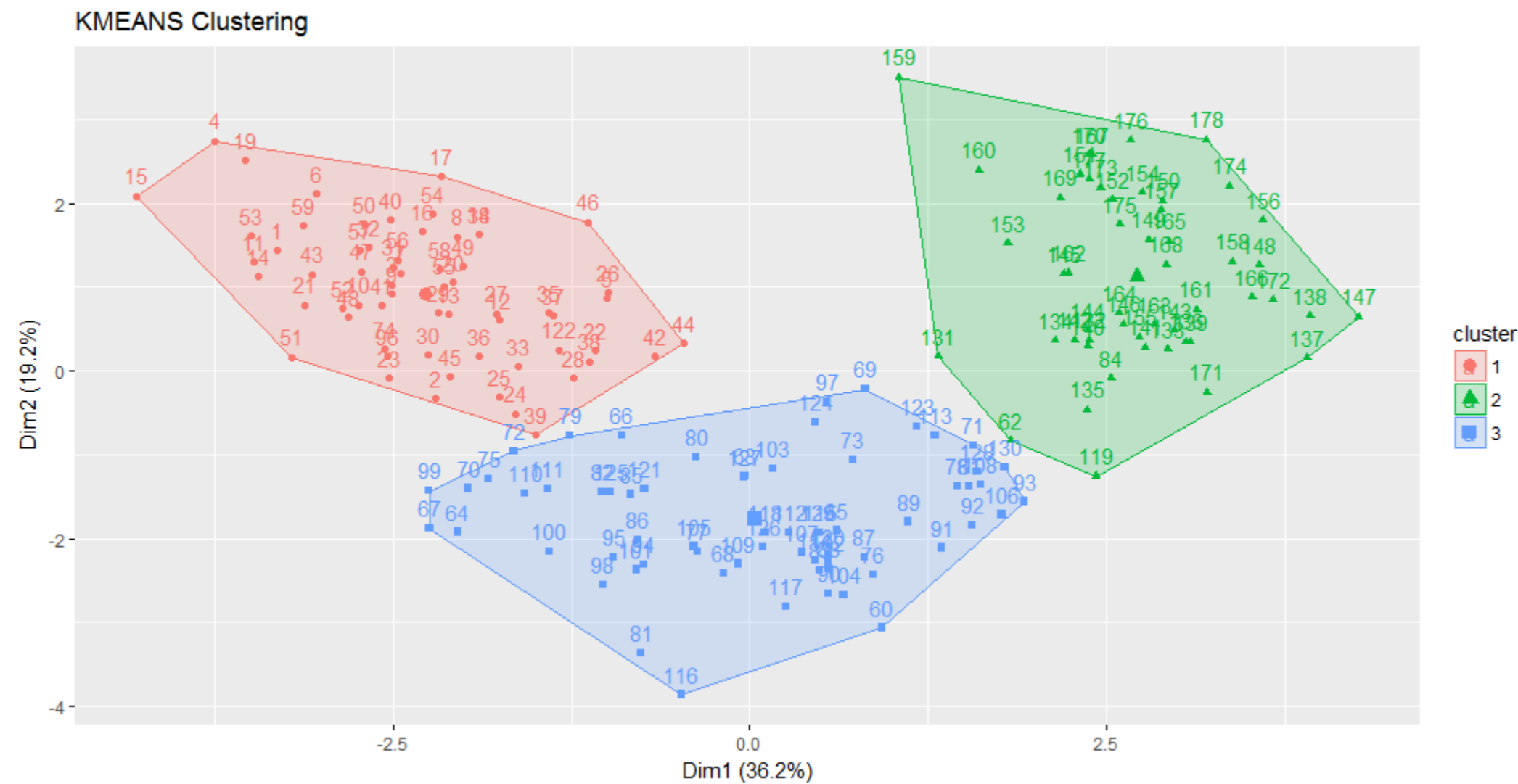
- Supervisadas – reconocimiento de patrones
- No supervisadas - predictivas

NO SUPERVISADO

EL PROGRAMA NO TIENE
NI IDEA DE LA SALIDA

IV y V- Machine Learning

El mapa de Técnicas



Se pueden dividir en:

- Supervisadas – reconocimiento de patrones
- No supervisadas - predictivas

IV – Reconocimiento de patrones

El mapa de Técnicas

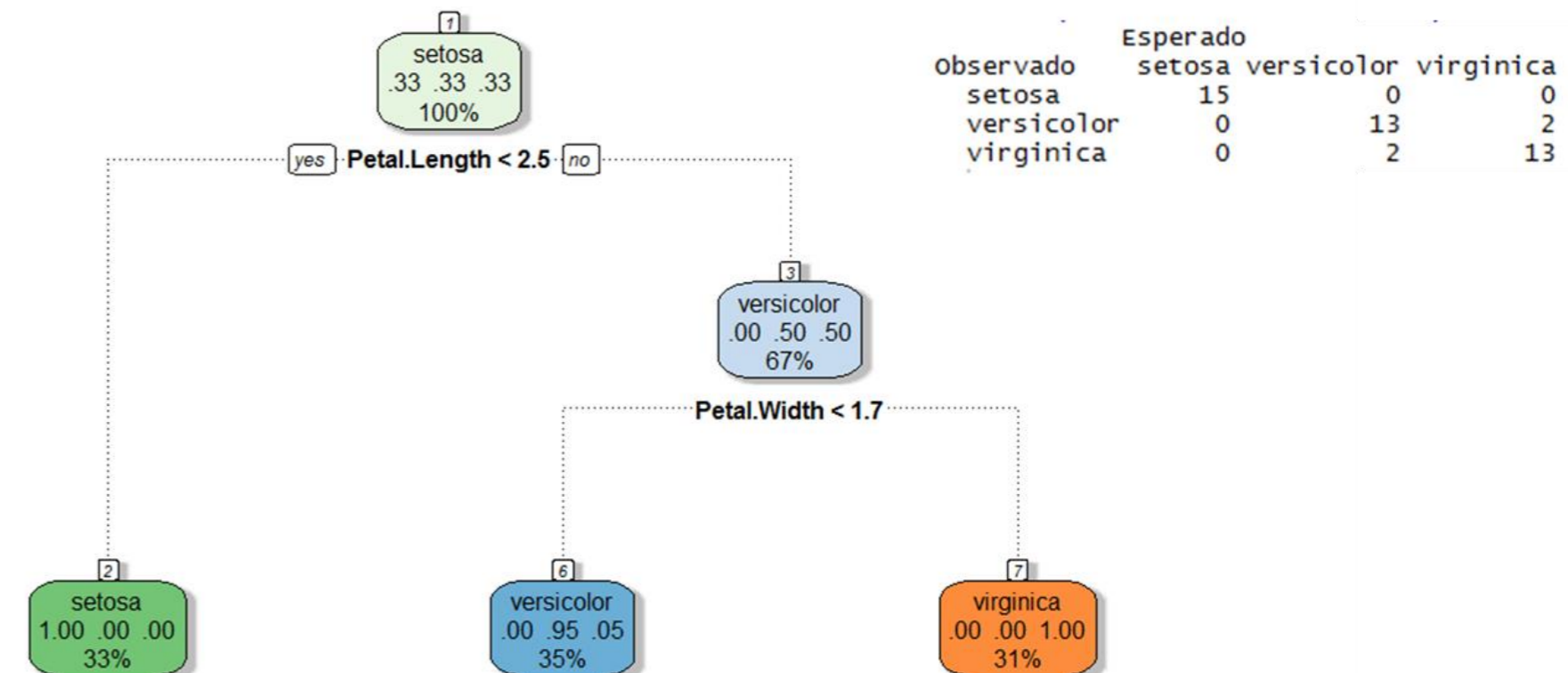


Fisher's Iris Data

Largo de sépalo ⇅	Ancho de sépalo ⇅	Largo de pétalo ⇅	Ancho de pétalo ⇅	Especies ⇅
5.0	2.0	3.5	1.0	<i>I. versicolor</i>
6.0	2.2	4.0	1.0	<i>I. versicolor</i>
6.0	2.2	5.0	1.5	<i>I. virginica</i>
6.2	2.2	4.5	1.5	<i>I. versicolor</i>
4.5	2.3	1.3	0.3	<i>I. setosa</i>
5.0	2.3	3.3	1.0	<i>I. versicolor</i>
5.5	2.3	4.0	1.3	<i>I. versicolor</i>
6.3	2.3	4.4	1.3	<i>I. versicolor</i>
4.9	2.4	3.3	1.0	<i>I. versicolor</i>
5.5	2.4	3.7	1.0	<i>I. versicolor</i>
5.5	2.4	3.8	1.1	<i>I. versicolor</i>
4.9	2.5	4.5	1.7	<i>I. virginica</i>
5.1	2.5	3.0	1.1	<i>I. versicolor</i>
5.5	2.5	4.0	1.3	<i>I. versicolor</i>
5.6	2.5	3.9	1.1	<i>I. versicolor</i>
5.7	2.5	5.0	2.0	<i>I. virginica</i>
6.3	2.5	4.9	1.5	<i>I. versicolor</i>
6.3	2.5	5.0	1.9	<i>I. virginica</i>
6.7	2.5	5.8	1.8	<i>I. virginica</i>
5.5	2.6	4.4	1.2	<i>I. versicolor</i>
5.7	2.6	3.5	1.0	<i>I. versicolor</i>
5.8	2.6	4.0	1.2	<i>I. versicolor</i>
6.1	2.6	5.6	1.4	<i>I. virginica</i>



- Se utilizan para tomar decisiones o para el uso de técnicas de clasificación
- El cálculo se basa en algoritmo de probar cortes y ver qué combinación es la ganadora



IV – Reconocimiento de patrones

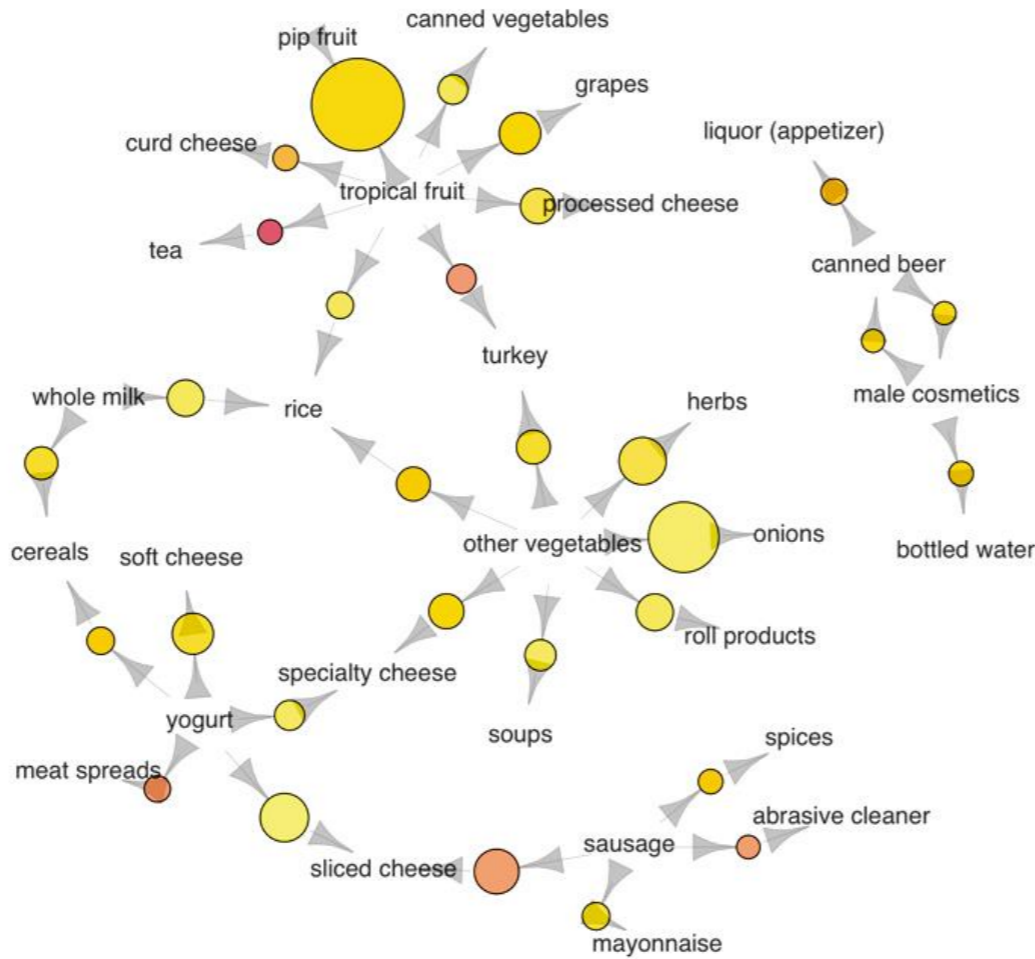
El mapa de Técnicas



$\text{Support} \{\text{🍎}\} = \frac{4}{8}$

Transaction 1	🍎 🍺 🥛 🍗
Transaction 2	🍎 🍺 🥛
Transaction 3	🍎 🍺
Transaction 4	🍎 🍏
Transaction 5	🍼 🍺 🥛 🍗
Transaction 6	🍼 🍺 🥛
Transaction 7	🍼 🍺
Transaction 8	🍼 🍏

Table 1. Example Transactions



Objetivo:

A partir de un histórico de datos podemos observar asociaciones entre variables.

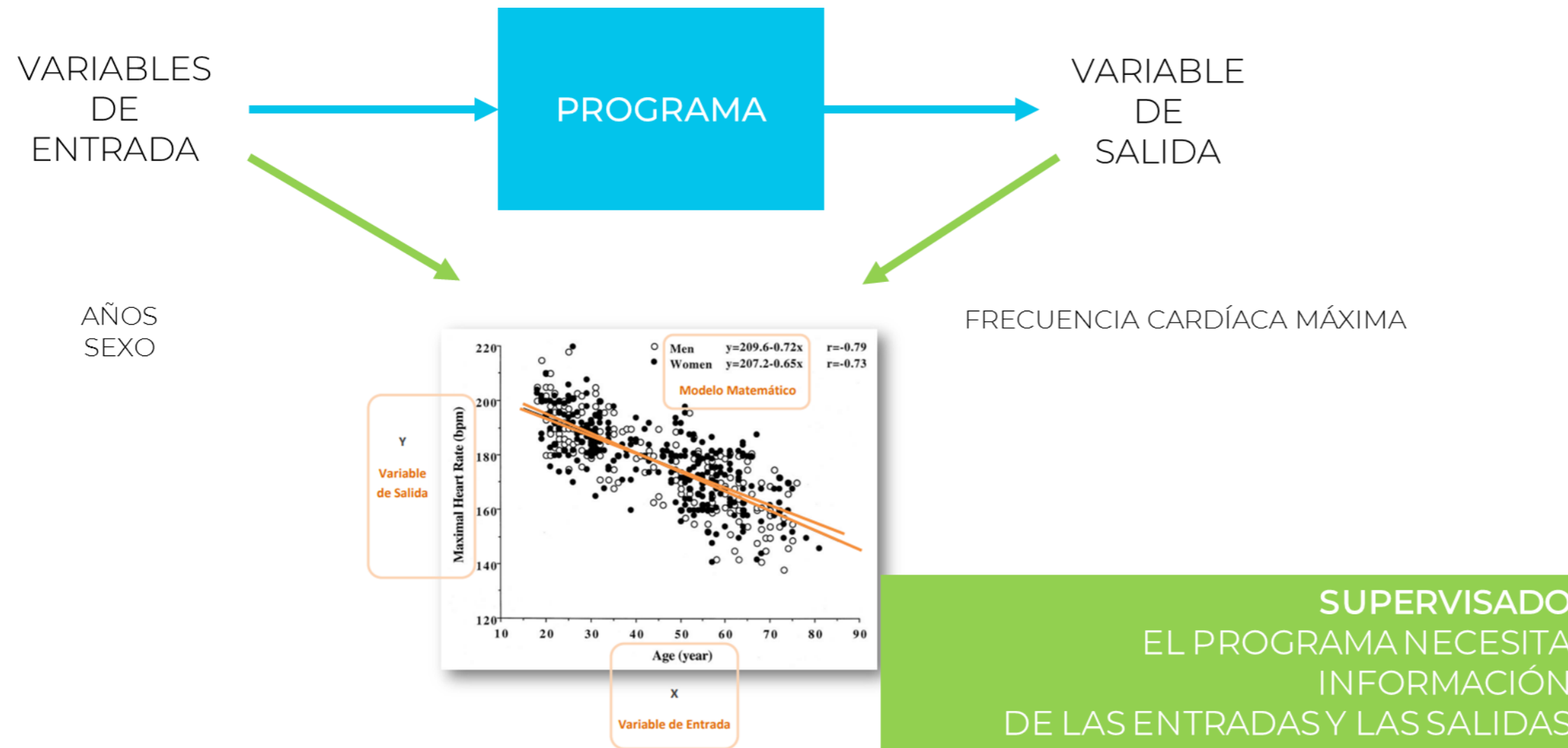
“Según el histórico de compras en supermercados se ha observado que muchos de los consumidores compran salsichas y queso en lonchas”

IV y V- Machine Learning

El mapa de Técnicas



- Predictivas- Supervisadas



Se pueden dividir en:

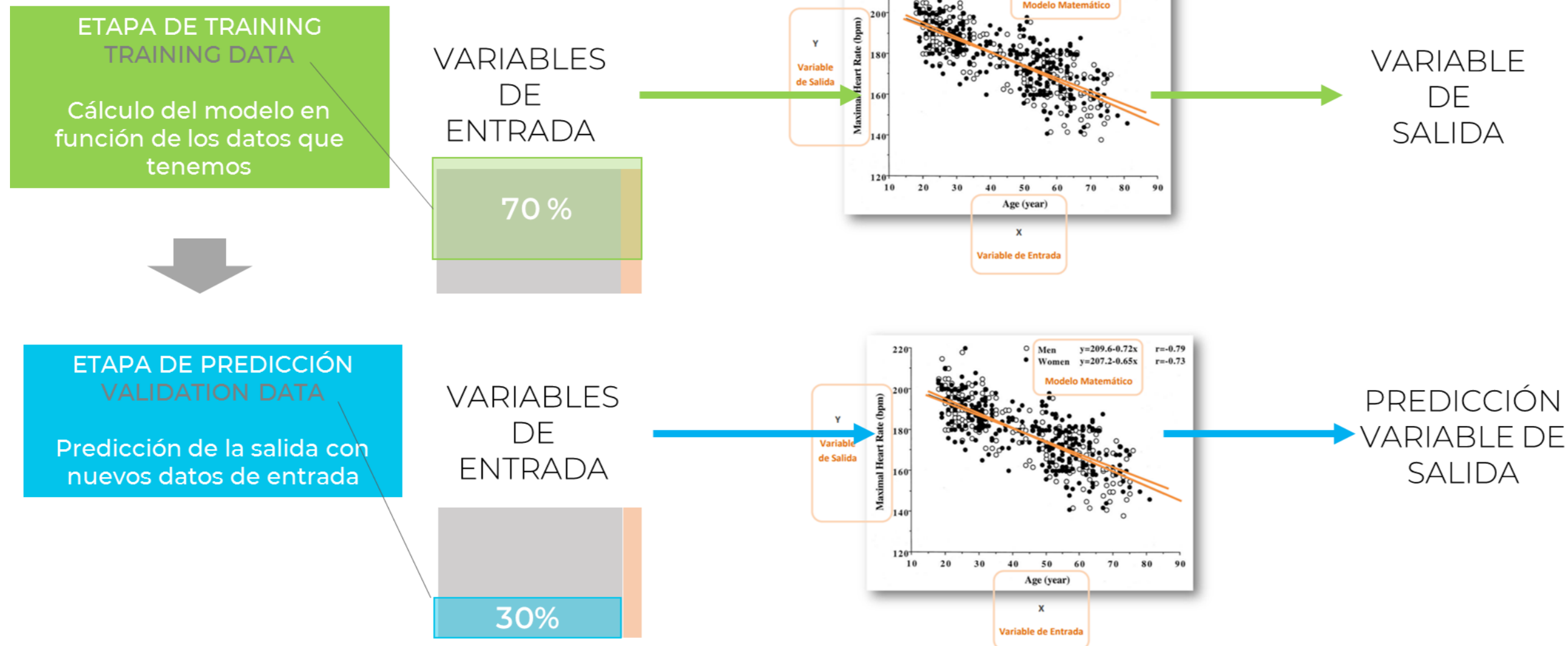
- Supervisadas – reconocimiento de patrones
- No supervisadas - predictivas**

IV y V- Machine Learning

El mapa de Técnicas



- Supervisadas – necesitamos entrenar el modelo

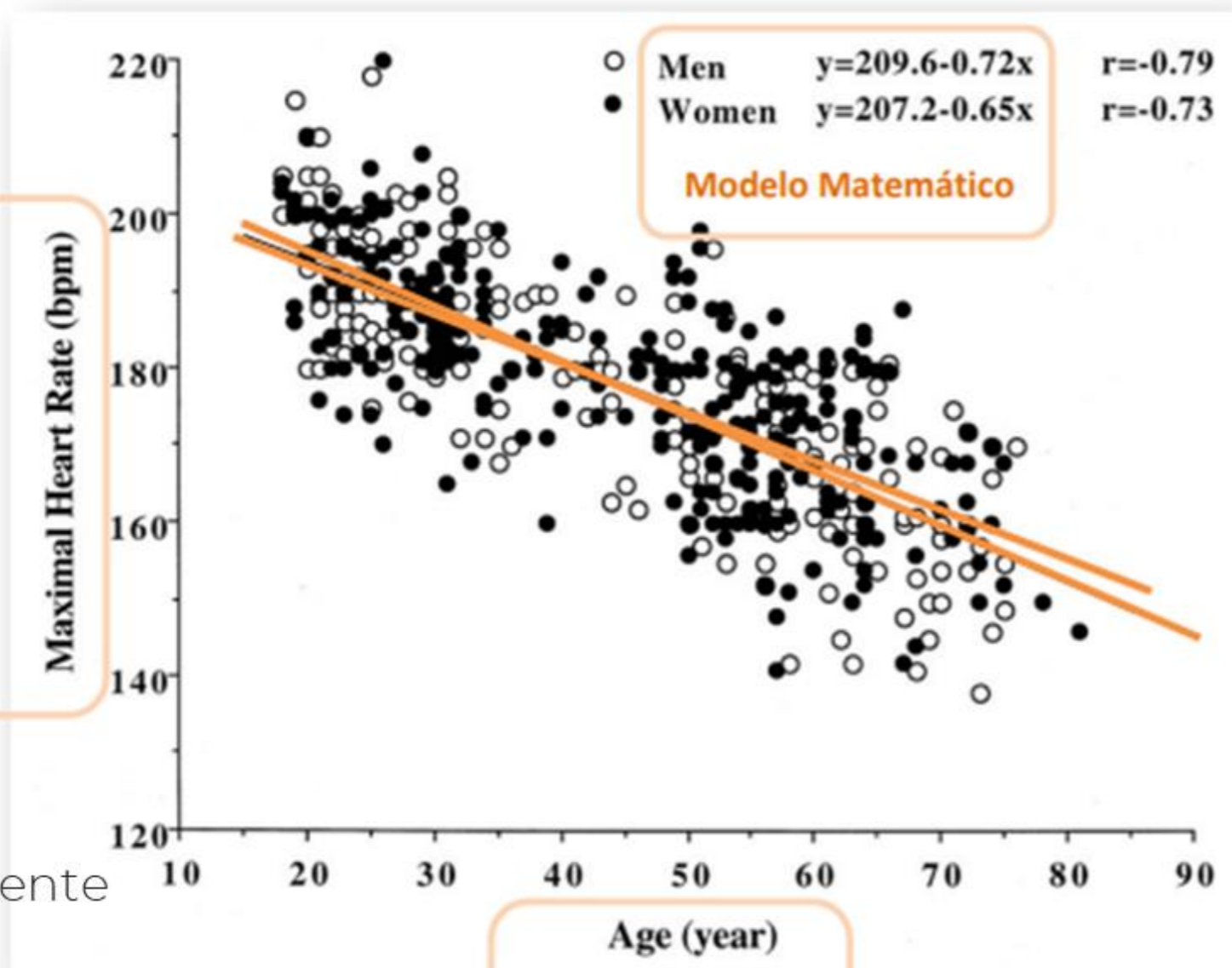


Se pueden dividir en:

- Supervisadas – reconocimiento de patrones
- **No supervisadas - predictivas**

V-Predictivas

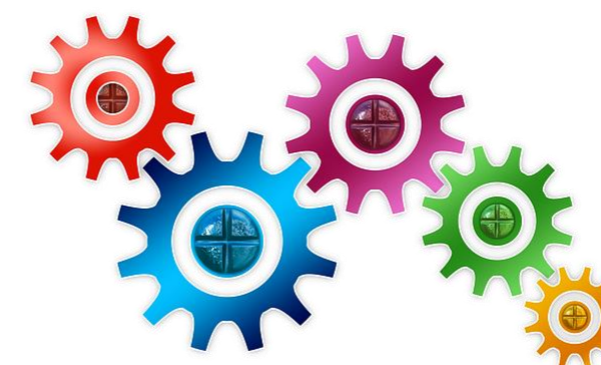
El mapa de Técnicas



Variable
Respuesta
O dependiente

Variable
de Entrada

Variable
Respuesta
O independiente



X - EDAD
Variable de
Entrada

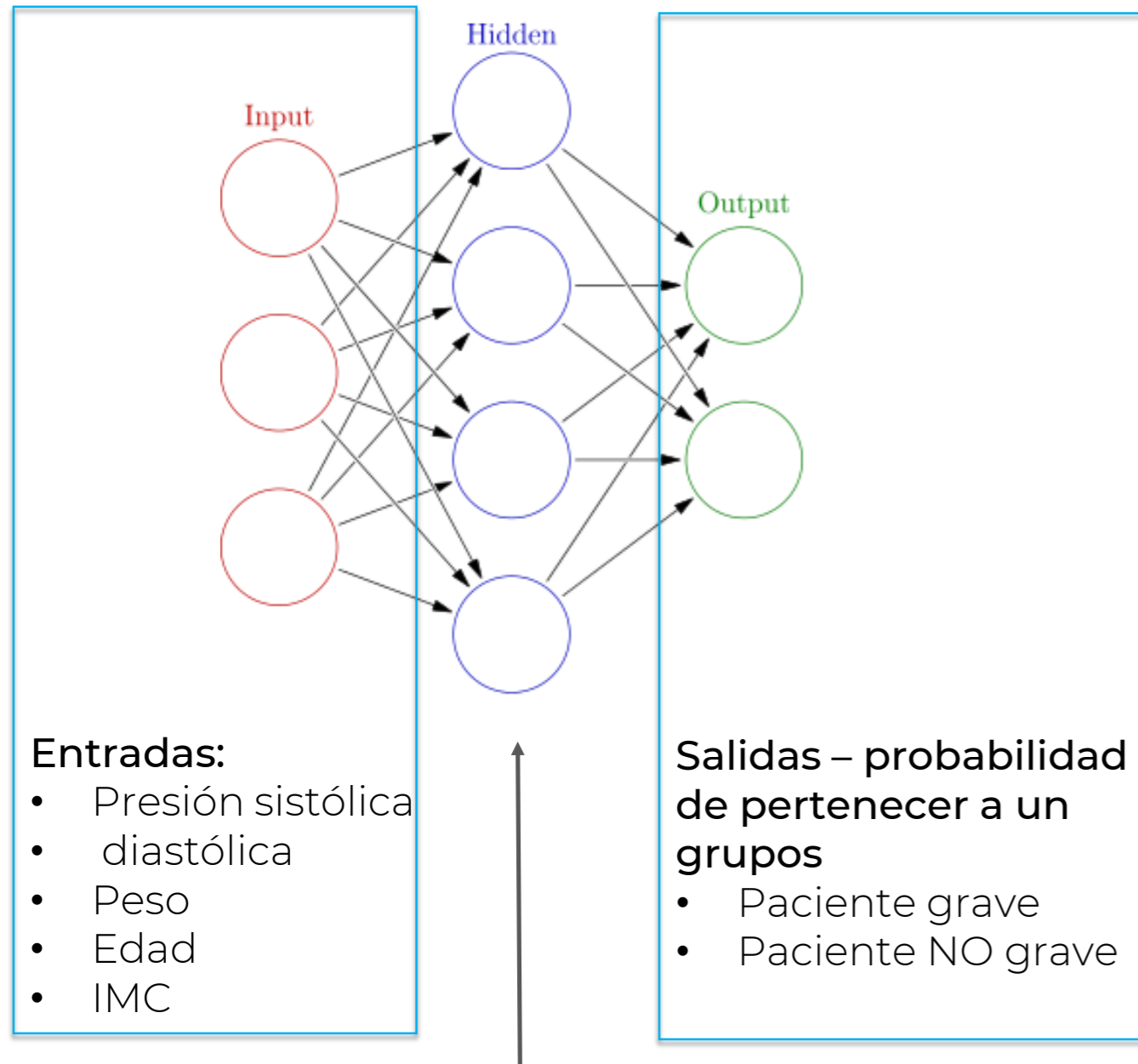
Men $y=209.6-0.72x$
Women $y=207.2-0.65x$
Modelo Matemático

Y - FREC.
CARDÍACA MÁX
Variable de
Salida

Variable Salida – Y	Frecuencia Cardíaca Máxima
Variable Entrada – X	Edad
Función matemática	<p>Men $y=209.6-0.72x$</p> <p>Women $y=207.2-0.65x$</p>

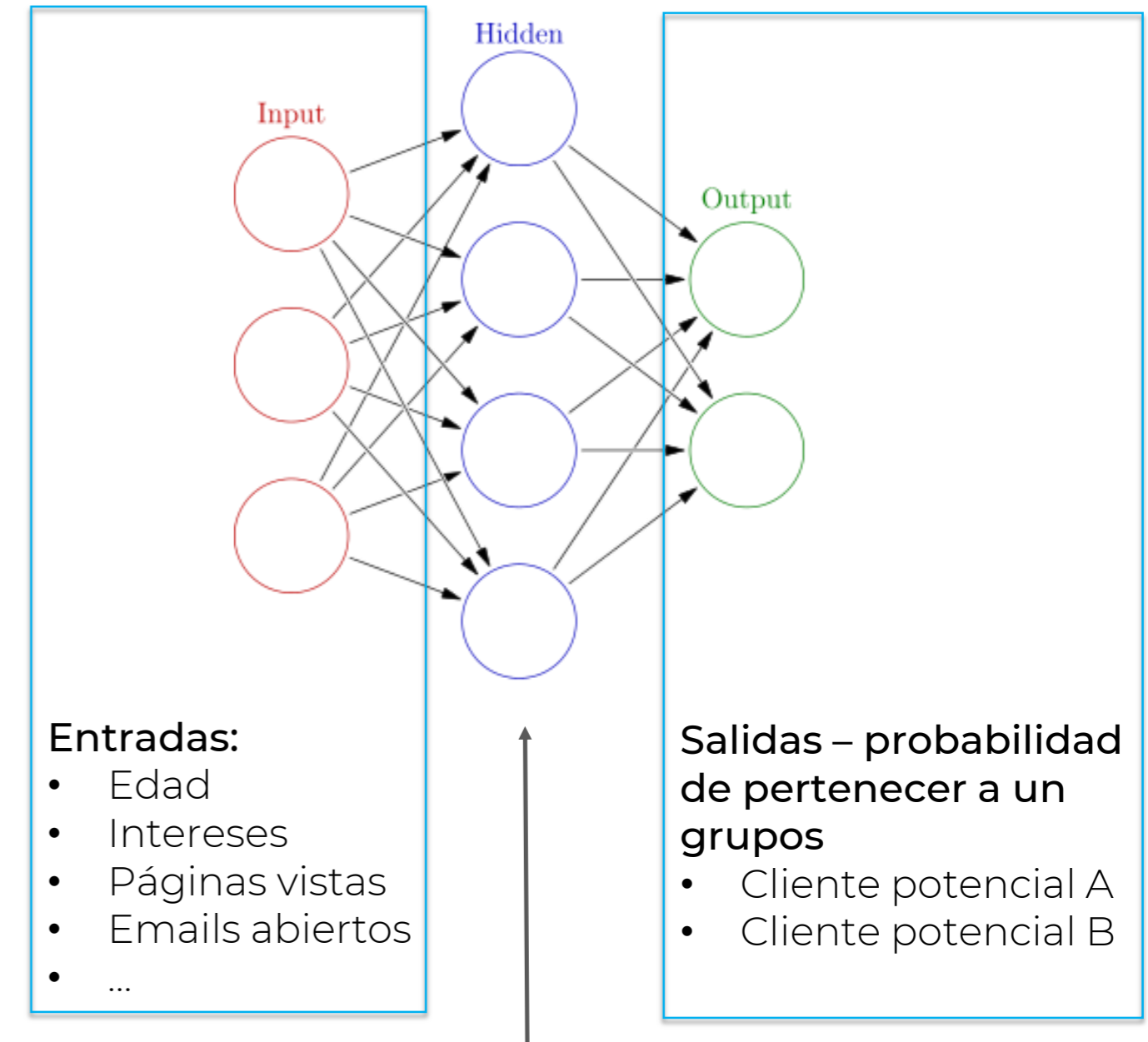
V - Predictivas

El mapa de Técnicas



Función Matemática
que explica los datos

Clasificador Red Neuronal



Función Matemática
que explica los datos

TAKE AWAY - El mapa de Técnicas

- I- Comparativas
- II- Relacionales
- III- Causales
- IV- Reconocimiento de Patrones
- V- Predictivas





El software de datos
que mejor se adapta
para tu caso

... y cuál de las opciones te
puedo recomendar





CASO A: ESTUDIOS CIENTÍFICOS

(SALUD, BIOLOGÍA, SOCIOLOGÍA, ESTUDIOS DE MERCADOS...)

- Relativamente pocos datos
- Análisis simples sin necesidad de modelos personalizados
- No es necesario la iteración o replicar el cálculo muchas veces
- La mayoría de artículos científicos

Software Estadístico a base de clics



CASO GENERAL: reporting rápido y resumen de datos

- Representar gráficos
- Crear tablas resumen
- Pequeños cálculos
- Informes y reporting rápidos

HOJA DE CÁLCULO –
HERRAMIENTAS DE
REPORTING



CASO B: CIENCIA DE DATOS

- Puede trabajar con grandes volúmenes de datos
- lectura de datos sea robusta y automatizada(no se introduce a mano)
- Análisis personalizados
- Necesidad de automatizarlos y robustez en el cálculo
- Flexibilidad

Software Estadístico a base de comandos (programación)



Opciones recomendables por aplicación

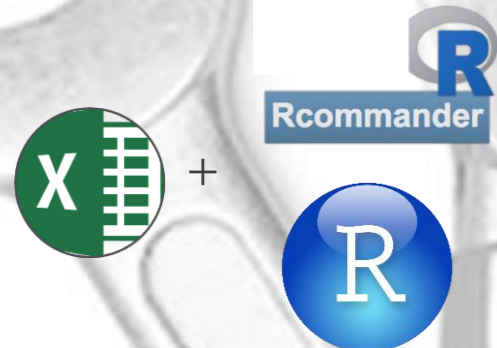
CASO GENERAL: reporting rápido y resumen de datos

HOJA DE CÁLCULO –
HERRAMIENTAS DE
REPORTING

Opción 1



Opción 2



CASO A: ESTUDIOS CIENTÍFICOS

(SALUD, BIOLOGÍA, SOCIOLOGÍA,
ESTUDIOS DE MERCADOS...)

Software Estadístico a base de
clics

Intuitivo

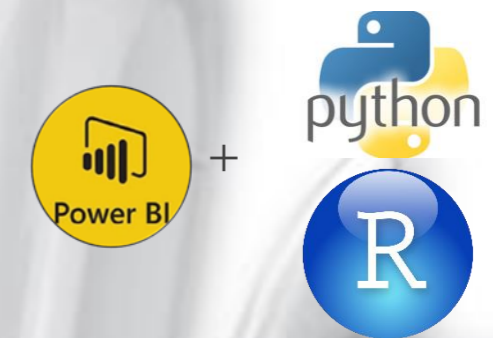


CASO B: CIENCIA DE DATOS

(SALUD, BIOLOGÍA, SOCIOLOGÍA,
ESTUDIOS DE MERCADOS...)

Software Estadístico a base de
comandos (programación)

Opción Avanzada



Un Sistema para tratar
tus datos paso a paso
con seguridad y
confianza

y ahorrar miles de horas de tu
tiempo





TÉCNICAS DE
ANÁLISIS DE
DATOS



SISTEMA

BASES
ESTADÍSTICAS



TÉCNICAS DE
ANÁLISIS DE
DATOS



SISTEMA

BASES
ESTADÍSTICAS

Etapas de un proyecto con Datos

LA ETAPA DE ANÁLISIS DE DATOS (ETAPA4 + ETAPA5 + ETAPA 6) ESTÁ DIRECTAMENTE RELACIONADA CON UN SÓLO OBJETIVO ALINEADO CON EL PROBLEMA DEFINIDO EN LA ETAPA 1

ETAPA1: EL PROBLEMA

- Identificar el problema
- Definir un objetivo de estudio

ETAPA2: RECOLECCIÓN

- Estrategia de muestreo
- Diseño experimental

ETAPA3: LIMPIEZA

- Igualar formatos
- Eliminar información errónea

ETAPA4: EXPLORACIÓN

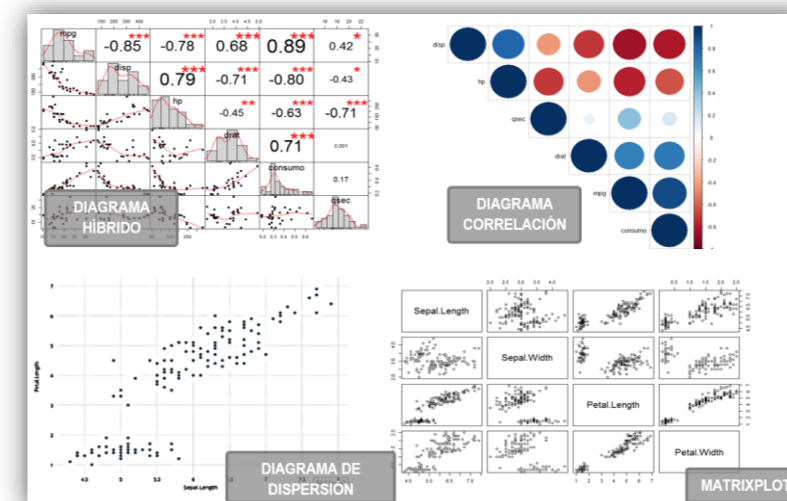
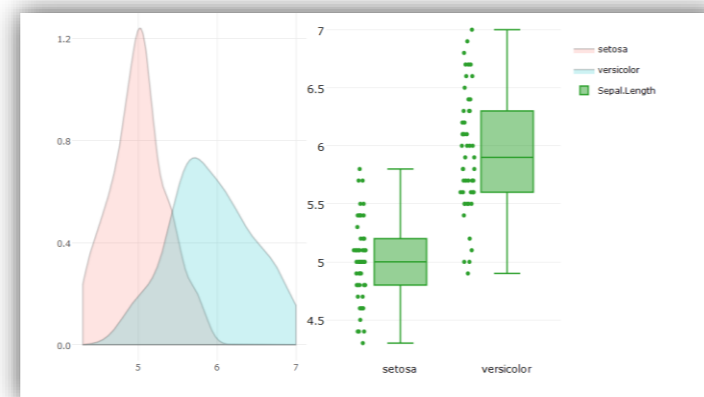
- Descripción de variables una a una
- Buscar diferencias entre grupos
- Forma de la distribución y proporciones

ETAPA5: ANÁLISIS

- Comparación de grupos
- Comparación de proporciones
- Modelos predictivos
- Relación entre variables

ETAPA6: CONCLUSIÓN

- Informe de resultados
- Interpretación y síntesis de resultados
- Objeciones y siguientes pasos



Contraste de Hipótesis
Correlación
Comparación de medias
Modelos predictivos
Clustering

Tabla con los valores de centralidad y dispersión (mediana [cuartil3] y el p-valor de la prueba Wilcoxon para comparar grupos)

	Setosa	Versicolor	p-valor	Setosa	Virginica	p-valor	Versicolor	Virginica	p-valor
Sepal.Length	5.00 [4.80 - 5.20]	5.90 [5.60 - 6.30]	8.35E-11	5.00 [4.80 - 5.20]	6.50 [6.20 - 6.90]	2.20E-16	5.90 [5.60 - 6.30]	6.50 [6.20 - 6.90]	5.87E-04
Sepal.Width	3.40 [3.20 - 3.68]	2.80 [2.53 - 3.00]	2.14E-10	3.40 [3.20 - 3.68]	3.00 [2.80 - 3.18]	7.10E-09	2.80 [2.53 - 3.00]	3.00 [2.80 - 3.18]	0.004572
Petal.Length	1.50 [1.40 - 1.58]	4.35 [4.00 - 4.60]	2.20E-16	1.50 [1.40 - 1.58]	5.55 [5.10 - 5.88]	2.20E-16	5.00 [4.80 - 5.20]	5.55 [5.10 - 5.88]	2.20E-16
Petal.Width	0.20 [0.20 - 0.30]	1.30 [1.20 - 1.50]	2.20E-16	0.20 [0.20 - 0.30]	2.00 [1.80 - 2.30]	2.20E-16	1.30 [1.20 - 1.50]	2.00 [1.80 - 2.30]	2.20E-16

Setosa vs Versicolor	Sepal.Length	8.35E-11
Setosa vs Versicolor	Sepal.Width	2.14E-10
Setosa vs Versicolor	Petal.Length	2.20E-16
Setosa vs Versicolor	Petal.Width	2.20E-16
Setosa vs Virginica	Sepal.Length	2.20E-16
Setosa vs Virginica	Sepal.Width	7.10E-09
Setosa vs Virginica	Petal.Length	2.20E-16
Setosa vs Virginica	Petal.Width	2.20E-16
Versicolor vs Virginica	Sepal.Length	5.87E-04
Versicolor vs Virginica	Sepal.Width	0.004572
Versicolor vs Virginica	Petal.Length	2.20E-16
Versicolor vs Virginica	Petal.Width	2.20E-16

Diferencias significativas entre todas las especies y todas las variables.
Las características que mejor apuntan a diferenciar poblaciones de especies son las dimensiones de los pétalos para todas las especies.

Las peor característica es el ancho del sépalo.

Y la longitud del pétalo funciona muy bien para diferenciar setosa del resto. No tan bien para diferenciar versicolor de virginica.

METODOLOGÍA

En la práctica ...

Herramientas y softwares

Largo de sépalo	Ancho de sépalo	Largo de pétalo	Ancho de pétalo	Especies
5.0	2.0	3.5	1.0	<i>I. versicolor</i>
6.0	2.2	4.0	1.0	<i>I. versicolor</i>
6.0	2.2	5.0	1.5	<i>I. virginica</i>
6.2	2.2	4.5	1.5	<i>I. versicolor</i>
4.5	2.3	1.3	0.3	<i>I. setosa</i>
5.0	2.3	3.3	1.0	<i>I. versicolor</i>
5.5	2.3	4.0	1.3	<i>I. versicolor</i>
6.3	2.3	4.4	1.3	<i>I. versicolor</i>
4.9	2.4	3.3	1.0	<i>I. versicolor</i>
5.5	2.4	3.7	1.0	<i>I. versicolor</i>
5.5	2.4	3.8	1.1	<i>I. versicolor</i>
4.9	2.5	4.5	1.7	<i>I. virginica</i>
5.1	2.5	3.0	1.1	<i>I. versicolor</i>
5.5	2.5	4.0	1.3	<i>I. versicolor</i>
5.6	2.5	3.9	1.1	<i>I. versicolor</i>
5.7	2.5	5.0	2.0	<i>I. versicolor</i>
6.3	2.5	4.9	1.5	<i>I. versicolor</i>
6.3	2.5	5.0	1.9	<i>I. versicolor</i>
6.7	2.5	5.8	1.8	<i>I. versicolor</i>
5.5	2.6	4.4	1.2	<i>I. versicolor</i>
5.7	2.6	3.3	1.0	<i>I. versicolor</i>
5.8	2.6	4.0	1.2	<i>I. versicolor</i>
5.1	2.6	5.6	1.4	<i>I. virginica</i>

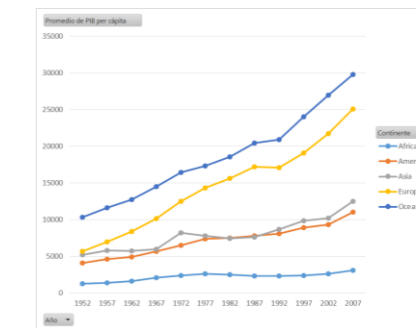
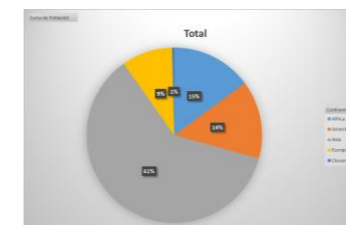
Paso 1 – Enfocar El Estudio

1. Entender tu tabla de datos
2. Definir el objetivo de análisis
3. Leer e interpretar tus base de datos

Nivel I - Descriptivo

Paso 2 – Explorar tu base de datos

1. Crear gráficos
2. Interpretarlos
3. Extraer las primeras respuestas



Nivel II y III – Inferencia

Paso 3 – Analizar los Datos

1. Responder a tus preguntas
2. Interpretar los resultados
3. Conclusiones

Paso 4 – Interpretar y Comunicar tus resultados

1. Responder a tus objetivos
2. Un esquema efectivo de comunicación

*Repeticiones
Definir Objetivos
de + fáciles a +
difíciles*

País	Año	Población	Continente	Esperanza de Vida	PIB per cápita
Afghanistan	1952	8425333	Asia	28.801	779.4453145
Afghanistan	1957	9240934	Asia	30.332	820.8530296
Afghanistan	1962	10267083	Asia	31.997	853.10071
Afghanistan	1967	11537966	Asia	34.02	836.1971382
Afghanistan	1972	13079460	Asia	36.088	739.9811058
Afghanistan	1977	14880372	Asia	38.438	786.11336
Afghanistan	1982	12881816	Asia	39.854	978.0114388
Afghanistan	1987	13867957	Asia	40.822	852.3959448
Afghanistan	1992	16317921	Asia	41.674	649.3413952
Afghanistan	1997	22227415	Asia	41.763	635.341351
Afghanistan	2002	25268405	Asia	42.129	726.7340548
Afghanistan	2007	31889923	Asia	43.828	974.5803384
Albania	1952	1282697	Europe	55.23	1601.056136
Albania	1957	1476505	Europe	59.28	1942.284244
Albania	1962	1728137	Europe	64.82	2312.888958
Albania	1967	1984060	Europe	66.22	2760.196931
Albania	1972	2263554	Europe	67.69	3313.422188
Albania	1977	2509048	Europe	68.93	3533.00391
Albania	1982	2780097	Europe	70.42	3630.880722
Albania	1987	3075321	Europe	72	3738.932735
Albania	1992	3326498	Europe	71.581	2497.437901
Albania	1997	3428038	Europe	72.95	3193.054604
Albania	2002	3508512	Europe	75.651	4604.211737
Albania	2007	3600523	Europe	76.423	5937.029526
Algeria	1952	9279525	Africa	43.077	2449.008185
Algeria	1957	10270856	Africa	45.685	3013.976023
Algeria	1962	11000948	Africa	48.303	2550.81688
Algeria	1967	12760499	Africa	51.407	3246.991771
Algeria	1972	14760787	Africa	54.518	4182.663766
Algeria	1977	17152804	Africa	58.014	4910.416756
Algeria	1982	20033753	Africa	61.368	5745.160213
Algeria	1987	23254956	Africa	65.799	5681.358539
Algeria	1992	26298373	Africa	67.744	5023.216647
Algeria	1997	29072015	Africa	69.152	4797.295051
Algeria	2002	31287142	Africa	70.994	5288.040382
Algeria	2007	33333216	Africa	72.301	6223.367465
Angola	1952	4232095	Africa	30.015	3520.610273
Angola	1957	4561361	Africa	31.999	3827.940465
Angola	1962	4826015	Africa	34	4269.276742



Paso 1 – Enfocar El Estudio

1. Entender tu tabla de datos
2. Definir el objetivo de análisis
3. Leer e interpretar tus base de datos

Variables

País – categórica

Año – Ordinal

Población – numérica discreta

Continente – Categórica

Esperanza de Vida – numérica continua

PIB per cápita – numérica contínua

Objetivos

Evolución de la población en los diferentes continentes

Evolución de la esperanza de vida en los diferentes continentes

Evolución del PIB per cápita en los diferentes continentes

Buscar la relación entre la Esperanza de vida y el PIB per cápita

1. Crear gráficos
2. Interpretarlos
3. Extraer las primeras respuestas

Objetivos

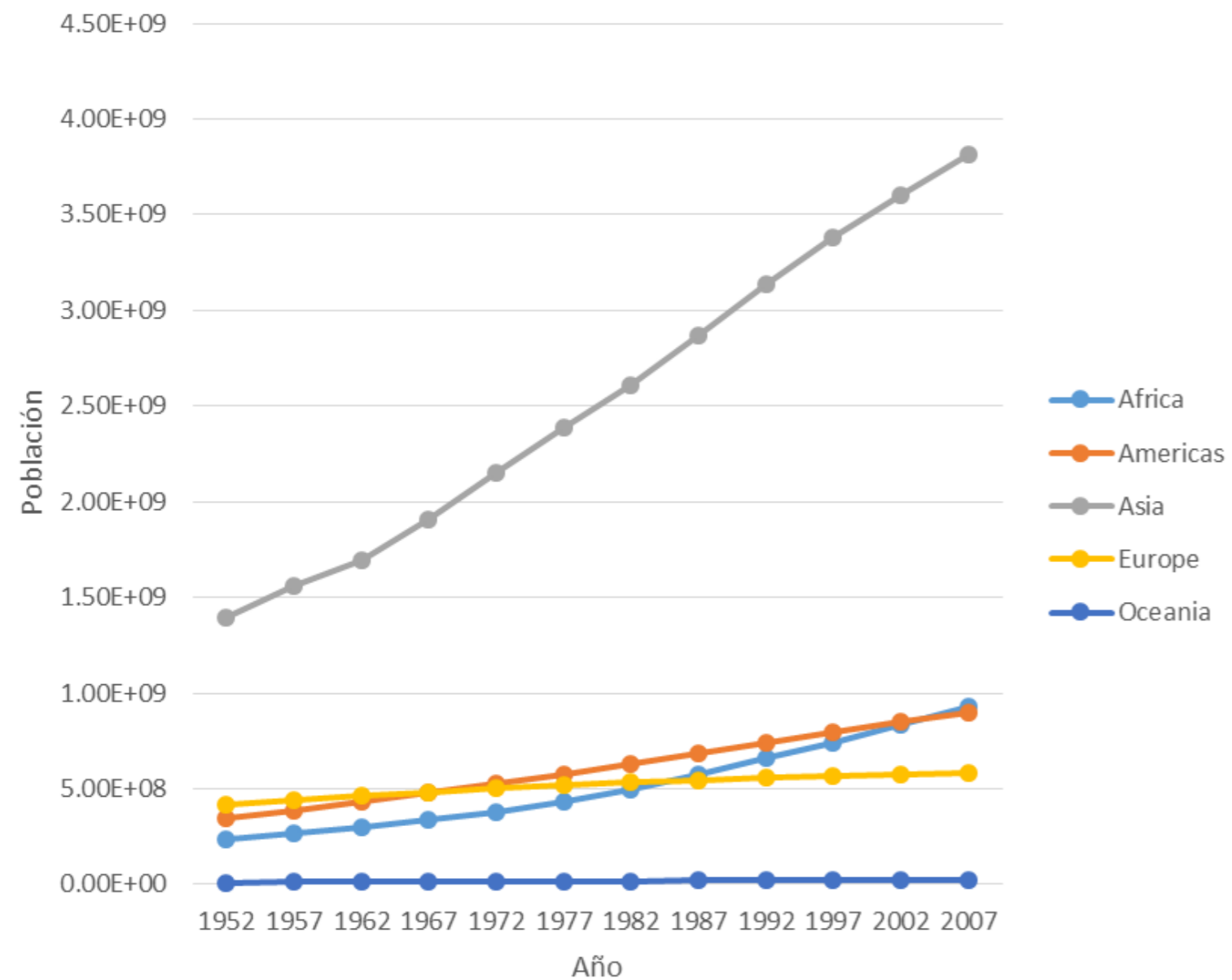
Evolución de la población en los diferentes continentes

Evolución de la esperanza de vida en los diferentes continentes

Evolución del PIB per cápita en los diferentes continentes

Buscar la relación entre la Esperanza de vida y el PIB per cápita

Evolución Población



1. Crear gráficos
2. Interpretarlos
3. Extraer las primeras respuestas

Objetivos

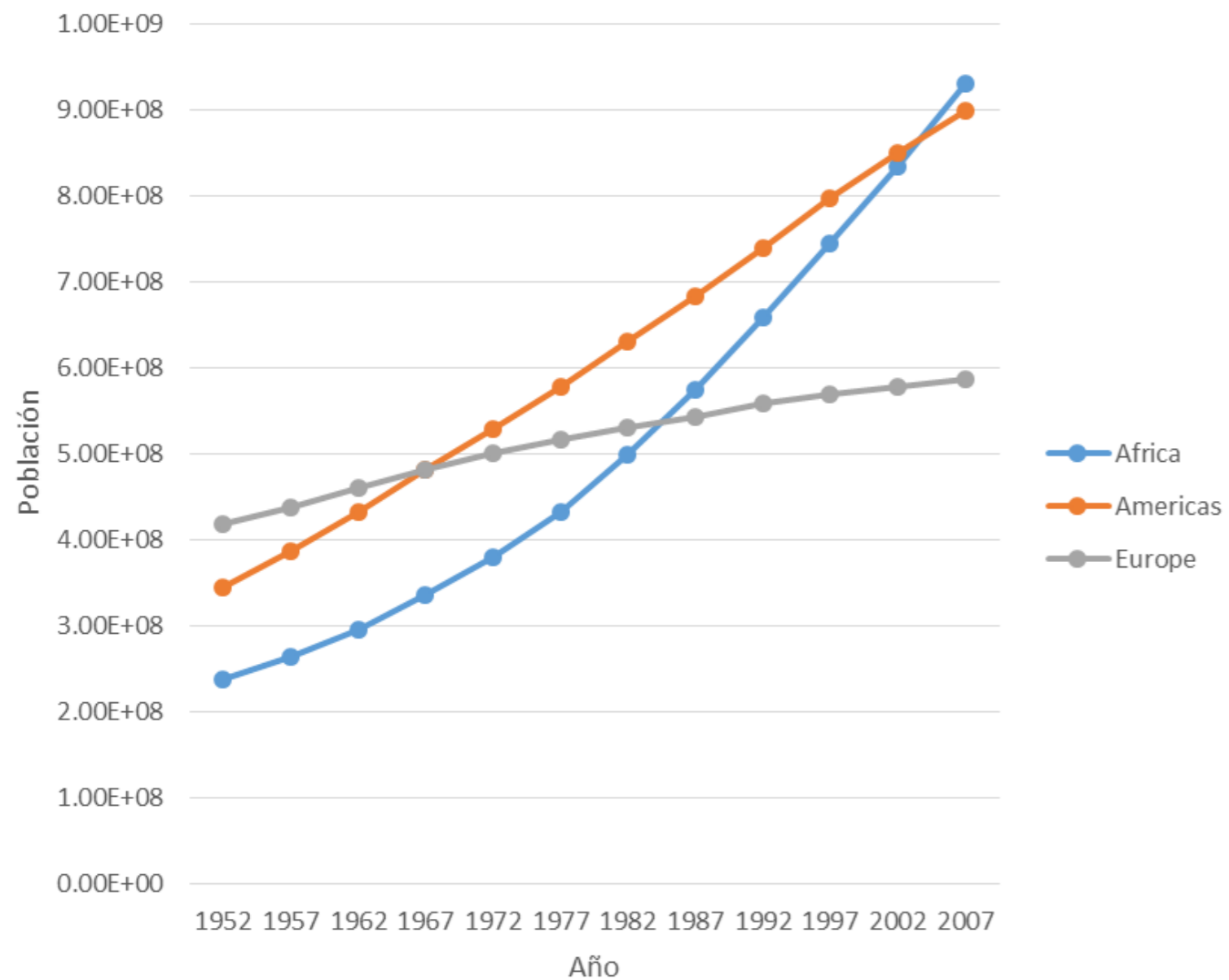
Evolución de la población en los diferentes continentes

Evolución de la esperanza de vida en los diferentes continentes

Evolución del PIB per cápita en los diferentes continentes

Buscar la relación entre la Esperanza de vida y el PIB per cápita

Evolución Población



2

Paso 2 – Explorar tu base de datos

1. Crear gráficos
2. Interpretarlos
3. Extraer las primeras respuestas

Objetivos

Evolución de la población en los diferentes continentes

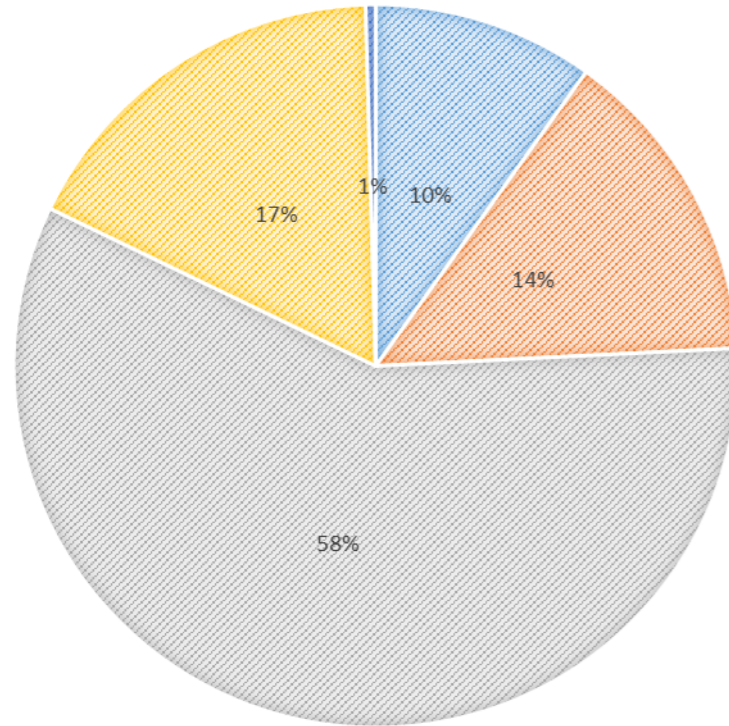
Evolución de la esperanza de vida en los diferentes continentes

Evolución del PIB per cápita en los diferentes continentes

Buscar la relación entre la Esperanza de vida y el PIB per cápita

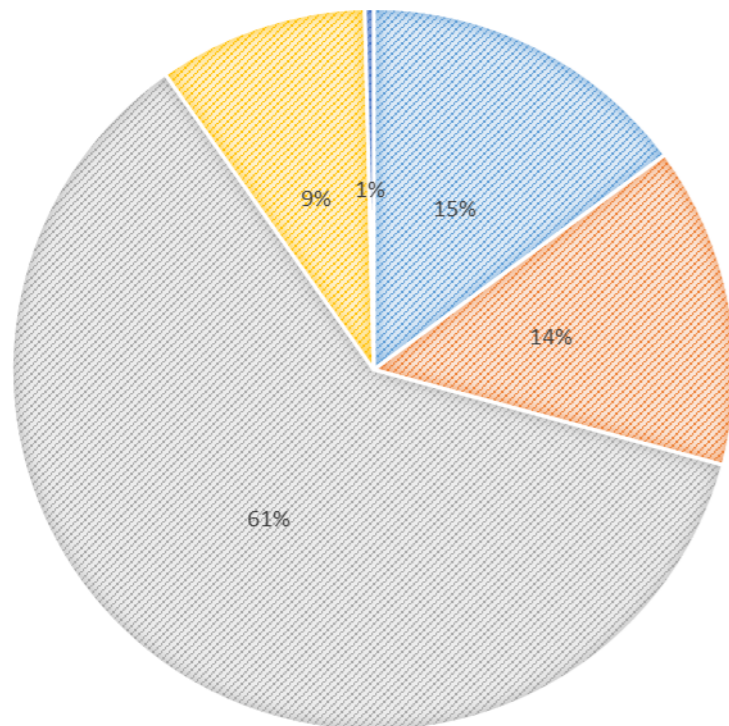
DISTRIBUCIÓN POBLACIÓN 1952

■ Africa ■ Americas ■ Asia ■ Europe ■ Oceania



DISTRIBUCIÓN POBLACIÓN 2007

■ Africa ■ Americas ■ Asia ■ Europe ■ Oceania



1. Crear gráficos
2. Interpretarlos
3. Extraer las primeras respuestas

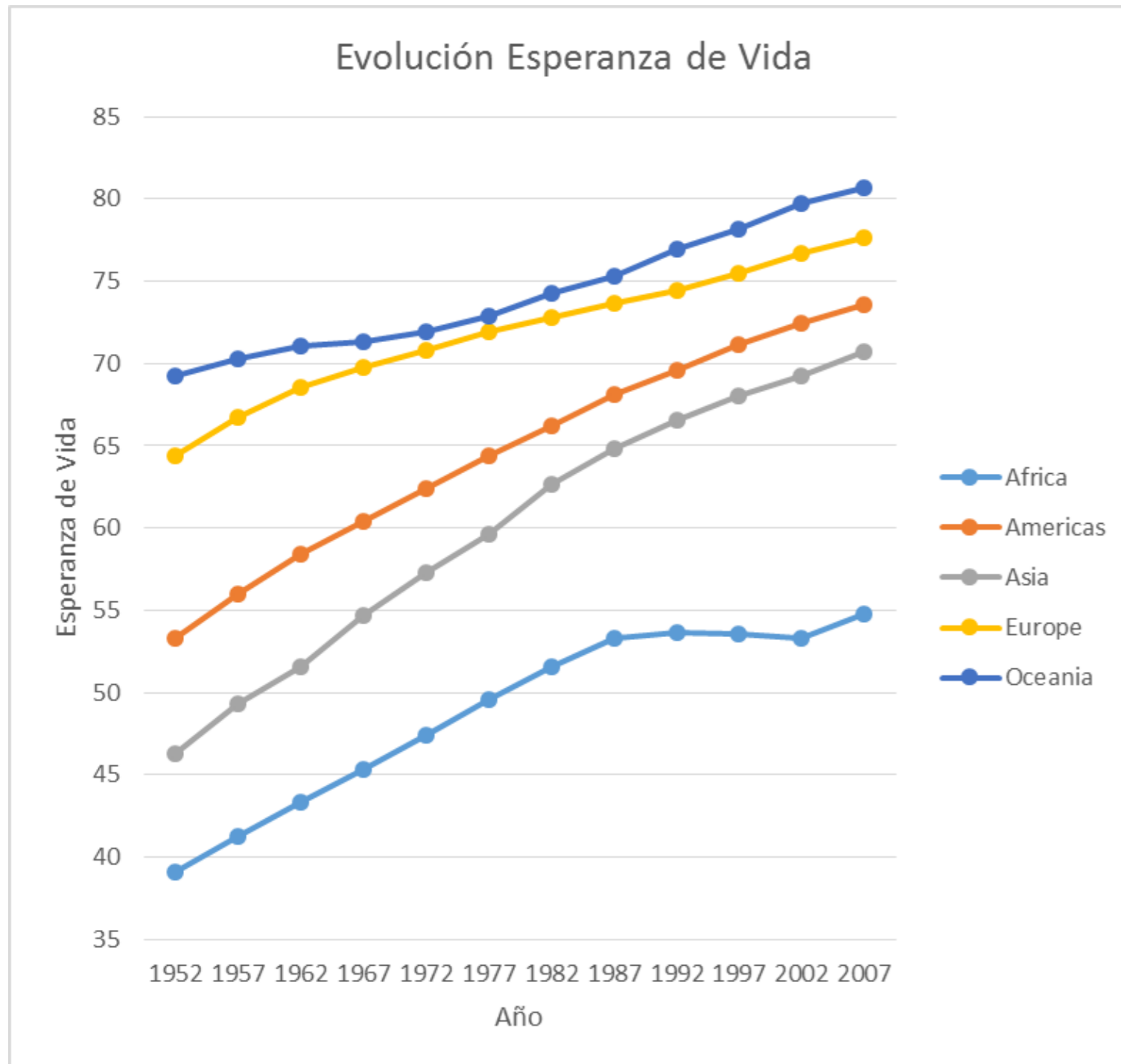
Objetivos

Evolución de la población en los diferentes continentes

Evolución de la esperanza de vida en los diferentes continentes

Evolución del PIB per cápita en los diferentes continentes

Buscar la relación entre la Esperanza de vida y el PIB per cápita



1. Crear gráficos
2. Interpretarlos
3. Extraer las primeras respuestas

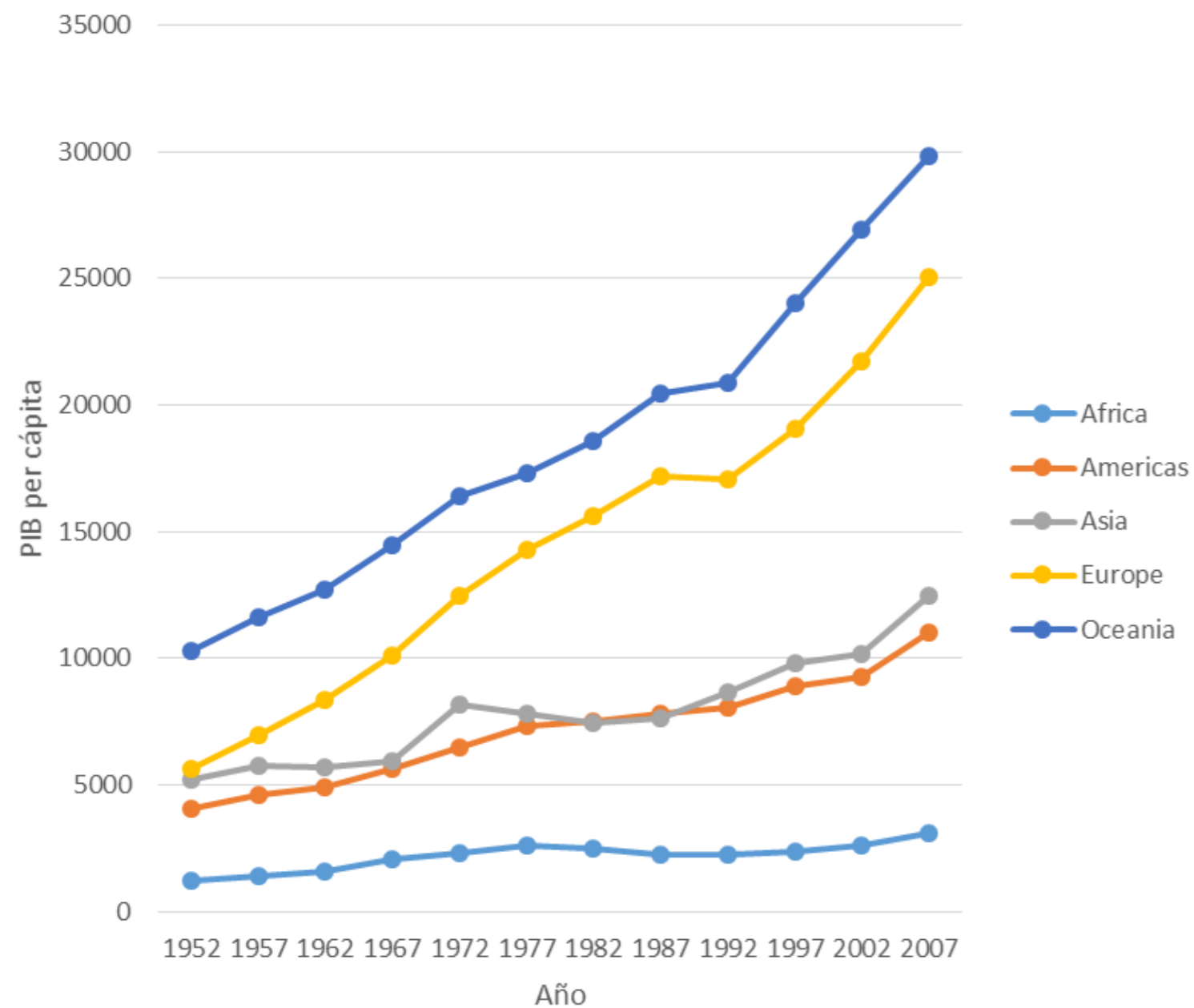
Objetivos

Evolución de la población en los diferentes continentes
Evolución de la esperanza de vida en los diferentes continentes

[Evolución del PIB per cápita en los diferentes continentes](#)

Buscar la relación entre la Esperanza de vida y el PIB per cápita

Evolución PIB per cápita



2

Paso 2 – Explorar tu base de datos

1. Crear gráficos
2. Interpretarlos
3. Extraer las primeras respuestas

Objetivos

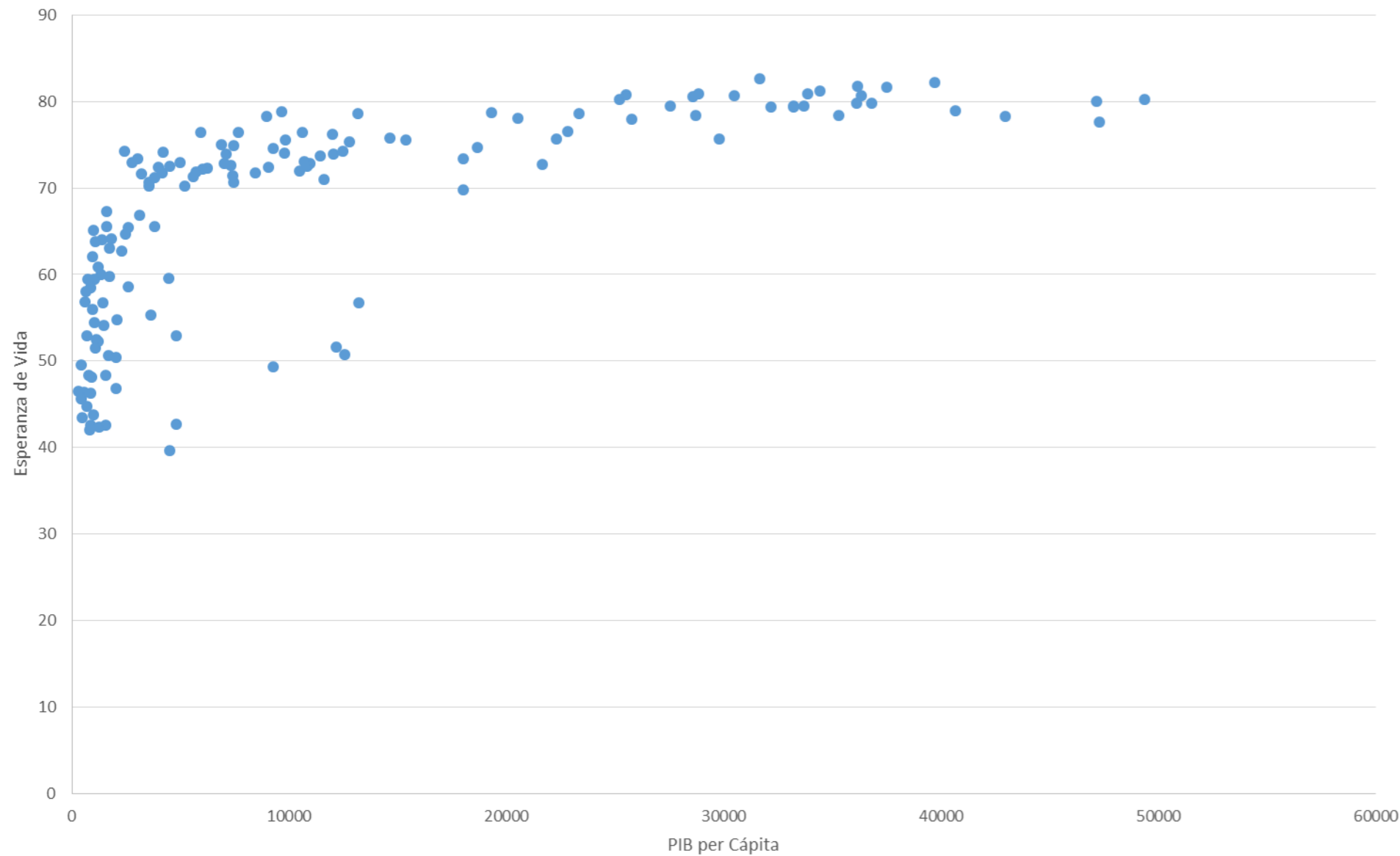
Evolución de la población en los diferentes continentes

Evolución de la esperanza de vida en los diferentes continentes

Evolución del PIB per cápita en los diferentes continentes

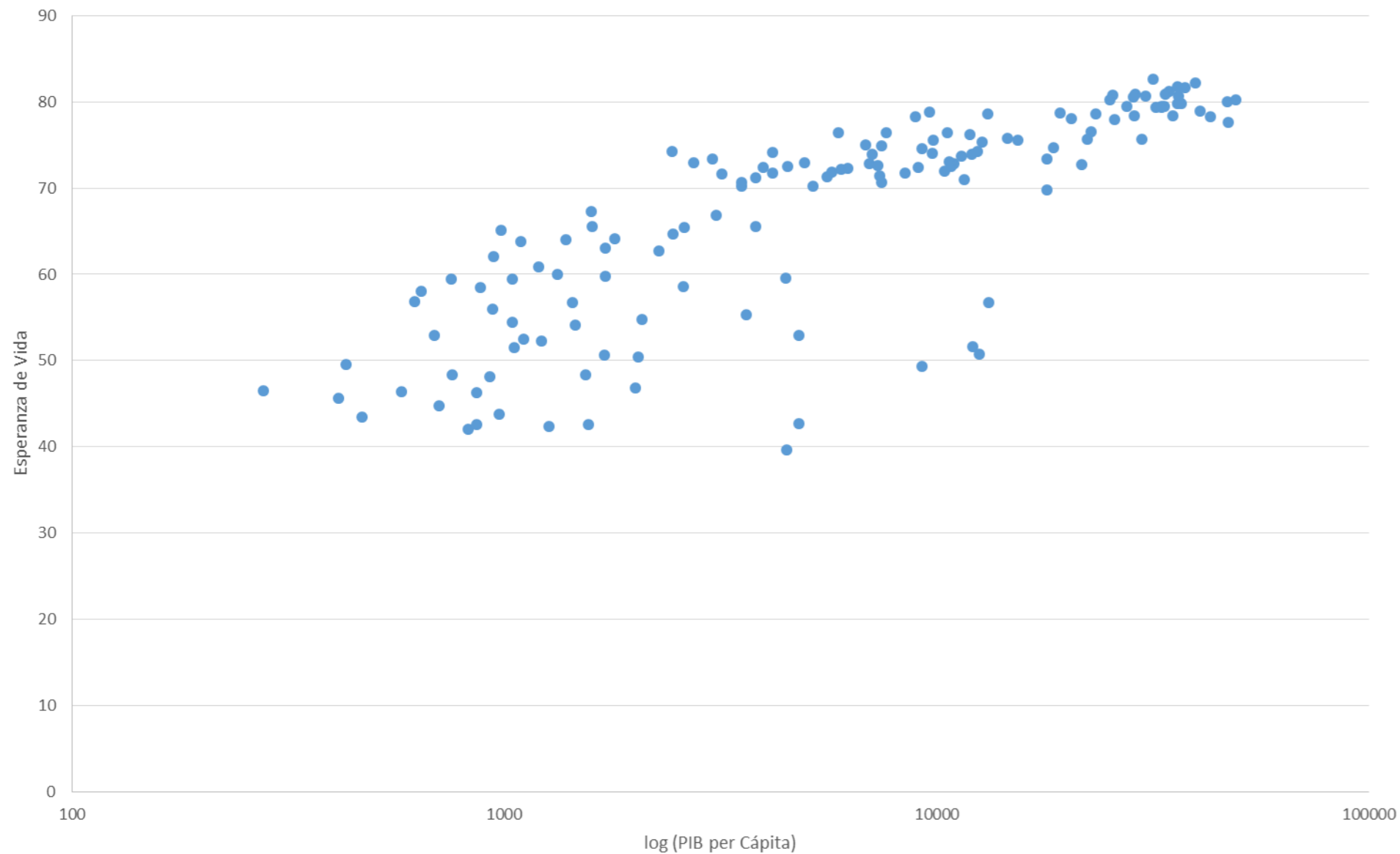
Buscar la relación entre la Esperanza de vida y el PIB per cápita

Esperanza de Vida vs PIB per Cápita 2007



1. Crear gráficos
2. Interpretarlos
3. Extraer las primeras respuestas

Esperanza de Vida vs PIB per Cápita 2007



Objetivos

Evolución de la población en los diferentes continentes
Evolución de la esperanza de vida en los diferentes continentes
Evolución del PIB per cápita en los diferentes continentes

Buscar la relación entre la Esperanza de vida y el PIB per cápita

1. Responder a tus preguntas
2. Interpretar los resultados
3. Conclusiones

Objetivos

Evolución de la población en los diferentes continentes

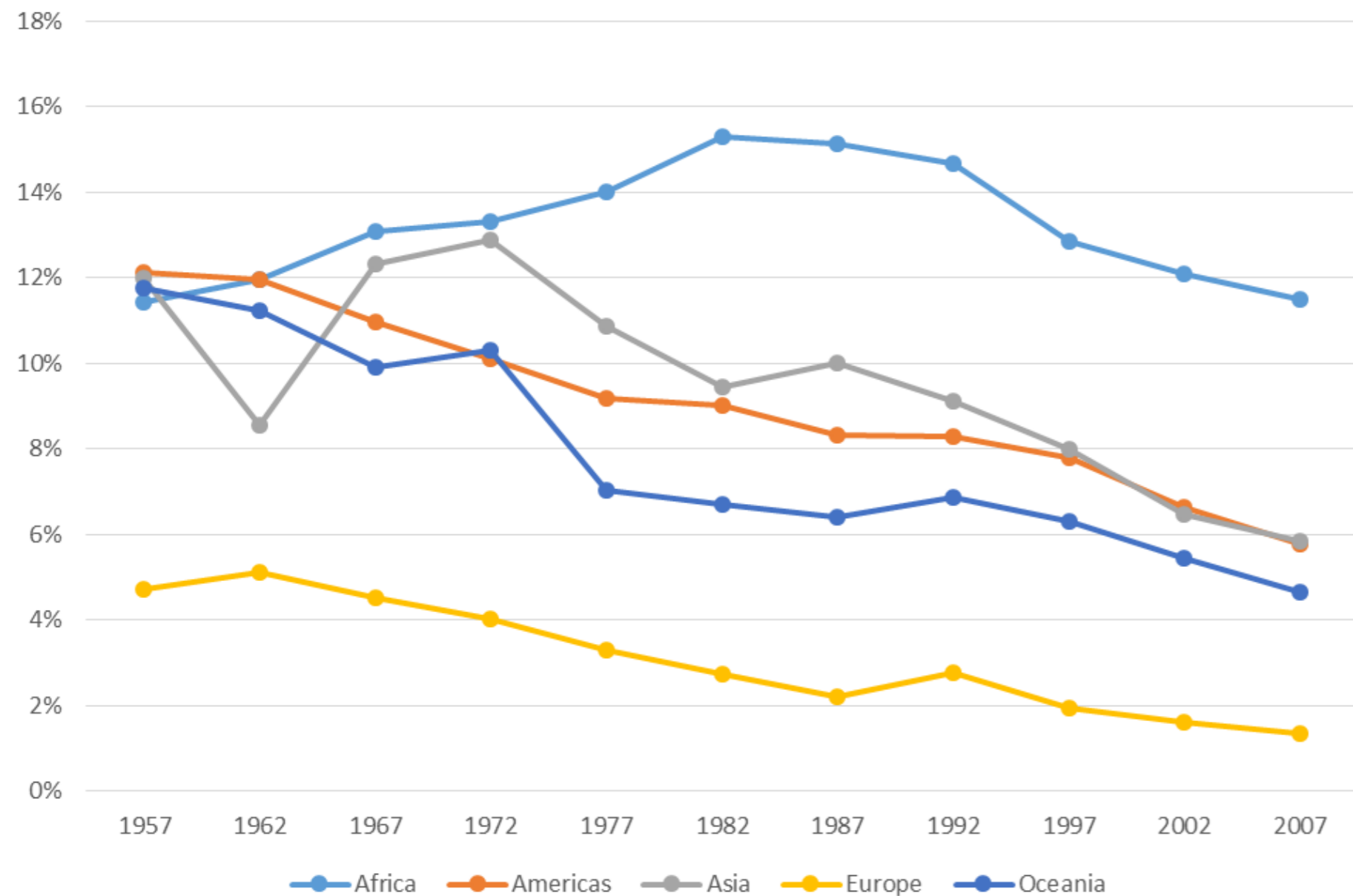
Evolución de la esperanza de vida en los diferentes continentes

Evolución del PIB per cápita en los diferentes continentes

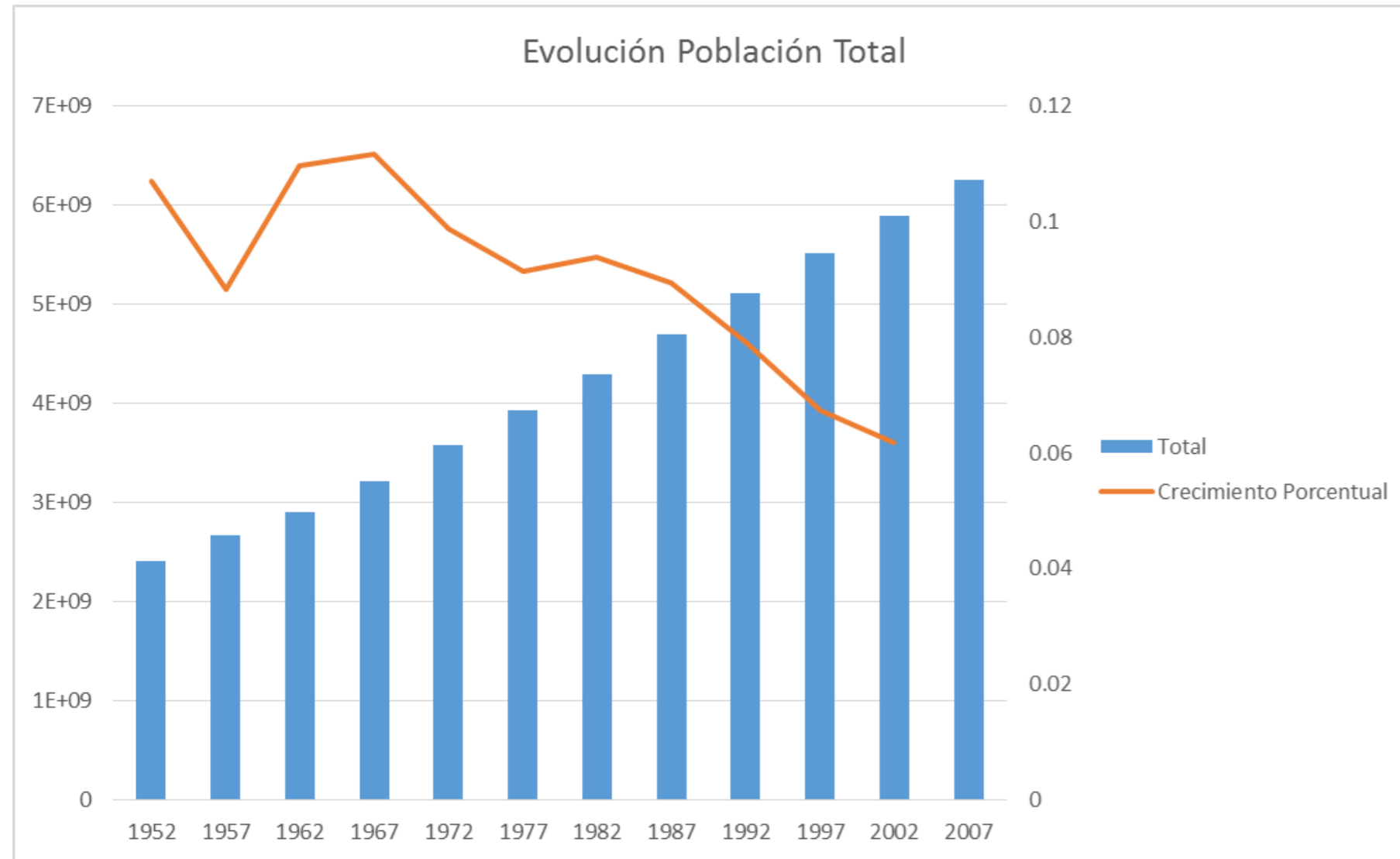
Buscar la relación entre la Esperanza de vida y el PIB per cápita

$$\text{Crecimiento Porcentual} = \frac{\text{Período de Ahora} - \text{Período de Antes}}{\text{Período de Antes}} \cdot 100$$

Evolución Crecimiento Porcentual de la Población



1. Responder a tus preguntas
2. Interpretar los resultados
3. Conclusiones



Objetivos

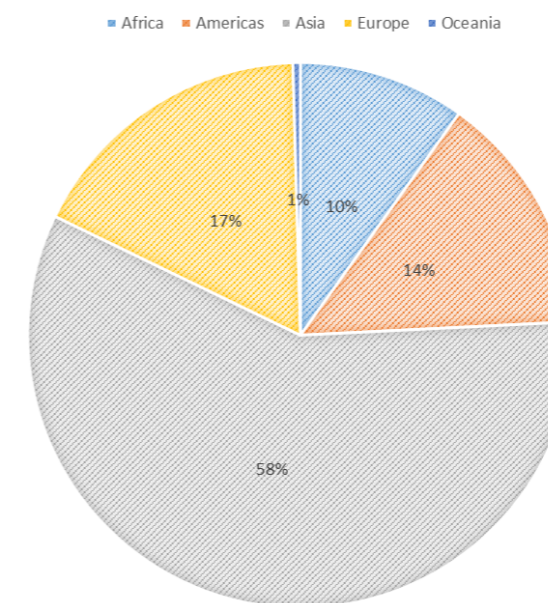
Evolución de la población en los diferentes continentes

Evolución de la esperanza de vida en los diferentes continentes

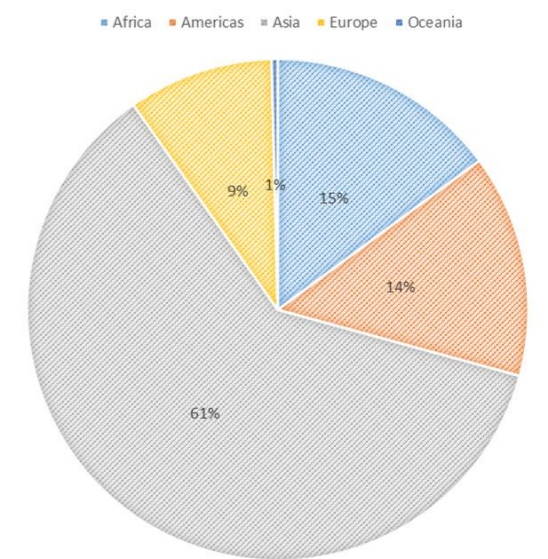
Evolución del PIB per cápita en los diferentes continentes

Buscar la relación entre la Esperanza de vida y el PIB per cápita

DISTRIBUCIÓN POBLACIÓN 1952

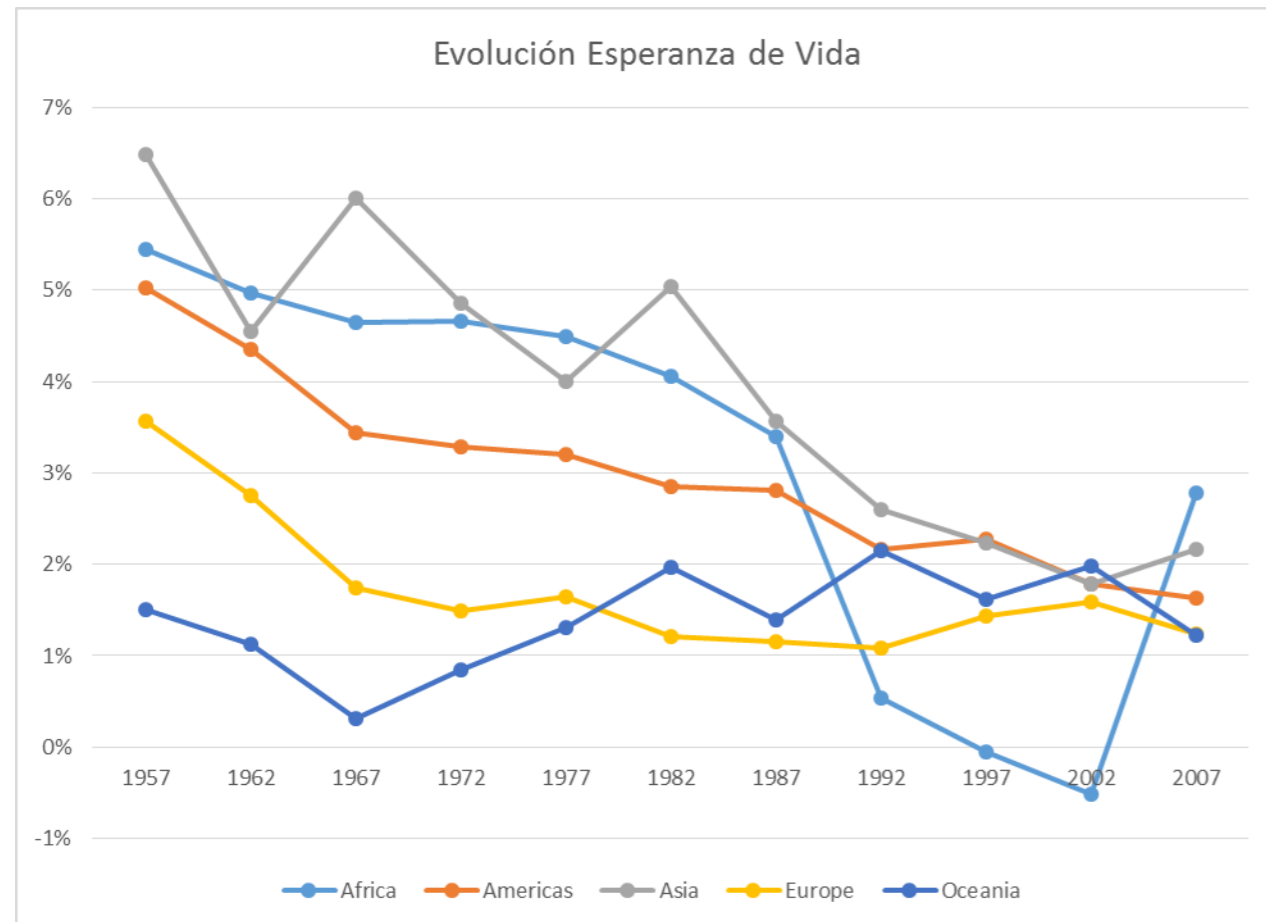


DISTRIBUCIÓN POBLACIÓN 2007



Crecimiento continuado de la población total
El ritmo de crecimiento va disminuyendo

1. Responder a tus preguntas
2. Interpretar los resultados
3. Conclusiones



El crecimiento de Esperanza de vida ha sido elevado de los 50 a los 68 a nivel mundial

El ritmo de crecimiento ha descendido hasta que en 2007 ha aumentado gracias al crecimiento de Asia y África

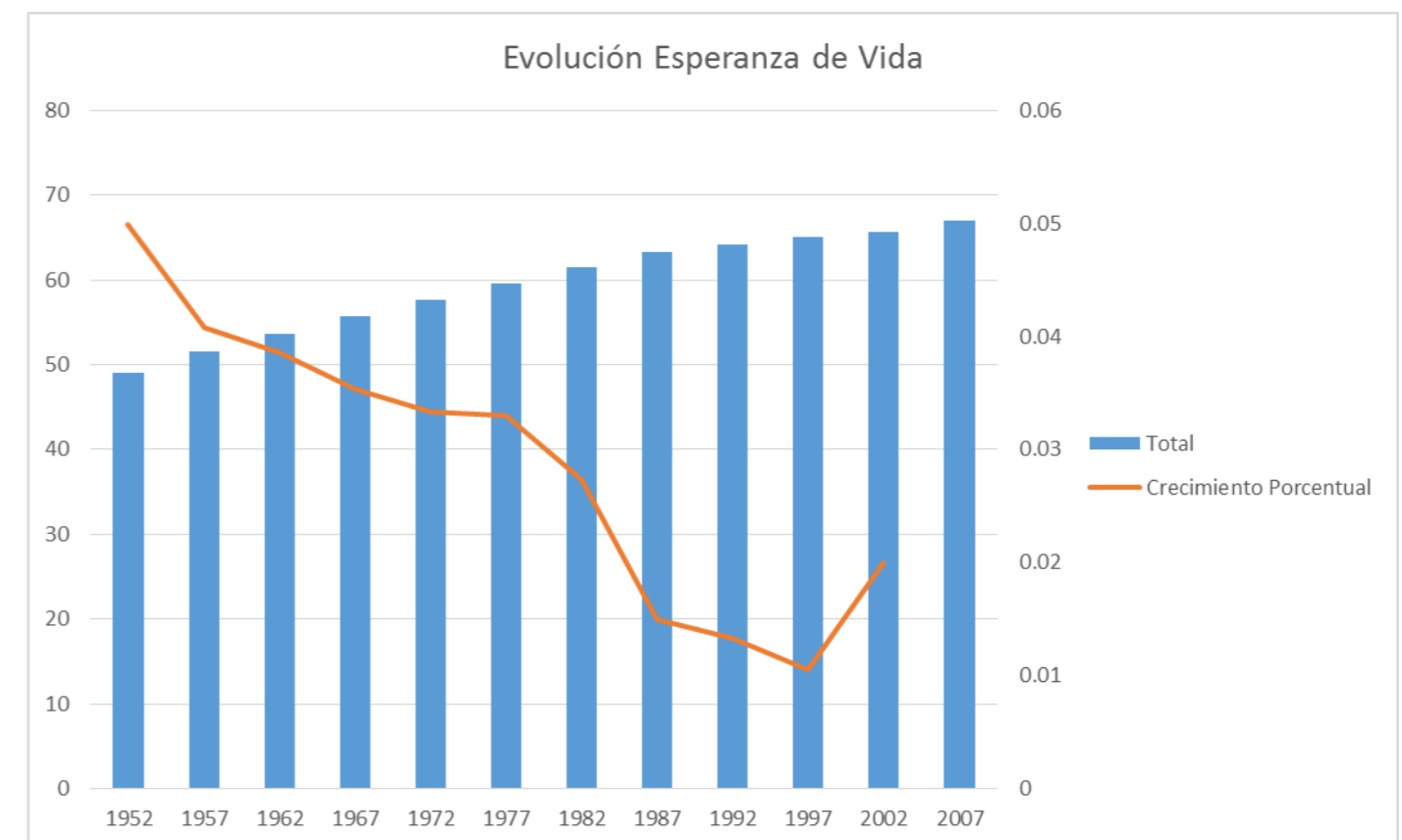
Objetivos

Evolución de la población en los diferentes continentes

Evolución de la esperanza de vida en los diferentes continentes

Evolución del PIB per cápita en los diferentes continentes

Buscar la relación entre la Esperanza de vida y el PIB per cápita



1. Responder a tus preguntas
2. Interpretar los resultados
3. Conclusiones

Objetivos

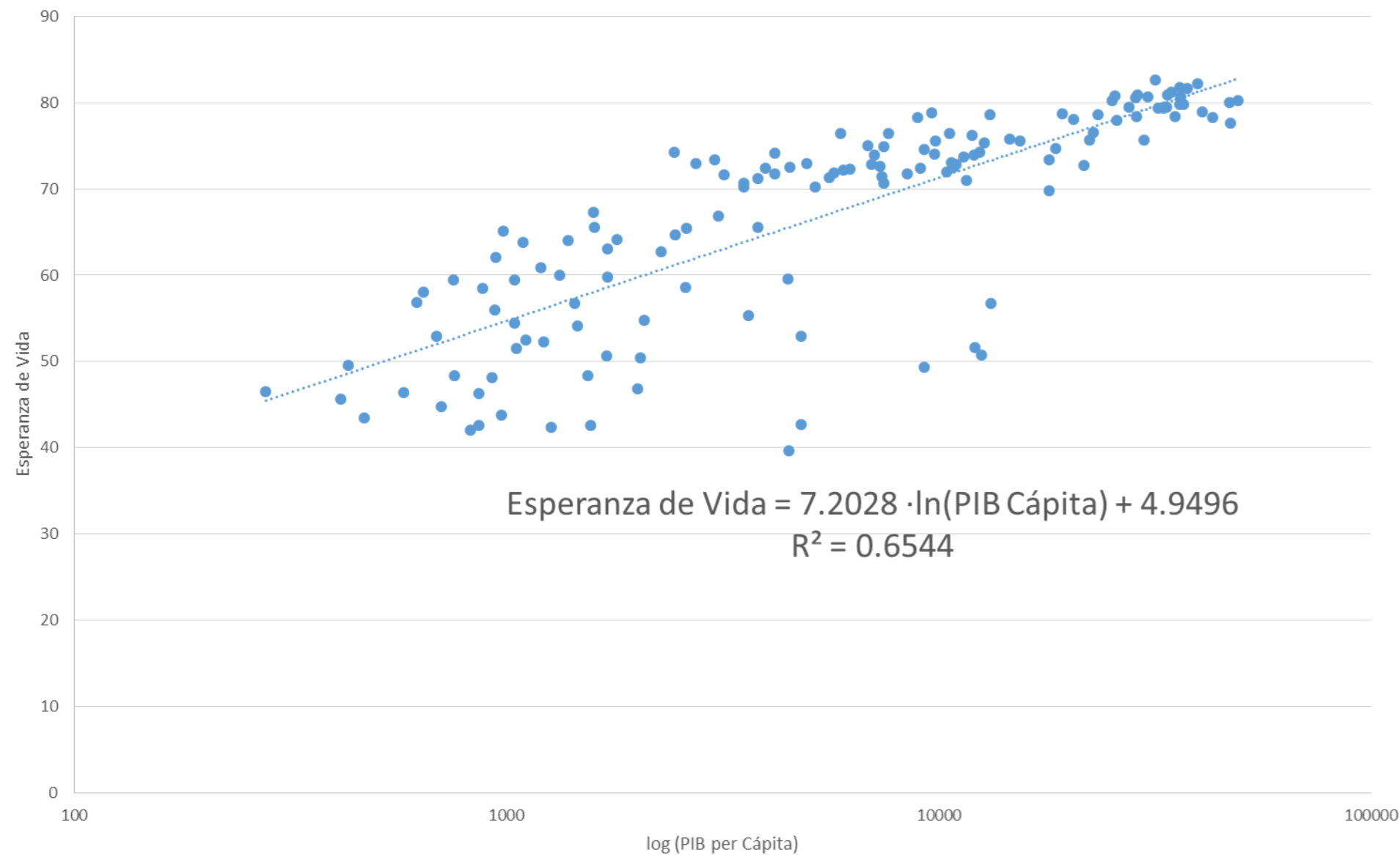
Evolución de la población en los diferentes continentes
Evolución de la esperanza de vida en los diferentes continentes
Evolución del PIB per cápita en los diferentes continentes

Buscar la relación entre la Esperanza de vida y el PIB per cápita

La asociación es clara entre la esperanza de Vida y el log del PIB per cápita

Esperanza de Vida = $7.2028 \cdot \ln(\text{PIB Cápita}) + 4.9496$
 $R^2 = 0.6544$

Esperanza de Vida vs PIB per Cápita 2007



20.13 +
4.63 PIB Cápita +
11.69 America +
10.11 Asia +
11.26 Europa +
12.92 Oceania

3

Paso 3 – Analizar los Datos

1. Responder a tus preguntas
2. Interpretar los resultados
3. Conclusiones

Objetivos

Evolución de la población en los diferentes continentes
Evolución de la esperanza de vida en los diferentes continentes
Evolución del PIB per cápita en los diferentes continentes

Buscar la relación entre la Esperanza de vida y el PIB per cápita

```
call:  
glm(formula = lifeExp ~ log(gdpPercap) + continent, data = data2007)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-19.4917	-2.3146	-0.0432	2.5498	14.8818

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	20.1376	4.0332	4.993	1.79e-06	***
log(gdpPercap)	4.6308	0.5274	8.780	6.14e-15	***
continentAmericas	11.6942	1.6546	7.068	7.46e-11	***
continentAsia	10.1144	1.4761	6.852	2.31e-10	***
continentEurope	11.2682	1.8936	5.951	2.14e-08	***
continentOceania	12.9293	4.5211	2.860	0.00491	**

signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 35.15473)

Null deviance: 20552 on 141 degrees of freedom
Residual deviance: 4781 on 136 degrees of freedom
AIC: 916.33

Number of Fisher Scoring iterations: 2

Modelo lineal Generalizado que
dependa del PIB per Cápita y el
Continente

$$\text{Esperanza de Vida} = 7.2028 \cdot \ln(\text{PIB Cápita}) + 4.9496$$
$$R^2 = 0.6544$$

1. Responder a tus objetivos
 2. Un esquema efectivo de comunicación
-

Contexto

Cita la problemática

Objetivos

Define tu objetivo alineado al problema que quieres resolver

Gráficos y Resultados

Muestra la información escondida (etapa de exploración)

Define tu metodología de análisis (etapa de análisis)

Muestra los resultados (etapa de análisis)

Conclusiones

Resume las conclusiones

Problemas y next steps

Problemas encontrados durante el estudio y siguientes pasos





PASO 1 – ENFOCA TU PROYECTO

PASO 2 – EXPLORA TUS DATOS

PASO 3 – ANALIZA CON UN
SOFTWARE y UN MAPA DE
TÉCNICAS

PASO 4 – COMUNICA TUS
RESULTADOS

1. Entender tu tabla de datos
 2. Definir el objetivo de análisis
 3. Leer e interpretar tus base de datos
-

Problema

Mejorar el tratamiento del dolor lumbar.

El 85% de la población los sufre a lo largo de la vida



Recolección



Grupo Control

Tratamiento Convencional
VS
Tratamiento Estándar



Grupo Investigación

Mediciones la
mejora antes
y después de
los
tratamientos

- 1. Entender tu tabla de datos
- 2. Definir el objetivo de análisis
- 3. Leer e interpretar tus base de datos

La tabla de datos

Nombre	Tipo	¿Qué característica observa?
Id Paciente	Cuantitativa Discreta	Identificador del Paciente. Número identificado.
Sexo	Cualitativa Nominal	Sexo (0 = Hombre 1 = Mujer)
Edad	Cuantitativa Discreta	Edad del Paciente
Peso	Cuantitativa Continua	Peso del paciente en kg
Altura	Cuantitativa Continua	Altura en metros
LC	Cualitativa Nominal	Tipo de Lumbociática (0 = NO Lumbociática, 1 = Lumbociática Derecha y 2 = Lumbociática Izquierda)
PC	Cualitativa Nominal	Pierna corta aparente (0 = No Pierna Corta, 1 = Pierna Corta Derecha, 2 = Pierna Corta Izquierda)
NHD	Cualitativa Ordinal	Número hernia discal (0 = Ninguna, 1 = Una, 2 = Dos)
ODIMes0	Cuantitativa Continua	Índice de Discapacidad de Oswestry Mes0 (Porcentaje Antes Tratamiento)
ODIMes1	Cuantitativa Continua	Índice de Discapacidad de Oswestry Mes1 (Porcentaje Después Tratamiento)
Grupo	Cualitativa Nominal	Tipo de Tratamiento (0 = Convencional, 1 = Investigación)
LF_Mes0	Cualitativa Ordinal	Límite Funcional Mes0 (0 = Mínima, 1 = Moderada, 2 = Intensa, 3 = Discapacidad, 4 = Máxima)
LF_Mes1	Cualitativa Ordinal	Límite Funcional Mes1 (0 = Mínima, 1 = Moderada, 2 = Intensa, 3 = Discapacidad, 4 = Máxima)
IMC_Grupos	Cualitativa Ordinal	Estados Nutricional (0 = desnutrido, 1 = Normal, 2 = Sobrepeso, 3 = Obesidad, 4 = Obesidad Mórbida)

	A	B	C	D	E	F	G	H	I	J	K
1	Id Paciente	Sexo	Edad	Peso	Altura	LC	PC	NHD	ODI Mes0	ODI Mes1	Grupo
2	1	0	44	64.5	1.63	0	0	0	44.3	35.2	0
3	2	0	33	67.52	2.02	0	0	0	38.9	29.4	0
4	3	0	47	72.72	1.6	0	0	0	39.3	31.7	0
5	4	0	34	71.83	1.85	0	0	0	33.4	23.6	0
6	5	0	41	52.08	1.67	0	0	0	42.7	34.6	0
7	6	0	38	72.24	1.61	1	2	1	69.4	55.1	0
8	7	0	47	77.98	1.57	0	0	0	31.3	22.4	0
9	8	0	60	60.54	1.71	1	2	1	61.2	47.9	0
10	9	0	75	89.12	1.98	1	2	1	65	52.3	0
11	10	0	40	68.88	1.54	1	2	1	68.6	54.1	0
12	11	0	71	86.48	1.78	2	1	0	63.9	55.2	0
13	12	0	51	60.13	1.65	0	0	0	34	26.8	0
14	13	0	46	74.17	1.75	0	0	0	36.3	27.6	0
15	14	0	57	69.93	1.82	1	2	0	71.8	56.9	0
16	15	0	56	89.46	1.89	0	0	0	74.6	66	0
17	16	0	72	109.2	2.08	1	2	0	65.3	53	0
18	17	0	64	66.33	1.62	1	2	0	65.5	58	0
19	18	0	71	65.79	1.67	2	1	0	60.2	54.7	0
20	19	0	56	72.21	2.02	0	0	0	68.3	62.8	0
21	20	0	35	61.85	1.59	0	0	0	58	50.1	0
22	21	0	35	63.42	1.78	0	0	0	25.2	20.2	0
23	22	0	73	72.27	1.69	2	1	0	66.6	57.7	0
24	23	0	46	63.3	1.38	0	0	0	48.2	38.9	0
25	24	0	62	64.33	1.78	1	2	0	67.1	61.5	0
26	25	0	79	73.55	1.71	2	1	1	72.6	49.6	0
27	26	0	63	65.29	1.66	1	2	1	62.3	54.6	0
28	27	0	64	66.55	1.31	1	2	1	64.8	58.7	0
29	28	0	63	69.34	1.88	2	1	1	67.6	43.1	0
30	29	0	31	66.01	1.68	2	1	1	71.7	53.5	0
31	30	0	55	71.57	1.8	2	1	1	67.7	46.2	0
32	31	0	36	64.81	1.78	0	0	0	36.8	27.1	0
33	32	0	73	60.42	1.64	1	2	1	73.3	61.9	0
34	33	0	77	88.9	1.87	1	2	1	68.9	62.1	0
35	34	0	66	50.15	1.44	1	2	1	72.7	62.2	0

1

Paso 1 – Enfocar El Estudio

- 1. Entender tu tabla de datos
- 2. Definir el objetivo de análisis
- 3. Leer e interpretar tus base de datos

• **Objetivo:** definir qué queremos obtener con los datos

OBJETIVO

Comparar dos tratamientos
Control vs investigación

La diferencia entre el ODI en el mes 0 y el mes 1 y analizar qué tratamiento mejora más a los pacientes

	A	B	C	D	E	F	G	H	I	J	K
1	Id Paciente	Sexo	Edad	Peso	Altura	LC	PC	NHD	ODI Mes0	ODI Mes1	Grupo
2	1	0	44	64.5	1.63	0	0	0	44.3	35.2	0
3	2	0	33	67.52	2.02	0	0	0	38.9	29.4	0
4	3	0	47	72.72	1.6	0	0	0	39.3	31.7	0
5	4	0	34	71.83	1.85	0	0	0	33.4	23.6	0
6	5	0	41	52.08	1.67	0	0	0	42.7	34.6	0
7	6	0	38	72.24	1.61	1	2	1	69.4	55.1	0
8	7	0	47	77.98	1.57	0	0	0	31.3	22.4	0
9	8	0	60	60.54	1.71	1	2	1	61.2	47.9	0
10	9	0	75	89.12	1.98	1	2	1	65	52.3	0
11	10	0	40	68.88	1.54	1	2	1	68.6	54.1	0
12	11	0	71	86.48	1.78	2	1	0	63.9	55.2	0
13	12	0	51	60.13	1.65	0	0	0	34	26.8	0
14	13	0	46	74.17	1.75	0	0	0	36.3	27.6	0
15	14	0	57	69.93	1.82	1	2	0	71.8	56.9	0
16	15	0	56	89.46	1.89	0	0	0	74.6	66	0
17	16	0	72	109.2	2.08	1	2	0	65.3	53	0
18	17	0	64	66.33	1.62	1	2	0	65.5	58	0
19	18	0	71	65.79	1.67	2	1	0	60.2	54.7	0
20	19	0	56	72.21	2.02	0	0	0	68.3	62.8	0
21	20	0	35	61.85	1.59	0	0	0	58	50.1	0
22	21	0	35	63.42	1.78	0	0	0	25.2	20.2	0
23	22	0	73	72.27	1.69	2	1	0	66.6	57.7	0
24	23	0	46	63.3	1.38	0	0	0	48.2	38.9	0
25	24	0	62	64.33	1.78	1	2	0	67.1	61.5	0
26	25	0	79	73.55	1.71	2	1	1	72.6	49.6	0
27	26	0	63	65.29	1.66	1	2	1	62.3	54.6	0
28	27	0	64	66.55	1.31	1	2	1	64.8	58.7	0
29	28	0	63	69.34	1.88	2	1	1	67.6	43.1	0
30	29	0	31	66.01	1.68	2	1	1	71.7	53.5	0
31	30	0	55	71.57	1.8	2	1	1	67.7	46.2	0
32	31	0	36	64.81	1.78	0	0	0	36.8	27.1	0
33	32	0	73	60.42	1.64	1	2	1	73.3	61.9	0
34	33	0	77	88.9	1.87	1	2	1	68.9	62.1	0
35	34	0	66	50.15	1.44	1	2	1	72.7	62.2	0

2

Paso 2 – Explorar tu base de datos

1. Crear gráficos
2. Interpretarlos
3. Extraer las primeras respuestas

- **Exploración/Descripción:** interpretar los datos

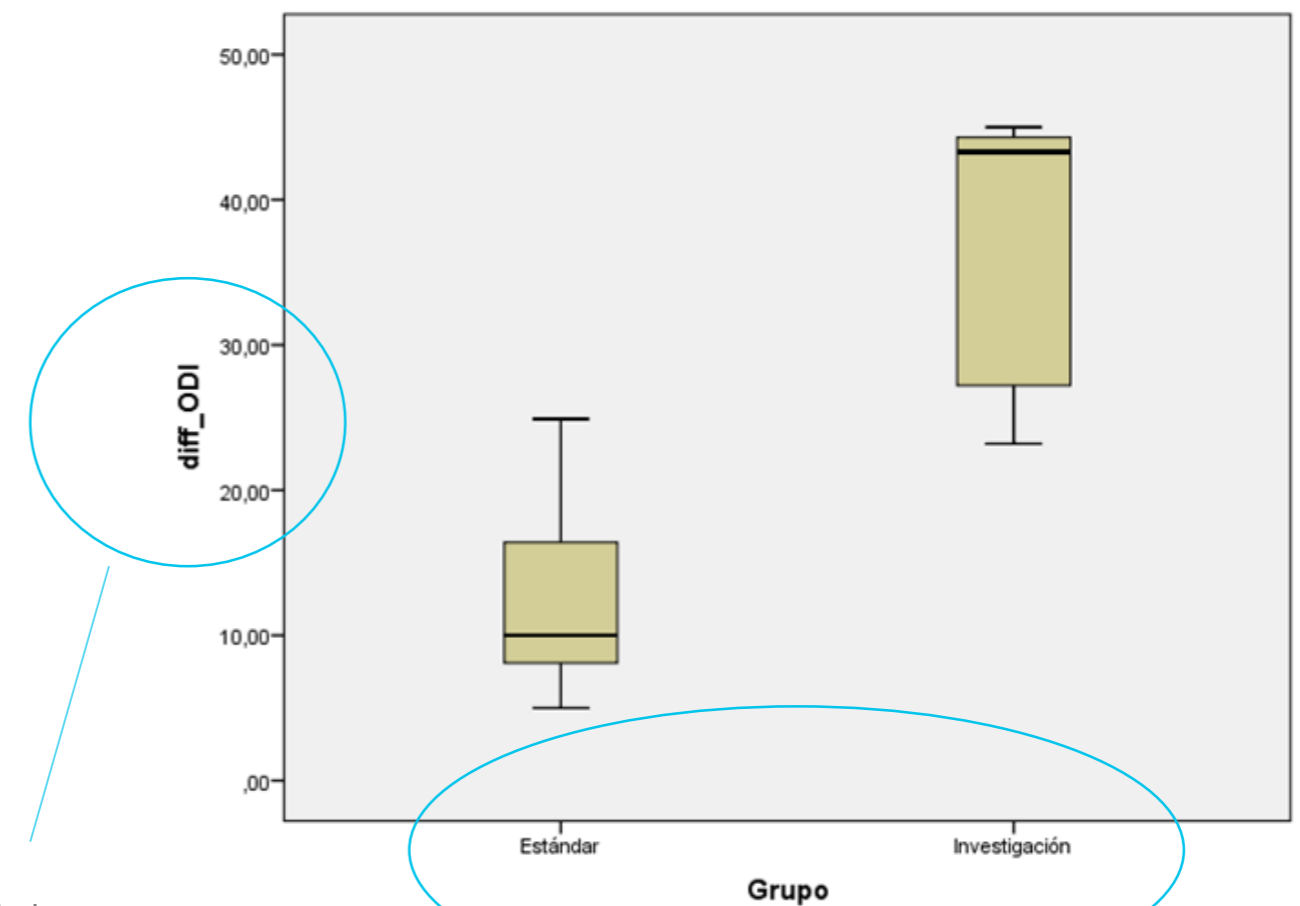
OBJETIVO

Comparar dos tratamientos

Control vs investigación

La diferencia entre el ODI en el mes 0 y el mes 1 y analizar qué tratamiento mejora más a los pacientes

Comparamos las distribuciones de la diferencia de ODI – limitación funcional



Variable dependientes
O variable respuesta

Grupos que comparamos
(variable independiente)

1. Responder a tus preguntas
2. Interpretar los resultados
3. Conclusiones

Comparación de medias

- **Análisis y conclusión:** aplicar técnicas de análisis

OBJETIVO

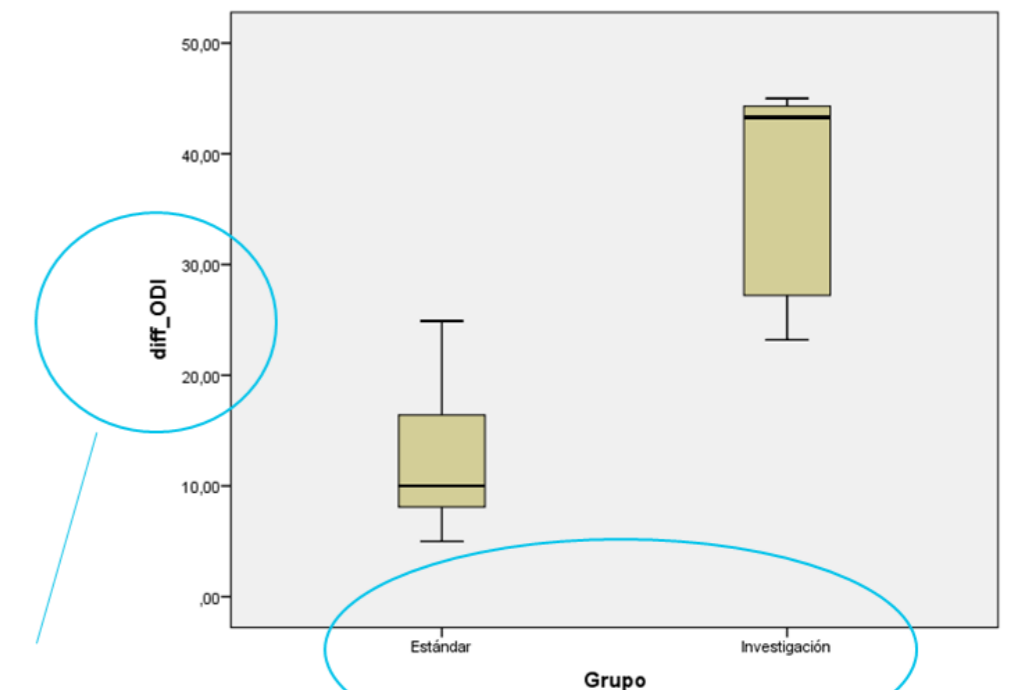
Comparar dos tratamientos

Control vs investigación

La diferencia entre el ODI en el mes 0 y el mes 1 y analizar qué tratamiento mejora más a los pacientes

Estadísticos de prueba ^a	
	diff_ODI
U de Mann-Whitney	27,000
W de Wilcoxon	5178,000
Z	-12,151
Sig. asintótica (bilateral)	,000

Detalles:
Test NO paramétrico
Grupos independientes



Variable dependientes
O variable respuesta

Grupos que comparamos
(variable independiente)

- 1. Responder a tus preguntas
- 2. Interpretar los resultados
- 3. Conclusiones

Comparación de medias

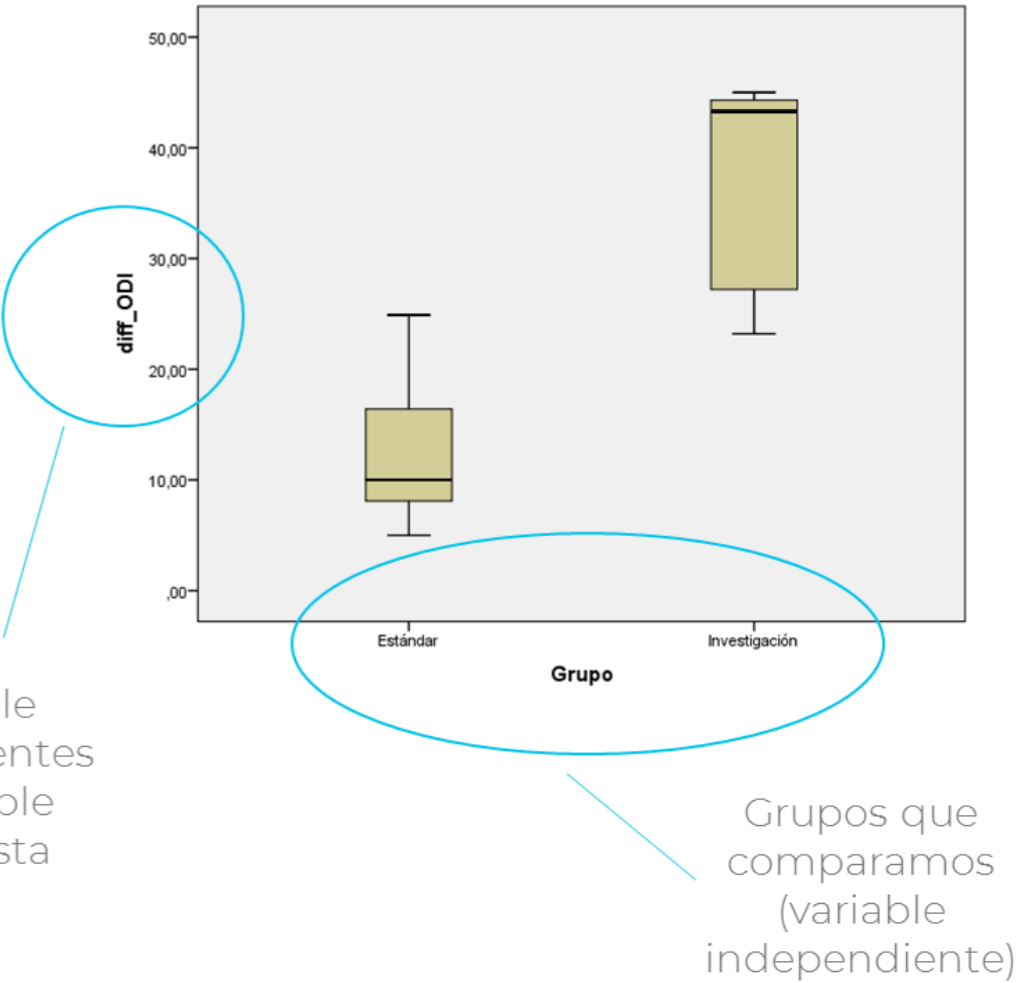
- **Análisis y conclusión:** aplicar técnicas de análisis

OBJETIVO

Comparar dos tratamientos
Control vs investigación

La diferencia entre el ODI en el mes 0 y el mes 1 y analizar qué tratamiento mejora más a los pacientes

Estadísticos de prueba ^a	
	diff_ODI
U de Mann-Whitney	27,000
W de Wilcoxon	5178,000
Z	-12,151
Sig. asintótica (bilateral)	,000
Detalles: Test NO paramétrico Grupos independientes	



Hay diferencias significativas en el efecto sobre la limitación funcional entre el tratamiento control y el de investigación

1. Responder a tus objetivos
 2. Un esquema efectivo de comunicación
-

Contexto

Cita la problemática

Objetivos

Define tu objetivo alineado al problema que quieres resolver

Gráficos y Resultados

Muestra la información escondida (etapa de exploración)

Define tu metodología de análisis (etapa de análisis)

Muestra los resultados (etapa de análisis)

Conclusiones

Resume las conclusiones

Problemas y next steps

Problemas encontrados durante el estudio y siguientes pasos





PASO 1 – ENFOCA TU PROYECTO

PASO 2 – EXPLORA TUS DATOS

PASO 3 – ANALIZA CON UN
SOFTWARE y UN MAPA DE
TÉCNICAS

PASO 4 – COMUNICA TUS
RESULTADOS



TAKE AWAY

- **Tu Temática + Software + Estadística + Pasión** son los ingredientes necesarios para que puedas ser un profesional imparable
- El Elemento que necesitas trabajar es un **ÁRBOL: bases estadísticas, técnicas de análisis de datos, sistema**
- Un **Sistema** de forma rápida y segura tiene los siguientes pasos: **Enfoque, descripción, análisis, conclusión**
- Los **softwares** de datos se clasifican en tres: **Genéricos, Análisis Científicos, Ciencia de Datos**. Y los recomendados son **los de Ciencia de Datos**





Jesús Porro - Experto en trading

Os presento a Jesús Porro, alumno de Analiza tus Datos.

Dónde está descubriendo y entendiendo muchos conceptos estadísticos que se aplican en sus cálculos para inversiones el bolsa.

Con el primer bloque ya está viendo claros muchas de las herramientas y nomenclatura en bolsa des del punto de vista estadístico. ¡Esto es fundamental!

() El vídeo está extraído en motivo del segundo congreso Investiga con Datos publicado en julio de 2019.*

Ya que en un futuro tendrá el control completo de sus cálculos y con su conocimiento en inversiones va a mejorar seguro sus resultados y para sus alumnos.

La estadística es una disciplina transversal, aplicable en muchas situaciones.

Jesús Porro es el creador de vivirdeltrading.com para ayudarte a vivir de las inversiones con un método diferente y práctico. Actualmente también ha creado un nuevo proyecto eldespertardelborrego.com para ayudarnos a los que nos sabemos de inversiones a invertir nuestro dinero a largo plazo con rentabilidades mucho más grandes que las que nos ofrecen los bancos.



(*) El vídeo está extraído en motivo del segundo congreso Investiga con Datos publicado en julio de 2019.

Gabriele Zunino - Datos y RRHH

Gabriele es un apasionado del comportamiento humano y la psicología en el entorno organizacional.

Se está especializando en la ciencia de los datos en el entorno de la psicología y recursos humanos.

Hoy Gabriele nos contará su experiencia sobre cómo empezar a especializarse en la ciencia de los datos desde el inicio. Es alumno de Analiza tus Datos y nos mostrará un pequeño ejemplo de asociación con tablas de contingencia:)

Está aprendiendo R y técnicas estadísticas para poder trabajar en un entorno organizacional para después lanzar su propio proyecto en torno a la ciencia de los datos. Se está formando en más cursos y recursos y me comentó: "Oye Jordi, Analiza tus Datos es muy barato por lo que realmente aprendes. Si lo comparo con otros cursos"

¡Gracias Gabriele!



() El vídeo está extraído en motivo del segundo congreso Investiga con Datos publicado en Julio de 2019.*

María Videgain - Investigadora Ingeniera Agrónoma

Os presento a María Videgain.

Alumna de Analiza tus Datos y doctorando. Es ingeniera agrónoma y está realizando la tesis doctoral en la universidad de Zaragoza.

En este vídeo nos muestra un estudio muy interesante donde aplica técnicas estadísticas multivariantes y hoy vamos a realizar un tutorial para que entiendas el paso a paso de cómo lo está realizando María con SPSS.

María, con Analiza tus Datos se está preparando para encontrar nuevas oportunidades y retos donde el análisis de datos sea una parte muy importante.
¡Seguro que le va a ir genial!

¡Gracias María!

Tienes 3 opciones

1

Tirar la toalla y conformarte

2

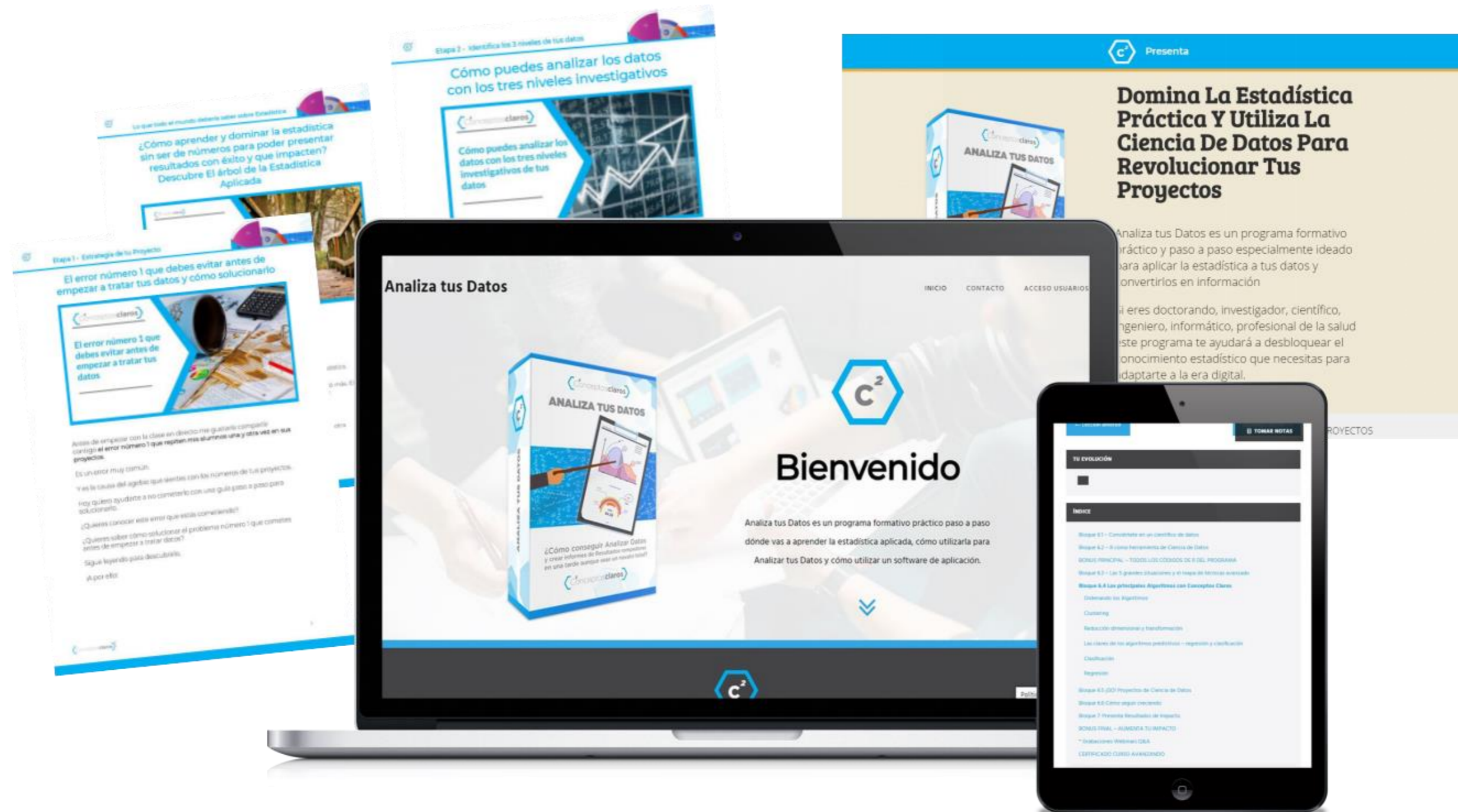
Seguir tú sólo esperando tener en algún momento mejores resultados

3

Dejar que nosotros te ayudemos a conseguir mejores resultados usando un sistema probado



Analiza tus datos



Analizatusdatos

ANTES

- Pierdes horas para sacar resultados de tus datos o no sabes ni cómo empezar
- Te cuesta entenderte con los datos y números de tus proyectos
- No estás seguro de si los resultados que obtienes tienen sentido
- No logras tener un mapa claro y cristalino para obtener resultados con éxito

DESPUÉS

- ✓ Eres rápido y enfocado a la hora de trabajar tus datos.
- ✓ Sabes qué técnica utilizar en cada caso y por qué. Dispones de un mapa de técnicas y cómo aplicarlas
- ✓ Abres nuevas opciones. Conoces nuevas maneras de enfocar tus proyectos mediante datos
- ✓ Aprovechas el 100% de los datos ya que eres dueño de una habilidad: la Ciencia de Datos



Nuestro **propósito** es crear **más innovadores**
capaces de utilizar sus datos para presentar
nuevos avances **con éxito** y consigan
mejorar el mundo gracias a su conocimiento



INNOVADORES AD



"Mi nombre es María Jesús, soy enfermera. Y tanto en la carrera como en distintos posgrados universitarios he tenido estadística.

Sin embargo, hasta que no conté con el asesoramiento de Jordi no comencé a comprenderla.

Y eso que tuve mis dudas al principio, por si iba a malgastar el dinero. Pero ha sido como una espiral, poco a poco vas aprendiendo de una manera práctica y directa.

Por eso Jordi y sus conceptos claros forman parte de los agradecimientos de mi tesis doctoral"

Mª JESÚS SAMANIEGO

PhD Modelos Meteorológicos - Físico



"Antes conocía la estadística y he realizado algún curso de inteligencia de negocio pero con el blog y sobre todo el curso de Jordi, estoy perdiendo el miedo a la práctica de la analítica avanzada. Espero poder afianzar mucho más esta sensación con lo que queda de curso.

De momento tener un método para atacar los objetivos analíticos y estoy empezando a entender esto del foco, que antes tenía muy "desenfocado"

Ni lo dudes un minuto. Estés donde estés. Hagas lo que hagas. Necesitas referencias y contexto. Y eso en nuestros días es analítica y la mejor forma de conseguirlo es "el curso completo de Analiza tus Datos"

MANUEL CAÑO

Análisis del entorno de la contratación pública electrónica en España



"Después de haber estudiado estadística en la universidad hace algún tiempo me di cuenta de que, como dice el dicho: "Yo había pasado por la universidad pero la universidad no había pasado por mí", y que a la hora de hacer uso de herramientas estadísticas para mí día a día la cosa no iba todo lo bien que debería.

El hecho de ver que puedes aplicar todo lo que enseña Jordi tanto en el ámbito personal (para emprender), como en el profesional (ser un mejor profesional en tu empresa y destacar) hace que veas posible llevar a cabo un montón de proyectos que antes no veía.

El curso es de bueno a buenísimo, yo lo estoy cursando ahora y estoy super enganchado, se aprende mucho, y además una de las cosas que mas valoro es que siempre tienes la sensación de estar acompañado."

REICHE TORAO

Calidad y medio ambiente



"Antes de llegar a Analiza tus Datos andaba un poco perdida sobre qué análisis estadístico aplicar, y sobre todo, como analizar los resultados.

Ahora tengo muy claro qué análisis se aplica en cada caso particular, y cómo ser capaz de expresar los resultados de una forma correcta.

El programa Analiza tus Datos es una gran inversión, se avanza con facilidad y además con las hojas de trabajo es como mejor se aprende, trabajando! Yo personalmente estoy muy satisfecha con los contenidos, el apoyo de Jordi pero sobre todo con los conocimientos adquiridos!

Si te lo estás pensando, es porque debes entrar en el programal"

ZALOA SANCHEZ

Estudiante de doctorado - profesora



"Yo súper encantada con el curso que seguí me ha servido un montón, me toco desaprender para aprender bien!!!!!!...gracias por ello"

JOHANA MUÑOZ

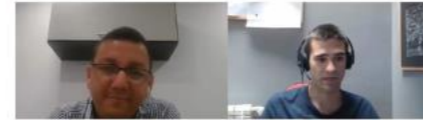
Estudiante de Doctorado - Ingeniera Forestal



"Con este curso he podido mejorar mi comprensión y capacidad de análisis con el uso adecuado de la estadística y R+RStudio, trabajo habitualmente con datos en salud y este curso me ha permitido entender de forma fácil conceptos, que para mí eran complejos, permitiendo abordar con más profundidad mis análisis."

NORBERTO VERA

Analista de Datos en Salud



"Al principio me ayudó mucho aprender conceptos que eran muy complejos para mí"

Norberto Vera

Freelance analista de datos en Salud

Colombia



"Conceptos estadísticos explicados con mimo"

José Casanova

Doctor e investigador Ingeniería Agrónoma

Huesca, España



"Analiza tus Datos me está cambiando la vida. Veo nuevas oportunidades que antes no veía"

Reiche Torao

Calidad y medio ambiente

Valencia, España



"Entiendo la lógica que hay detrás de la estadística para aplicarla a mis proyectos"

José de Jesús Rivera

Doctor e investigador

Barcelona, España



Pedro Margolles

Psicólogo y Enfermero creador del blog neoscientia.com



Fernando Ingeniero Agrónomo

Con este curso conseguí ver qué técnicas se utilizan en los diferentes casos del análisis estadístico de datos (cosa que nunca me había quedado demasiado claro). Otra cosa es haber visto el potencial que tiene el programa R. El training está muy bien estructurado y con mucho material adicional.



En

Hola compañeros. Aunque un poco tarde me voy a presentar. Soy sociologo y estoy haciendo un master en comunicacion politica. Encontre el curso de Jordi, y compre Analiza tu estudio. La verdad es que ha sido una buena inversion, me ha aclarado un monton de terminos y hasta he disfrutado con la estadistica. Un saludo a todos y felices fiestas

Me encanta

Comentar



Sí, perdón por el retraso, me presento. Soy economista pero hace tiempo que acabé la carrera. Me ha dado por volver a estudiar y ampliar mi formación y necesito refrescar conocimientos además de aprender a utilizar R. Vi el curso de Jordi y de momento me parece genial. Un salud a todos.

Me gusta · Responder · 2 d



NORBERTO VERA
Alumno Analiza tus Datos - Estudios Científicos
Analista y consultor de Datos en TI/ Salud - Freelance
"Con Jordi he aprendido conceptos estadísticos muy importantes y que no tenía asentados. Además he podido aprender a utilizar RStudio para poder presentar resultados mucho más potentes a mis clientes."

SALVADOR RAMOS
Ponente congreso Investiga con Datos 2019
Experto en Power BI - CEO de eldataanalistas.com
"Soy compañero y lector de Jordi des de hace un tiempo y realmente su dedicación y ganas de transmitir conceptos complejos y muy importantes a los demás es algo único. Con Jordi conseguí entender herramientas estadísticas muy importantes con rapidez"

QUIQUE ARRANZ
Ponente congreso Investiga con Datos 2019
CEO de eusejuda.com
"Jordi cada día se supera y la motivación y ganas de ayudarme a entender la ciencia de los datos es muy difícil que las encuentres en el mercado. Si realmente quieres aprender a entender cómo transformar datos en conocimiento Jordi te puede ayudar."

MARÍA JESÚS SAMANIEGO
Alumna Analiza tus Datos - Estudios Científicos
Doctora enfermera especializada en investigación
"hasta que no conté con el asesoramiento de Jordi no comencé a comprenderla. Y eso que tuve mis dudas al principio, por si iba a malgastar el dinero. Pero ha sido como una espiral, poco a poco vas aprendiendo de una manera práctica y directa. Por eso Jordi y sus conceptos claros forman parte de los agradecimientos de mi tesis doctoral."

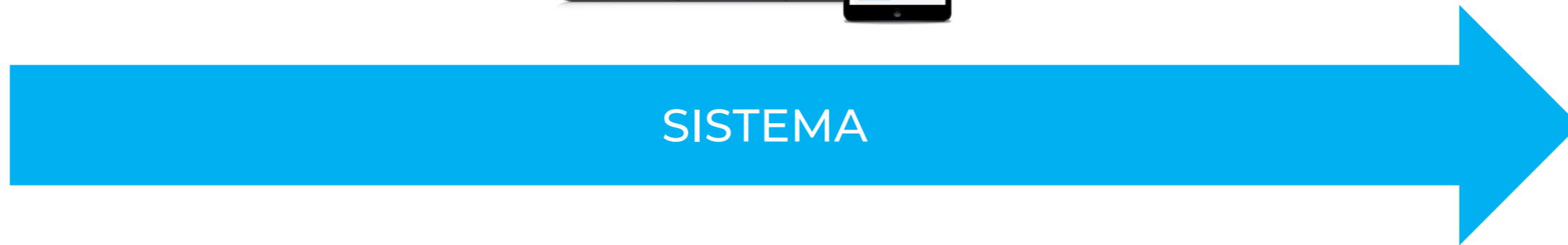
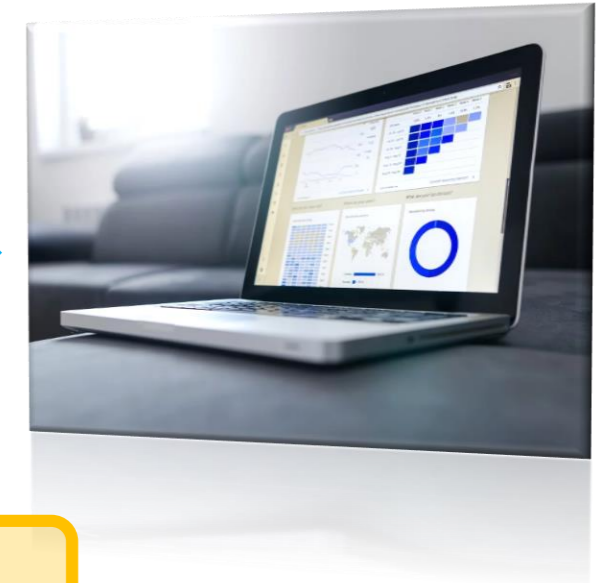
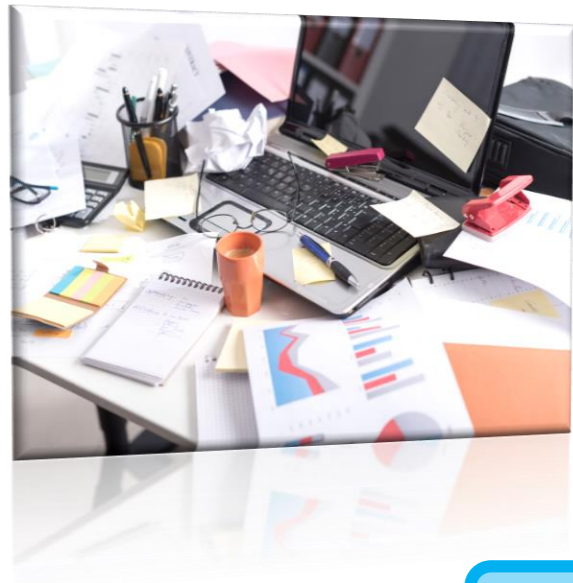
MANUEL CAÑO
Alumno Analiza tus Datos - Data Science
Análisis del entorno de la contratación pública electrónica en España
"Antes conocía la estadística y he realizado algún curso de inteligencia de negocio pero con el blog y sobre todo el curso de Jordi, he perdido el miedo a la práctica de la analítica avanzada."

JOSÉ DE JESÚS RIVERA
Alumno Analiza tus Datos - Estudios Científicos
Investigador y docente universitario
"Ahora sé la lógica que hay detrás de la estadística y de los números de mis investigaciones. Me dio el salto cualitativo que necesitaba para poder terminar mis estudios y dame cuenta de nuevas maneras de plantear nuestras preguntas de investigación"

DANI PEREZ
Ponente congreso Investiga con Datos 2019
Autor y CEO de objetivosanidad.com
"Conoció a Jordi a través de su blog conceptosclaros.com y me ha ayudado a entender conceptos importantes para analizar datos. Además con Jordi he podido participar en el congreso Investiga con Datos donde he compartido mis conocimientos sobre la figura del analista de fútbol. ¡Gracias!"

REICHE TORAO
Alumno Analiza tus Datos - Data Science
Analista de calidad y medio ambiente
"Después de pensármelo mucho compré el curso Analiza tus Datos - Completo/Data Science y realmente me está cambiando la vida. Ahora puedo ver nuevas oportunidades que antes no veía y cada vez tengo más claro qué tengo que hacer con los datos y las posibilidades que existen. Además con las sesiones con Jordi he conseguido aprender mucho con poca dedicación."

ZALOA SANCHEZ
Alumna Analiza tus Datos - Data Science
Ingeniera Mecánica, investigadora y profesora
"Con Jordi me ha ayudado a entender cómo aplicar técnicas estadísticas a los datos de mi doctorado. Ahora sé muy bien que mis publicaciones tienen mucho potencial y me sacará el doctorado en mucho menos tiempo al haberlo ido sola. Gracias por la dedicación"



SISTEMA

COMIENZO

BLOQUE 1
NUTRE TU
BASE
ESTADÍSTICA

BLOQUE 2
EMPIEZA
CON TU
SOFTWARE

BLOQUE 3
DEFINE TUS
OBJETIVOS

BLOQUE 4
DESCRIBE
TUS DATOS

BLOQUE 5
ANALIZA
TUS DATOS

BLOQUE 6
CIENCIA DE
DATOS

BLOQUE 7
PRESENTA
TUS
RESULTADOS

CIENTÍFICO DE
DATOS

ETAPA 1
RAICES PROFUNDAS

ETAPA 2
ANÁLISIS CIENTÍFICOS

ETAPA 3
DATA SCIENCE

HABILIDAD CIENCIA DE DATOS

