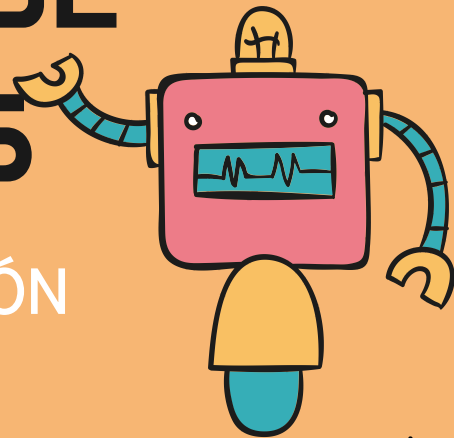
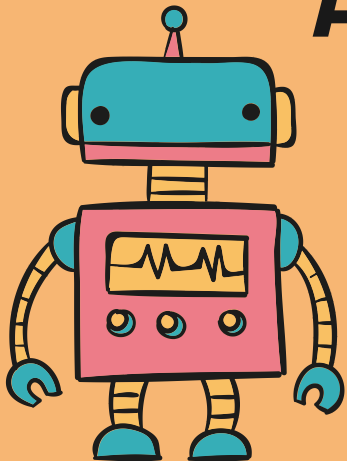


ARMADO Y USO DE BASE DE DATOS

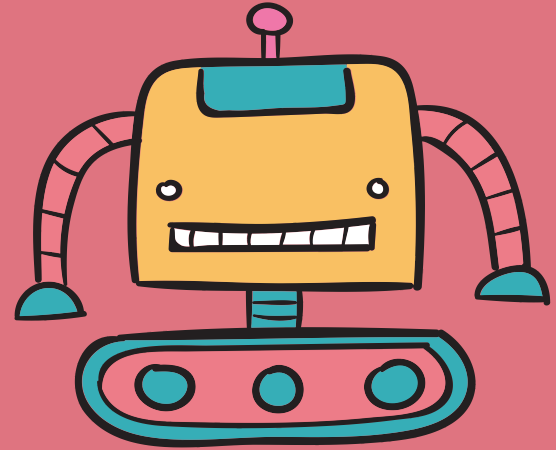
PARA PROYECTOS DE INVESTIGACIÓN



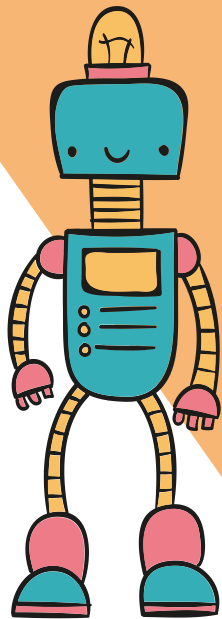
04.

Análisis y Visualización de Datos

Análisis exploratorio de datos. Herramientas y técnicas de análisis de datos. Visualización de datos: gráficos, tablas y representaciones visuales. Interpretación de resultados. Ejercicios prácticos de análisis y visualización de datos.



Lo que veremos en esta clase



01 Análisis Exploratorios

Tipos de variable, frecuencias y visualizaciones en Excel

03 Excel vs. Claude

Replicando los análisis con inteligencia artificial

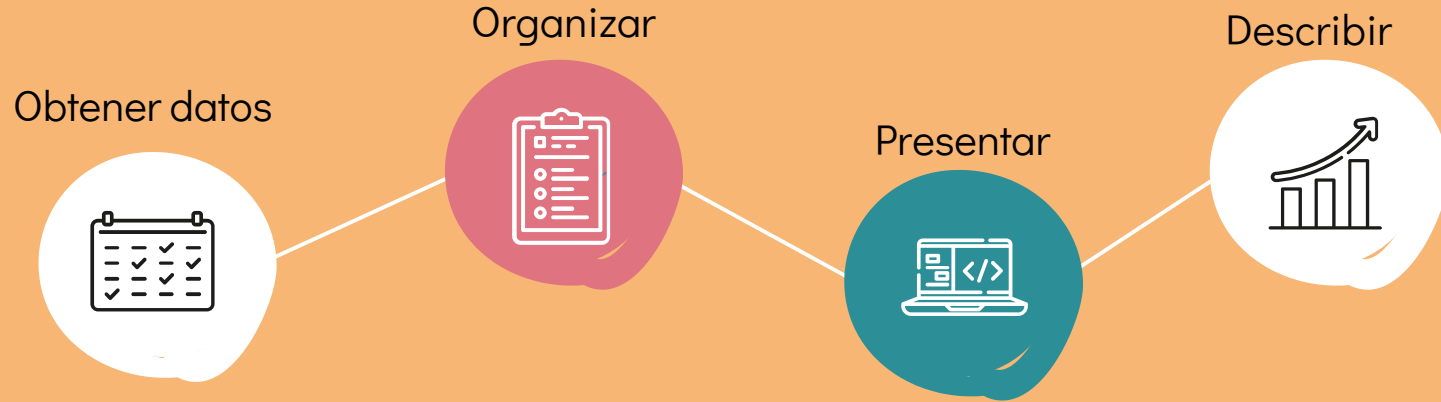
02 Prompts e IA

¿Qué es un prompt? Estructura y componentes básicos

04 Análisis Cualitativo

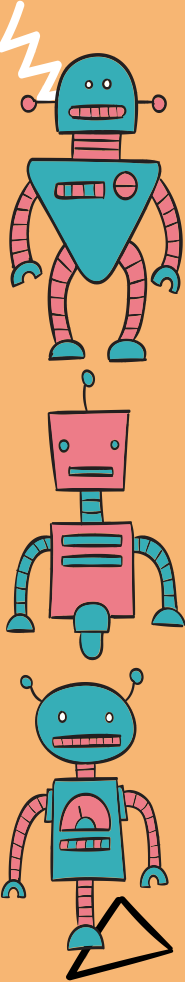
De la pregunta abierta al análisis asistido por IA

Estadística Descriptiva



Datos brutos → **Información**

+



Estadística Descriptiva

Indicadores

$$\frac{\sum ni}{n} = \bar{x}$$

Tablas

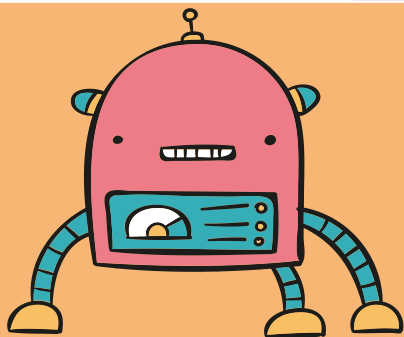
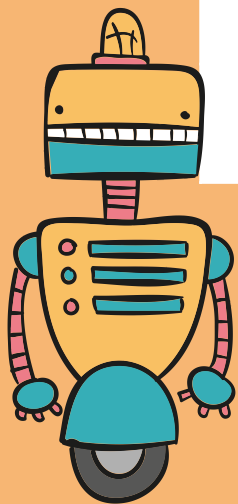
Sexo	Frec. Absoluta	Frec. Relativa	%
hembra	15	0.6	60
macho	10	0.4	40
Total	25	1	100

Gráficos



Repasando

Tipo de variable	Datos	Subtipo	Ejemplos	
CUALITATIVA	Categorías o atributos	Nominales	Sexo	macho-hembra
		Ordinales	Estadío	Larva, pupa, adulto
CUANTITATIVA	Numéricos	Discretas	Nº huevos	0, 1, 2, 3, ...
		Continuas	Longitud	0.3, 0.6, 0.5, ...



Variable Cualitativa Nominal



Indicador

- Frecuencias
- Porcentajes

Tablas

- Frecuencia absoluta
- Frecuencia relativa
- Frecuencia porcentual

Gráficos

- Barraas
- Diagrama circular

Variable Cualitativa Ordinal



Indicador

- Frecuencias
- Porcentajes



Tablas

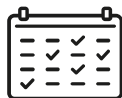
- Frecuencia absoluta
- Frecuencia relativa
- Frecuencia porcentual
- Frecuencia acumulada



Gráficos

- Barraas
- Diagrama circular?

Variable Cuantitativa Discreta



Indicador

- Medidas de tendencia central (media, mediana y moda)
- Medidas de dispersion (rango, recorrido IQ, desv. est.)
- Medidas de posición (min, max, cuartiles, percentiles)

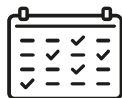
Tablas

- Frecuencia absoluta
- Frecuencia relativa
- Frecuencia porcentual
- Frecuencia acumulada

Gráficos

- Barraas
- Grafico de lineas
- Diagrama de dispersión

Variable Cuantitativa Continuas



Indicador

- Medidas de tendencia central (media, mediana y moda)
- Medidas de dispersion (rango, recorrido IQ, desv. est.)
- Medidas de posición (min, max, cuartiles, percentiles)

Tablas

- Frecuencia absoluta
- Frecuencia relativa
- Frecuencia porcentual
- Frecuencia acumulada

Gráficos

- Histograma
- Polígono de frecuencia
- Diagrama de dispersión
- Box-plot

Variables Categóricas Nominales

Base de Datos

ID	Sexo
1	macho
2	hembra
3	hembra
4	hembra
5	macho
6	hembra
7	hembra
8	macho
9	macho
10	macho
11	hembra
12	macho
13	hembra
14	macho
15	macho
16	hembra
17	macho
18	hembra
19	hembra
20	hembra
21	hembra
22	hembra
23	hembra
24	macho
25	hembra

Objetivo: determinar el porcentaje de machos eclosionados de la spp X.

Tabla de frecuencia

Sexo	Frec. Absoluta	Frec. Relativa	%
hembra	15	0,6	60
macho	10	0,4	40
Total	25	1	100

■ hembra ■ macho

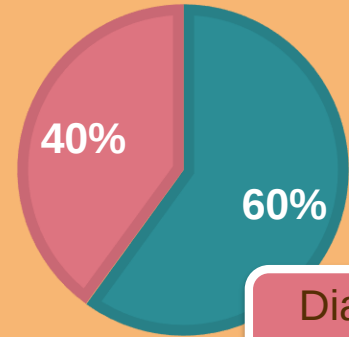
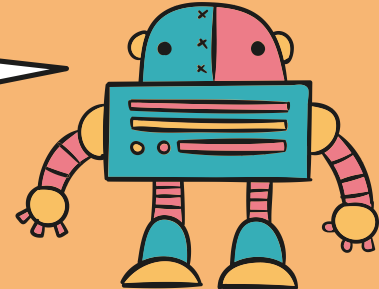


Diagrama Circular

¿Es necesario hacer un gráfico?



Variables Categóricas Nominales

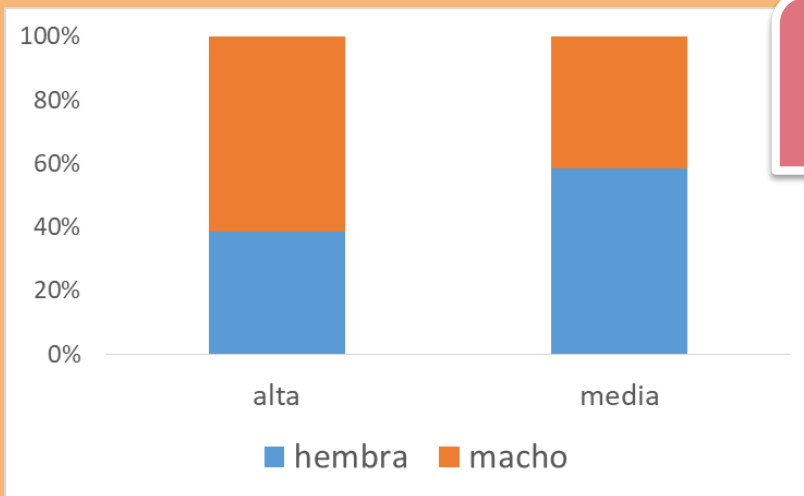
+



ID	Sexo	Temperatura
1	macho	alta
2	macho	alta
3	hembra	alta
4	macho	alta
5	macho	alta
6	hembra	alta
7	macho	alta
8	hembra	alta
9	macho	alta
10	macho	alta
11	hembra	alta
12	hembra	alta
13	macho	alta
14	hembra	media
15	hembra	media
16	macho	media
17	macho	media
18	macho	media
19	hembra	media
20	macho	media
21	macho	media
22	hembra	media
23	hembra	media
24	hembra	media
25	hembra	media

Objetivo: evaluar la influencia de temperatura en el porcentaje de machos eclosionados de la spp X.

Sexo	T° alta	%	T° media	%
hembra	5	38,5	7	58,3
macho	8	61,5	5	41,7



Columnas
100%
Apiladas



Variables Categóricas Ordinales

ID	Panta	Estadio
1	planta A	pupa
2	planta A	adulto
3	planta A	pupa
4	planta A	huevo
5	planta A	adulto
6	planta A	adulto
7	planta A	pupa
8	planta A	adulto
9	planta A	larva
10	planta A	larva
11	planta A	larva
12	planta B	pupa
13	planta B	huevo
14	planta B	larva
15	planta B	larva
16	planta B	pupa
17	planta B	larva
18	planta C	adulto
19	planta C	huevo
20	planta C	pupa
21	planta C	pupa
22	planta C	adulto
23	planta C	pupa
24	planta C	larva
25	planta C	larva

Objetivo: determinar la frecuencia de insectos según la fase de desarrollo.

Fases	planta A	planta B	planta C	planta D	Total
(I) Huevo	1	1	3	2	7
(II) Larva	3	3	5	3	14
(III) Pupa	3	2	5		10
(IV) Adulto	4		4	1	9
Total	11	6	17	6	40

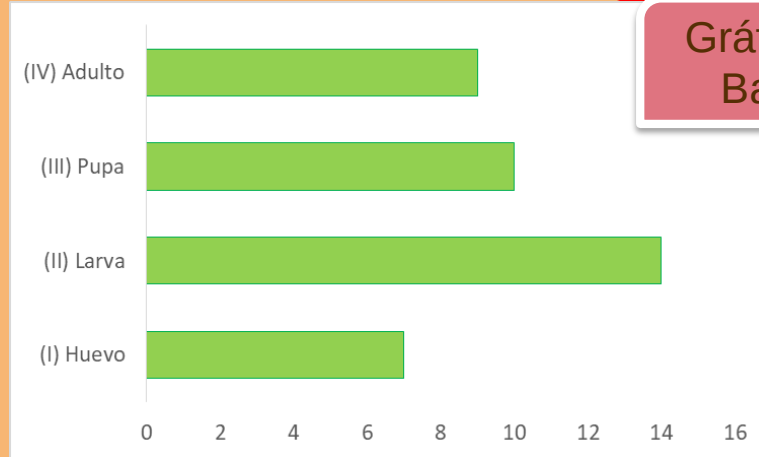
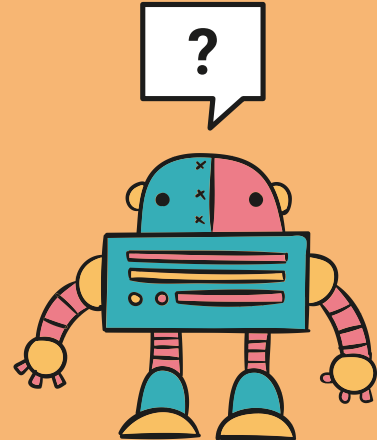
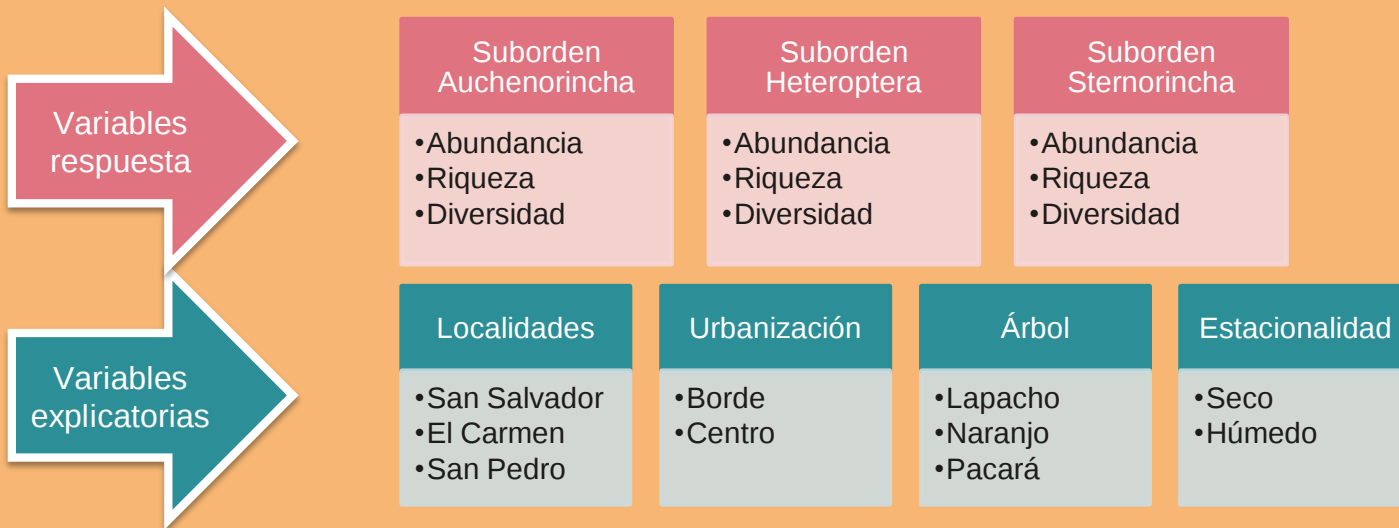


Gráfico de Barras



Variables Cuantitativas Discretas

Objetivo general: determinar y caracterizar la diversidad de Hemiptera asociada a tres especies del arbolado público de Jujuy



Variables Cuantitativas Discretas

+

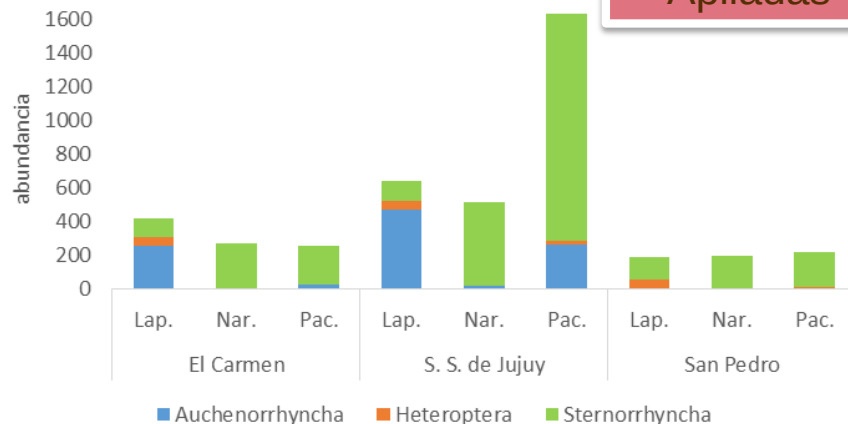
ID	n Aucheno	spp Aucheno	n Heterop	spp Heterop	n Sterno	spp Sterno	n total	spp total	Estacion	Localidad	Arbol	Urban
1	0	0	1	1	0	0	1	1	Seca	El Carmen	Lapacho	Borde
2	0	0	0	0	0	0	0	0	Seca	El Carmen	Lapacho	Borde
3	0	0	0	0	0	0	0	0	Seca	El Carmen	Lapacho	Borde
4	0	0	0	0	0	0	0	0	Seca	El Carmen	Lapacho	centro
5	0	0	1	1	45	2	46	3	Seca	El Carmen	Lapacho	centro
6	0	0	0	0	0	0	0	0	Seca	El Carmen	Lapacho	centro
7	0	0	4	2	28	5	32	7	Humeda	El Carmen	Lapacho	Borde
8	7	1	1	1	5	2	13	4	Humeda	El Carmen	Lapacho	Borde
9	110	1	2	2	7	3	119	6	Humeda	El Carmen	Lapacho	Borde
10	58	2	5	2	10	2	73	6	Humeda	El Carmen	Lapacho	centro
11	33	3	39	3	2	2	74	8	Humeda	El Carmen	Lapacho	centro
12	47	2	2	1	13	2	62	5	Humeda	El Carmen	Lapacho	centro
13	0	0	0	0	0	0	0	0	Seca	El Carmen	Naranjo	Borde
14	0	0	0	0	1	1	1	1	Seca	El Carmen	Naranjo	Borde
15	0	0	0	0	0	0	0	0	Seca	El Carmen	Naranjo	Borde

Variables Cuantitativas Discretas⁺

Objetivo específico: conocer la distribución de las abundancias de los subórdenes de Hemiptera en función de los árboles hospederos y localidad

Columnas Apiladas

Localidad	Arbol	Abundancia		
		Auchenorrhyncha	Heteroptera	Sternorrhyncha
El Carmen	Lap.	255	55	110
	Nar.	1	3	268
	Pac.	27	1	231
S. S. de Jujuy	Lap.	472	49	121
	Nar.	17	2	497
	Pac.	262	25	1345
San Pedro	Lap.	0	57	130
	Nar.	2	1	190
	Pac.	8	3	207



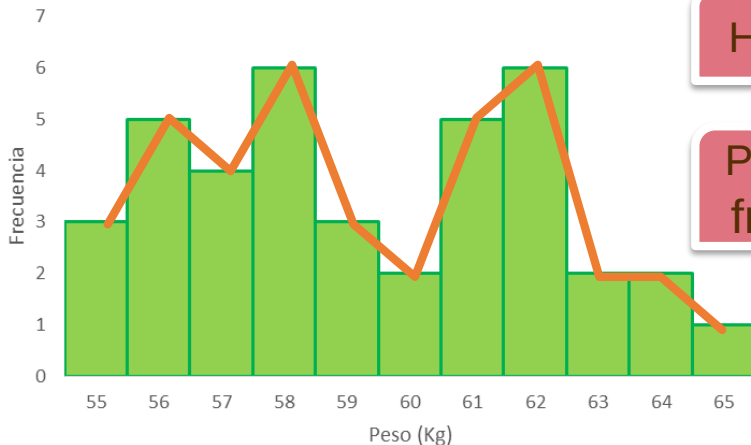
Variables Cuantitativas Continuas

Objetivo: determinar la frecuencia de pesos según el sexo en *Tayassu pecari*

Individuo	peso	Sexo	altura	score
1	56,50	hembra	45	2
2	63,12	macho	54	6
3	58,87	hembra	48	4
4	57,64	hembra	50	2
5	63,06	macho	53	4
6	61,67	hembra	54	6
7	60,50	macho	52	3
8	62,00	macho	52	6
9	65,42	macho	56	2
10	62,17	macho	58	5
11	61,62	macho	51	5
12	57,41	hembra	45	3
13	58,00	macho	46	4
14	59,74	hembra	54	3
15	62,44	macho	53	2
16	57,50	macho	49	5
17	62,67	macho	55	2
18	61,16	macho	51	4
19	62,04	macho	57	6
20	58,55	hembra	53	4
21	58,60	hembra	52	6
22	64,72	macho	55	1
23	64,41	macho	57	3
24	61,02	hembra	55	5
25	61,03	macho	50	1
26	56,18	hembra	48	5
27	55,51	hembra	43	1
28	60,00	hembra	48	5
29	58,54	macho	49	2

Peso (Kg)	Frec.
55-56	3
56-57	5
57-58	4
58-59	6
59-60	3
60-61	2
61-62	5
62-63	6
63-64	2
64-65	2
65-66	1

Agrupo los datos en intervalos de igual longitud: "intervalos de clase"



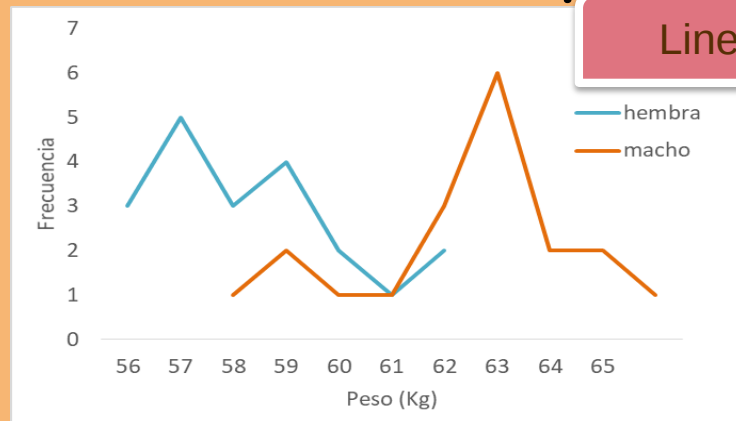
Histograma

Polígono de frecuencia

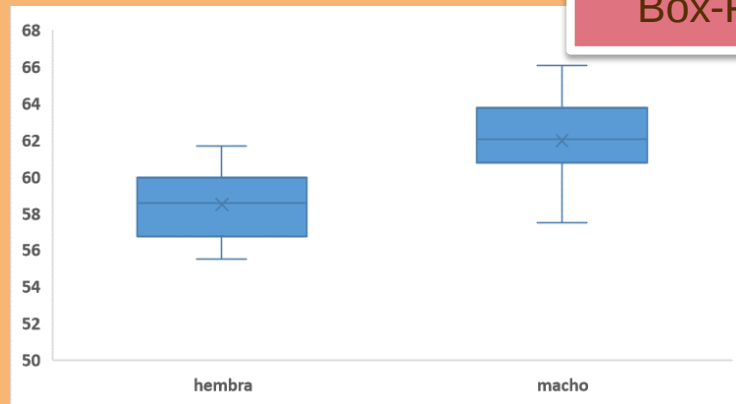
Variables Cuantitativas Continuas

Objetivo: determinar la frecuencia de pesos según el sexo en *Tayassu pecari*

Peso (Kg)	hembra	macho
55-56	3	
56-57	5	
57-58	3	1
58-59	4	2
59-60	2	1
60-61	1	1
61-62	2	3
62-63		6
63-64		2
64-65		2
65-66		1



Lineas



Box-Plot

Relación entre dos variables cuantitativas

Individuo	peso	Sexo	altura	score
1	56,50	hembra	45	2
2	63,12	macho	54	6
3	58,87	hembra	48	4
4	57,64	hembra	50	2
5	63,06	macho	53	4
6	61,67	hembra	54	6
7	60,50	macho	52	3
8	62,00	macho	52	6
9	65,42	macho	56	2
10	62,17	macho	58	5
11	61,62	macho	51	5
12	57,41	hembra	45	3
13	58,00	macho	46	4
14	59,74	hembra	54	3
15	62,44	macho	53	2
16	57,50	macho	49	5
17	62,67	macho	55	2
18	61,16	macho	51	4
19	62,04	macho	57	6
20	58,55	hembra	53	4
21	58,60	hembra	52	6
22	64,72	macho	55	1
23	64,41	macho	57	3
24	61,02	hembra	55	5
25	61,03	macho	50	1
26	56,18	hembra	48	5
27	55,51	hembra	43	1
28	60,00	hembra	48	5
29	58,54	macho	49	2

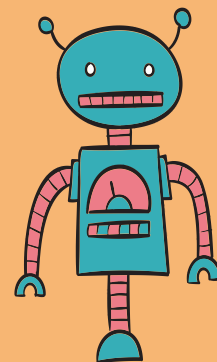
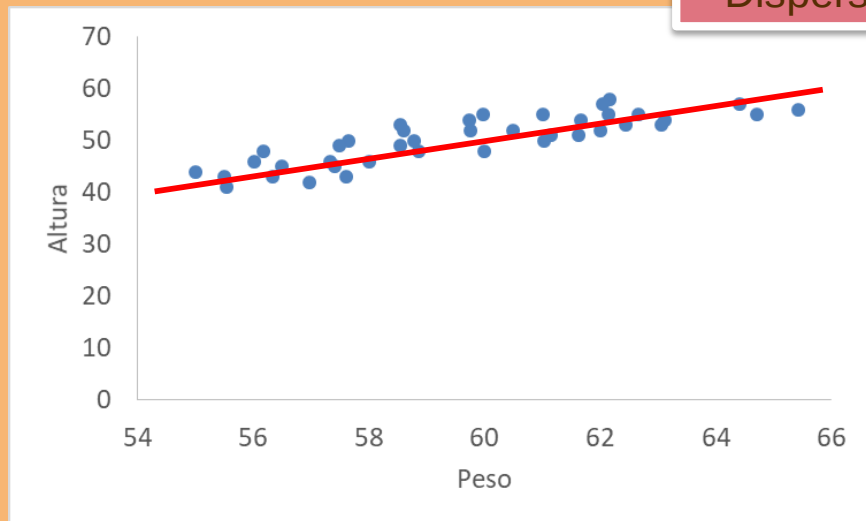


Diagrama de Caja y Bigotes (Box-Plot)

cobertura	ambiente	estacion
87,91	BS	Humedo
86,29	BS	Humedo
89,88	BS	Humedo
89,49	BS	Humedo
89,49	BS	Humedo
93,52	BS	Humedo
89,41	BS	Humedo
88,27	BS	Humedo
89,88	BS	Humedo
89,85	BS	Humedo
82,93	BS	Seco
86,49	BS	Seco
86,71	BS	Seco
86,05	BS	Seco
81,93	BS	Seco
81,15	HF	Humedo
90,75	HF	Humedo
87,59	HF	Humedo
86,17	HF	Humedo
88,86	HF	Humedo
87,25	HF	Humedo

Objetivo: determinar el porcentaje de cobertura según la estación y el ambiente

Diagrama de Caja y Bigotes (Box-Plot)

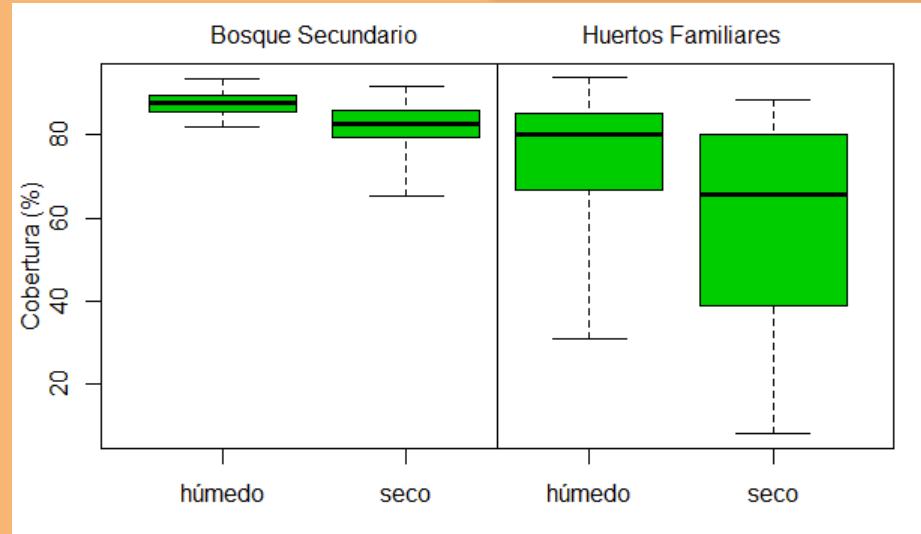
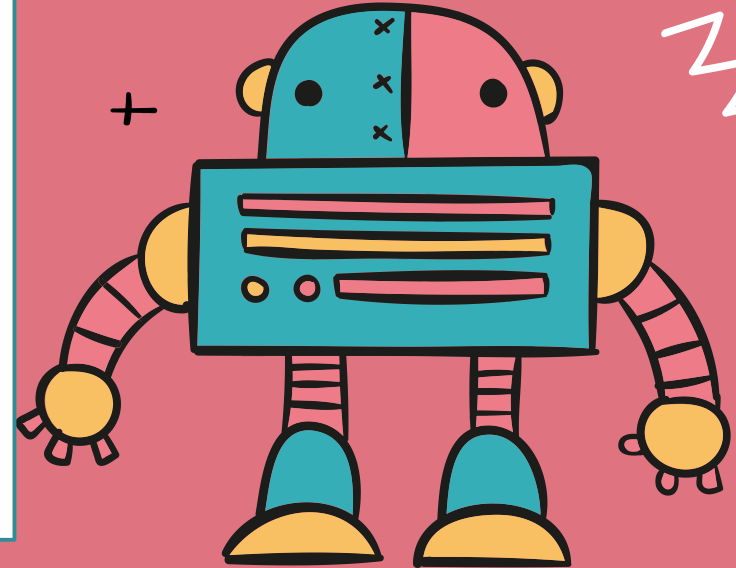
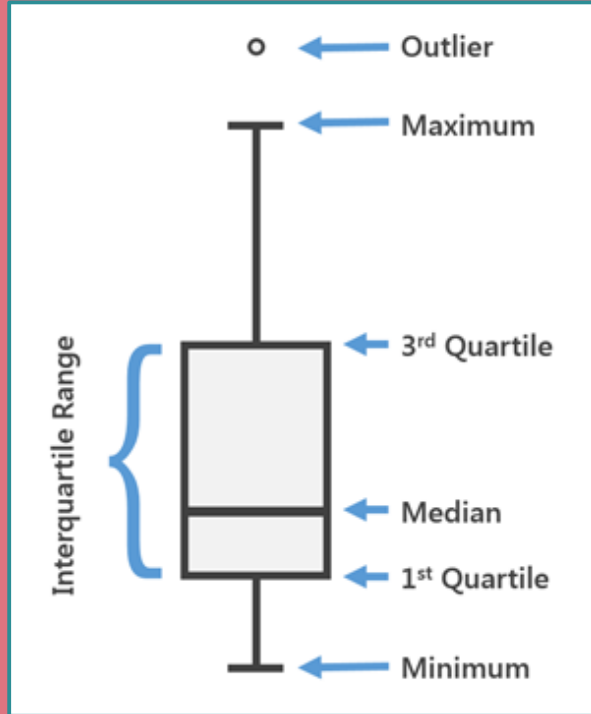
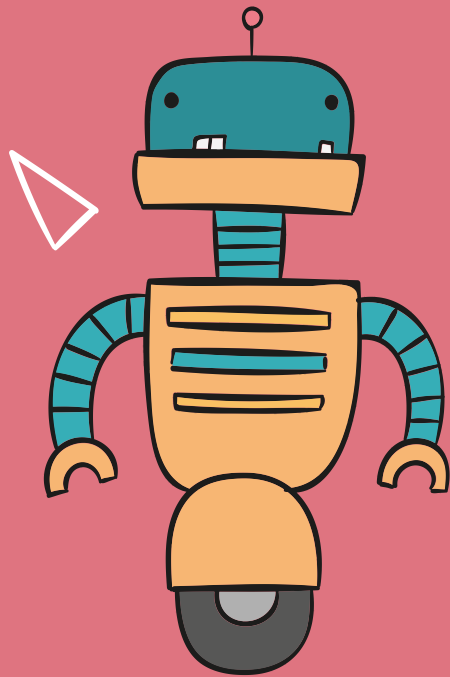


Diagrama de Caja y Bigotes (Box-Plot)





Hagámoslo con Excel

Trabajaremos con el archivo BD_graficos.xlsx — 4 casos en vivo

Los casos restantes: BD_graficos_resueltos.xlsx + videos tutoriales en el aula

Inteligencia Artificial

y Prompting

Herramientas de IA para el análisis de datos

¿Qué es un prompt?

Definición

Un prompt es la instrucción que le das a una IA para que haga algo específico.

La calidad del prompt determina directamente la calidad del resultado.

Analogía

Como una consigna de trabajo:
vaga → resultado genérico.
Clara y detallada →
resultado preciso y útil.

¿Por qué importa?

La IA responde exactamente a lo que le pedís.
No adivina intenciones.

Aprender a construir buenos prompts es una habilidad metodológica.

Los 4 componentes de un buen prompt

1. CONTEXTO

¿Quién sos y qué estás haciendo?

"Soy investigadora y trabajo con datos de pacientes pediátricos..."

2. TAREA

¿Qué querés que haga?
Verbos concretos:
calculá, generá, resumí,
clasificá, graficá.

3. FORMATO

¿Cómo querés el resultado?
tabla, Excel, gráfico,
lista numerada, texto...

4. RESTRICCIONES

¿Qué no debe hacer?
Ignorá filas vacías.
Solo describí, no interpretes.

Prompt débil vs. prompt bien armado

PROMPT DÉBIL

"Haceme un análisis de la base de datos."

- X Sin variable específica
- X Sin formato esperado
- X Sin contexto
- X La IA inventa qué hacer
- X Resultado genérico e irrelevante

PROMPT BIEN ARMADO

"Soy investigadora. Tengo datos de pacientes pediátricos con asma (Asma.xlsx, 25 casos). La variable Sexo es nominal. Genere: 1) tabla de frecuencia absoluta, relativa y %; 2) el gráfico más adecuado con su justificación. Indícame si hay valores vacíos."

- ✓ Contexto claro
- ✓ Tarea numerada
- ✓ Formato pedido
- ✓ Restricción

Regla práctica: si tu prompt tiene menos de 2 líneas, probablemente le está faltando algo.

Excel vs. Claude: mismo análisis, dos caminos

EN EXCEL (manual)

1. Abrís BD_graficos.xlsx
2. Identificás el tipo de variable
3. Insertás tabla dinámica
4. Armás tabla de frecuencia
5. Elegís e insertás el gráfico
6. Ajustás formato y etiquetas

Tiempo estimado: 10-15 min por variable

Control total del proceso

CON CLAUDE (asistido)

1. Adjuntás BD_graficos.xlsx
2. Escribís el prompt con:
 - tipo de variable
 - qué tabla querés
 - qué gráfico y justificación
3. Revisás y ajustás si es necesario

Tiempo estimado: 2-3 min por variable

Documentás el prompt como metodología

Ambos caminos son válidos. La diferencia clave: con IA el prompt queda como registro metodológico explícito.

Análisis Cualitativo

con Inteligencia Artificial

De la pregunta abierta al análisis asistido

El caso: BD_psico — de la pregunta abierta a la metodología

La base de datos

~200 estudiantes de
Psicología y Psicopedagogía

Pregunta abierta:

"Si tendrías que sugerir
cambios a las materias de
investigación, ¿qué propondrías?"

El problema

200 respuestas en texto libre.
¿Cómo las analizamos?

Necesitamos una metodología
antes de empezar.

→ La IA puede ayudar
a construirla.

El objetivo

Interpretar y caracterizar
las respuestas según
lo que proponen los
estudiantes.

No alcanza con leer las 200:
necesitamos categorías,
patrones, frecuencias.

Pipeline: objetivo → metodología → análisis → tablas → gráficos

1

Objetivo

Definís qué
querés saber
de los datos



2

Prompt Metodología

Pedís a la IA
que proponga
un método con
antecedentes



3

Metodología

La IA genera
tipología de
perfiles con
indicadores



4

Prompt Análisis

Aplicás la
metodología
a los datos
reales



5

Tablas y Gráficos

La IA produce
tablas, gráficos,
frecuencias,
nube de palabras

Mini-práctica individual

10 minutos · tu proyecto, tu prompt

Estoy trabajando en mi proyecto de investigación sobre [tema].

Mi objetivo es [completá en una línea].

Tengo datos de tipo [cualitativos / cuantitativos / mixtos]
sobre [fuente de datos].

Buscá antecedentes en la web sobre cómo se aborda
este tipo de análisis y proponé una metodología adecuada.

Completá los corchetes con tu proyecto real y probá el prompt en Claude ahora mismo.



Contenido Adicional

(si sobra tiempo)

Dos herramientas extra para el proyecto de investigación

Reemplazar códigos con diccionario de variables (EPH)

La Encuesta Permanente de Hogares usa códigos numéricos. La IA puede traducirlos al instante.

PROMPT

Tengo una base de datos de la Encuesta Permanente de Hogares (EPH) de Argentina. Las variables están codificadas numéricamente (por ejemplo, ESTADO: 1=Ocupado, 2=Desocupado, 3=Inactivo; CH04: 1=Varón, 2=Mujer).

Te adjunto la base y el diccionario de variables (o te lo pego a continuación). Reemplazá todos los códigos numéricos por sus etiquetas descriptivas en una nueva columna junto a cada variable original. Por ejemplo:

- junto a ESTADO agregá ESTADO_etiq con 'Ocupado', 'Desocupado', etc.
- junto a CH04 agregá CH04_etiq con 'Varón', 'Mujer'.

Devolveme el resultado como un archivo Excel con ambas columnas visibles.



Este mismo prompt sirve para cualquier base con códigos: censos, encuestas propias, INDEC. Solo cambiá el diccionario.

Creación de un dashboard a partir de una base de datos

Con un solo prompt podés pedirle a la IA que proponga y genere un panel de indicadores clave.

PROMPT

Tengo una base de datos en Excel con las siguientes variables:
[describí brevemente las variables principales de tu base].

Quiero construir un dashboard de análisis que muestre de forma visual y clara los principales indicadores. El dashboard debe incluir:

- 1) Un resumen de los indicadores clave (KPIs): totales, promedios, distribuciones.
- 2) Al menos 3 gráficos que representen las variables más importantes:
uno de barras o columnas, uno de torta o dona, y uno de líneas o dispersión.
- 3) Una tabla resumen con los resultados principales.

Primero proponé qué KPIs y gráficos incluirías según los datos disponibles y por qué. Luego generá el dashboard en una hoja nueva del mismo Excel.



Para un dashboard más interactivo, podés pedirle que use segmentaciones de datos (slicers) y tablas dinámicas conectadas.

+

¡Gracias!

javialtamirano1986@gmail.com

CREDITS: This presentation template was created by [Slidesgo](#), including icons by [Flaticon](#), infographics & images by [Freepik](#)

