



Estadística

Unidad 2: Organización y resumen de datos. Primera parte

Profesora Adjunta: Sandoval, Marisa Susana

Organización y resumen de datos

Distribuciones de frecuencias simples y agrupadas.
Definiciones básicas. Histograma, polígono de
frecuencias, ojiva acumulada. Ejercicios.

Recordemos que: Un método estadístico está compuesto por un conjunto de etapas que, en cierta medida, definen su secuencia completa.

Primera etapa: Definición del Problema.

Es importante delimitar una correcta delimitación del proyecto de investigación así como de los objetivos que persigue el trabajo, con el propósito de optimizar la utilización de los recursos.

Segunda etapa: Recopilación de los datos

Esta etapa se vincula con la captación de los datos. Es preciso planificar la recopilación, atendiendo a la correcta realización de un cuestionario, periodo al que se refiere la medición, metas, etc.

Tercera etapa: Organización de los datos

Interviene en esta etapa la corrección de los datos, donde se evalúa la consistencia de los datos para evitar errores que puedan incidir en las etapas posteriores. Además, de clasificar los datos.

Cuarta etapa: Presentación de los datos

Cuando se hace referencia a la presentación de la información se suelen mencionar los siguientes modos de hacerlo:

Tablas Estadísticas

Representaciones Gráficas

Tablas Estadísticas

Ejemplo de aplicación

Incidencia de los capítulos en la variación mensual del nivel general del ICC. Julio de 2022*

Nivel general y capítulos	Incidencia en el nivel general
	%
Nivel general	6,8
Materiales	5,53
Mano de obra	0,81
Gastos generales	0,47

Nota: la incidencia mide la participación que ha tenido cada apertura en la variación del nivel general del ICC respecto al mes anterior. Los totales por suma pueden no coincidir por redondeos en las cifras parciales.

Fuente: INDEC, Dirección Nacional de Estadísticas de Precios. Dirección de Índices de Precios de la Producción.

Partes Principales de las Tablas Estadísticas

- ✓ Título: debe ser preciso y comunicar concretamente de qué trata el estudio.
- ✓ Subtítulo: Es un complemento necesario del título que aclara algunos aspectos del trabajo.
- ✓ Encabezamiento: Describe los datos en las columnas.
- ✓ Concepto: Describe los datos contenidos en las filas.
- ✓ Notas al pie: Son aclaraciones para comprender cómo se efectuó el relevamiento o algún otro aspecto.

Representación Gráfica

Las herramientas gráficas es un instrumento para la divulgación /presentación de las conclusiones. La habilidad del investigador está en elegir el que más se adecue a los objetivos de presentación.

Por su parte, en el gráfico 1 se ha registrado la distribución de los establecimientos que permanecieron abiertos durante abril según su categoría hotelera.

Establecimientos abiertos según categoría hotelera. Abril de 2020



Fuente: INDEC, Dirección de Estadísticas Básicas de la Balanza de Pagos.

Los resultados que se presentan en el cuadro 1 corresponden a la información recogida por la EOHMC de los 154 hoteles que respondieron en los plazos solicitados.

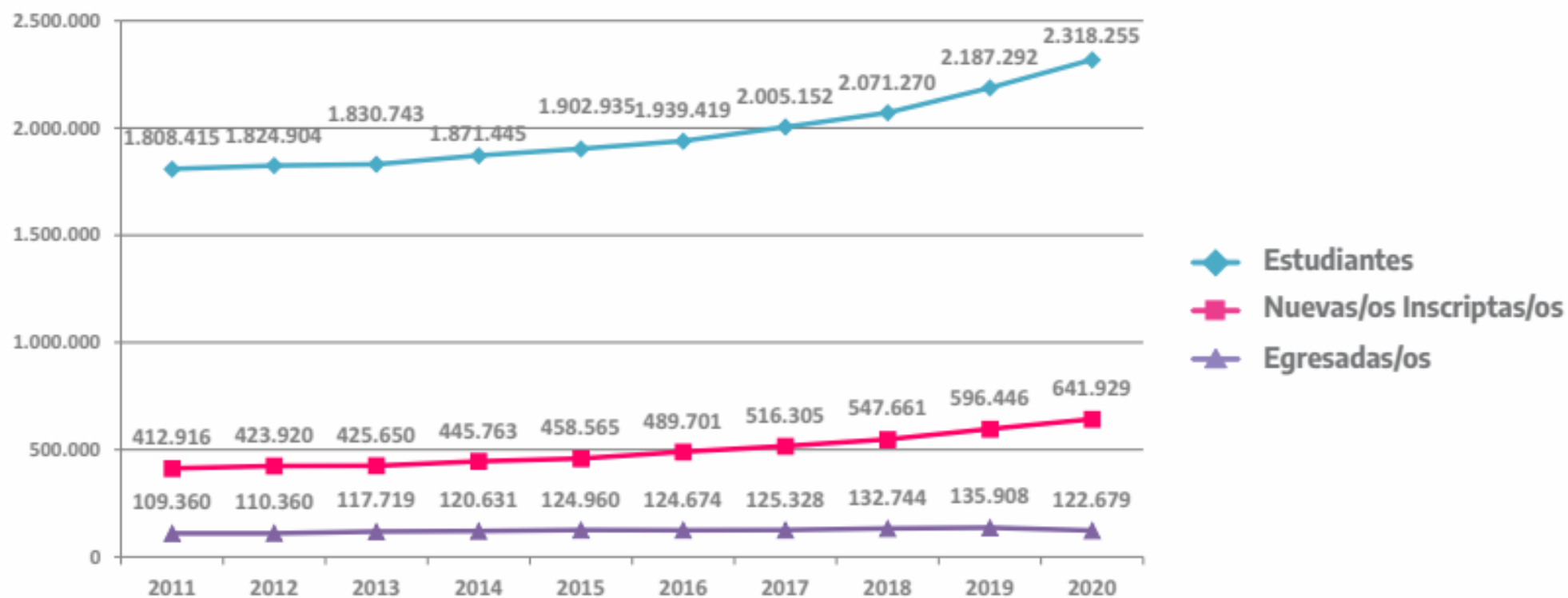
Tipos de gráficos: lineales

Gráfico lineal simple



Gráfico lineal compuesto

Evolución de la cantidad de estudiantes, nuevas/os inscriptas/os y egresadas/os de pregrado y grado. Años 2011 -2020



Fuente: Departamento de Información Universitaria – DNPelU - SPU.

Tipos de gráficos: De barras

Se utilizan rectángulos cuyas bases deben ser iguales. De esta manera, las magnitudes que se representan son proporcionales a la altura.

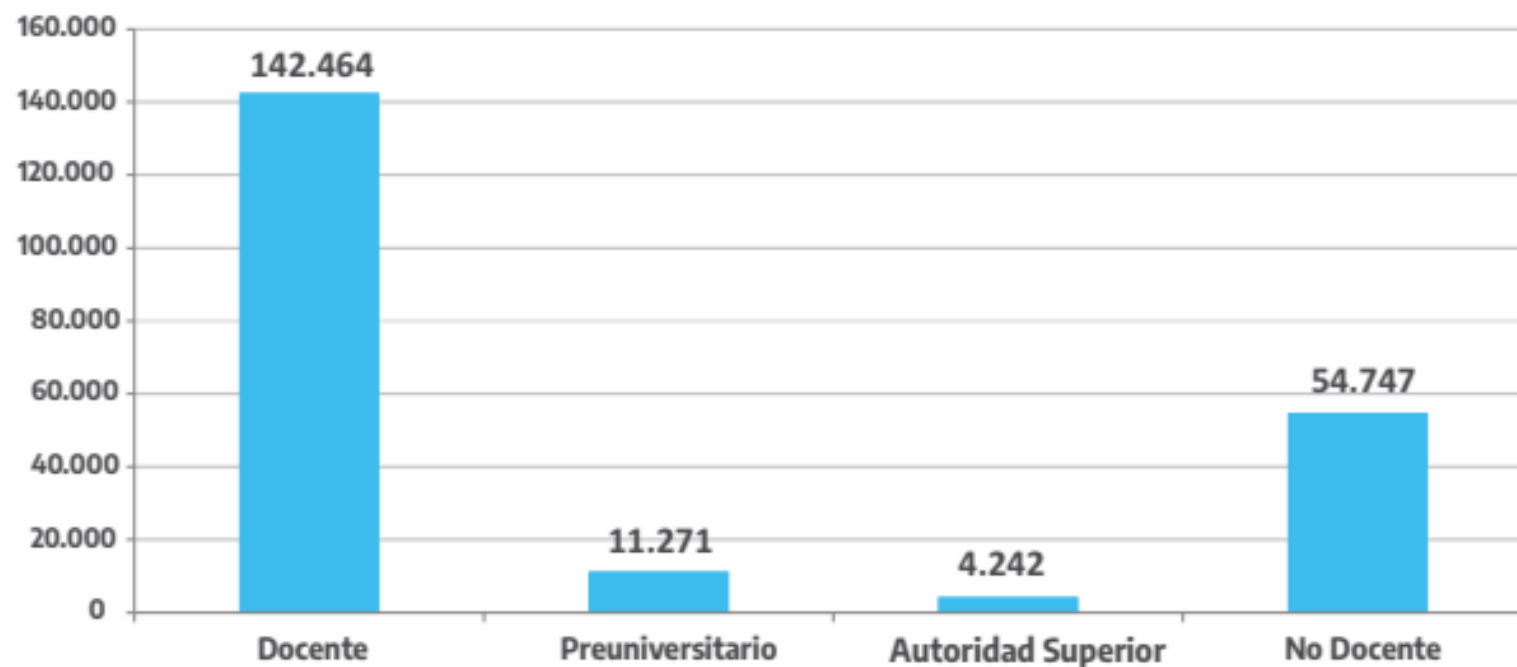
Criterios a tener en cuenta para la construcción son:

- Todas las barras de igual ancho.
- La longitud de cada barra representa la magnitud del dato.
- Las barras no deben ser excesivamente cortas y anchas, ni largas ni angostas.
- El espacio entre las barras debe ser menor que el ancho de las mismas.
- Las barras pueden ser verticales u horizontales.

Casos

Gráfico de barras simples (verticales)

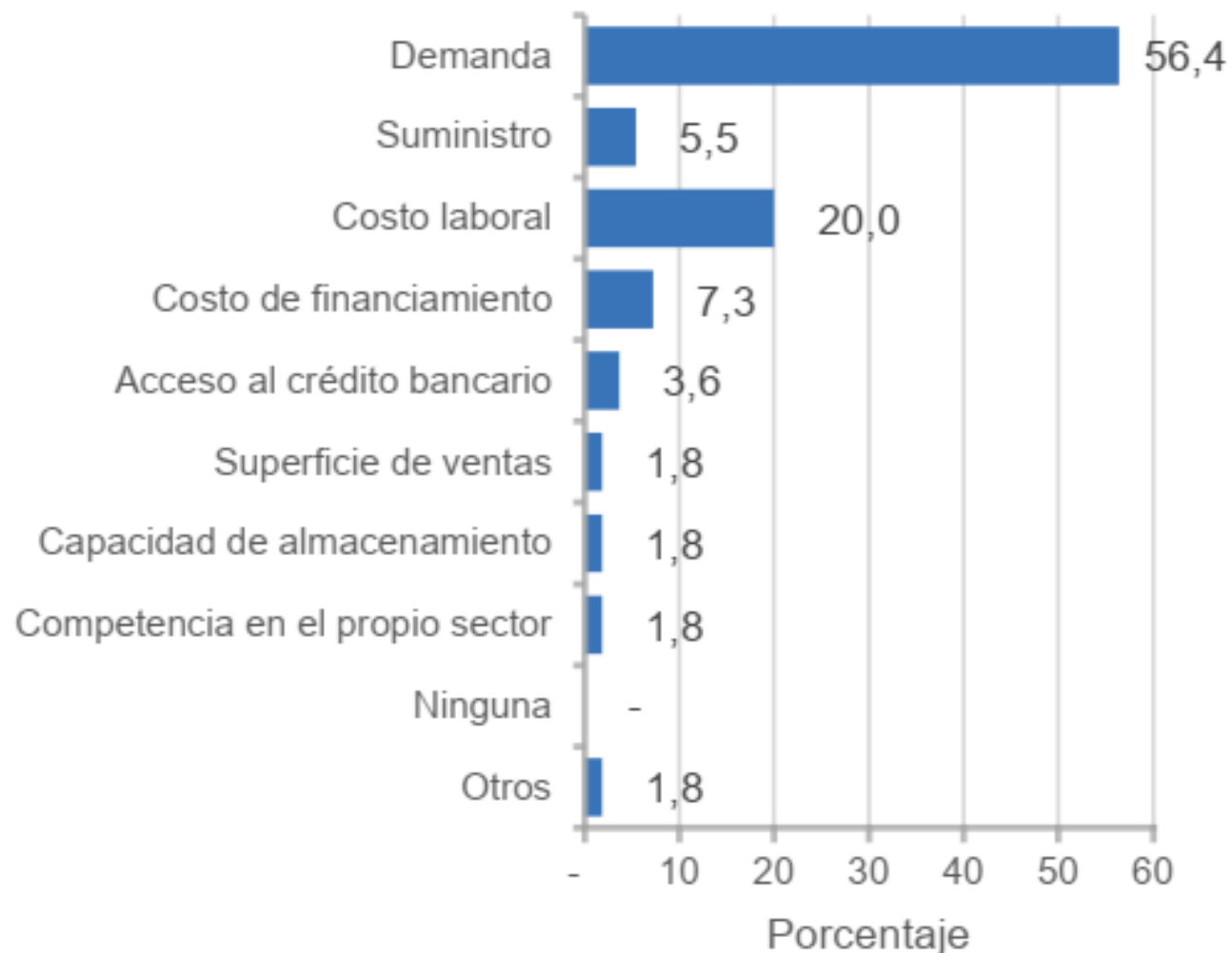
Cantidad de personas según escalafón.
Año 2021



Fuente: Sistema RHUN (Recursos Humanos de Universidades Nacionales), correspondiente al mes de septiembre 2021.
Departamento de Información Universitaria – DNPelU – SPU.-

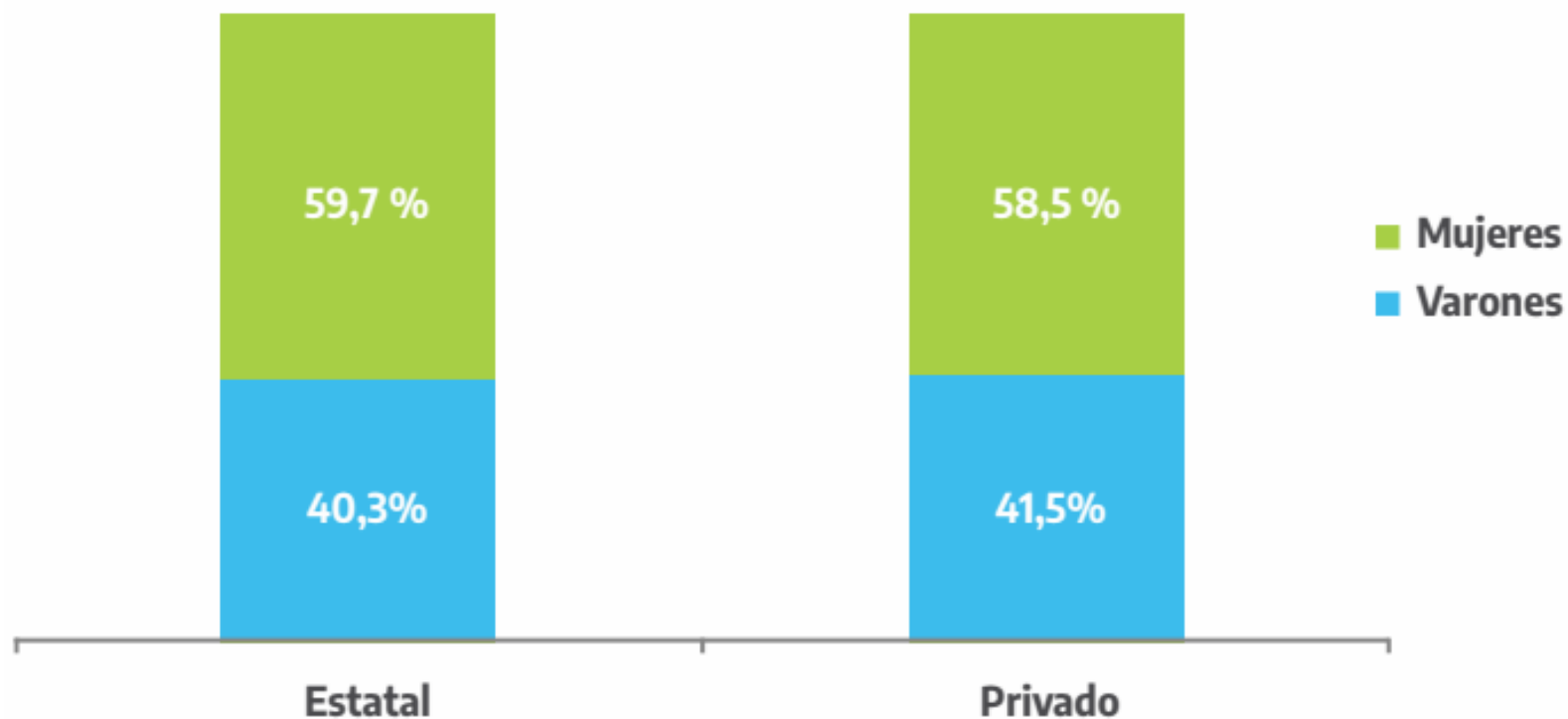
Gráfico de barras simples (horizontales)

¿Qué factores limitan su capacidad para aumentar la actividad comercial?



Gráficos de barras subdivididas

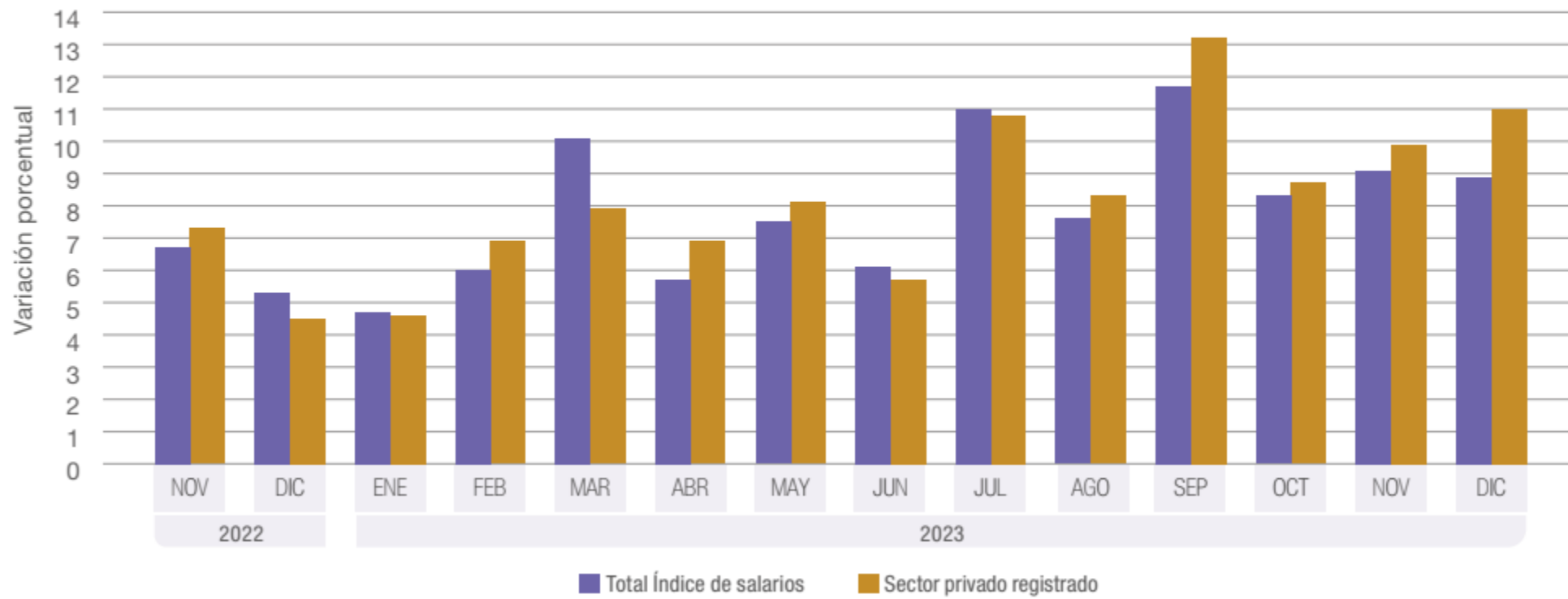
Estudiantes de pregrado y grado por género
según sector de gestión. Año 2020



Fuente: Departamento de Información Universitaria – DNPeIU - SPU.

Gráfico de barras agrupadas

Variación porcentual respecto al mes anterior. IS total y sector privado registrado

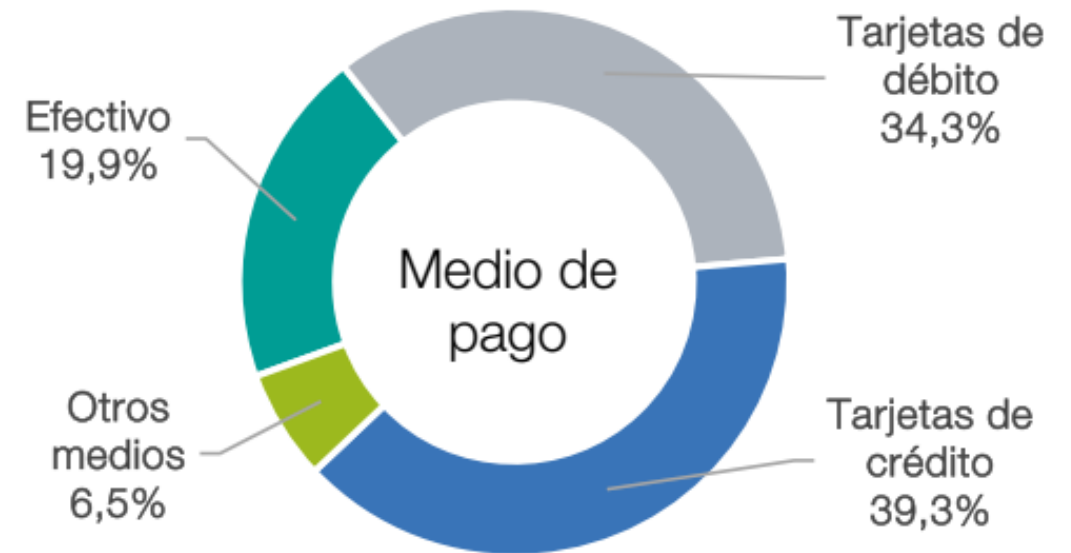


Tipos de gráficos: de sectores

El gráfico de sectores (tortas) es útil para representar contribuciones porcentuales a un total o la distribución de frecuencias de una variable cualitativa.

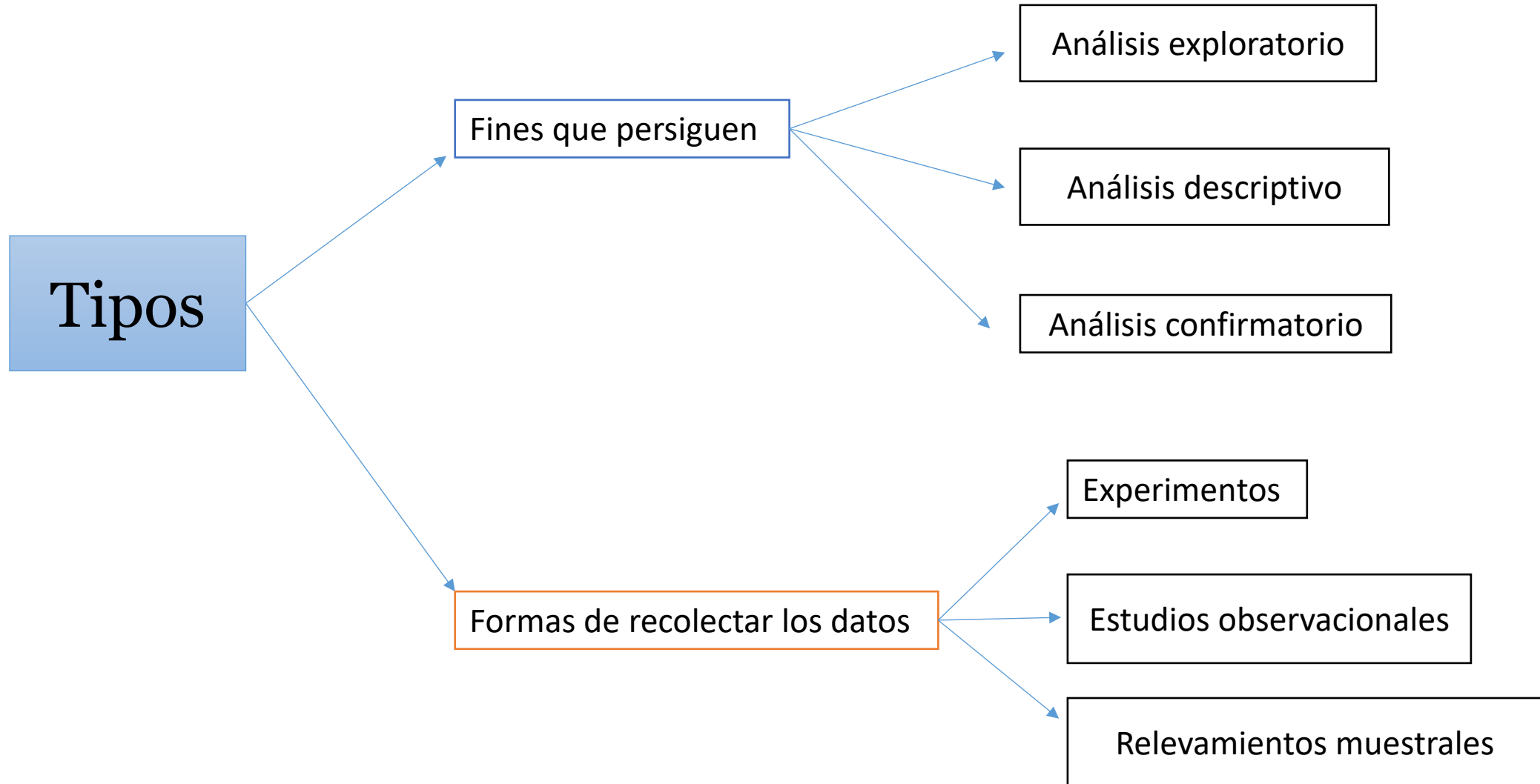
Instituto Nacional de Estadística y Censos –INDEC.
Comercio. Vol.8, nº 5. Encuesta de supermercados.
Diciembre de 2023,

Gráfico 3. Ventas totales a precios corrientes por medio de pago, composición porcentual.
Diciembre de 2023



Fuente: INDEC, Dirección Nacional de Estadísticas Económicas.
Dirección de Estadísticas de Comercio y Servicios.

Tipos de investigación



Niveles de medición

Tipo de variable	Escala	Propiedades matemáticas
Cualitativa	Nominal	1) Equivalencia
	Ordinal	1) Equivalencia 2) Jerarquía
Cuantitativa	Intervalo	1) Equivalencia 2) Jerarquía 3) Distancia 4) Proporción de intervalos
	De Razón	1) Equivalencia 2) Jerarquía 3) Distancia 4) Proporción de intervalos 5) Proporción entre dos valores

Observar detenidamente estas unidades de análisis



Método de
relevamiento:
encuesta

Variables

Nivel de medición

Conjunto de datos

1	Nivel de instrucción alcanzado	Unidad de Gestión	Agrupamiento	Área	Antigüedad laboral	¿Cuántas capacitaciones realizó?
2	Nivel universitario completo	UARG	Técnico-Profesional	Mantenimiento de Tecnologías	23	10
3	Nivel universitario completo	UARG	Administrativo	Asistencia Técnica y Reglamentaria	18	10
4	Nivel terciario completo	Rectorado	Administrativo	Secretaría de Ciencia y Tecnología	3	10
5	Nivel universitario completo	Rectorado	Técnico-Administrativo	Educación a Distancia	12	7
6	Nivel universitario incompleto	Rectorado	Administrativo	Educación a Distancia	3	13
7	Nivel universitario completo	UART	Administrativo	Secretaría de Extensión	16	10
8	Nivel universitario incompleto	UART	Administrativo	Secretaría Académica	17	6
9	Nivel universitario completo	Rectorado	Técnico-Profesional	Secretaría de Extensión Universitaria	13	3
10	Nivel universitario completo	UASJ	Administración	Secretaría Privada Decanato	1	4
11	Nivel universitario completo	UARG	Técnico profesional	Secretaría de Extensión	23	4
12	Nivel universitario completo	Rectorado	Técnico-Profesional	Productora de Contenidos Audiovisuales	1	4
13	Nivel universitario completo	Rectorado	Mantenimiento, producción y servicios generales	servicios generales	20	7
14	Nivel universitario incompleto	UART	Administrativo	Departamento de Personal	19	4
15	Nivel universitario completo	UACO	Administrativo	Asistencia legal y Técnica	7	51
16	Nivel universitario incompleto	UARG	administrativo	tesorería	18	6
17	Nivel universitario incompleto	UARG	Administrativo	Administración de Personal	25	3
18	Nivel universitario completo	UARG	Administrativo	Secretaría de Administración	6	5

Organización y Presentación

Datos Cualitativos

Variables Cualitativas

Nivel de medición nominal

Pregunta de encuesta: Unidad de gestión

Rectorado UARG Rectorado Rectorado UARG Rectorado Rectorado

UARG Rectorado Rectorado UARG Rectorado Rectorado

Rectorado UASJ Rectorado UARG Rectorado Rectorado

Rectorado UARG Rectorado UARG Rectorado UARG

UARG Rectorado UARG Rectorado UARG Rectorado

Rectorado UACO Rectorado UARG Rectorado UARG

UARG Rectorado UASJ Rectorado UARG Rectorado

Rectorado Rectorado UARG Rectorado UARG Rectorado

Rectorado UARG Rectorado UARG Rectorado UARG

UARG Rectorado UARG Rectorado UARG

Ejemplo: Distribución de la cantidad de cursantes organizados de acuerdo a la unidad de gestión.

Unidad	Cursantes
Rectorado	23
UACO	3
UARG	10
UART	3
UASJ	3
Total	42

Operador sumatoria: Σ

El símbolo Σ , sumatoria, se utiliza con frecuencia en las fórmulas estadísticas para facilitar la escritura.

Por ejemplo: en lenguaje coloquial se pide sumar números naturales de 1 hasta 10

En símbolos:

$$x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4, x_5 = 5, \\ x_6 = 6, x_7 = 7, x_8 = 8, x_9 = 9, x_{10} = 10$$

Con el operador sumatoria:

$$\sum_{i=1}^{10} x_i = 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 = 55$$

Frecuencia absoluta

El número de veces que se repite cada una de las k categorías de la variable se denomina frecuencia absoluta. Para hacer referencia a este tipo de frecuencias, se la representará con la notación f_i

La ley de cierre que debe verificar establece:

$$\sum_{i=1}^k f_i = n$$

Ejemplo 1: Distribución de la cantidad de cursantes organizados de acuerdo a la unidad de gestión.

Unidad (x_i)	Cursantes (f_i)
Rectorado	23
UACO	3
UARG	10
UART	3
UASJ	3
Total (n)	42

Frecuencia relativa

La frecuencia relativa de una categoría es la proporción de veces que ocurre dicha categoría. Se denota como:

$$f_{ri} = \frac{f_i}{n}$$

Se debe verificar la ley de cierre

$$\sum_{i=1}^k f_{ri} = 1$$

Tablas de frecuencias

Una tabla de frecuencias para variables cualitativas es una tabla que asocia a cada categoría de la variable con el número de veces que se repite dicha categoría.

X	f_i	f_{r_i}
x_1	f_1	f_{r_1}
x_2	f_2	f_{r_2}
...
x_k	f_k	f_{r_k}
total	n	1,00

Completando la información para la tabla anterior

Unidad (x_i)	Cursantes (f_i)	f_{r_i}	$f_{r_i}\%$
Rectorado	23	$23/42=0,55$	55
UACO	3	$3/42=0,07$	7
UARG	10	$10/42=0,24$	24
UART	3	$3/42=0,07$	7
UASJ	3	$3/42=0,07$	7
Total (n)	42	1,00	100

El orden con el que se presenta la información, ¿lo consideran claro para comunicar?



Criterios de orden para variables
cualitativas con nivel de medición
nominal

Criterios de orden

- Magnitud de la variable
- Alfabético
- Convencional
- Geográfico

¿Qué criterio se utilizó para ordenar las categorías?

Unidad (x_i)	Cursantes (f_i)	f_{r_i}	$f_{r_i}\%$
Rectorado	23	$23/42=0,55$	55
UARG	10	$10/42=0,24$	24
UACO	3	$3/42=0,07$	7
UART	3	$3/42=0,07$	7
UASJ	3	$3/42=0,07$	7
Total (n)	42	1,00	100



¿Cómo se interpretan las frecuencias en el contexto del problema?

¿Qué información brinda cada celda pintada?

Unidad (x_i)	Cursantes (f_i)	f_{r_i}	$f_{r_i}\%$
Rectorado	23	$23/42=0,55$	55
UARG	10	$10/42=0,24$	24
UACO	3	$3/42=0,07$	7
UART	3	$3/42=0,07$	7
UASJ	3	$3/42=0,07$	7
Total (n)	42	1,00	100

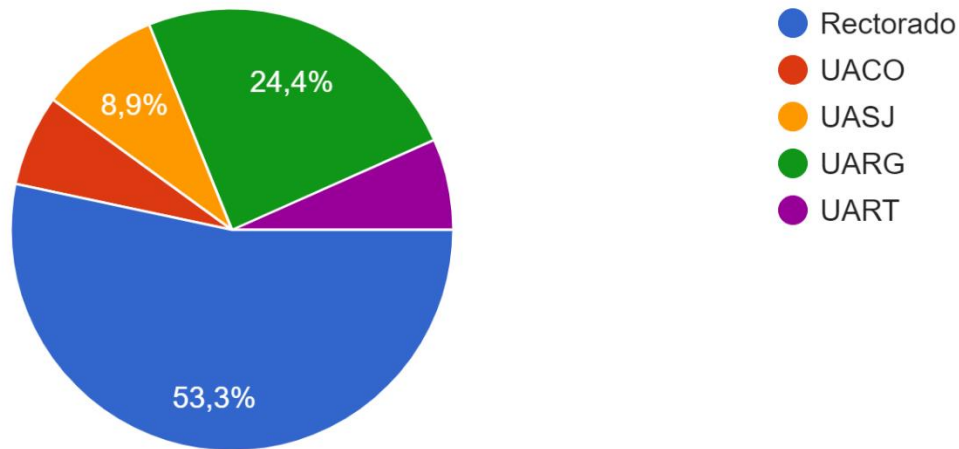


¿Qué gráfico resulta admisible para este tipo de variable?

Considerar el nivel de medición

Gráfico circular que brinda el formulario Google

Unidad de Gestión
45 respuestas



ÁNGULO CENTRAL QUE
CORRESPONDE A CADA
SECTOR CIRCULAR , SE
CALCULA MEDIANTE LA
SIGUIENTE FÓRMULA DE
TRABAJO:

$$\text{Ángulo central} = \frac{f_i \cdot 360^\circ}{n}$$

Variables Cualitativas

Nivel de medición ordinal

Frecuencias acumuladas

Distribución de frecuencias en las que se consignan como frecuencias de cada valor o categoría de la variable, la suma de su frecuencia absoluta con las frecuencias absolutas de todos los valores o categorías de la variable que la preceden en la distribución.

Tabla de frecuencia admisible para el tipo y nivel de medición

X	f_i	f_{r_i}	F_i	F_{r_i}
x_1	f_1	f_{r_1}	f_1	f_{r_1}
x_2	f_2	f_{r_2}	$f_1 + f_2$	$f_{r_1} + f_{r_2}$
...
x_k	f_k	f_{r_k}	n	1,00
total	n	1,00		

Ejemplo 2: El Censo Nacional de Población, Hogares y Viviendas 2001 ha relevado diferentes variables sobre características educacionales de la población. En este caso se presenta la siguiente información:

- a- Analice la estructura de la tabla. Determine variable, unidades de análisis y el tamaño del universo analizado. Indique la clasificación y nivel de medición de la variable
- b- Establezca el criterio de ordenamiento con el cual se presenta la información, ¿es adecuado?
- c- Complete la distribución de frecuencias

Máximo nivel de instrucción alcanzado	Habitantes
Sin instrucción	1.774
Primario incompleto	6.625
Primario completo	13.548
Secundario incompleto	18.379
Secundario completo	12.183
Superior no universitario incompleto	1.750
Superior no universitario completo	3.115
Superior universitario incompleto	3.390
Superior universitario completo	2.716

Departamento Güer Aike, Provincia de Santa Cruz. Población de 15 años o más según máximo nivel de instrucción alcanzado. Año 2001

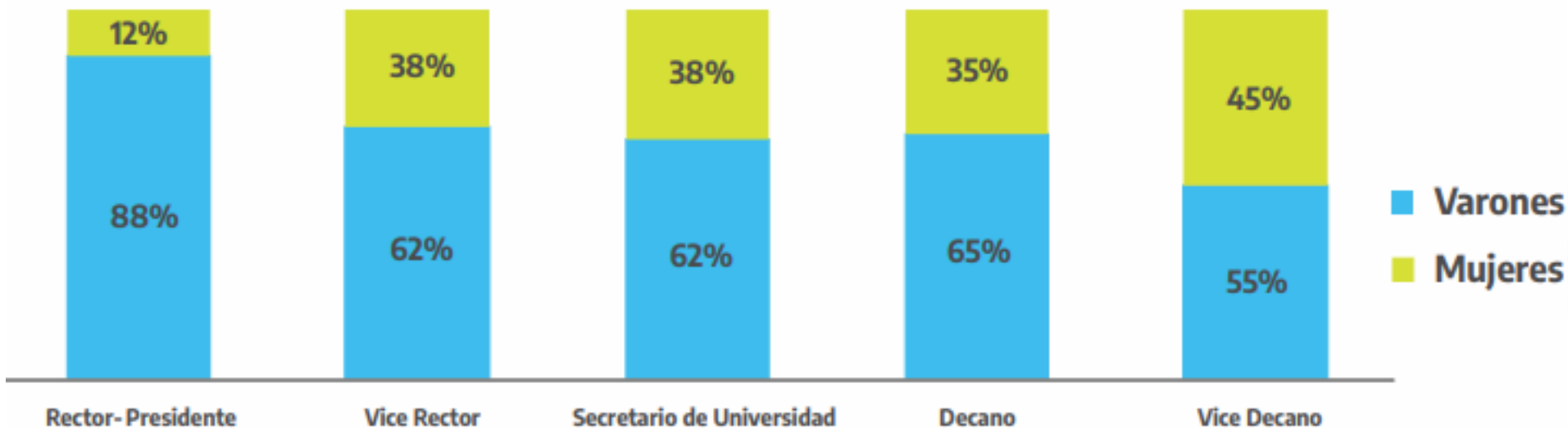
Fuente: INDEC. Censo Nacional de Población, Hogares y Viviendas 2001

Completando la información para la tabla anterior,
se genera la siguiente distribución

Máximo nivel de instrucción alcanzado	Habitantes	fri%	Fi	Fri%
Sin instrucción	1.774	2,8	1.774	2,8
Primario incompleto	6.625	10,4	$1.774 + 6.625 = 8.399$	13,2
Primario completo	13.548	21,3	$8.399 + 13.548 = 21.947$	34,5
Secundario incompleto	18.379	29,0	$21.947 + 18.379 = 40.326$	63,5
Secundario completo	12.183	19,2	$40.326 + 12.183 = 52.509$	82,7
Superior no universitario incompleto	1.750	2,8	$52.509 + 1.750 = 54.259$	85,5
Superior no universitario completo	3.115	4,9	$54.259 + 3.115 = 57.374$	90,4
Superior universitario incompleto	3.390	5,3	$57.374 + 3.390 = 60.764$	95,7
Superior universitario completo	2.716	4,3	$60.764 + 2.716 = 63.480$	100
Total	63,480	100		

Gráficas para datos categóricos con nivel de medición ordinal

Distribución porcentual de cargos de autoridades superiores según categoría y género. Año 2021



Fuente: Sistema RHUN (Recursos Humanos de Universidades Nacionales) , correspondiente al mes de septiembre de 2021. Departamento de Información Universitaria - DNPeIU – SPU.

