

Facultad de Ciencias Sociales
Escuela de Sociología
Estadística I
Prof. Paulina Lizama

GUÍA EJERCICIOS

1. Para cada una de las tablas siguientes, identifique:

- i. Variables involucradas**
- ii. Unidad de análisis y observación**
- iii. Nivel y escala de medición.**

- a) Independientemente de su posición política, ¿Ud. aprueba o desaprueba la forma como Michelle Bachelet y su equipo económico está manejando la economía?

	SEXO		POSICIÓN POLÍTICA			
	Hombres	Mujeres	Der. / C. Der.	Centro	Izq. / C. Izq.	NC/NS
Aprueba	35	33	18	35	56	22
Desaprueba	42	42	68	45	28	41

Fuente: CEP, Encuesta Nacional de Opinión Pública, Junio 2007

Variables Involucradas	Unidad de análisis (UA)	Unidad de Observación (UO)	Nivel de Medición	Escala de Medición
¿Ud. aprueba o desaprueba la forma como Michelle Bachelet y su equipo económico está manejando la economía?	La encuesta CEP, es una encuesta de opinión, por lo que la UA corresponde a las personas mayores de 18 años habitantes de Chile.	En las encuestas de opinión las UA y la UO coinciden. UO: personas mayores de 18 años habitantes de Chile, exceptuando Isla de Pascua	Cualitativa dicotómica	nominal, No Aplica
Sexo			Cualitativa dicotómica	nominal, No Aplica
Posición Política			Cualitativa ordinal	No Aplica

b)

Tabla 6. Diagnósticos informados en licencia médicas curativas sector minería

Autor-Empresa	Respiratorias	Osteomuscular	Digestivo	Mentales	Trauma	Total
Cantuarias ¹ Minera El Indio	28,9	18,8	19,6	8,0	8,3	83,6
Cantuarias ² Minera El Indio	23,8	17,5	23,9	6,1	12,2	83,5
Jiménez División Andina	12,1	23,8	S/I	6,0	17,3	59,2
Barrientos División Andina	26,8	22,3	7,0	10,2	11,9	78,2

Fuente: Rev. méd. Chile v.132 n.9 Santiago sep. 2004

Variables Involucradas	Unidad de análisis (UA)	Unidad de Observación (UO)	Nivel de Medición	Escala de Medición
Diagnóstico informado	¿Sobre qué se está indagando? → UA	¿Quién es el informante? → UO → Dpto. RRHH (también puede ser el Comité Paritario)	Cualitativa nominal	No Aplica
Autor-Empresa (minera)	→ Licencias médicas	→ Listado de licencias informadas	Cualitativa nominal	No Aplica

c)

		Quintil de ingreso autónomo					Total
		I	II	III	IV	V	
Tiene Computador	Si	5,6%	3,2%	12,7%	19,5%	30,9%	21,0%
	No	94,4%	96,8%	87,3%	80,5%	69,1%	79,0%
Total		100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Fuente: Encuesta CASEN, Santiago. 2000

Variables Involucradas	Unidad de análisis (UA)	Unidad de Observación (UO)	Nivel de Medición	Escala de Medición
Quintil de ingreso autónomo	La encuesta CASEN es una encuesta de hogares → UA	¿Quién es el informante? → UO → Jefe/a de Hogar	Cualitativa ordinal	No Aplica
Tiene computador	→ Hogares Pero también tiene módulos que buscan conocer a las personas (de todas las edades) habitantes de Chile → UA → Personas habitantes de Chile	o adulto/a (mayor de 18 años) capacitado/a para contestar	Cualitativa nominal, dicotómica	No Aplica

2. Para las siguientes preguntas identifique

- i) La o las variables que habría que digitar en una base de datos
- ii) Los atributos o rango de valores de cada variable
- iii)Cuál es el nivel de medición de cada variable.

- a) De acuerdo a su experiencia, ¿cómo calificaría el funcionamiento del Transantiago en los siguientes aspectos?

	Muy Malo	Malo	Regular	Bueno	Muy Bueno	
Forma de Pago	1	2	3	4	5	
Calidad del Servicio del Metro	1	2	3	4	5	
Calidad de los Buses	1	2	3	4	5	
Cercanía de los paraderos	1	2	3	4	5	
Cantidad de Transbordos	1	2	3	4	5	
Tiempos de Viaje	1	2	3	4	5	
Frecuencia de buses durante el día	1	2	3	4	5	
Frecuencia de buses durante la noche	1	2	3	4	5	

Fuente: CEP, Encuesta Nacional de Opinión Pública, Junio 2007

Variables a Digitar	Atributos	Nivel Medición
Forma de Pago		
Calidad del Servicio del Metro		
Calidad de los Buses	Muy Malo	Variables Cualitativas Ordinales
Cercanía de los paraderos	Malo	
Cantidad de Transbordos	Regular	
Tiempos de Viaje	Bueno	
Frecuencia de buses durante el día	Muy Bueno	
Frecuencia de buses durante la noche		

- b) A continuación se indican seis (6) cualidades personales que podrían considerarse deseables en un empleado. Ordénelas de acuerdo a su importancia, siendo uno (1) la menos importante y seis (6) la más importante.

Compromiso con la institución	
Iniciativa propia	
Buena apariencia profesional	
Responsabilidad	
Discreción	
Liderazgo	

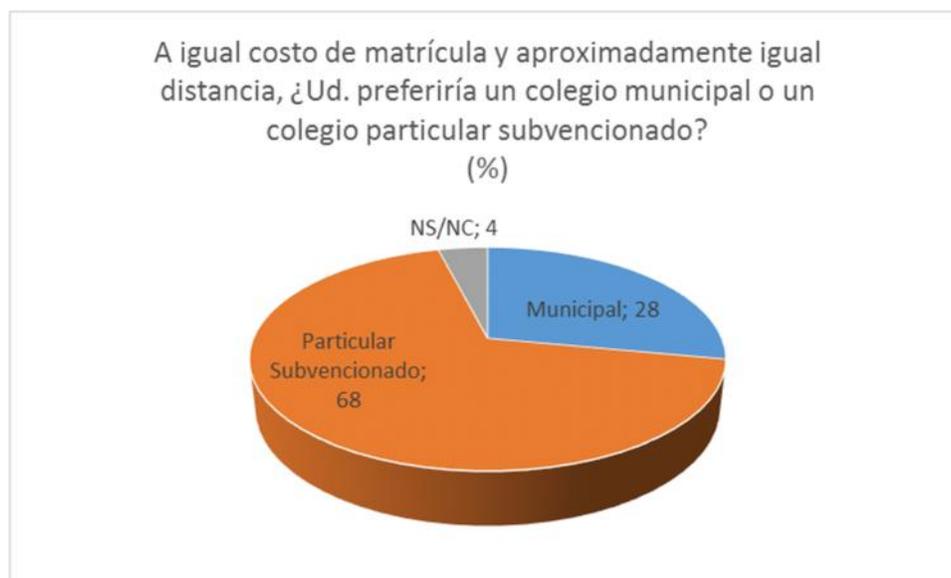
Fuente: Cuestionario de destrezas y ocupaciones en mayor demanda. www.net-empleopr.org

Variables a Digitar	Atributos	Nivel Medición
Compromiso con la institución	1 a 6	Cualitativas Ordinales
Iniciativa propia		
Buena apariencia profesional		
Responsabilidad		
Discreción		
Liderazgo		

3. Las siguientes tablas muestran datos de la Encuesta Nacional de Opinión Pública, realizada por el CEP, en diciembre de 2006. Seleccione una forma para graficar los datos y realice el gráfico. Justifique su elección

- a) A igual costo de matrícula y aproximadamente igual distancia, ¿Ud. preferiría un colegio municipal o un colegio particular subvencionado?

Tipo de Establecimiento	%
Municipal	28
Particular Subvencionado	68
NS/NC	4



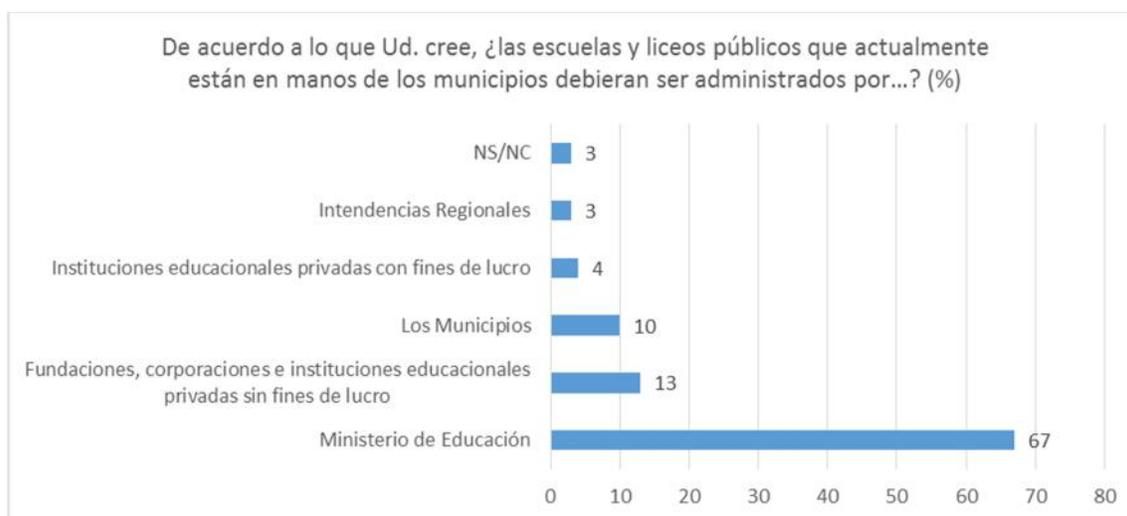
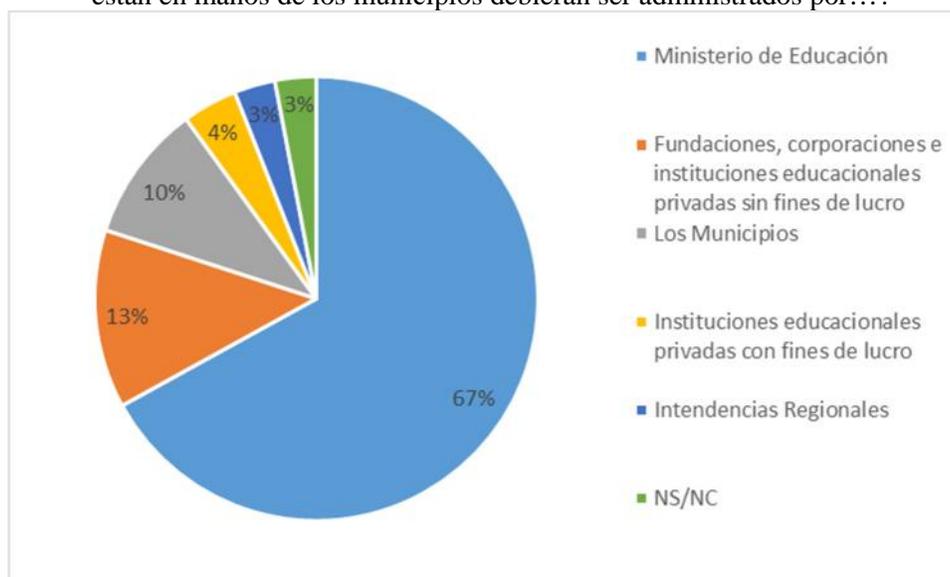
En este caso utilizamos gráfico de torta porque son pocas categorías y las diferencias entre ellas son importantes. Graficamos los No sabe/No Contesta, porque los datos están presentados en porcentajes, y estas categorías están incluidas en el cálculo de ellos.

- b) De acuerdo a lo que Ud. cree, ¿las escuelas y liceos públicos que actualmente están en manos de los municipios debieran ser administrados por...?

Tipo Administración	%
Ministerio de Educación	67
Fundaciones, corporaciones e instituciones educativas privadas sin fines de lucro	13
Los Municipios	10
Instituciones educativas privadas con fines de lucro	4
Intendencias Regionales	3
NS/NC	3

En este caso, podría hacerse tanto un gráfico de barra como de torta, veamos cómo quedaría cada uno:

De acuerdo a lo que Ud. cree, ¿las escuelas y liceos públicos que actualmente están en manos de los municipios debieran ser administrados por...?



Si bien el gráfico de torta permite tener, visualmente, una idea más clara de las diferencias entre las categorías, el gráfico de barra, permite unir, visualmente, cada categoría con el porcentaje de respuesta y debido al número de categorías pareciera que el segundo gráfico (barras) es más apropiado

4. La siguiente tabla, muestra los resultados para la pregunta 10 de la Encuesta Nacional de Opinión Pública UDP 2009.

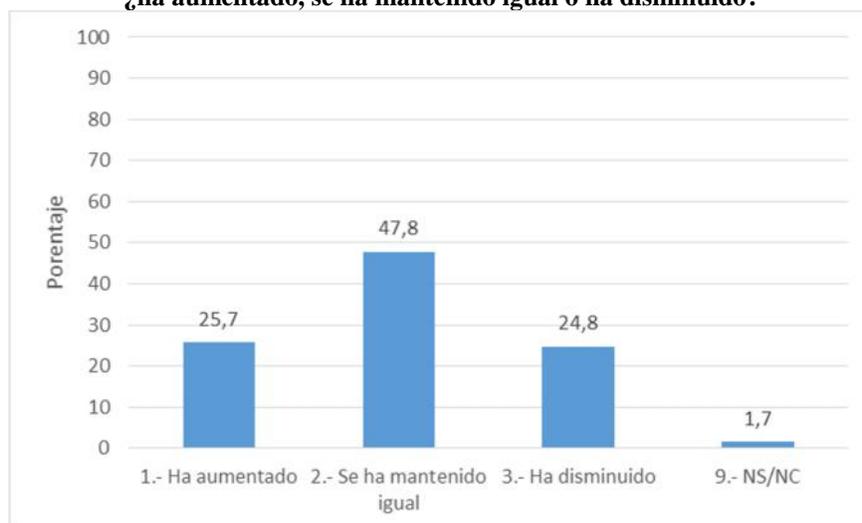
P10.- Considera Ud. que en los últimos 5 años el nivel de pobreza en Chile, ¿ha aumentado, se ha mantenido igual o ha disminuido?

	Frecuencia	Porcentaje
1.- Ha aumentado	335	25,7
2.- Se ha mantenido igual	622	47,8
3.- Ha disminuido	323	24,8
9.- NS/NC	22	1,7
Total	1302	100,0

Fuente: Encuesta Nacional de Opinión Pública UDP, 2009

i) Seleccione una forma para graficar los datos y justifique su elección.

P10.- Considera Ud. que en los últimos 5 años el nivel de pobreza en Chile, ¿ha aumentado, se ha mantenido igual o ha disminuido?



Debido a que la variable es ordinal, es recomendable utilizar un gráfico de barras ya que el gráfico de torta se pierde la idea de orden

ii) Indique la unidad de análisis y de observación de la encuesta UDP 2009.

La encuesta UDP, es una encuesta de opinión URBANA, por lo que la **Unidad de Análisis** corresponde a las personas mayores de 18 años habitantes de zonas urbanas Chile.

La **Unidad de Observación** es similar a la Unidad de Análisis: población de 18 años y más, residentes en 86 comunas de 20.000 habitantes de todas las regiones, excluyendo Aysén y Magallanes

5. Se tomó una muestra de 50 calificaciones de los resultados de un examen de admisión para primero medio en Colegio Particular. Los siguientes son los resultados

27	68	79	91	108
43	71	80	91	108
43	71	81	93	116
44	71	82	94	120
47	73	82	94	120
49	73	84	94	122
50	74	84	96	123
54	75	86	97	127
58	76	88	106	128
65	77	88	107	203

i) Realice gráfico tallo-hoja

2	7
3	
4	33479
5	048
6	58
7	1113345679
8	012244688
9	11344467
10	6788
11	6
12	002378
13	
14	
15	
16	
17	
18	
19	
20	3

ii) ¿De qué otra forma se podría graficar?

R: Otra forma de graficar sería utilizar un **histograma**, ya que “calificaciones”, aunque no es una variable continua, tiene un rango muy amplio, lo que permite utilizar este tipo de representación gráfica (algo similar a lo que pasa con edad o ingreso)

iii) ¿Qué podría decir respecto de las calificaciones a partir de lo que muestra el gráfico?

6. Para el siguiente conjunto de datos realice el gráfico tallo-hoja e interprete

Edad Según Sexo

		sexo		Total
		Mujeres	Hombres	
edad	15	1	1	2
	16	0	2	2
	17	1	0	1
	18	1	2	3
	20	1	2	3
	21	1	3	4
	22	1	1	2
	24	1	1	2
	25	2	1	3
	29	0	1	1
	30	0	2	2
Total		9	16	25

Alternativa 1

Mujeres		Hombre
5	1	5
	1	6 6
7	1	
8	1	8 8
0	2	0 0
1	2	1 1 1
2	2	2
4	2	4
5 5	2	5
	2	9
	3	0 0

Alternativa 2

1	5 5
1	6 6
1	7
1	8 8 8
2	0 0 0
2	1 1 1 1
2	2 2
2	4 4
2	5 5 5
2	9
3	0 0

7. Para las siguientes preguntas fueron extraídas del cuestionario del Latinobarómetro, en cada caso identifique

- i) La o las variables que habría que digitar en una base de datos
- ii) Los atributos o rango de valores de cada variable
- iii) Nivel y la escala de medición de cada variable

P13ST. (MOSTRAR TARJETA 1) Imagínese una escala de 10 peldaños, en que en el "1" se ubican las personas más pobres y en el "10" se ubican las personas "más ricas", ¿dónde se ubicaría Ud.? (MUESTRE ESCALA, ESPERE RESPUESTA Y ANOTE UNA EN P13ST.A) ¿Dónde se ubicarían sus padres? (MUESTRE ESCALA, ESPERE RESPUESTA Y ANOTE UNA EN P13ST.B) ¿Y dónde cree Ud. que se encontrarán sus hijos? (MUESTRE ESCALA, ESPERE RESPUESTA Y ANOTE UNA EN P13ST.C)

		1-10	NS	NR
P13ST.A	Personal		98	00
P13ST.B	Padres		98	00
P13ST.C	Hijos		98	00

P14ST/N (MOSTRAR TARJETA 2) ¿Está Ud. muy de acuerdo (1), de acuerdo (2), en desacuerdo (3) o muy en desacuerdo (4) con las siguientes afirmaciones? (LEA CADA AFIRMACIÓN Y MARQUE UNA RESPUESTA PARA CADA UNA)

MA DA ED MD NS/M

P14ST.A La democracia puede tener problemas, pero es el mejor sistema de gobierno 1 2 3 4 0

P14ST.B Cuando hay una situación difícil en (país), esta bien que el gobierno pase por encima de las leyes, el parlamento y/o las instituciones con el objeto de resolver los problemas 1 2 3 4 0

P14ST.C Los medios de comunicación deben poder publicar noticias sin temer a ser cerrados 1 2 3 4 0

P14ST.D Las decisiones del gobierno buscan privilegiar a algunos pocos 1 2 3 4 0

P14ST.E Que el presidente controle los medios de comunicación 1 2 3 4 0

P14NCC.F Algunas personas y/o grupos tienen tanta influencia que los intereses de la mayoría son ignorados . 1 2 3 4 0

P14NG El matrimonio entre homosexuales 1 2 3 4 0

VARIABLES	ATRIBUTOS	NIVEL DE MEDICIÓN
P13ST.A Personal	1 Personas más pobres A	Cualitativa Ordinal
P13ST.B Padres	10 Personas más ricas	
P13ST.B Hijos	98 No sabe – 00 No responde	
P14ST.A	1. Muy de acuerdo	Cualitativa Ordinal
5P14ST.B	2. De acuerdo	
P14ST.C	3. En desacuerdo	
P14ST.D	4. Muy en desacuerdo	
P14ST.E	0. NS/NR	
P14NCC.F		
P14NG		

8. Para las siguientes preguntas fueron extraídas del cuestionario de la Encuesta CASEN 2011, en cada caso identifique

- i) La o las variables que habría que digitar en una base de datos
- ii) Los atributos o rango de valores de cada variable
- iii) Nivel y la escala de medición de cada variable
- iv) Indique la unidad de análisis y de observación de la encuesta CASEN

Todas las personas	Personas que asistieron o asistieron a Educación Superior (e6.a=11, 12 ó 13)
<p>e6.a. ¿Cuál es el nivel más alto alcanzado o el nivel educacional actual? MOSTRAR TARJETA E6 "Nivel Educacional" Lea alternativas</p> <ol style="list-style-type: none"> 1. Nunca asistió 2. Jardín Infantil / Sala Cuna 3. Kinder / Prekinder 4. Educación Especial (Diferencial) 5. Primaria o Preparatoria (Sistema antiguo) 6. Educación Básica 7. Humanidades (Sistema Antiguo) 8. Educación Media Científico-Humanista 9. Técnica Comercial, Industrial o Normalista (Sistema Antiguo) 10. Educación Media Técnica Profesional 11. Técnico Nivel Superior 12. Profesional 13. Postgrado <p>e6.b. ¿Completó el nivel educacional anteriormente declarado?</p> <ol style="list-style-type: none"> 1. Si 2. No 7. No aplica (nivel 1 a 4 en e6.a) 	<p>e7. ¿Cuántos años dura la carrera que estudió o que estudia actualmente?</p> <p>→ Si no sabe anote 99</p> <p> Pregunte por la duración de la carrera de pre-grado a quienes reportan estudios de postgrado (e6.a=13)</p>

VARIABLES	ATRIBUTOS	NIVEL DE MEDICIÓN	ESCALA DE MEDICIÓN
Nivel Educacional	1. Nunca Asistió 2. Jardín Infantil / Sala Cuna 3. Kinder / Prekinder 4. Educación Especial (Diferencial) 5. Primaria o Preparatoria (Sistema Antiguo) 6. Educación Básica 7. Humanidades (Sistema Antiguo) 8. Técnica	Cualitativa Ordinal	
¿Completó el nivel educacional?	1. Si - 2. No - 7. N/A	Cualitativa Nominal, Dicotómica	
Duración de la carrera en años	1 a 7 años 99 NS	Cuantitativa de intervalo	Continua (ejemplo: 5 semestres → 1,5 años)

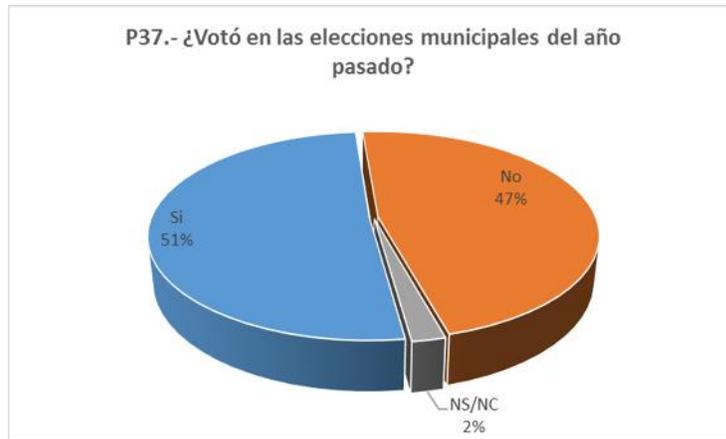
9. La siguiente tabla, muestra los resultados para la pregunta 37 de la Encuesta Nacional de Opinión Pública UDP septiembre 2013.

P37.- ¿Votó en las elecciones municipales del año pasado?

	Frecuencia	Porcentaje
Si	663	51,0
No	611	47,0
NS	6	,4
NC	21	1,6
Total	1300	100,0

Fuente: Encuesta Nacional de Opinión Pública UDP sept. 2013

i) Selecciones una forma para graficar los datos y justifique su elección



En este caso utilizamos gráfico de torta porque son pocas categorías y este tipo de gráfico permite visualizar de forma simple y clara la distribución de la variable. Además, debido a que los No Sabe y No Contesta tienen frecuencias muy bajas se juntaron en una sola categoría de manera de representarlos más claramente.

2. La siguiente tabla muestra la edad del primer consumo de alcohol para un grupo de 180 jóvenes entre 18 y 27 años: ¿A qué edad consumió alcohol por primera vez? (VER GUÍA II)

Edad (x_i)	Número Casos (f_i)			
11	9			
12	24			
13	27			
14	30			
15	27			
16	32			
17	10			
18	11			
19	5			
20	1			
22	2			
23	1			
27	1			
	180			

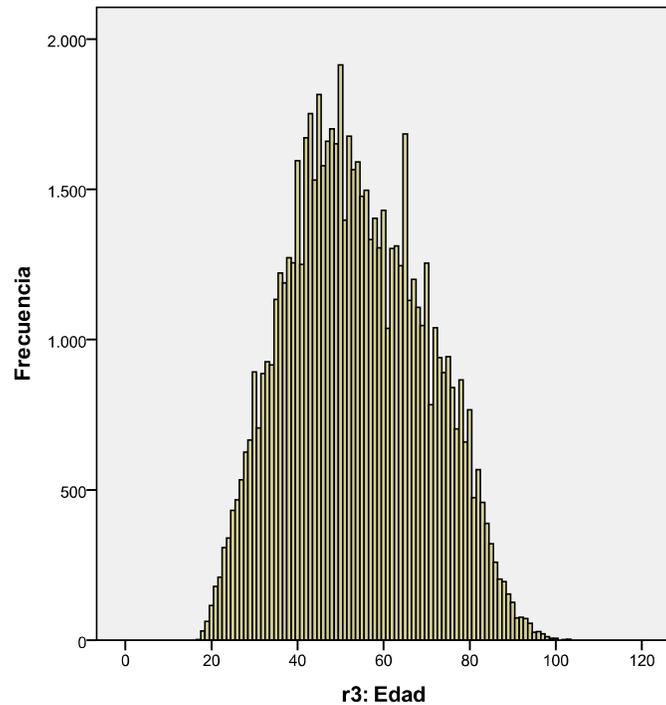
i) Complete la tabla de frecuencia: calcule frecuencia relativa, absoluta acumulada y relativa acumulada

- ii) Calcule medidas de tendencia central
- iii) Indique qué representación gráfica utilizaría para esta variable, justifique su elección
- iv) A partir de los datos presentados ¿Qué puede decir sobre la edad del primer consumo de alcohol de los jóvenes?

3. Los datos que se presentan a continuación, corresponde a la edad de los/as jefes/as de hogar, de acuerdo a los resultados de la encuesta CASEN 2009. *(VER GUÍA II)*

- i) A partir de los datos presentados ¿Qué puede decir respecto de la relación entre las medidas de tendencia central? ¿Está sesgada la distribución?
- ii) ¿Qué puede decir sobre la edad del de los/as jefes/as de hogar?

Edad Jefe/a de Hogar		
N	Válidos	71460
	Perdidos	0
Media		53,73
Mediana		53,00
Moda		50



4. Para la siguiente tabla de frecuencia:

x	f				
1	4				
2	4				
3	8				
4	7				
5	5				
6	10				
7	7				
8	5				
Total	50				

i) Complete la tabla con fr; F; Fr

x	f	fr	F	Fr
1	4	8%	4	8%
2	4	8%	8	16%
3	8	16%	16	32%
4	7	14%	23	46%
5	5	10%	28	56%
6	10	20%	38	76%
7	7	14%	45	90%
8	5	10%	50	100%
Total	50	100%		

ii) Calcule Medidas de tendencia central

x	f	X*f
1	4	4
2	4	8
3	8	24
4	7	28
5	5	25
6	10	60
7	7	49
8	5	40
Total	50	238

Media:

$$\bar{x} = \frac{(1*4) + (2*4) + (3*8) + (4*7) + (5*5) + (6*10) + (7*7) + (8*5)}{50}$$

$$\bar{x} = \frac{4 + 8 + 24 + 28 + 25 + 60 + 49 + 40}{50}$$

$$\bar{x} = \frac{238}{50} = 4,76$$

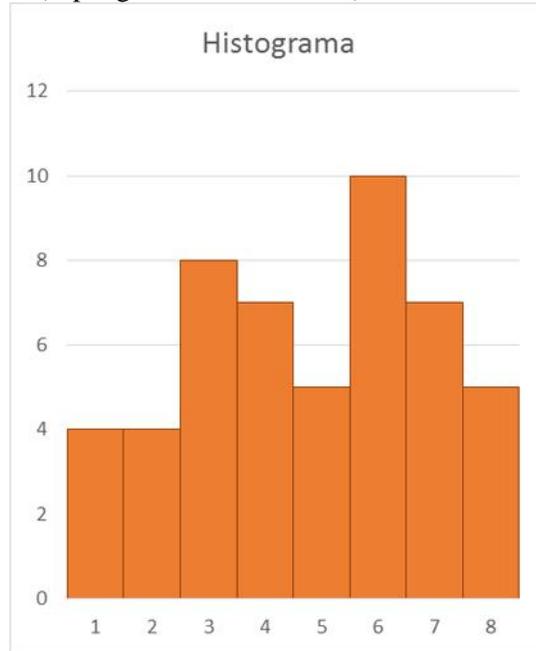
Mediana:

$$d(\tilde{x}) = \frac{50+1}{2} = 25,5$$

$$\text{Mediana } \tilde{x} = 5$$

Moda M = 6

iii) Dibuje el histograma (suponga variable continua)



5. En la fabricación de un cierto tipo de clavos, aparecen un cierto nº de ellos defectuosos. Se han estudiado 200 lotes de 500 clavos cada uno obteniendo:

Clavos defectuosos	nº de lotes	fr	F	Fr	x*fi
1	5	2,5	5	2,5%	5
2	15	7,5	20	10%	30
3	38	19,0	58	29%	114
4	42	21,0	100	50%	168
5	49	24,5	149	74,5%	245
6	32	16,0	181	90,5%	192
7	17	8,5	198	99%	119
8	2	1,0	200	100%	16
	200	100,0			889

- i) Complete la tabla de frecuencia
- ii) Calcule medidas de tendencia central
- iii) Calcule medidas de dispersión
- iv) Grafique

ii) Calcule medidas de tendencia central

Media: $\bar{x} = \frac{889}{200} = 4,45$

Mediana

$$d(\tilde{x}) = \frac{200+1}{2} = 100,5$$

Mediana $\tilde{x} = 4,5$

Moda: M = 5

iii) Calcule medidas de dispersión

$$\text{Rango} = 8 - 1 = 7$$

Varianza y Desviación Estándar.

Calculamos para muestra ya que sólo se estudian 200 lotes, que no es el total de clavos producidos.

Fórmula 1

Clavos defectuosos	nº de lotes	x_i -media	$(x_i$ -media) ²	$[(x_i$ -media) ²]* f_i
1	5	-3,45	11,87	59,34
2	15	-2,45	5,98	89,67
3	38	-1,45	2,09	79,34
4	42	-0,45	0,20	8,32
5	49	0,56	0,31	15,09
6	32	1,56	2,42	77,38
7	17	2,56	6,53	110,98
8	2	3,56	12,64	25,28
	200			465,40

$$\text{Varianza} = s^2 = \frac{465,4}{199} = 2,34$$

$$\text{Desviación Estándar} = \sqrt{2,34} = 1,53$$

Fórmula 2

Clavos defectuosos	nº de lotes	$f_i \cdot x_i$	x^2	$x^2 \cdot f_i$
1	5	5	1	5
2	15	30	4	60
3	38	114	9	342
4	42	168	16	672
5	49	245	25	1225
6	32	192	36	1152
7	17	119	49	833
8	2	16	64	128
	200	889		4417

$$\text{Varianza} = s^2 = \frac{4417 - \left(\frac{889^2}{200}\right)}{199} = 2,34$$

$$\text{Desviación Estándar} = \sqrt{2,34} = 1,53$$