



**UNIVERSIDAD DE BUENOS AIRES**

**ESTADÍSTICA - CÁTEDRA I**

**CÓDIGO 60**

**PROF. ADJUNTA A CARGO: DRA. MARÍA SILVIA GALIBERT**

**TRABAJO PRÁCTICO Nro. 1**

**AUTORES**

**Profesores y Docentes de la Cátedra I**





- a) A partir de las respuestas de quienes integran su comisión obtenga la matriz de datos manualmente o con un programa. Indique cuáles son las unidades de análisis. ¿Qué se entiende por matriz de datos y por unidades de análisis?
- b) Mencione y clasifique cada variable e indique su nivel de medición. Explique los conceptos de variable, categoría y de qué depende el nivel de medición. Ejemplifique.
- c) Obtenga manualmente las tablas de distribución de frecuencias absolutas, relativas y porcentuales de la edad y del sexo. Grafique ambas distribuciones. Defina frecuencia absoluta, relativa y porcentual.
- d) Obtenga manualmente las tablas de distribución de frecuencias absolutas y absolutas acumuladas del nivel socioeconómico. Grafique la primera.
- e) Realice manualmente una tabla con la distribución de frecuencias porcentuales de las variables P08 y P09. Grafíquelas. ¿Encuentra alguna tendencia general en el comportamiento u opinión sobre la puntualidad? Si fuera el caso ¿cuál?
- f) Obtenga manualmente la tabla de frecuencias para P11 agrupando el tiempo en intervalos de clase y gráfiquela (tome la cantidad de intervalos conveniente según los tiempos observados en su comisión).
- ¿En torno a qué valores estuvieron los tiempos que la comisión necesitó para resolver el cuestionario?
  - ¿Qué porcentaje de personas necesitaron a lo sumo 12,5 minutos para responder el cuestionario?
  - ¿Qué porcentaje de personas necesitaron por lo menos 14,5 minutos para responder el cuestionario?
- g) Utilice la base de datos de este problema y un programa informático (InfoStat, Statistix, EXCEL, GeoGebra, etc.) para hallar las distribuciones de frecuencia de las variables P1, P2, P3, P7, P8, P9 y P10. Grafíquelas. ¿Encuentra alguna tendencia general en el comportamiento y opinión sobre la puntualidad? Si fuera el caso ¿cuál?
- h) Halle una distribución de frecuencias para datos agrupados de la variable P11 y el correspondiente histograma.
- ¿En torno a qué valores estuvieron los tiempos que la comisión necesitó para resolver el cuestionario?
  - ¿Qué porcentaje de personas necesitaron a lo sumo 14 minutos para responder el cuestionario?
  - ¿Qué porcentaje de personas necesitaron por lo menos 17 minutos para responder el cuestionario?

2.- La siguiente tabla presenta las estadísticas de cierto centro terapéutico con respecto al tipo de trastorno diagnosticado en los pacientes que consultaron durante el último año.

DIAGNÓSTICO	PORCENTAJE
Trastorno alimentario	13
Trastorno de ansiedad	31
Distrés	24
Trastorno obsesivo-compulsivo	18
Trastorno de Sueño	9
Otros	5

- Mencione la variable que se está considerando y clasifíquela.
- Represente la información en un gráfico adecuado.
- ¿Por qué podría ser de interés disponer de la información que da la tabla?
- Halle las frecuencias absolutas de cada categoría suponiendo que al centro asistieron 800 personas el último año.
- Durante el año anterior, de un total de 990 personas que fueron a la consulta, 280 tuvieron trastornos de ansiedad. ¿Podría decirse que, con relación al año anterior, en ese centro se observa una tendencia decreciente de consultas por problemas de ansiedad? Justifique la respuesta.
- Para conceptualizar:** Compare la información que dan las frecuencias absolutas, las relativas y las porcentuales. ¿En qué casos es indistinto, preferible o indispensable usar una u otra?

3.- Cierta centro de estudios universitarios tiene que planificar las vacantes para cada carrera, lo que implica, entre otras cosas, determinar el equipamiento de cada aula en cuanto a butacas y eventualmente computadoras. Le informan que para el año entrante el 30% de inscriptos quiere ingresar a la carrera de Psicopedagogía, el 20% a Musicoterapia y el resto a Psicología. ¿Es esta información suficiente para la planificación? Justifique.

4.- Dos docentes de Estadística comentan los resultados obtenidos por sus estudiantes en el primer parcial que tomaron. Carla dice, gratamente sorprendida, que tuvo 3 alumnos con 10 y Tomás le contesta jocosamente: “Yo tuve 4 alumnos con 10, te gané”. Discuta en qué casos el “te gané” de Tomás puede sustentarse.

5.- Un profesor de castellano de primer año de escuela media toma una exhaustiva prueba diagnóstica a sus 40 alumnos. Los puntajes están en una escala de 0 a 100 y su distribución se muestra en la figura 1. Después de analizarlos, decide dedicar el primer mes de clase a un repaso general a fin de emparejar y mejorar el desempeño del grupo antes de comenzar a enseñar su programa específico. Después de dicho repaso vuelve a tomar una prueba diagnóstica similar a la anterior, cuya distribución de puntajes se muestra en la figura 2. a) Según los histogramas, ¿se obtuvieron resultados favorables a los objetivos del profesor? Fundamente su respuesta comparando dichos histogramas.

**Conceptualice:**

i) ¿Para qué tipo de variable es recomendable el uso de histogramas?

ii) Relacione la forma de las distribuciones representadas por los histogramas con la concentración o dispersión de los datos (homogeneidad o heterogeneidad de los grupos).

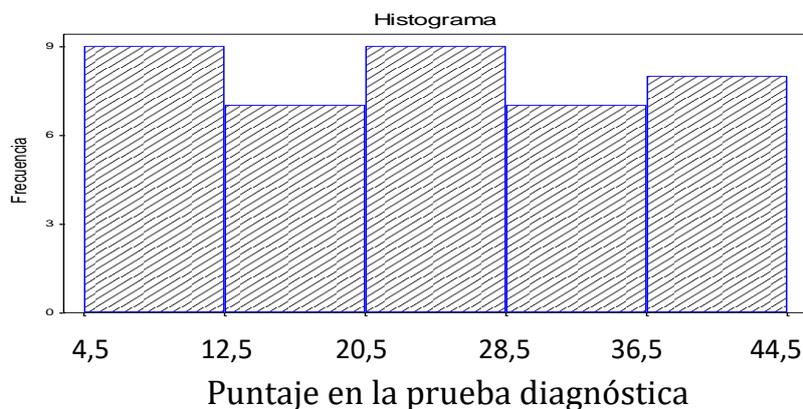


Figura 1. Distribución de los puntajes antes del repaso

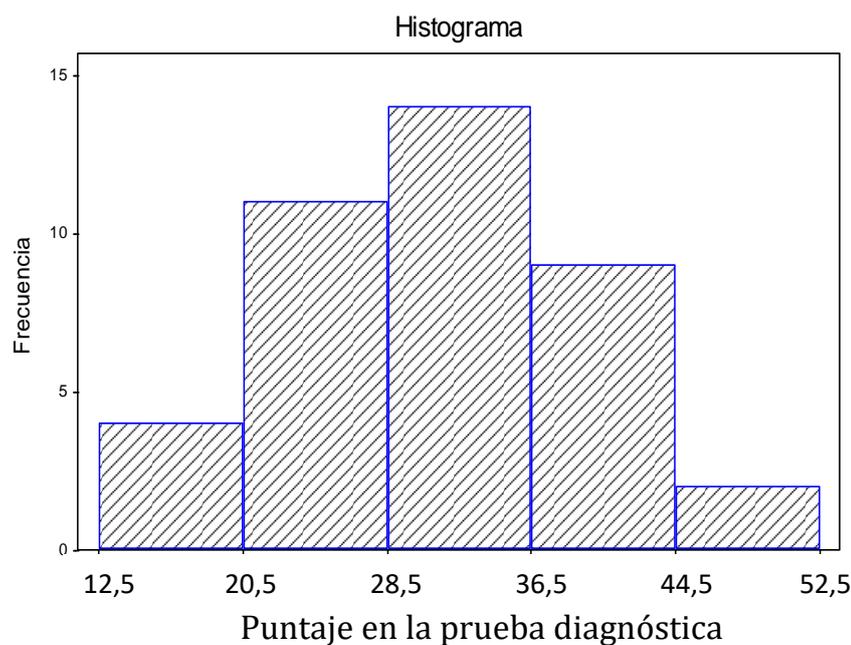


Figura 2. Distribución de los puntajes después del repaso

6.- El profesor aludido en el ejercicio 5 debe presentar los resultados en un informe donde no puede poner gráficos. Vuelque la información de los mismos en sendas tablas de frecuencias.

7.- En cierta investigación (datos ficticios) se registró, en décimas de segundo, el tiempo de reacción a un estímulo de 20 adolescentes entre 15 y 17 años y de 20 adultos entre 57 y 60 años. Los resultados se muestran respetivamente en los diagramas de tallo-hoja 1 y 2.

#### Tiempo de reacción de los adultos

Unidad de las hojas = 0,1

5 7 representa 5,7

##### Tallo Hojas

```

5 7
6 48
7 224677
8 13556899
9 237

```

#### Tiempo de reacción de los adolescentes

Unidad de las hojas = 0,1

5 7 representa 5,7

##### Tallo Hojas

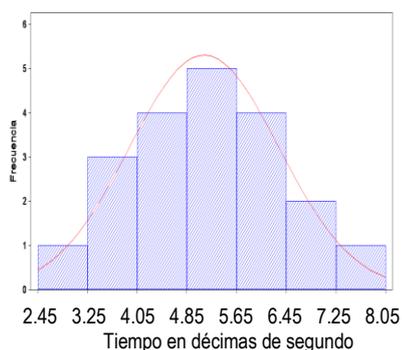
```

5 7
6 12556899
7 13778
8 0226
9 07

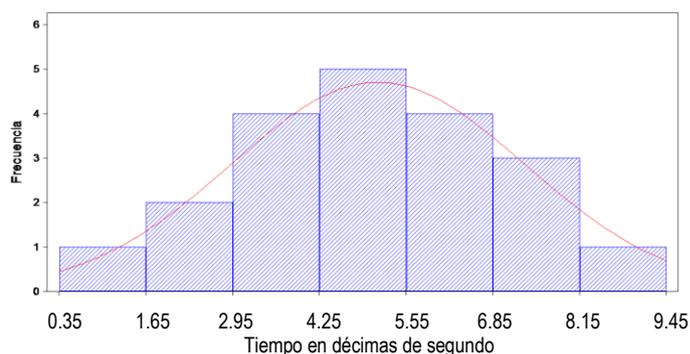
```

- I) ¿Qué puede concluir, en general, de la comparación de ambos grupos?
- II) Las 40 personas que participaron de la experiencia fueron luego entrenadas para mejorar su tiempo de reacción. Después del entrenamiento se midieron nuevamente los tiempos de reacción. Como los datos los analizó otro estadístico, utilizó histogramas en lugar de diagramas de tallo-hoja para presentar los resultados de esta etapa, los cuales se muestran a continuación.

### Tiempo de reacción de los adultos



### Tiempo de reacción de los adolescentes



- ¿Parece haber tenido el entrenamiento algún efecto sobre cada grupo? Si es así, ¿cuál?
- Compare las distribuciones de los tiempos entre ambos grupos después del entrenamiento.
- Conceptualice** las representaciones por histogramas y por diagramas de tallo-hoja. ¿Qué información similar proporcionan? ¿Qué ventaja tiene el diagrama de tallo-hoja con respecto al histograma?
- Obtenga manualmente, en un mismo par de ejes cartesianos, la ojiva de Galton para cada histograma y compare la forma de ambas ojivas. **Conceptualice:** ¿Qué información da el eje de ordenadas?

8.- Proponga una variable continua que pueda medir en su entorno. Tome 30 mediciones de la misma y represente la distribución observada con un diagrama de tallo-hoja y con un histograma.

*Nota. Como alternativa, en lugar de efectuar Ud. mismo las mediciones, puede buscar datos en alguna base como, por ejemplo, del sitio <https://data.world/datasets/psychology> o simularlos con una aplicación de generación de números aleatorios como <https://pinetools.com/es/generador-gaussiano-numeros-aleatorios>*

9.- Los datos del siguiente ejercicio son ficticios pero basados sobre las estadísticas del artículo de Fava, GA, Park SK, Sonino, N (2006). Treatment of recurrent depression. *Expert Review of Neurotherapeutics* 6 (11), 1735-1740.

En una clínica 20 pacientes recibieron terapia por haber padecido un episodio depresivo mayor y se les realizó un seguimiento de por vida a partir del alta. A continuación se indica la cantidad de recaídas de dichos pacientes en este tipo de episodios.

4 – 6 – 0 – 1 – 5 – 2 – 0 – 3 – 5 – 0 – 3 – 4 – 0 – 4 – 3 – 5 – 4 – 3 – 2 – 4

- Mencione la variable en cuestión y clasifíquela.
- Organice los datos en una tabla de frecuencias y represente la distribución con un gráfico adecuado.
- Obtenga en la misma tabla las frecuencias porcentuales acumuladas.
- ¿Qué porcentaje de pacientes tuvo: i) al menos una recaída, ii) a lo sumo una recaída?
- De los pacientes que tuvieron recaídas, ¿qué porcentaje sufrió más de una?
- Conceptualice:**
  - ¿Por qué no es adecuado en este caso un histograma?
  - ¿Podría haber utilizado un diagrama de tallo-hoja para organizar estos datos?
  - ¿Qué similitudes y diferencias hay entre un diagrama de tallo-hoja y un diagrama de bastones?

### 10.- Conceptualice:

- ¿Qué diferencia esencialmente a la distribución de frecuencias de una variable discreta de la de una continua?
- ¿Qué vinculación hay entre esta diferencia y sus respectivas formas de representación gráfica?

**11.-** Imagine que va a tomar datos para cierta investigación en un hospital.

- i) Mencione una variable cualitativa de nivel nominal y otra de nivel ordinal que pudieran ser de interés.
- ii) Mencione también las categorías de cada variable e invente una tabla de frecuencias para cada una de ellas; luego representelas con gráficos adecuados.

**Conceptualice:** En cada tabla de frecuencias Ud. puede hallar las frecuencias acumuladas pero esto no tiene sentido en el caso de la variable de nivel nominal y sí en el de la variable de nivel ordinal ¿Por qué?

**12.-** Se registró el porcentaje de deserción escolar de los estudiantes de primer año en cada escuela de enseñanza media de Gran Buenos Aires correspondiente al año pasado. Indique cuáles son las unidades de análisis.

**En los siguientes ítems elija la única opción correcta**

**13.-** En un diagrama circular, el ángulo de cada sector es directamente proporcional a la

- a. frecuencia del correspondiente valor de la variable.
- b. cantidad de unidades de análisis observadas.
- c. cantidad de categorías o valores de la variable.

**14.-** Como parte de un estudio de Planificación Comunitario, 50 personas fueron encuestadas sobre su Nivel de Escolaridad. Se encontró que la frecuencia absoluta de los que tenían estudios Secundarios Completos (SC) triplicaba a la frecuencia absoluta de los que tenían estudios Terciarios Completos (TC) (por ejemplo, SC = 15 y TC = 5). Entonces esta misma relación de 3 a 1 entre esos valores se mantendría también para sus frecuencias

- a. absolutas acumuladas.
- b. relativas.
- c. acumuladas porcentuales.

**15.-** Dada la distribución de frecuencias relativas de una variable, es posible conocer la distribución de frecuencias absolutas

- a. si se conoce el total de observaciones.
- b. se conozca o no el total de observaciones.
- c. si se conoce la distribución de frecuencias porcentuales.

**16.-** En la representación gráfica de las distribuciones de frecuencia, un histograma y un diagrama de barras anchas se utilizan

- a. independientemente del tipo de variable en cuestión.
- b. para variables continuas y cualitativas respectivamente.
- c. ambos como alternativa al diagrama circular.

**17.-** Cuando se desea comparar dos distribuciones provenientes de una cantidad distinta de datos la comparación debe basarse sobre las frecuencias

- a. absolutas o relativas.
- b. absolutas o porcentuales.
- c. relativas o porcentuales.

## NOTA: TEORÍA Y PRÁCTICA ACERCA DE LA CLASIFICACIÓN DE VARIABLES

### Niveles de medición

Los puntajes de muchos test de ejecución típica (como el del ejemplo del ejercicio 1, práctica 1) se obtienen a partir de escalas de tipo Likert que son esencialmente ordinales. Sin embargo, se les asigna un número a cada categoría de la escala y estos números se suman, lo cual sólo sería admisible si la escala para la evaluación de cada ítem fuera intervalar (unidad de medida objetiva y distancias bien definidas). Esta práctica supone una aproximación de las escalas ordinales por las intervalares, suponiendo una cierta equidistancia entre las categorías codificadas con números, lo cual es aceptado en el ambiente psicométrico, siempre que se tomen al menos 4 categorías y se las defina “más o menos” equidistantes. Los puntajes obtenidos resultan ser enteros; es decir, se genera una variable “discreta” pero en muchos casos representan un continuo teórico. Por tanto, en la mayoría de los casos, los puntajes que provienen de los tests son tratados como variables de nivel intervalar y continuas.

Los modelos son construcciones teóricas; cuando se habla de niveles de medición y variables discretas o continuas, se está aludiendo a modelos matemáticos. Al aplicarlos para abordar problemas prácticos, difícilmente puedan satisfacerse totalmente los requerimientos de dichos modelos; siempre habrá discrepancias entre la teoría y la práctica. Si esa discrepancia es tolerable, el modelo es útil; de lo contrario es necesario cambiar el modelo para ajustarlo a la realidad y describir el problema adecuadamente.

### Discretización de lo continuo por la medición

Teóricamente, la frecuencia de cada valor individual de una variable continua es cero; es decir, no son los valores individuales sino los intervalos los que captan frecuencia. En un conjunto de datos cada valor, obviamente, tiene frecuencia porque todo conjunto de datos es finito, por tanto todos los datos observados son “discretos”. Sin embargo, dichos datos pueden estar representando un continuo. El ejemplo más claro es el del tiempo y su medición. El tiempo es una variable continua pero al medirlo se discretiza. Si se observa un tiempo de reacción de 7,5 décimas de segundo, éste podría estar representando cualquier tiempo en torno a 7,5 que no se registró por la limitación de la precisión de la medida; el 7,5 podría estar representando cualquier valor entre 7,45 y 7,55. Lo mismo podría decirse de cualquier puntaje obtenido para medir algún rasgo; por ejemplo el puntaje proveniente de la “cantidad de palabras recordadas de una lista” como representante de cierto nivel de memoria de un sujeto. Estas variables pueden ser tratadas estadísticamente como discretas tal cual fueron observadas o puede querer rescatarse el carácter continuo del rasgo que representan.

### Referencias

Galibert, M., Hoyos Páez, C. y Alvarez Ponte, L. (2019). Construcción de un cuestionario para detectar la impuntualidad crónica. *XI Congreso Internacional de Investigación y Práctica Profesional en Psicología, Fac. Psicología, UBA*. Libro de resúmenes.

## APÉNDICE DE LA PRÁCTICA 1

Matriz de datos para una muestra de 30 individuos que respondieron el Cuestionario de Impuntualidad

CASO	P01	P02	P03	P04	P05	P06	P07	P08	P09	P10	P11
1	59	1	3	5	2	7	3	1	4	1	11.4
2	40	1	3	3	1	1	4	1	4	1	14
3	60	1	3	6	1	1	4	1	4	1	12.5
4	40	1	3	3	2	1	3	1	4	1	10.9
5	28	1	3	3	1	1	1	4	1	2	17.1
6	60	1	4	3	1	1	2	2	3	1	17.9
7	31	1	2	3	1	1	4	2	4	3	12.4
8	48	1	2	6	2	1	1	2	2	2	17.5
9	42	1	2	5	2	1	2	2	4	1	12.8
10	26	1	2	6	2	1	2	3	2	1	14.4
11	53	1	3	3	1	1	2	2	3	2	9
12	33	1	2	3	2	1	4	1	4	4	11.2
13	53	1	3	3	2	1	3	2	3	1	14.3
14	52	1	2	3	2	1	2	2	2	1	13.1
15	24	1	1	6	2	1	1	1	4	1	15.9
16	42	1	1	3	2	1	4	1	4	1	13.2
17	68	1	3	6	2	1	4	1	4	1	13.8
18	23	2	1	3	1	1	2	4	2	1	11.1
19	38	1	3	3	2	1	4	1	4	3	14.7
20	48	1	2	3	2	1	1	3	2	1	14.8
21	59	2	3	5	1	1	4	1	4	1	10.1
22	44	1	3	3	2	1	3	1	3	2	14.4
23	20	2	4	6	1	1	3	2	4	3	19.2
24	23	2	3	3	1	1	4	1	4	1	20.2
25	43	2	3	3	2	1	4	1	3	2	10
26	52	2	2	5	2	1	3	1	4	1	14.6
27	34	2	3	3	1	1	3	1	4	1	11.8
28	68	2	3	3	2	1	4	1	3	2	13.4
29	20	2	3	6	1	1	4	1	4	1	14.9
30	21	2	2	6	2	1	4	1	4	1	13.4

## RESOLUCIÓN

1.-

a) A partir de las respuestas de quienes integran su comisión obtenga la matriz de datos manualmente o con un programa. Indique cuáles son las unidades de análisis. ¿Qué se entiende por matriz de datos y por unidades de análisis?

La estructura de la matriz consiste en 12 columnas y tantas filas como personas respondieron el cuestionario. La primera columna de la matriz corresponde al número de cuestionario y las siguientes 11 a cada una de las preguntas formuladas (variables: P1, P2,...,P11). En cada fila de la matriz se halla la respuesta codificada de cada alumno de la comisión a cada una de las preguntas. La matriz es como la que figura en el Apéndice de la Práctica 1.

Las unidades de análisis son cada uno de los estudiantes de la comisión que respondieron al cuestionario.

b) Mencione y clasifique cada variable e indique su nivel de medición.

Variable	Tipo	Nivel de Medición
P1	Cuantitativa Discreta*	de Razón*
P2, P4, P5 y P6	Cualitativa	Nominal
P3, P7, P8, P9, P10	Cualitativa	Ordinal
P11	Cuantitativa	Continua

\* Por definición la Edad es la cantidad de años enteros desde el momento del nacimiento. De esta definición resulta que la variable es cuantitativa discreta y de nivel de razón: el 0 significa que no pasaron años desde el nacimiento. Pero si esta variable se piensa como una operacionalización del “tiempo que la persona lleva de existencia biológica”, entonces esa variable es cuantitativa continua (discretizada en su operacionalización por la edad) y el cero en el momento del nacimiento es arbitrario (ya que su vida biológica no comenzó en ese momento); por lo que sería de nivel intervalar (el cero absoluto corresponde al inicio mismo de su vida biológica). Ocurre como con la temperatura, donde el cero absoluto corresponde a los Kelvin (equivalentes a los  $-273,15^{\circ}$  Celsius), pero que en las otras escalas como de grados Celsius o Fahrenheit el cero es arbitrario.

b) Explique los conceptos de variable, categoría y de qué depende el nivel de medición. Ejemplifique. Ver Cap.1 de Bologna, Apunte de Explicaciones Complementarias de García Díaz y Clase Teórica 1.

c) Obtenga manualmente las tablas de distribución de frecuencias absolutas, relativas y porcentuales de la edad y del sexo. Grafique ambas distribuciones. Defina frecuencia absoluta, relativa y porcentual.

Lo resolveremos como si la matriz de datos fuera la del apéndice, aunque no corresponde a estudiantes, sino a población general.

Tabla de frecuencias para la variable Edad

Edad $x_i$	$f_i$	$f_i'$	$f_i\%$
20	2	0,067	6,7
21	1	0,033	3,3
23	2	0,067	6,7
24	1	0,033	3,3
26	1	0,033	3,3
28	1	0,033	3,3
31	1	0,033	3,3
33	1	0,033	3,3
34	1	0,033	3,3
38	1	0,033	3,3
40	2	0,067	6,7
42	2	0,067	6,7
43	1	0,033	3,3
44	1	0,033	3,3
48	2	0,067	6,7
52	2	0,067	6,7
53	2	0,067	6,7
59	2	0,067	6,7
60	2	0,067	6,7
68	2	0,067	6,7
Total	30	1	100,0

Tabla de frecuencias para la variable Sexo

Sexo	$f_i$	$f_i'$	$f_i\%$
Mujer	20	2/3	66,7
Varón	10	1/3	33,3
Total	30	1	100,0

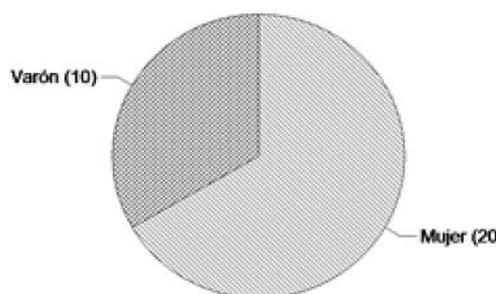
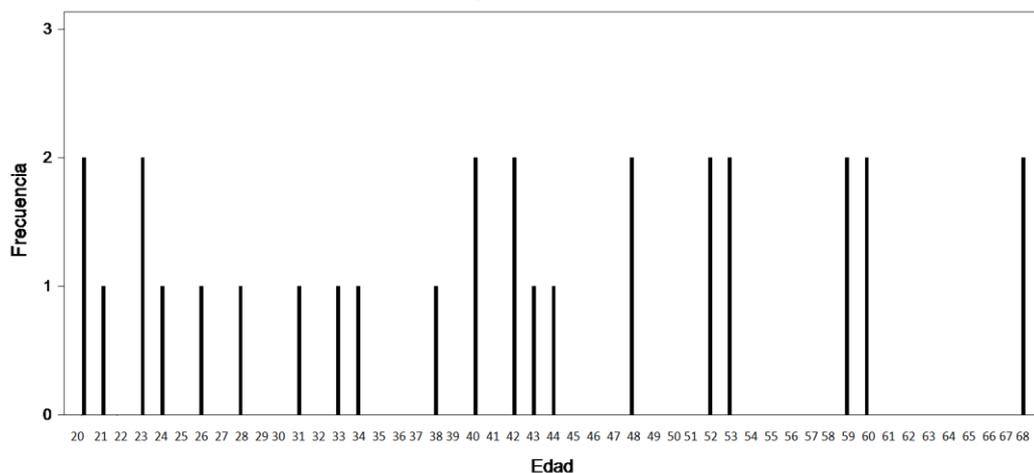
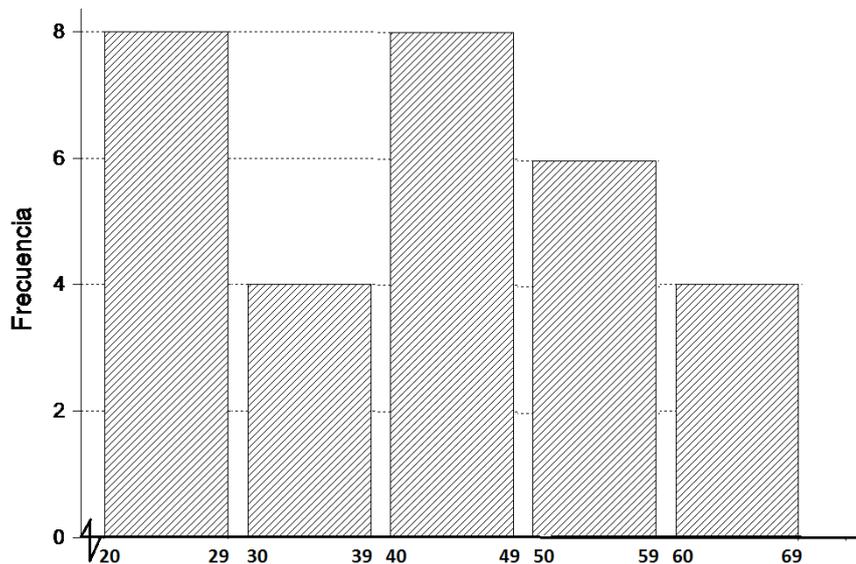


Diagrama de Bastones



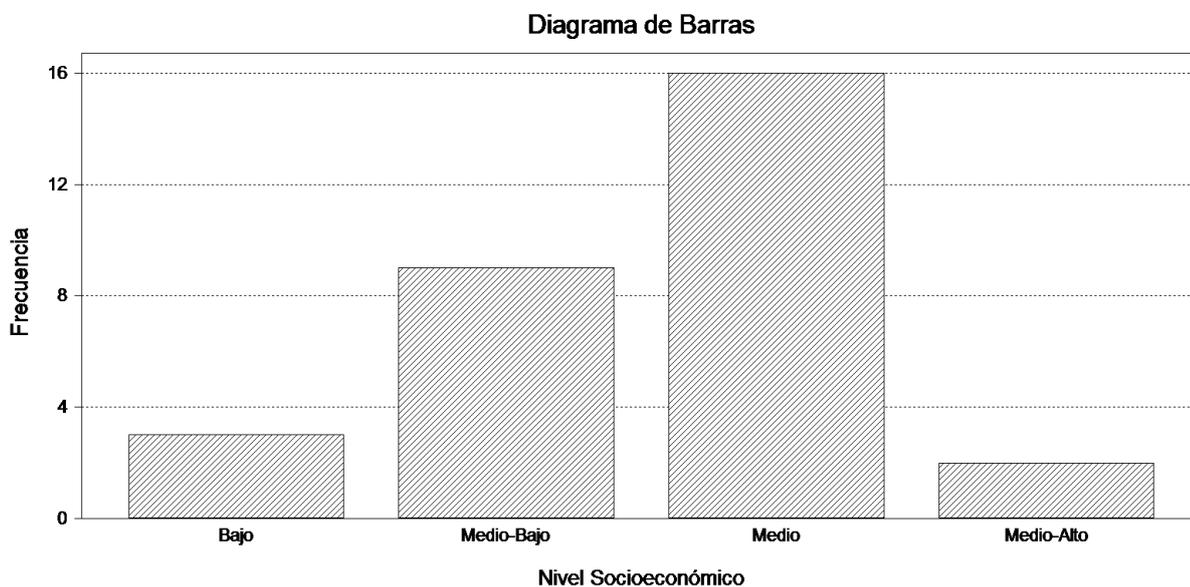
Nótese que en el caso de la edad, al corresponder a la población general, hubo un amplio espectro de edades, con lo que difícilmente coincidieron varios valores; la mayor frecuencia fue solamente 2. En esos casos la tabla no presenta mucha utilidad porque termina siendo prácticamente una enumeración de los valores observados de menor a mayor. Por eso se prefiere, para resumir y visualizar mejor la tendencia, presentar una tabla de frecuencia con los valores agrupados. Por ejemplo, por décadas:

Edad	$f_i$	$f_i'$	$f_i\%$
20 – 29	8	0,2667	26,67
30 – 39	4	0,1333	13,33
40 – 49	8	0,2667	26,67
50 – 59	6	0,2000	20,00
60 – 69	4	0,1333	13,33



d) Obtenga manualmente las tablas de distribución de frecuencias absolutas y absolutas acumuladas del nivel socioeconómico. Grafique la primera.

Nivel Socioeconómico	$f_i$	$F_i$
Bajo	3	3
Medio-Bajo	9	12
Medio	16	28
Medio-Alto	2	30
Total	30	

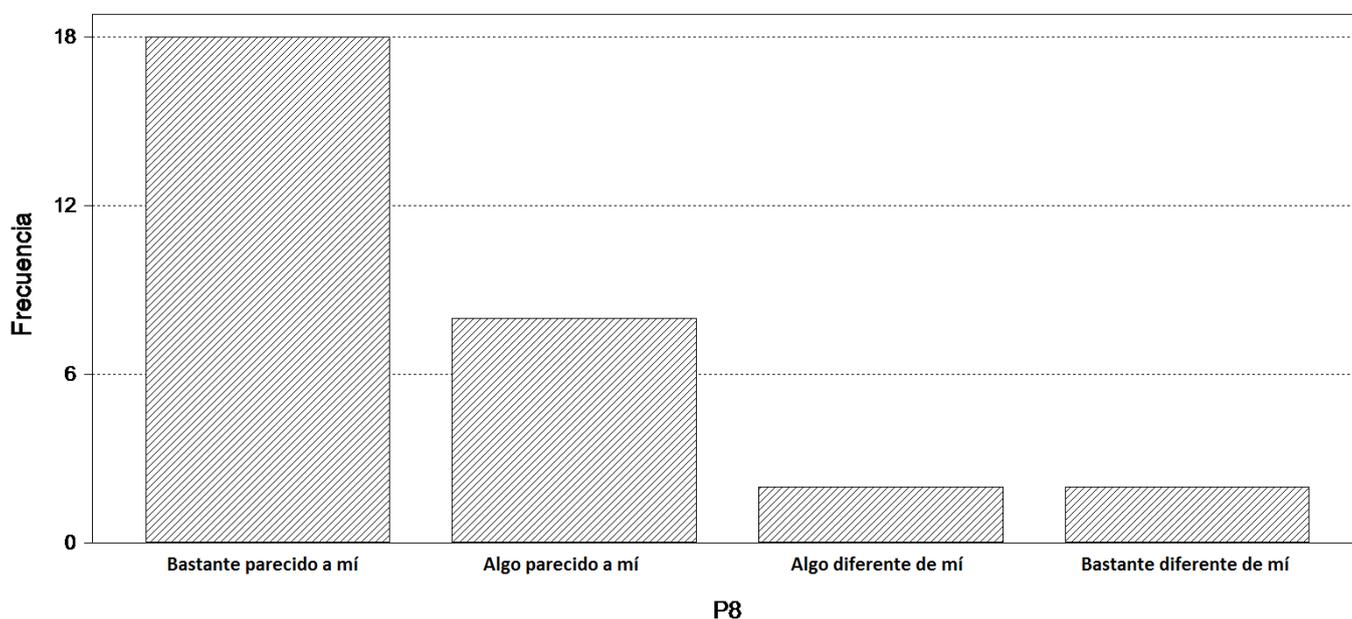


e) Realice manualmente una tabla con la distribución de frecuencias porcentuales de las variables P08 y P09. Grafíquelas. ¿Encuentra alguna tendencia general en el comportamiento u opinión sobre la puntualidad? Si fuera el caso ¿cuál?

P8: Opino que la impuntualidad es una falta de consideración hacia los demás.

$x_i$	$f_i$	$f_i\%$
Bastante parecido a mí	18	60
Algo parecido a mí	8	26,7
Algo diferente de mí	2	6,7
Bastante diferente de mí	2	6,7
Total	30	100

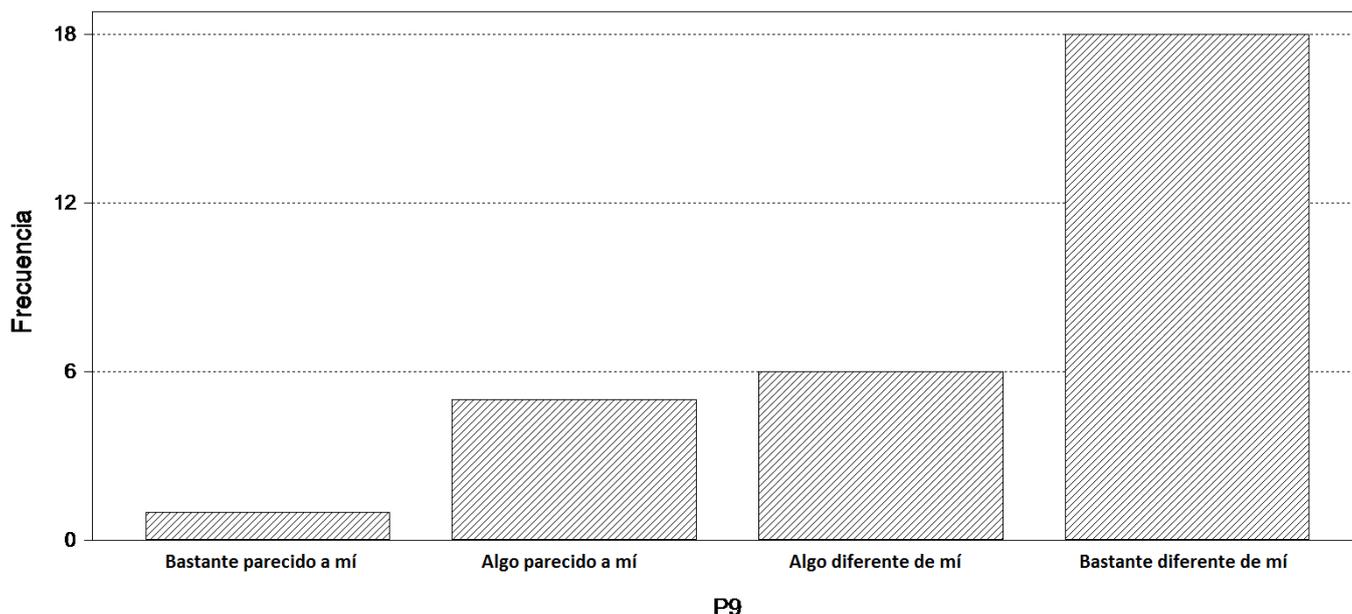
Diagrama de Barras



P9: Por lo general llego tarde a los lugares donde voy.

$x_i$	$f_i$	$f_i\%$
Bastante parecido a mí	1	3,3
Algo parecido a mí	5	16,7
Algo diferente de mí	6	20
Bastante diferente de mí	18	60
Total	30	100

Diagrama de Barras

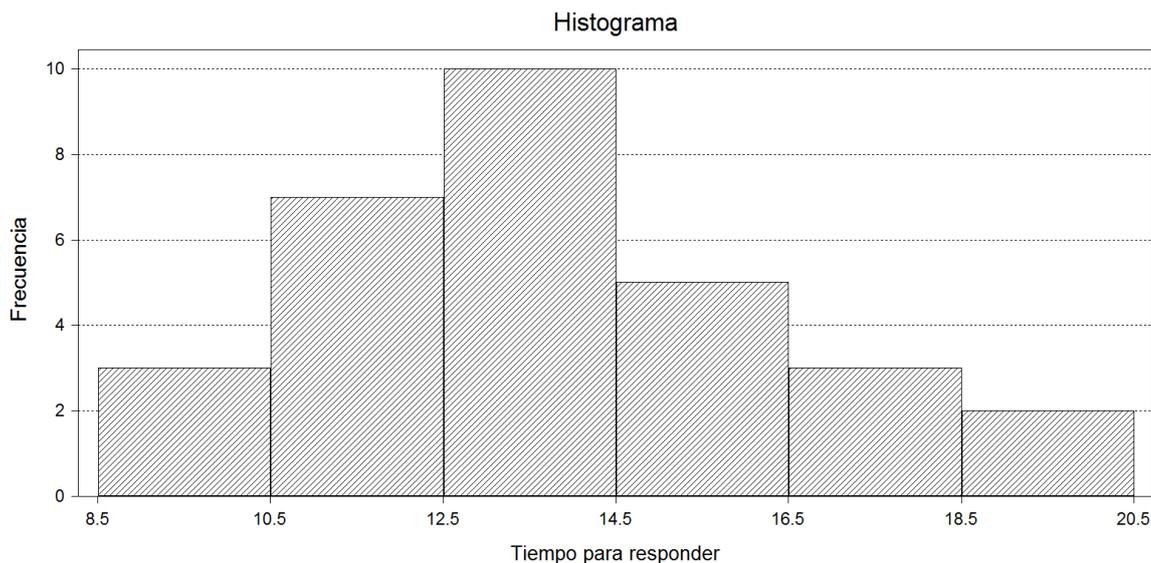


En los dos gráficos se observa una tendencia pronunciada en favor de la puntualidad. En el primer caso, mayoritariamente las personas tienden a calificar la impuntualidad como una desconsideración y, en el segundo caso, se autoperceben como puntuales. Ambas tendencias son congruentes ya que, cabe esperar que si una persona piensa que la impuntualidad es una desconsideración hacia los demás, trate ella misma de ser puntual. La recíproca no necesariamente es verdadera porque una persona podría ser puntual y atribuir las impuntualidades ajenas a otros factores. De todos modos, según estos resultados, que son reales, ambas tendencias van en el mismo sentido. El hecho de que la frecuencia aparezca distribuida casi de manera opuesta en ambos diagramas, se debe al sentido en que están hechas las preguntas. Si la pregunta 9 estuviera formulada al revés: Por lo general llego puntualmente a los lugares donde voy, probablemente ambos diagramas habría sido similares.

f) Obtenga manualmente la tabla de frecuencias simples y acumuladas para P11 agrupando el tiempo en intervalos de clase y gráfiquela (tome la cantidad de intervalos conveniente según los tiempos observados en su comisión).

- i) ¿En torno a qué valores estuvieron los tiempos que la comisión necesitó para resolver el cuestionario?
- ii) ¿Qué porcentaje de personas necesitaron a lo sumo 12,5 minutos para responder el cuestionario?
- iii) ¿Qué porcentaje de personas necesitaron por lo menos 14,5 minutos para responder el cuestionario?

Tabla de distribución de frecuencias para el Tiempo para responder el cuestionario				
Intervalos	$f_i$	$f_i\%$	$F_i$	$F_i\%$
8,5 10,5	3	10,0	3	10,0
10,5 12,5	7	23,3	10	33,3
12,5 14,5	10	33,3	20	66,7
14,5 16,5	5	16,7	25	83,3
16,5 18,5	3	10,0	28	93,3
18,5 20,5	2	6,7	30	100,0
Total	30	100,0		



i) En términos generales los tiempos se concentraron mayormente en torno a los 13 minutos aproximadamente; ya que la mayor frecuencia corresponde al intervalo cuyo valor central es 13,5 y el intervalo anterior (10,5 a 12,5) tiene más peso que el posterior (14,5 a 16,5). El 73,3% de los tiempos estuvieron comprendidos entre 10,5 y 16,5.

ii) A lo sumo significa 12,5 “como máximo 12,5” o “menor o igual a 12,5”. Esa información está en la columna de las frecuencias porcentuales acumuladas y es 33,3%.

iii) Por lo menos 14,5 significa “como mínimo 14,5” o “mayor o igual a 14,5”. Utilizando la misma columna hacemos  $100\% - 66,7\% = 33,3\%$ .

g) Utilice la base de datos de este problema y un programa informático (InfoStat, Statistix, EXCEL, GeoGebra, etc.) para hallar las distribuciones de frecuencia de las variables P1, P2, P3, P7, P8, P9 y P10. Grafíquelas. ¿Encuentra alguna tendencia general en el comportamiento y opinión sobre la puntualidad? Si fuera el caso ¿cuál?

Utilizaremos el programa Statistix

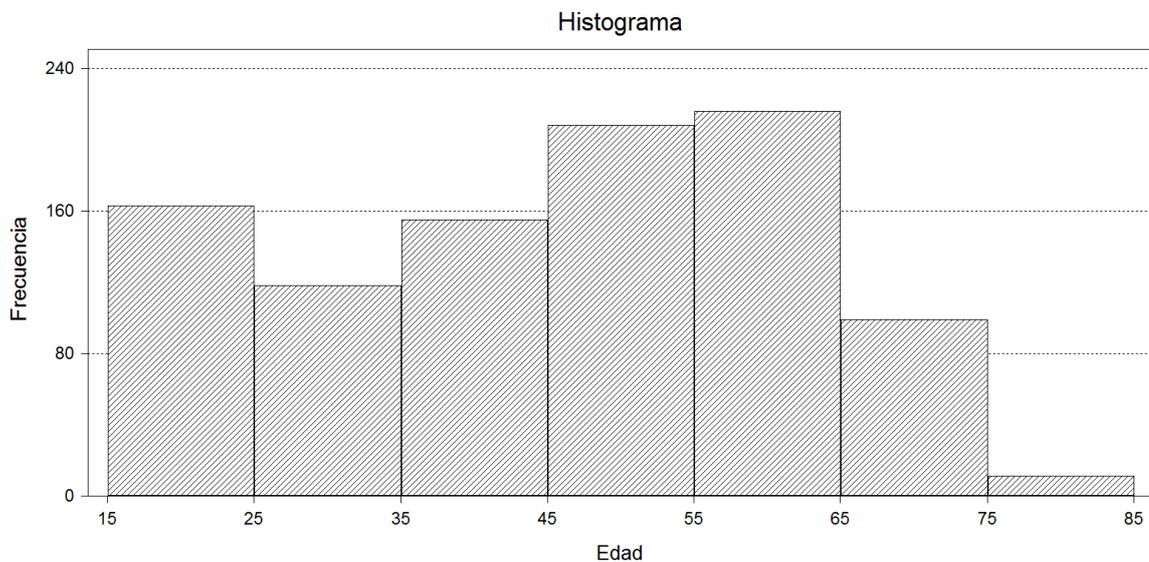
- 1) Abrir el programa y clicar en la pestaña File-Open, allí elegir el archivo “Base impunt larga”.
- 2) Clicar sucesivamente las pestañas Statistics – Summary Statistics – Frequency Distribution. Elegir de la ventana izquierda las variables y pasarlas a la ventana de la derecha. Clicar OK. Si se quiere agrupar en intervalos, hay que definir los valores mínimo y máximo, como también la amplitud del intervalo.
- 3) Si se desea que los valores de la variable aparezcan cualitativamente, ir previamente a las pestañas Data – Labels – Value Labels y etiquetar cada valor de la variable. En Variable Labels también se puede poner el nombre de las variables; por ejemplo P03: Sexo.
- 4) Para los gráficos, ir a las pestañas Statistics – Summary Statistics. Elegir Pie chart para el diagrama circular. Los diagramas de barras y los histogramas están bajo un único título de Histogram. Una vez Si la variable es continua hay que definir los valores mínimo, máximo y la amplitud del intervalo de igual manera que cuando se obtiene la tabla de datos agrupados. Los títulos de los diagramas se pueden cambiar entrando en la pestaña de Results - Titles

A continuación se muestran las salidas computacionales obtenidas con Statistix

Agrupamos la edad en intervalos por décadas. El programa “continúa” la variable y hace los intervalos contiguos. Cada extremo del intervalo es contado en el intervalo siguiente.

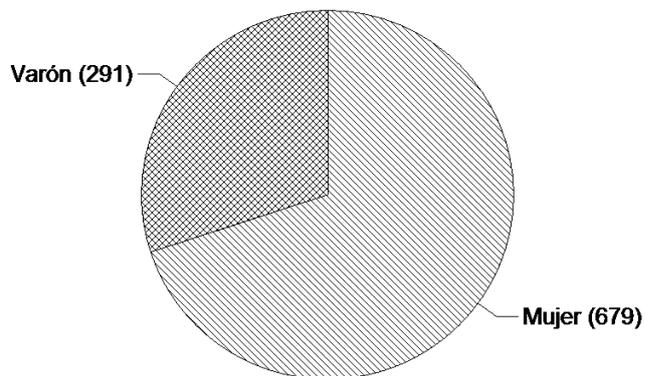
**Frequency Distribution of P01: Edad**

Low	High	Freq	Percent	Cumulative	
				Freq	Percent
15	25	163	16.8	163	16.8
25	35	118	12.2	281	29.0
35	45	155	16.0	436	44.9
45	55	208	21.4	644	66.4
55	65	216	22.3	860	88.7
65	75	99	10.2	959	98.9
75	85	11	1.1	970	100.0
Total		970	100.0		



**Frequency Distribution of P02: Sexo**

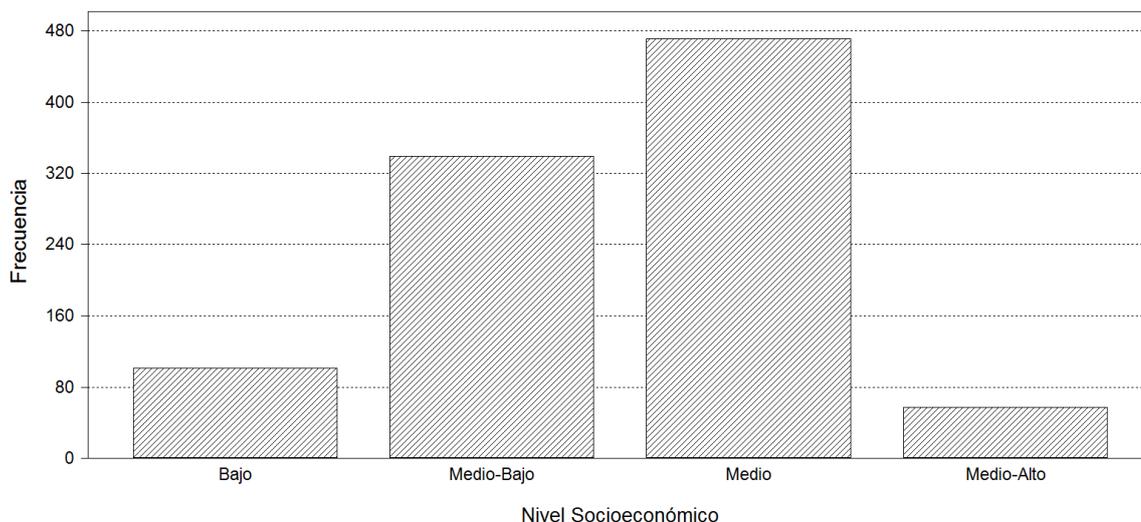
Value	Freq	Percent	Cumulative	
			Freq	Percent
Mujer	679	70.0	679	70.0
Varón	291	30.0	970	100.0
Total		970	100.0	



**Frequency Distribution of P03: Nivel Socioeconómico**

Value	Freq	Percent	Cumulative	
			Freq	Percent
Bajo	101	10.4	101	10.4
Medio-Bajo	339	35.0	440	45.5
Medio	471	48.7	911	94.1
Medio-Alto	57	5.9	968	100.0
Total	968	100.0		

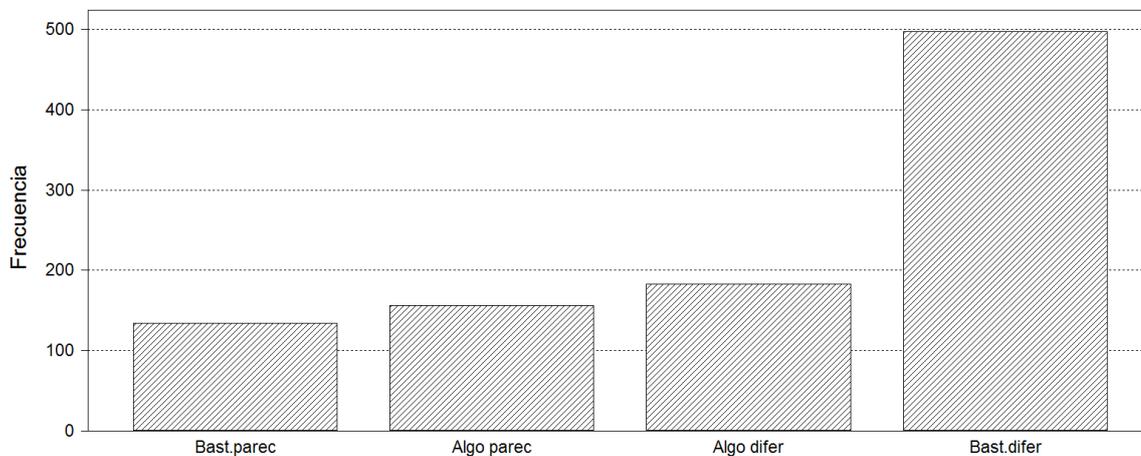
Diagrama de Barras



**Frequency Distribution of P07**

Value	Freq	Percent	Cumulative	
			Freq	Percent
Bast.parec	134	13.8	134	13.8
Algo parec	156	16.1	290	29.9
Algo difer	183	18.9	473	48.8
Bast.difer	497	51.2	970	100.0
Total	970	100.0		

Diagrama de Barras

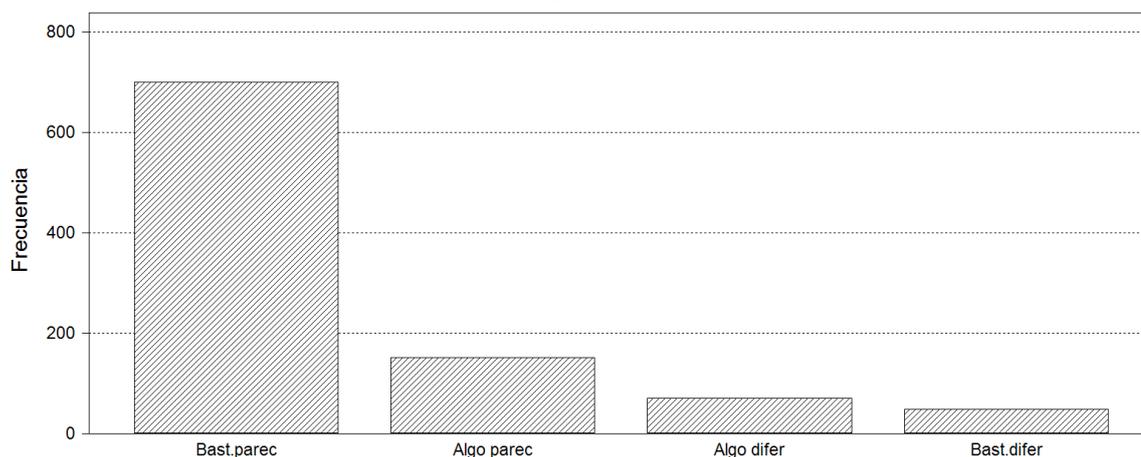


Antes de salir de casa doy tantas vueltas que termino llegando tarde

### Frequency Distribution of P08

Value	Freq	Percent	Cumulative	
			Freq	Percent
Bast.parec	700	72.2	700	72.2
Algo parec	151	15.6	851	87.7
Algo difer	71	7.3	922	95.1
Bast.difer	48	4.9	970	100.0
Total	970	100.0		

Diagrama de Barras

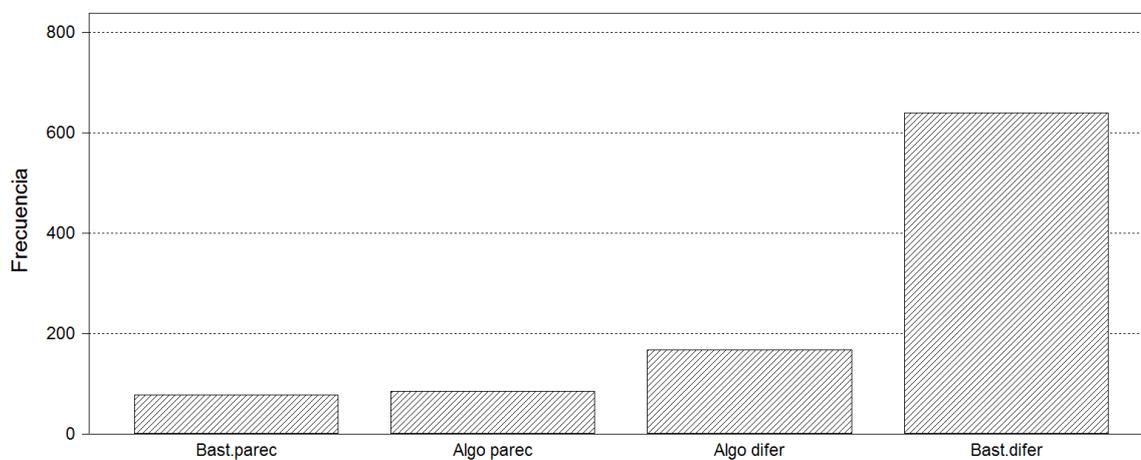


Opino que la impuntualidad es una falta de consideración hacia los demás.

### Frequency Distribution of P09

Value	Freq	Percent	Cumulative	
			Freq	Percent
Bast.parec	78	8.0	78	8.0
Algo parec	86	8.9	164	16.9
Algo difer	167	17.2	331	34.1
Bast.difer	639	65.9	970	100.0
Total	970	100.0		

Diagrama de Barras

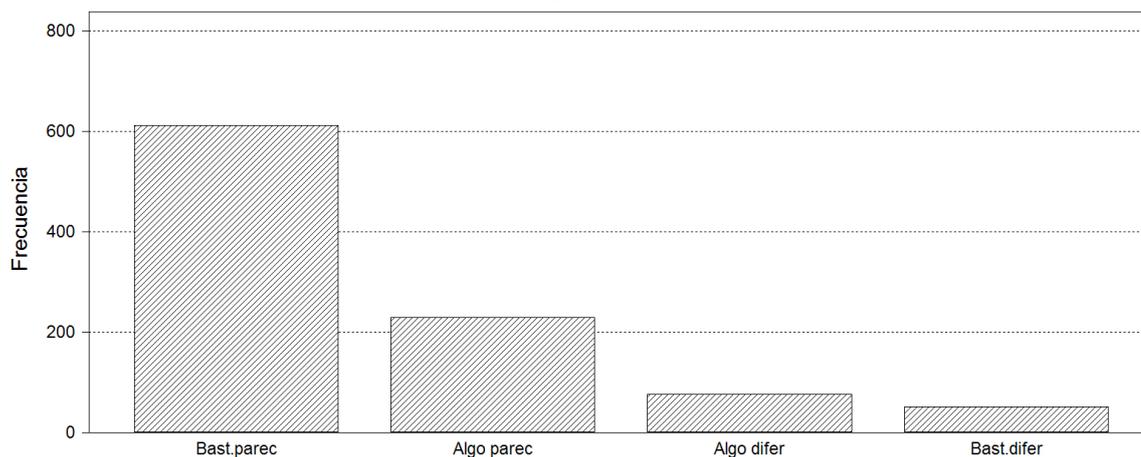


Por lo general llego tarde a los lugares donde voy.

### Frequency Distribution of P10

Value	Freq	Percent	Cumulative	
			Freq	Percent
Bast.parec	612	63.1	612	63.1
Algo parec	230	23.7	842	86.8
Algo difer	76	7.8	918	94.6
Bast.difer	52	5.4	970	100.0
Total	970	100.0		

Diagrama de Barras



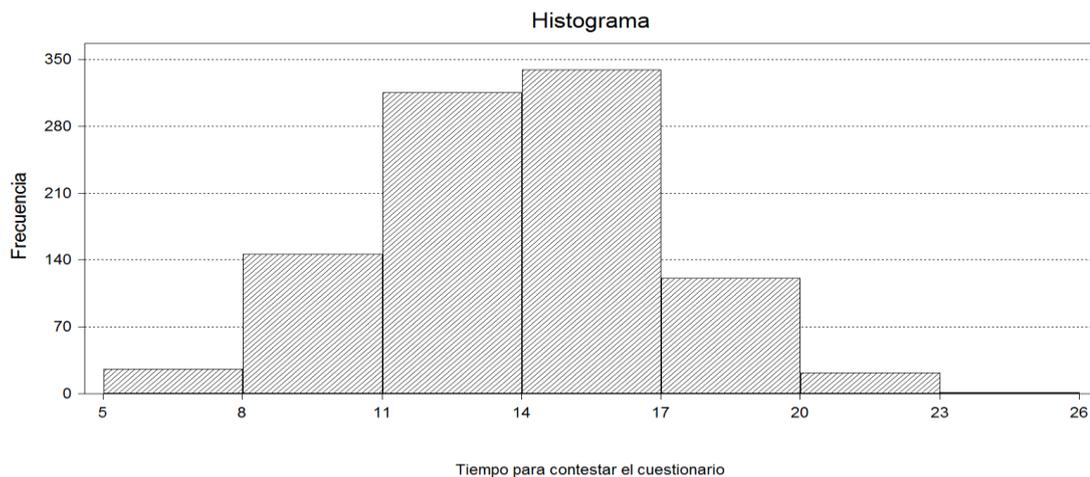
Me molesta que me hagan esperar cuando quedamos en un horario.

Las distribuciones de frecuencias de las respuestas a las cuatro preguntas son similares en cuanto a que se concentran más en los valores favorables a la valoración o práctica de la puntualidad. Puede concluirse que la tendencia de la población muestreada es a autoperibirse puntuales y a considerar la impuntualidad como un disvalor.

h) Halle una distribución de frecuencias para datos agrupados de la variable P11 y el correspondiente histograma.

### Frequency Distribution of P11

Low	High	Freq	Percent	Cumulative	
				Freq	Percent
5	8	26	2.7	26	2.7
8	11	146	15.1	172	17.7
11	14	315	32.5	487	50.2
14	17	339	34.9	826	85.2
17	20	121	12.5	947	97.6
20	23	22	2.3	969	99.9
23	26	1	0.1	970	100.0
Total		970	100.0		



i) ¿En torno a qué valores estuvieron los tiempos que la comisión necesitó para resolver el cuestionario?

Los tiempos estuvieron concentrados en torno a 14 minutos aproximadamente, con un gran porcentaje comprendido entre 11 y 17 minutos.

ii) ¿Qué porcentaje de personas necesitaron a lo sumo 14 minutos para responder el cuestionario?

El 50,2 % de las personas necesitó como máximo 14 minutos.

iii) ¿Qué porcentaje de personas necesitaron por lo menos 17 minutos para responder el cuestionario?

$100\% - 85,2\% = 14,8\%$  de personas necesitaron por lo menos 17 minutos.

2.- La siguiente tabla presenta las estadísticas de cierto centro terapéutico con respecto al tipo de trastorno diagnosticado en los pacientes que consultaron durante el último año.

DIAGNÓSTICO	PORCENTAJE
Trastorno alimentario	13
Trastorno de ansiedad	31
Distrés	24
Trastorno obsesivo-compulsivo	18
Trastorno de Sueño	9
Otros	5

a) Mencione la variable que se está considerando y clasifíquela.

Variable: Tipo de trastorno diagnosticado a los pacientes que consultaron en el último año. Es cualitativa medida en nivel nominal.

b) Represente la información en un gráfico adecuado.



c) ¿Por qué podría ser de interés disponer de la información que da la tabla?

Es información útil para epidemiología o, por ejemplo, para saber qué tipo de formación específica se ponderará de los profesionales del centro terapéutico.

d) Halle las frecuencias absolutas de cada categoría suponiendo que al centro asistieron 800 personas el último año.

DIAGNÓSTICO	PORCENTAJE	Frecuencia absoluta
Trastorno alimentario	13	104
Trastorno de ansiedad	31	248
Distrés	24	192
Trastorno obsesivo-compulsivo	18	144
Trastorno de Sueño	9	72
Otros	5	40
Total	100	800

- e) Durante el año anterior, de un total de 990 personas que fueron a la consulta, 280 tuvieron trastornos de ansiedad. ¿Podría decirse que, con relación al año anterior, en ese centro se observa una tendencia decreciente de consultas por problemas de ansiedad? Justifique la respuesta.

Para responder a ello es posible hacer uso de las frecuencias relativas o porcentuales, dado que se trata de grupos con diferente tamaño. Las 280 personas con trastorno de ansiedad representan el 28,3% del grupo de 990 pacientes que consultaron el año anterior. Según el dato de la en el último año el 31% de los consultantes recibió ese diagnóstico. Al comparar estos porcentajes se puede afirmar a nivel descriptivo que el porcentaje de personas con trastornos de ansiedad aumentó de un año al siguiente (de 28,3% a 31%); es decir que no hay una tendencia decreciente. Tampoco se puede hablar de una tendencia creciente (a pesar del resultado muestral) porque el cambio es pequeño y habría que verificar que no se deba a fluctuaciones azarosas (ese tema corresponde a inferencia estadística).

- f) **Para conceptualizar:** Compare la información que dan las frecuencias absolutas, las relativas y las porcentuales. ¿En qué casos es indistinto, preferible o indispensable usar una u otra?

Para responder cabalmente la pregunta es recomendable que releas el Capítulo 2 del libro de Bologna (2018) y el material didáctico con Explicaciones Complementarias a los Capítulos 1 y 2 del libro de Bologna de García Díaz (2020).

**3.-** Cierta centro de estudios universitarios tiene que planificar las vacantes para cada carrera, lo que implica, entre otras cosas, determinar el equipamiento de cada aula en cuanto a butacas y eventualmente computadoras. Le informan que para el año entrante el 30% de inscriptos quiere ingresar a la carrera de Psicopedagogía, el 20% a Musicoterapia y el resto a Psicología. ¿Es esta información suficiente para la planificación? Justifique.

Se necesita la información del total de inscriptos. De esta manera tomarán sentido los porcentajes para calcular de forma precisa la cantidad de equipamiento destinado a cada carrera.

**4.-** Dos docentes de Estadística comentan los resultados obtenidos por sus estudiantes en el primer parcial que tomaron. Carla dice, gratamente sorprendida, que tuvo 3 alumnos con 10 y Tomás le contesta jocosamente: “Yo tuve 4 alumnos con 10, te gané”. Discuta en qué casos el “te gané” de Tomás puede sustentarse.

Sólo puede sustentarse en los casos en que los cuatro alumnos de Tomás impliquen un porcentaje mayor sobre el total de la clase, que el porcentaje que significan los 3 alumnos de Carla sobre el total de sus respectivos alumnos.

**5.-** Un profesor de castellano de primer año de escuela media toma una exhaustiva prueba diagnóstica a sus 40 alumnos. Los puntajes están en una escala de 0 a 100 y su distribución se muestra en la figura 1. Después de analizarlos, decide dedicar el primer mes de clase a un repaso general a fin de emparejar y mejorar el desempeño del grupo antes de comenzar a enseñar su programa específico. Después de dicho repaso vuelve a tomar una prueba diagnóstica similar a la anterior, cuya distribución de puntajes se muestra en la figura 2. a) Según los histogramas, ¿se obtuvieron resultados favorables a los objetivos del profesor? Fundamente su respuesta comparando dichos histogramas.

El histograma que representa la distribución de los puntajes de la prueba diagnóstica previa al repaso

tiene valores que oscilan entre 4.5 y 44.5. Luego del repaso, el puntaje mínimo ascendió a 12.5 y el máximo resultó de 60.5; es decir que los puntajes aumentaron. Estos hallazgos son favorables a los objetivos del profesor porque muestran una tendencia de los alumnos a mejorar el desempeño en la prueba diagnóstica después del repaso. También logró el objetivo de emparejar el grupo, ya que los puntajes están más dispersos antes del repaso que después, donde se concentran más en torno a un valor medio.

**b) Conceptualice:**

i) ¿Para qué tipo de variable es recomendable el uso de histogramas?

Para elaborar la respuesta a este ítem se sugiere que revise el Capítulo 2 del libro de Bologna (2018), el material didáctico con Explicaciones Complementarias a los Capítulos 1 y 2 del libro de Bologna de García Díaz (2020) y la clase teórica 2.

ii) Relacione la forma de las distribuciones representadas por los histogramas con la concentración o dispersión de los datos (homogeneidad o heterogeneidad de los grupos).

El primer histograma es más “chato”, más plano, que el segundo. Eso significa que las frecuencias se distribuyen más equitativamente a lo largo de todo el rango de variación de la variable, mientras que en el segundo caso se concentra mayormente en valores centrales. El rango de variación (la distancia entre el mínimo y el máximo puntaje) en ambos casos es igual, de 40 puntos. Antes del repaso hay puntajes en todos los niveles, aproximadamente en igual proporción, como muestra la forma casi plana del histograma (distribución aproximadamente uniforme); en otras palabras, todos los puntajes se presentan más o menos en igual cantidad. Después del repaso, los puntajes muy bajos o muy altos son menos frecuentes y la mayoría corresponde a valores medios. Esto se muestra en la forma más apuntada del histograma. Al concentrarse la mayor frecuencia en un intervalo menor, los puntajes son más parecidos entre sí que en el caso anterior. Por tanto el grupo es más heterogéneo antes del repaso que después.

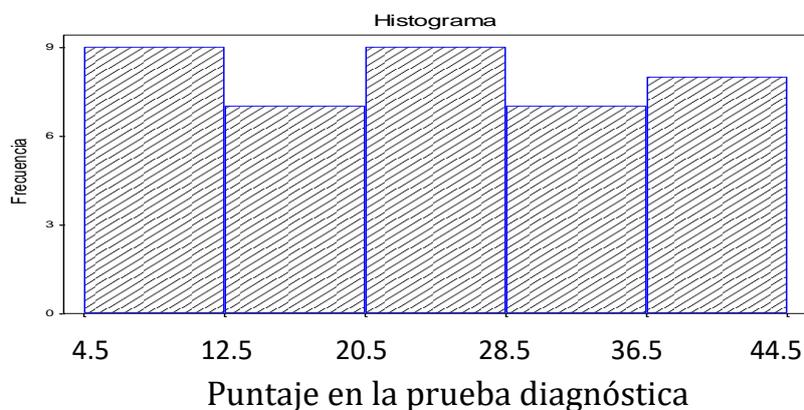


Figura 1. Distribución de los puntajes antes del repaso

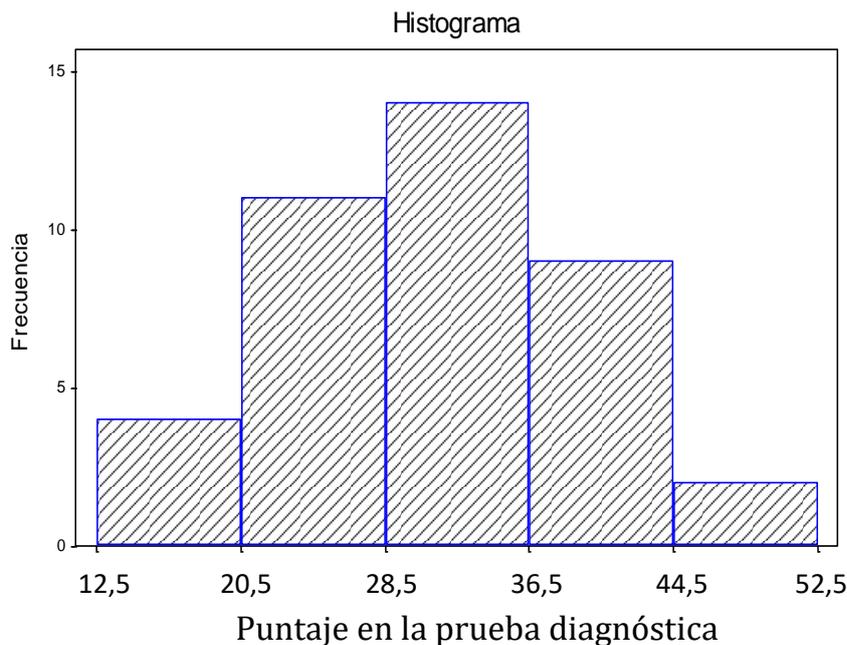


Figura 2. Distribución de los puntajes después del repaso

6.- El profesor aludido en el ejercicio 5 debe presentar los resultados en un informe donde no puede poner gráficos. Vuelque la información de los mismos en sendas tablas de frecuencias.

Distribución de Frecuencias de los puntajes antes del repaso

Puntaje en la prueba diagnóstica	Frecuencias			Frecuencias acumuladas		
	Absoluta	Relativa	Porcentual	Absoluta	Relativa	Porcentual
4,5 - 12,5	9	0,225	22,5	9	0,225	22,5
12,5 - 20,5	7	0,175	17,5	16	0,400	40,0
20,5 - 28,5	9	0,225	22,5	25	0,625	62,5
28,5 - 36,5	7	0,175	17,5	32	0,800	80,0
36,5 - 44,5	8	0,200	20,0	40	1	100,0
Total	40	1	100,0			

Distribución de Frecuencias de los puntajes después del repaso

Puntaje en la prueba diagnóstica	Frecuencias			Frecuencias acumuladas		
	Absoluta	Relativa	Porcentual	Absoluta	Relativa	Porcentual
12,5 - 20,5	4	0,100	10,0	4	0,100	10,0
20,5 - 28,5	11	0,275	27,5	15	0,375	37,5
28,5 - 36,5	14	0,350	35,0	29	0,725	72,5
42,5 - 50,5	9	0,225	22,5	38	0,950	95,0
50,5 - 58,5	2	0,050	5,0	40	1	100,0
Total	40	1	100,0			

7.- En cierta investigación (datos ficticios) se registró, en décimas de segundo, el tiempo de reacción a un estímulo de 20 adolescentes entre 15 y 17 años y de 20 adultos entre 57 y 60 años. Los resultados se muestran respetivamente en los diagramas de tallo-hoja 1 y 2.

**Tiempo de reacción de los adultos**

Unidad de las hojas = 0,1  
5 7 representa 5,7

Tallo	Hojas
5	7
6	48
7	224677
8	13556899
9	237

**Tiempo de reacción de los adolescentes**

Unidad de las hojas = 0,1  
5 7 representa 5,7

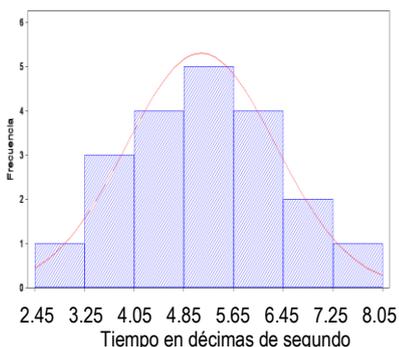
Tallo	Hojas
5	7
6	12556899
7	13778
8	0226
9	07

I) ¿Qué puede concluir, en general, de la comparación de ambos grupos?

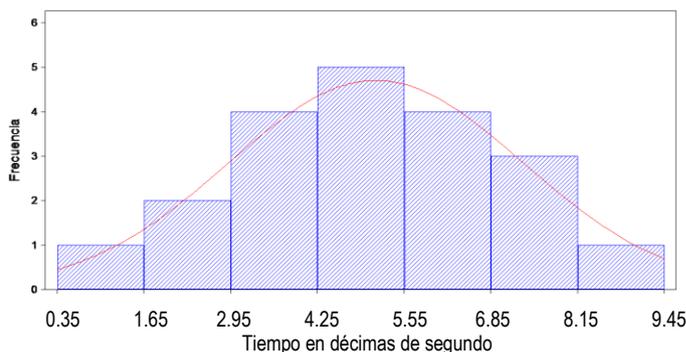
Los diagramas muestran que ambos grupos registraron los mismos valores mínimo (5,7 décimas de segundo) y máximo (9,7 décimas de segundos). No obstante, los datos pertenecientes a los adultos tendieron a concentrarse en los tallos 7 y 8 en tanto que los adolescentes se distribuyeron mayoritariamente en los tallos 6 y 7. Considerando estas distribuciones podría afirmarse que el grupo de adolescentes, en términos generales, mostró tiempos de reacción más rápidos que los adultos.

II) Las 40 personas que participaron de la experiencia fueron luego entrenadas para mejorar su tiempo de reacción. Después del entrenamiento se midieron nuevamente los tiempos de reacción. Como los datos los analizó otro estadístico, utilizó histogramas en lugar de diagramas de tallo-hoja para presentar los resultados de esta etapa, los cuales se muestran a continuación.

**Tiempo de reacción de los adultos**



**Tiempo de reacción de los adolescentes**



a) ¿Parece haber tenido el entrenamiento algún efecto sobre cada grupo? Si es así, ¿cuál?

Los resultados parecen señalar que el entrenamiento ha generado una reducción en los tiempos de reacción de ambos grupos. Además, las distribuciones de los histogramas muestran una forma más simétrica en comparación con los respectivos diagramas tallo-hoja de cada grupo. Para el grupo de los adultos en el gráfico de tallo-hoja se nota una concentración de los individuos en los valores de los tallos 7 y 8. Al observar el histograma correspondiente a este grupo se advierte que luego del entrenamiento los puntajes de estos individuos se concentraron mayoritariamente en los tres intervalos definidos entre 4,05 y 6,45. Esta

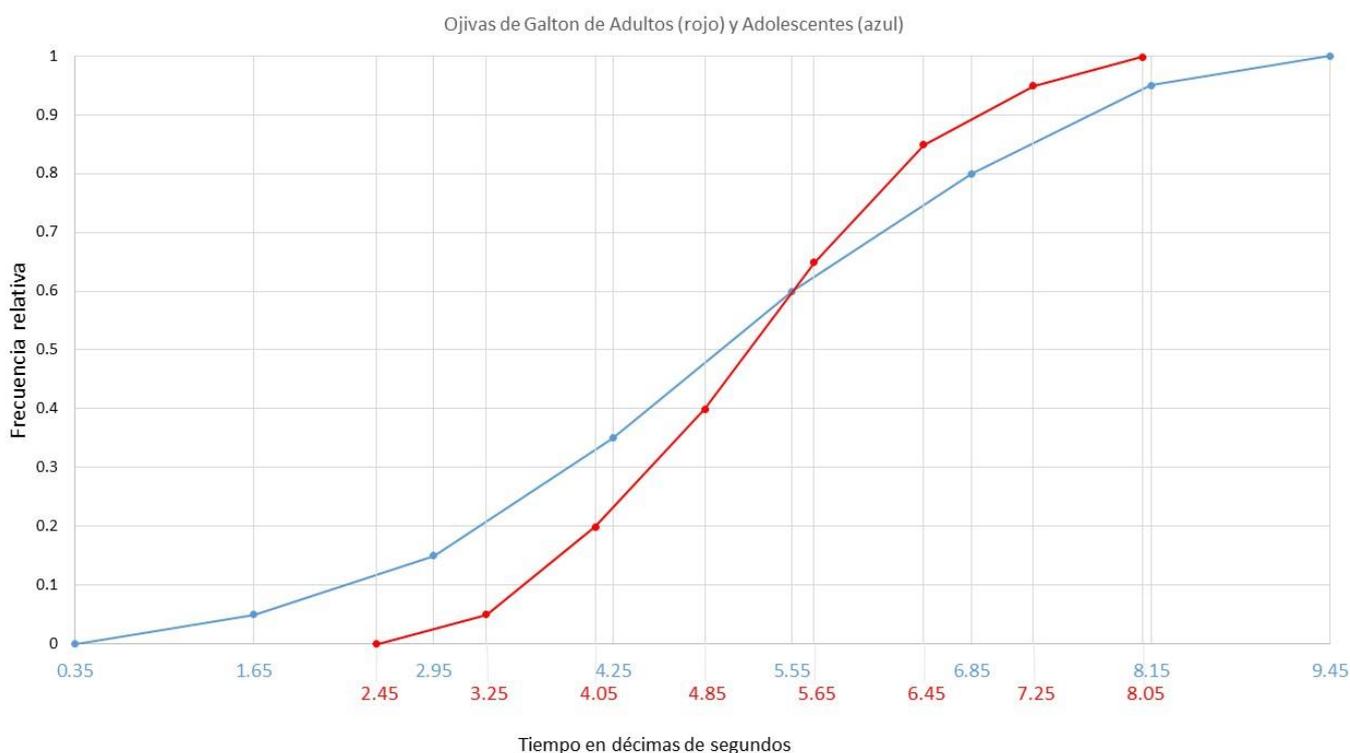
reducción también es apreciable en la comparación de los valores mínimo y máximo registrados antes y después del entrenamiento. Previamente a la práctica los adultos reportaron tiempos entre 5,7 y 9,7 décimas y luego de la práctica los valores observados oscilaron entre 2,45 y 8,05 décimas. En la misma línea, en el tallo-hoja se aprecia que los adolescentes tendieron a agruparse en los valores de tallo 6 y 7 antes del entrenamiento y luego del mismo la mayoría presentó valores entre 2,95 y 6,85. Además, sus tiempos mínimos y máximos fueron de 5,7 y 9,7 a 0,35 y 9,45, lo que indica, en términos generales, que el entrenamiento ha resultado eficaz también para este grupo.

b) Compare las distribuciones de los tiempos entre ambos grupos después del entrenamiento.

Al realizar la comparación de los dos histogramas con los puntajes después del entrenamiento resalta que el efecto de la práctica ha tenido un impacto diferente en los grupos. El histograma de los adolescentes muestra un valor mínimo de 0,35 mostrando que algunos sujetos han conseguido reducir sus tiempos de reacción de manera apreciable. Pero el valor máximo de 9,45 no se aleja tanto del máximo observado previo al entrenamiento (9,7), por lo que podría deducirse que el impacto no fue parejo para todos los sujetos. En el caso de los adultos también se produjo una situación similar, aunque menos pronunciada. Los valores mínimo y máximo luego del entrenamiento variaron entre 2,45 y 8,05, lo que muestra que este grupo se desempeñó de manera más homogénea.

c) **Conceptualice** las representaciones por histogramas y por diagramas de tallo-hoja. ¿Qué información similar proporcionan? ¿Qué ventaja tiene el diagrama de tallo-hoja con respecto al histograma? Para elaborar la respuesta a este ítem se sugiere que revise el Capítulo 2 del libro de Bologna (2018) y el material didáctico con Explicaciones Complementarias de García Díaz (2020).

d) Obtenga manualmente, en un mismo par de ejes cartesianos, la ojiva de Galton para cada histograma y compare la forma de ambas ojivas. **Conceptualice:** ¿Qué información da el eje de ordenadas?



Se observa que la ojiva correspondiente a los adultos es más empinada, crece más rápidamente que la de los adolescentes; eso tiene relación directa con el hecho de que las frecuencias están más concentradas en un rango menor de valores, por tanto la variación de frecuencia acumulada de un intervalo a otro es mayor. La ojiva correspondiente a los adolescentes es más suave porque hay más variabilidad.

Si se trabaja, como en este caso, con la frecuencia relativa, el eje de ordenadas indica la proporción de individuos con tiempos de reacción de a lo sumo el valor de abscisa considerado. Por ejemplo, para los adultos, puede afirmarse que una quinta parte de las personas (0,2) reacciona en a lo sumo 4,85 décimas de segundos.

8.- Proponga una variable continua que pueda medir en su entorno. Tome 30 mediciones de la misma y represente la distribución observada con un diagrama de tallo-hoja y con un histograma.

*Nota. Como alternativa, en lugar de efectuar Ud. mismo las mediciones, puede buscar datos en alguna base como, por ejemplo, del sitio <https://data.world/datasets/psychology> o simularlos con una aplicación de generación de números aleatorios como <https://pinetools.com/es/generador-gaussiano-numeros-aleatorios>*

Para apreciar la diferencia conceptual entre el histograma y el diagrama de tallo-hoja lo remitimos al material didáctico con Explicaciones Complementarias a los Capítulos 1 y 2 del libro de Bologna de García Díaz (2020).

9.- Los datos del siguiente ejercicio son ficticios pero basados sobre las estadísticas del artículo de Fava, GA, Park SK, Sonino, N (2006). Treatment of recurrent depression. *Expert Review of Neurotherapeutics* 6 (11), 1735-1740.

En una clínica 20 pacientes recibieron terapia por haber padecido un episodio depresivo mayor y se les realizó un seguimiento de por vida a partir del alta. A continuación se indica la cantidad de recaídas de dichos pacientes en este tipo de episodios.

4 - 6 - 0 - 1 - 5 - 2 - 0 - 3 - 5 - 0 - 3 - 4 - 0 - 4 - 3 - 5 - 4 - 3 - 2 - 4

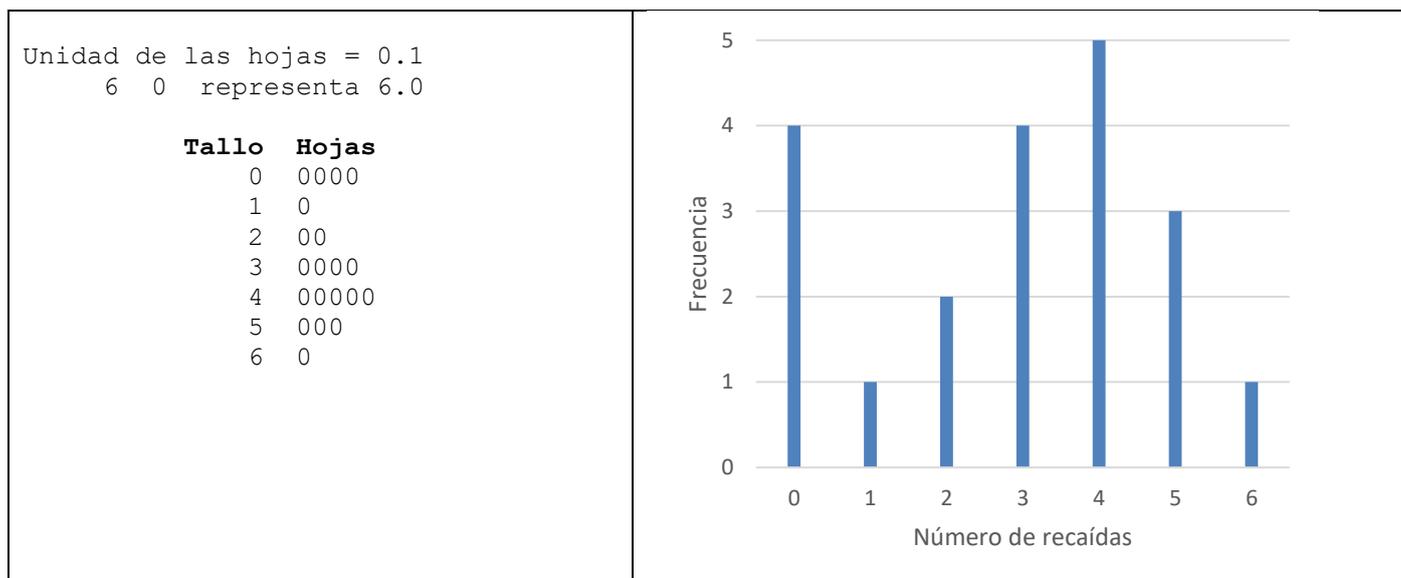
a) Mencione la variable en cuestión y clasifíquela.

La variable es “Número de recaídas registradas luego del alta tras haber recibido tratamiento por un episodio depresivo mayor”. Se trata de una variable cuantitativa discreta medida a nivel proporcional o de razón.

b) Organice los datos en una tabla de frecuencias y represente la distribución con un gráfico adecuado.

c) Obtenga en la misma tabla las frecuencias porcentuales acumuladas.

Número de recaídas	Frecuencias			Frecuencia Porcentual acumulada
	Absoluta	Relativa	Porcentual	
0	4	0,20	20	20
1	1	0,05	5	25
2	2	0,10	10	35
3	4	0,20	20	55
4	5	0,25	25	80
5	3	0,15	15	95
6	1	0,05	5	100
Total	20	1	100,0	



d) ¿Qué porcentaje de pacientes tuvo: i) al menos una recaída, ii) a lo sumo una recaída?

i) El 80% de los pacientes tuvo por lo menos una recaída. ii) El 25% de los pacientes presentó a lo sumo una recaída.

e) De los pacientes que tuvieron recaídas, ¿qué porcentaje sufrió más de una?

Al considerar los pacientes que tuvieron recaídas, se registra que un 93,75% sufrió más de una.

f) **Conceptualice:**

- ¿Por qué no es adecuado en este caso un histograma?
- ¿Podría haber utilizado un diagrama de tallo-hoja para organizar estos datos?
- ¿Qué similitudes y diferencias hay entre un diagrama de tallo-hoja y un diagrama de bastones?

Para orientar tu respuesta se han graficado aquí el diagrama de tallo-hoja y el diagrama de bastones. Para apreciar la diferencia conceptual entre ambos lo remitimos al material didáctico con Explicaciones Complementarias a los Capítulos 1 y 2 del libro de Bologna de García Díaz (2020).

**10.- Conceptualice:**

i) ¿Qué diferencia esencialmente a la distribución de frecuencias de una variable discreta de la de una continua?

ii) ¿Qué vinculación hay entre esta diferencia y sus respectivas formas de representación gráfica?

Para elaborar su respuesta se sugiere que recurra a la distinción entre las variables continuas y discretas que aparecen en el capítulo 1 de Bologna (2018), el material didáctico de García Díaz (2020) y las clases teóricas. Además, tenga en cuenta el apartado “Discretización de continuo por la medición” que aparece al final de la Práctica 1. Tener en mente estas nociones le permitirá articular mejor con las formas de representar gráficamente ambos tipos de variables.

**11.-** Imagine que va a tomar datos para cierta investigación en un hospital.

- Mencione una variable cualitativa de nivel nominal y otra de nivel ordinal que pudieran ser de interés.
- Mencione también las categorías de cada variable e invente una tabla de frecuencias para cada una de ellas; luego represéntelas con gráficos adecuados.

**Conceptualice:** En cada tabla de frecuencias Ud. puede hallar las frecuencias acumuladas pero esto no tiene sentido en el caso de la variable de nivel nominal y sí en el de la variable de nivel ordinal ¿Por qué?

Se sugiere que revise el concepto de frecuencia acumulada del capítulo 2 de Bologna (2018). Particularmente en el material didáctico de García Díaz (2020) encontrará una distinción entre acumuladas que siguen un sentido ascendente y descendente. Esta noción podría ayudar a pensar su respuesta a esta pregunta.

**12.-** Se registró el porcentaje de deserción escolar de los estudiantes de primer año en cada escuela de enseñanza media de Gran Buenos Aires correspondiente al año pasado. Indique cuáles son las unidades de análisis.

Las unidades de análisis son cada una de las escuelas de enseñanza media de Gran Buenos Aires. En esta investigación la variable en cuestión es el porcentaje de deserción escolar de los estudiantes de primer año de las escuelas de enseñanza media de Gran Buenos Aires. Los valores de dicha variable son los diferentes porcentajes de deserción y la frecuencia está determinada por la cantidad de escuelas de enseñanza media del Gran Buenos Aires en que se encuentra el mismo porcentaje. Por tanto la medición se efectúa sobre cada escuela de enseñanza media del Gran Buenos Aires.

**En los siguientes ítems elija la única opción correcta**

**13.-** En un diagrama circular, el ángulo de cada sector es directamente proporcional a la

- a. frecuencia del correspondiente valor de la variable.
- b. cantidad de unidades de análisis observadas.
- c. cantidad de categorías o valores de la variable.

**14.-** Como parte de un estudio de Planificación Comunitario, 50 personas fueron encuestadas sobre su Nivel de Escolaridad. Se encontró que la frecuencia absoluta de los que tenían estudios Secundarios Completos (SC) triplicaba a la frecuencia absoluta de los que tenían estudios Terciarios Completos (TC) (por ejemplo, SC = 15 y TC = 5). Entonces esta misma relación de 3 a 1 entre esos valores se mantendría también para sus frecuencias

- a. absolutas acumuladas.
- b. relativas.
- c. acumuladas porcentuales.

**15.-** Dada la distribución de frecuencias relativas de una variable, es posible conocer la distribución de frecuencias absolutas

- a. si se conoce el total de observaciones.
- b. se conozca o no el total de observaciones.
- c. si se conoce la distribución de frecuencias porcentuales.

**16.-** En la representación gráfica de las distribuciones de frecuencia, un histograma y un diagrama de barras anchas se utilizan

- a. independientemente del tipo de variable en cuestión.
- b. para variables continuas y cualitativas respectivamente.
- c. ambos como alternativa al diagrama circular.

**17.-** Cuando se desea comparar dos distribuciones provenientes de una cantidad distinta de datos la comparación debe basarse sobre las frecuencias

- a. absolutas o relativas.
- b. absolutas o porcentuales.
- c. relativas o porcentuales.