

INSTITUCIÓN EDUCATIVA EDUARDO FERNÁNDEZ BOTERO

ÁREA DE MATEMÁTICAS



ESTADÍSTICA II

MÓDULO GRADO 11°



Diseñado por:

Esp. María Cristina Marín Valdés

Amalfi (Ant)

2019

ESTADÍSTICA II

MÓDULO GRADO 11°



Diseñado por:

Esp. María Cristina Marín Valdés

ESTUDIANTE: _____

Institución Educativa “Eduardo Fernández Botero”

Área de Matemáticas

Amalfi

2019

CONTENIDOS

	Página
INTRODUCCIÓN	5
RESEÑA HISTÓRICA DE LA ESTADÍSTICA	6
IMPORTANCIA DE LA ESTADISTICA	8
ESQUEMA DE UNIDAD	10
CAPITULO I	
CONCEPTOS BÁSICOS DE ESTADÍSTICA	11
Conceptos de estadística	12
Técnicas de recolección de datos	17
Conteo y organización de datos	20
Frecuencia absoluta	20
Frecuencia relativa	22
Actividad de profundización No.1	24
Autoevaluación No.1	29
Tablas de frecuencia agrupadas en intervalos	31
Actividad de profundización No.2	33
Graficas estadísticas	34
Actividad de profundización No.3	37
Autoevaluación No.2	41
Medidas de tendencia central	46
Moda	46
Media aritmética	48
Mediana	52
Medidas de posición	59
Cuartiles	59
Deciles	68
Percentiles	71
Medidas de dispersión	73
Medidas de dispersión absoluta	74
Varianza	75
Desviación media	75
Desviación estándar	75
Medidas de dispersión relativa	81
Coeficiente de variación de Pearson	81
Actividad de profundización N.4	82
Autoevaluación No.3	87

CAPÍTULO II

PROBABILIDAD

Propiedades de la probabilidad

Probabilidad de eventos simples

Probabilidad de eventos independientes

Probabilidad condicional

CAPÍTULO III

TEORÍA DEL CONTEO

Permutaciones

Combinaciones

INTRODUCCIÓN

La estadística es una de las ramas más antiguas de las matemáticas, sus comienzos datan del antiguo Egipto, tiempo en el que los faraones lograron recopilar datos relativos a la población y a la riqueza de sus imperios. Los romanos, chinos y hebreos de la antigüedad hicieron más amplio el uso de la estadística, realizando censos de población aproximadamente cada 5 años, llevando minuciosas descripciones del número de nacimientos, matrimonios y defunciones ocurridas entre censo y censo.

En la actualidad la estadística nos brinda métodos y técnicas para organizar, analizar e interpretar información, convirtiéndose en un instrumento de investigación aplicable a todos los niveles de la ciencia y de la técnica. La industria, por ejemplo, precisa de información estadística para poder tomar decisiones en materia de inversión, planeación, ventas, producción, etc.; el Estado utiliza este instrumento para estimar la recaudación de impuestos, para control de precios y de productos, para proyectar la construcción de obras de infraestructura, para investigaciones en materia económica que le permitan aplicar nuevas y mejores políticas ajustadas a sus propósitos y metas.

Apreciado estudiante este módulo II es la continuidad de las temáticas estudiadas en el módulo I, contiene un mayor nivel de profundización, distribuyéndose en tres capítulos: Conceptos de estadística, es un compendio de conceptos básicos de recolección de información, tablas de distribución de frecuencias, representación gráfica, análisis de datos, medidas de centralización, medidas de posición y de dispersión, en el capítulo dos se profundiza el estudio de diferentes eventos y su probabilidad de ocurrencia, en el capítulo final se hace un estudio sobre teoría del conteo.

La metodología empleada tiene como propósito que los estudiantes dosifiquen su propio nivel de aprendizaje, con explicaciones narrativas de las diferentes temáticas, las cuales facilitan una mayor apropiación de los conceptos, el desarrollo de las actividades propuestas se hace de una manera activa, dinámica y participativa, buscando fortalecer el trabajo en equipo, con ejemplos y talleres basados en el contexto social de los educandos, pretendiendo que estos conocimientos les permitan desempeñarse de una manera más efectiva en diferentes pruebas censales y en la educación superior.

RESEÑA HISTÓRICA DE LA ESTADÍSTICA

La Estadística, como todas las ciencias, no surgió de improviso, sino mediante un proceso largo de desarrollo y evolución, desde hechos de simple recolección de datos hasta la diversidad y rigurosa interpretación de los datos que se dan hoy en día.

La palabra estadística proviene del latín “statisticus” que significa “del Estado”; es decir, correspondiente al gobierno. Por mucho tiempo, la estadística se refería a información numérica sobre los estados o territorios políticos. Como se conoce hoy en día, requirió de varios siglos para desarrollarse y de la intervención de muchas personas, teniendo como impulso la resolución de problemas prácticos planteados por la dinámica social de la época y teniendo siempre como objeto de estudio a la variación, es decir, la motivación la ha constituido el análisis de los valores que toman las diferentes variables de estudio a través de las cuales se analiza una población.

La historia de la estadística se puede resumir en tres etapas. A continuación se presentan los aspectos más importantes de cada una:

Primera Etapa (Los Censos): Desde que los pueblos se organizaron como Estados, sus gobernantes necesitaron estar informados sobre aspectos relativos a la cantidad o distribución de la información, nacimientos o defunciones, producción agrícola o ganadera, bienes muebles, bienes inmuebles, efectivos militares, etc., con el objeto de recaudar impuestos o de analizar las condiciones de vida de la población, la estadística se convierte entonces en un importante instrumento del Estado. Desde el momento en que se constituye una autoridad política, la necesidad de realizar inventarios de una forma regular a la población y las riquezas existentes en el territorio está ligada a la conciencia de soberanía y a los primeros esfuerzos administrativos.

Con base en los descubrimientos y sus evidencias sobre la recolección de datos referentes a población, bienes y producción, los orígenes de la estadística se remontan a civilizaciones muy antiguas tales como la Babilónica (5,000 años a.C.), Egipcia (3,000 años a. C.), China (2,200 años a.C.), Hindú (400 años a.C.), Romana (400 años a.C.), Griega (300 años a.C.). No hay que olvidar que fue un censo lo que motivó el viaje de José y María a Belén, trayecto en el cual nace Jesús.

Segunda Etapa: De la descripción de los conjuntos a la Aritmética Política: La estadística da un gran salto cualitativo a mediados del siglo XVII, debido a que los datos recopilados empiezan a ser utilizados por los bancos y por las nacientes compañías de seguros europeas; por otro lado, se inventa en Inglaterra el concepto de “Aritmética Política” y se empiezan a “matematizar” otras disciplinas, que hasta entonces eran sólo descriptivas, tales como la demografía, la economía y las ciencias sociales.

Para los aritméticos políticos, la estadística era el arte de gobernar, su función era de servir de ojos y oídos al gobierno. En esta época proliferan las tablas numéricas, lo cual permitió observar la frecuencia de distintos sucesos y el descubrimiento de leyes estadísticas. Son ejemplos notables los estudios de John Graunt sobre tablas de mortalidad y esperanza de vida, y los de Edmund Halley para resolver el problema de las rentas vitalicias de las compañías de seguros.



John Graunt encabeza una tendencia conocida como Estadística Investigadora. Buscaban fijar en números los fenómenos sociales y políticos. Para su tiempo esto fue atrevido, casi imposible; pero el mérito de ellos es de ser los primeros en buscar las leyes cuantitativas que rigen la sociedad.

Tercera Etapa (Estadística y Cálculo de Probabilidades): Otro impulso más al desarrollo de la estadística y la probabilidad es debido a los trabajos realizados por Jakob Bernoulli y Siméon Denis Poisson sobre las leyes de los grandes números. Este teorema fue el primer intento para deducir medidas estadísticas a partir de probabilidades individuales. Posteriormente surgen investigaciones sobre la teoría de los errores en las mediciones, teoría de los mínimos cuadrados, y la estadística logra con estos descubrimientos una relevancia científica.

Actualmente se puede decir que la Estadística es la ciencia que proporciona métodos para recopilar, organizar, presentar, resumir, analizar e interpretar información y poder tomar decisiones con cierto grado de confiabilidad. Hoy, la Estadística, junto con el cálculo de probabilidades, constituyen una rama fundamental de las matemáticas, con aplicaciones en casi todas las actividades humanas: física, astronomía, biología, genética, medicina, agricultura, química, y muchas más; en todas estas ciencias se hacen predicciones, encuestas, controles de calidad, estimaciones o verificaciones de hipótesis con respecto a parámetros poblacionales, todo ello ha permitido lograr avances científicos y tecnológicos; que a través de los años, han contribuido al desarrollo y bienestar social.

IMPORTANCIA DE LA ESTADÍSTICA

Los conceptos y argumentos de la estadística se utilizan en la actualidad en un gran número de ocupaciones. Las técnicas estadísticas constituyen una parte integral de las actividades de investigación en distintas áreas del saber humano. La persona que comprenda los conceptos estadísticos y su metodología obtendrá mejor provecho de ellos.

La estadística día con a día gana terreno en su aplicación en toda actividad humana por simple que ésta sea. La estadística se aplica en los programas de Gobierno, Ingeniería, Agronomía, Economía, Medicina, Biología, Psicología, Pedagogía, Sociología, Física, Astronomía, Educación, etcétera; no hay alguna ciencia que no la requiera o profesión que no la aplique. A continuación, se citan algunos ejemplos de la utilidad de la estadística:

1. En las agencias gubernamentales, tanto federales, estatales o municipales utilizan la estadística para realizar planes y programas para el futuro.
2. En el campo de la ingeniería se aplica en muchas de sus actividades tales como: La planeación de la producción, el control de calidad, las ventas.
3. En la Sociología se aplica para comparar el comportamiento de grupos socioeconómicos y culturales y en el estudio de su conducta.
4. En el campo económico su uso es fundamental para informar el desarrollo económico de una empresa o de un país que da a conocer los índices económicos relativos a la producción, a la mano de obra, índices de precios para el consumidor, las fluctuaciones del mercado bursátil, las tasas de interés, el índice de inflación, el costo de la vida, etcétera. Todos estos aspectos que se estudian, se reportan e informan, no solamente describen el estado actual de la economía, sino que trazan y predicen el camino de las futuras tendencias. Así mismo sirve a los encargados de las agencias, para tomar decisiones acertadas en sus operaciones.
5. En el campo demográfico la Estadística se aplica en los registros de los hechos de la vida diaria, tales como: nacimientos, defunciones, matrimonios, divorcios, adopciones. En materia de población los datos aportan una buena ayuda para fijar la política de estímulos al control de la natalidad, dirigir la inmigración o emigración, establecer los planes de lucha contra las enfermedades epidémicas o plagas que azotan los campos, etcétera.
6. En el campo educativo la Estadística contribuye al conocimiento de las condiciones fisiológicas, psicológicas y sociales de los alumnos y de los profesores. Al perfeccionamiento de los métodos de enseñanza, de evaluación, a la efectividad de programas de tutorías, la necesidad de reformas curriculares en función de los requerimientos sociales reales, etc.

7. En la industria la utilizan para el control de calidad, la implementación de incentivos a la producción, entre otros.
8. En la agricultura, se emplea en actividades como experimentos sobre la reproducción de plantas y animales entre otras cosas. También se usa la Estadística para determinar los efectos de clases de semillas, insecticidas y fertilizantes en el campo.
9. En la Biología se emplean métodos estadísticos para estudiar las reacciones de las plantas y los animales ante diferentes períodos ambientales y para investigar la herencia. Las leyes de Mendel sobre la herencia en donde los factores hereditarios se atribuyen a unidades llamadas genes y al estudio sistemático de los cruzamientos entre individuos portadores de genes diferentes, lo que ha permitido precisar de qué manera los genes se separan o se reúnen en las generaciones sucesivas. La verificación de las hipótesis formuladas por Mendel y sus continuadores necesitó el empleo de la Estadística.
10. En la medicina, los resultados que se obtienen sobre la efectividad de fármacos se analizan por medio de métodos estadísticos. Los médicos investigadores se ayudan del análisis estadístico para evaluar la efectividad de tratamientos aplicados. La Estadística también se aplica en el establecimiento y evaluación de los procedimientos de medida o clasificación de individuos con el propósito de establecer la especificidad y sensibilidad a las enfermedades.
11. En el Sector Salud, los técnicos de la salud la utilizan para planear la localización y el tamaño de los hospitales y de otras dependencias de sanidad. También se aplica en la investigación sobre las características de los habitantes de una localidad, sobre el diagnóstico y la posible fuente de un caso de enfermedad transmisible; sobre la proporción de personas enfermas en un momento determinado, de ciertos padecimientos de una localidad, sobre la proporción de enfermos de influenza en dos grupos, uno vacunado contra el padecimiento y el otro no. También se aplica en cualquier otro tipo de investigación similar a éste.
12. En la Psicología se aplican los conceptos y técnicas de la estadística para medir y comparar la conducta, las actitudes, la inteligencia y las aptitudes de las personas.
13. En los negocios se pueden predecir los volúmenes de venta, medir las reacciones de los consumidores ante los nuevos productos, probar la efectividad de una campaña publicitaria.
14. En la Física se utiliza la Estadística para obtener datos y probar hipótesis.
15. En el Deporte se ocupa para determinar el impacto de una nueva dieta alimenticia en el rendimiento de atletas o someter a prueba la efectividad de dos o más técnicas de ejercitación y práctica de un deporte.

ESQUEMA DE UNIDAD





CAPITULO

I

LOGROS:

INTERPRETA NOCIONES BÁSICAS RELACIONADAS CON EL MANEJO DE INFORMACIÓN COMO POBLACIÓN, MUESTRA, VARIABLE ALEATORIA, DISTRIBUCIÓN DE FRECUENCIAS, PARÁMETROS Y ESTADÍGRAFOS.

TABULA INFORMACIÓN ESTADÍSTICA PARA ANALIZARLA, GRAFICARLA Y DEDUCIR CONCLUSIONES.

CALCULA LAS MEDIDAS DE CENTRALIZACIÓN, POSICIÓN Y DISPERSIÓN EN UNA MUESTRA DE DATOS, REALIZANDO SU RESPECTIVO ANÁLISIS E INTERPRETACIÓN DE LA INFORMACIÓN.

INDICADORES DE LOGRO:

- Reconoce y aplica diferentes técnicas de recolección de datos.
- Identifica la población, muestra y variable en un estudio estadístico.
- Tabula y grafica información estadística.
- Analiza información estadística mediante su tabulación y representación gráfica.
- Calcula y analiza las medidas de centralización en datos estadísticos.
- Calcula y analiza las medidas de dispersión en datos estadísticos.
- Interpreta y deduce conclusiones de información estadística, partiendo de las medidas de centralización y dispersión.
- Aplica conceptos básicos de estadística en la solución de situaciones problema.
- Aplica conceptos de estadística en la formulación y elaboración de propuestas de investigación.

Capítulo 1: CONCEPTOS BÁSICOS DE ESTADÍSTICA

Definiciones de Estadística

Estadística: disciplina de las matemáticas que se refiere a los métodos de recolección, clasificación, presentación, de información para el análisis e interpretación de un conjunto de datos para la toma de decisiones.

Estadística: Ciencia que recoge, organiza, presenta, analiza e interpreta datos con el fin de propiciar una toma de decisiones más eficaz.

Estadística: es la rama de la matemática que se ocupa de recopilar datos, de organizarlos para una mejor comprensión del fenómeno que se desea estudiar y de analizarlos con un determinado objetivo.

Conceptos fundamentales de Estadística

Datos: concepto que proviene del latín “Datum” - lo que se dá – y que se refiere a un documento, información o testimonio que, después de procesado permite llegar a un conocimiento. El dato no tiene sentido en sí mismo, el sentido lo adquiere cuando se procesa y se interpreta a la luz de una teoría, un concepto o una hipótesis.

Datos crudos: datos no procesados, tal y como se recopilan en un estudio con un instrumento de recolección de información.

Tipos de datos

Los datos se pueden clasificar como: nominal, ordinal y numéricos.

Dato nominal: es una clasificación basada en categorías o nominaciones. Ejemplo: alto o mediano, sanos o enfermos, a favor o en contra.

Dato ordinal: son aquellos que el investigador ordena de más a menos o a la inversa, con el propósito de realizar comparaciones más equitativas, es decir, comparar el más alto de un grupo contra el más alto de otro grupo.

Dato numérico: son expresiones de valor, por ejemplo: el tiempo, la longitud, la distancia, el peso, entre otros.

CLASIFICACIÓN DE ESTADÍSTICA

Según el tipo de investigación

Estadística Descriptiva.

Tiene por objetivo fundamental la descripción numérica de un conjunto de datos. No generaliza las conclusiones obtenidas a otros grupos de datos.

Estadística Inferencial.

Usa la información aportada por una muestra para sacar conclusiones de la población de la cual ha sido extraída; siempre recordando que existe la probabilidad de hacerlo en forma errada

CONCEPTOS BÁSICOS:

POBLACIÓN: se denomina **población** al conjunto de individuos (personas, animales, cosas) sobre la cual se estudia una determinada característica. Ejemplo:

Datos que se recopilan	Población
Se registra el peso de los alumnos de un curso	Alumnos del curso
Se hace una encuesta en las viviendas de un barrio para determinar cuántas personas viven en cada una de ellas	Habitantes del barrio

El **tamaño de la población** es el número de individuos que la componen.

MUESTRA: Cuando el tamaño de una población es muy grande, se trabaja con una parte de ella llamada **muestra**.

Ejemplo:



A una consultora le encargan hacer un estudio acerca de cuál es la intención del voto de los ciudadanos de una ciudad en las próximas elecciones. Como no es posible encuestar a todos los ciudadanos, la consultora toma un grupo de 500 y sobre él analiza la variable. Con los datos recopilados sobre esta muestra, se puede hacer una proyección de los votos que obtendrá cada candidato.

Para que el estudio estadístico sea confiable, es muy importante que la selección de los individuos de una muestra resulte **representativa** de la población que se analiza, o sea, se deben obtener de ella resultados aproximadamente iguales a los que se hubieran obtenido considerando el total de la población.

La muestra se obtiene mediante un método de muestreo sobre el cual se realiza un estudio o comparación para luego generalizar los resultados a la población de la cual se extrajo la muestra. La muestra debe reunir las mismas características que el resto de la población y en consecuencia deberá representar a la misma de manera confiable.

Normalmente se recomienda que el tamaño de la muestra sea del 10% de la población definida, pero si es mayor es mejor, debido que incluye a más población y por tanto, asegura su representatividad.

MUESTREO Y CLASES DE MUESTREO

El muestreo es el procedimiento seguido para la selección de una muestra de una población determinada. Entre las clases de muestreo se destacan:

MUESTREO ALEATORIO O PROBABILÍSTICO

Este tipo de muestreo es uno de los más utilizados en las investigaciones. Se comprende que aquí todos los sujetos o elementos de la población pueden pasar a formar parte de la muestra, pues tienen la misma probabilidad de ser escogidos.

1. Muestreo aleatorio simple:

En este método de muestreo lo principal es tener una idea clara de cuantos sujetos serán necesarios para completar el tamaño de la muestra que se va a investigar, pues este consiste en que a cada sujeto de la población se le debe asignar un número, para que posteriormente mediante algún sorteo, o generando números aleatorios con ayuda de rifas o algún ordenador se logre dar a conocer los números de los sujetos seleccionados que serán tomados como muestra.

2. Muestreo sistemático:

Pues en este método como en el primero, los sujetos que conforman a la población son enumerados, pero en lugar de dar a conocer los que serán incluidos mediante el azar en este caso incluyen otra técnica. Posteriormente de que son enumerados, los investigadores dividen el total de la población que se presenta entre el total de sujetos que requieren para la muestra; para después elegir a uno de los primeros de estos sujetos al azar (ya enumerados) y posteriormente se va sumando de una forma secuencial o constante el mismo valor para elegir a los sujetos requeridos.

3. Muestreo estratificado:

Mediante este método de investigación, los investigadores dividen a la población en grupos o estratos que tengan relación o compartan características similares y posteriormente se selecciona al azar o aleatoriamente a los sujetos finales de los grupos o estratos formados. Este método es utilizado para cuando los investigadores pretenden que dentro de la muestra se encuentren incluidos todos los grupos de interés que puedan representar significancia para la investigación.

La distribución de la muestra tomada de la población es llamada “Fijación” y se pueden conocer los tres que se muestran a continuación:

1. **A fijación simple:** Donde cada grupo formado cuenta con la misma cantidad de número de elementos.
2. **A fijación proporcional:** Donde las distribuciones se hacen de acuerdo al tamaño de la población o sujetos de cada estrato.
3. **Afijación óptima:** Fijando la cantidad de sujetos que pueden ser admitidos en la muestra.
4. Muestreo por Conglomerados:

En este método, la población ya se encuentra dividida en grupos o estratos formados naturalmente y a partir de estos se toman la cantidad de sujetos que sean necesarios de cada uno para así formar la muestra. Este es similar al método por estratificación, sin embargo, se facilita más, pues los grupos ya están formados y solo se requieren los datos de esos elementos que lo integran.

II. MUESTREO NO ALEATORIO O NO PROBABILÍSTICOS

Mediante esta técnica de muestreo, en comparación del muestreo probabilístico las muestras no se recogen por medio de procesos donde a los sujetos se les brinden las mismas posibilidades de ser seleccionados.

Dentro de esta técnica existen cuatro tipos de muestreos:

1. Muestreo por cuotas:

Mediante este método los investigadores se encargan de incluir en la muestra solo a un grupo determinado de sujetos que cumplen con ciertos requisitos o condiciones específicas.

2. Muestreo intencional o de conveniencia:

Mediante este método, el investigador se encarga de elegir de acuerdo a sus propios criterios o alcances a los sujetos que formaran parte de la muestra.

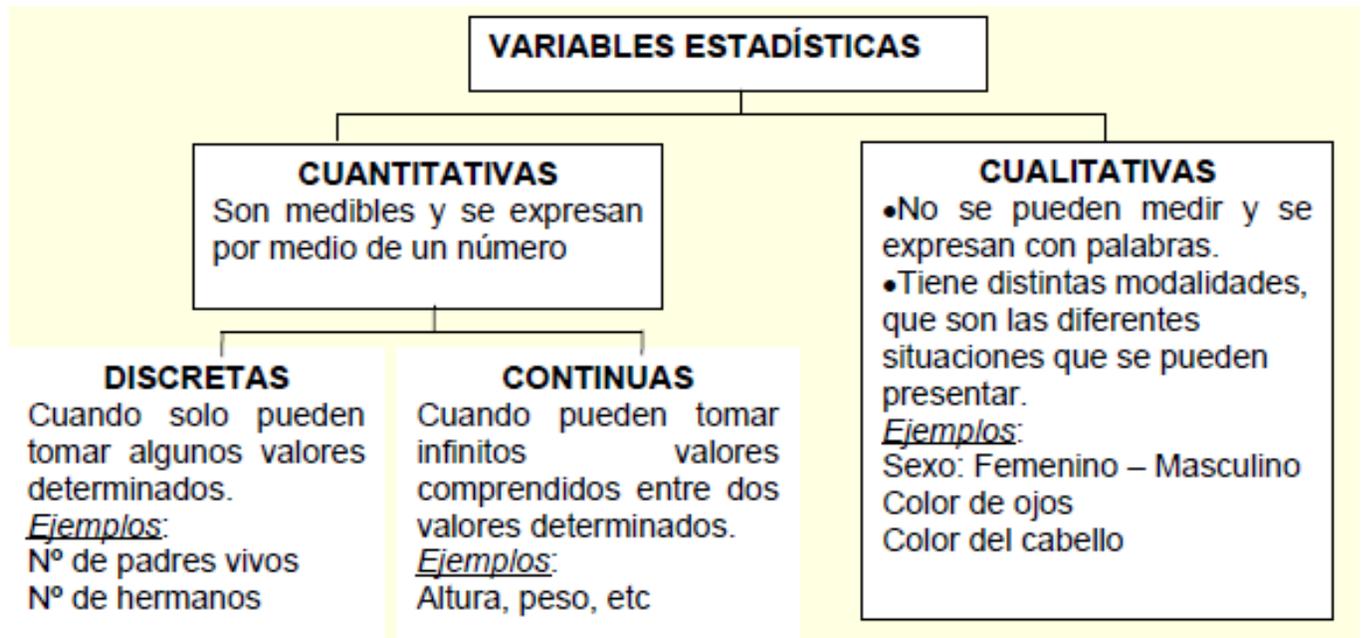
3. Muestreo por bola de nieve:

En este método se hace contacto con algún sujeto en específico y este se encargará de buscar a otros sujetos o nuevos participantes consecutivamente. Haciéndolo de esta forma hasta cumplir con el tamaño de muestra o cantidad de sujetos que se necesitan.

4. Muestreo discrecional:

Mediante este método, se aplica una técnica parecida a la del método intencional, pues el investigador recluta a los sujetos según sus criterios; pero en este caso, los busca de acuerdo a lo que piensen que estos sujetos puedan aportar a la investigación.

Variables: son los caracteres o cualidades de la población que es objeto de estudio o análisis. Pueden ser:



DATOS

Puede concebirse como información numérica necesaria para ayudarnos a tomar una decisión con más bases en una situación en particular.

TÉCNICAS DE RECOLECCIÓN DE DATOS

La recolección de datos se refiere al uso de una gran diversidad de técnicas y herramientas que pueden ser utilizadas por el analista para desarrollar los sistemas de información, los cuales pueden ser la entrevistas, la encuesta, el cuestionario, la observación, el diagrama de flujo y el diccionario de datos. Todos estos instrumentos se aplicarán en un momento en particular, con la finalidad de buscar información que será útil a una investigación en común. En la presente investigación trata con detalle los pasos que se debe seguir en el proceso de recolección de datos, con las técnicas ya antes nombradas. Las 5 principales técnicas de recolección de datos son:



ENTREVISTA: La entrevista es una conversación dirigida, con un propósito específico y que usa un formato de preguntas y respuestas. Se establece así un diálogo, pero un diálogo peculiar, asimétrico, donde una de las partes busca recoger informaciones y la otra se nos presenta como fuente de estas informaciones. Una entrevista es un dialogo en el que la persona (entrevistador), generalmente un periodista hace una serie de preguntas a otra persona (entrevistado), con el fin de conocer mejor sus ideas, sus sentimientos su forma de actuar.

Preparación de la Entrevista:

1. Determinar la posición que ocupa de la organización el futuro entrevistado, sus responsabilidades básicas, actividades, etc.
2. Preparar las preguntas que van a plantearse, y los documentos necesarios
3. Fijar un límite de tiempo y preparar la agenda para la entrevista.
4. Elegir un lugar donde se puede conducir la entrevista con la mayor comodidad.
5. Hacer la cita con la debida anticipación.

Tipo de preguntas:

***Preguntas abiertas:** Son aquellas preguntas que describen hechos o situaciones por parte del entrevistado con una gran cantidad de detalles que a juicio del entrevistado son importantes.

***Preguntas cerradas:** En las preguntas cerradas las respuestas posibles están cerradas al entrevistado, debido a que solamente puede responder con un número finito, tal como “ninguno”, “uno”, o “quince”. Una pregunta cerrada limita las respuestas disponibles al entrevistado.



OBSERVACIÓN: La observación es otra técnica útil para el analista en su proceso de investigación, consiste en observar a las personas cuando efectúan su trabajo, durante la cual el analista participa activamente, actúa como espectador de las actividades llevadas a cabo por una persona para conocer mejor su sistema. El propósito de la observación es múltiple, permite al analista determinar que se está haciendo, como se está haciendo, quien lo hace, cuando se lleva a cabo, cuánto tiempo toma, donde se hace y porque se hace.

Tipos de observación:

El analista puede observar de tres maneras básicas:

- Puede observar a una persona o actividad sin que el observado se dé cuenta y sin interactuar por parte del propio analista.
- El analista puede observar una operación sin intervenir para nada, pero estando la persona observada enteramente consiente de la observación.
- Se puede observar y estar en contacto con las personas observadas.

La interrogación puede consistir simplemente en preguntar respecto a una actividad específica, pedir una explicación, etc. La observación puede emplearse para verificar los

resultados de una entrevista, o bien como preparación de la misma. También es otra técnica valiosa para recopilar datos que implican relaciones. La observación tiende a adquirir mayor sentido al nivel técnico del procesamiento de datos, donde las tareas se cuantifican más fácilmente. Entre estas tareas encontramos la recopilación, acumulación y transformación de los datos.

Pasos de la observación:

1. Determinar y definir aquello que se va a observar.
2. Estimar el tiempo necesario de observación.
3. Obtener la autorización para llevar a cabo la observación.
4. Explicar a las personas que van a ser observadas lo que se va hacer y las razones para ello.



LA ENCUESTA: Una encuesta es un conjunto de preguntas normalizadas dirigidas a una muestra representativa de la población o instituciones, con el fin de conocer estados de opinión o hechos específicos. La intención de la encuesta no es describir los individuos particulares quienes, por azar, son parte de la muestra sino obtener un perfil compuesto de la población. Una "encuesta" recoge información de una "muestra.". La encuesta es un procedimiento a través del cual se recopila información. Forma parte del diseño de una investigación descriptiva.



CUESTIONARIO: Los cuestionarios proporcionan una alternativa muy útil para la entrevista; sin embargo, existen ciertas características que pueden ser apropiada en algunas situaciones e inapropiadas en otra. Al igual que la entrevistas, deben diseñarse cuidadosamente para una máxima efectividad. Un cuestionario es un instrumento de investigación con una serie de preguntas cuidadosamente organizadas con el fin de realizar un análisis estadístico de las respuestas.

Selección de formas para cuestionarios

Existen dos formas de cuestionarios para recabar datos: cuestionarios abiertos y cerrados, y se aplican dependiendo de si los analistas conocen de antemano todas las posibles respuestas de las preguntas y pueden incluirlas. Con frecuencia se utilizan ambas formas en los estudios de sistemas.

Cuestionario Abierto: Al igual que las entrevistas, los cuestionarios pueden ser abiertos y se aplican cuando se quieren conocer los sentimientos, opiniones y experiencias generales; también son útiles al explorar el problema básico, por ejemplo, un analista que utiliza cuestionarios para estudiar los métodos de verificación de crédito, es un medio. El formato

abierto proporciona una amplia oportunidad para quienes respondan escriba las razones de sus ideas. Algunas personas, sin embargo, encuentran más fácil escoger una de un conjunto de respuestas preparadas que pensar por sí mismas.

Cuestionario Cerrado: El cuestionario cerrado limita las respuestas posibles del interrogado. Por medio de un cuidadoso estilo en la pregunta, el analista puede controlar el marco de referencia. Este formato es el método para obtener información sobre los hechos. También fuerza a los individuos para que tomen una posición y forma su opinión sobre los aspectos importantes.



ANÁLISIS DOCUMENTAL: Su instrumento es la ficha de registro. Consiste en recabar información relacionada al tema de investigación de todas las fuentes disponibles: tesis, revistas, páginas web, libros, historias clínicas, expedientes judiciales, registro de ventas, historial de notas, documentales, etc. Estas fuentes deben ser válidas para citarse.

CLASES DE PREGUNTAS:

Preguntas abiertas: no se brindan opciones de respuesta. Ejemplo: ¿cuál es tu escritor favorito?, ¿cuál es tu equipo de fútbol favorito?, ¿qué opina del acoso sexual en lugares de trabajo?

Preguntas cerradas: El encuestador brinda las opciones de respuesta. Este tipo de encuestas facilitan el trabajo de obtención de datos o como se conoce comúnmente tabulación. Para nuestro estudio sólo utilizaremos preguntas cerradas.

CONTEO Y ORGANIZACIÓN DE DATOS

Los datos de un estudio estadístico se pueden organizar mediante tablas en las que se presente la variable que se analiza con los diferentes valores que puede tomar y se van registrando las respuestas de los encuestados. El número de veces que aparece una misma respuesta se conoce como **frecuencia absoluta**.

FRECUENCIAS

Frecuencia absoluta: es el número de veces que se repite una observación o valor de la variable (f).

ACTIVIDAD PRÁCTICA:

Encuesta No.1: Entretenimiento (cine)

Instrucción: en la siguiente encuesta seleccionar sólo una de las respuestas.

1. Sexo:

- a. Masculino ____
- b. Femenino ____

2. ¿Te gusta el cine?

- a. Sí ____
- b. No ____

3. ¿Cuál es el lugar de preferencia para observar algún tipo de película?

- a. Sala de cine ____
- b. Casa ____
- c. Otro sitio ____

4. Acostumbras observar películas con tu familia:

- a. Sí ____
- b. No ____
- c. Algunas veces ____

5.Cuál es tu género de películas favorito

- a. Acción ____
- b. Drama ____
- c. Comedia ____
- d. Infantil ____
- e. Suspenso ____
- f. Ciencia ficción ____
- g. Otro ____

TABLA PARA EL REGISTRO DE LA INFORMACIÓN:

PREGUNTA	CONTEO	FCIA. ABSOLUTA
1. Sexo		
a. Masculino		
b. Femenino		
2. Te gusta el cine		
a. Sí		
b. No		
3. ¿Cuál es el lugar de preferencia para observar algún tipo de película?		
a. Sala de cine		
b. Casa		
c. Otro		
4. Acostumbras observar películas con tu familia		
a. Sí		
b. No		
c. Algunas veces		
5.Cuál es tu género de películas favorito		
a. Acción		
b. Drama		
c. Comedia		
d. Infantil		
e. Suspenso		
f. Ciencia ficción		
g. Otro		

En la tabla anterior se puede observar que la cantidad de veces que se repitió una respuesta constituye su frecuencia absoluta.

Frecuencia relativa de una observación: es el cociente entre su frecuencia absoluta y el número total de observaciones realizadas (F_r)

$$F_r = \frac{F}{n}$$

La tabla que muestra la frecuencia relativa, absoluta y acumulada se llama **distribución de frecuencias**.

- La suma de las frecuencias absolutas es el total de observaciones.
- La suma de las frecuencias relativas es siempre 1.
- Si multiplicamos por 100 cada frecuencia relativa, obtenemos el **porcentaje** de cada valor de la variable (%).

Porcentaje = frecuencia relativa . 100

$$\% = F_r \times 100$$

Ejemplo de tabla de frecuencias:

Se ha encuestado a un grupo de personas, queriendo conocer el número de habitantes en su vivienda.

DATOS X_i	Frecuencia absoluta f_i	Frecuencia absoluta acumulada F_i	Frecuencia relativa		Frecuencia relativa acumulada	
			Decimal $h_i = \frac{f_i}{N}$	Porcentual $\%_i = 100 \times h_i$	Decimal H_i	Porcentual $\%A_i$
5	6	6	0,0150	1,5 %	0,0150	1,5 %
6	48	54	0,1200	12 %	0,1350	13,5 %
7	95	149	0,2375	23,75 %	0,3725	37,25 %
8	105	254	0,2625	26,25 %	0,6350	63,50 %
9	87	341	0,2175	21,75 %	0,8525	85,25 %
10	59	400	0,1475	14,75 %	1	100 %
	N = 400		1	100 %		



Actividad de profundización 1:

1. A un grupo de alumnos se le aplica una evaluación escrita, en la tabla se describe las notas Obtenidas:

Notas	F	Fr	%
1	2		
2	4		
3	3		
4	5		
5	6		
6	8		
7	5		
8	3		
9	2		
10	2		
Total			

- Completar la tabla.
- ¿Cuántos alumnos hay en el curso?
- ¿Cuántos aprobaron?
- ¿Cuántos reprobaron?

2. Conformar grupos de cuatro estudiantes.

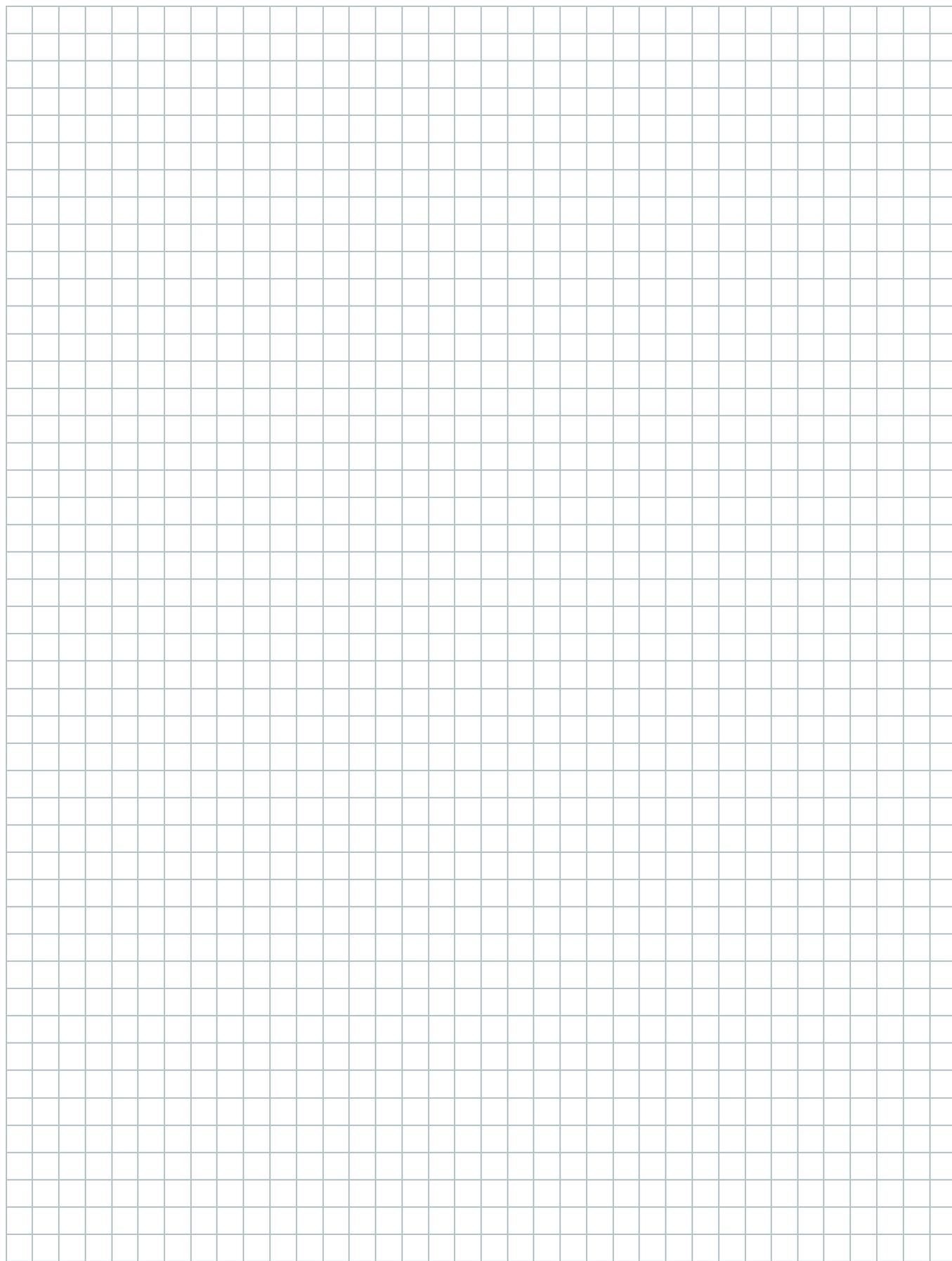
Seleccionar una muestra de veinte (20) personas, aplicando uno de los métodos de muestreo relacionados en el texto.

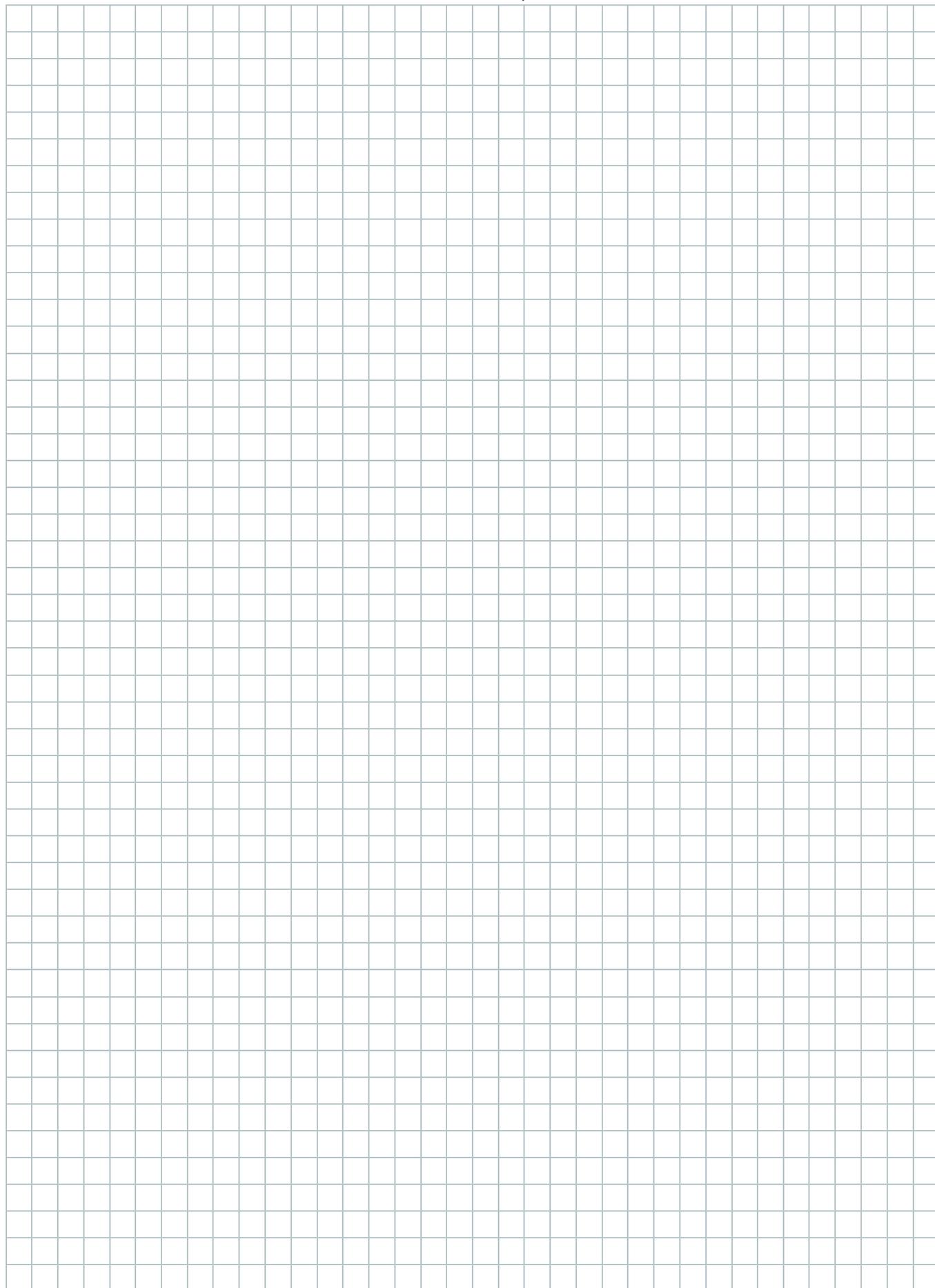
Aplicar encuesta sobre los temas “Estilos de vida saludable”, “hábitos de estudio”, “prácticas deportivas”, “uso de internet”, para lo cual deberán sacar las fotocopias respectivas.

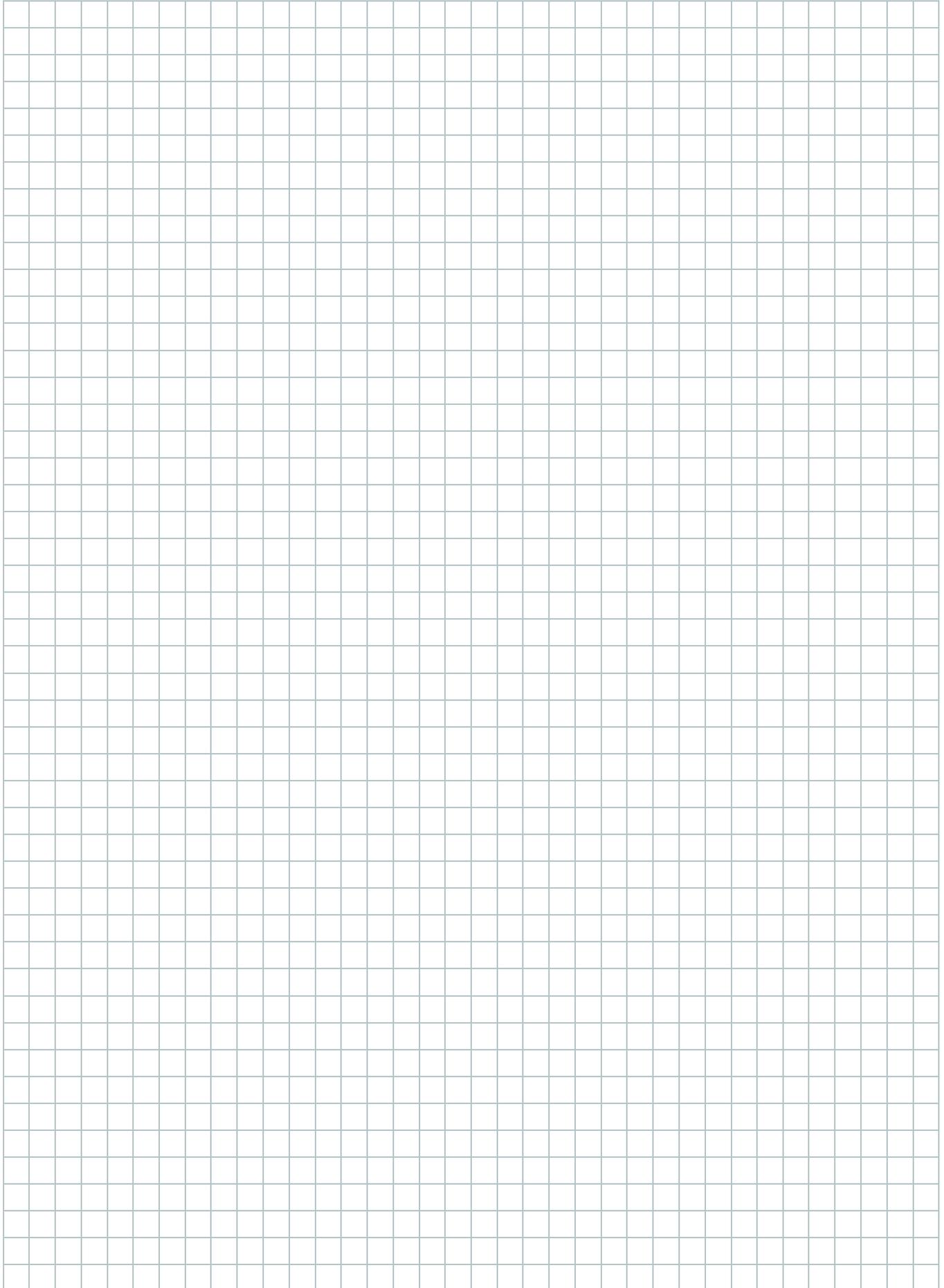
Después de aplicar las encuestas, estas serán numeradas de 1 a 20.

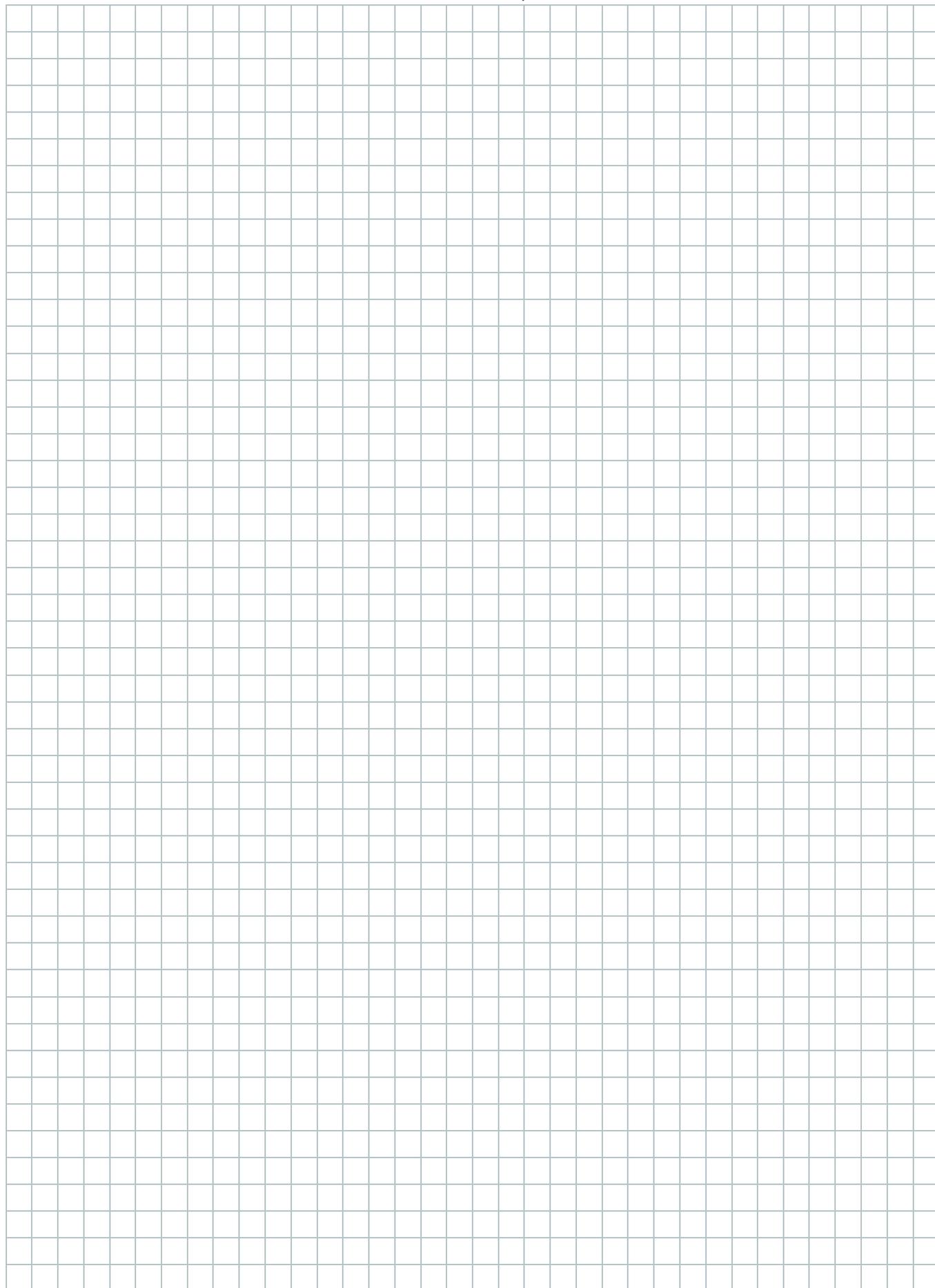
Con los datos de cada una de las encuestas realizar tablas de tabulación y frecuencias.

Con la información obtenida sacar conclusiones y exponerlas al grupo.









ENCUESTA No.2

Tema: Estilos de vida saludable

Instrucciones: en las siguientes preguntas escoger sólo una opción de respuesta y marcarla con una x

1. Sexo
 - a. Masculino _____
 - b. Femenino _____

2. Hace actividad física con regularidad (caminar, trotar, practicar algún deporte, asistir a gimnasio, etc)
 - a. Siempre _____
 - b. Frecuentemente _____
 - c. Algunas veces _____
 - d. Nunca _____

3. Realiza ejercicios que le ayudan al funcionamiento cardíaco (spinning, aeróbicos, rumba aeróbicos, entre otros).
 - a. Siempre _____
 - b. Frecuentemente _____
 - c. Algunas veces _____
 - d. Nunca _____

4. Practica ejercicios que le ayudan a estar tranquilo(a) (yoga, taichí, danza, meditación, auto relajación, entre otros).
 - a. Siempre _____
 - b. Frecuentemente _____
 - c. Algunas veces _____
 - d. Nunca _____

5. Incluye momentos de descanso en su rutina diaria:
 - a. Siempre _____
 - b. Frecuentemente _____
 - c. Algunas veces _____
 - d. Nunca _____

6. Consume de 4 a ocho vasos de agua diariamente:
 - a. Siempre _____
 - b. Frecuentemente _____
 - c. Algunas veces _____
 - d. Nunca _____

7. Duerme regularmente entre 7 y 8 horas diarias:
 - a. Siempre _____
 - b. Frecuentemente _____
 - c. Algunas veces _____
 - d. Nunca _____

ENCUESTA No.3**Tema:** Hábitos de estudio**Instrucciones:** en las siguientes preguntas escoger sólo una opción de respuesta y marcarla con una x

1. Sexo:

- a. Masculino _____
- b. Femenino _____

2. Grado:

- a. Sexto: _____
- b. Séptimo _____
- c. Octavo _____
- d. Noveno _____
- e. Décimo _____
- f. Undécimo _____

3. Tienes un lugar fijo para estudiar:

- a. Sí _____
- b. No _____

4. Tu lugar de estudio es suficientemente amplio y adecuado para estudiar:

- a. Sí _____
- b. No _____

5. Buscas ampliar el conocimiento de manera extraclase:

- a. Sí _____
- b. No _____

6. Cuánto tiempo dedicas semanalmente de manera extraclase a reforzar tu aprendizaje:

- a. No dedicas tiempo: _____
- b. Menos de una hora _____
- c. Entre 1 y 3 horas _____
- d. Más de 3 horas _____

7. Recibes acompañamiento de tus familiares para la realización de tareas escolares:

- a. Nunca _____
- b. Ocasionalmente _____
- c. Siempre _____

8. Cuando tienes problemas afectivos, estos inciden en tu rendimiento académico:

- a. Sí _____
- b. No _____
- c. Algunas veces _____

ENCUESTA No.4**Tema:** Prácticas deportivas**Instrucciones:** en las siguientes preguntas escoger sólo una opción de respuesta y marcarla con una x

1. Sexo:

- a. Masculino _____
- b. Femenino _____

2. Con qué frecuencia hace ejercicio o practica algún tipo de deporte:

- a. Una vez por semana _____
- b. Dos veces por semana _____
- c. Tres veces por semana _____
- d. Sólo en vacaciones _____

3. Cotidianamente practica deportes de tipo:

- a. Individual _____
- b. Colectivo _____

4.Cuál es la principal razón por la cual haces algún tipo de ejercicio:

- a. Por salud _____
- b. Por estar en forma _____
- c. Por diversión y pasar el tiempo _____
- d. Para disminuir el estrés _____
- e. Porque está de moda _____
- f. Para relacionarse con otras personas _____

5. La práctica de la actividad física acostumbras a realizarla con:

- a. Solo _____
- b. Familiares _____
- c. Amigos _____
- d. Mascota _____
- e. Conocidos _____

6. En qué tipo de instalaciones acostumbras a realizar la actividad física:

- a. Instalación pública _____
- b. Gimnasio _____
- c. Lugares abiertos (parques, calles, etc)
- d. Casa _____
- e. Otros _____

7. Qué carácter competitivo suele tener el ejercicio y/o deporte que practicas:

- a. Ligas o competiciones deportivas _____
- b. Campeonatos y torneos de fin de semana _____
- c. Compite sólo con amigos _____
- d. No tiene carácter de competencia _____

8. Qué importancia tiene para usted la ropa y el material deportivo:

- a. Ninguna _____
- b. Poca _____
- c. Bastante _____

ENCUESTA No.5**Tema:** Uso de internet**Instrucciones:** en las siguientes preguntas escoger sólo una opción de respuesta y marcarla con una x

1. Sexo:

- a. Masculino _____
b. Femenino _____

2. Te conectas diariamente a internet:

- a. Sí _____
b. No _____

3. Te conectas más de 10 horas a la semana:

- a. Sí _____
b. No _____

4. El horario en que con más frecuencia utilizas internet es:

- a. Mañana _____
b. Tarde _____
c. Noche _____

5. Te conectas a internet principalmente por motivos de:

- a. Académicos _____
b. Laborales _____
c. Ocio _____
d. Otros _____

6. Has dejado de hacer actividades que antes te gustaban por estar conectado a internet:

- a. Sí _____
b. No _____

7. Te pones de mal humor si no puedes conectarte:

- a. Sí _____
b. No _____

8. Si no estás conectado, te preocupa lo que está pasando en la red:

- a. Sí _____
b. No _____

9. Acostumbras a conseguir amistades por las redes sociales:

- a. Sí _____
b. No _____

10. Compartes información confidencial a través de internet:

- a. Sí _____
b. No _____
c. Algunas veces _____

11. Has notado que el uso de internet te está generando dependencia:

- a. Sí _____
b. No _____

12. Has disminuido el tiempo que dedicas a tu familia por estar conectado a internet:

- a. Sí _____
b. No _____



AUTOEVALUACIÓN No.1

Los alumnos de grado undécimo decidieron aplicar una encuesta entre los estudiantes de bachillerato, con el fin de conocer la opinión sobre las actividades culturales y deportivas que quisieran que se desarrollen en la institución. Para ello se seleccionó una muestra de 80 estudiantes de manera deliberada. Las preguntas base de la encuesta son:

- ☆ ¿Qué actividades culturales quisiera que se realizaran en el colegio?
- ☆ ¿Qué actividades deportivas desea que se realice para fomentar la actividad física en los estudiantes?

1. 20 de los 80 estudiantes expresaron su deseo de que en el colegio se realizaran “campeonatos”. La frecuencia absoluta de este dato es:

- a. 10 b. 20/80 c. 1/4 d. 20

2. Si de los 80 estudiantes 16 respondieron “aeróbicos” ¿cuál es el porcentaje que prefiere esta actividad:

- a. 8% b. 20% C. 10% d. 16%

3. Las respuestas a la primera pregunta fueron organizadas en la siguiente tabla. Calcular la frecuencia relativa (fr) y frecuencia porcentual (f%).

Actividades Culturales	Frecuencia absoluta (fi)	Frecuencia relativa (fr)	Frecuencia Porcentual f%
Noches de cine	35		
Festival de la canción	25		
Festival de poesía	20		
Total	80		

4. En cuanto a la pregunta sobre ¿Qué actividades deportivas desea que se realice para fomentar la actividad física en los estudiantes? Una de las respuestas fue “caminatas semanales”, con una frecuencia relativa de $\frac{7}{80}$, eso significa:

- a. 7 personas dieron esa respuesta.
- b. 9 personas dieron esa respuesta.
- c. 0,8 personas respondieron.
- d. 10 personas respondieron.

5. Las respuestas a la segunda pregunta están registradas en la siguiente tabla:

DEPORTE	ESTUDIANTES
Torneos bimestrales	30
Aeróbicos semanales	18
Ciclovías	8
Caminatas semanales	7
Sesiones de rumba aeróbicos	14
Sesiones de porrismo	3

Se pide elaborar la tabla de distribución de frecuencias:

Evento	Frecuencia absoluta	Frecuencia relativa		
		fracción	decimal	fr%
Torneos bimestrales				
Aeróbicos semanales				
Ciclovías				
Caminatas semanales				
Sesiones de rumba aeróbicos				
Sesiones de porrismo				
TOTAL				

TABLAS DE FRECUENCIA AGRUPADAS EN INTERVALOS

Para agrupar datos en intervalos se puede realizar de diversas maneras, de acuerdo a la conveniencia del investigador, sin embargo, estadísticamente se acostumbra a utilizar LA REGLA DE STURGES.

Regla de Sturges. En estadística descriptiva la **Regla de Sturges** es un criterio muy utilizado cuando se quiere realizar un histograma de frecuencias ya que con esta **regla** se calcula el número de clases (o intervalos) necesarios para representar fielmente los datos.

Procedimiento para agrupar los datos:

- a) Se ordenan los datos en forma creciente o decreciente (se acostumbra con mayor frecuencia hacerlo de manera creciente, es decir, de menor a mayor).
- b) Se determina el número mayor y el menor.
- c) Se calcula el rango, que es la diferencia entre el número mayor y el número menor.
- d) Se determina el número de intervalos que es igual $1 + 3,322 \log(n)$, donde n es el tamaño de la muestra. El número de intervalos debe ser un número entero, por lo tanto, se debe aproximar.
- e) Se calcula la amplitud del intervalo, que es igual a:

$$A = \frac{\text{Rango}}{\text{Número de intervalos}}$$

- f) El número se aproxima al entero próximo, porque el cálculo se está realizando con números enteros.
- g) Se determina el rango ampliado, que es igual al producto de la Amplitud * Número de intervalos.
- h) Se determina K , que es la diferencia del rango ampliado y del rango original. Si k es igual a 1, este valor se resta al valor mínimo. Si k es un número par, se divide por 2, uno de los números se resta al mínimo y el otro se suma al máximo. Si k es un número impar mayor que 1, se buscan dos números que al ser sumados se obtenga K y al restarlos el resultado sea 1.
El número menor se le resta al mínimo y el número mayor se le suma al máximo
- i) El primer número del intervalo se conoce como límite inferior y el segundo número se conoce como límite superior.
- j) La marca de clase es el promedio entre el límite inferior y el límite superior.
- k) Para hallar la frecuencia absoluta en el primer intervalo no se acostumbra a contar el límite superior, ya que se tendrá en cuenta en el siguiente intervalo.

EJEMPLO DE REGLA DE STURGES:

Se seleccionó una muestra de 30 estudiantes de un curso de Estadística, con el fin de conocer su peso en kilos.

74	67	94	70	69	61	71	79	47	85	82	55	65
88	52	58	76	57	72	66	48	56	63	71	60	64
68	83	74	92	(Martínez, 2013).								

A partir de los datos presentados construir la tabla de distribución de frecuencias agrupadas en intervalos.

Se determinan los intervalos

Para identificar la clase, se debe desarrollar los siguientes pasos:

a) Ordenar los datos de forma ascendente.

47	48	52	55	56	57	58	60	61	63	64	65	66
67	68	69	70	71	71	72	74	74	76	79	82	83
85	88	92	94									

b) Se determina el número mayor y el menor. Para este ejemplo el número mayor es 94 y el menor es 47.

c) Se calcula el rango, que es la diferencia entre el número mayor y el menor $94 - 47 = 47$.

d) Se determina el número de intervalos que es igual $1 + 3,322 \log(n)$, donde n es el tamaño de la muestra, para este ejemplo $n = 30$; número de intervalos = $1 + 3,3 * \log(30) = 5,907$, este valor se aproxima a 6.

Nota: el número de intervalos debe ser un número entero.

e) Se calcula la amplitud del intervalo, que es igual a:

$$A = \frac{\text{Rango}}{\text{Número de intervalos}} = \frac{47}{6} = 7,83333 \approx 8$$

Nota: El número se aproxima al entero próximo, porque el cálculo se está realizando con números enteros.

f) Se determina el rango ampliado, que es igual al producto de la Amplitud * Número de intervalos = $8 * 6 = 48$

g) Se determina K, que es la diferencia del rango ampliado y del rango original, $48 - 47 = 1$, este valor se resta al valor mínimo, $47 - 1 = 46$, ahora se inicia la construcción de los intervalos iniciando en el valor mínimo, con una amplitud de 8.

h) se inicia la construcción de los intervalos iniciando en el valor mínimo, con una amplitud de 8. En este caso construiremos una tabla con marca de clase y frecuencia absoluta.

Peso en kilos estudiantes	Marca de clase	Frecuencia absoluta
[46 – 54)	50	3
[54 – 62)	58	6
[62 – 70)	66	8
[70 – 78)	74	6
[78 – 86)	82	4
[86 – 94)	90	3
TOTAL		30

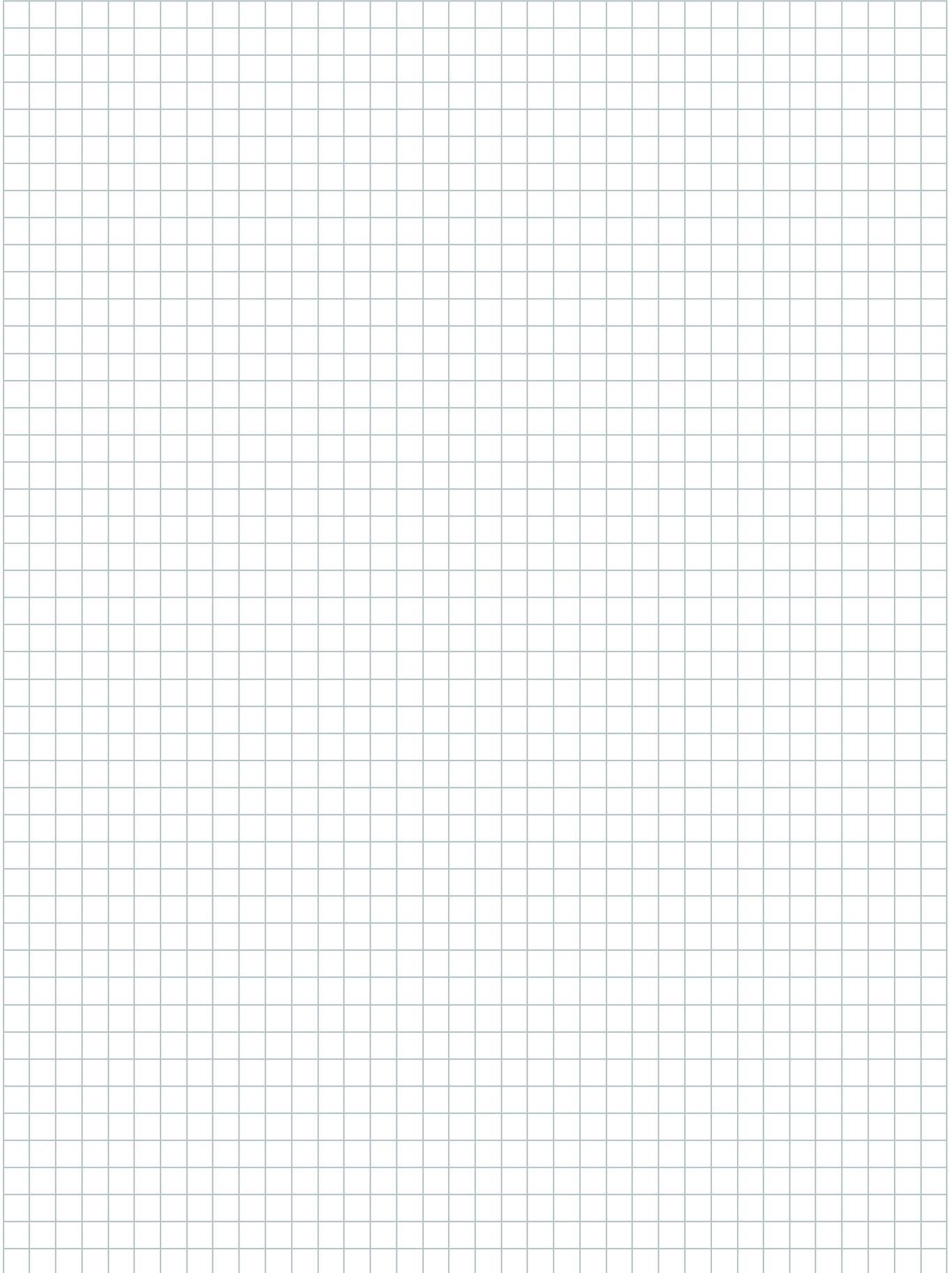
En los intervalos el corchete significa valor cerrado y el paréntesis valor abierto, esto significa, que si el número está con corchete fue incluido en ese intervalo y si está con paréntesis significa que no fue incluido en ese intervalo si no en el siguiente.



Actividad de profundización 2:

Dado el siguiente conjunto de datos: 105, 106, 105, 107, 109, 111, 110, 110, 107, 107, 104, 99, 103, 99, 103, 101, 100, 101, 100, 103, 98, 92, 97, 94, 95, 95, 93, 95, 95, 95, 91, 82, 91, 85, 90, 86, 87, 89, 87, 89

- Agrupar en intervalos haciendo uso de la regla de Sturges.
- Elaborar tabla de frecuencias con marca de clase, frecuencia absoluta, frecuencia acumulada, frecuencia relativa, porcentaje.



GRÁFICAS ESTADÍSTICAS

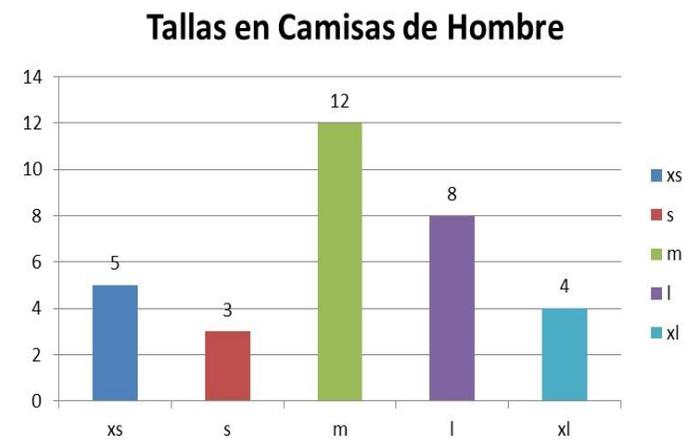
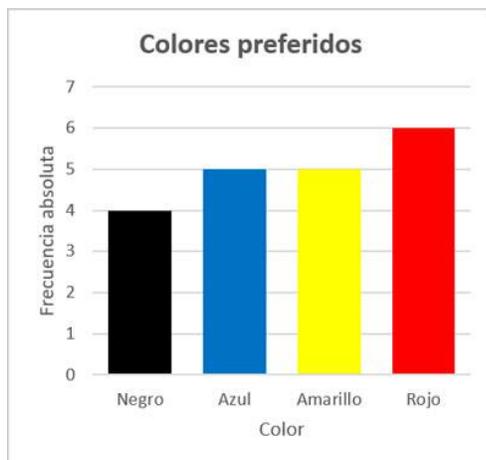
Las gráficas estadísticas permiten visualizar la información contenida en las tablas de manera rápida y sencilla. Existen muchos tipos de gráficas estadísticas, unas se emplean con variables cuantitativas y otras con variables cualitativas.

Se emplean para tener una representación visual de la totalidad de la información. Los gráficos estadísticos presentan los datos en forma de dibujo de tal modo que se pueda percibir fácilmente los hechos esenciales.

Existen varios tipos de gráficas estadísticas, sin embargo, en este texto se relacionaran algunos de los más comunes.

Diagrama de barras:

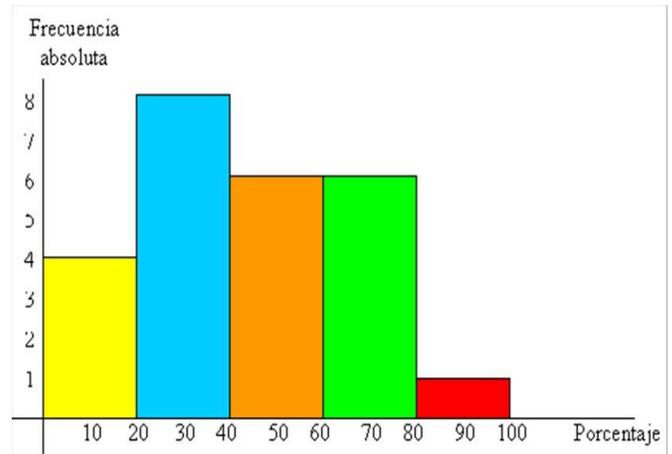
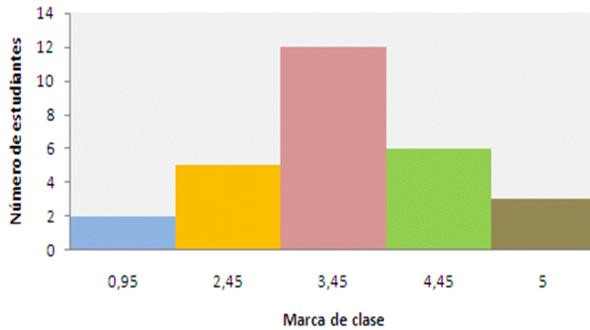
Se utiliza para representar datos cualitativos o datos cuantitativos de tipo discreto. Se representan sobre unos ejes de coordenadas, en el eje de abscisas ("X") se colocan los valores de la variable, y sobre el eje de ordenadas ("Y") las frecuencias absolutas, relativas, porcentajes o frecuencias acumuladas. Los datos se representan mediante barras de una altura proporcional a la frecuencia. Ej:



Histograma:

Se utiliza para de presentar datos cuantitativos de tipo continuo. Se representan sobre unos ejes de coordenadas, en el eje de abscisas ("X") se colocan los intervalos de los valores de la variable, y sobre el eje de ordenadas ("Y") las frecuencias absolutas, relativas, porcentajes o frecuencias acumuladas. Los datos se representan mediante barras pegadas unas a otras de una altura proporcional a la frecuencia. Ej:

Título: Histograma
Calificación obtenida en la prueba



Polígono de frecuencia:

Es un gráfico lineal, su construcción es similar al histograma; para su construcción se unen los puntos medios de cada clase, con sus respectivas frecuencias; de tal manera que al unir sus puntos medios por segmentos forman un polígono. Ej:

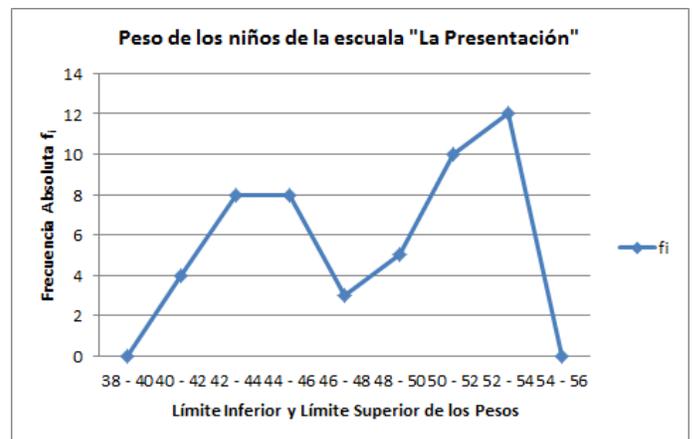
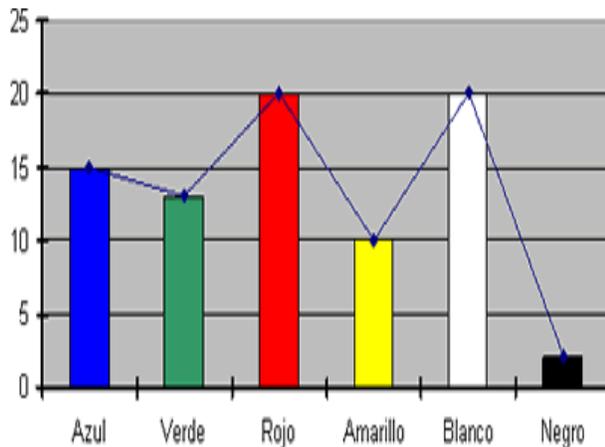
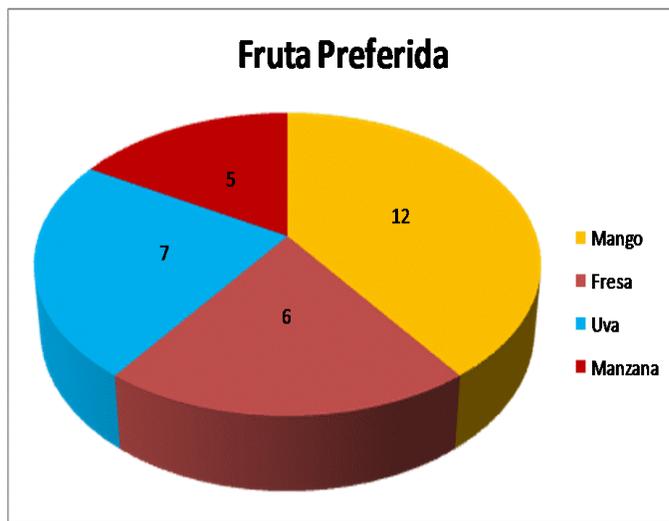


Diagrama de sectores o diagrama circular:

En un diagrama de este tipo, los 360° de un círculo se reparten proporcionalmente a las frecuencias de los distintos valores de la variable. Resultan muy adecuados cuando hay pocos valores, o bien cuando el carácter que se estudia es cualitativo. Ej:



Pictograma:

Los pictogramas representan los datos mediante un dibujo ilustrativo relacionado con el tema tratado. Los pictogramas se utilizan buscando hacer más amigable y entendible los informes estadísticos. Ej:

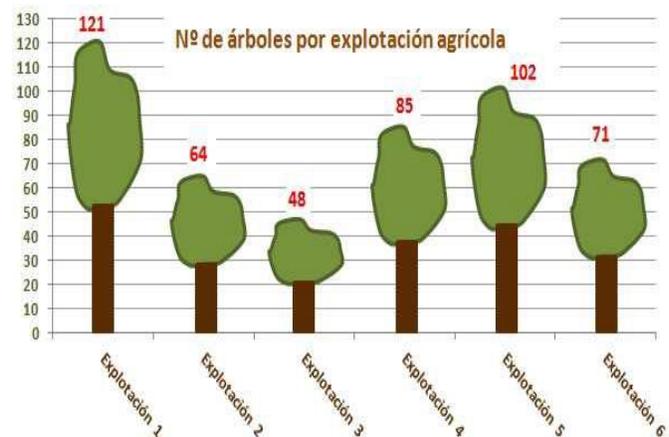


Diagrama de tallos y hojas:

El **diagrama "tallo y hojas"** permite obtener simultáneamente una distribución de frecuencias de la variable y su representación gráfica. Para construirlo basta separar en cada dato el último dígito de la derecha (que constituye la **hoja**) del bloque de cifras restantes (que formará el **tallo**). Esta representación de los datos es semejante a la de un histograma, pero además de ser fáciles de elaborar, presentan más información que estos.



Es especialmente útil para conjuntos de datos de tamaño medio (entre 20 y 50 elementos) y que sus datos no se agrupan alrededor de un único tallo.

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

El nombre de tallo y hojas hace referencia a la ramificación de una planta, siendo los dígitos delanteros los que marcan el tallo donde se encuentra el número y el dígito final la hoja.

Construcción del diagrama de tallo y hojas

Para construir el diagrama de tallo y hojas, debemos seguir los siguientes pasos:

1. Ordenar los datos.
2. Redondear los números (en el caso de que no lo estén) hasta tengan las cifras que queramos. Por ejemplo, si tenemos el número 3,62856 y queremos que tenga 2 dígitos la parte decimal, lo redondeamos a 3,63.
3. Dibujar una tabla con dos columnas, la primera columna para el tallo y la segunda para las hojas. Disponer todos los tallos en la primera columna en orden descendente. Cada tallo solo se escribe una vez.
4. Registrar en la segunda columna todas las hojas, en orden creciente, junto al tallo correspondiente.

Ejemplo:

La siguiente tabla representa la estatura de un grupo de estudiantes de natación, se relacionan en el diagrama de tallo.

Altura de los 40 alumnos de una clase										Tallo	Hoja
145	147	149	152	153	154	154	156	157	158	14	5 7 9
162	162	162	163	163	164	164	165	167	167	15	2 3 4 4 6 7 8
168	169	169	170	171	171	172	173	174	174	16	2 2 2 3 3 4 4 5 7 7 8 9 9
175	176	176	178	179	180	181	183	185	186	17	0 1 1 2 3 4 4 5 6 6 8 9
										18	0 1 3 5 6



Actividad de profundización 3:

1. Se preguntó a un grupo de estudiantes de grado once sobre su asignatura favorita, los resultados son los siguientes:

ASIGNATURA	CANTIDAD DE ESTUDIANTES
Español	12
Matemáticas	8
Inglés	5
Sociales	10
Química	5

Con la información anterior realizar:

- Tabla de frecuencias.
- Representar la información en un diagrama de barras.

2. La siguiente tabla representa las calificaciones obtenidas por 50 alumnos en la asignatura de matemáticas instrumental II en un examen.

50	53	54	55	59	60	60	60	61	61
62	62	63	65	66	68	68	68	69	71
73	73	74	74	75	75	75	75	76	77
78	78	78	79	79	82	82	84	85	87
88	88	89	90	93	93	94	95	95	99

Con la información anterior:

- Hallar rango.
- Calcular el tamaño del intervalo a través de la regla de Sturges.
- Elaborar tabla con frecuencia absoluta, acumulada, relativa y porcentaje.
- Representar la información en histograma.

3. Se consulta a un grupo de personas sobre su género cinematográfico favorito, los resultados se presentan en la siguiente tabla:

GÉNERO CINEMATROGRÁFICO	No. DE PERSONAS
Acción	20
Drama	15
Documental	4
Ciencia ficción	6
Comedia	16
Infantil	12
Suspenso	8
Terror	5
Histórico	4
Otro	10

- A. Elaborar tabla de frecuencias.
B. Representar la información en un diagrama circular.

4. Los miembros de una cooperativa de viviendas tienen las siguientes edades:

42, 60, 60, 38, 60, 63, 21, 66, 56, 57, 51, 57, 44, 45, 35, 30,
35, 47, 53, 49, 50, 49, 38, 45, 28, 41, 47, 42, 53, 32, 54, 38,
40, 63, 48, 33, 35, 61, 47, 41, 55, 53, 27, 20, 21, 42, 21, 39,
39, 34, 45, 39, 28, 54, 33, 35, 43, 48, 48, 27, 53, 30, 29, 53,
38, 52, 54, 27, 27, 43, 28, 63, 41, 23, 58, 56, 59, 60, 40, 24

A. Elaborar tabla de frecuencias agrupando los datos en intervalos.

B. Representar la información en un polígono de frecuencias.

5. Se preguntó a un grupo de jóvenes sobre su deporte favorito, los resultados son:

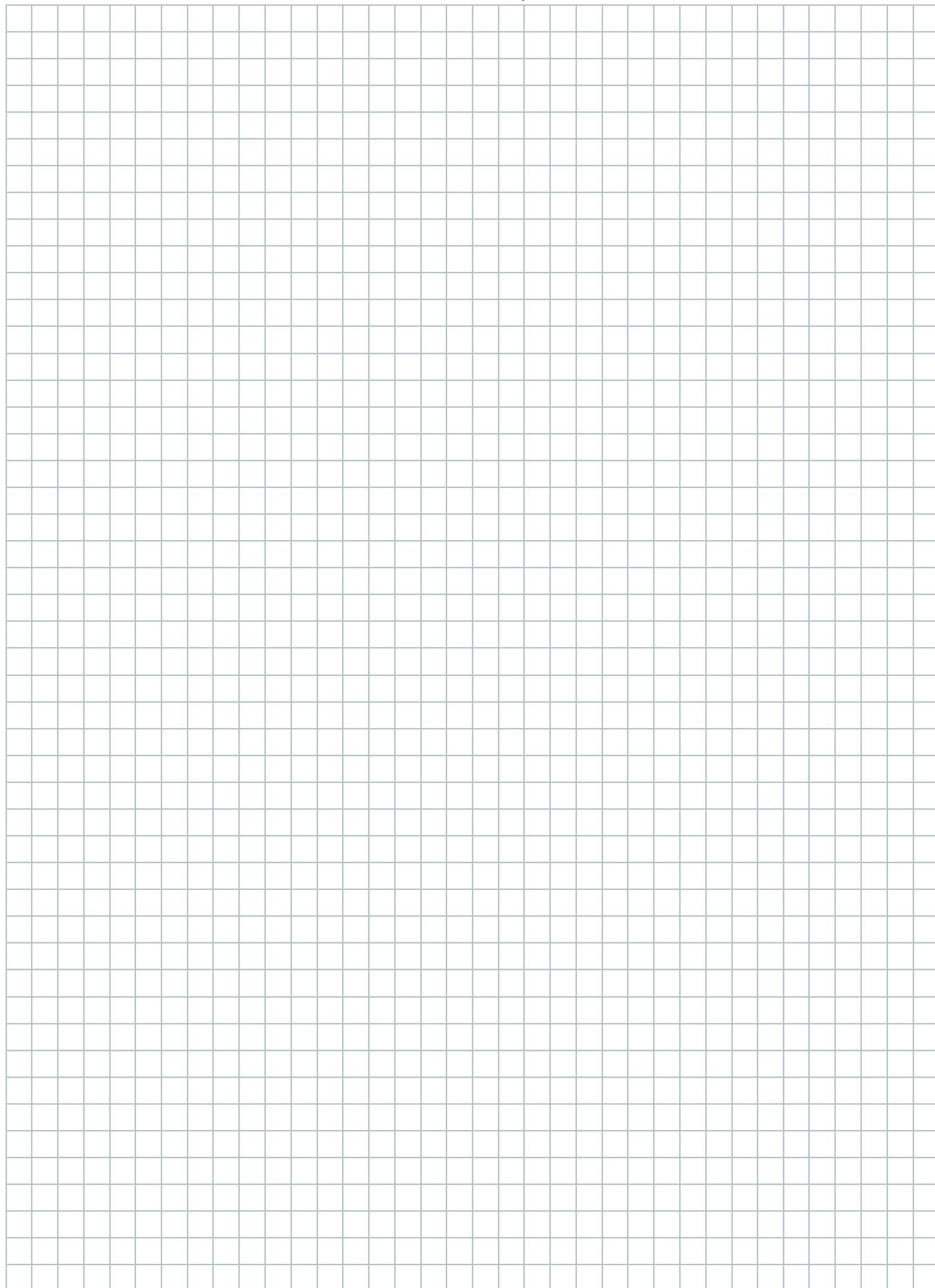
DEPORTE	No. PERSONAS
Fútbol	25
Baloncesto	10
Voleibol	5
Tenis	6
Boxeo	4

- A. Elaborar tabla de frecuencias.
- B. Representar la información en un pictograma.

6. En un sector de la localidad se obtuvo una muestra aleatoria del peso corporal de mujeres de la tercera edad para un estudio de determinación de la densidad ósea, los datos son los siguientes:

68, 72, 50, 70, 65, 83, 77, 78, 80, 93, 71, 74, 60, 84, 72, 84, 73, 81, 84, 92, 77, 57, 70, 59, 85, 74, 78, 79, 91, 102, 83, 67, 66, 75, 79, 82, 93, 90, 101, 80, 79, 69, 76, 94, 71, 97, 95, 83, 86, 69

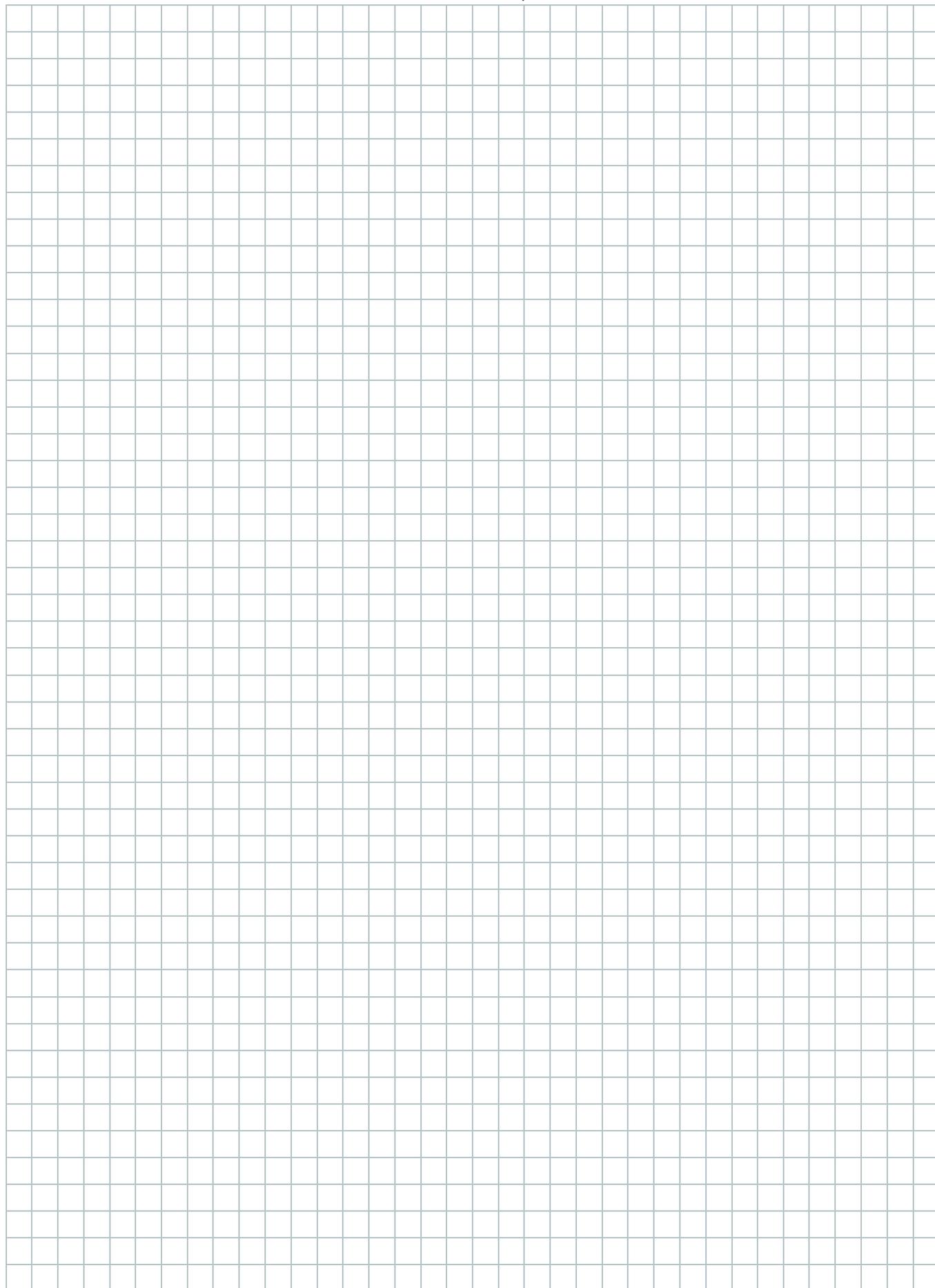
- A. Construir tabla con intervalos, marca de clase y frecuencias absoluta, relativa y porcentual.
- B. ¿Cuál es el dato que más se repite?
- C. Representar la información en un histograma.



7. Los siguientes datos corresponden a las edades que se obtuvieron de una encuesta realizada a mujeres atendidas de parto en un hospital de la localidad y que participaron en un estudio epidemiológico.

15, 38, 14, 13, 29, 25,
20, 13, 16, 32, 44, 39,
45, 46, 19, 23, 24, 18,
19, 20, 21, 18, 25, 33,
13, 18, 22, 24, 27, 27,

- A. Construir tabla con intervalos y frecuencias.
B. ¿Cuál es el límite inferior para la primera clase?
C. Representar la información en diagrama de barras y diagrama circular.
D. ¿Cuál es la edad que más se repite?



8. El número de hermanos de los alumnos de una clase es el siguiente:

0, 1, 0, 0, 3, 2, 1, 4, 0, 0, 1, 1, 2, 0, 1,

1, 2, 0, 1, 1, 2, 1, 3, 0, 0, 2, 1, 2, 3, 5,

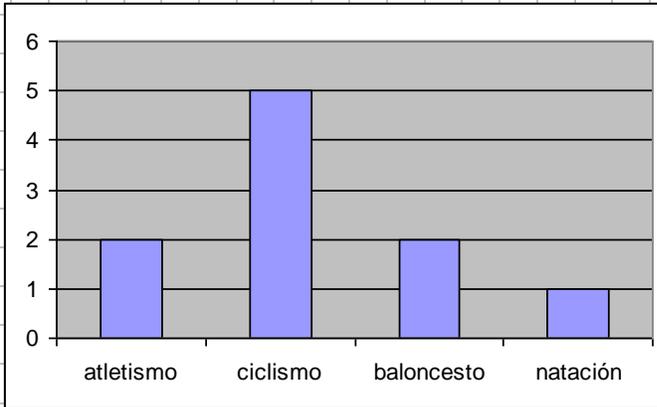
- A. Elabora una tabla de frecuencias en las que se incluyan: frecuencia absoluta, absoluta acumulada y relativas
- B. Representar la información en un diagrama de barras.
- C. ¿Qué porcentaje de alumnos son hijos únicos?
- D. ¿Cuántos alumnos tienen más de un hermano?

9. En una encuesta sobre vivienda se pregunta, entre otras cosas, cuántas personas viven en la casa, obteniéndose las siguientes respuestas:

**4, 4, 8, 1, 3, 2, 1, 3, 4, 2, 2, 7, 0, 3, 8, 0, 1, 5, 6, 4,
3, 3, 4, 5, 6, 8, 6, 2, 5, 3, 3, 5, 4, 6, 2, 0, 4, 3, 6, 1**

- A. Elaborar tabla de frecuencias.
- B. ¿Cuántas viviendas fueron objeto de estudio?
- C. ¿En cuántas de ellas no vive nadie?
- D. ¿Qué porcentaje de viviendas está ocupado por más de cinco personas?

10. A partir de la siguiente gráfica estadística de gustos deportivos:



- A. Realizar tabla de frecuencias.
- B. ¿A qué porcentaje de las personas no les gusta el ciclismo?
- C. Representar la información anterior en un diagrama circular.



AUTOEVALUACIÓN No.2

RESPONDE LAS PREGUNTAS 1 Y 2 CON BASE EN LA SIGUIENTE INFORMACIÓN:

En la siguiente tabla se registró la información de una encuesta realizada en el grado cuarto, donde cada estudiante seleccionó un sabor de helado de su preferencia.

SABORES	
Arequipe	▲ ▲ ▲
Chocolate	▲ ▲ ▲ ▲
Vainilla	▲ ▲
Ron con pasas	▲ ▲ ▲

Cada triángulo como este ▲ representa 3 estudiantes

- Según los datos de la tabla, el sabor de helado que prefiere la mayoría de estudiantes es:
 - Chocolate
 - Vainilla
 - Arequipe
 - Ron con pasas
- Según la información dada en la tabla, ¿Cuántos estudiantes en total respondieron la encuesta?
 - 9 estudiantes
 - 12 estudiantes
 - 36 estudiantes
 - 48 estudiantes

RESPONDE LAS PREGUNTAS 3 Y 4 CON BASE EN EL ENUNCIADO

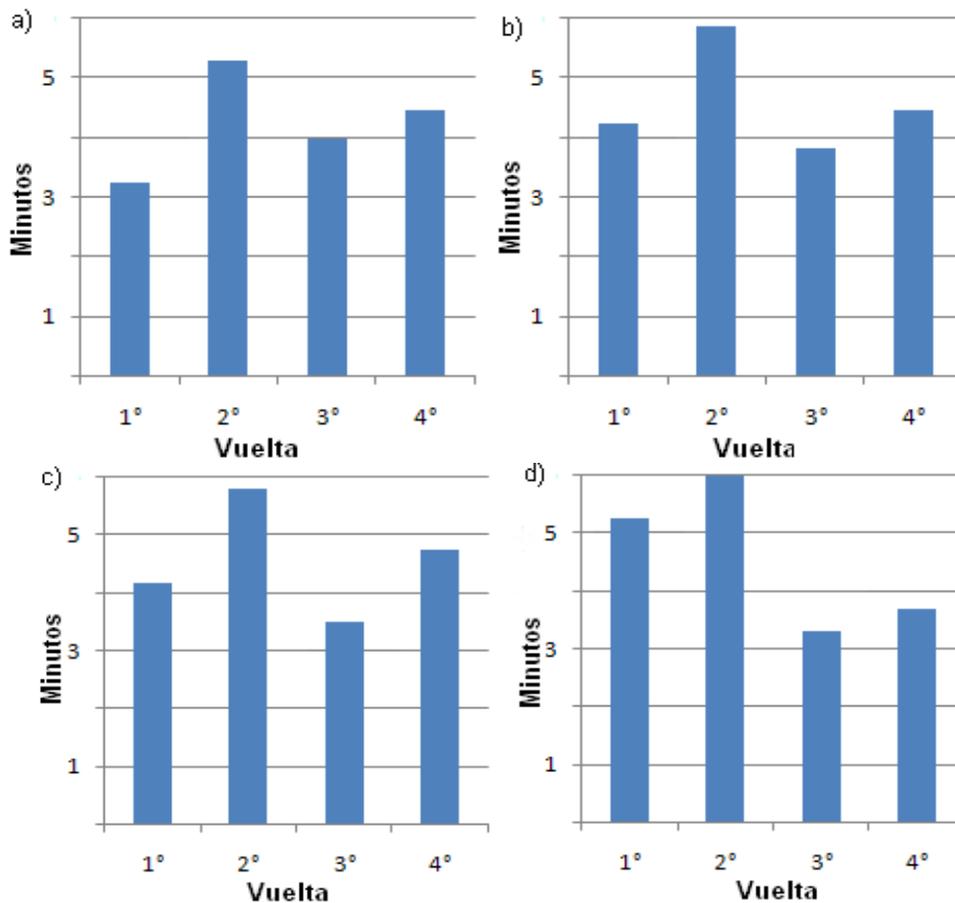
Luis dio cuatro vueltas alrededor de una pista de atletismo. El tiempo que él empleó en dar cada vuelta se muestra a continuación.

VUELTA	TIEMPO EMPLEADO
Primera vuelta	4 minutos y 15 segundos
Segunda vuelta	5 minutos y 45 segundos
Tercera vuelta	3 minutos y 30 segundos
Cuarta vuelta	4 minutos y 45 segundos

3. El tiempo total que empleo Luis en dar las cuatro vueltas alrededor de la pista fue de:

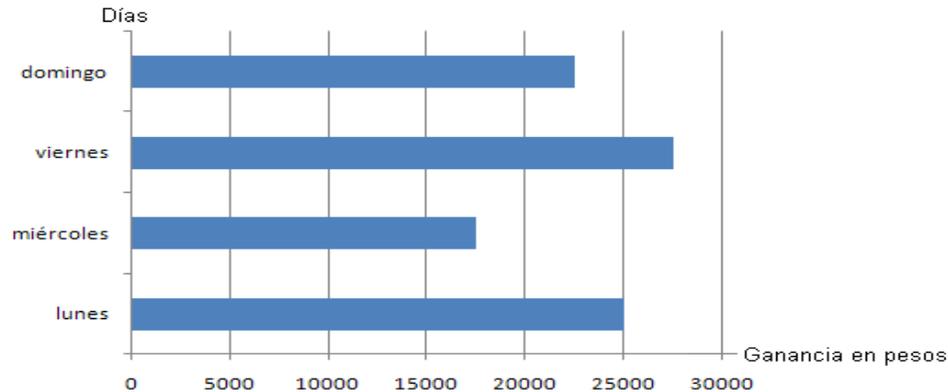
- a) 16 minutos
- b) 16 minutos y 30 segundos
- c) 17 minutos
- d) 18 minutos y 15 segundos

4. ¿En cuál de las siguientes gráficas se representan correctamente los datos de la tabla?



RESPONDE LAS PREGUNTAS 5 Y 6 CON BASE EN LA INFORMACIÓN DE LA GRÁFICA

Don Rubén, por cada 10 unidades que vende de un nuevo producto gana 2500 pesos. En la siguiente gráfica se muestra la ganancia que don Rubén obtuvo algunos días de la semana.



5. De acuerdo con la información de la gráfica, la venta del día domingo fue de:

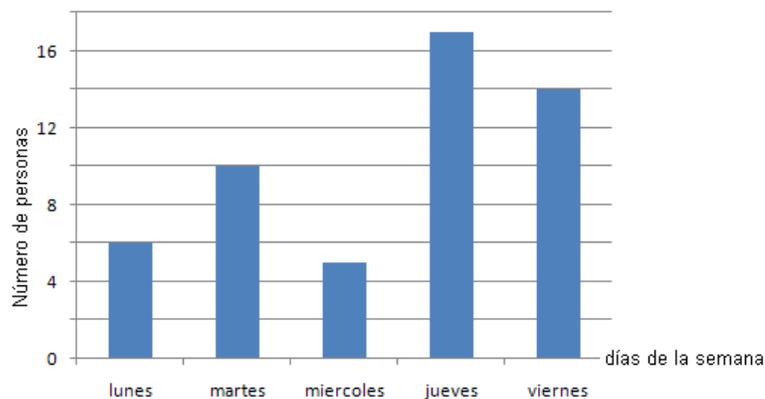
- a) 90 unidades
- b) 100 unidades
- c) 22500 unidades
- d) 2500 unidades

6. ¿Qué día vendió 109 unidades del producto?

- a) El lunes
- b) El miércoles
- c) El viernes
- d) El domingo

RESPONDE LAS PREGUNTAS 7 Y 8 CON LA SIGUIENTE INFORMACIÓN

La siguiente gráfica muestra la cantidad de personas atendidas en un centro médico durante una semana.



7. De acuerdo con la información de la gráfica es correcto afirmar que:

- a) El martes se atendieron menos personas que el jueves, pero más que el viernes.
- b) El viernes se atendieron más personas que el miércoles, pero menos que el jueves.
- c) El viernes se atendieron menos personas que el lunes, pero más que el jueves.
- d) El miércoles se atendieron más personas que el lunes, pero menos que el martes.

8. La tabla que representa la información dada en la gráfica es:

a)

DIA	NUMERO DE PERSONAS
Lunes	5
Martes	10
Miércoles	4
Jueves	14
Viernes	13

b)

DIA	NUMERO DE PERSONAS
Lunes	6
Martes	10
Miércoles	4
Jueves	15
Viernes	14

c)

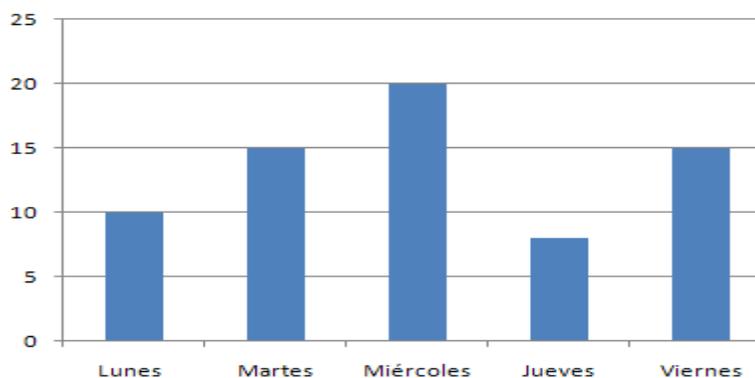
DIA	NUMERO DE PERSONAS
Lunes	5
Martes	10
Miércoles	5
Jueves	18
Viernes	13

d)

DIA	NUMERO DE PERSONAS
Lunes	6
Martes	10
Miércoles	5
Jueves	17
Viernes	14

RESPONDE LAS PREGUNTAS 9 Y 10 CON BASE EN LA GRÁFICA

Los niños y niñas de cuarto grado durante la semana recogieron envases de jugo reciclables. La información la registraron en la siguiente gráfica de barras.



9. El día en que se recogió la MENOR cantidad de envases fue:

- a) Lunes
- b) Miércoles
- c) Jueves
- d) Viernes

10. ¿Cuál de las siguientes tablas registra la información dada en la gráfica de barras?

a)

DIA	NUMERO DE ENVASES
Lunes	10
Martes	15
Miércoles	20
Jueves	8
Viernes	15

b)

DIA	NUMERO DE ENVASES
Lunes	2
Martes	3
Miércoles	20
Jueves	8
Viernes	3

c)

DIA	NUMERO DE ENVASES
Lunes	10
Martes	11
Miércoles	20
Jueves	10
Viernes	11

d)

DIA	NUMERO DE ENVASES
Lunes	2
Martes	3
Miércoles	4
Jueves	1
Viernes	3

Respuestas:

	1	2	3	4	5	6	7	8	9	10
a)	<input type="radio"/>									
b)	<input type="radio"/>									
c)	<input type="radio"/>									
d)	<input type="radio"/>									

MEDIDAS DE TENDENCIA CENTRAL

Las medidas de tendencia central, son medidas estadísticas que pretenden resumir en un solo valor a un conjunto de valores. Representan un centro en torno al cual se encuentra ubicado el conjunto de los datos. Las medidas de tendencia central más utilizadas son: media, mediana y moda.

Importancia de medidas de tendencia central

Las medidas de tendencia central (moda, media y mediana) sirven como puntos de referencia para interpretar las calificaciones que se obtienen en una prueba. En resumen, el propósito de las medidas de tendencia central es:

- Mostrar en qué lugar se ubica la persona promedio o típica del grupo.
- Sirve como un método para comparar o interpretar cualquier puntaje en relación con el puntaje central o típico.
- Sirve como un método para comparar el puntaje obtenido por una misma persona en dos diferentes ocasiones.
- Sirve como un método para comparar los resultados medios obtenidos por dos o más grupos.

Moda (Mo): la moda es el valor de la variable que más veces se repite. Si existe un solo valor que se repite, el conjunto tiene una sola moda y se llama UNIMODAL. Si existen dos valores que se repiten, el conjunto tiene dos modas, es BIMODAL. Si existen más de dos valores que se repiten, se dice que el conjunto tiene varias modas, en este caso se llama MULTIMODAL.

Moda para datos no agrupados

Ejemplo 1: Hallar la moda del siguiente conjunto de datos:

19, 1, 3, 4, 2, 5, 7, 6, 6, 6, 6, 6, 20, 17, 8, 18, 9, 10

En este conjunto el número que se repite es el 6, por tanto, la moda es **Mo = 6**

Ejemplo 2: Determinar la moda en el siguiente conjunto de datos que corresponden a las edades de niñas de un Jardín Infantil.

5, 7, 3, 3, 7, 8, 3, 5, 9, 5, 3, 4, 3.

La edad que más se repite es 3, por lo tanto, la Moda es 3 (**Mo = 3**)

Moda para datos agrupados

Para hallar la moda para datos agrupados, primeramente, se observa en columna las frecuencias, el valor más alto (clase con la mayor frecuencia.) Luego se halla la moda utilizando la siguiente fórmula:

$$Mo = li + \left(\frac{\Delta 1}{\Delta 1 + \Delta 2} \right) * A$$

Mo: Moda

Li: Limite inferior de la clase modal (clase con la mayor frecuencia)

$\Delta 1$: Frecuencia de la clase modal menos la frecuencia de la clase anterior.

$\Delta 2$: Frecuencia de la clase modal menos la frecuencia de la siguiente clase.

A: Amplitud o longitud del intervalo de clase.

Ejemplo 3: El peso en kg de un grupo de personas se presenta en la siguiente tabla. Encontrar la moda.

Edad	Frecuencia
[40,50]	12
[50,60]	20
[60,70]	35
[70,80]	39
[80,90]	4

Li: 70

Frecuencia de la clase (fi): 39

$\Delta 1$: $39 - 35 = 4$

$\Delta 2$: $39 - 4 = 35$

C: 10

$$Mo = Li + \left(\frac{\Delta 1}{\Delta 1 + \Delta 2} \right) * C$$

$$Mo = 70 + \left(\frac{4}{4 + 35} \right) * 10$$

$$Mo = 70 + \left(\frac{4}{39} \right) * 10$$

$$Mo = 70 + (0,10) * 10$$

$$Mo = 70 + 1$$

$$Mo = \mathbf{71}$$

Ejemplo 4: Encontrar la moda de los siguientes datos:

Intervalo	Frecuencia
[40,50]	12
[50,60]	8
[60,70]	5
[70,80]	3
[80,90]	2

Li: 40

Frecuencia de la clase (fi): 12

Δ_1 : 12 – 0 (se coloca 0 porque no hay clase anterior)

Δ_2 : 12 – 8 = 4

C: 10

$$Mo = Li + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) * C$$

$$Mo = 40 + \left(\frac{12}{12 + 4} \right) * 10$$

$$Mo = 40 + \left(\frac{12}{16} \right) * 10$$

$$Mo = 40 + (0,75) * 10$$

$$Mo = 40 + 7,5$$

$$Mo = \mathbf{47,5}$$

Media aritmética: Es aquella medida que se obtiene al dividir la suma de todos los valores de una variable por la frecuencia total. En palabras más simples, corresponde a la suma de un conjunto de datos dividida por el número total de dichos datos.

Media para datos no agrupados

$$\bar{X} = \frac{\text{Suma de todos los valores}}{\text{Cantidad total de datos}} = \frac{X_1 + X_2 + X_3 + X_4 + \dots + X_n}{N}$$

$$\bar{X} = \frac{x_i}{N}$$

Ejemplo 1: calcular el promedio de edad de un grupo de estudiantes que tienen los siguientes años: 15, 16, 14, 17 y 15.

$$\bar{X} = \frac{x_i}{N}$$

$$\bar{X} = \frac{15 + 16 + 14 + 17 + 15}{5} \rightarrow \bar{X} = \frac{77}{5} \rightarrow \bar{X} = \mathbf{15,4 \text{ años.}}$$

Ejemplo 2: en una serie de exámenes de matemáticas un estudiante obtuvo los siguientes aciertos: 4, 7, 7, 2, 5, 3. ¿cuál es el promedio del puntaje?

$$\bar{X} = \frac{x_i}{N}$$

$$\bar{X} = \frac{4 + 7 + 7 + 2 + 5 + 3}{6} \rightarrow \bar{X} = \frac{28}{6} \rightarrow \bar{X} \approx \mathbf{4,7}$$

Ejemplo 3: Se pregunta a un grupo de señoras sobre su peso en kg. Los resultados son: 58kg, 55 kg, 60 kg, 64 kg, 70 kg, 58 kg, 60 kg, 54 kg, 72 kg, 68 kg, 58 kg, 55 kg.

Solución: ordenando los datos en kg se obtiene: 54, 55, 55, 58, 58, 58, 60, 60, 64, 68, 70, 72.

$$\bar{X} = \frac{x_i}{N}$$

$$\bar{X} = \frac{54 + 55 + 55 + 58 + 58 + 58 + 60 + 60 + 64 + 68 + 70 + 72}{12} \rightarrow \bar{X} = \frac{732}{12}$$

$$\rightarrow \bar{X} = \mathbf{61 \text{ kg}}$$

Media para datos agrupados

Se calcula sumando todos los productos de **marca clase** con la frecuencia absoluta respectiva y su resultado dividirlo por el número total de datos:

$$\bar{X} = \frac{\text{Suma (marca clase } \times \text{ frecuencia absoluta)}}{\text{Total de datos}}$$

La **marca clase** de una tabla para datos agrupados en intervalos corresponde al promedio de los extremos de cada intervalo.

Ejemplo 1: En un test realizado a un grupo de 42 personas se han obtenido las puntuaciones que muestra la tabla. Calcular la puntuación media.

PUNTAJE	Fi
[10 – 20)	1
[20 – 30)	8
[30 – 40)	10
[40 – 50)	9
[50 – 60)	8
[60 – 70)	4
[70 – 80)	2
Total	42

Solución: lo primero es calcular la marca de clase:

PUNTAJE	Xi	Fi
[10 – 20)	15	1
[20 – 30)	25	8
[30 – 40)	35	10
[40 – 50)	45	9
[50 – 60)	55	8
[60 – 70)	65	4
[70 – 80)	75	2
Total		42

Después se realiza la multiplicación entre la marca de clase y la frecuencia absoluta y se suma el total de este resultado.

PUNTAJE	Xi	Fi	Xi . Fi
[10 – 20)	15	1	15
[20 – 30)	25	8	200
[30 – 40)	35	10	350
[40 – 50)	45	9	405
[50 – 60)	55	8	440
[60 – 70)	65	4	260
[70 – 80)	75	2	150
Total		42	1820

Finalmente se divide la suma total entre el número de datos:

$$\bar{X} = \frac{\sum xi}{N} \rightarrow \bar{X} = \frac{1820}{42} = 43,3$$

Ejemplo 2: En la siguiente tabla se representa la edad de un grupo de personas. Hallar la media.

EDAD	Fi
[13 – 15)	4
[15 – 17)	9
[17 – 19)	3
[19 – 21)	3
[21 – 23)	1
Total	20

Solución: lo primero es calcular la marca de clase:

EDAD	Xi	Fi
[13 – 15)	14	4
[15 – 17)	16	9
[17 – 19)	18	3
[19 – 21)	20	3
[21 – 23)	22	1
Total		20

Después se realiza la multiplicación entre la marca de clase y la frecuencia absoluta y se suma el total de este resultado.

EDAD	Xi	Fi	Xi . Fi
[13 – 15)	14	4	56
[15 – 17)	16	9	144
[17 – 19)	18	3	54
[19 – 21)	20	3	60
[21 – 23)	22	1	22
Total		20	336

Finalmente se divide la suma total entre el número de datos:

$$\bar{X} = \frac{\sum xi}{N} \rightarrow \bar{X} = \frac{336}{20} = 16,8$$

Mediana (Me): es el valor que ocupa la posición central en un conjunto de datos, que deben estar ordenados, de esta manera la mitad de las observaciones es menor que la mediana y la otra mitad es mayor que la mediana. Si existiesen dos elementos centrales, por tratarse de un N par, la mediana se calcularía con la media aritmética de los valores centrales correspondientes. La mediana se puede hallar sólo para variables cuantitativas.

Cálculo de la mediana con pocos datos

- Ordenamos los datos de menor a mayor.
- Si la serie tiene un número impar de medidas la mediana es la puntuación central de la misma.

2, 3, 4, 4, **5**, 5, 5, 6, 6 **Me= 5**

- Si la serie tiene un número par de puntuaciones la mediana es la media entre las dos puntuaciones centrales.

7, 8, **9, 10**, 11, 12 **Me= 9.5**

Ejemplo 1: Encontrar la mediana en el siguiente conjunto de datos: 4; 1; 4; 8; 5; 6; 9

Solución: lo primero es ordenar los datos (se puede hacer en orden creciente o decreciente), en este caso lo haremos en orden creciente.

1; 4; 4; 5; 6; 8; 9

Como son datos impares, se encontrará que el número que ocupe la posición central será el que corresponde a la mediana:

1; 4; 4; **5**; 6; 8; 9 **Me= 5**

Ejemplo 2: Encontrar la mediana en el siguiente conjunto de datos: 10; 13; 11; 9; 20; 7; 6; 4; 18; 16

Solución: lo primero es ordenar los datos (se puede hacer en orden creciente o decreciente), en este caso lo haremos en orden creciente.

4; 6; 7; 9; 10; 11; 13; 16; 18; 20.

Como son datos pares, se encontrará que en el centro estarán dos números:

4; 6; 7; 9; **10; 11**; 13; 16; 18; 20.

Posteriormente se calcula el promedio entre los dos números del centro (10; 11)

$$\text{Me} = \frac{10+11}{2} = \frac{21}{2} = 10,5$$

Ejemplo 3: Los sueldos de un grupo de trabajadores está dado en dólares y son los siguientes: 423, 425, 327, 440, 230, 255, 340, 550. Hallar la mediana de los sueldos:

Solución: lo primero es ordenar los datos (se puede hacer en orden creciente o decreciente), en este caso lo haremos en orden creciente.

230; 255; 327; 340; 423; 425; 440; 550

Como la cantidad de datos es par, se encontrará que en el centro estarán dos números:

230; 255; 327; **340; 423**; 425; 440; 550

Posteriormente se calcula el promedio entre los dos números del centro (340; 423)

$$\text{Me} = \frac{340+423}{2} = \frac{763}{2} = 381,5$$

En resumen, en caso de que la cantidad de datos sea impar, hay que buscar el número que se encuentre en la posición: $\frac{n+1}{2}$

Si la cantidad de datos es par, hay que promediar los números que se encuentran en las posiciones $\frac{n}{2}$ y $\frac{n}{2} + 1$

Cálculo de la mediana haciendo uso de la frecuencia acumulada

Cuando se tienen datos agrupados por intervalos, se puede calcular la mediana de la siguiente manera:

Se divide el total de datos entre dos.

En la columna de frecuencia acumulada se busca el valor igual o mayor al dato obtenido en el paso anterior.

El intervalo que se encuentre en este valor será el correspondiente a la mediana.

Ejemplo 1: La siguiente tabla representa el peso en kg de un grupo de personas estudiantes de música. Calcular la mediana.

Intervalo	fi
39 – 44	4
45 – 50	10
51 – 56	6
57 – 62	15
63 - 67	5
TOTAL	40

Solución: lo primero que se hará será calcular la frecuencia acumulada para los datos registrados.

Intervalo	fi	Fcia Acumulada
39 – 44	4	4
45 – 50	10	14
51 – 56	6	20
57 – 62	15	35
63 - 68	5	40
TOTAL	40	

Posteriormente se divide el total de datos entre dos: $\frac{40}{2}$, se obtiene como resultado el número 20. La frecuencia acumulada con esta característica (20), que corresponde al intervalo 51 – 56, por lo tanto, este intervalo es el correspondiente a la mediana.

Me= **[51 – 56]**

Teniendo en cuenta que la mediana es una medida de posición, si se desea calcular este valor con mayor exactitud, en un conjunto de datos agrupados en intervalos, se procede a calcularla haciendo uso de la siguiente fórmula:

Mediana en datos agrupados

$$Me = Li + A \left(\frac{\frac{n}{2} - \Sigma Fa - 1}{Fi} \right)$$

Me= mediana

Li= límite inferior

A= amplitud (se obtiene al restarle al límite superior el límite inferior)

n= total de datos

$\Sigma Fa - 1$ = suma de las frecuencias acumuladas inferiores a la clase mediana

Fi= frecuencia absoluta

Para una mejor comprensión lo haremos a través del siguiente ejemplo:

Ejemplo 2: Se tiene la siguiente tabla, la cual representa la cantidad de horas trabajadas por un grupo de obreros durante un mes:

Horas	Xi	Fi	Fa
55 – 60	57,5	5	5
60 – 65	62,5	18	23
65 – 70	67,5	20	43
70 – 75	72,5	50	93
75 – 80	77,5	17	110
80 – 85	82,5	16	126
85 – 90	87,5	4	130
TOTAL		130	

Solución: lo primero que se hace es dividir el total de datos entre 2.

$$\frac{130}{2} = 65$$

Este número se ubica en la tabla en la columna de la frecuencia acumulada, como en este caso no se encuentra este valor, entonces buscaremos el valor que le sigue (93)

Horas	X_i	F_i	F_a
55 – 60	57,5	5	5
60 – 65	62,5	18	23
65 – 70	67,5	20	43
70 – 75	72,5	50	93
75 – 80	77,5	17	110
80 – 85	82,5	16	126
85 – 90	87,5	4	130
TOTAL		130	

Al ubicarlo podemos determinar que el valor de la mediana se encuentra en el intervalo 70 – 75.

Ahora vamos a proceder a extraer los datos, para posteriormente reemplazar en la fórmula:

$$Li = 70$$

$$A = (75 - 70) = 5$$

$$n = 130$$

$$\Sigma Fa - 1 = 43$$

$$Fi = 50$$

$$\frac{n}{2} = 65$$

Reemplazando en la fórmula se obtiene:

$$Me = Li + A \left(\frac{\frac{n}{2} - \Sigma Fa - 1}{Fi} \right)$$

$$Me = 70 + 5 \left(\frac{65 - 43}{50} \right) \rightarrow Me = 70 + 5 \left(\frac{22}{50} \right) \rightarrow Me = 70 + 5(0,44) \rightarrow$$

$$Me = 70 + 2,2 \rightarrow \mathbf{Me = 72,2}$$

Ejemplo 3: Dada la siguiente tabla de frecuencias, hallar la mediana

Li - Ls	Fi	Fa
25 - 33	4	4
34 - 42	4	8
43 - 51	8	16
52 - 60	7	23
61 - 69	5	28
70 - 78	3	31
	n= 31	

Solución: lo primero que se hace es dividir el total de datos entre 2.

$$\frac{31}{2} = 15,5$$

Posteriormente ubicamos el valor anterior (15,5) en la tabla de frecuencia acumulada, en este caso el valor más próximo a 15,5 es 16 y se encuentra en el intervalo 43 – 51.

Li - Ls	Fi	Fa
25 - 33	4	4
34 - 42	4	8
43 - 51	8	16
52 - 60	7	23
61 - 69	5	28
70 - 78	3	31
	n= 31	

Ahora vamos a proceder a extraer los datos, para posteriormente reemplazar en la fórmula:

$$Li = 42,5 \text{ (como no se incluyen los intervalos, entonces se promedia } \frac{42+43}{2} = 42,5)$$

$$A = 9 \text{ (porque se incluyen los dos extremos)}$$

$$n = 31$$

$$\Sigma Fa - 1 = 8$$

$$Fi = 8$$

$$\frac{n}{2} = 15,5$$

Reemplazando en la fórmula se obtiene:

$$Me = Li + A \left(\frac{\frac{n}{2} - \Sigma Fa - 1}{Fi} \right) \rightarrow Me = 42,5 + 9 \left(\frac{15,5 - 8}{8} \right) \rightarrow$$

$$Me = 42,5 + 9 \left(\frac{7,5}{8} \right) \rightarrow Me = 42,5 + 9(0,9375) \rightarrow Me = 42,5 + 8,4375 \rightarrow$$

$$Me = 50,9375$$

Ejemplo 4: La siguiente tabla representa las edades de un grupo de estudiantes de música, calcular la mediana:

EDADES	Xi	Fi	Fa
13 – 15	14	4	4
15 – 17	16	9	13
17 – 19	18	3	16
19 – 21	20	3	19
21 - 23	22	1	20
TOTAL		20	

Solución: lo primero que se hace es dividir el total de datos entre 2.

$$\frac{20}{2} = 10$$

Este número se ubica en la tabla en la columna de la frecuencia acumulada, como en este caso no se encuentra este valor, entonces buscaremos el valor que le sigue (13)

Al ubicarlo podemos determinar que el valor de la mediana se encuentra en el intervalo 15 – 17.

EDADES	Xi	Fi	Fa
13 – 15	14	4	4
15 – 17	16	9	13
17 – 19	18	3	16
19 – 21	20	3	19
21 - 23	22	1	20
TOTAL		20	

Ahora vamos a proceder a extraer los datos, para posteriormente reemplazar en la fórmula:

$$Li = 15$$

$$A = (17 - 15) = 2$$

$$n = 20$$

$$\Sigma Fa - 1 = 4$$

$$Fi = 9$$

$$\frac{n}{2} = 10$$

Reemplazando en la fórmula se obtiene:

$$Me = Li + A \left(\frac{\frac{n}{2} - \Sigma Fa - 1}{Fi} \right)$$

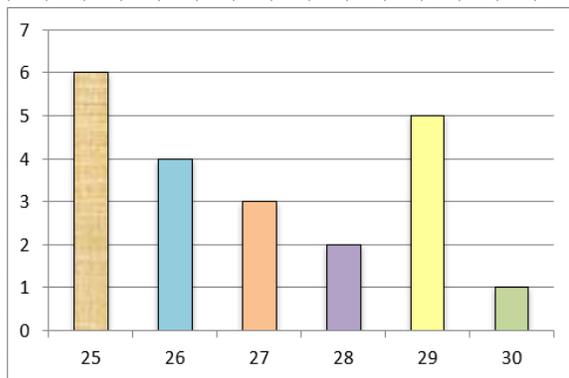
$$Me = 15 + 2 \left(\frac{10 - 4}{9} \right) \rightarrow Me = 15 + 2 \left(\frac{6}{9} \right) \rightarrow Me = 15 + 2(0,67) \rightarrow$$

$$Me = 15 + 1,34 \rightarrow \mathbf{Me = 16,34}$$



Actividad de profundización 4:

1. La siguiente grafica representa las edades de un grupo de empleados de una empresa. A partir de ella construir tabla de frecuencias, calcular moda, media y mediana.



2. El número de estrellas de los hoteles de una ciudad viene dado por la siguiente serie: 3, 3, 4, 3, 4, 3, 1, 3, 4, 3, 3, 3, 2, 1, 3, 3, 3, 2, 3, 2, 2, 3, 3, 3. Hallar la moda, la mediana y la media aritmética.

3. Un profesor de estadística desea dar a conocer el promedio de las notas finales de 10 alumnos de la clase. Las notas de los alumnos son:

7,2	8,1	6,4	9,0	8,5	9,0	7,5	8,8	4,2	10
-----	-----	-----	-----	-----	-----	-----	-----	-----	----

¿cuál es el promedio?

4. La siguiente tabla de frecuencia muestra el número de preguntas de 81 encuestados sobre un Test que consta de sólo seis preguntas.

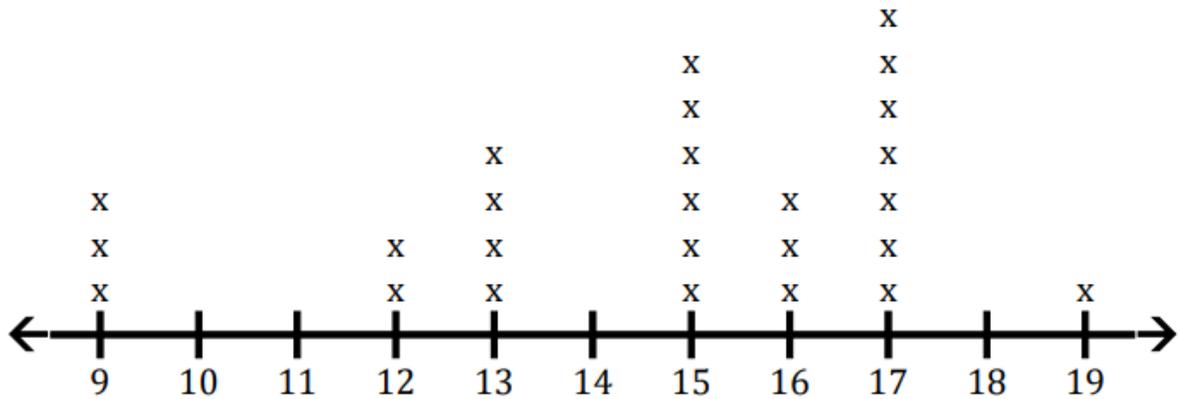
Preguntas acertadas	Número de Personas
1	15
2	13
3	8
4	19
5	21
6	5

Calcular la moda, media, mediana y representar la información en un diagrama de barras.

5. Los siguientes datos provienen del resultado de entrevistar a 30 personas sobre la marca de refresco que más consume a la semana. ¿Cuál es la moda?

Marca1	Marca2	Marca1	Marca1	Marca1	Marca3
Marca1	Marca3	Marca1	Marca2	Marca1	Marca1
Marca2	Marca1	Marca3	Marca3	Marca2	Marca1
Marca1	Marca1	Marca1	Marca3	Marca1	Marca2
Marca3	Marca1	Marca3	Marca3	Marca2	Marca3

6. De acuerdo al siguiente diagrama, responder:

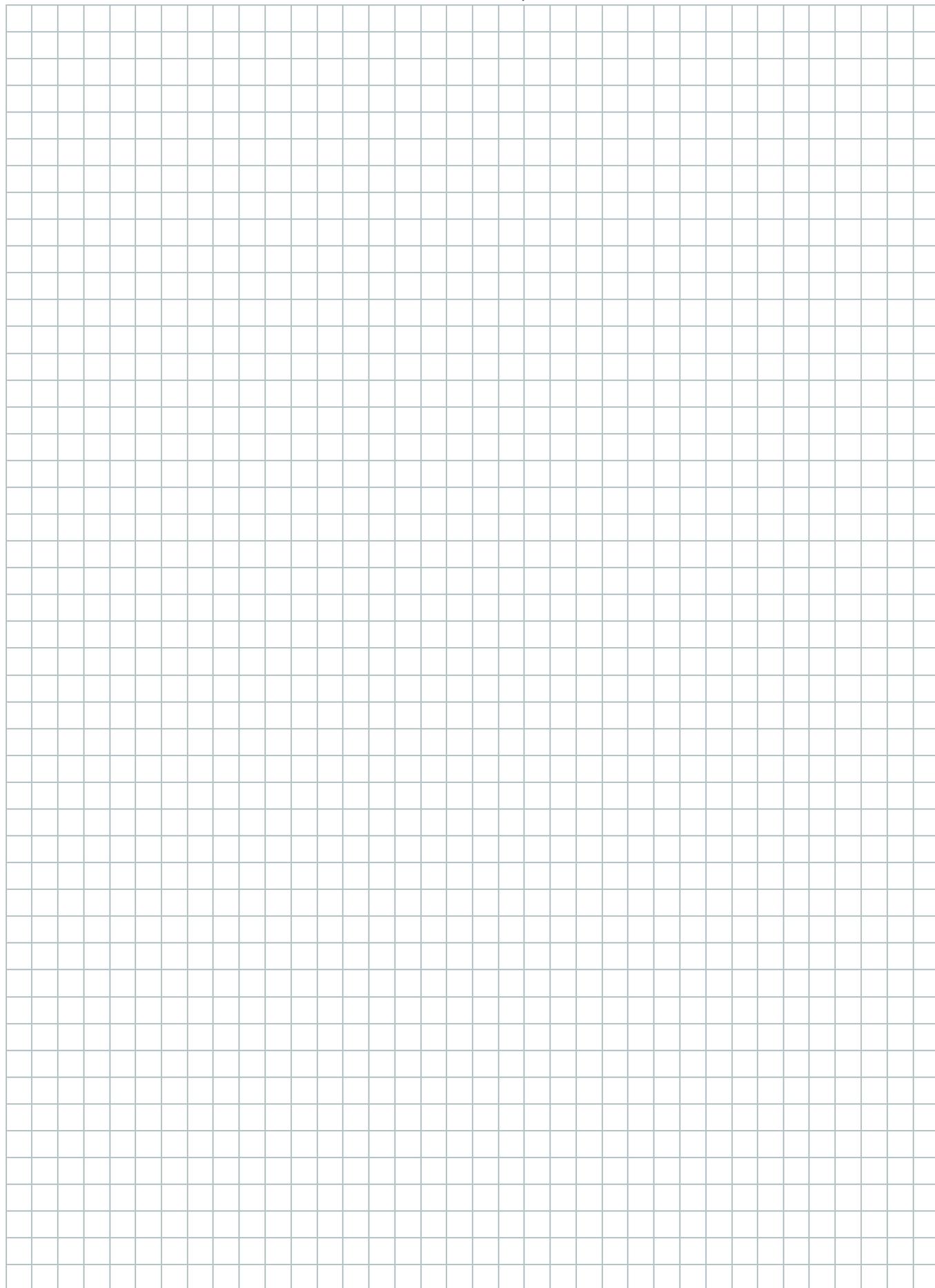


- A. Determinar el valor máximo, el valor mínimo y el rango de datos.
- B. Calcular la moda, media y mediana para datos no agrupados.
- C. ¿Cuántos valores son mayores que 12?
- D. ¿Cuántos valores son menores que 11?

7. De acuerdo al siguiente gráfico de estructura arbórea, realizar:

tallo	hoja
12	6
13	1 5 8 9 9 9
14	
15	0 1 2 5 6 8 8
16	2 3
17	1 1 5 9
18	
19	5 8
20	0 6 8
21	0 8

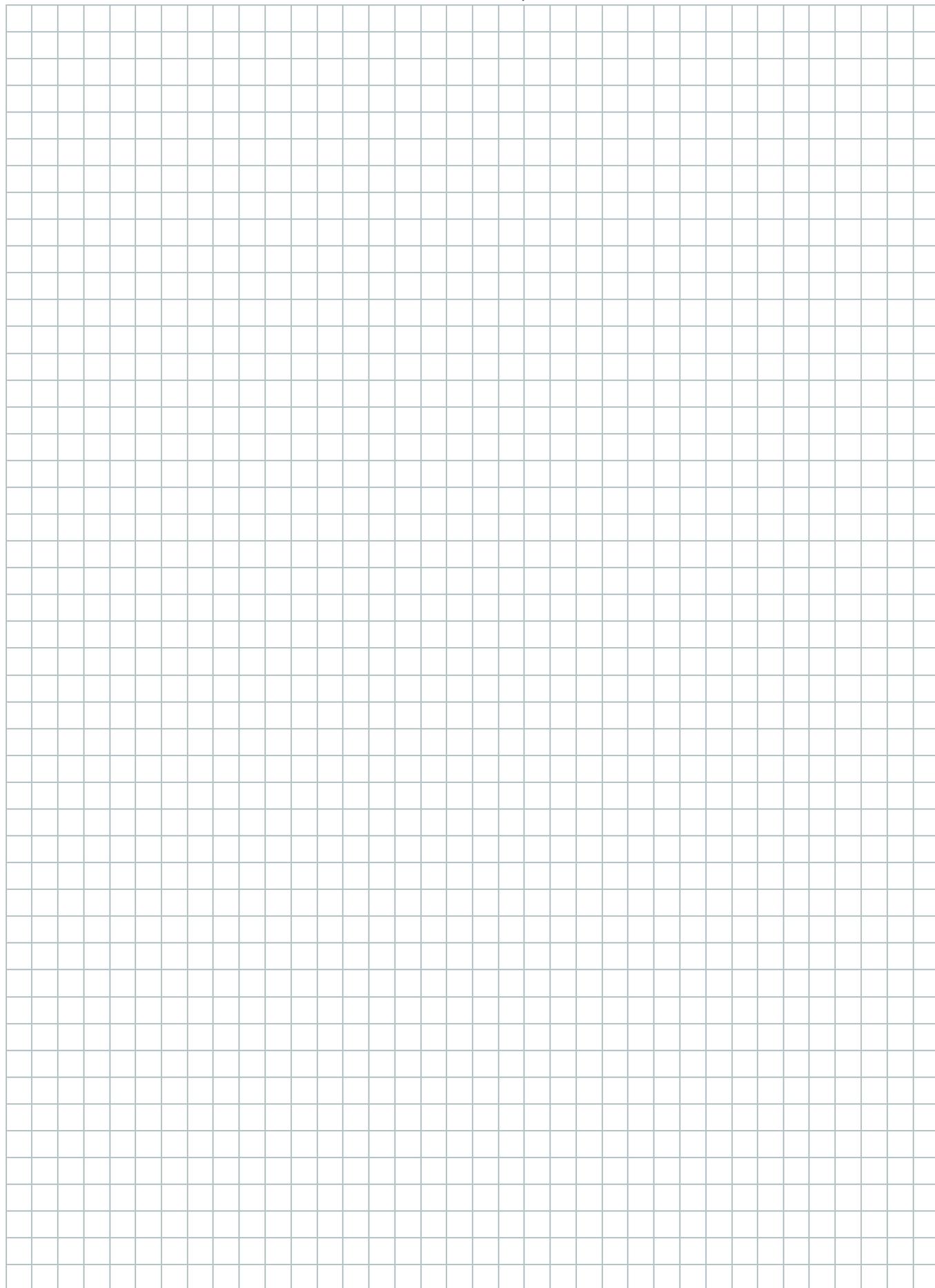
- Determinar el valor mínimo, el valor máximo y el rango de los datos.
- Construir tabla con intervalos haciendo uso de la regla de Sturges.
- Calcular moda, media y mediana para datos agrupados.
- ¿Cuántos valores son menores que 154?



8. De acuerdo al siguiente gráfico de estructura arbórea, realizar:

tallo	hoja
6	0 0 2 2 4 5 5 6 6 6 7 7 7 8 8 9 9
7	0 0 1 1 2 3 3 5 6 7 7 8 9
8	0 0 2 3 4 4 5 6 6 7 9 9 9 9
9	0 0 3 3 3 4 6 7 7 8 8 9 9 9 9
10	2
11	9
12	0 0 1 2 3 4 5 6 6 7 7
13	0 1 1 1 1 2 2 2 2 2 3 5 5 5 6 7 7 8 8 9
14	0 0 0 0 2 3 4 4 5 5 6 6 6 7 7 8 9 9 9 9
15	3 4

- A. Determinar el valor mínimo, el valor máximo y el rango de los datos.
- B. Construir tabla con intervalos haciendo uso de la regla de Sturges.
- C. Calcular moda, media y mediana para datos agrupados.
- D. ¿Cuántos valores son mayores que 111?
- E. ¿Cuántos valores son menores que 81?



MEDIDAS DE POSICIÓN

Las medidas de posición equivalen a los valores que puede tomar una variable caracterizados por agrupar a cierto porcentaje de observaciones en la muestra o población. Las medidas de posición son ideales para obtener información adicional a partir de datos resumidos, es decir, que presentan pérdida de información por agrupamiento en intervalos de clase.

Las medidas de posición son indicadores estadísticos que muestran la frecuencia acumulada hasta un valor k cualquiera.

CUARTILES (Q_k)

Los cuartiles son medidas estadísticas de posición que tienen la propiedad de dividir la serie estadística en cuatro grupos de números iguales de términos. Los cuartiles representan los valores 25%, 50% y 75%, que serían los cuartiles Q_1 , Q_2 y Q_3 . El cuartil Q_2 siempre es análogo a la mediana.

Pasos para calcular cuartiles:

Ordenar los datos de menor a mayor

Calcular frecuencias absolutas y frecuencias relativas acumuladas.

Aplicar la fórmula para cuartiles:

$$\frac{Kn}{4}$$

Ejemplo 1: La siguiente tabla representa las edades de un grupo de estudiantes de natación, calcular los cuartiles Q_1 , Q_2 y Q_3

Edad	Fi	Fa
13	3	3
14	14	17
15	23	40
16	10	50
17	5	55
18	4	59
19	1	60
TOTAL	60	

Solución: lo primero que se hace es identificar las variables de la fórmula:

$$\frac{Kn}{4}$$

K= número del cuartil

n= total de datos

Posteriormente se procede a calcular cada uno de los cuartiles

Q1:

$$Q_1 = \frac{Kn}{4} \rightarrow Q_1 = \frac{1.60}{4} \rightarrow Q_1 = \frac{60}{4} \rightarrow Q_1 = 15$$

El valor de 15 es la posición en la cual se encuentra el dato, entonces se ubica este valor en la columna de la frecuencia acumulada, si no está exactamente, entonces se utiliza el valor siguiente.

Edad	Fi	Fa
13	3	3
14	14	17
15	23	40
16	10	50
17	5	55
18	4	59
19	1	60
TOTAL	60	

En este caso el dato que se encuentra en la posición número 17 es **14**, entonces, este es el cuartil

Q2:

$$Q_2 = \frac{Kn}{4} \rightarrow Q_2 = \frac{2.60}{4} \rightarrow Q_2 = \frac{120}{4} \rightarrow Q_2 = 30$$

El valor de 30 es la posición en la cual se encuentra el dato, entonces se ubica este valor en la columna de la frecuencia acumulada, si no está exactamente, entonces se utiliza el valor siguiente.

Edad	Fi	Fa
13	3	3
14	14	17
15	23	40
16	10	50
17	5	55
18	4	59
19	1	60
TOTAL	60	

En este caso el dato que se encuentra en la posición número 40 es **15**, entonces, este es el cuartil

Q3:

$$Q_3 = \frac{Kn}{4} \rightarrow Q_3 = \frac{3 \cdot 60}{4} \rightarrow Q_3 = \frac{180}{4} \rightarrow Q_3 = 45$$

El valor de 45 es la posición en la cual se encuentra el dato, entonces se ubica este valor en la columna de la frecuencia acumulada, si no está exactamente, entonces se utiliza el valor siguiente.

Edad	Fi	Fa
13	3	3
14	14	17
15	23	40
16	10	50
17	5	55
18	4	59
19	1	60
TOTAL	60	

Ejemplo 2: Calcular los cuartiles en el siguiente conjunto de datos:

2, 3, 4, 5, 5, 7, 8, 10, 12, 12, 14, 15, 16, 16, 16, 20, 22, 22

Solución: En este ejercicio vamos a calcular los cuartiles de una manera más sencilla. Lo primero que se hace es verificar que los datos se encuentran ordenados de menor a mayor.

Después de verificar el orden, se procede a ubicar el Q_2 que es igual a la mediana. En este caso la cantidad de datos son pares, por lo tanto, se ubicarán los dos valores que quedan en el centro del conjunto de datos:

Q_2
 2, 3, 4, 5, 5, 7, 8, 10, 12, | 12, 14, 15, 16, 16, 16, 20, 22, 22

El cuartil 2 queda ubicado entre estos datos:

$$Q_2 = \frac{12 + 12}{2} \rightarrow Q_2 = \frac{24}{2} \rightarrow Q_2 = 12$$

Para hallar el cuartil 1 se calcula la mediana entre el primer bloque de números:

Q_1 Q_2
 2, 3, 4, 5, 5, 7, 8, 10, 12, | 12, 14, 15, 16, 16, 16, 20, 22, 22

El cuartil No.1 es el dato **5**.

Para hallar el cuartil 3 se calcula la mediana entre el segundo bloque de números:

Q_1 Q_2 Q_3
 2, 3, 4, 5, 5, 7, 8, 10, 12, | 12, 14, 15, 16, 16, 16, 20, 22, 22

El cuartil No.3 es el dato **16**.

Vamos a verificar el cálculo de los cuartiles haciendo uso de la tabla de frecuencias para demostrar la exactitud de ambos métodos:

Se realiza la tabla de frecuencias con los datos suministrados:

Número	Fi	Fa
2	1	1
3	1	2
4	1	3
5	2	5
7	1	6
8	1	7
10	1	8
12	2	10
14	1	11

15	1	12
16	3	15
20	1	16
22	2	18
TOTAL	18	

Q1:

$$Q_1 = \frac{Kn}{4} \rightarrow Q_1 = \frac{1 \cdot 18}{4} \rightarrow Q_1 = \frac{18}{4} \rightarrow Q_1 = 4,5$$

El valor de 4,5 es la posición en la cual se encuentra el dato, entonces se ubica este valor en la columna de la frecuencia acumulada, si no está exactamente, entonces se utiliza el valor siguiente.

Número	Fi	Fa
2	1	1
3	1	2
4	1	3
5	2	5
7	1	6
8	1	7
10	1	8
12	2	10
14	1	11
15	1	12
16	3	15
20	1	16
22	2	18
TOTAL	18	

En este caso el dato que se encuentra en la posición número 5 es 5, entonces, este es el cuartil

Q2:

$$Q_2 = \frac{Kn}{4} \rightarrow Q_2 = \frac{2 \cdot 18}{4} \rightarrow Q_2 = \frac{36}{4} \rightarrow Q_2 = 9$$

El valor de 9 es la posición en la cual se encuentra el dato, entonces se ubica este valor en la columna de la frecuencia acumulada, si no está exactamente, entonces se utiliza el valor siguiente.

Número	Fi	Fa
2	1	1
3	1	2
4	1	3
5	2	5
7	1	6
8	1	7
10	1	8
12	2	10
14	1	11
15	1	12
16	3	15
20	1	16
22	2	18
TOTAL	18	

En este caso el dato que se encuentra en la posición número 10 es **12**, entonces, este es el cuartil.

Q3:

$$Q_3 = \frac{Kn}{4} \rightarrow Q_3 = \frac{3 \cdot 18}{4} \rightarrow Q_3 = \frac{54}{4} \rightarrow Q_3 = 13,5$$

El valor de 13,5 es la posición en la cual se encuentra el dato, entonces se ubica este valor en la columna de la frecuencia acumulada, si no está exactamente, entonces se utiliza el valor siguiente.

Número	Fi	Fa
2	1	1
3	1	2
4	1	3
5	2	5
7	1	6
8	1	7
10	1	8
12	2	10
14	1	11
15	1	12
16	3	15
20	1	16
22	2	18
TOTAL	18	

Ejemplo 3: En el grado once, la profesora de matemáticas realiza un examen para verificar el conocimiento que tienen sus estudiantes sobre diversas temáticas de estadística, la cantidad de aciertos de los estudiantes se representan en la siguiente tabla:

RESPUESTAS	Fi	Fa
0 - 10	5	5
10 - 20	9	14
20 - 30	25	39
30 - 40	32	71
40 - 50	21	92
50 - 60	8	100
TOTAL	100	

Con esta información, calcular los cuartiles Q_1 , Q_2 y Q_3

Solución: En este ejercicio tenemos unos datos agrupados en intervalos y para calcular los cuartiles se emplea la siguiente fórmula:

$$Q_k = Li + A \left(\frac{\frac{Kn}{4} - Fa - 1}{Fa - Fa - 1} \right)$$

Ahora vamos a proceder a extraer los datos que requiere la fórmula en cada uno de los cuartiles, iniciando con el cálculo de la posición:

Q1:

$$Q_1 = \frac{Kn}{4} \rightarrow Q_1 = \frac{1.100}{4} \rightarrow Q_1 = \frac{100}{4} \rightarrow Q_1 = 25$$

RESPUESTAS	Fi	Fa
0 - 10	5	5
10 - 20	9	14
20 - 30	25	39
30 - 40	32	71
40 - 50	21	92
50 - 60	8	100
TOTAL	100	

$$Fa = 39$$

$$Fa - 1 = 14$$

$$Li = 20$$

$$Ls = 30$$

$$A = 10$$

Reemplazando en la fórmula se obtiene:

$$Q_k = Li + A \left(\frac{\frac{K_n}{4} - Fa - 1}{Fa - Fa - 1} \right)$$

$$Q_1 = 20 + 10 \left(\frac{25 - 14}{39 - 14} \right) \rightarrow Q_1 = 20 + 10 \left(\frac{11}{25} \right) \rightarrow Q_1 = 20 + 10(0,44)$$

$$Q_1 = 20 + 4,4 \rightarrow Q_1 = 24,4$$

Q2:

$$Q_2 = \frac{Kn}{4} \rightarrow Q_2 = \frac{2 \cdot 100}{4} \rightarrow Q_2 = \frac{200}{4} \rightarrow Q_2 = 50$$

RESPUESTAS	Fi	Fa
0 - 10	5	5
10 - 20	9	14
20 - 30	25	39
30 - 40	32	71
40 - 50	21	92
50 - 60	8	100
TOTAL	100	

$$Fa = 71$$

$$Fa - 1 = 39$$

$$Li = 30$$

$$Ls = 40$$

$$A = 10$$

Reemplazando en la fórmula se obtiene:

$$Q_k = Li + A \left(\frac{\frac{K_n}{4} - Fa - 1}{Fa - Fa - 1} \right)$$

$$Q_2 = 30 + 10 \left(\frac{50 - 39}{71 - 39} \right) \rightarrow Q_2 = 30 + 10 \left(\frac{11}{32} \right) \rightarrow Q_2 = 30 + 10(0,34)$$

$$Q_2 = 30 + 3,4 \rightarrow Q_2 = 33,4$$

Q3:

$$Q_3 = \frac{Kn}{4} \rightarrow Q_3 = \frac{3 \cdot 100}{4} \rightarrow Q_3 = \frac{300}{4} \rightarrow Q_3 = 75$$

RESPUESTAS	Fi	Fa
0 - 10	5	5
10 - 20	9	14
20 - 30	25	39
30 - 40	32	71
40 - 50	21	92
50 - 60	8	100
TOTAL	100	

$$Fa = 92$$

$$Fa - 1 = 71$$

$$Li = 40$$

$$Ls = 50$$

$$A = 10$$

Reemplazando en la fórmula se obtiene:

$$Q_k = Li + A \left(\frac{\frac{Kn}{4} - Fa - 1}{Fa - Fa - 1} \right)$$

$$Q_3 = 40 + 10 \left(\frac{75 - 71}{92 - 71} \right) \rightarrow Q_3 = 40 + 10 \left(\frac{4}{21} \right) \rightarrow Q_3 = 40 + 10(0,19)$$

$$Q_3 = 40 + 1,9 \rightarrow Q_3 = 41,9$$

DECILES (D_k)

Son los valores de la variable que dividen en 10 partes iguales la distribución: los deciles se representan por las letras $D_1, D_2, D_3, D_4, \dots, D_9$.

Pasos para calcular deciles:

Ordenar los datos de menor a mayor

Calcular frecuencias absolutas y frecuencias relativas acumuladas.

Aplicar la fórmula para deciles:

$$\frac{Kn}{10}$$

Ejemplo 1: La siguiente tabla representa las edades de un grupo de estudiantes de natación, calcular el decil No. 6

Edad	Fi	Fa
13	3	3
14	14	17
15	23	40
16	10	50
17	5	55
18	4	59
19	1	60
TOTAL	60	

Solución: lo primero que se hace es identificar las variables de la fórmula:

$$\frac{Kn}{10}$$

K= número del decil

n= total de datos

Posteriormente se procede a calcular cada uno de los deciles

D₆:

$$D_6 = \frac{Kn}{10} \rightarrow D_6 = \frac{6 \cdot 60}{10} \rightarrow D_6 = \frac{360}{10} \rightarrow D_6 = 36$$

El valor de 36 es la posición en la cual se encuentra el dato, entonces se ubica este valor en la columna de la frecuencia acumulada, si no está exactamente, entonces se utiliza el valor siguiente.

Edad	Fi	Fa
13	3	3
14	14	17
15	23	40
16	10	50
17	5	55
18	4	59
19	1	60
TOTAL	60	

En este caso el dato que se encuentra en la posición número 40 es **15**, entonces, este es el decil No.6

Ejemplo 2: Calcular el valor del decil 8 en el siguiente conjunto de datos:

2, 5, 5, 7, 10, 13, 32, 34, 20, 8, 3, 0

Solución: En este ejercicio lo primero que se realiza es ordenar los datos de menor a mayor

0, 2, 3, 5, 5, 7, 8, 10, 13, 20, 32, 34

se hace es identificar las variables de la fórmula:

$$\frac{Kn}{10}$$

K= número del decil (8)

n= total de datos (12)

Posteriormente se procede a calcular el decil No.8

D8:

$$D_8 = \frac{Kn}{10} \rightarrow D_8 = \frac{8 \cdot 12}{10} \rightarrow D_8 = \frac{96}{10} \rightarrow D_8 = 9,6 \approx 10$$

El valor de 9,6 es la posición en la cual se encuentra el dato, en este caso por ser decimal no se encuentra, entonces se aproxima a 10 y se ubica este valor en la serie (20)

0, 2, 3, 5, 5, 7, 8, 10, 13, 20, 32, 34

9 10

Ejemplo 3: En el grado once, la profesora de matemáticas realiza un examen para verificar el conocimiento que tienen sus estudiantes sobre diversas temáticas de estadística, la cantidad de aciertos de los estudiantes se representan en la siguiente tabla:

RESPUESTAS	Fi	Fa
0 - 10	5	5
10 - 20	9	14
20 - 30	25	39
30 - 40	32	71
40 - 50	21	92
50 - 60	8	100
TOTAL	100	

Con esta información, calcular el decil No.2

Solución: En este ejercicio tenemos unos datos agrupados en intervalos y para calcular los deciles se emplea la siguiente fórmula:

$$D_k = Li + A \left(\frac{\frac{Kn}{10} - Fa - 1}{Fa - Fa - 1} \right)$$

Ahora vamos a proceder a extraer los datos que requiere la fórmula, iniciando con el cálculo de la posición:

$$\frac{Kn}{10} = \frac{2 \cdot 100}{10} = \frac{200}{10} = 20$$

RESPUESTAS	Fi	Fa
0 - 10	5	5
10 - 20	9	14
20 - 30	25	39
30 - 40	32	71
40 - 50	21	92
50 - 60	8	100
TOTAL	100	

$$Fa = 39$$

$$Fa - 1 = 14$$

$$Li = 20$$

$$Ls = 30$$

$$A = 10$$

Reemplazando en la fórmula se obtiene:

$$D_2 = Li + A \left(\frac{\frac{K_n}{10} - Fa - 1}{Fa - Fa - 1} \right)$$

$$D_2 = 20 + 10 \left(\frac{20 - 14}{39 - 14} \right) \rightarrow D_2 = 20 + 10 \left(\frac{6}{25} \right) \rightarrow D_2 = 20 + 10(0,24)$$

$$D_2 = 20 + 2,4 \rightarrow D_2 = 22,4$$

PERCENTILES (P_k)

Los percentiles representan los valores de la variable que están por debajo de un porcentaje, el cual puede ser un valor de 1% a 100% (en otras palabras, el total de los datos es dividido en 100 partes iguales).

La notación empleada será P_k , donde k es equivalente al porcentaje de datos acumulados, y P_k es el valor de la variable que representa dicho porcentaje. Por ejemplo, P_5 es el valor de la variable que deja por debajo el 5% de los datos. P_{78} será entonces el valor que agrupa el 78% de los datos.

Pasos para calcular percentiles:

Ordenar los datos de menor a mayor

Calcular frecuencias absolutas y frecuencias relativas acumuladas.

Aplicar la fórmula para percentiles:

$$\frac{Kn}{100}$$

Ejemplo 1: La siguiente tabla representa las edades de un grupo de estudiantes de natación, calcular percentil No. 5:

Edad	Fi	Fa
13	3	3
14	14	17
15	23	40
16	10	50
17	5	55
18	4	59
19	1	60
TOTAL	60	

Solución: lo primero que se hace es identificar las variables de la fórmula:

$$\frac{Kn}{100}$$

K= número del percentil

n= total de datos

Posteriormente se procede a calcular el dato de posición

P5:

$$P_5 = \frac{Kn}{100} \rightarrow P_5 = \frac{5.60}{100} \rightarrow P_5 = \frac{300}{100} \rightarrow P_5 = 3$$

El valor de 3 es la posición en la cual se encuentra el dato, entonces se ubica este valor en la columna de la frecuencia acumulada, si no está exactamente, entonces se utiliza el valor siguiente.

Edad	Fi	Fa
13	3	3
14	14	17
15	23	40
16	10	50
17	5	55
18	4	59
19	1	60
TOTAL	60	

En este caso el dato que se encuentra en la posición número 3 es **13**, entonces, este es el percentil No.6.

Ejemplo 2: En el grado once, la profesora de matemáticas realiza un examen para verificar el conocimiento que tienen sus estudiantes sobre diversas temáticas de estadística, la cantidad de aciertos de los estudiantes se representan en la siguiente tabla:

RESPUESTAS	Fi	Fa
0 - 10	5	5
10 – 20	9	14
20 – 30	25	39
30 – 40	32	71
40 – 50	21	92
50 – 60	8	100
TOTAL	100	

Con esta información, calcular el percentil No.71

Solución: En este ejercicio tenemos unos datos agrupados en intervalos y para calcular los deciles se emplea la siguiente fórmula:

$$P_k = Li + A \left(\frac{\frac{Kn}{100} - Fa - 1}{Fa - Fa - 1} \right)$$

Ahora vamos a proceder a extraer los datos que requiere la fórmula, iniciando con el cálculo de la posición:

$$\frac{Kn}{100} = \frac{71 \cdot 100}{100} = \frac{7100}{100} = 71$$

RESPUESTAS	Fi	Fa
0 - 10	5	5
10 - 20	9	14
20 - 30	25	39
30 - 40	32	71
40 - 50	21	92
50 - 60	8	100
TOTAL	100	

En este caso la frecuencia acumulada es exacta, por lo tanto, el valor del percentil es el límite superior de este intervalo (40).

MEDIDAS DE DISPERSIÓN

Las medidas de dispersión en cambio miden el grado de dispersión de los valores de la variable. Dicho en otros términos las medidas de dispersión pretenden evaluar en qué medida los datos difieren entre sí. También son llamadas medidas de variabilidad y muestran la variabilidad de una distribución, indicando por medio de un número si las diferentes puntuaciones de una variable están muy alejadas de la media. Cuanto mayor sea ese valor, mayor será la variabilidad, y cuanto menor sea, más homogénea será a la media.

Así se sabe si todos los casos son parecidos o varían mucho entre ellos. Las medidas de dispersión nos informan sobre cuánto se alejan del centro los valores de la distribución.

CARACTERÍSTICAS DE MEDIDAS DE DISPERSION

- Muestran la variabilidad de una distribución, indicando por medio de un número, si las diferentes puntuaciones de una variable están muy alejadas de la media.
- Nos informan sobre cuánto se alejan del centro los valores de la distribución.
- Cuanto mayor sea ese valor, mayor será la variabilidad, cuanto menor sea, más homogénea será a la media.
- Se sabe si todos los casos son parecidos o varían mucho entre ellos.
- Indican si esos datos están próximos entre sí o sí están dispersos.
- Indican cuan esparcidos se encuentran los datos.
- Permiten apreciar la distancia que existe entre los datos a un cierto valor central e identificar la concentración de los mismos en un cierto sector de la distribución; es decir, permiten estimar cuan dispersas están dos o más distribuciones de datos.

TIPOS DE MEDIDAS DE DISPERSION

Medidas de Dispersión Absolutas

- Rango o Recorrido
- Desviación media
- Desviación estándar o típica
- Varianza

Medidas de Dispersión Relativas

- Coeficiente de variación



MEDIDAS DE DISPERSIÓN ABSOLUTAS

Rango o Recorrido

El **Rango** es la diferencia entre el mayor y el menor de los datos de una distribución estadística.

$$R = V_{max} - V_{min}$$

Características de Rango o Recorrido:

- Solo suministra información de los extremos de la variable.
- Informa sobre la distancia entre el mínimo y máximo valor observado.
- Se limita su uso a una información inicial.

Varianza

La **Varianza** es la media aritmética del cuadrado de las desviaciones respecto a la media de una distribución estadística.

Características de la Varianza:

- Es siempre un valor no negativo, que puede ser igual o distinta de 0.
- La Varianza es la medida de dispersión cuadrática optima por ser la menor de todas.
- Si a todos los valores de la variable se le suma una constante la varianza no se modifica.
- Si todos los valores de la variable se multiplican por una constante, la varianza queda multiplicada por el cuadrado de dicha constante.

Desviación Media

La **Desviación Media** es la diferencia entre cada valor de la variable estadística y la media aritmética.

La desviación media se representa por:

$$DM = \frac{\sum |X_i - \bar{X}|}{n}$$

DM= desviación media

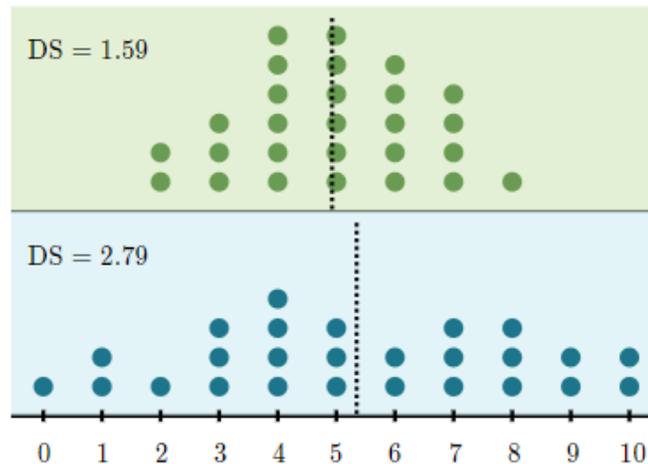
Características de la Deviación Media:

- Todas las observaciones se usan en el cálculo.
- No tiene la influencia debido a los valores altos y bajos.
- Es un poco difícil trabajar con los valores absolutos.

Desviación Estándar o Típica

La desviación estándar mide la dispersión de una distribución de datos. Entre más dispersa está una distribución de datos, más grande es su desviación estándar.

Por ejemplo, la distribución azul en la parte de abajo tiene una desviación estándar mayor que la distribución verde de arriba:



La desviación estándar no puede ser negativa. Una desviación estándar cercana a 000 indica que los datos tienden a estar más cerca a la media (se muestra por la línea punteada). Entre más lejos estén los datos de la media, más grande es la desviación estándar.

La **Desviación estándar** es la raíz cuadrada de la varianza; es decir, la raíz cuadrada de la media de los cuadrados de las puntuaciones de desviación.

Se aplican dos fórmulas para hallar la varianza y desviación estándar, cuando se trata de una población y cuando se trata de una muestra.

Población:

$$\sigma^2 = \frac{\sum(x - \bar{x})^2}{N}$$

Muestra:

$$S^2 = \frac{\sum(x - \bar{x})^2}{n-1}$$

σ^2 = varianza con una población

S^2 = varianza con una muestra

Σ = significa “suma de”

X = es un valor de un conjunto de datos.

\bar{X} = es la media del conjunto de datos

N = es el número de datos (cuando es en una muestra se representa por la n minúscula)

Características de la Desviación Estándar o Típica:

Es afectada por el valor de cada observación.

Como consecuencia de considerar desviaciones cuadráticas pone mayor énfasis en las desviaciones extremas que en las demás desviaciones.

Ejemplo 1:

Se preguntó las edades a un grupo de niños, los resultados son: 5, 6, 6, 7, 8

Calcular la desviación

Solución: En este ejercicio lo primero que haremos será calcular la media o promedio de las edades:

$$\bar{x} = \frac{5 + 6 + 6 + 7 + 8}{5} = \frac{32}{5} = 6,4$$

Posteriormente vamos a calcular la varianza como una población:

$$\sigma^2 = \frac{\sum(x - \bar{x})^2}{n}$$

$$\sigma^2 = \frac{(5-6,4)^2 + (6-6,4)^2 + (6-6,4)^2 + (7-6,4)^2 + (8-6,4)^2}{5} \rightarrow$$

$$\sigma^2 = \frac{(-1,4)^2 + (-0,4)^2 + (-0,4)^2 + (0,6)^2 + (1,6)^2}{5} \rightarrow$$

$$\sigma^2 = \frac{1,96 + 0,16 + 0,16 + 0,36 + 2,56}{5} \rightarrow$$

$$\sigma^2 = \frac{5,2}{5} \rightarrow \sigma^2 = 1,04$$

Finalmente calculamos la desviación estándar, la cual se obtiene al extraer la raíz cuadrada de la varianza.

$$\sigma^2 = 1,04 \rightarrow \sigma = \sqrt{1,04} \rightarrow \sigma = 1,02$$

Ejemplo 2:

En los siguientes datos, calcular la varianza y la desviación estándar para una población y para una muestra. 6, 2, 3, 1

Solución: En este ejercicio lo primero que haremos será calcular la media o promedio de los datos:

$$\bar{x} = \frac{6 + 2 + 3 + 1}{4} = \frac{12}{4} = 3$$

Posteriormente vamos a calcular la varianza como una población:

$$\sigma^2 = \frac{\sum(x - \bar{x})^2}{n}$$

$$\sigma^2 = \frac{(6-3)^2 + (2-3)^2 + (3-3)^2 + (1-3)^2}{4} \rightarrow$$

$$\sigma^2 = \frac{(3)^2 + (-1)^2 + (0)^2 + (-2)^2}{4} \rightarrow$$

$$\sigma^2 = \frac{9 + 1 + 0 + 4}{4} \rightarrow$$

$$\sigma^2 = \frac{14}{4} \rightarrow \sigma^2 = 3,5$$

Finalmente calculamos la desviación estándar, la cual se obtiene al extraer la raíz cuadrada de la varianza.

$$\sigma^2 = 3,5 \rightarrow \sigma = \sqrt{3,5} \rightarrow \sigma = 1,87$$

Para calcular la varianza y desviación estándar en una muestra, sólo se resta 1 a los datos y se representa la varianza por la letra S

$$S^2 = \frac{14}{3} \rightarrow S^2 = 4,67$$

Finalmente calculamos la desviación estándar, la cual se obtiene al extraer la raíz cuadrada de la varianza.

$$S^2 = 4,67 \rightarrow S = \sqrt{3,5} \rightarrow S = 1,87$$

Ejemplo 3:

Calcular la varianza y desviación estándar como una muestra para el siguiente conjunto de datos agrupados en intervalos, los cuales corresponden a la cantidad de respuestas correctas en un examen de matemáticas aplicado a 100 estudiantes:

RESPUESTAS	x	fi
0 – 10	5	6
10 – 20	15	19
20 – 30	25	45
30 – 40	35	22
40 - 50	45	8
TOTAL		100

Se aplican dos fórmulas para hallar la varianza y desviación estándar, cuando se trata de una población y cuando se trata de una muestra.

Población:

$$\sigma^2 = \frac{\sum(x - \mu)^2}{N} \cdot fi$$

Muestra:

$$S^2 = \frac{\sum(x - \bar{x})^2}{n-1} \cdot fi$$

σ^2 = varianza con una población

S^2 = varianza con una muestra

Σ = significa “suma de”

X= es un valor de un conjunto de datos.

\bar{X} = es la media del conjunto de datos

μ = es la forma como se representa el promedio para una población.

N= es el número de datos.

Solución: En este ejercicio lo primero que haremos será calcular la media o promedio de los datos, para ello agregaremos a la tabla una nueva columna, con la multiplicación de $x \cdot f_i$ y de esta manera podemos obtener de manera más sencilla el promedio:

DATOS	x	f _i	x.f _i
0 – 10	5	6	30
10 – 20	15	19	285
20 – 30	25	45	1125
30 – 40	35	22	770
40 - 50	45	8	360
TOTAL		100	2570

$$\bar{X} = \frac{\sum x \cdot f}{n} \rightarrow \bar{X} = \frac{2570}{100} \rightarrow \bar{X} = 25,7$$

Posteriormente vamos a calcular la varianza como una muestra, en este caso completaremos la tabla con dos columnas más, en las cuales se hallarán las partes de la fórmula que vamos a requerir y de esta manera se facilita un poco más el proceso. En la primera columna se resta el dato de X con el promedio y posteriormente se eleva al cuadrado, en la última columna se multiplica el dato anterior por la frecuencia absoluta.

DATOS	x	f _i	x.f _i	$(x - \bar{x})^2$	$(x - \bar{x})^2 \cdot f_i$
0 – 10	5	6	30	428,49	2570,94
10 – 20	15	19	285	114,49	2175,31
20 – 30	25	45	1125	0,49	22,05
30 – 40	35	22	770	86,49	1902,78
40 - 50	45	8	360	372,49	2979,92
Σ		100	2570	1002,45	9651

Se halla la varianza para una muestra:

$$S^2 = \frac{\sum (x - \bar{x})^2}{n - 1} \cdot f_i \rightarrow S^2 = \frac{9651}{99} \rightarrow S^2 = 97,48$$

Finalmente se calcula la desviación estándar:

$$S^2 = 97,48 \rightarrow S = \sqrt{97,38} \rightarrow S = 9,87$$

MEDIDAS DE DISPERSIÓN RELATIVAS

Coeficiente de Variación de PEARSON

El Coeficiente de Variación de PEARSON es una de las más significativas y lo podemos definir, como el cociente entre la desviación estándar y la media aritmética de una distribución, sirve para calcular el nivel de desviación de una serie de datos con respecto al valor promedio o media aritmética.

Características del Coeficiente de Variación de PEARSON:

- Se calcula como cociente entre la desviación típica y la media.
- Es un porcentaje que permite comparar el nivel de dispersión de dos muestras.

El coeficiente de variación permite determinar si los datos son homogéneos o si son heterogéneos; si el coeficiente es menor al 25% son homogéneos, si es mayor este porcentaje, entonces son heterogéneos.

$$Cv = \frac{S}{\bar{X}} \cdot 100$$

Cv= coeficiente de variación.

S= desviación estándar

\bar{X} = Media o promedio

Ejemplo 1:

En un ejemplo anterior se obtuvo un promedio de 25,7 en los datos y una desviación estándar de 9,87. Calcular el coeficiente de variación

Solución: Si se desea calcular el coeficiente de variación, entonces, se divide la desviación estándar entre X.

$$Cv = \frac{S}{\bar{X}} \rightarrow Cv = \frac{9,87}{25,7} \rightarrow Cv = 0,38$$

Si se desea dar el coeficiente de variación en porcentaje, entonces se multiplica el valor obtenido en la división, por 100.

$$Cv = 0,38 \times 100 \rightarrow Cv = 38\%$$

En este caso se encuentra que los datos son heterogéneos, es decir son muy dispersos con respecto a la media.

Ejemplo 2:

En un ejercicio en el cual se calcula la desviación estándar, se obtiene un valor igual a 8,4 y un promedio de datos de 68,75. Calcular la desviación estándar.

Solución: para calcular el coeficiente de variación, entonces, se divide la desviación estándar entre X.

$$Cv = \frac{S}{\bar{X}} \cdot 100 \rightarrow Cv = \frac{8,4}{68,75} \cdot 100 \rightarrow Cv = 0,12 \cdot 100 \rightarrow Cv = 12\%$$

En este ejercicio se obtuvo un coeficiente de variación de 12%, lo que significa que los datos se pueden considerar homogéneos.

Ejemplo 3:

En un ejercicio en el cual se calcula la desviación estándar, se obtiene un valor igual a 15,3 y un promedio de datos de 25,5. Calcular la desviación estándar.

Solución: para calcular el coeficiente de variación, entonces, se divide la desviación estándar entre X.

$$Cv = \frac{S}{\bar{X}} \cdot 100 \rightarrow Cv = \frac{15,3}{25,5} \cdot 100 \rightarrow Cv = 0,6 \cdot 100 \rightarrow Cv = 60\%$$

En este ejercicio se obtuvo un coeficiente de variación de 60%, lo que significa que los datos se pueden considerar heterogéneos.



Actividad de profundización 5:

Las calificaciones de historia del arte de 40 alumnos de una clase vienen dadas por la siguiente tabla:

Calificación	1	2	3	4	5	6	7	8	9
Fcia. Absol.	2	2	4	5	8	9	3	4	3

Construir tabla de frecuencias.

Calcular las medidas de tendencia central.

Calcular Q_3 , D_7 , P_{40}

Representar la información en un histograma.

7. La siguiente tabla muestra los rangos de edades de un grupo de personas asistentes a una clase de hidroterapia. Con la información allí suministrada, calcular:

Intervalos	[38 – 44)	[44 – 50)	[50 – 56)	[56 – 62)	[62 – 68)	[68 – 74)	[74-80)
Fi	7	8	15	25	18	9	6

Medidas de tendencia central

Cuartil Q_1

Decil D_5

Percentil P_{48}

Representar la información en un diagrama circular.

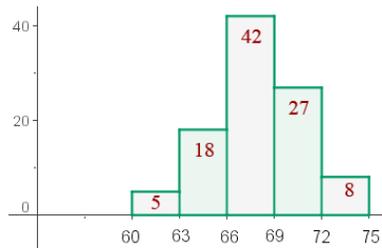
8. Calcular los CUARTILES en el siguiente conjunto de datos:

1, 9, 5, 4, 6, 2, 9, 7, 6

9. Calcular los CUARTILES de acuerdo a la información suministrada en la siguiente tabla:

Intervalo	Fi
[10, 15)	3
[15, 20)	5
[20, 25)	7
[25, 30)	4
[30, 25)	2
TOTAL	21

10. El histograma de la distribución correspondiente al peso de 100 alumnos de bachillerato es el siguiente:



¿A partir de qué valores se encuentran el 25% de los alumnos más pesados y cuánto es este valor?

11. Calcular la varianza y desviación estándar para la siguiente población: **1, 4, 7, 2, 6**.
Calcular la varianza y desviación estándar para una población.

12. La siguiente tabla representa las edades de un grupo de personas pertenecientes al grupo de canto “Los juglares”. Con esta información hallar:

EDAD EN AÑOS	Fi
10 – 15	7
15 – 20	9
20 – 25	12
25 - 30	2
TOTAL	30

Varianza.

Desviación estándar para una muestra.

Coefficiente de variación

13. La siguiente tabla representa el peso en kilogramos de un grupo de estudiantes de una escuela rural del municipio de Amalfi:

Peso (kg)	Fi
30 – 35	2
35 – 40	4
40 – 45	8
45 – 50	5
50 - 55	1
Σ	20

Con la información anterior hallar:

Medidas de tendencia central

Cuartil Q1

Decil D_8

Percentil P_{50}

Varianza

Desviación estándar para una población

Coefficiente de variación

Representar la información en un diagrama circular.

14. Se pregunta a un grupo de 20 personas sobre cuántas veces por semana va a cine. Los resultados son los siguientes:

0	1	0	1	0
1	2	1	3	1
2	0	2	1	1
3	4	2	2	3

Calcular el coeficiente de variación.

15. La siguiente tabla muestra los niveles de humedad relativa en una ciudad, durante los 10 primeros días de los meses de septiembre y octubre:

Septiembre	77	78	67	64	60	74	70	71	68	57
Octubre	78	75	73	76	67	77	75	76	67	58

Con la información anterior hallar:

Varianza

Desviación estándar para una población

Coefficiente de variación

Representar la información en un polígono de frecuencias.



AUTOEVALUACIÓN No.3

☞ Responder las preguntas 1 a 3 con base en la siguiente información: se tiene las notas de 11 alumnos en un examen de matemática: **10; 12; 09; 12; 08; 14; 12; 10; 11; 12; 08.**

1. ¿Cuál es la moda?

- A. 8 B. 10 C. 11 D. 12

2. ¿Cuál es la mediana?

- A. 9 B. 10,5 C. 11 D. 10

3. Si se elimina la nota mayor, ¿cuál es la mediana de las notas restantes?

- A. 10 B. 10,5 C. 11 D. 11,5

4. En la I.E. Eduardo Fernández Botero se realiza anualmente las olimpiadas matemáticas, las cuales se dividen en diferentes fases. En la primera fase se aplica a los estudiantes una prueba de 20 ejercicios. Se tiene a continuación los puntajes de los mejores 20 estudiantes, los cuales accederán a la fase 2: **16 18 20 21 19 19 20 18 17 18 21 16 21 19 16 16 17 18 16 18** se puede decir entonces que la moda es:

- A. Unimodal B. Bimodal C. Amodal D. Multimodal

5. En una muestra de datos 1; 2; 2; 3; 4; 4; 5; 5; 6; 7; X; el promedio es 4. Entonces, el valor de "X" es.

- A. 2 B. 3 C. 4 D. 5

5. Los puntajes obtenidos por un estudiante en unas pruebas virtuales de matemáticas son: 48, 52, 65, 38 y 54, por ende, su media aritmética es:

- A. 52,3 B. 60,6 C. 48,6 D. 51,4

Las preguntas de la 7 a la 11 se responden con base en la siguiente serie de datos, que muestran el número de horas que los estudiantes de la Institución utilizan a hacer uso de internet semanalmente, para el estudio se trabajó con los estudiantes de grado undécimo:

5, 12, 6, 3, 1, 0, 5, 10, 6, 5, 6, 2, 2, 5, 6, 10, 5, 5, 6, 2, 15, 12, 10, 5, 8, 12, 10, 14, 4, 8, 6, 12, 10

6. La población de la investigación es:

- A. Los estudiantes del grado undécimo.
- B. Los estudiantes que utilizan internet.
- C. Todos los estudiantes de la institución.
- D. Los estudiantes que les gusta internet.

7. La muestra de la investigación es:

- A. Los estudiantes del grado undécimo.
- B. Los estudiantes que utilizan internet.
- C. Todos los estudiantes de la institución.
- D. Los estudiantes que les gusta internet.

8. La moda de la tabla de frecuencia es:

- A. 6 horas
- B. 5 horas
- C. 10 horas
- D. 12 horas

9. La media de la tabla de frecuencias es:

- A. 6,9 horas
- B. 5,9 horas
- C. 7,9 horas
- D. 8,9 horas

10. La mediana de la tabla de frecuencias es:

- A. 6 horas
- B. 5 horas
- C. 10 horas
- D. 12 horas

11. La tabla de la figura muestra la distribución de frecuencias de las masas de 30 personas, escogidas al azar en un local de comida rápida, ¿Cuál es la media aritmética de las masas de estas 30 personas?

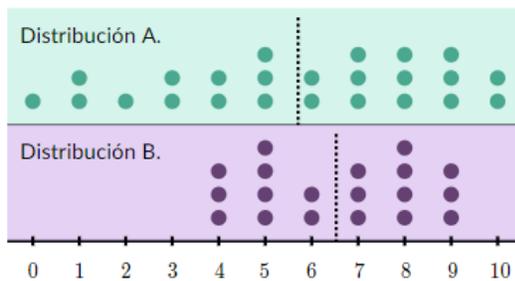
Masas (kg)	Frecuencias
(0 – 20]	2
(20 – 40]	3
(40 – 60]	3
(60 – 80]	12
(80 – 100]	8
(100 – 120]	2

- A. 66 kg
- B. 67 kg
- C. 68 kg
- D. 69 kg

12. El promedio de Andrea en 2 pruebas de Biología es 5. ¿qué nota debería sacarse en la siguiente prueba de manera que su promedio sea de 5,5 en las tres pruebas?

- A. 5 B. 5,5 C. 6,6 D. 6

13.Cuál de las distribuciones de datos que se muestran a continuación tiene la desviación estándar más grande:



- A. Distribución A B. Distribución B

De acuerdo a la información suministrada en la tabla responder las preguntas 15 a 17:

Dato	Fi
1	2
2	4
3	3
4	5
5	4

14. El valor del cuartil 2 es:

- A. 6 B. 3 C. 9 D. 3,5

15. El valor del decil 3 es:

- A. 3,5 B. 5,4 C. 6 D. 2

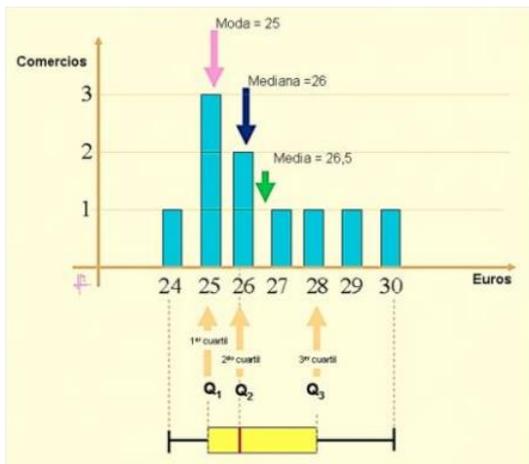
16. El valor del percentil 85 es:

- A. 5 B. 4 C. 18 D. 15,3

17. Si el coeficiente de variación es de 20, quiere decir:

- A. Los datos tienen un promedio de 0,2.
 B. Los datos se encuentran dispersos en un 20%.
 C. El estudio consideró sólo un 0,2 de los datos.
 D. No se puede determinar.

18. De acuerdo a la imagen, el valor del tercer cuartil es:



- A. 28 B. 25 C. 28,5 D. 26,5

19. Dados varios datos, necesitamos calcular el 75%. Este lo podemos encontrar utilizando el concepto de:

- A. Percentiles y cuartiles.
 B. Mediana y cuartiles.
 C. Deciles y percentiles.
 D. Cuartiles y deciles.

20. En datos agrupados, para calcular las medidas de posición se necesita de:

- A. La frecuencia relativa B. La frecuencia absoluta
 C. La frecuencia acumulada D. La marca de clase.

21. De acuerdo a la tabla, el decil 5 es:

Intervalos	x_i	f_i	h_i	F_i
[0; 2)	1	6	0,12	6
[2; 4)	3	11	0,22	17
[4; 6)	5	10	0,20	27
[6; 8)	7	6	0,12	33
[8; 10]	9	10	0,20	43
[10; 12]	11	7	0,14	50
Totales		50	1	

- A. 5,6 B. 27 C. 0,20 D. 10

22. Dados los datos: **2, 5, 3, 6, 7, 4, 9**. El decil 5 es igual a:

- A. La moda y la mediana.
- B. La mediana, el decil 10, el cuartil 3 y el percentil 5.
- C. La mediana, la moda y la media.
- D. La mediana, el decil 5, el cuartil 2 y el percentil 50

23. Pedro presenta las siguientes notas en 5 pruebas de aptitud numérica: 4, 9, 5, 4, 8. ¿cuál es el valor de la desviación típica?

- A. 2,3
- B. 2,4
- C. 2,35
- D. 2,80

24. El coeficiente de variación de los grupos A y B es 0,09 y 0,08 respectivamente, por lo tanto:

- A. El grupo B presenta mayor variabilidad en los datos.
- B. El grupo A presenta mayor variabilidad en los datos.
- C. Este valor no indica variación.
- D. Los grupos tienen la misma variabilidad.

RESPUESTAS AUTOEVALUACIÓN No.3

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D

WEBGRAFÍA

<https://www.gestiopolis.com/tipos-de-muestreo-estadistica/>

<https://gabriellebet.files.wordpress.com/2013/01/tecnicas-de-recoleccic3b3n4.pdf>

<https://www2.uned.es/111044/examenes/EJERESUS01.htm>

<file:///C:/Users/USER/Downloads/mariangelvivenzioguzman-170325060725.pdf>

<http://bibliotecavirtual.lasalleurubamba.edu.pe/Estadistica/res/pdf/estadisticadescriptivavARIABLES2.pdf>

<http://ocw.usal.es/ciencias-sociales-1/estadistica-i/contenidos/%20Tema3.pdf>

<http://www.economia.unam.mx/profesor/barajas/estadis/parte2.pdf>

<https://www.superprof.es/apuntes/escolar/matematicas/estadistica/descriptiva/ejercicios-de-cuartiles.html>

<https://es.slideshare.net/PatriciaCastillo31/medidas-de-dispersin-49617288>

<https://es.slideshare.net/IkaroVillar/medidas-de-dispersin-8415860>

<https://www.thatquiz.org/es/preview?c=grt9ncjd&s=nv3dea>

<https://es.scribd.com/doc/90344813/Pruebas-Saber-de-Medidas-de-Tendencia-Central-y-ad-de-Funciones-Grado-11c2b0>