

ESTADÍSTICA DISTRIBUCIÓN DE FRECUENCIAS Y REPRESENTACIÓN GRÁFICA.

SEGUNDA PARTE

AUTOR: PABLO VARGAS PEREIRA

FEBRERO: 2021



San Marcos

Contenido

Introducción	2
Representación gráfica de los datos	3
I. Gráfico de barras	7
II. Histogramas	9
III. Ojivas	12
IV. Polígonos de frecuencia	15
V. Gráfico circular	17
VI. Gráfico de dispersión	19
VII. Gráfico de cajas	23
Referencias bibliográficas	27



Introducción

Los gráficos, son el constructo de estadística más utilizado a nivel cotidiano, por todas las personas. Se observan en diversos campos, como las noticias, los diarios, redes sociales, entre otras. Además, son los que más se mal interpretan o se manipulan. Por esta razón es importante analizar cada uno de estos tipos de gráficos, para aprender a construirlos y a leerlos de forma correcta.

Representación gráfica de los datos

Este apartado corresponde a los gráficos, quienes junto con las tablas de frecuencias, "son la parte fundamental de las investigaciones estadísticas, informes provenientes de encuestas y notas de prensa" (Abad, Huapaya, 2009, p. 3).

De acuerdo con Abad, Huapaya (2009), la utilización de gráficos:

consiste en la utilización de puntos, líneas y figuras que sirven para mostrar magnitudes, asociadas a una escala de medición, de manera que se facilita la comparación e interpretación de los datos estadísticos, sin que necesariamente se incluyan los valores numéricos. (p. 7)

Aprender a confeccionar un gráfico es importante, pero aún más relevante que se aprenda a interpretar correctamente la información que se presenta en este. Esto se debe a que muchos software en el mercado pueden elaborar gráficas estadísticas, sin necesidad de que el investigador conozca el cómo o el porqué se realizan, mas, si es importante que este reconozca cual es la interpretación correcta con respecto a la información que se presenta, los objetivos de la investigación y la población con la que se cuenta.

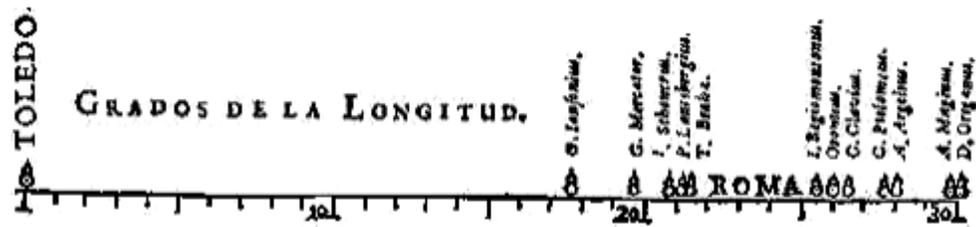
Un punto importante, lo mencionan Abad, Huapaya (2009) ya que el gráfico estadístico debe realizarse o efectuarse teniendo en cuenta la utilidad y el usuario al que le va a llegar la información. Independientemente si se realiza por parte de la persona que realiza el estudio o si se ejecuta en un software.

Con respecto a la historia del uso de los gráficos, Abad, Huapaya (2009) mencionan:

Los inicios de la elaboración gráfica de datos empezaron en el siglo XVII, mostrándose avances que prepararon el terreno para el desarrollo de los gráficos modernos.

En 1637 René Descartes publicó tres libros sobre física: Geometría, Dióptrica y Meteoros cuya introducción Discurso del Método lo hizo famoso. En Geometría creó el sistema de coordenadas cartesianas que sentó las bases del dibujo técnico y científico.

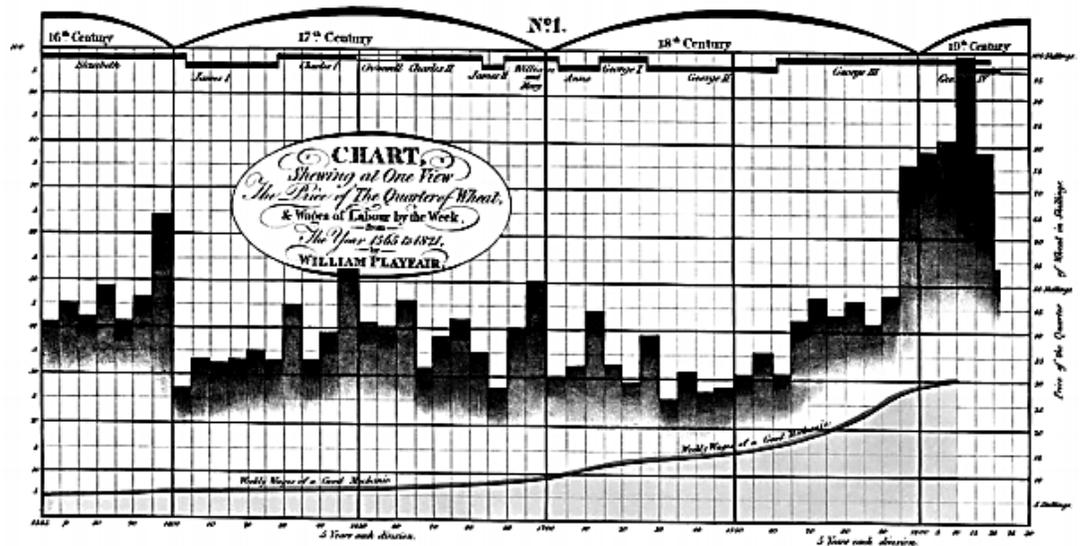
Este sistema estableció la relación entre la línea representada y la ecuación que la define. En 1644 Michael F. Van Langren produjo la que probablemente es la primera representación de datos estadísticos, al mostrar las variaciones en la determinación de la longitud entre la ciudad de Toledo (España) y Roma (Italia).



(p. 9)

Así, inició la evolución de los gráficos, hasta que en 1786, surge el primer gráfico de barras de la siguiente manera:

Playfair, expone su idea de que los gráficos permiten una comunicación más eficiente que las tablas de frecuencia. Playfair publicó el libro titulado *The Commercial and Political Atlas* (1786), el cual contiene 43 gráficos de series de tiempo y por primera vez, es usado un gráfico de barras. En 1801, utiliza el primer gráfico de sectores en su obra *Playfair's Statistical Breviary*.



(Abad, Huapaya, 2009, p. 10)

Por otra parte, cuando se realiza un gráfico es necesario tomar en consideración características relevantes, en cuanto a formato y unificación de criterios, de manera que toda persona que tenga acceso a la información pueda leerla sin importar el contexto.

Estos elementos son: título, escala, ejes, leyenda, fuente.

1. Título

Es la inscripción que se coloca en la parte superior de este, después del número de gráfico, con el propósito de dar a conocer las variables y las características que contiene. (Abad, Huapaya, 2009)

"El título expresará el contenido del gráfico en forma ordenada, clara y breve, evitando la descripción excesiva o la brevedad extrema en la descripción del contenido de la información." (Abad, Huapaya, 2009, p. 15)

Ejemplo:

Gráfico 1: Distribución de participantes en las fiestas de zapote 2019, de acuerdo con el sexo.

Note que para que se evidencie una mayor comprensión del título se debe responder a qué, cómo, dónde y cuándo.

2. Escala

"Es la línea recta segmentada que representa la escala de medición a la que corresponden los datos estadísticos del gráfico." (Abad, Huapaya, 2009, p. 20)

Estos, pueden escogerse de acuerdo con la naturaleza de los datos, sin embargo, se recomienda que abarque la mayor cantidad de datos, por ejemplo:

Si los datos que se tienen son los siguientes: 1, 2, 4, 5, 8, 10, 13, 18, 15, 20, 22, 25. Por lo tanto se presenta una escogencia de escala adecuada y otra inadecuada.

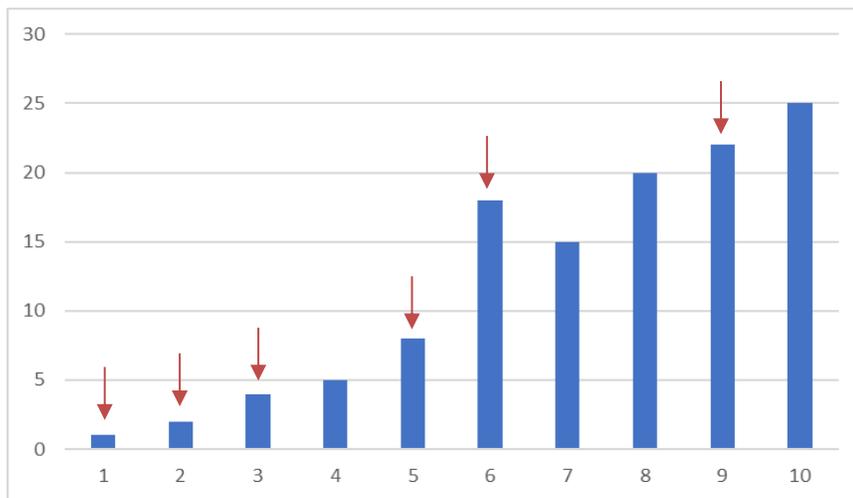
Adecuada	Una escala de dos en dos, se abordan 7 datos.
Inadecuada	Una escala de cinco en cinco, se abordan solo 4 datos.

La mayor diferencia radica en que un gráfico será más fácil de leer cuanto más datos se puedan visualizar en el eje.

Ejemplo con la escala de 5 en 5:

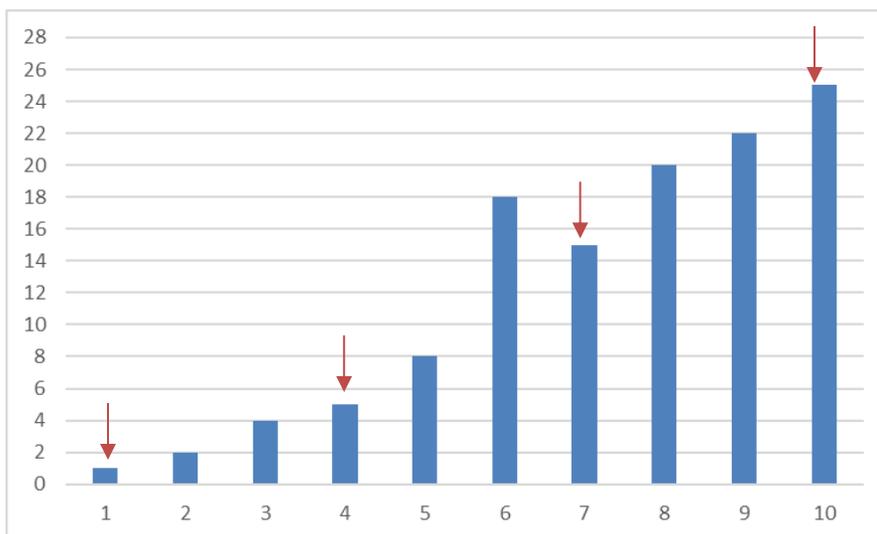
En este ejemplo se denota con una  flecha las columnas que no se denota claramente cuál es el valor del dato.





Ejemplo con la escala de 2 en 2.

En este ejemplo se denota con una  flecha las columnas que no se denota claramente cuál es el valor del dato, sin embargo, se puede inferir rápidamente cuál es el valor real.



3. Título de los ejes

Para una mayor facilidad para el lector, se recomienda que los ejes estén rotulados con un indicio de la información que contiene.

4. Leyenda

"Es la descripción de la simbología utilizada, sea ésta mediante colores, densidades de color, sombreados o tipos de línea usados para diferenciar conceptos a los que se refieren los datos estadísticos." (Abad, Huapaya, 2009, p. 20)

I. Gráfico de barras

1. Gráfico de barras

En el gráfico de barras, los datos se representan por medio de rectángulos de igual base sobre el eje de conceptos; en tanto que la longitud del otro lado corresponde al valor del dato, según la escala utilizada en el eje de valores.

No existen reglas estrictas aplicables a los gráficos de este tipo, pero como normas generales de presentación se indican las siguientes:

- El ancho de la barra debe ser uniforme para todas las barras del diagrama.
- La longitud de la barra debe ser proporcional a la cantidad que representa.
- El espacio de separación entre barras por cada concepto debe ser constante. (Abad, Huapaya, 2009, p. 25)

Se le llama gráfico de barras a los que están en posición vertical. Estos son muy utilizados en la vida cotidiana, sin embargo, es importante reconocer cuáles son las situaciones más óptimas para poder aplicarlo.

Este gráfico se utiliza, para comparar cantidades y para representar las tablas de frecuencias, sobretodo, la frecuencia absoluta, relativa o porcentual.

Ejemplos:



En este ejemplo se puede notar que se comparan los niveles educativos de una Institución, con respecto a las horas permitidas para los trabajos de medio tiempo.

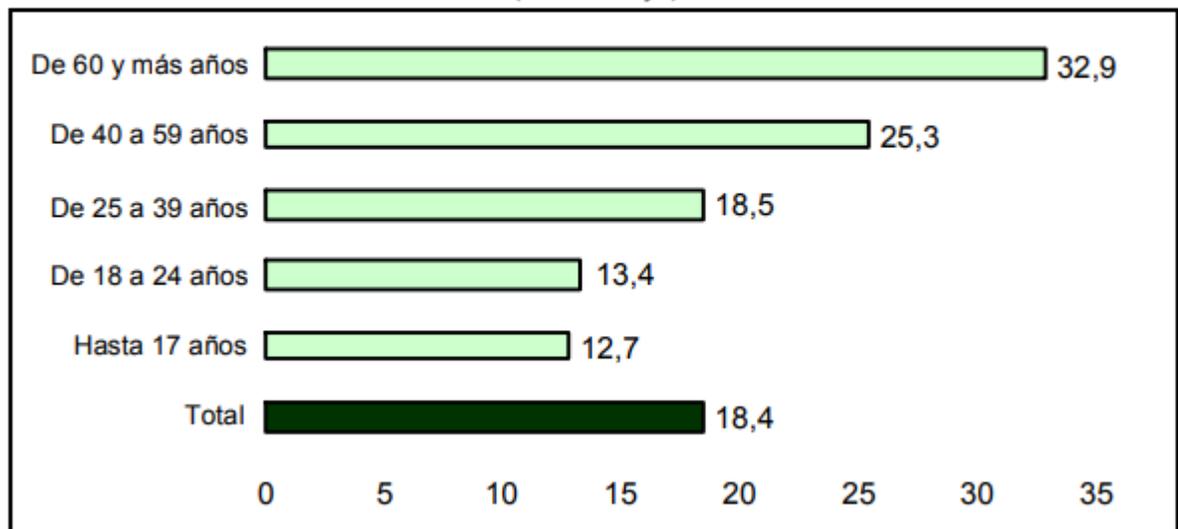
El gráfico, presenta el título, escala, título de los ejes, columnas del mismo ancho y espaciado entre barras constante.

2. Gráfico de columnas

Este gráfico tiene el mismo estilo y forma de construcción que la anterior, sin embargo, las columnas se colocan de forma horizontal.

"Representan valores discretos, a base de trazos horizontales, aislados unos de otro. Se utilizan cuando los textos correspondientes a cada categoría son muy extenso." (Abad, Huapaya, 2009, p. 25)

PERÚ: POBLACIÓN QUE APRENDIÓ QUECHUA, AYMARA U OTRA LENGUA NATIVA EN SU NIÑEZ, SEGÚN GRUPO DE EDAD, 2006
(Porcentaje)



Fuente: Abad, Huapaya (2009). Guía para la presentación de gráficos estadísticos.

Muchos autores prefieren, el uso de este tipo de gráficos ya que son más fáciles de leer, ya que las personas leen siempre de izquierda a derecha.

El gráfico anterior relaciona la cantidad de personas que aprendieron una lengua nativa, de acuerdo con la edad. Se puede notar además, que en el gráfico se presenta un título, la escala, esta vez en el eje horizontal, título de los ejes, barras del mismo ancho y espacio entre barras constante.

II. Histogramas

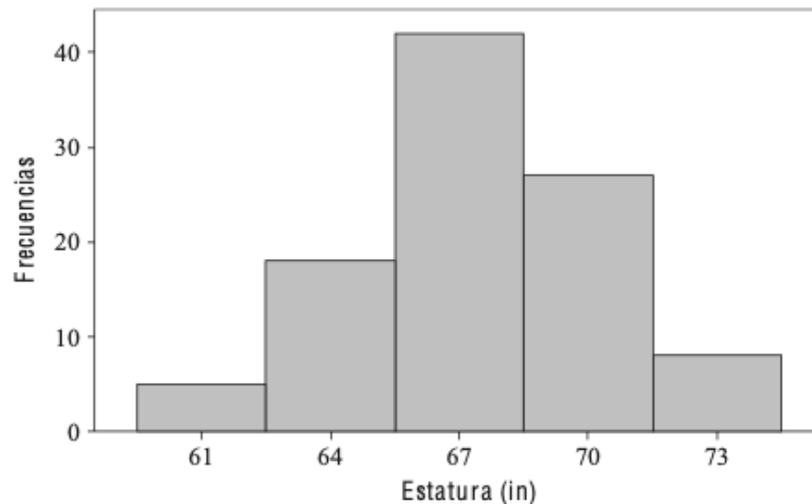
Una representación de datos es la realización de un histograma, la cual es “una gráfica con barras de la misma anchura, dibujadas una junto a la otra (sin espacios entre sí)” (Triola, 2013, p. 55). Como punto importante en esta interpretación es la utilización de datos continuos, ya que no se debe tener pérdida de información entre los datos. A manera de explicación factible se tiene el siguiente ejemplo:

- Entre el número 3 y el 4 se tiene un infinita cantidad de números entre los mismos.

El histograma consiste en un conjunto de rectángulos, que de acuerdo con Spiegel y Stephens (2009) se construyen con:

- Las bases sobre un eje horizontal (el eje X), las cuales deben coincidir los centros con las marcas de clase de longitudes iguales a la amplitud del intervalo de clase.
- Las áreas proporcionales a las frecuencias de clase.

A manera de ejemplo se tiene el siguiente histograma, ya que éste cumple con las características antes mencionadas:



Histograma. Fuente: Spiegel y Stephens (2009)

Con respecto al histograma anterior se logra tener un panorama más claro de cómo se debe realizar.

La construcción del mismo inicia:

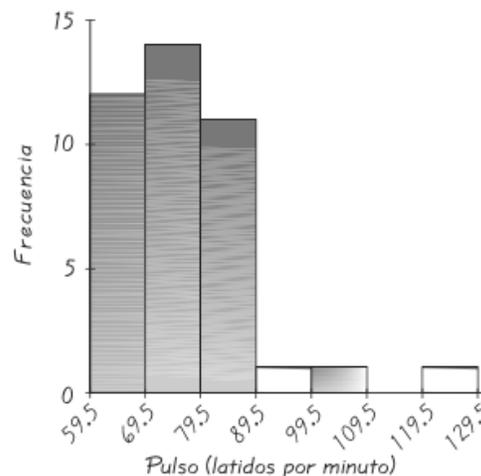
- Identificando los datos que el investigador posee, dado que esta representación se adecua a datos continuos, tal y como se muestra en la figura anterior.
- Luego de esto se tiene que reconocer las variables a considerar en los ejes de coordenadas (X o Y) , cabe destacar que, en esta representación analítica los datos deben ser cuantitativos.

Con respecto al histograma anterior, se visualizan la variable de frecuencia y de estatura de manera cuantitativa, es decir, se logra una relación entre la representación cuantitativa de la frecuencia y la estatura de las personas.

- Posteriormente se da una división analítica, debido que la construcción del gráfico se puede realiar por medio del punto medio o por los límites de clase.

Límites de clase:

Primeramente se mantiene la construcción del histograma por medio de los límites de clase como se muestra en la siguiente figura. Esta opción es la más adecuada en la realización de un histogama, ya que brinda una mayor cantidad de información, en otras palabras, brinda el dato de donde inicia y termina una clase así como la barra que la persona observa.



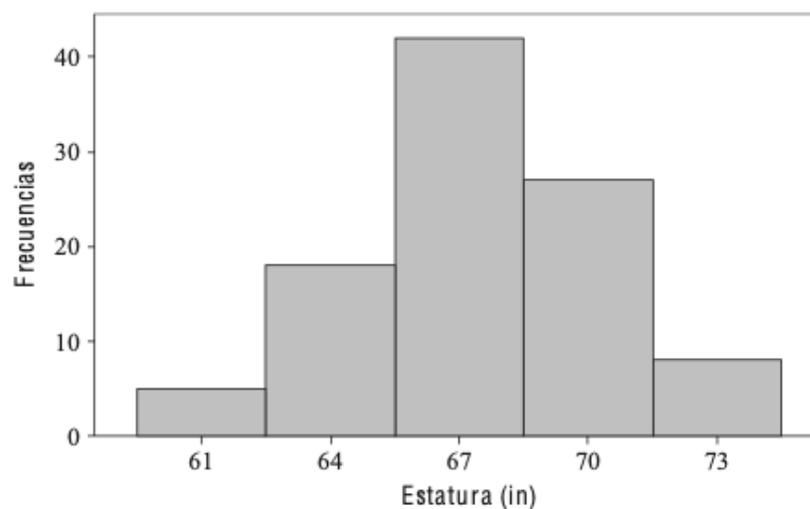
Histograma. Fuente: Triola (2013) Estadística

Como punto importante se observa en la figura anterior, que el histograma utiliza datos continuos, característica fundamental de esta representación. Además de esto, se reconoce una ventaja al momento de leerlo, dado que se observa con mayor facilidad la tendencia de los datos.

Marca de clase

En segundo punto se explica la utilización del punto medio en la construcción del histograma, cabe destacar que este método no es muy utilizado, debido que no brinda gran información como el anterior, sin embargo, es muy útil para realizar polígonos de frecuencia.

Esta alternativa se relaciona con la siguiente figura, dado que en las columnas se logra una coincidencia con los centros y las marcas de clase; esto con el fin de poseer longitudes iguales a la amplitud del intervalo de clase.



Histograma. Fuente: Spiegel y Stephens (2009)

Por último, se tiene el análisis del histograma que se desarrolla por medio del ancho de los rectángulos, el cual representa el intervalo del rango de los datos y la altura que simboliza la cantidad de veces que aparece ese valor dentro de ese intervalo. Además de esto, el modelo de las alturas de los respectivos rectángulos identifican la distribución de los valores de los datos.

Por ejemplo en el caso anterior, se analiza que la estatura de (67) tiene una mayor frecuencia con respecto a lo antes mencionado. Asimismo, se logra observar una alta distribución de los valores, ya que los rectángulos tienen un cambio considerable de tamaño.



III. Ojivas

Antes de iniciar con este gráfico es necesario conocer una nueva definición, la cual puede ser unida a la tabla de frecuencia. Esta se llama frecuencia porcentual acumulada.

Frecuencia porcentual acumulada

Esta es similar a la frecuencia acumulada, solamente que suma cada frecuencia porcentual con todas las anteriores. Por ejemplo:

Hijos	Frecuencia Absoluta	Frecuencia Acumulada (Abs)	Frecuencia Relativa	Frecuencia Porcentual	Frecuencia Acumulada (%)
2	3	3	$\frac{3}{9} = 0,33$	33%	33%
3	4	3+4=7	$\frac{4}{9} = 0,44$	44%	33+44=77%
4	2	3+4+2=9	$\frac{2}{9} = 0,23$	23%	33+44+23=100%
Total	9		1	100%	

Ojiva

De acuerdo con Quintana (2019) "es un gráfico que muestra las frecuencias acumuladas menores de cada frontera inferior o superior de clase." (p. 4)

Con respecto a este gráfico, se presenta las siguientes consideraciones:

- Utiliza las frecuencias acumuladas y las frecuencias porcentuales acumuladas.
- Se trabaja con variables continuas.
- Se construye con respecto al límite inferior de cada clase y el último de los superiores.

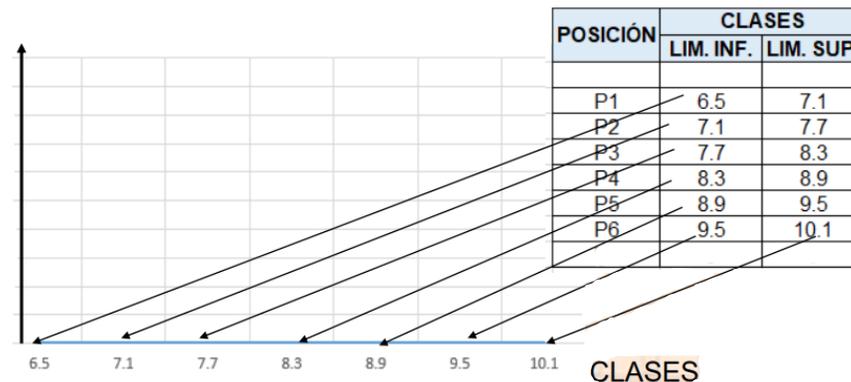
A continuación, con base en la enumeración de Quintana (2019), se explicarán los pasos para la construcción, ya que con esto se comprende de una mejor manera el tipo de gráfico.

Ejemplo:

En el curso de Estadística, en la U San Marcos, se tienen 22 alumnos, y las notas finales se evidencian en la siguiente tabla de frecuencias.

Notas	Frecuencia Absoluta	Frecuencia Acumulada
[6.5 , 7.1 [1	1
[7.1 , 7.7 [2	3
[7.7 , 8.3 [4	7
[8.3 , 8.9 [7	14
[8.9 , 9.5 [5	19
[9.5 , 10.1 [3	22
Total	22	

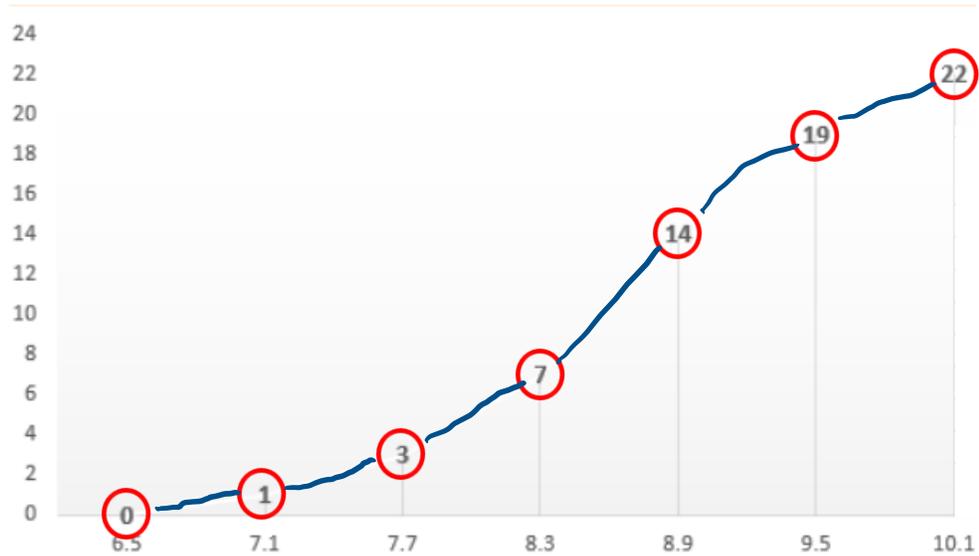
Paso 1: Colocar los límites inferiores en el horizontal (eje x) y el último de los límites superiores.



Quintana (2019). Construcción de una ojiva.

Paso 2: En el eje vertical (eje y) se colocan las frecuencias acumuladas. Se puede hacer una escala adecuada para esto.

Paso 3: Colocar los valores de cada una de las frecuencias acumuladas.



Quintana (2019). Construcción de una ojiva.

Paso 4: Se traza una línea continua entre las ojivas de manera que se represente también la continuidad de la variable.

Paso 5: se colocan los requerimientos de todo gráfico, título, escala y título de ejes.

Paso 6: al mismo nivel que las ojivas pero al lado derecho, paralelo al eje y se colocan los porcentajes de la frecuencia porcentual acumulada, tal y como se observa en la imagen.



Quintana (2019). Construcción de una ojiva.

La Interpretación de un gráfico de ojiva es muy simple, una vez comprendido el proceso para la construcción, por ejemplo de acuerdo con el gráfico anterior se puede deducir que 14 estudiantes tuvieron una nota, en el curso de estadística, menor a 8.9. Lo que representa un 63,6%.

IV. Polígonos de frecuencia

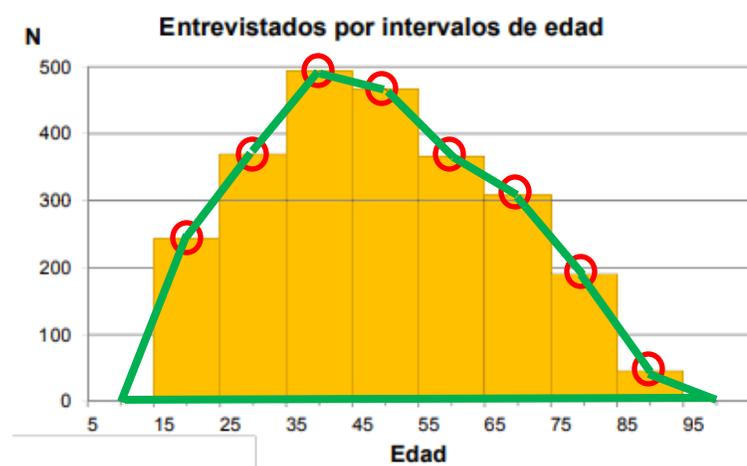
Una vez comprendido el histograma, se deriva de este, el polígono de frecuencias. Este se utiliza, de acuerdo con De la Puente (2015), para "el estudio, comprensión y descripción de las variables numéricas, conjuntamente con los estadísticos de tendencia central, dispersión y forma." (p. 44)

Para comprender, de forma indónea esta representación, es necesario recordar el concepto de marca de clase.

$$M_c = \frac{\text{límite inferior} + \text{límite superior}}{2}$$

Este concepto corresponde al punto medio de cada una de las clases, por lo que, es utilizado por el polígono de frecuencias. El cual se forma por "la línea quebrada que resulta de unir los puntos medios o marcas de clase de las caras opuestas a la base de cada intervalo o rectángulo." (De la Puente, 2015, p. 44)

Al nombrarse como polígono debe cumplir con la definición de este: "Un polígono es una figura geométrica plana cerrada, formada completamente por segmentos de línea que se intersecan en sus extremos, formando vértices." (Palma, 2020, p. 1) Es decir, el polígono de frecuencias debe ser una figura cerrada, por lo que de De la Puente (2015) menciona que debe iniciar y culminar en la frecuencia cero, de manera que se pueda crear la figura cerrada, tal y como se puede observar en la siguiente imagen.

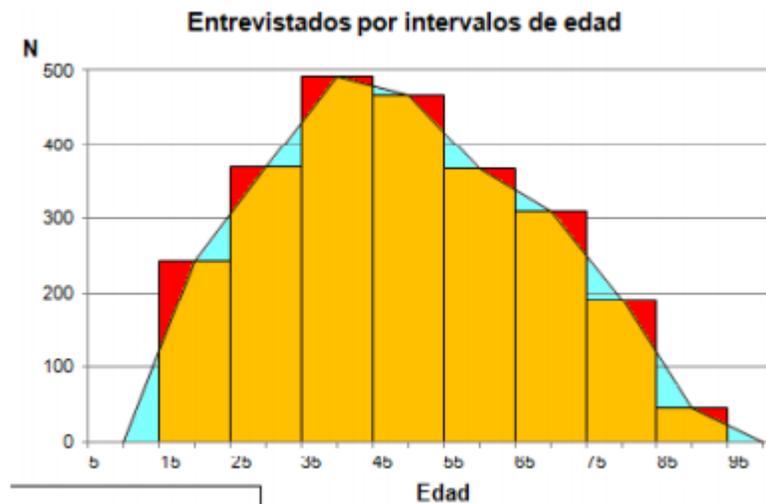


Fuente: De la Puente (2015). Diagrama de barras, histograma, polígono de frecuencias y gráfico de dispersión

En el gráfico se puede observar el histograma, representado con color rojo; las marcas de clase, representada por los círculos color rojo; el polígono, representado por el color verde y las características típicas de un gráfico: título, título de los ejes y escala.

La analogía entre el histograma y el polígono de frecuencias puede realizarse debido a que se pueden homologar los espacios vacíos entre los rectángulos y los lados del polígono, con los espacios sobrantes entre los lados del polígono y el rectángulo.

De la Puente lo evidencia por medio de la siguiente representación, donde la suma de todas las áreas de los triángulos color celeste es igual a la suma del área de todos los triángulos color rojo.



Fuente: De la Puente (2015). Diagrama de barras, histograma, polígono de frecuencias y gráfico de dispersión

Este gráfico es muy utilizado en las áreas de las ciencias económicas y en las ciencias sociales, pues con esto se llevan a cabo comparaciones de resultados entre diversas categorías, con base en las variables continuas.

Normalmente, este tipo de gráfico se utiliza para realizar una comparación cruzada de diversas variables dentro del mismo dibujo.

Dentro de las características de estos se encuentra, que no se muestran las frecuencias acumuladas, salvo que se realice un polígono de frecuencias acumuladas. Además, se prefiere para el tratamiento de datos que sean cuantitativos con variable continua.

Por último, el área bajo la curva representa al 100% de los datos, pues está diseñado para correlacionarse con el área de los rectángulos que comprenden al histograma.

V. Gráfico circular

El gráfico circular permite “ver la distribución interna de los datos que representan un hecho, en forma de porcentajes sobre un total” (García, 2009, p.37) .

De acuerdo con la definición anterior se identifica una ventaja en esta representación, ya que el análisis se realiza con mayor factibilidad al observar el comportamiento de uno de los datos con respecto al total.

Ejemplo:

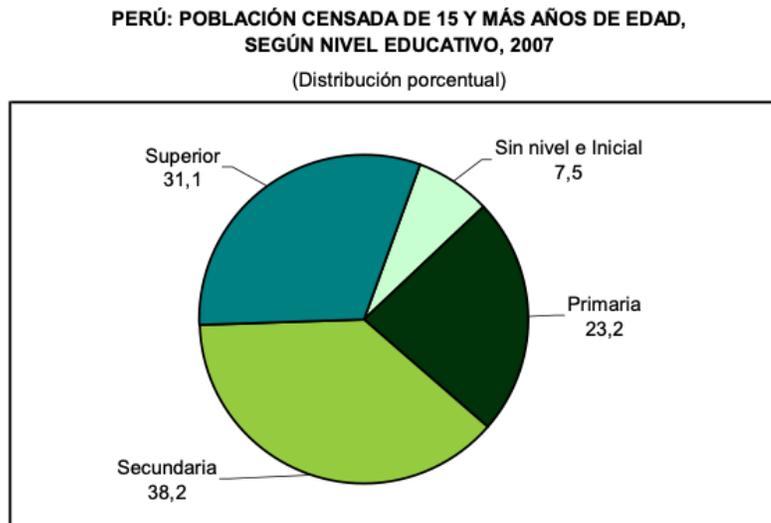


Gráfico circular. Fuente: (García, 2009, p.37)

Se destaca el gráfico anterior, dado que se reconoce un análisis sencillo. Además, se reconoce el uso estadístico en los porcentajes de la población en diferentes categorías según su nivel educativo.

Características:

Este gráfico posee las siguientes características de acuerdo con el Ministerio de Trabajo y Seguridad Social (2013):

- Más sencillo de representar
- Factibilidad en interpretación
- Útil en la composición o participación relativa (porcentual) de una característica o hecho
- Muestra la distribución de los datos respecto a una sola variable

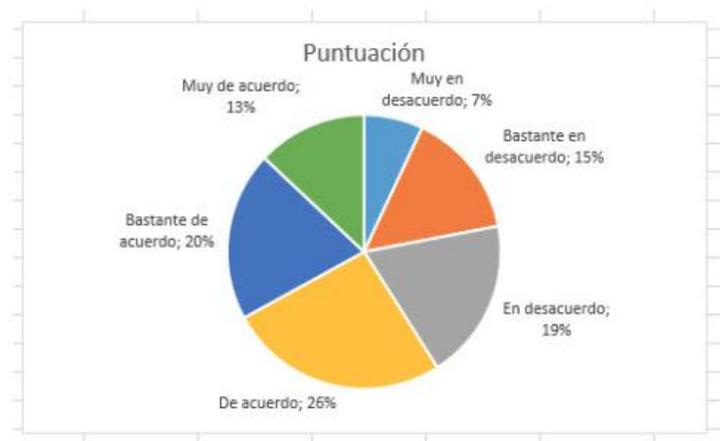
El gráfico circular o de pastel, se utiliza para la representación de porcentajes o

porciones. Como punto importante se menciona que, en esta interpretación se incorporan más de cuatro elementos en comparación; esto porque con menor cantidad de elementos no se logra concatenar un buen análisis comparativo de los datos.

Forma correcta de construirlos

A manera de explicación, se brindan ejemplos para lograr entender cuando se debe utilizar esta representación de datos.

Ejemplo correcto:



Fuente: <https://ayudaexcel.com/grafico-circular-no-mejor-opcion/>

Con respecto al gráfico anterior, se observa la clasificación de una porción con respecto a las demás, por lo que se puede concluir que los datos están correctamente adecuados al gráfico circular. Además, de que se tiene más de 4 datos en un medio comparativo.

Ejemplo incorrecto:



Fuente: <https://www.elartedepresentar.com/2011/11/4-formas-de-mentir-con-graficos->

de-datos-en-una-presentacion/

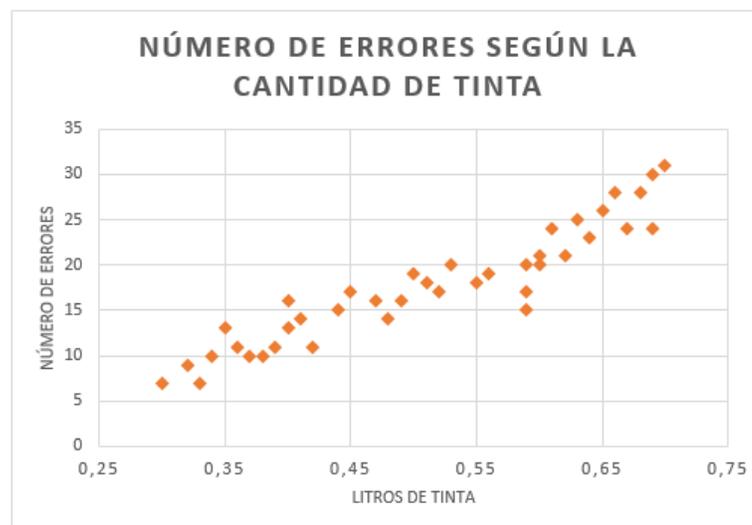
Con respecto al gráfico anterior, se observa una mala distribución de las regiones en el gráfico circular, ya que el apartado de otros posee un mayor porcentaje, sin embargo, tiene una menor región con respecto a la de Apple, que no mantiene un porcentaje mayor.

VI. Gráfico de dispersión

De acuerdo con Abad, Huapaya (2009)

Este tipo de gráfico es de mucha utilidad en los estudios que tienen muchas observaciones, y muestra la asociación o relación que existe entre dos variables unidas a un mismo fenómeno. En este gráfico cada observación debe aparecer situada en el plano como un simple punto, cuya posición es fijada en relación a cada una de las variables. Las escalas van precisadas en cada uno de los ejes rectangulares del dibujo, una para cada variable; asentándose mediante el mismo principio todas las demás observaciones que constituyen los datos del problema. (p.47)

Este gráfico suele ser difícil de interpretar, puesto que si se tienen muchos datos, resulta complejo visualizar las coordenadas exactas de dato que se representa. Sin embargo, este gráfico es más útil para realizar una comparación de la tendencia que presenta un grupo de datos y así con esto realizar conclusiones sobre el comportamiento de estos.



Fuente: Betancourt (2016) Cómo hacer un diagrama de dispersión.

En este gráfico se puede notar gran cantidad de puntos, los cuales corresponden a los litros de tinta con respecto al número de errores en la impresión, sin embargo, es muy

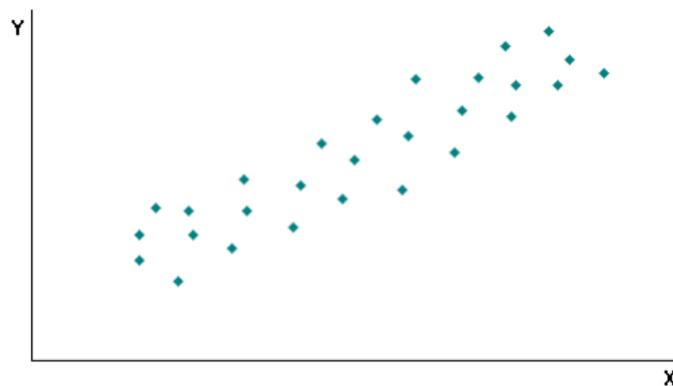
poco visible cuál es el valor exacto de cada uno de los datos.

Lo que se puede notar, de forma más accesible es que los datos tienen un comportamiento creciente, es decir, se puede deducir que a mayor cantidad de litros de tinta se presenten mayor es la cantidad de número de errores presentes en las impresiones.

La interpretación anterior se fortalece gracias a la defición de los siguientes conceptos

Correlación positiva

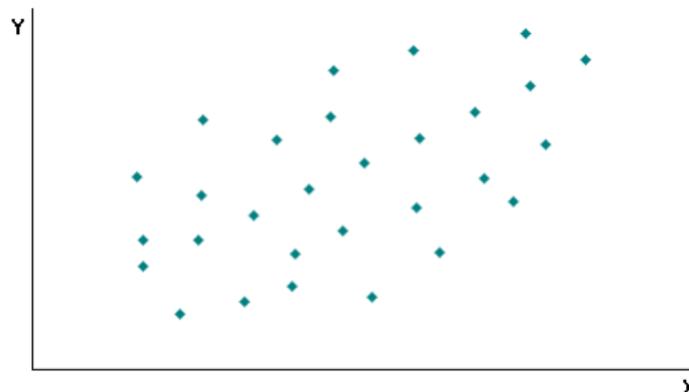
"A un crecimiento de X (causa) corresponde un crecimiento de Y (efecto). Controlando la evolución de los valores de X , quedan controlados los valores de Y ." (Domenech, sf, p. 9)



Fuente: Domenech. sf. Diagrama de correlación - dispersión.

Correlación positiva débil

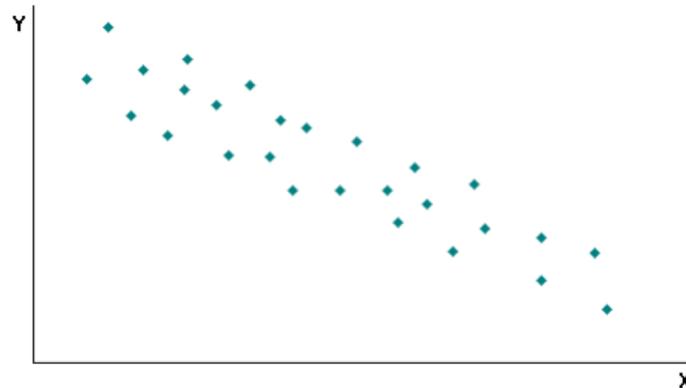
A un crecimiento de X se observa una tendencia a crecer de Y , pero se presume que existen otras causas de dependencia. (Domenech, sf, p. 9)



Fuente: Domenech. sf. Diagrama de correlación - dispersión.

Correlación negativa

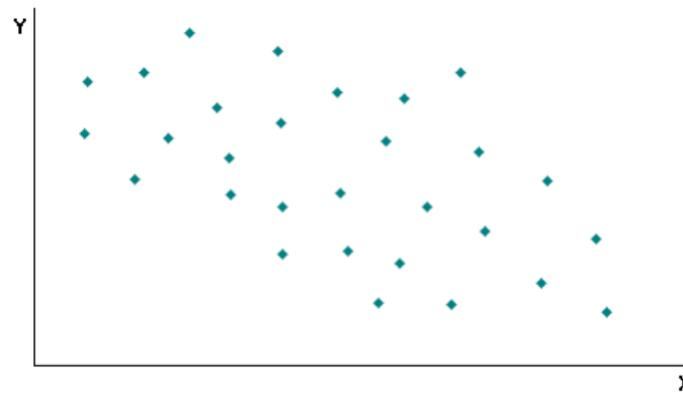
" A un crecimiento de X se observa una tendencia a disminuir de Y
Correlación negativa débil." (Domenech, sf, p. 10)



Fuente: Domenech. sf. Diagrama de correlación - dispersión.

Correlación negativa débil

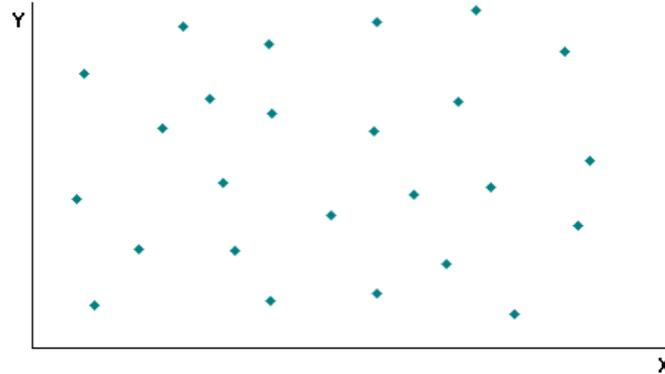
" A un crecimiento de X se observa una tendencia a disminuir de Y, pero se presume que existen otras causas de dependencia" (Domenech, sf, p. 10)



Fuente: Domenech. sf. Diagrama de correlación - dispersión.

Correlación nula

Esta última se presenta cuando no se evidencia una relación entre ambas variables.



Fuente: Domenech. sf. Diagrama de correlación - dispersión.

Ejercicio 1¹

La Dirección de una mina está preocupada por el alto porcentaje de indisponibilidad de sus máquinas cargadoras. Encarga al Jefe de Mantenimiento que analice si está influyendo la antigüedad de dichas máquinas en su porcentaje de indisponibilidad. Para ello, recoge la información de la fecha de compra y del porcentaje de indisponibilidad de cada máquina y la traslada a la siguiente tabla.

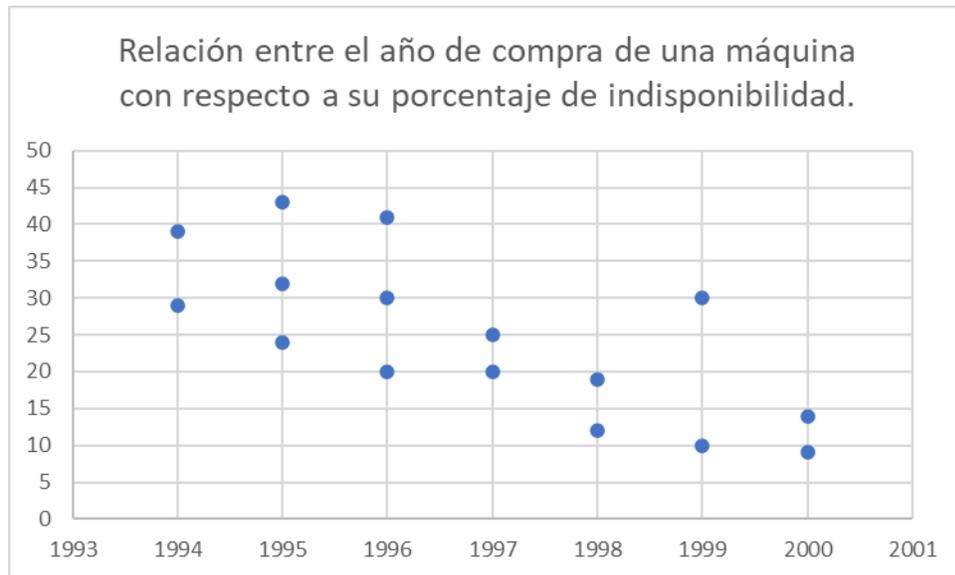
Máquina	Fecha de compra	% Indisponibilidad
C-0037	1994	29
C-0038	1994	39
C-0039	1995	24
C-0040	1995	32
C-0041	1995	43
C-0042	1996	20
C-0043	1996	41
C-0044	1996	30
C-0045	1997	20
C-0046	1997	25
C-0047	1998	12
C-0048	1998	19
C-0049	1999	10
C-0050	1999	30
C-0051	2000	9
C-0052	2000	14

Utilice un gráfico de dispersión para colaborar con la resolución de la situación problema.

¹ Tomado de: Domenech. sf. Diagrama de correlación - dispersión.

Solución

Con respecto a los datos presentados en la tabla y utilizando el software Excel se tiene el siguiente gráfico que representa los datos planteados.



Fuente: datos del Ejercicio 1

Se puede notar una tendencia decreciente en los puntos planteados, por lo que se deduce que la correlación es negativa, esto indica que mientras una variable aumenta la otra disminuye.

Por lo tanto se puede determinar que a más reciente el modelo de la máquina que se compra, menor es el porcentaje de indisponibilidad de esta. O lo que es análogo, a más antigua sea la adquisición de la máquina mayor es el porcentaje de indisponibilidad.

VII. Gráfico de cajas

Este tipo de gráfico es muy utilizado para observar la variabilidad de los datos, sin embargo antes de comenzar con el estudio de este, es necesaria la definición de ciertos conceptos de gran importancia.

Mínimo

Es el valor menor, con respecto a todo el conjunto de datos. Se representa como: *mín*

Máximo

Es el valor mayor, con respecto a todo el conjunto de datos. Se representa como: *máx*



Cuartil 1

Corresponde al punto equivalente al 25% de los datos, por lo que si n es el número total de datos, entonces:

$$Q_1 = n \cdot \frac{1}{4}$$

Si el número de datos es par este resultado indica que se debe elegir el número promedio entre el resultado de Q_1 y el dato siguiente.

Si el número de datos es impar este resultado indica que se debe elegir el número que se encuentra en la posición Q_1 .

Cuartil 2

Corresponde al punto equivalente al 50% de los datos, por lo que si n es el número total de datos, entonces:

$$Q_2 = n \cdot \frac{1}{2}$$

Nuevamente, si el número de datos es par, este resultado indica que se debe elegir el número promedio entre el resultado de Q_2 y el dato siguiente.

Si el número de datos es impar este resultado indica que se debe elegir el número que se encuentra en la posición Q_2 .

Cuartil 3

Corresponde al punto equivalente al 75% de los datos, por lo que si n es el número total de datos, entonces:

$$Q_3 = n \cdot \frac{3}{4}$$

Si el número de datos es par este resultado indica que se debe elegir el número promedio entre el resultado de Q_3 y el dato siguiente.

Si el número de datos es impar este resultado indica que se debe elegir el número que se encuentra en la posición Q_3 .

Cajas

Las cajas, son rectángulos delimitados por el cuartil 1, 2 y 3.

Recorrido Intercuartílico

El recorrido intercuartílico está dado por la diferencia entre el cuartil 3 y el cuartil 1. Gracias a esta medida se pueden analizar los datos sin la influencia de los datos extremos.

$$R_I = Q_3 - Q_1$$

Ejemplo de construcción

La siguiente tabla representa las edades de 20 personas:

Persona	Edad	Persona	Edad	Persona	Edad	Persona	Edad
1	36	6	20	11	39	16	40
2	25	7	36	12	24	17	33
3	37	8	45	13	29	18	24
4	24	9	31	14	23	19	34
5	39	10	31	15	41	20	40

Para iniciar, los datos deben estar ordenados en forma ascendente.

20, 23, 24, 24, 24, 25, 29, 31, 31, 33, 34, 36, 36, 37, 39, 39, 40, 40, 41, 45

Claramente el mínimo es 20 y el máximo es 45

Cálculo de cuartiles:

Cuartil 1: (Q_1)

Como $n = 20$ entonces $\frac{n}{4} = \frac{20}{4} = 5$, lo que significa que el primer cuartil inicia en el promedio entre el dato en la posición 5 y la posición 6, esto por ser un número de datos par.

$$Q_1 = \frac{24 + 25}{2} = 24,5$$

Cuartil 2: (Q_2)

Como $n = 20$ entonces $\frac{n}{2} = \frac{20}{2} = 10$, lo que significa que el segundo cuartil inicia en el promedio entre el dato en la posición 10 y 11, por ser una cantidad par de datos

$$Q_2 = \frac{33 + 34}{2} = 33,5$$

Cuartil 3: (Q_3)

Como $n = 20$ entonces $\frac{3n}{4} = \frac{3 \cdot 20}{4} = 15$, lo que significa que el tercer cuartil inicia en el promedio entre el dato en la posición 15 y la 16, esto por ser una cantidad par de datos.

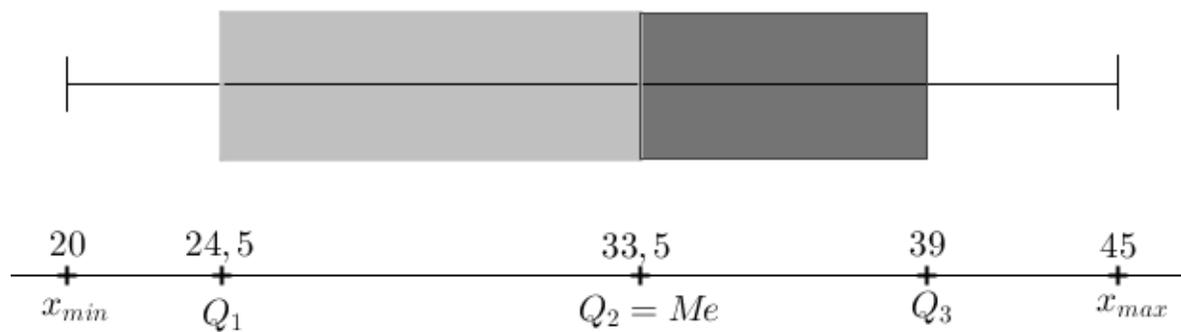
$$Q_3 = \frac{39 + 39}{2} = 39$$

Recorrido intercuartílico:

$$R_I = Q_3 - Q_1$$

$$R_I = 39 - 24,5 = 14,5$$

Dibujo del gráfico



Referencias bibliográficas

- Abad, P. Huapaya, E. (2009). Guía Para La Presentación De Gráficos Estadísticos. Centro de Investigación y Desarrollo. Lima. Perú.
- Betancourt, D. F. (26 de julio de 2016). Cómo hacer un diagrama de dispersión: Ejemplo en calidad. <https://www.ingenioempresa.com/diagrama-de-dispersion>
- De la Puente, C. (2015). Diagrama de barras, histograma, polígono de frecuencias y gráfico de dispersión. <https://eprints.ucm.es/35985/13/GR%C3%81FICOS.pdf>
- Palma, P. (2020). Guía de aprendizaje: polígonos. <https://www.colegioconcepcionsanpedro.cl/wp-content/uploads/2020/03/Guia-N%C2%BA1-7%C2%BAB-Geo-2020.pdf>
- Quintana, O. (2019). Construcción de una ojiva. https://www.uaeh.edu.mx/docencia/P_Presentaciones/b_sahagun/2019/oql-estadistica.pdf
- Triola, M. (2013). Estadística. 11° Ed. Pearson. México
- Spiegel, M. Stephens, L. (2009). Estadística. 4° Ed. Mc Graw Hill. México. <http://up-rid2.up.ac.pa:8080/xmlui/bitstream/handle/123456789/1476/Estad%C3%ADstica.%20Serie%20Schaum-%20Murray%20R.%20Spiegel.pdf?sequence=1&isAllowed=y>



www.usanmarcos.ac.cr

San José, Costa Rica