

TRATAMIENTO DE LA INFORMACIÓN.

Best (1969) señala que los datos recogidos mediante algunos de los medios, procedimientos o técnicas indicados precedentemente, deben ser elaborados y clasificados con arreglo a ciertos criterios de sistematización, para proceder luego al recuento de los mismos conforme al sistema más adecuado o factible. Se procede luego al tratamiento estadístico matemático de los datos.

El trabajo con los datos.

Según Encinas (1993), a quien sigo en este apartado, el trabajo con los datos comprende: la codificación que consiste en ubicar cada dato dentro de una determinada categoría; la tabulación que consiste en disponer los datos, organizada y sistemáticamente, en tablas, cuadros o gráficos, según el diseño propuesto para los efectos del cómputo, análisis e interpretación; y, el tratamiento *estadístico de los datos*.

a. La codificación.

Es una técnica mediante la cual se categorizan los datos obtenidos. Prácticamente, consiste en sustituir los datos obtenidos, por símbolos -generalmente numéricos- con el objeto de agruparlos en un número limitado de categorías, las cuales se seleccionan a partir de algún criterio de clasificación que puede estar contenido en el enunciado del problema o de la hipótesis.

La técnica de codificación supone un juicio valorativo sobre los datos que el investigador debe realizar teniendo como base un marco común de referencia respecto del material que va a codificar, para tener el mismo criterio en todos los que sean similares.

"El valor de la categorización de los datos depende totalmente de la exactitud de las categorías utilizadas. Estas deben ser bien definidas desde el punto de vista conceptual y deben ser significativas para los propósitos de la investigación" (Selltiz, 1970).

La exactitud con que pueden ser definidas las categorías depende de una serie de condiciones tales como: el tipo de problema de investigación; la naturaleza de las situaciones a las que se refieren los datos, y la naturaleza misma de los datos, ya sean elaborados o no elaborados. En términos generales, se puede afirmar que cuanto más estructurado se halla el material a codificar y más simples las categorías utilizadas, mayor será la fiabilidad de la codificación.

Cuando los datos provienen de un instrumento estructurado, por ejemplo un cuestionario, cada índice o grupo de preguntas determina directamente la categoría correspondiente.

Los datos cuantificados no ofrecen mayores dificultades en el proceso de elaboración: en cambio, los datos cualitativos exigen un tratamiento según su propia naturaleza, a fin de lograr una elaboración sistemática y objetiva.

Algunas dificultades que surgen en la codificación son consecuencia de procedimientos inadecuados en la recogida de datos; por ejemplo, preguntas mal formuladas. También se deben a la falta de claridad y precisión en la definición de las categorías propuestas. Por esta razón, es conveniente definir las categorías en términos de índices aplicables a los datos.

b. Matriz de datos.

Es un modo de ordenar los datos en un esquema o tabla de doble entrada.

En la elaboración de la matriz de datos se consideran las unidades de análisis - sujetos u objetos motivo de estudio-, las variables, que corresponden a las cualidades o características que se investigan, y los valores, que son los resultados que se obtienen respecto de cada unidad de análisis, en relación con la variable que se investiga.

La estructura de la matriz de datos corresponde a una tabla en la cual las unidades de análisis se disponen en columna (vertical) y las variables en línea horizontal, de suerte que en el cruce de ambas direcciones se anotan los valores.

La matriz de datos permite reducir, ordenar y objetivar los datos recogidos.

La matriz de datos puede y debe elaborarse durante el proceso de planeamiento de la investigación a fin de precisar los datos que se requieran y las operaciones que se realizarán con ellos.

c. La tabulación.

Respecto de la tabulación hay que tener en cuenta lo siguiente:

1) Tabulación lineal y cruzada.

La tabulación es lineal cuando se procesa el número de casos correspondientes a una sola categoría; por ejemplo, el número de profesores de educación secundaria según el grado o título profesional que poseen; y tabulación cruzada cuando se refiere al número de casos que ocurren conjuntamente en dos o más categorías.- por ejemplo, el número de profesores de educación secundaria según el grado o título profesional y los años de servicio en la docencia.

2) Tabulación manual y mecánica.

Según la forma en que se realiza la tabulación, ésta puede ser, manual o mecánica.

La tabulación manual es posible cuando se procesa un número manejable de casos y la serie de categorías es pequeña. En caso contrario, o cuando se precisa de tabulación cruzada, se recurre al procesamiento mecánico que hoy recibe el nombre de informatizado o computarizado.

3) Escalas de medida.

En el estudio de los hechos es necesario establecer diferencias respecto a su naturaleza y a sus características; es decir, establecer distinciones cualitativas. Pero, como casi siempre se requiere hacer mayores precisiones que impliquen diferencias de grado o nivel, se utilizan las distinciones cuantitativas, expresadas mediante series numéricas teniendo en cuenta la correspondencia entre las propiedades del sistema numérico y los elementos y relaciones de carácter empírico. Esta es la razón que fundamenta la construcción de escalas de medida.

Técnicamente, la elaboración de escalas válidas, confiables y precisas requiere, en primer lugar, describir el factor que se desea evaluar y, en segundo término, identificar las características que guardan relación con el factor descrito.

Las que se aplican en las investigaciones educacionales son tres: nominales, ordinales y de intervalo. La escala de razones o de cocientes sólo se utiliza en las ciencias exactas.

a) Escala nominal o de clasificación.

Es aquella que ordena a los individuos o a los objetos en categorías diferentes, teniendo en cuenta aspectos cualitativos; de ese modo, quedan agrupados según una igualdad o equivalencia referida a una determinada característica. Para identificar a los miembros de una categoría se les asigna símbolos, generalmente, numéricos.

Ejemplo: se puede construir una escala nominal a partir de una relación de alumnos que pertenecen a distintos tipos de planteles, asignando un número diferente a cada con. junto de alumnos de cada tipo de plantel. Al conjunto de alumnos de los colegios estatales se les asigna el n° 1, al de los colegios particulares, el n° 2, al de los colegios cooperativos el n° 3, y así sucesivamente. Estos números no indican ninguna relación jerárquica sino el tipo de colegio al que pertenecen los alumnos.

La única relación que existe entre las categorías es que son diferentes entre sí.

Se utilizan escalas nominales cuando se necesita clasificar a grupos o conjuntos de sujetos u objetos; por ejemplo, tratándose del alumnado, podría ser: por años de ingreso al plantel, por horarios de clase, por especialidad etc.

Según Ander-Egg (1972), la escala nominal tiene ciertas propiedades básicas:

- ✓ entre los sujetos u objetos clasificados puede o no existir una relación de equivalencia;
- ✓ los números que se utilizan sólo distinguen orden de posiciones de determinadas características y no relación temática;
- ✓ con ellos no puede realizarse ninguna operación aritmética.

b) Escala ordinal o de orden jerárquico.

Esta escala determina la posición de sujetos u objetos con relación a una determinada característica o cualidad según grados o niveles; pero sin referirse a la distancia que hay entre las posiciones.

Para aplicar esta escala se necesita contar con un criterio empírico respecto del atributo común a todos los elementos del conjunto, a los cuales se pretende ordenar.

Los datos también se pueden ordenar utilizando términos "ordinales verbales" tales como: bueno, regular y malo; pero en este caso, sólo se tiene una estimación cualitativa.

Los órdenes de rango "representan escalas ordinales y son los más comúnmente empleados en investigaciones educacionales. La fijación del rango que ocupa un alumno en la graduación de su clase implica el uso de una escala ordinal" (Travers, 1971).

El orden de rango indica una posición de un sujeto u objeto en la escala con relación a otros sujetos u objetos; pero no expresa una medida exacta.

Las propiedades de la escala ordinal son:

- ✓ entre los objetos ordenados puede existir una relación de rango mayor, igual o menor;
- ✓ con los datos de esta escala no se puede realizar ninguna operación aritmética.

Generalmente, la escala ordinal se utiliza para clasificar a los alumnos a partir de ciertas características, tales como: rendimiento escolar, intereses, nivel socio-económico, actitudes, etc.

c) Escala de intervalo o de distancias iguales.

Es la escala que ordena a los sujetos u objetos según la magnitud de la característica que se estudia y que, además, establece intervalos iguales.

Es conveniente señalar que los intervalos son iguales desde el punto de vista del instrumento de medición, pero no, respecto de la capacidad que se está midiendo. En el ejemplo propuesto la distancia entre 120 y 129 es la misma que existe entre 70 y 79; no obstante, la diferencia de inteligencia entre los estudiantes que están ubicados en el intervalo 70 y 79 no es tan grande como la diferencia de inteligencia entre los estudiantes ubicados en el intervalo 120 y 129.

Según Ander-Egg (1972), las propiedades de este tipo de escalas son:

- ✓ entre los objetos ordenados existe una relación de mayor, igual o menor;
- ✓ la utilización de números indica relaciones entre los objetos, y distancias entre los intervalos, que cuando son numéricamente iguales representan distancias también iguales en la cualidad o característica medida;
- ✓ el punto cero de la escala es arbitrario y convencional, por ello, cuando se comparan dos escalas, es de gran importancia haber establecido el mismo punto de partida.

En investigación educacional es frecuente utilizar escalas de intervalo, pues la mayoría de los tests de rendimiento se construyen considerando que los puntajes representan escalas de intervalo.

d. El tratamiento estadístico de los datos.

Previamente al análisis de los datos cuantificados, dice Ander-Egg (1972), es necesario, casi siempre, realizar algún tipo de tratamiento estadístico, el cual consiste en "reducir los datos en una expresión única según valores y atributos iguales, lo cual permite el conocimiento eficaz y rápido de un universo estadístico".

1) Cuantificación y descripción.

Las personas y las cosas pueden describirse en términos cualitativos basados en la presencia o ausencia de alguna cualidad o característica. El color del pelo, la raza, la nacionalidad y el sexo son elementos cualitativos y pueden ser expresados en frecuencia u orden jerárquico de aparición, en fracciones o en porcentajes del total. Las características cuantitativas son descritas en términos de magnitud o cantidad del factor presente en una situación. El peso, el tamaño, el coste o la cantidad son descripciones cuantitativas. Las medidas cualitativas pueden responder a la pregunta *cómo*. Las medidas cuantitativas responden a la cuestión de *cuánto*. Ambos tipos de descripción son importantes en la investigación y constituyen el cuerpo de los datos.

Cuando en una situación se comparan elementos que no son homogéneos es aconsejable encontrar una base para la comparación. Un método es transformar las frecuencias en porcentajes. Otra base válida y útil es ordenar por la frecuencia, el tamaño o el grado de progreso en el acercamiento hacia una norma establecida. Cuando los datos se clasifican en orden jerárquico puede ser usada la ponderación. Un primer rango vale más que el segundo, éste más que el tercero, etc. Por ejemplo, si hemos de puntuar cinco pruebas se suelen asignar los puntos como sigue:

1.a.....	5	puntos
2.a.....	4	puntos
3.a.....	3	puntos
4.a.....	2	puntos
5.a.....	1	punto

Utilizando este sistema inverso de puntuación los datos que han de ser ordenados pueden manejarse cuantitativamente con valores bien ponderados cuando las clasificaciones son complicadas.

2) Clasificación.

Cuando se analizan las características o respuestas de un grupo numeroso, resulta generalmente satisfactorio describirlo como un todo. En los tipos de análisis sencillos donde el grupo descrito es suficientemente homogéneo no es necesaria la fragmentación en subgrupos.

Pero en muchas situaciones, el cuadro del grupo total no es claro. Hay demasiadas diferencias dentro del mismo que tienden a difuminar su aspecto y hacen difícil una descripción con sentido. En tales casos puede ser útil dividirlo en categorías o clases con alguna característica común distintiva, que puede ser importante a efectos del análisis.

Esos grupos más homogéneos ofrecen aspectos que pueden llevar a ciertas generalizaciones. Pueden incluso conducir a conclusiones del tipo de causa-efecto. Este proceso de clasificación en categorías distintivas se halla implicado en toda ciencia.

Conviene recordar, sin embargo, que, una categoría sólo es distinta cuando es considerada de cierta manera o con determinado propósito. El no reconocer este hecho puede llevar fácilmente a la hipersimplificación.

a) Comparaciones.

Muchos estudios emplean la clasificación de datos en dos categorías (dicotómica). Estas categorías, como hombres o mujeres, muchachos o muchachas, casados o solteros, blancos o negros, elemental o secundario, urbano o rural, maestros buenos o malos, bachilleres o universitarios, sirven con frecuencia para buscar aspectos importantes en un grupo, pero que faltan en el otro. Esta lista podría ampliarse indefinidamente, limitada sólo por los propósitos de la investigación y la naturaleza de los datos.

Cuando las categorías se establecen sobre la base de puntuaciones en tests, por una graduación o alguna otra medida cuantitativa, es aconsejable comparar los primeros con los últimos omitiendo aquellos que se hallan cerca del centro en la distribución establecida para la comparación. Es habitual comparar el primer 25 por 100 con el último 25 por 100, o el 10 por 100 de los primeros con el 10 por 100 de los últimos. Los casos que se hallan precisamente por encima o por debajo del punto medio hacen que se difumine una diferencia

significativa que pudiera existir. Al eliminar la parte central se consigue un contraste más agudo.

Las comparaciones no son siempre dicotómicas. A veces es conveniente dividir el grupo numeroso en otras varias categorías. El número de éstas se determinará por el de características significativas diferentes que el investigador pueda identificar y por su habilidad para manejar las complejas relaciones que estas clasificaciones hacen posibles.

b) Criterios externos de comparación.

Además de las comparaciones que pueden ser hechas entre los subgrupos o *dentro* del grupo mayor, todo el grupo o los subgrupos pueden ser analizados en términos de algunos criterios externos. Naturalmente, hay que suponer que se dispone de sistemas de medida válidos y de confianza para efectuar estas comparaciones. Estos "instrumentos de medida" pueden consistir en tests tipificados, fichas de puntuación, escalas de clasificación, cálculos de frecuencia y métodos de medida físicos o psicológicos. Algunos de estos criterios externos son:

- ✓ *Condiciones, prácticas o rendimientos dominantes* de una o varias unidades de trabajo diferentes, pero comparables, Esta comparación puede hacerse con otra u otras comunidades, escuelas o clases buscando grupos que representen las mejores condiciones de experiencia y típicas; o grupos equivalentes que han sido seleccionados de acuerdo con ciertas variables, dejando para comparación una variable o un limitado número de ellas.

- ✓ *Lo que los expertos consideran como las mejores condiciones o procedimientos.* Los expertos pueden constituir un equipo especialmente seleccionado para este propósito; pueden asociarse con un grupo de profesionales en el campo correspondientes quienes se supone familiarizados con las características a examinar, o trabajar independientemente. Se suelen aceptar como criterios los juicios de autoridades reconocidas que han publicado sus opiniones.
- ✓ *Lo que determina como normas apropiadas un grupo profesional, un departamento que concede títulos, una comisión o un cuerpo científico deliberativo.* Esas normas pueden presentarse en forma de listas de objetivos o pueden ser medidas cuantitativas de las condiciones necesarias para concesión del título o aprobación.
- ✓ *Leyes o normas que han sido estatuidas o promulgadas por un cuerpo legislativo o cuasi-legislativo.* Las regulaciones para la obtención del título de maestro, las normas técnicas sobre construcciones escolares o los reglamentos sobre salud y seguridad proporcionan criterios apropiados para la comparación.
- ✓ *La demostración obtenida en la investigación.* Los factores sometidos a un análisis pueden ser examinados a la luz de principios confirmados por la investigación científica que hayan sido publicados y aceptados generalmente.
- ✓ *Opinión pública.* Aunque no siempre apropiadas como un criterio de lo que "debería ser", las opiniones o pareceres del "hombre de la calle" son aprovechables, en ocasiones, como base para la comparación.

3) Criterios de reducción.

Para la reducción de los datos se suelen tener en cuenta dos criterios opuestos:

- ✓ que los datos sean, en lo posible, lo menos numerosos, y
- ✓ que los datos no den una representación errónea de los múltiples caracteres del conjunto.

Los principales procedimientos para esta reducción consisten en determinar:

- ✓ las medidas de tendencia central o de posición, y
- ✓ las medidas de dispersión o de variabilidad.

4) Medidas de tendencia central o de posición.

Las medidas de tendencia central o deposición: media, moda y mediana, Representan el valor medio del conjunto y se utilizan según el criterio que impone el objetivo que se persigue con el estudio. Estas medidas se aplican cuando se necesita conocer las características relevantes de la muestra que pueden haber influido en los resultados obtenidos.

La Media (\bar{X}) es el valor promedio de todos los términos de una serie o conjunto, el cual es afectado por los valores extremos. Es la medida que mejor caracteriza la tendencia de la serie, aunque no discrimina a ésta en ningún aspecto. Por otra parte, es la medida más importante y la más utilizada.

La *Moda* (M_o) expresa la puntuación que en una distribución se presenta con la mayor frecuencia. Tiene la ventaja de que puede determinarse fácilmente; pero es la menos segura, pues "cuando se extraen muestras sucesivas del mismo universo, la magnitud de la moda fluctúa en mayor medida que la mediana o la media" (Van Dalen y Meyer (1971)). Los valores extremos no producen ningún efecto en la moda porque no guardan ninguna relación con esta medida.

La *Mediana* (M_d) es la medida que indica el valor central que se da en una serie o distribución de frecuencias. Es la puntuación que divide las posiciones en partes iguales, de modo que la mitad de las puntuaciones son mayores que la mediana y la otra mitad, menores que ella. Permite obtener una estimación adecuada en la mayoría de los casos, particularmente, cuando en la distribución, los puntajes extremos presentan poca frecuencia, ya que no la afecta la magnitud de estos puntajes, sino, únicamente, la frecuencia de los mismos.

5) Medidas de dispersión o de variabilidad.

Son las que expresan la forma como los valores de una variable se distribuyen en torno a las medidas de tendencia central; indican en términos generales, el "grado de las diferencias individuales alrededor de la tendencia central" (Anastasi, 1980) Se utilizan debido a que las medidas de tendencia central "no son suficientes para representar adecuadamente a un conjunto" (García Hoz y Ferrer, 1966), siendo necesario emplear una medida de dispersión al mismo tiempo que una medida de tendencia central.

Las medidas de dispersión más utilizadas en investigación educacional son la desviación estándar, la desviación cuartil, la desviación total, amplitud o rango y, como medida de dispersión relativa, el coeficiente de variación.

La *desviación estándar* (σ) es el valor numérico, que indica la forma como se distribuyen los términos de una serie, con respecto a la media. La desviación estándar es la medida, más precisa y la que más se utiliza en investigación educacional.

Así como la mediana divide a la distribución de frecuencias en dos grupos iguales, existen otros tres puntos llamados cuartiles que la dividen en cuatro grupos.

- ✓ El *cuartil superior* (Q_3) es el punto de la distribución por debajo del cual se agrupa el 75% de las puntuaciones; el *cuartil inferior* (Q_1) es el punto por debajo del cual se agrupa el 25% de las puntuaciones. El *cuartil* (Q_2) coincide con la mediana.
- ✓ La *desviación cuartil* (Q) es la mitad de la diferencia entre el cuartil más alto y el más bajo de una distribución. Es de mayor utilidad que el rango pues muestra mayor estabilidad; generalmente, se le utiliza con la mediana.
- ✓ La *desviación total*, amplitud o rango (A) es la medida más simple, expresa la diferencia entre la puntuación más alta y la más baja de una distribución. Se le utiliza conjuntamente con otras medidas de variabilidad ya que, por si sola, no discrimina la índole de la dispersión.

- ✓ El *coeficiente de variación* (CV) establece la relación entre la desviación estándar y la media. Puede representarse en: tanto por 1, tanto por 100 ó tanto por 1,000.

El propósito de cada una de estas medidas es "indicar cuán semejantes o cuán distintos son los individuos en el grupo con respecto a una característica determinada" (Selltiz, 1970). Indican la magnitud en que varían los términos en el grupo.

Respecto a la función de la estadística en investigación, es conveniente recordar con Best (1969), que la estadística "es la servidora -no la dueña- de la lógica; un medio -no un fin- de la investigación. A menos que los supuestos básicos sean válidos, los datos idóneos, cuidadosamente recogidos, registrados y tabulados y el análisis e interpretación sean lógicos, la estadística no puede contribuir a la búsqueda de la verdad".

En el caso de la *investigación operativa*, puente entre la acción educativa y la investigación pedagógica o protagónica, recogidos los datos, se procede a su estudio, análisis y clasificación. - De este modo surgen, necesariamente, algunas ideas acerca del mejor procedimiento para resolver el problema o al menos para planear un estudio más profundo que ayude a resolverlo. - Sobre estos datos e ideas se formula una hipótesis; pero esta hipótesis es siempre una *actividad, un plan de acción escolar o educativo*, que ha de llevarse a efecto en las condiciones ordinarias de la escuela o escuelas sometidas a control tan riguroso como sea posible. No se trata, pues, de una prueba aislada y esterilizada dentro del hacer escolar, es la acción de la propia escuela que somete a observación sistemática su propia actividad para perfeccionarla.

Los resultados de esta acción investigadora, o, si se quiere, de esta investigación activa, son siempre normas operacionales, principios de acción inmediatamente aplicables, a la tarea ordinaria de la escuela de la cual ha surgido.