



Universidad de Buenos Aires
Facultad de Ciencias Sociales
Carrera de Sociología

**Documento de Cátedra 63: EDICIÓN Y CARGA DE CUESTIONARIOS.
CONSIDERACIONES PRÁCTICAS PARA ALUMNOS DE LA MATERIA
METODOLOGÍA DE LA INVESTIGACIÓN SOCIAL.**

**Cecilia Fraga (cecifrag@yahoo.com.ar), Mercedes Krause
(merkrause@gmail.com) y Manuel Riveiro (manox3@yahoo.com.ar)**

Este Documento de Cátedra forma parte de una serie que tiene como propósito contribuir a la formación de alumnos de la Carrera de Sociología de la Universidad de Buenos Aires en cuestiones vinculadas al diseño y realización de investigaciones científicas en el campo de las ciencias sociales. Su contenido complementa los textos de metodología de lectura obligatoria y optativa incluidos en cada uno de los tres cursos que se dictan en la Carrera.

Estos documentos son material de uso interno y no pueden ser incorporados a ediciones impresas ni reproducidos comercialmente.

La Cátedra solicita a los usuarios de estos Documentos que citen a sus autores indicando las referencias completas, es decir: autores, fecha, título, número y tipo de documento (traducción, resumen, elaboración propia, etc.). En los casos en que el usuario utilice sólo parte del documento, haciendo referencia a algunos de los autores/obras originales allí incluidos, consignar que fue tomado de nuestro Documento de Cátedra. Por ejemplo:

E. O. Wrigth (1985), *Classes*, London: Verso, citado en Documento de Cátedra II.1., Plotno, G., Lederman, F. & Krause, M. (2007) “Escalas Ocupacionales”.

Introducción

Este estudio es parte del proyecto UBACyT S051 “El análisis de clase: formación histórica y enfoques micro-sociales” dirigido por la Dra. Ruth Sautu. El proyecto comprende varios estudios de miembros del equipo que investigan objetivos específicos vinculados a la temática general, siguiendo la tendencia actual de articular temas y metodologías. El mismo se propone estudiar los cambios históricos en la estructura de clase argentina durante la segunda mitad del siglo XX y principios del siglo XXI, mediante el análisis de datos provenientes de fuentes documentales y estadísticas. También, desde una perspectiva micro-social, busca comparar experiencias, orientaciones psicosociales y auto-identificaciones de clase de familias de clase media y popular. En este caso el método de investigación consiste en historias de familia que permiten analizar el micro-cosmos de las redes de relaciones sociales y económicas que sostienen las diferencias de clase en el mundo de la vida cotidiana. El propósito general de ambas perspectivas es articular la reconstrucción macro-histórica de las clases y los procesos micro-sociales para comprender de manera integrada la movilidad social de la sociedad argentina y sus desigualdades durante los últimos 60 años.

En el marco de la materia Metodología de la Investigación Social I, II, & III, Cátedra Sautu, Facultad de Ciencias Sociales, UBA y como parte de la formación de los alumnos de la Carrera de Sociología en investigación en ciencias sociales; se lleva a cabo desde el primer cuatrimestre de 2009 el estudio *Las clases sociales en el Área Metropolitana de Buenos Aires. Trayectorias de movilidad social intergeneracional y connubio*. Los objetivos del mismo son: el análisis de trayectorias de movilidad social intergeneracional y la selección de parejas en términos de homogamia/heterogamia de clase, de familias que residen actualmente en el Área Metropolitana de Buenos Aires.

Miembros de la cátedra diseñaron un cuestionario que recolecta información sobre: la trayectoria ocupacional y educativa de la familia del encuestado y de su cónyuge; el recorrido ocupacional, educativo y las migraciones de los abuelos; la opinión respecto a las relaciones familiares y a los ámbitos de frecuentación social y la opinión acerca de la existencia de diferencias sociales y de clase en nuestro país. También se realizaron

preguntas acerca de la auto-identificación de clase y se recogieron datos socio-demográficos del encuestado/a y los miembros del hogar.¹

El cuestionario fue aplicado por alumnos de la materia Metodología de la Investigación Social I durante el primer y el segundo cuatrimestre de 2009. En el contexto del curso de Metodología de la Investigación Social II que se focaliza en el desarrollo de estrategias cuantitativas de investigación, las encuestas fueron editadas por docentes de la materia, quienes, posteriormente, construyeron una matriz de datos.

Teniendo en cuenta lo anterior, y con vistas a profundizar en diferentes fases de la práctica investigativa, los objetivos de este documento de cátedra son: señalar la importancia del proceso de edición de encuestas en la construcción del dato, y proveer a los alumnos de algunas consideraciones prácticas para la carga de los datos y construcción de la matriz de datos. La carga de la información contenida en los cuestionarios proveerá la información para la elaboración del Trabajo Práctico Aplicado en el cual los alumnos – conformando grupos- harán un breve análisis en función de sus propios objetivos generales y específicos.

A partir del cuestionario elaborado *Las clases sociales en el Área Metropolitana de Buenos Aires. Trayectorias de movilidad social intergeneracional y connubio* el trabajo de los docentes y alumnos consistió en:

1. La selección de la muestra y el trabajo de campo
2. La edición de los cuestionarios
3. La carga de datos construidos

Una vez finalizada la tarea de carga de la matriz, en subsiguientes clases prácticas, se realizará de manera conjunta entre los docentes y alumnos el proceso de codificación de las ocupaciones. Para esto se contará con la Ponencia: “Reconstruyendo la estructura de clases: la cocina del método por encuestas” (Casteñeira et.al, 2008), y el manual de códigos de la Clasificación Internacional Uniforme de Ocupaciones (1988-OIT).

¹ La organización por secciones del cuestionario y la batería de preguntas para la indagación de las ocupaciones siguió el camino trazado por las encuestas de estratificación social en la Argentina llevadas a cabo por el Centro de Estudios de Opinión Pública (CEDOP), dirigido por el Prof. Raúl Jorrat, del Área de Estratificación Social del Instituto Gino Germani, UBA.

A continuación se detallan las secciones del documento de cátedra. Primero se describe brevemente la muestra y el trabajo de campo realizados. A continuación se detalla el proceso de edición de las encuestas y por último se señalan algunas indicaciones prácticas para la carga de la matriz de datos.

Muestra y trabajo de campo

Los datos provienen de una encuesta realizada durante 2009 a personas de 25-64 que estén trabajando o busquen trabajo, que vivan en pareja y residan en el AMBA a una muestra de 300 casos por cuotas de sexo, edad y nivel educativo. Un requisito fue que el encuestado no fuera alumno de la Carrera de sociología y que tampoco fuera un conocido del alumno.

En el paradigma cuantitativo, los supuestos epistemológicos que se refieren a la relación que se establece entre el investigador y el fenómeno a ser estudiado, el investigador mantiene una independencia de aquello que investiga. De este modo, las perspectivas del investigador, de los sujetos investigados y la audiencia están por fuera de la realidad a estudiar (Creswell, 1994; Sautu, 2005).

Enmarcado en este supuesto epistemológico, como mencionamos en la Introducción de este documento de cátedra, el trabajo de campo se hizo en el primer y segundo cuatrimestre de 2009 por alumnos de la materia Metodología de la Investigación Social I. En el espacio de las clases prácticas de la materia, se instruyó previamente a los alumnos-encuestadores en la administración del cuestionario y para que no incidan en las respuestas.

Procedimiento de campo

El trabajo de campo constituye una de las etapas más importantes en el proceso de investigación. La calidad de los datos depende, por un lado, de las decisiones teórico-metodológico (momento de la operacionalización de los conceptos teóricos hasta los ítems en el cuestionario). Y, por el otro, de la correcta aplicación del instrumento debidamente supervisada, tarea llevada a cabo por los ayudantes de las clases prácticas.

Cada alumno fue provisto de cuatro o tres cuestionarios numerados y con una cuota asignada. La selección del entrevistado no tenía mayor requisito que encontrar a una

persona que, cumpliendo con la cuota, estuviera dispuesto a conceder la encuesta. Con fines de supervisión, se debía consignar el día y la hora de la encuesta y de ser posible el teléfono de los encuestados.

La fase del trabajo de campo puede ser tan determinante en el desarrollo del estudio como la definición clara de los objetivos, las variables relevantes a investigar y la adecuada construcción de los instrumentos de recolección de datos. En este sentido, en la aplicación del instrumento de recolección y el trabajo de campo en su conjunto se pone a prueba la validación interna de una investigación.

Secciones del Cuestionario de Encuesta

El cuestionario consta de **5 secciones**, luego de la primera carilla, donde se completan los datos de los integrantes del hogar (nombre, relación con el encuestado, sexo, edad, sexo y jefatura del mismo). En **la primera sección** se preguntan, a dos columnas, **datos socio-demográficos sobre el encuestado/a y su cónyuge**, básicamente la trayectoria educativa y ocupacional, con la intención de construir la clase social “objetiva”. Luego, en **la segunda sección**, la más extensa, se indaga la **trayectoria ocupacional y educativa del padre y madre del encuestado**, (o personas que se hayan desempeñado como tales), y a continuación se realizan la misma batería de preguntas respecto al **padre y madre del cónyuge**. Esto con el fin de ver la movilidad educacional y social. La sección termina con el **recorrido ocupacional, educativo y las migraciones para los abuelos paternos y maternos del encuestado/a**. Asimismo se pregunta por el origen inmigratorio familiar y étnico del encuestado y su pareja, y por la religión en el que fue criado. Una breve **tercera sección** está destinada a conocer la opinión del respondiente respecto a sus **relaciones familiares y a los ámbitos de frecuentación social**, mientras que en **la cuarta sección**, menos “cerrada” que las anteriores, busca conocer **la opinión del encuestado** respecto de su percepción acerca de **la existencia de diferencias sociales y de clase en nuestro país**, con qué aspectos considera que se hallan asociados y los factores que considera relevantes para ascender socialmente. También se realizan preguntas acerca de su **auto-identificación de clase**. Por último, se realizan algunas preguntas sobre el ingreso total del mes del hogar y del encuestado y se recogen algunos datos de la vivienda.

Edición de los cuestionarios

La edición de las encuestas consistió en una revisión de las respuestas del cuestionario, de la articulación entre los países previamente establecidos para las distintas opciones de respuesta y en un control/evaluación de la calidad de la información que poseían las entrevistas. El objetivo era evitar el arrastre de errores que luego se trasladarían a la base de datos y a un posterior análisis. Se buscaba detectar las inconsistencias y superarlas. Parte de la tarea de la edición es detectar elementos no claros del relevamiento del cuestionario (Castañeira et al, 2008).

El proceso de edición en una investigación que utiliza el método por encuesta se constituye en uno de los elementos centrales a la hora de construir el dato ya que hace a la fiabilidad y validez que permiten obtener medidas con mayor precisión.

El problema de la validez es básico en toda investigación, pudiendo ser de distinto tipo y sesgar de muchas formas la investigación. Por ejemplo, puede suceder que en torno a la calidad de la operacionalización de las variables, los indicadores no se trasladen de manera satisfactoria a las preguntas del cuestionario; o que la forma en que están redactadas estas últimas no sea la más conveniente tanto en relación a la operacionalización como en relación con el público encuestado.

La edición permite sortear aquellos errores voluntarios o involuntarios que cometa el encuestador a la hora de recoger los datos en el campo. En este sentido, el proceso de edición, se convierte en una especie de “extensión” del trabajo de campo. La edición consiste en un proceso de toma de decisiones luego de salir al campo que elimine las inconsistencias que pueden poseer los datos o el propio instrumento de recolección. Aunque se haya realizado previamente un testeado del cuestionario, siempre es necesario volver a revisar los cuestionarios y sus posibles inconsistencias para resolverlas en el proceso de edición previo a la codificación y análisis de los datos.

El proceso de Edición: principales dificultades y las decisiones tomadas

Queriéndose aproximar a los estudios sobre estratificación social, el cuestionario da una buena aproximación a las trayectorias de movilidad social intergeneracional y conubio en el Área Metropolitana de Buenos Aires.

La ventaja de un cuestionario extenso es que brinda mucha información y por lo tanto, permite advertir mayor cantidad de inconsistencias al momento de la edición. Asimismo, permite a través de un conjunto de baterías de preguntas, completar información escasa, lo cual contribuye a la supervisión.

Sin embargo, al momento de la edición de las encuestas surgieron varios problemas, ligados tanto a la interpretación de conceptos por parte de los encuestados como también a la falta de información recolectada por parte de los encuestadores.

Dentro de la primera clase de problemas encontramos la pregunta acerca de quién es el jefe de hogar (p003jho). Se suponía que el/la encuestado/a indicaría, según su percepción, quién es el jefe/a del hogar; sin embargo, muchos encuestados se negaron a elegir una de las opciones o directamente a contestar la pregunta. Visto esto, se decidió cambiar el sistema de categorías que se suponía iba a ser: 1. Encuestado/a; 2. Cónyuge; 3. Primer hijo/a; 4. Segundo/a hijo/a; etc. Dichas opciones fueron reemplazadas por un sistema de categorías que contempla las respuestas múltiples que encontramos en los cuestionarios: 1. Encuestado; 2. Cónyuge; 3. Ambos; 4. Otro (especificar).

Otro concepto alrededor del cual surgieron problemas de interpretación fue el de habitaciones del hogar (p099). En dicha pregunta encontramos que muchos encuestados diferenciaban entre “habitaciones” y “ambientes” en términos inmobiliarios. Decidimos entonces contemplar sólo el número de habitaciones concretas, ya que el fin de esa pregunta es llegar a comparar el número de habitaciones con el número de miembros del hogar y así conformar el índice de hacinamiento del hogar.

Respecto a la condición de ocupación del/a encuestado/a y su cónyuge (p016e y p016c), se agregaron dos categorías que no habían sido contempladas al momento de la confección del cuestionario, completando el sistema de categorías de la siguiente forma: 1. Trabaja ahora; 2. Desocupado/a; 3. Ama de casa; 4. Jubilado/a. Vale aclarar que en el caso de las amas de casa, no fueron tenidas en cuenta las respuestas a la batería de preguntas sobre la ocupación principal porque se las considera inactivas. No obstante, se procuró mantener un código de respuesta diferenciado de los desocupados, para no perder el dato de la cantidad de mujeres amas de casa². En cambio, en el caso de los jubilados sí se tomaron

² En el caso de las madres y abuelas, al no existir la pregunta sobre Condición de ocupación, se les asignó el código de respuesta 9995.

en cuenta las respuestas referidas a su ocupación principal, a pesar de considerarlos también inactivos, ya que su última ocupación principal (no actual), nos servirá para clasificar su posición de clase.

Dentro del segundo tipo de problemas que encontramos a la hora de editar los cuestionarios, podemos nombrar varios casos en los que la información recolectada se encontraba incompleta y hubo que reconstruirla o darla por perdida.

Un primer ejemplo es el de la situación conyugal y la cantidad de hijos del/a encuestado/a y su cónyuge (p014 y p015). En dichos casos, cuando faltaba la información respectiva al cónyuge, se supuso que era igual a la del/a encuestado/a. Luego, en el caso de los padres del/a encuestado/a y su cónyuge, vimos que el problema no era la falta de información acerca del cónyuge sino más bien la incongruencia entre las dos respuestas. Una combinación frecuente resultó ser la de 2 (Casado) y 5 (Viuda), que recodificamos como 98 (No corresponde) y 5 (Viuda) al interpretar que el padre no estaba vivo y que se había relevado el último estado civil. Asimismo, se consideró como fallecidos a todos los padres del/a encuestado/a y su cónyuge que no hubieran respondido a la pregunta sobre su situación conyugal.

Otro ejemplo en este sentido son las preguntas sobre educación. En primer lugar, consideramos que de no responder a la pregunta sobre si continúa estudiando, la respuesta era 2 (No). En segundo lugar, se creó la nueva variable “Años de educación” a partir de la batería de preguntas sobre educación, cuyas categorías van de 0 a 18 años. En el caso de encontrarse incompleta la respuesta sobre el último año aprobado, se supuso un promedio para cada nivel educativo de la siguiente manera: Primario incompleto = 4 años; Secundario incompleto = 9 años; Terciario incompleto = 13 años; Universitario incompleto = 15 años.

Por último, algunas respuestas fueron completadas por los editores a partir de las respuestas a la misma pregunta efectuada de forma abierta. Por ejemplo: el origen familiar del/a encuestado/a y su cónyuge (p067b y p068b), qué es más importante para identificar a qué clase social pertenece una persona (p085b) y la clase social de pertenencia (p091).

Instructivo de carga: la lógica del cuestionario y la carga de los datos

La matriz de datos es un modo de ordenar los datos de manera que sea visible la forma tripartita de los mismos: las unidades de análisis, las dimensiones o variables y los valores. De esta forma, permite de manera clara y sencilla la clasificación, la comparación y una mirada integral del fenómeno estudiado. Todas las celdas deben tener un valor:

Unidad de análisis	Variable 1	Variable 2	Variable 3
unidad 1	valor	valor	valor
unidad 2	valor	valor	valor
unidad 3	valor	valor	valor

Indicaciones generales

La matriz de datos está dividida en la Hoja 1, la Hoja 2 y la Hoja 3. Es aconsejable que inicien y finalicen la carga de una hoja de manera continua para no perderse en el orden de la carga de la información. La **Hoja 1** incluye la información sociodemográfica sobre el/la encuestado/a y su cónyuge, y sobre los padres de ambos. Asimismo, se recolecta información sobre la ocupación, la educación y las migraciones de cada de ellos. En la **Hoja 2** se vuelven a repetir las mismas secciones para los abuelos paternos y maternos del/a encuestado/a. En la **Hoja 3** encontramos las preguntas sobre el origen étnico familiar, a continuación se realiza una batería de preguntas que indaga acerca del connubio y la comensalidad, y por último se indaga la opinión acerca de la existencia de las clases sociales y la autoidentificación de clase.

Las **variables marcadas con color amarillo** refieren a preguntas abiertas. Los casilleros que corresponden al resto de las preguntas fueron validadas. Es decir, que en ellos sólo es posible insertar ciertas opciones de respuestas según el tipo de variable de que se trate. Estas opciones se pueden encontrar en un menú desplegable que se encuentra haciendo click en cada casillero.

El nombre de cada variable respeta el orden de las preguntas que aparecen en el cuestionario especificando también la persona sobre la cual se está recolectando información. Por ejemplo:

- e = encuestado
- c = cónyuge
- ep = padre del encuestado
- em = madre del encuestado
- cp = padre del cónyuge
- cm = madre del cónyuge
- aop = abuelo paterno del encuestado
- aap = abuela paterna del encuestado
- aom = abuelo materno del encuestado
- aam = abuela materna del encuestado

Por ejemplo: la variable p058aop refiere a la pregunta 58 del cuestionario que recolecta información sobre el abuelo paterno del encuestado.

El orden de carga del cuestionario se realiza de forma vertical siguiendo el orden de las preguntas para cada miembro de la familia. Por ejemplo: se carga desde la pregunta 4 hasta la 22 para el encuestado y luego se carga desde la pregunta 4 a la 22 para el cónyuge.

Indicaciones específicas

A continuación detallamos los valores que asumen las respuestas, que no siempre se encuentran especificadas en el cuestionario.

En general	<ul style="list-style-type: none"> - Sí = 1 - No = 2 - NS/NC = 98 - No se aplica/ tachado = 99
Sexo	<ul style="list-style-type: none"> - Varón = 1 - Mujer = 2
Jefe de hogar	<ul style="list-style-type: none"> - Encuestado = 1

	<ul style="list-style-type: none"> - Cónyuge = 2 - Ambos = 3 - Otro = 4 (especificar)
Condición de ocupación	<ul style="list-style-type: none"> - Trabaja ahora = 1 - Desocupado = 2 - Ama de casa = 3 - Jubilado = 4
Preguntas sobre posición ocupacional (p022, p040, p057 y p066)	Para registrar la cantidad de empleados del patrón, gerente, jefe de oficina o capataz se agregó una variable. Por ejemplo a la variable p022e (posición ocupacional del encuestado), se le suma la variable p022ce (cantidad de empleados/ obreros a cargo del encuestado).
Edad/ año de nacimiento de los padres y abuelos del encuestado y de los padres del cónyuge	Se registra el año de nacimiento, no la edad.
Escalómetro de cinco segmentos (p082)	<ul style="list-style-type: none"> - Más alto = 1 - etc. - Más bajo = 5
Preguntas ordenadas jerárquicamente (p085b y p087)	Respetar el orden jerárquico indicado en la edición. Por ejemplo para la pregunta 087, se construyeron dos variables: p0871 es la primera opción de respuesta elegida y p0872 es la segunda opción de respuesta elegida. Las categorías para ambas variables pueden variar entre 1 y 4.

Bibliografía

Castañeira et al. (2008) “Reconstruyendo la estructura de clases: la cocina del método por encuestas”, presentado en I Encuentro Latinoamericano de Metodología de las Ciencias Sociales “Cambios y continuidades sociales y políticas en Argentina y la región en las últimas décadas. Desafíos para el conocimiento social”, Jornadas de Sociología de la UNLP, diciembre.

Creswell, J.W (1994) *Research design: Qualitative & quantitative approaches*, Sage: Thousand Oaks.

Sautu, R., Boniolo P., Dalle, P. y Elbert, R. (2005) *Manual de Metodología*. Clacso: Buenos Aires.