

# Recogida de datos: Introducción

Curso selectivo de nuevo ingreso

Francisco Hernández

SG de recogida de datos

10 de junio de 2016

# Recogida de datos

- Es el proceso de recopilación de la información requerida para cada una de las unidades a investigar
  - Encuestas
  - Censos
- Recopilación de información a través de fuentes administrativas o técnicas de big data: individual o agregada : a) Administración Pública , b) Otras fuentes c) Redes sociales
- La SGRD recopila la información a través de censos y encuestas

# Unidades a considerar

- Sociales: (consideración de variables nucleares)
  - Viviendas // hogares
  - Personas // Colectivos de personas:
    - inmigrantes, estudiantes , personas que acaban de finalizar un máster
- Económicas
  - Empresas (Unidades legales)
  - Establecimientos
  - Centros de cotización
  - Explotaciones agrícolas

# Métodos básicos de recogida (I)

- Auto-cumplimentación: El informante responde el mismo a la preguntas del cuestionario
  - Sobre un cuestionario en papel – Paper and Pencil Interviewing ( PAPI) ( Ejemplo: ECH- fase correo)
  - Sobre un soporte electrónico - Computer Assited Self Interviewing (CASI ) (CAWI – Web) ( Ejemplo: IPRI, IASS, Anual de empresas, ...)
    - Cuestionario muy bien probado
    - Casi seguro que requiera depuración la información recogida
    - La información en soporte electrónico será más consistente
      - Se amplia información en otras charlas

# Métodos básicos de recogida (II)

- Asistida por entrevistador en entrevista personal y presencial
  - Sobre papel – Ejemplo EPF , IPC
  - Con un ordenador o tablet – CAPI ( Computer Assisted Personal Interviewing) (Ejem: EPA; ECV)
- Asistida por entrevistador a través del teléfono
  - Se puede hacer en papel, o por medio del teléfono CATI – Computer Assited Telephone Inteviewing) ( Ejem, EPA , TIC-H; ETR

# Uso de entrevistadores: Algunas reflexiones

- Se garantiza una mejor interpretación de los conceptos utilizados en la encuesta
- Suele haber mejor tasa de respuesta que si solo se envía el cuestionario
- Cuando es vista personal hay un mayor coste ( dietas y desplazamientos, y menor ratio)
- Para usar el teléfono: a) se debe disponer de los correspondientes números, b) si la entrevista es larga o con listas largas cansa al informante

# Combinación de métodos

- La eficiencia en la recogida obliga a combinar diversos métodos
  - Una posible estrategia: Empezar por los más baratos, y seguir por los más caros
  - Condicionantes:
    - Tipo de marco disponible ( Ine usa marco de listas)
      - El marco y la información existente en el mismo, como por ejemplo los teléfonos y/o las direcciones correctas
    - Coste y tiempo
      - Uso de proxy
    - Carga a los informantes: Hay compromisos de no aumentar

# Errores en las encuestas

- Errores de muestreo
- Errores ajenos al muestreo : Muchos de ellos se producen / manifiestan durante la recogida
  - Errores de cobertura: Sobre el marco. Se miden a través de las incidencias con las unidades en recogida
  - Errores de medida: Son las diferencias entre la respuesta registrada y el verdadero valor
  - Errores de no respuesta:
    - A) Total. La unidad no responde ( Se intenta minimizar)
    - B) Parcial: No responde a alguna pregunta
  - Errores del proceso: Por ejemplo errores de codificación o de depuración, durante la recogida



# Principales herramientas (I)

- Se precisa ajustar los costes a las calidad de la encuesta
- Se necesita construir / disponer de herramientas estándar para la recogida
  - Para la cumplimentación de los cuestionarios
  - Para gestionar todo el proceso de recogida:  
Cuando nos dirigimos, cuántas veces, cómo fue la respuesta, ...
    - Sería una buena práctica proporcionar la información recogida cuanto antes al servicio promotor

# Principales herramientas (II)

- Encuestas económicas
  - Históricos ( en desaparición)
    - ARCE – Recogida CAWI
      - Xml
    - GRECO
  - Actualmente se está en el paso a IRIA
    - Fase CAWI – Casi está sustituido ARCE
    - Fase de control de todo el proceso de recogida: IRIA  
Datacollection – Empieza a finales de este año
  - En el futuro ( 2018?) todo el proceso estará en IRIA

# Principales herramientas (III)

- Encuestas a hogares
  - SIGUE: Sistema de recogida para la EPA : Consiste en la aplicación CAPI y CATI
    - En un futuro próximo se pasará a IRIA
  - ARZ- TIC-H Aplicación del SIGUE en el TIC-H y la ECH
  - En las demás encuestas cada cuestionario se hace de forma independiente,
  - Las nuevas encuestas se están diseñado en IRIA: Ejemplo EMIDE ( movilidad de estudiantes) y EILU ( Encuesta de Inserción de Universitarios) – Métodos CAWI y CATI
  - En un futuro todo en IRIA

# Delegaciones provinciales (I)

- EL INE recoge la información a través de las delegaciones provinciales
- Además de recoger datos gestionan el padrón, censos electoral y MNP
- Hay una delegación por provincia, incluidas Ceuta y Melilla
- Los TSGySC (Técnico Superior de Gestión y Servicios Comunes- EE) (entrevistadores) que son recolectores de información primaria

# Delegaciones provinciales (II)

- Los IE se dedican a la supervisión directa de los EE.
  - También apoyan al DEE en sus funciones.
  - Pueden sustituir al EE en vacaciones, enfermedades o por necesidades de servicio.
- Los DEE son los encargados de organizar y controlar la recogida en la delegación:
  - Velan por : a) la calidad de la información recogida y b) el cumplimiento de plazos
- Los delegados: Son los responsables de todo el trabajo de la delegación
  - Tienen la firma de las comisiones de servicios
- Todas las delegaciones recogen tanto encuestas económicas como de hogares: Las primeras con especialización, y las segunda generalmente la provincia

# Cupos de trabajo (I)

- Cupo de trabajo: Cuestionarios que un EE debe recoger de una operación estadística, normalmente por mes.
- En las encuestas económicas coyunturales será por meses (IPRI- 374 cuest/mes)
- En las encuestas económicas estructurales el periodo es mayor, se le han reclamaciones ( EIAE – 100 c/mes)
- En las encuestas de hogares el cupo suele ser por secciones
- Los inspectores se suelen asignar en base al número de EE ( 5 EE por 1EE, algo más en hogares para fuera de oficina)

# Cupos de trabajo (II)

- Estos cupos son revisados periódicamente, pudiendo aumentar el cupo en función de por ejemplo
  - N° de cuestionarios recogidos por Internet
  - Posibilidad de depuración selectiva
- Desde el año 2013 se han liberado recursos internalizando la ECH, ECV y TIC-H
  - Modificación sustancial de las condiciones de trabajo
- Se ha unido en la recogida de económicas a los coyunturales unos pocos cuestionarios de estructural

# Recogidas encuestas económicas

- La recogida se realiza desde oficina, y el cuestionario lo cumplimenta el informante
  - Las delegaciones especializadas por encuestas, recogiendo cada una determinada zona de España.
    - Así no todas las delegaciones recogen la muestra de su provincia
  - Con el tiempo estas zonas se han mezclando, en aras a tener que ajustar los recursos a las nuevas situaciones
  - Internet tiene unos niveles muy altos de recogida
  - Se utiliza la denominación URCE para la Unidades de Recogida de encuestas estructurales
  - La depuración es fundamental en la calidad de esta recogida
    - Suele haber recontactos con las empresas



# Recogida encuesta de hogares

- Hay recogida presencial y por teléfono.
  - En las nuevas encuestas se ha implementado por Internet - CAWI
- Para la recogida por teléfono no se asignan las unidades, se trabaja por colas
  - Estos centro CATI se encuentran por agrupación de provincias y trabajan por colas, no hay asignación directa
    - Hay 7 centros CATIs: turno de mañana y de tarde
- En la recogida presencial cada provincia recoge su muestra:
  - menor desplazamiento y mejor conocimiento de la provincia
  - Estos EE tienen unos Complementos de disponibilidad horaria

# Programación anual (I)

- La planificación de la recogida se hace de forma anual
- Se recoge en un documento que se cuelga en Sinodo
  - Consultar los diferentes documento en Sinodo (SG/ General)
  - Se consideran todas las operaciones estadísticas que manifiestan las unidades promotoras
  - Se de finen los tamaños a considerar. Los objetivos suelen venir de los Reglamentos
  - Se ajustan los recursos disponibles a las muestras
  - Si no hay recursos internos suficiente se procede a la externalización.

# Programación anual (II)

- Muchas encuestas no sufren variación del año anterior
- Para los cambios o nuevas operaciones se estudian los cupos y sus posibles repartos
- Se reasignan muestras si hubiera cambios en los recursos de algunas delegaciones.
- Aquellos trabajos que no se pueden acometer se externalizan
  - Hay que controlar /inspeccionar a las empresas contratadas

# Algunos retos en el medio plazo

- Estandarización de los procesos
  - Incorporación de las OE a la plataforma de recogida IRIA
    - » Centralización inmediata de la información recogida, con un mayor control y disponibilidad inmediata de los datos por servicios promotores
    - » Internalización del software utilizado
    - » La recogida que se haga en papel pasarla a soporte electrónico
- Mayor uso de fuentes administrativas
- Implementación de mejoras en la depuración
- Impulso del uso del CAWI en la recogida de los hogares
  - Uso de nuevos dispositivos
- Mayor implantación de los análisis cualitativos de los cuestionarios