



# **TRABAJO DE FIN DE GRADO**

## **Diseño de Encuestas y Muestreo en la Encuesta de Población Activa**

(Surveys Design and Sampling in the Active Population Survey)

**Autor:** D<sup>a</sup>. Yolanda Simón Carmona.

**Tutor/es:** D<sup>a</sup>. Eva María Artés Rodríguez.

**Grado en Administración y Dirección de empresas**

Facultad de Ciencias Económicas y Empresariales

UNIVERSIDAD DE ALMERÍA

Curso Académico: 2013 / 2014

Almería, Septiembre de 2014

# ÍNDICE

Resumen.....	3
Introducción.....	3
Capítulo 1: Importancia de la información estadística.....	5
Capítulo 2: El proceso de investigación.....	6
1. Esquema general del proceso de obtención, elaboración y análisis de resultados.....	6
2. Variables y escalas de medida.....	7
Capítulo 3: Fuentes de información estadística de la EPA.....	8
1. Elaboración de datos primarios.....	8
2. Importancia de las fuentes secundarias.....	8
Capítulo 4: Cuestionario y encuesta.....	9
1. Tipos de preguntas que aparecen en la EPA.....	9
2. Tipos de variables.....	12
Capítulo 5. Muestreo.....	13
1. Población y muestra.....	14
2. Métodos de muestreo.....	14
3. Tamaño de la muestra y errores de muestreo.....	16
4. Estimadores de las pequeñas áreas.....	19
Capítulo 6. Encuesta de Población Activa.....	19
1. Datos técnicos de la encuesta.....	19
2. Difusión de los resultados.....	22
3. Diferencias entre el paro registrado por el Servicio Público de Empleo y el registrado por el INE a través de la EPA.....	23
4. Microdatos para la explotación de la EPA.....	24
5. La reponderación como novedad metodológica de la EPA.....	26
6. Últimos datos de la EPA.....	26
7. Actualización del Censo de Población y Viviendas.....	28
8. Importancia de la EPA.....	28
9. Aplicación práctica. Análisis por series temporales de la evolución del número de parados en el período 2007-14.....	29
Conclusiones.....	34
Bibliografía.....	36
Anexo.....	38

## ***RESUMEN.***

El propósito de este trabajo es conocer la metodología de investigación de una de las principales encuestas económicas elaboradas en España, la Encuesta de Población Activa (EPA), haciendo referencia a su objetivo y a los métodos de muestreo que utiliza. También se analiza la importancia de la información estadística, las variables, las escalas y los estimadores en las pequeñas áreas que se llevan a cabo.

Uno de los objetivos principales es conocer de forma sintética los datos técnicos de la encuesta, y su importancia en la Economía y en la Sociedad.

Otro aspecto que se analiza en este trabajo es la diferencia entre el paro registrado por el Servicio Público de Empleo Estatal (INEM) y el paro registrado por el Instituto Nacional de Estadística (INE) a través de la EPA, cuestión que ha provocado en los últimos años confusiones en algunos medios de comunicación.

Finalmente se comentan los cambios y novedades metodológicas de la EPA. Se desea destacar con esto la importancia de una correcta metodología estadística en cualquier investigación económica, en vías de proporcionar estimaciones más precisas y minimizar en la medida de lo posible los errores. Se ha incluido también como aplicación práctica un análisis por series temporales.

## ***INTRODUCCIÓN.***

La Encuesta de Población Activa (EPA) la realiza el INE desde el 1964, es una encuesta continua llevada a cabo trimestralmente, está dirigida a investigar las características socioeconómicas de la población. La metodología del diseño actual han sufrido diferentes cambios, pero siempre dirigidas a una mejora en su realización. Las modificaciones se han producido en el tipo de variables que son recogidas, en la operacionalización de las mismas y en el trabajo de campo. Estos cambios han tenido que ser necesarios para poder adaptar los datos a las normas de organismos como la Organización Internacional del Trabajo o de EUROSTAT. En 1999 se produjo la última reforma de la encuesta, en ella se concretó que la difusión de los resultados se produjese en siete capítulos, dedicados a la población total, activos, ocupados, parados, inactivos, hogares y evaluación de resultados. Por otro lado, con el nuevo cuestionario se ha visto necesario añadir más tablas y aumentar el nivel de desagregación de determinadas características.

En el capítulo uno de este trabajo se describe la importancia de la información estadística, y se explica el principal objetivo de la EPA y su ámbito de difusión.

En el capítulo dos se expone un esquema del proceso general de obtención, elaboración y análisis de los resultados en una investigación, así como las distintas variables y escalas de medidas que pueden aparecer en un cuestionario, estableciendo ejemplos concretos de la EPA en cada una de ellas.

En el capítulo tres se hace referencia a las distintas fuentes de información que se utilizan para la investigación a la hora de realizar la EPA. Se distinguen las fuentes de información primarias y las fuentes de información secundarias. Las primeras son escritas por las personas que desean llevar a cabo una investigación, es decir son elaboradas subjetivamente. En cambio, las segundas son fuentes de información estadística ya elaboradas y disponible. Éstas suelen ser utilizadas porque suponen un menor coste de tiempo y de dinero.

En el capítulo cuatro se habla sobre el cuestionario de población activa, incluyendo los tipos de preguntas que hay y las distintas posibilidades de respuesta, proporcionando ejemplos.

En el capítulo cinco se hace referencia a los tipos de muestreo, el tamaño de la muestra, los errores de muestreo y los estimadores en las pequeñas áreas. La EPA tiene un muestreo bietápico con estratificación en las unidades de primera etapa. En las unidades de segunda etapa, las viviendas familiares, la información es recogida a las familias. El tamaño de la muestra actual es de 3.588 secciones y pueden existir errores de muestreo o errores ajenos al muestreo.

Por último en el capítulo seis, se explican los datos técnicos de la EPA, definiendo a toda la población a la que va dirigida. También se describe el proceso de difusión de los datos, indicando donde se pueden encontrar los resultados obtenidos con el cuestionario y la importancia de los mismos. A continuación, se lleva a cabo la justificación de porqué los datos que obtiene el Servicio Público de Empleo sobre el paro son diferentes a los obtenidos por el INE a través de la EPA.

En el siguiente apartado se expone la posibilidad de trabajar con los ficheros de microdatos, esto proporciona la ventaja de que la explotación sea mucho más flexible, ya que, permite hacer todo tipo de análisis estadísticos, y relacionar variables que no siempre

se pueden cruzar en otras aplicaciones on line. Se exponen los datos más recientes obtenidos con la EPA en el primer trimestre de 2014. En último lugar, se ha llevado a cabo un análisis por series temporales del número de parados desde el periodo 2007-14.

## **CAPÍTULO 1: IMPORTANCIA DE LA INFORMACIÓN ESTADÍSTICA.**

El principal objetivo de la Encuesta de Población Activa es tener conocimiento del grado de actividad económica del país. Su propósito es proporcionar información de las diferentes categorías poblacionales relacionadas con el mercado laboral y en función de diferentes variables conseguir clasificaciones de esas categorías.

Los objetivos de la EPA no pueden obtenerse con otras fuentes que proporcionan información relacionada con la actividad económica como el Censo, el Paro registrado o las Encuestas de Salarios.

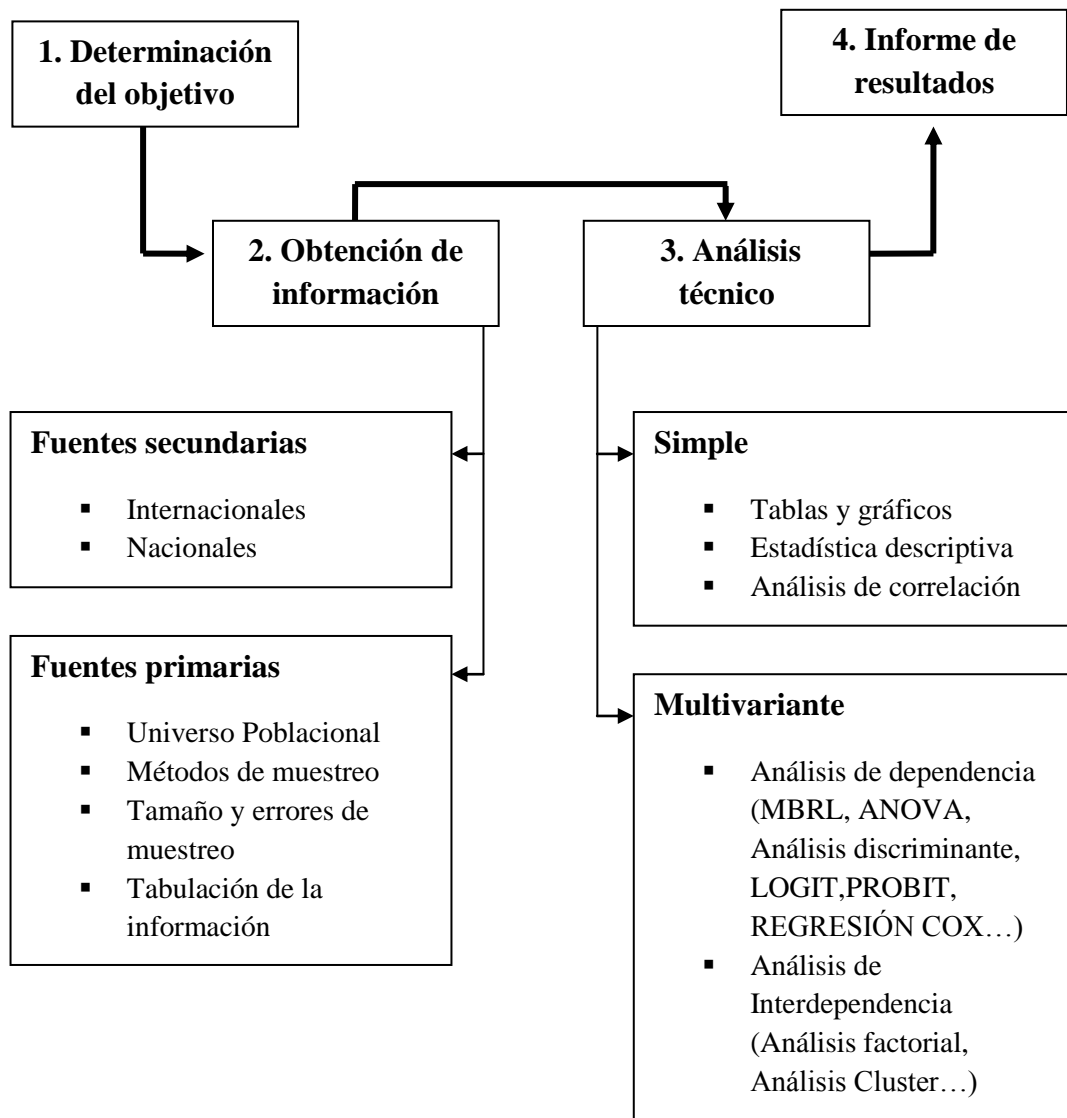
Los resultados de la EPA están diseñados para ser a nivel nacional. Pero para las Comunidades Autónomas y las provincias también se ofrece información sobre las diferentes características analizadas.

La población económicamente activa se define según la aportada por la Oficina Internacional del Trabajo (OIT), y se considera esta como: *“El conjunto de personas, que en un periodo de referencia dado, suministran mano de obra para la producción de bienes y servicios económicos o que están disponibles y hacen gestiones para incorporarse a dicha producción”*.(Instituto Nacional de Estadística, 2014).

## CAPÍTULO 2: EL PROCESO DE INVESTIGACIÓN.

### 1. Esquema general del proceso de obtención, elaboración y análisis de resultados.

Figura 2.1. Esquema General.



Fuente: Vicens 2005.

## **2. Variables y escalas de medida.**

Se considera una variable al fenómeno que estamos estudiando. Las variables pueden ser cuantitativas o cualitativas, en este último caso se llaman atributos.

El análisis estadístico nos permite resumir y organizar los datos obtenidos de fuentes de información secundaria o primaria, de manera que se pueda utilizar para elaborar un proyecto o informe robusto y completo que garantice el beneficio de la ayuda que exige la Encuesta de Población Activa.

### **2.1. Cualitativas.**

#### *2.1.1. Nominal.*

La información se clasifica en categorías no numéricas mutuamente excluyentes, entre las cuales no se puede establecer ninguna relación de orden, y no aparece ningún orden que sirva de referencia. Ejemplos que aparecen en la EPA: Sexo, Nacionalidad, Estado Civil.

#### *2.1.2. Ordinal.*

Se diferencian con las variables nominales en que en este tipo de escalas si se puede establecer algún tipo de orden, existiendo por lo tanto, algún punto de referencia para tal ordenación. Algunos ejemplos que podemos encontrar en la EPA son: Edad por tramos, Situación Económica.

### **2.2. Cuantitativas.**

#### *2.2.1. Discreta.*

Son aquellas variables que muestran separaciones en la escala de valores que pueden tomar. Es decir, son variables que presentan un número finito o infinito de valores, de manera que no cubre todos los posibles valores numéricos entre dos datos. Ejemplos de variables discretas que aparecen en la Encuesta de Población Activa son: Número de hijos, Número de meses de un contrato.

#### *2.2.2. Continuas.*

Son aquellas variables que pueden tomar un valor cualquiera que esté dentro de un intervalo dado. Por lo tanto, entre dos valores que son observables, va a existir un tercer valor intermedio que también podría ser tomado por la variable continua. Estas variables no pueden ser medidas con exactitud, ya que, el valor observado depende en gran medida

de la precisión de los instrumentos de medición. En la Encuesta de Población Activa un ejemplo de variable continua es: Ingreso de una familia. (Barbero, 2007).

## **CAPÍTULO 3: FUENTES DE INFORMACIÓN ESTADÍSTICA DE LA EPA.**

### **1. Elaboración de datos primarios.**

Una fuente primaria es una fuente que es considerada material de primera mano en lo relacionado con el fenómeno que se desea investigar. Son escritas durante el periodo de tiempo estudiado por la persona que está directamente envuelta en el evento.

La recopilación de la información debe estar bastante estructurada, para poder obtener una máxima calidad de los datos y que se puedan llevar a cabo decisiones acertadas.

La elaboración de los datos primarios puede ser realizada por distintos enfoques de investigación como son la observación, la encuesta o el experimento.

La observación consiste en obtener información a partir de personas, acciones o situaciones. La encuesta, en cambio, debe ser clara y lo más concisa posible, realizándose a un gran número de personas, con la finalidad de tener una mayor certeza de que los datos obtenidos son válidos. La validez de esta fuente depende mucho de la honestidad de los encuestadores y encuestados. Por último el experimento se lleva a cabo para recopilar información causal, lo que estudia son relaciones de causa y efecto eliminando las explicaciones competidoras de resultados observados.

Los métodos de contacto que utilizan los enfoques de investigación son llevados a cabo mediante el correo, el teléfono o con contacto personal, realizándose a una muestra a través de instrumentos de investigación como cuestionarios o instrumentos mecánicos. La función principal de las fuentes primarias es servir como elemento de contraste ente la teoría y la realidad. Entre los tipos de fuentes primarias se pueden destacar: Documentos originales, trabajos creativos, novelas, poesía, autobiografías. (Alvira, 2001).

### **2. Importancia de las Fuentes Secundarias.**

Las fuentes secundarias rediseñan la información que está publicada en las fuentes primarias. Son las más utilizadas en la investigación, informes y justificaciones de proyectos porque suponen un ahorro de tiempo. Sin embargo hay que tener en cuenta que



utilizar fuentes secundarias supone estar condicionado al conocimiento o investigación teórica del investigador.

Las fuentes secundarias pueden ser de dos tipos: Web (vía electrónica) o Soporte de papel.

No se pueden conocer todas las fuentes secundarias puesto que existen una gran cantidad.

( Alvira, 2001).

## **CAPÍTULO 4: CUESTIONARIO Y ENCUESTA.**

Los datos primarios en general se obtienen mediante la comunicación o mediante la observación. La diferencia entre ambas consiste en que la comunicación necesita a una persona o entidad para proporcionar un cuestionario y obtener así la información. En cambio, en la observación, esto no es necesario ni tampoco se realizan preguntas.

El cuestionario de la EPA abarca una serie de preguntas sobre los datos generales del encuestado y más preguntas divididas en ocho categorías relacionadas con la encuesta: Relación con la Actividad Económica, Características del empleo, Búsqueda de Empleo, Experiencia Profesional, Enseñanza y Formación, Inscripción en la Oficina Pública de Empleo, Situaciones Diversas (en la actualidad) y Situación hace un año.

(Pérez, 2013).

### **1. Tipos de preguntas que aparecen en la EPA.**

#### **1.1. Cerradas.**

Estas preguntas tienen establecidas sus opciones de respuesta, por ejemplo un “Sí”, un “No” o “No sabe”. El entrevistado sólo puede contestar a una opción por lo que son preguntas muy fáciles de responder y codificar. Las respuestas “Sí” son codificadas con un 1, las respuestas “No” con 2 y “No sabe” como 0. Un ejemplo es el siguiente:

*C1: En la semana de referencia, ¿Tenía más de un empleo o trabajó al menos una hora en otro empleo?*

- Sí.

- No.

- No sabe.

## **1.2. Abiertas.**

Consisten en que el usuario encuestado responda con sus propias palabras a la pregunta formulada. Sólo contienen la pregunta y no establecen previamente ninguna clase de respuesta, dejando esta por tanto al libre arbitrio del que realiza la encuesta. Estas preguntas presentan el inconveniente de que precisan más esfuerzo para el entrevistado y más tiempo para el entrevistador. Además son más difíciles de codificar puesto que son codificadas agrupando las respuestas similares en cada código, lo supone mayor probabilidad de error en el registro.

Por otro lado, estas preguntas pueden ser muy acertadas cuando el investigador no puede prever las posibles respuestas a una pregunta o cuando lo que se pretende es conocer un valor numérico. Una pregunta abierta que aparece en la EPA es la siguiente:

*C13: ¿Qué nivel académico requiere su puesto o qué categoría profesional tiene? ———*

## **1.3. Semicerradas o semiabiertas.**

El entrevistado puede elegir entre unas posibles respuestas que se presentan o una respuesta abierta denominada normalmente como “Otros” donde debe contestar con sus propias palabras. Un ejemplo que se puede encontrar en la EPA es:

*B6: En su trabajo principal, ¿Cuál de las siguientes es su situación profesional?.*

- *Miembro de una cooperativa.*
- *Empresario con asalariados.*
- *Trabajador independiente o empresario sin asalariados.*
- *Ayuda en la empresa o negocio familiar.- Asalariado sector público.*
- *Asalariado sector privado.*
- *Otra situación, especificar: ———*

## **1.4. Introductivas.**

Estas preguntas son realizadas en la primera parte del cuestionario. Normalmente contienen los datos personales del entrevistado y pretenden despertar el interés del mismo.

Un ejemplo de la EPA es:

*A5: ¿Cuál es su estado civil legal?*

- *Soltero.*
- *Casado.*

- *Viudo.*
- *Separado o divorciado.*
- *No sabe.*

### **1.5. De recuerdo.**

El objetivo de estas preguntas son obtener información sobre el recuerdo del entrevistado en lo relacionado a un hecho pasado. Ejemplo de una pregunta de recuerdo en la EPA es el siguiente:

*J1: ¿Realizó algún trabajo remunerado, por cuenta propia o como asalariado, en algún momento del año pasado?*

- *Sí.*
- *No.*
- *No sabe.*

### **1.6. De control.**

Son realizadas para comprobar la veracidad y honestidad de los entrevistados. Se pueden realizar preguntando una misma cosa de dos maneras distintas o comparar dos respuestas cuyas preguntas se contradicen. Un ejemplo que aparece en la EPA consiste en preguntar si se ha buscado empleo en las últimas semanas de manera diferente:

*D1: En las 4 últimas semanas, hasta el domingo de la semana de referencia, ¿Ha tratado de encontrar algún empleo o ha hecho alguna gestión para crear su propia empresa o negocio?. Considere cualquier tipo de empleo aunque sea de unas pocas horas.*

- *Sí.*
- *No.*
- *No sabe.*

*D5: ¿De qué formas ha buscado empleo en las cuatro últimas semanas? (No se leen las opciones, se señalan todas las que indique el entrevistado. Si dice alguna que no está en la lista, se anota en otra forma).*

- 1. Poniéndose en contacto con la Oficina Pública de empleo.*
- 2. Poniéndose en contacto con alguna oficina de empleo privada o apuntándose a una bolsa de trabajo.*
- 3. Solicitándolo al empresario directamente o presentando su currículum.*

4. *A través de amigos, parientes o de un sindicato.*
5. *A través de anuncios u ofertas de empleo, incluido Internet.*
6. *Haciendo gestiones para crear su propio negocio.*
7. *Realizando exámenes o entrevistas.*
8. *Ha estado esperando los resultados de una solicitud de trabajo o de una oposición.*
9. *Ha estado esperando la llamada de una Oficina Pública de empleo.*
10. *Otra forma:* \_\_\_\_\_

(Instituto Nacional de Estadística, 2014; Marban, 2014).

## **2. Tipos de variables.**

### **2.1. Respuesta simple.**

Son aquellas preguntas que solo pueden ser marcadas por una opción de respuesta. Por ejemplo:

*D16: ¿En cuál de las siguientes situaciones se encontraba antes de empezar a buscar empleo?*

- *Trabajando.*
- *Estudiando.*
- *Realizando las tareas del hogares.*
- *Atendiendo otras responsabilidades familiares.*
- *Otra situación.*
- *No sabe.*

### **2.2. Respuesta múltiple.**

Son preguntas en las cuales se pueden marcar varias opciones de respuesta. Ejemplo:

*D17: ¿Por cuál de los siguientes motivos busca otro empleo? (se leen las opciones por orden hasta haber señalado tres, las demás no se leen).*

- *Por desear un empleo más adaptado a su formación.*
- *Por inseguridad en su empleo actual.*
- *Por considerar su actividad actual como provisional.*

- Por complementar el empleo que tiene.
- Porque desea un trabajo con mayor horario.
- Porque desea un trabajo con menos horario.
- Porque desea un trabajo con mayores ingresos por horas.
- Por mejorar sus condiciones de bienestar, tiempo de transporte, horario.
- Por otras causas.
- No sabe.

### 2.3. Respuesta filtro.

Son aquellas preguntas donde la opción de respuesta te lleva a otra pregunta, sirven para obtener más información de un conjunto homogéneo de población. Por ejemplo:

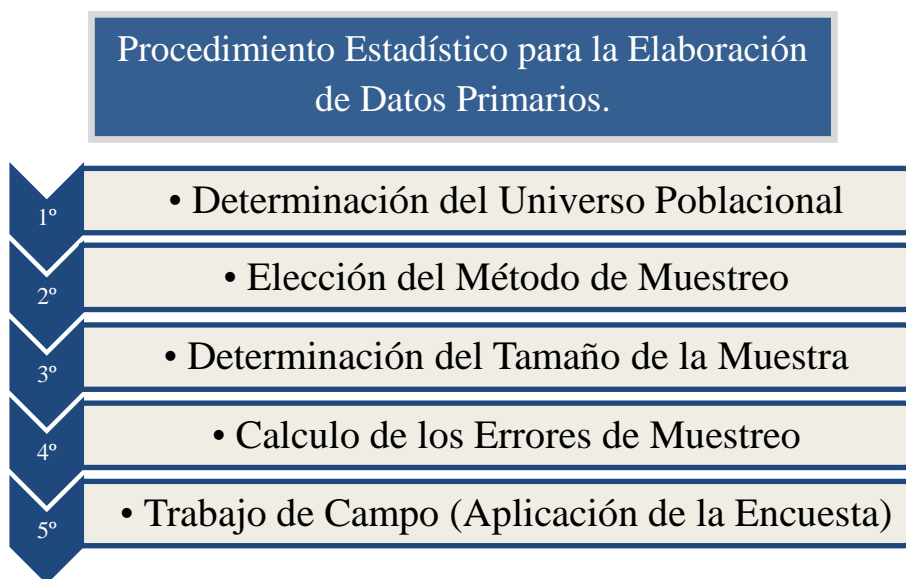
*B1: La semana de referencia, de lunes a domingo, ¿Trabajó aunque sólo fuera una hora?*

- Si.
- No (Pasar a la pregunta 3).
- No sabe (Pasar a la pregunta 3).

( Instituto Nacional de Estadística, 2014; Marban, 2014).

## CAPÍTULO 5. MUESTREO.

Una vez que se ha definido el cuestionario, se deben de realizar las siguientes etapas para que la elaboración de datos primarios sea completada:



Fuente: Elaboración propia.

## **1. Población y Muestra.**

El análisis estadístico puede extenderse o no, a todo el conjunto de elementos que participan en una investigación. La población se define como el conjunto de individuos con características comunes. En cambio la muestra es un subconjunto representativo del total de la población. En un análisis estadístico se suele analizar una muestra, ya que supone menos coste.

Cuando la investigación se dirige a toda la población, se realiza un censo, en cambio a una muestra se lleva a cabo una encuesta. (Scheaffer y otros, 2007).

## **2. Métodos de Muestreo.**

El muestreo consiste en la elección de elementos de la población que pueden que sean tenidos en cuenta en otras etapas como el tamaño de la muestra, la estratificación o en el momento del cálculo del error muestral.

### **2.1. Tipos de Muestreo.**

#### *2.1.1. Aleatorios.*

- 1º Muestreo Aleatorio Simple.
- 2º Muestreo Aleatorio por Sistemático.
- 3º Muestreo Aleatorio Estratificado.
- 4º Muestreo Aleatorio por Conglomerados.

Este método es el utilizado por los investigadores de la EPA para la elección de la vivienda en la última etapa.

#### *2.1.2. No aleatorios.*

- 5º Muestreo Aleatorio por Cuotas.

### **2.2. Muestreo en la EPA.**

El tipo de muestreo que lleva a cabo la EPA es bietápico con estratificación en las unidades de primera etapa. Las unidades de primera etapa son las secciones censales y las de segunda etapa son las viviendas familiares habitadas principales y los alojamientos fijos como las chabolas o cuevas.

Las unidades de primera etapa, las secciones censales, son estratificadas según un doble criterio:

1. Criterio geográfico (de estratificación): Según la provincia y el tipo de municipio (en función de la importancia demográfica) del que forman parte, las secciones son agrupadas en estratos.

2. Criterio socioeconómico (de subestratificación): Dentro de cada estrato geográfico las secciones censales son agrupadas en subestratos según la categoría socioeconómica de los hogares que se encuentran en la sección.

Para la formación de los estratos es necesario considerar los siguientes tipos de municipios:

1. Municipios autorrepresentados: Siempre hay secciones en la muestra dada su importancia dentro de la provincia.

2. Municipios correpresentados: Son los que dentro de la misma provincia forman parte de un grupo de municipios demográficamente similares y que son representados en común.

Se obtienen muestras autoponderadas, puesto que en cada estrato cualquier vivienda tiene la misma probabilidad de ser seleccionada. Por lo tanto, la selección de la muestra es seleccionada con una probabilidad proporcional al número de viviendas que se pueden encontrar en los últimos datos del Censo o Padrón.

Dentro de cada sección las viviendas son seleccionadas a través de un muestreo sistemático con arranque aleatorio. La información es recogida a la gente que tenga su residencia en las mismas.

La encuesta se realiza en un periodo trimestral, y la muestra se distribuye uniformemente a lo largo de todo el periodo. Debido a esto cada una de las secciones son visitadas en una de las 13 semanas del trimestre.

La muestra total está dividida en tres submuestras mensuales independientes, representativas cada una de ellas de toda la población. En los trimestres posteriores, la muestra de secciones permanece igual. Sin embargo, con la finalidad de evadir el cansancio de las familias que participan, la muestra es renovada parcialmente en una sexta parte.

Por lo tanto la muestra total de secciones está dividida en seis grupos, denominados turnos de rotación, renovándose cada trimestre las viviendas de las secciones que pertenecen a un turno determinado.

De esta forma una vivienda seleccionada se mantiene en la muestra y es entrevistada durante seis trimestres consecutivos.

(Scheaffer, y otros, 2007; Instituto Nacional de Estadística, 2014).

### **3. Tamaño de la Muestra y Errores de Muestreo.**

Para determinar el tamaño de la muestra hay que tener en cuenta lo siguiente:

1° Tipo de Muestreo utilizado.

2° El nivel de confianza → El nivel de confianza óptimo es de 95%. Esto es que el nivel de significación es de un 5%.

3° Error de Muestreo.

4° Tamaño de la población.

5° Varianza Poblacional.

#### **3.1. Tamaño de muestra en la EPA.**

El tamaño de la muestra actual es de 3.588 secciones y 18 viviendas por sección, excepto en las provincias de Madrid, Barcelona, Sevilla, Valencia y Zaragoza, en las cuales el número de entrevistas por sección es de 22. En el caso de Andalucía el número de secciones es de 585.

El tamaño de la encuesta se lleva a cabo mediante la aplicación de un procedimiento de mínima varianza para coste fijo. Se parte de un presupuesto (Q) y a partir de él se procede a determinar el número de secciones(n) y el número de viviendas (m) que minimizan la varianza de las estimaciones.

Para ello se utiliza una función de coste de tipo lineal y la expresión del coeficiente de variación para una proporción en el muestreo de conglomerados con submuestreo.

#### **3.2. Tipos de errores de muestreo en la EPA.**

1. Errores de muestreo: se producen cuando se utiliza una parte de la población para estimar las características de todo el conjunto de la población. Si el tamaño de la muestra es mayor, el error de muestreo disminuye, también influyen en él la naturaleza de la variable y las características del diseño.



La desviación típica es un buen indicador del error de muestreo, que se estima mediante la siguiente fórmula:

$$D(\hat{X}) = \sqrt{\frac{\sum_{i=1}^r (\hat{x}_i - \hat{x})^2}{r}}$$

Siendo:

r = el número de reiteraciones utilizado.

$x_i$  = la estimación obtenida con la i-ésima reiteración aplicando el mismo proceso de estimación que para la muestra completa.

$x$  = la estimación obtenida con la muestra completa.

El error de muestreo nos facilita un intervalo numérico donde existe una cierta confianza, medida en términos de probabilidad, de contener el verdadero valor de la característica estimada. Este intervalo tiene una confianza del 95%.

Cuando se precise tomar una decisión, el usuario debe tener en cuenta a las estimaciones con un error de muestreo elevado, y considerar si le puede ser útil o no.

2. Errores ajenos al muestreo: Están formados por todos los errores que se producen en las diferentes fases de la realización de una encuesta, se incluye desde el diseño del cuestionario, hasta la publicación de los resultados finales (definiciones deficientes, errores de codificación, grabación, etc).

La EPA lleva a cabo un estudio de la falta de respuesta originada por las negativas de las unidades informantes. Si se produce una falta de respuesta parcial (por ejemplo algunos datos de la persona que ha sido encuestada faltan) se produce una imputación automática. En cambio si la falta de respuesta es total (por ejemplo una vivienda encuestable por la ausencia o por ser inaccesible en el momento de la entrevista) el tratamiento recibido es diferente según sea la primera entrevista o posterior.

Si se trata de la primera entrevista las viviendas ausentes son visitadas de nuevo. Si siguen siendo inaccesibles, se vuelve a visitar a la semana siguiente. Por otro lado, las negativas de las familias se sustituyen aleatoriamente por otras viviendas de la misma sección. En

cambio, si la negativa se produce en la segunda entrevista no se sustituye, sino que son visitadas en los trimestres posteriores.

Las viviendas no encuestables, estas son las que se destinan a fines distintos a los de la vivienda familiar, vacías y de temporada, son visitadas el trimestre siguiente por si ha pasado a ser encuestable.

La falta de respuesta total de los entrevistados posteriores a la primera entrevista, consiste en repetir los datos de la vivienda obtenidos en el trimestre anterior. Pero esto solo se realiza la primera vez que se produce el incidente.

La EPA tiene en cuenta los errores de cobertura y los errores de contenido. Los errores de cobertura se producen por la omisión o inclusión errónea de unidades en la encuesta y los errores de contenido aparecen cuando alguna característica del entrevistado está recogida por error en el cuestionario. Estos errores son evaluados mediante una encuesta de evaluación que consiste en seleccionar una submuestra de viviendas donde se repite la entrevista. Los datos recogidos en la primera entrevista o entrevista original se comparan con los recogidos en la segunda o entrevista repetida. Con esto se llega a indicadores tanto de la cobertura de personas y viviendas, como de los errores de contenido de las características más importantes analizadas en la encuesta.

Los entrevistadores que el INE tiene en cada una de sus 52 delegaciones contactan por teléfono o visitan a las viviendas para realizarles la entrevista. Las encuestas son recogidas con un ordenador que está diseñado para controlar los ordenadores más importantes, sobre todo los errores de flujo.

Los entrevistadores son supervisados por los encargados de grupo bajo la dirección de un funcionario especializado.

Cuando la entrevista es finalizada y revisada la información es remitida por teleproceso a los Servicios Centrales del INE, en donde se trata informáticamente para garantizar su coherencia interna, asignar los factores de elevación y obtener tablas y series de resultados.

(Scheaffer, y otros, 2007; Instituto Nacional de Estadística, 2014).

#### **4. Estimadores de las pequeñas áreas.**

Las estimaciones de las pequeñas áreas (como por ejemplo las islas Canarias) o dominios son tratados de diferente manera. Las pequeñas áreas comprenden subconjuntos poblacionales con un tamaño inferior al considerado en el diseño de la encuesta.

La información auxiliar utilizada para dar estimaciones puede proceder de la propia encuesta (de áreas ajenas o que contienen a la considerada), de otras encuestas, de censos o de registros administrativos. (Instituto Nacional de Estadística, 2014).

## **CAPÍTULO 6. ENCUESTA DE POBLACIÓN ACTIVA.**

### **1. Datos Técnicos de la Encuesta.**

La Encuesta de Población Activa se lleva realizando desde 1964, aunque su metodología vigente en la actualidad es la de 2005. Es una investigación continua, con un periodo de tres meses, por muestreo, dirigida a todas las viviendas familiares residentes en territorio nacional.

Las viviendas familiares son las que son utilizadas todo el año o la mayor parte de él como vivienda habitual. Son excluidos de este grupo a los hogares colectivos (hospitales, hoteles, orfanatos, cuarteles, conventos, etc.). Sin embargo sí son incluidos los directores de los centros, conserjes y porteros, ya que son familias que forman un grupo independiente que residen en estos establecimientos.

Los hombres que cumplen el servicio militar, estudiantes que están en internados, enfermos ingresados en sanatorios y detenidos en cárceles se incluyen en sus viviendas de origen, ya que son sus familiares los que dan información sobre ellos. Estas categorías son subestimadas algunas veces, puesto que puede que los lazos con sus familias se hayan debilitado o porque no sean declaradas al entrevistador.

La población excluida de la encuesta está representada por menos del 1% del total.

El propósito de la EPA es obtener datos de la fuerza de trabajo y de sus diversas categorías (ocupados y parados) y de la población inactiva.

La población activa son las personas que tienen 16 o más años que, durante la semana de referencia, suministran mano de obra para la producción de bienes y servicios o están

disponibles y en condiciones de incorporarse a dicha producción. Está compuesta por ocupados y parados.

Los ocupados son personas de 16 o más años que durante la semana de referencia han trabajado al menos una hora, a cambio de una retribución (salario, jornal, beneficio empresarial, etcétera) en dinero o especie. También son ocupados quienes teniendo trabajo han estado temporalmente ausentes del mismo por enfermedad, vacaciones, etcétera.

No son ocupadas las personas que cuidan su hogar sin ningún tipo de remuneración, aquellas que prestan servicios sociales o de carácter benéfico que no están remunerados, en general son todas aquellas personas que lleven a cabo actividades sin un fin lucrativo. Tampoco se consideran ocupados los trabajadores ocasionales, estacionales o discontinuos en la época de menor actividad que no hayan trabajado en la semana de referencia.

Dentro de los ocupados hay una situación especial, los subempleados por insuficiencia de horas, son ocupados que desean trabajar más horas, que están disponibles para trabajar durante las dos semanas siguientes a la de referencia (o que declaran que no están disponibles debido al periodo de preaviso) y cuyas horas efectivas trabajadas en la semana de referencia, tanto en el empleo principal como en el empleo secundario, son inferiores a las horas semanales que trabajan los ocupados a tiempo completo de la rama de actividad en la que el subempleo tiene su empleo principal.

Los parados en cambio son las personas de 16 o más años que durante la semana de referencia han estado sin trabajo, disponibles para trabajar y buscando trabajo activamente. Se consideran parados también quienes ya han encontrado un trabajo y están a la espera de incorporarse a él, en un plazo de dos semanas a partir del domingo de la semana de referencia, siempre que verifiquen las dos primeras condiciones.

Los parados se dividen en parados que buscan su primer empleo o en parados que han trabajado anteriormente.

La población inactiva incluye a todas las personas de 16 y más años, no clasificadas como ocupadas o paradas, durante la semana de referencia. Son las siguientes:

- Personas que se ocupan de su hogar.
- Estudiantes.
- Jubilados o prejubilados.

- Personas que reciben una pensión distinta de la de jubilación y prejubilación.
- Personas que realizan sin remuneración trabajos sociales, actividades de tipo benéfico etc.
- Personas incapacitadas para trabajar.
- Personas que no ejercen ninguna actividad económica, pero reciben ayuda pública o privada y todas aquellas que no estén incluidas en ningún apartado anterior.

Los hombres que cumplen el servicio militar se consideraban población contada aparte, hasta el cuarto trimestre de 2001, es decir, no eran incluidos entre los activos ni entre los inactivos, independientemente de que en la semana de referencia trabajaran o no. Desde el 1 de Enero de 2002 se elimina el servicio obligatorio militar, debido a esto, esta categoría se suprime también de la encuesta.

España está dividida en más de 30.000 secciones censales, la EPA selecciona de entre estas a 3.588 secciones censales. En cada una de estas se entrevista un promedio de 18 hogares, con excepción de las siguientes provincias, donde se entrevistan a 22 hogares, Barcelona, Madrid, Sevilla, Valencia y Zaragoza.

De esta forma el tamaño de la muestra es de 65.000 viviendas aproximadamente, lo que supone obtener información de unas 200.000 personas. Los datos se recogen a las familias por entrevista personal y telefónica, llevada a cabo por entrevistadores fijos del Instituto Nacional de Estadística adscritos a las delegaciones provinciales del mismo. Siendo el periodo de referencia de la información la semana anterior a la entrevista.

La información es cuidadosamente depurada y procesada informáticamente y los resultados son obtenidos al cabo de un mes y medio de ser finalizado el trabajo de campo.

El tipo de muestreo es bietápico con estratificación en las unidades de primera etapa. Las unidades de primera etapa son las secciones censales y las de segunda etapa son las viviendas familiares habitadas. (Como se ha comentado anteriormente con más detalle en el capítulo 5 punto 2). (Instituto Nacional de Estadística, 2014).

## **2. Difusión de los resultados.**

La EPA pretende ser una radiografía del mercado laboral en España en el momento en el que se recoge los datos, debido a esto en el mismo día de la salida de la encuesta se dispone, en formato de papel y en la página web del INE, un resumen de los resultados más relevantes denominándose Nota de Prensa.

También cada trimestre se publica:

- Principales Resultados (papel y publicación electrónica en disquete).
- Resultados Detallados (sólo publicación electrónica en disquete).

Y anualmente son elaboradas las publicaciones:

- Tablas Anuales.
- Encuesta de Migraciones (EMI).
- Evaluación de la calidad de los datos.

Por otro lado, también son publicados los Módulos ad hoc de temática específica, para conocer con más detalle algunos aspectos del mercado laboral, los resultados están referidos al segundo trimestre de cada año y forman parte del sistema de módulos anexos a la Encuesta de Fuerza de Trabajo (EFT) de la Oficina de Estadística de la Unión Europea (EUROSTAT).

Con periodicidad variable se dispone de la Estadística de Flujos. Análisis de resultados que analiza el comportamiento de distintos grupos (jóvenes menores de 25 años, asalariados con contrato temporal, parados de larga duración, inactivos...) a partir del enlace de ficheros trimestrales de la encuesta.

Con carácter no periódico son elaborados Proyecciones de tasas de actividad con distintos horizontes de predicción y ámbito territorial (nacional y comunidad autónoma).

Por último, es posible solicitar el fichero final de la encuesta con los datos individuales anonimizados así como peticiones a medida de tablas o series específicas que necesite el usuario (siempre que sea permitido por el tamaño de la muestra).

En la página del INE también se pueden consultar resultados de las publicaciones trimestrales citadas. Estos datos son de carácter transversal y longitudinal, lo que permite analizar el mercado laboral en referencia a un trimestre determinado y también a lo largo de un periodo de tiempo más amplio y observar su evolución.

El análisis longitudinal también puede encontrarse en la página web del Banco de Datos TEMPUS (divisiones PAE Y RDE) donde se incluye un gran número de series temporales de resultados. (Instituto Nacional de Estadística, 2014).

### **3. Diferencias entre el paro registrado por el Servicio Público de Empleo y el registrado por el INE a través de la EPA.**

Es muy importante no confundir la cifra de paro registradas en el Servicio Público de empleo Estatal con las que son proporcionadas por la Encuesta de Población Activa (EPA).

Las diferencias entre ambas se deben tanto a la distinta metodología utilizada como a los colectivos que son considerados. El Servicio Público de Empleo Estatal (antes INEM) elabora una estadística administrativa para registrar el paro en base a todos los inscritos que se hayan obtenido en las distintas Oficinas Públicas de Empleo en demanda de trabajo el último día del mes.

En cambio la metodología de la EPA consiste en elaborar trimestralmente por el INE (Instituto Nacional de Estadística) una encuesta a las personas que están en edad de trabajar y son residentes en hogares familiares. Además el paro sigue criterios internacionales definidos por la OIT así como la normativa de la UE. Los colectivos considerados son las personas entre 16 y 74 años que no tienen empleo por cuenta propia o por cuenta ajena en la semana de referencia (la semana anterior a la de la encuesta), buscar activamente un empleo cuatro semanas anteriores a la entrevista y estar disponible para trabajar en las dos semanas siguientes a la entrevista.

Por otro lado, hay varias razones en las cuales se puede dar un paro superior en la EPA, ya que no todos los parados considerados en la EPA están inscritos en las oficinas de empleo ya que la obligación de la inscripción es limitada en los casos en los que el desempleado tiene derecho a percibir una prestación por desempleo o beneficiarse de las medidas de políticas activas de empleo. En cambio son considerados parados en la EPA, todos los no ocupados que declaran buscar activamente un empleo y estar disponibles para trabajar en el plazo de dos semanas después de la encuesta.

En cambio, puede ser registrado un paro superior por el Servicio Público Estatal en lugar de que el paro sea mayor en la EPA, ya que, no todas las personas que son inscritas en la oficina pública de empleo son clasificadas por la EPA como personas en paro, puesto que, una persona para considerarse como parada en la EPA tiene que declarar en la entrevista que el único método de búsqueda de trabajo es la inscripción en una oficina pública de empleo y si la inscripción ha sido hecha hace más de un mes, ha tenido contacto con la oficina en las últimas cuatro semanas. La inscripción en la oficina puede ser por un motivo diferente a la búsqueda de empleo, como obtener un beneficio en el cual sea obligatoria la inscripción o prejubilados que no buscan empleo ni están disponibles para trabajar y están inscritos para percibir una prestación de desempleo. También se puede dar el caso de personas inscritas en la oficina pero que no estén disponibles para trabajar en el plazo de quince días, y aún así el Servicio de Empleo las incluye en las cifras de paro registrado.

Estas dos estadísticas son las únicas que permiten medir el paro en España. Hasta 2004 la cifra de paro de la EPA superaba a la del paro registrado, pero 2005 la situación cambió debido a los cambios metodológicos de la EPA. En 2008 se vuelve a la situación anterior a la de 2005 y actualmente el paro registrado es inferior al paro de la EPA.

( Pérez, 2013; Navas, 2014).

En mi opinión los datos que son aportados por la EPA para definir el paro en España tienen más credibilidad que las cifras que aparecen en el Servicio de Empleo, ya que la EPA incluye exclusivamente a los colectivos que están buscando activamente un empleo y en cambio, en el Servicio Estatal de empleo hay inscripciones por motivos diferentes a la búsqueda de empleo.

Sin embargo, también hay que decir que ambas fuentes no son sustitutivas, sino complementarias.

#### **4. Microdatos para la explotación de la EPA.**

El INE ofrece la posibilidad de explotar los microdatos de la EPA, estos se encuentran en una gran cantidad de ficheros en la página web del INE, (<http://www.ine.es/prodyser/microdatos.htm>). Hay que descargarlos para poder acceder a ellos.

Los ficheros de microdatos son los ficheros de las encuestas que realiza el INE, con los datos detallados a nivel individual, para diferenciarlos de los ficheros de datos agregados,



cuyos resultados se pueden obtener en el apartado herramientas de explotación de datos disponibles en su web (base de datos Tempus).

Trabajar con los ficheros de microdatos proporciona la ventaja de que la explotación es mucho más flexible, ya que, permite hacer todo tipo de análisis estadísticos, y relacionar variables que no siempre se pueden cruzar en otras aplicaciones on line. Pero para esto se necesita mayor formación en el uso de software, así como mejores conocimientos de estadística, para poder conocer las posibilidades, y las limitaciones de trabajar con la EPA.

En el fichero Excel aparece la información sobre la estructura de los datos, hay dos tipos de registros; es decir, la información que es recogida del cuestionario se graba de dos formas diferentes. Por un lado están los registros para las personas que legalmente pueden estar ocupadas o paradas (de 16 años o más). Y por otro lado, los registros para personas menores de esa edad. La principal diferencia entre ambas es la cantidad de variables, para menores de 16 años, es mucho menor.

Los microdatos también sirven para explotar la encuesta como encuesta a viviendas, posibilidad poco explotada en los resultados que publica el INE sobre esta cuestión porque es una explotación muy compleja. Se puede conocer por ejemplo en cuantos hogares el sustentador principal y su cónyuge están en paro, u ocupados, o cualquier otro tipo de relación. Esto tiene más relevancia sociológica, porque permite estudiar la situación de una persona en el contexto de su familia.

En la explotación de una submuestra de la EPA los márgenes de error de los resultados pueden llegar a ser tan grandes que es difícil encontrar diferencias estadísticamente puesto que se dispone de pocos casos. Una forma de responder este problema es agregar varios trimestres de la encuesta, en lo que se conoce como roll-sample.

Otra posibilidad que ofrece la EPA es explotarla en tanto que panel de viviendas. Una vivienda seleccionada en un trimestre se vuelve a entrevistar durante cinco trimestres más. Por lo tanto es posible explotar un panel de hogares, y estudiar variaciones a lo largo de un periodo de tiempo, como por ejemplo, cuánto tarda un parado en encontrar trabajo o durante cuánto tiempo permanecen todos los miembros activos de una familia en el paro. Pero hay que tener en cuenta que la muestra se hace con viviendas, no con personas o familias. Por lo tanto, si durante el periodo de observación, la familia es mudada de vivienda y entra otra familia, se pasa a entrevistar a la nueva familia, y la pista de la familia

saliente es perdida. Por ello hay que intentar calcular el tamaño de los posibles sesgos que se puedan producir. (Martínez, 2005).

### **5. La reponderación como novedad metodológica de la EPA.**

La reponderación de los factores de elevación tiene como finalidad mostrar el peso que cada grupo de edad representa para la población.

Los factores de elevación son ajustados de manera que los resultados obtenidos con la encuesta sobre la población, en cada comunidad autónoma y por provincias, sean iguales a las cifras reales de la población.

La reponderación solo podrá ser llevada a cabo si se dispone de una variable de población externa a la de la encuesta con una periodicidad que permita ser utilizada trimestralmente y que sea fiable. La variable que se utiliza es el Padrón Continuo ya que permite disponer de las cifras de población frecuentemente y son consideradas de calidad.

De esta forma se puede ver a los entrevistados clasificados dentro de su tramo de edad que representatividad tienen exactamente en el total de la población. Con esta técnica algunos resultados importantes de la encuesta cambian, como es la cifra de ocupación, puesto que hay mayor ocupación en el tramo de edad entre los 25 a 54 años. (Instituto Nacional de Estadística, 2014).

### **6. Últimos datos de la EPA.**

La Encuesta de Población Activa lleva a principales conclusiones sobre la situación de España. En el INE, en nota de prensa, se pueden encontrar los resultados sobre el primer trimestre de 2014, los últimos resultados señalan una mejora del empleo, ya que se ha reducido en 79.600 personas en los últimos 12 meses. La variación anual es del -0,47%, la menos negativa desde el tercer trimestre de 2008 y la tasa de empleo ha subido ocho centésimas en el último año. En cuanto a la variación trimestral del empleo es del -1,08% (del -0,10% en términos desestacionalizados). La tasa de empleo ha descendido en el último trimestre 41 centésimas, situándose en el 44,05%.

Otro dato de interés es que la ocupación baja en 184.600 personas este trimestre, hasta un total de 16.950.600. El descenso de ocupados es el menor en un primer trimestre desde 2008. La ocupación baja en el empleo privado (195.800 personas) y aumenta en el empleo

público (11.100). Llegando a la conclusión de que en el año la ocupación se reduce en 48.400 en el sector privado y en 31.200 en el público.

La agricultura es el sector que muestra un aumento de la ocupación (31.800 más), en cambio en el sector servicios, la construcción y la industria bajan respectivamente (126.000, 48.200 y 42.200). Y la comunidad de Andalucía es la que más crece en la ocupación respecto al trimestre anterior (41.700 más). Los mayores descensos se dan en la Comunidad Valenciana (40.900 ocupados menos), Cataluña (-38.100) e Islas Baleares (-36.300 menos).

Sin duda el dato más importante que se obtiene en la EPA es el número de parados, el cual este trimestre disminuye en 2.300 personas y se sitúa en 5.933.300. La variación trimestral del desempleo es del -0,04%. En un año la cifra total de desempleados se ha reducido en 344.900 personas llegándose al mayor descenso del paro desde el tercer trimestre de 2005. Situándose las mayores bajas en Andalucía (59.200 menos), Canarias (9.300 menos) y Comunidad de Madrid (-8.600) y los mayores incrementos en Islas Baleares (18.800 más) y Galicia (16.100).

El número de activos se ha visto descendido este trimestre en 187.000 hasta 22.883.900. La tasa de actividad se sitúa en el 59,46%, siendo la más baja desde el segundo trimestre de 2007. Dándose los descensos de activos mayormente en la Comunidad de Madrid (38.900 activos menos), Cataluña (-37.400) y Comunidad Valenciana (-26.100). [9] (Instituto Nacional de Estadística, 2014).

Según una noticia del periódico El País las nuevas altas de trabajadores en el sector público desde 2010 son escasas porque la oferta de empleo de las administraciones está congelada excepto en algunas ramas que tienen una tasa de reposición del 10% (en la Seguridad Social) y no en la mutua de funcionarios (MUFACE). Esto supone que todo el incremento de afiliación en el sector público no implica creación de empleo. En las ramas de Seguridad Social más vinculadas a la Administración es donde el aumento fue de unos 70.000 afiliados.

El empleo indefinido sigue cayendo (210.000) mientras el temporal repunta (152.500); los contratos por horas suben (55.700) y los de jornada completa siguen cayendo (-135.200).

Por otro lado, los hogares que tienen todos sus miembros en paro son casi dos millones y los que no tienen ningún ingreso llega los 736.900. En cuanto a los desocupados de larga

duración, son 3,6 millones (el 61,6%). Estas personas llevan un año o más sin empleo y tienen más difícil volver a trabajar. (Gómez, 2014).

## **7. Actualización del Censo de Población y Viviendas.**

Los últimos datos de la EPA se han obtenido debido a la base de población donde se puede encontrar la información de los Censos de Población y Viviendas. Esta base de población se ha cambiado ahora y en ella se encuentra los Censos de Población y Viviendas de 2011 más actual. El primer trimestre de 2014 es el primer período que ha utilizado esta base. Los cambios se hacen cuando son detectados cambios significativos en lo relacionado al volumen de la población y en las poblaciones utilizadas en la EPA, donde se ha comprobado al obtener los resultados del Censo de 2011. Además las modificaciones también están previstas en la metodología de la encuesta y son recomendado por el EUROSTAT.

En 2005 se produjo la revisión del Censo que se utilizaba anteriormente, también como consecuencia del aumento de población que se produjo en aquel año.

La EPA consiste en trasladar los resultados obtenidos de la muestra a toda la población, estos resultados son considerados fiables y cercanos a la realidad debido al método de muestreo utilizado. De esta forma si los datos del censo se ven variados las cifras absolutas van a variar.

La EPA ha cambiado reflejando mayor población en España, y por tanto, mayor número de población en Edad de trabajar. Un resultado importante que ha variado con el nuevo Censo es la cifra del paro. La EPA utiliza a la población inmigrante como un dato más de análisis. En el Censo antiguo este grupo tenía mucho peso en la población, lo que se traducía a un mayor paro. En cambio, el nuevo Censo muestra que la población inmigrante es menor que en el censo anterior, uno de los motivos por lo que el paro en la población total se ha visto disminuido. (Grima, 2014).

## **8. Importancia de la EPA.**

La EPA es elaborada por el INE para medir el mercado laboral de España, y además nuestro país está obligado a realizar la misma metodología que todos los países de la Unión Europea, según ordena las normas del EUROSTAT. Por lo tanto, la EPA ofrece resultados homologados y que pueden ser comparados internacionalmente.

Por otro lado, el objetivo del INE es responder a las demandas de sus usuarios privados y públicos, principalmente al Consejo Superior de Estadística. En él se dictaminan los planes estadísticos cuatrienales y los programas anuales.

La EPA no es una encuesta sencilla puesto que conlleva más de 700.000 entrevistas al año, lo que supone un coste de 10 millones de euros, y la repercusión anual en medios de comunicación es de unos 70 millones de euros. El valor social neto de la EPA es alrededor de 60 millones de euros. Todo esto se traduce en aportar a los ciudadanos, a las autoridades públicas y a los agentes económicos y sociales, información de calidad, imparcial y fiable que permite explicar la realidad socioeconómica y su evolución en el tiempo.

Por otro lado, comentar algunos cambios de la EPA que no son acertados. Por ejemplo, el INE ofrecía todos los datos sobre la edad de las personas desagregada, pero ahora se proporciona en grupos quinquenales. Con esto algunos estudios educativos comprendidos entre los jóvenes de 16 y 20 años, cursando estudios post-obligatorios secundarios o en la Universidad, ya no pueden hacerse.

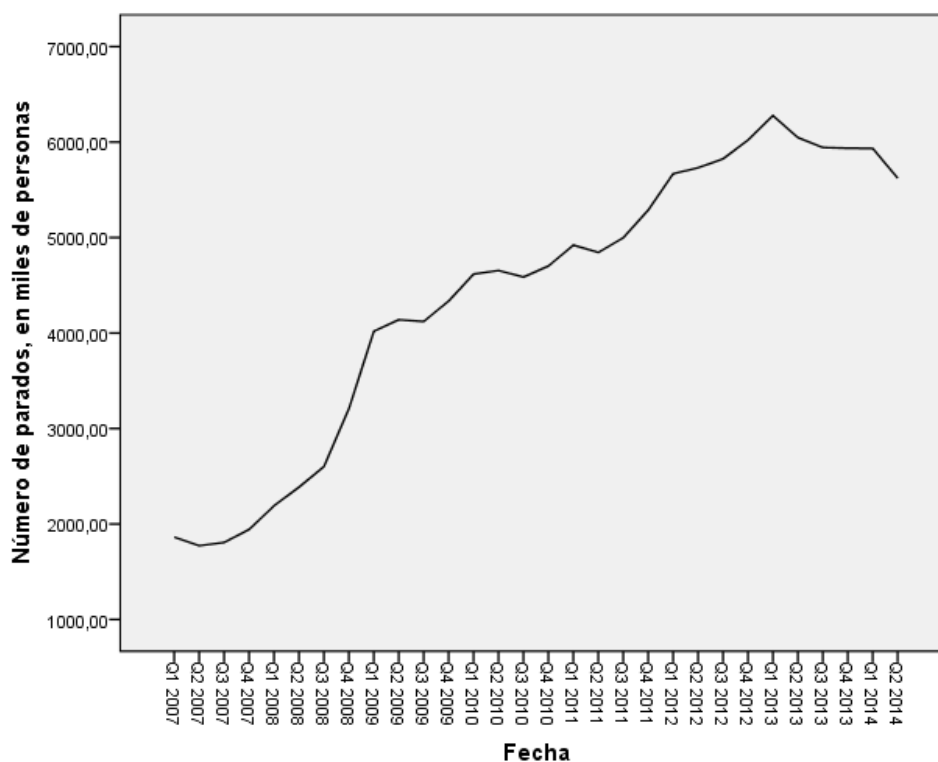
Y por último, como he comentado anteriormente, no es facilitado el factor de elevación de las viviendas, lo que dificulta el estudio de los hogares. (Álvarez e Izquierdo, 2013).

### **9. Aplicación práctica. Análisis por series temporales de la evolución del número de parados en el período 2007-14.**

Este análisis estadístico se ha llevado a cabo con el SPSS, siendo su objetivo analizar la serie temporal del número de parados en el periodo estudiado (2007-14), se ha elegido este periodo porque el 2007 fue justo el año previo al comienzo de la última crisis económica en España, hecho que afecta de forma significativa a esta variable.

El método del ajuste analítico ha permitido ver que la serie temporal sigue una tendencia lineal y creciente. También se ha estudiado la influencia que tiene la componente estacional sobre la serie, y mediante la desestacionalización se ha obtenido el gráfico que nos indica como sería la serie si no existiera esta influencia estacional sobre la misma.

**Gráfica 6.1.Serie temporal del número de parados en miles de personas 2007-14.**



Fuente: Elaboración propia con SPSS.

El gráfico muestra la serie temporal del número de parados en miles de personas desde 2007 al 2014. Una serie temporal es un conjunto de valores observados a lo largo del tiempo. Por lo tanto, el orden de observación es muy importante, ya que, lo que se pretende ver es como evoluciona una variable.

La serie temporal tiene una periodicidad trimestral puesto que los datos que la componen están recogidos por la Encuesta de Población Activa trimestralmente. Con ella se consiguen ver las regularidades que se han producido en el pasado en el número de parados y además permite predecir el comportamiento futuro del mismo.

A la vista del gráfico el número de parados en este periodo tiende a aumentar con el paso de los años. Se puede observar que es una serie suavizada, ya que, se han producido pocas fluctuaciones o variaciones en este tiempo, consecuencia de los pocos picos que se aprecian.

El aumento más pronunciado se puede ver desde el último trimestre de 2007 (1.942.000 de personas) hasta el primer trimestre de 2009 (4.018.200 personas), aumentando durante el 2008 en 2.000.000 de personas aproximadamente, esto fue consecuencia de la crisis económica estallada en ese año. A partir de ahí la variable estudiada ha ido aumentando

pero gradualmente hasta llegar a la cifra más alta del número de parados (6.278,2 de personas) en el primer trimestre de 2013.

Cifra que hay que destacar porque coincide con el pico más elevado de toda la gráfica, a partir de ahí hay una tendencia a decrecer hasta el segundo trimestre del 2014. Esto parece demostrar que el paro está disminuyendo, aunque estas cifras puede que no sean del todo reales, (como se ha comentado con más detalle en el punto 7 de este capítulo), consecuencia de la actualización del censo en 2014, donde aparecen recogidos un menor número de inmigrantes, colectivo que también analiza la EPA y que debido a esto a la hora de trasladar los resultados de la muestra al conjunto de la población, el paro puede verse disminuido.

**Tabla 6.1. Modelo de regresión lineal.**

**Coefficientes<sup>a</sup>**

Modelo	Coefficientes no estandarizados		Coefficientes tipificados	t	Sig.
	B	Error típ.	Beta		
1 (Constante)	-1424539,289	149130,500		-9,552	,000
YEAR, not periodic	710,866	74,194	,974	9,581	,000

a. Variable dependiente: parados\_mean

Fuente: Elaboración propia con SPSS.

Por otro lado, al observar la serie temporal, se aprecia que los datos se pueden ajustar a una recta, por eso se ha llevado a cabo el método del ajuste analítico. Para poder aplicar este método se necesita obtener la media de cada año. Por lo tanto, los datos que se han utilizado son el número promedio de parados en cada año, desde el 2007 hasta el 2013. (El año 2014 no se ha considerado puesto que aún no se tienen los datos para obtener la media).

El método del ajuste analítico ha permitido obtener la ecuación de la recta de la tendencia, ya que, lo que se quiere es obtener un modelo matemático que explique la variable dependiente (y), número de parados promedio a partir de la variable independiente (x), año.

$$Y = -1424539.289 + 710.866X$$

La pendiente de la ecuación es positiva, lo que indica la tendencia creciente de la serie.

**Tabla 6.2. Coeficiente de correlación lineal y coeficiente de determinación lineal.**

Resumen del modelo.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,974 <sup>a</sup>	,948	,938	392,59902

a. Variables predictoras: (Constante), YEAR, not periodic

Fuente: Elaboración propia con SPSS.

El valor de R cuadrado (Coeficiente de determinación lineal) permite conocer la forma funcional de la tendencia de la recta y conocer la bondad del ajuste del modelo lineal. El coeficiente toma valores entre 0 y 1, cuanto más próximo sea a 1, el ajuste será mejor y por tanto más fiables serán las predicciones que con él se realice. En este caso al ser 0.948, muy próximo a 1, se puede afirmar que la tendencia es claramente lineal.

**Tabla 6.3. Índices de Variación Estacional.**

Factores estacionales.

Nombre de la serie: Número de parados, en miles de personas

Período	Factor estacional (%)
1	102,6
2	100,5
3	97,7
4	99,2

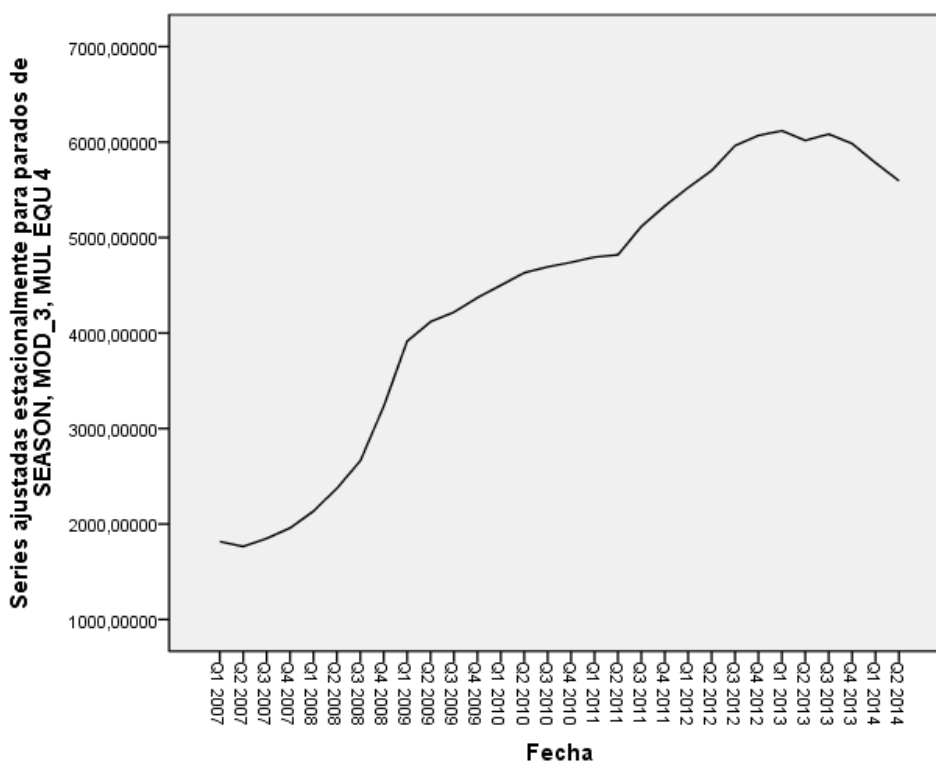
Fuente: Elaboración propia con SPSS.

A continuación se ha estudiado la componente estacional de la serie temporal, para estudiar la influencia del trimestre en el número de parados. Para ello se utilizan los índices de variación estacional (IVE), estos recogen el aumento o disminución porcentual que la componente estacional produce en cada trimestre sobre la variable. Un valor del índice igual a 100 indica que no existe influencia estacional sobre la variable.



A la vista de los índices de variación estacional observados en el cuadro anterior, se puede afirmar que la estación influye muy poco en el aumento o disminución del número de parados, ya que, todos rondan el 100%. El primer trimestre (IVE=102.6) tiene una influencia estacional positiva que hace aumentar el número de parados en un 2.6 % y en el tercer trimestre (IVE=97.7) tiene una influencia estacional negativa, que hace disminuir en un 3.3% el número de parados, un dato muy bajo, ya que, este periodo coincide con los meses de verano (Julio, Agosto y Septiembre) donde el empleo debería aumentar considerablemente respecto del resto de estaciones.

**Gráfico 6.2. Serie desestacionalizada del número de parados en el periodo 2007-14.**



Fuente: Elaboración propia con SPSS.

A partir de los Índices de Variación Estacional se obtiene la grafica de la serie desestacionalizada. Esta gráfica se ha obtenido eliminando de los valores de la serie original (gráfica 6.1.) la componente estacional, de esta manera se obtiene una serie más suavizada aún que la original.

Se puede observar que hay pocas diferencias entre ambas gráficas, esto es debido a que existe poca influencia estacional en el número de parados como he comentado anteriormente.

## ***CONCLUSIONES.***

La Encuesta de Población Activa, se trata de una investigación dirigida hacia las familias con el objetivo de conseguir datos de la población relacionados con el posible número de personas ocupadas, activas, inactivas y desempleadas. Utiliza una muestra para extrapolar los datos al resto de la población y su periodicidad es trimestral. Sus datos se utilizan en las comparaciones internacionales y son los que la Comisión Europea acepta como válidos.

Obtiene mucha información del mercado laboral, ya que la adquiere de 400 unidades distintas, de entre ellas se pueden destacar por ejemplo el número de horas extras llevadas a cabo, el número de desempleados que hacen algún tipo de curso, o la motivación para trabajar, especificadas por edad, sexo o sector profesional.

Además, a parte de proporcionar el grado de ocupación de los entrevistados, permite conocer si la ocupación es informal, regulada o si es una actividad sumergida, también ofrece información para saber si la jornada es diaria o semanal, pero este no es el propósito de la Encuesta de Población Activa. Esto hace que la veracidad de las respuestas de los entrevistados sea mayor. Por ejemplo si una persona imparte clases en su casa y ocasionalmente, la EPA considera a esta persona como ocupada, ya que ha trabajado al menos una hora en la semana de referencia a cambio de un sueldo o salario. También considera como ocupados a aquellas personas que se encuentran en vacaciones o en paro estacional por un contrato fijo discontinuo.

Por otro lado, también considera como parados a personas cuya ocupación es ser estudiantes pero están buscando activamente un empleo esporádico para obtener una ganancia económica. Y a personas jubiladas que pretenden volver al mercado laboral por diversos motivos.

Por lo tanto, uno de los factores más importantes que mide la EPA es la situación de paro en España, y como consecuencia la caída de la población activa. Por eso la encuesta es bastante útil, ya que estos datos son el motor para conocer la situación económica del país y para saber el crecimiento del mismo. Según los datos de la EPA la población activa se ha reducido en más de 400.000 personas en el último año. También la EPA muestra los emigrantes e inmigrantes que ha habido en el trimestre estudiado.

En el INE se pueden encontrar los ficheros de microdatos donde son realizadas las encuestas y se muestran los datos detallados a nivel individual, esto permite explotar los datos y con ellos hacer todo tipo de análisis estadísticos, y relacionar variables.

Es importante distinguir los resultados que muestra el INE a través de la EPA y los mostrados por el INEM, ya que son distintos. Los datos que son aportados por el Servicio Estatal de empleo muestran inscripciones diferentes a la búsqueda de empleo, en cambio la EPA incluye exclusivamente a los colectivos que están buscando activamente un empleo, por este motivo y por todo lo que se ha analizado en el punto tres del capítulo seis se considera más “creíbles” los resultados de esta última.

Por otro lado, que el número de ocupados en la EPA (16.950.600) supere al número de afiliados a la Seguridad Social (16.430.053), hace pensar que en España se produce un claro volumen de actividad sumergida, es decir personas que están trabajando pero no cotizan en la Seguridad Social.

Por último, la Encuesta de Población Activa está sometida a los problemas que presentan los errores de muestreo, los errores ajenos al muestreo y de los procedimientos de encuestación (cuestionario, trabajo de campo, etc.). Por lo tanto sus datos no están exentos de problemas y es imprescindible valorar sus resultados con las debidas cautelas así como hacer todos los esfuerzos por mejorarla.

## ***BIBLIOGRAFÍA.***

Los libros que se ha utilizado para la realización de este trabajo es la siguiente:

- Alvira Martín, F. (2001). *La encuesta: una perspectiva general metodológica*. Madrid: CIS.
- Barbero, M. I. (2007). *Métodos de elaboración de escalas*. Madrid: UNED.
- Scheaffer, R. Mendenhal, W. y Ott R.(2007). *Elementos del Muestreo*. Paraninfo.

Las páginas web utilizadas son:

- Álvarez, F. e Izquierdo, G. (2013). Reivindicando a la EPA. El país. Recuperado el 10 de Mayo de 2014, de [http://elpais.com/elpais/2013/02/12/opinion/1360684794\\_978758.html](http://elpais.com/elpais/2013/02/12/opinion/1360684794_978758.html)
- Gómez, M. (2014). España sigue destruyendo empleo en el primer trimestre de 2014. El país. Recuperado el 12 de Mayo de 2014, de [http://economia.elpais.com/economia/2014/04/29/actualidad/1398754752\\_267733.html?rel=rosEP](http://economia.elpais.com/economia/2014/04/29/actualidad/1398754752_267733.html?rel=rosEP)
- Grima, C. (2014). ¿Qué pasa con la nueva EPA?.El diario. Recuperado el 2 de Agosto de 2014, de [http://www.eldiario.es/andalucia/desdeelsur/pasa-nueva-EPA\\_6\\_238736151.html](http://www.eldiario.es/andalucia/desdeelsur/pasa-nueva-EPA_6_238736151.html)
- Instituto Nacional de Estadística. (2014). Encuesta de población activa 2005. Recuperado 17 de Mayo de 2014, de <http://www.ine.es/inebaseDYN/epa30308/docs/epacues05.pdf>
- Instituto Nacional de Estadística. (2014). Descripción de la encuesta. Recuperado el 9 de Mayo de 2014, de <http://www.ine.es/epa02/descripcion%20encuesta.pdf>
- Instituto Nacional de Estadística. (2014). Notas de prensa. Recuperado el 14 de Mayo de 2014, de <http://www.ine.es/daco/daco42/daco4211/epa0114.pdf>
- Instituto Nacional de Estadística. (2014). Novedades metodológicas. Recuperado el 15 de Mayo de 2014, de <http://www.ine.es/epa02/repondera.htm>
- Marban, V. (2014). La encuesta estadística. Recuperado el 6 de Agosto de 2014. de [http://www2.uah.es/vicente\\_marban/ASIGNATURAS/SOCIOLOGIA%20ECONOMICA/TEMA%205/tema%205.pdf](http://www2.uah.es/vicente_marban/ASIGNATURAS/SOCIOLOGIA%20ECONOMICA/TEMA%205/tema%205.pdf)

- Martínez, J. (2005). Explotación de la encuesta de población activa. Recuperado el 20 de Mayo de 2014 de <http://josamaga.webs.ull.es/jsmg-epa.pdf>
- Navas, J. (2014). ¿Por qué el paro registrado en el INEM es diferente al dato de la EPA?. El mundo. Recuperado el 21 de Mayo de 2014, de <http://www.elmundo.es/elmundo/2012/05/03/economia/1336064669.html>
- Pérez, B. (2013). Así se cuece la EPA, la encuesta más controvertida del INE. La información. Recuperado el 2 de Agosto de 2014, de [http://noticias.lainformacion.com/mano-de-obra/mercado-laboral/asi-se-cuece-la-epa-la-encuesta-mas-controvertida-del-ine\\_iUXDm6fng3GIzxAdTuFaa7/](http://noticias.lainformacion.com/mano-de-obra/mercado-laboral/asi-se-cuece-la-epa-la-encuesta-mas-controvertida-del-ine_iUXDm6fng3GIzxAdTuFaa7/)
- Pérez, I. (2013). ¿Por qué el paro de la EPA es distinto al paro registrado?. Seguridad social activa. Recuperado el 21 de Mayo de 2014, de [http://www1.seg-social.es/ActivaInternet/Opinion/REV\\_026935](http://www1.seg-social.es/ActivaInternet/Opinion/REV_026935)

## **ANEXO.**

### **ANEXO 1. Encuesta de Población Activa (EPA). Serie histórica (Datos en miles de personas).**

#### **Datos revisados.**

<b>Trimestre</b>	<b>Activos</b>	<b>Ocupados</b>	<b>Parados</b>	<b>Tasa de actividad (en %)</b>	<b>Tasa de paro (en %)</b>
1T 2014	22.883,9	16.950,6	5.933,3	59,46	25,93
4T 2013	23.070,9	17.135,2	5.935,6	59,86	25,73
3T 2013	23.173,4	17.230,0	5.943,4	60,04	25,65
2T 2013	23.207,9	17.160,6	6.047,3	60,00	26,06
1T 2013	23.308,4	17.030,2	6.278,2	60,18	26,94
4T 2012	23.360,4	17.339,4	6.021,0	60,23	25,77
3T 2012	23.491,9	17.667,7	5.824,2	60,55	24,79
2T 2012	23.489,5	17.758,5	5.731,0	60,50	24,40
1T 2012	23.433,0	17.765,1	5.667,9	60,31	24,19
4T 2011	23.440,3	18.153,0	5.287,3	60,29	22,56
3T 2011	23.482,5	18.484,5	4.998,0	60,44	21,28
2T 2011	23.466,2	18.622,0	4.844,2	60,44	20,64
1T 2011	23.347,3	18.426,2	4.921,2	60,16	21,08
4T 2010	23.377,1	18.674,9	4.702,2	60,25	20,11
3T 2010	23.404,4	18.819,0	4.585,4	60,37	19,59
2T 2010	23.406,4	18.751,1	4.655,3	60,41	19,89
1T 2010	23.270,5	18.652,9	4.617,7	60,09	19,84
4T 2009	23.225,4	18.890,4	4.335,0	59,99	18,66
3T 2009	23.219,8	19.098,4	4.121,4	60,05	17,75
2T 2009	23.293,4	19.154,2	4.139,6	60,30	17,77
1T 2009	23.302,6	19.284,4	4.018,2	60,39	17,24
4T 2008	23.262,1	19.055,3	3.206,8	60,35	13,79
3T 2008	23.157,1	20.556,4	2.600,7	60,23	11,23
2T 2008	23.032,6	20.646,9	2.385,7	60,07	10,36
1T 2008	22.810,4	20.620,0	2.190,5	59,67	9,60

4T 2007	22.659,9	20.717,9	1.942,0	59,47	8,57
3T 2007	22.559,6	20.753,4	1.806,2	59,50	8,01
2T 2007	22.354,1	20.580,9	1.773,2	59,23	7,93
1T 2007	22.130,8	20.267,5	1.863,2	58,90	8,42
4T 2006	22.014,8	20.195,4	1.819,4	58,88	8,26
3T 2006	21.857,8	20.091,0	1.766,9	58,74	8,08
2T 2006	21.726,0	19.891,6	1.834,4	58,63	8,44
1T 2006	21.521,3	19.578,4	1.942,8	58,30	9,03
4T 2005	21.369,5	19.509,2	1.860,3	58,08	8,71
3T 2005	21.205,6	19.422,1	1.783,5	57,86	8,41
2T 2005	21.129,7	19.160,6	1.969,1	57,86	9,32
1T 2005	20.857,4	18.736,1	2.121,3	57,38	10,17
4T 2004	20.667,7	18.490,8	2.176,9	57,12	10,53
3T 2004	20.488,9	18.289,1	2.199,8	56,93	10,74
2T 2004	20.266,5	18.018,9	2.247,6	56,55	11,09
1T 2004	20.080,0	17.770,2	2.309,8	56,24	11,50
4T 2003	20.017,1	17.740,5	2.276,7	56,25	11,37
3T 2003	19.893,5	17.646,0	2.247,5	56,15	11,30
2T 2003	19.639,2	17.423,2	2.216,0	55,65	11,28
1T 2003	19.421,2	17.092,7	2.328,5	55,28	11,99
4T 2002	19.224,3	16.991,9	2.232,4	54,95	11,61
3T 2002	19.115,3	16.919,3	2.196,0	54,92	11,49
2T 2002	18.870,2	16.766,9	2.103,3	54,48	11,15
1T 2002	18.635,1	16.482,3	2.152,8	54,06	11,55
4T 2001	18.292,0	16.348,2	1.943,8	53,41	10,63
3T 2001	18.164,0	16.294,3	1.869,7	53,22	10,29
2T 2001	17.932,1	16.076,3	1.855,8	52,73	10,35
1T 2001	17.814,7	15.866,3	1.948,3	52,57	10,94

