

---

# INDICADORES ESTADÍSTICOS ASOCIADOS A LA ENCUESTA DE CONDICIONES DE VIDA

---

TRABAJO FIN DE GRADO



2022

VÍCTOR TOSCANO DURÁN  
GRADO EN ESTADÍSTICA, FACULTAD DE MATEMATICAS

## *Resumen*

La encuesta de condiciones de vida proporciona información sobre la calidad de vida en nuestro país y permite abordar el análisis de la pobreza y de la desigualdad de una manera homologable y anual. Actualmente, es un tema de relevancia debido a la reciente crisis económica que ha vivido el país en 2008 y que se ve perjudicada hoy en día por otros sucesos de gran escala como puede ser la pandemia Covid-19. Estas características sobre nuestra sociedad son medibles gracias a los indicadores proporcionados por dicha encuesta. El objetivo es conocer la situación del País en la última década mediante la información que nos proporciona esta encuesta y se realiza el análisis de la desigualdad y de la pobreza por comunidades autónomas, así como un análisis de su evolución en el intervalo de tiempo del 2009 al 2020.

## *ABSTRACT*

The Living Conditions Survey provides information on the quality of life in our country and allows us to address the analysis of poverty and inequality in a comparable and annual way. It is currently a relevant topic due to the recent economic crisis that the country has experienced in 2008 and which is currently undermined by other large-scale events such as the current Covid-19 pandemic. These characteristics of our society are measurable thanks to the indicators provided by this survey. The aim is to know the situation of the country in the last decade through the information provided by this survey and an analysis of inequality and poverty by autonomous communities is carried out, as well as an analysis of its evolution in the time interval from 2009 to 2020.

---

# ÍNDICE

---

## Contenido

1	Encuesta de condiciones de vida .....	4
1.1	Introducción general del objetivo de esta encuesta .....	4
1.2	Metodología general en España a grandes rasgos.....	5
1.2.1	Evolución temporal .....	5
1.2.2	Ámbito de investigación y tipo de muestreo. ....	6
1.2.3	Cuestionario .....	8
1.2.4	Error de Muestreo.....	8
2	Definiciones e Indicadores asociados a la encuesta .....	10
2.1	Definiciones.....	10
2.1.1	Vivienda y tipos .....	10
2.1.2	Personas y tipos.....	11
2.1.3	Otras definiciones .....	12
2.2	Pobreza y Desigualdad .....	15
2.3	Principales indicadores.....	16
2.3.1	Lista de Indicadores.....	17
2.3.2	Indicadores de desigualdad.....	20
3	Análisis de los Indicadores en España .....	28
3.1	Serie temporal de los principales indicadores en España .....	28
3.1.1	Renta neta media por persona.....	28
3.1.2	Umbral de pobreza.....	30
3.1.3	Tasa riesgo de pobreza AROPE.....	31
3.1.4	Ratio s80s20 .....	36
3.1.5	Índice y coeficiente de Gini .....	36
3.2	Principales Indicadores en las Comunidades Españolas en el 2009. ....	38

3.2.1	Curva de Lorentz .....	39
3.2.2	Renta neta media .....	41
3.2.3	Tasa de Pobreza .....	43
3.2.4	Coeficiente Gini .....	46
3.2.5	Análisis Clúster con respecto a los indicadores de condiciones de vida y desigualdad .....	47
3.2.6	Evaluación situación 2009.....	51
3.3	Principales Indicadores en las Comunidades Españolas en el 2020 .....	52
3.3.1	Curva de Lorentz .....	52
3.3.2	Renta neta media .....	54
3.3.3	Tasa de pobreza .....	57
3.3.4	Coeficiente Gini .....	60
3.3.5	Análisis Clúster con respecto a los indicadores de condiciones de vida y desigualdad. ....	61
3.3.6	Evaluación situación 2020.....	64
3.4	Evolución principales indicadores desigualdad y condiciones de vida 2009-2020.....	64
4	Bibliografía .....	72

# 1 Encuesta de condiciones de vida

## 1.1 Introducción general del objetivo de esta encuesta

El objetivo del presente trabajo es el estudio y análisis de la encuesta de condiciones de vida y de sus principales indicadores en España durante el periodo de 2008 a 2020, haciendo énfasis en estudiar cómo ha evolucionado desde el inicio de la crisis económica estos indicadores y la desigualdad en España.

La Encuesta de Condiciones de Vida (ECV), en inglés “European Statistics on Income and Living Conditions” (EU-SILC), es una operación estadística armonizada que se realiza de manera anual en todos los países de la Unión Europea, relativa a las estadísticas comunitarias sobre la renta y las condiciones de vida.

El reglamento europeo especifica los objetivos, las áreas cubiertas, el marco temporal y más aspectos relacionados.

El objetivo fundamental que se busca conseguir con la ECV es la de generar una fuente de referencias estadísticas que permita comparar la distribución de los ingresos y la exclusión social en el ámbito europeo, así como generar una medida de desigualdad para la sociedad. La ECV proporciona información muy importante para desarrollar políticas socioeconómicas en el contexto del mercado único. Esta operación estadística es flexible en relación con las fuentes estadísticas y se recomienda el uso de las existentes. La ECV va dirigida a recoger datos sobre las privaciones de los hogares, las condiciones de los hogares y el indicador de riesgo de pobreza o de exclusión social armonizado.

Estos datos presentan tanto la dimensión transversal como la longitudinal, aunque se da prioridad a la dimensión transversal. Este doble planteamiento tiene grandes ventajas ya que la componente longitudinal (en el caso español se realizará a lo largo de cuatro años) permite estudiar los cambios que se producen en ciertas personas con el paso del tiempo y poder detectar su reacción a dichos cambios. Por otro lado, gracias a la componente transversal podemos sacar conclusiones de calidad en lo que a su actualidad y comparabilidad se refiere.

En conclusión, la ECV nos ayuda a conocer la situación de la población con la finalidad de ayudar en la creación de políticas públicas para poder mejorar las condiciones de la población como pueden ser la exclusión social, ingresos, rentas, etc.

## 1.2 Metodología general en España a grandes rasgos

### 1.2.1 Evolución temporal

La ECV comienza en España en 2004 y es una encuesta de periodicidad anual, es decir los ficheros de microdatos se generan con una periodicidad anual. Es una encuesta que se realiza de manera presencial, con la excepción de 2020 y 2021 debido a la pandemia sanitaria causada por el Covid-19, en los que se ha realizado de manera telemática la recogida de información, ya sea a través de llamada telefónica o internet.

A partir del año 2005 se introdujeron módulos adicionales en la componente transversal para temas de interés especial que pueden repetirse cada cuatro o más años.

#### *Modificación de la definición de “hijos dependientes”.*

Hasta el año 2007, se consideraban **hijos dependientes** a la población de edad menor de 16 años y a la población entre 16 y 24 años que vivía con al menos una de sus figuras paternas y no tenía actividad económica se le consideraba hijos dependientes. Esto cambió en 2008, considerando **hijos dependientes** a la población menor de 18 años y a la población entre 18 y 24 años que vive al menos con una de sus figuras paternas y que es inactiva económicamente.

#### *Incorporación de registros administrativos.*

A partir de 2013 se empezó a recoger información de fuentes fiscales y administrativa, que complementa la información recogida en el trabajo de campo, lo cual produjo una ruptura en la serie temporal de esta estadística en la forma de obtener ingresos del hogar.

#### *Investigación exhaustiva por comunidades.*

Aunque el INE presenta y difunde datos por Comunidades Autónomas, otras regiones, al igual que en Andalucía, incluyen entre sus actividades estadísticas la explotación y publicación de datos de la ECV para su ámbito territorial. Así, por ejemplo, la Comunidad de Navarra en su IV Plan de Estadística (Años 2011-2016) incluye entre sus actividades estadísticas la realización de la “Encuesta de Condiciones de Vida”.

El Instituto de Estadística y Cartografía de Andalucía (IECA) obtiene los resultados publicados en este trabajo a partir de la explotación de los ficheros de microdatos anonimizados de la Encuesta de Condiciones de Vida proporcionados por el INE. Una vez recibido, el IECA explota el fichero de microdatos para obtener la información estadística de Andalucía. (IECA, 2017)

Hay que destacar el caso de Cataluña, ya que en el año 2015, el INE y el Instituto de Estadística de Cataluña (Idescat) atendiendo a la necesidad de aumentar la fiabilidad estadística de los resultados presentados y de ofrecer información territorial más detallada, decidieron colaborar para ampliar la muestra en el ámbito territorial catalán. El acuerdo se hace efectivo a partir de la ECV del año 2016. En ese mismo año la operación estatal contribuyó con la información correspondiente a 1.275 hogares y 3.264 individuos en Cataluña, mientras que la operación adicional gestionada por el Idescat aportó 1.751 hogares y 4.094 individuos. Simultáneamente a la ampliación, el Idescat y el INE acordaron añadir dos nuevas preguntas en el cuestionario del hogar (sobre la necesidad del hogar de pedir ayuda por su situación económica) y otra en el cuestionario individual (para conocer la situación de las personas que tienen limitaciones para realizar las actividades de la vida diaria). (Instituto de Estadística de Cataluña, s.f.)

Finalmente, hay que tener en cuenta que esta operación estadística se llamaba “Estadística de distribución personal de la renta y riesgo de pobreza antes de 2016.

### 1.2.2 Ámbito de investigación y tipo de muestreo.

#### *Ámbito poblacional*

El ámbito poblacional de esta encuesta lo constituye las personas miembros de hogares privados que residen en viviendas familiares principales.

La población seleccionable para la investigación exhaustiva son aquellos miembros del hogar con 16 o más años el 31 de diciembre del año anterior al de la entrevista.

#### *Ámbito geográfico*

El ámbito geográfico lo constituye todo el territorio español.

#### *Ámbito temporal*

El ámbito temporal difiere en función de la variable objeto de estudio: semana anterior a la entrevista, último mes, último año etc. Por otro lado, se recoge información biográfica de los entrevistados, lo que indica que el periodo de referencia será diferente para cada individuo.

## Muestreo

La ECV es una encuesta anual con un diseño de **panel rotante** (los hogares permanecen durante un número determinado de períodos en la muestra y luego son sustituidos) en el que la muestra la forman cuatro submuestras independientes, cada una de las cuales es un panel de cuatro años de duración. Cada año se renueva la muestra en uno de los paneles.

El tipo de muestreo usado en estas submuestras es un **muestreo bietápico** con estratificación en las unidades de primera etapa. Las unidades de la primera etapa son las secciones censales y las de la segunda etapa son las viviendas familiares principales. Las secciones se seleccionan dentro de cada estrato con probabilidad proporcional a su tamaño y las viviendas se seleccionan, dentro de cada sección, con igual probabilidad mediante muestreo sistemático con arranque aleatorio, lo que conduce a muestras autoponderadas en cada estrato.

En cada Comunidad Autónoma se diseña una muestra independiente que la represente, con el fin de producir resultados significativos para el siguiente nivel de desagregación: Las unidades de primera etapa (secciones) se agrupan en estratos de acuerdo con el tamaño del municipio al que pertenecen. La muestra complementaria llevada a cabo para Cataluña por el Idescat se lleva a cabo mediante la selección de viviendas usando un muestreo aleatorio estratificado en dos etapas.

Para poder cumplir con los objetivos de esta encuesta y proporcionar resultados con un nivel de fiabilidad aceptable a nivel nacional y a nivel de las comunidades, se selecciona una muestra de 16000 viviendas distribuidas en 2000 secciones censales y a su vez la muestra se distribuye entre las comunidades asignando una parte uniforme (40% de las secciones) y otra proporcional al tamaño de las comunidades.

Por lo tanto, se destacan como aspectos importantes a tener en cuenta de la ECV los siguientes:

- La encuesta como hemos dicho es anual y cada año se denomina ciclo de la encuesta.
- La muestra está dividida en 4 grupos de rotación y se denomina periodo al número de años que lleva un grupo de rotación en la muestra. Estos periodos van del 1 al 4 siendo el primer periodo el año que un grupo de rotación entra en la muestra y el 4 el último que permanece en la misma.
- En cada ciclo se renueva una cuarta parte de la muestra
- En el primer ciclo de la ECV toda la muestra obviamente es nueva y estará en el periodo 1.

- Una cuarta parte de las unidades seleccionadas en la muestra del primer ciclo solo permanecen un año, otra cuarta parte dos y otra cuarta parte tres años.

### 1.2.3 Cuestionario

El cuestionario correspondiente a la encuesta de condiciones de vida se divide en 4 partes:

1. **Mesa de Composición del hogar:** Se recoge información básica de las personas que ocupan la vivienda y su objetivo es determinar las personas que forman parte del hogar.
2. **Ficha del hogar:** Se recoge información general de las personas que componen el hogar, pero con la excepción de que se solo se realiza a las personas que cumplen los requisitos de ser miembros del hogar. (Complementa a la 1 parte del cuestionario)
3. **Cuestionario del Hogar:** Se recoge información sobre la vivienda encuestada (la unidad de observación). Se realiza a la persona responsable del hogar.
4. **Cuestionario individual:** Se recoge información más detallada sobre cada miembro del hogar (mayor de 16 años el año anterior a la entrevista)

En la página web del INE podemos acceder a estos cuestionarios y observar su contenido: <https://www.ine.es/metodologia/t25/t2530453cues.htm> .

### 1.2.4 Error de Muestreo

Como es conocido, el error de muestreo mide la imprecisión que se comete al estimar una característica de la población de estudio (parámetro) mediante el valor obtenido a partir de una parte o muestra de esa población (estadístico). Este error depende de muchos factores como pueden ser del método de extracción de la muestra, su tamaño etc. Podemos definir una expresión generalizada del error de muestreo:

$$\text{Error de muestreo} = \text{Var}(\theta) \quad (\theta \text{ estimado})$$

donde  $\theta$  es el estadístico de interés y por lo tanto la variabilidad de dicho estadístico en el muestreo determinará el error de muestreo.

Como toda encuesta, la ECV tiene asociado errores de muestreo. Como consecuencia, las diferencias que se observan en los resultados de la encuesta en dos momentos de tiempo pueden deberse tanto a cambios reales del objeto bajo estudio como al error de muestra, es decir no toda variación observada se debe a un cambio real. En conclusión, a mayor error de muestreo de una variable, mayor será la probabilidad de que el cambio sea explicado por dicho error y no por un cambio real.

En España, el INE calcula dicho error a partir de estimaciones territoriales de las principales variables bajo estudio de la ECV, siguiendo las recomendaciones de Eurostat y del grupo de trabajo Net-SILC2, usando la linealización del error y enfoque del cluster final (ultimate cluster), el cual explica que el cálculo de la varianza de error de muestreo solo se debe a la variación entre los totales de las unidades primarias de muestreo, es decir las secciones censales. Este método es conocido como el método de expansión de Taylor.

Junto con el estadístico de interés se proporcionan otras medidas del error que son de utilidad y ayudan a la interpretación de este. Entre ellas, las más interesantes son:

- **El Coeficiente de Variación**, como una medida relativa del error que permite comparar precisiones entre distintos grupos o poblaciones.
- **Intervalo de Confianza al 95%**, basado en la distribución en el muestreo del estadístico.
- **Error relativo al 95% de confianza**, que se obtiene al multiplicar el percentil 1,96 por el Coeficiente de Variación. Este error relativo nos permite hablar en términos de puntos porcentuales del valor de la estimación.

## 2 Definiciones e Indicadores asociados a la encuesta

La ECV pretende poner de relieve los principales indicadores estadísticos utilizados que nos permiten obtener conclusiones a partir de dicha encuesta, así como algunas de sus propiedades más relevantes y discutir sobre los valores obtenidos en un conjunto de datos reales.

La información que se recoge en la Encuesta sobre los hogares y personas que las componen queda detallada en los cuestionarios anteriormente citados. [[Sección 1.2.3](#)]

Para alcanzar los objetivos según los reglamentos europeos, se necesita definir una serie de conceptos que veremos en los siguientes apartados.

### 2.1 Definiciones

#### 2.1.1 Vivienda y tipos

Una **vivienda** es un recinto cuya funcionalidad es la de ser habitado por personas o se constituye como la residencia habitual de alguna persona. No se considera recinto aquel que está totalmente dedicado a otro fin (por ejemplo, locales).

Podemos identificar dos tipos de vivienda:

- **Viviendas colectivas;** las cuales están destinadas a ser habitadas por un grupo de personas sometidas a una autoridad o régimen, sin lazos familiares ni de convivencias.
- **Vivienda familiar;** las cuales están destinadas a ser habitada por una o varias personas sin la necesidad de estar unidas por parentesco, y que no constituyen un hogar colectivo (se verá posteriormente su significado).

Los **alojamientos** son un tipo de vivienda familiar con la particularidad de ser móvil o improvisada, o bien que no había sido destinada en un principio a fines habitables, pero constituye la residencia de una o varias personas.

Los alojamientos pueden ser: **Fijos**, constituidos por las viviendas semipermanentes, por chabolas y chozas (recintos destinados a ser una vivienda sin ningún pilar y construido con materiales de desecho) y por otros residuos no destinados a ser una vivienda pero que a pesar de eso vive gente en su interior (molinos, garajes etc.). **Móviles**, son aquellos que han sido construidos para ser transportados o forman un recinto móvil, y sirven de residencia a una o varias personas (barcos, yates etc).

Podemos definir el concepto de vivienda familiar principal a una vivienda familiar utilizada como residencia habitual de uno o más hogares, esto quiere decir que aquellas viviendas familiares que son de temporada como una casa en la playa, las cuales no son residencia habitual, no pueden ser consideradas vivienda familiar principal.

Según la manera en la que las personas residan en una vivienda se definen tres tipos de hogares:

1. **Privado**, se define como la persona o grupo de personas que habitan en común una vivienda familiar principal y consumen o comparten alimentos u otros bien con cargo a un mismo presupuesto. Solo estos hogares son objeto de investigación en la ECV.
2. **Colectivo**, se define como el grupo de personas que reside en una vivienda colectiva no institucional como pueden ser hoteles, residencias militares, colegios mayores etc y por las viviendas compartidas por más de cinco personas independientes entre sí. En este segundo grupo de hogares también se incluyen las instituciones y está formada por el conjunto de personas que reside en una vivienda colectiva institucional (asilos, residencias de ancianos, instituciones sanitarias etc).
3. **Hogar panel**, es todo hogar privado que resida en una vivienda de la muestra.

### 2.1.2 Personas y tipos

En primer lugar, vamos a definir el siguiente concepto:

- **Miembro de un hogar privado:** Conjunto de personas que tienen la misma residencia habitual y participan de un presupuesto común. Una persona se considera **residente habitual** de una vivienda si la tiene como domicilio y pasa la mayor parte de su descanso diaria. Además, en caso de convivir con otros, debe compartir presupuesto ya que si no lo hace constituye otro hogar dentro del mismo domicilio

A continuación, se establece una clasificación de las personas que tienen relevancia dentro de esta encuesta:

- **Persona panel y corresidente**, son aquellos miembros de un hogar panel en periodo 1, que son mayores de 13 años el 31 de diciembre del año anterior al de la entrevista. El resto de las personas (menores de 14) miembros de un hogar panel son corresidentes.

- **Persona encuestable**, es aquello que es miembro del hogar y tiene 16 años o más el 31 de diciembre del año anterior a la entrevista. En caso contrario se considera menor y por lo tanto no puede cumplimentar el cuestionario individual.
- **Antiguo miembro**, hace referencia a las personas que actualmente no son miembros de un determinado hogar ni fue registrado como miembro de dicho hogar en la encuesta anterior, pero ha residido en dicho hogar durante un mínimo de tres meses a lo largo del período de referencia de los ingresos.
- **Persona responsable de la vivienda**, es aquel miembro del hogar que tiene a su nombre la propiedad o el contrato de arrendamiento. Si el hogar se ha cedido, el responsable será la persona a quien se le haya sido cedido. Si dos personas comparten responsabilidad la mayor es considerada primera responsable y si es compartida la responsabilidad por más de dos personas, las dos mayores son los responsables.
- **Informante del hogar**. El informante idóneo para ser entrevistado y que debe dar la información general del hogar es la persona responsable de la vivienda, en el caso de que haya más de una, debe ser alguna de ellas. Si dicha persona no fuese posible que contestara a la encuesta, debe ser alguien que pueda hablar en nombre del hogar. Existen ciertos casos, donde se permite que el informante sea una persona no-miembro del hogar como puede suceder en hogares formados por personas mayores o incapacitados que viven solos.

### 2.1.3 Otras definiciones

Como bien hemos dicho en el resumen y en la introducción, los objetivos fundamentales que tiene la encuesta son el estudio de la distribución de la renta y el riesgo de pobreza y de la exclusión social, así como la desigualdad. Además, se recogen información sobre los ingresos de los hogares privado, la distribución de sus miembros, las fuentes de ingresos, las condiciones de la vivienda etc.

Para alcanzar estos objetivos de acuerdo con los reglamentos europeos definidos, se definen una serie de conceptos que nos va a servir de ayuda conocer de antemano.

#### 2.1.3.1 Ingresos brutos

Renta total, monetaria y no monetaria, percibida por el hogar durante el período de referencia, antes de deducir el impuesto de la renta, los impuestos sobre el patrimonio y las cotizaciones de los asalariados, autónomos, y desempleados (en su caso) en la Seguridad Social obligatoria. Incluye las cotizaciones sociales de los empresarios en la Seguridad Social, así como las transferencias recibidas entre los hogares.

### *2.1.3.2 Ingresos netos*

Renta bruta menos el impuesto sobre la renta, los impuestos sobre el patrimonio y las cotizaciones obligatorias de los asalariados, autónomos y desempleados (en su caso) en la Seguridad Social obligatoria, las de los empresarios y las transferencias enviadas entre hogares. Por su parte, en el cálculo de las rentas netas del hogar se incorporan las rentas procedentes de esquemas privados de pensiones no relacionados con el trabajo.

NOTA: HAY QUE TENER EN CUENTA QUE EL PERÍODO TEMPORAL DE REFERENCIA SOBRE LOS INGRESOS DE LOS HOGARES ES EL DEL AÑO ANTERIOR AL DE LA ENCUESTA. EN EL CÁLCULO DE LAS ESTADÍSTICAS E INDICADORES SOBRE LOS OBJETIVOS (POBREZA, RENTA, DESIGUALDAD) QUE TENEMOS SE UTILIZA SIEMPRE LA RENTA NETA.

### *2.1.3.3 Renta por persona*

Renta que se calcula dividiendo la renta neta/sucia total del hogar por el número de miembros.

### *2.1.3.4 Renta por unidad de consumo*

Renta que se obtiene dividiendo la renta neta/sucia total del hogar entre el número de unidades de consumo. Una vez calculada la renta por unidad de consumo del hogar se adjudicará este valor a cada uno de sus miembros.

### *2.1.3.5 Unidad de consumo*

Para una mejor comparabilidad de los ingresos de los diferentes tipos de hogares se utiliza internacionalmente el concepto unidad de consumo. El número de unidades de consumo de cada hogar se calcula utilizando la escala de la OCDE modificada, que es una escala de equivalencia que convierte el número de miembros residentes en el hogar en número de unidades de consumo, que asigna un peso de 1 en el primer adulto, de 0,5 en el resto de los adultos y de 0,3 en los menores de 14 años.

### *2.1.3.6 Umbral de pobreza*

Nivel de ingreso mínimo necesario para adquirir un adecuado nivel de vida. El umbral de riesgo de pobreza por unidad de consumo se fija como el 60% de la mediana de los ingresos netos por unidad de consumo del conjunto de individuos de una sociedad. La mediana es el valor que, ordenando todos los individuos de menor a mayor ingreso, deja una mitad por debajo de ese valor y la otra mitad por encima. El umbral de pobreza se establece en euros.

El valor del umbral de pobreza de cada hogar depende del tamaño del hogar y de las edades de sus miembros. Se calcula multiplicando el umbral por unidad de consumo por el número de unidades de consumo del hogar.

### *2.1.3.7 Población en riesgo de pobreza*

Proporción de la población que vive en un hogar con una renta por debajo de lo que se considera nivel de vida medio en la sociedad en la que se reside. Un hogar (y por tanto todas las personas que forman parte de este) está en situación de riesgo de pobreza cuando su renta neta por unidad de consumo es inferior al umbral de pobreza.

### *2.1.3.8 Condiciones de la vivienda*

Circunstancias que caracterizan a la vivienda, como el régimen de tenencia, los gastos medios mensuales y los equipamientos, entre otros.

- Propiedad: cuando algún miembro del hogar es el propietario de la vivienda.
- Alquiler: cuando algún miembro del hogar satisface para su uso una cantidad en metálico o en especie en calidad de alquiler, ya sea a precio de mercado o inferior, e independientemente de que exista o no un contrato.
- Cesión gratuita: cuando la vivienda la cede la empresa u organización en la que trabaja algún miembro del hogar, u otros hogares o instituciones.
- Gastos totales de la vivienda: incluyen el alquiler (en caso de que la vivienda esté alquilada), los intereses de la hipoteca (en caso de que la vivienda sea de propiedad con pagos hipotecarios pendientes) y otros gastos asociados al mantenimiento de la vivienda (comunidad, agua, electricidad, gas, seguro del continente y algunas tasas e impuestos municipales).
- Capacidad para llegar a fin de mes: recoge la percepción de los entrevistados de su dificultad o facilidad para llegar a fin de mes. Esta medida es de naturaleza subjetiva.

### *2.1.3.9 Alquiler imputado*

Componente no monetario de los ingresos del hogar. Se puede aplicar a los hogares que no pagan un alquiler completo por ser propietarios o por tener una vivienda alquilada a un precio inferior al del mercado o gratuito. El valor que se imputa es el equivalente a lo que se pagaría en el mercado por una vivienda similar a la ocupada, menos cualquier alquiler realmente abonado.

Se obtiene combinando la cantidad obtenida mediante un método objetivo con el alquiler subjetivo que proporciona el informante del hogar. El método objetivo consiste en imputar la media del alquiler real de las viviendas con otras de características similares en cuanto a número de habitaciones, tipo de vivienda, grado de urbanización de la zona y tiempo de ocupación de la vivienda. Asimismo, se deducen de los ingresos totales del hogar los intereses de los préstamos solicitados para la compra de la vivienda principal.

## 2.2 Pobreza y Desigualdad

Además de estas definiciones, como en este trabajo nos vamos a centrar más en la pobreza, la desigualdad y exclusión social, vamos a definir brevemente estos conceptos. (Llera Lasobras, 2019)

**Pobreza:** Existen diversas definiciones de la pobreza dadas por diferentes autores a lo largo del tiempo. En 1901, se definió el concepto de pobreza básica y viene determinada cuando los ingresos totales no son suficientes para poder hacer cargo de todas las necesidades indispensables para mantener nuestra salud y físico de manera eficiente.

En 1995, el consejo europeo definió un concepto general de la pobreza, indicando que las personas, familias y grupos de personas pobre son aquellos que están excluidos de las necesidades mínimas de consumo y actividades de vida aceptable debido a la escasez de recursos.

Así mismo, se definió la **pobreza económica**, que hace referencia a los individuos que no alcanzan la libertad suficiente para satisfacer el hambre, conseguir nutrientes suficientes, vestir dignamente, tener una vivienda aceptable etc.

Por último, debemos tener en cuenta que la pobreza se puede clasificar de diversas maneras según los datos que se tengan y la manera en que vayan a ser analizados:

- **Pobreza objetiva:** se define utilizando información sobre variables objetivas de los hogares, como el nivel de ingresos o gastos, o los equipamientos de la vivienda.
- **Pobreza subjetiva:** se define a partir de la visión de los habitantes del hogar, de su situación y sus necesidades.
- **Pobreza absoluta:** falta de bienes y servicios básicos y primordiales para vivir; la alimentación, vivienda o vestido.
- **Pobreza relativa:** se define comparando el nivel medio de una población. Esta medida se establece a partir de los ingresos o gastos, y supone una carencia relativa respecto a la disponibilidad de bienes habituales.

Más adelante se introdujo el termino de exclusión social para establecer las desigualdades entre la población.

En 2016, se definió la **exclusión social**, como la ruptura del vínculo social que se asocia a la privación de los derechos sociales, es decir ciertos individuos son empujados al margen de la sociedad y se les impide participar plenamente debido a su pobreza, falta de competencias básicas etc.

En 2002, se determinaron tres grados de exclusión social, dependiendo de la gravedad de la situación y actitud de la persona en cuestión.

- **Exclusión de 1º grado (pobres):** situación donde la vida del individuo está en peligro: pobreza extrema, personas sin techo, infancia desprotegida.
- **Exclusión de 2º grado (marginados):** situaciones que pueden derivar a la exclusión de primer grado o pueden subsanarse. Forman parte los desempleados, ancianos, habitantes de pueblos rurales con pocas oportunidades, enfermos de SIDA, discapacitados, minorías étnicas, inmigrantes.
- **Exclusión de 3º grado (desviados):** la actitud de la persona y las acciones que ejecuta contribuyen a su exclusión (ruptura de normas morales o legales). Se incluye a los reclusos, alcohólicos, drogodependientes, prostitución.

La **desigualdad social** fue un término que se incluyó más adelante para poder vincular a la sociedad en función de la falta de equilibrio entre dos o más personas. Según el sociólogo, Eduardo López Aranguren, es el resultado de los problemas sociales, y no puede observarse únicamente como un fenómeno natural.

Es de gran importancia destacar que varios sociólogos indican que la desigualdad proviene del estado natural del ser humano y se hace hincapié en que el ser humano se compara con sus semejantes y ve sus diferencias, y entonces es cuando se pierde la igualdad del ser humano. Hoy en día, la desigualdad está implementada en nuestra sociedad a gran escala y se debe a que ha sido establecido estructuralmente debido a los sistemas políticos y económicos, de ahí la dificultad para que sea combatida.

## 2.3 Principales indicadores

En este apartado vamos a hacer hincapié en los indicadores de calidad de vida. En mayo de 2013, Eurostat publicó por primera vez una sección explicando un primer conjunto de Indicadores de calidad de vida. Posteriormente en 2017 se amplió este grupo de Indicadores de calidad de vida, donde dicha información viene recogida al completo en el Grupo de Trabajo de Expertos de Calidad de Vida. La cantidad de indicadores que se incluye en dicha publicación es enorme, por lo que en este trabajo incluiré los más relevantes para el análisis de algunas dimensiones. (Grupo de Expertos de Calidad de Vida, 2017)

Existen 9 dimensiones posibles para los indicadores existentes:

- Condiciones materiales de vida

- Trabajo
- Salud
- Educación
- Ocio y relaciones sociales
- Seguridad física y personal
- Gobernanza y derechos básicos
- Entorno y medioambiente
- Experiencia general de la vida.

Todos los indicadores existentes en estas 9 dimensiones se obtienen fundamentalmente tanto de la ECV como de la Encuesta de Población Activa (EPA).

Según el informe del Comité del Sistema Estadístico Europeo Sponsorship Group, estos indicadores que nos permiten conocer y medir la calidad de vida deben cumplir ciertos principios:

- Medir resultados, pero no los procesos intermedios.
- La recogida de datos debe estar basada en observaciones individuales, siendo inválido las observaciones agregadas.
- Medir desigualdades a través de estadísticos de distribución.
- Detectar desigualdades en diferentes grupos de población.

Estos indicadores de calidad de vida están generalizados a toda la población, aunque existen algunas limitaciones para disponer de información debido a las características de la ECV, como puede ser para la población menor de 16 años.

Además, los indicadores engloban todo el territorio nacional y sus desagregaciones geográficas por Comunidades Autónomas.

Los datos sobre estos Indicadores son extraídos de manera anual comenzando en 2004, año en el que se inició la ECV, fuente principal de información sobre los indicadores de calidad de vida. Debido a los cambios que la ECV presento en 2013, la serie de datos existente sobre algunos de estos indicadores sufrió una ruptura. Para algunos indicadores se han construidos series temporales anuales desde 2004 hasta 2012 y otra serie desde 2008 hasta 2020. Debemos tener en cuenta que estas series no son comparables.

### 2.3.1 Lista de Indicadores

A continuación, se van a citar los indicadores de calidad más importantes de la ECV y si es conveniente realizar algún comentario sobre ellos.

**Renta disponible del hogar:** Es la suma de ingresos en los miembros del hogar.

$$Renta\ disponible\ del\ hogar = \sum_{i=1}^n ingresos_i\ siendo\ i = miembros\ del\ hogar$$

**Renta media del hogar:** Es calculado mediante la división de la renta neta del conjunto de los hogares y el número total de hogares. Esta definición incluye el alquiler imputado, que es aplicado a los hogares en los que no se paga un alquiler completo. Se imputa el equivalente al alquiler que se pagaría en el mercado por una vivienda de similares características.

$$Renta\ media\ del\ hogar = \frac{Renta\ neta\ del\ conjunto\ de\ hogares}{Número\ total\ de\ hogares}$$

**Renta media por persona:** Igual que el anterior, pero siendo dividido por el número total de personas.

$$Renta\ media\ por\ persona = \frac{Renta\ neta\ del\ conjunto\ de\ hogares}{Número\ total\ de\ personas}$$

**Renta media por unidad de consumo:** Igual, pero se realiza la división entre el número total de unidades de consumo que componen dichos hogares.

$$Renta\ media\ por\ U.C = \frac{Renta\ neta\ del\ conjunto\ de\ hogares}{Número\ total\ de\ U.C}$$

**Umbral de pobreza:** El umbral de riesgo de pobreza es calculado cada año a razón de la distribución de ingresos del año anterior. Este umbral se fija en el 60% de la mediana de los ingresos por unidad de consumo de las personas. El valor de este umbral depende del tamaño del hogar y la edad de sus miembros. Además, nos permite calcular la tasa de riesgo de pobreza.

**Tasa de Riesgo de pobreza:** que corresponde al porcentaje de personas que vive en hogares considerados riesgos de pobreza, es decir es el porcentaje de personas que está por debajo del umbral de pobreza.

% personas por debajo del umbral de pobreza

**Brecha o intensidad del riesgo de pobreza:** Muestra la distancia entre la mediana de los ingresos de la población en riesgo de pobreza y el umbral de pobreza. Nos proporciona información sobre si las personas tienden a acercarse o distanciarse del umbral. Este indicador se expresa en términos porcentuales con relación al umbral.

**Población con privación material:** Se entiende por concepto de privación material a la carencia de las condiciones de la vivienda, la disponibilidad de los bienes duraderos, los retrasos en los pagos y la falta de capacidad para hacer frente a las necesidades materiales básicas. De esta manera, la población con privación material es aquella que representa la privación de algunos de los ítems siguientes:

- Poder pagar sin retrasos gastos relacionados con la vivienda (hipoteca o alquiler, recibos de gas, comunidad...) o de compras aplazadas.
- Poder ir de vacaciones al menos una semana al año.
- Poder comer una carne, pollo o pescado (o el equivalente vegetariano) al menos cada dos días.
- Poder afrontar gastos imprevistos.
- Poder permitirse un coche.
- Poder mantener la vivienda a una temperatura adecuada.
- Poder permitirse un ordenador personal.

**Población con privación material severa:** Población que no puede cumplir 4 o más ítems de la lista anterior añadiéndole dos ítems:

- Poder permitirse un teléfono.
- Poder permitirse una lavadora.

**Intensidad de trabajo del hogar:** Se calcula mediante la división de la suma total de los meses en los que cada miembro del hogar en edad de trabajar durante el año anterior a la encuesta entre la suma total de los meses en las que cada miembro ha podido trabajar durante dicho año. La edad de trabajar está formada por los individuos de 18 a 59 años, excluyendo los de 18 a 24 años que sean estudiantes. El rango de esta medida es desde 0 a 1, donde 0 es que no se ha trabajado en dicho hogar y cercano a 1 es cuando cada individuo ha trabajado prácticamente todos los meses posibles.

**Población en hogares con intensidad de trabajo escasa:** Todos los individuos del hogar cuyos adultos tienen una intensidad de trabajo menor de 0,20.

**Tasa AROPE:** Esta tasa surgió de la necesidad de incorporar un indicador que recogiera situaciones de vulnerabilidad de manera única. Hace referencia al porcentaje de población que se encuentra en riesgo de pobreza y/o exclusión social. Define de manera agrupada a las personas que cumplen uno o más de los siguientes requisitos (indicadores definidos anteriormente):

- Se encuentra en riesgo de pobreza.
- Se encuentra en situación de privación material severa.
- Viven en hogares con intensidad de trabajo escasa.

Esta tasa no es homogénea para todas las personas y presenta importantes diferencias si es desagregada según sexo, edad, nacionalidad etc.

**S80/S20:** también denominada ratio quintil. Es el cociente de los ingresos totales del 20% de las personas con ingresos más elevados (percentil 80 y superior en la distribución de la renta) y los ingresos totales del 20% de las personas con ingresos más bajos (percentil 20 e inferior en la distribución de la renta). En el caso de que este índice sería 1 (su valor mínimo) nos indicaría que todo el mundo tiene los mismos ingresos, es decir nos indica la igualdad absoluta.

$$\frac{\sum_i^P \text{ingresos}_i}{\sum_j^C \text{ingresos}_j}$$

$P = 20\%$  personas con ingresos más elevados

$C = 20\%$  personas con ingresos más bajos

### 2.3.2 Indicadores de desigualdad

A continuación, se explicará con mayor detalle los indicadores de desigualdad que son utilizados en la ECV. Algunos de estos indicadores serán objetivo de estudio de este trabajo, así como de realizar ciertos análisis estadísticos que permitan introducir el conocimiento e interpretación de dichos indicadores.

No se compararán de manera numérica el valor obtenido para estos coeficientes debido a que los valores no son comparables entre sí, y por lo tanto carece de sentido hacer la comparación. Lo que si podremos realizar y observar será la evolución de estos indicadores de desigualdad de manera global o de manera desagregada para cada uno de los indicadores ya que las evoluciones entre ellos sí que resultan comparables.

#### 2.3.2.1 Curva de Lorenz

En primera instancia, vamos a hablar sobre la curva de Lorenz, la cual es una herramienta muy valiosa tanto en el análisis de la distribución de la renta sobre una cierta población tanto como al tratar de ordenar en términos de desigualdad.

En 1905, M.O. Lorenz propuso estas curvas para analizar la desigualdad de la renta y riqueza, destacando que el abombamiento de dicha curva es un indicador de desigualdad existente en la distribución.

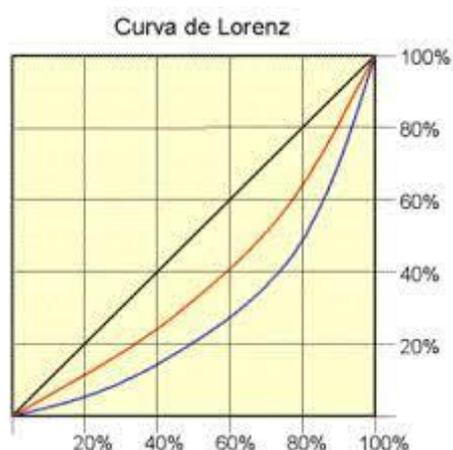
La curva de Lorenz se construye de la siguiente manera: Sea  $x$  una distribución de renta en el espacio  $D$ , a partir de ella se construyen los porcentajes acumulados de individuos y de rentas repartidas, teniendo en cuenta que los vectores tienen sus vectores ordenados de forma ascendente, son no negativas y denotando a la media aritmética por  $\bar{x}$ :

$$p_0 = 0; p_i = \frac{i}{N}, \text{ con } i = 1, 2, \dots, N$$

$$q_0 = 0; q_i = \frac{1}{N \cdot \bar{x}} * \sum_{j=1}^i x_j \text{ con } i = 1, 2, \dots, N$$

Una vez obtenido estos porcentajes, se obtiene la curva de Lorenz denotada por  $L(p)$ , que es la poligonal que une los puntos del conjunto  $(p_i, q_i \text{ con } i = 0, 1, \dots, N)$ .

De manera más sencilla es una representación gráfica de la desigualdad en el reparto de la renta existente en un determinado territorio (país, CC. AA, ciudad etc.) y en ella se sitúa en el eje X los acumulados de población y en el Y los acumulados de renta. Obviamente está inscrita en el cuadrado de la unidad, de manera que la situación de igualdad viene determinada por la cercanía de la curva a la diagonal del cuadrado unidad donde esta curva es inscrita y el abombamiento indica un aumento de la desigualdad, es decir cuanto mayor sea el área entre la recta de la diagonal y la curva mayor será la desigualdad. Veamos un gráfico de ejemplo que nos permita ver lo citado:



Aquí podemos observar la diagonal y dos curvas de Lorenz, donde la curva naranja es más cercana a la línea de igualdad y la azul indica un aumento de la desigualdad al tener mayor área entre dicha línea y la curva. Con otras palabras, la población representada por la línea azul tiene una distribución de la renta más desigualada que la población representada por la línea naranja.

Esta definición que hemos aportado a la curva de Lorenz es descriptiva y puede generalizarse al caso de que la renta sea una variable aleatoria  $X$ , no negativa, con esperanza matemática  $\mu$  y cuya función de distribución es  $F(x)$ :

$$p = F(x) = \int_0^x dF(t)$$

$$q = L[F(x)] = \frac{1}{\mu} * \int_0^x t * dF(t)$$

La curva de Lorenz presenta varias propiedades, donde cabe destacar que la pendiente de  $L(p)$  viene determinada por la siguiente función en el caso de ser derivable:

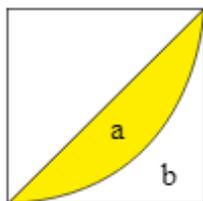
$$t(p) = \frac{F^{-1}(p)}{\mu}, p \in [0,1]$$

En el aspecto estadístico, sean  $x$  e  $y$  dos distribuciones de renta del espacio  $D$ , entonces se dirá que  $x$  es menos desigual que  $y$  en el sentido de Lorenz cuando la curva de Lorenz encierra completamente a la de  $x$ . (Iritani y Kuga,1983).

Además, a partir de la curva de Lorenz podemos calcular varios indicadores adicionales como puede ser el índice de Gini.

### 2.3.2.2 Índice y Coeficiente de Gini

En 1912, C. Gini propuso el coeficiente de Gini con la finalidad de medir la desigualdad a partir de la diferencia media entre las rentas de la distribución. Se calcula como el cociente entre el área comprendida entre la diagonal de perfecta igualdad y la curva de Lorenz ( $A$ ) y el área del triángulo creado a partir de la propia diagonal ( $A+B$ ).



$$\text{Coeficiente de Gini} = \frac{A}{(A + B)}$$

Este coeficiente toma valores en el intervalo  $[0,1]$ , donde en situaciones de máxima igualdad de distribución de la renta en la población valdría 0, lo que significaría que la

curva de Lorenz sería la misma diagonal. En cambio, a medida que aumenta la desigualdad, es decir la curva de Lorenz se aleja de la diagonal, el coeficiente de Gini se acercaría a 1 y nos estaría dejando ver que la desigualdad es grande, es decir los individuos más pobres disfrutan de una distribución muy baja y los más ricos de una distribución de renta muy alta. Podemos decir que el Coeficiente de Gini es un indicador inverso, es decir a mayor valor de dicho indicador mayor es la desigualdad.

El índice de Gini, no es más, que el coeficiente de Gini expresado en porcentajes, es decir, es igual al Coeficiente de Gini multiplicado por 100.

$$\text{Indice Gini} = \text{Coeficiente Gini} \times 100$$

Hay problemas en la interpretación de dicho coeficiente, ya que el mismo valor puede obtenerse de diversas curvas de Lorentz de distribución y debe tenerse en cuenta la estructura demográfica de la población que estemos analizando y obtengamos los valores.

Un enfoque alternativo es definir el coeficiente de Gini como la mitad de la diferencia media absoluta relativa, que es matemáticamente equivalente a la definición basada en la curva de Lorenz.

$$\text{Coeficiente Gini} = \frac{(\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|)}{2n^2 \bar{x}}$$

Donde n es el tamaño de la población y  $x_i$  es la renta del individuo i y  $\bar{x}$  es el promedio de la renta en dicha población.

Si la distribución de la renta en la población se da como una función de probabilidad continua  $p(x)$ , entonces el coeficiente de Gini es de nuevo la mitad de la diferencia absoluta de medias relativa:

$$\text{Coeficiente Gini} = \frac{1}{2\mu} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x)p(y)|x - y|dx dy$$

Donde  $\mu$  es la media de la distribución de la renta, y los límites inferiores de integración pueden ser reemplazados por cero cuando los ingresos de toda la población sean positivos.

Este indicador es el más usado debido a su calidad de medición y su facilidad tanto para el cálculo como la interpretación. Su principal problema es que no cumple con la propiedad de Descomposición Aditiva, es decir en el caso de las comunidades autónomas, si calculamos la media del coeficiente de Gini para las regiones no

obtendremos el valor de dicho coeficiente para el conjunto estatal. El coeficiente de Gini es estrictamente S-cóncavo

### 2.3.2.3 Índice de Theil

Esta medida de desigualdad fue propuesta por Theil en 1967 y fue derivada de la noción de entropía de la teoría de la información. Esta medida es diferente al resto. Sea  $X$  la probabilidad de que ocurra un determinado evento, el contenido de la información  $h(x)$  denota que el evento bajo estudio y debe ser una función decreciente de  $x$ , es decir, cuanto más improbable sea que ocurra dicho evento más interesante será conocer que dicho evento ha ocurrido realmente.

Una fórmula que verifica esta propiedad y que es utilizada con bastante frecuencia es la función logaritmo del inverso de  $X$ :  $h(x) = \log(1 \div x)$

La entropía de dicho evento puede denotarse como la suma de los contenidos de información  $h(x)$  de cada evento ponderado según sus correspondientes probabilidades:

$$H(x) = \sum_{i=1}^n \log\left(\frac{1}{x_i}\right)$$

La entropía generalmente fue diseñada para medir el desorden, pero en este caso se interpreta como una medida de desigualdad, ya que si  $X$  se interpreta como la parte de ingreso que va a un individuo entonces  $H(x)$  parece una medida de igualdad cuando cada valor de  $X$  es igual a  $1/n$  siendo  $n$  el número de eventos posibles.

Si restamos la entropía de una distribución del ingreso a partir de su valor máximo que corresponde al logaritmo de los posibles eventos obtenemos el índice de Theil, que viene expresado por la siguiente fórmula:

$$T = \log(n) - H(x) = \sum_{i=1}^n x_i * \log(n * x_i)$$

Como hemos comentado, la entropía en general estudia el desorden, pero en este caso cuanto mayor sea la entropía menor será la desigualdad, por lo que tomaremos la entropía a grandes valores como algo bueno en este caso.

Además, este indicador presenta una ventaja con respecto al coeficiente de Gini y es que se puede descomponer de forma aditiva, por lo que nos informa sobre la desigualdad en distintos grupos formados. Esto lo hace ser un índice bastante

interesante ya que nos permite conocer la inequidad que se genera entre los grupos formados en porcentajes de desigualdad y qué porcentaje proviene de las diferencias de ingresos en el interior de los mismos grupos.

De manera más automática, existe otra fórmula para el cálculo del índice de theil:

$$T = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{\mu} \ln \frac{y_i}{\mu}, \text{ si } c = 1$$

Donde n, representa la población bajo estudio en este caso, la variable y corresponde al ingreso por hogar y  $\mu$  es la media de los ingresos por hogar. El termino c, nos indica el peso que es otorgado a la parte más baja de la distribución de ingresos, en el caso de  $c = 1$  es mejor para observar los cambios en hogares con menos ingresos. Con tal de que los resultados no estén afectados por el número de observaciones, se realiza el siguiente ajuste una vez se hayan calculado los valores: Índice de Theil / log(N)).

Mencionamos también el índice L de theil que se define mediante la siguiente formula:

$$T_L = \frac{1}{n} \sum_{i=1}^n \log \frac{u}{y_i}$$

Este índice nos representa la disentropía de la distribución del ingreso por persona con respecto a la entropía máxima (lo que se logra con total igualdad).

#### 2.3.2.4 Índice de Atkinson

También conocida como medida de desigualdad de Atkinson, es una medida desarrollada por el economista británico Anthony Barnes Atkinson, que mide la desigualdad en los ingresos de la población. Su principal utilidad es la posibilidad de determinar qué extremos de la distribución contribuye más a la desigualdad observada, refiriéndose a extremo de la distribución a los más ricos o pobres, es decir a los extremos del eje x en la curva de Lorentz.

Atkinson definió lo que él llama “la renta equivalente igualmente distribuida de una distribución dada de una renta total, y se define como el nivel de renta per cápita que, de ser disfrutado por todos, haría que el bienestar total fuera exactamente igual a el bienestar total generado por la distribución real de la renta.

El índice de Atkinson viene definido por la siguiente formula:

$$Atk(e) = 1 - \left( \frac{1}{n} \sum_{i=1}^n \left( \frac{y_i}{u} \right)^{1-e} \right)^{1/[1-e]}, \text{ si } e \geq 0 \text{ y } \neq 1$$

Donde las variables representan las mismas variables que en los índices previamente descritos. En este índice se introduce el coeficiente “e” que representa el rechazo a la desigualdad, es decir, cuanto mayor sea “e”, la ponderación de los ingresos será mayor en los hogares con menos ingresos y observaremos los cambios en la parte más baja de la distribución de la renta. Al igual que en el coeficiente de Gini a la hora del procesamiento de los datos las observaciones negativas pueden ser transformadas a valor 0 para mayor facilidad a la hora de realizar cálculos y conclusiones.

Para “e” = 0, (sin aversión a la desigualdad), el bienestar social marginal del ingreso es invariante al ingreso, es decir, los aumentos marginales en el ingreso producen tanto bienestar social ya sea que vayan a un individuo pobre o rico. En este caso, el ingreso equivalente de bienestar distribuido equitativamente es igual al ingreso medio, y el índice de Atkinson es cero.

Para “e” = ∞ (aversión infinita a la desigualdad) el bienestar social marginal de los ingresos del individuo más pobre es infinitamente mayor que cualquier individuo incluso ligeramente más rico, y la función de bienestar social de Atkinson es igual al ingreso más pequeño de la muestra. En este caso, el índice de Atkinson es igual al ingreso medio menos el ingreso más pequeño, dividido por el ingreso medio. Como en las grandes distribuciones típicas de ingresos los ingresos de cero o casi cero son comunes, el índice de Atkinson tenderá a ser uno o muy cercano a uno para valores grandes de “e”.

El índice de Atkinson varía entonces entre 0 y 1 y es una medida de la cantidad de utilidad social que se obtiene mediante la redistribución completa de una distribución de ingreso dada, para el parámetro “e” dado. Bajo el estándar ético utilitario y algunos supuestos restrictivos (una población homogénea y elasticidad constante de utilidad de sustitución), “e” es igual a la elasticidad ingreso de utilidad marginal del ingreso.

Este índice también cumple la propiedad de descomposición aditiva y el principio de transferencias.

### 2.3.2.5 Otros indicadores

Hay que mencionar que hay más indicadores como el de Dalton, la varianza y el coeficiente de variación...

Vamos a escribir únicamente la fórmula que nos permite calcular estos otros indicadores:

$$Var = \frac{\sum_{i=1}^n (\mu - y_i)^2}{n}$$

$$CV = \frac{Var^{1/2}}{\mu}$$

$$D = \frac{\sum_{i=1}^n U(y_i)}{nU(\mu)}$$

## 3 Análisis de los Indicadores en España

### 3.1 Serie temporal de los principales indicadores en España

Este apartado consistirá en el análisis y representación de las series temporales en base a 2013, año en el que se produjeron cambios en la obtención de los datos, por lo que mostraremos la evolución año a año que han tenido los principales indicadores de la ECV y de desigualdad en España desde el año 2008 hasta el año 2020.

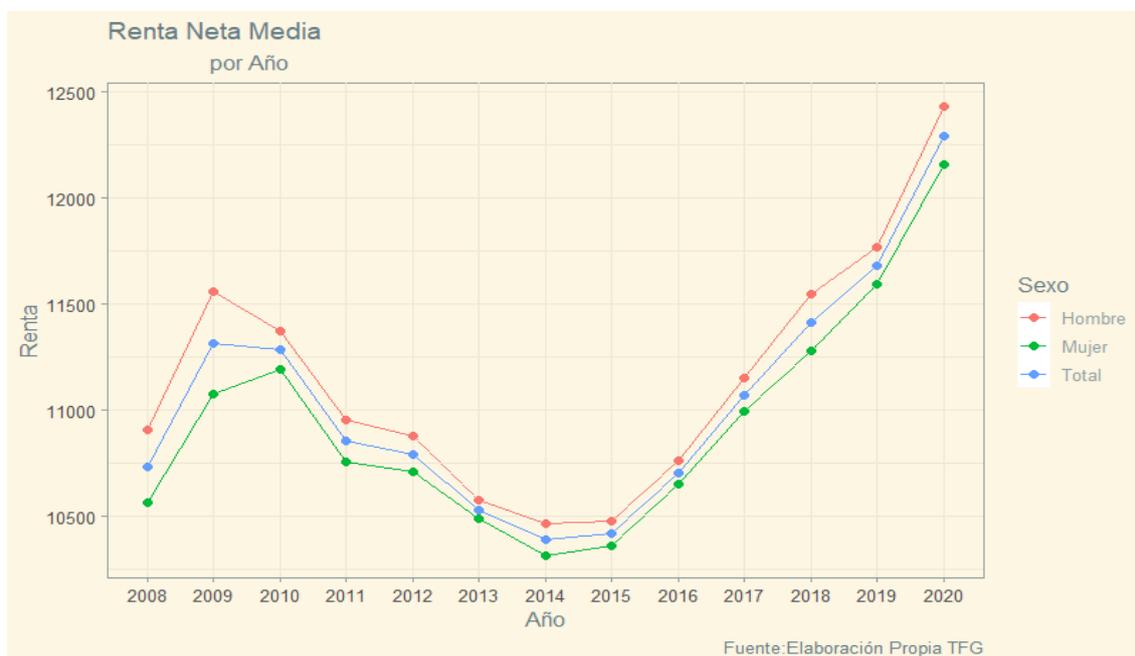
Estas series temporales de cada indicador han sido obtenidas en la página web del INE, (INE Resultado Nacionales ECV, s.f.), la cual nos permite acceder a archivos Excel con los correspondientes valores de los indicadores año tras año.

El método utilizado para su representación ha consistido en la importación de las series temporales de cada indicador en el programa R-studio, donde posteriormente se ha elaborado los gráficos de cada indicador de manera individual a través de la librería “ggplot2” (Wickham, ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics, 2022), que nos ofrece grandes mecanismos para la representación de datos que nos permitan comentar dichas series temporales.

Comenzaremos analizando la principal medida recogida en la ECV y de la cuál a partir de ella se realizará todo el estudio y los cálculos posteriores de los indicadores y tasas a estudiar, la renta (Resultados nacionales).

*NOTA: TODOS LOS PAQUETES USADOS DEL PROGRAMA R-STUDIO SE ENCUENTRAN REFERENCIADOS EN LA BIBLIOGRAFÍA.*

### 3.1.1 Renta neta media por persona



En primer lugar, podemos observar como la evolución de la renta neta media por individuo (en euros) es similar para ambos sexos, la cual se caracteriza por un incremento en los dos primeros años de la renta neta media en las mujeres mientras que en los hombres y de manera global en el segundo año disminuyó. Desde el año 2010 hasta el año 2014 la renta neta media fue disminuyendo lentamente hasta llegar a su valor más bajo con una renta neta media por individuo algo menor de 10.500 euros anuales, lo que sería equivalente a un sueldo mensual menor de 1000 euros al mes. Esta bajada seguramente tiene como causa la grave crisis que comenzó en nuestro país en el año 2007. A partir de 2015 hasta el último año del que tenemos información, 2020, la renta neta media en nuestro país ha comenzado a aumentar nuevamente, lo que parece indicar que cada año, la población española consigue generar más dinero, generando cada individuo una renta neta media de aproximadamente 12250 euros al año.

Viendo la actual tendencia creciente de nuestro país con respecto a este indicador, a simple vista podemos decir que nuestra población está teniendo mayor capacidad económica para hacer gasto de las principales necesidades humanas, aunque cuando observemos la evolución de los indicadores de desigualdad y tasas de pobreza,

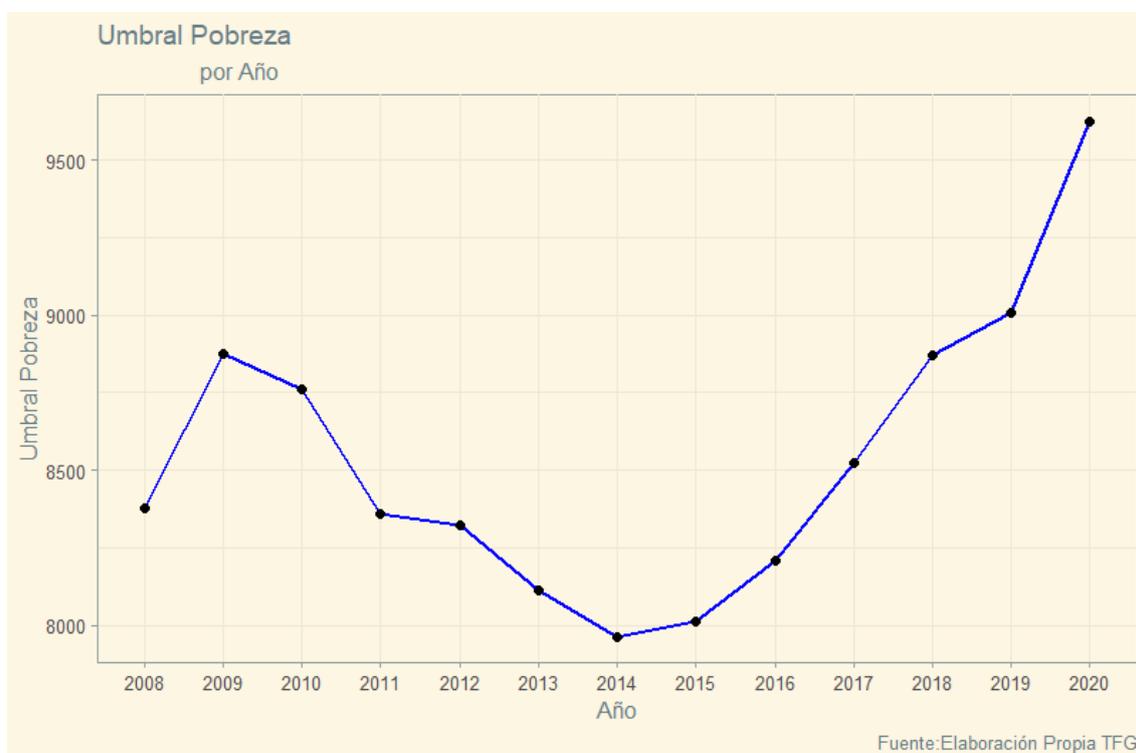
podremos detectar que el aumento de renta no es equivalente a que las condiciones de vida sean mejor en España.

Otra cosa que podemos observar de gran interés es la aún existente diferencia entre la renta neta de los hombres y mujeres, donde los hombres aún siguen obteniendo una renta neta algo mayor en nuestro país.

Si comparamos la renta neta media por individuo en nuestro país con la que obtiene un individuo en otro país europeo, seguimos estando aún por debajo de las principales potencias europeas como Alemania, Holanda etc. (Renta neta media países europeos, s.f.)

### 3.1.2 Umbral de pobreza

En el siguiente gráfico podemos observar a primera vista que la evolución que ha presentado el umbral de pobreza ha sido idéntica a la de la renta neta media anual por persona, por lo tanto, para poder analizar y comentar si la población española tiene cada año mayores problemas para poder adquirir un adecuado nivel de vida y llegar a los ingresos necesarios para ello tendremos que comparar la evolución año a año tanto de la renta como del umbral.



Calculamos la evolución de estos indicadores y mediante la librería “mmtable2” (Moran, 2021) elaboramos el siguiente gráfico.

Evolución Renta y Umbral												
con respecto año año anterior (%)												
	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
<b>renta</b>	5.41	-0.3	-3.78	-0.58	-2.45	-1.33	0.27	2.77	3.42	3.05	2.35	5.24
<b>umbral</b>	5.94	-1.28	-4.63	-0.44	-2.48	-1.88	0.62	2.47	3.82	4.09	1.56	6.85

Elaboración propia TFG

Lo más destacable que podemos observar, es cómo ambas variables tienen una relación directa, ya que cuando una aumenta la otra también lo hace y viceversa. Pero lo realmente importante que podemos detectar es que cuando la renta neta media tiende a crecer, el umbral de pobreza crece en mayor porcentaje con respecto a la renta neta media, lo que nos indica que el nivel mínimo de ingreso para poder tener un nivel de vida adecuado crece más rápidamente que los ingresos de las personas, lo cual nos permite predecir que la tendencia del número de personas en nuestro país que tienen dificultades para tener un nivel de vida adecuado está aumentando considerablemente.

### 3.1.3 Tasa riesgo de pobreza AROPE

A continuación, observaremos la evolución de la Tasa de riesgo de pobreza (AROPE) en España. Estamos representando el porcentaje de personas en España que anualmente se encuentran en riesgo de pobreza y/o exclusión social.

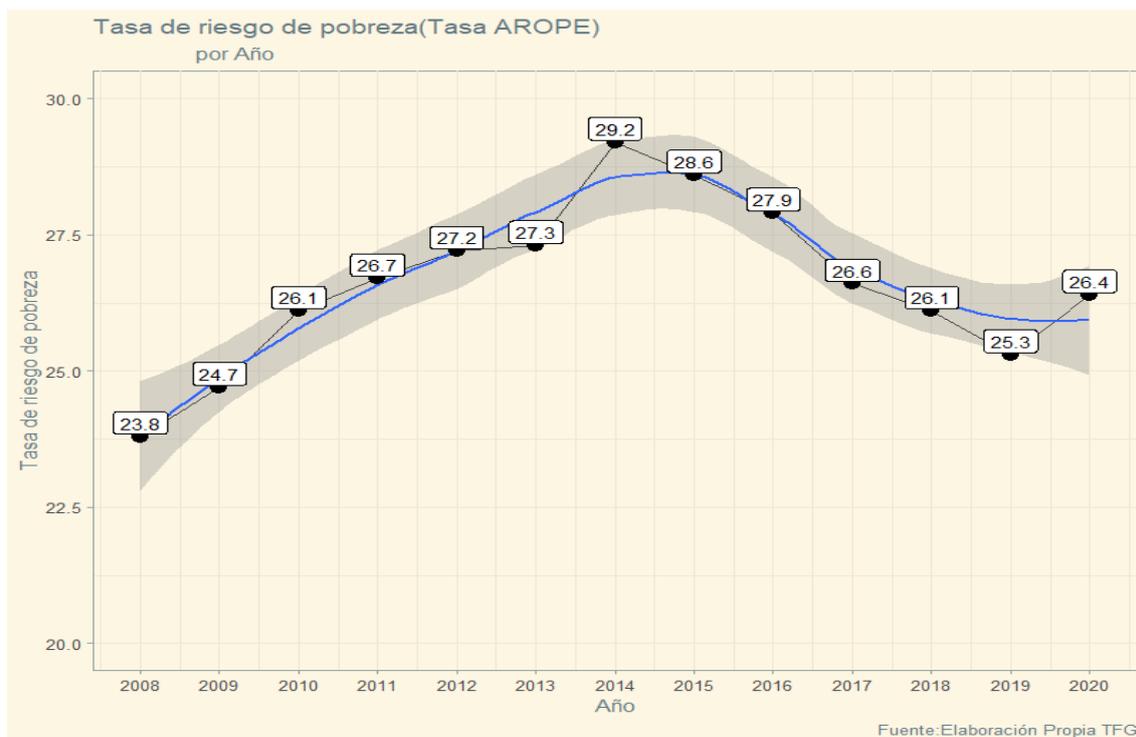


A primera vista podemos observar como la tasa de riesgo de pobreza en España durante la última década indica que en torno a una cuarta parte de la población española se encuentra en riesgo de pobreza y/o exclusión social, lo que es algo bastante elevado.

En los primeros años de esta década la evolución de esta tasa ha ido en aumento, obteniendo su mayor valor en el año 2014 con un 29% de la población en tasa de pobreza. Esto es un indicativo del enorme impacto que tuvo la crisis de 2007 en España y como año tras año repercutió en este indicador.

A partir de 2015 comenzó a verse una bajada en esta tasa, lo que comenzó a indicar que el país empezaba a recuperarse de la crisis y la pobreza en nuestro país volvía a bajar, no obstante, la disminución era lenta.

En el 2020 esta tasa volvió a subir y veremos si en el futuro esta tasa continúa creciendo debido a la pandemia Covid-19 y a la evolución que presenta el País (9 país en la UE con mayor tasa de pobreza AROPE) (Tasa AROPE países europeos, s.f.). La actual guerra desarrollada en Europa va a traer grandes consecuencias con respecto a la pobreza en nuestro país, debido al aumento de los precios en las principales necesidades de la población y la llegada de refugiados de Ucrania.



En este gráfico podemos observar la línea suavizada ajustada a los datos. Algo que nos permite interpretar lo anterior con mayor facilidad y donde podemos observar que todos estos valores entran dentro de lo esperado con respecto a su evolución ya que están dentro de los límites. Además, podemos observar como en el año 2020 parece ser que la evolución será creciente o se mantendrá constante.

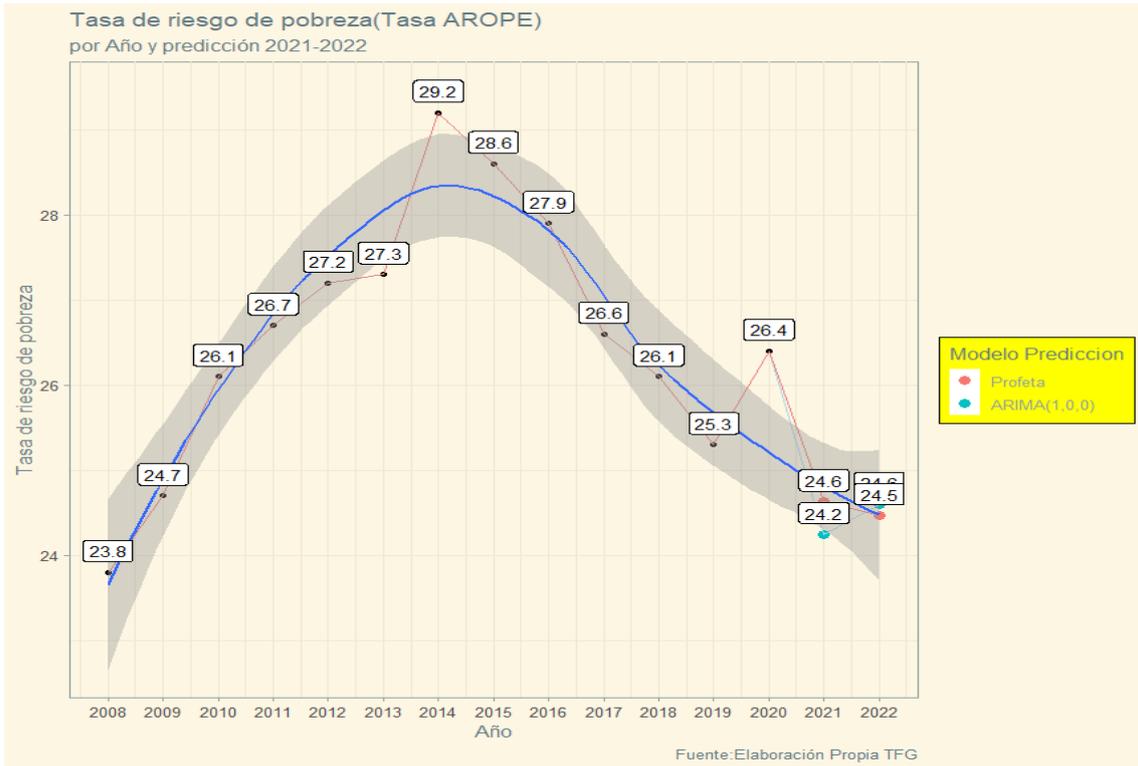
A continuación, mostraremos el mismo gráfico con la tendencia que ha seguido este indicador en dichos años y la predicción para los dos próximos años.

Hemos obtenido estas predicciones a través del uso de series temporales mediante la librería “tidymodels” de R (Kuhn & Wickham, 2019). Hemos realizado el ajuste a los datos y las predicciones mediante dos modelos diferentes: ARIMA(1,0,0) (mejor modelo recomendado por la función “auto.arima” de la librería “forecast”) y el modelo profeta que nos proporciona la librería “tidymodels”, donde podemos observar que para el año 2021 las predicciones con respecto a la tasa de riesgo en nuestro país tienden a una bajada enorme. En cambio, para el año 2022 está prevista una nueva subida realizando la predicción con el ARIMA (1,0,0) y en cambio con el modelo profeta se mantiene prácticamente igual la tasa de riesgo de pobreza.

Hay que destacar que el valor de estas predicciones nos indica que España volvería a alcanzar el menor valor de este indicador en la última década, algo que sería bastante positivo si estuviera en lo correcto.

Estas predicciones no son de gran confianza debido a que ha sido realizada únicamente teniendo en cuenta la serie temporal de la evolución de dicha tasa en los últimos años mientras que no tiene en cuenta la gran cantidad de variables que influyen en el cálculo de esta tasa y como varía en función de la situación española y europea. Por lo tanto, podemos asegurar que esta predicción no se cumplirá y que además en su efecto viendo la situación del Covid-19 y la guerra como hemos comentado anteriormente crecerá notablemente, pudiendo incluso superar el récord que se obtuvo en el año 2014.

Otra cosa que podemos destacar del siguiente gráfico es que al ajustar una línea suavizada a los datos incluyendo las predicciones, el valor de la tasa de riesgo en 2014 y 2019, son valores realmente sorprendentes, ya que no se encuentran dentro del intervalos de confianza de dicha línea.



Vamos a ver y comentar también la evolución de los componentes de esta tasa de riesgo de pobreza, pero solo desde 2013 a 2020 debido a que no disponemos de los datos de estos componentes en base a 2013 antes de dicho año.



Todos estos componentes tienen un comportamiento muy similar a la tasa de riesgo, lo cual es esperado ya que estos componentes nos dan una pista sobre la propia tasa de riesgo. Hay que destacar el gran porcentaje de población española que sufre

privación material severa, que se mantiene bastante similar a partir del año 2014 con aproximadamente un 22% de la población española que sufre este fenómeno, es decir casi  $\frac{1}{4}$  de la población española tiene carencia material de permitirse algunos de los ítems que son considerados como “necesidades humanas”.

España se caracteriza por tener pocos hogares con una intensidad de trabajo baja, en torno al 6% en la última década, y a pesar de esta característica existe un enorme % de personas en riesgo de pobreza. Es algo chocante ya que el tener una intensidad de trabajo estándar no parece ser un indicativo de tener una buena estabilidad económica y que sea suficiente para cubrir las principales necesidades.

En el 2020 a raíz del Covid-19 podemos observar el gran aumento que supuso en los hogares tener una intensidad de trabajo baja. Esto se debe principalmente al enorme número de ERTES que se realizaron e incluso despidos.

### 3.1.4 Ratio s80s20



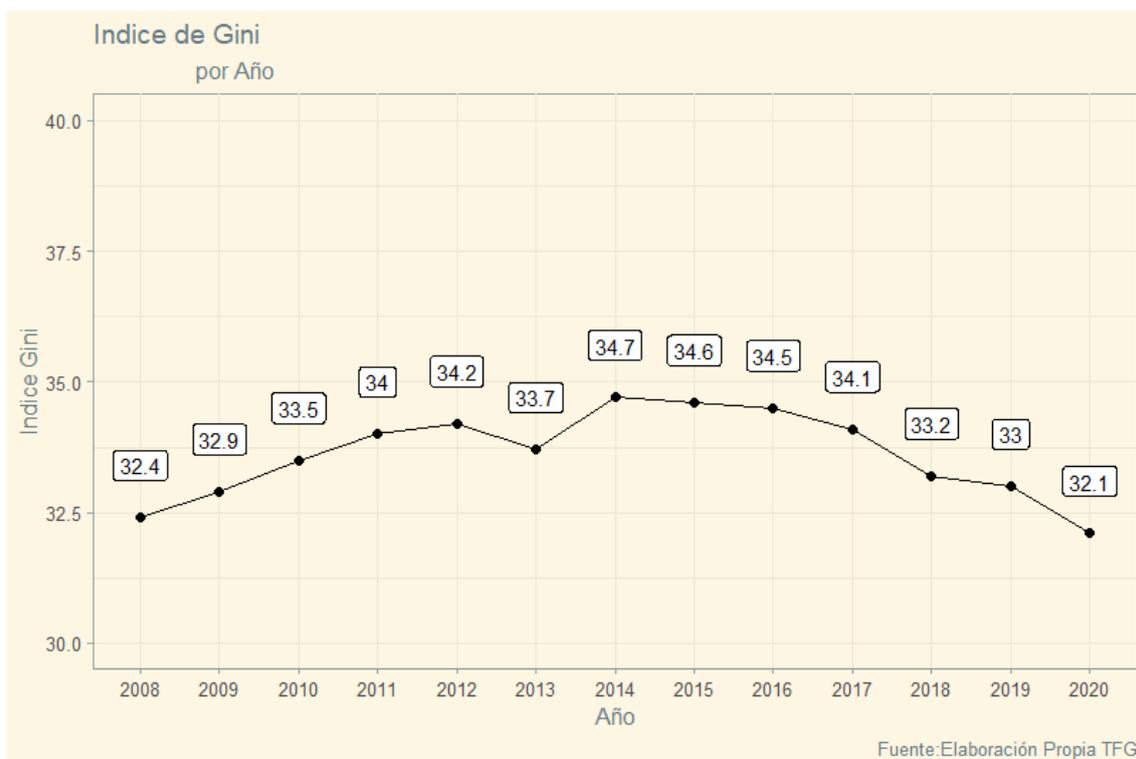
En este gráfico podemos observar la ratio s80/s20 y a simple vista nos indica, la desigualdad que existe en nuestro país a la hora de la distribución de la renta de los más ricos en comparación del 20% de los más pobres, ya que el 20% de los más ricos tienen una distribución de la renta aproximadamente 6,5 veces mayor que el 20% de

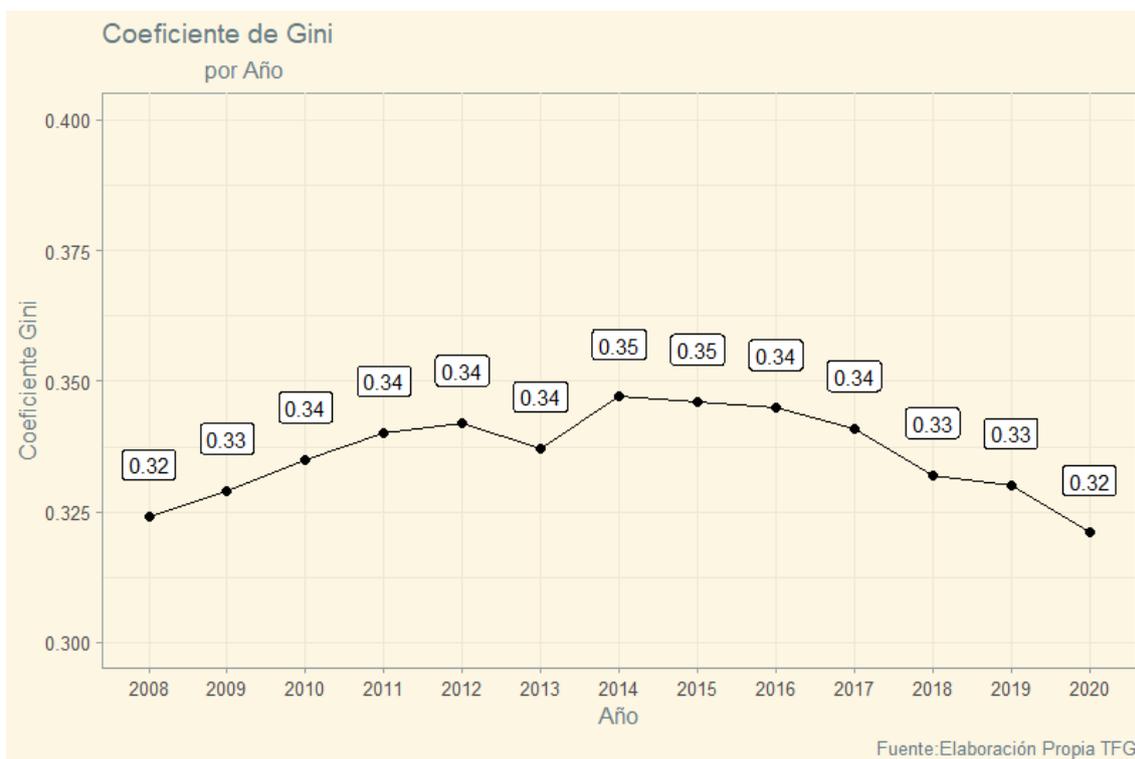
los más pobres. Cabe indicar que es bastante alto el valor de este indicador con respecto a otros países de la UE, ya que España es el noveno país con mayor ratio s80s20. (Ratio s80s20 países europeos., s.f.)

Además, en los últimos años la tendencia de esta ratio tiende a la baja, aunque con una tendencia prácticamente escasa, siendo esto algo positivo en el sentido de la igualdad de nuestra población. Obviamente no nos interesa que esta ratio se acerque lo máximo posible a 1 y que esto se deba a que los sueldos son muy bajos, ya que esto indicaría igualdad, pero también degradación de nuestros sueldos.

### 3.1.5 Índice y coeficiente de Gini

A continuación, vamos a dar una primera interpretación y visualización del que considero uno de los indicadores más importante de este trabajo, por eso lo presentare en sus dos respectivas formas, Índice de Gini y Coeficiente de Gini. La interpretación es la misma para ambos indicadores, por lo tanto, pueden observar cualesquiera de las dos gráficas que le resulte más sencillo a la hora de analizar y entender este indicador. Les recuerdo que el Índice de Gini no es más que el Coeficiente multiplicado por 100.





La evolución de este indicador es muy similar al de los anteriores, con la diferencia de que persigue dos tendencias en este intervalo de tiempo. A partir del 2008 hasta 2014, seguramente fruto de la crisis sufrida la desigualdad de la distribución de la renta en nuestro país estaba en un aumento considerable. A partir de 2014 comenzó a disminuir lentamente y en el 2017 se produjo una aceleración de esta disminución, lo que nos permite decir que la desigualdad sufrida en años previos está siendo recuperada. Esto último es algo contradictorio ya que como hemos podido observar el porcentaje de población en tasa de riesgo está aumentando en los últimos años y por lo tanto si la tasa de riesgo aumento y el coeficiente de Gini disminuye, todo cabe indicar que nuestra población es cada año más pobre, aunque los ricos sufren más la bajada de su renta en comparación con los pobres. Efectivamente esto lo vimos en el gráfico de la ratio s80/s20, de ahí que todos estos indicadores están altamente relacionados entre sí a la hora de interpretar la distribución de la renta y desigualdad que sufre nuestra población. España forma parte del “top 10” países europeos con mayor desigualdad. (Indice Gini países europeos, s.f.)

Se realizará el análisis de los principales indicadores en las distintas comunidades en el año 2020 (año más reciente del que podemos recopilar información de la ECV) y compararemos la evolución que ha tenido las comunidades autónomas españolas en los diferentes indicadores en este periodo teniendo como base de la comparación el año 2009.

## 3.2 Principales Indicadores en las Comunidades Españolas en el 2009.

En este apartado, el objetivo será exponer los valores de los principales indicadores según las comunidades autónomas de España con la misión de evaluar e identificar cuáles de ellas tienen mejores condiciones con respecto a las otras.

Para poder realizar este apartado, el mecanismo ha sido mucho más complejo que en el apartado anterior, ya que no disponemos del cálculo oficial de estos indicadores por comunidades. A partir de los micro ficheros de datos disponibles en el INE, en primer lugar, he realizado el tratamiento y depuración de los datos, seguidos posteriormente de la ordenación y cálculo de los indicadores bajo estudio desglosados por comunidades. En el anexo se incluirán los pdf proporcionados por el INE y por el EUROSTAT que nos definen las diferentes variables que forman estos micro ficheros y como se realiza a partir de ellas el cálculo de los indicadores. (Anais Santourian, 2014) y (Encuesta de Condiciones de Vida. Ficheros de usuario del año 2020 de la encuesta, 2021)

He trabajado en primer lugar con los micro ficheros del año 2009, ya que es el primer año del que podemos obtener estos micro ficheros en base a 2013. Posteriormente haremos lo mismo con los del año 2020 y poder así ver la evolución de las comunidades con respecto a los indicadores de condiciones de vida y desigualdad.

Estos micro ficheros se dividen en 4 para cada año:

- Datos básicos del hogar
- Datos básicos de las personas
- Datos detallados del hogar
- Datos detallados de las personas

Para poder tener datos reales sobre la proporción de España fue de vital importancia el factor de ponderación de las personas.

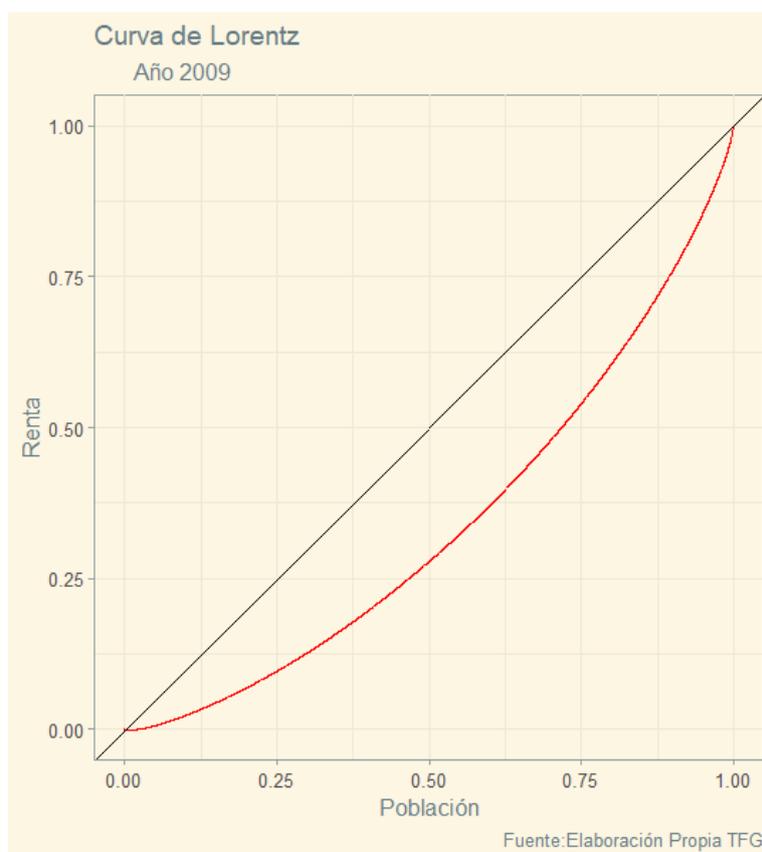
En primer lugar, representaremos el mapa en función de los valores de cada indicador en las comunidades. Esto nos permitirá tener una visión muy sencilla de interpretar y que nos permitirá detectar las comunidades más homogéneas o similares en cuanto a estos indicadores. Estos mapas han sido realizados gracias a la librería “highcharts”. (Kunst, 2022)

Por último, se ha realizado un análisis de conglomerados de las comunidades españolas en función de los indicadores calculados (renta neta media anual por persona, tasa de pobreza AROPE e índice de Gini), con el fin de obtener conglomerados que nos muestren cuales son las comunidades más homogéneas con respecto a estos indicadores y sus respectivas características. Este análisis de conglomerados se ha realizado con la ayuda de las siguientes librerías: “cluster” (Maechler, 2022) y “Nbclust” (Charrad & Ghazzali, s.f.).

Todo este análisis que ha sido explicado se ha realizado a través del programa R-studio.

### 3.2.1 Curva de Lorentz

Antes que nada, debido a que en el apartado anterior hemos reflejado la serie temporal en referencia a los indicadores, vamos a representar la curva de Lorenz con respecto a los ingresos de la población española.



A partir de la gráfica de Lorenz, la cual nos representa en el eje X el porcentaje acumulado de la población y en el eje Y el porcentaje acumulado de los ingresos

obtenidos por dicha población de entre toda la generada, podemos observar como la situación del país no es ni mucho menos cercana a la de la igualdad, ya que la curva presenta cierto abombamiento sobre la diagonal que nos representa el caso de igualdad absoluta.

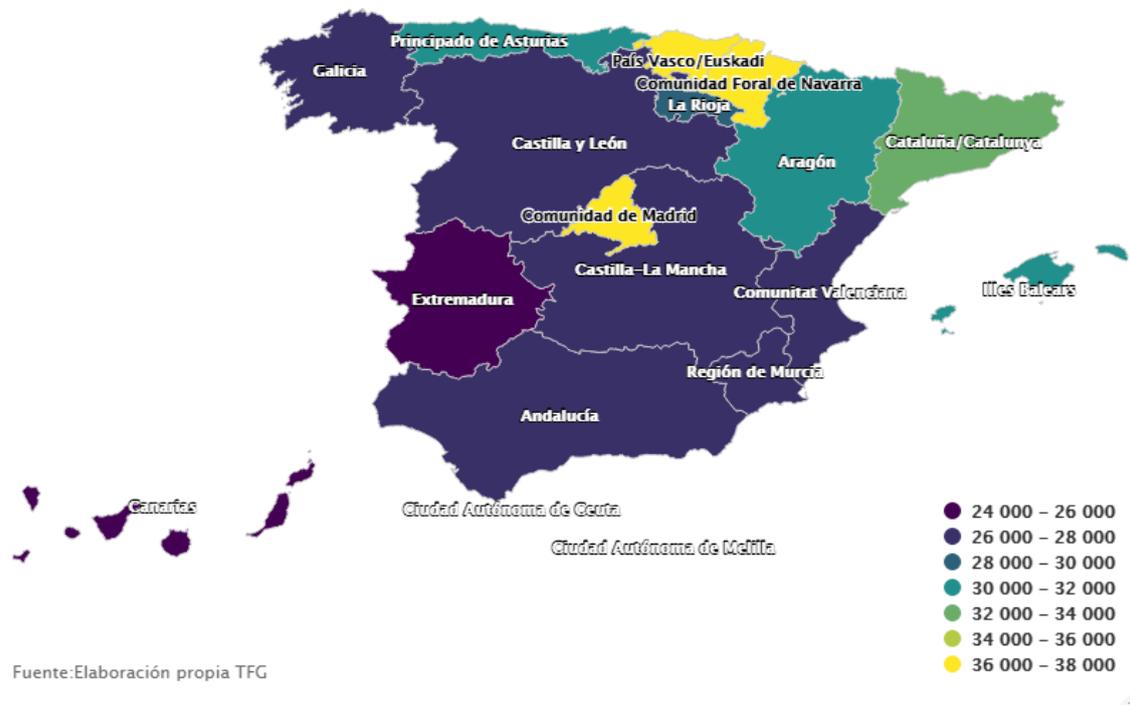
Algunos datos curiosos que podemos extraer de esta gráfica de Lorentz son los siguientes:

- En torno al 25 % de la población más rica de España genera el 50% de la renta existente en 2009.
- El 25% de la población más pobre de España obtienen únicamente aproximadamente el 10% de la renta existente en 2009.
- Si dividimos el país en dos grupos de idéntico tamaño, siendo el grupo “pobre”, el 50% de la población con ingresos más bajos, y siendo el grupo “rico”, el 50% de la población con ingresos más elevados, entonces podemos garantizar que el grupo pobre únicamente obtienen una cuarta parte de la renta generada y el grupo rico obtienen tres cuartas partes de la renta generada.

### 3.2.2 Renta neta media

En primer lugar, vamos a representar la renta neta media por hogar de cada comunidad autónoma. Cabe destacar que la Comunidad Autónoma española que percibe una renta neta media anual menor por hogar es Extremadura con 24 mil euros anuales y las que perciben una renta media mayor por hogar son el País Vasco y la Comunidad Foral de Navarra con 37,5 mil euros anuales.

## Renta neta media por hogar por Comunidad Autónoma 2009



Algunos datos interesantes que podemos observar en este gráfico son:

- Tan solo la Comunidad Autónoma de Extremadura y Canarias generan una renta neta media por debajo de 26000 euros en el año 2009
- La mayoría de las Comunidades Autónomas generan entre 26 mil euros y 32 mil euros por hogar en el año 2009.
- Se puede observar cómo existe una tendencia de mayor renta neta media por hogar en las Comunidades Autónomas del Norte, mientras que las Comunidades Autónomas más peninsulares y del sur están por debajo.
- Tan solo el País Vasco y Navarra están al mismo nivel que la capital española, Madrid.

A continuación, vamos a mostrar el número de personas que residen por hogar en cada comunidad autónoma, simplemente como ayuda para visualizar posteriormente la renta neta media por persona y la variación que genera dicho número de individuos entre la renta neta media por hogar y por persona en cada comunidad.

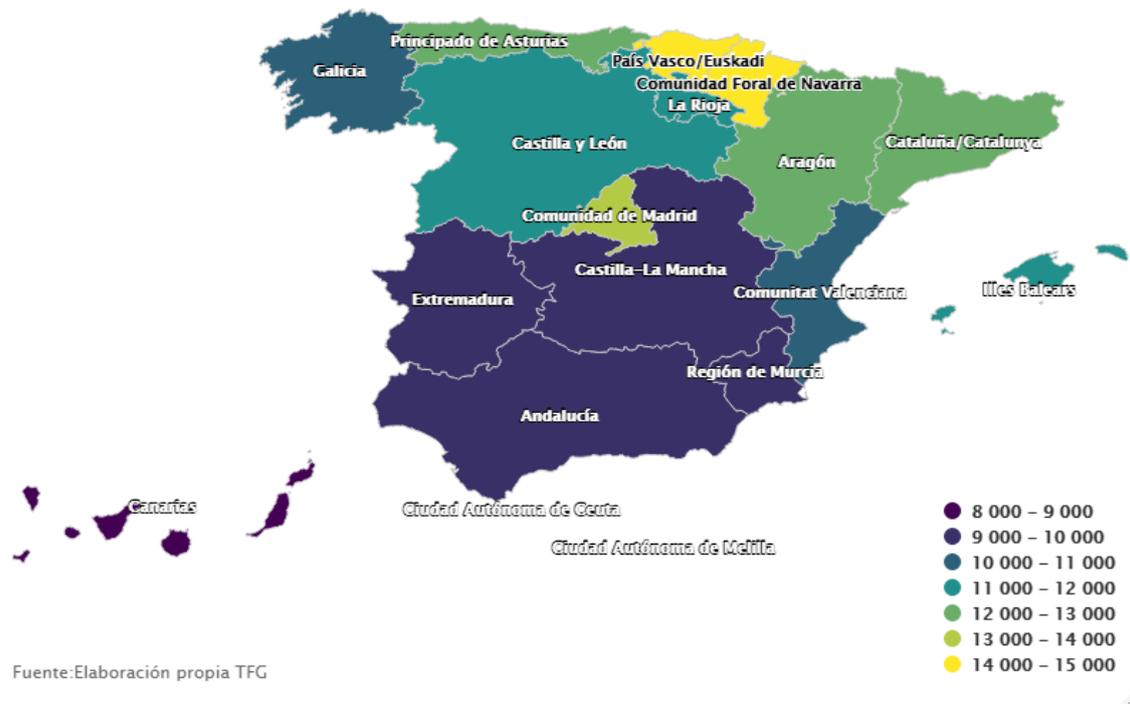
## Número medio de personas por hogar por Comunidad Autónoma



Hay que destacar que tanto Ceuta como Melilla, son las únicas comunidades con un número mayor de 3 personas por hogar, destacando Ceuta con unas 3,62 personas por hogar. El resto de las comunidades del territorio español tienen entre 2 y 3 personas por hogar.

Ahora sí, representamos la renta neta media por persona en cada comunidad y podemos detectar algunas curiosidades.

## Renta neta media por persona por Comunidad Autónoma



Lo primero que podemos observar rápidamente y de interés es que la mayor renta por persona se la adjudican las comunidades del País Vasco y Navarra, quedando la capital algo por debajo de estas dos debido a un mayor número de personas por hogar por norma general. Además, si generamos una línea horizontal a la altura de la capital podemos observar como las comunidades por debajo de dicha línea tienen una renta por persona inferior a las comunidades por encima.

Con respecto a Ceuta y Melilla, también hay bastante diferencia, ya que Ceuta es la comunidad con menor renta por persona (8,7 mil euros) mientras que Melilla está más próximo a la media con una renta de 11 mil euros por persona.

### 3.2.3 Tasa de Pobreza

Vamos a analizar la tasa de pobreza con respecto a los hogares, es decir el porcentaje de hogares por Comunidad Autónoma española que se encuentran en riesgo de pobreza, y además representaremos también la tasa de pobreza AROPE, es decir el

porcentaje de personas que o bien se encuentran en riesgo de pobreza, en situación de privación material severa o que vive en hogares con intensidad de trabajo escasa.

De esta manera podremos ver las diferencias existentes entre los hogares que se encuentran únicamente en riesgo de pobreza e individuos con situación desfavorecida por otra circunstancia o la acumulación de ellas.

En primer lugar, mostraremos la tasa de hogares en riesgo de pobreza:

Tasa de hogares en pobreza por Comunidad Autónoma



Fuente:Elaboración propia TFG

Donde podemos observar de manera similar al mapa con referencia a la renta neta media por persona y por hogar que las comunidades autónomas por debajo de la línea horizontal de Madrid presentan un mayor porcentaje de hogares en riesgo de pobreza que las comunidades del norte de España.

Existe una gran relación con respecto a dicho gráfico, ya que a mayor renta neta media por persona menor es el porcentaje de hogares en riesgo de pobreza, es decir estas dos variables la renta neta media por persona y la tasa de riesgo de pobreza están fuertemente relacionadas, teniendo una relación inversa, esto es que, al aumentar la

renta neta media por persona de una comunidad, su porcentaje de hogares en riesgo de pobreza disminuye y viceversa.

Ceuta y Extremadura son las comunidades autónoma con mayor porcentaje de hogares en riesgo de pobreza de España, con un 36 %, esto quiere que, de 100 hogares de dichas comunidades, 36 de ellos están en riesgo de pobreza.

Por el lado contrario, la comunidad Foral de Navarra presenta el menor porcentaje de hogares en riesgo de pobreza con tan solo un 8,12% seguido del País Vasco con un 12%, lo que nos indica que debemos fijarnos en el modelo de economía que siguen estas dos Comunidades para el bienestar y una buena condición de vida de nuestra población.

Otra interesante comparación es que Cataluña, considerada la segunda capital española presenta una tasa algo menor que la capital, siendo esta tasa de un 14% en Cataluña y un 15% en Madrid, a pesar de que Madrid presenta una renta neta media mayor por persona.

Tasa de pobreza AROPE por Comunidad Autónoma



Fuente:Elaboración propia TFG

Al ser la tasa AROPE, el cumplimiento de uno de los tres ítems mencionados anteriormente, los valores se concentran más. Con este gráfico podemos obviamente concluir que las Comunidades Autónomas del sur de España presentan una mayor pobreza que los del Norte, ya que las comunidades del sur presentan una tasa de pobreza AROPE de en torno al 32% de media y las comunidades del norte presentan una tasa de en torno al 18%. Es decir, el Norte tiene aproximadamente menos de la mitad de los individuos en riesgo de pobreza que el Sur, indicando la enorme desigualdad existente y la pobreza que hay en el sur con respecto al país.

**NOTA: ESTAMOS CONSIDERANDO EL NORTE Y EL SUR, DIVIDIENDO EL PAÍS EN LA LÍNEA HORIZONTAL GENERADA A LA ALTURA DE LA CAPITAL ESPAÑOLA.**

### 3.2.4 Índice Gini

A través del índice de Gini podemos detectar la desigualdad existente en nuestro país con respecto a los ingresos de nuestra población. Usaremos el índice de Gini, ya que desde mi punto de vista me resulta más cómodo trabajar con valores desde el 0 al 100 a la hora de sacar interpretaciones y conclusiones.

Coeficiente de Gini por Comunidad Autónoma



Como ya vimos en el gráfico de la serie temporal sobre el coeficiente de Gini en España desde el 2008 al 2020, este coeficiente es de un 33% en España y podemos observar que los valores están más centrados con respecto a la media que la tasa de pobreza.

Evidentemente, las interpretaciones que habíamos realizado anteriormente se confirman en este mapa, donde el Sur presenta mayor desigualdad que el Norte.

Hay que destacar la brutal desigualdad existente en Ceuta, con un índice de Gini de 41, lo que hace a esta comunidad sin ninguna duda como la más pobre de toda España y donde mayor desigualdad hay. España debe tomar nota de estos indicadores en esta comunidad y tratar de mejorar su situación.

### 3.2.5 Análisis Clúster con respecto a los indicadores de condiciones de vida y desigualdad

A continuación, realizaremos un análisis cluster (también conocido como análisis de conglomerados) de las Comunidades Autónomas en referencia a estos principales indicadores, de manera que nos permita detectar cuales son las Comunidades más homogéneas y además cuales son las Comunidades con mejores condiciones de vida en el año 2009.

El análisis de conglomerados lo constituye un conjunto de técnicas mediante las cuales se clasifican objetos o casos en grupos relativamente homogéneos llamados conglomerados o clusters.

El propósito de este apartado será el de organizar las comunidades autónomas en grupos homogéneos con respecto a los indicadores de dichas comunidades asociados a la encuesta de condiciones de vida y desigualdad en el año 2009.

En este tipo de estudios, se siguen los siguientes pasos:

1. Formular el problema. En esta etapa, el investigador identifica y selecciona las variables en las que basará la agrupación, ya que si se incluyen variables irrelevantes, pueden distorsionar el problema. En nuestro caso, las variables en las que se basarán las agrupaciones serán el valor de los siguientes indicadores: Renta neta media por persona, índice de gini y tasa de pobreza AROPE.

2. Seleccionar una medida de similitud. Se necesita una forma de medir la diferencia o semejanza entre observaciones u objetos, la forma en que generalmente se hace es en términos de la distancia entre cada par de casos; cuando la distancia es menor se considera que los casos son más parecidos entre sí. Existen diversas maneras de calcular la distancia, las que se aplican con mayor frecuencia son: la distancia euclidiana que es la raíz cuadrada de la suma de las diferencias al cuadrado entre los valores de dos casos para cada variable; la distancia euclidiana al cuadrado, para esto

sería el valor como se explicó anteriormente, antes de calcular la raíz cuadrada; la distancia de Manhattan o de calles urbanas entre dos casos es la suma de los valores absolutos de la diferencia entre observaciones para cada variable; la distancia de Chebychev entre dos objetos es el valor absoluto de la diferencia máxima entre los valores para cualquier variable. Cuando las variables se miden en unidades muy diferentes, antes de agrupar los casos, se recomienda estandarizar los datos para eliminar la influencia de la unidad de medición. En nuestro caso, al estar las variables de agrupación medidas en distintas unidades, estandarizaremos los datos y posteriormente calcularemos las distancias mediante la distancia euclídea.

3. Seleccionar un procedimiento de agrupamiento. Para este paso hay diversos métodos. Los procedimientos de conglomerados se dividen en jerárquicos y no jerárquicos de acuerdo al procedimiento de agrupación. Los conglomerados jerárquicos se caracterizan porque dicha jerarquía se hace en forma de árbol.

Los procedimientos jerárquicos pueden ser por aglomeración o por división. Siguiendo el método de conglomerados por aglomeración, la formación de los conglomerados, se puede hacer de diferentes maneras; en la investigación de mercados se aplican tres métodos que son: el de enlace, el que utiliza la varianza y el de centroides.

Los procedimientos no jerárquicos se conocen como agrupación de  $k$  medias (  $k$  means clustering).

4. Decidir el número de conglomerados a conservar.

5. Interpretar y elaborar un perfil de los conglomerados. En esta etapa se procederá a determinar las características de cada conglomerado que se conservará.

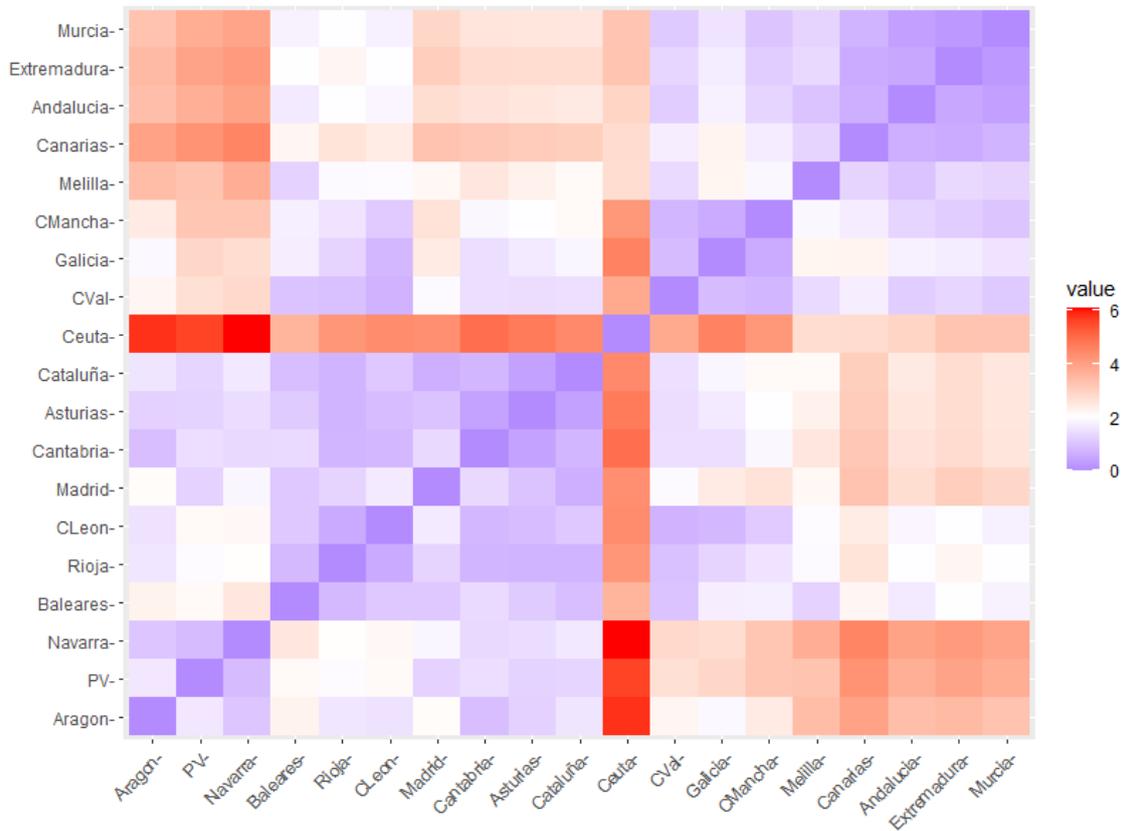
En nuestro caso, aplicaremos el método de clasificación no jerárquico, conocido como  $k$  means clustering y para medir la distancia se calculará la distancia euclídea al cuadrado al cuadrada. Además el número de grupos a obtener se ha calculado con la función "NbClust" de la librería "NbClust", que nos indica el número óptimo de cluster obtenido mediante varios índices formados por la combinación de distintos métodos de clustering y distancias.

Se formarán 4 grupos de comunidades autónomas con medidas de desigualdad y pobreza similares, ya que las variables de agrupación son la renta neta media por persona, la tasa de pobreza AROPE y el índice de Gini, las cuales serán estandarizadas para eliminar la influencia de las distintas unidades de medición de estos indicadores.

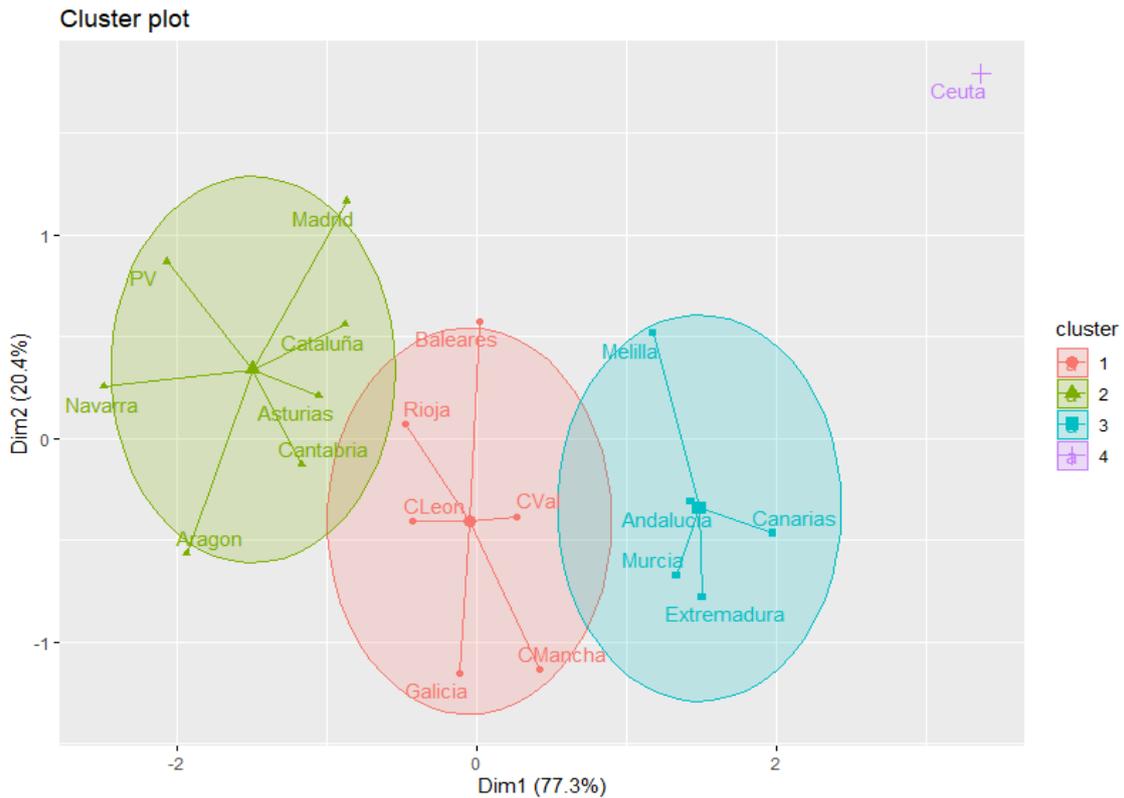
Los gráficos aportados en esta sección se realizan de manera automática mediante la librería "factoextra", la cuál extrae y visualiza los resultados de análisis de datos multivariantes.

En el siguiente gráfico podremos observar las distancias existentes entre las Comunidades Autónomas españolas en función de los valores de dichos indicadores, y Podemos observar que existen diferencias, es decir comunidades cuyos valores en los indicadores presentan bastante disimilitud y en cambio entre otras comunidades

existe similitudes, lo que nos permite observar que merece la pena realizar un análisis de cluster. Podemos destacar el caso de Ceuta como ya hemos comentado anteriormente que presenta unos valores extremos de los indicadores en comparación con otras comunidades.



A continuación, mostaremos el gráfico que nos muestra la formación de los 4 clusters, con las Comunidades Autonomas que forman cada uno de ellos, además del centroide de cada cluster y la distancia a cada comunidad que lo forma.



La formación de cada cluster es la siguiente:

1. Islas Baleares, La Rioja, Comunidad Valenciana, Castilla y León, Galicia y Castilla y La Mancha.
2. País Vasco, Navarra, Madrid, Cataluña, Aragón, Cantabria y Asturias.
3. Melilla, Andalucía, Canarias, Murcia y Extremadura.
4. Ceuta.

Además cada cluster se caracteriza por:

1. Una renta neta media por persona aproximadamente de unos 11 mil euros, una tasa de pobreza AROPE del 24% y un índice de Gini de 0,31.
2. Una renta neta media por persona aproximadamente de unos 13 mil euros, una tasa de pobreza AROPE del 16% y un índice de Gini de 0,3.
3. Una renta neta media por persona aproximadamente de unos 9,5 mil euros, una tasa de pobreza AROPE del 35% y un índice de Gini de 0,33.
4. Una renta neta media por persona aproximadamente de unos 8,5 mil euros, una tasa de pobreza AROPE del 37% y un índice de Gini de 0,42.

De esta manera podemos concretar que las comunidades del sur (clusters 3 y 4) tienen una renta neta media por persona mucho por debajo que las comunidades del norte, además de una tasa de pobreza mucho mayor, destacando el caso más extremo, que es el de la comunidad autónoma de Ceuta, que es con diferencia la más pobre y donde más desigualdad existe. Excluyendo Ceuta, el resto de España presenta una

desigualdad muy homogénea (valores de coeficiente de gini bastante similares, es decir existe poca varianza con respecto a este indicador).

En las comunidades del norte, donde mejores condiciones de vida se alcanzan, destacan algunas comunidades por una condición de vida aún superior como es el caso de las comunidades que forman el cluster 2 debido a la mayor renta neta media por persona y al menor porcentaje de población que están en riesgo de pobreza y/o exclusión social.

Si hubiéramos realizado un cluster de dos grupos, se dividiría España en dos partes, marcadas como venimos viendo en los anteriores mapas por una línea horizontal a la altura de Madrid, donde el Norte destacaría por unas mejores condiciones de vida que el Sur con respecto a estos 3 indicadores con una gran diferencia.

### 3.2.6 Evaluación situación 2009

Mediante el análisis realizado sobre la situación española en el año 2009 en referencia a los indicadores de condiciones de vida hemos visto la enorme cantidad de personas que están en riesgo de pobreza y/o exclusión social, así como la desigualdad existente. Si comparamos nuestros datos con las potencias europeas, estamos muy por debajo de ellas con una renta mucho inferior y una pobreza elevada, por lo que sería conveniente realizar un cambio que mejore la situación de nuestra población.

Hacer hincapié en la situación de vida que presenta la población española que habita en las comunidades autónomas del sur de España, las cuales presentan con diferencia las peores condiciones de vida, por lo que habría que tratar de convertir su situación en algo más similar al comportamiento de las del Norte. Más preocupante es aún la situación de Ceuta.

## 3.3 Principales Indicadores en las Comunidades Españolas en el 2020

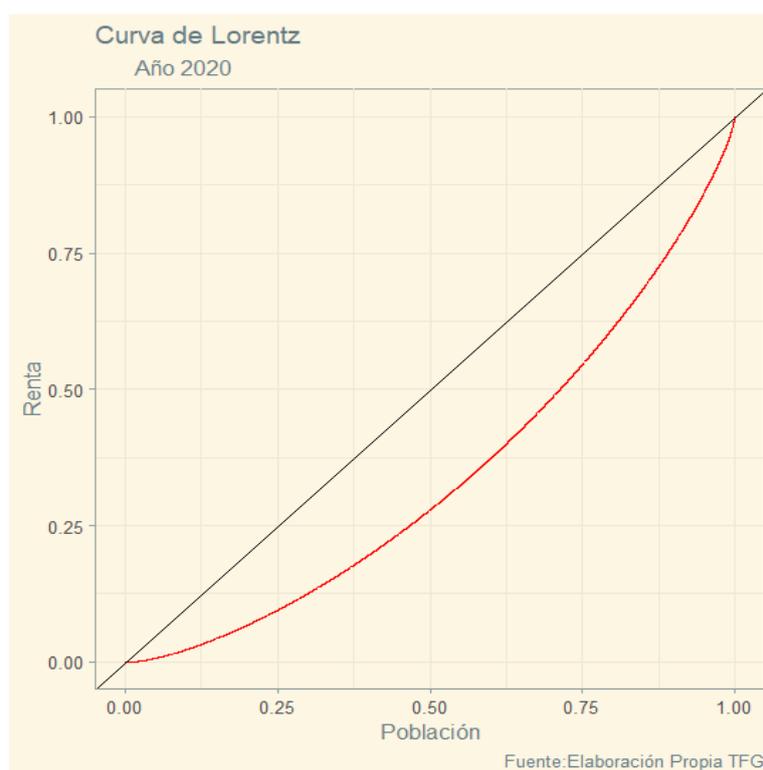
En este apartado, de manera similar al apartado anterior, el objetivo será exponer los valores de los principales indicadores según las comunidades autónomas de España con la misión de evaluar e identificar cuáles de ellas tienen mejores condiciones con respecto a las otras.

La estructura y manera de trabajar con los datos ha sido similar a la del apartado anterior ([Principales Indicadores en las Comunidades Españolas en el 2009.](#)), representando el mapa del territorio español con sus comunidades en función de los

valores de cada indicador para poder interpretar y detectar las comunidades que presentan en el año 2020 mejores condiciones de vida, y que por lo tanto representan a nuestra población en la actualidad. Como ya comentamos anteriormente, el año 2020 es el último año del que tenemos información sobre la encuesta de condiciones de vida en el INE.

### 3.3.1 Curva de Lorentz

En primer lugar, vamos a representar la curva de Lorentz con respecto a los ingresos de la población española en el año 2020.



A partir de la gráfica de Lorentz, la cual nos representa en el eje X el porcentaje acumulado de la población y en el eje Y el porcentaje acumulado de los ingresos obtenidos por dicha población de entre toda la generada, podemos observar claramente, de manera similar a la del año 2009 como la situación del país no es ni mucho menos cercana a la de la igualdad, ya que existe cierto abombamiento de la curva de Lorentz que nos representa la existencia de desigualdad en la distribución de la renta en la población española.

Algunos datos curiosos que podemos extraer de la curva de Lorentz con respecto al año 2020 son los siguientes:

- En torno al 25% de la población más rica de España obtienen el 50% de la renta generada en el año 2020.

- El 25% de la población más pobre de España obtienen únicamente aproximadamente el 10% de la renta generada en el año 2020.
- Si dividimos el país en dos grupos de idéntico tamaño, siendo el grupo “pobre”, el 50% de la población con ingresos más bajos, y siendo el grupo “rico”, el 50% de la población con ingresos más elevados, entonces podemos observar una diferencia bestial, donde el grupo pobre únicamente obtiene una cuarta parte de la renta generada y el grupo rico obtiene tres cuartas parte de la renta generada.

Si comparamos estos datos obtenidos con los del año 2009 son exactamente los mismos, por lo que nos hace pensar que la desigualdad con respecto a la obtención de ingresos no ha variado a lo largo de dichos años. Posteriormente analizaremos ambas curvas en un mismo gráfico para ver si la desigualdad ha aumentado o disminuido algo en este intervalo de tiempo.

### 3.3.2 Renta neta media

Representaremos la renta neta media por hogar de cada comunidad autónoma y la renta neta media por individuo, además del número de individuos medio por hogar en cada comunidad, ya que nos sirve como ayuda a la hora de visualizar la renta por hogar y por individuo debido a que en función la renta neta media por hogar y del número de individuos por hogar se calcula la renta neta media por persona.

En primer lugar, representaremos el mapa de comunidades en función de la renta neta media por hogar, donde Extremadura se mantiene como la comunidad española con menor renta neta media por hogar en España al igual que en 2009, pero habiendo bajado esta renta considerablemente ya que es de 22250 euros anuales, es decir cerca de 2000 euros por debajo menos que en 2009. Las comunidades donde existe una renta neta media por hogar mayor continúan siendo las mismas que en el año 2009, y se suma a ellas Melilla.

## Renta neta media por hogar por Comunidad Autónoma



Al igual que en el año 2009, la segunda comunidad con menor renta neta por hogar es Canarias, aunque es prácticamente similar a la comunidad de Andalucía con respecto este indicador

La mayoría de las comunidades generan de manera neta una renta de entre 28 y 32 mil euros anuales por hogar, y además continúa existiendo una tendencia de mayor renta neta por hogar en las comunidades del norte de España.

A continuación, vamos a mostrar el número de personas que residen por hogar en cada comunidad autónoma, dónde destaca la comunidad autónoma de Melilla la cuál es la única que presenta un número mayor de 3 personas por hogar.

Además, podemos observar como la tendencia del número de personas por hogar es a la baja por norma general en España ya que si lo comparamos con el año 2009 el mapa es más apagado en el 2020 (están en las mismas escalas y cortes la leyenda), y también podemos observar como el número de personas por hogar tiende a ser algo mayor en el sur de España.

## Número medio de personas por hogar por Comunidad Autónoma

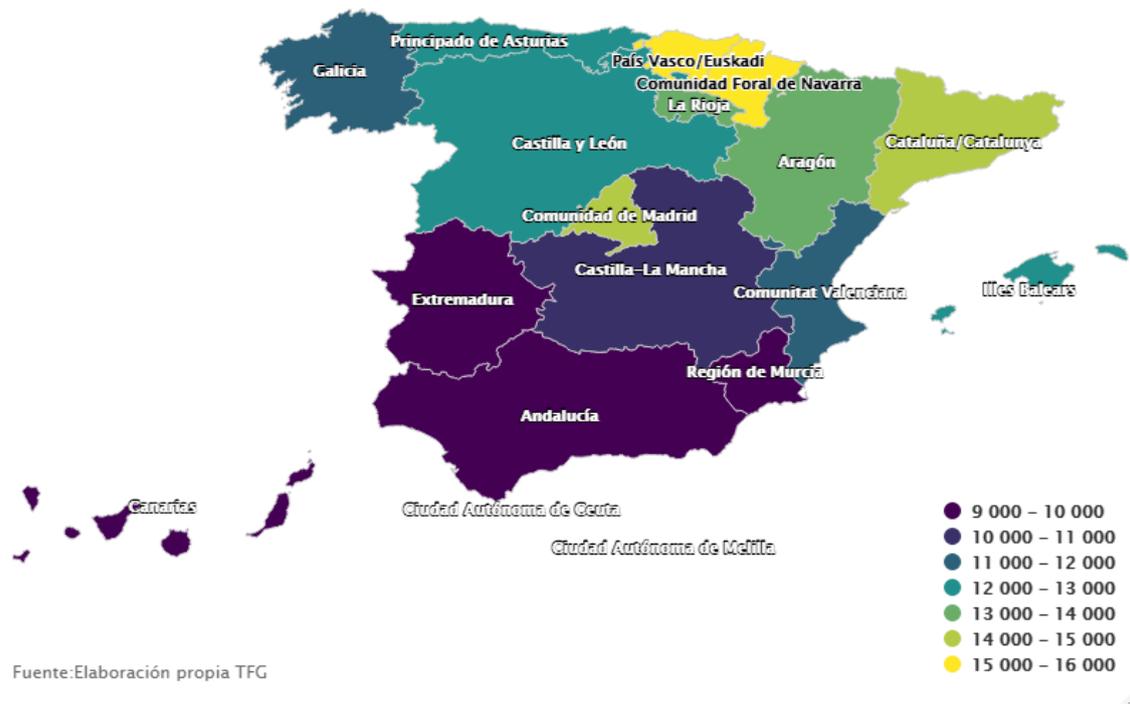


Haciendo referencia a la renta neta media por persona, observaremos algo que nos sorprenderá bastante con respecto a 2009 y esto es debido a los dos gráficos anteriores que están fuertemente relacionados con este.

Al haber aumentado la renta neta media por hogar, o mejor dicho, ser esa la presente tendencia y la disminución del número de personas por hogar, podemos predecir que la renta neta media por persona debería incrementar.

Efectivamente, la renta neta media por persona ha aumentado, pero cuidado al comparar ambos mapas (2020 y 2009) debido a que las escalas han cambiado ya que la menor renta neta por persona es mayor que la menor de 2009 y al igual pasa con la mayor. Por lo tanto, efectivamente la renta neta por persona ha aumentado, pero podemos prever que la población española está reduciendo el hecho de construir familias, es decir tener hijos, y también mayor número de personas deciden vivir de manera solitaria.

## Renta neta media por persona por Comunidad Autónoma



Las comunidades con mayor renta neta media por persona continúan siendo el País Vasco y Navarra, al igual que las comunidades del norte presenta en este indicador valores mayores que en el sur, lo que parece ser evidente que este problema de desigualdad sigue estando presente.

En cambio, Ceuta ha dejado de ser la comunidad con una renta neta por persona menor, a pesar de seguir siendo de las más pobres en este sentido con una renta de casi 10 mil euros anuales por persona. Este puesto es ocupado actualmente por la comunidad de Extremadura con una renta de 9100 euros anuales por persona.

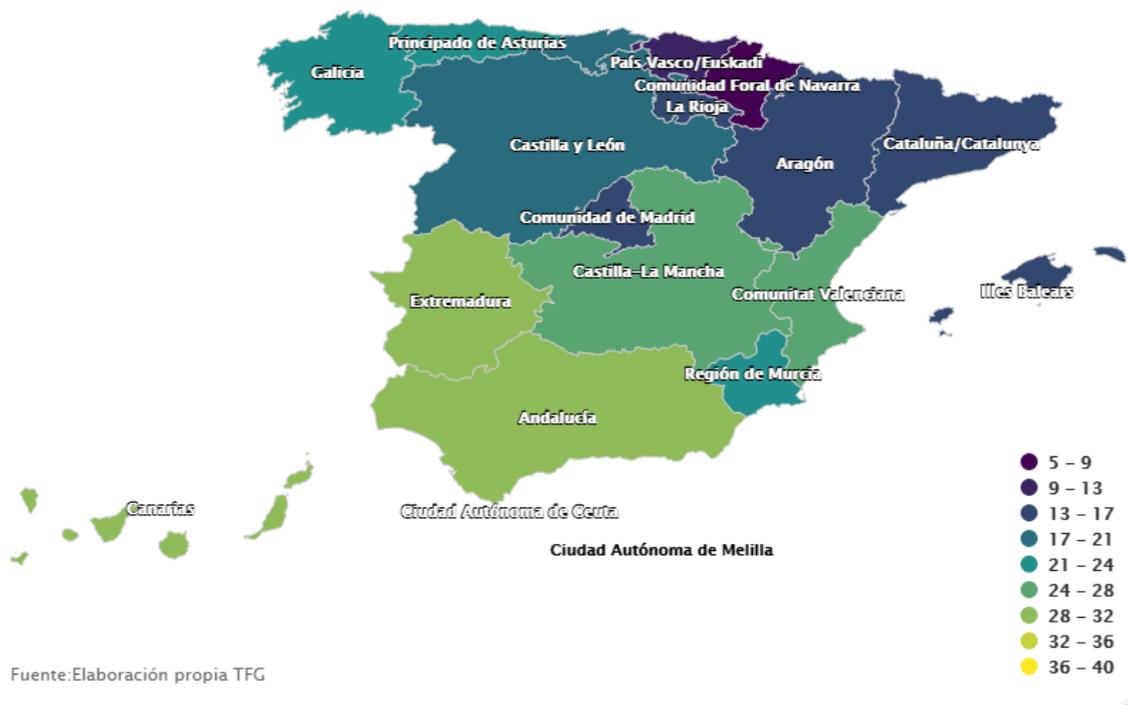
### 3.3.3 Tasa de pobreza

En este punto, vamos a analizar y representar la tasa de pobreza con respecto a los hogares y la tasa de pobreza AROPE, para conocer tanto el porcentaje de hogares que se encuentran en riesgo de pobreza por comunidad como el porcentaje de personas por comunidad que cumplen algunas de los requisitos para formar parte de la pobreza

según el modelo AROPE, es decir persona en riesgo de pobreza, con situación de privación material severa y/o vive en hogar con intensidad de trabajo escasa.

En primer lugar, mostraremos la tasa de hogares en riesgo de pobreza, usando las mismas escalas y cortes que en el gráfico del año 2009.

Tasa de hogares en pobreza por Comunidad Autónoma



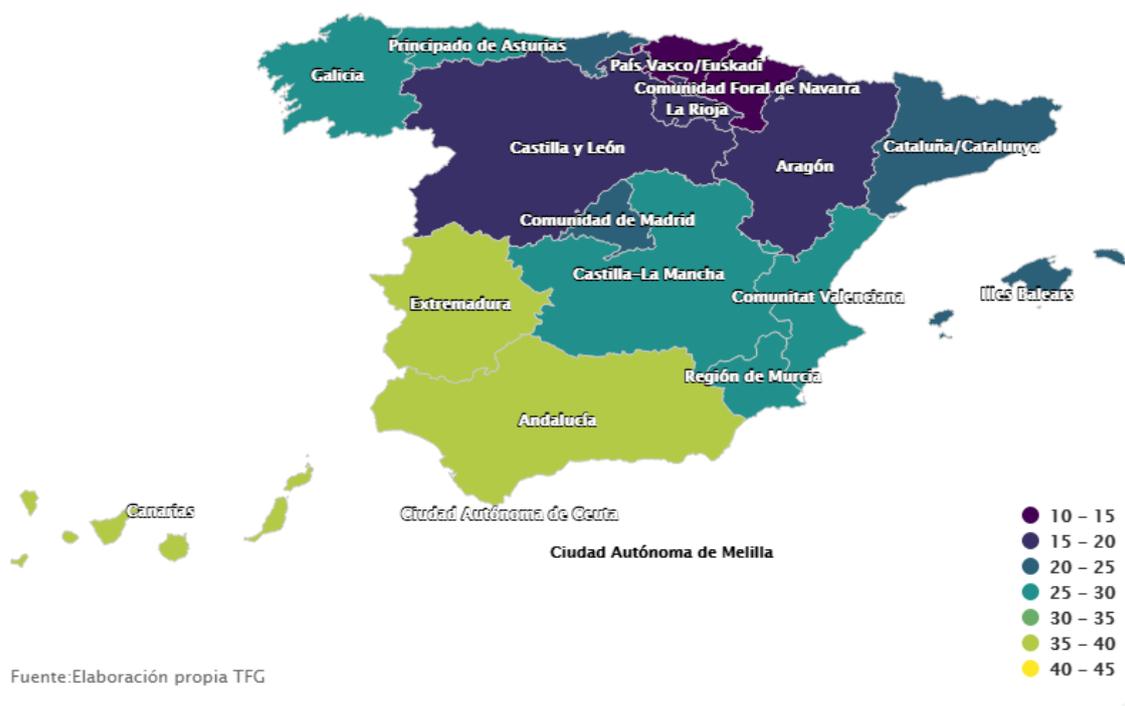
En primer lugar y más destacable se puede observar que la tasa de hogares en pobreza mayor es inferior a la del 2009 (debido a que no hay ninguna comunidad en color amarillo). En el año 2009, la comunidad con mayor tasa de riesgo de pobreza por hogar eran Ceuta y Extremadura con un 38% y en el año 2020 es Melilla con un 35% de hogares en pobreza.

Además, en el norte podemos observar cómo los colores son menos vivos que en el gráfico del año 2009 por lo que la tasa de pobreza de los hogares parece tener una tendencia disminuyente, algo que es bastante bueno. Seguimos pudiendo ver como las comunidades del norte presenta una tasa de pobreza mucho menor que el sur y como este indicador está relacionado con la renta neta, ya que a mayor renta neta por hogar/persona menor tasa de pobreza se presenta. Como la tendencia de la renta está en auge con respecto al año 2009 entonces la tendencia de la tasa de pobreza es disminuyente.

Las comunidades con menor tasa de hogares en pobreza son las 2 que mayor renta neta tienen (Navarra, PV) destacando Navarra con un 8% de hogares en pobreza.

Si analizamos la tasa de pobreza AROPE, hay algo chocante a primera vista que no hace referencia a lo que llevamos visto hasta ahora en el año 2020 y es la presencia de un valor en este indicador muy elevado. La comunidad de Melilla presenta una tasa de pobreza AROPE de un 42,5%, valor que no había sido alcanzado en el año 2009, lo que parece significar que existe un abundante número de personas con privación material severa y/o que vive en hogar con intensidad de trabajo muy baja, factor que se puede deber al incremento en los precios de algunos bienes necesarios y la pandemia que comenzó en marzo de 2020 debido al Covid-19.

Tasa de pobreza AROPE por Comunidad Autónoma



Al igual que en el año 2009, la comentada situación que viven en el sur de España continúa siendo notable.

Las comunidades de Canarias, Andalucía, Ceuta, Melilla y Extremadura continúan siendo las más pobres y desfavorecidas del territorio español con una diferencia abismal.

De esta manera, una breve conclusión que desde mi punto de vista es conveniente indicar es que el gobierno español debe empezar a tomar manos en el asunto para eliminar estos valores tan abismales de pobreza en el sur, implantando leyes o cambios tomando como ejemplo las comunidades del norte como Navarra y el PV, consiguiendo así que el territorio español además de presentar menos desigualdad sea menos pobre y las condiciones de vida mejoren.

### 3.3.4 Índice Gini

Vamos a analizar la desigualdad existente en nuestro país en el año 2020 que ya hemos visto que es existente a través de la curva de Lorenz en gran medida. Para ello volveremos a usar el índice de Gini.

Coefficiente de Gini por Comunidad Autónoma



Algo obvio a partir de los resultados obtenidos por los demás indicadores es que la comunidad autónoma con mayor desigualdad es la de Melilla con un 44%. El resto de las comunidades se concentran más o menos en torno al 30% con respecto a este indicador y por lo tanto el del país en general sea algo mayor incentivado por Ceuta (40%) y Melilla.

Destacable la gran desigualdad existente en comunidades tan influyentes e importantes como Madrid y Cataluña.

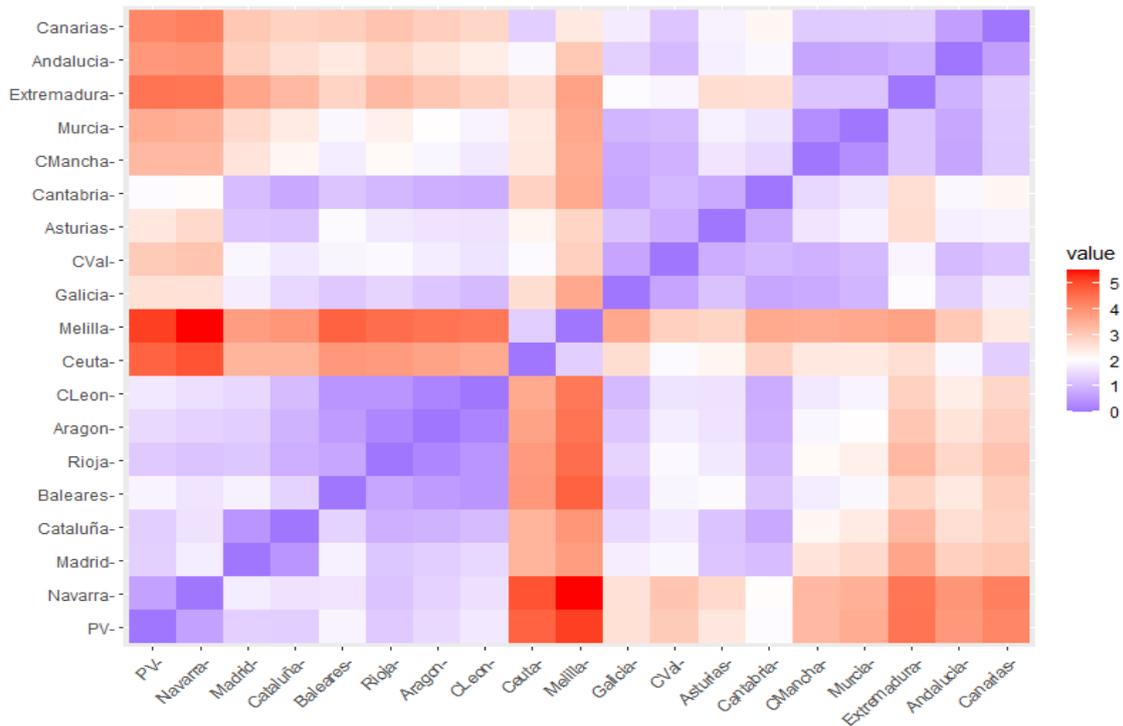
Lo que en este gráfico no parece tan evidente es la situación de desigualdad mayor en el sur (sin tener en cuenta Ceuta y Melilla) que, en el Norte, algo sorprendente con respecto a lo visto en mapas anteriores, siendo en 2009 esta desigualdad notable en el sur mayor que en el norte.

### 3.3.5 Análisis Clúster con respecto a los indicadores de condiciones de vida y desigualdad.

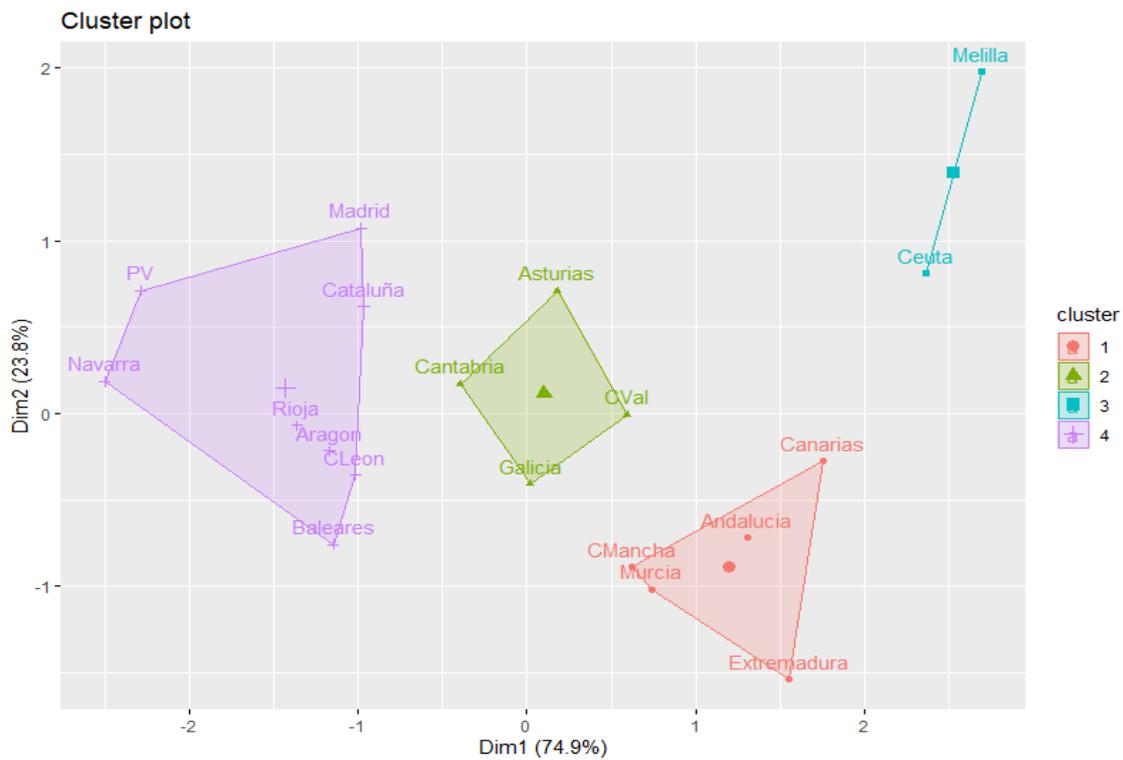
El propósito de este apartado será el de organizar las comunidades autónomas en grupos homogéneos con respecto a los indicadores de dichas comunidades asociados a la encuesta de condiciones de vida y desigualdad en el año 2020. Para ello se hará uso del análisis clúster que ya explicamos anteriormente en el mismo caso, pero para la ECV del 2009. ([Análisis Clúster con respecto a los indicadores de condiciones de vida y desigualdad](#))

En este caso, se aplicará el mismo método: No jerárquico conocido como k means clustering y para medir la distancia se calculará la distancia euclídea al cuadrado. El número de grupos a obtener se ha calculado con la función “NbClust”, donde se ha obtenido al igual que en el realizado para el año 2009, 4 grupos de comunidades autónomas con medidas de desigualdad y pobreza similares, donde las variables de agrupación son la renta neta media por persona, tasa de pobreza AROPE y índice de Gini. Estas variables serán estandarizadas para eliminar la influencia de las distintas unidades de medición de estos indicadores.

Para observar las distancias existentes entre las comunidades autónomas españolas en función de los indicadores usados está el siguiente gráfico, donde a mayor intensidad del color rojo mayor diferencia existe entre las comunidades en función de las condiciones de vida y a mayor intensidad del color azul mayor similitud existe. Este gráfico es solamente orientativo y nos permite saber que es apropiado realizar un clúster ya que existen comunidades homogéneas y distintas con respecto a los indicadores de desigualdad y pobreza.



En el siguiente gráfico, mostraremos los clústeres que se han formado una vez realizado el análisis de conglomerados con los indicadores ya estandarizados. Se han formado efectivamente 4 clústeres y en cada uno de ellos encontraremos las comunidades autónomas que lo forman, además del centroide y la distancia de que cada comunidad al centroide del grupo al que pertenece.



Se han formado los siguientes 4 clusters:

1. Castilla la Mancha, Murcia, Andalucía, Canarias y Extremadura.
2. Asturias, Cantabria, Galicia y Valencia.
3. Ceuta y Melilla
4. PV, Madrid, Cataluña, Navarra, Rioja, Aragón, Baleares y Castilla y León.

Cada cluster se caracteriza por los siguientes aspectos:

1. Una renta neta media por persona aproximadamente de unos 10 mil euros, una tasa de pobreza AROPE del 34% y un índice de Gini de 0,30.
2. Una renta neta media por persona aproximadamente de unos 12 mil euros, una tasa de pobreza AROPE del 26% y un índice de Gini de 0,32.
3. Una renta neta media por persona aproximadamente de unos 10,5 mil euros, una tasa de pobreza AROPE del 38% y un índice de Gini de 0,42.
4. Una renta neta media por persona aproximadamente de unos 14 mil euros, una tasa de pobreza AROPE del 17% y un índice de Gini de 0,27.

De esta manera podemos concretar que las comunidades del sur tienen unas condiciones de vida mucho por debajo que las comunidades del norte, además de estar en riesgo de pobreza y/o exclusión social con mucha más frecuencia, destacando el caso extremo de las comunidades de Ceuta y Melilla que son con diferencia las comunidades más pobres y donde más desigualdad existe de toda España.

En las comunidades del norte (clúster 4) es donde mejores condiciones de vida se alcanzan, excepto en las comunidades que forman el clúster número 2, las cuales viven una situación algo peor, pero a la vez algo mejor que el sur, es decir tienen una situación más neutra, sin tener las mejores condiciones ni las peores.

Si hubiéramos realizado un clúster de dos grupos se dividiría el país en dos regiones: sur y norte a excepto de las comunidades del clúster 2 que pasarían a ser más similares al clúster de comunidades del sur, por lo que es evidente las diferentes condiciones de vida en España cortada por la línea horizontal en Madrid.

### 3.3.6 Evaluación situación 2020

Al igual que en el año 2009 la situación de España continúa siendo deficiente con respecto a las de las potencias europeas, debido al enorme número de personas en riesgo de pobreza y/o exclusión social y la enorme desigualdad existente.

La brecha respecto a algunas comunidades sigue siendo abismal, con lugares en España con unas condiciones de vida muy malas, por lo que deberíamos fijarnos en las comunidades que si tienen condiciones mejores y tratar de realizar cambios en las comunidades más afectadas y mejorar su situación.

Aunque debido a la situación actual del Covid-19 y la guerra existente en Europa, no parece ser que estos indicadores vayan a ir por buen camino, sino todo lo contrario.

### 3.4 Evolución principales indicadores desigualdad y condiciones de vida 2009-2020

En este apartado veremos con gráficos apropiados la evolución de los principales indicadores de desigualdad y condiciones de vida que han tenido en el intervalo de tiempo desde el año 2009 al año 2020, de manera que sea más fácil detectar la evolución presente.

Se realizará por una parte la evolución de los indicadores principales con respecto al territorio español y posteriormente se analizará la evolución de cada indicador por comunidades autónomas, para poder detectar correctamente e interpretar algunas de los comentarios sobre el comportamiento que hemos podido observar en los anteriores apartados.

Para la realización de estos gráficos, diseñados especialmente para analizar evolución de ciertos aspectos, he usado el gráfico de pendientes en “ggplot2” y la librería adicional “CGPfunctions”, que nos proporciona la función “newggslopegraph” que nos generará dicho gráfico de pendientes.

En primer lugar, veremos la evolución en el territorio español de los principales indicadores sin la desagregación por comunidades autónomas

## Evolución de los Indicadores de Pobreza y desigualdad en España

2009-2020. (Umbral y Renta Neta en miles de euros)



Analizando uno a uno estos indicadores, podemos observar que han aumentado tres de ellos:

- Renta neta media por persona
- Tasa de pobreza AROPE
- Umbral de pobreza)

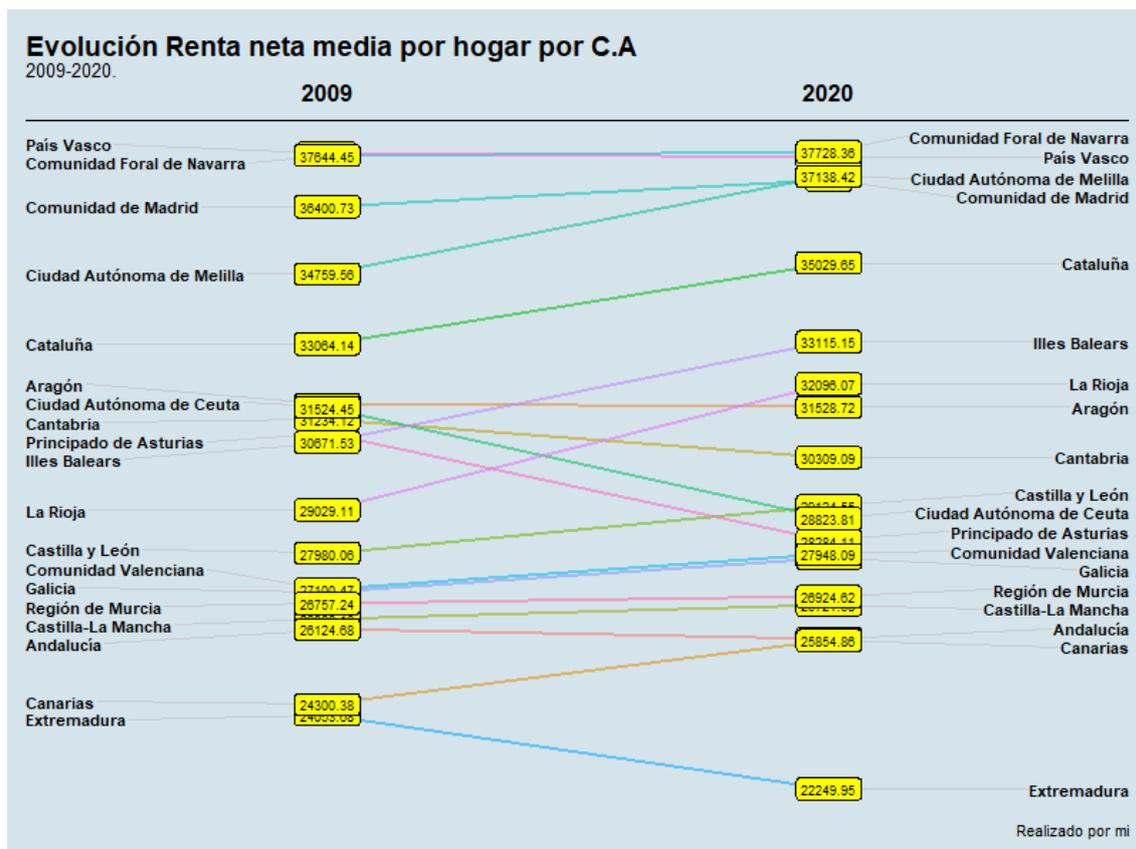
Los otros dos indicadores restantes, han disminuido aunque prácticamente podemos considerar que se han mantenido constantes:

- Índice de Gini
- Ratio s80s20

Es evidente que la población Española genera una renta neta media mayor que en el 2009 con un aumento de casi mil euros de media y a la vez el umbral de pobreza ha aumentado cerca de 800 euros de media, por lo que en el año 2020 para estar en riesgo de pobreza y/o exclusion social necesitas tener una renta neta media mayor que en el año 2009. Si se calculan los porcentajes de incremento, el umbral de pobreza ha crecido más que la renta neta media por persona, lo que justifica el crecimiento de la población Española que esta en riesgo de pobreza mediante la tasa de pobreza AROPE, la cuál ha crecido en un 1,7%. El ratio s80s20 prácticamente se ha mantenido constante y el índice de gini ha disminuido muy poco.

En conclusión, el territorio Español esta en una situación peor que en el 2009 y esto se debe a los incrementos de los bienes principales y necesarios para poder tener una vida decente, sin el crecimiento necesario en los salarios, añadido al aumento de impuestos. Existe una enorme desigualdad entre los más ricos y más pobres en este país, donde el 20% de las personas con más ingresos generan casi 6 veces más que el 20% de las personas con ingresos más bajos.

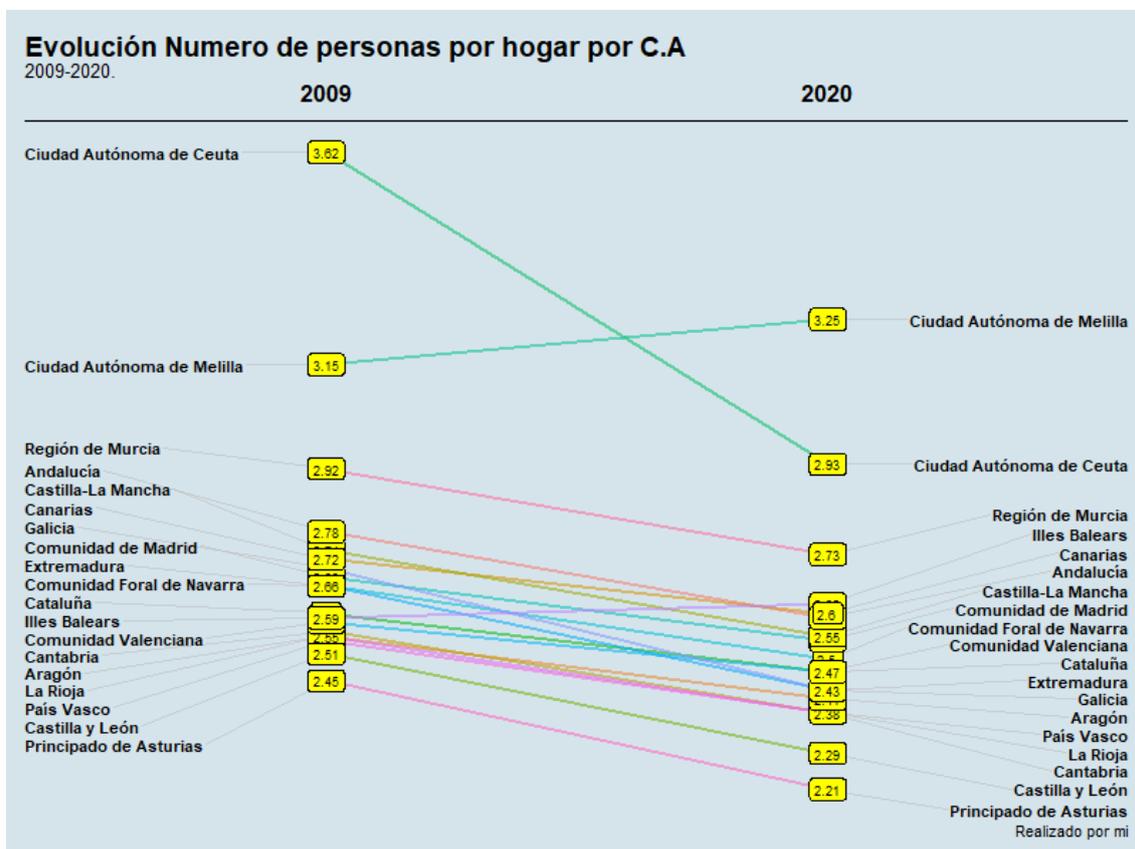
A continuación, comentaremos la evolución de las comunidades autónomas con respecto a algunos de los principales indicadores de desigualdad y pobreza.



Observando la evolución de la renta neta media por hogar en las diferentes comunidades, se observa que las comunidades donde los hogares tenían una mayor renta siguen siendo los mismos en 2020 al igual que los que tenían menor renta en 2009 siguen siendo los mismos en 2020.

Muy destacable el caso de la comunidad de Extremadura que siendo la comunidad con menor renta neta por hogar en 2009, en el año 2020 aún es mucho menor, siendo la que menor renta neta por hogar tiene con una diferencia abismal. Otras comunidades que han presentado un gran descenso con respecto a este indicador son el Principado de Asturias y Ceuta. Otras comunidades como Canarias, La Rioja, las Islas Baleares, Melilla y Cataluña han presentado un gran aumento en este indicador.

Con respecto al número de personas por hogar la tendencia ha sido la misma para prácticamente todas las comunidades españolas, las cuales han reducido su número medio en gran cantidad, a excepción de Melilla que se ha convertido en la comunidad con más número de personas por hogar. La comunidad con menor promedio es la de Asturias. Lo podemos ver en el siguiente gráfico:

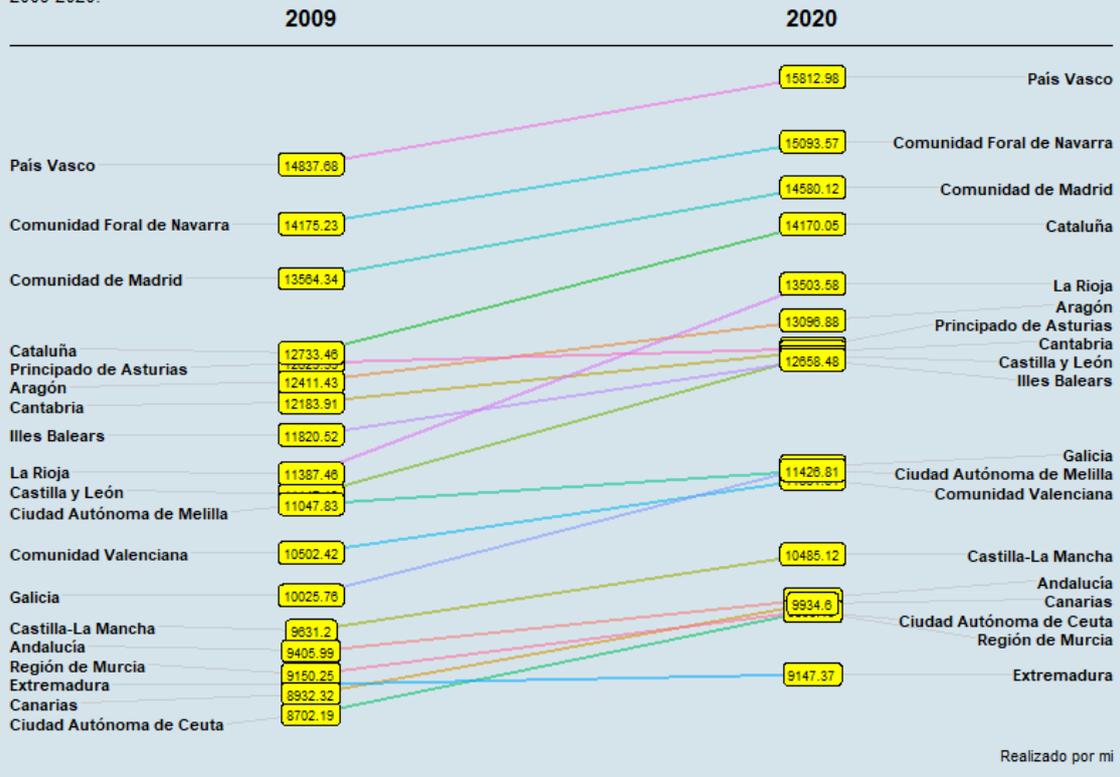


Una vez observado estos dos gráficos vamos a ver la evolución de la renta neta media por persona, que debido a la gran tendencia disminuyente del número de personas por hogar se observará el crecimiento de la renta neta por persona.

Efectivamente, como podemos ver en el gráfico a continuación, la renta por persona ha crecido en todas las comunidades excepto en Asturias y Extremadura que se ha mantenido constante.

La comunidad con mayor renta por persona continua siendo el PV con diferencia y la comunidad de Extremadura se ha convertido en la que tiene menor renta ya que Ceuta y Canarias han aumentado considerablemente su renta neta por persona. Galicia, Cataluña y La Rioja son las comunidades con un mayor aumento de este indicador y se puede concretar, que tanto en el año 2009 como en el año 2020, las comunidades con menor renta por persona son las del sur de España.

## Evolución renta neta media por persona por C.A. 2009-2020.

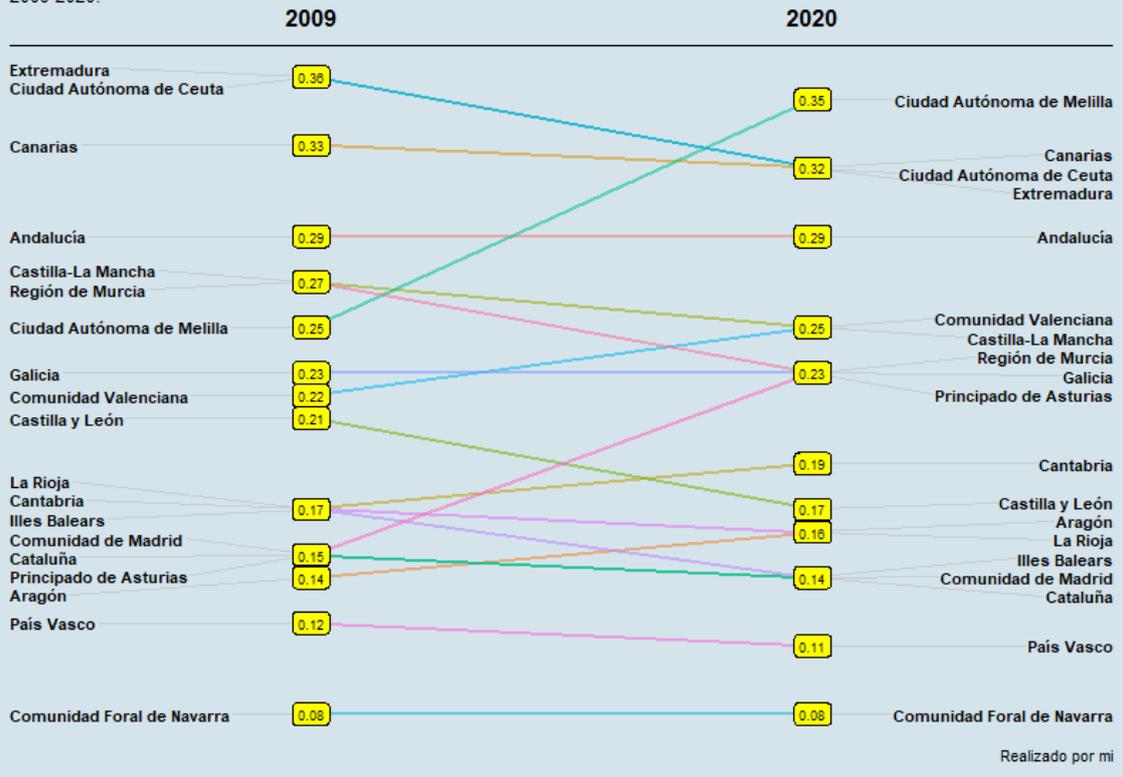


En los siguientes dos gráficos analizaremos tanto la evolución de la tasa de hogares en pobreza como la tasa de pobreza AROPE, las cuales son muy similares pero la tasa AROPE siempre será algo mayor debido a que basta con cumplir uno de tres requisitos, siendo uno de ellos el requisito para la tasa de pobreza básica.

En la tasa de pobreza básica, podemos destacar el impactante aumento de la tasa de pobreza en Melilla con un 10% de aumento y convirtiéndose en la comunidad con mayor tasa de hogares en pobreza. También es destacable el gran aumento de Asturias que ha pasado de un 15% en el 2009 a un 23% en 2020.

En el resto de comunidades la tendencia ha sido de mantenerse más o menos constante o disminuir levemente, lo que parece hasta el momento una buena señal.

## Evolución Tasa de hogares en Pobreza por C.A 2009-2020.



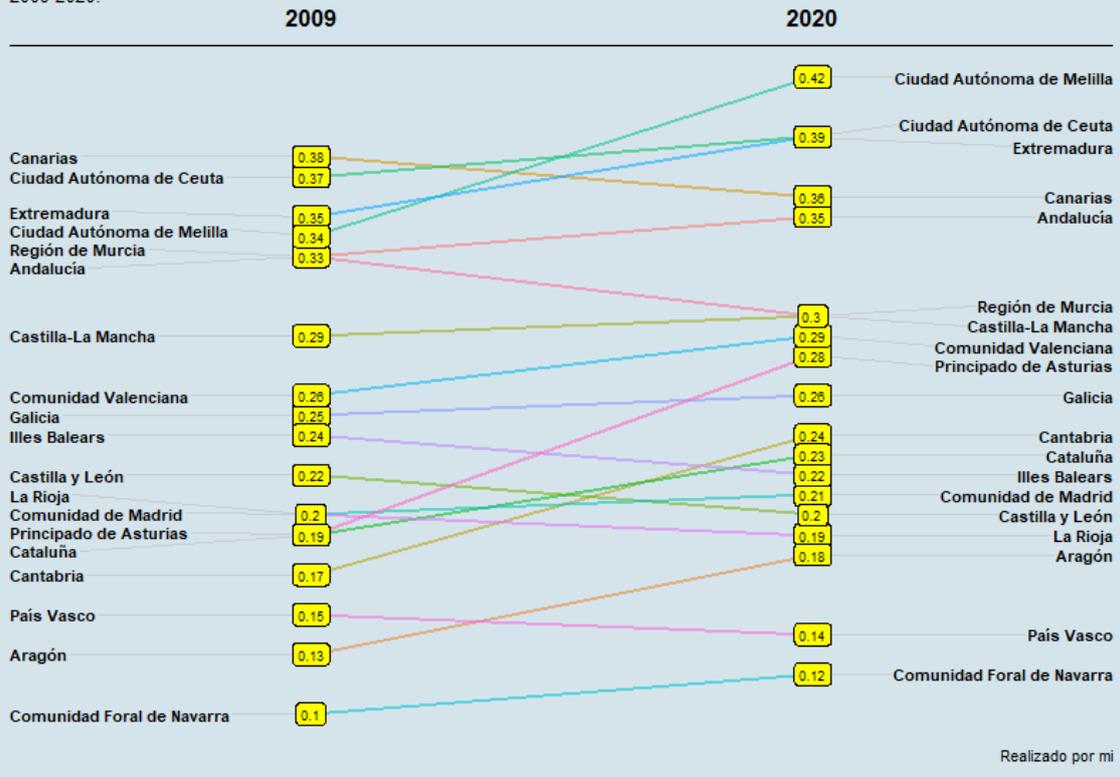
Con respecto a la tasa de pobreza AROPE, la mayoría de las comunidades han presentado un aumento en este indicador bastante grande que refleja la grave situación que se vive en España y que la población Española presenta malas condiciones de vida. Destacar al igual que en el anterior gráfico el aumento en la comunidad de Melilla con un dato nefasto, ya que el 42% población esta en riesgo de pobreza.

Tan solo 4 comunidades han presentado un descenso en este indicador, destacando el caso de Murcia. Navarra sigue siendo la comunidad con menor número de personas en riesgo de pobreza.

También se puede confirmar la pesimá situación del sur de España, las cuales presentan mayor tasa de pobreza que en el norte.

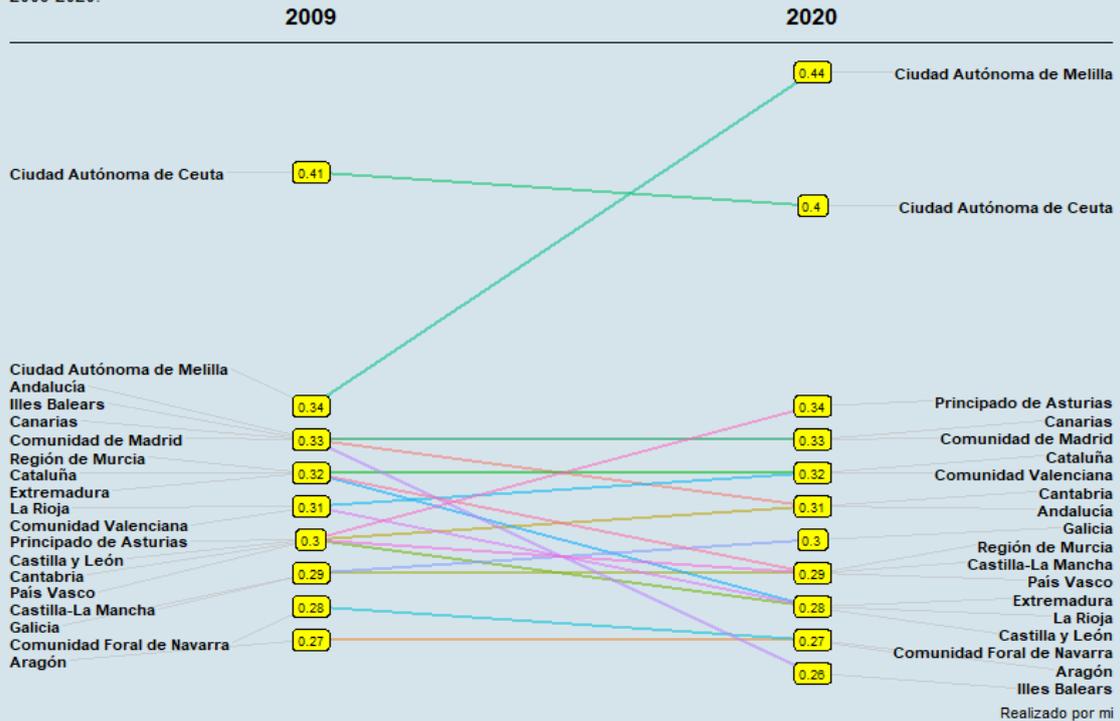
Lo que podemos observar mediante la relación de estos dos gráficos, es que a pesar de que el porcentaje de hogares en riesgo de pobreza se mantienen más o menos constante, el porcentaje de personas en riesgo de pobreza, en situación de privación material severa y/o que vive en hogares con intensidad de trabajo escasa está en auge en la mayoría de comunidades.

## Evolución Tasa de pobreza AROPE por C.A. 2009-2020.



Por últimos, vamos a analizar la evolución que ha presentado la desigualdad a partir de la distribución de la renta de la población Española.

## Evolución Índice de Gini por C.A. 2009-2020.



En el año 2009, ya existía una gran desigualdad con respecto a los ingresos en las comunidades, el cual más o menos se ha mantenido similar en el año 2020 en referencia a sus valores, habiendo grandes cambios comunidad por comunidad. Melilla al igual que ha tenido un gran aumento en la tasa de pobreza, su desigualdad ha crecido abismalmente, pasando de 0,34 a 0,44 medidos en el índice gini y se ha situado como la comunidad con mayor desigualdad, seguida de Ceuta con un 0,40. Ambas comunidades tienen los valores más extremos de este indicador con gran diferencia de las demás.

Las islas Baleares ha presentado un enorme descenso en cuanto a este indicador que ha pasado de ser una de las comunidades con mayor desigualdad a ser la que menor desigualdad presenta en referencia a los ingresos. El resto de comunidades se han mantenido más o menos constante o con leves bajadas o subidas.

## 4 Bibliografía

- Almató, M., Cortés, Á., A.Rey, R., Simón, D., & Werckmeister, A. (2015). *Análisis cuantitativo de la desigualdad de renta regional de España*.
- Anais Santourian, E. N. (2014). *Working paper with the description of the 'Income and living conditions dataset'*. Eurostat.
- Charrad, M., & Ghazzali, N. (s.f.). *NbClust: Determining the Best Number of Clusters in a Data Set*. Obtenido de <https://cran.r-project.org/web/packages/NbClust/index.html>
- Economipedia. (s.f.). *Curva Lorenz*. Obtenido de <https://economipedia.com/definiciones/curva-de-lorenz.html>
- (2021). *Encuesta de Condiciones de Vida. Ficheros de usuario del año 2020 de la encuesta*. INE.
- Eurostat. (s.f.). Obtenido de <https://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do>
- Eurostat. (s.f.). Obtenido de <http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do>
- Eurostat. (s.f.). Obtenido de <http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do>
- Eurostat. (s.f.). Obtenido de <https://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do>
- EUROSTAT. (s.f.). Recuperado el 2022, de <https://ec.europa.eu/eurostat/web/income-and-living-conditions/methodology>
- Foster, J., & Sen, A. (1973). On economic inequality. En A. Sen, & J. Foster, *On economic inequality* (págs. 24-46). Oxford.
- Grupo de Expertos de Calidad de Vida. (2017). *Final report of the expert group on quality of life indicators*. Eurostat.
- Iannone, R. (2022). *gt: Easily Create Presentation-Ready Display Tables*. Obtenido de <https://cran.r-project.org/web/packages/gt/index.html>
- IECA. (2017). *Memoria Técnica de la Actividad "Encuesta de Condiciones de Vida en Andalucía"*. IECA. Obtenido de <https://www.juntadeandalucia.es/institutodeestadisticaycartografia/ecv/>
- Imedio Olmedo, L. (2005). *Estadística Española*.

- INE. (2022). *Encuesta de condiciones de vida*. Obtenido de [https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica\\_C&cid=1254736176807&menu=ultiDatos&idp=1254735976608](https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176807&menu=ultiDatos&idp=1254735976608)
- INE *Resultado Nacionales ECV*. (s.f.). Obtenido de <https://www.ine.es/dynt3/inebase/es/index.htm?padre=1927&capsel=1922>
- Instituto de Estadística de Cataluña*. (s.f.). Obtenido de Idescat: <https://www.idescat.cat/pub/?id=ecv&lang=es>
- Instituto Nacional Estadística. (2005). *Encuesta de Condiciones de Vida. Metodología*. Madrid.
- Kassambara, A. (2020). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. Obtenido de <https://cran.r-project.org/web/packages/factoextra/index.html>
- Kuhn, M., & Wickham, H. (2019). *tidymodels: Easily Install and Load the 'Tidymodels' Packages*. Obtenido de <https://cran.r-project.org/web/packages/tidymodels/index.html>
- Kunst, J. (2022). *highcharter: A Wrapper for the 'Highcharts' Library*. Obtenido de <https://cran.r-project.org/web/packages/highcharter/index.html>
- Llera Lasobras, S. (2019). *Análisis de la pobreza y exclusión social en España*. TFG.
- Maechler, M. (2022). *cluster: "Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al*. Obtenido de <https://cran.r-project.org/web/packages/cluster/index.html>
- Moran, I. (2021). *Create and combine tables with a ggplot2/patchwork syntax*. Obtenido de <https://github.com/ianmoran11/mtable2>
- Núñez Velázquez, J. J. (2006). La desigualdad económica medida a través de las curvas de Lorenz. *Revista de métodos cuantitativos para la economía y la empresa*, 67-108.
- Powell, C. (2020). *CGPfunctions: Powell Miscellaneous Functions for Teaching and Learning Statistics*. Obtenido de <https://cran.r-project.org/web/packages/CGPfunctions/index.html>
- S.Handcock, M. (2022). *reldist: Relative Distribution Methods*. Obtenido de <https://cran.r-project.org/web/packages/reldist/index.html>
- Wickham, H. (2021). *tidyverse: Easily Install and Load the 'Tidyverse'*. Obtenido de <https://cran.r-project.org/web/packages/tidyverse/index.html>
- Wickham, H. (2022). *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. Obtenido de <https://cran.r-project.org/web/packages/ggplot2/index.html>
- Wikipedia. (s.f.). *Coeficiente de Gini*. Obtenido de [https://en.wikipedia.org/wiki/Gini\\_coefficient](https://en.wikipedia.org/wiki/Gini_coefficient)
- Wikipedia. (s.f.). *Curva de Lorenz*. Obtenido de [https://en.wikipedia.org/wiki/Lorenz\\_curve](https://en.wikipedia.org/wiki/Lorenz_curve)

Wikipedia. (s.f.). *Índice Atkinson*. Obtenido de [https://en.wikipedia.org/wiki/Atkinson\\_index](https://en.wikipedia.org/wiki/Atkinson_index)

Wikipedia. (s.f.). *Índice Theil*. Obtenido de [https://en.wikipedia.org/wiki/Theil\\_index](https://en.wikipedia.org/wiki/Theil_index)

Zeileis, A. (2014). *Ineq: Measuring Inequality, Concentration, and Poverty*. Obtenido de <https://cran.r-project.org/web/packages/ineq/index.html>