

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS MATEMÁTICAS
MÁSTER EN ESTADÍSTICAS OFICIALES E INDICADORES SOCIALES Y
ECONÓMICOS



TRABAJO FIN DE MÁSTER
2021-2022

Uso de Datos de Comercio Exterior para la Depuración e Imputación de la Cuenta Satélite de Flujos de Materiales

Autor:

Alberto Ismael Usero Lara

Tutores UCM:

David Salgado Fernández

Elena Rosa Pérez

Tutora INE:

Julia Maestro López



08 de septiembre de 2022

Resumen/Abstract

Resumen: En un contexto actual, en el que la presión política y medioambiental sobre los recursos materiales es máxima, se plantea de absoluta importancia la elaboración de unas estadísticas de calidad en este sentido para una correcta toma de decisiones. Este trabajo tiene su origen, por tanto, en dar un impulso a este aspecto en la compilación de la cuenta de importaciones y exportaciones físicas de materiales (cuentas de flujos de materiales) mediante el uso de una nueva fuente de datos. Para ello, primeramente, se ha elaborado la metodología oportuna para la gestión de unas estadísticas procedentes de registros administrativos, efectuando las consultas necesarias con los expertos en la materia. A continuación, se han expuesto los diferentes procesos de depuración e imputación desempeñados en el análisis de datos incongruentes de la base de datos utilizada. Posteriormente, se han puntualizado los avances y los desafíos contemplados con respecto a lo que estaba previamente confeccionado. Y para concluir, se han expuesto unas futuras líneas de actuación que son contempladas para la continuación del proyecto iniciado.

Abstract: Nowadays, the political and environmental pressure on material resources came to the maximum. It is of absolute importance the elaboration of quality statistics on this matter for a correct decision making. This work has its origin in giving an impulse to this key aspect: quality improvement of the physical imports and exports of materials compilation account (material flow accounts) using a new data source. Firstly, the appropriate methodology for the management of statistics from administrative records has been developed with the inevitable expert consultation. Secondly, the different edit and imputation processes have been performed in the analysis of incongruous data from the used database. Thirdly, it has been pointed out the progress and the challenges faced regarding what was previously completed. To conclude, some future lines of action that are considered for the continuation of the initiated project have been revealed.

Índice General

Declaración de no plagio.....	I
Resumen/Abstract	II
Índice de Tablas	IV
1. Introducción.	1
2. Cuentas de Flujos de Materiales para el Total de la Economía.....	4
Dimensiones en la cuenta de flujos de materiales.	6
Flujos de materiales entre una economía y el resto del mundo.	7
Clasificación de materiales en las Importaciones y Exportaciones físicas.....	9
3. Fuente de datos: Estadísticas elaboradas a partir de registros administrativos.	12
Importaciones y Exportaciones físicas: De datos anuales a datos mensuales máxima desagregación AEAT.	14
4. Marco Teórico.....	19
Experiencia en depuración y características adyacentes a las bases de datos de comercio exterior.	19
Depuración de datos: Edits, outliers e imputaciones.....	23
5. Metodología y Resultados.....	28
Metodología.	28
Resultados.	36
6. Conclusiones y Trabajo Futuro.	45
Bibliografía.	47
Anexo.....	51
A1. Clasificación de materiales en la CFM-TE según el tipo de flujo.....	51
A2. Diseño de registro ficheros datos comercio exterior AEAT.	56
A3. Códigos de variables Series Mensuales Máxima Desagregación.....	58
A4. Edits y resumen comext_Schema.....	70

Índice de Tablas

Tabla 2.1. Definición de los flujos físicos de materiales de la CFM-TE.	5
Tabla 2.2. Indicadores absolutos derivados de los Flujos directos de Materiales.	6
Tabla 2.3. Resumen de transacciones en comercio internacional y tratamiento en CFM-TE. (Eurostat 2017a). Caso de España como país que reporta la información.	8
Tabla 3.1. Resumen variables Series Mensuales Máxima Desagregación. Códigos de variables incluidos en A3.	17
Tabla 5.1. Resumen tamaño archivos series mensuales máxima desagregación.	28
Tabla 5.2. Resumen de paquetes utilizados en programa R.	30
Tabla 5.3. Resumen de pasos en la lectura de archivos.	31
Tabla 5.4. Resumen de pasos en el procesamiento de datos.	34
Tabla 5.5. Variación porcentual entre los datos obtenidos de máxima desagregación sin realizar tratamiento alguno y realizando el tratamiento de atípicos por el RIQ e imputando por la mediana para las Exportaciones.	37
Tabla 5.6. Variación porcentual entre los datos obtenidos de máxima desagregación sin realizar tratamiento alguno y realizando el tratamiento de atípicos por el RIQ e imputando por la mediana para las Importaciones.	38
Tabla 5.7. Variación porcentual entre los datos publicados por el INE y los datos obtenidos de máxima desagregación sin realizar tratamiento alguno para las Exportaciones.	39
Tabla 5.8. Variación porcentual entre los datos publicados por el INE y los datos obtenidos de máxima desagregación sin realizar tratamiento alguno para las Importaciones.	41

1. Introducción.

Desde que un 13 de marzo de 2020 la Organización Mundial de la Salud oficializara la existencia de una pandemia causada por la COVID-19, hasta la actualidad, hemos podido observar cómo han ido sobreviniendo una serie de circunstancias que han puesto de manifiesto la importancia de la interconexión comercial mundial y cuan graves problemas originan los desajustes en la misma. Algunos ejemplos son el aumento en la demanda global cuando la mayoría de los sectores productivos aún se encontraban bajo restricciones por la pandemia, [el acopio de materias primas llevado a cabo por parte de China](#), [las tensiones en el transporte marítimo](#) o la invasión de Ucrania.

Las repercusiones en el plano del comercio exterior que tuvieron estas circunstancias y decisiones fueron variadas en el tiempo. Al principio pudimos observar cómo aumentos en los precios y problemas de abastecimiento de algunas materias primas (p.ej. los semiconductores, los productos químicos, los plásticos o la madera) provocaron la imposibilidad de atender los pedidos en los plazos requeridos (Kataryniuk et al., 2021). Y justo cuando se atisbaba una cierta recuperación gradual de las consecuencias de la pandemia, llegó una nueva crisis de materiales con la guerra en Ucrania. Siendo especialmente destacable los récords en la subida en los precios del petróleo y del gas natural por parte de materias primas energéticas (Alonso et al., 2022).

Sin lugar a duda, estas crisis han evidenciado la gran dependencia actual de la Unión Europea en una serie de materias primas clave y que le ha conducido a replantearse las políticas adoptadas en materia de comercio exterior. Para una correcta toma de decisiones las autoridades gubernamentales se apoyan en las llamadas Estadísticas Oficiales, unas estadísticas que deben ser de calidad y decisivas en momentos tan complicados como los expuestos anteriormente.

Una [calidad en las Estadísticas Oficiales](#) que debe provenir de diferentes aspectos. No se trata solo de los resultados que se obtienen o de los procesos que las producen, sino también el entorno en el que operan las instituciones productoras de estas estadísticas. Así, desde Europa, las autoridades de producción estadística se encuentran absolutamente comprometidas con la calidad al promover unas mejores prácticas en las estadísticas mediante principios, indicadores o métodos. Esta política de calidad se encuentra establecida por Reglamento [223/2009](#) del Parlamento Europeo y del Consejo de 11 de marzo de 2009, y asegurada mediante el [Código](#)

[de Buenas Prácticas de las Estadísticas Europeas](#) (ES CoP, por sus siglas en inglés) y el [Marco Común de Calidad del Sistema Estadístico Europeo](#) (ESS QAF, por sus siglas en inglés)¹.

La creación de Estadísticas Oficiales de calidad centra el objetivo de este trabajo, el cual ha sido realizado en el contexto de una Beca de Postgrado en el Instituto Nacional de Estadística en la Subdirección General de Estadísticas de Medioambiente, Agrarias y Financieras. Mis tareas han estado relacionadas en el entorno de las [Cuentas de Flujos de Materiales para el Total de la Economía \(CFM-TE\)](#), en inglés [Economy-wide material flow accounts \(EW-MFA\)](#). Estas cuentas no solo proveen de información sobre la interacción entre la economía de un país y el medioambiente, sino que también reportan sobre los intercambios de materiales con el resto de economías, esto es, información sobre cuán dependientes somos del exterior.

Las CFM-TE se encuentran integradas dentro de la denominada Contabilidad Medioambiental (CMA), cuyo principal objetivo sería el de medir la contribución del medioambiente a la economía y la presión que esta última ejercería sobre los recursos naturales. Estas estadísticas se presentan de gran importancia, puesto que la presión que estamos haciendo sobre el medioambiente es, hoy en día, un problema mayor que cualquier retraso que pudiera haberse producido en algunas entregas durante la ya mencionada crisis de los semiconductores.

En este contexto, se hace absolutamente necesario una combinación de lo expuesto anteriormente, esto es, unas estadísticas medioambientales y de calidad. Hasta ahora hemos estado determinando la salud de nuestras economías meramente a través de indicadores que exclusivamente tenían en cuenta lo producido, en concreto, hemos estado aplicando en exclusiva el indicador del Producto Interior Bruto (PIB) y derivados. Es hora de no usar únicamente estos indicadores, sino también tener en cuenta las repercusiones que tienen la producción llevada a cabo. Puesto que una producción desmesurada y sin control nos puede estar haciendo pobres en realidad, especialmente si tenemos en cuenta la desolación de los recursos naturales (Iriarte, 2021).

A través de la CMA y de los indicadores que se pueden elaborar a partir de ella, podemos conocer la dependencia de materiales de nuestra economía y conocer cuáles de estos materiales provienen del exterior, esto provee de una información de calidad de cara a la toma de decisiones en políticas ambientales. Políticas que, en este sentido, han ido tomando fuerza en estos años como por ejemplo el Pacto Verde Europeo (con el objetivo de lograr una neutralidad climática para 2050), la Estrategia Española de Economía Circular (EEEC, España

¹ El ES CoP es el eje del marco de calidad de las estadísticas a nivel europeo. Estableciendo una serie de principios para el desarrollo, producción y difusión de unas estadísticas europeas de calidad. Se tratan de 16 principios los cuales son complementados por el ESS QAF para asegurar su correcta implementación. El ESS QAF tiene como principal objetivo de acompañar a la ES CoP facilitando una serie de herramientas más detalladas para su implementación.

Circular 2030) (como solución al problema de la Economía Lineal, para que los materiales y recursos se puedan mantener en la economía por el mayor tiempo posible) o dentro de la Agenda 2030 con los Objetivos de Desarrollo Sostenible (como serían el 12. Producción y consumo responsables, el 13. Acción por el clima o el 15. Vida de ecosistemas terrestres).

Con este trabajo se espera contribuir en todo este proceso de toma de decisiones medioambientales, proveyendo de una mejora de la calidad (acuracidad, oportunidad, precisión) de los datos publicados por el Instituto Nacional de Estadística (INE) en las CFM-TE, relativos a las Importaciones y Exportaciones físicas de materiales. Se va a realizar mediante una automatización de la lectura y puesta a punto de una nueva fuente de datos con información más desagregada sobre las transacciones de comercio exterior ([datos máxima desagregación](#), cuya información tiene su origen en fuentes de Registros Administrativos). Todo ello para, posteriormente, poder realizar una exploración, detección y corrección de datos incongruentes de calidad. El análisis y los diferentes procesos de depuración e imputación han sido realizados con paquetes de R diseñados para su uso en la producción de estadísticas oficiales.

Para abordar este objetivo hemos estructurado el trabajo de la siguiente manera: el capítulo 2 expone unos puntos básicos sobre las CFM-TE y determinados puntos claves de la cuenta de importaciones y exportaciones físicas de materiales para situar al lector en un contexto necesario. En el capítulo 3 se muestra en que marco se sitúan los registros administrativos en las Estadísticas Oficiales. Adicionalmente, hablamos de los registros que originan la fuente de datos utilizada en este proyecto y exponemos las diferentes variables que son generadas a través de ellos. En el capítulo 4 describimos, por un lado, las distintas dificultades y líneas de actuación que otros organismos han tenido en la compilación de la cuenta de importaciones y exportaciones físicas. Por otro lado, situamos un marco teórico sobre la detección de valores atípicos y la manera de imputarlos. El capítulo 5 desarrolla punto por punto las diferentes acciones que se han llevado a cabo para la elaboración de este trabajo y la obtención de unos resultados, los cuales son presentados como una comparativa con los datos que actualmente se encuentran publicados por el INE. Y finalmente, en el capítulo 6, exponemos en primer lugar las diferentes conclusiones que hemos ido obteniendo a lo largo del proyecto, y, en segundo lugar, el trabajo futuro que consideramos necesario para la consecución de unos objetivos futuros.

2. Cuentas de Flujos de Materiales para el Total de la Economía.

Las Cuentas de Flujos de Materiales ([CFM-TE](#)) se encuentran integradas en el [Sistema de Contabilidad Ambiental y Económica 2012 – Marco Central](#) (en inglés: System of Environmental Economic Accounting 2012 – Central Framework: SEEA-CF 2012). La SEEA-CF 2012 es un marco conceptual que describe las interacciones entre la economía y el medioambiente. Distingue entre cuentas de flujos físicos y cuentas de activos físicos a la par de permitir el desarrollo de indicadores, agregados, así como la predisposición del uso y por tanto de la disponibilidad de recursos provenientes de la naturaleza.

Las CFM-TE estaría ubicada en cuentas de flujos físicos y se trataría de una estadística de síntesis, realizada en España por el Instituto Nacional de Estadística (INE)² por mandato de la Oficina de Estadística de la Unión Europea (Eurostat), a partir de resultados obtenidos de diversas fuentes. Se encuentra recogida en el [Plan Estadístico Nacional 2021-2024](#), por lo que estaríamos hablando de una operación de realización obligatoria.

Su compilación viene regulada a través del Reglamento (UE) [2011/0691R](#) del Parlamento Europeo y del Consejo, de 6 de Julio de 2011, relativo a las cuentas económicas europeas medioambientales (L192 22/07/2011).³

Las CFM-TE registran los flujos físicos de materiales en la economía nacional (interacción de la economía nacional con el medioambiente, así como los intercambios de materiales con otras economías, es decir, con el resto del mundo). La cuenta abarcaría los materiales sólidos, gaseosos y líquidos, excepto los grandes volúmenes de agua y aire⁴.

Las CFM-TE son denominadas cuentas satélites, dado que integran de manera coherente con el Sistema de Cuentas Nacionales (SCN): Una cuenta satélite sería aquella en la que se ha realizado una expansión del marco central de la contabilidad nacional y cuyo principal objetivo sería la de centrar el foco en ciertos aspectos que, aun encontrándose en el propio marco, lo hacen de manera limitada. El análisis del medioambiente, con las CFM-TE se beneficia de este modo tanto de la inclusión de elementos incluidos en el marco central, como de elementos complementarios (como son las unidades físicas) (SCN 2008, párr. 2.166).

² [Ficha Inventario Operaciones Estadística \(IOE\) actual: 30086](#)

³ Completado por [2016/0172R](#), [2016/0335D](#), [2015/2174R](#).

⁴ Los flujos de agua representan unas cantidades de una magnitud mayor que el resto de los materiales. Incluirlo distorsionaría la cuenta y haría algo complicado el estudio de patrones en el resto de los materiales. En el caso de los flujos de aire, ciertos gases son incluidos en los Elementos de Balance.

Cuando hablamos de flujos de materiales nos referimos a aquellas interacciones que tienen lugar dadas las operaciones llevadas a cabo por las unidades residentes⁵ de una nación. Todo ello, independientemente de donde se ubiquen geográficamente dichas unidades. Respetando así el principio de residencia y, por tanto, manteniendo la coherencia con el SCN.

La periodicidad de las CFM-TE (entradas, salidas y variaciones de las existencias de materiales en una economía) es anual. Son transmitidas (establecido en Reglamento) 24 meses después del año de referencia en los datos (Maestro et al., 2021). Y son publicadas en unidades de masa (toneladas en el caso de la publicación del INE, kilotoneladas en la de Eurostat).

Tabla 2.1. Definición de los flujos físicos de materiales de la CFM-TE.	
Extracción Nacional (EN)	Materiales sólidos, líquidos y gaseosos (excl. aire y agua) extraídos del medio natural en el territorio de una nación y que son usados como insumos para producción o consumo en la economía nacional.
Importaciones y Exportaciones físicas de materiales	Todas las mercancías importadas o exportadas entre economías nacionales, que sean metodológicamente correcta su inclusión en la cuenta y que son reportadas en unidades físicas de masa.
Output Nacional Procesado a la Naturaleza (ONP)	Materiales que entran al medio natural, tanto en el mismo momento en que se produce los procesos de producción o de consumo, como una vez concluidos estos procesos. Estaríamos hablando de las emisiones al aire o al agua y también de los vertidos de residuos, como aquellos materiales que causan contaminación difusa.
Elementos de Balance (de entrada y de salida)	Estimación de los flujos de materiales que están incluidos en un lado del balance de masa, pero no se encuentran en el otro lado de este.

Su importancia va desde la oportunidad que da a los países para llevar un control sobre la presión medioambiental ejercida sobre ciertos recursos medioambientales, hasta la

⁵ Una unidad económica sería residente de un país cuando presenta su centro de interés económico en el territorio geográfico de ese país, esto sería considerado de esta manera si se ha dedicado, al menos por un año, a realizar actividades económicas en ese lugar (Maestro et al., 2021).

dependencia que pueden tener de ciertos recursos materiales que son comercializados con el resto del mundo (Eurostat 2001, párr. 537). Estas, y otras muchas conclusiones, son accesibles gracias a la combinación de los resultados de los flujos mostrados en la tabla 2.1 en los distintos indicadores que mostramos en la tabla 2.2. [Indicadores per cápita, de productividad y de intensidad de materiales](#) son obtenidos a través de los datos absolutos.

Tabla 2.2. Indicadores absolutos derivados de los Flujos directos de Materiales.	
Input Directo de Materiales (IDM)	EN + Importaciones Físicas
Consumo Nacional de Materiales (CNM)	EN + Importaciones Físicas – Exportaciones Físicas
Balance Comercial Físico (BCF)	Importaciones Físicas – Exportaciones Físicas
Acumulación Neta de Materiales (ANM)	CNM + Elementos de Balance de entrada – (OPN + Elementos de Balance de salida)

Dimensiones en la cuenta de flujos de materiales.

En un sentido más específico, y para una mejor comprensión de la cuenta, esta puede ser entendida como un marco estadístico contable que persigue describir la manera en que una economía interactúa en diversas dimensiones. Interesará observar una u otra dimensión en función de la investigación que se quiera llevar a cabo.

Una dimensión podría verse en las interacciones que ocurren entre la propia economía y el entorno natural. En ella encontramos del lado de la entrada de materiales la EN. La cual vendrían a ser los **inputs de materiales** que tienen su origen en una extracción del medioambiente llevada a cabo por los seres humanos (de manera manual o a través de medios tecnológicos) (Eurostat 2001, párr. 3.10). Mientras que, del lado de la salida, en cambio, tendríamos la cuenta del OPN. Está serían aquellos outputs de materiales que son emitidos al aire, a la tierra o al agua. Esto es, aquellos materiales que son considerados como **residuos** una vez se ha producido algún proceso en la economía.

Una segunda dimensión puede verse en el estudio de las entradas y salidas de materiales en su sentido más amplio. Las entradas a la economía nacional incluirían las tablas de EN, las Importaciones Físicas y los Elementos de Balance de Entrada. Mientras que las salidas de materiales son recogidas en las tablas de ONP, las Exportaciones Físicas y los Elementos de Balance de Salida.

Otro interés puede recaer en el estudio de la interacción que acontece entre la economía en cuestión y el resto de las economías (resto del mundo). Se trataría de la dimensión de los flujos

físicos comerciales (flujos de **productos**⁶) que ocurren en el comercio internacional entre economías. Los inputs recogidos en esta dimensión se corresponden con el conjunto de importaciones físicas realizadas por una economía, mientras que los outputs serían las exportaciones físicas (Eurostat, 2018).

Flujos de materiales entre una economía y el resto del mundo.

Esta dimensión es la que centra el interés de este trabajo y, por ello, la que vamos a desarrollar más detenidamente. Como bien decíamos en el apartado anterior, esta dimensión de las CFM-TE se basaría en la compilación de los flujos físicos comerciales que acontecen de manera anual en una economía. En un sentido más técnico, estaríamos hablando de aquellas transacciones realizadas por unidades residentes de un país con unidades no residentes⁷, lo cual se correspondería con las transacciones entre la economía nacional informante y el resto del mundo.

El criterio principal para establecer que se ha realizado una transacción es el **cambio de propiedad**. Este cambio de propiedad entre un residente y un no residente se trataría de una exportación y al revés sería una importación. Esto es mantenido para casi todo tipo de transacciones comerciales, existiendo la excepción de aquellos *bienes enviados al extranjero para ser procesados*. En este caso, la CFM-TE deberá tenerlos en cuenta, aunque no se produzca un cambio de propiedad y no se mantengan en Cuentas Nacionales (Eurostat 2017a). Todo ello, siguiendo el Sistema de Contabilidad Ambiental y Económica - Marco Central (SCAE-MC)⁸.

A modo de resumen, indicamos en la tabla 2.3 los diferentes tipos de transacciones categorizadas y si debe o no estar presente en la contabilización de importaciones y exportaciones físicas en la CFM-TE.

⁶ Que tienen masa y volumen y que pueden ser representados en unidades de masa.

⁷ Siguiendo la coherencia con respecto a Cuentas Nacionales, en la CFM-TE se sigue el criterio de que una unidad es residente de una nación cuando presenta su principal centro de interés económico en el territorio económico de esa nación. Esto es, cuando la unidad desarrolla sus actividades económicas en ese territorio por un periodo de un año o más (Eurostat 2021, párr.1.61). Por ello, la definición de economía nacional no se corresponde con la de actividades que tienen lugar en el territorio nacional. Algunas actividades pueden desarrollarse más allá del territorio nacional (p.ej. el transporte aéreo y marítimo internacional), pero contar para la economía nacional por ser unidad residente (Eurostat 2018, párr. 42).

⁸ El SCAE-MC es un estándar estadístico internacional con el que medir el MA y la relación de este con la economía. Trata la medición de tres grupos principales: Flujos medioambientales, Existencias de activos medioambientales y la Actividad económica relacionada con el MA. La información medioambiental puede integrarse con la información económica gracias a que la SCAE-MC presenta unos conceptos alineados con el SCN.

Tabla 2.3. Resumen de transacciones en comercio internacional y tratamiento en CFM-TE. (Eurostat 2017a). Caso de España como país que reporta la información.		
Tipo de Transacción	Definición de la transacción	Tratamiento en CFM-TE
Transacciones Intrastat	Productos comercializados entre España y otros Estados Miembro en los cuales se produce el cambio de propiedad entre los mismos.	Deben incluirse
Transacciones Extrastat	Productos comercializados entre España y países Extracomunitarios en los cuales se produce el cambio de propiedad entre los mismos.	Deben incluirse
Reexportaciones	Productos que fueron producidos en un país extranjero, son importados a España (cambian de propiedad a la de una unidad de residencia), pero que resultan ser exportados sin una transformación sustancial.	Deben incluirse
Tránsito	Productos cuyas formalidades aduaneras son realizadas en el extranjero y que únicamente atraviesan España para acabar en otro lugar.	No deben incluirse
Cuasi-Tránsito	<p>En el caso de los bienes que entran: Productos no comunitarios que entran en España, en el que un no residente realiza el despacho aduanero antes de que la mercancía sea enviada a otro Estado Miembro.</p> <p>En el caso de los bienes que salen: Productos que llegan (de otro estado miembro) a España, y nuevamente un no residente realiza el despacho aduanero para la exportación a un Estado No Miembro.</p>	No deben incluirse

<p>Productos para ser procesados</p>	<p>En el caso de los bienes que entran: Una unidad residente recibe los bienes para ser procesados (no se produce cambio de propiedad) para una vez realizada la transformación vuelvan a ser enviados al país del que vinieron. Bienes que fueron enviados para ser procesados en el extranjero y que son recibidos una vez se ha realizado la transformación por una unidad no residente.</p> <p>En el caso de los bienes que salen: Una unidad residente de España envía bienes que fueron recibidos para su procesamiento (nunca tuvo la propiedad). Una unidad residente envía bienes a una unidad no residente para que los bienes sean procesados (no se produce cambio de propiedad).</p>	<p>Deben Incluirse</p>
<p>Productos comprados por unidades residentes en el extranjero. Productos vendidos a unidades no residentes en España.</p>	<p>Bienes que entran: Unidad residente de España que obtiene productos de no residentes fuera del territorio. Bienes que salen: Unidad residente que vende productos a no residentes en España.</p>	<p>Debe incluirse</p>
<p>Productos bajo comercialización</p>	<p>Unidad residente de España adquiere productos de un no residente y los vende de nuevo a un no residente sin que los bienes entren en territorio español.</p>	<p>No deben incluirse</p>

Clasificación de materiales en las Importaciones y Exportaciones físicas.

La CFM-TE clasifica los flujos físicos de dos maneras, una sería la descomposición según el tipo de flujo (flujos descritos en tabla 2.1) y otra sería en función del tipo de material del que estemos hablando.

Así, los diferentes flujos de materiales se encuentran a su vez clasificados según el tipo de material de acuerdo con las clasificaciones recogidas en el Anexo III, sección 5 del Reglamento (UE) [2011/0691R](#) del Parlamento Europeo y del Consejo, de 06/07/2011.

Como cualquier otra clasificación, la de materiales, trata de un conjunto organizado de categorías mutuamente excluyentes. Presentadas en una jerarquía donde las principales categorías de flujos de materiales vienen clasificadas al nivel de 1 dígito (categoría de material) y a partir de ahí se va especificando cada vez más de que tipo de material se trata, 2 dígitos (clase de material), 3 dígitos (diferentes grupos de materiales de la misma clase) y 4 dígitos (subgrupo de materiales).⁹

Adicionalmente, para el envío a Eurostat (y con carácter voluntario), las importaciones y exportaciones físicas deben ser clasificadas en función de su estado de manufacturación. Por lo que adicionalmente se cuenta con una clasificación donde existe un único nivel con tres categorías que diferencian los productos según la fase, o estado de manufacturación, en el que se encuentren: Productos Básicos, Productos Semi-Elaborados y Productos Acabados.¹⁰

Esto último es relevante en cuanto a la posibilidad de comparativa de los materiales obtenidos a través del comercio exterior (importaciones y exportaciones físicas) con los materiales presentes en la EN. Todos aquellos productos englobados dentro de la categoría de Productos Brutos son altamente comparables de igual a igual con los materiales presentes en la EN. Por el contrario, en medida que el producto ha ido sufriendo un procesamiento mayor, esa comparativa deja de ser posible.

Esto último es de gran importancia e iría ligado a la pequeña peculiaridad con respecto a cómo son clasificadas las importaciones y exportaciones físicas en esta cuenta. Como bien ha sido descrito con anterioridad, son obtenidas a través de flujos de productos y no de materiales per se. Entonces, ¿cómo es obtenida esa equivalencia entre un producto y un tipo de material?

Por una parte, decir que la principal fuente de información para elaborar esta cuenta son las [estadísticas de comercio exterior](#)¹¹. Estas presentan información en unidades físicas y monetarias para las transacciones realizadas y recogen la información de las transacciones acontecidas en función del producto comercializado con los distintos países, utilizando una

⁹ Las clasificaciones de materiales en la CFM-TE, según el tipo de flujo y material puede ser consultada en el Anexo, apartado A1.

¹⁰ Productos básicos (crudos o sin procesar): Producidos en la agricultura, silvicultura, pesca o minería. Productos Semi-Elaborados: Productos que han sufrido algún tipo de procesamiento, pero que aún no se corresponden a Productos Acabados.

Productos Acabados: Productos que no van a sufrir un mayor procesamiento o sucesivas transformaciones.

¹¹ En el caso de España, estas estadísticas son elaboradas por la Agencia Estatal de Administración Tributaria.

clasificación de los productos, que en este caso se trataría de la clasificación por [Nomenclatura Combinada \(NC\)](#), la cual se trataría de la clasificación usada a nivel Comunitario para recopilar y procesar los datos de este tipo.

Por otra parte, la equivalencia es elaborada mediante una correspondencia entre la clasificación de materiales de la CFM-TE y los diferentes productos de la NC. Esto es así debido a que un producto presenta más de un tipo de material y fue establecido que el tipo de material asignado a un producto en concreto sería aquel que se presenta de manera predominante (mayor peso físico) en el mismo. Esta correspondencia, al igual que la existente entre el estado de manufacturación y los productos, están recogidas en una serie de [Anexos](#) que son actualizados por Eurostat de manera periódica.

3. Fuente de datos: Estadísticas elaboradas a partir de registros administrativos.

Las fuentes de datos manifiestan una diferencia clave que las distingue entre sí y esa diferencia va a ser la finalidad que presentan. Por ejemplo, las [fuentes de datos administrativas](#) se diferencian de las fuentes estadísticas en que los datos que son recogidos por instituciones del sector gubernamental se utilizan con fines administrativos.

En la actualidad, las Estadísticas Oficiales podrían ser nutridas a través de diversas fuentes de datos. La importancia de la fuente de datos utilizada es que se encuentre en línea con la finalidad estadística que se quiera llevar a cabo. Teniendo esto en cuenta, a parte de las fuentes de datos administrativos se tiene constancia de otras dos fuentes de datos más, reconocidas en las Estadísticas Oficiales, como son las encuestas y los datos digitales.

Las encuestas son ampliamente utilizadas en las instituciones generadoras de estadísticas gubernamentales. Una encuesta es una investigación de las características de una población mediante la información recogida a una muestra de esta (OCDE, 2007).

Mientras que los datos digitales, cuya utilización para realizar Estadísticas Oficiales aún se encuentra en una fase no muy avanzada, podrían ser definidos como *“el conjunto de unidades y datos derivados de una fuente digital, siendo esto un sistema de información digital”* (Salgado & Oancea, 2020).

Como apuntábamos anteriormente, las estadísticas elaboradas a partir de registros administrativos son aquellas desarrolladas gracias a una fuente de datos inicial cuyo principal interés es la de suplir fines administrativos. Existirán casos en los que los fines administrativos no se encuentren alineados con los estadísticos, y por lo que tanto en el INE como en Otras Autoridades Nacionales (OANs)¹² sea necesaria la elaboración de unos registros estadísticos en base a la información administrativa.

Podemos observar como el uso de registros administrativos viene amparado e impulsado por Eurostat por estar presentes en puntos del ES CoP de la siguiente manera:

- Principio 2. Mandato de recogida de datos y acceso a los datos.

Las autoridades estadísticas no solo se encuentran autorizadas a la utilización de datos administrativos por ley, sino que desde un principio deben participar en su desarrollo para poder alinearlos a los fines estadísticos marcados.

¹² Es considerado como OAN a toda institución que ejerza una autoridad pública, que tenga la producción de estadísticas incluidas entre sus responsabilidades y que, a nivel nacional, deba proporcionar estadísticas elaboradas con fines europeos establecidos en el Programa Estadístico Europeo. (Eurostat, 2014).

- Principio 8. Procedimientos estadísticos adecuados.

Es necesario una cooperación entre los titulares de los datos administrativos y la autoridad (INE u OAN) que quiera explotarlos con fines estadísticos, para garantizar la calidad de estos.

Las definiciones y conceptos recogidos en los datos administrativos serían una buena aproximación a aquellos buscados en los propósitos estadísticos, siempre asegurándose mediante consultas a los responsables de estos datos.

- Principio 9. Carga no excesiva para los encuestados.

Los organismos responsables de las estadísticas deben controlar la carga de respuesta sobre los encuestados, teniendo un objetivo claro de reducirla paulatinamente. Este fin va muy ligado al uso de las fuentes de datos administrativos como alternativa a aquellas encuestas que puedan acarrear ciertas molestias a los informantes.

- Principio 10. Utilización de recursos de manera eficiente.

El uso de fuentes de datos administrativos presenta un fuerte potencial y así debe ser tenido en cuenta a la hora de aplicar las tecnologías de la información necesarias para su procesamiento estadístico.

En España, el anterior presidente del INE, Juan Manuel Rodríguez Póo, apuntaba en el [Pleno celebrado del Consejo Superior de Estadística \(CSE\)¹³ con fecha de 16 de diciembre de 2021](#) la importancia que está adquiriendo la información procedente de registros administrativos y las colaboraciones mantenidas con la AEAT en esta materia. Especialmente en la época en la que fue celebrada, puesto que durante ese año el trabajo de los servicios estadísticos estuvo marcada, al igual que el resto, por la crisis originada por COVID-19.

Además, continuando en la misma dirección, podemos ver como se ha producido una modificación de la Ley 12/1989, de 9 de mayo, de la Función Estadística Pública del artículo 10 en esta línea: “1. *Las estadísticas para fines estatales deberán tener como fuente prioritaria de información los datos contenidos en los registros administrativos, entendiendo por tales los recogidos en archivos o directorios del sector público*” (Disposición final segunda de la Ley 13/2022, de 7 de julio, General de Comunicación Audiovisual).

¹³ El CSE forma junto al Comité Interterritorial de Estadística (CITE) y la Comisión Interministerial de Estadística (CIME) los tres órganos consultivos de los servicios estadísticos de la Administración del Estado.

Importaciones y Exportaciones físicas: De datos anuales a datos mensuales máxima desagregación AEAT.

Como apuntábamos con anterioridad, la fuente de datos utilizada para la elaboración de las estadísticas de Importaciones y Exportaciones físicas son las [estadísticas de comercio exterior](#) producidas a partir de registros por la AEAT.

Son dos las fuentes de datos que se utilizan para la elaboración de las estadísticas de comercio exterior y su principal diferenciación se encuentra en el agente con quien se ha realizado la transacción (importación o exportación), esto es, si el agente es comunitario ([Metodología Intrastat](#)) o si el agente es extracomunitario ([Metodología Extrastat](#)).

Esta división a la hora de realizar la recopilación de datos proviene del momento en que se formalizó la creación del Mercado Común Europeo. A partir de 1993 se produjo una supresión de los trámites aduaneros tal y como se estaban llevando a cabo entre los Estados Miembro y por tanto se produjo una supresión también de las declaraciones que aportaban la información necesaria para elaborar estas estadísticas. A partir de entonces, se lleva a cabo un procedimiento diferente y ampliamente generalizado para la obtención de la información necesaria de países comunitarios: declaraciones de Intrastat.

Por un lado, para el comercio extracomunitario, la información para la elaboración de la estadística de comercio exterior es obtenida a través del Documento Único Administrativo (DUA). Este documento se trata de un formulario que es utilizado para la declaración de mercancías en la aduana (importaciones y exportaciones) entre los países comunitarios y los extracomunitarios. El DUA puede ser transmitido de manera electrónica (establecido en el capítulo 6º Presentación de DUAs por transmisión electrónica de datos (EDI) de la [Resolución de 15 de diciembre de 2009, del Departamento de Aduanas e Impuestos Especiales de la Agencia Estatal de la Administración Tributaria](#)).

Por otro lado, para el comercio comunitario, existe una singularidad por la que el responsable no deberá suministrar información sobre las transacciones. Esto ocurrirá si no supera un cierto umbral de valor estadístico determinado por el Ministerio de Economía. En el caso en el que se deba transmitir la información sobre la operación llevada a cabo, esta podrá ser aportada mediante DUA o a través de un formulario electrónico ubicado en la página web de la AEAT. La información también puede ser proporcionada a través de un impreso oficial ([Resolución de 27 de enero de 2009, de la Presidencia de la Agencia Estatal de Administración Tributaria](#)).

A partir de estos datos obtenidos a través de fuentes expuestas, la AEAT elabora diferentes productos en los cuales recoge la información relativa al comercio exterior de diferente forma, para que así el usuario pueda elegir la serie que mejor se adapte a la finalidad estadística que

tenga. En la web de la AEAT se presentan: Series Mensuales Máxima Desagregación, Series Mensuales Territoriales, Series Resúmenes Mensuales y Series Históricas Anuales (tanto territorial como resumen). A continuación, vamos a desarrollar de un modo más extenso las que han sido exploradas por el INE para la elaboración de las estadísticas sobre importaciones y exportaciones físicas de materiales para las CFM-TE. Además, en el anexo A2 se puede encontrar los diseños de registro de todas ellas.

- Series Históricas Anuales (resumen).

Alguna de la información más relevante que es recogida en esta serie sería el **Año** cuando se ha realizado la transacción, el **País** con el que se ha comercializado, la **Posición Estadística** (Código de NC) para poder realizar la debida identificación del producto que es comercializado, el peso en masa del producto comercializado en kg (**Peso Importación** por un lado y por otro el **Peso Exportación**), unidades suplementarias (unidades de medida suplementarias para ciertos productos, **Unidades Importación** y **Unidades Exportación**) o el valor monetario de la transacción en céntimos de euro (**Valor Importación** y **Valor Exportación**).

Estas series anuales han sido tradicionalmente las utilizadas para elaborar las estadísticas de importaciones y exportaciones físicas en los envíos a Eurostat y en las publicaciones en la web del INE para los años de referencia de 2008 a 2019. La principal desventaja que presenta el uso de esta serie sería la demora de dos años en la publicación de los datos por parte de la AEAT. Por ejemplo, para el año 2021 los últimos datos disponibles eran los de comercio exterior del año 2019. Lo que dificultaba la posibilidad de enviar un dato de calidad para las estimaciones avance de las importaciones y exportaciones físicas para el año 2020.

Así se tomó la decisión de utilizar los datos de las Series Resúmenes Mensuales para ese cometido, puesto que se publican los datos con simplemente un par de meses de retraso, contando por tanto con todos los datos para poder llevar a cabo el avance del año 2020.

- Series Resúmenes Mensuales.

Esta serie se presenta de un modo similar a la expuesta en el punto anterior. La principal diferencia con las series anuales es que en las mensuales no se ha realizado la agregación por mes en la comercialización de un cierto producto, con un cierto país. Por ello esta base de datos presenta un mayor nivel de desagregación. Se han agregado las transacciones realizadas para un cierto producto, con un cierto país en el mismo mes de un año en concreto. Por ello la única variable adicional con respecto a la serie

anterior sería la de **Mes** en la que se especifica en que mes ha acontecido la transacción.

Mediante el uso de esta serie se consiguió llevar a cabo una mejora en la calidad de las estadísticas mediante una mejora de la Oportunidad, puesto que nos permitió publicar datos con un menor intervalo entre el momento de difusión y el final del periodo de referencia con la precisión que proporciona el tener los datos disponibles para la totalidad del año 2020.

El principal problema que descubrimos, y que también se encontraba presente en los datos anuales, era que para ciertas transacciones teníamos dato para el valor monetario, pero en cambio el dato del peso físico nos aparecía como dato ausente. Siendo este último el más relevante para la consecución del producto estadístico objetivo.

Es por ello por lo que decidimos explorar la idea de utilizar la base de datos de máxima desagregación. Estas series no presentan la problemática de dato ausente expuesta en las series mensuales y anuales, quizá por tener un menor grado de manipulación que el resto.

- **Series Mensuales Máxima Desagregación.**

El uso de esta serie supone un incremento considerable de la acuracidad del dato obtenido a la hora de realizar los agregados, puesto que los datos se encuentran en el punto de máxima desagregación. Mientras que en la Series Históricas Anuales y en las Series Resúmenes Mensuales se ha realizado agregaciones por ciertas variables, en las Series Mensuales Máxima Desagregación los datos que se presentan serían de las transacciones según como han acontecido en el mes, de acuerdo con la información proporcionada por la AEAT. La serie ha sufrido tratamiento con respecto al dato bruto obtenido de los registros administrativos, pero no de agregación sino transformación de algunas variables (como la conversión a euros de ciertas transacciones llevadas a cabo en otra moneda) o de eliminación de observaciones (como las transacciones tránsito que únicamente atravesarían nuestro país para terminar en otro lugar).

Para que el lector se pueda hacer una idea de lo que ha supuesto el cambio en el tratamiento de una base de datos a otra podemos hacer una comparativa de tamaños para el año 2020. La Serie Resumen Mensual para este año presentaba una cantidad de 2.328.284 filas para un total de 10 variables. Mientras que la Serie Mensual Máxima Desagregación para el mismo año tendría una cantidad de 67.825.765 filas para un total de 30 variables.

Gracias a la realización de diversas consultas con la AEAT y con ayuda de la metodología hemos podido tener conocimiento de gran variedad de las variables que entendemos de interés en esta serie y parte de las cuales exponemos en la tabla 3.1.

Tabla 3.1. Resumen variables Series Mensuales Máxima Desagregación. Códigos de variables incluidos en A3.	
FLUJO	Identificación del desplazamiento de bienes o servicios (Importación o Exportación).
AÑO	Año natural de la admisión por el Departamento de Aduanas e Impuestos Especiales del correspondiente Documento único administrativo (DUA) referido a la exportación o importación de la mercancía, si se trata de comercio con terceros países; y el año natural de la expedición o de introducción de las mercancías desde o al territorio estadístico español, en el caso de operaciones intracomunitarias.
MES	Mes natural de la admisión por el Departamento de Aduanas e Impuestos Especiales del correspondiente Documento único administrativo (DUA) referido a la exportación o importación de la mercancía, si se trata de comercio con terceros países; y el mes natural de la expedición o de introducción de las mercancías desde o al territorio estadístico español, en el caso de operaciones intracomunitarias.
PROVINCIA ADUANA RECINTO	Identificación de la provincia donde se sitúa la aduana de frontera, receptora de la mercancía por primera vez en territorio español en caso de importación o emisora de la mercancía al extranjero en caso de exportación.
POSICION ESTADISTICA	Identificación de Códigos NC/TARIC. Códigos NC (8 dígitos). Códigos TARIC (10 dígitos).
TIPO DECLARACION	Identificación del tipo de declaración presentada: Intrastat o DUA
PAIS ORIGEN DESTINO	Se trataría, en el caso de las Importaciones, del último país donde se ha producido la última modificación significativa en el producto, independientemente de donde haya sido enviado por última vez. Y, en el caso de las Exportaciones, se trataría del país de destino en la transacción.
PAIS PROCEDENCIA EXPEDICION	Solo tiene sentido en Importaciones. Ultimo país desde donde se manda la mercancía independientemente de la última transformación.
PROVINCIA ORIGEN DESTINO	Indica, en el caso de las introducciones, la provincia española de destino, es decir, aquélla a la que finalmente llegan las mercancías, bien para su consumo,

	bien para ser objeto de operaciones de transformación. En el caso de las expediciones, la procedencia, es decir, aquella provincia española en la que las mercancías hayan sido obtenidas, elaboradas, o hayan sido objeto de operaciones de montaje, ensamblaje, transformación, reparación o mantenimiento, o en su defecto aquella en que haya tenido lugar el proceso de comercialización, o aquella desde la que se hayan expedido las mercancías.
REGIMEN ADUANERO SOLICITADO	Regímenes aduaneros pertinentes que aplica el solicitante.
REGIMEN ADUANERO PRECEDENTE	Régimen aduanero pertinente que precedía la solicitud final.
PESO	Peso en gramos de la mercancía transportada. Masa propia de la mercancía desprovista de todos sus envases y embalajes
UNIDADES	Número de unidades estadísticas, para todos aquellos códigos de la Nomenclatura Combinada que exigen esta información, de acuerdo con las unidades de medida que refiere, para cada código.
VALOR ESTADISTICO	El valor que tendría la mercancía en el momento de entrar, si se trata de introducciones, o salir, si se trata de expediciones, en/del territorio estadístico español, deduciendo únicamente los impuestos que graven el consumo, pero no otros e imputando asimismo la parte proporcional de los gastos de transporte y seguro del trayecto realizado hasta el punto en que las mercancías entran o salen del territorio español.
MODO TRANSPORTE FRONTERA	Modo en que la mercancía ha sido transportada hasta la frontera.
MODO TRANSPORTE INTERIOR	Modo en que la mercancía ha sido transportada una vez pasada la frontera.
ZONA INTERCAMBIO UE_3os	Identificación de la relación entre las partes que intercambian: Si el intercambio es con un país perteneciente a la Unión Europea (C) o si en cambio es con un tercero (T).
NATURALEZA TRANSACCION	Identificación de la Naturaleza de la Transacción realizada.
CONDICIONES ENTREGA	Identificación de las condiciones de entrega de la mercancía.

4. Marco Teórico.

Haciendo una revisión del principal documento de referencia para las cuentas de Importaciones y Exportaciones físicas de materiales (Eurostat, 2018: *Economy-wide material flow accounts*) vemos que obtenemos ciertas referencias a distintas fuentes de datos que nos permiten la realización del primer input (como son las [European statistics on international trade in goods](#) o las bases de datos de las estadísticas nacionales de comercio exterior) o ciertas transacciones que no deberían ser recogidas en la cuenta (como son las transacciones denominadas “Transit” o las “Quasi-Transit”). Pero desde la guía no se realiza exposición alguna, o recomendación, de la forma en la que se debería tratar la base de datos una vez localizadas las ciertas incongruencias que puedan haberse cometido con los registros administrativos.

Por ello, en este apartado trataremos de recopilar diferentes procedimientos que se puedan desempeñar para la depuración de las bases de datos que supongan una mejora en la calidad en cuanto a la elaboración de las estadísticas de Importaciones y Exportaciones físicas. Realizando, en primer lugar, una revisión bibliográfica en búsqueda de procedimientos o métodos que se hayan desempeñado en otros Institutos Nacionales para este fin.

Experiencia en depuración y características adyacentes a las bases de datos de comercio exterior.

A pesar de que el trabajo con registros administrativos está a la orden del día en el entorno de la estadística oficial, se habla también de que aún no se habría establecido una teoría sobre su utilización. O al menos, no tan extendida como la teoría en torno al muestreo y a las estadísticas generadas a través de encuestas, en donde las teorías sobre la probabilidad y la inferencia estadística han sido desarrolladas durante décadas en las universidades y en los institutos de estadística.

Puesto que desde la literatura (Wallgren & Wallgren, 2014) se nos trasmite que una gran calidad en los sistemas administrativos es una condición más que necesaria, pero que no es la única para poder satisfacer unas estadísticas de calidad. Nos vemos en la necesidad de indagar que otros requerimientos son necesarios para finalmente poder obtener un producto de calidad cuya fuente hayan sido los registros.

Igualmente, en la documentación se habla de la importancia de las bases legales que legislan los sistemas nacionales de administración. Puesto que los datos que van a ser obtenidos en los registros van a depender de manera directa de la norma que recoja la legislación. Esto iría de

alguna manera ligado también a la finalidad que presentan los registros y si se adaptan a la finalidad estadística que nosotros tenemos.

Siguiendo esta línea, vemos que las estadísticas de comercio exterior son elaboradas por la AEAT en base a dos estadísticas diferentes por lo que podemos ver en la [metodología](#). Una sería la Estadística de Comercio Extracomunitario ([Extrastat](#)), mientras que la otra sería la Estadística de Intercambio de Bienes entre Estados de la Unión Europea ([Intrastat](#)). Ambas estadísticas tienen finalidades que apuntan al mismo objetivo, como es la recogida de información estadística, por un lado, de las importaciones y exportaciones y por otro de las llegadas y expediciones¹⁴.

Así las bases legales que regulan tanto Intrastat, como Extrastat son los Reglamentos [N.º 638/2004 de 31 de marzo de 2004](#), modificado por el [N.º 222/2009 de 11 de marzo de 2009](#) y el [N.º 471/2009 de 60 de mayo de 2009](#), de manera respectiva. En ellos se observa que la principal fuente de información para abastecer las estadísticas va a estar en la necesidad de elaboración del Documento Único Administrativo (DUA) que presentan los individuos que se encuentran involucrados en la entrada o salida de mercancías en un territorio. Este tendría una finalidad fiscal y aduanera sirviendo de base, entre otras muchas funciones, para la declaración tributaria. Aquellos, que, por su condición, no sean objeto de DUA en Intrastat, aportarán la información estadística necesaria a través de un medio diferente.

En los documentos legales expuestos en el párrafo anterior se puede observar de manera más concreta la información estadística necesaria que se debe aportar, a través de los medios facilitados, a las autoridades europeas. Se aprecia que tanto en Intrastat, como en Extrastat se requiere la recogida de información de variables clave para nuestros intereses en la cuenta de flujos. Nos permite, entre otras muchas cosas, la identificación de los productos (a cierto nivel de exactitud) mediante los códigos de la Nomenclatura Combinada (NC), la identificación del país de origen de la mercancía y el país de destino o la contabilización de la cantidad física de la mercancía involucrada en la transacción, algo esencial para poder desarrollar nuestro objetivo en las CFM-TE.

La principal circunstancia que se nos plantea como una problemática vendría a la hora de estudiar los principales objetivos de las estadísticas Intrastat y Extrastat. En el caso de Intrastat encontramos referencia a establecer “*un marco común para la elaboración sistemática de estadísticas comunitarias de intercambios de bienes entre Estados miembros*” (R. CE N.º 638, 2004), dejando bastante margen en cuanto a qué estadísticas se elaborarían a raíz de los

¹⁴ Al estar unificadas ambas estadísticas en las bases de datos producidas, en el resto del trabajo hablaremos de importaciones y exportaciones a todas las transacciones que tienen lugar con países extranjeros de manera independiente si se trata de países comunitarios o extracomunitarios.

registros. Pero con el caso de Extrastat veríamos que las estadísticas en cuestión van muy enfocadas a cubrir informaciones de índole económica como la “*participación del euro en el comercio internacional de mercancías*” o la “*elaboración de la balanza de pagos y las cuentas nacionales*” (R. CE N.º 471, 2009).

Todo lo expuesto anteriormente hace sospechar que la información recopilada por los registros es tratada con una finalidad en línea con las estadísticas con un sentido más monetario. Además, encontramos evidencia de este hecho en otro país, como es el caso de Países Bajos. El Instituto Nacional de Estadística Neerlandés (Centraal Bureau voor de Statistiek, CBS) decidió no usar de forma directa los datos que describían la información física recogida en las estadísticas de comercio exterior por varios motivos. Se decidió estimar para poder estar en línea con las estadísticas monetarias elaboradas en Cuentas Nacionales y debido a que la calidad de la información física no era la apropiada (Delahaye & Nootenboom, 2008) (Delahaye & Nootenboom, 2009). En España, nos transmiten de forma similar este hecho a través de una consulta realizada a la AEAT, que se encarga de los procedimientos Intrastat y Extrastat en nuestro país. La información que recibe una depuración y un tratamiento y presenta una mayor calidad es la relativa a los datos monetarios de las transacciones en comercio exterior, pasando un poco por alto la información en materia física.

El Instituto Nacional de Estadística Estonio (Statistikaamet) nos narra otra circunstancia en la cual se deben de realizar las estimaciones de cantidades físicas para ciertas categorías. Allí, las estadísticas de comercio exterior, Intrastat y Extrastat son elaboradas por organismos públicos distintos. Las correspondientes al comercio comunitario serían elaboradas por el Statistikaamet, mientras que las de comercio extracomunitario serían elaboradas por la Oficina de Impuestos y Aduanas. Es en la parte Intrastat donde presentan ciertos valores ausentes (empresas con poca actividad comercial exterior no declaran información, empresas directamente no responden o responden tarde). Las estimaciones de las cantidades físicas son realizadas a través de las estimaciones de datos monetarios efectuadas en los datos ausentes. Se estima en proporcionalidad, pero como también nos transmiten, hay cierta inexactitud puesto que se realizan estimaciones para las empresas pequeñas con los datos aportados por grandes empresas con suficientes transacciones que rebasan el umbral de exención (Oras et al., 2010).

Y es que la normativa Intrastat posee esta peculiaridad para todos los países. Quedan excluidos de aportar información por las transacciones comunitarias realizadas, aquellos responsables cuyos valores anuales de comercio intracomunitario se queden por debajo de un umbral (R. CE N.º 638, 2004). En España el umbral de exención se sitúa en la cantidad de 400.000 euros anuales actualmente, tanto para las importaciones como para las exportaciones. Teniendo que

empezar a declarar el implicado en el mes que lo sobrepase, esto es, las operaciones que hayan tenido lugar en ese mes y los siguientes (Orden HFP/1480/2021, de 28 de diciembre).

Por último, vemos el caso del Instituto Nacional de Estadística Sueco (Statistiska centralbyrån, SCB), el cual nos traslada la complejidad que supuso el reportar los datos a un nivel de desagregación de 8 dígitos en la NC en las estadísticas de comercio exterior, por lo que decidieron, para poder mantener unas estadísticas de calidad, reducirlo a 6 dígitos (Carlsson et al., 2006). Los expertos en materia de CFM-TE también nos exponen otra problemática que presenta este sistema de códigos de la NC y es que es un sistema que cambia año tras año, dificultando el seguimiento de ciertos bienes en el tiempo. Es decir, un producto que tenía asociado un código para un cierto año podría tener asociado otro en el año siguiente por la propia naturaleza del sistema. Al tratarse de un sistema que tiene como objetivo establecer el tipo de derecho de aduana que se aplica a un producto en concreto y el tratamiento estadístico que va a tener la carga, todo ello a nivel europeo, va a permitir su comparabilidad entre los diferentes estados miembro.

En el INE, el principal elemento a tener en cuenta sobre este apartado de la CFM-TE, y que pretendemos mejorar a través de este proyecto, sería la adecuación de una nueva base de datos para la elaboración de las estadísticas de Importaciones y Exportaciones físicas. Como hemos comentado anteriormente, los datos proporcionados para los años de referencia de 2008 a 2020 han sido obtenidos a partir de datos anuales de comercio exterior (o mensuales en el caso de los avances). Estas bases de datos presentan información agregada según los códigos de NC de las importaciones y exportaciones que han tenido lugar entre España y el resto del mundo. La depuración para el control de incongruencias en estas bases de datos ha sido realizada hasta ahora a través de depuración manual por expertos en la materia. En el procedimiento se supervisaba la evolución del precio por kilogramo de las transacciones agrupadas según los códigos de producto y el país de destino u origen. Se revisaba que el precio por kilogramo no pasara un umbral establecido y si este sobrepasaba y el error era obvio se corregía la incongruencia del dato de manera manual.

Mediante el uso de la nueva base de datos de máxima desagregación se pretende tener un mayor control sobre los agregados, pudiendo tratar de identificar transacciones que por la naturaleza de la CFM-TE no deben ser incluidas y que sospechamos pueden estar siendo incorporadas en las bases anuales y mensuales. Adicionalmente, al contar con un nivel tan desagregado en la cuenta, vamos a poder llevar a cabo una mejor depuración de datos al presentar una información más abundante.

En consecuencia, teniendo en cuenta todo lo expuesto hasta ahora, entendemos que las principales problemáticas que se va a presentar en la compilación de la estadística de

importaciones y exportaciones físicas en las CFM-TE se centran en lo desatendida que puede llegar a estar la información física de las transacciones internacionales y lo complicado que va a resultar el trabajar con unas bases de datos de un tamaño descomunal y con las que no se ha trabajado hasta ahora. El esfuerzo va a recaer en la automatización de la lectura de archivos y su posterior depuración con la información monetaria, la cual presenta una mayor calidad y ha tenido un mejor tratamiento. Todo ello combinado con consultas periódicas con la entidad responsable de la producción de las estadísticas de comercio exterior, la AEAT.

Depuración de datos: Edits, outliers e imputaciones.

Una de las cosas que hemos podido advertir a la hora de investigar sobre los diferentes procedimientos que se pueden desempeñar en la depuración de datos ha sido que la bibliografía al respecto está muy extendida para los casos de las encuestas. Indagando y desarrollándose en esta cuestión, atendiendo de cerca los casos en las encuestas sociales y las económicas e incluso tratando por encima el caso de los censos. Algo que no ocurre, o no de manera tan extendida, en el caso de trabajar con estadísticas que tienen su origen en los registros administrativos, puesto que se trata de una materia un tanto novedosa y debido a que presentan ciertas características considerablemente alejadas de las encuestas (De Waal et al., 2011).

En uno de los pocos textos en los que se trata este asunto, (Wallgren & Wallgren, 2014), se nos habla que existen dos principales dimensiones en el tratamiento a realizar. Uno sería el análisis de duplicados y otro sería la revisión de los valores de las variables.

El tema del análisis de duplicados es, sin lugar a duda, absolutamente necesario por la propia naturaleza que tienen las estadísticas producidas a partir de la recopilación y combinación de datos de registros administrativos. Un ejemplo de esto lo encontraríamos en las propias estadísticas de comercio exterior las cuales centran el interés de este trabajo. En ellas, la parte de información comunitaria es recopilada a través de formulario Intrastat o bien a través del DUA. Por lo que podríamos encontrarnos que algún agente implicado haya puesto en práctica ambas vías para aportar la misma información. Por tanto, se precisará de este proceso en la parte de depuración del trabajo a realizar.

Para el análisis de las variables, ponen el foco a la importancia de un contacto regular con la oficina responsable del tratamiento de los Registros, puesto que los responsables del tratamiento de los mismos datos, pero con fines estadísticos en el INE, deben de conocer las complejas normas adyacentes a los fines administrativos que tenga el registro en cuestión. Algo que es preciso debido a que los datos obtenidos en los Registros Administrativos se

consiguen antes del desarrollo de unos metadatos con finalidad estadística, “*la información detrás de los datos*”. Aun así, los Registros presentan unos metadatos que pueden estar en línea con las definiciones encontradas en la estadística pública (Salgado & Oancea, 2020).

Adicionalmente a las recomendaciones expuestas en la literatura sobre los registros, consideramos que los criterios de depuración de datos para variables que provengan de encuestas pueden ser de utilidad en este estudio. En consecuencia, examinaremos de igual manera los diferentes planteamientos indicados para la depuración de datos en encuestas para su utilización en este proyecto, en la medida en que esto sea posible teniendo en cuenta las posibles limitaciones.

Observamos que se destacan diferentes métodos a la hora de efectuar la depuración de datos, todo ello en relación con cómo interactúan en el proceso diversos factores como son el grado de participación o el equilibrio entre la experta en la materia y la computación, o si el estudio se realiza de una parte de los datos (más influyentes) o del conjunto completo (De Waal et al., 2011).

Identificamos que los posibles métodos más orientados a este proyecto van a ser, por un lado, la depuración interactiva de datos y, por otro lado, la depuración automática. En la primera se utilizan los grandes conocimientos de la experta en la materia a lo largo del proceso de depuración, presentando por tanto una gran calidad al final del proceso. Pero, por otro lado, al tratarse de un gran volumen de datos, se debe de tentar la mayor automatización del proceso en aquellos apartados donde sea posible. Puesto que con tal cantidad de información la intervención del experto en ciertos aspectos va a ser considerablemente compleja. Se va a perder en flexibilidad en cuanto a la interacción con la experta, pero se ganaría en la productibilidad a la hora de procesar ciertas operaciones (De Waal et al., 2011).

Una ventaja, a simple vista, que pueden presentar las encuestas con respecto a los registros administrativos es la supervisión en la etapa en la que se están recogiendo los datos. Esto ayuda a prevenir el traspaso de errores entre fases, los errores serían solventados en la recogida, evitando su llegada a la fase de procesamiento. Pero los registros administrativos también ayudan a prevenir, en su particular forma, que ciertos errores se traspasen entre fases del proceso estadístico. Nuevamente encontramos un ejemplo de este suceso en los registros de origen de las estadísticas que tiene interés este trabajo. Los registros de aduana tienen su interés en la compilación de las transacciones realizadas y en la posible carga impositiva que pueda tener la operación. De no manifestarse con la veracidad necesaria, el error podría acarrear sanciones económicas para el declarante, lo que asegura en cierto modo que los datos no contengan apenas errores en lo que respecta a lo monetario (Wallgren & Wallgren, 2014).

La trascendencia de la depuración de datos va a recaer en una buena localización de las incoherencias y en la posterior decisión que se tome sobre estas. Esto es, si eliminar las observaciones o realizar una imputación de los valores detectados como contradictorios. Para la detección de los datos incongruentes vemos que podremos proceder de diferentes formas.

Una de ellas sería establecer una serie de normas o restricciones que se deben de cumplir en las cuantías que presentan las variables. Si algunas de las observaciones violaran estas normas, por tratarse de un valor ausente o simplemente incongruente, se marcaría la observación para, posteriormente, tomar una decisión sobre el valor contradictorio. En este caso estaríamos hablando de establecer unas normas sobre los edits (Luzi et al., 2007).

Otra manera de detectar posibles incongruencias que se están produciendo sería observando aquellos valores que sobresalgan del conjunto total de observaciones, estos valores son considerados como atípicos (outliers) y de la misma manera que en el caso anterior necesitarían de una atención especial puesto que podrán distorsionar el análisis que se desea desempeñar.

Una definición más completa de outlier sería la de aquella observación cuyo valor se “*desvía lo suficiente del resto como para levantar sospechas de que su valor haya sido generado por un mecanismo diferente*” (Hawkins, 1980). A la hora de elegir que técnica se deben utilizar para la localización de estos valores, se debe tener en cuenta si el análisis va a ser univariante o multivariante.

Para el caso de un análisis univariante de outliers, como es el caso de este proyecto, encontramos en la bibliografía diferentes métodos como los siguientes:

- Valores que sobresalgan el rango de la **media más o menos tres desviaciones típicas**. Este método tiene su base en la propia naturaleza que se puede observar en la distribución normal, donde prácticamente la totalidad de los datos se encuentran recogidos en ese rango. Algunos de los problemas que presenta la ejecución de este método sería que, por ejemplo, el análisis se ve considerablemente afectado por la presencia de los propios outliers, ya que estos incidirían tanto en los parámetros de la media como en la desviación típica (Leys et al., 2013).
- Una alternativa robusta sería el uso de la **Desviación Mediana Absoluta (median absolute deviation, MAD)**. Esta medida ofrece una ventaja en cuanto a robustez en su uso, debido a que en este caso la mediana no se vería afectada por la presencia de outliers.

$$MAD = b M_i(|x_i - M_j(x_j)|)$$

Donde la x_i son las n observaciones presentes en el estudio, M_i es la mediana de la serie y b sería una constante la cual está unida al principio de normalidad que seguirían los datos, por norma general $b = 1,4826$.

Para la fijación de unos umbrales, al igual que ocurría en el caso de la media y desviación típica, se tiene que establecer un valor para asignar al MAD según la rigurosidad que se quiera seguir. Siendo más rigurosos, como en el caso anterior, los valores identificados como outliers serían los que se salen del intervalo $[M - 3 MAD, M + 3 MAD]$ (Leys et al., 2013).

- Otra alternativa, basada en el diagrama de caja de Tukey, es la de señalar esas observaciones que sobresalgan de los umbrales generados mediante el **Rango Intercuartílico (RIQ)**. El RIQ estaría formado por el primer cuartil $Q_1 = x_{[n/4]}$ y el tercer cuartil $Q_3 = x_{[3n/4]}$ de la forma que $RIQ = Q_3 - Q_1$. Del mismo como que ocurría anteriormente se deben establecer unos umbrales en función del rigor con la que se quiera hacer el análisis. Replicando la exhaustividad de los casos anteriores, marcaríamos como atípicos los valores que no estén incluidos en $[Q_1 - 3 RIQ, Q_3 + 3 RIQ]$ (Rousseeuw & Hubert, 2011).

Una vez elegida la manera de marcar aquellas observaciones que sobresalen los patrones establecidos, habría que decidir qué hacer con ellas. Se podría optar por no incluirlas, pero teniendo la opción de contar con variables auxiliares, entendemos que se podría conseguir una buena depuración de los atípicos imputando su valor para poder continuar con el mismo número de datos.

Para las imputaciones se observan diferentes técnicas para encontrar un valor que sustituya al valor detectado como incongruente o en su defecto el valor que se encuentre ausente. Una vez que no podemos contactar con la persona nuevamente, como es nuestro caso, para tratar de corregir la incongruencia, nos quedan una serie de técnicas para realizar la imputación.

Por un lado, tendríamos técnicas que se podrían utilizar siempre y cuando no se disponga de información auxiliar como es la técnica de hot-deck aleatorio (el donante pertenece a los datos actuales) o la de cold-deck aleatorio (si el donante pertenece a otra base de datos) (Luzi et al., 2007). Por otro lado, contando con información de variables auxiliares, nos centraremos en aquellas técnicas que nos permitan sacar provecho de ello. No siempre es posible tener acceso a información auxiliar para poder realizar imputaciones. En nuestro caso sí que disponemos de esta información (variable monetaria) y, además, como hemos expuesto anteriormente, refleja de forma más confiable la información sobre las transacciones al estar sujetos a posibles sanciones.

Una técnica que utiliza unidades similares, como la expuesta en el apartado anterior, pero que también contaría con la ayuda de una variable auxiliar sería la de imputar por vecino más cercano. Con esta técnica se obtendrá el dato imputado gracias a una unidad la cual minimiza la distancia según la variable auxiliar (De Waal et al., 2011).

Otras técnicas, basadas en modelos, serían las imputaciones por regresión y sus variantes a la hora de imputar por razón o imputar por la media. La imputación por regresión sería la más general y a partir de la cual se pueden observar los casos particulares de imputación por razón o por media. Los valores imputados se obtendrían a partir de

$$\hat{y}_i = \hat{\beta}_0 + x_1\hat{\beta}_1 + \dots + x_k\hat{\beta}_k$$

Donde x_1 a x_k serían las variables auxiliares y $\hat{\beta}_0$ a $\hat{\beta}_k$ serían las estimaciones por mínimos cuadrados (Luzi et al., 2007). La imputación por razón sería un caso especial de lo expuesto anteriormente con la particularidad de que solamente tendríamos una variable auxiliar. Las imputaciones de este tipo son incesantes si realmente existiera una relación proporcional lineal entre la variable auxiliar y la objetivo.

5. Metodología y Resultados.

Metodología.

Como se expuso en la sección 3 de este trabajo, los datos con los que se van a trabajar son las [estadísticas de comercio exterior](#) producidas a partir de registros por la AEAT. En concreto las bases de datos de las series mensuales de máxima desagregación.

Para poder producir la serie completa de la cuenta bajo reglamento necesitamos los datos al menos desde el año 2008. Desde la página web de la AEAT, para esta base de datos, únicamente se disponen de datos a partir del año 2016. Es por ello que, en vistas de la organización de la reunión anual de seguimiento del Convenio de Colaboración entre la AEAT y el INE en materia estadística, se decide trasladar la necesidad de contar con el resto de los años, esto es datos desde 2008 hasta 2015.

Para la realización de este proyecto no ha sido posible disponer de datos que no se encontraban en la página web. Aun así, conocemos de la disponibilidad de estos archivos puesto que se ha trabajado con las series anuales para la elaboración de la cuenta para los años 2008 a 2020 y el reglamento Intrastat entró en vigor a partir del año 2004. Por todo ello, se va a realizar el estudio para este proyecto con los datos a partir de 2016 y una vez que se dispongan de los datos desde 2008 se replicará la metodología empleada para la obtención de la serie completa.

Una problemática inicial que nos encontramos con las series mensuales de máxima desagregación ha sido que, debido a la desmesurada cantidad de información que viene recogida, se presentó como un reto computacional para su lectura y posterior procesado. Para que el lector pueda hacerse una idea, exponemos un resumen en la tabla 5.1 para los años disponibles en la página web de la AEAT. Cada año presenta la información en 120 archivos divididos en 10 archivos por cada mes y con un total de 30 variables.

Año	Tamaño totalidad archivos en .txt (en gigabytes)	Número de transacciones recogidas en el año
2016	9,86	46.480.087
2017	10,7	50.560.884
2018	14,0	66.348.478
2019	15,8	74.528.543
2020	14,4	67.825.717
2021	19,0	89.768.918

Seguidamente, nos encontramos con la problemática de la comprensión de las diversas operaciones que puedan estar desarrollándose en las estadísticas de comercio exterior con los registros administrativos en la AEAT. Ante la falta de unos metadatos con información exhaustiva sobre ello, la primera fase del proyecto consistió en el entendimiento de las variables presentes en las series mensuales de máxima desagregación.

A parte del estudio en profundidad de los reglamentos relativos a los registros a través de los que se obtienen los datos con finalidad administrativa, entendimos la necesidad de establecer el contacto con los responsables de tratamiento en la AEAT para poder realizar las consultas oportunas y necesarias, tal y como se refleja en la teoría (Wallgren & Wallgren, 2014). Lo más importante en esta parte preliminar del proyecto consistía en el entendimiento de la mayoría de las variables presentes en las bases de datos. Todo ello para poder cerciorarnos de que la nueva fuente de información nos iba a permitir cubrir las necesidades oportunas para el objetivo de mejora en la cuenta.

La consulta sobre las variables tenía el objetivo principal de localizar aquellas transacciones que tienen un cuasi-tránsito en España. Algunas de las variables consultadas fueron la de NATURALEZA TRANSACCIÓN, PAÍS ORIGEN/DESTINO, PAÍS PROCEDENCIA/EXPEDICIÓN o RÉGIMEN ADUANERO. Gracias a esta primera consulta nos fue posible completar la información recogida en la tabla 3.1 sobre gran parte de la base de datos.

Existían dos situaciones en las cuales se daban las transacciones cuasi-tránsito. En primer lugar, cuando un país comunitario hace el despacho para la exportación a un país no comunitario. En segundo lugar, aquellos productos que entran en España procedentes de un país extracomunitario y una vez despachados en España son enviados a otro estado miembro.

El experto en materia de la AEAT nos traslada la complejidad del asunto. Sobre todo, el del primer caso, puesto que primeramente un país comunitario realiza la declaración Intrastat en España y posteriormente el DUA de exportación. Por tanto, existe una importación UE y luego una salida a un país tercero registradas en la estadística de tal manera que no se encuentran conectadas, no se sabe si la mercancía se ha quedado en España o si se ha exportado a un país tercero.

Por el contrario, para el segundo caso, sí es posible la localización de ciertas transacciones cuasi-tránsito gracias a la variable de REGIMEN ADUANERO SOLICITADO, a través del código 42. Literalmente este código nos transmite que se trataría de un “*despacho a consumo con despacho a libre práctica simultáneo de mercancías objeto de entrega exenta del IVA a otro Estado Miembro*”. Por todo ello, y tras confirmación por parte del experto, procederemos

a eliminar del análisis todas aquellas Importaciones que vienen identificadas con este régimen número 42.

Otras observaciones que son identificadas en la base de datos y las cuales son susceptibles de eliminar fueron las que vienen identificadas con el REGIMEN ADUANERO SOLICITADO de 80, esto es, “*despacho a consumo (tributos REF) de mercancía comunitaria previamente vinculada a un depósito REF*”. Puesto que entendemos que esta transacción hace referencia a una transacción que tuvo lugar en el pasado y que tiene su fin en depósito, decidimos eliminar estas observaciones también para evitar duplicidad.

Una vez aclaradas las circunstancias expuestas sobre algunas de las variables, nos dispusimos a realizar una lectura completa y un tratamiento de las bases de datos para su posterior procesamiento. Este trabajo del que hablamos ha sido realizado mediante el uso del software estadístico libre R, en el cual hemos utilizado los paquetes expuestos en la tabla 5.2.

Tabla 5.2. Resumen de paquetes utilizados en programa R.	
Paquete	Objetivo
data.table (Dowle et al., 2019)	Administración de los datos de gran tamaño de los que disponemos (uniones, agregaciones, tratamiento de variables...)
fastReadfwf (Salgado et al., 2018)	Lectura rápida de archivos que presentan un ancho fijo mediante la utilización de un esquema de entrada.
validate (van der Loo et al., 2022)	Detección de ciertas incongruencias en la base de datos mediante la confrontación de los datos frente a unas reglas de validación preestablecidas por consenso con experto en la materia.
openxlsx (Schauberger et al., 2021) readxl (Wickham et al., 2022)	Importación y creación de archivos Excel en la interacción con R. El nexo que existe entre Eurostat y el INE en el intercambio de información es realizado mediante archivos Excel: En formato .xlsx se encuentran los anexos proporcionados por la institución europea para la clasificación de materiales y también la comunicación de datos por parte del INE en las recopilaciones anuales.

En primer lugar, antes de la elaboración del código para la lectura de los archivos .txt de la AEAT, se procedió a la elaboración del esquema necesario para la rápida lectura con el paquete fastReadfwf.: *comext_Schema.xlsx*. Este ha sido elaborado a partir de los diseños de registro aportados por la AEAT. En el esquema especificamos el nombre de las variables

analizadas, la anchura, posición inicial, posición final en el archivo de texto, el tipo de variable que se trata (carácter o numérica), una breve descripción y una expresión regular para cada una de ellas. Esto último nos va a permitir detectar si en alguna de las bases de datos analizadas se estuviera introduciendo algún dato o carácter que no tuviéramos recogido previamente. Detectando así posibles incoherencias o valores para los cuales no tuviéramos un conocimiento previo.

A la par también elaboramos un archivo .txt en el que recogemos una serie de reglas de validación combinando un conjunto de variables para detectar posibles contradicciones: *edits_general.txt*.

Tanto los *edits_general.txt* como un resumen del *comext_Schema.xlsx* pueden ser observados en el anexo A4. Del *comext_Schema.xlsx* añadimos únicamente para aquellas variables utilizadas a lo largo del proyecto y para aquellas que consideramos que pueden tener utilidad en el futuro.

Una vez elaboramos los archivos expuestos anteriormente, y junto con los .txt de las series mensuales de máxima desagregación, nos preparamos para realizar la lectura completa de los archivos con el objetivo de almacenarlos en un formato adecuado al programa utilizado para, posteriormente, realizar el procesamiento de datos necesario. Exponemos de manera resumida en las tablas 5.3 y 5.4 los pasos llevados a cabo en el programa R para la obtención de las tablas finales de la cuenta de importaciones y exportaciones físicas de materiales.

Tabla 5.3. Resumen de pasos en la lectura de archivos.

En primer lugar, realizamos una primera lectura de los datos para proceder a realizar un análisis de datos duplicados. Puesto que la base de datos contiene una elevada cantidad de información, vamos a aprovechar gran cantidad de variables para hacer lo más específicas posibles las transacciones y así eliminar aquellas que se repitan. Para ello, seleccionamos de variables en la base de datos las siguientes:

"FLUJO", "ANO", "MES", "PROVINCIA_ADUANA_RECINTO", "FECHA_ADMISION_DOCUMENTO", "POSICION_ESTADISTICA", "PAIS_ORIGEN_DESTINO", "PAIS_PROCEDENCIA_EXPEDICION", "PROVINCIA_ORIGEN_DESTINO", "REGIMEN_ADUANERO_SOLICITADO", "REGIMEN_ADUANERO_PRECEDENTE", "PESO", "UNIDADES", "VALOR_ESTADISTICO", "MODO_TRANSPORTE_FRONTERA", "MODO_TRANSPORTE_INTERIOR", "NATURALEZA_TRANSACCION", "CONDICIONES_ENTREGA".

El número de observaciones duplicadas encontradas, y eliminadas, en cada año se muestran a continuación:

2016	6.044.125
2017	6.678.846
2018	9.368.713
2019	11.178.961
2020	11.099.359
2021	16.451.661

Una vez concluido el procedimiento con los valores duplicados, pasamos a observar si algunas de las observaciones incumplen con los edits establecidos. De ser así analizamos cuáles son esas observaciones y procederíamos a tomar una decisión sobre las mismas.

En los años analizados quedan señalizadas las observaciones que no deberían estar, que son las de REGIMEN_ADUANERO_SOLICITADO de códigos 80 y 42, las cuales eliminamos antes de guardar el archivo final.

El número de observaciones identificadas, y eliminadas, de cada código son las siguientes en cada año:

	“80”	“42”
2016	6.697	73.607
2017	5.221	68.301
2018	3.407	66.997
2019	2.704	63.517
2020	3.002	78.415
2021	4.586	79.957

Realizamos una validación de las variables para comprobar que no se esté incluyendo algún código del cual no tenemos constancia o por error de manera individual en cada una de las variables utilizadas para el análisis. Los códigos no detectados previamente, pero que realmente deben permanecer al conjunto son añadidos en el *comext_Schema.xlsx* antes de proseguir con la lectura.

Esta circunstancia expuesta se ha dado sobre todo para códigos de países que previamente no se encontraban en los textos de referencia.

Una vez el conjunto de datos ha sido validado por los edits y para cada una de las variables, se procede a realizar una serie de acciones pertinentes para su posterior procesado:

- La variable monetaria VALOR_ESTADISTICO es transformada en VALOR_EUR para disponer el dato en euros.
- El PESO es transformado en PESO_KG, pasando la variable física de gramos a kilogramos.

- Obtenemos el valor por kilogramo mediante la creación de la variable EUR_KG. Esta es calculada dividiendo la variable VALOR_EUR entre PESO_KG.
- De la variable POSICION_ESTADISTICA únicamente nos quedamos con los primeros 8 dígitos puesto que así creamos la variable CN_CODE correspondiente a los códigos de la nomenclatura combinada presentes en la base de datos.
- Finalmente, hacemos un breve estudio de los datos ausentes presentes en cada año. Obteniendo que los datos que presentan un valor de 0 en la variable física de PESO_KG (dato ausente en nuestra investigación) se corresponden en gran medida con el código de NC de 27160000 correspondiente a “Energía eléctrica”, la cual, por su condición, no va a suponer transacciones con movimientos de material físico. El resto de los valores ausentes encontrados son residuales y los exponemos a continuación indicando el año analizado en la primera columna:

2016	1 transacción con peso ausente para el código 71039900. Este código recogería aquellos productos de “ <i>piedras preciosas y semipreciosas, trabajadas, pero sin ensartar, montar o engarzar...</i> ”. Se corresponde al código de flujos de materiales 3.9: otros minerales no metálicos. Esta categoría incluirá productos muy dispares por lo que omitimos la imputación del dato.
2017	0 datos ausentes en la variable PESO_KG.
2018	0 datos ausentes en la variable PESO_KG.
2019	1 transacción con peso ausente con código 30021900. Este código recogería productos denominados como “ <i>productos inmunológicos</i> ”. Presente en el código de flujos de materiales 5: otros productos. Nuevamente, no imputaremos por ser una categoría muy genérica y por tratarse de algo no relevante.
2020	0 datos ausentes en la variable PESO_KG.
2021	3 transacciones con peso ausente con código 99900000. Este código se corresponde a la categoría de flujos de materiales 5: otros productos. Además, se va a acabar excluyendo por la naturaleza de la cuenta de flujos de materiales. Por lo que no vamos a imputar ninguno de los valores ausentes para este año.

Estas transacciones no van a suponer un cambio relevante en absoluto, por lo que los sacaremos del análisis realizado en los siguientes apartados. Todo ello para que no distorsionen en absoluto las acciones puestas en práctica en el procesamiento de datos.

Finalmente guardamos los archivos finales para cada año en formato .rds para una posterior lectura y tratamiento en R.

Tabla 5.4. Resumen de pasos en el procesamiento de datos.

En primer lugar, se emprende la toma de ciertas decisiones junto a los expertos en la materia en cuanto a la depuración de la base de datos que afecta de lleno a este apartado.

Por un lado, se toma la decisión de efectuar el análisis de los valores atípicos mediante el diagrama de caja de Tukey, esto es, se van a señalar las observaciones que sobresalgan de los umbrales generados mediante el Rango Intercuartílico (RIQ). Esta decisión consensuada con el grupo de expertos se sustenta en la simplicidad del procedimiento y en la experiencia previa en otros análisis realizados para la propia cuenta. Se opta por un sentido estricto a la hora de señalar los valores atípicos y se decide usar un factor de $3 \cdot \text{RIQ}$ para establecer los umbrales.

Las distribuciones que se van a observar van a ser de la variable numérica EUR_KG, diferenciando por las variables "FLUJO", "CN_CODE", "PAIS_ORIGEN_DESTINO". La lógica detrás de esta forma de actuar es que van a ser completamente distintos (en calidad o precio, factores que afectan al precio por kg) los productos importados de los exportados e incluso dentro de las importaciones y exportaciones va a ser muy distinto el país desde donde se importe o exporte el producto. Un ejemplo para ayudar a la comprensión de este hecho es el que observamos con el código 2042982. Este código se corresponde a productos en relación con el vino. Se puede observar en los datos del año 2016 una gran cantidad de importación de productos encasillados en esta categoría con origen de Suecia y unos pocos con origen de Francia. Los productos originarios de Suecia presentan un menor precio por kilogramo que los procedentes de Francia, pero agrupados por países no difieren tanto. Si no se tuviera en cuenta el país las transacciones originarias de Francia para este grupo de productos serían identificadas como datos atípicos.

Una vez identificados los datos atípicos estos deberán ser imputados. Como se ha señalado desde la teoría, el dato en sí que presenta mayores complicaciones es el dato físico PESO_KG. Se decide a través de consulta con expertos que nos serviremos del dato monetario para poder realizar la imputación. Se va a obtener la mediana (med) de la variable EUR_KG agrupada en "FLUJO", "CN_CODE" y "PAIS_ORIGEN_DESTINO" sin tener en cuenta los valores atípicos. Para estos datos, se va a obtener un nuevo PESO_KG* imputado mediante el uso de la mediana y la variable auxiliar de VALOR_EUR: $\text{VALOR_EUR}/\text{med}$.

Por otro lado, ante la falta de respuesta por parte de la AEAT sobre la disponibilidad de los datos desde el año 2000, se va a suponer que la serie de la que se va a disponer en un futuro va a estar formada por datos a partir de 2008. Se toma por tanto la decisión de hacer la depuración del año 2008 de manera individual, la depuración del año 2009 junto con la del 2008 ya depurado, la depuración de 2010 con 2008 y 2009 ya depurados. Así se continuará hasta el

final de la serie, empleando 3 años para la depuración del último año en cuestión. La lógica que tiene este procedimiento es que al disponerse de una serie histórica se va a poder comparar las transferencias de productos en el tiempo. Pero no se va a poder comparar toda la serie al estar involucrados factores como la volatilidad en los precios o la propia naturaleza de los códigos de nomenclatura combinada, que van cambiando cada tiempo.

Al disponer únicamente de los datos presentes en la página web, se emplean los años entre 2016 y 2021 como ejemplo de cómo se replicaría la serie entre 2008 y 2021:

Año depurado	Procedimiento
2016	Únicamente con los datos del año 2016.
2017	Con los datos del año 2016 depurados y con los del año 2017.
2018	Con los del año 2016 y 2017 depurados y con los del año 2018.
2019	Con los del año 2017 y 2018 depurados y con los del año 2019.
2020	Con los del año 2018 y 2019 depurados y con los del año 2020.
2021	Con los del año 2019 y 2020 depurados y con los del año 2021.

El procedimiento mediante el que se ha realizado la depuración ha sido el siguiente:

- Lectura de los datos necesarios para el año a depurar. Para el año 2016 únicamente se tiene en cuenta el año en sí, pero para el resto de los años se tienen en cuenta datos con una depuración realizada.
- Se realiza la lectura del anexo 4 proporcionado por Eurostat en donde aparecen la relación existente entre los códigos de la Nomenclatura Combinada y las categorías de flujos de materiales.
- Se sacan del análisis ciertas transacciones que puedan distorsionar el procedimiento. Dentro de estas se encuentran las identificadas como Energía Eléctrica, los escasos datos ausentes de PESO_KG y los datos que aparecen a 0 en la variable monetaria. Más adelante serán añadidos para el cálculo de los agregados.
- Se identifican las observaciones que sobrepasan el umbral y son marcadas. Aquellas transacciones de las que no se disponen de suficientes datos (< 20, consulta con expertos) para un análisis de datos atípicos no van a entrar a formar parte del análisis y son añadidas sin imputar.
- Se obtiene la mediana sin tener en cuenta las observaciones identificadas como atípicos y se imputa el dato de PESO_KG de las que si se han marcado.
- Guardamos el archivo con el tratamiento realizado para analizar años posteriores.
- Se cruza la base de datos con la del anexo 4 a través de la variable de CN_CODE para tener identificado cada código con su equivalente en categorías en la cuenta de flujos.
- Finalmente, agrupamos según los códigos de la cuenta de flujos y obtenemos los

resultados en formato .xlsx para el posterior trabajo en Excel para el envío a EUROSTAT de los resultados obtenidos.

El número de datos atípicos han sido los siguientes, según el año:

Año depurado	Número de valores identificados como atípicos.	Número de observaciones incluidas en análisis
2016	2.110.760	40.352.567
2017	2.728.188	43.808.516
2018	3.427.046	56.909.361
2019	4.170.722	63.283.361
2020	4.174.007	56.644.941
2021	5.153.370	73.232.714

Resultados.

A continuación, se van a mostrar los resultados obtenidos en la realización del proyecto y que van a suponer las dos líneas de actuación con las que se va a trabajar en un futuro próximo para el desarrollo de la cuenta. Los resultados se van a mostrar como variaciones porcentuales.

En concreto, observamos, por un lado, la variación que existe en la serie de máxima desagregación una vez aplicado el análisis de datos atípicos en la tabla 5.5 para las exportaciones físicas de materiales y en la tabla 5.6 para las importaciones. Y, por otro lado, se examina la variación porcentual entre los datos publicados por el INE y los datos obtenidos de máxima desagregación sin realizar tratamiento alguno para las exportaciones en la tabla 5.7 y para las importaciones en la tabla 5.8.

El primero de los resultados tiene su fundamento en observar la repercusión que tendría la implementación de la técnica estadística seleccionada para ello. Mientras que el segundo tiene un razonamiento basado en observar la repercusión del trabajo previo al análisis de datos ausentes. Esto es, por ejemplo, el análisis de duplicados o la eliminación de aquellas observaciones catalogadas como cuasi-tránsito. Esto final debido a que los datos publicados en el INE para los años mostrados fueron obtenidos mediante las series anuales ya agregadas y con un mínimo grado de depuración.

Tabla 5.5. Variación porcentual entre los datos obtenidos de máxima desagregación sin realizar tratamiento alguno y realizando el tratamiento de atípicos por el RIQ e imputando por la mediana para las Exportaciones.	2016	2017	2018	2019	2020	2021
1.1.1. Cereales	4,5%	3,3%	1,3%	1,2%	3,5%	3,9%
1.1.2. Raíces y tubérculos	2,7%	3,6%	7,9%	5,2%	0,8%	1,9%
1.1.3. Cultivos azucareros	0,1%	0,1%	0,2%	0,4%	-1,3%	7,1%
1.1.4. Leguminosas	0,0%	0,9%	0,1%	-0,5%	1,9%	3,5%
1.1.5. Frutos secos	0,1%	0,5%	0,2%	-0,3%	-0,4%	0,4%
1.1.6. Cultivos oleaginosos	12,5%	10,1%	19,5%	-1,1%	0,8%	2,5%
1.1.7. Hortalizas	1,2%	2,1%	1,5%	1,8%	2,6%	2,9%
1.1.8. Frutas	1,0%	1,4%	1,8%	1,2%	1,3%	1,4%
1.1.9. Fibras	-0,1%	0,0%	-0,5%	-0,1%	0,2%	0,8%
1.1.10. Otros cultivos	0,3%	0,4%	1,0%	0,6%	0,6%	1,4%
1.2.1.1. Paja	10,8%	3,1%	10,2%	6,4%	6,7%	8,5%
1.2.1.2. Otros residuos de cultivos						
1.2.2.1. Cultivos forrajeros	0,5%	0,3%	1,4%	0,8%	0,7%	1,4%
1.3.1. Madera en bruto	3,8%	3,4%	4,1%	2,1%	2,6%	4,7%
1.3.2. Leña y otras extracciones	1,3%	2,7%	-3,1%	-0,3%	1,0%	4,9%
1.4.1. Capturas de peces	0,3%	0,6%	1,3%	0,8%	0,2%	0,7%
1.4.2. Resto de animales y plantas acuáticos	0,8%	1,4%	1,4%	2,1%	1,8%	1,4%
1.5.1. Animales vivos no incluidos en 1.4.	0,7%	1,0%	0,2%	0,8%	1,3%	1,0%
1.5.2. Carne y preparados cárnicos	0,8%	0,8%	1,5%	1,8%	1,1%	1,3%
1.5.3. Productos lácteos, huevos de aves y miel	1,7%	1,7%	1,2%	1,2%	1,6%	1,0%
1.5.4. Otros productos de animales (fibras de animales, piel, pelo, cueros, etc.)	2,9%	2,2%	6,2%	2,2%	1,8%	6,5%
1.6. Productos principalmente de biomasa	2,3%	3,2%	4,2%	4,8%	6,7%	7,6%
2.1. Minerales y concentrados de hierro, hierro y acero	1,2%	2,2%	1,0%	0,9%	1,4%	4,6%
2.2.1. Cobre	-0,2%	0,2%	0,6%	-1,1%	-1,6%	0,8%
2.2.2. Níquel	0,0%	0,6%	0,3%	6,4%	-0,1%	9,7%
2.2.3. Plomo	0,0%	0,0%	0,3%	-0,2%	0,0%	0,1%
2.2.4. Zinc	0,3%	-1,1%	-1,1%	-1,7%	-3,8%	-2,2%
2.2.5. Estaño	0,3%	0,3%	43,2%	0,5%	-0,5%	3,0%
2.2.6. Oro, plata, platino y otros metálicos preciosos	0,2%	0,7%	1,3%	8,5%	0,6%	1,1%
2.2.7. Bauxita y otros aluminios	2,4%	1,2%	0,8%	2,2%	2,4%	2,5%
2.2.8. Uranio y torio	0,1%	3,1%	19,2%	14,0%	18,2%	19,1%
2.2.9. Otros metales	0,4%	2,5%	14,7%	1,0%	-3,9%	3,1%
2.3. Productos principalmente de metales	2,2%	2,8%	3,4%	3,7%	-2,6%	2,8%
3.1. Piedras de construcción u ornamentales (excepto pizarra)	3,8%	3,2%	3,6%	3,6%	3,8%	4,2%
3.2. Creta y dolomita	3,9%	6,1%	2,0%	7,1%	3,9%	-3,0%
3.3. Pizarra	0,9%	3,2%	2,4%	6,8%	0,2%	8,0%
3.4. Minerales para fertilizantes y químicos	0,7%	0,7%	1,2%	-0,5%	-0,5%	0,1%
3.5. Sal	1,7%	3,4%	0,0%	1,0%	1,2%	42,9%
3.6. Caliza y yeso	6,1%	6,2%	9,6%	2,8%	7,2%	4,1%

3.7. Arcillas y caolín	3,1%	4,2%	4,3%	5,4%	5,5%	5,9%
3.8. Arena y grava	3,7%	5,5%	4,8%	5,3%	8,0%	13,0%
3.9. Otros minerales no metálicos	-6,7%	-10,1%	1,2%	-1,2%	22,4%	13,5%
3.10. Productos principalmente de minerales no metálicos	0,6%	1,0%	1,3%	3,5%	1,3%	2,4%
4.1.1. Lignito	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
4.1.2. Hulla y antracita	0,5%	1,1%	-1,4%	-0,2%	-0,1%	-1,0%
4.1.3. Arenas de alquitrán y pizarras bituminosas	-2,7%	0,0%	0,0%	0,0%	0,0%	-0,1%
4.1.4. Turba	1,2%	1,4%	101,4%	2,0%	7,2%	15,0%
4.2.1. Crudos de petróleo y gas natural licuado	-0,1%	1,7%	0,4%	0,0%	-2,4%	1,9%
4.2.2. Gas natural	0,2%	0,4%	0,9%	-0,4%	-4,6%	0,0%
4.3. Productos principalmente de combustibles fósiles	1,1%	1,4%	1,5%	1,5%	1,6%	2,4%
5. Otros productos	2,2%	2,1%	3,5%	3,4%	4,1%	4,8%
6. Residuos exportados para su tratamiento final y eliminación	0,0%	0,0%	0,1%	0,0%	0,0%	2,8%

Tabla 5.6. Variación porcentual entre los datos obtenidos de máxima desagregación sin realizar tratamiento alguno y realizando el tratamiento de atípicos por el RIQ e imputando por la mediana para las Importaciones.	2016	2017	2018	2019	2020	2021
1.1.1. Cereales	0,6%	-0,5%	-3,3%	0,5%	0,6%	6,7%
1.1.2. Raíces y tubérculos	3,2%	3,4%	4,4%	4,3%	4,6%	1,7%
1.1.3. Cultivos azucareros	0,5%	0,8%	-1,0%	-0,3%	-0,1%	2,5%
1.1.4. Leguminosas	0,1%	0,3%	0,2%	0,4%	0,7%	2,0%
1.1.5. Frutos secos	-0,3%	1,4%	1,2%	1,7%	1,4%	0,7%
1.1.6. Cultivos oleaginosos	0,4%	0,0%	0,3%	-0,1%	-0,1%	0,2%
1.1.7. Hortalizas	3,9%	5,4%	4,1%	1,1%	2,2%	2,4%
1.1.8. Frutas	-0,5%	1,2%	0,3%	1,9%	1,6%	1,9%
1.1.9. Fibras	1,1%	1,9%	0,3%	0,2%	0,5%	0,9%
1.1.10. Otros cultivos	0,0%	-0,1%	0,6%	0,5%	0,9%	1,3%
1.2.1.1. Paja	0,0%	16,1%	-4,9%	-0,5%	2,9%	-5,3%
1.2.1.2. Otros residuos de cultivos						
1.2.2.1. Cultivos forrajeros	0,0%	3,7%	0,1%	-0,8%	2,3%	0,5%
1.3.1. Madera en bruto	7,4%	8,9%	8,9%	9,8%	5,8%	10,3%
1.3.2. Leña y otras extracciones	-0,2%	1,3%	0,4%	-2,9%	2,9%	0,8%
1.4.1. Capturas de peces	0,3%	0,6%	1,0%	0,7%	0,4%	0,7%
1.4.2. Resto de animales y plantas acuáticos	1,3%	0,7%	0,9%	0,7%	0,6%	0,7%
1.5.1. Animales vivos no incluidos en 1.4.	2,3%	2,2%	2,1%	3,5%	2,0%	2,7%
1.5.2. Carne y preparados cárnicos	-0,1%	0,5%	0,7%	0,6%	-0,3%	-0,2%
1.5.3. Productos lácteos, huevos de aves y miel	0,2%	-1,4%	0,1%	0,4%	0,9%	1,0%
1.5.4. Otros productos de animales (fibras de animales, piel, pelo, cueros, etc.)	4,8%	2,8%	1,8%	1,3%	2,0%	1,5%
1.6. Productos principalmente de biomasa	0,5%	0,7%	1,4%	2,6%	3,2%	2,8%

2.1. Minerales y concentrados de hierro, hierro y acero	1,4%	1,4%	1,4%	1,8%	0,2%	3,6%
2.2.1. Cobre	0,3%	0,4%	1,0%	0,6%	2,0%	0,4%
2.2.2. Níquel	0,0%	0,1%	0,3%	0,8%	0,6%	0,0%
2.2.3. Plomo	1,3%	2,0%	0,7%	0,3%	0,2%	4,4%
2.2.4. Zinc	0,0%	0,0%	0,0%	0,1%	0,9%	-0,2%
2.2.5. Estaño	0,1%	0,3%	1,5%	0,3%	0,1%	2,8%
2.2.6. Oro, plata, platino y otros metálicos preciosos	0,1%	0,2%	-2,7%	7,4%	1,7%	1,2%
2.2.7. Bauxita y otros aluminios	1,8%	0,5%	0,3%	0,3%	0,7%	4,2%
2.2.8. Uranio y torio	1,6%	6,6%	2,1%	1,0%	0,6%	-5,7%
2.2.9. Otros metales	0,6%	-0,1%	1,1%	3,5%	9,1%	1,0%
2.3. Productos principalmente de metales	92,9%	368,4%	630,6%	34,2%	21,7%	3,7%
3.1. Piedras de construcción u ornamentales (excepto pizarra)	26,4%	2,8%	2,8%	3,4%	2,6%	2,2%
3.2. Creta y dolomita	471,0%	466,4%	391,5%	517,1%	7,3%	2,8%
3.3. Pizarra	0,3%	2,5%	0,2%	0,8%	0,7%	0,3%
3.4. Minerales para fertilizantes y químicos	0,3%	0,7%	2,2%	10,2%	0,9%	16,9%
3.5. Sal	6,1%	3,5%	12,0%	2,8%	2,0%	-3,3%
3.6. Caliza y yeso	-23,8%	-1,1%	-11,1%	-17,9%	3,8%	-2,1%
3.7. Arcillas y caolín	1,0%	0,5%	0,6%	-0,1%	-2,8%	0,2%
3.8. Arena y grava	5,3%	7,1%	13,3%	8,5%	23,0%	83,7%
3.9. Otros minerales no metálicos	1,9%	8,6%	3,4%	11,2%	10,4%	12,9%
3.10. Productos principalmente de minerales no metálicos	6,7%	3,3%	1,7%	0,7%	4,2%	4,2%
4.1.1. Lignito	0,0%	18,3%	2,3%	39,7%	144,8%	124,5%
4.1.2. Hulla y antracita	2,0%	1,4%	0,0%	0,3%	0,3%	7,7%
4.1.3. Arenas de alquitrán y pizarras bituminosas	0,0%	0,1%	1,0%	466,0%	144,1%	24,8%
4.1.4. Turba	6,6%	5,8%	8,3%	4,4%	7,7%	5,1%
4.2.1. Crudos de petróleo y gas natural licuado	0,1%	0,2%	0,4%	0,1%	-1,1%	8,4%
4.2.2. Gas natural	0,5%	0,0%	0,8%	0,0%	0,7%	0,3%
4.3. Productos principalmente de combustibles fósiles	1,6%	2,0%	2,5%	1,9%	2,0%	4,1%
5. Otros productos	3,0%	3,4%	3,9%	3,7%	6,9%	6,8%
6. Residuos importados para su tratamiento final y eliminación	799,4%	23889,9%	370329,9%	198477,9%	200814,6%	0,5%

Tabla 5.7. Variación porcentual entre los datos publicados por el INE y los datos obtenidos de máxima desagregación sin realizar tratamiento alguno para las Exportaciones.	2016	2017	2018	2019
1.1.1. Cereales	-2,3%	-13,9%	-12,8%	-1,8%
1.1.2. Raíces y tubérculos	-0,5%	-0,3%	-0,5%	-0,6%
1.1.3. Cultivos azucareros	-2,3%	0,0%	-0,7%	-0,5%
1.1.4. Leguminosas	-5,7%	-29,2%	-29,2%	-2,1%
1.1.5. Frutos secos	-9,1%	-10,8%	-11,8%	-5,8%
1.1.6. Cultivos oleaginosos	-0,2%	-0,6%	-0,6%	-0,3%
1.1.7. Hortalizas	-1,2%	-8,1%	-7,6%	-1,5%

1.1.8. Frutas	-1,3%	-12,3%	-11,3%	-1,6%
1.1.9. Fibras	-0,1%	-0,9%	-1,5%	-0,8%
1.1.10. Otros cultivos	-3,8%	-3,4%	-3,6%	-4,7%
1.2.1.1. Paja	0,0%	0,0%	0,0%	-0,3%
1.2.1.2. Otros residuos de cultivos				
1.2.2.1. Cultivos forrajeros	-4,2%	-46,4%	-45,1%	-1,6%
1.3.1. Madera en bruto	-1,5%	-8,3%	-8,1%	-0,6%
1.3.2. Leña y otras extracciones	-0,3%	-0,5%	-0,3%	-0,2%
1.4.1. Capturas de peces	-8,9%	-10,9%	-10,9%	-2,8%
1.4.2. Resto de animales y plantas acuáticos	-21,6%	-20,0%	-19,2%	-0,9%
1.5.1. Animales vivos no incluidos en 1.4.	-0,4%	-0,8%	-0,6%	-0,3%
1.5.2. Carne y preparados cárnicos	-1,0%	-1,6%	-1,5%	-1,1%
1.5.3. Productos lácteos, huevos de aves y miel	-0,3%	-6,3%	-6,2%	-1,1%
1.5.4. Otros productos de animales (fibras de animales, piel, pelo, cueros, etc.)	-0,9%	-3,8%	-4,2%	-0,8%
1.6. Productos principalmente de biomasa	-17,5%	-34,5%	-31,2%	-4,9%
2.1. Minerales y concentrados de hierro, hierro y acero	-0,5%	-4,3%	-5,2%	-0,4%
2.2.1. Cobre	-1,2%	-10,3%	-8,3%	-1,1%
2.2.2. Níquel	0,0%	-0,3%	-1,6%	-0,1%
2.2.3. Plomo	0,0%	0,0%	0,0%	0,0%
2.2.4. Zinc	0,0%	-0,9%	-1,7%	-0,1%
2.2.5. Estaño	-12,1%	-5,9%	-5,5%	-6,2%
2.2.6. Oro, plata, platino y otros metálicos preciosos	0,5%	0,0%	0,0%	0,0%
2.2.7. Bauxita y otros aluminios	-0,1%	-0,6%	-0,7%	-0,3%
2.2.8. Uranio y torio	-1,1%	-1,3%	-0,3%	0,0%
2.2.9. Otros metales	-4,0%	-3,5%	-2,5%	-1,8%
2.3. Productos principalmente de metales	-28,1%	-35,8%	-36,0%	-5,0%
3.1. Piedras de construcción u ornamentales (excepto pizarra)	-0,1%	-30,3%	-31,0%	-0,3%
3.2. Creta y dolomita	-1,2%	-1,7%	-1,0%	-1,2%
3.3. Pizarra	20,7%	-0,5%	-0,1%	-0,6%
3.4. Minerales para fertilizantes y químicos	-0,5%	-21,8%	-21,9%	-2,3%
3.5. Sal	-1,3%	-1,7%	-2,1%	-3,8%
3.6. Caliza y yeso	-0,1%	-0,7%	-1,1%	-0,3%
3.7. Arcillas y caolín	-1,2%	-5,4%	-6,6%	-1,3%
3.8. Arena y grava	-1,6%	-1,3%	-0,9%	-0,6%
3.9. Otros minerales no metálicos	-0,6%	3,8%	-0,3%	-0,1%
3.10. Productos principalmente de minerales no metálicos	-0,7%	-9,7%	-12,6%	-0,9%
4.1.1. Lignito	0,4%	12,5%	0,0%	0,1%
4.1.2. Hulla y antracita	-0,1%	-0,9%	-1,6%	-0,3%
4.1.3. Arenas de alquitrán y pizarras bituminosas	0,0%	0,0%	0,0%	0,0%
4.1.4. Turba	-4,8%	-4,8%	-4,0%	-2,9%
4.2.1. Crudos de petróleo y gas natural licuado	-26,6%	-22,7%	-19,0%	2,3%
4.2.2. Gas natural	-2,7%	-2,1%	-2,8%	-3,6%
4.3. Productos principalmente de combustibles fósiles	-8,1%	-10,1%	-14,0%	-6,5%
5. Otros productos	-18,0%	-11,0%	-19,6%	-34,7%
6. Residuos exportados para su tratamiento final y eliminación	0,0%	0,0%	-31,8%	0,0%

Tabla 5.8. Variación porcentual entre los datos publicados por el INE y los datos obtenidos de máxima desagregación sin realizar tratamiento alguno para las Importaciones.	2016	2017	2018	2019
1.1.1. Cereales	-0,6%	-1,4%	-1,8%	-0,9%
1.1.2. Raíces y tubérculos	-2,3%	-1,7%	-1,7%	-2,8%
1.1.3. Cultivos azucareros	-0,6%	-1,3%	-0,9%	-1,2%
1.1.4. Leguminosas	-4,9%	-5,2%	153,3%	-0,5%
1.1.5. Frutos secos	-11,4%	-13,9%	-13,4%	-7,9%
1.1.6. Cultivos oleaginosos	-0,3%	-0,5%	-0,5%	-0,6%
1.1.7. Hortalizas	-4,7%	-8,8%	-9,6%	-6,2%
1.1.8. Frutas	-7,4%	-19,0%	-19,8%	-8,9%
1.1.9. Fibras	-2,2%	-1,1%	-1,5%	-1,5%
1.1.10. Otros cultivos	-3,7%	-4,0%	-4,2%	-3,7%
1.2.1.1. Paja	0,0%	-1,4%	0,0%	-0,1%
1.2.1.2. Otros residuos de cultivos				
1.2.2.1. Cultivos forrajeros	-0,2%	-48,2%	-48,4%	0,0%
1.3.1. Madera en bruto	-6,1%	-14,7%	-12,9%	-1,7%
1.3.2. Leña y otras extracciones	-6,1%	-5,3%	-3,7%	-3,8%
1.4.1. Capturas de peces	-8,9%	-12,4%	-13,0%	-4,8%
1.4.2. Resto de animales y plantas acuáticos	-25,3%	-25,5%	-25,4%	-3,5%
1.5.1. Animales vivos no incluidos en 1.4.	-0,1%	-1,9%	-1,6%	-0,1%
1.5.2. Carne y preparados cárnicos	-1,7%	-5,6%	-6,4%	-1,5%
1.5.3. Productos lácteos, huevos de aves y miel	-2,5%	-10,0%	-10,5%	-2,7%
1.5.4. Otros productos de animales (fibras de animales, piel, pelo, cueros, etc.)	-0,7%	-3,2%	-4,3%	-0,8%
1.6. Productos principalmente de biomasa	-4,0%	-11,9%	-13,0%	-1,3%
2.1. Minerales y concentrados de hierro, hierro y acero	-0,3%	-5,9%	-6,1%	-0,2%
2.2.1. Cobre	-18,4%	-34,4%	-27,7%	-20,4%
2.2.2. Níquel	65,5%	-4,0%	-3,1%	-1,5%
2.2.3. Plomo	-0,4%	-1,4%	26,0%	-0,8%
2.2.4. Zinc	-4,4%	-11,0%	-4,1%	0,0%
2.2.5. Estaño	-2,4%	-7,2%	-5,3%	-1,2%
2.2.6. Oro, plata, platino y otros metálicos preciosos	-1,0%	-0,2%	382,8%	-0,6%
2.2.7. Bauxita y otros aluminios	-1,9%	-2,1%	-1,7%	-0,9%
2.2.8. Uranio y torio	0,0%	-0,3%	-0,8%	-1,1%
2.2.9. Otros metales	-0,1%	-0,4%	-0,6%	-1,6%
2.3. Productos principalmente de metales	-20,8%	-30,9%	-30,3%	-4,2%
3.1. Piedras de construcción u ornamentales (excepto pizarra)	-3,9%	-15,5%	-15,9%	-2,8%
3.2. Creta y dolomita	-1,1%	-0,1%	-0,1%	-0,1%
3.3. Pizarra	-0,1%	-0,4%	0,0%	-1,4%
3.4. Minerales para fertilizantes y químicos	-0,6%	-25,6%	-25,1%	-0,7%
3.5. Sal	-0,8%	-0,7%	-1,2%	-0,7%
3.6. Caliza y yeso	-1,2%	110,1%	-3,2%	-4,4%
3.7. Arcillas y caolín	-0,9%	-12,7%	-16,0%	-0,8%
3.8. Arena y grava	-2,5%	-1,4%	-1,5%	-1,7%
3.9. Otros minerales no metálicos	-0,9%	-2,3%	-2,7%	-0,9%

3.10. Productos principalmente de minerales no metálicos	-8,7%	-31,1%	-30,3%	-3,2%
4.1.1. Lignito	0,0%	0,0%	-3,3%	-3,4%
4.1.2. Hulla y antracita	0,0%	0,0%	0,0%	-0,1%
4.1.3. Arenas de alquitrán y pizarras bituminosas	0,0%	23,4%	0,0%	0,0%
4.1.4. Turba	-5,4%	-7,4%	-6,4%	-5,8%
4.2.1. Crudos de petróleo y gas natural licuado	6,3%	5,9%	6,3%	6,8%
4.2.2. Gas natural	0,0%	-0,1%	0,0%	0,0%
4.3. Productos principalmente de combustibles fósiles	-4,1%	-8,8%	-10,7%	-3,1%
5. Otros productos	-8,1%	-13,5%	-14,2%	-6,0%
6. Residuos importados para su tratamiento final y eliminación	-40,4%	-33,2%	-21,0%	-22,8%

Mirando en primer lugar la tabla 5.5 podemos ver como el procedimiento de análisis e imputación de datos ausentes tiene cierta relevancia en casi todos los grupos de materiales en las exportaciones, salvo para los casos de las categorías de “Otros Productos” y de “Residuos exportados para su tratamiento final y eliminación”.

En la categoría de “Biomasa y productos de la biomasa observamos situaciones aisladas en los casos de “Raíces y tubérculos”, “Otros productos de animales” o “Productos principalmente de biomasa”. Estas dos últimas categorías abarcan productos considerablemente heterogéneos y con menos similitudes que en el resto de las categorías. Por lo que es entendible que se produzcan desajustes ante la gran variedad de productos almacenados en una misma categoría de materiales.

Otro caso distinto es el que se observa para la categoría de la “Paja”, en la cual se están produciendo cambios con cierta relevancia durante toda la serie analizada. Algo que tiene que ser tenido en cuenta y ser evaluado con especial detenimiento.

En el caso de los “Minerales y concentrados de metal, primarios y procesados” encontramos que en el “Estaño” se produce una variación anormal en el año 2018, la cual va a requerir de especial atención. Al igual que el caso del “Uranio y torio” para los cuales la variación en el tiempo desde 2018 puede ser catalogada de relevante.

Pero es en la categoría de “Minerales no metálicos, primarios y procesados” donde vemos que se están produciendo la mayoría de los cambios en proporción a una categoría de materiales. Casi todas las subcategorías presentan variaciones relevantes teniendo grandes cambios en el año 2021 para la “Sal” y en el año 2020 para “Otros minerales no metálicos”.

Finalmente, vemos que el mayor de los cambios en las exportaciones por la metodología puesta en práctica va a ocurrir en la de “Combustibles fósiles, primarios y procesados”, en concreto en el caso de la “Turba”, donde en el año 2018 la variación es de un 101,4%, algo descomunal.

En segundo lugar, en el caso de las importaciones, en la tabla 5.6, se puede ver como el efecto del análisis de atípicos tienen menos repercusión en la categoría de “biomasa y productos de la biomasa” que en el caso de las exportaciones. Aisladamente estaría ocurriendo algún hecho en el caso de la “Paja” para el año 2017 y también detectamos un efecto constante en el tiempo en el caso de la “Madera en bruto”.

Algo parecido en lo que al número de subcategorías de materiales afectadas se refiere ocurrirá en el caso de los “Minerales y concentrados de metal, primarios y procesados”. De gran relevancia también los cambios que vemos que se producen en el tiempo para la categoría de “Productos principalmente de metales”. Algo similar a lo que expresábamos con categorías que son muy genéricas en la parte de las exportaciones. En esta categoría se estarán incluyendo grandes variedades de productos con poca relación entre ellos, distorsionando el proceso de localización e imputación de atípicos.

Nuevamente en la categoría de “Minerales no metálicos, primarios y procesados” se puede observar que en el caso de las importaciones también sería el que mayor número de categorías sufrirán cambios abismales por el procedimiento aplicado de análisis de outliers. Especialmente cambios considerables lo que ocurre con la “Creta y dolomita”, cambios por encima del 300% desde 2016 hasta 2019. A esto se añaden los cambios que ocurrirán en la “Caliza y yeso” o en la “Arena y grava”

Por todo ello, esta categoría de “Minerales no metálicos, primarios y procesados” va a requerir de una especial atención en los análisis futuros.

Finalmente, observamos nuevas diferencias con respecto a lo que ocurría en las exportaciones en las categorías de “Combustibles fósiles” y “Residuos importados para su tratamiento final y eliminación”. De presentar escasos casos de cambio en las exportaciones a tener cambios abismales en las importaciones. Casos del “Lignito” y las “Arenas de alquitrán y pizarras bituminosas” donde se presentan cambios extremos a partir de 2019 hasta el final de la serie en ambos casos. Caso aparte la circunstancia que se observa en la categoría de “Residuos exportados para su tratamiento final y eliminación” en prácticamente toda la serie, cambios exagerados que nos indican que algo está pasando en el análisis con esta categoría de una forma que no deja lugar a dudas.

La segunda parte de los resultados, mostrados en las tablas 5.7 y 5.8, van a mostrar la influencia de haber realizado un tratamiento previo al análisis de datos atípicos con la base de datos de máxima desagregación. En este tratamiento varias de las transacciones son eliminadas por considerarse duplicadas (al repetirse los datos en las variables mostradas en la metodología) o por tratarse de transacciones cuasi-tránsito. De ahí la lógica de ver que en la mayoría de las categorías de materiales se reduce la cantidad total.

Solo mostramos los casos observados entre 2016 y 2019 debido a que para los años 2020 y 2021 aún no se encuentran publicados los datos definitivos y no se considera oportuno la exposición de las variaciones sobre algo que aún no ha sido difundido. Por ello exponemos los años mencionados y más adelante se replicará el ejercicio para el resto de los años una vez contemos con los datos proporcionados por la AEAT.

Por ello, van a destacar y van a requerir de una investigación más detallada aquellos casos en los que se está produciendo un aumento de las cantidades en lugar de una reducción. Puede que en estos casos esté influyendo la depuración manual realizada a la hora de publicar, algún error en el tratamiento o en su defecto algunas transacciones que no se estén teniendo en cuenta en las series anuales y que si deban de considerarse.

Los pocos casos en los que vemos esta situación serian, para las exportaciones, la categoría de la “Pizarra” en el año 2016 con un aumento del 20,7%, el Lignito en el año 2017 con un aumento del 12,5% y los “Crudos de petróleo y gas natural licuado” en el año 2019 con un 2,3%.

En el caso de las importaciones estos casos estarían en las “Leguminosas” con un aumento del 153,3% en 2018, el “Níquel” en 2016 con un aumento del 65,5%, el “Plomo” en 2018 con un 26,0%, “Oro, plata, platino y otros metálicos preciosos” en 2018 con un 382,8%, “Caliza y yeso” en 2017 con un 110,1% de aumento, “Arenas de alquitrán y pizarras bituminosas” en 2017 con un 23,4% y finalmente un porcentaje de entorno al 6% de aumento en todos los años para los “Crudos de petróleo y gas natural licuado”.

Otra situación particular es la similitud en la variación para los años 2017 y 2018 que también nos sorprende. Que varíen las cantidades en una proporción tan similar tanto en las exportaciones como en las importaciones para esos años en concreto es algo que llama nuestra atención y que requerirá de un análisis particular.

6. Conclusiones y Trabajo Futuro.

La principal finalidad de este proyecto ha consistido en la automatización del proceso de lectura y la preparación de una nueva fuente de datos para su posterior depuración, con lo que se pretende impulsar un aumento de la calidad en la elaboración de la CFM-TE de las importaciones y exportaciones físicas de materiales. La fuente de datos empleada tiene su origen en los registros administrativos, explicándose en el trabajo las diferentes ventajas que tiene la creación de productos estadísticos a través de ellos, al igual que también los requisitos necesarios para que su uso se corresponda con la creación de unos productos estadísticos de calidad. Posteriormente se ha reflexionado sobre las peculiaridades de la compilación de la cuenta, sobre los diferentes métodos para detectar incongruencias en la misma y las distintas formas para imputar los valores detectados como contradictorios. Después hemos aplicado una metodología consensuada con los expertos en materia en los datos de los años de 2016 a 2021 para ilustrar como sería el procedimiento de compilación de la cuenta. Todo ello ha sido realizado en un ordenador con la capacidad de memoria RAM de 32GB, con el programa R.

Obtenemos diversas conclusiones sobre el trabajo realizado hasta ahora mediante esta nueva forma de compilación de las importaciones y exportaciones físicas de materiales. La primera de ellas tiene que ver con el salto de calidad que supone la utilización de la nueva fuente de datos. Con los datos de las series mensuales de máxima desagregación se consigue una mejora de la oportunidad puesto que los datos de comercio exterior de este tipo son publicados por la AEAT con solo 2 meses de demora al mes donde se produjo el suceso. Adicionalmente, al tener tal nivel de desagregación se nos permite presentar un dato más cercano a la realidad en el caso de las importaciones, pudiendo haber detectado las operaciones cuasi-tránsito que se producen. En definitiva, el principal beneficio de haber podido procesar tal volumen de información ha sido el disponer de la capacidad de decisión a la hora de trabajar con las transacciones de comercio exterior. Y no lo que nos ocurría con los datos agregados anualmente, donde se desconocía que datos estaban siendo incluidos en los agregados por código de NC y país involucrado, en los que además se presentaban datos ausentes en el peso, algo que en las de máxima desagregación no ocurre apenas.

Otra de las conclusiones tiene que ver con la forma en la que se encuentran identificados los productos en las estadísticas de comercio exterior y el análisis implementado de datos atípicos. Como señalábamos anteriormente, los productos se encuentran identificados por medio de los códigos de NC. En la mayoría de los casos existe una suficiente precisión en el código para tratarse de productos homogéneos en los que un aumento del precio suponga un aumento del valor de la transacción. Pero esta ventaja supondrá un pequeño inconveniente en ciertos casos en los que se incluye en un mismo código varios tipos de productos dispares en sus

características en cuanto a peso y valor. Un ejemplo lo encontramos en productos de joyería incluidos en una misma categoría en la que a pesar de tener el mismo peso, uno tenga mayor valor que el otro. Además, recordemos también que cada producto es asignado a una categoría de flujo de material según el material más pesado que contiene. Quizá alguna pieza más pequeña pueda estar desatando el valor del producto y no se corresponda en absoluto con contenido más pesado.

Mediante los resultados obtenidos en el análisis de datos atípicos hemos detectado ciertas subcategorías donde se estarían produciendo grandes desequilibrios y que pueden ser el origen de la necesidad de una técnica distinta a la empleada. Para estos casos donde los códigos de NC no supongan la agrupación de productos homogéneos, técnicas como la del análisis de conglomerados podrían ayudar a detectar y a agrupar productos más similares dentro de cada código de NC. La base de datos de las series mensuales de máxima desagregación presenta una cuantía numerosa de variables para ayudar en este cometido por lo que el trabajo futuro consistirá en determinar si está técnica puede ser ejecutada con las variables con las que contamos.

Finalmente, también entendemos de absoluta necesidad en la parte de trabajo futuro el hecho de que se establezca una colaboración más estrecha con la AEAT, en la conveniencia de aclarar ciertos puntos como los siguientes:

- Esclarecer la metodología para la estimación de aquellas transacciones que por reglamento no deben ser declaradas en régimen Intrastat. Hablamos de aquellos sujetos que no hayan excedido el umbral de exención de 400.000 euros anuales.
- Fijar la posibilidad de un envío de datos habitual adicionalmente al envío de los datos para poder completar la serie que nos compete.
- Precisar la metodología que se está llevando a cabo en la elaboración de las estadísticas de comercio exterior originadas mediante de registros administrativos. Esto a raíz de la numerosa cuantía de datos duplicados que encontramos en el análisis.
- Explicación sobre que puede estar ocurriendo con aquellas transacciones en las que la variable monetaria aparece dato de 0. En consultas anteriores se nos había comentado de la calidad de esta variable por lo que nos extraña esta situación.

Todo este trabajo se encuentra enmarcado en la necesidad de unas estadísticas medioambientales de calidad para acometer todos los desafíos que se vislumbran en el futuro. Uno de los objetivos futuros desde las CFM-TE trataría sobre la ampliación de las estadísticas expuestas en este trabajo a las Cuentas de Flujos Materiales-Unidades de Material Equivalente. Para ello es fundamental la continuación de la línea que sigue este trabajo, esto es, el desarrollo de unas estadísticas de importaciones y exportaciones físicas de calidad.

Bibliografía.

Alonso, Odendahl, Sánchez, & Viani. (2022). Recuadro 2. Posibles canales de impacto económico en la UEM del conflicto bélico en Ucrania. *Boletín Económico / Banco de España*, 30–32.

<https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/InformesBoletinesRevistas/BoletinEconomico/22/T1/Fich/be2201-it-Rec2.pdf>.

Bryan, J. [aut, cre], Wickham, H. [aut] et al. Paquete: readxl. Read excel files.

<https://CRAN.R-project.org/package=readxl>

Carlsson, A., Wadeskog, A., Palm, V., Kanlén, F. & Statistics Sweden. (2006). Material Flow Accounts and Policy. Data for Sweden 2004. UNSTAT.

https://unstats.un.org/UNSD/envaccounting/ceea/archive/MFA/Sweden_MFA_2007.pdf

Celestino, F., Maestro López, J., & Martín Salvador, L. (2021). Las cuentas medioambientales físicas en el INE: situación actual y perspectivas futuras. *Cuadernos Económicos De ICE*, (101). <https://doi.org/10.32796/cice.2021.101.7189>.

De Waal, T., Pannekoek, J., Scholtus, S., & Waal, T. D. (2011). Handbook of statistical data editing and imputation. John Wiley & Sons, Incorporated.

Delahaye, R. & Nootenboom, L. (2008, diciembre). Economy-wide material flow accounts in the Netherlands. CBS. [https://www.cbs.nl/-/media/imported/onze%20diensten/methoden/dataverzameling/aanvullende%20onderzoeksbes](https://www.cbs.nl/-/media/imported/onze%20diensten/methoden/dataverzameling/aanvullende%20onderzoeksbeschrijvingen/documents/2010/44/2008-mfa-netherlands.pdf)

[chrijvingen/documents/2010/44/2008-mfa-netherlands.pdf](https://www.cbs.nl/-/media/imported/onze%20diensten/methoden/dataverzameling/aanvullende%20onderzoeksbeschrijvingen/documents/2010/44/2008-mfa-netherlands.pdf). Project number: 206691.

Delahaye, R. & Nootenboom, L. (2009, diciembre). Material flow accounts in the Netherlands, time series 1996–2006. CBS. [https://www.cbs.nl/-/media/imported/onze-](https://www.cbs.nl/-/media/imported/onze-diensten/methoden/dataverzameling/aanvullende-onderzoeksbeschrijvingen/documents/2010/44/2009-mfa-netherlands.pdf?la=nl-nl)

[diensten/methoden/dataverzameling/aanvullende-](https://www.cbs.nl/-/media/imported/onze-diensten/methoden/dataverzameling/aanvullende-onderzoeksbeschrijvingen/documents/2010/44/2009-mfa-netherlands.pdf?la=nl-nl)
[onderzoeksbeschrijvingen/documents/2010/44/2009-mfa-netherlands.pdf?la=nl-nl](https://www.cbs.nl/-/media/imported/onze-diensten/methoden/dataverzameling/aanvullende-onderzoeksbeschrijvingen/documents/2010/44/2009-mfa-netherlands.pdf?la=nl-nl). Project number: 207545

Dowle, M. [aut, cre] et al. (2019): Paquete: data.table. Extension of 'data.frame'

<https://CRAN.R-project.org/package=data.table>

Eurostat (2001): *Economy-wide Material Flow Accounts and Derived Indicators – A Methodological Guide*. Luxembourg

<http://ec.europa.eu/eurostat/documents/1798247/6191533/3-Economy-wide-material-flow-accounts...-A-methodological-guide-2000-edition.pdf/9dfae42d-0831-4522-9fe5-571785f8fecf>.

Eurostat (2014): 21st Meeting of the European Statistical System Committee. Item 12 of the agenda. Other National Authorities (ONAs) – practical considerations.

https://ec.europa.eu/eurostat/documents/13019146/13574152/ESSC_2014_21_EN_12_Other_National_Authorities.pdf/f7f456b1-8670-a578-7f57-e332c326b08b?t=1648713280420.

Eurostat (2017a): *Further clarifying the conceptual treatment of physical imports and exports in economy-wide material flow accounts*. Technical Note EEEA/2017/02

<http://ec.europa.eu/eurostat/documents/1798247/6191533/Technical+Note+on+physical+trade>.

Eurostat (2018): Economy-wide material flow accounts. HANDBOOK. Luxembourg

<https://ec.europa.eu/eurostat/documents/3859598/9117556/KS-GQ-18-006-EN-N.pdf/b621b8ce-2792-47ff-9d10-067d2b8aac4b?t=1537260841000>.

Eurostat (2020): User guide on European statistics on international trade in goods.

HANDBOOK. Luxembourg. <https://ec.europa.eu/eurostat/documents/3859598/12137783/KS-GQ-20-012-EN-N.pdf/f982fc06-3ff8-d37b-298f-9c76c843ae52?t=1608633443374>.

Eurostat (2021): The European System of Accounts — ESA 2010 — interactive version

<https://ec.europa.eu/eurostat/esa2010/>.

Hawkins, D. M. (1980). Identification of outliers (Ser. Monographs on applied probability and statistics). Chapman & Hall.

Iriarte Goñi, I. (2021, 27 julio). Si se mide mal, las decisiones serán equivocadas: las medidas del crecimiento. Conversación sobre Historia.

<https://conversacionsobrehistoria.info/2021/07/27/si-se-mide-mal-las-decisiones-seran-equivocadas-las-medidas-del-crecimiento/>.

Kataryniuk, Ró Lopezosa, & Sánchez Carretero. (2021). Recuadro 3. Los cuellos de botella del sector manufacturero de la zona del euro. Boletín Económico / Banco de España, 30–34.

<https://repositorio.bde.es/bitstream/123456789/17635/1/be2103-it-Rec3.pdf>.

Ley 13/2022, de 7 de julio, General de Comunicación Audiovisual. BOE núm. 163, de 8 de julio de 2022, páginas 96114 a 96220 <https://www.boe.es/eli/es/l/2022/07/07/13>

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median.

<https://www.sciencedirect.com/science/article/pii/S0022103113000668>.

<https://doi.org/10.1016/j.jesp.2013.03.013>

Luzi, O., De Waal, T., Hulliger, B., Di Zio, M., Pannekoek, J., Kilchmann, D., Guarnera, U., Hoogland, J., Manzari, A., & Tempelman, C. (2007). Recommended Practices for Editing and

Imputation in Cross-Sectional Business Surveys.

<https://ec.europa.eu/eurostat/documents/64157/4374310/30-Recommended+Practices-for-editing-and-imputation-in-cross-sectional-business-surveys-2008.pdf/6e51b229-8628-422d-8c4c-7ede411e107f>

Naciones Unidas et al. (2014): Sistema de Contabilidad Ambiental y Económica 2012 – Marco Central https://seea.un.org/sites/seea.un.org/files/seea_cf_final_sp.pdf.

OECD (2007). Organización para la Cooperación y el Desarrollo Económicos. Glossary of Statistical Terms. OECD Publishing.

https://ec.europa.eu/eurostat/ramon/coded_files/OECD_glossary_stat_terms.pdf.

Oras, K., Grüner, E., & Statistics Estonia. (2010). Final report Economy Wide Material Flow Account. CIRCABC. <https://circabc.europa.eu/sd/a/a5df9394-2c83-48a5-808b-6f39ee8083cd/08%20EE%20329%20MFA%20pdf.pdf>.

Orden HFP/1480/2021, de 28 de diciembre, por la que se aprueban el contenido y los plazos de presentación de la declaración de intercambio de bienes dentro de la Unión Europea (Intrastat) y se determinan las condiciones generales y el procedimiento para su presentación por medios telemáticos. BOE núm. 313, de 30 de diciembre de 2021, páginas 166939 a 166945.

<https://www.boe.es/eli/es/o/2021/12/28/hfp1480>

Resolución de 27 de enero de 2009, de la Presidencia de la Agencia Estatal de Administración Tributaria, para la elaboración de las estadísticas de intercambios de bienes entre Estados miembros (sistema Intrastat). BOE núm. 36, de 11 de febrero de 2009, páginas 14454 a 14488.

[https://www.boe.es/eli/es/res/2009/01/27/\(2\)](https://www.boe.es/eli/es/res/2009/01/27/(2)).

Resolución de 15 de diciembre de 2009, del Departamento de Aduanas e Impuestos Especiales de la Agencia Estatal de la Administración Tributaria, en la que se recoge las instrucciones para la formalización del Documento Único Administrativo (DUA). BOE núm. 314, de 30 de diciembre de 2009, páginas 111612 a 111911. [https://www.boe.es/eli/es/res/2009/12/15/\(4\)](https://www.boe.es/eli/es/res/2009/12/15/(4)).

Rousseeuw, P. & Hubert, M. (2011). Robust statistics for outlier detection. Wiley Interdisc. Rev.: Data Mining and Knowledge Discovery. 1. 73-79. DOI: 10.1002/widm.2.

<http://i2pc.es/coss/Docencia/SignalProcessingReviews/Rousseeuw2011.pdf>

Salgado, D., & Oancea, B. (2020). On new data sources for the production of official statistics. arXiv preprint arXiv:2003.06797. <https://arxiv.org/abs/2003.06797>.

Salgado, D. [aut, cre], Poch, J. [ctb], Barragan, S. [ctb] & Fernandez, J. [ctb] (2018): Paquete: fastReadfwf. Fast read of fwf files with arbitrary input schema. <https://github.com/david-salgado/fastReadfwf>

Schauberger, P. [aut, cre], Walker, A. [aut], Braglia, L. [ctb], Sturm, J. [ctb] et al. Paquete: openxlsx. Read, write and edit xlsx files. <https://CRAN.R-project.org/package=openxlsx>

SCN (2008): Sistema de Cuentas Nacionales 2008. Naciones Unidas, la Comisión Europea, la Organización de Cooperación y Desarrollo Económicos, el Fondo Monetario Internacional y el Grupo del Banco Mundial, 2008.

http://unstats.un.org/unsd/publication/seriesf/SeriesF_2Rev5e.pdf.

Unión Europea. Reglamento (CE) N.º 638/2004 del parlamento europeo y del consejo, de 31 de marzo de 2004, sobre estadísticas comunitarias relativas al comercio exterior con terceros países y por el que se deroga el Reglamento (CE) N.º 3330/91 del Consejo. Diario Oficial de la Unión Europea L 102/1, 7 de abril de 2004. <https://www.boe.es/doue/2004/102/L00001-00008.pdf>.

Unión Europea. Reglamento (CE) N.º 222/2009 del parlamento europeo y del consejo, de 11 de marzo de 2009, por el que se modifica el Reglamento (CE) N.º 638/2004 sobre las estadísticas comunitarias de intercambios de bienes entre Estados miembros. Diario Oficial de la Unión Europea L 87/160, 31 de marzo de 2009. <https://www.boe.es/doue/2009/087/L00160-00163.pdf#:~:text=REGLAMENTO%20%28CE%29%20No222%2F2009%20DEL%20PARLAMENTO%20EUROPEO%20Y%20DEL,estad%C3%ADsticas%20comunitarias%20deintercambios%20de%20bienes%20entre%20Estados%20miembros>.

Unión Europea. Reglamento (CE) N.º 471/2009 del parlamento europeo y del consejo, de 6 de mayo de 2009, sobre estadísticas comunitarias relativas al comercio exterior con terceros países y por el que se deroga el Reglamento (CE) N.º 1172/95 del Consejo. Diario Oficial de la Unión Europea L 152/23, 16 de junio de 2009. <https://www.boe.es/doue/2009/152/L00023-00029.pdf>.

Van der Loo, M. [cre, aut], de Jonge, E. [aut] & Hsieh, P. [ctb]. Paquete: validate. Data Validation Infrastructure. <https://CRAN.R-project.org/package=validate>

Wallgren, B., & Wallgren, A. (2014). Register-based statistics: Statistical methods for administrative data. John Wiley & Sons, Incorporated.

Anexo

A1. Clasificación de materiales en la CFM-TE según el tipo de flujo.

Extracción Nacional
1. Biomasa
1.1. Cultivos primarios
1.1.1. Cereales
1.1.2. Raíces y tubérculos
1.1.3. Cultivos azucareros
1.1.4. Leguminosas
1.1.5. Frutos secos
1.1.6. Cultivos oleaginosos
1.1.7. Hortalizas
1.1.8. Frutas
1.1.9. Fibras
1.1.10. Otros cultivos
1.2. Residuos de cultivos usados, cultivos forrajeros y biomasa pastada
1.2.1. Residuos de cultivos usados
1.2.1.1. Paja
1.2.1.2. Otros residuos de cultivos
1.2.2. Cultivos forrajeros y biomasa pastada
1.2.2.1. Cultivos forrajeros
1.2.2.2. Biomasa pastada
1.3. Madera
1.3.1. Madera en bruto
1.3.2. Leña y otras extracciones
M1.3. Incremento neto de masa forestal en pie
1.4. Captura de peces, plantas/animales acuáticos, caza y recolección
1.4.1. Peces capturados
1.4.2. Resto de animales y plantas acuáticos
1.4.3. Caza y recolección
2. Minerales metálicos (mineral en bruto)
2.1. Minerales de hierro
2.2. Minerales metálicos no férricos
2.2.1. Cobre
M.2.2.1. Cobre (contenido de metal)
2.2.2. Níquel
M.2.2.2. Níquel (contenido de metal)
2.2.3. Plomo
M.2.2.3. Plomo (contenido de metal)
2.2.4. Zinc
M.2.2.4. Zinc (contenido de metal)
2.2.5. Estaño
M.2.2.5. Estaño (contenido de metal)
2.2.6. Minerales de oro, plata y metales preciosos
M.2.2.6. Minerales de oro, plata y metales preciosos (contenido de metal)
2.2.7. Minerales de bauxita y otros minerales de aluminio
M.2.2.7. Minerales de bauxita y otros minerales de aluminio (contenido de metal)
2.2.8. Minerales de uranio y torio
M.2.2.8. Minerales de uranio y torio (contenido de metal)
2.2.9. Otros minerales metálicos
M.2.2.9. Otros minerales metálicos (contenido de metal)

3. Minerales no metálicos
3.1. Piedras de construcción u ornamentales (excl. Pizarra)
3.1.1. Arenisca
3.1.2. Basalto
3.1.3. Fonolita
3.1.4. Granito
3.1.5. Margas
3.1.6. Mármol
3.1.7. Ofita
3.1.8. Pórfidos
3.1.9. Serpentina y dunita
3.1.10. Otras
3.2. Creta y dolomía
3.2.1. Creta
3.2.2. Dolomía
3.3. Pizarra
3.4. Minerales para fertilizantes y químicos
3.4.1. Baritina
3.4.2. Cloruro potásico
3.4.3. Espato-f ácido
3.4.4. Espato-f metalúrgico
3.4.5. Espato-flúor cerámico
3.4.6. Óxidos de hierro
3.4.7. Pirita
3.4.8. Sales potásicas
3.5. Sal
3.5.1. Sal gema
3.5.2. Sal manantial
3.5.3. Sal marina
3.6. Piedra caliza y yeso
3.6.1. Piedra caliza
3.6.2. Yeso
3.7. Arcillas y caolín
3.7.1. Arcilla refractaria
3.7.2. Attapulgita
3.7.3. Bentonita
3.7.4. Caolín bruto
3.7.5. Caolín lavado
3.7.6. Sepiolita
3.7.7. Arcillas comunes
3.8. Arenas y grava
3.8.1. Arenas y grava
3.8.2. Sílice y arenas silíceas
3.9. Otros minerales
3.9.1. Cuarzita
3.9.2. Cuarzo
3.9.3. Esteatita
3.9.4. Feldespato
3.9.5. Glauberita
3.9.6. Magnesita
3.9.7. Mica
3.9.8. Piedra pómez
3.9.9. Thenardita

3.9.10. Trípoli
3.9.11. Otros
3.10. Tierra excavada, sólo si es usada
4. Combustibles fósiles
4.1. Carbón y otros recursos energéticos sólidos
4.1.1. Lignito
4.1.2. Hulla y antracita
4.1.3. Arenas de alquitrán y pizarras bituminosas
4.1.4. Turba
4.2. Recursos energéticos líquidos y gaseosos
4.2.1. Crudos de petróleo y gas natural licuado
4.2.2. Gas natural

Importaciones y Exportaciones Físicas
1. Biomasa y productos de la biomasa
1.1. Cultivos primarios y procesados
1.1.1. Cereales
1.1.2. Raíces y tubérculos
1.1.3. Cultivos azucareros
1.1.4. Leguminosas
1.1.5. Frutos secos
1.1.6. Cultivos oleaginosos
1.1.7. Hortalizas
1.1.8. Frutas
1.1.9. Fibras
1.1.10. Otros cultivos
1.2. Residuos de cultivos usados, cultivos forrajeros y biomasa pastada
1.2.1. Residuos de cultivos (usados), primarios y procesados
1.2.1.1. Paja
1.2.2.2. Otros residuos de cultivos
1.2.2. Cultivos forrajeros y biomasa pastada
1.2.2.1. Cultivos forrajeros
1.3. Madera y productos de la madera, primarios y procesados
1.3.1. Madera en bruto
1.3.2. Leña y otras extracciones
1.4. Capturas de peces y otras plantas/animales acuáticos, primarios y procesados
1.4.1. Capturas de peces
1.4.2. Resto de animales y plantas acuáticos
1.5. Animales vivos no incluidos en 1.4. y productos de animales
1.5.1. Animales vivos no incluidos en 1.4.
1.5.2. Carne y preparados de carne
1.5.3. Productos lácteos, huevos de aves y miel
1.5.4. Otros productos de animales (fibras de animales, piel, pelo, cueros, etc.)
1.6. Productos principalmente de biomasa
2. Minerales y concentrados de metal, primarios y procesados
2.1. Minerales y concentrados de hierro, hierro y acero
2.2. Minerales y concentrados metálicos no féreos
2.2.1. Cobre
2.2.2. Níquel
2.2.3. Plomo
2.2.4. Zinc

2.2.5. Estaño
2.2.6. Oro, plata, platino y otros metales preciosos
2.2.7. Bauxita y otros aluminios
2.2.8. Uranio y torio
2.2.9. Otros metales
2.3. Productos principalmente de metales
3. Minerales no metálicos, primarios y procesados
3.1. Piedras de construcción u ornamentales (excepto pizarra)
3.2. Creta y dolomita
3.3. Pizarra
3.4. Minerales para fertilizantes y químicos
3.5. Sal
3.6. Caliza y yeso
3.7. Arcilla y caolín
3.8. Arena y grava
3.9. Otros minerales no metálicos
3.10. Productos principalmente de minerales no metálicos
4. Combustibles fósiles, primarios y procesados
4.1. Carbón y otros recursos energéticos sólidos
4.1.1. Lignito
4.1.2. Hulla y antracita
4.1.3. Arenas de alquitrán y pizarras bituminosas
4.1.4. Turba
4.2. Recursos energéticos líquidos y gaseosos
4.2.1. Crudos de petróleo y gas natural licuado
4.2.2. Gas natural
4.2.3. Ajuste por el principio de residencia
4.2.3.1. Combustible para transporte terrestre
4.2.3.2. Combustible para transporte marítimo
4.2.3.3. Combustible para transporte aéreo
4.3. Productos principalmente de combustibles fósiles
5. Otros productos
6. Residuos importados para su tratamiento final y eliminación

Output nacional procesado a la naturaleza (OPN)
7.1 Emisiones al aire
7.1.1 Dióxido de Carbono (CO ₂)
7.1.1.1 Dióxido de Carbono (CO ₂) de la combustión de biomasa
7.1.1.2 Dióxido de Carbono (CO ₂) excluida la combustión de biomasa
7.1.2 Metano (CH ₄)
7.1.3 Monóxido de dinitrógeno (N ₂ O)
7.1.4 Óxidos de nitrógeno (NO _x)
7.1.5 Hidrofluorocarburos (HFCs)
7.1.6 Perfluorocarburos (PFCs)
7.1.7 Hexafluoruro de azufre
7.1.8 Monóxido de carbono (CO)
7.1.9 Compuestos Orgánicos Volátiles Distintos del Metano (COVDM)
7.1.A Dióxido de azufre (SO ₂)
7.1.B Amoníaco (NH ₃)
7.1.C Metales Pesados
7.1.D Contaminantes Orgánicos Persistentes (COP)
7.1.E Partículas (p.ej. PM ₁₀ , Polvo)

7.1.F Otras emisiones a la Atmósfera
7.2 Vertido de residuos en el medio ambiente
7.3 Emisiones al agua
7.3.1 Nitrógeno (N)
7.3.2 Fósforo (P)
7.3.3 Metales Pesados
7.3.4 Otras sustancias y materiales (orgánicos)
7.3.5 Vertido de materiales en el mar
7.4. Uso disipativo de productos
7.4.1. Abono orgánico (estiércol)
7.4.2. Fertilizantes sintéticos-inorgánicos
7.4.3. Lodos de depuradora
7.4.4. Compost
7.4.5. Fitosanitarios
7.4.6. Semillas
7.4.7. Sal y otros materiales de deshielo esparcidos por las carreteras (incluida la gravilla)
7.4.8. Disolventes, gas de la risa y otros
7.5. Pérdidas por disipación

Elementos de Balance
8.1 Elementos de Balance: Entrada
8.1.1 Oxígeno para los procesos de combustión (O ₂)
8.1.2 Oxígeno para la respiración de los seres humanos y el ganado; respiración bacteriana de los residuos sólidos y las aguas residuales (O ₂)
8.1.3 Nitrógeno para el proceso Haber-Bosch (N)
8.1.4 Necesidades de agua para la producción nacional de bebidas exportadas (H ₂ O)
8.2 Elementos de Balance: Salida
8.2.1 Vapor de agua de la combustión (H ₂ O)
8.2.1.1 Vapor de agua del contenido de humedad en los combustibles (H ₂ O)
8.2.1.2 Vapor de agua de los componentes de hidrógeno oxidado de los combustibles (H ₂ O)
8.2.2 Gases procedentes de la respiración de los seres humanos y del ganado (CO ₂ y H ₂ O), y de la respiración bacteriana de los residuos sólidos y de las aguas residuales (H ₂ O)
8.2.2.1 Dióxido de carbono (CO ₂)
8.2.2.2 Vapor de Agua (H ₂ O)
8.2.3 Pérdidas de agua contenida en la biomasa (H ₂ O)

A2. Diseño de registro ficheros datos comercio exterior AEAT.

Diseño de registro de los ficheros de datos de las Series Históricas Anuales (resumen).

Columna	Variable	Longitud
1-2	AÑO	2
4-5	PAÍS	2
7-14	POSICIÓN ESTADÍSTICA	8
16-27	PESO IMPORTACIÓN EN KG	12
29-43	VALOR IMPORTACIÓN CÉNTIMOS EURO	15
45-56	UNIDADES IMPORTACIÓN	12
58-69	PESO EXPORTACIÓN EN KG	12
72-85	VALOR EXPORTACIÓN CÉNTIMOS EURO	15
87-98	UNIDADES EXPORTACIÓN	12

Diseño de registro de los ficheros de datos de las Series Resúmenes Mensuales.

Columna	Variable	Longitud
1-2	AÑO	2
4-5	MES	2
7-8	PAÍS	2
10-17	POSICIÓN ESTADÍSTICA	8
19-30	PESO IMPORTACIÓN EN KG	12
32-46	VALOR IMPORTACIÓN CÉNTIMOS EURO	15
48-59	UNIDADES IMPORTACIÓN	12
61-72	PESO EXPORTACIÓN EN KG	12
74-88	VALOR EXPORTACIÓN CÉNTIMOS EURO	15
90-101	UNIDADES EXPORTACIÓN	12

Diseño de registro de los ficheros de datos de las Series Mensuales Máxima Desagregación.

Columna	Variable	Longitud
1-1	FLUJO	1
2-3	AÑO	2
4-5	MES	2
6-7	PROVINCIA DE ADUANA-RECINTO	2
8-19	FILLER	12
20-25	FECHA DE ADMISION DEL DOCUMENTO	6
26-37	POSICION ESTADISTICA	12
38-38	TIPO DECLARACIÓN I/D	1
39-46	CODIGOS ADICIONALES	8
47-65	FILLER	19
66-66	FILLER	1
67-69	PAIS ORIGEN/DESTINO	3
70-72	PAIS DE PROCEDENCIA/EXPEDICION	3
73-75	FILLER	3
76-77	PROVINCIA DE ORIGEN/DESTINO	2
78-79	FILLER	2
80-82	FILLER	3
83-84	REGIMEN ADUANERO SOLICITADO	2
85-86	REGIMEN ADUANERO PRECEDENTE	2
87-89	FILLER	3

90-104	PESO	15
105-119	UNIDADES	15
120-131	VALOR ESTADISTICO	12
132-143	VALOR FACTURA	12
144-146	PAIS DIVISA	3
147-158	FILLER	12
159-159	CONTENEDOR	1
160-164	REGIMEN DE TRANSPORTE	5
165-165	MODO DE TRANSPORTE EN FRONTERA	1
166-166	MODO DE TRANSPORTE INTERIOR	1
167-169	NACIONALIDAD MEDIO TRANSPORTE	3
170-170	FILLER	1
171-171	ZONA INTERCAMBIO (UE o TERCEROS)	1
172-172	FILLER	1
173-174	NATURALEZA DE LA TRANSACION	2
175-177	CONDICIONES DE ENTREGA	3
178-183	CONTINGENTE	6
184-189	PREFERENCIA ARANCELARIA	6
190-201	FLETE	12
202-224	FILLER	23
225-226	PROVINCIA DOMICILIO FISCAL	2

A3. Códigos de variables Series Mensuales Máxima Desagregación.

PROVINCIA	1	ALAVA/ARABA
ADUANA	2	ALBACETE
RECINTO	3	ALICANTE/ALACANT
PROVINCIA	4	ALMERIA
ORIGEN	5	AVILA
DESTINO	6	BADAJEZ
	7	BALEARES/BALEARIS
	8	BARCELONA
	9	BURGOS
	10	CACERES
	11	CADIZ
	12	CASTELLON/CASTELLÓ
	13	C. REAL
	14	CORDOBA
	15	A CORUÑA
	16	CUENCA
	17	GIRONA
	18	GRANADA
	19	GUADALAJARA
	20	GUIPÚZCOA/GIPUZKOA
	21	HUELVA
	22	HUESCA
	23	JAEN
	24	LEON
	25	LLEIDA
	26	LA RIOJA
	27	LUGO
	28	MADRID
	29	MALAGA
	30	MURCIA
	31	NAVARRA
	32	OURENSE
	33	ASTURIAS
	34	PALENCIA

	36	PONTEVEDRA
	37	SALAMANCA
	39	CANTABRIA
	40	SEGOVIA
	41	SEVILLA
	42	SORIA
	43	TARRAGONA
	44	TERUEL
	45	TOLEDO
	46	VALENCIA
	47	VALLADOLID
	48	VIZCAYA/BIZKAIA
	49	ZAMORA
	50	ZARAGOZA
PAIS ORIGEN DESTINO	AD	ANDORRA
PAIS PROCEDENCIA EXPEDICION	AE	EMIRATOS ÁRABES UNIDOS (Abu Dabi, Dubai, Sharya, Ayman, Umm al-Qaiwain, Ras al-Jaima y Fuyaira)
	AF	AFGANISTÁN
	AG	ANTIGUA Y BARBUDA
	AI	ANGUILA
	AL	ALBANIA
	AM	ARMENIA
	AN	ANTILLAS NEERLANDESAS
	AO	ANGOLA (incluido Cabinda)
	AQ	ANTÁRTIDA
	AR	ARGENTINA
	AS	SAMOA AMERICANA
	AT	AUSTRIA
	AU	AUSTRALIA
	AW	ARUBA
	AZ	AZERBAIYÁN
	BA	BOSNIA-HERZEGOVINA
	BB	BARBADOS
	BD	BANGLADESH
	BE	BÉLGICA
	BF	BURKINA FASO (Alto Volta)
	BG	BULGARIA
	BH	BAHRÉIN
	BI	BURUNDI
	BJ	BENÍN
	BL	SAN BARTHELEMY
	BM	BERMUDAS
	BN	BRUNÉI (BRUNÉI DARUSSALAM)
	BO	BOLIVIA

BQ	PAÍSES BAJOS (parte caribeña: Bonaire, San Eustaquio y Saba; antes en las Antillas Neerlandesas)
BR	BRASIL
BS	BAHAMAS
BT	BUTÁN
BV	BOUVET, ISLA
BW	BOTSUANA
BY	BIELORRUSIA (BELARÚS)
BZ	BELICE
CA	CANADÁ
CC	COCOS, ISLA DE (KEELING)
CD	CONGO, REPÚBLICA DEMOCRÁTICA DEL (Zaire)
CF	CENTROAFRICANA, REPÚBLICA
CG	CONGO
CH	SUIZA (Incluidos el territorio alemán de Büsingen y el municipio italiano de Campione de Italia)
CI	COSTA DE MARFIL
CK	COOK, ISLAS
CL	CHILE
CM	CAMERÚN
CN	CHINA
CO	COLOMBIA
CR	COSTA RICA
CU	CUBA
CV	CABO VERDE, REPÚBLICA DE
CW	CURAÇAO (antes en Antillas Neerlandesas)
CX	NAVIDAD, ISLA
CY	CHIPRE
CZ	CHECA, REPÚBLICA
DE	ALEMANIA (Incluida la Isla de Helgoland)
DJ	YIBUTI
DK	DINAMARCA
DM	DOMINICA
DO	DOMINICANA, REPÚBLICA
DZ	ARGELIA
EC	ECUADOR (incluidas las Islas Galápagos)
EE	ESTONIA
EG	EGIPTO
EH	SÁHARA OCCIDENTAL
ER	ERITREA
ES	ESPAÑA
ET	ETIOPÍA
FI	FINLANDIA (Incluidas las Islas Aland)
FJ	FIJI
FK	MALVINAS, ISLAS (FALKLANDS)
FM	MICRONESIA, FEDERACIÓN DE ESTADOS DE (Yap, Kosrae, Truk, Pohnpei)
FO	FEROE, ISLAS
FR	FRANCIA (Incluidos los departamentos franceses de

	ultramar: Reunión, Guadalupe, Martinica y Guayana Francesa)
GA	GABÓN
GB	REINO UNIDO (Gran Bretaña e Irlanda del Norte)
GD	GRANADA (incluidas las Islas Granadinas del Sur)
GE	GEORGIA
GH	GHANA
GI	GIBRALTAR
GL	GROENLANDIA
GM	GAMBIA
GN	GUINEA
GQ	GUINEA ECUATORIAL
GR	GRECIA
GS	GEORGIA DEL SUR Y LAS ISLAS SANDWICH DEL SUR
GT	GUATEMALA
GU	GUAM
GW	GUINEA-BISSAU
GY	GUYANA
HK	HONG-KONG
HM	HEARD Y MCDONALD, ISLAS
HN	HONDURAS (incluidas Islas del Cisne)
HR	CROACIA
HT	HAITÍ
HU	HUNGRÍA
ID	INDONESIA
IE	IRLANDA
IL	ISRAEL
IN	INDIA
IO	OCÉANO ÍNDICO, TERRITORIO BRITÁNICO DEL (Archipiélago de Chagos)
IQ	IRAQ
IR	IRÁN
IS	ISLANDIA
IT	ITALIA (Incluido Livigno)
JM	JAMAICA
JO	JORDANIA
JP	JAPÓN
KE	KENIA
KG	KIRGUISTÁN
KH	CAMBOYA
KI	KIRIBATI
KM	COMORAS (Gran Comora, Anjouan y Mohéli)
KN	SAN CRISTÓBAL Y NIEVES (Saint Kitts y Nevis)
KP	COREA DEL NORTE (República Popular Democrática de Corea)
KR	COREA DEL SUR (República de Corea)
KW	KUWAIT
KY	CAIMÁN, ISLAS

KZ	KAZAJSTÁN
LA	LAOS (LAO)
LB	LÍBANO
LC	SANTA LUCÍA
LI	LIECHTENSTEIN
LK	SRI LANKA
LR	LIBERIA
LS	LESOTHO
LT	LITUANIA
LU	LUXEMBURGO
LV	LETONIA
LY	LIBIA
MA	MARRUECOS
MD	MOLDAVIA
ME	MONTENEGRO
MG	MADAGASCAR
MH	MARSHALL, ISLAS
MK	MACEDONIA (Antigua República Yugoslava)
ML	MALI
MM	MYANMAR (Antigua Birmania)
MN	MONGOLIA
MO	MACAO
MP	MARIANAS DEL NORTE, ISLAS
MR	MAURITANIA
MS	MONTSERRAT
MT	MALTA (Incluidos Gozo y Comino)
MU	MAURICIO (Isla Mauricio, Isla Rodrígues, Islas Agalega y Cargados, Carajos Shoals (Islas San Brandón))
MV	MALDIVAS
MW	MALAWI
MX	MÉXICO
MY	MALASIA ((Malasia Peninsular y Malasia Oriental: Sarawak, Sabah y Labuán)
MZ	MOZAMBIQUE
NA	NAMIBIA
NC	NUEVA CALEDONIA (Incluidas las islas Lealtad: Maré, Lifou y Ouvéa)
NE	NÍGER
NF	NORFOLK, ISLA
NG	NIGERIA
NI	NICARAGUA (incluidas las Islas del Maíz)
NL	PAÍSES BAJOS (parte europea)
NO	NORUEGA (Incluidos la Isla Jan Mayen y el archipiélago Svalbard)
NP	NEPAL
NR	NAURU
NZ	NUEVA ZELANDA
OM	OMÁN
PA	PANAMÁ (incluida la antigua Zona del Canal)

PE	PERÚ
PF	POLINESIA FRANCESA (Islas Marquesas, Isla de la Sociedad (incluido Tahiti), Islas Gambier, Islas Tuamotú e islas Australes incluida la Isla de Clipperton)
PG	PAPÚA NUEVA GUINEA
PH	FILIPINAS
PK	PAKISTÁN
PL	POLONIA
PM	SAN PEDRO Y MIQUELÓN
PN	PITCAIRN (incluidas las Islas Henderson, Ducie y Oeno)
PS	TERRITORIO PALESTINO OCUPADO (Cisjordania y Franja de Gaza)
PT	PORTUGAL (Incluidos los Archipiélagos de las Azores y de Madeira)
PW	PALAU
PY	PARAGUAY
QA	QATAR
QP	Elementos de código asignados por el usuario
QQ	Elementos de código asignados por el usuario
QR	Elementos de código asignados por el usuario
QS	Elementos de código asignados por el usuario
QU	OTROS PAÍSES O TERRITORIOS NO RELACIONADOS
QV	Elementos de código asignados por el usuario
QW	Elementos de código asignados por el usuario
QY	Elementos de código asignados por el usuario
QZ	Elementos de código asignados por el usuario
RO	RUMANÍA
RU	RUSIA (FEDERACIÓN DE)
RW	RUANDA
SA	ARABIA SAUDÍ
SB	SALOMÓN, ISLAS
SC	SEYCHELLES
SD	SUDÁN
SE	SUECIA
SG	SINGAPUR
SH	SANTA ELENA (Incluidos la Isla de la Ascensión y el Archipiélago Tristán da Cunha)
SI	ESLOVENIA
SK	ESLOVAQUIA
SL	SIERRA LEONA
SM	SAN MARINO
SN	SENEGAL
SO	SOMALIA
SR	SURINAM
SS	SUDÁN DEL SUR
ST	SANTO TOMÉ Y PRÍNCIPE
SV	SALVADOR, EL
SX	SAN MARTÍN (parte meridional; antes en Antillas Neerlandesas)

SY	SIRIA (REPÚBLICA ÁRABE)
SZ	SUAZILANDIA
TC	TURCAS Y CAICOS, ISLAS
TD	CHAD
TF	TIERRAS AUSTRALES FRANCESAS
TG	TOGO
TH	TAILANDIA
TJ	TAYIKISTÁN
TK	TOKELAU, ISLAS
TL	TIMOR LESTE
TM	TURKMENISTÁN
TN	TÚNEZ
TO	TONGA
TR	TURQUÍA
TT	TRINIDAD Y TOBAGO
TV	TUVALU
TW	TAIWÁN
TZ	TANZANIA (REPÚBLICA UNIDA DE) (Tanganica e islas de Zanzibar y Pemba)
UA	UCRANIA
UG	UGANDA
UM	MENORES ALEJADAS DE LOS EE. UU, ISLAS (Baker, Howland, Jarvis, Johnston, Kingman Reef, Midway, Navassa, Palmira y Wake)
UN	NIUE, ISLA
US	ESTADOS UNIDOS DE AMÉRICA
UY	URUGUAY
UZ	UZBEKISTÁN
VA	VATICANO, CIUDAD DEL (Santa Sede)
VC	SAN VICENTE Y LAS GRANADINAS
VE	VENEZUELA
VG	VÍRGENES BRITÁNICAS, ISLAS
VI	VÍRGENES DE LOS EE.UU, ISLAS
VN	VIETNAM
VU	VANUATU
WF	WALLIS Y FUTUNA, ISLAS (incluida la Isla Alofi)
WS	SAMOA (Samoa Occidental)
XC	Elementos de código asignados por el usuario
XI	Elementos de código asignados por el usuario
XK	Elementos de código asignados por el usuario
XL	Elementos de código asignados por el usuario
XS	Elementos de código asignados por el usuario
YE	YEMEN (Yemen del Norte y Yemen del Sur)
YT	MAYOTTE (Gran Tierra y Pamandzi)
ZA	SUDÁFRICA
ZM	ZAMBIA
ZW	ZIMBABWE
REGIMEN ADUANERO	01: Despacho a libre práctica de mercancías con reexpedición simultánea a

<p>SOLICITADO</p> <p>REGIMEN ADUANERO PRECEDENTE</p>	<p>territorio no IVA.</p> <p>07: Despacho a libre práctica de mercancías con inclusión simultánea.</p> <p>40: Despacho a libre práctica con despacho simultáneo a consumo de mercancías.</p> <p>41: Despacho a consumo con despacho a libre práctica simultáneo de mercancías incluidas en el régimen de perfeccionamiento activo (sistema de reintegro) (a).</p> <p>42: Despacho a consumo con despacho a libre práctica simultáneo de mercancías objeto de entrega exenta del IVA a otro Estado Miembro.</p> <p>43: Despacho a consumo con despacho a libre práctica simultáneo de mercancías en el marco de la aplicación.</p> <p>44: Destino final.</p> <p>45: Despacho a libre práctica y entrada parcial para consumo de mercancías sujetas al IVA o a impuestos especiales e inclusión de las mismas en un depósito.</p> <p>46: Importación de productos transformados obtenidos a partir de mercancías equivalentes en el marco del régimen de perfeccionamiento pasivo antes de la exportación.</p> <p>48: Despacho a consumo de productos de sustitución en el marco del régimen de perfeccionamiento pasivo, intercambios estándar con importación anticipada.</p> <p>49: Despacho a consumo en parte del TAU donde sea de aplicación la Directiva IVA de mercancía comunitaria procedente de otra parte del territorio donde no sea de aplicación.</p> <p>51: Inclusión de mercancías en el régimen de perfeccionamiento activo.</p> <p>53: Inclusión de mercancías en el régimen de importación temporal.</p> <p>54: Perfeccionamiento activo en otro Estado miembro (sin que las mercancías se hayan despachado a libre práctica en él) (a).</p> <p>61: Reimportación con despacho simultáneo a libre práctica y despacho a consumo de mercancías.</p> <p>63: Reimportación con despacho a consumo y despacho a libre práctica simultáneo de mercancías objeto de una entrega exenta del IVA a otro Estado miembro.</p> <p>68: Reimportación con entrada parcial para consumo y despacho a libre práctica simultáneo e inclusión en un régimen de depósito distinto del</p>
--	---

	<p>régimen de depósito aduanero.</p> <p>10: Exportación definitiva.</p> <p>11: Exportación de productos transformados obtenidos a partir de mercancías equivalentes.</p> <p>21: Exportación temporal en el marco del régimen de perfeccionamiento pasivo, si no está cubierta por el código 22.</p> <p>22: Exportación temporal distinta de la contemplada para los códigos 21 y 23.</p> <p>23: Exportación temporal con vistas a una reintroducción posterior sin transformar.</p> <p>31: Reexportación.</p> <p>71: Inclusión en el régimen de depósito aduanero.</p> <p>73: Vinculación de bienes al depósito distinto del aduanero.</p> <p>76: Inclusión de mercancías de la Unión en el régimen de depósito aduanero de conformidad con el artículo 237, apartado 2, del Código.</p> <p>77. Fabricación de mercancías de la Unión bajo la vigilancia aduanera de las autoridades aduaneras y sometida a controles aduaneros previamente a la exportación y el pago de restituciones por exportación.</p> <p>78: Inclusión de mercancías en una zona franca.</p> <p>91: Inclusión en régimen de transformación bajo control aduanero.</p> <p>92: Transformación bajo control aduanero en otro Estado miembro (sin que las mercancías se hayan despachado a libre práctica en él).</p> <p>95: Inclusión de mercancías de la Unión en un régimen de depósito distinto de un régimen de depósito aduanero cuando no se haya pagado el IVA ni, en su caso, impuestos especiales.</p> <p>96. Inclusión de mercancías de la Unión en un régimen de depósito distinto de un régimen de depósito aduanero cuando no se haya pagado el IVA ni, en su caso, impuestos especiales y esté suspendido el pago de los demás impuestos.</p> <p>80: Despacho a consumo (tributos REF) de mercancía comunitaria previamente vinculada a un depósito REF.</p>
<p>MODO TRANSPORTE FRONTERA</p> <p>MODO TRANSPORTE INTERIOR</p>	<p>1: Transporte marítimo</p> <p>2: Transporte por ferrocarril</p> <p>3: Transporte por carretera</p> <p>4: Transporte por vía aérea</p> <p>5: Envíos postales</p>

	<p>7: Instalaciones fijas de transporte (oleoductos, gaseoductos, red eléctrica, etc.)</p> <p>8: Transporte por navegación interior</p> <p>9: Propulsión propia</p>
NATURALEZA TRANSACCION	<p>Combinación códigos columna A y columna B.</p> <p>COLUMNA A</p> <ol style="list-style-type: none"> 1. Transacciones que supongan un traspaso de propiedad real o previsto de residentes a no residentes a cambio de una compensación financiera o de otro tipo (salvo las transacciones indicadas en los puntos 2, 7 y 8). 2. Devolución y sustitución gratuita de mercancías después del registro de la transacción original. 3. Transacciones que supongan un traspaso de propiedad sin compensación financiera o en especie (por ejemplo, los envíos de ayuda). 4. Operaciones con vistas a un perfeccionamiento (2) bajo contrato (no se traspasa la propiedad al contratista). 5. Operaciones posteriores al perfeccionamiento bajo contrato (no se traspasa la propiedad al contratista). 6. Transacciones específicas registradas con fines nacionales. 7. Operaciones en el marco de proyectos comunes de defensa u otros programas intergubernamentales de producción conjunta. 8. Transacciones que impliquen el suministro de materiales de construcción y de equipo técnico en el marco de un contrato general de construcción o ingeniería civil para el que no sea necesaria una facturación separada, sino que se establezca una factura para la totalidad del contrato. 9. Otras transacciones que no puedan clasificarse en otros códigos. <p>COLUMNA B</p> <ol style="list-style-type: none"> 1. Compraventa en firme 2. Entrega para venta a la vista o de prueba, para consignación o con la mediación de un agente comisionado 3. Trueque (compensación en especie) 4. Arrendamiento financiero (alquiler-compra) (1) 9. Otras 1. Devolución de mercancías 2. Sustitución de mercancías devueltas 3. Sustitución (por ejemplo, en garantía) de mercancías no devueltas

	<p>9. Otras</p> <p>1. Mercancías suministradas en el marco de programas de ayuda promovidos o financiados parcial o totalmente por la Unión Europea.</p> <p>2. Otras ayudas gubernamentales.</p> <p>3. Otras ayudas (privadas, organizaciones no gubernamentales, etc.)</p> <p>1. Mercancías destinadas a regresar al Estado miembro de expedición inicial</p> <p>2. Mercancías no destinadas a regresar al Estado miembro de expedición inicial</p> <p>4. Reparación o mantenimiento a título gratuito</p> <p>5. Reparación o mantenimiento a título oneroso.</p> <p>1. Mercancías que regresan al Estado miembro de expedición</p> <p>2. Mercancías que no regresan al Estado miembro de expedición</p> <p>4. Reparación o mantenimiento a título gratuito</p> <p>5. Reparación o mantenimiento a título oneroso</p> <p>1. Alquiler, préstamo o arrendamiento operativo</p> <p>2. Otros usos temporales</p> <p>1. Alquiler, préstamo y arrendamiento operativo por un período superior a veinticuatro meses</p> <p>9. Otras</p>
CONDICIONES ENTREGA	<p>EXW Franco fabrica</p> <p>FCA Franco transportista</p> <p>FAS Franco al costado del buque</p> <p>FOB Franco a bordo</p> <p>CFR Coste y flete (C & F)</p> <p>CIF Coste, seguro y flete (CIF)</p> <p>CPT Porte pagado hasta</p> <p>CIP Porte y seguro pagado hasta</p> <p>DAF Franco frontera</p> <p>DES Franco ex ship</p> <p>DEQ Franco muelle</p> <p>DDU Franco sin despachar en aduana</p> <p>DDP Franco despachado en aduana</p> <p>XXX Condiciones de entrega distintas de las anteriores</p>

FUENTE	<p>https://www.agenciatributaria.es/static_files/AEAT/Aduanas/Contenidos_Privados/Estadisticas_Comercio_Exterior/Tabla_codigos/Fich_cod.xlsx</p> <p><u>RESOLUCIÓN DEL DEPARTAMENTO DE ADUANAS E IMPUESTOS ESPECIALES DE 15 DE DICIEMBRE DE 2009 (BOE 30 DE DICIEMBRE DE 2009)</u></p> <p><u>RESOLUCIÓN DEL DEPARTAMENTO DE ADUANAS E IMPUESTOS ESPECIALES DE 29 MARZO DE 2022 (BOE 6 ABRIL DE 2022)</u></p>
--------	---

A4. Edits y resumen comext_Schema.

Edits

edits for physical imports and exports

rules:

-

```
expr: if (FLUJO == "E"){REGIMEN_ADUANERO_SOLICITADO %in% c("10", "11", "21", "22", "23", "31", "NA")}
```

name: regimen_export

label: regimen exportacion

description: |

Si el flujo es exportación los códigos del régimen aduanero deben ser los indicados.

-

```
expr: if (FLUJO == "I"){REGIMEN_ADUANERO_SOLICITADO %in% c("01", "07", "40", "41", "43", "44", "45", "46", "48", "49", "51", "53", "54", "61", "63", "68", "91", "NA")}
```

name: regimen_import

label: regimen importacion

description: |

Si el flujo es importación los códigos del régimen aduanero deben ser los indicados.

Resumen comext_Schema

variable	width	initial Pos	final Pos	type	valueRegEx
FLUJO	1	1	1	char	I E
ANO	2	2	3	char	[0-2][0-9]
MES	2	4	5	char	[0-1][0-9]
PROVINCIA_ADUANA_RECINTO	2	6	7	char	[0-5][0-9]
FECHA_ADMISION_DOCUMENTO	6	20	25	char	(2[0-1](0[1-9] 1[0-2])(0[1-9] [1-2][0-9] 3[0-1]))
POSICION_ESTADISTICA	12	26	37	char	[0-9]{8,10}
TIPO_DECLARACION	1	38	38	char	I D
CODIGOS_ADICIONALES	8	39	46	char	(0[AEIVG][0-9] R001 R002 R099 ^\\$)

					AD AE AF AG AI AL AM AN AO AQ
					AR AS AT AU AW AZ BA BB BD BE
					BF BG BH BI BJ BL BM BN BO BQ
					BR BS BT BV BW BY BZ CA CC CD
					CF CG CH CI CK CL CM CN CO CR
					CU CV CW CX CY CZ DE DJ DK DM
					DO DZ EC EE EG EH ER ES ET FI FJ
					FK FM FO FR GA GB GD GE GH GI
					GL GM GN GQ GR GS GT GU GW
					GY HK HM HN HR HT HU ID IE IL
					IN IO IQ IR IS IT JM JO JP KE KG
					KH KI KM KN KP KR KW KY KZ LA
					LB LC LI LK LR LS LT LU LV LY MA
					MD MG MH MK ML MM MN MO
					MP MR MS MT MU MV MW MX
					MY MZ NA NC NE NF NG NI NL
					NO NP NR NU NZ OM PA PE PF
					PG PH PK PL PM PN PS PT PW
					PY QA QP QQ QZ RO RU RW
PAIS_ORIGEN_DESTINO	3	67	69	char	SA SB SC SD SE SG SH SI S

					K SL SM SN SO SR SS ST SV S X SY SZ TC TD TF TG TH TJ T K TM TN TO TL TR TT TV T W TZ UA UG UM US UY UZ V A VC VE VG VI VN VU WF W S XC XI XL YE YT XS XK M E ZA ZM ZW QR QS QU QW QV QY
PROVINCIA_ORIGEN_DEST INO	2	76	77	char	([0-5][0-9] ^\$)
REGIMEN_ADUANERO_SO LICITADO	2	83	84	char	00 01 07 10 11 21 22 23 31 40 41 42 43 44 45 46 48 49 51 53 54 61 63 68 71 73 76 77 78 80 91 95 96 ^\$
REGIMEN_ADUANERO_PRE CEDENTE	2	85	86	char	00 01 07 10 11 21 22 23 31 40 41 42 43 44 45 46 48 49 51 53 54 61 63 68 71 73 76 77 78 80 91 95 96 ^\$
PESO	15	90	104	num	[0-9]{1,15}
UNIDADES	15	105	119	num	[0-9]{1,15}
VALOR_ESTADISTICO	12	120	131	num	[0-9]{1,12}
VALOR_FACTURA	12	132	143	num	[0-9]{1,12}
CONTENEDOR	1	159	159	char	0 ^\$
REGIMEN_TRANSPORTE	5	160	164	char	A B X Y Z ^\$
MODO_TRANSPORTE_FRO NTERA	1	165	165	char	[1-9] ^\$
MODO_TRANSPORTE_INTE RIOR	1	166	166	char	[1-9] ^\$
ZONA_INTERCAMBIO_UE_ 3os	1	171	171	char	C T

NATURALEZA_TRANSACCI ON	2	173	174	char	(1[1-4]) 19 (2[1-3]) 29 (3[1-3]) (4[1-2]) (4[4-5]) (5[1-2]) (5[4-5]) (6[1-2]) (6[4-5]) 7 8 9 99 15 9
CONDICIONES_ENTREGA	3	175	177	char	EXW FCA FAS FOB CFR CIF CPT CIP DAF DES DEQ DDU DDP XXX DAP DAT DPU