

Memoria

Análisis exploratorio de datos

Julia María Martínez Tapia

www.linkedin.com/in/juliamariamartineztapia

Evolución del resultado contable y la base imponible del Impuesto sobre Sociedades en España (EDA I)

Datos

Los datos utilizados para el análisis se han extraído de las estadísticas del Impuesto sobre Sociedades elaboradas por el servicio de estudios tributarios y estadísticos de la Agencia Estatal de la Administración Tributaria (AEAT).

En particular, se ha utilizado el informe anual de recaudación tributaria para el año 2019 y se han analizado datos del periodo impositivo 2005 al 2018, ambos inclusive.

Se puede acceder a los datos utilizados pinchando [aquí](#)

Recursos utilizados

- Notebook de Jupyter.
- Numpy
- Pandas
- Matplotlib
- Seaborn
- Plotly

Cuestiones destacables

Me he encontrado con obstáculos a la hora de analizar las estadísticas de la Agencia Tributaria.

En particular, he detectado ciertas inconsistencias para las que no encuentro explicación en la metodología y que me han obligado a ceñir el alcance del análisis a las compañías con base imponible positiva.

Análisis de los ajustes fiscales del Impuesto sobre Sociedades en España (EDA II, EDA III, EDA IV)

Datos

Los datos utilizados para el análisis se han extraído de las estadísticas relativas al Impuesto sobre Sociedades elaboradas por el servicio de estudios tributarios y estadísticos de la Agencia Estatal de la Administración Tributaria (AEAT).

En particular, se ha utilizado la estadística de cuentas anuales no consolidadas del Impuesto sobre Sociedades para el 2018, sección ajustes al resultado contable de sociedades no financieras (excluyendo aseguradoras, entidades de crédito y IIC), apartado relativo a las sociedades no financieras por Dimensión de empresa y Sector (excluyendo entidades transparentes, exentas y cooperativas).

Se puede acceder a los datos utilizados en el presente estudio pinchando [aquí](#).

La actividad de cada sociedad se obtiene de la actividad principal que figura en el Censo de grandes empresas, de la actividad declarada en el modelo anual de IVA (modelo 390) y de la información de epígrafes del impuesto de actividades económicas, la cual determina el sector al que pertenece. En función de esta actividad las sociedades se clasifican en 10 agrupaciones sectoriales en términos de la nueva nomenclatura de actividad de la CNAE 2009.

- Agricultura, Ganadería, Silvicultura y Pesca (A G,S y P)
- Industria extractiva, Energía y Agua (Indus Extract, E y A)
- Industria (Industria)
- Construcción y Actividades inmobiliarias (Constr, Inmob)
- Comercio, Reparaciones y Transporte (Comercio, rep, transp)
- Información y Comunicaciones (Info, comunic)
- Entidades financieras y aseguradoras (Finac, asegur)
- Servicios a Empresas (Serv empresas)
- Servicios sociales (Serv sociales)
- Otros servicios personales y de ocio (Otros serv Ocio)

En el siguiente [enlace](#) se puede encontrar el desglose de las anteriores agrupaciones sectoriales.

Precisamente por este nivel de desglose de las agrupaciones sectoriales se ha preferido utilizar las estadísticas de cuentas anuales no consolidadas en vez de las estadísticas de cuentas anuales consolidadas, que desglosan los datos en cinco agrupaciones sectoriales. Dado que el presente análisis se limita a los ajustes fiscales practicados a

la base imponible, se considera que las potenciales diferencias deberían ser inmateriales.

Recursos utilizados

- Notebook de Jupyter.
- Numpy
- Pandas
- Matplotlib
- Seaborn
- Plotly

Cuestiones destacables

El trabajo desarrollado en el EDA IV no está terminado. El objetivo es crear una página en streamlit donde sea posible:

- Seleccionar un sector económico, y
- subir los datos de una empresa concreta,

para obtener un análisis de las diferencias entre los datos de la empresa y el promedio del sector.