Modelo Básico 25 10 2023

## Introducción

En este documento hay una primera explicación sobre las variables que Deoleo nos ha facilitado. En seguida hay las graficas de correlaciones de estas variables, y al final los outputs de modelos y Regresión.

Muchas de las variables disponibles se han desplazado (shifted) porque **la variable objetivo**, es decir, el precio del aceite de oliva ponderado por los volúmenes, **tiene una correlación más alta con un rezago** .

Con estas Variables ha sido entrenado un modelo de regresión lineal multivariada usando el algoritmo de OLS y el **R2 del modelo es = 0.508**

Al final ha sido calculado un Backtesting out of sample: que ha generado los siguientes resultados :

|  |  |  |
| --- | --- | --- |
|  | **MAPE** | **MSFE** |
| **1** | 26.25395 | 0.98547263 |
| **2** | 26.75042 | 1.0262363 |
| **3** | 27.21195 | 1.06316232 |
| **4** | 27.66905 | 1.09934695 |
| **5** | 28.09845 | 1.13360668 |
| **6** | 28.48876 | 1.1645186 |

Con un error de previsión del 26% por un mes vista y del 28% por 6 meses vista.

Como sugieren los resultados esto modelo no es muy robusto y tiene una capacidad previsional limitada.

### Variables usadas:

Este documento muestra las variables útiles facilitadas por Deoleo en formato excel, que se reducen al conjunto de variables físicas tales como Existencias de importación y exportación, etc.

El otro conjunto de variables excel está formado por precios de productos que se correlacionan casi perfectamente con la variable objetivo, es decir, el precio mensual ponderado del aceite de oliva (variable VIRGEN\_EXTRA\_EUR\_kg) y, por tanto, no pueden utilizarse como variables explicativas.

Por último, las variables físicas que se muestran tienen un historial que comienza en octubre de 2005

A continuación, se describen brevemente estas variables con el suporte de unas graficas:

A white and blue document with black text

Description automatically generated

Hay otros datos que nos ha señalado Deoleo, que pueden incluirse en el análisis en una fase posterior y que son :

### Variables “Deoleo” que falta de incluir

1. Datos meteroelógicos, para los que se acaban de indicar varios enlaces, pero son datos con otra fragmentación a nivel territorial, y habría que establecer un rascador de datos para continuar con la fase de ingesta de datos

2. Datos con granularidad anual sobre el consumo mundial de importación y exportación dividido en macrozonas, en formato pdf. Estos datos pueden y serán incluidos en un análisis posterior.

3. Datos de Stock Italia no han sido incluidos en el análisis porque los datos que nos han pasado tienen un histórico desde Abril 2021, sin embargo online hay datos desde el 2010, haría falta descargarlos y organizarlos para incluirlos.

### Matriz de correlación

En seguida hay la matriz de correlación de estas variables:

A screenshot of a computer screen

Description automatically generated

Después del análisis exploratorio de las variables, que se incluye a continuación por completitud, se calcularon retrasos para varias variables y se creó la nueva variable ''PRODUCTION\_HARVEST\_OLD', que representa la cantidad total recolectada durante el año anterior."

Como se puede observar en los gráficos de la variable 'PRODUCTION\_HARVEST', la cosecha del año anterior está correlacionada positivamente con el precio del aceite de oliva virgen extra, mientras que la cosecha actual está correlacionada negativamente

Es interesante considerar este factor porque sabemos que la cantidad producida el año anterior influye en la producción del año siguiente y, posiblemente, también en el precio.

A continuación, incluyo una tabla en la que se registran todos los desplazamientos (shifting) que se han realizado para cada variable

Transformaciones aplicadas = IMPORTS = 6 substituida

|  |  |
| --- | --- |
| **Variable** | **Lags** |
| PRODUCTION HARVEST | -3 |
| PRODUCTION HARVEST\_OLD | 12 |
| EXTERNAL DEMAND | 13 |
| INTERNAL DEMAND | 13 |
| EXPORT | 12 |
| PRODUCTION | 6 |

Después de las transformaciones de las variables considerando los rezagos la correlación ha subido por muchas variables obteniendo esta matriz

A blue and pink squares with white text

Description automatically generated

Se puede notar que los valores de correlación para todas las variables a las que se les aplicaron desplazamientos suben. Un próximo paso seria analizar la autocorrelación de estas variables que han sido transformadas para evaluar si se pueden dejar en el modelo ambas variables, antes y después de la transformación.

### Graficas De variable

Graficas de Variable IMPORTS

|  |  |
| --- | --- |
|  |  |
|  | |
| A graph with red and blue lines  Description automatically generated | |
|  | |
|  | |

Graficas de Variable INNER\_CONS

|  |  |
| --- | --- |
|  |  |
|  | |
|  | |
|  | |
|  | |

Graficas de Variable PRODUCTION\_HARVEST

|  |  |
| --- | --- |
|  |  |
|  | |
|  | |
|  | |
|  | |

Graficas de Variable EXTERNAL\_DEMAND

|  |  |
| --- | --- |
|  |  |
|  | |
|  | |
|  | |
|  | |

Graficas de Variable EXIS\_INIC

|  |  |
| --- | --- |
|  |  |
|  | |
|  | |
|  | |
|  | |

Graficas de Variable EXPORTS

|  |  |
| --- | --- |
|  |  |
|  | |
|  | |
|  | |
|  | |

Graficas de Variable INTERNAL DEMAND

|  |  |
| --- | --- |
|  |  |
|  | |
|  | |
|  | |
|  | |

Graficas de Variable PRODUCTION

|  |  |
| --- | --- |
|  |  |
|  | |
|  | |
|  | |
|  | |

## Output modelo basico regression simple

A screenshot of a computer screen

Description automatically generated

El major modelo que se ha encontrado es con estas variable, solo se han sacado desde el análisis inicial de las correlaciones la variable de producción, poque teniendo muchos valores = 0 no presentaba una relación lineal con la variable objetivo y la variable “External\_Demand” que sería la demanda total porque se suman los consumos internos a las exportaciones, esta variable estaba muy correlacionada a la variable de export que sigue en el modelo.

Aquí

El R2 = 508

Aqui pongo los MAPE y los MSFE del modelo obtenidos con un backtesting out of sample dejando fuera de la muestra 50 observaciones, adonde la metodología elegida es un expanding.

A table with numbers and numbers

Description automatically generated