

Obtención de un modelo econométrico para explicar el consumo de alcohol diario en Europa

MEMORIA PRESENTADA POR:
Mariia Kuts

GRADO DE ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESAS

Convocatoria de defensa: Septiembre de 2018

*En memoria de mi padre que fue quien hizo que
esto fuera posible y el que siempre estaba allí.
¡Gracias!*

Índice

Introducción y Objetivos	4
Motivación	4
Hipótesis	5
Objetivos	6
Metodología y Técnicas	7
Metodología	7
Técnicas	10
Análisis univariante.....	10
Análisis Bivariante.....	13
Análisis Multivariante.....	15
Predicciones.....	20
Cálculos, resultados y explotación	23
Análisis Univariante	23
Resumen estadístico y cajas y bigotes.....	23
Análisis Bivariante	35
Gráfico x e y, modelo de regresión simple y valor R.....	35
Multicolinealidad.....	46
Análisis Multivariado	52
Modelo teórico propuesto.....	52
Interpretación de los parámetros.....	52
Modelo ajustado.....	53
Validación del modelo.....	55
Normalidad de los residuos.....	58
Heterocedasticidad.....	60
Autocorrelación.....	62
Puntos anómalos y puntos influyentes.....	64
Solución modelo.....	66
Predicciones	67
Conclusiones y Futuras líneas de investigación	70
Conclusiones	70
Futuras líneas de investigación	70
Bibliografía	71

Introducción y Objetivos

Motivación

El consumo de alcohol está presente todos los días en nuestras vidas, pero hay personas que están atrapados en un círculo vicioso de Alcohol a grandes cantidades y a lo largo de cada uno de los días de su vida, ¿una vida muy triste?

Hoy en día, incluso podemos llegar a ver jóvenes de 14, 15 años o incluso más jóvenes iniciarse en el consumo, aunque sea solo los fines de semana, pero esto es muy perjudicial para su salud, ya que su cuerpo todavía está en fase de crecimiento. Una iniciación tan temprana puede generar dependencia e incluso dentro de unos años podrían considerarse alcohólicos.

La decisión de realizar este proyecto sobre el Consumo de alcohol diario fue tomada ya que es un tema que afecta a todos, porque, aunque no nos damos cuenta, pero muchas de las personas que consumen a diario es por evadirse de sus problemas principalmente, incluso lo consideran su forma de vivir. Pero no debería de ser así, por esta razón se decide buscar una serie de variables que a nuestra consideración puede afectar a ese consumo abusivo.

Hipòtesis

- Explicar la variable Consumo de alcohol cada día a partir de otras variables explicativas.
- Verificar que el porcentaje de víctimas de homicidios intencionados esté relacionado con el consumo de alcohol diario, y por lo tanto si los homicidios subieran también subiría el consumo de alcohol.
- Comprobar que la variable población esté relacionada con el consumo de alcohol, por consiguiente, si sube la población debería de subir el consumo de alcohol.
- Determinar si el PIB puede ser un indicador del consumo de alcohol, influyendo de manera que mayor PIB menor será el consumo.
- Distinguir si la temperatura puede ejercer una motivación para que las personas beban a diario, suponiendo que a menor temperatura mayor será el consumo.
- Establecer cuál es la influencia de la tasa de desempleo o el nivel educativo sobre el consumo de alcohol, siendo la suposición que a menor tasa de desempleo menor consumo, y a mayor nivel educativo disminuirá el consumo de alcohol.

Objetivos

El objetivo primordial de este proyecto radica en examinar y estudiar la conducta del consumo de alcohol diario a nivel europeo para conseguir un modelo matemático que explique el consumo mediante las variables explicativas. Para poder lograr esto es necesario seguir los siguientes pasos:

- Efectuar un análisis descriptivo de las variables que logren aclarar el consumo diario de alcohol.
- Verificar la presencia de relación lineal entre las variables independientes y el consumo de alcohol diario.
- Realizar el estudio de la multicolinealidad para así determinar si hay alguna relación entre las variables explicativas.
- Hacer con las variables explicativas una validación del modelo econométrico, frente al consumo diario de alcohol, tras haber realizado la multicolinealidad.
- Localizar los países que tengan un comportamiento diferente del resto.
- Con el modelo definitivo habrá que realizar las predicciones, ya que de esta forma se sabrá cómo se va a comportar un país se cambia alguna de sus variables.

Metodología y Técnicas

Metodología

En primer lugar, hay que destacar que en este apartado se enunciarán y justificarán cada uno de los pasos que se van a seguir en el proyecto.

Como se ha comentado anteriormente, la variable “y”, en este caso, será el porcentaje de personas que ingieren alcohol cada día en diferentes países. En cambio, las variables explicativas serán:

- PIB:** [1] Es el Producto Interior Bruto de un país, es decir, el valor total de los bienes y servicios producidos en un país.
- PIB per cápita:** [2] Producto Interior Bruto per cápita es un indicador que mide la relación existente entre el nivel de un país y su población.
- Temperatura:** Esta variable pretende explicar la temperatura media de cada uno de los países a lo largo del año.
- % estudios excluyendo infantil:** Porcentaje de habitantes que poseen cualquier tipo de estudios descartando los de Educación Infantil
- % estudios universitarios:** Porcentaje de personas que tienen estudios universitarios.
- Tasa de desempleo:** [3] Es el nivel de desocupación en relación con la población activa.
- % víctimas de homicidio intencionados:** Porcentaje de habitantes que fueron asesinados deliberadamente.
- Trabajos temporales:** [4] Cantidad de trabajadores que son contratados solo por un periodo de tiempo específico.
- Población:** (Según definición) Es un grupo formado por las personas que viven en un determinado lugar.
- Esperanza de vida:** [5] Es un índice que toma en cuenta para determinar cuánto se espera que viva una persona en un contexto social determinado.
- % inmigrantes:** [6] Porcentaje de personas que llegan a un país diferente de aquel del que procede con el objetivo de establecerse en él.

Es necesario aclarar, también, que todos los [7] datos de cada una de las variables citadas anteriormente corresponden al año 2014, excepto, la variable “% inmigrantes” que son del año 2015, ya que el recuento de estos en la mayoría de los países se efectuaba cada 5 años.

A continuación, se presentarán todos los datos que fueron empleados para realizar este trabajo.

Tabla 1: Tabla base de datos Eurostat.

Pais	Consuma cada día	PIB	PIB per capita	Temperatura	% estudios exc infantil	% estudios universitarios	Tasa de desempleo	% victimas homicidio intencional	Trabajos temporales	Población	Esperanza vida	% Inmigrantes
Bélgica	14,2	369	120	9,50	0,27	0,03207	8,5	0,000018	8,6	11.311.117,00	81,4	0,1227
Bulgaria	8,9	92	46	10,20	0,18	0,02676	10,3	0,000016	5,2	7.153.784,00	74,5	0,0143
Republica Checa	9,5	250	86	4,30	0,20	0,02429	5,8	0,000007	8,9	10.553.843,00	78,9	0,0379
Dinamarca	11,3	198	127	8,20	0,27	0,03292	6,2	0,000013	8,6	5.707.251,00	80,7	0,103
Alemania	9,3	2798	125	9,00	0,19	0,02111	4,9	0,000009	13,1	82.175.684,00	81,2	0,1461
Estonia	2,0	28	76	5,20	0,22	0,03081	6,4	0,000031	3,1	1.315.944,00	77,4	0,1538
Irlanda	2,0	174	137	10,00	0,26	0,03265	10,8	0,0000002	10	4.726.286,00	81,4	0,1579
España	15,3	1149	90	15,96	0,21	0,02542	23,6	0,000007	24,1	46.440.099,00	83,3	0,126
Croacia	9,1	68	58	10,85	0,19	0,02462	16,8	0,000009	16,9	4.190.669,00	77,9	0,1377
Italia	14,1	1620	97	15,70	0,18	0,01798	12,3	0,000008	13,6	60.665.551,00	83,2	0,0954
Chipre	4,2	19	81	19,70	0,20	0,02345	16,6	0,000013	19	843.319,00	82,3	0,2312
Letonia	1,4	35	64	6,10	0,21	0,02945	10,1	0,000033	3,3	1.968.957,00	74,5	0,1336
Lituania	0,6	61	75	6,25	0,23	0,03942	10,2	0,000055	2,8	2.888.558,00	74,7	0,0471
Hungría	7,3	185	68	11,55	0,20	0,02325	7,4	0,000015	10,8	9.830.485,00	76	0,0457
Malta	7,4	11	90	19,54	0,17	0,01539	5,6	0,000013	7,7	450.415,00	82,1	0,092
Austria	6,2	305	129	11,80	0,20	0,02065	5,6	0,000005	9,2	8.700.471,00	81,6	0,1715
Polonia	1,8	717	68	8,60	0,21	0,03089	8,2	0,000014	28,3	37.967.209,00	77,8	0,0163
Portugal	24,2	220	77	16,78	0,20	0,02136	13,6	0,000009	21,4	10.341.330,00	81,3	0,081
Rumania	3,8	304	55	11,15	0,19	0,01909	6,4	0,000015	1,5	19.760.314,00	75	0,0115
Eslovenia	7,7	47	83	11,10	0,20	0,02569	9,4	0,000007	16,5	2.064.188,00	81,2	0,1143
Eslovaquia	3,1	115	77	10,70	0,19	0,02085	12,3	0,000013	8,8	5.426.252,00	77	0,0327
Finlandia	2,7	166	111	5,25	0,26	0,04073	9	0,000016	15,4	5.487.308,00	81,3	0,0576
Suecia	2,7	330	124	7,60	0,25	0,02509	7,5	0,000009	13,1	9.851.017,00	82,3	0,1665
Reino Unido	7,5	1931	108	9,90	0,23	0,02344	5,7	0,000008	5,5	65.385.556,00	81,4	0,1307
Islandia	0,6	11	118	4,45	0,31	0,04203	4,4	0,000003	13,4	332.529,00	82,9	0,1128
Noruega	2,1	248	175	6,00	0,25	0,03586	3,8	0,000006	8,7	5.213.985,00	82,2	0,1423

Tras buscar, analizar e interpretar los datos, se procede a realizar los cálculos siguiendo el orden descrito a continuación:

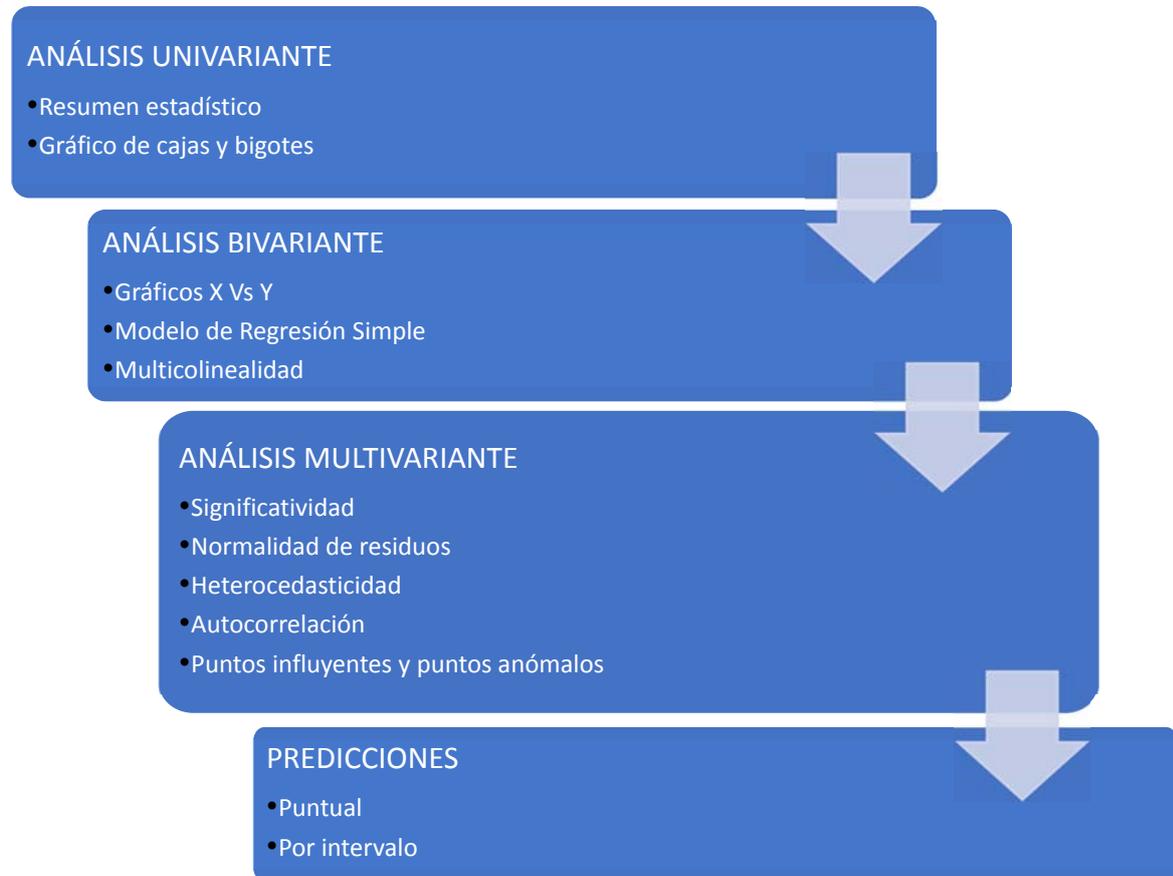


Imagen 1: Notación de Procesos a seguir.

Técnicas

Análisis univariante

Resumen estadístico

En primer lugar, se desarrollará el análisis univariante, [8] que consiste en el análisis de cada una de las variables estudiadas por separado, es decir el análisis estará basado en una sola variable.

Para poder realizar este estudio es necesario antes que nada desarrollar un resumen estadístico, éste consiste en una serie de indicadores más importantes como son:

- Recuento, número de países a analizar (**n**).
- Promedio o media, consiste principalmente en realizar una media aritmética que es la suma de los valores de la variable dividida entre el número total de países (**n**).

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{N}$$

- Mediana, se basa primordialmente en clasificar los datos de menor a mayor, para así poder seleccionar un valor medio entre estos, dejando el mismo número de valores tanto a la derecha como a la izquierda.

$$Me = L_i + \frac{\frac{N}{2} - F_{i-1}}{f_i} \cdot a_i$$

- Desviación estándar, indica lo esparcidos que están los puntos si los comparamos con la media.

$$\sigma = \sqrt{\frac{\sum(X_i)^2 \cdot f_i}{N} - (\bar{X})^2}$$

- Mínimo, es el valor más pequeño de cada variable.
- Máximo, es el valor más grande de cada variable.
- Los cuartiles son una división de la muestra en cuatro partes iguales.
 - Cuartil inferior o primer cuartil es el valor que omite por debajo el 25% de los datos.
 - Cuartil superior o tercer cuartil omite por debajo el 75% de los datos.
- Sesgo, [9] es el grado de asimetría de una distribución, es decir, cuanto se aparta de la simetría. Hay tres tipos de asimetría que son los siguientes:



Imagen 2: Tipos de asimetría.

$$CA_F = \frac{\sum_{i=1}^N (x_i - \bar{x})^3}{N S^3}$$

- Curtosis, [10] indica la cantidad de datos que hay cercanos a la media de manera que, a mayor grado de curtosis, más apuntada será la forma de la curva. También posee tres tipos:

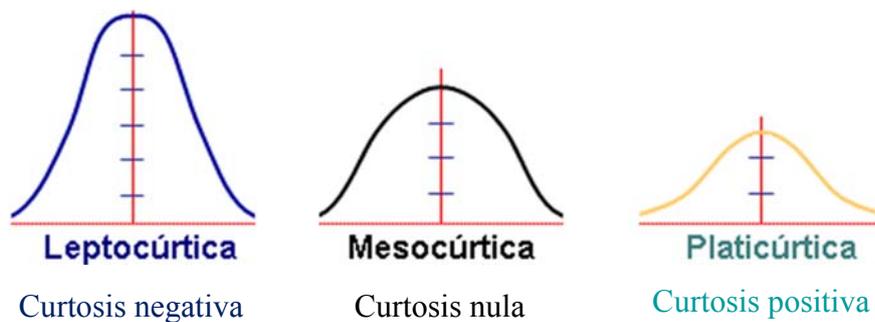


Imagen 3: Tipos de curtosis.

$$Curtosis = \frac{\sum_{i=1}^N (x_i - \bar{x})^4}{N S^4}$$

Gráfico de cajas y bigotes

Un gráfico de cajas y bigotes como se puede observar en la imagen 4 es donde podemos examinar gráficamente mediante la utilización de los cuartiles distintos grupos de datos numéricos. En los lados salientes llamados “bigotes” es donde se pueden ver claramente los puntos que tienen una variación a los de los cuartiles. En cuanto a los puntos atípicos se suelen representar completamente fuera, tanto de los cuartiles como de los bigotes, es decir se interpretan como puntos individuales que normalmente poseen un comportamiento distinto a los demás.

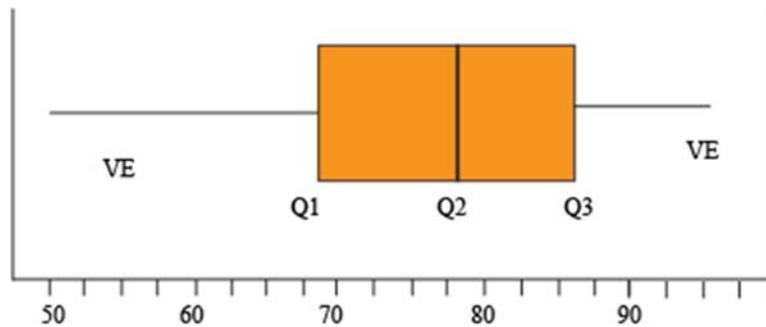


Imagen 4: Gráfico de Cajas y Bigotes.

Anàlisis Bivariante

En cuanto al análisis bivariante, este pretende explicar una variable (y) en función de otra u otras (x).

Gràfics X Vs Y

Con el gráfico X e Y lo que se pretende es representar dos variables (X e Y) para así poder determinar si existen alguna relación entre estas dos. Como ya se ha comentado anteriormente la Y es el consumo de alcohol cada día y se representará en el eje vertical, mientras que las X, eje horizontal, será cada una de las variables explicativas que se reproducirán de una en una en gráficos completamente distintos.

Modelo de regresión simple

Tras el gráfico de X e Y se presentará una función con la incògnita X para poder sustituir esta X por cualquier valor que se presente tanto en el futuro como en el pasado.

Multicolinealidad

La multicolinealidad es cuando todas las X, variables explicativas, están relacionadas, esto puede inducir varios problemas en el modelo, ya que significaría que diversas variables están proporcionando la misma información.

Para saber si el modelo econométrico tiene problemas de multicolinealidad se deben de seguir tres pasos, que se detallan a continuación:

- 1) Matriz de correlación: se trata de una matriz cuadrada y proporcionada que posee en su diagonal principal de 1, pero fuera de esta, representa los coeficientes de correlación entre las variables. Para poder localizar si hay o no problemas de multicolinealidad, se debe de mirar por debajo de la diagonal principal y buscar si aparece algún valor por encima de |0,7| significará que las dos variables ofrecen la misma información.
- 2) Matriz de correlación inversa: al igual que la matriz de correlación es una matriz cuadrada y proporcionada, pero esta no presenta una diagonal de 1, ya que esta matriz tiene en cuenta todas las variables, no solo dos. Por lo tanto, se comparará una variable con todas las otras. En este caso se debe de analizar la diagonal principal en busca de valores superiores a 10, este hecho indicará que hay problemas de multicolinealidad y por lo tanto hay por lo menos dos variables que están muy relacionadas entre sí.
- 3) Índice de acondicionamiento: Se fundamenta en los autovalores máximos y mínimos de la matriz de autocorrelación.

$$IC = \sqrt{\frac{\text{Autovalor}_{max}}{\text{Autovalor}_{min}}}$$

Tras realizar este cálculo hay que comprobar el resultado para saber si hay multicolinealidad o no. Si el IC es mayor o igual a 10 se entenderá que si hay multicolinealidad.

Hay que destacar que es imprescindible realizar los tres pasos, descritos anteriormente, para saber si existen problemas de multicolinealidad. También es importante saber que, aunque

solo uno muestre problemas y los demás no habrá que interpretarlo como una clara existencia de problemas.

Al encontrarse problemas de multicolinealidad, hay que solucionarlos, para ello el primer paso es fijarse en la matriz de correlación inversa, exactamente en los valores superiores a 10, que serán los principales candidatos para eliminar, pero siempre se empezará por eliminar el valor más grande. En el caso de que el valor más grande no tenga una diferencia de más de una unidad con el valor siguiente más grande se deberán de plantear dos modelos omitiendo en cada uno de ellos una de las variables. De esta forma, a continuación, se realizará una regresión múltiple para cada uno de los modelos planteados y se elegirá el modelo con un R^2 mayor.

Análisis Multivariante

El análisis multivariante o regresión lineal múltiple intenta obtener un modelo matemático donde se logre explicar la variable dependiente (y) mediante las variables explicativas (x).

Significatividad

Tras obtener un modelo matemático válido y haber eliminado todas las variables que tenían multicolinealidad se desarrollará la regresión múltiple para así conocer si las variables son significativas o no.

Con la ayuda de P-Valor se elaborará tanto un contraste de hipótesis para cada una de las variables (contraste de parámetro) como el contraste de modelo.

$$\left. \begin{array}{l} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{array} \right\} \alpha = 0,05$$

$$\left. \begin{array}{l} H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0 \\ H_1: \text{Algún } \beta_i \neq 0 \end{array} \right\} \alpha = 0,05$$

En ambas hipótesis planteadas si P-Valor $\geq \alpha$ (0,05) se aceptará H_0 y se concluirá que el modelo o los parámetros NO son significativos.

Normalidad de residuos

En primer lugar, para poder analizar la normalidad de los residuos debemos de obtener el gráfico probabilístico.

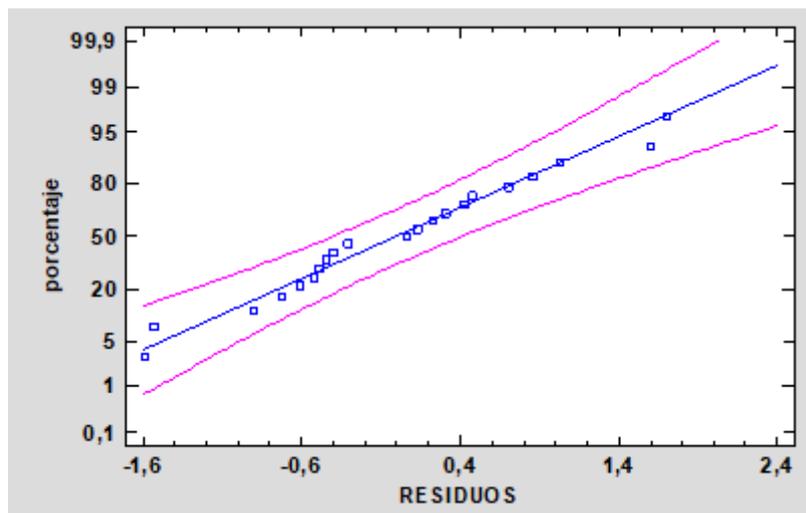


Imagen 5: Papel probabilístico normal.

En la imagen 5 se puede observar una distribución normal, ya que todos los puntos azules están muy cercanos a la línea de tendencia y ninguno de ellos se sale del intervalo de confianza que son las dos líneas rosas. Si hubiera algún punto por fuera de este intervalo se hablaría de la posibilidad de existencia de puntos anómalos o puntos influyentes.

En segundo lugar, se elabora un histograma:

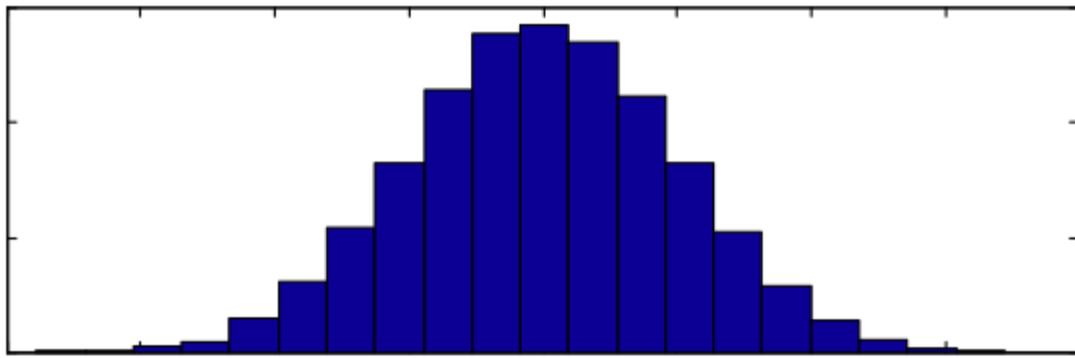


Imagen 6: Histograma.

Si este gráfico posee una forma de campana de Gauss, como en la imagen 6, este hecho podría sugerir que la gran parte de los residuos se encuentran en la media, por ello se afirmaría la existencia de la normalidad de residuos.

En tercer lugar, la normalidad de los residuos se podría determinar mediante el contraste de hipótesis. Con la ayuda de P-Valor que proporciona Statgraphics se elabora un test de normalidad, en el cual nos proporciona cuatro P-Valores de cuatro test distintos, que son: Chi-Cuadrado, Estadístico W de Shapiro-Wilk, Valor-Z para asimetría y Valor-Z para curtosis. De estos cuatro test se elegirá el más restrictivo y se plantearán las hipótesis:

$$\left. \begin{array}{l} H_0: \text{Los residuos son normales} \\ H_1: \text{Los residuos no son normales} \end{array} \right\} \alpha = 0,05$$

Si $P\text{-Valor} \geq \alpha$ (0,05) se aceptará H_0 y se concluirá que los residuos se distribuyen normalmente.

Heterocedasticidad

La heterocedasticidad [11] aparece cuando los errores no son constantes a lo largo de toda la muestra. Es decir, en los modelos de regresión lineal se dice que hay elasticidad cuando la varianza de los errores no es igual en todas las observaciones realizadas. En caso de que haya heterocedasticidad no se cumple uno de los requisitos básicos de las hipótesis de los modelos lineales.

$$\sigma^2 = \frac{\text{Var } U_i}{X_i}$$

Para poder determinar si el modelo econométrico posee heterocedasticidad se utiliza el contraste de hipótesis, para realizar este es necesario haber guardado los residuos del modelo inicial en Statgraphics, tras este paso ya se podrá realizar el contraste, se ejecutaría de la siguiente forma:

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

Para poder saber si existe o no heterocedasticidad hay que analizar el P-valor, si este es mayor o igual a α (0,05) se aceptará H_0 y se concluirá que la variable no genera problemas de heterocedasticidad.

Autocorrelación

En el caso de la autocorrelación, se analizará la existencia o no de la misma, mediante el contraste de hipótesis de autocorrelación de primer orden y mediante los gráficos FAS y FAP.

En cuanto al contraste de hipótesis de primer orden se plantea de la siguiente forma:

$$\left. \begin{array}{l} H_0: \rho_1 = 0 \\ H_1: \rho_1 \neq 0 \end{array} \right\} \alpha = 0,05$$

El contraste anterior se ha resolver mediante el P-Valor del Estadístico de Durbin Watson. Por lo tanto, si este es mayor o igual a α (0,05) se aceptará la H_0 y se concluirá que no existen problemas de autocorrelación.

En consideración a los gráficos FAS (Función de Autocorrelación Simple) y FAP (Función de Autocorrelación Parcial) son los gráficos que sirven para determinar la presencia de autocorrelación

El FAS [12] es todo el conjunto de los coeficientes de autocorrelación simple ρ_s , y mide la relación lineal que existe entre el residuo en el instante de tiempo t , e_t , y el residuo en el instante $t-s$, e_{t-s} .

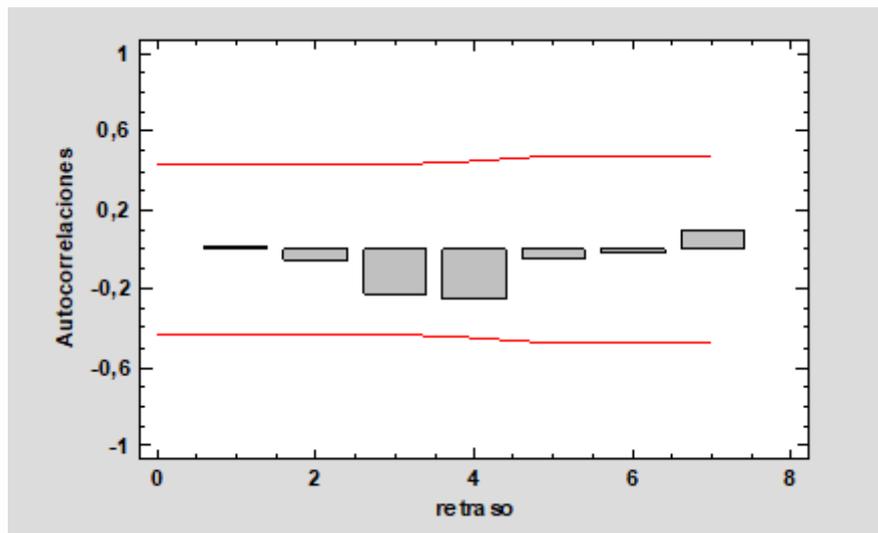


Imagen 7: FAS (Función de Autocorrelación Simple).

El FAP [12] es el conjunto de coeficiente de autocorrelación parcial y mide la relación directa que existe entre el residuo de tiempo t , e_t , y el residuo en el instante $t-s$, e_{t-s} , eliminando el posible efecto de los residuos intermedios o ambos.

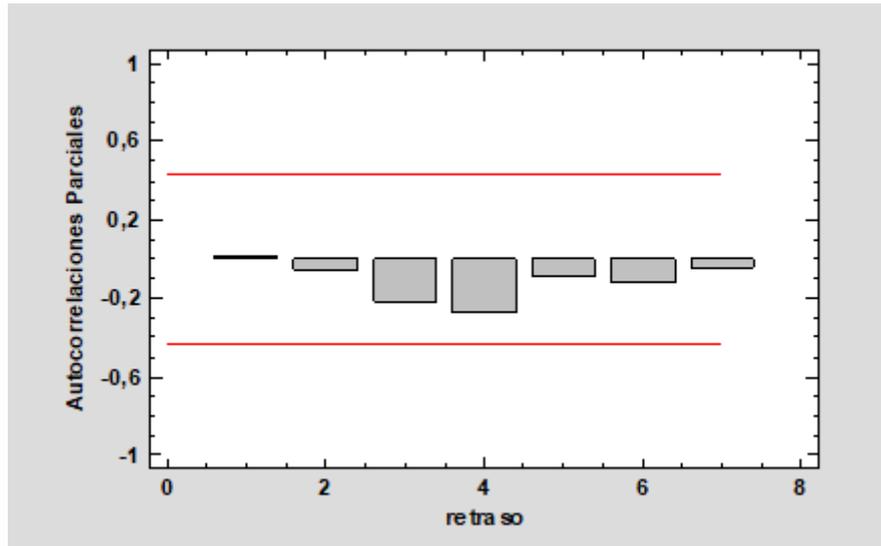


Imagen 8: FAP (Función de Autocorrelación Parcial).

Puntos influyentes y puntos anómalos

Como se ha comentado anteriormente los puntos influyentes y los puntos anómalos son detectados cuando se realiza la regresión simple. Cuando existe un punto que se aleja de la línea de tendencia y tampoco se encuentra entre los intervalos de confianza, como en la siguiente imagen:

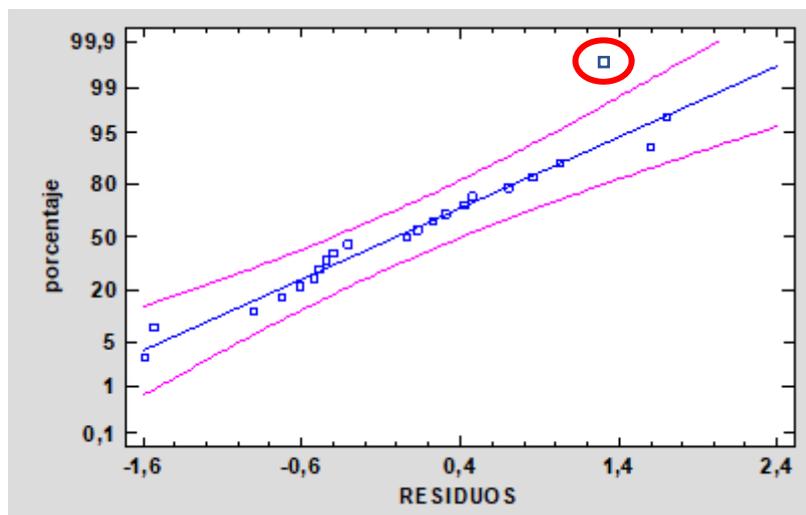


Imagen 9: Punto influyente, representación.

Hay que destacar, que los puntos influyentes pueden ser tanto a priori como a posteriori, generalmente Statgraphics dice exactamente de que puntos se trata, pero, hay que tener clara una serie de información que nos proporciona el programa son:

En el caso de los puntos priori:

Si $n_{ii} > 2\tilde{n}$, donde " \tilde{n} " es la influencia media, se trata de un punto influyente a priori.
 Si $n_{ii} < 2\tilde{n}$, no es un punto influyente a priori.

En el caso de los puntos influyentes a posteriori:

Si $|DFIS| \geq 2\sqrt{\bar{n}}$, se trata de un punto influyente a posteriori.

Si $|DFIS| \leq 2\sqrt{\bar{n}}$, no se trata de un punto influyente a posteriori.

En cuanto a los puntos anómalos, el Statgraphics calculará el residuo estudentizado, mediante el cual se puede detectar si hay puntos anómalos o no y si son candidatos para eliminar. Siempre que el residuo estudentizado, en valor absoluto, sea mayor que dos habrá un punto anómalo, pero si es superior a tres significa que no solo hay un punto anómalo, sino que también es candidato para eliminar del modelo.

Predicciones

La finalidad principal tras realizar el modelo econométrico es la predicción, ya que de esta forma se puede predecir un valor medio ajustado, a partir del modelo ajustado.

Las predicciones se pueden realizar mediante dos formas:

- Predicción por intervalos: Esta indica que la predicción para una variable puede ir entre un máximo y mínimo. Existe una gran probabilidad de que el valor de “y” realmente se encuentre entre esos dos valores.
- Predicción puntual: Suele ser generalmente la media de máximo y mínimo de los intervalos.

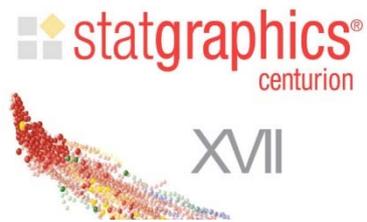
Asignaturas del grado

Las asignaturas cursadas que han sido útiles para este trabajo son:

- Econometría
- Estadística

Softwares empleados

Los softwares empleados en el proyecto son:



Statgraphics centuri3n XVII



Microsoft Excel 2018



Microsoft Word 2018



Google Chrome 2018



Microsoft Paint 2018

Cálculos, resultados y explotación

Análisis Univariante

Resumen estadístico y cajas y bigotes

Consumidores de alcohol cada día: Mide el porcentaje de personas que consumen alcohol a diario del país.

Tabla 2: Resumen estadístico para en consumo de alcohol diario durante el año 2014.

Recuento	26
Promedio	6,88462
Mediana	6,75
Desviación Estándar	5,62464
Mínimo	0,6
Máximo	24,2
Cuartil Inferior	2,1
Cuartil Superior	9,3
Sesgo	1,29248
Curtosis	2,10775

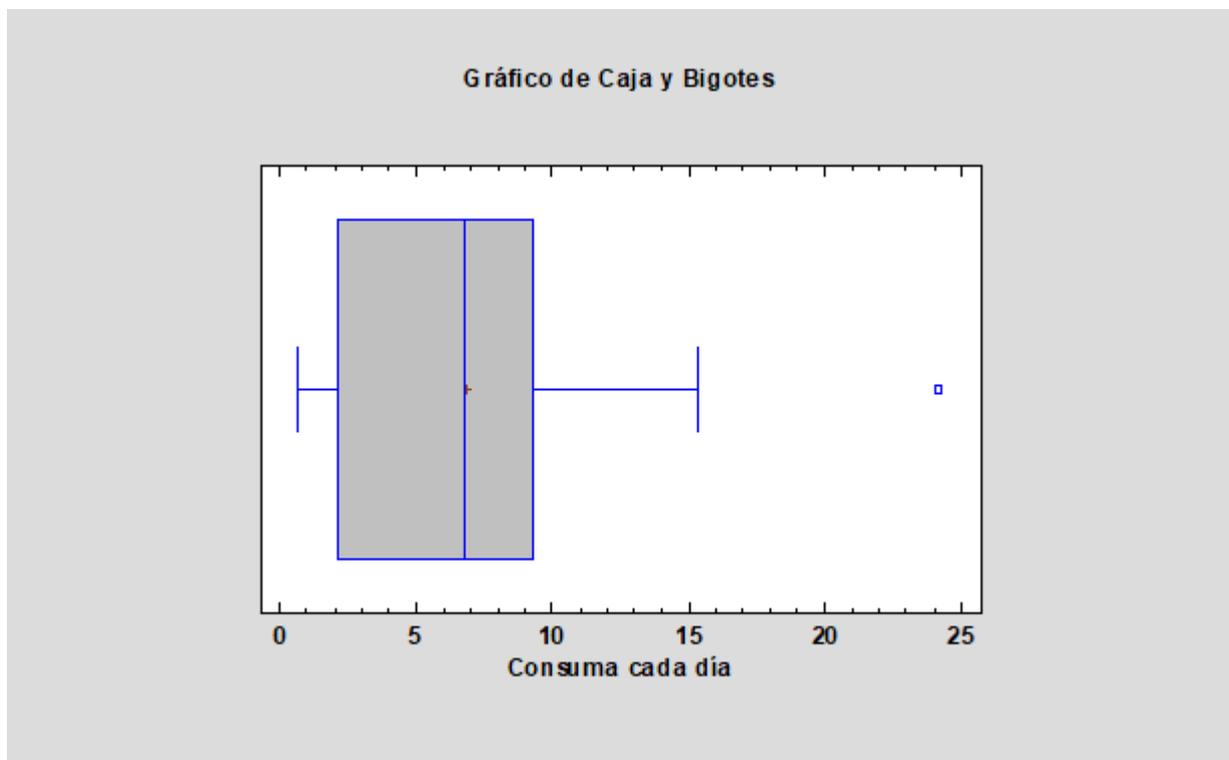


Imagen 10: Gráfico de Cajas y Bigotes del consumo de alcohol diario en el año 2014.

El punto anómalo corresponde a Portugal.

PIB

Tabla 3: Resumen estadístico para el PIB en el año 2014.

Recuento	26
Promedio	440,423
Mediana	191,5
Desviación Estándar	685,332
Mínimo	11,0
Máximo	2798,0
Cuartil Inferior	61,0
Cuartil Superior	330,0
Sesgo Estandarizado	4,96527
Curtosis	5,43361

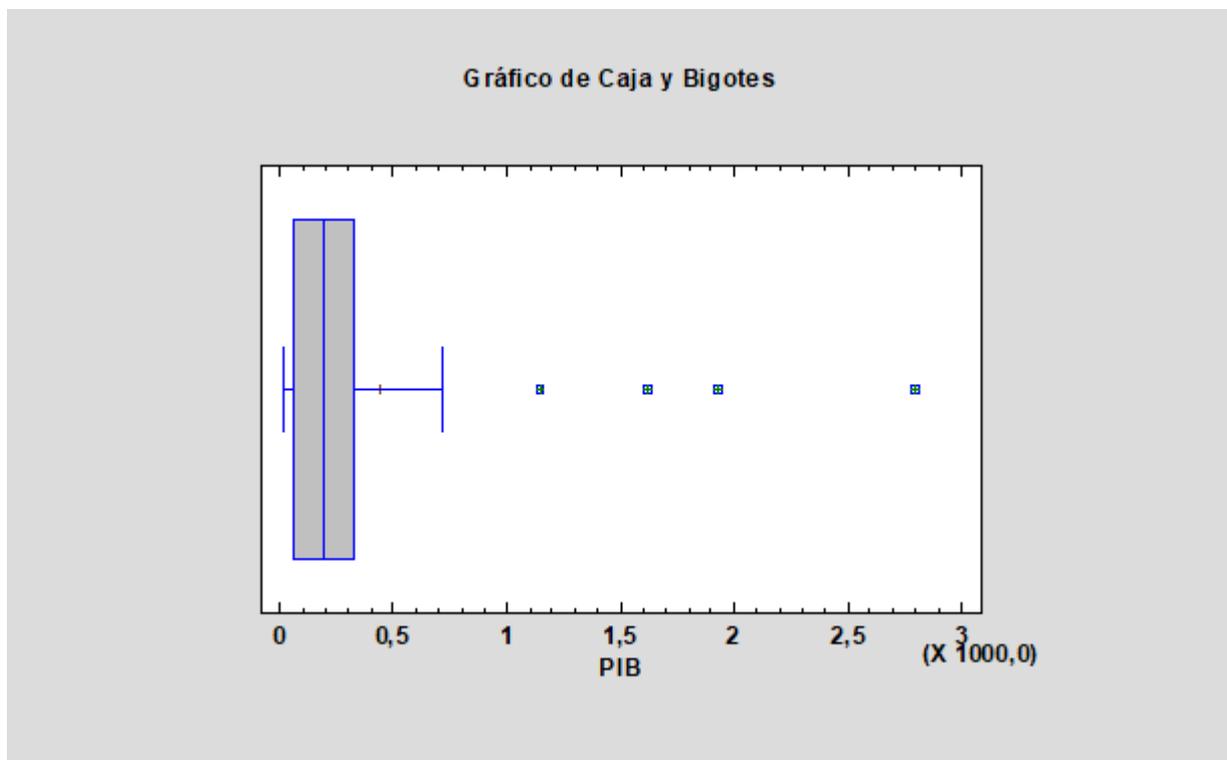


Imagen 11: Gráfico de Cajas y Bigotes para el PIB en el año 2014.

Los puntos anómalos de izquierda a derecha son: España, Italia, Reino Unido y Alemania.

PIB per cápita

Tabla 4: Resumen estadístico para el PIB per cápita en el año 2014.

Recuento	26
Promedio	94,8077
Mediana	88,0
Desviación Estándar	30,5523
Mínimo	46,0
Máximo	175,0
Cuartil Inferior	75,0
Cuartil Superior	120,0
Sesgo	0,643756
Curtosis	0,211997

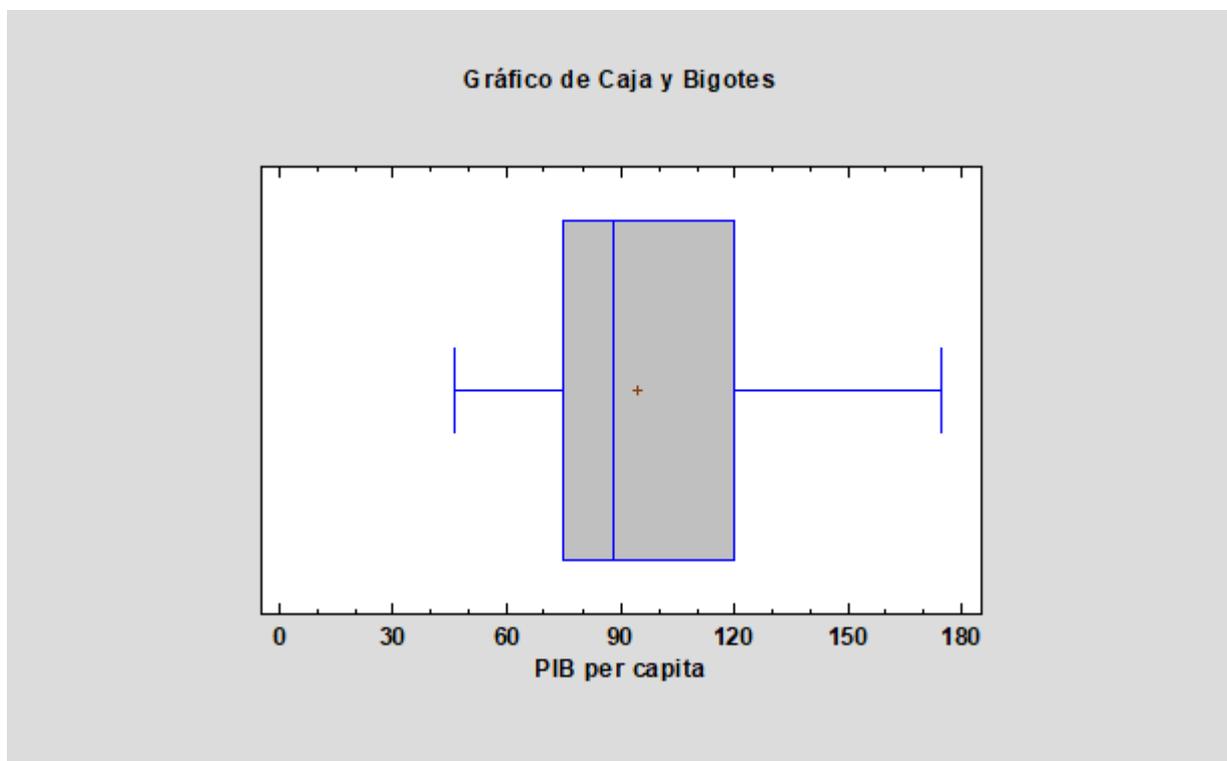


Imagen 12: Gráfico de Cajas y Bigotes para el PIB per cápita en el año 2014.

No hay puntos anómalos.

Temperatura

Tabla 5: Resumen estadístico para la temperatura en el año 2014.

Recuento	26
Promedio	10,2069
Mediana	9,95
Desviación Estándar	4,35331
Mínimo	4,3
Máximo	19,7
Cuartil Inferior	6,25
Cuartil Superior	11,55
Sesgo	0,753745
Curtosis	0,0103827

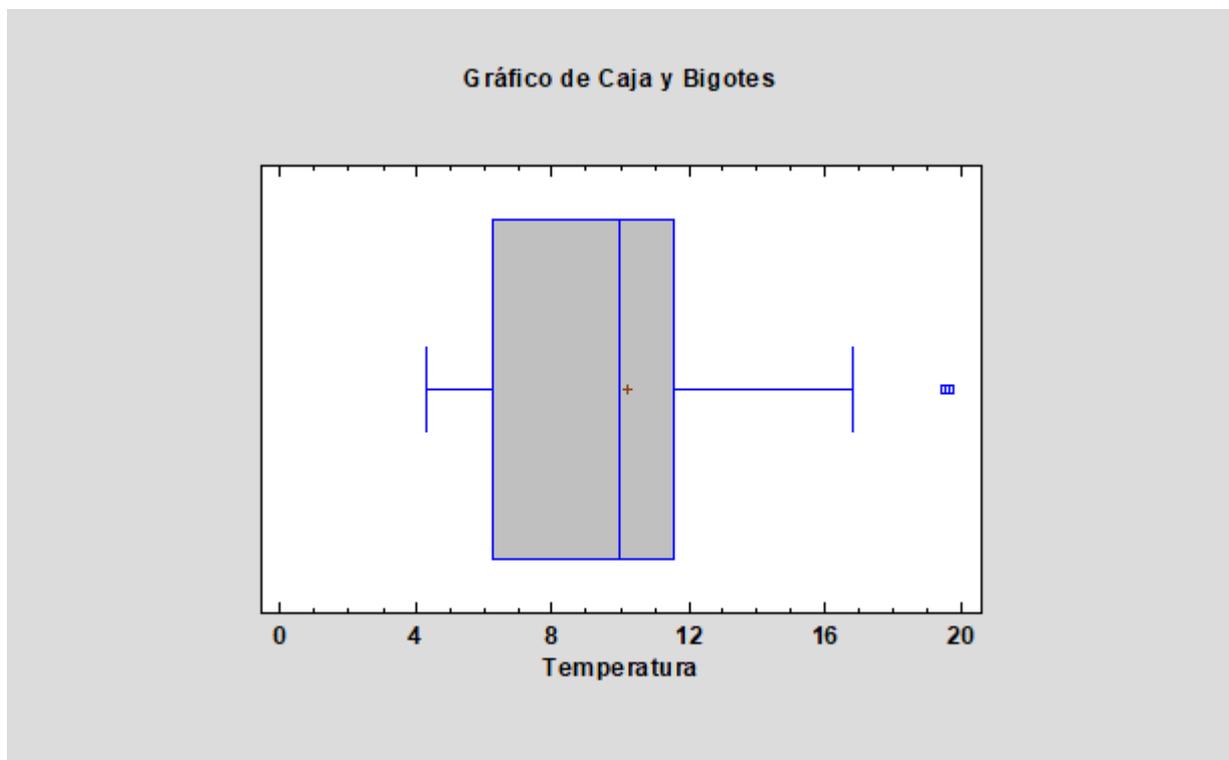


Imagen 13: Gráfico de Cajas y Bigotes para la temperatura en el año 2014.

El punto anómalo es Chipre.

% estudios excluyendo infantil

Tabla 6: Resumen estadístico para el porcentaje de habitantes con cualquier tipo de estudios excluyendo infantil en el año 2014.

Recuento	26
Promedio	0,218077
Mediana	0,205
Desviación Estándar	0,0347585
Mínimo	0,17
Máximo	0,31
Cuartil Inferior	0,19
Cuartil Superior	0,25
Sesgo	0,954962
Curtosis	0,329452

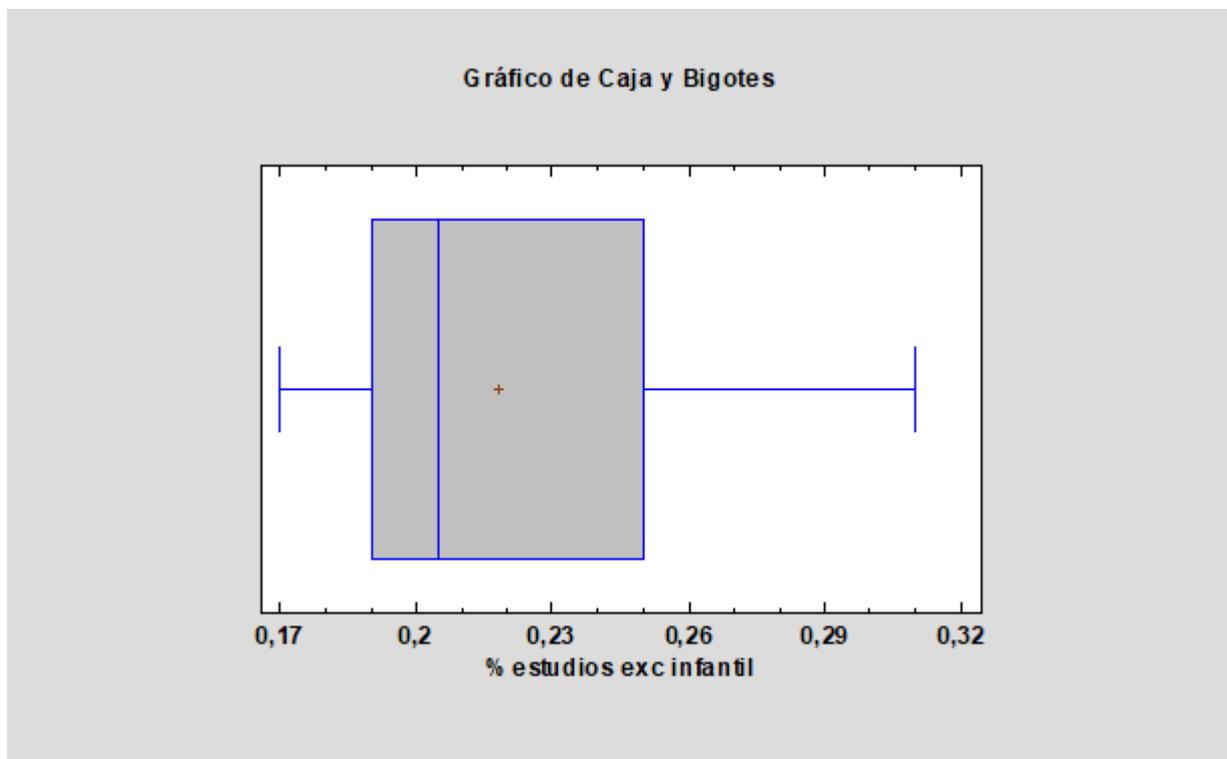


Imagen 14: Gráfico de cajas y Bigotes para el porcentaje de habitantes con cualquier tipo de estudios excluyendo infantil en el año 2014.

No existen puntos anómalos.

% estudios universitarios

Tabla 7: Resumen estadístico para el porcentaje de habitantes con estudios universitarios en el año 2014.

Recuento	26
Promedio	0,0271258
Mediana	0,025255
Desviación Estándar	0,00709568
Mínimo	0,01539
Máximo	0,04203
Cuartil Inferior	0,02136
Cuartil Superior	0,03207
Sesgo	0,595766
Curtosis	-0,348166

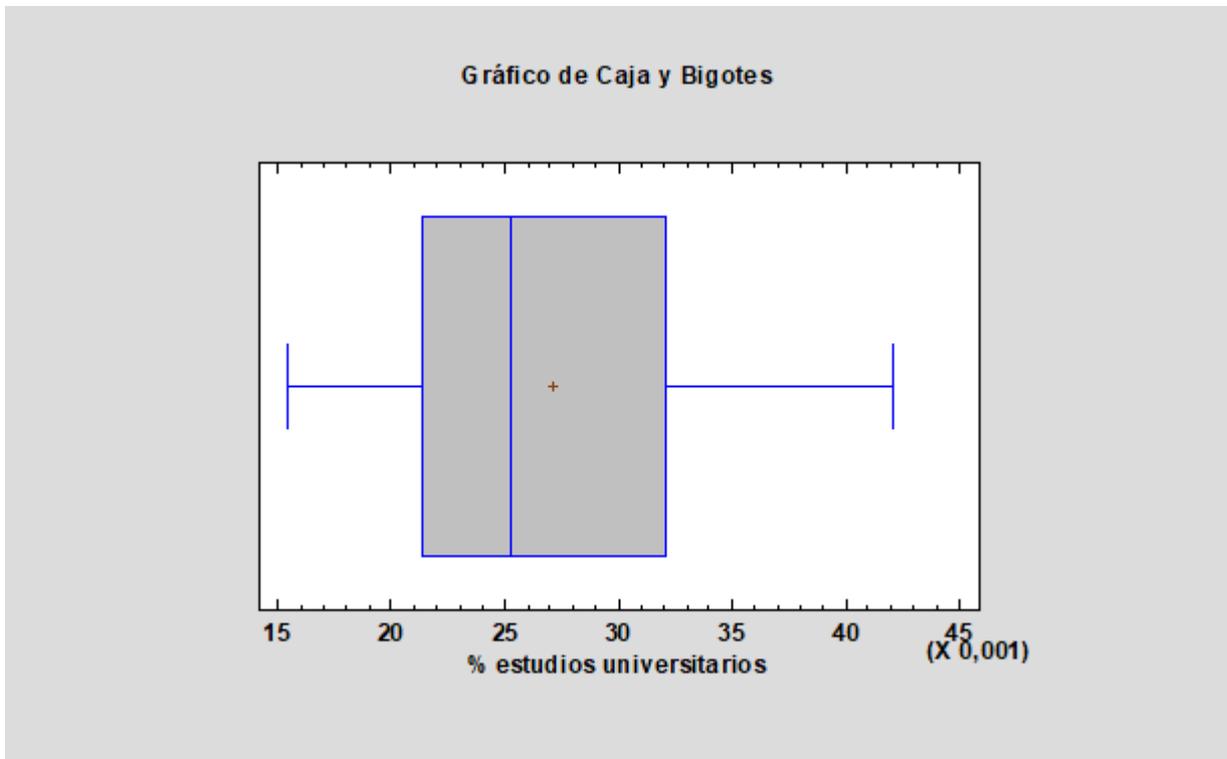


Imagen 15: Gráfico de Cajas y Bigotes para el porcentaje de habitantes con estudios universitarios en el año 2014.

No existen puntos anómalos.

Tasa de desempleo

Tabla 8: Resumen estadístico para la tasa de desempleo en el año 2014.

Recuento	26
Promedio	9,28462
Mediana	8,35
Desviación Estándar	4,546
Mínimo	3,8
Máximo	23,6
Cuartil Inferior	5,8
Cuartil Superior	10,8
Sesgo	1,49669
Curtosis	2,71924

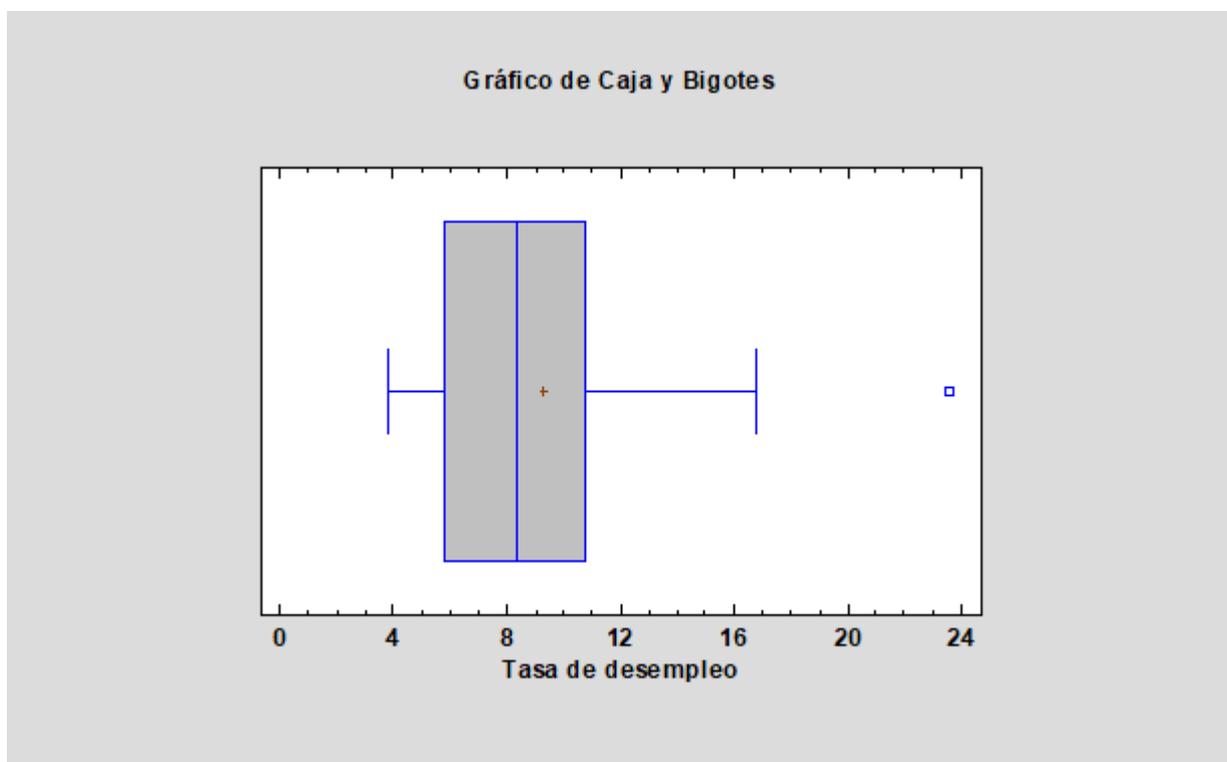


Imagen 16: Gráfico de Cajas y Bigotes para la tasa de desempleo en el año 2014.

El punto anómalo es España.

% víctimas de homicidios intencionados

Tabla 9: Resumen estadístico para el porcentaje de víctimas de homicidio intencional en el año 2014.

Recuento	26
Promedio	0,0000135385
Mediana	0,000011
Desviación Estándar	0,0000112008
Mínimo	0
Máximo	0,000055
Cuartil Inferior	0,000007
Cuartil Superior	0,000015
Sesgo	2,40033
Curtosis	7,12504

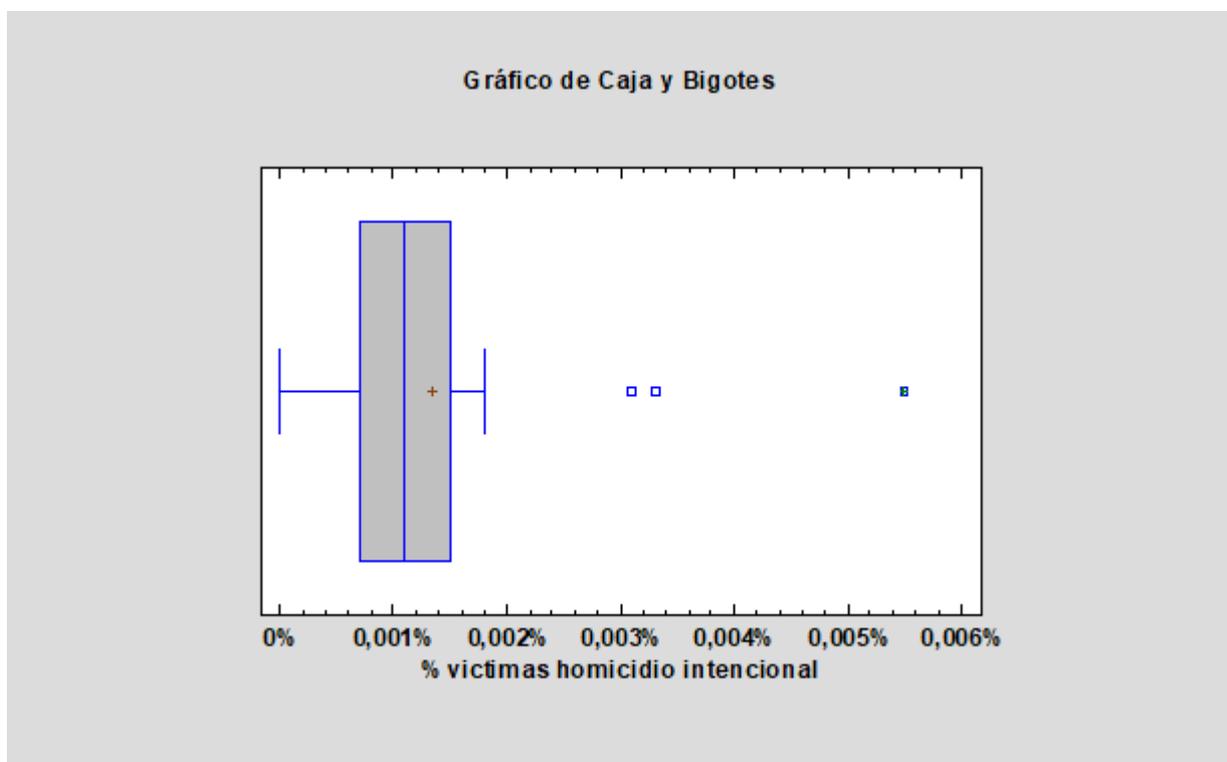


Imagen 17: Gráfico de Cajas y Bigotes para el porcentaje de víctimas por homicidio intencional en el año 2014.

Los puntos anómalos de izquierda a derecha son: Estonia, Letonia y Lituania.

Trabajos temporales

Tabla 10: Resumen estadístico para los trabajos temporales en el año 2014.

Recuento	26
Promedio	11,4423
Mediana	9,6
Desviación Estándar	6,71727
Mínimo	1,5
Máximo	28,3
Cuartil Inferior	7,7
Cuartil Superior	15,4
Sesgo	0,750326
Curtosis	0,331552

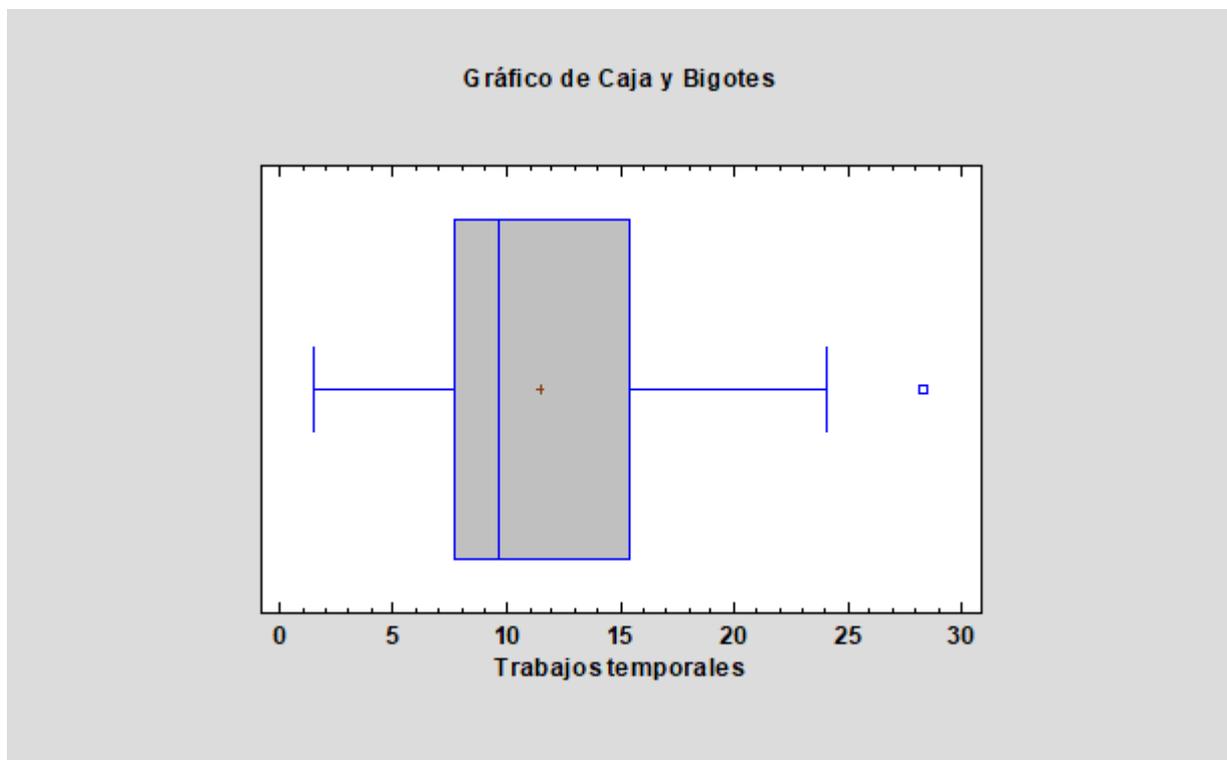


Imagen 18: Gráfico de Cajas y Bigotes para los trabajos temporales en el 2014.

El punto anómalo es Polonia.

Población

Tabla 11: Resumen estadístico para la población en el año 2014.

Recuento	26
Promedio	1,61828E7
Mediana	6,43052E6
Media Geométrica	6,69024E6
Mínimo	332529,
Máximo	8,21757E7
Cuartil Inferior	2,88856E6
Cuartil Superior	1,13111E7
Sesgo	1,88297
Curtosis	2,53552

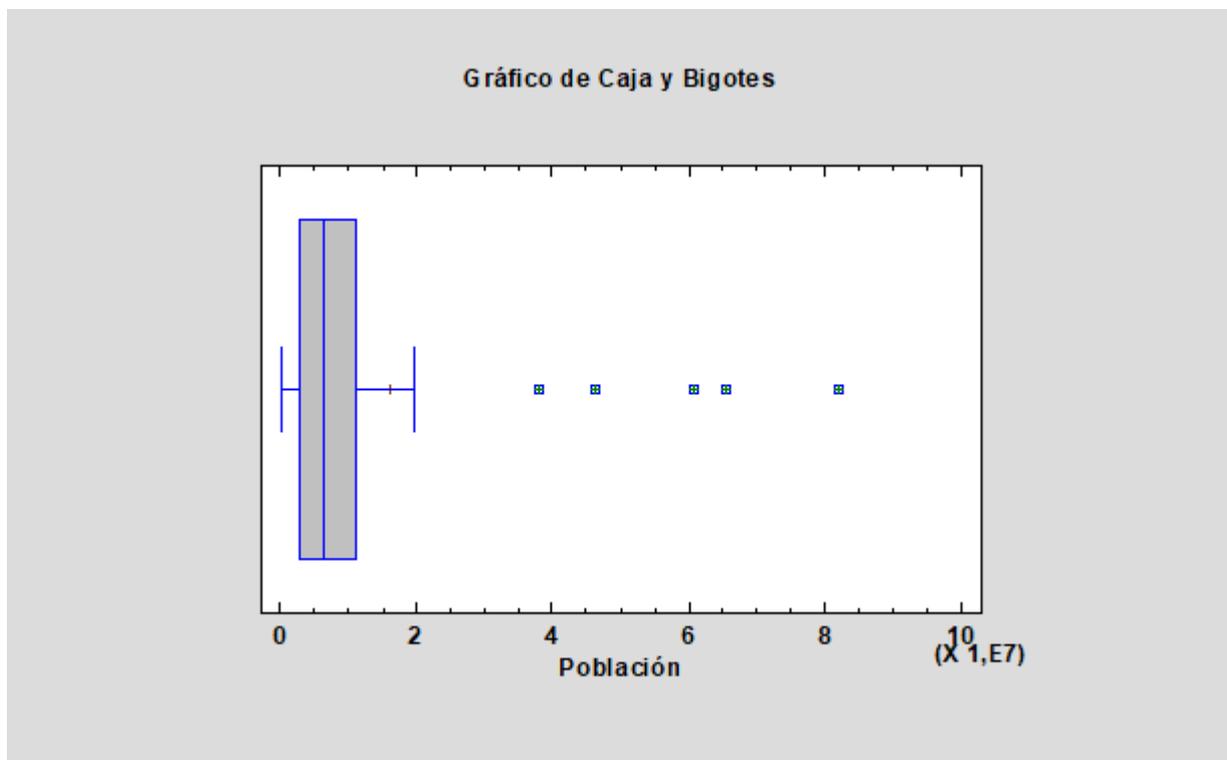


Imagen 19: Gráfico de Cajas y Bigotes para la población en el año 2014.

Los puntos anómalos son: Polonia, España, Italia, Reino Unido y Alemania.

Esperanza de vida

Tabla 12: Resumen estadístico para la esperanza de vida en el año 2014.

Recuento	26
Promedio	79,75
Mediana	81,25
Desviación Estándar	2,95801
Mínimo	74,5
Máximo	83,3
Cuartil Inferior	77,4
Cuartil Superior	82,1
Sesgo	-0,693348
Curtosis	-1,00728

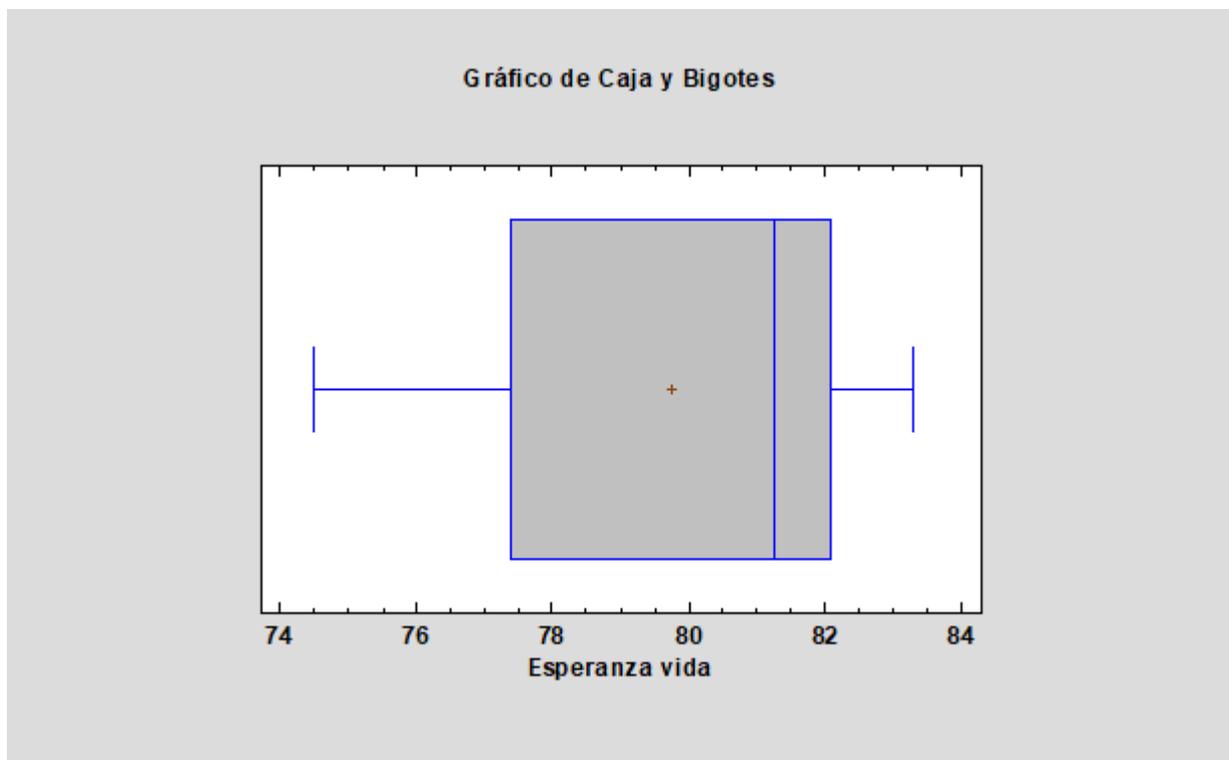


Imagen 20: Gráfico de Cajas y Bigotes para la esperanza de vida en el año 2014.

No hay puntos anómalos.

% inmigrantes

Tabla 13: Resumen estadístico para el porcentaje de inmigrantes en el año 2015.

Recuento	26
Promedio	0,103154
Mediana	0,11355
Desviación Estándar	0,0568821
Mínimo	0,0115
Máximo	0,2338
Cuartil Inferior	0,0471
Cuartil Superior	0,1423
Sesgo	0,0300615
Curtosis	-0,426722

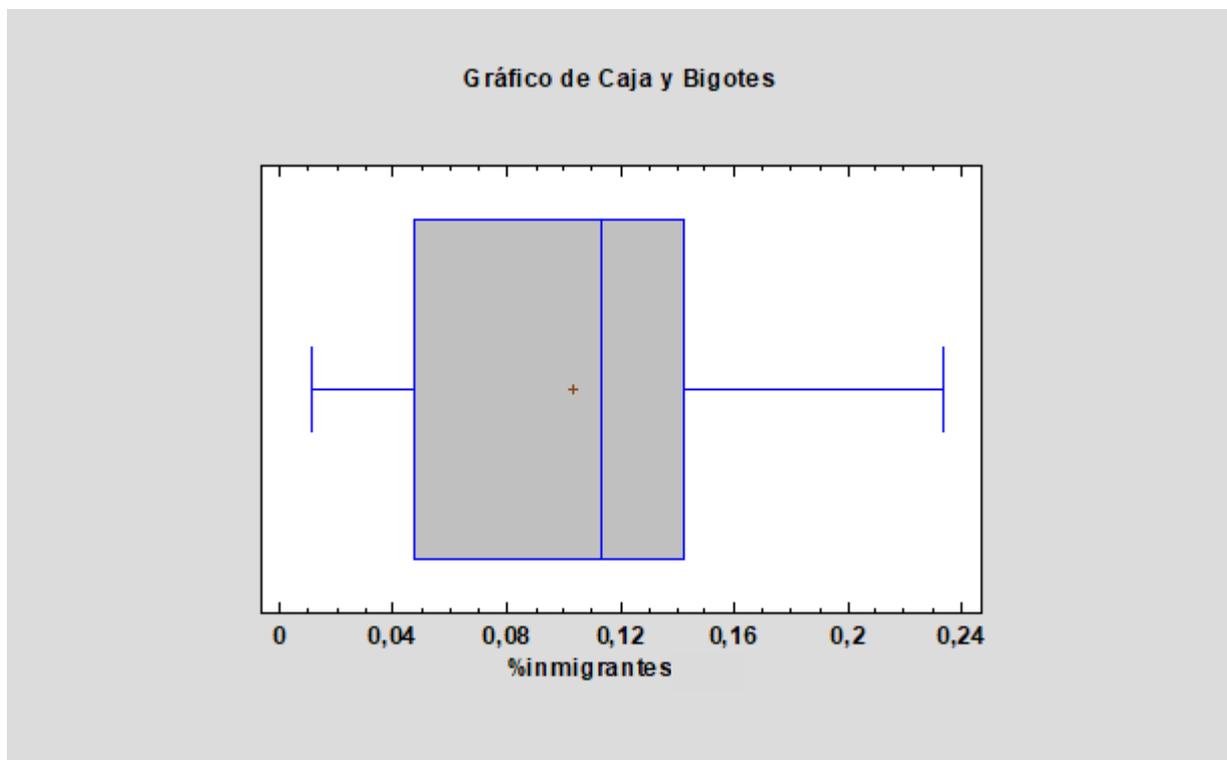


Imagen 21: Gráfico de Cajas y Bigotes para el porcentaje de inmigrantes en el año 2015.

No hay puntos anómalos.

Anàlisis Bivariante

Gràfico x e y, modelo de regresión simple y valor R

PIB

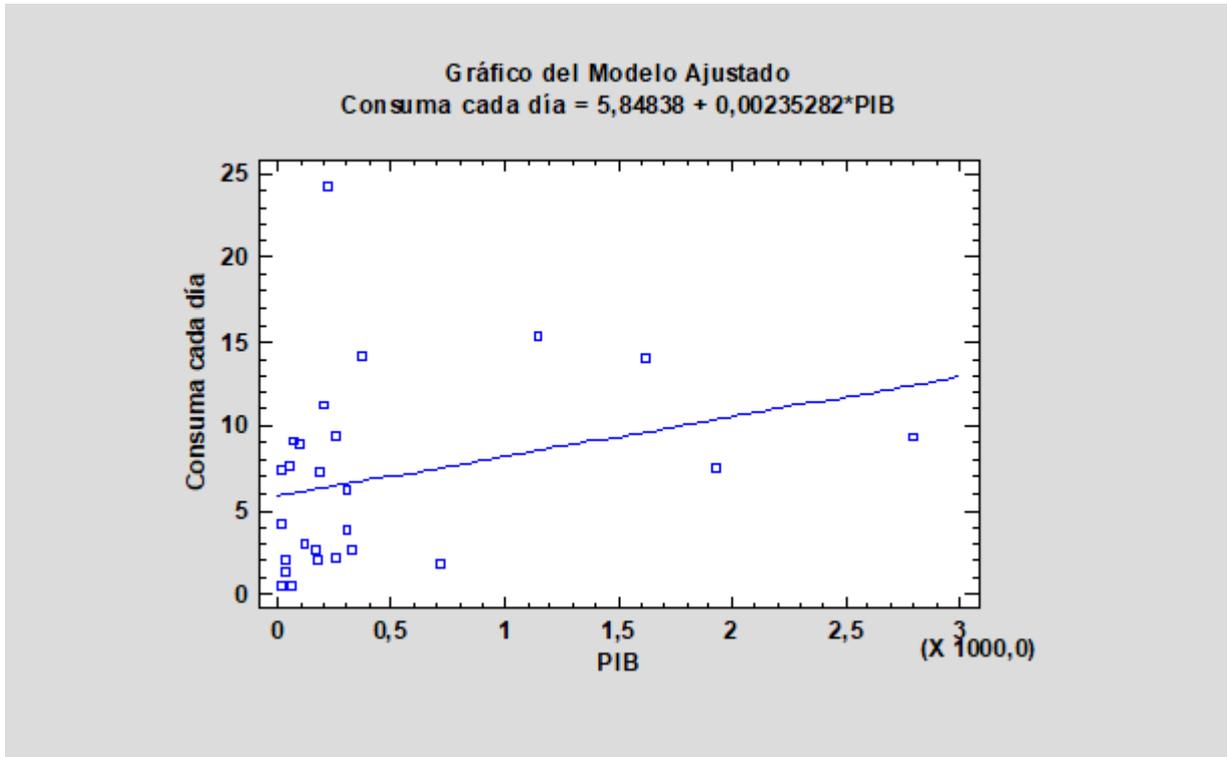


Imagen 22: Gràfico X-Y para el PIB en el año 2014.

Modelo de regresión simple: consumo cada día = 5,84838 + 0,00235282 · PIB

Coefficiente de Correlación = **0,286678**, por lo tanto, se puede decir que las variables no están relacionadas entre sí.

PIB per cpita

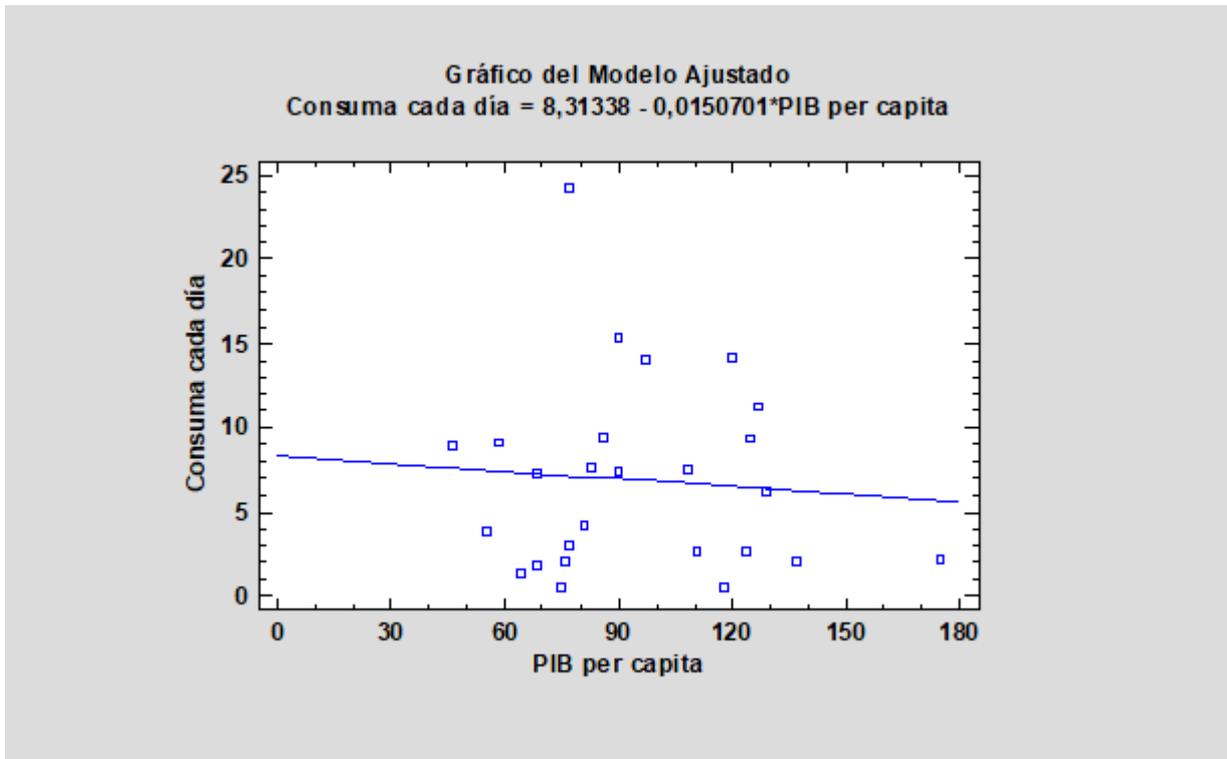


Imagen 23: Grfico X-Y para el PIB per cpita en el ao 2014.

Modelo de regresin simple: Consumo cada da = 8,31338 – 0,0150701 · PIB per cpita

Coefficiente de Correlacin = -0,0818589, por lo tanto, se puede afirmar que las variables no estn relacionadas.

Temperatura

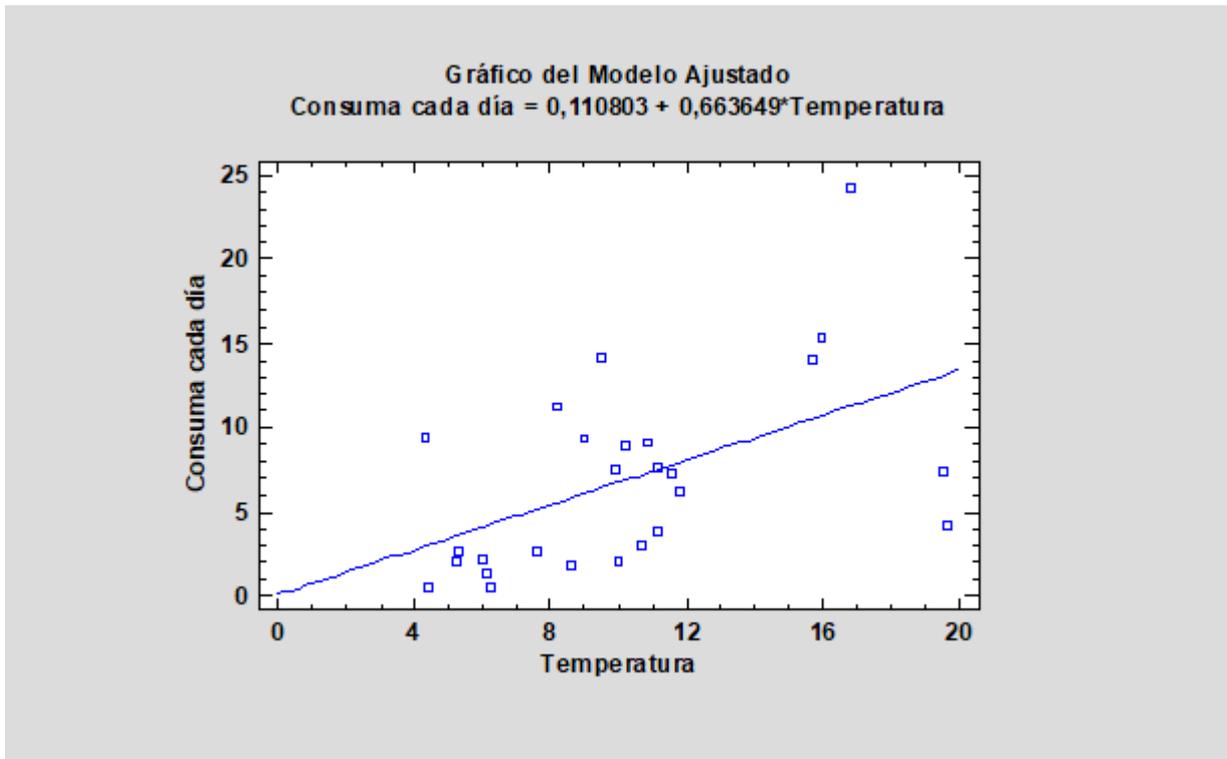


Imagen 24: Gráfico X-Y para la temperatura en el año 2014.

Modelo de regresión simple: $\text{Consumo cada día} = 0,110803 + 0,663649 \cdot \text{Temperatura}$

Coefficiente de Correlación = **0,513646**, por lo tanto, se puede decir que las variables están relacionadas moderadamente de forma directa.

% estudios excluyendo infantil

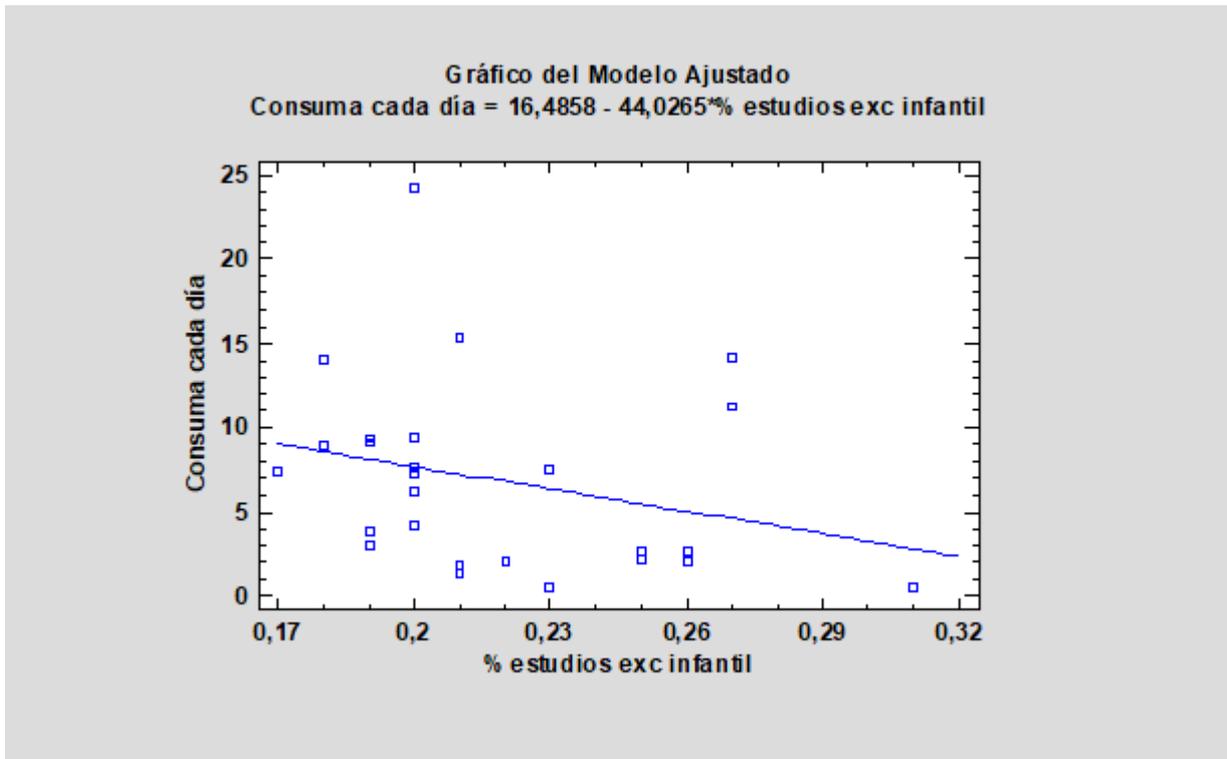


Imagen 25: Gráfico X-Y para el porcentaje de habitantes con cualquier tipo de estudios excluyendo infantil en el año 2014.

Modelo de regresión simple: Consumo cada día = 16,4858 – 44,0265 · (% estudios exc infantil)

Coefficiente de Correlación = **-0,27207**, por lo tanto, se confirma que ambas variables no están relacionadas.

% estudios universitarios

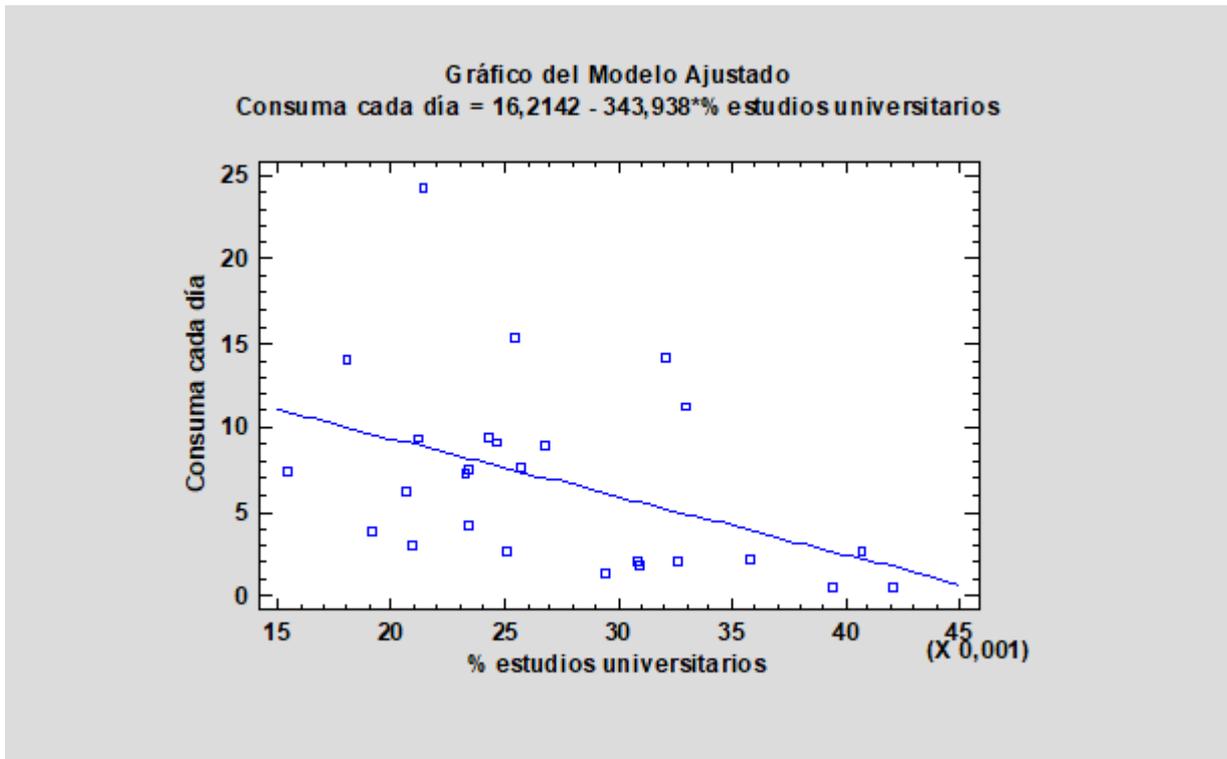


Imagen 26: Gráfico X-Y para el porcentaje de habitantes con estudios universitarios en el año 2014.

Modelo de regresión simple: Consumo cada día = 16,2142 - 343,938 · (%estudios universitarios)

Coefficiente de Correlación = -0,433889, por lo tanto, se considera que estas dos variables están moderadamente relacionadas de forma inversa.

Tasa de desempleo

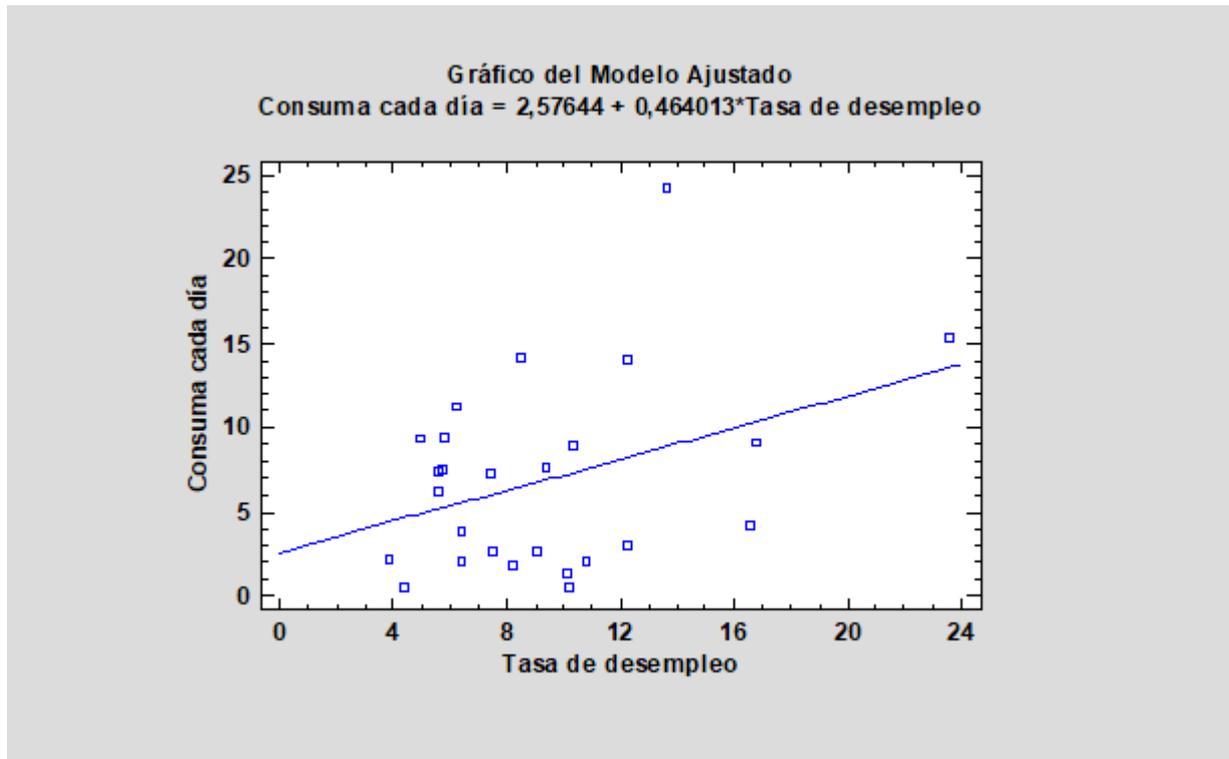


Imagen 27: Gráfico X-Y para la tasa de desempleo en el año 2014.

Modelo de regresión simple: Consumo cada día = 2,57644 + 0,464013 · (Tasa de desempleo)

Coefficiente de Correlación = 0,375029, por lo tanto, se puede decir que las variables no están relacionadas.

% víctimas homicidio intencional

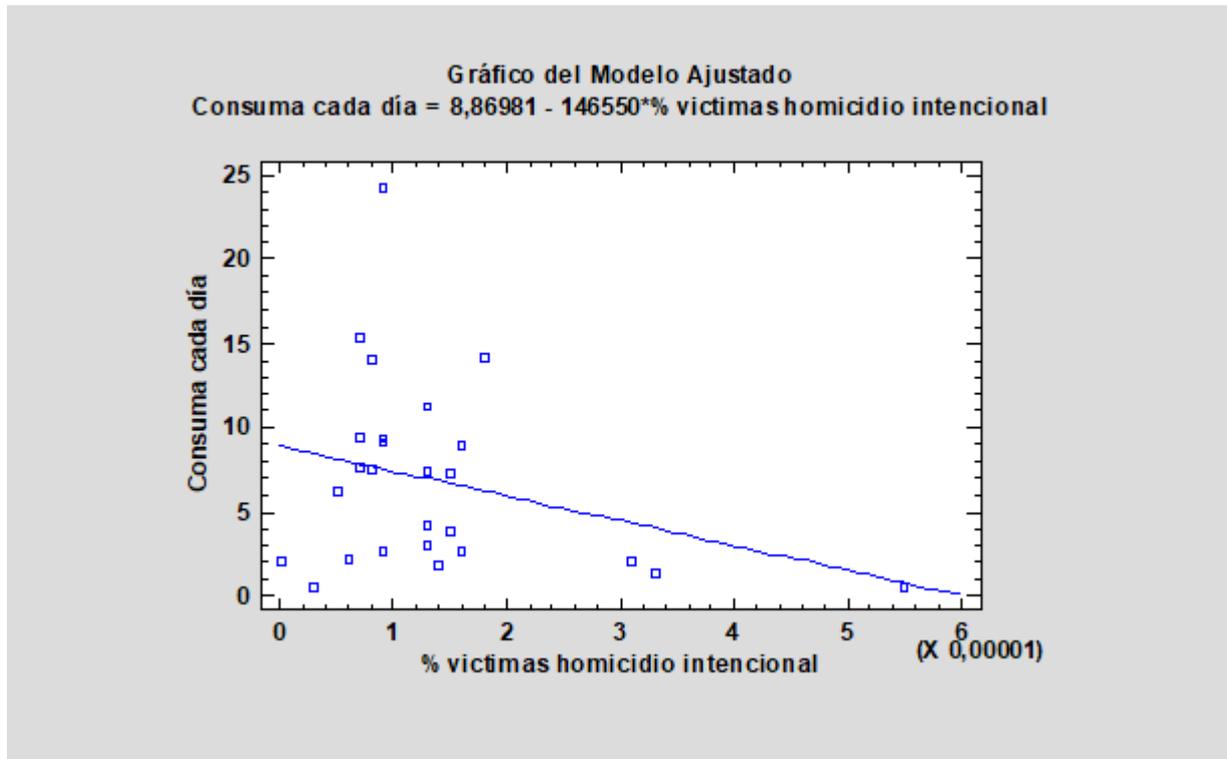


Imagen 28: Gráfico X-Y para el porcentaje de víctimas de homicidio intencional en el año 2014.

Modelo de regresión simple: Consumo cada día = 8,8681 – 146550 · (% víctimas homicidio intencional)

Coefficiente de Correlación = **-0,291588**, se puede concluir que las variables no están relacionadas.

Trabajos temporales

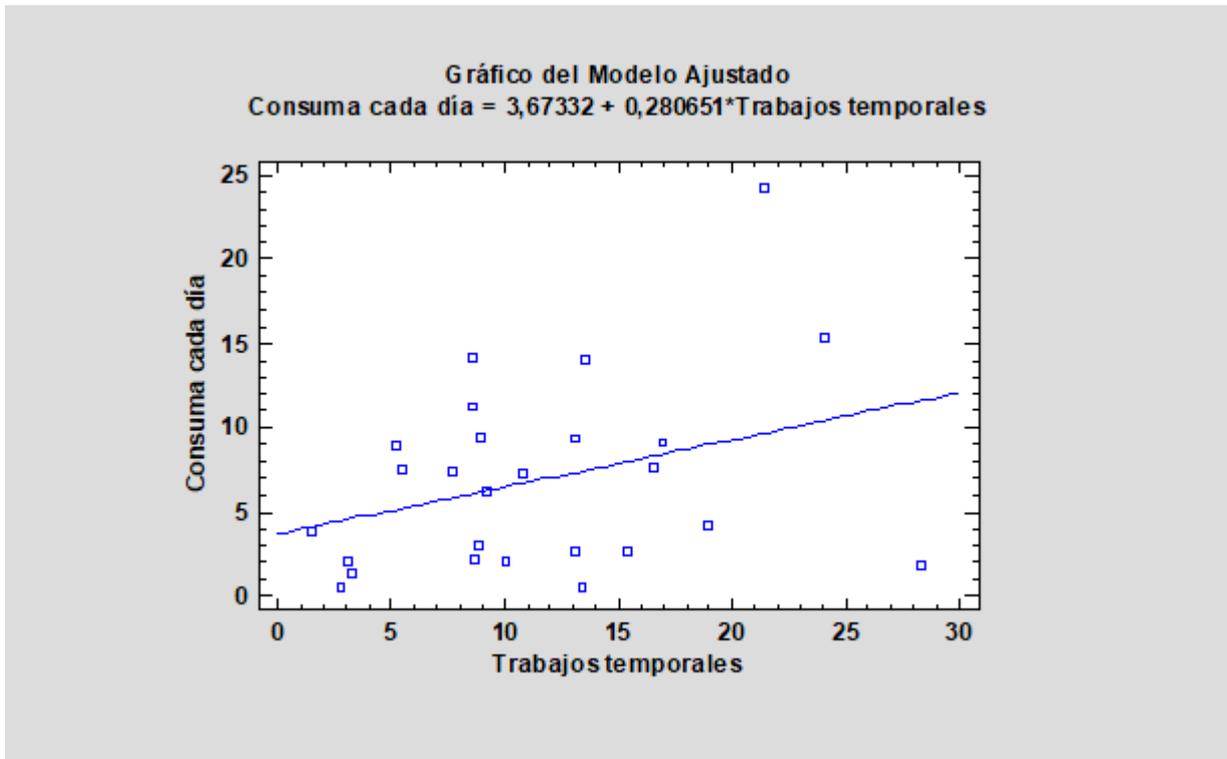


Imagen 29: Gráfico X-Y para los trabajos temporales en el año 2014.

Modelo de regresión simple: Consumo cada día = $3,67332 + 0,280651 \cdot \text{Trabajos temporales}$

Coefficiente de Correlación = $0,33517$, por lo tanto, se concluye que las variables no están relacionadas.

Población

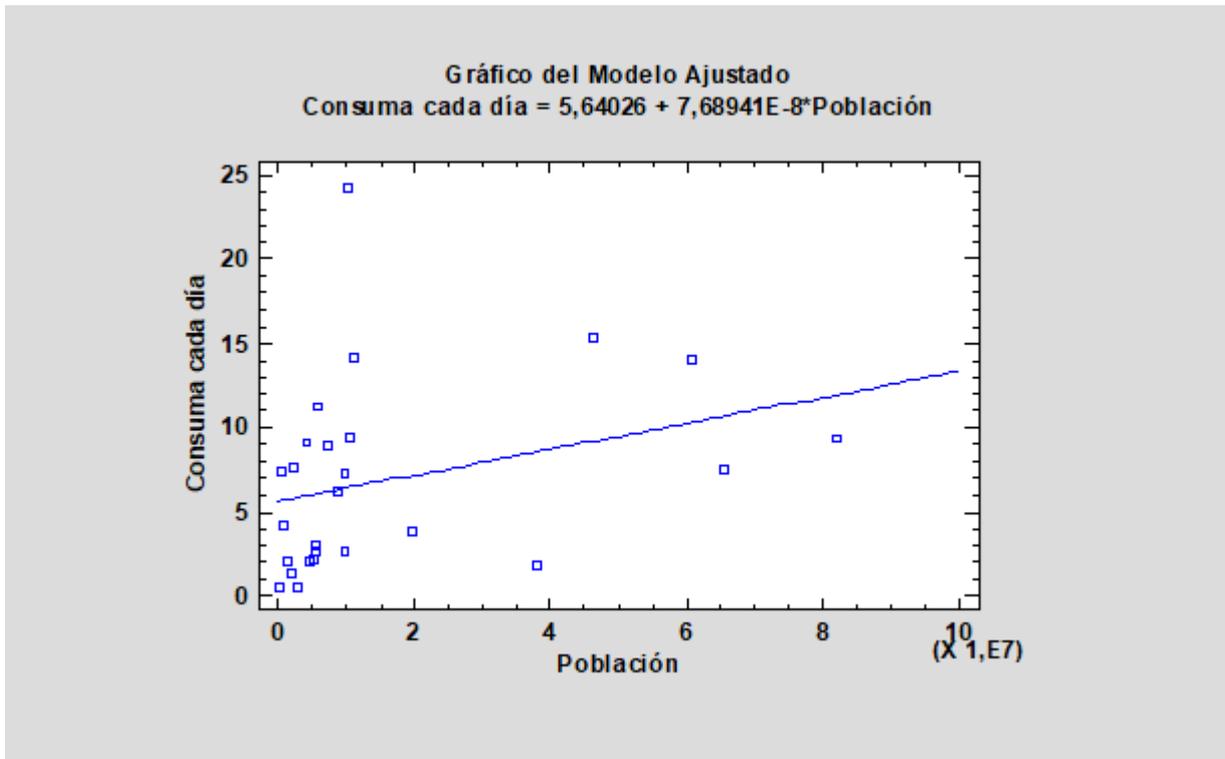


Imagen 30: Gráfico X-Y para la población en el año 2014.

Modelo de regresión simple: Consumo cada día = $5,64026 + 7,68941 \cdot 10^{-8} \cdot \text{Población}$

Coefficiente de Correlación = **0,308532**, por lo tanto, las variables no están relacionadas.

Esperanza de vida

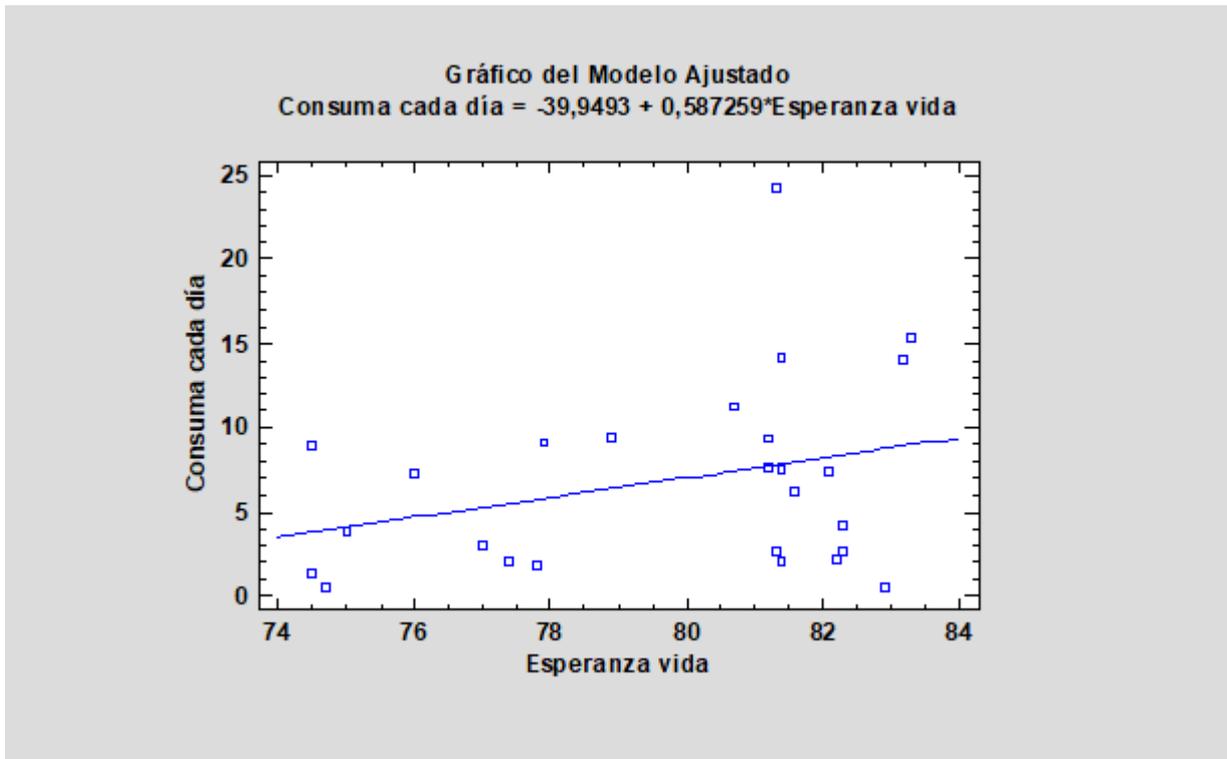


Imagen 31: Gráfico X-Y para la esperanza de vida en el año 2014.

Modelo de regresión simple: Consumo cada día = $-39,9493 + 0,587259 \cdot (\text{Esperanza de vida})$

Coefficiente de Correlación = $0,308841$, por lo tanto, las variables no están relacionadas.

% inmigrantes

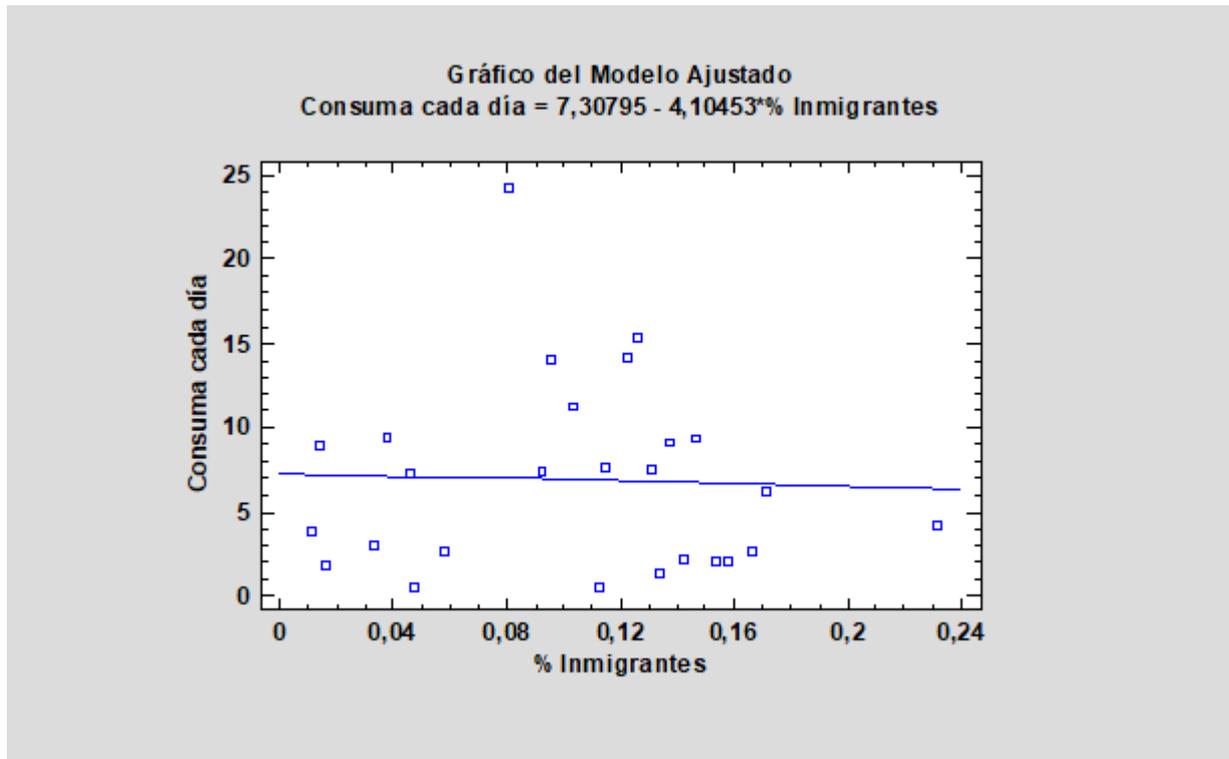


Imagen 32: Gráfico X-Y para el porcentaje de inmigrantes en el 2015.

Modelo de regresión simple: Consumo cada día = 7,30795 – 4,10453 · (%inmigrantes)

Coefficiente de Correlación = -0,0413509, por lo tanto, las variables no están relacionadas.

Multicolinealidad

Matriz de correlación

Tabla 14: Matriz de correlación (una variable frente a otra).

	PIB	PIB per capita	Temperatura	% estudios exc infantil	% estudios universitario	Tasa de desempleo	% victimas homicidio intencional	Trabajos temporales	Población	Esperanza vida	% Inmigrantes
PIB	1,00	0,23	0,10	-0,18	-0,31	-0,03	-0,23	0,16	0,98	0,31	0,12
PIB per capita	0,23	1,00	-0,22	0,60	0,31	-0,39	-0,41	-0,01	0,11	0,69	0,49
Temperatura	0,10	-0,22	1,00	-0,56	-0,69	0,52	-0,26	0,36	0,15	0,31	0,18
% estudios exc infantil	-0,18	0,60	-0,56	1,00	0,81	-0,28	-0,04	-0,04	-0,23	0,32	0,22
% estudios universitarios	-0,31	0,31	-0,69	0,81	1,00	-0,16	0,31	-0,04	-0,34	-0,03	-0,01
Tasa de desempleo	-0,03	-0,39	0,52	-0,28	-0,16	1,00	0,00	0,51	0,04	0,06	0,13
% victimas homicidio intencional	-0,23	-0,41	-0,26	-0,04	0,31	0,00	1,00	-0,43	-0,23	-0,64	-0,22
Trabajos temporales	0,16	-0,01	0,36	-0,04	-0,04	0,51	-0,43	1,00	0,23	0,46	0,10
Población	0,98	0,11	0,15	-0,23	-0,34	0,04	-0,23	0,23	1,00	0,25	0,01
Esperanza vida	0,31	0,69	0,31	0,32	-0,03	0,06	-0,64	0,46	0,25	1,00	0,57
% Inmigrantes	0,12	0,49	0,18	0,22	-0,01	0,13	-0,22	0,10	0,01	0,57	1,00

Matriz de correlación inversa

Tabla 15: Matriz de correlación inversa (una variable frente al resto).

PIB	61,6	-5,3	2,5	5,0	-1,4	1,6	-1,1	2,9	-59,9	-1,6	-5,7
PIB per capita	-5,3	5,7	0,2	-0,5	-1,5	1,0	1,0	1,1	4,6	-3,2	-0,4
Temperatura	2,5	0,2	4,6	0,5	3,0	-1,3	-1,1	-0,1	-1,5	-2,7	0,2
% estudios exc infantil	5,0	-0,5	0,5	6,9	-5,5	0,4	1,1	1,5	-4,7	-2,2	-0,7
% estudios universitarios	-1,4	-1,5	3,0	-5,5	9,1	-1,3	-3,0	-2,3	2,7	0,6	1,1
Tasa de desempleo	1,6	1,0	-1,3	0,4	-1,3	2,6	0,1	-0,7	-1,8	0,3	-0,9
% victimas homicidio intencional	-1,1	1,0	-1,1	1,1	-3,0	0,1	3,1	1,2	0,5	1,1	-0,5
Trabajos temporales	2,9	1,1	-0,1	1,5	-2,3	-0,7	1,2	3,1	-3,3	-2,0	-0,1
Población	-59,9	4,6	-1,5	-4,7	2,7	-1,8	0,5	-3,3	59,9	0,9	5,9
Esperanza vida	-1,6	-3,2	-2,7	-2,2	0,6	0,3	1,1	-2,0	0,9	7,2	-0,9
% Inmigrantes	-5,7	-0,4	0,2	-0,7	1,1	-0,9	-0,5	-0,1	5,9	-0,9	2,5

Índice de acondicionamiento

Tabla 16: Tabla autovalores.

Eigenvalor
3,31
2,98
1,77
1,14
0,87
0,34
0,28
0,16
0,09
0,06
0,01

$$IC = \frac{\sqrt{\text{Autovalor max.}}}{\sqrt{\text{Autovalor mín.}}} = \frac{\sqrt{3.31}}{\sqrt{0.01}} = 20.14$$

Matriz de correlación

Tabla 17: Matriz de correlación (una variable frente a otra).

	PIB per capita	Temperatura	% estudios exc infantil	% estudios universitarios	Tasa de desempleo	% víctimas homicidio intencional	Trabajos temporales	Población	Esperanza vida	% Inmigrantes
PIB per capita	1,00	-0,22	0,60	0,31	-0,39	-0,41	-0,01	0,11	0,69	0,49
Temperatura	-0,22	1,00	-0,56	-0,69	0,52	-0,26	0,36	0,15	0,31	0,18
% estudios exc infantil	0,60	-0,56	1,00	0,81	-0,28	-0,04	-0,04	-0,23	0,32	0,22
% estudios universitarios	0,31	-0,69	0,81	1,00	-0,16	0,31	-0,04	-0,34	-0,03	-0,01
Tasa de desempleo	-0,39	0,52	-0,28	-0,16	1,00	0,00	0,51	0,04	0,06	0,13
% víctimas homicidio intencional	-0,41	-0,26	-0,04	0,31	0,00	1,00	-0,43	-0,23	-0,64	-0,22
Trabajos temporales	-0,01	0,36	-0,04	-0,04	0,51	-0,43	1,00	0,23	0,46	0,10
Población	0,11	0,15	-0,23	-0,34	0,04	-0,23	0,23	1,00	0,25	0,01
Esperanza vida	0,69	0,31	0,32	-0,03	0,06	-0,64	0,46	0,25	1,00	0,57
% Inmigrantes	0,49	0,18	0,22	-0,01	0,13	-0,22	0,10	0,01	0,57	1,00

Matriz de correlación inversa

Tabla 18: Matriz de correlación inversa (una variable frente a otras).

PIB per capita	5,23	0,42	-0,03	-1,67	1,13	0,91	1,39	-0,54	-3,31	-0,86
Temperatura	0,42	4,51	0,32	3,03	-1,41	-1,07	-0,25	0,92	-2,67	0,41
% estudios exc infantil	-0,03	0,32	6,47	-5,35	0,25	1,20	1,28	0,13	-2,05	-0,27
% estudios universitarios	-1,67	3,03	-5,35	9,06	-1,28	-3,03	-2,27	1,31	0,52	0,97
Tasa de desempleo	1,13	-1,41	0,25	-1,28	2,51	0,14	-0,75	-0,27	0,34	-0,77
% víctimas homicidio intencional	0,91	-1,07	1,20	-3,03	0,14	3,04	1,26	-0,55	1,07	-0,63
Trabajos temporales	1,39	-0,25	1,28	-2,27	-0,75	1,26	2,95	-0,47	-1,88	0,20
Población	-0,54	0,92	0,13	1,31	-0,27	-0,55	-0,47	1,55	-0,65	0,40
Esperanza vida	-3,31	-2,67	-2,05	0,52	0,34	1,07	-1,88	-0,65	7,13	-1,09
% Inmigrantes	-0,86	0,41	-0,27	0,97	-0,77	-0,63	0,20	0,40	-1,09	1,98

Índice acondicionamiento

Tabla 19: Tabla autovalores.

Eigenvalor
3,06
2,94
1,34
1,00
0,73
0,34
0,27
0,16
0,09
0,06

$$IC = \sqrt{\frac{\text{Autovalor max.}}{\text{Autovalor mín.}}} = \sqrt{\frac{3.06}{0.06}} = 7.07$$

Matriz de correlación

Tabla 20: Matriz de correlación (una variable frente a otra).

	PIB per capita	Temperatura	% estudios exc infantil	Tasa de desempleo	% victimas homicidio int	Trabajos temporales	Población	Esperanza vida	% Inmigrantes
PIB per capita	1,00	-0,22	0,60	-0,39	-0,41	-0,01	0,11	0,69	0,49
Temperatura	-0,22	1,00	-0,56	0,52	-0,26	0,36	0,15	0,31	0,18
% estudios exc infantil	0,60	-0,56	1,00	-0,28	-0,04	-0,04	-0,23	0,32	0,22
Tasa de desempleo	-0,39	0,52	-0,28	1,00	0,00	0,51	0,04	0,06	0,13
% victimas homicidio intencional	-0,41	-0,26	-0,04	0,00	1,00	-0,43	-0,23	-0,64	-0,22
Trabajos temporales	-0,01	0,36	-0,04	0,51	-0,43	1,00	0,23	0,46	0,10
Población	0,11	0,15	-0,23	0,04	-0,23	0,23	1,00	0,25	0,01
Esperanza vida	0,69	0,31	0,32	0,06	-0,64	0,46	0,25	1,00	0,57
% Inmigrantes	0,49	0,18	0,22	0,13	-0,22	0,10	0,01	0,57	1,00

Matriz de correlación inversa

Tabla 21: Matriz de correlación inversa (una variable frente a otras).

PIB per capita	4,93	0,98	-1,01	0,90	0,36	0,98	-0,29	-3,21	-0,68
Temperatura	0,98	3,50	2,10	-0,99	-0,06	0,51	0,48	-2,84	0,09
% estudios exc infantil	-1,01	2,10	3,32	-0,51	-0,58	-0,06	0,90	-1,74	0,30
Tasa de desempleo	0,90	-0,99	-0,51	2,33	-0,28	-1,07	-0,08	0,41	-0,64
% victimas homicidio intencional	0,36	-0,06	-0,58	-0,28	2,03	0,51	-0,12	1,25	-0,31
Trabajos temporales	0,98	0,51	-0,06	-1,07	0,51	2,39	-0,14	-1,75	0,44
Población	-0,29	0,48	0,90	-0,08	-0,12	-0,14	1,36	-0,73	0,26
Esperanza vida	-3,21	-2,84	-1,74	0,41	1,25	-1,75	-0,73	7,10	-1,15
% Inmigrantes	-0,68	0,09	0,30	-0,64	-0,31	0,44	0,26	-1,15	1,87

Índice de acondicionamiento

Tabla 22: Tabla autovalores.

<i>Eigenvalor</i>
2,95
2,49
1,12
0,88
0,71
0,33
0,27
0,16
0,09

$$IC = \sqrt{\frac{\text{Autovalor max.}}{\text{Autovalor mín.}}} = \sqrt{\frac{2.95}{0.09}} = 5.86$$

Análisis Multivariado

Modelo teórico propuesto

$$\begin{aligned} \text{Consumo de alcohol cada día} = & \beta_0 + \beta_1 \cdot \text{PIB per cápita} + \beta_2 \cdot \text{Temperatura} + \\ & + \beta_3 \cdot (\% \text{estudios excluyendo infantil}) + \beta_4 \cdot \text{Tasa de desempleo} + \beta_5 \cdot \\ & \cdot \text{Víctimas homicidio intencional} + \beta_6 \cdot \text{Trabajos temporales} + \beta_7 \cdot \\ & \cdot \text{Población} + \beta_8 \cdot \text{Esperanza de vida} + \beta_9 \cdot (\% \text{Inmigrantes}) + U \end{aligned}$$

Interpretación de los parámetros

β_0 : Valor medio de la variable Consumo alcohol cada día, cuando las variables explicativas valen 0.

β_1 : Incremento medio de la variable Consumo alcohol cada día cuando aumentamos en 1 unidad PIB per cápita, y el resto de las variables explicativas permanecen constantes.

β_2 : Incremento medio de la variable Consumo alcohol cada día cuando aumentamos en 1 unidad Temperatura, y el resto de las variables explicativas permanecen constantes.

β_3 : Incremento medio de la variable Consumo de alcohol cada día cuando aumentamos en 1 unidad estudios excluyendo infantil, y el resto de las variables explicativas permanecen constantes.

β_4 : Incremento medio de la variable Consumo de alcohol día cuando aumentamos en 1 unidad Tasa de desempleo, y el resto de las variables explicativas permanecen constantes.

β_5 : Incremento medio de la variable Consumo de alcohol día cuando aumentamos en 1 unidad Víctimas homicidio intencional, y el resto de las variables explicativas permanecen constantes.

β_6 : Incremento medio de la variable Consumo alcohol cada día cuando aumentamos en 1 unidad Trabajos temporales, y el resto de las variables explicativas permanecen constantes.

β_7 : Incremento medio de la variable Consumo alcohol cada día cuando aumentamos en 1 unidad Población, y el resto de las variables explicativas permanecen constantes.

β_8 : Incremento medio de la variable Consumo alcohol cada día cuando aumentamos en 1 unidad Esperanza de vida, y el resto de las variables explicativas permanecen constantes.

β_9 : Incremento medio de la variable Consumo alcohol cada día cuando aumentamos en 1 unidad Inmigrantes, y el resto de las variables explicativas permanecen constantes.

U: Parte de los datos que no se pueden explicar con las variables explicativas que se han

elegido.

Modelo ajustado

$$\begin{aligned}
 \text{Consumo cada día} = & -65,9014 - 0,021798 \cdot (\text{PIB per cápita}) + \\
 & +0,299299 \cdot (\text{Temperatura}) - 8,10204 \cdot (\% \text{ estudios exc infantil}) + 0,348475 \cdot \\
 & \cdot (\text{Tasa de desempleo}) - 34194,4 \cdot (\% \text{ víctimas homicidio intncionl}) - \\
 & -0,129715 \cdot (\text{Trabajos temporales}) + 4,03708 \cdot 10^{-8} \cdot \text{Población} + 0,940791 \cdot \\
 & \cdot \text{Esperanza de vida} - 33,0124 \cdot (\% \text{ inmigrantes})
 \end{aligned}$$

La R^2 de este modelo es 46,24%, se puede decir que es un valor muy bajo, por lo tanto, si nos fijamos en el siguiente gráfico:

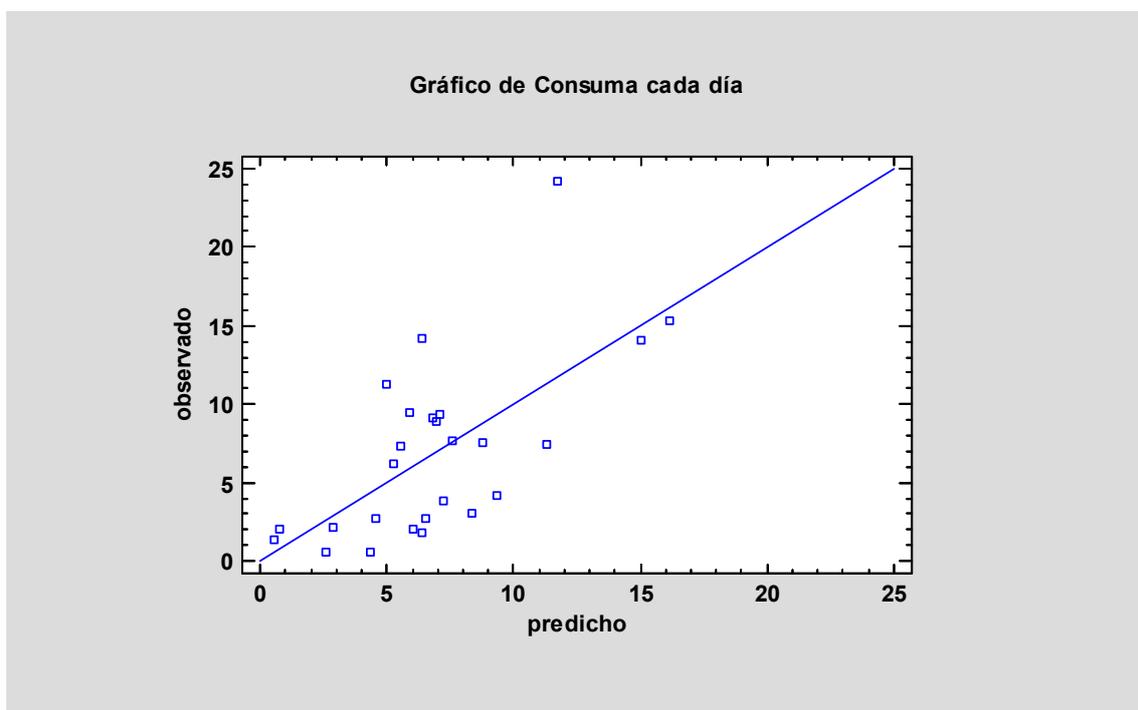


Gráfico 33: Gráfico de datos predicho frente a datos observados (a)

Se ve claramente que hay varios puntos que no se comportan igual al resto, en este caso se decide eliminarlos, para así intentar mejorar el R^2 del modelo.

Tras analizar cada uno de los puntos, se decide eliminar los siguientes países: Portugal, Bélgica, Dinamarca, Hungría y Eslovaquia.

Después de hacer esta operación el modelo queda así:

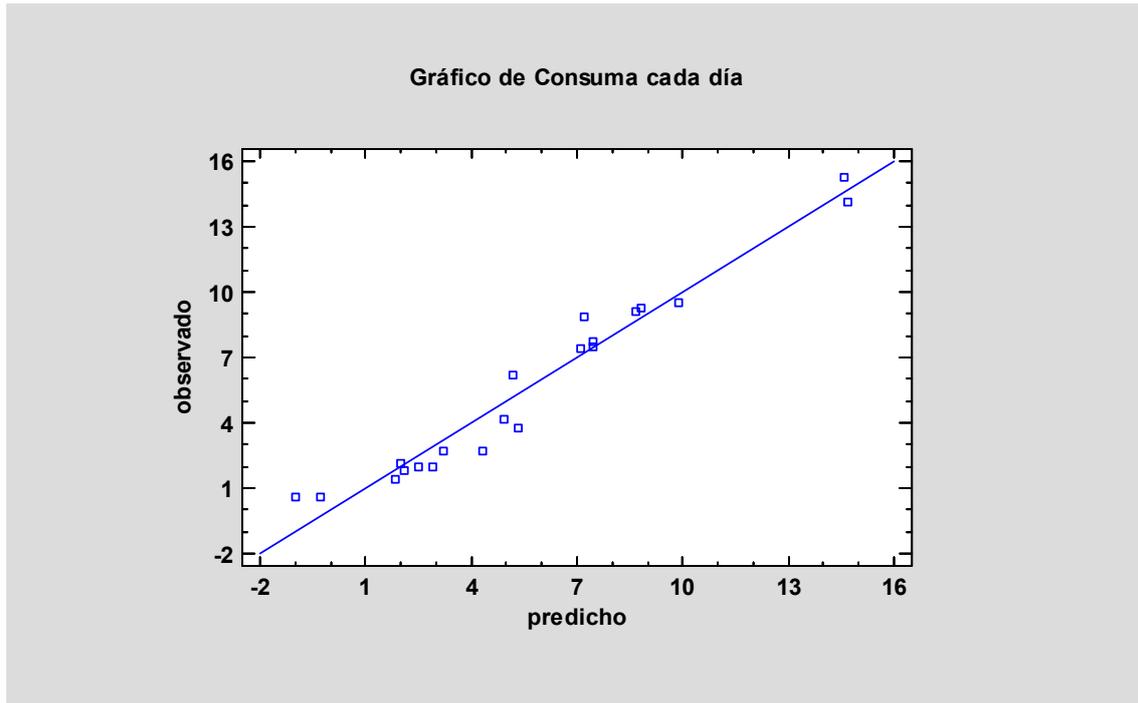


Gráfico 34: Gráfico de datos predicho frente a datos observados (b)

Se ve claramente que tras esta eliminar esos puntos el modelo ha mejorado, así como su R^2 , que representa un 95,72%.

A continuación, se representará el modelo ajustado válido:

$$\begin{aligned}
 \text{Consumo cada día} = & -75,9804 - 0,0313616 \cdot (\text{PIB per capita}) - 0,55782 \cdot \\
 & \cdot (\text{Temperatura}) - 105,602 \cdot (\% \text{ estudios exc infantil}) + 0,626896 \cdot \\
 & \cdot (\text{Tasa de desempleo}) - 90486,3 \cdot (\% \text{ víctimas homicidio intencional}) - \\
 & -0,295901 \cdot (\text{Trabajos temporales}) + 3,90543 \cdot 10^{-8} \cdot (\text{Población}) + 1,42125 \cdot \\
 & \cdot (\text{Esperanza Vida}) - 21,8422 \cdot (\% \text{ inmigrantes})
 \end{aligned}$$

Validación del modelo

Significatividad del parámetro

Tabla 23: Parámetros del modelo.

PARÁMETROS	Valor-P
Constante	0,0007
PIB per capita	0,1141
Temperatura	0,0007
% estudios exc infantil	0,0000
Tasa de desempleo	0,0000
% víctimas homicidio intencional	0,0155
Trabajos temporales	0,0004
Población	0,0108
Esperanza vida	0,0001
% Inmigrantes	0,0048

Contraste β_0

$$\left. \begin{array}{l} H_0: \beta_0 = 0 \\ H_1: \beta_0 \neq 0 \end{array} \right\} \alpha = 0,05$$

$P_Valor=0,0007 < \alpha$, por lo tanto, rechazamos H_0 y concluimos que la constante es significativa.

Contraste β_1

$$\left. \begin{array}{l} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{array} \right\} \alpha = 0,05$$

$P_Valor=0,1140 > \alpha$, por lo tanto, aceptamos H_0 y concluimos que la variable PIB per cápita no es significativa.

Contraste β_2

$$\left. \begin{array}{l} H_0: \beta_2 = 0 \\ H_1: \beta_2 \neq 0 \end{array} \right\} \alpha = 0,05$$

$P_Valor=0,0007 < \alpha$, por lo tanto, rechazamos H_0 y concluimos que la variable Temperatura es significativa.

Contraste β_3

$$\left. \begin{array}{l} H_0: \beta_3 = 0 \\ H_1: \beta_3 \neq 0 \end{array} \right\} \alpha = 0,05$$

$P_Valor=0,0000 < \alpha$, por lo tanto, rechazamos H_0 y concluimos que la variable % estudios excluyendo infantil es significativa.

Contraste β_4

$$\left. \begin{array}{l} H_0: \beta_4 = 0 \\ H_1: \beta_4 \neq 0 \end{array} \right\} \alpha = 0,05$$

$P_Valor = 0,0000 < \alpha$, por lo tanto, rechazamos H_0 y concluimos que la variable Tasa de desempleo es significativa.

Contraste β_5

$$\left. \begin{array}{l} H_0: \beta_5 = 0 \\ H_1: \beta_5 \neq 0 \end{array} \right\} \alpha = 0,05$$

$P_Valor=0,0155 < \alpha$, por lo tanto, rechazamos H_0 y concluimos que la variable % víctimas de homicidio intencional es significativa.

Contraste β_6

$$\left. \begin{array}{l} H_0: \beta_6 = 0 \\ H_1: \beta_6 \neq 0 \end{array} \right\} \alpha = 0,05$$

$P_Valor= 0,0004 < \alpha$, por lo tanto, rechazamos H_0 y concluimos que la variable Trabajos temporales es significativa.

Contraste β_7

$$\left. \begin{array}{l} H_0: \beta_7 = 0 \\ H_1: \beta_7 \neq 0 \end{array} \right\} \alpha = 0,05$$

$P_Valor=0,0108 < \alpha$, por lo tanto, rechazamos H_0 y concluimos que la variable Población es significativa.

Contraste β_8

$$\left. \begin{array}{l} H_0: \beta_8 = 0 \\ H_1: \beta_8 \neq 0 \end{array} \right\} \alpha = 0,05$$

$P_Valor=0,0001 < \alpha$, por lo tanto, rechazamos H_0 y concluimos que la variable Esperanza de vida es significativa.

Contraste β_9

$$\left. \begin{array}{l} H_0: \beta_9 = 0 \\ H_1: \beta_9 \neq 0 \end{array} \right\} \alpha = 0,05$$

$P_Valor=0,0048 < \alpha$, por lo tanto, rechazamos H_0 y concluimos que la variable % inmigrantes es significativa.

Significatividad del modelo

Tabla 24: Contraste del modelo.

FUENTE	Valor-P
Modelo	0,000

$$\left. \begin{array}{l} H_0: \beta_1=\beta_2=\beta_3=\beta_4=\beta_5=\beta_6=\beta_7=\beta_8=\beta_9=0 \\ H_1: \text{Algún } \beta_i \neq 0 \end{array} \right\} \alpha = 0,05$$

$P_Valor=0,0000 < \alpha$, por lo tanto, se rechaza la hipótesis nula y concluimos que el modelo es significativo.

Normalidad de los residuos

Gráfico Probabilístico

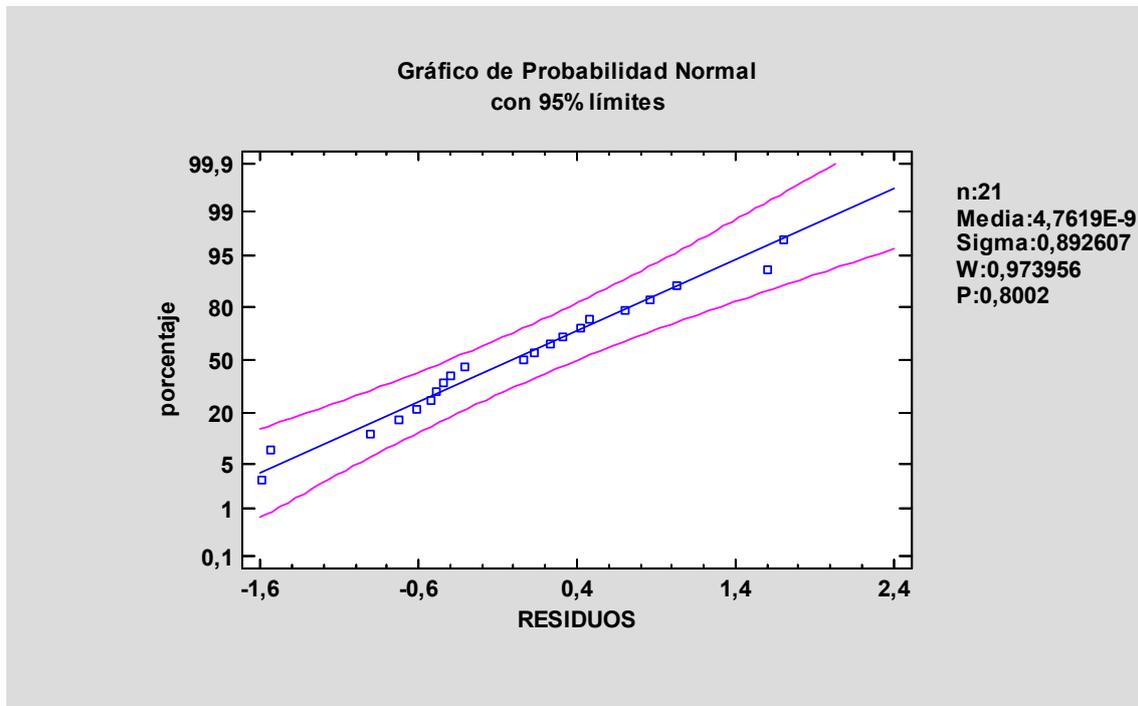


Gráfico 35: Gráfico probabilístico de residuos.

En este caso, parece que los Residuos siguen una distribución normal, ya que los puntos se sitúan en la línea de normalidad, hay que destacar también que ninguno de estos sale del intervalo de confianza que son las dos líneas rosas.

Histograma

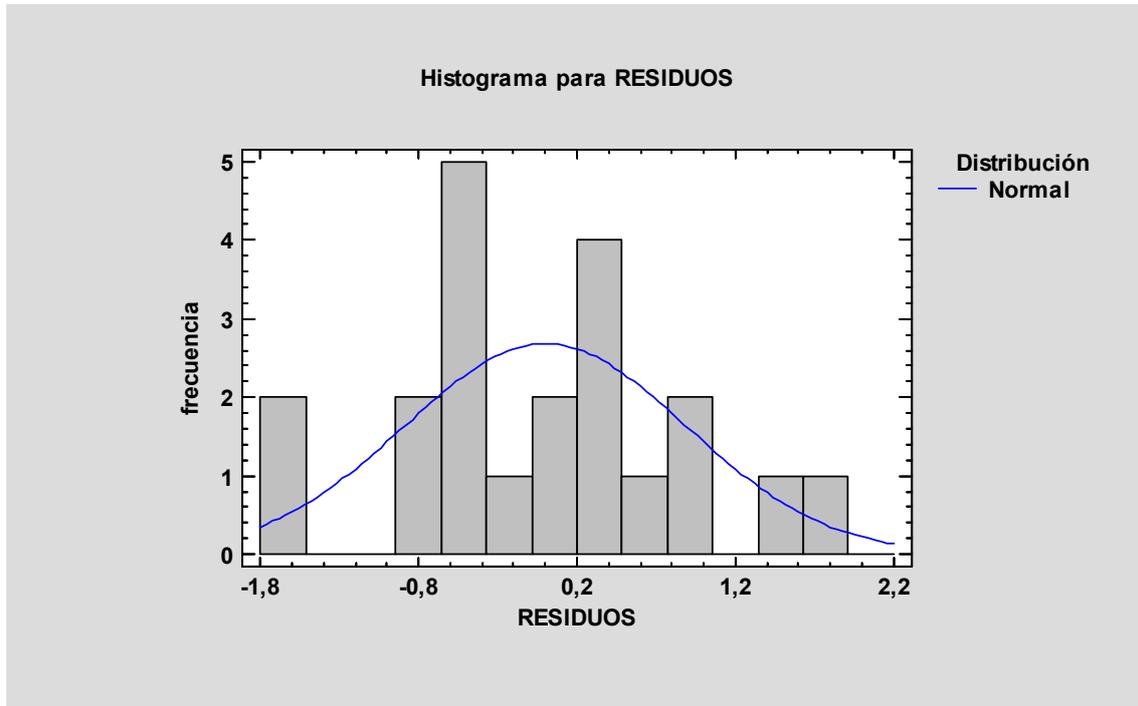


Gráfico 36: Histograma de residuos.

Al observar el histograma se puede decir que no parece que siga una distribución normal, ya que hay varias lagunas en las colas y en el centro, así como también hay picos muy significativos que no están en la media.

Test de normalidad

Tabla 25: Tabla de los test de normalidad.

PRUEBA	Valor-P
Chi-Cuadrado	0,738859
Estadístico W de Shapiro-Wilk	0,800166
Valor-Z para asimetría	0,826365
Valor-Z para curtosis	0,914957

$$\left. \begin{array}{l} H_0: \text{Los residuos son normales} \\ H_1: \text{Los residuos no son normales} \end{array} \right\} \alpha = 0,05$$

$0,738859 > \alpha$, por lo tanto, aceptamos la H_0 y se concluye que los residuos son Normales.

Heterocedasticidad

$$\begin{aligned} \text{Residuos}^2 = & 4,59965 - 0,00617803 \cdot \text{PIB per cápita} + 0,091299 \cdot \text{Temperatura} - \\ & 22,2252 \cdot (\% \text{ estudios exc infantil}) + 0,0189143 \cdot \\ & \cdot \text{Tasa de desempleo} - 31542,8 \cdot (\% \text{ víctimas homicidio intencional}) - \\ & -0,0453753 \cdot \text{Trabajos temporales} - 2,48974 \cdot 10^{-9} \cdot \text{Población} - 0,09105 \cdot \\ & \cdot \text{Esperanza vida} - 8,04361 \cdot (\% \text{ inmigrantes}) \end{aligned}$$

Tabla 26: Tabla de los P-Valores de cada variable.

PARÁMETRO	Valor-P
PIB per capita	0,6560
Temperatura	0,3259
% estudios exc infantil	0,0622
Tasa de desempleo	0,7744
% víctimas homicidio intencional	0,2041
Trabajos temporales	0,3183
Población	0,7964
Esperanza vida	0,6161
% Inmigrantes	0,1068

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P_Valor 0,6560 > α aceptamos H_0 y concluimos que la variable PIB per cápita no genera problemas de heterocedasticidad.

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P_Valor 0,3259 > α aceptamos H_0 y concluimos que la variable Temperatura no genera problemas de heterocedasticidad.

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P_Valor 0,0622 > α aceptamos H_0 y concluimos que la variable % de estudios excluyendo infantil no genera problemas de heterocedasticidad.

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P_Valor 0,7744 > α aceptamos H_0 y concluimos que la variable Tasa de desempleo no genera problemas de heterocedasticidad.

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P_Valor $0,2041 > \alpha$ aceptamos H_0 y concluimos que la variable % de víctimas de homicidios intencionados no genera problemas de heterocedasticidad

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P_Valor $0,3183 > \alpha$ aceptamos H_0 y concluimos que la variable Trabajos temporales no genera problemas de heterocedasticidad

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P_Valor $0,7964 > \alpha$ aceptamos H_0 y concluimos que la variable Población no genera problemas de heterocedasticidad

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P_Valor $0,6161 > \alpha$ aceptamos H_0 y concluimos que la variable Esperanza de vida no genera problemas de heterocedasticidad

$$\left. \begin{array}{l} H_0: \sigma^2 = \text{Constante} \\ H_1: \sigma^2 \neq \text{Constante} \end{array} \right\} \alpha = 0,05$$

P_Valor $0,1068 > \alpha$ aceptamos H_0 y concluimos que la variable % de inmigrantes no genera problemas de heterocedasticidad

Autocorrelación

FAS

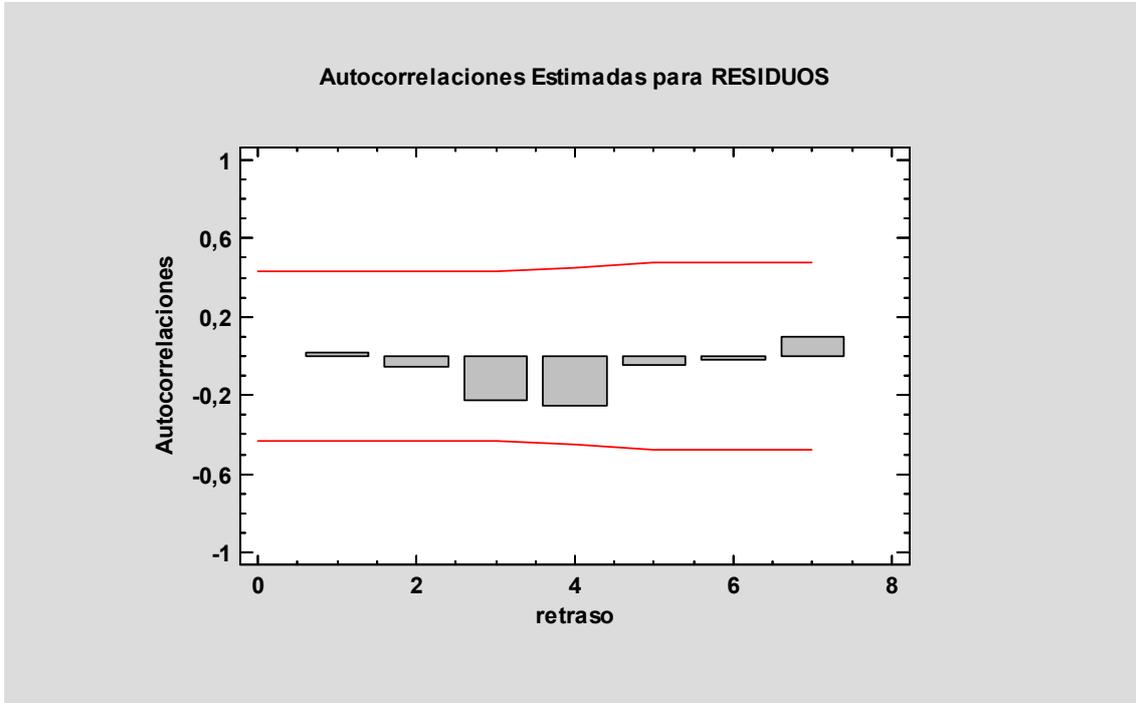


Gráfico 37: FAS.

FAP

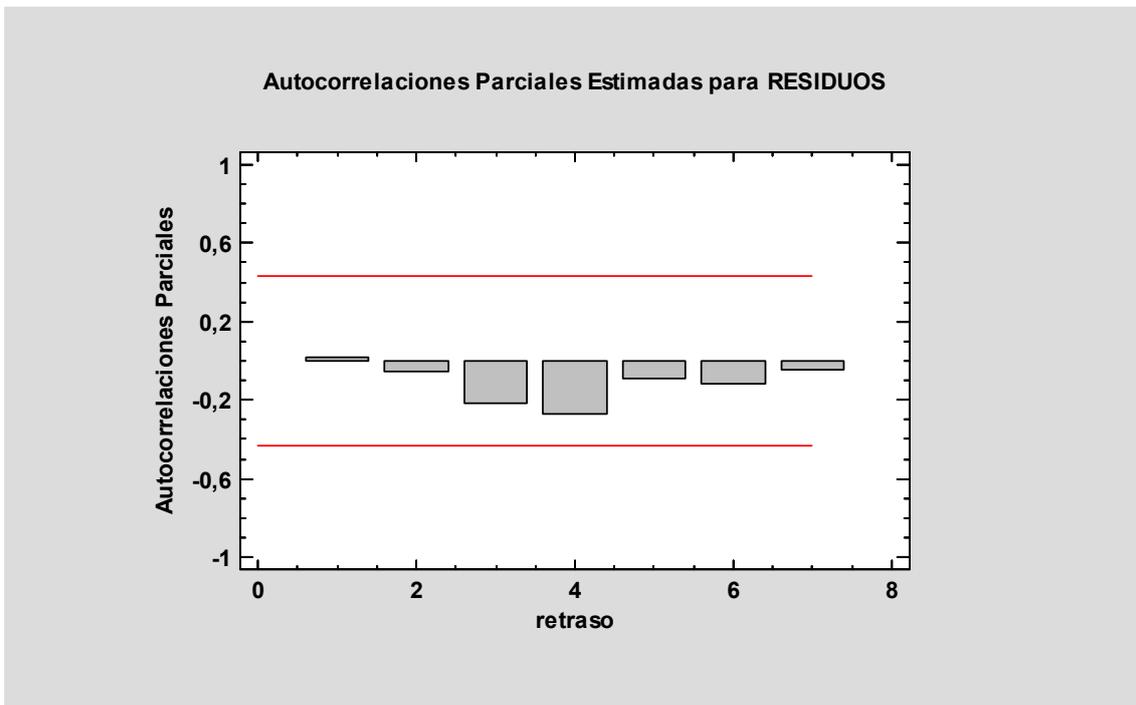


Gráfico 38: FAP.

Puesto que ningún factor de autocorrelación, en ambos gráficos, supera los límites, se puede concluir que no hay autocorrelación de ningún orden.

Test de autocorrelación (Estadístico de Durbin-Watson)

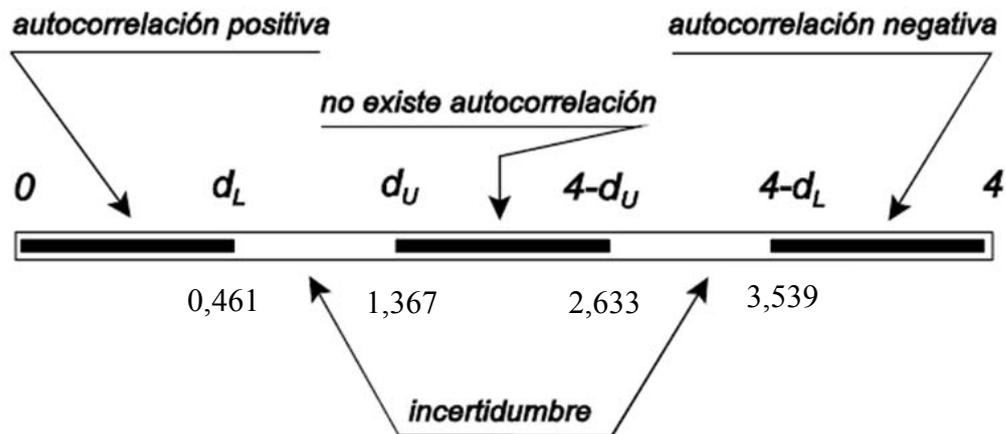


Gráfico 39: Prueba de Durbin-Watson.

Estadístico Durbin-Watson = 1,78379 (P=0,2086)

En este caso se puede decir que no existe autocorrelación, ya que si situamos el valor del estadístico de Durbin Watson (1.78379) en el gráfico anterior, se ve claramente que se situaría en la zona donde no existe autocorrelación. También se puede confirmar la no existencia de esta, mediante **el contraste de hipótesis**, que se detallará a continuación.

$$\left. \begin{array}{l} H_0: \rho_1 = 0 \\ H_1: \rho_1 \neq 0 \end{array} \right\} \alpha = 0,05$$

P_valor 0,2086 > α aceptamos H_0 y concluimos que no existen problemas de autocorrelación de 1º orden.

Puntos anómalos y puntos influyentes

Puntos anómalos

Tabla 27: Tabla residuo estudentizado.

FILA	Residuo Estudentizado
1	2,01
20	2,97

Tanto Bulgaria (fila 1) como Islandia (fila 20), se consideran puntos anómalos, aunque no son candidatos para eliminar, ya que su residuo no es mayor a 3.

Puntos influyentes

Tabla28: Puntos influyentes.

FILA	INFLUENCIA	DFITS
1	0,369438	1,53821
11	0,743556	2,555
14	0,855443	-1,62337
15	0,472883	-1,87224
20	0,654015	4,08133

Influencia media de un solo punto = 0,47619

Tabla 29: Cálculos del límite de puntos influyentes a priori y a posteriori.

A priori	$n_{ii} > 2\bar{n}$	0,95
A posteriori	$ DFIS \geq 2\sqrt{\bar{n}}$	1,38

- **Puntos influyentes a priori**

Tabla 30: Tabla resumen de los puntos influyentes a priori.

PROVINCIA	FILA	n_{ii}	
Bulgaria	1	$0,369438 < 2\bar{n}$	No es un punto influyente a priori
Lituania	11	$0,743556 < 2\bar{n}$	No es un punto influyente a priori
Polonia	14	$0,855443 < 2\bar{n}$	No es un punto influyente a priori
Rumania	15	$0,472883 < 2\bar{n}$	No es un punto influyente a priori
Islandia	20	$0,654015 < 2\bar{n}$	No es un punto influyente a priori

- **Puntos influyentes a posteriori**

Tabla 31: Tabla resumen de los puntos influyentes a posteriori.

PROVINCIA	FILA	DFITS	
Bulgaria	1	1,53821>2vñ	Sí es un punto influyente a posteriori
Lituania	11	2,555>2vñ	Sí es un punto influyente a posteriori
Polonia	14	1,62337>2vñ	Sí es un punto influyente a posteriori
Rumania	15	1,87224>2vñ	Sí es un punto influyente a posteriori
Islandia	20	4,08133>2vñ	Sí es un punto influyente a posteriori

Solución modelo

Se quita la variable PIB per cápita, ya que no era significativa, con este nuevo modelo planteado se comprueba que todas las variables son significativas:

$$\begin{aligned} \text{Consumo de alcohol cada día} = & \beta_0 + \beta_1 \cdot \text{Temperatura} + \\ & + \beta_2 \cdot (\% \text{estudios excluyendo infantil}) + \beta_3 \cdot \text{Tasa de desempleo} + \beta_4 \cdot \\ & \cdot \text{Víctimas homicidio intencional} + \beta_5 \cdot \text{Trabajos temporales} + \beta_6 \cdot \\ & \cdot \text{Población} + \beta_7 \cdot \text{Esperanza de vida} + \beta_8 \cdot (\% \text{Inmigrantes}) + U \end{aligned}$$

Tabla 32: Tabla con los P-Valores para cada una de las variables.

PARÁMETROS	Valor-P
Constante	0,0013
Temperatura	0,0017
% estudios exc infantil	0,0000
Tasa de desempleo	0,0000
% víctimas homicidio intencional	0,0243
Trabajos temporales	0,0009
Población	0,0212
Esperanza vida	0,0001
% Inmigrantes	0,0022

Tras esta solución el modelo final presenta un R^2 de 94,572% y queda de la siguiente forma:

$$\begin{aligned} \text{Consumo cada día} = & -60,6269 - 0,502979 \cdot \text{Temperatura} - 109,878 \cdot \\ & \cdot (\% \text{ estudios exc infantil}) + 0,6673 \cdot \text{Tasa de desempleo} - 87726,3 \cdot \\ & \cdot (\% \text{ víctimas homicidio intencional}) - 0,270422 \cdot \\ & \cdot \text{Trabajos temporales} + 3,60859 \cdot 10^{-8} \cdot \text{Población} + 1,19251 \cdot \\ & \cdot \text{Esperanza de vida} - 24,8876 \cdot (\% \text{ inmigrantes}) \end{aligned}$$

Predicciones

Partiendo del modelo final, se realizarán seis predicciones, para saber cuál será el consumo de alcohol en España en 2014.

Primera Predicción: Aumento de la variable % estudios exc infantil de 0,21 a 0,30

Tabla 33: Predicción puntual y pronóstico inferior y superior de la variable porcentaje de habitantes con cualquier tipo de estudios excluyendo infantil.

	Ajustado	Inferior 95,0%	Superior 95,0%
España	4,87505	0,157804	9,59231

- Predicción puntual

Consumo cada día = 4.87505

- Predicción por intervalo

$$[0,157804 \leq \text{Consumo cada día} \leq 9.59231]_{95\%}$$

Por lo tanto, se concluye que si el gobierno lograra aumentar el porcentaje de cualquier tipo de estudios excluyendo el infantil a un 30% la tasa de consumo de alcohol bajaría considerablemente pasando de un 15,3% a un 4,88%. Éste hecho favorecería claramente al nivel cultural de España, pero perjudicaría notablemente a los supermercados, bares, o cualquier establecimiento que venda alcohol.

Segunda Predicción: Diminución de la variable tasa de desempleo de 23,6 a 18

Tabla 34: Predicción puntual y pronóstico inferior y superior de la variable tasa de desempleo.

	Ajustado	Inferior 95,0%	Superior 95,0%
España	11,0272	7,76392	14,2905

- Predicción puntual

Consumo cada día = 11,0272

- Predicción por intervalo

$$[7,76392 \leq \text{Consumo cada día} \leq 14,2905]_{95\%}$$

Si la tasa de desempleo lograra bajar a un 18% la tasa de consumo de alcohol también lo haría, situándose en un 11,03%. España se vería favorecida, ya que al bajar la tasa de desempleo significaría que hay mayor cantidad de población trabajando, éste hecho favorece directamente al consumo de las personas en bienes y servicios. Como hecho perjudicial sería el mismo comentado anteriormente.

Tercera Predicción: Disminución de la variable % víctimas homicidio intencionado de 0,000007 a 0,000001

Tabla 35: Predicción puntual y pronóstico inferior y superior de la variable porcentaje de víctimas por homicidio intencional.

	Ajustado	Inferior 95,0%	Superior 95,0%
España	15,2904	11,6637	18,9172

- Predicción puntual

Consumo cada día = 15,2904

- Predicción por intervalo

$$[11,6637 \leq \text{Consumo cada día} \leq 18,9172]_{95\%}$$

Se ve claramente que, aunque se intentara bajar el porcentaje de víctimas de homicidios intencionados de un 0,0007% a 0,0001% la tasa de consumo de alcohol cada día apenas cambiaría, ya que pasaría de 15,3 a 15,3.

Cuarta Predicción: Aumento de la variable trabajos temporales de un 24,1 a 26

Tabla 36: Predicción puntual y pronóstico inferior y superior de la variable trabajos temporales.

	Ajustado	Inferior 95,0%	Superior 95,0%
España	14,2503	10,6618	17,8388

- Predicción puntual

Consumo cada día = 14,2503

- Predicción por intervalo

$$[10,6618 \leq \text{Consumo cada día} \leq 17,8388]_{95\%}$$

En el hipotético caso de que los trabajos aumentaran a un 26%, el consumo de alcohol cada día disminuiría, esto se debe principalmente a que actualmente existe bastante paro, por lo que ese pequeño aumento de trabajos temporales significaría que bastantes personas han salido de una situación precaria, aunque sea temporalmente.

Quinta Predicción: Aumento de la variable población de 46440099 habitantes a 50000000

Tabla 37: Predicción puntual y pronóstico inferior y superior de la variable población.

	Ajustado	Inferior 95,0%	Superior 95,0%
España	14,8925	11,2723	18,5128

- Predicción puntual

Consumo cada día = 14,8925

- Predicción por intervalo

$$[11,2723 \leq \text{Consumo cada día} \leq 17,5128]_{95\%}$$

Si el gobierno lograra poder aumentar la población, esto se podría hacer mediante incentivos, como por ejemplo “cheque bebe”, a padres primerizos o a familias con más de un hijo, el consumo de alcohol disminuiría afectando así a todos los establecimientos que vendan alcohol.

Sexta predicción: Aumento de la variable % inmigrantes de un 0,126 a un 0,2

Tabla 38: Predicción puntual y pronóstico inferior y superior de la variable porcentaje de inmigrantes.

	Ajustado	Inferior 95,0%	Superior 95,0%
España	12,9224	9,33065	16,5142

- Predicción puntual

Consumo cada día = 12,9224

- Predicción por intervalo

$$[9,33065 \leq \text{Consumo cada día} \leq 16,5142]_{95\%}$$

Si el porcentaje de inmigrantes aumentara el consumo de alcohol cada día bajaría notoriamente de un 15,3% a un 12,92%. Esto parece no tener ninguna lógica, ya que aunque vengan más inmigrantes los españoles no dejarán de beber.

Conclusiones y Futuras líneas de investigación

Conclusiones

- Se han encontrado varios problemas de multicolinealidad, el primero de enfrentaba a la variable PIB y la variable Población donde finalmente PIB era eliminado del modelo. En segundo lugar, también existían problemas entre las variables porcentaje de estudios excluyendo infantil y porcentaje de estudios universitarios, esto significa que ambas variables explican lo mismo, por lo tanto, se decide eliminar porcentaje de estudios universitarios.
- Al hacer el modelo econométrico se eliminan los siguientes países: Portugal, Bélgica, Dinamarca, Hungría y Eslovaquia ya que estos se comportan de forma distinta al resto de los países.
- Para conseguir un modelo válido se ha tenido que eliminar PIB per cápita porque esta variable no era significativa.
- El modelo explica un 94.57% de los datos.
- Para poder disminuir el consumo de alcohol diario hay que subir la temperatura, el porcentaje de habitantes con estudios excluyendo infantil, los trabajos temporales y los inmigrantes.
- Para disminuir el consumo de alcohol diario hay que disminuir la tasa de desempleo y la esperanza de vida.
- Hay que destacar también las variables que afectan menos al consumo de alcohol diario son: el porcentaje de víctimas de homicidios intencionados y la población.

Futuras líneas de investigación

Las futuras líneas de investigación podrían ser:

- Este proyecto al haber sido realizado solamente a través de la estructura lineal, los futuros trabajos podrían realizarse mediante distintas estructuras, como podría ser: semilogarítmica, logarítmica, inversa y logarítmica inversa.
- Desarrollar un proyecto idéntico, pero teniendo en cuenta datos más actualizados, de esta forma se podrá ver si las predicciones se cumplen.
- En este proyecto, los problemas de la multicolinealidad se resuelven mediante el método de la eliminación de las variables, por el cual se pierde información y datos que podrían llegar a ser importantes, por lo tanto, en las futuras líneas de investigación se podrá utilizar el método de componentes principales.
- Efectuar un estudio interno de los países que fueron eliminados: Portugal, Bélgica, Dinamarca, Hungría y Eslovaquia. De esta forma se podrán conocer más datos sobre los factores internos que los puede llevar a consumir alcohol cada día.

Bibliografía

- [1] <https://www.ig.com/es/glosario-trading/definicion-de-pib>
- [2] <http://economipedia.com/definiciones/renta-pib-per-capita.html>
- [3] <http://economipedia.com/definiciones/tasa-de-desempleo-paro.html>
- [4] http://www.ilo.org/global/topics/non-standard-employment/WCMS_536960/lang-es/index.htm
- [5] <https://deconceptos.com/ciencias-naturales/esperanza-de-vida>
- [6] <https://www.significados.com/inmigrante/>
- [7] http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=hlth_ehis_a1e&lang=en
- [7] http://ec.europa.eu/eurostat/statistics-explained/images/c/c9/GDP_at_current_market_prices%2C_2006_and_2014-2016_YB17.png
- [7] http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=crim_hom_soff&lang=en
- [7] <https://www.datosmacro.com/demografia/homicidios>
- [7] http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=educ_uoe_enra01&lang=en
- [7] <http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do>
- [7] <http://www.belgica.org/clima-en-belgica/>
- [7] <https://es.climate-data.org/location/6424/>
- [7] https://es.wikipedia.org/wiki/Geograf%C3%ADa_de_la_Rep%C3%ABlica_Checa
- [7] <https://www.101viajes.com/dinamarca/tiempo-dinamarca>
- [7] <https://www.elclima-enelmundo.com/2013/09/clima-estonia-europa-info.html>
- [7] <https://www.ireland.com/es-es/about-ireland/discover-ireland/irelands-weather/>
- [7] <https://www.guiaviajes.org/grecia-clima/#>
- [7] <http://www.aemet.es/es/noticias/2015/01/climatico2014>
- [7] <https://www.guiaviajes.org/croacia-clima/>
- [7] <https://es.climate-data.org/location/1185/>
- [7] <https://www.guiaviajes.org/chipre-clima/>
- [7] <https://www.guiaviajes.org/letonia-clima/>
- [7] <https://www.guiaviajes.org/lituania-clima/>
- [7] <http://www.luxemburgo.net/clima-de-luxemburgo/>
- [7] <https://www.guiaviajes.org/hungria-clima/>
- [7] <http://www.viajarmalta.com/clima-en-malta.php>
- [7] <https://www.guiaviajes.org/austria-clima/>
- [7] <https://www.guiaviajes.org/polonia-clima/>
- [7] [https://www.pordata.pt/Portugal/Temperatura+m%C3%A9dia+do+ar+\(m%C3%A9dia+anual\)-1067](https://www.pordata.pt/Portugal/Temperatura+m%C3%A9dia+do+ar+(m%C3%A9dia+anual)-1067)
- [7] <https://www.guiaviajes.org/rumania-clima/>
- [7] <https://www.guiaviajes.org/eslovenia-clima/>
- [7] <https://www.guiaviajes.org/eslovaquia-clima/>
- [7] <https://www.guiaviajes.org/finlandia-clima/>
- [7] <https://www.guiaviajes.org/suecia-clima/>
- [7] <https://es.climate-data.org/location/22/>
- [7] <https://www.guiaviajes.org/islandia-clima/>

- [7] <https://www.guiaviajes.org/noruega-clima/>
- [7] <https://www.guiaviajes.org/turquia-clima>
- [7] <https://www.datosmacro.com/paro?anio=2014>
- [7] <https://www.datosmacro.com/demografia/esperanza-vida>
- [7] <https://www.datosmacro.com/demografia/migracion/inmigracion>
- [8] <http://www.eumed.net/libros-gratis/2006c/203/2o.htm>
- [9] <https://prezi.com/pucotkzhuy7/medidas-de-simetria-o-sesgo/>
- [10] <http://www.universoformulas.com/estadistica/descriptiva/curtosis/>
- [11] <http://economipedia.com/definiciones/heterocedasticidad.html>
- [12] Chirivella V (2005) <<Econometría>>. En: P.I.D 12059-C, Valencia, España