



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

***Estudio de los factores  
comerciales de la empresa  
Sollutia S.L. mediante modelos  
de regresión***

---

**MEMORIA PRESENTADA POR:**

*José María Pastor Nebot*

GRADO DE ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESAS

Convocatoria de defensa: Noviembre de 2016

## ÍNDICE

1.	Introducción y objetivos.....	4
2.	Metodología y Técnicas.....	6
2.1.	Datos .....	6
2.2.	Variables.....	6
2.2.1.	Variable dependiente .....	6
2.2.2.	Variables independientes.....	6
2.3.	Metodología .....	7
3.	Cálculos, Resultados, y Explotación .....	11
3.1.	Análisis univariante de las variables cuantitativas .....	11
3.1.1.	Análisis de Una Variable - Beneficio Anuales .....	11
3.1.2.	Análisis de Una Variable - Gasto servidor anual .....	13
3.1.3.	Análisis de Una Variable - Ganancias de hosting .....	15
3.1.4.	Análisis de Una Variable - Coste dominio/s .....	17
3.1.5.	Análisis de Una Variable - Dominios contratados .....	19
3.1.6.	Análisis de Una Variable - Valor de la web.....	20
3.1.7.	Análisis de Una Variable - productos contratados anualmente.....	22
3.2.	Análisis univariante de las variables cualitativas .....	24
3.2.1.	Tipo de hosting.....	24
3.2.2.	Tipo de dominio .....	25
3.2.3.	Tipo web.....	25
3.3.	Análisis bivariante variables cuantitativas .....	26
3.3.1.	Regresión Simple - Beneficio Anuales vs. Gasto servidor anual .....	26
3.3.2.	Regresión Simple - Beneficio Anuales vs. Coste hosting.....	27
3.3.3.	Regresión Simple - Beneficio Anuales vs. Coste dominio/s .....	28
3.3.4.	Regresión Simple - Beneficio Anuales vs. Dominios contratados .....	29
3.3.5.	Regresión Simple - Beneficio Anuales vs. Valor de la web.....	30
3.3.6.	Regresión Simple - Beneficio Anuales vs. Productos contratados anualmente..	31
3.4.	Análisis multicolinealidad.....	31
3.4.1.	Análisis Multivariado .....	31
3.4.2.	Análisis Multivariado 2 .....	33
3.4.3.	Análisis Multivariado 3 .....	35



3.4.4.	Conclusión multicolinealidad .....	36
3.5.	Formulación del modelo teórico propuesto .....	37
3.6.	Regresión múltiple .....	38
3.6.1.	Validación del modelo.....	39
3.6.1.1.	Regresión múltiple .....	39
3.6.1.2.	Significadidad .....	39
3.6.1.3.	Normalidad de residuos .....	41
3.6.1.4.	Heterocedasticidad .....	43
3.6.1.5.	Autocorrelación.....	43
3.6.1.6.	Puntos anómalos y puntos influyentes .....	45
3.6.1.7.	Conclusión regresión múltiple.....	45
3.6.1.8.	Regresión Múltiple - Beneficio Anuales .....	48
3.6.1.9.	Predicción .....	49
4.	Conclusiones y futuras líneas de investigación .....	51
4.1.	Futuras líneas de investigación .....	51
5.	Bibliografía .....	53
6.	Anexo .....	54
	Servicios que Sollutia Ofrece.....	54
	Diseño web.....	54
	Hosting y dominios.....	54
	Marketing online.....	55
	[1] Tabla 6.1 Datos de los clientes de Sollutia S.L. Fuente: La empresa Sollutia.....	56

## 1. Introducción y objetivos

La empresa que se ha escogido se denomina Sollutia, S.L. es una de las empresas de Alcoy con más experiencia en sector de internet, ya que llevan dedicándose desde el año 1999 al desarrollo de soluciones integrales de Internet para empresas e instituciones. Está compuesta por un equipo profesional multidisciplinario, y fue constituida por Don Sergio Grau García. Se dedican principalmente al asesoramiento, asistencia y responsabilidad de los proyectos, desde la fase de diseño web hasta su completa realización y puesta en marcha.

Las razones para elegirlos las dejan muy claras en la web:

1. **Cercanía y confianza.** Ya que llevan 17 años de experiencia en el mundo de internet que lo avalan todos sus clientes y la cercanía del servicio y de sus profesionales que garantizan la satisfacción de cada cliente.
2. **Productos personalizados.** No venden webs a kilos, estudian cada caso y ofrecen la mejor solución basándose en programación propia, sin depender de sistemas de terceros que limiten sus posibilidades.
3. **Atención al cliente.** Saben lo importante que es sentirse atendido por personas que entiendan las necesidades de cada cliente. Por lo tanto piden que se pongan en contacto por teléfono, email o personalmente en sus oficinas para solucionar cualquier cosa lo antes posible.
4. **Ponen todos sus servicios en el aire para ayudar a cualquier cliente o futuro cliente.** Desde diseño web hasta el alojamiento web pasando por la maquetación y el servicio de correo electrónico.

Realizan proyectos de calidad y personalizados para cada cliente. La empresa se encuentra estrechamente conectada con el cliente para que su página web sea el reflejo del producto que vende.

Por ello la filosofía de esta empresa consiste en:

**"La buena atención al cliente es la garantía y**

**Confianza de un trabajo bien hecho."**



No obstante la empresa requiere una labor constante de investigación a fin de intentar optimizar sus recursos para mejorar un servicio de los que ofrece y para ello deseamos saber qué servicio nos otorga más beneficios.

Por ello, el objetivo y motivación de esta investigación econométrica para el trabajo fin de grado se trata de plasmar de qué servicio proviene más beneficios y de donde proceden los beneficios anuales de la empresa Sollutia, S.L. mediante una muestra aleatoria de 50 clientes de la empresa.

El trabajo se ha elaborado con información y datos otorgados por la empresa. Estos datos son totalmente confidenciales, es decir, no se dirá en ningún momento el nombre del cliente, solo podemos decir que se ha escogido la muestra entre empresas grandes, medianas y pequeñas.

Este estudio se basara en la obtención de un modelo de regresión, para poder definir de donde provienen beneficios anuales por cliente, según los productos contratados. Se ha estudiado la empresa y conseguido obtener distintas variables, aunque finalmente se han escogido 6 variables explicativas y 3 variables ficticias.

## 2. Metodología y Técnicas

### 2.1. Datos

Para la realización de este trabajo, los datos que se han utilizado provienen de la empresa Sollutia. La muestra de datos empleada abarca desde enero a diciembre de 2015.

Los programas informáticos utilizados para la elaboración del modelo han sido Microsoft Excel, que se ha utilizado para organizar los datos obtenidos de las bases de datos SIS (Sistema integrado de Sollutia) y el programa Statgraphics, para realizar el tratamiento estadístico de esos datos.

### 2.2. Variables

#### 2.2.1. Variable dependiente

La variable dependiente o endógena que se va emplear en el modelo se trata de los beneficios anuales de la empresa, surgen de los pagos de los clientes por los servicios que Sollutia les proporciona. Esta es la definición de la variable:

- **Beneficios Anuales:** Consisten en los cobros que le otorgan los clientes por servicios prestados.

Cabe destacar que consiste de una muestra de 50 clientes escogidos al azar, que estamos en época de crisis, que existe más competencia, y que anteriormente los beneficios eran mayores a los de hoy en día.

#### 2.2.2. Variables independientes

A continuación, se va a describir las variables explicativas o independientes que previamente se han considerado para elaborar el modelo econométrico más adecuado, y estudiar cuales de ellas son las que más afectan a los beneficios de la empresa.

Las variables son las siguientes:

- **Gastos Servidor Anual:** Toda página web debe estar alojada en un servidor para que funcione y por ello, cada cliente que tiene una web debe pagar su espacio en el servidor, dependiendo cuanta información o tipo de web ocupan más o menos.

- **Coste hosting/s:** Es el coste que le repercute al cliente por tener un servicio técnico 24 horas al día, los 365 días al año, este servicio consiste en que por ejemplo no caiga la web en ningún momento, servicio antispam y que le funcione correctamente su correo en cualquier minuto, ya que si hablamos de negocios por internet en cualquier momento puede enviar un cliente potencial un pedido y si falla se podría perderlo.
- **Coste dominio/s:** Es el coste que cliente paga a la empresa para la adquisición, renovación y administración de cualquier dominio.
- **Números de dominios contratados:** Esta variable consiste en tener en cuenta la cantidad de dominios que tiene cada empresa contratada.
- **Tipo de dominios contratados:** Consiste en tener en cuenta los diferentes tipos de variables de dominios.
- **Tipo de web contratada:** Esta variable diferencia los 4 tipos de web que existe en el mercado desde una web informativa hasta una tienda virtual.
- **Valor de la web:** Cada tipo de web tiene un valor diferente, sobre todo si se escoge el tipo de web 3 ya que es a medida y cada proyecto es diferente.
- **Productos contratados:** Con esta variable contabilizaremos los productos contratados por cada cliente anualmente.
- **Tipo hosting:** Esta variable diferencia los 6 tipos de espacio que existe en la empresa.

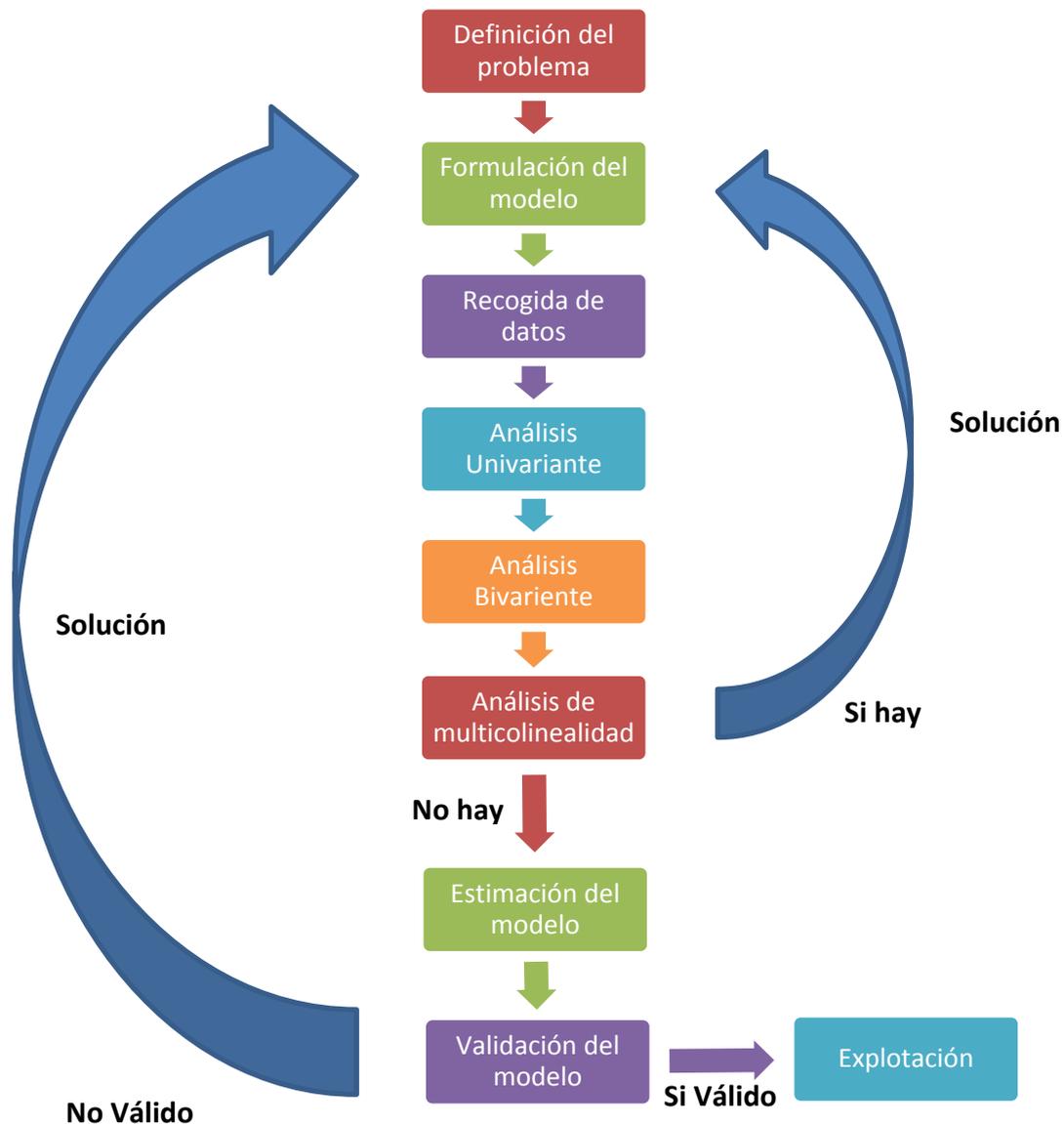
Todas las variables explicativas del modelo las he extraído de la base de datos de la empresa Sollutia S.L.

### 2.3. Metodología

El objetivo del análisis de la regresión lineal es analizar un modelo que pretende explicar el comportamiento de una variable (Variable endógena, o dependiente), que denominaremos por el nombre de Y, utilizando la información proporcionada por los valores tomados por un conjunto de variables (explicativas, exógenas o independientes), que denotaremos por  $X_1$ ,  $X_2$ , ...,  $X_n$ .

La metodología o plan de trabajo que seguiremos en el análisis estadístico de un modelo de regresión simple consiste en varias partes como podemos ver en el **Gráfico 2.1** la estructura de un análisis de regresión para poder valorar si un modelo es válido o no es válido.

Inicialmente debemos observar la situación de la empresa y que deseamos estudiar de ella, en este caso deseamos saber qué beneficios podemos obtener en la situación que un cliente desee contratar los servicios de la empresa Sollutia, S.L.



**Gráfico 2.1.** Esquema de regresión. **Fuente:** Elaboración propia

Después se formulará mediante las variables escogidas un modelo de regresión. A continuación, se reunirán los datos numéricos que se han escogido como muestra, se adjuntarán a una hoja de Excel para su ordenación y cómo podemos ver en la **tabla 6.1. Anexo [1]** más adelante esta tabla la volcaremos en el programa Statgraphics Centurion XVI para empezar el estudio que se desea hallar. Y así, se podrán obtener y utilizar herramientas estadísticas que nos permitan tomar decisiones objetivas y razonables, ya que se necesita que el modelo se ajuste a unas determinadas hipótesis.

Más tarde se hará un análisis de cada variable individualmente y esto se realizará mediante el análisis univariante. Este análisis se centra en dos aspectos en la tendencia de la distribución y en su dispersión, para poder analizar se extraerá de Statgraphics un resumen estadístico de cada variable cuantitativa, un diagrama de caja y bigotes y un histograma. Para las variables cualitativas se va a coger los datos individualmente de cada y vamos a colocar gráficamente en forma circular.

Posteriormente del análisis univariante se continuará con otro análisis que consiste en el análisis bivariante. Se utiliza para estudiar el efecto de una variable frente otra variable, para llevar a cabo este estudio se debe hacer una regresión múltiple, y así se podrá saber si existe relación entre ellas o no. Ya que cuanto más relación exista más explicara la variable exógena a la variable endógena.

A continuación del anterior análisis se procederá a hacer el análisis de multicolinealidad que consiste en una situación en la que se presenta una fuerte correlación entre variables explicativas del modelo.

Posteriormente se hará la propuesta de una formulación para resolver el modelo teórico en el que se basa mediante la variable dependiente y las variables independientes.

Seguidamente se realizará un análisis de significación, que consiste en analizar cada variable independiente para saber si afectan verdaderamente a la variable dependiente, así se valorará si la variable vale la pena tenerla para que explique el modelo o eliminarla por no explicar el modelo muy poco.

Más tarde se hará el análisis de normalidad de residuos que consiste en el objetivo de buscar relaciones entre las variables mediante pruebas paramétricas, es necesario comprobar si las variables en cuestión cumplen con los requisitos necesarios para este tipo de pruebas, serían:

- Distribución normal de los residuos
- Uniformidad de la varianza
- Independencia de los datos

Este análisis se guía mediante la prueba Chi-cuadrado. Consecutivamente del análisis anterior se hará el análisis para comprobar la existencia de heterocedasticidad en las perturbaciones de  $U_i$ , ya que esto implica que la varianza de las perturbaciones no sea constante, y el incumplimiento de una de la hipótesis del método de los mínimos cuadrados ordinarios, MCO.

Más adelante se continuará el análisis mediante las pruebas de Autocorrelación, que ayudan a percibir la relación entre las perturbaciones. Se dice que en caso de existir autocorrelación, deja de cumplirse una de las hipótesis del modelo de mínimos ordinarios.

Definitivamente el último análisis consistirá en analizar los puntos anómalos y puntos influyentes para poder saber definitivamente que puntos tendríamos que obviar y que puntos son influyentes en nuestro modelo.

Finalmente se realizará una regresión múltiple que nos proporcionará definitivamente la formulación del modelo para hacer una predicción y así poder dar definitivamente unas conclusiones.

### 3. Cálculos, Resultados, y Explotación

#### 3.1. Análisis univariante de las variables cuantitativas

##### 3.1.1. Análisis de Una Variable - Beneficio Anuales

#### Resumen Estadístico para Beneficio Anuales

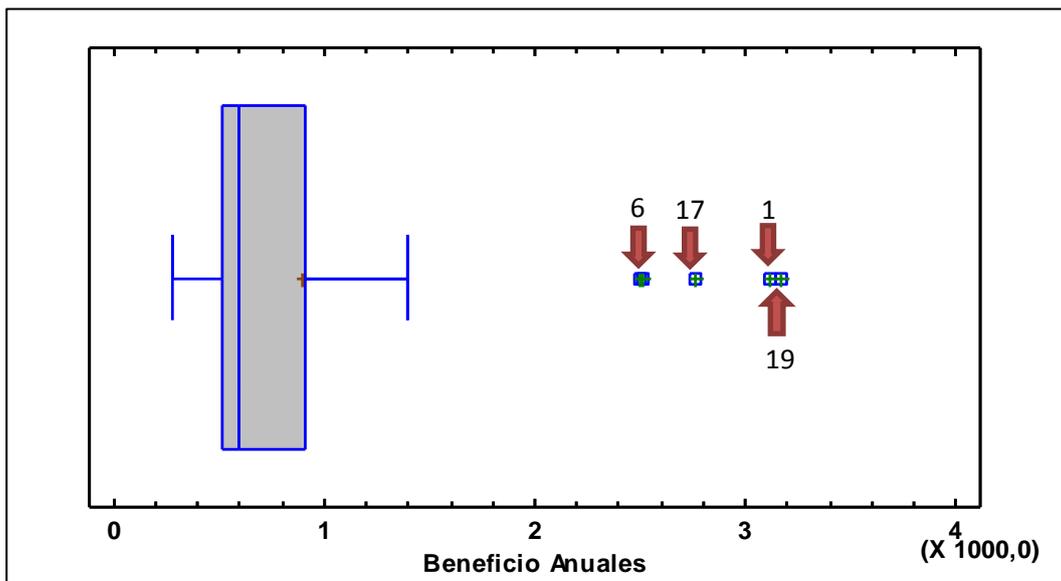
Los datos que a continuación se comentan, se trata de los beneficios anuales de la empresa Sollutia S.L. Estos datos se consideran como la **Variable Y**, esta variable se mide en euros. Como se observa en la **tabla 3.1**, se ha hecho un análisis univariante de la variable cuantitativa "**Beneficios Anuales**". Se puede dar relevancia a algunos estadísticos como puede ser la media que se encuentra en unos 890€ y que se encuentra a la par de la desviación típica. También podemos observar que los rangos del máximo y mínimo oscilan entre 276€ y 3.174,50€.

Datos	
Promedio	891,596
Mediana	599,25
Varianza	501305,
Desviación Estándar	708,029
Coficiente de Variación	79,4114%
Mínimo	276,0
Máximo	3174,5
Rango	2898,5
Sesgo Estandarizado	6,3327
Curtosis Estandarizada	5,79714

**Tabla 3.1.** Resumen Estadístico para Beneficio Anuales. **Fuente:** Elaboración propia

#### Caja y bigotes

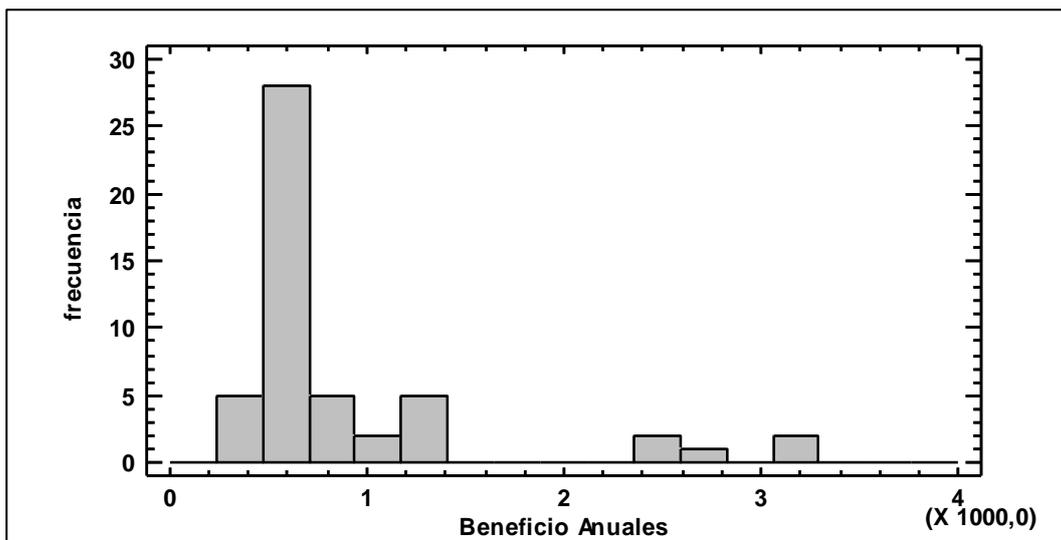
Los diagramas de caja y bigotes son muy útiles al hacer comparaciones gráficas entre conjuntos de datos, ya que tienen un gran impacto visual y son fáciles de comprender. Como se puede observar en el gráfico 3.1, existen varios puntos anómalos. Solo comentaremos los que más apartados se encuentra como es el caso de la empresa 1 que otorga un volumen de beneficios de 3.116,5 € al año y la empresa 19 que nos concede un beneficio de 3.174,5 € al año. Esto sucede, porque son empresas grandes y con un gran volumen de ventas por internet, cosa que las otras empresas son más pequeñas tienen un volumen pequeño en venta por internet o simplemente solo quieren otorgar información de ellas en la red. Por ello, nos permiten obtener de ellas unos beneficios entre 500€ y 1.000€ al año.



**Gráfico 3.1.** Gráfico de cajas y bigotes de Beneficios Anuales. **Fuente:** Elaboración propia

## Histograma

Un histograma representa la distribución de frecuencias de una variable cuantitativa continua. Se construye representando sobre cada intervalo un rectángulo de área proporcional a la frecuencia de la categoría. Cuando los intervalos tienen igual longitud (opción más usual), la altura del rectángulo es proporcional a la frecuencia. Como podemos ver en el gráfico 3.2 existe un comportamiento común entre la mayoría de las empresas, excepto unas cuantas que se sobresalen de las demás, esto ocurre ya que generan beneficios más grandes que las otras.



**Gráfico 3.2.** Histograma de Beneficios Anuales. **Fuente:** Elaboración propia

### 3.1.2. Análisis de Una Variable - Gasto servidor anual

#### Resumen Estadístico para Gasto servidor anual

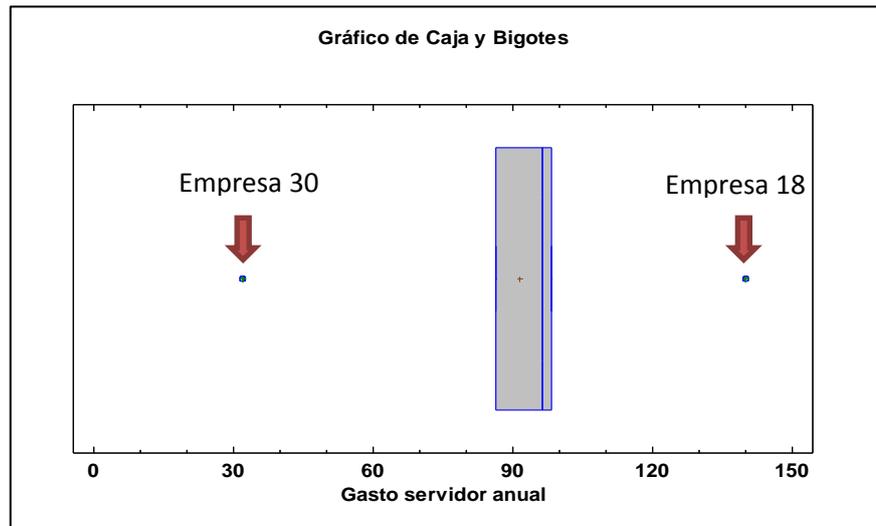
Los datos que se pueden ver posteriormente en la **tabla 3.2** se tratan de los gastos que ocasionan al dar el servicio de servidor a las empresas, se trata de una **Variable  $X_1$** , como bien mencionamos en el “**apartado 2.2.2 variables independientes**” este servicio se debe ofrecer casi obligatoriamente, ya que si una web no está alojada en un servidor no puede estar en online y el cliente deberá buscar este servicio en la competencia. Estos datos se miden mediante euros. Mediante un análisis univariante de “**Gasto servidor anual**”, se puede dar relevancia algunos datos como puede ser la desviación típica que no es nada parecida al promedio y los rangos del máximo y mínimo oscilan entre 31,97€ y 140€.

Datos	
Promedio	92,3178
Mediana	96,44
Varianza	1103,88
Desviación Estándar	33,2247
Coefficiente de Variación	35,9894%
Mínimo	31,97
Máximo	140,0
Rango	108,03
Sesgo Estandarizado	-1,67235
Curtosis Estandarizada	-0,0749512

**Tabla 3.2.** Resumen Estadístico para Gasto servidores anuales. **Fuente:** Elaboración propia

#### Caja y bigotes

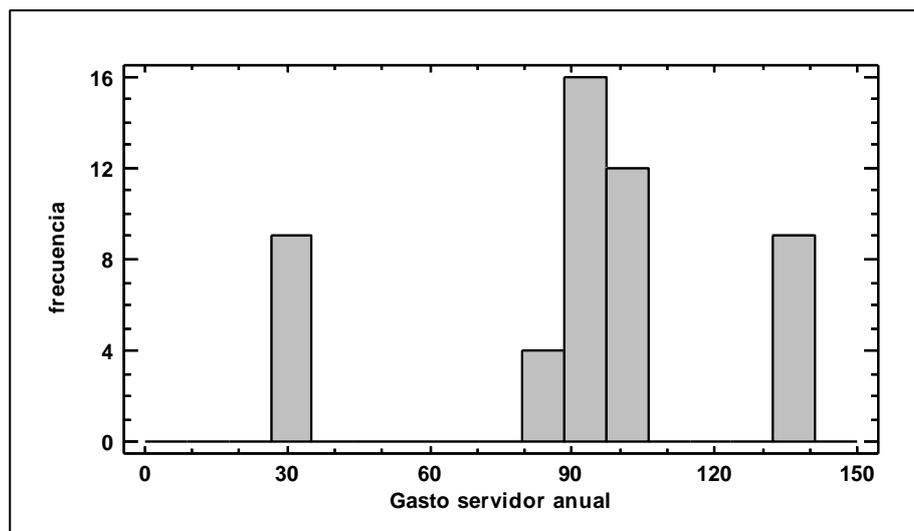
Como se puede observar en gráfico 3.3 existen varios puntos anómalos. Solo se comentará al igual que en el apartado anterior los puntos anómalos más apartados. En este gráfico se puede ver que el punto o empresa 30 se encuentra por debajo de la media, ya que se encuentran en un servidor más barato porque empresa tiene una página web simplemente informativa de tipo 1 lo que nos indica que esta página solo tendrá información y alguna imagen por ello necesita poco espacio (es una de las web más antiguas que tiene la empresa). En el lado contrario se ve que existe un punto como es la empresa 18 que se encuentran de manera viceversa a las anteriormente nombradas ya que necesitan un servidor más potente y más grande para cumplir sus necesidades porque estamos hablando de una web de tipo 3, que tiene mucha información, con accesos a usuarios donde estos pueden subir documentos e imágenes, con una galería de fotos, un blog...



**Gráfico 3.3.** Gráfico de cajas y bigotes de Gasto servidores anuales. **Fuente:** Elaboración propia

## Histograma

Como bien se ha visto anteriormente en la representación del gráfico 3.3, existen varios tipos de servidor pero existen generalmente 3 tipos de servidor que son los más contratados, por ello existe en el gráfico 3.4 varias áreas proporcionales con diferente longitud, como es un servidor pequeño que se produce un coste anual de 31,97€ al año donde estarán alojadas las web que necesitan menos espacio. Después se podrá observar este gráfico 3.4 que existe como una agrupación, ya que hay varios servidores que se encuentran alrededor entre los 80€ y 110€ al año que hay alojaríamos las web medianas que sería las de tipo 2 y finalmente se arrendaría un servidor muy potente con mucho espacio que se consideraría un coste de 140€ al año donde podríamos los proyectos web de tipo 3 y 4.



**Gráfico 3.4.** Histograma de Gasto servidores anuales. **Fuente:** Elaboración propia

### 3.1.3. Análisis de Una Variable - Ganancias de hosting

#### Resumen Estadístico para Ganancias de hosting

Los datos que se ven en la **tabla 3.3** se tratan del coste que nos ocasiona dar el servicio de asistencia técnica, este servicio lo obtiene el cliente cuando contrata el servicio hosting con la empresa, esta variable se puede definir qué cuanto mayor tipo de hosting contrate y menos tiempo de asistencia técnica anualmente más beneficios obtiene la empresa. Y por ello se ha nombrado la **Variable X<sub>2</sub>**, esta variable también la medimos euros.

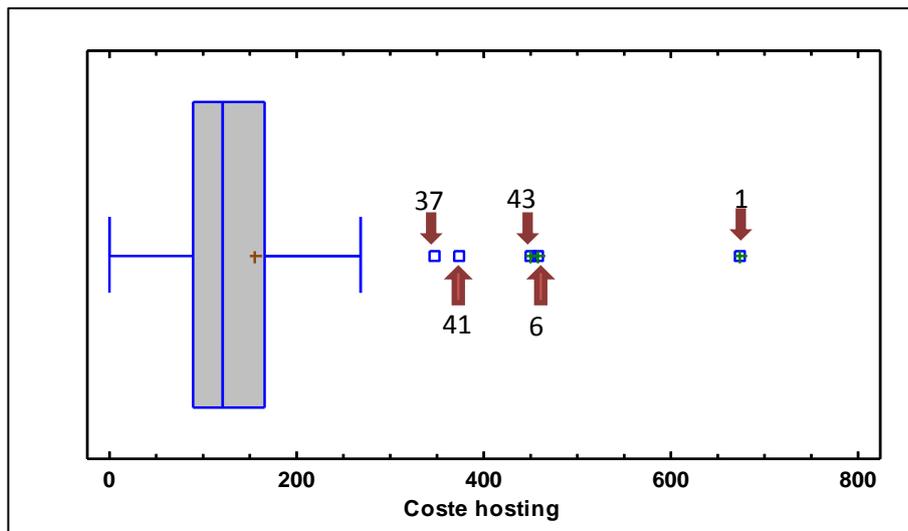
A través de un análisis univariante de la **Variable X<sub>2</sub>** ("**Ganancias de hosting**"), se hará una pequeña mención a los datos más relativos que vemos en la tabla 3.3 que como puede ser el **promedio** que se encuentra en 154,382 € y la **desviación típica** que se encuentra cerca. Y finalmente que los rangos máximos y mínimos oscilan entre el 0 y 674.

Datos	
Promedio	154,382
Mediana	120,5
Varianza	15015,6
Desviación Estándar	122,538
Coficiente de Variación	79,3734%
Mínimo	0
Máximo	674,0
Rango	674,0
Sesgo Estandarizado	6,66353
Curtosis Estandarizada	9,41758

**Tabla 3.3.** Resumen Estadístico para coste hosting. **Fuente:** Elaboración propia

#### Caja y bigotes

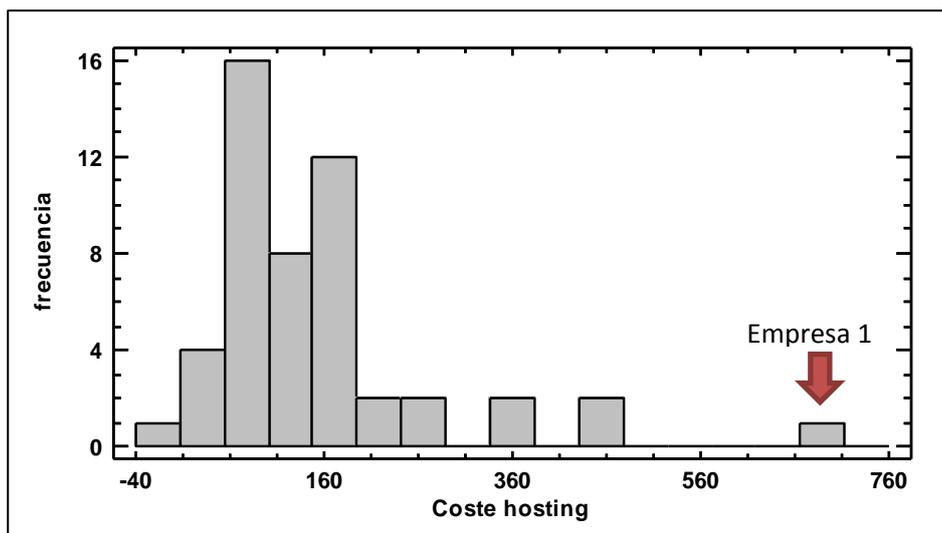
Si se observa el **gráfico 3.5** de cajas y bigotes se obtiene la información que existen varios puntos anómalos, se nombraran por orden de alejamiento de menor rango a mayor como es la empresa 37, 41, 43, 6 y 1 como uno datos más relevante se comentara que el punto más alejado de la media es la empresa 1 ya que consiste en una web de tipo 3 eso quiere decir que es una página web hecha a medida con mucha información sobre la empresa, una galería de fotos, con una intranet protegida mediante usuario y contraseñas, en ella los usuarios pueden subir fotos, documentación e incluso escribir mails o internamente pueden comunicarse mediante chats, por ello se le ofrece un hosting totalmente personalizado para él que se nombra "**Plan Prestigie**".



**Gráfico 3.5.** Gráfico de cajas y bigotes de coste hosting. *Fuente: Elaboración propia*

## Histograma

En el **gráfico 3.6** se observa que la altura del rectángulo es proporcional con más frecuencia. Como se puede observar existe un comportamiento común similar a una distribución normal entre todas las empresas, ya que la tarifa de hosting más vendida de la empresa es el plan profesional que tiene un coste para la empresa que lo contrata de 120€. Excepto varias empresas que se sobre salen, que comparando la representa en el gráfico 3.5 de cajas y bigote podemos ver como en todo momento la empresa 1 esta apartada.



**Gráfico 3.6.** Histograma de coste hosting. *Fuente: Elaboración propia*

### 3.1.4. Análisis de Una Variable - Coste dominio/s

#### Resumen Estadístico para Coste dominio/s

Como se puede ver en la tabla 3.4 de a continuación se puede observar la variable “**Coste de dominio**” que es **Variable  $X_3$**  se mide en euros y es una variable cuantitativa. Como en los puntos anteriores se ha hecho un análisis univariante y se ha hallado que existe una media de 106,005€, que justamente en este caso no se encuentra a la par de su desviación típica y un rango máximo y mínimo que oscila entre 29,5€ y 772,40€.

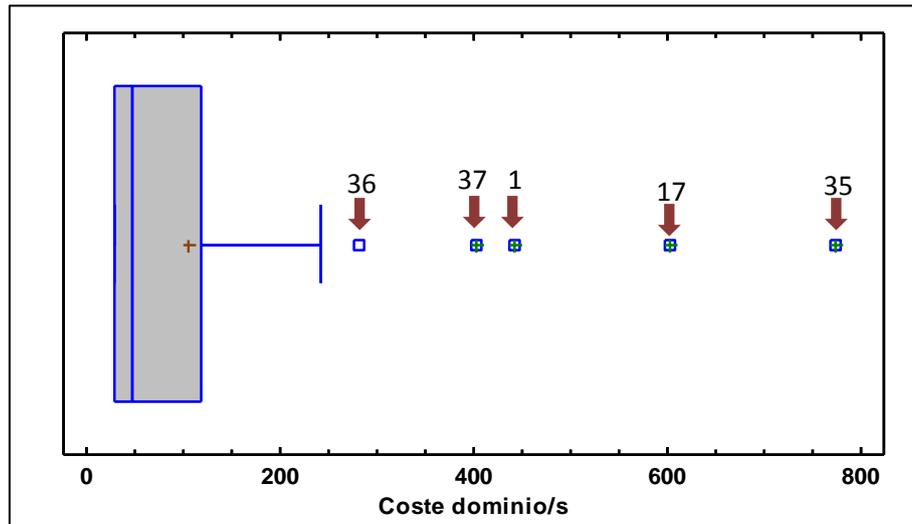
Datos	
Promedio	106,005
Mediana	48,5
Varianza	22938,7
Desviación Estándar	151,455
Coefficiente de Variación	142,876%
Mínimo	29,5
Máximo	772,4
Rango	742,9
Sesgo Estandarizado	8,47091
Curtosis Estandarizada	13,1513

*Tabla 3.4. Resumen Estadístico para coste dominio. Fuente: Elaboración propia*

#### Caja y bigotes

A juzgar por el **gráfico 3.7** de cajas y bigotes se obtiene que existan varios puntos anómalos, como son la empresa 1, 17, 35, 36 y 37. En este caso se nombraran las dos empresa más alejadas o anómalas son la empresa 35 que al año paga por sus 9 dominios 772,40€ y la empresa 17 que paga por 10 dominios 602,00€.

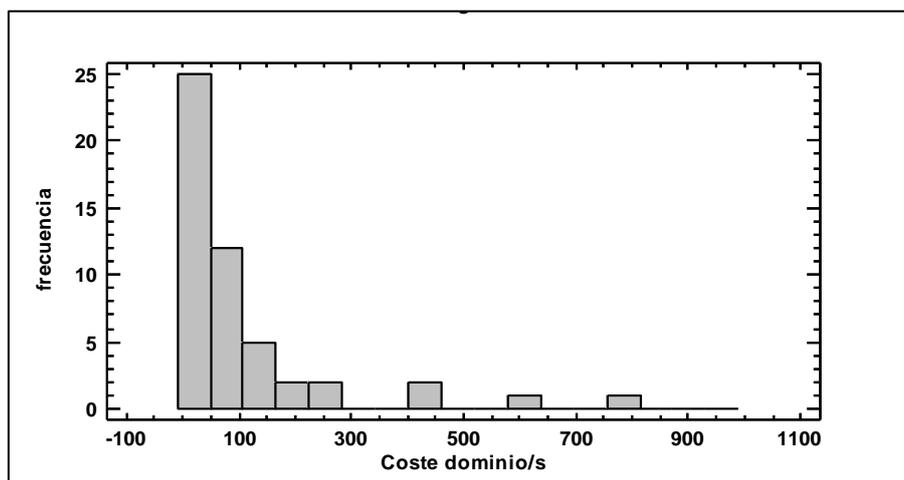
Se puede observar que la empresa 17 tiene contratado un domino más que la empresa 37 esto ocurre porque existen diferentes tipos de dominios y no tienen un precio estándar ya que por ejemplo el .com y .es que son los más comunes tienen un precio de 29,50€ cosa que dominios con diferentes características pueden valer más, como por ejemplo el dominio que está representa la unión europea como es el .eu este dominio puede llegar a valer hasta 50€, por no hablar de dominios más característicos como puede ser el .tv o el .video todos dominios pueden valer mucho más.



**Gráfico 3.7.** Gráfico de cajas y bigotes de coste dominio. *Fuente:* Elaboración propia

### Histograma

Como bien se observa en el histograma se representa esta variable con una la distribución asimétrica a la derecha. Vemos que el coste de dominio que más se repite es en un intervalo entre el 0 a 150 euros que representan empresas que tienen en su poder como máximo 2, a partir de ese valor las empresas pueden tener más de 2 dominios e incluso tener un dominio específico.



**Gráfico 3.8.** Histograma de coste dominio. *Fuente:* Elaboración propia

### 3.1.5. Análisis de Una Variable - Dominios contratados

#### Resumen Estadístico para Dominios contratados

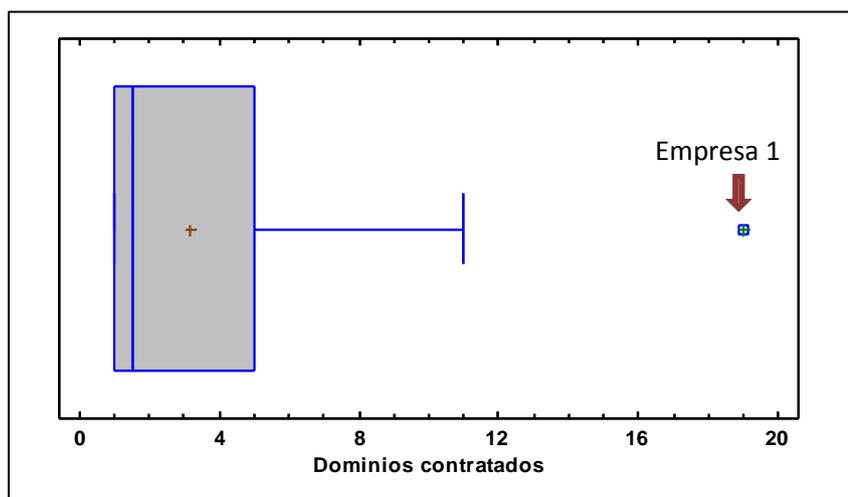
Como se puede observar en la tabla 3.5 se ha hecho un análisis univariante a la **Variable X<sub>4</sub>** y gracias a ello se ha conseguido obtener estos datos, si se repasa que los datos más importantes que se pueden encontrar en la tabla existe un promedio 3,2 podemos ver que se acerca a la par la desviación. También podemos ver que los rangos se encuentran como mínimo 1 y como máximo en 19.

Datos	
Promedio	3,2
Mediana	1,5
Varianza	12,6939
Desviación Estándar	3,56285
Coficiente de Variación	111,339%
Mínimo	1,0
Máximo	19,0
Rango	18,0
Sesgo Estandarizado	6,89092
Curtosis Estandarizada	10,2541

**Tabla 3.5.** Resumen Estadístico para dominio contratados. **Fuente:** Elaboración propia

#### Caja y bigotes

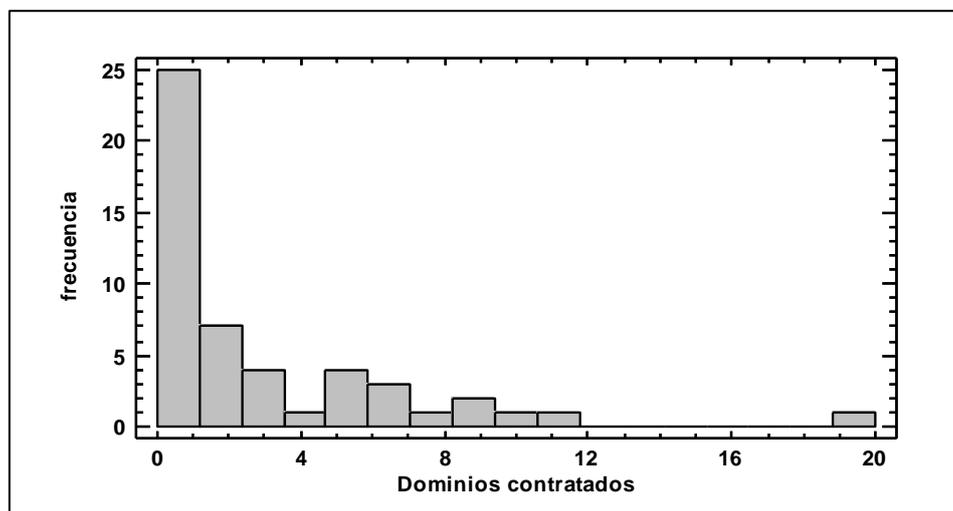
Como se puede ver en el diagrama de caja y bigotes todas las empresas contratan entre 1 y 5 dominios, se puede observar en gráfico 3.9 que existe solo un punto anómalo que es la empresa 1 que tiene contratado con la empresa Sollutia, S.L. 19 dominios diferentes.



**Gráfico 3.9.** Gráfico de cajas y bigotes de dominio contratados. **Fuente:** Elaboración propia

## Histograma

Como se puede ver en el gráfico 3.10 el histograma representa el volumen medio de dominios contratados por los clientes. Analizando visualmente el gráfico se puede observar como la moda en esta variable es tener 1 dominio que normalmente es el “.com” o “.es”, ya que observamos que la mitad de los clientes de la muestra tienen un dominio.



**Gráfico 3.10.** Histograma de dominios contratados. **Fuente:** Elaboración propia

### 3.1.6. Análisis de Una Variable - Valor de la web

#### Resumen Estadístico para Valor de la web

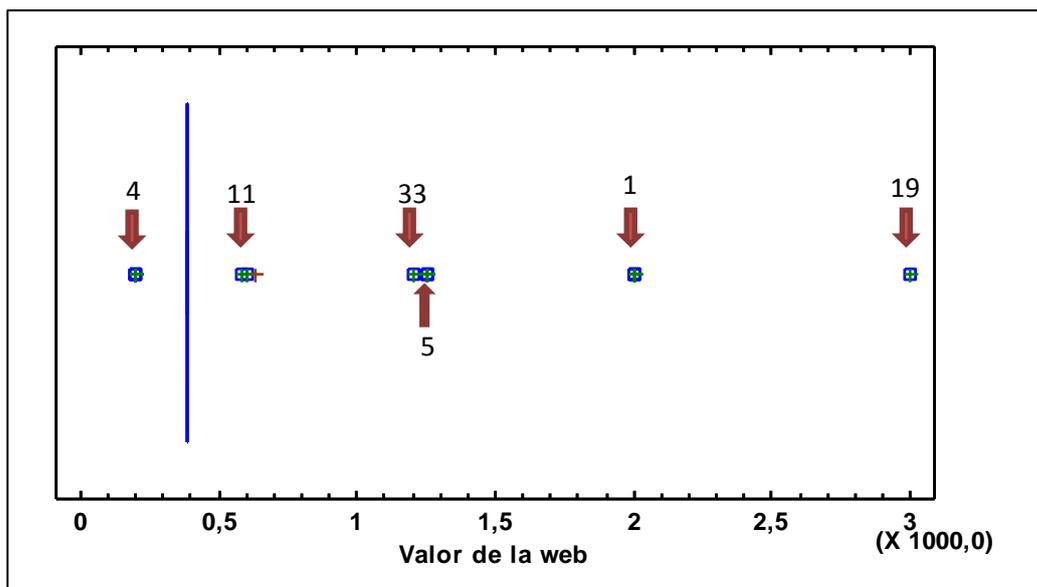
Como se observa en la tabla 3.6, se ha hecho un análisis univariante a los 50 datos de la variable cuantitativa “dominio contratados” la consideraremos como la variable  $X_5$  y la mediremos en euros y se ha hallado un promedio de 631,21 euros, como se puede observar en esta variable la desviación no concuerda con la media, esto es debido a que existe una dispersión de los datos muy elevada. También podemos observar que los rangos del máximo y mínimo oscilan entre 195€ y 2.995€.

Datos	
Promedio	631,21
Mediana	390,0
Varianza	354176,
Desviación Estándar	595,127
Coficiente de Variación	94,2835%
Mínimo	195,0
Máximo	2995,0
Rango	2800,0
Sesgo Estandarizado	6,85284
Curtosis Estandarizada	7,61962

**Tabla 3.6.** Resumen Estadístico para valor de la web. **Fuente:** Elaboración propia

## Caja y bigotes

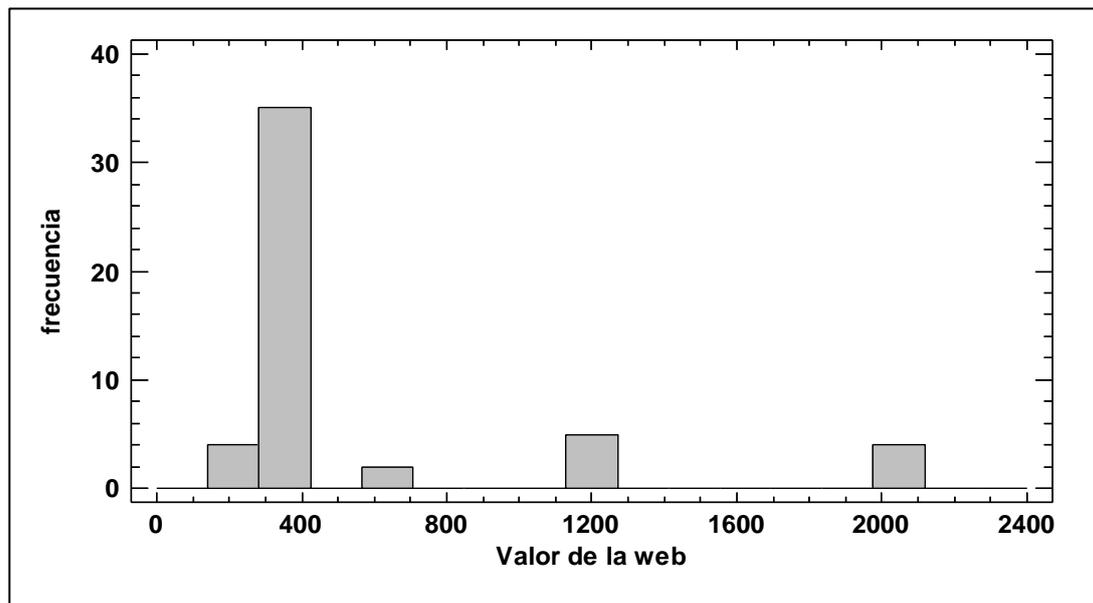
Como se puede observar en gráfico 3.11 en la muestra de estas 50 empresas existen varios puntos anómalos anualizando según su valor web como son las empresas 1, 4, 5, 11 y 32. Estas empresas se diferencian porque cada cliente tiene unas necesidades diferentes por eso existe un tipo general como pueden ser las web de tipo 1 y 2 que son páginas web informativas. Pero en el momento que el cliente escoge proyectos web como son webs de tipo 3 o 4 son proyectos a medida y llegan a tener un valor de más 800€.



**Gráfico 3.11.** Gráfico de cajas y bigotes de valor de la web. **Fuente:** Elaboración propia

## Histograma

Como se puede ver en el histograma del gráfico 3.12 está representando que existe una gran frecuencia en el valor de web en la muestra de las 50 empresas. Ya que el producto estrella de la empresa Sollutia son las páginas web de tipo 2, ya que son las más vendidas por su precio tan competitivo ya que tienen un valor aproximado de 400€ este precio ofrece un web totalmente personalizada, un panel de administración para la contratación de la web y así gestionar todos sus contenidos, también ofrece la posibilidad de hacer secciones ilimitadas en la web y la posibilidad de ampliación de diferentes módulos. Gracias a este gráfico se puede observar fácilmente representado.



**Gráfico 3.12.** Histograma de valor de la web. *Fuente:* Elaboración propia

### 3.1.7. Análisis de Una Variable - productos contratados anualmente

#### Resumen Estadístico para productos contratados anualmente

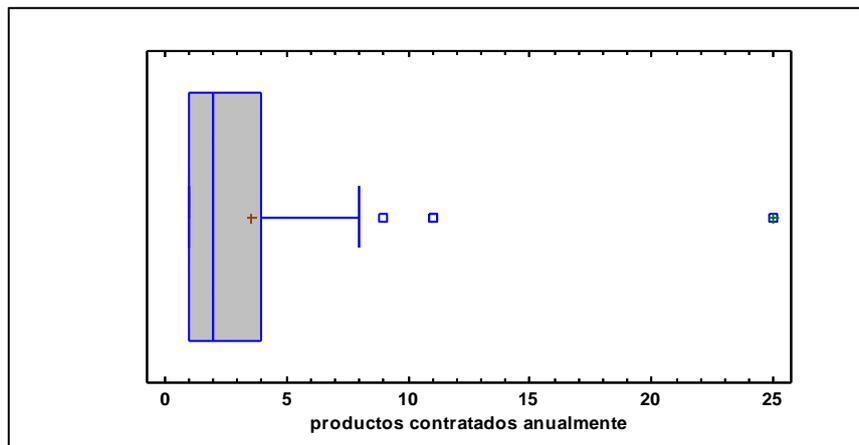
Como se observa en la tabla 3.7, se ha hecho un análisis univariante a los 50 datos de la variable cuantitativa “dominio contratados” que la consideraremos como la variable  $X_6$  y se ha hallado un promedio de 3,52 que se acerca a la desviación típica de la variable. Como se ha observado en las tablas anteriores podemos también tener en cuenta los rangos máximos y mínimos que fluctúan en 1 hasta 25.

Datos	
Promedio	3,52
Mediana	2,0
Varianza	16,9078
Desviación Estándar	4,1119
Coefficiente de Variación	116,815%
Mínimo	1,0
Máximo	25,0
Rango	24,0
Sesgo Estandarizado	9,64309
Curtosis Estandarizada	21,2552

**Tabla 3.7.** Resumen Estadístico para productos contratados anualmente. *Fuente:* Elaboración propia

## Caja y bigotes

Como se puede observar en gráfico 3.13 las empresas normalmente tienen varios productos contratados anualmente, habitualmente se podría decir que las empresas tienen como mínimo entre 2 o 3 productos ya que son necesarios que son la web, el hosting, el dominio existe varios puntos anómalos, como se puede ver la empresa 1, 17 y 34. Si se presta atención podemos ver gracias al gráfico que la empresa 1 es la empresa que más productos tienen contratados en la muestra.



**Gráfico 3.13.** Gráfico de cajas y bigotes de productos contratados anualmente. **Fuente:** Elaboración propia

## Histograma

Como bien hemos comentado anteriormente un histograma representa la distribución de frecuencias de una variable cuantitativa continua. Se construye representando proporcionalmente la altura del rectángulo es proporcional a la frecuencia. Como bien sea mencionado y se ha podido observar visualmente en el gráfico 3.13 las empresas normalmente tienen contratados 2 o 3 productos, por ello en gráfico 3.14 podemos ver otra forma gráfica que consigue nos explica la misma situación.

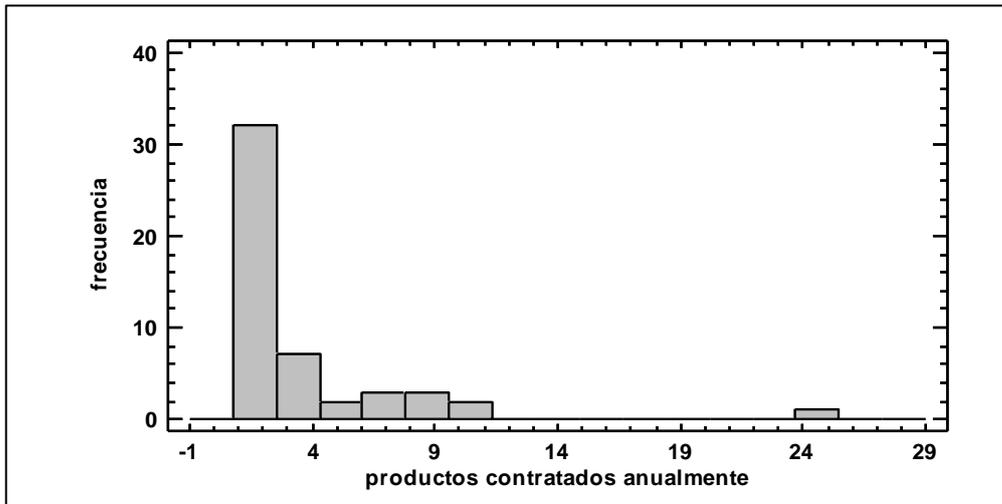


Gráfico 3.14. Histograma de productos contratados anualmente. Fuente: Elaboración propia

## 3.2. Análisis univariante de las variables cualitativas

### 3.2.1. Tipo de hosting

Como bien hemos comentado anteriormente sobre en el tema del espacio no todos los clientes de Sollutia tienen las mismas necesidades, por eso existen diferentes pack de hosting como es el de tipo 1 que tiene un espacio muy limitado o hasta el de tipo 6 que tiene una gran capacidad este tipo de producto no se usa para una página web de empresas sino para tiendas online con mucha variedad de productos. Como se puede ver en el gráfico 3.15 la variabilidad de los contratos de hosting que existen en la muestra observando el tanto por ciento de cada tipo de hosting.

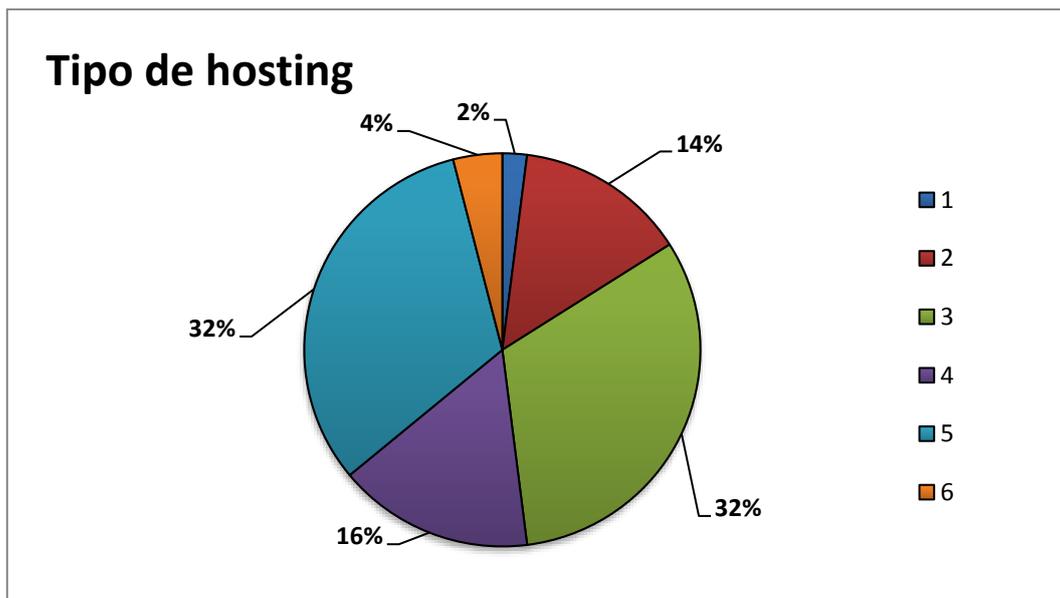
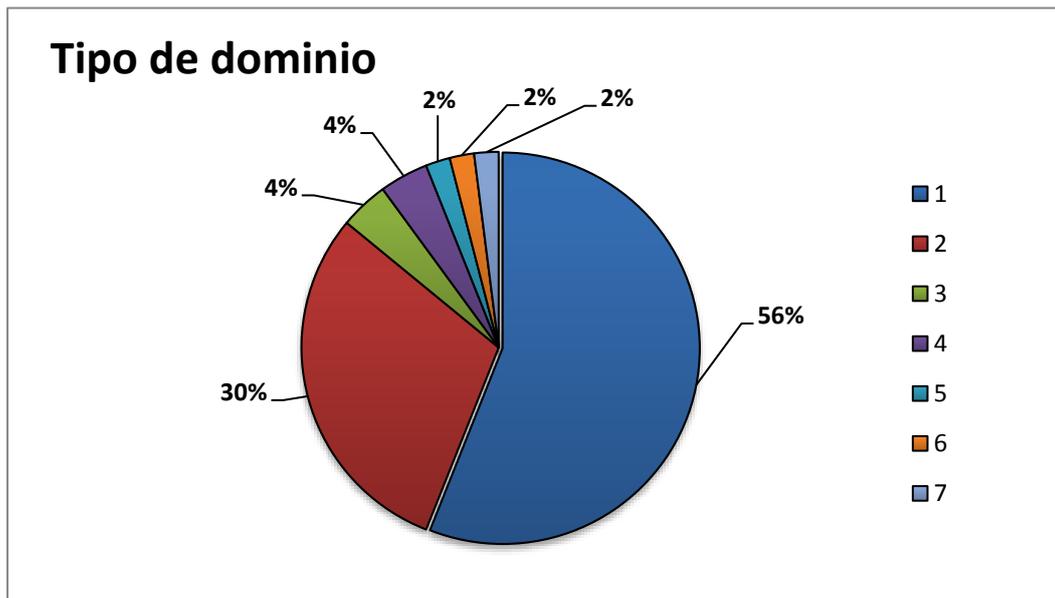


Gráfico 3.15. Tipos de hosting. Fuente: Elaboración propia

### 3.2.2. Tipo de dominio

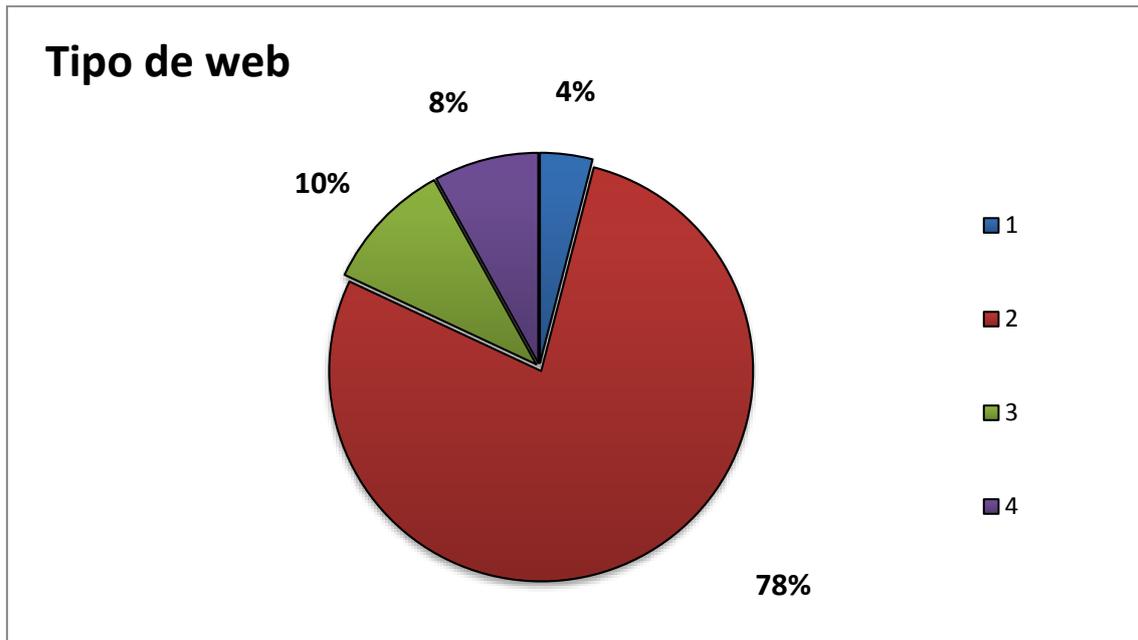
Observando el **gráfico 3.16** podemos interpretar que normalmente o la moda es cada cliente tiene un solo dominio. Aunque existe más de una empresa que si tiene más de uno, esto ocurre porque tienen otros negocios, hacen concursos, desean posicionar su página o simplemente quieren que su web que está alojada en Inglaterra o sirve para vender productos en ese país también tenga un dominio inglés como puede él “.uk” y es cuando ocurre que se tenga más de uno.



**Gráfico 3.16.** Tipo de dominio. *Fuente:* Elaboración propia

### 3.2.3. Tipo web

Teniendo en cuenta el **Gráfico 3.17** se ve que de las 4 modalidades que existen de webs, el producto 2 es el más vendido anualmente. Ya que es un tipo de web corporativo, es decir, es un producto bastante frecuente para cualquier empresa que desea empezar a estar en internet consiste en un producto económico, fácil de usar y que permite hacer casi todo lo que uno desea el cliente excepto la venta de productos online que entonces estaríamos hablando del producto 4 ya es un producto más específico y necesita una elaboración mucho más confeccionada hablando en términos de programación.



*Gráfico 3.17. Tipos de web. Fuente: Elaboración propia*

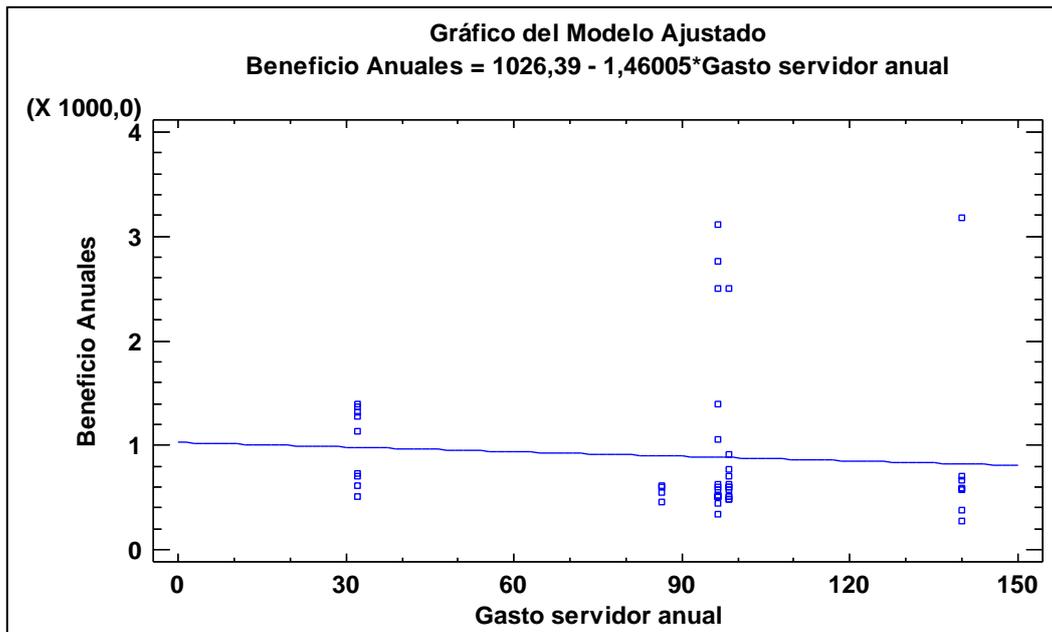
### 3.3. Análisis bivalente variables cuantitativas

Este apartado se centrará en los diagramas de dispersiones. Pueden representar cualquier combinación por parejas siempre manteniendo la variable dependiente y cambiando la variable independiente

#### 3.3.1. Regresión Simple - Beneficio Anuales vs. Gasto servidor anual

- **Variable dependiente:** Beneficio Anuales
- **Variable independiente:** Gasto servidor anual

Como se puede ver mediante este análisis bivalente entre la variable dependiente beneficios anuales y la variable independiente gastos del servidor anual, existe un coeficiente de correlación de  $-0,0685138$ , esto quiere decir que entre estas variables no existe relación entre ellas, aunque mediante el gráfico 3.18 podemos ver que existe una pequeña curva hacia la izquierda esto quiere decir pequeña relación indirecta.

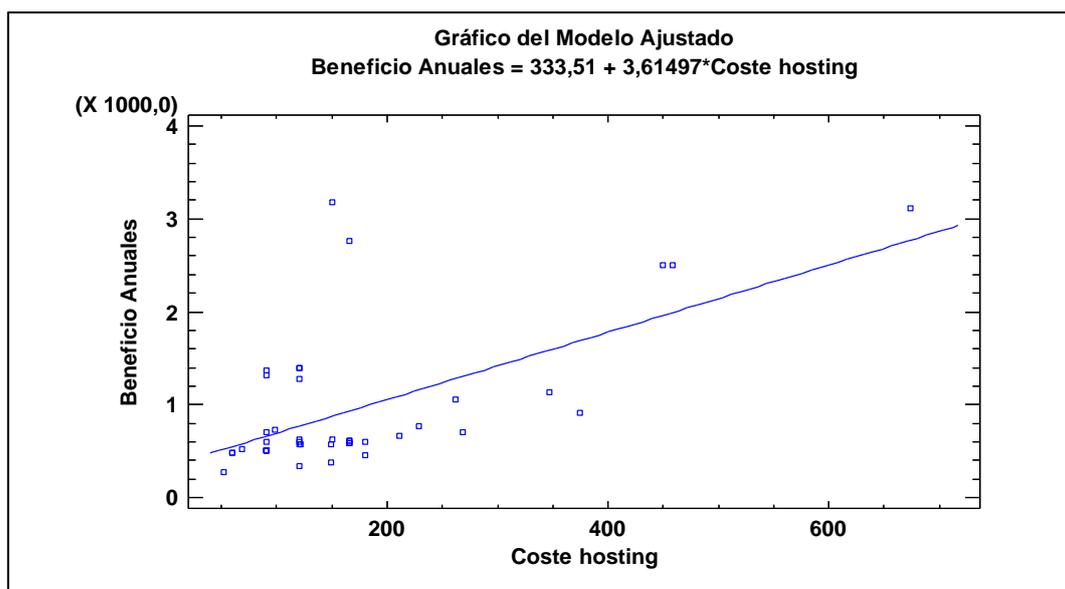


**Gráfico 3.18.** Regresión Simple - Beneficio Anuales vs. Gasto servidor anual. **Fuente:** Elaboración propia

### 3.3.2. Regresión Simple - Beneficio Anuales vs. Coste hosting

- **Variable dependiente:** Beneficio Anuales
- **Variable independiente:** Coste hosting

Si observamos este análisis bivalente entre la variable dependiente beneficios anuales y la variable independiente coste de hosting se haya que existe un coeficiente de correlación de 0,625641, esto quiere decir que entre estas variables existe algo de relación entre ellas, aunque mediante el gráfico 3.19 podemos ver que se forma una recta ascendente que nos indica que existe un relación directa con la variable Y.

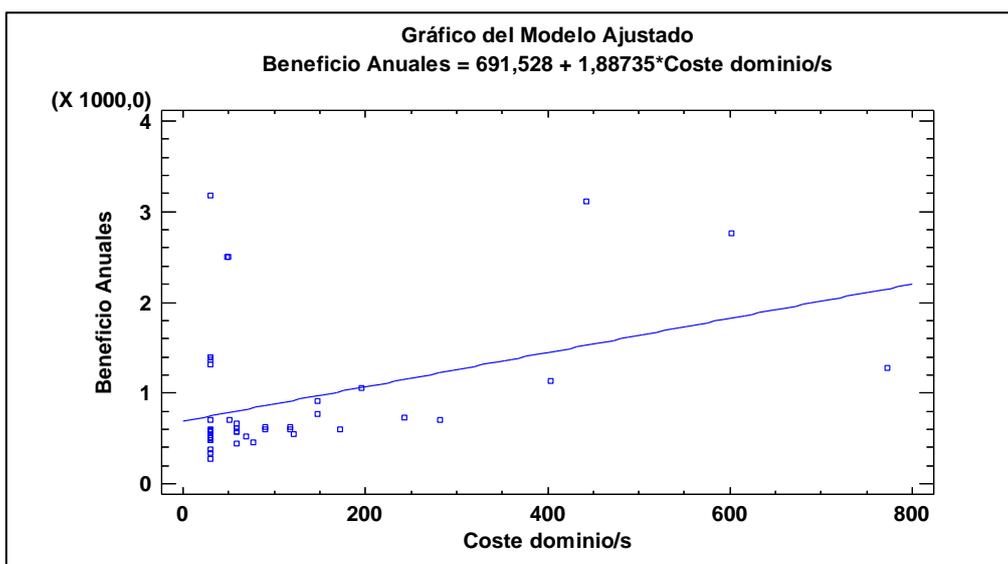


**Gráfico 11.** Regresión Simple - Beneficio Anuales vs. Coste hosting. **Fuente:** Elaboración propia

### 3.3.3. Regresión Simple - Beneficio Anuales vs. Coste dominio/s

- **Variable dependiente:** Beneficio Anuales
- **Variable independiente:** Coste dominio/s

Mediante un análisis bivariante entre la variable dependiente beneficios anuales y la variable independiente coste de dominio/s se haya que existe un coeficiente de correlación de 0,403726, esto nos indica que existe entre algo de relación, y observando el gráfico 3.20 podemos ver que existe recta ascendente lo que indica que existe una relación directa con la variable Y.

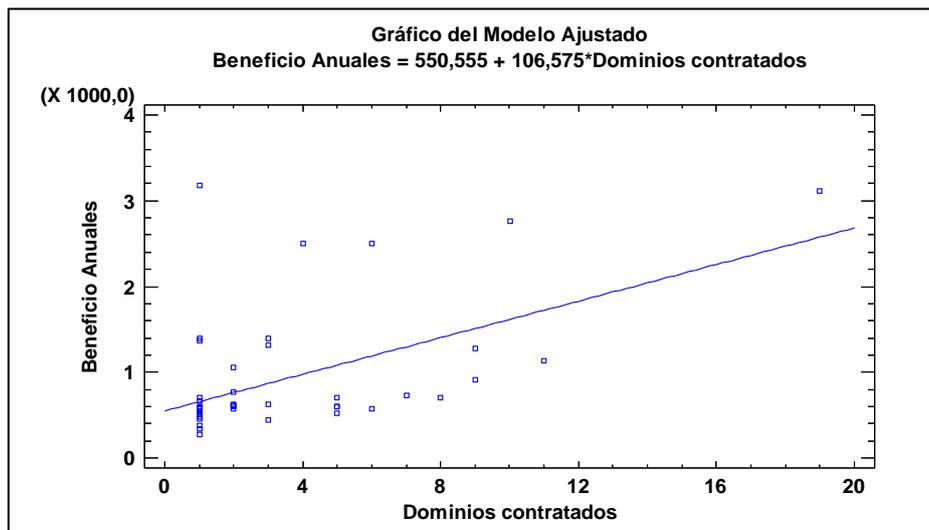


**Gráfico 3.20.** Regresión Simple - Beneficio Anuales vs. Coste dominio/s. **Fuente:** Elaboración propia

### 3.3.4. Regresión Simple - Beneficio Anuales vs. Dominios contratados

- **Variable dependiente:** Beneficio Anuales
- **Variable independiente:** Dominios contratados

Escogiendo la variable dependiente beneficios anuales y la variable independiente dominio/s contratados hacemos un análisis bivalente en que observamos que existe un coeficiente de correlación de 0,536294, ello nos indica que existe algo de relación entre las variables, y observando el gráfico 3.21 podemos ver que existe recta totalmente ascendente por lo que nos indica que existe una relación directa con la variable beneficios anuales.

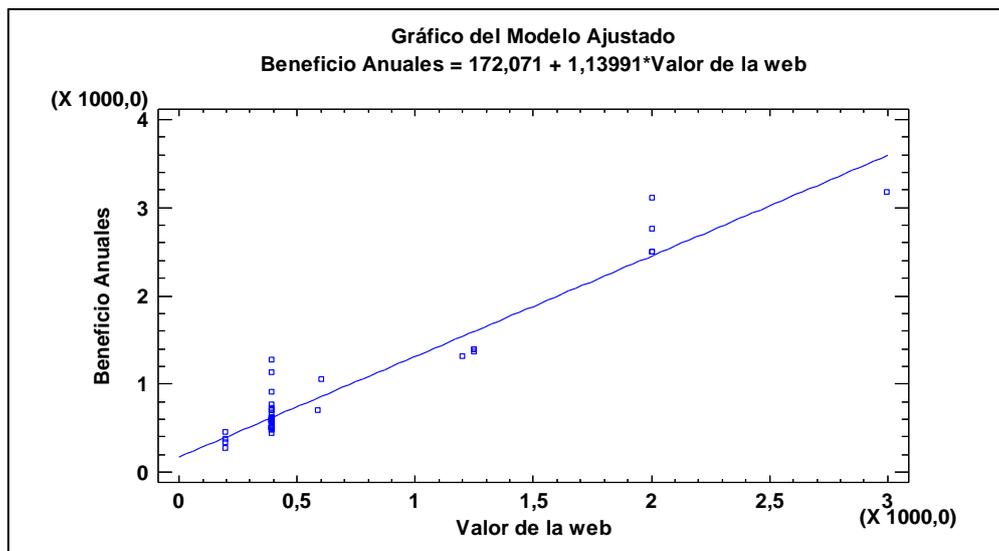


**Gráfico 3.21.** Regresión Simple - Beneficio Anuales vs. Dominios contratados. **Fuente:** Elaboración propia

### 3.3.5. Regresión Simple - Beneficio Anuales vs. Valor de la web

- **Variable dependiente:** Beneficio Anuales
- **Variable independiente:** Valor de la web

Como podemos ver mediante un análisis bivalente entre la variable dependiente beneficios anuales y la variable independiente valor de la web podemos ver que es la variable independiente con más relación mediante su coeficiente de correlación que es de 0,958144, por ello podemos observar que en el gráfico 3.22 existe una curva totalmente inclinada ascendentemente con una relación directa positiva.

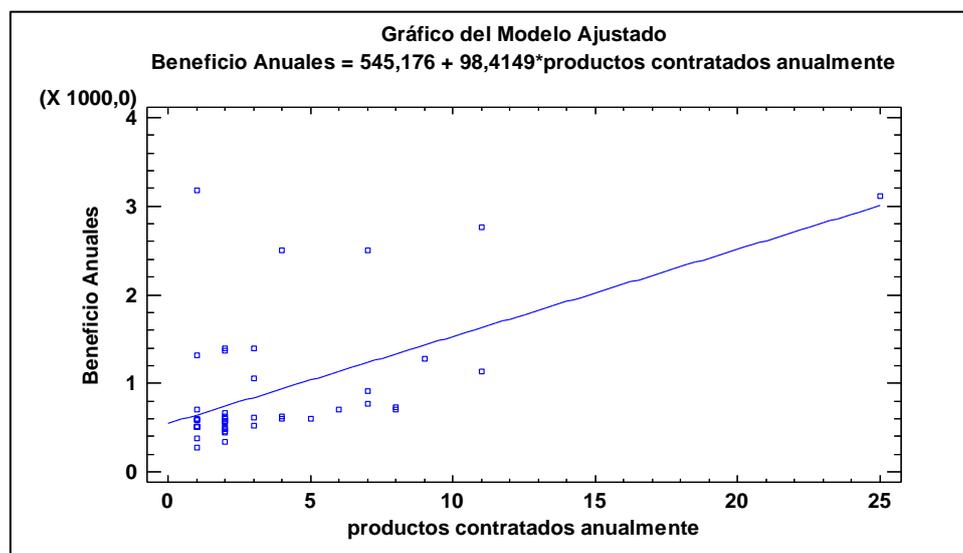


**Gráfico 3.22.** Regresión Simple - Beneficio Anuales vs. Valor de la web. **Fuente:** Elaboración propia

### 3.3.6. Regresión Simple - Beneficio Anuales vs. Productos contratados anualmente

- **Variable dependiente:** Beneficio Anuales
- **Variable independiente:** Productos contratados anualmente

Realizando un análisis bivalente entre la variable dependiente beneficios anuales y la variable independiente productos contratados anualmente podemos observar que existe una relación algo de relación gracias a su coeficiente de correlación que es de 0,571548. Gracias al gráfico 3.23 existe una curva totalmente inclinada ascendentemente con una relación directa positiva.



**Gráfico 3.23.** Regresión Simple - Beneficio Anuales vs. Productos contratados anualmente. **Fuente:** Elaboración propia

## 3.4. Análisis multicolinealidad

### 3.4.1. Análisis Multivariado

Por ello vamos a analizar nuestras variables cuantitativas independientes que son:

- Gasto servidor anual - GSA
- Coste hosting - CHosting
- Coste dominio/s - CDominio
- Dominios contratados - DContrat
- Valor de la web - VWeb
- Productos contratados anualmente - PCA

### Matriz a analizar

Como podemos ver posteriormente hemos hecho un análisis multivariado mediante la combinación de las variables cuantitativas de nuestro modelo y de ellas hemos conseguido sacar esta matriz que vemos a continuación:

	<i>GSA</i>	<i>CHosting</i>	<i>CDominio</i>	<i>DContrat</i>	<i>VWeb</i>	<i>PCA</i>
<i>GSA</i>	1	0,05	-0,35	-0,22	0,00	-0,23
<i>CHosting</i>	0,05	1	0	0,64	0,47	0,69
<i>CDominio</i>	-0,35	0,27	1	0,73	0,17	0,73
<i>DContrat</i>	-0,22	0,64	<b>0,73</b>	1	0,32	0,94
<i>VWeb</i>	0,00	0,47	0,17	0,32	1	0,35
<i>PCA</i>	-0,23	0,69	<b>0,73</b>	<b>0,94</b>	0,35	1

### FASE 1. Matriz de correlación

Como podemos ver en la **tabla 3.8** hacemos la primera fase para ello, se nos muestra las correlaciones comparando cada variable contra la otra excluyendo nuestra variable Y. En esta tabla observaremos los coeficientes marcados en rojo para analizarlos mediante su rango. Estos coeficientes de correlación van de -1 a +1, y miden la fuerza de la relación lineal entre las variables. Los coeficientes con valor por debajo de |0,5| indican correlaciones débiles y los coeficientes por encima de |0,7| nos indican correlaciones elevadas, y por tanto problema de multicolinealidad.

	Gastos servidor anual	Coste hosting	Coste dominio	Dominios contratados	Valor de la web	Productos contratados anualmente
	GSA	CHosting	CDominio	DContrat	VWeb	PCA
GSA	1	0,05	-0,35	-0,22	0,00	-0,23
CHosting	<b>0,05</b>	1	0,27	0,64	0,47	0,69
CDominio	<b>-0,35</b>	<b>0,27</b>	1	0,73	0,17	0,73
DContrat	<b>-0,22</b>	<b>0,64</b>	<b>0,73</b>	1	0,32	0,94
VWeb	<b>0,00</b>	<b>0,47</b>	<b>0,17</b>	<b>0,32</b>	1	0,35
PCA	<b>-0,23</b>	<b>0,69</b>	<b>0,73</b>	<b>0,94</b>	<b>0,35</b>	1

Tabla 3.8. Matriz de correlación. Fuente: Elaboración propia

### FASE 2. Matriz inversa de correlación

Como se puede ver en la **tabla 3.9** se hace una segunda fase en la que mediante la tabla 3.8 se calcula la matriz inversa de la matriz de correlación, en la que ahora no se analiza una variable contra otra sino ahora se está analizando contra el resto de variables a la vez. En esta tabla se tiene que observar los coeficientes marcados en rojo para analizarse mediante su rango. Los coeficientes con valor por debajo de 10 indican que no existe multicolinealidad, pero si el coeficiente da un rango mayor de 10 entonces indican problema de multicolinealidad.

<i>GSA</i>	<b>1,20</b>	-0,37	0,29	-0,12	0,02	0,44
<i>CHosting</i>	-0,37	<b>2,83</b>	1,16	-0,27	-0,58	-2,43
<i>CDominio</i>	0,29	1,16	<b>2,90</b>	-1,09	-0,05	-1,81
<i>DContrat</i>	-0,12	-0,27	-1,09	<b>8,48</b>	0,06	-7,01
<i>VWeb</i>	0,02	-0,58	-0,05	0,06	<b>1,29</b>	-0,07
<i>PCA</i>	0,44	-2,43	-1,81	-7,01	-0,07	<b>10,67</b>

<b>GSA</b>	<b>1,20</b>	-0,37	0,29	-0,12	0,02	0,44
<b>Chosting</b>	-0,37	<b>2,83</b>	1,16	-0,27	-0,58	-2,43
<b>Cdominio</b>	0,29	1,16	<b>2,90</b>	-1,09	-0,05	-1,81
<b>Dcontrat</b>	-0,12	-0,27	-1,09	<b>8,48</b>	0,06	-7,01
<b>Vweb</b>	0,02	-0,58	-0,05	0,06	<b>1,29</b>	-0,07
<b>PCA</b>	0,44	-2,43	-1,81	-7,01	-0,07	<b>10,67</b>

**Tabla 3.9.** Matriz de correlación inversa. **Fuente:** Elaboración propia

### FASE 3. Índice de acondicionamiento

Para saber cuál es su índice de acondicionamiento se ha cogido el autovalor máximo y el mínimo y mediante la fórmula matemática que se ve a continuación se otorga un índice de acondicionamiento del 7,46.

	<b>Autovalor</b>
<b>1</b>	<b>3,2869</b>
<b>2</b>	1,24719
<b>3</b>	0,758956
<b>4</b>	0,479072
<b>5</b>	0,168811
<b>6</b>	<b>0,0590685</b>

$$IC = \sqrt{\frac{\text{Autovalor Maximo}}{\text{Autovalor Minimo}}} = \sqrt{\frac{3,2869}{0,0590685}} = 7,46$$

<b>IC</b>	<b>7,46</b>
-----------	-------------

**Tabla 3.10.** Valores de representados de mayor a menor e índice de acondicionamiento. **Fuente:** Elaboración propia

### 3.4.2. Análisis Multivariado 2

En esta segunda parte se vuelves a dar los mismos pasos que anteriormente, pero esta vez se suprime la variable cuantitativa que ha dado multicolinealidad como ha sido “**Productos contratados anualmente**”.

#### Matriz a analizar

	<b>GSA</b>	<b>CHosting</b>	<b>CDominio</b>	<b>DContrat</b>	<b>VWeb</b>
<b>GSA</b>	1	0,05	-0,35	-0,22	0,00
<b>CHosting</b>	0,05	1	0	0,64	0,47
<b>CDominio</b>	-0,35	0,27	1	0,73	0,17
<b>DContrat</b>	-0,22	0,64	<b>0,73</b>	1	0,32
<b>VWeb</b>	0,00	0,47	0,17	0,32	1

### FASE 1. Matriz de correlación

Como se puede ver en la **tabla 3.10** se recrea la primera fase para analizar variables marcadas en rojo para analizarlos mediante su rango. Como se puede ver en la tabla, sigue existiendo multicolinealidad en la variable cuantitativa “dominio contratados”.

	Gastos servidor anual	Coste hosting	Coste dominio	Dominios contratados	Valor de la web
	GSA	Chosting	Cdominio	Dcontrat	Vweb
GSA	1	0,05	-0,35	-0,22	0,00
Chosting	0,05	1	0,27	0,64	0,47
Cdominio	-0,35	0,27	1	0,73	0,17
Dcontrat	-0,22	0,64	0,73	1	0,32
Vweb	0,00	0,47	0,17	0,32	1

**Tabla 3.10.** Matriz de correlación. Fuente: Elaboración propia

### FASE 2. Matriz inversa de correlación

Observando la **tabla 3.11** se haya que incluso calculando la matriz inversa de la matriz de correlación sigue existiendo multicolinealidad en la variable cuantitativa “dominios contratados” como anteriormente.

$$\begin{matrix}
 GSA \\
 CHosting \\
 CDominio \\
 DContrat \\
 VWeb
 \end{matrix}
 \begin{pmatrix}
 1,18 & -0,27 & 0,36 & 0,16 & 0,02 \\
 -0,27 & 2,28 & 0,75 & -1,87 & -0,60 \\
 0,36 & 0,75 & 2,59 & -2,27 & -0,06 \\
 0,16 & -1,87 & -2,27 & 3,88 & 0,02 \\
 0,02 & -0,60 & -0,06 & 0,02 & 1,29
 \end{pmatrix}$$

GSA	1,18	-0,27	0,36	0,16	0,02
Chosting	-0,27	2,28	0,75	-1,87	-0,60
Cdominio	0,36	0,75	2,59	-2,27	-0,06
Dcontrat	0,16	-1,87	-2,27	3,88	0,02
Vweb	0,02	-0,60	-0,06	0,02	1,29

**Tabla 3.11.** Matriz de correlación inversa. Fuente: Elaboración propia

### FASE 3. Índice de acondicionamiento

En la tabla 3.12 se puede ver que finalmente se calcula de nuevo el autovalor máximo y el autovalor mínimo, para obtener el índice de acondicionamiento mediante la fórmula matemática que vemos a continuación. Se otorga un índice de acondicionamiento del 3,95.

	Autovalor
1	2,39647
2	1,2462
3	0,728543
4	0,475466
5	0,15332

$$IC = \sqrt{\frac{\text{Autovalor Maximo}}{\text{Autovalor Minimo}}} = \sqrt{\frac{2,39647}{0,15332}} = 3,95$$

IC	3,95
----	------

**Tabla 3.12.** Valores de representados de mayor a menor e índice de acondicionamiento. **Fuente:** Elaboración propia

### 3.4.3. Análisis Multivariado 3

Finalmente se vuelve hacer este análisis de multicolinealidad hasta que finalmente no aparezca multicolinealidad, y por ello, se ha vuelto a quitar otra variable como es la variable cuantitativa “dominios contratados”.

#### Matriz a analizar

$$\begin{matrix} & & \mathbf{GSA} & \mathbf{CHosting} & \mathbf{CDominio} & \mathbf{VWeb} \\ \mathbf{GSA} & & 1 & 0,05 & -0,35 & 0 \\ \mathbf{CHosting} & & 0,05 & 1 & 0,27 & 0,47 \\ \mathbf{CDominio} & & -0,35 & 0,27 & 1 & 0,17 \\ \mathbf{VWeb} & & 0 & 0,47 & 0,17 & 1 \end{matrix}$$

### FASE 1. Matriz de correlación

Viendo la **tabla 3.13** se puede ver que no existe variable cuantitativa con problemas de multicolinealidad puesto que ningún coeficiente de correlación es mayor a |0,7|.

	Gasto servidor anual	Coste hosting	Coste dominio/s	Valor de la web
	GSA	Chosting	Cdominio	Vweb
GSA	1	0,05	-0,35	0,00
Chosting	0,05	1	0,27	0,47
Cdominio	-0,35	0,27	1	0,17
Vweb	0,00	0,47	0,17	1

**Tabla 3.13.** Matriz de correlación. **Fuente:** Elaboración propia

### FASE 2. Matriz inversa de correlación

Juzgando la **tabla 3.14** se puede hallar que mediante el cálculo de la matriz inversa no existe multicolinealidad.

$$\begin{matrix} \text{GSA} \\ \text{CHosting} \\ \text{CDominio} \\ \text{VWeb} \end{matrix} \begin{pmatrix} 1,17 & -0,19 & 0,46 & 0,02 \\ -0,19 & 1,38 & -0,34 & -0,59 \\ 0,46 & -0,34 & 1,26 & -0,05 \\ 0,02 & -0,59 & -0,05 & 1,29 \end{pmatrix}$$

<b>GSA</b>	<b>1,17</b>	<b>-0,19</b>	<b>0,46</b>	<b>0,02</b>
<b>Chosting</b>	-0,19	<b>1,38</b>	-0,34	-0,59
<b>Cdominio</b>	0,46	-0,34	<b>1,26</b>	-0,05
<b>Vweb</b>	0,02	-0,59	-0,05	<b>1,29</b>

**Tabla 3.14.** Matriz de correlación inversa. **Fuente:** Elaboración propia

### FASE 3. Índice de acondicionamiento

En la tabla 3.15 se halla que finalmente no existe en ningún caso multicolinealidad así que se vuelve hacer el cálculo para obtener realmente el índice de acondicionamiento del 1,86.

	<b>Autovalor</b>
<b>1</b>	<b>1,66</b>
<b>2</b>	1,24
<b>3</b>	0,64
<b>4</b>	<b>0,47</b>

$$IC = \sqrt{\frac{\text{Autovalor Maximo}}{\text{Autovalor Minimo}}} = \sqrt{\frac{1,66}{0,47}} = 1,86$$

<b>IC</b>	<b>1,86</b>
-----------	-------------

**Tabla 3.15.** Valores de representados de mayor a menor e índice de acondicionamiento. **Fuente:** Elaboración propia

#### 3.4.4. Conclusión multicolinealidad

Como se ha podido observar en este apartado del análisis de multicolinealidad se ha hallado que entre las variables del modelo existía multicolinealidad.

En la primera fase se ha hecho una comprobación de “variable vs variable” y gracias a ello se apreció que existían dos variables con problema de multicolinealidad, las cuales son “Coste de dominio” y “Dominios contratados” aunque no solo con la fase 1 se puede dar fe que existe multicolinealidad. Así que a continuación se ha hecho la fase 2 en la que se trata de poner una variable contra todas las demás (“Variable  $X_1$  vs Variable  $X_2$  + Variable  $X_3$  +...”) y gracias a ello se ha confirmado que realmente donde existía la multicolinealidad en la variable “Productos contratados anualmente” y por ello se elimina de nuestro modelo esta variable. Finalmente se

calcula el índice de acondicionamiento que nos da 7,46 en este modelo con todas las variables antes de eliminarla, ya ahora se volverá a repetir este análisis hasta que no de multicolinealidad.

Con el nuevo modelo se volverá hacer de nuevo el análisis de multicolinealidad repitiendo los mismos pasos y como se ha visto anteriormente en la primera fase nos ha dado problema en la variable “Coste de dominio”, a continuación, ha dado en la segunda fase otra vez problemas en la misma variable así que definitivamente se puede concluir que esta variable debe ser eliminada. Y finalmente se ha acabado el análisis calculando el índice de acondicionamiento que da un 3,95.

Y finalmente, por última vez se prueba si existen problemas de multicolinealidad y se ve que tanto en la fase 1 como en la fase 2 no existe en ninguna variable que genere problemas. Por ello, se da por libre el modelo de multicolinealidad y se quedaría sin estas las dos variables que se han eliminado. El índice de acondicionamiento con estas variables se quedaría en 1,86.

### 3.5. Formulación del modelo teórico propuesto

En este apartado se va a explicar el modelo teórico en el que se basa para resolver este estudio econométrico. Siguiendo con el estudio, si se considera que la variable endógena o dependiente es “**Beneficios anuales**” que a continuación se nombrará “**Y**” y a las variables exógenas o explicativas las nombraremos:

- **Gastos servidor anual:** “GSA”
- **Coste de hosting:** “CHosting”
- **Coste de dominio/s:** “CDomino”
- **Valor de Web:** “VWeb”

Los parámetros  $\beta$  del modelo de regresión recogen (cuantifican) la relación existente entre cada variable explicativa  $X_i$  y la variable explicada  $Y$ . A continuación, observaremos esta representación:

$$\text{Benéficos Anuales} = \beta_0 + \beta_1 \text{GSA} + \beta_2 \text{CHosting} + \beta_3 \text{CDomino} + \beta_4 \text{VWeb} + U$$

**Y (variable endógena):** este valor se determina por las variables exógenas que hay en el modelo. El objetivo es buscar y obtener los valores para explicar  $Y$ .

$\beta_0$ : Valor promedio del beneficio anual cuando el resto de variables explicativas son cero.

$\beta_1$ : Incremento de los beneficios anuales cuando aumenta en una unidad la variable GSA y el resto de variables explicativas permanecen constantes.

$\beta_2$ : Incremento de los beneficios anuales cuando aumenta en una unidad la CHosting y el resto de variables explicativas permanecen constantes.

$\beta_3$ : Incremento de los beneficios anuales cuando aumenta en una unidad el CDominio y el resto de variables explicativas permanecen constantes.

$\beta_4$ : Incremento de los beneficios anuales cuando aumenta en una unidad la Vweb y el resto de variables explicativas permanecen constantes.

**U (perturbación)**: Todas las variables explicativas que no se han tenido en cuenta para explicar los beneficios y sí que servirían para explicarlo.

### 3.6. Regresión múltiple

Como se ha mencionado en el apartado anterior, se utiliza el método de la regresión múltiple cuando se desea estudiar la posible relación entre varias variables independientes (predictores o explicativas) y otra variable dependiente (criterio, explicada o respuesta).

En primer lugar, se utilizarán todas las variables del modelo para hacer este cálculo utilizando la herramienta Statgraphics para saber si estas variables explican los beneficios.

Se debe tener en cuenta el **estadístico R-Cuadrado**, ya que aporta realmente en cuanto se adapta el modelo a los datos mediante las variables.

Al ver la regresión múltiple, se observa un  $R^2$  de 100% de explicación. Se va a separar las variables para poder saber cada una el tanto por ciento que nos puede indicar.

Variables % de explicación:

- El **estadístico R-Cuadrada** indica que la variable **GSA** explica un **0,469415%** de la variabilidad en Beneficio Anuales. El error estándar del estimado muestra que la desviación estándar de los residuos es 713,685.
- El **estadístico R-Cuadrada** indica que la variable **CHosting** explica **39,1426%** de la variabilidad en Beneficio Anuales.
- El **estadístico R-Cuadrada** indica que la variable **CDomino** explica **16,2995%** de la variabilidad en Beneficio Anuales. El error estándar del estimado muestra que la desviación estándar de los residuos es 654,474.
- El **estadístico R-Cuadrada** indica que la variable **VWeb** explica **91,8041%** de la variabilidad en Beneficio Anuales. El error estándar del estimado muestra que la desviación estándar de los residuos es 204,799.

Como se puede ver en los porcentajes la variable “GSA” no explica para nada el modelo, ni un 1% de él y que la variable “VWeb” lo explica casi en su totalidad. Viendo estos dos porcentajes mediante el estadístico R-Cuadrada se piensa que se debería descartar estas dos variables ya que con una no se podría explicar el modelo y con la otra sería fácilmente explicable, por ello se va a utilizar las variables “CHosting” y la “CDominio” para poder explicar este modelo econométrico.

En segundo lugar, como bien se ha comentado antes, se va a calcular la regresión lineal múltiple para describir la relación entre Beneficio Anuales y las 2 variables independientes “CHosting” y la “CDominio”. A continuación, se observará esta representación:

$$\text{Beneficio Anuales} = 269,426 + 3,21911 * \text{Coste hosting} + 1,18107 * \text{Coste dominio/s}$$

### 3.6.1. Validación del modelo

#### 3.6.1.1. Regresión múltiple

Teniendo en cuenta la R-Cuadrada de este modelo, nos indica que el **modelo es explicado en un 45,0561%** de la variabilidad de los Beneficio Anuales, y por ello no es difícil explicarlo totalmente.

#### 3.6.1.2. Significadidad

**Variable dependiente:**

Beneficio Anuales

**Variables independientes:**

Coste hosting

Coste dominio/s

Parámetro	Estimación	Error	Estadístico	Valor-P
		Estándar	T	
CONSTANTE	269,426	125,924	2,13958	0,0376
Coste hosting	3,21911	0,649048	4,95974	0,0000
Coste dominio/s	1,18107	0,525125	2,24912	0,0292

**Tabla 3.16.** Tabla de significadidad. **Fuente:** Elaboración propia

### Análisis de Varianza

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	1,10676E7	2	5,53378E6	19,27	0,0000
Residuo	1,34964E7	47	287157,		
Total (Corr.)	2,45639E7	49			

**Tabla 3.17.** Tabla de significadidad de análisis de varianza. **Fuente:** Elaboración propia

- **Estimación puntual de los parámetros del modelo**

$$\text{Beneficio Anuales} = 269,426 + 3,21911 * \text{Coste hosting} + 1,18107 * \text{Coste dominio/s}$$

- **Estimación por intervalos de confianza al 95% de los parámetros del modelo.**
  - $\beta_0$ : [16,098 ; 522,753]<sub>95%</sub>
  - $\beta_1$ : [1,91339 ; 4,52483]<sub>95%</sub>
  - $\beta_2$ : [0,124652 ; 2,23749]<sub>95%</sub>
- **Determinar si el modelo en sí resulta significativo. Se realiza la prueba utilizando el P-Value.**

### CONTRASTE DEL MODELO

$$H_0: B_1=B_2=0$$

$$H_1: \text{Algún } B_i \neq 0$$

Como P-valor es igual a 0.00, se rechaza H0. Por tanto, se concluye que el modelo resulta significativo.

- **Significación de los parámetros del modelo. Se realiza la prueba utilizando el P-Valor.**

**CONTRASTE DE  $\beta_0$**

$$H_0: B_0=0$$

$$H_1: B_0 \neq 0$$

**CONTRASTE DE  $\beta_1$**

$$H_0: B_1=0$$

$$H_1: B_1 \neq 0$$

**CONTRASTE DE  $\beta_2$**

$$H_0: B_2=0$$

$$H_1: B_2 \neq 0$$

**CONTRASTE DE  $\beta_0$ :** Como P-valor es igual a 0.0376, se rechaza H0. El  $\beta_0$  es significativo para el modelo.

**CONTRASTE DE  $\beta_1$ :** Como P-valor es igual a 0.000, se rechaza H0. El  $\beta_1$  es significativo para el modelo. La variable CHosting es significativa para el modelo.

**CONTRASTE DE  $\beta_2$ :** Como P-valor es igual a 0.0292, se rechaza  $H_0$ . El  $\beta_2$  es significativo para el modelo.

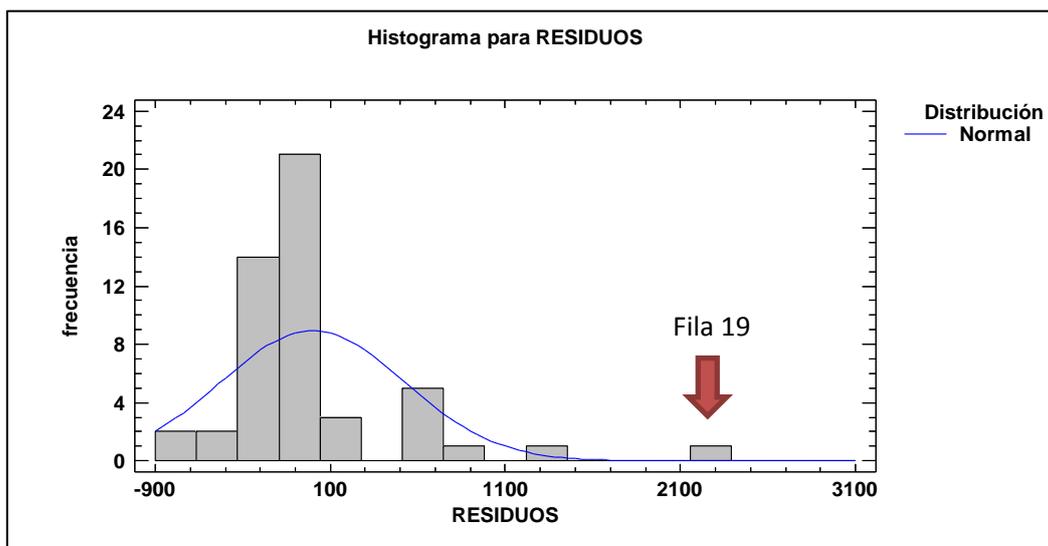
- **Determinar los valores del coeficiente de determinación y del coeficiente de determinación corregido, y explicar su significado.**

El *R-cuadrado* que nos aporta el modelo mediante estas dos variables es bastante bajo es de un 45,0561%. Se puede decir que, aunque las variables sean significativas, estas variables solo pueden llegar a explicar una pequeña parte del modelo. Por ello, más adelante se propondrá una mejora del modelo para que se ajuste mejor.

### 3.6.1.3. Normalidad de residuos

En este apartado se analizará la normalidad de residuos del modelo, aquí se nos muestra los resultados de una distribución normal a los datos de **residuos**. Observando el gráfico 3.24 se puede evaluar visualmente que también la distribución normal se ajusta al modelo.

Aun así, para asegurarse adecuadamente se va a eliminar la fila 19, ya que es la más alejada de los residuos. (Gráfico 3.25)



**Gráfico 3.24.** Gráfico de Probabilidad Normal - RESIDUOS. **Fuente:** Elaboración propia

#### Pruebas de Normalidad para RESIDUOS

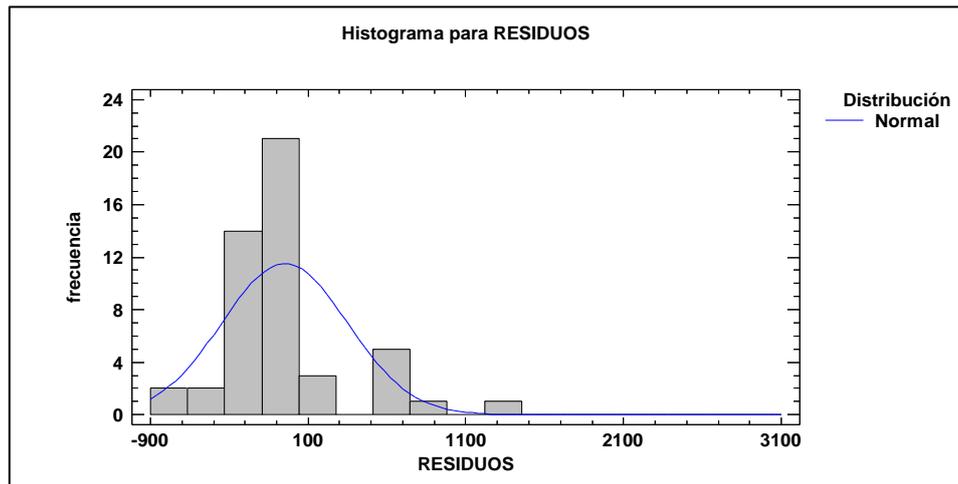
Observando la tabla 3.18 de las pruebas de normalidad se puede evaluar si los residuos son normales o no mediante el valor-P. Si es mayor a  $\alpha$  aceptamos  $H_0$  y si es menor lo rechazamos.

Prueba	Estadístico	Valor-P
Valor-Z para curtosis	2,1802	<b>0,0292428</b>

**Tabla 3.18.** Pruebas de Normalidad para residuos. **Fuente:** Elaboración propia

La prueba en la que más nos se va a enfocar se trata de la **Valor-Z para curtosis** estandarizada que busca si la forma de la distribución es más plana o picuda que la distribución normal. Viendo su valor-P que es de 0,0292 es menor que  $\alpha$  (0,05), y se puede decir que los residuos no son normales.

Si se vuelve a hacer de nuevo el análisis de normalidad mediante Statgraphics y eliminando la fila 19(Empresa 19) se puede observar el **gráfico 3.25** como visualmente los residuos se ajustan mucho más a la distribución normal.



**Gráfico 3.25.** Gráfico de Probabilidad Normal – RESIDUOS (Eliminando fila 19). **Fuente:** Elaboración propia

### Pruebas de Normalidad para RESIDUOS

Observando la tabla 3.17 de las pruebas de normalidad se puede evaluar si los residuos son normales o no mediante el valor-P. Si es mayor a  $\alpha$  se acepta  $H_0$ , es decir, los residuos son normales.

Prueba	Estadístico	Valor-P
Valor-Z para curtosis	1,91526	0,0554592

**Tabla 3.19.** Pruebas de Normalidad para residuos. **Fuente:** Elaboración propia

### 3.6.1.4. Heterocedasticidad

Posteriormente para comprobar la heterocedasticidad se realiza un modelo que relaciona los residuos con las variables explicativas.

$$\text{RESIDUOS}^2 = 21326,0 + 529,695 \cdot \text{Coste hosting} + 492,123 \cdot \text{Coste dominio/s}$$

**Variable dependiente:**

Residuos<sup>2</sup>

**Variables independientes:**

Coste hosting

Coste dominio/s

$H_0 \sigma^2 = \text{cte}$	<b>no existe heterocedasticidad</b>
$H_1 \sigma^2 \neq \text{cte}$	<b>existe heterocedasticidad</b>

Parámetro	Estimación	Error	Estadístico	Valor-P
		Estándar	T	
CONSTANTE	21326,0	63766,7	0,334438	0,7396
Coste hosting	529,695	326,842	1,62064	<b>0,1119</b>
Coste dominio/s	492,123	265,14	1,85609	0,0699

**Tabla 3.20.** Pruebas de heterocedasticidad con residuos al cuadrado. **Fuente:** Elaboración propia

No se ve que las variables presenten problemas de heterocedasticidad.

P-valor >  $\alpha$   $\longrightarrow$  0,1119 > 0,05 por lo tanto se acepta  $H_0$

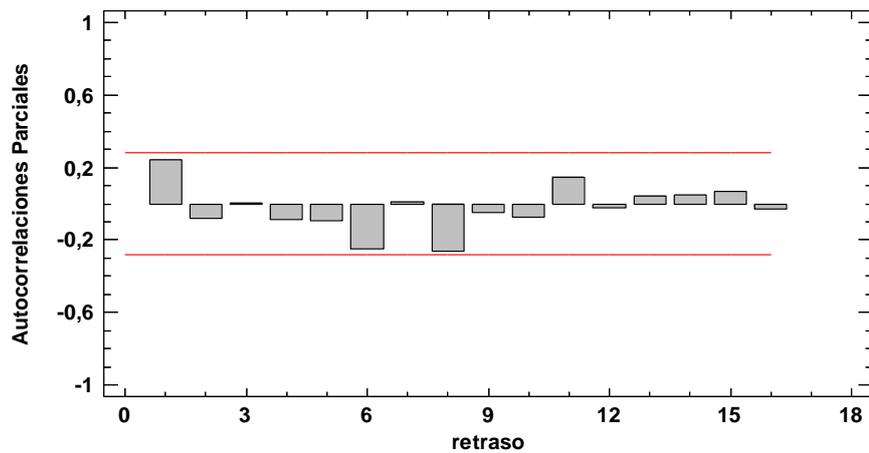
P-valor >  $\alpha$   $\longrightarrow$  0,0699 > 0,05 por lo tanto se acepta  $H_0$

Se concluye mediante un contraste de hipótesis que las variables explicativas no producen heterocedasticidad en el modelo.

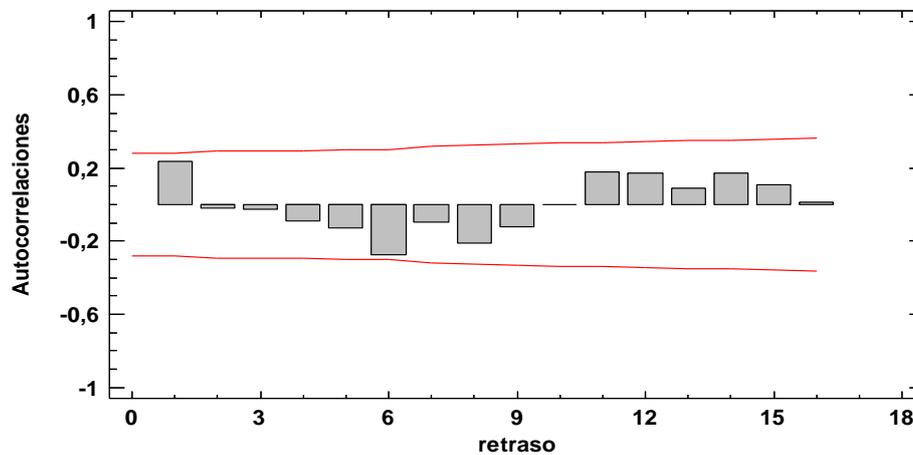
### 3.6.1.5. Autocorrelación

A continuación, se va a realizar mediante los gráficos de residuos **FAS** (Función de autocorrelación simple), que consiste en un conjunto de coeficientes de autocorrelación  $\rho_s$  y mide la relación lineal entre el residuo a  $e_t$  y  $e_{t-s}$  y **FAP** (Función de autocorrelación parcial), que mide lo mismo que FAS pero hace una pequeña corrección eliminando relaciones entre residuos intermedios.

Para su comprobación y detección se debe observar si existe presencia un sobrepaso de los límites, ya que si es así existirán problemas de autocorrelación en las variables.



**Gráfico 3.26.** Autocorrelaciones parciales estimadas FAP de residuos. **Fuente:** Elaboración propia



**Gráfico 3.27.** Autocorrelaciones parciales FAS de residuos. **Fuente:** Elaboración propia

Se concluye que mediante la observación del gráfico 3.26 “FAP” y 3.27 “FAS” se puede decir que no existe autocorrelación de ningún orden, ya que ningún factor de Autocorrelación supera los límites de autocorrelación.

### 3.6.1.6. Puntos anómalos y puntos influyentes

Son los valores atípicos que afectan sobre la estimación de parámetros del modelo.

#### Puntos Influyentes

Fila	Empresa	Influencia	Distancia de Mahalanobis	DFITS
1	1	0,412304	31,9941	0,445386
6	6	0,171639	8,7594	0,989789
17	17	0,253124	14,9496	2,3086
34	33	0,463778	39,6711	-1,11323
36	35	0,122515	5,58301	-0,72249
40	39	0,087617	3,53429	-0,559722
42	41	0,163063	8,178	0,980527

**Tabla 3.21.** Puntos influyentes a priori y puntos influyentes a posteriori. **Fuente:** Elaboración propia

Influencia media de un solo punto = **0,0612245**

Influencia media por 2 = **0,122449**

Se puede hacer una comprobación de puntos de influencia observando si los puntos a analizar doblan la media de influencia de un solo punto, en este caso podemos observar que las empresas con influencia a priori del modelo serían las empresas número 1, 6, 17, 33, 35 y 41.

#### Puntos a posteriori

Influencia media de un solo punto = **0,0612245**

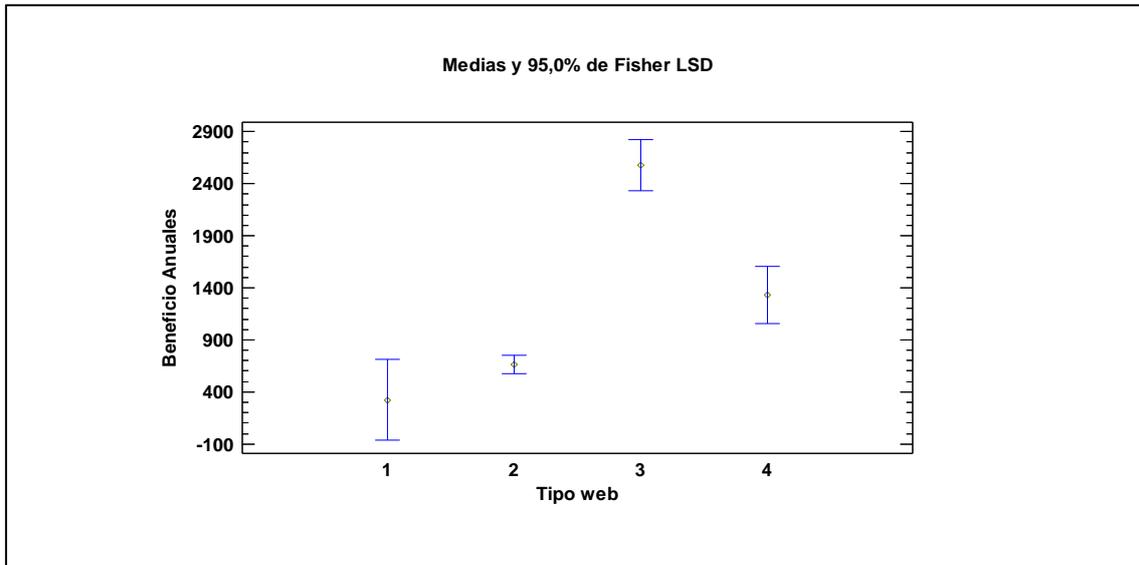
Influencia media  $|DFITS| \geq 2 \sqrt{n} = 2\sqrt{0,0612245} = \mathbf{0,4948717005}$

Se puede hacer una comprobación de puntos a posteriori observando si los puntos a analizar en valor absoluto son mayores a dos veces la raíz de la influencia media, en este caso podemos observar que las empresas con influencia a posteriori del modelo serían la empresa número 6, 17, 33, 35, 39 y 41.

### 3.6.1.7. Conclusión regresión múltiple

Teniendo en cuenta la R-Cuadrada de este modelo nos indica que el **modelo es explicado en un 60,26%** de la variabilidad en Beneficio Anuales y por ello no resulta muy eficiente para explicar los datos.

Por tanto, al ver que existe un R-cuadrado bajo se van a añadir variables cualitativas en el modelo para ver si al añadirlas el R-cuadrado aumenta. Por ellos, utilizaremos el programa Statgraphics para proceder a una regresión múltiple y así observando el **grafico 3.28** podremos valorar si se podría utilizar esta variable cualitativa en el modelo como es la variable “tipo de web”.



**Gráfico 3.28.** Gráfico de medias de tipos de web. **Fuente:** Elaboración propia

En **gráfico 3.28** se puede catalogar que existen 4 tipos de webs, pero se van englobaremos en 3 tipos ya que el tipo 1 y 2 es una web profesional, el tipo 3 se trata de una web a medida y el tipo 4 se trata de una tienda virtual.

Para añadir estas variables cualitativas al modelo vamos a darle un valor a cada una para diferenciarlas y clasificarlas. Por ello a continuación cuando hagamos una regresión múltiple para saber si mejora el modelo o no se tendrá que pensar que cuando se esté hablando de páginas web hechas a medida, es decir de tipo 3 se le denominara como “d1” y obtendrá valor 1 si es de tipo 3 y en caso contrario tomara valor 0 que consistirá que la página web del cliente es de tipo 1 o 2 e igual con la otra variable cuando se hable de tiendas virtuales es decir de tipo 4 se le denominara “d2” y obtendrá valor 1 si es de tipo 4 y en caso contrario tomara valor 0 que consistirá que la página web del cliente es de tipo 1 o 2.

- **Codificación de las variables cualitativas.**

**$d_1$**   
**1 = tipo web 3**  
**0 = Otro tipo**

**$d_2$**   
**1 = tipo web 4**  
**0 = Otro tipo**

**$d_3$**   
**1 = tipo web 1 o 2**  
**0 = Otro tipo**

- **Cuadro codificación.**

Tipo Web	$d_1$	$d_2$	$d_3$
1	0	0	1
2	0	0	1
3	1	0	0
4	0	1	0

**Tabla 3.22.** Cuadro codificación. **Fuente:** Elaboración propia



Finalmente, se va a recalcular la regresión lineal múltiple para describir la relación entre Beneficio Anuales y las 2 variables independientes “CHosting” y la “CDominio”, más la variable cualitativa “Tipo de web”.

### 3.6.1.8. Regresión Múltiple - Beneficio Anuales

**Variable dependiente:**

Beneficio Anuales

**Variabes independientes:**

Coste hosting

Coste dominio/s

d1

d2

Parámetro	Estimación	Error		Estadístico T	Valor-P
		Estándar			
CONSTANTE	292,624	66,3129		4,41278	0,0001
Coste hosting	2,05799	0,360109		5,7149	0,0000
Coste dominio/s	0,837926	0,273963		3,05853	0,0038
d1	1185,45	165,872		7,14679	0,0000
d2	582,839	141,904		4,10727	0,0002

**Tabla 3.23.** Cuadro de regresión múltiple de significatividad de las variables. **Fuente:** Elaboración propia

Como se puede ver en la tabla 3.23 se puede expresar que todas las variables actuales tanto cuantitativas, como cualitativas son todas significativas para el modelo, es decir, se sabe que simplemente que pueden explicarlo.

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	1,60519E7	4	4,01296E6	55,28	0,0000
Residuo	3,19408E6	44	72592,7		
Total (Corr.)	1,92459E7	48			

R-cuadrada = **83,4039** por ciento

**Tabla 3.24.** Cuadro de análisis de la varianza y R-cuadrada. **Fuente:** Elaboración propia

El estadístico R-Cuadrada nos indica que el modelo está ajustado y explica **83,40%** de la variabilidad en Beneficio Anuales. Esto quiere decir, que las variables explican considerablemente el modelo econométrico. Por lo tanto, hemos pasado que antes sin las variables cualitativas explicábamos muy poco el modelo ya que era un **60,26%** y ahora mediante esta variable hemos conseguido aumentar el porcentaje un **23,14%** podemos concluir que al mejorar hemos conseguido definir para finalmente poder hacer una predicción.

Representación de la fórmula de nuestro modelo finalmente:

$$\text{Beneficio Anuales} = 292,624 + 2,05799 \cdot \text{Coste hosting} + 0,837926 \cdot \text{Coste dominio/s} + 1185,45 \cdot d1 + 582,839 \cdot d2$$

### 3.6.1.9. Predicción

Para poder obtener valores fuera de la muestra, se ha realizado una predicción de la variable endógena. Por ello, se han estado haciendo las comprobaciones pertinentes anteriormente para conseguir obtener un buen comportamiento del modelo durante el periodo de la muestra para poder realizar una predicción fiable.

A continuación, se va a emplear el modelo para predecir, por ello vamos a crear el “Cliente 51” con nuevo proyecto web que consistirá en un hosting profesional, un nuevo dominio y una web profesional.

#### Resultados de la Regresión para Beneficio Anuales

	Ajustado	Inferior 95,0%	Superior 95,0%
<i>Fila</i>		<i>LC para Pronóstico</i>	<i>LC para Pronóstico</i>
50	1749,75	1110,17	2389,33

**Tabla 3.25.** Resultados de la predicción **Fuente:** Elaboración propia

Observando la tabla 3.25 podemos concluir que el pronóstico de nuestra predicción es que obtendríamos del cliente 51 unos beneficios de 1.749,75 euros con un intervalo de confianza entre 1.110,17€ y 2.389,33€.

Si se presta atención a la formulación del modelo se puede ver que la variable “d1” se basa en páginas webs a medida o de tipo 3, ya que es el producto que más beneficios nos puede aportar a la empresa porque es un tipo de web totalmente personalizado. Por ello, la empresa se debería replantear como diseñar la venta de este producto mediante un plan de marketing que consistiría en un análisis y diagnóstico de la situación de la empresa mediante un análisis externo observando el Macroentorno y el Microentorno, y un análisis interno para así diagnosticar la situación actual.

Gracias a este análisis la empresa podrá definir unos objetivos anuales para aumentar la venta del producto de tipo 3 y así se podrá obtener más beneficios. Para cumplir los objetivos la empresa deberá plantear una estrategia de segmentación como por ejemplo nos dirigiremos a un segmento de mercado de empresas medianas y grandes con un poder adquisitivo medio y posicionamiento como podría ser una estrategia de precios más por más, ya que al añadir una característica extra, le proporciona una serie de ventajas al producto que aumentan la calidad y precio de éste.

A continuación, utilizaremos la herramientas del Marketing Mix, compuesto inicialmente por las “4P”: Product (Producto), Price (Precio), Place (Emplazamiento) y Promotion (Promoción). Que la utilizaremos para hacer un análisis de estrategia internamente.



Finalmente acabaríamos haciendo un plan de acción definiendo en cada momento que se hacer, el instrumento que vamos a usar para comunicarlo a nuestros clientes potenciales, el responsable que deberá encargarse que se cumpla, el plazo que deseamos que se lleve a cabo y el coste aproximado de la acción. Al final se deberá hacer una reunión para ver si se han cumplido los objetivos.

## 4. Conclusiones y futuras líneas de investigación

A lo largo del presente trabajo y mediante los modelos econométricos planteados, se ha podido comprobar las diferentes relaciones que existe entre algunas de las variables del sector de servicio web.

Para ello se procedió a formular un modelo donde la variable dependiente o endógena era los beneficios anuales de la empresa, que posteriormente fui relacionando con las variables exógenas que supuse en un primer momento iban a estar más correlacionadas, como el gastos servidor anual, coste hosting/s, coste dominio/s, números de dominios contratados, tipo de dominios contratados, tipo de web contratada, valor de la web y productos contratados.

Pero al final se ha comprobado que mediante la significatividad de una variable como es VWeb (Valor de la web) se podía explicar el modelo en un 91%, es decir, que mediante esta variable hubiéramos acabado fácilmente este estudio, pero para poder estudiar más este modelo hemos escogido otras variables menos significativas pero que al juntar variables 2 variables cuantitativas y 1 cualitativa se ha hecho la comprobación mediante una regresión múltiple y se ha conseguido que explique un 83%. Esto ha permitido poder hacer una predicción sobre el comportamiento de los beneficios bajo la demanda de un cliente potencial.

### 4.1. Futuras líneas de investigación

Hoy en día, con la evolución tecnológica a la que estamos expuestos somos más propensos a tener una gran cantidad de “herramientas” que nos dan la posibilidad de orientar nuestras actividades diarias buscando la máxima eficiencia. Como por ejemplo aplicaciones para hacer deporte.

Pues en el paso de los años y más después de estar durante un año y medio dentro de este sector como es el de internet pienso que una de las futuras líneas de investigación que podría ofrecer este proyecto consistiría en la evolución en la demanda de aplicaciones para empresas, ya que es un mercado que está en auge y que dentro de un tiempo toda empresa tendrá en aplicación que podrían utilizarse para trabajar para sus propios trabajadores por medio de tabletas e incluso con el propio móvil personal o para darse a conocer a los clientes potenciales.

#### **¿En qué sentido podríamos desarrollar una aplicación para una empresa?**

Pongamos el ejemplo de una empresa que tiene diferente fases de producción y que incluso exista varios productos con diferentes composiciones como puede ser una empresa química para la creación de cremas para el cuerpo e incluso ponemos que tiene un contrato con un gran supermercado y simplemente la composición está mal hecha y no se ha detectado en ninguna fase en la línea de producción, existirá una perdida millonaria ya que si esta muestra se vendiera seguramente habrían quejas de los consumidores y seguro que el contrato con este gran supermercado se rompería.



Pues gracias a una aplicación que la empresa proporcionaría a sus trabajadores, se podría llegar mediante un código como puede ser el “QR o bidi”, que en la siguiente fase simplemente escaneando este código se podría validar como que la mezcla está bien hecha o simplemente el contenedor que lleva esta mezcla está equivocado con otra, y así evitar errores de producción. Es más, gracias a esto se podría llegar a personalizar un usuario de cada trabajador y saber la producción que se debe hacer durante toda semana, las horas que ha hecho de más durante esa semana e incluso se podría poner un aparatado en el que el trabajador mediante GPS puede validar la hora de entrada en la empresa y la hora de salida.



## 5. Bibliografía

- Gujarati, D. (1997). *Econometría*. ed. *Me Graw Hill*.
- Johnston, J. D., Johnston, J. J., & Dinardo, J. (2001). *Métodos de econometría* (No. 330.18/J72eE/2001).
- Díaz, M., & Llorente, M. M. (1998). *Econometría*. *Madrid, Pirámide*.
- Rodríguez, G., & Álvarez, F. J. M. (1996). *El análisis de los datos. Metodología de la Investigación Cualitativa*.
- Espina, P. Z. *MODELOS ECONOMÉTRICOS*.
- Peña, D. (1993). *Estadística, modelos y métodos: Fundamentos 2. Modelos lineales y series temporales*. Alianza Editorial.
- Peña, D. (2002). *Análisis de datos multivariantes* (Vol. 24). Madrid: McGraw-Hill.
- Sollutia S.L. URL: <http://www.sollutia.com/>

## 6. Anexo

### Servicios que Sollutia Ofrece

#### Diseño web

- **Tipo 1 (Sollutia Webpack Básico):** Se trata de un tipo de web cooperativa con el deseo de estar en internet pero más como una carta de presentación de sus servicios y su localización.
- **Tipo 2 (Sollutia Webpack Avanzado):** Se trata de un tipo de web cooperativa en el que la empresa desea no solo una carta de presentación si no enseñar su experiencia en proyectos realizados anteriormente para que su cliente potencial sepa lo que la empresa le puede ofrecer.
- **Tipo 3 (Diseño web a medida):** Se trata de un tipo de web específica en la que la empresa desea una web específica para clientes específicos como puede ser de una empresa de mecanizados que tiene un apartado en su web para que el cliente pueda subir los planos de una máquina que desea que le presupuesten para su futura construcción.
- **Tipo 4 (Tiendas virtuales Sollutia Shop):** Este tipo de web consiste en una tienda virtual en la que se puede adaptar a cualquier sector de venta online, como por ejemplo el caso de una farmacia de barrio que desea vender online.

#### Hosting y dominios

- **Hosting: o Alojamiento web:** Se refiere al lugar que guarda una página web y donde se gestionan los dominios y el correo electrónico de los clientes ya que el espacio de los buzones está hospedado en el espacio contratado para la página web.
- **Dominios:** Un dominio de Internet es una red de identificación que es asociada a un grupo de dispositivos o equipos conectados a la red Internet. El propósito principal de los nombres de dominio en Internet y del sistema de nombres de dominio (DNS), es traducir las direcciones IP de cada nodo activo en la red, a términos memorizables y fáciles de encontrar.  
Por ejemplo:
  - **.com:** son los dominios más extendidos en el mundo. Sirven para cualquier tipo de página web, temática.
  - **.es:** para servicios de España
  - **.eu:** para países de Unión Europea
  - **.video:** Sirve para una temática web específica
  - ...



## Marketing online

- **Posicionamiento en buscadores:** Análisis de criterios de búsqueda que identifican al negocio o al sector.
- **Creación de contenidos:** Redacción de contenidos optimizados para la web de los clientes con todas las combinaciones de los criterios de búsqueda a posicionar.
- **Campañas publicitarias:** Inversiones en publicidad en internet. Google, Facebook y Microsoft con el objetivo de obtener el máximo alcance a bajo coste con resultados inmediatos.
- **Redes sociales:** Contenidos que la empresa desea ofrecer y mediante imágenes o videos, ya que las redes sociales son tu campo de batalla.
- **Servicios de Google:** Como puede ser desde publicidad, hasta herramientas para tu empresa.
- **Servicio community manager personal:** Que consiste en estrategias de comunicación, marketing on-line y redes sociales personalizadas con gestor de cuenta, análisis, social media manager, gestión de perfiles en redes sociales.

[1] Tabla 6.1 Datos de los clientes de Sollutia S.L. Fuente: La empresa Sollutia

Clientes	Beneficio Anuales	Gasto servidor anual	Coste hosting	Tipo de hosting	Coste dominio/s	Dominios contratados	Cuantos Tipos de dominio	Tipo web	Valor de la web	productos contratados anualmente
1	3116,50	96,44	674,00	5	442,50	19	6	3	2000,00	25
2	449,00	96,44	0,00	5	59,00	3	2	2	390,00	2
3	510,00	96,44	90,50	3	29,50	1	1	2	390,00	2
4	344,50	96,44	120,00	4	29,50	1	1	2	195,00	2
5	1400,00	96,44	120,50	4	29,50	3	1	4	1250,00	2
6	2507,00	96,44	459,00	2	48,00	6	3	3	2000,00	7
7	596,15	98,33	34,45	2	171,70	5	4	2	390,00	5
8	510,00	96,44	90,50	3	29,50	1	1	2	390,00	1
9	510,00	96,44	90,50	3	29,50	1	1	2	390,00	1
10	600,00	96,44	120,50	4	89,50	2	2	2	390,00	2
11	1058,39	96,44	262,31	6	196,08	2	2	2	600,00	3
12	569,00	96,44	149,50	3	29,50	2	2	2	390,00	2
13	526,50	96,44	68,00	5	68,50	5	4	2	390,00	3
14	510,00	96,44	90,50	3	29,50	1	1	2	390,00	1
15	628,50	96,44	120,50	3	118,00	3	2	2	390,00	4
16	510,00	96,44	90,50	5	29,50	1	1	2	390,00	1
17	2767,50	96,44	165,50	4	602,00	10	7	3	2000,00	11
18	705,50	140,00	90,50	3	29,50	1	1	2	585,50	1
19	3174,50	140,00	150,00	5	29,50	1	1	3	2995,00	1
20	569,50	140,00	120,50	5	59,00	6	2	2	390,00	2
21	659,50	140,00	210,50	6	59,00	1	1	2	390,00	2
22	585,00	140,00	165,50	5	29,50	1	1	2	390,00	1
23	276,00	140,00	51,50	1	29,50	1	1	1	195,00	1
24	585,00	140,00	165,50	5	29,50	1	1	2	390,00	1



25	373,50	140,00	149,00	3	29,50	1	1	1	195,00	1
26	585,00	140,00	165,50	5	29,50	1	1	2	390,00	1
27	614,50	86,29	165,50	5	59,00	2	2	2	390,00	3
28	600,00	86,29	180,50	5	29,50	1	1	2	390,00	1
29	545,30	86,29	34,45	2	120,85	1	1	2	390,00	2
30	453,00	86,29	180,50	5	77,50	1	1	2	195,00	2
31	1400,00	31,97	120,50	5	29,50	1	1	4	1250,00	3
32	1370,00	31,97	90,50	3	29,50	1	1	4	1250,00	2
33	1320,00	31,97	90,50	3	29,50	3	2	3	1200,00	1
34	510,00	31,97	90,50	3	29,50	1	1	2	390,00	1
35	1282,40	31,97	120,00	4	772,40	9	2	2	390,00	9
36	706,69	31,97	34,45	2	282,24	5	2	2	390,00	6
37	1140,46	31,97	347,03	4	403,43	11	2	4	390,00	11
38	731,08	31,97	98,40	2	242,68	7	2	2	390,00	8
39	614,50	31,97	165,50	5	59,00	2	1	2	390,00	2
40	630,00	98,33	150,50	5	89,50	2	2	2	390,00	2
41	912,50	98,33	375,00	2	147,50	9	3	2	390,00	7
42	570,00	98,33	121,00	4	59,00	1	1	2	390,00	2
43	2499,00	98,33	450,00	4	49,00	4	2	2	2000,00	4
44	710,35	98,33	269,00	3	51,35	8	5	2	390,00	8
45	510,00	98,33	90,50	3	29,50	1	1	2	390,00	1
46	479,50	98,33	60,00	5	29,50	1	1	2	390,00	2
47	766,00	98,33	228,50	3	147,50	2	2	2	390,00	7
48	479,50	98,33	60,00	2	29,50	1	1	2	390,00	2
49	510,00	98,33	90,50	3	29,50	1	1	2	390,00	1
50	598,50	98,33	90,50	3	118,00	5	1	2	390,00	4