

Estudio de la Potencialidad del Comercio Electrónico en la Región de Murcia mediante un Modelo de Regresión Logística

(*) BERNAL GARCÍA, JUAN JESÚS

(**) MARTÍNEZ MARIA-DOLORES, SOLEDAD MARÍA

(***) SÁNCHEZ GARCÍA, JUAN FRANCISCO

Dpto. de Métodos Cuantitativos e Informáticos. Universidad Politécnica de Cartagena.

E-mails: (*) juanjesus.bernal@upct.es - (**) soledad.martinez@upct.es - (***) jf.sanchez@upct.es

RESUMEN

A través de una investigación de carácter empírico se han determinado las características que impulsan o retraen a una empresa murciana a decidirse a plantear políticas de comercio electrónico. Este trabajo muestra, de las variables planteadas, aquellas que los agentes decisores consideran de mayor importancia, bien como objetivos a alcanzar o como barreras a solventar, detectando además que existen ciertas características intrínsecas a la organización que también ejercen una fuerte influencia y que no dependen de las decisiones que se puedan tomar en el seno de la misma.

Palabras clave: modelos, regresión logística, e-commerce

Study of the potentiality of the electronic commerce in the Region of Murcia by means of a model of logistic regression

ABSTRACT

Through an investigation of empirical character one has determined the characteristics that they impel or contain to a Murcia's company to be decided to raise policies of electronic commerce in its company. This work shows the raised variables, that the agents who make the decisions, consider of greater importance or like objectives to reach or barriers to resolve, detecting in addition that certain intrinsic characteristics to the organization exist that also exert a strong influence and that do not depend on the decisions that can be taken in the same one.

keywords: modelling, logistic regression, e-commerce

JEL classification: C5 (Econometric Modeling)

Artículo recibido en septiembre de 2004 y aceptado para su publicación en febrero de 2006.

Artículo disponible en versión electrónica en la página www.revista-eea.net, ref.: e-24111.

1. INTRODUCCIÓN¹

Apenas sin darnos cuenta estamos siendo partícipes de un gran cambio en la Sociedad, de una nueva Revolución, de un modo diferente de entender el bienestar, la economía, la forma de hacer y concebir las cosas cotidianas. Se han modificado desde los hábitos en el quehacer diario, hasta las estructuras sociales, económicas y políticas más profundas, envolviéndonos en un proceso continuo de crecimiento y aprendizaje. Todo ello impulsado por el nacimiento de diferentes innovaciones a nivel tecnológico que todavía están en continuo desarrollo.

La incorporación de la tecnología de redes, de Internet, y en consecuencia del comercio electrónico, supone un impacto directo sobre la forma habitual de realizar los procesos, del contexto social en el que se desenvuelven, además de ser un generador de nuevas necesidades a la hora de gestionar las organizaciones. Para ello las empresas deben de incorporar medios técnicos que puedan convivir con los ya existentes de forma coordinada.

Esta nueva concepción social ha llegado a las organizaciones empresariales, a los clientes o consumidores potenciales y también al resto de organismos involucrados. El comercio electrónico es una herramienta importante en esta Revolución aún latente, y abre un horizonte muy amplio de posibilidades que hay que saber aprovechar y seleccionar. Pero para ello es fundamental conocer cuál es la situación de partida y qué obstáculos o alicientes vamos a encontrar durante este camino.

2. MARCO DE LA INVESTIGACIÓN

Para centrar un tema tan amplio y que además puede resultar muy variable si se prolonga su estudio en el tiempo, nos planteamos como objeto principal de la investigación las prácticas que sobre comercio electrónico se realizaban en las empresas de la Región de Murcia. Escogimos realizar el trabajo de campo en esta región porque los datos que teníamos acerca de la misma eran todos a nivel nacional y muy escasos. Éstos presentan a la Región de Murcia en una posición poco adelantada frente a otras Comunidades Autónomas, tanto a nivel de implantación, como de desarrollo y evolución del comercio electrónico.

Por ejemplo, la Región de Murcia se sitúa en el duodécimo lugar de las veinte regiones europeas con mejores expectativas de crecimiento en el 2004, según un estudio efectuado a más de cien mil empresas por todas las Cámaras de Comercio de Europa. Las ventas previstas en ese año se igualaron a la media española duplicando

¹ Queremos agradecer de forma expresa el trabajo realizado por los evaluadores de este artículo por sus sugerencias y comentarios para mejorar el mismo.

el promedio europeo, considerando sobre todo un incremento en el mercado nacional cercano al 68%². Por otro lado, respecto al estado de la tecnología en Murcia, en un primer sondeo efectuado por la Dirección General de Economía de la Consejería de Hacienda³ para conocer el consumo de equipamiento tecnológico se expone que hay un 30,6% de familias en esta región que disponen de ordenador, es decir, una de cada tres viviendas, mientras que en España el ratio se sitúa en el 32,5%. Además, de éstos, solamente el 11% de los murcianos está conectado a Internet distanciándonos por tanto del 14,3% existente de media en el total del país.

Fijándonos en los informes de n-Economía⁴ sobre penetración mundial de las TIC, nos encontramos que el número de usuarios de Internet a nivel mundial se aproxima a los 683 millones siendo Estados Unidos el país con un mayor porcentaje, alrededor de 185 millones de personas, siendo por tanto la penetración de un 63,2%, mientras que Suecia es el país que presenta un mayor índice, situado en el 75,8%. La media europea de penetración en la red es del 45%. En dichos informes también se indica que en España la comunidad con un porcentaje de población internauta mayor respecto al total es Madrid con un 29,6% seguida de Cataluña con el 28,3% y del País Vasco con el 24,7%.

Si seguimos analizando el contexto en el que vamos a desarrollar la investigación, comprobamos que según datos del Observatorio de Información Tecnológica europea EITO, en España se invierten unos 262 euros al año por habitante en Tecnologías de la Información, un 40% de la media europea y un 17,4% respecto al promedio estadounidense, desfase que afecta claramente a las pymes. La Asociación Española de Empresas de Tecnologías de la Información (SEDISI) a su vez, indica que el gasto español en Tecnologías de la Información debería alcanzar al menos un incremento del 15% anual durante los próximos cuatro años para que nuestro país pudiera alcanzar la media de los países europeos. Aún así el 53% de las pymes realizan comercio electrónico porque, según ellas, reduce los tiempos de aprovisionamiento y obtienen un ahorro en costes⁵. Según el II Barómetro de la Pyme Española de 2003⁶, solamente el 10,59% de estas empresas considera a las Nuevas Tecnologías como estratégicas para desarrollar su negocio. El correo electrónico, con un 51,26%, es lo más utilizado, usándolo después como un medio más de promoción. Este estudio también indica que solamente el 14,18% de las pymes realiza comercio electrónico. El informe "La Sociedad de la Información en España" presentado por Telefónica nos muestra que

² Según este estudio son Dinamarca y Austria los países que superan a España en la mejoría esperada en el negocio, duplicando la media de la Unión Europea y las que peores previsiones presentan son Italia y Alemania.

³ Extracto del Diario La Opinión de Murcia, Domingo 17 de Febrero de 2002.

⁴ www.n-economia.es

⁵ *Dealer World* 22/09/2003.

⁶ *American Express* elaboró este estudio mediante encuesta a 1.270 pymes de las Comunidades Autónomas de Madrid, Cataluña, Valencia, País Vasco y Andalucía.

el 64,5%⁷ de las pymes españolas cuenta con acceso a Internet mientras que el 24% tiene su propio sitio.

Otro dato a nivel regional referente al gasto en inversión es que, por ejemplo en Cataluña, era del 1,27% del PIB, lo que supera la media española alejándose de comunidades como Madrid con el 1,9% o el País Vasco con el 1,32%. Estas tres comunidades son las líderes en cuanto a inversión en España, y la Región de Murcia se encuentra muy alejada de esas cifras.

Todo este contexto nos llevó a intentar descubrir los factores que podían estar obstaculizando el desarrollo del comercio electrónico en las empresas murcianas, y al mismo tiempo cuáles podían ser potenciadores de un crecimiento del mismo.

Para ello, planteamos una sistemática de encuestas que nos sirvió para solventar uno de los principales problemas con los que contábamos: la ya mencionada carencia de datos específicos existente en esta región. Es cierto que existe información de forma general acerca de todas las Comunidades Autónomas, pero también lo es la escasez de datos adecuados al caso de la Región de Murcia, ya que no existía información de campo para realizar ningún tipo de examen fiable y eficaz sobre los mismos.

Algunos de los objetivos perseguidos y cuyos resultados planteamos en este artículo son realizar una investigación, sobre todo de carácter empírico, que nos permitiera por un lado aproximarnos a la situación de las empresas de la Región de Murcia respecto al Comercio Electrónico, obteniendo datos reales y fiables que nos posibilitaran, a su vez, caracterizar de forma general a las empresas que finalmente deciden hacer comercio electrónico en base a una serie de variables. El valor de dichas variables dependerá a su vez de los agentes decisores en dichas empresas; es decir de que éstos consideren que tienen una mayor importancia como objetivos o barreras a salvar en su negocio. También pretendemos detectar si las características intrínsecas de la empresa murciana influyen a la hora de llevar a término políticas de Comercio Electrónico en la misma.

3. MARCO TEÓRICO DEL ANÁLISIS DE REGRESIÓN LOGÍSTICA

El tema propuesto en nuestro estudio, consistía claramente en el análisis de un fenómeno que no presenta datos de carácter continuo sino discretos, por este motivo los modelos de respuesta cualitativa⁸ son los que dan respuesta a cuestiones donde la variable dependiente es discreta y presenta un carácter dicotómico. La pregunta que contenía en el cuestionario los datos referentes a si una empresa realiza o no comercio

⁷ Este porcentaje es algo menor si nos referimos a un estudio del DMR Consulting situándose en un 62% aproximadamente.

⁸ *Ameniya Takeshi "Modelos de respuesta cualitativa: un examen"* Cuadernos Económicos del ICE. N° 39. 1988/2

electrónico, y que denotaremos a partir de ahora como P6, tiene esta característica, por lo que aunque lo más habitual es utilizar técnicas de regresión lineal, en esta ocasión no eran aplicables debido a que la variable dependiente presenta dos posibles valores únicamente. Tampoco empleamos el análisis discriminante debido a que éste no permite utilizar variables independientes de tipo categórico, lo que representaba nuevamente un inconveniente, ya que las variables relacionadas con nuestra variable objeto de estudio aparecen evaluadas en la encuesta mediante escalas *Likert*, y son estimadas de forma cualitativa por el sujeto según las considere de mayor o menor importancia. Por este motivo se utiliza un análisis de regresión logística, ya que podemos distinguir claramente la variable dependiente entre los que *SÍ* realizan alguna política de comercio electrónico y los que *NO*.

Recordamos aquí que el objetivo principal que pretende resolver la regresión logística es el intentar modelar cómo influye en la probabilidad de aparición de un suceso la presencia o no de diferentes factores y el nivel que éstos presenten. Como ya hemos sugerido, las variables independientes pueden tener un carácter cualitativo o cuantitativo. La caracterización de la variable respuesta obtiene el valor 0 cuando no se presenta el suceso, en nuestro modelo cuando la empresa no realiza ninguna actividad de comercio electrónico, y el valor 1 cuando sí ocurre el fenómeno tratado, es decir cuando sí que realiza algún tipo de práctica en torno al comercio electrónico.

Pretendemos conocer y cuantificar si existe relación entre la existencia o no de políticas de comercio electrónico en función de una variable, como puede ser por ejemplo la antigüedad de la empresa. Aunque ya hemos explicado por qué no la utilizamos, comprobamos que si presentamos el modelo según una regresión lineal obtendríamos algo similar a esta expresión:

$$SI/NO \text{ Ce} = \alpha + \beta (\text{antigüedad})$$

Por medio de la estimación de mínimos cuadrados ordinarios a partir de los datos que tengamos podemos conseguir el valor de los parámetros α y β , aunque utilizando este sistema alcanzaríamos resultados no válidos ya que no podemos asegurar que las probabilidades del modelo realmente sean tales, porque si modificamos los valores de la variable independiente, en nuestro ejemplo la variable antigüedad, obtendremos resultados en esa ecuación distintos a los que presenta la variable dependiente, originando probabilidades que no son posibles y a su vez varianzas negativas.

Utilizando la probabilidad p de que una empresa realice comercio electrónico como variable de respuesta podemos operar con la siguiente función:

$$h \frac{p}{1-p}$$

Utilizando esta expresión sí que obtenemos una variable que puede obtener cualquier valor. Si planteamos una ecuación de regresión lineal como la anterior obtenemos:

$$\ln \frac{p}{1-p} = \alpha + \beta (\text{antigüedad})$$

De ahí podemos alcanzar mediante cálculos algebraicos la siguiente expresión:

$$\text{PrSI/NOCE} = \frac{1}{1 + e^{(-\hat{\alpha} - \hat{\beta}(\text{antigüedad}))}}$$

Esta ecuación corresponde con un modelo logístico en donde el número de factores a tener en cuenta puede ser más de uno, incorporándose al mismo con su correspondiente coeficiente. A dichos factores se les denomina *factores de riesgo*⁹. El modelo permite conocer la importancia y la contribución de determinadas características y posibilita estimar o predecir la magnitud del riesgo global de que ocurra el suceso cuando concurren dos o más factores. El modelo de regresión logística se utiliza para predecir la probabilidad estimada $P(Y)$ de que la variable (Y) presente uno de los valores posibles que son 1, es decir que sí ocurre el suceso, y 0, en caso contrario, en función de los valores que presenten las variables independientes.

También es conveniente conocer mediante esta técnica de regresión la relación de los coeficientes estimados en el modelo. Se realiza mediante la determinación de un parámetro de cuantificación de riesgo denominado *odds ratio (OR)*. Este parámetro asociado a un suceso es la *razón* entre la probabilidad de que ocurra dicho suceso frente a la probabilidad de que no ocurra:

$$\text{odds} = \frac{p}{1-p}$$

Al logaritmo de este cociente mismo se le conoce como *Logit estimado de L*:

$$L = \ln \left(\frac{p}{1-p} \right)$$

El valor de L puede oscilar entre los valores $-\infty$ para la probabilidad igual a 0 y $+\infty$ cuando vale 1. Si este valor aumenta también lo hace la probabilidad estimada de que un valor haga que se cumpla el suceso.

Cuando contamos en la regresión con una variable independiente que también tenga carácter dicotómico, su coeficiente está directamente relacionado con el *odds ratio* de que exista ese factor o que no exista:

$$(\text{OR}) = \exp(\beta)$$

siendo $\exp(\beta)$ la medida de cuantificación del riesgo en función de que dicha variable esté presente o no, si suponemos constantes al resto de variables presentes en el modelo.

⁹ Albert J. Novell. Análisis de Regresión Logística. Cuadernos Metodológicos. CIS. 1995.

Si en la regresión se presentan variables independientes que sean numéricas, el *OR* lo que cuantifica es el cambio de riesgo cuando se pasa de un valor a otro en esa variable, siempre y cuando el resto permanezca constante:

$$(OR) = \exp(\beta^*(X_2 - X_1))$$

Si el coeficiente obtenido en una variable es positivo, obtenemos entonces un *odds ratio* mayor que 1, lo que implica que esta variable es considerada tal y como hemos mencionado antes como un *factor de riesgo*, mientras que si β es negativo el *odds ratio* será menor a 1 por lo que se le denominará *factor de protección*.

Las características de un modelo *logit* por tanto se puede decir que son¹⁰:

- La probabilidad P va de 0 a 1 y por lo tanto L están entre $-\infty$ y $+\infty$, por lo que los modelos *logit* no tienen límites.
- Aunque L sea lineal para una determinada variable independiente, las probabilidades de ésta no lo son.
- β es la pendiente que mide el cambio en L por cada unidad que cambia su variable independiente correspondiente.

Los parámetros β de la ecuación de regresión logística se pueden estimar por el método de máxima verosimilitud, que se basa en la estimación de los valores de β poblacionales que maximizan la función logística para el conjunto de valores muestrales, reflejando de la mejor forma posible los datos observados en la muestra¹¹. Al contrario de lo que ocurre en los modelos lineales de probabilidad, en donde la influencia que tienen las variables explicativas sobre la probabilidad de que ocurra el suceso depende sólo de los coeficientes, con los modelos de regresión logística conseguimos que también dependa de las variables independientes, ya que los coeficientes estimados miden la relación lineal existente entre el índice Li ($I_i = x'_i \beta$) y las variables independientes (x_{ij}).

Para comprobar la bondad del ajuste se utilizan¹²:

- **-2 Log Likelihood (-2LLO) y Goodness of fit.** Estos dos estadísticos contrastan como hipótesis nula que el modelo es significativo y como hipótesis alternativa que no lo es. Para el estadístico *-2LLO* el modelo será válido cuando tenga una probabilidad asociada mayor a 0,05; sin embargo para el de *Goodness of fit* el modelo será significativo cuando tenga una probabilidad asociada inferior a 0,05. Este último estadístico se calcula por la siguiente expresión:

$$\sum_{i=1}^N \frac{E_i^2}{P_i(1 - P_i)}$$

¹⁰ Greene, W. H. *Análisis Econométrico*. Prentice Hall. 1998. Madrid.

¹¹ Podemos encontrar el algoritmo de estimación de parámetros desarrollado en diversos manuales dedicados a la Econometría como por ejemplo el de Damorarn Gujarati "*Econometría*". 2ª Edición. McGraw-Hill. 1981.

¹² Santos Peña, J., Muñoz Alamillos, A. Juez Martel, P. y Guzmán Justicia, L. "*Diseño y tratamiento estadístico de encuestas para estudios de mercado*". Ed. Centro de Estudios Ramón Areces, SA. 1999.

Siendo:

E_i es el i-ésimo residuo calculado por la diferencia entre la probabilidad observada y la estimada.

P_i o probabilidad estimada del i-ésimo caso

P o probabilidad observada

Distribuyéndose como una χ^2 con n-2 grados de libertad.

- **Test de Wald.** Se utiliza para conocer si las variables introducidas en el análisis son o no válidas. Si es inferior a 0,05 diremos que la variable es significativa y válida para el modelo, aunque una variable no se ha de rechazar simplemente porque este estadístico haya señalado que no es significativo. Se determina por la siguiente expresión:

$$WALD = \frac{\beta_i}{E(\beta_i)}$$

siendo EE el error estándar del coeficiente. Este estadístico se distribuye como una χ^2 con un grado de libertad.

- **Tabla de aciertos** o porcentaje de elementos de la muestra que eligen la opción predicha en el modelo, por lo que a mayor porcentaje de casos clasificados correctamente mejor será el mismo.
- **Prueba de Hosmer-Lemeshow** (1989) que consiste en dividir el recorrido de la probabilidad en deciles de riesgo. Por ejemplo la probabilidad de que si exista una política de comercio electrónico en la empresa $\leq 0,1$; $\leq 0,2$ y así sucesivamente hasta ≤ 1 , calculando así la distribución esperada y observada, contrastándose mediante una prueba de χ^2 .

Con respecto a los datos necesarios para que una regresión logística sea óptima podemos plantear que el tamaño de la muestra que se necesita ha de ser al menos de $10 \cdot (k+1)$ casos para un modelo con k variables independientes, lo que significa 10 casos al menos por cada variable que interviene en el modelo. Como podemos comprobar el tamaño muestral de nuestro análisis es correcto, ya que en el modelo final permanecen 15 variables independientes y el tamaño muestral que consideramos es de 450, cifra considerablemente superior a los 160 datos que según este ratio se considerarían óptimos.

Hemos de tener en cuenta si existe multicolinealidad para así determinar finalmente si el modelo es adecuado o no, es decir, hay que comprobar si existe una correlación entre las variables independientes, ya que de ser así ésta puede producir coeficientes estimados sesgados y errores estándar elevados, lo que alterará el valor de las probabilidades y también los del estadístico de Wald. Para diagnosticar si efectivamente existe multicolinealidad podemos inspeccionar la matriz de correlación entre las variables independientes, o bien examinar la matriz de correlación de los coeficientes estimados. Pese a todo, para solventar este tipo de problemas se pueden excluir variables independientes utilizadas con anterioridad en el modelo, o aumentar el tamaño de la muestra.

En la regresión logística se puede decidir el método para seleccionar las variables que se incluirán o no en el modelo, atendiendo a criterios teóricos o estadísticos. En el primer caso se aconseja la introducción forzosa de todas las variables mientras que en el segundo se establecen métodos de selección por pasos. Para ello existen distintos tipos de algoritmos¹³ que resumimos de forma muy esquemática:

1.- *Método forward*. Incorpora progresivamente de forma secuencial las variables hacia delante. Es decir, va añadiendo las variables independientes una a una, manteniendo en el modelo aquellas que resultan significativas para un determinado nivel que normalmente se fija en el 0,05 y descarta el resto. La incorporación de las variables se realiza mediante un orden preestablecido en función de las que tienen una mayor significación estadística (mayor nivel de correlación), deteniéndose el proceso cuando al añadir una variable independiente nueva no se aprecia mejora en la bondad del ajuste a los datos muestrales.

2.- *Método backward*. Se eliminan progresivamente de forma secuencial las variables hacia atrás. Al principio del análisis se tienen en cuenta todas las variables, construyendo lo que se denomina *modelo saturado* que ajusta todos los parámetros a los datos, pero que incluye variables que no tienen que ver con la variable dependiente objeto de estudio. Progresivamente se van eliminando variables comenzando por la que menos contribuye a la bondad del ajuste. Este es el sistema que hemos elegido para nuestro análisis empírico.

3.- *Método stepwise*. Este método combina los dos anteriores ya que va incluyendo paso a paso las variables de forma secuencial mientras que el nivel de significación estadística se encuentre en los límites señalados.

Para terminar este repaso teórico, vamos a centrarnos de nuevo en las variables independientes que, como hemos mencionado, pueden ser cualitativas binarias o categóricas, cuantitativas o continuas. En un modelo de regresión logística la utilización de variables cualitativas es bastante habitual aunque no pueden ser contempladas tal y como se presentan en los datos, ya que no sería adecuado el introducir los números asignados en una escala a cada categoría. Porque si a una categoría le asignamos al tabular las respuestas el valor 0, a otra un 1 y a otra un 2, no implica que la última categoría tenga un efecto dos veces superior a la de la primera; simplemente es una escala. Por esta razón se crean tantas variables dicotómicas como número de respuestas posibles existan, menos una. Estas variables son conocidas como *dummy* o *variables internas, indicadoras o variables diseño*, y se crean mediante la asignación de valores 0 y 1. En estas variables el coeficiente asignado en la ecuación de regresión obtenida para cada variable *dummy* es igual al *odds ratio* de esta categoría con respecto al nivel de referencia utilizado, es decir la variable interna que presenta todo ceros y que se excluye del análisis.

¹³ Ameniya Takeshi "Modelos de respuesta cualitativa: un examen". Cuadernos Económicos del ICE. Nº 39 1988/2

4. DETERMINACIÓN DE LA MUESTRA

Tabla 1. Cálculo del Tamaño Muestral

Cálculo del Tamaño Muestral			
1. No se conoce el tamaño de la población			
Probabilidad	95%		
Alfa (α)	0,05		
Nivel confianza (Z)	97,50%	1,96	(Ajustado a la normal)
Variación positiva (p)	50%		
Variación negativa (q)	50%		
Error o precisión (E)	4,00%	Error o precisión (E)	4,62%
Tamaño muestral población no conocida (n)	600	empresas	450
2. Si se conoce el tamaño de la población			
Tamaño población (N)	100.000	empresas	
Tamaño muestral población si es conocida (N)	597	empresas	448
Proporción presente (P)	0,00596646		
Tolerancia de la muestra	0,006179429		
Intervalo de confianza (n) =	0,012145889	<i>superior</i>	<input type="checkbox"/> 1.214
(P \pm tolerancia de la muestra)	-0,000212969	<i>inferior</i>	<input type="checkbox"/> - 22

Tabla 2

Total Encuestas Obtenidas	537
Total Encuestas Válidas	450
× Web	152
× Fax	21
× Correo Electrónico	219
× Entrevista no guiada	58

La presente investigación se realizó sobre una muestra de empresas de la Región de Murcia escogidas de una población cuyo tamaño desconocíamos a priori. Por esta razón para un nivel de confianza del 97,5% ajustado a una curva normal, con un 95% de probabilidad y asumiendo un 4% de error, obtuvimos que sería necesario un tamaño muestral de 600 empresas. Si asumimos un mayor error, un 4,62% obtenemos una

muestra para una población no conocida de 450 empresas (*Tabla 1*)¹⁴. Según el DIRCE del INE la población total de empresas en este periodo de tiempo estaba en torno a las 75.973 empresas, por lo que el error soportado sería del 4,61% para este mismo número de encuestas. Se consiguieron un total de 537 respuestas, de las que finalmente resultaron válidas 450, que se obtuvieron por los medios que se presentan en la *Tabla 2*.

El cuestionario se estructuró de la siguiente manera:

1. Datos generales, para conocer detalles de las personas y/o empresas que estábamos entrevistando tales como el sexo del encuestado, el sector al que pertenecía la empresa, la antigüedad en la misma y a su vez también de la empresa, la categoría según el número de trabajadores, y el conocimiento que estimaba que tenían respecto a temas de comercio electrónico tanto el encuestado como el nivel general de los trabajadores de su empresa.

2. Cuestiones en torno a las acciones y barreras que fomentarían o pueden resultar un inconveniente en el desarrollo del Comercio Electrónico en la Región de Murcia, así como el tipo de empresa que consideran más idónea. Esta parte consta de 2 preguntas de valoración cualitativa sobre una escala de Likert de 1 a 5, una pregunta de jerarquización de factores respecto a las respuestas previamente planteadas en una cuestión anterior, y otra pregunta de valoración cualitativa sobre cinco perfiles dados.

3. Este bloque contiene la pregunta que creemos el eje principal de nuestra encuesta y que cuestiona si efectivamente se está realizando una política de comercio electrónico en la empresa consultada. Si esta respuesta resultaba ser afirmativa, se pasaba a una pregunta abierta en donde se interrogaba sobre qué es lo que se hace y que nos permitía saber lo que efectivamente se entiende en la empresa murciana por realizar una política de comercio electrónico. Si la respuesta era negativa, se iba a otra pregunta en la que se indicaba mediante una escala de evaluación de 1 a 10 el grado en el que se encuentra su empresa en la actualidad para iniciarla.

4. Pasamos a un cuarto bloque en el que se presentan dos preguntas sobre cuestiones más concretas acerca de objetivos y barreras que pueden afectar directamente a las empresas murcianas, más otras dos cuestiones acerca del perfil del consumidor murciano sobre los factores que pueden favorecer o ejercer de barreras para realizar transacciones a través de la red. Estas cuatro preguntas presentan distintos ítems que se debían de evaluar mediante una escala Likert de 1 a 5. En este bloque se presenta una última pregunta acerca de las características consideradas clave en una página web para generar confianza y/o fidelidad en el consumidor, cuya valoración se presentaba mediante una escala Likert de 1 a 10.

5. Se efectuaron por último dos preguntas: una de estimación cuantitativa puntual de carácter absoluto, con la que se pretendía estimar una propiedad cuantitativa futura

¹⁴ Estos cálculos se han realizado mediante la hoja de cálculo Microsoft® Excel siguiendo la formulación de la Teoría de Muestras tradicional para el cálculo del tamaño muestral.

sobre la fecha real del despegue del comercio electrónico en las empresas de la Región de Murcia, y otra cuestión para conocer las peticiones que harían las empresas respecto a estos temas a la Administración Regional, proponiendo en ambos casos unos ítems pero dejando uno abierto para nuevas sugerencias.

La reiterada utilización del sistema de evaluación mediante escalas de tipo Likert, se fundamenta en el hecho de que dicho sistema aporta de forma clara y contundente una mejora en la jerarquización clásica cuando se presenta un número elevado de ítems a evaluar, tal y como se planteaba en nuestro caso. Además la evaluación dentro de un intervalo determinado es más sencilla para el entrevistado de completar frente a una ordenación de distintas proposiciones sobre su opinión respecto a un determinado tema, sobre todo si dicho proceso se realiza de forma reiterada.

Por otro lado, se ha de tener en cuenta que con este tipo de valuaciones se obtienen respuestas sobre escalas cualitativas a las que podemos asignar un valor cuantitativo. Además, los valores que otorga el encuestado a cada uno de los ítems considerados nos permite, a su vez, realizar una jerarquización consiguiendo así información importante acerca de lo que realmente piensa. Además este sistema fue el que nos dio un resultado más adecuado en el análisis con expertos que se realizó previamente, por lo que se adoptó en el nuevo proceso de la investigación.

Algunas de las cuestiones planteadas tienen una escala de evaluación superior, de uno a diez. Estas preguntas fueron diseñadas así de forma expresa, ya que valoraban bien conocimientos por lo que esta escala resulta más habitual en los ámbitos docentes, o bien ítems que representaban atributos y/o características de una página web, con las que se pretendía realizar un análisis basado en un modelo multiatributo siguiendo técnicas de descomposición, como el análisis de componentes principales. El obtener un mayor grado de segregación y de información entre los distintos atributos nos permitiría conseguir un mejor ajuste en este análisis.

Al estar compuesta la encuesta en su mayoría por variables de tipo categórico, se realizaron los correspondientes análisis de bondad del ajuste, independencia y fiabilidad de las respuestas obtenidas, no presentando ninguna de las variables ajuste a ninguna distribución conocida, una vez realizadas las pruebas de Chi-cuadrado de Pearson, Binomial o de Kolmogorov-Smirnov en función del tipo de variable que estuviéramos evaluando.

Destacar que del total de empresas encuestadas, el 33,33% afirma realizar alguna actividad de comercio electrónico, aunque si desglosamos aún más sus respuestas podemos concluir que e-commerce completo solamente es efectuado por el 52,38% de las que contestaron de forma positiva. Resaltar que mediante estas respuestas hemos podido detectar que, en muchas ocasiones, para una empresa el realizar comercio electrónico implica simplemente tener una presencia en la web mediante una página muy simple, insertar su nombre en una base de datos del sector o utilizar el correo electrónico para ponerse en contacto con los clientes. Es cierto que estas actividades pueden ser sólo los primeros pasos dentro del comercio electrónico, pero también lo es que estas empresas al menos demuestran un cierto interés inicial con respecto al

uso de esta tecnología en sus negocios y que al menos han conseguido vencer las reticencias iniciales que existen respecto a un medio desconocido.

5. Modelización de la probabilidad de que una PYME murciana realice comercio electrónico

El cuestionario se estructuró en varios grupos que incluían diferentes variables según el tema a consultar. Para esta modelización se han escogido: las preguntas 8 y 10, que se refieren a la importancia de una lista de 12 objetivos y de 11 barreras, respectivamente, a la hora de que una empresa en la Región decida instaurar políticas de comercio electrónico; las preguntas 9 y 11 que instan a evaluar los factores más importantes como estímulos (7 ítems) o barreras (10 ítems) para que el consumidor se decida a realizar transacciones en la red. La relación de variables predictoras introducidas en el modelo inicialmente se presenta en la *Tabla 3*.

Tabla 3. Variables Independientes introducidas en el modelo saturado

P.2 Barreras básicas nivel regional	P.8 Objetivos a alcanzar por las empresas	P.10 Barreras de entrada para las empresas
<ul style="list-style-type: none"> • Falta de recursos tecnológicos • Desconocimiento de los usos de nuevas tecnologías • Falta de legalización y normativa • Factores psicológicos contrarios a la realización de operaciones online • Alto nivel de incertidumbre del medio/mercado • Miedo ante posibles fallos en la seguridad de la transacción • Inadecuación de los productos ofertados • Posible carencia de clientes entre los usuarios de Internet • Falta de apertura socio-económica de la Región 	<ul style="list-style-type: none"> • Posibilidad de ahorro en costes • Venta • Estandarización de productos • Disminución de tiempo en las operaciones • Acortamiento de las cadenas de distribución • Adaptación a la demanda y mejor segmentación • Ventaja competitiva ante los competidores • Apertura a una presencia global • Nuevas oportunidades de negocio y de mercados • Buscar información • Generar nuevos clientes • Publicidad / Promoción 	<ul style="list-style-type: none"> • Inexperiencia en el ámbito de las nuevas tecnologías • Idioma • Formación • Seguridad • Legislación • Costes asociados • Falta de desarrollo tecnológico • Inadecuación de los productos • Inexistencia de una concentración adecuada de consumidores • Hábitos de los consumidores potenciales • Arraigo de la cultura empresarial reacia a los cambios
P. 9 Factores que facilitan la compra en el consumidor	P. 11 Factores negativos para comprar por Internet en el consumidor	
<ul style="list-style-type: none"> • Comodidad • Disminución del precio • Disminución del tiempo • Personalización • Anonimato • Producto difícil de encontrar o no disponible en tiendas físicas • Acceso a productos de carácter internacional 	<ul style="list-style-type: none"> • Dificultad de acceso a Internet • Lentitud en el acceso a las direcciones buscadas • Coste añadido de la tecnología • Falta de legislación • Falta de información • Sensación de inseguridad • Falta de privacidad • Desconfianza del medio • Incertidumbre sobre la veracidad de la información • Incertidumbre sobre cómo se realizará el proceso 	

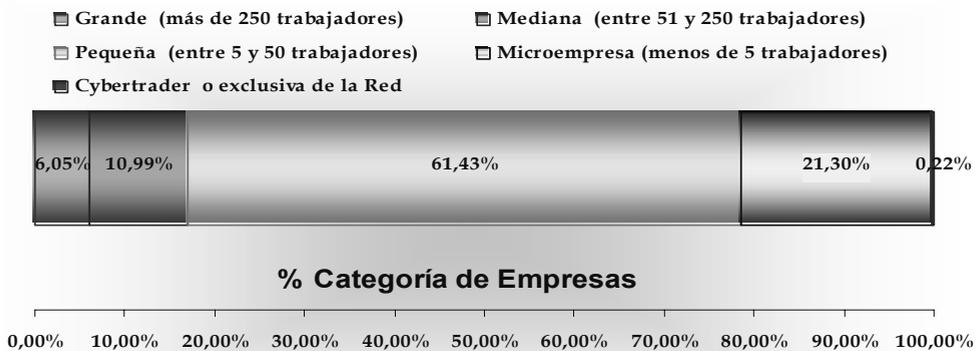
La nomenclatura utilizada para cada una de estas variables se compone mediante el número de pregunta del cuestionario en el que se encuadran, más el número de *ítem* por orden de entrada en la lista. También se incluye una pequeña descripción para favorecer la comprensión de las mismas.

Por otro lado, existen preguntas de carácter general que describen el tipo de empresa sobre el que estamos trabajando y que consideramos importantes para introducir también como variables independientes, tales como el *Sector* al que pertenece la empresa, la *Categoría* en la que se sitúa la misma y la *Antigüedad*. La variable *Sector* es una variable consistente en tres opciones que desglosamos en la *Tabla 4*, y la *Categoría* a su vez se divide en cinco posibilidades de elección que presentamos en la *Figura 1*. Así, nos encontramos con un conjunto de variables independientes, casi todas de tipo categórico por estar evaluadas mediante escalas Likert 1 a 5 o categorías específicas, con la excepción de la pregunta *Antigüedad* de la empresa, que se considera como cuantitativa.

Tabla 4

Sector		Frecuencia	%
Válidos	1. Servicios	244	54,2
	2. Industria	157	34,9
	3. Construcción	49	10,9
	Total	450	100,0

Figura 1



Se han realizado diversas regresiones, que de forma paulatina han ido incorporando variables predictoras a la regresión. Así pudimos comprobar que la pregunta que incluía variables que nos proporcionan información acerca de los elementos que se consideran que pueden influir negativamente para que el consumidor compre, sí que ha contribuido significativamente a la mejora en el modelo de regresión planteado. Sin embargo las

variables planteadas en la cuestión propuesta a valuación acerca de los factores que inciden en el consumidor favorablemente para que éste compre por la red, han resultado ser no relevantes para nuestra regresión. Esta no inserción de variables tales como comodidad, disminución del precio y del tiempo, personalización, anonimato, etc..., como covariables independientes de la regresión, puede estar ocasionada por el hecho de que son percepciones que la empresa considera que logra tener el consumidor y que consiguen incentivarle a realizar transacciones en la red, pero que realmente no pueden ser manejadas por la empresa, por lo que no van a tener un peso específico a la hora de incidir sobre si ésta realiza políticas de comercio electrónico o no.

Por otro lado las barreras con las que se encontrará un consumidor a la hora de acceder a nuestro negocio en la red, o en realidad la percepción que de ellas tiene la empresa, sí que resultan un motivo suficiente para que el negocio se instale en la web o no.

Tal y como mencionamos en el inicio de este artículo, en la regresión logística se puede decidir el método para seleccionar las variables, y en distintas fases de la investigación se han seguido las distintas sistemáticas de análisis utilizando diferentes métodos de introducción de variables. Finalmente se ha optado por el método hacia atrás (*backward*) que excluye las variables basándose en la razón de verosimilitudes (*RV*).

Definitivamente se han introducido las variables correspondientes a las preguntas *Sector Codificado*, *Categoría de la Empresa*, *Antigüedad de la Empresa*, y todas las de las preguntas 8, 10 y 11. Hacer hincapié de nuevo en el hecho de la necesidad de categorizar las variables que obtienen valoraciones cualitativas obtenidas mediante escalas.

Tabla 5

	Nº variables	α de Cronbach
P.1	8	0,8349
P.2	9	0,7673
P.8	12	0,8596
P.9	7	0,7356
P.10	11	0,8428
P. 11	10	0,8311
P.12	20	0,9539

En la *Tabla 5* se presenta el *parámetro alfa o de Cronbach* obtenido para cada una de las preguntas que han sido objeto del análisis de regresión logística presentado. Este parámetro o coeficiente de fiabilidad se puede interpretar como el *coeficiente de correlación R^2* de un análisis de regresión lineal, por lo que cuanto más próximo se encuentre a uno mejor será la consistencia interna entre las variables contrastadas. Depende del número de elementos de la escala. Aunque en general todas las preguntas tienen un buen coeficiente, las preguntas 2 y 9 son las que obtienen un valor inferior (0,7). Tal vez éste también sea un motivo por el que al realizar la regresión éstas cuestiones sean las que no han incorporado ninguna variable al modelo final. Podemos afirmar por lo tanto, que existe un elevado nivel de consistencia interna entre las variables de nuestro estudio.

Utilizamos el contraste Indicador para categorizar las variables independientes, mediante el cual cada categoría se compara con la categoría señalada de referencia, y ésta última no se evalúa, asignándole zeros como valor.

En nuestro análisis se ha utilizado la categoría última como referencia, ya que queríamos introducir expresamente dentro del mismo el sector al que pertenece la empresa con la categoría uno, es decir la que nos arroja información acerca de las empresas pertenecientes al sector servicios, debido a la importancia y el peso que distintos estudios adjudican a dicho sector en la Región de Murcia¹⁵. Respecto a la muestra analizada, en donde se partía de un total de 450 casos introducidos para el análisis, finalmente se contó con 434 ya que 16 se consideraron no válidos.

En la *Tabla 6* se presentan los estadísticos de bondad de ajuste global para el modelo estimado, que solamente pueden ser tenidos en cuenta de forma orientativa, debido a que la variable de objeto de estudio es una variable categórica, lo que puede llevar a que dichos coeficientes sean moderados, dando la sensación de que el modelo no es adecuado cuando en realidad tal vez esto no sea así. A continuación se muestra la *matriz de confusión*, o lo que es lo mismo, la clasificación de los casos en el *paso cero* (*Tabla 7*).

Tabla 6

Historial de iteraciones (a,b,c)			
Iteración		-2 log de la verosimilitud	Coefficientes
			Constante
Paso 0	1	553,021	-,664
	2	552,955	-,690
	3	552,955	-,690

- a) En el modelo se incluye una constante.
 b) -2 log de la verosimilitud inicial: 552,95
 c) La estimación ha finalizado en el número de iteración 3 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Tabla 7

Tabla de clasificación(a,b)					
Observado		Pronosticado			
		P6		% correcto	
		0	1		
Paso 0	P6	0	289	0	100,0
		1	145	0	,0
Porcentaje global					66,6

- a) En el modelo se incluye una constante.
 b) El valor de corte es ,500

En dicha matriz de confusión se cruzan los resultados observados para la variable dependiente con los pronosticados por el modelo nulo. Es por esta razón por lo que todos los casos son clasificados según pertenezcan a la categoría en la que existe un mayor número de datos, lo que en nuestro caso supone que se ajusten a la categoría cero en un 100%, es decir, a la de las empresas que no realizan comercio electrónico alguno, mientras que para la categoría uno el porcentaje de ajuste es cero y el ajuste global es de un 66,6%. Incidir que la codificación interna utilizada para separar la

¹⁵ Según el DIRCE del INE el número de empresas dedicadas al comercio y al resto de servicios es aproximadamente del 76,69% en la Región de Murcia.

dicotomía de la variable dependiente es de un cero cuando *No* se realiza ninguna actividad de comercio electrónico en la empresa, y un uno cuando *Sí* existe alguna política respecto a este tema. Es importante destacar que la categoría con el valor uno es imprescindible en la interpretación final del modelo de regresión, porque lo que en último lugar obtendremos la probabilidad de que la variable dependiente sea uno en función de una serie de variables independientes y una constante.

El análisis realizado nos ofrece además una tabla con la información de las variables no incluidas en la ecuación en el *paso cero*, pero ya que en este paso son todas las que en principio introdujimos para su estudio, no trasladamos aquí esos resultados. Dicha tabla también nos informa del nivel crítico asociado a cada variable. Si el nivel de significación resulta inferior a 0,05, se puede rechazar la hipótesis nula de que dicha variable no contribuye a explicar de forma significativa el comportamiento de la variable estudiada. Es decir, que dicha variable independiente puede ser incorporada mejorando así el ajuste final del modelo, por lo que nos va a permitir detectar a priori cuáles de las variables, que aún no han sido valoradas, van a poder ser las que finalmente se encuentren incluidas en el modelo.

Pasando a los resultados obtenidos en los siguientes pasos, en primer lugar obtenemos una tabla denominada *historial de iteraciones*, que resulta muy extensa por lo que no la presentamos aquí, y a la que se van incorporando por el método de introducción las variables, seleccionando las que se consideran significativas para el modelo y excluyendo las que no lo son. En total se han realizado 22 pasos hasta conseguir que el modelo no siga mejorando mediante la inclusión de una nueva variable.

En la *Tabla 8* se presentan las pruebas denominadas *ómnibus* sobre los coeficientes del modelo o contrastes de ajuste global, pero en esta ocasión para cada uno de los pasos utilizados en la introducción de las variables. El valor de *Chi-cuadrado* permite contrastar que la mejora obtenida en el ajuste es nula, por lo que en cada paso nos informa de cuál es la variación producida en el ajuste como consecuencia de la incorporación o eliminación de una variable nueva.

Antes de seguir, realizaremos una aclaración que consideramos importante. Es preciso tener en cuenta que utilizamos el método de elección de variables hacia atrás o *backward* con extracción por máxima verosimilitud. Esto significa que partimos del modelo saturado en el que se incluyen todas las variables y vamos eliminando aquellas que no incorporan significación al ajuste del modelo. Por esta razón un valor de *Chi-cuadrado* negativo indica que ha disminuido este valor con respecto al paso anterior, lo cual en este tipo de metodología es lo correcto. El modelo consigue realizar el mejor ajuste en 22 pasos, y se puede observar en la *Tabla 9*, que presenta los estadísticos de ajuste global del modelo, que éste va mejorando progresivamente, disminuyendo a cada paso el valor de los mismos, conforme se eliminan variables no significativas del análisis. Hemos de fijarnos en esta referencia porque como ya se ha comentado antes, el valor de estos estadísticos de ajuste puede que nos lleve a cometer errores de apreciación, ya que pueden parecer elevados pese a que el ajuste

sea bueno. El problema reside en la caracterización de la variable dependiente objeto de estudio, puesto que al ser dicotómica implica que la cuantificación de los estadísticos puede no ser la apropiada.

Otra forma de medir la bondad del ajuste es mediante la prueba de *Hosmer-Lemeshow*, estadístico muy utilizado en el análisis de regresión logística. Suele ser muy útil en el supuesto que existan numerosas variables independientes como ocurre en este caso, lo que puede hacer que invalide la utilización de otros estadísticos de ajuste más habituales. La tabla de contingencias origina el estadístico de *Chi-cuadrado de Hosmer-Lemeshow*, donde se presentan los deciles denominados de riesgo, en los que dicha prueba divide a la muestra a partir de las probabilidades pronosticadas para cada uno de los pasos efectuados en el análisis. Cada uno de estos *deciles de riesgo* indica el número de casos que pertenecen a cada categoría de la variable dependiente y el número de casos que para esa categoría pronostica el modelo obtenido en ese paso.

Conseguiremos así el número observado y el número pronosticado. En la *Tabla 10* presentamos la tabla de contingencias para el *último paso del modelo*.

Por comparación entre las observaciones y los valores esperados se calcula el estadístico *chi-cuadrado de Pearson*. Presentamos en la *Tabla 11* los datos que muestra esta prueba para el Paso 22. Como se puede apreciar el estadístico toma un valor de 9,046 y tiene un nivel crítico de 0,338, por lo que no se puede rechazar la hipótesis nula que plantea el mismo de igualdad de distribución de los datos en todos los deciles de riesgo en los que divide la muestra, por lo que podemos concluir que los datos sí que se ajustan al modelo.

Al igual que antes con el resto de pruebas de bondad del ajuste, tenemos que indicar que el valor del estadístico puede ser muy elevado, ya que es sensible al tamaño de la muestra, pero al mismo tiempo la *prueba de Hosmer-Lemeshow* es más eficiente con muestras grandes, por lo que es un inconveniente difícil de salvar.

La nueva *matriz de confusión* generada muestra el resultado de la clasificación para cada uno de los pasos. Hemos utilizado el punto de corte habitual en este tipo de regresiones, que es del 0,5. Con este procedimiento introduciendo todas las variables originales que ya se han mencionado, se obtiene el porcentaje global mejor, llegando a explicar un 79,3% de todos los casos y el 90,3% de aquellas empresas que deciden no realizar ninguna política de comercio electrónico, frente al 57,2% de las que eligen sí realizar alguna (*Tabla 12*).

Tabla 8. Pruebas ómnibus sobre los coeficientes del modelo

		Chi-cuadrado	gl	Sig.			Chi-cuadrado	gl	Sig.
Paso 1	Paso	247,857	139	,000	Paso 12	Bloque	-3,638	4	,457
	Bloque	247,857	139	,000		Modelo	223,536	95	,000
	Modelo	247,857	139	,000		Paso	223,536	98	,000
Paso 2	Paso	-,689	4	,953	Paso 13	Bloque	-4,118	4	,390
	Bloque	247,168	135	,000		Modelo	219,418	91	,000
	Modelo	247,168	138	,000		Paso	219,418	94	,000
Paso 3	Paso	-,788	4	,940	Paso 14	Bloque	-3,677	4	,451
	Bloque	246,381	131	,000		Modelo	215,741	87	,000
	Modelo	246,381	134	,000		Paso	215,741	90	,000
Paso 4	Paso	-1,032	4	,905	Paso 15	Bloque	-5,188	4	,269
	Bloque	245,349	127	,000		Modelo	210,553	83	,000
	Modelo	245,349	130	,000		Paso	210,553	86	,000
Paso 5	Paso	-,765	4	,943	Paso 16	Bloque	-5,438	4	,245
	Bloque	244,583	123	,000		Modelo	205,115	79	,000
	Modelo	244,583	126	,000		Paso	205,115	82	,000
Paso 6	Paso	-1,670	4	,796	Paso 17	Bloque	-6,699	4	,153
	Bloque	242,913	119	,000		Modelo	198,416	75	,000
	Modelo	242,913	122	,000		Paso	198,416	78	,000
Paso 7	Paso	-2,455	4	,653	Paso 18	Bloque	-6,204	4	,184
	Bloque	240,458	115	,000		Modelo	192,212	71	,000
	Modelo	240,458	118	,000		Paso	192,212	74	,000
Paso 8	Paso	-2,659	4	,616	Paso 19	Bloque	-5,801	4	,215
	Bloque	237,799	111	,000		Modelo	186,411	67	,000
	Modelo	237,799	114	,000		Paso	186,411	70	,000
Paso 9	Paso	-3,181	4	,528	Paso 20	Bloque	-5,435	4	,246
	Bloque	234,618	107	,000		Modelo	180,976	63	,000
	Modelo	234,618	110	,000		Paso	180,976	66	,000
Paso 10	Paso	-2,976	4	,562	Paso 21	Bloque	-6502	4	,165
	Bloque	231,642	103	,000		Modelo	174,474	49	,000
	Modelo	231,642	106	,000		Paso	174,474	62	,000
Paso 11	Paso	-4,469	4	,346	Paso 22	Bloque	-6,157	4	,188
	Bloque	227,174	99	,000		Modelo	168,317	55	,000
	Modelo	227,174	102	,000		Paso	168,317	58	,000

Tabla. 9

Resumen de los modelos			
Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	305,098(a)	,435	,604
2	305,787(a)	,434	,603
3	306,575(a)	,433	,601
4	307,607(a)	,432	,599
5	308,372(a)	,431	,598
6	310,042(a)	,429	,595
7	312,497(a)	,425	,591
8	315,156(a)	,422	,586
9	318,337(a)	,418	,580
10	321,313(a)	,414	,574
11	325,782(a)	,408	,566
12	329,420(a)	,403	,559
13	333,538(a)	,397	,551
14	337,215(a)	,392	,544
15	342,402(a)	,384	,534
16	347,840(a)	,377	,523
17	354,539(a)	,367	,509
18	360,743(a)	,358	,497
19	366,544(b)	,349	,485
20	371,979(b)	,341	,473
21	378,481(b)	,331	,460
22	384,638(b)	,321	,446

a) La estimación ha finalizado en el número de iteración 20 porque se han alcanzado las iteraciones máximas. No se puede encontrar una solución definitiva.

b) La estimación ha finalizado en el número de iteración 6 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Tabla. 11

Prueba de Hosmer y Lemeshow			
Paso	Chi-cuadrado	gl	Sig.
1	8,207	8	,414
2	5,695	8	,681
3	10,191	8	,252
4	12,896	8	,115
5	6,320	8	,611
6	8,322	8	,403
7	11,855	8	,158
8	5,090	8	,748
9	12,679	8	,123
10	8,517	8	,385
11	9,065	8	,337
12	7,595	8	,474
13	7,557	8	,478
14	6,391	8	,603
15	9,243	8	,322
16	4,975	8	,760
17	11,918	8	,155
18	14,613	8	,067
19	7,140	8	,522
20	5,094	8	,748
21	5,203	8	,736
22	9,046	8	,338

Tabla 10. Tabla de Contingencias - Prueba de Hosmer-Lemeshow

		P6 Realizan o No CE = 0		P6 Realizan o No CE = 1		Total
		Observado	Esperado	Observado	Esperado	
Paso 22	1	43	42,341	0	,659	43
	2	40	40,926	3	2,074	43
	3	38	39,246	5	3,754	43
	4	38	37,061	5	5,939	43
	5	37	33,886	6	9,114	43
	6	28	29,313	15	13,687	43
	7	26	25,190	17	17,810	43
	8	23	21,080	20	21,920	43
	9	8	14,350	35	28,650	43
	10	8	5,607	39	41,393	47

Tabla 12. Tabla de clasificación-Matriz de Confusión

Observado			Pronosticado		
			P6 REALIZAN YA CE		Porcentaje correcto
			0	1	
Paso 1	P6 Realizan o No CE	0	259	30	89,6
		1	42	103	71,0
	Porcentaje global				83,4
Paso 2	P6 Realizan o No CE	0	260	29	90,0
		1	41	104	71,7
	Porcentaje global				83,9
Paso 3	P6 Realizan o No CE	0	258	31	89,3
		1	42	103	71,0
	Porcentaje global				83,2
Paso 4	P6 Realizan o No CE	0	258	31	89,3
		1	43	102	70,3
	Porcentaje global				82,9
Paso 5	P6 Realizan o No CE	0	261	28	90,3
		1	43	102	70,3
	Porcentaje global				83,6
...
Paso 20	P6 Realizan o No CE	0	261	28	90,3
		1	53	92	63,4
	Porcentaje global				81,3
Paso 21	P6 Realizan o No CE	0	263	26	91,0
		1	55	90	62,1
	Porcentaje global				81,3
Paso 22	P6 Realizan o No CE	0	261	28	90,3
		1	62	83	57,2
	Porcentaje global				79,3

6. PRINCIPALES RESULTADOS

Presentamos a continuación un cuadro que recoge la información más relevante del análisis, que nos indica las variables incluidas en la ecuación con la estimación de los coeficientes, y el nivel crítico o de significación de las mismas (*Tabla 13a, b y c*). Las variables que finalmente sí se encuentran en la ecuación son:

- a) De las variables consideradas de información general sobre la empresa:
 - **ANTIGÜED** : Antigüedad de la empresa, variable de tipo cuantitativo.
 - **SECTORCO**: Sector Codificado, sector al que pertenece la empresa.
- b) De las variables extraídas de la pregunta 8 que cuestionaba los objetivos más importantes para la empresa:
 - **P8.1_AHO** : Posibilidad de Ahorro en Costes.
 - **P8.2_VEN** : Venta
 - **P8.6_ADA** : Adaptación de la Demanda y Mejor Segmentación.
 - **P8.7_VEN** : Ventaja Competitiva ante los competidores.
 - **P8.10_BU** : Buscar Información.
- c) De las variables extraídas de la pregunta 10 que valoraba los factores barrera de entrada para las empresas murcianas:
 - **P10.1_IN** : Inexperiencia en el ámbito de las Nuevas Tecnologías.
 - **P10.7_FA** : Falta de Desarrollo Tecnológico.
 - **P10.11_A** : Arraigo de la Cultura Empresarial reacia a los cambios.
- d) De las variables extraídas de la pregunta 11 que valoraba de nuevo los factores barrera pero en esta ocasión para que un consumidor se decida a realizar transacciones por la red:
 - **P11.1_DI** : Dificultad de acceso a Internet.
 - **P11.2_LE** : Lentitud en el acceso a las direcciones buscadas.
 - **P11.3_CO** : Coste añadido de la Tecnología.
 - **P11.9_NO** : Incertidumbre sobre la Veracidad de la Información.
 - **P11.10_I** : Incertidumbre sobre cómo se Realizará el Proceso.
- e) Más el término **Constante**.

En las *Tablas 13a, b y c* se incluyen los intervalos de confianza para cada odds ratio y que en la tabla aparecen como $Exp(\beta)$ para el 95% de confianza, lo que nos permite contrastar la hipótesis nula de que un *odds ratio* vale 1 para la población. Si el intervalo de confianza no incluye este valor, podemos rechazar dicha hipótesis, lo que implica que la variable independiente sí que tiene un efecto significativo en el análisis. También nos informa acerca de cada una de las variables y su importancia relativa. Las que tienen intervalos que se encuentran solapados tienen un efecto similar, mientras que las variables cuyos intervalos no coincidan con otros, significa que

tienen un efecto significativamente diferente en su aportación al modelo.

Para comprender mejor el significado de esta tabla, conviene indicar que la primera fila de cada variable categórica implicada en el modelo, en donde aparece solamente el nombre de la misma, por ejemplo *SECTORCO*, ofrece un contraste de la variable considerándola de forma global. Si éste no fuera significativo, o lo que es lo mismo, su nivel crítico fuese superior a un valor 0,05, no tendría ningún sentido el seguir intentando conocer lo que significan el resto de variables categorizadas en las que se descompone dicha variable.

Tabla 13a. Variables de la ecuación

	β	E.T.	Wald	gl	Sig.	Exp(β)	I.C. 95,0% para EXP(β)		
							Inferior	Superior	
Paso 22(a)	ANTIGÜED	0,024	0,010	6,553	1	0,010	1,025	1,006	1,044
	SECTORCO			10,212	2	0,006			
	SECTORCO(1)	1,757	0,568	9,548	1	0,002	5,793	1,901	17,651
	SECTORCO(2)	1,317	0,580	5,148	1	0,023	3,732	1,196	11,643
	P8.1_AHO			8,300	4	0,081			
	P8.1_AHO(1)	1,645	0,673	5,983	1	0,014	5,182	1,387	19,367
	P8.1_AHO(2)	-0,316	0,508	0,385	1	0,535	0,729	0,269	1,976
	P8.1_AHO(3)	0,111	0,447	0,062	1	0,804	1,118	0,465	2,685
	P8.1_AHO(4)	0,016	0,376	0,002	1	0,966	1,016	0,487	2,123
	P8.2_VEN			10,255	4	0,036			
	P8.2_VEN(1)	-1,580	0,895	3,114	1	0,078	0,206	0,036	1,191
	P8.2_VEN(2)	-1,285	0,579	4,921	1	0,027	0,277	0,089	0,861
	P8.2_VEN(3)	-0,475	0,435	1,196	1	0,274	0,622	0,265	1,457
	P8.2_VEN(4)	0,227	0,365	0,386	1	0,534	1,255	0,614	2,565
	P8.6_ADA			19,337	4	0,001			
	P8.6_ADA(1)	2,335	0,689	11,497	1	0,001	10,332	2,679	39,847
	P8.6_ADA(2)	2,158	0,603	12,816	1	0,000	8,654	2,655	28,205
	P8.6_ADA(3)	0,739	0,483	2,341	1	0,126	2,093	0,812	5,392
	P8.6_ADA(4)	1,127	0,458	6,054	1	0,014	3,087	1,258	7,576
	P8.7_VEN			33,551	4	0,000			
	P8.7_VEN(1)	-3,178	0,885	12,886	1	0,000	0,042	0,007	0,236
	P8.7_VEN(2)	-2,626	0,646	16,512	1	0,000	0,072	0,020	0,257
	P8.7_VEN(3)	-2,086	0,449	21,590	1	0,000	0,124	0,052	0,299
	P8.7_VEN(4)	-1,951	0,404	23,373	1	0,000	0,142	0,064	0,313

En el caso de la variable *Sector* de la Empresa *Codificado*, la significación es de 0,006 por lo que sí es interesante pasar al estudio de las dos variables categóricas que la sustituirán en la ecuación, y que son *SECTORCO(1)* y *SECTORCO(2)*.

El contraste que hemos elegido para las variables categóricas ha sido el sistema Indicador, por lo que los β estimados representan a cada categoría de referencia de la variable. Si el coeficiente de β es significativo, implica que la categoría a la que

representa es bastante diferente a la categoría de referencia elegida. Por ejemplo, para la variable P8.1_AHO, que mide la variable independiente del objetivo de la empresa como nuevo sistema para incrementar las ventas, podemos observar que las variables P8.2_VE(3) y P8.2_VE(4) tienen unos niveles de significación muy altos, 0,274 y 0,534 respectivamente, por lo que podemos asegurar que la probabilidad de que la venta sea un objetivo muy importante no difiere de la probabilidad de la categoría de referencia.

Tabla 13b. Variables de la ecuación

	β	E.T.	Wald	gl	Sig.	Exp(β)	I.C. 95,0% para EXP(β)	
							Inferior	Superior
P8.10_BU			14,749	4	0,005			
P8.10_BU(1)	-1,183	0,784	2,278	1	0,131	0,306	0,066	1,424
P8.10_BU(2)	-2,078	0,647	10,309	1	0,001	0,125	0,035	0,445
P8.10_BU(3)	-0,952	0,376	6,428	1	0,011	0,386	0,185	0,806
P8.10_BU(4)	-0,958	0,366	6,854	1	0,009	0,384	0,187	0,786
P10.1_IN			10,209	4	0,037			
P10.1_IN(1)	1,897	1,235	2,357	1	0,125	6,663	0,592	75,024
P10.1_IN(2)	-1,169	0,671	3,041	1	0,081	0,311	0,083	1,156
P10.1_IN(3)	-1,038	0,438	5,621	1	0,018	0,354	0,150	0,835
P10.1_IN(4)	-0,479	0,353	1,840	1	0,175	0,620	0,310	1,237
P10.7_FA			9,167	4	0,057			
P10.7_FA(1)	-2,620	1,560	2,821	1	0,093	0,073	0,003	1,549
P10.7_FA(2)	0,344	0,619	0,308	1	0,579	1,410	0,419	4,740
P10.7_FA(3)	0,159	0,438	0,132	1	0,716	1,172	0,497	2,765
P10.7_FA(4)	0,812	0,395	4,213	1	0,040	2,252	1,037	4,888
P10.11_A			7,683	4	0,104			
P10.11_A(1)	0,620	0,778	0,635	1	0,426	1,859	0,404	8,548
P10.11_A(2)	0,226	0,529	0,182	1	0,670	1,253	0,444	3,533
P10.11_A(3)	1,066	0,413	6,647	1	0,010	2,903	1,291	6,525
P10.11_A(4)	0,706	0,380	3,455	1	0,063	2,027	0,962	4,269
P11.1_DI			9,511	4	0,050			
P11.1_DI(1)	-0,776	0,779	0,993	1	0,319	0,460	0,100	2,118
P11.1_DI(2)	0,563	0,552	1,039	1	0,308	1,756	0,595	5,181
P11.1_DI(3)	-0,672	0,500	1,807	1	0,179	0,511	0,192	1,360
P11.1_DI(4)	-0,270	0,486	0,308	1	0,579	0,764	0,295	1,979

Además del nivel de significación, hemos de tener en cuenta el valor del *odds ratio* calculado en la columna *Exp(β)*. Los coeficientes positivos que sean significativos indican que la razón de ventas es mayor a uno. Si comprobamos que el *odds ratio* es menor que la unidad, podemos indicar que la probabilidad de que sí realice comercio electrónico es menor en esa variable que en la categoría de referencia. Por ejemplo, la variable categórica P11.1_DI(4), es decir aquella que mide en nuestro modelo a los sujetos que han dado una valoración de cuatro sobre cinco a la variable

que estudiaba la *Dificultad en el Acceso a Internet* como barrera para el consumidor, obtiene un nivel crítico de 0,579, mientras que su razón de ventajas es de 0,764, lo que implica que la probabilidad de que la variable dependiente tome el valor uno, es decir de que la empresa que considera esta barrera con un valor de cuatro, es menor que la que indica que esta barrera obtiene una importancia de cinco, observación que parece lógica, ya que si la empresa considera que esta barrera es bastante importante, y que por lo tanto va a afectar a sus clientes potenciales de forma negativa, va a estar menos dispuesta a lanzarse a negociar por la red.

Tabla 13c. Variables de la ecuación

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95,0% para EXP(B)	
							Inferior	Superior
P11.2_LE			7,952	4	0,093			
P11.2_LE(1)	2,094	0,931	5,059	1	0,024	8,121	1,309	50,379
P11.2_LE(2)	0,958	0,600	2,553	1	0,110	2,608	0,805	8,450
P11.2_LE(3)	0,263	0,488	0,290	1	0,590	1,300	0,500	3,383
P11.2_LE(4)	-0,063	0,480	0,017	1	0,895	0,939	0,366	2,405
P11.3_CO			7,727	4	0,102			
P11.3_CO(1)	-1,654	0,737	5,044	1	0,025	0,191	0,045	0,810
P11.3_CO(2)	0,031	0,543	0,003	1	0,954	1,032	0,356	2,993
P11.3_CO(3)	-0,235	0,457	0,264	1	0,607	0,790	0,322	1,938
P11.3_CO(4)	0,075	0,473	0,025	1	0,874	1,078	0,427	2,723
P11.9_NO			11,887	4	0,018			
P11.9_NO(1)	-2,458	1,654	2,209	1	0,137	0,086	0,003	2,189
P11.9_NO(2)	0,474	0,769	0,380	1	0,538	1,606	0,356	7,245
P11.9_NO(3)	0,720	0,487	2,192	1	0,139	2,055	0,792	5,334
P11.9_NO(4)	-0,634	0,381	2,777	1	0,096	0,530	0,252	1,118
P11.10_I			9,258	4	0,055			
P11.10_I(1)	1,652	1,182	1,951	1	0,162	5,216	0,514	52,947
P11.10_I(2)	1,411	0,692	4,154	1	0,042	4,100	1,056	15,925
P11.10_I(3)	0,334	0,501	0,446	1	0,504	1,397	0,524	3,729
P11.10_I(4)	1,016	0,415	5,998	1	0,014	2,763	1,225	6,233
Constante	-2,147	0,757	8,038	1	0,005	0,117		

La significación de cada coeficiente se evalúa a partir del *estadístico de Wald* que también aparece en esta tabla, y que nos permite contrastar la hipótesis nula de que el coeficiente vale cero en la población. Dicho estadístico se obtiene calculando el cociente entre el *valor estimado del coeficiente (B)* y el *error típico cometido (E.T.)* y elevando el resultado obtenido al cuadrado.

Cuando encontramos un coeficiente de las variables con signo negativo, esto implica que el incremento en cualquiera de ellas disminuye la probabilidad de que una empresa murciana realice comercio electrónico. Señalar además, que en este tipo de análisis, los coeficientes del modelo se expresan en las unidades originales de las variables independientes.

Una vez obtenidos los coeficientes y las variables que sí se incluyen en el modelo, podemos construir una ecuación lineal (PR) que transformándola a la escala logística nos permitirá obtener las probabilidades pronosticadas. Para facilitar la estimación de dicha probabilidad, hemos planteado el modelo en una hoja de cálculo de *Microsoft® Excel 2003* cuyo aspecto presentamos en la *Figura 2*, con el fin de que un empresario, o la Administración, introduciendo como datos iniciales los valores de las variables que se encuentran en el modelo, pueda realizar un pronóstico acerca de la probabilidad de que su empresa con esa percepción de las barreras y objetivos a alcanzar vaya a realizar comercio electrónico. Para que en las variables categóricas, como por ejemplo el *Sector* al que pertenece la empresa, sólo se introduzcan los datos codificados tal y como el modelo los admite hemos utilizado la herramienta de Validación de las celdas para impedir que se introduzcan datos extraños, pudiendo por lo tanto escogerlos de una lista desplegable en donde se presentan las opciones posibles a elegir.

Figura 2

Variables XI	B	codifi	B'XI
Constante	-2,14747423	1	-2,14747423
AMORRUP	0,024298308	10	0,24298308
SECTORCOI	1,756575764	1	1,756575764
SECTORCO2	1,317005102	0	0
PR.1_AHOR3	1,645253382	0	0
PR.1_AHOR2	-0,316543373	0	0
PR.1_AHOR1	0,111209282	0	0
PR.1_AHOR4	0,0162213	0	0
PR.2_VEN10	-1,500035901	0	0
PR.2_VEN9	-1,295046657	0	0
PR.2_VEN8	-0,474545811	0	0
PR.2_VEN4	0,226764629	0	0
PR.6_ADA1	2,335213456	1	2,335213456
PR.6_ADA2	2,160023313	0	0
PR.6_ADA3	0,738621464	1	0,738621464
PR.6_ADA4	1,127168491	0	0
PR.7_VEN10	-3,177646047	0	0
PR.7_VEN2	-2,620002551	0	0
PR.7_VEN3	-2,060003373	0	0
PR.7_VEN4	-1,951168416	0	0
PR.10_BUN1	-1,163418952	0	0
PR.10_BUN2	-2,077639748	0	0
PR.10_BUN3	-0,952197551	1	-0,952197551
PR.10_BUN4	-0,957900147	0	0
P10.1_BU1	1,896522902	0	0
P10.1_BU2	-1,16942143	0	0
P10.1_BU3	-1,037919416	0	0
P10.7_FAC1	0,476897554	0	0
P10.7_FAC2	-2,618895008	0	0
P10.7_FAC3	0,343661742	0	0
P10.7_FAC4	0,19001852	0	0
P10.7_FAC5	0,811668776	0	0
P10.11_AJ1	0,620380001	0	0
P10.11_AJ2	0,225690311	0	0
P10.11_AJ3	1,956572924	0	0
P10.11_AJ4	0,7064305	1	0,7064305
P11.1_EA1	-0,776193956	0	0
P11.1_EA2	0,662078962	1	0,662078962
P11.1_EA3	-0,671897242	0	0
P11.1_EA4	-0,203607725	0	0
P11.2_LE10	2,994499202	0	0
P11.2_LE2	0,95946610	1	0,95946610
P11.2_LE3	0,262715317	0	0
P11.2_LE4	-0,95306468	0	0
P11.3_CO1	-1,654270774	0	0

Así, a modo de ejemplo, podemos saber que la probabilidad de que una empresa murciana del sector servicios, que tenga 10 años de antigüedad realice comercio electrónico es del 77,43% (*Figura 3*) siempre que presente entre los objetivos a alcanzar

mediante el comercio electrónico la venta, la ventaja competitiva que adquiere frente a los competidores y el posible ahorro en costes con una valoración máxima, mientras que le resulte más o menos indiferente el objetivo de búsqueda de información y la adaptación de la demanda, que valore con un 5 en su empresa la barrera de falta de desarrollo tecnológico, y con un 4 el arraigo de la cultura empresarial reacia en cambios y la inexperiencia en el ámbito de las nuevas tecnologías, mientras que considere muy importantes las barreras que tienen los consumidores por el coste añadido de la tecnología y la incertidumbre que les genera la realización del proceso, creyendo que es un poco menor la incertidumbre en la veracidad de la información y valorando con un dos la dificultad de acceso y la lentitud de la red.

Si con las mismas opciones consideramos que la empresa pertenece al Sector Industrial comprobamos que su probabilidad desciende a un 68,86%, mientras que si pertenece al sector de la Construcción esta probabilidad se reduce aún más situándose en el 37,20%.

Se han realizado tiradas para comprobar que el modelo se ajusta a los datos, obteniendo de forma general que las empresas pertenecientes al sector servicios son las que efectivamente presentan la probabilidad más alta de realizar actividades de comercio electrónico en la Región de Murcia, seguidas de las del sector industrial y por último, el sector de la construcción.

Figura 3. Probabilidad de que una empresa realice Comercio Electrónico

Generales	Antigüedad de la empresa	10
	Sector	Servicios
Objetivos	Posible ahorro en costes	5
	Venta	5
	Adaptar demanda y mejor segmentación	3
	Ventaja competitiva ante competidores	5
	Buscar información	3
Barreras empresa	Inexperiencia ámbito NT	4
	Falta desarrollo tecnológico	5
	Arraigo cultura empresarial reacia a cambios	4
Barreras consumidor murciano	Dificultad de acceso a Internet	2
	Lentitud en el acceso	2
	Coste añadido de la tecnología	5
	Incertidumbre veracidad de la información	4
	Incertidumbre realización proceso	5
Probabilidad de que sí realice Comercio electrónico esta empresa:		77,43%

También obtenían probabilidades más elevadas las empresas de mayor antigüedad, y las que menos valoraban la falta de veracidad del proceso o de la información, y en general las barreras a las que se enfrenta el consumidor, frente a las

que consideraban con un alto grado de valoración las barreras propias que tienen que salvar. Indicar que para bastantes de las empresas que afirmaron no realizar políticas de comercio electrónico en la muestra, al calcular con sus datos la probabilidad de que sí pudieran realizar alguna actividad en torno a este tema, los resultados no eran excesivamente bajos cuando ponderan como objetivos importantes a alcanzar la *Venta* y el alcanzar una *Ventaja Competitiva* frente a los competidores, además de los ítems anteriormente mencionados.

7. CONCLUSIONES

Mediante esta investigación y utilizando la metodología de encuestas hemos obtenido datos reales de empresas de la Región de Murcia que nos indican cuál es su posición en temas de comercio electrónico. Esto a su vez nos ha facilitado establecer las variables que pueden ser consideradas más relevantes y que pueden ayudar a la implantación de esta forma de hacer negocios, o por el contrario, servir de obstáculo en el desarrollo de los mismos. Para ello, se han analizado las mismas mediante un modelo de regresión logística que indica la probabilidad con la que una empresa murciana pueda llegar a realizar acciones respecto al comercio electrónico, y que se han implementado en una herramienta informática, a la que tanto las organizaciones empresariales como la Administración Regional tienen fácil acceso, como es una aplicación de hoja de cálculo, lo que les permitirá realizar previsiones acerca de cómo la empresa debe de enfrentarse a la red.

Mediante el análisis previo realizado acerca del contexto en el que se desenvuelve la Región de Murcia y contrastándolo con los datos obtenidos con nuestro trabajo de campo, podemos afirmar que esta región arrastra una fuerte carencia tecnológica con respecto a otras regiones españolas, que le resulta muy complicado de solventar. Si bien es cierto que existe un crecimiento económico latente, éste no resulta suficiente para alcanzar un nivel óptimo en temas de infraestructuras y de desarrollo tecnológico¹⁶, que garantice que las empresas regionales puedan acceder a ellos con facilidad y por este motivo en nuestra investigación se ponen de manifiesto como barreras que se han de superar. Dichas carencias también se detectan a nivel de usuarios, ya que el acceso de los murcianos a Internet se sitúa en un nivel medio, comenzando a ser impulsado en el año 2003 por el “*Plan Región de MurciaSI*” llevado a cabo por la Administración Regional, ya que en años anteriores quedábamos relegados a los últimos lugares de la tabla. Pese a todo, cabe subrayar que Murcia destaca por ser una Comunidad muy dinámica en torno a estos temas, tanto a nivel nacional como europeo. Aunque la estructura del

¹⁶ Rodríguez Cohard, Juan Carlos y Bernal Jurado, Enrique “E-Commerce and territorial development in th Objective-1 spanish regions”.

tejido empresarial se conforme por PYMES principalmente, esto no es un obstáculo, ya que sectores como los de Turismo y Ocio, así como el Sector Servicios en general que tiene un peso muy relevante en esta Región, son también los que cuentan con un mayor potencial para su futuro desarrollo por la red. Además, las empresas murcianas pertenecientes al sector industrial, se caracterizan en su mayoría por su gran capacidad exportadora, y se encuentran en una disposición favorable para seguir o para comenzar a introducir parte de su negocio tradicional a través de los canales electrónicos.

De los datos obtenidos finalmente podemos considerar que para que el comercio electrónico sea una realidad para esta región en un futuro cercano, se deberían de mejorar las infraestructuras, facilitar la incorporación de la tecnología en las empresas, cuidar el entorno virtual para adecuarlo a las necesidades de los clientes potenciales, mejorar e incorporar sistemas de formación y asesoramiento de los agentes implicados, encontrar mecanismos que refuercen la seguridad desde todos los aspectos que ésta conlleva, y potenciar desde la Administración el uso de la red tanto en sus actividades de gestión administrativa para con los ciudadanos y empresas, como en sus procesos comerciales como organización.

BIBLIOGRAFÍA

- ABRAIRA SANTOS, V.; PÉREZ DE VARGAS LUQUE, A. "Métodos multivariantes en bioestadística". Ed. Centro de Estudios Ramón Areces. Madrid 1996.
- AMERICAN EXPRESS "II Barómetro de la Pyme Española, 2003"
- BERNAL GARCÍA, J.J. "Búsqueda de herramientas informáticas de productividad aplicables a los métodos cuantitativos, para la toma de decisiones en la empresa. Un ensayo de sistematización." Trabajo de Investigación para Cátedra de Universidad. Cartagena 2000.
- BERNAL GARCÍA, JJ; MARTÍNEZ MARÍA DOLORES, S Y SÁNCHEZ GARCÍA, JF. Proyecto "CEREMU-PYMES". Proyecto de Investigación sobre el desarrollo del negocio electrónico en las PYMES de la Región de Murcia. Instituto de Fomento de la Región de Murcia y Universidad Politécnica de Cartagena, 2004.
- CARM. "Economía de la Región de Murcia". www.carm.es/carm/region/economia/economia.jsp
- DACCACH, J.C.. "Factores de éxito para el eCommerce". <http://www.tiendasurbanas.com/ecommerce/3088.htm>
- DATA MINING INSTITUTE. "Introducción al análisis Multivariante". 1997-2004 www.estadistico.com/arts.html?20020527
- DIRCE 2003 y 2004-INE.

- DIRECTORIO DE EMPRESAS. Cámaras de Comercio e Instituto de Fomento <http://directorio.camaras.org> y www.infocarm.es
- DOCUMENTOS DE TRABAJO DE LA UNIVERSIDAD AUTÓNOMA DE QUERÉTARO. "Teoría de muestreo". www.uaq.mx/matemáticas/estadísticas/xu5.html
- ENCUESTA GLOBAL COMMERCENET año 2000 sobre Barreras e Inhibidores al Comercio Electrónico. <http://www.commercenet.org>
- ESPACIO PYME "Informe Creación de Empresas" Septiembre de 2001.
- GONDAR NORES, J.E. Data Mining Intitute SL "SPSS y Modelado Estadístico". www.estadistico.com/arts.html?20001113
- GREENE, W. H. "Análisis Econométrico". Ed. Prentice Hall. Madrid, 1998.
- GUJARATI, D. "Econometría". 2ª Edición. McGraw-Hill. 1981.
- HOSMER, D.W; LEMESHOW, S. Applied Logistic Regresión. Ed. John Wiley. Nueva York, 1989.
- INE Y AECE. "Encuesta sobre el uso de las Tecnologías de la Información y la Comunicación y Comercio Electrónico 2002". www.ine.es
- INE. "Encuesta sobre innovación Tecnológica en las empresas, 2003". www.ine.es
- INE. Anuario Estadístico 2003. www.ine.es
- INFORME DE TELEFÓNICA. La Sociedad de la Información en España. 2003.
- INFORME PENETRACIÓN REGIONAL DE LA NUEVA ECONOMÍA www.n-economia.com
- LECCIÓN SOBRE ANÁLISIS MULTIVARIANTE. Universidad de Zaragoza www.5campus.org/leccion/anamul
- LIZASOAIN, L; JOARISTI. L. Gestión y análisis de datos con SPSS versión 11. Universidad del País Vasco. Ed. Thomson-Paraninfo. Madrid, 2003.
- LÓPEZ, C. "Modelos de Negocio en Internet (I, II, III y IV)". <http://www.gestiopolis.com/canales/emprendedora/articulosn0%208/modelosdenegocioeninternet1.htm> (2, 3 y 4)
- MALDONADO, R., KHAINATA Y WEB DESIGNERS. "Algunas pautas para evaluar una página web" The Franchise MarketPlace, basado en los estudios de Survesite: Online Market Research and Web Site Evaluation. <http://www.franchipolis.com/doc/herra/h3.2.asp> www.surveysite.com
- MARTÍNEZ MARÍA DOLORES, S. Tesis Doctoral "El Comercio Electrónico: Análisis y Propuestas de Potenciación. Aplicación al Caso de la Región de Murcia." Director: Bernal García, J.J. Universidad Politécnica de Cartagena, 2004.
- MUÑOZ REYES, A. M, JIMÉNEZ GOMERO, M.D, MUÑOZ GARCÍA, J., PINO MEJÍAS, R. "e-Encuestas Probabilísticas II. Los métodos de Muestreo Probabilística" Estadística Española, Vol. 44, núm 151, 2002.

- MÚRIA ALBIOL, J. GIL SAURA, R. Preparación, Tabulación y Análisis de Encuestas para Directivos. ESIC Editorial, 1998.
- NEWBOLD, P. Estadística para los Negocios y la Economía. Ed. Prentice, 1997.
- PARDO MERINO, A Y RUIZ DÍAZ, M.A. SPSS 11. Guía para el análisis de datos. Ed. McGraw-Hill, 2002.
- PEÑA SÁNCHEZ DE RIVERA, D. Estadística. Modelos y Métodos. Volumen 2. Ed. Alianza. Madrid, 1987.
- PÉREZ, C. Técnicas Estadísticas con SPSS. Ed. Prentice Hall, Madrid, 2001.
- PULIDO SAN ROMAN, A. Y LÓPEZ, A. "Una propuesta metodológica para evaluar Regionalmente la Economía Digital" XXVII Reunión de Estudios Regionales. Cambios Regionales en la UE y Nuevos Retos Territoriales.
- PULIDO, A "El reto de la nueva información TIC" Revista Fuentes Estadísticas, Nº 51, Febrero de 2001. www.fuentesestadisticas.com/Numero51/paginas/4-5.htm
- PULIDO, A. Fichas n-economía "Impacto de la Economía Digital sobre las Respuestas de las Empresas a las Fuerzas del Mercado" "Guía Estratégica para empresas en el Dominio de la Información Digital" "Una nueva interpretación de la Ley de los Rendimientos" "Comprender la Nueva Economía y sus Efectos Macroeconómicos. Enfoques conexos con la Nueva Economía" www.n-economia.com
- RAPPA, M. "Business Models on the Web" . http://ecommerce.ncsu.edu/business_models.html
- RAYPORT, J.F. "The Truth about Internet Business Models" Harvard University <http://www.strategy-business.com/pdf/099301.pdf>
- RODRÍGUEZ COHARD, JUAN CARLOS Y BERNAL JURADO, ENRIQUE. "E-Commerce and territorial development in th Objective-1 spanish regions", European Regional Science Association in its series ERSa conference papers with number ersa02p326.
- RUIZ SOLER, M. "Metodología de encuestas". Curso del Doctorado de UCLA. http://campus-virtual.uma.es/campus/jur_soc/rsoler/material/UCLA2/sld057.htm
- SANTOS PEÑA, J., MUÑOZ ALAMILLOS, A. JUEZ MARTEL, P. Y GUZMÁN JUSTICIA, L. "Diseño y tratamiento estadístico de encuestas para estudios de mercado". Ed. Centro de Estudios Ramón Areces, SA. 1999.
- SIERRA Y ALVAREZ. "El método científico", La Habana, s.n. versión electrónica 1996, citado en El observador digital versión electrónica.
- TAKESHI, A. "Modelos de respuesta cualitativa: un examen" Cuadernos Económicos del ICE. Nº 39. 1988/2.
- TASCOPTT, DON. "La Era de los Negocios Electrónicos. Como generar utilidades en la economía digital" Ed. Mc Graw-Hill, 1999.
- DEALER WORLD. Septiembre, 2003 www.idg.es/computerworld/noticia.asp?id=32194
- DIARIO. "La Opinión de Murcia" www.laopiniondemurcia.es

