

ESTUDIO DE LA RELACIÓN ENTRE EL IMPUESTO PAGADO POR LOS
CONTRIBUYENTES DE INDUSTRIA Y COMERCIO DEL DISTRITO CAPITAL,
SUS INGRESOS Y SU ACTIVIDAD ECONÓMICA DURANTE EL PRIMER
BIMESTRE DE 2013

ALVAREZ BASTO JORGE TULIO
PALACIOS PALACIOS SMITH
TARQUINO BOJACA RICARDO

FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES
DEPARTAMENTO DE CIENCIAS BÁSICAS
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA
BOGOTÁ D.C.
2016

ESTUDIO DE LA RELACIÓN ENTRE EL IMPUESTO PAGADO POR LOS
CONTRIBUYENTES DE INDUSTRIA Y COMERCIO DEL DISTRITO CAPITAL,
SUS INGRESOS Y SU ACTIVIDAD ECONÓMICA DURANTE EL PRIMER
BIMESTRE DE 2013

ALVAREZ BASTO JORGE TULIO
PALACIOS PALACIOS SMITH
TARQUINO BOJACA RICARDO

Trabajo de grado para optar el título de:
Especialista en Estadística Aplicada

Director de trabajo de grado
LUIS ANTONIO HERRERA PÉREZ
M.Sc en Ciencias Estadística

FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES
DEPARTAMENTO DE CIENCIAS BÁSICAS
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA
BOGOTÁ D.C.
2016

Nota de aceptación

Firma del Presidente del jurado

Firma del jurado

Firma del jurado

Las directivas de la Fundación Universitaria Los Libertadores, los jurados calificadores y el cuerpo docente no son responsables de los criterios e ideas expuestas en el presente documento. Estos corresponden únicamente a los autores.

Con todo nuestro sincero y gran amor,
A nuestras esposas Rosmira, Yailuz y
Andrea.

A nuestros hijos, Kody, Kiara; Evan
Ghelyn, Milton Smith; y Camilo, a
nuestros Padres, a nuestros hermanos,
por su paciencia y comprensión,
porque prefirieron sacrificar su tiempo
para que nosotros pudiéramos cumplir
con nuestros sueños. Nos inspiraron a
ser mejores cada día, ahora podemos
decir que este trabajo de grado lleva
mucho de ustedes, gracias por estar
siempre a nuestro lado.

AGRADECIMIENTOS

Expresamos nuestros sinceros agradecimientos a nuestros padres, por habernos apoyado en todo momento, por sus consejos, por su ejemplo de perseverancia y constancia, por sus valores, por la motivación constante que nos han permitido ser personas de bien, a nuestros hijos por ser las personas que nos inspiran enfrentar todos los retos que encontramos día a día, a nuestros hermanos por ser nuestros ejemplos a seguir, a nuestras esposas por ser nuestras confidentes y amigas en nuestra vida diaria y a nuestros docentes de manera muy especial por hacer posible la comprensión de tantos conocimientos que sin su orientación no habría sido posible podernos apropiarnos de ellos, por su amistad, su paciencia, su dedicación y su entrega para alcanzar esta meta.

De igual manera agradecer a nuestro Director de trabajo de grado, M.Sc. Luis Antonio Herrera Pérez por su visión crítica de muchos aspectos cotidianos de la vida, por su rectitud en su profesión como docente, por sus consejos, que ayudan a formarnos como personas e investigadores.

También a nuestros jefes de trabajo, los cuales nos han motivado durante nuestra formación profesional.

Son muchas las personas que han formado parte de nuestra vida profesional a las que nos encantaría agradecerles su amistad, consejos, apoyo, ánimo y compañía en los momentos más difíciles de nuestras vidas. Algunas están aquí con nosotros y otras en nuestros recuerdos y en nuestros corazones, sin importar en donde estén queremos darles las gracias por formar parte de nosotros, por todo lo que nos han brindado y por todas sus bendiciones.

Para ellos: Muchas gracias y que Dios los bendiga.

CONTENIDO

	Pág.
1. INTRODUCCIÓN	14
2. OBJETIVOS	16
2.1 OBJETIVO GENERAL	16
2.2 OBJETIVOS ESPECIFICOS	16
3. MARCO DE REFERENCIA	17
3.1 ANTECEDENTES	17
3.2 MARCO CONCEPTUAL	17
3.2.1 Impuesto de industria, comercio, avisos y tableros (ICA)	17
3.2.2 Análisis de correlación	18
3.2.3 Análisis de regresión	18
3.2.4 Regresión múltiple con variables categóricas	19
4. MARCO METODOLÓGICO	21
4.1 TIPO DE ESTUDIO	21
4.2 POBLACIÓN	21
4.3 VARIABLES DE ESTUDIO	21
4.4 DISEÑO ESTADÍSTICO	21
4.5 PROCESAMIENTO Y ANÁLISIS DE LA INFORMACIÓN	22
5. RESULTADOS	23
5.1 ANÁLISIS ESTADÍSTICO DESCRIPTIVO	23
5.1.1 Variable impuesto pagado	23
5.1.2 Variable ingreso declarado	24
5.1.3 Tipo de actividad	24
5.1.4 Box plot variable impuesto por actividad	25
5.1.5 Box plot variable ingreso por actividad	26
5.1.6 Correlación entre las variables impuesto e ingreso	26
5.2 AJUSTE DE UN MODELO DE REGRESIÓN LINEAL MÚLTIPLE	27
5.2.1 Modelo 1	28
5.2.1.1 Validación de supuestos del modelo 1	29
5.2.2 Modelo 2	30
5.2.2.1 Validación de supuestos del modelo 2	31
5.2.3 Modelo 3	33
5.2.3.1 Validación de supuestos del modelo 3	35
6. CONCLUSIONES	37
7. RECOMENDACIONES	38
BIBLIOGRAFÍA	39

LISTA DE TABLAS

	Pág.
Tabla 1 Configuración dummy de variable categórica	19
Tabla 2 Descripción de variables de estudio	21
Tabla 3 Estadística descriptiva de la variable impuesto	23
Tabla 4 Estadística descriptiva de la variable ingreso	24
Tabla 5 Caracterización de los contribuyentes por actividad	24
Tabla 6 Coeficiente de correlación de Pearson entre las variables impuesto e ingreso	27
Tabla 7 Correlación de Pearson entre las variables impuesto e ingreso por actividad	27
Tabla 8 Resumen del modelo 1	28
Tabla 9 Coeficientes del modelo 1	28
Tabla 10 Estadísticos sobre los residuos modelo 1	29
Tabla 11 Prueba de normalidad modelo 1	30
Tabla 12 Resumen del modelo 2	31
Tabla 13 Coeficientes del modelo 2	31
Tabla 14 Estadísticos sobre los residuos modelo 2	32
Tabla 15 Prueba de normalidad modelo 2	32
Tabla 16 Resumen del modelo 3	34
Tabla 17 Coeficientes del modelo 3	34
Tabla 18 Comparación parámetros modelo 2 vs modelo 3	35
Tabla 19 Estadísticos sobre los residuos modelo 3	35
Tabla 20 Pruebas de normalidad modelo 3	36

LISTA DE GRÁFICAS

	Pág.
Gráfica 1 Cantidad y porcentaje de contribuyentes por actividad	25
Gráfica 2 Box plot de la variable impuesto	25
Gráfica 3 Box plot de la variable ingreso	26
Gráfica 4 Diagrama de dispersión impuesto vs ingreso	26
Gráfica 5 Dispersión pronostico tipificado vs residuo estudentizado modelo 1	30
Gráfica 6 Dispersión pronostico tipificado vs residuos estudentizado modelo 2	33
Gráfica 7 Dispersión pronostico tipificado vs residuos estudentizado modelo 3	36

GLOSARIO

ACTIVIDAD COMERCIAL: Es actividad comercial, la destinada al expendio, compraventa o distribución de bienes y mercancías, tanto al por mayor como al por menor y las demás actividades definidas como tales por el Código de Comercio, siempre y cuando no estén consideradas por la ley como actividades industriales o de servicios.

ACTIVIDAD DE SERVICIO: Es actividad de servicio, toda tarea, labor o trabajo ejecutado por persona natural o jurídica o por sociedad de hecho, sin que medie relación laboral con quien lo contrata, que genere una contraprestación en dinero o en especie y que se concrete en la obligación de hacer, sin importar que en ella predomine el factor material o intelectual.

ACTIVIDAD FINANCIERA: Son la actividades desarrolladas por los establecimientos de crédito, las sociedades de servicios financieros y las sociedades de capitalización, vigiladas por la Súper Intendencia Financiera.

ACTIVIDAD INDUSTRIAL: Es actividad industrial, la producción, extracción, fabricación, manufactura, confección, preparación, reparación, ensamblaje de cualquier clase de materiales y bienes y en general cualquier proceso de transformación por elemental que éste sea.

CORRELACIÓN: En probabilidad y estadística, la correlación indica la fuerza y la dirección de una relación lineal y proporcionalidad entre dos variables cuantitativas. Se considera que dos variables cuantitativas están correlacionadas cuando los valores de una de ellas varían sistemáticamente con respecto a los valores homónimos de la otra.

ERARIO: Conjunto de haberes, bienes y rentas pertenecientes al estado.

HECHO GENERADOR: El hecho generador del impuesto de industria y comercio está constituido por el ejercicio o realización directa o indirecta de cualquier actividad industrial, comercial o de servicios en la jurisdicción del Distrito Capital de Bogotá, ya sea que se cumplan de forma permanente u ocasional, en inmueble determinado, con establecimientos de comercio o sin ellos¹.

ICA: industria, comercio, avisos y tableros.

INGRESOS BRUTOS: Es el total de los ingresos obtenidos.

¹ (Estatuto tributario de Bogotá Decreto 352 de 2002)

INGRESOS NETOS: Son los ingresos después de descontar, las exportaciones, actividades exentas, rebajas y descuentos.

MÍNIMOS CUADRADOS: En estadística, los mínimos cuadrados ordinarios (MCO) o mínimos cuadrados lineales es el nombre de un método para encontrar los parámetros poblacionales en un modelo de regresión lineal. Este método minimiza la suma de las distancias verticales entre las respuestas observadas en la muestra y las respuestas del modelo.

MODELO ESTADÍSTICO: Es una expresión simbólica en forma de igualdad o ecuación que se emplea en todos los diseños experimentales y en la regresión para indicar los diferentes factores que modifican la variable de respuesta.

REGRESIÓN LINEAL: En estadística la regresión lineal o ajuste lineal es un método matemático que modela la relación entre una variable dependiente Y, las variables independientes X y un término aleatorio ϵ^2

REGRESIÓN SIMPLE: La regresión lineal simple es un método matemático que modela la relación entre una variable dependiente Y, una variable independiente X y un término aleatorio ϵ^3

REGRESIÓN MÚLTIPLE: La regresión lineal múltiple es un método matemático que modela la relación entre una variable dependiente Y, variables independientes X y un término aleatorio ϵ^4

² George C Canavos, Probabilidad y Estadística Métodos y Aplicaciones MC GRAW HILL 1998.

³ George C Canavos, Probabilidad y Estadística Métodos y Aplicaciones MC GRAW HILL 1998.

⁴ George C Canavos, Probabilidad y Estadística Métodos y Aplicaciones MC GRAW HILL 1998.

RESUMEN

Uno de los rubros importantes en el Distrito Capital para financiar El Plan de Desarrollo de Bogotá es el recaudo por medio de los impuestos, dentro de estos, se encuentra el impuesto de industria, comercio, avisos y tableros, que en adelante se denominara (ICA), que representa cerca del 50% de los recaudos anuales⁵, no se sabe con certeza como estos varían, en la medida que los ingresos de los contribuyentes aumenten o disminuyan, afectando positiva o negativamente los diferentes programas que se deben ejecutar en el Plan de Desarrollo.

El presente estudio aborda esta problemática y estudio de la relación entre el impuesto pagado por los contribuyentes de industria y comercio del Distrito Capital, sus ingresos y su actividad económica durante el primer bimestre de 2013.

Para esto se tomó la información de la totalidad de los contribuyentes de cada una de las cuatro actividades desarrolladas (industrial, comercial, servicios y financieros) que tributaron en el Distrito Capital en el periodo enero - febrero de 2013 se calculó y analizó el coeficiente de correlación lineal de Pearson para las variables Impuesto e Ingreso en conjunto, y para cada una de las actividades. Finalmente, se ajustó un modelo de regresión lineal múltiple para estudiar la relación que tienen las variables ingreso y actividad económica con la variable impuesto declarado.

Los resultados encontrados indican que el valor del coeficiente de correlación de Pearson, entre la variable impuesto e ingreso de manera conjunta es de 0.979 mostrando una relación fuerte y directamente proporcional entre estas. Igualmente, el modelo 2, sugerido por ser el más consistente de los tres, explica un 99,1% la relación de la variabilidad, de la variable dependiente impuesto a partir de la variable independiente ingreso, según la actividad desarrollada por el contribuyente.

Palabras Clave: Correlación, Regresión Lineal Múltiple, Variable Dummy

5

http://bibliotecadigital.ccb.org.co/bitstream/handle/11520/1298/3522_2008_2cartilla_distritales.pdf?sequence=1

ABSTRACT

One of the important items in the Capital District to finance The Development Plan of Bogotá is the collection through taxes, within these, is the industry and trade tax notices and boards (ICA), which represents about 50% Of the annual revenues, it is not known with certainty how these vary, to the extent that the income of the taxpayers increases or decreases, affecting positively or negatively the different programs that must be executed in the Development Plan.

The present study deals with this problem and study of the relationship between the tax paid by the taxpayers of industry and commerce of the Capital District, its income and its economic activity during the first two months of 2013.

For this, the information of all the taxpayers of each of the four activities (industrial, commercial, services and financial) that were taxed in the Capital District in the period January-February 2013 was calculated and analyzed. Linear correlation of Pearson for the variables Tax and Income as a whole, and for each one of the activities. Finally, a multiple linear regression model was fitted to study the relationship between income and economic activity variables with the declared tax variable.

The results show that the value of the Pearson correlation coefficient, between the tax and income variable, is 0.979, showing a strong and directly proportional relationship between them. Likewise, model 2, suggested to be the most consistent of the three, explains the relation of the variability, of the dependent variable imposed by the independent income variable, according to the activity developed by the taxpayer.

Key Words: Correlation, Multiple Linear Regression, Variable Dummy

1. INTRODUCCIÓN

En el Distrito Capital actualmente se encuentran tributando en forma bimestral aproximadamente 93.000 contribuyentes en el ICA, sin embargo, en muchas ocasiones no existe un soporte estadístico que permita establecer la relación existente entre los ingresos declarados y el impuesto pagado por estos contribuyentes.

Si bien existe información referente a los impuestos recibidos anualmente, no se conoce algún estudio que aborde el tema de la relación existente entre los ingresos, la actividad desarrollada y los impuestos pagados. Lo anterior teniendo en cuenta la dificultad que surge a la hora de contar con información confiable suministrada por los mismos contribuyentes, hecho que incentiva la evasión y la elusión al no tener herramientas de control que permitan determinar a partir de los ingresos obtenidos, el valor del impuesto pagado.

El desconocimiento de un modelo que permita observar como es la relación existente entre los ingresos, la actividad desarrollada y los impuestos pagados, no permite conocer con certeza como afectan a las finanzas del Distrito Capital y por ende al plan de desarrollo, los cambios presentados en el entorno económico por el aumento o la disminución de los ingresos de los contribuyentes.

El estudio y desarrollo del presente trabajo es de gran importancia debido a que la financiación del plan de desarrollo de Bogotá, se compone de los siguientes rubros; Transferencias (19%), Recursos de Capital (27%), No tributarios (6%) y Tributarios (48%). Del mismo modo el ICA representa el 50% de los Ingresos Tributarios. Esta razón establece la necesidad de determinar del ingreso total obtenido por los contribuyentes, el porcentaje que estos destinan al pago del impuesto de industria y comercio en el Distrito Capital, e igualmente conocer cómo se distribuyen estos ingresos en las cuatro actividades desarrolladas por los contribuyentes.⁶

Igualmente, se pretende que este documento sea de permanente consulta y que sirva como guía al realizar proyecciones sobre el cambio de los impuestos recibidos anualmente, cuando los ingresos recibidos por los contribuyentes en cada una de sus actividades, sufren algún tipo de variación.

Por esta razón, se desarrolla el presente trabajo con el fin de estudiar estadísticamente la relación entre el impuesto declarado y pagado, el ingreso neto gravable y el tipo de actividad económica de los contribuyentes en el Distrito Capital, específicamente aquellos registrados durante el primer bimestre de 2013.

⁶ <http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=5437>

Para cumplir con este propósito, se tomó la totalidad de los contribuyentes del ICA que tributaron en el Distrito Capital en el periodo enero - febrero de 2013.

2 .OBJETIVOS

Teniendo en cuenta los contribuyentes del distrito capital que tributaron con el ICA, durante el primer bimestre de 2013, se mencionan a continuación el Objetivo General y los Objetivos específicos del presente trabajo.

2.1 GENERAL

Estudiar la relación entre el impuesto pagado por los contribuyentes de ICA, su ingreso y su actividad económica desarrollada en el en el Distrito Capital en el periodo enero - febrero de 2013.

2.2 ESPECIFICOS

- Realizar una caracterización de los contribuyentes teniendo en cuenta el impuesto pagado, ingreso declarado y la actividad correspondiente.
- Analizar el grado de correlación entre el impuesto pagado y el ingreso declarado.
- Ajustar un modelo de regresión lineal múltiple que permita estudiar la relación entre el impuesto pagado, el ingreso declarado y la actividad correspondiente, a partir de las estimaciones de sus parámetros.

3. MARCO DE REFERENCIA

3.1 ANTECEDENTES

Algunos trabajos que se han desarrollado en torno al estudio del ICA son: Base gravable en el impuesto de industria y comercio (Gerencie.com Bogotá, 20, diciembre, 2010), Concepto N° 016461 Impuesto de industria y comercio (Ministerio de Hacienda y Crédito Público, 7 de mayo de 2014), Factores determinantes del recaudo de los impuestos Predial e Industria y Comercio en el Distrito de Cartagena de Indias, (Alean Pico, Ernesto Acosta, Rodolfo Enrique Matos Abril de 2009). Los principales resultados de estos trabajos se encaminaron en mostrar la necesidad de compilar y simplificar la normatividad existente para este impuesto y de esta manera aumentar el número de contribuyentes.

Por su parte Erick Céspedes y John Sepúlveda en el año (2008) en su estudio Relación entre el PIB de Bogotá y el impuesto de industria, comercio, avisos y tableros: un análisis VAR estructural⁷, se propusieron estimar las reformas tributarias que aumentaron significativamente el recaudo del impuesto de industria, comercio, avisos y tableros ICA en Bogotá a través de la metodología (VAR)⁸, para el periodo 1961-2004. A manera de resultado los autores encontraron que no todas las reformas del ICA han incrementado significativamente el recaudo de impuestos en Bogotá.

3.2 MARCO CONCEPTUAL

A continuación se presentan algunos conceptos fundamentales que facilitan la comprensión del presente trabajo.

3.2.1 Impuesto de industria, comercio, avisos y tableros ICA

El ICA es un gravamen, considerado como un impuesto directo, de carácter municipal que grava toda actividad industrial, comercial, financiera o de servicios que se realice en la jurisdicción del Distrito Capital de Bogotá.

La base gravable del ICA se liquida con base en los ingresos netos del contribuyente obtenidos durante el periodo a gravar. Para determinarlo se resta de la totalidad de los ingresos ordinarios y extraordinarios, los correspondientes a actividades exentas y no sujetas, así como las devoluciones, rebajas y descuentos, las exportaciones y la venta de activos fijos. El ICA, es un importante rubro de ingresos tributarios para Bogotá. Durante los últimos años ha representado en promedio, alrededor del 50% de los ingresos tributarios y por su estructura, recoge el dinamismo de la actividad

⁷ http://www.usta.edu.co/otras_pag/revistas/r_cife/Cife%2013/17.pdf

⁸ Valor en Riesgo (VAR por sus siglas en inglés) forma para medir el riesgo de mercado de un activo o portafolio

económica de la capital. Desde el punto de vista legal, en Colombia la Constitución Política en sus artículos 313, 362 y 363 desarrolla los principios en que se funda los impuestos; igualmente la autorización legal del impuesto de industria y comercio y su complementario el impuesto de avisos y tableros, autorizados por la Ley 97 de 1913, la Ley 14 de 1983 y los decretos ley 1333 de 1986 y 1421 de 1993⁹.

3.2.2 Análisis de correlación

El coeficiente de correlación de Pearson es una medida de la relación lineal existente entre dos variables X e Y, se calcula mediante la expresión que se muestra a continuación en la ecuación 1.

$$r = \frac{Cov(x,y)}{S_x S_y}, -1 \leq r \leq 1 \quad (1)$$

En ese orden en el análisis de correlación se estudia el grado de asociación o relación entre las variables, ya que la correlación es un coeficiente que permite evaluar la fuerza de asociación entre dos variables; para ello debe considerarse el valor del coeficiente cuyos límites son -1 a 1, una correlación que se acerque a 1 indica que el valor de Y aumenta a medida que aumenta el valor de X. En caso de que el valor se acerque a -1 indica que la correlación es negativa o inversa; esto es, el valor de Y disminuye a medida que aumenta el valor de X. Finalmente, si el valor es cercano a cero, esto indica que la relación entre X e Y produce una línea horizontal, y por tanto, no hay relación lineal entre las variables es decir, la variable dependiente Y permanece constante a medida que cambia X¹⁰.

3.2.3 Análisis de regresión

El análisis de regresión consiste en emplear métodos que permitan determinar la relación funcional entre dos o más variables y en ese sentido analizar la correspondencia entre las variables, estimar el efecto marginal que tiene una variable sobre la otra, predecir observaciones futuras, cuyos parámetros son estimados a partir del método de mínimos cuadrados ordinarios, para obtener la expresión de un modelo de regresión lineal simple de la forma:

$$y = \beta_0 + \beta_1 x + e \quad (2)$$

La ecuación 2 es la relación poblacional según la cual se hace regresión de Y sobre X además $\beta_0 + \beta_1 X$, es la proporción determinística de la relación, mientras que e representa el carácter aleatorio de la variable dependiente y por tanto denota el término de error en la expresión. Así mismo β_0 Intercepto y β_1 pendiente son

⁹ Secretaría Distrital de Hacienda. Estatuto Tributario de Bogotá (Decreto 352 de 2002)

¹⁰ George C Canavos, Probabilidad y Estadística Métodos y Aplicaciones MC GRAW HILL 1998.

parámetros que permanecen desconocidos y se pueden estimar con los datos obtenidos¹¹, esto puede ser expresado del siguiente modo.

$$y = b_0 + b_1x + e \quad (3)$$

Donde b_0 y b_1 son estimaciones de β_0 y β_1 respectivamente y e es el término aleatorio, de este modo para las predicciones el modelo de regresión toma la forma:

$$\hat{y} = b_0 + b_1x \quad (4)$$

En donde \hat{y} es el valor estimado de Y, b_0 es el intercepto o termino constante del modelo y b_1 es la pendiente o razón cambio de la variable independiente en la recta de regresión estimada; eso quiere decir que \hat{y} sería el valor a pagar en el impuesto según el modelo de regresión.

3.2.4 Regresión Múltiple con variables categóricas

En este caso se pretende modelar la relación entre una variable dependiente y un conjunto de variables independientes o predictoras, con la finalidad de poder pronosticar el valor de la variable dependiente, a partir del conocimiento del conjunto de variables predictoras. En la regresión lineal múltiple se modela la relación entre una variable dependiente y dos o más variables independientes mediante una función lineal; lo que significa que en regresión múltiple se asume que la variable Y se puede relacionar con las k variables X a través de un modelo estimado de la forma:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_kx_k + e \quad (5)$$

En donde \hat{y} es el valor estimado para la variable dependiente. b_k son valores estimados para los coeficientes poblacionales, por tanto b_i es la cantidad por las cual \hat{y} cambiará si x_k cambia en una unidad, asumiendo que las otras variables regresoras se mantiene constantes, teniendo en cuenta además que x_k son las variables regresoras y e es el término aleatorio asociado al error.

En este caso, la variable categórica se convierte en una variable dummy que toma la configuración mostrada en la siguiente tabla.

Actividad	X ₁	X ₂	X ₃	X ₄
X ₁	1	0	0	0
X ₂	0	1	0	0
X ₃	0	0	1	0
X ₄	0	0	0	1

Tabla1. Configuración dummy de variable categórica

¹¹ Regresión simple pág. 328/ Estadística aplicada a los negocios y la economía /Allen L webster tercera edición/ Mc Graw Hill ED

El uno y el cero indican la actividad a la que pertenece el contribuyente, así mismo se puede observar que no se obtiene solo un valor de β para la variable actividad, por el contrario se tiene un parámetro para cada categoría de la variable. β_1 ; β_2 ; β_3 ; β_4 , los cuales miden el efecto de cada sector sobre los impuestos. Es decir que aunque el modelo solo tenga una ecuación, se trata de un conjunto de 4 ecuaciones donde hay una para cada actividad, a continuación la ecuación muestra el modelo de regresión estimado para predecir en nuestro caso el valor del impuesto Y a pagar por un contribuyente, a partir de unos ingresos X y dependiendo la actividad desarrollada.

$$\text{impuesto pagado} = \beta_0 + \beta_1(\text{ingresos}) + \beta_2(\text{financiero}) + \beta_3(\text{industrial}) + \beta_4(\text{servicios}) + e \quad (6)$$

A partir del ajuste de la recta de regresión presentada anteriormente, será posible predecir el valor de un impuesto "y" que se debe pagar, a partir de un valor "x" de los ingresos declarados por un contribuyente en cada una de sus actividades; además, permite hacer una predicción del comportamiento de las variables estudiadas en un determinado punto o momento.

4. MARCO METODOLÓGICO

4.1 TIPO DE ESTUDIO

El presente trabajo es de tipo correlacional y explicativo, ya que estudia y analiza la relación entre el impuesto pagado, el ingreso declarado y la actividad económica del contribuyente.

4.2 POBLACIÓN

La población objeto de estudio corresponde a 93.033 contribuyentes del distrito capital que tributaron durante el segundo bimestre de 2013.

4.3 VARIABLES DE ESTUDIO

A continuación se presenta una breve descripción de las variables consideradas en el presente trabajo.

Nombre de variable	Abreviatura	Tipo	Unidad de medida
Ingreso declarado	Ingreso	Cuantitativa	Pesos colombianos
impuesto pagado	Impuesto	Cuantitativa	Pesos colombianos
Actividad del contribuyente	Actividad	Cualitativa	Categorías: Comercial, financiero, industrial, servicios

Tabla 2. Descripción de variables de estudio

Cabe precisar que la recolección de los datos se obtiene mediante el aplicativo Sistema de Información Tributaria, desarrollado en Oracle, el cual contiene la información de los valores pagados por los contribuyentes y sus ingresos.

4.4 DISEÑO ESTADÍSTICO

Para dar alcance a los objetivos planteados en el presente trabajo de grado, se realiza inicialmente un análisis descriptivo de cada una de las tres variables en estudio. Posteriormente se presentan gráficos de caja de la variable Impuesto e Ingreso para cada una de las actividades, un análisis de correlación entre las variables impuesto pagado e ingreso declarado en conjunto y luego para cada actividad. Finalmente, se realiza el ajuste de un modelo de regresión lineal múltiple para estudiar la relación entre el impuesto pagado y el ingreso declarado para cada actividad.

4.5 PROCESAMIENTO Y ANÁLISIS DE LA INFORMACIÓN

El almacenamiento, procesamiento y análisis de información se realizó haciendo uso de los software Excel 2010 y los programas estadísticos R V 3.1.1 y SPSS V 22, SAS ('Local', X64_VSPRO).

5. RESULTADOS

A continuación se presentan los resultados obtenidos con el desarrollo del presente trabajo. De manera específica se presenta inicialmente una descripción de la población de estudio de acuerdo a las variables impuesto, ingreso y actividad.

Posteriormente y con el ánimo de estudiar la relación entre las variables mencionadas anteriormente, se realiza y presenta un análisis de correlación y el ajuste de un modelo de regresión lineal múltiple.

5.1. ANÁLISIS ESTADÍSTICO DESCRIPTIVO

El presente análisis se hace teniendo en cuenta las variables impuesto, ingreso y actividad.

5.1.1 Variable impuesto pagado

En la tabla 3, se presenta una caracterización de la población de estudio, teniendo en cuenta la variable Impuesto.

	IMPUESTO PAGADO
Media	3.246.966
Error típico	164.673
Mediana	241.000
Moda	1.000
Desviación estándar	50.227.403
Varianza de la muestra	2.522.791.976.877.300
Curtosis	4.406
Coficiente de asimetría	58
Rango	5.418.620.000
Mínimo	1.000
Máximo	5.418.621.000
Suma	302.074.993.000
Cuenta	93.033

Tabla 3. Estadística descriptiva de la variable impuesto

De acuerdo a la información de la tabla 3, se puede evidenciar que el valor del impuesto pagado por los contribuyente, está en un rango que varía desde \$1.000 hasta \$5.418.621.000 con un promedio de \$3.246.966 y una desviación de \$50.227.403, la mediana nos indica que un 50% de los valores está por debajo de \$241.000 y el otro 50% por encima de ese valor.

5.1.2 Variable ingreso declarado

	INGRESO DECLARADO
Media	313.684.716
Error típico	15.316.152
Mediana	24.657.000
Moda	10.000.000
Desviación estándar	4.671.627.107
Varianza de la muestra	21.824.099.824.658.900.000
Curtosis	4.449
Coficiente de asimetría	58
Rango	530.835.979.000
Mínimo	67.000
Máximo	530.836.046.000
Suma	29.183.030.171.000
Cuenta	93.033

Tabla 4. Estadística descriptiva de la variable ingreso

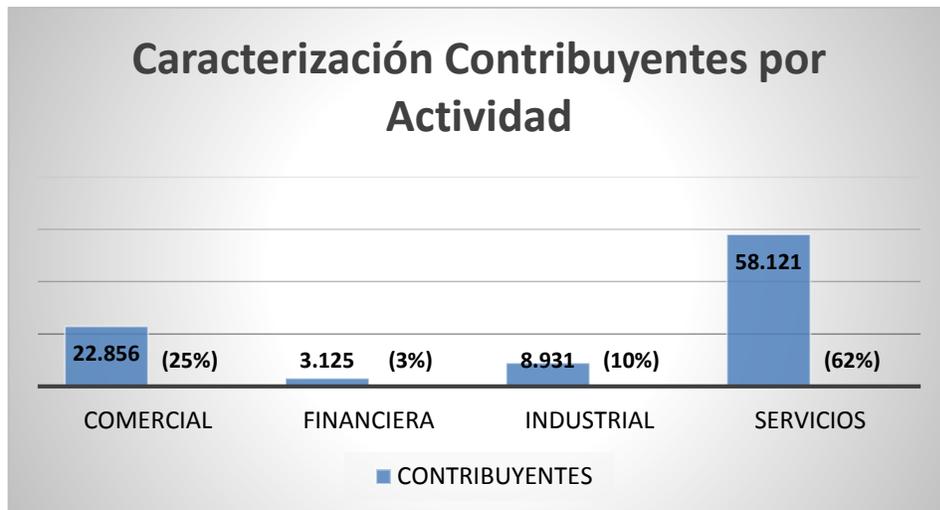
Como se puede observar en tabla 4, los valores del ingreso percibido por los contribuyentes, se encuentran en un intervalo de valores que varía desde \$ 67.000 hasta \$530.836.046.000 con un promedio de \$313.684.716 pesos colombianos y una desviación de \$4.671.627.107 pesos colombianos, igualmente se puede observar que un 50% de los contribuyentes obtiene un ingreso por un valor superior a \$ 24.657.000 y que el otro 50% un ingreso inferior a este valor.

5.1.3 Tipo de actividad

A continuación se realiza una descripción de la cantidad y el porcentaje de los contribuyentes por cada actividad.

ACTIVIDAD	CANTIDAD	%
Servicios	58.121	62%
Comercial	22.856	25%
Industrial	8.931	10%
Financiero	3.125	3%
Total	93.033	100%

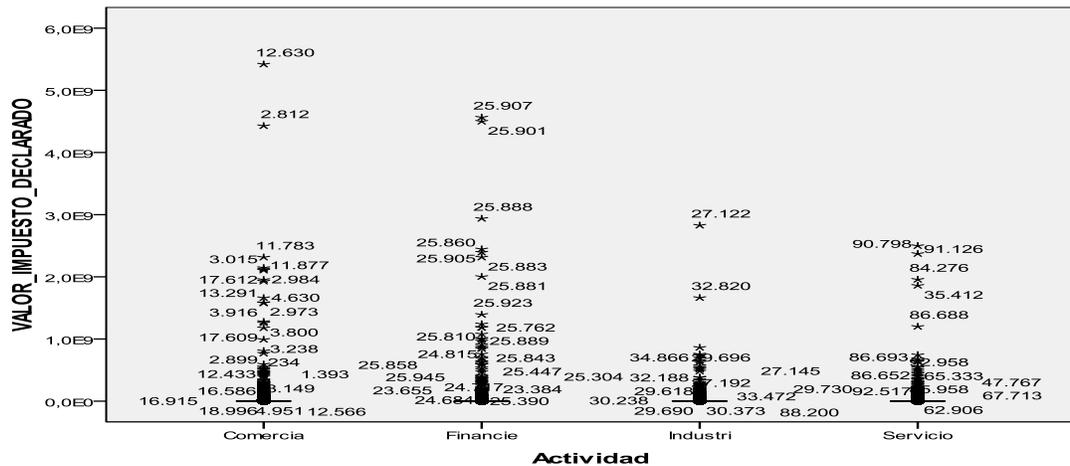
Tabla 5. Caracterización de los contribuyentes por actividad



Gráfica 1. Cantidad y porcentaje de contribuyentes por actividad

De acuerdo a la gráfica 1, se observa que la actividad que presenta el mayor número de contribuyentes en el distrito capital es la de servicios, con 58.121 lo que representa un porcentaje de participación del 62%, seguida de la actividad comercial con 22.856 para un 25%. Por último, la actividad industrial con 8.931 y financiera con 3.125 son las que menos contribuyentes presentan, estas actividades tienen participaciones del 10% y del 3% respectivamente en el total de contribuyentes.

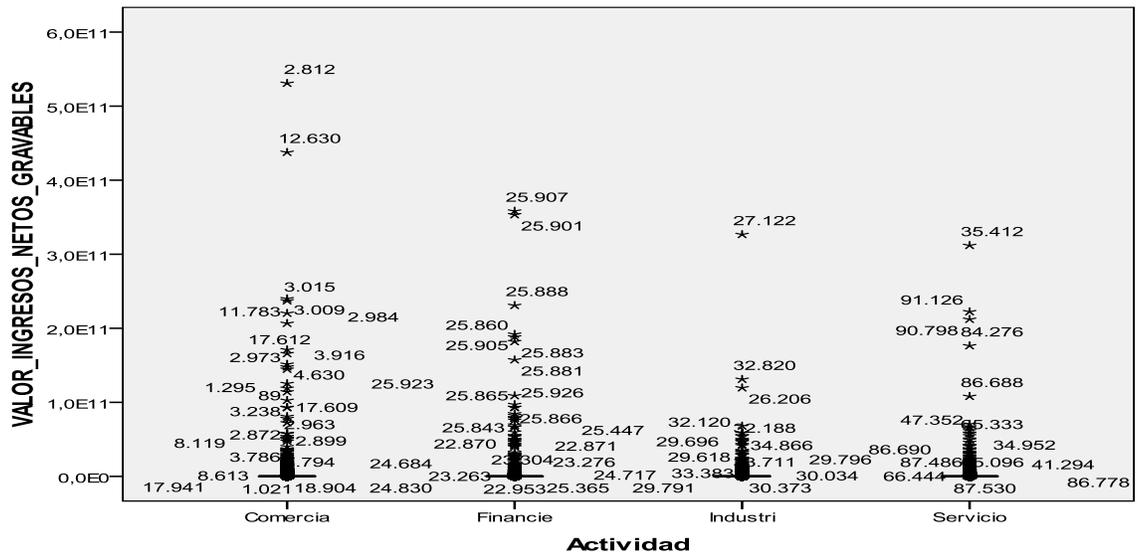
5.1.4 Box plot variable impuesto por actividad



Gráfica 2. Box plot de la variable impuesto

La gráfica 2, muestra el comportamiento de la distribución de los datos en la variable impuesto, identificando inicialmente en todas las actividades datos atípicos, además de la dispersión de algunos registros, evidenciándose más en las actividades comercial y financiera.

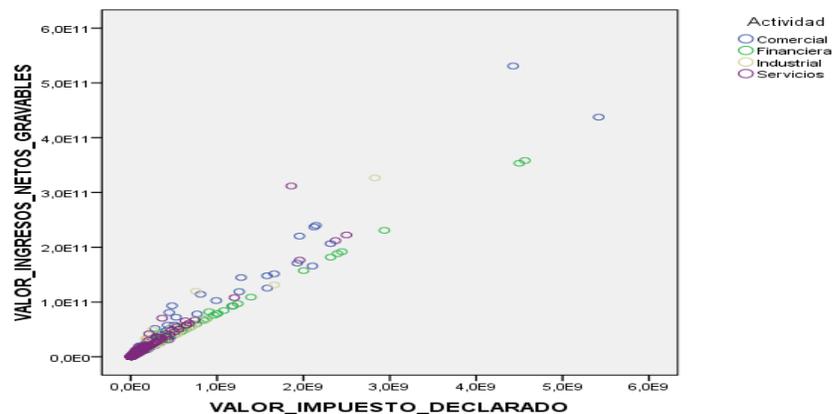
5.1.5 Box plot variable ingreso por actividad



Gráfica 3. Box plot de la variable ingreso

La gráfica 3, muestra el comportamiento de la distribución en los datos de la variable ingreso, las actividades que presentan mayor dispersión, son la comercial y la financiera.

5.1.6 Correlación entre las variables Impuesto e ingreso



Gráfica 4. Diagrama de dispersión impuesto vs ingreso

El diagrama de dispersión muestra una tendencia lineal y una posible correlación positiva entre las variables impuesto e ingreso y de esta forma, el valor impuesto declarado aumentaría proporcionalmente cuando exista un incremento en el valor

de la variable ingreso. Lo anterior se puede apreciar en conjunto, y realizando el análisis de manera separada para cada actividad.

Para medir cuantitativamente el grado de relación de las variables utilizamos el coeficiente de correlación de Pearson.

	IMPUESTO	INGRESO
IMPUESTO	1	0.979
INGRESO	0.979	1

Tabla 6. Coeficiente de correlación de Pearson entre las variables impuesto e ingreso

De acuerdo con los datos de la tabla 6, se puede observar que el coeficiente de correlación lineal de Pearson entre las variables impuesto e ingreso es de 0.979. Lo anterior indica de manera descriptiva, que existe una correlación lineal fuerte y directamente proporcional entre las variables impuesto e ingreso. Así a medida que aumenta el ingreso, genera un aumento en el impuesto.

Posteriormente, se procedió a calcular el coeficiente de correlación de Pearson entre las variables impuesto e ingreso por actividad. Los resultados se muestran en la siguiente tabla.

Actividad	Coeficiente de correlación lineal de Pearson entre las variables impuesto e ingreso
Comercial	0,98236246
Financiero	0,99926602
Industrial	0,97684774
Servicios	0,96608018

Tabla 7. Correlación de Pearson entre las variables impuesto e ingreso por actividad

En la tabla 7, se puede observar que el coeficiente de correlación lineal de Pearson entre las variables impuesto e ingreso por actividad es mayor a 0.96, en todos los casos. Lo anterior indica de manera descriptiva, que existe una correlación lineal fuerte y directamente proporcional entre las variables impuesto e ingreso. Así a medida que aumenta el ingreso, genera un aumento en el recaudo del impuesto.

5.2 AJUSTE DE UN MODELO DE REGRESIÓN LINEAL MÚLTIPLE

A continuación se presentan los resultados obtenidos con el análisis de regresión lineal múltiple aplicado a las variables en estudio.

Específicamente se presenta el procedimiento realizado para la selección de un modelo adecuado, que permita estudiar la relación entre el impuesto pagado (variable dependiente), el ingreso y la actividad económica de los contribuyentes (variables independientes).

5.2.1 MODELO 1

Inicialmente se ajusta un primer modelo (modelo 1), escogiendo como variable dependiente impuesto y como variables independientes ingresos y actividad. Los resultados obtenidos con el ajuste de este modelo se presentan a continuación:

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
1	,979 ^a	,958	,958	1,034E7	1,908

Tabla 8. Resumen del modelo 1

De acuerdo con la tabla 8, el modelo ajustado explicaría un 95,8% de la variabilidad del impuesto a partir del ingreso y la actividad.

Modelo		Coeficientes no estandarizados	
		B	Error típ.
1	(Constante)	-96602,862	68478,140
	VALOR_INGRESOS_NETO S_GRAVABLES	,011	,000
	financiero	2779530,556	197340,702
	industrial	74195,309	129030,773
	servicios	-88738,720	80757,614

Tabla 9. Coeficientes del modelo 1

Teniendo en cuenta los resultados presentados en la tabla 9, se tiene que el modelo 1 ajustado estaría dado por:

Valor impuesto declarado = $-96.602,862 + 0.011(\text{ingreso}) + 2.779.530,556$ (financiero) + $74.195,309$ (industrial) – $88.738,720$ (servicios).

Para el sector comercial el modelo es $Y = -96.602,862 + 0.011(\text{ingreso})$, lo que implica que por el aumento de un peso en los ingresos, el pago del impuesto aumentaría en 0,11 pesos; así mismo, si el contribuyente pertenece a la actividad financiera el pago del impuesto aumentaría en promedio 2.779.530,556 pesos; si pertenece a la actividad industrial el pago promedio del impuesto aumentaría en 74.195,309 pesos y si pertenece al sector de servicios el pago promedio del impuesto disminuye en -88.738,720 pesos.

5.2.1 .1 Validación de supuestos del Modelo 1

A continuación se muestran los resultados obtenidos, con el propósito de validar el modelo 1, haciendo uso de gráficos y de las diferentes pruebas en cada caso.

Independencia. El estadístico DW es un valor comprendido entre 0 y 4, para valores de DW “próximos” a 2 no se rechaza la hipótesis nula H_0 que establece que los residuales no son independientes¹². Por el contrario, para valores de DW “alejados” de 2, sí se rechaza H_0 ; de esta forma, se asume independencia para valores entre 1,5 y 2,5, en la tabla 8 resumen del modelo uno, se puede observar un coeficiente Durbin Watson de 1,908, con lo cual se puede asumir la independencia de los residuos en el modelo 1.

	Mínimo	Máximo	Media	Desviación típica	N
Valor pronosticado	-184637,05	5,58E9	3246966,06	4,915E7	93033
Valor pronosticado tip.	-,070	113,498	,000	1,000	93033
Error típico de valor pronosticado	42889,512	3854956,750	61370,529	44493,981	93033
Valor pronosticado corregido	-184640,25	5,77E9	3247766,84	4,928E7	93033
Residual	-1,419E9	8,175E8	,000	1,034E7	93033
Residuo típ.	-137,243	79,059	,000	1,000	93033
Residuo estud.	-140,660	83,079	,000	1,026	93033
Residuo eliminado	-1,491E9	9,027E8	-800,787	1,090E7	93033
Residuo eliminado estud.	-158,524	86,343	,000	1,078	93033
Dist. de Mahalanobis	,601	12930,118	4,000	67,112	93033
Distancia de Cook	,000	467,413	,012	1,780	93033
Valor de influencia centrado	,000	,139	,000	,001	93033

Tabla 10. Estadísticos sobre los residuos modelo 1

De acuerdo con la tabla 10, los residuales del modelo presentan un valor mínimo de -1,419E9 así mismo un valor máximo de 8,175E8; un promedio 0,000, y una desviación de 1,034E7.

Normalidad. Para validar este supuesto se realizó la prueba Kolmogorov – Smirnov, de hipótesis de normalidad, estos resultados se muestran a continuación.

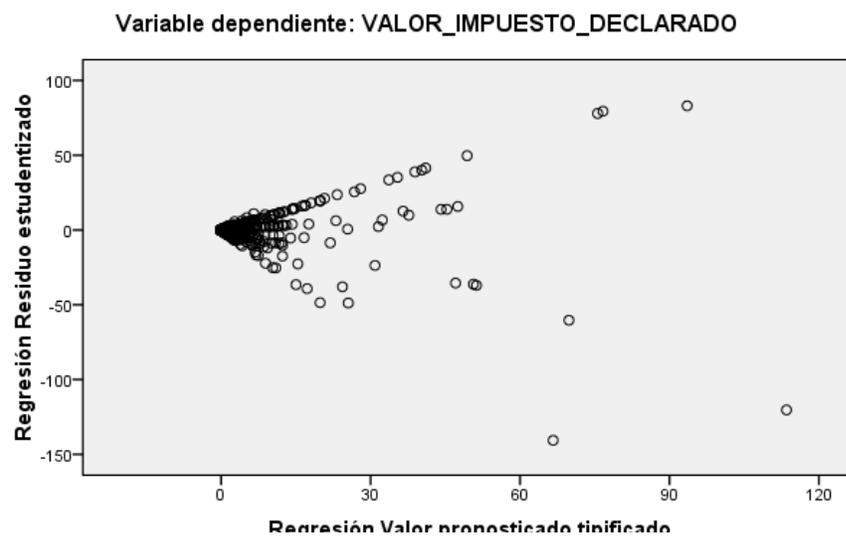
¹² George C Canavos, Probabilidad y Estadística Métodos y Aplicaciones MC GRAW HILL 1998.

Kolmogorov-Smirnov ^a		
Estadístico	gl	Sig.
,441	93033	,000

Tabla 11. Prueba de normalidad modelo 1

En la tabla 11, se evidencia con un nivel de significancia del 5% los residuales no se distribuyen normalmente.

Homocedasticidad. Para observar si los residuales tienen un comportamiento homocedástico, se realiza un gráfico de dispersión entre los valores pronosticados tipificados con los residuos estudentizados. Los resultados correspondientes a los residuales del modelo 1 se presentan en la gráfica 5.



Gráfica 5. Dispersión pronostico tipificado vs residuo estudentizado modelo 1

La gráfica 5 muestra un comportamiento de los residuales en forma de embudo, permitiendo concluir que los residuales no tienen varianza constante.

Luego de realizar el análisis del modelo 1, y debido al problema que se presenta con la homogeneidad de varianza y el supuesto de normalidad, se procedió a realizar una transformación de las variables impuesto e ingreso, originando así un modelo 2.

5.2.2 MODELO 2

En este caso y con el ánimo de solventar el problema de homogeneidad de varianza y normalidad de los errores que presenta el modelo 1, se ajustó un nuevo modelo 2

teniendo en cuenta como variable dependiente el logaritmo natural del impuesto, y como variables independientes el logaritmo natural del ingreso y las actividades; financiera, industrial, servicios y comercial .

Los resultados obtenidos se presentan a continuación:

Modelo 2	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
	,995 ^a	,991	,991	,17691	,003

Tabla 12. Resumen del modelo 2

De acuerdo con la tabla 12, el modelo ajustado explica un 99,1% de la variabilidad del logaritmo natural del impuesto a partir del logaritmo natural del ingreso, según la actividad desarrollada por el contribuyente.

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	-4,422	,006		-777,046	,000
	LnIngreso	,996	,000	,988	3119,873	,000
	financiero	-,038	,003	-,004	-11,273	,000
	industrial	-,026	,002	-,004	-11,791	,000
	servicios	-,191	,001	-,050	-137,551	,000

Tabla 13. Coeficientes del modelo 2

Teniendo en cuenta los resultados presentados en la tabla 13, se tiene que el modelo 2 ajustado estaría dado por:

$\text{Ln (Valor impuesto declarado)} = -4,422 + 0,996 (\text{Ln ingresos}) - 0,038 (\text{financiero}) - 0,026(\text{industrial}) - 0,191(\text{servicios}).$

Para el sector comercial el modelo es, $Y = -4,442 + 0.996 \text{ Ln (ingreso)}$, esto implica que por el incremento de un peso en el logaritmo natural del ingreso, el logaritmo natural del impuesto aumentaría en 0,996 pesos; además con un nivel de ingreso igual en cada una de las actividades desarrolladas, los contribuyentes que ejercen la actividad de servicios son los que pagan en promedio el menor valor en el impuesto con una disminución de -0,191 pesos respecto de los que ejercen la actividad comercial; lo sigue los contribuyentes que ejercen la actividad industrial con una disminución en promedio de -0,026 pesos; igualmente los contribuyentes

que ejercen la actividad financiera pagan en promedio en el impuesto -0,038 pesos respecto a la actividad comercial.

5.2.2.1 Validación de supuestos del Modelo 2

Independencia. La tabla 12 resumen del modelo 2, muestra un coeficiente Durbin Watson de 0,003 relativamente lejano a 2 y evitando el supuesto de independencia. Sin embargo, vale la pena recordar que el presente estudio es de carácter descriptivo y poblacional.

	Mínimo	Máximo	Media	Desviación típica	N
Valor pronosticado	6,4492	22,4545	12,4961	1,83671	93033
Valor pronosticado tip.	-3,292	5,422	,000	1,000	93033
Error típico de valor pronosticado	,001	,004	,001	,001	93033
Valor pronosticado corregido	6,4491	22,4546	12,4961	1,83671	93033
Residual	-,82144	,51168	,00000	,17691	93033
Residuo típ.	-4,643	2,892	,000	1,000	93033
Residuo estud.	-4,643	2,892	,000	1,000	93033
Residuo eliminado	-,82155	,51173	,00000	,17692	93033
Residuo eliminado estud.	-4,644	2,893	,000	1,000	93033
Dist. de Mahalanobis	,601	56,378	4,000	5,861	93033
Distancia de Cook	,000	,001	,000	,000	93033
Valor de influencia centrado	,000	,001	,000	,000	93033

Tabla 14 Estadísticos sobre los residuos modelo 2

De acuerdo con la tabla 14, los residuales del modelo 2 presentan un valor mínimo de -0,82144; el máximo es de 0,51168; una media de 0,000000 y una desviación de 0,17691.

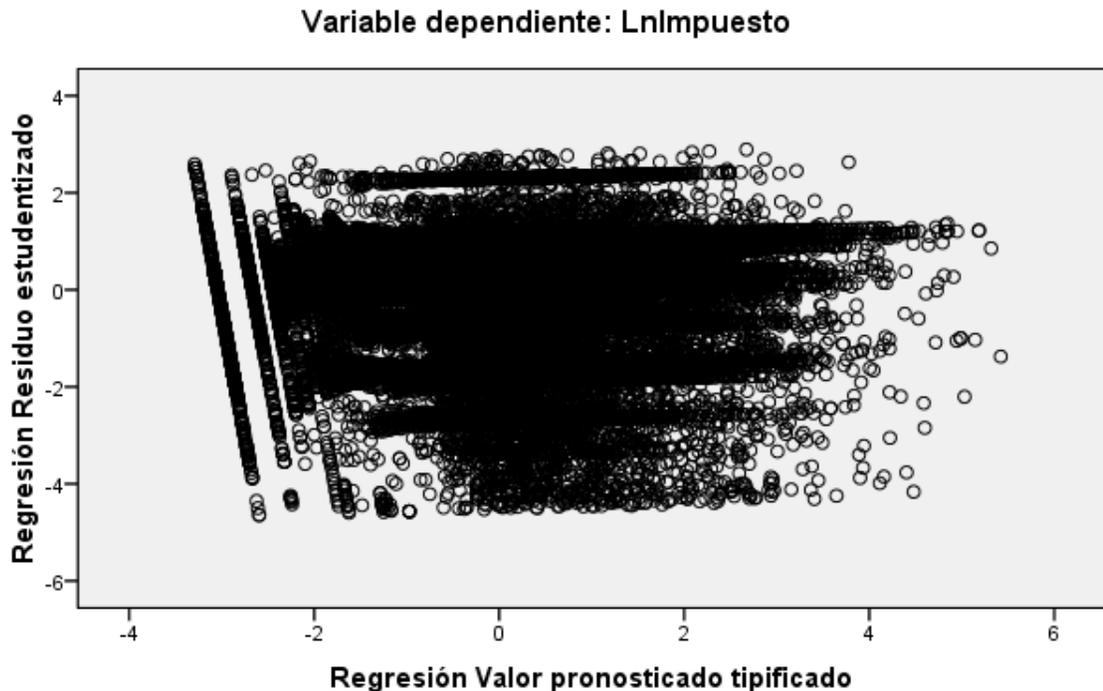
Normalidad. Para validar este supuesto se realizó prueba de Kolmogorov-Smirnov con la cual se verifica la hipótesis sobre el cumplimiento del supuesto de normalidad de los residuales. Los resultados relacionados se muestran en la siguiente tabla.

	Kolmogorov-Smirnov ^a		
	Estadístico	gl	Sig.
Unstandardized Residual	,219	93033	,000

Tabla 15 prueba de normalidad modelo 2

En la tabla 15, se evidencia con un nivel de significancia del 5% los residuales no se distribuyen normalmente.

Homocedasticidad. Al igual que en el caso anterior se realiza un gráfico de dispersión entre los valores pronosticados con los residuos tipificados. Los resultados correspondientes a los residuales del modelo 2, se presentan en la gráfica 6.



Gráfica 6. Dispersión pronostico tipificado vs residuos estudentizado modelo 2

En la gráfica 6, se evidencia una mejoría en el comportamiento de la homogeneidad de la varianza de los residuales respecto al modelo 1.

No obstante, y con el ánimo de analizar la posible influencia de datos atípicos en los resultados del modelo 2, incluyendo algún cambio sustancial en las estimaciones de sus parámetros, se ajustó un modelo 3 sustrayendo del análisis aquellos individuos cuyos residuales fueran mayores que 3 y menores que -3.

5.2.3 MODELO 3

En este caso y para el ajuste de este modelo, se utilizaron las mismas variables del modelo 2, pero ahora eliminando aquellos individuos cuyos residuales estudentizados, eran mayores que 3 y menores que -3. Los resultados obtenidos para este modelo se presentan a continuación:

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
1	,996 ^a	,992	,992	,16892	,010

Tabla 16. Resumen del modelo 3

De acuerdo con la tabla 16, el modelo ajustado explicaría un 99,2% de la variabilidad del impuesto a partir del ingreso y la actividad desarrollada por el contribuyente.

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	-4,442	,006		-783,406	,000
	Ln ingreso	,997	,000	,989	3135,842	,000
	financiero	-,042	,003	-,004	-12,610	,000
	industrial	-,022	,002	-,004	-10,125	,000
	servicios	-,184	,001	-,049	-133,169	,000

Tabla 17. Coeficientes del modelo 3

Teniendo en cuenta los resultados presentados en la tabla 17, se tiene que el modelo 3 ajustado estaría dado por:

$$\text{Ln impuesto} = - 4.442 + 0,997 \text{ Ln (ingresos)} - 0,042 \text{ (financiero)} - 0,022 \text{ (industrial)} - 0,184 \text{ (servicios)}$$

De acuerdo al modelo 3, Para el sector comercial el modelo es, $Y = -4,442 + 0.997 \text{ Ln (ingreso)}$, esto implica que por el incremento de un peso en el logaritmo natural del ingreso, el logaritmo natural del impuesto aumentaría en 0,997 pesos; además con un nivel de ingreso igual en cada una de las actividades desarrolladas, los contribuyentes que ejercen la actividad de servicios son los que pagan en promedio el menor valor en el impuesto con una disminución de -0,184 pesos respecto de los que ejercen la actividad comercial; lo sigue los contribuyentes que ejercen la actividad industrial con una disminución en promedio de -0,022 pesos; igualmente los contribuyentes que ejercen la actividad financiera pagan en promedio en el impuesto -0,042 pesos respecto a la actividad comercial.

Con el fin de analizar los cambios que presentan los modelos 2 y 3 en las estimaciones de sus parámetros, se construye y presenta a continuación la siguiente tabla:

Parámetros	Modelo 2	Modelo 3
Constante	-4,442	-4.442
Ln ingreso	0,996	0,997
Financiero	-0 ,038	-0,042
Industrial	-0,026	-0,022
Servicio	-0,191	-0,184

Tabla 18. Comparación parámetros modelo 2 vs modelo 3

Como se puede ver en la tabla 18, no se presentan cambios sustanciales en las estimaciones de los parámetros.

5.2.3.1 Validación de Supuestos del Modelo 3

Independencia. La tabla 16, resumen del modelo 3 muestra un coeficiente Durbin Watson de 0,010 con lo cual no se puede asumir la independencia de los residuos.

	Mínimo	Máximo	Media	Desviación típica	N
Valor pronosticado	6,8496	22,4678	12,4979	1,82504	86782
Valor pronosticado tip.	-3,095	5,463	,000	1,000	86782
Error típico de valor pronosticado	,001	,004	,001	,001	86782
Valor pronosticado corregido	6,8495	22,4679	12,4979	1,82504	86782
Residual	-,81678	,49712	,00000	,16891	86782
Residuo típ.	-4,835	2,943	,000	1,000	86782
Residuo estud.	-4,836	2,943	,000	1,000	86782
Residuo eliminado	-,81690	,49718	,00000	,16892	86782
Residuo eliminado estud.	-4,836	2,943	,000	1,000	86782
Dist. de Mahalanobis	,588	56,050	4,000	5,814	86782
Distancia de Cook	,000	,001	,000	,000	86782
Valor de influencia centrado	,000	,001	,000	,000	86782

Tabla 19. Estadísticos sobre los residuos modelo 3

De acuerdo con la tabla 19, los residuales del modelo 3 presentan un valor mínimo de -0,81678; el valor máximo es de 0,49712; una media de 0,000000 y una desviación de 0,16891.

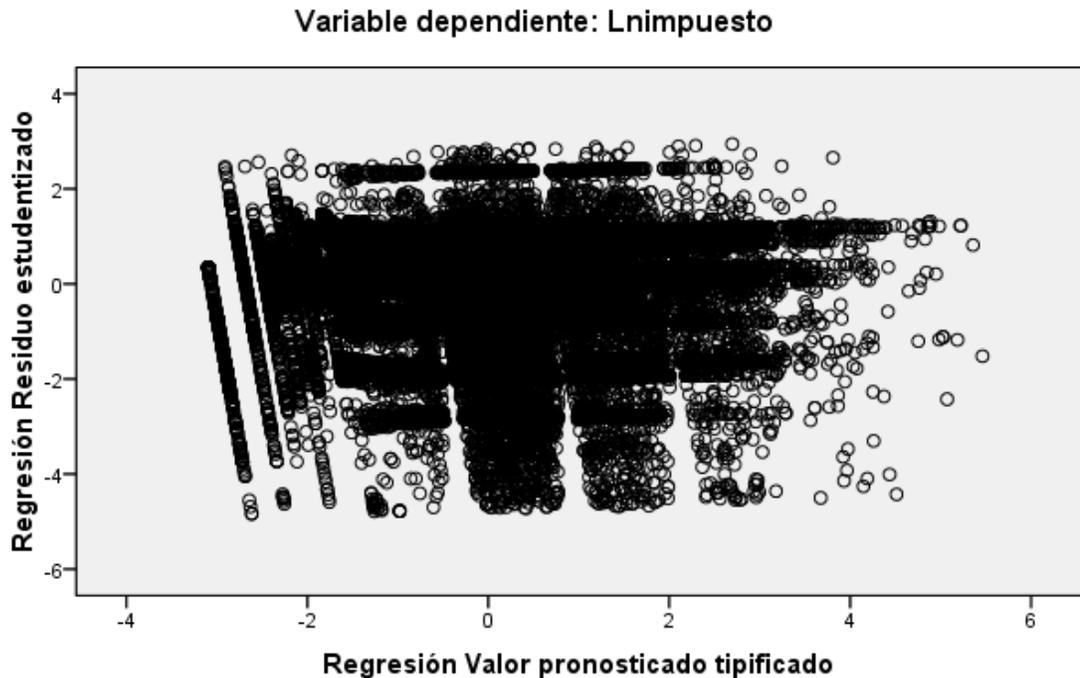
Normalidad. Para validar este supuesto se realizó prueba de hipótesis de normalidad, estos resultados se muestran en la siguiente tabla.

	Kolmogorov-Smirnov ^a		
	Estadístico	gl	Sig.
Unstandardized Residual	,233	86782	,000

Tabla 20. Pruebas de normalidad modelo 3

Basándose en el resultado presentado en la tabla 20, y con un nivel de significancia del 5% se evidencia que los residuales no se distribuyen normalmente.

Homocedasticidad. Se realiza un gráfico de dispersión entre los valores pronosticados con los residuos tipificados. Los resultados correspondientes a los residuales del modelo 3, se presentan en la gráfica 7.



Gráfica 7. Dispersión pronostico tipificado vs residuos estudentizado modelo 3

La gráfica 7, muestra que el supuesto de homocedasticidad, presenta gran similitud con relación al modelo 2.

Por todo lo anterior y en vista de que no se presentaron cambios sustanciales en las estimaciones de los parámetros del modelo 2 y 3, se decide escoger entonces el modelo 2, cuya ecuación, como ya se presentó anteriormente, está dada por:

$$\text{Ln (Valor impuesto declarado)} = -4,422 + 0,996 (\text{Ln ingresos}) - 0,038 (\text{financiera}) - 0,026(\text{industrial}) - 0,191(\text{servicios}).$$

6. CONCLUSIONES

- Existe una correlación fuerte directa, entre los ingresos y el valor a pagar en el impuesto de (0,995), evidenciando que un aumento en los ingresos genera un aumento en el impuesto pagado.
- En cada una de las actividades (financiero, industrial, servicios y comercial), las variables impuesto pagado e ingreso declarado presentan una correlación lineal fuerte y directamente proporcional, ya que el coeficiente de correlación lineal de Pearson es mayor a (0,96608018), indicando que a medida que haya un aumento o disminución en el ingreso ocurre un cambio similar en el impuesto.
- Mediante el ajuste de un modelo de regresión lineal múltiple y teniendo como variable dependiente el logaritmo natural del impuesto y como variables independientes el logaritmo natural del ingreso y cada una de las actividades, se logra concluir que; por cada peso que aumenta el logaritmo natural del ingreso, el logaritmo natural del impuesto aumenta en 0,996 pesos.
- Según los datos del modelo 2, por cada peso que aumenta el logaritmo natural del ingreso, el logaritmo natural del impuesto aumenta en 0,996 pesos en el sector comercial; además con ingresos iguales, si el contribuyente pertenece a la actividad financiera disminuye el pago de impuesto en -0,038 pesos; si pertenece a la actividad industrial el pago disminuye -0,026 pesos y si pertenece al sector industrial el pago disminuye en -0,191 pesos.
- Se evidencia que un aumento en los ingresos de los contribuyentes favorece positivamente el recaudo del impuesto, hecho que trae como resultado un aumento en la capacidad para financiar los planes de desarrollo del Gobierno Distrital.
- Ninguno de los 3 modelos ajustados, cumplen con la totalidad de los supuestos (Independencia, Normalidad y Homocedasticidad), por lo anterior no es posible aplicar uno de ellos para predecir comportamientos futuros.

7. RECOMENDACIONES

- Debido al incumplimiento de los supuestos se recomienda utilizar un modelo lineal generalizado.
- La utilización de modelos econométricos en estudios futuros de igual o similar interés de investigación.
- Realizar un análisis de influencia en los modelos ajustados, para verificar si una observación es demasiado influyente sobre las demás y de esta manera tener mayor certeza en la selección del modelo apropiado para realizar predicciones.

BIBLIOGRAFIA

Estadística aplicada a los negocios y la economía /Allen L webster tercera edición/
Mc Graw Hill ED

(Alean, Acosta y Matos, 2009). *Factores determinantes del recaudo de los impuestos Predial e Industria y Comercio en el Distrito de Cartagena de Indias.*

http://bibliotecadigital.ccb.org.co/bitstream/handle/11520/1298/3522_2008_2cartilla_distritales.pdf?sequence=1

DIRECCIÓN DISTRITAL DE IMPUESTOS. Informe consolidado de recaudo (3, febrero, 2014)

George C Canavos, Probabilidad y Estadística Métodos y Aplicaciones MC GRAW HILL 1998

(Ibarra, 2011) metodología de la investigación Internet:
(<http://metodologiadelainvestigacinsiis.blogspot.com/2011/10/tipos-de-investigación-exploratoria.html>)

Secretaría Distrital de Hacienda. Estatuto Tributario de Bogotá (Decreto 352 de 2002)
<http://www.shd.gov.co/shd/informes-de-gestion>

William Mendenhall Estadística para Administración y Economía 1.981 ISBN 968-7270-13-6

(Zuluaga, 2009). *Boletín número 12 del Ministerio de Hacienda y Crédito Público de la República de Colombia.*

<http://www.minhacienda.gov.co/portal/page/portal/HomeMinhacienda/asistenciaentidadesterritoriales/Cundinamarca/ViabilidadFiscal/Bogota%20junio%202013.pdf>

<http://www.gerencie.com/impuesto-de-industria-y-comercio.html>

<http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=5437>

http://www.usta.edu.co/otras_pag/revistas/r_cife/Cife%2013/17.pdf

<https://ideas.repec.org/p/col/000162/008835.html>