

CIENCIA DE DATOS PARA LA TOMA DE DECISIONES GERENCIALES

JUAN MANUEL CANDELO VIÁFARA
CARLOS FERNANDO OSORIO ANDRADE
EDWIN ARANGO ESPINAL

Catalogación en la publicación

Candelo Viáfara, Juan Manuel, autor

Ciencia de datos para la toma de decisiones gerenciales / Juan Manuel Candelo Viáfara, Carlos Fernando Osorio Andrade, Edwin Arango Espinal -- Primera edición. -- Bogotá: Ecoe Ediciones, 2025.

104 páginas, il., col.; 23.5x32 cm

(Consulta, información y materias interdisciplinarias. Ciencia y análisis de datos: generalidades).

Incluye datos curriculares de los autores -- Incluye referencias bibliográficas.

ISBN 978-958-508-485-8 (impreso) -- 978-958-508-486-5 (pdf) --978-958-508-487-2 (digital)

1.Gestión empresarial -- procesamiento de datos 2.Métodos estadísticos para la gestión 3.Ciencia de datos -- modelado 4. Sistemas de apoyo a decisiones I.Candelo Viáfara, Juan Manuel, autor II.Osorio Andrade, Carlos Fernando, autor III.Arango Espinal, Edwin, autor

CDD: 658.4038011 Ed. 23

nla-



Área: Consulta, información y materias interdisciplinarias

Subárea: Ciencia y análisis de datos: generalidades



Área: Economía, finanzas, empresa y gestión

Subárea: Toma de decisiones en la gestión de las empresas

ECOE
EDICIONES

© Juan Manuel Candelo Viáfara^[1]

© Carlos Fernando Osorio Andrade^[1]

© Edwin Arango Espinal^[1]

© Ecoe Ediciones S.A.S.

info@ecoeediciones.com

www.ecoeediciones.com

Carrera 19 # 63 C 32

Teléfono: (+57) 321 226 46 09

Bogotá, Colombia

► Cita sugerida: Candelo Viáfara, J. M., Osorio Andrade, C. F., y Arango Espinal, E. (2025). *Ciencia de datos para la toma de decisiones gerenciales*. Ecoe Ediciones.

Primera edición: Bogotá, 2025

ISBN: 978-958-508-485-8

e-ISBN (PDF): 978-958-508-486-5

e-ISBN (ePUB): 978-958-508-487-2

Directora editorial: Ana María Rueda G.

Coordinadora de producción editorial:

Alejandra Rondón Forero

Asistente administrativo editorial: Edwin Corzo S.

Corrección de estilo: Camila Gil Sierra

Diagramación: Paula Andrea Cubillos

Carátula: Natalia Herrera F.

Impresión: Carvajal Soluciones de Comunicación S.A.S. Carrera 69 # 15 - 24

^[1]Profesores asociados de la Facultad de Ciencias de la Administración en la Universidad del Valle, Cali - Colombia.

La reproducción total de esta obra, ya sea en formato físico o digital, está estrictamente prohibida sin la autorización expresa del titular de los derechos. Asimismo, cualquier reproducción parcial de este libro, con o sin fines comerciales, en formato físico o digital, requiere la autorización previa.

RESUMEN EJECUTIVO

Ciencia de datos para la toma de decisiones gerenciales es una obra que busca conectar la complejidad técnica del análisis de datos con su aplicación práctica en el contexto empresarial. Este libro se distingue por su enfoque didáctico y aplicado, diseñado específicamente para gerentes, administradores y tomadores de decisiones que necesitan comprender y utilizar de manera efectiva las herramientas de análisis de datos en su gestión diaria.

La obra está estructurada en cuatro capítulos fundamentales que abordan las técnicas más relevantes para el análisis empresarial: 1) “Modelos de corte transversal”: introduce la regresión lineal simple desde una perspectiva práctica, utilizando ejemplos cotidianos de una gerencia naciente como la relación entre presupuesto publicitario y ventas. Se explican detalladamente los supuestos del modelo y su interpretación en contextos empresariales, lo que permite a los lectores comprender cuándo y cómo aplicar estas técnicas en sus organizaciones. 2) “Datos de conteo”: aborda el análisis de variables discretas frecuentes en el entorno empresarial, como el número de visitantes, ventas diarias o incidencias. Se presentan los modelos de Poisson y binomial negativa, con ejemplos prácticos que facilitan su comprensión y aplicación. 3) “Modelos binarios”: se enfoca en el análisis de decisiones dicotómicas (sí/no) mediante modelos logit, utilizando casos prácticos como la evaluación de candidatos en recursos humanos o decisiones de inversión. Esta sección resulta especialmente útil para la toma de decisiones basadas en probabilidades. 4) “Series de tiempo”: explora técnicas para analizar datos temporales, fundamentales para

la previsión y planificación empresarial. Se incluyen modelos ARIMA, regresión cuantílica y vectores autorregresivos, con aplicaciones prácticas en análisis de mercados financieros y predicción de variables económicas.

Esta obra resulta particularmente relevante en el contexto actual, donde la toma de decisiones basada en datos se ha convertido en un factor clave para la competitividad empresarial. El texto ofrece las herramientas necesarias para que los gerentes aprovechen el poder de los datos en sus decisiones estratégicas, sin necesidad de ser expertos en estadística o programación.

El libro está diseñado tanto para profesionales en ejercicio como para estudiantes de programas de administración y negocios, brindándoles una base sólida para la toma de decisiones *basada en datos* en el entorno empresarial moderno.

Palabras clave: Ciencia de datos, análisis empresarial, toma de decisiones, modelos estadísticos, gestión basada en datos.

CONTENIDO

RESUMEN EJECUTIVO	v
CAPÍTULO 1. MODELOS DE CORTE TRANSVERSAL.....	1
1.1. Introducción a la regresión simple.....	2
1.2. Importancia y aplicaciones en administración y negocios	4
1.3. Supuestos del modelo de regresión simple.....	5
1.3.1 Supuesto de linealidad.....	5
1.3.2 Supuesto de independencia.....	7
1.3.3 Supuesto de homocedasticidad	8
1.3.4 Supuesto de normalidad	10
1.3.5 Supuesto de multicolinealidad.....	14
1.4. Aplicación e interpretación del modelo de regresión	16
1.4.1 Interpretación de resultados	17
1.4.2 Evaluando la calidad del modelo.....	17
1.4.3 Interpretando una salida de resultados en un software estadístico ...	18
1.5. ¿Son importantes los modelos de corte transversal?.....	19
CAPÍTULO 2. DATOS DE CONTEO	21
2.1. Introducción a los datos de conteo.....	22
2.2. Determinando si los datos son de conteo.....	23
2.3. Rechazando el supuesto de normalidad	24
2.4. Distribuciones comunes para datos de conteo	25
2.4.1 Distribución de Poisson.....	25
2.4.2 Distribución binomial negativa.....	26

2.4.3 Selección de la distribución adecuada.....	26
2.5. Supuestos de los modelos de conteo	26
2.5.1 Homocedasticidad	26
2.5.2 Ausencia de multicolinealidad	27
2.5.3 Independencia de observaciones	27
2.6. Aplicación del modelo de conteo	28
2.7. Los datos de conteo y su importancia en la toma de decisiones	32
CAPÍTULO 3. MODELOS BINARIOS.....	35
3.1. Introducción a los modelos logit.....	36
3.2. Características de la variable dependiente binaria.....	37
3.3. Supuestos de la regresión binaria	40
3.4. Aplicación del modelo logit	41
3.5. Los modelos binarios y su desarrollo en la modernidad empresarial	43
CAPÍTULO 4. SERIES DE TIEMPO	45
4.1. Una serie de tiempo.....	46
4.2. Modelos autorregresivos integrados de medias móviles (ARIMA).....	53
4.3. Regresión cuantílica en series de tiempo.....	63
4.4. Modelo VAR.....	68
4.5. Series de tiempo: una complejidad sencilla de abordar	78
CONCLUSIONES GENERALES.....	81
REFERENCIAS	83

ÍNDICE DE FIGURAS

Figura 1.	Proceso de carga de bases de datos.....	2
Figura 2.	Regresión lineal simple: presupuesto de publicidad vs. ventas.....	3
Figura 3.	Gráfico de dispersión: presupuesto publicitario vs. ventas.....	6
Figura 4.	Comando para el modelo de regresión.....	6
Figura 5.	Comando para el modelo de regresión.....	7
Figura 6.	Gráfico de residuos: verificación de homocedasticidad.....	8
Figura 7.	Gráfico de residuos: valores ajustados vs. residuos (relación no lineal)....	9
Figura 8.	Comando para el supuesto de homocedasticidad.....	10
Figura 9.	Visualización del resultado	10
Figura 10.	Distribución de los errores de predicción (ventas y presupuesto publicitario).....	11
Figura 11.	Gráfico Q-Q para los errores de predicción (ventas y presupuesto publicitario).....	12
Figura 12.	Comando para el gráfico Q-Q	12
Figura 13.	Comando para el gráfico Q-Q	13
Figura 14.	Comando para la transformación logarítmica	13
Figura 15.	Comando para VIF.....	16
Figura 16.	Distribución de visitantes en un parque de atracciones	22
Figura 17.	Comando para el supuesto Shapiro-Wilk.....	25
Figura 18.	Comando para la regresión de Poisson.....	29

Figura 19.	Comando para la regresión binomial negativa.....	29
Figura 20.	Comando para tasas de incidencia.....	32
Figura 21.	Distribución de asistencia a la fiesta de fin de año.....	37
Figura 22.	Relación no lineal entre años de experiencia y probabilidad de éxito....	38
Figura 23.	Comparación entre relación lineal y curva sigmoide	39
Figura 24.	Comando logit	42
Figura 25.	Carga de datos en el <i>software</i>	46
Figura 26.	Selección de las variables a estudiar en el <i>software</i> EViews®	47
Figura 27.	Creación de una gráfica en EViews®.....	47
Figura 28.	Configuración del gráfico en EViews®	48
Figura 29.	Configuración del gráfico de series normalizadas en EViews®	49
Figura 30.	Precios del petróleo WTI y tasa de cambio del peso colombiano frente al dólar	50
Figura 31.	Precios del petróleo WTI y tasa de cambio del peso colombiano frente al dólar, normalizados.....	51
Figura 32.	Proceso de transformación de las variables en EViews®	52
Figura 33.	Notación para la transformación de una serie en retornos en EViews®	52
Figura 34.	Retornos WTI y COP.....	53
Figura 35.	COP con transformación logarítmica.....	55
Figura 36.	Proceso para hallar la raíz unitaria.....	56
Figura 37.	Diferencia del logaritmo COP	57
Figura 38.	Función de autocorrelación en EViews®	58
Figura 39.	Funciones de autocorrelación y autocorrelación parcial.....	59
Figura 40.	Función de autocorrelación y autocorrelación parcial de COP diferenciada	60
Figura 41.	Proceso de estimación del modelo ARIMA en EViews®	61
Figura 42.	Selección del método para la estimación del modelo ARIMA en EViews®	61
Figura 43.	Configuración del modelo ARIMA en EViews®	62
Figura 44.	Asignación de variables para la regresión cuantílica	64
Figura 45.	Asignación del método para la regresión cuantílica	65
Figura 46.	Estimación de los cuantiles en la regresión cuantílica.....	66
Figura 47.	Pestaña de resultados de la regresión cuantílica.....	66
Figura 48.	Coefficientes de la regresión cuantílica de WTI respecto al COP.....	67
Figura 49.	Selección de variables para el modelo VAR	69
Figura 50.	Identificación del número de rezagos en EViews®	71
Figura 51.	Selección de la función de autocorrelación en EViews®	72
Figura 52.	Creación de la prueba de raíz inversa en EViews®.....	73
Figura 53.	Prueba de raíces inversas del polinomio característico AR	74

Figura 54.	Creación de la función impulso-respuesta en EViews®	74
Figura 55.	Configuración de la función impulso-respuesta en EViews®	75
Figura 56.	Respuesta COP a WTI	76
Figura 57.	Respuesta Gas a WTI	76
Figura 58.	Selección de la descomposición de varianza en EViews®	77
Figura 59.	Configuración de la descomposición de varianza en EViews®	77

ÍNDICE DE TABLAS

Tabla 1.	Valores del factor de inflación de la varianza VIF para las variables independientes.....	15
Tabla 2.	Resultados del modelo de regresión	18
Tabla 3.	Resultados de la prueba de normalidad de Shapiro-Wilk para la variable dependiente	24
Tabla 4.	Resultado del modelo binomial negativo.....	30
Tabla 5.	Estimación del modelo binomial negativo.....	30
Tabla 6.	Tasas de incidencia basadas en los coeficientes del modelo binomial negativo.....	32
Tabla 7.	Ejemplo de resultados del modelo logit.....	41
Tabla 8.	Prueba de Dickey-Fuller.....	56
Tabla 9.	Prueba de Dickey-Fuller para la diferencia de LCOP	57
Tabla 10.	Estimación del modelo ARIMA	62
Tabla 11.	Coefficientes de la regresión del WTI respecto al COP en los diferentes deciles.....	68
Tabla 12.	Pruebas para elegir el número de rezagos	71
Tabla 13.	Prueba de autocorrelación.....	73
Tabla 14.	Descomposición de la varianza	78

Sistema de Información en Línea



Al final del libro encontrará la información para ingresar al **Sistema de Información en Línea -SIL-** donde podrá acceder a los siguientes recursos complementarios:

- » Video introductorio (Introducción.mp4): presenta los contenidos del libro.
- » Presentación en PowerPoint sobre sistemas de información: diseñada como guía de estudio y herramienta de apoyo en entornos académicos y ejecutivos. Su estructura modular permite una lectura dinámica e incremental, abordando conceptos clave.
- » Archivo Excel interactivo (Laboratorio Regresión): contiene macros, fórmulas y visualizaciones para experimentar con modelos de regresión simple y logística. Este recurso permite al lector trabajar con datos reales o simulados y comprender de forma práctica el análisis estadístico. Incluye una guía de uso.

Estos materiales profundizan en los temas abordados en los capítulos 1, 2 y 3 del libro, proporcionando una experiencia de aprendizaje dinámica y aplicada.