

Análisis de simulaciones de mercado con ayuda de inteligencia artificial

Ana Paula González Muñoz
Dennis Daniel González Durán
Frank H. Perez Fleita
C312

1. Introducción

Las simulaciones basadas en agentes han emergido como una poderosa herramienta en el campo de la economía y la teoría de mercados para modelar sistemas complejos y dinámicos. En el contexto de los mercados, estas simulaciones permiten la representación de interacciones entre actores heterogéneos, como empresas, consumidores y proveedores, cuyos comportamientos individuales y decisiones afectan al sistema en su conjunto. Este enfoque contrasta con los modelos económicos tradicionales, que tienden a suponer racionalidad perfecta y equilibrio estático, mientras que las simulaciones basadas en agentes capturan las dinámicas emergentes y adaptativas de los mercados reales.

El uso de simulaciones en el estudio de mercados ofrece ventajas significativas. Permiten analizar la evolución de precios, la competencia entre empresas, la dinámica de la oferta y demanda, así como los efectos de cambios en el entorno económico. Además, estas simulaciones pueden incorporar la toma de decisiones descentralizada y el comportamiento adaptativo de los agentes, lo que facilita la investigación de fenómenos como la competencia imperfecta y las estrategias comerciales.

En este contexto, las empresas actúan como agentes económicos que producen y venden bienes o servicios, buscando maximizar sus beneficios. Por otro lado, los consumidores buscan adquirir productos que maximicen su satisfacción o utilidad, mientras que los suministradores proporcionan a las empresas los insumos necesarios para la producción. Las interacciones entre estos agentes forman un mercado en el que los precios, volúmenes de transacción y otros parámetros emergen como resultado de las decisiones y estrategias individuales de los actores involucrados.

El presente informe describe una simulación de mercado basada en agentes que modela estas interacciones entre empresas, clientes y suministradores. A través de esta simulación, se exploran dinámicas clave, como la competencia por recursos, las estrategias de fijación de precios y la adaptación a cambios en la oferta y la demanda. Este enfoque proporciona una visión detallada y flexible de cómo los comportamientos locales de los agentes pueden generar patrones globales y propiedades emergentes dentro del mercado.

2. Modelación del problema

2.1. Características del ambiente

1. **Accesibilidad:** Los agentes tienen acceso completo a los datos generales del mercado como los precios de los productos y la popularidad de las compañías.
2. **No determinista:** Al estar compuesta la simulación por diversos agentes que buscan lo mejor para si mismos, la variación del mercado a cada decisión no será única, debido a que esta depende de las diferentes acciones que realizan el conjunto de agentes. Por ejemplo, si una empresa decide bajar sus precios no significa que el mes siguiente tenga mejores ventas, dado que su competencia puede disminuir más sus precios.
3. **No episódico:** Las empresas pueden adaptar sus decisiones actuales a las consecuencias que las mismas tendrán en el futuro. Por ejemplo, cada empresa toma la decisión de como dividir su presupuesto basándose en como creen que afectará esta distribución en sus ventas futuras.

4. **Dinámico:** Aunque un agente en específico no realice ninguna acción, el resto de los integrantes del mercado sí lo harán, por lo que el medio se mantendrá en constante cambio.
5. **Continuo:** Los agentes deben adaptarse constantemente a las condiciones cambiantes del mercado. Esta necesidad de adaptación y decisión rápida contribuye a la evolución del mercado, ya que cada decisión puede tener un impacto significativo en el comportamiento del mismo.

2.2. Elementos del ambiente modelados

1. **Oferta de las empresas:** Representan una información crucial para el mercado. Aquí se almacenan los precios a los que las empresas venden los productos así como el inventario de los mismos.
2. **Popularidad de las compañías:** Representa que tan conocidos son los productos en el mercado. Refleja los resultados de las inversiones en *marketing* y como se afecta la popularidad con el paso del tiempo.

2.3. Otros elementos modelados

1. **Actitud de los clientes:** Representa el comportamiento de la población el cual influye directamente en la forma en la que compran. Por ejemplo, los clientes populistas priorizarán realizar sus compras en las empresas que tengan los productos más populares.

3. Sistema multiagente

En este caso se trabaja con un sistema compuesto por 3 tipos de agentes: clientes, empresas y suministradores, donde cada uno cuenta con conocimientos, deseos y acciones diferentes. La existencia de estos, así como sus interacciones determinan el comportamiento del mercado.

3.1. Características de los agentes

1. **Reactivos:** Los agentes son reactivos debido a que operan basándose en sus percepciones actuales del entorno para tomar decisiones y actuar de manera oportuna. Cada agente percibe el estado actual del mercado y toma una decisión informada para cumplir con sus objetivos. Por ejemplo, si un cliente nota un cambio en los precios puede optar por comprar una menor cantidad de un producto o comprar el mismo a otra empresa.
2. **Proactivos:** Son proactivos porque, en lugar de simplemente reaccionar a los cambios en el entorno, toman la iniciativa para anticipar y prepararse para posibles escenarios. Por ejemplo, las empresas tratan de controlar su destino al no limitarse a responder a eventos como ocurren. En cambio, toman medidas para influir en sus ventas futuras, seleccionando estrategias y acciones que les permitan tener una mayor ganancia.
3. **Sociables:** Los agentes reciben información de otros y la utilizan para tomar decisiones y cumplir sus objetivos. Por ejemplo, para que una empresa pueda producir necesita llevar a cabo un proceso de negociación con los suministradores, mediante el cual las empresas buscan minimizar el costo de los subproductos que necesita y los proveedores maximizar sus ganancias.

3.1.1. Estructura de los agentes

La estructura BDI (Belief-Desire-Intention) es un modelo de agente inteligente que se utiliza en la simulación de sistemas complejos, como los del mercado, debido a su capacidad para manejar múltiples tipos de agentes con diferentes roles y objetivos. Este modelo se basa en tres componentes principales:

- **Beliefs (Creencias):** Representan el conocimiento del agente sobre el mundo. En el contexto del mercado, esto incluye información sobre los precios de los productos y la popularidad de los mismos.

- **Desires (Deseos):** Son los objetivos que el agente quiere alcanzar. Para una empresa, podría ser administrar correctamente su presupuesto para maximizar sus ganancias a futuro y para un cliente, satisfacer su necesidad de un producto.
- **Intentions (Intenciones):** Son los planes que el agente tiene para alcanzar sus deseos. Esto incluye las acciones específicas que el agente tomará. Por ejemplo, comprar productos baratos o ajustar los precios de los productos.

3.2. Agente Empresa

En la simulación de mercado, las empresas son modeladas como agentes que operan bajo una arquitectura BDI, lo que les permite percibir su entorno, generar deseos en función de sus objetivos económicos y tomar decisiones adaptativas para maximizar sus ganancias. Estos agentes siguen un proceso de toma de decisiones que integra sus creencias actuales sobre el entorno, la generación de opciones estratégicas, la selección de intenciones y la ejecución de acciones basadas en estas intenciones, ajustando continuamente su comportamiento a medida que se desarrolla el mercado.

3.2.1. Creencias del Agente Empresa

Las creencias de las empresas incluyen información clave que obtienen al observar su entorno de mercado. Las empresas perciben los precios de sus competidores, el comportamiento reciente de sus ventas y la popularidad de sus productos entre los clientes. Esta percepción les permite evaluar el rendimiento de sus estrategias comerciales y ajustar su enfoque. La revisión continua de estas creencias les facilita adaptarse a las dinámicas del mercado y mejorar sus decisiones sobre producción, precios y marketing. A medida que el mercado evoluciona, las creencias del agente se actualizan para reflejar estos cambios.

3.2.2. Generación de Deseos

Basándose en sus creencias, las empresas desarrollan una serie de deseos que responden a su principal objetivo: maximizar las ganancias. Los deseos de las empresas incluyen la administración eficiente de su presupuesto, la gestión de subproductos y la optimización de la producción. Cada uno de estos deseos está orientado a garantizar que los recursos de la empresa se utilicen de manera óptima, con el fin de mejorar su posición competitiva y aumentar su rentabilidad. La generación de deseos también considera la necesidad de mantenerse competitivas mediante la adaptación a las condiciones cambiantes del mercado.

3.2.3. Filtrado de Intenciones

A partir de los deseos generados, el agente empresa selecciona las intenciones que mejor se alinean con sus creencias actuales y su contexto competitivo. Entre las posibles intenciones que puede elegir, se encuentran decisiones como la distribución del presupuesto entre los distintos productos que ofrece, la planificación de la inversión en marketing para aumentar la popularidad de un producto, la negociación con los proveedores para adquirir subproductos, la producción de bienes y el ajuste de precios de los productos en función de la demanda y la competencia. Este proceso de filtrado de intenciones asegura que las empresas seleccionen las estrategias más adecuadas a sus circunstancias específicas.

3.2.4. Ejecución de Acciones

Una vez que se ha seleccionado la intención más adecuada, el agente procede a ejecutar las acciones correspondientes. Por ejemplo, si el agente decide ajustar los precios de sus productos, lleva a cabo un análisis basado en las creencias actuales sobre el comportamiento de los competidores y la demanda del mercado. Si elige invertir en marketing, se destinan los recursos necesarios para aumentar la visibilidad del producto y, en consecuencia, su popularidad entre los clientes. Asimismo, las empresas pueden negociar con los proveedores para asegurar la disponibilidad de subproductos necesarios para la producción. Este ciclo de toma de decisiones se repite continuamente, lo que permite a las empresas adaptarse a las fluctuaciones del mercado de manera proactiva y eficiente.

3.3. Agente Cliente

En la simulación de mercado, los clientes son modelados como agentes que, utilizando una arquitectura BDI, actúan para satisfacer sus necesidades de productos de acuerdo con sus restricciones presupuestarias. Cada cliente tiene un presupuesto mensual, que representa su capacidad de gasto, y este presupuesto se asigna de acuerdo a distintos criterios que dividen a los clientes en diferentes grupos o sectores de la población. Este enfoque permite modelar a los clientes con diferentes niveles de ingresos, lo que influye en la cantidad de recursos que cada uno puede destinar a diferentes productos.

3.3.1. Creencias del Agente Cliente

Los clientes, basándose en su percepción del mercado y su actitud (populista, tacaño, random o precavido), forman creencias sobre los productos disponibles. Estas creencias incluyen factores como los precios de los productos y su popularidad. La información que los clientes recopilan del entorno les permite adaptar sus decisiones de compra según las condiciones del mercado.

3.3.2. Generación de Deseos

A partir de sus creencias y de las restricciones de su presupuesto, los clientes generan deseos orientados a la satisfacción de sus necesidades. Estos deseos guían sus decisiones sobre cómo distribuir su presupuesto entre los diferentes productos. En función de su necesidad de un producto específico, los clientes priorizan sus compras, buscando maximizar la satisfacción personal dentro de las limitaciones impuestas por sus ingresos.

3.3.3. Filtrado de Intenciones

Los deseos se traducen en intenciones que determinan las estrategias de compra que los clientes seguirán para adquirir los productos. Las intenciones de los clientes pueden variar dependiendo de la importancia que le den a factores como el precio y la popularidad del producto. Entre las estrategias que pueden adoptar están priorizar productos de bajo costo, optar por aquellos con mayor popularidad o encontrar un equilibrio entre precio y popularidad. Algunos clientes incluso pueden tomar decisiones de compra aleatorias. El filtrado de intenciones permite que cada cliente seleccione la estrategia de compra más adecuada a su contexto particular.

3.3.4. Ejecución de Acciones

Una vez seleccionadas sus intenciones, los agentes clientes ejecutan las acciones necesarias para adquirir los productos. Esto incluye decidir qué productos comprar, de qué empresas y en qué cantidad, tomando en cuenta su presupuesto y las necesidades previamente identificadas. A través de estas decisiones, los clientes representan la demanda del mercado, influyendo en las dinámicas de oferta y precio de los productos.

3.4. Agente Suministrador

En el contexto de la simulación de mercado, el agente suministrador representa a los proveedores que proporcionan los insumos necesarios para la producción de las empresas. A diferencia de los agentes cliente y empresa, el agente suministrador no sigue una arquitectura BDI, ya que su comportamiento es puramente reactivo y negociador. Su objetivo principal es maximizar sus beneficios mediante la negociación directa con las empresas, evaluando ofertas y generando contraofertas en función de las condiciones actuales del mercado.

3.4.1. Proceso de Negociación

El agente suministrador recibe propuestas de las empresas interesadas en adquirir productos específicos. Utilizando lógica difusa, el agente evalúa cada oferta recibida para decidir si la acepta, la rechaza o propone una contraoferta. Este enfoque flexible de la negociación permite que el agente interprete las condiciones de la oferta, como el precio y la cantidad solicitada, en términos relativos en lugar de absolutos. En caso de que el suministro disponible sea inferior a la cantidad solicitada, el agente ajusta la oferta a la cantidad disponible y propone una cantidad alternativa en la negociación.

4. Integración de inteligencia artificial

El uso de la inteligencia artificial (IA) en las simulaciones de mercado permite mejorar la precisión y realismo de los simuladores, debido a que los agentes pueden tomar decisiones que sin el uso de la IA serían muy difíciles de modelar. En la modelación planteada se hace uso de la inteligencia artificial sobre todo en los agentes inteligentes implementados y algunas de sus funcionalidades.

4.1. Conocimiento/Razonamiento usando lógica difusa

La lógica difusa es una extensión de la lógica clásica que permite manejar incertidumbre y vaguedad en la toma de decisiones. A diferencia de la lógica tradicional, en la que las variables son binarias (verdadero o falso), la lógica difusa trabaja con grados de verdad, lo que permite modelar situaciones más complejas y realistas. Este enfoque es particularmente útil en sistemas donde las decisiones no son claramente blanco o negro, sino que involucran una escala de valores intermedios, como sucede en muchos contextos empresariales.

En nuestra simulación de mercado, la lógica difusa se utiliza para resolver problemas de decisión que involucran variables ambiguas, como la asignación de presupuesto entre producción y marketing, y el ajuste de precios de los productos. A continuación, se describen los procesos donde la lógica difusa juega un papel clave:

4.1.1. Asignación de Presupuesto entre Producción y Marketing

Una vez que las empresas deciden cuánto presupuesto destinar a un producto, enfrentan la compleja tarea de decidir qué porcentaje de ese presupuesto se asignará a la producción y cuánto al marketing. Este es un proceso que depende de múltiples factores que no siempre son fáciles de cuantificar. Para abordar este desafío, las empresas en la simulación utilizan un sistema basado en lógica difusa.

Las empresas evalúan dos variables difusas:

- Ventas del producto, que se clasifican en conjuntos difusos como "bajas", "promedio." "altas".
- Popularidad del producto, que también se divide en "baja", "promedio." "alta".

Con base en estas evaluaciones, se aplican un conjunto de reglas difusas que determinan qué porcentaje del presupuesto debe destinarse a la producción del producto. Los porcentajes de asignación se describen en términos difusos como "muy bajo", "bajo", "medio", "alto." "muy alto". Por ejemplo, si las ventas son altas y la popularidad del producto también es alta, la regla difusa puede indicar que una mayor parte del presupuesto debería destinarse a la producción. En cambio, si las ventas son bajas y la popularidad es baja, podría asignarse un porcentaje menor a la producción y destinar más recursos al marketing para mejorar la visibilidad del producto.

4.1.2. Ajuste de Precios

Además, la lógica difusa también se emplea en el proceso de ajuste de precios, que es crucial para que las empresas mantengan su competitividad. El sistema difuso tiene en cuenta dos factores principales:

- Popularidad del producto, que puede ser baja, promedio o alta.
- Calidad de las ventas, clasificada igualmente en "baja", "promedio." "alta".

Basándose en estas variables, se definen reglas que determinan qué tan alto debería ser el margen de ganancia aplicado al producto, expresado en términos como "muy bajo", "bajo", "medio", "alto." "muy alto". Por ejemplo, si un producto tiene una alta popularidad y las ventas han sido excelentes, la lógica difusa podría sugerir un margen de ganancia más alto, aprovechando la demanda elevada. En cambio, si la popularidad y las ventas son bajas, el sistema podría recomendar un margen más bajo para estimular la demanda a través de precios más competitivos.

4.1.3. Evaluación de Ofertas de los Agentes Suministradores

Para decidir si aceptar una oferta o generar una contraoferta, el agente suministrador utiliza lógica difusa basada en las condiciones del mercado y la oferta específica. Este proceso considera factores como el precio ofertado y la cantidad solicitada, permitiendo que el agente tome decisiones adaptativas.

Evaluación del Precio Basado en la Cantidad Solicitada: La cantidad solicitada se clasifica en términos difusos como "baja", "media." "alta." en relación con el stock disponible. Basándose en esta clasificación, el agente ajusta el precio de la oferta para maximizar su rentabilidad; una cantidad baja tiende a un precio alto, mientras que una cantidad alta permite un descuento.

Ajuste de la Cantidad Basado en el Precio Ofrecido: Al recibir una propuesta de precio, el agente suministrador evalúa este valor en términos de "bajo", "medio." "alto" respecto al precio mínimo aceptable. Este análisis permite al agente ajustar la cantidad de productos ofrecidos, siendo más flexible en la cantidad a medida que el precio aumenta. Si la oferta resulta aceptable, el agente la acepta; de lo contrario, propone una contraoferta ajustada.

4.2. Búsqueda de la mejor estrategia de división de presupuesto

En problemas de toma de decisiones complejos, como la distribución de recursos en un entorno empresarial, es fundamental encontrar una estrategia óptima que permita maximizar los beneficios. En nuestra simulación, las empresas deben decidir cómo distribuir su presupuesto entre una gama de productos, lo que constituye un desafío dado el número de variables y la incertidumbre del mercado. Para abordar esta tarea, hemos implementado un algoritmo genético, una técnica de optimización inspirada en la evolución natural que es eficaz para resolver problemas de búsqueda en grandes espacios de soluciones.

4.2.1. Representación de los Individuos

En nuestro modelo, cada individuo dentro del algoritmo genético representa una posible solución al problema de distribución de presupuesto entre productos. Los genes de cada individuo son una serie de porcentajes que indican qué proporción del presupuesto total se dedicará a cada producto. Así, el genoma completo de un individuo define una estrategia de distribución de recursos. Estos genes no se eligen al azar, sino que están sujetos a la evolución y selección natural dentro del algoritmo.

4.2.2. Función de Ajuste (Fitness)

La función de ajuste o fitness es un componente clave del algoritmo genético, y en nuestro caso, se basa en una simulación mensual que permite evaluar el rendimiento de cada distribución de presupuesto. Para cada individuo, se simula el comportamiento de la empresa durante el mes siguiente, considerando la inversión en producción y marketing de cada producto. El valor de la función de ajuste es la ganancia total obtenida al final de la simulación, de modo que los individuos con distribuciones de presupuesto más rentables tendrán un mayor valor de fitness. Esto asegura que las mejores estrategias se seleccionen para evolucionar en generaciones futuras.

4.2.3. Cruce (Crossover)

El siguiente paso es el cruce, donde los individuos de mayor fitness se combinan para generar nuevas soluciones. En nuestro caso, se utilizan dos metaheurísticas de cruce:

- Selección del mejor gen: Para cada par de productos (genes), se selecciona el valor que haya generado más ganancias en relación a la inversión en la simulación. Esto permite crear nuevos individuos que retienen las mejores características de sus "padres".
- Cruce ponderado: En lugar de seleccionar directamente el mejor gen, se ponderan los genes de ambos padres, dando más peso a los individuos con mayor valor de fitness. Esto introduce una mayor diversidad en la población, al tiempo que sigue priorizando las características de los individuos más exitosos.

4.2.4. Mutación

La mutación es un proceso que introduce variabilidad en la población y evita que el algoritmo quede atrapado en óptimos locales. En nuestra simulación, la mutación consiste en identificar el producto que genera el menor retorno sobre la inversión (es decir, que produce menos ganancias en comparación con el dinero invertido) y eliminarlo del mercado. Este proceso garantiza que las empresas puedan ajustar su portafolio de productos de manera dinámica, descartando productos no rentables.

4.2.5. Evolución y Optimización

A lo largo de varias generaciones, el algoritmo genético selecciona, cruza y muta a los individuos, evolucionando hacia estrategias de distribución de presupuesto cada vez más eficientes. Este proceso iterativo permite que las empresas encuentren una distribución óptima de sus recursos, maximizando así sus ganancias en el entorno competitivo del mercado.

4.3. Procesamiento de lenguaje natural

El procesamiento de lenguaje natural (PLN) se utiliza en esta simulación para generar resúmenes mensuales detallados del comportamiento de los agentes dentro del mercado. Estos resúmenes, generados automáticamente tras cada simulación, proporcionan una visión estructurada y clara de las dinámicas clave, incluyendo patrones de inversión, comportamiento de compra y resultados financieros. Esto facilita la interpretación de datos extensos y complejos de manera efectiva.

4.3.1. Implementación del PLN en la Simulación

El sistema utiliza un modelo de PLN avanzado, en este caso Gemini-1.5-flash, que emplea el contexto y los logs generados durante la simulación para producir un resumen estructurado de cada ciclo mensual. Tras cada simulación, el proceso sigue los siguientes pasos:

- Preparación del Contexto: El contexto necesario para generar el resumen se encuentra en un archivo JSON, que contiene las instrucciones específicas para el modelo. Este archivo incluye un formato que define cómo estructurar el análisis y resaltar los aspectos clave:
 - Inversiones de cada empresa en áreas como producción y marketing.
 - Patrones de Compra entre los clientes, especificando las tendencias de preferencia y patrones de compra según precio o popularidad.
 - Ganancias y Deudas de cada empresa, reflejando los resultados financieros mensuales en términos monetarios precisos.

Generación de Resumen: Utilizando este contexto y el log mensual generado en la simulación, el modelo de PLN sintetiza la información relevante en un archivo markdown (.md). Esta respuesta se guarda en la carpeta de summaries con el nombre `summary_simulation.i.md`, donde `i` representa el número del mes, permitiendo un fácil acceso a los resultados de cada simulación.

Formato del Resumen: El análisis para cada mes se genera en una estructura estándar que asegura coherencia y facilita la comparación de tendencias mensuales. Los informes incluyen:

Inversiones: Enumera las áreas principales de inversión de cada empresa, indicando los productos priorizados en producción. **Patrones de Compra:** Describe las preferencias de los clientes hacia productos específicos y las tendencias de compra según precio y popularidad. **Ganancias y Deudas:** Muestra las ganancias o pérdidas exactas de cada empresa. Este uso del PLN permite que cada simulación produzca un resumen claro y conciso de los eventos del mercado, ahorrando tiempo y proporcionando una herramienta de análisis que facilita la comprensión de las tendencias de comportamiento de los agentes y las fluctuaciones del mercado.

5. Resultados y experimentos

Con el objetivo de recopilar información de interés de la simulación y realizar posteriormente análisis estadísticos con los datos se llevaron a cabo 30 simulaciones aleatorias de 12 meses cada una, cuyos resultados se guardaron en el archivo `simulation_results.csv`

	ganancia_empresa	min_salary	mean_salary	customer_attitude_stingy \
count	30.000000	30.000000	30.000000	30.000000
mean	93327.005000	1235.066667	1737.900000	0.236647
std	54557.957599	440.513880	440.03874	0.156630
min	25550.950000	516.000000	984.000000	0.002492
25%	43822.722500	918.000000	1421.500000	0.114774
50%	85486.695000	1229.000000	1806.500000	0.199276
75%	124864.285000	1595.500000	2025.750000	0.315817
max	217976.640000	1980.000000	2730.000000	0.562364

	customer_attitude_populist	customer_attitude_cautious \
count	30.000000	30.000000
mean	0.263881	0.238107
std	0.151525	0.130974
min	0.007571	0.037169
25%	0.152637	0.133276
50%	0.240249	0.231598
75%	0.344872	0.299804
max	0.590138	0.548662

Las variables extraídas y a partir de las cuales se trabajará a continuación son:

Del mercado:

- Customer attitude stingy: por ciento de la población que prioriza los precios más bajos
- Customer attitude populist: por ciento de la población que prioriza los productos más populares
- Customer attitude cautious: por ciento de la población que realiza compras de manera precavida
- Customer attitude random: por ciento de la población que realiza sus compras de manera random
- Marketing cost: dinero necesario para aumentar una unidad de popularidad en un producto
- Lose popularity: unidades de popularidad que pierde cada producto al final de cada mes
- Min salary: salario mínimo de los clientes de la población
- Mean salary: salario medio de los clientes de la población

De los resultados de las simulaciones:

- Ganancia de la empresa: dinero obtenido por la empresa al final del año
- Precio promedio producto_: para cada producto se almacena el precio promedio al que fue vendido
- Precio mínimo producto_: para cada producto se almacena el precio mínimo al que fue vendido
- Precio máximo producto_: para cada producto se almacena el precio máximo al que fue vendido
- Unidades producidas producto_: cantidad total de unidades producidas para cada producto
- Unidades vendidas producto_: cantidad total de unidades vendidas para cada producto
- Popularidad inicial producto_: popularidad que tiene cada producto después del primer mes
- Popularidad final producto_: popularidad que tiene cada producto después del último mes

5.1. Análisis inicial e hipótesis sobre los datos

Se comenzó el análisis describiendo el conjunto de datos con el objetivo de ver en qué rango de valores trabajaremos. Estos fueron los resultados:

5.1.1. Impacto del salario medio de los clientes en las ganancias de la empresa

Estamos evaluando la hipótesis de que el salario medio de los clientes influye positivamente en las ganancias de la empresa, bajo el supuesto de que los consumidores con ingresos más altos pueden permitirse gastar más, lo que podría llevar a un aumento en las ventas y, en última instancia, en las ganancias. Para verificar esta relación, hemos utilizado el coeficiente de correlación de Pearson, que mide la fuerza y la dirección de la relación lineal entre dos variables continuas: el salario medio y las ganancias de la empresa.

	customer_attitude_random	lose_popularity	marketing_cost \
count	30.000000	30.000000	30.000000
mean	0.261365	12.466667	65.833333
std	0.139918	4.717782	17.661732
min	0.060012	5.000000	31.000000
25%	0.125644	8.500000	53.250000
50%	0.259542	13.000000	64.000000
75%	0.339630	15.750000	81.250000
max	0.572027	20.000000	98.000000

	precio_medio_product_1 ...	precio_maximo_product_2 \
count	30.000000 ...	30.000000
mean	190.300000 ...	214.500000
std	186.134017 ...	85.876716
min	0.000000 ...	118.000000
25%	43.000000 ...	137.250000
50%	133.500000 ...	198.500000
75%	332.500000 ...	274.250000
max	762.000000 ...	479.000000

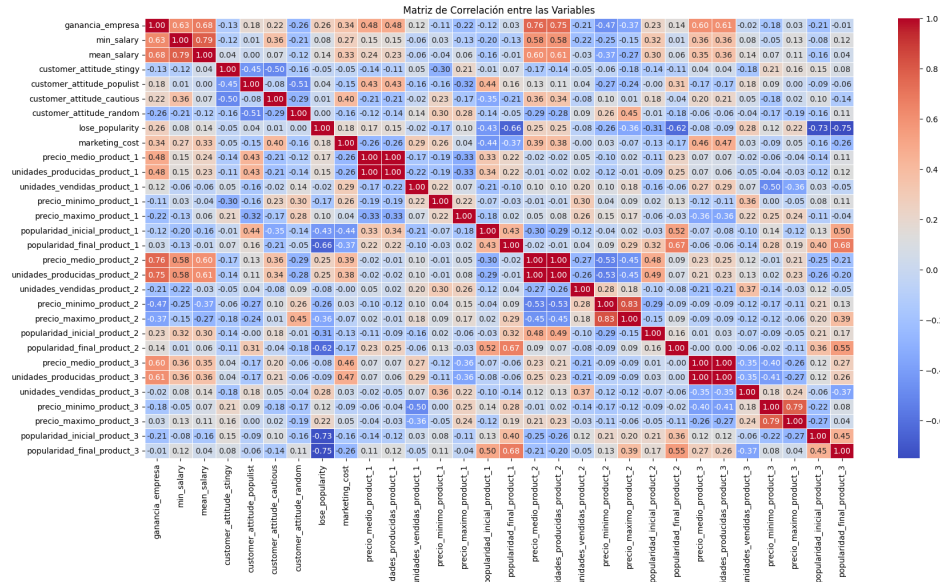
	popularidad_inicial_product_2	popularidad_final_product_2 \
count	30.000000	30.000000
mean	75.300000	27.300000
std	7.857129	25.365127
min	60.000000	0.000000
25%	69.250000	0.000000
50%	75.000000	22.500000
75%	81.750000	44.250000
max	88.000000	82.000000

	precio_medio_product_3	unidades_producidas_product_3 \
count	30.000000	30.000000
mean	196.600000	213.966667
std	197.265199	196.835566
min	0.000000	20.000000
25%	54.750000	70.500000
50%	129.000000	149.500000
75%	253.750000	275.750000
max	844.000000	839.000000

	unidades_vendidas_product_3	precio_minimo_product_3 \
count	30.000000	30.000000
mean	57.100000	135.727920
std	14.730569	49.746202
min	27.000000	63.000000
25%	47.000000	104.000000
50%	63.000000	117.750000
75%	63.000000	158.205556
max	80.000000	250.300000

	precio_maximo_product_3	popularidad_inicial_product_3 \
count	30.000000	30.000000
mean	201.166667	73.433333
std	76.098431	9.814428
min	63.000000	60.000000
25%	136.750000	65.250000
50%	174.000000	71.000000
75%	271.500000	80.750000
max	334.000000	95.000000

	popularidad_final_product_3
count	30.000000
mean	26.900000
std	27.889375
min	0.000000
25%	0.000000
50%	19.000000
75%	50.000000
max	86.000000



El resultado del coeficiente de correlación ha sido 0.6818, lo que indica una relación positiva moderadamente fuerte entre ambas variables. Esto sugiere que a medida que el salario medio aumenta, también tienden a aumentar las ganancias de la empresa. Además, el valor de $p = 0.000$ indica que la correlación es estadísticamente significativa, lo que significa que la probabilidad de que esta relación ocurra por azar es prácticamente nula ($p < 0.05$). En resumen, los resultados respaldan nuestra hipótesis de que los salarios más altos se asocian con mayores ganancias para la empresa.

5.1.2. Impacto del precio del marketing medio de los clientes en las ganancias de la empresa

Estamos evaluando la hipótesis de que el precio del marketing influye en las ganancias de la empresa bajo el supuesto de que los costos en marketing afectan la visibilidad y popularidad de los productos, lo que potencialmente podría influir en las ventas y, por ende, en las ganancias. Para explorar esta relación, utilizamos el coeficiente de correlación de Pearson.

El resultado del coeficiente de correlación ha sido 0.3446, lo que indica una **relación positiva débil** entre ambas variables. Esto sugiere que, aunque hay una tendencia de que mayores costos en marketing se asocien con mayores ganancias, la relación no es fuerte. Además, el p-valor de 0.0622 está cerca pero ligeramente por encima del umbral típico de significancia ($p < 0.05$), lo que significa que no podemos concluir con suficiente confianza estadística que exista una correlación significativa. En resumen, aunque parece haber una leve relación positiva entre el costo de marketing y las ganancias, los resultados no son lo suficientemente significativos como para respaldar la hipótesis de una influencia clara del marketing en las ganancias.

5.2. Correlación entre variables

La matriz de correlación es una herramienta fundamental para comprender las relaciones entre las variables de un conjunto de datos. Nos permite identificar de manera rápida y visual si existen asociaciones lineales fuertes, débiles o inexistentes entre diferentes variables. En el contexto de un análisis de mercado, por ejemplo, puede revelar cómo factores como el salario medio, el costo del marketing o las actitudes de los clientes influyen en las ganancias de la empresa, los precios o las ventas. Además, la matriz de correlación puede ayudarnos a detectar multicolinealidad (cuando varias variables están altamente correlacionadas entre sí), lo cual es clave si queremos construir modelos predictivos robustos y confiables. Con esta información, es posible tomar decisiones informadas sobre qué variables merecen un análisis más detallado o cuál podría tener un impacto directo en los resultados del negocio.

Análisis de la matriz de correlación: Variables fuertemente correlacionadas: Por ejemplo, lose popularity y popularidad final del producto 3 tienen un coeficiente de correlación de -0.75, lo que indica

OLS Regression Results						
Dep. Variable:	ganancia_empresa		R-squared:	0.573		
Model:	OLS		Adj. R-squared:	0.437		
Method:	Least Squares		F-statistic:	4.216		
Date:	Wed, 09 Oct 2024		Prob (F-statistic):	0.00436		
Time:	04:24:44		Log Likelihood:	-356.51		
No. Observations:	30		AIC:	729.0		
Df Residuals:	22		BIC:	740.2		
Df Model:	7					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	-6.329e+04	3.29e+04	-1.921	0.068	-1.32e+05	5033.107
min_salary	19.3277	33.178	0.583	0.566	-49.479	88.134
mean_salary	61.4246	32.492	1.890	0.072	-5.959	128.808
customer_attitude_stingy	-3.816e+04	3.93e+04	-0.970	0.343	-1.2e+05	4.34e+04
customer_attitude_populist	2.509e+04	4.17e+04	0.601	0.554	-6.15e+04	1.12e+05
customer_attitude_random	-5.384e+04	4.41e+04	-1.204	0.241	-1.44e+05	3.83e+04
customer_attitude_cautious	2824.4850	6.22e+04	0.045	0.964	-1.26e+05	1.32e+05
lose_popularity	1839.0885	1654.402	1.112	0.278	-1591.932	5270.109
marketing_cost	283.5534	528.392	0.537	0.597	-812.265	1379.372
=====						
Omnibus:	7.865	Durbin-Watson:	2.470			
Prob(Omnibus):	0.020	Jarque-Bera (JB):	6.209			
Skew:	-0.898	Prob(JB):	0.0448			
Kurtosis:	4.320	Cond. No.	1.23e+19			

una fuerte correlación negativa. Esto es evidente ya que mientras mayor sea la pérdida de popularidad mensual, los productos tenderán a bajar su popularidad en todo el ciclo. Variables con correlación moderada: Por ejemplo, precio máximo del producto 2 y unidades vendidas del producto 2 tienen un coeficiente de correlación de 0.18. Esto indica que a medida que mientras mayor sea las unidades vendidas del producto el precio de los productos tiende a aumentar.

5.3. Predicción a partir de los datos

5.3.1. Predicción de la ganancia de la empresa

Predecir las ganancias de una empresa utilizando variables como el salario mínimo, el salario medio y las actitudes de los clientes es crucial para la toma de decisiones estratégicas. Estas variables pueden influir en el comportamiento de compra de los consumidores y, por ende, en el rendimiento financiero de la empresa. Al entender cómo factores como el costo del marketing y la popularidad de los productos afectan las ganancias, se puede optimizar las estrategias de precios y marketing. Esto permite a la empresa anticiparse a cambios en el mercado, maximizar sus ingresos y mejorar su competitividad. Para esto, se utilizará una regresión lineal clásica.

Interpretación de los resultados:

El modelo de regresión lineal ajustado para predecir las ganancias de la empresa presenta un valor de R-squared de 0.573, lo que indica que aproximadamente el 57.3 % de la variabilidad en las ganancias puede ser explicada por las variables independientes incluidas en el modelo. Este es un resultado razonable, pero sugiere que todavía hay un 42.7 % de variación en las ganancias que no se explica, lo que puede estar influenciado por otros factores no incluidos en el análisis.

El valor de F (4.216) y su p-valor asociado (0.00436) indican que el modelo en su conjunto es significativo, lo que sugiere que al menos una de las variables predictivas tiene un efecto sobre las ganancias. Sin embargo, al observar los coeficientes de cada variable, notamos que muchos de ellos tienen valores p que superan el umbral de significancia comúnmente aceptado de 0.05, lo que sugiere que no son estadísticamente significativos en este modelo.

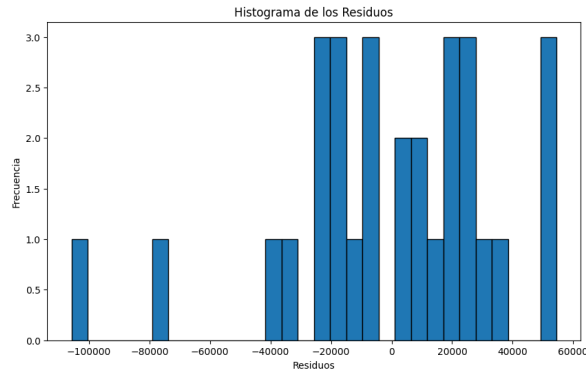
Coefficiente de min_salary: 19.33 ($p = 0.566$) indica que por cada unidad de aumento en el salario mínimo, se espera que las ganancias aumenten en aproximadamente 19.33 unidades, aunque este efecto no es significativo.

Coefficiente de mean_salary: 61.42 ($p = 0.072$) sugiere un impacto positivo en las ganancias, con un p-valor cercano al umbral de significancia. Esto implica que el salario medio podría estar relacionado con un aumento en las ganancias, aunque se requiere más investigación para confirmarlo.

Aunque el modelo de regresión muestra un cierto grado de explicación de las ganancias de la empresa, muchas de las variables incluidas no son estadísticamente significativas. Esto puede sugerir que se debe explorar un modelo más complejo, considerar variables adicionales o analizar interacciones entre variables. Las predicciones realizadas por el modelo indican que las ganancias reales varían, lo que puede ser útil para la toma de decisiones, pero también resaltan la necesidad de un análisis más profundo para identificar los factores que realmente impactan las ganancias.

Cálculo del error cuadrático medio:

Los resultados del modelo de regresión lineal muestran un Error Cuadrático Medio (ECM) de 1,228,869,515.28, lo que indica que, en promedio, las predicciones de las ganancias de la empresa se desvían de los va-



lores reales en esta magnitud. Este alto valor de ECM sugiere que hay una considerable variabilidad en las predicciones del modelo, lo que puede indicar que el modelo no está capturando adecuadamente las relaciones entre las variables. Por otro lado, el Coeficiente de Determinación (R^2) es 0.57, lo que significa que aproximadamente el 57% de la variabilidad en las ganancias puede ser explicada por las variables incluidas en el modelo. Si bien esto indica que el modelo tiene cierta capacidad para explicar las ganancias, también sugiere que hay otros factores no considerados que podrían influir en los resultados. En conjunto, estos resultados indican que el modelo tiene limitaciones y que se podrían explorar variables adicionales o diferentes enfoques para mejorar la precisión de las predicciones.

El modelo no nos brinda los mejores resultados, esto es sin duda debido a la forma de generar los mismos a través de simulaciones aleatorias.

Analizaremos los residuos para verificar los supuestos del modelo. Esto nos ayudará a entender si nuestro modelo es adecuado para los datos.

La media de los residuos es cero y la suma de los residuos es cero

Media de los residuos: $-2.2361443067590396 \times 10^{-10}$. Esto indica que, en promedio, los residuos (es decir, los errores del modelo) son cero. En otras palabras, el modelo predice los valores observados con precisión en promedio.

Suma de los residuos: $-6.708432920277119 \times 10^{-9}$. La suma de los residuos igual a cero es un resultado matemático esperado debido a la definición de los residuos como las diferencias entre los valores observados y los valores predichos. Dado que cada residuo es la diferencia entre un valor observado y un valor predicho, la suma de todos los residuos será cero. Este resultado no tiene un significado directo en términos de la bondad del ajuste del modelo, sino que es simplemente una consecuencia de cómo se calculan los residuos.

Distribución de los errores

$D = 0.5333333333333333$: Este es el valor de la estadística D, que mide la máxima diferencia entre la función de distribución acumulativa (CDF) de los residuos y la CDF de una distribución normal teórica. Un valor de D más cercano a 0 indica que los residuos se distribuyen más cerca de una distribución normal. Sin embargo, un valor de D por sí solo no es suficiente para determinar si los residuos son normalmente distribuidos, ya que la distribución de los residuos puede variar significativamente dependiendo del tamaño de la muestra y la naturaleza de los datos.

$p\text{-value} = 1.6935887182599826 \times 10^{-8}$: Este es el valor p asociado con la estadística D. El valor p es la probabilidad de obtener una estadística D tan grande o más grande que la observada, asumiendo que los residuos se distribuyen normalmente. Un valor p muy pequeño indica que es muy improbable obtener una estadística D tan grande o más grande que la observada si los residuos se distribuyen normalmente. Por lo tanto, un valor p tan pequeño sugiere que los residuos no se distribuyen normalmente.

Dado que el valor p es mucho menor que el nivel de significancia comúnmente utilizado (por ejemplo, 0.05), rechazamos la hipótesis nula de que los residuos se distribuyen normalmente. Esto significa que los residuos del modelo no parecen estar normalmente distribuidos, lo cual es una preocupación en el contexto de la regresión lineal, ya que la suposición de normalidad de los residuos es fundamental para la validez de muchos de los resultados del modelo, incluyendo las pruebas de significancia de los coeficientes y la interpretación de los intervalos de confianza para los coeficientes.

Independencia de los residuos La estadística de Durbin-Watson se utiliza para probar la

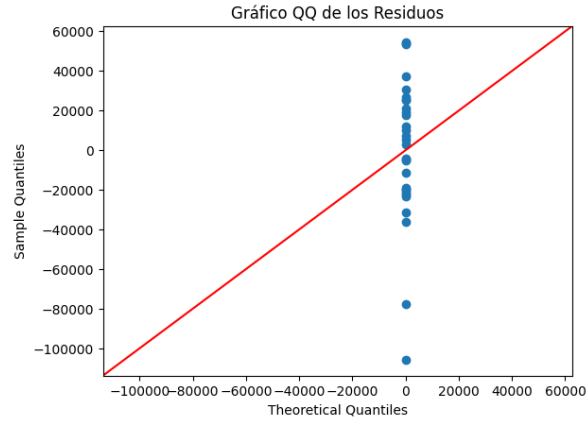


Figura 1: Gráfico QQ

ausencia de autocorrelación en los residuos de un modelo de regresión. Dado que el resultado es 2.4697754701730843 (un valor cercano a 2), indica que no hay autocorrelación en los residuos del modelo. Esto es una buena señal, ya que la ausencia de autocorrelación en los residuos es una suposición clave en la regresión lineal y puede mejorar la precisión de las estimaciones de los coeficientes del modelo.

Supuesto de homocedasticidad

Los resultados obtenidos del test de Breusch-Pagan se utilizan para determinar si hay heterocedasticidad en los residuos de un modelo de regresión. La heterocedasticidad se refiere a la variación no constante de la varianza de los errores a través de los niveles de las variables explicativas. Aquí está la interpretación de cada resultado:

Lagrange multiplier statistic: 10.746313268380169. Este valor es la estadística de prueba del test de Breusch-Pagan. En el contexto de este test, una estadística de prueba alta sugiere que hay heterocedasticidad en los residuos del modelo.

p-value: 0.2164918633088121, Este valor es la probabilidad de obtener una estadística de prueba al menos tan extrema como la observada, asumiendo que la hipótesis nula de homocedasticidad (varianza constante de los errores) es verdadera. Un valor p-value muy pequeño (menor que 0.05) indica que es muy poco probable obtener una estadística de prueba tan extrema como la observada si la hipótesis nula es verdadera, lo que lleva a rechazar la hipótesis nula y concluir que hay heterocedasticidad en los residuos del modelo. En nuestro caso dado que el p-value es mayor que 0.05 no tenemos evidencias suficientes para rechazar la hipótesis nula de que hay homocedasticidad en los residuos del modelo.

f-value: 1.7541641705140416. Este valor es el estadístico F del test de Breusch-Pagan, que es una medida de la relación entre la varianza de los residuos y la varianza de los residuos ajustados por los predictores del modelo. Un valor F alto sugiere que hay heterocedasticidad.

f p-value: 0.1480115295751364. Este valor es la probabilidad de obtener un estadístico F al menos tan extrema como el observado, asumiendo que la hipótesis nula de homocedasticidad es verdadera. Al igual que con el valor p, un valor p-value muy pequeño indica que es muy poco probable obtener un estadístico F tan extrema como el observado si la hipótesis nula es verdadera, lo que lleva a rechazar la hipótesis nula y concluir que hay heterocedasticidad en los residuos del modelo. En nuestro caso dado que el f p-value es mayor que 0.05 no tenemos evidencias suficientes para rechazar la hipótesis nula de que hay homocedasticidad en los residuos del modelo.

En resumen, los resultados indican que no hay evidencia suficiente para rechazar la hipótesis nula de homocedasticidad. Por lo que asumimos que el modelo cumple el supuesto de homocedasticidad

6. Conclusiones

Las simulaciones basadas en agentes ofrecen una poderosa herramienta para analizar mercados complejos, permitiendo observar cómo las decisiones individuales de empresas, clientes y suministradores generan dinámicas emergentes. Las empresas, utilizando la arquitectura BDI (Beliefs, Desires, Intentions) junto con lógica difusa y algoritmos genéticos, pueden adaptar sus estrategias de produc-

ción, precios y marketing, optimizando sus recursos para maximizar beneficios. Los consumidores, por su parte, muestran diversos patrones de comportamiento que influyen en la demanda, mientras los suministradores negocian eficientemente los insumos necesarios para mantener la estabilidad del mercado.

Este enfoque integral permite modelar interacciones reales del mercado de manera flexible y precisa, proporcionando una plataforma para estudiar la evolución de precios, competencia y estrategias comerciales. La integración de inteligencia artificial ha sido clave para mejorar la precisión de las simulaciones, facilitando la exploración de escenarios futuros y permitiendo una mejor toma de decisiones frente a entornos cambiantes.