International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

Modeling Retail Store Location Selection using Agent-Based Simulation: A Data-Driven Decision Support Approach

Damla DEMİR
Applied Science Faculty – Management Information Systems Department
Altınbaş University
Istanbul, Turkey

Abstract - This study develops a simulation-based model. Various marketing techniques were used to support this model's adaptation to current competitive conditions and its ability to determine the most effective strategic positioning for reaching customers. The model was trained considering the location and frequency of competing stores, as well as potential negative external factors (such as low visibility in side streets or insufficient customer traffic).

The dataset used for training the model consists entirely of synthetic data. The model's structure was designed based on the assumptions of various hypotheses developed for the application. An evaluation of the corresponding simulation scenario was conducted for each hypothesis.

This study hopes that this proposed model, created using the NetLogo program, will benefit academics, researchers, industry professionals, and anyone interested in this topic.

Keywords – simulation, agent-based model, marketing

I. INTRODUCTION

Retail is a location-oriented industry. Leading retailers utilize map-based analytical tools to determine growth strategies, analyze customer preferences, and effectively manage inventory and supply networks. The retail sector is a dynamic economic environment characterized by intense competition and rapidly changing consumer behavior, making strategic decision-making a critical factor for business success. Among these decisions, selecting the most appropriate physical store location is one of the most important. A suitable store location can directly improve sales performance by increasing customer accessibility, whereas an inappropriate location may lead to significant financial losses and competitive disadvantages. Therefore, location selection in the retail sector should not be regarded merely as a geographical choice but as a multidimensional decision-making problem that requires the simultaneous evaluation of demand characteristics, competitive intensity, customer behavior, and environmental factors.

Traditional approaches to location selection are generally based on statistical analyses, optimization techniques, or Geographic Information Systems (GIS). Although these methods provide valuable insights, they are often insufficient to fully represent the complex interactions and dynamic nature of real-world customer behavior. In particular, the relationships among customer mobility, competitor locations, and environmental characteristics exhibit nonlinear patterns that are difficult to capture using conventional analytical techniques. In this context, simulation-based approaches have emerged as an effective research method for understanding the behavior of complex systems and evaluating the outcomes of different scenarios.

Simulation models provide simplified yet behaviorally meaningful representations of real-world systems, enabling researchers to observe system performance under various assumptions and experimental conditions. Among these approaches, Agent-Based Modeling (ABM) offers a suitable framework for analyzing the interactions among individual actors and examining how these interactions generate system-level outcomes. This modeling approach is particularly appropriate for retail location selection problems, where customer movements, competitor locations, and environmental factors interact simultaneously. NetLogo, one of the most widely used platforms for developing agent-based simulation models, allows researchers to visualize spatial and behavioral systems while providing an interactive environment for analyzing complex relationships.

The primary objective of this thesis is to develop a simulation-based model for investigating the retail store location selection problem. Within the scope of the study, a simulation environment based on synthetic data is constructed to evaluate potential store locations by considering customer spatial distribution, competitor locations, environmental characteristics (such as locations on main streets or side streets), and other strategic factors. The developed model enables the examination of multiple scenarios, and the performance of the system is evaluated according to the proposed research hypotheses.

Another objective of this study is to examine the extent to which the findings obtained from the simulation model are consistent with real-world conditions and to provide an analytical perspective for decision-makers involved in retail location planning. The proposed model aims to improve the

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

understanding of the relationships among customer density, competitive intensity, and spatial factors, thereby supporting the strategic planning processes of retail businesses.

In conclusion, this study addresses the retail location selection problem through a simulation-based approach with the aim of contributing to the existing literature from a methodological perspective while proposing a practical decision-support model. By employing an agent-based simulation model developed using synthetic data, the study provides an alternative analytical framework for evaluating different strategic scenarios and supports more informed decision-making in retail location planning.

II. LITERATURE REVİEW

This chapter provides a comprehensive review of the literature that establishes the theoretical and methodological foundation of the thesis, which focuses on simulation-based retail location selection. The literature review is structured around three main pillars. First, the strategic importance of location selection in the retail sector and the spatial dynamics influencing this decision are discussed. Second, the theoretical evolution of location analysis is examined, ranging from classical spatial theories to contemporary approaches to spatial competition. Finally, the chapter explores simulation techniques and Agent-Based Modeling (ABM), which have emerged as important methodological tools for representing the complex, dynamic, and multi-actor nature of retail systems.

A. Retail Location Selection and Its Importance

The retail sector is a dynamic field of economic activity that establishes a direct connection between producers and final consumers, shapes household consumption patterns at the microeconomic level, and serves as one of the major drivers of economic growth at the macroeconomic level.[1]. By enabling consumers to access goods and services in a convenient, timely, and cost-effective manner, retail businesses operate in an increasingly competitive market environment shaped by globalization and digital transformation. Under such competitive conditions, maintaining financial sustainability and achieving a competitive advantage largely depend on the effectiveness of long-term strategic decisions made by management. Among these decisions, selecting an appropriate store location is one of the most critical factors influencing business success.[2].

For retail businesses, location selection cannot be regarded merely as the acquisition of commercial property or the choice of a physical site. Instead, it represents a complex multi-criteria decision-making problem requiring the simultaneous evaluation of numerous heterogeneous factors, including local demographics, customer accessibility, market potential, competitive intensity, transportation networks, rental and investment costs, and future urban development projections.[3]. An appropriate location can maximize store visibility and customer attraction, thereby improving sales performance and enhancing brand value. Conversely, an unsuitable location selected without comprehensive analysis may weaken the connection between the business and its target customers, reduce operational efficiency, and ultimately lead to substantial financial losses or even business failure. Unlike pricing strategies, promotional activities, or product assortment decisions, which can be modified relatively quickly, location decisions involve considerable sunk costs and are therefore extremely difficult to reverse in the short term.[4].

The influence of retail location on business performance and organizational survival has been extensively investigated from multidisciplinary perspectives. Early studies in retail geography demonstrated that store location is one of the primary determinants of consumer spatial behavior.[5]. When making shopping decisions, consumers evaluate factors such as accessibility, travel time, transportation costs, parking availability, and the attractiveness of the surrounding environment, either consciously or intuitively. Consequently, retail businesses must analyze not only current market conditions but also the daily mobility patterns, spatial perceptions, and socioeconomic shopping preferences of their target customers when determining store locations.

Another critical factor affecting retail location decisions is the competitive structure of the market. The spatial distribution and density of competing businesses within a trade area largely determine market saturation.[6]. In central business districts or popular shopping centers characterized by intense competition, retailers benefit from existing customer traffic and agglomeration economies while simultaneously facing aggressive price competition that may reduce profit margins. Conversely, peripheral areas with limited competition may provide opportunities to dominate the local market; however, insufficient customer demand and higher logistical costs may prevent stores from achieving optimal sales performance. Therefore, retailers should evaluate not only existing competitive conditions but also potential competitive responses when making expansion or store-opening decisions.

The physical and micro-environmental characteristics of a location also play an essential role in store performance. Factors such as storefront width, sign visibility, pedestrian and vehicle traffic patterns, and surrounding infrastructure directly influence customer conversion rates.[7]. Stores located along major roads and commercial streets generally benefit from higher levels of spontaneous customer traffic, whereas businesses situated on secondary streets often depend on strong customer loyalty or destination-oriented shopping behavior to remain competitive. In addition, access to public transportation, walkability, accessibility for individuals with disabilities, and the complementary nature of neighboring businesses all contribute significantly to store performance at the local level.

Historically, retail location decisions relied primarily on managerial intuition, previous experience, and heuristic judgment. However, the increasing saturation of retail markets, shrinking profit margins, and the growing availability of large-scale data have encouraged businesses to replace intuition-based approaches with more analytical and evidence-based decision support systems.[8]. Accordingly, both researchers and practitioners have developed various analytical approaches, including Geographic Information Systems (GIS), multi-criteria decision-making techniques such as the Analytic

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

Hierarchy Process (AHP) and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), as well as spatial statistical models to improve location selection decisions.

Recent advances in data analytics, artificial intelligence, and computer simulation have further transformed retail location analysis. These developments enable businesses not only to analyze historical spatial data but also to evaluate market dynamics, population mobility, and competitive scenarios within simulated environments before making investment decisions. Retail systems exhibit the characteristics of complex adaptive systems, involving nonlinear interactions, feedback mechanisms, and substantial uncertainty. Customer decisions, competitor strategies, and unexpected market changes continuously reshape competitive conditions. Consequently, traditional analytical techniques, such as regression models or simple evaluation checklists, are often insufficient to capture these dynamic interactions. To address this methodological limitation, simulation-based approaches, particularly Agent-Based Modeling (ABM), have emerged as powerful methodologies capable of representing the temporal evolution of retail systems and the spatial dynamics of market competition with a high degree of realism.

B. From Classical Approaches to Spatial Competition in Location Selection

The literature on retail location selection has a long-standing and evolving theoretical foundation that seeks to explain the role of space in shaping economic and social activities. Over time, this perspective has evolved from mechanistic approaches, in which space was regarded merely as a transportation cost or physical distance, to more dynamic frameworks that incorporate consumer behavior, psychological preferences, and the strategic interactions among competing firms.

C.Spatial Organization and Central Place Theory

The origins of location analysis can be traced back to the work of Von Thünen who explained agricultural land-use patterns based on transportation costs and distance from a central market.[9]. However, the first comprehensive theory explaining the spatial hierarchy of retail and service activities as well as the organization of urban settlements was Central Place Theory, developed by the German geographer Walter Christaller (1933)[10]. To explain why goods and services are distributed in particular spatial patterns, Christaller introduced two fundamental concepts into the literature: threshold population and range.

The threshold population refers to the minimum market size or customer base required for a retail business to operate profitably by covering its fixed and variable costs. The range represents the maximum distance or travel cost that consumers are willing to incur to purchase a particular product or service. High-value or infrequently purchased goods, such as luxury jewelry or automobiles, generally require a larger threshold population and have a wider range. In contrast, frequently purchased convenience goods, such as bread or basic groceries, have relatively small threshold populations and shorter ranges. Based on the interaction between these two concepts, Christaller demonstrated that retail centers would organize themselves hierarchically into hexagonal market areas, ensuring complete spatial coverage across a homogeneous landscape. Lösch (1954) later extended this framework by introducing greater flexibility into market structures and allowing for more complex spatial arrangements of economic regions. Despite their significant contributions, these theories rely on assumptions such as homogeneous population distribution and perfectly rational consumers, limiting their ability to explain the complexity and asymmetry of real-world retail markets.

D. Probabilistic Attraction and Consumer Behavior

The geometric and deterministic framework proposed by Christaller and Lösch was criticized for assuming that consumers would always choose the nearest shopping destination. This assumption overlooked the multidimensional nature of consumer decision-making and the psychological factors influencing shopping behavior. Recognizing that the attractiveness of retail centers extends beyond physical distance, Reilly (1931) introduced the Law of Retail Gravitation, which explained how two competing urban centers divide the trade area between them according to their population sizes and the squared distances separating them from surrounding communities. Nevertheless, one of the most influential methodological advances in retail location analysis was achieved through David Huff's (1964) Probabilistic Gravity Model.

Drawing inspiration from Newton's law of universal gravitation, Huff adapted the concept of gravitational attraction to consumer choice behavior. In this model, the probability that a consumer located at a specific origin selects a particular retail store is expressed as a probabilistic function. According to the model, the attractiveness of a retail store is directly related to its positive characteristics, such as store size, product assortment, and brand image, while being inversely related to the travel distance or travel time required to reach the store, raised to a distance-decay parameter. Consequently, larger and more attractive stores are expected to attract customers from broader geographic areas, whereas increasing travel costs reduce the likelihood of store selection.

Mathematically, the Huff Model can be expressed as follows:

$$P_{ij} = \frac{\frac{S_j}{D_{ij}^{\lambda}}}{\sum_{k=1}^{n} \frac{S_k}{D_{ik}^{\lambda}}}$$

In this equation, P_{ij} denotes the probability that a consumer located at point *i* chooses store *j*; S_j represents the size or attractiveness of store *j*; D_{ij} denotes the distance between location *i* and store *j*; and λ is the distance-decay (friction) parameter. The Huff Model constitutes one of the principal theoretical foundations of the customer decision-making mechanism adopted in the simulation-based framework developed in this thesis. Rather than assuming that consumers always select the nearest store, the model represents store choice as a probabilistic process determined by the balance between store attractiveness and the resistance imposed by travel distance. This probabilistic formulation provides a more realistic representation of consumer behavior in retail

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

environments. Subsequently, Lakshmanan and Hansen (1965) [11] extended the Huff Model to market share analysis, thereby broadening its application in retail planning and location analysis.

E. Spatial Competition and Strategic Agglomeration

Another fundamental factor influencing customer traffic and retail attractiveness is the presence of competitors and their spatial strategies within the market. Contrary to classical economic theories, the tendency of competing firms to locate in close proximity rather than dispersing geographically to divide the market evenly is explained by Harold Hotelling's [12] Principle of Minimum Differentiation.

Using a game-theoretic framework, later interpreted as a Nash equilibrium, Hotelling analyzed the competition between two ice cream vendors selling homogeneous products along a linear market represented by a beach. If the vendors initially position themselves at one-quarter and three-quarters of the beach, transportation costs are minimized and overall social welfare is maximized. However, each vendor has an incentive to move toward the center in order to capture a larger share of customers from the competitor's market area while retaining its own customer base. As a result, the stable equilibrium occurs when both competitors locate side by side at the midpoint of the market, leading to geographical clustering.

This principle explains the spatial concentration frequently observed in retail industries, where fast-food restaurants, bank branches, automobile dealerships, and similar businesses tend to cluster along the same streets or commercial districts.[13], further extended Hotelling's model to multidimensional spatial settings and multiple-firm environments, demonstrating that spatial agglomeration may also benefit consumers engaged in multi-purpose shopping by creating attractive commercial destinations.

F. Synthesis of Classical Spatial Theories

The historical and theoretical perspectives discussed above demonstrate that each theory contributes to understanding a different dimension of retail location selection. Table 1. provides a comparative summary of these theories, highlighting their underlying assumptions, contributions to retail location analysis, and their role in developing rule-based simulation models.

TABLE I. COMPARATIVE SUMMARY OF CLASSICAL SPATIAL LOCATION THEORIES

Theory / Model	Key Researchers	Main Focus and Assumptions	Contribution to Retail Location Analysis	Role in Agent-Based Modeling (ABM)
Central Place Theory	Christaller (1933); Lösch (1954)	Homogeneous space, threshold population, and range.	Explains the spatial hierarchy of retail centers and the geometric structure of market areas.	Defines the initial capacity limits and spatial structure of the simulated environment (grid).
Law of Retail Gravitation	Reilly (1931)	Market allocation based on population size and distance between competing urban centers.	Provides a mathematical framework for delineating trade areas.	Determines transition probabilities between macro-level market regions for customer agents.
Probabilistic Gravity Model	Huff (1964); Lakshmanan & Hansen (1965)	Store choice is positively associated with store attractiveness and negatively associated with travel distance.	Models uncertainty and probabilistic consumer choice behavior.	Serves as the primary stochastic decision algorithm governing store selection by customer (buyer) agents.

The concept of the minimum market threshold introduced by Christaller (1933), the probabilistic consumer choice mechanism proposed by Huff (1964), and the spatial competition model developed by Hotelling (1929) together constitute the theoretical foundation of modern retail geography. The integration of these theories provides the conceptual basis for the behavioral rules and decision-making algorithms employed in contemporary computer-based Agent-Based Modeling (ABM) simulations. By combining principles of spatial hierarchy, probabilistic consumer behavior, and strategic competition, these theoretical frameworks enable simulation models to represent retail systems in a realistic and dynamic manner.

III. SIMULATION AND AGENT-BASED MODELING

Simulation is a powerful analytical approach that involves constructing mathematical and logical representations of complex real-world systems or dynamic processes that are difficult, costly, or hazardous to observe directly. By conducting experiments on these computational models over time, researchers can examine system behavior under different parameter settings and evaluate the potential outcomes of alternative scenarios [14]. Simulation is particularly valuable for analyzing complex systems whose analytical or closed-form solutions are mathematically intractable. It enables researchers and decision-makers to evaluate what-if scenarios within a risk-free virtual environment, eliminating the need for costly or disruptive interventions in real-world systems. By accurately representing both micro- and macro-level system

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

behaviors, simulation models provide meaningful insights into complex interactions and dynamic processes, thereby supporting strategic decision-making while reducing uncertainty.[15].

Simulation techniques are particularly well suited for analyzing systems characterized by high levels of uncertainty, stochastic processes, dynamic feedback mechanisms, and nonlinear relationships. Traditional econometric and analytical approaches often simplify the complexity of real-world systems in order to obtain mathematically tractable solutions, which may reduce the realism and predictive capability of the resulting models. In contrast, simulation models create behaviorally meaningful digital representations of real-world systems, enabling researchers to investigate system dynamics in greater depth and evaluate alternative policies or investment strategies before implementation.

Simulation methodologies are commonly classified into three major paradigms: Discrete Event Simulation (DES), System Dynamics (SD), and Agent-Based Modeling (ABM). Discrete Event Simulation models systems in which state changes occur at specific points in time and is widely applied to manufacturing, logistics, and operational processes such as production lines. System Dynamics adopts a macro-level perspective by representing systems through stocks, flows, and feedback loops, allowing researchers to investigate long-term behavioral trends. In contrast, Agent-Based Modeling (ABM) adopts a bottom-up perspective, focusing on how macro-level system behavior emerges from interactions among autonomous individuals and between those individuals and their environment.[16].

Agent-Based Modeling has become one of the most influential simulation methodologies for representing complex adaptive systems. Within this framework, a system is modeled as a dynamic population of heterogeneous and autonomous entities capable of making independent decisions. These entities, referred to as agents, possess their own attributes, memory, bounded rationality, and behavior governed by predefined decision rules, typically represented as if–then rules.[17]. One of the defining characteristics of ABM is its ability to demonstrate how relatively simple local interactions among individual agents can generate complex, nonlinear, and often unexpected patterns at the system level.[18]. Furthermore, agents are adaptive rather than static; they may update their strategies, learn from previous experiences, and modify their behavior in response to environmental feedback.

The conceptual architecture of an agent-based model consists of three fundamental components: agents, the environment, and interaction rules.

- Agents: Autonomous entities capable of making independent decisions based on predefined characteristics such as income level, brand preference, purchasing power, or travel speed. Within a retail context, both consumers and retail stores may be represented as agents.
- Environment: The spatial or conceptual setting in which agents are located, move, and interact. Depending on the research objectives, the environment may consist of a two-dimensional cellular grid or a realistic geographic representation derived from Geographic Information Systems (GIS).
- Interaction Rules: Algorithmic mechanisms specifying how agents interact with one another and with their environment. These rules determine how agents communicate, exchange information, respond to environmental conditions, and make decisions such as purchasing products, changing locations, or adapting their strategies.

The continuous interactions among these three components give rise to complex system-level phenomena that are difficult to predict using conventional analytical models. Examples include spatial monopolization, customer clustering, and market collapse, all of which emerge from nonlinear interactions among individual agents.

Agent-Based Modeling has been widely applied across numerous disciplines, including sociology, artificial life, economics (particularly Agent-Based Computational Economics, ACE), urban and regional planning, and marketing science, owing to its flexibility and ability to capture human behavior.[19]. In particular, ABM has proven highly effective for simulating heterogeneous consumer preferences, brand loyalty dynamics, innovation diffusion, and spatial or price-based competition among firms, overcoming many of the limitations associated with traditional statistical approaches.[20]. Rather than assuming a homogeneous consumer population, ABM enables researchers to model each consumer individually, accounting for personalized budget constraints, behavioral preferences, and spatial limitations.

IV. METHODOLOGY

This study develops an Agent-Based Modeling (ABM) simulation framework to analyze retail store location selection and spatial competition dynamics. A synthetic dataset approach was adopted to parameterize the simulation model. In retail location analysis, obtaining empirical data such as real-time pedestrian traffic, household income distributions, and the commercial performance of competing businesses is often challenging due to high acquisition costs, commercial confidentiality, and data accessibility constraints. Furthermore, empirical datasets generally represent static observations of a specific geographic region and therefore provide limited support for examining alternative market scenarios or dynamic feedback processes. The use of a synthetic grid-based environment in this study eliminates dependence on the characteristics of a particular location, enabling the proposed decision-support framework to be adapted to different retail sectors and geographical settings. The synthetic dataset and the behavioral rules assigned to agents were not generated arbitrarily; instead, they were designed based on established theories of spatial interaction, particularly the Huff Gravity Model, and widely accepted models of consumer location choice. Consequently, the simulation provides a methodologically controlled environment in which the causal relationships among consumer behavior, travel distance, and competitive intensity can be investigated while minimizing the influence of external environmental factors.

The simulation model was implemented in NetLogo and consists of three principal components:

- Agents (Consumers): Each consumer agent possesses a set of individual attributes, including income level,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

movement speed, and preference coefficient. Consumer agents make purchasing decisions according to the satisficing principle, selecting alternatives that satisfy predefined decision criteria rather than seeking globally optimal solutions.

- Retail Stores: Stores are represented as fixed entities located at predefined coordinates. Each store is characterized by attributes such as attractiveness, price index, and service capacity.
- Environment: The simulation environment is represented by a two-dimensional synthetic grid designed to approximate the spatial characteristics of a dense metropolitan retail environment..

G. Store Selection Mechanism

The decision-making process used by consumer agents is based on the following components:

- Distance Cost: The Euclidean distance or Manhattan distance between the consumer's current location (L_i) and the store location (L_j) is calculated.
- Store Attractiveness Function: The probability of selecting a particular store depends positively on the store's attractiveness (S_j) and negatively on the travel distance (d_{ij}). The decision mechanism is based on the Huff Gravity Model:

$$P_{im} = \frac{S_m / d_{im}^{\beta}}{\sum_{j=1}^{n} (S_j / d_{ij}^{\beta})}$$

where β denotes the distance-decay parameter, reflecting the sensitivity of consumers to travel distance.

H. Simulation Process

During each simulation tick (time step), the following sequence of operations is performed:

1. Search: Consumer agents scan the surrounding environment to identify available retail stores.
2. Evaluation: Each agent calculates a utility score for the candidate stores based on travel distance and store-specific characteristics.
3. Movement and Store Selection: The consumer moves toward and selects the store with the highest utility value.
4. Feedback: Following the shopping decision, the consumer's satisfaction level is updated according to store conditions such as inventory availability and congestion level. These updated states influence subsequent decision-making processes throughout the simulation.

I.Dataset and Calibration

To enhance the validity of the proposed model, several datasets representing key characteristics of the retail environment were incorporated into the simulation framework. These include:

- Population Density: Demographic distribution representing the spatial concentration of potential consumers within the study area.
- Transportation Networks: The influence of major transportation corridors and arterial roads on retail accessibility and location attractiveness.
- Competitor Analysis: The impact of the spatial proximity of competing stores (agglomeration effect) on sales performance.

V. SIMULATION IMPLEMENTATION

The simulation framework developed in this study is based on a data-driven model that is dynamically updated through user interaction. The primary objective of the model is to observe system behavior under different parameter settings and to present these behaviors through graphical outputs that facilitate interpretation and analysis. To achieve this objective, the system was designed around two complementary components: the computational model operating in the background and the graphical user interface (GUI).

From the computational perspective, the simulation follows an event-driven architecture. User commands, such as the "go" button, initiate the main simulation loop, during which the model updates its variables and recalculates all relevant performance measures at every iteration. Maintaining the consistency of model variables throughout this process is essential for ensuring reliable simulation results. The source code was organized into modular procedures responsible for distinct tasks, including data updating, graphical visualization, and simulation control. This modular design simplifies debugging, improves code readability, and enhances the maintainability and scalability of the simulation model.

The graphical user interface provides an interactive control panel through which users can observe the behavior of the simulation in real time. Various graphical components display the outputs generated during model execution. In particular, plots such as the "Revenue Trend" chart visualize changes in key performance indicators over time. During model development, careful synchronization between the graphical interface and the underlying source code proved essential. Attempting to reference a plot that had not been defined within the interface resulted in the "no such plot" runtime error. Consequently, all graphical objects and their corresponding code references were carefully matched to ensure stable system execution.

The input dataset was preprocessed before being incorporated into the simulation. Raw data obtained from external sources were transformed into a structure compatible with the simulation environment through data cleaning, standardization of variable names, and numerical transformations. Variables such as store counts, revenue indicators, and other operational measures were normalized before being used within the model, thereby enabling meaningful comparisons among variables measured on different scales. This preprocessing stage improved both the consistency and interpretability of the simulation outputs.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

Once the simulation is executed, user commands continuously trigger the iterative update mechanism, allowing both numerical indicators and graphical visualizations to be refreshed at each simulation step. Consequently, the proposed framework functions not merely as a static analytical tool but as a dynamic decision-support system capable of illustrating the evolution of the retail environment over time. The continuous updating of graphical outputs enables users to monitor system dynamics and evaluate the effects of alternative scenarios interactively.

The proposed framework is an agent-based simulation model developed to investigate the relationship between retail location selection and store performance. The fundamental modeling approach represents each retail store as an autonomous agent whose revenue is influenced by both environmental conditions and internal operational characteristics. Through this structure, the simulation captures the multidimensional and interactive nature of real-world retail systems.

The model is organized into three principal layers. The first layer consists of patches, representing the spatial environment. Each patch is assigned an economic score, which reflects the economic attractiveness of that particular location. Initially, each patch receives a randomly generated value ranging between 80 and 140, and these values are visualized using different shades of green within the graphical interface. This visualization enables users to observe the spatial heterogeneity of economic attractiveness across the simulated environment. Consequently, the environmental layer forms the foundation of the model's retail location selection mechanism.

The second layer consists of store agents. Each store is designed to represent a real-world retail business, and a fixed number of 200 store agents are initialized within the simulation. Each store is randomly assigned to a patch, thereby ensuring that stores are distributed across different economic regions. This design enables the model to examine the influence of location selection on store performance. Each store is characterized by a set of attributes, including product variety, marketing expenditure, customer traffic, store size, employee efficiency, store age, distance to competitors, and the number of promotional campaigns. These variables represent normalized and adapted versions of the key attributes contained in the retail dataset obtained from Kaggle. Rather than being directly incorporated into the simulation, the multidimensional retail indicators contained in the original dataset were transformed into behavioral parameters that govern the decision-making processes of the agents.

The third layer represents the model's output variable, revenue. Revenue is calculated using a multifactor function that simultaneously considers both store-specific operational characteristics and the economic attractiveness of the surrounding environment. The calculate-revenue procedure constitutes the core computational mechanism of this process. Within this procedure, each explanatory variable is assigned a predefined weight reflecting its relative contribution to store performance. Variables such as marketing expenditure, customer traffic, and employee efficiency are assigned relatively higher coefficients, thereby exerting a greater influence on revenue generation. In contrast, product variety and store size receive moderate weights. Furthermore, the economic score of the location is incorporated directly into the revenue function, allowing the influence of location quality to be explicitly represented within the model. Consequently, the revenue calculation follows a weighted linear scoring approach while operating as a dynamic mechanism that is continuously updated throughout the simulation.

The operation of the model is organized around two principal procedures: setup and go. During the setup phase, the simulation environment is initialized by resetting the system, generating the economic scores associated with each patch, and randomly distributing store agents across the spatial environment. Initial values are then assigned to all store attributes, after which the first revenue calculation is performed. This stage represents the initial integration of the dataset into the simulation environment. Although the variables were inspired by the characteristics of the Kaggle retail dataset; however, all initial values were synthetically generated for simulation purposes. This approach enables the model to evaluate a wide range of alternative scenarios while maintaining consistency with the characteristics of the reference dataset.

The go procedure represents the temporal evolution of the simulation. During each simulation tick (time step), customer traffic associated with each store is subject to small stochastic fluctuations, promotional activities are updated, and revenue is recalculated based on the modified conditions. This mechanism was intentionally designed to capture the uncertainty and variability observed in real-world retail environments. For example, increases or decreases in customer traffic directly influence store revenue; however, these changes are governed by stochastic rather than deterministic processes, allowing the model to produce more realistic behavioral dynamics.

One of the most important components of the graphical user interface is the "Revenue Trend" plot, which displays the average revenue of all stores as a time series throughout the simulation. At each iteration, the calculated average-revenue value is plotted, enabling users to monitor the overall performance of the retail system rather than focusing solely on individual stores. An important implementation requirement is that the plot name defined within the graphical interface must exactly match the corresponding reference used in the source code. Otherwise, the simulation generates a "no such plot" runtime error. Consequently, synchronization between the graphical interface and the underlying code constitutes a critical design requirement for ensuring stable model execution.

The adaptation of the dataset to the simulation environment represents one of the most important stages of model development. Rather than incorporating the Kaggle retail dataset directly into the simulation, the dataset underwent feature engineering to ensure compatibility with the agent-based modeling framework. Retail variables describing sales performance, customer behavior, store characteristics, and operational metrics were transformed into behavioral parameters governing agent decisions and interactions. As a result, the dataset evolved from serving as a static analytical resource into a dynamic mechanism that drives agent behavior throughout the simulation. This transformation established a closed-loop simulation process in which data → parameters →

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

agent behavior → model outputs, allowing the system to continuously evolve through interactions among its constituent components.

In summary, the proposed model addresses the retail location selection problem within a multivariate simulation environment, enabling the simultaneous examination of the effects of both environmental and operational factors on store revenue. The primary strength of the model lies in its integration of data-driven parameterization with an agent-based simulation framework. Consequently, the model functions not only as a tool for analyzing historical data but also as a dynamic decision-support system capable of generating and evaluating alternative scenarios under varying market conditions.

To evaluate the effectiveness of the proposed model, three research hypotheses were developed and examined independently through separate simulation scenarios.

- Scenario 1 (H1): Stores located on main streets generate higher average revenue than stores located on side streets.
- Scenario 2 (H2): Stores implementing promotional campaigns attract more customers than stores without promotional activities.
- Scenario 3 (H3): Stores located in areas with high economic attractiveness generate higher revenue than stores located in areas with low economic attractiveness.

The first hypothesis investigates whether the location of a retail store on a main street or a side street influences customer movement patterns and, consequently, store performance. To evaluate this hypothesis, the necessary interface components, including control buttons and graphical displays, were implemented in NetLogo, while the underlying model was configured by defining the characteristics of the agents and the simulation environment (grid structure). Figure 3.1 illustrates the graphical user interface developed for the simulation model.

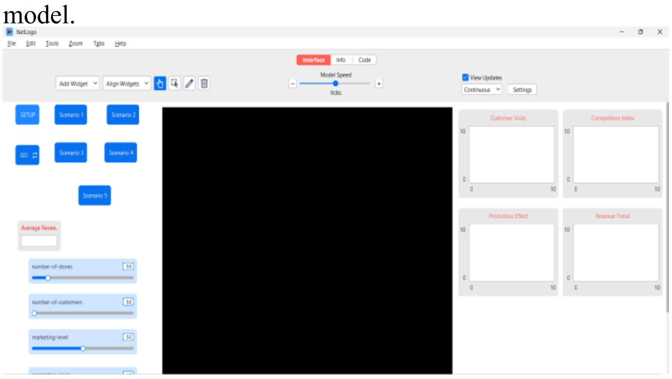

Fig. 1. Interface of simulation.

After pressing the “SETUP” button, the visualization of the simulation environment was generated as shown below.

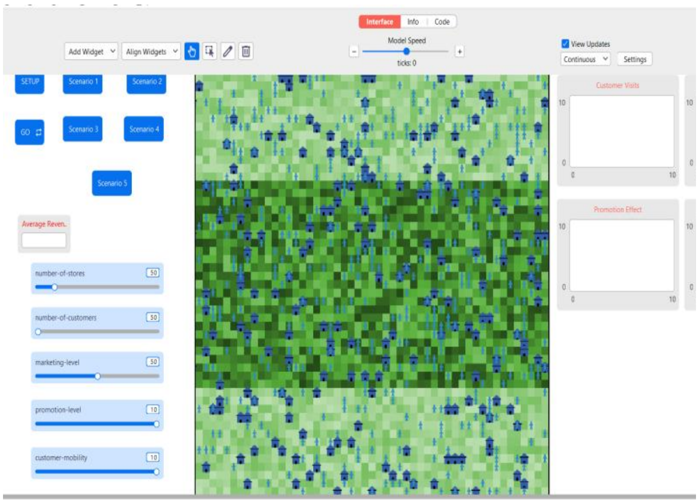

Fig. 2. Setting up the simulation environment

The middle section (the dark green area) represents the main street, while the lower and upper sections (the light green areas) represent the side streets.

Fig. 3. Location of shops, main street, side streets and customers.

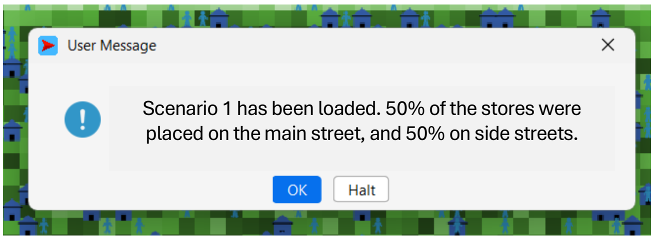

Fig. 4. Prerequisite for the simulation environment designed for hypothesis 1

After this step, we click the "Go" button. Once the button is pressed, the simulation starts running.

- Customers (represented by blue agents) begin moving randomly across the map.
- They visit stores, and each store starts generating revenue based on its real-time customer traffic.
- Stores with revenues exceeding 300 units immediately change to light green, while stores with lower revenues turn red.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

The revenue generated by the stores is displayed in the Revenue Trend graph. As the simulation progresses, two separate curves appear on the graph:

- Main Street Stores (the average revenue curve of the 50% of stores located on the main street)
- Side Street Stores (the average revenue curve of the 50% of stores located on the side streets)

After several hundred ticks (time steps), the simulation is stopped by pressing the Go button again.

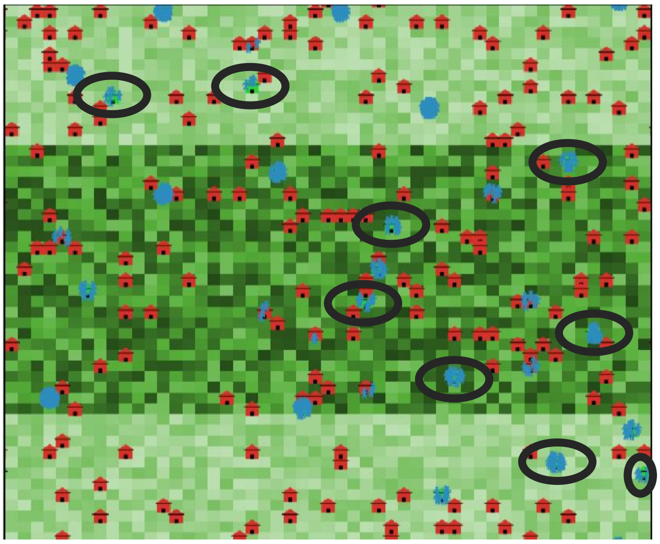

Fig. 5. The final state of the simulation environment tested as a result of hypothesis 1.

As a result of the simulation, stores that saw an increase in revenue turned green and were circled. Customer circulation increased around these green-colored stores, and significant clustering occurred.

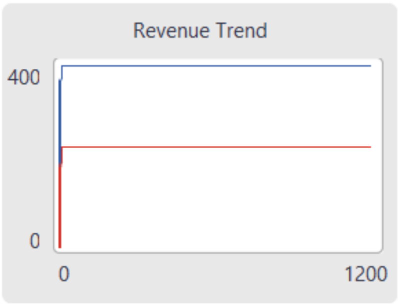

Fig. 6. Income balance of hypothesis 1

As can be clearly seen from the graph, the blue line at the top represents the main street stores. Immediately after the simulation begins, it rises rapidly and reaches a stable equilibrium at above 400 revenue units.

In contrast, the red line, which represents the side street stores, remains at only 200–230 revenue units.

The economic advantage provided by location and the higher pedestrian/customer traffic on the main streets significantly improve store performance. Even when operational variables such as marketing expenditure and employee efficiency are kept constant or assigned randomly, the revenue of main street stores can be approximately twice as high as that of side street stores. This finding provides strong agent-based simulation evidence supporting the well-known retail principle of "Location, Location, Location."

In our second hypothesis, half of the stores implemented promotional campaigns and discount strategies, while the remaining half chose not to run any promotions and instead relied solely on their existing customer traffic.

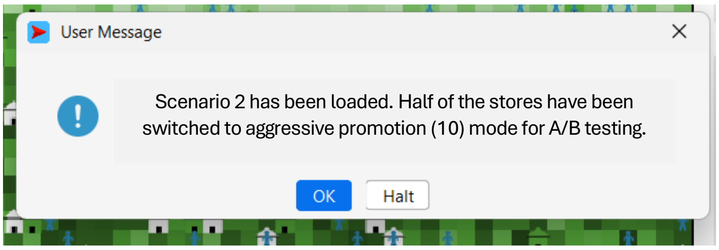

Fig. 7. Preconditions of the simulation environment for hypothesis 2

Fig. 8. Categorization of stores as white stores and green stores

In the simulation environment activated by pressing the “Go” button, it was observed that customers mostly gathered around the green stores.

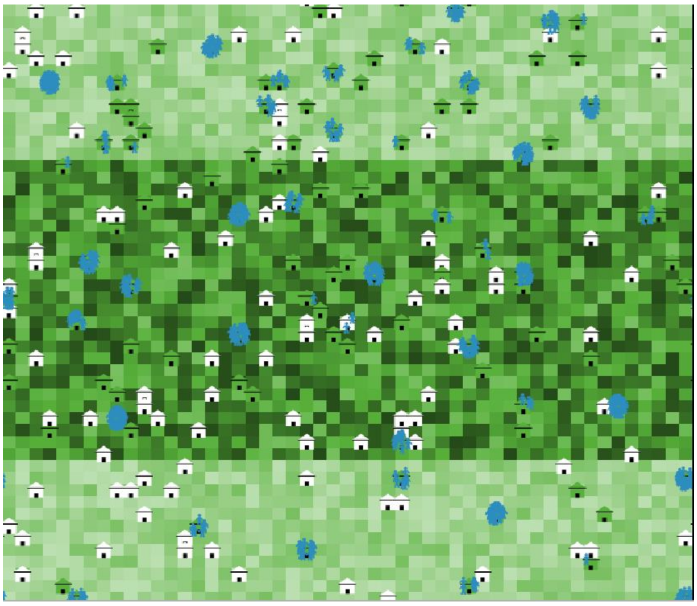

Fig. 9. Customer distributions during simulation execution

The primary objective of Scenario 2 is to create an operational A/B test among the stores.

- Group A (Green Stores): Stores implementing aggressive promotional campaigns, allocating promotional budgets, and having the highest customer attraction coefficient.

- Group B (White Stores): Passive stores that do not offer any promotions and rely solely on their organic customer traffic.

Exactly 50% of the stores are represented as green (aggressive promotion) and 50% as white (no promotion), and they are distributed uniformly across the map. Since no simulation time (ticks) has elapsed yet, the customer agents (blue agents) are waiting at random locations. This represents a neutral starting point, ensuring that the experiment is ready to test the independent variable without bias.

When we press the Go button to start the simulation, we can clearly observe that the blue customer agents stop wandering randomly and begin clustering around the green stores that offer promotions. In contrast, the areas surrounding the white stores remain almost completely empty.

The aggressive promotion strategy completely outweighs the disadvantage of distance (location) in the customers' decision-making process. Even when a non-promotional white store is located closer to them, customers prefer traveling to the more distant green stores because of the stronger financial incentive created by the promotional offers.

These results indicate that, in highly competitive retail environments, promotional activities are as powerful as location in attracting and concentrating customer traffic. Stores that do not implement promotional strategies (the white stores) rapidly lose both market share and customer traffic, placing them at a significant financial disadvantage.

In our third hypothesis, the simulation environment is divided into three income segments. The northern region represents high-income customers, the central region represents middle-income customers, and the southern region represents low-income customers.

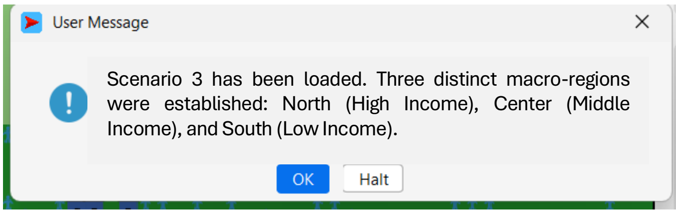

Fig. 10. Segmentation of the environment

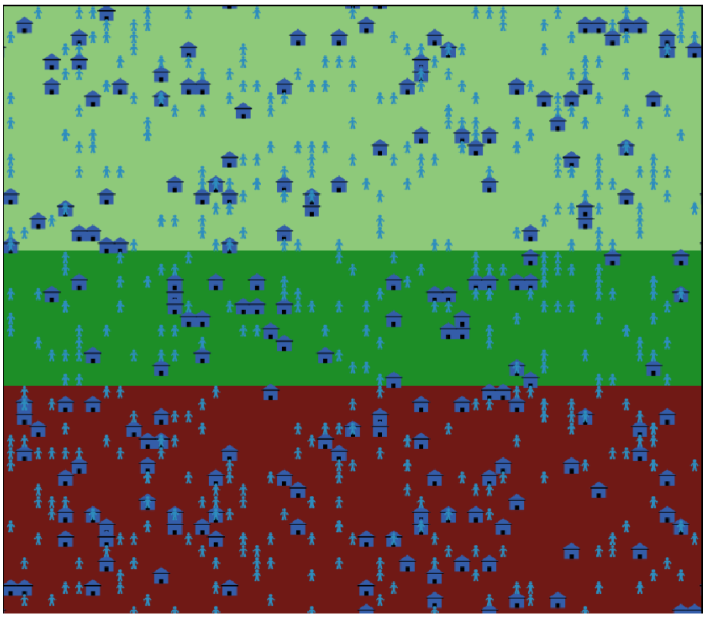

Fig. 11. Grid partitioning

Although the stores (blue buildings) appear to be distributed uniformly across the map, the background colors clearly define the boundaries of the three income zones. Since no simulation time has elapsed yet, all stores have the same initial economic status (blue), and customer agents are evenly distributed across all regions.

- Northern Region (Light Green – High Income): Represents the premium ("A+") area, where both the economic score and the location attractiveness potential are at their highest, resulting in the greatest customer spending capacity.

- Central Region (Green – Middle Income): Represents a balanced market environment characterized by moderate purchasing power and average economic conditions.

- Southern Region (Dark Brown/Red – Low Income): Represents the area with the lowest economic score and, consequently, the lowest customer spending potential.

After pressing the "Go" button, the simulation runs as follows.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

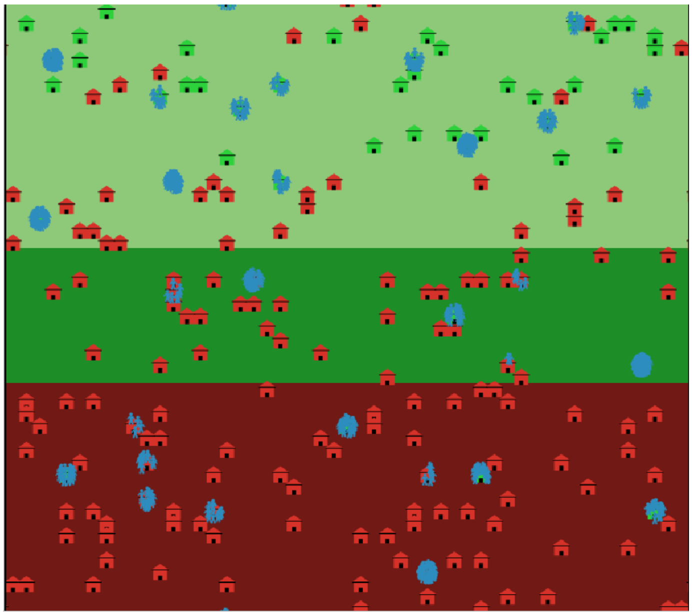

Fig. 12. Customer distributions after the simulation runs

Once the simulation begins, the changes in store colors provide a clear visualization of the underlying economic reality.

- Green Stores in the Northern Region: The stores located in the northern region rapidly turn light green, indicating that the higher purchasing power of customers in this area enables these stores to surpass the profitability threshold within a short period of time.
- Concentration of Red Stores in the Southern Region: Nearly all stores in the southern region remain red. Although the number of customers (blue agents) is comparable to that of the northern region, their lower purchasing power (economic score) is insufficient to raise store revenues above the profitability threshold.
- Transition in the Central Region: The central region represents an intermediate market, where red stores remain the majority, while some stores gradually turn green. This indicates that retailers in this area operate close to the profitability threshold, with performance fluctuating around the break-even point.

VI. RESULTS

This study addressed the dynamic and multidimensional nature of strategic retail location selection by adopting a simulation-based approach that extends beyond traditional analytical methods. The model developed in this thesis simulates the effects of dynamic factors—including customer behavior, demographic mobility, competitive intensity, and logistics networks—on the success of retail locations, thereby providing predictive and optimized outcomes for future decision-making. The findings demonstrate that, unlike static analytical models, simulation-based approaches can rapidly adapt to changing market conditions and serve as powerful decision-support tools that minimize uncertainty and reduce decision-making risks.

From an academic perspective, this thesis contributes to the literature on retail analytics and location theory by addressing existing research gaps and demonstrating the effectiveness of simulation methodologies in modeling complex market ecosystems. Beyond providing a theoretical framework, the study proposes a practical methodology capable of reducing the financial risks associated with poor location decisions in the retail industry, where investment costs are typically substantial. In this respect, the research offers an innovative reference for both industry practitioners and researchers seeking to integrate data-driven decision-making into modern retail strategies.

For future research, this study provides a scalable foundation that can be further extended through the integration of new datasets and advanced computational algorithms. Incorporating real-time digital footprint data, psychometric consumer analyses, or deep learning–based demand forecasting models into the simulation framework has the potential to further improve its predictive performance. Overall, this thesis demonstrates that retail location selection should not be regarded merely as a geographical decision, but rather as an integral component of a continuously evolving consumer and market simulation ecosystem.

VII. RECOMMENDATIONS FOR FUTURE STUDIES

First, the simulation parameters used in this study were based on synthetic data and variables derived from the existing literature. Future research could enhance the realism of the simulation by integrating real-world data sources such as anonymous location data collected from mobile devices, credit card transaction records, social media interactions, Google Maps traffic density data, and real-time information obtained from IoT sensors.

Furthermore, the decision-making mechanisms of customer agents can be improved by incorporating more sophisticated behavioral models. Factors such as price sensitivity, brand loyalty, income level, lifestyle, psychographic characteristics, and seasonal shopping patterns could be integrated into the model to provide a more accurate representation of consumer behavior.

Another promising direction is the integration of artificial intelligence techniques into the simulation model. In particular, deep learning, reinforcement learning, and evolutionary optimization algorithms could enable stores to evolve into intelligent agents capable of autonomously optimizing dynamic pricing, inventory management, promotional strategies, and location selection decisions.

In addition, future studies may model competitive market dynamics in greater detail. Factors such as new store openings, store closures, price competition, advertising campaigns, and market entry and exit dynamics could be incorporated into the simulation to evaluate their effects on store performance under changing competitive conditions.

The simulation framework could also be extended beyond physical retail stores to include e-commerce and omnichannel retail environments. Incorporating variables such as online orders, delivery times, warehouse locations, and last-mile logistics would provide a more comprehensive and realistic representation of today's retail ecosystem.

Finally, the proposed model could be applied to different cities, countries, and business sectors to evaluate its generalizability. Similar simulation studies could be conducted for supermarkets, restaurant chains, pharmacies, bank branches, healthcare facilities, and other service industries to assess the model's applicability and performance across diverse retail and service environments.

ACKNOWLEDGMENT

The author would like to express sincere gratitude to her parents, sibling, and family for their unwavering love, encouragement, and continuous support throughout this research. The author also gratefully acknowledges Altınbaş University for providing the opportunity and academic environment to conduct this study. Finally, the author extends her sincere appreciation to Assoc. Prof. Dr. Fatma Sebnem Akal Ilkhan for her invaluable guidance, constructive feedback, and continuous support throughout the completion of this research.

REFERENCES

[1] Jones, K., & Simmons, J. (1990). The retail environment. Routledge

[2] Ghosh, A., & McLafferty, S. L. (1987). Location strategies for retail and service firms. Lexington Books.

[3] Brown, S. (1992). Retail location: A micro-scale perspective. Avebury.

[4] Craig, C. S., Ghosh, A., & McLafferty, S. (1984). Models of the retail location process: A review. Journal of Retailing, 60(1), 5-36.

[5] Birkin, M., Clarke, G., & Clarke, M. (2002). Retail geography and intelligent GIS. John Wiley & Sons.

[6] Hernandez, T., & Bennison, D. (2000). Retail location planning in an era of corporate growth. International Journal of Retail & Distribution Management, 28(8), 357-366. https://doi.org/10.1108/09590550010337391

[7] Davies, R. L. (1976). Marketing geography: With special reference to retailing. Methuen.

[8] Clochoatti, A., Santos, J., Alturas, B., & Postolache, O. (2020). Retail site location selection using GIS and multi-criteria decision analysis. International Journal of Agile Systems and Management, 13(2), 145-168.

[9] Von Thünen, J. H. (1826). Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie. Wirtschaft & Finan.

[10] Christaller, W. (1933). Die zentralen Orte in Süddeutschland. Gustav Fischer. (English translation by C. W. Baskin, 1966, Central places in southern Germany. Prentice-Hall).

[11] Lakshmanan, T. R., & Hansen, W. G. (1965). A retail market potential model. Journal of the American Institute of Planners, 31(2), 134-143. https://doi.org/10.1080/01944366508978155.

[12] Hotelling, H. (1929). Stability in competition. The Economic Journal, 39(153), 41-57. https://doi.org/10.2307/2224214

[13] Eaton, B. C., & Lipsey, R. G. (1975). The principle of minimum differentiation reconsidered: Some new developments in the theory of spatial competition. The Review of Economic Studies, 42(1), 27-49. https://doi.org/10.2307/2296817

[14] Law, A. M. (2015). Simulation modeling and analysis (5th ed.). McGraw-Hill.

[15] Banks, J., Carson, J. S., Nelson, B. L., & Nicol, D. M. (2010). Discrete-event system simulation (5th ed.). Prentice Hall.

[16] Macal, C. M., & North, M. J. (2010). Tutorial on agent-based modeling and simulation. Journal of Simulation, 4(3), 151-162. https://doi.org/10.1057/jos.2010.3

[17] Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences, 99(suppl 3), 7280-7287. https://doi.org/10.1073/pnas.082080899

[18] Epstein, J. M. (2006). Generative social science: Studies in agent-based computational modeling. Princeton University Press.

[19] Tesfatsion, L. (2006). Agent-based computational economics: A constructive approach to economic theory. In L. Tesfatsion & K. L. Judd (Eds.), Handbook of computational economics (Vol. 2, pp. 831-880). Elsevier.

[20] Rand, W., & Rust, R. T. (2011). Agent-based modeling in marketing: Guidelines for rigor. International Journal of Research in Marketing, 28(3), 181-193. https://doi.org/10.1016/j.ijresmar.2011.04.002