Decision Support System for Retail and Logistics Center Site Selection Based on Open Data

小組成員:

1111646 梁凱晰 組長

1111651 翁嘉筠 組員

1111721 林楦嘻 組員

1111744 黃英綺 組員

1111750 葉明蓁 組員

1. Topic

Decision Support System for Retail and Logistics Center Site Selection Based on Open Data

2. Abstract

The main goal of this research is to develop a "Decision Support System (DSS) for Retail and Logistics Site Selection Based on Open Data." The system is designed to help companies make better decisions when choosing new locations. By using data instead of personal judgment, it reduces uncertainty and risk, and improves the accuracy of site selection decisions. As data-driven decision-making (DDDM) becomes more popular, more businesses are realizing the importance of data analysis in planning and strategy. At the same time, open data makes it easier and cheaper for companies to access reliable information, making it more practical and flexible to build effective decision support systems.

Our system mainly uses data from the government open data platform and integrates five key indicators: population structure, transportation convenience, rental cost, competition density, and regional development potential.

We first integrate and standardize the data to build a regional evaluation database. Then, we use the AHP method to analyze the weights of each indicator, and combine it with the TOPSIS method to score and rank the candidate locations.

This approach considers both qualitative and quantitative factors, making the results more accurate and reliable. The system first integrates and standardizes data from different sources to build a region-level analysis database, which serves as the foundation for model analysis.

In the model design, we use AHP to collect decision-makers' opinions on the importance of each indicator and calculate their respective weights. Then, we apply TOPSIS to evaluate and rank the candidate locations, identifying which ones best meet the ideal criteria. This approach combines subjective judgment with objective data, making the results more reasonable and trustworthy.

The system architecture is divided into four main modules:
(1) **Data Integration Module** – automatically collects and updates government open data.
(2) **Database Management Module** – handles data cleaning, normalization, and structured storage.
(3) **Multi-Criteria Decision-Making Module** – performs AHP weight calculation and TOPSIS ranking algorithms.
(4) **Result Visualization Module** – presents each region's overall score, strengths, and limitations through GIS maps and interactive dashboards.

We hope that this system can help retail businesses open new stores, logistics companies choose warehouse locations, and chain brands redesign their store layouts. By doing so, the decision-making process becomes more data-driven, efficient, and reliable.

3. **Background, significance and motivation**

In today's competitive world of globalization and digitalization, the strategy behind where a company chooses to set up its locations has become one of the key factors affecting business performance. For the retail industry, store locations directly impact customer traffic and sales performance, while for the logistics industry, the placement of warehouses and distribution centers affects the overall efficiency and cost of the supply chain.

In traditional site selection processes, decisions often rely on the management team's past experience and personal judgment. While this can quickly provide an initial direction, it often overlooks multiple important factors—such as market changes, cost limitations, and external environmental influences. Without a comprehensive analysis, decisions tend to be based on subjective opinions rather than systematic or data-driven reasoning, which increases investment risks and long-term operational costs.

In recent years, with the growing availability of government open data and continuous advancements in information technology, companies can now more easily access various types of geographic and industry-related information — such as population structure, transportation convenience, rental levels, competition density, and regional development potential.

These open datasets have become valuable references for businesses when making location and strategic planning decisions. By analyzing consumer demographics, companies can adjust their strategies accordingly. Integrating and analyzing such data not only reduces data collection costs but also makes the decision-making process more transparent and scientific, aligning with the modern business mindset of being "data-driven." However, while open data provides diverse information, its sources are often scattered and its formats inconsistent. Without proper integration and analytical mechanisms, it remains difficult to apply this data directly to strategic decision-making.

Therefore, the main motivation of our research is to develop a Decision Support System (DSS) for Retail and Logistics Site Selection based on government open data. By applying Multi-Criteria Decision-Making (MCDM) methods, the system helps businesses transform complex, multivariable data into actionable decision insights. Specifically, the study uses the Analytic Hierarchy Process (AHP) to determine the weight of each criterion and integrates the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) to rank potential locations, identifying the most promising and beneficial sites.

The development of our system focuses on three main aspects:

1. Theoretical Aspect:
By combining AHP and TOPSIS, the system verifies the feasibility of applying multi-criteria decision-making models to open data and provides a standardized decision-making process.

2. Practical Aspect:

Through automated data collection and geographic visualization features, businesses can obtain more accurate market analysis results in less time, improving both site selection efficiency and decision quality.

3. Social Aspect:

The application of open data promotes the reuse of public information and enhances the value of government data, supporting the goals of smart cities and sustainable development (SDGs).

In summary, the motivation behind our research is to create a data-centered system that integrates open data, decision science, and information technology. Our goal is to design a flexible decision support platform that can be applied across multiple fields, providing valuable insights and references for different industries when planning new locations or business sites.

4. Literature review or related works

Research and Application Cases:

1.Decision support system for location selection of convenience stores and retail facilities using optimization techniques and GIS（Golovnin & Igonina, 2021）

The study developed a DSS specifically for convenience stores and retail facilities, combining GIS with open map data (using OpenStreetMap as the data source) to identify spatial potential areas, and applying the gradient descent method for optimization analysis.

ResearchGate

2.A Decision Support System for Business Location Based on Open GIS Technology and Data（Ghita, 2014）

This paper proposes a commercial site selection DSS framework based on open GIS technologies and open data. It integrates theoretical concepts, empirical research, and open GIS/open data into a unified practical tool.

IDEAS/RePEc

3.Decision Support System for Store Location Selection in Retailing（Döner et al., 2025）

The study focuses on site selection for retail stores, applying fuzzy AHP, TOPSIS, and a set covering optimization model to build a DSS, using bookstores as the case study.

University of Strathclyde

4.Data-driven decision support systems in retail（Jayasundara et al., 2025）

This review explores the technical foundations, practical applications, and challenges of data-driven DSS in retail contexts.

ResearchGate

5.Logistics center site selection by ANP/BOCR analysis（Peker, 2016）

The study applies the Analytic Network Process (ANP) and the BOCR (Benefits, Opportunities, Costs, Risks) analysis method for logistics center site selection.

ACM Digital Library

6.Decision Support Systems for Smarter and Sustainable Logistics（Guerlain et al., 2019）

This paper discusses the application of DSS in urban construction logistics and supply chains, including topics such as the location planning of Urban Consolidation Centres (UCC).

MDP

7.A Novel Approach for Data-Driven Automatic Site Recommendation and Selection（Baumbach et al., 2016）

This paper proposes an automated site selection method based on up to 200 influencing factors, including socioeconomic, geographic, ecological, and company-specific requirements.

arXiv

8.Leveraging Urban Big Data for Informed Business Location Decisions: A Case Study of Starbucks in Tianhe District, Guangzhou City（Xiang et al., 2023）

The study uses urban big data to analyze the spatial distribution and site selection factors of Starbucks stores in Tianhe District, Guangzhou.

arXiv

9.Logistics center site selection by ANP/BOCR analysis: : A case study of Turkey

Using ANP and BOCR for choosing logistics center locations.

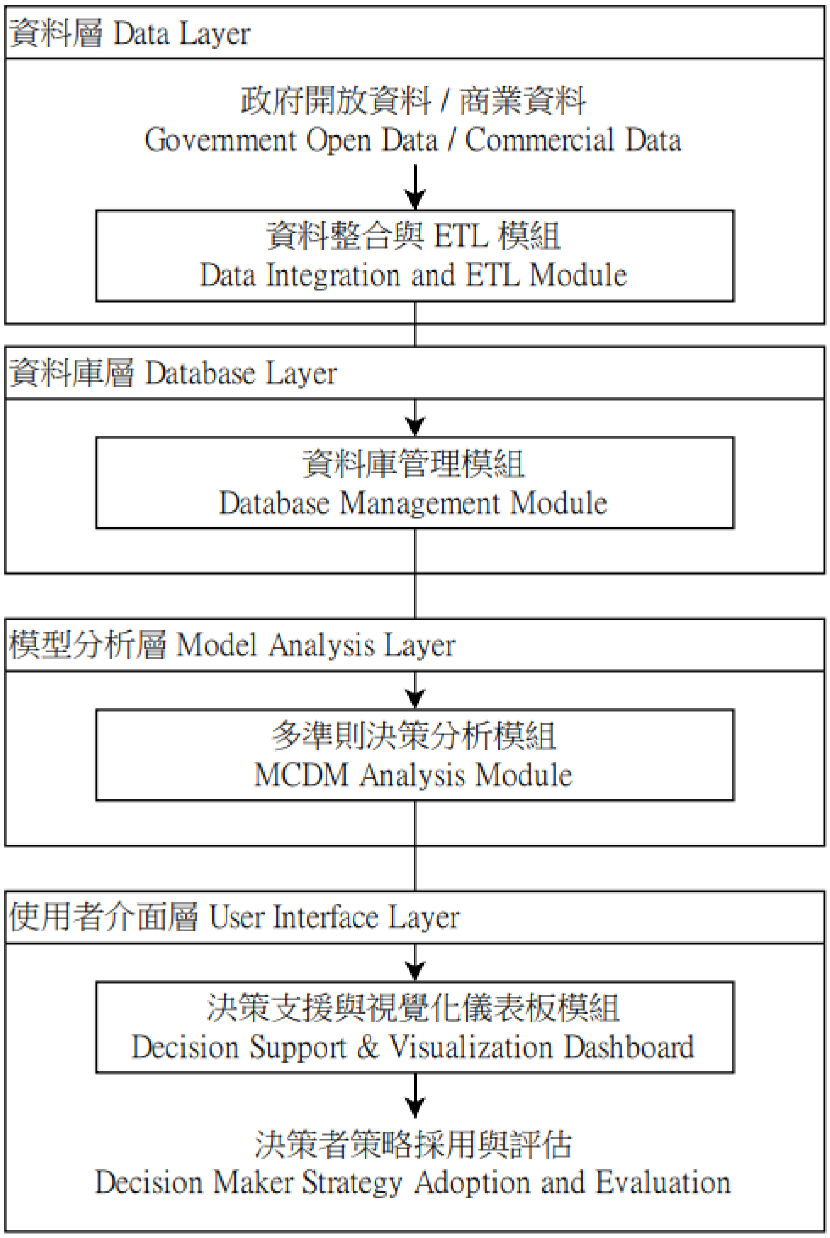
ACM Digital Libra

5. **Proposed framework of DSS**

The overall goal of this decision support system is not just to serve as a data analysis tool, but to act as a complete solution that integrates data processing, model computation, and interactive visualization. It helps businesses make location decisions based on data rather than relying solely on management experience.

The core idea is simple: turn complex data into an easy-to-understand map, and turn difficult decisions into simple actions.

The system framework is divided into four main layers: Data Layer、Database Layer 、Model Analysis Layer、User Interface Layer.



模組一：資料整合與 ETL 模組（Data Integration and ETL Module）

The system’s “information agent” is mainly in charge of turning external data into analyzable data. It handles the ETL process to make sure the data used for analysis is clean and ready to use.

Data Extraction (Extract):
The main data sources come from the government open data platform, including demographic information (such as age structure, median income, and population density), transportation facilities (like bus stops and MRT stations), the number of registered companies, and industry distribution.
In addition, simple web crawlers or APIs are used to collect data from sources like Google Maps for competitor branch locations and public rental listing websites for market rent information.

Data Transformation (Transform):
The raw data comes in many formats, such as CSV, JSON, or text-based addresses. All data must be converted into a standardized format, and missing or abnormal values need to be handled properly.

Data Loading (Load):
After cleaning and organizing, the processed data is imported into a central database for later analysis.

模組二：資料庫管理模組（Database and Management Module）

The system’s “central library” is responsible for storing all the cleaned and organized data. It manages the data systematically, making it easy to access and analyze later on.

From a technical perspective, to support spatial queries and map visualization, the ideal approach is to use a database with GIS capabilities, such as **PostgreSQL with PostGIS**. This setup allows for fast spatial computations, such as finding “the nearest MRT station” or calculating “the business density within a specific area.”

模組三：多準則決策分析模組（MCDM Analysis Module）

The system’s “brain” is responsible for finding the best site selection plan based on different decision goals. It uses two methods — **AHP (Analytic Hierarchy Process)** and **TOPSIS (Technique for Order Preference by Similarity to Ideal Solution)** — combining subjective preferences with objective data to make the decisions both reasonable and convincing.

AHP: This method is mainly used to calculate the weights of each evaluation criterion. The system guides users through pairwise comparisons (for example: “Which is more important, foot traffic or rental cost?”). Through this interactive process, subjective preferences are converted into numerical weights, establishing the relative importance of each decision criterion.

TOPSIS: After the weights are determined, the system compares each location's distance to the "ideal" and "worst" conditions. The closer a location is to the ideal condition, the higher its score. Finally, the system generates an objective ranking report of all candidate locations.

模組四：決策支援與視覺化儀表板模組（Decision Support & Visualization Dashboard）

The system's "face" is responsible for presenting the analysis results to decision-makers in a simple and intuitive way.

- **Interactive GIS Map:**
  The map displays all candidate locations, with colors representing different recommendation levels. Decision-makers can click on any location to view detailed information, such as population data, rental costs, and competitor distribution.

- **Heatmap:**
  The system provides heatmaps from different perspectives, such as **population density hotspots** or **rental cost hotspots**, helping decision-makers quickly understand the spatial distribution of market potential.

- **What-If Analysis:**
  Users can instantly adjust the weights of different indicators — for example, increasing the importance of "cost" or "transportation." The system will immediately recalculate and update the map results, allowing managers to simulate the optimal locations under different decision scenarios.

- **Automatic Report Generation:**
  The system can export PDF reports that include maps, indicators, scores, and recommendations, making it convenient for direct use in internal meetings and decision discussions.

6. **Methods**

Our implementation follows a "runnable and reproducible" principle. Based on the four modules described above, we develop and integrate the whole DSS so every step—from data extraction and processing, to analysis, to results display—is clear and traceable. Below are the concrete steps and technologies used in each stage.

1. Data Preparation

Start by pulling structured datasets from the Taiwan government open data portal (data.gov.tw): demographics, transportation, business registrations, and industry distribution. Use the Google Maps API or simple in-house crawlers to add public competitor locations and rental price info.
Save all raw data in CSV/JSON, and record both the source and the fetch timestamp.

Batch-geocode addresses (via a Geocoding API) into latitude/longitude. For failed geocodes, fix them with fuzzy matching or manual checks.

Clean the data by: filling small missing gaps with the area median, removing obvious errors (e.g., negative or extreme outliers), and standardizing field names and time formats.

2. Feature Design and Model Analysis

(1) Database Management Module

For each candidate location, the system generates both spatial and non-spatial features — including population count, number of nearby stores (competitors), average rent within 300m/500m/1km radii, and the distance to the nearest MRT or bus station.
All numerical fields are normalized using **Min-Max scaling** or **z-score normalization** to remove unit differences and ensure comparability between indicators.

(2) Multi-Criteria Decision Analysis Module

The weighting process follows a simplified **AHP** concept: decision-makers can use sliders or a 1–5 rating scale in the interface to set the importance of each criterion. The system then automatically normalizes these inputs into a weight vector whose total equals 1.

For ranking, the system applies the **TOPSIS** procedure:

1. Normalize all indicator values.
2. Multiply by the corresponding weights to create a weighted matrix.
3. Calculate each candidate location's distance from the ideal and negative-ideal solutions.
4. Rank locations based on their closeness scores.

Core computations are implemented through modular functions — such as normalize, apply_weights, and topsis_score — and the output includes both individual indicator scores and total scores for transparency and easy validation.

3. Result Presentation and Interactive Operations

The system uses an **interactive map interface** (built with **Mapbox**) as the main visualization tool. Candidate locations are marked on the map, with color or marker size indicating their recommendation level. Clicking a marker displays the location's key indicators and TOPSIS score.

Multiple **heatmap layers** (e.g., population density, rental distribution, and competition hotspots) allow users to analyze market conditions from different perspectives.

The **what-if simulation** feature lets users adjust weights in real time, triggering backend recalculations and updating the map dynamically. To balance performance, recalculations can be limited to the top *N* ranked sites or use caching strategies.

Finally, the system supports **report exports** in both **CSV** and **PDF** formats — including map snapshots and ranking tables — which can be directly attached to reports or used for decision-making discussions.

This approach emphasizes a **repeatable and traceable process**. By combining simplified weight settings with TOPSIS ranking, along with spatial features and interactive visualization, the goal is to build a functional decision support system within the given timeframe and produce **verifiable analytical reports**.

7. Job assignment and descriptions

Leader 1111646 梁凱晰:Literature review or related works、Methods、References

Member 1111651 翁嘉筠:Literature review or related works、Proposed framework of DSS、Expected benefits of DSS、References

Member 1111721 林楦嘻:Literature review or related works、Methods、References

Member 1111744 黃英綺:Literature review or related works、Proposed framework of DSS、Expected benefits of DSS、References

Member 1111750 葉明蓁:Literature review or related works、Abstract、Background, significance and motivation、Gantt Chart

8. **Gantt Chart**

Requirement Analysis												
Data Collection and Integration												
System Design												
Model Construction												
System Prototype Development												
Testing and Refinement												
Final Report												
	week 7	**week 8**	**week 9**	**week 10**	**week 11**	**week 12**	**week 13**	**week 14**	**week 15**	**week 16**	**week17**	**week18**

9. Expected benefits of DSS

This system is designed to assist businesses in selecting locations for new stores or logistics centers. Its main purpose is to **help companies make smarter location decisions through data**, transforming the traditional experience-based approach into a more **scientific and efficient process**.

The expected benefits can be divided into the following aspects:

(1) Improved Decision Quality and Technological Support

By applying multi-criteria decision analysis methods, the system can transform decision-makers'subjective preferences into adjustable weight ratios and use objective data to rank potential locations. This helps reduce reliance on experience or intuition alone. As a result, decisions become more consistent and reproducible, while the overall process gains stronger scientific grounding and transparency.

(2) Improved Data Integration and Analysis Efficiency

Through automated data extraction (ETL) and a centralized database management module, the system effectively integrates data from multiple sources — including government open data platforms, Google Maps, and rental price websites. This greatly reduces the time and effort required for manual data collection and organization, ensures data accuracy and timeliness, and enables decision-makers to access complete and consistent information more quickly, thereby improving operational decision efficiency.

(3) Strategic Planning and Investment Risk Reduction

The system can perform comprehensive evaluations across different regions, helping businesses identify the most beneficial areas for new store development or logistics center placement. This reduces the likelihood of poor investment decisions or resource waste. Through the **what-if simulation** feature, managers can instantly adjust indicator weights to evaluate location outcomes under various strategies, enhancing flexibility in investment planning and strengthening risk management capabilities.

(4) Enhanced Visualization and Communication Efficiency

By presenting analysis results through **interactive GIS maps and heatmaps**, decision-makers can intuitively understand the strengths and weaknesses of each region. The system also automates report generation, which supports decision meetings, cross-department communication, and strategic proposal discussions within the company.

(5) **Open Data Reuse and Enhanced Social Value**

This system demonstrates the practical value of open data in industrial decision-making, promoting the circulation and transformation of public information. For governments, it helps showcase the real benefits of open data policies; for businesses, it provides access to valuable decision-making insights at a low cost. Moreover, the system's framework can be extended to other domains — such as healthcare facility planning, educational institution placement, or community resource allocation — thereby supporting urban development and sustainable governance.

This system not only improves decision-making efficiency and accuracy at the enterprise level but also promotes the broader application of open data at the societal level. It achieves the goal of “data-driven decision-making and intelligence-driven efficiency”, providing the retail and logistics industries with a sustainable and reusable analytical foundation for strategic planning.

References

1. https://www.researchgate.net/publication/357178081_Decision_support_system_for_location_selection_of_convenience_stores_and_retail_facilities_using_optimization_techniques_and_GIS
2. https://www.researchgate.net/publication/222402824_Past_Present_and_Future_of_Decision_Support_Technology
3. https://www.researchgate.net/publication/283073405_A_Survey_of_Big_Data_Research
4. https://ideas.repec.org/a/mgt/youmgt/v12y2014i2p101-120.html?utm_source=chatgpt.com
5. https://pureportal.strath.ac.uk/en/publications/decision-support-system-for-store-location-selection-in-retailing?utm_source=chatgpt.com
6. https://www.mdpi.com/2071-1050/11/10/2762?utm_source=chatgpt.com
7. Government Open Data Platform, "Open Data Platform (data.gov.tw)," [Online]. Available: https://data.gov.tw/ [Accessed: Oct. 20, 2025].
8. Google, "Geocoding API — Google Maps Platform," [Online]. Available: https://developers.google.com/maps/documentation/geocoding/overview [Accessed: Oct. 20, 2025].
9. PostgreSQL Global Development Group, "PostgreSQL Documentation," [Online]. Available: https://www.postgresql.org/docs/ [Accessed: Oct. 20, 2025].
10. PostGIS Development Team, "PostGIS — spatial and geographic objects for PostgreSQL," [Online]. Available: https://postgis.net/documentation/ [Accessed: Oct. 20, 2025].
11. GeoPandas community, "GeoPandas documentation," [Online]. Available: https://geopandas.org/ [Accessed: Oct. 20, 2025].
12. Mapbox, "Mapbox Documentation," [Online]. Available: https://docs.mapbox.com/ [Accessed: Oct. 20, 2025].