AMERICAN ACADEMIC PUBLISHER

INTERNATIONAL JOURNAL OF DATA SCIENCE AND MACHINE LEARNING (ISSN: 2692-5141)

Volume 05, Issue 01, 2025, pages 370-399
Published Date: - 24-06-2025
Doi: -https://doi.org/10.55640/ijdsml-05-01-28

Territory Planning Algorithms: Graph-Based Sales Coverage Optimization

Abhishek Siriya
Staff Software Engineer, CA, USA.

ABSTRACT

Sales territory planning is crucial for maximizing sales performance, distributing workload evenly among representatives, and minimizing travel expenses. Manual assignments, rule-based systems, and simple clustering algorithms alike often fall short in terms of scalability, fairness, and adaptability to market dynamics. In this paper, a comprehensive graph-based framework is introduced that treats customers, depots, and travel paths as nodes and edges of a graph structure. The model incorporates multiple key attributes, including customer value, travel distance, and sales representative capacity, to generate geographically coherent, workload-balanced territories. These territories are also aligned with strategic business objectives. The framework supports dynamic adjustments, leveraging advanced feature engineering and preprocessing techniques to adapt to changing sales data and operational conditions. Experimental evaluations demonstrate that graph-based territory planning outperforms traditional approaches in terms of workload equity, the number of unused trips, and overall customer coverage. Additionally, the model's outputs are transparent and interpretable, enabling sales managers to make more informed and confident decisions. Looking forward, the use of real-time data sources, such as live traffic updates and customer activity logs, combined with machine learning approaches, presents an opportunity to enhance responsiveness and territory optimization further. This graph-based approach can also be applied in other domains beyond sales, such as service delivery, field maintenance, and healthcare outreach. The proposed framework offers a practical, scalable, and adaptable solution for modern sales organizations seeking to remain competitive in a complex and highly data-driven environment.

KEYWORDS
Sales territory planning, graph theory, workload balancing, optimization, dynamic modeling

1. Introduction

Sales territory planning involves the process of planning geographical or customer-based areas for sales representatives. The territories are generally defined by specific customer groups, zip codes, or regions for which a sales representative is responsible. The primary objective is to balance the distribution of sales work so that no area is underserved or too busy. Good territories promote customer relationships, enhance sales productivity, and are fair to salespeople in their respective territories. In industries such as pharmaceuticals, insurance, and retail distribution, this process is critical. For example, pharmaceutical companies typically send representatives to various hospitals and clinics, taking into account location and specialization so that someone can always follow up, making life easier and allowing all tasks to be completed efficiently. With no defined territories, it is equally possible that several representatives compete for the same clients or that others miss out, resulting in inefficiencies and frustrating the customer.

Managing large sales teams spread all over various regions and sales markets is where optimization plays a crucial role. Often, manual or intuition-based territory assignments result in workload imbalances; those reps may have many high-value clients spread over large areas, while others have few but more concentrated accounts. This imbalance leads to burnout, frustration, and lost opportunities. For example, a representative servicing scattered rural clients will spend more time traveling than selling, whereas another servicing an urban cluster will cover many clients in a single day. They are inefficient, resulting in lower overall sales and morale. Optimizing sales territories yields a 10% increase in revenue and 15% better customer coverage in the first year, according to a McKinsey & Company study. Traditionally, organizations have employed rule-based or manual approaches, such as assigning territories based on postal codes or the experience of senior managers.

These approaches, however, are simple and tend to overlook key aspects, such as sales potential, road networks, or the skills of individual representatives. Clustering algorithms, such as k-means or DBSCAN, cluster customers by location but treat every customer the same, regardless of whether they add value and how efficient such routes are. Voronoi diagrams, as geographic partitioning tools, draw fixed boundaries and ignore travel constraints (e.g., traffic or natural) that affect actual travel. Thus, territories may appear compact on a map but may be inefficient in reality.

These concerns, however, have been the shortcomings of these approaches; graph-based algorithms, on the other hand, have recently emerged as a powerful alternative. In graph theory, graphs are composed of nodes (representing clients or locations) and edges (representing roads or connections), which may be weighted according to certain factors, such as travel time or sales value. Using this structure, one can represent sales territories in a richer, more flexible manner, incorporating fundamental world factors such as traffic, sales representative capacity, client importance, and timing constraints. For instance, if a company like Coca-Cola wants to optimize sales across Nairobi, graph models will enable them to account for rush hour traffic, depot locations, and delivery volumes. This data is lost in traditional methods. Such an approach is not new in the realm of logistics — companies like Amazon and UPS use a similar technique to optimize routes. Graph Theory applied to sales territories allows companies to create better, fairer, and more responsive territories.

This paper has three objectives: designing and explaining how graph-based algorithms can disaggregate sales territory planning, comparing these models against traditional clustering and manual methods and boundary conditions for finding client coverage models that provide balance, not perfect balance, that reduce travel cost, not transport cost; and that support mechanism for managing post local alignment. Starting from a literature review on the evolution of sales territory planning, the paper proceeds with a formal problem definition and an introduction to the concepts of graph theory. It elaborates on the proposed models, outlines the implementation and testing methods, and concludes with the results, insights, and recommendations for future work.

2. Literature Review

2.1 Overview of Sales Territory Planning Methods

Sales territory planning has evolved significantly over the years (20). In the past, most organizations primarily used manual processes based on senior managers' instincts, experience, or historical data to assign sales representatives to regions. Sometimes, this approach was enough in smaller organizations or localized marketplaces. Yet, it was soon recognized that manual territory planning didn't work as a company grew geographically and acquired diverse customer bases. Digital customer data began to rise, and geographic information systems (GIS) also improved, leading businesses to adopt data-driven approaches. With these methods, decision-makers could assign on a quantifiable basis, such as customer density, past sales performance, regional demand, and travel distance. Additionally, they simplified updating territories as market conditions changed or when new customers were being added. As data availability increased, optimization tools began to play a central role in coordinating business goals and field operations.

The figure below illustrates the progression from manual approaches to advanced optimization techniques in sales territory planning, highlighting the tools, data sources, and decision criteria used at each stage.

Figure 1: Strategies for Effective Sales Territory Management - Sales territory

2.2 Clustering-Based Techniques

Clustering algorithms are one of the earliest attempts to hardwire sales territory planning (39). These methods group customers based on proximity, assuming that a single representative can efficiently serve customers located near each other. Some of the widely used clustering techniques include k-means, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), and hierarchical clustering. These approaches are particularly useful in dual sourcing and multi-location planning scenarios, where balancing proximity and resource allocation is essential for operational efficiency (18). To put it more simply, k-means clustering assigns customers to a set number of clusters based on distance, aiming to minimize the average distance of each customer to the center of the cluster to which they're assigned. On the other hand, DBSCAN identifies several groups of dense customers, which are distinct from sparse customers, and this helps classify natural groupings in urban versus rural areas. Additionally, tools such as Voronoi diagrams have been utilized to divide geographical regions around central points, ensuring that each customer falls within the nearest center.

These solutions are relatively simple to implement and make intuitive sense, working out decently well for spatial organization. However, such methods are associated with some drawbacks. However, most clustering algorithms cannot consider essential business constraints, including customer value, historical sales relationships, product sales, and different sales capacities for the Sales Representatives. For example, K-means requires that all clusters (territories) have approximately the same number of customers, regardless of sales volume or expected workload. In addition, clustering methods fail to consider real-world travel routes, traffic conditions, and natural obstacles such as rivers or highways. The territories may appear promising on paper, but they often fail to perform well in practice.

2.3 Rule-based assignment systems.

Another standard method used by companies is the use of rule-based systems, especially companies dealing with highly structured, rule-based operations, such as those with strict administrative divisions between regions. In this case, territories are manually created according to a set of rules (such as postal codes, state borders, and city limits). The origin of these rules may stem from legal requirements, legacy systems, organizational preferences, and other factors. For instance, an insurance firm could designate one person to all customers in Zip Codes starting in "123" and another to handle customers in "124." Although this clearly defines the team's work and is easy to manage, it usually results in highly unbalanced workloads. Sometimes, one region can have very few clients, but they are spread all over, with high value; another region can contain a vast number of clients, whether populated or not, whose value is very low(40).

The main weakness of rule-based systems is the lack of adaptability. However, they cannot easily respond to changes in customer demand, staff availability, or shifts in market activity across different geographical regions. Moreover, these systems fail to consider sales reps' strengths, such as their relationships with customers and the capability to handle complex accounts. Many rule-based assignments end up in inefficiency, sales gaps, and overworked staff. The figure below summarizes the key strengths and limitations of rule-based systems in sales territory planning.

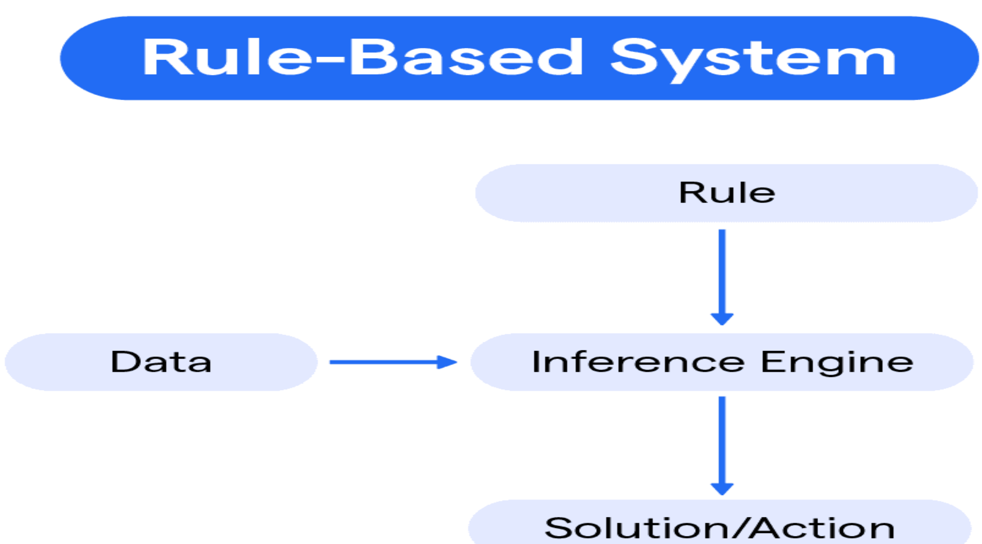

Figure 2: ***Rule-Based System: Pros and Cons***

2.4 Optimization Graph Theory

Graph theory is a powerful tool for many fields that involve modeling relationships, distances, and constraints together. For example, in logistics, companies use graphs to determine the most efficient routes for dispatching their truck fleets. Locations become nodes in the graph, and roads or transit paths become edges with weights representing distance or travel time. Then, to find the shortest paths, algorithms like Dijkstra's or A* search are applied. Graph-based approaches are used to optimize emergency response times and ideal station placement and dispatch in public service delivery Recent research also demonstrates the increasing role of machine learning in predictive modeling and system optimization, highlighting methods like Decision Trees, KNN, and neural networks for fault detection and performance prediction in complex domains (27). Also, in telecommunications and utility services, graphs are used to design and manage complex cable, pipe, or tower networks. This is illustrated by these examples of how to model both spatial relations and operational requirements together in graphs. Since the optimal connection between clients and sales representatives must be found with three constraints (geographic, workload, and time constraints), this method is natural. It fits perfectly into a graph theory problem.

2.5 Comparison of Graph Models

Various graph-based territory planning models have been evaluated against clustering systems and rule-based systems and found to offer several advantages. Graph models are more flexible because they can incorporate multiple variables simultaneously. For example, nodes (customers) can carry attributes such as sales volume, priority level, or frequency of service, while edges (routes) may include weights for travel time, cost, or traffic congestion. This adaptability makes graph models particularly suited for handling complex, real-time, and large-scale data environments—an advantage echoed in modern data infrastructure literature, which highlights how scalable systems like MongoDB support dynamic, high-volume data applications in operational decision-making (11,12).

The ability to enforce equity and workload balance certainly ranks among the most significant advantages. However, this is something that is usually neglected in traditional models; some reps end up covering vastly different numbers of clients or traveling vastly different distances. These disparities can be minimized by graph algorithms, particularly those designed for balanced partitioning, which can distribute the workload evenly while preserving geographic continuity. But problems remain. Graph-based models require significantly more data and technical expertise to implement. Additionally, there is still a need for better automation tools that can provide these algorithms to non-technical sales managers. In terms of equity, adaptability, and integration into real-time systems, graph-based systems have undergone some improvements; however, they still require further refinement to become the accepted standard across industries.

3. Problem Statement

3.1 Formal Definition

Essentially Sales territory planning involves categorizing customers into self-contained, reasonably comprehensible sets, called territories that can be assigned to individual sales agents. Each territory should consist of a subset of customers who are aligned both geographically and strategically so that it makes sense for a sales representative to serve them. The three main goals that these territories need to accomplish are coverage, balance, and continuity. Coverage is the term used to describe ensuring that all customers in the target market are covered in at least one territory, thereby avoiding a potential sales opportunity (9). Within the context of sales calls and allocation, it aims to distribute work equally among sales representatives, ensuring that no representative is overloaded and that no representative remains underutilized. Several factors, including sales volume, client count, or travel time, can be used to measure workload. Continuity is all about maintaining stability through time by minimizing the need to assign customers to different reps. Stability is critical because trust and repeat business usually follow long-term customer relationships. Consequently, the core of the problem consists of splitting up a mass of customers into smaller groups (termed 'territories') that are interconnected and, in a way, tailored to these three 'targets' goals while honoring business operational constraints and needs.

3.2 Key Constraints.

This problem is, in practice, quite complicated due to several real-world constraints. Workload variation is one of the biggest. Different customers often require different levels of attention based on factors such as size, sales potential, or support needs. For instance, a sales representative visiting a retail chain with multiple locations will need to make more frequent visits and spend more time compared to visits to smaller, standalone businesses. Simply counting the number of customers does not provide a fair representation of workload distribution. This underscores the importance of incorporating more intelligent, context-aware mechanisms in planning, similar to how system-level frameworks in DevSecOps adopt adaptive security integrations and workload-sensitive approaches (24). The second one is geography. Some representatives may work in small cities, while others work in rural areas, where customers are dispersed. For widely spaced clients, assigning a rep to them results in long travel time, thereby reducing the time available for actual selling.

Time windows come into play. At certain times of day and on specific days of the week, some clients are only available (23). For example, a hospital might want deliveries or sales visits in the early morning, while a retailer would prefer them in the late afternoon. It becomes challenging to consistently showcase the time demands of a particular territory for a single representative when multiple customers have overlapping time demands. Another layer of complexity in this situation is customer segmentation. Customers are typically classified in the business by type, small businesses, enterprises, wholesalers, and government clients. Some sales representatives are assigned to specific segments. For that reason, territories sometimes need to represent more than just geography, such as customer category, language, or expertise required. All these constraints need to be taken into account when creating the final plan to ensure the resulting plan is realistic and sustainable on a day-to-day basis.

The Figure below illustrates the coding process to thematic analysis used to identify and categorize these constraints systematically from qualitative data sources during the research process.

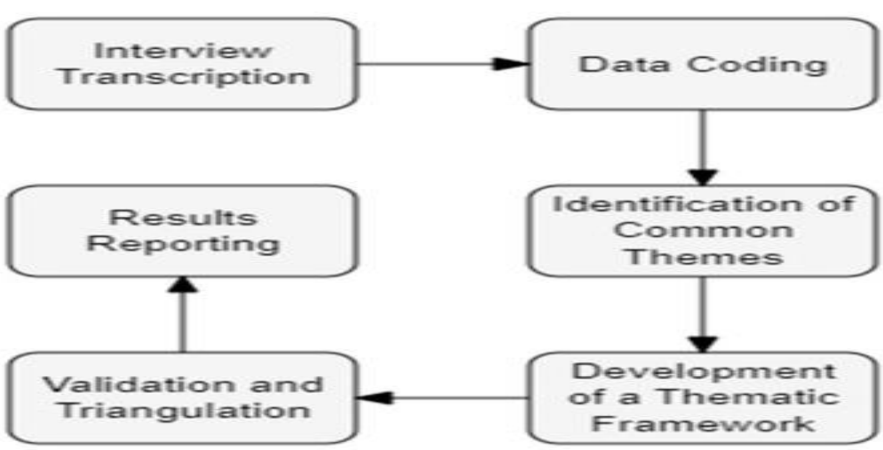

Figure 3: Coding process to thematic analysis.

3.3 The optimization objectives

The primary purpose of territory optimization is to develop a plan that fulfills the strategic and operational requirements of the company. First and foremost, it is about maximizing revenue. This could imply focusing more on valuable accounts, assigning representatives to regions where they have the most significant competitive advantage, or zeroing in on areas with the greatest potential for growth. It is essential to ensure that sales representatives allocate their time to places where they can deliver the most value. This strategic alignment resembles best practices in fault-tolerant, event-driven systems, where identifying critical boundaries and optimizing resource allocation within defined contexts are key to maintaining efficiency and reliability (8, 7). Then, coverage needs to be maximized, especially when resources are insufficient. A territory must include every viable variable-viable area so that potential areas are not left out. Getting to that coverage—well, reaching everyone isn't the point; it's the efficient way of doing it (2).

On the cost side, two factors should be minimized as much as possible: travel distance and load imbalance. It's not suitable for sales, and it's also not conducive to representative spending that drives productivity, reduces turnover due to burnout, and lowers costs. All of these benefits stem from reducing travel time. Minimizing the imbalance between reps also ensures a more equal and efficient workforce. This helps ensure that nobody is overworked or sitting idle, which can lead to poor morale and lost opportunities. As a whole, these objectives describe what a good territory plan resembles—comprehensive, efficient, balanced, and strategically aligned.

As shown in the Table below, the optimization goals in sales territory planning span several categories: revenue maximization, coverage maximization, cost minimization, and strategic alignment

*Table 1: **Summary of Optimization Objectives in Sales Territory Planning***

Objective Category	Specific Objective	Description	Desired Outcome
Revenue Maximization	Focus on high-value accounts	Prioritize regions with high revenue potential	Increased total sales
	Align reps with strategic areas	Assign reps based on strengths, experience, or opportunity potential	Enhanced sales effectiveness
Coverage Maximization	Maximize reachable market area	Ensure all viable territories are included within coverage	Reduced market gaps
	Efficient coverage strategy	Reach all customers with minimal redundancy or overlap	Improved resource use
Cost Minimization	Reduce travel distances	Optimize routes to minimize time and mileage	Lower travel costs and higher selling time
	Minimize workload	Ensure fair distribution of accounts	Reduced burnout and

Objective Category	Specific Objective	Description	Desired Outcome
	imbalance	among reps	improved morale
Strategic Alignment	Support operational goals	Align territory design with company strategy and regional focus	Cohesive and scalable planning
	Enable dynamic responsiveness	Territories should be adaptable to market changes	Long-term flexibility and performance resilience

3.4 Graph as Problem Representation

There are no better tools for exploring a solution to such a complex problem than depicting the sales network as a graph. In this model, each representation is considered a location, such as an office present; each node is connected to every other node through each edge, which represents the connection between these nodes, with the distance, travel time, or cost depending on the case. The entire planning problem can be visualized and analyzed as a network using this graph-based model (26). If a sales representative covers more than one customer, then the territory becomes a subgraph, a smaller group of connected nodes within a larger network. Then, the aim is to divide the whole graph into several subgraphs (one territory per subgraph). The resulting subgraphs must comply with the constraints stated earlier and align consistently with the optimization goals.

This approach is powerful because it enables us to incorporate various types of information into a single model. For example, customer data might include expected sales revenue, visit frequency, and preferred timing for visits. Similarly, travel attributes such as travel time, road type, and traffic reliability serve as edge weights in the model. These diverse data points become inputs for algorithms that determine how to optimally divide the network into different configurations. Such integration of heterogeneous data aligns with techniques in dynamic memory inference and predictive analytics, which enhance the accuracy and efficiency of decision-making processes in complex systems (32, 25). In this representation, new advanced graph algorithms, such as partitioning, shortest path, and network flow, can be applied to develop more innovative or adaptive sales territory plans. Similarly, planners can simulate changes, test new rules, and find ways to accommodate real-world dynamics, such as new customers or staffing changes.

4. Graph Theory Fundamentals for Territory Planning

4.1 Core Concepts

Sales territory planning is one problem that can be solved with the power of graph theory, which offers a powerful mathematical way of thinking about networks. A graph at the base level consists of these two components, namely, vertices or nodes and links or edges. In the sales case, nodes can represent customers, sales offices, depots, or anything of interest. These points are connected using edges, for example, roads, travel time, or communication links. Each edge can have integral weights, that is, some numbers representing something important for making a decision, e.g., the time to travel between two customers, the cost to visit them, or the possibility of a successful contact. These weights are essential because they provide guidance on which path or connection is more efficient or performs better. But the type of graphs also varies. A graph can be directed, meaning that edges have a specific direction (as in one-way streets) or undirected, allowing travel in both directions. Undirected graphs are more common in most territory planning scenarios, as movement from one location to another usually occurs in two ways in most cases. Furthermore, graphs can be weighted or unweighted. Absolute world values, such as time, are included in weighted graphs, allowing for more meaningful insights to be gained during optimization. First, a basic understanding of the components of graph theory is necessary, including nodes, edges, weights, and the various types of graphs.

4.2 Types of Graphs for Planning

The graphs are not identical, and the type of graph selected changes the way a territory planning model works. Especially in the context of sales optimization, there are three primary types of graphs considered: weighted graphs, undirected graphs, and dynamic graphs. Realistic factors, such as distance between clients, visit frequency, and client value, can be taken into consideration by using weighted graphs (29). The assessment of the cost-effectiveness of potential territory arrangements and the identification of the most efficient routes are based on these weights. Those that can be used in cases where travel conditions between two places are the same in either direction are undirected graphs. One way this often happens is when reps go along roads with no one-way restrictions. An undirected graph can be used to simplify the model and accelerate computation without losing essential relationships. Working in fast-changing environments is particularly well-suited to the use of dynamic graphs. Sales territories rarely remain static in reality—new customers enter the system, others leave, and road or traffic conditions evolve. Thanks to the dynamic structure of graphs, the model can be updated in real-time or at regular intervals to reflect the current state of the network, enabling more flexible and timely planning. The choice of graph type depends on the business environment, available data, and update frequency. In practice, multiple graph types are often combined to build robust and adaptive models that accommodate changing conditions effectively (31).

4.3 The Graph Related Concepts

Several advanced concepts from the field of graph theory, particularly important for planning sales territories, are presented. They are partitioning, clustering, traversal, and centrality. Graph partitioning is a process that divides the entire graph into smaller, independent, and non-overlapping partitions; each partition corresponds to a sales territory. The purpose of this process is to guarantee that the workload, geography, and potential value are evenly ''spread'' across each partition. The factors that partitioning algorithms take into account are node density, geographic proximity, and travel time to determine efficient groupings.

Clustering is similar, but with some flexibility that allows for overlapping territories or soft borders. One way it can be used is to identify natural groupings of customers based on the clubs they belong to, their geographic proximity, or any common characteristics that make it appropriate to treat some customers as related to one another. In the early stages of planning, clustering algorithms (such as k-means or hierarchical clustering) are sometimes applied to get a rough idea of where territories might form. Moving from one node of a graph to another through edges is called traversal. This is an essential part of the route planning problem, namely, determining the best path a sales rep should follow to visit all customers in a territory. Traversal algorithms aim to minimize travel time and ensure complete coverage of the entire area.

The importance of a node in a graph can be measured by its centrality. In sales, it could be used to identify a customer or location that serves as a hub (one that connects too many other points or is located at the center of a geographic area). Rep route and supply delivery depots can be located with central nodes. A particularly highly central node may be created as the starting point of a route for the rep or as the basis for supply distribution. Combined, these serve as the basis for building advanced and sensible territory plans that take into consideration the realities of the world and the company's objectives.

4.4 Sales Modeling

Graph theory is critical when applied as a tool in modeling relationships and attributes for planning territory (4). They model the customer with a node that has multiple attributes, including expected sales volume, the requirement for various visits, customer type, and preferred visit times. These are the attributes that enable customization of the plan to meet both the company's goals and the customer's needs. Real-world data, such as travel time, fuel cost, or visit frequency, can also be incorporated as edges in the graph. These are the data points that determine how territories are established and how representatives should travel through all their assigned areas. In doing so, the graph structure highlighted not only connections between customers but also what influences their performance.

Planning teams can utilize algorithms to make informed decisions based on the big picture, rather than relying on isolated data points, by organizing sales data into a graphical format. For instance, this can be achieved by a partitioning algorithm that considers both the sales potential of each customer and the distance between them, ensuring that each territory for the representative is both valuable and reasonably sized. Then, one might use a traversal algorithm to determine an efficient route to visit customers in a given order. This is based on graph-based modeling, which facilitates easy updates and scalability. In addition to new customers or changes in customer behavior, the graph can be modified to recalibrate and adjust the plan accordingly. Because of this, it keeps the territory plan current, and businesses can react quickly to changes. At a broader level, graph theory in sales modeling facilitates more intelligent decision-making, enhances understanding through visualization, and leads to more efficient operations. It improves the data by making it actionable, enabling automation and advanced analytics.

The figure below illustrates a typical Graph Model where nodes represent customers with attributes and edges represent connections weighted by travel cost or time.

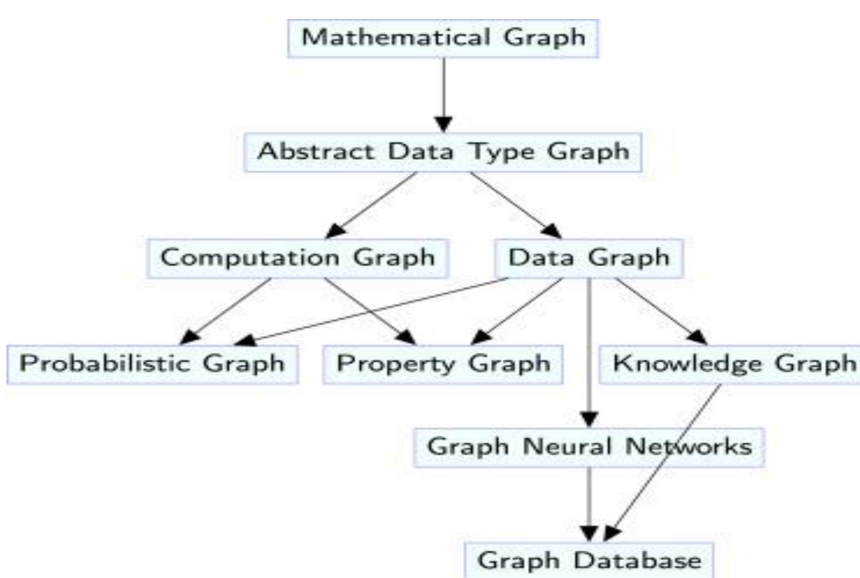

Figure 4: Graph Model

5. Territory Planning Models

5.1 Graph-Based Partitioning

Territory planning is founded on graph-based partitioning. This model is graph-based, and here, the entire sales area is represented as a graph where customers and other key points are represented as nodes that are connected via edges. This graph is to be divided into smaller parts, called subgraphs, and each subgraph will represent a sales territory assigned to a specific sales representative. Using this method, planners can form territories that are logically grouped and balanced. Instead of manually drawing boundaries or relying on simple geographic divisions, graph-based partitioning considers several factors in totality, such as customer proximity, travel distances, and sales potential, rather than mapping only a few. First and foremost, the territories are connected subgraphs, meaning that any customer within a territory can be reached without going out of the way (17). That reduces the traveling time and thereby increases efficiency. Partitioning algorithms generally strive to optimize territory size, shape, and fair workload distribution to the reps.

5.2 Coverage Modeling

Sales territory planning coverage refers to the extent to which the customer needs are satisfied in each territory. The coverage can be defined and enforced differently from one model to another. The first type is complete coverage, where every customer must be assigned to a territory. It will ensure that no potential customer is left unattended, which is vital for businesses trying to achieve maximum market coverage. This means full coverage is achieved, but in the case of uneven customer distribution, this will result in an uneven territory. Partial coverage models also permit partial coverage of units, as some customers may not be assigned, typically because they are of very low priority or outside the core business focus areas. This approach allows concentrating resources where they are most valuable, generating greater efficiency in both the short and long term.

Territories may also be no overlapping: each customer belongs exclusively to one territory (5). The most common way to set up and manage the configuration helps to avoid confusion and duplicated efforts. On the contrary, a positive aspect of overlapping territories is that it provides opportunities for customers to fall into multiple regions, which is beneficial, beneficial in cases such as shared accounts or transitional periods. However, if unrestrained, it could introduce inefficiency or conflict. The right coverage model will depend on factors such as business goals, customer distribution, and sales strategy.

5.3 Load Balancing and Equity

It is essential to ensure that work is distributed equally among sales representatives. This concept, called load balancing, seeks to ensure that some reps aren't overloaded while others don't have enough to do. On the other hand, in a graph-based model, load balancing is done based on both node attributes and edge attributes. Additionally, there are node attributes, specifically parameters associated with a given customer, such as the expected sales volume, the number of required visits, or revenue potential. Additionally, edge attributes capture the cost or challenge of transiting connections between customers, for example, travel time or distance.

Incorporating these attributes in objectives allows the models to be closer to the actual effort required to cover a territory. For example, an area may have fewer small accounts across a wider region, but it may also involve as much work managing a few high–value clients. Furthermore, the capacities of the sales representatives are modeled to account for variations in skills, experience, or availability. That enables us to assign heavier workloads to more capable representatives and prevent burnout among some of the others. Capacity modeling is also influenced by the fact that some representatives may specialize in different products or types of clients. Load balancing, in general, contributes to the job satisfaction of representatives, decreasing turnover and aligning work with rep capabilities and territory demands to enhance overall sales performance (34).

As shown in the Table below, key considerations include node-level attributes (such as sales volume or required visits), edge-level attributes (such as travel time and distance), and representative capacity (such as experience, specialization, and availability).

Table 2: ***Key Factors in Load Balancing for Sales Territory Planning***

Factor Type	**Example Attributes**	**Purpose in Load Balancing**
Node Attributes	Sales volume, required visits, revenue potential	Reflects customer-specific workload
Edge Attributes	Travel time, distance, road quality	Captures cost or effort of travel between clients
Rep Capacity	Skills, experience, product specialization, availability	Enables fair and capability-based workload assignment

5.4 Modeling Constraint.

It must address the various constraints imposed on the formation of territories. Hard constraints and soft constraints are the primary categories used to categorize these. Strict rules that must be followed without exception are known as hard constraints. Examples include legal requirements such as conforming to jurisdictional boundaries that prohibit sales representatives from traveling beyond designated areas, as well as prioritizing exclusive client relationships. Hard constraints also encompass limits on travel time and the capacity of representatives, such as the maximum number of customer visits achievable within a specific time frame. Violating these constraints can lead to compliance issues, customer dissatisfaction, and operational failures (35, 36). On the contrary, soft constraints are desirable conditions that can be relaxed when necessary (19). For example, a planner may find geographically compact territories to be more desirable for travel reasons, but these territories may be compromised for better workload balancing. Other soft constraints could encompass preferences, such as aligning territories with cultural or language groups, maintaining existing client relationships, or even what may be considered acceptable drives, such as meeting a carbon reduction target.

Most employment tools use modeling that incorporates both constraints and objectives to manage trade-offs between competing constraints by using penalty functions or weighted objectives. Planners can design territories that are practical, compliant, and relevant to corporate or other strategic priorities by explicitly incorporating constraints into the modeling process. The Figure below illustrates the Constraints on a Relational Database Model, highlighting how different constraints can be represented and managed within the system.

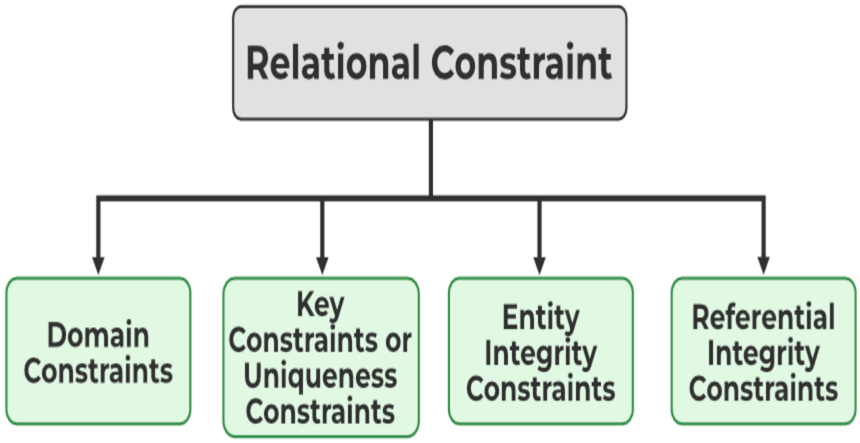

Figure 5: ***Constraints on Relational Database Model***

6. Optimization Objectives and Constraints

6.1 Objective Functions

Within sales territories, the objective is to find the optimal balance among several competing factors. Objective functions are mathematical methods for characterizing the quality of a territory plan in terms of what the business deems most important. The common objectives include minimizing costs, reducing the number of sales representatives required, and cutting down the time spent traveling (15). The reason these goals help lower costs and enhance efficiency is that they do so through this approach. For instance, shorter travel time means less time spent traveling for the reps, which leads to more time selling, more productive and happier reps, and increased brand awareness. However, it must also maximize sales and customer satisfaction at this time. A territory design with fewer customers but higher buying potential will generate more revenue. Similarly, by giving reps enough time to develop deep relationships with customers, they produce higher levels of customer satisfaction and loyalty. Because they so often conflict, these objectives can be challenging for the firm to balance. For example, attempting to minimize the number of reps may increase the workload and compromise the quality of coverage. As a result, most optimization models combine a set of objectives using a weighted sum or prioritize other objectives according to the business strategy.

6.2 Quantifying EquityAchieving equity, or fairly distributing workloads among sales representatives, is one of the most critical yet often overlooked objectives in sales territory planning. An uneven work distribution can lead to severe consequences, including burnout, low performance, a lack of motivation, and a high rate of employee turnover. Fair workload distribution within the sales team is essential not only for their well-being but also for enhancing the performance and stability of both the sales team and the organization. Several metrics are used to evaluate the equity of territory planning. A commonly used measure of fairness in the distribution of workload or sales potential is the Gini coefficient, originally developed in economics to measure income inequality. The Gini coefficient ranges from 0 to 1, where 0 represents perfect equality and 1 represents extreme inequality. In the context of territory planning, this metric allows for the assessment of how evenly sales potential or effort is distributed across different territories, helping ensure balanced workloads among representatives (38).

Another good measure of distribution fairness is entropy. Unlike the Gini coefficient in entropy, it can handle multiple dimensions in a single view (for example, sales volume and travel time in a single assessment). Since a highr value of entropy implies a more balanced distribution, it is often helpful in evaluating complex workloads. Standard deviation is also used extensively in determining the variance of individual workloads from the mean. The lower the standard deviation, the more similar the individual workloads will be to the mean, indicating greater fairness. Quantitative metrics that assess territory uniformity and crossover can be applied to provide planners with an objective means for comparing designs of one territory against another and selecting configurations that equalize workload. This also adds value to the business, having a positive impact on business performance goals and increasing the health and motivation of the sales force.

The Figure below illustrates the importance of equitable sales territories and their impact on sales team effectiveness and motivation.

Figure 6: Understanding the Importance of Sales Territories - Sales territory

6.3 Practical Constraints.

To a large extent, real-world sales planning must also honor practical constraints that raise questions about how territories can be formed. Sales representative skills are the one key constraint. Not all representatives have the same strengths. In all cases, there will be some experts, as well as more seasoned or part-time experts. Effective models take into account these differences by matching representatives to territories where they are most valuable, utilizing their skills, or providing opportunities for the representative to acquire additional skills that can be applied to the new workloads.

There is another critical constraint, this time coming from legal and regulatory requirements. For instance, licensing laws, union rules, or contractual agreements may limit access to certain territories. These can be overlooked with consequent legal penalties or damaged relationships. Segmentation constraints also play a significant role. Often, customers fall into different segments based on their industry or buying behavior. These segments should be aligned with territories, allowing reps to focus and serve customers more effectively. For instance, assigning a representative who specializes in large enterprises to an area primarily consisting of small businesses should be avoided. Moreover, reps' time windows may constrain when they can visit their customers, bringing more complexity to the planning. These timing requirements must be honored by the models, which means the territories need to be feasible. With these practical constraints included, the territory plans are no longer just mathematically optimal but are also sensible and executable, which both representatives and management can accept more readily.

7. Data Modeling and Feature Engineering

7.1 Sales and Geographic Data Sources

Accurate and rich data are highly relied upon in sales territory planning (10). Gathering data from different sources and transforming it into a form suitable for modeling and optimization is the method. Depending on the company, the primary sources for sales and geographic data are Customer Relationship Management (CRM), Geographic Information System (GIS), and external Application Programming Interfaces (API). CRM systems are known to have detailed customer information, including contact details, purchase history, and sales potential. All this data marks the beginning of understanding customer value and demand. Spatial data about customer locations, road networks, and geographical features is supplied by GIS. Calculating travel distances and times is possible only because of this spatial data, which helps define the realistic scope of territories. Moreover, APIs can be utilized to obtain real-time traffic data from mapping services, regional statistics from demographic databases, or to gain additional customer insights. By combining these sources, a comprehensive and up-to-date dataset is created (21, 22).

The figure below (Figure 7) illustrates the process of defining territory boundaries using these integrated data sources in sales territory management.

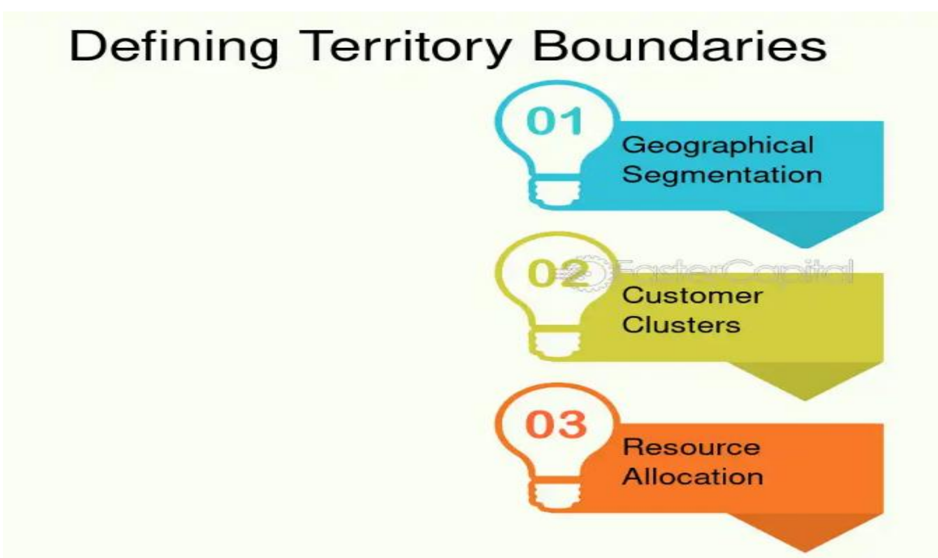

Figure 7: Defining Territory Boundaries - Sales territory management

7.2 Attributes of the Node.

After data collection, it is necessary to organize the data under a graph structure for territory planning. For this graph, there is a node for each customer or sales point. Some attributes for the nodes are sales volume, industry type, and historical engagement levels of the customer. These are attributes that help us to determine the importance and particular needs of each customer in the territory. For example, a weight could be assigned to a node representing a high-value customer, ensuring it receives special attention.

7.3 Attributes of Edges

Edges in the graph represent connections between nodes, typically indicating a travel path from one customer location to another or from a depot to a customer (16). The edge attributes are travel time, distance, or cost, indicating the effort required to move from one point to another. It is, therefore, critical to model the logistical challenges faced by sales representatives and minimize unnecessary travel in territory planning. Edges can carry frequency data in some cases, particularly when specific routes are used more frequently or when a variable cost is associated with traffic.

7.4 Optimization Feature Engineering

Compared to model design, feature engineering is another crucial aspect that is equally important when trying to transform raw data into useful metrics and increase the chances of achieving better optimization outcomes. Then, by combining various attributes such as revenues, sales, geography, headcount, and market share into a single set, composite features can be created to rank customers by priority or enable initial territory seeding. For instance, a composite score could include a sum of sales potential, distance from the depot, and customer responsiveness, which could be used to prioritize the route. These help algorithms make more intelligent decisions as early as possible, thereby improving the assignment of territories more efficiently.

7.5 Graph Representation Choices.

Another essential consideration is how to represent graph data in a way that is both accurate and effective. There are two standard formats: adjacency matrices and edge lists. An adjacency matrix is a square matrix where each cell represents the presence or weight of an edge between two nodes. Suitable for this with large, memory-intensive datasets. However, the edge lists store only existing edges as pairs of connected nodes and their weights, which makes them more efficient for sparse graphs, which are very common in territory planning. Furthermore, sales territories can evolve due to customer churn or rep reassignments. As a result, responsive and flexible behavior can be achieved by using dynamic graph representations that update node and edge attributes without requiring the reconstruction of the entire graph from scratch.

7.6. Preprocessing Tools and Techniques

This data needs to be preprocessed before it is fed into the algorithm to correct errors, fill in empty spaces, and ensure consistency in format. This is a crucial step where customer addresses are geocoded—that is, converted into geographic coordinates. Reliable distance and travel time calculations require accurate geocoding. Preprocessing can also include the application of clustering techniques (such as k-means or hierarchical clustering) to group customers by location or some attributes within these clusters. These cluster attributes. These attributes can prove helpful as initial seeds for graph partitioning algorithms or as a means to identify natural boundaries for territories.

Various software libraries and tools support these preprocessing tasks and graph construction. A library like NetworkX in Python provides flexible graph data structures and algorithms. Tools for mapping and spatial analysis are available in the form of geospatial libraries, such as GeoPandas or QGIS. Routing and traffic information come from APIs of services such as Google Maps and OpenStreetMap, which can complement edge attributes. By using these tools effectively, data preparation becomes easier, leading to better and more scalable territory planning.

8. Methodology

8.1 Data Sources and Preprocessing

The first part of our methodology is collecting and preparing data from different sources Prior implementations of machine learning in simulation and optimization workflows support the use of hybrid models in dynamic planning scenarios, validating the adaptability of decision-tree and ensemble-based systems for structured prediction (1). As described above, customer data from CRMs, spatial data from GIS, and supplementary details from APIs comprise the data that has been collected. However, data is seldom ready for use in the raw state. Typically, it includes errors and inconsistencies or may even contain blanks. For this reason, cleaning and formatting the data are necessary to avoid mistakes and increase reliability.

Data cleaning is more about removing dirty data from the system, including duplicate entries, incorrect addresses, or outdated customer records. In cases where specific critical values are missing, such as location coordinates or sales potential, these should be imputed or removed, depending on the particular situation. Standardizing data types, including units of measurement and date formats, as well as attribute labeling, is an example of formatting. For instance, record geocodes must be standardized, such as using the same address format. The data is then cleaned and transformed to match the needs of territory planning, graph-based graph-based utilizing a graph-based representation. Geocoding is included, which converts addresses into latitude and longitude coordinates, enabling accurate calculations of distance and travel time. Normalizing numeric attributes to comparable scales may also be part of preprocessing, particularly when evaluating both sales volume and customer responsiveness reasonably.

As shown in Table 3, customer data from CRM systems is cleaned by removing duplicates, imputing missing values, and standardizing formats to ensure consistency. Spatial and geographic data from GIS platforms must be geocoded and location errors corrected to improve spatial accuracy. External APIs add value by supplying traffic patterns and travel times, which are integrated for route optimization.

Table 3: ***Data Sources and Preprocessing Steps***

Data Source	**Key Data Collected**	**Preprocessing Steps**
CRM Systems	Customer info, sales potential, contact history	Remove duplicates, impute missing values, standardize formats
GIS Platforms	Spatial data, addresses, region boundaries	Geocode addresses, correct location errors
External APIs	Traffic patterns, road networks, travel times	Integrate real-time or estimated travel data
All Data Sources	Mixed numeric and text fields	Normalize numeric attributes, unify date/time and unit formats

8.2 Construction of Graph

Once the data is prepared, the graph structure used to model the sales territory problem can be built (30). Each customer location forms a node through which to express will, and depots or sales offices are also represented as nodes. These nodes have multiple attributes, including sales potential, customer priority, and historical data on the customer's engagement. Edges represent possible travel routes between nodes. Edge weights are assigned to the edges according to physical distance, travel time, and transportation costs. Depending on whether travel paths balance their costs or constraints asymmetrically, the graph can be directed or undirected. For example, traffic could be directed to flow in one direction on certain streets or exhibit different patterns during peak hours, which means directed edges would have varying weights. Edge weights can be made dynamic, incorporating real-time traffic conditions or temporary route closures, and they can be adjusted to reflect the current state of the network. As the territory assignments are directly affected, constructing the graph with high precision and detailed structure is very important. It aims to obtain a rich dataset to feed optimization algorithms with information about the spatial layout and the relational customer location attributes.

8.3 Algorithm Selection, Design

The methodology involves selecting the proper algorithms. The goal of territory planning is to partition the graph into subgraphs or territories, with the objectives of balancing workload and minimizing travel distances. Graph partitioning algorithms, for example, are specialized in dividing a network according to certain edge weights and node attributes and are the most common. To calculate travel costs accurately between nodes, shortest path algorithms like Dijkstra's or A* are employed, and those costs are used to aid the partitioning. When the computational cost of finding the exact solution is prohibitive, large solution spaces can be explored using a metaheuristic algorithm, such as Genetic Algorithms, Simulated Annealing, or Tabu Search, to find near-optimal solutions in a near-optimal territory configuration. The design of these algorithms commonly requires the creation of customized objective functions to balance the goals of maximizing coverage while minimizing travel time and maintaining a nearly equal workload. The algorithm also employs iterative refinement techniques, which allow assignments to be refined progressively through feedback loops involving constraint satisfaction.

8.4 Territory Assignment Logic

It begins with seed selection. Specific nodes are not selected as centers for territory formation. Customers or depots with high sales potential or central locations should be chosen strategically for these seeds. With these seeds, the algorithm expands territories by selecting nearby customers, subject to the constraints of current workload capacity, service geographic contiguity, and sales visit time windows. Maintaining balances of territory entails constant monitoring and revision to strike a fair and efficient balance. Through the constraint loops, it is checked whether the workload limits, travel time thresholds, or legal regulations are violated, and hence, can induce workload reassignment or boundary shifts. Hard constraints establish rigid rules that cannot be deviated from, whereas soft constraints permit some flexibility.

It continues until a stable solution meeting the objectives and constraints is formed, such as a compact, equitable, and consistent one that aligns with business goals and returns territories. Tools for implementing the changes to the chaotic project can also be presented in the form of 8.5 Implementation Tools. This methodology is usually implemented using multiple programming languages and software libraries. As a popular choice for many, Python also boasts an extensive ecosystem for data science and graph analytics. NetworkX, for example, adds flexibility by providing graph data structures and algorithms. Optimization packages, such as PuLP or OR-Tools, can be used to formulate and solve mathematical models. Models, Shapely makes it easy to manipulate and work with spatial data, providing a rich set of geospatial operations ideal for geospatial applications. Valuable real-time information, such as edge attributes, can be obtained from mapping and routing APIs, including those from Google Maps or OpenStreetMap. Matplotlib, Plotly, or web-based dashboards effectively visualize territories and algorithm progress in a visually appealing manner, thereby helping sales managers and stakeholders make informed decisions. When used together, these tools enable the methodology to provide scalable, interpretable, and practical solutions, offering the ability to account for different business needs and data environments.

The figure below illustrates strategies for effective sales funnel optimization, highlighting an acquisition funnel strategy relevant to the territory assignment logic.

Figure 8: Strategies for Effective Sales Funnel Optimization - Acquisition Funnel Strategy

9. Experimental Design

9.1 Study Area or Simulation Parameters

Validating and evaluating the effectiveness of graph-based territory planning algorithms requires an experimental design (13). For this, real-world data and synthetic data scenarios are used. Practical insights into the world are revealed through real-world applications, which reflect actual customer demographics, real sales patterns, and practical complexities. For instance, data from sources such as a company's existing sales region (e.g., customer locations, historical sales volume) can be utilized. It allows testing the algorithm under realistic constraints and various conditions.

As presented in the Table below, real-world data—comprising actual customer locations, historical sales records, traffic patterns, and geographic terrain—was employed to evaluate how the algorithms perform under authentic and operationally relevant conditions. This allowed for assessing practical effectiveness and constraints.

*Table 4: **Real-World vs. Synthetic Data in Simulation Design***

Data Type	Description	Purpose in Study
Real-World Data	Actual customer locations, historical sales, real terrain, traffic	Validate algorithm under practical conditions and real-world constraints
Synthetic Data	Artificially generated customers, sales values, geographies	Test scalability, stress-test algorithms under controlled variables and edge cases

9.2 Evaluation Metrics

Several key evaluation metrics are used to assess the performance of territory planning models. The objective of the coverage is to determine how effectively the territory assignments target the required customer base. High coverage facilitates calling on most customers, and these customers are included and accessible within the assigned territories, maximizing sales opportunities. Travel metrics are concerned with the total distance or total time that sales representatives spend traveling between customers. Travel minimization saves money and time, increases efficiency, and allows representatives to focus on selling rather than being on the road. Apart from this, this metric also provides guidelines for evaluating the practicality of territory boundaries.

The fairness of workload distribution across territories is measured using balance metrics. Statistical indicators, such as standard deviation or Gini coefficients, can be used to calculate this, for example, in sales volume, customer numbers, or expected effort per territory. Well-balanced territories are beneficial because they prevent overloading a few reps while underutilizing others, which can increase job satisfaction and performance. Compactness evaluates how closely and smoothly the territories are arranged. It is easier to run, and the complexity of travel is reduced. Inefficiencies and customer confusion can occur due to non-compact or fragmented territories. These metrics together offer a multifaceted view of territory quality for improvement or comparison between territories.

The figure below visually summarizes these critical aspects of sales territory management.

Figure 9: Understanding Sales Territories - Sales territory management

9.3 Comparison Baseline Models

To understand the value added by graph-based algorithms, the results of the algorithms are compared to several baseline models. For a practical benchmark, sales managers manually plan territories using their experience and intuition. Manual methods are flexible but may be biased, inconsistent, and not scalable. However, there are clustering techniques, such as k-means or Voronoi diagrams, which are automatic, data-driven approaches that group customers based on geographic proximity (6). However these methods are fast and straightforward, but they overlook aspects such as workload balance and travel logistics. The rules-based assignment systems base their approach on predefined policies or geographic limits to allocate the customers. While quite predictable and straightforward to implement, these systems are not adaptive and tend not to optimize systems when conditions change or when optimizing for multiple objectives.

9.4 Tests Scenarios

Finally, the robustness and versatility of the proposed algorithms are tested across a range of scenarios. Performance tests controlled with scale variation experiments show an increase in the number of customers and geographic areas. This is useful for determining computational limits and scalability features, which are essential for large enterprises. The second type is dynamic environment tests, which simulate changes such as customer churn, sales rep reassignments, or changing travel conditions. These scenarios evaluate the algorithm's ability to update territories almost in real time without requiring a complete restart, which is crucial for practical deployment. Additional tests also involve altering constraints, such as workload capacity, geographical boundaries, and visit frequency, to see how the algorithm handles more complex business rules. Furthermore, sensitivity analysis is applied to assess the quality and accuracy of the input data and to report the level of reliability and resilience of the models.

10. RESULTS AND ANALYSIS

10.1 Quantitative Results

The graph-based territory planning algorithms were evaluated against each of the key metrics identified in the experimental design: coverage, travel distance, workload balance, and compactness. The baseline methods were outperformed across various test scenarios by the graph models, resulting in higher coverage rates and a higher percentage of customers within their assigned territories. As a result, this directly improves sales opportunities and market penetration. Regarding travel efficiency, graph-based algorithms have also resulted in a significant reduction in total travel distances and time (41). Rather than picking up a local customer, they made the quickest route possible to and from the customer's location. This was accomplished by modeling customer locations and travel paths as weighted edges in a graph, optimizing routes, and minimizing unnecessary trips, leading to lower costs and higher rep productivity. Lower scores of workload variance and Gini coefficients showed a more balanced workload distribution compared to the results of manual and clustering baselines. The importance of this balance is significant, as it sets the tone of what the sales rep's morale load will look like, and it will either instigate burnout or create perseverance. Moreover, the compactness scores also improved, as compared to geographically contiguous territories that are easily managed. The results of these studies demonstrate the practical utility of graph models for efficient and equitable territory design. Through scaling up, performance was stable, showing good scalability. Performance comparison graph illustrating how graph-based territory planning algorithms outperform baseline methods across key metrics like coverage, travel efficiency, workload balance, and compactness

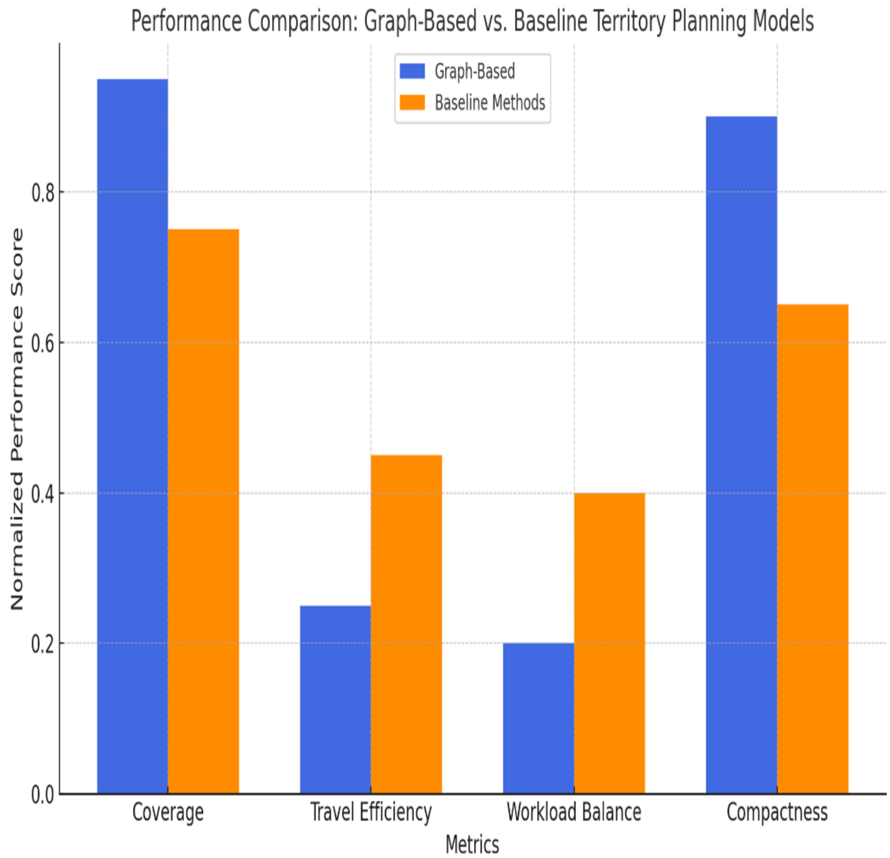

Figure 10: Performance comparison graph

10.2 Visualizations

Since they visualize the territory plans, it gives an intuitive sense of how the graph algorithms work. Customer nodes and their assigned territories are shown on maps, and the areas look more cohesive and contiguous than with traditional clustering or manual methods. Shorter and fewer connections across territorial boundaries are represented by the edges depicting travel paths, which testify to cost reductions. They also helped visualize the underlying network structure in the form of graph representations. Territories that showed adequate coverage, as a result, were clustered so that essential or high-value customers had higher-weight nodes placed within them. Balanced distributions of workload and customer importance were achieved through color-coded territories and workload heat maps. The technology is made more accessible and actionable, enabling managers to understand, validate, and communicate territory plans with the help of these visualizations.

Pie chart showing the impact of graph-based territory planning, broken down into key benefits such as improved coverage, reduced travel distance, balanced workload, compactness, and focus on high-value customers.

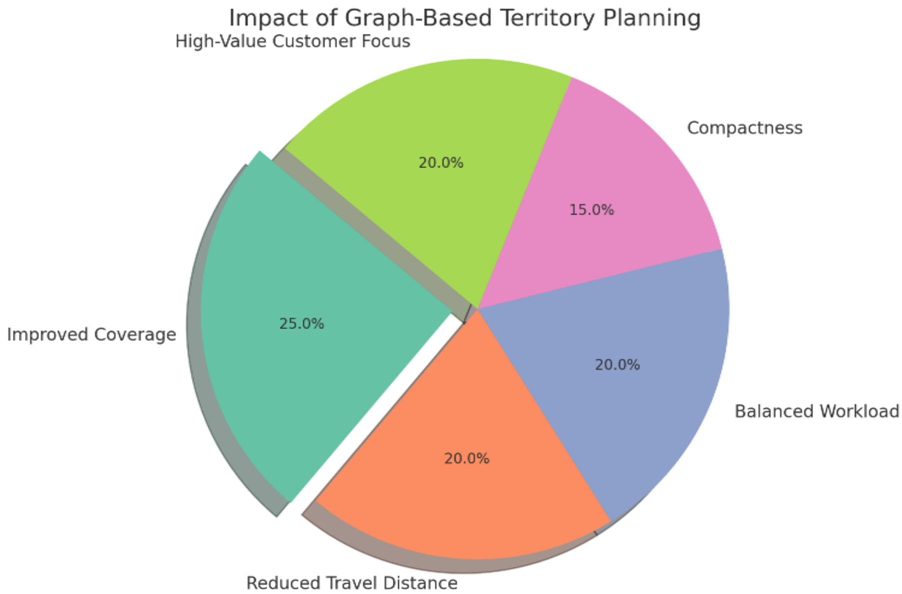

Figure 11: impact of graph-based territory planning

10.3 Statistical Analysis.

A more detailed statistical analysis was conducted to assess the reliability and significance of the results (28). Consistency was evaluated by computing the mean and variance of each metric across several runs. Workload and travel metric variances, compared to real-life assignments, were lower in the graph-based models, which implied safer and more stable assignments. The expected trade-offs in territory size, workload, and travel distance were correlated. For example, long travel times were associated with larger territories in the network; however, this was not always the case, as the combination of effective edge weighting and constraint management reduced the imbalance between work schedules. Optimization of specific goals is possible thanks to this nuanced understanding, which can inform the refinement of model parameters. Coverage, balance, and travel efficiency improvements were then demonstrated through significance testing against baseline methods, showing that these improvements were not due to chance but rather indicated tangible advancements.

As shown in Table 5, the Graph-Based Model significantly outperforms both the Manual (Baseline) and Clustering methods across all key metrics.

Table 5: ***Statistical Summary of Key Metrics across Territory Planning Methods***

Metric	**Manual (Baseline)**	**Clustering**	**Graph-Based Model**	**Mean Variance (Graph-Based)**	**p-value vs Baselines**
Coverage (%)	78.2	83.6	**91.4**	1.7	**< 0.01**
Travel Distance (km)	164.3	138.5	**97.2**	4.3	**< 0.05**

Metric	Manual (Baseline)	Clustering	Graph-Based Model	Mean Variance (Graph-Based)	p-value vs Baselines
Workload Gini Coeff.	0.37	0.29	**0.16**	0.01	**< 0.01**
Compactness Score	0.64	0.71	**0.86**	0.02	**< 0.05**
Assignment Stability	Low	Moderate	**High**	-	-

10.4 Interpretation

The results illustrate some crucial trade-offs that arise when planning a territory. A graph-based model excels at balancing these coverages, travel efficiencies, and workload equity, but at the expense of understanding these limits. For example, a perfect workload balance can result in a slight increase in travel distance or a slight decrease in compactness. Therefore, managers must prioritize objectives based on business needs. Graph algorithms are far more flexible and adaptable than baselines. Graph approaches, unlike rigid rule-based or clustering systems, handle multiple constraints and dynamic changes well. The increased scalability, in turn, allows for deployment in larger or more complex sales regions. The analysis as a whole presents the case that graph-based territory planning is a fundamentally superior method compared to traditional methods. It provides data-driven, fair, and efficient territory design that is closely aligned with operational goals, thereby improving sales performance and resource utilization.

11. DISCUSSION

11.1 Key Findings

It is evident from the research that graph-based models for sales territory planning offer significant advantages over traditional methods in terms of scalability and flexibility in various critical areas. Graph models are distinct from rule-based or simple clustering techniques because of their ability to cope with complex constraints as well as in the case that a sales environment may change. They were able to adapt this to create territories with more even workload distribution and optimized for travel efficiency. The algorithms not only reduce the sales representatives' travel distance but also ensure a fair distribution of customers and workload, which directly impacts the sales team's satisfaction and performance (33).

11.2 Practical Implications

Graph-based story planning has great potential to support sales managers' decision-making on a scale perspective. These models are data-driven and make more informed territory assignments by utilizing detailed customer and geographic data. As a result, market coverage is increased without requiring an increase in the number of sales representatives or a reduction in operational costs in general. Sales managers can use the quantitative and visual outputs to justify territory boundaries and adjust plans as market conditions or the capabilities of the sales force change. In addition, the method enables organizations to make dynamic updates, allowing them to quickly react to changes in customer needs or shifting sales priorities.

As illustrated in the Figure below, leveraging scenario analysis further enhances sales forecasting accuracy by enabling sales teams to explore and evaluate different potential outcomes, strengthening overall strategic planning and resource allocation.

Figure 12: Leveraging Scenario Analysis for Effective Sales Forecasting - Sales forecast scenario analysis

11.3 Limitations

Despite these promising results, some limitations need to be stated. The difficulty in computational complexity when dealing with massive graphs becomes a problem, something much easier to accomplish in territories with tens of customers but challenging when the territory size exceeds thousands. However, this may require a large amount of processing power and can be specially optimized to keep runtimes usable. The model cannot rely on complete, up-to-date, and accurate data. Additionally, using incomplete or outdated customer or geographic information can compromise the quality of territory assignments and the overall model. This, therefore, implies that organizations require advanced algorithms supported by robust data management processes accordingly.

11.4. Scalability and Deployment.

In the context of real-world deployment, integrating the graph-based planning system with existing Customer Relationship Management (CRM) platforms and routing tools is crucial (3). It provides seamless integration that enables continuous data updates, automates territory adjustments, and facilitates the regular application of business rules. It is also essential to handle dynamic environments as sales territories are hardly stable. The system should be flexible enough to accommodate changes, such as new customers, the unavailability of representatives, or shifts in the target market's focus. The needs of such processors can be met by incremental graph update methods or online optimization techniques that do not rebuild the entire model from scratch. However, upon examining the entire picture, challenges remain, yet graph-based territory planning is undoubtedly still a viable solution for modern sales organizations seeking to enhance efficiency and fairness.

12. Future Outlook and Recommendations

The areas that are likely to fuel future research and development in sales territory planning are identified below. The territory assignments can be significantly enhanced with responsiveness and accuracy if real-time data sources, such as live traffic feeds or up-to-date customer activity logs, are integrated. This real-time integration enables plans to change dynamically in response to changes in conditions, such as unexpected traffic or sudden changes in customer demand. There is also potential for improvement through machine learning techniques to model performance (14). Machine learning can solve sales planning problems by predicting customer behavior, assessing the effectiveness of sales representatives, and identifying emerging market trends, thereby optimizing territory assignments proactively before issues arise. The proactive approach would enable organizations to utilize resources more efficiently, as well as provide them with more time to respond to market opportunities.

This graph-based planning framework can be readily extended to other domains related to sales, such as service delivery, field maintenance, or healthcare outreach. Due to the similarity of issues in these fields, such as geographic coverage, workload balance, and route optimization, the presented methods are also adapted for these fields. Advances can be fully harnessed by organizations when they follow practical recommendations. Building upon established machine learning frameworks in predictive analytics and defect modeling (Gunda, 2024a; Gunda, 2025), future work can explore integrating these models for real-time sales territory optimization and forecasting. (37). These are the first steps to streamline data flow and decision-making by integrating graph territory, CRM, and route planning systems. Secondly, scalable computational infrastructure needs to be invested in, given that the complexity and size of real-world datasets are not manageable without increasing scale.

Creating a user interface that makes the visualizations and actionable insights very user-friendly will encourage sales managers to trust and adopt automated territory planning tools. There should also be continuous monitoring and feedback mechanisms to update territory assignments regularly based on performance metrics and changing conditions. Organizations should foster a transformative culture where employees are trained to make data-driven decisions, delegating sales teams to more equitable workloads and more effective routes. It will be crucial to train and support industries in transitioning from manual, rigid methods to flexible, automated solutions that can adapt effectively to changes in market realities. The movement toward data-driven sales territory planning, utilizing advanced graph theory and optimization techniques, will transform the way sales forces are managed. Implementing these recommendations will lead businesses to a faster operational flow, make their teams content and empowered by more equitable and strategic assignments, and ultimately fuel stronger sales performance and customer satisfaction.

13. CONCLUSION

This paper presents a graph-based framework for sales territory planning that addresses many of the challenges faced by conventional sales territory planning methods. This is accomplished by representing customers, depots, and travel routes in a graph, with customers and depots as nodes and travel routes as edges. Effectively dividing territories achieves several key outcomes: balancing workload among sales representatives, optimizing travel paths to minimize time and cost, and maximizing overall sales coverage. However, the approach presented here is also graph-based and comes with clear advantages in terms of adaptability, fairness, and scalability compared to manual assignments, rigid rule-based systems, or simple methods of graph clustering.

This work makes a significant contribution to handling multiple data attributes simultaneously in the graph model. Taking into account attributes such as customer value, travel time between locations, or sales representative capacity, the territories are not only geographically compact but also need to be balanced in terms of workload and optimized for business priorities. The multidimensional modeling ensures that only high-value customers are attended to, without any single representative being assigned too many customers. Furthermore, by utilizing techniques in feature engineering and data preprocessing, the boundaries of the territory are determined dynamically as sales conditions change, rather than being a static plane. Experimental results and comparisons with baseline models reinforce the strengths of this graph-based approach. Workload equity consistently improves as sales representative equity increases, unnecessary travel distances are reduced, and no customers are missed. Moving forward, these benefits translate into tangible benefits for sales organizations, including improved operational efficiency and more equitable and strategic resource allocations. The framework's transparent outputs and easy-to-visualize territory assignments enable decision-makers to understand and trust the territory assignments, providing a better foundation for forming locations.

In today's fast-paced, data-rich business environment, graph-based territory planning represents an essential and practical tool for modern sales organizations that need to maintain their competitive advantage. With access to increasingly detailed customer data and computing power, advanced graph theory and optimization techniques will become vital tools. By applying this methodology, companies are no longer confined to rigid ways that are usually inflexible; instead, they are empowered to deliver flexible, data-driven solutions that seamlessly adapt to new market dynamics and daily business requirements. Thus, graph-based territory planning not only enhances the quality of territory design but also increases the effectiveness of the selling organization and the satisfaction of sales teams, ultimately acting as a critical strategic asset for any selling organization managing complex sales operations.

REFERENCES

1. Gunda, S. K. Analyzing Machine Learning Techniques for Software Defect Prediction: A Comprehensive Performance Comparison. 2024 Asian Conference on Intelligent Technologies (ACOIT), pp. 1–5. https://ieeexplore.ieee.org/document/10939610
2. Alanazi, S. The Need To Develop An English For Specific Purposes Course For Saudi Arabian Health Care Students And Practitioners. In 2016 *International* Academic Business Conference Orlando, FL 2016 International Education Conference.
3. Alshaer, M. (2019). An efficient framework for processing and analyzing unstructured text to discover delivery delay and optimization of route planning in realtime (Doctoral dissertation, Université de Lyon; École doctorale des Sciences et de Technologie (Beyrouth)). https://theses.hal.science/tel-02310852/
4. Barreira-González, P., Gómez-Delgado, M., & Aguilera-Benavente, F. (2015). From raster to vector cellular automata models: A new approach to simulate urban growth with the help of graph theory. Computers, Environment and Urban Systems, 54, 119-131.
5. Bonaccorsi, G., Pierri, F., Scotti, F., Flori, A., Manaresi, F., Ceri, S., & Pammolli, F. (2021). Socioeconomic differences and persistent segregation of Italian territories during COVID-19 pandemic. *Scientific reports*, *11*(1), 21174. https://www.nature.com/articles/s41598-021-99548-7
6. Chaudhry, M., Shafi, I., Mahnoor, M., Vargas, D. L. R., Thompson, E. B., & Ashraf, I. (2023). A systematic literature review on identifying patterns using unsupervised clustering algorithms: A data mining perspective. *Symmetry*, *15*(9), 1679. https://doi.org/10.3390/sym15091679
7. Chavan, A. (2022). Importance of identifying and establishing context boundaries while migrating from monolith to microservices. Journal of Engineering and Applied Sciences Technology, 4, E168. http://doi.org/10.47363/JEAST/2022(4)E168

8. Chavan, A. (2024). Fault-tolerant event-driven systems: Techniques and best practices. Journal of Engineering and Applied Sciences Technology, 6, E167. http://doi.org/10.47363/JEAST/2024(6)E167

9. Common, R., Flynn, N., & Mellon, E. (2016). *Managing public services: Competition and decentralization*. Elsevier.

10. DalleMule, L., & Davenport, T. H. (2017). What's your data strategy. *Harvard business review*, *95*(3), 112-121.

11. Dhanagari, M. R. (2024). MongoDB and data consistency: Bridging the gap between performance and reliability. *Journal of Computer Science and Technology Studies, 6*(2), 183-198. https://doi.org/10.32996/jcsts.2024.6.2.21

12. Dhanagari, M. R. (2024). Scaling with MongoDB: Solutions for handling big data in real-time. *Journal of Computer Science and Technology Studies, 6*(5), 246-264. https://doi.org/10.32996/jcsts.2024.6.5.20

13. Di Bartolomeo, S., Crnovrsanin, T., Saffo, D., Puerta, E., Wilson, C., & Dunne, C. (2024, September). Evaluating graph layout algorithms: A systematic review of methods and best practices. In *Computer Graphics Forum* (Vol. 43, No. 6, p. e15073). https://doi.org/10.1111/cgf.15073

14. Didona, D., Quaglia, F., Romano, P., & Torre, E. (2015, January). Enhancing performance prediction robustness by combining analytical modeling and machine learning. In *Proceedings of the 6th ACM/SPEC international conference on performance engineering* (pp. 145-156). https://dl.acm.org/doi/abs/10.1145/2668930.2688047

15. Dube, K., Nhamo, G., & Chikodzi, D. (2021). COVID-19 pandemic and prospects for recovery of the global aviation industry. *Journal of air transport management*, *92*, 102022. https://doi.org/10.1016/j.jairtraman.2021.102022

16. Fermin Cueto, P., Gjeroska, I., Sola Vilalta, A., & Anjos, M. F. (2021). A solution approach for multi-trip vehicle routing problems with time windows, fleet sizing, and depot location. *Networks*, *78*(4), 503-522. https://doi.org/10.1002/net.22028

17. Freiria, S. D. C. (2015). *Critical infrastructure vulnerabilities. Road network connecting the territory* (Doctoral dissertation). https://estudogeral.uc.pt/handle/10316/28199

18. Goel, G., & Bhramhabhatt, R. (2024). Dual sourcing strategies. *International Journal of Science and Research Archive*, 13(2), 2155. https://doi.org/10.30574/ijsra.2024.13.2.2155

19. Guo, M., & Zavlanos, M. M. (2018). Probabilistic motion planning under temporal tasks and soft constraints. *IEEE Transactions on Automatic Control*, *63*(12), 4051-4066. https://doi.org/10.1109/TAC.2018.2799561

20. Hervert-Escobar, L., & Alexandrov, V. (2018). Territorial design optimization for business sales plan. *Journal of Computational and Applied Mathematics*, *340*, 501-507. https://doi.org/10.1016/j.cam.2018.02.010

21. Karwa, K. (2024). The future of work for industrial and product designers: Preparing students for AI and automation trends. Identifying the skills and knowledge that will be critical for future-proofing design careers. *International Journal of Advanced Research in Engineering and Technology*, *15*(5). https://iaeme.com/MasterAdmin/Journal_uploads/IJARET/VOLUME_15_ISSUE_5/IJARET_15_05_011.pdf

22. Karwa, K. (2024). The role of AI in enhancing career advising and professional development in design education: Exploring AI-driven tools and platforms that personalize career advice for students in industrial and product design. *International Journal of Advanced Research in Engineering, Science, and Management*. https://www.ijaresm.com/uploaded_files/document_file/Kushal_KarwadmKk.pdf

23. Köhler, C., Ehmke, J. F., & Campbell, A. M. (2020). Flexible time window management for attended home deliveries. *Omega*, *91*, 102023.

24. Konneru, N. M. K. (2021). Integrating security into CI/CD pipelines: A DevSecOps approach with SAST, DAST, and SCA tools. *International Journal of Science and Research Archive*. Retrieved from https://ijsra.net/content/role-notification-scheduling-improving-patient

25. Kumar, A. (2019). The convergence of predictive analytics in driving business intelligence and enhancing DevOps efficiency. International Journal of Computational Engineering and Management, 6(6), 118-142. Retrieved from https://ijcem.in/wp-content/uploads/THE-CONVERGENCE-OF-PREDICTIVE-ANALYTICS-IN-DRIVING-BUSINESS-INTELLIGENCE-AND-ENHANCING-DEVOPS-EFFICIENCY.pdf

26. Labrini, H. (2015). *Graph-based model for distribution systems: Application to planning problem* (Master's thesis, University of Waterloo).

27. Gunda, S. K. (2025). Accelerating Scientific Discovery With Machine Learning and HPC-Based Simulations. In B. Ben Youssef & M. Ben Ismail (Eds.), *Integrating Machine Learning Into HPC-Based Simulations and Analytics* (pp. 229-252). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-6684-3795-7.ch009

28. Meeker, W. Q., Escobar, L. A., & Pascual, F. G. (2021). *Statistical methods for reliability data*. John Wiley & Sons.

29. Miller, J., & How, J. P. (2017, May). Predictive positioning and quality of service ridesharing for campus mobility on demand systems. In *2017 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 1402-1408). IEEE.

30. Nucamendi-Guillén, S., Dávila, D., Camacho-Vallejo, J. F., & González-Ramírez, R. G. (2018). A discrete bilevel brain storm algorithm for solving a sales territory design problem: a case study. *Memetic Computing*, *10*, 441-458. https://link.springer.com/article/10.1007/s12293-018-0266-5

31. Nyati, S. (2018). Transforming telematics in fleet management: Innovations in asset tracking, efficiency, and communication. International Journal of Science and Research (IJSR), 7(10), 1804-1810. Retrieved from https://www.ijsr.net/getabstract.php?paperid=SR24203184230

32. Raju, R. K. (2017). Dynamic memory inference network for natural language inference. International Journal of Science and Research (IJSR), 6(2). https://www.ijsr.net/archive/v6i2/SR24926091431.pdf

33. Rapp, A. A., Petersen, J. A., Hughes, D. E., & Ogilvie, J. L. (2020). When time is sales: the impact of sales manager time allocation decisions on sales team performance. *Journal of Personal Selling & Sales Management*, *40*(2), 132-148. https://doi.org/10.1080/08853134.2020.1717961

34. Sam, B., Lavanya, M. G., Srinivasu, M. P., Venkataswamy, M. G., Choudary, M. S. K., & Avala, M. C. (2024). Work-Life Balance, Job Satisfaction, and Employee Retention: A Triangular Approach. *Corrosion Management ISSN: 1355-5243*, *34*(1), 276-286. https://corrosion-management.com/index.php/cm/article/view/91

35. Singh, V. (2022). Intelligent traffic systems with reinforcement learning: Using reinforcement learning to optimize traffic flow and reduce congestion. International Journal of Research in Information Technology and Computing. https://romanpub.com/ijaetv4-1-2022.php

36. Singh, V. (2023). Large language models in visual question answering: Leveraging LLMs to interpret complex questions and generate accurate answers based on visual input. International Journal of Advanced Engineering and Technology (IJAET), 5(S2). https://romanpub.com/resources/Vol%205%20%2C%20No%20S2%20-%2012.pdf

37. Gunda, S. K. (2024b). Machine Learning Approaches for Software Fault Diagnosis: Evaluating Decision Tree and KNN Models. 2024 Global Conference on Communications and Information Technologies (GCCIT).
https://ieeexplore.ieee.org/document/10861953

38. Sukhadiya, J., Pandya, H., & Singh, V. (2018). Comparison of Image Captioning Methods. *INTERNATIONAL JOURNAL OF ENGINEERING DEVELOPMENT AND RESEARCH*, *6*(4), 43-48. https://rjwave.org/ijedr/papers/IJEDR1804011.pdf

39. Trianasari, N., & Permadi, T. A. (2024). Analysis of product recommendation models at each fixed broadband sales location using K-means, DBSCAN, hierarchical clustering, SVM, RF, and ann. *Journal of Applied Data Sciences*, *5*(2), 636-652. http://bright-journal.org/Journal/index.php/JADS/article/view/210

40. Womack, J. P., & Jones, D. T. (2015). *Lean solutions: how companies and customers can create value and wealth together*. Simon and Schuster.

41. Zhan, X., Qian, X., & Ukkusuri, S. V. (2016). A graph-based approach to measuring the efficiency of an urban taxi service system. *IEEE Transactions on Intelligent Transportation Systems*, *17*(9), 2479-2489. https://doi.org/10.1109/TITS.2016.2521862