J. Umm Al-Qura Univ. Eng.Archit. (2023) 14:172–187
https://doi.org/10.1007/s43995-023-00029-x

ORIGINAL ARTICLE

The effect of data science on urban sustainability through the optimization of demand-responsive transportation

Raed Nayif A. Alahmadi[1] · Abdulaziz Alzahrani[2]

Received: 8 June 2023 / Accepted: 24 August 2023 / Published online: 13 September 2023

Abstract
The study investigated the viability of implementing a demand-responsive transportation (DRT) system within the premises of Al-Baha University to facilitate the punctual arrival of students to their classes and ensure their convenient departure from the campus. This study aims to demonstrate and elucidate how data mining enhances the efficiency of DRT systems. The case study investigates the potential improvements in DRT operations through the utilization of data mining techniques. The research employs descriptive and prescriptive techniques for mining trip planning data. The database maintained by the Deanship of Admission and Registration et al.-Baha University in Al-Baha City, Saudi Arabia, has identified 50 designated locations for student pick-up. The research establishes a comprehensive database that classifies pick-up points as spatial vectors, incorporating longitude, latitude, and counts of students. Cluster analysis and multidimensional scaling techniques are employed to minimize travel times and enhance operational efficiency by consolidating pick-up points. A completely adaptable DRT system prioritizes user satisfaction over operational efficiency, facilitating superior flexible services with fewer buses. The University Transport Department can implement Fixed Route systems in regions with significant population density and substantial demand. Conversely, in areas with a lower population, where Fixed Route systems may be less efficient, the department can opt for DRT systems. Additionally, the findings demonstrate the potential cost-saving benefits of implementing a hybrid system that integrates the university and off-campus student housing. The results indicate that providing fully flexible DRT services to students' harms efficiency and user experience. A comparative analysis between the Fixed route system and the DRT system reveals that the latter has the potential to offer enhanced service to students, utilizing an equivalent number of vehicles while reducing the number of rides required, all while maintaining a higher degree of flexibility in service provision. According to the recommendation of this study, it is advisable to implement a hybrid fixed route and DRT system. However, to assess the feasibility of such a system in a university setting in the Kingdom of Saudi Arabia (KSA), it is necessary to conduct a comprehensive viability analysis to identify any potential challenges or concerns.

Keywords Demand-responsive transportation (DRT) · Urban sustainability · Connectivity · Data mining · GIS

Abbreviations

DRT	Demand-responsive transportation
GHG	Greenhouse gases
UK	United Kingdom
CB	Customized bus
CO_2	Carbon dioxide
CO	Carbon monoxide
HC	Hydrocarbon
NO_2	Nitrogen oxide
PM	Particulate matter
KDD	Knowledge discovery in databases

✉ Abdulaziz Alzahrani
azahran@bu.edu.sa

Raed Nayif A. Alahmadi
rnaif@bu.edu.sa

1 Department of Civil Engineering, Faculty of Engineering, Al-Baha University, Al-Baha, Saudi Arabia

2 Department of Architecture, Faculty of Engineering, Al-Baha University, Al-Baha, Saudi Arabia

1 Research objectives

DRT is the most user-friendly among all the different types of public bus transportation systems; since their routes are flexible, the buses can off-schedule to pick up and drop off passengers, but whether this approach is cost-effective and efficient for their profit-making objective is yet to be determined. Therefore, this study adopts the Al-Baha University, Al Baha City, Saudi Arabia, as its case study by studying the application of data mining to improve the operational efficiency of DRT systems. Hence the proposed study has the following objectives:

Based on the purpose of this study, the objectives of the study are:

a. To conduct a literature review on previous studies about DRT to provide further understanding to the researcher regarding the role of urban design and its impact in enhancing the operational efficiency of DRT systems.
b. To identify the shortest distance for travel between DRT pick-up points.
c. To establish the impact of DRT on the economy and the environment.
d. To demonstrate a brief investigation of the current transportation system et al.-Baha University

2 The emerge of demand-responsive transport systems

DRT systems emerged by Cole and Merritt in 1968 as an urban transportation system that solves significant issues in parts of urban areas and small cities [13]. Urban transportation issues occur where travel demand is too small to support any transit service at all, and as such, it would serve as a hybrid between an ordinary bus and a taxi that picks up passengers at their doors or a nearby bus stop shortly after they have requested the service through the telephone. DRT started its popularity in the field in 1970 when The Massachusetts Institute of Technology (MIT) conducted the First Annual Conference on demand-responsive transportation systems to review various researchers' findings [6]

In the USA, the various forms of DRT identified earlier have evolved into the mainstream DRT as operated today in terms of the services offered, the management approach, and the techniques used to manage them, especially with the emergence of information and communication technology [16]. Extant studies have reviewed the operation of DRT systems in the USA. Navidi et al. extensively compared DRT and conventional public transport with a dynamic routing algorithm in an agent-based traffic simulation exercise [38]. They reported that replacing conventional public transport with DRT will improve mobility by decreasing passengers' perceived travel time without extra cost under specific situations. Rahimi, Amirgholy, and Gonzales [42] showed a continuum approximation model for analyzing the operating cost of DRT systems in large urban networks. The model highlighted how demand and the DRT system's characteristics influence significant cost components, including fleet, vehicle hours, and vehicle miles traveled. The results of their study highlighted a precise estimate of the cost of operation of the paratransit system in New Jersey. After achieving this result, they advanced a systematic technique for assessing the efficiency of policy implementations for DRTs [42].

According to Leiren and Skollerud [30], evidence from areas where the public transport authorities have replaced conventional bus services with DRT services shows that such new transport solutions are popular among users. Coutinho et al. [14] presented a case study of Mokumflex, a 12-month DRT pilot package that replaced conventional bus service in low-density areas of Amsterdam, the Netherlands. They leveraged their close relationship with the private enterprise that ran the service alongside the local bus operator. They performed an empirical pre-and post-operation comparison using selected intermodal indicators: travel distances, ridership, costs, Greenhouse Gases (GHG) emissions, and population's perception. They reported that the comparison of the pre-and post-introduction of the Mokumflex showed that ridership fell from 78.1 passengers/day to 15.9 passengers/day and based on the "demand- tailored" tailored nature, passenger-km reduced, even more, going from 1252.8 km/day to 136.6 km/day, this shows a reduction in costs and GHG emissions per passenger.

The population's perception reported that the system enjoyed a positive appraisal. Also, [15] explored the provision of DRT services in the United Kingdom (UK) through a national survey to identify the services that are working well and why. Their findings suggest increasing roles for voluntary and private sector stakeholders who will be able to introduce and efficiently manage smaller vehicles, which their study also showed to significantly influence the size of the operation (in terms of seats offered). Hence, vehicle size is an essential factor in the operation of DRT in rural areas in the UK.

Countries in Asia have also adopted a DRT system where conventional public transport was not feasible. For example, China introduced a DRT system called Customized Bus (CB) in 2013 in Qingdao, Beijing. The CB was regarded as an innovative mode of DRT systems that provides advanced, attractive, and user-oriented service to a specific type of commuters by combining their travel-demand pattern using online information platforms, such as the Internet, telephone, and smartphone [54]. Now, the CB runs in all cities in China. For example, Beijing launched 164 CB lines that run on online reservations on Thursdays.

Although inadequate data shows that DRT is being operated in Africa and demands responsive transport in Africa [36], noted that DRT service (Jitney) had been used since 1926 in South Africa as a kombi taxi. They explained that despite the long history of the Jitney, there had been little formal theoretical work on the subject. However, Coetzee et al. [12] presented a demand and supply model that uses the current passenger demand in the service network to calculate the current operations' total operating cost, revenue, and carbon footprint and then compared the results with the optimized and rationalized operations. The comparison showed that the DRT model potentially contributes to significant operating cost reduction, reduces carbon footprint, and increases profitability by running more efficiently than any other alternative. Demand-responsive transport represents a reliable and efficient alternative to public transport to help reduce urban traffic congestion, improve traffic safety, and alleviate energy consumption and greenhouse gas emission problems worldwide.

3 The impact of demand-responsive transport system on urban sustainability and connectivity

An urbanized society cannot be imagined without constant communication and movement. The relatively new concept of "sustainable mobility" symbolizes a person's freedom in his spatial movement, which does not harm the environment and other people's health [33]. Furthermore, the interests of future generations are considered when demand-responsive transport systems are designed to satisfy the needs of urban citizens regarding connectivity and sustainability.

Globalization and large-scale urbanization, combined with technical innovation and the digital revolution, are rapidly changing the world around, the structure of the economy, and the way people live [10]. The transition process to an economy of a new technological order will intensify in the foreseeable future. As a result, a comfortable urban environment, the quality of which is one-third dependent on mobility conditions, becomes a more valuable capital and a source of economic growth [21]. At the highest political level, urbanization is recognized as a significant positive trend, which means a reassessment of the role of cities in the economy and their contribution to social and innovative development.

In turn, cities consume 60–80% of all energy produced globally and are the main culprits in climate change. The transport sector is responsible for 23% of greenhouse gas emissions, with most urban travel [37]. If climate change had been considered an environmental problem a few decades ago, civilization would be concerned about this global threat.

The complex nature of recommendations in urban planning activities allows architects to consider national and local priorities and conditions. It predetermines the need beyond sectoral and administrative boundaries [31]. The main goal of the city authorities is not to manage the economic complex but to take care of people, including those with limited physical and financial capabilities.

Habitat III calls on the world community to more actively form special legislation focused on cities, use modern tools for integrated planning, maintain a balance of competencies and responsibilities of local authorities, and strive to ensure the financial sustainability of urban development [21]. A decade ago, an UN-Habitat report assessed urban planning as a means of addressing the challenges facing cities in the twenty-first century. A well-reasoned conclusion was drawn: the existing planning system is often a source of urban problems, as it is not adequately aimed at improving the environment and people's living conditions [22]. Furthermore, it is noted that the most apparent drawback of the developed master plans is that they cannot make the city comfortable for life and attractive for businesses and investors because they are not sufficiently integrated into the city management system [26]. In addition, they operate separately from city services and the budget process [23]. Consequently, it is necessary to change the existing approaches to urban planning to increase its responsibility for implementing the SDGs and ensuring sustainable urban development.

A balanced distribution of residences of people and social and economic activity centers is the basis for the sustainable development of the city and its transport system. This approach ensures people's movement and forms the city's architectural appearance [53]. Moreover, it helps regulate the use of the territory, ensuring its connectivity and permeability [58]. Implementing the concept of transit-oriented design, including urban development territories, taking into account access to public transport, and intelligent architecture of demand-responsive transport systems can reduce the ecological footprint, curb the process of urban sprawl, and their dependence on automobiles [20]. This approach will make cities greener and more livable.

In many countries, transport authorities are increasingly embracing a new integrating role beyond traditional sectoral boundaries to encompass mobility management in general. They are engaged in creating conditions for cycling and walking and developing new types of services, including short-term rental of cars, bicycles, scooters, and hoverboards, along with others [28]. In addition, the transport sector is forced to adapt to a rapidly changing environment, including people's expectations and technological innovations and the implementation of digitalization in all areas of activity.

At the same time, the management paradigm is changing. There is a transition from making industry decisions to integrated transport and urban planning, focused on the needs of people and the city, increasing its investment attractiveness. Therefore, in European countries, more and more attention is paid to integrating public transport into the urban environment [55]. Transport organizations often become public space planners, not just mobility service providers [39]. A comfortable city for life implies a wide range of personalized mobility services integrated with traditional modes of public transport based on digital technologies, mutual exchange of information and innovative solutions, knowledge, and experience working with passengers.

4 The value of demand-responsive transport systems

DRT systems are a form of collective travel alternative that is flexible and potentially complements the fixed transportation system [1]. As part of the solutions to problems faced by urban planners, DRT systems have contributed highly to the mobility needs of the population that lives in the geographical area where they have been introduced and have impacted these communities in different ways. According to Ronald et al. [44], DRT systems can decrease congestion, reduce the environmental effects of transport, and provide different transportation options to communities that are limited in their travel options. Van Oort et al. [52] proposed the 5Es Framework, which examined the value of public transport based on economic, environmental, and social impacts, including its effectiveness, which refers to safely and reliably transporting people while reducing congestion; and efficiency, which refers to judicious use of limited resources. However, DRT systems develop a special transport arrangement to serve a low-density and relatively dispersed population where fixed pick-up public transportation is not financially feasible [1].

When the 5Es Framework is extended to DRT systems, it is apparent that they are a cost-efficient alternative for densely populated rural areas from an economic perspective. Although costs may increase initially at commencement, the operating costs have been identified to decrease with time as the number of passengers increases [11]. Studies evidence this in DRT's system performance and productivity. For example [19], investigated the effects of dynamic and stochastic variations in travel time on scheduling reliability and system performance. They performed various experiments on a DRT system in Edmonton, Canada. Their study showed that dynamic and stochastic variations in travel times significantly affected the quality of the schedules. Hence, the travel time variation in routing can also be counted to achieve a more reliable and efficient system. Similarly, Schlüter et al. [46] evaluated the practicality of DRT systems in rural and urban areas. Their results showed that automated DRT systems could reduce transportation's economic and environmental costs in the interaction of rural and urban areas.

Regarding road safety, Brake, Nelson, and Mulley [7] suggested considerable opportunities to apply DRT systems to reduce road traffic and increase road safety, thereby saving operating costs while increasing the level of service provided to passengers. DRT systems contribute significantly to road safety. Coutinho et al. [14] reported in their study that the examined DRT system could save more than 85% of expenditure on road safety. DTR was essentially due to the reduced mileage and operating timeframe, which increased the system's overall efficiency compared with the previous system being examined. Hence, DRT systems continue to witness increased patronage due to flexible, customer- and demand-oriented business models, encouraging increased ridership [25].

From an environmental perspective, DRT systems have been recommended as essential solutions to reducing pollutant emissions caused by vehicles in transportation [2]. Prud'homme et al. [41] suggested that vehicles in traffic have various types of environmental effects. They explained that vehicle flows cause noise and visual impacts on neighborhoods due to their fuel combustion natures. This includes cars, trucks, and other road vehicles that emit pollutant substances such as carbon dioxide (CO_2), carbon monoxide (CO), hydrocarbon (HC), nitrogen oxide (NO_2), and particulate matter (PM). Hence, Prud'homme et al. [41] investigated the impacts of DRTs on pollutant emissions by running several simulations. The simulation results showed that the deviation time supports the grouping rate and decreases the vehicle quantity and pollutant emissions.

Concerning travel behavior, DRT has a significant impact on travel behavior based on the study reported by Ryley et al. [45], which showed that airport and station access DRT services recorded increased market share and showed more potential than the employment shuttle. However, DRT systems must compete with the car in the domain under study, and the success of cars as alternatives depends on available parking spaces. Hence, DRT systems can positively impact travel behavior due to the potential for reducing the number of vehicles on the road through vehicle sharing, reducing carbon emissions, and improving air quality within the home and work environments [1].

5 Data mining process

Data mining is the process of discovering patterns and extracting valuable information within datasets, data warehouses, or other information repositories; while data mining and knowledge discovery in databases (KDD) are frequently treated as synonyms, Knowledge Discovery in Databases is actually an iterative process comprises of a few steps leading from raw data collections to some form of new knowledge [9]. Zaïane [56] explained that the KDD process involves data cleaning, integration, selection, transformation, mining, pattern evaluation, and knowledge representation. This actually means that data mining is a part of the knowledge discovery process.

Fayyad et al. [17] show that data mining mainly performs descriptive, classification, and predictive tasks in a dataset. The descriptive task includes Class Description, Mining of Associations, and Mining of Clusters. According to Fayyad et al. [17], the Class Description refers to the relevant data to be associated with user-defined concepts such as discrimination or characterization that are normally retrieved by a data query through a comparison or summarization module to extract the essence of the data at different levels of abstractions. For instance, in a company, the classes of items for sale include cartons and bottles, while the customer summary will include England-based and Wales-based and can be called from the database [17].

Mining of Association is used to identify patterns using the association and correlations rule to uncover interesting statistical correlations between associated attribute-value pairs or between two item sets to analyze if they have positive, negative, or no effect on each other [24]. Finally, the Mining of Clusters refers to creating a group of very similar objects that are highly different from those in other clusters [43]. They explained that clustering is also called unsupervised classification because class labels do not dictate the classification. Instead, it includes the identification of attributes that are used to separate a set of objects in a certain way. Meanwhile, the objects in the same group are more similar than those in other groups regarding specific attributes [18]. An example of this location is a road network where the request for taxi services is usually not satisfied.

Classification and prediction are one of the most commonly performed tasks in data mining; the classification analyses the attribute in a dataset to classify future objects and develop a better understanding of the classes of the objects in the database, classification approaches normally use a training set where all objects are already associated with known class labels, and the classification algorithm learns from the training set and builds a model [59]. According to Zheng [59], this method mainly uses algorithms such as regression, Bayesian network, K nearest neighbor, support vector machine, and even neural networks for analysis. For example, the request decline in a taxi service can be investigated by generating the level of decline in different areas in a road network; this can also produce weather conditions, traffic volume, type of vehicles, etc., that are also related to the area, this can produce the trends and normalities that can be extended to prediction for future use. Predictions are similar to classifications, but data are classified based on behavior or values predicted in the future. Examples of predictive tasks, for example, are to predict a reduction in the number of customers in the near future and stock price predictions in the next three months [34].

According to Chamatkar and Butey [9], there are two major types of predictions; the first one involves trying to predict some unavailable data values or pending trends, or predict a class label for some data, while the second involves classification. Once a classification model is built based on a training set, the class label of an object can be foreseen based on the attribute values of the object and the attribute values of the classes. Other types of tasks in data mining include Outlier and Evolution analysis, and Outliers are data elements that cannot be grouped in a given class or cluster [4]. According to Arnold and Beyer [4], these are usually considered noise and discarded in some applications and can reveal important knowledge in other domains that are not mainly related to the course of study. Outlier analysis may be applied to investigate fraudulent use of credit cards by adopting methods to help detect purchases of suspiciously large amounts of a product for a particular account number as compared to regular purchases made by the same account [24].

On the other hand, evolution analysis refers to studying time-related data that changes in time. It mainly highlights trends in data, which consent to characterizing, comparing, classifying, or clustering time-related data [50]. Although data mining has been described as a process in KDD, Padhy et al. [40] described the steps involved in data mining as including the following;

1. Understand the application domain: A proper understanding of the application domain is necessary to appreciate the data mining outcomes desired by the user. It is also essential to assimilate and take advantage of available prior knowledge to maximize the chance of success.
2. Collect and create the target dataset: Data mining relies on the availability of suitable data that relate to the underlying variety, order, and structure of the problem being analyzed. Therefore, the collection of a dataset that captures all the possible situations that are relevant to the problem being analyzed is crucial.
3. Clean and transform the target dataset: Raw data contain many errors and inconsistencies, such as noise, outliers, and missing values. An essential element of this process is the de-duplication of data records to produce a non-redundant dataset, for example, collecting information from public sequence databases such as raw CCTV footage, satellite pictures, or GPS coordinate of vehicles.
4. Select features, and reduce dimensions: Even after the data have been cleaned up in terms of eliminating duplicates, inconsistencies, missing values, and so on, there may still be noise that is irrelevant to the problem being analyzed. These noise attributes may confuse subsequent data mining steps, produce irrelevant rules and associations, and increase computational costs. It is, therefore wise to perform a dimension reduction or feature selection step to separate those attributes that are pertinent from those that are irrelevant.

5. Apply data mining algorithms: Now we are ready to apply appropriate data mining algorithms, association rules discovery, sequence mining, classification tree induction, clustering, and so on, to analyze the data.
6. Interpret, evaluate, and visualize patterns: After the algorithms above have produced their output, it is still necessary to examine the output to interpret and evaluate the extracted patterns, rules, and models. We can only derive new insights into the analyzed problem through this interpretation and evaluation process. The final phase is visually representing the discovered knowledge to the user. This essential step uses visualization techniques to help users understand and interpret the data mining results.

These steps were highlighted by Zaki and Wong [57], as highlighted in Fig. 1, and they explained that the data mining process outspreads the functionality of just collecting data from a source. It also includes all the processes that have been mentioned. Hence it can be concluded that data mining's primary function is encircled in knowledge discovery.

Zaki and Wong [57] explained that the goals of data mining are essential to achieve reliable predictions which can be descriptively presented understandably (Fig. 1). This research will strive to achieve this objective.

6 Methodology

Demand-responsive transport has always been attributed to improving the available transportation system. This study intends to propose a destination-specific demand transport system in a university environment to specifically serve female students at the University of Al-Baha, Saudi Arabia, and reduce the travel time and cost of transportation for the students that attend the university from their various homes. Students' data was collected from the university database to identify their residences and potential pick-up locations to determine pick-up times based on students' requests. The platform would serve as a database for operating the demand-responsive transportation service. From the submitted request by the interested student, a routing plan would be created to provide an adequate stop for picking up the student. The methodology that is intended for this work is elaborated below.

7 Problem formulation

The project's initial phase focuses on understanding the project objectives and requirements of the problem. Before building the model, data will be assembled to establish a relationship between all the attributes that can be deducted from the results. It starts with defining the parameters involved in the study and describing the mathematical notations, objective function, and constraints involved [8]. Hence the goal is given as a set of demands to optimize the routes of the different vehicles of the fleet according to objective functions that will be discussed in this next chapter. Therefore, the objective is to solve several optimization problems with different values of the parameters to identify for which of their values the objectives are reached [47].

These objective functions involve the arrival times, walking time, and cost from the consumer perspective, while in the operator's perspective, it may include waiting time and maximizing the operation's profit. Constraints are the factor in the transport system that renders the operation infeasible. This includes vehicle capacity, the time each stop needs to be visited, and connectivity between the vehicle depot and student location. This approach was adopted by Tom and Mohan [51] to resolve the selection of routes and associated frequencies to achieve a desired objective within specific

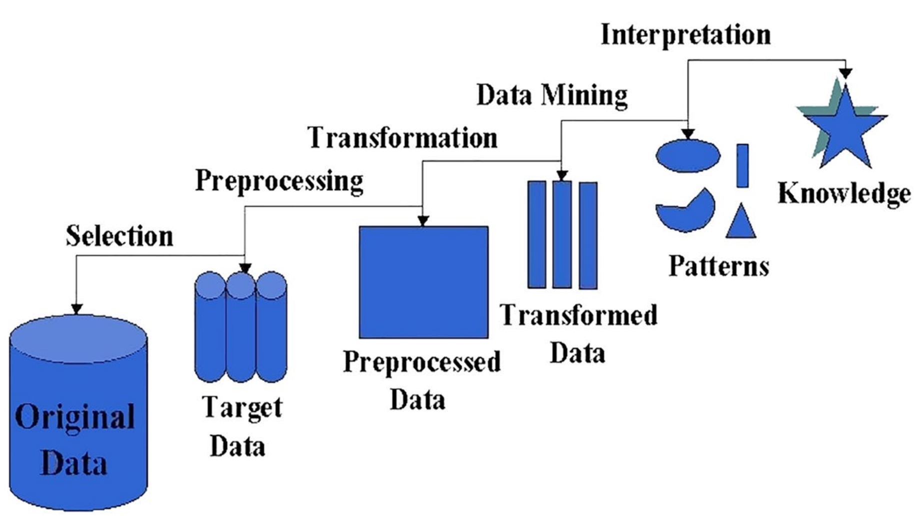

Fig. 1 The steps in data mining. Adapted from data mining techniques. WSPC/lecture notes series [57]

operational constraints. The success they achieved has guided the researcher to adopt this approach.

8 Data collection

Data collection has become ever more dynamic than in the past wherein researchers rely on manual collection such as questionnaires or mechanical devices such as loop detectors, magnetic sensors, image processors, laser sensors, or magnetic sensors [29, 32]. Researchers continuously collect, monitor, analyze, summarize, and visualize relevant information from social interactions and user-generated content in various domains such as business, public administration, politics, or consumer decision-making [27, 49]. In the last decade, sources for data collection have expanded from the conventional questionnaire and interviews to include non-conventional forms such as big data that are collected with the aid of Information and Communication Technology (ICT) systems which involve extracting data from databases or harvested from real-life systems to measure specific variables [45]. Examples of databases include sales data, market data, company raw materials database, university database of students in each department, and real-life sources such as social media platforms. Social media has experienced tremendous growth in the past decade and has become a veritable source of research as there are over 2.7 billon active users on the Facebook, Instagram, YouTube, WhatsApp, and other platforms [48].

These ICT-enabled data collection methods have also been extended to research in demand-based mobility [35], which is getting increased attention worldwide. Hence, data for this study has been collected from the University of Al-Baha student registration database. The student data collected identifies students by their Student ID, their department, the city where they live, including the area and the street, and the start and end times of their days at the university. This data will be simulated in order to achieve the objectives of the research, which are:

a. To provide further understanding regarding the role of data mining in enhancing the operational efficiency of DRT systems.
b. To demonstrate how data mining can be applied to develop an effective and efficient DRT system.

9 Data analysis

The simulation of DRT systems is essential to ascertain the start time, origin, and destination of each request, including the capacity and positioning of each vehicle in the fleet [3]. The data collected from the University of Al-Baha student registration database will be stimulated, and the results will be interpreted. The data was submitted to Via Transportation, Inc platform, which applied its Transit Assessment Methodology and data mining algorithm to the data, and the results were analyzed. This would enable the design of suitable time-table and routes for picking up students at predetermined stops. The data analysis would also be extended by predicting students' waiting and arrival times during their travel times.

Hence, the procedure for data analysis in this study is as follows:

Discovering data refers to the study's data analysis to identify existing trip patterns between student residences and the campus and a review of Al-Baha University's service quality.

1. Data discovery: This refers to the analysis of the data collected for this study to identify existing trip patterns between the student's residences and the campus; it also includes implementing a review of the quality of service that Al-Baha University can provide.
2. Generate demand scenarios: Identify the factors in the data to generate estimated demand patterns.
3. Service design and simulations: Set the required parameters that are of value to the Al-Baha University, and then implement the simulation.
4. Iteration and optimization: Review the simulation results and conduct a sensitivity analysis to highlight the results of varying scenarios based on changing parameters. Conclusions will be drawn from recommendations (Via Transportation Inc, 2021).

10 Case study: Al-Baha University

Al-Baha University was founded in 2006 to ease the burden on big metropolis institutions and serve society. As a result, the demand for higher education in Saudi Arabia's south has grown. Al-Baha University comprises 16 faculties and colleges in Al-Mandaq, Al-Makhwah, Baljurashi, and Qilwah. The main campus is 3 km2 and connects to King Fahd roads in Al-Aqiq (Fig. 2).

23,080 students, including 264 international students, attend Al-Baha University. There are also 1706 academic staff members (Top Universities, 2020). The institution does not give students housing. Students must thus either locate local housing or travel every day. To increase campus mobility, the institution requires a comprehensive transportation infrastructure that includes the DRT.

Al-Baha University provides transportation exclusively for female students traveling from several governorates of Al-Baha's to the university's colleges spread throughout the region. The university's transportation department identified

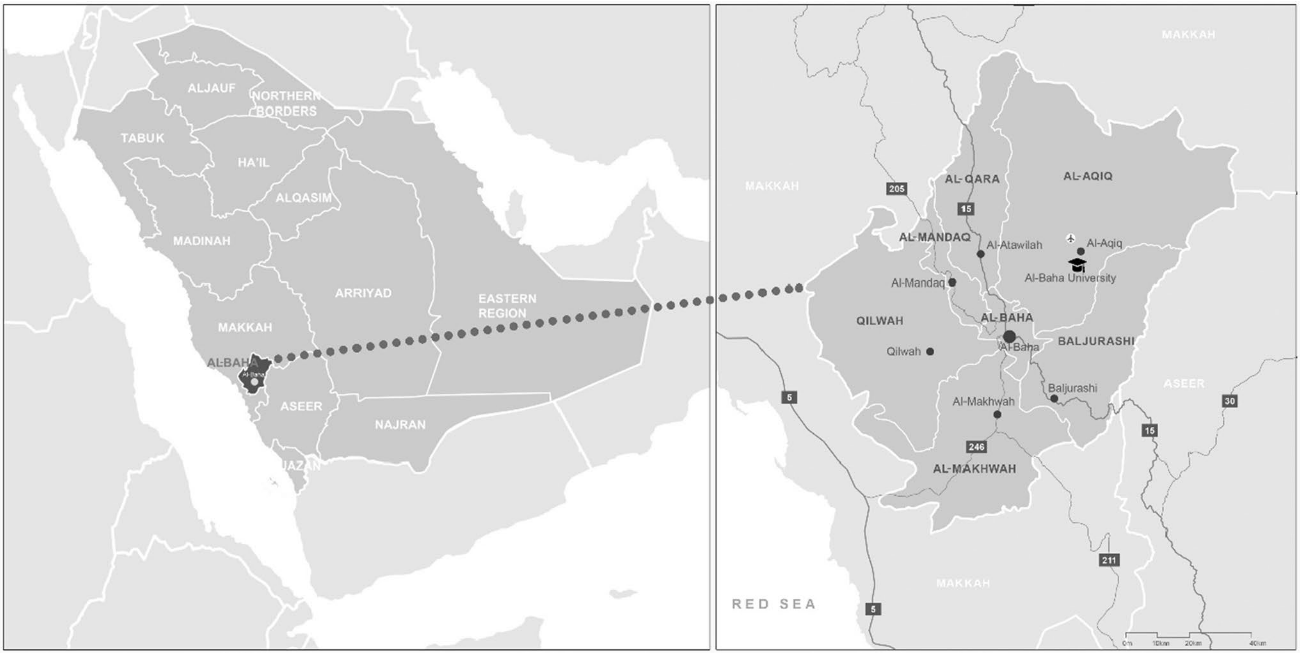

Fig. 2 Al-Baha is located in the southwest of Saudi Arabia. Al-Baha University is towards the northeast side of Al-Baha region, near the airport. Adapted from Al-Baha Report (2022)

several deficiencies in the university's transportation system in a report. The report identified university transportation system weaknesses.

- As per the contract, there are fewer minibusses than 2016 models.
- Frequent bus breakdowns
- Maintenance delays
- Buses lack university logo stickers
- Lack of driver parking at bus stops
- Spare parts delays

The research recognized through the provided reports that absences are a significant problem since they impact students' academic performance. The actual and theoretical mismatch, which refers to differences between the reported number of bus rides and the number taken, was also noted in the study. Buses are sent inconsistently to various sectors, and beneficiaries are distributed unevenly. As a result, there are variations between the actual and reported numbers of beneficiaries. The current cost estimate was derived from the contract for transporting only female students of Al-Baha University in 2019. The number of beneficiaries of transportation is 3987 female students who were able to get to their destinations through the transport service. According to the terms of the agreement, the total cost of the contract reached with the company that provides transport services was SAR 6.3 million ($1.67 million). It covered the operation of 84 buses during the academic year.

In light of this, it is necessary to remedy the report's shortcomings by establishing a comprehensive transportation infrastructure that serves both on-campus and off-campus students. This should include innovative technologies and apps that improve and optimize transportation.

11 Findings

Based on the study's objectives, a problem would be formulated involving students requesting transport pick-up from the University of Al-Baha, where it is proposed that the institution's transportation operator provide vehicles to transport students between lecture halls and their residences. Each bus uses the same route to pick up and drop off students at a specific time. Each bus has a maximum capacity and travel distance. No partial pick-ups are allowed at each pick-up or drop-off point.

This is similar to the school bus vehicle routing difficulty [5]. It involves picking accessible bus stations, assigning students to available buses, and planning bus routes. The journeys must be independent, except at the depot node, and the number of students on a single route cannot exceed bus capacity. Vehicles do not have to return to the central depot after their last customer. Those who return must visit the same consumers in reverse order. Therefore, a single origin point has several pathways, not a collection of points. The university transport department's report highlights several challenges, including those bus models manufactured

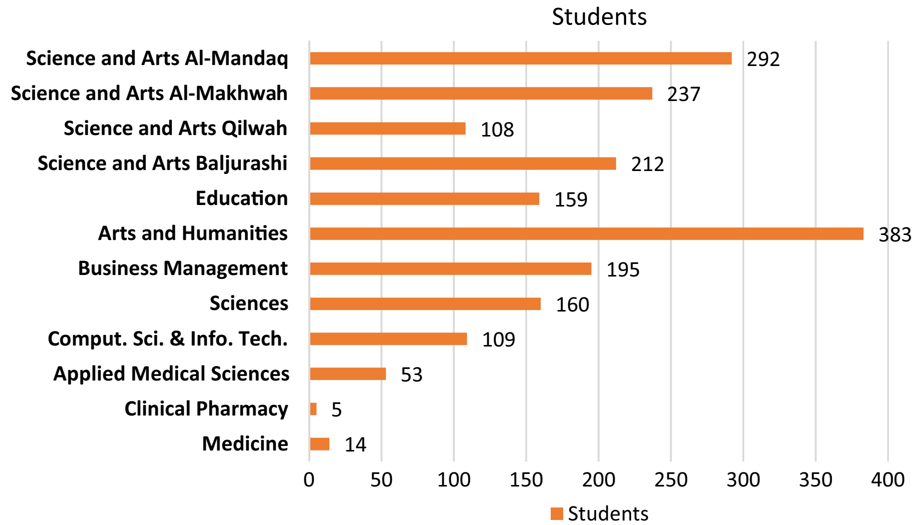

Fig. 3 Composition of Students who commute to the university by bus in Al-Baha University Colleges

before 2016 are considered a weakness, the requirement for an increased number of minibuses, and the frequency of bus breakdowns. There are delays in the completion of maintenance work, the delivery of spare parts, and the requirement for additional parking spaces for drivers in the areas frequented by buses.

Students in this study must attend lectures et al.-Baha University to fulfill course requirements. Currently, the institution operates a Fixed Route Service with 78 of 85 contracted vehicles at the cost of SAR 6961, which is equal to $ 1855.74.

In the rest of this section, the research question will be answered. How does data mining in DRT optimize the operations of a DRT company that ensures Al-Baha University students arrive on time and efficiently?

This study used data from Al-Baha University's student database on 1,927 students offering various courses. The data were analyzed with *ViaViewer*™ (Via, 2020), a data mining software developed by Via Transportation, Inc. in New York, United States. Analysis of data with *ViaViewer*™ reveals the following about the data: The characteristics of the data are based on the following headings: Student ID, College, City, Street, Semester, reference number, Day (day of the week they have lectures), Lecture Start Time, and Lecture End Time.

12 Discovering data

As shown in Fig. 3, each college's makeup within the university is represented. Figure 4, which depicts the start and end times for each class, and Fig. 5, which depicts the end times for each class, reflect the main conclusions from the characteristics of the data based on the start and end times.

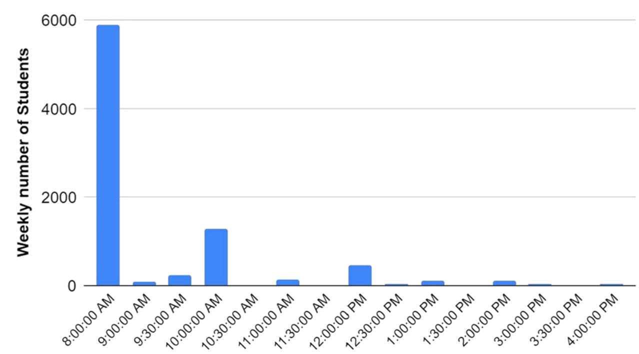

Fig. 4 Class start times. Adapted from Via. (2020). Introducing Via-Viewer™: Our sophisticated public transportation modeling engine. New York, NY: Via Transportation Inc

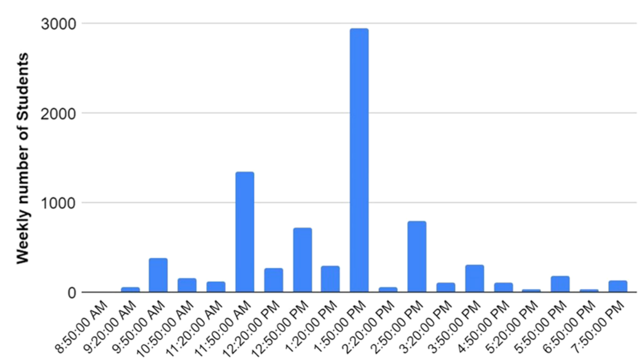

Fig. 5 Class end times. Adapted from Via. (2020). Introducing Via-Viewer™: Our sophisticated public transportation modeling engine. New York, NY: Via Transportation Inc

The main conclusions drawn from the data based on the start and end times shown in Figs. 4 and 5 are as follows:

1. Demand for transportation is consistent at the beginning and end of classes every day of the week, excluding weekends, because a stable student population needs to travel to campus.

2. Most classes begin at 8 am, with a smaller number beginning at 10 am.

3. Most students finish their classes around 2:00 pm, with a significant number finishing at 1:00 pm.

13 Data scenarios

A close look at the data reveals that daily demand is not variable. As a result, Monday, the day with the highest demand, is chosen as the Simulation Day, and as a result, a typical Monday morning time is chosen for the Simulation (Fig. 6). In addition, two arrival times were established: 8 am and 10 am, to allow for flexibility and to guarantee service level effectiveness.

14 Service design simulations

For this study, two service designs—the Fixed Route design and the DRT design—were both simulated. For the Fixed Route design, the University of Al-Baha's demand requirements are taken into account, the booking method is based on a Fixed Route Model, the arrival hours are between 8 and 10, the number of vehicles is not anticipated to exceed 78, and walking distances are covered by Stop-to-Stop service. Also, the following parameters were established for the *ViaViewer*™ algorithm by Via Transportation Inc. in 2021:

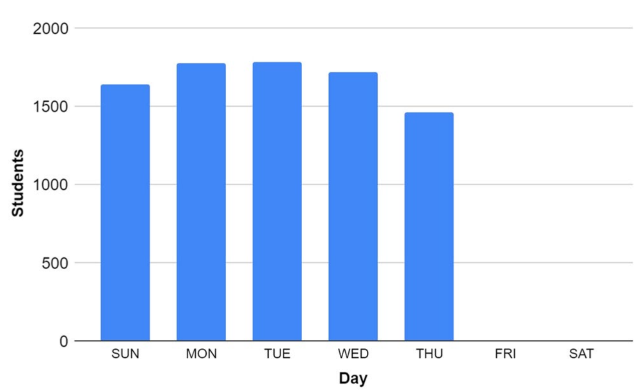

Fig. 6 Low demand variability: Monday is a high-demand day. Adapted from Via. (2020). Introducing ViaViewer™: Our sophisticated public transportation modeling engine. New York, NY: Via Transportation Inc

- Maximum drop-off arrival window: 30–60 min
- Two hundred meters is the intended average walking distance.
- Ten minutes is the intended average detour time.

Regarding the DRT design, the service hours are between 7:00 am and 2:00 am, the vehicle supply should not exceed 85, and the walking distances are Corner to Corner service. The following parameters were established for the *ViaViewer*™ algorithm by Via Transportation Inc. in 2021 (Fig. 7):

- Maximum permitted drop-off arrival time: 45 min.
- Maximum pick-up lateness permitted: 45 min.
- Two hundred meters is the intended average walking distance.
- Ten minutes is the intended average detour time.

As shown in Figs. 7 and 8, the locations of pick-ups are represented by blue dots, and orange dots represent the locations of pick-up requests. The dots in the grey area indicate upcoming reservations for which payment has not yet been received. The black dots represent vehicles, and the length of their "tails" indicates their travel speed.

Table 1 summarizes the simulation exercise's results; two simulation outcomes are presented: Option 1 and Option 2. First, there is significant and stable demand for the DRT system, with simulated demand of between 1500 and 1600 students for Option 1 and Option 2, even though an actual daily demand of 1450 students is assured. The average detour time, a measure of service quality, is between 5 and 10 min for both alternatives, while the average trip duration between pick-up and drop-off is between 50 and 60 min. Finally, the buses' peak utilization rate is 7–8 rides per vehicle hour.

Option 1 and Option 2 are both reasonable solutions for the Al-Baha University Transportation Department to adopt because students arrive on campus at least 30 min before courses begin, as shown in Fig. 9. While Fig. 9 shows the drop-off time, Fig. 10 shows the classes start time, and closer inspection reveals that students whose class start time is 8.00 am were dropped off at 7.30 am, while those whose class start time is 9.00 am were dropped off at 8.00 am, giving them an hour to prepare for their classes.

Figure 11 illustrates the number of students in one vehicle at the start of the day, and an analysis of the graph reveals that some vehicles have very few students at a time, implying that Al-Baha University should use smaller vans or ride-hailing services to reduce the cost of operating a fixed route service.

Based on the different fixed routes within the parameters defined for the students' home and university addresses, the *ViaViewer*™ algorithm successfully used 68 vehicles to service two-morning options for all students

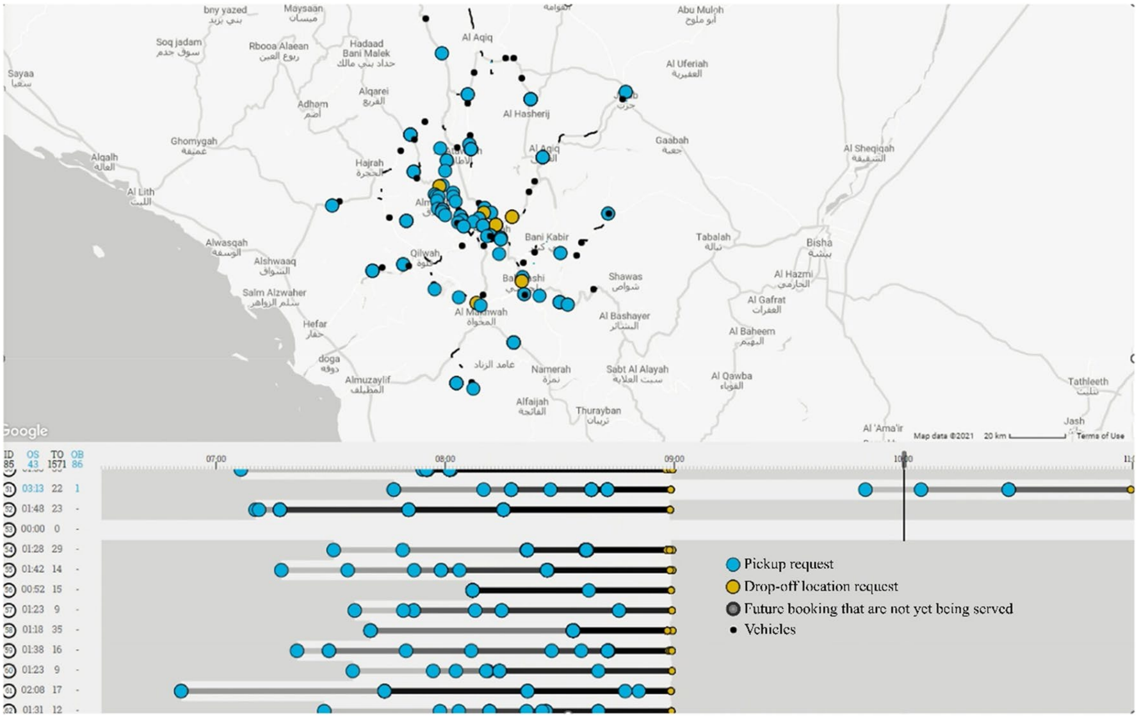

Fig. 7 The simulation demonstrates how the arrival on the fixed route took place at 8:00 and 10:00 in the morning. Adapted and run through Via Algorithm. (2020). Introducing ViaViewer™: Our sophisticated public transportation modeling engine. New York, NY: Via Transportation Inc

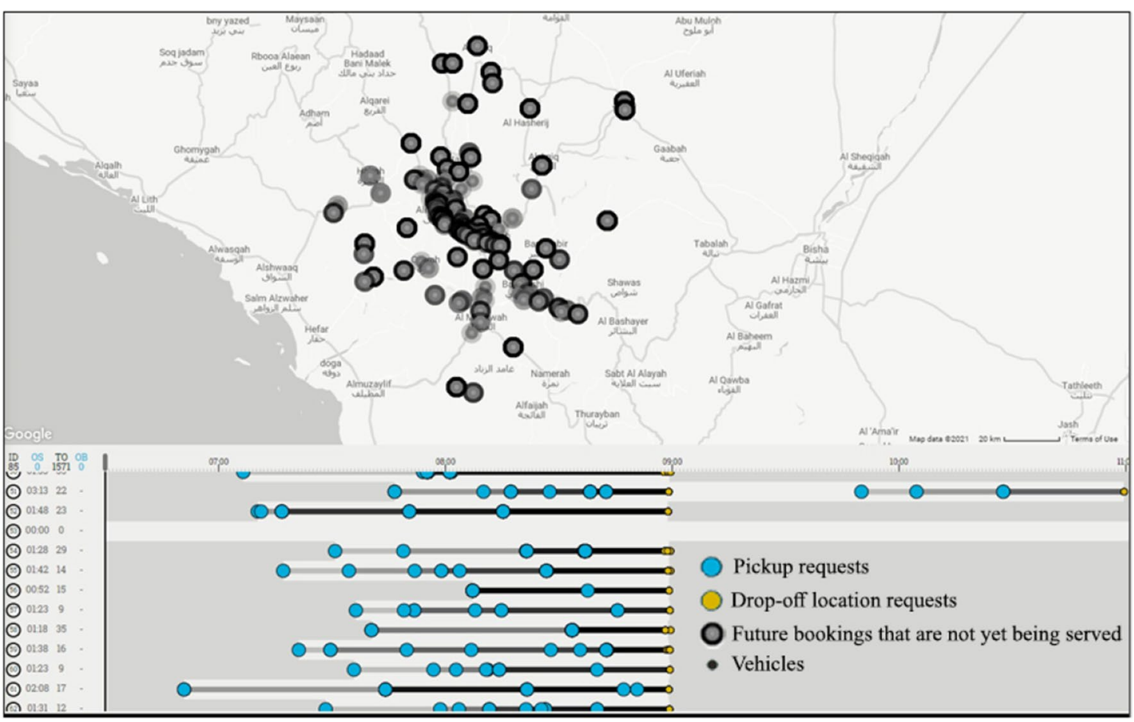

Fig. 8 The simulation demonstrates how the fixed route booking for service begins in the morning and how the buses must pick up students as a stop-to-stop service to complete their routes. Adapted and run through Via Algorithm. (2020). Introducing ViaViewer™: Our sophisticated public transportation modeling engine. New York, NY: Via Transportation Inc

Table 1 Iteration and Optimization: Simulation results

Particulars	Simulation features	Option 1	Option 2
Supply	Peak vehicle count	74	68
Demand	Total rides served in AM shift (Actual demand: 1450)	1500–1600	1500–1600
Quality of service	Maximum allowed drop-off arrival time in minutes	30	60
Quality of service	Average detour time in minutes	5–10	5–10
Quality of service	Average trip duration (pick-up to the drop-off in minutes)	50–60	50–60
Efficiency	Peak utilization rate (rides/vehicle-hour)	7–8	7–8

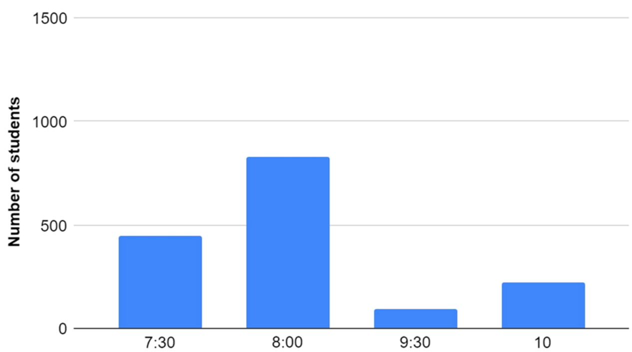

Fig. 9 Drop-off time. Adapted from Via. (2020). Introducing ViaViewer™: Our sophisticated public transportation modeling engine. New York, NY: Via Transportation Inc

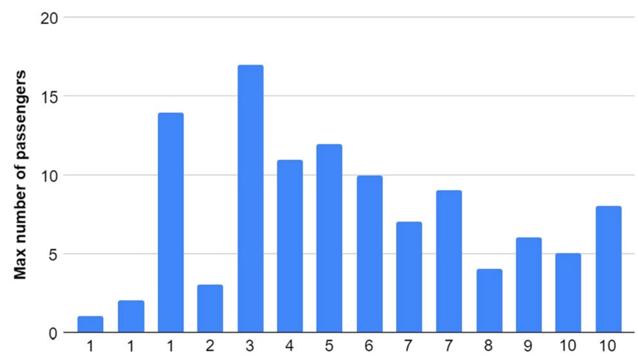

Fig. 11 Number of passengers per bus. Adapted from Via. (2020). Introducing ViaViewer™: Our sophisticated public transportation modeling engine. New York, NY: Via Transportation Inc

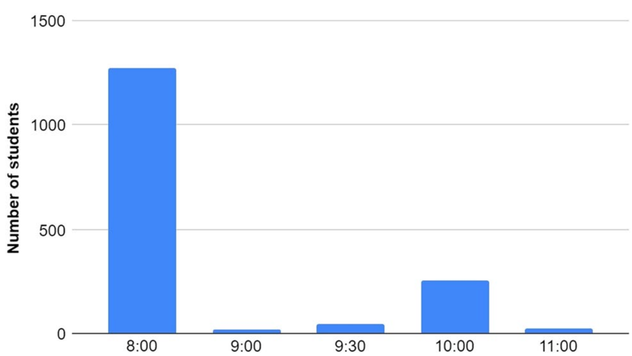

Fig. 10 Classes start time. Adapted from Via. (2020). Introducing ViaViewer™: Our sophisticated public transportation modeling engine. New York, NY: Via Transportation Inc

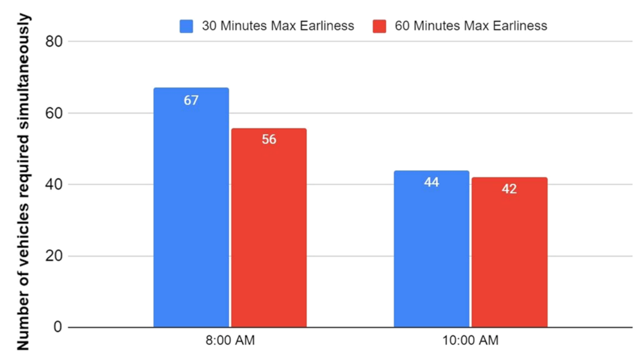

Fig. 12 Arrival before the commencement of classes. Adapted from Via. (2020). Introducing ViaViewer™: Our sophisticated public transportation modeling engine. New York, NY: Via Transportation Inc

traveling to the university, representing a significant reduction from the current service. However, it is crucial to note that some cars from the 8.00 am shift will not be able to make a round journey, that would allow them to be available for the 10.00 am shift. Therefore, more vehicles will be needed for the 10.00 am shift than for the 8.00 am shift (Fig. 12). Given that demand created by students from the university to their residences has a comparable distribution but slightly lower peaks, it can be concluded that the afternoon shift will not require more vehicles than the morning shift.

The study outlined the necessary improvements that should be made to the way that student transport is handled. This study is based on the current situation in the university transfer of students based on fixed routes and timing. The success of this application is dependent on two primary factors: the first of these is spatial consideration, and the second is organizational consideration. The location of the student's residence and the venue where the lectures will take place are the elements that make up the space. When it comes to the particulars, these include the time of the classes, which is determined by certain factors such as the appropriate time to arrive before the class begins and the right time to stay after the class has ended. The variation in these variable variables directly impacts the number of buses utilized for students' transportation.

Regarding the financial aspect, the existing Fixed Route service includes 85 contracted cars, of which 78 provide services, and the monthly operating cost per vehicle is $1853. Table 2 summarizes the Fixed Route scenario's financial simulation results. In Option 1, fewer buses (74) than are currently run (78) can be deployed with a monthly savings of $7412, while Option 2 highlights a more efficient operation running with far fewer buses (68) with a monthly savings of $18,530, with a total monthly technical fee of $18,700 as compared to $20,350 for Option 1 representing a monthly savings of $1870.

This study does not consider other potential direct and indirect cost savings because doing so would go beyond the scope of the research and call for additional investigation into topics such as the following:

- Seats that can be used by users other than female students (female employees, for example), in addition to female students.
- Measuring the reduced carbon emissions produced by buses.
- Estimation of cost reductions related to road maintenance and expansion of capacity.
- An evaluation of the university's parking lots' reduced capacities concerning occupancy.

Table 2 Financial implications: simulation results

Particulars	Option 1	Option 2
Supply of buses	74	68
Buses reduced	4	10
Monthly savings in $	$7412	$18,530
Monthly technical fee/bus	$275	$275
Total monthly technical fee	$20,350	$18,700

15 Demand generation scenarios: DRT

Table 3 displays the findings of the second simulation in this study, as described earlier, which concerns the simulation of the DRT service design. This study aims to build a DRT system to serve university students and staff within a 30-mile radius. Therefore, as shown earlier, a fully flexible DRT model parameters simulation was created. The scenario comprised students requesting a trip in any area without following a time-table, using virtual stops instead of "fixed" bus stops. For example, students might request trips from their homes to the university to arrive on time for their first class and return home after their last class. The DRT service is flexible and suited to each student's travel needs because they can request a trip anytime, giving them multiple options rather than a few selected ones.

The simulation results show that providing students with fully flexible DRT services leads to a trade-off between efficiency and user experience. This is because a comparison of the Fixed route system and the DRT system reveals that the DRT system can offer superior service to the students, using the same number of vehicles, and a smaller number of rides can be provided at a more flexible service. This conclusion supports Armellini and Bieker-Walz's [3] assessment that the DRT system has potential and good advantages due to reduced journey times from the city center to Vechelde, Denstorf, and Lamme. The results support Jain et al. (2014), who showed significant differences in trip time for each passenger when the bus (fixed) and taxi (flexible) journeys are compared, with taxi rides averaging 534 s and bus rides 1441 s, mainly because taxis use more direct routes. Moreover, The Schlüter et al. case study from 2021, which evaluated the DTR system in rural and urban areas, showed that automated DRT systems could lower transportation's economic and environmental costs in the interaction of rural and urban areas despite the limited travel time options.

Table 3 Iteration and optimization: simulation results

Particulars	Simulation features	Option 1	Option 2
Supply	Peak vehicle count	74	85
Demand	Total rides served per day	2300–2500	2500–2700
Quality of service	Average assignment time (requested pick-up vs. assigned pick-up in minutes)	11–13	10–12
Quality of service	Average trip duration (pick-up to the drop-off in minutes)	90–100	90–100
Quality of service	Average utilization rate (rides/vehicle-hour)	3–4	3–4
Efficiency	% Shared ride duration time (shared/total time with the minimum of 1 rider on board)	93–98%	93–98%

16 Conclusion and recommendation

This study was conducted to determine the feasibility of a DRT system et al.-Baha University so that students and staff can arrive on time for lectures and leave campus comfortably. Thus, they are not distracted by poor transportation services. In order to achieve the study's objectives, the following problems were posed:

- To provide further details concerning how data mining can improve the DRT systems' operational efficiency.
- To illustrate how data mining can be employed to create a DRT system that is both effective and efficient.

The above objectives were attained based on simulation, analysis, and interpretation of the collected data during the study's implementation. The first objective was reached throughout the research process but started with a thorough literature review. This began with an in-depth review of DRT Systems' historical development, which revealed that the system emerged in the early twentieth century and that most countries have adopted it in several variants. The literature review also focuses on the value of Demand-Responsive Transport Systems, highlighting their impact on urban sustainability and connectivity and the system's potential contribution to transportation in terms of effective and efficient mobility. Poor demand, poor dispatched routing and timetabling, and environmental factors, raise serious privacy concerns. The environmental, economic, social, effectiveness and efficiency, road safety, and travel behavior effects of DRT systems on local economies were also discussed. Intelligent technology and data mining in DRT systems were discussed. The literature review concluded with an assessment of DRT systems in KSA universities.

The second objective was achieved by defining three tasks to be completed successfully:

1. Determine the best morning and evening shift times based on the student's time-table to maximize convenience.
2. Optimize bus routes to use the same or fewer buses without making student rides too long.
3. Define the optimal bus capacity for a future DRT system to transport Al-Baha University students and staff to and from their homes.

The first task was completed by offering two bus arrival times (8.00 am and 10.00 am) before classes and two departure times (12.00 pm and 2.00 pm) after classes. These arrival and departure times maximize convenience for students so they are on time for classes and can quickly return home after class.

The second task was completed, resulting in an improved efficiency (lower average utilization rate) of 3–4 rides per hour for both DRT options and 7–8 rides per hour for both Fixed Route options. DRT is more efficient than Fixed Route. The DRT will have lower operating costs and a smaller environmental footprint than the Fixed Route System.

The third task was fulfilled, and the Financial Implications of the Fixed Routes scenarios show that fewer buses can be deployed to the Fixed routes than are currently operated. For example, option 1 saves $7412 per month using four fewer buses, while Option 2 saves $18,530 per month using ten fewer buses.

Deploying a fully flexible DRT system leads to a trade-off between efficiency and user experience, which means high-quality flexible service can be achieved with fewer buses. The University Transport Department can combine the Fixed Route Option and the DRT Option on its routes by operating Fixed Route systems in areas with high population density and high demand and DRT systems in areas with fewer people where it may be inefficient to operate Fixed Route systems.

This study investigates the feasibility of a DRT system at Al-Baha University and suggests using a Fixed Route system in high-demand areas and a DRT system in low-demand areas to help students arrive on time for class. Furthermore, based on this study's recommendation to operate a hybrid Fixed Rout and DRT system, an expanded viability analysis should be conducted to identify the following:

- The potential issues of such a hybrid system in a university setting in the KSA.
- The potential efficiency (cost savings) that such a system may produce on a broader scale deployment of the hybrid system between the university and off-campus student housing.

Acknowledgements This research is part of a project entitled "The Effect of Data Science on Urban Sustainability Through the Optimization of Demand-Responsive Transportation." This project was funded by the Deanship of Scientific Research, Al-Baha University, KSA (Grant No. 1442/035). The assistance of the deanship is gratefully acknowledged.

Author contributions The authors confirm contribution to the paper as follows: study conception and design: RNAA. Data collection: RNAA and AA. Analysis and interpretation of results: RNAA and AA. Draft manuscript preparation: AA. All authors reviewed the results and approved the final version of the manuscript.

Funding The Deanship of Scientific Research et al.-Baha University in Saudi Arabia funded this study (Grant no. 1442/035).

Availability of data and materials On a case-by-case basis, at the request of the corresponding authors, additional data may be made available to others.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that they are relevant to the content of this article.

References

1. Alonso-González MJ, Liu T, Cats O, Van Oort N, Hoogendoorn S (2018) The potential of demand-responsive transport as a complement to public transport: an assessment framework and an empirical evaluation. Transp Res Rec 2672(8):879–889
2. Ambrosino G, Nelson JD, Romanazzo M (2004) Demand responsive transport services: towards the flexible mobility agency. ENEA, Italian National Agency for New Technologies, Energy and the ….
3. Armellini MG, Bieker-Walz L (2020) Simulation of a Demand Responsive Transport feeder system: a case study of Brunswick. In: SUMO conference proceedings, vol 1
4. Arnold DV, Beyer H-G (2003) On the effects of outliers on evolutionary optimization. In: International conference on intelligent data engineering and automated learning, pp 151–160
5. Bektaş T, Elmastaş S (2007) Solving school bus routing problems through integer programming. J Oper Res Soc 58(12):1599–1604
6. Board HR (1973) Demand-responsive transportation systems. In: Proceedings of the demand-responsive transportation conference, vol 121
7. Brake JF, Nelson JD, Mulley C (2006) Good practice guide for demand responsive transport services using telematics. University of Newcastle upon Tyne, Newcastle upon Tyne
8. Carić T, Galić A, Fosin J, Gold H, Reinholz A (2008) A modelling and optimization framework for real-world vehicle routing problems. IntechOpen
9. Chamatkar AJ, Butey PK (2014) Importance of data mining with different types of data applications and challenging areas. J Eng Res Appl 4(5):38–41
10. Chein F, de Xavier Pinto CC (n.d.) Transport Infrastructure, urbanization and shipping costs: an analysis of the effects of an exogenous transport cost reduction on regional development
11. Chia D (2008) Policies and practices for effectively and efficiently meeting ADA paratransit demand (Issue 74). Transportation Research Board
12. Coetzee J, Zhuwaki N, Blagus D (2019) Demand-responsive transit design methods and applications for minibus Taxi hybrid models in South Africa
13. Cole LM (1968) Tomorrow's transportation: new systems for the urban future, vol 62. US Government Printing Office
14. Coutinho FM, van Oort N, Christoforou Z, Alonso-González MJ, Cats O, Hoogendoorn S (2020) Impacts of replacing a fixed public transport line by a demand responsive transport system: Case study of a rural area in Amsterdam. Res Transp Econ 83:100910
15. Davison L, Enoch M, Ryley T, Quddus M, Wang C (2014) A survey of demand responsive transport in Great Britain. Transp Policy 31:47–54
16. Fares N, Lyahyaoui A, Sedqui A (2019) Sustainable smart cities: optimization of demand responsive transport by using data science tools. In: International conference on advanced intelligent systems for sustainable development, pp 669–676
17. Fayyad U, Piatetsky-Shapiro G, Smyth P (1996) From data mining to knowledge discovery in databases. AI Mag 17(3):37
18. Friedman JH, Meulman JJ (2004) Clustering objects on subsets of attributes (with discussion). J R Stat Soc Ser B (Stat Methodol) 66(4):815–849
19. Fu L (1999) Improving paratransit scheduling by accounting for dynamic and stochastic variations in travel time. Transp Res Rec 1666(1):74–81
20. Haas A, Kriticos S, Lippolis N (2021) How transport reforms have impacted urban connectivity in Kenya. International Growth Center. https://www.theigc.org/blogs/how-transport-reforms-have-impacted-urbanconnectivity-kenya. Accessed 12 Mar 2022
21. HABITAT (2016) The New Urban Agenda. Habitat
22. HABITAT (2021) Mobility and transport. The UN-habitat. https://unhabitat.org/topic/mobility-and-transport. Accessed 12 Mar 2022
23. Haggiag A (2017) Expanded urbanization will change the face of transportation. https://www.masstransitmag.com/technology/article/12374977/expanded-urbanization-will-change-the-face-of-transportation. Accessed 12 Mar 2022
24. Han J, Pei J, Kamber M (2011) Data mining: concepts and techniques. Elsevier, Amsterdam
25. Imhof S, Frölicher J, von Arx W (2020) Shared autonomous vehicles in rural public transportation systems. Res Transp Econ 83:100925
26. Kasraian D, Maat K, van Wee B (2019) The impact of urban proximity, transport accessibility and policy on urban growth: a longitudinal analysis over five decades. Environ Plann B Urb Anal City Sci 46(6):1000–1017
27. Kavanaugh A, Fox EA, Sheetz S, Yang S, Li LT, Whalen T, Shoemaker D, Natsev P, Xie L (2011) Social media use by government: from the routine to the critical. In: Proceedings of the 12th annual international digital government research conference: digital government innovation in challenging times, pp 121–130
28. Lang N-S, Wachtmeister A (2020) This is how cities can overcome their growing transport pains. World Economic Forum. https://www.weforum.org/agenda/2020/11/this-is-how-to-overcome-cities-growing-mobility-pains/. Accessed 16 Mar 2022
29. Leduc G (2008) Road traffic data: Collection methods and applications. Working papers on energy, transport and climate change, vol 1, no 55, pp 1–55
30. Leiren MD, Skollerud K (2016) Demand responsive transport and citizen experiences: insights from rural Norway. In: Paratransit: shaping the flexible transport future. Emerald Group Publishing Limited
31. Liu T-Y, Su C-W, Qin M, Zhang X-Y (2020) What came first, transportation or urbanization? Singap Econ Rev 1–16
32. Ma Q, Liu W, Sun D, Lar S-U (2012) Traffic condition online estimation using multi-source data. J Comput Inf Syst 8(6):2627–2635
33. Mediterranean U for the (2021) Transport & urban development. union for the Mediterranian. https://ufmsecretariat.org/publication-speech/regional-integration-progress-report/. Accessed 8 Apr 2022
34. Meiryani AS (2019) Functions, processes, stages and application of data mining. Int J Sci Technol Res 8(7)

35. Mewari R, Singh A, Srivastava A (2015) Opinion mining techniques on social media data. Int J Comput Appl 118(6)
36. Munitz SM (1991) Demand responsive transport: an economic study of the jitney in the South-Western Cape, 1926–1990. University of Cape Town
37. Narayanaswami S (2016) Urban transportation: trends, challenges and opportunities
38. Navidi Z, Ronald N, Winter S (2018) Comparison between ad-hoc demand responsive and conventional transit: a simulation study. Public Transport 10(1):147–167
39. Nieuwenhuijsen MJ (2016) Urban and transport planning, environmental exposures and health-new concepts, methods and tools to improve health in cities. Environ Health 15(1):161–171
40. Padhy N, Mishra DP, Panigrahi R (2012) The survey of data mining applications and feature scope. ArXiv:1211.5723
41. Prud'homme J, Josselin D, Aryal J (2011) Quantitative analysis of pollutant emissions in the context of demand responsive transport. In: International conference on computational science and its applications, pp 439–453
42. Rahimi M, Amirgholy M, Gonzales EJ (2018) System modeling of demand responsive transportation services: evaluating cost efficiency of service and coordinated taxi usage. Transport Res Part E Logist Transport Rev 112:66–83
43. Rahman M, Watanobe Y (2019) An efficient approach for selecting initial centroid and outlier detection of data clustering. In: Advancing technology industrialization through intelligent software methodologies, tools and techniques. IOS Press, pp 616–628
44. Ronald N, Thompson R, Haasz J, Winter S (2013) Determining the viability of a demand-responsive transport system under varying demand scenarios. In: Proceedings of the sixth ACM SIGSPATIAL international workshop on computational transportation science, pp 7–12
45. Ryley TJ, Stanley PA, Enoch MP, Zanni AM, Quddus MA (2014) Investigating the contribution of demand responsive transport to a sustainable local public transport system. Res Transp Econ 48:364–372
46. Schlüter J et al (2020) A stochastic prediction of minibus taxi driver behaviour in South Africa. Humanit Soc Sci Commun 7(1):1–12
47. Sedighizadeh D, Mazaheripour H (2018) Optimization of multi objective vehicle routing problem using a new hybrid algorithm based on particle swarm optimization and artificial bee colony algorithm considering Precedence constraints. Alex Eng J 57(4):2225–2239
48. Statista (2021) Most popular social networks worldwide as of January 2021, ranked by number of active users. Statista Inc. https://www.Statista.Com/Statistics/272014/Global-Social-Networks-Rankedby-Number-of-Users/. Accessed 14 Apr 2022
49. Stieglitz S, Dang-Xuan L (2013) Social media and political communication: a social media analytics framework. Soc Netw Anal Min 3:1277–1291
50. Teller A, Veloso M (1995) Program evolution for data mining. Int J Expert Syst Res Appl 8(3):213–236
51. Tom VM, Mohan S (2003) Transit route network design using frequency coded genetic algorithm. J Transp Eng 129(2):186–195
52. Van Oort N, van der Bijl R, Verhoof F (2017) The wider benefits of high quality public transport for cities. In: European transport conference, pp 1–17
53. van Wee B, Kamargianni M, Shiftan Y (2018) Preparing for the new era of transport policies. Elsevier Science & Technology
54. Wang J, Yamamoto T, Liu K (2019) Role of customized bus services in the transportation system: insight from actual performance. J Adv Transport 2019
55. World C (2019) Urbanization and transportation. https://connectedworld.com/may-urbanization-and-transportation-archive/. Accessed 19 Mar 2022
56. Zaïane OR (1999) Principles of knowledge discovery in databases. Department of Computing Science, University of Alberta, 20
57. Zaki MJ, Wong L (2003) Data mining techniques. WSPC/lecture notes series
58. Zhao X, Mahendra A, Godfrey N, Dalkmann H, Rode P, Floater G (2020) Unlocking the power of urban transport systems for better growth and a better climate
59. Zheng Z (2018) Applications of data mining techniques in transportation safety study. North Dakota State University