Applied Decision Analytics Volume 2, Issue 1 (2026) 283-294

Applied Decision Analytics

Journal homepage: www.ada-journal.org
ISSN: 3104-2945

Business Analytics for Ride-Hailing Platforms: Demand Forecasting and Ride-Completion Modeling

Sara Debeljević[1,*], Filip Peovski[2]

[1] Faculty of Organizational Sciences, University of Belgrade, Serbia
[2] Department of Mathematics and Statistics, Faculty of Economics – Skopje, Ss. Cyril and Methodius University in Skopje, Republic of North Macedonia

ARTICLE INFO

Article history:
Received 26 December 2025
Received in revised form 26 January 2026
Accepted 1 March 2026
Available online 22 April 2026

Keywords:
Business Analytics, User Behavior, Ride-hailing platforms, Demand Forecasting, Service Quality

ABSTRACT

Digital ride-hailing platforms operate in dynamic environments characterized by fluctuating demand and variable user behavior. Business analytics plays a key role in supporting data-driven decision-making and operational planning in such systems. This study examines the application of descriptive and predictive analytics to analyze ride demand and ride completion behavior on a digital transportation platform. The empirical analysis is based on a dataset containing ride-level and user-related information. Descriptive analytics is used to preprocess the data, generate summary statistics, and identify temporal demand patterns. Predictive analytics is applied to forecast short-term ride demand using time-series methods and to model ride completion as a binary outcome using logistic regression. The results indicate that ride demand exhibits recurring temporal patterns suitable for short-horizon forecasting, with exponential smoothing achieving improved accuracy compared to a naïve approach. In addition, waiting time is identified as a key factor influencing ride completion probability. The findings demonstrate that business analytics can support proactive demand management, improved resource allocation, and enhanced understanding of user behavior in ride-hailing services.

1. Introduction

Digital ride-hailing platforms have become widely adopted and highly competitive, requiring continuous adaptation to rapidly changing demand conditions. External factors such as weather variability, seasonal patterns, and unexpected events contribute to frequent demand fluctuations, increasing operational uncertainty. In this context, service quality, fleet allocation, and user satisfaction depend directly on the platform's ability to balance supply and demand in real time [1-3].

To support these operational requirements, ride-hailing platforms increasingly rely on data-driven business analytics. Evidence-based decision-making enables firms to forecast demand, anticipate user behavior, reduce uncertainty, and optimize the allocation of operational resources.

* *Corresponding author.*
E-mail address: saradebeljevic@gmail.com

https://doi.org/10.66972/ada21202625

In environments characterized by continuous variability, analytical models provide a structured approach for designing pricing strategies, repositioning drivers, and maintaining a balance between profitability and user experience [4-7].

Recent research emphasizes the importance of applying both descriptive and predictive analytics in digital service platforms. Descriptive analytics enables the cleaning, structuring, and summarization of historical data, allowing organizations to identify temporal patterns and performance deviations [6, 8-9]. Predictive analytics extends this capability by estimating future outcomes, such as demand intensity and service completion probability, thereby supporting proactive operational planning rather than reactive decision-making.

Despite the growing adoption of analytics-driven solutions, several challenges remain. Inaccurate data preprocessing limited temporal coverage, and inappropriate model selection may lead to unreliable forecasts and misinformed managerial decisions. In ride-hailing systems, such limitations can result in increased waiting times, higher cancellation rates, and inefficient resource utilization, ultimately affecting both business performance and user satisfaction [10-12].

The present study addresses this gap by applying a combined descriptive and predictive analytics framework to evaluate business performance and user behavior in a digital ride-hailing platform. The empirical analysis is conducted on a dataset containing detailed information on Uber rides performed in 2024, including timestamps, location attributes, behavioral indicators, and ride outcomes. Descriptive analytics is used to identify temporal demand patterns and key performance indicators, while predictive analytics is applied to forecast short-term demand and model ride completion as a binary outcome [4, 7-8].

The objective of this research is to identify critical performance drivers, assess the effectiveness of forecasting techniques, and examine behavioral factors influencing ride completion. Specifically, the study seeks to answer the following research questions: (i) which temporal patterns characterize ride demand and support performance evaluation, (ii) which predictive model provides the most accurate demand forecasts, and (iii) how analytical results can be translated into actionable managerial recommendations. To address these questions, three hypotheses are formulated: H1, Uber ride data exhibit sufficient consistency to support analytical modeling; H2, exponential smoothing and naïve forecasting provide accurate short-term demand predictions [5, 11, 13-14]; and H3, logistic regression effectively predicts ride completion probability based on operational and behavioral drivers [15-17].

The remainder of the paper is organized as follows. Section 2 introduces the business analytics framework and reviews demand management mechanisms in ride-hailing platforms. Section 3 presents empirical results, including descriptive and predictive analytics. Section 4 concludes the paper with managerial implications and directions for future research.

2. Business Analytics Framework

Business analytics represents a systematic approach to transforming raw data into actionable insights that support evidence-based decision-making. It integrates statistical analysis, quantitative modeling, optimization techniques, and information technologies to improve organizational performance and managerial effectiveness. By enabling organizations to analyze historical data and anticipate future outcomes, business analytics reduces uncertainty and supports proactive strategic planning [4-7].

The concept of business analytics has evolved from traditional business intelligence, which primarily focused on reporting and descriptive summaries of past performance. Unlike business intelligence, business analytics emphasizes analytical modeling and optimization, enabling organizations to move beyond retrospective analysis toward prediction and decision support. As a result, analytics-driven organizations gain competitive advantage through faster response times, improved resource allocation, and more accurate performance evaluation [5, 18-19].

Business analytics is commonly classified into four complementary categories (descriptive, diagnostic, predictive, and prescriptive analytics.), as illustrated in Figure 1. Descriptive analytics focuses on summarizing historical data and answering the question of what has occurred. Diagnostic analytics extends this analysis by examining why certain outcomes occurred, often through correlation and variance analysis. Predictive analytics aims to estimate future outcomes using statistical models and machine learning techniques, while prescriptive analytics provides recommendations for optimal actions based on forecasting and simulation models.

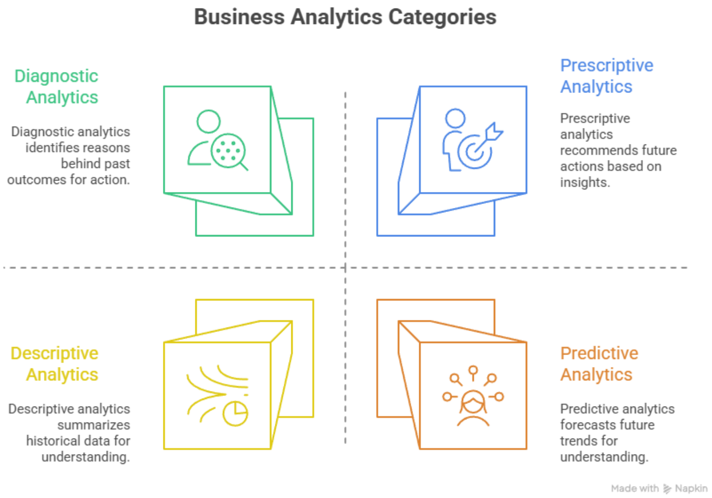

Fig. 1. Types of business analytics (adapted from [5])

In digital ride-hailing platforms, descriptive and predictive analytics play a particularly important role. Descriptive analytics enables platforms to identify demand patterns, peak periods, and performance deviations, while predictive analytics supports short-term demand forecasting and behavioral modeling. Together, these approaches form the analytical foundation for operational decision-making in environments characterized by high variability and real-time constraints.

2.1 Descriptive Analytics in Digital Ride-Hailing Platforms

Descriptive analytics constitutes the initial stage of analytical inquiry and provides a structured overview of historical data. Its primary purpose is to clean, organize, summarize, and visualize data in a form suitable for further analysis. Through descriptive analytics, organizations gain insight into past system behavior, identify regularities, and detect anomalies that may influence performance [9, 19].

In the context of ride-hailing platforms, descriptive analytics is used to examine ride volumes, temporal demand patterns, cancellation rates, and user behavior indicators. Typical analytical tasks include data preprocessing, aggregation across time intervals, and computation of summary statistics such as averages, dispersion measures, and frequency distributions. Visualization techniques further enhance interpretability by enabling rapid identification of trends, seasonal effects, and irregular demand fluctuations.

A well-executed descriptive analysis establishes the foundation for subsequent predictive modeling. By ensuring data consistency and revealing key performance indicators, descriptive analytics reduce the risk of biased or unreliable forecasts and supports meaningful interpretation of analytical results [4, 6].

2.2 Predictive Analytics and Demand Forecasting

Predictive analytics focuses on estimating future system behavior based on historical observations. In ride-hailing platforms, predictive models are primarily applied to forecast short-term demand and to estimate the probability of service completion. These forecasts are essential for driver allocation, capacity planning, and service quality management.

Time-series forecasting methods are commonly used to model demand dynamics, as ride requests exhibit temporal dependence, trend behavior, and recurring seasonal patterns. Various forecasting approaches may be applied, ranging from simple benchmark models to more advanced smoothing techniques. Because no single forecasting method performs optimally under all conditions, comparative evaluation of multiple models is a standard analytical practice [11, 13-14, 20].

In addition to demand forecasting, predictive analytics is applied to classification problems such as modeling ride completion outcomes. By estimating completion probabilities, platforms can quantify cancellation risk and identify operational conditions associated with service failure. These insights support proactive intervention strategies and improve overall system efficiency [15, 17].

2.3 Evaluation of Predictive Models

The effectiveness of predictive analytics depends on the accuracy and reliability of the employed models. Model evaluation is therefore a critical component of the analytical process. Forecasting accuracy is typically assessed by comparing predicted values with observed outcomes using standardized error measures [21].

Commonly used evaluation metrics include mean squared error (MSE), mean absolute percentage error (MAPE), and relative performance measures that compare advanced models with benchmark approaches. Lower error values indicate better predictive performance and improved responsiveness to temporal variation.

By systematically evaluating forecasting and classification models, organizations can select methods that best balance accuracy, interpretability, and operational feasibility. This ensures that analytical outputs can be effectively translated into managerial decisions and operational improvements [6, 8, 22].

3. Results

This section presents the empirical results obtained through the application of descriptive and predictive analytics to the Uber ride dataset for 2024. The results are organized into two main parts. First, descriptive analytics results are reported to examine data quality, demand patterns, service characteristics, and user behavior. Second, predictive analytics results are presented, focusing on demand forecasting performance and ride completion modeling.

Applied Decision Analytics
Volume 2, Issue 1 (2026) 283-294

3.1 Results of Descriptive Analytics

Descriptive analytics is applied to explore historical ride data, assess data quality, identify operational patterns, and evaluate service performance indicators. The analysis covers data completeness, distributional properties, temporal demand patterns, vehicle utilization, payment behavior, and cancellation dynamics.

3.1.1 The effect of data completeness and booking status logic

Before conducting further analysis, the internal consistency and completeness of the dataset were examined. The dataset contains 300 ride observations recorded during 2024, including both completed and non-completed bookings.

Null values were analyzed in relation to booking status. It was confirmed that null entries occur only in logically valid cases, such as rides with status No Driver Found or cancelled rides, where variables such as ride duration, distance, fare value, ratings, or payment method are not applicable. No invalid or inconsistent missing values were detected among completed rides. An example of logically valid null values associated with non-completed bookings is illustrated in Figure 2.

D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
Booking	Custom	Vehicle	Pickup I	Drop Lo	Avg VTA	Avg CTA	Cancell	Reason	Cancell	Driver C	Incomp	Incomp	Booking	Ride Dis	Driver R	Custom	Paymen
No Driver F	"CID19821	eBike	Palam Viha	Jhilmil	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID78736	Go Sedan	Noida Sect	Noida Sect	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID75681	Auto	Vidhan Sat	AIIMS	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID57175	Premier Se	Sadar Baza	Mehrauli	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID99658	Uber XL	Anand Viha	Dwarka Se	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID40517	Go Mini	Adarsh Na	Gwal Paha	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID16283	Uber XL	Madipur	DLF City C	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID83352	Bike	IFFCO Chc	Dilshad Ga	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID57367	Go Mini	Vishwavidy	MG Road	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID35158	eBike	Madipur	Saket	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID77921	Auto	Nawada	Jahangirpu	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID85816	Go Mini	Mansarova	Manesar	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID49987	Premier Se	Bahadurga	Uttam Nag	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID76943	Bike	Punjabi Ba	Ghitorni Vi	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID36567	Go Mini	Chanakyap	Munirka	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID32779	Auto	Satguru Ra	Sikanderpu	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID63075	Auto	GTB Nagar	Jasola	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID81928	Go Mini	Noida Film	Kirti Nagar	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID24611	eBike	Ashok Parl	Lok Kalyan	null	null	null	null	null	null	null	null	null	null	null	null	null
No Driver F	"CID14661	eBike	Seelampur	Udyog Bha	null	null	null	null	null	null	null	null	null	null	null	null	null

Fig. 2. An example of logically valid null values associated with non-completed bookings

These findings confirm that the dataset is internally consistent and suitable for subsequent descriptive and predictive analysis.

3.1.2 The effect of outliers on continuous ride metrics

Outlier detection was conducted for key continuous variables: waiting time for driver arrival (VTAT), ride duration (CTAT), booking value, and ride distance. Tukey's criterion with parameter $k = 1.5$ was applied. Table 1 reports quartiles and Tukey bound for the analyzed continuous variables.

Table 1
Quartiles and Tukey bounds for continuous variables

	VTAT	CTAT	Booking Value	Ride Distance
Q1	5.4	20.65	204	13.24
Q2	8.2	28.1	401	24.4
Q3	11.1	36.3	680.5	36.945
IQR	5.7	15.65	476.5	23.705
Q1-IQR*K	-3.15	-2.825	-510.75	-22.3175
Q3+IQR*K	19.65	59.775	1395.25	72.5025

Outliers were detected only for the Booking Value variable, indicating the presence of a small number of high-fare rides. Given that only five observations exceeded the Tukey bounds out of 300 total rides, and that these values may reflect legitimate premium or long-distance trips, all outliers were retained for further analysis.

Figure 3 and Figure 4 display box-and-whisker plots for VTAT and Booking Value.

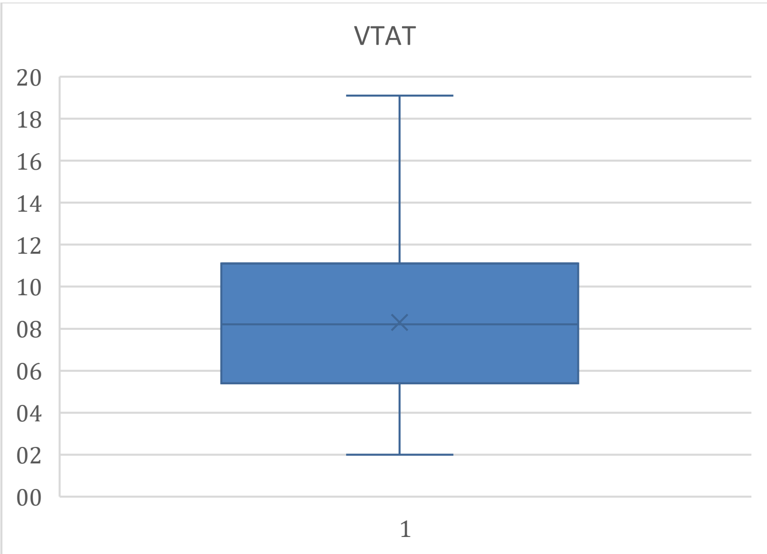

Fig. 3. Box-and-whisker plots for VTAT

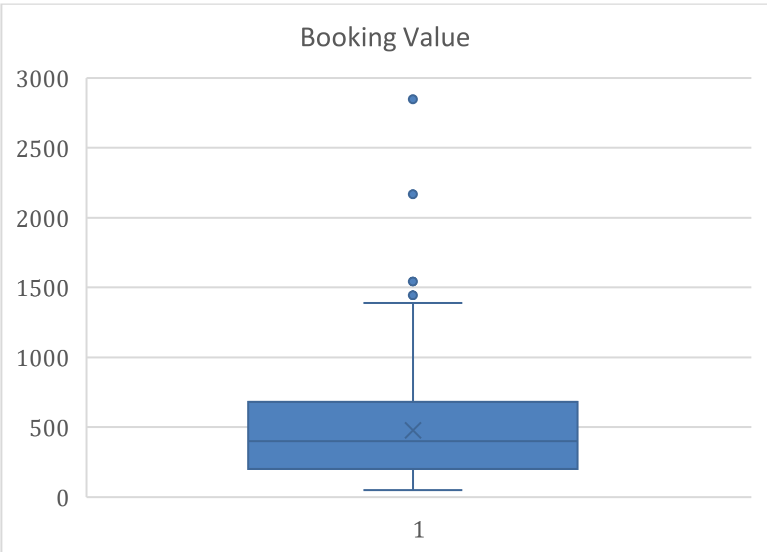

Fig. 4. Box-and-whisker plots for Booking Value

3.1.3 The effect of vehicle type on service utilization and waiting time

Vehicle type plays a central role in ride-hailing service configuration and operational planning. To evaluate vehicle utilization, the number of bookings by vehicle category was analyzed. Table 2 shows the distribution of bookings by vehicle type.

Table 2
Number of bookings by vehicle type

Row Labels	Count of Booking ID
Auto	78
Bike	38
eBike	26
Go Mini	59
Go Sedan	59
Premier Sedan	32
Uber XL	8
Grand Total	300

Table 2 shows the distribution of vehicle usage. The most frequently used vehicle categories are Auto, Go Mini, and Go Sedan, indicating strong demand for standard, low-cost ride options.

Average waiting time was then evaluated across vehicle types.

Table 3
Average driver arrival time (VTAT) by vehicle type

Row Labels	Average of VTAT
Auto	8.448298693
Bike	8.500042829
eBike	8.478422179
Go Mini	8.468101184
Go Sedan	8.401596442
Premier Sedan	8.438749926
Uber XL	8.575761974
Grand Total	8.456351971

Table 3 illustrates that average waiting time is relatively uniform across vehicle categories, with values close to 8.5 minutes. Slightly longer waiting times were observed for Uber XL, reflecting limited vehicle availability.

3.1.4 The effect of time on ride demand patterns

Monthly demand patterns were examined to identify temporal fluctuations in ride requests. Table 4 summarizes monthly ride demand during 2024.

Table 4
Monthly ride demand during 2024

Row Labels	Count of Booking ID	Row Labels	Count of Booking ID
Jan	13	Jul	30
Feb	25	Avg	25
Mar	26	Sep	29
Apr	24	Okt	17
Maj	32	Nov	18
Jun	37	Dec	24
Grand Total	300		

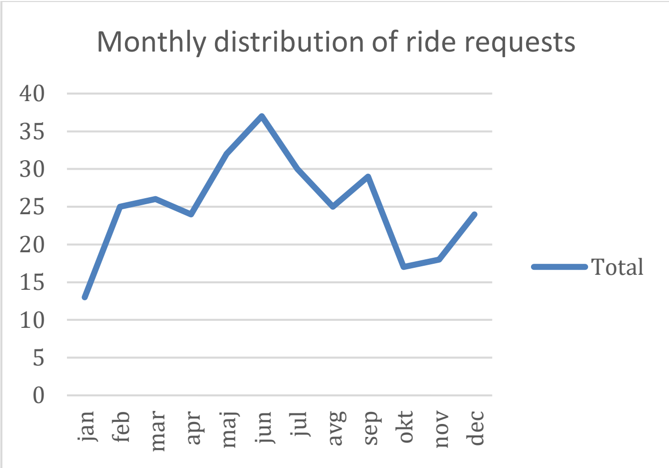

Fig. 5. Monthly distribution of ride requests

Figure 5 presents the monthly distribution of ride requests. Demand peaks are observed during late spring and summer months (May–July), while lower demand occurs during colder months, particularly January, October, and November.

These findings suggest the presence of seasonal effects driven by weather conditions, tourism, and holiday-related mobility patterns.

3.1.5 The effect of payment methods on transaction behavior

Payment preferences were analyzed for completed rides to evaluate digital adoption and transaction behavior. Table 5 presents the distribution of payment methods for completed rides.

Table 5
Distribution of payment methods

Row Labels	Count of Booking ID
Cash	47
Credit Card	12
Debit Card	16
Uber Wallet	29
UPI	107
Grand Total	211

Table 5 illustrates that UPI is the dominant payment method, followed by cash and Uber Wallet. Card-based payments remain comparatively limited, highlighting potential opportunities for partnerships and digital incentive programs.

3.1.6 The effect of booking status and cancellation reasons on service performance

Booking outcomes were examined to assess overall service reliability. Table 6 reports on the distribution of booking statuses.

Table 6
Distribution of booking statuses

Row Labels	Count of Booking ID
Cancelled by Customer	20
Cancelled by Driver	49
Completed	189
Incomplete	22
No Driver Found	20
Grand Total	300

Figure 6 shows that 63% of rides were successfully completed, while cancellations by drivers and customers account for a non-negligible share.

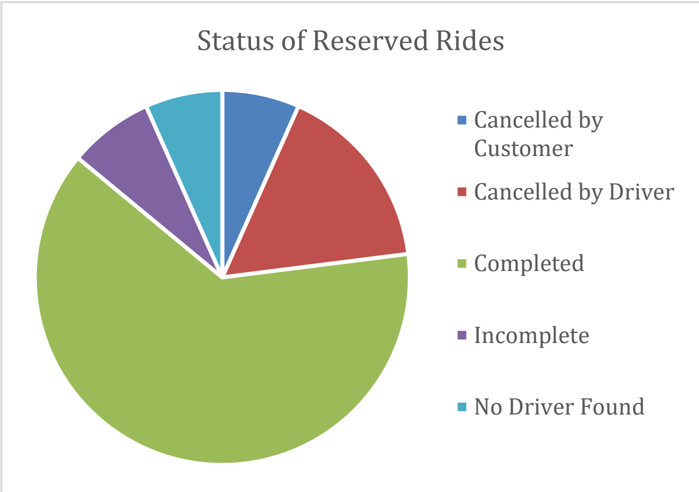

Fig. 6. Status of Reserved Rides

3.2 Results of Predictive Analytics

Predictive analytics was applied to forecast future ride demand and estimate the probability of successful ride completion.

3.2.1 The effect of forecasting method on demand prediction accuracy

Several forecasting models were applied to monthly ride demand data, including Forecast.ETS, naïve forecasting, average method, Holt's method, and ARIMA. Model performance was evaluated on training and test sets using RMSE, MAPE, and Theil's U statistics.

Based on the applied forecasting methods, ride demand for the first upcoming quarter of 2025 is expected to remain relatively stable. All approaches considered indicate the absence of pronounced short-term fluctuations, suggesting a steady demand level over the forecast horizon. In particular, the average forecasting method projects a constant demand of approximately 25 rides per month across the three-month period, providing a representative estimate of the expected demand level. The demand forecast obtained using the average method is illustrated in Figure 7.

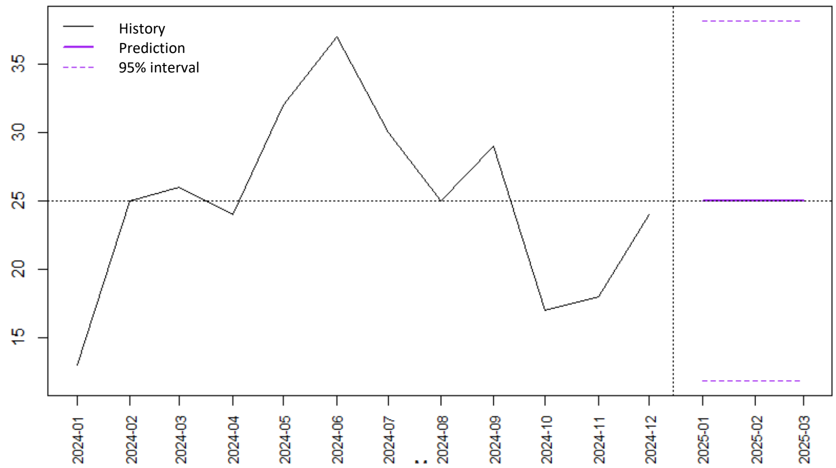

Fig. 7. Average method

Table 7 reports forecast error metrics and Theil's U statistics for the evaluated models. Results indicate that the average method achieved the lowest forecast error on the test set, outperforming more complex models under a short historical time horizon. ARIMA demonstrated strong in-sample performance but exhibited reduced generalization on test data, suggesting sensitivity to limited temporal coverage. Holt's method showed the weakest performance, likely due to the absence of sufficient seasonal information. Overall, the findings confirm that simpler benchmark models can provide more stable demand forecasts when data availability is limited.

Table 7
Forecasting error metrics and Theil's U statistics

Method	**Dataset**			**Theil's U**
		RMSE	**MAPE**	
Naïve	Training	6.40	19.48	-
	Test	9.83	50.84	1.00
Holt	Training	12.45	48.00	-
	Test	17.38	90.17	1.768
ARIMA	Training	2.97	7.39	-
	Test	10.90	54.01	1.108
Average	Training	6.25	21.9	-
	Test	7.75	39.29	0.789

3.2.2 The effect of operational and behavioral drivers on ride completion

Ride completion was modeled as a binary outcome using logistic regression. The dataset was split into training (80%) and test (20%) sets.

Model performance was evaluated using accuracy, precision, recall, and F1-score. Table 8 presents classification performance metrics for the logistic regression model.

Table 8
Classification performance metrics

Accuracy	**Precision**	**Recall**	**F1**
0.829	0.850	0.971	0.907

These results indicate strong classification performance, particularly in identifying completed rides. The model demonstrates high recall, suggesting effective detection of successful trips, while maintaining balanced precision.

4. Conclusions

This study applied a combined descriptive and predictive analytics framework to analyze business performance and user behavior on a digital ride-hailing platform using Uber ride data from 2024. The primary objective of the research was to identify demand patterns, evaluate service performance, assess forecasting accuracy, and examine factors influencing ride completion.

The descriptive analytics results revealed clear temporal demand patterns, with higher ride volumes during late spring and summer months and lower demand during colder periods. Vehicle utilization analysis showed a strong preference for standard and low-cost vehicle types, while waiting times remained relatively stable across categories. Cancellation analysis indicated that a substantial share of unsuccessful rides is driven by driver-side constraints and communication-related issues, highlighting areas for operational improvement. Payment method analysis confirmed a strong shift toward digital transactions, particularly UPI-based payments.

Predictive analytics results demonstrated that, under conditions of limited historical data, simpler forecasting approaches such as the average and naïve methods can outperform more complex models. While ARIMA and Holt's methods captured certain dynamics of the time series, their predictive accuracy was constrained by the short observation window. These findings emphasize the importance of aligning model complexity with data availability. Furthermore, logistic regression proved effective in predicting ride completion probability, achieving high recall and balanced overall performance, making it suitable for identifying successful rides and potential operational risks.

Overall, the findings confirm that business analytics can provide actionable insights for demand planning, service optimization, and evidence-based decision-making in ride-hailing platforms. However, the study is limited using a single-year dataset, which restricts the ability to fully capture long-term seasonal effects. Future research should incorporate multi-year data, additional contextual variables such as weather and real-time traffic conditions, and advanced machine learning models to further improve forecasting accuracy and operational recommendations.

Acknowledgement
No grant or financial support was provided for this research.

Conflicts of Interest
The authors declare no conflicts of interest.

References

[1] Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324
[2] Cachon, G. P., Daniels, K. M., & Lobel, R. (2017). The role of surge pricing on a service platform with self-scheduling capacity. *Manufacturing & Service Operations Management, 19*(3), 368–384. https://doi.org/10.1287/msom.2017.0623
[3] Castillo, J. C., Knoepfle, D., & Weyl, E. G. (2024). Matching and pricing in ride-hailing: Wild goose chases and how to solve them. *Management Science, 71*(5), 4377–4395. https://doi.org/10.1287/mnsc.2022.00096
[4] Albright, S. C., & Winston, W. L. (2017). *Business analytics: Data analysis and decision making* (6th ed.). Cengage Learning.
[5] Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning*. Harvard Business School Press.
[6] Liu, S. (2023). A review on business analytics: Definitions, techniques, applications and challenges. *Mathematics, 11*(4), 899. https://doi.org/10.3390/math11040899
[7] Sharda, R., Delen, D., & Turban, E. (2018). *Business intelligence, analytics, and data science: A managerial perspective* (4th ed.). Pearson.
[8] Delen, D., & Ram, S. (2018). Research challenges and opportunities in business analytics. *Journal of Business Analytics, 1*(1), 2–12. https://doi.org/10.1080/2573234X.2018.1507324
[9] Varsha, P. S., & Karan, A. (2023). Descriptive analytics and data visualization in e-commerce. In S. Akter & S. F. Wamba (Eds.), *Handbook of Big Data Research Methods* (pp. 86–104). Edward Elgar Publishing. https://doi.org/10.4337/9781800373713.00012
[10] Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007
[11] Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting, 20*(1), 5–10. https://doi.org/10.1016/j.ijforecast.2003.09.015
[12] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539
[13] Gardner, E. S. (2006). Exponential smoothing: The state of the art – Part II. *International Journal of Forecasting, 22*(4), 637–666. https://doi.org/10.1016/j.ijforecast.2006.03.005
[14] Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and applications* (3rd ed.). John Wiley & Sons.

[15] Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). John Wiley & Sons. https://doi.org/10.1002/9781118548387

[16] Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (5th ed.). John Wiley & Sons. https://doi.org/10.1002/9780470542811

[17] Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. *MIS Quarterly, 35*(3), 553–572. https://doi.org/10.2307/23042796

[18] Delen, D., & Demirkan, H. (2013). Data, information and analytics as services. *Decision Support Systems, 55*(1), 359–363. https://doi.org/10.1016/j.dss.2012.05.044

[19] Waller, M. A., & Fawcett, S. E. (2013). Click here for a data scientist. *Journal of Business Logistics, 34*(2), 249–252. https://doi.org/10.1111/jbl.12024

[20] Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). John Wiley & Sons. https://doi.org/10.1002/9781118619193

[21] Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting, 22*(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001

[22] Wolniak, R. (2023). The applications of usage of business analytics in Industry 4.0. *Scientific Papers of Silesian University of Technology, Organization and Management Series, 179*, 730–741. https://doi.org/10.29119/1641-3466.2023.179.38