WAVETEC

2023

Data Science Division

Strategic Banking

Insights on Customer Experience

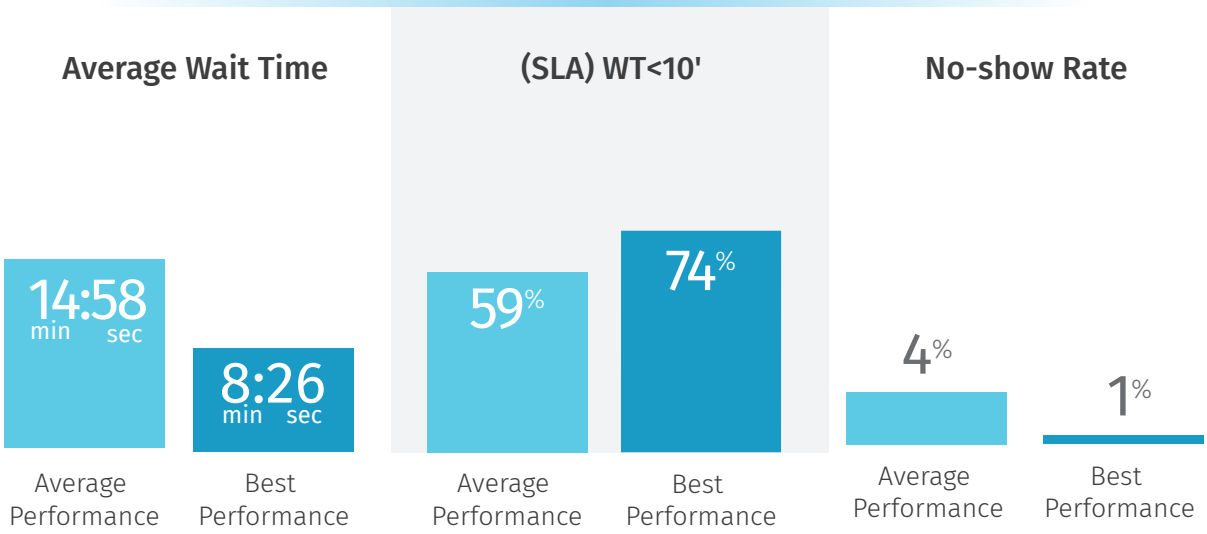
Benchmarking Analysis

wavetec.com | us@wavetec.com

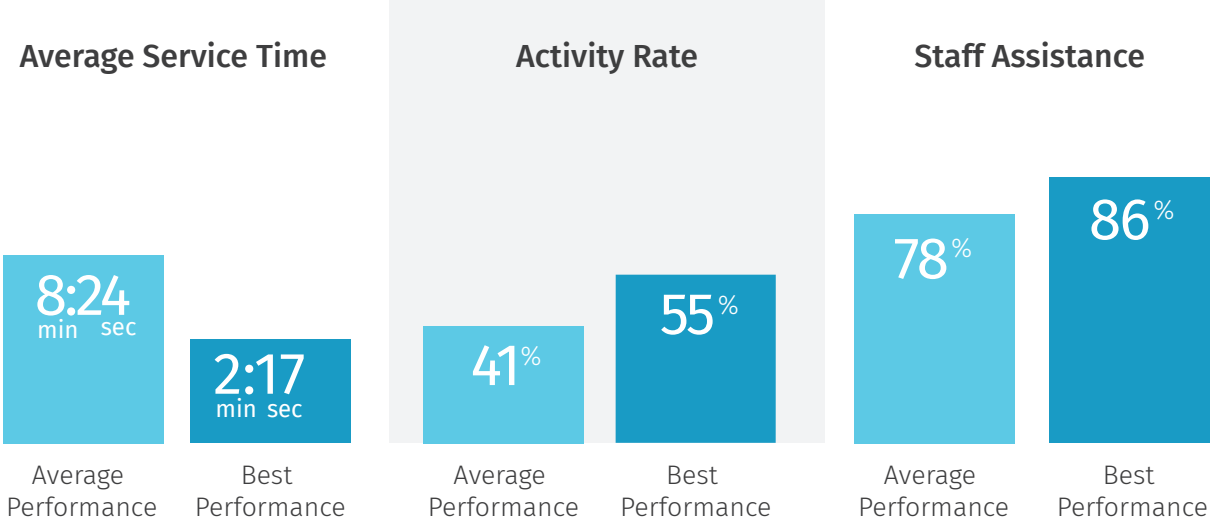
EXECUTIVE SUMMARY

This benchmarking report includes benchmarking data for bank branch performance metrics, or KPIs, surrounding the queuing and appointment experience, such as Waiting Time, No-Show Rates, Teller Efficiency, and more. The report is the result of over 20 years working with banks across the world. Moreover, the report incorporates an introduction to simulations and how data science can be used to make decisions to improve customer experience and tellers' efficiency. Refer to the following pages to see a breakup of the below summarized insights (e.g. by KPI by type of service). For a regional approach (e.g. North America, South America, Asia, Africa, etc.), please email us at **us@wavetec.com**

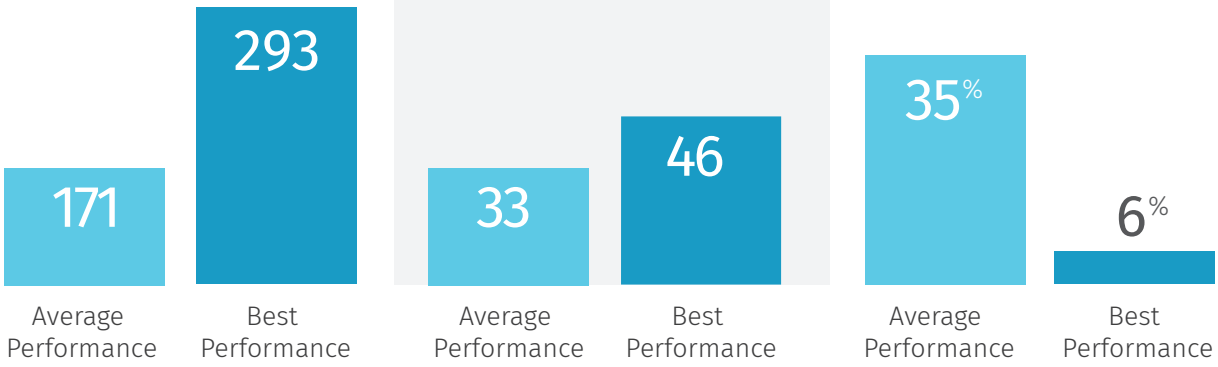
SERVICE QUALITY



STAFF EFFICIENCY



BRANCH TRAFFIC



FROM INSIGHTS TO ACTIONS. By running a simulation of over 6,000 scenarios, a bank branch could reduce waiting time by 30% without increasing the number of staff.

	Staff Availability	Tellers	Wait Time
Scenario A	45%	4	03:23 min
Scenario B	60%	4	02:02 min

For a personalized region specific report, email **us@wavetec.com**

INTRODUCTION

How is our branch experience compared to other banks? How can we do better?

Commercial banks have faced great adversities in the last couple of years: a global health crisis that led to lock-downs, web channel-forced digital transformation, post-pandemic effects, inflation, and the advent of fintechs. All these changes have generated re-adaptations to the way banks are doing business. There are several ways to cope, but one way to reroute may be to examine how others are doing. Thus, having a benchmark on the branch network performance of other banks becomes fundamental to understanding where we stand. After understanding where we stand, a question arises; can we do better? What should we do?

In an era of ample data and information, data science has become a fundamental ally in determining a plan of action.

The objective of this report is to shed some light on these two topics: **determining a benchmark** for commercial banks, and demonstrating some use cases on how data science can help **improve branch performance**. In the following section, we introduce a number of metrics regarding different areas of the banking industry. These are to be used as a reference for banks to understand whether they are on the correct path or if they need to improve some aspect of their operations. In the last section, we show how banks can use data science to improve their branch performance.

Benchmarking Banking Performance Worldwide:

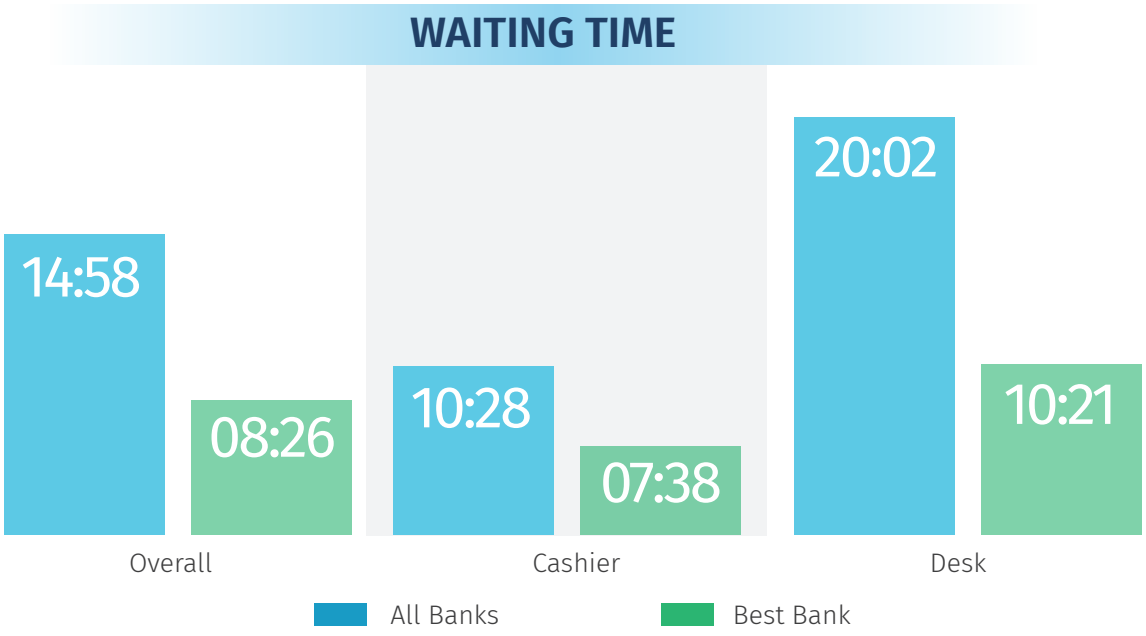
An Analytical Overview

SERVICE QUALITY

How can **commercial banks know** if they are providing **adequate service** to their customers without asking them?

Service quality can be measured in different ways. One of them is the **average waiting time**. As waiting time is "dead" time for customers, all banks try to minimize it. One problem with this metric is that it can be affected by extreme values. For example, if we have 10 clients waiting 5 minutes and 1 client waiting 60 minutes, the average will be 10 minutes which is not representative of the waiting time for most individuals. For that reason, banks complement this metric with the **Service Level Agreement (SLA is the percentage of customers that wait less than a target wait time).** In the previous example, if our target is 7 minutes, then the SLA will be 91%; this is more representative than the waiting time metric. Both metrics have their pros and cons:

Metric	Waiting Time	Service Level Agreement (SLA)
Pros	Average Customer	Fails with extreme values
Cons	Extreme values do not affect it	No information about customers that do not achieve the target



Reduce customer wait time without extra cost

Another metric that gives an idea of the service quality is the **no-show rate (NSR)** which measures the percentage of ticket no-shows over the issued tickets. The importance of this metric lies in giving us a client perspective of the service quality (implies action from their side) and an idea of the willingness to wait. For example, let's assume that banks A and B have the same waiting time and SLA but differ in the churn rate. Evaluating the service, we could say that they are similar, but one of them is less adapted to clients' needs. In one case, clients are more impatient (they value their time more), therefore, the bank could be affecting its revenues (non-per ceived incomes because of higher abandonment).

In the following chart, we include the benchmarks of these 3 metrics, split by cashier and desk:

	Waiting TIme (Best Waiting Time)	SLA* (Best SLA)	No-Show Rate** (Best No-Show rate)
Overall	14:58 (08:26)	59% (74%)	4% (1%)
Cashier	10:28 (07:38)	67% (75%)	3% (1%)
Desk	20:02 (06:37)	50% (80%)	4% (1%)

*Tickets with WT<10' - (SLA= Service level agreement)
**No-show tickets / issued tickets

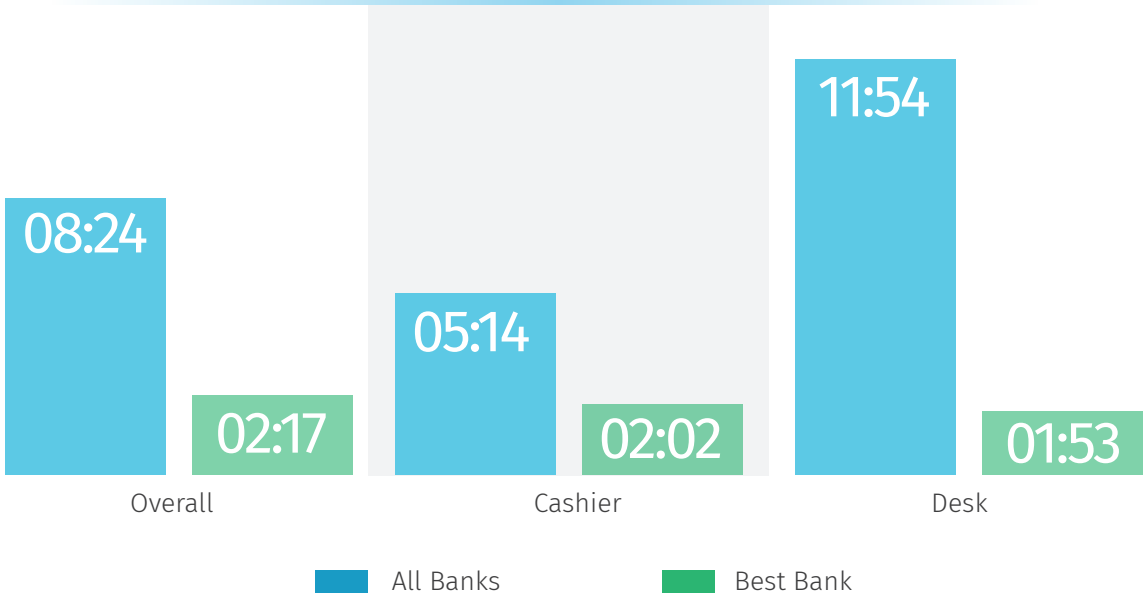
As it can be observed, there is **a generalized area of opportunity to improve the service at branches, especially at the desk counters where the SLA stands at 50 % and the No-Show Rate at 4%.**

STAFF EFFICIENCY

After analyzing the service quality and client satisfaction, one must consider bank staff. How can commercial banks know if their staff is efficient? The main metrics used for this purpose are the average service time and activity rate. In the first one, we calculate the average time an operator (teller/cashier/banker) takes to deliver the service while in the second one, we measure the percentage of time an operator is servicing a customer over the total working hours. A 100% activity rate is not ideal either because it will translate into a long queue and excess waiting time.

The activity rate is 41%, which means that for every hour the branch opens, a worker provides services for 25 minutes.

SERVICE TIME



Another metric that may be considered with a Queue Management System (QMS) is the average delay in calling the next ticket. This metric measures the average time between the end of the last ticket and the calling of the next one. Surprisingly, this delay can be higher than a minute in certain banks. Lastly, a metric that complements preceding ones is the assistance of each operator measured by the number of days they register in the QMS per month over business days. When service quality is not satisfactory, all these staff efficiency metrics shed light on potential causes of poor performance. Below are the benchmarks of these metrics in the banking industry:

	Average Service Time (Best Service Time)	Activity Rate (Best Activity Rate)	Delay in Calling the Next Ticket** (Lowest Delay)	Staff Assistance (Best Assistance)
Overall	8:24 (02:17)	41% (55%)	00:37 (00:05)	78% (86%)
Cashier	5:14 (02:02)	48% (69%)	00:18 (00:04)	–
Desk	11:54 (01:53)	34% (51%)	00:58 (00:15)	–

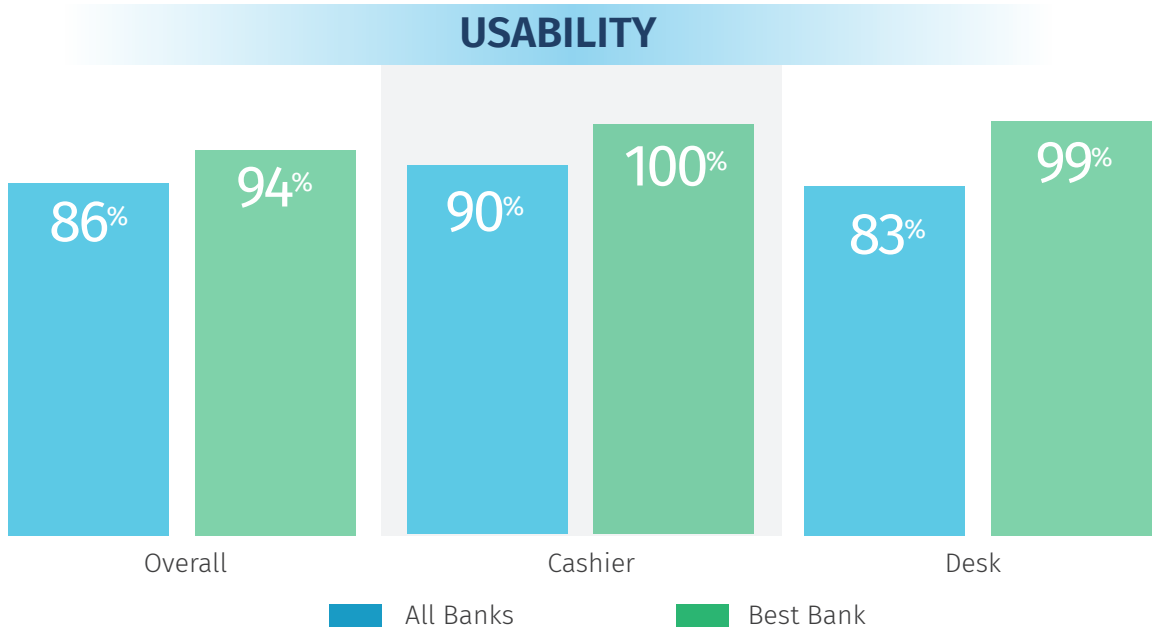
*Total service time / total working time
**The average time between tickets
***Days registered in QMS / business days

SYSTEM USE

Having a Queue Management System is the first step in measuring branch performance, and the KPIs discussed in the previous sections. However, it is important to ensure that the QMS deployed is being used correctly. Otherwise the insights might not be representative and may lead to incorrect actions. This is where two important questions must be asked:

Is the Queue Management System (QMS) being used in the branches?
Is it being used the right way?

To answer the first question, **we need to measure the usability level of the QMS per branch.** One way would be to measure the number of days that at least one ticket is issued from the QMS over the business days of the month. We can also do it hourly or combine different metrics to construct one KPI.



In the second question, **inadequate use of the system could be because of manipulation and/or lack of training (or tellers' resistance to change).** In the first case, operators could try to manipulate the system as they know they would be evaluated by their performance. The second case usually happens due to ignorance and lack of interest.

One way to measure the level of manipulation is by the Fast Ticket Rate which is the percentage of tickets that are served in less than 5 seconds. For example, an operator can start the service and immediately stop the service with the objective of reducing the average service time. In case of misinformation, branches can use the QMS for issuing and printing tickets but not for calling them. In this case, most tickets will not have a standard status such as "served", "transferred" or "no-show." Most of them will be under N/A status.

Below are the benchmarks of these metrics in the banking industry:

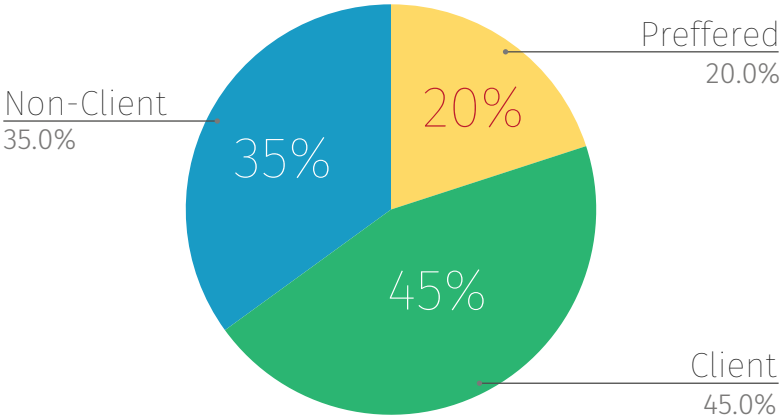
	Usability (Best usability)	Fast Tickets ** (Lowest fast tickets)	% NA Tickets *** (Lowest NA tickets)
Overall	86% (94%)	6% (1%)	5% (1%)
Cashier	90% (95%)	5% (2%)	3% (1%)
Desk	83% (90%)	7% (1%)	9% (1%)

*Days issuing tickets/business days
**Percentage of tickets served under 5 seconds
***N/A tickets / issued tickets

TRAFFIC

Regarding the number and type of customers that arrive at commercial bank branches, we can assume some standard metrics; tickets issued per day per branch, tickets served per day per operator, and distribution of segments between preferred, client, and non-client tickets.

Ticket Segments



Below are the benchmarks of these metrics in the banking industry:

	Tickets per Branch* (Highest tickets per branch)	Tickets per Operator** (Highest tickets per operator)	% Preferred Tickets Preferred tickets/ issued tickets	% Client Tickets Client tickets/ issued tickets	% Client Tickets Client tickets/ issued tickets
Overall	171 (293)	33 (46)	20%	45%	35%
Cashier	70%	44 (64)	19%	40%	41%
Desk	30%	19 (28)	22%	51%	27%

BUSINESS OPPORTUNITIES

Not all transactions are equally profitable for commercial banks. Usually, **desk services yield higher profits than counter ones.** Identifying these high-profit transactions is of great importance to banks.

Then, we compare the wait time, SLA, and no-show rate of these tickets with to those of the less profitable tickets. This would give one an idea of whether the bank is prioritizing these high-profit tickets.

High-profit transactions such as account openings, mortgages, and loans abandon the queue 14% more than other tickets.

Below are the benchmarks of these metrics in the banking industry:

	High-Profit Tickets* (Best HPT)	Waiting Time Others vs. HPT** (Highest difference)	No-Show Rate HPT vs. Others*** (Lowest difference)
Scenario A	13%	12%	14%
	(46%)	(31%)	(29%)

*Percentage of desk tickets that belong to the following categories: open account, mortgage, loan, investment, and product advisory.
**Waiting time others / Waiting Time High-profit tickets - 1
***No-Show rate High-Profit tickets / No-Show rate others - 1

For a personalized region specific report, email **us@wavetec.com**

WAVETEC®

A Data Science Approach to Attaining Branch Performance Objectives

OBJECTIVES

How can **data science improve branch performance** and help banks stand out from competitors?

Data science can help improve branch performance in many ways. Banks must identify use cases, form hypotheses, and test them with relevant data science tools. Data science can be used to gain knowledge about behaviors and processes, write algorithms that process large amounts of information quickly and efficiently, and guide data-driven decision-making. The following use cases demonstrate how banks can introduce data science to reduce wait times, achieve SLA targets, and segment customers

Reduce Customer Wait Time Without Extra Cost

Staff Reallocation

Using discrete-event simulation(DES), **a bank can match branch network staff needs with the bank's waiting time targets.** This technique involves duplicating a business day in a branch with their regular metrics (service time, customer arrivals, segment distribution, activity rate, etc.) and calculating the average wait time according to the number of tellers. Then, this imitation is replicated randomly thousands of times to obtain a robust outcome.

Multiple scenarios with varying numbers of tellers can be run and compared. The more tellers are added, the less the waiting time will be, ceteris paribus, and the simulation will indicate the impact of additional staff on the waiting time, so that a cost/benefit-based decision can be made. This set of scenarios allows banks to choose the scenario

that reaches their target waiting time and identifies the staff needed. Finally, **banks can reallocate workers between branches according to shortages and surpluses (and locations), to reduce the overall wait time.**

Improve Staff Productivity

Staff availability is the percentage of time a worker is available to provide services over working hours. Usually, this metric is under 100% because of breaks, trainings, meetings, and other back-office activities. In this case, we can use Simulation Analysis to detect the improvement in wait times by incrementing the staff availability rate (and reducing other activities).

	Staff Availability	Tellers	Wait TIme
Scenario A	45%	4	03:23 min sec
Scenario B	60%	4	02:02 min sec
Scenario C	60%	3	03:36 min sec

For a personalized region specific report, email **us@wavetec.com**

For example, a bank could reduce the wait time by 30% incrementing the staff availability from 45% to 60% (bankers' average availability is 50%).

Instead of reducing the wait time, another potential outcome of incrementing staff availability is to reduce the staff needed and maintain a similar wait time (cost reduction action).

Auto-Call Implementation

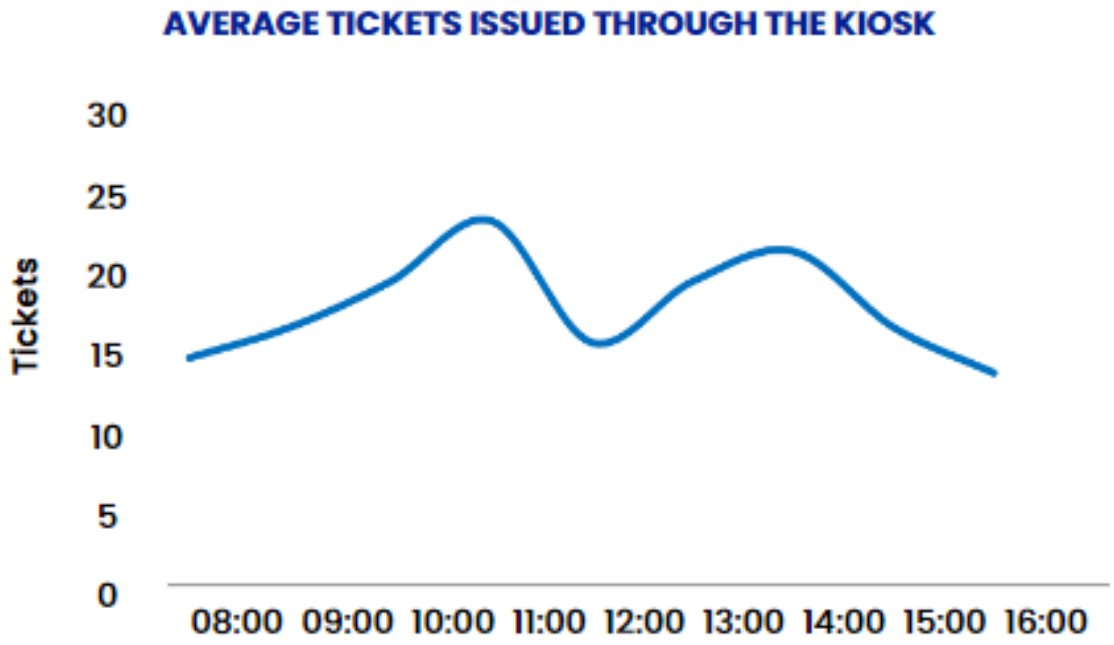
Some queue management system-feature called auto-call. Its function is to **automatically call the next ticket at the end of the preceding one**. This would eradicate potential teller delay between tickets. This delay impacts not only the next customer's wait time, but also that of others in the queue. Using Simulation Analysis, a bank can measure the impact of implementing a feature like auto-call. This simulation **is important so banks can understand whether the action will deliver satisfactory results. Implementation is challenging** because employees are not used to working in this way. By implementing such auto-call functionality, we have observed significant improvements in wait times, i.e., up to 10% reduction.

Achieve Targeted SLA

Flatten the Curve with Appointments

A well-known fact is that branch traffic moves from peaks to valleys and differs based on the day of the week. During the pandemic, many banks started implementing an appointment system, and nowadays, most banks utilize dual ticketing.

These appointments can be arranged to flatten the traffic curve complementing footfall valleys with extra slots of appointments (and zero or a few appointments for peak footfalls hours).

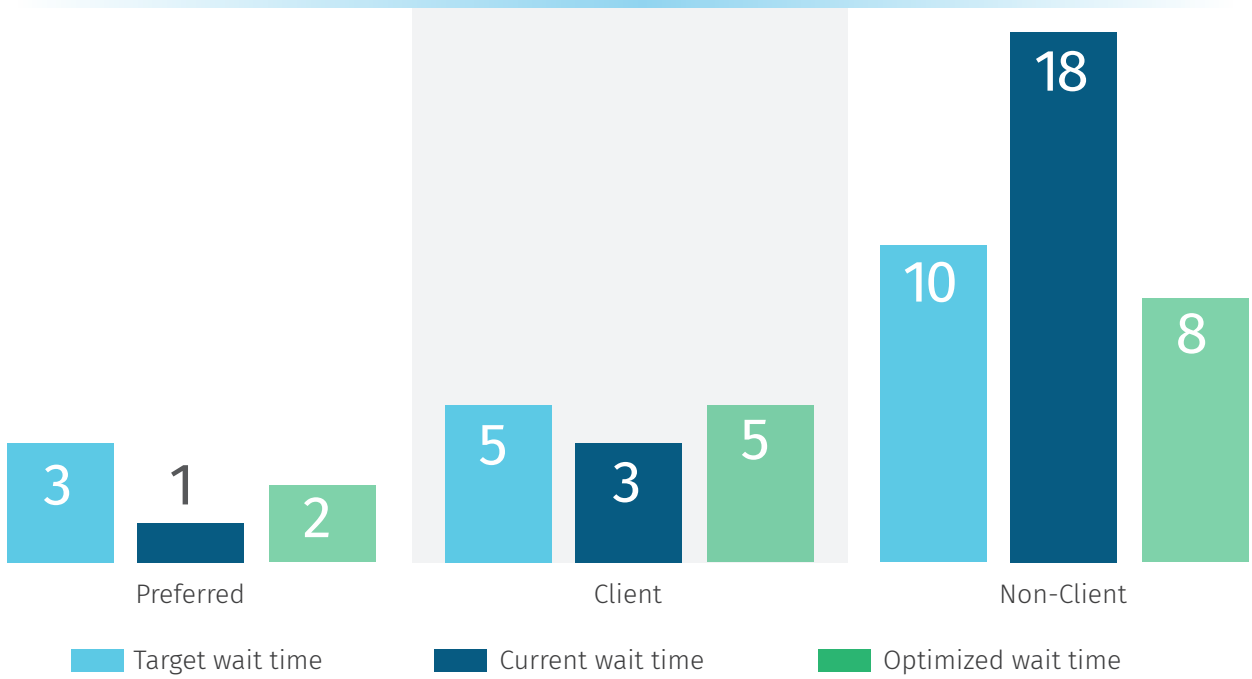


By using machine learning analysis, banks can predict the traffic of their branch network. Different models such as ridge regression, random forest, or neural network, can be used to forecast the expected number of customers that will visit each branch per hour. Machine learning predicts this by producing output of an algorithm after it has been trained on a historical dataset. Having an accurate model will help us detect the peak hours and valleys of traffic footfalls. Thus, the bank will be able to **fine-tune the slots for appointments to flatten the curve** and improve the SLA by not congesting the system.

Optimize Calling Profile According to Customer Segments Priorities

Most banks have different customer segments (preferred, client, and non-client) which they use to set different attention priorities. For that, banks configure tellers' calling logic to follow a particular order to give higher priority to preferred customers (there are some cases where banks do not assign a calling logic according to their targets). Yet, **defining the calling logic for all tellers in the branch to reach targeted waiting time per segment is not a straightforward task.**

WAITING TIME OF CUSTOMER SEGMENTS UNDER 2 SCENARIOS (MIN.)



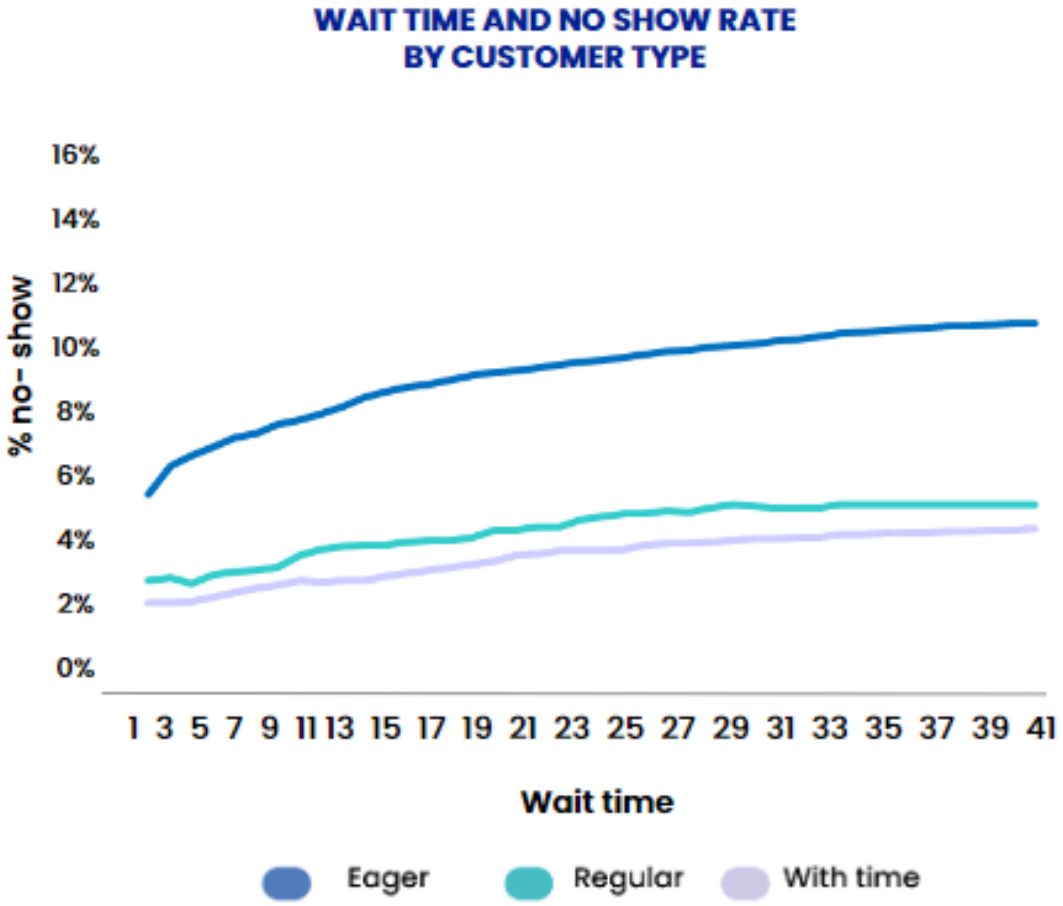
For example, bank wait time targets were 3, 5, and 10 minutes for preferred, clients and non-clients respectively. A branch with 4 tellers assigned the same calling profile for all tellers (priority: first serve the higher segment priority) but the results 1, 3, and 18 minutes did not fit the target. Simulations were run to identify the best combination of profiles to better achieve the target. In this case, the results showed that 2 tellers should be configured as FIFO (first-in, first-out) and 2 as priority profiles to obtain waiting times more in line with the targets 3, 5, and 8 minutes.

Customer Segmentation

Eager customers vs Non-Eager Customers

When a customer goes to the bank with the intention of purchasing a new product (investment, loan, new account) but leaves the queue before being served due to the long waiting time, it translates into forgone income. To avoid these situations, the bank must identify and segment customers by their level of eagerness/distress, since each

one's willingness to wait is different. The following chart shows 3 customer segments of a bank according to the willingness to wait



Understanding the relationship between wait times and no-show per segment helps the bank implement actions regarding prioritizing the different customer segments to diminish impact on revenue. Another potential comparison is the extra cost of hiring more workers (to reduce the wait time and abandons) against the forgone revenue because of abandoned tickets.

WAVETEC®

Wavetec is a market leader in customer experience solutions globally with extended experience in a number of markets and industries. We provide innovative answers to complex questions by leveraging technology and intelligent practices.

With over 200 employees spanned across 9 global offices, Wavetec specializes in queue management systems, financial market displays, self-service kiosks, and WhatsApp solutions, with an extensive clientele, including, but not limited to, Nike, HSBC, Emirates, Ikea, and Unilever.

If you would like to discuss this further, please email us at **us@wavetec.com** so that a meeting with our data science team can be scheduled. We are happy to help, regardless of whether you use Wavetec queue & appointment management systems.

Global Offices:

Barcelona | Dubai | Karachi | Lima | London | México | Nairobi | New York | Riyadh | Santiago