

# MODELO PARA LA MEJORA DEL SERVICIO DE ATENCIÓN AL CLIENTE MEDIANTE LA TEORÍA DE COLAS: CASO DE ESTUDIO DE UNA AGENCIA BANCARIA

LETICIA SIU LENG LEÓN LAZO  
20172281@aloe.ulima.edu.pe  
<https://orcid.org/0000-0002-4766-2728>  
Universidad de Lima, Perú

LORENA PAOLA VIVANCO VIVANCO  
20171715@aloe.ulima.edu.pe  
<https://orcid.org/0000-0002-2762-2980>  
Universidad de Lima, Perú

## RESUMEN

El análisis de las líneas de espera en las diferentes agencias bancarias nace por el descontento de la población por las largas colas y un tiempo promedio de más de 8,539 minutos de espera para recibir el servicio respectivo, tomando en cuenta el día y hora con mayor demanda. Por eso es necesario un equilibrio entre la capacidad y la oferta. El objetivo de esta investigación es aplicar la teoría de colas como base fundamental para modelar el sistema con la ayuda del *software* Arena (16.1 Student Version), y analizar su comportamiento para identificar las posibles mejoras y lograr una mayor calidad de servicio al cliente. Así se simuló el comportamiento de las líneas de espera, definiendo los recursos, los tiempos de espera, el tiempo entre llegadas, las entidades involucradas, entre otros. Además, se utilizó un enfoque cuantitativo de alcance explicativo y descriptivo para la obtención de datos. De esta manera, al realizar las réplicas en el *software* con el modelo propuesto, se consiguieron resultados positivos, tales como la reducción de tiempo en cola en un 52,61 %; pero también se identificaron factores que incrementan ese tiempo, como el hecho de que el cliente no finaliza el proceso y se retira porque desconoce ciertos requisitos para la operación que desea realizar. Por ello, la mejora planteada también ayuda a disminuir el número de personas que se retiran del proceso, cuyo porcentaje pasó de 14,52 % a 4,032 %. Finalmente, la existencia de diferentes variables en los clientes que no califican un buen servicio de atención con un menor tiempo de espera limita los resultados obtenidos.

PALABRAS CLAVE: teoría de colas, servicio de atención, simulación, agencia bancaria

## MODEL FOR IMPROVING CUSTOMER SERVICE THROUGH QUEUING THEORY: CASE STUDY OF A BANKING AGENCY

### **ABSTRACT**

**ABSTRACT.** This article analyses waiting lines at a bank branch where customers described long queues and waiting time (8,539 minutes on average) as the main problem with the bank's service. For research purposes, we modeled a system, based on queuing theory, in Arena software (16.1 Student Version) to identify possible improvements to customer service. The waiting line simulation considered resources, waiting times, time between arrivals, and entities involved. We gathered data using an explanatory and descriptive quantitative approach. The proposed model offered positive results, such as a 52,61 % reduction in queue time. The model also allowed us to identify different factors affecting queue time, such as customers who do not finish the process and leave because they did not comply with all the requirements for a given service. The proposed improvement also helps to reduce the number of people who withdraw from the process from 14,52 % to 4,032 %. Different variables that determine poor customer service reviews despite shorter waiting time limits the results obtained.

**KEYWORDS:** queuing theory, customer service, simulation, bank agency

## 1. INTRODUCCIÓN

Actualmente, la experiencia bancaria es definida por muchas personas como tediosa y se ha vuelto un símbolo de insatisfacción por las extensas líneas de espera, lo cual tiene como consecuencia que la mayoría de los clientes prefieran evitarlas de todas las maneras posibles. Por otra parte, los productos y servicios que ofrecen las diferentes agencias bancarias guardan una gran similitud entre ellos. Por ello, la decisión de cuál de ellas elegir depende muchas veces de la calidad del servicio al cliente (Entriger, 2020).

La teoría de colas es una herramienta que se caracteriza por su enfoque sistemático y por su aplicación en modelos que analizan el comportamiento de las líneas de espera (Peter & Sivasamy, 2019). Por lo general, la mala gestión de dichas líneas se debe a la disparidad entre la capacidad de la empresa al brindar el servicio y la demanda. Por eso, el objetivo principal al aplicar la teoría de colas es determinar la capacidad apropiada y la estabilidad del sistema para poder garantizar el equilibrio y una buena calidad del servicio al cliente (Kittsteiner & Moldovanu, 2005).

Los modelos matemáticos basados en la teoría de colas presentan una alta variedad, que difiere en sus supuestos: tiempo de llegada del cliente, capacidad de la cola, tiempo de espera, número de servidores, entre otros (Dorda et al., 2019). A partir de la información obtenida, la simulación puede ser utilizada como una técnica para ejecutar estos modelos, con resultados y predicciones relativamente rápidas, representando así la situación con las modificaciones propuestas de manera concisa (Portilla et al., 2010).

Existen muchos escenarios en los que, para recibir un servicio, el usuario debe esperar para ser atendido. Las agencias bancarias no son ajenas a esta realidad, pues constituyen uno de los servicios más utilizados y en los cuales se generan las mayores molestias por sus largas colas y la demora en sus tiempos de espera. Esto depende de distintos factores: el día, el horario, la disponibilidad de los recursos, etcétera. Por ejemplo, en los días de fin de mes, la afluencia de usuarios suele ser mayor; o en las horas de cambio de turno del personal de caja, se reduce la capacidad de atención y aumenta el tiempo de espera del servicio. En la mayoría de los casos, esto se origina porque la capacidad del servicio es menor que la demanda (López Hung & Joa Triay, 2018). La teoría de colas busca comprender y determinar el funcionamiento de las colas, examinando los diferentes escenarios para diseñar estrategias que permitan alcanzar la capacidad idónea para brindar un servicio de calidad (Cárdenas Estrada et al., 2019). Con la aplicación de un *software* es factible simular y representar visualmente el comportamiento del sistema, lo cual permite cambiar los parámetros que definen el proceso de la simulación, de manera que facilita la experimentación (Vallejos et al., 2017).

En este contexto, resulta crucial evaluar la satisfacción del cliente con el servicio proporcionado, lo que se puede realizar de dos maneras: una medición objetiva del grado de cumplimiento respecto a la calidad requerida y las encuestas dirigidas al cliente

(Morán Ruiz, 2017). Por ejemplo, los modelos matemáticos son aplicados mediante el concepto de teoría de colas como representación de los comportamientos de las líneas de espera en diferentes escenarios (Vega de la Cruz, 2017). De esta manera, se encuentra una estabilidad entre el número de los clientes que se encuentren esperando y la cantidad de cajas bancarias que llevan a cabo el servicio.

El problema de las largas e interminables colas en diferentes agencias bancarias no es nuevo. En muchas ocasiones, los medios de comunicación han informado acerca de la aglomeración de personas en las agencias por falta de organización e información. Los usuarios que se encuentran en cola no solo se forman para realizar sus cobros en la agencia, también los que carecen de información del bono otorgado por el gobierno, por ejemplo (Redacción RPP, 2021). En este caso, es fundamental destacar la relevancia de establecer una efectiva comunicación con el cliente, la información que se ofrece y cómo se relaciona con la empresa, debido a que la satisfacción del cliente se ve reflejada en los resultados del crecimiento de una organización. Por ello, el cliente será el factor decisivo para obtener un impacto positivo en varios aspectos, tales como una mejor relación cliente-empresa, satisfacción, lealtad, competitividad, entre otras (Robayo, 2017).

El objetivo de este estudio es identificar posibles mejoras en el servicio de atención al cliente en una agencia bancaria de Lima Metropolitana, Perú, utilizando la teoría de colas y el *software* de modelación Arena. Se busca medir e identificar las variables más influyentes en el servicio brindado, representar el comportamiento del sistema y predecir posibles impactos. Se aborda el problema de las largas colas y tiempos de espera en las agencias bancarias, y se destaca la importancia de la comunicación con el cliente y la satisfacción como factores clave para el crecimiento de una organización.

El artículo sigue la siguiente estructura: la sección 2 examina el estado del arte, la sección 3 describe la metodología, la sección 4 expone los hallazgos del estudio o resultados, la sección 5 se dedica a la discusión de los mismos y, por último, la sección 6 aborda las conclusiones del trabajo.

## **2. ESTADO DEL ARTE**

### **2.1 La teoría de colas y el servicio al cliente**

Cada vez son más las industrias de servicios en las que se forman las denominadas líneas de espera para la atención, las cuales llegan a ser un impedimento para alcanzar la satisfacción del cliente. Según Cárdenas Estrada et al. (2019), las situaciones de espera surgen cuando los clientes se acercan solicitando un servicio, pero este no se encuentra disponible, lo cual los lleva a decidir si desean esperar para recibirlo; si la espera se extiende excesivamente, los clientes podrían elegir retirarse, lo que resultaría en pérdidas para la empresa. La teoría de colas tiene como principal objetivo lograr una

estabilidad en el sistema y establecer una capacidad de servicio suficiente que asegure un balance entre el factor cuantitativo, relacionado con los costos del sistema, y el factor cualitativo, relacionado con la satisfacción del cliente con el servicio brindado (Portilla et al., 2010).

La estructura de un sistema de colas se compone de los siguientes elementos. Los clientes que requieren un servicio emergen en el transcurso de tiempo en una fuente de entrada. Posteriormente, ingresan al sistema para unirse a una cola de espera. En un momento específico, se elige un cliente de la cola para recibir el servicio, utilizando alguna regla conocida como disciplina de la cola. Se proporciona el servicio solicitado por el cliente a través de un mecanismo de atención, se le atiende y posteriormente el cliente abandona el sistema de colas (Hillier & Lieberman, 2010).

Proporcionarle un servicio de calidad satisfactorio al cliente constituye uno de los principales retos para toda organización, pues supone un proceso de alcanzar la satisfacción del cliente con el servicio proporcionado. Las empresas se esfuerzan considerablemente por ofrecer a los clientes una variedad de actividades que cuentan con una esencia única, con el objetivo de lograr que se sientan satisfechos y perciban que el valor invertido por su adquisición es correspondido (García, 2016). El servicio prestado por las agencias bancarias tiene un rol cada vez más importante en la sociedad moderna; la globalización ha motivado a las instituciones financieras a competir a través de la inversión tecnológica e innovación como una estrategia para destacarse y diferenciarse con el objetivo de garantizar la lealtad de los clientes (Morillo et al., 2011).

## **2.2 Antecedentes de la teoría de colas en diferentes escenarios**

Desde el punto de vista de López Hung y Joa Triay (2018), el uso de la teoría de colas en diferentes escenarios ayuda a consolidar una base teórica tanto del tipo de servicio como del diseño, que puede lograrse mediante modelos matemáticos. Vinculando a lo anteriormente expuesto, se encontraron los siguientes aportes al estado del arte del presente artículo.

En primer lugar, al aplicarse la metodología mencionada en una agencia bancaria brasileña, se demostró que una buena relación entre la demanda y la capacidad tiene como resultado un tiempo de espera menor para los clientes, así como la obtención de información necesaria para proponer mejoras al sistema, lo que abarca a la empresa y la satisfacción del cliente (Entriger, 2020). Por consiguiente, Wang et al. (2020) enfatizan que se debe determinar una estrategia óptima de las colas, ya que presentan muchos factores, lo que dificulta la minimización de las líneas de espera; esto es posible mediante la modelación de los sistemas con base en la teoría de colas para simularlas. En síntesis, para diseñar un modelo con enfoque de la teoría de colas, es necesario obtener los datos que tengan la mayor influencia en el comportamiento del sistema para lograr el

establecimiento de un perfil o estrategia de mejora de la situación actual (Longaray et al., 2016).

En un contexto similar, como las líneas de espera originadas en el sector de salud, Song et al. (2015) encuentran que la gestión de colas permite que los pacientes tengan un menor tiempo de espera, de manera que se genere una mayor capacidad para brindar el servicio médico y se mejore su calidad. Asimismo, en un centro de ortopedia, fue posible determinar una alta probabilidad de que los tiempos en la cola sean de cinco minutos en promedio, por lo que se trazaron estrategias para reducirlo y mejorar la calidad de la estancia (Vega de la Cruz et al., 2017). De igual manera, Peter y Sivasamy (2019) señalan que la teoría de colas demuestra eficacia en este tipo de sistemas de gestión, aportando un marco de ciertos aspectos para la satisfacción de los pacientes, a pesar de que existen situaciones en la que no será posible la reducción de tiempos. En este sentido, Komashie et al. (2015) destacan que es de gran importancia determinar los factores de mayor impacto en la satisfacción del cliente.

La teoría de colas busca determinar el funcionamiento y analizar el comportamiento de las líneas de espera en distintos escenarios para plantear estrategias que permitan lograr la capacidad óptima para brindar un buen servicio (Cárdenas Estrada et al., 2019). En el sector de servicios alimentarios, se analizó el caso de Burger Ranch, un restaurante que presentaba un problema porque su capacidad de atención no era suficiente para la cantidad de clientes que atraía a ciertas horas del día, por lo cual perdía ventas, ya que los clientes simplemente se retiraban u optaban por otras opciones. Según Dorda et al. (2019), con un buen diseño para la prestación de un servicio idóneo, se puede estimar el número óptimo de servidores que deberían estar en el sistema para minimizar las pérdidas económicas; y satisfacer las necesidades del cliente brindándole un mejor servicio de calidad y eficacia. En el sector industrial, se halló el caso de una empresa de máquinas que buscaba optimizar el número de reparadores con respecto a los costes del sistema, debido a que se presentaban continuamente averías, fallos u otros problemas que necesitaban ser solucionados para cumplir con las funciones respectivas; y para optimizar la gestión de sus procesos, se aplicó la teoría de colas.

Otras investigaciones sobre las líneas de espera y la atención en el servicio brindado al cliente se enfocan en estrategias para la organización de los clientes. En el primer estudio, se buscaba un mayor orden en el sistema de las colas, por lo cual se desarrolló el diseño de un modelo para asignar la prioridad y el orden en el que los clientes llegaban para recibir un servicio. Para Kittsteiner y Moldovanu (2016), hay dos razones para otorgar prioridad a algunos clientes: la primera, porque puede conducir a una reducción en los costos generales de las colas; y la segunda, porque los clientes estarán dispuestos a pagar un poco más con tal de ser atendidos en el menor tiempo posible, con lo cual la prestadora de servicios se beneficia debido al incremento de ingresos. Otro estudio

relacionado aborda la implementación de precios cíclicos a un servicio; un ejemplo de ello es el caso de Estados Unidos, donde muchos parques temáticos ofrecen descuentos de marzo a octubre, la temporada alta de turismo (Huang et al., 2019). Se demostró que, al aplicar precios cíclicos, tanto los clientes como el proveedor se ven beneficiados, ya que la prestadora de servicios incrementa su rentabilidad y no se ve afectada la calificación del cliente. El último estudio comprende el desarrollo de estrategias para reducir los costos por el tiempo de espera mediante la implementación de prioridad de colas aplicada a una agencia de servicios. De acuerdo con Gavirneni y Kulkarni (2016), la segmentación de clientes es el concepto por el cual los precios se cobran en función de la capacidad disponible para determinar mejor lo que un cliente está dispuesto a pagar por un producto o servicio; ello permite a los proveedores aumentar sus ganancias, hacer eficiente el sistema y, al mismo tiempo, otorgar al cliente un buen servicio.

Finalmente, Dorda et al. (2019) agregan que es de vital importancia designar el servicio para cada cliente que ingrese al sistema, ya que a partir de esta información se puede optimizar el número de servidores para mejorar la eficiencia y evitar pérdidas de tiempo valioso mediante los modelos matemáticos ejecutados en una simulación. Estas simulaciones se caracterizan por su rapidez en la ejecución y su alta precisión en comparación con otros métodos, tomando en cuenta las diferentes variaciones de los factores y sus niveles de clasificación (Zengin et al., 2016).

### **3. METODOLOGÍA**

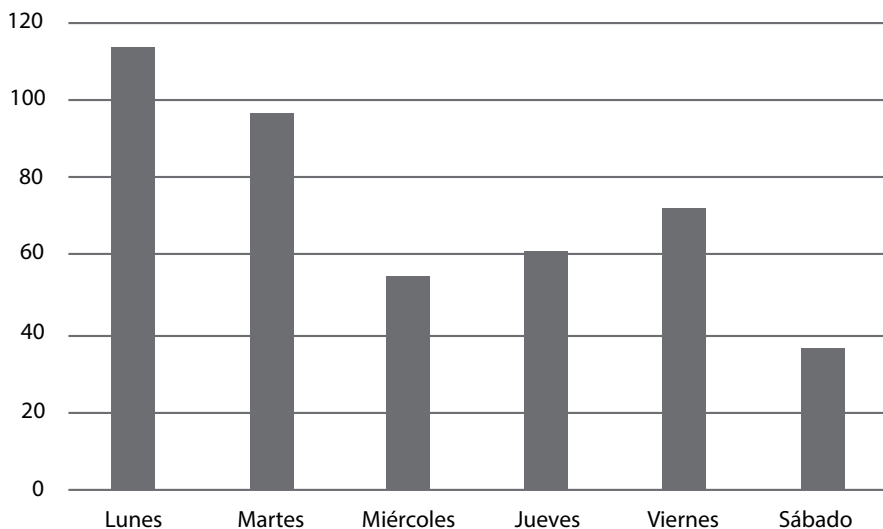
En la presente investigación, se empleó un enfoque cuantitativo con un alcance explicativo y descriptivo. Se describe el método por el cual se realizó la obtención de datos; además, se incluyeron conceptos relacionados con el diseño de modelos y la aplicación de colas, que son relevantes para el desarrollo de este trabajo (Entriger, 2020).

#### **3.3 Recopilación de datos**

La recolección de datos para establecer la base del modelo de simulación se hizo mediante la toma de tiempos por turnos durante dos meses. La población objetivo fueron los clientes de una agencia bancaria en el área de servicio de atención. De esta manera, la distribución que tienden a seguir se analizó mediante la herramienta Input Analyzer de Arena. Cabe destacar que se tuvo en cuenta las horas y los días pico. La Figura 1 presenta la cantidad de arribos promedio en la semana, que, en su mayoría, son los días lunes y martes; para este estudio, se escogió entre ambos el día lunes.

**Figura 1**

*Arribos promedio a lo largo de la semana*



**Tabla 1**

*Composición de los usuarios en una agencia bancaria en un día pico*

Tipo de cliente	Porcentaje (%)
Cliente 1 (regular)	75,14
Cliente 2 (preferencial)	24,86

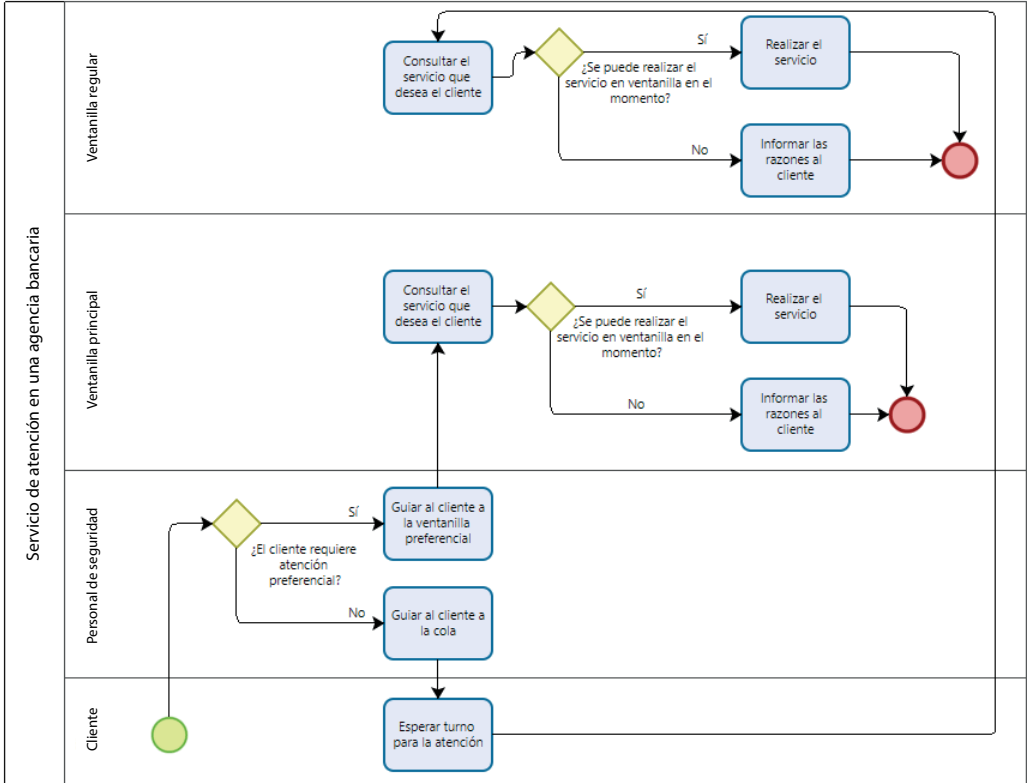
En el caso de la hora pico, se observó que a partir de las cinco de la tarde aumentaba la demanda, pues solo falta una hora antes de que cierre la agencia bancaria. De este modo, los usuarios que llegan a última hora generan las aglomeraciones y hacen que el avance de la cola se vuelva más lento. Es importante indicar que, para el estudio, se asume que los agentes de ventanilla tienen una misma eficiencia, puesto que todos son capacitados por igual y al momento del reclutamiento el personal es calificado por sus competencias y habilidades.



### 3.2 Diseño del modelo y definición de data

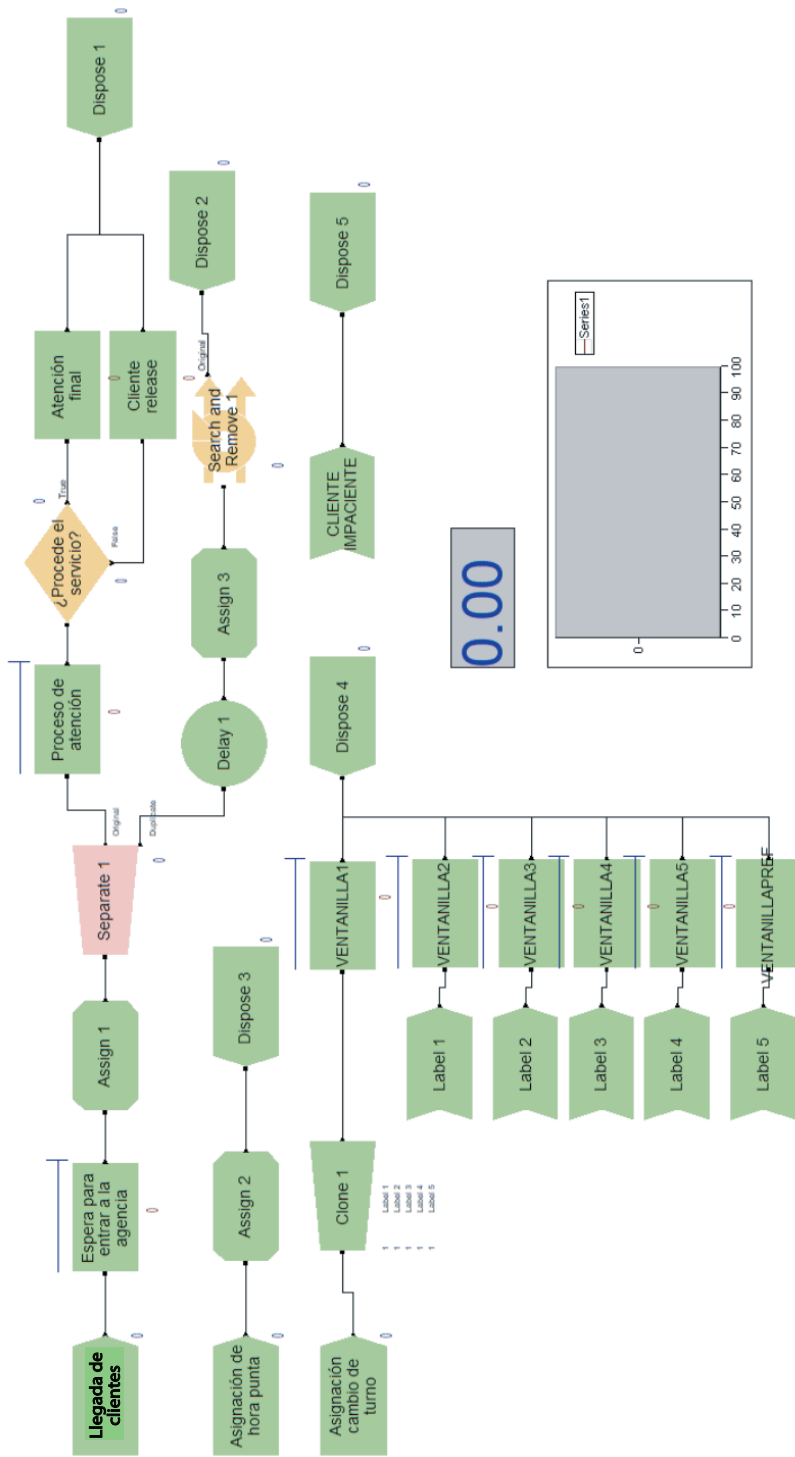
Figura 2

Diagrama de procesos



**Figura 3**

Modelo del sistema actual



En primer lugar, para modelar el escenario actual del sistema en ventanilla de la agencia bancaria, se tomó en cuenta los datos del día lunes, que fue seleccionado como día pico. En la información recolectada, se observó que existen dos tipos de clientes, el regular y preferencial, los cuales se designan como cliente 1 y cliente 2, respectivamente. El horario de atención del banco comprende desde las 9:00 horas hasta las 18:00 horas, con un aumento del número de clientes a partir de las 17:00 horas. Por otro lado, no se considera una hora específica de almuerzo, ya que existe un doble turno; sin embargo, durante el cambio de turno, que es realizado por todos los cajeros, se presenta una espera adicional para los clientes de aproximadamente diez minutos como máximo. En segundo lugar, ante la llegada de los clientes a la agencia bancaria, el agente de seguridad los guía a la cola para ser atendidos en ventanilla según el tipo de cliente. El banco cuenta con seis ventanillas regulares y una preferencial, que atiende exclusivamente a los clientes designados; no obstante, cabe la posibilidad de que si el cajero de usuario regular está libre, el cliente preferencial pueda usarlo, y viceversa. Asimismo, se observó que después de una espera de entre 25 y 30 minutos como máximo, los clientes se retiran. En tercer lugar, al iniciarse el servicio, se consulta si este procede, ya que existen casos en los cuales no es posible brindar el servicio por ventanilla o no es necesario. Finalmente, los clientes que no reciben el servicio se retiran, y los que permanecen, salen del sistema luego de su consulta.

**Tabla 2**

*Definición de data actual*

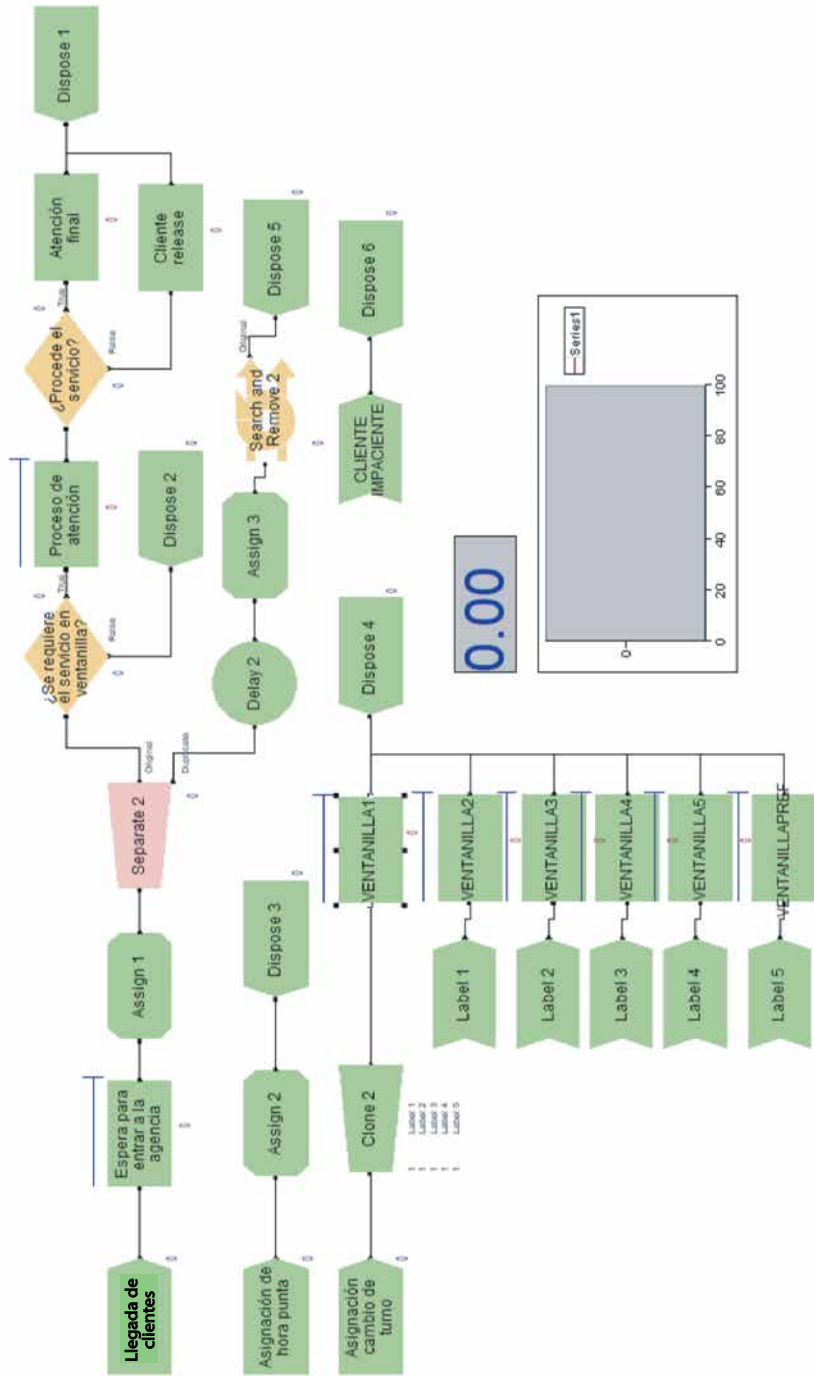
Módulo	EntityType	Descripción	Expresión
Llegada de clientes	Cliente	Cliente que ingresa al sistema.	Unif(Vmin, Vmax, 1)  First Creation: Unif(3,5)
Asignación de hora punta	HoraPunta	Ingresar nuevos valores para la llegada de clientes.	First Creation: 8
Asignación cambio de turno	CTurno	Define los cambios de turno por ventanilla.	First Creation: 5
Assign1	Variable y atributo	Define el tipo, variable tiempo y número de clientes.	Variable, CUENTA, CUENTA+1 Attribute, NUMERO, CUENTA Attribute, TIEMPO, Unif(25,30) Attribute, TIPOCLIENTE, disc(0.75,1,1,2)

*(continúa)*

(continuación)

Módulo	EntityType	Descripción	Expresión
Assign2	Variable	Definición de nueva distribución en hora punta.	Variable,Vmin,1 Variable,Vmax,2
Assign3	Variable	Selecciona el número de clientes que sobrepasan el tiempo máximo de espera.	Variable,BUSQUEDA, NUMERO
Separate1	Duplicate Original	Duplicar al cliente para simular a los que se retiran después de 10 a 15 minutos de espera.	50%, 1# duplicate
Delay1	Delay	Define a los clientes que deciden retirarse.	Delaytime: TIEMPO
Search and remove	Search a Queue	Los clientes definidos en el delay1 se retiran del sistema.	BUSQUEDA==NUMERO Destination Label CLIENTE IMPACIENTE
Clone1	Clone	Clona el recurso de ventanillas para organizar los cambios de turno.	1,Label1/2,Label2... /5,Label5
Cliente impaciente	Label	Retira a los clientes impacientes.	CLIENTE IMPACIENTE
¿Procede el servicio?	Decide	Separa a los clientes que proceden con el servicio en ventanilla.	Percent true: 60%
Espera para entrar a la agencia	Seize Delay Release	El cliente realiza una consulta al vigilante.	Unif(1,2) Resource,VIGILANTE,1
Proceso de atención	Seize Delay	Según el tipo de cliente, se brinda el servicio inicial.	Set, VENTANILLAS, 1, Specific member, TIPOCLIENTE Unif(9,11)
Atención final	Delay Release	Se realiza la operación final deseada por el cliente.	Set, VENTANILLAS, 1, Specific member, TIPOCLIENTE Unif(15,18)
ClienteRelease	Delay Release	Módulo para liberar a la entidad cliente.	Set, VENTANILLAS, 1, Specific member, TIPOCLIENTE
VENTANILLA 1-5	Seize Delay Release	Módulos para realizar el cambio de turno por ventanilla.	Resource, VENTANILLA,1
VENTANILLA PREF.			Constant(10)

**Figura 4**  
Propuesta de modelo para el proceso de simulación



Al analizar el escenario original, se determinó que el sistema puede mejorarse tanto en ahorro de tiempos como en ofrecer una mejor calidad de servicio al cliente, estableciendo un mejor balance entre la demanda y la disponibilidad de recursos con los que cuenta (Entriger, 2020). Esto se puede lograr introduciendo al sistema una nueva persona encargada de preguntar a las personas por el servicio que buscan e informarles la manera en la que puede llevarse a cabo. Por lo tanto, se reduce el número de personas en cola, el porcentaje de clientes que ya no puede proceder con el servicio por ventanilla y los que se retiran por esperar demasiado en la cola. El nuevo personal orientará a los usuarios en cola si el servicio que buscan recibir se realiza en ventanilla o se puede hacer por otro medio como plataforma o cajero; de esta manera, se ahorrará tiempo de espera haciendo cola en ventanilla para recibir un servicio al que es posible acceder por otro medio.

**Tabla 3**

*Cambios en la definición de data*

Módulo	Inputdata	Descripción	Expresión
Decide EscOriginal Procede el servicio	2-way chance	Se refiere a si puede atender el servicio de atención en ventanilla.	65% Percent true (no varía).
Decide EscMejora Procede el servicio	2-way chance	Se refiere a si puede atender el servicio de atención en ventanilla.	85% Percent true.

### 3.4 Técnicas e instrumentos

En la aproximación al comportamiento de colas en el servicio al cliente, se tuvo en cuenta que, para realizar el análisis, se necesitó determinar diferentes probabilidades, entre ellas, la existencia o no de las líneas de espera, el tiempo que el cliente decide permanecer en la cola, el rendimiento del proceso y la satisfacción por parte del cliente (López Hung & Joa Triay, 2018). En este sentido, la teoría de colas proporciona una base teórica y el modelamiento o simulación de los sistemas, el cual puede variar según la complejidad del escenario que se va a investigar y el objetivo con el que se realiza; en este caso, es del tipo *input-output*, ya que se insertan datos de entrada para obtener los resultados (Entriger, 2020).

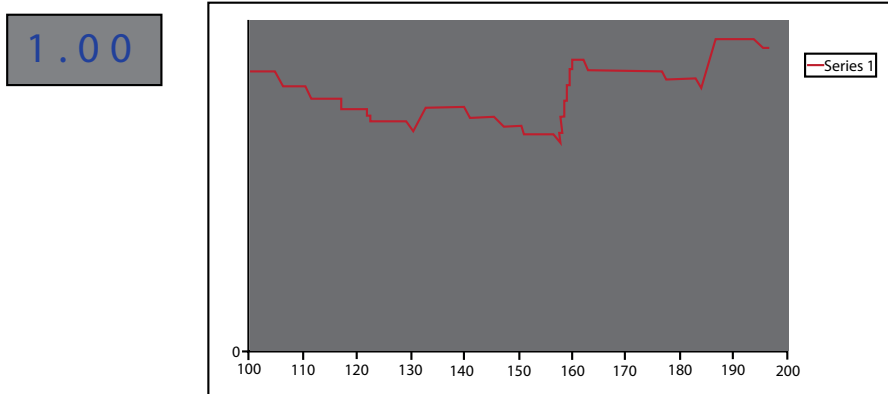
En tal virtud, se utilizó la simulación y diseño de modelos para diseñar y evaluar dos escenarios: el primero describe la situación original y el segundo comprende los cambios sugeridos basados en la teoría de colas; posteriormente, se comparan ambos escenarios mediante el traslape de intervalos de confianza para comprobar la mejora. En este contexto, el instrumento principal es el *software* para modelación de eventos discretos, en este caso, Arena Software Simulation (16.1 Student Version).

### 3.5 Criterios de validez

Para procesar los datos relacionados con el comportamiento de las líneas de espera en una agencia bancaria de Lima Metropolitana, se definió la data de entrada en el simulador: el horario de atención, el porcentaje por tipo de cliente, los tiempos entre llegadas, los cambios por turno, el tiempo de servicio, la proporción de clientes que procedieron con el servicio y el número de ventanillas. Posteriormente, el simulador realizó un total de 30 réplicas con un 95 % de confianza para lograr una representación más exacta respecto a la realidad, y en el reporte se obtuvo los indicadores para evaluar el nuevo escenario con la propuesta de mejora. Asimismo, con ayuda del módulo "Variable" se pudo observar en tiempo real el comportamiento de la simulación (véase la Figura 5).

**Figura 5**

*Validación de modelo vs. realidad chart plot*



Como se puede observar en el *chart plot*, el número de personas en cola empieza a aumentar según las variables mencionadas en el proceso, con lo que se comprueba que el comportamiento de la cola se asemeja al real.

## 4. RESULTADOS

Tras el procesamiento de la data en el Arena Software Simulation para la evaluación de tiempos antes realizada, se obtuvieron los siguientes reportes, los cuales muestran resultados en dos tipos de escenarios: el escenario original y el escenario óptimo. A continuación, se presentan los resultados del primer escenario.

**Tabla 4**

*Escenario original: waiting time*

<i>Waiting time</i>	Media
Proceso de atención. <i>Queue</i>	8,539

**Tabla 5**

*Escenario mejorado: number waiting*

<i>Number waiting</i>	Media
Proceso de atención. <i>Queue</i>	18

En el escenario original, se encontró que el tiempo promedio de espera en cola es de 8,539 minutos, lo cual indica que cuando un cliente va a una agencia bancaria y hace la cola respectiva para realizar algún tipo de operación le va a tomar el tiempo mencionado. De acuerdo con Belás y Gabčová (2016), el reconocimiento de las necesidades financieras de un cliente es esencial para la gestión eficaz de la venta de productos y servicios, pues se obtiene como resultado una experiencia satisfactoria. Por ello, adicionalmente, se considera que el proceso de atención tiene una duración inicial aproximada de diez minutos; antes de definir si es posible realizar la operación en ventanilla, en el caso de que se continúe, se agregan entre quince a dieciocho minutos en promedio; por tanto, el tiempo en sistema del cliente se incrementa, generando incomodidad y su eventual retiro antes de recibir el servicio. Para evitar esto, se debe tener en cuenta la atención limitada que se brinda para trazar las estrategias de diseño de las formas de servicio (Canyakmaz & Boyaci, 2021). Por otro lado, el tiempo de espera puede incrementarse o reducirse, pues es afectado por la hora en que el cliente va al banco, la operación que va a realizar y el número de ventanillas disponibles.

Del mismo modo, en la Tabla 6, se muestra el número de personas que se retiraron de cola por esperar más de un tiempo determinado, que se encontró que era de entre 20 y 25 minutos. Con respecto a ello, se calculó que casi un 14,51 % de los usuarios que ingresa a la agencia bancaria se termina retirando debido al alto tiempo de espera.

**Tabla 6**

*Número de personas que se retiran de la cola*

Número de personas que ingresan al sistema	124
Número de personas que se retiran de la cola	18



Sobre esta base, la investigación busca lograr la optimización del servicio, así como la eficiencia de los recursos implicados dentro del sistema para mejorar el proceso de atención al cliente de la agencia bancaria y así reducir el tiempo de espera en cola (Nascimento et al., 2021). Debido a esto, se incorporó un nuevo recurso: un personal que pregunte a los clientes en cola la operación que van a realizar, lo que permite agilizar los movimientos y ahorrar tiempo, pues existen otros medios para hacer las mismas operaciones, ya sea mediante la plataforma digital, la banca por internet o la aplicación del banco. Otras posibilidades, entre otras, son que dicha operación no se realice en ventanilla o que el cliente solo haya ido para hacer una consulta, y por eso muchos clientes no culminan el servicio y se retiran sin realizar operación alguna. De igual manera, el resultado del tiempo de espera por la atención depende de quien la ejecute; en otras palabras, el empleado a cargo de la ventanilla también es responsable de la demora a la hora de prestar el servicio (Entriger, 2020). Finalmente, todo lo mencionado genera un mayor tiempo de espera innecesario en la cola y molestias en los clientes que realmente van a hacer uso del servicio. En la Tabla 7, se muestran los resultados con la aplicación de las mejoras respectivas.

**Tabla 7**

*Propuesta de sistema: waiting time*

<i>Waiting time</i>	Media
Proceso de atención. <i>Queue</i>	4,0461

**Tabla 8**

*Propuesta de sistema: number waiting*

<i>Number waiting</i>	Media
Proceso de atención. <i>Queue</i>	5

En los resultados del escenario óptimo, se hallaron mejoras significativas. El tiempo de espera promedio en cola ahora es de 4,0461 minutos y el número de personas promedio en cola es de 5. El tiempo promedio de espera en cola disminuyó en más de la mitad; y el número de personas en cola se redujo en más de 10 personas. Igualmente, con las nuevas medidas, el porcentaje de personas que procede y culmina el servicio pasó de un 85,48 % a un 95,96 %.

Los resultados demuestran que un pequeño cambio puede ser significativo para la optimización de un proceso, y ello se evidencia en el tiempo de espera menor que en el escenario original. Como señalan Cárdenas Estrada et al. (2019), no solo se agiliza el proceso de atención, sino que también el cliente se siente bien y a gusto con el servicio

que está recibiendo. Entriger (2020) afirma que un mal servicio al cliente puede generar diferentes reacciones de su parte, como disminuir su preferencia hacia la tienda, cambiar de producto o marca, provocar quejas o incluso hacer *marketing* negativo con personas cercanas.

Con respecto al número de personas que se retiraron de la cola, este disminuyó significativamente. Con estos resultados se demuestra que casi todas las personas que ingresan al banco llegan a realizar sus respectivas operaciones y no se retiran, lo que se debe principalmente a la disminución del tiempo de espera en la cola. En la Tabla 9, se observa que con las mejoras aplicadas el porcentaje de personas que se retiraron de la cola pasó de 14,52 % a 4,032 % con respecto a las que ingresaron a la agencia bancaria.

**Tabla 9**

*Número de personas que se retiran de la cola*

Número de personas que ingresan al sistema	124
Número de personas que se retiran de la cola	5

En ese sentido, para comprobar si las diferencias eran significativas, se eligió como indicador el “AverageWaitTime: TAVG(Proceso de atención.Queue.WaitingTime)”, el cual es la medida del tiempo que esperan los clientes antes de ingresar a ventanilla. Así se generaron los cálculos para definir intervalos independientes de confianza para cada escenario (véase la Tabla 10).

**Tabla 10**

*Intervalos de confianza*

	Límite inferior	Límite superior
Escenario actual	1,9	19,36
Propuesta de modelo	1,96	10,08

**Figura 6**

*Prueba con Output Analyzer*



Se observó que los intervalos de confianza presentan una superposición entre ellos, por lo que fue necesario utilizar la herramienta Output Analyzer para realizar una prueba pareada. Esta dio como resultado extremos del intervalo de las diferencias pareadas positivos (+;+), lo que indica que el escenario mejorado es considerado superior por sus valores. En ese sentido, se logró la disminución del número de personas que se retiran del sistema y una atención en un tiempo más corto, mejorando la calidad de servicio, así como una reducción de tiempos, aumentando la productividad (Entriger, 2020).

De esta manera, basándonos en los conceptos que aplica la teoría de colas, hemos determinado la forma en que operan las líneas de espera, teniendo en cuenta diferentes escenarios, y es posible aplicar mejoras al sistema para alcanzar un balance con el objetivo de brindar un mejor servicio (Cárdenas Estrada et al., 2019). Asimismo, el *software* permitió la simulación del comportamiento del sistema, sin restricción al cambiar los parámetros para la experimentación respecto a las diferencias en los tiempos del indicador analizado (Vallejos et al., 2017). Por lo tanto, se diseñó un modelo de simulación que permite medir la mejora del servicio al cliente mediante la aplicación de la teoría de colas y el uso del *software* Arena (16.1 Student Version).

## 5. DISCUSIÓN DE LOS RESULTADOS

Después de mostrar de manera detallada los resultados, es preciso resaltar los hallazgos de esta investigación. En primer lugar, con respecto al tiempo promedio de espera en cola, en el primer escenario, se detalla cómo es el funcionamiento de las líneas de espera, y se comprueba que el tiempo y el número de personas en cola tienden a aumentar porque una gran cantidad de personas, una vez dentro del sistema, no terminan el proceso y se retiran; a ello se añade el hecho de que los clientes preferenciales tienen una prioridad más alta y pasan a la ventanilla designada. Muchas veces, los usuarios que requieren servicios de una agencia bancaria desconocen que las operaciones que quieren realizar se pueden llevar a cabo a través de otros medios, como la página web, agentes, aplicativos, cajeros, plataformas digitales, entre otros. Sin embargo, no deja de ser necesaria la asistencia presencial a las agencias bancarias para operaciones específicas que solo se pueden realizar por ese medio. Por ello, es indispensable para la agencia bancaria brindar un buen servicio y de calidad a sus usuarios, principalmente en cuanto a las esperas en cola que suelen ser un motivo de molestia para los clientes.

Por ejemplo, para abrir una cuenta de ahorros no es necesario ir a la agencia, pues hoy en día existe la opción de apertura a través de la página web del banco solo con el DNI, y tan solo se tendría que ir para recoger la tarjeta. Del mismo modo, se puede hacer un depósito a través de los agentes o los cajeros; y si se es cliente del banco, se puede realizar a través de transferencia, ya sea por la página web o el aplicativo.

Mencionando esto, una vez aplicada la propuesta de mejora del sistema, se observa que el tiempo y las personas en cola se han minimizado considerablemente, pues se optó por implementar en el sistema un nuevo personal que tenga la función de preguntar a las personas en cola la operación que van a realizar y, de ese modo, asesora a los usuarios que no terminen la operación o a los que puedan realizarlas por otros medios de manera más rápida y eficiente, con el resultado de evitar hacer colas innecesarias.

Si se compara los resultados obtenidos con los de otros estudios similares, mencionados en la sección sobre el estado del arte, se encuentra que, según Dorda et al. (2019), es relativamente importante designar primero el servicio que va a recibir el cliente. Por eso, se realizó algo similar en el presente estudio, pues los usuarios no reciben el servicio sin previamente haber sido consultados por la operación que van a hacer —esta es la función del nuevo personal— y de esta manera se optimizan los tiempos de espera. Así se consigue que el cliente esté satisfecho.

En ese sentido, se analizó el sistema para determinar el proceso que representaba un problema para lograr la reducción de tiempos y un mejor servicio de calidad; se determinó la causa raíz y, de acuerdo con esta, se trazaron estrategias para mejorar el sistema (Vega de la Cruz et al., 2017). De esta manera, la teoría de colas demostró su eficiencia para cumplir los objetivos establecidos en este caso de estudio; sin embargo, existen aspectos que pueden variar el punto de vista de cada cliente respecto a la calidad del servicio, los cuales afectan al tiempo de la propuesta (Peter & Sivasamy, 2019).

## 6. CONCLUSIONES

El desarrollo del modelo de simulación en una agencia bancaria logró replicar el comportamiento de la línea de espera mediante la aplicación de la teoría de colas como base, y se pudo comprobar la mejora de atención al cliente por la reducción de tiempos de espera, lo cual fue validado mediante un *chart plot*.

Asimismo, con la propuesta se logró la reducción del porcentaje de usuarios que se retiraban por molestia respecto al excesivo tiempo de espera en cola, debido a que el modelo planteado disminuye este lapso realizando una consulta a los clientes mientras esperan ser atendidos. Estos disminuyeron en más de un 50 %, pasando de 14,52 % a 4,032 %.

De igual manera, cabe destacar que existen diferentes variables en los clientes, las cuales no solo dependen de un menor tiempo de atención para calificar un servicio como de calidad. También es necesario asumir la eficiencia de los agentes de ventanilla, ya que existen variados factores que podrían aumentar o disminuir su eficiencia personal, errores en el sistema, entre otros. Estos estarían limitando los resultados del presente estudio.

Adicionalmente, con el uso del *software* Arena (16.1 Student Version) se demostró consistencia con el sistema real, ya que en los datos del reporte se obtuvieron tiempos casi similares a los conseguidos en el estudio para el análisis de datos.

Para finalizar, los resultados demostraron la mejora del sistema al adicionar una entidad extra, encargada de ayudar y guiar a los clientes mientras esperan en cola, reduciendo el número de personas en espera y aumentando la calidad del servicio; sin embargo, por el factor de los costos, se recomendaría activar este recurso a partir de la hora pico establecida.

## REFERENCIAS

- Belás, J., & Gabčová, L. (2016). The relationship among customer satisfaction, loyalty and financial performance of commercial banks. *Ekonomie a Management*, 19(1), 132-147. DOI: 10.15240/tul/001/2016-1-010. <https://doi.org/10.15240/tul/001/2016-1-010>
- Canyakmaz, C., & Boyaci, T. (2021). Queueing systems with rationally inattentive customers. *Manufacturing & Service Operations Management*, 25(1), 1-22. DOI: 10.1287/msom.2021.1032. <https://doi.org/10.1287/msom.2021.1032>
- Cárdenas Estrada, R., Pérez Pin, M., Tejada Solórzano, A., & Cevallos Torres, L. (2019). Aplicación de un modelo híbrido de teoría de colas y algoritmo evolutivo para medir la optimización en el servicio de atención al cliente en un local de comida rápida. *Ecuadorian Science*, 3(1), 15-22. <https://doi.org/10.26911/issn.26028077vol3iss1.2019pp15-22p>.
- Dorda, M., Teichmann, D., & Graf, V. (2019). Optimisation of service capacity based on queueing theory. *MM Science Journal*, October, 2975-2981. [https://doi.org/10.17973/mmsj.2019\\_10\\_201889](https://doi.org/10.17973/mmsj.2019_10_201889)
- Entriger, T. C. (2020). Simulation and analysis of queues in banks: A case study of an agency in the Southern State of Rio de Janeiro. *Independent Journal of Management & Production*, 11(3), 892-907. <https://doi.org/10.14807/ijmp.v11i3.1074>
- García, A. (2016). Cultura de servicio en la optimización del servicio al cliente. *Telos*, 18(3), 381-398. <https://www.redalyc.org/pdf/993/99346931003.pdf>
- Gavirneni, S., & Kulkarni, V. G. (2016). Self-selecting priority queues with Burr distributed waiting costs. *Production and Operations Management*, 25, 979-992. <https://doi.org/10.1111/poms.12520>
- Hillier, F., & Lieberman, G. (2010). *Introducción a la investigación de operaciones* (9.ª ed.). The McGraw-Hill Companies.

- Huang, F., Guo, P., & Wang, Y. (2019). Cyclic pricing when customers queue with rating information. *Production and Operations Management*, 28(10), 2471-2485. <https://doi.org/10.1111/poms.13052>
- Kittsteiner, T., & Moldovanu, B. (2005). Priority auctions and queue disciplines that depends on processing time. *Management Science*, 51(2), 236-248. <https://doi.org/10.1287/mnsc.1040.0301>
- Komashie, A., Mousavi, A., Clarkson, J., & Young, T. (2015). An integrated model of patient and staff satisfaction using queuing theory. *Journal of Translational Engineering in Health and Medicine*, 3. <https://doi.org/10.1109/JTEHM.2015.2400436>
- Longaray, A. A., Munhoz, P. R., Castelli, T. M., & Santos, A. C. (2016). Analysis of queue discipline in services to citizens in social security agencies: A case study. *Revista de Gestão em Sistemas de Saúde*, 5(1), 105-119.
- LópezHung, E., & JoaTriay, L. (2018). Teoría de colas aplicada a la estudi del sistema de servicio de una farmacia. *Revista Cubana de Informática Médica*, 10(1), 3-15. [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1684-18592018000100002&lng=es&tlng=en](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1684-18592018000100002&lng=es&tlng=en)
- Morán Ruiz, E. (2017). La calidad del servicio al cliente externo administrativo vinculado al incremento de las ventas en el grupo Santillana. *Espirales. Revista Multidisciplinaria de Investigación*, 9, 82-104. <http://www.revistaespirales.com/index.php/es/article/view/221/168>
- Morillo, M. del C., Morillo, M. C., & Rivas, D. (2011). Medición de la calidad del servicio en las instituciones financieras a través de la escala de Servqual. *Contaduría y Administración*, 234, 101-130. [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S018610422011000200006](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S018610422011000200006)
- Nascimento, M. A. R., dos Santos, L. M., da Silva, A. M., Bueno, R. C., Machado, S. T., & Tanaka, W. Y. (2021). Discrete event simulation applied to single queue management: A case study at a bank agency. *Independent Journal of Management & Production*, 12(9), S831-S842. <https://doi.org/10.14807/ijmp.v12i9.1632>
- Peter, P., & Sivasamy, R. (2019). Queueing theory techniques and its real applications to health care systems – Outpatient visits. *International Journal of Healthcare Management*, 14, 114-122. <https://doi.org/10.1080/20479700.2019.1616890>
- Portilla, L. M., Arias Montoya, L., & Fernández Henao, S. A. (2010). Análisis de líneas de espera a través de teoría de colas y simulación. *Scientia et Technica*, 17(46), 56-61.
- Redacción RPP. (2021, 17 de febrero). *Largas colas afuera de las sedes del Banco de la Nación en primer día de pago del Bono de S/600*. RPP Noticias. <https://rpp.pe/lima/actualidad/coronavirus-en-peru-largas-colas-afuera-de-las-sedes->

del-banco-de-la-nacion-en-primer-dia-de-pago-del-bono-de-s-600-noticia-1321387

- Robayo, A. (2017). *La importancia del servicio al cliente y el reflejo de las ventas en una empresa* [Tesis de grado, Universidad Militar Nueva Granda]. Repositorio Institucional UMNG. <https://repository.unimilitar.edu.co/bitstream/handle/10654/16203/RobayoRodriguezAlejandro2017.pdf?sequence=1&isAllowed=y>
- Song, H., Tucker, A., & Murrell, K. (2015). The diseconomies of queue pooling: An empirical investigation of emergency department length of stay. *Management Science*, 61(12), 3032-3053. <http://dx.doi.org/10.1287/mnsc.2014.2118>
- Vallejos, R. D., Mariño, S., & Alfonzo, P. L. (2017). Teoría de colas. Propuesta de un simulador didáctico. *Revista Publicando*, 13(1), 5-20. <https://core.ac.uk/download/pdf/236645186.pdf>
- Vega de la Cruz, L. O., Leyva Cardeñosa, E., Pérez Pravia, M., & Tapia Claro, I. I. (2017). La teoría de colas en la consulta de ortopedia. *Revista Cubana de Ortopedia y Traumatología*, 31(2), 1-13. [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S0864-215X2017000200002&lng=es&tlng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-215X2017000200002&lng=es&tlng=es)
- Wang, X.-L., Wen, Q., Zhang, Z.-J., & Ren, M. (2020). The optimal queuing strategy for airport taxis. *IEEE Acces*, 8, 208232-208239. <https://doi.org/10.1109/ACCESS.2020.3038176>
- Zengin, I., Vardakas, J., Zorba, N., & Verikoukis, C. (2016). Analysis and quality of service evaluation of a fast charging station for electric vehicles. *Energy*, 112, 669-678.