



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CAMPUS AGRESTE
NÚCLEO DE TECNOLOGIA
CURSO DE ENGENHARIA DE PRODUÇÃO

WINICIUS ANTONIO SOUZA SILVA

**PLANEJAMENTO DA CAPACIDADE DE ATENDIMENTO EM UMA
EMPRESA DE SERVIÇOS PARA VEÍCULOS PESADOS UTILIZANDO A
TEORIA DAS FILAS**

Caruaru

2023

WINICIUS ANTONIO SOUZA SILVA

**PLANEJAMENTO DA CAPACIDADE DE ATENDIMENTO EM UMA
EMPRESA DE SERVIÇOS PARA VEÍCULOS PESADOS UTILIZANDO A
TEORIA DAS FILAS**

Trabalho de Conclusão de Curso apresentado à Coordenação do Curso de Engenharia de Produção do Campus Agreste da Universidade Federal de Pernambuco – UFPE, na modalidade de monografia, como requisito parcial para a obtenção do título de Bacharel em Engenharia de Produção.

Área de concentração: Pesquisa Operacional

Orientador (a): Profº. Dr. Isaac Pergher

Caruaru

2023

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Silva, Winicius Antonio Souza.

Planejamento da capacidade de atendimento em uma empresa de serviços para veículos pesados utilizando a teoria das filas / Winicius Antonio Souza Silva. - Caruaru, 2023.

68 : il., tab.

Orientador(a): Isaac Pergher

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro Acadêmico do Agreste, Engenharia de Produção, 2023.

Inclui referências, apêndices.

1. Teoria das Filas. 2. Serviços. 3. Serviços em veículos pesados. 4. Ordem de serviço. I. Pergher, Isaac. (Orientação). II. Título.

620 CDD (22.ed.)

WINICIUS ANTONIO SOUZA SILVA

**PLANEJAMENTO DA CAPACIDADE DE ATENDIMENTO EM UMA
EMPRESA DE SERVIÇOS PARA VEÍCULOS PESADOS UTILIZANDO A
TEORIA DAS FILAS**

Trabalho de Conclusão de Curso apresentado à Coordenação do Curso de Engenharia de Produção do Campus Agreste da Universidade Federal de Pernambuco – UFPE, na modalidade de monografia, como requisito parcial para a obtenção do título de Bacharel em Engenharia de Produção.

Aprovada em: 30 / 08 / 2023

BANCA EXAMINADORA

Prof. Dr. Isaac Pergher (Orientador)
Universidade Federal de Pernambuco

Prof. Dr. Lucimário Gois (Examinador Interno)
Universidade Federal de Pernambuco

Prof. Dr. Walton Coutinho (Examinador Interno)
Universidade Federal de Pernambuco

AGRADECIMENTOS

Gostaria de expressar minha gratidão mais sincera aos meus pais, Josineide Rita e Valter Antônio, que sempre me ajudaram e me encorajaram durante toda a minha vida acadêmica. Eles me inspiraram, motivaram e guiaram. Eu não teria chegado até aqui sem o seu amor, apoio, carinho e dedicação. Eu gostaria de agradecer por todos os esforços que fizeram por mim, desde a minha infância até a conclusão deste trabalho.

Agradeço aos meus amigos da faculdade, Mirelly Lima, Silvio Dyego e Thaylon Gomes, que compartilharam comigo tantos momentos de aprendizado, alegria e amizade. Vocês foram fundamentais para a minha formação profissional e pessoal. Muito obrigado por todos os conselhos, apoio, risadas, desafios e oportunidades que tivemos juntos. Vocês não são apenas amigos; vocês são uma família que eu escolhi e que sempre vou guardar no meu coração.

Quero expressar minha gratidão sincera à Dilma Santos, mãe do meu amigo Silvio Dyego, que me recebeu em sua casa em momentos de necessidade com tanto amor e generosidade. Ela foi como uma segunda mãe para mim, pois me ajudou, me apoiou e me cuidou. Sou muito grato por ter conhecido essa maravilhosa pessoa, que agora considero parte da minha família. Muito obrigado por todas as suas ações a meu favor.

Desejo agradecer ao meu supervisor de estágio João Pedro, que me forneceu apoio e orientação durante todo o período em que este trabalho foi realizado. Além de ser um profissional competente, atencioso e paciente, ele me forneceu informações valiosas sobre o campo em que atuo. Ele me permitiu desenvolver minhas habilidades, superar os obstáculos e maximizar meu potencial. Ele forneceu um exemplo de virtude, dedicação e excelência que eu sempre respeitarei e admirarei. Muito obrigado por todas as suas ações a meu favor.

Gostaria de expressar minha profunda gratidão ao meu orientador do TCC, Prof^o. Dr. ^o Isaac Pergher, que me guiou e apoiou durante todo o desenvolvimento deste trabalho. Ele foi um mestre, um amigo e um parceiro, que me ensinou muito sobre a minha área de pesquisa. Ele me auxiliou a superar os desafios, a aperfeiçoar as minhas ideias e a concretizar o meu projeto. Ele foi um modelo de sabedoria, ética e excelência, que sempre vou honrar e respeitar. Muito obrigado por tudo o que fez por mim.

Conhecemos o mundo apenas pela janela da mente.
Quando a mente está agitada, o mundo também está.
Quando a mente está em paz, o mundo também está.
Conhecer a nossa mente é tão importante quanto tentar mudar o mundo.
(Haemin Sunim, 2017, p. 25).

RESUMO

As filas são eventos que afetam diretamente a percepção do cliente em relação ao serviço prestado e por conseguinte, no desempenho da organização. Neste sentido, este trabalho utiliza a teoria das filas para analisar o comportamento das filas em um processo de serviços de alinhamento e balanceamento em veículos pesados, visando dimensionar a capacidade do atendimento. Inicialmente foi realizado um mapeamento do estado atual das atividades e foram coletados dados referentes as taxas de chegada e de atendimento, considerando um período entre novembro de 2022 e janeiro de 2023. Para o planejamento de capacidade, foram considerados três cenários alternativos que foram modelados na linguagem Python v.3.10. O cenário atual de atendimento segue os pressupostos de um modelo M/M/1. A partir deste cenário são propostos três cenários alternativos que modificam a taxa de atendimento, o número de atendentes e o grau de variabilidade do sistema. O cenário escolhido reflete os objetivos táticos que a empresa pretende alcançar em nível de satisfação dos seus clientes.

Palavras-chave: Serviços; Teoria das Filas; Serviços em veículos pesados; Ordem de serviço.

ABSTRACT

Queues are events that directly affect the customer's perception of the service provided and consequently, the performance of the organization. In this sense, this work uses Queue Theory to analyze the behavior of queues in a service process of alignment and balancing in heavy vehicles, aiming to dimension the capacity of the service. Initially, a mapping of the current state of the activities was carried out and data related to the arrival and service rates were collected, considering a period between November 2022 and January 2023. For capacity planning, three alternative scenarios were considered that were modeled in Python language v.3.10. The current service scenario follows the assumptions of an M/M/1 model. From this scenario, three alternative scenarios are proposed that modify the service rate, the number of attendants and the degree of variability of the system. The chosen scenario reflects the tactical objectives that the company intends to achieve in terms of customer satisfaction.

Keywords: Services; Queue Theory; Services in heavy vehicles; Service order.

LISTA DE FIGURAS

Figura 1 -	Componentes principais de uma fila	19
Figura 2 -	Sistema Canal único, Atendimento único.....	20
Figura 3 -	Sistema Canal único, Atendimento múltiplo	20
Figura 4 -	Sistema Canal múltiplo, atendimento único	21
Figura 5 -	Sistema Canal múltiplo, atendimento múltiplo	22
Figura 6 -	Etapas do Processo	38
Figura 7 -	Etapas do Processo de Atendimento	41

LISTA DE QUADROS

Quadro 1 -	Notação de Fila 1	25
Quadro 2 -	Notação de Fila 2	25

LISTA DE TABELAS

Tabela 1 -	Parâmetros usados para calcular as amostras de entradas	43
Tabela 2 -	Dados coletados para chegada de clientes	44
Tabela 3 -	Parâmetros usados para calcular as amostras de atendimento	44
Tabela 4 -	Dados coletados para tempos de atendimento	45
Tabela 5 -	Resultados dos cenários	49

LISTA DE ABREVIATURAS

FIFO	<i>First In First Out</i> (primeiro a entrar primeiro a sair)
IA	Inteligência Artificial
IBGE	Instituto Brasileiro de Geografia e Estatística
LIFO	<i>Last in First Out</i> (último a entrar é o primeiro a sair)
OS	Ordem de Serviço
PIB	Produto Interno Bruto
PRI	Prioridade
SIRO	Servidor Atende Quem Chega em Ordem Aleatória

SUMÁRIO

1	INTRODUÇÃO	14
1.1	OBJETIVO GERAL	15
1.2	OBJETIVOS ESPECÍFICOS	15
1.3	JUSTIFICATIVA	15
2	REFERENCIAL TEÓRICO.....	17
2.1	FILAS EM SERVIÇOS.....	17
2.2	TIPOS DE FILAS.....	19
2.3	TEORIA DAS FILAS.....	22
2.4	NOTAÇÃO DE KENDALL	25
2.4.1	Modelo M/M/1	26
2.4.2	Modelo M/Ek/1.....	28
2.4.3	Modelo M/M/s	29
2.4.4	Modelo M/D/1	30
2.5	SÍNTESE DO ESTADO DA ARTE.....	30
3.	METODOLOGIA	36
3.1	TIPO DE PESQUISA.....	36
3.2	COLETA DE DADOS	36
3.3	ANÁLISE DOS DADOS.....	37
3.4	MÉTODO DE TRABALHO	37
4	ESTUDO DE CASO	40
4.1	DESCRIÇÃO DA EMPRESA.....	40
4.2	SITUAÇÃO ATUAL DO SISTEMA.....	40
4.3	DADOS COLETADOS E DEFINIÇÃO DO MODELO DE FILAS	43
5.	ANÁLISE DOS RESULTADOS	46
5.1	CENÁRIO ATUAL.....	46

5.2	CENÁRIO 1	47
5.3	CENÁRIO 2	47
5.4	CENÁRIO 3	48
5.5	COMPARAÇÃO DOS CENÁRIOS	48
5.4.1	Comparação entre cenários	49
5.4.2	Escolha do melhor cenário.....	52
6	CONSIDERAÇÕES FINAIS	55
	REFERÊNCIAS.....	56
	APÊNDICE A – CÓDIGO DE DETERMINAÇÃO DA AMOSTRA DO NÚMERO DE CHEGADAS	61
	APÊNDICE B – CÓDIGO DE DETERMINAÇÃO DA AMOSTRA DO TEMPO DE ATENDIMENTO	62
	APÊNDICE C – CÓDIGO TESTE DE KOLMOGOROV-SMIRNOV POISSON	63
	APÊNDICE D - CÓDIGO TESTE DE KOLMOGOROV-SMIRNOV EXPONENCIAL.....	64
	APÊNDICE E – CÓDIGO PYTHON MODELO M/M/1 CENÁRIO ATUAL	65
	APÊNDICE F – CÓDIGO PYTHON MODELO M/M/1 CENÁRIO 1	66
	APÊNDICE G – CÓDIGO PYTHON MODELO M/M/S CENÁRIO 2	67
	APÊNDICE H – CÓDIGO PYTHON MODELO M/D/1 CENÁRIO 3	68

1 INTRODUÇÃO

As organizações têm desenvolvido planos que norteiam as ações em diferentes períodos de planejamento, a fim de obter uma posição de destaque em relação aos competidores. Em sistemas produtivos de serviços, um dos elementos que influenciam o nível de satisfação dos clientes diz respeito ao tempo gasto aguardando na fila para receber o serviço. Neste sentido, as empresas precisam desenvolver planos de capacidade que comportem diferentes volumes de chegada, de modo a não causar desistências, longas filas ou tempos de atendimento elevados.

Esse problema, combinado com a alta competitividade e o aumento da diversidade de serviços, torna o desenvolvimento de estratégias para sustentar a atividade econômica das empresas muito difícil. O atendimento aos requisitos do cliente torna-se um dos diferenciais competitivos. Como a satisfação do cliente está ligada à retenção do cliente e, portanto, à rentabilidade, a satisfação do cliente está se tornando um dos pilares do marketing, de acordo com Rosa e Kamakura (2001).

Existem diferentes abordagens disponíveis na literatura para subsidiar o estudo de filas e uma delas é denominada teoria das filas (Hillier e Lieberman, 2010). Segundo Costa (2018), essa técnica é bem consolidada na literatura e tem sido aplicada em diferentes contextos, sejam de processos de serviços ou de processos de manufatura. De modo geral, tem-se um conjunto de equações que seguem determinados pressupostos, que são padronizados pela notação de Kendall. Assim, o estudo se inicia pela verificação destes pressupostos à luz daqueles observados no problema.

As filas em serviços de balanceamento e alinhamento de veículos são um problema que afeta tanto os clientes quanto os prestadores de serviço. Os clientes podem ficar insatisfeitos com a demora e a falta de conforto na espera, enquanto os prestadores podem perder clientes e receita por não atenderem à demanda.

Este trabalho concentra-se em uma empresa que fornece serviços de alinhamento e balanceamento de veículos pesados, como caminhões e ônibus, que são procedimentos essenciais para garantir a segurança, o conforto e a economia ao conduzir esses veículos. Ao ajustar o ângulo e a distribuição de peso das rodas, ajuda-se a evitar desgastes irregulares nos pneus, vibrações na direção e desvios na

trajetória. Esta empresa deseja melhorar o nível de serviço oferecido no processo de geração de ordens de serviço, visando principalmente à redução dos tempos de espera na fila.

Neste sentido, este trabalho utiliza a teoria das filas para subsidiar o planejamento da capacidade da empresa sob estudo, a fim de contribuir com ganhos de competitividade, por meio da redução dos tempos de espera na fila.

1.1 OBJETIVO GERAL

Utilizar a teoria das filas para determinar a capacidade de atendimento em uma empresa prestadora de serviços de balanceamento e alinhamento em veículos pesados.

1.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos são:

- Realizar uma pesquisa bibliográfica, buscando, selecionando e analisando as informações publicadas sobre o tema de estudo, a fim de fundamentar teoricamente o trabalho;
- Realizar um mapeamento das atividades relacionadas ao serviço de atendimento;
- Coletar dados determinísticos e estocásticos relacionados as taxas de chegadas e atendimento;
- Construir os cenários que refletem diferentes modelos de filas;
- Escolher o cenário a ser implantado à luz dos objetivos táticos da empresa.

1.3 JUSTIFICATIVA

Esta pesquisa se justifica pelo estudo de filas em uma empresa situada na zona metropolitana do Recife-PE, que presta serviços de balanceamento e alinhamento em veículos pesados. As filas afetam a percepção do cliente em relação ao serviço prestado e, portanto, devem ser consideradas nos planos de determinação da capacidade do sistema. A teoria das filas, que sustenta esse trabalho, tem sido amplamente utilizada em diferentes contextos e também em serviços, como demonstrado na seção 2.5. Além da consolidação desta teoria, que pode ser

observada na literatura, os modelos de filas são de fácil aplicação e proporcionam resultados que contribuem com processos de melhoria contínua nas operações sob análise.

Segundo Lima (2010), as filas são uma realidade comum em muitos setores de serviços, como bancos, supermercados, hospitais, etc. Representam um desequilíbrio entre a demanda do cliente e a capacidade de atendimento do sistema. As filas podem afetar a satisfação do cliente, a qualidade do serviço e a capacidade de funcionamento das organizações. É por isso que é importante gerir eficazmente as filas e reduzir os tempos de espera, otimizar a utilização dos recursos e melhorar a compreensão dos clientes sobre o serviço oferecido.

Os processos de serviço no Brasil possuem uma participação significativa no produto interno bruto (PIB) e na geração de empregos. Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE) de 2022, os processos de serviços representavam 74% do PIB do Brasil. O principal setor econômico do Brasil, o setor de serviços, cresceu 2,9% em 2022, principalmente no primeiro semestre. O setor de serviços abrange uma ampla gama de empresas, incluindo pequenas empresas e instituições financeiras e de ensino. No primeiro trimestre de 2021, o setor de serviços subiu 0,4% em relação ao quarto trimestre de 2020. No entanto, o setor teve uma queda de 0,8% na comparação com o mesmo período de 2020.

Segundo o IBGE em 2022, a participação dos processos de serviços no PIB de Pernambuco foi de 60,1%. O setor de serviços, que abrange uma ampla gama de empresas, desde pequenas empresas até instituições financeiras e de ensino, cresceu 3,2% no ano de 2022, recuperando-se da queda de 4,6% em 2020. O crescimento do PIB de Pernambuco em 2022, de 4,5%, foi impulsionado principalmente pelo setor de serviços, que superou a média nacional de 3,8%.

2 REFERENCIAL TEÓRICO

Este capítulo explora os conceitos da teoria das filas, conduzindo estudos sobre temas-chave e pesquisadores da área. Além disso, destaca elementos importantes que caracterizam cada assunto.

2.1 FILAS EM SERVIÇOS

Para Magalhães (1996), a fila é um processo estocástico que descreve a dinâmica do sistema de atendimento a um grupo de clientes que chegam aleatoriamente, solicitam horários de atendimento aleatórios e saem do sistema após serem atendidos. Ela pode ser interpretada tanto pelos clientes que estão na espera de um serviço quanto pelos servidores que estão sendo atendidos. Além disso, é importante observar que a condição de existência de uma fila é a incapacidade do sistema de atender a demanda de forma eficaz.

Estudar as características de um sistema de filas pode ser importante para a possibilidade de implementar mudanças que melhorem o serviço oferecido. A demanda por um serviço pode chegar a um ponto em que supera sua capacidade de atendimento, resultando em filas e tempos de espera prolongados para os usuários. Os usuários podem ficar irritados e insatisfeitos com a prestação do serviço por causa dessa espera.

Segundo Corrêa e Caon (2002), esperar significa reconhecer que o tempo do consumidor é menos valioso do que o do indivíduo que espera. O gerente de serviços deve ser empático em garantir que os funcionários não coloquem suas necessidades acima das necessidades dos clientes. As filas em todo o mundo precisam ser melhoradas para satisfazer os clientes sem prejudicar os prestadores de serviço e ainda gerar lucro. As empresas podem tomar medidas com o objetivo de diminuir os efeitos prejudiciais desses sentimentos, mas não mudam a duração real da espera.

A experiência de esperar em uma fila é diretamente influenciada pelo local e pelo tempo de espera. Por exemplo, uma pessoa que está na fila à espera de um hospital por quinze minutos pode ouvir um paciente que está sendo atendido gritar ocasionalmente. O paciente terá a impressão de que este período foi muito longo, bem maior do que realmente foi, porque é totalmente desgastante. No entanto, mesmo que o tempo transcorrido na sala de espera seja maior do que na situação anteriormente descrita, a espera não se torna tão traumatizante para o paciente e

pode passar relativamente despercebida se o ambiente tiver revistas com temas intrigantes e um televisor que atraia a atenção.

Existem algumas empresas que se destacam com a preocupação em melhorar a experiência da fila de espera dos seus clientes. Por exemplo:

- Uma empresa chamada Mutant Customer Experience fornece soluções para otimizar o atendimento ao cliente em vários canais, como filas de espera presenciais e digitais. Para melhorar a experiência do cliente em fila de espera, eles sugerem fazer análises da fila atual, criar fóruns para receber sugestões, fornecer entretenimento e informações e usar tecnologias como chatbots e Inteligência artificial (IA) (Mutant Customer Experience, 2023).
- A Take Blip é uma plataforma de chat que permite o desenvolvimento de chatbots para vários canais, como WhatsApp, Facebook Messenger e Telegram, entre outros. Eles demonstram como os chatbots podem ajudar a administrar as filas de atendimento, fornecendo autoatendimento, controle de filas e filas fluidas. Os clientes podem receber atendimento mais rápido, claro e eficiente (BLIP, 2023).
- A Filah produz soluções de gestão de filas para atendimento presencial e online. Eles usam a teoria das filas, para prever congestionamentos, esperas potenciais e formação de filas. Como resultado, eles podem melhorar o atendimento, reduzir custos e aumentar a satisfação do cliente (Filah, 2023).
- A Nubank empresa de serviços financeiros. Eles usam um sistema de fila inteligente que prioriza os casos mais urgentes e complexos e envia mensagens de texto para informar o tempo estimado de espera e o nome do atendente que vai resolver o problema (Nubank, 2023).
- O aplicativo móvel da Starbucks, uma empresa de cafeterias, permite que os consumidores façam pedidos antecipados e evitem filas nas lojas. O aplicativo permite que os clientes paguem pelo celular e acumulem pontos no programa de fidelidade, além de mostrar o tempo estimado de preparação e entrega do pedido (Starbucks Brasil, 2023).
- A Disney, uma empresa de entretenimento, está investindo em uma série de iniciativas com o objetivo de tornar as filas de espera em seus

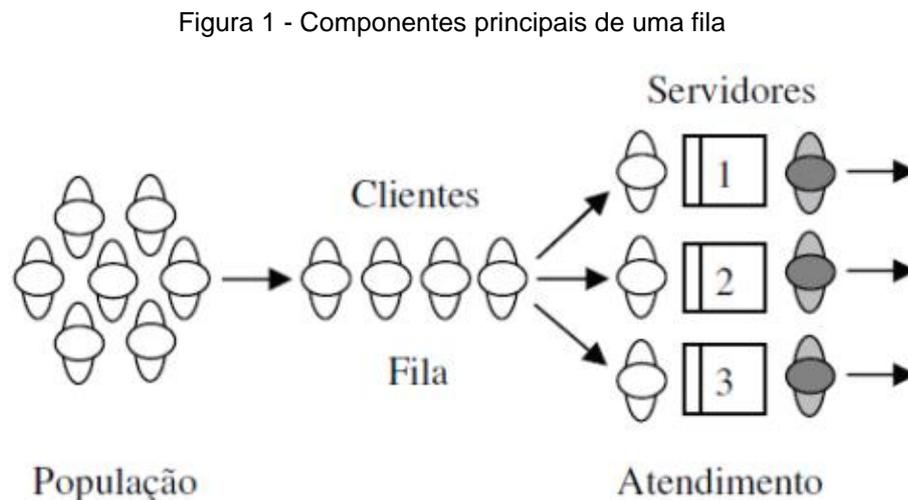
parques temáticos mais interativas e divertidas. Eles reduzem a percepção do tempo de espera e atraem os visitantes com conteúdo como música, cenários, personagens, jogos, vídeos e até mesmo cheiros (Natália Grandi, 2017).

As filas precisam ser melhoradas para satisfazer os clientes sem prejudicar os prestadores de serviço e ainda gerar lucro. As empresas podem tomar medidas com o objetivo de diminuir os efeitos prejudiciais desses sentimentos, mas não mudam a duração real da espera. Elas apresentam um impacto significativo na qualidade do serviço, então os gestores devem prestar a elas atenção especial.

Segundo Fogliatti (2007), a teoria das filas é uma área do conhecimento que ajuda a estudar e analisar filas. Usando modelos matemáticos, é possível calcular o número médio de clientes no sistema, a probabilidade de ociosidade do sistema, a utilização média dos recursos, o tempo médio que cada cliente passa na fila, o tamanho médio da fila e a probabilidade de que um determinado número de clientes esteja presente no sistema. É possível obter as características específicas de uma fila usando esses parâmetros.

2.2 TIPOS DE FILAS

Segundo Camelo et al. (2010), o cliente, a fila e o canal de atendimento são os componentes principais de uma fila.

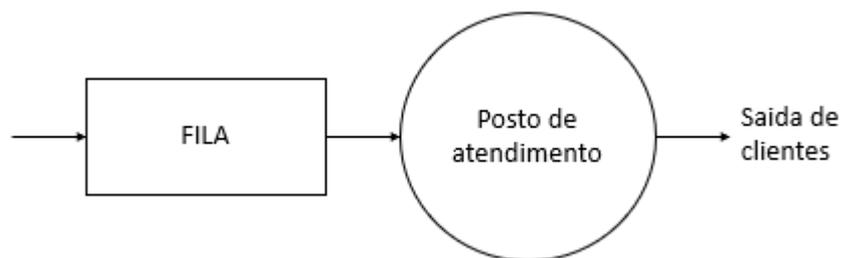


Fonte: Camelo et al. (2010).

O modelo de atendimento do sistema determina a forma das filas. Isso pode incluir:

Para Moreira (2007), uma fila é chamada de canal único quando tem apenas uma instalação de atendimento. Essa instalação pode ser composta por vários postos em série que fornecem uma parte do atendimento ou por um único posto que fornece todo o atendimento. Situações como essa ocorrem em um teatro onde apenas um guichê vende ingressos. Alternativamente, em um estacionamento com apenas um manobrista, como mostrado na figura 2.

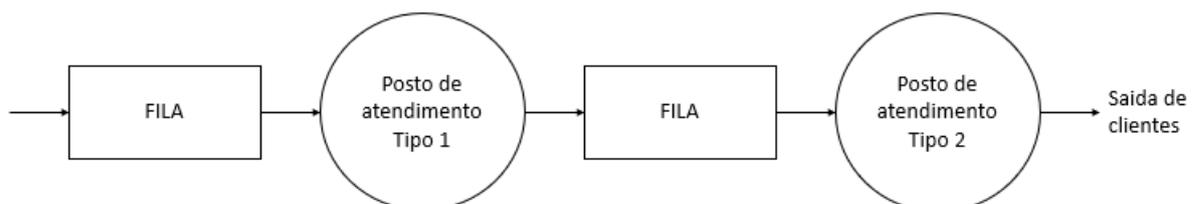
Figura 2 - Sistema Canal único, Atendimento único



Fonte: Moreira (2007).

Segundo Moreira (2007), canal único, atendimento múltiplo é a segunda situação de fila, onde a fila continua sequencialmente porque há apenas uma sequência de postos de atendimento em série. Um exemplo dessa situação é encontrado em restaurantes self-service, onde os clientes recebem bandejas de comida com diferentes tipos de comida, cada uma das quais serve como um ponto de atendimento, como mostrado na figura 3.

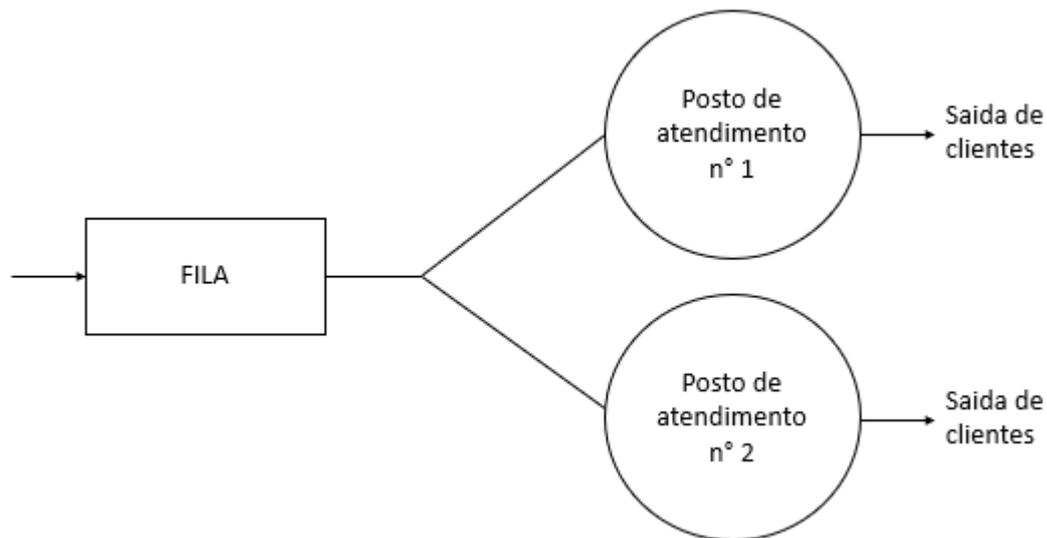
Figura 3 - Sistema Canal único, Atendimento múltiplo



Fonte: Moreira (2007).

De acordo com Moreira (2007), uma fila de canal múltiplo é criada quando duas ou mais instalações de atendimento são colocadas em paralelo e cada uma opera de forma independente das demais. Cada instalação pode ter um posto isolado ou vários postos conectados juntos. Situações de filas com canal múltiplo, atendimento único, ocorrem quando a fila é única, mas o cliente tem a opção de ir para um posto vago. Todos os postos oferecem o mesmo tipo de serviço. Um exemplo dessas circunstâncias são as que ocorrem nos caixas de bancos e nas lojas, como mostrado na figura 4.

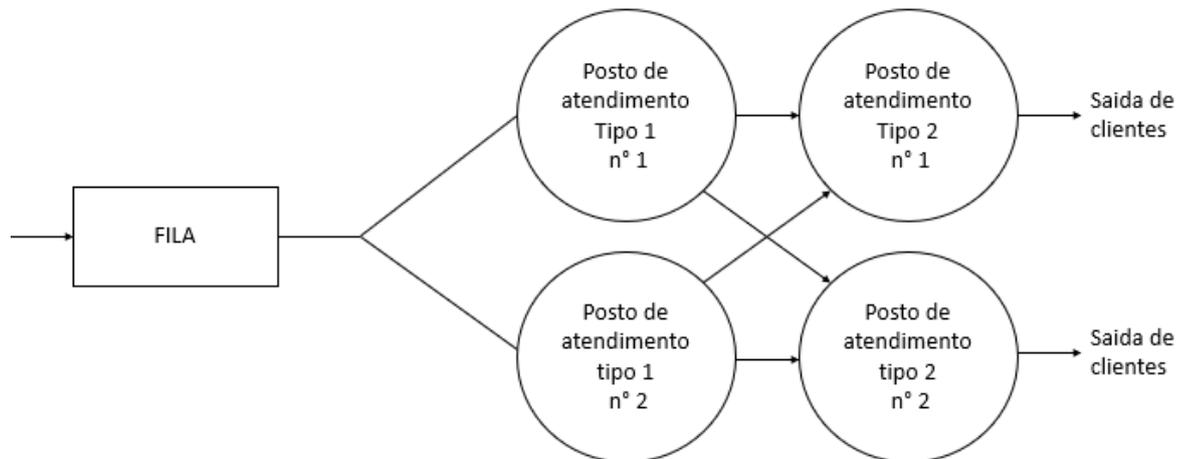
Figura 4 - Sistema Canal múltiplo, atendimento único



Fonte: Moreira (2007).

Também existe a fila que apresenta a característica de canal múltiplo, atendimento múltiplo. Nesse caso, há mais de um estágio de atendimento e o cliente é atendido por um posto vago em cada estágio. Em cada estágio, há vários postos de atendimento que desempenham a mesma tarefa. Como exemplo, suponha que se possa usar o leitor do consultório pediátrico ambulatório, mas admitindo que as etapas iniciais no consultório possam ser realizadas por duas enfermeiras de forma independente, ou seja, o cliente é atendido por quem estiver disponível no momento; de forma semelhante, uma consulta pode ser realizada por dois médicos, dirigindo-se ao cliente para quem estiver disponível no momento, como mostrado na figura 5.

Figura 5 - Sistema Canal múltiplo, atendimento múltiplo



Fonte: Moreira (2007).

Cada tipo de fila tem suas próprias vantagens e desvantagens devido às várias possibilidades do processo de chegada do cliente, do serviço, da disciplina e da configuração da fila. A teoria das filas é uma abordagem matemática que pode ajudar na melhoria do processo de seleção e ajuste desses fatores em problemas de filas de espera.

2.3 TEORIA DAS FILAS

A gestão de filas nos sistemas apresentados nas figuras 1-5 pode melhorar o serviço ao cliente e evitar congestionamento, erros e atrasos nas entregas. A teoria das filas é um dos vários métodos discutidos na literatura que pode auxiliar no alcance deste objetivo. Essa teoria surgiu em 1909, quando o engenheiro dinamarquês Agner Krarup Erlang teve que resolver um problema comum com a quantidade de circuitos necessários para fornecer um atendimento adequado para a *Copenhagen Telephone Company*. (FOGLIATTI, 2007). Erlang foi capaz de entender que a matemática resolveria outro problema porque era racional: quantos operadores de telefone são necessários para atender um certo número de chamadas. É importante destacar que durante esse período, a maioria das centrais telefônicas empregava funcionários para atender às chamadas telefônicas e conectar os fios telefônicos às tomadas elétricas das placas com circuitos. A partir da Segunda Guerra Mundial, a teoria das filas foi aplicada em outros problemas de filas.

As áreas de economia, administração e processamento de fluxos têm usado essa abordagem desde então. Ela resolveu vários problemas, como congestionamento de tráfego; escoamento de fluxo de carga em terminais; carregamento e descarregamento de veículos; comunicação de computadores; formação de estoques; e escoamento de fluxo de processamento de informações (FOGLIATTI; COSTA MATTOS, 2007). O objetivo principal da teoria das filas é criar modelos matemáticos capazes de prever o comportamento de sistemas de prestação de serviços. Em geral, é admitido que as solicitações ocorrem em momentos aleatórios e que o tempo de atendimento para cada pedido também é uma variável aleatória (MARINS, 2011).

A teoria das filas é um ramo da probabilidade que estuda o fenômeno da formação de filas usando análises matemáticas com variáveis mensuráveis (Guedes e Araujo, 2013). Modelos podem ser usados para analisar o comportamento de um sistema que fornece serviços cuja demanda é aleatória, o que permite a identificação de desperdícios e gargalos, tornando viáveis propostas de melhorias para atender às necessidades dos clientes e tornar o sistema economicamente viável.

Carrion (2007) afirma que a teoria das filas visa melhorar o desempenho de um sistema reduzindo os custos operacionais. Para chegar ao objetivo de otimizar o comportamento dos modelos de filas de espera, é necessário analisar os resultados produzidos pelas fórmulas adequadas para cada modelo. Estes resultados devem permitir a análise de uma situação específica. Isso pode ser feito manualmente em vez de usar os dados de entrada nas fórmulas, ou podem ser obtidos usando um programa de computador.

Para Taha (2008), a teoria da fila não é uma ferramenta de otimização. Em vez disso, ela determina a situação da fila usando medidas de desempenho (taxa de chegada, taxa de serviço, tempo de espera, taxa de utilização, entre outras) visando obter melhorias para o sistema em questão. As filas de espera causam problemas para todas as empresas que trabalham com processamento de fluxos e têm limitações de atendimento, diz Corrêa (2005). E afirma que todas as organizações que processam fluxos e enfrentam limitações de recursos enfrentam o problema de filas de espera.

Telecomunicações, bancos, call centers, hospitais, aeroportos, restaurantes, supermercados e outras áreas de prestação de serviços são exemplos concretos da implementação dessa abordagem, conforme afirmado por Brown et al. (2002).

Quando o sistema está em desequilíbrio, ou seja, quando a capacidade de atendimento é menor do que a demanda de serviço, é possível identificar a formação de filas. Para identificar problemas no sistema, as medidas de desempenho da teoria das filas visam dimensionar esse possível desequilíbrio.

Como afirmado por Bronson (1985), os cinco componentes que compõem os sistemas de filas são o modelo de chegada dos usuários, o modelo de serviço, o número de atendentes, a capacidade do estabelecimento de atender e a ordem em que os usuários são atendidos. Este modelo de chegadas calcula a quantidade de tempo que os clientes levam para chegar ao estabelecimento. O modelo de serviços estabelece um cronograma para fornecer serviços ao consumidor. O número de atendentes representa a disponibilidade simultânea de pessoas ou equipamentos. A capacidade do sistema, com o número máximo de usuários atendidos e na(s) fila(s), representa a capacidade do sistema. A ordem na qual os clientes são atendidos é conhecida como disciplina na fila. Assim, o estudo da teoria das filas tem como objetivo, entre outros, determinar:

- $N(t)$: o número de clientes na fila no instante t ($t \geq 0$);
- $P_n(t)$: a probabilidade de n clientes no sistema de filas estarem presentes no instante t , dado o número no instante 0;
- P_n : a probabilidade de que exatamente n clientes se encontrem no sistema de filas;
- P_0 : Probabilidade do sistema está vazio;
- K : é o número de atendentes no sistema de filas;
- λ_n : é a taxa média de chegada (número de chegadas esperado por unidade de tempo) de novos clientes quando n clientes se encontram no sistema.
- μ_n : é a taxa média de atendimento, número de clientes que estão sendo atendidos num determinado período de tempo;
- L_q : Número esperado na fila é o número de cliente que aguarda atendimento.
- L : Número esperado no sistema, é o número esperado de clientes que estejam esperando na fila ou que estejam sendo atendido.
- W_q : Tempo esperado na fila, é o tempo individual que cada cliente aguarda na fila, excluindo o tempo de atendimento.

- W : Tempo de sistema esperado é o tempo de espera no sistema incluindo o tempo de atendimento para cada cliente individual.

2.4 NOTAÇÃO DE KENDALL

As figuras 1-5 ilustram a existência de diferentes tipos de sistemas de filas, que exigirão modelos matemáticos específicos para gerenciamento. Como resultado, é necessário padronizar a identificação desses sistemas com base nos modelos matemáticos. Para descrever e comparar diferentes sistemas de filas, os pesquisadores usam a notação de Kendall como padronização. Ela é um sistema padrão para classificar e descrever um nó de filas. Em 1953, David George Kendall propôs a notação, que consiste em vários símbolos que descrevem as características do sistema de filas. (A/B/C/D/E) são as cinco partes da notação. Cada letra representa um elemento específico do sistema de filas. A notação de Kendall é de suma importância porque permite que os pesquisadores descrevam e comparem diferentes sistemas de filas, de acordo com Costa (2018). Além disso, ela ajuda na identificação de falhas no sistema de fila e na avaliação do desempenho do sistema.

Quadro 1 - Notação de Fila 1

Siglas	Explicações
A	Distribuição do tempo entre chegadas sucessivas
B	Distribuição do tempo de atendimento
C	Número de postos de atendimento em paralelo
D	Capacidade física do sistema
E	Disciplina de atendimento

Fonte: Autor (2023).

A maioria das pessoas usa as seguintes siglas para descrever as distribuições mais comuns:

Quadro 2 - Notação de Fila 2

Siglas	Explicações
D	Representa uma distribuição determinística ou degenerada
M	Representa uma distribuição exponencial
E_k	Representa uma distribuição de Erlang do tipo k

G	Representa uma distribuição geral (não específica)
-----	--

Fonte: Autor (2023).

Um sistema de filas com disciplina de atendimento FIFO (primeiro a chegar é o primeiro a sair) e capacidade ilimitada (infinita) é admitido por alguns autores que às vezes simplificam a notação de Kendall, nos casos das siglas D e E.

Com a disciplina da fila podendo ser classificada em PRI (Prioridade), FIFO (First In First Out), LIFO (Last In First Out) ou SIRO (Servidor Atende Quem Chega em Ordem Aleatória). Maiores detalhes sobre as disciplinas de atendimento podem ser encontrados em Taha (2008).

2.4.1 Modelo M/M/1

Um dos modelos de teoria das filas mais comumente utilizados é o modelo M/M/1, que mostra um sistema com um servidor e as chegadas dos clientes e os tempos de serviço seguem uma distribuição exponencial com taxa λ uma fila única que usa o FIFO e uma estação de serviço.

Este modelo consiste em um processo estocástico cujo espaço de estados é o conjunto $X = \{1,2,3 \dots n\}$ com n estados ou elementos em que o valor corresponde ao número de clientes no sistema, incluindo aqueles que estão sendo servidos. Os estados mudam a taxas constantes λ e μ .

Algumas equações básicas que podem ser aplicadas na resolução de problemas desse tipo, podem ser representadas por:

A probabilidade de um sistema estar ocupado ou da utilização de um sistema é representado por:

$$\rho = \frac{\lambda}{\mu} \quad (1)$$

Onde: ρ = Representa a utilização do sistema, λ = Representa a taxa de chegada por unidades em um determinado período de tempo, μ = Representa a taxa de atendimento por unidade em um determinado período de tempo.

Quando $\lambda/\mu < 1$, as equações a seguir serão válidas, portanto, temos:

A probabilidade de o sistema estar vazio, ou seja, com zero unidades no sistema é dada pela equação (2):

$$P_0 = 1 - \frac{\lambda}{\mu} \quad (2)$$

O número esperado de clientes na fila, pode ser obtido por (3):

$$L_q = \frac{\lambda^2}{\mu(\mu + \lambda)} \quad (3)$$

O número previsto de clientes esperando no sistema é dado por (4):

$$L = \frac{\lambda}{\mu - \lambda} \quad (4)$$

O tempo de espera na fila é oferecido pela equação (5):

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} \quad (5)$$

O tempo de espera no sistema, predisposto por (6):

$$W = \frac{\lambda}{\mu - \lambda} \quad (6)$$

A probabilidade de existirem n unidades no sistema é ofertado pela equação (7):

$$P_n = P_0 \left(\frac{\lambda}{\mu}\right)^n \quad (7)$$

A probabilidade de existirem mais de k unidades no sistema, dado por (8):

$$P(n > k) = \left(\frac{\lambda}{\mu}\right)^{k+1} \quad (8)$$

A probabilidade de haver n clientes no sistema é oferecido pela equação (9):

$$P(n) = \left(\frac{\lambda}{\mu}\right)^n \left(\frac{\mu - \lambda}{\mu}\right) \quad (9)$$

A probabilidade de 1 unidade demorar mais de t unidades de tempo no sistema é dado por (10):

$$P(T > t) = e^{-\mu(1-\rho)t} \quad (10)$$

2.4.2 Modelo M/E_k/1

A distribuição de Erlang é um outro tipo de distribuição de tempo de atendimento. Foi chamado de Erlang Distribution em homenagem a Agner Krarup Erlang, o criador da teoria das filas. A função probabilística de densidade do Erlang é:

$$f(t) = \frac{(\mu k)^k}{(k-1)!} t^{(k-1)} e^{-k\mu t}, \text{ para } t \geq 0 \quad (11)$$

Na distribuição Erlang, os parâmetros μ e k são usados, μ é a taxa média de atendimento e k é o parâmetro que determina o grau de variabilidade dos tempos de atendimento em relação à média. A distribuição de Erlang pode ser considerada uma família de distribuições porque existe uma distribuição para cada k . O parâmetro de distribuição k deve ser inteiro, além de que ambos tanto k quanto μ devam ser positivos.

Nesse caso, pode-se ver que a distribuição exponencial é um exemplo único da distribuição de Erlang, com $K = 1$ e $K = \infty$, com os valores k fornecendo a média $= 1/\mu$, variância $= 1/(k\mu^2)$ e moda $= (k-1)/(k\mu)$. O valor inteiro de k que se aproxime das estimativas pode ser encontrado usando as estimativas da média e variância de uma distribuição empírica de tempos de atendimento.

Para este modelo temos as seguintes equações:

O número de clientes aguardando na fila é dado por (12):

$$L_q = \frac{(1+k)\lambda^2}{2k\mu(\mu-\lambda)} \quad (12)$$

O tempo de espera na fila é oferecido pela equação (13):

$$W_q = \frac{(1+k)\lambda}{2k\mu(\mu-\lambda)} \quad (13)$$

O Tempo de espera no sistema é ofertado pela equação (14):

$$W = W_q + \frac{1}{\mu} \quad (14)$$

O Número de clientes aguardando no sistema é dada por (15):

$$L = \lambda W \quad (15)$$

2.4.3 Modelo M/M/s

Um sistema de filas com uma única fila e vários servidores idênticos pode ser descrito pelo modelo M/M/s de teoria das filas. Nesse sistema, as chegadas dos clientes seguem um processo de Poisson e os tempos de atendimento têm uma distribuição exponencial. O número médio de clientes na fila, o tempo médio de espera na fila, a probabilidade de um cliente esperar na fila e outros fatores são medidos com esse modelo para avaliar o desempenho dos sistemas de filas. Em 1917, o matemático dinamarquês Agner Krarup Erlang propôs o modelo M/M/s como uma generalização do modelo M/M/1, que considerava apenas um servidor.

As equações para o modelo M/M/s são as seguintes:

A taxa de utilização do sistema é dada por (16):

$$\rho = \frac{\lambda}{s\mu} \quad (16)$$

A probabilidade de que haja n clientes no sistema ($n \geq 0$) é dada por (17):

$$P_n = \begin{cases} \frac{(\lambda/\mu)^n}{n!} P_0, n < s \\ \frac{(\lambda/\mu)^n}{s! s^{n-s}} P_0, n \geq s \end{cases} \quad (17)$$

A probabilidade de que um cliente que chega ao sistema tenha que esperar na fila é dada por (18):

$$P_w = P_s = \frac{(\lambda/\mu)^s}{s! (1 - \rho)} P_0 \quad (18)$$

O número médio de clientes na fila é dado pela equação (19):

$$L_q = P_w \frac{\rho}{1 - \rho} = \frac{(\lambda/\mu)^s}{s! (1 - \rho)^2} P_0 \quad (19)$$

O número médio de clientes no sistema é oferecido pela equação (20):

$$L_s = L_q + \frac{\lambda}{\mu} = L_q + s\rho \quad (20)$$

O tempo médio de espera na fila é ofertado pela equação (21):

$$W_q = \frac{L_q}{\lambda} = P_w \frac{1}{\mu(1-\rho)} \quad (21)$$

O tempo médio de espera no sistema é dado por (22):

$$W_s = W_q + \frac{1}{\mu} = W_q + \frac{s}{s\mu - \lambda} \quad (21)$$

2.4.4 Modelo M/D/1

O modelo M/D/1 é um modelo de fila de espera com tempos de serviço constantes (determinísticos) e um único servidor. Geralmente é usado em sistemas com um único prestador de serviços e tempos de serviço constantes. Por exemplo, em um comutador de pacotes de comprimento fixo.

Segundo Hillier (2013), o modelo M/D/1 é o caso único do modelo M/G/1, onde $\sigma^2 = 0$, deste modo a equação de Pollaczek-Khintchine se reduz a equação (29):

$$L_q = \frac{\rho^2}{2(1-\rho)} \quad (22)$$

Conforme as características do sistema em estudo, existem modelos de filas adicionais que podem ser usados. Esses modelos podem ser verificados em Hillier (2013).

2.5 SÍNTESE DO ESTADO DA ARTE

Neste tópico, mostraremos como usar a teoria das filas para resolver problemas em diferentes situações. O foco do trabalho está em serviços de alinhamento e

balanceamento, mas há poucas pesquisas sobre esse tema na área. Então, vamos explorar como a teoria das filas pode ser aplicada em outros campos e quais são os benefícios e limitações que os pesquisadores descobriram.

Mendes (2019) analisou o serviço de entrega de uma pizzaria em Sumé-PB. As distribuições de probabilidade foram calculadas, os tempos entre a entrada de pedidos no sistema e o tempo de atendimento foram coletados. Em seguida, foram analisados três cenários: o primeiro estava relacionado à situação atual, o segundo sugeriu a contratação de um motoboy e o terceiro analisou como reduzir os erros no processo. Os resultados mostraram que com o auxílio da teoria das filas e outras ferramentas o cenário três é o mais adequado para a organização, o que significa que os processos de entrega podem ser melhorados. Para atingir este último objetivo, foram utilizadas ferramentas de qualidade para identificar as causas, priorizar os problemas e estabelecer um plano de ação.

Domingos (2019) examinou uma variedade de modelos de filas de espera e apresentou alguns dos seus usos nos serviços de saúde, que geralmente enfrentam atrasos no atendimento de pacientes. Foram explorados alguns artigos em revistas internacionais como o *Journal of Medical Systems*, o *European Journal of Operational Research* e o *Journal of Health Care Management Sciences*. O trabalho é dividido em três partes principais. A primeira aborda os fundamentos da teoria das probabilidades e dos processos estocásticos, com ênfase em cadeias de Markov. A segunda aborda modelos de filas de espera, especialmente os markovianos, e discute questões relevantes de ponto de vista das aplicações, como o número médio de clientes no sistema (ou na fila), o tempo médio que um cliente leva para entrar no sistema (ou na fila). Por fim, a revisão de artigos publicados em revistas internacionais como o *Journal of Medical Systems*, o *European Journal of Operational Research* e as ciências de gestão de saúde apresenta aplicações práticas de filas de espera nos serviços de saúde. Para os resultados encontrados foi que os modelos de filas de espera podem ser usados para analisar e otimizar vários aspectos dos serviços de saúde, incluindo dimensionamento de capacidade, alocação de recursos, programação de consultas e gestão da demanda.

Toporowicz et al. (2018) realizou um estudo onde aplicou a teoria das filas em um restaurante universitário localizado em Ponta Grossa, PR. Este restaurante tem muitas reclamações devido ao tamanho das filas que são formadas durante o horário de almoço. Uma fila de alunos faz o pagamento e vai para as duas mesas de self-

service no sistema do restaurante atual. A análise quantitativa do sistema permitiu a proposta de uma sugestão de melhoria que visa determinar o ritmo de reposição da mistura principal nas cubas de comida, onde o gargalo do sistema está localizado. O tamanho da fila e o tempo de espera foram reduzidos em até 95% no período estudado com os cálculos adequados e considerando o tempo médio de reposição da mistura nos pontos de atendimento e a capacidade das cubas de comida. Esses resultados mostram que a teoria das filas é eficaz em estudos de atendimento no setor de serviços e pode aumentar significativamente a satisfação dos usuários.

Rossana (2009) trouxe um estudo com o objetivo de investigar como os sistemas de filas de espera se comportam ao longo do tempo e quando estão em equilíbrio. Sendo dividido em três capítulos principais, no primeiro capítulo, conceitos básicos de probabilidade, estatística e processos estocásticos são apresentados, além das condições e características necessárias para formar um sistema de filas de espera. Vários tipos de sistemas de filas de espera markovianos, ou seja, com propriedades de memória nula, são desenvolvidos no segundo capítulo. O número esperado de clientes no sistema e na fila, bem como o tempo esperado para um cliente permanecer no sistema e na fila após o equilíbrio do sistema, estão entre as muitas características de cada modelo que são examinadas. Além disso, são mostrados gráficos e comparações entre os modelos. No terceiro capítulo, são discutidos sistemas de filas de espera que não são markovianos, ou seja, que possuem dependência temporal. Embora a pesquisa seja menos aprofundada, ela sempre busca identificar as características que foram identificadas nos modelos markovianos. Os resultados da dissertação mostram que os sistemas de filas de espera são úteis para modelar e analisar situações reais que envolvem chegadas e serviços aleatórios, como atendimento ao cliente, tráfego rodoviário, redes de computadores, etc. As fórmulas de Little, que relacionam o tempo médio e o número médio de clientes no sistema, e os processos de nascimento e morte, que descrevem a evolução do número de clientes no sistema, também são aplicadas na dissertação.

Silva et al. (2014) apresentou um artigo buscando avaliar o índice de qualidade na prestação de serviços em uma filial de uma grande rede varejista usando os componentes da teoria das filas. O artigo é descritivo-exploratório e usa uma abordagem quantitativa. Durante 15 dias, os autores coletaram e trataram informações sobre o tempo de espera e atendimento dos clientes em vários caixas da loja. Eles calcularam métricas de desempenho do sistema como taxa média de

chegada, taxa média de atendimento, número médio de clientes na fila, tempo médio de espera na fila, número médio de clientes no sistema e tempo médio no sistema usando a teoria das filas. Os resultados mostraram que a empresa estudada apresentou resultados satisfatórios em relação à sua condição de serviço de automóveis, pois os clientes esperaram em média menos de dois minutos na fila e receberam assistência em menos de três minutos. Além disso, os índices coletados permitiram a reconstrução da situação da organização, que pode ser usada para comparar com o objetivo da organização. Com isso, a conclusão de que a teoria das filas pode ser usada como um indicador de qualidade de serviços, pois permite quantificar os elementos intangíveis da satisfação do cliente e fornece informações úteis para a tomada de decisão e a gestão do sistema. Eles sugeriram que estudos futuros podem examinar outras partes da rede varejista e tipos de serviços adicionais.

Coelho (2016) aplicou a teoria das filas em busca de analisar o desempenho de uma unidade bancária na cidade de João Pessoa/PB e encontrar problemas e oportunidades de melhoria no atendimento ao cliente usando ideias da teoria das filas e Simulação. Durante um período de trinta dias, foram coletados e tratados dados sobre o número de chegadas e saídas dos clientes, tempos de espera e tempo de serviço em vários guichês bancários. Além disso, pode-se encontrar medidas de desempenho do sistema como a taxa média de chegada, a taxa média de serviço, o número médio de clientes na fila, o tempo médio de espera na fila, o número médio de clientes no sistema e o tempo médio no sistema usando os modelos matemáticos da teoria das filas. Os resultados mostraram que a capacidade de atendimento do banco estudado era insatisfatória; os clientes esperaram em média mais de dez minutos na fila e foram atendidos em mais de cinco minutos. Além disso, os índices mostraram que o fluxo de clientes e o tempo de serviço variam muito, o que resulta em filas longas e demoradas. Além disso, o autor realizou uma simulação computacional para avaliar uma variedade de cenários de mudança na estrutura do sistema. Esses cenários incluíam mudanças no tipo de serviço, na distribuição dos clientes e no número de guichês. Os resultados da simulação mostraram que a implementação de um sistema de triagem de clientes baseado em tipo de serviço, com guichês específicos para cada tipo, seria a melhor maneira de melhorar o desempenho do sistema. Essa mudança reduziria drasticamente o tempo de espera e o tempo gasto no sistema dos clientes, ao mesmo tempo em que melhoraria a eficiência e a qualidade do atendimento. Portanto se conclui que a teoria das filas e a

Simulação podem ser usadas como ferramentas estratégicas na gestão de operações em serviços, pois permitem avaliar o comportamento do sistema, identificar problemas e sugerir soluções.

Fernando (2010) buscou analisar o fluxo de atendimento e o número de atendentes em um supermercado usando as ideias da teoria das filas para maximizar o tempo de espera dos clientes e reduzir os custos operacionais. Durante quinze dias, dados foram coletados e tratados sobre o número de chegadas e saídas dos clientes, tempo de espera e serviço em vários caixas do supermercado. Foram calculadas várias medidas de desempenho do sistema, como a taxa média de chegada, a taxa média de serviço, o número médio de clientes na fila, o tempo médio de espera na fila, o número médio de clientes no sistema e o tempo médio no sistema, utilizando modelos matemáticos da teoria das filas. Os resultados mostraram que o supermercado estudado apresentou resultados satisfatórios em termos de sua capacidade de atendimento, pois os clientes esperaram em média menos de três minutos na fila e foram atendidos em menos de quatro minutos. Além disso, as filas são curtas e rápidas devido à baixa variabilidade no fluxo de clientes e no tempo de serviço. Além disso, o autor realizou uma análise econômica com o objetivo de estimar os custos e os benefícios de aumentar ou diminuir o número de caixas em operação. A análise econômica revelou que, sem comprometer significativamente o tempo de espera dos clientes, o supermercado poderia economizar cerca de R\$ 10.000,00 por mês se reduzisse o número de caixas em operação de 10 para 8. Por outro lado, se o supermercado aumentasse o número de caixas em operação para doze, aumentaria os custos do supermercado em cerca de R\$ 15.000,00 por mês, sem melhorar significativamente o tempo de espera dos clientes. A conclusão é que a teoria das filas pode ser usada como uma ferramenta gerencial para a gestão do fluxo de atendimento em supermercados. Isso se deve ao fato de que permite avaliar o desempenho do sistema, identificar pontos problemáticos e sugerir soluções. Ele propôs que futuros estudos possam examinar outros tipos de lojas e outros fatores que impactam a satisfação dos clientes.

Com base nos estudos apresentados, é possível observar os benefícios da aplicação da teoria das filas em diversas áreas. Isso permite, entre outros aspectos, coletar informações sobre o desempenho do sistema, no que diz respeito ao número de clientes na fila, à capacidade de atendimento, ao custo de operação, entre outros,

de modo a subsidiar o desenvolvimento de planos para a melhoria do processo de filas.

3. METODOLOGIA

3.1 TIPO DE PESQUISA

Quanto à abordagem, esta pesquisa é caracterizada como pesquisa quantitativa, visto que os dados coletados e posteriormente analisados são numéricos. Segundo Jacobsen (2009), a pesquisa quantitativa preocupa-se com a quantificação dos dados coletados e utiliza-se de técnicas estatísticas para a medição dos dados.

No que tange aos objetivos, esta pesquisa é considerada exploratória, pois conforme Gil (2002), o objetivo é aumentar a familiaridade com o ambiente ou objeto de estudo e torná-lo mais explícito para que as hipóteses possam ser construídas.

Referente aos procedimentos técnicos é possível classificar este estudo como:

- Pesquisa bibliográfica: este tipo de pesquisa perpassa todos os momentos do trabalho acadêmico e é utilizado em todas as pesquisas. Consiste na seleção e análise de materiais que sejam acessíveis ao público e que possam fornecer dados relevantes e atuais sobre o tema a ser pesquisado (VERGARA, 2010). Para fazer a fundamentação teórica desta pesquisa de estágio, foram utilizados livros e artigos que abordam os temas concernentes às questões norteadoras propostas sobre a teoria das filas, que é um ramo da probabilidade que estuda a formação de filas, através de análises matemáticas precisas e propriedades mensuráveis das filas.
- Documental: pois utiliza relatórios e documentos arquivados da empresa em estudo, visando coletar informações que auxiliem no entendimento da situação atual das atividades de atendimento dos clientes.
- Estudo de caso: Segundo Gil (2002), os estudos de casos são utilizados tanto em pesquisas exploratórias quanto descritivas e explicativas, com o objetivo de estudar profundamente e minuciosamente um ou poucos objetos de um assunto para permitir a obtenção de um conhecimento abrangente e detalhado.

3.2 COLETA DE DADOS

Diz respeito à descrição das técnicas utilizadas para a coleta de dados (primários e secundários). Como técnicas, foram realizados acompanhamentos das chegadas de clientes, onde essas chegadas eram arquivadas em um arquivo Excel. O período de coleta dos dados foi de novembro de 2022 a janeiro de 2023. A coleta foi realizada durante oito semanas, de segunda a sexta, com o horário de coleta de oito horas por dia, das 08:00 às 12:00 e das 13:00 às 17:00. Durante esse período, foi possível concluir a coleta de 40 dias, calculando o número de clientes atendidos por dia.

3.3 ANÁLISE DOS DADOS

Nesta pesquisa, os dados serão analisados de forma quantitativa. Knechtel (2014) afirma que a análise quantitativa atua sobre um assunto baseado em uma teoria e possui a capacidade de quantificar variáveis relacionadas a um objeto de estudo específico. Em outras palavras, a análise quantitativa busca estatisticamente quantificar as informações e opiniões obtidas.

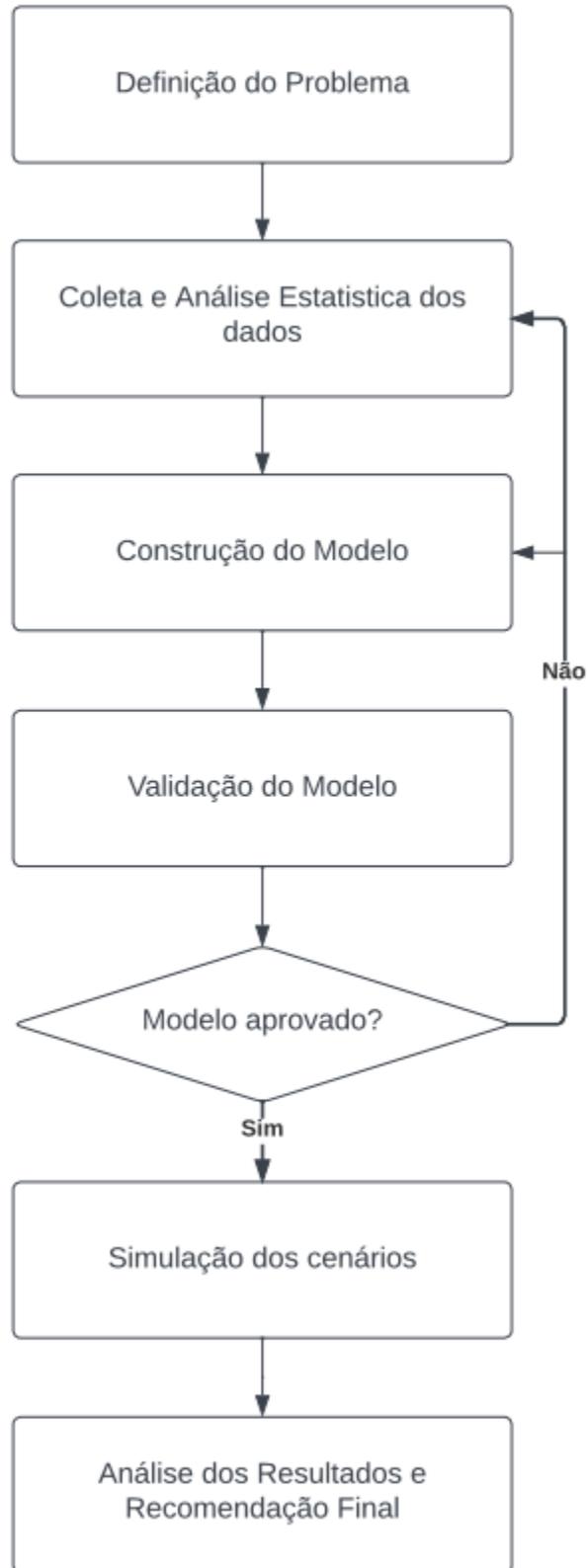
A análise dos dados ocorrerá à luz dos conceitos do referencial teórico, visando determinar possíveis pontos de divergência, nos quais serão fundamentadas as sugestões de melhoria, atendendo, deste modo, ao objetivo geral desta pesquisa. Além disso, este trabalho fará uso da estatística descritiva, para determinar o tamanho da amostra e obter uma melhor compreensão dos dados coletados. Em seguida, serão realizados os testes de aderência para determinar o tipo de distribuição de probabilidade que as amostras coletadas seguem. Isso será feito por meio do teste de Kolmogorov-Smirnov, com um nível de significância alfa (α) = 0.05. Para tanto, será necessário o uso da linguagem Python v.3.10.12.

No que diz respeito à validação do modelo base da teoria das filas, este trabalho realizará uma comparação qualitativa entre a saída do modelo e os dados observados no sistema real.

Os códigos implementados via Python são apresentados nos apêndices C e D.

3.4 MÉTODO DE TRABALHO

Figura 6 - Etapas do Processo



Fonte: Autor (2023).

- Identificar o problema e delinear os objetivos do estudo foram os primeiros passos. Assim, foi realizado um mapeamento dos processos para coletar informações que ajudariam a entender e resolver os problemas.
- Em seguida, os dados em campo foram coletados de novembro de 2022 a janeiro de 2023. Essas informações são usadas para explicar os processos, definir as distribuições de probabilidade para as variáveis consideradas neste estudo.
- Na próxima etapa, o modelo de equações de filas adequado para as características do sistema alvo de estudo é selecionado e aplicado aos dados coletados.
- A fim de validar o modelo criado na etapa anterior, o quarto passo faz uma comparação qualitativa entre as saídas obtidas e os dados do sistema real. Se não for aprovado, várias etapas devem ser tomadas, como verificar a existência de erros de lógica no modelo e aumentar a amostra de dados.
- Após a aprovação do modelo, simulações são realizadas considerando diferentes cenários. O último passo fornece uma recomendação final sobre o assunto de pesquisa e os pressupostos que sustentam as recomendações.

4 ESTUDO DE CASO

4.1 DESCRIÇÃO DA EMPRESA

A empresa estudada é uma indústria que comercializa pneus novos, pneus de carga, realiza recapagem de pneus, bem como os serviços do *Truck Center*. Sua sede principal, e a que será abordada no estudo, está localizada na zona metropolitana do Recife-PE. Além disso, hoje ela está atuando em mais cinco estados (Paraíba, Rio Grande do Norte, Ceará, Piauí e Maranhão), atuante no mercado há mais de 22 anos. Outros tipos de produtos, como rodas, câmaras de ar e protetores, também podem ser encontrados, além de serviços de alinhamento e balanceamento.

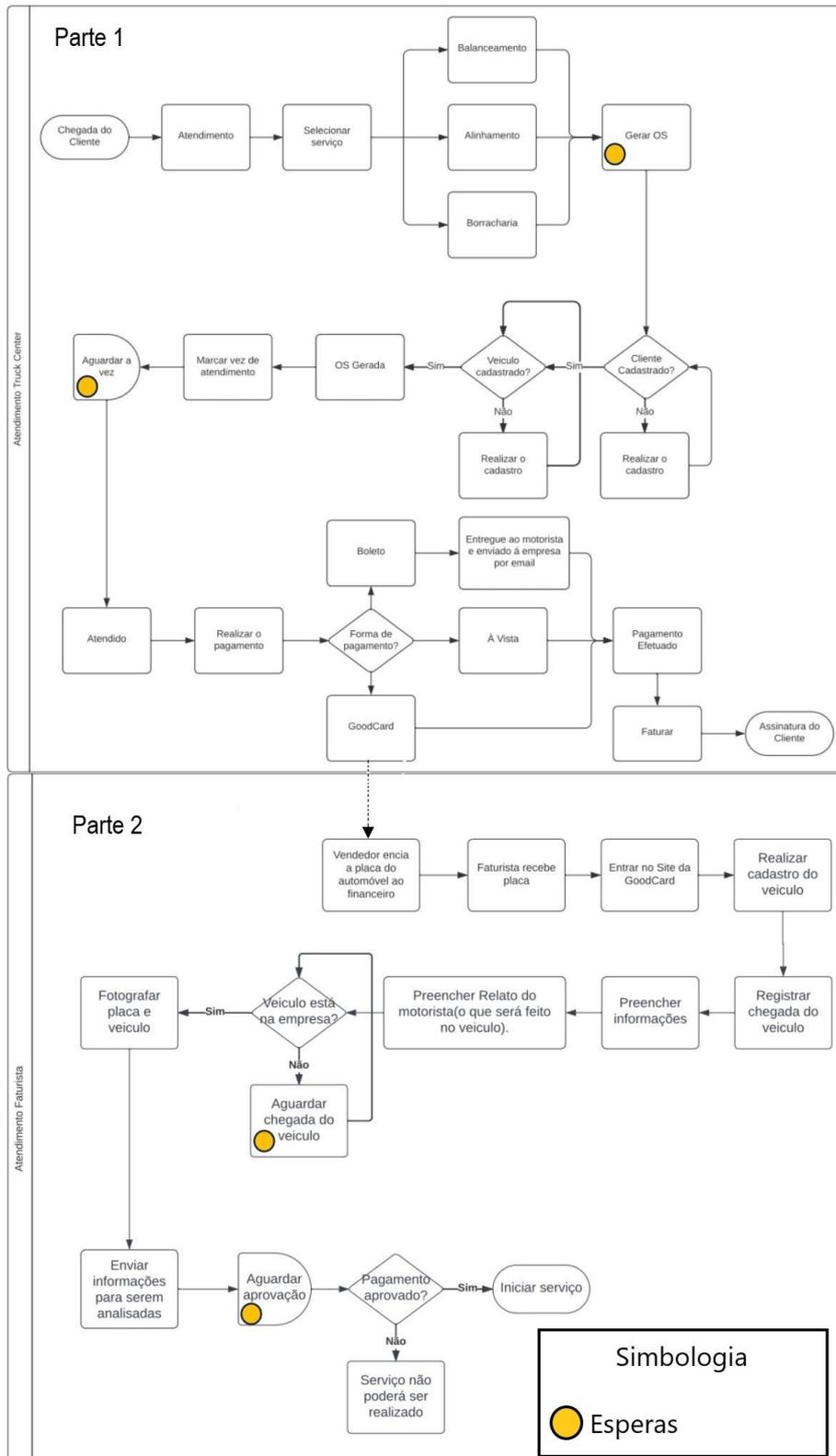
Além do mais, também é oferecido ao mercado de transporte uma variedade de ferramentas de gestão, garantindo o mais moderno e eficiente sistema de soluções integradas buscando aumentar o rendimento por quilômetro dos pneus. A cada etapa do processo de produção, todos os pneus são examinados e seus dados são registrados no sistema. Isso garante a rastreabilidade perfeita e dá ao transportador alguns benefícios, como: controle completo de todas as atividades de movimentação e vida útil dos pneus; identificação dos fatores que contribuem para uma baixa performance; e diagnóstico que ajuda na tomada de decisões.

A empresa também presta serviços de alinhamento e balanceamento de veículos pesados, como caminhões e ônibus, que são procedimentos essenciais para garantir a segurança, o conforto e a economia na condução desses veículos. Ajudam a evitar desgastes irregulares nos pneus, vibrações na direção e desvios na trajetória, ajustando o ângulo e a distribuição de peso das rodas. Com uma equipe qualificada e equipamentos modernos, a empresa atende entre 20 e 25 clientes por dia, oferecendo um serviço de qualidade e confiança.

4.2 SITUAÇÃO ATUAL DO SISTEMA

Para os processos de chegada e atendimento dos clientes existem algumas etapas no processo. Essas etapas são apresentadas na figura 7.

Figura 7 - Etapas do Processo de Atendimento



Fonte: Autor (2023).

A primeira parte do serviço (Parte 1 do fluxograma da Fig. 7) começa quando o cliente chega à empresa. O cliente é atendido pelo vendedor e recebe todas as informações pertinentes aos tipos de serviços prestados na empresa e seus respectivos valores. Após escolher um serviço, o cliente precisa criar uma ordem de serviço (OS) para que ele seja iniciado. A OS contém as informações sobre o veículo, o cliente e quais serviços serão realizados. Ela será repassada para o alinhador, o funcionário que realiza os alinhamentos e balanceamentos dos veículos, para que ele possa saber quais os serviços a serem realizados. A OS só poderá ser gerada caso o cliente e seu veículo estejam cadastrados na empresa; caso contrário, será necessário realizar esse cadastro. Após a criação da OS, o cliente deverá aguardar a vez de seu veículo ser atendido pelo alinhador.

Ao ser atendido, deverá efetuar o pagamento do serviço, escolhendo entre três formas de pagamento.

- À vista: o cliente paga o serviço com dinheiro vivo ou cartão.
- Boleto: forma de pagamento disponível apenas para clientes antigos, em que o boleto será enviado para o e-mail do motorista do veículo e para a empresa dona do veículo que irá efetuar o pagamento.
- GoodCard: é uma bandeira de cartão que oferece mais opções de pagamento para o consumo e tem um portal com várias funcionalidades, como o extrato do seu cartão, o envio de saldo por e-mail, além de uma rede credenciada com mapas de localização e alguns descontos nos estabelecimentos.

O estudo em questão focará apenas nos processos de serviços quando os clientes selecionam a forma de pagamento *GoodCard* (Parte 2 do fluxograma Fig. 7). Esse tipo de pagamento é o mais problemático para a empresa, pois há muitos clientes que não quitam os serviços contratados. Ao selecionar esse tipo de pagamento, a empresa deve enviar as informações do cliente e seu veículo para o site da *GoodCard*, uma vez que, para efetuar o serviço solicitado pelo cliente, a empresa do cartão deverá liberar esse pagamento. Para realizar esse processo, o faturista deve entrar no site do cartão e realizar o cadastro do veículo e sua chegada. Em seguida, deverá ser preenchido um formulário contendo algumas informações, sendo elas: a quilometragem atual do veículo, sua placa, nome e telefone do motorista, além de preencher com o relato do motorista que diz respeito aos serviços que serão realizados no veículo. Além disso, também é necessário enviar uma foto do

veículo e de sua placa. Essas fotos devem mostrar de forma clara que o veículo se encontra na empresa prestadora do serviço de alinhamento e balanceamento. Caso ainda não esteja na empresa, existe a necessidade de aguardar sua chegada. Algumas vezes, acontece de o serviço ser realizado antes da liberação do cartão e, no processo da fotografia do veículo, o cliente já ter ido embora. Isso acarreta que o pagamento não seja aprovado para um serviço já realizado, trazendo prejuízos para a empresa.

Após todas as etapas concluídas, será necessário aguardar um retorno da empresa do cartão *GoodCard*, confirmando ou rejeitando o pagamento. Em casos de rejeição do pagamento, o serviço não pode ser iniciado; logo, o cliente deverá escolher outra forma de pagamento. Para os casos de confirmação, o serviço já pode ser iniciado no veículo. Após o serviço ser realizado, o cliente deverá assinar um documento confirmando que o serviço foi realizado pela empresa.

4.3 DADOS COLETADOS E DEFINIÇÃO DO MODELO DE FILAS

Inicialmente, foram coletados dados preliminares referentes ao número de chegadas de clientes e ao tempo de atendimento. Com base nesses dados, utilizou-se a distribuição t de Student para calcular o tamanho da amostra a ser utilizada. Em seguida, foi utilizado o teste de Komogorov-Smirnov para executar o ajuste de distribuições de probabilidade.

Em relação ao número de chegadas de clientes no sistema, a Tabela 1 ilustra os parâmetros para determinação do tamanho de amostra a ser coletada em cada semana, considerando um período entre novembro de 2022 e janeiro de 2023.

Tabela 1 - Parâmetros usados para calcular as amostras de entradas

Parâmetros	Valores
Nível de significância	0,05
Erro padrão da amostra	3
Tamanho da amostra inicial	10
Desvio padrão amostral	2,875
Tamanho sugerido da amostra	5

Fonte: Autor (2023).

A Tabela 2 ilustra os dados obtidos e as estatísticas descritivas.

Tabela 2 - Dados coletados para chegada de clientes

Semanas	Segunda	Terça	Quarta	Quinta	Sexta
Semana 1	11	12	8	6	8
Semana 2	12	5	11	6	5
Semana 3	8	10	12	11	6
Semana 4	13	13	8	9	7
Semana 5	8	13	11	7	5
Semana 6	8	11	13	13	8
Semana 7	9	9	5	12	7
Semana 8	9	7	9	12	10
Média	9,75	10	9,63	9,50	7
Mediana	9	10,50	10	10	7
Moda	8	10; 13	8; 11	6; 12	5; 7; 8
Desvio padrão	1,98	2,88	2,62	2,88	1,69
Coeficiente de variação	20,31%	28,78%	27,17%	30,30%	24,15%

Fonte: Autor (2023).

Em seguida, foi utilizado o teste de Kolmogorov-Smirnov para testar a hipótese nula que preconiza uma distribuição de Poisson para os dados coletados. A um $\alpha = 0.05$, faltam elementos que corroborem para a rejeição desta hipótese. Assim, pode-se tomar como verdade que os dados amostrais seguem uma distribuição de Poisson.

Os parâmetros usados para determinar o tamanho de amostra a ser coletada estão listados na Tabela 3 em relação ao tempo de atendimento dos clientes.

Tabela 3 - Parâmetros usados para calcular as amostras de atendimento

Parâmetros	Valores
Nível de significância	0,05
Erro padrão da amostra	1
Tamanho da amostra inicial	8
Desvio padrão amostral	0,1942
Tamanho sugerido da amostra	1

Fonte: Autor (2023).

A Tabela 4 ilustra os dados obtidos e as estatísticas descritivas tempo de atendimento dos clientes em horas.

Tabela 4 - Dados coletados para tempos de atendimento

Coletas	Tempos de atendimento em horas
Coleta 1	0,89
Coleta 2	0,765
Coleta 3	0,667
Coleta 4	0,92
Coleta 5	0,529
Coleta 6	1,137
Coleta 7	0,995
Coleta 8	0,733
Coleta 9	0,617
Coleta 10	0,731
Média	0,7984
Mediana	0,749
Moda	Amodal
Desvio padrão	0,1853
Coeficiente de variação	23,21%

Fonte: Autor (2023).

Em seguida, é testada a hipótese que preconiza uma distribuição exponencial das amostras. A um $\alpha = 0.05$, o resultado obtido pelo teste de Kolmogorov-Smirnov mostra que faltam elementos que corroborem para a rejeição desta hipótese. Os métodos empregados nesta seção foram implementados via Python e estão apresentados nos apêndices A-D.

Com base no sistema descrito na seção 4.2, os dados coletados para o número de chegadas e para o tempo de atendimento revelam que o modelo de filas que mais se adequa aos objetivos deste estudo é o M/M/1.

5. ANÁLISE DOS RESULTADOS

Este capítulo discute os dados coletados e os resultados. A situação do cenário atual do sistema é apresentada primeiro, seguida dos cenários adotados.

5.1 CENÁRIO ATUAL

Este cenário examina o sistema de entrega atual da empresa usando o modelo M/M/1, que conta com um faturista.

Considerando que, durante o período de 40 dias de coleta de dados, 367 clientes procuraram os serviços da empresa e que a jornada de trabalho diária é de oito horas, o que resulta em um total de 320 horas de trabalho, a razão entre as duas variáveis dará como resultado a taxa de chegada (λ) por horas: 367/320, que resultará em 1,15 clientes por hora. Isso significa que a previsão de chegada de clientes no sistema é de 9,12 clientes por dia. Ainda assim, é importante destacar que isso varia, já que durante os dias de coleta de dados houve casos em que chegaram até 13 clientes e outros em que apenas cinco clientes apareceram no mesmo período.

A taxa de atendimento (μ) é de 1,2523 atendimentos por hora, com a média de 0,7984 horas de duração do atendimento. Multiplicando a taxa de 1,2523 por oito horas, um servidor consegue atender 10 clientes por dia. Como resultado, é possível observar a ineficiência do sistema, pois a taxa de atendimento em alguns dias é significativamente menor do que a taxa de chegada.

Para que o sistema seja considerado equilibrado, ρ deve ser menor ou igual a 1. Como a taxa de utilização (ρ) é de 0,918, o sistema atual é considerado equilibrado. Portanto, o modelo M/M/1 é válido.

A média do número de clientes na fila de espera (L_q) é de 10,33, ao mesmo tempo em que a quantidade média de clientes no sistema (L) é de 11,241. Cada cliente espera em média (W_q) 8,977 horas ou 538,62 minutos (quinhentos e trinta e oito minutos e trinta e sete segundos) na fila. O tempo médio de espera do cliente no sistema (W) é de 9,775 horas ou 586,5 minutos (quinhentos e oitenta e seis minutos e trinta segundos). A probabilidade de que um cliente não precise esperar na fila (P_0) é de 0,082, ou 8,2%. Em outras palavras, existem 8,2% de chances de que nenhum cliente esteja no sistema.

5.2 CENÁRIO 1

Diante dos resultados da situação atual, é evidente que o sistema em estudo não funciona bem. O sistema apresenta um desempenho insatisfatório, pois os clientes enfrentam longos tempos de espera na fila e no atendimento. Além disso, há uma grande quantidade de clientes na fila e no sistema, o que sobrecarrega a capacidade do serviço. A probabilidade de um cliente não precisar esperar na fila é muito baixa, apenas 8,2%. Como resultado, há duas maneiras de aumentar a eficiência do sistema. Uma é limitar a quantidade de atendimento diário de clientes, o que não seria uma boa opção para a empresa em termos de receita, além de impedir o crescimento da empresa. A outra é colocar um pré-requisito para que o cliente forneça um relato do que está sendo feito no veículo; caso o cliente não o faça, ele terá que fazê-lo na empresa ou esperar que o alinhador o forneça. O atendimento dura em média 0,55 horas quando o motorista chega com um relatório pronto, enquanto pode levar mais de 0,9 horas quando o relatório não está pronto.

Como resultado, o cenário 1 que também usará o modelo M/M/1, espera a chegada do cliente com o relato pronto para avaliar o comportamento do sistema como um todo após a implementação dessa mudança. Neste caso, as medidas de desempenho do sistema foram calculadas novamente, mantendo o mesmo λ de 1,15 clientes por hora e modificando apenas o μ , que agora é $1/0,55 = 1,82$ clientes por hora. Os resultados obtidos são: L_q de 1,085, L de 1,716, W_q com 0,943 horas ou 56,58 minutos (cinquenta e seis minutos e trinta e cinco segundos), W com 1,493 horas ou 89,58 minutos (oitenta e nove minutos e trinta e cinco segundos), e P_0 de 0,368. Logo, existe 36,8% de chances de que nenhum cliente esteja no sistema.

5.3 CENÁRIO 2

Com base nos resultados obtidos na situação atual, será analisada a criação de um segundo cenário. Desta forma, o cenário dois usaria as mesmas informações da situação atual que usaram o modelo M/M/1. Entretanto, agora seria usado o modelo M/M/S com a adição de mais um faturista que trabalharia em conjunto com o primeiro. Com a adição de um faturista, haveria uma melhora do sistema, pois com dois servidores, a taxa de atendimento aumenta e o tempo médio de espera diminui. Isso

significa que mais clientes podem ser atendidos em menos tempo. Em seguida, seriam analisados os resultados dos dois cenários buscando entender qual se adequa melhor ao sistema da empresa.

Como resultado, o cenário 2 agora com dois faturistas. As medidas de desempenho do sistema, mantendo o valor de $\lambda = 1,15$ clientes por hora e apenas alterando o valor da taxa de atendimento, que agora será o dobro de antes, sendo $\mu = 2,5046$ clientes por hora. Os resultados foram os seguintes: L_q de 0,245; L de 1,164; W_q de 0,213 horas ou 12,78 minutos (doze minutos e quarenta e sete segundos); W de 1,012 horas ou 60,72 minutos (sessenta minutos e quarenta e três segundos); e P_0 de 0,327 indica 32,7% de probabilidade de que nenhum cliente esteja presente no sistema.

5.4 CENÁRIO 3

Um terceiro cenário será analisado, por se tratar de uma tarefa rotineira que é realizada para todos os clientes com pouca variação do que é realizado, existe uma tendência de que o tempo necessário para realizar o atendimento tenha pouca variação. Portanto, o modelo M/D/1 geralmente dispõe de uma boa representação para situações como esta, tendo em vista que é suposto que os tempos de atendimento sejam iguais.

Segundo Hillier, como o modelo M/D/1 é um caso em que o modelo M/G/1 tem a variância igual a zero, pode-se assumir que as equações sejam as mesmas para ambos, exceto para o L_q , como mencionado no tópico 2.4.4.

O cenário 3 agora seguindo que as taxas de atendimento são constantes, usando as mesmas medidas de desempenho do sistema do cenário atual, com $\lambda = 1,15$ clientes por hora e apenas alterando o valor da taxa de atendimento, $\mu = 1,2523$ clientes por hora. Os resultados foram os seguintes: L_q de 5,162; L de 6,080; W_q de 4,488 horas ou 269,28 minutos (duzentos e sessenta e nove minutos e dezessete segundos); W de 5,287 horas ou 317,22 minutos (trezentos e dezessete minutos e treze segundos); e P_0 de 0,082 indica 8,2% de chances de que nenhum cliente esteja presente no sistema.

5.5 COMPARAÇÃO DOS CENÁRIOS

A Tabela 5 apresenta um resumo de todos esses resultados para que seja simplificado o entendimento e leitura.

Tabela 5 - Resultados dos cenários

Variáveis	Cenário atual	Cenário 1	Cenário 2	Cenário 3
Quantidade de servidores	1	1	2	1
Taxa de chegada dos clientes (h)	1,15	1,15	1,15	1,15
Taxa de atendimento dos clientes (h)	1,2523	1,82	1,2523	1,2523
Taxa de utilização	0,918	0,632	0,459	0,918
Número médio de clientes na fila de espera	10,33	1,085	0,245	5,162
Número médio de clientes no sistema	11,241	1,716	1,164	6,080
Tempo médio de espera na fila (h)	8,977	0,943	0,213	4,488
Tempo médio de espera no sistema (h)	9,775	1,493	1,012	5,287
Probabilidade de um cliente não ter que esperar na fila	0,082	0,368	0,371	0,082

Fonte: Autor (2023).

5.4.1 Comparação entre cenários

Ao comparar a μ do cenário atual com o cenário 1, observamos que a taxa de atendimento no cenário atual era de 1,2523 clientes por hora, enquanto no cenário 1 era de 1,82 clientes por hora. Isso resultou em um atendimento de 14 clientes por dia, 4 a mais do que no cenário anterior, que era de 10 clientes, resultando em um acréscimo de 40% na produtividade.

A taxa de utilização caiu de 0,918 (cenário atual) para 0,632 (cenário 1). Houve uma queda em L de 11,241 (cenário atual) para 1,716 (cenário 1), representando uma redução de 84,72%. L_q reduziu de 10,33 (cenário atual) para 1,085 (cenário 1), apresentando uma queda de aproximadamente 89,5%. Ao comparar o W_q do cenário atual com o cenário 1, constata-se uma redução no tempo de 8,977 horas ou 538,62 minutos (quinhentos e trinta e oito minutos e trinta e sete segundos) no cenário atual

para 0,943 horas ou 56,58 minutos (cinquenta e seis minutos e trinta e cinco segundos) no cenário 1. Isso representa uma diminuição no tempo de 89,5%. Já em relação ao W , caiu de 9,775 horas ou 586,5 minutos (quinhentos e oitenta e seis minutos e trinta segundos) no cenário atual para 1,493 horas ou 89,58 minutos (oitenta e nove minutos e trinta e cinco segundos) no cenário 1. Isso indica uma redução no tempo de 84,7%. A probabilidade de P_0 aumentou de 8,2% para 36,8%.

Ao considerar os cenários 1 e 2, percebe-se que o cenário 1, que inclui um pré-requisito para que o cliente envie um relato com informações sobre o que está sendo feito no veículo, é mais vantajoso do que o cenário atual. Isso aumenta a probabilidade de o cliente não encontrar fila e reduz o tempo médio de espera no sistema. Com a redução desses tempos de atendimento, haveria um grande aumento na satisfação dos clientes com os serviços prestados, além das grandes melhorias encontradas nos tempos de espera tanto em filas quanto no sistema.

Em relação ao cenário atual e ao cenário 2, ao atentar-se à μ no cenário atual e no cenário 2, ambos serão de 1,2523 clientes por hora, já a taxa de utilização diminuiu de 91,8% no cenário atual para 45,9% no cenário 2. O número médio L caiu de 11,241 no cenário atual para 1,164 no cenário 2, tendo uma redução em torno de 89,6%, e o L_q caiu de 10,33 para 0,245 no cenário 2, com uma baixa de 97,6%. Ao comparar o tempo médio W_q no cenário atual com o cenário 2, repara-se uma redução de 8,977 horas ou 538,62 minutos (quinhentos e trinta e oito minutos e trinta e sete segundos) no cenário atual para 0,213 horas ou 12,78 minutos (doze minutos e quarenta e sete segundos) no cenário 2. Isso indica uma queda no tempo de 97,6%. O tempo W baixou de 9,775 horas ou 586,5 minutos (quinhentos e oitenta e seis minutos e trinta segundos) no cenário atual para 1,012 horas ou 60,72 minutos (sessenta minutos e quarenta e três segundos) no cenário 2. Isso mostra uma queda no tempo de 89,6%. A probabilidade de um cliente não precisar esperar na fila P_0 aumentou de 8,2% para 62,7%.

Ao comparar o cenário atual com o cenário 2, é possível observar diversas diferenças que contribuem para a superioridade do cenário 2. Portanto, do mesmo modo que o cenário 1, o cenário 2 também apresenta resultados melhores que o cenário atualmente implantado. Desta forma, será necessária uma comparação entre os dois cenários, tanto o 1 quanto o 2, para selecionar o melhor entre eles.

Ao comparar a μ do cenário atual com o cenário 3, não haverá nenhuma distinção entre os dois, pois foram usados os mesmos valores para ambos os modelos. A taxa de utilização de 0,918 se manteve igual para ambos os cenários. L_q apresentou uma baixa pela metade de seu valor, indo de 10,33 no cenário atual para 5,162 no cenário 3, caindo cerca de 50%, e L também registrou uma queda, indo de 11,241 no cenário atual para 6,080 no cenário 3, a baixa representa 45,9% do seu valor inicial. Ao comparar o tempo médio W_q no cenário atual com o cenário 3, observamos que o tempo do cenário atual decresceu de 8,977 horas ou 538,62 minutos (quinhentos e trinta e oito minutos e trinta e sete segundos) para 4,488 horas ou 269,28 minutos (duzentos e sessenta e nove minutos e dezessete segundos) no cenário 3. Isso indica uma queda no tempo de 50%. Em relação ao W , também teve uma redução, saindo de 9,775 horas ou 586,5 minutos (quinhentos e oitenta e seis minutos e trinta segundos) no cenário atual para 5,287 horas ou 317,22 minutos (trezentos e dezessete minutos e treze segundos) no cenário 3. Isso também mostra uma economia de 50% no tempo. Em relação à probabilidade de P_0 se manteve igual com 8,2%.

Ao considerar ambos os cenários, percebe-se que o cenário 3 tem uma grande queda no tempo de espera do cliente tanto na fila quanto no sistema, devido a essas baixas muito grandes. O cenário 3 se torna uma possível alternativa para ser aplicada, porém ainda é um cenário inferior aos cenários 1 e 2.

Como demonstrado na Tabela 5, a taxa μ (cenário 1) é de 1,82 clientes por hora, que é capaz de atender uma demanda de 14 clientes. No cenário 2, a taxa de atendimento μ será de 1,2523 clientes por hora, que a torna capaz de atender 10 clientes por dia, mas como se trata de 2 atendentes, logo será possível o atendimento de até 20 clientes em um dia de trabalho. Como resultado, a produtividade aumentou em 42,86%, com uma capacidade de atendimento de 20 clientes diários, em vez de 14.

A taxa ρ apresenta uma redução de 63,2% no cenário 1 para 45,9% no cenário 2. O número médio L reduziu de 1,716 do cenário 1 para 1,164 no cenário 2, apresentando um decréscimo de 32,2% de seu valor do cenário 1, e o número L_q reduziu de 1,085 para 0,245 no cenário 2, diminuindo 77,4%. Com a comparação do tempo médio W_q , observamos uma queda de 0,943 horas ou 56,58 minutos (cinquenta e seis minutos e trinta e cinco segundos) no cenário 1 para 0,213 horas ou

12,78 minutos (doze minutos e quarenta e sete segundos) no cenário 2. Isso indica uma redução no tempo de 77,4%, representando uma grande economia de tempo. O tempo W caiu de 1,493 horas ou 89,58 minutos (oitenta e nove minutos e trinta e cinco segundos) no cenário 1 para 1,012 horas ou 60,72 minutos (sessenta minutos e quarenta e três segundos) no cenário 2. Isso mostra uma queda no tempo de 32,2%. A probabilidade de P_0 aumentou de 36,68% para 37,1%.

5.4.2 Escolha do melhor cenário

A redução do tempo de espera do cliente é um fator muito importante no estudo em questão, pois a satisfação do cliente e o sucesso do negócio dependem da redução do tempo de espera. Em alguns casos, os clientes ficam frustrados e irritados se tiverem que esperar muito tempo na fila, o que prejudica a empresa. Os clientes muitas vezes também podem optar por não voltar à empresa ou até mesmo compartilhar suas experiências negativas com outras pessoas. Outro fator prejudicial à empresa em relação a grandes tempos de espera é que, por pressão dos clientes, o serviço é iniciado pelo alinhador antes mesmo de ser aprovado pelo faturista. Devido a isso, quando se chega na etapa de fotografar o veículo para enviar para análise comprovando que o mesmo se encontra no local onde o serviço será prestado, o serviço de alinhamento ou balanceamento já foi finalizado pelo alinhador e o motorista já foi embora da empresa. Em razão de casos como este, o pagamento acaba não sendo aprovado por falta de provas de que o veículo se encontrava no local informado, mas o serviço já foi realizado, o que acarreta em não recebimento da empresa sobre o serviço prestado. Tendo em mente estes motivos, o tempo que o cliente deve esperar é um elemento significativamente influente na escolha do melhor cenário.

Quanto maior a probabilidade de um cliente não ter que esperar na fila, melhor será para a empresa. Esta medida tem diversos impactos positivos tanto para o prestador do serviço quanto para o cliente. Para a prestadora dos serviços, essa probabilidade será usada posteriormente para medir o nível de satisfação dos clientes, a utilização dos recursos, o custo operacional e o lucro. Para o cliente, influenciará na percepção da qualidade do serviço, no tempo disponível, no estresse e na fidelidade. Portanto, como a maioria dos clientes da empresa abordada no estudo são caminhoneiros, que na grande maioria dos casos estão com pressa para retornar aos

seus trabalhos, é de suma importância que o prestador do serviço busque maximizar essa probabilidade, desde que seja economicamente viável e atenda às expectativas dos clientes.

No atual período de crescimento da empresa, havia dias em que a demanda de clientes pelos seus serviços era muito alta. O único faturista que realizava o serviço não dava conta de atender a todos, e precisava pedir ajuda a outra pessoa para cumprir os pedidos dos clientes. Esses acontecimentos evidenciam a necessidade de um segundo faturista para essa área.

Com base nos resultados e nas análises feitas para os quatro cenários, observa-se que o cenário atual foi o que apresentou o pior desempenho em relação aos demais, apresentando uma alta taxa de utilização e altos tempos de atendimento e espera, seguido pelo cenário 3, que também apresentou uma taxa de utilização elevada e grandes tempos de atendimento. O cenário 1 mostra resultados superiores, porém, o pré-requisito que será necessário para sua aplicação pode acabar gerando a insatisfação e o descontentamento dos clientes, porque pode acarretar em sua busca pelos serviços prestados pelo concorrente.

Já o cenário 2 é o que revela resultados mais positivos do que ambos os cenários anteriores, além de expor ótimos resultados em relação aos tempos de espera tanto na fila quanto no sistema dos clientes, dispõe da maior probabilidade de o cliente não ter que enfrentar fila quando chega à empresa, que são dois pontos cruciais para a empresa em questão. Também traz a inclusão de um novo faturista, o que de fato será algo que beneficiará a empresa a curto e longo prazo, pois a empresa está em uma onda de crescimento, tendo cada vez mais uma maior quantidade de clientes em busca dos serviços de alinhamento e balanceamento. Sem a inclusão de mais um faturista em um futuro próximo, a empresa não seria capaz de atender todos os clientes. Logo, com a execução do cenário 2, esse problema futuro não existiria, não sendo necessário realizar outro estudo em busca de uma solução para o mesmo problema em questão.

A empresa deve analisar a implementação de um faturista no processo para trabalhar em conjunto com o atual, pois isso pode trazer benefícios como maior agilidade, qualidade e eficiência. No entanto, antes de tomar essa decisão, a empresa deve fazer uma análise dos custos que esse novo faturista irá trazer, como salário e impostos. Além disso, a empresa deve considerar os encargos trabalhistas e sociais

que incidem sobre o salário, como FGTS, INSS, 13º salário, férias, vale-transporte, vale-refeição, entre outros. Esses custos podem representar até 70% do valor do salário bruto.

Portanto, a empresa deve fazer uma avaliação criteriosa dos custos e benefícios de contratar um novo faturista, levando em conta o impacto financeiro e operacional dessa decisão.

6 CONSIDERAÇÕES FINAIS

Considerando uma empresa prestadora de serviços de balanceamento e alinhamento em veículos pesados, este trabalho aplicou a teoria das filas para determinação da capacidade de atendimento. Através de um mapeamento do estado atual das atividades, foram levantadas as principais características dessas atividades. Também foram coletados dados relativos às taxas de chegada e de atendimento, considerando um período entre novembro de 2022 e janeiro de 2023. Para o planejamento de capacidade, foram considerados três cenários alternativos que foram modelados na linguagem Python em sua versão 3.10.12.

O cenário atual da prestação de serviço de atendimento segue os pressupostos para ser representado por um modelo M/M/1. A partir deste cenário, são propostos três cenários alternativos que modificam a taxa de atendimento, o número de atendentes e o grau de variabilidade do sistema. O cenário 1, que também utilizará o modelo M/M/1, consiste em implementar um pré-requisito para que o cliente forneça um relato do que deverá ser feito no veículo. No cenário 2, foi usado o modelo M/M/S com a adição de mais um faturista que trabalhará em conjunto com o primeiro. Já no cenário 3, que recorreu ao modelo M/D/1, foi suposto que os tempos de atendimento seriam iguais.

O cenário 2 foi o escolhido, pois atendeu a três objetivos que a empresa pretende alcançar, sendo eles: a redução dos tempos de espera dos clientes tanto na fila quanto no sistema, a busca por eliminar casos em que os caminhoneiros, por meio de pressão nos alinhadores, acabam tendo a prestação do serviço feita antes da liberação do faturista, e o aumento da probabilidade de o caminhoneiro não enfrentar fila quando chega na empresa, tendo em vista que os principais clientes são caminhoneiros que estão com pressa para voltar ao trabalho, agilizando assim o processo e gerando uma maior satisfação do cliente. E a adição de um faturista que trabalhará em conjunto com o atual, pois a empresa está em um grande processo de crescimento e, em um futuro próximo, um único faturista, como está no cenário atual, não será capaz de atender todos os clientes.

Por fim, como sugestão de trabalhos futuros, recomenda-se o uso da abordagem Lean Office visando reduzir a variabilidade do tempo de atendimento e também considerar fatores econômicos de longo prazo no planejamento de capacidade.

REFERÊNCIAS

- AGÊNCIA BRASIL. **Serviços avançam e comércio recua na participação no PIB desde 1947**. 3 dez. 2020. Disponível em: <<https://agenciabrasil.ebc.com.br/economia/noticia/2020-12/servicos-avancam-e-comercio-recua-na-participacao-no-pib-desde-1947>>. Acesso em: 29 jul. 2023.
- ANDRADE, Eduardo Leopoldino de. **Introdução à Pesquisa Operacional: métodos e modelos para análise das decisões**. Rio de Janeiro: LTC, 2015. Biblioteca Digital de Teses e Dissertações da UFCG. Disponível em: <<http://dspace.sti.ufcg.edu.br:8080/jspui/handle/riufcg/12890>>. Acesso em: 18 jun. 2023.
- BRONSON, Richard. **Pesquisa operacional**. São Paulo: McGraw-Hill do Brasil, 1985.
- BROWN, L. et al. **Statistical Analysis of a Telephone Call Center: a Queueing-science Perspective**. Wharton Financial Institutions Center, Pennsylvania, n. 4, p.1-37, 5 out. 2004.
- CARRIÓN, Edwin Arturo. **Teoria das Filas como Ferramenta para Análise de Desempenho de Sistemas de Atendimento: Estudo do Caso de Um Servidor da UECE**. Dissertação de Mestrado. Fortaleza – Ceará; Centro Federal de Educação Tecnológica do Ceará – CEFET-CE. 2007.
- COELHO, L. A. **Teoria das filas e simulação: contribuições estratégicas da aplicação em uma unidade bancária na cidade de João Pessoa/PB**. João Pessoa: UFPB, 2016. Disponível em: <<https://repositorio.ufpb.br/jspui/bitstream/123456789/1998/1/LAC06092017.pdf>>. Acesso em: 18 jun. 2023.
- CORRÊA, H. L. **Administração de produção e de operações: manufatura e serviços: uma abordagem estratégica**. São Paulo: Atlas, 2005.
- CORRÊA, HENRIQUE L.; CAON, M. **Gestão de serviços: lucratividade por meio de operações e de satisfação dos clientes**. São Paulo: Atlas, 2002.
- COSTA, Luciano Cajado. **Teoria das Filas**. Disponível em: <http://www.deinf.ufma.br/~mario/grad/filas/TeoriaFilas_Cajado.pdf>. Acesso em: 10 jan. 2023.
- Domingos, B. (2018). **Aplicação da Teoria das Filas para melhorar o nível de serviço em um Restaurante Universitário em Ponta Grossa - PR** [Resumo]. **Queueing Theory Application in the improvement of the Customer Service in a University Restaurant in Ponta Grossa - PR**. Disponível em: <https://www.researchgate.net/publication/323919196_Aplicacao_da_Teoria_das_Filas_para_melhorar_o_nivel_de_servico_em_um_Restaurante_Universitario_em_Ponta_Grossa_PR_Resumo_Queueing_Theory_Application_in_the_improvement_of_the_Customer_Service_in_a_Uni>. Acesso em: 17 jun. 2023.

ÉPOCA NEGÓCIOS ONLINE. **Starbucks avança para o Nordeste do Brasil e terá cafeteria em Salvador em 2023.** São Paulo: Globo, 2022. Disponível em: <<https://epocanegocios.globo.com/Empresa/noticia/2022/09/epoca-negocios-starbucks-avanca-para-o-nordeste-do-brasil-e-tera-cafe-teria-em-salvador-em-2023.html>>. Acesso em: 14 jun. 2023.

FERREIRA FILHO, Virgílio José Martins. **Processos Estocásticos e Teoria das Filas.** Programa de Engenharia da Produção (PEP), Universidade Federal do Rio de Janeiro, Rio de Janeiro, [s. d.].

FERREIRA, M. C.; SANTOS, A. C.; SILVA, A. L. **Análise do perfil dos artigos científicos sobre gestão de projetos publicados no ENEGEP entre 2009 e 2013.** 2014. Disponível em: <https://abepro.org.br/biblioteca/enegep2014_TN_STO_196_113_25111.pdf>. Acesso em: 18 jun. 2023.

FERREIRA, Mariana. **Esperar nas filas na Disney é muito mais divertido do que você pensa. Viagem e Turismo,** 13 jan. 2017. Disponível em: <<https://viagemeturismo.abril.com.br/materias/esperar-nas-filas-na-disney-e-muito-mais-divertido-do-que-voce-pensa/>>. Acesso em: 14 jun. 2023.

FILAH. **Teoria das filas: o que é e como aplicar na gestão do seu negócio?** Disponível em: <<https://blog.filah.com.br/teoria-das-filas/>>. Acesso em: 13 jun. 2023.

FOGLIATTI, MARIA CRISTINA; MATTOS, NELI MARIA COSTA; **Teoria de Filas,** 2007.

FOLHA DE S. PAULO. **Setor de serviços, o principal do PIB, tem alta de 0,4% no primeiro trimestre.** 1 jun. 2021. Disponível em: <<https://www1.folha.uol.com.br/mercado/2021/06/setor-de-servicos-o-principal-do-pib-tem-alta-de-04-no-primeiro-trimestre.shtml>>. Acesso em: 29 jul. 2023.

G1 PERNAMBUCO. **PIB de Pernambuco tem aumento de 5,8% na comparação entre segundo trimestre de 2021 e mesmo período de 2020, diz governo.** 2 set. 2021. Disponível em: <<https://g1.globo.com/pe/pernambuco/noticia/2021/09/02/pib-de-pernambuco-tem-aumento-58percent-na-comparacao-entre-segundo-trimestre-de-2021-e-mesmo-periodo-de-2020-diz-governo.ghtml>>. Acesso em: 29 jul. 2023.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa.** 4. ed. São Paulo: Atlas, 2002.

GUEDES, Debora Barbosa; ARAUJO, Anna Cristina de. **gestão de filas: um estudo de caso em torno da qualidade dos serviços numa agência bancária da região metropolitana do recife - pe.** 2013. Disponível em: <http://www.abepro.org.br/biblioteca/enegep2013_tn_sto_177_014_22639.pdf>. Acesso em: 15 jun. 2023.

HILLIER, F.; LIEBERMAN, G. **Introdução a Pesquisa Operacional**. 8. ed. Porto Alegre: McGraw-Hill, 2010.

IBGE. **Contas Nacionais Trimestrais**: Indicadores de Volume e Valores Correntes - 1º trimestre de 2021. Rio de Janeiro, 2021. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv101873_informativo.pdf>. Acesso em: 29 jul. 2023.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **O que é o PIB?** Disponível em: <<https://ibge.gov.br/explica/pib.php/>>. Acesso em: 29 jul. 2023.

JACOBSEN, Alessandra de Linhares. **Gestão por Resultados, Produtividade e Inovação**. Florianópolis, UFSC, 2009.

KNECHTEL, Maria do Rosário. **Metodologia da pesquisa em educação: uma abordagem teórico-prática dialogada**. Curitiba: Intersaberes, 2014. 193 p.

LIMA, E. R. A.; ALMEIDA, A. T. **Gestão de filas de espera em serviços: uma abordagem multicritério**. *Gestão & Produção*, v. 17, n. 4, p. 729-744, 2010.

Magalhães, M. N. (1996). **Introdução à rede de filas**. Edição da ABE- Associação Brasileira de Estatística

MAISTER, David H. The psychology of waiting lines. **The Service Encounter: Managing Employee/Customer Interaction in Service Business**, Eds. John A. Czepiel, Michael R. Solomon e Carol F. Suprenant. Lexington, MA: Lexington Books, p. 113-23. 1985. Disponível em: <www.davidmaister.com>. Acesso em: 15 jun. 2023.

MARINS, Fernando Augusto Silva. **Introdução a Pesquisa Operacional**. São Paulo: Cultura Acadêmica: Universidade Estadual Paulista, Pró - Reitoria de Graduação, 2011. Disponível em: <http://www.cairu.br/biblioteca/arquivos/Ciencia/Introducao_pesquisa_operacional.pdf>. Acesso em: 10 jun. 2023.

MOORE, Jeffrey H.; WEATHERFORD, Larry R. Tomada de decisão em administração com planilhas eletrônicas. 6. ed. Porto Alegre: Bookman, 2005. **MUTANT. Experiência do cliente em fila: como melhorar?** Disponível em: <<https://mutant.com.br/blog/experiencia-do-cliente-em-fila/>>. Acesso em: 12 jun. 2023.

PEREIRA, Cláudia Sofia Gomes. **A importância da comunicação interna na gestão de recursos humanos: um estudo de caso**. Funchal: Universidade da Madeira, 2011. 104 p. Disponível em: <<https://digituma.uma.pt/bitstream/10400.13/48/1/MestradoCI%C3%A1udiaPereira.pdf>>. Acesso em: 18 jun. 2023.

ROSA, F.; KAMAKURA, W.A. **Pesquisas de satisfação de clientes e efeito halo: interpretações equivocadas?** In: ENANPAD, 25., 2001 Curitiba. Anais... Campinas: ANPAD, 2001. CD-ROM.

SAMEJIMA, K. MATD49-**Estatística não paramétrica** - 5–Testes de Aderência. Salvador: UFBA, s.d. Disponível em: <<https://est.ufba.br/sites/est.ufba.br/files/kim/matd49-aula05-aderencia.pdf>>. Acesso em: 22 fev. 2023.

SILVA, A. C.; SILVA, C. A.; SILVA, J. A.; SILVA, L. A.; SILVA, M. A. **Aplicação da Teoria das Filas em um Supermercado**. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 35., 2015, Fortaleza. Anais... Fortaleza: ABEPRO, 2015. Disponível em: <https://abepro.org.br/biblioteca/tn_stp_211_252_28194.pdf>. Acesso em: 16 jun. 2023.

SILVA, A. L.; SCHAEFFER, L. **Teoria de filas: analisando o fluxo de atendimento e o número de atendentes em um supermercado**. Porto Alegre: UFRGS, 2010. Disponível em: <<https://lume.ufrgs.br/handle/10183/29741>>. Acesso em: 18 jun. 2023.

SILVA, João Mário Rofino da. **Análise da aplicação da teoria das filas no atendimento ao cliente em uma agência bancária**. 2018. 56 f. Trabalho de Conclusão de Curso (Graduação em Engenharia de Produção) - Universidade Federal de Campina Grande, Campina Grande, 2018. Disponível em: <<https://www.todamateria.com.br/referencias-abnt/>>. Acesso em: 10 mai. 2023.

SILVA, Júlio Queiroz Batista da; SILVA, José Vieira da. **Avaliação do tempo de espera em filas de supermercados utilizando a teoria das filas**. Revista Produção Online, Florianópolis, v. 17, n. 4, p. 1289-1312, dez. 2017. Disponível em: <<https://www.normasabnt.org/referencias-bibliograficas/>>. Acesso em: 10 mai. 2023.

SILVA, Luiz Carlos da; SILVA, Maria Cristina da; SILVA, Paulo Roberto da. **Teoria das filas aplicada ao atendimento bancário**. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 35., 2003, Natal. Anais... Natal: SBPO, 2003. p. 1-10. Disponível em: <<https://rockcontent.com/br/talent-blog/referencia-bibliografica-abnt/>>. Acesso em: 11 mai. 2023.

SILVA, Maria. **A importância da leitura**. Repositório. Disponível em: <<https://repositorium.sdum.uminho.pt/handle/1822/65341>>. Acesso em: 18 jun. 2023.

STARBUCKS COFFEE COMPANY. **Aplicativo Starbucks Brasil**. Disponível em: <<https://www.starbucks.com.br/aplicativo-starbucks-brasil>>. Acesso em: 13 jun. 2023.

SUNIM, Haemin. **As coisas que você só vê quando desacelera**. Tradução de Fabiano Morais. Rio de Janeiro: Sextante, 2017.

TAHA, H. A. **Pesquisa Operacional: Uma Visão Geral**. 8. ed. São Paulo: Pearson Prentice Hall, 2008.

TAKE NET. **Fila de atendimento: como reduzir e otimizar o tempo de espera?** Disponível em: <<https://www.take.net/blog/atendimento/fila-de-atendimento/>>. Acesso em: 13 jun. 2023.

VERGARA, Sylvia Constant. **Gestão da Qualidade**. Editora FGV. 3ª Edição. Rio de Janeiro, 2006.

APÊNDICE A – CÓDIGO DE DETERMINAÇÃO DA AMOSTRA DO NÚMERO DE CHEGADAS

```

▶ import math
  from scipy.stats import t

  def tamanho_amostra(nivel_significancia, numero_replicacoes_inicial, desvio_padrao, erro_absoluto):

      n0 = numero_replicacoes_inicial
      graus_liberdade = n0 - 1
      valor_t_student = t.ppf(nivel_significancia/2, graus_liberdade)
      z = valor_t_student
      p = 1 - nivel_significancia / 2
      q = 1 - p
      e = erro_absoluto

      n = (((z*desvio_padrao)/e)**2)

      print('Nível de significância: {}'.format(nivel_significancia))
      print('Erro absoluto aceitável: {}'.format(e))
      print('Número de replicações inicial: {}'.format(numero_replicacoes_inicial))
      print('Desvio-padrão: {}'.format(desvio_padrao))
      print('Valor t student: {}'.format(valor_t_student))
      print('Tamanho da amostra sugerido: {}'.format(math.ceil(n)))

      return math.ceil(n)

[3] nivel_significancia = float(input("Digite o nível de significância desejado: "))
      numero_replicacoes_inicial = int(input("Digite o número da amostra inicial: "))
      desvio_padrao = float(input("Digite o desvio-padrão amostral: "))
      erro_absoluto = float(input("Digite o erro padrão da amostra: "))

      Digite o nível de significância desejado: 0.05
      Digite o número da amostra inicial: 10
      Digite o desvio-padrão amostral: 2.875
      Digite o erro padrão da amostra: 3

▶ tamanho_amostra(nivel_significancia, numero_replicacoes_inicial, desvio_padrao, erro_absoluto)

      Nível de significância: 0.05
      Erro absoluto aceitável: 3.0
      Número de replicações inicial: 10
      Desvio-padrão: 2.875
      Valor t student: -2.262157162740992
      Tamanho da amostra sugerido: 5
      5

```

APÊNDICE B – CÓDIGO DE DETERMINAÇÃO DA AMOSTRA DO TEMPO DE ATENDIMENTO

```

7 import math
  from scipy.stats import t

  def tamanho_amostra(nivel_significancia, numero_replicacoes_inicial, desvio_padrao, erro_absoluto):

      n0 = numero_replicacoes_inicial
      graus_liberdade = n0 - 1
      valor_t_student = t.ppf(nivel_significancia/2, graus_liberdade)
      z = valor_t_student
      p = 1 - nivel_significancia / 2
      q = 1 - p
      e = erro_absoluto

      n = (((z*desvio_padrao)/e)**2)

      print('Nível de significância: {}'.format(nivel_significancia))
      print('Erro absoluto aceitável: {}'.format(e))
      print('Número de replicações inicial: {}'.format(numero_replicacoes_inicial))
      print('Desvio-padrão: {}'.format(desvio_padrao))
      print('Valor t student: {}'.format(valor_t_student))
      print('Tamanho da amostra sugerido: {}'.format(math.ceil(n)))
      return math.ceil(n)

```

```

[5] nivel_significancia = float(input("Digite o nível de significância desejado: "))
     numero_replicacoes_inicial = int(input("Digite o número da amostra inicial: "))
     desvio_padrao = float(input("Digite o desvio-padrão amostral: "))
     erro_absoluto = float(input("Digite o erro padrão da amostra: "))

```

```

Digite o nível de significância desejado: 0.05
Digite o número da amostra inicial: 8
Digite o desvio-padrão amostral: 0.1942
Digite o erro padrão da amostra: 1

```

```

▶ tamanho_amostra(nivel_significancia, numero_replicacoes_inicial, desvio_padrao, erro_absoluto)

Nível de significância: 0.05
Erro absoluto aceitável: 1.0
Número de replicações inicial: 8
Desvio-padrão: 0.1942
Valor t student: -2.3646242510103
Tamanho da amostra sugerido: 1
1

```

APÊNDICE C – CÓDIGO TESTE DE KOLMOGOROV-SMIRNOV POISSON

```

▶ from scipy.stats import poisson
from scipy.stats import kstest
import numpy as np

dados = np.array([11, 12, 8, 6, 8,
                  12, 5, 11, 6, 5,
                  8, 10, 12, 11, 6,
                  13, 13, 8, 9, 7,
                  8, 13, 11, 7, 5,
                  8, 11, 13, 13, 8,
                  9, 9, 5, 12, 7,
                  9, 7, 9, 12, 10])

# Estimativa do parâmetro lambda
mu = dados.mean()

# Distribuição de Poisson com lambda estimado
poisson_dados = poisson(mu)

# Teste de Kolmogorov-Smirnov
kstest_result = kstest(dados, poisson_dados.cdf)

# Encontrando o P-Valor
pvalor = kstest_result.pvalue

# Imprimindo o P-Valor e a média
print('O P-Valor do teste é: {:.3f}'.format(pvalor))
print('A média dos dados é: {}'.format(mu))

# Definindo o valor de Alpha como 0.05
alpha = 0.05

if pvalor > alpha:
    print("Como o P-Valor é maior que Alpha, logo os dados seguem uma distribuição de Poisson.")
else:
    print("Como o P-Valor é menor que Alpha, logo os dados não seguem uma distribuição de Poisson.")

# Neste código, a base dados é o conjunto de dados que você deseja testar
# e alpha é o nível de significância que você deseja usar para
# o teste de Kolmogorov-Smirnov (geralmente é definido como 0,05).
# O código estima o parâmetro lambda da distribuição de Poisson
# usando a média dos dados e gera uma distribuição de Poisson com esse valor estimado.
# Em seguida, ele usa o teste de Kolmogorov-Smirnov para comparar a distribuição
# gerada com os dados reais e determinar se eles seguem uma distribuição de Poisson.

```

O P-Valor do teste é: 0.227
 A média dos dados é: 9.175
 Como o P-Valor é maior que Alpha, logo os dados seguem uma distribuição de Poisson.

APÊNDICE D - CÓDIGO TESTE DE KOLMOGOROV-SMIRNOV EXPONENCIAL

```
import numpy as np
import scipy.stats as stats

def teste_exponencial(dados):
    '''Função que testa se um conjunto de dados segue uma distribuição exponencial'''
    # Calculando o parâmetro lambda da distribuição exponencial
    lamb = 1 / np.mean(dados)

    # Calculando a estatística e o p-valor do teste de Kolmogorov-Smirnov
    ks_stat, p_value = stats.kstest(dados, 'expon', args=(0, lamb))
    print('O p_valor é de: {:.3f}'.format(p_value))

    # Definindo um nível de significância alpha
    alpha = 0.05

    # Comparar o p-valor com o alpha e retornar o resultado
    if p_value > alpha:
        return "Como o p_valor é maior que alpha = 0.05, logo conjunto de dados segue uma distribuição exponencial."
    else:
        return "Como o p_valor é menor que alpha = 0.05, logo conjunto de dados não segue uma distribuição exponencia

# Usando a função
dados = np.array([0.89, 0.765, 0.667, 0.92, 0.529,
                  1.137, 0.995, 0.733, 0.617, 0.731])
resultado = teste_exponencial(dados)
print(resultado)
```

O p_valor é de: 0.055520
Como o p_valor é maior que alpha = 0.05, logo conjunto de dados segue uma distribuição exponencial.

APÊNDICE E – CÓDIGO PYTHON MODELO M/M/1 CENÁRIO ATUAL

```
[1] def mm1(lamda, mu):

    '''Implementação de um modelo de fila M/M/1 em Python,
    a taxa média de chegada de clientes (lamda) e a
    taxa média de serviço (mu) em um sistema de fila.
    Os números médios de clientes na fila,
    números médios de clientes no sistema,
    tempos médios de espera na fila e tempos médios de espera no sistema
    são todos calculados pelo código usando essas taxas. Para calcular as
    métricas para vários sistemas de fila, com os valores dessas variáveis
    podendo serem ajustados.'''

    ro = lamda / mu
    Lq = ((lamda ** 2) / (mu*(mu - lamda)))
    L = lamda / (mu-lamda)
    Wq = (lamda/(mu*(mu-lamda)))
    W = 1 / (mu - lamda)
    P0 = 1 - ro
    return ro, Lq, L, Wq, W, P0
```

```
[2] lamda = float(input("Informe o valor da taxa média de chegada: "))
    mu = float(input("Informe o valor da taxa média de serviço: "))
```

```
Informe o valor da taxa média de chegada: 1.15
Informe o valor da taxa média de serviço: 1.2523
```

```
[3] ro, Lq, L, Wq, W, P0 = mm1(lamda, mu)
```

```
print(f"Com uma taxa média de chegada de {lamda} e uma taxa média de serviço de {mu}, temos:")
print(f"Taxa de utilização: {ro:.3f}")
print(f"Número médio de clientes na fila: {Lq:.3f}")
print(f"Número médio de clientes no sistema: {L:.3f}")
print(f"Tempo médio de espera na fila: {Wq:.3f} horas")
print(f"Tempo médio de espera no sistema: {W:.3f} horas")
print(f"Probabilidade de um cliente não ter que esperar na fila: {P0:.3f}")
```

```
Com uma taxa média de chegada de 1.15 e uma taxa média de serviço de 1.2523, temos:
Taxa de utilização: 0.918
Número médio de clientes na fila: 10.323
Número médio de clientes no sistema: 11.241
Tempo médio de espera na fila: 8.977 horas
Tempo médio de espera no sistema: 9.775 horas
Probabilidade de um cliente não ter que esperar na fila: 0.082
```

APÊNDICE F – CÓDIGO PYTHON MODELO M/M/1 CENÁRIO 1

```
[ ] def mm1(lamda, mu):

    '''Implementação de um modelo de fila M/M/1 em Python,
    a taxa média de chegada de clientes (lamda) e a
    taxa média de serviço (mu) em um sistema de fila.
    Os números médios de clientes na fila,
    números médios de clientes no sistema,
    tempos médios de espera na fila e tempos médios de espera no sistema
    são todos calculados pelo código usando essas taxas. Para calcular as
    métricas para vários sistemas de fila, com os valores dessas variáveis
    podendo serem ajustados.'''

    ro = lamda / mu
    Lq = ((lamda ** 2) / (mu*(mu - lamda)))
    L = lamda / (mu-lamda)
    Wq = (lamda/(mu*(mu-lamda)))
    W = 1 / (mu - lamda)
    P0 = 1 - ro
    return ro, Lq, L, Wq, W, P0
```

```
[ ] lamda = float(input("Informe o valor da taxa média de chegada: "))
mu = float(input("Informe o valor da taxa média de serviço: "))
```

```
Informe o valor da taxa média de chegada: 1.15
Informe o valor da taxa média de serviço: 1.82
```

```
ro, Lq, L, Wq, W, P0 = mm1(lamda, mu)
```

```
[ ] print(f"Com uma taxa média de chegada de {lamda} e uma taxa média de serviço de {mu}, temos:")
print(f"Taxa de utilização: {ro:.3f}")
print(f"Número médio de clientes na fila: {Lq:.3f}")
print(f"Número médio de clientes no sistema: {L:.3f}")
print(f"Tempo médio de espera na fila: {Wq:.3f} horas")
print(f"Tempo médio de espera no sistema: {W:.3f} horas")
print(f"Probabilidade de um cliente não ter que esperar na fila: {P0:.3f}")
```

```
Com uma taxa média de chegada de 1.15 e uma taxa média de serviço de 1.82, temos:
Taxa de utilização: 0.632
Número médio de clientes na fila: 1.085
Número médio de clientes no sistema: 1.716
Tempo médio de espera na fila: 0.943 horas
Tempo médio de espera no sistema: 1.493 horas
Probabilidade de um cliente não ter que esperar na fila: 0.368
```

APÊNDICE G – CÓDIGO PYTHON MODELO M/M/S CENÁRIO 2

```
[ ] from math import factorial
```

```
[ ] def calculo_p0(lamda, mu, S):
    p0 = 0
    for n in range(S):
        p0 += ((lamda/mu)**n)/factorial(n)
    p0 += (((lamda/mu)**S)/(factorial(S)*(1-(lamda/(S*mu)))))
    p0 = 1/p0
    return p0
```

```
[ ] lamda = float(input("Informe o valor da taxa média de chegada: "))
    mu = float(input("Informe o valor da taxa média de serviço: "))
    S = int(input("Informe o número de servidores: "))
```

```
Informe o valor da taxa média de chegada: 1.15
Informe o valor da taxa média de serviço: 1.2523
Informe o número de servidores: 2
```

```
[ ] P0 = calculo_p0(lamda, mu, S)
```

```
[ ] def mms(lamda, mu, S):

    '''Implemente um modelo de fila M/M/S em Python para um sistema de fila,
    usando a taxa média de chegada de clientes (lamda), a taxa média de serviço (mu)
    e a quantidade de servidores (S). Essas taxas são usadas pelo código para calcular
    os números médios de clientes na fila, números médios de clientes no sistema,
    tempos médios de espera na fila e tempos médios de espera no sistema.
    para calcular as métricas de diferentes sistemas de fila, com a possibilidade
    de alterar os valores das variáveis.'''
    P0
    ro = (lamda / (S * mu))
    Lq = (((lamda/mu) ** (S + 1)) / (factorial(S)*(S)*((1 - ro) ** 2))) * P0
    L = Lq + (lamda / mu)
    Wq = Lq / lamda
    W = Wq + (1 / mu)
    return ro, Lq, L, Wq, W
```

```
[ ] ro, Lq, L, Wq, W = mms(lamda, mu, S)
```

```
print(f"Com uma taxa média de chegada de {lamda}, uma taxa média de serviço de {mu} e {S} servidores temos:")
print(f"Taxa de utilização: {ro:.3f}")
print(f"Número médio de clientes na fila: {Lq:.3f}")
print(f"Número médio de clientes no sistema: {L:.3f}")
print(f"Tempo médio de espera na fila: {Wq:.3f} horas")
print(f"Tempo médio de espera no sistema: {W:.3f} horas")
print(f"Probabilidade de um cliente não ter que esperar na fila: {P0:.3f}")
```

```
Com uma taxa média de chegada de 1.15, uma taxa média de serviço de 1.2523 e 2 servidores temos:
Taxa de utilização: 0.459
Número médio de clientes na fila: 0.245
Número médio de clientes no sistema: 1.164
Tempo médio de espera na fila: 0.213 horas
Tempo médio de espera no sistema: 1.012 horas
Probabilidade de um cliente não ter que esperar na fila: 0.371
```

APÊNDICE H – CÓDIGO PYTHON MODELO M/D/1 CENÁRIO 3

```
[1] def md1(lamda, mu):
    '''Calcula os seguintes resultados para o modelo M/D/1 de Teoria das Filas:
    * Taxa de ocupação (ro)
    * Tamanho médio da fila (Lq)
    * Tamanho médio da fila e do sistema (L)
    * Tempo médio de espera na fila (Wq)
    * Tempo médio de espera no sistema (W)
    * Probabilidade de encontrar a fila vazia (P0)

    Parâmetros:
    * lambda_ (float): Taxa de chegada.
    * mu (float): Taxa de atendimento.

    Retorno:
    * ro (float): Taxa de utilização.
    * Lq (float): Tamanho médio da fila.
    * L (float): Tamanho médio da fila e do sistema.
    * Wq (float): Tempo médio de espera na fila.
    * W (float): Tempo médio de espera no sistema.
    * P0 (float): Probabilidade de encontrar a fila vazia.'''
    ro = (lamda / mu) # Taxa de utilização
    Lq = ((ro**2) / (2 * (1 - ro))) # número médio na fila
    L = Lq + ro # número médio no sistema
    Wq = Lq / lamda # tempo médio na fila
    W = Wq + (1 / mu) # tempo médio no sistema
    P0 = 1 - ro # probabilidade do sistema estar vazio
    return ro, Lq, L, Wq, W, P0
```

```
[2] lamda = float(input("Informe o valor da taxa média de chegada: "))
    mu = float(input("Informe o valor da taxa média de serviço: "))
```

```
Informe o valor da taxa média de chegada: 1.15
Informe o valor da taxa média de serviço: 1.2523
```

```
[3] ro, Lq, L, Wq, W, P0 = md1(lamda, mu)
```

```
print(f"Com uma taxa média de chegada de {lamda} e uma taxa média de serviço de {mu}, temos:")
print(f"Taxa de utilização: {ro:.3f}")
print(f"Número médio de clientes na fila: {Lq:.3f}")
print(f"Número médio de clientes no sistema: {L:.3f}")
print(f"Tempo médio de espera na fila: {Wq:.3f} horas")
print(f"Tempo médio de espera no sistema: {W:.3f} horas")
print(f"Probabilidade de um cliente não ter que esperar na fila: {P0:.3f}")
```

```
Com uma taxa média de chegada de 1.15 e uma taxa média de serviço de 1.2523, temos:
Taxa de utilização: 0.918
Número médio de clientes na fila: 5.162
Número médio de clientes no sistema: 6.080
Tempo médio de espera na fila: 4.488 horas
Tempo médio de espera no sistema: 5.287 horas
Probabilidade de um cliente não ter que esperar na fila: 0.082
```