

UNIVERSIDADE FEDERAL DA PARAÍBA – UFPB

Centro de Ciências Sociais Aplicadas – CCSA

Graduação em Administração – GADM

**APLICAÇÃO DE TEORIA DAS FILAS EM UMA UNIDADE DA
DEFENSORIA PÚBLICA DA UNIÃO**

ANDRÉ CAVALCANTI FLORENTINO

João Pessoa

Agosto 2025

ANDRÉ CAVALCANTI FLORENTINO

**APLICAÇÃO DE TEORIA DAS FILAS EM UMA UNIDADE DA
DEFENSORIA PÚBLICA DA UNIÃO**

Trabalho de Curso apresentado como parte dos
requisitos necessários à obtenção do título de
Bacharel em Administração, pelo Centro de
Ciências Sociais Aplicadas, da Universidade
Federal da Paraíba / UFPB.

Docente Orientadora: Profa. Dra. Pamela Adelino
Ramos Albertins

João Pessoa

Agosto 2025

Catálogo na publicação
Seção de Catalogação e Classificação

F633a Florentino, André Cavalcanti.

Aplicação de Teoria das Filas em uma unidade da
Defensoria Pública da União / André Cavalcanti
Florentino. - João Pessoa, 2025.

22 f. : il.

Orientação: Pamela Adelino Ramos Albertins.
TCC (Graduação) - UFPB/CCSA.

1. Pesquisa Operacional. 2. Teoria das filas. 3.
Defensoria Pública da União. 4. Serviço público. 5.
Administração Pública. I. Albertins, Pamela Adelino
Ramos. II. Título.

UFPB/CCSA

CDU 658

Folha de aprovação

Trabalho apresentado à banca examinadora como requisito parcial para a Conclusão de Curso do Bacharelado em Administração.

Aluno: André Cavalcanti Florentino

Trabalho: Aplicação de Teoria das Filas em uma unidade da Defensoria Pública da União

Área da pesquisa: Administração da Produção

Data de aprovação: 21/08/2025

Banca examinadora

Profª. Dra. Pamela Adelino Ramos Albertins

Orientadora

Prof. Dr. André Gustavo Carvalho Machado

Avaliador 1

Prof. Dr. Fábio Walter

Avaliador 2

RESUMO

O objetivo deste estudo foi analisar o comportamento da fila no atendimento primário na Defensoria Pública da União em João Pessoa - PB, utilizando a Teoria das Filas como ferramenta de avaliação do desempenho. Trata-se de uma pesquisa aplicada e quantitativa e caracterizada como estudo de caso, com coleta de dados *in loco* durante um dia considerado normal e típico de funcionamento da unidade, seguindo o modelo M/M/3. Foram mensurados indicadores como taxa de utilização, probabilidade de o sistema estar vazio, número médio de clientes na fila, tempos médios de espera e de permanência no sistema. Os resultados apontaram taxa de utilização de 35,62%, probabilidade de 31,65% de não haver clientes no sistema, fila praticamente inexistente ($L_q = 0,055$) e tempo médio de espera de 0,91 segundos. Constatou-se elevada ociosidade, indicando possibilidade de redistribuição de recursos para otimizar a eficiência sem comprometer a qualidade do atendimento. Conclui-se que o sistema apresenta agilidade acima da demanda observada, sendo recomendada a ampliação da amostragem e simulações futuras para subsidiar decisões gerenciais.

Palavras-chave: Pesquisa Operacional, Teoria das filas, Defensoria Pública da União, Serviço público, Administração Pública.

ABSTRACT

The objective of this study was to analyze the behavior of the primary care queue at the Federal Public Defender's Office in João Pessoa, Paraíba, using Queuing Theory as a performance assessment tool. This is an applied, quantitative study characterized as a case study, with on-site data collection during a normal and typical operating day for the unit, following the M/M/3 model. Indicators such as utilization rate, probability of the system being empty, average number of customers in the queue, average wait times, and dwell time in the system were measured. The results indicated a 35.62% utilization rate, a 31.65% probability of no customers in the system, a virtually non-existent queue ($L_q = 0.055$), and an average wait time of 0.91 seconds. High idleness was observed, indicating the possibility of reallocating resources to optimize efficiency without compromising service quality. The conclusion is that the system performs agility above the observed demand, and sampling expansion and future simulations are recommended to support management decisions.

Keywords: Operational Research, Queueing Theory, Public Defender's Office of the Union, Public Service, Public Administration.

LISTA DE FIGURAS

Figura 1 - Unidade da DPU pesquisada.....	14
---	----

LISTA DE TABELAS

Tabela 1 - Equações utilizadas.....	13
Tabela 2 - Intervalo entre chegadas (valores em minutos).....	15
Tabela 3 - Duração dos atendimentos (valores em minutos).....	16
Tabela 4 - Tempo de espera na fila (valores em minutos).....	17

LISTA DE EQUAÇÕES

Equação 1 - Taxa de utilização do sistema.....	17
Equação 2 - Probabilidade de não ter nenhum cliente no sistema.....	17
Equação 3 - Número médio de clientes aguardando na fila do sistema.....	18
Equação 4 - Tempo médio que um cliente aguarda na fila.....	18
Equação 5 - Número médio de clientes no sistema.....	18
Equação 6 - Tempo médio de espera no sistema.....	18

SUMÁRIO

1. INTRODUÇÃO.....	10
2. REFERENCIAL TEÓRICO.....	10
2.1 Pesquisa Operacional.....	10
2.2 Teoria das Filas.....	11
2.3 Filas nos serviços de atendimento público.....	13
3. METODOLOGIA.....	13
3.1 Amostra.....	13
3.2 Objetivo da pesquisa.....	15
3.3 Coleta de dados.....	15
4. ANÁLISE DOS RESULTADOS.....	15
5. CONSIDERAÇÕES FINAIS.....	19
REFERÊNCIAS.....	20

1. INTRODUÇÃO

A Pesquisa Operacional (PO) faz uso de métodos analíticos e científicos, podendo ser um modelo matemático, estatístico ou algoritmos computacionais para atingir seu objetivo final de investigar um problema e encontrar sua solução ótima de maneira sistêmica, segundo Andrade (2002), auxiliando na tomada de decisão por parte dos(as) gestores(as).

Como apontado por Hillier e Lieberman (2013) esse campo de estudo foi desenvolvido no Reino Unido enquanto a segunda guerra mundial se desenrolava, e inicialmente era aplicada militarmente visando diminuir custos logísticos. Após o período bélico, esse conhecimento começou a ser utilizado em áreas distintas, como finanças, logística e produção, tendo seu uso ampliado para a otimização de processos organizacionais.

Dentro desse campo matemático, existe uma área de estudo conhecida como Teoria das Filas que se preocupa com a dinâmica das linhas de espera, ou filas. Apesar de não ser uma técnica de otimização, sua aplicação é importante para analisar sistemas onde o tempo de chegada e tempo na fila são essenciais, como em bancos, hospitais, aeroportos, portos, fábricas e centrais de atendimento. Utilizando análises matemáticas e informações mensuráveis (como frequência e tempo de chegada, quantidade de canais de atendimento), é possível projetar um sistema mais eficiente através desses estudos, reduzindo tempo de espera e custos operacionais, aumentando a satisfação de clientes, como exposto por Jiayi (2023).

O presente trabalho tem como objetivo investigar a dinâmica da fila do atendimento primário em uma unidade da Defensoria Pública da União na cidade de João Pessoa, Paraíba, segundo o modelo M/M/s (chegadas e atendimentos exponenciais com “s” servidores).

Este trabalho surgiu da necessidade de entender as constantes queixas dos servidores em relação à demanda do atendimento que ouvi durante os dois anos em que fui estagiário da organização. A relevância do estudo se apoia nos pilares sociais, buscando melhoria de qualidade de vida para funcionários e assistidos, e acadêmicos, por não terem sido encontradas pesquisas sobre a Teoria das Filas realizadas em atendimentos de Defensorias.

O artigo está organizado nas seguintes seções: esta seção 1 apresentou uma curta introdução sobre a Pesquisa Operacional, a Teoria das Filas, a justificativa e relevância da pesquisa. A seção 2 fornece uma breve revisão da literatura sobre as áreas de conhecimento abordadas no estudo. Na seção 3 é apresentada a metodologia da pesquisa. Ela descreve como o universo e a amostra foram escolhidos, o tipo de pesquisa realizada, como os dados foram coletados, quais etapas foram seguidas e o processamento dos dados. Também será discutido o objetivo da pesquisa. As seções 4 e 5 versam sobre os resultados obtidos e as conclusões com base nos resultados, a contribuição deste estudo e suas limitações, respectivamente.

2. REFERENCIAL TEÓRICO

Nesta seção serão apresentados, primeiramente, os conceitos sobre a Pesquisa Operacional (PO) e sua importância para investigação e melhoria dos sistemas. Em seguida, será abordado sobre a Teoria das Filas, que tem como objetivo prever como as filas são formadas e analisar suas características. Por fim, é falado sobre a formação de filas em serviços públicos e como isso impacta a percepção do público sobre a satisfação no atendimento.

2.1 Pesquisa Operacional

A PO, sigla pela qual é conhecida, se desenvolveu na Inglaterra durante a Segunda Guerra Mundial para impedir que rotas civis e comerciais fossem atacadas pelos nazistas e também resolver problemas estatísticos e logísticos com fins militares, como transporte de mantimentos, alocação de tropas e simulação de ataques. Ela se utiliza de conceitos analíticos

com o objetivo de otimizar um sistema através da redução de custos ou maximização de lucro, melhor distribuição de recursos, identificar gargalos e dar base para decisões organizacionais.

Do fim da guerra até os tempos atuais, a PO vem se popularizando em diferentes setores, como transporte, produção e telecomunicações, de acordo com Gupta *et al.* (2021). Um estudo de Pesquisa Operacional normalmente consiste em seis fases: formulação do problema, construção do modelo do sistema, cálculo da solução por meio do modelo, teste do modelo e da solução, estabelecimento de controles da solução e implementação e monitoramento da solução (Ramos; Pereira, 2023).

Os modelos de problemas da pesquisa operacional são categorizados em determinísticos, híbridos ou estocásticos, como apontado pelo IFORS (2020). Alguns exemplos em cada uma dessas categorias são os modelos determinísticos que se subdivide em otimização linear (sendo eles a programação linear, programação inteira, modelos de transporte, modelos de rede e tomada de decisão multicritério), otimização não-linear (por exemplo o modelo clássico, o modelo de busca e a programação não-linear), modelos híbridos (como a metodologia PERT-CPM, a programação dinâmica, os modelos de inventário e a simulação) e modelos estocásticos (como os modelos de análise de decisão, modelos de Markov e modelos de filas).

Para que seja possível desenvolver, em uma organização, um estudo que utilize pesquisa operacional, é necessário seguir por certas etapas, procedendo de modo adequado, a fim de alcançar os resultados almejados pelo estudo. [...] Sendo assim, a pesquisa inicia por compreender quais são os fatores relacionados ao problema que será tratado no estudo.

(Melo, 2020, p. 19-20)

Para esse estudo iremos utilizar o modelo estocástico que analisa a formação de filas e suas características, mais conhecido como Teoria das Filas.

2.2 Teoria das Filas

Quando se fala em “fila” não estamos nos referindo apenas a clientes esperando por atendimento em uma loja ou banco, esse termo também é aplicado para filas de barcos esperando para carga ou descarga em um porto, aviões aguardando pelo pouso ou decolagem em um aeroporto e até mesmo arquivos esperando para serem processados por um computador. É importante comentar que a aplicação dessa teoria não é utilizada para a otimização de um sistema, mas para medir o desempenho e a eficiência do sistema, servindo de base para auxiliar no planejamento por parte dos gestores.

Atrasos e esperas podem significar ineficiência no sistema e perdas significativas monetárias, de clientes e de produção. Entretanto, nem todas as filas são problemas, dessa forma é preciso achar o meio termo entre os custos dos serviços e o tempo de espera para o atendimento ser realizado, como afirma Hillier e Lieberman (2013). O objetivo essencial é equilibrar os custos de espera (custo do local adequado para espera na fila, perda de negócios por eventual desistência por causa da espera, entre outros) com os gastos relacionados para manter a capacidade de atendimento (contratação de mais funcionários, instalação de novos balcões, etc). Um caso para ensino onde foi aplicada a Teoria das Filas no atendimento de uma lotérica expõe de maneira didática o que ocorre quando o sistema está sobrecarregado e como isso afeta a satisfação e bem-estar dos clientes e operadores do caixa (Santos, 2024).

A existência da fila propriamente dita não é um problema, mas seu congestionamento é, e a não-formação de uma, ou seja, muita ociosidade dos servidores, pode ser bastante cara para ser mantida pela organização. Segundo Ramos, Melo e Gonçalves (2023), o gerenciamento de sistema de filas e seu processo decisório deve levar em consideração todos os fatores que auxiliem os gestores na tomada de decisão.

Segundo Hillier e Lieberman (2013), o mecanismo de um sistema de fila funciona com a fonte ou população gerando os clientes que chegam a um estabelecimento. Nesse sistema eles podem se dirigir ao servidor e iniciar o serviço imediatamente caso não haja fila, ou esperar, caso todos os balcões estejam ocupados. Quando um serviço é finalizado, o próximo cliente na fila é chamado. Quando a fila ficar vazia, o sistema estará ocioso até que o próximo cliente chegue.

Como exposto por Leonelli (2021), as características das filas são: população, podendo ser infinita (quando a chegada de clientes pode ocorrer a qualquer momento, podendo exceder a capacidade do sistema) ou finita (quando a quantidade possível de clientes é limitada); disciplina da fila (referente a ordem de como o atendimento é realizado), levando em consideração regras como PEPS (primeiro que entra vai ser primeiro atendido), UEPS (último que entra vai ser primeiro atendido), prioridade de atendimento e as características de chegada e atendimento, que definem se o padrão de chegada é controlado ou não, se a chegada dos clientes é individual ou em grupos e o processo de decisão dos clientes para entender se irão escolher entre entrar ou não na fila e sua disposição em permanecer ou não no sistema.

Filas com prioridade de atendimento são divididas entre preemptiva ou não-preemptiva. O primeiro define que clientes de maior prioridade devem ser atendidos imediatamente, interrompendo o atendimento de quem esteja sendo atendido no momento, e o segundo define que o cliente prioritário deverá passar na frente da fila, mas sem interromper o atendimento que já está sendo realizado, como aponta Hillier e Lieberman (2013).

Sobre os parâmetros, são duas maneiras de abordar a chegada de clientes no sistema, uma é em relação ao número de clientes que chegam em um intervalo de tempo λ e o tempo que se passa entre uma chegada e outra ($1/\lambda$). Para o atendimento aos clientes, ou duração do serviço, se utiliza μ para representar a quantidade de atendimentos realizados em um intervalo de tempo (taxa média de atendimentos) e $(1/\mu)$ como o tempo médio de atendimento.

Conforme exposto por Corrêa e Corrêa (2022), a medição do Desempenho do Sistema adota as seguintes medidas de desempenho: Percentual de tempo que o sistema fica ocioso ou ocupado ρ , número médio de clientes no sistema em um determinado tempo L , o número médio de clientes na fila L_q , tempo médio gasto pelo cliente no sistema W , tempo médio que cada cliente gasta na fila de espera W_q e a probabilidade de existir n clientes no sistema P_n .

Para medir o desempenho do sistema levando em consideração a taxa de utilização do sistema ρ , a equação $\rho = \lambda/\mu$ é utilizada quando tem-se apenas um único servidor atendendo e $\rho = \lambda/s\mu$ quando há mais de um servidor. Caso $\rho < 1$, significa que a fila está sob controle, mas se $\rho > 1$, então a fila está fora do controle e cresce sem limites. A equação $1-\rho$ nos fornece a taxa de ociosidade do sistema. Outras medidas de desempenho vão depender do tipo de sistema que a organização utiliza.

Virgillito (2018) expõe que a nomenclatura adotada para padronizar um tipo de problema nos casos em que os tamanhos da fila e população sejam infinitos segue o modelo A/B/s, em que A significa o tipo de distribuição de chegada, B o tipo de distribuição da duração do atendimento e s a quantidade de servidores disponíveis para realizar o serviço. Nesse sistema, percebeu-se através da observação que o modelo mais adequado ao cenário encontrado é o M/M/s, onde os tempos de chegada se distribuem de forma exponencial, com M significando chegadas de *Poisson*, onde o mesmo autor informa que essa é uma variável aleatória discreta, significando que “devido à aleatoriedade das chegadas, o modelo matemático que melhor descreve os chamados “eventos raros” é a distribuição de probabilidades discreta de Poisson” (Virgillito, 2018, p. 307). A duração de cada atendimento é também exponencial e conta com 3 servidores para realizar o atendimento. As equações utilizadas para este modelo estão descritas na Tabela 1.

Tabela 1 - Equações utilizadas

Numeração	Nome	Equação
Equação 1	Taxa de utilização do sistema	$\rho = \frac{\lambda}{s\mu}$
Equação 2	Probabilidade de não ter nenhum cliente no sistema	$P_0 = \left[\sum_{n=0}^{s-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^s}{s!} \left(\frac{s\mu}{s\mu - \lambda} \right) \right]^{-1}$
Equação 3	Número médio de clientes aguardando na fila do sistema	$Lq = \frac{P_0(\lambda/\mu)^{s+1}}{(s-1)! \times (s - \frac{\lambda}{\mu})^2}$
Equação 4	Tempo médio que um cliente aguarda na fila do sistema	$Wq = \frac{Lq}{\lambda}$
Equação 5	Número médio de clientes no sistema	$L = Lq + \frac{\lambda}{\mu}$
Equação 6	Tempo médio de espera no sistema	$W = Wq + \frac{1}{\mu}$

Fonte: Adaptado de Corrêa e Corrêa (2022)

2.3 Filas nos serviços de atendimento público

No Brasil, a prioridade em filas é um direito previsto na lei N° 14.626/23, que atualizou as Leis anteriores N° 10.048/00 e N° 10.205/01 para incluir outros grupos preferenciais, como informa Prado (2024). Dessa maneira, o atendimento deve priorizar pessoas com deficiência física ou intelectual, idosos, gestantes e lactantes, tutores com crianças de colo, pessoas com obesidade e doadores de sangue.

Filas excessivas no atendimento em serviço público costumam ser frequentes, e como forma de impor limite a isso, algumas cidades definem normas para que a duração da espera de um cliente não ultrapasse 30 minutos em bancos até o início do serviço, como informado pelo Jusbrasil (2025). Balcões de atendimento ao público são um bom cenário para exemplificar a aplicação da teoria das filas. Eles frequentemente lidam com fluxo de alta demanda, recursos limitados e procedimentos complexos, como apontado por Jiayi (2023). É importante existir uma gestão desses sistemas para que não haja comprometimento no serviço, como interrupções, demoras excessivas, ineficiência, desperdício de recursos e insatisfação dos consumidores.

3. METODOLOGIA

Nesta seção será abordada a metodologia empregada no estudo, incluindo a classificação da amostra, o método e instrumento de coleta, objetivo da pesquisa e a coleta e análise dos dados obtidos.

3.1 Amostra

Pela sua natureza, este estudo é classificado como pesquisa exploratória, pois é um método científico que tem como principal objetivo gerar conhecimento através da análise de uma situação real para entender um problema, sendo possível apresentar soluções e foi realizada em uma unidade da Defensoria Pública da União, situada na cidade de João Pessoa (PB).

A Defensoria Pública da União é uma autarquia federal, o que significa que é “um serviço com autonomia, criado por lei, com personalidade jurídica, patrimônio e receita próprios e que possui gestão administrativa e financeira descentralizada” (Brasil, 1967). As autarquias são regidas pelo Artigo 5º do Decreto-Lei nº 200/67, de acordo com o Conselho Nacional do Ministério Público. A defensoria pública oferece atividades de assistência jurídica para pessoas em situação de hipossuficiência (falta de recursos financeiros para acessar a justiça), entidades sem fins lucrativos e associações comunitárias. Entre suas principais funções está a orientação jurídica adequada, defesa judicial e extrajudicial, defender direitos individuais e coletivos e conscientização da população, como informado no portal da Defensoria Pública da União.

Figura 1 - Unidade da DPU pesquisada



Fonte: Dados da pesquisa (2025).

A DPE-PB (2025) aponta que o serviço prestado pelas defensorias à população é de grande importância como função essencial para garantir a democratização do acesso à justiça de pessoas em vulnerabilidade social, fundamentada pela inafastabilidade da jurisdição prevista no Artigo 5º da Constituição Federal, inciso LXXIV e na Lei Complementar Nº80, de 12 de Janeiro de 1994, que organiza a Defensoria Pública da União e dos territórios e prescreve normas gerais para sua organização nos estados, como informado pelo portal do Tribunal Superior Eleitoral.

Como dito anteriormente, as câmaras municipais costumam legislar para propor normas de regulamentação em prol da defesa do consumidor, definindo tempo máximo de espera e condições mínimas de salubridade. No caso da DPU, quem legisla e tem jurisdição na fiscalização sobre seu funcionamento é a União. Por ter sido criada relativamente a pouco tempo, ainda não há nenhuma norma que regule o tempo de espera e as condições de conforto na fila. Esse estudo busca analisar a eficiência do sistema na unidade de João Pessoa para sugestões à gestão local.

Nesse estudo, o modelo utilizado e que mais se adequa a realidade da DPU é o M/M/3, os clientes chegam ao local e se dirigem à fila para serem atendidos por um dos três servidores, onde irão realizar o cadastro (para novos assistidos), realizar a consulta do processo (para pessoas que já estão sendo acompanhadas pela defensoria), buscar orientações ou anexar documentos ao processo.

3.2 Objetivo da pesquisa

O principal objetivo deste trabalho é entender a dinâmica da fila no atendimento primário na unidade da DPU em João Pessoa, PB.

Entre as variáveis observadas pode-se destacar a formação de filas ou ociosidade, a eficiência do sistema, a taxa de utilização do sistema, a probabilidade de ter ou não clientes no sistema, o número médio de clientes aguardando na fila, o tempo médio que um cliente aguarda na fila, o número médio de clientes no sistema e o tempo médio de espera no sistema, que são estratégicos para a gestão organizacional e satisfação de clientes e servidores.

3.3 Coleta de dados

A coleta de dados foi realizada *in loco*, dentro da própria autarquia federal (com permissão dos gestores), durante um dia típico de serviço, sendo composto por dois turnos (manhã e tarde) durante uma quarta-feira, 21 de maio de 2025, de 08:00 da manhã até 03:00 da tarde, contabilizando 7 horas de funcionamento do sistema. Utilizou-se cronômetros digitais para contabilizar o tempo decorrido da chegada do cliente no sistema até a conclusão do serviço e uma tabela física para anotar os dados de chegada, permanência e saída. Esses dados foram utilizados para aplicar nos cálculos que envolvem a Teoria das Filas. O dia foi escolhido devido a normalidade da demanda.

A observação não-participante (onde o autor atua apenas como analisador e não há interferência no recolhimento de dados) foi a técnica adotada para o levantamento das informações necessárias para as equações.

A metodologia da pesquisa adotada é quantitativa, para esse caso levou-se em consideração o número de clientes ao longo de um dia e a periodicidade de chegada e permanência, não sendo necessário um contato mais aprofundado ou pessoal com os indivíduos presentes no sistema.

Esse trabalho configura-se como estudo de caso com base em observação empírica do sistema de atendimento da DPU-PB e tem como objetivo analisar a dinâmica da fila no atendimento primário.

A instituição disponibilizou dois dias para a realização desta pesquisa. Entretanto, no primeiro dia houve muitos imprevistos como falta de energia elétrica, queda do sistema e comportamento de fila atípico devido a chuvas intensas na região. Foi permitido então que o pesquisador voltasse em outro dia típico de atendimento.

Para o tratamento dos dados foram utilizadas fórmulas matemáticas do assunto Teoria das Filas, seguindo o modelo M/M/3, e apresentadas na Tabela 1. Para auxílio nos cálculos foi utilizada planilha eletrônica.

4. ANÁLISE DOS RESULTADOS

O fluxo de atendimento segue de tal maneira: O sistema, identificado como Atendimento, possui 3 servidores disponíveis e dedicados aos diversos tipos de serviços, podendo ser uma consulta processual, digitalização e anexação de documentos, edição de narrativa, dentre outros casos que, a depender da situação, afetam o tempo de atendimento, independente da agilidade do operador. Para o atendimento, foram observados os valores presentes na Tabela 2.

Tabela 2 - Intervalo entre chegadas (valores em minutos)

Cliente	Intervalo	Momento		Cliente	Intervalo	Momento
1	9	8:09		13	24	10:58

2	14	8:23		14	7	11:05
3	10	8:33		15	12	11:17
4	0	8:33		16	56	12:13
5	11	8:44		17	51	13:04
6	23	9:07		18	3	13:07
7	7	9:14		19	47	13:54
8	42	9:56		20	3	14:00
9	14	10:10		21	6	14:06
10	15	10:25		22	14	14:20
11	8	10:33		23	2	14:22
12	1	10:34				

Fonte: Elaborado pelo autor (2025).

O valor médio dos intervalos acima é de aproximadamente 16,43 minutos, e portanto o sistema funcionou com um ritmo médio de 3,6508 chegadas por hora (λ).

A Tabela 3 a seguir apresenta a duração em minutos de cada atendimento por cliente:

Tabela 3 - Duração dos atendimentos (valores em minutos)

Cliente	Intervalo		Cliente	Intervalo
1	12		13	23
2	4		14	6
3	35		15	16
4	6		16	6
5	21		17	42
6	26		18	34
7	6		19	25
8	23		20	23
9	20		21	7
10	34		22	9
11	9		23	8
12	9			

Fonte: Elaborado pelo autor (2025).

O valor médio dos dados apresentados na Tabela 3 acima é de 17,5652 minutos, portanto, podemos dizer que os servidores têm capacidade de atender 3,4158 clientes por hora (μ).

Na Tabela 4 podem ser encontrados os tempos que cada cliente esperou individualmente na fila até o atendimento iniciar. Os clientes 1, 3, 6, 8, 10, 11, 12, 13, 14, 18, 19, 21, 22 e 23 tiveram que esperar em fila, com um Tempo Médio na fila = 6,478 minutos. Não se sabe o motivo que ocasionou a espera desses assistidos em particular.

No sistema, a Taxa de Utilização foi calculada utilizando a Equação 1 a seguir:

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

$$= 3,6508 \div (3 \times 3,4158) = 0,35626598 \approx 35,62\%$$

Uma taxa de utilização menor que 1 ($\rho < 1$, onde 1 = 100%) é necessária para que o sistema encontre o equilíbrio, e no atendimento da DPU-PB o sistema se encontrou ocupado em aproximadamente 35,62% do tempo de operação, ou seja, com ociosidade próxima a 64,38% do tempo. Esse alto índice de ociosidade pode indicar subutilização dos recursos do sistema.

Tabela 4 - Tempo de espera na fila (valores em minutos)

Cliente	Intervalo		Cliente	Intervalo
1	1		13	24
2	0		14	5
3	23		15	0
4	0		16	0
5	0		17	0
6	12		18	1
7	0		19	2
8	39		20	0
9	0		21	3
10	6		22	2
11	27		23	3
12	1			

Fonte: Elaborado pelo autor (2025).

A Probabilidade de não ter nenhum cliente no sistema foi calculada através da Equação 2 a seguir:

$$P0 = \left[\sum_{n=0}^{s-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^s}{s!} \left(\frac{s\mu}{s\mu - \lambda} \right) \right]^{-1} \quad (2)$$

$$= \left[\frac{(3,6508/3,4158)^0}{0!} + \frac{(3,6508/3,4158)^1}{1!} + \frac{(3,6508/3,4158)^2}{2!} x \left(\frac{3 x 3,4158}{3 x 3,4158 - 3,6508} \right) \right]^{-1}$$

Com o resultado obtido $P_0 = 0,316500537 \approx 31,65\%$, foi encontrada essa probabilidade do sistema de atendimento estar vazio (sem clientes na fila ou em atendimento com um dos três servidores dessa fase). Significa que o sistema tem uma probabilidade relativamente alta de nunca estar vazio, com pelo menos um assistido em atendimento, indicando que há boa chance de ter algum atendimento em andamento na maior parte do tempo.

O número médio de clientes aguardando na fila do sistema foi obtido com a Equação 3 a seguir:

$$Lq = \frac{P_0(\lambda/\mu)^{s+1}}{(s-1)! x (s - \frac{\lambda}{\mu})^2} \quad (3)$$

$$= \frac{0,316500537(3,6508/3,4158)^{3+1}}{(3-1)! x (3 - \frac{3,6508}{3,4158})^2} = 0,05536960861$$

O resultado encontrado foi um número extremamente pequeno, praticamente zero. Na prática, isso significa que quase nunca há clientes em espera, a fila está quase sempre vazia. No tópico 2.2 foi informado que a não formação de fila pode significar um alto custo a ser mantido pela organização.

O tempo médio que um cliente aguarda na fila do sistema foi calculado com a Equação 4 a seguir:

$$Wq = \frac{Lq}{\lambda} \quad (4)$$

$$= \frac{0,05536960861}{3,6508} \approx 0,0152$$

Ou aproximadamente 0,91 segundos, se multiplicar por 60. Quase não há tempo de espera, o processo de atendimento é iniciado por um dos três servidores quase sempre no mesmo instante que o cliente chega ao sistema.

Pelo lado dos clientes, essa é uma situação ideal, já que seriam prontamente atendidos ao chegarem ao sistema. Já na ótica da gestão, pode significar que os recursos direcionados para o atendimento primário poderiam ser realocados para outras áreas caso haja necessidade.

O número médio de clientes no sistema foi obtido através da Equação 5 a seguir:

$$L = Lq + \frac{\lambda}{\mu} \quad (5)$$

$$= 0,05536960861 + \frac{3,6508}{3,4158} \approx 1,1241$$

Significa dizer que o sistema analisado está quase sempre com aproximadamente 1,12 clientes, incluindo tanto os que estão em atendimento quanto os da fila.

O tempo médio de espera no sistema utilizou a Equação 6 a seguir:

$$W = Wq + \frac{1}{\mu} \quad (6)$$

$$= 0,01516643163 + \frac{1}{3,4158} \approx 0,3079236188$$

$W = 0,3079236188$ horas, e ao multiplicar por 60 obtemos seu valor = 17,56 segundos.

Evidencia que o sistema possui atendimento rápido, em que em algumas das situações observadas no dia da pesquisa, o comparecimento ao atendimento se deu por questões burocráticas, como a entrega de algum documento para anexar ao processo. A DPU não se opõe ao atendimento virtual, entretanto, como informado no tópico 3.1 deste trabalho, a organização atende a população em hipossuficiência, e alguns dos assistidos podem não ter familiaridade ou confiança em sistemas virtuais.

Em suma, a DPU atualmente goza de recursos para atender a demanda de atendimento presencial em dias típicos. Caso a instituição opte pela redução de custos, poderiam focar no ganho de eficiência ao reduzir a capacidade ociosa, com a realocação de pessoal para áreas críticas, como a minuta e acompanhamento dos processos judiciais realizado pelos cartórios da própria autarquia ou realizando rodízios dos servidores para que possam contornar a oscilação da capacidade durante o período de férias sem afetar o serviço.

5. CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo a investigação da dinâmica da fila do atendimento na unidade da Defensoria Pública da União em João Pessoa-PB através da Teoria das Filas e contribuiu, através da demonstração da sua aplicação no atendimento do serviço público, com indicadores que poderão ser utilizados para embasar decisões gerenciais e comparativos com estudos futuros em outras unidades.

Apesar de ser apresentada uma ociosidade relativamente alta, é importante ressaltar que o sistema investigado está sujeito a sazonalidades na demanda que não são previsíveis, como por exemplo, quando algum entendimento é decidido nos Supremos Tribunais ou quando as prefeituras de cidades da região enviam caravanas para a capital com intuito de seus cidadãos resolverem pendências, são variáveis que sobrecarregam os servidores em dias atípicos. Devido ao nível de acesso permitido pelo tipo de organização estudada foi possível catalogar apenas um dia comum de atendimento, não sendo possível analisar o padrão de longo prazo do sistema, sendo esta a limitação identificada.

Entre as sugestões, poderiam ser realizadas observações em dias distintos, investigando dias atípicos com muito ou pouco movimento, levando em consideração sazonalidades semanais, mensais e anuais (como clima e eventos atípicos), acrescentar simulações computadorizadas para testar diferentes cenários e analisar o custo-benefício da diminuição ou redistribuição dos servidores.

Através dessa pesquisa pude aplicar na prática o conhecimento e ferramentas aprendidas ao longo da graduação em um local onde estagiei e tive vivência, e poderei devolver para a sociedade o investimento.

Por fim, espero que esse estudo sirva de apoio para que a Defensoria Pública da União possa aliar teoria e prática em seus processos e que continue garantindo agilidade e qualidade nos atendimentos aos cidadãos.

REFERÊNCIAS

ANDRADE, E. L. Introdução à pesquisa operacional: métodos e modelos para análise de decisões. Rio de Janeiro: LTC, 2002.

BRASIL. Decreto-Lei nº 200, de 25 de fevereiro de 1967. Dispõe sobre a organização da Administração Federal, estabelece diretrizes para a Reforma Administrativa e dá outras providências. **Diário Oficial da União**: Seção 1, Suplemento, Brasília, 27 fev. 1967. Disponível em: https://www.planalto.gov.br/ccivil_03/decreto-lei/del0200.htm. Acesso em: 12 ago. 2025.

CORRÊA, H. L.; CORRÊA, C. A. Administração de produção e operações: manufatura e serviços: uma abordagem estratégica. 5. ed. São Paulo: Atlas, 2022. 480 p.

DPE-PB - DEFENSORIA PÚBLICA DO ESTADO DA PARAÍBA. A Defensoria Pública. João Pessoa: Defensoria Pública do Estado da Paraíba, 2025. Disponível em: <https://defensoria.pb.def.br/a-defensoria-publica/>. Acesso em: 12 dez. 2024.

GUPTA, S.; MODGIL, S.; BHATTACHARYYA, S. *et al.* Artificial intelligence for decision support systems in the field of operations research: review and future scope of research. **Ann Oper Res** 308, 215–274 (2022). <https://doi.org/10.1007/s10479-020-03856-6>

HILLIER, F. S.; LIEBERMAN, G. J. Introdução à Pesquisa Operacional. 9. ed. Porto Alegre: AMGH Editora Ltda, 2013.

IFORS - INTERNATIONAL FEDERATION OF OPERATIONAL RESEARCH SOCIETIES. What is Operational Research? c2020. Disponível em: <https://www.ifors.org/what-is-or/>. Acesso em: 09 dez. 2024.

JIAYI, J. Application of Queuing Theory in Public Service Counters. ICIAAI 2023. P. Kar *et al.* (eds.), **Proceedings of the 2023 International Conference on Image, Algorithms and Artificial Intelligence** (ICIAAI 2023), Advances in Computer Science Research 108, https://doi.org/10.2991/978-94-6463-300-9_52

JUSBRASIL. Tempo de espera no banco: o que fazer em caso de demora. Jusbrasil. Disponível em: <https://www.jusbrasil.com.br/artigos/tempo-de-espera-no-banco-o-que-fazer-em-caso-de-demora/173949669>. Acesso em: 25 ago. 2025.

LEONELLI, M. Simulation and Modelling to Understand Change. Madrid: School of Human Sciences and Technology, IE University, 2021. Disponível em: https://bookdown.org/manuele_leonelli/SimBook/. Acesso em: 08 ago. 2025.

MELO, M. F. M. de. Aplicação de programação linear através do Solver em uma microempresa de serviços de beleza. 2020. Trabalho de Conclusão de Curso (Bacharelado em Administração) — Universidade Federal da Paraíba, Centro de Ciências Sociais Aplicadas, João Pessoa, 2020. Disponível em: <https://repositorio.ufpb.br/jspui/bitstream/123456789/19561/1/MFMM22022021.pdf>. Acesso em: 03 ago. 2025.

PRADO, F. Atendimento prioritário: veja quem deve furar filas e se sentar em ônibus. Florianópolis: PROCON-SC, 2024. Disponível em: <https://www.procon.sc.gov.br/atendimento-prioritario-veja-quem-deve-furar-filas-e-sentar-em-onibus/>. Acesso em: 11 mar. 2025.

RAMOS, P. A.; MELO, M. F. M. de; GONÇALVES, H. S. Aplicando pesquisa operacional no contexto de uma microempresa de serviços de beleza. **Revista Produção Online**, Florianópolis, v. 22, n. 1, p. 2209–2234, jan. 2023. DOI: 10.14488/1676-1901.v22i1.4401. Disponível em: <https://www.producaoonline.org.br/rpo/article/view/4401>. Acesso em: 12 ago. 2025

RAMOS, P. A.; PEREIRA, F. Use of Transportation Problem in a Water Company. 29th IJCIEOM – International Joint Conference on Industrial Engineering and Operations Management, Lusófona University, Lisbon, Portugal, 2023. Disponível em: https://portal.abepro.org.br/ijcieom/restrito/arquivos/icieom2023/FULL_0044_37771.pdf. Acesso em: 12 ago. 2025.

SANTOS, M. P. A. Por que a fila não anda?: uma casa lotérica em tempos de pandemia de COVID-19. 2024. Caso para Ensino (Trabalho de Curso de Graduação em Administração) — Universidade Federal da Paraíba, Centro de Ciências Sociais Aplicadas, João Pessoa, 19 abr. 2024. Disponível em: <https://repositorio.ufpb.br/jspui/handle/123456789/33415>. Acesso em: 25 ago. 2025.

VIRGILLITO, S. B. Pesquisa operacional: métodos de modelagem quantitativa para a tomada de decisões. 1. ed. São Paulo: Saraiva, 2018. 366 p.