



**UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS I
CENTRO DE CIÊNCIAS E TECNOLOGIAS
DEPARTAMENTO DE ESTATÍSTICA
CURSO DE ESTATÍSTICA**

AMAURI DOS SANTOS LIMA NETO

**CADASTRO NACIONAL DE OBRAS: UMA APLICAÇÃO DA ANÁLISE
DE SOBREVIVÊNCIA.**

**CAMPINA GRANDE
2024**

1 Resumo

Este trabalho apresenta uma aplicação da análise de sobrevivência para avaliar o tempo de conclusão das obras públicas na Paraíba, utilizando dados do Cadastro Nacional de Obras (CNO). Com foco na melhoria da infraestrutura e desenvolvimento regional, a pesquisa utiliza o estimador de Kaplan-Meier para analisar 3.327 observações, revelando que a maioria das obras é concluída nos primeiros 2.500 dias, com um tempo mediano de conclusão de 177 dias. Os resultados indicam que a probabilidade de conclusão diminui com o aumento do tempo de execução, e não há diferença significativa entre obras regidas pelos próprios donos e terceirizadas sugerindo a necessidade de mais variáveis e uma abordagem mais complexa para otimizar recursos e melhorar a gestão e transparência das obras públicas.

2 Introdução

As obras públicas desempenham um papel fundamental no desenvolvimento econômico e social de qualquer região. Elas abrangem uma ampla gama de projetos, desde a construção de estradas, pontes e aeroportos até hospitais, escolas e sistemas de saneamento básico. A importância dessas obras reside em sua capacidade de melhorar a qualidade de vida da população, facilitar o transporte e a mobilidade, desenvolver a economia local e promover a inclusão social.

Na Paraíba, um estado situado na Região Nordeste do Brasil, as obras públicas podem ter um impacto significativo na modernização da infraestrutura e na promoção do desenvolvimento sustentável. Investimentos em infraestrutura são essenciais para superar desafios históricos de desenvolvimento e para aumentar a competitividade do estado em nível nacional e internacional. Além disso, tais obras podem contribuir para a redução das desigualdades regionais, impulsionando o crescimento econômico em áreas menos desenvolvidas.

A duração das obras públicas é um indicador preciso para avaliar o progresso no desenvolvimento da infraestrutura (4). O tempo de conclusão de projetos pode afetar diretamente a eficiência e a eficácia das políticas públicas. Obras que são concluídas dentro do prazo e do orçamento previsto indicam uma gestão eficaz e uma alocação eficiente de recursos, enquanto atrasos podem resultar em custos adicionais, transtornos para a população e um impacto negativo na credibilidade das instituições responsáveis.

No caso da Paraíba, monitorar e entender o tempo de execução das obras públicas é essencial para garantir que os benefícios esperados sejam entregues à população de forma apropriada. Isso inclui a melhoria das condições de vida, a geração de empregos e a atração de investimentos. Além disso, a transparência e a responsabilidade na execução dessas obras são importantes para fortalecer a confiança pública e garantir que os recursos sejam utilizados de maneira justa e eficiente.

3 Problemas e objetivos

Diante disso, a importância das obras públicas e a gestão do tempo de execução dessas obras são aspectos críticos para o desenvolvimento sustentável e a melhoria da infraestrutura na Paraíba, contribuindo para o bem-estar geral da sociedade e para o progresso econômico do estado.

Será utilizado o banco de dados(3) que trás informações sobre as obras do governo cadastradas no Cadastro Nacional de Obras (CNO) entre os anos de 1900 até 2021 e é composto por 24 variáveis das quais só serão utilizadas 3, e apenas para o estado da Paraíba:

- **data_inicio**
- **situacao**
- **data_situacao**

Com essas variáveis teremos acesso ao tempo em que a obra demorou para ser finalizada ou interrompida de alguma forma, desse modo nosso evento de interesse é o término da obra, enquanto as censuras serão obras paralisadas ou suspensas. Assim Aplicaremos métodos da análise de sobrevivência nos dados tratados para estimar a função de sobrevivência e conseguiremos algumas informações tais como: o tempo médio até a conclusão de uma obra pública, a probabilidade de que ela seja concluída em n meses e qual período a obra tem maior probabilidade de ser finalizada. Portanto o objetivo geral será extrair informações sobre as obras públicas afim de entender o comportamento probabilístico da curva de sobrevivência por meio do estimador de kaplan-Meier.

4 Metodologia

A análise de sobrevivência é um conjunto de técnicas estatísticas utilizadas para analisar dados onde o resultado de interesse é o tempo até a ocorrência de um evento específico. Esses eventos podem ser de diversos tipos, como a morte de um paciente, falha de um componente, recuperação de uma doença, entre outros.

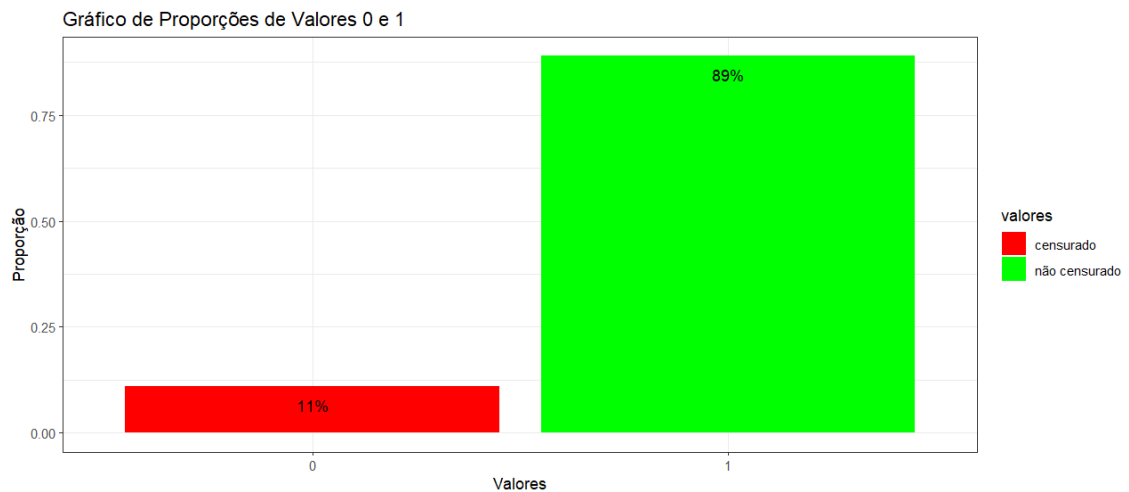
Neste trabalho será utilizado o estimador para a função de sobrevivência de Kaplan-Meier proposto por Kaplan e Meier em 1958 que é dado pela seguinte expressão.(2)

$$\hat{S}(t) = \frac{\text{no. de observações que não falharam até o tempo } t}{\text{no. total de observações no estudo}}$$

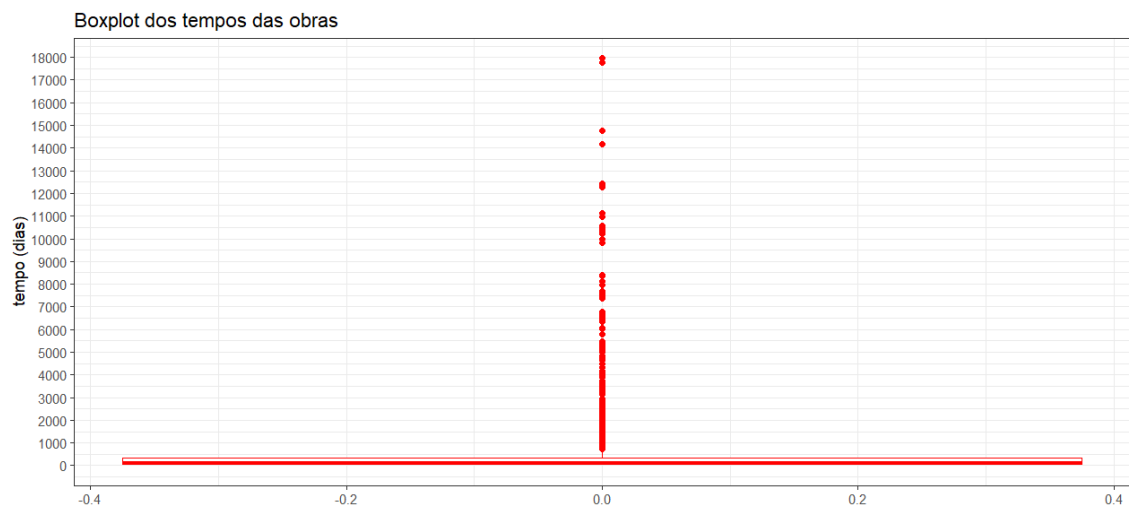
em que $\hat{S}(t)$ é uma função escada com degraus nos tempos observados de falha de tamanho $1/n$, e n é o tamanho da amostra. Se existirem empates em um certo tempo t , o tamanho do degrau fica multiplicado pelo número de empates e na sua construção, considera tantos intervalos de tempo quantos forem os números de falhas distintas. Os limites dos intervalos de tempo são os tempos de falha da amostra.

5 Resultados

Desta forma, foi criada uma coluna chamada status indicando se a observação sofreu ou não censura de acordo com as obras no estado da Paraíba, totalizando 3327 observações.



O gráfico acima mostra a relação das proporções do status das observações evidenciando o percentual de censuras nos dados.



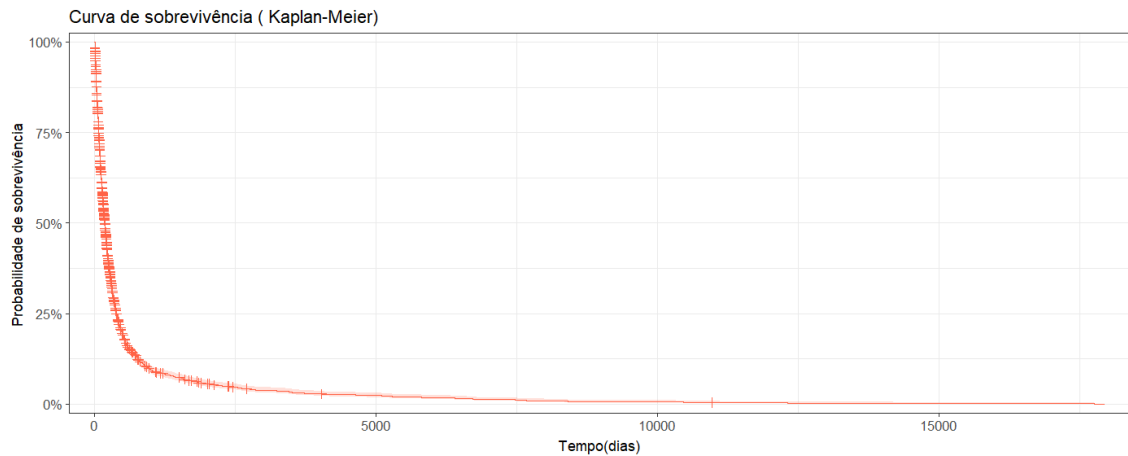
O Boxplot mostra uma grande assimetria na distribuição dos tempos das obras e evidência a presença de outliers.

Tabela de Dados de Sobrevivência

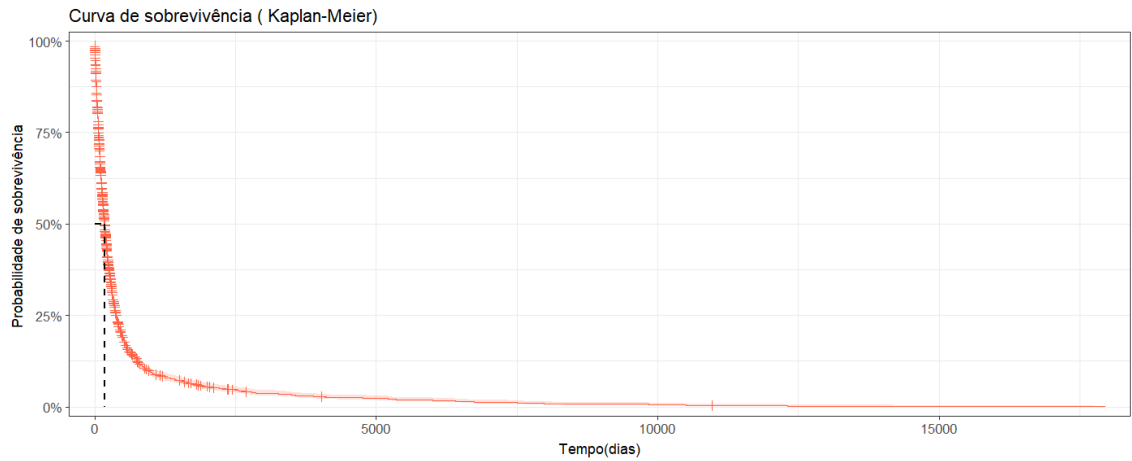
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
1	3327	53	0.9841	0.00217	0.9798	0.9883
10	3122	97	0.9544	0.00364	0.9473	0.9616
50	2615	475	0.8081	0.00690	0.7947	0.8218
100	2068	504	0.6506	0.00840	0.6344	0.6673
500	521	1359	0.1911	0.00729	0.1774	0.2060
1000	240	254	0.0948	0.00560	0.0844	0.1065
1500	182	54	0.0732	0.00504	0.0640	0.0838
2000	132	42	0.0560	0.00450	0.0478	0.0656
2500	106	21	0.0469	0.00419	0.0394	0.0559
3000	84	21	0.0376	0.00382	0.0308	0.0459
3327	79	5	0.0354	0.00372	0.0288	0.0435

Tabela 5.1

A tabela acima trás um resumo de informações a respeito do modelo de Kaplan-Meier com os tempos, taxa de risco, estimativas da função de sobrevivência, o desvio padrão e os limites para os respectivos intervalos de confiança. Que serão utilizados para extrair os resultados citados acima.



Observeando o gráfico da curva de sobrevivência, verificou-se um decaimento rápido em seu comportamento, ou seja, a grande maioria das obras tem uma maior probabilidade de ser concluída nos primeiros 2500 dias.



No dia 177, metade das obras já aviam sido concluidas.

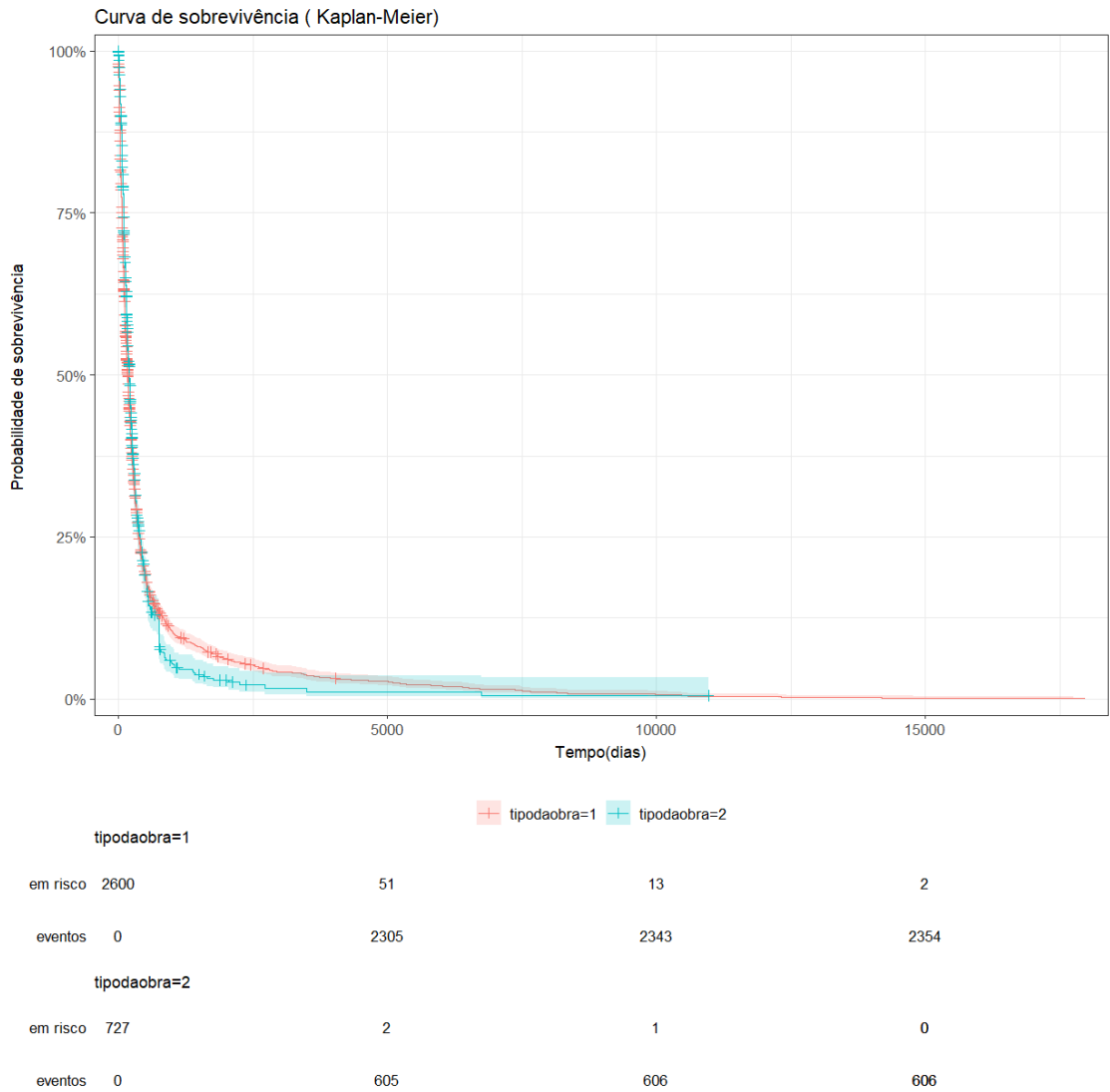
Tabela de Dados sobre o modelo

n	events	rmean*	se(rmean)	median	0.95LCL	0.95UCL
3327	2962	547	28.6	177	166	184

Tabela 5.2

Na tabela acima observa-se algumas informações importantes sobre o modelo de kaplan-Meier onde das 3327 obras 2962 foram concluídas com uma média 547 dias.

Para melhor visualização do contexto da obra, foi criada uma nova variável de acordo com a qualificacao_responsavel onde os códigos 70 e 57 foram interpretadas como obras regidas pelos próprios donos ou do tipo 1, e os demais códigos foram interpretados como obras terceirizadas ou do tipo 2 e novamente foi aplicado o estimador de Kaplan-Meier para estimar as funções de sobrevivência para cada grupo, e foi usado o teste log-rank para verificar se há alguma diferença estatística entre as curvas.



O gráfico acima mostra o comportamento das curvas por grupo, porém ao ser aplicado o teste de log-rank as curvas não mostraram ser estatisticamente diferentes com um p-valor de 0.4813.

6 Conclusão

A análise revelou padrões importantes no tempo de conclusão das obras públicas, identificando fatores críticos que influenciam a duração dos projetos. Através da modelagem de sobrevivência, foi possível observar que quanto mais tempo a obra demora para ser concluída menor sua probabilidade de conclusão e apesar de não ter uma diferença significativa entre obras terceirizadas e próprias é possível destacar que sobre as obras próprias em que a responsabilidade da obra é do proprietário do imóvel ou do dono da obra todas as obras foram finalizadas. Para um maior entendimento geral desse comportamento seria necessário uma maior quantidade de variáveis e uma análise mais complexa, esses achados têm implicações práticas significativas para a gestão e planejamento de obras públicas. Governos e gestores de projetos podem utilizar essas informações para melhorar a alocação de recursos, otimizar processos administrativos, e priorizar intervenções em áreas que apresentam maiores atrasos. também é de grande importancia criar um rigor sobre os dados, pois em diversos casos existem inconsistências a respeito da área onde obras de 0,01 m^2 chegaram a durar 476 dias.

Tabela com dados inconsistentes

nome empresarial	unidade medida	área	diferença dias	status	tipodoadora
Empresa A	m^2	00,1	450	1	2
Empresa A	m^2	00,1	476	1	2
Empresa A	m^2	00,1	304	1	2
Empresa B	Outra	00,1	241	1	1
Empresa C	Outra	00,1	1581	1	2
Empresa D	Outra	00,1	20	0	1
Empresa A	m^2	00,1	90	1	2

Tabela 6.1

REFERÊNCIAS

- COLOSIMO, E. A., AND GIOLO, S. R. *Análise de sobrevivência aplicada*. Editora Blucher, 2021.
- DAHIS, R., CARABETTA, J., SCOVINO, F., ISRAEL, F., AND OLIVEIRA, D. Data basis (base dos dados): Universalizing access to high-quality data. *Available at SSRN 4157813* (2022).
- GUIMARÃES, I. F. G., DO VALE, C. M., ET AL. Causas de atrasos na conclusão de obras públicas: uma análise comparativa por pesquisa bibliométrica e estudo de caso. *Gestão e Projetos: GeP 14*, 1 (2023), 190–218.

.1 Anexo

Variáveis:

- **id_pais:** Contém o código do país onde está localizada a obra e é representado por três dígitos numéricos.
- **sigla_uf:** Contém a sigla do estado onde a obra está localizada ou o nome da cidade se a obra for localizada em outro país.
- **id_municipio:** Contém o código de referência de cada município segundo o IBGE e é composto por 7 dígitos.
- **id_municipio_rf:** Contém o antigo código de referência do IBGE e varia de 1 a 9997.
- **cep:** Identifica o CEP de onde a obra está localizada.
- **id_cno:** Contém o código de registro no sistema nacional de obras e é composto por 12 dígitos.
- **id_cno_vinculado:** Contém o código da obra vinculada, se ela for vinculada.
- **data_registro:** Contém a data que a obra foi registrada no sistema.
- **data_inicio:** Representa a data em que a obra foi iniciada.
- **data_inicio_responsabilidade:** Representa a data de registro no responsável pela obra.
- **situacao:** Representa a situação atual da obra e assume os seguintes valores: 01 - NULA, 02 - ATIVA, 03 - SUSPENSA, 14 - PARALISADA e 15 - ENCERRADA
- **data_situacao:** Representa a data em que a situação da obra foi atualizada pela última vez.
- **tipo_logradouro:** Representa o tipo de logradouro onde a obra está localizada. ex: rua, avenida e entre outros.
- **logradouro:** Representa o nome do logradouro onde a obra está localizada.
- **numero_logradouro:** Representa o número de referência do logradouro onde a obra está localizada.
- **bairro:** Contém o nome do bairro onde a obra está localizada.
- **caixa_postal:** Representa o código de caixa postal da obra, caso tenha.

- **complemento:** Contém o complemento de endereço da obra caso necessário.
- **ni_responsavel:** Número de identificação do responsável pela obra (Se for CPF o campo virá em branco).
- **qualificacao_responsavel:** Qualificação do responsável pela obra:
 - 0070 - Proprietário do Imóvel
 - 0057 - Dono da Obra
 - 0064 - Incorporador de Construção Civil
 - 0053 - Pessoa Jurídica Construtora
 - 0111 - Sociedade Líder de Consórcio
 - 0109 - Consórcio
 - 0110 - Construção em nome coletivo
- **nome_responsavel:** Representa o nome do responsável pela obra.
- **nome_empresarial:** Representa o nome da empresa responsável pela obra.
- **unidade_medida:** Representa a unidade de medida da obra.
- **area:** Representa a medida em área da obra.