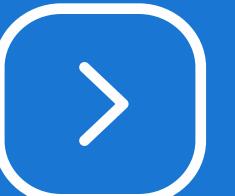


# Análise de Internações Hospitalares do SUS

Período: 12/2017 à 07/2019



**Fonte de Dados:** <http://tabnet.datasus.gov.br/cgi/sih/sxdescr.htm>

**Tratamento de Dados:** Limpeza dos dados, exclusão dos somatórios para não influenciar nas etapas posteriores. Renomeação dos Estados (Eliminação de caracteres especiais), separação de Estados e Regiões, renomeação das colunas.

**Missing:** Durante o tratamento de dados tinha períodos faltantes em 2018: Janeiro, Fevereiro, Junho e Outubro e em 2019: Março e Maio. Ao verificar os dados novamente, no mês de Janeiro de 2019 não tinha o Estado do Acre, onde foi feito a inserção manualmente por se tratar apenas de uma linha.

Para a imputação de dados poderia ter utilizado, Mediana, Média, método de Missing Forest e KNNimputer. Como é o primeiro modelo para o projeto, foi utilizado Média Móvel para imputação dos dados ausentes.

**Feature Engineering:** Conforme o desenvolvimento do projeto foi preciso criar algumas features para trazer mais segurança nos valores finais do modelo. Criação das features de Total de Complemento Federal e Gestor, Prorrogações, Região, Número de Leitos Ocupados.

**Análise Exploratória dos Dados (EDA):** Essa etapa é muito importante para extração de insights para solucionar os problemas relacionados as internações hospitalares do SUS. Durante o projeto foi possível obter mais insights sobre os dados, trazendo uma separação por Estado, Região e durante os períodos analisados.

Foi utilizado a biblioteca Pandas para manipulação e análise de dados e Matplotlib e Seaborn para visualização de dados e criação de gráficos.

**Modelagem de Dados:** No modelo de previsão foi feito teste de estacionaridade de Dickey-Fuller e para as previsões futuras foi utilizado Auto ARIMA.

**Avaliações Finais:** A avaliação final do modelo foi feita comparando os valores previstos dos reais. Conforme as previsões dos Valores Totais de Internações, Número de Óbitos e o Valor Médio AIH, foi possível estimar os valores gastos para os próximos seis meses.

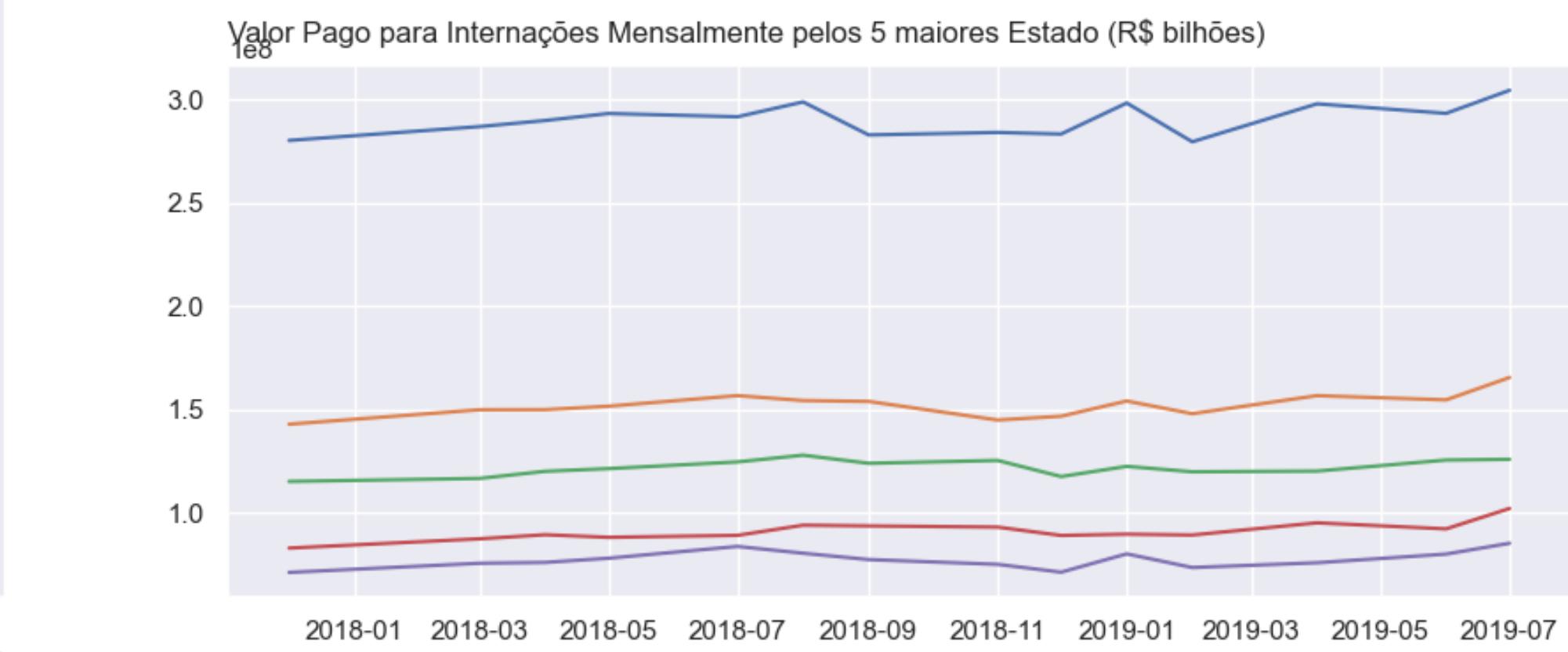
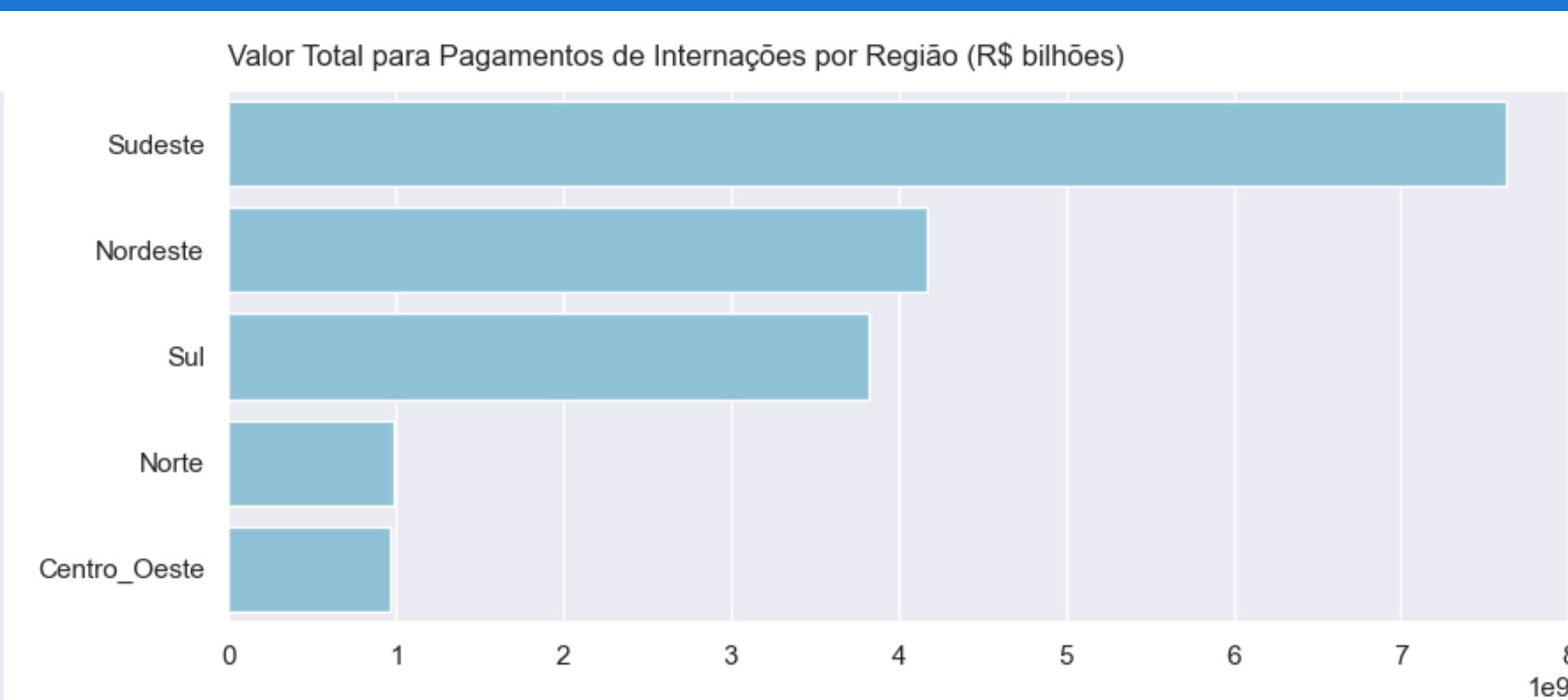
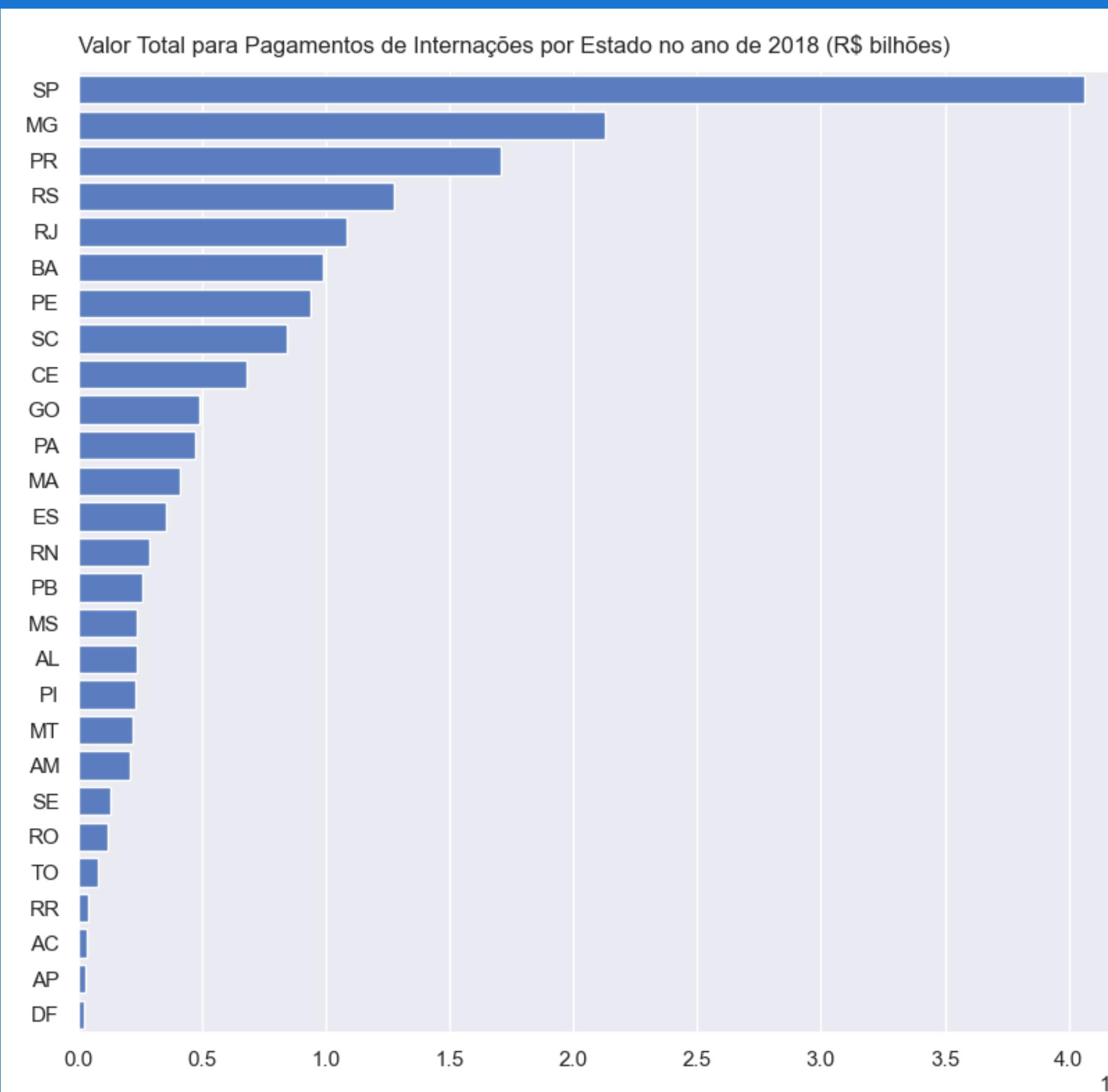


## Valor Total Destinado para Pagamento de Internações

O Sudeste teve o maior valor total pago de investimentos.

Esses valores são perceptíveis quando analisado ao longo do tempo.

O estado de São Paulo recebeu 53,27% do orçamento para internações na região Sudeste.



Considerando que o Sudeste é a região mais populosa do Brasil.

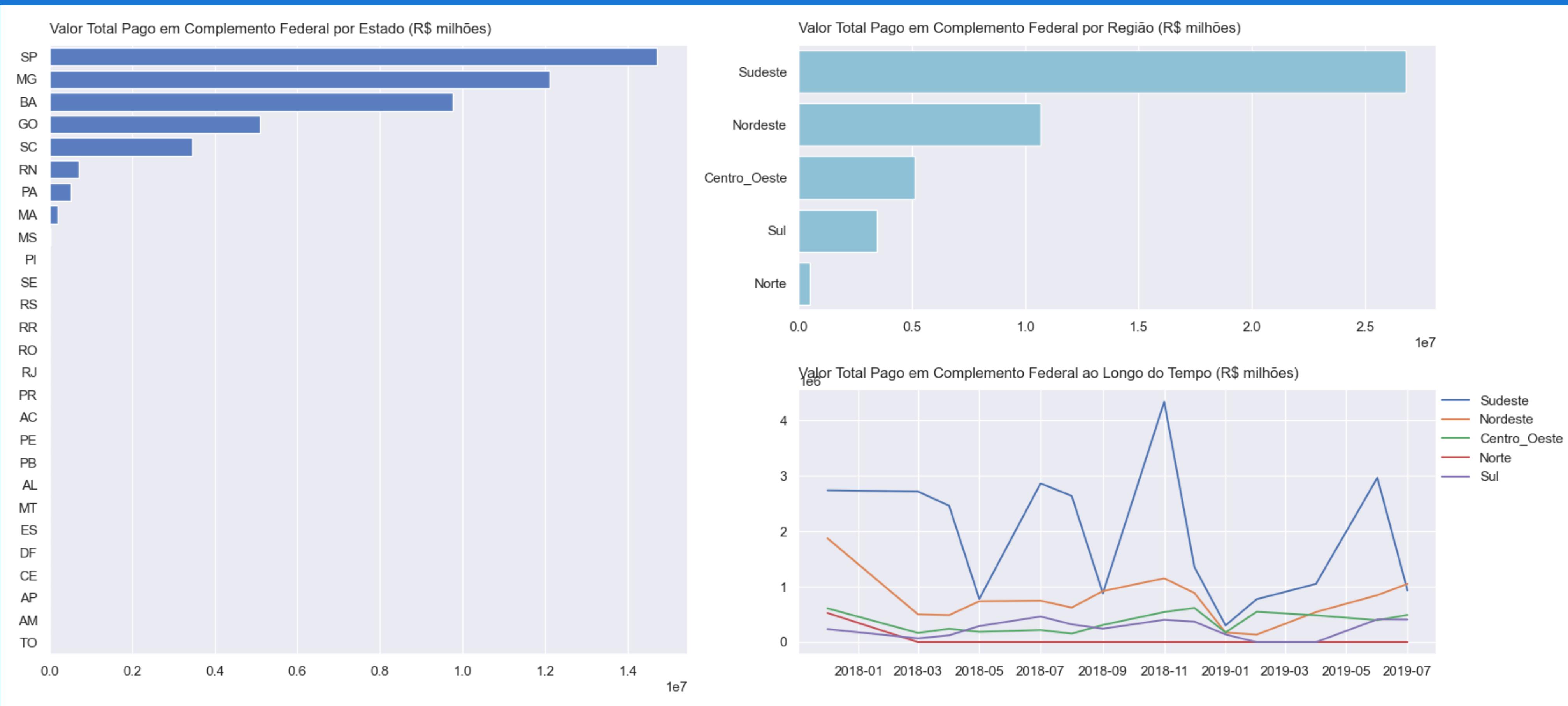
Consequentemente apresenta maiores gastos com internações.



## Valor Total pago em Complemento Federal

Complemento Federal é o somatório dos Serviços Hospitalares + Profissionais

O Sudeste seguido do Nordeste tiveram os maiores repasses do governo em complemento federal.  
As outras regiões possuem um comportamento semelhante.



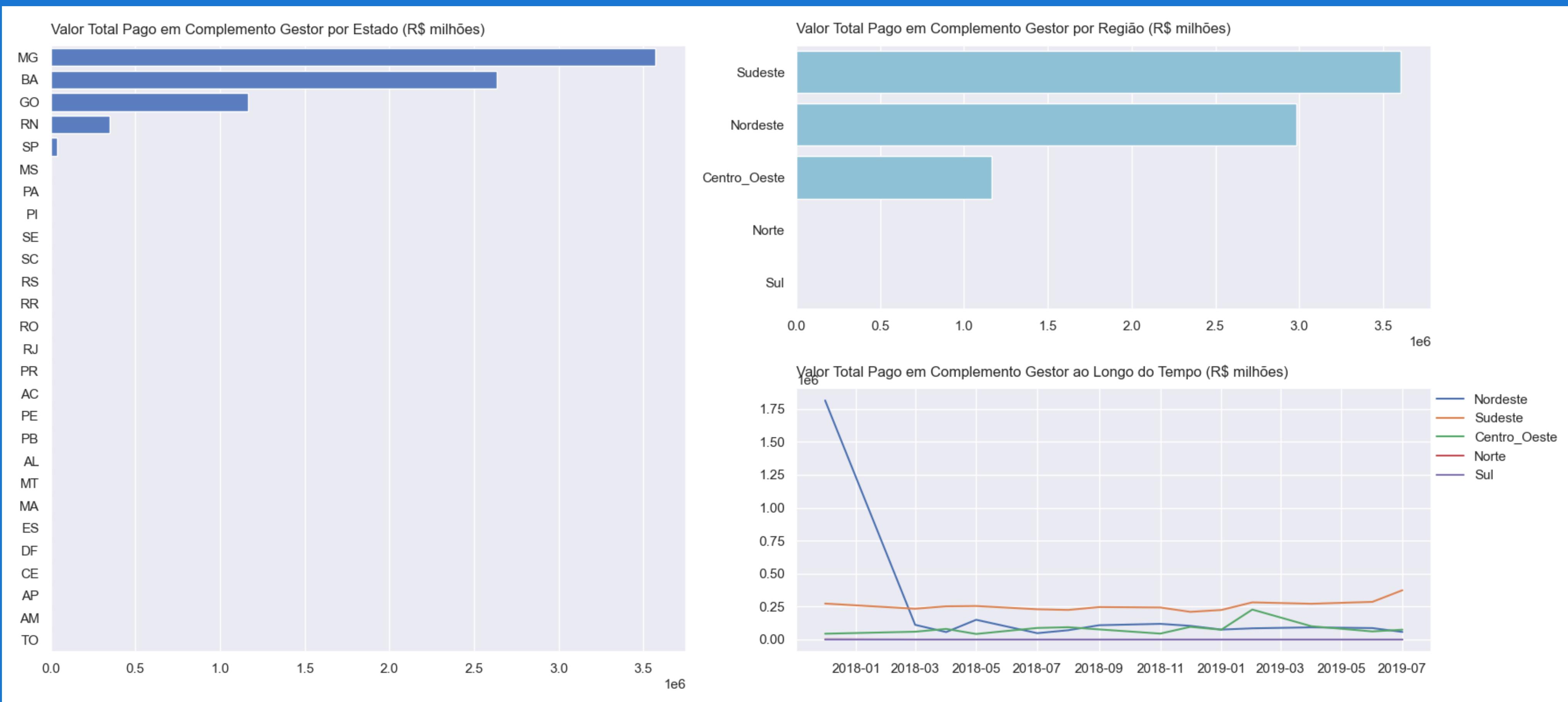
O mês de Agosto/2018 e Junho/2019 teve um aumento elevado em repasses na Região Sudeste.



## Valor Total pago em Complemento Gestor

Complemento Gestor é o somatório dos Serviços Hospitalares + Profissionais

O Sudeste seguido do Nordeste tiveram os maiores repasses pelos gestores.



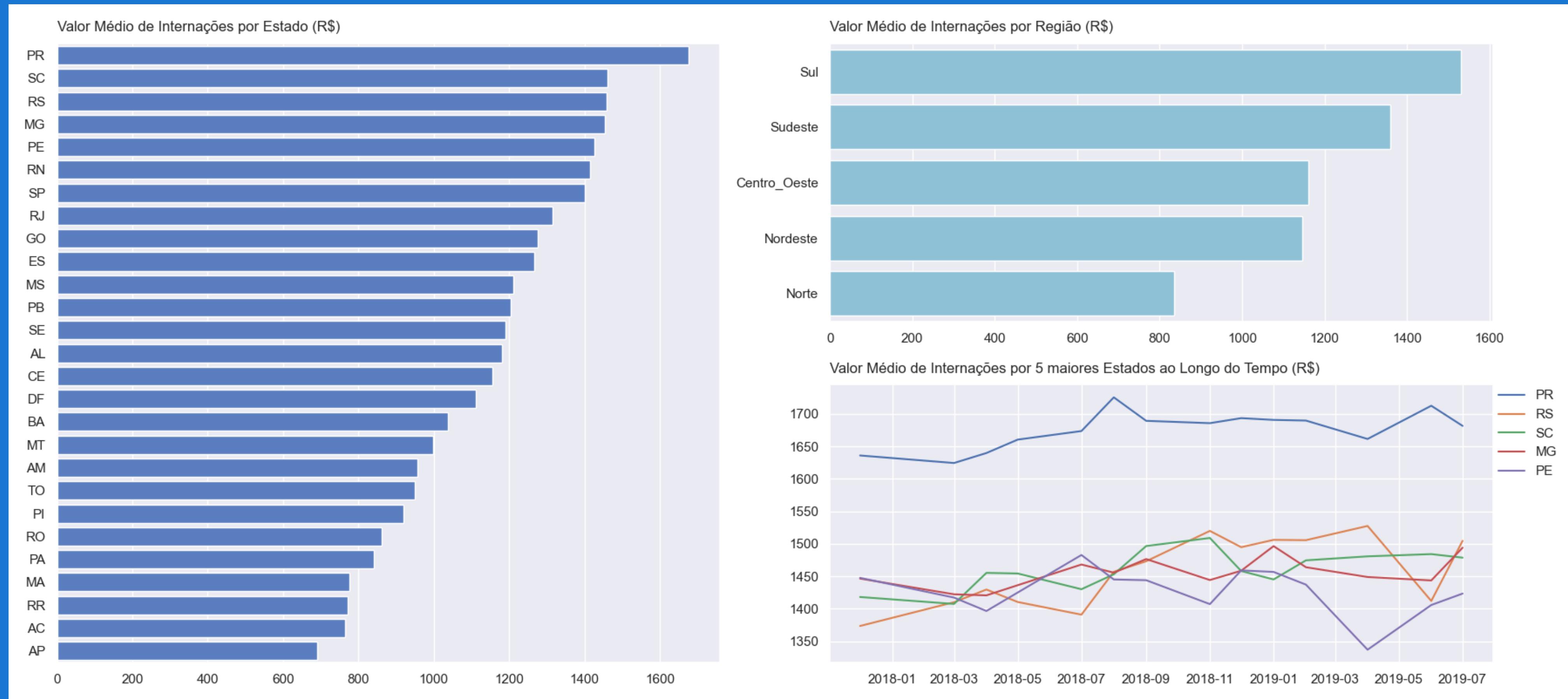
No mês de Dezembro/2017 o Nordeste teve um valor elevado em repasses.

Em seguida houve um comportamento semelhante com as demais regiões ao longo do tempo.



## Valor Médio de Internações

Paraná, Santa Catarina e Rio Grande do Sul tiveram as maiores médias de valores pagos em internações, respectivamente.



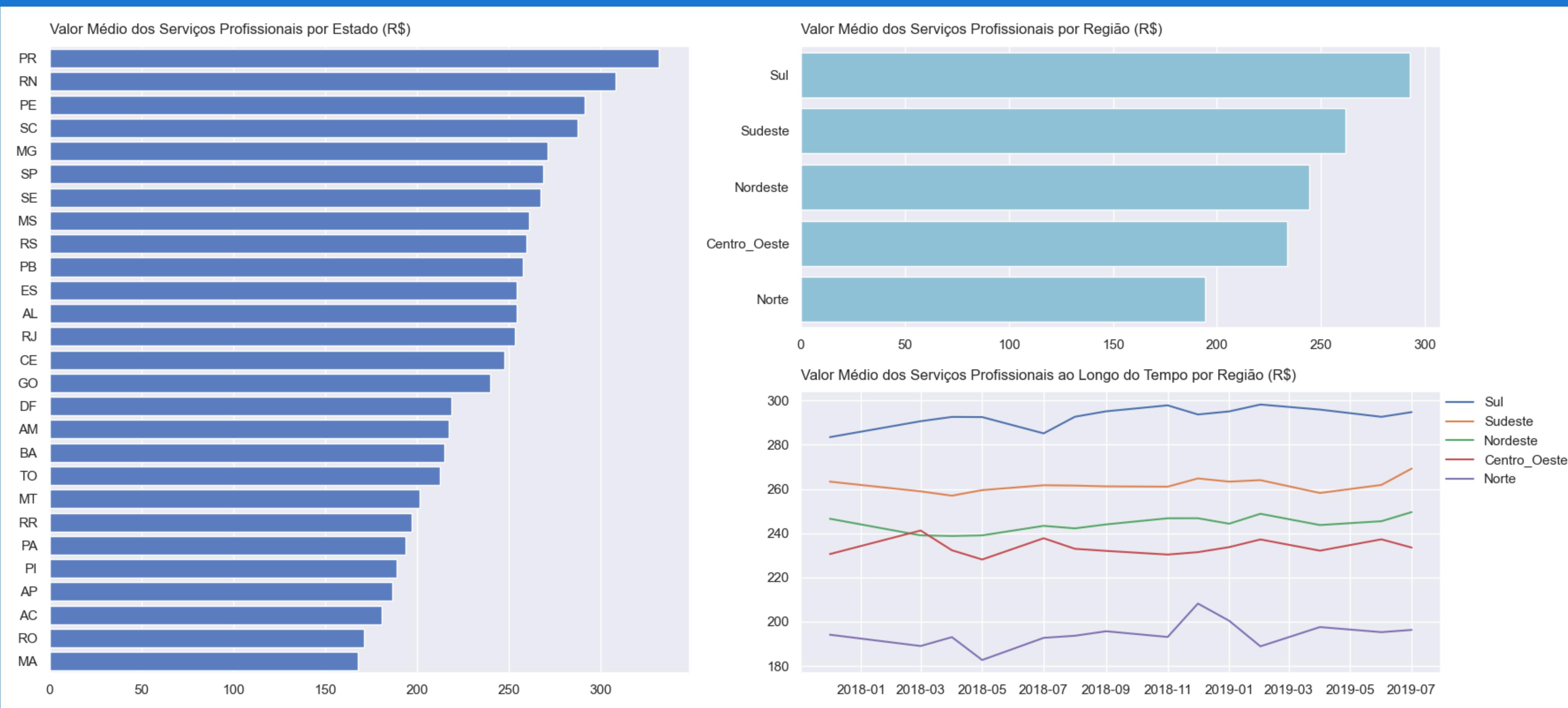
Consequentemente a região Sul teve o maior valor médio em internações.



## Valor Médio dos Serviços Profissionais

A região Sul e Sudeste tiveram as maiores médias de valores pagos em serviços profissionais.

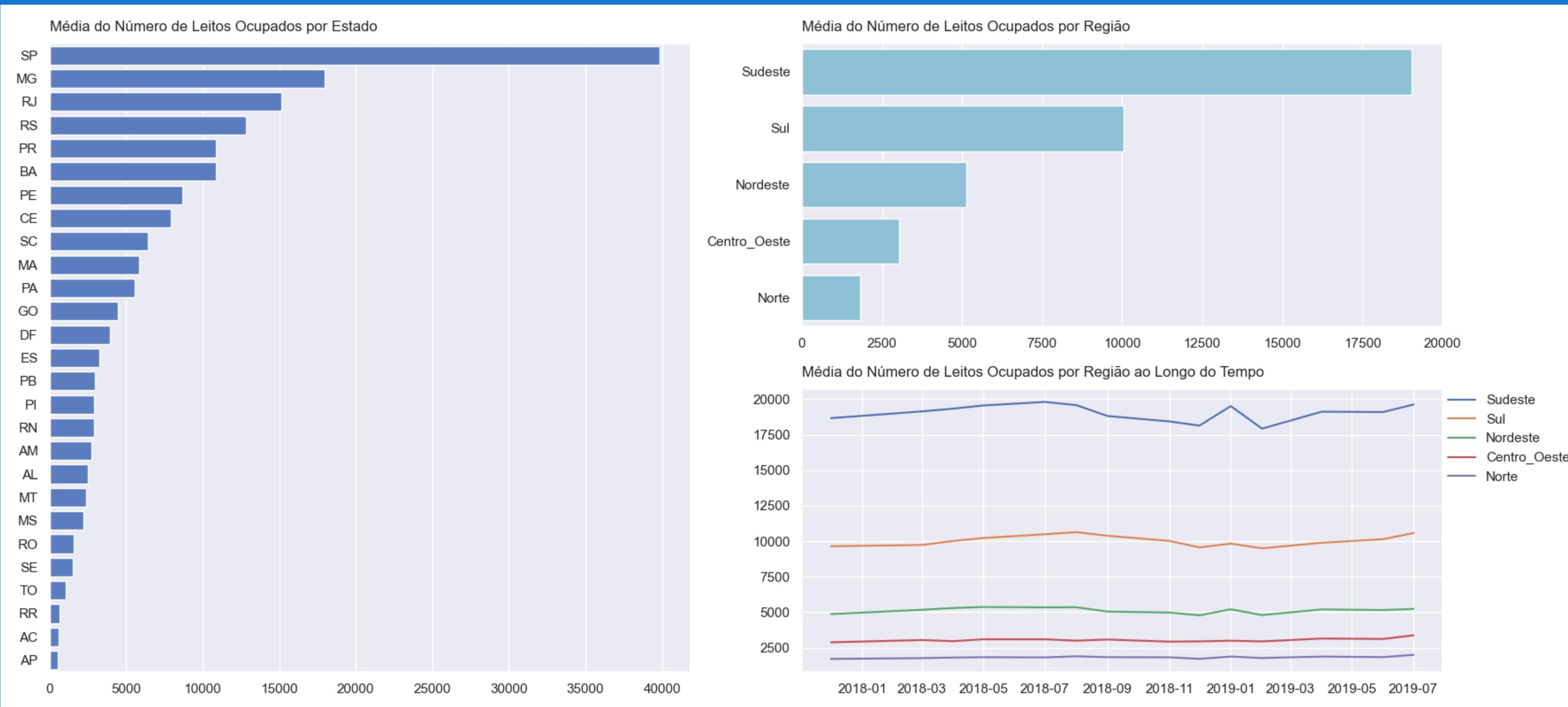
Já o Paraná foi o estado com maior valor médio pago.



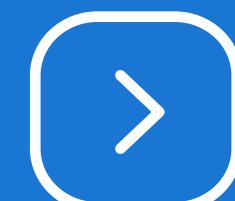
## Média do Número de Leitos Ocupados

A região **Sudeste** teve a maior média de leitos ocupados.

Consequentemente, o estado de **São Paulo, Minas Gerais e Rio de Janeiro** tiveram as **maiores média de leitos ocupados**.



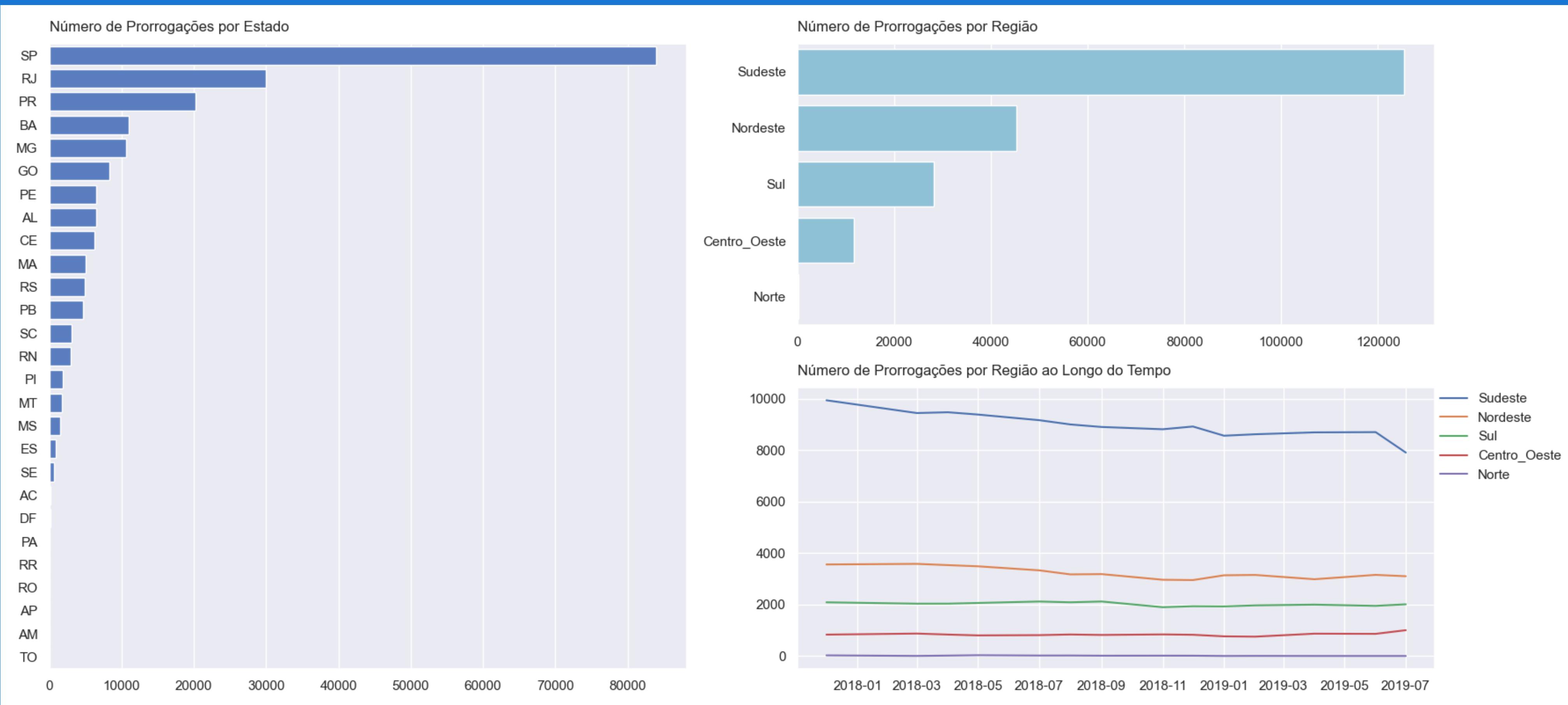
Todas as regiões tiveram um comportamento semelhante ao longo do tempo.



## Número de Prorrogações

O estado de São Paulo apresenta um número elevado de prorrogações.

A região Sudeste teve o maior número de prorrogações de internações.



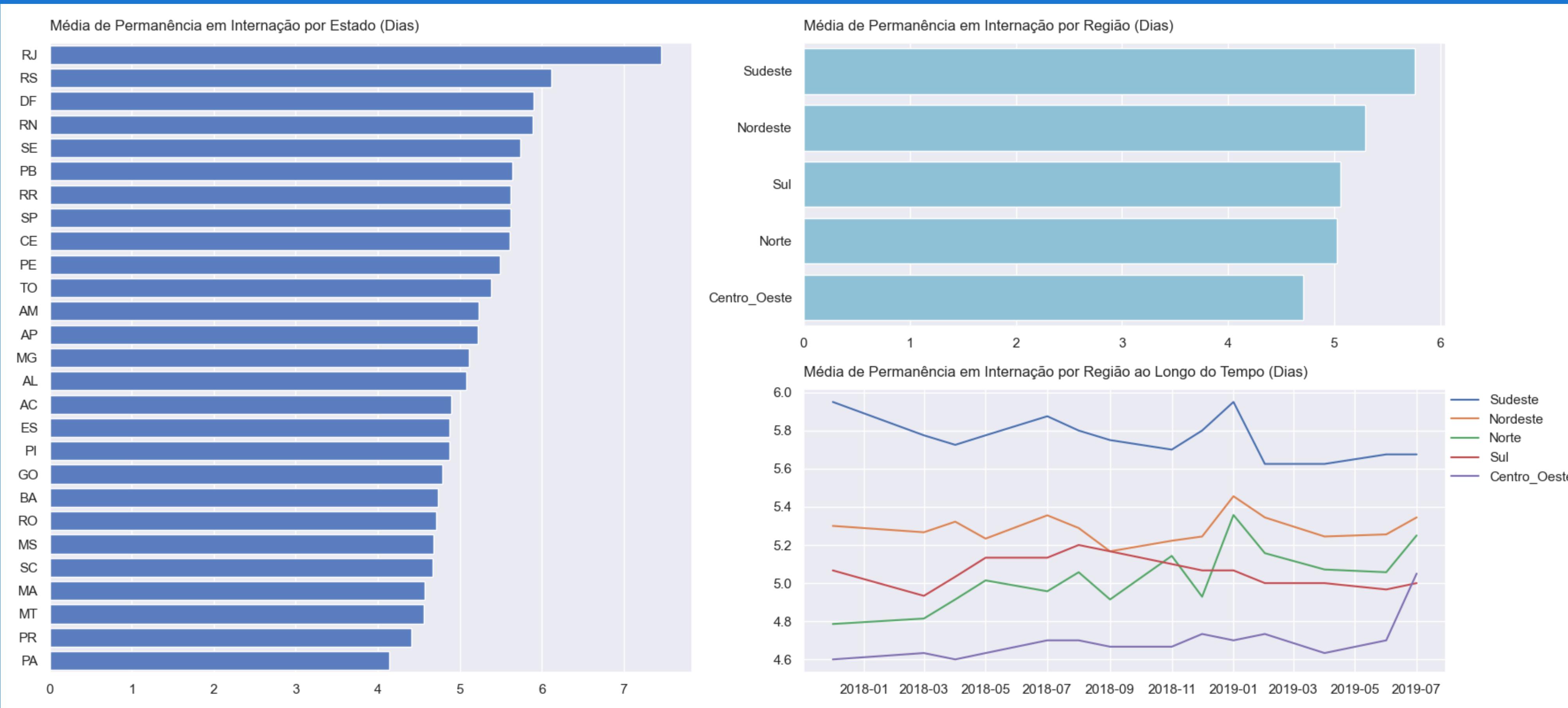
No entanto as demais regiões ao longo do tempo teve um comportamento semelhante no número de prorrogações.



## Média de Permanência em Internações

É perceptível que ao longo do tempo as regiões do Sudeste, Nordeste e Norte tiveram um comportamento parecido no gráfico.

O estado do Rio de Janeiro lidera com a maior média de permanência em internações.



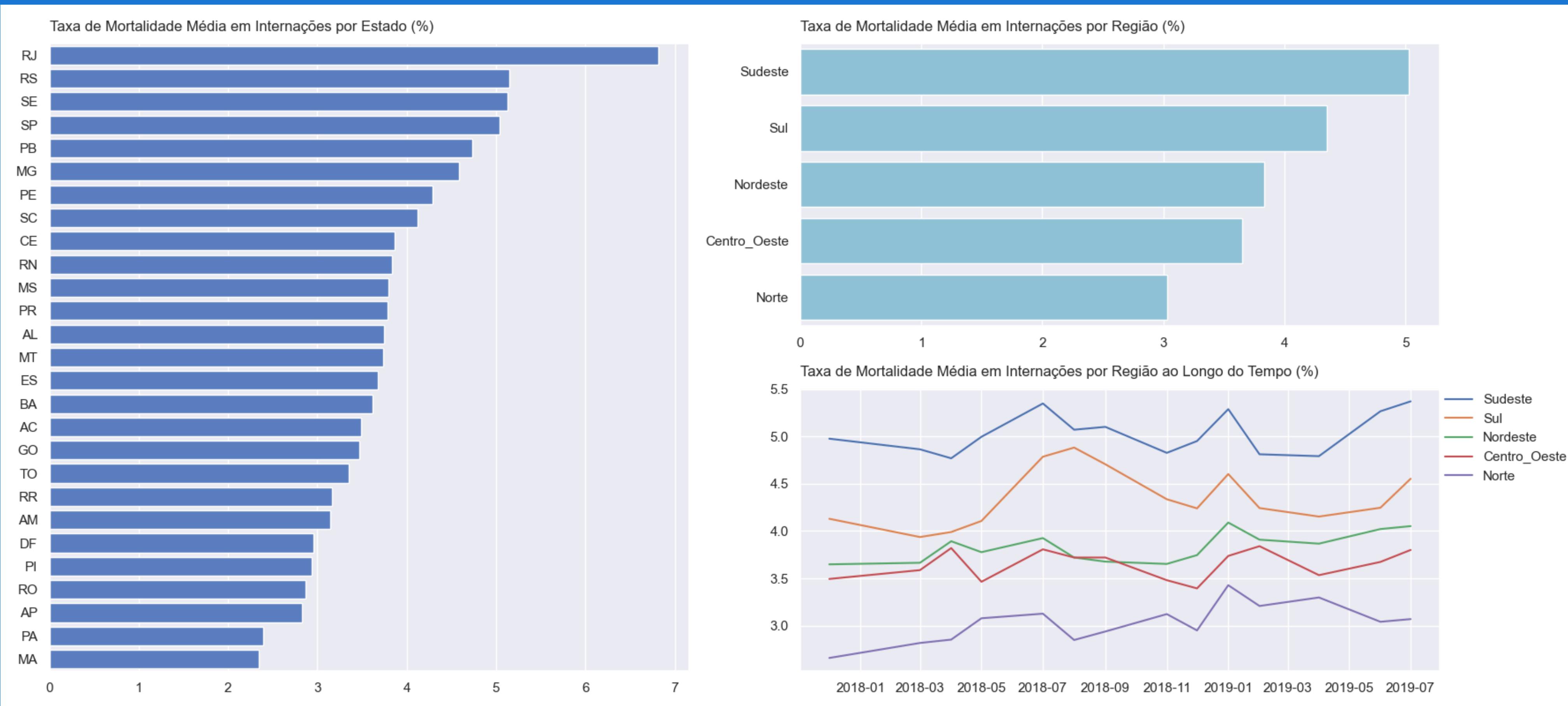
Houve um aumento semelhante no número de permanência em internações no mês de Dezembro/2018 e Janeiro/2019.



## Taxa de Mortalidade Média em Internações

O estado do Rio de Janeiro também lidera na taxa de mortalidade, assim como, na média de permanência nas internações.

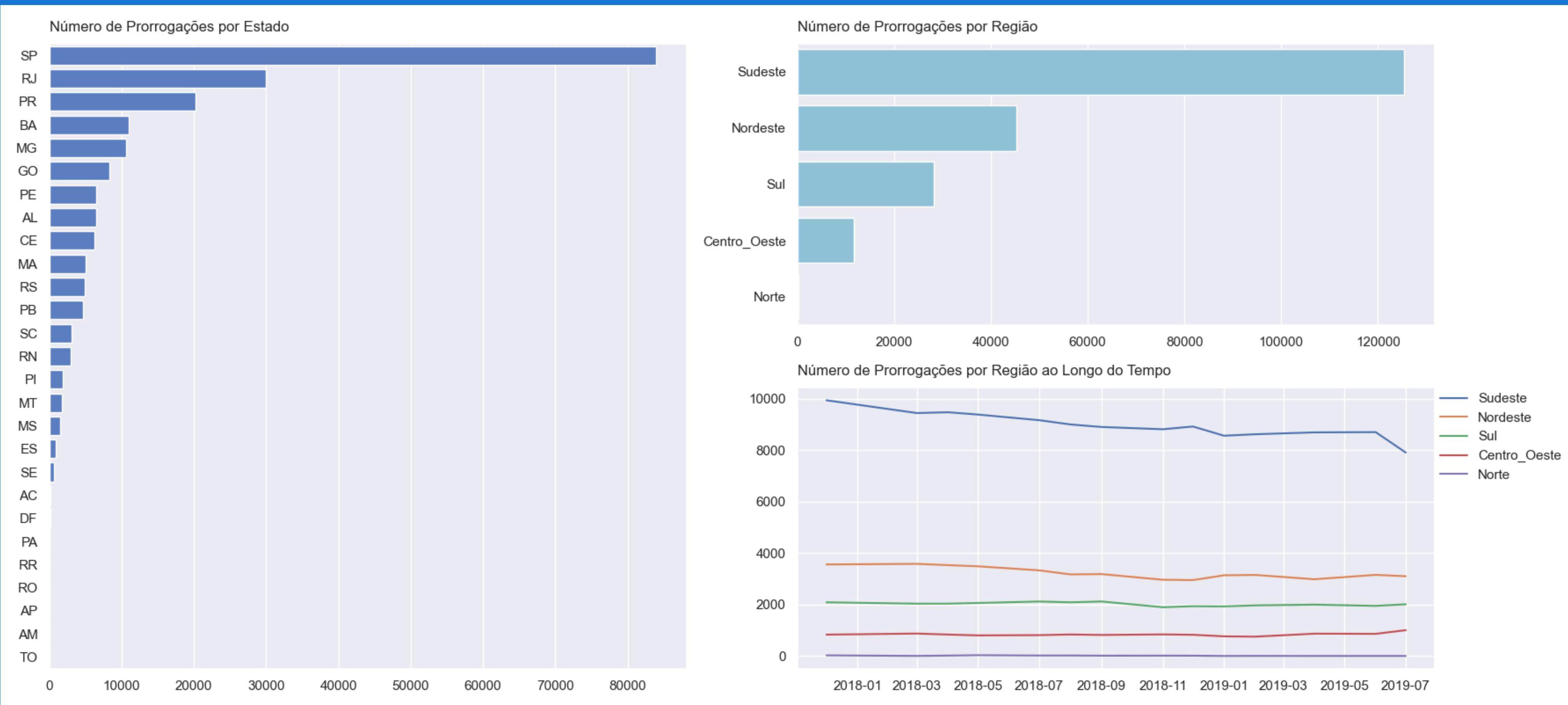
A taxa de mortalidade foi maior nos estados do Sudeste e Sul. Com picos no mês de Julho/2018 e Janeiro/2019.



## Número de Prorrogações

O estado de São Paulo apresenta um número elevado de prorrogações.

A região Sudeste teve o maior número de prorrogações de internações.

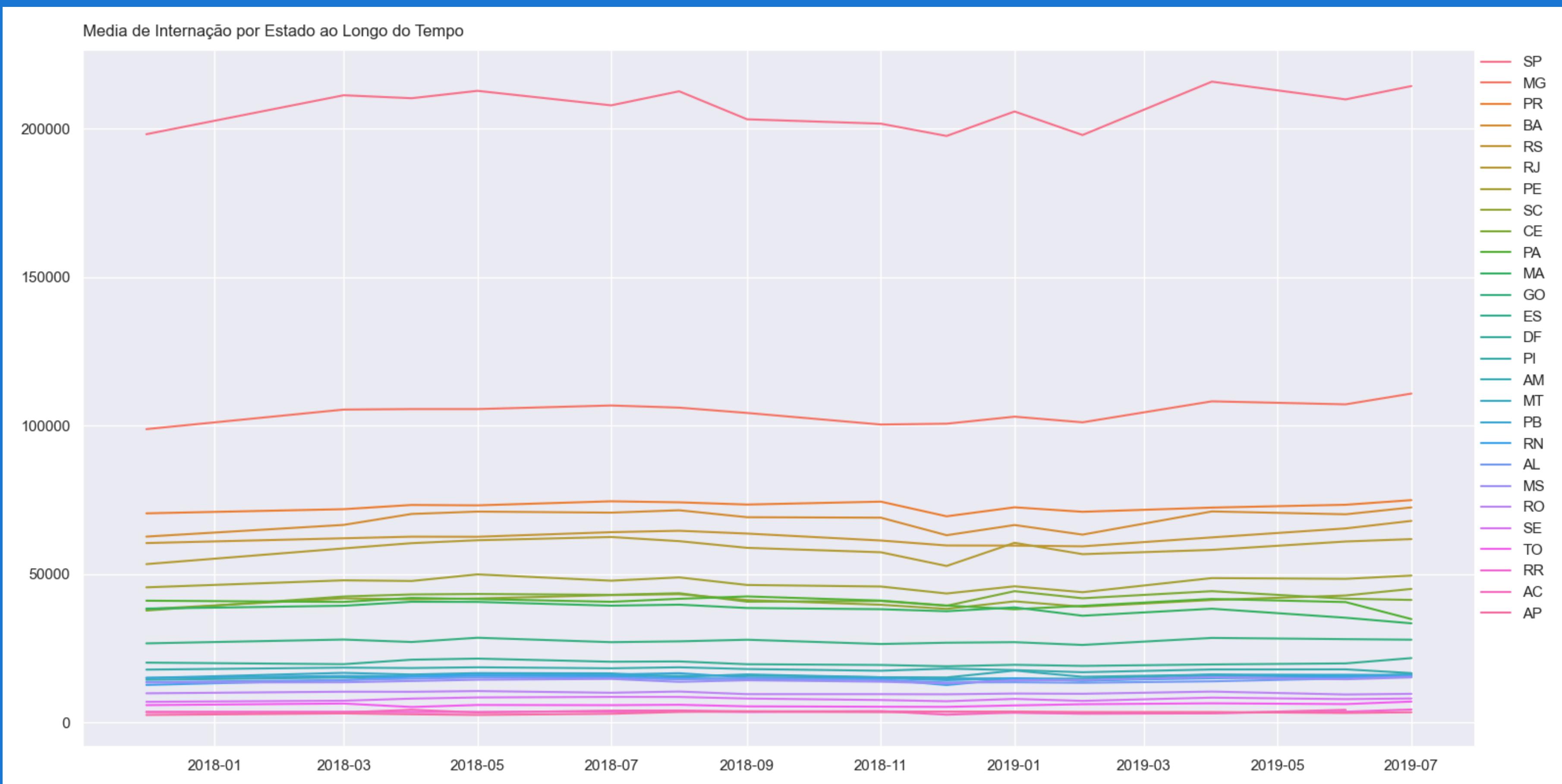


No entanto as demais regiões ao longo do tempo teve um comportamento semelhante no número de prorrogações.



## Número de Internação por Região

O número de internação teve um **comportamento semelhante em todos os estados ao longo do tempo.**



## Imputação de Dados Ausentes

### Estimação do Número de Internações e Valor Total

#### Número de Internações

	Real	Previsto	Erro Percentual
data			
2019-02-01	928729.00	968117.00	4.24
2019-03-01	940689.25	939654.00	-0.11
2019-04-01	993606.00	959513.25	-3.43
2019-05-01	950850.25	921687.00	-3.07
2019-06-01	986651.00	887696.00	-10.03
2019-07-01	1018125.00	930820.00	-8.58

**Imputação de Dados:** Como dito anteriormente, por questões de simplicidade e funcionalidade foi feito a imputação dos períodos faltantes através da **Média Móvel**, que para efeito de resultados performa melhor quando comparado com a média, por exemplo.

Para avaliar as previsões do modelo foi feito a comparação com o valor **Real**, **Previsto** e gerado o **Erro Percentual** para cada previsão.

#### Valor Total

	Real	Previsto	Erro Percentual
data			
2019-02-01	1207865832.37	1281092045.40	6.06
2019-03-01	1229407235.81	1240008017.63	0.86
2019-04-01	1280487320.11	1258876917.03	-1.69
2019-05-01	1240123472.18	1225458658.58	-1.18
2019-06-01	1282522786.61	1192137192.44	-7.05
2019-07-01	1352327454.58	1255676111.65	-7.15

Conforme apresentado é perceptível que o erro tende a **aumentar** com o **distanciamento** do dataset.

Por fim, os dados foram inseridos ao dataset para efetuar as **previsões finais**.



## Número de Internações, Óbitos e Valor Médio AIH

	Previsão Internação	Previsão Óbitos	Previsão Valor Médio AIH
2019-08-01	984759.00	42556.76	1171.29
2019-09-01	996719.25	43795.14	1176.72
2019-10-01	1049636.00	44808.36	1160.78
2019-11-01	1006880.25	44225.62	1189.79
2019-12-01	1042681.00	47407.93	1189.09
2020-01-01	1074155.00	50458.93	1207.86

Com os valores previstos do número de **Internação** e **Valor Médio AIH** nas etapas anteriores.

Foi possível estimar o valor total gasto em internações para os próximos 6 meses.

	Previsão Internação	Previsão Valor Médio AIH	Total - Bilhões de R\$
2019-08-01	984759.00	1171.29	1.15
2019-09-01	996719.25	1176.72	1.17
2019-10-01	1049636.00	1160.78	1.22
2019-11-01	1006880.25	1189.79	1.20
2019-12-01	1042681.00	1189.09	1.24
2020-01-01	1074155.00	1207.86	1.30

Como mencionado na etapa de projeto foi aplicado o modelo **Auto ARIMA** para fazer previsões das séries temporais.

O modelo gera automaticamente os melhores parâmetros para as previsões do projeto.



**Questionamentos:** Por que os custos dos serviços profissionais são mais caro na região Sul do Brasil? Custo de vida elevado na região consequentemente o valor médio de internação tem um custo elevado.

Por que o Rio de Janeiro possui a maior média de permanência em internação e uma taxa de mortalidade elevada? Qual a relação entre esses números?

Para os próximos estudo deve ser investigado com mais cautela e ter disponível outras variáveis como: população, tipo de internações, dados econômicos etc. Cada região tem suas particularidades, assim como, foi visto no Sul que teve um elevado custo de internações.

**Conclusões:** Conforme apresentado no estudo de previsão, o número de internação e valor pago pode aumentar nos meses analisados. Tendo em vista essa informação, o governo deve entrar com ações como programas de saúde voltados para a conscientização de epidemiologias e doenças mais frequentes a fim de mitigar a necessidade de internações, mobilizar a sociedade para o cuidado da saúde etc. Oferecer segurança para a população, assim como mais fiscalizações para evitar acidentes causados por alcoolismo entre outros problemas.

Prestar um serviço de qualidade nos primeiros atendimentos para que não haja retornos e buscas excessivas da população por diagnósticos eficazes.

O estado de São Paulo concentra os maiores números de internações, provavelmente por ter uma população numerosa. É preciso entender os motivos das internações. Quais os recursos básicos de saúde para essa população? Qual o tipo de internação, cirurgias? Contratar mais médicos reduz o número de internação? Sabemos que a demanda é grande para poucos profissionais. Se tiver um melhor acompanhamento na internação talvez o paciente retorne para casa mais rápido.

Assim como os principais estados e capitais tem os melhores hospitais e recursos. A maior parte das cidades ao redor recorre a essas capitais para atendimento. Dessa forma o estado precisa ter pontos focais de atendimento isso pode reduzir o número de internações e a população ter de perto um atendimento de qualidade. Na maioria das vezes conseguir um bom atendimento é muito complicado.



