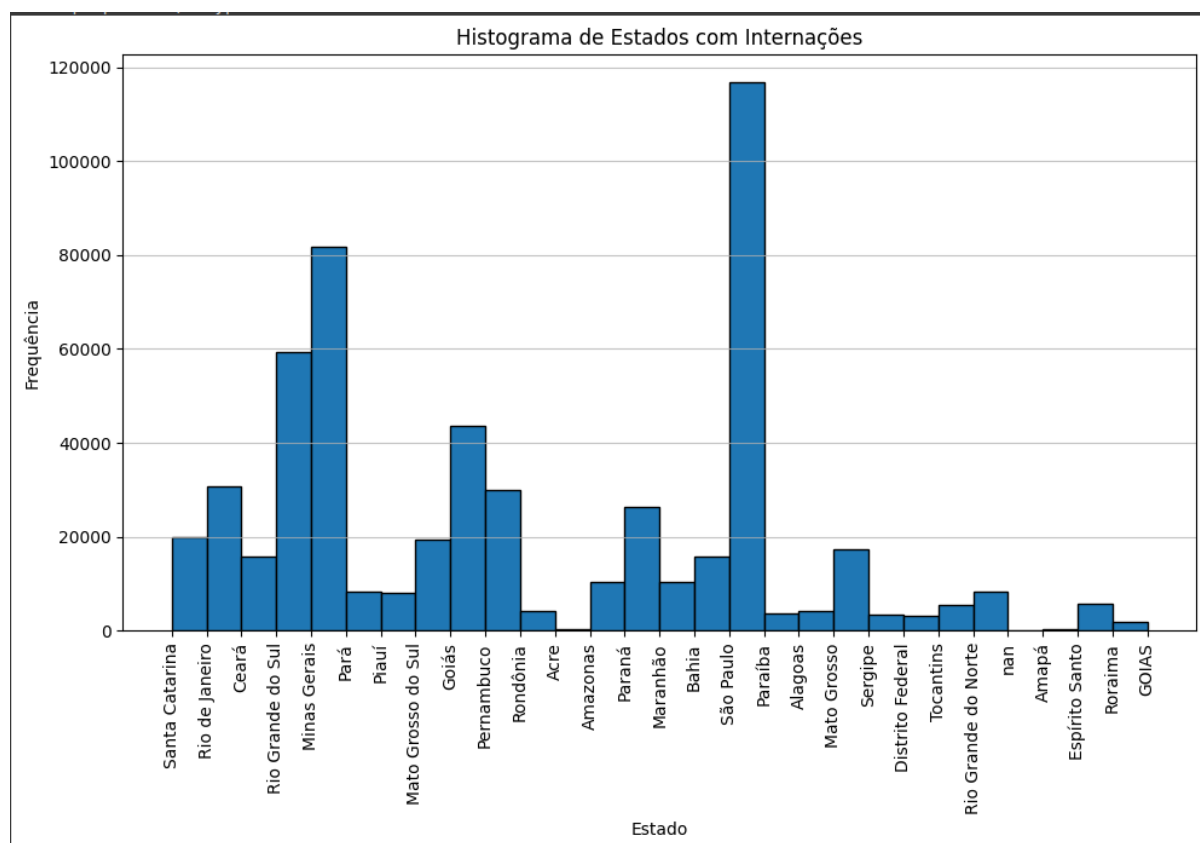


Relatório da análise da base de dados "Registro da ocupação hospitalar COVID-19" de 2020

Frequência absoluta, relativa e histograma:

Inicialmente, foi criada uma análise descritiva da base que guiou todo o processo até o final. Podemos notar estados com grandes números de ocupação hospitalar: São Paulo, Minas Gerais e Rio Grande do Sul são os mais destacáveis, conforme a figura abaixo:



Medidas de tendência central: Ocupações da UTI por suspeita de COVID-19

Pedro M. Rampazo (SP3100065)

Análise e Desenvolvimento de Sistemas - Noturno

Foi criado um gráfico boxplot para verificação da concentração do número de ocupações por suspeita de COVID-19. Podemos notas que no intervalo de 0 a 200 internações se encontram a maior quantidade de suspeitas, conforme mostrado no gráfico:

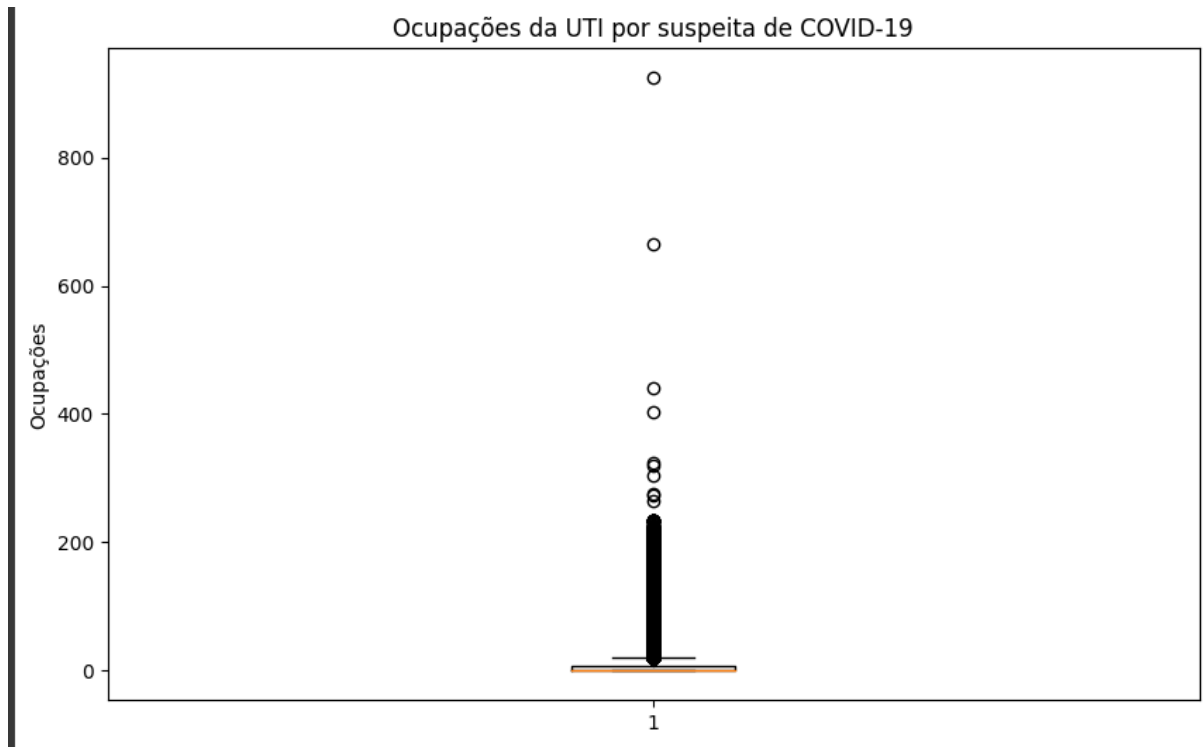
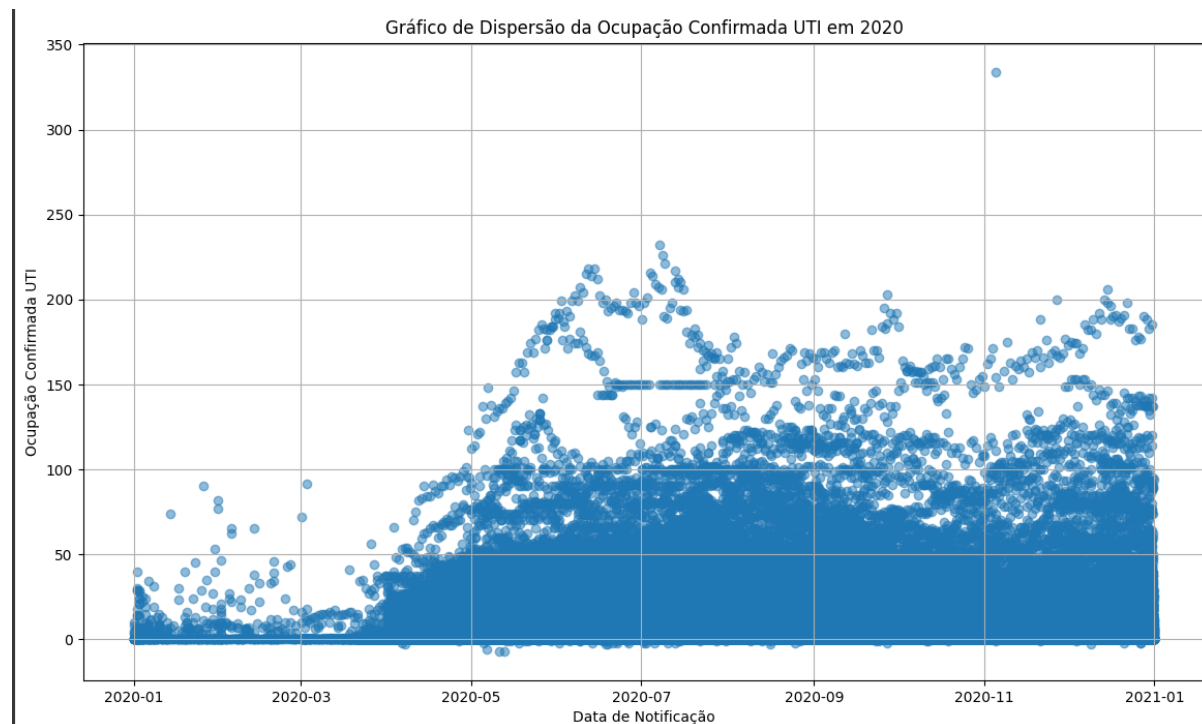


Gráfico de dispersão durante o ano de 2020

Com o gráfico, podemos notar o crescimento exacerbado dos casos de ocupação da UTI para casos de COVID-19, conforme mostrado:



Probabilidade: Teorema de Bayes

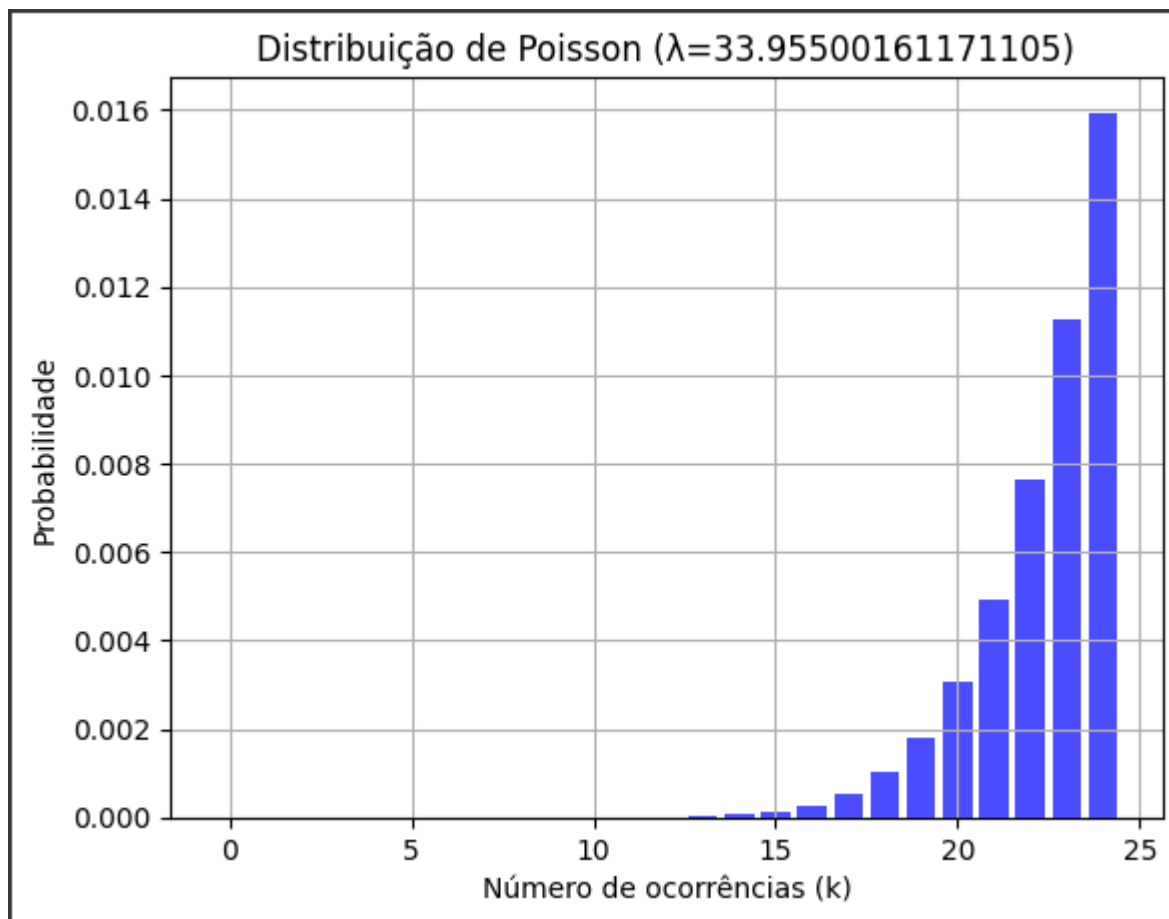
A base possuía muitos valores sobre as saídas das ocupações, então foi realizada uma análise para saber sobre, todas as que ocorreram no ano de 2020, qual foi a probabilidade de um óbito ter sido confirmado. Segue resultado abaixo:

```
Porcentagem de Confirmados: 39.94925824869184
Porcentagem de Óbitos: 0.6953386246173019
Saída confirmada por óbito: 0.43408678694246566
Resultado: 0.007555517239347032
```

Teorema de Poisson: Probabilidade de alta no próximo ano

Foi gerada uma estimativa baseada na média de pacientes que tiveram alta (aproximadamente 33 pacientes), foi analisado qual a probabilidade de 25 pacientes tiver alta no próximo ano. Segue resultado abaixo:

```
P(X = 0) = 1.7927930792586576e-15
P(X = 1) = 6.087429189569214e-14
P(X = 2) = 1.0334933397149958e-12
P(X = 3) = 1.1697422671905133e-11
P(X = 4) = 9.929650141935102e-11
P(X = 5) = 6.74322573146267e-10
P(X = 6) = 3.816104009665775e-09
P(X = 7) = 1.8510831114094023e-08
P(X = 8) = 7.856691253914682e-08
P(X = 9) = 2.964155157659883e-07
P(X = 10) = 1.0064789315570283e-06
P(X = 11) = 3.1068176130156518e-06
P(X = 12) = 8.790999754769882e-06
P(X = 13) = 2.2961416218597066e-05
P(X = 14) = 5.568963747925939e-05
P(X = 15) = 0.0001260627820242575
P(X = 16) = 0.00026752887292565163
P(X = 17) = 0.0005343496065511598
P(X = 18) = 0.001007991208425662
P(X = 19) = 0.001801386479299157
P(X = 20) = 0.0030583040403958594
P(X = 21) = 0.004944986600987786
P(X = 22) = 0.007632137636655953
P(X = 23) = 0.011267358511019676
P(X = 24) = 0.015940965683391643
```



T-Student: Teste de Hipóteses

Nessa ocasião, foi testado o número de estabelecimentos de saúde que estavam nos municípios de São Paulo e Rio de Janeiro. Após uma análise, constatou-se que não há diferença significativa do número de estabelecimentos analisados nessa base. Conforme o resultado:

```
Média de SP: 0.4992413884346323
Média de RJ: 0.500412488047933
Desvio padrão de SP: 0.5057834196678361
Desvio padrão de RJ: 0.5044894156051053
Estatística t: -0.28705851139790306
P-valor: 0.7740684504874324
Não rejeitar a hipótese nula (H0). As médias das ocorrências não são significativamente diferentes.
```

Qui-quadrado: Teste de significância

Nessa ocasião, foi testado se os números de ocupações confirmadas na UTI têm relação direta com o número de saídas confirmadas por óbitos. E não podemos concluir que há uma relação direta, conforme cálculo abaixo:

```
Estatística t: 11.822138836028412
P-valor: 2.4652869497521483e-25
Rejeitar a hipótese nula (H0). As médias das colunas são significativamente diferentes.
```