

Alunos: Lucas Fernandes E Jeremias Oliveira

Turma: 5º Período A

Faculdade: CESAR School

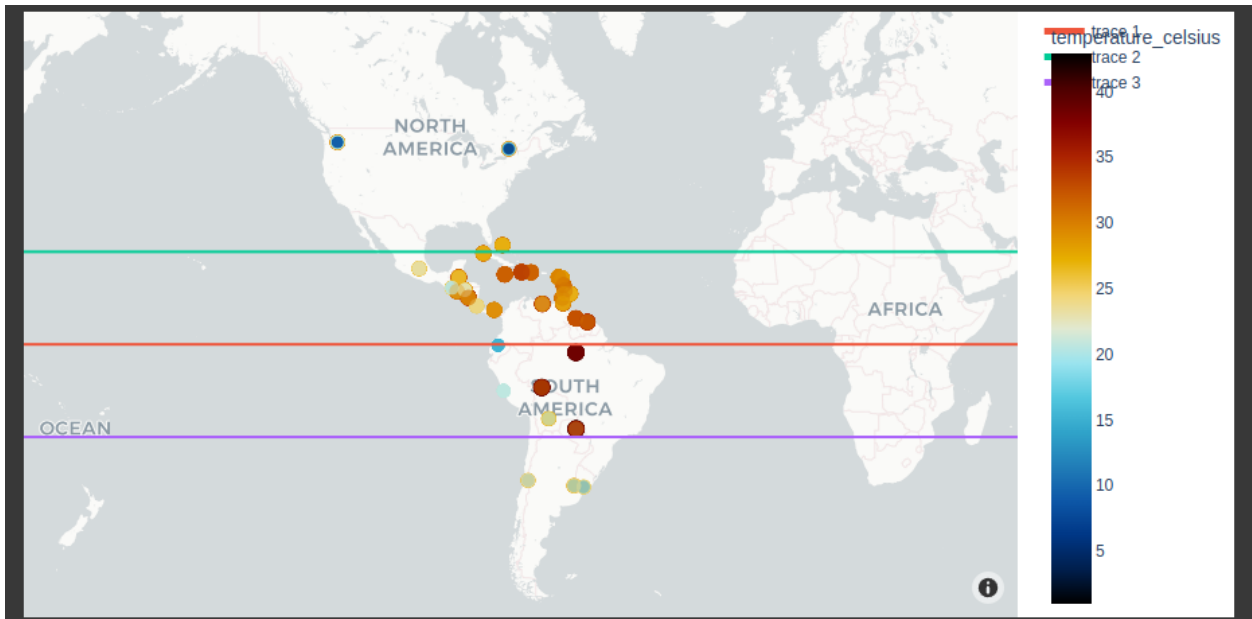
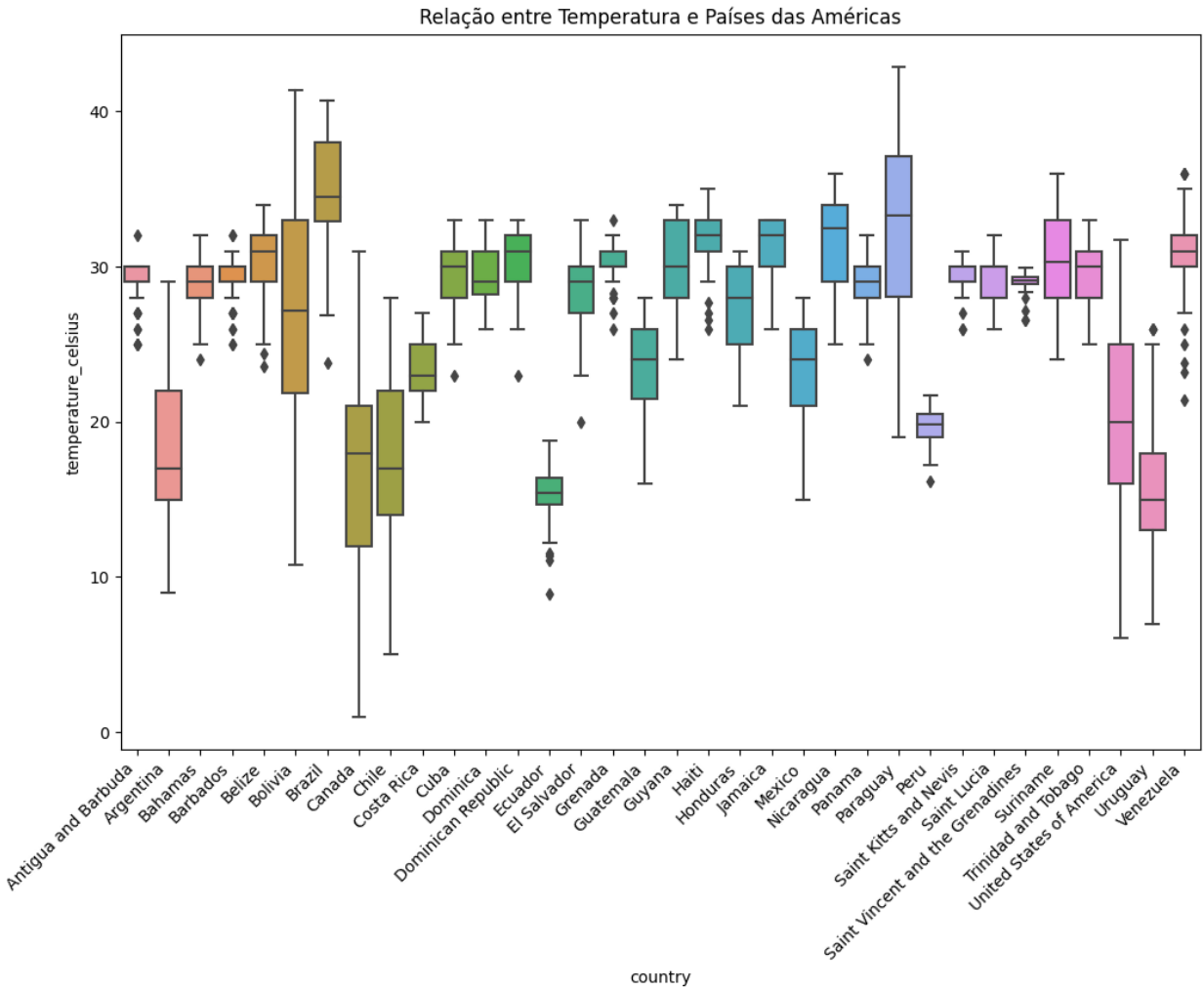
Professor: Guilherme Fernando Cavalcanti Pereira

Algumas informações essenciais a respeito do dataset:

- Este dataset compreende a cidades da maior parte do planeta e suas informações climáticas (qualidade de ar, temperatura, lua, sol e etc...)
- Vimos que o Brasil é uma exceção pois o ponto central da área está situado mais próximo a Manaus, no norte do país.
- As datas e horas vão de: **2023-08-29 14:00** até **2023-11-05 22:00**
- Todas as referências externas estão ao final do documento

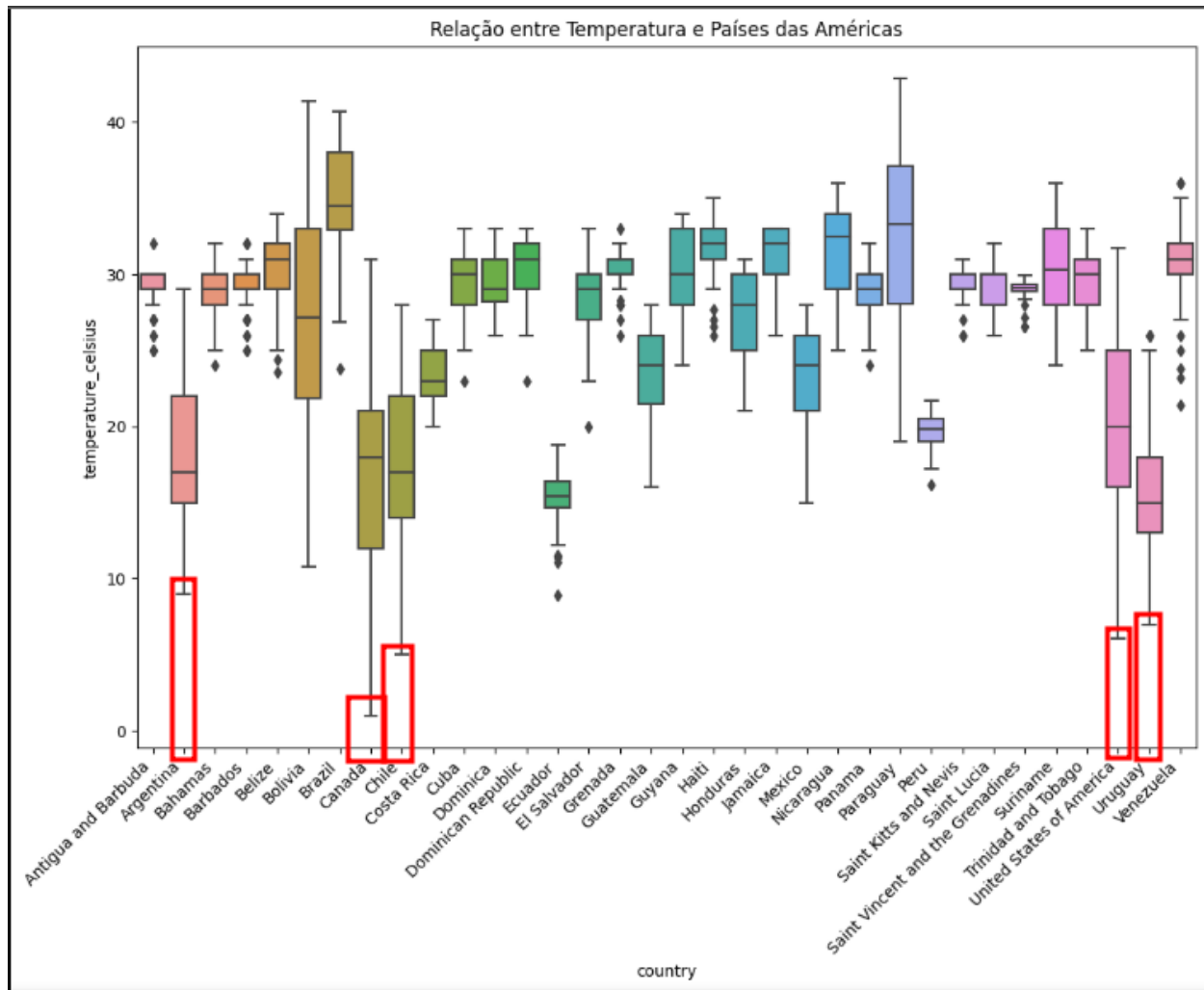
Estatística Descritiva (2,0)

- **Box plot**
 - Relacionando as temperaturas médias das américas com sua distribuição nos trópicos.





Utilizamos os boxplots para comparar a temperatura dos países das américas e entender baseado em suas geolocalização, o motivo de alguns desses países terem a temperatura mais baixa que a maioria dos países das américas. A América está dividida tanto próximo a linha do ártico (América do Norte), quanto abaixo do Trópico de Capricórnio (boa parte da América do Sul) assim também como sua maior parte está concentrada na faixa da linha do Equador, entre Trópico de Câncer e o Trópico de Capricórnio.



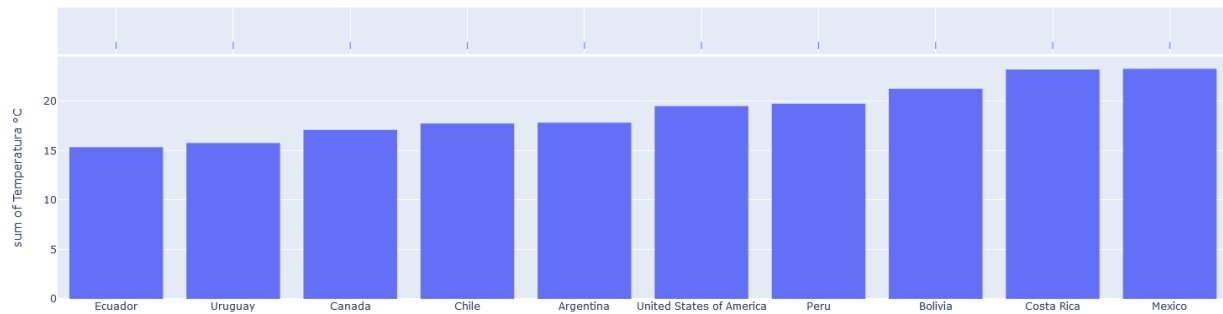
Demarcado então estão as temperaturas que tiveram sua média/mínima como a menor possível durante o tempo que os dados foram coletados. Nisso podemos olhar mais uma vez na segunda imagem desta seção que, todos os países plotados no mapa de calor com temperaturas médias mais baixas estão mais próximos do polo, locais com menor incidência solar, assim como mais próximos dos polos.

- **Histograma**

Nos histogramas, selecionamos as tendências gerais dos países das Américas e focamos também no plano nacional. Selecionamos no plano dos países das Américas os 10 países com uma média de temperaturas mais baixas e os 10 países com média de temperaturas mais altas.

Temperaturas mais baixas:

Histograma dos países de temperaturas mais baixas das Americas



Nota-se que: no histograma o país com a temperatura média mais baixa trata-se do Equador, porém, isso deve-se a onde ocorrem as coletas dos dados dessa temperatura, no Equador esses dados são obtidos em Quito, no Canadá, país consideravelmente mais frio em um plano geral, a coleta ocorre em Ottawa.

Porque Quito e frio?

Devido à sua posição ao equador e a sua alta altitude, **Quito** tem um clima subtropical de altitude, caracterizado por temperaturas frias constantes ao longo do ano com duas estações, uma chuvosa e outra seca. A temperatura anual em **Quito** é de 15,8°C. As temperaturas variam pouco durante o ano. 20 de set. de 2014



guiaviagem.org

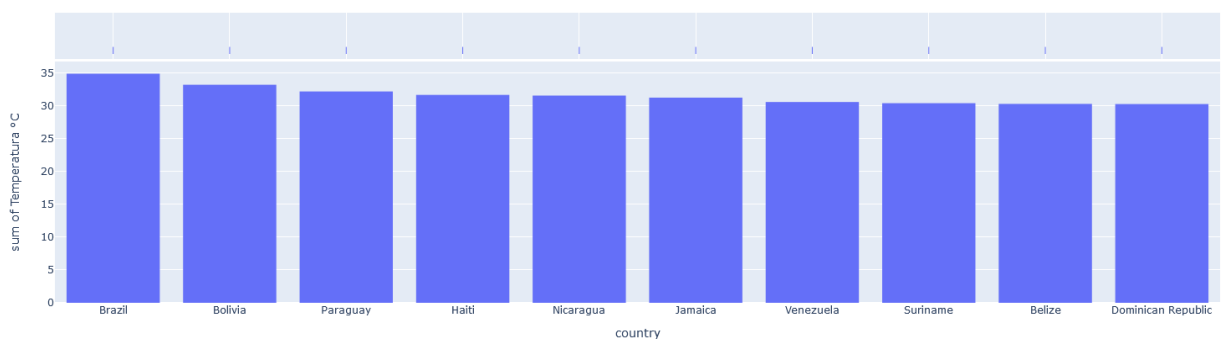
<https://www.guiaviagem.org> > ... > Equador

Quito clima: quando ir para Quito - Guia Viagem

<https://www.guiaviagem.org/quito-clima/>

Temperaturas mais altas:

Histograma dos países de temperaturas mais altas das Americas

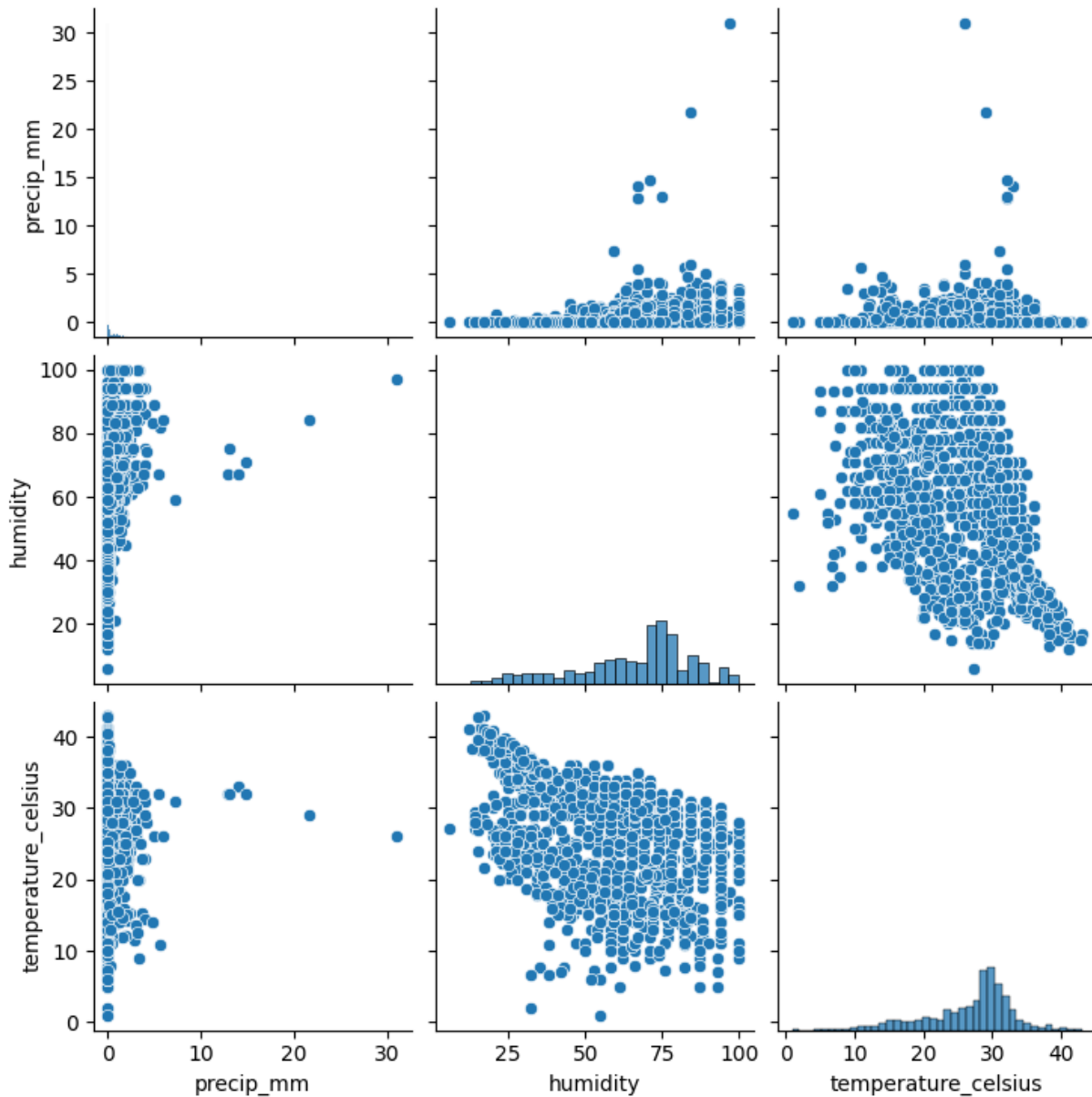


Nota-se que: em primeiro lugar no histograma, o país com temperaturas médias mais altas é o Brasil, na região do Amazonas

- Scatter plot

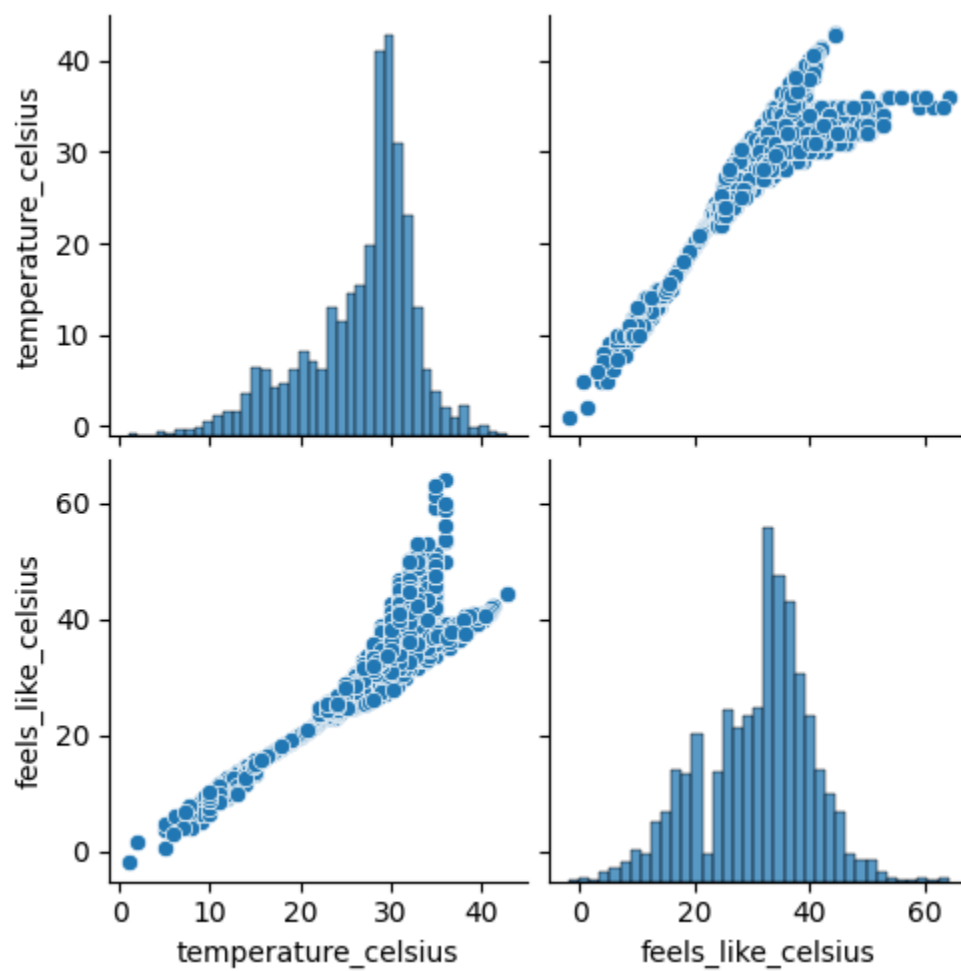
No scatter plot, foram feitas diversas conjecturas, para observar se há alguma tendência ou alguma padronização na distribuição dos dados para definir posteriormente se há ou não correlação entre essas informações e se há uma distribuição que preveja esses dados. As conjecturas feitas foram:

Conjecturas sobre o que afeta as precipitações:



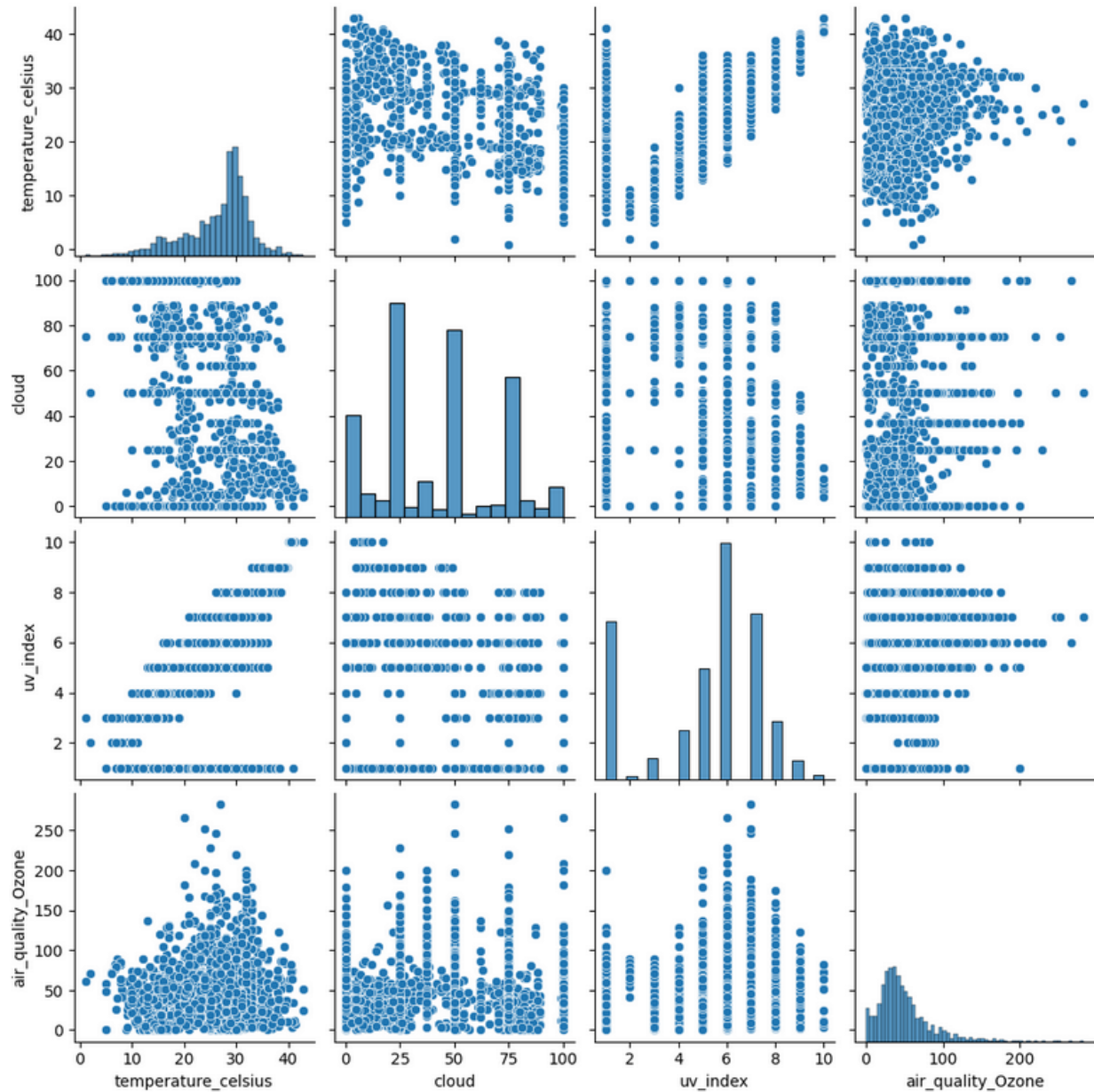
Dados: Precipitação em mm, umidade e temperatura em Celsius

Conjecturas sobre correlação da sensação térmica em celsius e temperatura em celsius:



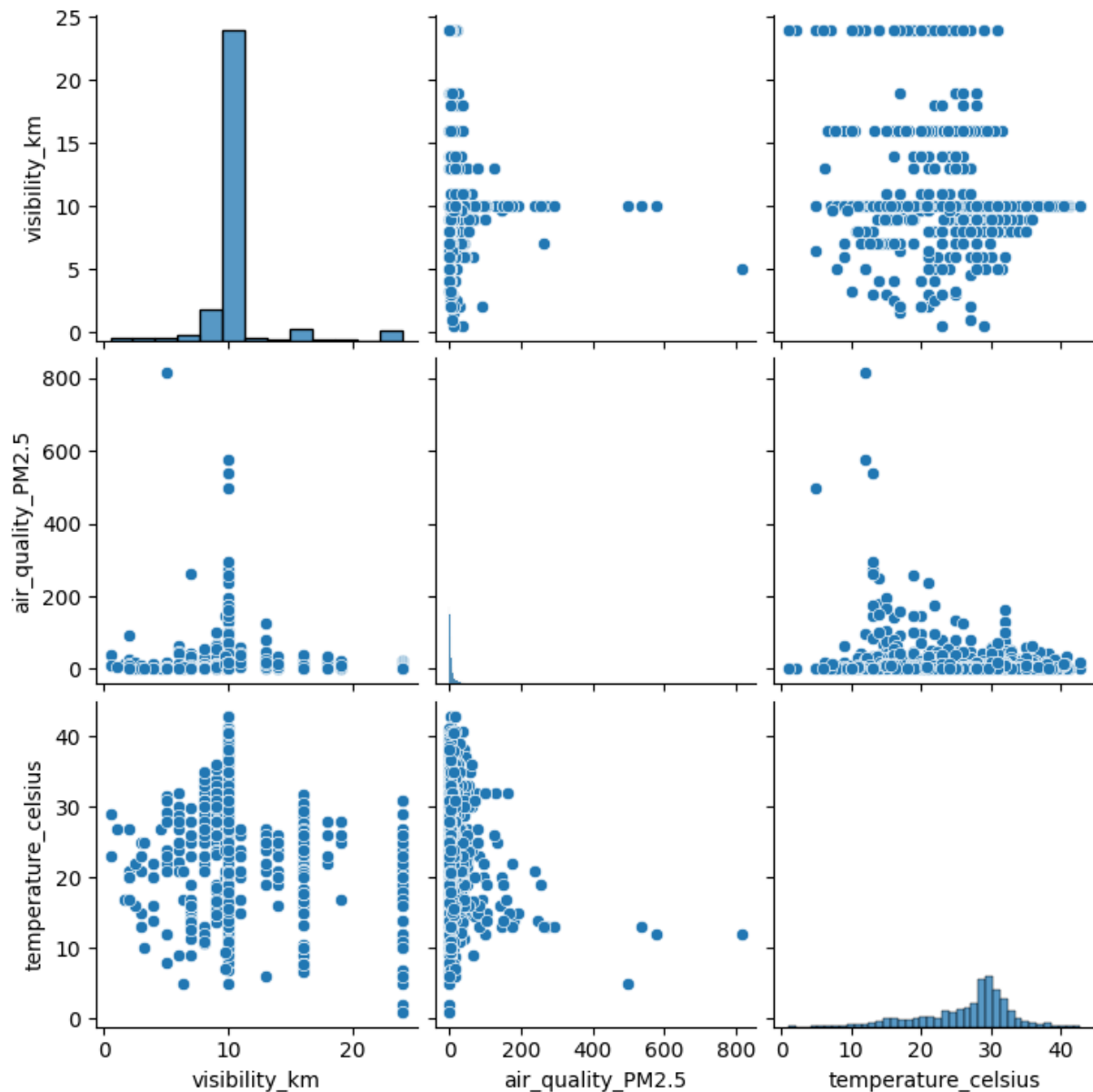
Dados: Temperatura em celsius e Sensação térmica em Celsius

Conjecturas sobre Temperatura em celsius de acordo com porcentagem de nuvens cobrindo o céu e index UV de acordo com temperatura e cobertura do céu, além da qualidade do ar com ozônio:



Dados: Index UV, temperatura em celsius, quantidade de nuvens cobrindo a porcentagem do céu e qualidade de ar com poluição de ozônio

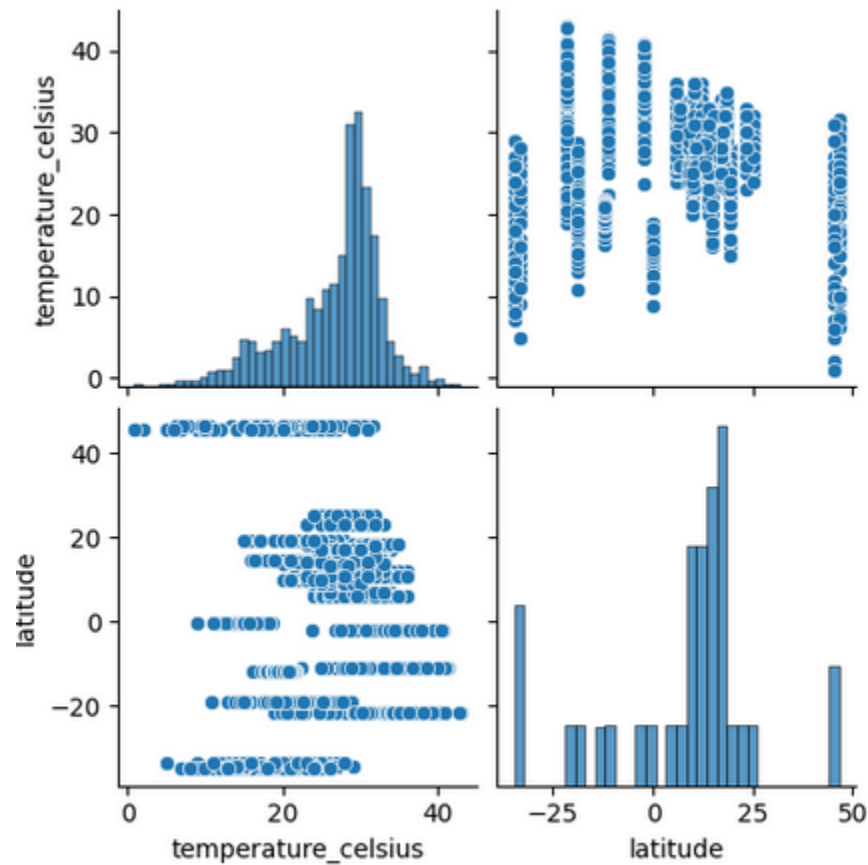
Conjecturas sobre o que afeta a distância da visibilidade:



Dados: Visibilidade em Km, qualidade do ar baseado no padrão PM2.5

(<https://ww2.arb.ca.gov/resources/inhalable-particulate-matter-and-health>, 12 µg/m³ ou abaixo é considerável um padrão saudável de nível de poluição de partículas solidas e liquidas) e temperatura em celsius

Conjecturas sobre a latitude afetar a temperatura do local:



- **Mapa de calor para correlação (heatmap)**

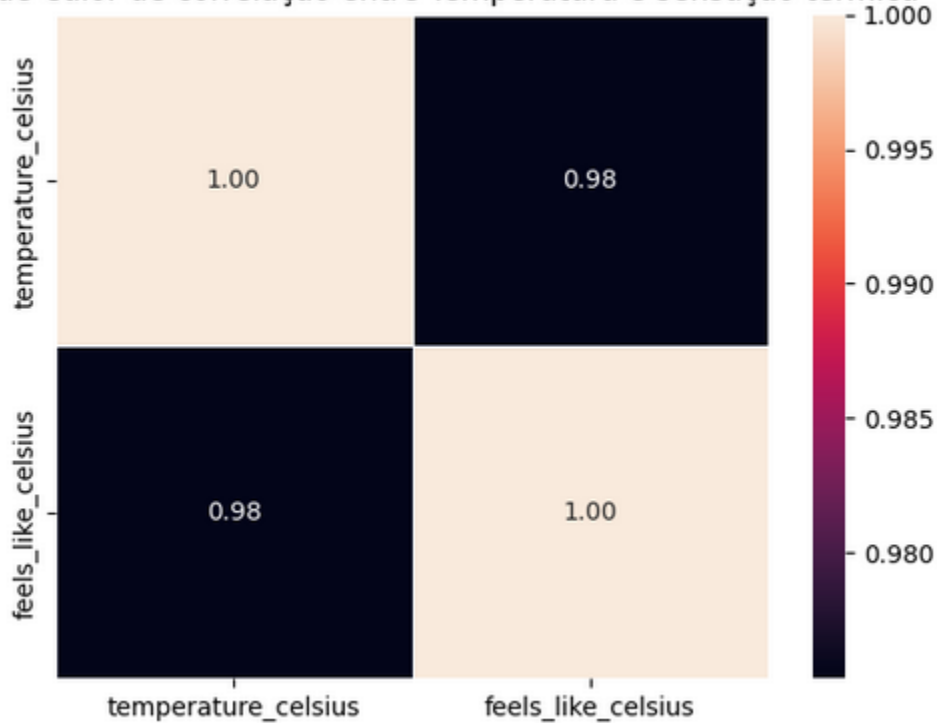
O mapa de calor em conjunto as correlações confirmou o nível de fator correlação, se haviam correlações falsas e também reforçou os dados do scatterplot.

Mapa de calor Temperatura e sensação térmica:

	temperature_celsius	feels_like_celsius
temperature_celsius	1.000000	0.975349
feels_like_celsius	0.975349	1.000000



Mapa de Calor de correlação entre Temperatura e sensação térmica

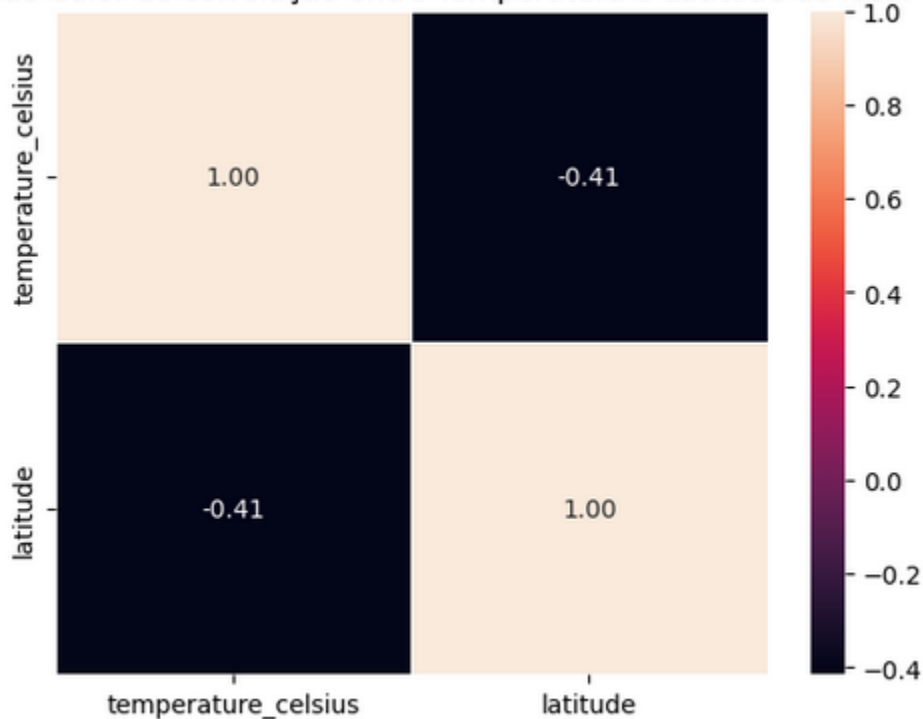


Conclusão: Sensação térmica e Temperatura são extremamente correlacionáveis, pois um deriva quase diretamente do outro.

Mapa de calor que correlaciona os dados de latitude e temperatura em celsius:

	temperature_celsius	latitude
temperature_celsius	1.000000	-0.414807
latitude	-0.414807	1.000000

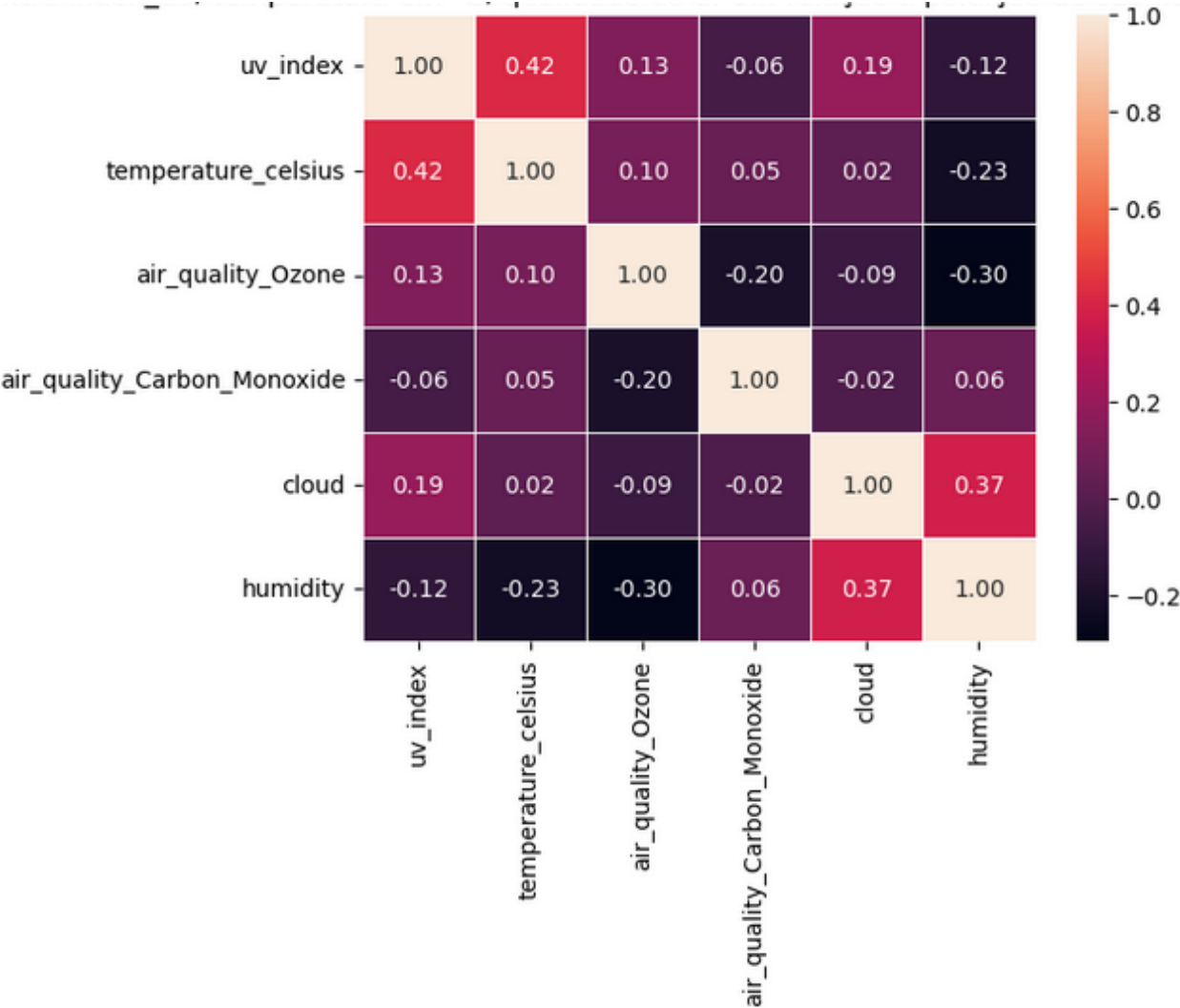
Mapa de Calor de correlação entre Temperatura e Latitude do local



Conclusão: Há uma correlação entre os dados, não tão alta. Isso deve-se ao fato de que a latitude não é o único fator que influencia na temperatura do local, portanto não levanta um fator de correlação tão alto como esperado quando se foi feita a conjectura, há de levar-se outros fatores em consideração, como: altitude, massas de ar, correntes marítimas, continentalidade e maritimidade

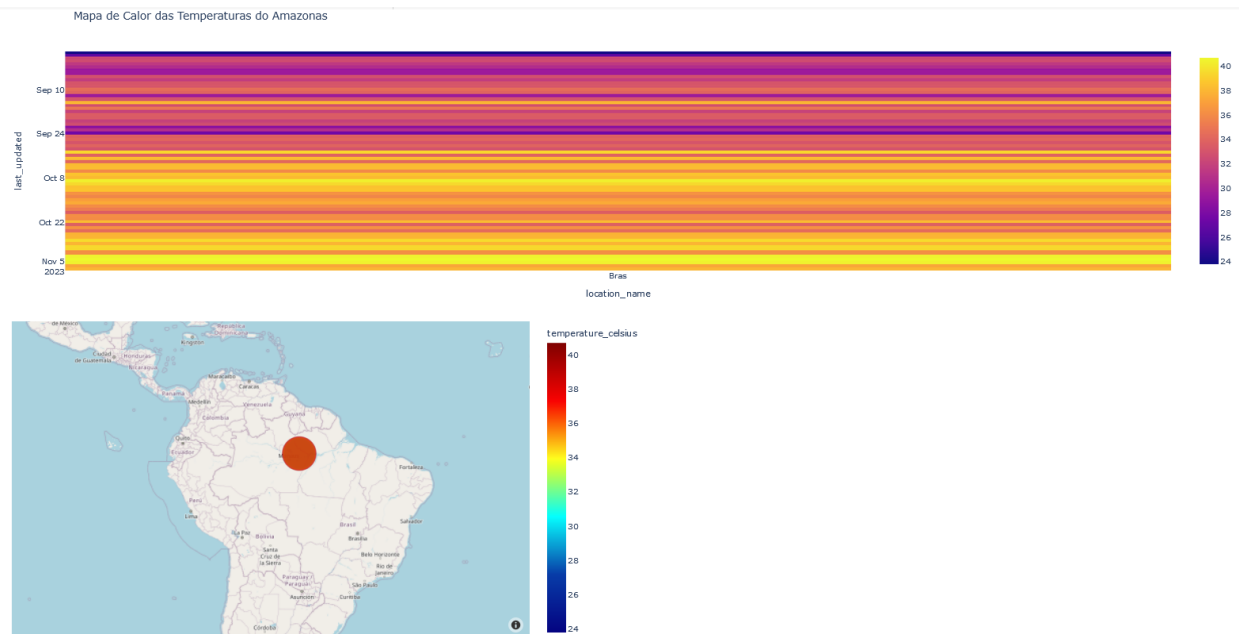
Mapa de calor que correlaciona: Index UV, temperatura em celsius, Qualidade do ar em ozônio, qualidade do ar em monóxido de carbono, porcentagem de nuvens que cobre o céu e umidade

	uv_index	temperature_celsius	air_quality_Ozone	air_quality_Carbon_Monoxide	cloud	humidity
uv_index	1.000000	0.415342	0.128291	-0.058859	0.191778	-0.124990
temperature_celsius	0.415342	1.000000	0.100403	0.051465	0.021196	-0.234230
air_quality_Ozone	0.128291	0.100403	1.000000	-0.201248	-0.090714	-0.296887
air_quality_Carbon_Monoxide	-0.058859	0.051465	-0.201248	1.000000	-0.024757	0.060324
cloud	0.191778	0.021196	-0.090714	-0.024757	1.000000	0.373084
humidity	-0.124990	-0.234230	-0.296887	0.060324	0.373084	1.000000



Conclusão: Temperatura em celsius e index UV tem uma pequena correlação, pois a intensidade solar que define o index UV, porém dentre as correlações, somente há esta e a correlação de nuvens com umidade, já que no ciclo da água, o mecanismo que gera nuvens precisa de umidade, com um fator de correlação minimamente considerável no heatmap.

Mapa de calor das temperaturas no Amazonas

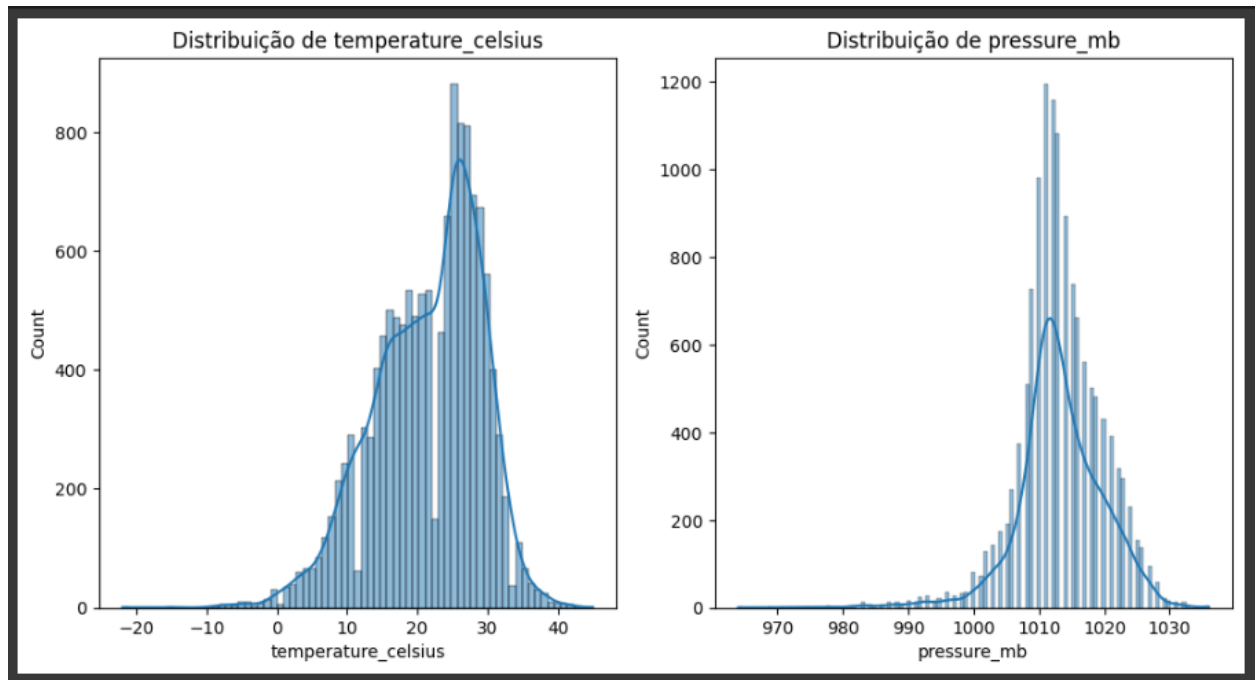


Conclusão: É visto uma tendência de aumento de temperatura no mapa de calor das temperaturas no Amazonas em função do tempo. Além disso presenciamos um momento de seca histórica no Amazonas, convergindo para uma problematização com imprevisibilidade de fim, trazendo tragédias, perdas econômicas e perdas humanas que afetam não só a população local, mas toda a nação.

REF.: (<https://g1.globo.com/am/amazonas/noticia/2023/10/25/seca-historica-sobe-para-60-numero-de-cidades-em-situacao-de-emergencia-no-amazonas.ghtml>,
<https://g1.globo.com/am/amazonas/noticia/2023/10/26/com-seca-recorde-porto-que-atende-zona-franca-de-manaus-fica-mais-de-um-mes-sem-receber-navios-cargueiros.ghtml>)

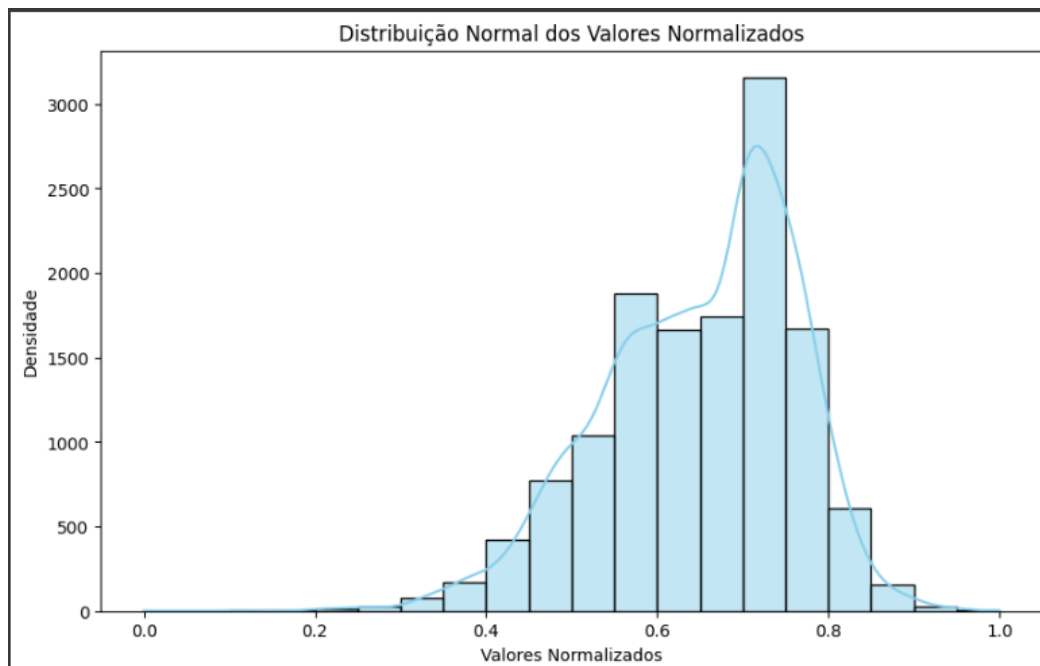
Normal Padrão (1,0)

Como visto em sala de aula, nem todos dados irão seguir uma distribuição normal, com isso foi necessário verificar quais dados do nosso dataframe iria gerar uma distribuição normal (ou mais próxima disto), a verificação se dá pelos gráficos abaixo para as colunas:

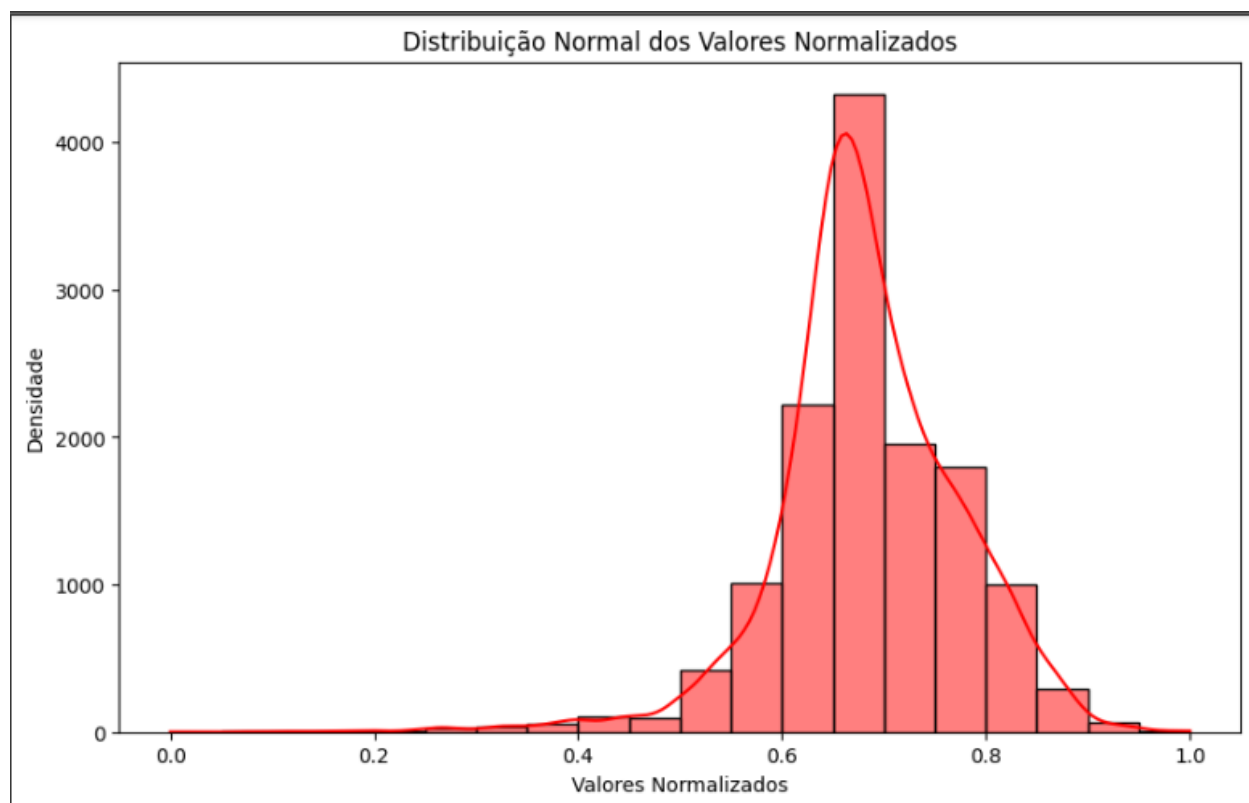


Basicamente aplicamos a normalização gerando para as colunas **temperature_celsius** e também **pressure_mb**, abrangendo todos os países.

Para temperature_celsius:



Para pressure_mb:



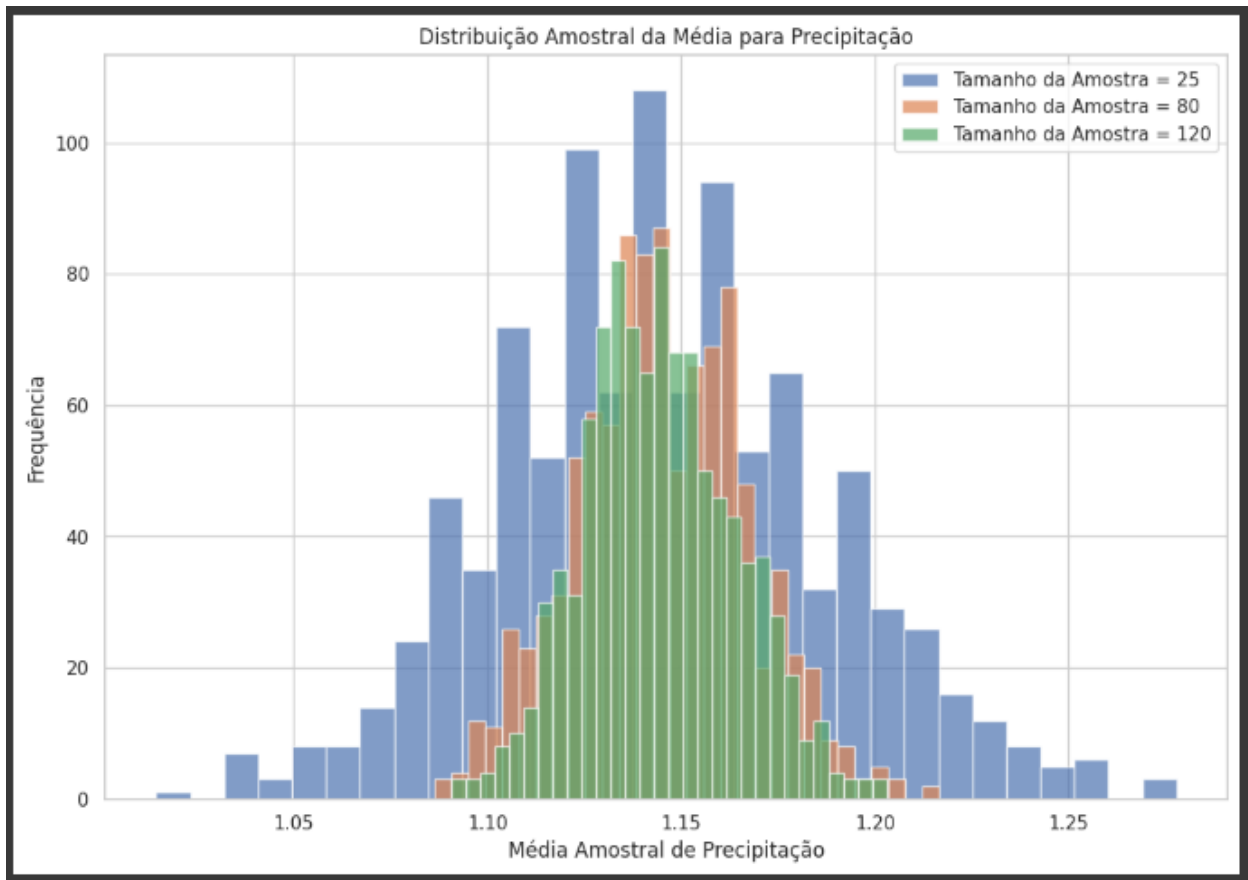
E como pedido pela questão, segue os valores normalizados nas colunas.

	normalized_pressure_mb	normalized_temperature_celsius
0	0.555556	0.758209
1	0.583333	0.731343
2	0.694444	0.746269
3	0.708333	0.480597
4	0.722222	0.701493
...
13440	0.652778	0.791045
13441	0.680556	0.701493
13442	0.763889	0.571642
13443	0.652778	0.711940
13444	0.694444	0.665672

13445 rows × 2 columns

Distribuições Amostrais (1,5)

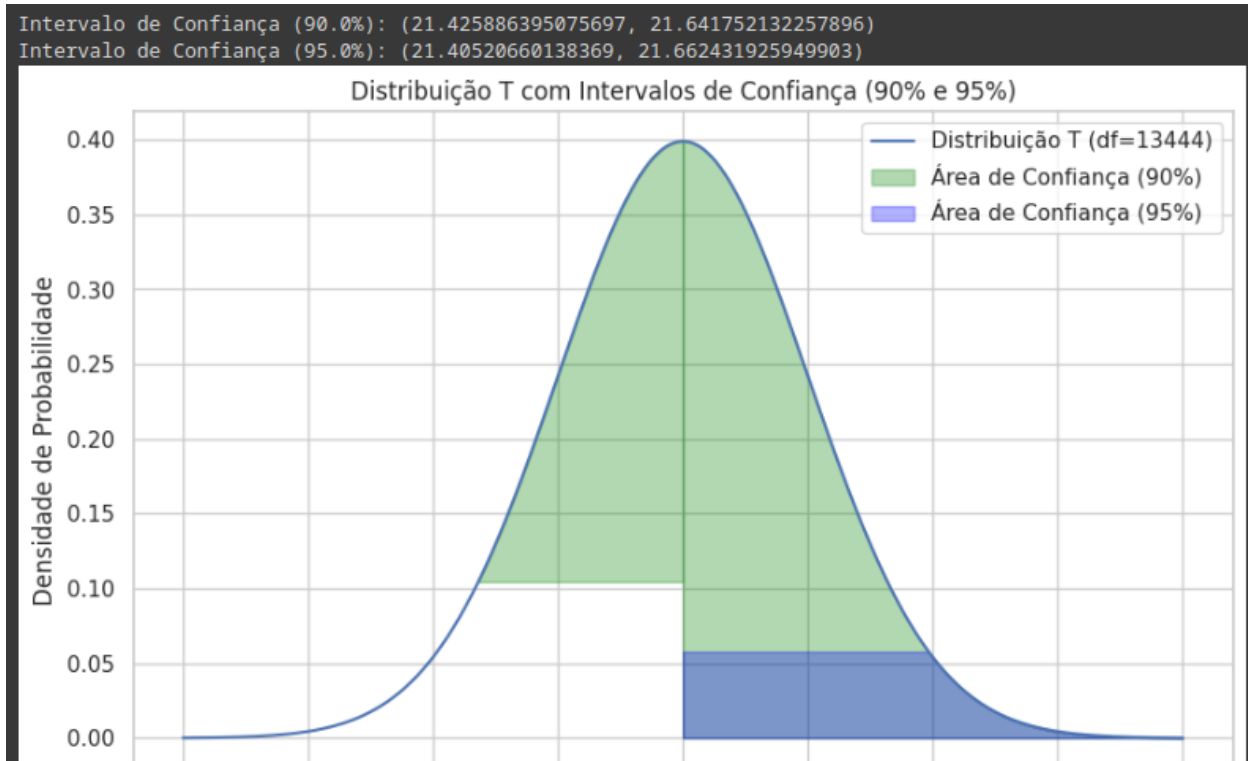
Utilizamos 3 amostras populacionais para o evento de precipitação por milímetro, sendo amostras 25, 80 e 120, resultando no seguinte gráfico:



Não conseguimos plotar as linhas de distribuição amostral. Porém este gráfico aproxima-se de uma distribuição normal, sendo, a amostra laranja e verde com uma dispersão não tão grande, enquanto a dispersão da precipitação do gráfico azul é mais alta.

Intervalo de Confiança (0,5)

Este foi o resultado para intervalo de confiança quando aplicado a coluna temperature_celsius



Plotamos o resultado em T-student pois a coluna se comporta como normal padrão como visto anteriormente, mas não está claro o desvio populacional (sigma).

Referências:

<https://ww2.arb.ca.gov/resources/inhalable-particulate-matter-and-health>

<https://www.indoorairhygiene.org/pm2-5-explained/>

<https://www.guiaviagem.org/quito-clima/>

<https://g1.globo.com/am/amazonas/noticia/2023/10/25/seca-historica-sobe-para-60-numero-de-cidades-em-situacao-de-emergencia-no-amazonas.ghtml>