

# Relatório de análise dos dados do Cepagri

## Integrantes

- **19164** - Bruno Arnone Franchi
- **19191** - Nicolas Denadai Schmidt

## Introdução

Este relatório descreve as análises feitas por nosso grupo sobre os dados meteorológicos do Cepagri/Unicamp, no período correspondido de 01/01/2015 a 31/12/2020. Os dados foram lidos de um arquivo CSV, disponível aqui.

```
# fazer download do arquivo, se necessário
if (!file.exists("cepagri.csv")) {
  download.file("http://ic.unicamp.br/~zanoni/cepagri/cepagri.csv", "cepagri.csv")
}

# ler dataset .csv
names <- c("horario", "temp", "vento", "umid", "sensa")
cepagri <- read.csv("cepagri.csv", header = FALSE, sep = ";", col.names = names)
```

## Tratamento de dados

Fizemos um tratamento inicial aos tipos de dados, para que a coluna de `horario` fosse lida como Data; e a coluna `temp` fosse, assim como as outras, lida como Número. Além disso, removemos as entradas com erros de leitura.

```
# criar dataframe, tratar tipos de coluna e erros
df <- data.frame(cepagri)
# ignorar warning de leituras N/A, vamos corrigir isso na próxima linha
suppressWarnings({
  df <- df %>%
    mutate(horario = as_datetime(df$horario, format = "%d/%m/%Y-%H:%M")) %>%
    mutate(temp = as.numeric(df$temp))
})
df <- na.omit(df)
```

Depois, extraímos o período de tempo que vamos estudar: do primeiro dia de 2015 ao último dia de 2020.

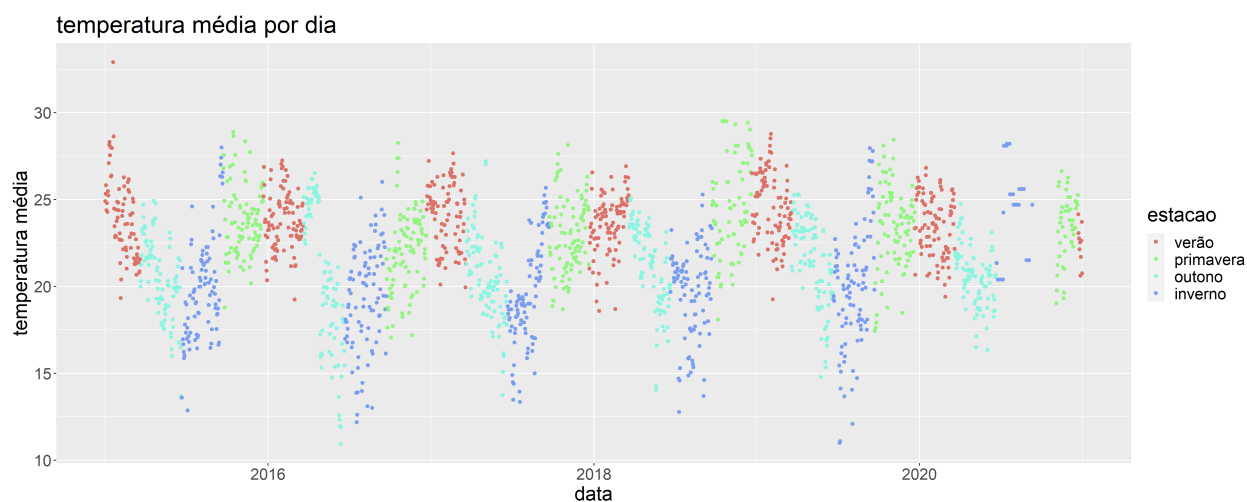
```
# extrair dados entre 2015-01-01 e 2020-12-31
df <- df[
  df$horario >= "2015-01-01" &
  df$horario <= "2020-12-31",
]
```

## Estações do Ano

Nessa análise, esperamos encontrar relações entre a estação do ano e as leituras de temperatura, umidade, e vento.

estacao	temp	umid	vent
inverno	20.24914	54.32983	29.73646
outono	20.84724	71.22079	27.96673
primavera	23.39928	68.81081	30.20303
verão	23.90506	76.50467	26.06102

## Temperatura



Ao analisar as médias de temperatura, podemos claramente relacioná-las às estações do ano, visto que o verão se destaca com médias mais altas; e o inverno com médias mais baixas.

Enquanto isso, primavera e verão aparecem como épocas de “transição” entre os dois extremos, com as temperaturas caindo ao longo do outono, e se elevando ao longo da primavera.

## Umidade



Assim como na análise de temperatura, é fácil perceber diferenças entre as estações: em geral, o verão apresenta umidade maior e o inverno umidade menor, com as estações restantes apresentando dados similares entre as duas outras estações.

## Vento



Ao analisar o gráfico de vento médio visualmente, não se percebe diferenças significativas entre cada estação.

Vamos confirmar a baixa correlação entre vento e os outros dados quando analisarmos a matriz de correlação entre eles.

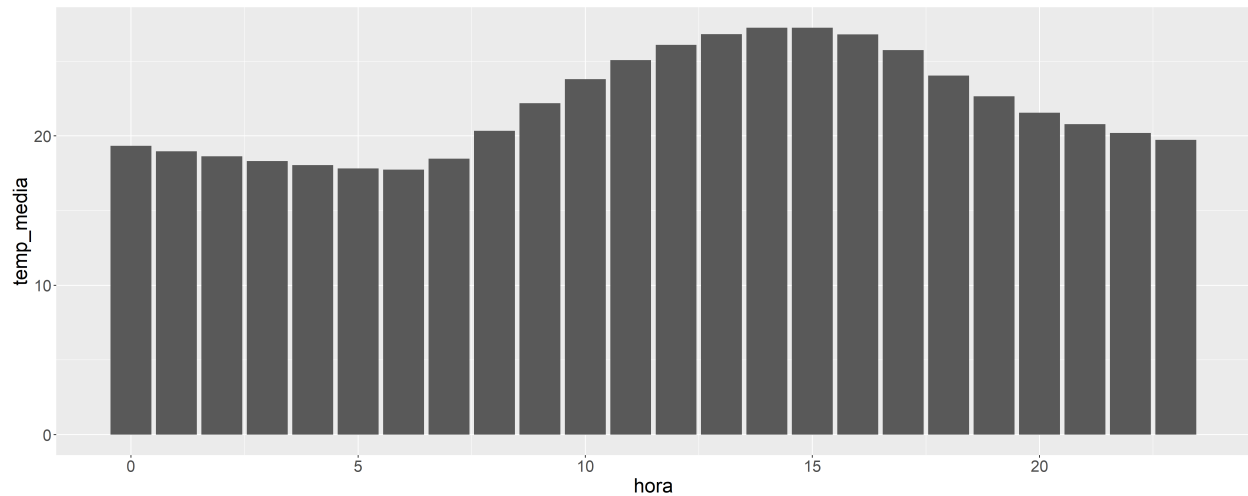
## Ao Longo do Dia

Como nossa segunda análise, decidimos estudar quais eram as temperaturas, taxas de umidade e velocidades do vento médias por hora no período em questão, como visto nos gráficos a seguir:

hora	temp	umid	vent
0	19.33292	79.80218	33.02413
1	18.97275	80.73524	30.97338
2	18.63220	81.58349	29.27986
3	18.31903	82.28488	27.88427
4	18.04292	82.92194	26.60783
5	17.81882	83.47220	25.67924
6	17.74545	83.81083	24.79510
7	18.47384	75.99745	24.32861
8	20.33807	77.13613	24.79870
9	22.19465	70.81765	25.60517
10	23.79054	65.23323	26.46951
11	25.07912	60.81781	27.11706
12	26.09774	57.23029	27.32643
13	26.81262	54.44897	27.61757
14	27.24223	52.58335	27.78395
15	27.23571	52.20967	27.82197
16	26.79769	53.22239	27.77451
17	25.74017	56.56589	26.88287
18	24.03912	61.91580	26.94156

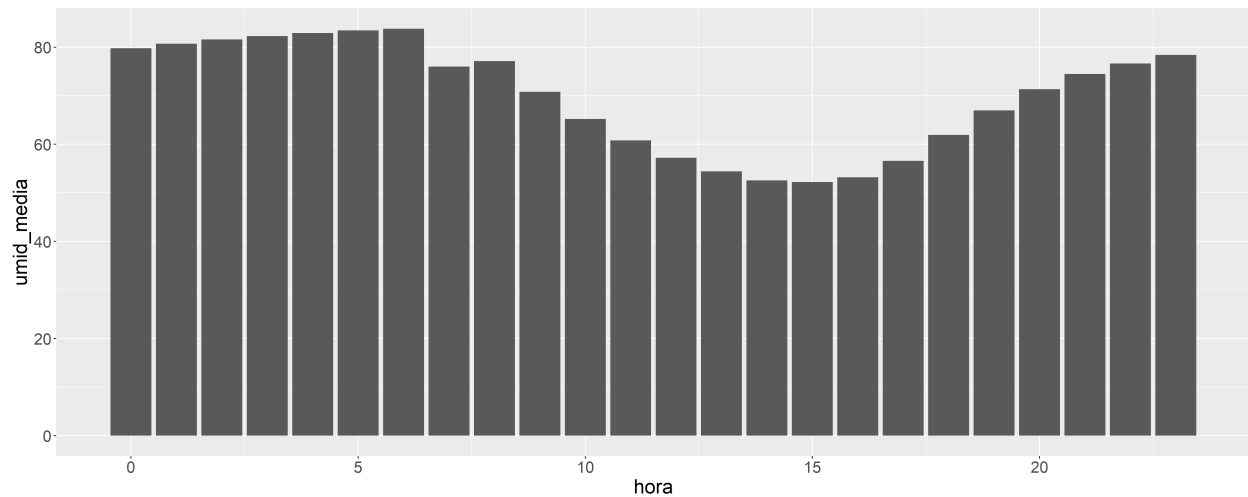
hora	temp	umid	vent
19	22.64236	66.97897	29.26832
20	21.54621	71.32457	32.07470
21	20.78054	74.46206	33.78218
22	20.19568	76.66723	34.52091
23	19.73766	78.40029	34.06435

## Temperatura



Ao analisar o gráfico da temperatura média por hora, descobrimos que a temperatura tende a aumentar progressivamente das 6:00 da manhã até as 15:00 da tarde, quando atinge seu pico, e então diminuir até as 6:00 da manhã do dia seguinte.

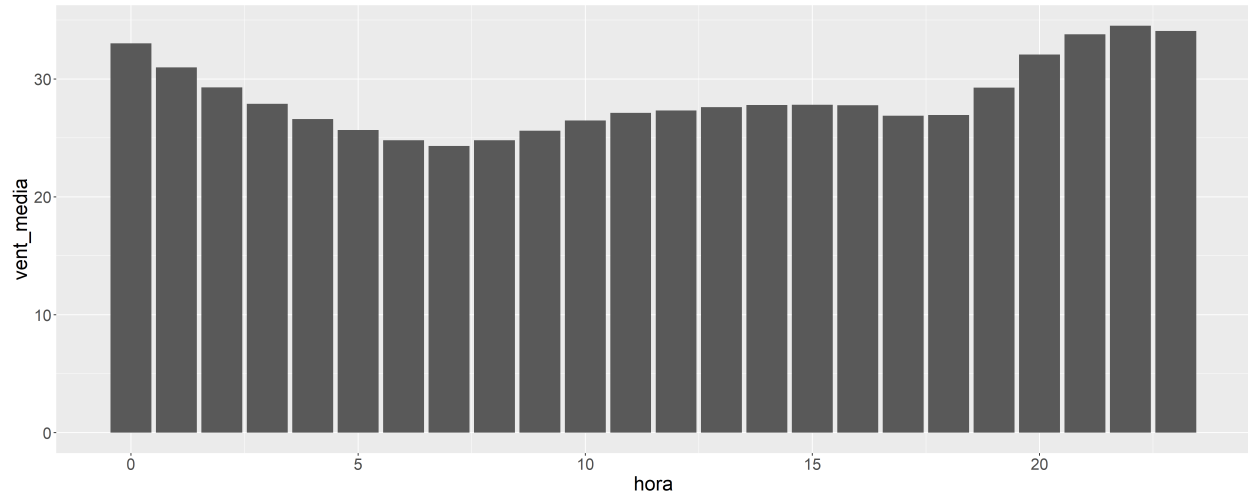
## Umidade



Em contrapartida, ao analisar o gráfico de umidade média por hora, podemos observar que seu comportamento é o oposto do esperado após realizar a análise das relações entre as estações do ano, e a temperatura, umidade e velocidade do vento, onde a umidade tendia a ser maior quanto maior fosse a temperatura. Aqui,

vemos que a umidade encontra-se em seu pico as 6:00 da manhã, de onde diminui até atingir seu mínimo às 15:00 da tarde, para então aumentar até as 6:00 da manhã do dia seguinte.

## Vento



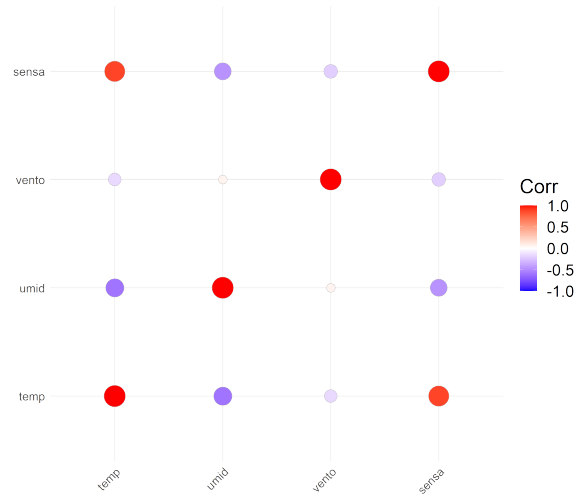
O comportamento da velocidade do vento observada no gráfico de velocidade do vento média por hora por outro lado, é bem curioso, a velocidade do vento tende a atingir seu ponto mínimo as 7:00, permanece sem grandes mudanças até as 18:00 da tarde, e então aumenta até atingir seu ponto máximo às 22:00 da noite, diminuindo, gradualmente, até as 7:00 do próximo dia, sendo assim, maior durante o período noturno, e menor durante o dia.

## Correlação entre Dados

A melhor forma de encontrar relações entre dados coletados é com uma matriz de correlação.

Vamos, então, gerar uma matriz de correlação entre todas as colunas de dados coletados.

	temp	umid	vento	sensa
temp	1.0000000	-0.5973763	-0.1618342	0.8811474
umid	-0.5973763	1.0000000	0.0619120	-0.4715382
vento	-0.1618342	0.0619120	1.0000000	-0.2100982
sensa	0.8811474	-0.4715382	-0.2100982	1.0000000

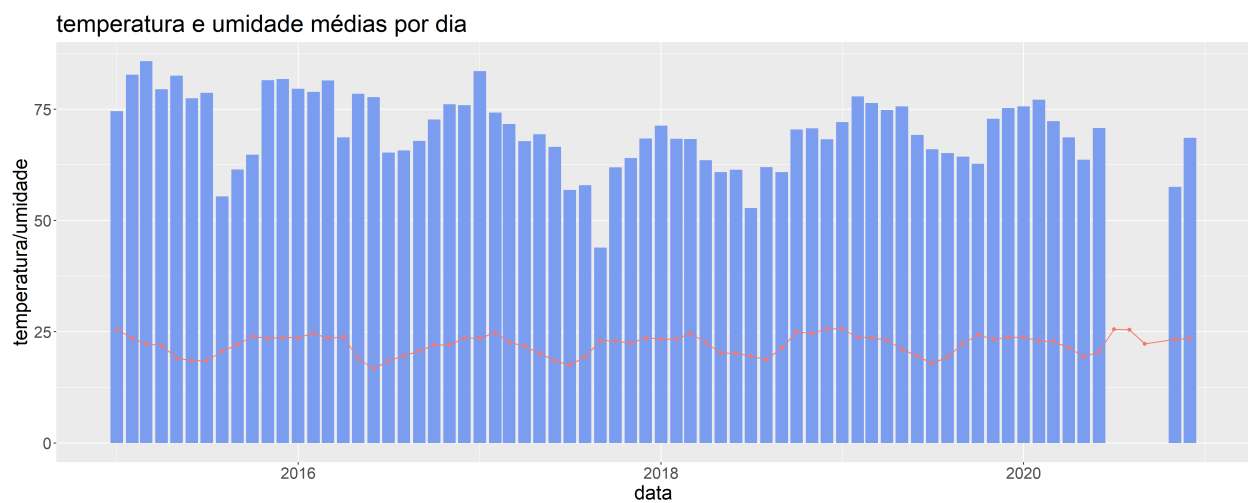


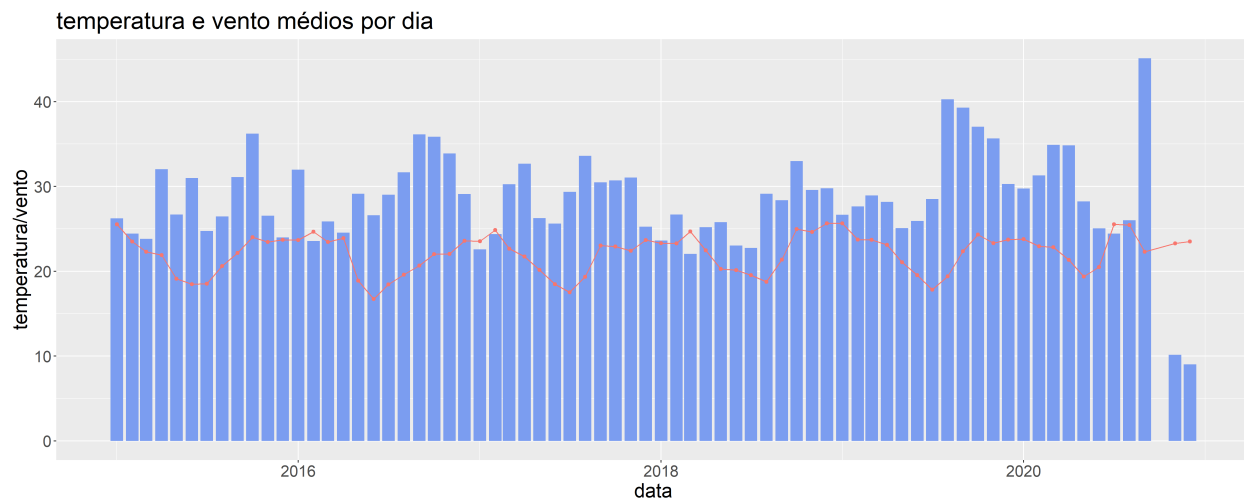
## Temperatura, Umidade e Vento

Ao analisar a correlação entre umidade/vento e temperatura, podemos notar que:

- **umidade:** correlação de **-60%**
- **vento:** correlação de **-16%**

Com isso, podemos concluir que tanto o vento quanto a umidade tendem a ocorrer em conjunto com baixas temperaturas, sendo a relação temperatura-umidade mais forte do que a relação temperatura-vento. O mesmo pode ser observado nos gráficos a seguir:





## Sensação Térmica

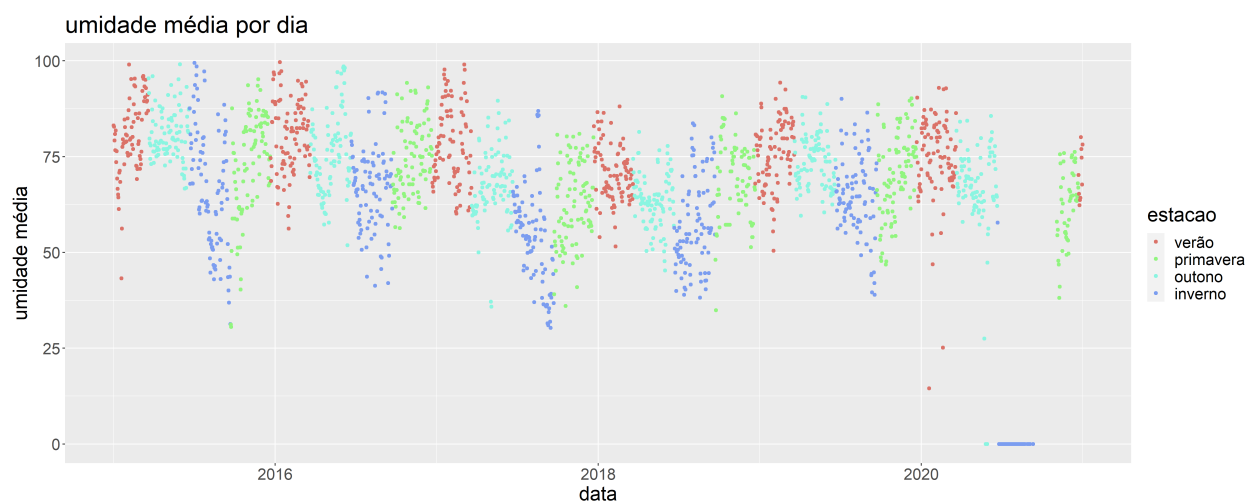
A primeira vista, quando analisamos a matriz podemos perceber que a correlação em maior destaque é que a temperatura afeta diretamente a sensação térmica, sendo aproximadamente 88% relacionadas. Isso faz sentido, porque, se estiver mais calor, as pessoas provavelmente vão sentir mais calor.

Além dessa, podemos encontrar outras correlações:

- **aprox. 88%** temperatura: maior temperatura causa maior sensação térmica
- **aprox. -58%** umidade: menor umidade causa maior sensação térmica
- **aprox. -21%** vento: menos vento causa maior sensação térmica

## Crise Hídrica de 2014-2015

Houve uma grande e famosa crise hídrica em todo o Brasil nos anos de 2014 e 2015. Levando em conta que temos em mãos os dados do CEPAGRI de 2015, um dos anos da crise, decidimos analisar a umidade no ano de 2015 usando o gráfico de umidade diária.



Vemos no gráfico que a umidade no começo de 2015 estava comparativamente alta com relação ao mesmo período dos demais anos da análise. Tal resultado nos deixou surpresos, uma vez em que acreditávamos que

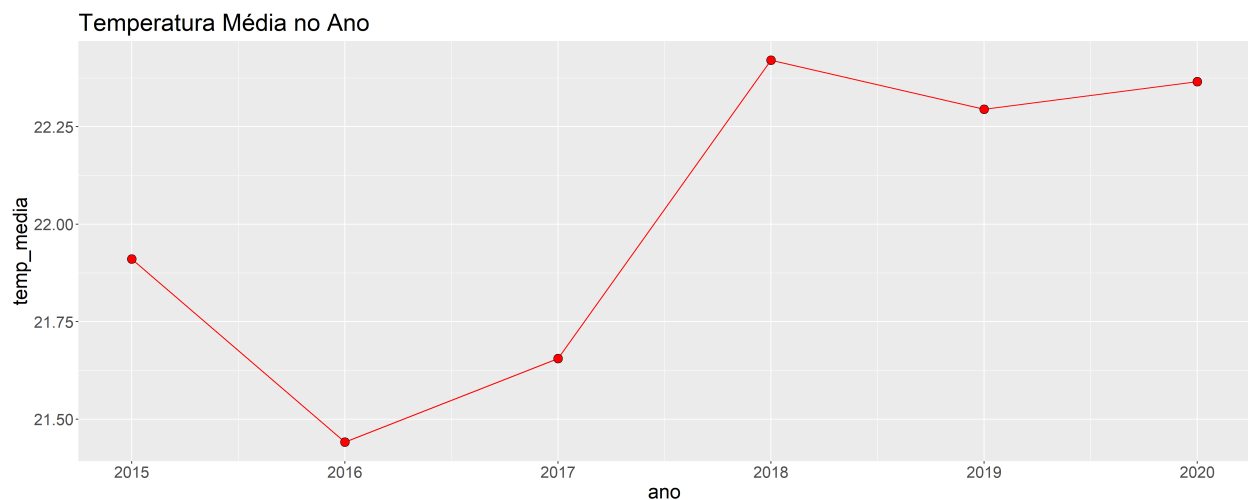
a umidade relativa do ar tenderia a ser menor com relação à um ano sem crise hídrica. Talvez a água que estava em falta tivesse evaporado e se encontrava no ar, o que explicaria a elevada umidade do ar, embora deixaria questões sobre o por que da falta de chuvas.

## Média Anual

Nós já fizemos análises da média dos valores em cada dia analisado, da média por hora, e decidimos realizar a análise da média por ano, montando gráficos de linha para facilitar nosso trabalho.

ano	temp_media	umid_media	vent_media
2015	21.91070	75.46971	27.76617
2016	21.44157	73.97826	29.82182
2017	21.65546	65.28839	28.55130
2018	22.42030	64.86642	26.57498
2019	22.29458	70.96481	31.14706
2020	22.36567	69.02364	25.63963

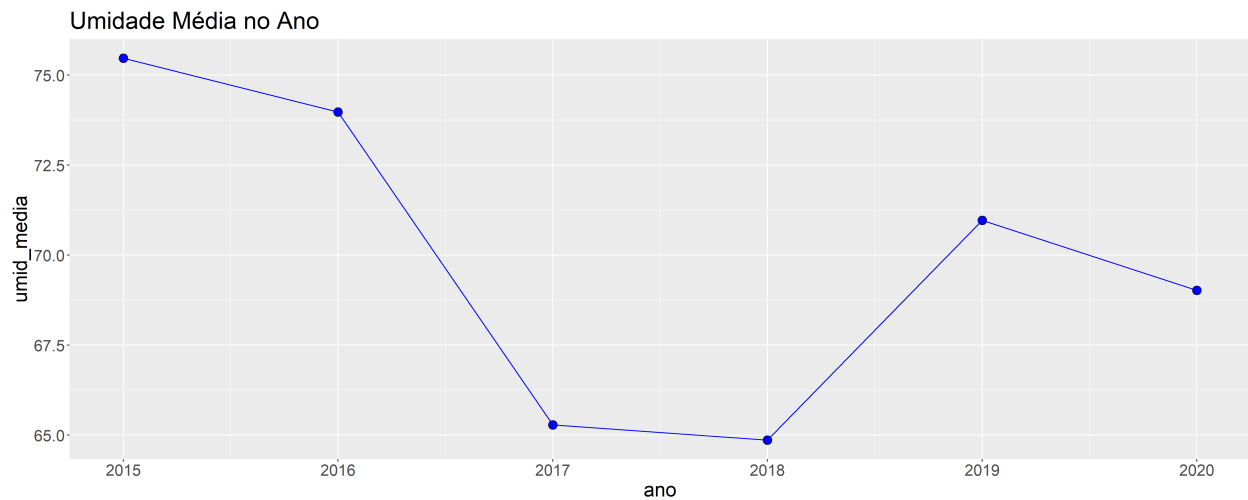
## Média Anual da Temperatura



Antes de analisarmos o gráfico da média anual da temperatura, pensamos que, com tantas discussões a cerca do aquecimento global, o gráfico tenderia a aumentar progressivamente com o passar dos anos, talvez com a variação de décimos de graus Celsius. Para nossa surpresa, o gráfico não seguiu nenhum padrão, com variações de quase 1 grau entre o ano de menor e o de maior média, sendo respectivamente 2016 e 2018, sendo muito inconstante. Percebemos que não seria possível obter nenhum resultado consistente sobre o aquecimento global tendo como base os dados de apenas 6 anos.

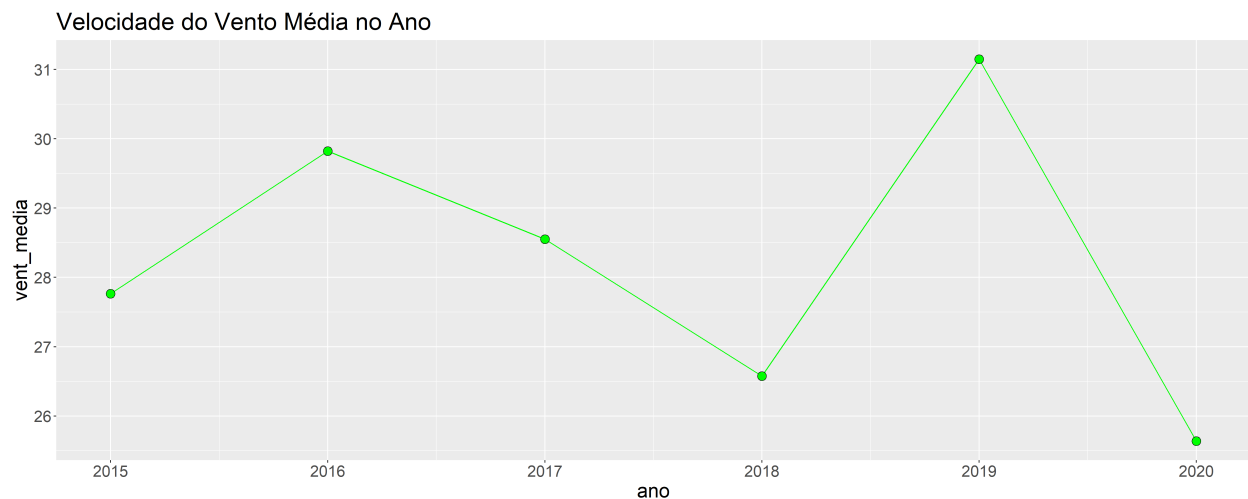


## Média Anual da Umidade



Nos anos estudados, a umidade estava em seu ápice no ano de 2015, estando acima de 75% de umidade relativa, tendo uma pequena queda em 2016, para algo em torno de 74%, uma grande queda em 2017, atingindo algo pouco acima de 65%, continuando a queda em 2018, até atingir pouco menos de 65%, subindo drasticamente para algo em torno de 71% em 2019, tendo uma pequena queda para algo próximo à 69% em 2020. Chegamos à conclusão de que variações pequenas, de 0% a 2,5% na umidade de dois anos é uma ocorrência comum, enquanto variações de mais de 2,5%, embora mais raras, ainda podem ocorrer.

## Média Anual da Velocidade do Vento



Analisando o gráfico, chegamos à conclusão que a velocidade média do vento está sujeita a grandes variações, como as ocorridas entre 2018 e 2019 e entre 2019 e 2020, no entanto, devido à termos poucos dados, não é possível analisar as tendências da velocidade média do vento nem estimar valores para suas variações.