Data sets:

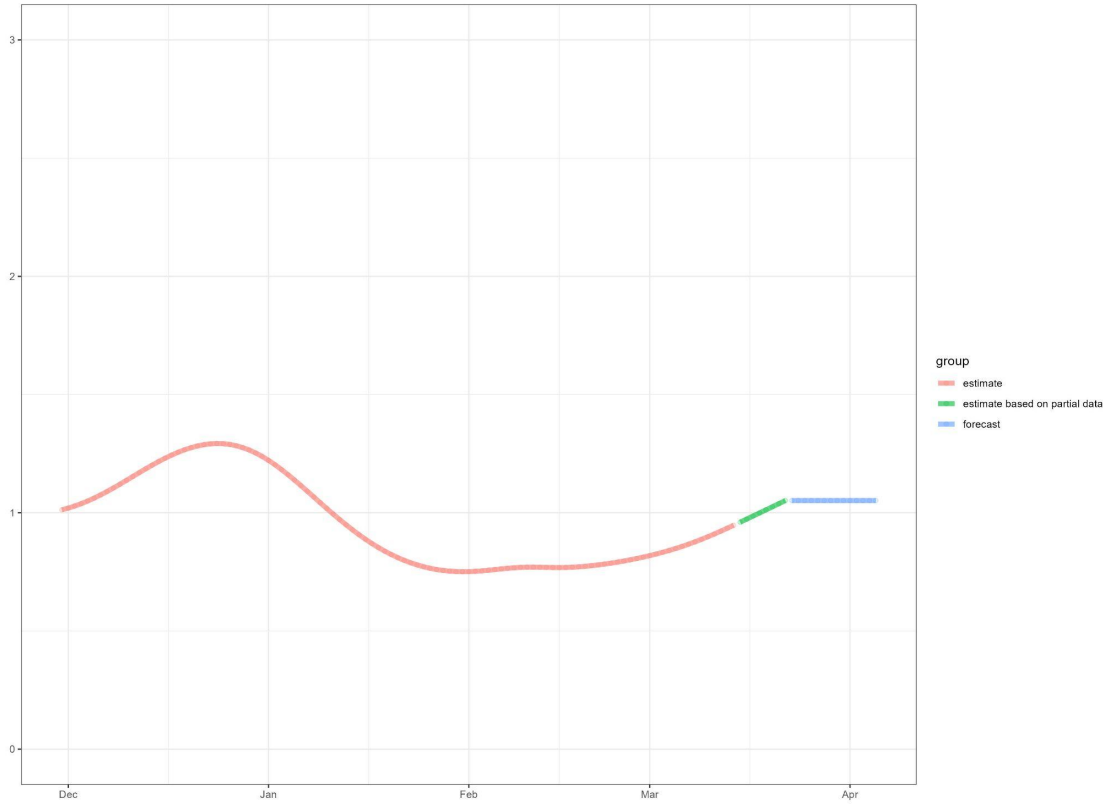
1. Calibrating COVID-19 community transmission risk levels to reflect infection prevalence (Run and update the project from GitHub)

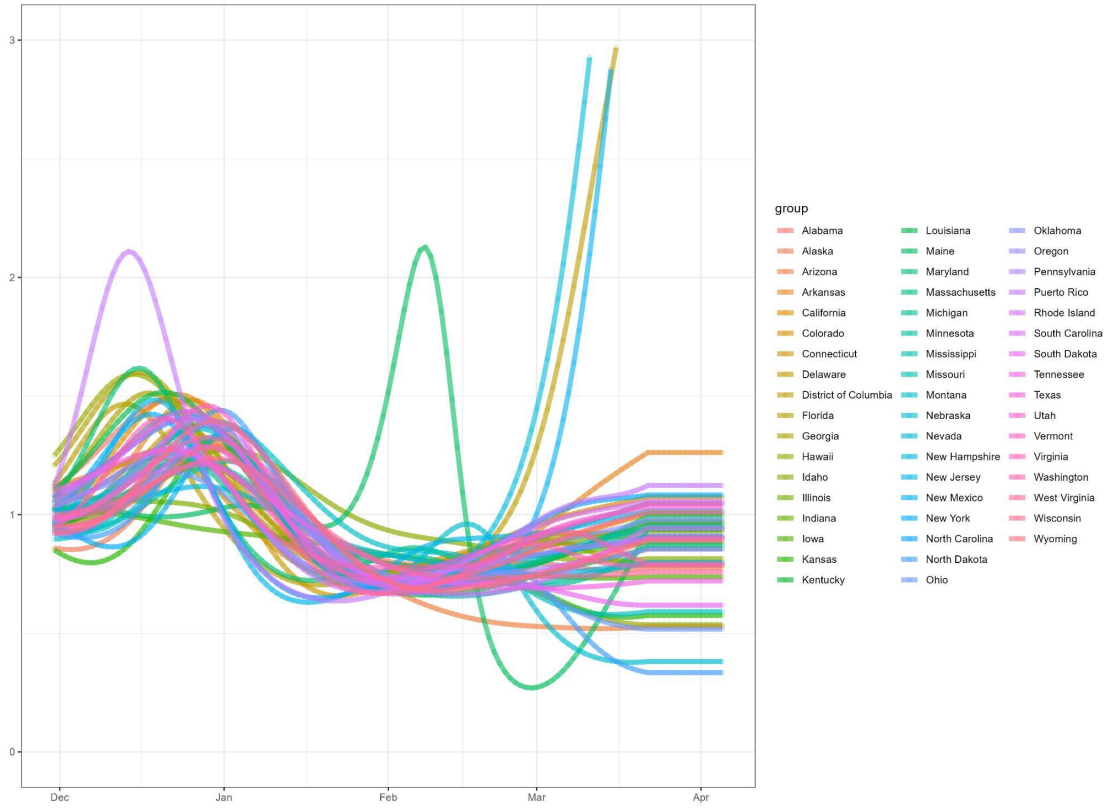
- start.date = "2020-07-01" to “2023-02-13” daily
- State level

2. Harvard National and Subnational Estimates of the Covid-19 Reproduction Number (R) for the United States of America Based on Test Results

- start.date = "2021-11-30" to “2022-04-05” daily (limited time)
- State level
- Only estimate types have been used

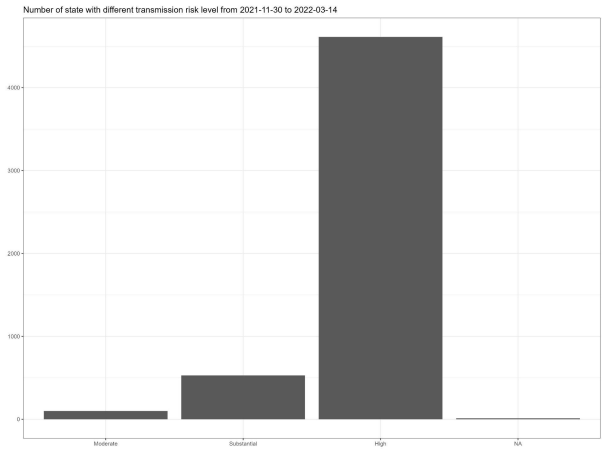
```
   state                date                                      type          median            mean
Length:6603        Min.   :2021-11-30   estimate                     :5459   Min.   :0.2622   Min.   :0.2697
Class :character   1st Qu.:2021-12-31   estimate based on partial data: 416   1st Qu.:0.7556   1st Qu.:0.7633
Mode  :character   Median :2022-02-01   forecast                     : 728   Median :0.8865   Median :0.9015
                   Mean   :2022-02-01                                        Mean   :0.9588   Mean   :0.9720
                   3rd Qu.:2022-03-05                                        3rd Qu.:1.0768   3rd Qu.:1.0827
                   Max.   :2022-04-05                                        Max.   :4.3365   Max.   :4.6928
       sd             lower_90          lower_50         lower_20         upper_20         upper_50
Min.   :0.02238   Min.   :0.07467   Min.   :0.1918   Min.   :0.2368   Min.   :0.2905   Min.   :0.3362
1st Qu.:0.06985   1st Qu.:0.60259   1st Qu.:0.6923   1st Qu.:0.7335   1st Qu.:0.7793   1st Qu.:0.8116
Median :0.09614   Median :0.71377   Median :0.8056   Median :0.8504   Median :0.9149   Median :0.9727
Mean   :0.14121   Mean   :0.76911   Mean   :0.8776   Mean   :0.9274   Mean   :0.9921   Mean   :1.0512
3rd Qu.:0.13919   3rd Qu.:0.91965   3rd Qu.:1.0116   3rd Qu.:1.0528   3rd Qu.:1.1093   3rd Qu.:1.1734
Max.   :2.02461   Max.   :2.23625   Max.   :3.2091   Max.   :3.8719   Max.   :4.8905   Max.   :5.9191
   upper_90
Min.   :0.4424
1st Qu.:0.8965
Median :1.1080
Mean   :1.2199
3rd Qu.:1.3602
Max.   :8.6069
```


3
2
1
0
Dec
Jan
Feb
Mar
Apr
group
estimate
estimate based on partial data
forecast

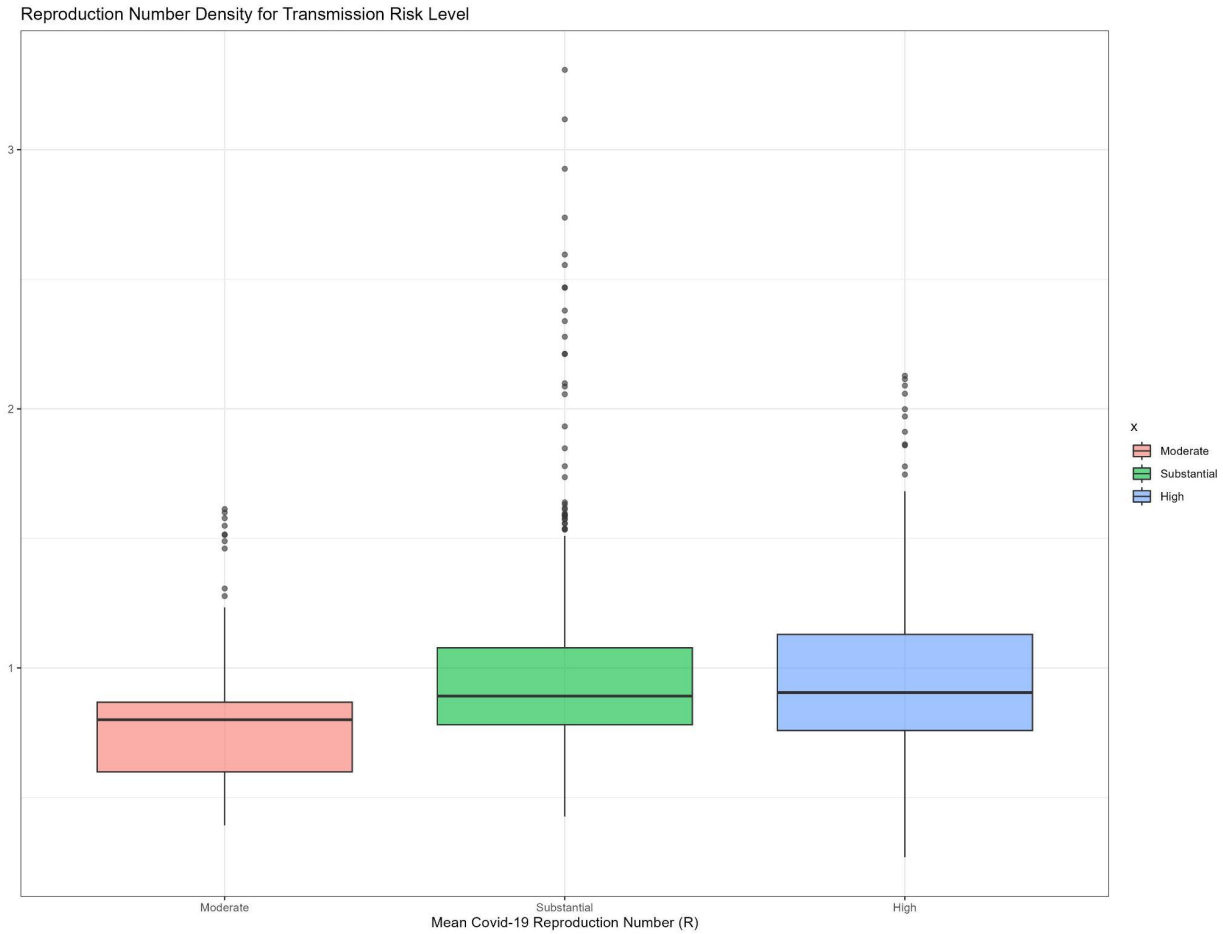

3
2
1
0
Dec
Jan
Feb
Mar
Apr
group
Alabama
Alaska
Arizona
Arkansas
California
Colorado
Connecticut
Delaware
District of Columbia
Florida
Georgia
Hawaii
Idaho
Illinois
Indiana
Iowa
Kansas
Kentucky
Louisiana
Maine
Maryland
Massachusetts
Michigan
Minnesota
Mississippi
Missouri
Montana
Nebraska
Nevada
New Hampshire
New Jersey
New Mexico
New York
North Carolina
North Dakota
Ohio
Oklahoma
Oregon
Pennsylvania
Puerto Rico
Rhode Island
South Carolina
South Dakota
Tennessee
Texas
Utah
Vermont
Virginia
Washington
West Virginia
Wisconsin
Wyoming

- Number of states with different transmission risk levels from 2021-11-30 to 2022-03-14
  There is no low level state in that specific time interrval

	CDCLevelCommTrans	n
	<fct>	<int>
1	Moderate	100
2	Substantial	529
3	High	4612
4	NA	12



- Tukey Method (the pairwise comparison of risk levels)




```
> fit = lm(mean ~ CDCLevelCommTrans, data=merged.data)
> TukeyHSD(aov(fit))
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = fit)

$CDCLevelCommTrans
                          diff         lwr          upr     p adj
Substantial-Moderate  0.17199589  0.10372531  0.240266466 0.0000000
High-Moderate         0.14187557  0.07817110  0.205580050 0.0000006
High-Substantial     -0.03012032 -0.05760711 -0.002633527 0.0276058
```