Figures, Tables, and Important Values

Ariane Stark

Data Editing

Important Values

Quartile Values

Spearman Rank Correlation Cases and Prop. Black Residentd

```
#race
cor(x=all_the_data_pull_cutoff$`total cases`,
    y=all_the_data_pull_cutoff$Total_Black/
      all_the_data_pull_cutoff$Total_Pop,
    method="spearman")
```

```
## [1] 0.4493985
```

```
cor.test(x=all_the_data_pull_cutoff$`total cases`,
    y=all_the_data_pull_cutoff$Total_Black/
      all_the_data_pull_cutoff$Total_Pop,
    method="spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  all_the_data_pull_cutoff$`total cases` and all_the_data_pull_cutoff$Total_Black/all_the_data_p
## S = 2779051013, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.4493985
```

Spearman Rank Correlation Cases and Prop. Households in Poverty

```
#poverty
x_case <- all_the_data_pull_cutoff$`total cases`
y_pov <- all_the_data_pull_cutoff$Total_Households_Below_Poverty/
      (all_the_data_pull_cutoff$Total_Households_Above_Poverty+
         all_the_data_pull_cutoff$Total_Households_Below_Poverty)
```

```r
cor(x=x_case[-c(1814)],
    y=y_pov[-c(1814)],
    method="spearman")
```

```
## [1] -0.04914091
```

```r
cor.test(x=x_case[-c(1814)],
    y=y_pov[-c(1814)],
    method="spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  x_case[-c(1814)] and y_pov[-c(1814)]
## S = 5290233922, p-value = 0.006076
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.04914091
```

MAE Wilcoxon Tests

Wilcoxon rank sum between proportion of black constituent quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_1_mae and quart_2_mae
## W = 173277, p-value < 2.2e-16
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_1_mae and quart_3_mae
## W = 101894, p-value < 2.2e-16
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_1_mae and quart_4_mae
## W = 108750, p-value < 2.2e-16
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_3_mae and quart_4_mae
## W = 306424, p-value = 0.7352
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 1 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_1_race_quart_1 and case_quart_1_race_quart_2
## W = 30054, p-value = 0.001876
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_1_race_quart_1 and case_quart_1_race_quart_3
## W = 14052, p-value = 2.005e-06
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_1_race_quart_1 and case_quart_1_race_quart_4
## W = 11432, p-value = 1.595e-10
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 2 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_2_race_quart_1 and case_quart_2_race_quart_2
## W = 22173, p-value = 0.02819
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_2_race_quart_1 and case_quart_2_race_quart_3
## W = 10671, p-value = 3.572e-07
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_2_race_quart_1 and case_quart_2_race_quart_4
## W = 16340, p-value = 0.0002577
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 3 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
```

```
## data:  case_quart_3_race_quart_1 and case_quart_3_race_quart_2
## W = 11556, p-value = 0.009097
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_3_race_quart_1 and case_quart_3_race_quart_3
## W = 8962, p-value = 4.621e-07
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_3_race_quart_1 and case_quart_3_race_quart_4
## W = 8308, p-value = 1.754e-05
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 4 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_4_race_quart_1 and case_quart_4_race_quart_2
## W = 2478, p-value = 0.3154
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_4_race_quart_1 and case_quart_4_race_quart_3
## W = 4146, p-value = 0.1204
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_4_race_quart_1 and case_quart_4_race_quart_4
## W = 3000, p-value = 0.0006125
## alternative hypothesis: true location shift is not equal to 0
```

RMAE Wilcoxon Tests

Wilcoxon rank sum between proportion of black constituent quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_1_rmae and quart_2_rmae
## W = 293350, p-value = 0.2393
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_1_rmae and quart_3_rmae
## W = 290569, p-value = 0.1363
## alternative hypothesis: true location shift is not equal to 0

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_1_rmae and quart_4_rmae
## W = 222676, p-value < 2.2e-16
## alternative hypothesis: true location shift is not equal to 0

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  quart_3_rmae and quart_4_rmae
## W = 230807, p-value = 2.887e-16
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 1 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_1_race_quart_1 and case_quart_1_race_quart_2
## W = 34835, p-value = 0.5787
## alternative hypothesis: true location shift is not equal to 0

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_1_race_quart_1 and case_quart_1_race_quart_3
## W = 21029, p-value = 0.5765
## alternative hypothesis: true location shift is not equal to 0

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_1_race_quart_1 and case_quart_1_race_quart_4
## W = 12860, p-value = 1.232e-07
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 2 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_2_race_quart_1 and case_quart_2_race_quart_2
## W = 26228, p-value = 0.4504
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_2_race_quart_1 and case_quart_2_race_quart_3
## W = 16610, p-value = 0.3087
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_2_race_quart_1 and case_quart_2_race_quart_4
## W = 16328, p-value = 0.0002473
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 3 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_3_race_quart_1 and case_quart_3_race_quart_2
## W = 14352, p-value = 0.6376
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_3_race_quart_1 and case_quart_3_race_quart_3
## W = 13443, p-value = 0.9551
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_3_race_quart_1 and case_quart_3_race_quart_4
## W = 8858, p-value = 0.0003257
## alternative hypothesis: true location shift is not equal to 0
```

Within Covid Case Quartile 4 Compare Across Race Quartiles

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_4_race_quart_1 and case_quart_4_race_quart_2
## W = 2616, p-value = 0.121
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_4_race_quart_1 and case_quart_4_race_quart_3
## W = 5553, p-value = 0.2804
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  case_quart_4_race_quart_1 and case_quart_4_race_quart_4
## W = 4364, p-value = 0.4485
## alternative hypothesis: true location shift is not equal to 0
```

Summary Statistics Case Quartile

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.271   4.763   8.144   8.474  11.987  23.814
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   53.32  104.81  159.10  289.31  298.12 6227.05
```

Summary Statistics By Race Quartile

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.271   6.373  13.941  24.617  27.390 671.542
```

```
##     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
##    2.068   24.559   51.458  157.423  152.212 4675.458
```

Regression Equations

```
regressionData<-data.frame(all_the_data_pull_cutoff$FIPS_CODE,
                           all_the_data_pull_cutoff$Total_Pop,
                 all_the_data_pull_cutoff$Total_Black/all_the_data_pull_cutoff$Total_Pop,
                 all_the_data_pull_cutoff$`total cases`,
                 all_the_data_pull_cutoff$Total_Households_Below_Poverty/
                   (all_the_data_pull_cutoff$Total_Households_Below_Poverty+
                      all_the_data_pull_cutoff$Total_Households_Above_Poverty),
                    (all_the_data_pull_cutoff$`total cases`/
                       all_the_data_pull_cutoff$Total_Pop)*1,
                 all_the_data_pull_cutoff$`1 week ahead MAE`,
                 all_the_data_pull_cutoff$`4 week ahead MAE`
                 ) %>%
  rename("Tot_Pop"=2) %>%
  rename("Tot_Cases" = 4) %>%
  rename("Prop_Black" = 3) %>%
  rename("Prop_Pov" = 5) %>%
  rename("Prop_Cases" = 6) %>%
  rename("One_Week_MAE" = 7) %>%
  rename("Four_Week_MAE" = 8)




summary(lm(One_Week_MAE~Tot_Cases,data = regressionData))
```

```
##
## Call:
## lm(formula = One_Week_MAE ~ Tot_Cases, data = regressionData)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1421.30   -14.82    -8.90     1.95  1639.73
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.972e+01  1.163e+00   16.95   <2e-16 ***
## Tot_Cases   3.914e-03  1.799e-05  217.55   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 62.16 on 3115 degrees of freedom
## Multiple R-squared:  0.9382, Adjusted R-squared:  0.9382
## F-statistic: 4.733e+04 on 1 and 3115 DF,  p-value: < 2.2e-16
```

```
summary(lm(One_Week_MAE~Tot_Cases+Prop_Black*Prop_Pov,
           data = regressionData))
```

```
##
## Call:
## lm(formula = One_Week_MAE ~ Tot_Cases + Prop_Black * Prop_Pov,
##     data = regressionData)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1385.83   -13.01    -8.07     2.38  1645.45
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.158e+01  3.530e+00   3.279  0.00105 **
## Tot_Cases            3.895e-03  1.827e-05 213.148  < 2e-16 ***
## Prop_Black           1.422e+02  2.428e+01   5.855 5.28e-09 ***
## Prop_Pov             3.261e+01  2.271e+01   1.436  0.15107
## Prop_Black:Prop_Pov -5.279e+02  1.018e+02  -5.184 2.32e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 61.85 on 3111 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.9389, Adjusted R-squared:  0.9389
## F-statistic: 1.196e+04 on 4 and 3111 DF,  p-value: < 2.2e-16
```

Figures

Figure 1

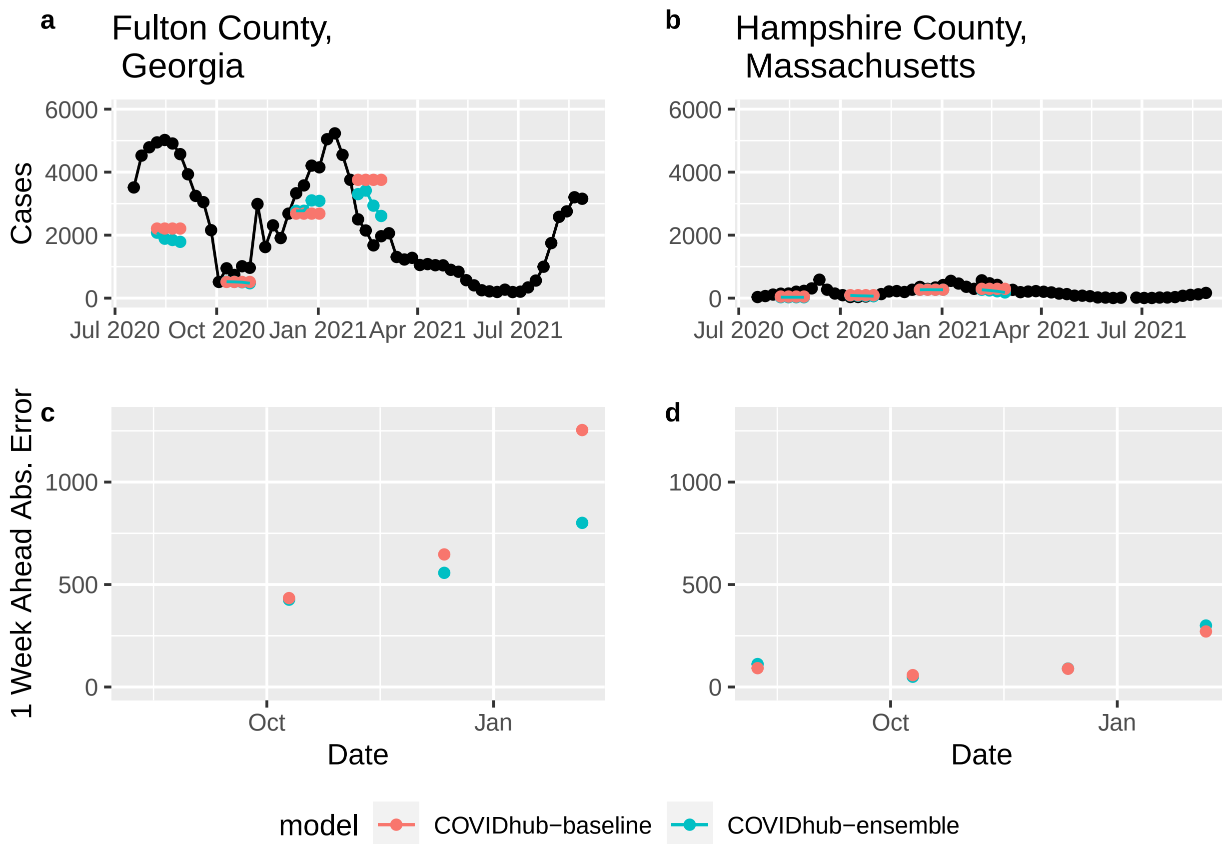


Figure 2

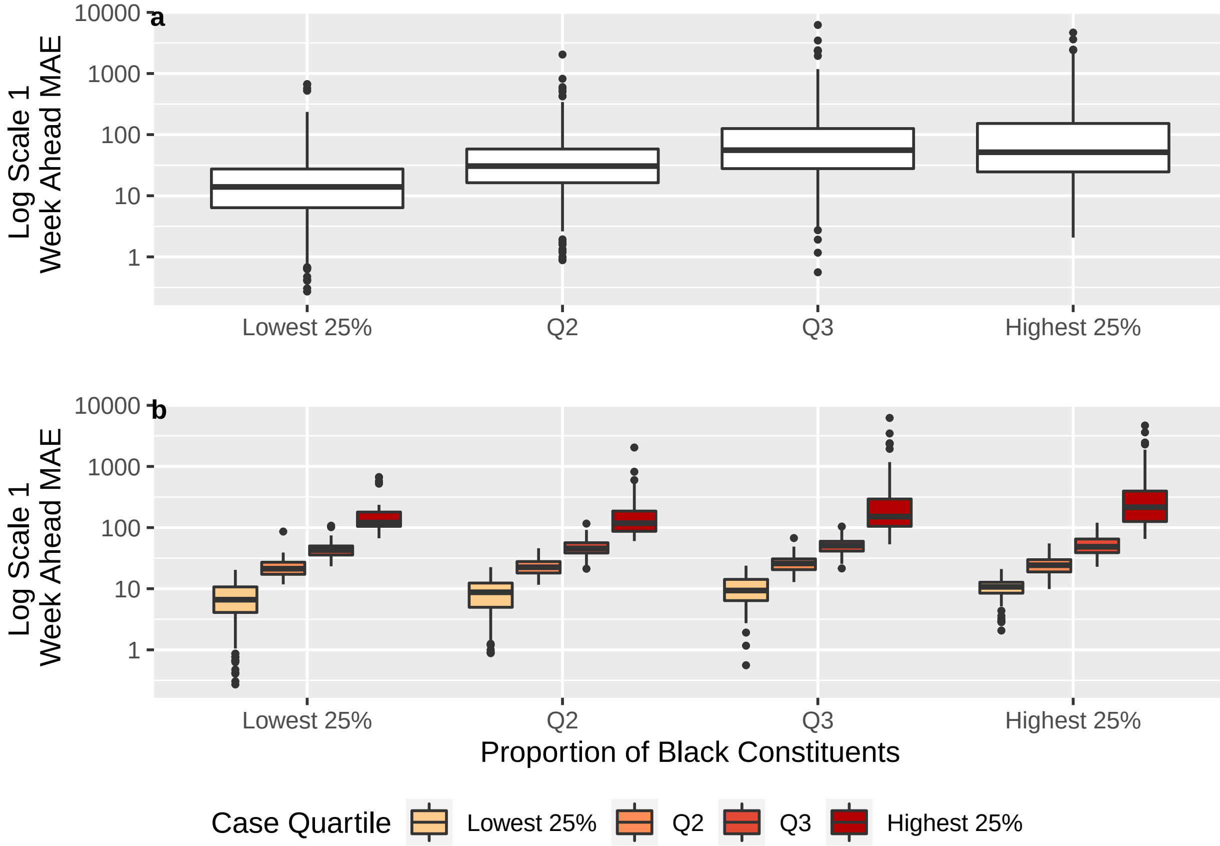

a
b
10000
1000
100
10
1
Log Scale 1
Week Ahead MAE
Lowest 25%
Q2
Q3
Highest 25%
Proportion of Black Constituents
Case Quartile
Lowest 25%
Q2
Q3
Highest 25%

Figure 3: 1 and 4 week MAE and RMAE for Racial Demographics

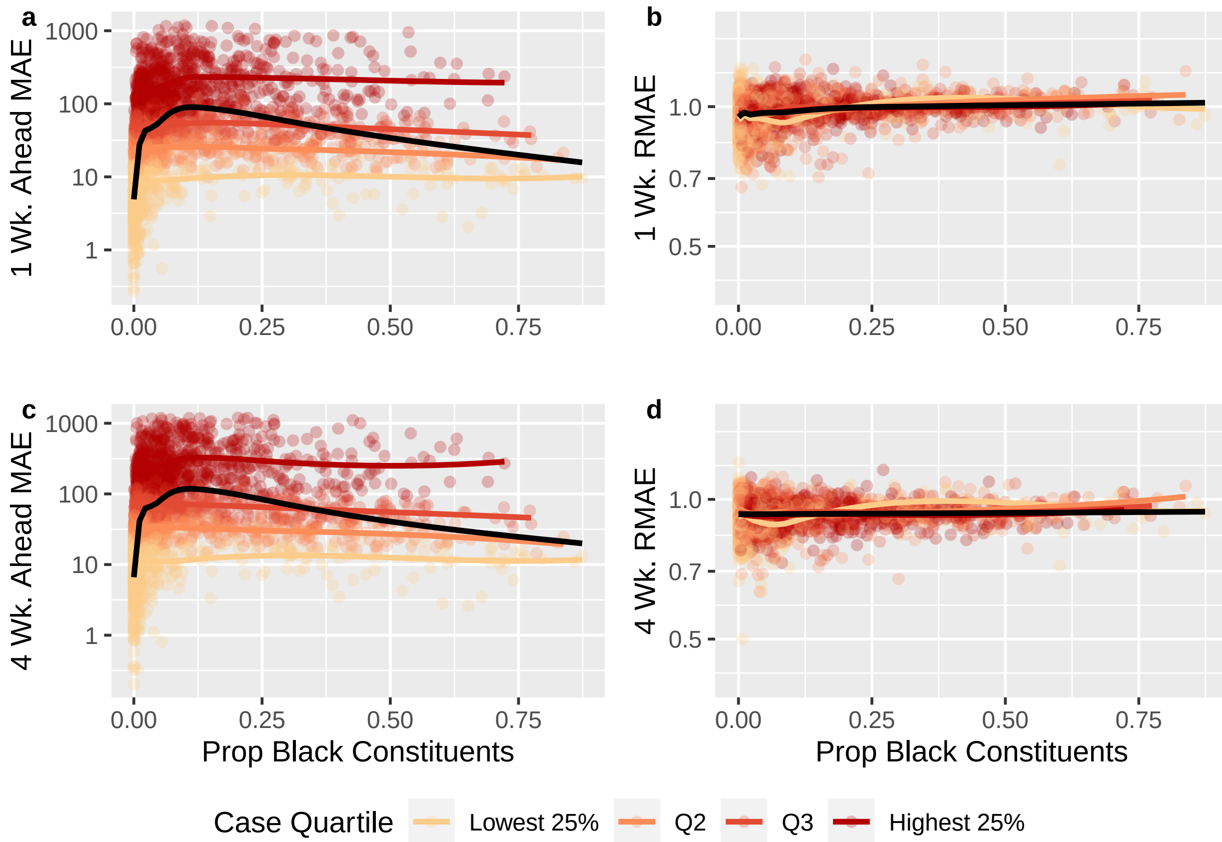
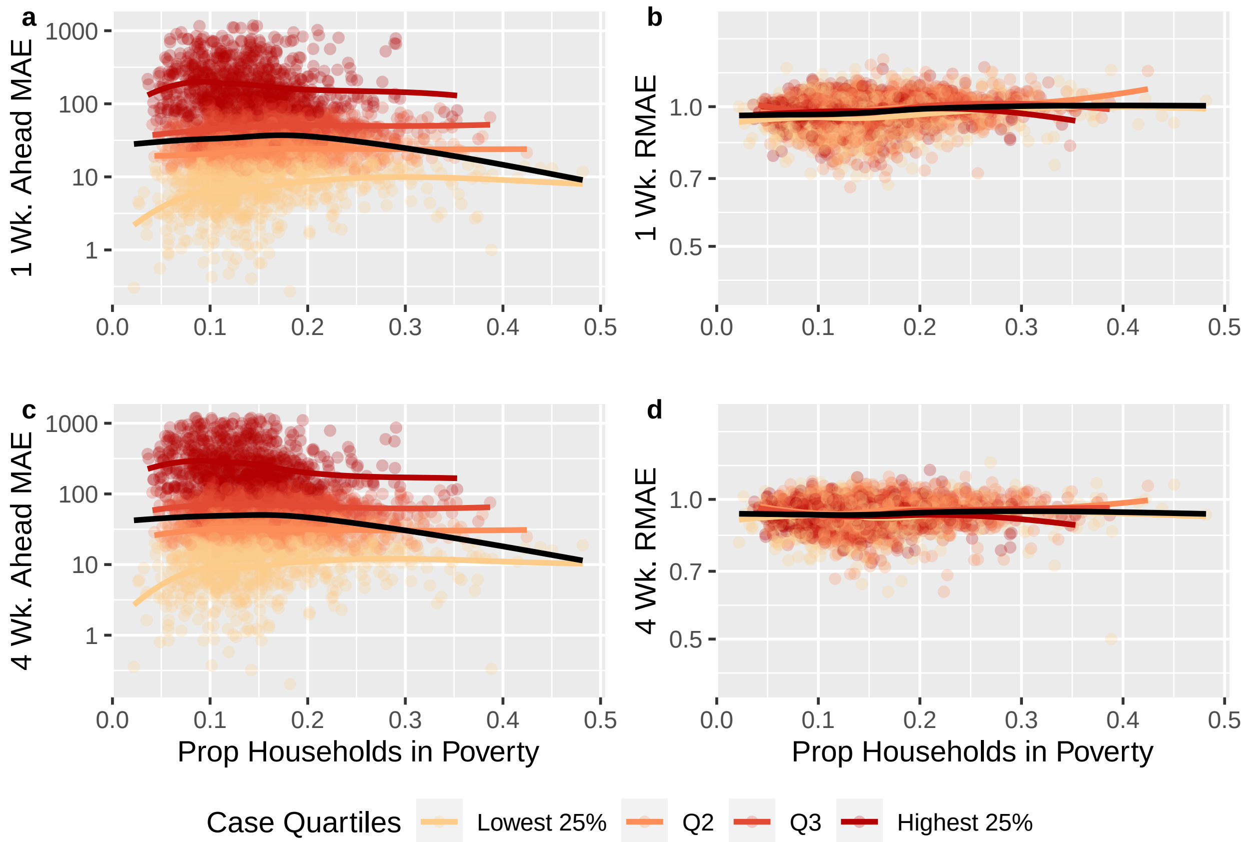


Figure 4: 1 and 4 week MAE and RMAE for Poverty



Tables

Table 1

	R Squared	Adj. R Squared
1 wk Ahead MAE as a funct. of Prop. Black and Prop. Pov	0.0472	0.0463
1 wk Ahead MAE as a funct. of Total Population	0.8711	0.8711
1 wk Ahead MAE as a funct. of Total Cases	0.9382	0.9382
1 wk Ahead MAE as a funct. of Cases and Minority Prop.	0.9389	0.9389
4 wk Ahead MAE as a funct. of Prop. Black and Prop. Pov	0.0405	0.0396
4 wk Ahead MAE as a funct. of Total Population	0.9437	0.9437
4 wk Ahead MAE as a funct. of Total Cases	0.9767	0.9767
4 wk Ahead MAE as a funct. of Cases and Minority Prop.	0.9768	0.9768

County Group based on Quartile of Total Cases		County Group Based on Quartile of % Black Residents				Overall
		Lowest 25%	Q2	Q3	Highest 25%	
		[0.0% , 0.7%]	(0.7% , 2.3%]	(2.3% , 10.3%]	(10.3% , 87.4%]	
Lowest 25%	[8,1945]	6.61	8.746	9.339	10.72	8.144
Q2	(1945,4906]	21.229	22.424	25.89	24.22	22.966
Q3	(4906,13002]	42.822	45.61	50.542	48.551	47.373
Highest 25%	(13002,1948984]	120.915	117.847	151.203	215.237	159.102
Overall		13.941	30.542	55.695	51.458	32.746

County Group based on Quartile of Total Cases		County Group Based on Quartile of % Black Residents				Overall
		Lowest 25%	Q2	Q3	Highest 25%	
		[0.0% , 0.7%]	(0.7% , 2.3%]	(2.3% , 10.3%]	(10.3% , 87.4%]	
Lowest 25%	[8,1945]	0.956	0.96	0.941	0.994	0.961
Q2	(1945,4906]	0.992	0.987	0.986	1.009	0.992
Q3	(4906,13002]	0.986	0.982	0.984	1.003	0.99
Highest 25%	(13002,1948984]	0.987	0.972	0.971	0.99	0.979
Overall		0.974	0.979	0.977	0.997	0.983

	Adj. R Squared
1 wk Ahead MAE as a funct. of Prop. Black and Prop. Pov	0.0463
1 wk Ahead MAE as a funct. of Total Population	0.8711
1 wk Ahead MAE as a funct. of Total Cases	0.9382
1 wk Ahead MAE as a funct. of Cases and Minority Prop.	0.9389
4 wk Ahead MAE as a funct. of Prop. Black and Prop. Pov	0.0396
4 wk Ahead MAE as a funct. of Total Population	0.9437
4 wk Ahead MAE as a funct. of Total Cases	0.9767
4 wk Ahead MAE as a funct. of Cases and Minority Prop.	0.9768

Table 2 - code

Table 2- MAE pvalues

Table 2 - LaTex

Table 3 - code

Table 3- RMAE pvalues

Table 3 - LaTex

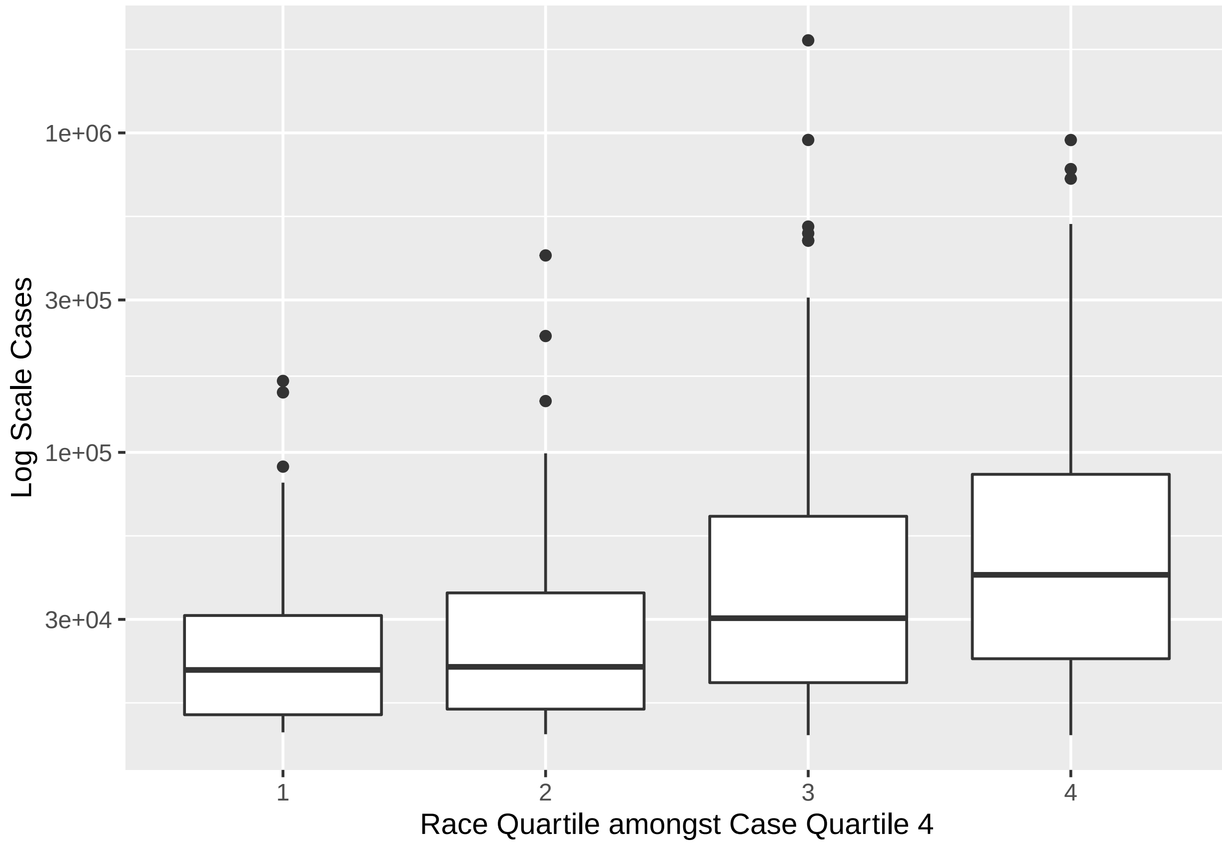
Potential Supplementary Materials

Counts by Quartile Quartile Breakdown

```
##              Prop. Black Quartile
## Case Quartile lowest 25%  Q2  Q3 highest 25% Overall
##   lowest 25%         410 175  99          96     780
##   Q2                 220 229 142         188     779
##   Q3                 118 236 227         198     779
##   highest 25%         32 139 311         297     779
##   Overall            780 779 779         779    3117
```

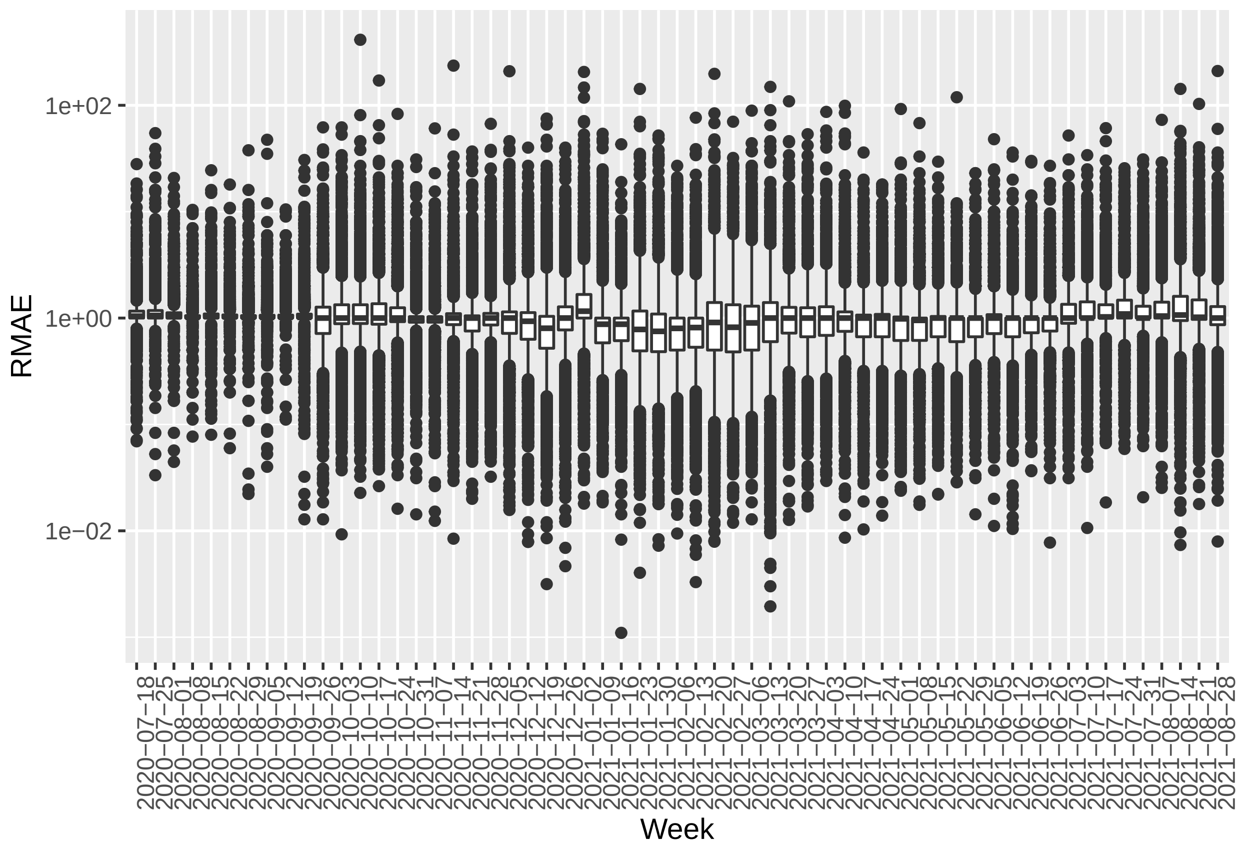
Figure 2 just 4th case quartile

```
all_the_data_pull_cutoff %>%
  filter(`Case Quartile`==4) %>%
  ggplot()+
  geom_boxplot(aes(x=as.factor(`black_prop_quart`),
      y=`total cases`)) +
  xlab("Race Quartile amongst Case Quartile 4") +
  ylab("Log Scale Cases") +
  scale_y_log10()
```

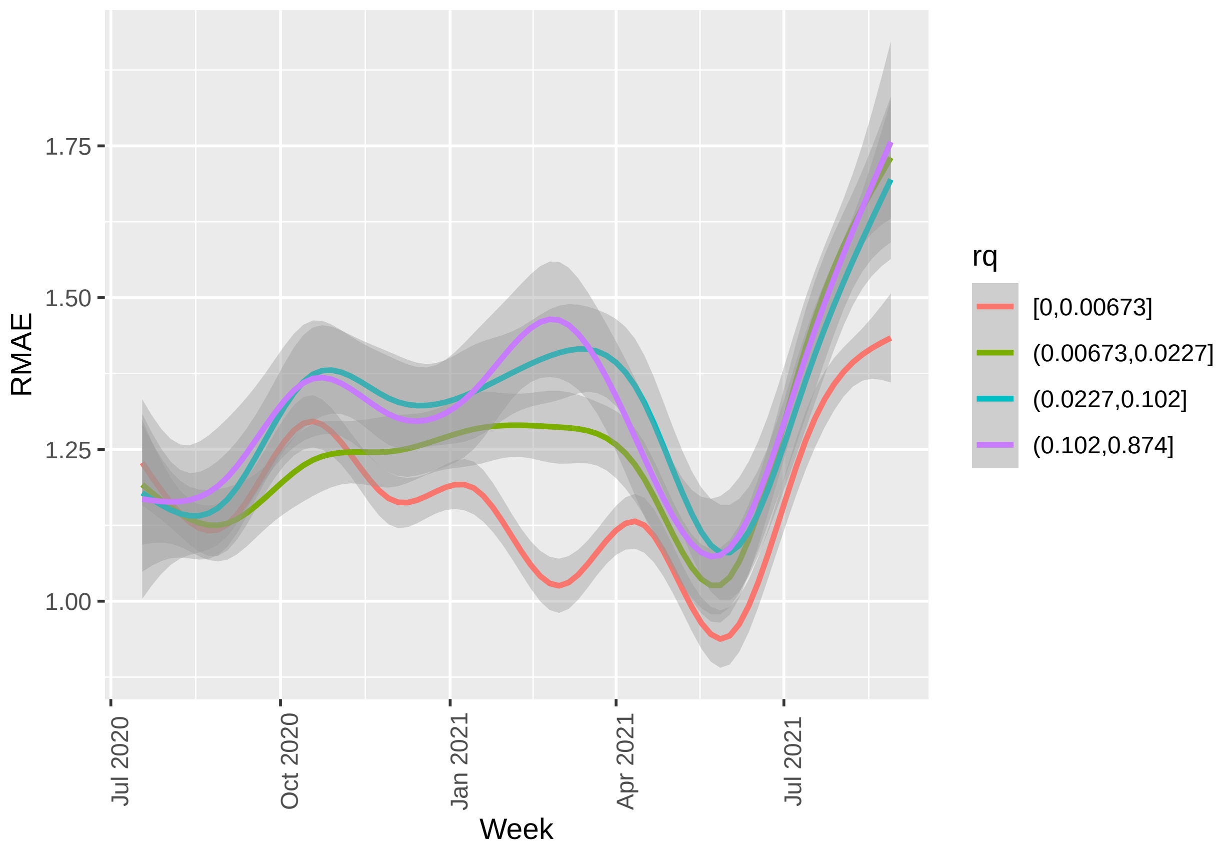


Exploratory Analysis of RMAE

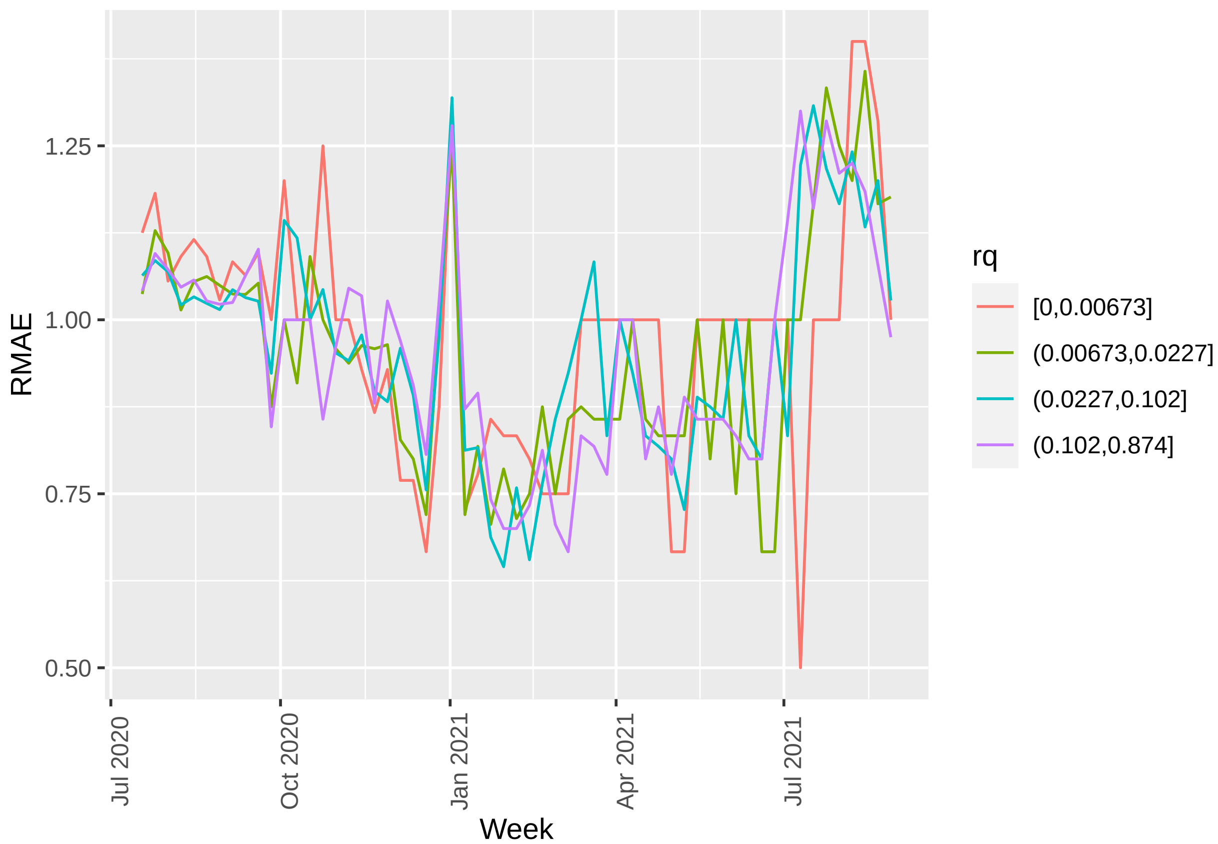
```
all_the_data_by_week_relative_pull_cutoff %>%
  ggplot(aes(x=as.factor(`last date of week`),
             y=`1 week ahead absolute error`/`1 week ahead absolute error base`))+
  geom_boxplot() +
  scale_y_log10()+
  theme(axis.text.x = element_text(angle = 90))+
  xlab("Week")+
  ylab("RMAE")
```



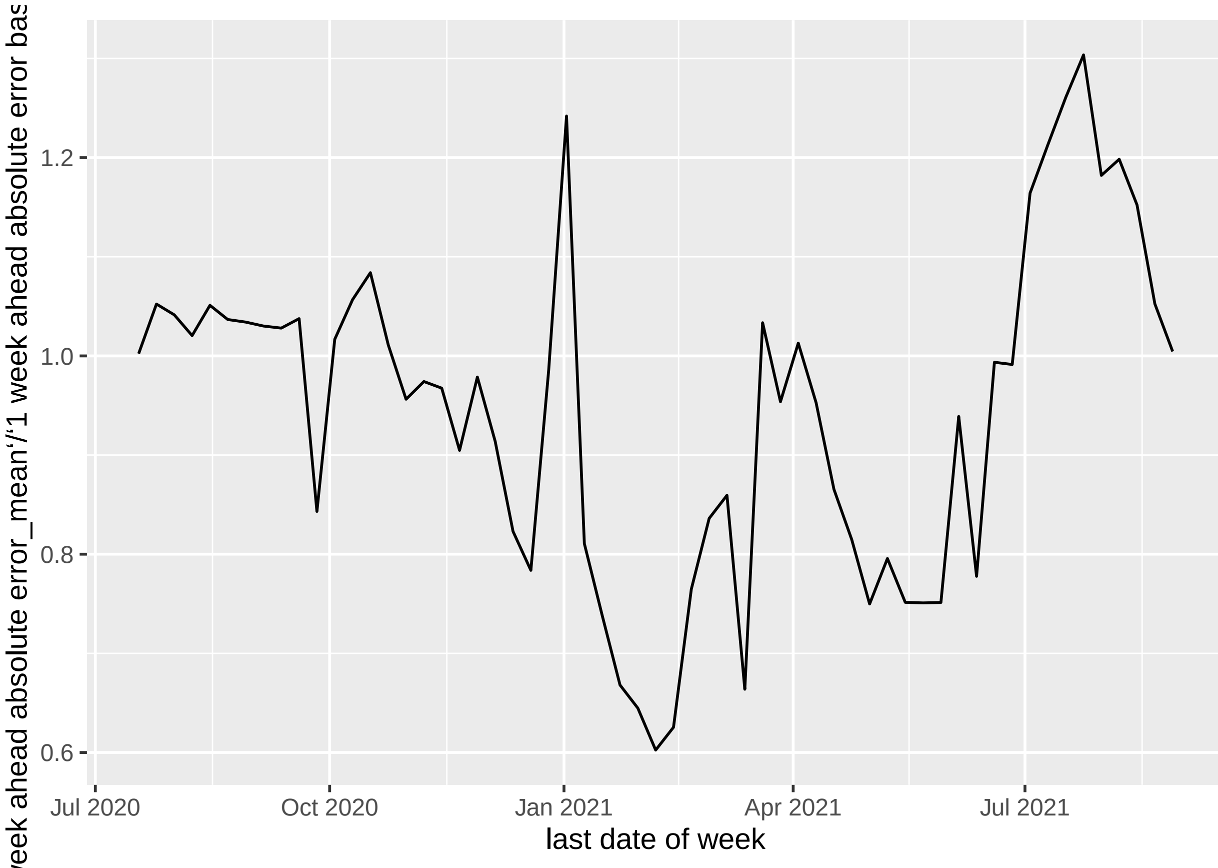
```
all_the_data_by_week_relative_pull_cutoff %>%
  mutate(rq = cut_number(Total_Black/Total_Pop,n=4)) %>%
  ggplot(aes(x=`last date of week`,
             y=`1 week ahead absolute error`/`1 week ahead absolute error base`,
             color = rq))+
  geom_smooth() +
  #scale_y_log10()+
  theme(axis.text.x = element_text(angle = 90))+
  xlab("Week")+
  ylab("RMAE")
```



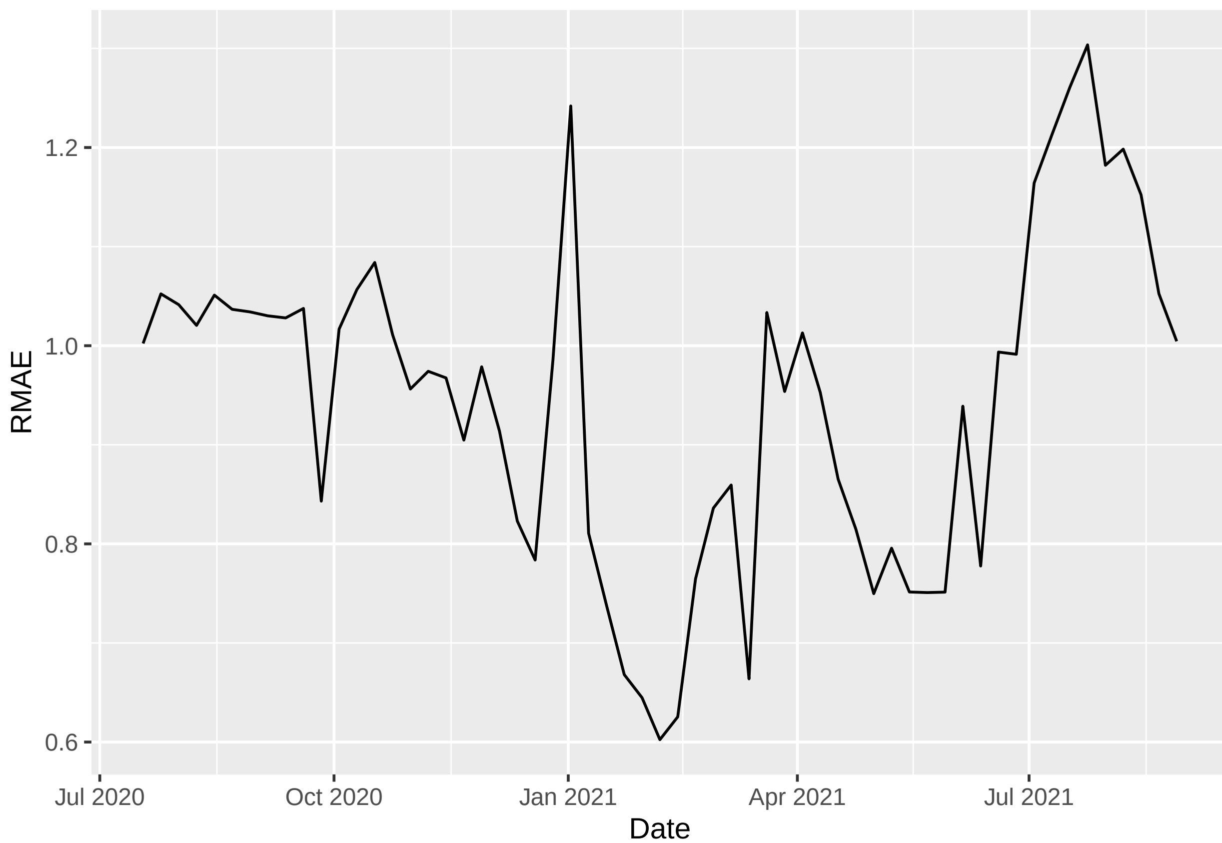

```
all_the_data_by_week_relative_pull_cutoff %>%
  mutate(rq = cut_number(Total_Black/Total_Pop,n=4)) %>%
  group_by(`last date of week`,rq) %>%
  mutate(median_e =median(`1 week ahead absolute error`,na.rm = TRUE),
         median_b =median(`1 week ahead absolute error base`,na.rm = TRUE)) %>%
  ungroup() %>%
  ggplot(aes(x=`last date of week`,
             y=median_e/median_b,
             color = rq))+
  geom_line() +
  #scale_y_log10()+
  theme(axis.text.x = element_text(angle = 90))+
  xlab("Week")+
  ylab("RMAE")
```




```
all_the_data_by_week_relative_pull_cutoff %>%
  group_by(`last date of week`) %>%
  summarise_at(vars("1 week ahead absolute error",
                    "1 week ahead absolute error base"),
               list("mean" = mean),
               na.rm = TRUE) %>%
  ggplot(aes(x=`last date of week`,y=`1 week ahead absolute error_mean`/
               `1 week ahead absolute error base_mean`))+
  geom_line()
```




```
all_the_data_by_week_relative_pull_cutoff %>%
  group_by(`last date of week`) %>%
  summarise_at(vars("1 week ahead absolute error",
                    "1 week ahead absolute error base"),
               list("sum" = sum),
               na.rm = TRUE) %>%
  ggplot(aes(x=`last date of week`,y=`1 week ahead absolute error_sum`/
               `1 week ahead absolute error base_sum`))+
  geom_line() +
  ylab("RMAE")+
  xlab("Date")
```



RMAE
0.6
0.8
1.0
1.2
Jul 2020
Oct 2020
Jan 2021
Apr 2021
Jul 2021
Date