project

Kavya Gajjar

```
folder <- "D:\\MS DS\\DS 5110\\project\\Final project\\data\\"
df<- read_csv(paste(folder,"us_state_vaccinations.csv",sep=""))
globe_data <- read_csv(paste(folder,"country_vaccinations.csv",sep=""))



data <- globe_data %>%
  filter(iso_code == 'USA') %>%
  filter(date <= '2021-03-30') %>%
  select(c('date','people_vaccinated'))

library(imputeTS)
data$people_vaccinated <- na.interpolation(data$people_vaccinated,
                                           option = "linear")



test_data <- globe_data %>%
  filter(iso_code == 'USA') %>%
  filter(date > '2021-03-30') %>%
  select(c('date','people_vaccinated'))

data$date <- strptime(data$date, "%Y-%m-%d" )
data$date <- as.POSIXct(data$date)
data$people_vaccinated <- as.numeric(data$people_vaccinated)

myts<- ts(data$people_vaccinated, frequency = 1,
          start = as.Date('2020-12-20'), end = as.Date('2021-03-30'))

my_df_ts <- data.frame(vaccination = myts, as.numeric(time(myts)))
names(my_df_ts) <- c("vaccination", "time")



# Then we can create a model using tslm
# We can model using trend, season and random
mymodel <- tslm(vaccination ~ trend,my_df_ts)

# And forecast using this same model
# We are going to predict the next 10 years
# We can see the ascending trend
test_fc <- forecast(mymodel,h=12)
test_fc
```

```
##       Point Forecast    Lo 80     Hi 80    Lo 95     Hi 95
## 18717       84127578 77089160  91165997 73302771  94952386
## 18718       85071409 78028933  92113886 74240361  95902458
## 18719       86015240 78968630  93061851 75177834  96852647
## 18720       86959071 79908250  94009892 76115189  97802953
## 18721       87902902 80847794  94958010 77052427  98753377
## 18722       88846733 81787262  95906203 77989548  99703917
## 18723       89790564 82726655  96854473 78926553 100654574
## 18724       90734395 83665972  97802818 79863441 101605348
## 18725       91678226 84605213  98751238 80800213 102556238
## 18726       92622056 85544378  99699734 81736869 103507243
## 18727       93565887 86483469 100648306 82673410 104458365
## 18728       94509718 87422484 101596952 83609835 105409602
```

```
test_data$pred <- test_fc$mean


RMSE(test_data$people_vaccinated, test_data$pred)
```

```
## [1] 19136010
```

```
my_fc <- forecast(mymodel, h=200)
my_fc
```

```
##       Point Forecast     Lo 80     Hi 80     Lo 95     Hi 95
## 18717       84127578  77089160  91165997  73302771  94952386
## 18718       85071409  78028933  92113886  74240361  95902458
## 18719       86015240  78968630  93061851  75177834  96852647
## 18720       86959071  79908250  94009892  76115189  97802953
## 18721       87902902  80847794  94958010  77052427  98753377
## 18722       88846733  81787262  95906203  77989548  99703917
## 18723       89790564  82726655  96854473  78926553 100654574
## 18724       90734395  83665972  97802818  79863441 101605348
## 18725       91678226  84605213  98751238  80800213 102556238
## 18726       92622056  85544378  99699734  81736869 103507243
## 18727       93565887  86483469 100648306  82673410 104458365
## 18728       94509718  87422484 101596952  83609835 105409602
## 18729       95453549  88361425 102545673  84546144 106360954
## 18730       96397380  89300291 103494469  85482339 107312421
## 18731       97341211  90239082 104443340  86418419 108264003
## 18732       98285042  91177798 105392285  87354384 109215700
## 18733       99228873  92116440 106341305  88290235 110167510
## 18734      100172704  93055008 107290399  89225972 111119435
## 18735      101116534  93993502 108239566  90161595 112071474
## 18736      102060365  94931923 109188808  91097104 113023626
## 18737      103004196  95870269 110138123  92032501 113975892
## 18738      103948027  96808542 111087512  92967784 114928270
## 18739      104891858  97746742 112036974  93902954 115880762
## 18740      105835689  98684869 112986509  94838012 116833365
## 18741      106779520  99622922 113936117  95772958 117786081
## 18742      107723351 100560903 114885798  96707792 118738909
## 18743      108667182 101498812 115835551  97642514 119691849
## 18744      109611012 102436647 116785377  98577125 120644900
```

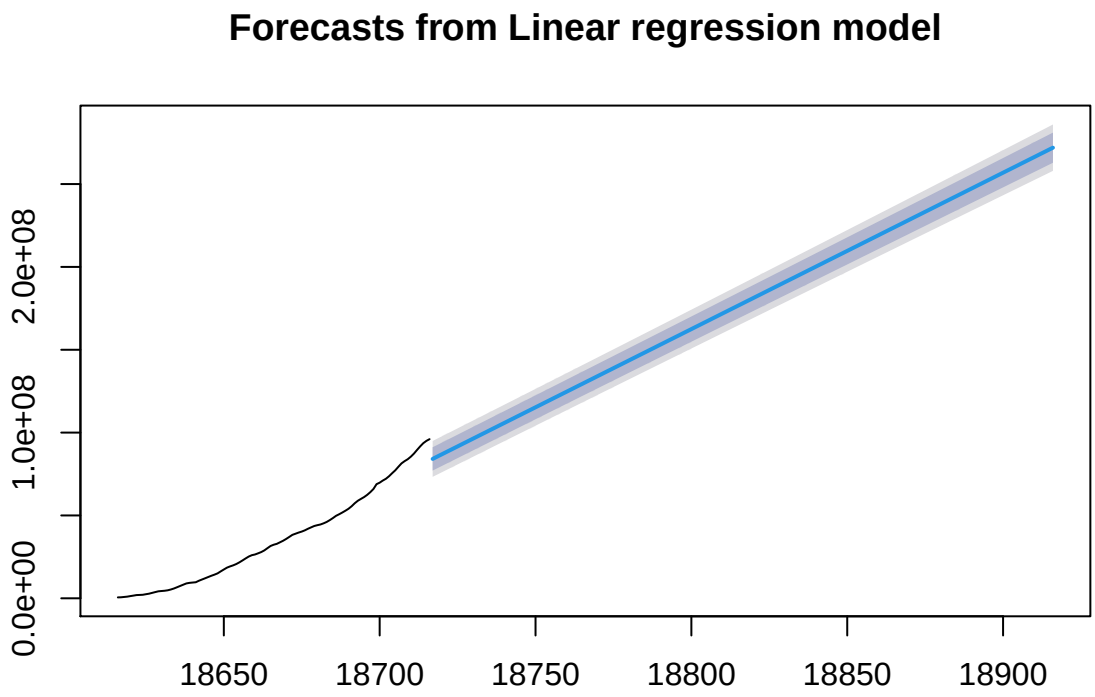
```
## 18745      110554843 103374411 117735276  99511624 121598062
## 18746      111498674 104312103 118685246 100446013 122551335
## 18747      112442505 105249722 119635288 101380291 123504719
## 18748      113386336 106187270 120585402 102314459 124458213
## 18749      114330167 107124746 121535587 103248517 125411817
## 18750      115273998 108062152 122485844 104182465 126365530
## 18751      116217829 108999486 123436172 105116304 127319353
## 18752      117161660 109936749 124386570 106050034 128273285
## 18753      118105490 110873941 125337040 106983655 129227326
## 18754      119049321 111811063 126287580 107917167 130181475
## 18755      119993152 112748114 127238190 108850572 131135733
## 18756      120936983 113685095 128188871 109783868 132090098
## 18757      121880814 114622006 129139622 110717057 133044571
## 18758      122824645 115558848 130090442 111650138 133999152
## 18759      123768476 116495620 131041332 112583113 134953839
## 18760      124712307 117432322 131992291 113515980 135908633
## 18761      125656138 118368956 132943319 114448742 136863533
## 18762      126599968 119305520 133894417 115381397 137818540
## 18763      127543799 120242016 134845583 116313947 138773652
## 18764      128487630 121178443 135796817 117246391 139728869
## 18765      129431461 122114802 136748120 118178730 140684192
## 18766      130375292 123051093 137699491 119110965 141639619
## 18767      131319123 123987315 138650930 120043095 142595151
## 18768      132262954 124923471 139602437 120975121 143550787
## 18769      133206785 125859558 140554011 121907043 144506526
## 18770      134150616 126795579 141505652 122838861 145462370
## 18771      135094446 127731532 142457361 123770577 146418316
## 18772      136038277 128667418 143409136 124702190 147374365
## 18773      136982108 129603238 144360978 125633700 148330517
## 18774      137925939 130538992 145312887 126565108 149286770
## 18775      138869770 131474679 146264861 127496414 150243126
## 18776      139813601 132410300 147216902 128427619 151199583
## 18777      140757432 133345856 148169008 129358723 152156141
## 18778      141701263 134281346 149121180 130289726 153112800
## 18779      142645094 135216770 150073417 131220628 154069559
## 18780      143588924 136152130 151025719 132151430 155026419
## 18781      144532755 137087425 151978086 133082133 155983378
## 18782      145476586 138022655 152930518 134012736 156940436
## 18783      146420417 138957820 153883014 134943240 157897594
## 18784      147364248 139892922 154835574 135873645 158854851
## 18785      148308079 140827959 155788199 136803952 159812205
## 18786      149251910 141762933 156740887 137734161 160769658
## 18787      150195741 142697843 157693638 138664272 161727209
## 18788      151139572 143632690 158646453 139594286 162684857
## 18789      152083402 144567474 159599331 140524203 163642602
## 18790      153027233 145502195 160552272 141454023 164600443
## 18791      153971064 146436853 161505276 142383747 165558381
## 18792      154914895 147371449 162458341 143313375 166516415
## 18793      155858726 148305983 163411469 144242907 167474545
## 18794      156802557 149240455 164364659 145172344 168432769
## 18795      157746388 150174865 165317911 146101687 169391089
## 18796      158690219 151109213 166271224 147030934 170349503
## 18797      159634050 152043501 167224598 147960088 171308011
## 18798      160577880 152977727 168178034 148889147 172266613
```

```
## 18799      161521711 153911893 169131530 149818114 173225309
## 18800      162465542 154845998 170085086 150746987 174184098
## 18801      163409373 155780043 171038703 151675767 175142979
## 18802      164353204 156714028 171992380 152604455 176101953
## 18803      165297035 157647953 172946117 153533051 177061019
## 18804      166240866 158581818 173899914 154461555 178020177
## 18805      167184697 159515624 174853770 155389968 178979426
## 18806      168128528 160449370 175807685 156318290 179938766
## 18807      169072358 161383058 176761659 157246521 180898196
## 18808      170016189 162316687 177715691 158174662 181857717
## 18809      170960020 163250258 178669782 159102713 182817328
## 18810      171903851 164183771 179623932 160030674 183777028
## 18811      172847682 165117225 180578139 160958547 184736817
## 18812      173791513 166050622 181532404 161886330 185696696
## 18813      174735344 166983961 182486727 162814025 186656662
## 18814      175679175 167917243 183441107 163741632 187616717
## 18815      176623006 168850468 184395544 164669151 188576860
## 18816      177566836 169783636 185350037 165596583 189537090
## 18817      178510667 170716747 186304588 166523928 190497407
## 18818      179454498 171649802 187259194 167451186 191457810
## 18819      180398329 172582801 188213857 168378358 192418300
## 18820      181342160 173515744 189168576 169305444 193378876
## 18821      182285991 174448631 190123350 170232444 194339538
## 18822      183229822 175381463 191078180 171159359 195300285
## 18823      184173653 176314240 192033065 172086189 196261116
## 18824      185117484 177246962 192988005 173012935 197222032
## 18825      186061314 178179629 193943000 173939596 198183033
## 18826      187005145 179112242 194898049 174866174 199144117
## 18827      187948976 180044800 195853152 175792668 200105285
## 18828      188892807 180977304 196808310 176719079 201066535
## 18829      189836638 181909755 197763521 177645407 202027869
## 18830      190780469 182842152 198718786 178571653 202989285
## 18831      191724300 183774495 199674104 179497817 203950782
## 18832      192668131 184706786 200629475 180423899 204912362
## 18833      193611962 185639024 201584900 181349900 205874023
## 18834      194555792 186571209 202540376 182275820 206835765
## 18835      195499623 187503341 203495906 183201660 207797587
## 18836      196443454 188435422 204451487 184127419 208759490
## 18837      197387285 189367450 205407120 185053098 209721472
## 18838      198331116 190299427 206362805 185978698 210683535
## 18839      199274947 191231352 207318542 186904218 211645676
## 18840      200218778 192163226 208274330 187829660 212607896
## 18841      201162609 193095049 209230169 188755023 213570195
## 18842      202106440 194026821 210186058 189680308 214532572
## 18843      203050270 194958542 211141999 190605515 215495026
## 18844      203994101 195890213 212097989 191530644 216457558
## 18845      204937932 196821834 213054030 192455697 217420168
## 18846      205881763 197753406 214010121 193380673 218382854
## 18847      206825594 198684927 214966261 194305572 219345616
## 18848      207769425 199616399 215922451 195230395 220308454
## 18849      208713256 200547822 216878690 196155143 221271369
## 18850      209657087 201479195 217834978 197079815 222234358
## 18851      210600918 202410520 218791315 198004413 223197423
## 18852      211544748 203341797 219747700 198928935 224160562
```

```
## 18853      212488579 204273025 220704134 199853383 225123775
## 18854      213432410 205204205 221660616 200777757 226087063
## 18855      214376241 206135337 222617146 201702058 227050424
## 18856      215320072 207066421 223573723 202626285 228013859
## 18857      216263903 207997458 224530348 203550439 228977366
## 18858      217207734 208928448 225487020 204474521 229940947
## 18859      218151565 209859390 226443739 205398530 230904599
## 18860      219095396 210790286 227400505 206322468 231868324
## 18861      220039226 211721135 228357318 207246333 232832120
## 18862      220983057 212651938 229314176 208170128 233795987
## 18863      221926888 213582695 230271082 209093851 234759925
## 18864      222870719 214513406 231228032 210017504 235723934
## 18865      223814550 215444071 232185029 210941086 236688014
## 18866      224758381 216374691 233142071 211864599 237652163
## 18867      225702212 217305265 234099159 212788042 238616382
## 18868      226646043 218235794 235056291 213711415 239580670
## 18869      227589874 219166279 236013468 214634720 240545027
## 18870      228533704 220096719 236970690 215557956 241509453
## 18871      229477535 221027114 237927957 216481123 242473947
## 18872      230421366 221957465 238885267 217404223 243438509
## 18873      231365197 222887772 239842622 218327255 244403139
## 18874      232309028 223818036 240800020 219250220 245367836
## 18875      233252859 224748255 241757462 220173117 246332601
## 18876      234196690 225678432 242714948 221095948 247297432
## 18877      235140521 226608565 243672476 222018712 248262329
## 18878      236084352 227538655 244630048 222941411 249227292
## 18879      237028182 228468703 245587662 223864043 250192322
## 18880      237972013 229398708 246545319 224786611 251157416
## 18881      238915844 230328670 247503018 225709113 252122576
## 18882      239859675 231258591 248460759 226631550 253087800
## 18883      240803506 232188470 249418543 227553923 254053089
## 18884      241747337 233118306 250376368 228476231 255018443
## 18885      242691168 234048102 251334234 229398476 255983860
## 18886      243634999 234977856 252292142 230320657 256949340
## 18887      244578830 235907569 253250091 231242775 257914884
## 18888      245522660 236837240 254208080 232164830 258880491
## 18889      246466491 237766872 255166111 233086823 259846160
## 18890      247410322 238696462 256124182 234008753 260811892
## 18891      248354153 239626013 257082294 234930621 261777686
## 18892      249297984 240555523 258040445 235852427 262743541
## 18893      250241815 241484993 258998637 236774172 263709458
## 18894      251185646 242414424 259956868 237695856 264675436
## 18895      252129477 243343815 260915138 238617479 265641475
## 18896      253073308 244273167 261873449 239539041 266607574
## 18897      254017138 245202479 262831798 240460543 267573734
## 18898      254960969 246131753 263790186 241381986 268539953
## 18899      255904800 247060988 264748613 242303368 269506232
## 18900      256848631 247990184 265707078 243224692 270472571
## 18901      257792462 248919342 266665582 244145956 271438968
## 18902      258736293 249848461 267624124 245067162 272405424
## 18903      259680124 250777543 268582705 245988309 273371939
## 18904      260623955 251706587 269541322 246909398 274338511
## 18905      261567786 252635593 270499978 247830429 275305142
## 18906      262511616 253564562 271458671 248751403 276271830
```

```
## 18907      263455447 254493493 272417401 249672319 277238576
## 18908      264399278 255422388 273376169 250593178 278205378
## 18909      265343109 256351246 274334973 251513981 279172238
## 18910      266286940 257280066 275293814 252434727 280139153
## 18911      267230771 258208851 276252691 253355417 281106125
## 18912      268174602 259137599 277211605 254276051 282073153
## 18913      269118433 260066311 278170555 255196629 283040236
## 18914      270062264 260994987 279129541 256117153 284007375
## 18915      271006095 261923627 280088562 257037621 284974568
## 18916      271949925 262852231 281047620 257958034 285941817
```

```r
plot(my_fc)
```

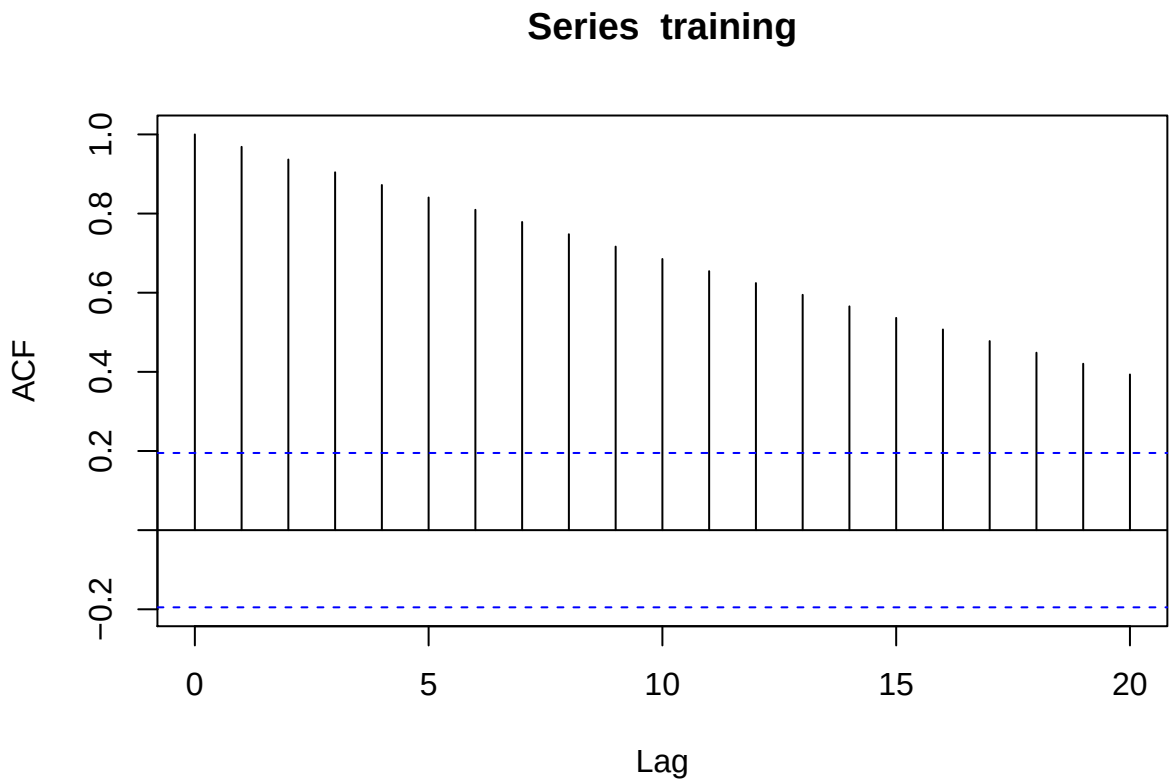



```r
pop <- 331400000
herd <- pop * 0.9


num = 0
for(i in 1:length(my_fc$mean)){
  if (my_fc$mean[i] >= herd){
    num = i
    break
  }
}
```

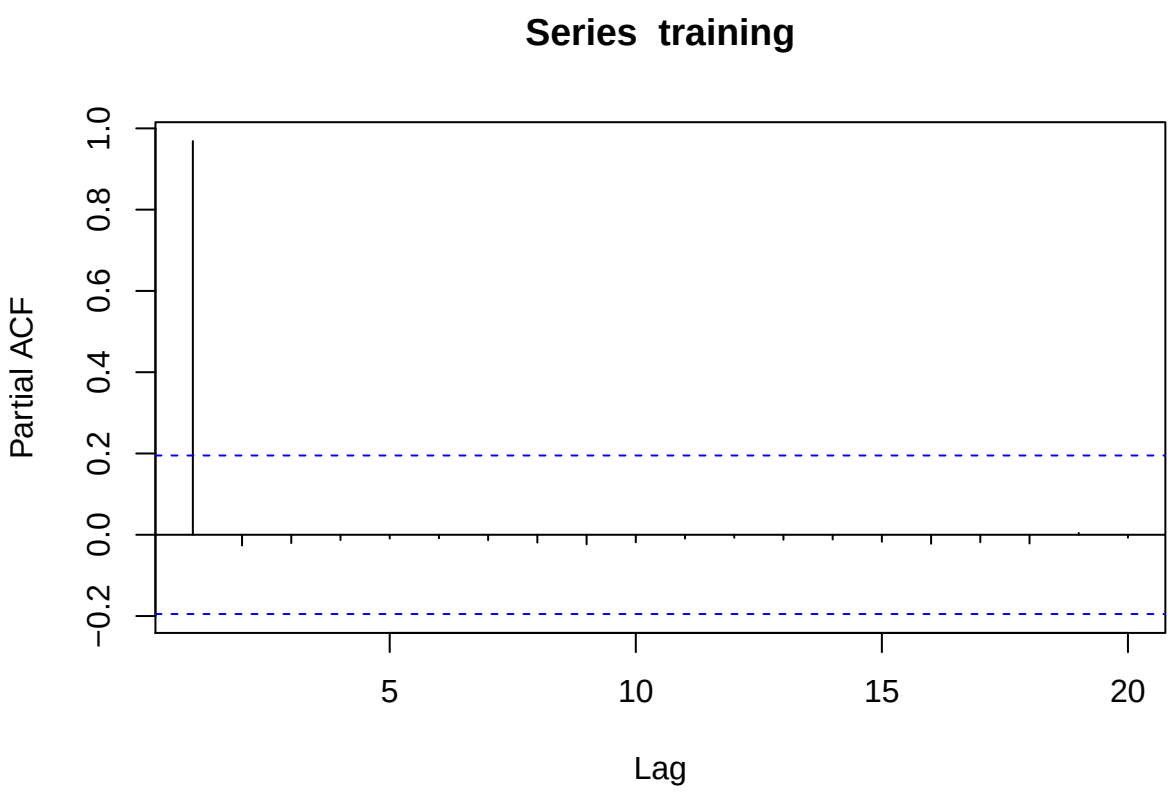
```r
date <- as.Date('2021-04-11')
final <- date + num
final = format(final, "%d,%B, %Y")
print(paste("Herd Immunity can be achieved by ",
            final, sep=""))
```

```
## [1] "Herd Immunity can be achieved by 11,April, 2021"
```

```r
library(forecast)
library(MLmetrics)

training <- window(data$people_vaccinated)

#90- train 10- test
#rmse
acf(training)
```



```
pacf(training)
```



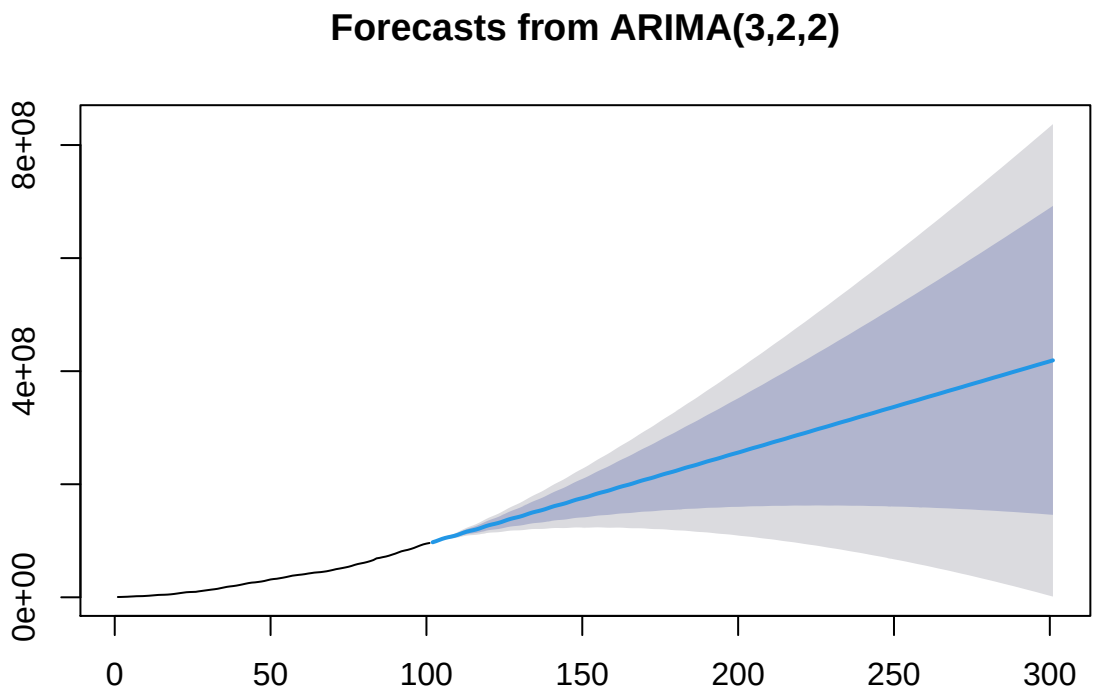

```
arima_optimal = auto.arima(training)

test_fc <- forecast(arima_optimal,h=12)

test_data$pred <- test_fc$mean
test_data
```

```
## # A tibble: 12 x 3
##    date       people_vaccinated       pred
##    <date>                 <dbl>      <dbl>
##  1 2021-03-31          97593290  97385821.
##  2 2021-04-01          99565311  99127463.
##  3 2021-04-02         101804762 101213836.
##  4 2021-04-03         104213478 103265812.
##  5 2021-04-04         106214924 104972249.
##  6 2021-04-05         107515428 106268977.
##  7 2021-04-06         108301234 107402121.
##  8 2021-04-07         109995734 108728746.
##  9 2021-04-08         112046611 110455952.
## 10 2021-04-09         114436039 112490011.
## 11 2021-04-10         117142879 114513988.
## 12 2021-04-11         119242902 116226242.
```

```
RMSE(test_data$people_vaccinated, test_data$pred)
```

```
## [1] 1563172
```

```
my_fc <- forecast(arima_optimal, h = 200)
plot(my_fc)
```




```
pop <- 331400000
herd <- pop * 0.9


for(i in 1:length(my_fc$mean)){
  if (my_fc$mean[i] >= herd){
    num = i
    break
  }
}


arima_optimal
```

```
## Series: training
## ARIMA(3,2,2)
##
```

```
## Coefficients:
##          ar1      ar2      ar3      ma1     ma2
##       0.7767  -0.4027  -0.4502  -1.2620  0.7975
## s.e.  0.1003   0.1197   0.0952   0.0666  0.0991
##
## sigma^2 estimated as 6.766e+10:  log likelihood=-1373.98
## AIC=2759.95   AICc=2760.87   BIC=2775.52
```

```
# imputation
# seasonality and trend
# model working
# p,d, q , arima
# sarima
# tslm doesn't work for us!
date <- as.Date('2021-04-11')
final <- date + num
final = format(final, "%d,%B, %Y")
print(paste("Herd Immunity can be achieved by ",
            final, sep=""))
```

```
## [1] "Herd Immunity can be achieved by 15,August, 2021"
```