Skeleton_of_Outputs

3/19/2020

This is a document for a code skeleton for the outputs section of the project.

Loading in relevant data sources for this:

```
#setwd('~/Dropbox/COVID-BaltimoreCity//')

require(socialmixr)
require(xlsx)
require(magrittr)
require(stringr)
require(reshape2)
require(dplyr)
require(ggplot2)
require(tidyr)
#load in data, subset incident cases/day (eventually should be symptomatic incident cases)
data_infected<- read.csv("SEIR_results_test.csv")

data_incident<- data_infected%>%
  select(incid1,incid2,incid3,incid4,incid5,incid6,incid7,incid8,incid9,incid10,incid11,
         incid12,incid13,incid14)

#load some sort of matrix that designates probabilities for symptomatics being severe,
#then conditional on being severe the probability of being hospitalized, or dead, then
#conditional on being hospitalized the prob of being in ICU and then conditional on being in ICU
#the prob of being on a vent and then a global probability of death in each category.

prob_severe<- c(seq(from=0.05, by = 0.05, length=12),max(seq(from=0.05, by = 0.05, length=12)),
                      max(seq(from=0.05, by = 0.05, length=12)))
prob_death<- c(0.002, 0.0020, 0.002, 0.004, 0.004,
               0.004, 0.013,0.013,0.036,0.036,0.08,
               0.15,0.036,0.08)
prob_hospitalized<- c(seq(from=0.3, by = 0.05, length=12),max(seq(from=0.3, by = 0.05, length=12)),
                      max(seq(from=0.3, by = 0.05, length=12)))
prob_ICU<- c(seq(from=0.1, by = 0.05, length=12),max(seq(from=0.3, by = 0.05, length=12)),
                      max(seq(from=0.3, by = 0.05, length=12)))
prob_vent<-c(seq(from=0.1, by = 0.05, length=12),max(seq(from=0.3, by = 0.05, length=12)),
                      max(seq(from=0.3, by = 0.05, length=12)))

prob_matrix<- cbind(prob_severe,prob_death,prob_hospitalized,prob_ICU,
                    prob_vent)
```

Now calculate the number in of severe, hospitalized, ICU and vented infections at each timepoint (day) in each group (age groups 1-12 and healthcare workers and homeless).

```
#calculate number of incident cases that get classified as severe,  of those,
#number that die, of those remaining, those thatget hospitalized , of that then those that get classifi
#this is for each age group (+ HCW and homeless)
#note this categorization is NOT a disease progression, a person gets assigned based on the worst state

counts_array<- array(NA, dim=c(nrow(data_incident),ncol(data_incident),ncol(prob_matrix)))
```

```
for(i in 1:nrow(data_incident)){
  for(j in 1:ncol(data_incident)){
      counts_array[i,j,1]<- data_incident[i,j]*prob_matrix[j,1]
      counts_array[i,j,2]<- counts_array[i,j,1]*prob_matrix[j,2]
      counts_array[i,j,3]<- (counts_array[i,j,1]-counts_array[i,j,2])*prob_matrix[j,3]
      counts_array[i,j,4]<- counts_array[i,j,3]*prob_matrix[j,4]
      counts_array[i,j,5]<- counts_array[i,j,4]*prob_matrix[j,5]

  }
}
```

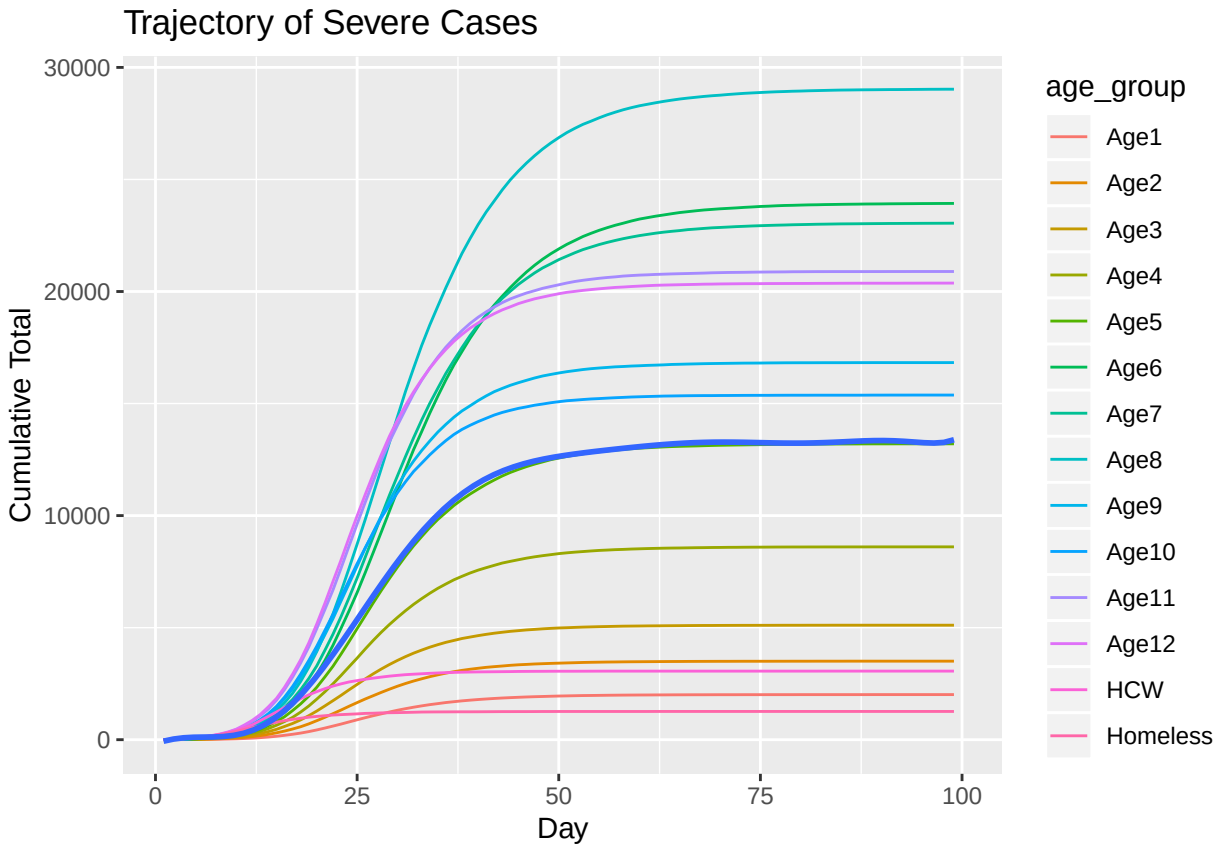
Now load in a dataset that gives estimates of how long people stay hospitalized, in ICU or on vents Compute a running total of each category.

```
#for now we are supposing we get a point estimate to determine how long people stay in each of the stat

length_of_stay<- c(NA,NA,8,9,10)

cumulative_array<- array(NA, dim=c(nrow(data_incident),ncol(data_incident),ncol(prob_matrix)))


  for(j in 1:ncol(data_incident)){
    cumulative_array[1,j,]<- counts_array[1,j,]

    for(i in 2:nrow(data_incident)){
      cumulative_array[i,j,1]<-cumulative_array[i-1,j,1]+counts_array[i,j,1]
      cumulative_array[i,j,2]<-cumulative_array[i-1,j,2]+counts_array[i,j,2]
    }

    for(k in 2:nrow(data_incident)){
      if(k<=length_of_stay[3]){
      cumulative_array[k,j,3]<- cumulative_array[k-1,j,3]+counts_array[k,j,3]
      }
      else {
        cumulative_array[k,j,3]<-cumulative_array[k-1,j,3]+
          counts_array[k,j,3]-counts_array[k-length_of_stay[3],j,3]
      }
    }
  for(l in 2:nrow(data_incident)){
      if(l <= length_of_stay[4]){
        cumulative_array[l,j,4]<- cumulative_array[l-1,j,4]+counts_array[l,j,4]
      }
    else{
      cumulative_array[l,j,4]<-cumulative_array[l-1,j,4]+
          counts_array[l,j,4]-counts_array[l-length_of_stay[4],j,4]
    }
  }
    for(m in 2:nrow(data_incident)){
    if(m<=length_of_stay[5]){
      cumulative_array[m,j,5]<- cumulative_array[m-1,j,5]+counts_array[m,j,5]
    }
    else{
      cumulative_array[m,j,5]<-cumulative_array[m-1,j,5]+
          counts_array[m,j,5]-counts_array[m-length_of_stay[5],j,5]
```

```
    }
  }
}
```

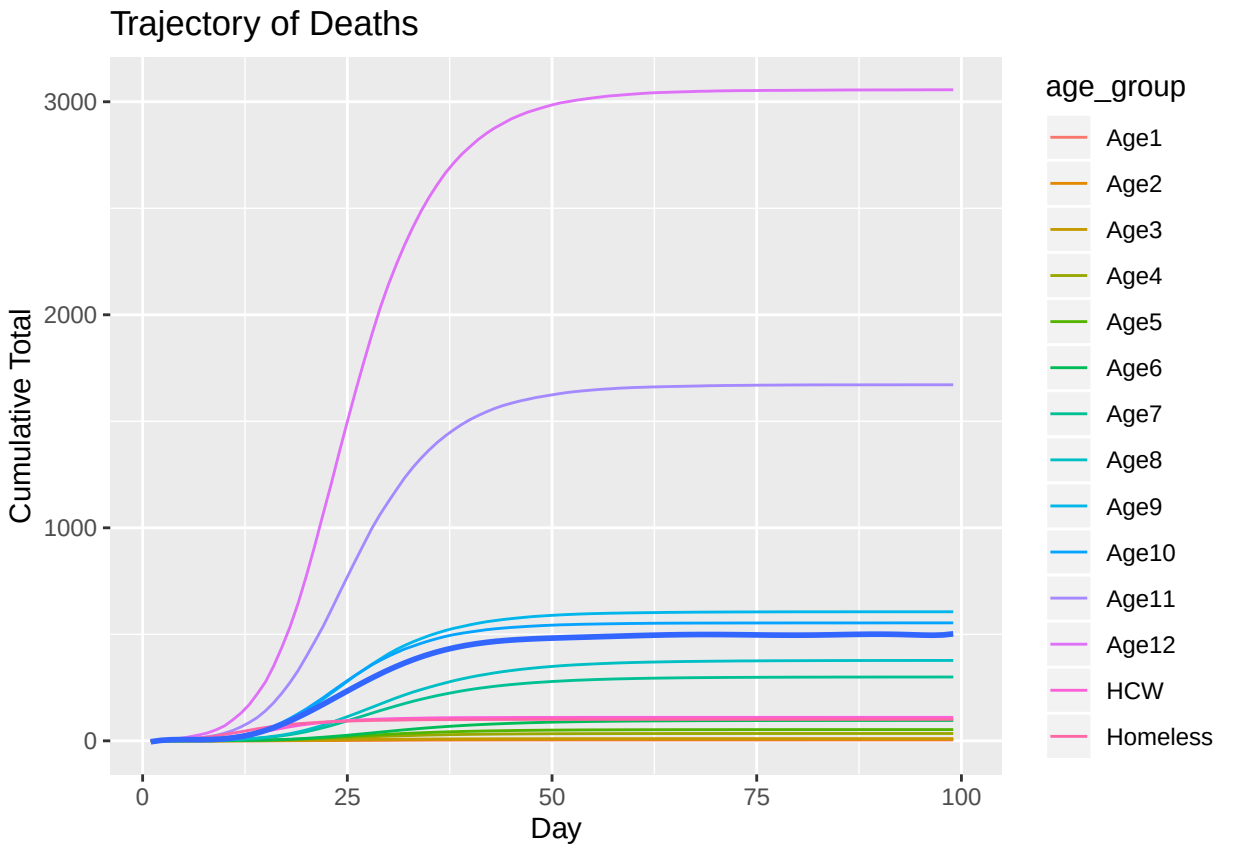
Plot cumulative values over 99 days.

```
#return to long format for each state
#severe

severe_wide<-as.data.frame(cbind(rep(1:100),cumulative_array[,,1]))
colnames(severe_wide)<- c("Day","Age1","Age2","Age3","Age4","Age5","Age6","Age7",
                          "Age8","Age9","Age10","Age11","Age12","HCW","Homeless")

severe_long<- gather(severe_wide,age_group, cumulative_total, Age1:Homeless,factor_key=TRUE)



plt_severe<- ggplot(severe_long)+
  geom_line(aes(y= cumulative_total, x=Day,color=age_group))+
  stat_smooth(aes(y= cumulative_total,x=Day), method=lm, formula=y~poly(x,10), se=FALSE)+
  labs(title="Trajectory of Severe Cases",y="Cumulative Total")

plt_severe
```



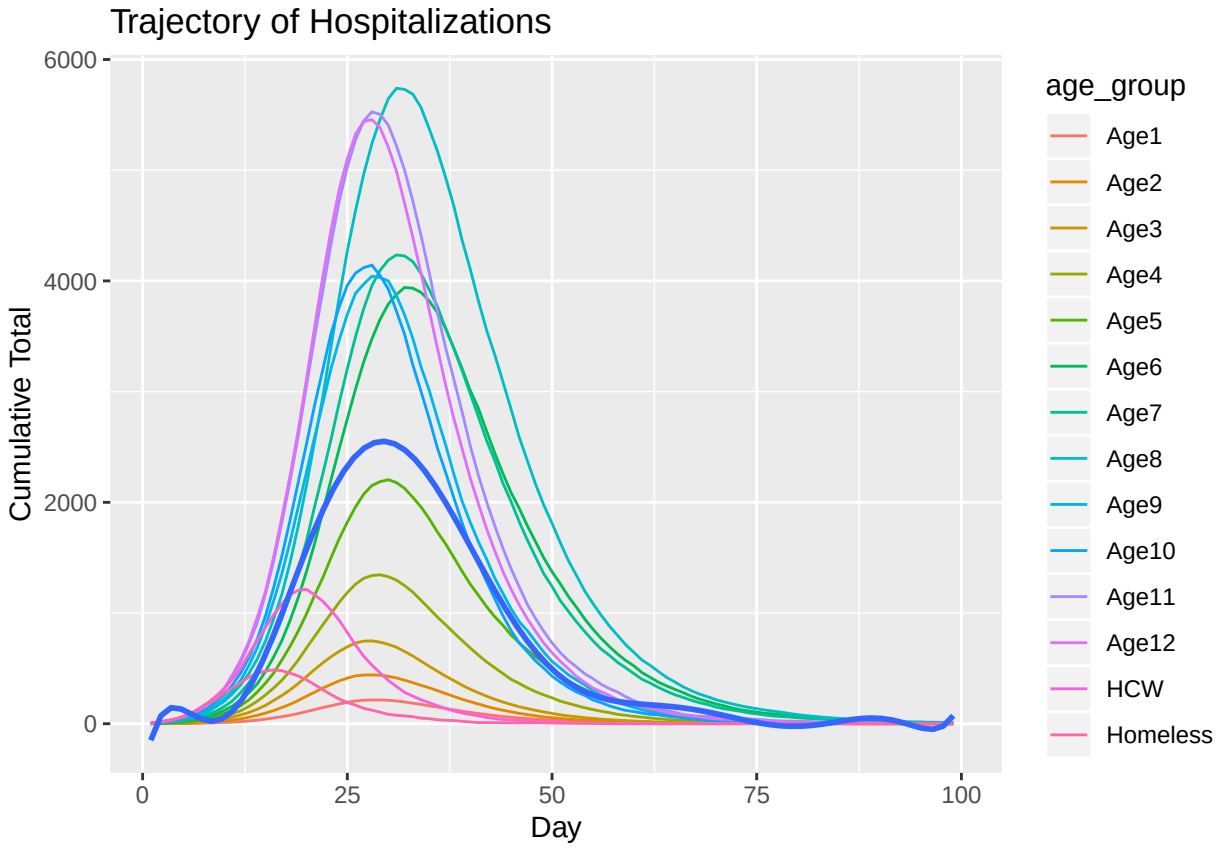

```
###############################
deaths_wide<-as.data.frame(cbind(rep(1:100),cumulative_array[,,2]))
colnames(deaths_wide)<- c("Day","Age1","Age2","Age3","Age4","Age5","Age6","Age7",
                          "Age8","Age9","Age10","Age11","Age12","HCW","Homeless")
```

```
deaths_long<- gather(deaths_wide,age_group, cumulative_total, Age1:Homeless,factor_key=TRUE)



plt_deaths<- ggplot(deaths_long)+
  geom_line(aes(y= cumulative_total, x=Day,color=age_group))+
  stat_smooth(aes(y= cumulative_total,x=Day), method=lm, formula=y~poly(x,10), se=FALSE)+
  labs(title="Trajectory of Deaths",y="Cumulative Total")

plt_deaths
```

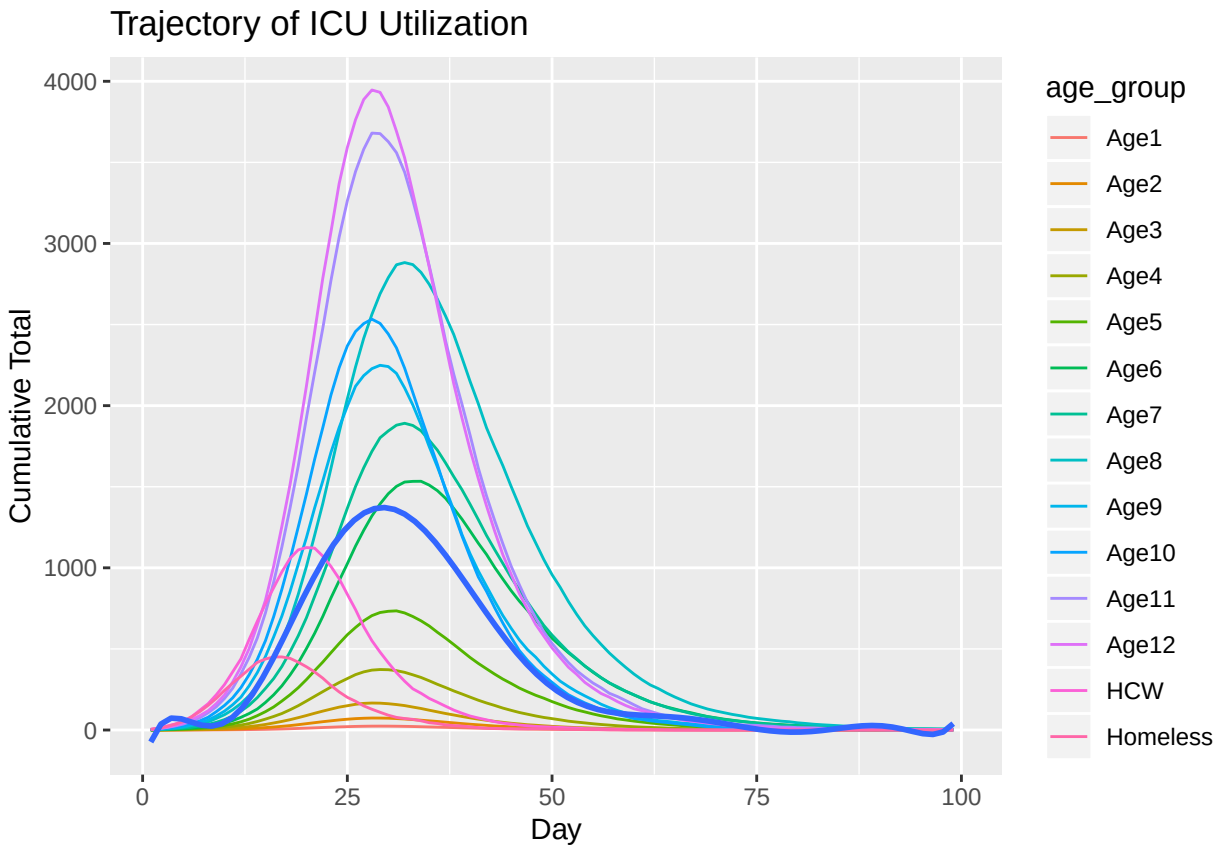



```
#################################


hosp_wide<-as.data.frame(cbind(rep(1:100),cumulative_array[,,3]))
colnames(hosp_wide)<- c("Day","Age1","Age2","Age3","Age4","Age5","Age6","Age7",
                        "Age8","Age9","Age10","Age11","Age12","HCW","Homeless")

hosp_long<- gather(hosp_wide,age_group, cumulative_total, Age1:Homeless,factor_key=TRUE)



plt_hosp<- ggplot(hosp_long)+
  geom_line(aes(y= cumulative_total, x=Day,color=age_group))+
  stat_smooth(aes(y= cumulative_total,x=Day), method=lm, formula=y~poly(x,10), se=FALSE)+
```

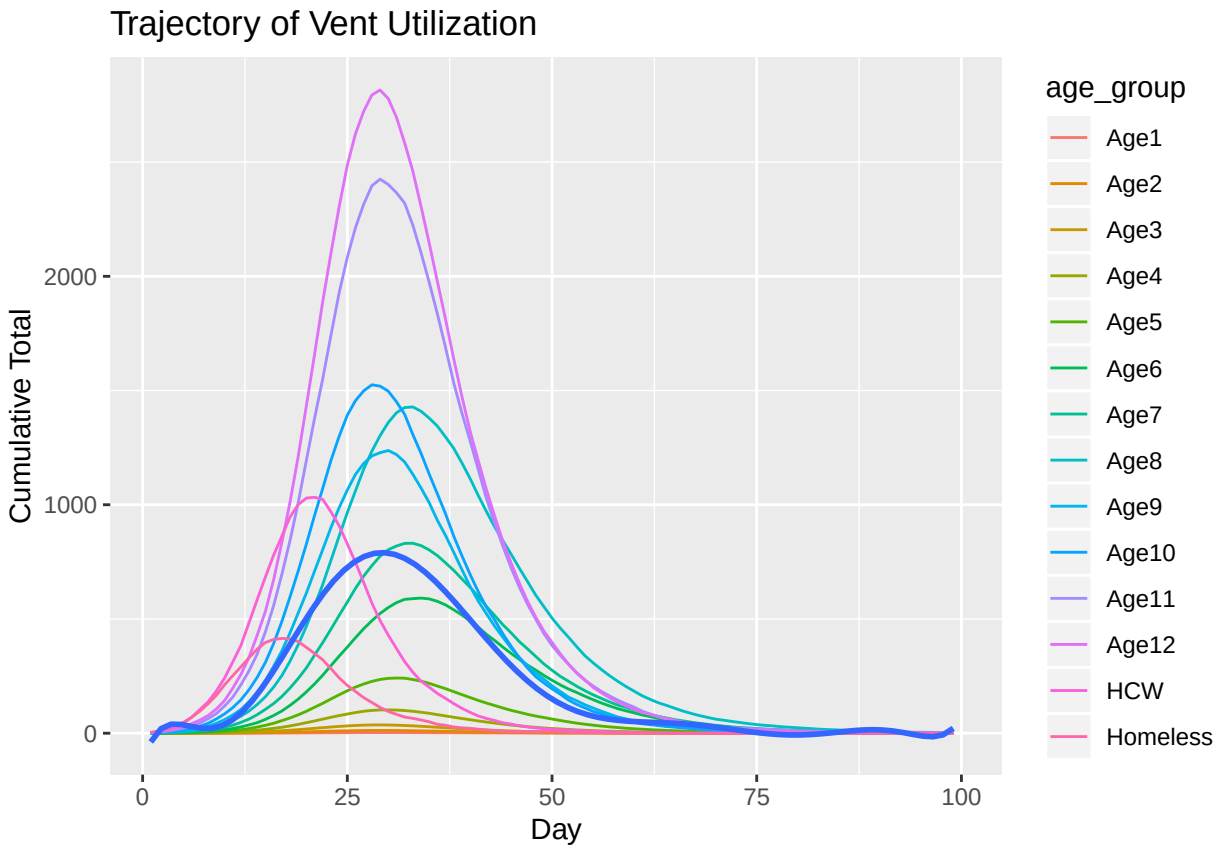
```
  labs(title="Trajectory of Hospitalizations",y="Cumulative Total")

plt_hosp
```




```
####################################



ICU_wide<-as.data.frame(cbind(rep(1:100),cumulative_array[,,4]))
colnames(ICU_wide)<- c("Day","Age1","Age2","Age3","Age4","Age5","Age6","Age7",
                       "Age8","Age9","Age10","Age11","Age12","HCW","Homeless")

ICU_long<- gather(ICU_wide,age_group, cumulative_total, Age1:Homeless,factor_key=TRUE)



plt_ICU<- ggplot(ICU_long)+
  geom_line(aes(y= cumulative_total, x=Day,color=age_group))+
  stat_smooth(aes(y= cumulative_total,x=Day), method=lm, formula=y~poly(x,10), se=FALSE)+
  labs(title="Trajectory of ICU Utilization",y="Cumulative Total")

plt_ICU
```




```
################################################


vent_wide<-as.data.frame(cbind(rep(1:100),cumulative_array[,,5]))
colnames(vent_wide)<- c("Day","Age1","Age2","Age3","Age4","Age5","Age6","Age7",
                        "Age8","Age9","Age10","Age11","Age12","HCW","Homeless")

vent_long<- gather(vent_wide,age_group, cumulative_total, Age1:Homeless,factor_key=TRUE)



plt_vent<- ggplot(vent_long)+
  geom_line(aes(y= cumulative_total, x=Day,color=age_group))+
  stat_smooth(aes(y= cumulative_total,x=Day), method=lm, formula=y~poly(x,10), se=FALSE)+
  labs(title="Trajectory of Vent Utilization",y="Cumulative Total")

plt_vent
```


Trajectory of Vent Utilization
Cumulative Total
2000
1000
0
0
25
50
75
100
Day
age_group
Age1
Age2
Age3
Age4
Age5
Age6
Age7
Age8
Age9
Age10
Age11
Age12
HCW
Homeless