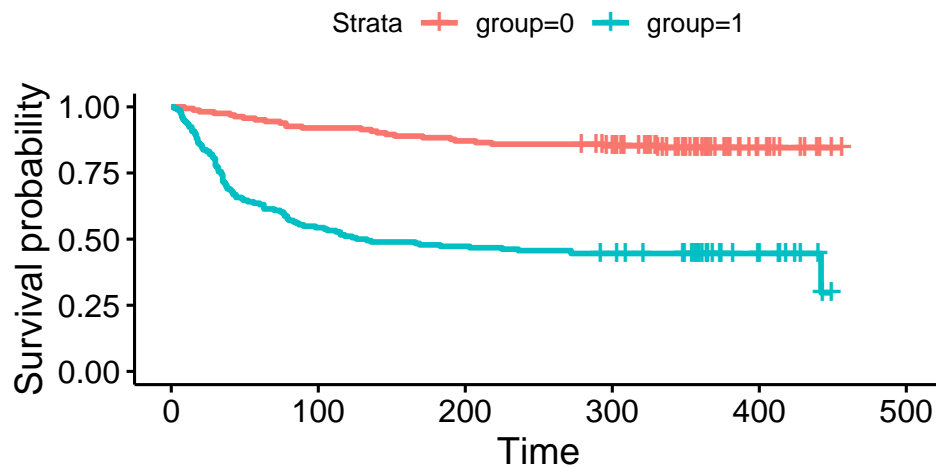# Data_viz

Arm and Guy

April 19, 2021

```
urlfile="https://raw.githubusercontent.com/gthampak/Arm_Guy_MATH150_Project/main/HELPdata.csv"

HELPdata <- read_csv(url(urlfile))

HELPdata_survfit <- survfit(Surv(dayslink, linkstatus) ~ group, data=HELPdata)

ggsurvplot(HELPdata_survfit, conf.type = "TRUE") +
  ggtitle("KM-curve")
```



```
data(HELPfull)

table(HELPdata$dayslink)
```
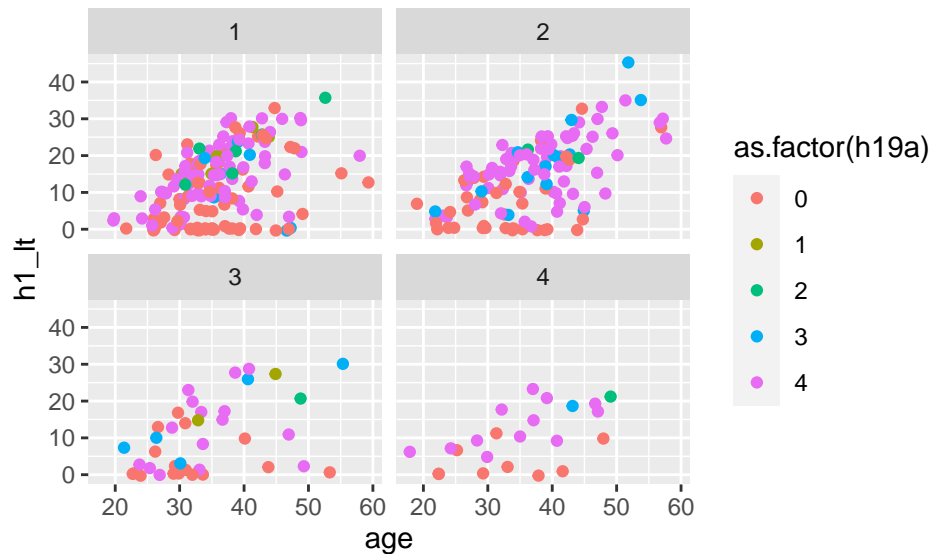
```
##
##    2    4    6    7    8    9   10   11   12   14   15   16   17   18   19   20   21   22   24   26
##    1    1    2    3    2    2    1    1    2    3    1    2    3    3    2    1    2    1    1    2
##   28   29   30   31   32   34   35   36   37   38   40   41   42   43   44   48   49   52   56   57
##    2    2    5    1    3    2    5    1    2    2    2    1    2    1    2    1    2    1    1    1
##   60   63   64   70   74   76   77   78   79   80   83   85   87   89   90   98  104  106  112  115
##    1    3    1    1    2    1    2    2    2    1    1    1    1    1    1    1    1    1    1    2
##  118  123  126  129  133  136  140  146  152  166  169  171  183  190  193  203  207  218  225  236
##    1    1    1    1    1    2    1    1    1    1    1    1    1    1    1    1    1    1    1    1
##  272  279  289  292  293  294  296  300  302  303  306  307  308  309  318  321  322  324  326  329
##    2    1    1    1    2    1    1    1    1    2    1    1    1    1    1    1    1    1    1    1
##  330  331  334  337  343  345  348  349  350  353  354  356  357  358  359  361  363  364  365  367
```

```
##    1   2   1   1   1   2   2   2   1   1   1   2   1   2   2   3   1   2 132   1
## 368 370 373 374 376 377 379 380 382 386 387 393 398 399 400 405 407 410 413 414
##    2   2   1   1   1   1   1   1   1   1   1   2   1   2   1   1   1   1   1   3
## 418 424 428 431 439 440 441 442 443 449 456
##    1   1   2   1   1   1   1   1   1   3   1
```

```r
ggplot(HELPdata) +
  geom_jitter(aes(x = age, y = h1_lt, color = as.factor(h19a))) +
  facet_wrap(~race)
```
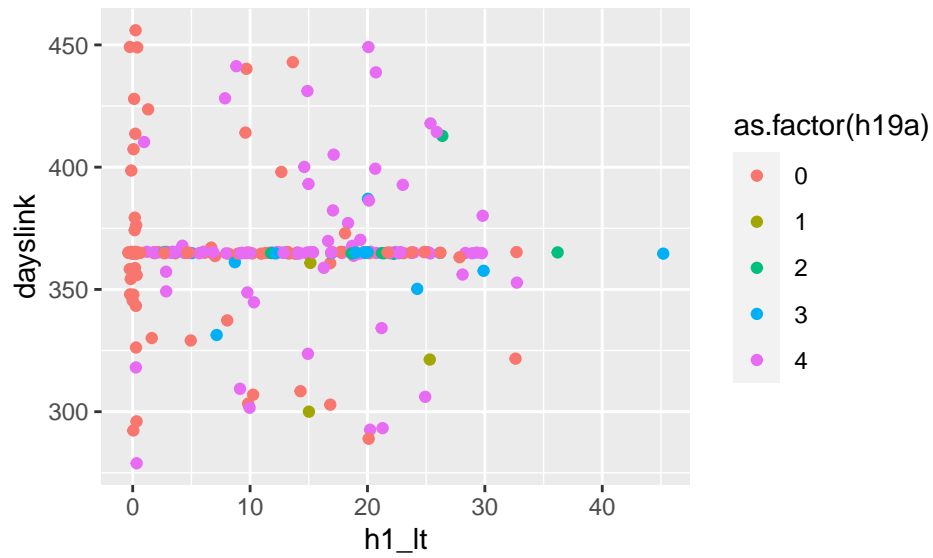


```r
HELPdata2 <- HELPdata %>%
  filter(linkstatus == 0) %>%
  mutate(dayslink_blck = case_when(
    dayslink <= 0 ~ "0",
    dayslink <= 25 ~ "0-25",
    dayslink <= 50 ~ "25-50",
    dayslink <= 100 ~ "50-100",
    dayslink <= 150 ~ "100-150",
    dayslink <= 200 ~ "150-200",
    dayslink <= 250 ~ "200-250",
    dayslink <= 300 ~ "250-300",
    dayslink <= 350 ~ "300-350",
    dayslink <= 400 ~ "350-400",
    dayslink <= 450 ~ "400-450",
    dayslink <= 500 ~ "450-500"
  ))
```

```r
table(HELPdata2$dayslink_blck)
```
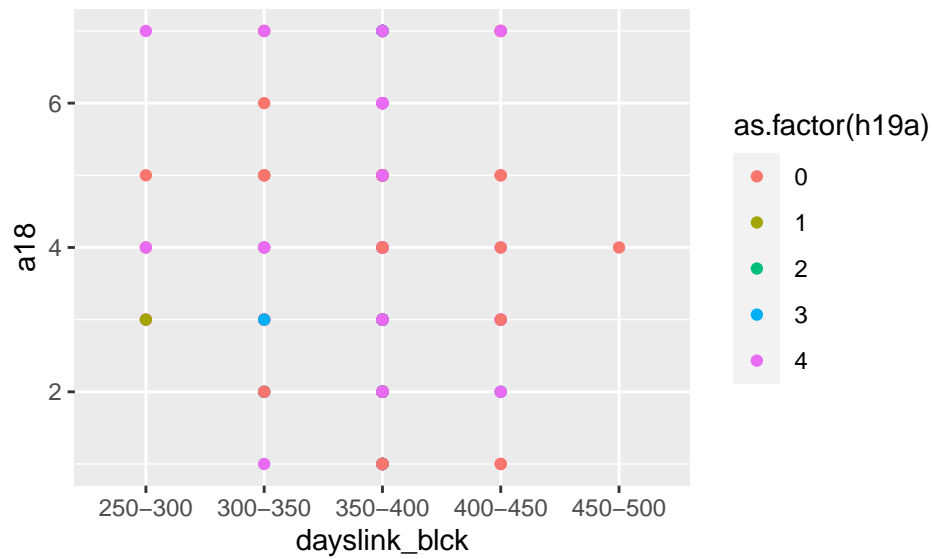
```
##
## 250-300 300-350 350-400 400-450 450-500
##       7      25     167      19       1
```
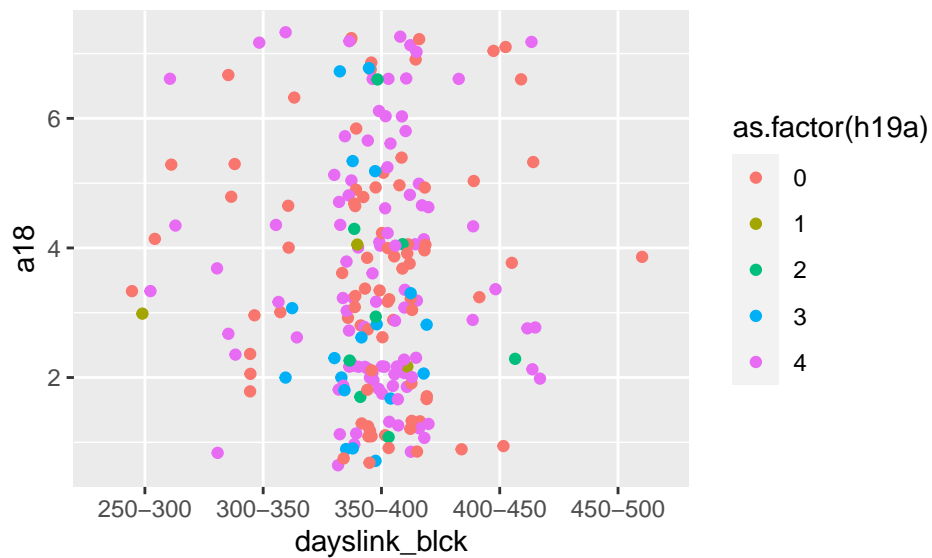
```r
?HELPrct
```

```r
ggplot(HELPdata2) +
  geom_jitter(aes(x = h1_lt, y = dayslink, color = as.factor(h19a)))
```
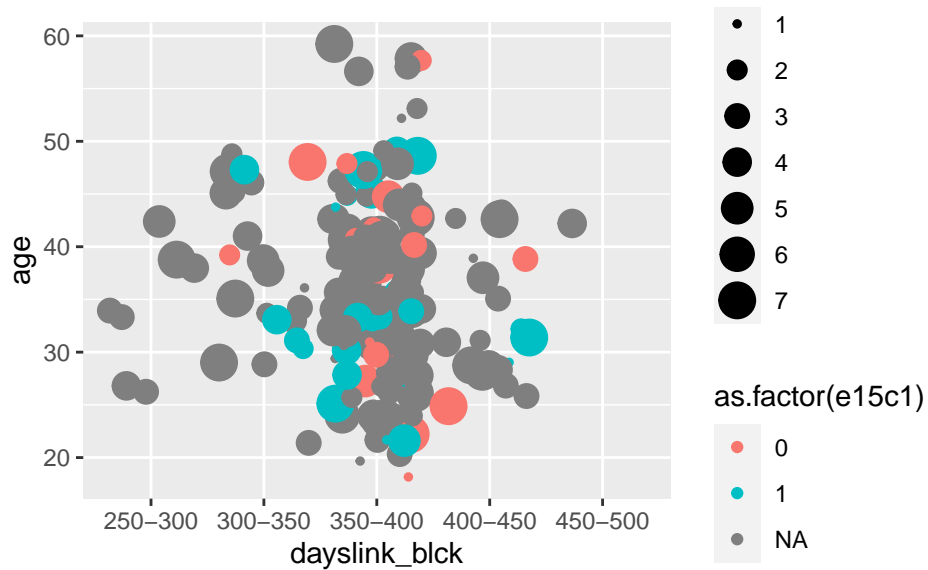
```
ggplot(HELPdata2) +
  geom_point(aes(x = dayslink_blck, y = a18, group = race, color = as.factor(h19a)))
```
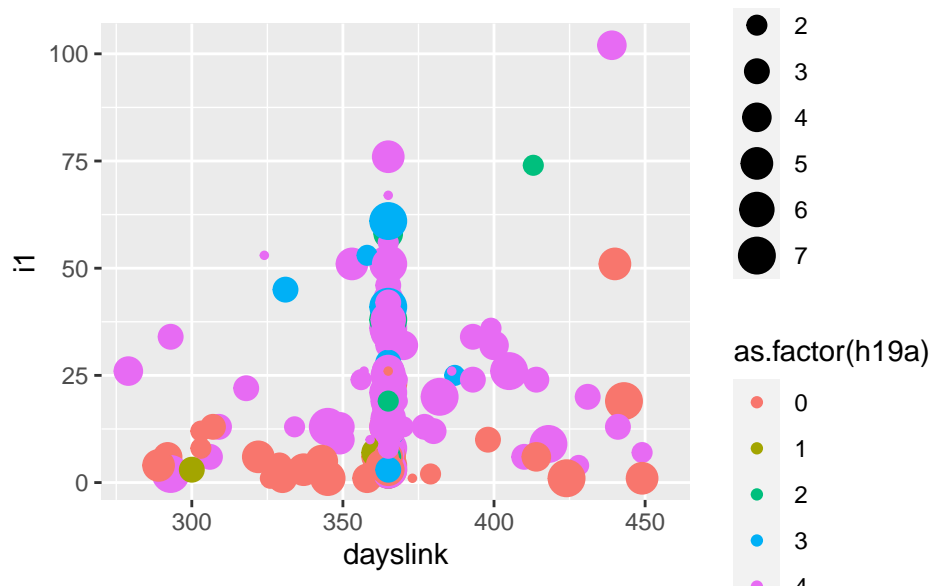


```
ggplot(HELPdata2) +
  geom_jitter(aes(x = dayslink_blck, y = a18, group = race, color = as.factor(h19a)))
```

```
ggplot(HELPdata2) +
  geom_jitter(aes(x = dayslink_blck, y = age, size = a18, color = as.factor(e15c1)))
```



```
ggplot(HELPdata2) +
  geom_point(aes(x = dayslink, y = i1, size = a18, color = as.factor(h19a)))
```

```
ggplot(HELPdata2) +
  geom_jitter(aes(x = dayslink_blck, y = i1, size = a18, color = as.factor(h19a)))
```