DAR F21 Project Status Notebook Assignment 4

DeFi

Jaimin Vyas

10-01-2021

Contents


```
#import libraries
if (!require("ggplot2")) {
  install.packages("ggplot2")
  library(ggplot2)
}
```

```
## Loading required package: ggplot2
```

```
if (!require("knitr")) {
  install.packages("knitr")
  library(knitr)
}
```

```
## Loading required package: knitr
```

```
if (!require("RColorBrewer")) {
  install.packages("RColorBrewer")
  library(RColorBrewer)
}
```

```
## Loading required package: RColorBrewer
```

```
if (!require("tidyverse")) {
  install.packages("tidyverse")
  library(tidyverse)
}
```

```
## Loading required package: tidyverse
```

```
## Warning in system("timedatectl", intern = TRUE): running command 'timedatectl'
## had status 1
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v tibble  3.1.5     v dplyr   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.0.2     v forcats 0.5.1
## v purrr   0.3.4
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
if (!require("xts")) {
  install.packages("xts")
  library(xts)
}
```

```
## Loading required package: xts
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
##
## Attaching package: 'xts'
```

```
## The following objects are masked from 'package:dplyr':
##
##     first, last
```

```
if (!require("dplyr")) {
  install.packages("dplyr")
  library(dplyr)
}
if (!require("devtools")) {
  install.packages("devtools")
  library(devtools)
}
```

```
## Loading required package: devtools
```

```
## Loading required package: usethis
```

```
if (!require("ggfortify")) {
  install.packages("ggfortify")
  library(ggfortify)
}
```

```
## Loading required package: ggfortify
```

```
if (!require("lubridate")) {
  install.packages("lubridate")
  library(lubridate)
}
```

```
## Loading required package: lubridate
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```
if (!require("survival")) {
  install.packages("survival")
```

```
  library(survival)
}
```

```
## Loading required package: survival
```

```
if (!require("zoo")) {
  install.packages("zoo")
  library(zoo)
}
```

```
borrows <- read_csv('../../Data/borrows.csv')
```

```
## Rows: 94977 Columns: 12
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (6): borrowRateMode, onBehalfOf, pool, reserve, user, type
## dbl (6): amount, borrowRate, timestamp, reservePriceETH, reservePriceUSD, am...
```

```
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
deposits <- read_csv('../../Data/deposits.csv')
```

```
## Rows: 370559 Columns: 10
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (6): amount, onBehalfOf, pool, reserve, user, type
## dbl (4): timestamp, reservePriceETH, reservePriceUSD, amountUSD
```

```
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
raw_df <- read_rds('../../Data/transactions.Rds')
```

```
borrows <- raw_df %>%
  filter(type=="borrow")

deposits <- raw_df %>%
  filter(type=="deposit")

depositBorrow <- left_join(deposits,borrows,by="onBehalfOf") %>%
  dplyr::rename(depositTime=timestamp.x) %>%
  dplyr::rename(borrowTime=timestamp.y) %>%
  group_by(onBehalfOf) %>%
  dplyr::summarise(timeDiff=case_when(min(borrowTime)-min(depositTime)>0 ~ (min(borrowTime)-min(deposit
  mutate(status=case_when(timeDiff==max(borrows$timestamp)/86400 ~ 0, timeDiff>0 ~ 1, TRUE ~ 0)) %>%
  select(onBehalfOf,timeDiff,status)

depositBorrow
```

```
## # A tibble: 44,327 x 3
##    onBehalfOf    timeDiff status
##         <dbl>       <dbl>  <dbl>
##  1    2.58e33 18858.           0
##  2    6.66e34     0.00260      1
```

```
##  3    4.87e35 18858.           0
##  4    8.01e35 18858.           0
##  5    1.36e37     0.00113      1
##  6    8.09e37 18858.           0
##  7    1.54e39     0.00295      1
##  8    3.73e40     0.0169       1
##  9    5.82e40 18858.           0
## 10    6.66e40 18858.           0
## # ... with 44,317 more rows
```

```
unique(depositBorrow$status)
```

```
## [1] 0 1
```

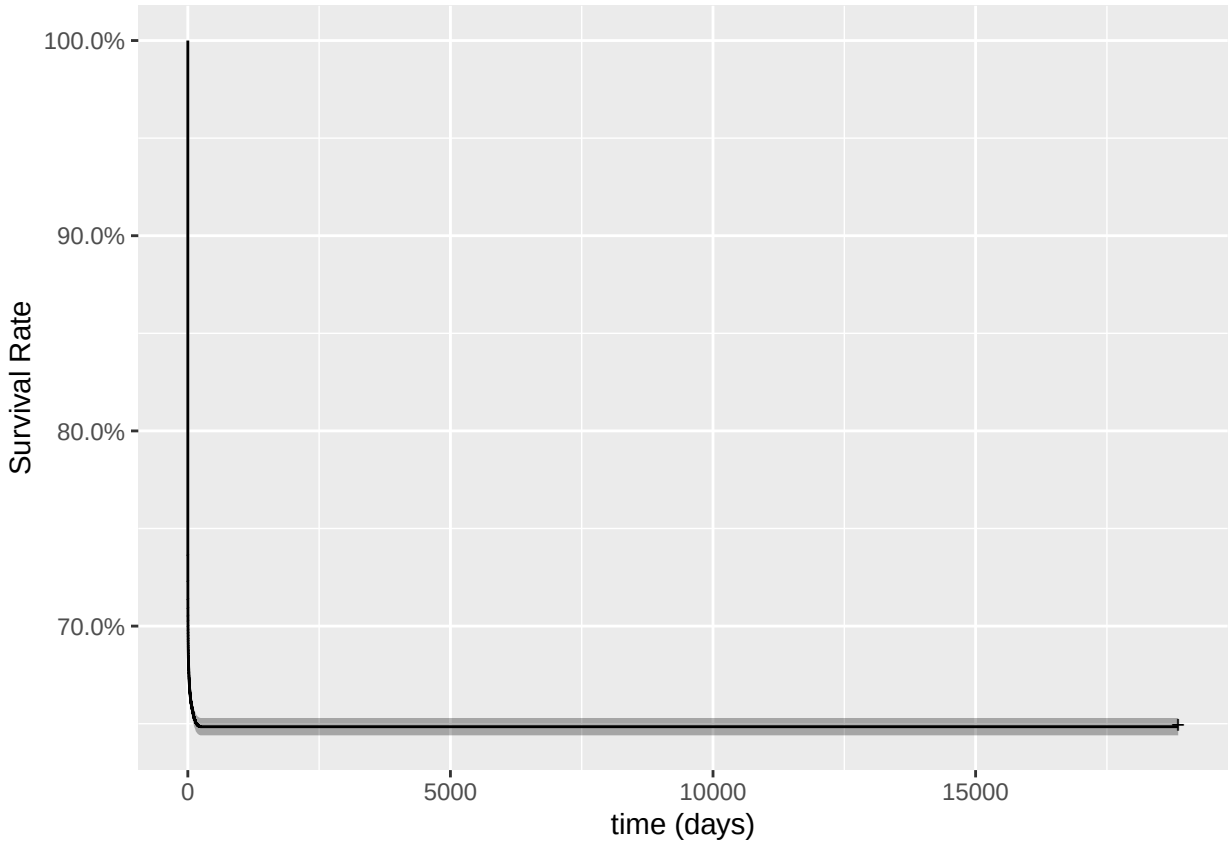
```
km <- with(depositBorrow, Surv(timeDiff, status))
head(km,80)
```

```
##  [1] 1.885787e+04+ 2.604167e-03  1.885787e+04+ 1.885787e+04+ 1.134259e-03
##  [6] 1.885787e+04+ 2.951389e-03  1.689815e-02  1.885787e+04+ 1.885787e+04+
## [11] 2.743056e-03  3.503472e-02  1.885787e+04+ 1.885787e+04+ 1.885787e+04+
## [16] 1.885787e+04+ 1.885787e+04+ 8.360292e+01  1.885787e+04+ 1.885787e+04+
## [21] 1.885787e+04+ 4.386574e-03  1.412037e-03  7.664352e-01  3.787037e-02
## [26] 1.885787e+04+ 9.375000e-04  1.885787e+04+ 6.759259e-03  1.885787e+04+
## [31] 1.885787e+04+ 1.648495e+00  4.629630e-04  1.885787e+04+ 1.885787e+04+
## [36] 1.885787e+04+ 1.885787e+04+ 1.885787e+04+ 1.885787e+04+ 1.885787e+04+
## [41] 5.081019e-03  1.885787e+04+ 1.885787e+04+ 7.222222e-03  1.885787e+04+
## [46] 1.885787e+04+ 1.885787e+04+ 2.234954e-02  1.885787e+04+ 1.885787e+04+
## [51] 1.885787e+04+ 1.885787e+04+ 1.885787e+04+ 6.595439e+01  1.885787e+04+
## [56] 1.885787e+04+ 7.627315e-03  2.662037e-03  1.885787e+04+ 8.956829e-01
## [61] 1.885787e+04+ 1.885787e+04+ 1.885787e+04+ 6.025463e-02  1.885787e+04+
## [66] 1.885787e+04+ 1.885787e+04+ 2.410880e-02  1.885787e+04+ 1.885787e+04+
## [71] 1.885787e+04+ 1.885787e+04+ 1.885787e+04+ 1.956019e-03  1.885787e+04+
## [76] 1.885787e+04+ 1.885787e+04+ 1.844907e-02  1.885787e+04+ 1.885787e+04+
```

```
km_fit <- survfit(Surv(timeDiff, status) ~ 1, data=depositBorrow)
summary(km_fit, times = c(1,30,60,90*(1:10)))
```

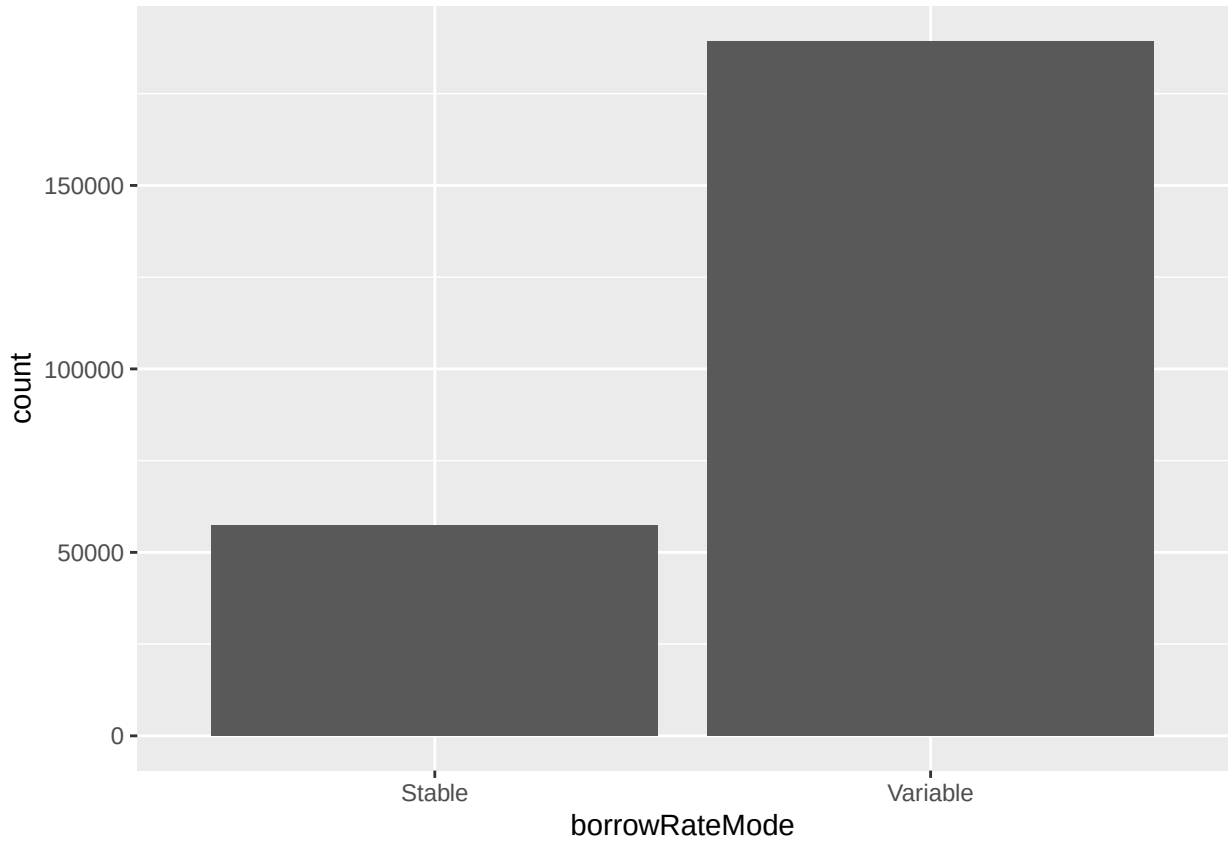
```
## Call: survfit(formula = Surv(timeDiff, status) ~ 1, data = depositBorrow)
##
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     1  31895   12432    0.720 0.00213        0.715        0.724
##    30  29694    2201    0.670 0.00223        0.666        0.674
##    60  29300     394    0.661 0.00225        0.657        0.665
##    90  29108     192    0.657 0.00226        0.652        0.661
##   180  28790     318    0.649 0.00227        0.645        0.654
##   270  28744      46    0.648 0.00227        0.644        0.653
##   360  28744       0    0.648 0.00227        0.644        0.653
##   450  28744       0    0.648 0.00227        0.644        0.653
##   540  28744       0    0.648 0.00227        0.644        0.653
##   630  28744       0    0.648 0.00227        0.644        0.653
##   720  28744       0    0.648 0.00227        0.644        0.653
##   810  28744       0    0.648 0.00227        0.644        0.653
##   900  28744       0    0.648 0.00227        0.644        0.653
```

```
autoplot(km_fit,xlab="time (days)",ylab="Survival Rate",title="Survival Rate of Deposit before the firs
```

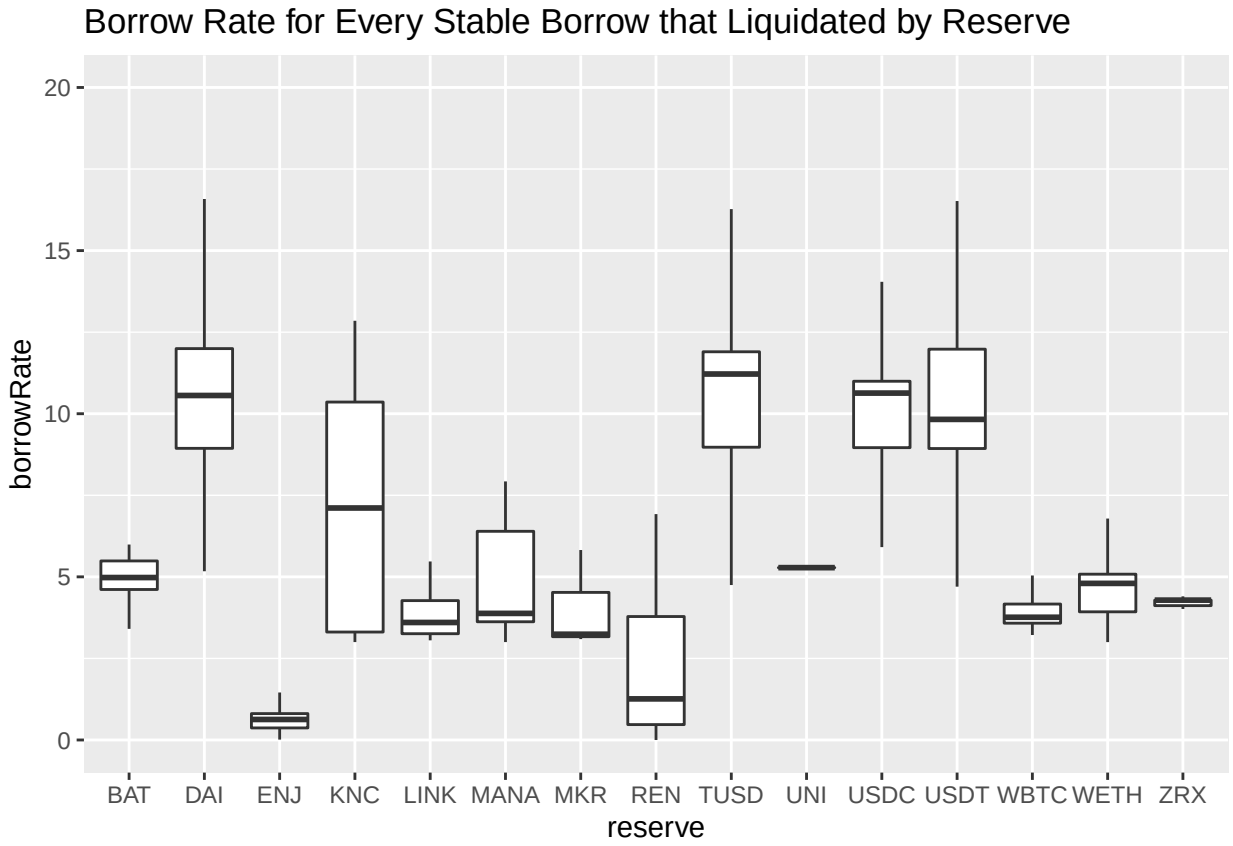



```
liquidations <- raw_df %>%
  filter(type=="liquidation")
```

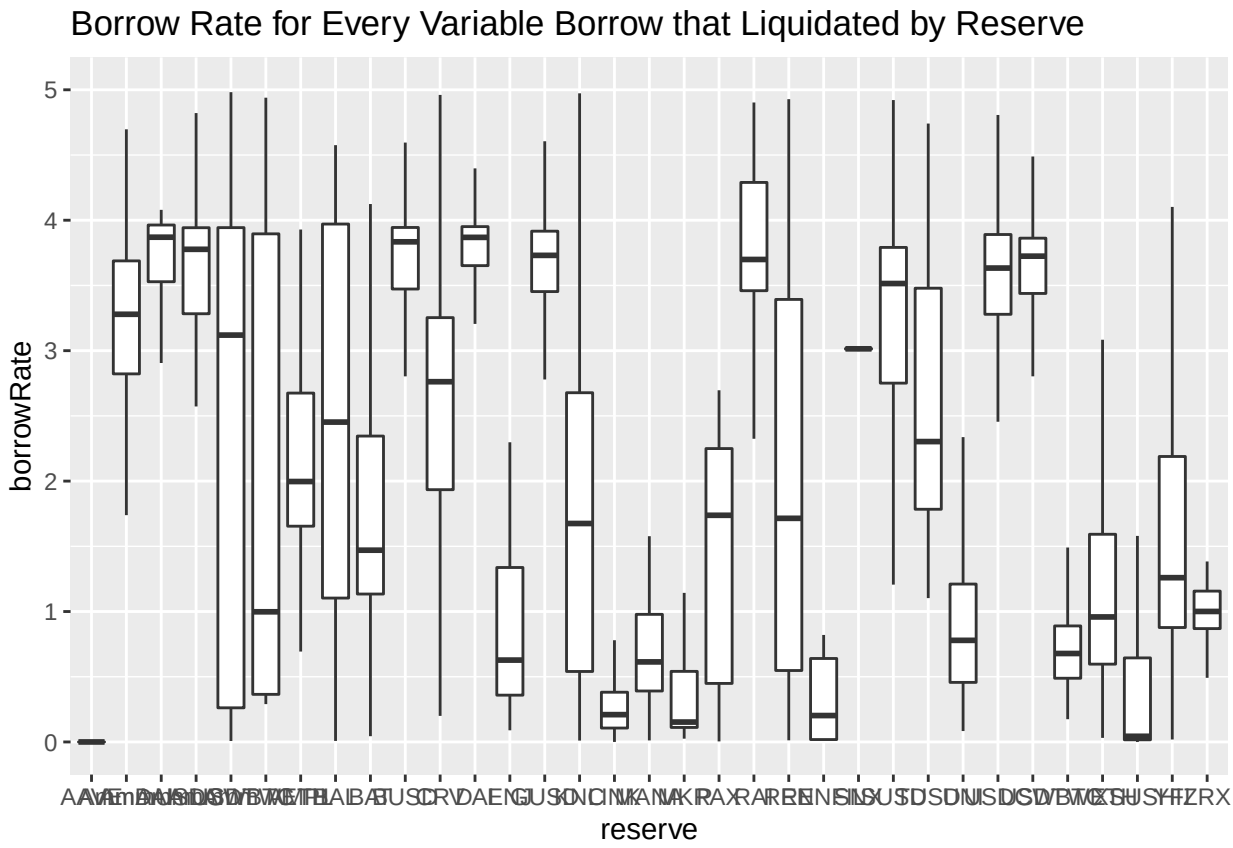
```
borrows2 <- raw_df %>%
  filter(type=="borrow") %>%
  select(onBehalfOf,borrowRateMode,borrowRate,reserve,amountUSD,timestamp) %>%
  dplyr::rename(user=onBehalfOf)

liquids <- raw_df %>%
  filter(type=="liquidation") %>%
  select(user)

liquidModes <- left_join(borrows2,liquids,by="user")

ggplot(liquidModes,aes(x=borrowRateMode)) + geom_bar()
```



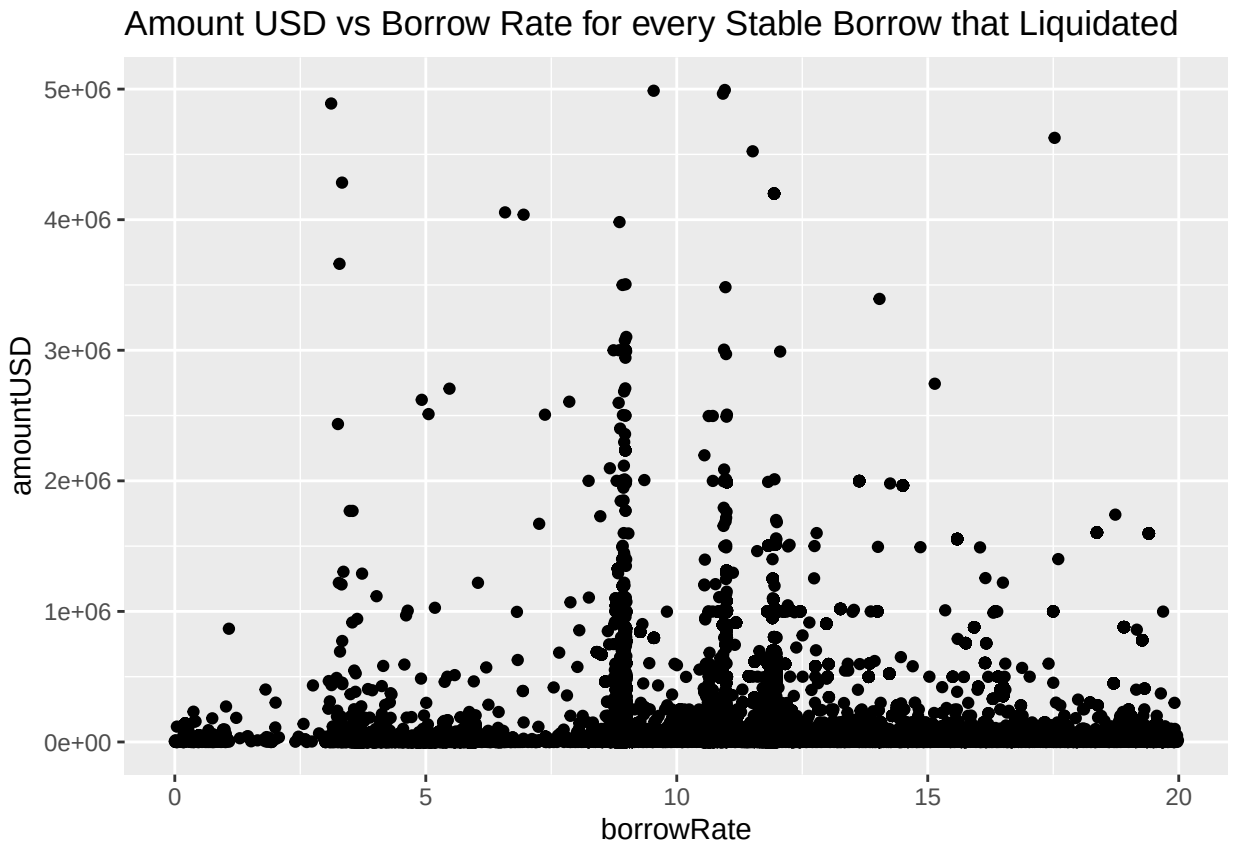

```
liquidStable <- liquidModes %>%
  filter(borrowRateMode=="Stable" & amountUSD <= 5000000 & borrowRate <= 20) %>%
  select(borrowRate,reserve,amountUSD,timestamp)

liquidVariable <- liquidModes %>%
  filter(borrowRateMode=="Variable" & amountUSD <= 25000000 & borrowRate <= 5) %>%
  select(borrowRate,reserve,amountUSD,timestamp)

ggplot(liquidStable,aes(x=reserve,y=borrowRate)) + geom_boxplot(outlier.shape=NA) + ggtitle("Borrow Rat
```



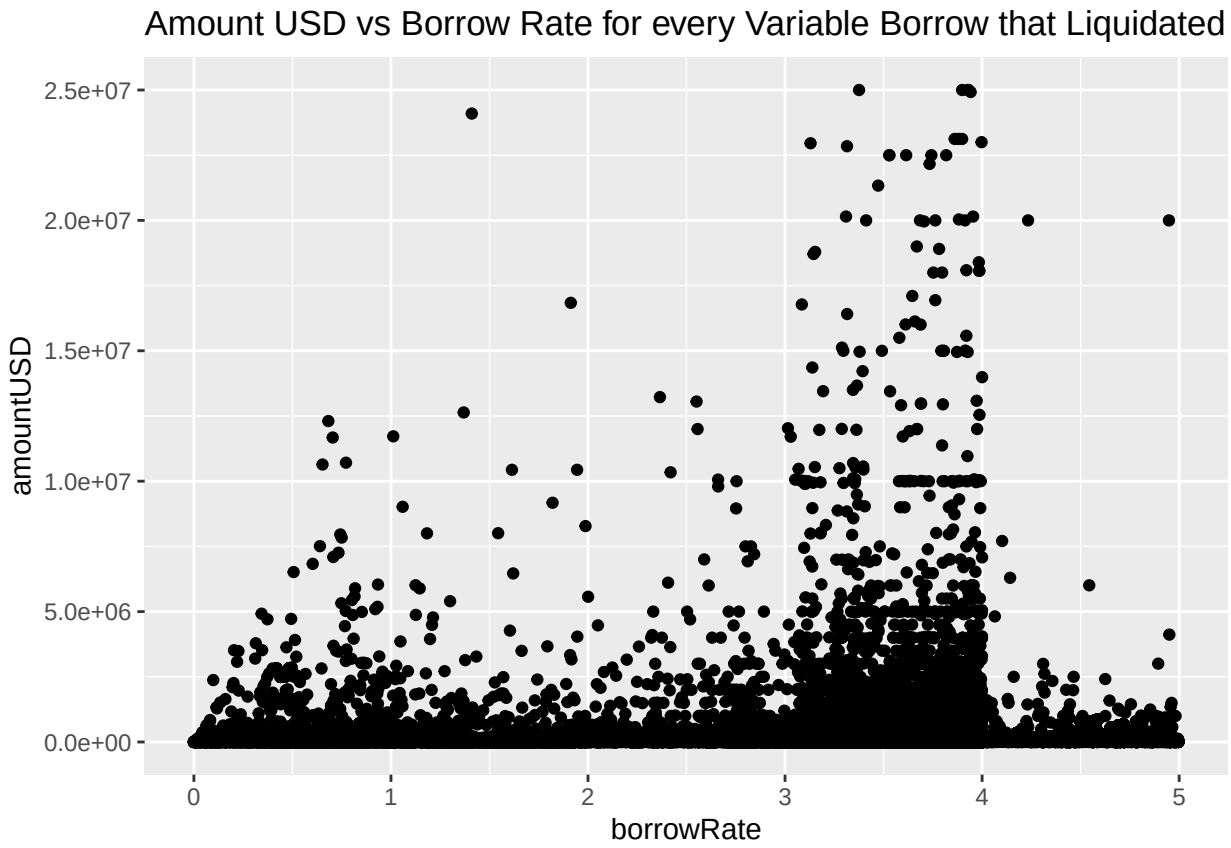

```
ggplot(liquidVariable,aes(x=reserve,y=borrowRate)) + geom_boxplot(outlier.shape=NA) + ggtitle("Borrow Ra
```



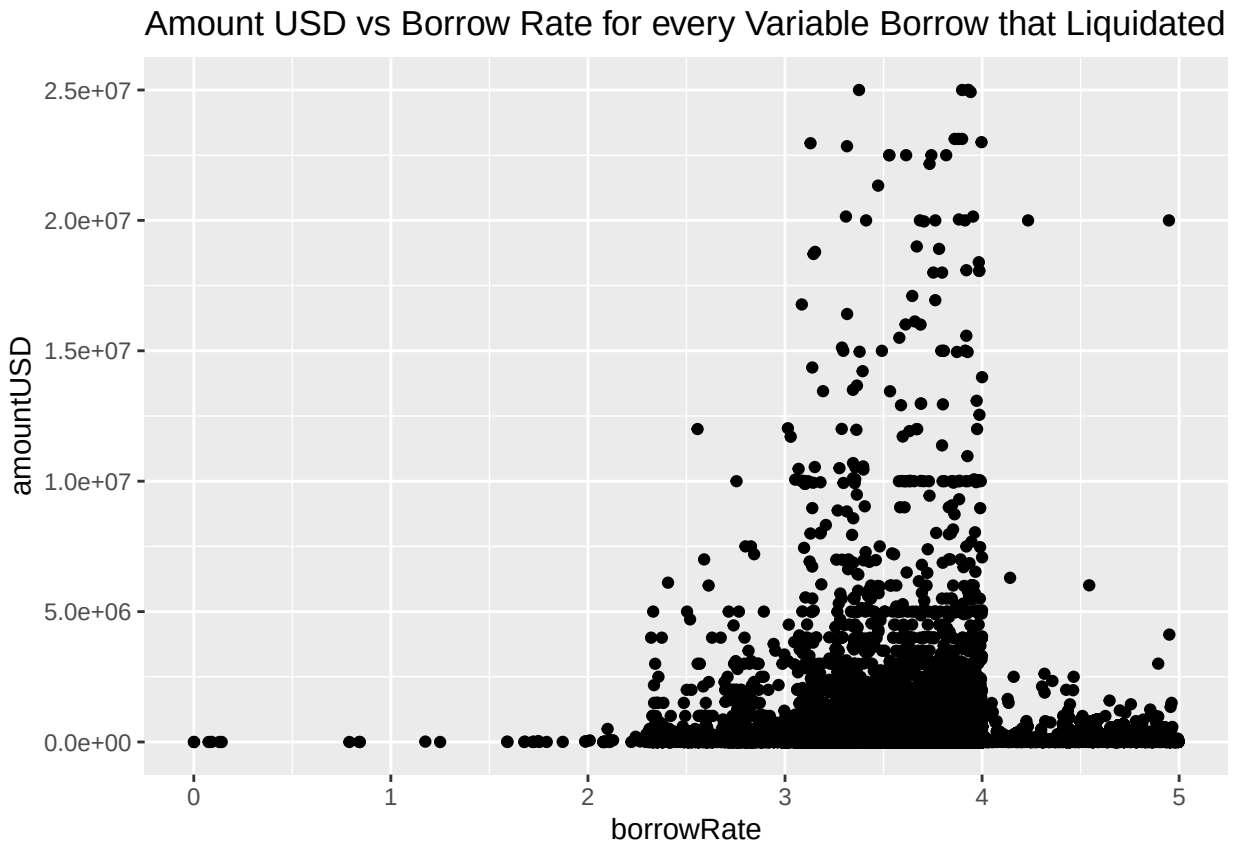

```
ggplot(liquidStable,aes(x=borrowRate,y=amountUSD)) + geom_point() + ggtitle("Amount USD vs Borrow Rate
```



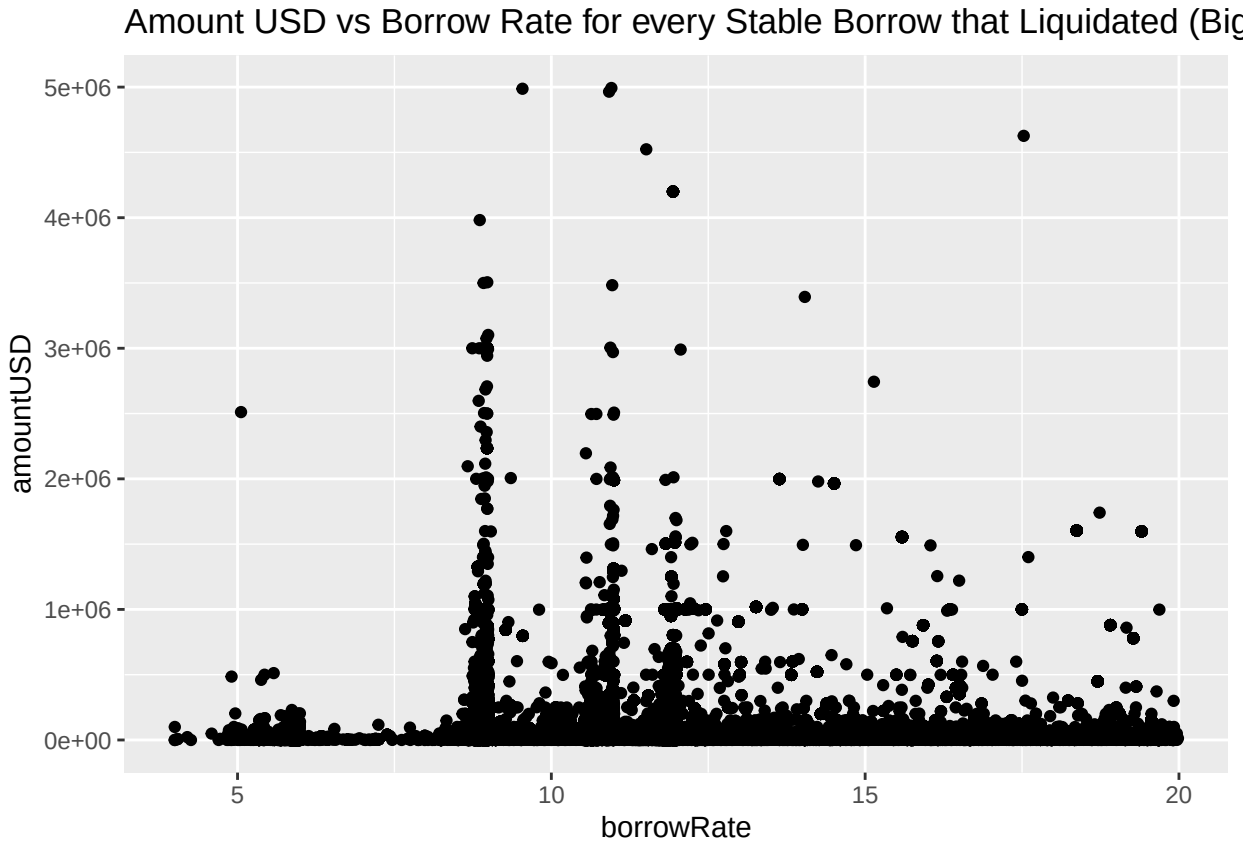

```
ggplot(liquidVariable,aes(x=borrowRate,y=amountUSD)) + geom_point() + ggtitle("Amount USD vs Borrow Rat
```




```
big_liquidVariable <- liquidVariable %>%
  filter(reserve=="USDC" | reserve=="USDT" | reserve=="DAI")

big_liquidStable <- liquidStable %>%
  filter(reserve=="USDC" | reserve=="USDT" | reserve=="DAI")

ggplot(big_liquidVariable,aes(x=borrowRate,y=amountUSD)) + geom_point() + ggtitle("Amount USD vs Borrow
```

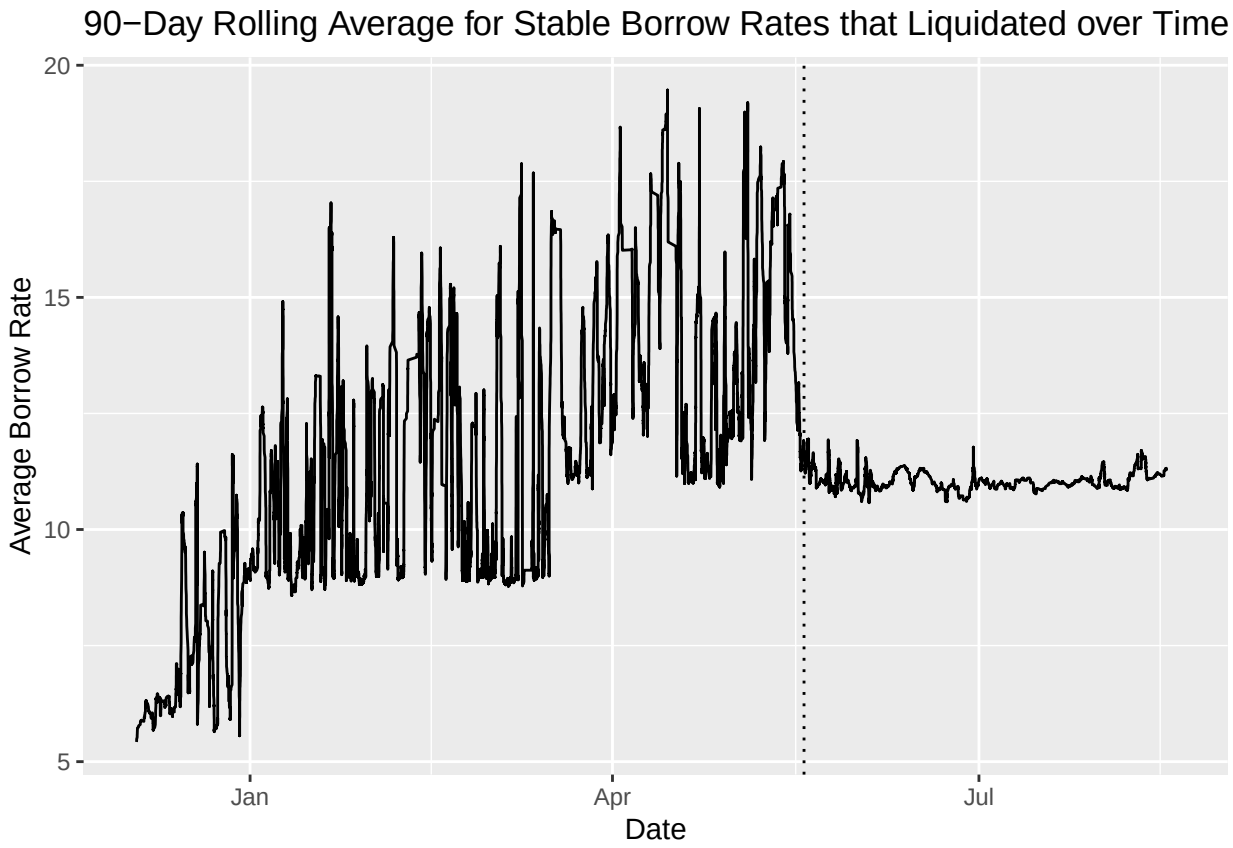



```
ggplot(big_liquidStable,aes(x=borrowRate,y=amountUSD)) + geom_point() + ggtitle("Amount USD vs Borrow Ra
```



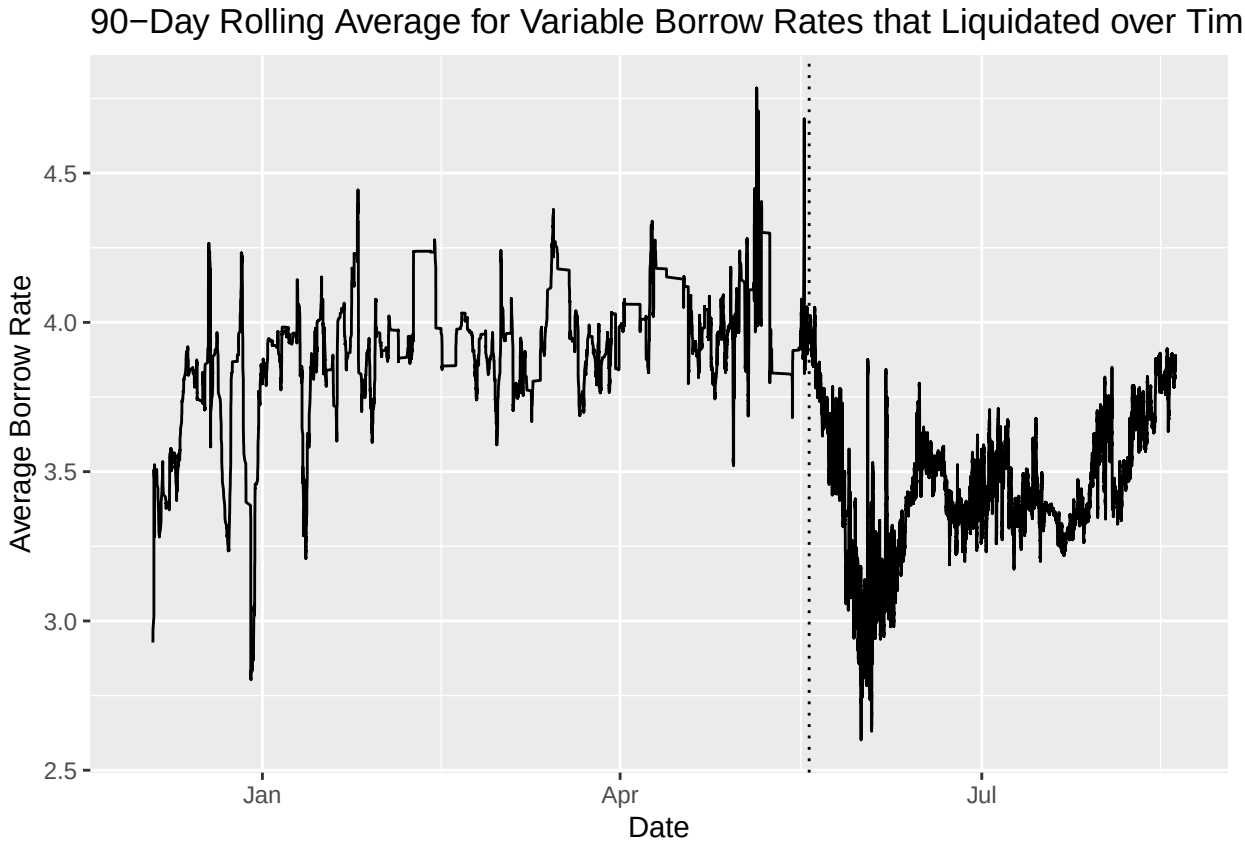

```
roll_big_liquidStable <- big_liquidStable %>%
  dplyr::arrange(timestamp) %>%
  dplyr::mutate(rollingAvg=zoo::rollmean(borrowRate,k=1,fill=NA)) %>%
  dplyr::mutate(rollingAvg21=zoo::rollmean(borrowRate,k=21,fill=NA)) %>%
  dplyr::mutate(rollingAvg90=zoo::rollmean(borrowRate,k=90,fill=NA))

ggplot(roll_big_liquidStable, aes(x=as_datetime(timestamp),y=rollingAvg90)) + geom_line() + xlab("Date")
```

```
## Warning: Removed 89 row(s) containing missing values (geom_path).
```




```
roll_big_liquidVariable <- big_liquidVariable %>%
  dplyr::arrange(timestamp) %>%
  dplyr::mutate(rollingAvg=zoo::rollmean(borrowRate,k=1,fill=NA)) %>%
  dplyr::mutate(rollingAvg21=zoo::rollmean(borrowRate,k=21,fill=NA)) %>%
  dplyr::mutate(rollingAvg90=zoo::rollmean(borrowRate,k=90,fill=NA))

ggplot(roll_big_liquidVariable, aes(x=as_datetime(timestamp),y=rollingAvg90)) + geom_line() + xlab("Dat
```

```
## Warning: Removed 89 row(s) containing missing values (geom_path).
```



Weekly Work Summary

- RCS ID: vyasj2
- Project Name: DeFi
- Summary of work: Corrected survival analysis graph and did additional analysis on borrow rates and their impact on liquidation
- Summary of GitHub commits
  - Branch name: dar-vyasj2
  - analysis.Rmd - complete analysis of past two weeks
- Presentations *https://docs.google.com/presentation/d/1eVkf-xFs-Wm57odNED6eebXvjaKUZdA-JZ2O7r2B7Co/edit#slide=id.gf71b1f45de_1_3

Personal Contribution

- I realized that my survival analysis was incorrect, because the line went all the way down to y=0, which is not what we want. So, I tweaked the code a little bit and used dplyr entirely to get the correct dataset, and now it looks more correct in the sense that it levels off at around 75%, meaning that 25% of all depositors borrowed soon after, and the remaining 75% of depositors did not borrow at all. This could account for yield farmers, because they do not borrow, they only loan their money out to accrue an interest over time.
- I then wanted to find some other variables that might affect liquidation, and I settled on looking at borrow rates, since the total amount borrowed (in ETH) is included in the health factor calculation, which is the component that determined whether or not someone liquidates. To get my data I split the original data into borrows and liquidations, and did a left join so that I got all the users that borrowed

and then liquidated. I decided to plot a few different plots, first finding the frequency of each type of borrow, Stable and Variable rate. Then, I split the data up into the two types, and plot some different graphs using that data. First I made a dotplot of the amountUSD vs the borrow rate, so I could get a general idea of how much was borrowed at what rate, that ended up being liquidated. I then split up the data into what I call the Big 3 - USDC, USDT, and DAI, since almost all borrows were done using one of those three reserves. Finally, I calculated a rolling average of 1 day, 21 days, and 90 days, and graphed the 90 day rolling average. From these graphs, I found an interesting result. I found that after some date in May (which we now know to be the day that China announced their crypto ban), the average borrow rate that led to liquidation leveled out, but the average variable rate behaved the exact opposite - it started jumping all over the place.

Future Steps

- The next steps would be to interpret this result, and possibly see if there are other aspects in the data that were impacted by this event as well.