DAR F21 Project Status Notebook:

DeFi

Roman Vakhrushev (vakhrr)

10/03/2021

Contents

Biweekly Work Summary . . . 1
Personal Contribution . . . 1
Discussion of Primary Findings . . . 1

Biweekly Work Summary

- RCS ID: vakhrr
- Project Name: Blockchain DeFi
- These two weeks I was working on analyzing coins for liquidation
- First week I started by analyzing the relationship between principal and collateral in liquidation transactions
- Second week I had to fix several errors in the initial analysis and then I made some small analysis on the transactions related to liquidation over time
- Branch: dar-vakhrr

Personal Contribution

All contributions were completed by me.

Discussion of Primary Findings

What did you want to know?

I wanted to study how the distribution of coins for liquidation looks like, understand what coins are used more as collateral and principal and what coins are not used at all. I also wanted to see how liquidation transactions are distributed over time and verify if there are any interesting trends there.

How did you go about finding it?

I decided to analyze the data using pairs of coins (collateral, principal). I had to create new data frames out of the initial data frame, extended them with new columns, and built graphs and tables out of it. For analysis of transactions over time I also created new data frames and build visualizations.

What did you find?

```
#data collection as always
df<-read_rds('../../Data/transactions.Rds')
# Use deplyr to drop NA reserves, add the counts and then kep only the top 20
reservecoins <- df %>%  drop_na(reserve) %>%
count(reserve) %>%
```

```r
arrange(-n) %>%
head(20)
```

```r
#Create cointype function to label all coins by types correctly,
#This is useful for visualizations

coinType <- function(coin) {
    #stable_coins <- list("USDC","USDT","DAI","BUSD","SUSD","GUSD","TUSD")
    if(str_contains(coin,"USD",ignore.case = TRUE))
    {
      result = "stable"
    }
    else if(str_contains(coin,"DAI",ignore.case = TRUE))
    {
      result = "stable"
    }
    else
    {
      result = "non-stable"
    }
    return(result)
}

#Main data frame for analysis of liquidations, includes information on total and
#average values of collateral and principal

LiquidSummary <- df%>% filter(type == "liquidation") %>% group_by(collateralReserve, principalReserve)

LiquidSummary <- df%>% filter(type == "liquidation") %>% group_by(collateralReserve, principalReserve) %
```

```
## `summarise()` has grouped output by 'collateralReserve'. You can override using the `.groups` argumen
```

```r
#These are just ordered versions of LiquidSummary too see what coins are used
#more for liquidations
LiquidSummary <- LiquidSummary[order(-LiquidSummary$total_usd_princ),]
print(head(LiquidSummary,5))
```

```
## # A tibble: 5 x 13
## # Groups:   collateralReserve [3]
##   collateralReserve principalReserve avg_usd_princ total_usd_princ avg_usd_collat
##   <fct>             <fct>                    <dbl>           <dbl>          <dbl>
## 1 WETH              USDC                   104443.       89925227.       1.10e-13
## 2 WETH              USDT                    87256.       64045874.       9.16e-14
## 3 WBTC              USDC                   206599.       51029921.       2.27e+ 5
## 4 WETH              DAI                     68264.       46009992.       7.17e-14
## 5 LINK              USDC                    44420.       24608887.       4.89e+ 4
## # ... with 8 more variables: total_usd_collat <dbl>, avg_eth_princ <dbl>,
## #   total_eth_princ <dbl>, avg_eth_collat <dbl>, total_eth_collat <dbl>,
## #   collateralType <chr>, principalType <chr>, n <int>
```

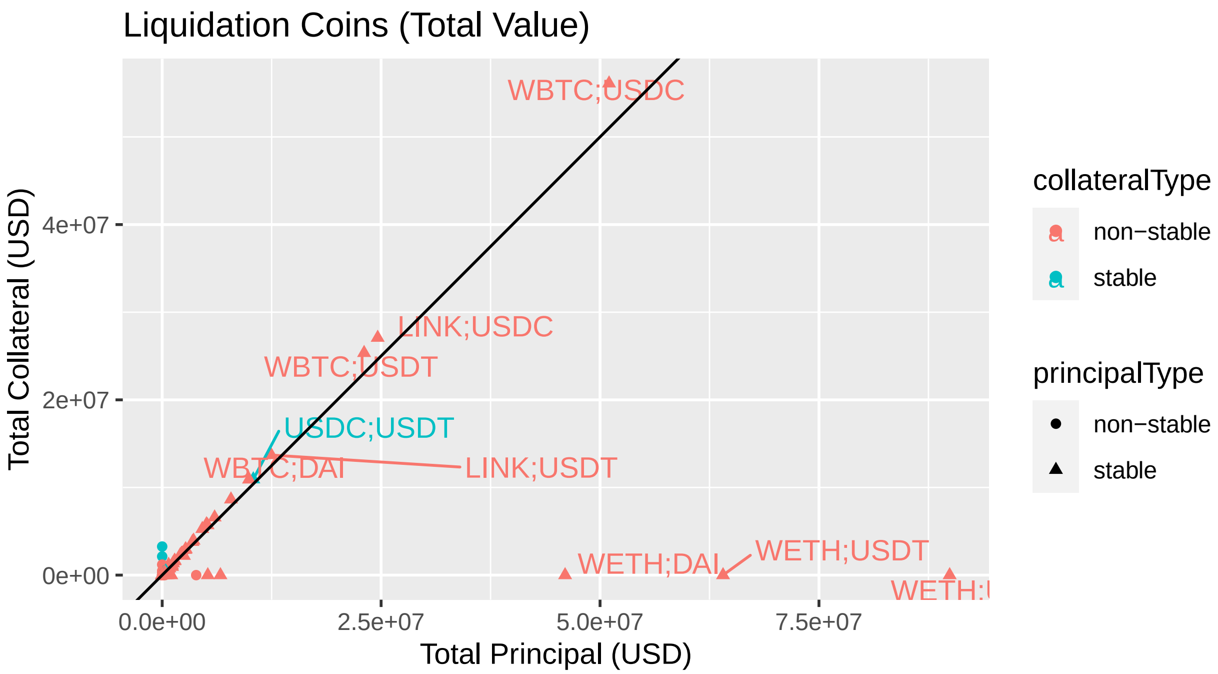
```r
LiquidSummary <- LiquidSummary[order(-LiquidSummary$total_usd_collat),]
print(head(LiquidSummary,5))
```

```
## # A tibble: 5 x 13
## # Groups:   collateralReserve [3]
##   collateralReserve principalReserve avg_usd_princ total_usd_princ avg_usd_collat
```

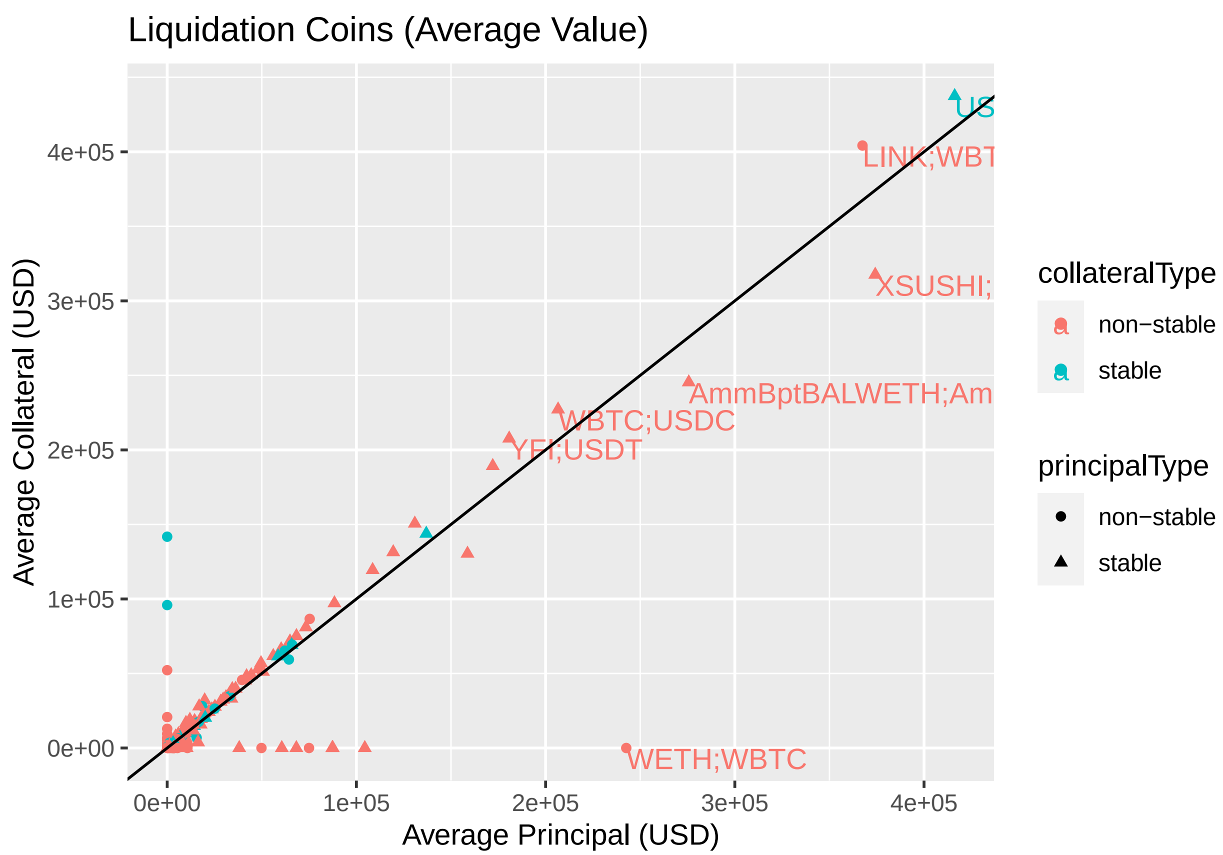
```
##   <fct>             <fct>                         <dbl>          <dbl>         <dbl>
## 1 WBTC              USDC                        206599.       51029921.       227173.
## 2 LINK              USDC                         44420.       24608887.        48893.
## 3 WBTC              USDT                        172065.       23056656.       189227.
## 4 LINK              USDT                         41244.       12455699.        45359.
## 5 USDC              USDT                        416198.       10404949.       437447.
## # ... with 8 more variables: total_usd_collat <dbl>, avg_eth_princ <dbl>,
## #   total_eth_princ <dbl>, avg_eth_collat <dbl>, total_eth_collat <dbl>,
## #   collateralType <chr>, principalType <chr>, n <int>
```

These tables provide the initial analysis for liquidation coins. The first table shows the first five entries of the data frame ordered by total principal (in USD), while the second table shows the same but for total collateral (in USD). From these, we can initially draw some basic conclusion. First, almost all combinations of coins have non-stable coin for collateral and stable coin for principal. Second, we can observe popular coins for collateral and principal, these are WBTC, WETH for collateral and USDC, USDT for principal. Lastly, we see that different combinations of coins, even ordered, are still quite different. For example, combinations (WETH,USDT) and (WBTC,USDC) are very different in average value for principal while quite close in total value for principal.

```
#ggplotly() alternatives are commented out to ease reading

#Build plots for total principal/ total collateral (USD)

p <- ggplot(LiquidSummary,aes(x = total_usd_princ, y = total_usd_collat,color = collateralType, shape =

p <- p + coord_equal(ratio = 1)

#Build plots for total principal/ total collateral (ETH)
p_eth <- ggplot(LiquidSummary,aes(x = total_eth_princ, y = total_eth_collat,color = collateralType, shap



p_eth <- p_eth + coord_equal(ratio = 1)


LiquidNCollat <- LiquidSummary %>% group_by(collateralReserve) %>% summarize(total_liquidations = sum(n)

LiquidNCollat <- LiquidNCollat[order(-LiquidNCollat$total_liquidations),]

LiquidNPrinciple <- LiquidSummary %>% group_by(principalReserve) %>% summarize(total_liquidations = sum

LiquidNPrinciple <- LiquidNPrinciple[order(-LiquidNPrinciple$total_liquidations),]

p
```




```
#This graph is almost the same, so commented out
#p_eth

#ggplotly(p)

p_avg <- ggplot(LiquidSummary,aes(avg_usd_princ,avg_usd_collat,color = collateralType, shape = principal

#ggplotly(p2)

p_avg
```



The two graphs support our observations from the tables above. Additionally, we can see some other interesting trends.

From the graph on the total value we can deduce that almost all combinations of coins (except for erroneous WETH combinations) lie above the diagonal identity line. This means that in total value, collateral is always higher than principal (the ratio is actually about 1.1-1.15 for most combinations).

However, the graph on the average values shows something even more interesting: there are some combinations of coins (not including WETH ones), where average principal is higher than average collateral (for example (XSUSHI,BUSD)). This is very interesting thend that cannot be explained from just mathematics: if every transaction has a higher principal than collateral then the total should also behave the same way, which implies average should also follow the same trend. Aaron proposed an interesting explanation to this: if there are multiple borrows with different types of collateral/principle taken for the same user simultaneously, the health factor might be computed based on different coins, which would potentially allow for this situation to happen. However, this still does not explain why liquidators chose these coins because this is not profitable economically.

```
#Show transactions, where collateral<principal (exclude WETH for now).
dfst <- df %>% filter(type == "liquidation") %>% filter(collateralReserve != "WETH") %>% filter(principa

count(dfst)
```

```
##     n
## 1 169
```

```
#Show just random 10 of those (exclude some data)
dfst %>% select(collateralReserve,principalReserve,amountUSDCollateral,amountUSDPincipal) %>% head(10)
```

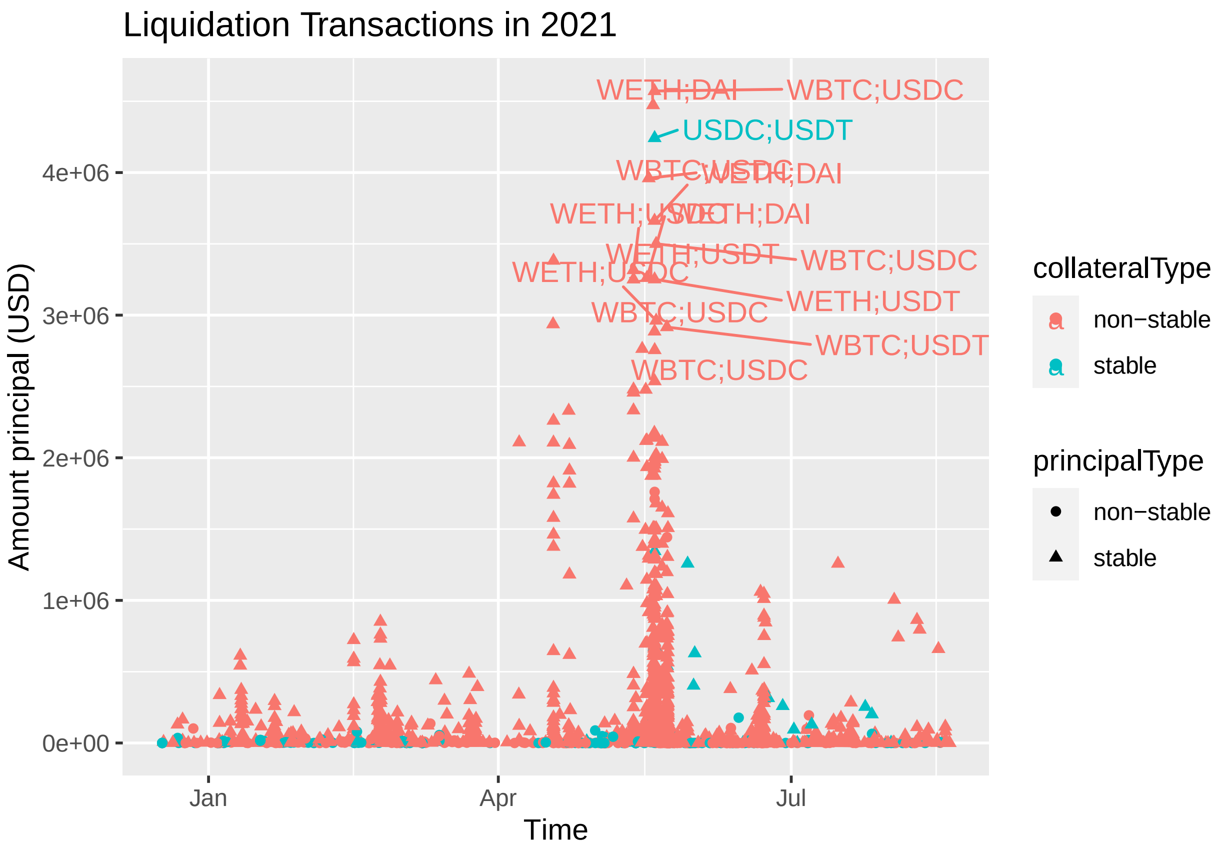
```
##    collateralReserve principalReserve amountUSDCollateral amountUSDPincipal
## 1                ENJ             USDT        1.886249e+04        55144.5691
```

```
## 2      AmmBptBALWETH          AmmUSDC        9.963162e+05      1198318.9516
## 3                ZRX              DAI        3.698066e+03         8339.8518
## 4      AmmBptBALWETH           AmmDAI        1.481574e+02          150.9005
## 5      AmmBptBALWETH           AmmDAI        4.627767e+02          513.0650
## 6               WBTC             GUSD        8.947895e+02         1350.9114
## 7                DAI              ENJ        2.845527e+03         2889.4629
## 8             XSUSHI              DAI        2.871827e+04        30281.5916
## 9            AmmWETH           AmmDAI        4.262961e-14        40599.6245
## 10           AmmWETH          AmmUSDT        3.277788e-16          312.1703
```

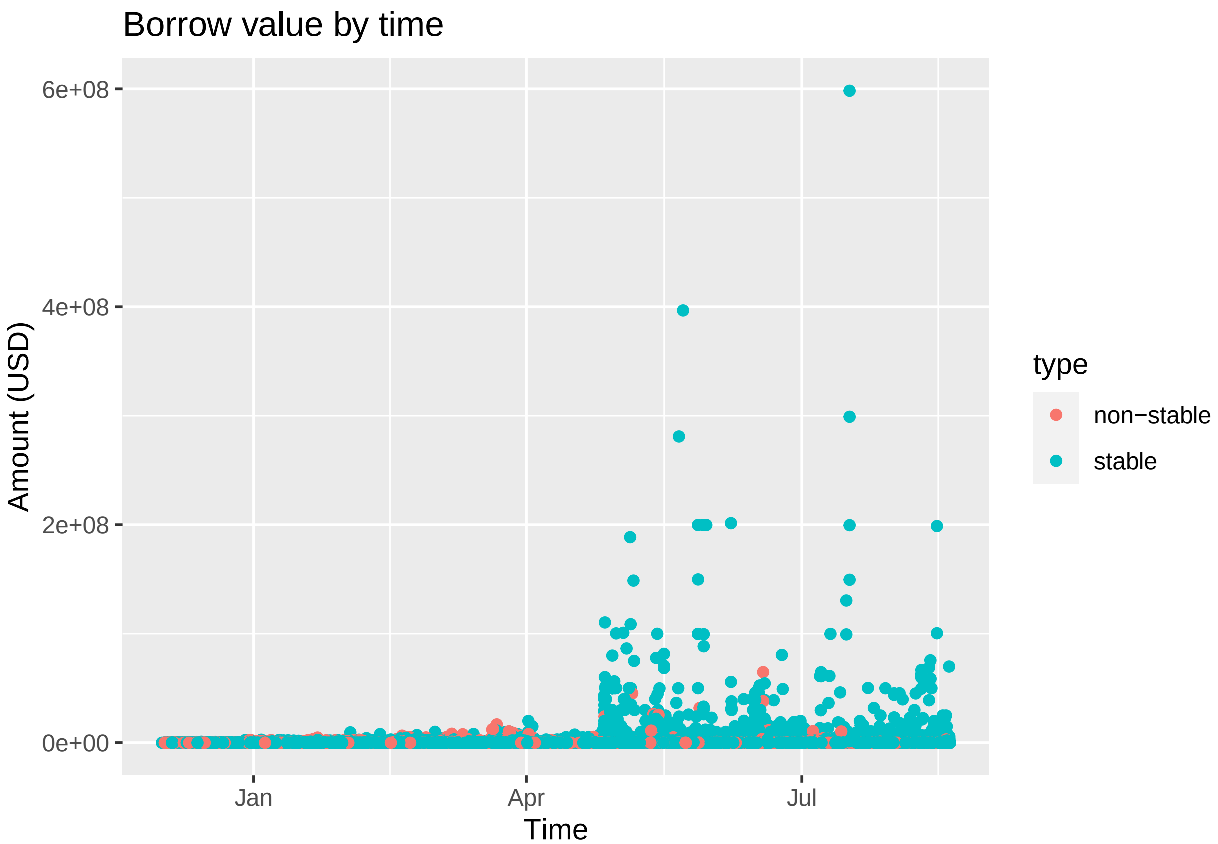
Continuing this discussion on strange behavior of some coin combinations, we can take a look at individual liquidation transactions. This includes all transactions where collateral<principal and this also excludes WETH. Turns out there are 169 transaction like that. Almost all transactions include at least one unpopular coin (which means this particular coin combination is also unpopular), which might explain this partially. This behavior is to be analyzed in details next week.

```r
#df<-df %>% filter(type == "liquidation") %>% group_by(timestamp, reserveUSDPrincipal) %>% filter(colla

#df <- df %>% mutate(timestamp = as_date(timestamp))

#LiquidSummary2 and dfl are dataframes for analyzing data over time

LiquidSummary2 <- LiquidSummary %>% mutate(collat_princ_rat = total_usd_collat/total_usd_princ)

#LiquidSummary2

dfl<-df %>% filter(type == "liquidation")

dfl$collateralType <- mapply(coinType, dfl$collateralReserve)

dfl$principalType <- mapply(coinType, dfl$principalReserve)

#Change timestamp to show actual dates instead of timestamps
dfl<-dfl %>% mutate(timestamp = as_datetime(timestamp, tz = "UTC"))

#head(dfl,20)

#make a plot for liquidations over time (color and shape appropriate coin types)
p <- ggplot(dfl,aes(x = timestamp, y = amountUSDPincipal,color = collateralType, shape = principalType)

p
```

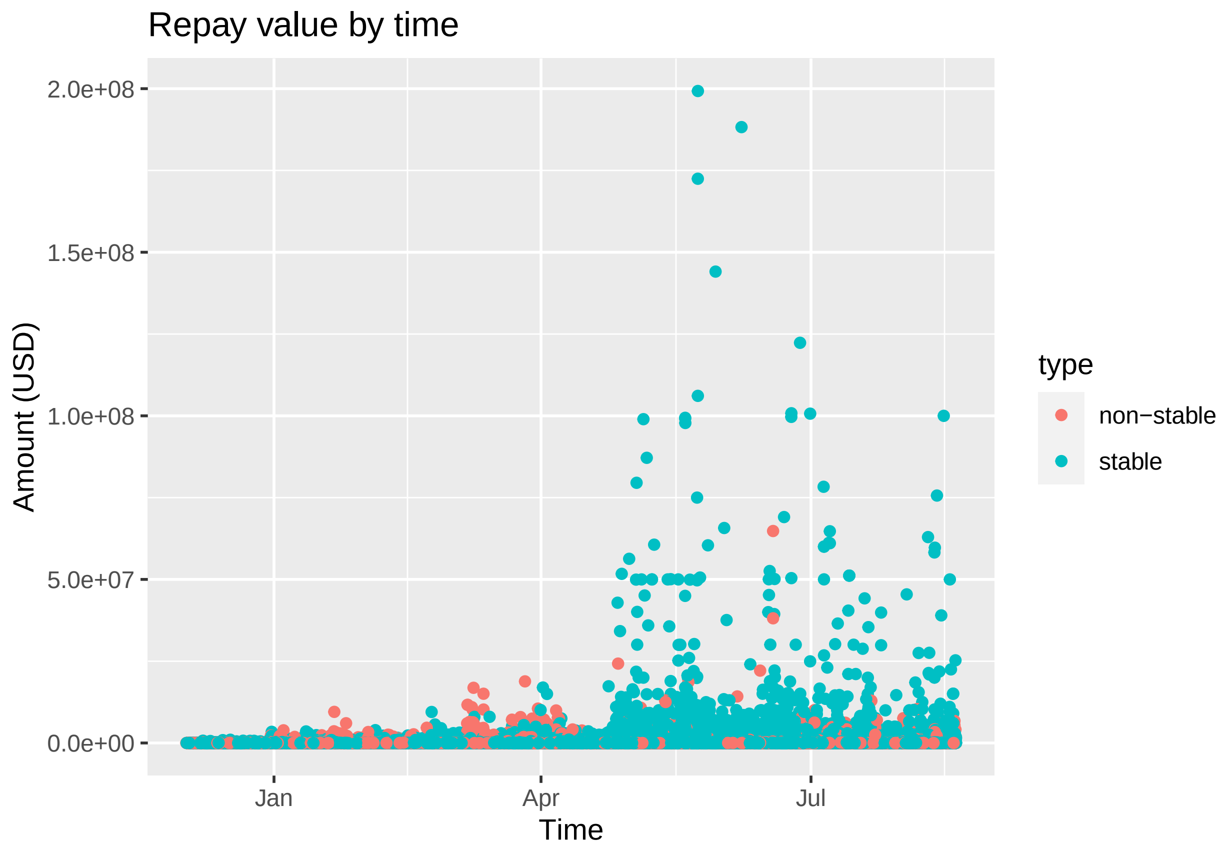
```
## Warning: ggrepel: 24 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```




```r
#Same plots to be done for p2 and p3

BorrowSummary <- df %>% filter(type == "borrow")

BorrowSummary$type <- mapply(coinType,BorrowSummary$reserve)

p2 <- ggplot(BorrowSummary,aes(x = as_datetime(timestamp, tz = "UTC"), y = amountUSD, color = type)) + g

p2
```




```
RepaySummary <- df %>% filter(type == "repay")

RepaySummary$type <- mapply(coinType,RepaySummary$reserve)

p3 <- ggplot(RepaySummary,aes(as_datetime(timestamp, tz = "UTC"), y = amountUSD, color = type)) + geom_p

p3
```



These graphs show different types of transactions over time. I was focusing more on liquidations, but also made similar graphs for borrow and repay for comparison. The liquidation graph looks very interesting. In general, it is more or less flat, sometimes we see some spikes, but there is a big spike around the middle of May, 2021. This period is very different on the graph: we see both a lot more transactions than usual and also a lot of transactions of high volume. This hypothesis for this event is ban of cryptocurrency operations by China, which happened on May 18. It looks like WETH and WBTC coins got liquidated a lot, but liquidations actually occurred in many different coin combinations.

Unlike the liquidation graph, borrow and repay graphs look differently. We see some increase in total number of transactions (and also value) starting from early May, but we do not see so high spikes and also transaction behavior remains the same after May unlike liquidation transactions, which got back to normal level after May. The reason could be high volatility of liquidation market (even slight difference causes great changes to the market).