CAD Time Stamp Analysis

1 CAD Time Stamp Analysis

The following analysis is desinged to look at the time stamps related to service calls for the past 13 months and determine the significance of event sequencing and investigate the challenges presented when calls for service are manipulated after the first call closure.

1.1 Author's Note

The data was harvested from the CentralSquare Enterprise CAD Reporting Database. The two database tables that were queried are Response_Master_Incident and Response_Master_Incident_Ext. These two tables contain all of the time stamps needed for the analysis. For this analysis, the data consists of call information, including the method of call reception, priority number, and the problem type from 01 January 2024 through 31 January 2025. This dataset consists of 178780 rows prior to any cleaning of adjustments.

1.2 Data Cleaning and Preparation

To ensure the data will be useful to us, the first step is cleaning. In this instance, after extracting the data from the archive database, cleaning was done in the Excel file, including removing the "NULL" values and replacing them with empty cells. This forces R to insert its own NULL value and allows it to correctly assign the data types to the columns. Here is a view of that so you can see what we have.

```
tibble [178,780 × 50] (S3: tbl_df/tbl/data.frame)
 $ ID              : num [1:178780] 2659640 2659641 2659642 2659643 2659644 ...
 $ Reception       : chr [1:178780] "E-911" "E-911" "E-911" "E-911" ...
 $ Problem         : chr [1:178780] "911 TRANSFER VSP" "BLS EMERGENCY" NA NA ...
 $ Priority        : num [1:178780] 7 3 NA NA 2 NA 2 1 NA NA ...
 $ Agency          : chr [1:178780] "LAW" "FIRE" "DEC" "DEC" ...
 $ Jurisdiction    : chr [1:178780] "Alexandria PD" "200 ALX" "DEC" "DEC" ...
 $ Response_Date   : POSIXct[1:178780], format: "2024-01-01 00:01:09" "2024-01-01 00:02:05" ...
 $ Clock_Start     : POSIXct[1:178780], format: "2024-01-01 00:01:09" "2024-01-01 00:02:05" ...
 $ Phone_Start     : POSIXct[1:178780], format: "2024-01-01 00:01:08" "2024-01-01 00:02:04" ...
 $ Fixed_Phone_Start: POSIXct[1:178780], format: "2024-01-01 00:01:08" "2024-01-01 00:02:04" ...
 $ First_Keystroke : POSIXct[1:178780], format: "2024-01-01 00:01:09" "2024-01-01 00:02:05" ...
 $ DELTA_RD_CST    : num [1:178780] 0 0 0 0 0 0 22 0 0 0 ...
 $ DELTA_RD_TPU    : num [1:178780] -1 -1 0 0 0 0 0 -2 0 0 ...
 $ DELTA_RD_FTPU   : num [1:178780] -1 -1 0 0 0 0 NA -2 0 0 ...
 $ DELTA_RD_FCTK   : num [1:178780] 0 0 0 0 4 0 0 9 0 0 ...
 $ DELTA_CST_TPU   : num [1:178780] -1 -1 0 0 0 0 -22 -2 0 0 ...
 $ DELTA_CST_FTPU  : num [1:178780] -1 -1 0 0 0 0 NA -2 0 0 ...
 $ DELTA_CST_FCTK  : num [1:178780] 0 0 0 0 4 0 -22 9 0 0 ...
```

```
$ DELTA_TPU_FTPU   : num [1:178780] 0 0 0 0 0 0 NA 0 0 0 ...
$ DELTA_TPU_FCTK   : num [1:178780] 1 1 0 0 4 0 0 11 0 0 ...
$ DELTA_FTPU_FCTK  : num [1:178780] 1 1 0 0 4 0 NA 11 0 0 ...
$ Viewed           : POSIXct[1:178780], format: NA "2024-01-01 00:04:33" ...
$ Queued           : POSIXct[1:178780], format: "2024-01-01 00:01:44" "2024-01-01 00:03:11" ...
$ Fixed_Queued     : POSIXct[1:178780], format: "2024-01-01 00:01:44" "2024-01-01 00:03:11" ...
$ DELTA_QUEUES     : num [1:178780] 0 0 0 0 0 0 0 0 0 0 ...
$ Phone_End        : POSIXct[1:178780], format: "2024-01-01 00:01:45" "2024-01-01 00:08:57" ...
$ Fixed_Phone_End  : POSIXct[1:178780], format: "2024-01-01 00:01:45" "2024-01-01 00:08:57" ...
$ DELTA_CALLTAKING : num [1:178780] 0 0 0 0 0 0 0 0 0 0 ...
$ First_Dispatch   : POSIXct[1:178780], format: NA "2024-01-01 00:03:15" ...
$ First_Ack        : POSIXct[1:178780], format: NA "2024-01-01 00:05:13" ...
$ First_Enroute    : POSIXct[1:178780], format: NA "2024-01-01 00:05:13" ...
$ First_Arrived    : POSIXct[1:178780], format: NA "2024-01-01 00:10:32" ...
$ First_Clear      : POSIXct[1:178780], format: NA "2024-01-01 01:40:47" ...
$ Call_Closed      : POSIXct[1:178780], format: "2024-01-01 00:01:45" "2024-01-01 01:40:48" ...
$ Fixed_Call_Closed: POSIXct[1:178780], format: "2024-01-01 00:01:45" "2024-01-01 01:40:48" ...
$ STEP_1           : num [1:178780] -1 -1 0 0 0 0 -22 -2 0 0 ...
$ STEP_2           : num [1:178780] 1 1 0 0 4 0 0 11 0 0 ...
$ STEP_2A          : num [1:178780] NA 148 NA NA 19 NA 22 57 NA NA ...
$ STEP_2B          : num [1:178780] NA -82 NA NA -2 NA -22 -9 NA NA ...
$ STEP_3           : num [1:178780] 35 66 37 25 17 197 0 48 19 101 ...
$ CALL_TIME        : num [1:178780] 37 413 37 25 175 197 0 350 19 102 ...
$ STEP_4           : num [1:178780] NA 4 NA NA 21 NA 3 24 NA NA ...
$ STEP_5           : num [1:178780] NA 118 NA NA 2 NA 94 18 NA NA ...
$ STEP_6           : num [1:178780] NA 0 NA NA 0 NA 0 0 NA NA ...
$ STEP_7           : num [1:178780] NA 319 NA NA 193 NA 272 224 NA NA ...
$ STEP_8           : num [1:178780] NA 5415 NA NA 3210 ...
$ STEP_9           : num [1:178780] NA 0 NA NA 20928 ...
$ STEP_10          : num [1:178780] NA 5415 NA NA 24138 ...
$ TOTAL_TIME_A     : num [1:178780] 36 5922 37 26 24375 ...
$ TOTAL_TIME_B     : num [1:178780] 36 5922 37 26 24375 ...
```

This leaves us well positioned because we now have the character (chr) columns, numeric (num) columns, and time stamps (POSIXct) defined. The numeric columns will be expressed as BIGINT in the background which is what we want for now.

1.2.1 Column Definitions

The columns for this dataset are defined as follows:

> **Important**
>
> Note that any negative values in the deltas occur when the time for the rightmost variable is earlier than the time recorded for the leftmost variable.

1. ID: *This is the ID column from the CAD. We will use it to isolate specific calls for further analyses.*
2. Reception: *This is how the call entered DECC.* ***'Null'*** *Values will be changed to 'Not Recorded' for analytical purposes.*
3. Problem: *This is the final problem nature as updated by the parties working the call*

4. Priority: *This is the relative importance level assigned by DECC. This will be used for further analyses*
5. Agency: *This indicates if the call was for LAW, FIRE, or DEC.*
6. Jurisdiction: *This indicates for whom the call was run.*
7. Response_Date: *This, per CS, is the timestamp indicating when the ID was created and is frequently used, by them, as the **start** of the call.*
8. Clock_Start: *This should be, per CS, the **actual** starting point for the call.*
9. Phone_Start: *This timestamp identifies when the CAD acknowledged the start of the phone call for the event.*
10. Fixed_Phone_Start: *This timestamp also identifies when the CAD acknowledged the start of the phone call for the event, but is deemed **immutable** by the software.*
11. First_Keystroke: *This timestamp identifies when the CAD software registers the first keystroke related to the event*
12. DELTA_RD_CST: *This is the difference, hence Δ, between the Response_Date and ClockStartTime columns.*
13. DELTA_RD_TPU: *This is the difference between the Response_Date and the Time_PhonePickUp columns.*
14. DELTA_RD_FTPU: *This is the difference between the Response_Date and the Fixed_Time_PhonePickUp columns.*
15. DELTA_RD_FCTK: *This is the difference between the Response_Date and the Time_FirstCallTakingKeystroke columns.*
16. DELTA_CST_TPU: *This is the difference between the ClockStartTime and the Time_PhonePickUp columns.*
17. DELTA_CST_FTPU: *This is the difference between the ClockStartTime and the Fixed_Time_PhonePickUp columns.*
18. DELTA CST_FCTK: *This is the difference between the ClockStartTime and the Time_FirstCallTakingKeystroke columns.*
19. DELTA_TPU_FTPU: *This is the difference between the Time_PhonePickUp and Fixed_Time_PhonePickUp columns.*
20. DELTA_TPU_FCTK: *This is the difference between the Time_PhonePickUp and Time_FirstCallTakingKeystroke columns.*
21. DELTA_FTPU_FCTK: *This is the difference between the Fixed_Time_PhonePickUp and Time_FirstCallTakingKeystroke columns.*
22. Viewed: *This is defined by the vendor as the time the call was first viewed in the Call Taking Form.*
23. Queued: *This is the time when the call has been released to be dispatched by the calltaker.*
24. Fixed_Queued: *This is the **fixed** time when the call has been released to be dispatched by the calltaker.*
25. DELTA_QUEUES: *This is the difference between the Time_CallEnteredQueue and Fixed_Time_CallEnteredQueue columns.*
26. Phone_End: *This is the time when the caller is released.*
27. Fixed_Phone_End: *This is the **fixed** time when the caller is released.*
28. DELTA_CALLTAKING: *This is the difference between the Time_CallTakingComplete and Fixed_Time_CallTakingComplete columns.*
29. First_Dispatched: *This is the time when the dispatcher assigned the first unit to the call.*
30. First_Ack: *This is the time when the first unit acknowledges receiving the dispatched call.*
31. First_Enroute: *This is the time the first dispatched unit marks or is marked 'En Route' for the assigned call.*
32. First_Arrived: *This is the time the first dispatched unit marks or is marked 'Arrived' for the assigned call.*
33. First_Clear: *This is the time that the first unit clears or leaves the call after arrival.*
34. Call_Closed: *This is the time that the final unit is cleared from the call, closing the event out.*

35. Fixed_Call_Closed: *This is the **fixed** time that the final unit is cleared from the call, closing the event out.*
36. STEP_1: *This is the time difference between the ClockStartTime and the Time_PhonePickUp columns.*
37. STEP_2: *This is the time difference between the Time_PhonePickUp and the Time_FirstCallTakingKeystroke columns.*
38. STEP_2A: *This is the time difference between the Time_PhonePickUp and TimeCallViewed columns.*
39. STEP_2B: *This is the time difference between the TimeCallViewed and Time_FirstCallTakingKeystroke columns.*
40. STEP_3: *This is the time difference between the Time_FirstCallTakingKeystroke and Time_CallEnteredQueue columns.*
41. CALL_TIME: *This is the time difference between the Time_PhonePickUp and Time_CallTakingComplete columns.*
42. STEP_4: *This is the time difference between the Time_CallEnteredQueue and Time_First_Unit_Assigned columns.*
43. STEP_5: *This is the time difference between the Time_First_Unit_Assigned and TimeFirstUnitDispatchAcknowledged columns.*
44. STEP_6: *This is the time difference between the TimeFirstUnitDispatchAcknowledged and Time_First_Unit_Enroute columns.*
45. STEP_7: *This is the time difference between the Time_First_Unit_Enroute and Time_First_Unit_Arrived columns.*
46. STEP_8: *This is the time difference between the Time_First_Unit_Arrived and TimeFirstCallCleared columns.*
47. STEP_9: *This is the time difference between the TimeFirstCallCleared and Time_CallClosed columns.*
48. STEP_10: *This is the time difference between the Time_First_Unit_Arrived and Time_CallClosed columns.*
49. TOTAL_TIME_A: *This is the time difference between the Response_Date and Time_CallClosed columns.*
50. TOTAL_TIME_B: *This is the time difference between the ClockStartTime and Time_CallClosed columns.*

> **Note**
>
> Please note that there should not be, if calls are not re-opened, any delta between a time stamp and its 'Fixed' counterpart.

In the list above, the Steps columns are meant to follow a call from the time the clock starts to the time the clock stops. I've identified a possible 10 step timing chain that would allow us to step through the history of a call. There are multiple possible steps for the second step depending on management preferences. Once we have established the chain time wise, then we can use these steps as a tool to investigate specific events and to get better analyses and insights from processes overall. This can also assist in trend analyses. Where there are no entries for some of these time points, we may need to create alternative elapsed time points to create our steps. Those decisions are part of this project's goals.

You will note there are some columns duplicated because they are meant to be used for different analyses using the same dataset. Traditionally, we shouldn't have to do this, but I chose to make this distinction for ease of use.

1.3 Data Analysis

1.3.1 Missing Data

Looking at the dataset, here is a graph of missing data:
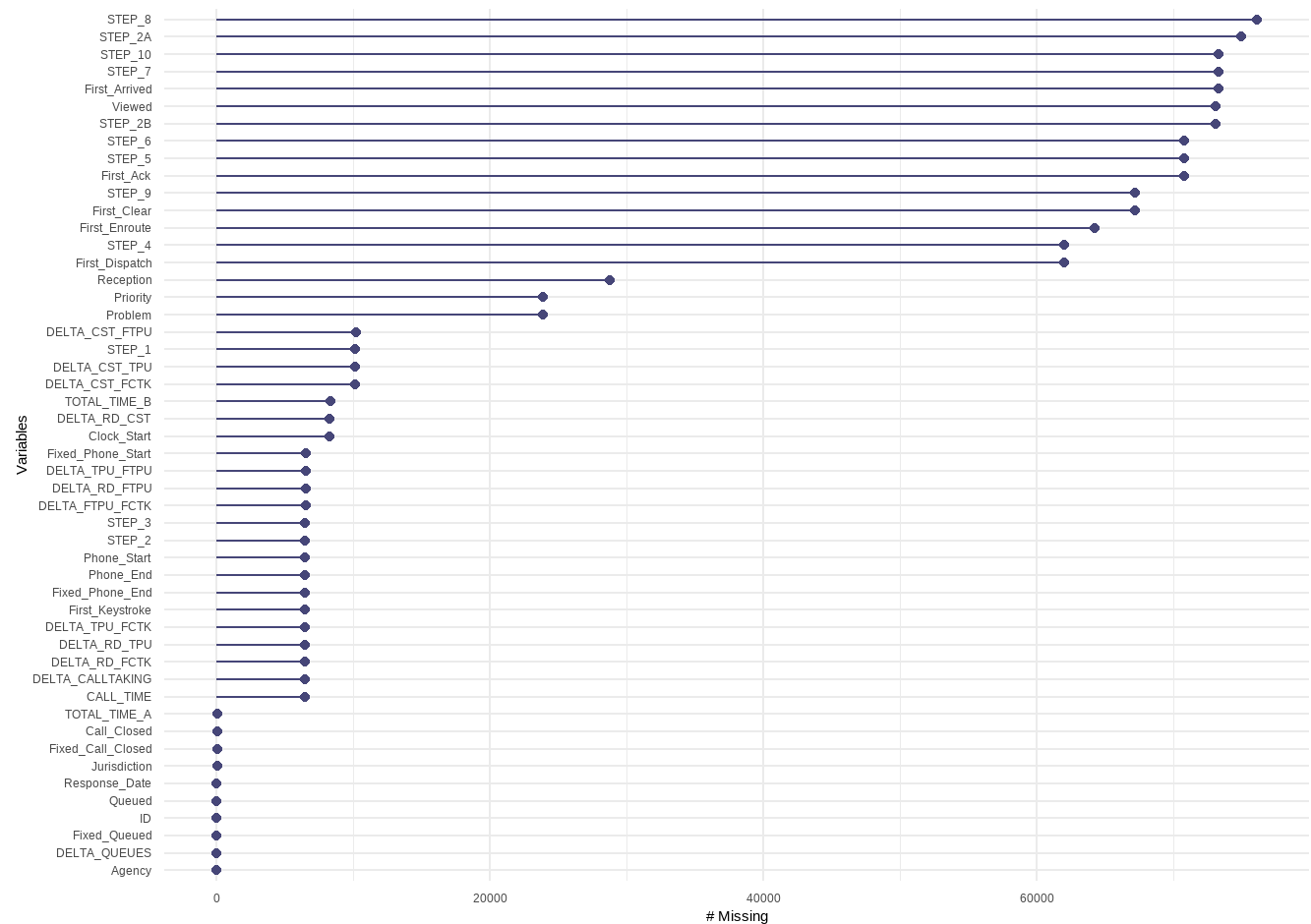
The graph shows that we have over 70,000 calls where we never recorded an arrival to the event. We might need to add a disposition for the calls to see if there is a correlation between missing values and disposition codes. I've included a printout of the number of missing rows by column to assist in the numbers.

ID	Reception	Problem	Priority
0	28779	23858	23861
Agency	Jurisdiction	Response_Date	Clock_Start
0	77	0	8241
Phone_Start	Fixed_Phone_Start	First_Keystroke	DELTA_RD_CST
6450	6516	6450	8241
DELTA_RD_TPU	DELTA_RD_FTPU	DELTA_RD_FCTK	DELTA_CST_TPU
6450	6516	6450	10168
DELTA_CST_FTPU	DELTA_CST_FCTK	DELTA_TPU_FTPU	DELTA_TPU_FCTK
10234	10168	6516	6450
DELTA_FTPU_FCTK	Viewed	Queued	Fixed_Queued
6516	73125	0	0
DELTA_QUEUES	Phone_End	Fixed_Phone_End	DELTA_CALLTAKING
0	6450	6450	6450
First_Dispatch	First_Ack	First_Enroute	First_Arrived
62044	70820	64274	73291
First_Clear	Call_Closed	Fixed_Call_Closed	STEP_1

67202	95	82	10168
STEP_2	STEP_2A	STEP_2B	STEP_3
6450	75013	73125	6450
CALL_TIME	STEP_4	STEP_5	STEP_6
6450	62044	70821	70823
STEP_7	STEP_8	STEP_9	STEP_10
73291	76129	67204	73292
TOTAL_TIME_A	TOTAL_TIME_B		
95	8334		

With this information, we will need to address some of the NULL values in order to determine if there are any correlations. Immediately, you can see that the steps where the arrival time is missing also have missing values.

We can repeat the graph to ensure that there are no longer any missing values for these four columns.
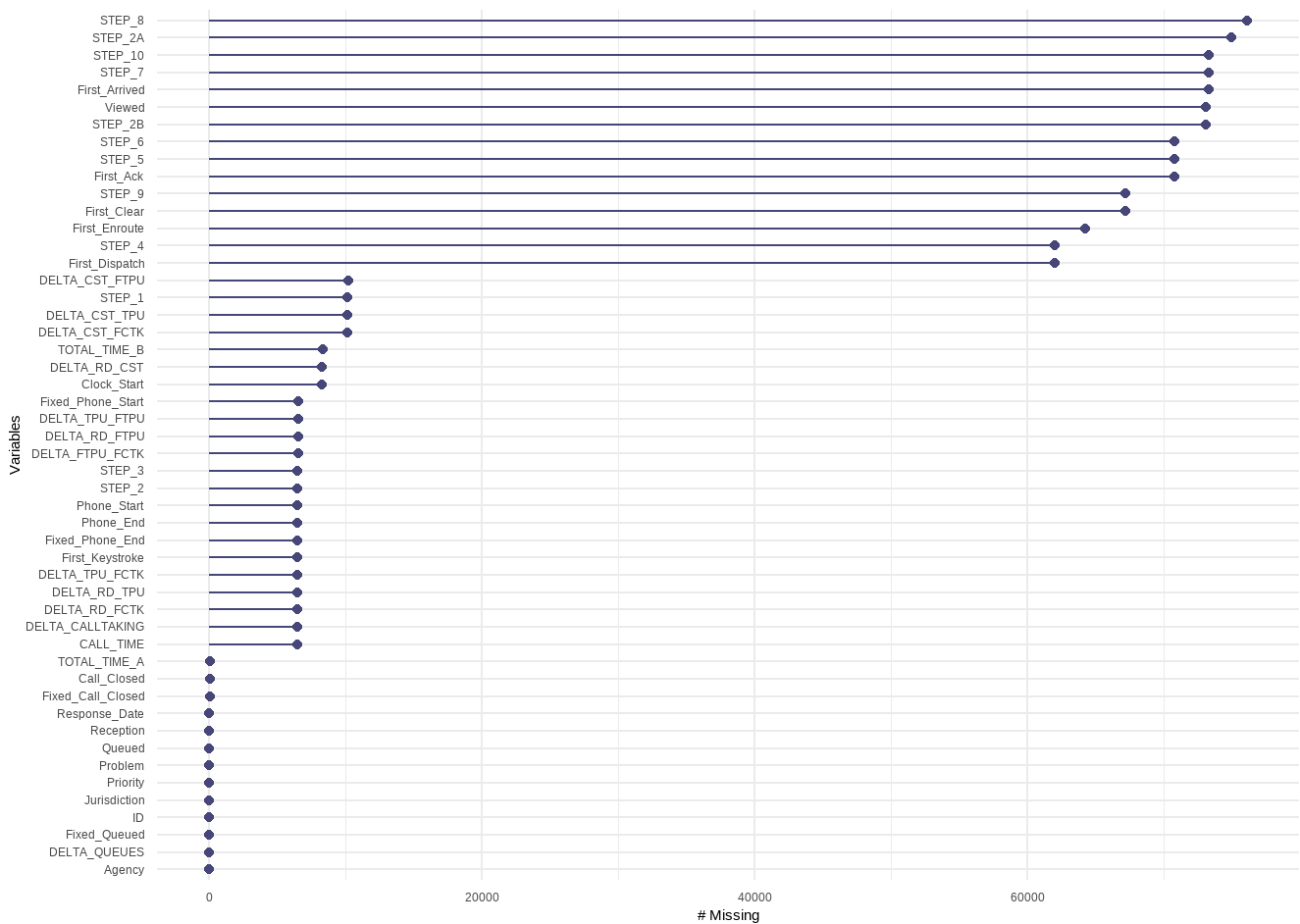
Now that we have those addressed, let's look at how many negative values we have across the thirteen month period. For this, I'm going to create a variable for each of the numeric columns, except Priority, and calculate a percentage of the data we do have that is negative and missing. That table can be found below:

Column	Negative Rows	NA Rows	Neg Percentage	NA Percentage
DELTA_RD_CST	4	8241	0	4.61

Column	Negative Rows	NA Rows	Neg Percentage	NA Percentage
DELTA_RD_TPU	86934	6450	50.45	3.61
DELTA_RD_FTPU	86934	6516	50.47	3.64
DELTA_RD_FCTK	6959	6450	4.04	3.61
DELTA_CST_TPU	99001	10168	58.72	5.69
DELTA_CST_FTPU	98935	10234	58.7	5.72
DELTA_CST_FCTK	19026	10168	11.28	5.69
DELTA_TPU_FTPU	0	6516	0	3.64
DELTA_TPU_FCTK	0	6450	0	3.61
DELTA_FTPU_FCTK	0	6516	0	3.64
DELTA_QUEUES	0	0	0	0
DELTA_CALLTAKING	0	6450	0	3.61
STEP_1	99001	10168	58.72	5.69
STEP_2	0	6450	0	3.61
STEP_2A	2	75013	0	41.96
STEP_2B	105565	73125	99.91	40.9
STEP_3	582	6450	0.34	3.61
CALL_TIME	586	6450	0.34	3.61
STEP_4	2369	62044	2.03	34.7
STEP_5	67	70821	0.06	39.61
STEP_6	5578	70823	5.17	39.61
STEP_7	2	73291	0	41
STEP_8	9591	76129	9.34	42.58
STEP_9	11	67204	0.01	37.59
STEP_10	6	73292	0.01	41
TOTAL_TIME_A	3	95	0	0.05

Column	Negative Rows	NA Rows	Neg Percentage	NA Percentage
TOTAL_TIME_B	113	8334	0.07	4.66

Quick notes: We should not use TimeCallViewed in the process. There are too many missing values