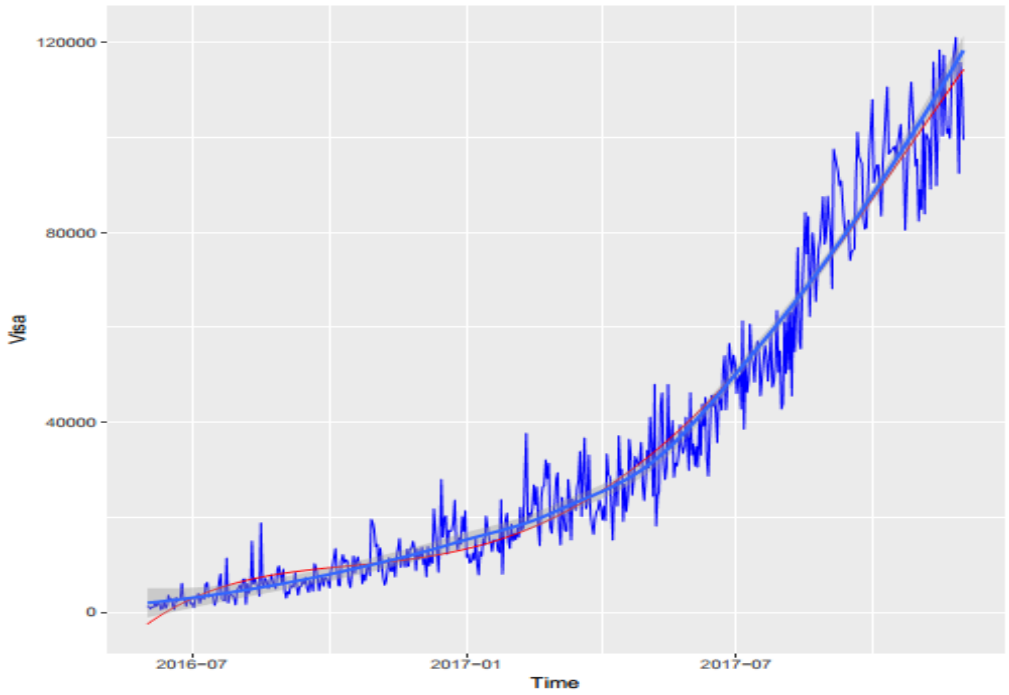
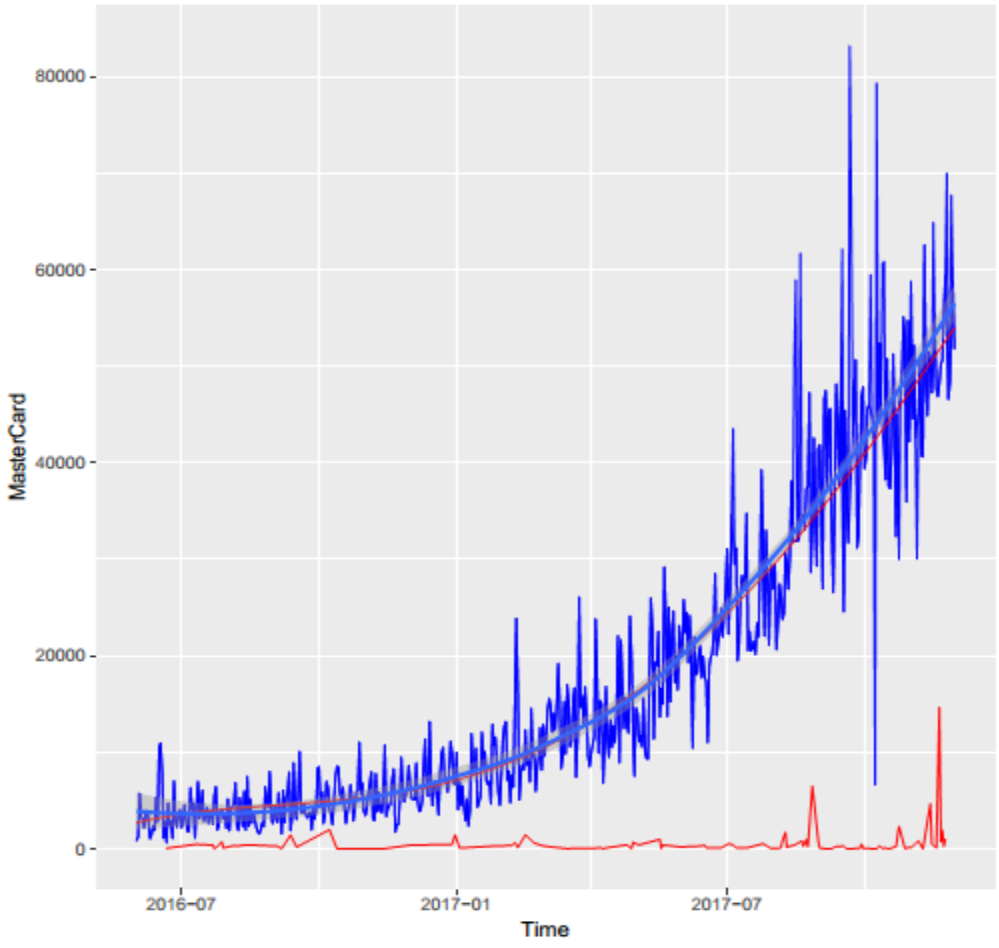
Report #1

1. *The obtained data was processed and split by payment systems.*

A) merchant #1



B) merchant #2



2. *Both series are non-stationary, as shown by formal Dickey-Fuller tests:*

merchant #1

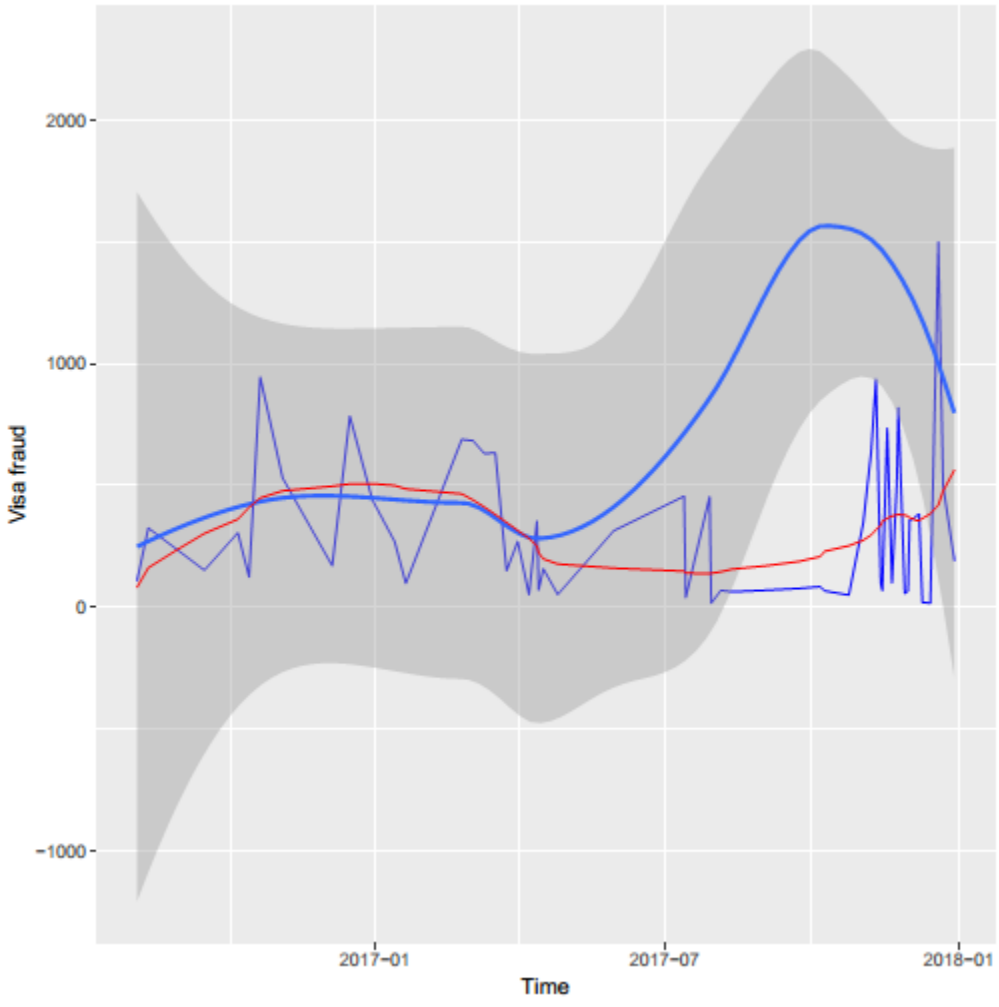
```
> adf.test(o.visa.ts)

        Augmented Dickey-Fuller Test

data:  o.visa.ts
Dickey-Fuller = -0.97282, Lag order = 8, p-value = 0.9431
alternative hypothesis: stationary
```

merchant #2

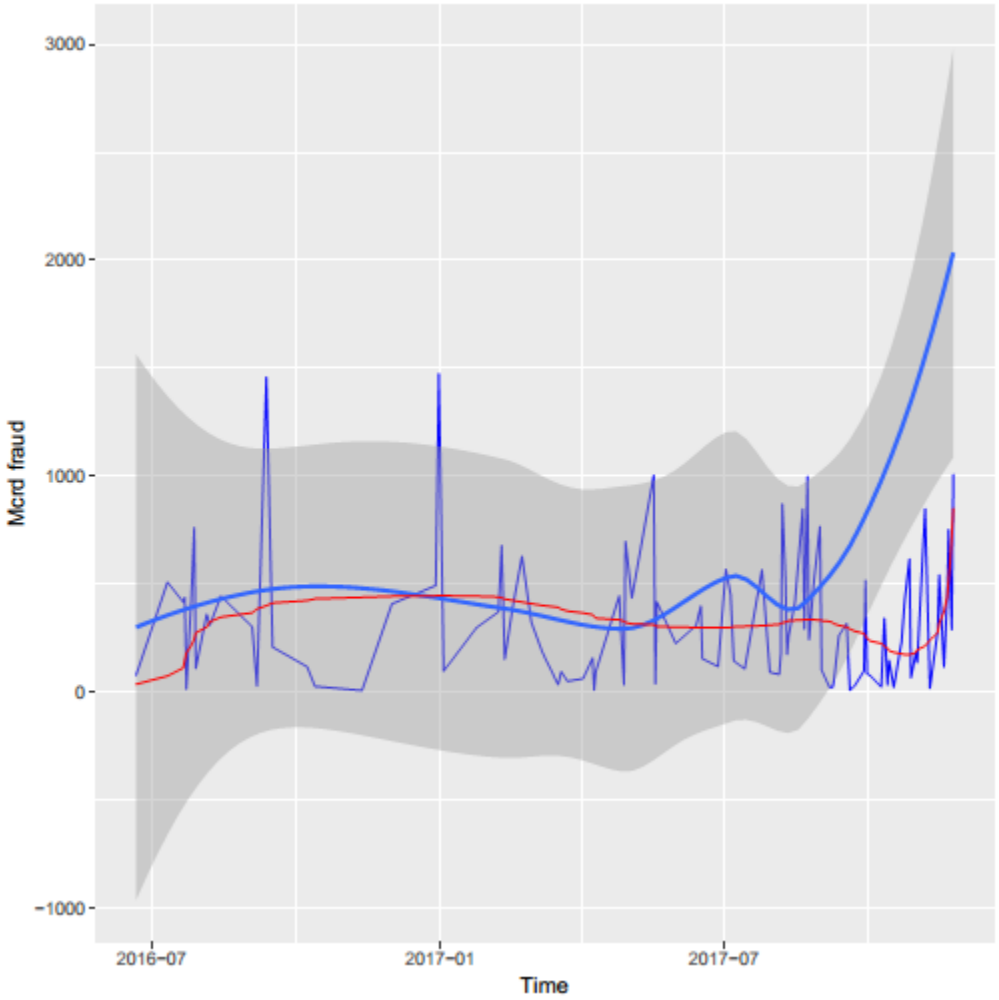
```
> adf.test(o.mcrd.ts)

        Augmented Dickey-Fuller Test

data:  o.mcrd.ts
Dickey-Fuller = -1.5554, Lag order = 8, p-value = 0.7665
alternative hypothesis: stationary
```

For frauds.

merchant #1)



merchant #2)



In both graphs, the bright blue curve is offered by the built-in analyzer, red curve is obtained in the process of current analysis manually.

Very important:

Despite the fact that the Dickey-Fuller test showed high estimate for non-stationarity of time series, both of the standard procedures for language R, decompose () and stl(), failed to decompose none of series to components, namely -
Trend + Seasonality + Random noise. It can tell about the absence of seasonal component.

```
> decompose(o.visa.ts)
Error in decompose(o.visa.ts) :
  the time series has no or less than two periods
```

```
> stl(o.visa.ts,s.window = "periodic")
Error in stl(o.visa.ts, s.window = "periodik") :
  the series is not periodic or has less than two periods
```

Similar result was obtained for data merchant #2*. It was decided to obtain manual trend and forecast based on it.*

merchant #1)

```
> summary(visa.fraud.fit)

Call:
lm(formula = o.visa.fraud.ts1~v+I(v^3)+I(v^4)+I(v^5)+I(v^6)-1)

Residuals:
    Min      1Q  Median      3Q     Max
-388.09 -199.33  -50.06  155.90 1079.89

Coefficients:
         Estimate Std. Error t value Pr(>|t|)
v       8.306e+01  1.666e+01   4.987 7.49e-06 ***
I(v^3) -5.636e-01  1.711e-01  -3.293  0.00180 **
I(v^4)  2.903e-02  1.008e-02   2.879  0.00581 **
I(v^5) -5.442e-04  2.121e-04  -2.566  0.01326 *
I(v^6)  3.537e-06  1.521e-06   2.325  0.02408 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 291.3 on 51 degrees of freedom
Multiple R-squared:  0.6047,   Adjusted R-squared:  0.5659
F-statistic:  15.6 on 5 and 51 DF,  p-value: 2.682e-09
```

The estimated formula explains almost 57% of the variance, which should be considered a real estimate. All coefficients are statistically significant.

merchant #2)

```
> summary(mcrd.fraud.fit)

Call:
lm(formula = o.mcrd.fraud.ts1 ~ v + I(v^2) + I(v^3) + I(v^4) - 1)

Residuals:
    Min      1Q  Median      3Q     Max
-382.90 -207.45  -57.29  166.22 1135.98

Coefficients:
         Estimate Std. Error t value Pr(>|t|)
v       3.442e+01  1.057e+01   3.257  0.00155 **
I(v^2) -9.164e-01  5.647e-01  -1.623  0.10784
I(v^3)  7.961e-03  9.405e-03   0.846  0.39939
I(v^4) -1.740e-05  4.904e-05  -0.355  0.72348
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 311.9 on 97 degrees of freedom
Multiple R-squared:  0.5306,   Adjusted R-squared:  0.5113
F-statistic: 27.41 on 4 and 97 DF,  p-value: 3.13e-15
```

The estimated formula explains almost 51% of the variance, which should also be considered a real estimate.

The time series was detrended – trend was removed from consideration. The series became stationary.

merchant #1)

```
> adf.test(visa.fraud.fit$residuals,alternative = "stationary")

3.   Augmented Dickey-Fuller Test
```

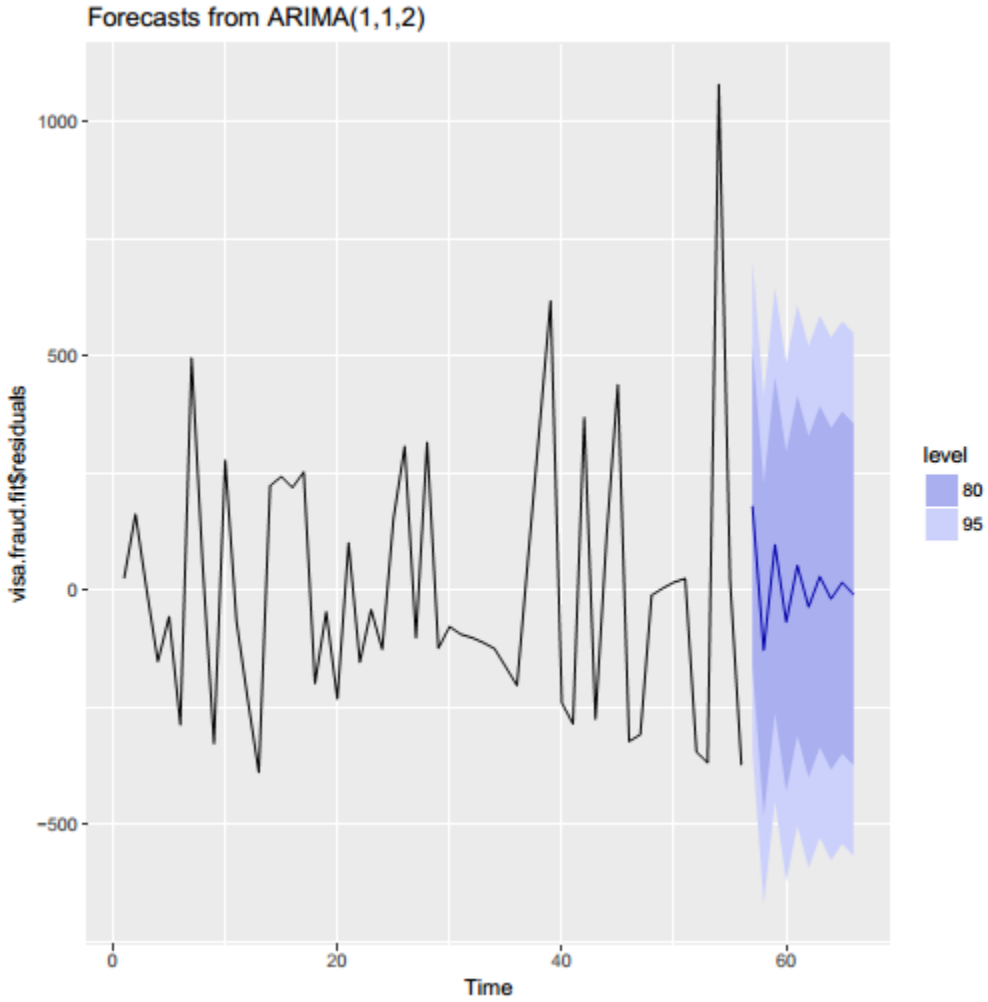
```
4.
5. data:  visa.fraud.fit$residuals
6. Dickey-Fuller = -5.0359, Lag order = 3, p-value = 0.01
7. alternative hypothesis: stationary
```

merchant #2)

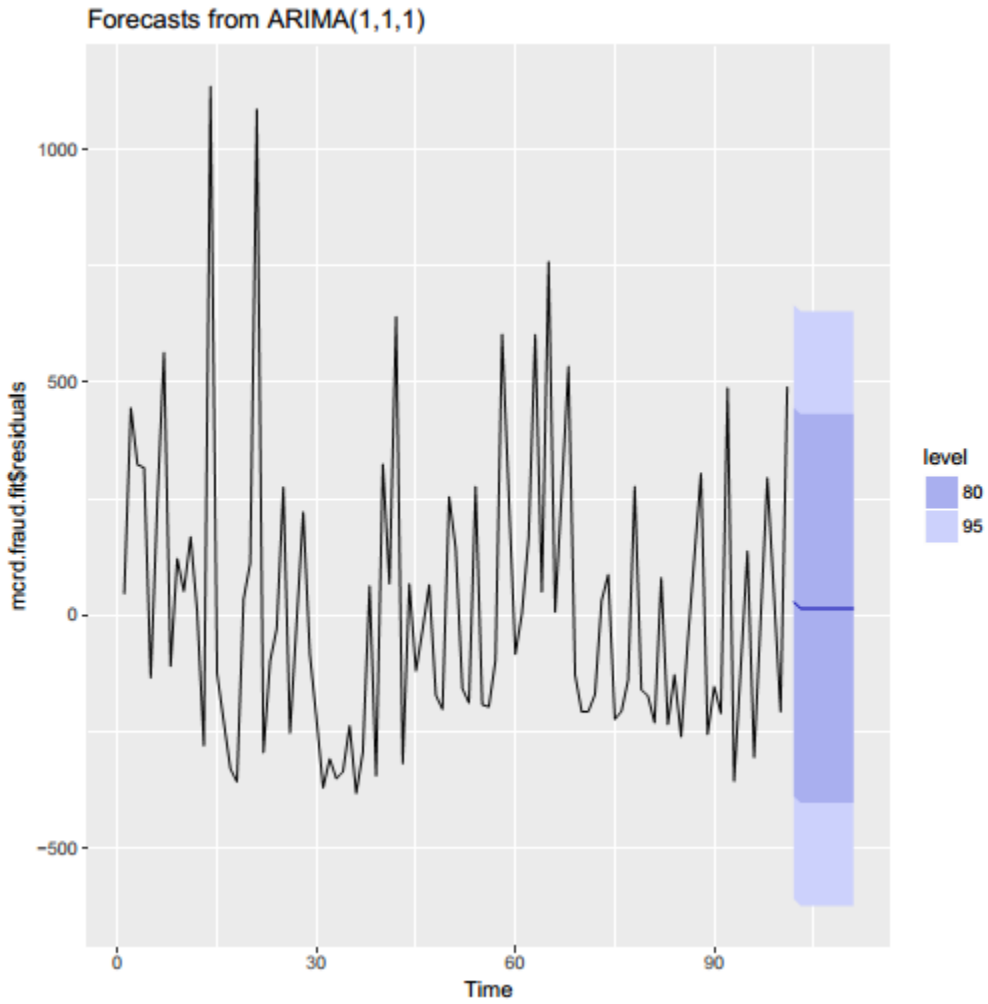
```
> adf.test(mcrd.fraud.fit$residuals,alternative = "stationary")

        Augmented Dickey-Fuller Test

data:  mcrd.fraud.fit$residuals
Dickey-Fuller = -3.3772, Lag order = 4, p-value = 0.06214
alternative hypothesis: stationary
```

3. Preliminary forecasting

merchant #1)



merchant #2)



Following diagrams show the forecast results for the stationary part of the frauds. In general, the model may have (at first estimate) a form:

Forecast = Trend + Stat.