

# A Network Performance Index in Ghana

By  
Nii Narku Quaynor

# A Network Performance Index

- The report is an experience with the Ghana Open Data Initiative (GODI) portal while analyzing selected data
- The author analyzed data published by the NCA on congestion and call drop in various mobile networks
- In particular what relationship, if any exists in the networks between call drops and congestion in the infrastructure
- A national index combining the performance of all the networks is defined in process for call drops and congestion

# Background

- September 2011 - Ghana signed the open government partnership
- NITA released its first data set on [data.gov.gh](http://data.gov.gh) portal in 2012
- Over 800 data sets have since been published at the portal; includes data from NCA

# The NCA Data

- The source data was in excel spreadsheet format
- The data covered MTN, Vodafone, Tigo, Airtel and Expresso networks
- The information content was in text format and was on % of call drops and congestion
- Small data set of about 104 records (2010-2012)
- Requirement is not to exceed 1% congestion, 3% call drops

data.gov.gh

REPUBLIC OF GHANA

Ghana Open Data Initiative Open Data Portal beta

Search the site Search

DATASETS APPS METRICS COMMUNITIES GH DATA SITES WW DATA SITES VISUALIZATION

Improved access to health facilities and service delivery as health facility geo-data is "opened"

1069 Datasets  
17 Departments participating

7 Apps  
2 Mobile Apps  
6 Departments participating

Regional/National Out of School for Primary Schools: 2012-2013

Visualization of the day: Click here to enlarge

# The Data



B110

	A	B	C	D	E	F	G	H
1	<b>Trends in Quality of Services - Call Drop from 2010-2012</b>							
2								
3	<b>Month</b>	<b>City/Town</b>	<b>MTN</b>	<b>Vodafone</b>	<b>Tigo</b>	<b>Airtel</b>	<b>Expresso</b>	<b>Allowable Threshold</b>
4	Feb-10	Cape Coast	2%	0%	10%	32%	0%	3%
5	Feb-10	Elmina	2%	0%	0%	63%	0%	3%
6	Feb-10	Winneba	4%	0%	20%	53%	0%	3%
7	Feb-10	Sekondi Takoradi	0%	1%	0%	0%	0%	3%
8	May-10	Kumasi	4%	5%	11%	10%	1%	3%
9	May-10	Obuasi	11%	0%	2%	16%	3%	3%
10	May-10	Mampong	10%	0%	0%	0%	0%	3%

A1

	A	B	C	D	E	F	G	H
1	<b>Trends in Quality of Services - Congestion from 2010-2012</b>							
2								
3	<b>Month</b>	<b>City/Town</b>	<b>MTN</b>	<b>Vodafone</b>	<b>Tigo</b>	<b>Airtel</b>	<b>Expresso</b>	<b>Allowable Threshold</b>
4	Feb-10	Cape Coast	9%	1%	13%	1%	20%	1%
5	Feb-10	Elmina	30%	0%	2%	0%	6%	1%
6	Feb-10	Winneba	0%	0%	18%	3%	21%	1%
7	Feb-10	Sekondi Takoradi	3%	0%	2%	0%	18%	1%
8	May-10	Kumasi	5%	7%	10%	2%	7%	1%
9	May-10	Obuasi	2%	2%	0%	3%	3%	1%
10	May-10	Mampong	7%	0%	0%	0%	21%	1%
11	May-10	Konongo	0%	7%	6%	0%	0%	1%
12	Jun-10	Accra	3%	17%	27%	12%	10%	1%

# The Analysis Environment and process

- Author is a student of Machine Learning and uses the R programming environment with a console on Mac
- The data from the portal is saved in csv format and loaded in R; then cleaned; the text entries are converted to dates and numbers
- One provider, Espresso, had  $> 10\%$  missing data and was excluded in the analysis

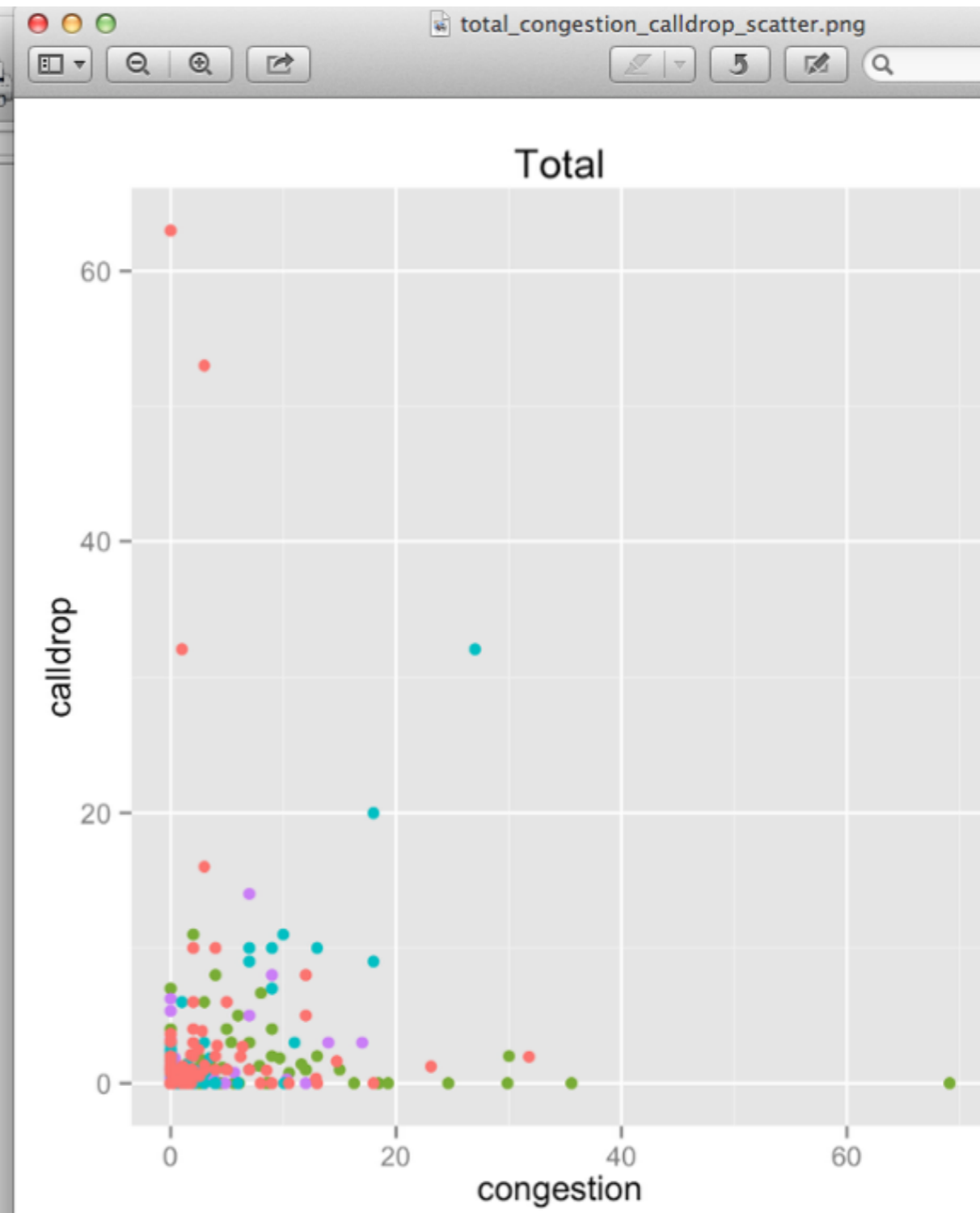
# The R Environment

```

57 #for Airtel
58 congestion$Airtel<-gsub("%", "", congestion$Airtel)
59 congestion$Airtel<-as.numeric(congestion$Airtel)
60
61 #for Expresso
62 #congestion$Expresso<-gsub("%", "", congestion$Expresso)
63 #congestion$Expresso<-as.numeric(congestion$Expresso)
64
65
66 #look at congestion vs call drops
67 #MTN
68 calldrop.congestion.mtn<-data.frame(calldrop.congestion.mtn)
69 names(calldrop.congestion.mtn)<-c("calldrop", "congestion")
70 scatter.mtn<-ggplot(calldrop.congestion.mtn, aes(calldrop, congestion))
71 ggsave(plot=scatter.mtn, filename="/Users/quentin/Desktop/total_congestion_calldrop_scatter.png")
72
73 #Vodafone
74 calldrop.congestion.vodafone<-data.frame(calldrop.congestion.vodafone)
75 names(calldrop.congestion.vodafone)<-c("calldrop", "congestion")
76 scatter.vodafone<-ggplot(calldrop.congestion.vodafone, aes(calldrop, congestion))
77 ggsave(plot=scatter.vodafone, filename="/Users/quentin/Desktop/vodafone_congestion_calldrop_scatter.png")
78
79 #Tigo
80 calldrop.congestion.tigo<-data.frame(calldrop.congestion.tigo)
81 names(calldrop.congestion.tigo)<-c("calldrop", "congestion")
82 scatter.tigo<-ggplot(calldrop.congestion.tigo, aes(calldrop, congestion))
83 ggsave(plot=scatter.tigo, filename="/Users/quentin/Desktop/tigo_congestion_calldrop_scatter.png")
84
85 #Airtel
86 calldrop.congestion.airtel<-data.frame(calldrop.congestion.airtel)
87 names(calldrop.congestion.airtel)<-c("calldrop", "congestion")
88 scatter.airtel<-ggplot(calldrop.congestion.airtel, aes(calldrop, congestion))
89 ggsave(plot=scatter.airtel, filename="/Users/quentin/Desktop/airtel_congestion_calldrop_scatter.png")
90
91 #total scatter
92 congestion.total<-congestion$MTN
93 congestion.operator<-array("MTN", length(congestion.total))
94 congestion.total<-c(congestion.total, congestion.operator)
95 congestion.temp<-array("Vodafone", length(congestion.total))
96 congestion.operator<-c(congestion.operator, congestion.temp)
97 congestion.total<-c(congestion.total, congestion.operator)
98 congestion.temp<-array("Tigo", length(congestion.total))
    
```

```

> summary(calldrop.op$MTN)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000  0.000   0.000   1.015  1.182  11.000
>
> summary(calldrop.op$Vodafone)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000  0.000   0.000   1.227  0.455  17.000
>
> summary(congestion$MTN)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000  0.000   1.495   4.899  5.465  69.100
>
> summary(congestion$Vodafone)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000  0.000   0.000   1.227  0.455  17.000
>
> summary(congestion$Tigo)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000  0.000   0.000   1.969  1.505  27.000
>
> summary(congestion$Airtel)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000  0.000   0.065   2.752  3.000  31.810
    
```



# Analysis Approach

- Study congestion and call drop separately using suitable techniques for analyzing a single variable: statistics, time series, histograms, density plots
- Employ Principal Component Analysis (PCA) to derive an index about the mean for a combined index of four operators for congestion and call drops
- Use techniques suitable to find relationship among multiple variables to study congestion and call drops: scatterplots, curve fitting for each operator and for combined set of operators



# Congestion Statistics

```
> summary(congestion$MTN)
  Min. 1st Qu. Median      Mean 3rd Qu.      Max.
  0.000  0.000  1.495  4.899  5.465 69.100

> summary(congestion$Vodafone)
  Min. 1st Qu. Median      Mean 3rd Qu.      Max.
  0.000  0.000  0.000  1.227  0.455 17.000

> summary(congestion$Tigo)
  Min. 1st Qu. Median      Mean 3rd Qu.      Max.      NA's
  0.000  0.000  0.000  1.969  1.505 27.000         1

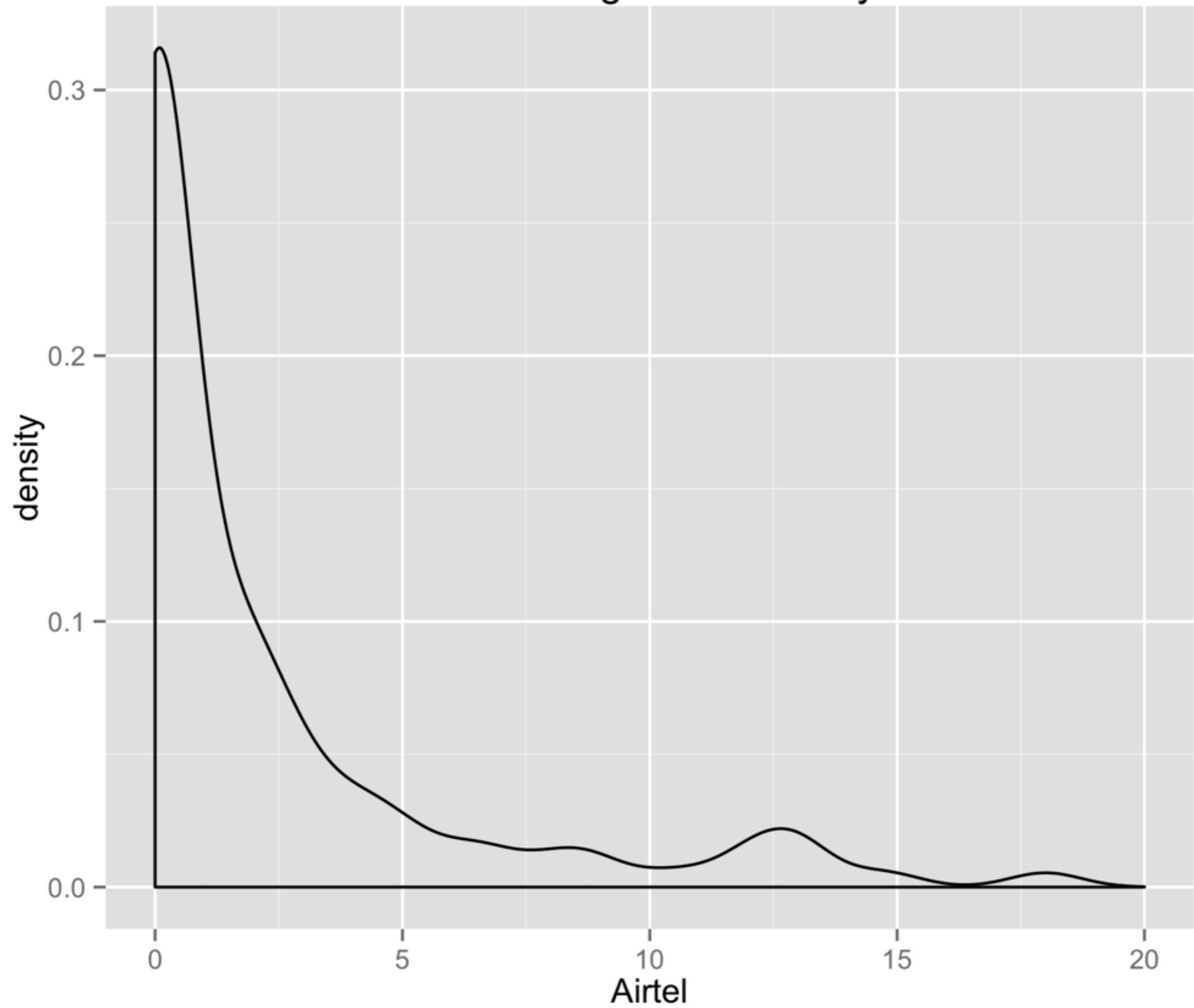
> summary(congestion$Airtel)
  Min. 1st Qu. Median      Mean 3rd Qu.      Max.
  0.000  0.000  0.065  2.752  3.000 31.810
```

Fail

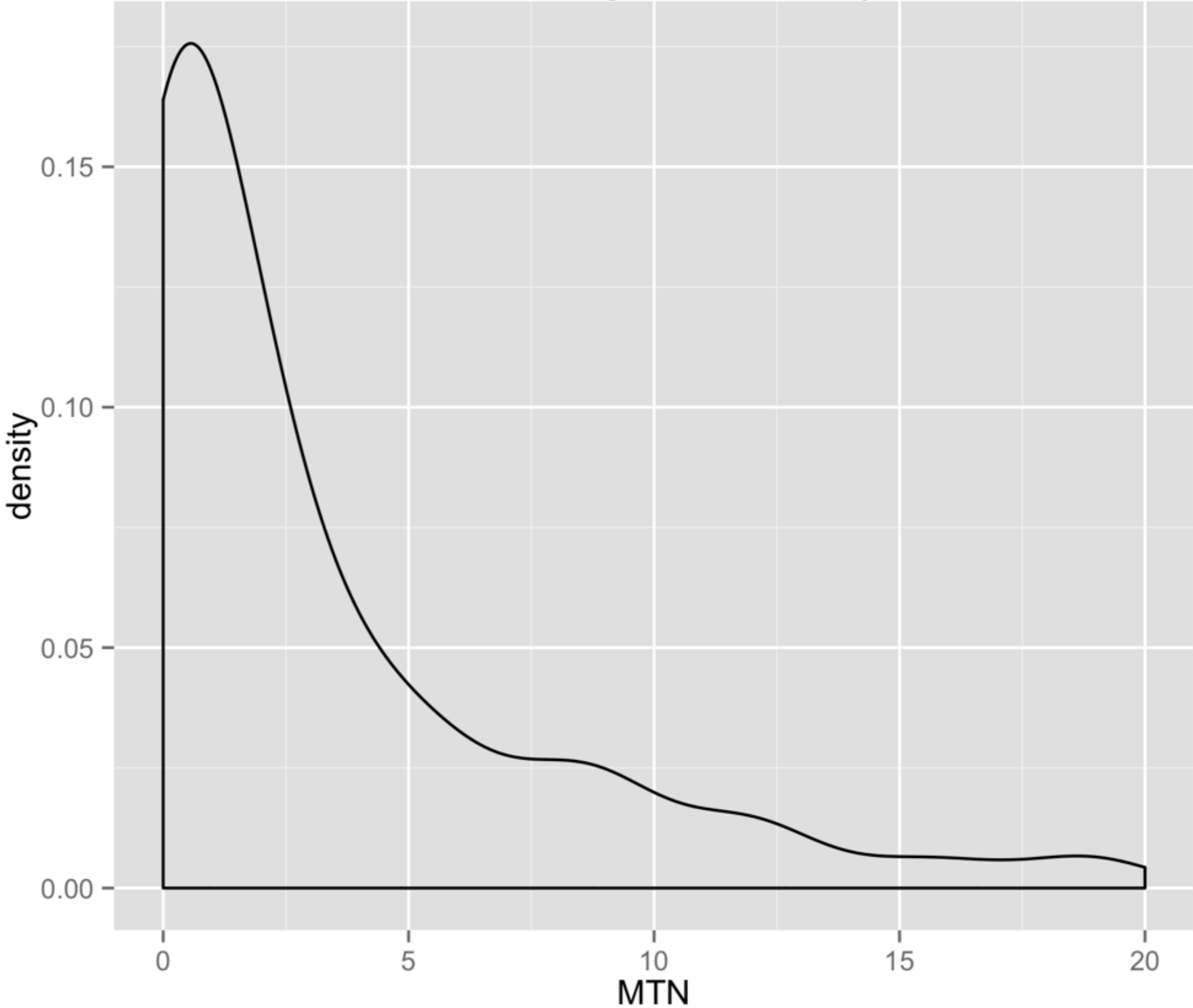
Pass

# Congestion Density plots

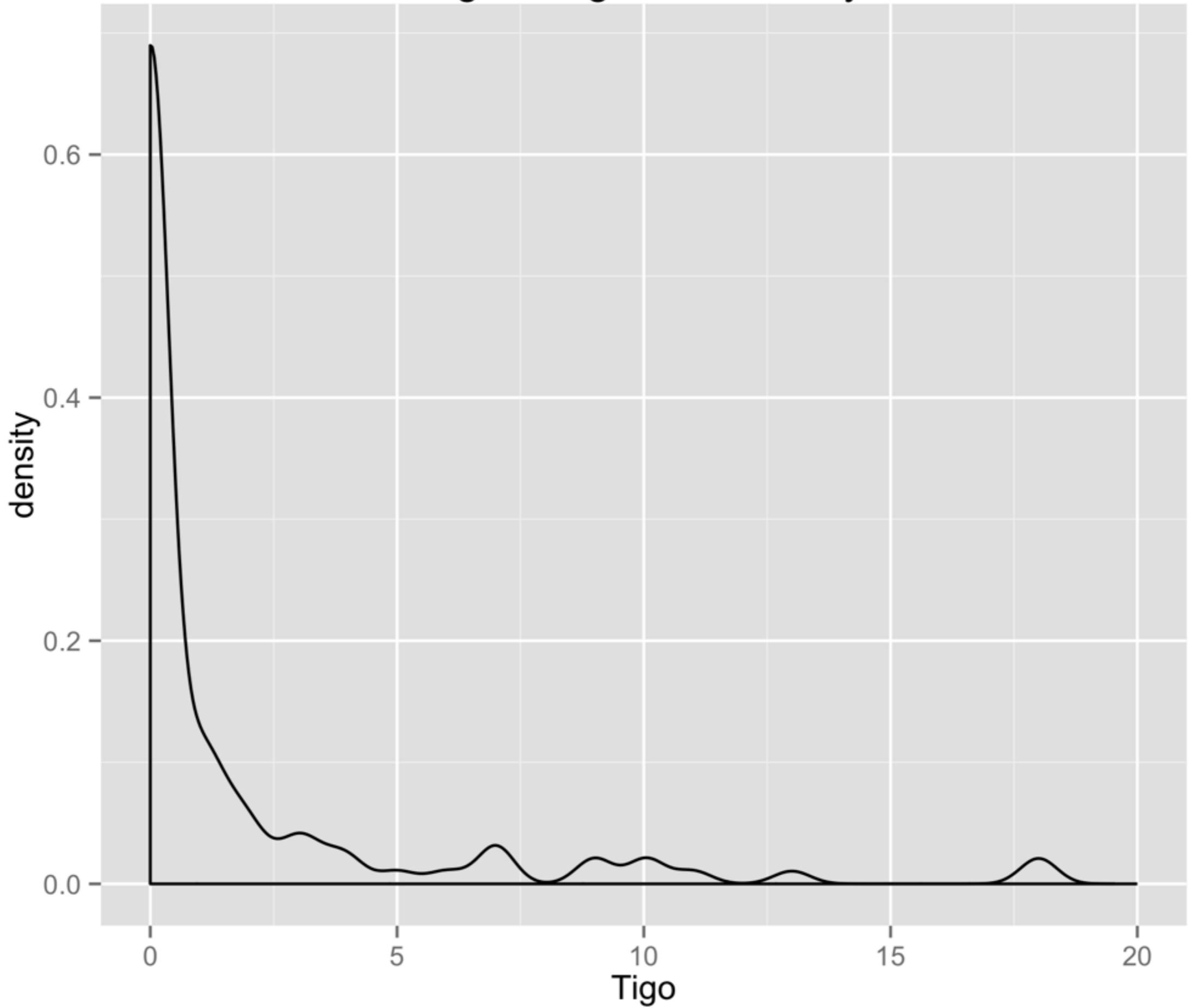
# Airtel Congestion Density



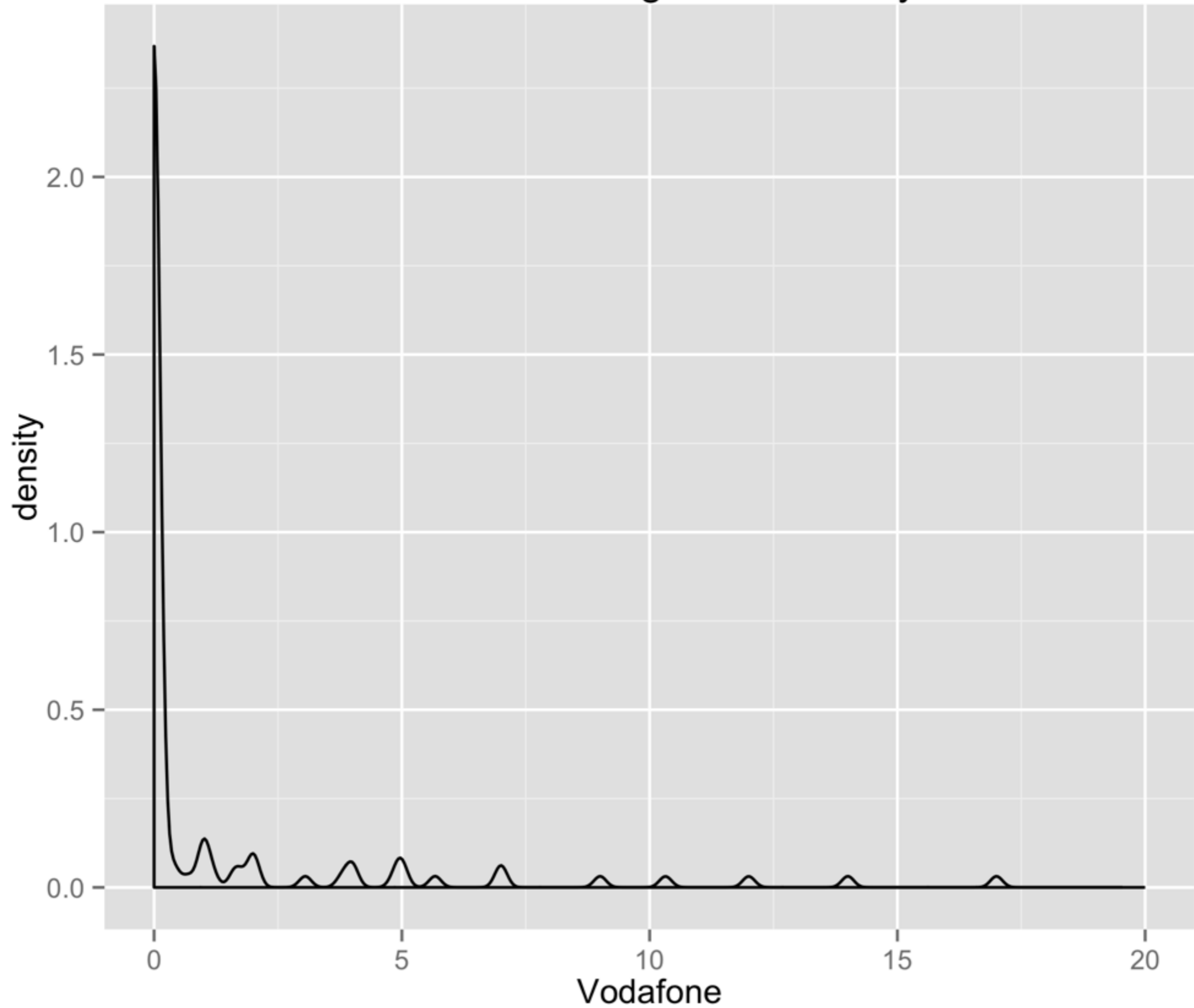
# MTN Congestion Density



# Tigo Congestion Density



# Vodafone Congestion Density



# Congestion Density plots

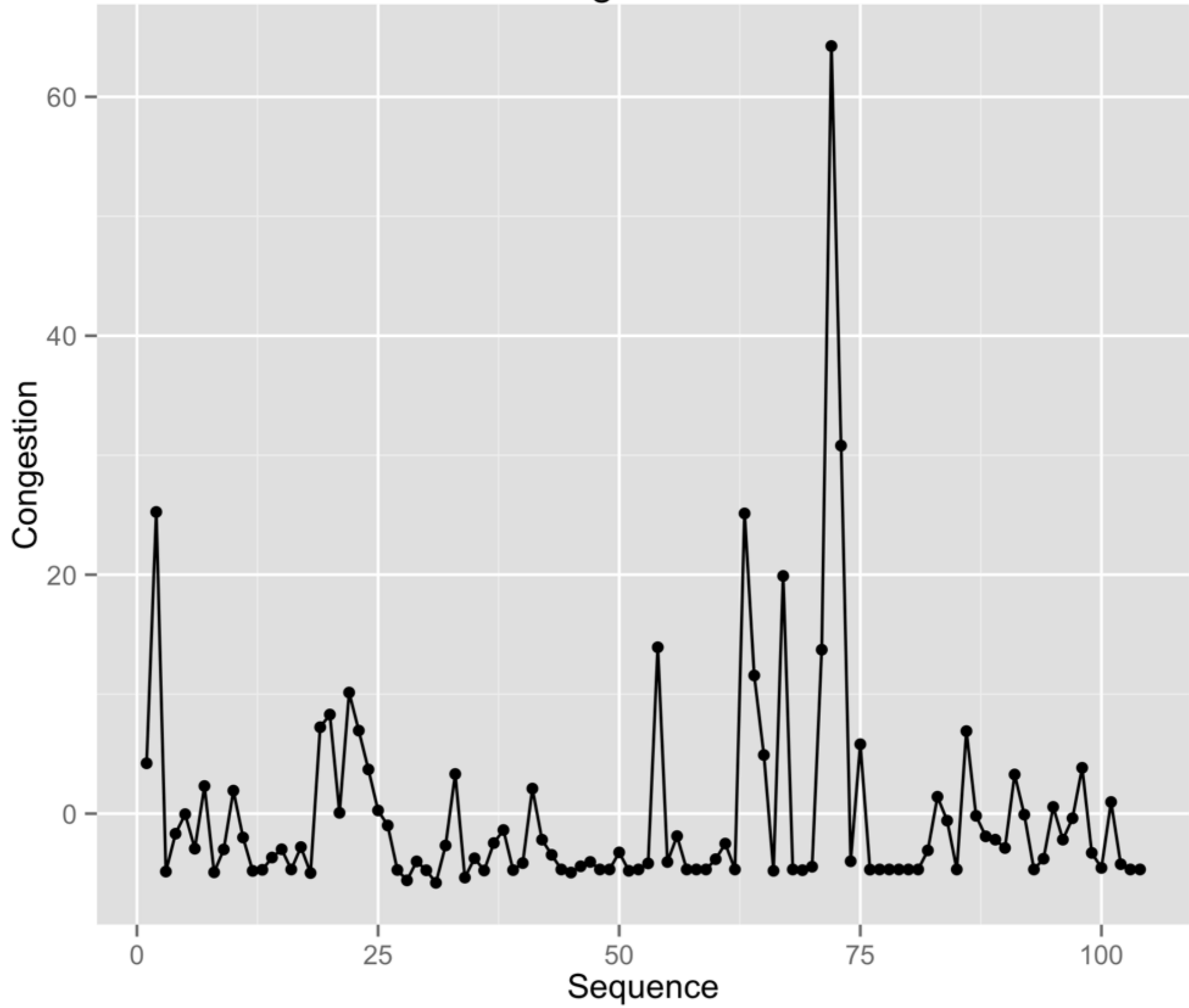
- Vodafone and Tigo are skewed to 0-1%
- MTN has a tail at 20%
- Airtel is tails at  $< 14\%$

# Congestion Index

- Use PCA to combine the four variables to create a national index on congestion about the mean of the data
- In practice one studies the density of the principal component loadings looking to see all positive or all negative

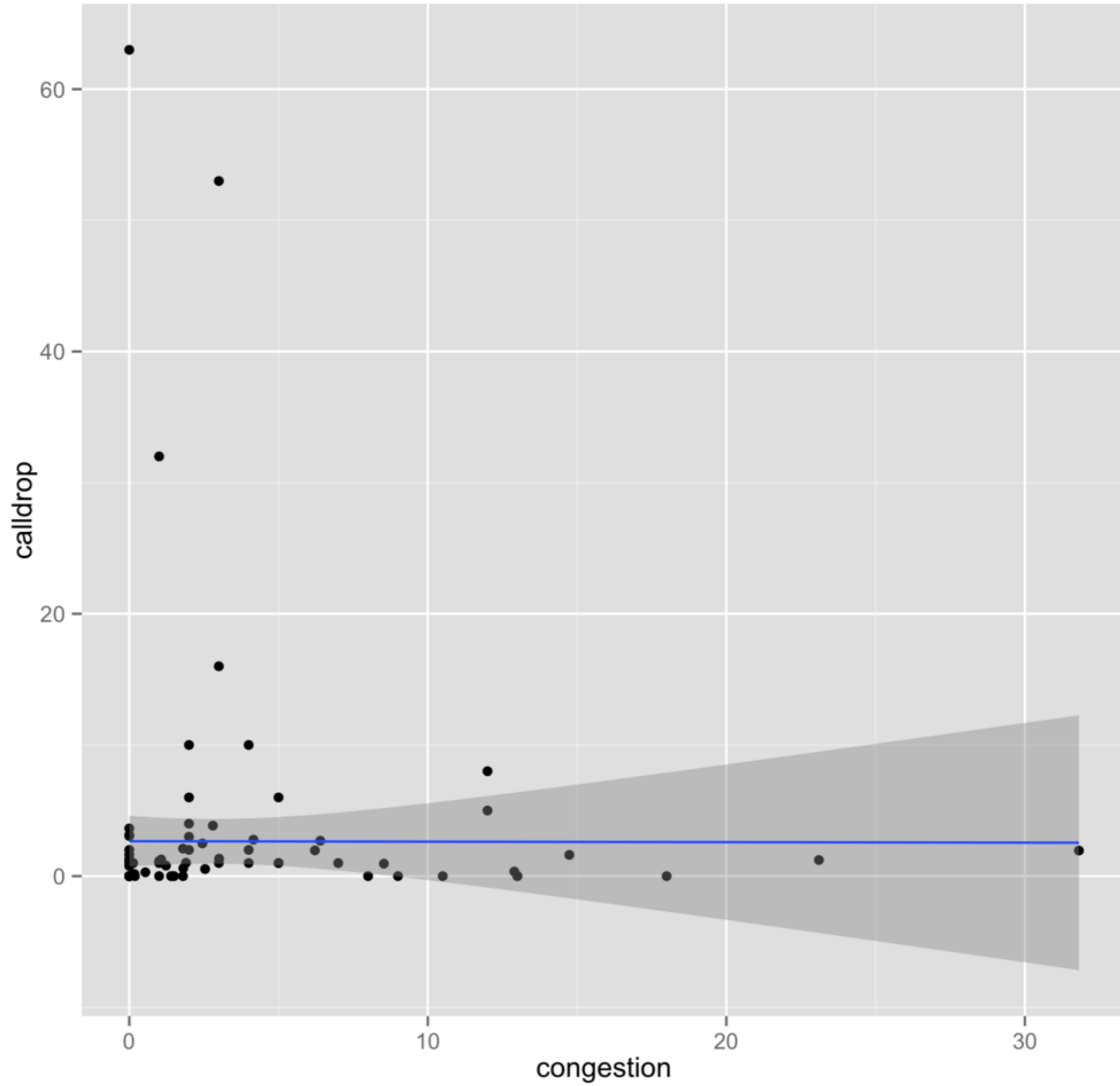


# NII Congestion Index

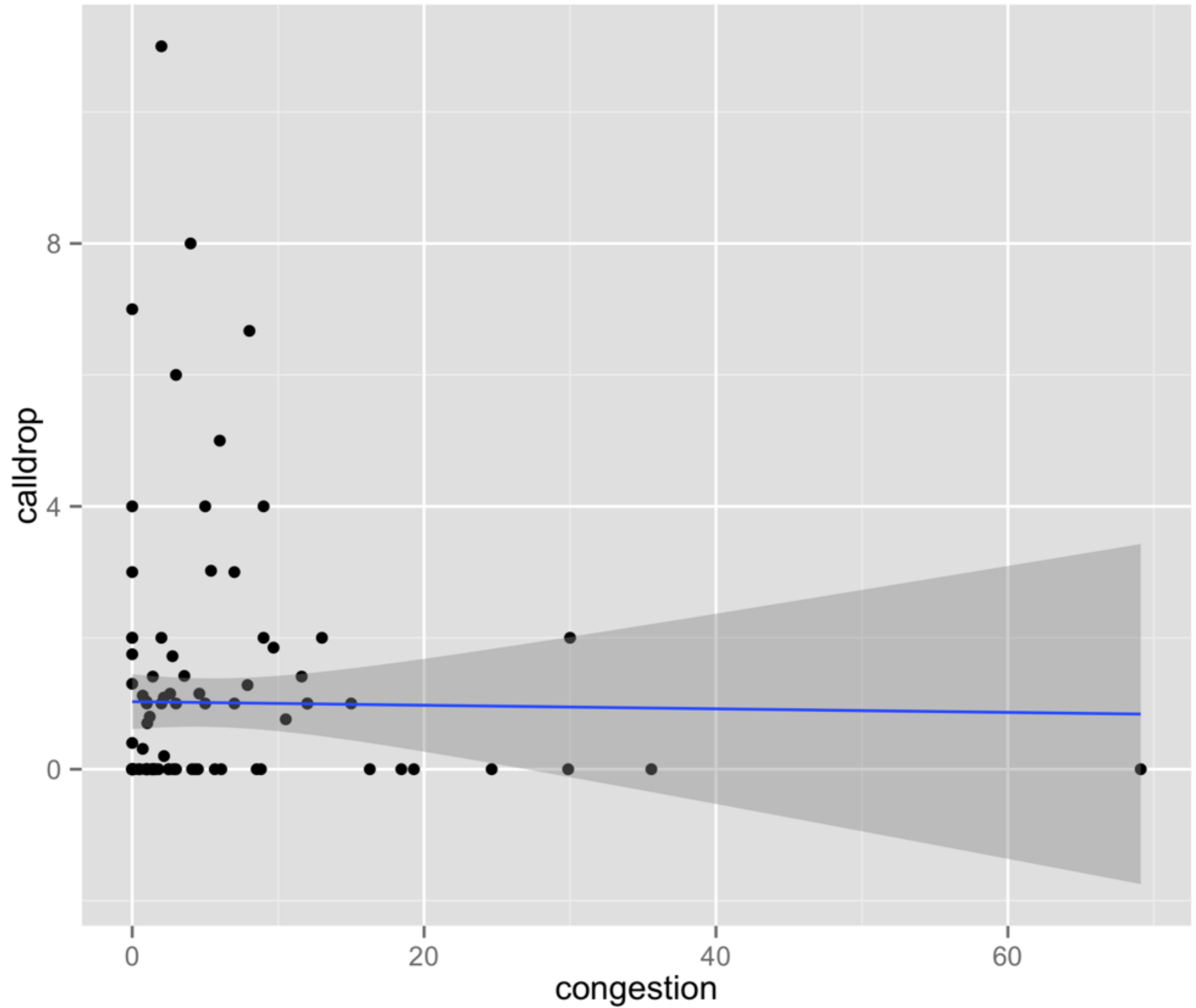


# Congestion and Call drops

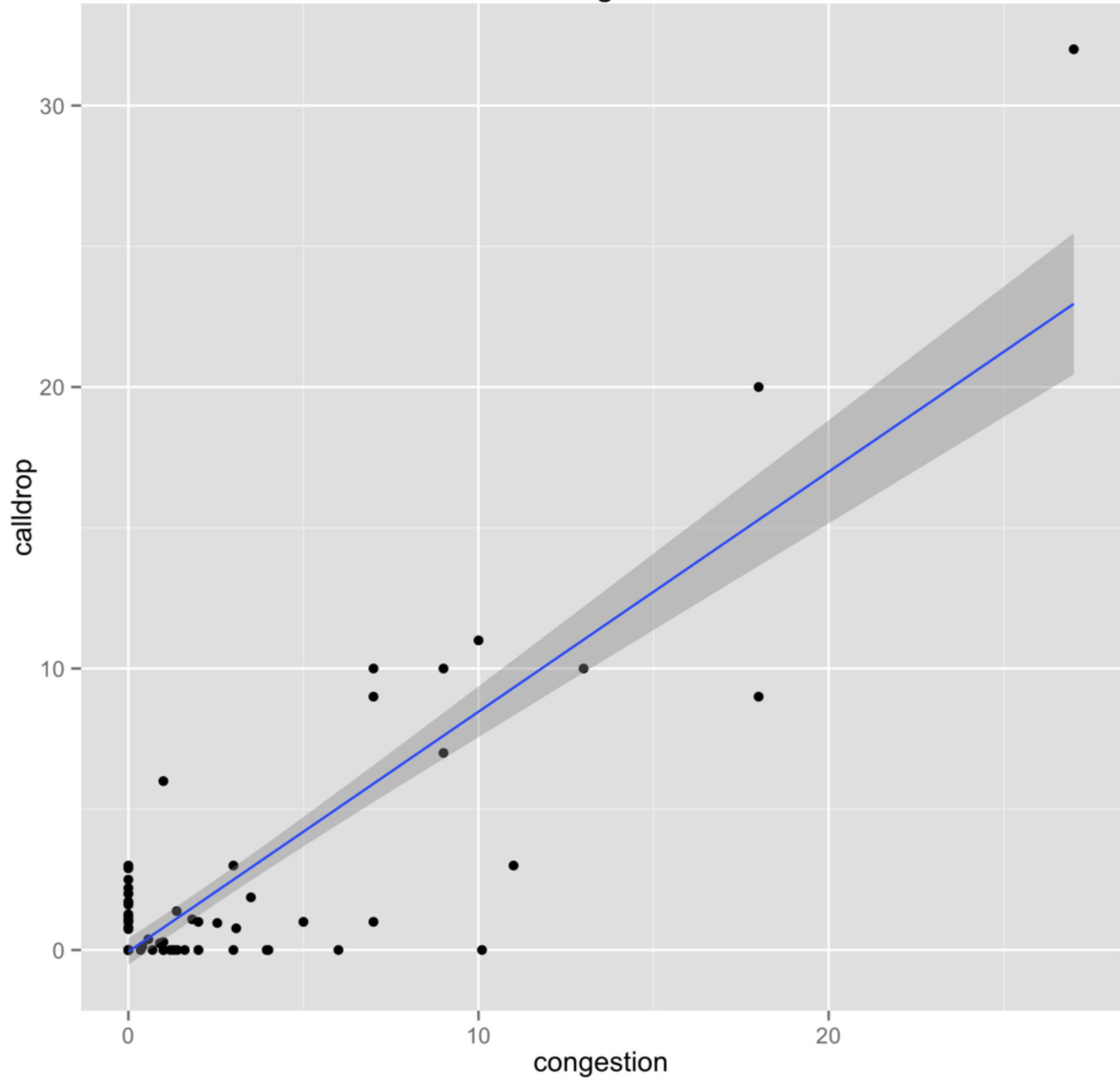
# Airtel



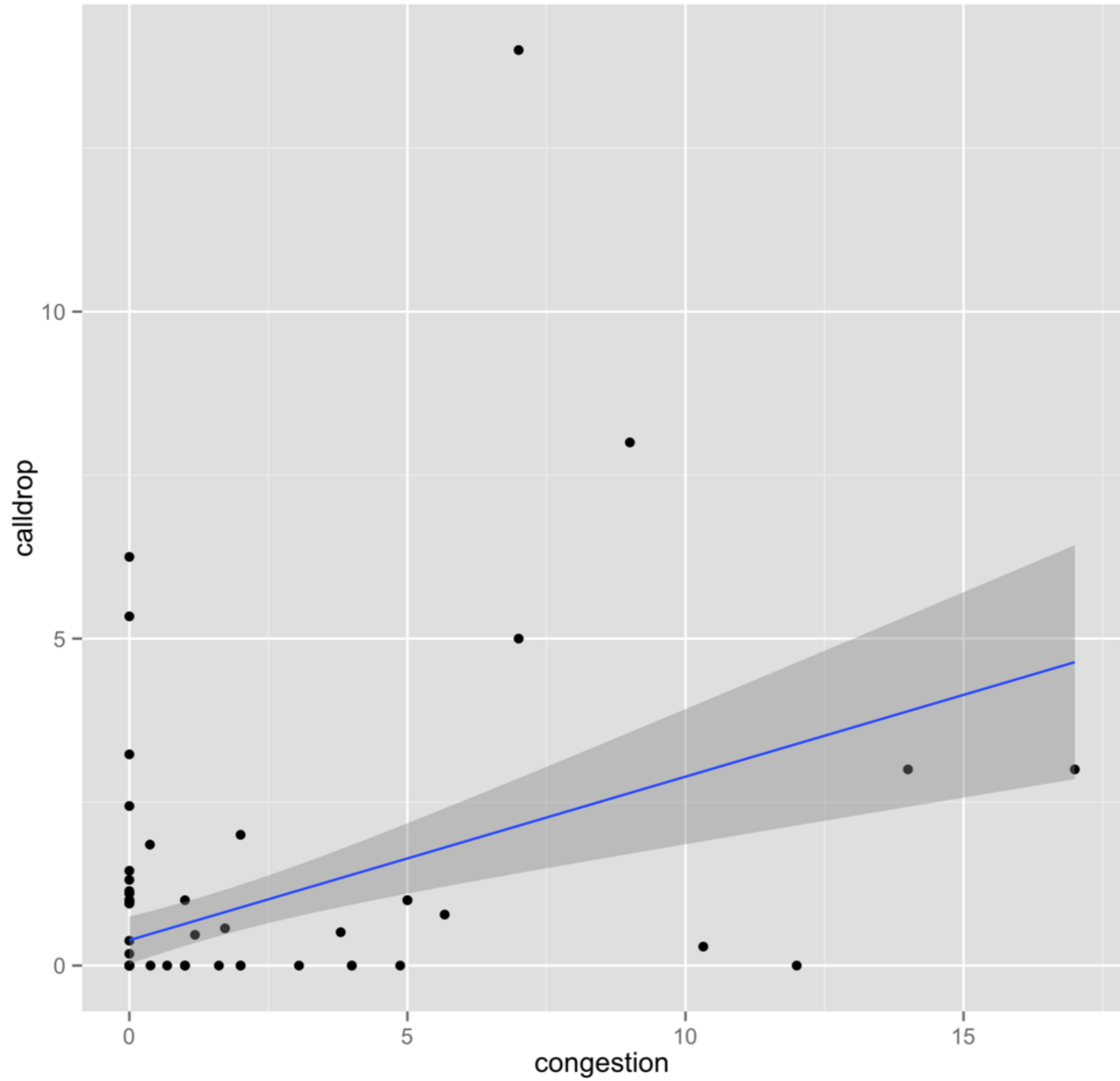
# MTN



# Tigo



# Vodafone



- Tigo and Vodafone show a relationship between call drops and congestion
- MTN and Airtel on the other hand show no meaningful relationship

# Conclusion

- NCA to release larger volume of data for analysis and to be more complete
- NCA to define more precisely the performance criteria for limits
- The 1% congestion limit proved difficult for providers while 3% call drops was easily met by all providers



# Conclusion(2)

- More current data would be enhance the study
- Make regulations on the definition and calculation of limits to better govern the mobile telecommunication sector
- GODI as a platform for evidence based analysis and reporting has been successful
- An independent scientist, without permission, has analyzed the performance of the national network