# Progress update 1/10/2015 - Fauzy Bin Che Yayah

## Data Exploration

* Acquiring dataset by zone

Dataset structure

a <- read.csv("table\_struct.csv")  
names(a)

## [1] "tt\_row\_id" "tt\_num"   
## [3] "tt\_type" "tt\_sub\_type"   
## [5] "status" "severity"   
## [7] "important\_message" "appointment\_flag"   
## [9] "nova\_account\_name" "nova\_subscriber\_num"   
## [11] "nova\_account\_num" "package\_row\_id"   
## [13] "created\_by" "category"   
## [15] "symptom\_error\_code" "priority"   
## [17] "product" "sub\_product"   
## [19] "package\_name" "network\_tt\_id"   
## [21] "swap\_order\_num" "cause\_category"   
## [23] "cause\_code" "resolution\_code"   
## [25] "closure\_category" "resolution\_team"   
## [27] "service\_affected" "service\_order\_num"   
## [29] "btu\_type" "owner"   
## [31] "owner\_name" "group\_owner"   
## [33] "owner\_position" "btu\_platform"   
## [35] "dp\_location" "created\_date"   
## [37] "pending\_verify\_date" "closed\_by"   
## [39] "closed\_date" "source"   
## [41] "installed\_date" "description"   
## [43] "repeat\_ticket\_count" "follow\_up\_ticket\_count"  
## [45] "fdp\_device\_name" "fdp\_site\_name"   
## [47] "olt\_site\_name" "exchange"   
## [49] "timestamp" "contact\_id"   
## [51] "contact\_name" "contact\_office\_phone"   
## [53] "contact\_mobile\_phone" "contact\_home\_phone"   
## [55] "contact\_email\_addr" "due\_date"   
## [57] "part\_num" "network\_layer"   
## [59] "network\_row\_id" "asset\_id"   
## [61] "ptt" "zone"   
## [63] "service\_point\_id"

select \* from nova\_trouble\_ticket limit 1

Total Zone = 53

* ZONE AIR ITAM
* ZONE BANGI
* ZONE BANGSAR
* ZONE BANTING
* ZONE BATU
* ZONE BATU PAHAT
* ZONE BAYAN BARU
* ZONE BINTULU
* ZONE BUKIT ANGGERIK
* ZONE BUKIT MERTAJAM
* ZONE BUKIT RAJA
* ZONE BUTTERWORTH
* ZONE CYBERJAYA
* ZONE GOMBAK
* ZONE IPOH
* ZONE KAJANG
* ZONE KEPONG
* ZONE KERAMAT
* ZONE KINRARA
* ZONE KL CENTRAL
* ZONE KLANG
* ZONE KOTA KINABALU SELATAN
* ZONE KOTA KINABALU UTARA
* ZONE KUCHING
* ZONE KULIM
* ZONE LANGKAWI
* ZONE MALURI
* ZONE MELAKA UTARA
* ZONE MIRI
* ZONE N. SEMBILAN UTARA
* ZONE PANDAN
* ZONE PELANGI
* ZONE PERLIS
* ZONE PETALING JAYA
* ZONE PUCHONG
* ZONE SEBERANG JAYA
* ZONE SENAI
* ZONE SG PETANI
* ZONE SHAH ALAM
* ZONE SIBU
* ZONE SKUDAI PONTIAN
* ZONE STAMPIN
* ZONE SUBANG JAYA
* ZONE TAMAN PETALING
* ZONE TAMPOI
* ZONE TAR
* ZONE TASEK
* ZONE TASIK AMPANG
* ZONE TDI
* ZONE TELUK INTAN
* ZONE TERENGGANU SELATAN
* ZONE TERUNTUM

## Encoding

* Re-encoding the dataset

## Sampling

* Finding the independent variables and dependent variable.
* Sampling method
* Using the independent variables for prediction