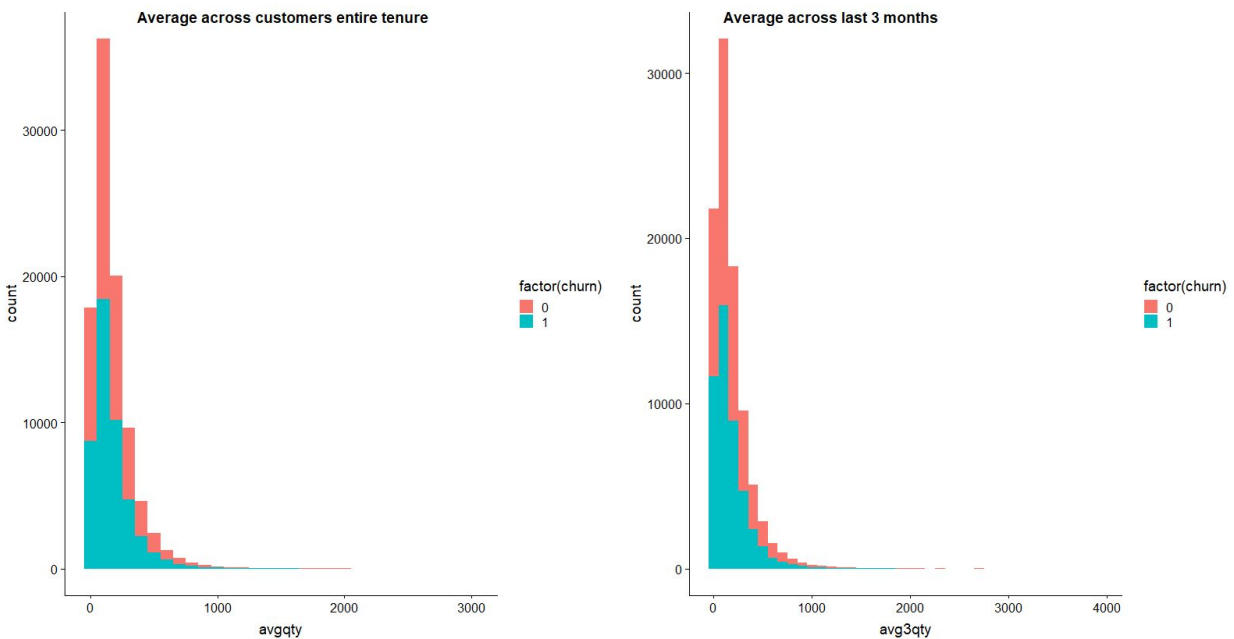


## APPENDIX

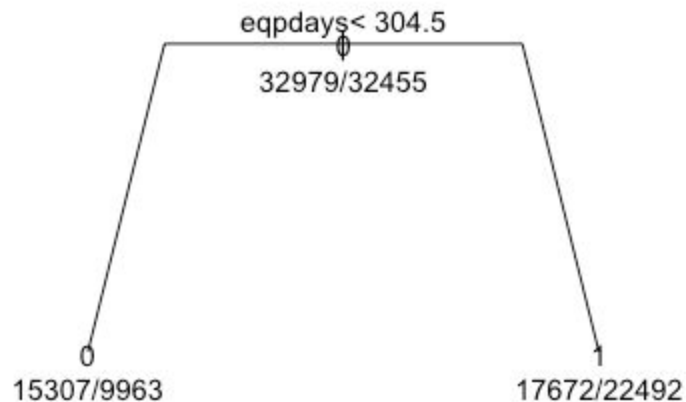
Data URL:

<https://www.kaggle.com/abhinav89/telecom-customer/home>

**Exhibit 2:** Distribution of churn customers over average monthly number of calls in entire customer lifetime and average monthly number of calls in the previous 3 months



**Exhibit 3:** Results of cart model with churn as dependent variable and the rest of variables in clean dataset as dependent variables



#### Exhibit 4: Results of logistic regression model after lasso

```
call:
glm(formula = My ~ ., family = binomial(link = "logit"), data = training_set2
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-3.110	-1.137	-0.585	1.131	2.781

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	3.319e-01	8.771e-02	3.784	0.000154	***
mou_Mean	-7.371e-04	5.504e-05	-13.392	< 2e-16	***
totmrc_Mean	-2.322e-03	4.756e-04	-4.883	1.05e-06	***
da_Mean	1.234e-02	4.437e-03	2.780	0.005431	**
ovrmou_Mean	3.042e-04	3.251e-04	0.936	0.349406	
ovrrev_Mean	5.466e-03	9.972e-04	5.481	4.23e-08	***
roam_Mean	5.327e-03	1.117e-03	4.770	1.85e-06	***
change_mou	-4.390e-04	4.421e-05	-9.931	< 2e-16	***
change_rev	2.401e-03	2.925e-04	8.206	2.28e-16	***
drop_vce_Mean	4.366e-03	1.531e-03	2.851	0.004364	**
drop_dat_Mean	-1.116e-02	1.335e-02	-0.836	0.403260	
recv_sms_Mean	1.044e-02	8.693e-03	1.201	0.229788	
custcare_Mean	-1.489e-03	1.886e-03	-0.789	0.430036	
threeway_Mean	-4.225e-02	9.560e-03	-4.420	9.89e-06	***
mou_cvce_Mean	-5.909e-05	7.951e-05	-0.743	0.457348	
iwylis_vce_Mean	-1.432e-03	7.191e-04	-1.991	0.046469	*
peak_vce_Mean	-6.647e-04	2.694e-04	-2.468	0.013602	*
mou_peav_Mean	-6.712e-05	1.266e-04	-0.530	0.596060	
drop_blk_Mean	4.526e-03	9.122e-04	4.962	6.97e-07	***
callwait_Mean	-1.366e-03	2.408e-03	-0.567	0.570573	
months	-1.685e-02	1.555e-03	-10.836	< 2e-16	***
uniqusubs	1.898e-01	1.602e-02	11.848	< 2e-16	***
actvsubs	-1.741e-01	2.193e-02	-7.941	2.01e-15	***
new_cellu	-1.462e-02	2.450e-02	-0.597	0.550632	
new_cellu	-5.813e-02	2.846e-02	-2.043	0.041073	*
crc1scodA2	3.590e-01	9.144e-02	3.926	8.63e-05	***
crc1scodAA	4.776e-02	2.421e-02	1.973	0.048485	*
crc1scodB	3.994e-02	4.254e-02	0.939	0.347818	
crc1scodBA	6.374e-02	3.000e-02	2.125	0.033622	*
crc1scodC2	-3.430e-01	1.979e-01	-1.733	0.083022	.
crc1scodC5	-2.647e-01	3.121e-01	-0.848	0.396376	
crc1scodD	-1.332e-01	1.670e-01	-0.798	0.425062	
crc1scodD4	-2.023e-01	1.582e-01	-1.279	0.200749	
crc1scodD5	-1.425e+00	3.665e-01	-3.888	0.000101	***
crc1scodDA	-9.030e-02	4.701e-02	-1.921	0.054769	.
crc1scodE	-4.137e-01	1.435e-01	-2.883	0.003944	**
crc1scodE2	-1.718e+00	7.695e-01	-2.233	0.025544	*
crc1scodE4	-5.903e-01	1.094e-01	-5.398	6.73e-08	***
crc1scodEA	-2.782e-01	3.756e-02	-7.409	1.28e-13	***
crc1scodEM	5.858e-01	3.776e-01	1.552	0.120752	
crc1scodG	1.820e-01	1.693e-01	1.075	0.282319	
crc1scodH	-1.767e+00	1.086e+00	-1.626	0.103865	
crc1scodJF	7.278e-01	1.873e-01	3.886	0.000102	***
crc1scodO	-9.473e-01	5.453e-01	-1.737	0.082327	.
crc1scodTP	9.458e+00	4.115e+01	0.230	0.818214	
crc1scodZA	5.381e-02	4.781e-02	1.126	0.260336	
asl_flagv	-2.282e-01	3.218e-02	-7.091	1.33e-12	***
avgmou	7.269e-04	6.957e-05	10.448	< 2e-16	***
avgqty	1.632e-04	1.718e-04	0.950	0.342171	
areaCALIFORNIA.NORTH.AREA	4.826e-02	3.723e-02	1.296	0.194892	
areaCENTRAL.SOUTH.TEXAS.AREA	-9.953e-02	4.378e-02	-2.274	0.022992	*
areaDALLAS.AREA	-1.452e-02	3.866e-02	-0.376	0.707177	
areaDC.MARYLAND.VIRGINIA.AREA	-1.729e-01	3.585e-02	-4.823	1.42e-06	***
areaHOUSTON.AREA	-1.421e-02	4.307e-02	-0.330	0.741506	
areaMIDWEST.AREA	-1.414e-01	3.641e-02	-3.884	0.000103	***
areaNEW.ENGLAND.AREA	5.352e-02	3.859e-02	1.387	0.165512	
areaNEW.YORK.CITY.AREA	4.011e-02	3.078e-02	1.303	0.192642	
areaNORTH.FLORIDA.AREA	9.021e-02	4.213e-02	2.141	0.032262	*
areaNORTHWEST.ROCKY.MOUNTAIN.AREA	3.169e-01	4.275e-02	7.414	1.23e-13	***

areaOHIO. AREA	-1.207e-01	4.125e-02	-2.925	0.003440	**
areaPHILADELPHIA. AREA	3.400e-02	5.415e-02	0.628	0.530065	
areaSOUTH. FLORIDA. AREA	2.117e-01	4.768e-02	4.440	9.01e-06	***
areaSOUTHWEST. AREA	7.309e-02	3.693e-02	1.979	0.047816	*
areaTENNESSEE. AREA	-1.241e-01	5.177e-02	-2.397	0.016538	*
dualbandT	-7.053e-02	4.171e-02	-1.691	0.090823	.
refurb_newN	-1.878e-01	2.686e-02	-6.991	2.73e-12	***
hnd_price	-1.491e-03	1.985e-04	-7.510	5.90e-14	***
phones	2.798e-02	1.416e-02	1.976	0.048152	*
models	6.359e-02	2.186e-02	2.909	0.003627	**
hnd_webcapUNKW	4.814e-02	3.331e-02	1.445	0.148371	
hnd_webcapWCMB	-4.520e-02	3.077e-02	-1.469	0.141795	
maritalB	9.623e-02	3.448e-02	2.790	0.005263	**
maritalM	-5.328e-03	2.277e-02	-0.234	0.814965	
maritalU	1.211e-01	2.497e-02	4.850	1.24e-06	***
forgrntv1	-6.799e-02	3.523e-02	-1.930	0.053599	.
ethnicB	2.358e-01	7.134e-02	3.305	0.000951	***
ethnicC	-7.219e-01	2.014e-01	-3.584	0.000339	***
ethnicD	1.460e-01	8.822e-02	1.655	0.097880	.
ethnicH	7.983e-02	2.816e-02	2.835	0.004582	**
ethnicJ	5.494e-02	4.994e-02	1.100	0.271265	
ethnicN	-4.247e-02	2.167e-02	-1.960	0.049995	*
ethnicO	3.433e-01	4.306e-02	7.973	1.55e-15	***
ethnicP	-3.005e-01	1.313e-01	-2.290	0.022050	*
ethnicR	1.238e-01	8.162e-02	1.516	0.129455	
ethnics	-2.085e-02	2.778e-02	-0.751	0.452833	
ethnicX	-5.413e-01	3.209e-01	-1.687	0.091698	.
ethnicZ	-3.315e-01	4.327e-02	-7.662	1.83e-14	***
kid0_2U	-1.582e-01	4.261e-02	-3.714	0.000204	***
kid3_5U	-5.767e-02	3.973e-02	-1.452	0.146566	
kid16_17U	-9.314e-02	2.818e-02	-3.306	0.000947	***
creditcdN	7.453e-02	2.274e-02	3.278	0.001045	**
eqpdays	1.054e-03	5.706e-05	18.465	< 2e-16	***

---

signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

**Exhibit 5: Data Dictionary**

<b>Continuous Variables</b>	<b>Explanation</b>	<b>% Missing</b>
ADJMOU	Billing adjusted total minutes of use over the life of the customer	0.000%
ADJQTY	Billing adjusted total number of calls over the life of the customer	0.000%
ADJREV	Billing adjusted total revenue over the life of the customer	0.000%
ATTEMPT_MEAN	Mean number of attempted calls	0.000%
ATTEMPT_RANGE	Range of number of attempted calls	0.000%
AVG3MOU	Average monthly minutes of use over the previous three months	0.000%
AVG3QTY	Average monthly number of calls over the previous three months	0.000%
AVG3REV	Average monthly revenue over the previous three months	0.000%
AVG6MOU	Average monthly minutes of use over the previous six months	2.839%
AVG6QTY	Average monthly number of calls over the previous six months	2.839%
AVG6REV	Average monthly revenue over the previous six months	2.839%
AVGMOU	Average monthly minutes of use over the life of the customer	0.000%
AVGQTY	Average monthly number of calls over the life of the customer	0.000%
AVGREV	Average monthly revenue over the life of the customer	0.000%
BLCK_DAT_MEAN	Mean number of blocked (failed) data calls	0.000%

BLCK_DAT_RANGE	Range of number of blocked (failed) data calls	0.000%
BLCK_VCE_MEAN	Mean number of blocked (failed) voice calls	0.000%
BLCK_VCE_RANGE	Range of number of blocked (failed) voice calls	0.000%
CALLFWDV_MEAN	Mean number of call forwarding calls	0.000%
CALLFWDV_RANGE	Range of number of call forwarding calls	0.000%
CALLWAIT_MEAN	Mean number of call waiting calls	0.000%
CALLWAIT_RANGE	Range of number of call waiting calls	0.000%
CC_MOU_MEAN	Mean unrounded minutes of use of customer care (see CUSTCARE_MEAN) calls	0.000%
CC_MOU_RANGE	Range of unrounded minutes of use of customer care calls	0.000%
CCRNDMOU_MEAN	Mean rounded minutes of use of customer care calls	0.000%
CCRNDMOU_RANGE	Range of rounded minutes of use of customer care calls	0.000%
CHANGE_MOU	Percentage change in monthly minutes of use vs previous three month average	0.891%
CHANGE_REV	Percentage change in monthly revenue vs previous three month average	0.891%
COMP_DAT_MEAN	Mean number of completed data calls	0.000%
COMP_DAT_RANGE	Range of number of completed data calls	0.000%
COMP_VCE_MEAN	Mean number of completed voice calls	0.000%
COMP_VCE_RANGE	Range of number of completed voice calls	0.000%
COMPLETE_MEAN	Mean number of completed calls	0.000%

COMPLETE_RANGE	Range of number of completed calls	0.000%
CUSTCARE_MEAN	Mean number of customer care calls	0.000%
CUSTCARE_RANGE	Range of number of customer care calls	0.000%
DA_MEAN	Mean number of directory assisted calls	0.357%
DA_RANGE	Range of number of directory assisted calls	0.357%
DATOVR_MEAN	Mean revenue of data overage	0.357%
DATOVR_RANGE	Range of revenue of data overage	0.357%
DROP_BLK_MEAN	Mean number of dropped or blocked calls	0.000%
DROP_BLK_RANGE	Range of number of dropped or blocked calls	0.000%
DROP_DAT_MEAN	Mean number of dropped (failed) data calls	0.000%
DROP_DAT_RANGE	Range of number of dropped (failed) data calls	0.000%
DROP_VCE_MEAN	Mean number of dropped (failed) voice calls	0.000%
DROP_VCE_RANGE	Range of number of dropped (failed) voice calls	0.000%
EQPDAYS	Number of days (age) of current equipment	0.001%
INONEMIN_MEAN	Mean number of inbound calls less than one minute	0.000%
INONEMIN_RANGE	Range of number of inbound calls less than one minute	0.000%
IWYUS_VCE_MEAN	Mean number of inbound wireless to wireless voice calls	0.000%
IWYLIS_VCE_RANGE	Range of number of inbound wireless to wireless voice calls	0.000%
MONTHS	Total number of months in service	0.000%

MOU_CDAT_MEAN	Mean unrounded minutes of use of completed data calls	0.000%
MOU_CDAT_RANGE	Range of unrounded minutes of use of completed data calls	0.000%
MOU_CVCE_MEAN	Mean unrounded minutes of use of completed voice calls	0.000%
MOU_CVCE_RANGE	Range of unrounded minutes of use of completed voice calls	0.000%
MOU_MEAN	Mean number of monthly minutes of use	0.357%
MOU_OPKD_MEAN	Mean unrounded minutes of use of off-peak data calls	0.000%
MOU_OPKD_RANGE	Range of unrounded minutes of use of off-peak data calls	0.000%
MOU_OPKV_MEAN	Mean unrounded minutes of use of off-peak voice calls	0.000%
MOU_OPKV_RANGE	Range of unrounded minutes of use of off-peak voice calls	0.000%
MOU_PEAD_MEAN	Mean unrounded minutes of use of peak data calls	0.000%
MOU_PEAD_RANGE	Range of unrounded minutes of use of peak data calls	0.000%
MOU_PEA_V_MEAN	Mean unrounded minutes of use of peak voice calls	0.000%
MOU_PEA_V_RANGE	Range of unrounded minutes of use of peak voice calls	0.000%
MOU_RANGE	Range of number of minutes of use	0.357%
MOU_RVCE_MEAN	Mean unrounded minutes of use of received voice calls	0.000%
MOU_RVCE_RANGE	Range of unrounded minutes of use of received voice calls	0.000%

MOUIWY LISV_MEAN	Mean unrounded minutes of use of inbound wireless to wireless voice calls	0.000%
MOUIWY LISV_RANGE	Range of unrounded minutes of use of inbound wireless to wireless voice calls	0.000%
MOUOWY LISV_MEAN	Mean unrounded minutes of use of outbound wireless to wireless voice calls	0.000%
MOUOWY LISV_RANGE	Range of unrounded minutes of use of outbound wireless to wireless voice calls	0.000%
OWY LIS_VCE_MEAN	Mean number of outbound wireless to wireless voice calls	0.000%
OWY LIS_VCE_RANGE	Range of number of outbound wireless to wireless voice calls	0.000%
OPK_DAT_MEAN	Mean number of off-peak data calls	0.000%
OPK_DAT_RANGE	Range of number of off-peak data calls	0.000%
OPK_VCE_MEAN	Mean number of off-peak voice calls	0.000%
OPK_VCE_RANGE	Range of number of off-peak voice calls	0.000%
OVRMOU_MEAN	Mean overage minutes of use	0.357%
OVRMOU_RANGE	Range of overage minutes of use	0.357%
OVRREV_MEAN	Mean overage revenue	0.357%
OVRREV_RANGE	Range of overage revenue	0.357%
PEAK_DAT_MEAN	Mean number of peak data calls	0.000%
PEAK_DAT_RANGE	Range of number of peak data calls	0.000%
PEAK_VCE_MEAN	Mean number of inbound and outbound peak voice calls	0.000%
PEAK_VCE_RANGE	Range of number of inbound and outbound peak voice calls	0.000%
PLCD_DAT_MEAN	Mean number of attempted data calls placed	0.000%
PLCD_DAT_RANGE	Range of number of attempted data calls placed	0.000%



PLCD_VCE_MEAN	Mean number of attempted voice calls placed	0.000%
PLCD_VCE_RANGE	Range of number of attempted voice calls placed	0.000%
RECY_SMS_MEAN	Mean number of received SMS calls	0.000%
RECV_SMS_RANGE	Range of number of received SMS calls	0.000%
RECV_VCE_MEAN	Mean number of received voice calls	0.000%
RECV_VCE_RANGE	Range of number of received voice calls	0.000%
RETDAYS	Number of days since last retention call	96.017%
REV_MEAN	Mean monthly revenue (charge amount)	0.357%
REV_RANGE	Range of revenue (charge amount)	0.357%
RMCALLS	Total number of roaming calls	85.777%
RMMOU	Total minutes of use of roaming calls	85.777%
RMREV	Total revenue of roaming calls	85.777%
ROAM_MEAN	Mean number of roaming calls	0.357%
ROAM_RANGE	Range of number of roaming calls	0.357%
THREWAY_MEAN	Mean number of three way calls	0.000%
THREWAY_RANGE	Range of number of three way calls	0.000%
TOTCALLS	Total number of calls over the life of the customer	0.000%
TOTMOU	Total minutes of use over the life of the customer	0.000%
TOTMRC_MEAN	Mean total monthly recurring charge	0.357%
TOTMRC_RANGE	Range of total monthly recurring charge	0.357%
TOTREV	Total revenue	0.000%

UNAN_DAT_MEAN	Mean number of unanswered data calls	0.000%
UNAN_DAT_RANGE	Range of number of unanswered data calls	0.000%
UNAN_VCE_MEAN	Mean number of unanswered voice calls	0.000%
UNAN_VCE_RANGE	Range of number of unanswered voice calls	0.000%
VCEOVR_MEAN	Mean revenue of voice overage	0.357%
VCEOVR_RANGE	Range of revenue of voice overage	0.357%
<b>Category Variables</b>	<b>Explanation</b>	<b>% Missing</b>
ACTVSUBS	Number of active subscribers in household	0.000%
ADULTS	Number of adults in household	23.019%
AGE1	Age of first household member	1.732%
AREA	Geographic area	0.040%
ASL_FLAG	Account spending limit	0.000%
CAR_BUY	New or used car buyer	1.732%
CARTYPE	Dominant vehicle lifestyle	68.412%
CHILDREN	Children present in household	65.928%
CHURN	Instance of churn between 31-60 days after observation date	0.000%
CRCLSCOD	Credit class code	0.000%
CREDITCD	Credit card indicator	1.732%
CRTCOUNT	Adjustments made to credit rating of individual	96.500%
CSA	Communications local service area	0.000%

DUALBAND	Dualband	0.001%
DWLLSIZE	Dwelling size	38.308%
DWLLTYPE	Dwelling unit type	31.909%
ETHNIC	Ethnicity roll-up code	1.732%
FORGNTVL	Foreign travel dummy variable	1.732%
HND_PRICE	Current handset price	0.847%
HHSTATIN	Premier household status indicator	37.923%
HND_WEBCAP	Handset web capability	0.001%
INCOME	Estimated income	25.436%
INFOBASE	InfoBase match	22.079%
KID0_2	Child 0 - 2 years of age in household	1.732%
KID3_5	Child 3 - 5 years of age in household	1.732%
KID6_10	Child 6 - 10 years of age in household	1.732%
KID11_15	Child 11 - 15 years of age in household	1.732%
KID16_17	Child 16 - 17 years of age in household	1.732%
LOR	Length of residence	30.190%
MARITAL	Marital status	1.732%
MODELS	Number of models issued	0.001%
NEW_CELL	New cell phone user	0.000%
NUMBCARS	Known number of vehicles	49.366%
OWNRENT	Home owner/renter status	33.706%
PCOWNER	PC owner dummy variable	81.534%
PHONES	Number of handsets issued	0.001%
PRE_HND_PRICE	Previous handset price	57.515%

PRIZM_SOCIAL_O NE	Social group letter only	7.388%
PROPTYPE	Property type detail	71.788%
REFURB_NEW	Handset: refurbished or new	0.001%
RV	RV indicator	1.732%
TRUCK	Truck indicator	1.732%
UNIQSUBS	Number of unique subscribers in the household	0.000%