



CUSTOMER CHURN RATE ANALYSIS

PRESENTED BY LINYA (LEAH) LIU AT SEPTEMBER 21, 2019

This slide is protected by copyright law and may not be reproduced, in whole or in part, without the express written consent of the author.

METHODOLOGIES

R-LANGUAGE PROCEDURE

- Step 1: Install packages & Read data from csv files
- Step 2: Transform variables to numeric form (no=1, yes=0)
- Step 3: Begin exploratory data analysis
- Step 5: Forward elimination
- Step 6: Logit regression & prediction

FORMULA

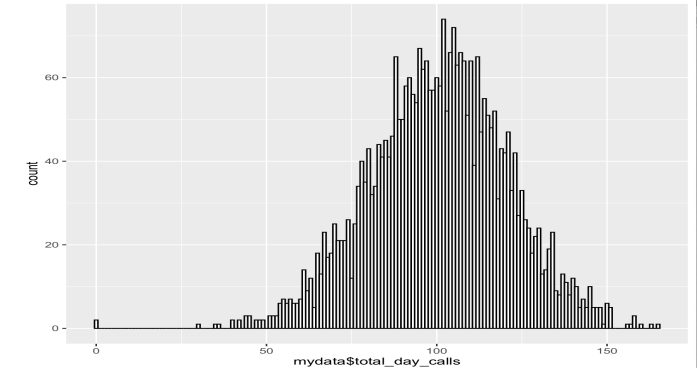
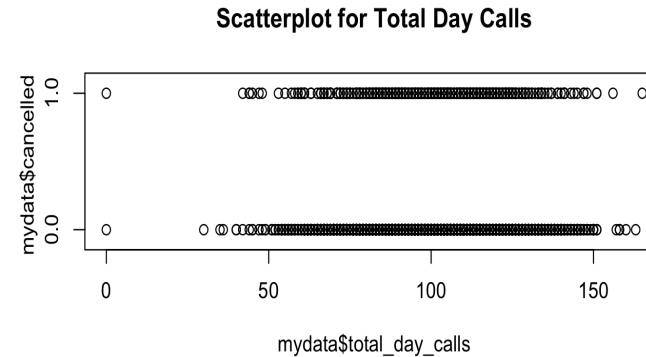
$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

OUTCOME

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-7.161862	0.456949	-15.673	< 2e-16 ***
voice_mail_plan	-1.010683	0.173294	-5.832	5.47e-09 ***
total_charges	0.076140	0.006696	11.371	< 2e-16 ***
customer_support_calls	0.502574	0.045805	10.972	< 2e-16 ***

GRAPHS



RECOMMENDATIONS

PREDICTION OF CANCELLATION STATUS

Based on the Excel confidence prediction, 4 out of the ten is estimated to cancel the plan, but 6 of them is going to stay.

RECOMMENDATIONS FOR CUSTOMER RETENTION

According to AIC estimators, voice_mail_plan, total_charges, and customer_support_calls are mostly significant. With a small standard error, the risk of variation is low. Therefore, the company should focus on improve these three categories.

SOURCES

1Instructions_Fall_2019_Data_Challenge.pdf

2 WCAI Data Challenge Fall 2019.xlsx

* AIC is an estimator for out-of-sample deviance, which lower indicates a better model

```
> summary(mydata)
```

```
X...customer_id account_age_days international_plan voice_mail_plan
Min. : 0.0 Min. : -999.00 Min. : 0.000000 Min. : 0.0000
1st Qu.: 825.5 1st Qu.: 74.00 1st Qu.: 0.000000 1st Qu.: 0.0000
Median : 1667.0 Median : 100.00 Median : 0.000000 Median : 0.0000
Mean : 1664.5 Mean : 99.64 Mean : 0.09743 Mean : 0.2805
3rd Qu.: 2500.5 3rd Qu.: 127.00 3rd Qu.: 0.000000 3rd Qu.: 1.0000
Max. : 3332.0 Max. : 243.00 Max. : 1.000000 Max. : 1.0000
```

```
number_vmail_messages total_day_calls total_eve_calls total_night_calls
Min. : 0.000 Min. : 0.0 Min. : 0.0 Min. : 33.0
1st Qu.: 0.000 1st Qu.: 87.0 1st Qu.: 87.0 1st Qu.: 87.0
Median : 0.000 Median : 101.0 Median : 100.0 Median : 100.0
Mean : 8.196 Mean : 100.5 Mean : 100.1 Mean : 100.2
3rd Qu.: 20.000 3rd Qu.: 114.0 3rd Qu.: 113.0 3rd Qu.: 114.0
Max. : 51.000 Max. : 165.0 Max. : 170.0 Max. : 175.0
```

```
total_intl_calls total_minutes total_charges customer_support_calls
Min. : 0.000 Min. : 284.3 Min. : 22.93 Min. : 0.000
1st Qu.: 3.000 1st Qu.: 531.4 1st Qu.: 52.44 1st Qu.: 1.000
Median : 4.000 Median : 593.8 Median : 59.45 Median : 1.000
Mean : 4.484 Mean : 591.6 Mean : 59.42 Mean : 1.571
3rd Qu.: 6.000 3rd Qu.: 652.4 3rd Qu.: 66.47 3rd Qu.: 2.000
Max. : 20.000 Max. : 885.0 Max. : 96.15 Max. : 9.000
```

```
cancelled X
Min. : 0.0000 Mode: logical
1st Qu.: 0.0000 NA's: 3151
Median : 0.0000
Mean : 0.1426
3rd Qu.: 0.0000
Max. : 1.0000
NA's : 10
```

```
> sapply(mydata, sd)
```

```
X...customer_id account_age_days international_plan
965.3157460 55.7983856 0.2965886
voice_mail_plan number_vmail_messages total_day_calls
0.4493372 13.7302144 20.0988579
total_eve_calls total_night_calls total_intl_calls
19.8372614 19.5707439 2.4655707
total_minutes total_charges customer_support_calls
89.9661205 10.5011898 1.3216198
cancelled X
NA NA
```

```
> cormatrix <- round(cor(mydata), digits=2)
```

```
> cormatrix
```

```
X...customer_id account_age_days international_plan
```

```
X...customer_id 1.00 0.01 -0.01
account_age_days 0.01 1.00 0.01
international_plan -0.01 0.01 1.00
voice_mail_plan 0.02 -0.01 0.00
```

```
number_vmail_messages 0.03 -0.02 0.01
```

```
total_day_calls -0.02 0.03 0.00
```

```
total_eve_calls 0.00 -0.01 0.01
```

```
total_night_calls -0.02 0.01 0.02
```

```
total_intl_calls 0.02 0.02 0.02
```

```
total_minutes 0.01 -0.01 0.03
```

```
total_charges 0.02 -0.01 0.05
```

```
customer_support_calls 0.00 0.01 -0.03
```

```
cancelled NA NA NA
```

```
X NA NA NA
```

```
voice_mail_plan number_vmail_messages total_day_calls
```

```
X...customer_id 0.02 0.03 -0.02
```

```
account_age_days -0.01 -0.02 0.03
```

```
international_plan 0.00 0.01 0.00
```

```
voice_mail_plan 1.00 0.96 -0.01
```

```
number_vmail_messages 0.96 1.00 0.00
```

```
total_day_calls -0.01 0.00 1.00
```

```
total_eve_calls -0.01 -0.01 0.00
```

```
total_night_calls 0.01 0.01 -0.02
```

```
total_intl_calls 0.00 0.01 0.01
```

```
total_minutes 0.02 0.02 0.01
```

```
total_charges 0.01 0.02 0.01
```

```
customer_support_calls -0.02 -0.01 -0.02
```

```
cancelled NA NA NA
```

```
X NA NA NA
```

```
total_eve_calls total_night_calls total_intl_calls
```

```
X...customer_id 0.00 -0.02 0.02
```

```
account_age_days -0.01 0.01 0.02
```

```
international_plan 0.01 0.02 0.02
```

```
voice_mail_plan -0.01 0.01 0.00
```

```
number_vmail_messages -0.01 0.01 0.01
```

```
total_day_calls 0.00 -0.02 0.01
```

```
total_eve_calls 1.00 0.01 0.02
```

```
total_night_calls 0.01 1.00 0.00
```

```
total_intl_calls 0.02 0.00 1.00
```

```
total_minutes 0.01 0.03 0.00
```

```
total_charges 0.02 0.03 0.01
```

```
customer_support_calls 0.01 -0.02 -0.02
```

```
cancelled NA NA NA
```

```
X NA NA NA
```

```
total_intl_calls 0.00 0.01 -0.02
total_minutes 1.00 0.89 -0.03
total_charges 0.89 1.00 -0.02
customer_support_calls -0.03 -0.02 1.00
cancelled NA NA NA
X NA NA NA
```

```
cancelled X
```

```
X...customer_id NA NA
```

```
account_age_days NA NA
```

```
international_plan NA NA
```

```
voice_mail_plan NA NA
```

```
number_vmail_messages NA NA
```

```
total_day_calls NA NA
```

```
total_eve_calls NA NA
```

```
total_night_calls NA NA
```

```
total_intl_calls NA NA
```

```
total_minutes NA NA
```

```
total_charges NA NA
```

```
customer_support_calls NA NA
```

```
cancelled 1 NA
```

```
X NA 1
```

Start: AIC=1709.62
cancelled ~ 1

	Df	Deviance	AIC
+ international_plan	1	261.94	1561.9
+ total_charges	1	264.92	1587.0
+ customer_support_calls	1	265.26	1589.9
+ total_minutes	1	268.80	1619.3
+ voice_mail_plan	1	276.73	1683.8
+ number_vmail_messages	1	277.81	1692.4
+ total_intl_calls	1	279.35	1704.7
+ account_age_days	1	279.82	1708.4
+ total_day_calls	1	279.97	1709.6
<none>		280.22	1709.6
+ total_eve_calls	1	280.01	1709.9
+ total_night_calls	1	280.05	1710.2

Step: AIC=1561.91
cancelled ~ international_plan

	Df	Deviance	AIC
+ customer_support_calls	1	246.30	1427.3
+ total_charges	1	247.60	1439.0
+ total_minutes	1	250.69	1466.5
+ voice_mail_plan	1	258.58	1535.3
+ number_vmail_messages	1	259.52	1543.3
+ total_intl_calls	1	260.82	1554.4
+ account_age_days	1	261.51	1560.3
+ total_day_calls	1	261.69	1561.8
<none>		261.94	1561.9
+ total_eve_calls	1	261.79	1562.6
+ total_night_calls	1	261.84	1563.1

Step: AIC=1427.28
cancelled ~ international_plan + customer_support_calls

	Df	Deviance	AIC
+ total_charges	1	231.33	1290.2
+ total_minutes	1	234.40	1319.4
+ voice_mail_plan	1	242.98	1399.2
+ number_vmail_messages	1	243.84	1407.0
+ total_intl_calls	1	245.34	1420.7
+ account_age_days	1	245.86	1425.3
+ total_day_calls	1	246.00	1426.6
<none>		246.30	1427.3
+ total_eve_calls	1	246.15	1428.0
+ total_night_calls	1	246.16	1428.1

+ number_vmail_messages	1	228.64	1266.2
+ total_intl_calls	1	230.31	1282.4
+ account_age_days	1	230.87	1287.8
+ total_day_calls	1	231.09	1289.8
<none>		231.33	1290.2
+ total_eve_calls	1	231.23	1291.2
+ total_night_calls	1	231.31	1292.0
+ total_minutes	1	231.33	1292.2

Step: AIC=1257.47
cancelled ~ international_plan + customer_support_calls +
total_charges +
voice_mail_plan

	Df	Deviance	AIC
+ total_intl_calls	1	226.76	1249.9
+ account_age_days	1	227.28	1254.9
+ number_vmail_messages	1	227.41	1256.2
+ total_day_calls	1	227.51	1257.2
<none>		227.74	1257.5
+ total_eve_calls	1	227.66	1258.7
+ total_night_calls	1	227.72	1259.3
+ total_minutes	1	227.74	1259.5

Step: AIC=1249.91
cancelled ~ international_plan + customer_support_calls +
total_charges +
voice_mail_plan + total_intl_calls

	Df	Deviance	AIC
+ account_age_days	1	226.27	1247.1
+ number_vmail_messages	1	226.40	1248.3
+ total_day_calls	1	226.52	1249.6
<none>		226.76	1249.9
+ total_eve_calls	1	226.68	1251.1
+ total_night_calls	1	226.75	1251.7
+ total_minutes	1	226.76	1251.9

+ total_eve_calls	1	226.18	1248.2
+ total_night_calls	1	226.26	1249.0
+ total_minutes	1	226.27	1249.1

Step: AIC=1245.23
cancelled ~ international_plan + customer_support_calls +
total_charges +
voice_mail_plan + total_intl_calls + account_age_days +
number_vmail_messages

	Df	Deviance	AIC
+ total_day_calls	1	225.66	1245.2
<none>		225.88	1245.2
+ total_eve_calls	1	225.79	1246.4
+ total_night_calls	1	225.86	1247.1
+ total_minutes	1	225.87	1247.2

Step: AIC=1245.15
cancelled ~ international_plan + customer_support_calls +
total_charges +
voice_mail_plan + total_intl_calls + account_age_days +
number_vmail_messages +
total_day_calls

	Df	Deviance	AIC
<none>		225.66	1245.2
+ total_eve_calls	1	225.58	1246.3
+ total_night_calls	1	225.65	1247.0
+ total_minutes	1	225.66	1247.1