**CAPSTONE PROJECT ASSIGNMENT BY AVIJIT MALLICK (JIG15684)**

**PATH: D:\Users\Jig15684\Practice\Avijit Mallick Graded assignment\AVIJIT MALLICK CAPSTONE PROJECT.**

/\*\*\*\*\*Import File from source location\*\*\*\*\*/

libname SAS "Y:\Practice\Graded assignment\Capstone project";

**run**;

**PROC** **IMPORT** DATAFILE = "Z:\Assignments\Graded Assignment\Topic 13 - Final Case Study Implementation\telecomfinal.csv"

OUT = SAS.cap

DBMS = csv REPLACE;

**RUN**;

**proc** **contents** data = sas.cap;

**run**;

/\*\*\*\*\*Our main target variable is the CHURN and since it is having

categorical variable so we will go for logistic regression\*\*\*\*\*/

/\*\*\*\*\*making another datasheet that will be more relevent to churn thus simplyfy our problem\*\*\*\*\*/

**data** sas.cap1 (keep = Customer\_ID ACTVSUBS ADJREV ADJMOU AVGMOU

AVGREV AVGQTY AGE1 AGE2 BLCK\_DAT\_MEAN CALLWAIT\_MEAN

CALLWAIT\_RANGE CHANGE\_MOU CHILDREN CHURN COMP\_VCE\_MEAN CUSTCARE\_MEAN CSA DATOVR\_MEAN DA\_MEAN

DROP\_BLK\_MEAN DROP\_DAT\_MEAN DROP\_VCE\_MEAN EQPDAYS FORGNTVL HND\_WEBCAP INCOME MARITAL MONTHS MOU\_MEAN

PRIZM\_SOCIAL\_ONE OCCU1 OVRMOU\_MEAN OVRREV\_MEAN RETDAYS REV\_MEAN ROAM\_MEAN TOTCALLS

TOTREV WRKWOMAN ASL\_FLAG DWLLTYPE REFURB\_NEW MTRCYCLE truck HND\_PRICE MODELS NUMBCARS);

set sas.cap;

**RUN**;

**proc** **contents** data = sas.cap1;

**run**;

|  |
| --- |
| The WPS System |

The CONTENTS Procedure

| Data Set Name | CAP1 |
| --- | --- |
| Member Type | DATA |
| Engine | WPD |
| Created | 20OCT2017:11:39:53 |
| Last Modified | 20OCT2017:11:39:53 |
| Observations | 66297 |
| Variables | 48 |
| Indexes | 0 |
| Observation Length | 317 |
| Deleted Observations | 0 |
| Data Set Type |  |
| Label |  |
| Compressed | NO |
| Sorted | NO |
| Data Representation | Little endian, IEEE Windows |
| Encoding | wlatin1 Windows-1252 Western |

| **Engine/Host Dependent Information** | |
| --- | --- |
| Data Set Page Size | 4096 |
| Number of Data Set Pages | 5526 |
| First Data Page | 1 |
| Max Obs Per Page | 12 |
| Obs In First Data Page | 12 |
| Data Set Diagnostic Code | 0009 |
| File Name | Y:\Practice\Graded assignment\Capstone project\CAP1.wpd |
| WPD Engine Version | 3 |
| Large Data Set Support | no |

| **Alphabetic List of Variables and Attributes** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Number** | **Variable** | **Type** | **Len** | **Pos** | **Format** | **Informat** | **Label** |
| 48 | Customer\_ID | Num | 8 | 272 | BEST12. | BEST32. | Customer\_ID |
| 26 | actvsubs | Num | 8 | 152 | BEST12. | BEST32. | actvsubs |
| 44 | adjmou | Num | 8 | 240 | BEST12. | BEST32. | adjmou |
| 46 | adjrev | Num | 8 | 256 | BEST12. | BEST32. | adjrev |
| 22 | age1 | Num | 8 | 120 | BEST12. | BEST32. | age1 |
| 23 | age2 | Num | 8 | 128 | BEST12. | BEST32. | age2 |
| 17 | asl\_flag | Char | 1 | 283 | $1. | $1. | asl\_flag |
| 15 | avgmou | Num | 8 | 104 | BEST12. | BEST32. | avgmou |
| 16 | avgqty | Num | 8 | 112 | BEST12. | BEST32. | avgqty |
| 47 | avgrev | Num | 8 | 264 | BEST12. | BEST32. | avgrev |
| 36 | blck\_dat\_Mean | Num | 8 | 192 | BEST12. | BEST32. | blck\_dat\_Mean |
| 9 | callwait\_Mean | Num | 8 | 56 | BEST12. | BEST32. | callwait\_Mean |
| 10 | callwait\_Range | Num | 8 | 64 | BEST12. | BEST32. | callwait\_Range |
| 2 | change\_mou | Num | 8 | 8 | BEST12. | BEST32. | change\_mou |
| 38 | children | Char | 3 | 305 | $3. | $3. | children |
| 37 | churn | Num | 8 | 200 | BEST12. | BEST32. | churn |
| 14 | comp\_vce\_Mean | Num | 8 | 96 | BEST12. | BEST32. | comp\_vce\_Mean |
| 39 | csa | Char | 9 | 308 | $9. | $9. | csa |
| 8 | custcare\_Mean | Num | 8 | 48 | BEST12. | BEST32. | custcare\_Mean |
| 40 | da\_Mean | Num | 8 | 208 | BEST12. | BEST32. | da\_Mean |
| 41 | datovr\_Mean | Num | 8 | 216 | BEST12. | BEST32. | datovr\_Mean |
| 3 | drop\_blk\_Mean | Num | 8 | 16 | BEST12. | BEST32. | drop\_blk\_Mean |
| 42 | drop\_dat\_Mean | Num | 8 | 224 | BEST12. | BEST32. | drop\_dat\_Mean |
| 43 | drop\_vce\_Mean | Num | 8 | 232 | BEST12. | BEST32. | drop\_vce\_Mean |
| 28 | dwlltype | Char | 3 | 293 | $3. | $3. | dwlltype |
| 7 | eqpdays | Num | 8 | 40 | BEST12. | BEST32. | eqpdays |
| 27 | forgntvl | Num | 8 | 160 | BEST12. | BEST32. | forgntvl |
| 25 | hnd\_price | Num | 8 | 144 | BEST12. | BEST32. | hnd\_price |
| 20 | hnd\_webcap | Char | 4 | 288 | $4. | $4. | hnd\_webcap |
| 6 | income | Char | 3 | 280 | $3. | $3. | income |
| 21 | marital | Char | 1 | 292 | $1. | $1. | marital |
| 24 | models | Num | 8 | 136 | BEST12. | BEST32. | models |
| 4 | months | Num | 8 | 24 | BEST12. | BEST32. | months |
| 1 | mou\_Mean | Num | 8 | 0 | BEST12. | BEST32. | mou\_Mean |
| 30 | mtrcycle | Num | 8 | 168 | BEST12. | BEST32. | mtrcycle |
| 31 | numbcars | Char | 3 | 298 | $3. | $3. | numbcars |
| 29 | occu1 | Char | 2 | 296 | $2. | $2. | occu1 |
| 13 | ovrmou\_Mean | Num | 8 | 88 | BEST12. | BEST32. | ovrmou\_Mean |
| 11 | ovrrev\_Mean | Num | 8 | 72 | BEST12. | BEST32. | ovrrev\_Mean |
| 18 | prizm\_social\_one | Char | 3 | 284 | $3. | $3. | prizm\_social\_one |
| 19 | refurb\_new | Char | 1 | 287 | $1. | $1. | refurb\_new |
| 32 | retdays | Char | 2 | 301 | $2. | $2. | retdays |
| 12 | rev\_Mean | Num | 8 | 80 | BEST12. | BEST32. | rev\_Mean |
| 35 | roam\_Mean | Num | 8 | 184 | BEST12. | BEST32. | roam\_Mean |
| 5 | totcalls | Num | 8 | 32 | BEST12. | BEST32. | totcalls |
| 45 | totrev | Num | 8 | 248 | BEST12. | BEST32. | totrev |
| 33 | truck | Num | 8 | 176 | BEST12. | BEST32. | truck |
| 34 | wrkwoman | Char | 2 | 303 | $2. | $2. | wrkwoman |

/\*\*\*\*\*change char variable to numeric variable and check for missing values and outliers \*\*\*\*\*/

**proc** **means** n nmiss min max stddev data = sas.cap3;

**run**;

|  |
| --- |
| The WPS System |

The MEANS Procedure

| **Summary statistics** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Label** | **N** | **N Miss** | **Minimum** | **Maximum** | **Std Dev** |
| mou\_Mean | mou\_Mean | 66012 | 0 | 0 | 12206.75 | 535.9083838 |
| change\_mou | change\_mou | 66012 | 0 | -3875.00 | 31219.25 | 281.3724438 |
| drop\_blk\_Mean | drop\_blk\_Mean | 66012 | 0 | 0 | 489.6666667 | 15.5763714 |
| months | months | 66012 | 0 | 6.0000000 | 61.0000000 | 9.7835160 |
| totcalls | totcalls | 66012 | 0 | 0 | 98874.00 | 3956.97 |
| eqpdays | eqpdays | 66012 | 0 | 0 | 1823.00 | 252.0841740 |
| custcare\_Mean | custcare\_Mean | 66012 | 0 | 0 | 365.6666667 | 5.1828111 |
| callwait\_Mean | callwait\_Mean | 66012 | 0 | 0 | 212.6666667 | 5.6187427 |
| callwait\_Range | callwait\_Range | 66012 | 0 | 0 | 192.0000000 | 4.6556102 |
| ovrrev\_Mean | ovrrev\_Mean | 66012 | 0 | 0 | 896.0875000 | 30.0665910 |
| rev\_Mean | rev\_Mean | 66012 | 0 | -6.1675000 | 3843.26 | 47.1063875 |
| ovrmou\_Mean | ovrmou\_Mean | 66012 | 0 | 0 | 4320.75 | 97.5332057 |
| comp\_vce\_Mean | comp\_vce\_Mean | 66012 | 0 | 0 | 1894.33 | 120.4987043 |
| avgmou | avgmou | 66012 | 0 | 0 | 7040.13 | 446.2562528 |
| avgqty | avgqty | 66012 | 0 | 0 | 2708.29 | 172.3364470 |
| age1 | age1 | 66012 | 0 | 0 | 99.0000000 | 21.8824196 |
| age2 | age2 | 66012 | 0 | 0 | 99.0000000 | 23.7244130 |
| models | models | 66012 | 0 | 1.0000000 | 16.0000000 | 0.9107142 |
| hnd\_price | hnd\_price | 66012 | 0 | 9.9899979 | 499.9899902 | 60.9977975 |
| actvsubs | actvsubs | 66012 | 0 | 0 | 11.0000000 | 0.6337366 |
| forgntvl | forgntvl | 66012 | 0 | 0 | 1.0000000 | 0.2326232 |
| mtrcycle | mtrcycle | 66012 | 0 | 0 | 1.0000000 | 0.1143528 |
| truck | truck | 66012 | 0 | 0 | 1.0000000 | 0.3891064 |
| roam\_Mean | roam\_Mean | 66012 | 0 | 0 | 3685.20 | 17.1065565 |
| blck\_dat\_Mean | blck\_dat\_Mean | 66012 | 0 | 0 | 413.3333333 | 1.7764224 |
| churn | churn | 66012 | 0 | 0 | 1.0000000 | 0.4264401 |
| da\_Mean | da\_Mean | 66012 | 0 | 0 | 159.3900000 | 2.2211249 |
| datovr\_Mean | datovr\_Mean | 66012 | 0 | 0 | 242.8725000 | 2.5160844 |
| drop\_dat\_Mean | drop\_dat\_Mean | 66012 | 0 | 0 | 207.3333333 | 1.0244546 |
| drop\_vce\_Mean | drop\_vce\_Mean | 66012 | 0 | 0 | 221.6666667 | 9.1005853 |
| adjmou | adjmou | 66012 | 0 | 0 | 232855.10 | 8931.48 |
| totrev | totrev | 66012 | 0 | 3.7500000 | 27321.50 | 882.6329284 |
| adjrev | adjrev | 66012 | 0 | 2.4000000 | 27071.30 | 869.6355445 |
| avgrev | avgrev | 66012 | 0 | 0.4800000 | 902.3800000 | 36.5690869 |
| Customer\_ID | Customer\_ID | 66012 | 0 | 1000001.00 | 1099994.00 | 29249.48 |
| child |  | 66012 | 0 | 0 | 1.0000000 | 0.4277679 |
| account |  | 66012 | 0 | 0 | 1.0000000 | 0.3575368 |
| dwelt |  | 66012 | 0 | 0 | 1.0000000 | 0.3723554 |
| refurb |  | 66012 | 0 | 0 | 1.0000000 | 0.3449399 |
| urban |  | 66012 | 0 | 0 | 1.0000000 | 0.4266995 |
| city |  | 66012 | 0 | 0 | 1.0000000 | 0.3763257 |
| suburban |  | 66012 | 0 | 0 | 1.0000000 | 0.4672415 |
| Town |  | 66012 | 0 | 0 | 1.0000000 | 0.3561839 |
| married |  | 66012 | 0 | 0 | 1.0000000 | 0.4635297 |
| marital\_A |  | 66012 | 0 | 0 | 1.0000000 | 0.2223151 |
| unmarried |  | 66012 | 0 | 0 | 1.0000000 | 0.4819425 |
| marital\_S |  | 66012 | 0 | 0 | 1.0000000 | 0.3838346 |
| marital\_B |  | 66012 | 0 | 0 | 1.0000000 | 0.2577645 |
| webcap |  | 66012 | 0 | 0 | 1.0000000 | 0.3322756 |
| cars |  | 66012 | 0 | 1.0000000 | 3.0000000 | 0.4487111 |
| retencall |  | 2132 | 63880 | 0 | 99.0000000 | 22.6607102 |
| wrkwoman1 |  | 66012 | 0 | 0 | 1.0000000 | 0.3317577 |
| income1 |  | 66012 | 0 | 1.0000000 | 9.0000000 | 1.8931091 |

/\*\*\*\*\*Some more data preparation \*\*\*\*\*/

**data** sas.cap2;

set sas.cap1;

if children = "Y" then child = 1;

else if children = "NA" then child = 0;

If asl\_flag = "Y" then account = 1;

else if asl\_flag = "N" then account = 0;

if dwlltype = "M" then dwelt = 1;

else if dwlltype = "S" then dwelt = 0;

if refurb\_new = "N" then refurb = 1;

else if refurb\_new = "R" then refurb = 0;

if PRIZM\_SOCIAL\_ONE = "U" then urban = 1;

else urban = 0;

if PRIZM\_SOCIAL\_ONE = "C" then city = 1;

else city = 0;

if PRIZM\_SOCIAL\_ONE = "S" then suburban = 1;

else suburban = 0;

if PRIZM\_SOCIAL\_ONE = "T" then Town = 1;

else town = 0;

if marital = "M" then married = 1;

else married = 0;

if marital = "A" then marital\_A = 1;

else marital\_A = 0;

if marital = "U" then unmarried = 1;

else unmarried = 0;

if marital = "S" then marital\_S = 1;

else marital\_S = 0;

if marital = "B" then marital\_B = 1;

else marital\_B = 0;

if hnd\_webcap = "WCMB" then webcap =1;

else if hnd\_webcap = "WC" then webcap =0;

cars = numbcars\*1;

retencall = retdays\*1;

if wrkwoman = 'Y' then wrkwoman1 = 1;

else wrkwoman1 = 0;

income1 = income\*1;

drop children asl\_flag dwlltype refurb\_new PRIZM\_SOCIAL\_ONE marital numbcars wrkwoman income

hnd\_webcap retdays;

**run**;

/\*Imputing the missing values with the averages of the variable\*/

**data** sas.cap3;

set sas.cap2;

if mou\_mean = "." then delete;

if eqpdays <0 then delete;

if change\_mou = "." then change\_mou= -9.19;

if eqpdays = "." then eqpdays = 377.14;

if ovrrev\_Mean = "." then ovrrev\_Mean = 13.26;

if rev\_mean = "." then rev\_mean = 59.08;

if ovrmou\_Mean = "." then ovrmou\_Mean = 40.18;

if age1 = "." then age1 = 31.39;

if age2 = "." then age2 = 21.13;

if models = "." then models = 1.57;

if hnd\_price = "." then hnd\_price = 105.16;

if forgntvl = "." then forgntvl = 0.058;

if mtrcycle = "." then mtrcycle = 0.0134;

if truck = "." then truck = 0.1897;

if roam\_mean = "." then roam\_mean = 1.26;

if datovr\_Mean = "." then datovr\_Mean = 0.254;

if da\_Mean = "." then da\_Mean = 0.904;

if child = "." then child = 0;

if dwelt = "." then dwelt = 0.281;

if webcap = "." then webcap = 0.857;

if income1 = "." then income1 = 5.79;

if cars = "." then cars = 1.57;

**run**;

|  |
| --- |
| The WPS System |

The MEANS Procedure

| **Summary statistics** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Label** | **N** | **N Miss** | **Mean** | **Minimum** | **Maximum** | **Std Dev** |
| mou\_Mean | mou\_Mean | 66012 | 0 | 528.7549076 | 0 | 12206.75 | 535.9083838 |
| change\_mou | change\_mou | 66012 | 0 | -9.3048647 | -3875.00 | 31219.25 | 281.3724438 |
| drop\_blk\_Mean | drop\_blk\_Mean | 66012 | 0 | 10.1898847 | 0 | 489.6666667 | 15.5763714 |
| months | months | 66012 | 0 | 18.7001303 | 6.0000000 | 61.0000000 | 9.7835160 |
| totcalls | totcalls | 66012 | 0 | 2908.36 | 0 | 98874.00 | 3956.97 |
| eqpdays | eqpdays | 66012 | 0 | 377.1605163 | 0 | 1823.00 | 252.0841740 |
| custcare\_Mean | custcare\_Mean | 66012 | 0 | 1.8836575 | 0 | 365.6666667 | 5.1828111 |
| callwait\_Mean | callwait\_Mean | 66012 | 0 | 1.8724121 | 0 | 212.6666667 | 5.6187427 |
| callwait\_Range | callwait\_Range | 66012 | 0 | 1.9031691 | 0 | 192.0000000 | 4.6556102 |
| ovrrev\_Mean | ovrrev\_Mean | 66012 | 0 | 13.2123071 | 0 | 896.0875000 | 30.0665910 |
| rev\_Mean | rev\_Mean | 66012 | 0 | 59.0528640 | -6.1675000 | 3843.26 | 47.1063875 |
| ovrmou\_Mean | ovrmou\_Mean | 66012 | 0 | 40.1731750 | 0 | 4320.75 | 97.5332057 |
| comp\_vce\_Mean | comp\_vce\_Mean | 66012 | 0 | 112.0060848 | 0 | 1894.33 | 120.4987043 |
| avgmou | avgmou | 66012 | 0 | 488.8693695 | 0 | 7040.13 | 446.2562528 |
| avgqty | avgqty | 66012 | 0 | 175.5299635 | 0 | 2708.29 | 172.3364470 |
| age1 | age1 | 66012 | 0 | 31.3890799 | 0 | 99.0000000 | 21.8824196 |
| age2 | age2 | 66012 | 0 | 21.1372323 | 0 | 99.0000000 | 23.7244130 |
| models | models | 66012 | 0 | 1.5650639 | 1.0000000 | 16.0000000 | 0.9107142 |
| hnd\_price | hnd\_price | 66012 | 0 | 105.2047623 | 9.9899979 | 499.9899902 | 60.9977975 |
| actvsubs | actvsubs | 66012 | 0 | 1.3532540 | 0 | 11.0000000 | 0.6337366 |
| forgntvl | forgntvl | 66012 | 0 | 0.0584813 | 0 | 1.0000000 | 0.2326232 |
| mtrcycle | mtrcycle | 66012 | 0 | 0.0134878 | 0 | 1.0000000 | 0.1143528 |
| truck | truck | 66012 | 0 | 0.1902744 | 0 | 1.0000000 | 0.3891064 |
| roam\_Mean | roam\_Mean | 66012 | 0 | 1.2605024 | 0 | 3685.20 | 17.1065565 |
| blck\_dat\_Mean | blck\_dat\_Mean | 66012 | 0 | 0.0293482 | 0 | 413.3333333 | 1.7764224 |
| churn | churn | 66012 | 0 | 0.2389414 | 0 | 1.0000000 | 0.4264401 |
| da\_Mean | da\_Mean | 66012 | 0 | 0.9033520 | 0 | 159.3900000 | 2.2211249 |
| datovr\_Mean | datovr\_Mean | 66012 | 0 | 0.2540201 | 0 | 242.8725000 | 2.5160844 |
| drop\_dat\_Mean | drop\_dat\_Mean | 66012 | 0 | 0.0429467 | 0 | 207.3333333 | 1.0244546 |
| drop\_vce\_Mean | drop\_vce\_Mean | 66012 | 0 | 6.0434769 | 0 | 221.6666667 | 9.1005853 |
| adjmou | adjmou | 66012 | 0 | 7624.28 | 0 | 232855.10 | 8931.48 |
| totrev | totrev | 66012 | 0 | 1033.83 | 3.7500000 | 27321.50 | 882.6329284 |
| adjrev | adjrev | 66012 | 0 | 961.9400261 | 2.4000000 | 27071.30 | 869.6355445 |
| avgrev | avgrev | 66012 | 0 | 58.0809304 | 0.4800000 | 902.3800000 | 36.5690869 |
| Customer\_ID | Customer\_ID | 66012 | 0 | 1050477.51 | 1000001.00 | 1099994.00 | 29249.48 |
| child |  | 66012 | 0 | 0.2411228 | 0 | 1.0000000 | 0.4277679 |
| account |  | 66012 | 0 | 0.1504726 | 0 | 1.0000000 | 0.3575368 |
| dwelt |  | 66012 | 0 | 0.2811454 | 0 | 1.0000000 | 0.3723554 |
| refurb |  | 66012 | 0 | 0.8619645 | 0 | 1.0000000 | 0.3449399 |
| urban |  | 66012 | 0 | 0.2393656 | 0 | 1.0000000 | 0.4266995 |
| city |  | 66012 | 0 | 0.1707871 | 0 | 1.0000000 | 0.3763257 |
| suburban |  | 66012 | 0 | 0.3219869 | 0 | 1.0000000 | 0.4672415 |
| Town |  | 66012 | 0 | 0.1490941 | 0 | 1.0000000 | 0.3561839 |
| married |  | 66012 | 0 | 0.3125341 | 0 | 1.0000000 | 0.4635297 |
| marital\_A |  | 66012 | 0 | 0.0521420 | 0 | 1.0000000 | 0.2223151 |
| unmarried |  | 66012 | 0 | 0.3668272 | 0 | 1.0000000 | 0.4819425 |
| marital\_S |  | 66012 | 0 | 0.1795734 | 0 | 1.0000000 | 0.3838346 |
| marital\_B |  | 66012 | 0 | 0.0715627 | 0 | 1.0000000 | 0.2577645 |
| webcap |  | 66012 | 0 | 0.8579302 | 0 | 1.0000000 | 0.3322756 |
| cars |  | 66012 | 0 | 1.5683423 | 1.0000000 | 3.0000000 | 0.4487111 |
| retencall |  | 2132 | 63880 | 35.8255159 | 0 | 99.0000000 | 22.6607102 |
| wrkwoman1 |  | 66012 | 0 | 0.1259165 | 0 | 1.0000000 | 0.3317577 |
| income1 |  | 66012 | 0 | 5.7897787 | 1.0000000 | 9.0000000 | 1.8931091 |

**proc** **freq** data = sas.cap3;

table churn;

**run**;

|  |
| --- |
| The WPS System |

The FREQ Procedure

| **churn** | | | | |
| --- | --- | --- | --- | --- |
| **churn** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| 0 | 50239 | 76.11 | 50239 | 76.11 |
| 1 | 15773 | 23.89 | 66012 | 100.00 |

/\*\*\*\*\*New variables generated to make model better:\*\*\*/

**data** sas.cap4 ;

set sas.cap3;

charge\_per\_call= totrev/totcalls;

charge\_per\_min=adjrev/adjmou;

**run**;

**data** sas.cap4;

set sas.cap4;

if charge\_per\_call = "." then delete;

if charge\_per\_min = "." then delete;

**run**;

/\* data partition for data validation\*/

**proc** **surveyselect** data = sas.cap4

method=srs out= sas.sample samprate=0.7 outall;

**run**;

|  |
| --- |
| The WPS System |

The SURVEYSELECT Procedure

| Selection Method | Simple Random Sampling |
| --- | --- |

| Input Data Set | SAS.cap4 |
| --- | --- |
| Random Number Seed | 13737 |
| Sample Size | 19803 |
| Selection Probability | 0.70000909 |
| Sampling Weight | 1.42855288 |
| Output Data Set | SAS.sample |

**data** sas.train sas.valid;

set sas.cap4;

if *ranuni* (100)<0.7 then output sas.train;

else output sas.valid;

**run**;

**proc** **means** n nmiss mean min max stddev data = sas.train;

**run**;

|  |
| --- |
| The WPS System |

The MEANS Procedure

| **Summary statistics** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Label** | **N** | **N Miss** | **Mean** | **Minimum** | **Maximum** | **Std Dev** |
| mou\_Mean | mou\_Mean | 46135 | 0 | 528.1888236 | 0 | 12206.75 | 534.3842595 |
| change\_mou | change\_mou | 46135 | 0 | -8.5797160 | -2835.25 | 31219.25 | 289.7559970 |
| drop\_blk\_Mean | drop\_blk\_Mean | 46135 | 0 | 10.1584625 | 0 | 327.3333333 | 15.2636709 |
| months | months | 46135 | 0 | 18.6646797 | 6.0000000 | 60.0000000 | 9.7660639 |
| totcalls | totcalls | 46135 | 0 | 2898.66 | 0 | 98874.00 | 3945.12 |
| eqpdays | eqpdays | 46135 | 0 | 375.8862902 | 0 | 1813.00 | 251.3743541 |
| custcare\_Mean | custcare\_Mean | 46135 | 0 | 1.8748311 | 0 | 327.3333333 | 5.0229318 |
| callwait\_Mean | callwait\_Mean | 46135 | 0 | 1.8733789 | 0 | 212.6666667 | 5.6725730 |
| callwait\_Range | callwait\_Range | 46135 | 0 | 1.8869188 | 0 | 164.0000000 | 4.5102857 |
| ovrrev\_Mean | ovrrev\_Mean | 46135 | 0 | 13.1293554 | 0 | 890.7625000 | 29.5877816 |
| rev\_Mean | rev\_Mean | 46135 | 0 | 58.9403433 | -5.8625000 | 3843.26 | 47.4951141 |
| ovrmou\_Mean | ovrmou\_Mean | 46135 | 0 | 39.9414174 | 0 | 4320.75 | 97.1839945 |
| comp\_vce\_Mean | comp\_vce\_Mean | 46135 | 0 | 111.9914526 | 0 | 1894.33 | 119.5919830 |
| avgmou | avgmou | 46135 | 0 | 488.8330645 | 0 | 7040.13 | 445.3923022 |
| avgqty | avgqty | 46135 | 0 | 175.4147597 | 0 | 2580.25 | 171.7336626 |
| age1 | age1 | 46135 | 0 | 31.4472236 | 0 | 98.0000000 | 21.8376520 |
| age2 | age2 | 46135 | 0 | 21.1889936 | 0 | 99.0000000 | 23.7017665 |
| models | models | 46135 | 0 | 1.5661428 | 1.0000000 | 11.0000000 | 0.9090509 |
| hnd\_price | hnd\_price | 46135 | 0 | 105.2310859 | 9.9899979 | 499.9899902 | 60.8701192 |
| actvsubs | actvsubs | 46135 | 0 | 1.3557169 | 0 | 11.0000000 | 0.6373331 |
| forgntvl | forgntvl | 46135 | 0 | 0.0582216 | 0 | 1.0000000 | 0.2321480 |
| mtrcycle | mtrcycle | 46135 | 0 | 0.0133226 | 0 | 1.0000000 | 0.1136569 |
| truck | truck | 46135 | 0 | 0.1920369 | 0 | 1.0000000 | 0.3905337 |
| roam\_Mean | roam\_Mean | 46135 | 0 | 1.2803855 | 0 | 3685.20 | 19.3414775 |
| blck\_dat\_Mean | blck\_dat\_Mean | 46135 | 0 | 0.0229544 | 0 | 122.3333333 | 0.7662069 |
| churn | churn | 46135 | 0 | 0.2393411 | 0 | 1.0000000 | 0.4266859 |
| da\_Mean | da\_Mean | 46135 | 0 | 0.9096191 | 0 | 159.3900000 | 2.2698974 |
| datovr\_Mean | datovr\_Mean | 46135 | 0 | 0.2602259 | 0 | 242.8725000 | 2.6627846 |
| drop\_dat\_Mean | drop\_dat\_Mean | 46135 | 0 | 0.0463350 | 0 | 207.3333333 | 1.1805374 |
| drop\_vce\_Mean | drop\_vce\_Mean | 46135 | 0 | 6.0583794 | 0 | 221.6666667 | 9.1137569 |
| adjmou | adjmou | 46135 | 0 | 7610.73 | 0 | 232855.10 | 8951.82 |
| totrev | totrev | 46135 | 0 | 1031.90 | 3.7500000 | 27321.50 | 887.4926365 |
| adjrev | adjrev | 46135 | 0 | 960.0468070 | 2.4000000 | 27071.30 | 875.0101388 |
| avgrev | avgrev | 46135 | 0 | 58.0718060 | 0.4800000 | 902.3800000 | 36.5585679 |
| Customer\_ID | Customer\_ID | 46135 | 0 | 1050514.32 | 1000002.00 | 1099994.00 | 29257.97 |
| child |  | 46135 | 0 | 0.2432643 | 0 | 1.0000000 | 0.4290580 |
| account |  | 46135 | 0 | 0.1512951 | 0 | 1.0000000 | 0.3583402 |
| dwelt |  | 46135 | 0 | 0.2792257 | 0 | 1.0000000 | 0.3715284 |
| refurb |  | 46135 | 0 | 0.8597161 | 0 | 1.0000000 | 0.3472852 |
| urban |  | 46135 | 0 | 0.2363715 | 0 | 1.0000000 | 0.4248575 |
| city |  | 46135 | 0 | 0.1719302 | 0 | 1.0000000 | 0.3773239 |
| suburban |  | 46135 | 0 | 0.3224450 | 0 | 1.0000000 | 0.4674173 |
| Town |  | 46135 | 0 | 0.1513601 | 0 | 1.0000000 | 0.3584034 |
| married |  | 46135 | 0 | 0.3124092 | 0 | 1.0000000 | 0.4634807 |
| marital\_A |  | 46135 | 0 | 0.0525631 | 0 | 1.0000000 | 0.2231621 |
| unmarried |  | 46135 | 0 | 0.3676818 | 0 | 1.0000000 | 0.4821794 |
| marital\_S |  | 46135 | 0 | 0.1779560 | 0 | 1.0000000 | 0.3824798 |
| marital\_B |  | 46135 | 0 | 0.0721795 | 0 | 1.0000000 | 0.2587877 |
| webcap |  | 46135 | 0 | 0.8588293 | 0 | 1.0000000 | 0.3311688 |
| cars |  | 46135 | 0 | 1.5675752 | 1.0000000 | 3.0000000 | 0.4493591 |
| retencall |  | 1499 | 44636 | 35.3609073 | 0 | 99.0000000 | 22.2288100 |
| wrkwoman1 |  | 46135 | 0 | 0.1255663 | 0 | 1.0000000 | 0.3313635 |
| income1 |  | 46135 | 0 | 5.7912754 | 1.0000000 | 9.0000000 | 1.8946044 |
| charge\_per\_call |  | 46129 | 6 | 0.8222219 | 0.0352659 | 279.0400000 | 3.4297749 |
| charge\_per\_min |  | 46125 | 10 | 0.3173401 | 0.0100187 | 229.0600000 | 2.4716728 |

**proc** **freq** data = sas.valid;

table churn;

**run**;

|  |
| --- |
| The WPS System |

The FREQ Procedure

| **churn** | | | | |
| --- | --- | --- | --- | --- |
| **churn** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| 0 | 15146 | 76.20 | 15146 | 76.20 |
| 1 | 4731 | 23.80 | 19877 | 100.00 |

**proc** **freq** data = sas.train;

table churn;

**run**;

|  |
| --- |
| The WPS System |

The FREQ Procedure

| **churn** | | | | |
| --- | --- | --- | --- | --- |
| **churn** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| 0 | 35093 | 76.07 | 35093 | 76.07 |
| 1 | 11042 | 23.93 | 46135 | 100.00 |

**proc** **contents** data= sas.train;

**run**;

|  |
| --- |
| The WPS System |

The CONTENTS Procedure

| Data Set Name | TRAIN |
| --- | --- |
| Member Type | DATA |
| Engine | WPD |
| Created | 20OCT2017:11:50:06 |
| Last Modified | 20OCT2017:11:50:06 |
| Observations | 46135 |
| Variables | 57 |
| Indexes | 0 |
| Observation Length | 451 |
| Deleted Observations | 0 |
| Data Set Type |  |
| Label |  |
| Compressed | NO |
| Sorted | NO |
| Data Representation | Little endian, IEEE Windows |
| Encoding | wlatin1 Windows-1252 Western |

| **Engine/Host Dependent Information** | |
| --- | --- |
| Data Set Page Size | 4096 |
| Number of Data Set Pages | 5128 |
| First Data Page | 1 |
| Max Obs Per Page | 9 |
| Obs In First Data Page | 9 |
| Data Set Diagnostic Code | 0009 |
| File Name | Y:\Practice\Graded assignment\Capstone project\TRAIN.wpd |
| WPD Engine Version | 3 |
| Large Data Set Support | no |

| **Alphabetic List of Variables and Attributes** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Number** | **Variable** | **Type** | **Len** | **Pos** | **Format** | **Informat** | **Label** |
| 37 | Customer\_ID | Num | 8 | 272 | BEST12. | BEST32. | Customer\_ID |
| 45 | Town | Num | 8 | 336 |  |  |  |
| 39 | account | Num | 8 | 288 |  |  |  |
| 20 | actvsubs | Num | 8 | 152 | BEST12. | BEST32. | actvsubs |
| 33 | adjmou | Num | 8 | 240 | BEST12. | BEST32. | adjmou |
| 35 | adjrev | Num | 8 | 256 | BEST12. | BEST32. | adjrev |
| 16 | age1 | Num | 8 | 120 | BEST12. | BEST32. | age1 |
| 17 | age2 | Num | 8 | 128 | BEST12. | BEST32. | age2 |
| 14 | avgmou | Num | 8 | 104 | BEST12. | BEST32. | avgmou |
| 15 | avgqty | Num | 8 | 112 | BEST12. | BEST32. | avgqty |
| 36 | avgrev | Num | 8 | 264 | BEST12. | BEST32. | avgrev |
| 26 | blck\_dat\_Mean | Num | 8 | 192 | BEST12. | BEST32. | blck\_dat\_Mean |
| 8 | callwait\_Mean | Num | 8 | 56 | BEST12. | BEST32. | callwait\_Mean |
| 9 | callwait\_Range | Num | 8 | 64 | BEST12. | BEST32. | callwait\_Range |
| 52 | cars | Num | 8 | 392 |  |  |  |
| 2 | change\_mou | Num | 8 | 8 | BEST12. | BEST32. | change\_mou |
| 56 | charge\_per\_call | Num | 8 | 424 |  |  |  |
| 57 | charge\_per\_min | Num | 8 | 432 |  |  |  |
| 38 | child | Num | 8 | 280 |  |  |  |
| 27 | churn | Num | 8 | 200 | BEST12. | BEST32. | churn |
| 43 | city | Num | 8 | 320 |  |  |  |
| 13 | comp\_vce\_Mean | Num | 8 | 96 | BEST12. | BEST32. | comp\_vce\_Mean |
| 28 | csa | Char | 9 | 442 | $9. | $9. | csa |
| 7 | custcare\_Mean | Num | 8 | 48 | BEST12. | BEST32. | custcare\_Mean |
| 29 | da\_Mean | Num | 8 | 208 | BEST12. | BEST32. | da\_Mean |
| 30 | datovr\_Mean | Num | 8 | 216 | BEST12. | BEST32. | datovr\_Mean |
| 3 | drop\_blk\_Mean | Num | 8 | 16 | BEST12. | BEST32. | drop\_blk\_Mean |
| 31 | drop\_dat\_Mean | Num | 8 | 224 | BEST12. | BEST32. | drop\_dat\_Mean |
| 32 | drop\_vce\_Mean | Num | 8 | 232 | BEST12. | BEST32. | drop\_vce\_Mean |
| 40 | dwelt | Num | 8 | 296 |  |  |  |
| 6 | eqpdays | Num | 8 | 40 | BEST12. | BEST32. | eqpdays |
| 21 | forgntvl | Num | 8 | 160 | BEST12. | BEST32. | forgntvl |
| 19 | hnd\_price | Num | 8 | 144 | BEST12. | BEST32. | hnd\_price |
| 55 | income1 | Num | 8 | 416 |  |  |  |
| 47 | marital\_A | Num | 8 | 352 |  |  |  |
| 50 | marital\_B | Num | 8 | 376 |  |  |  |
| 49 | marital\_S | Num | 8 | 368 |  |  |  |
| 46 | married | Num | 8 | 344 |  |  |  |
| 18 | models | Num | 8 | 136 | BEST12. | BEST32. | models |
| 4 | months | Num | 8 | 24 | BEST12. | BEST32. | months |
| 1 | mou\_Mean | Num | 8 | 0 | BEST12. | BEST32. | mou\_Mean |
| 23 | mtrcycle | Num | 8 | 168 | BEST12. | BEST32. | mtrcycle |
| 22 | occu1 | Char | 2 | 440 | $2. | $2. | occu1 |
| 12 | ovrmou\_Mean | Num | 8 | 88 | BEST12. | BEST32. | ovrmou\_Mean |
| 10 | ovrrev\_Mean | Num | 8 | 72 | BEST12. | BEST32. | ovrrev\_Mean |
| 41 | refurb | Num | 8 | 304 |  |  |  |
| 53 | retencall | Num | 8 | 400 |  |  |  |
| 11 | rev\_Mean | Num | 8 | 80 | BEST12. | BEST32. | rev\_Mean |
| 25 | roam\_Mean | Num | 8 | 184 | BEST12. | BEST32. | roam\_Mean |
| 44 | suburban | Num | 8 | 328 |  |  |  |
| 5 | totcalls | Num | 8 | 32 | BEST12. | BEST32. | totcalls |
| 34 | totrev | Num | 8 | 248 | BEST12. | BEST32. | totrev |
| 24 | truck | Num | 8 | 176 | BEST12. | BEST32. | truck |
| 48 | unmarried | Num | 8 | 360 |  |  |  |
| 42 | urban | Num | 8 | 312 |  |  |  |
| 51 | webcap | Num | 8 | 384 |  |  |  |
| 54 | wrkwoman1 | Num | 8 | 408 |  |  |  |

**proc** **contents** data= sas.valid;

**run**;

|  |
| --- |
| The WPS System |

The CONTENTS Procedure

| Data Set Name | VALID |
| --- | --- |
| Member Type | DATA |
| Engine | WPD |
| Created | 20OCT2017:11:50:06 |
| Last Modified | 20OCT2017:11:50:06 |
| Observations | 19877 |
| Variables | 57 |
| Indexes | 0 |
| Observation Length | 451 |
| Deleted Observations | 0 |
| Data Set Type |  |
| Label |  |
| Compressed | NO |
| Sorted | NO |
| Data Representation | Little endian, IEEE Windows |
| Encoding | wlatin1 Windows-1252 Western |

| **Engine/Host Dependent Information** | |
| --- | --- |
| Data Set Page Size | 4096 |
| Number of Data Set Pages | 2210 |
| First Data Page | 1 |
| Max Obs Per Page | 9 |
| Obs In First Data Page | 9 |
| Data Set Diagnostic Code | 0009 |
| File Name | Y:\Practice\Graded assignment\Capstone project\VALID.wpd |
| WPD Engine Version | 3 |
| Large Data Set Support | no |

| **Alphabetic List of Variables and Attributes** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Number** | **Variable** | **Type** | **Len** | **Pos** | **Format** | **Informat** | **Label** |
| 37 | Customer\_ID | Num | 8 | 272 | BEST12. | BEST32. | Customer\_ID |
| 45 | Town | Num | 8 | 336 |  |  |  |
| 39 | account | Num | 8 | 288 |  |  |  |
| 20 | actvsubs | Num | 8 | 152 | BEST12. | BEST32. | actvsubs |
| 33 | adjmou | Num | 8 | 240 | BEST12. | BEST32. | adjmou |
| 35 | adjrev | Num | 8 | 256 | BEST12. | BEST32. | adjrev |
| 16 | age1 | Num | 8 | 120 | BEST12. | BEST32. | age1 |
| 17 | age2 | Num | 8 | 128 | BEST12. | BEST32. | age2 |
| 14 | avgmou | Num | 8 | 104 | BEST12. | BEST32. | avgmou |
| 15 | avgqty | Num | 8 | 112 | BEST12. | BEST32. | avgqty |
| 36 | avgrev | Num | 8 | 264 | BEST12. | BEST32. | avgrev |
| 26 | blck\_dat\_Mean | Num | 8 | 192 | BEST12. | BEST32. | blck\_dat\_Mean |
| 8 | callwait\_Mean | Num | 8 | 56 | BEST12. | BEST32. | callwait\_Mean |
| 9 | callwait\_Range | Num | 8 | 64 | BEST12. | BEST32. | callwait\_Range |
| 52 | cars | Num | 8 | 392 |  |  |  |
| 2 | change\_mou | Num | 8 | 8 | BEST12. | BEST32. | change\_mou |
| 56 | charge\_per\_call | Num | 8 | 424 |  |  |  |
| 57 | charge\_per\_min | Num | 8 | 432 |  |  |  |
| 38 | child | Num | 8 | 280 |  |  |  |
| 27 | churn | Num | 8 | 200 | BEST12. | BEST32. | churn |
| 43 | city | Num | 8 | 320 |  |  |  |
| 13 | comp\_vce\_Mean | Num | 8 | 96 | BEST12. | BEST32. | comp\_vce\_Mean |
| 28 | csa | Char | 9 | 442 | $9. | $9. | csa |
| 7 | custcare\_Mean | Num | 8 | 48 | BEST12. | BEST32. | custcare\_Mean |
| 29 | da\_Mean | Num | 8 | 208 | BEST12. | BEST32. | da\_Mean |
| 30 | datovr\_Mean | Num | 8 | 216 | BEST12. | BEST32. | datovr\_Mean |
| 3 | drop\_blk\_Mean | Num | 8 | 16 | BEST12. | BEST32. | drop\_blk\_Mean |
| 31 | drop\_dat\_Mean | Num | 8 | 224 | BEST12. | BEST32. | drop\_dat\_Mean |
| 32 | drop\_vce\_Mean | Num | 8 | 232 | BEST12. | BEST32. | drop\_vce\_Mean |
| 40 | dwelt | Num | 8 | 296 |  |  |  |
| 6 | eqpdays | Num | 8 | 40 | BEST12. | BEST32. | eqpdays |
| 21 | forgntvl | Num | 8 | 160 | BEST12. | BEST32. | forgntvl |
| 19 | hnd\_price | Num | 8 | 144 | BEST12. | BEST32. | hnd\_price |
| 55 | income1 | Num | 8 | 416 |  |  |  |
| 47 | marital\_A | Num | 8 | 352 |  |  |  |
| 50 | marital\_B | Num | 8 | 376 |  |  |  |
| 49 | marital\_S | Num | 8 | 368 |  |  |  |
| 46 | married | Num | 8 | 344 |  |  |  |
| 18 | models | Num | 8 | 136 | BEST12. | BEST32. | models |
| 4 | months | Num | 8 | 24 | BEST12. | BEST32. | months |
| 1 | mou\_Mean | Num | 8 | 0 | BEST12. | BEST32. | mou\_Mean |
| 23 | mtrcycle | Num | 8 | 168 | BEST12. | BEST32. | mtrcycle |
| 22 | occu1 | Char | 2 | 440 | $2. | $2. | occu1 |
| 12 | ovrmou\_Mean | Num | 8 | 88 | BEST12. | BEST32. | ovrmou\_Mean |
| 10 | ovrrev\_Mean | Num | 8 | 72 | BEST12. | BEST32. | ovrrev\_Mean |
| 41 | refurb | Num | 8 | 304 |  |  |  |
| 53 | retencall | Num | 8 | 400 |  |  |  |
| 11 | rev\_Mean | Num | 8 | 80 | BEST12. | BEST32. | rev\_Mean |
| 25 | roam\_Mean | Num | 8 | 184 | BEST12. | BEST32. | roam\_Mean |
| 44 | suburban | Num | 8 | 328 |  |  |  |
| 5 | totcalls | Num | 8 | 32 | BEST12. | BEST32. | totcalls |
| 34 | totrev | Num | 8 | 248 | BEST12. | BEST32. | totrev |
| 24 | truck | Num | 8 | 176 | BEST12. | BEST32. | truck |
| 48 | unmarried | Num | 8 | 360 |  |  |  |
| 42 | urban | Num | 8 | 312 |  |  |  |
| 51 | webcap | Num | 8 | 384 |  |  |  |
| 54 | wrkwoman1 | Num | 8 | 408 |  |  |  |

/\*\*\*\*\*\*doing iteration method of proc logistic\*\*\*///

**proc** **logistic** data = sas.train descending outmodel = sas.traindmm;

model churn = Customer\_ID Selected Town account actvsubs adjmou adjrev

age1 age1 age2 avgmou avgmou avgqty avgrev blck\_dat\_Mean callwait\_Mean

callwait\_Range cars change\_mou child city comp\_vce\_Mean da\_Mean

datovr\_Mean drop\_blk\_Mean drop\_dat\_Mean dwelt eqpdays forgntvl hnd\_price

income1 marital\_A marital\_B marital\_S married models months mou\_Mean

mtrcycle ovrmou\_Mean refurb retencall rev\_Mean roam\_Mean suburban

totcalls totrev truck unmarried urban webcap wrkwoman1 charge\_per\_call charge\_per\_min;

score out = sas.train\_dmp;

**run**;

**proc** **logistic** data = sas.train descending outmodel = sas.traindmm;

model churn =Customer\_ID Town account actvsubs adjmou

/\* adjrev\*/ age1 age1 /\*age2\*/ avgmou avgqty avgrev

/\*blck\_dat\_Mean\*/ /\* callwait\_Mean\*/ callwait\_Range

/\*cars\*/ change\_mou child /\*city\*/ comp\_vce\_Mean

da\_Mean /\*datovr\_Mean\*/ drop\_blk\_Mean /\*drop\_dat\_Mean\*/

dwelt eqpdays /\* forgntvl\*/ hnd\_price income1 /\*marital\_A\*/

marital\_B marital\_S married models months mou\_Mean /\*mtrcycle\*/ /\*ovrmou\_Mean\*/

refurb retencall rev\_Mean roam\_Mean suburban totcalls totrev truck unmarried urban webcap wrkwoman1;

score out = sas.train\_dmp;

**run**;

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*tried number of iterationbut found this most suitable\*\*\*\*\*\*\*\*\*\*/

**proc** **logistic** data = sas.train descending outmodel = sas.traindmm;

model churn =/\*Customer\_ID Selected Town account\*/ actvsubs /\*adjmou adjrev age1 age1\*/ /\*age2\*/

avgmou /\*avgqty \*/avgrev /\*blck\_dat\_Mean\*/ /\* callwait\_Mean callwait\_Range \*/

/\*cars\*/ change\_mou /\*child city\*/ comp\_vce\_Mean /\*da\_Mean datovr\_Mean

drop\_blk\_Mean\*/ /\*drop\_dat\_Mean dwelt\*/ eqpdays /\* forgntvl hnd\_price income1\*/

/\*marital\_A marital\_B marital\_S married models \*/ months mou\_Mean /\*mtrcycle\*/

/\*ovrmou\_Mean refurb retencall\*/ /\*rev\_Mean\*/ roam\_Mean /\*suburban\*/ totcalls /\*totrev

truck unmarried urban webcap\*/ wrkwoman1 custcare\_Mean drop\_vce\_Mean

ovrrev\_Mean charge\_per\_call charge\_per\_min;

score out = sas.train\_dmp;

**run**;

|  |
| --- |
| The WPS System |

The LOGISTIC Procedure

| **Model Information** | |
| --- | --- |
| Data Set | SAS.train |
| Response Variable | churn |
| Number of Response Levels | 2 |
| Model | Binary logit |
| Optimisation Technique | Fisher's scoring |

| Number of Observations Read | 46135 |
| --- | --- |
| Number of Observations Used | 46125 |

| **Response Profile** | | |
| --- | --- | --- |
| **Ordered Value** | **churn** | **Total Frequency** |
| 1 | 1 | 11037 |
| 2 | 0 | 35088 |
| **Probability modeled is churn='1'.** | | |

| **Model Convergence Status** |
| --- |
| Convergence criterion (GCONV=1e-008) satisfied. |

| **Model Fit Statistics** | | |
| --- | --- | --- |
| **Criterion** | **Intercept only** | **Intercept and Covariates** |
| AIC | 50762.917 | 49837.717 |
| SC | 50771.656 | 49986.282 |
| -2 Log L | 50760.917 | 49803.717 |

| **Testing Global Null Hypothesis: BETA=0** | | | |
| --- | --- | --- | --- |
| **Test** | **Chi-Square** | **DF** | **Pr > Chi-Square** |
| Likelihood Ratio | 957.2001 | 16 | <.0001 |
| Score | 938.8853 | 16 | <.0001 |
| Wald | 901.4944 | 16 | <.0001 |

| **Analysis of Maximum Likelihood Estimates** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Parameter** | **DF** | **Estimate** | **Standard Error** | **Wald Chi-Square** | **Pr > ChiSq** |
| Intercept | 1 | -1.4085 | 0.0451 | 975.2439 | <.0001 |
| actvsubs | 1 | 0.0774 | 0.0171 | 20.5346 | <.0001 |
| avgmou | 1 | 0.000838 | 0.000075 | 123.8730 | <.0001 |
| avgrev | 1 | -0.00242 | 0.000544 | 19.8386 | <.0001 |
| change\_mou | 1 | -0.00012 | 0.000047 | 6.6993 | 0.0096 |
| comp\_vce\_Mean | 1 | -0.00071 | 0.000200 | 12.6919 | 0.0004 |
| eqpdays | 1 | 0.00104 | 0.000054 | 372.1155 | <.0001 |
| months | 1 | -0.00963 | 0.00170 | 32.0234 | <.0001 |
| mou\_Mean | 1 | -0.00083 | 0.000064 | 166.3655 | <.0001 |
| roam\_Mean | 1 | 0.00581 | 0.00125 | 21.6378 | <.0001 |
| totcalls | 1 | 0.000015 | 4.668E-6 | 9.7298 | 0.0018 |
| wrkwoman1 | 1 | -0.1318 | 0.0341 | 14.9696 | 0.0001 |
| custcare\_Mean | 1 | -0.0121 | 0.00319 | 14.2610 | 0.0002 |
| drop\_vce\_Mean | 1 | 0.0107 | 0.00159 | 45.1801 | <.0001 |
| ovrrev\_Mean | 1 | 0.00666 | 0.000489 | 185.7116 | <.0001 |
| charge\_per\_call | 1 | -0.0189 | 0.00936 | 4.0888 | 0.0432 |
| charge\_per\_min | 1 | 0.0154 | 0.0113 | 1.8411 | 0.1748 |

| **Odds Ratio Estimates** | | | |
| --- | --- | --- | --- |
| **Effect** | **Point Estimate** | **Lower 95% Wald Confidence Limit** | **Upper 95% Wald Confidence Limit** |
| actvsubs | 1.080 | 1.045 | 1.117 |
| avgmou | 1.001 | 1.001 | 1.001 |
| avgrev | 0.998 | 0.997 | 0.999 |
| change\_mou | 1.000 | 1.000 | 1.000 |
| comp\_vce\_Mean | 0.999 | 0.999 | 1.000 |
| eqpdays | 1.001 | 1.001 | 1.001 |
| months | 0.990 | 0.987 | 0.994 |
| mou\_Mean | 0.999 | 0.999 | 0.999 |
| roam\_Mean | 1.006 | 1.003 | 1.008 |
| totcalls | 1.000 | 1.000 | 1.000 |
| wrkwoman1 | 0.877 | 0.820 | 0.937 |
| custcare\_Mean | 0.988 | 0.982 | 0.994 |
| drop\_vce\_Mean | 1.011 | 1.008 | 1.014 |
| ovrrev\_Mean | 1.007 | 1.006 | 1.008 |
| charge\_per\_call | 0.981 | 0.963 | 0.999 |
| charge\_per\_min | 1.015 | 0.993 | 1.038 |

| **Association of Predicted Probabilities and Observed Responses** | | | |
| --- | --- | --- | --- |
| Percent Concordant | 59.7 | Somer's D | 0.204 |
| Percent Discordant | 39.3 | Gamma | 0.206 |
| Percent Tied | 1.0 | Tau-a | 0.074 |
| Pairs | 387266256 | c | 0.602 |

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*rank\*\*\*\*\*\*\*\*/

**proc** **rank** data = sas.train\_dmp out = sas.decile groups = 10 ties=mean;

var p\_1;

ranks decile;

**run**;

**proc** **sort** data= sas.decile;

by descending p\_1;

**run**;

/\* export this file to creat lift chart\*/

**proc** **export** data = sas.decile

outfile = "Y:\Practice\Graded assignment\Capstone project\capston.csv"

dbms = csv replace;

**run**;

/\* Trying with complete data for creating lift chart by importing the original file again\*/

/\*Scoring is done to creat lift chart\*/

**PROC** **IMPORT** DATAFILE = "Z:\Assignments\Graded Assignment\Topic 13 - Final Case Study Implementation\telecomfinal.csv"

OUT = SAS.scoring

DBMS = csv REPLACE;

**RUN**;

**data** sas.scoring1 (keep = Customer\_ID ACTVSUBS ADJREV ADJMOU AVGMOU

AVGREV AVGQTY AGE1 AGE2 BLCK\_DAT\_MEAN CALLWAIT\_MEAN

CALLWAIT\_RANGE CHANGE\_MOU CHILDREN CHURN COMP\_VCE\_MEAN CUSTCARE\_MEAN CSA DATOVR\_MEAN DA\_MEAN

DROP\_BLK\_MEAN DROP\_DAT\_MEAN DROP\_VCE\_MEAN EQPDAYS FORGNTVL HND\_WEBCAP INCOME MARITAL MONTHS MOU\_MEAN

PRIZM\_SOCIAL\_ONE OCCU1 OVRMOU\_MEAN OVRREV\_MEAN RETDAYS REV\_MEAN ROAM\_MEAN TOTCALLS

TOTREV WRKWOMAN ASL\_FLAG DWLLTYPE REFURB\_NEW MTRCYCLE truck HND\_PRICE MODELS NUMBCARS);

set sas.scoring;

**RUN**;

/\*\*\*\*\*change char variable to numeric variable and check for missing values and outliers \*\*\*\*\*/

**data** sas.scoring2;

set sas.scoring1;

if children = "Y" then child = 1;

else if children = "NA" then child = 0;

If asl\_flag = "Y" then account = 1;

else if asl\_flag = "N" then account = 0;

if dwlltype = "M" then dwelt = 1;

else if dwlltype = "S" then dwelt = 0;

if refurb\_new = "N" then refurb = 1;

else if refurb\_new = "R" then refurb = 0;

if PRIZM\_SOCIAL\_ONE = "U" then urban = 1;

else urban = 0;

if PRIZM\_SOCIAL\_ONE = "C" then city = 1;

else city = 0;

if PRIZM\_SOCIAL\_ONE = "S" then suburban = 1;

else suburban = 0;

if PRIZM\_SOCIAL\_ONE = "T" then Town = 1;

else town = 0;

if marital = "M" then married = 1;

else married = 0;

if marital = "A" then marital\_A = 1;

else marital\_A = 0;

if marital = "U" then unmarried = 1;

else unmarried = 0;

if marital = "S" then marital\_S = 1;

else marital\_S = 0;

if marital = "B" then marital\_B = 1;

else marital\_B = 0;

if hnd\_webcap = "WCMB" then webcap =1;

else if hnd\_webcap = "WC" then webcap =0;

cars = numbcars\*1;

retencall = retdays\*1;

if wrkwoman = 'Y' then wrkwoman1 = 1;

else wrkwoman1 = 0;

income1 = income\*1;

drop children asl\_flag dwlltype refurb\_new PRIZM\_SOCIAL\_ONE marital numbcars wrkwoman income

hnd\_webcap retdays;

**run**;

/\*Imputing the missing values with the averages of the variable\*/

**data** sas.scoring3;

set sas.scoring2;

if mou\_mean = "." then delete;

if eqpdays <0 then delete;

if change\_mou = "." then change\_mou= -9.19;

if eqpdays = "." then eqpdays = 377.14;

if ovrrev\_Mean = "." then ovrrev\_Mean = 13.26;

if rev\_mean = "." then rev\_mean = 59.08;

if ovrmou\_Mean = "." then ovrmou\_Mean = 40.18;

if age1 = "." then age1 = 31.39;

if age2 = "." then age2 = 21.13;

if models = "." then models = 1.57;

if hnd\_price = "." then hnd\_price = 105.16;

if forgntvl = "." then forgntvl = 0.058;

if mtrcycle = "." then mtrcycle = 0.0134;

if truck = "." then truck = 0.1897;

if roam\_mean = "." then roam\_mean = 1.26;

if datovr\_Mean = "." then datovr\_Mean = 0.254;

if da\_Mean = "." then da\_Mean = 0.904;

if child = "." then child = 0;

if dwelt = "." then dwelt = 0.281;

if webcap = "." then webcap = 0.857;

if income1 = "." then income1 = 5.79;

if cars = "." then cars = 1.57;

charge\_per\_call= totrev/totcalls;

charge\_per\_min=adjrev/adjmou;

**run**;

**proc** **means** n nmiss mean min max stddev data = sas.scoring3;

**run**;

/\*validating with the above export file with the previously created sas.traindmm file\*/

**proc** **logistic** inmodel = sas.traindmm;

score data = sas.scoring3 out = sas.scoring4;

**run**;

**proc** **rank** data = sas.train\_dmp out = sas.decile1 groups = 10 ties=mean;

var p\_1;

ranks decile;

**run**;

**proc** **sort** data= sas.decile1;

by descending p\_1;

**run**;

/\* Exporting the scored dataset to excel for chart creation and further model validation \*/

**proc** **export** data = sas.decile1 outfile = "Y:\Practice\Graded assignment\Capstone project\capstonvarify.csv"

dbms = csv replace;

**run**;

**LIFT CHART**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Decile** | **Predicted** | **Actual** | **Cum Pred** | **Cum Actual** | **Cum Pred %** | **Cum Actual%** | **Fitting Value** |
| 9 | 16459.67 | 3816.495 | 16459.67 | 3816.495 | 43% | 10% | 4.31 |
| 8 | 13690 | 3816.495 | 30149.67 | 7632.99 | 79% | 20% | 3.95 |
| 7 | 1260.6 | 3816.495 | 31410.27 | 11449.485 | 82% | 30% | 2.74 |
| 6 | 1180.23 | 3816.495 | 32590.5 | 15265.98 | 85% | 40% | 2.13 |
| 5 | 1115.32 | 3816.495 | 33705.82 | 19082.475 | 88% | 50% | 1.77 |
| 4 | 1055.93 | 3816.495 | 34761.75 | 22898.97 | 91% | 60% | 1.52 |
| 3 | 994.82 | 3816.495 | 35756.57 | 26715.465 | 94% | 70% | 1.34 |
| 2 | 923 | 3816.495 | 36679.57 | 30531.96 | 96% | 80% | 1.20 |
| 1 | 831.93 | 3816.495 | 37511.5 | 34348.455 | 98% | 90% | 1.09 |
| 0 | 653.45 | 3816.495 | 38164.95 | 38164.95 | 100% | 100% | 1.00 |
|  | 38164.95 | 38164.95 |  |  |  |  |  |