

DETAILED DATA DESCRIPTION

THE INSURANCE COMPANY (TIC) 2000

(c) Sentient Machine Research 2000

DISCLAIMER

This dataset is owned and supplied by the Dutch datamining company Sentient Machine Research, and is based on real world business data. You are allowed to use this dataset and accompanying information for non commercial research and education purposes only. It is explicitly not allowed to use this dataset for commercial education or demonstration purposes. For any other use, please contact Peter van der Putten, info@smr.nl.

This dataset has been used in the CoIL Challenge 2000 datamining competition. For papers describing results on this dataset, see the TIC 2000 homepage: <http://www.wi.leidenuniv.nl/~putten/library/cc2000/>

REFERENCE

P. van der Putten and M. van Someren (eds). CoIL Challenge 2000: The Insurance Company Case.
Published by Sentient Machine Research, Amsterdam. Also a Leiden Institute of Advanced Computer Science Technical Report 2000-09. June 22, 2000.

RELEVANT FILES

TICDATA2000.txt:

Dataset to train and validate prediction models and build a description (5822 customer records). Each record consists of 86 attributes, containing sociodemographic data (attribute 1-43) and product ownership (attributes 44-86). The sociodemographic data is derived from zip codes. All customers living in areas with the same zip code have the same sociodemographic attributes. Attribute 86, "CARAVAN: Number of mobile home policies", is the target variable.

TICEVAL2000.txt:

Dataset for predictions (4000 customer records). It has the same format as TICDATA2000.txt, only the target is missing. Participants are supposed to return the list of predicted targets only. All datasets are in tab delimited format.

The meaning of the attributes and attribute values is given below.

TICTGTS2000.txt

Targets for the evaluation set.

DATADictionary

Nr	Name	Description	Domain
----	------	-------------	--------

1	MOSTYPE	Customer Subtype	see L0
---	---------	------------------	--------

2	MAANTHUI	Number of houses	1 ? 10
---	----------	------------------	--------

3	MGEMOMV	Avg size household	1 ? 6
---	---------	--------------------	-------

4 MGEMLEEF Avg age see L1

5 MOSHOOFD Customer main type see L2

6 MGODRK Roman catholic see L3

7 MGODPR Protestant ...

8 MGODOV Other religion

9 MGODGE No religion

10 MRELGE Married

11 MRELSA Living together

12 MRELOV Other relation

13 MFALLEEN Singles

14 MFGEKIND Household without children

15 MFWEKIND Household with children

16 MOPLHOOG High level education

17 MOPLMIDD Medium level education

18 MOPLLAAG Lower level education

19 MBERHOOG High status

20 MBERZELF Entrepreneur

21 MBERBOER Farmer

22 MBERMIDD Middle management

23 MBERARBG Skilled labourers

24 MBERARBO Unskilled labourers

25 MSKA Social class A

26 MSKB1 Social class B1

27 MSKB2 Social class B2

28 MSKC Social class C

29 MSKD Social class D

30 MHHUUR Rented house

31 MHKOOP Home owners

32 MAUT1 1 car
33 MAUT2 2 cars
34 MAUT0 No car
35 MZFONDS National Health Service
36 MZPART Private health insurance
37 MINKM30 Income < 30.000
38 MINK3045 Income 30-45.000
39 MINK4575 Income 45-75.000
40 MINK7512 Income 75-122.000
41 MINK123M Income >123.000
42 MINKGEM Average income
43 MKOOPKLA Purchasing power class
44 PWAPART Contribution private third party insurance see L4
45 PWABEDR Contribution third party insurance (firms) ...
46 PWALAND Contribution third party insurance (agriculture)
47 PPERSONAUT Contribution car policies
48 PBESAUT Contribution delivery van policies
49 PMOTSCO Contribution motorcycle/scooter policies
50 PVRAAUT Contribution lorry policies
51 PAANHANG Contribution trailer policies
52 PTRACTOR Contribution tractor policies
53 PWERKT Contribution agricultural machines policies
54 PBROM Contribution moped policies
55 PLEVEN Contribution life insurances
56 PPERSONG Contribution private accident insurance policies
57 PGEZONG Contribution family accidents insurance policies
58 PWAOREG Contribution disability insurance policies
59 PBRAND Contribution fire policies
60 PZEILPL Contribution surfboard policies

61 PPLEZIER Contribution boat policies
 62 PFIETS Contribution bicycle policies
 63 PINBOED Contribution property insurance policies
 64 PBYSTAND Contribution social security insurance policies
 65 AWAPART Number of private third party insurance 1 - 12
 66 AWABEDR Number of third party insurance (firms) ...
 67 AWALAND Number of third party insurance (agriculture)
 68 APERSAUT Number of car policies
 69 ABESAUT Number of delivery van policies
 70 AMOTSCO Number of motorcycle/scooter policies
 71 AVRAAUT Number of lorry policies
 72 AAANHANG Number of trailer policies
 73 ATRACTOR Number of tractor policies
 74 AWERKT Number of agricultural machines policies
 75 ABROM Number of moped policies
 76 ALEVEN Number of life insurances
 77 APERSONG Number of private accident insurance policies
 78 AGEZONG Number of family accidents insurance policies
 79 AWAOREG Number of disability insurance policies
 80 ABRAND Number of fire policies
 81 AZEILPL Number of surfboard policies
 82 APLEZIER Number of boat policies
 83 AFIETS Number of bicycle policies
 84 AINBOED Number of property insurance policies
 85 ABYSTAND Number of social security insurance policies
 86 CARAVAN Number of mobile home policies 0 - 1

L0:

	Value	Label
1	1	High Income, expensive child
2	2	Very Important Provincials

3	3	High status seniors
4	4	Affluent senior apartments
5	5	Mixed seniors
6	6	Career and childcare
7	7	Dinki's (double income no kids)
8	8	Middle class families
9	9	Modern, complete families
10	10	Stable family
11	11	Family starters
12	12	Affluent young families
13	13	Young all american family
14	14	Junior cosmopolitan
15	15	Senior cosmopolitans
16	16	Students in apartments
17	17	Fresh masters in the city
18	18	Single youth
19	19	Suburban youth
20	20	Ethnically diverse
21	21	Young urban have-nots
22	22	Mixed apartment dwellers
23	23	Young and rising
24	24	Young, low educated
25	25	Young seniors in the city
26	26	Own home elderly
27	27	Seniors in apartments
28	28	Residential elderly
29	29	Porchless seniors: no front yard
30	30	Religious elderly singles
31	31	Low income catholics
32	32	Mixed seniors
33	33	Lower class large families
34	34	Large family, employed child
35	35	Village families
36	36	Couples with teens 'Married with children'
37	37	Mixed small town dwellers
38	38	Traditional families
39	39	Large religious families
40	40	Large family farms
41	41	Mixed rurals

L1:

- 1 20-30 years
- 2 30-40 years
- 3 40-50 years
- 4 50-60 years
- 5 60-70 years
- 6 70-80 years




L2:


- 1 Successful hedonists
- 2 Driven Growers
- 3 Average Family
- 4 Career Loners
- 5 Living well
- 6 Cruising Seniors
- 7 Retired and Religeous
- 8 Family with grown ups
- 9 Conservative families
- 10 Farmers

L3:

- 0 0%
- 1 1 - 10%
- 2 11 - 23%
- 3 24 - 36%
- 4 37 - 49%
- 5 50 - 62%
- 6 63 - 75%
- 7 76 - 88%
- 8 89 - 99%
- 9 100%


L4:

- 0 f 0
- 1 f 1  49
- 2 f 50  99
- 3 f 100  199

4 f 200  499

5 f 500  999

6 f 1000  4999

7 f 5000  9999

8 f 10.000 - 19.999

9 f 20.000 - ?