The SAS System

The FREQ Procedure

Frequency	Table 1 of Embarked by Survived			
	Controlling for Pclass=1			
		Survived		
	Embarked	0	1	Total
	C	21	53	74
	Q	1	1	2
	S	42	66	108
	Total	64	120	184
	Frequency Missing = 2			

Frequency	Table 2 of Embarked by Survived			
	Controlling for Pclass=2			
		Survived		
	Embarked	0	1	Total
	C	7	8	15
	Q	1	1	2
	S	82	74	156
	Total	90	83	173

Frequency	Table 3 of Embarked by Survived			
	Controlling for Pclass=3			
		Survived		
	Embarked	0	1	Total
	C	23	18	41
	Q	18	6	24
	S	229	61	290
	Total	270	85	355

The SAS System

The FREQ Procedure

Frequency	Table of FamilySize by Survived		
	Survived		
FamilySize	0	1	Total
1	274	130	404
2	63	76	139
3	40	53	93
4	6	21	27
5	8	3	11
6	19	3	22
7	8	4	12
8	6	0	6
Total	424	290	714

The SAS System

The FREQ Procedure

Frequency	Table of IsAlone by Survived		
	Survived		
IsAlone	0	1	Total
0	150	160	310
1	274	130	404
Total	424	290	714

The SAS System

The LOGISTIC Procedure

Model Information	
Data Set	WORK.CLEAN_TRAIN
Response Variable	Survived
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	714
Number of Observations Used	712

Response Profile		
Ordered Value	Survived	Total Frequency
1	1	288
2	0	424

Probability modeled is Survived='1'.

Note: 2 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information			
Class	Value	Design Variables	
Sex	female	1	
	male	0	
Embarked	C	1	0
	Q	0	1
	S	0	0
Pclass	1	1	0
	2	0	1
	3	0	0

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	962.904	650.464
SC	967.472	691.577
-2 Log L	960.904	632.464

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	328.4394	8	<.0001
Score	286.4666	8	<.0001
Wald	193.9872	8	<.0001

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
Sex	1	138.8087	<.0001
Age	1	25.0674	<.0001
Pclass	2	65.1445	<.0001
Embarked	2	2.5381	0.2811
FamilySize	1	8.7623	0.0031
IsAlone	1	2.6214	0.1054

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.3006	0.4732	0.4035	0.5253
Sex	female	1	2.6154	0.2220	138.8087	<.0001
Age		1	-0.0415	0.00829	25.0674	<.0001
Pclass	1	1	2.3721	0.2983	63.2392	<.0001
Pclass	2	1	1.1631	0.2503	21.5982	<.0001
Embarked	C	1	0.3854	0.2733	1.9888	0.1585
Embarked	Q	1	-0.3559	0.5538	0.4129	0.5205
FamilySize		1	-0.3186	0.1076	8.7623	0.0031
IsAlone		1	-0.4907	0.3031	2.6214	0.1054

Odds Ratio Estimates		

Effect	Point Estimate	95% Wald Confidence Limits	
Sex female vs male	13.673	8.849	21.126
Age	0.959	0.944	0.975
Pclass 1 vs 3	10.720	5.974	19.235
Pclass 2 vs 3	3.200	1.959	5.226
Embarked C vs S	1.470	0.861	2.512
Embarked Q vs S	0.701	0.237	2.074
FamilySize	0.727	0.589	0.898
IsAlone	0.612	0.338	1.109

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	85.9	Somers' D	0.718
Percent Discordant	14.0	Gamma	0.719
Percent Tied	0.1	Tau-a	0.346
Pairs	122112	c	0.859

The SAS System

The HPFOREST Procedure

Performance Information	
Execution Mode	Single-Machine
Number of Threads	4

Data Access Information			
Data	Engine	Role	Path
WORK.CLEAN_TRAIN	V9	Input	On Client

Model Information		
Parameter	Value	
Variables to Try	2	(Default)
Maximum Trees	100	
Actual Trees	100	
Inbag Fraction	0.6	(Default)
Prune Fraction	0	(Default)
Prune Threshold	0.1	(Default)
Leaf Fraction	0.00001	(Default)
Leaf Size Setting	1	(Default)
Leaf Size Used	1	
Category Bins	30	(Default)
Interval Bins	100	
Minimum Category Size	5	(Default)
Node Size	100000	(Default)
Maximum Depth	20	(Default)
Alpha	1	(Default)
Exhaustive	5000	(Default)
Rows of Sequence to Skip	5	(Default)
Split Criterion	.	Variance
Preselection Method	.	Loh
Missing Value Handling	.	Valid value

Number of Observations	

Type	N
Number of Observations Read	714
Number of Observations Used	714

Baseline Fit Statistics	
Statistic	Value
Average Square Error	0.241

Fit Statistics			
Number of Trees	Number of Leaves	Average Square Error (Train)	Average Square Error (OOB)
1	16	0.15709	0.18910
2	32	0.14511	0.16845
3	49	0.14620	0.18075
4	60	0.14749	0.18109
5	82	0.14329	0.17065
6	97	0.14206	0.16398
7	117	0.14052	0.15992
8	135	0.13900	0.15710
9	151	0.13811	0.15545
10	167	0.13751	0.15356
11	184	0.13747	0.15256
12	207	0.13699	0.15219
13	226	0.13684	0.15086
14	249	0.13639	0.15035
15	265	0.13608	0.14992
16	283	0.13641	0.15052
17	302	0.13624	0.15016
18	322	0.13558	0.14968
19	341	0.13573	0.14930
20	360	0.13546	0.14872
21	376	0.13558	0.14883
22	393	0.13556	0.14829
23	411	0.13546	0.14739
24	420	0.13549	0.14715

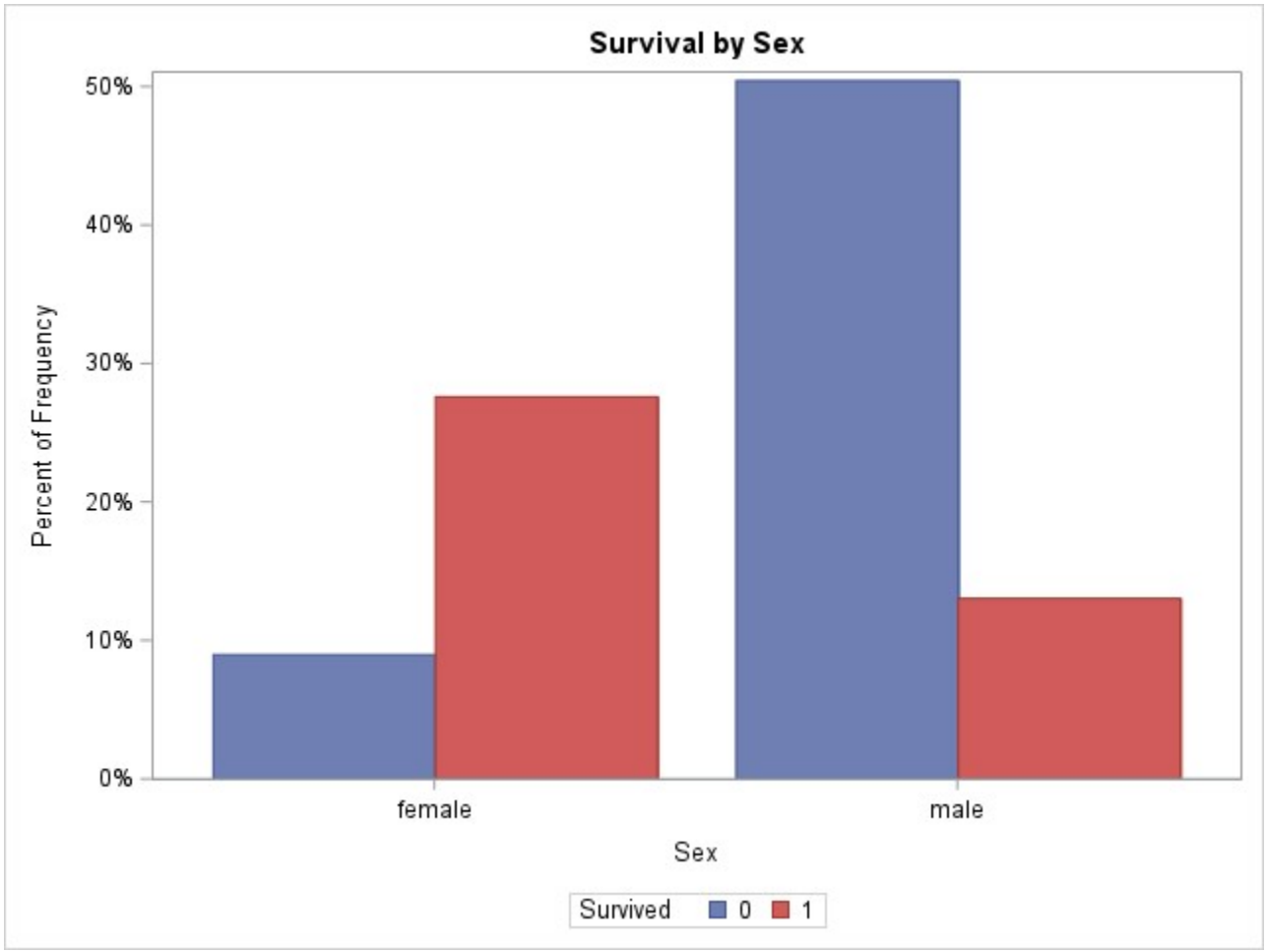
25	**439**	0.13553	0.14689
26	**455**	0.13523	0.14677
27	**469**	0.13544	0.14689
28	**487**	0.13508	0.14685
29	**509**	0.13508	0.14686
30	**529**	0.13484	0.14667
31	**541**	0.13503	0.14679
32	**559**	0.13485	0.14665
33	**573**	0.13520	0.14679
34	**586**	0.13510	0.14677
35	**609**	0.13520	0.14707
36	**632**	0.13482	0.14685
37	**651**	0.13469	0.14663
38	**668**	0.13451	0.14636
39	**689**	0.13444	0.14627
40	**708**	0.13445	0.14636
41	**726**	0.13444	0.14629
42	**746**	0.13430	0.14626
43	**757**	0.13435	0.14645
44	**774**	0.13433	0.14649
45	**787**	0.13444	0.14652
46	**812**	0.13446	0.14666
47	**828**	0.13453	0.14659
48	**843**	0.13455	0.14640
49	**861**	0.13451	0.14638
50	**881**	0.13452	0.14624
51	**888**	0.13472	0.14636
52	**908**	0.13470	0.14646
53	**928**	0.13451	0.14629
54	**939**	0.13457	0.14642
55	**957**	0.13458	0.14646
56	**976**	0.13460	0.14645
57	**992**	0.13471	0.14666
58	**1007**	0.13467	0.14644
59	**1024**	0.13462	0.14653

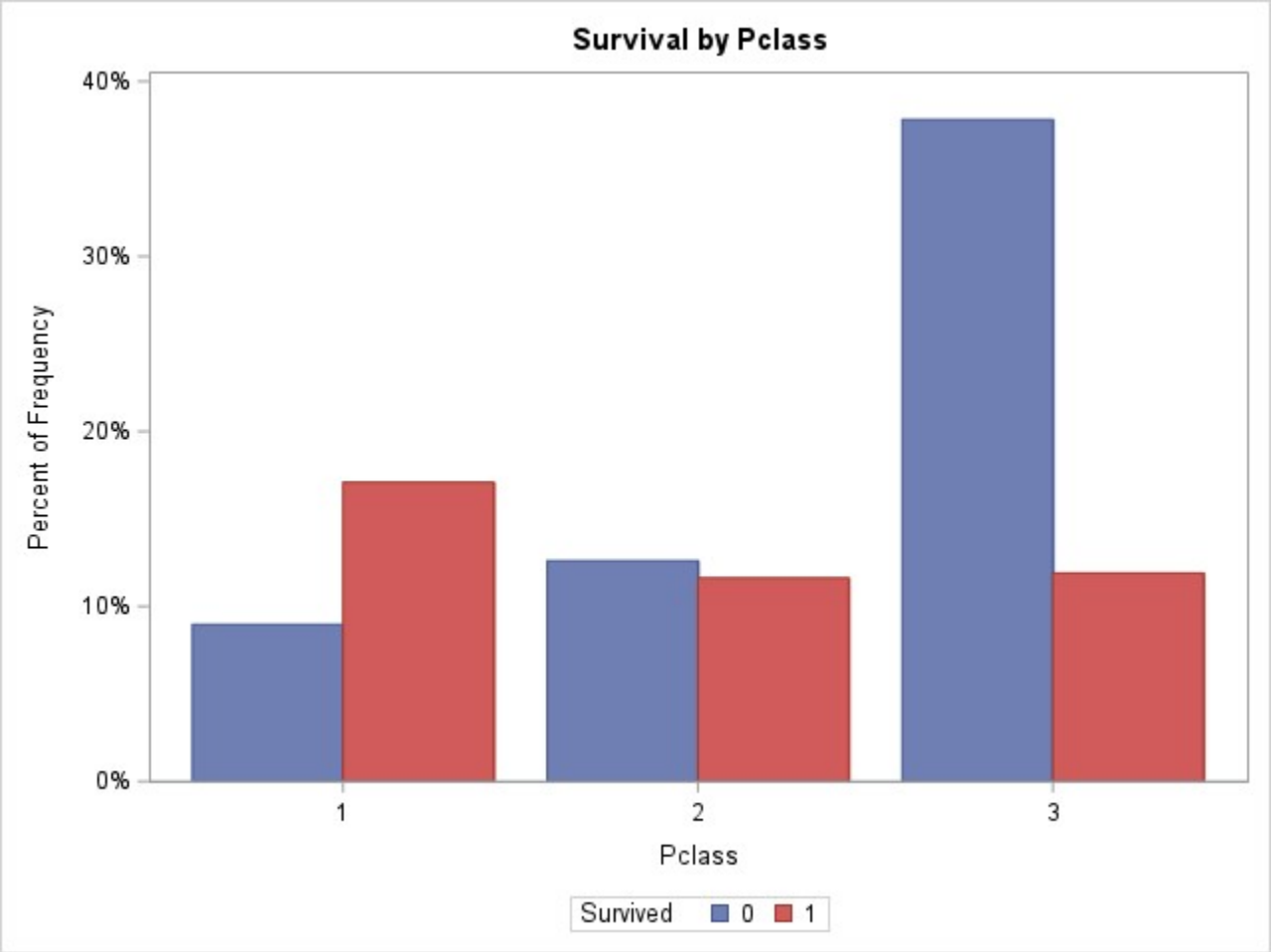
	1047	0.13454	0.14648
61	1064	0.13457	0.14644
62	1085	0.13453	0.14645
63	1106	0.13446	0.14634
64	1122	0.13452	0.14641
65	1144	0.13455	0.14636
66	1164	0.13448	0.14632
67	1177	0.13450	0.14626
68	1197	0.13419	0.14614
69	1220	0.13410	0.14595
70	1242	0.13388	0.14562
71	1265	0.13385	0.14553
72	1281	0.13372	0.14555
73	1296	0.13375	0.14545
74	1325	0.13356	0.14537
75	1338	0.13363	0.14542
76	1361	0.13362	0.14539
77	1382	0.13364	0.14538
78	1402	0.13368	0.14552
79	1417	0.13382	0.14559
80	1434	0.13388	0.14556
81	1444	0.13408	0.14574
82	1460	0.13416	0.14595
83	1480	0.13412	0.14599
84	1503	0.13417	0.14590
85	1520	0.13417	0.14578
86	1541	0.13413	0.14581
87	1558	0.13418	0.14593
88	1575	0.13415	0.14581
89	1599	0.13405	0.14580
90	1622	0.13402	0.14566
91	1634	0.13410	0.14564
92	1655	0.13406	0.14557
93	1680	0.13393	0.14561
94	1692	0.13402	0.14565

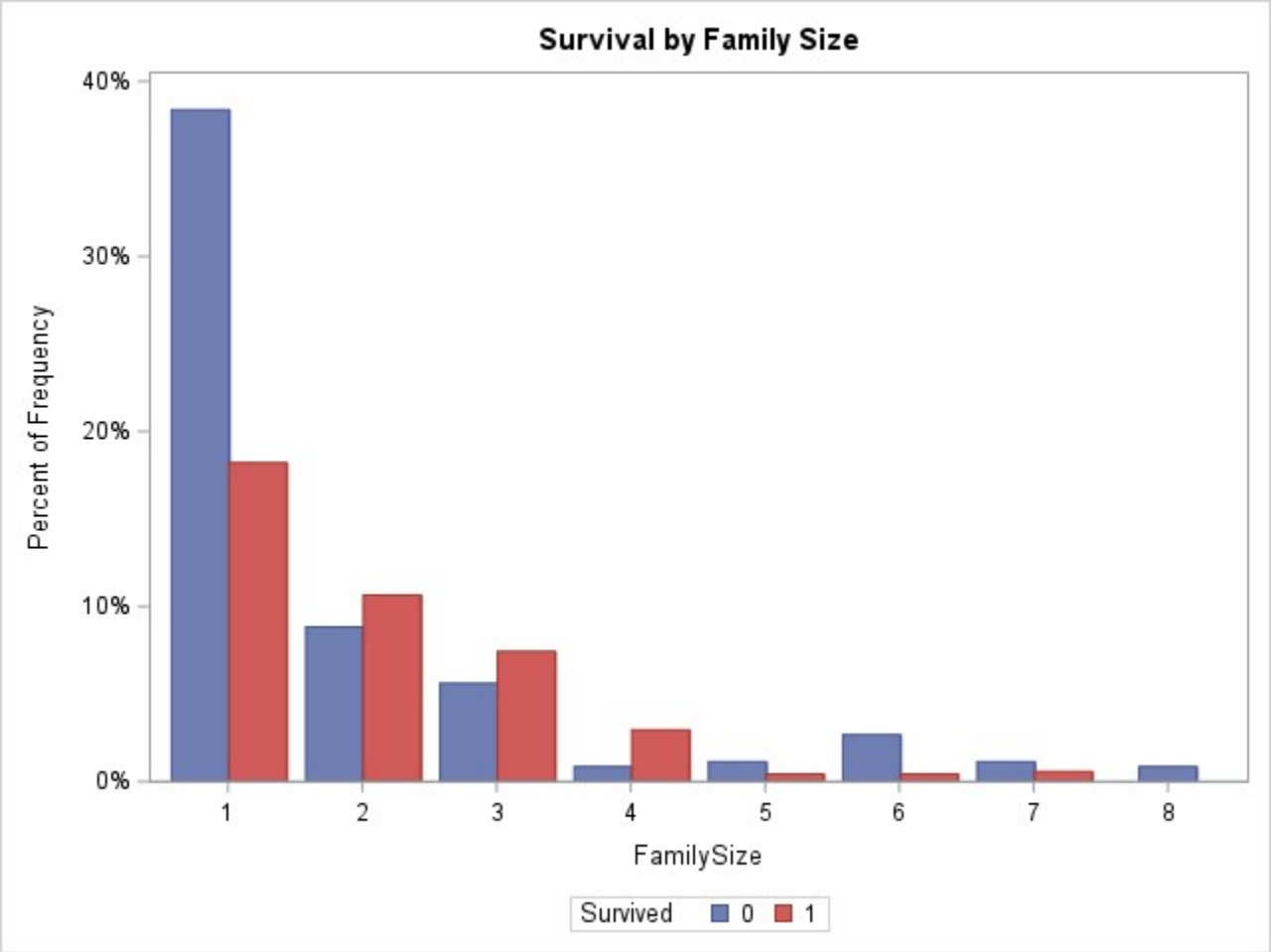
	1711	0.13410	0.14564
96	1734	0.13398	0.14547
97	1748	0.13394	0.14544
98	1772	0.13382	0.14537
99	1792	0.13378	0.14536
100	1817	0.13378	0.14542

Loss Reduction Variable Importance

Variable	Number of Rules	MSE	OOB MSE	Absolute Error	OOB Absolute Error
Sex	227	0.056540	0.05611	0.112073	0.111614
Pclass	421	0.023017	0.01939	0.043769	0.039773
FamilySize	345	0.006948	0.00267	0.013901	0.009625
IsAlone	230	0.004152	0.00246	0.006150	0.004518
Embarked	322	0.004494	0.00161	0.007543	0.004574
Age	172	0.001045	-0.00512	0.012341	0.007297


Survival by Sex
50%
40%
30%
20%
10%
0%
Percent of Frequency
female
male
Sex
Survived 0 1


Survival by Pclass
40%
30%
20%
10%
0%
Percent of Frequency
1
2
3
Pclass
Survived 0 1


Survival by Family Size
40%
30%
20%
10%
0%
Percent of Frequency
1
2
3
4
5
6
7
8
FamilySize
Survived
0
1

Survival by Family Size

The FREQ Procedure

Frequency	Table of FamilyGroup by Pclass				
		Pclass			
	FamilyGroup	1	2	3	Total
	1	86	93	225	404
	2	94	78	87	259
	5	6	2	43	51
	Total	186	173	355	714

Survival by Family Size

The FREQ Procedure

Frequency	Table of FamilySize by Pclass			
	Pclass			
FamilySize	1	2	3	Total
1	86	93	225	404
2	63	34	42	139
3	24	31	38	93
4	7	13	7	27
5	2	1	8	11
6	4	1	17	22
7	0	0	12	12
8	0	0	6	6
Total	186	173	355	714