

*Journal of Korean
Data & Information Science Society
2002, Vol. 13, No.1 pp. 1 - 16*

Implementation of On-Line Survey and Analysis System based on Database Structure

**Hee Chang Park¹⁾ • Ki Seong Nam²⁾ • Hee Jae Kim³⁾ •
Gum Min Song⁴⁾ • Ho Min Myung**

Abstract

In this paper, we suggest and implement an on-line survey and analysis system based on the database structure. We can do trends analysis in this system and reduce the number of function because we can treat basic algorithms on the database.

Keywords : , ,

1.

, WWW (World Wide Web)
가 , .
(Online survey)

WWW
(DB ; Data Base) 가

(integrated data) (minimal redundancy)
(query)

-
1. Associate Professor, Department of Statistics, Changwon National University, Changwon, Kyungnam, 641-773, Korea
E-mail : hcpark@sarim.changwon.ac.kr
 2. Lecturer, Department of Statistics, Changwon National University, Changwon, Kyungnam, 641-773, Korea
 3. Lecturer, Department of Sociology, Changwon National University, Changwon, Kyungnam, 641-773, Korea
 4. Graduate Student, Department of Statistics, Changwon National University, Changwon, Kyungnam, 641-773, Korea

(real time accessibility),
(continuous evolution),
(continuous sharing)

(, ,) 가
가
(content reference)

가

(2000)

(http://www.ahanet.co.kr)

가

Solution, Market Sight

TSS WebSurvey, Survey
, TSS WebSurvey (http://www.tssonline.com)
, 가 , 가
(Off-line)

(2001)

, ‘ (1999)
, ‘ (1999)

1999) CATS, ISP, 가

‘ (1992, 1994,
‘West (1998)
. ‘Schward(1997)

StatVillage

(1998)

2

, 3

가

, 5

, 4

2.

(query)

가

- 1) 가 가
- 2) 가
- 3) 2) 가 ()
(data mining)
- 4) ()
- 5) 가가 (questionnaire making) 가
- 6) 가 가 가
- 7) 가 ()

3.

3.1

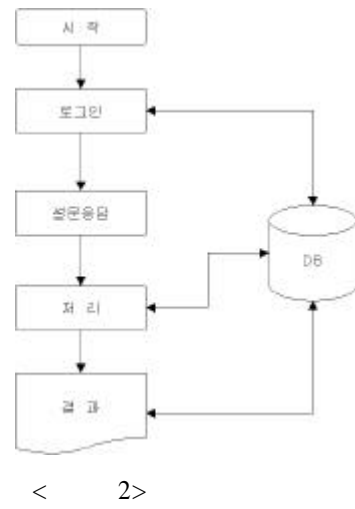
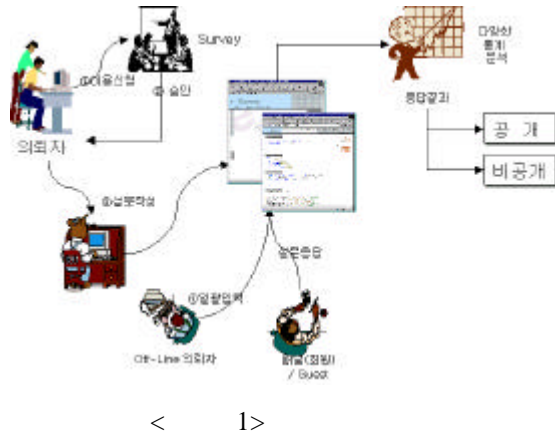
Server/ColdFusion Studio, PL/SQL, GCC, Java, HTML ColdFusion Application
 Linuxs, Windows NT, Windows 98 , Unix,
 ODBC / Oracle 7.3, 8.0, 8i / Sybase System 11 / Sybase
 Adaptive Server 11.5 / Informix 7.3 and 9 / IBM DB2 5.0 and 6.1 .
 SPSS

3.2

가

가

< 1> < 2>



3.3

< 1>

가

가

CFID CFTOKEN

(log-in)

가

.(< 2> < 3>)

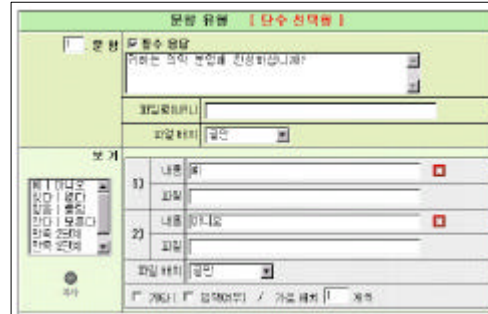
Case

가

가



< 4 >



< 5 >

Survey Data - Microsoft Internet Explorer

주소: http://cm.ftmaker.co.kr/basic_stat/all_basic_stat.cfm?requestTimeOn=327683tabid_n=3

기본 통계 진숴 보기

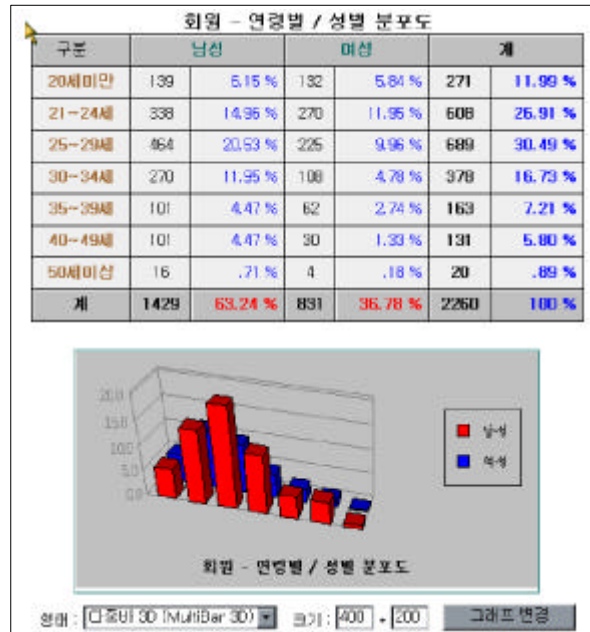
DB Query Time : 892 ms / Data Calculating Time : 41077 ms

설문 번호	1. 경험	2. 연령	3. 소용량소	4. 안화생량	5. 전취도	6. 경제사정	7. 지출	8. 피양수입	9. 행복도
자료 종류	명목척도	등간척도	명목척도	등간척도	등간척도	사양척도	등간척도	등간척도	비행척도
Sum	.0000	442,496,0000	.0000	55,450,0000	32,380,0000	.0000	1,525,316,0000	2,768,600,0000	85,424,0000
유효 응답자 (N)	11,201	11,201	11,201	11,201	11,201	11,201	11,201	11,201	11,201
a1. Arithmetic Mean	.0000	39,5054	.0000	4,949	2,9367	.0000	136,1767	247,1629	7,6432
a2. Geometric Mean	.0000	38,17100	.0000	2,9108	2,73660	.0000	128,97143	238,54401	7,58594
a3. Harmonic Mean	.0000	36,85795	.0000	1,4341	2,96707	.0000	118,36859	230,17429	7,52543
a4. Quadratic Mean	.0000	52095	.0000	1,1844	1,16886	.0000	1,17222	1,95490	28975
a5. Median	.0000	38,0000	.0000	3,0000	3,0000	.0000	138,0000	250,0000	8,0000
a6. Mode	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
a7. Quartiles	Q1 : .0000 Q2 : .0000 Q3 : .0000	Q1 : 31,0000 Q2 : 39,0000 Q3 : 48,0000	Q1 : .0000 Q2 : .0000 Q3 : .0000	Q1 : 2,0000 Q2 : 3,0000 Q3 : 4,0000	Q1 : 2,0000 Q2 : 3,0000 Q3 : 4,0000	Q1 : .0000 Q2 : .0000 Q3 : .0000	Q1 : 95,0000 Q2 : 130,0000 Q3 : 180,0000	Q1 : 200,0000 Q2 : 250,0000 Q3 : 300,0000	Q1 : 7,0000 Q2 : 8,0000 Q3 : 8,0000
a8. Trimmed Mean	.0000	38,71594	.0000	3,2442	2,84717	.0000	128,29584	238,01107	7,6275
a9. Min	.0000	24,0000	.0000	1,0000	1,0000	.0000	60,0000	190,0000	6,0000
a10. Max	.0000	58,0000	.0000	5,0000	4,0000	.0000	260,0000	490,0000	8,0000
a11. Range	.0000	34,0000	.0000	4,0000	3,0000	.0000	200,0000	290,0000	2,0000
a12. Quadrile Deviation	.0000	8,50000	.0000	1,00000	1,00000	.00000	42,50000	50,00000	5,00000
a13. Mean Deviation	.0000	8,68015	.0000	1,0234	1,1079	.0000	43,29251	51,82188	6,8413
a14. Variance	.0000	104,6487	.0000	1,34067	1,74727	.0000	2,004,99904	4,375,11268	67294
a15. Standard Deviation	.0000	10,22975	.0000	1,15783	1,32045	.0000	51,07920	66,14464	8,2041
a16. Coefficient of Variation	.0000	25,99499	.0000	26,95957	28,97734	.0000	37,48011	26,79945	12,22234
a17. Coefficient of Quartile Deviation	.0000	21,79487	.0000	33,33333	33,33333	.0000	32,66231	20,00000	8,25000
a18. Coefficient of Mean Deviation	.0000	21,86250	.0000	31,89959	28,75288	.0000	31,79141	30,32230	10,41932
a19. Skewness	.0000	21649	.0000	-15881	-14984	.0000	50836	48294	-11773
a20. Kurtosis	.0000	1,85044	.0000	2,68027	2,93307	.0000	2,95488	2,54398	2,12262

< 7 >

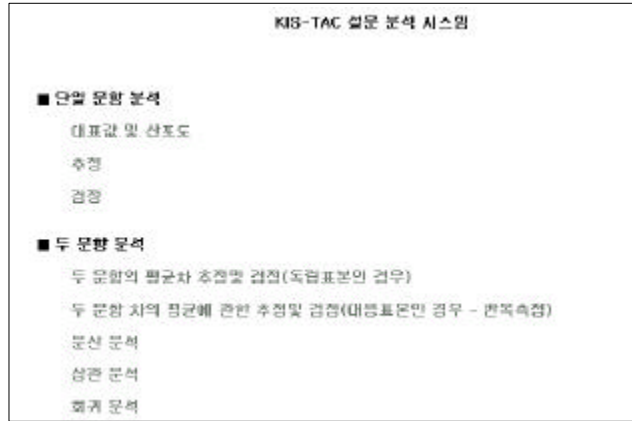


< 8 >



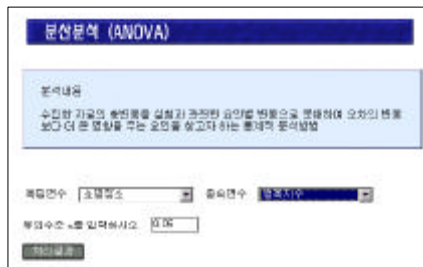
< 9 >

가



< 10>

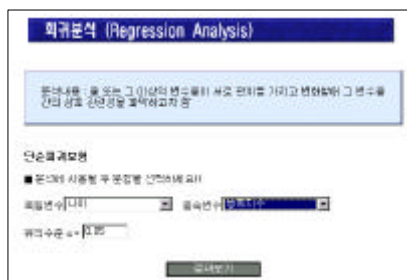
< 11> < 14>



< 11>



< 12>



< 13>



< 14>

- < 30>)

5.

, 가

OFF-Line

가

1. , , , , , (1999).
2. , , , (2001).
3. , , (1999). On-line Survey ,
4. , (1998).
, Vol. 5, 755-765.
5. , , (1992). ,
6. , , 5 , 1 , 81-92.
(1994). , , 7 ,
2 , 179-288.
7. , (1999). Window ,
8. (2000). (Internet Survey) , Internet Survey
Workshop ,
9. Dan Tapscott(1997). Net Generation, MacGraw-Hill.
10. Schwarz, C. J.(1997). StatVillage:An On-line, WWW-Accessible,
Hypothetical City Based on Real Data for Use in an Introductory Class in
Survey Sampling, Journal of Statistics Education, Vol. 5, No. 2.
<http://www.amstat.org/publications/jse>
11. SPSS(2000). SPSS Statistical Algorithms,

<http://www.spss.co.kr/tech/stat/Algorithms.htm>

12. West, R. W., Ogden, R. T. and Rossini, A. J.(1998). Statistical Tools on the World Wide Web, The American Statistician, Vol. 52, No. 3, 257- 262.

성별에 따른 신뢰도의 차에 대한 검증

귀무가설 H_0 : 성별 남성과 여성에 따라 신뢰도의 모평균이 같다.
 대립가설 H_1 : 성별 남성과 여성에 따라 신뢰도의 모평균이 다르다.

검정통계량: 1.021954
 기근적: |Z| = 2.062035
 귀무가설 H_0 를 기각한다.

결과 해석
 성별 남성과 여성에 따라 신뢰도의 모평균이 같다고 볼 수 있다.

< 15 >

분산분석 (ANOVA)

▲성별(성)에 대한 행복지수의 분산분석표

성별	평균값	자유도	평균제곱	F값
여성	2.760000	4	0.690000	0.736215
남자	19.000000	20	0.950000	
합	71.760000	24		

$H_0: \mu_1 = \mu_2 = \dots = \mu_k$
 $H_1: \mu_1 \neq \mu_2 \neq \dots \neq \mu_k$

$F_{(4, 20)} = 0.736215$
 $F_{(4, 20, 0.0500)} = 2.866109$
 $F < F_{(4, 20, 0.0500)}$ [미분포]
 귀무가설 H_0 를 기각한다.

< 16 > 1

결과 해석
 쇼핑질소에 따른 행복지수 차이는 없다고 볼 수 있다.

쇼핑질소인자 1에 대한 행복지수인 평균: 7.833333
 쇼핑질소인자 2에 대한 행복지수인 평균: 8.000000
 쇼핑질소인자 3에 대한 행복지수인 평균: 7.500000
 쇼핑질소인자 4에 대한 행복지수인 평균: 7.000000
 쇼핑질소인자 5에 대한 행복지수인 평균: 7.666667

쇼핑질소인자 1: 백화점
 쇼핑질소인자 2: 대형시점
 쇼핑질소인자 3: 할인점
 쇼핑질소인자 4: 동네수퍼
 쇼핑질소인자 5: 기타

< 17 > 2

상관분석 (Correlation Analysis)

표본상관계수 (Sample correlation coefficient)

나이(N)과 행복지수의 표본상관계수

$r = 0.187046$

결과 해석
 표본상관계수 r의 값이 1에 가까우면 양의 상관관계로 강한 양의 영향을 나타내며, -1에 가까우면 음의 상관관계로 강한 음의 영향을 나타낸다. r의 값이 0에 가까우면 두 변량에 강한 영향을 받다고 보기 어렵다. 여기서 r의 값이 0.187046은 상관관계의 어떠한 관계가 서로 밀접하게 나타낸다는 것을 뜻한다. 하지만 0에 가까운 경우라도 r의 값을 확인 할지라도 적절하거나 강등으로 관계할 수 있다.

< 18 >

모선관계수의 검정

귀무가설 $H_0: \rho = 0.000$
 대립가설 $H_1: \rho \neq 0.000000$
 검정통계량(Z) = -1.888701, 기근적(Z) = 1.960995

▲ 귀무가설 H_0 기각

결과 해석
 모집단의 내(외)국 행복지수에 대한 상관계수는 0.500000라고 볼 수 있다.

< 19 >

회귀분석 (Regression Analysis)

단순회귀모형 (Simple regression model)

독립변수(X): 내(외)국행복지수(Y): 행복지수

β_0, β_1 의 최소제곱추정값(least square estimator)

$b_0 = 6.966024$
 $b_1 = 0.017039$

추정된 회귀직선 $y = 6.966024 + 0.017039x$

< 20 > 1

ANOVA table (Analysis of variance table)

요인	제곱합	자유도	제곱평균	F0
회귀	SSR = 0.761297	1	MSR = 0.761297	MSR/MSE = 0.839853
잔차	SSE = 20.956703	23	MSE = 0.912987	
합계	SST = 21.760000	24		

추정된 회귀직선 모형
 $H_0: \beta_1 = 0$ (회귀 관계가 없다)
 $H_1: \beta_1 \neq 0$ (회귀 관계가 있다)
 $F_0 = 0.839853, F = 2.069376$
 $\therefore F_0 < F$ 이므로 귀무가설 채택

< 21 > 2

결과 해석
 종속변수의 변동은 X의 변화에 영향을 받지 않는다.

결정계수(Coefficient of determination)
 $R^2 = 0.034985$
 Adjusted R square = -0.006971

결과 해석
 종속변수의 변동이 매우 회귀직선을 사용하면 3.4% 정도를 나카려 의해 설명할 수 있다.

< 22 > 3

$\beta_0, \beta_1, E(y)$ 의 추정과 모형 및 잔차분석
 기설정된 값 입력이새요!

독립변수 나이: 일 때의 종속변수 행복지수의 추정값 $\hat{y}_0 =$

검정 방법: 양측검정 좌측검정 우측검정

결과 출력보기

< 23 > 4

회귀분석 (Regression Analysis)

모수의 추정, 모형 및 잔차

β_0, β_1 의 최소제곱추정량(least square estimator)

$b_0 = 6.966624$
 $b_1 = 0.017039$

추정된 회귀직선 $y = 6.966624 + 0.017039x$

$\text{adj} R^2 = 0.999504$

< 24 > 5

β_0 의 추정과 검정 (유의수준 $\alpha = 0.0500$)

β_0 의 신뢰구간: $[-0.021574 \leq \beta_0 \leq 0.052652]$

귀무가설 $H_0: \beta_0 = 0$, 대립가설 $H_1: \beta_0 \neq 0$

검정통계량(T) = 0.913155, 기각역(α) = 2.069376

\therefore 귀무가설 H_0 채택

결과 해석
 회귀 직선의 기울기는 0이다.

< 25 > 6

β_0 의 신뢰구간: $(-5.390223 \leq \beta_0 \leq 8.543025)$

귀무가설 $H_0: \beta_0 = 0$, 대립가설 $H_1: \beta_0 \neq 0$

검정통계량(T) = 0.145297, 기각역(α) = 2.069376

\therefore 귀무가설 H_0 기각

결과 해석
 회귀 직선의 기울기는 0이 아니다.

< 26 > 7

$E(y)$ 의 신뢰구간(est. y = 7.562814)
 $\rightarrow x = 35.000000$ 일때, $t = 7.130724 = E(y) = 7.995245$

$\alpha = 0.050000$ 일때, 귀무가설 $H_0: \beta_0 = 0.000000$, 대립가설 $H_1: \beta_0 \neq 0.000000$

검정통계량(T) = -340.760163, 기각역(α) = 2.069376

\therefore 귀무가설 H_0 기각

결과 해석
 독립변수 나이 $x = 35.000000$ 일때, 종속변수 행복지수 \hat{y}_0 는 00.0000000000가 아니다.

< 27 > 8

잔차분석

항	잔차	표준화잔차	스튜던트화된잔차
1	1.4330416	1.922266	1.541215
2	0.784430	0.921003	0.690103
3	0.147355	0.194237	0.102903
4	-0.296867	-0.396344	-0.421340
5	0.579327	0.600006	0.639404
6	1.149355	1.300796	1.263115
7	-0.443698	-0.421792	-0.433942
8	-0.682255	-0.714036	-0.729634
9	1.286359	1.352156	1.481027
10	-0.51888	-0.498704	-0.523960
11	0.587434	0.639590	0.677933
12	-0.546218	-0.696196	-0.716951

< 28 > 9

성별 별(男) 투표결과에 대한 표본

투표 결과	1	2	3	4	5	합계
1	2 33.33% 23.00% 8.00% 2.40	4 56.67% 40.00% 15.00% 2.40	2 33.33% 30.00% 8.00% 2.40	1 25.00% 10.00% 4.00% 1.00	1 33.33% 10.00% 1.00% 1.20	10 40.00%
2	4 88.89% 28.89% 18.00% 3.60	2 33.33% 13.33% 8.00% 3.60	4 88.89% 26.67% 16.00% 3.60	3 75.00% 30.00% 12.00% 3.40	2 88.89% 13.33% 3.00% 1.80	16 60.00%
합계	6 24.00%	6 24.00%	6 24.00%	4 16.00%	3 12.00%	26

< 29>

1



< 30>

2

< 1>

: Member_Gubun ()

Logical Name	Physical Name	Data Type	
-	mi_Serial	Integer	PK, Index
	mi_mgSerial	Varchar2(100)	
	mi_EMail	Integer	
가	mi_Name	Integer	

< 2>

: Member_Info ()

Logical Name	Physical Name	Data Type	
-	mi_Serial	Integer	PK, Index
-	mi_mgSerial	Integer	FK
	mi_EMail	Varchar2(20)	
	mi_Name	Varchar2(20)	
- -	mi_LogID	Varchar2(10)	Index
- -	mi_LogPW	Varchar2(10)	
가	mi_RequestDate	Date	
가	mi_ApproveDate	Date	
	mi_IdentityNo	Varchar2(14)	
CFID	mi_CurrentCFID	Varchar2(20)	Cookie
CFT OKEN	mi_CurrentCFT OKEN	Varchar2(20)	

< 3>

: Member_Survey_Item ()			
Logical Name	Physical Name	Data Type	
-	msi_Serial	Integer	PK, Index
-	msi_csitSerial	Integer	FK
-	msi_msgSerial	Integer	FK
	msi_Content	Varchar(250)	
가 -	msi_StartDate	Date	
가 -	msi_EndDate	Date	
	msi_Weight	Integer	
-가 Case	msi_RowCount	Integer	
	msi_CurrentState	Integer	0: / 1: / 2:
	msi_ResponseCount	Integer	
	msi_SelNumChar	Integer	0: / 1: / 2:
	msi_CharLength	Integer	
-	msi_Mileage	Integer	
	msi_ResultOpen	Integer	0: / 1:
	msi_LastResponse	Date	

< 4> Case

: Member_Survey_Case (Case)			
Logical Name	Physical Name	Data Type	
Case-	msc_Serial	Integer	PK, Index
-	msc_msiSerial	Integer	FK
	msc_CaseOrder	Integer	
Case 가	msc_Weight	Integer	
Case	msc_Content	Varchar(250)	
	msc_ResponseCount	Integer	

< 5>

: Member_Survey_Result ()			
Logical Name	Physical Name	Data Type	
-	msr_Serial	Integer	PK, Index
-	msr_miSerial	Integer	FK
-	msr_msiSerial	Integer	FK
Case-	msr_mscSerial	Varchar(250)	FK
	msr_ResponseDate	Date	
-	msr_ShortText	Varchar2(100)	
-	msr_LongText	Varchar2(2000)	

< 6>

: Common_Survey_Item ()			
Logical Name	Physical Name	Data Type	
-	csi_Serial	Integer	PK, Index
-	csi_csitSerial	Integer	FK
-	csi_msgSerial	Integer	FK
	csi_Content	Varchar(250)	
가 -	csi_StartDate	Date	
가 -	csi_EndDate	Date	
	csi_Weight	Integer	
-가 Case	csi_RowCount	Integer	
	csi_CurrentState	Integer	0: /1: /2:
	csi_ResponseCount	Integer	
	csi_SelNumChar	Integer	0: /1: /2:
	csi_CharLength	Integer	
-	csi_Mileage	Integer	
	csi_ResultOpen	Integer	0: /1:
	csi_LastResponse	Date	

< 7> Case

: Common_Survey_Case (Case)			
Logical Name	Physical Name	Data Type	
Case-	csc_Serial	Integer	PK, Index
-	csc_msiSerial	Integer	FK
	csc_CaseOrder	Integer	
Case 가	csc_Weight	Integer	Case 가
Case	csc_Content	Varchar(250)	
	csc_ResponseCount	Integer	

< 8>

: Common_Survey_Result ()			
Logical Name	Physical Name	Data Type	
-	csr_Serial	Integer	PK, Index
-	csr_miSerial	Integer	FK
-	csr_cslSerial	Integer	FK
-	csr_msiSerial	Integer	FK
Case-	csr_mscSerial	Varchar(250)	FK
	csr_ResponseDate	Date	
-	csr_ShortText	Varchar2(100)	
-	csr_LongText	Varchar2(2000)	

< 9>

: Survey_List ()			
Logical Name	Physical Name	Data Type	
-	sl_Serial	Integer	PK, Index
-	sl_Title	Varchar2(250)	Index
-	sl_StartDate	Date	Index
	sl_EndDate	Date	
-	sl_Intro	Varchar2(2000)	

< 10>

: Survey_Restrict ()			
Logical Name	Physical Name	Data Type	
-	sr_Serial	Integer	PK, Index
-	sr_cslSerial	Integer	FK
-	sr_msiSerial	Integer	FK
Case-	sr_mscSerial	Integer	FK

< 11>

: Survey_Item ()			
Logical Name	Physical Name	Data Type	
-	csl_Serial	Integer	PK, Index
-	csl_msiSerial	Integer	FK, Index
-	csl_csiSerial	Integer	FK
-	csl_Title	Integer	FK
- Order	csl_StartDate	Integer	