Group 7

Current Status of Smartphone Overdependency among Teenagers

201431960 Park Dong Seok
201431 Oh
201431 Lee

Contents

1. Motivation for topic selection

Currently, most people around us are really into their smart phone. Especially, teenagers can`t control use. Most of them tend to get overdependency on using smartphone. So we want to find the cause of teenager`s smartphone overdependency.

2. Data Introduction

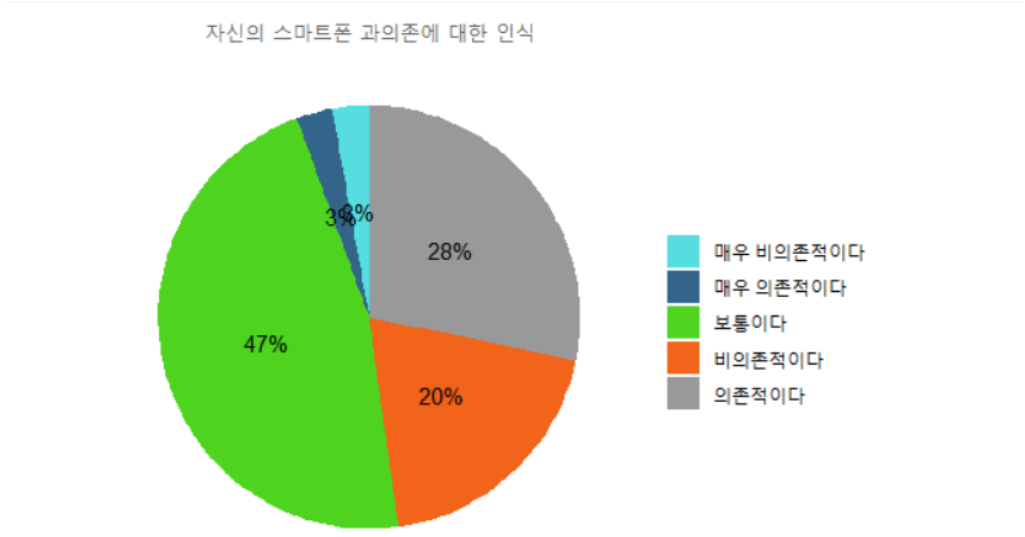
We got dataset from Microdata Integrated Service(mdis.kostat.go.kr) and we picked only teenager`s responses. So we have 56 variables, and each variable has 4886 records.

List of variables is below and Likert 8 point, 5 point, 4 point scales and T/F scales are mixed.

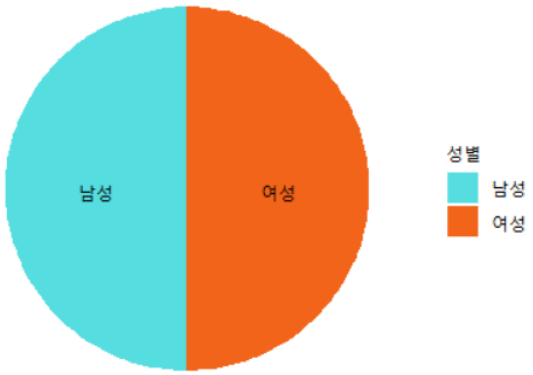
1	문1. 스마트폰 콘텐츠별 이용여부-1) 뉴스 검색
2	문1. 스마트폰 콘텐츠별 이용여부-2) 학업/업무용 검색
3	문1. 스마트폰 콘텐츠별 이용여부-3) 상품/서비스 정보 검색
4	문1. 스마트폰 콘텐츠별 이용여부-4) 교통 및 위치정보 검색
5	문1. 스마트폰 콘텐츠별 이용여부-5) 기타 일반적인 웹서핑
6	문1. 스마트폰 콘텐츠별 이용여부-6) 게임
7	문1. 스마트폰 콘텐츠별 이용여부-7) 성인용 콘텐츠
8	문1. 스마트폰 콘텐츠별 이용여부-8) 영화, TV, 동영상
9	문1. 스마트폰 콘텐츠별 이용여부-9) 음악
10	문1. 스마트폰 콘텐츠별 이용여부-10) 라디오, 팟캐스트
11	문1. 스마트폰 콘텐츠별 이용여부-11) 전자책(e-book), 웹툰, 웹소설 등
12	문1. 스마트폰 콘텐츠별 이용여부-12) 사행성 게임
13	문1. 스마트폰 콘텐츠별 이용여부-13) 이메일
14	문1. 스마트폰 콘텐츠별 이용여부-14) 메신저(카카오톡, 라인 등)
15	문1. 스마트폰 콘텐츠별 이용여부-15) SNS (카카오스토리, 밴드, 인스타그램, 페이스북, 블로그 등)
16	문1. 스마트폰 콘텐츠별 이용여부-16) 새로운 친구 만남
17	문1. 스마트폰 콘텐츠별 이용여부-17) 상품/서비스 구매
18	문1. 스마트폰 콘텐츠별 이용여부-18) 상품/서비스 판매
19	문1. 스마트폰 콘텐츠별 이용여부-19) 금융(스마트폰 뱅킹, 주식 거래 등)
20	문1. 스마트폰 콘텐츠별 이용여부-20) 생활관리(가계부 달력, 목표관리, 일기 등)
49	문2. 일, 학업, 가사 관련 목적 스마트폰 이용
50	문2. 일, 학업, 가사 이외 목적 스마트폰 이용
51	문2. 스마트폰 이용시간 과한 정도
61	문3. 청소년(만10~19세) 스마트폰 과의존 척도-1
62	문3. 청소년(만10~19세) 스마트폰 과의존 척도-2
63	문3. 청소년(만10~19세) 스마트폰 과의존 척도-3
64	문3. 청소년(만10~19세) 스마트폰 과의존 척도-4
65	문3. 청소년(만10~19세) 스마트폰 과의존 척도-5
66	문3. 청소년(만10~19세) 스마트폰 과의존 척도-6
67	문3. 청소년(만10~19세) 스마트폰 과의존 척도-7
68	문3. 청소년(만10~19세) 스마트폰 과의존 척도-8
69	문3. 청소년(만10~19세) 스마트폰 과의존 척도-9
70	문3. 청소년(만10~19세) 스마트폰 과의존 척도-10

91	문4. 나의 스마트폰 과의존에 대한 인식
96	문6. 인공지능 서비스 이용 경험
112	문9. 1599-0075 인지여부
113	문9. 1599-0075 서비스 이용여부
114	문9-1. 서비스 도움정도
115	문10. 스마트폰 과의존 예방교육 경험
116	문10-1. 스마트폰 과의존 예방교육 도움정도
118	문11. 향후 전문상담기관 이용 의향
119	문11. 향후 예방교육 이용 의향
120	문12. 우리사회 스마트폰 과의존 문제 심각성에 대한 인식
142	문16. 디지털 사용 역량 1) 온라인상에서 필요한 정보 및 콘텐츠를 잘 찾을 수 있다
143	문16. 디지털 사용 역량 2) 온라인상엥서 획득한 정보가 신뢰할만한지 판단할 수 있다
144	문16. 디지털 사용 역량 3) 온라인상에서 사회문제에 대해 인식하고 참여할 수 있다
145	문16. 디지털 사용 역량 4) 디지털 콘텐츠를 제작하고 편집할 수 있다
146	문16. 디지털 사용 역량 5) 온라인상의 개인정보보호, 프라이버시에 관련한 이슈를 파악하고 있다
147	문16. 디지털 사용 역량 6) 온라인상의 정보를 이용하여 학업, 직억 관련 활동을 하고 있다
148	문16. 가족/사회와의 관계 1) 가족들은 나를 위해 지원을 아끼지 않는다
149	문16. 가족/사회와의 관계 2) 내가 어려움에 처했을 때 도와줄 친구나 지인들이 많은 편이다
150	문16. 가족/사회와의 관계 3) 우리 사회는 개인들에게 공정한 기회와 혜택을 제공한다
151	문17. 인간관계
152	문17. 일/학업
153	문17. 건강
154	문17. 소비활동
155	문17. 여가
156	문17. 성취
157	문17. 전반
158	문18. (청소년만 응답) 고민하는 문제(1순위)
159	문18. (청소년만 응답) 고민하는 문제(2순위)
160	문19. (청소년만 응답) 주 상담자
173	성별

3. EDA



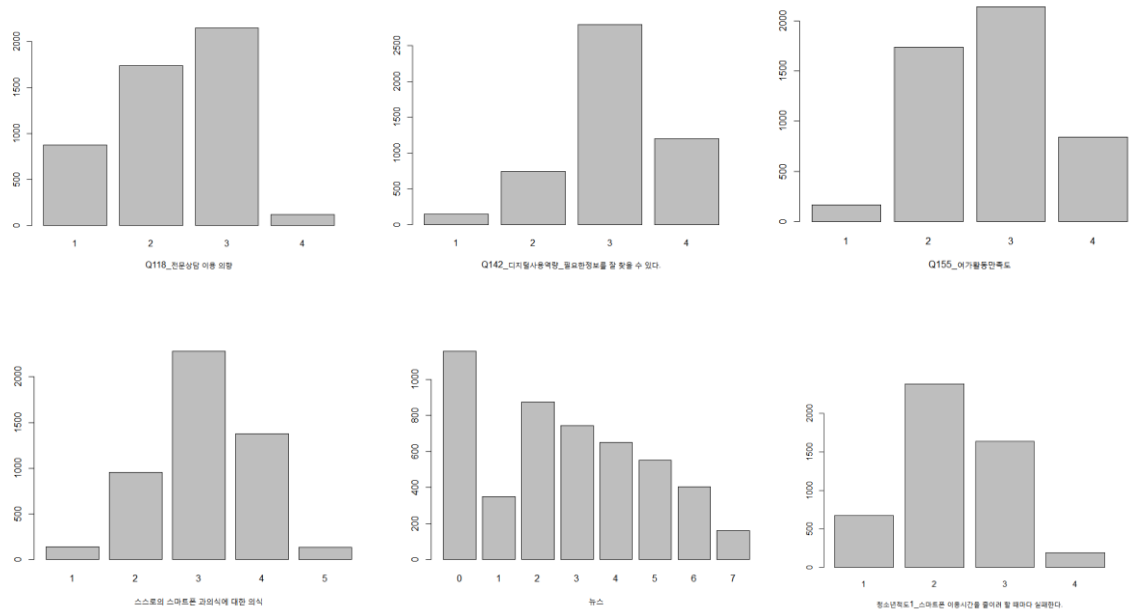
Pie Chart about Recognition of smartphone overdependence

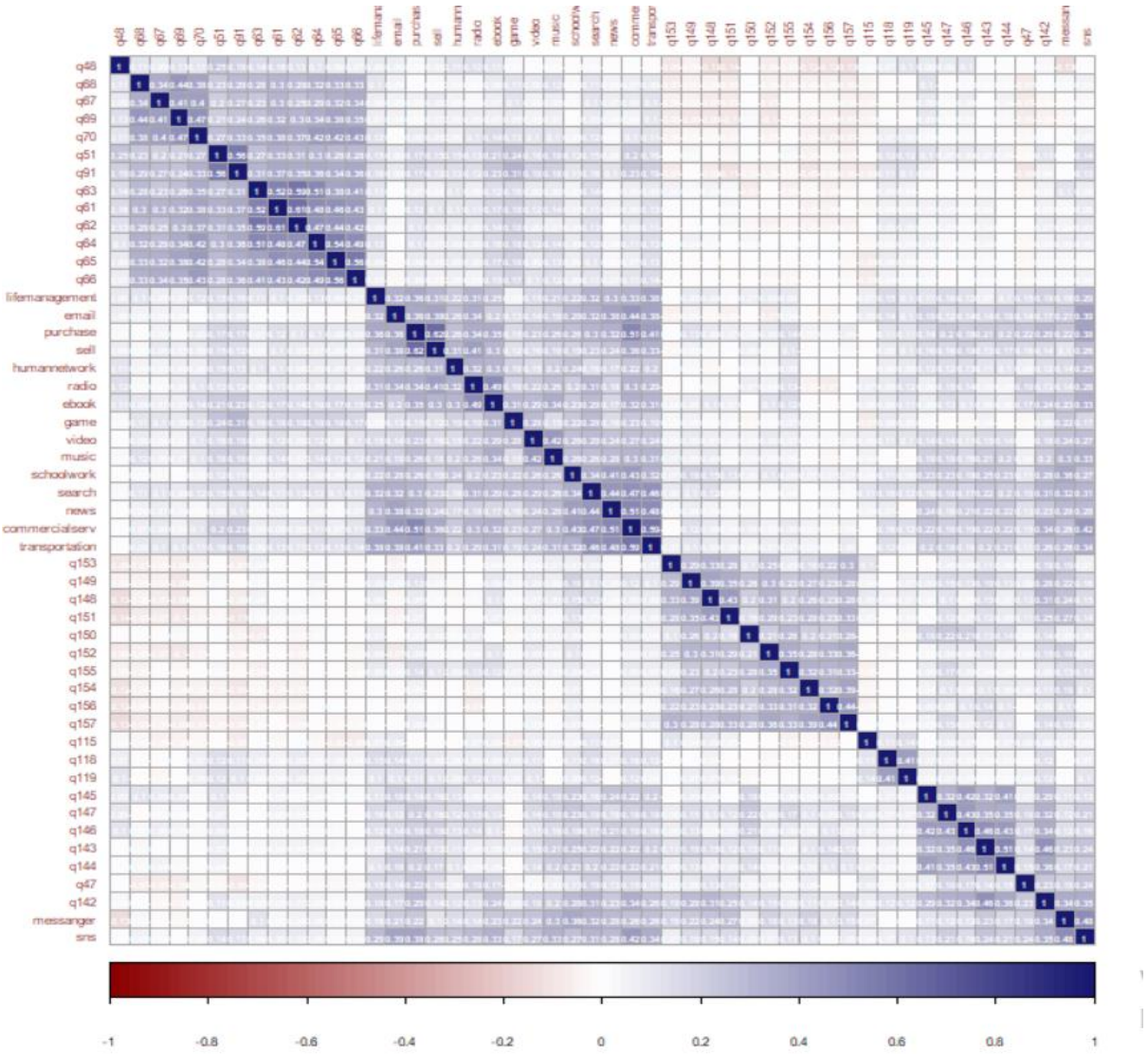


gender

Pie Chart of gender ratio

There are the bar plots of our main variables used our analysis





This is our heat map about all of variables. You can see low correlations between variables.

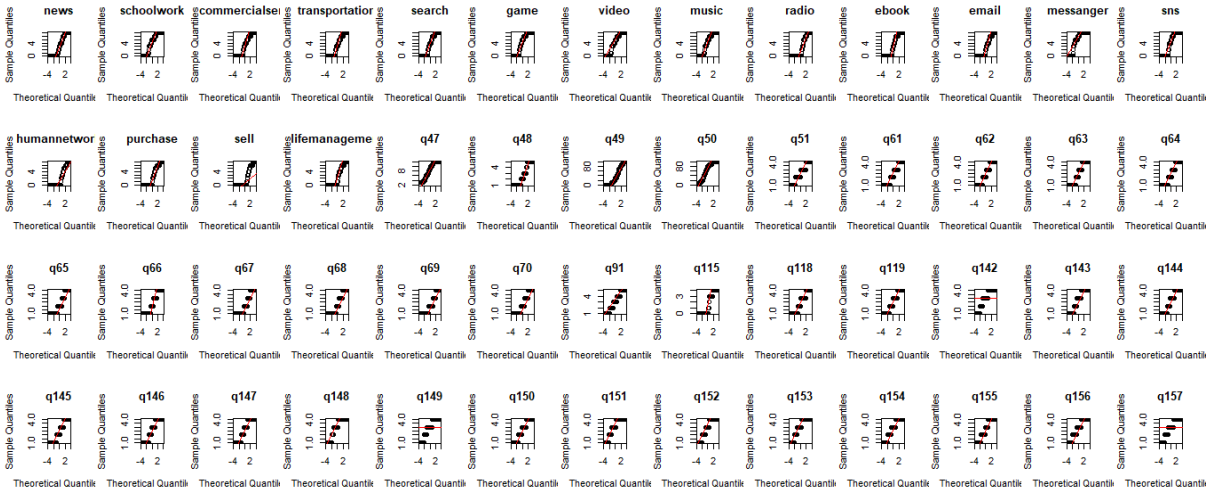
```
> cortest.bartlett(R=cor(data_int2)
$chisq
[1] 74492.22

$p.value
[1] 0
```

Bartlett result shows that P-value is under 0.05. so it is okay to do factor analysis.

<Normality Test>

Because The data set depend on Likert Scale, Can't do Nomality test exactly like p-value,

Confirm Nomarlity using qqplot, skewness and kurtosis.



If qqplot look ideal and absolute values of skewness is not over 1.5 , decide data satisfy Normality.

All date we use satisfy the Normality.

4. Factor Analysis

We tested the appropriateness of factor analysis.

```
$chisq
[1] 74492.22

$p.value
[1] 0

$df
[1] 1225
```

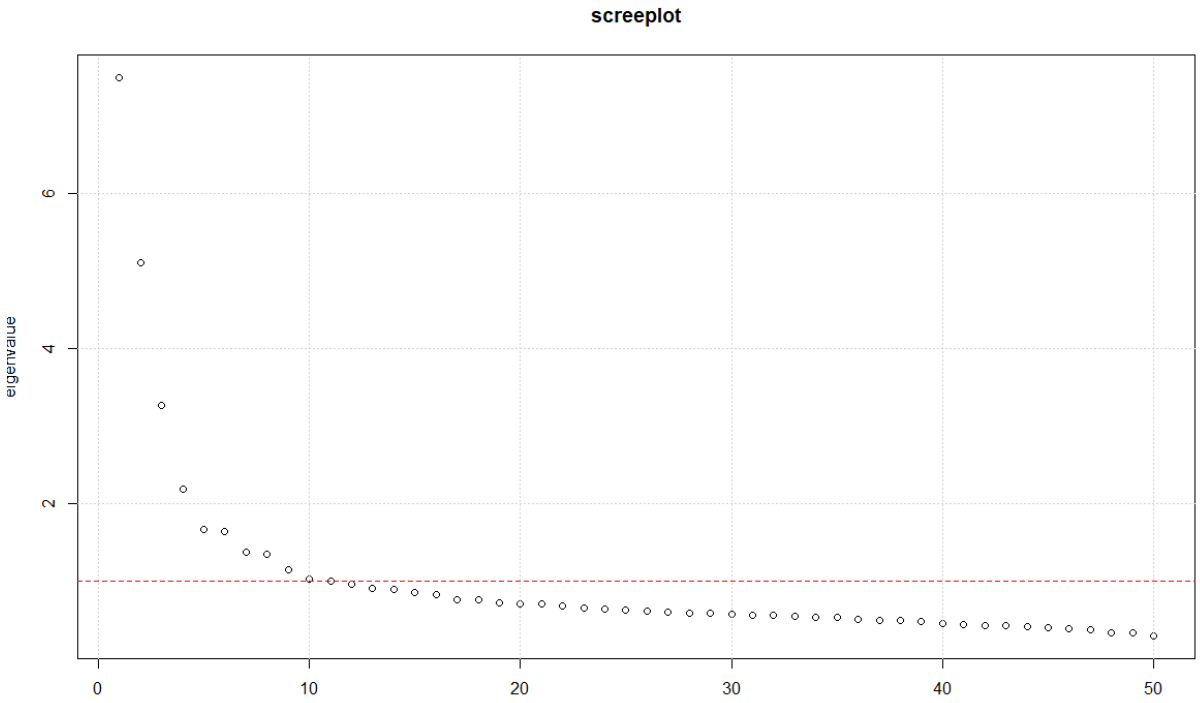
Because the dataset case are 4886 , took 100 samples and test again.

```
$chisq
[1] 1534.228

$p.value
[1] 3.298646e-09

$df
[1] 1225
```

At All dataset , MSA = 0.9 ,

```
> KMO(data_int2)
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = data_int2)
Overall MSA =  0.9
MSA for each item =
          news     schoolwork commercialserv transportation         search           game
          0.91           0.93          0.92           0.93           0.93           0.87
         video          music         radio          ebook          email      messanger
          0.90           0.90          0.87           0.91           0.91           0.87
           sns   humannetwork       purchase           sell lifemanagement            q47
          0.91           0.92          0.88           0.85           0.93           0.88
           q48            q51           q61            q62            q63            q64
          0.84           0.89          0.92           0.89           0.89           0.93
           q65            q66           q67            q68            q69            q70
          0.91           0.93          0.92           0.92           0.90           0.94
           q91           q115          q118           q119           q142           q143
          0.89           0.70          0.73           0.69           0.92           0.89
          q144           q145          q146           q147           q148           q149
          0.90           0.89          0.88           0.91           0.88           0.91
          q150           q151          q152           q153           q154           q155
          0.85           0.89          0.90           0.86           0.89           0.86
          q156           q157
```

At All dataset Scree plot , decide the number of factor is 5.



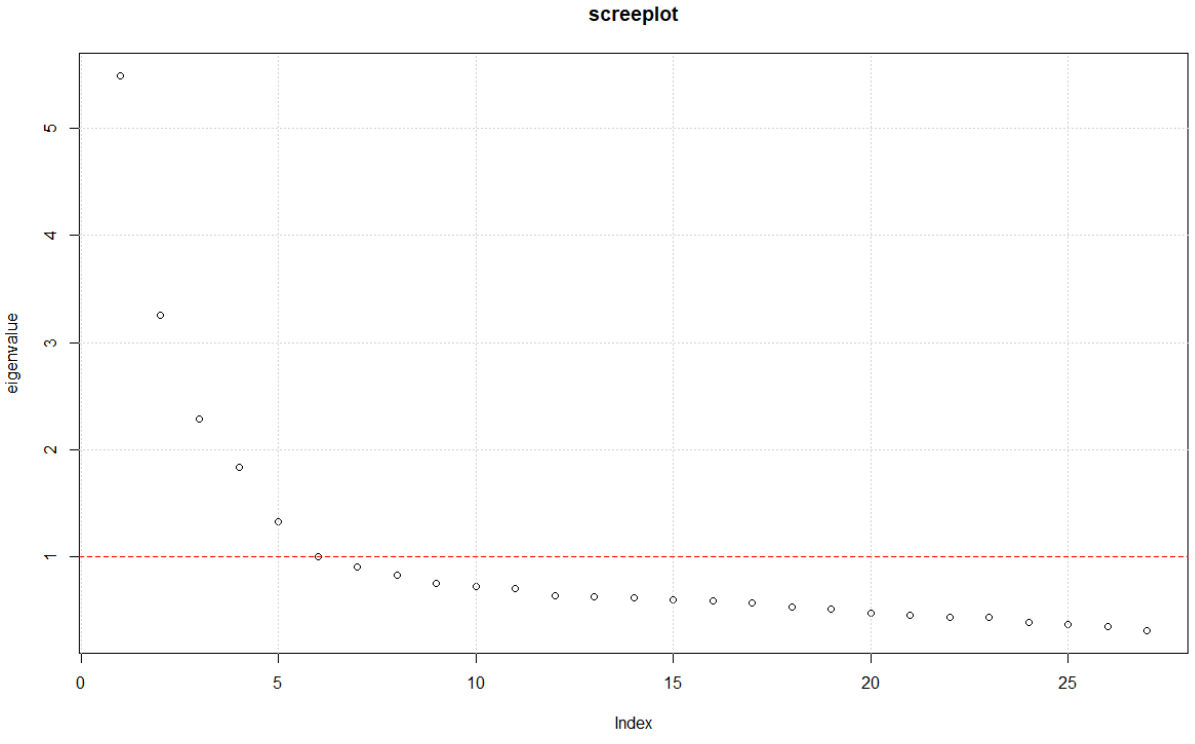
We removed variables that have little communality.

```
ff$communality
        news     schoolwork commercialserv transportation         search          radio
   0.5567321      0.4354488      0.6098075      0.5177883      0.4838527      0.5359395
       ebook          email      messanger            sns       purchase           sell
   0.4393771      0.4651678      0.5210875      0.4295287      0.5783429      0.6200737
         q61            q62            q63            q64            q65            q70
   0.6046737      0.6201864      0.5871988      0.5910434      0.5213778      0.4171056
        q143           q144           q145           q146           q148           q151
   0.5720276      0.6084046      0.5055807      0.6052063      0.4701452      0.5059928
        q152           q157           q155
   0.4861422      0.4572778      0.4493124
```

After removing, remain 27, all above 0.4

```
Overall MSA =  0.86
MSA for each item =
         news     schoolwork commercialserv transportation         search          radio
         0.90           0.92           0.90           0.92           0.93           0.83
        ebook          email      messanger            sns       purchase           sell
         0.87           0.91           0.83           0.87           0.85           0.81
          q61            q62            q63            q64            q65            q70
         0.87           0.83           0.84           0.86           0.86           0.91
         q143           q144           q145           q146           q148           q151
         0.84           0.84           0.85           0.83           0.79           0.77
         q152           q157           q155
         0.79           0.79           0.79
```

MSA = 0.86. Factor model is suitable



Other method lowers the communality. So choose Principal factor method.

Not rotate-varimax

	PC1	PC2	PC3	PC4	PC5	h2	u2	com
news	0.59	-0.17	-0.24	0.06	-0.35	0.56	0.44	2.2
schoolwork	0.56	-0.14	0.02	0.05	-0.31	0.44	0.56	1.7
commercialserv	0.73	-0.12	-0.20	-0.11	-0.13	0.61	0.39	1.3
transportation	0.68	-0.05	-0.15	-0.11	-0.14	0.52	0.48	1.2
search	0.63	-0.03	-0.11	-0.05	-0.27	0.48	0.52	1.5
radio	0.51	0.01	-0.18	-0.36	0.35	0.54	0.46	2.9
ebook	0.53	0.09	-0.02	-0.34	0.19	0.44	0.56	2.1
email	0.55	-0.16	-0.35	-0.13	-0.01	0.47	0.53	2.0
messanger	0.50	-0.17	0.26	-0.08	-0.41	0.52	0.48	2.8
sns	0.60	-0.15	0.02	-0.15	-0.14	0.43	0.57	1.4
purchase	0.65	-0.07	-0.13	-0.21	0.29	0.58	0.42	1.8
sell	0.55	-0.05	-0.23	-0.20	0.47	0.62	0.38	2.7
q61	0.29	0.70	0.15	0.06	0.04	0.60	0.40	1.5
q62	0.28	0.72	0.13	0.08	-0.02	0.62	0.38	1.4
q63	0.25	0.69	0.17	0.06	-0.13	0.59	0.41	1.5
q64	0.29	0.69	0.16	0.07	0.00	0.59	0.41	1.5
q65	0.25	0.66	0.11	0.08	0.05	0.52	0.48	1.4
q70	0.23	0.60	0.02	0.04	0.07	0.42	0.58	1.3
q143	0.44	-0.20	0.14	0.56	0.06	0.57	0.43	2.4
q144	0.42	-0.23	0.04	0.60	0.12	0.61	0.39	2.2
q145	0.38	-0.15	-0.03	0.56	0.16	0.51	0.49	2.2
q146	0.40	-0.19	0.04	0.57	0.29	0.61	0.39	2.6
q148	0.21	-0.16	0.60	-0.11	-0.15	0.47	0.53	1.6
q151	0.18	-0.17	0.65	-0.09	-0.12	0.51	0.49	1.4
q152	0.15	-0.20	0.62	-0.15	0.15	0.49	0.51	1.6
q157	0.14	-0.23	0.60	-0.07	0.13	0.46	0.54	1.5
q155	0.19	-0.19	0.46	-0.21	0.34	0.45	0.55	3.1

	PC1	PC2	PC3	PC4	PC5
SS loadings	5.49	3.25	2.29	1.84	1.32
Proportion Var	0.20	0.12	0.08	0.07	0.05
Cumulative Var	0.20	0.32	0.41	0.48	0.53
Proportion Explained	0.39	0.23	0.16	0.13	0.09
Cumulative Proportion	0.39	0.62	0.78	0.91	1.00

Rotate-varimax

Standardized loadings (pattern matrix) based upon correlation matr

	RC1	RC2	RC5	RC3	RC4	h2	u2	com
news	0.71	-0.02	0.10	-0.12	0.18	0.56	0.44	1.2
schoolwork	0.62	0.05	0.05	0.12	0.19	0.44	0.56	1.3
commercialserv	0.66	0.05	0.38	-0.02	0.14	0.61	0.39	1.7
transportation	0.62	0.10	0.34	0.01	0.12	0.52	0.48	1.7
search	0.65	0.13	0.18	0.00	0.10	0.48	0.52	1.3
radio	0.19	0.08	0.70	0.02	-0.03	0.54	0.46	1.2
ebook	0.27	0.19	0.55	0.14	-0.06	0.44	0.56	1.9
email	0.49	-0.07	0.43	-0.17	0.11	0.47	0.53	2.4
messanger	0.62	0.03	-0.06	0.35	0.04	0.52	0.48	1.6
sns	0.56	0.03	0.28	0.18	0.08	0.43	0.57	1.8
purchase	0.32	0.07	0.66	0.07	0.15	0.58	0.42	1.6
sell	0.15	0.04	0.75	-0.03	0.17	0.62	0.38	1.2
q61	0.03	0.77	0.08	0.01	0.02	0.60	0.40	1.0
q62	0.06	0.78	0.03	-0.02	0.01	0.62	0.38	1.0
q63	0.11	0.76	-0.06	0.01	-0.04	0.59	0.41	1.1
q64	0.05	0.77	0.04	0.02	0.02	0.59	0.41	1.0
q65	0.00	0.72	0.07	-0.02	0.03	0.52	0.48	1.0
q70	0.01	0.63	0.12	-0.09	0.01	0.42	0.58	1.1
q143	0.23	0.03	-0.02	0.15	0.70	0.57	0.43	1.3
q144	0.19	-0.02	0.02	0.05	0.75	0.61	0.39	1.1
q145	0.14	0.02	0.06	-0.03	0.69	0.51	0.49	1.1
q146	0.06	0.01	0.15	0.06	0.76	0.61	0.39	1.1
q148	0.20	0.02	-0.10	0.65	0.00	0.47	0.53	1.2
q151	0.16	0.02	-0.11	0.68	0.01	0.51	0.49	1.2
q152	-0.03	-0.03	0.09	0.69	0.04	0.49	0.51	1.1
q157	-0.02	-0.06	0.04	0.66	0.10	0.46	0.54	1.1
q155	-0.10	-0.05	0.31	0.58	0.05	0.45	0.55	1.6

	RC1	RC2	RC5	RC3	RC4
SS loadings	3.51	3.38	2.57	2.41	2.32
Proportion Var	0.13	0.13	0.10	0.09	0.09
Cumulative Var	0.13	0.26	0.35	0.44	0.53
Proportion Explained	0.25	0.24	0.18	0.17	0.16
Cumulative Proportion	0.25	0.49	0.67	0.84	1.00

Mean item complexity = 1.3
Test of the hypothesis that 5 components are sufficient.

The rotation method made the distinction between factors more appropriate.

The cumulative ratio explained by the factors is 52%.

We conclude that

F1: Use of information and communication

F2: Scale of smartphone overdependence

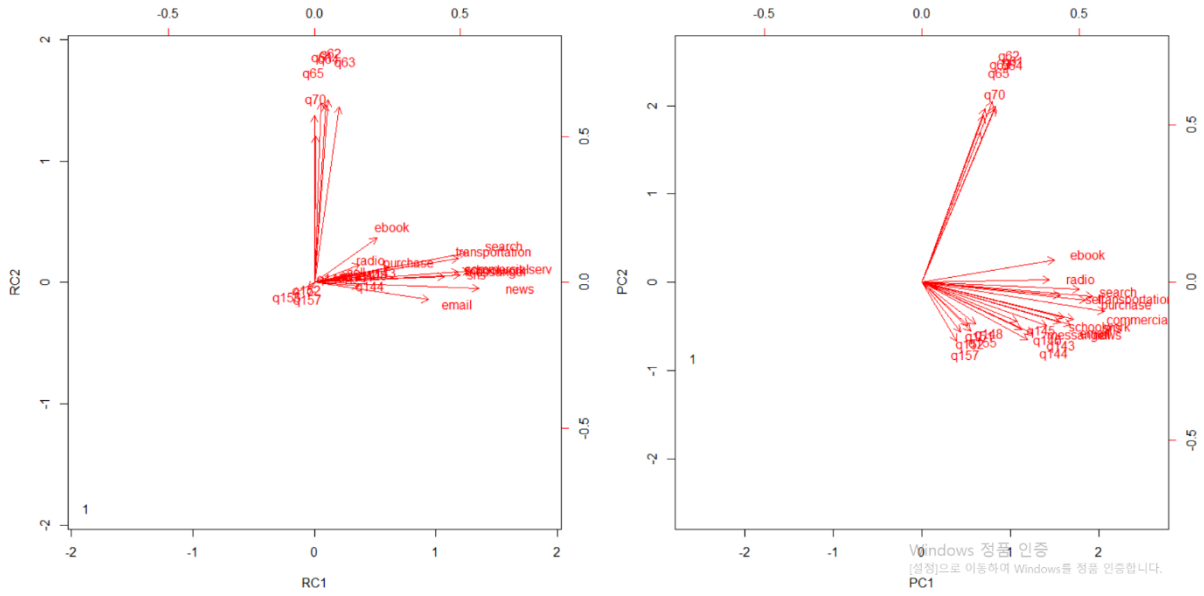
F3: Special uses such as online transactions, media

F4: psychological satisfaction

F5: Online knowledge level or ability to use information

Through Cronbach Alpha, It can be seen that each factor is well bound.

```
> psych::alpha(F1)$total[1]
 raw_alpha
  0.818449
> psych::alpha(F2)$total[1]
 raw_alpha
 0.8735507
> psych::alpha(F3)$total[1]
 raw_alpha
 0.6566015
> psych::alpha(F4)$total[1]
 raw_alpha
 0.7208613
> psych::alpha(F5)$total[1]
 raw_alpha
 0.7668005
```



At the Biplot, left thing is varimax.

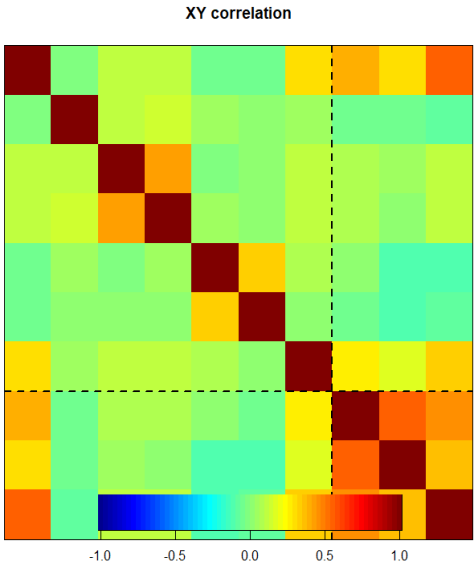
After rotation, the factors are more clearly distinguished.

5. Canonical Correlation Analysis(CCA)

Before canonical correlation analysis, Group X of independent variables is consists of 2 Derived variables and Q91 and Group Y of dependent variables one derived variables and Q51,Q115,Q118,Q119,Q151,Q157. We used canonical correlation analysis between them.

We made 3 derived variables. With 1 in the Y group is mean of Q61,62,63,65 and 66. With 2 is mean of Q68,69 and 70 and xx in the X group is mean of game, video, music and ebook.

This is Sample correlation matrix of the groups.



	Canonical Correlation	Adjusted Canonical Correlation	Approximate Standard Error	Squared Canonical Correlation
1	0,624355	0,623585	0,008730	0,389819
2	0,145619	.	0,014004	0,021205
3	0,055117	.	0,014264	0,003038

Test of H0: The canonical correlations in the current row and all that follow are zero

Likelihood Ratio	Approximate F Value	Num DF	Den DF	Pr > F
0,59542793	131,94	21	14002	<,0001
0,97582150	10,01	12	9754	<,0001
0,99696210	2,97	5	4878	0,0110

As a result, we can see only two canonical correlation coefficients are meaningful. A first canonical correlation coefficient is 0.624 and a second is 0.146

Multivariate Statistics and F Approximations					
S=3 M=1.5 N=2437					
Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.59542793	131.94	21	14002	<.0001
Pillai's Trace	0.41406172	111.58	21	14634	<.0001
Hotelling-Lawley Trace	0.66356901	154.04	21	10189	<.0001
Roy's Greatest Root	0.63885744	445.19	7	4878	<.0001
NOTE: F Statistic for Roy's Greatest Root is an upper bound.					

Even though Wilks` lamda value is a little bit big, But P-value is smaller then 0.05. it means that we can do canonical analysis.

The CANCORR Procedure

Canonical Correlation Analysis

Standardized Canonical Coefficients for the VAR Variables			
	V1	V2	V3
XX	0.3371	0.1607	-0.2343
q51	0.8100	0.0482	0.0215
q115	-0.1267	0.0957	-0.1894
q118	0.0545	0.0863	0.3112
q119	0.0264	0.1060	0.4893
q151	-0.1214	0.8999	-0.3800
q157	-0.0677	0.1109	0.7354

Sample canonical variables of X group are below

$\widehat{v1}$ = 0.3371xx+ 0.8100q51 -0.1267q115 +0.0545q118 +0.0264q119 -0.1214q151 -0.0677q154

$\widehat{v2}$ = 0.1607xx+ 0.0482q51 +0957q115 +0.0863q118 +0.1060q119 -0.8999q151 -1109q154

Standardized Canonical Coefficients for the WITH Variables			
	W1	W2	W3
q91	0.8204	-0.2978	0.7198
with1	0.2400	1.1686	-0.4133
with2	0.0978	-0.9131	-0.7945

Sample canonical variables of Y group are below

$\widehat{w1}$ = 0.8204q91+ 0.2400with1 -0.0978with2

$\widehat{w2}$ = -0.2978q91+ 1.1686with1 -0.9131with2

The CANCORR Procedure

Canonical Structure

Correlations Between the VAR Variables and Their Canonical Variables			
	V1	V2	V3
XX	0.5645	0.2560	-0.1607
q51	0.9266	0.0811	0.0276
q115	-0.1307	0.1745	-0.1100
q118	0.1831	0.1474	0.4805
q119	0.1578	0.2129	0.5733
q151	-0.1550	0.9526	-0.1507
q157	-0.1435	0.4124	0.6167

First, as seeing canonical variables of independent variables group, V1 has high correlation with Q51. So we can say V1 means how long teenagers use their smartphone. We can see V2 has higher correlation with Q151 so we can say V2 means satisfaction about their relationship.

Correlations Between the WITH Variables and Their Canonical Variables			
	W1	W2	W3
q91	0,9629	-0,1108	0,2459
with1	0,6591	0,5326	-0,5310
with2	0,5295	-0,3774	-0,7597

As seeing canonical variables of dependent variables group, W1 has strong correlation with Q91. W2 has strong correlation with with1 and minus with2.

Correlations Between the VAR Variables and the Canonical Variables of the WITH Variables			
	W1	W2	W3
XX	0,3525	0,0373	-0,0089
q51	0,5785	0,0118	0,0015
q115	-0,0816	0,0254	-0,0061
q118	0,1143	0,0215	0,0265
q119	0,0985	0,0310	0,0316
q151	-0,0968	0,1387	-0,0083
q157	-0,0896	0,0601	0,0340

Correlations Between the WITH Variables and the Canonical Variables of the VAR Variables			
	V1	V2	V3
q91	0,6012	-0,0161	0,0136
with1	0,4115	0,0776	-0,0293
with2	0,3306	-0,0550	-0,0419

V1 has correlation with Q51 and xx of variables Y.

W1 has correlation with Q91.

V2 and W2 have low correlation with all of their variables. So we decided not to choose them.

Standardized Variance of the VAR Variables Explained by					
	Their Own Canonical Variables			The Opposite Canonical Variables	
Canonical Variable Number	Proportion	Cumulative Proportion	Canonical R-Square	Proportion	Cumulative Proportion
1	0,1819	0,1819	0,3498	0,0636	0,0636
2	0,2015	0,3833	0,0207	0,0042	0,0678
3	0,1648	0,5481	0,0029	0,0005	0,0683

According to explanation ratio, V1 explains about 18% of variance of variables X. W1 explains about 55% of variance of variables Y

The result of this analysis is that only how long teenagers use their smartphone can affect their self-consciousness of their smartphone overdependency

6. Classification · Discrimination Analysis

Set Q91(Perception of smartphone overdependency) as a criterion. If the value of Q91 is 1 or 2, set to 0, If the value is 3, set to 1. If the value is 4 or 5, set to 2. 0 means smartphone dependency is low. 1 means smartphone dependency is normal. 2 means smartphone dependency is high. Set Q61, Q62, Q63, Q64, Q65, Q70 as independent variables.

6.1 Test of Homogeneity of Covariance Matrices

```
        Box's M-test for Homogeneity of Covariance Matrices

data:  data3
Chi-Sq (approx.) = 424.44, df = 42, p-value < 2.2e-16
```

Use Box's M-test to select the appropriate analysis method. P-value is smaller than significance level 0.05, null hypothesis H0 cannot be rejected. So, quadratic discriminant function is selected as an analysis method.

6.2 Classification Analysis by Quadratic Discriminant Function

Number of Observations and Percent Classified into y				
From y	1	2	3	Total
1	815 74.63	180 16.48	97 8.88	1092 100.00
2	1090 47.83	594 26.06	595 26.11	2279 100.00
3	352 23.23	199 13.14	964 63.63	1515 100.00
Total	2257 46.19	973 19.91	1656 33.89	4886 100.00
Priors	0.33333	0.33333	0.33333	

Error Count Estimates for y				
	1	2	3	Total
Rate	0.2537	0.7394	0.3637	0.4522
Priors	0.3333	0.3333	0.3333	

Group 1 is the group of people who think their smartphone overdependency level is low. Group 2 is the group of people who think their smartphone overdependency level is normal. Group 3 is the group of people who think their smartphone overdependency level is high. Misclassification rate of Group 1, Group 2, Group 3 is 25.37%, 73.94%, and 36.37%. Simple misclassification rate is 51.43%(2513/4886). If the posterior probability is same, misclassification rate is 45.22%.

6.2.1 Classification Analysis by Quadratic Discriminant Function – Cross-validate method

The DISCRIM Procedure
Classification Summary for Calibration Data: WORK.FOURS
Cross-validation Summary using Quadratic Discriminant Function

Number of Observations and Percent Classified into y

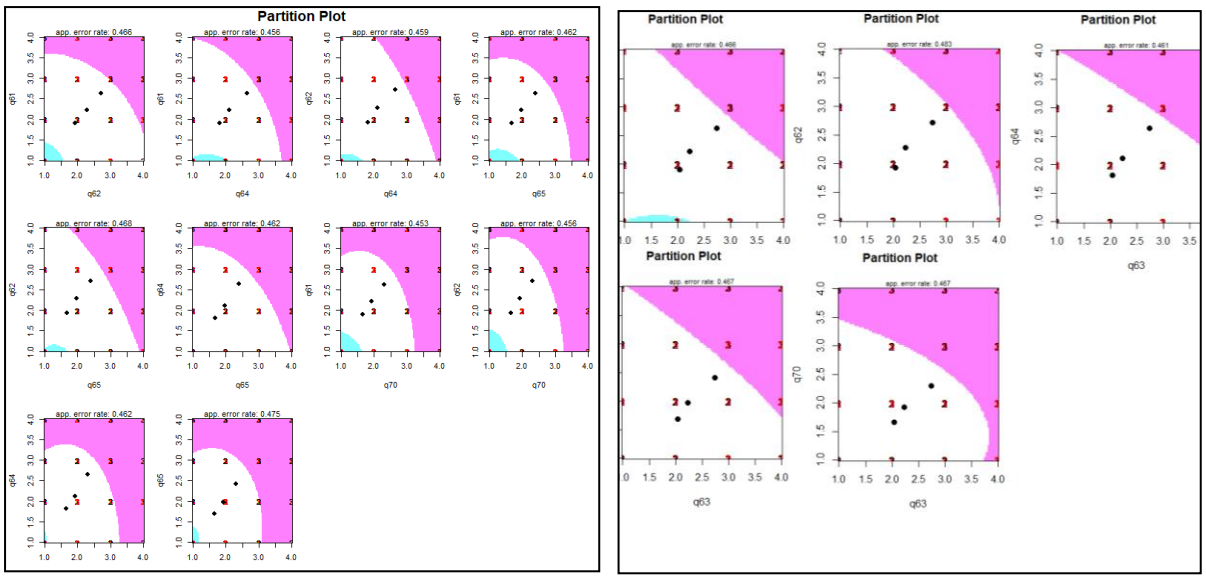
From y	1	2	3	Total
1	806 73.81	189 17.31	97 8.88	1092 100.00
2	1135 49.80	543 23.83	601 26.37	2279 100.00
3	352 23.23	205 13.53	958 63.23	1515 100.00
Total	2293 46.93	937 19.18	1656 33.89	4886 100.00
Priors	0.33333	0.33333	0.33333	

Error Count Estimates for y

	1	2	3	Total
Rate	0.2619	0.7617	0.3677	0.4638
Priors	0.3333	0.3333	0.3333	

Misclassification rate of Group 1, Group 2, Group 3 is 26.19%, 76.17%, 36.77%. Simple misclassification rate is 52.78%(2579/4886). If the posterior probability is same, misclassification rate is 46.38%. Misclassification rate of cross-validate method is little bit higher than quadratic discriminant function. Group 2 has a high misclassification rate because respondents answered Q91 as 3(normal) when they were not recognized their overdependency level exactly. As a result, over 50% of respondents are misclassified, and 75% of people who answered Q91 as 3 have differences between estimated smartphone overdependency level and their perception.

6.2.2 Partimat plots of Classification Analysis



Black characters are well-classified observations, and red characters are misclassified observations.

6.3 Logistic Classification Analysis (1)

Set Q91(Perception of smartphone overdependency) as a criterion. If the value of Q91 is 1 or 2, set to 0. If the value is 3 or 4 or 5, set to 1. Set Q61, Q62, Q63, Q64, Q65, Q70 as independent variables. 0 means smartphone dependency of respondent is below average, and 1 means smartphone dependency of respondent is over average. Variables are selected by stepwise selection, and all variables are significant.

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.275703   0.021604  12.762  < 2e-16 ***
q61          0.064286   0.010210   6.296 3.31e-10 ***
q62          0.053756   0.009345   5.752 9.35e-09 ***
q63         -0.020117   0.008517  -2.362 0.018216 *
q64          0.043483   0.008800   4.941 8.02e-07 ***
q65          0.032326   0.008798   3.674 0.000241 ***
q70          0.057536   0.008901   6.464 1.12e-10 ***
```

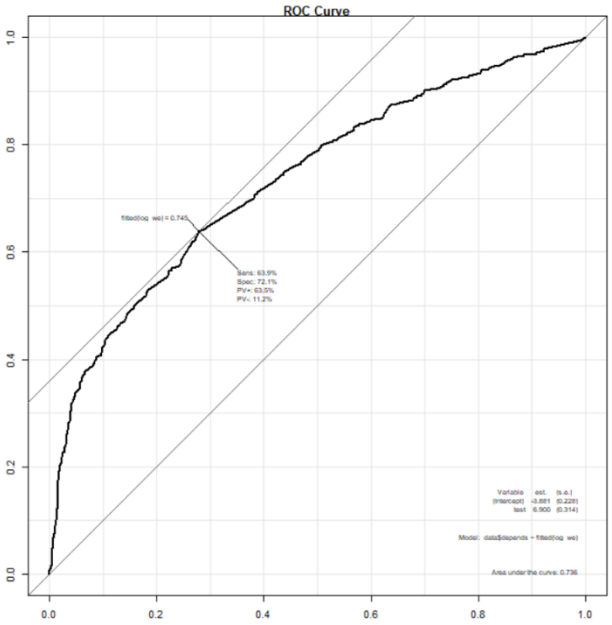
Logistic discriminant function :

$$\ln\left[\frac{f_1(x)}{f_2(x)}\right] = 0.275703 + 0.064286q61 + 0.053756q62 - 0.020117q63 + 0.043483q64 + 0.032326q65 + 0.057536q70$$

Odds ratio is exp(coefficients). If respondent failed to reduce smartphone use, overdependence odds increase by 1.066 times. If respondent cannot adjust smartphone use, overdependence odds increase by 1.055 times. If respondent cannot observe smartphone usage time, overdependence odds decrease by 0.98 times. If respondent cannot concentrate on his/her work when smartphone is next to him/her, overdependence odds increase 1.044 times. If respondent keep thinking smartphone, overdependence odds increase 1.032 times. If smartphone makes respondent difficult to study, overdependence odds increase 1.059 times.

Logistic Classification	Predict 0	Predict 1	Sum
Actual 0	9	1083	1092
Actual 1	17	3777	3794
Sum	26	4860	4886

Classification Matrix by Logistic discriminant function. Simple misclassification rate is 22.51%. Sensitivity is 64.9% and specificity is 72.1%.



ROC Curve by Logistic Classification function

The AUC value is 0.736, which is acceptable.

6.4 Logistic Classification Analysis (2)

Set Q91(Perception of smartphone overdependence) as a criterion. If the value of Q91 is 1 or 2, set to 0 . If the value is 3 or 4 or 5, set to 1. Set game, purchase, ebook, music, radio as independent variables. 0 means smartphone dependence of respondent is below average, and 1 means smartphone dependence of respondent is over average. Variables are selected by stepwise selection, and all variables are significant.

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.412293   0.017035  24.203  < 2e-16 ***
game         0.058373   0.002941  19.851  < 2e-16 ***
purchase     0.021668   0.002889   7.500 7.52e-14 ***
ebook        0.014819   0.002794   5.304 1.18e-07 ***
music        0.009674   0.003023   3.200  0.00138 **
radio       -0.007072   0.003231  -2.189  0.02863 *
```

Logistic discriminant function :

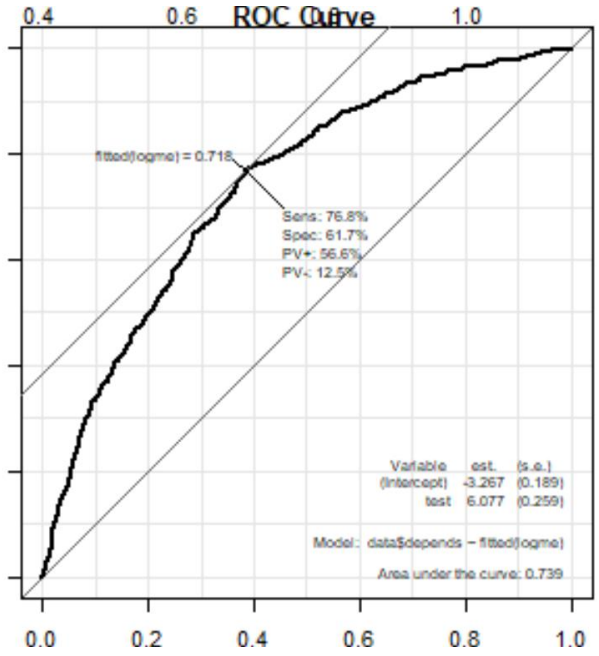
$$\ln\left[\frac{f_1(x)}{f_2(x)}\right] = 0.412293 + 0.058373game + 0.021668purchase + 0.014819ebook + 0.009674music - 0.007072radio$$

Odds ratio is exp(coefficients). If respondent plays games long time on smartphone, overdependence odds increase by 1.06 times. If respondent purchases products many times on smartphone, overdependence odds increase by 1.002 times. If respondent read e-books long time on smartphone, overdependence odds increase by 1.014 times, If respondent listen to music many times on smartphone, overdependence odds increase 1.01 times. If respondent don't listen

to radio on smartphone, overdependence odds decrease 0.99 times.

Logistic Classification	Predict 0	Predict 1	Sum
Actual 0	214	878	1092
Actual 1	135	3659	3794
Sum	349	4537	4886

Classification Matrix by Logistic discriminant function. Simple misclassification rate is 20.73%. Sensitivity is 76.8% and specificity is 61.7%.



ROC Curve by Logistic Classification function

The AUC value is 0.739, which is acceptable.

7. Conclusion

Before the analysis, we chose smartphone usage time, perception of smartphone overdependency, mentality such as life fulfillment as factors of teenagers' smartphone overdependency. After the analysis, teenagers who are using smartphone long time have high smartphone overdependency level. Other factors do not have significant influence on smartphone overdependency.

Many teenagers are not recognized how much they depend on smartphone exactly and they are not interested in smartphone overdependence Therefore, it will be necessary to promote smartphone overdependency counseling services such as 'Smart Rest Center' and expand smartphone overdependency education so that teenagers can be interested in smartphone dependency.