정량적 데이터분석과 연구 HW1 (2019/10/9)

2019521110 김진솔

1. 연구방법 및 결과
   1. 자료의 수집 및 표본의 특성
      1. 자료의 수집

본 연구의 실증적인 분석을 위한 데이터 수집을 위해 설문조사를 실시하였다. 본 설문조사 기간은 2019년 9월 2일부터 10월 7일까지 약 4주간이며, 설문대상은 알라딘을 통해 온라인거래를 한 번 이상 한 경험이 있는 일반인을 대상으로 하였다. 총 210개의 설문지가 배포되었고, 이 중 170개의 응답지를 회수하여 81.0%의 회수율을 보였다. 그 중 문항에 성실하게 응답하지 않았다고 판단된 9개의 설문지를 제외하고 161개의 유효한 설문지를 도출하여 연구 분석에 사용되었다.

* + 1. 표본의 특성

본 연구의 설문에 참가한 응답자에 관한 인구통계학적 분포는 아래 [표1]과 같으며, 응답자의 성별, 연령 그리고 인터넷 사용경험 기간 정보를 제시하였다.

[표 ] 응답자의 인구통계학적 분포

|  |  |  |  |
| --- | --- | --- | --- |
| 변수 | 분포 | 표본수(명) | 구성비(%) |
| 성별 | 남 | 60 | 37.27 |
| 여 | 101 | 62.73 |
| 연령(세)  (mean=26.89, s.d.=7.78) | 10 – 19 | 22 | 13.66 |
| 20 – 29 | 81 | 50.31 |
| 30 – 39 | 47 | 29.19 |
| 40 – 49 | 9 | 5.59 |
| ≥ 50 | 2 | 1.24 |
| 인터넷 사용경험 기간(년)  (mean=5.86, s.d.=2.42) | ≤ 3 | 26 | 16.15 |
| 3 – 7 | 101 | 62.73 |
| 8 – 11 | 34 | 21.12 |
| 합계 | | 161 | 100 |

* 1. 자료의 기술적 및 측정도구 분석
     1. 측정도구의 타당성 분석

본 연구에서는 측정항목의 타당성을 검증하기 위해 요인추출 방법으로 주성분 분석(principal component analysis)를 수행하였으며, 요인회전으로는 베리맥스(varimax) 직각회전 방식을 사용하였다. 고유값(eigenvalue)이 1.0 이상이며, 요인 적재량(self factor loading)의 값이 0.6 이상 그리고 다른 요인 적재량(other factor loading)의 값이 0.4 미만인 요인을 추출하였다. 요인분석 결과는 아래 [표 2]와 같다.

요인 적재량 값이 0.301로 측정 기준인 0.6에 미치지 못하는 측정항목 REP5는 분석에서 제외하였다. 또한 측정항목 INQ2는 요인 적재량이 0.671로 측정 기준에 만족하지만, 다른 요인과의 적재량이 0.505로 측정 기준인 0.4보다 큰 값을 보여 분석에서 제외하였다. 마지막으로 모든 5개의 요인에 대한 eigen value값은 1 이상으로 나타나고, 누적분산은 78.745%로 나타나 요인들에 대한 타당성이 만족스러운 것으로 판단되었다.

[표 ] 요인분석 결과

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 변수 | 측정항목 | 성분 | | | | |
| TRU | SYQ | REP | INQ | STA |
| 신뢰  (Trust) | TRU1 | **0.856** | 0.181 | 0.206 | 0.038 | 0.092 |
| TRU2 | **0.846** | 0.231 | 0.247 | 0.238 | 0.128 |
| TRU3 | **0.802** | 0.156 | 0.214 | 0.292 | 0.104 |
| TRU4 | **0.839** | 0.253 | 0.202 | 0.254 | 0.096 |
| TRU5 | **0.810** | 0.219 | 0.294 | 0.251 | 0.076 |
| 시스템 품질  (System Quality) | SYQ1 | 0.209 | **0.738** | 0.108 | 0.252 | 0.113 |
| SYQ2 | 0.168 | **0.847** | 0.122 | 0.250 | 0.177 |
| SYQ3 | 0.249 | **0.831** | 0.067 | 0.233 | 0.133 |
| SYQ4 | 0.191 | **0.819** | 0.206 | 0.255 | 0.151 |
| SYQ5 | 0.131 | **0.749** | 0.241 | 0.200 | 0.110 |
| 평판  (Reputation) | REP1 | 0.276 | 0.180 | **0.828** | 0.233 | 0.152 |
| REP2 | 0.279 | 0.172 | **0.868** | 0.180 | 0.175 |
| REP3 | 0.260 | 0.155 | **0.837** | 0.234 | 0.211 |
| REP4 | 0.218 | 0.131 | **0.804** | 0.264 | 0.184 |
| REP5 | 0.111 | 0.275 | **0.310** | 0.021 | -0.174 |
| 정보 품질  (Information Quality) | INQ1 | 0.307 | 0.252 | 0.250 | **0.687** | 0.091 |
| INQ2 | 0.217 | 0.505 | 0.120 | **0.671** | 0.222 |
| INQ3 | 0.269 | 0.352 | 0.261 | **0.648** | 0.266 |
| INQ4 | 0.170 | 0.255 | 0.253 | **0.820** | 0.121 |
| INQ5 | 0.255 | 0.324 | 0.223 | **0.779** | 0.084 |
| 거래 보안  (Structural Assurance) | STA1 | 0.138 | 0.133 | 0.144 | 0.101 | **0.772** |
| STA2 | 0.069 | 0.088 | 0.195 | 0.092 | **0.872** |
| STA3 | -0.044 | 0.176 | 0.111 | 0.098 | **0.886** |
| STA4 | 0.218 | 0.089 | 0.008 | 0.148 | **0.821** |
| Eigen Value | | 4.302 | 4.264 | 3.620 | 3.437 | 3.276 |
| 공분산(%) | | 17.926 | 17.767 | 15.082 | 14.321 | 13.649 |
| 누적분산(%) | | 17.926 | 35.693 | 50.775 | 65.097 | 78.745 |

* + 1. 신뢰도 분석과 상관관계 분석

본 연구에서는 측정항목 간에 내적 일관성이 있는지 확인하기 위해 신뢰도 분석(reliability analysis)을 실시하였다. 일반적으로 Cronbach’s alpha 계수가 0.6 이상이면 신뢰도가 비교적 높다고 판단한다. 검증된 변수들의 항목들에 대한 Cronbach’s alpha 계수가 0.893 ~ 0.956의 수준을 보여 변수들의 신뢰도가 높은 것으로 평가된다.

각 변수들 간의 관련성 확인을 위해 피어슨 계수를 이용한 상관관계 분석(correlation analysis)을 실시하였고, 그 결과는 [표 3]과 같다. 일반적으로 상관계수의 절대값이 0.6 이하면 적절한 상관관계, 0.6 이상이면 강한 상관관계를 갖고 있다고 판단한다. 본 연구에서는 변수 INQ와 TRU, INQ와 REP 그리고 SYQ와 INQ 간의 상관계수가 0.6 이상으로 강한 상관관계를 보이고 있다. 이러한 상관관계 분석 결과 다중공선성이 의심되어 다중공선성 진단(multicollinearity test)을 실시하였다. 그 결과 모든 요인에 대해 분산팽창지수(VIF)가 10 미만, 상태지수(condition index)가 30 미만으로 나타나 다중공선성에 문제가 없는 것을 확인하였다([표 4], [표 5]).

[표 ] 신뢰도 및 상관관계 분석 결과

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 변수 | 신뢰도 (Cronbach’s α) | 평균 | 표준편차 | TRU | REP | STA | INQ | SYQ |
| TRU | 0.954 | 5.47 | 1.01 | 1 |  |  |  |  |
| REP | 0.956 | 4.91 | 1.15 | 0.586\*\* | 1 |  |  |  |
| STA | 0.893 | 4.19 | 1.21 | 0.296\*\* | 0.377\*\* | 1 |  |  |
| INQ | 0.906 | 5.07 | 1.02 | 0.612\*\* | 0.607\*\* | 0.376\*\* | 1 |  |
| SYQ | 0.927 | 4.98 | 1.10 | 0.522\*\* | 0.465\*\* | 0.350\*\* | 0.659\*\* | 1 |

[표 ] 공선성 통계량

|  |  |  |
| --- | --- | --- |
| 모형 | 공선성 통계량 | |
| 공차 | VIF |
| REP | 0.602 | 1.661 |
| STA | 0.810 | 1.234 |
| INQ | 0.446 | 2.242 |
| SYQ | 0.550 | 1.818 |

종속변수: TRU

[표 ] 공선성 진단 확인

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 구분 | 고유값 | 상태지수 | 분산비율 | | | | |
| (상수) | REP | STA | INQ | SYQ |
| 1 | 4.886 | 1.000 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2 | 0.050 | 9.846 | 0.02 | 0.03 | 0.99 | 0.02 | 0.04 |
| 3 | 0.027 | 13.404 | 0.20 | 0.77 | 0.00 | 0.00 | 0.15 |
| 4 | 0.024 | 14.276 | 0.76 | 0.02 | 0.01 | 0.04 | 0.36 |
| 5 | 0.013 | 19.652 | 0.02 | 0.19 | 0.00 | 0.94 | 0.46 |

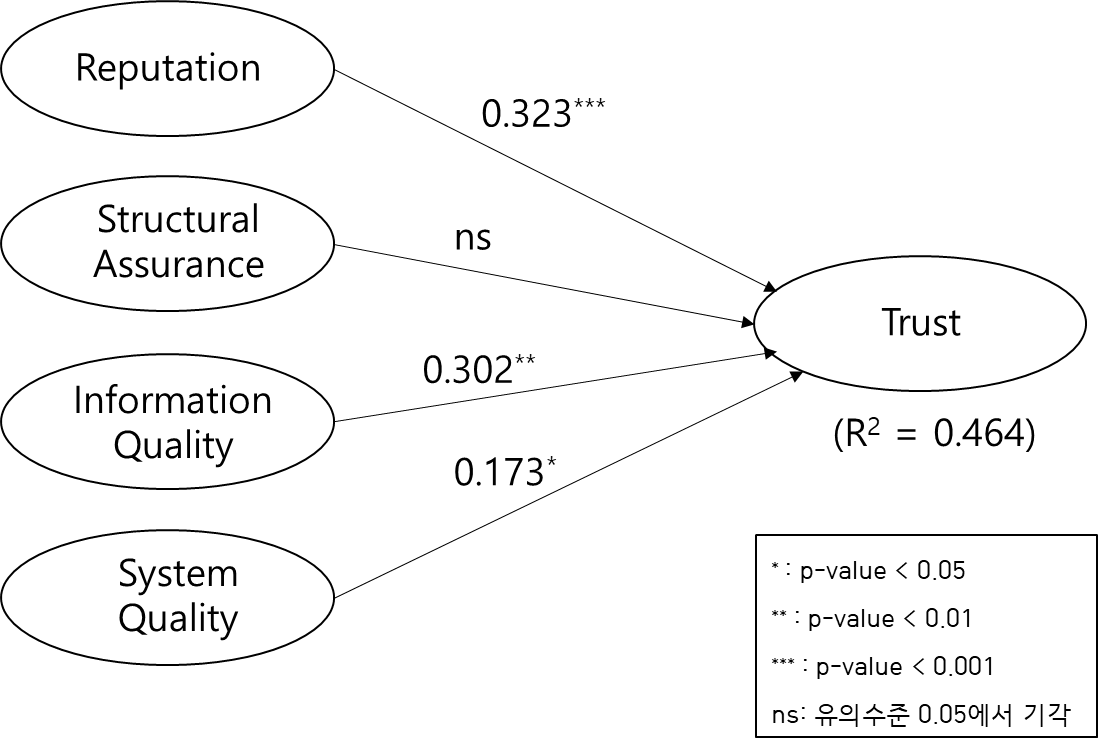
* 1. 가설의 검증 및 결과

본 연구에서는 연구 모형 및 가설을 검증하기 위해 유의확률 0.001 수준에서 다중회귀분석을 사용하였고 그 결과는 [표 6]와 같다. 다중회귀분석 결과에 제시되어 있는 beta값은 독립변수가 종속변수에 미치는 영향강도를 나타내고, 이 강도가 클수록 t값이 크게 나타난다. 변수 REP와 INQ 그리고 SYQ의 표준화 계수는 각각 0.323, 0.302, 0.173의 수치를 보이며 유의확률 0.001 유의수준에서 유의한 영향을 미치는 것으로 판단되어 가설이 채택되었다. 반면에 변수 STA는 통계적으로 유의한 영향을 끼치지 않았다. 본 분석결과에 따른 모형은 수정된 R2값이 0.461로 46.1%의 설명력을 가진다.

[표 ] 다중회귀분석 결과

|  |  |  |  |
| --- | --- | --- | --- |
| 독립변수 | 표준화 계수 | t | 유의확률 |
| beta |
|  |  | 5.318 | 0.000 |
| REP | 0.323 | 4.274 | 0.000 |
| STA | 0.000 | 0.006 | 0.995 |
| INQ | 0.302 | 3.443 | 0.001 |
| SYQ | 0.173 | 2.183 | 0.031 |

종속변수: TRU



[그림 ] 연구 모형 분석 결과