

주택 사업 분석 시스템 구축: 서울지역 아파트 가격 데이터를 중심으로

김태훈* · 이희석* · 김재윤** · 전진오*** · 이은식****

Implementing an Analysis System for Housing Business Based on Seoul Apartment Price Data

Taehun Kim* · Heeseok Lee* · Jaeyoon Kim** · Jino Jun*** · Eunsik Lee****

Abstract

The price structure of housing market varies depending upon market price policy rather than low or high price policy because of IMF. The object of this study is to develop an analysis system for analyzing housing market and its demand. The analysis system consists of four major categories: macro index analysis, market decision analysis, housing market analysis, and consumer analysis. We model each category by using a variety of techniques such as generalized linear model, categorical analysis, bubble analysis, drill-down analysis, price sensitivity meter analysis, optimum price index analysis, profit index measurement analysis, correspondence analysis, conjoint analysis, and multidimensional scaling analysis. Seoul apartment data is analyzed to demonstrate the practical usefulness of the system.

* 한국과학기술원 테크노경영대학원

** 솔토스 피엠

*** A.M.I.

**** 한국 마이크로데이터

1. 서론

전 산업의 지각변동을 야기한 국제통화기금(International Monetary Fund) 구제금융으로 인해서 주택시장환경의 급변화가 예상되고 있다. 주택가격이 저가정책(Low Price Policy)이나 고가정책(High Price Policy)보다는 시가정책(Market Price Policy)으로 변화하고 있다. 이 외에 주택 마케팅 환경 변화요인은 다음과 같다. 첫째, 공급자 중심 시장이 주택보급률 상승으로 인해서 수요자 중심 시장으로 전환될 것으로 예상된다. 둘째, 소득감소로 소유별, 지역별, 규모별로 주택수요구조가 변화할 것이다. 실수요자 중심의 시장이 투자수요를 상회할 것이다. 셋째, 소비자욕구의 다양한 변화로 인해서 연령별, 지역별, 소득계층별로 시장세분화가 가속될 것이다. 마지막으로, 분양가 자율화 전면 실시로 인한 업체간 차별화 경쟁이 심화될 것이다. 즉, 부동산 시장 개방으로 외국자본유입 및 업체간의 경쟁이 예상되기 때문이다. 이상의 향후 주택업체의 마케팅 변화 예상방향을 정리하면 <표 1>과 같다.

<표 1> 주택업체의 마케팅변화 예상

시기 구분	IMF 구제금융 이전	IMF 구제금융 이후
대상	• 불특정 다수 수요자 • 수요자	• 수요특성별 수요자 대상 • 핵심 우량 수요자 중심
목표	• 일반적 가치률 목표	• 특성별 수요자의 핵심가치를 목표
방식	• 대중에 효과적인 방식 • 일시 대량 투입 마케팅	• 특성별 수요자에 효과적인 방식 • 장기관계 기초 지속적 투입 마케팅
조사	• 일정시점 소비자 특성 파악	• 계속적인 변화 동향과 방향
수익	• 불특정 일반수요 원천	• 특성별 특정수요 원천

이에 따라 주택시장에 대한 수요를 세분화할 수 있는 대응전략이 필요하게 되었으며, 이를 위해서 마케팅 변화에 부합하는 과학적이고 체계적인 시장분석 및 수요분석을 개발하는 것이 필요하다.

본 연구 목적은 다음과 같다. 첫 번째, 기존 주택 사업방식 및 마케팅 전략접근에 있어서의 정성적이고 주관적인 한계를 극복하는 것이다. 두 번째, 시장세분화 및 수요세분화 분석을 통해서 주택시장구조의 심층적 분석을 간과한, 일방적인 소비자 접근 조사방식을 탈피하는 것이다. 세 번째, 정보시스템의 활용을 통해서, 시장분석 및 소비자 분석의 효율성을 강화하고 조사결과의 체계적인 분석과 데이터베이스로 관리하는 것이다. 네 번째, 주택사업의 추진에 있어 사전에 제기되는 시장상황판단, 소비자분석, 상품구성전략의 마케팅 접근 검토사항들을 명료화함으로써, 신속하고 객관적인 의사결정의 도구로서 역할을 강화하는 것이다. 이상의 배경과 목적에 의거해서 본 연구는 거시지표 분석, 시장규모판단분석, 주택시장분석 및 소비자 분석의 네 가지 범주(Category)로 분류된다. 분류된 범주에 대응하는 연구모형이 설계되고, 연구모형에 따른 정보 시스템을 개발하는 것이 본 연구의 주 목적이다. 또한 본 시스템은 패턴 인식(Pattern Recognition)보다는 통계 및 수리적 기법(Statistical and Mathematical Technique)에 초점을 둔 통계적 데이터 마이닝 시스템[Berry and Linoff 1997; Chen, et al. 1996; Elder and Pregibon 1996; Hou 1999]의 일종이다.

이하의 논문 구성은 다음과 같다. 2절에서는 주택가격모형에 대한 관련 연구를 대상지역, 변수, 분석기법, 목적, 특징의 구분 항목으로 비교하였으며, 3절에서는 2절에서 비교된 구분항목으로 본 연구 모형을 제시하였다. 4절에서는 개발된 분석 시스템을 자세히 설명하였다. 5절에서는 본 연구의 한계와 향후 연구 방향이 기술되었다.

2. 기존 연구

주택가격 산정에는 비용 접근법(Cost Approach), 시장 접근법(Market Analysis), 소득 접근법(Income Approach) 및 계량적 평가 기법의 네 유형의 방법이 사용된다[정홍주 1995, 이경우 1998]. 이들 기법에는 각각의 장단점이 있으나, 주택가격과 관련

된 기존 연구들은 대부분이 계량적 평가기법중의 하나인 특성가격 접근법(Hedonic Price Approach)을 통한 매매가격의 예측 및 특성가격 추정을 가능하게 하는 주택가격모형 설정에 초점을 맞추고 있다. 경험적연구에서 주로 이용되는 특성가격모형은 다음의 식 (1.1)과 같이 모형화된다[윤대식, 윤성순 1995]. 특성가격모형에 영향을 미치는 변수들 (H_i) 을 설명변수로 포함시킬 것인가 하는 것은 분석대상지역의 제반상황과 자료를 면면히 검토한 후에 결정된다.

$$R_i = \alpha_0 + \sum_{i=1}^n r_i H_i + e_i \quad (1.1)$$

($i=1, 2, 3, \dots, n$)

R_i = 주택가격 또는 임대료
 H_i = 주택의 특성을 나타내는 변수들의 벡터
 e_i = 오차항

1990년도 이후의 기존 연구는 <표 2>에 정리되어 있으며, 1980년 말에서 1990년 초까지의 기존 연구는 김명호[1994]에 정리되어 있다. 서울지역을

대상으로 한 연구[김용운 1990, 이진목 1991, 정홍주 1995, 조주현 1998]는 서울시 전체 혹은 서울 일부를 대상으로 회귀분석 기법으로 분석하였다. 김명호[1994]와 이경우 [1998]는 각각 청주시와 부산, 경남 지역에 대한 특성파악에 중점을 두고 있다. 실제 현장에서는 주택 사업 분석을 대부분 정성적(Qualitative)으로 판단하고 있는 실정이다. 반면에 <표 2>와 같은 수리적 모형을 가지는 정량적인(Quantitative) 연구도 많이 진행중이지만, 이러한 연구들은 한 시점의 단면적(Cross-sectional) 연구에 치중하고 있기 때문에 지속적으로 사용되고 의사결정지원(Decision Support)을 위한 정보 시스템 관점에서는 거의 실현되어져 있지 않으며, 지속적인 데이터의 보충이 불가능하다. 기존 연구를 본 연구와 비교하면, 주로 기존연구들은 본 연구의 주택시장분석 범주에 초점을 두었으며, 분석대상지역의 단면적 연구에 치중하였고, 정보시스템으로는 구현되지 못하였다.

<표 2> 주택가격모형의 관련 연구

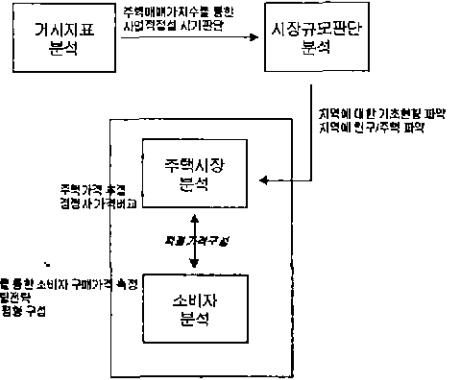
구분 연구자	목적	분석기법	변 수	대 상 지 역	특 징
이경우 [1998]	분석 대상 지역 의 특성 파악	• 분산분석	• 아파트 특성 (3개) • 인구통계학적 특성(6개) • 가격차등 특성 (2개)	• 부산, 경남 지역 아파트	• 단면적 연구 (설문조사 238명)
손재영 [1998]		• 회귀분석	• 이자율 변수 (2개) • 통화변수 (3개) • 소득변수 (2개) • 물가변수 (2개)	• 전국	• 시계열 연구 • 지가지수에 대한 분석
조주현 [1998]		• 회귀분석	• 단지면적 • 용적율 • 화장실수 • 최초입주일	• 서울 (성북, 종로, 중구 제외)	• 시간대별, 평형별 단면적 연구
정홍주 [1995]		• 회귀분석 (Box-Cox 변환)	• 주거특성 (5개) • 지역특성 (6개) • 환경특성 (1개)	• 서울지역 한강변 아파트 (156개)	• 단면적 연구
김명호 [1994]		• 회귀분석	• 물리적 특성 • 택지적 특성 • 접근성 • 만족도	• 청주시 자가주택 (700개)	• 단면적 연구
이진목 [1991]		• 회귀분석 (Log 변환)	• 지역 • 규모	• 서울지역 7개 아파트	• 시계열 연구 • 매매 가격 및 전세가격 • 매매가격과 전세가격간 모형
김용운 [1990]		• 회귀분석 • 요인분석	• 건축적 구성요소 (5개) • 입지적 구성요소 (6개) • 근린 사회환경 (2개)	• 서울시 (종로구, 중구, 성북구 제외), 파천시 아파트 (315개)	• 단면적 연구 (매매 및 전세 가격 요소에 관한 설문조사 315개)

3. 본 연구 모형

격구성표를 제시한다.

본 연구는 손쉽게 이용 가능한 시계열 (Time-variant) 자료와 단면적 자료에 대해서 모형을 설계하고, 정보시스템으로 구현하였다. 주택가격을 하나의 모형으로 고정시킨 것이 아니라, 거시지표 분석, 시장규모판단분석, 주택시장분석, 소비자분석으로 나누어서 의사결정에 이용될 수 있게 하였다. 이 연구모형을 (그림 1)과 같이 약할 수 있다.

(그림 1)은 네 가지 범주의 연구모형과 절차를 보여준다. 거시지표분석을 통하여 사업적정시기를 파악하고, 사업하려는 지역에 대한 기초 통계를 시장규모판단분석을 통해 수행하며, 주택시장분석을 통해서 주택가격을 추정한다. 최종적으로 소비자분석을 모델하우스 등의 조사를 통해서 소비자 구매가격, 평형 등을 분석한다. 주택시장분석과 소비자분석의 결과를 통해서 의사결정자에게 적정가



(그림 1) 본 연구 모형의 절차

<표 3>은 (그림 1)의 연구모형을 상세히 기술한 것이며 <표 2>와 대비되는 표이다. 네 가지 범주의 목적에 따라서 분석기법을 적용하였으며,

<표 3> 본 연구의 모형

범주	구분	목적	분석기법	변수	대상지역	특징
범주	거시지표분석	<ul style="list-style-type: none"> 사업적정시기 판단 - 거시지표의 경향과 주택시장과의 연관성 파악 거시적 사업 위험 검토 	<ul style="list-style-type: none"> 일반화 선형모형 (분산 레그된 회귀 분석) 	<ul style="list-style-type: none"> 13년간 거시지표 자료 (20개 변수) 주택매매가치수 	<ul style="list-style-type: none"> 서울 (강북과 강남으로도 분류), 경기, 인천 (5개 지역) 	<ul style="list-style-type: none"> 시계열 연구 월 모형 및 분기모형
	시장규모판단 분석	<ul style="list-style-type: none"> 기초 현황파악 및 다양한 관점의 자료 분류 인구 및 주택 현황 파악 	<ul style="list-style-type: none"> 범주형 분석 비블분석 및 드릴다운 분석 	<ul style="list-style-type: none"> 연도, 입지 (두개의 변수), 인구수 (전입 및 전출), 세대수, 연령, 종류별 주택수, 주택의 점유형태 	<ul style="list-style-type: none"> 서울 전지역 	<ul style="list-style-type: none"> 시계열 및 단면적 연구
	주택시장분석	<ul style="list-style-type: none"> 지역아파트 속성 분석을 통한 지역 시장특성 파악 동일 및 인근 지역의 경쟁 공급 업체의 상품 속성 비교분석 	<ul style="list-style-type: none"> 일반화 선형모형 (공분산분석, 분산분석, 회귀분석) 	<ul style="list-style-type: none"> 연도, 입지 (두개의 변수), 브랜드, 평형, 준공년도 (연식), 세대수 평당매매가/전세가 	<ul style="list-style-type: none"> 서울 전지역 아파트 (25개구 530개동 2808개 아파트) 	<ul style="list-style-type: none"> 시계열 및 단면적 연구 특성가격 접근법
	소비자 분석	<ul style="list-style-type: none"> 소비자 행태분석을 통한 마케팅 전략을 파악 소비자 구매의향 측정 (희망 평형 및 가격) 적정 공급 평형 효용값에 의한 경쟁업체의 상품 포지셔닝 브랜드 마케팅 전략 시장 점유율 분석 적정공급가격분석 	<ul style="list-style-type: none"> 적정가격구성 분석 가격민감도 측정분석 이익지수법 분석 대응분석 컨조인트분석, 회귀분석, BTL 분석 다차원척도분석 	<ul style="list-style-type: none"> 인구 통계학적 배경 항목 및 본조사 항목(7개 부분) -41개 변수 측정(속성별 인지도, 중요도, 만족도 등) 	<ul style="list-style-type: none"> 서울시 송파구, 강남구, 광진구, 강동구(4개구) 	<ul style="list-style-type: none"> 단면적 연구 (설문조사 400명) 시장 접근법

분석에는 다양한 기법이 사용되었지만 일반화 선형모형(Generalized Linear Model)[SAS Institute 1982]에 초점을 두고 있다. 각 범주에 맞는 변수들이 틀리며, 변수들의 자료는 KDI, 통계청, 민간경제연구소, 부동산뱅크 등 자료제공업체로부터 유, 무상으로 제공받았다. <표 2>는 대부분 단면적 연구의 일회성으로 끝나고 있지만, 본 연구는 정보시스템으로 구현하기 위해서 시계열 모형에 중점을 두고 있다. 특히 세부적 분석을 위해서 소비자 분석에서 400명을 대상으로 설문조사를 시행하였다.

일반적으로 주택의 구분은 <표 4>와 같으며 [이경우 1998], 본 연구는 주택 중에서도 데이터 수집과 분석이 용이한 아파트에 중점을 두고 있다. 세부적으로 설명하기 위해서 실제 아파트 건설 계획에 사용된 서울시 송파구 1동에 대한 데이터를 사용하였다.

<표 4> 주택의 구분

거처	주택	단독주택	
		공동주택	다세대주택
			연립주택 아파트
		비거주용 건물내의 건축물	
주택이외의 거처			

3.1 거시지표분석

거시지표 분석은 장기 및 단기적인 거시지표 예측을 통해서 유리한 사업시기 판단근거를 마련하는 것이다. 즉, 주택사업에 영향을 미치는 각종 거시지표들의 경향과 주택시장과의 상호연관성을 파악하고, 사업시점의 거시적 동향 파악을 통해서 사업위험을 관리하는데 목적이 있다. 고려된 거시지표 변수로는, 월자료인 주가지수, 총통화, 외환보유액, 소비자물가, 제조업가동률지수, 일반대출금리, 자가변동율 등 17개 변수와 분기자료인 지가변동율, 경상가계최종소비, 불변가계최종소비의 3개 변수이다. 서울, 강북, 강남, 경기, 인천의 5개 지역별로 86년부터 98년까지 주택매매가지수를

가지고 회귀분석에 적합시켜 보았다. 회귀진단(Regression Diagnostics)결과 다중공선성과 자기상관(Auto-correlation)이 존재하여서, 변수 삭제를 통한 다중공선성 제거를 수행하였다. 자기상관은 Koyck 의 분산 래그(Distributed Lag)[백운봉 1990; Wonnacott 1981]를 사용하여 제거하였으며, 식 (1.2) 과 같은 모형으로 적합되었다. 식 (1.2)에서 t는 현시점을 말하며, t-1은 t 시점의 전(前)시점을 나타내며, 독립변수는 <표 5>에 표시되어 있다.

$$Y_t = \lambda Y_{t-1} + \beta_0 X_{t1} + \beta_1 X_{t2} + \beta_2 X_{t3} + \dots + e_t^*$$

where $e_t^* = e_t - \lambda e_{t-1}$

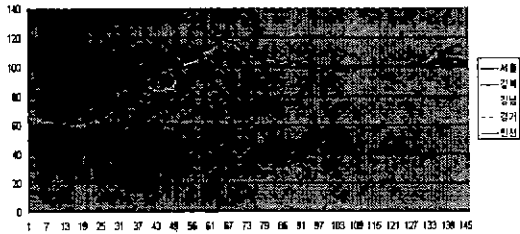
Y_t = 주택매매가지수, X_{t1} = 독립변수

<표 5> 월자료의 주택매매가지수에 관한 모형

지역	종속변수	독립변수
서울	주택매매가지수 (1986 ~ 1998)	서울 전시점 주택매매가지수, 주가지수, 제조업가동율지수, 국내건설수주액, 경상수지, 연월말기준환율, 건축허가면적, 실업률
강북		강북 전시점 주택매매가지수, 제조업가동율지수, 국내건설수주액, 경상수지, 연월말기준환율, 건축허가면적, 실업률월별추이, 일반대출금리
강남		강남 전시점 주택매매가지수, 주가지수, 제조업가동율지수, 경상수지, 외환보유액, 건축허가면적
경기		경기 전시점 주택매매가지수, 경상수지, 건축허가면적
인천		인천 전시점 주택매매가지수, 주가지수, 제조업가동율지수, 국내건설수주액, 경상수지, 건축허가면적, 실업률월별추이, 일반대출금리

대부분의 거시지표 독립변수가 월자료이지만 지가변동율 같은 자료는 분기로 발표되는 자료이기 때문에, 월자료 모형과 분기자료 모형으로 존재하게 되었다. <표 5>는 월자료 모형에 적합된 독립변수들이다. 각 지역의 독립변수에는 전시점의 주택매매가지수가 독립변수로 들어간다. 설명계수 (R^2)은 대부분 90% 이상으로 나타났다.

지역별 월자료에 대한 단일시계열 분석결과는 박스-젠킨스의 계절형 ARIMA (p,d,q)*(P,D,Q)s 로 적합되지만, 시스템 개발의 난해함때문에 제외하였다. (그림 2)처럼 각 지역의 매매 추이가 동일함을 알 수 있다. 그 결과는 인자분석을 통해서도 확인할 수 있었다.

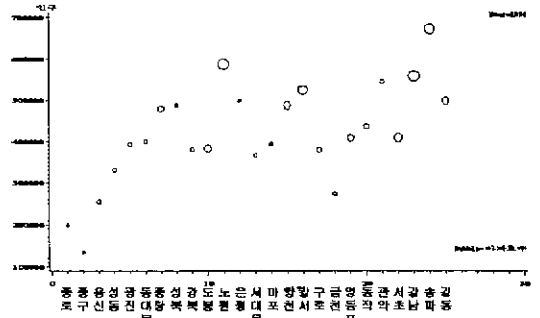


(그림 2) 지역별 월자료 주택매매지수 추이

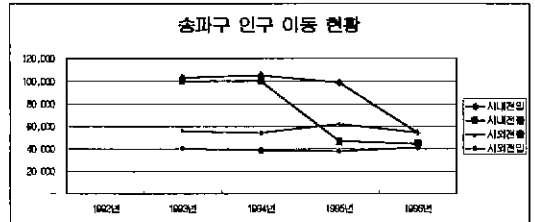
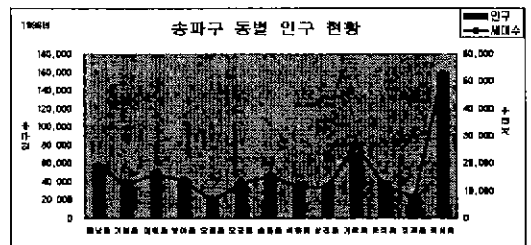
3.2 시장규모판단분석

시장규모판단분석은 인구 통계학적 분석으로, 서울시 25개구 530개동 기초현황자료를 분석하고 다양한 관점으로 자료를 분류하기 위한 것이다. 구체적으로 구, 동별 인구 및 세대수 연도별 증감 현황, 구, 동별 인구 현황, 구별 광역 인구이동 현황, 지역별 구간 인구이동 현황, 주택점유형태, 주택 종류, 구별 평형별 아파트 현황으로 나타난다.

(그림 3)는 서울시 구별 버블 분석(Bubble Analysis)을 보여준다. 분석 결과로 성북, 은평, 관악구는 상대적으로 다른 구에 비해서, 인구대비 아파트 수가 적음을 알 수 있다. 이에 대한 드릴 다운 분석(Drill-down Analysis)의 필요성이 대두된다. 반대로 송파구는 가장 많은 인구수에 도 불구하고 노원, 강남구보다는 적은 아파트 수를 가진다. 이에 대한 드릴 다운 분석으로, 송파구 동별 인구현황 및 송파구에 대한 인구이동 현황 추이가 (그림 4)에 나타나있다. 송파구에서 잠실동이 가장 많은 세대수를 가지며 요문동이 가장 적은 세대수를 가지고 있다. 과거의 인구 이동을 보기 위해서, 선택된 기간('93~'96)의 인구이동은 1996년에 급속히 줄어들었음을 알 수 있다.



(그림 3) 서울시 구별 버블 분석



(그림 4) 송파구 동별 인구현황 및 송파구에 대한 인구 이동 현황 추이

3.3 주택시장분석

식 (1.1)의 H_t 를 구와 동의 입지, 브랜드, 평형, 준공년도로 나타내었다. 분석결과 준공년도는 의미 없는 변수로 판단되어 삭제하였다. 전체자료가 불균형자료(Unbalanced Data)이기 때문에 일반화 선형모형을 써서 분석하였다[Speed et al. 1978]. 식 (1.1)을 일반화 선형모형으로 나타내면 다음의 식 (1.3)과 같다.

$$y = X\beta + e \tag{1.3}$$

〈표 6〉 송파구 모형의 모수 추정표

상 수 항			211.5		
동	1	-88.6140113	평 형 (동)	1	12.4168841
	2	18.6853397		2	11.2135712
	3	241.7235485		3	9.3372446
	4	-103.6286849		4	15.1307256
	5	160.1106195		5	8.6666141
	6	-85.7569447		6	13.1455870
	7	149.0728533		7	5.6927781
	8	297.7489731		8	0
	9	13.0699311		9	10.9937639
	10	-481.7591050		10	23.6194887
	11	-114.3639344		11	14.4157629
	12	-95.8725341		12	14.0739213
	13	-20.4906464		13	12.2849103
	14	0		14	10.8846154
브랜드	A	99.0271467	평 형 (브랜드)	A	-3.4927781
	B	164.1212059		B	-3.7762841
	C	-17.1382301		C	2.2567768
	D	206.0194492		D	-6.4000649
	E	74.8091057		E	-3.7049690
	F	527.5140113		F	-14.2168841
	G	983.4084410		G	-25.4594867
	H	644.1658472		H	-3.0699315
	I	175.1286849		I	-7.3807256
	J	99.1689168		J	-2.7207634
	K	74.6695755		K	-7.1725583
	L	533.7394227		L	-11.3363801
	M	89.0886448		M	-1.9491627
	N	663.8686468		N	-20.1243104
	O	563.1257971		O	-11.4570624
	P	303.9261754		P	-7.7686639
	Q	48.6358596		Q	-1.6576032
	R	841.2591050		R	-24.8694887
	S	199.7000000		S	-5.7846154
	T	292.0732697		T	-8.3874700
U	138.5994475	U	-6.0006375		
V	169.3656464	V	-6.1599103		
W	0	W	0		

예를 들어, 송파구 한 곳만의 분석결과는 <표 6>과 같다. 식 (1.3)에서 y 는 243×1 의 평당 매매가의 벡터이며, β 는 3×1 의 모수 벡터이며 e 는 243×1 의 오차 벡터이다. X 는 243×4 의 계획행렬(Design Matrix)으로써 동(입지), 브랜드, 평형으로 구성되어 있다. 동과 브랜드는 명목형 변수이고 평형은 연속형 변수이기 때문에, 불균형 자료의 공분산 분석 (Analysis of Covariance)을 수행하였다. 공분산 분석에서 기울기의 동질성 검정 (Homogeneity-of-slopes Test)을 채택하지 못하고 분리된 기울기 모형이 채택되었다. 이로 인해 평형간의 기울기는 상이함을 알 수 있었다. <표 6>에서 평형(동)과 평형(브랜드)

는 동과 브랜드 별로 다른 기울기 모수추정값을 나타낸다. 송파구의 1동, A브랜드, 24평형의 평당가 추정을 계산하면 다음의 식 (1.4)와 같다. 실제 값은 417만원이기 때문에 오차는 18만원이다. 이는 같은 1동, A브랜드, 21평, 24평의 평당가가 429, 417로써, 둘의 편차가 12만원 나고 A브랜드 자체의 표준오차가 126.437936 으로 상당히 크기 때문에 18만원 오차가 나는 것이다.

$$\hat{y} = 211.5 - 88.61 + 99 + 12.42 * 24 - 3.49 * 24 = 436 \text{만원} \quad (1.4)$$

시장진입을 위해서는 실제값이 없기 때문에 비교하기는 힘들지만 추정은 할 수 있고 의사결정자가 판단할 수 있는 정보를 제공한다. 1동, B사, 33평형의 평당가 추정은 식 (1.5)과 같이 할 수 있다. 이 같은 분석은 경쟁업체에 적용된 실제 분양가와 비교하여 수정 및 보완할 수 있다.

$$\hat{y} = 211.5 - 85.76 + 175.13 + 13.15 - 7.38 * 33 = 491.46 \text{만원} \quad (1.5)$$

송파구 모형의 설명계수는 84%로 나왔다. 한 개의동에 하나의 브랜드만 있는 경우가 하나있는데, 이 경우의 모수추정은 불가능하기 때문에 그 자료를 삭제하였으며, 세대수는 카운트(Count)로 사용되었다. 이런 모형을 기반으로 해서 서울시 전체에 적용하면 모수 추정과 최소제곱평균값 (Least Square Mean : LSMean; 보정평균)[성내경 1991; Montgomery 1984]을 구할 수 없었기 때문에, 정보의 손실 (Loss of Information)을 감수하더라도 평형자료를 10구간으로 범주화시킨후 분석을 하였다. 그 결과 불균형자료의 분산 분석(Analysis of Variance)이 수행되었으며 설명계수는 75%로 낮아졌다. 서울지역별 최소제곱평균은 (그림 5)와 같이 나타난다. 강남구가 858 만원의 최고 평당가를 가지며 강북구가 453만원의 최저 평당가를 가짐을 알 수 있다.

식 (1.3)에서 계획 행렬의 각 변수를 하나씩 분리해서, 세부적인 분석이 가능하게 하였다. 평형별, 준공년도별로는 평당가격에 대해서 회귀분석

	YLSMEAN
강남구 평균	858,2385292
강동구 평균	620,667235
강북구 평균	453,2658267
강서구 평균	594,1617156
관악구 평균	610,982664
광진구 평균	654,8769367
구로구 평균	495,5005486
금천구 평균	503,55904
노원구 평균	500,1399667
도봉구 평균	494,4415717
동대문구 평균	573,2792238
동작구 평균	630,495424
마포구 평균	616,5982511
서대문구 평균	647,704853
서초구 평균	802,295435
성동구 평균	600,7666008
성북구 평균	569,5630077
송파구 평균	678,9564885
양천구 평균	652,6583239
영등포구 평균	591,0830008
용산구 평균	744,9989771
은평구 평균	593,43901
종로구 평균	629,936285
중구 평균	495,76177
중랑구 평균	545,274066
전체 평균	618,6513598

(그림 5) 서울지역별 최소제품평균값

을 시행하였다. 구별, 동별, 브랜드별로는 평당가격에 대한 분산분석을 시행하였다. 이것을 통해서 평당가를 기준으로 타 속성의 포지션닝과 특징들을 분석할 수 있다. 또한 각 독립변수내에서 다중비교(Multiple Comparisons)기법중의 하나인 본페로니 t 검정(Bonferroni's t Test) 실시하여 어느

모평균들간에 차이가 있는지를 알아보았으며 그룹화를 실시하였다. (그림 6)은 브랜드별 평당가 분석에서, 본페로니 t 검정의 108개 결과 일부와 그룹화 결과를 보여준다. 전체 비교에 대한 유의 수준은 0.25로 고정시켰다. 이에 대한 108개의 쌍별비교에 대한 수정된 유의 수준(Corrected Significance Level)은 0.0002(=0.25/108)이다. 이 유의 수준이 작은 값이기 때문에 대부분은 유의하지 않다는 결과를 가진다((그림 6)의 왼쪽 그림 참조). (그림 6)의 오른쪽은 본페로니 t 검정을 이용해서 브랜드를 그룹화 시킨 결과이다. 결과의 해석을 두 가지로 나눌 수 있다. 첫째, 카테고리 부분의 22, 2, 1, 18, 15, 16, 30, 40, 14, 11, 9, 13, 10, 54 브랜드를 상위그룹으로, 12, 54 브랜드를 중위그룹으로, 53 브랜드를 하위그룹으로 처리하는 것, 둘째, 22 브랜드만을 상위그룹으로 2, 1, 18, 15, 16, 30, 40, 14, 11, 9, 13, 10, 54, 12, 52 브랜드를 중위그룹으로, 53 브랜드를 하위그룹으로 처리하는 것이다.

3.4 소비자분석

소비자분석은 투자수요자보다 실수요자를 선정하기 위해서 송파구 1동 모델하우스 내에서 이루어진 설문조사 결과를 토대로 이루어졌으며, 아파트 속성별 소비자 행태분석을 통한 마케팅 전략을 마련하는 목적이 있다. 설문조사를 계속해서 데이터

40 ~ 13	-73,0283	-0,9122	0,3627 No
40 ~ 11	-21,3492	-0,1373	0,8909 No
40 ~ 54	-224,5395	-3,1065	0,0021 No
2 ~ 14	-87,4062	-0,7510	0,4534 No
2 ~ 10	-200,9597	-1,4299	0,1541 No
2 ~ 13	-155,6767	-1,1676	0,2442 No
2 ~ 11	-103,9976	-0,5515	0,5818 No
2 ~ 54	-307,1879	-2,3848	0,0179 No
14 ~ 10	-113,5535	-1,1593	0,2476 No
14 ~ 13	-68,2705	-0,7821	0,4350 No
14 ~ 11	-16,5913	-0,1041	0,9172 No
14 ~ 54	-219,7816	-2,7400	0,0066 No
10 ~ 13	45,2930	0,3851	0,7005 No
10 ~ 11	96,9622	0,5454	0,5860 No
10 ~ 54	-106,2281	-0,9448	0,3458 No
13 ~ 11	51,6791	0,3002	0,7643 No
13 ~ 54	-151,5111	-1,4670	0,1438 No
11 ~ 54	-203,1903	-1,2048	0,2296 No

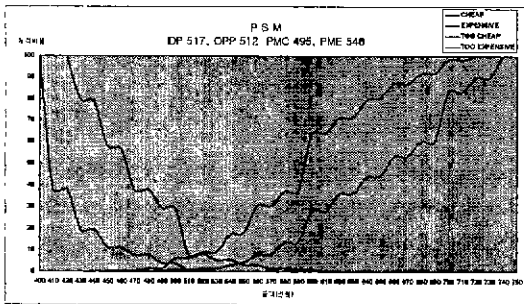
Corrected significance level : 0,0002

Means Ranking and Categories Grouping using Bonferroni's test:

Categories	Mean	Grouping
22	529,63477 A	
2	532,36607 A	B
1	562,20673 A	B
18	571,70168 A	B
15	588,93281 A	B
16	589,14261 A	B
30	591,1541 A	B
40	615,01444 A	B
14	619,77231 A	B
11	636,36364 A	B
9	649,68005 A	B
13	688,04277 A	B
10	733,32581 A	B
54	839,5539 A	B
12	944,73594	B
52	976,26425	B
53	1359,5111	C

(그림 6) 브랜드별 평당가 분석에서 본페로니 t 검정 결과 일부와 그룹화 결과

베이스로 저장하다면 시계열적 분석을 수행할 수 있을 것이다. 조사원을 통해서 400명의 설문지를 회수하였다. 설문지 구성은 조사배경과 협조기술 부분, 인구 통계학적 배경항목, 그리고 본조사 항목으로 구성되었다. 본조사 항목은 크게 아파트 상표 지수/연상, 이상적 아파트, 아파트에 대한 태도, 아파트 구매과정, 현 아파트 평가, 속성 조합간 선호 평가, 가격 민감도 측정(Price Sensitivity Meter; PSM)으로 구성되어 있다. 설문의 분석 결과는 적정평형 및 적정가격, 상품 포지셔닝(Product Positioning), 선호도(Preference)순으로 행하였다. 적정평형 구성전략으로써 향후구입 평형을 피벗 테이블(Pivot Table)을 통해서 분석하였다. 적정가격 구성(Otimum Price Index; OPT)분석에서는 PSM, 이익지수법(Profit Index Measurement; PIM), 일반화선형모형 가격의 조합을 통해 적정 가격대를 분석하였다. 34평형에 대한 PSM의 결과는 (그림 7)과 같다.



(그림 7) PSM 결과

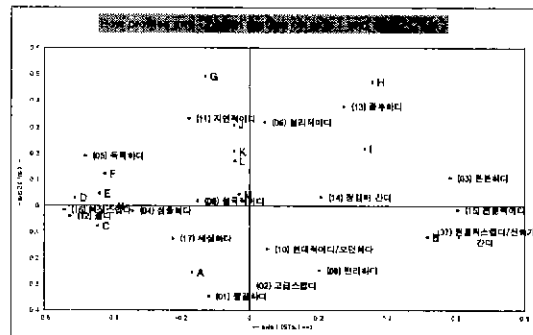
PSM은 소비자인지 시세가격(IDP), 소비자인지 적정가격(OPP), 소비자 가격수용가격대 (PMC~PME)로 구성되어 있다. 조사결과 가격이 시세가격에 비해서 저평가되어 있었기 때문에 시세가격을 반영해서 재평가하였다. 또한 이익 지수법을 사용하여 소비자 구매 의향율과 공급자의 단위 이익의 점점 분석을 통해 소비자 요구와 공급자 이익이 일치하는 이상점점가격을 분석 할 수 있었으며, 주택시장분석의 일반화 선형모형을 이용한 가격도 추정하였다. 적정가격구성을 송파구 1동, A브랜드,

34평형에 대해서 표로 정리하면 <표 7>과 같다.

(표 7) 적정가격구성 표

구분	PSM		이익 지수법 가격	일반화 선형모형 가격
	소비자인지 적정가격	소비자인지 수용가격		
평당가 (만원)	555	591	543	654
판 단 내 용	• 최저가격 • 분양율100%	• 유효마케팅에 의한 유 보가격	• 손익 가격	• 이상가격 • 주택시장 분석

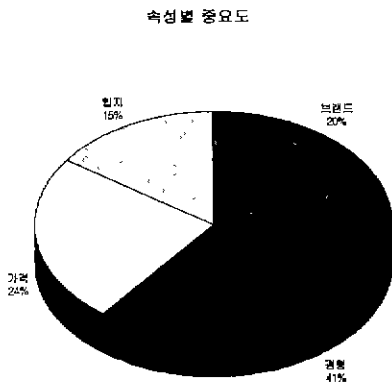
상품 포지셔닝과 속성에 대한 마케팅 전략에의 응용을 위해서 대응분석(Correspondence Analysis)[허명희 1989]을 수행하였다. 그 결과는 (그림 8)에 나타난다. 결과적으로 A브랜드의 이미지 속성은 “깔끔하다”, “고급스럽다”, “세심하다”, “편리하다”, “현대적이다”으로 나타나고 있다. 경쟁사인 B브랜드는 “튼튼하다”, “믿음직스럽다”, “전문적이다”라는 속성에 강세를 띄고 있다. 그러므로 A브랜드는 고급적, 현대적 이미지의 차별화 전략에 있어 우위에 있으며 이 쪽으로 브랜드 이미지의 인지를 강화하는 데 주력할 수 있다.



(그림 8) 대응 분석 결과

소비자들의 아파트 선호도에 있어 어떤 속성들을 중요하게 여기는 지를 알기 위해서 컨조인트 분석(Conjoint Analysis)을 시행하였다. 속성들로서는 브랜드, 평형, 가격, 입지를 고려하였다. 평형과 가격은 연속형 변수이기 때문에 범주화시키는

것보다는 현실적으로 타당한 고정값(Fixed Value)을 주었다. 브랜드 수준은 4이고 입지 수준은 3으로 두었으며, 총 가지수는 4종류수로 제한하였는데, 이는 컨조인트 관련 항목 뿐만 아니라 여러 가지 다른 설문항목이 있었기 때문이며, 추후 실험 계획(Experimental Design)을 라틴 방격법 혹은 직교 설계로 설계함이 타당할 것이다[허명희 1994]. 400명 표본으로부터의 회귀분석한 결과를 가지고 진입지역 상품구성 및 타업체 상품구성을 비교할 수 있으며, 이를 통한 점유율 예측을 행할 수 있다. 또한 이를 바탕으로 BTL 분석을 행할 수 있다. BTL 분석은 근접 경쟁업체를 파악하는 데 이용하는 것이다. 컨조인트 속성별 중요도 분석은 (그림 9)에 나타나있다. 분석결과 소비자는 평형 > 가격 > 브랜드 > 입지 순의 비중을 가지고 선택하는 경향을 보인다. 즉, 상품 구성시 적정 평형구성과 가격 전략이 우선시 되어야 한다. 소비자의 아파트 속성별 중요도 및 만족도 분석을 다차원 척도법(Multi Dimensional Scaling)에 의해서 시행하였다 시행결과 만족도가 떨어지는 소비자 군에 대해서 집중 마케팅을 통해 개선해야할 속성들을 파악하였다.



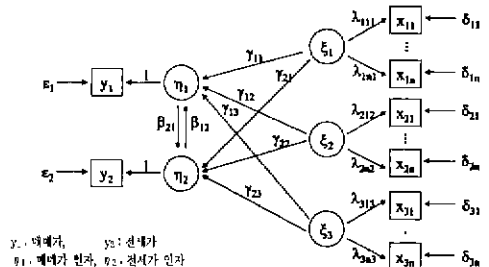
(그림 9) 속성별 중요도

3.5 추후 과제 : 구조방정식 모형

본 연구는 자료의 가용성(Availability)측면을 강조함으로써, 모형의 정밀도라는 점에서 단점을 지

니고 있다. 이러한 가용성을 떠나, 주택시장 분석에서, 독립변수들을 추가시켜 모형화하는 연구가 이루어져야 한다. 이를 위하여 구조방정식 모형(Structural Equation Model)[Jöreskog 1989]을 고려하여 입지인자, 브랜드인자 및 건축인자를 구성하는 변수들을 추가시켜서 인자들의 구조모형을 도출할 수도 있다.

(그림 10)은 주택시장 자료의 구조방정식 초기 완전 모형(Initial Full Model)을 나타낸다. ϵ_1 는 측정변수 y_1 에 대한 측정오차를 나타내며, δ_1 는 측정변수 x_1 에 대한 측정오차를 나타낸다. y_1 은 매매가 변수를 나타내며, y_2 는 전세가 변수를 나타낸다. 내생변수(Endogenous Variable) η_1 은 매매가 인자로서 측정변수로 y_1 을 가지며 그 때 요인계수(Factor Loading)는 1로 고정시켰다. 내생변수 η_2 전세가 인자를 나타내며, 또한 요인계수는 1을 가진다. 내생변수간의 경로계수는 β_{ij} 로 표현되어서 적절한 값을 추정할 수 있다. ζ_1 은 입지 인자를 나타내는 외생변수(Exogenous Variable)인데, 측정변수로서 통근거리, 지형, 도로망, 교육시설 등을 나타내는 x_{11} 로 표현되어진다. ζ_2 는 브랜드 인자, ζ_3 는 건축 인자를 나타낸다. 브랜드 인자는 시공회사, 회사이미지, 사업년수 등을 나타내는 x_{21} 로 표현되며, 건축인자는 용적율, 평형, 건축년도 등을 나타내는 x_{31} 로 표현된다. λ_{ijk} 는 외생변수와 측정변수간의 요인계수를 나타내며 추정할 모수이다. γ_{ij} 는 외생변수가 내생변수에 가지는 경로계수를 나타낸다.



y_1 : 매매가, y_2 : 전세가
 ϵ_1 : 매매가 원차, δ_1 : 전세가 원차
 x_1 : 통근거리, 지형, 도로망, 정거장 도보거리, 교육시설, 시공회사, 소음 및 악취정도, 위험시설, 권역시설, 상수도/하수시설, 옥외시설 등
 x_2 : 용적율, 평형, 건축년도, 방할, 층수, 방수, 난방방식, 욕실, 지시, 전방, 보온장터 등
 x_3 : 시공회사, 회사이미지, 사업년수, 회사만족도, 회사인지도 등
 ζ_1 : 입지인자, ζ_2 : 브랜드 인자, ζ_3 : 건축 인자

(그림 10) 주택시장 자료의 구조방정식 초기 완전 모형

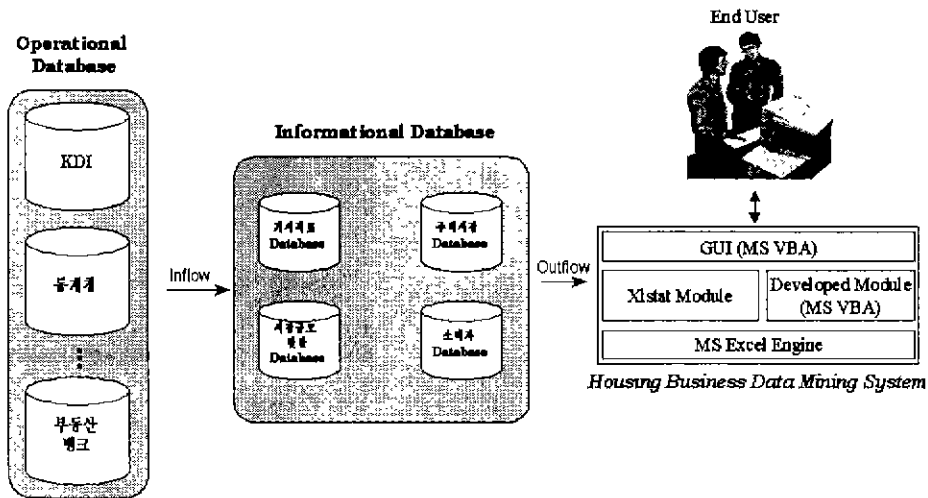
이러한 모형은 가용성과 정밀도간의 흥정(Trade-off) 관계가 있다. 이러한 모형을 정보 시스템으로 구현할 때, 수많은 변수들의 데이터를 얻기가 힘들 뿐 아니라, 모형을 쉽게 도출하더라도 데이터를 유지 및 보수하는 측면에서 비용이 많이 들 수 있고 사용하기가 힘들어질 수 있다. 즉, 실제 현장(Real Field)에서는 적은 변수를 사용해서, 적은 비용의 유지 및 보수가 이루어지는 모형을 선호하며, 이 또한 정보 시스템 구현에 쉽게 적용되어 질 수 있다. (그림 10)과 같은 모형은 주택시장 분석의 프레임워크(Framework)로 활용될 수 있으며, 연구로써도 이루어져야 한다. 다만 본 연구에서는 현실 적용성에 중점이 주어졌으므로 본 모형은 배제하였다.

4. 시스템 구현

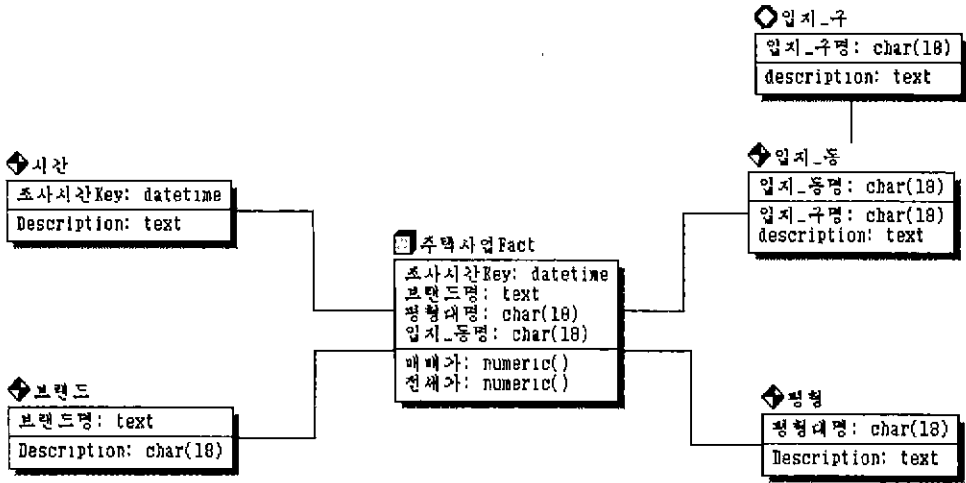
주택 사업 분석 시스템(Analysis System for Housing Business)은 현재 마이크로소프트 엑셀 엔진(Excel Engine)[Albright et al. 1999]과 액세스 엔진(Access Engine)[Roman et al. 1999]기반으로 운영되고 있으며, 시스템 구성도는 (그림 11)과 같다. 운영 데이터베이스(Operational Database)는 KDI, 통계청, 부동산 بانک 등에서 받는 데이터베이스이며, 정보 데이터베이스(Informational Database)는 거시경제 데이터베이스, 주택시장 데이터베이스, 시장규모 변동 데이터베이스, 소비자 데이터베이스로 구성된다. 이 데이터베이스는 GUI(MS VBA)를 통해 접근되며, Xlstat 모듈과 Developed Module(MS VBA)를 사용하여 분석이 수행된다. 이 시스템은 MS Excel 엔진을 기반으로 구축된다.

이며, 정보 데이터베이스(Informational Database) [이희석 외 1997]는 거시지표 데이터베이스, 시장 규모판단 데이터베이스, 주택시장 데이터베이스, 소비자 데이터베이스의 네 개로 구성되어 있다. 어플리케이션 컴포넌트(Application Component)는 주택 시장 분석 시스템으로 구성되어 있다. 주택 시장 분석시스템에서, 다변량 통계분석과 같은 고급통계분석은 엑셀에서 제공되지 않기 때문에 VBA(Visual Basic for Application)를 이용한 개발 모듈(Module)과 상용 Xlstat 모듈에 의해서 실행 가능하게 되었다. GUI(Graphical User Interface)는 VBA 로 자체 개발하였다. <표 3>에 나타난 네 범주는 정보 데이터베이스내의 데이터베이스로 표현되어 졌으며, 표에 나타난 목적에 맞게끔 설계된 분석기법들이 Xlstat와 개발모듈에 의해서 일종의 데이터 마이닝 시스템으로 구현되어 졌다.

(그림 11)에서 거시지표, 시장규모판단 및 소비자 데이터베이스는 ERD(Entity Relationship Diagram)로 표현되어있고 주택시장 데이터베이스는 스타모양 스키마(Star-shaped Schema)의 형태를 띠고 있다. 즉, 매매가, 전세가의 수치 변수를 가지고 있는 주택사업 사실 테이블(Fact Table)과 조사시간, 브랜드, 평형대, 입지(구와 동)의 차원



(그림 11) 주택 사업 분석 시스템 구성도



(그림 12) 주택시장 데이터베이스의 스타 모양 스키마

테이블(Dimension Table)로 구성되어 있다(그림 12) 참조)[Kimball, 1998].

주택 사업 분석 시스템 구성도는 데이터 웨어하우스 아키텍처(Data Warehouse Architecture)[Lee & Kim 1998]형태로 구성되어 있으며, 현재 추출 컴포넌트(Extracting Component)와 메타데이터(Metadata Component)[이회석 외 1999]는 존재하지 않고 있다. 현재 거시지표 데이터베이스, 시장규모판단 데이터베이스, 주택시장 데이터베이스, 소비자 데이터베이스의 정보 데이터베이스(Informational Database)는 데이터 마트(Data Mart)형태로 구성되어 있으며, 데이터 원천(Source)추가와 스키마 발전에 따라서 부족한 컴포넌트는 개발되어야 한다. 이러한 데이터 마트를 가짐으로서 OLAP(On-line Analytic Processing)도구와 같은 다차원분석도구에 기반을 두고서 의사결정을 위한 유연한 아키텍처가 제공 될 것이다.

모든 데이터베이스의 자료는 계속해서 추가되고, 주택시장 데이터베이스의 자료는 새롭게 추가 및 갱신되어서 모형의 모수값이 수정되어 나타난다. 과거의 주택시장자료도 저장되어서 시계열 분석모형을 추가하기 위한 기반을 제공할 것이다. 어플리케이션 컴포넌트(Application Component)에 해당하는 주택 사업 분석 시스템은 데이터 접근

도구(Data Access Tool)이나 OLAP 도구에 비해서 고급기능을 수행한다. 데이터 마트간의 결합으로 인해서, 주택 정보를 체계적으로 저장할 수 있으며, 네 가지 범주에 맞는, 다양한 분석을 통해서 사업 경영자에게 전략적 의사결정에 도움을 줄 수 있을 것이다.

주화면은 (그림 13)과 같다. 주화면에서 거시지표분석, 시장규모판단, 주택시장분석 및 소비자분석의 네 가지 범주가 화면에 나타나며, 해당 범주를 클릭하면 해당 범주의 분석을 행할 수 있다.



(그림 13) 주택 사업 분석 시스템의 주화면

거시지표분석에는 월별, 분기별 매매가지수에 의한 분석이 나타난다. 월별 5개 지역별 회귀분석

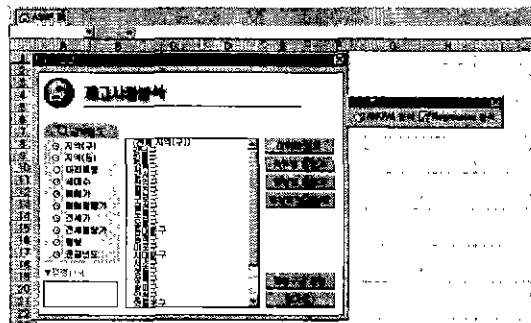
과 분기별 5개 지역별 회귀분석이 나타나서 예측을 할 수 있다.

주화면에서 시장규모판단분석을 선택하면, “년도별 인구 및 세대 증감추이”, “지역별 인구 및 세대현황” 등 7개의 구분(Section)이 나타난다(그림 14)참조).

인구 및 세대 증감 추이			
구	연도	인구	세대
1구	1990		
	1991		
2구	1991		
	1992		
3구	1993		
	1994		
4구	1995		
	1996		
5구	1996		
	1995		
6구	1996		
	1995		
7구	1996		
	1995		
8구	1996		
	1995		
9구	1996		
	1995		
10구	1996		
	1995		
11구	1996		
	1995		
12구	1996		
	1995		
13구	1996		
	1995		
14구	1996		
	1995		
15구	1996		
	1995		
16구	1996		
	1995		
17구	1996		
	1995		
18구	1996		
	1995		
19구	1996		
	1995		
20구	1996		
	1995		
합계			
총계			
1996	379,595	-4%	121,326
1995	479,958		127,462

(그림 14) 시장규모판단분석

주택시장분석은 “재고시장분석” 화면, “경쟁구조 분석” 화면, “브랜드속성 효율치 분석” 화면 및 “GLM” 분석 화면으로 구성되어 있다. “재고시장분석”을 선택하면(그림 15)와 같이 나타난다. “GLM 분석” 화면은 시장진입에 대한 분석, 최소제공평균 등을 나타내어서 예측 평당가와 순위 등을 보여준다.



(그림 15) 재고시장분석 화면

소비자분석은 “테이블 분석” 화면, “PSM” 화면, “Conjoint 분석” 화면, “Correspondence Analysis” 화면 및 “주용도+만족도 분석” 화면 및 “파일변환” 화면으로 구성되어 있다. “Conjoint 분석” 화면을 선택하면(그림 16)과 같이 컨조인트 분석 초

Conjoint 분석					
시뮬레이션	필터	정형	가격	입지	
1	성격	35 평	500만원	산	10
2	연립	25 평	450만원	하	10
3	단독	45 평	550만원	중	10
4	연립	35 평	600만원	상	10

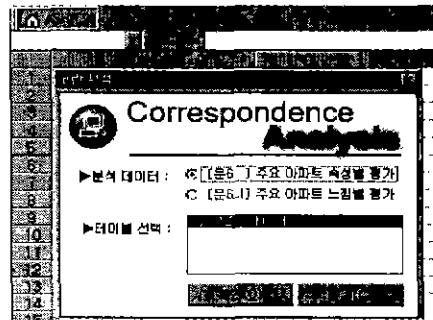
차이제곱	값의 계관	
	400	[분석 시작]

	필터	변수	CHOICE
1번	CHOICE	1	25.00 MAX
2번	CHOICE	1	25.00 MAX
3번	CHOICE	1	25.00 MAX
4번	CHOICE	1	25.00 MAX
총계		4	100.00

BTL	필터	변수	CHOICE
BTL 1	1.00 PART NORTH 1	1.0	필요도 1
BTL 2	1.00 PART NORTH 2	1.0	필요도 2
BTL 3	1.00 PART NORTH 3	1.0	필요도 3
BTL 4	1.00 PART NORTH 4	1.0	필요도 4
총계	4.00	4.0	

(그림 16) Conjoint 분석 화면

기 화면이 나타나며 분석시작 버튼을 누르면 분석을 시작하고 결과는 “점유율 예측 차트”, “BTL 차트”, “속성별 중요도 차트”에 나타난다. “Correspondence Analysis” 화면을 선택하면(그림 17)과 같이 나타난다.



(그림 17) Correspondence Analysis 화면

본 시스템의 사용자는 사용자 수준[Poe 1997] 중에서도 약간의 데이터 마이닝 지식이 있는 파워 사용자(Power User)가 되어야 한다. 즉, 시스템이 초기 단계이기 때문에 사용자가 쉽게 사용하기가 힘들며, 정해진 경로를 따라 분석하는 것도 있지만 새로운 발견을 위해서 비즈니스 분석을 해야하기 때문에 파워 사용자가 되어야 하며, 교육 또한 필요하게 되었다. 경영자들은 마케팅, 시장조사 등

에 대한 연구와 국내에서는 처음으로 아파트 수요 예측 프로그램을 개발, 적용하였다는 데 자부심을 가지고 있다. 아파트의 가격, 평형구성 등이, 본 시스템을 통해 분석된 자료를 토대로 결정되며 중요한 소비자의 의견까지 직접 반영한다는 데 의의를 두고 있다.

본 정보시스템은 분석 모형을 먼저 설계한 후, 그 모형에 기반(Model-based)해서 정보시스템을 구축한 상하향식(Top-down Approach)이기 때문에, 추가적인 연구모형들이 시스템에 추가하기 쉬울 것으로 평가되고 있다. 실제 모형 설명부분의 구성과 그림들이 실제 시스템의 로직(Logic)과 화면들로 이루어져 있다. 또한 기존 연구의 대부분이 본 연구의 “주택시장분석” 범주에 한정되어 있으며, 단면적이면서 일회적인 분석으로 끝나고 있다. 그러나 본 시스템은 시계열적이면서 반복적으로 정보시스템을 이용해서 분석할 수 있으며, 데이터의 관리와 지속적인 데이터 추가가 가능하다. 실제로 본 시스템은 프로젝트의 일환으로 개발된 것으로써, 건축부문에서 의사결정도구로 활용되고 있다. 개념적으로, 본 시스템을 데이터 마이닝의 일종으로 기술하였으며, 특히 통계적 데이터 마이닝 시스템에 초점을 두어 기술되었다.

5. 결 론

본 연구는 질적 대상을 수량화하는 정량적 분석 기법을 도입하였고 이를 정보시스템으로 구축하였기 때문에 지속적으로 사용되고 확장 가능할 것이다. 또한, 본 연구는 주택 분야에 적용되었으며, 데이터 마이닝 시스템을 통합한 데이터 웨어하우스 아키텍처를 이루고 있다.

본 연구는 주택시장 전반에 대한 이론적 기반과 정성적 차원에서 현 주택 시장의 근본적 이해가 부족한 점을 제외하고도 다음과 같은 한계를 가진다. 첫째, 전세가를 수집하였으나 모형에는 포함시키지 않고 전세가와 매매가와의 관계를 파악하지 못한 점, 둘째, 아파트를 제외한 공동 주택 및 단독 주택등을 포함한 전체 주택시장을 대상으로 한

분석이 아니었던 점과 서울지역에 한정된 점, 셋째, 입지를 나타내는 변수로 동과 구만으로 나타내었다는 점, 넷째, 주택 사업 분석 시스템의 불완전성, 다섯째, 다른 기관에서 예측한 자료를 거시 지표분석에 대입해서 사용하기 때문에 자료의 신뢰성 측면에서 단기에측에 적합하며 장기예측에는 적합하지 않다는 점이다.

추후 진행되어야할 연구 과제는 다음과 같다. 첫째, 상용 패키지 SAS 또는 소스코드가 공개된 ViSta[Young, 1996]와 같은 언어를 도입함으로써 시스템의 안정성을 높일 것이다. 둘째, 주택사업자료에 대한 새로운 모형을 고려하고 시계열자료 분석 등이 가능한 모듈이 개발되어야 한다. 시계열자료들이 ARIMA(p,d,q)*(P,D,Q)s에 적합되고 있으나 구현하지는 못했다. 셋째, 전국 아파트 자료에 대한 분석과 모형화 및 전세가와 매매가간의 관계를 들 수 있다. 또한 주택시장 자료의 구조방정식 초기 완전 모형을 연구해야 한다. 넷째, 데이터 웨어하우스 아키텍처 중에서 메타데이터 컴포넌트와 추출 컴포넌트가 개발되어야 한다.

참 고 문 헌

- [1] 김명호, “주택가격 결정요인에 관한 연구”, 단국대학교 지역개발학과 박사학위논문, 1994.
- [2] 김용운, 「서울시 아파트 가격과 전세금의 분석에 관한 연구」, 서울대학교 도시공학과 석사학위논문, 1990.
- [3] 백운봉, 「SAS 일반선형모형 분석」, 자유아카데미, 1990.
- [4] 성내경, 「SAS/STAT-분산분석」, 자유아카데미, 1991.
- [5] 손재영, 김관영, “우리나라 부동산 가격추이의 특징과 전망”, 「국토계획」, 제33권, 제1호, 1998.
- [6] 윤대식, 윤성순, 「도시모형론」, 홍문사, 1995.
- [7] 이경우, 「아파트 특성이 가격결정에 미치는 영향에 관한 실증적 연구: 방향, 경관, 층수를 중심으로」, 경남대학교 경영학과 석사학위논문, 1998.

- [8] 이진목, 「주택시장의 가격구조에 관한 연구 : 서울지역 APT의 매매가격과 전세가격에 대한 실증분석을 중심으로」, 건국대학교 행정대학원 부동산학과 석사학위논문, 1991.
- [9] 이희석, 홍의기, 김태훈, "데이터 웨어하우스를 이용한 임원정보시스템 구축용 아키텍처", 「경영정보학연구」, 제7권, 제1호, 1997, pp.7-24.
- [10] 정홍주, 「아파트 가격결정보형에 관한 실증연구 : 서울지역 한강변 아파트를 중심으로」, 건국대학교 대학원 부동산학과 석사학위논문, 1995.
- [11] 이희석, 조남철, 손명호, 김태훈, "데이터 웨어하우스 메타데이터 스키마 구현 : 홍사 사례", 「Information Systems Review」, 1999년 11월 22일(계재 승인).
- [12] 조주현, "주거밀도가 주택가격에 미치는 영향에 관한 연구", 「건국대학교 사회과학논총」, 1998.
- [13] 허명희, 「SAS 최적척도법」, 자유아카데미, 1994.
- [14] 허명희, "대응분석의 이론과 응용", 「응용통계」, Vol.4, No.2, 1989, pp.23-31.
- [15] Albright, A. S., Winston, W. L., and Zappe, C., *Data Analysis and Decision Making With Microsoft Excel*, Wadsworth Pub. Co., 1999.
- [16] Berry, M. and Linoff, G., *Data Mining Techniques : for Marketing, Sales, and Customer Support*, John Wiley & Sons, 1997.
- [17] Chen, M., Han J. and Yu, P. S., "Data Mining : An Overview from a Database Perspective," *IEEE Transactions on Knowledge and Data Engineering*, Vol.8, No.6, 1996, pp.866-883.
- [18] Elder, J. and Pregibon, D., "A Statistical Perspective on Knowledge Discovery in Databases," *Advances in Knowledge Discovery and Data Mining*, AAAI Press/The MIT Press, pp.83-115, 1996.
- [19] Hou, W., "A Framework for Statistical Data Mining with Summary Tables," *Proc. of 11th IEEE International Conference on Scientific and Statistical Database Management*, 1999, pp.14-23.
- [20] Jöreskog, K. G. and Sörbom, D., *LISREL 7 User's Reference Guide*, Scientific Software Incorporation, 1989.
- [21] Kimball, R., Reeves, L., Ross, M. and Thornthwaite, W., *The Data Warehouse Lifecycle Toolkit : Expert Methods for Designing, Developing, and Deploying Data Warehouses*, John Wiley & Sons, 1998.
- [22] Lee, H. and Kim, T., "A metastore-based data warehouse development methodology," *International Conference on Digital Library and Knowledge*, 1998, pp.448-474.
- [23] Montgomery, D. C., *Design and Analysis of Experiments*, 2nd Edition, John Wiley & Sons, 1984.
- [24] Poe, V., *Building a Data Warehouse for Decision Support*, 2nd Edition, John Wiley & Sons, 1997.
- [25] Roman, S., Roman, S., and Petruscha R., *Access Database Design & Programming*, O'Reilly & Associates, 1999.
- [26] SAS Institute, *SAS User's Guide : Statistics*, 1982 Edition. SAS Institute Inc., 1982.
- [27] Speed, F. M., Hocking, R. R. and Hackney, O. P., "Methods of analysis of linear models with unbalanced data," *American Statistical Association*, Vol.73, 1978, pp.105-112.
- [28] Wonnacott, T. H. and Wonnacott, R. L., *Regression : A Second Course in Statistics*, John Wiley & Sons, 1981.
- [29] Young, F. W., "ViSta : the visual statistics system," *Working Paper*, 1996.

■ 저자소개



김 태 훈

공동저자 김태훈은 고려대 통계학학사 및 동 대학원 석사를 취득하였으며, 현재는 KAIST 테크노경영대학원 경영공학 박사과정을 이수중이다. 주요 관심분야는 데이터 웨어하우스 구축, 분산 데이터베이스 설계, 메타데이터 관리, 데이터 마이닝 설계 등이다.



이 희 석

공동저자 이희석은 서울대 산업공학학사, KAIST 산업공학석사, 그리고 University of Arizona에서 경영정보시스템을 전공하여 경영학박사학위를 취득하였으며 University of Nebraska at Omaha에서 MIS 강의를 하였다. 현재는 KAIST 테크노경영대학원에서 경영정보분야를 연구하고 있다. 주요 관심분야는 지식관리, 데이터베이스, 기업정보시스템, 정보전략, ERP 등이다.



김 재 윤

공동저자 김재윤은 숭실대학교에서 학사를 취득하였으며, (주) J.U.M 해외마케팅팀 및 (주) 솔렉스플랜닝 기획팀 과장으로 근무하였으며, 현재는 부동산 컨설팅 업체인 솔토스피엠에서 과장으로 재직하고 있다. 주요 관심분야는 주택마케팅, 마케팅 리서치 등이다.



전 진 오

공동저자 전진오는 고려대학교 독문학학사, 동대학원 신문방송학 석사를 취득하였으며, (주) 동서리서치 연구부 선임연구원으로 근무하였으며, 현재는 리서치 업체인 A.M.I. (Asia Market Intelligence) Korea 에서 선임연구원으로 재직하고 있다. 주요 관심분야는 마케팅 리서치, 소비자행태 분석 등이다.

이 은 식

공동저자 이은식은 한양대학교 경영학학사를 취득하였으며, 현재는 한국 마이크로데이터를 운영하고 있다. 주요 관심분야는 ERP, 클라이언트/서버 구축, 데이터 웨어하우스 구축 등이다.