



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년09월30일
(11) 등록번호 10-1990799
(24) 등록일자 2019년06월13일

(51) 국제특허분류(Int. Cl.)
G06Q 10/06 (2012.01) G06F 17/18 (2006.01)
G06Q 10/04 (2012.01)
(52) CPC특허분류
G06Q 10/0637 (2013.01)
G06F 17/18 (2013.01)
(21) 출원번호 10-2017-0184151
(22) 출원일자 2017년12월29일
심사청구일자 2017년12월29일
(56) 선행기술조사문헌
KR100883827 B1
(뒷면에 계속)

(73) 특허권자
울산과학기술원
울산광역시 울주군 언양읍 유니스트길 50
(72) 발명자
김정섭
울산광역시 울주군 언양읍 유니스트길 50 울산과
학기술원
박주현
울산광역시 울주군 언양읍 유니스트길 50 울산과
학기술원
(뒷면에 계속)
(74) 대리인
특허법인지원

전체 청구항 수 : 총 15 항

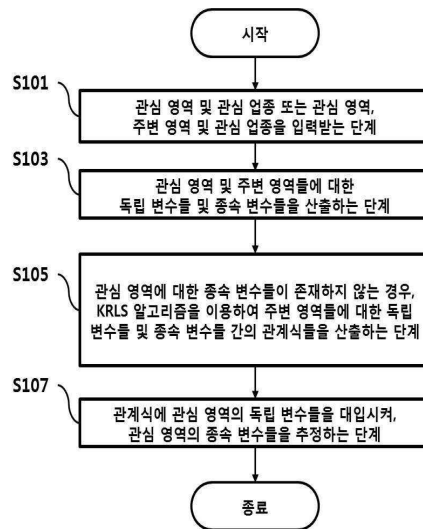
심사관 : 박장환

(54) 발명의 명칭 상권업종별 매출액 및 생존율 예측 방법 및 컴퓨터 판독 가능 기록매체

(57) 요약

본 발명의 일실시예에 따른 상권업종별 매출액 및 생존율 예측 방법은 관심 영역, 상기 관심 영역의 주변 영역들 및 관심 업종을 입력받는 단계; 기저장된 데이터베이스로부터 추출된 데이터들을 가공하여 상기 관심 영역 및 주변 영역들에 대한 독립 변수들 및 종속 변수들을 산출하는 단계; 및 상기 관심 영역 내에 상기 관심 업종이 존재하지 않는 경우, 상기 주변 영역들에 대한 독립 변수들 및 종속 변수들과 관심 영역들에 대한 독립 변수들에 기반하여, 상기 관심 영역에 대한 종속 변수들을 추정하는 단계를 포함한다.

대표도 - 도1



(52) CPC특허분류

G06Q 10/04 (2013.01)

(72) 발명자

최단비

울산광역시 울주군 언양읍 유니스트길 50 울산과학기술원

윤성보

울산광역시 울주군 언양읍 유니스트길 50 울산과학기술원

(56) 선행기술조사문헌

KR1020160050271 A*

정대석 외 1인. “상권 업종별 분포 및 매출 영향 요인 분석”. GRI연구논총. 2014년 8월, 제16권, 제2호, pp.101-122.*

Jens Hainmueller 외 1인. ‘Kernel Regularized Least Squares: Reducing Misspecification Bias with a Flexible and Interpretable Machine Learning Approach’. Political Analysis. October 10, 2013, pp.143-168.*

KR1020110045220 A*

*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호 2015R1C1A1A01053740

부처명 과학기술정보통신부

연구관리전문기관 한국연구재단

연구사업명 신진연구사업

연구과제명 한국형 젠트리피케이션의 진단 및 예측을 위한 행위자기반 모의실험 기술 개발

기 여 율 1/1

주관기관 울산과학기술원

연구기간 2017.11.01 ~ 2018.10.31

명세서

청구범위

청구항 1

컴퓨터에,

관심 영역, 상기 관심 영역에 대응하는 특정 영역들 및 관심 업종을 입력받는 단계;

기저장된 데이터베이스로부터 추출된 데이터들을 가공하여 상기 특정 영역에 대한 독립 변수들, 상기 특정 영역의 관심 업종에 대한 종속 변수들 및 상기 관심 영역에 대한 독립 변수들을 산출하는 단계;

상기 관심 영역 내에 상기 관심 업종이 존재하지 않는 경우, 상기 특정 영역에 대한 독립 변수들과 상기 특정 영역의 관심 업종에 대한 종속 변수들 간의 관계식들을 산출하는 단계; 및

상기 산출된 관계식에 상기 관심 영역에 대한 독립 변수들을 적용시켜, 상기 관심 영역의 관심 업종에 대한 종속 변수들을 추정하는 단계;

를 실행시키기 위하여,

매체에 저장된 컴퓨터 프로그램.

청구항 2

제1항에 있어서,

상기 관계식들을 산출하는 단계는,

KRLS(Kernel-based Regularized Least Squares) 알고리즘을 이용하여 상기 특정 영역에 대한 독립 변수들과 상기 특정 영역의 관심 업종에 대한 종속 변수들 간의 관계식들을 산출하는 단계;

를 포함하는,

매체에 저장된 컴퓨터 프로그램.

청구항 3

제2항에 있어서,

상기 관계식들을 산출하는 단계는,

상기 특정 영역에 대한 독립 변수들과 상기 특정 영역의 관심 업종에 대한 종속 변수들 간의 매칭결과들을 생성하는 단계; 및

상기 KRLS 알고리즘을 상기 생성된 매칭결과들에 적용시킴으로써, 상기 특정 영역에 대한 독립 변수들과 상기 특정 영역의 관심 업종에 대한 종속 변수들 간의 관계식들을 산출하는 단계를 포함하는,

매체에 저장된 컴퓨터 프로그램.

청구항 4

제2항에 있어서,

상기 관계식들을 산출하는 단계는,

상기 KRLS 알고리즘을 이용하여 상기 특정 영역에 대한 독립 변수들과 상기 특정 영역의 관심 업종에 대한 종속

$$y = \sum_{i=1}^n C_i * K(x_i, x)$$

변수들 간의 관계식 을 산출하는 단계를 포함하며,

여기서, y 는 종속 변수들, x_1 내지 x_n 은 독립 변수들, $K(x_i, x)$ 은 특정 독립변수(X_i)와 나머지 독립변수들의 커널 밀도 함수를 나타내며, C_1 내지 C_n 은 상기 KRLS 알고리즘을 통해 산출된 상수들인, 매체에 저장된 컴퓨터 프로그램.

청구항 5

제4항에 있어서,

상기 추정하는 단계는,

상기 산출된 종속 변수들 각각에 대한 관계식의 x_1 내지 x_n 에 대응되는 상기 관심 영역에 대한 독립 변수들을 대입시킴으로써, 상기 관심 영역의 관심 업종에 대한 종속 변수들을 추정하는 단계를 포함하는,

매체에 저장된 컴퓨터 프로그램.

청구항 6

제1항에 있어서,

상기 독립 변수들 및 종속 변수들을 산출하는 단계는,

상기 기저장된 데이터베이스에서 상기 관심 영역 및 특정 영역에 포함된 건물들의 층수 및 용적률을 추출하는 단계;

상기 관심 영역에 포함된 각 건물의 층수와 용적률을 곱한 값들을 합하여 상기 관심 영역에 대한 독립 변수들 중 가중 평균 층수를 산출하는 단계; 및

상기 특정 영역별로, 상기 특정 영역에 포함된 건물들 각각의 층수와 용적률을 곱한 값들을 합하여 상기 특정 영역에 대한 독립 변수들 중 가중 평균 층수를 산출하는 단계를 포함하는,

매체에 저장된 컴퓨터 프로그램.

청구항 7

제6항에 있어서,

상기 독립 변수들 및 종속 변수들을 산출하는 단계는,

상기 기저장된 데이터베이스에서 상기 관심 영역 및 특정 영역의 면적들 및 상기 관심 영역 및 특정 영역에 포함된 건물들의 면적들을 추출하는 단계;

상기 관심 영역에 포함된 건물들의 면적들의 합을 상기 관심 영역의 면적으로 나누어 상기 관심 영역에 대한 독립 변수 중 건폐율을 산출하는 단계; 및

상기 특정 영역별로, 상기 특정 영역에 포함된 건물들의 면적들의 합을 상기 특정 영역의 면적으로 나누어 상기 특정 영역에 대한 독립 변수 중 건폐율을 산출하는 단계를 포함하는,

매체에 저장된 컴퓨터 프로그램.

청구항 8

제7항에 있어서,

상기 독립 변수들 및 종속 변수들을 산출하는 단계는,

상기 관심 영역에 대한 독립변수들 중 건폐율 및 가중 평균 층수를 곱하여 상기 관심 영역에 대한 독립 변수들 중 연면적을 산출하는 단계; 및

상기 특정 영역별로, 상기 특정 영역에 대한 독립변수들 중 건폐율 및 가중 평균 층수를 곱하여 상기 특정 영역에 대한 독립 변수들 중 연면적을 산출하는 단계를 포함하는,

매체에 저장된 컴퓨터 프로그램.

청구항 9

제1항에 있어서,

상기 독립 변수들 및 종속 변수들을 산출하는 단계는,

상기 기저장된 데이터베이스로부터 추출된 상기 특정 영역 각각에 포함된 관심 업종에 대한 사업자 정보 및 휴폐업 정보를 추출하는 단계; 및

상기 추출된 특정 영역 각각에 포함된 관심 업종에 대한 사업자 정보 및 휴폐업 정보에 기반하여, 상기 특정 영역 각각에 포함된 관심 업종에 대한 종속 변수들 중 생존율을 산출하는 단계를 포함하는,

매체에 저장된 컴퓨터 프로그램.

청구항 10

제9항에 있어서,

상기 생존율을 산출하는 단계는,

상기 사업자 정보 및 휴폐업 정보를 통해, 상기 특정 영역 각각에 존재하는, 그리고 존재했던 관심 업종의 수에 대한 정보들을 추출하는 단계; 및

상기 특정 영역별로, 상기 추출된 관심 업종의 수에 대한 상기 추출된 관심 업종의 수 중 상기 특정 영역 각각에 일정기간 이상 존재하고 있고, 일정기간 이상 존재했던 관심 업종의 수의 비를 상기 생존율로 산출하는 단계를 포함하는,

매체에 저장된 컴퓨터 프로그램.

청구항 11

제10항에 있어서,

상기 관심 업종의 수에 대한 정보들을 추출하는 단계는,

상기 사업자 정보 및 휴폐업 정보에 포함된 주소 정보를 통해 위도 및 경도를 추출하는 단계; 및

상기 추출된 위도 및 경도를 상기 특정 영역의 위도 및 경도와 비교함으로써, 각각의 특정 영역에 존재하는, 그리고 존재했던 관심 업종의 수에 대한 정보들을 추출하는 단계를 포함하는,

매체에 저장된 컴퓨터 프로그램.

청구항 12

제1항에 있어서,

상기 독립 변수들 및 종속 변수들을 산출하는 단계는,

상기 기저장된 데이터베이스로부터 상기 관심 영역 및 상기 특정 영역 각각에 포함된 필지들의 지가(Land Price) 및 면적들을 추출하는 단계;

상기 추출된 관심 영역 및 특정 영역 각각에 포함된 필지들의 면적들의 가중치를 산출하는 단계;

상기 관심 영역에 포함된 필지들 각각의 지가 및 면적의 가중치를 곱한 값들을 더하여 상기 관심 영역에 대한 독립 변수들 중 지가를 산출하는 단계; 및

상기 특정 영역별로, 상기 특정 영역 각각에 포함된 필지들의 지가 및 면적의 가중치를 곱한 값들을 더하여 상기 특정 영역 각각에 대한 독립 변수들 중 지가를 산출하는 단계를 포함하는,

매체에 저장된 컴퓨터 프로그램.

청구항 13

제12항에 있어서,

상기 면적들의 가중치를 산출하는 단계는,

상기 관심 영역에 포함된 필지들의 면적들을 상기 관심 영역의 면적으로 나누어 상기 관심 영역에 포함된 필지들의 면적의 가중치를 산출하는 단계; 및

상기 특정 영역별로, 상기 특정 영역 각각에 포함된 필지들의 면적들을 상기 특정 영역의 면적으로 나누어, 상기 특정 영역 각각에 포함된 필지들의 면적의 가중치를 산출하는 단계를 포함하는,

매체에 저장된 컴퓨터 프로그램.

청구항 14

제1항에 있어서,

상기 독립 변수들은,

다중공선성 관계에 있는 변수들을 포함하는,

매체에 저장된 컴퓨터 프로그램.

청구항 15

제1항에 있어서,

상기 기저장된 데이터베이스는,

상기 관심 영역 및 특정 영역에 포함된 관심 업종들에 대한 종속 변수들 및 상기 관심 영역 및 특정 영역에 대한 독립 변수들을 산출하기 위한 데이터들을 포함하는,

매체에 저장된 컴퓨터 프로그램.

청구항 16

삭제

발명의 설명

기술분야

[0001] 본 발명은 상권업종별 매출액 및 생존율 예측 방법 및 컴퓨터 판독 가능한 기록매체에 관한 것으로서, KRLS(Kernel-based Regularized Least Squares) 통계기법을 이용한 상권업종별 매출액 및 생존율 예측 방법 및 컴퓨터 판독 가능 기록매체이다.

배경기술

[0002] 2016년 기준으로 창업 후 3년 이내 폐업률은 85%에 육박하면서, 생활밀착형 또는 골목상권에 해당하는 자영업자의 폐업률이 증가하고 있다.

[0003] 폐업률이 증가하는 원인은 대형 프랜차이즈 및 창업 전문가에 비하여 생활밀착형 또는 골목상권에 해당하는 자영업자의 경우에는 상권 분석의 전문성이 떨어지기 때문이다.

[0004] 현재 우리마을 가게 상권분석 서비스에서 생활밀착형 자영업자들에게 상권 분석 정보를 제공하고 있지만, 소극적인 정보만을 제공하고 있는 실정이다.

[0005] 이러한 문제를 해결하기 위하여 종래에는 내/외부 데이터를 이용하여 지역 내 업종 매출 추정 방정식을 구축하고, 구축한 매출 추정 방정식에 의해 상권 내에서 특정 업종의 총 매출을 추정하는 상권 내에서의 업종별 매출 추정 방법이 개시되었다.

[0006] 하지만, 단순선형회귀분석(Ordinary Least Square, OLS)을 이용하여 매출액을 추정하는 경우는 변수 간 관계를 선형으로 가정하기 때문에 예측력이 떨어지고, 다중공선성관계에 있는 변수들을 고려할 수 없다는 문제점이 있다.

[0007] 또한, 매출액만을 추정하고 있을 뿐, 다양한 변수들을 통한 자영업의 생존율을 추정하지 못하므로, 폐업률을 줄이기 위한 실질적인 상권 분석이 이루어지지 않고 있는 문제점이 있다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) KR 10-1078344 B1

발명의 내용

해결하려는 과제

[0009] 본 발명은 KRLS 통계기법을 이용함으로써, 관심 영역에 관심 업종이 존재하지 않는 경우에도 주변 영역들에 대한 변수들에 기반하여, 관심 영역에 대한 관심 업종의 매출액 및 생존율을 추정할 수 있는 상권업종별 매출액 및 생존율 예측 방법 및 컴퓨터 판독 가능 기록매체이다.

과제의 해결 수단

[0010] 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법은 관심 영역, 상기 관심 영역의 주변 영역들 및 관심 업종을 입력받는 단계; 기저장된 데이터베이스로부터 추출된 데이터들을 가공하여 상기 관심 영역 및 주변 영역들에 대한 독립 변수들 및 종속 변수들을 산출하는 단계; 및 상기 관심 영역 내에 상기 관심 업종이 존재하지 않는 경우, 상기 주변 영역들에 대한 독립 변수들 및 종속 변수들과 관심 영역들에 대한 독립 변수들에 기반하여, 상기 관심 영역에 대한 종속 변수들을 추정하는 단계를 포함할 수 있다.

[0011] 상기 추정하는 단계는, KRLS(Kernel-based Regularized Least Squares) 알고리즘을 이용하여 상기 주변 영역들에 대한 종속 변수들 및 독립 변수들 간의 관계식들을 산출하는 단계; 및 상기 산출된 관계식에 상기 관심 영역에 대한 독립 변수들을 적용시켜, 상기 관심 영역에 대한 종속 변수들을 추정하는 단계를 포함할 수 있다.

[0012] 상기 관계식들을 산출하는 단계는, 상기 주변 영역들 각각에 대한 종속 변수들 및 독립 변수들 간의 매칭결과들을 생성하는 단계; 및 상기 KRLS 알고리즘을 상기 생성된 매칭결과들에 적용시킴으로써, 상기 주변 영역들에 대한 종속 변수들 및 독립 변수들 간의 관계식들을 산출하는 단계를 포함할 수 있다.

[0013] 상기 관계식들을 산출하는 단계는, 상기 KRLS 알고리즘을 이용하여 상기 주변 영역들에 대한 종속 변수들 및 독

$$y = \sum_{i=1}^n C_i * K(x_i, x)$$

립 변수들 간의 관계식 을 산출하는 단계를 포함하며, 여기서, y는 종속 변수들,

$K(x_i, x)$ 은 특정 독립변수(x_i)와 나머지 독립변수들의 커널 밀도 함수를 나타내며, C_1 내지 C_n 은 상기 KRLS 알고리즘을 통해 산출된 상수들이다.

[0014] 상기 추정하는 단계는, 상기 산출된 종속 변수들 각각에 대한 관계식의 x_1 내지 x_n 에 대응되는 상기 관심 영역에 대한 독립 변수들을 대입시킴으로써, 상기 관심 영역에 대한 종속 변수들을 추정하는 단계를 포함할 수 있다.

[0015] 상기 독립 변수들 및 종속 변수들을 산출하는 단계는, 상기 기저장된 데이터베이스에서 상기 관심 영역 및 주변 영역들에 포함된 건물들의 층수 및 용적률을 추출하는 단계; 상기 관심 영역에 포함된 각 건물의 층수와 용적률 가중치를 곱한 값들을 합하여 상기 관심 영역에 대한 독립 변수들 중 가중 평균 층수를 산출하는 단계; 및 상기 주변 영역들별로, 상기 주변 영역에 포함된 건물들 각각의 층수와 용적률 가중치를 곱한 값들을 합하여 상기 주변 영역에 대한 독립 변수들 중 가중 평균 층수를 산출하는 단계를 포함할 수 있다.

[0016] 상기 독립 변수들 및 종속 변수들을 산출하는 단계는, 상기 기저장된 데이터베이스에서 상기 관심 영역 및 주변 영역들의 면적들 및 상기 관심 영역 및 주변 영역들에 포함된 건물들의 면적들을 추출하는 단계; 상기 관심 영역에 포함된 건물들의 면적들의 합을 상기 관심 영역의 면적으로 나누어 상기 관심 영역에 대한 독립 변수 중 건폐율을 산출하는 단계; 및 상기 주변 영역들별로, 상기 주변 영역에 포함된 건물들의 면적들의 합을 상기 주변 영역의 면적으로 나누어 상기 주변 영역에 대한 독립 변수 중 건폐율을 산출하는 단계를 포함할 수 있다.

[0017] 상기 독립 변수들 및 종속 변수들을 산출하는 단계는, 상기 관심 영역에 대한 독립변수들 중 건폐율 및 가중 평균 층수를 곱하여 상기 관심 영역에 대한 독립 변수들 중 연면적을 산출하는 단계; 및 상기 주변 영역들별로, 상기 주변 영역에 대한 독립변수들 중 건폐율 및 가중 평균 층수를 곱하여 상기 주변 영역에 대한 독립 변수들 중 연면적을 산출하는 단계를 포함할 수 있다.

[0018] 상기 독립 변수들 및 종속 변수들을 산출하는 단계는, 상기 기저장된 데이터베이스로부터 추출된 상기 주변 영역들 각각에 포함된 관심 업종에 대한 사업자 정보 및 휴폐업 정보를 추출하는 단계; 및 상기 추출된 주변 영역들 각각에 포함된 관심 업종에 대한 사업자 정보 및 휴폐업 정보에 기반하여, 상기 주변 영역들 각각에 포함된 관심 업종에 대한 종속 변수들 중 생존율을 산출하는 단계를 포함할 수 있다.

[0019] 상기 관심 업종의 수에 대한 정보들을 추출하는 단계는, 상기 사업자 정보 및 휴폐업 정보에 포함된 주소 정보를 통해 위도 및 경도를 추출하는 단계; 및 상기 추출된 위도 및 경도를 상기 주변 영역들의 위도 및 경도와 비교함으로써, 각각의 주변 영역에 존재하는, 그리고 존재했던 관심 업종의 수에 대한 정보들을 추출하는 단계를 포함할 수 있다.

[0020] 상기 독립 변수들 및 종속 변수들을 산출하는 단계는, 상기 기저장된 데이터베이스로부터 상기 관심 영역 및 상기 주변 영역들 각각에 포함된 필지들의 지가(Land Price) 및 면적들을 추출하는 단계; 상기 추출된 관심 영역 및 주변 영역들 각각에 포함된 필지들의 면적들의 가중치를 산출하는 단계; 상기 관심 영역에 포함된 필지들 각각의 지가 및 면적의 가중치를 곱한 값들을 더하여 상기 관심 영역에 대한 독립 변수들 중 지가를 산출하는 단계; 및 상기 주변 영역들별로, 상기 주변 영역들 각각에 포함된 필지들의 지가 및 면적의 가중치를 곱한 값들을 더하여 상기 주변 영역들 각각에 대한 독립 변수들 중 지가를 산출하는 단계를 포함할 수 있다.

[0021] 상기 면적들의 가중치를 산출하는 단계는, 상기 관심 영역에 포함된 필지들의 면적들을 상기 관심 영역의 면적으로 나누어 상기 관심 영역에 포함된 필지들의 면적의 가중치를 산출하는 단계; 및 상기 주변 영역들별로, 상기 주변 영역들 각각에 포함된 필지들의 면적들을 상기 주변 영역의 면적으로 나누어, 상기 주변 영역들 각각에 포함된 필지들의 면적의 가중치를 산출하는 단계를 포함할 수 있다.

[0022] 상기 기저장된 데이터베이스는, 상기 관심 영역 및 주변 영역들에 포함된 관심 업종들에 대한 종속 변수들 및 상기 관심 영역 및 주변 영역에 대한 독립 변수들을 산출하기 위한 데이터들을 포함할 수 있다. 또한 위 변수들은 다중공선성 관계에 있을 확률이 높다. 예를 들어, 특정 공간의 건물 층수가 높은 경우는 건물의 연면적이나

용적률, 지가 또한 높을 가능성이 높기 때문이다.

[0023] 상기 생존율을 산출하는 단계는, 상기 사업자 정보 및 휴폐업 정보를 통해, 상기 주변 영역들 각각에 존재하는, 그리고 존재했던 관심 업종의 수에 대한 정보들을 추출하는 단계; 및 상기 주변 영역들별로, 상기 추출된 관심 업종의 수에 대한 상기 추출된 관심 업종의 수 중 상기 주변 영역들 각각에 일정기간 이상 존재하고 있고, 일정기간 이상 존재했던 관심 업종의 수의 비를 상기 생존율로 산출하는 단계를 포함할 수 있다.

[0024] 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법을 실행하기 위한 컴퓨터 프로그램을 기록한 컴퓨터 판독 가능한 기록 매체를 포함할 수 있다.

발명의 효과

[0025] 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법 및 컴퓨터 판독 가능 기록매체에 따르면, KRLS 통계기법을 이용하기 때문에 다중공선성 관계에 있는 지표들을 통해서 매출액 및 생존율을 추정할 수 있는 효과가 있다.

[0026] 또한, KRLS 통계기법을 이용하기 때문에, 관심 영역에 관심 업종이 없더라도, 주변 영역에 대한 변수들에 기반하여, 관심 영역에 대한 관심 업종의 매출액 및 생존율을 추정할 수 있는 효과가 있다.

도면의 간단한 설명

[0027] 본 발명에 관한 이해를 돕기 위해 상세한 설명의 일부로 포함되는, 첨부 도면은 본 발명에 대한 실시예를 제공하고, 상세한 설명과 함께 본 발명의 기술적 특징을 설명한다.

도 1은 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법의 순서도를 간략히 도시한 도면이다.

도 2는 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법에 있어서, 관심 업종의 예시를 도시한 도면이다.

도 3은 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법에 있어서, 관심 영역 및 주변 영역들의 예시를 도시한 도면이다.

도 4는 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법에 있어서, 단순선형회귀분석(OLS)과 KRLS 알고리즘을 이용하여 매출액을 업체 수를 통해 추정한 데이터이다.

도 5는 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법에 있어서, 단순선형회귀분석과 KRLS 알고리즘의 예상 매출액을 비교한 데이터이다.

도 6은 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법에 있어서, 단순선형회귀분석과 KRLS 알고리즘의 예상 생존율을 비교한 데이터이다.

도 7은 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법에 있어서, 관심 영역 및 주변 영역에 대한 매출액과 보조 지표들의 통계를 산출한 데이터이다.

발명을 실시하기 위한 구체적인 내용

[0028] 본 명세서에서 제1 및/또는 제2 등의 용어는 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 즉, 구성요소들을 상기 용어들에 의해 한정하고자 함이 아니다.

[0029] 본 명세서에서 '포함하다' 라는 표현으로 언급되는 구성요소, 특징, 및 단계는 해당 구성요소, 특징 및 단계가 존재함을 의미하며, 하나 이상의 다른 구성요소, 특징, 단계 및 이와 동등한 것을 배제하고자 함이 아니다.

[0030] 본 명세서에서 단수형으로 특정되어 언급되지 아니하는 한, 복수의 형태를 포함한다. 즉, 본 명세서에서 언급된 구성요소 등은 하나 이상의 다른 구성요소 등의 존재나 추가를 의미할 수 있다.

[0031] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함하여, 본 명세서에서 사용되는 모든 용어들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자(통상의 기술자)에 의하여 일반적으로 이해되는 것과 동일한 의미이다.

[0032] 즉, 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하

는 의미인 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

- [0033] 이하에서는, 첨부된 도면을 참조하여 본 발명의 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법 및 컴퓨터 판독 가능 기록매체에 대해 상세하게 설명한다.
- [0034] 도 1은 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법의 순서도를 간략히 도시한 도면이다.
- [0035] 도 1을 참조하면, 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법은 입력받는 단계(S101), 독립 변수들 및 종속 변수들을 산출하는 단계(S103), 관계식들을 산출하는 단계(S105) 및 종속 변수들을 추정하는 단계(S107)를 포함할 수 있다.
- [0036] 입력받는 단계(S101)는 관심 영역, 주변 영역 및 관심 업종을 입력받는 단계이다.
- [0037] 예컨대, 관심 영역 및 주변 영역들은 블록단위로 설정될 수 있고, 관심 업종은 편의점, 호프집 및 분식집 등의 업종으로 설정될 수 있다.
- [0038] 주변 영역들은 선택된 시, 도, 군, 동 등의 행정자치구 내의 모든 블록들로 설정되거나, 개별적으로 선택된 복수 개의 블록들로 설정될 수 있다.
- [0039] 독립 변수들 및 종속 변수들을 산출하는 단계(S103)는 기저장된 데이터베이스로부터 추출된 데이터들을 가공하여, 관심 영역 및 주변 영역들에 대한 독립 변수들 및 종속 변수들을 산출하는 단계이다.
- [0040] 기저장된 데이터베이스는 외부의 서버로서, 기저장된 데이터베이스로부터 추출한 데이터들을 가공하여 관심 영역 및 주변 영역들에 대한 독립 변수들 및 관심 업종에 대한 종속 변수들을 산출할 수 있다.
- [0041] 예컨대, 독립 변수들은 가중 평균 층수, 건폐율, 연면적, 상권 여부, 대학가 여부, 역세권 여부, 업체 수, 지가(Land Price), 필지 용도의 다양성, 주거, 상업 및 공공 필지 비율 및 나이별 거주, 유동 및 직장인구 등을 포함할 수 있고, 종속 변수들은 관심 영역 및 주변 영역들에 대한 관심 업종의 매출액 또는 생존율 중 적어도 하나를 포함할 수 있다.
- [0042] 즉, 독립 변수들은 영역에 대한 변수들로서, 관심 업종의 존재 여부에 영향을 받지 않는다. 반면에, 종속 변수들은 특정 영역에 대한 관심 업종의 매출액 및 생존율이므로, 선택된 특정 영역에 관심 업종이 존재하지 않으면 종속 변수들은 산출되지 않는다.
- [0043] 종속 변수들 중 생존율은 관심 영역에 존재하고 있거나 과거에 존재했었던 관심 업종의 수에 대한 관심 영역에 일정기간 이상 존재하고 있거나 과거에 존재했었던 관심 업종의 수의 비이다. 예컨대, 생존율은 A블록에 존재한 편의점의 수에 대한 A블록에서 3년 이상 유지된 편의점의 수에 대한 비일 수 있다.
- [0044] 관심 영역에 대한 블록 내 평균 층수는, 기저장된 데이터베이스에서 관심 영역에 포함된 건물들의 층수 및 용적률을 추출하고, 각 건물의 층수와 용적률을 곱한 값들을 합하여 관심 영역에 대한 가중 평균 층수를 산출할 수 있다.
- [0045] 단순히 관심 영역에 포함된 건물들의 층수의 평균을 산출할 경우, 건물의 용적을 고려하지 않게 되는 문제가 생기기 때문에, 건물들의 용적률을 가중치로하여 관심 영역에 대한 가중 평균 층수를 산출할 수 있다.
- [0046] 관심 영역에 대한 블록의 건폐율은 기저장된 데이터베이스에서 관심 영역의 면적 및 관심 영역에 포함된 건물들의 면적을 추출하고, 건물들의 면적의 합을 관심 영역의 면적으로 나누어 산출될 수 있다.
- [0047] 관심 영역에 대한 연면적은 관심 영역에 대한 건폐율 및 가중 평균 층수를 곱하여 산출될 수 있다.
- [0048] 관심 영역에 대한 상권 여부, 대학가 여부, 및 역세권 여부는 관심 영역이 주변 일반상업지구, 대학 및 지하철역으로부터 일정 거리 이내에 있는지 여부를 통해 산출될 수 있다.
- [0049] 관심 영역에 대한 지가는 기저장된 데이터베이스로부터 관심 영역에 포함된 필지들의 지가 및 면적들을 추출하고, 관심 영역에 포함된 필지들의 면적의 가중치를 산출한다. 여기서 필지들의 면적의 가중치는 필지들의 면적들을 관심 영역의 면적으로 나누어 산출될 수 있다.
- [0050] 그 후, 산출된 필지들 면적의 가중치들 각각을 대응되는 필지들의 지가와 곱한 값들을 더하여 관심 영역에 대한 지가가 산출될 수 있다.

- [0051] 관심 영역에 대한 필지 용도의 다양성은 각 필지의 용도가 일반상업지역, 제3종일반주거지역인지 등 용도가 다양한지 여부에 따라 필지 용도의 다양성을 지수화하여 산출될 수 있다.
- [0052] 관심 영역에 대한 주거, 상업 및 공공 필지 비율은 관심 영역의 면적에서 주거, 상업, 공공 필지들 각각이 차지하는 비율을 통해 산출될 수 있다.
- [0053] 관심 영역에 대한 세대별 거주, 유동 및 직장 인구는 기저장된 데이터베이스로부터 추출할 수 있다.
- [0054] 기저장된 데이터베이스에서 추출된 데이터들 중 관심 영역에 포함된 관심 업종에 대한 사업자 정보 및 휴폐업 정보가 존재하는 경우, 사업자 정보 및 휴폐업 정보에 기반하여 관심 영역에 포함된 관심 업종에 대한 생존율이 산출될 수 있다.
- [0055] 먼저, 사업자 정보 및 휴폐업 정보에 포함된 주소를 통해 위도 및 경도를 추출하고, 추출된 위도 및 경도를 관심 영역에 대응시켜 비교함으로써, 관심 영역에 존재하거나 또는 존재했던 관심 업종의 수를 산출한다.
- [0056] 그 후, 관심 영역에 존재하거나 또는 존재했던 관심 업종의 수에 대한 일정 기간 이상 관심 영역에 존재하거나 또는 존재했던 관심 업종의 비를 산출함으로써, 관심 영역에 포함된 관심 업종에 대한 생존율이 산출될 수 있다.
- [0057] 관심 영역에 포함된 관심 업종에 대한 매출액은 기저장된 데이터베이스로부터 추출될 수 있다.
- [0058] 이상에서 설명한 관심 영역에 대한 독립 변수들 각각을 추출 또는 산출하는 방법을 주변 영역들 각각에 적용시킴으로써, 주변 영역들 각각에 대한 독립 변수들을 산출할 수 있다.
- [0059] 관계식을 산출하는 단계(S105)는 관심 영역에 관심 업종이 없는 경우, KRLS 알고리즘을 이용하여 주변 영역들에 대한 종속 변수들 및 독립 변수들 간의 관계식을 산출하는 단계이다.
- [0060] KRLS(Kernel-based Regularized Least Squares)는 Hainmueller & Hazlett가 제시한 알고리즘으로 유사한 공변량 값을 가진 관측자는 평균적으로 유사한 결과를 가질 것이라는 가정에 바탕한 통계기법이다.
- [0061] KRLS 알고리즘은 변수들 간의 선형성을 가정하지 않고 다중공선성 관계의 독립변수들을 모두 활용 할 수 있기 때문에, 종래의 상권분석에서 사용된 단순선형회귀분석(OLS)보다 정확하게 상권업종별 매출액 및 생존율을 추정할 수 있다.
- [0062] 단순선형회귀분석(Ordinary Least Square, OLS)의 경우 변수가 가지는 한계 효과가 공변량 영역(covariate space)에서 일정하다는 가정에 기반한 통계기법으로, 단순선형회귀분석의 경우 해석이 직관적이고 방법 자체가 쉬운 편이기 때문에, 많은 추정 모형에서 사용되고 있다.
- [0063] 다만, 단순선형회귀분석은 독립 변수와 종속 변수 간의 선형적 모형을 추정할 수 밖에 없는 문제가 있어, 상권업종별 매출액 및 생존율을 추정하는 데는 부적절하다.
- [0064] 예컨대, A블록에 옷가게가 1개가 있는 경우보다, 옷가게가 2개 또는 3개 있는 경우 A블록에 포함된 옷가게들의 매출액이 더 잘 나올 것이다. 하지만, 옷가게가 일정 수준 이상으로 많아진다면, 옷가게 각각의 매출액은 떨어질 수 밖에 없는데, 단순선형회귀분석은 이러한 비선형적 모형을 추정할 수 없는 문제가 있다.
- [0065] 또한, 단순선형회귀분석은 상권 분석에서는 다중공선성 관계에 있는 변수들을 인자로서 사용할 수 없는 문제가 있다.
- [0066] 예컨대, 10대 거주인구와 20대 거주인구는 밀접한 관련이 있기 때문에, 단순선형회귀분석(Ordinary Least Square, OLS)에서 10대 거주인구와 20대 거주인구를 인자로 사용할 경우, 10대 거주인구와 20대 거주인구 간의 연관성에 의하여 도출된 결과에 치명적인 오류가 생길 수 있다.
- [0067] KRLS 알고리즘을 이용하여 관계식을 산출하기 위해서, 먼저 주변 영역들 각각에 대한 종속 변수들 및 독립 변수들 간의 매칭결과들을 추출하고, KRLS 알고리즘을 매칭결과들에 적용시킴으로써, 주변 영역들에 대한 독립 변수들 및 주변 영역들에 포함된 관심 영역에 대한 종속 변수들 간의 관계식을 산출할 수 있다.

[0068] KRLS 알고리즘을 이용하여 산출된 주변 영역들에 대한 종속 변수들 및 독립 변수들 간의 관계식은 다음과 같다.

$$y = \sum_{i=1}^n C_i * K(x_i, x)$$

[0069]

[0070] 여기서, y 는 종속 변수들, x_1 내지 x_n 은 독립 변수들, $K(x_i, x)$ 은 특정 독립 변수(x_i)와 나머지 독립변수 간의 커널 밀도 함수를 나타내며, C_1 내지 C_n 은 KRLS 알고리즘을 통해 산출된 상수들이다.

[0071] KRLS 알고리즘을 주변 영역들에 적용함으로써, 종속 변수들 각각에 대한 관계식들이 산출될 수 있으며, 종속 변수들에 대한 관계식들 각각은 C_1 내지 C_n 의 상수 범위에서 차이가 있다.

[0072] 종속 변수들을 추정하는 단계(S107)는 산출된 관계식들에 관심 영역에 대한 독립 변수들을 대입함으로써, 관심 업종에 대한 종속 변수들을 추정하는 단계이다.

[0073] 예컨대, 독립 변수들은 가중 평균 층수, 유동인구 등 관심 영역 또는 주변 영역들에 관련된 변수들이기 때문에, 관심 영역에 관심 업종이 존재하지 않더라도 독립 변수들은 존재한다.

[0074] 따라서, KRLS 알고리즘을 통해 산출된 종속 변수들에 대한 관계식들에 관심 영역에 대한 독립 변수들을 대입시킴으로써, 관심 영역에 대한 종속 변수들을 추정할 수 있다.

[0075] 즉, 관심 영역에 대한 독립 변수들을 KRLS 알고리즘을 통해서 산출된 매출액 관계식의 x_1 내지 x_n 에 대입시키면 관심 영역에 대한 관심 업종의 예상 매출액이 추정되고, KRLS 알고리즘을 통해서 산출된 생존율 관계식의 x_1 내지 x_n 에 대입시키면 관심 영역에 대한 관심 업종의 예상 생존율이 추정된다.

[0076] 단순선형회귀분석(OLS) 통계기법을 통해 산출되는 관계식은 종속 변수(y)와 독립 변수들(x_1 내지 x_n)이 선형 관계를 가져야 하기 때문에, 비선형적인 관계를 고려할 수 없는 문제가 없지만, KRLS 알고리즘은 커널 밀도 함수를 활용하여 비선형적 관계를 고려할 수 있으므로, 단순선형회귀분석(OLS)에 비하여 정확하게 예상 매출액 및 생존율을 추정할 수 있다.

[0077] 또한, 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법은 추출하는 단계(S103)이후에 관심 영역에 대한 종속 변수들이 존재하는지를 비교하는 것과 병렬적으로, 기저장된 데이터베이스에서 추출한 종속 변수 및 독립 변수들 간의 통계를 제공할 수 있다.

[0078] 예컨대, 사용자가 관심 영역으로 A블록, 주변 영역으로 B, C 및 D블록, 관심 업종으로 편의점업을 선택한 경우, A, B, C 및 D블록에 포함된 편의점업의 매출액과 독립 변수들 간의 통계, 생존율과 독립 변수들 간의 통계가 제공될 수 있다.

[0079] 도 2는 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법에 있어서, 관심 업종의 예시를 도시한 도면이다.

[0080] 도 3은 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법에 있어서, 관심 영역 및 주변 영역들의 예시를 도시한 도면이다.

[0081] 사용자는 관심 업종들(101)에서 복수의 관심 업종들을 선택할 수 있으며, 관심 영역(103) 및 주변 영역들(105)을 선택할 수 있다.

[0082] 예컨대, 사용자가 관심 업종(101)으로 편의점을 선택하였으나, 선택한 관심 영역(103)에 편의점업이 없는 경우, KRLS 알고리즘을 이용하여 주변 영역들(105) 중 편의점업이 존재하는 주변 영역들로부터 편의점업에 대한 매출액 및 생존율에 대한 관계식들을 추출할 수 있다.

[0083] 추출된 관계식들에 관심 영역(103)의 보조 지표들을 대입시킴으로써, 사용자가 선택한 관심 영역에 관심 업종이 존재하지 않더라도, 관심 영역에서 관심 업종이 존재할 경우의 매출액 및 생존율을 추정할 수 있다.

[0084] 도 4는 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법에 있어서, 단순선형회귀분석(OLS)과 KRLS 알고리즘을 이용하여 매출액을 업체 수를 통해 추정할 데이터이다.

[0085] 도 4를 참조하면, 단순선형회귀분석을 이용한 업체 수에 대한 매출액 그래프(401)는 업체 수가 증가할수록, 매

출액이 증가하는 선형 형태를 나타내지만, KRLS를 이용한 업체 수에 대한 매출액 그래프(403)은 업체 수와 매출액의 비선형적 관계를 추정할 수 있다.

- [0086] 예컨대, A 블록에 포함된 분식집이 1개 있는 것 보다는 2, 3개 있을 경우 매출액이 증가하지만, 분식집이 일정 수준 이상으로 많아지게 되면 분식집들 각각의 매출액은 감소하게 된다. 단순선형회귀분석은 이렇게 분식집들이 일정 수준 이상 많아지는 경우에도 매출액이 증가된다고 추정하기 때문에, 상권업종별 매출액 및 생존율을 추정하는데 한계가 있다.
- [0087] 반면, KRLS는 업체 수와 매출액의 비선형적관계를 추정할 수 있기 때문에, 상권업종별 매출액 및 생존율을 적절하게 추정할 수 있다.
- [0088] 도 5는 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법에 있어서, 단순선형회귀분석과 KRLS 알고리즘의 예상 매출액을 비교한 데이터이다.
- [0089] 도 5를 참조하면, 단순선형회귀분석을 이용한 예상 매출액 그래프(501)보다 KRLS를 이용한 예상 매출액 그래프(503)의 예상 매출액이 $y=x$ 직선(빨간색 직선)에 더 가깝게 모여 있다.
- [0090] 이를 통해서, KRLS를 이용하여 예상 매출액을 추정하는 것이 OLS를 이용하여 예상 매출액을 추정하는 것보다 실제 매출액을 더 잘 예측하는 것을 알 수 있다.
- [0091] 도 6은 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법에 있어서, 단순선형회귀분석과 KRLS 알고리즘의 예상 생존율을 비교한 데이터이다.
- [0092] 도 6을 참조하면, 단순선형회귀분석을 이용한 예상 생존율 그래프(601)보다 KRLS를 이용한 예상 생존율 그래프(603)의 예상 생존율이 $y=x$ 직선(빨간색 직선)에 더 가깝게 모여 있다.
- [0093] 이를 통해서, KRLS를 이용하여 예상 생존율을 추정하는 것이 단순선형회귀분석을 이용하여 예상 생존율을 추정하는 것보다 실제 생존율을 더 잘 예측하는 것을 알 수 있다.
- [0094] 도 7은 본 발명의 일 실시예에 따른 상권업종별 매출액 및 생존율 예측 방법에 있어서, 관심 영역 및 주변 영역에 대한 매출액과 보조 지표들의 통계를 산출한 데이터이다.
- [0095] 도 7을 참조하면, 편의점 업과 연령별 거주인구 간의 상관관계에 대한 통계가 산출될 수 있다.
- [0096] 연령별 거주인구에 대한 독립 변수들은 다중공선성 관계에 있는 변수들이기 때문에, 단순선형회귀분석에서는 사용될 수 없는 변수들이지만 KRLS 알고리즘을 이용하면 각 독립 변수와 종속 변수들 간의 통계를 추출할 수 있다.
- [0097] 추출된 통계를 보면, 10대 및 50대 거주인구는 편의점의 매출액과 큰 관련이 없지만, 20대, 30대 및 60대 거주인구가 증가함에 따라 편의점의 매출액이 증가하는 패턴을 보이지만, 40대 거주인구는 증가할수록 편의점 매출액은 감소하는 패턴을 알 수 있다.
- [0098] 본 명세서에서 설명된 실시예들에 관한 예시적인 모듈, 단계, 과정, 로직 블록, 수단, 단계 또는 이들의 조합은 전자 하드웨어(코딩 등에 의해 설계되는 디지털 설계), 소프트웨어(프로그램 명령을 포함하는 다양한 형태의 애플리케이션) 또는 이들의 조합에 의해 구현될 수 있다. 하드웨어 및/또는 소프트웨어 중 어떠한 형태로 구현되는지는 사용자 단말에 부여되는 설계상의 제약에 따라 달라질 수 있다. 또한, 본 명세서에서 설명된 기능들은 하드웨어, 소프트웨어, 펌웨어 또는 이들의 조합에 의해 실행될 수 있다. 소프트웨어에 의해 실행되는 경우, 그 기능은 하나 이상의 명령이나 코드로서 컴퓨터-판독가능 저장 매체에 저장되거나 이를 통해 전송될 수 있다. 컴퓨터-판독가능 저장 매체는 컴퓨터에 의해 액세스 가능한 임의의 이용가능한 매체를 총괄적으로 지칭한다.
- [0099] 비록 본 명세서에서의 설명은 예시적인 몇 가지 양상으로 나타났지만, 다양한 수정이나 변경이 후술되는 특허청구범위에 의해 정의되는 범주로부터 이루어질 수 있으며, 본 발명의 기술적인 보호범위는 다음의 특허청구범위에 의하여 정해져야 할 것이다.

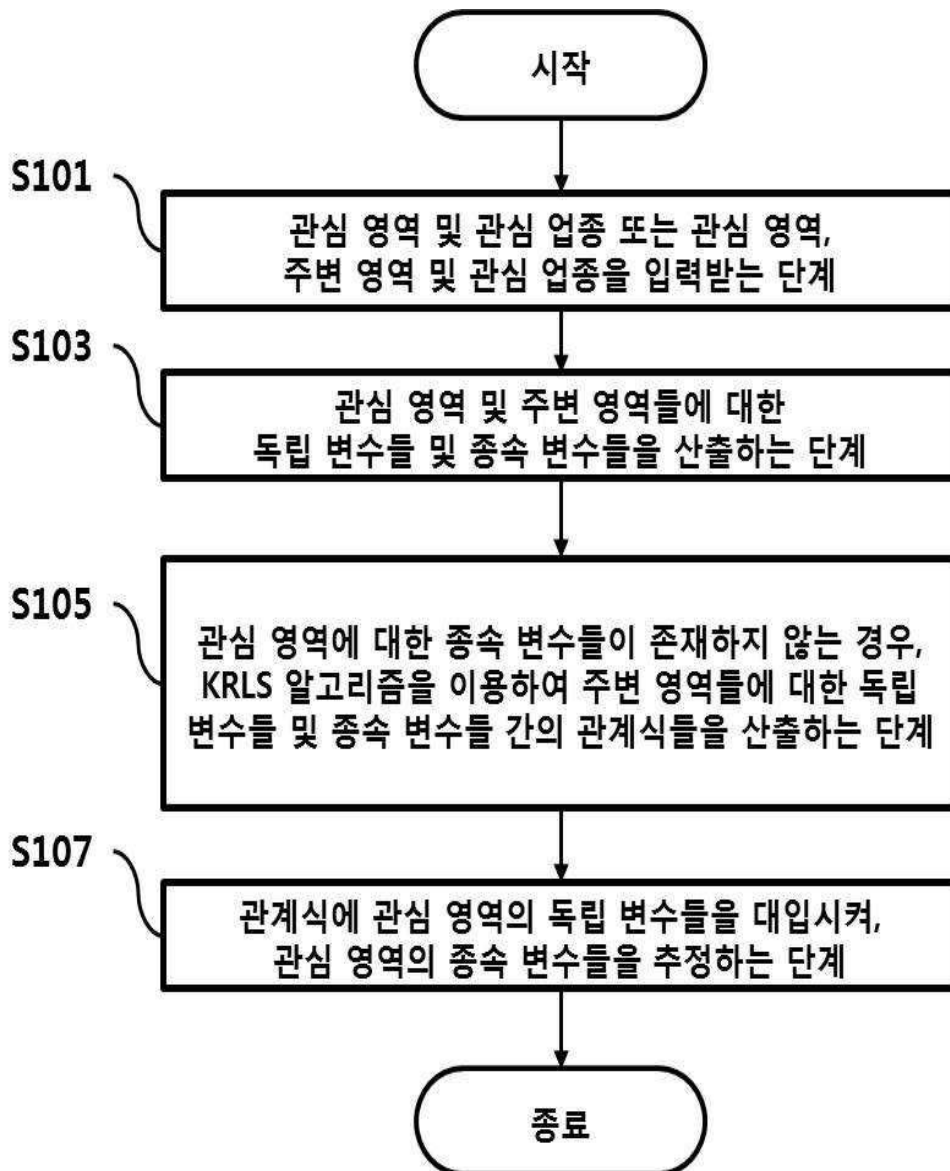
부호의 설명

- [0100] 101 : 관심 업종
- 103 : 관심 영역
- 105 : 주변 영역들

- 401 : 단순선형회귀분석을 이용한 업체 수에 대한 매출액 그래프
- 403 : KRLS를 이용한 업체 수에 대한 매출액 그래프
- 501 : 단순선형회귀분석을 이용한 예상 매출액 그래프
- 503 : KRLS를 이용한 예상 매출액 그래프
- 601 : 단순선형회귀분석을 이용한 예상 생존율 그래프
- 603 : KRLS를 이용한 예상 생존율 그래프

도면

도면1

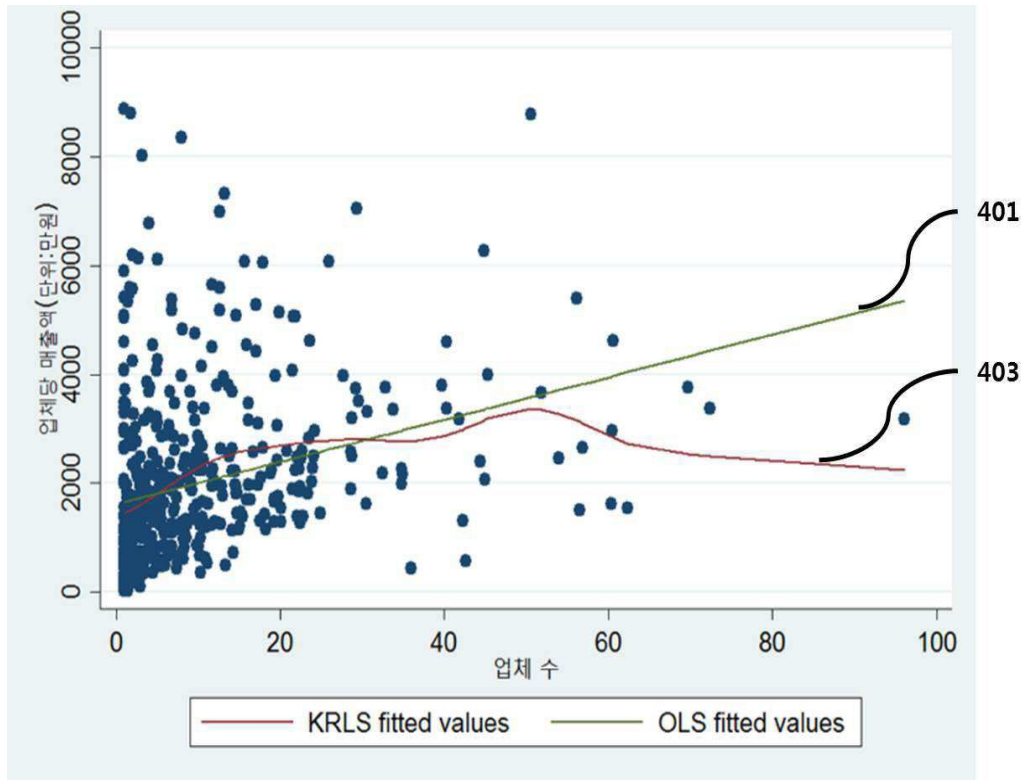


도면2

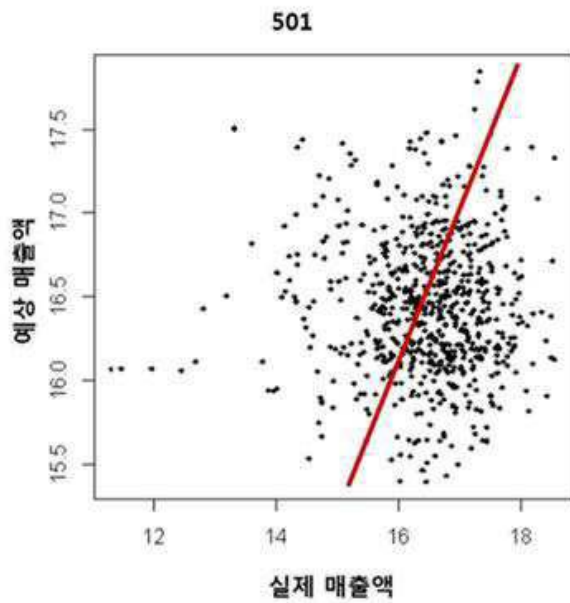
101

 한식음식점	 일식집	 본식집	 치킨집	 호프간이주점
 중국집	 양식집	 패스트푸드점	 제과점	 커피음료
 노래방	 미용실	 입시보습학원	 한의원	 PC방 PC방
 피부관리실	 외국어학원	 노인요양시설	 당구장	 네일숍
 예체능학원	 자동차수리	 골프연습장	 헬스클럽	 보육시설
 자동차미용	 여관업	 인테리어	 일반의원	 세탁소
 부동산중개업	 치과의원			
 슈퍼마켓	 편의점	 컴퓨터판매수리	 휴대폰	 정육점
 과일채소	 의류점	 패션잡화	 약국	 문구점
 화장품				

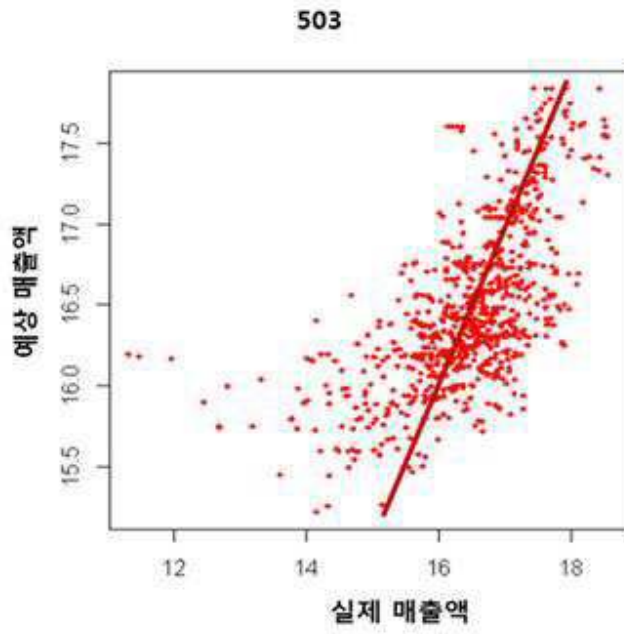
도면4



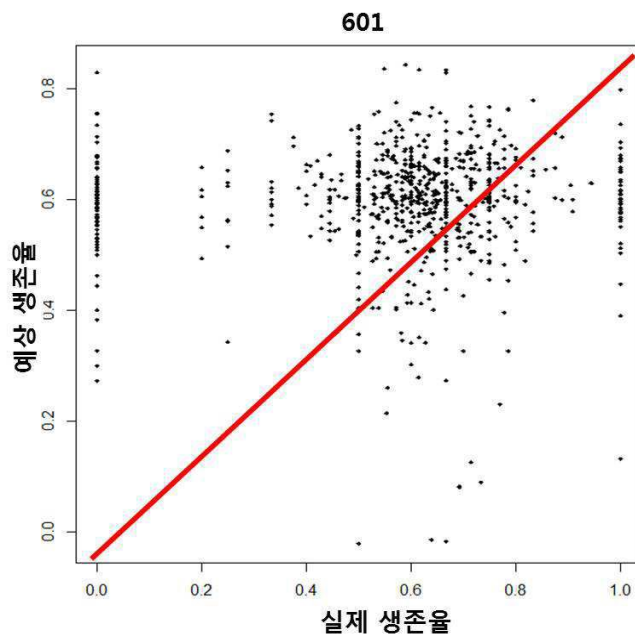
도면5a



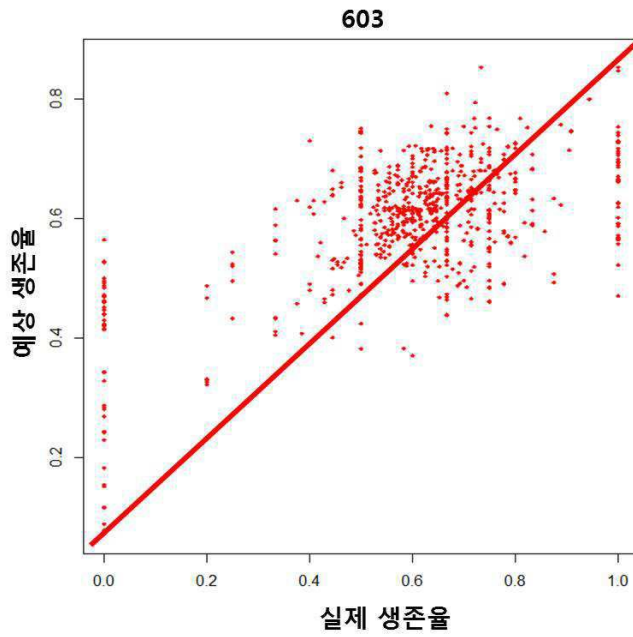
도면5b



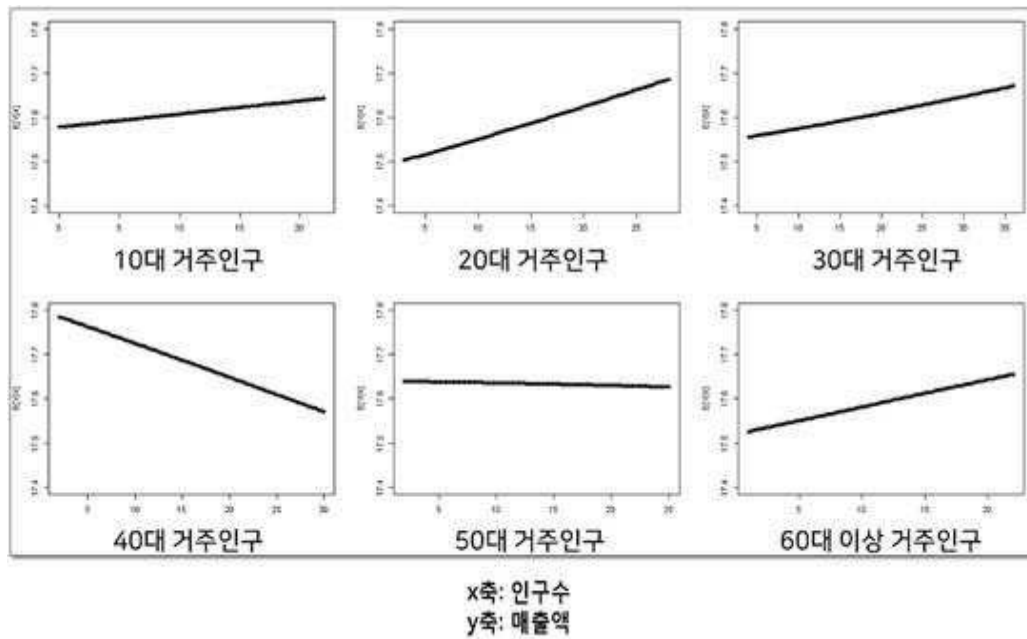
도면6a



도면6b



도면7



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 13, 1줄

【변경전】

제12에 있어서,

【변경후】

제12항에 있어서,