

Prediction of Game Platform Industry through Correlation Analysis between Metascore and Sales Volume

메타스코어와 판매량 간의 상관관계 분석을 통한 게임 플랫폼 산업 예측

Bong Hyun Kim¹

김봉현¹

¹ Professor, Department of Computer Engineering, Seowon University, Republic of Korea,
bhkim@seowon.ac.kr

Abstract: The gaming platform industry has changed and grown rapidly over the past few decades. Initially starting from simple 2D gaming methods, it is currently evolving into various forms and technologies, including converting cloud-based triple-A games into streaming services. These rapid changes in the gaming platform industry are contributing to the current situation of continuous technological innovation, full-scale competition, and an increase in the global user base. Efforts are increasing to analyze and predict growth indicators such as consumer demands, technical possibilities, and commercialization intersections in the game platform industry by analyzing the core elements of these changes in the game platform industry and reflecting them in the game industry. Therefore, in this paper, we conducted a study to analyze the correlation between game metascore and game sales volume to predict changes in the game platform industry and lead to game commercialization. In order to achieve this, games with comparatively high metascores and sales volume were found, positive/negative correlations between mutual indicators were verified, and the degree of relevance was assessed based on correlation strength. In addition, by exploring the relationship between metascore and sales under specific conditions, we derived data results that can support game developers and game companies in understanding the relationship between them.

Keywords: Game Platform Industry, Regression Analysis, Scraping, Metascore, Sales Volume

요약: 게임 플랫폼 산업은 지난 수십 년 동안 빠르게 변화하고 성장해왔다. 초기의 간단한 2D 게임에서 시작하여 현재는 클라우드 기반의 트리플-A 게임을 스트리밍 서비스로 전환시키는 방식까지 다양한 형태와 기술로 발전하고 있다. 이와 같은, 게임 플랫폼 산업의 빠른 변화는 지속적인 기술적 혁신, 전업적인 경쟁, 전 세계적인 사용자 기반의 증가 등에 기인하고 있다. 이러한 게임 플랫폼 산업 변화의 핵심 요소들을 분석하여 게임 플랫폼 산업에서의 소비자 요구, 기술적 가능성, 사업화 교차점 등의 성장 지표를 분석, 예측하여 게임 산업에 반영하려는 노력이 증대되고 있다. 따라서, 본 논문에서는 게임 플랫폼 산업의 변화를 예측하고 게임 사업화를 주도할 수 있도록 게임 메타스코어와 게임 판매량 사이의 상관관계를 분석하는 연구를 수행하였다. 이를 위해, 상호 지표간의 양/음 상관성을 확인하고, 메타스코어와 판매량이 상대적으로 높은 게임을 식별하여 상관관계 강도별 관련성 정도를 파악하였다. 또한, 특정 조건에서의 메타스코어와 판매량의 관계를 탐색하여 게임 개발자나

Received: September 05, 2023; 1st Review Result: October 08, 2023; Accepted: December 26, 2023

게임 기업에서 상호간의 연관성을 이해하는데 지원할 수 있는 데이터 결과를 도출하였다.

핵심어: 게임 플랫폼 산업, 회귀분석, 스크래핑, 메타스코어, 판매량.

1. 서론

현대사회에서 게임 산업은 빠르게 변화하는 기술, 다양화되고 세분화된 시장, 그리고 복잡해지는 사회적 상호 작용의 교차시기에 있는 상황이다. 게임 산업은 지난 몇 년 동안 급격한 성장을 경험하고 있다. 게임 산업은 더 이상 단순한 엔터테인먼트로 여겨지지 않고 있으며, 기술적 혁신, 문화적 변화, 경제적 기회의 중심지로 급부상하고 있다[1]. 최근의 게임 산업은 스마트폰의 보급률이 급증하면서 모바일 게임 시장이 빠르게 성장하고 있는 추세이다. 간편하고 접근성이 좋은 모바일 게임은 많은 사용자들을 끌어들이고 있으며, 이는 광고 및 인앱 구매를 통한 수익 증대로 이어지고 있다. 또한, 클라우드 기술의 발전으로 게임을 로컬 디바이스에 다운로드할 필요 없이 인터넷을 통해 스트리밍하는 것으로 변화되고 있다. 이러한 기술의 흐름속에서 구독 기반 모델은 사용자들에게 다양한 게임에 대한 저렴한 액세스를 제공하며, 이는 전통적인 게임 구매 방식에 변화를 가져오고 있다[2][3].

최근 AR/VR(증강현실/가상현실) 기술이 계속해서 발전하고 있으며, 이는 게임 경험에 혁신을 가져오고 있다. 이러한 기술은 사용자가 게임에 보다 몰입하고 대화형으로 경험할 수 있게 하여, 새로운 시장을 형성하고 있다. 이와 함께, e-스포츠도 전 세계적으로 인기를 끌고 있으며, 대규모 토너먼트와 리그가 많은 관중들을 끌어들이고 있다. 이러한 현상은 스폰서십, 광고, 상품 판매 등을 통해 새로운 수익 흐름을 생성하고 있다[4]. 또한, 개발 툴과 플랫폼의 발전 덕분에 소규모 스튜디오나 개인 개발자들도 시장에 진입할 수 있게 되었다. 인디 게임은 종종 독창적이며 혁신적인 컨셉트로 많은 게이머들의 관심을 끌어들이고 있다. 이러한 게임 산업의 발전속에서 많은 게임들이 초기 구매 비용을 낮추거나 무료로 제공하고 있으며, 게임 내에서 소액 결제를 통해 수익을 얻는 모델을 채택하고 있다. 이러한 방식은 지속적인 수익 생성을 가능하게 하지만, 일부 사용자들 사이에서 논란이 되기도 한다[5][6]. 게임 산업의 성장과 함께, 여러 정부 기관들은 게임 내 내용, 중독, 개인정보 보호 등과 관련된 규제를 강화하고 있는 실정이다.

현대 게임 플랫폼 산업은 전례 없는 다양성과 기술적 복잡성을 특징으로 하고 있다. 게임 산업의 성장과 발전에 따라 게임 판매량을 결정짓는 요인 역시 복잡해지고 있다. 게임 관련 기술의 발전으로 인해 게임의 질적인 면모가 향상되고 있으며, 이러한 현상으로 구매 결정에 영향을 미치고 있다. 특히, 클라우드 게이밍, AR/VR 기술 등의 신기술 채택은 신규 사용자를 유치하고 기존 사용자의 관심을 증대시키고 있다. 게임은 창의적이고 효과적인 마케팅 전략으로 대중의 인식을 높이고 제품에 대한 호기심을 자극하고 있다[7]. 특히, 게임 판매 분야에서 소셜 미디어 캠페인, 인플루언서 마케팅, 게임 리뷰 등의 전략이 게임의 초기 판매량 및 지속적인 관심에 큰 영향을 미치고 있다. 이와 함께, 게임 커뮤니티와 참여도는 게임의 판매 성공을 좌우하는 지표가 되고 있으며, 소비자의 문화적, 지역적 선호는 제품 디자인과 마케팅 전략의 개인화에 중요한 역할을 하고 있다. 또한, 사회적 이슈와 문화적 트렌드는 게임 콘텐츠의 수용 가능성을 결정짓고, 판매량에 직접적인 영향을 미치고 있으며, 경제적 안정성과 구매력은 소비자의 지출을 결정하는 주요 요인이 되고 있다[8][9].

게임 개발 및 출시는 많은 비용과 시간을 필요로 한다. 따라서 게임 개발자나 게임 회사는 자원을 효율적으로 운용하고, 성공적인 게임을 개발하고자 한다. 이러한 이유로 메타스코어와 판매량 간의 상관관계를 파악함으로써, 게임의 품질과 상업적 성과 사이의 연관성을 이해할 수 있으며, 이를 통해 게임 개발 및 마케팅 전략을 수립하는 데 도움을 줄 수 있고, 품질 높은 게임의 개발과 성공적인 게임 판매 전략을 마련할 수 있다[10].

따라서, 본 논문에서는 게임 플랫폼 산업의 변화를 예측하고 게임 사업화를 주도할 수 있도록 게임 메타스코어와 게임 판매량 사이의 상관관계를 분석하는 연구를 수행하였다. 게임 메타스코어는 전문 평론가들의 평가를 종합한 지표이고, 판매량은 게임의 상업적인 성과를 나타내는 지표이다. 이러한 지표들을 통해 게임의 평가와 상업적 성과 사이 간의 관계를 파악하면, 메타스코어 점수가 높은 게임이 더 많은 판매량을 기록하는지에 관한 여부를 조사할 수 있다. 이를 위해, 상호 지표간의 양/음 상관성을 확인하고, 메타스코어와 판매량이 상대적으로 높은 게임을 식별하여 상관관계 강도별 관련성 정도를 파악하였다. 또한, 특정 조건에서의 메타스코어와 판매량의 관계를 탐색하여 게임 개발자나 게임 기업에서 상호간의 연관성을 이해하는데 지원할 수 있는 데이터 결과를 도출하였다.

2. 관련 연구

2.1 메타스코어(Metascore)

메타스코어는 게임, 영화, TV 프로그램, 음악 앨범 등 다양한 엔터테인먼트 콘텐츠의 평가를 종합하여 제공하는 Metacritic 웹사이트의 평가 체계를 말한다. 이 체계는 특히 비디오 게임 산업에서 널리 인정받고 있으며, 게임의 전반적인 수용도와 품질을 판단하는 데 중요한 역할을 하고 있다. 메타스코어 생성 과정은 다음과 같다.

- ① **리뷰 수집** : Metacritic은 다양한 게임 리뷰 웹사이트, 잡지, 뉴스 출처 등에서 전문가 리뷰를 수집한다. 이 리뷰들은 각기 다른 평가 체계를 사용할 수 있지만, Metacritic은 이를 자체 평가 체계에 맞추어 표준화한다.
- ② **점수 변환**: 원본 리뷰의 평가 체계가 다를 수 있으므로, Metacritic은 모든 리뷰 점수를 100점 만점 체계로 변환한다. 이 과정에서 원본 리뷰의 의도를 최대한 정확하게 반영하려고 노력한다.
- ③ **가중 평균**: 단순히 모든 리뷰를 동일하게 취급하는 것이 아니라, Metacritic은 각 출처의 편집 품질, 신뢰도 등을 고려하여 특정 출처의 리뷰에 더 큰 가중치를 부여할 수 있다. 이 가중치는 내부 알고리즘에 의해 결정되며, 공개되지 않는다.
- ④ **메타스코어 생성**: 위의 과정을 거쳐 계산된 평균 점수가 메타스코어가 된다. 이 점수는 0점에서 100점 사이의 값으로, 게임의 전반적인 평가를 대표한다.

메타스코어의 색상 코드는 초록색(81-100, 대체로 긍정적인 리뷰), 노란색(61-80, 평균적인 리뷰), 주황색(40-60, 대체로 부정적인 리뷰), 빨간색(0-39, 극도로 부정적인 리뷰)으로 표현한다. 이러한 색상코드는 메타스코어의 웹사이트 및 관련 응용 프로그램에서 평가를 시각적으로 구분하기 위해 사용되고 있다. 색상코드는 게임뿐만 아니라 영화, TV 프로그램, 음악 앨범 등의 평가에서도 일반적으로 사용되는 방식이다. 이 점수 외에도 Metacritic은 사용자 리뷰를 기반으로 한 "유저 스코어"도 제공한다. 유저 스코어는 일반 사용자들이 게임에 대해 직접 평가한 것으로, 전문가 리뷰와 사용자 의견 사이의 차이점을 확인할 수 있다. 그러나 메타스코어에는 몇 가지 주의할 점이 있다.

첫째, 일부 리뷰어는 출시 직후 게임을 리뷰하므로, 후속 업데이트나 패치로 개선된 경우 이를 반영하지 않을 수 있다. 둘째, 개별 게임 경험은 주관적이므로, 메타스코어가 모든 사용자의 경험을 대변하지는 않는다. 따라서, 구매 결정을 내리기 전에 여러 리뷰를 읽고, 가능하다면 게임을 직접 체험해 보는 것이 좋다.

2.2 회귀분석(Regression Analysis)

회귀 분석은 통계학과 데이터 과학에서 사용되는 중요한 기법 중 하나로, 변수들 사이의 관계를 설명하고 예측하는 데 사용되고 있는 기법이다. 이 방법은 종속 변수(목표 변수)와 하나 이상의 독립 변수(예측 변수) 사이의 관계를 모델링한다[11]. 회귀 분석은 경제학, 과학 연구, 엔지니어링, 그리고 데이터 과학 등 다양한 분야에서 활용되며, 데이터 패턴을 이해하고, 예측 모델을 만들고, 미래 결과를 예측하는 데 사용되고 있다.

회귀 분석은 가설 설정, 데이터 수집, 모델 적합, 모델 평가, 예측 및 결론의 순서로 모델링이 진행된다. 가설 설정 단계는 회귀 분석을 시작하기 전에, 특정 독립 변수가 종속 변수에 미치는 영향에 대한 가설을 세우는 과정이다. 데이터 수집 단계는 연구 대상이 되는 현상에서 데이터를 수집하는 것으로, 수집된 데이터는 종속 변수와 잠재적인 독립 변수들의 관측값을 포함하고 있다. 모델 적합 단계는 수집된 데이터를 사용하여 회귀 모델을 적합시키는 것으로, 독립 변수의 계수(예: 선형 회귀의 기울기)를 추정하는 과정을 포함한다. 모델 평가 단계는 결정 계수(R-squared), p-값, 잔차 분석 등을 포함하여 통계적 기법을 사용하여 모델의 적합도를 평가하는 과정이다. 마지막으로, 예측과 결론 단계는 최종 모델을 사용하여 새로운 데이터에 대한 예측을 만들거나, 변수들 사이의 관계에 대한 결론을 도출하는 과정이다[12][13].

회귀 분석 유형에는 단순 선형 회귀(Simple Linear Regression), 다중 선형 회귀(Multiple Linear Regression), 다항 회귀(Polynomial Regression), 로지스틱 회귀(Logistic Regression)가 있다. 단순 선형 회귀는 한 개의 독립 변수를 사용하여 종속 변수를 예측하는 기법으로, 직선을 사용하여 변수들 사이의 관계를 나타낸다. 다중 선형 회귀는 둘 이상의 독립 변수를 기반으로 종속 변수를 예측하는 기법으로, 여러 개의 예측 변수를 고려하여 종속 변수의 변동성을 더 구체적으로 나타낼 수 있다. 다항 회귀는 독립 변수의 거듭제곱을 포함하여 비선형 관계를 캡처하는 기법이다. 로지스틱 회귀는 종속 변수가 이진(binary)인 경우(예: 발생/미발생, 성공/실패 등) 사용되는 기법이다[14][15].

3. 데이터 분석 및 예측

본 논문에서는 게임 플랫폼 산업에서의 게임 판매 예측을 위해 메타스코어와 판매량 간의 상관관계 분석을 수행하였다. 이를 위해, 데이터 크롤링, 데이터 정제 및 처리에 관한 데이터 수집 단계와 상호 간의 상관관계에 관한 데이터 분석 단계 등을 포함하여 연구를 수행하였다. 데이터 크롤링을 위해 게임 메타스코어 분석기관인 메타크리틱에서 게임 메타스코어와 판매량 데이터를 수집하고, 이를 활용하여 분석을 진행하였으며, 연구 범위는 스팀 플랫폼 내에 있는 게임 중 메타스코어가 제시된 게임에 한해서 수행하였다.

3.1 데이터 수집

연구의 핵심 데이터는 메타스코어 데이터와 판매량 데이터이다. 게임 메타스코어

데이터는 메타크리틱과 같은 게임 리뷰 사이트에서 메타스코어 점수를 크롤링하거나, 공식 API를 활용하여 데이터를 수집하였다. 또한, 게임 판매량 데이터는 게임 개발자, 게임 퍼블리셔, 또는 시장 조사 업체에서 제공하는 판매량 데이터를 활용하였다.

게임 메타스코어 데이터를 얻기 위해 메타크리틱을 참조하고, 게임 판매량 데이터를 얻기 위해 게임 대표 퍼블리셔인 벨브 사의 스팀 플랫폼을 참조하였다. 또한, 게임 시장 조사 및 데이터 분석을 위해 게임 시장의 판매량 추이를 조사하는 NPD(National Purchase Diary Panel Inc.) 그룹을 참조하였다. 스팀 플랫폼은 다른 플랫폼에 비해 거대한 게임 라이브러리를 보유하고 있으며, 사용자 커뮤니티와 소셜 기능, 편리한 게임 관리, 다중 플랫폼 지원 등의 독특한 장점을 보유하고 있어 다양한 게임 데이터 기반의 분석이 가능하다. 따라서, 다양한 게임 플랫폼들과 게임 분석 기관들 중 스팀 플랫폼, NPD 그룹의 데이터를 기반으로 연구를 수행하였으며, 게임 메타스코어 점수를 얻기 위한 게임 웹 스크래핑을 수행하였다. 스크래핑은 웹 페이지에서 데이터를 추출하는 과정을 말하며, 스크래핑을 통해 웹 페이지의 텍스트, 이미지 등의 다양한 정보를 수집하는 기법을 말한다. 웹 스크래핑을 하려면 첫 번째로 웹 페이지에 접근해야 하고, 두 번째로 웹 페이지의 내용을 파싱해야 한다. 세 번째로는 데이터를 이 프로젝트의 범위에 맞게 추출하고, 마지막으로 데이터를 가공하고 저장하는 과정을 거쳐야 한다. 이때 파이썬 과정에서 ‘BeautifulSoup4’와 ‘requests’를 사용한다.

메타크리틱의 링크를 기준으로 모든 게임 평점 링크를 스크래핑하여 CSV(Comma-Separated Values) 파일로 저장하면, 게임 링크 내의 게임 명칭, 발매일, 장르, 게시자, 개발자, 메타스코어, 구매 비용 등을 크롤링할 수 있게 된다. 따라서 스크래핑을 통해 얻은 CSV 파일을 기반으로 종류를 분할하여 재스크래핑을 시도하면, 분할한 종류에 따라 데이터를 수집할 수 있게 된다. 주의할 점으로는 메타크리틱에서 봇으로 감지하는 것을 피하기 위해 ‘user-agent’ 라이브러리를 이용하여, 웹 페이지로의 접근을 허용시켜야 웹 스크래핑이 가능하다는 것이 있다. 웹 페이지를 스크래핑하여 필요한 데이터를 종류에 맞게 크롤링을 수행하면, CSV 파일로 명시된다. 생성된 링크 CSV 파일을 토대로 XLSX 파일을 크롤링하면, 연구에서 사용될 게임 메타스코어와 판매량 등의 추이를 알아낼 수 있으므로, DESC.py 실행으로 추가 식별 데이터를 수집하였다.

	A	B	C	D	E	F	G	H	I
1	Name	Release Date	Genre	Publisher	Meta Score	Total Critic	User Rate	User Rate Count	
2	Disco Elysium	March 30, 2019	Role-Playing	ZA/UM	97	11	8.3	865	
3	Half-Life 2	November 16, 2004	Action, Shooter	Valve Software	96	81	9.2	12634	
4	Grand Theft Auto V	April 13, 2013	Modern, Action	Rockstar Games	96	57	7.9	9026	
5	Out of the Gate	March 23, 2019	Sports, Strategy	Sports Interactive	96	5	2.6	809	
6	The Orange Box	October 1, 2006	Action, Shooter	EA Games	96	34	9.1	2253	
7	Half-Life	November 1, 1998	Action, Shooter	Sierra Entertainment	96	24	9.1	5645	
8	BioShock	August 21, 2007	Action, Shooter	2K Games	96	44	8.6	5969	
9	Baldur's Gate	September 14, 1994	Role-Playing	Interplay	95	30	9.1	1567	
10	Divinity: The Fallen Angel	August 31, 1999	Role-Playing	Larian Studios	95	5	8.9	337	
11	Persona 5	October 2, 2017	Role-Playing	Sega, Atlus	95	13	8.4	370	
12	Portal 2	April 18, 2011	Action, Puzzle	Valve Software	95	52	9.1	9139	
13	The Elder Scrolls V: Skyrim	November 11, 2011	Role-Playing	Bethesda	94	32	8.4	12742	
14	Command & Conquer: Tiberian Dawn	September 24, 1994	Sci-Fi, Strategy	Virgin Interactive	94	5	8.9	489	
15	Mass Effect	January 26, 2007	Role-Playing	Electronic Arts	94	55	8.9	5646	
16	Grand Theft Auto: San Andreas	May 12, 2004	Action, Adventure	Rockstar Games	94	30	8.7	3240	
17	Sid Meier's Civilization II	February 2, 1991	Strategy	MicroProse	94	7	8.7	506	
18	Quake	June 22, 1996	Action, Shooter	id Software	94	9	8.7	1016	
19	BioShock Infinite	March 25, 2013	Action, Shooter	2K Games	94	68	8.6	9818	

[그림 1] 메타크리틱 게임 웹 스크래핑 결과

[Fig. 1] Metacritic Game Web Scraping Results

이와 같이, 메타크리틱 내에 있는 모든 게임의 명칭, 출시일, 장르, 배급사 그리고 메타스코어를 크롤링을 통해 추출할 수 있으며, 결과 XLSX 파일은 [그림 1]과 같다. 다음으로, 게임 판매량과의 상관관계 분석을 위해 스팀 플랫폼의 누적 판매량을 수집하였다. 수집된 데이터들은 메타크리틱의 PC, 콘솔 등 게임 플랫폼 전부의 점수를 매긴 메타스코어와 상이한 것을 이유로 주제 종목을 PC 게임에 치중하고, 메타크리틱에 있는 게임들로 분류를 해야하기 때문에 다음과 같은 코드를 사용하여 데이터 필터링 과정을 거친다. 데이터 필터링은 수집된 데이터에서 원하는 정보만을 추출하기 위해 불필요한 데이터를 제거하거나 걸러내는 과정을 의미한다. 이는 크롤링된 데이터의 품질을 향상시키고, 분석 또는 활용 목적에 맞는 정확한 데이터를 얻을 수 있도록 도와주는 역할을 수행한다.

메타크리틱의 메타스코어를 가진 게임들과 스팀 플랫폼에 개시된 게임들은 서로 상이할 수 있기 때문에, 데이터 수작업을 거친다. 데이터 수작업이란 특정한 요구사항에 맞추어 정확한 데이터를 수집해야 할 때 주로 사용하고, 이는 PC 플랫폼을 필터로 하여 정리한다. 수집된 모든 데이터들을 수작업으로 분류하고 나면 데이터 수집 과정을 끝낼 수 있게 된다. 추가로 수집한 데이터를 정제하여 필요한 정보만을 추출한 과정으로, 게임의 메타스코어와 판매량 데이터를 추출하고 일치시키며, 불필요한 정보를 제거하는 과정을 수행하였다. 또한, 데이터의 일관성을 확인하고 결측치나 이상치 등을 처리하여 데이터의 품질을 향상시켰다. 메타크리닉과 스팀 플랫폼을 통해 수집된 데이터들의 예외를 처리하는 과정은 [그림 2]와 같다.

```

6 "On game link 569, Error : 'NoneType' object has no attribute 'text'"
7 "On game link 582, Error : 'NoneType' object has no attribute 'text'"
8 "On game link 583, Error : 'NoneType' object has no attribute 'text'"
9 "On game link 598, Error : 'aggregateRating'"
10 "On game link 630, Error : 'NoneType' object has no attribute 'text'"
11 "On game link 635, Error : 'NoneType' object has no attribute 'text'"
12 "On game link 709, Error : 'NoneType' object has no attribute 'text'"
13 "On game link 710, Error : 'NoneType' object has no attribute 'text'"
14 "On game link 738, Error : 'NoneType' object has no attribute 'text'"
15 "On game link 749, Error : 'NoneType' object has no attribute 'text'"
16 "On game link 792, Error : 'NoneType' object has no attribute 'text'"
17 "On game link 793, Error : 'NoneType' object has no attribute 'text'"
18 "On game link 800, Error : 'aggregateRating'"
19 "On game link 818, Error : 'NoneType' object has no attribute 'text'"
20 "On game link 872, Error : 'NoneType' object has no attribute 'text'"

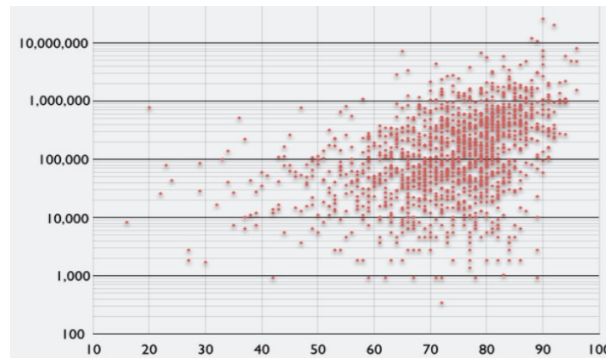
```

[그림 2] 수집 데이터에 대한 예외 처리 과정

[Fig. 2] Exception Handling Process for Collected Data

3.2 데이터 분석

데이터 분석을 위해 3.1 절의 데이터 수집 단계에서 메타크리틱 내의 메타스코어와 게임 명칭을 포함하여 수집하였다. 또한, 스팀 플랫폼 외 NPD 그룹에서도 페이지 별 게임 명칭과 판매량을 각각 크롤링하여 XLSX 파일로 정리하였다. 메타스코어는 자동 취합을 진행하고, 판매량은 수동 취합을 진행하여 데이터의 결측치와 이상치를 확인하여 정리하였다. 이를 통해, 메타스코어를 x축으로, 판매량을 y축으로 하는 산점도를 그려 상관관계를 시각적으로 확인하였다. 데이터 포인트들이 어떤 패턴을 보이는지, 양의 상관관계인지 음의 상관관계인지 등을 파악할 수 있었다. [그림 3]은 산점도 분석을 통해 시각화한 결과를 나타낸 것이다.

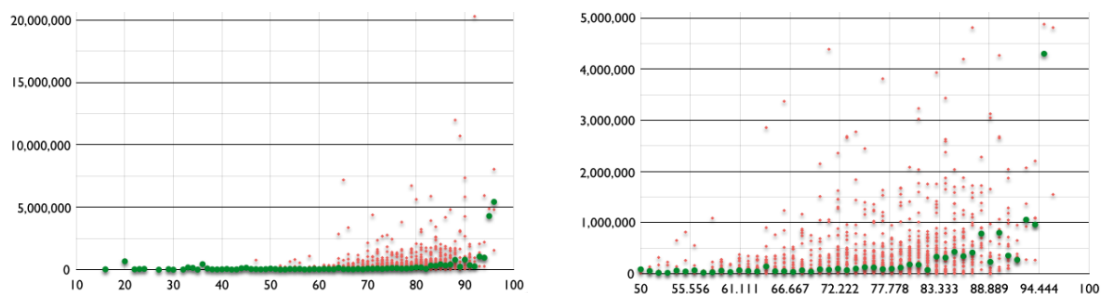


[그림 3] 산점도 분석 결과 그래프

[Fig. 3] Scatter Plot Analysis Result Graph

메타크리틱에서 80점을 받은 게임은 약 5만 6천개의 예상 판매량을 가질 가능성이 크며 3십 7만 5천개의 예상 판매량을 얻을 수 있다. 이는 모든 게임 분류에서의 메타크리틱 점수는 판매량 스펙트럼 전체에 걸쳐 감소할 수 있다는 결과를 도출하였다. 메타스코어 70점을 받은 중간 게임은 약 8만 9천개의 예상 사본을 판매량으로 가진다. 메타스코어 80점이면 중간 판매량은 약 1십 8만 2천점까지 증가하고, 이는 백분위수에서 동일한 상승 추세를 도출하였다. 하지만, 메타스코어 64점을 평균으로 받은 게임들은 82점을 평균으로 받은 게임보다 판매량을 훨씬 더 잘 기록하는 경향을 나타냈으며, 메타스코어 89점인 게임은 점수가 83점에서 88점 사이인 게임보다 더 잘 팔리는 경향을 나타냈다.

또한, [그림 4]는 메타스코어를 독립 변수 X로 판매량을 종속 변수 Y로 하여 회귀 분석을 수행한 결과로 메타스코어와 판매량을 선형적인 관계로 모델링 한 그래프이다. 회귀분석 결과로 얻은 회귀 식을 통해 메타스코어가 주어졌을 때 예상되는 판매량과 실제 판매량 간의 간극을 초록색 데이터 포인트를 이용하여 나타내었다. 출시 및 판매일의 문제점을 최소화하기 위한 수단으로 일부 게임의 판매량에 상당한 영향을 미칠 수는 있지만 단순히 출시된 지 얼마 되지 않아 판매량이 적거나 다른 플랫폼보다 훨씬 나중에 스팀 플랫폼에서 출시될 수 있으므로 스팀 플랫폼의 독점치 혹은 최소 6개월 된 게임을 먼저 살펴보는 동일한 분석을 하지 않았을 경우, 이상치가 발생할 수 있기 때문에 재출시가 더 높은 등급의 게임에 초점을 맞춰 결과 그래프가 도출되었다.

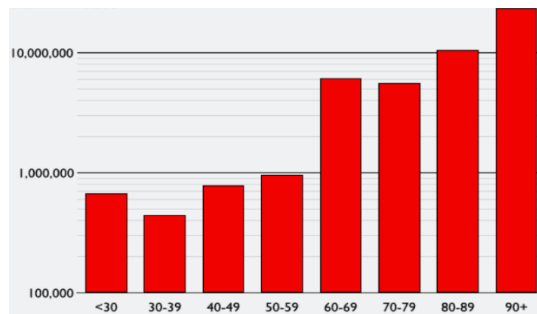


[그림 4] 산점도 결과를 선형 회귀 분석으로 변형한 결과(좌)와 평균 메타스코어 지표를 나눈 선형 회귀 분석 결과(우)

[Fig. 4] The Result of Transforming the Scatter Plot Result into Linear Regression Analysis(left) and the Result of Linear Regression Analysis Dividing the Average Metascore Indicator(right)

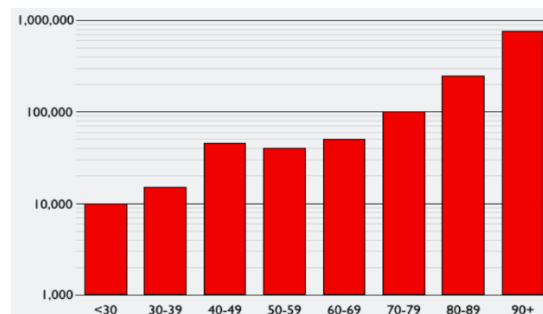
전반적인 메타스코어와 판매량 간의 데이터 예측은 선형회귀 방법 중 ARS(Automatic Regression Selection)에 의한 상관관계 분석 그래프를 이용하여 데이터를 분석할 수 있으며, 메타스코어와 판매량 간의 관계를 통계적으로 검정할 수 있다. 예를 들어, "메타스코어가 높은 게임은 판매량이 더 높다"라는 가설을 세우고, 이를 검정하여 가설이 통계적으로 유의미한지 판단할 수 있다.

흥미로운 것은 NBA2K23이라는 농구 게임은 데이터 상 가장 높은 등급의 게임에 대한 낮은 순위를 나타냈지만, 메타스코어 60점 이상으로 평가된 데이터에 있는 90개 게임은 모두 추정치에 따라 최소 5만장을 판매했다. 이러한 높은 메타스코어 점수를 받은 종목 중 상당수는 스팀 플랫폼 또는 NPD 그룹의 플랫폼을 이용하는 것이 아닌 다른 플랫폼 및 배포 방법을 통해서도 사용될 수 있으므로 메타스코어 반응 또한 낮은 등급의 게임 중 일부에 대해 일종의 상한선을 형성하는 것으로 볼 수 있다. 메타스코어 점수가 76점 미만인 60개 게임 중 어느 것도 추정치에 따르면 백만개 이상의 판매량을 기록하지 못한 것을 알 수 있다. 또한, 판매 실적의 일반적인 증가로 메타스코어 범위 전반에 걸쳐 균일하지 못하다는 것을 알 수 있으며 메타스코어 40점인 게임은 5만개의 판매량을 기록한 60점의 중간판매량을 기대해볼 수 있다는 것을 알 수 있다. 다시 말하면, 개별 게임에도 이러한 일반적인 추세를 상당 부분 쫓아가며, 나쁜 메타스코어 점수가 실패를 보장하는 것보다 더 이상 판매 성공을 보장하지 않는다는 점에 주목할 가치가 있다는 것을 알 수 있다.



[그림 5] 최대판매량과 메타스코어의 상관관계 분석 결과 그래프

[Fig. 5] Graph of Correlation Results between Highest Sales Volume and Metascore

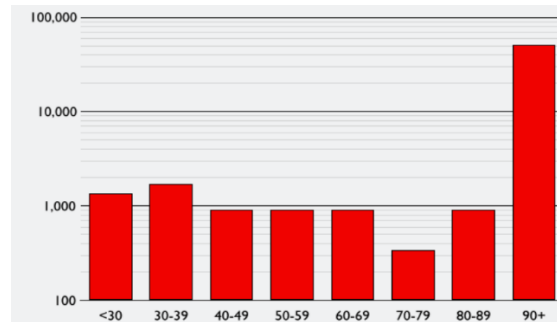


[그림 6] 중간판매량과 메타스코어의 상관관계 분석 결과 그래프

[Fig. 6] Graph of Correlation Results between Median Sales Volume and Metascore

메타스코어와 최소 판매량 간의 식별차는 [그림 7]과 같으며, 메타스코어 90점 이상을 획득할 수 있다면, 스팀 플랫폼 외 NPD 그룹의 플랫폼 내에서도 최소한의 판매 실적을

보장할 수 있다는 말이 된다. 이는 사업적으로 안정성과 예측 가능성을 제공할 수 있으며, 몇 가지 이점은 다음과 같다. 첫째, 안정성으로 최소한의 판매 실적이 보장되면, 기본적인 운영 비용 및 고정 비용을 충당하는 데 필요한 수익을 확보할 수 있다. 이는 비즈니스의 안정성을 높이고, 잠재적인 손실을 줄여준다. 둘째, 예측 가능성으로 최소한의 판매 실적이 보장되면 매출의 일정한 기초가 확보되기 때문에, 예측이 가능해진다. 이는 재고 관리, 운영 계획, 재무 예측 등과 같은 비즈니스 결정을 내리는 데 도움이 된다.



[그림 7] 최소판매량과 메타스코어의 상관관계 분석 결과 그래프

[Fig. 7] Graph of Correlation Results between Minimum Sales Volume and Metascore

4. 결론

게임 산업은 지난 수십 년 동안 급속도로 변화하고 성장해왔다. 초기의 간단한 2D 게임에서 시작하여 현재는 클라우드 기반의 게임까지 다양한 형태로 발전하였다. 게임 플랫폼 산업의 빠른 변화는 지속적인 기술적 혁신, 전업적인 경쟁, 전 세계적인 사용자 기반의 증가에 기여하였다. 이러한 변화의 핵심 요소들을 분석하여 게임 플랫폼 산업에서의 소비자 요구와 기술적 가능성, 사업화 교차점 등의 성장 지표를 반영하려는 노력이 증대되고 있다.

따라서, 본 논문에서는 게임 플랫폼 산업의 변화를 예측하고 게임 사업화를 주도할 수 있도록 게임 메타스코어와 게임 판매량 사이의 상관관계를 분석하는 연구를 수행하였다. 이를 위해, 상호 지표간의 양/음 상관성을 확인하고, 메타스코어와 판매량이 상대적으로 높은 게임을 식별하여 상관관계 강도별 관련성 정도를 파악하였다. 분석 결과, 메타스코어가 높을수록 최대판매량, 중간판매량, 최소판매량을 보장받을 수 있었다. 즉, 메타스코어가 60점 이상일 경우, 최대판매량 500만개 이상을 보장하였고, 메타스코어 40점 이상일 경우, 중간판매량 3만개 이상을 보장하였다. 다만, 최소판매량의 경우, 메타스코어 90점 이상을 제외하고는 대부분의 메타스코어 점수대에서 최소판매량이 1천개 내외로 보장되는 것을 확인할 수 있었다. 이와 같은 결과를 토대로, 최소한의 판매 실적은 비즈니스의 안정성과 장기적인 성장을 위해 중요한 지표로 판단할 수 있었다. 이를 위해 효율적인 운영, 고객 관계 유지, 마케팅 전략, 제품/서비스 개선 등 다양한 요소를 고려하여 최소한의 판매 실적을 유지하는 것이 유리하다는 결과를 도출하였다.

References

- [1] J. E. Bae, S. I. Kim, A Development Plan of Serious Game Based on Virtual Reality Through the Trend Analysis on

- Internal/External Game Industry, *Journal of Digital Design*, (2014), Vol.14, No.3, pp.737-748.
DOI: 10.17280/JDD.2014.14.3.073
- [2] E. H. Cho, Monetization of Smartphone Games in South Korea, Japan, and China, *Journal of Korea Game Society*, (2017), Vol.17, No.2, pp.75-84.
DOI: 10.7583/JKGS.2017.17.2.75
- [3] J. Balakrishnan, M. D. Griffiths, Loyalty towards online games, gaming addiction, and purchase intention towards online mobile in-game features, *Computers in Human Behavior*, (2018), Vol.87, pp.238-246.
DOI: 10.1016/j.chb.2018.06.002
- [4] L. A. Schneider, D. L. King, P. H. Delfabbro, Family factors in adolescent problematic Internet gaming, *Journal of Behavioral Addictions*, (2017), Vol.6, No.3, pp.321-333.
DOI: 10.1556/2006.6.2017.035
- [5] S. G. Oh, The Effect of Game Platform Evaluation and Flow Experience on Player Loyalty in Mobile Game Application, *The Journal of the Korea Contents Association*, (2020), Vol.20, No.3, pp.235-244.
DOI: 10.5392/JKCA.2020.20.03.235
- [6] J. E. Kim, H. J. Lee, Development Trends of Mobile Augmented Reality Game Platform and User Interface, *Journal of Korea Game Society*, (2017), Vol.17, No.6, pp.7-18.
DOI: 10.7583/JKGS.2017.17.6.7
- [7] Y. S. Su, W. L. Chiang, C. T. J. Lee, H. C. Chang, The effect of flow experience on player loyalty in mobile game application, *Computers in Human Behavior*, (2016), Vol.63, pp.240-248.
DOI: 10.1016/j.chb.2016.05.049
- [8] W. Admiraal, J. Huizenga, S. Akkerman, G. Ten Dam, The concept of flow in collaborative game-based learning, *Computers in Human Behavior*, (2011), Vol.27, No.3, pp.1185-1194.
DOI: 10.1016/j.chb.2010.12.013
- [9] Y. M. Kim, A Study on the Influence of Game Broadcasting Content Factors and Communicator Factors on Immersion and Viewing Intention : Focusing on e-sports game broadcasting contents, *Journal of Korea Game Society*, (2021), Vol.21, No.3, pp.39-52.
DOI: 10.7583/JKGS.2021.21.3.39
- [10] S. B. Hyun, Y. J. Kim, C. J. Park, Development of a Game Content Based on Metaverse Providing Decision Tree Algorithm Education for Middle School Students, *The Journal of the Korea Contents Association*, (2022), Vol.22, No.4, pp.106-117.
DOI: 10.5392/JKCA.2022.22.04.106
- [11] K. K. Sung, Study on the Delivery Demand Correlation Analysis through Analysis of Delivery Order Quantity, *Journal of Innovation Industry Technology*, (2023), Vol.1, No.2, pp.7-14.
DOI: 10.60032/JIIT.2023.1.2.007
- [12] J. S. Kim, Y. S. Kim, Development of a Model for Calculating the Construction Duration of Urban Residential Housing Based on Multiple Regression Analysis, *LHI Journal of Land, Housing, and Urban Affairs*, (2021), Vol.12, No.4, pp.93-101.
DOI: 10.5804/LHIJ.2021.12.4.93
- [13] H. Choi, A Study on Satisfaction Survey Based on Regression Analysis to Improve Curriculum for Big Data Education, *Journal of the Korean Society of Industry Convergence*, (2019), Vol.22, No.6, pp.749-756.
DOI: 10.21289/KSIC.2019.22.6.749
- [14] K. H. Lee, J. H. Lee, J. J. Kim & Y. G. Kim, Relation between Regression-based Food Additives and Carcinogenesis using Big Data Analysis, *The Journal of The Institute of Internet, Broadcasting and Communication*, (2018), Vol.18, No.1, pp.203-210.
DOI: 10.7236/JIIBC.2018.18.1.203
- [15] S. Y. Lee & J. H. Kim, Logistic regression analysis of newspaper readers characteristics affecting regular subscription, *The Korean Journal of Applied Statistics*, (2019), Vol.32, No.5, pp.653-669.
DOI: 10.5351/KJAS.2019.32.5.653