

1. 요약

본 프로젝트에서는 2019년 12월 13일부터 2022년 11월 8일까지 지니 마켓에서 판매된 상품 데이터를 바탕으로 RFM 분석, 연도별과 월별 매출 시각화, 결제 방법과 결제 금액 간 연관성 분석 그리고 시간대별 이용자 수 분석 등을 진행하였다. RFM 분석 결과를 토대로 프로모션을 진행할 때 이벤트 대상이 될 만한 주요 상품들을 알아보았다. 연도별 매출과 월별 매출 분석, 그리고 시간대별 이용자 수 분석을 통해 프로모션을 진행할 적정 시기를 도출하였다. 마지막으로 결제 방법과 결제 금액 간 연관성 분석을 통해 구매자들이 선호하는 결제 방식을 확인하고 해당 결제 방식을 이용한 프로모션을 제안하였다.

2. 프로젝트 목적

학원 전문 쇼핑몰인 지니 마켓에서는 학원 선생님과 학부모, 학생 등을 대상으로 교구뿐만 아니라 모바일 쿠폰, 생활용품까지 폭넓게 판매 중이다. 지니 마켓에서 더 많은 수익을 창출할 수 있도록 수익성이 높은 기업과 소비자의 구매 성향을 월별 매출과 결제방법 등을 토대로 분석하려고 한다. 이를 바탕으로 이벤트 대상 품목을 선정하고 적정 시기를 파악하여 새로운 이벤트 기획 방향을 설정하고자 한다.

3. 분석 방법

이번 분석에 활용된 데이터는 2019년 12월 13일부터 2022년 11월 8일까지 지니 마켓에서 판매된 상품 데이터이다. 총 218601개의 데이터가 있고 각 열은 주문번호, 업체명, 상품명, 제조사, 주문수량, 판매금액, 결제방법, 주문일자, 처리상태, 초도상태, 제작문구 내역, 할부기간을 나타낸다. 주문번호가 같은 데이터들은 같은 시점에 동일한 사용자로부터 구매된 상품들이다.

초도상품과 제작문구 내역은 분석에 사용하지 않는 변수이므로 데이터에서 제거하였다.

분석은 판매자 수익을 기준으로 하여 처리상태가 구매 확정인 경우만을 추출한 데이터를 가지고 아래와 같은 분석을 수행하였다.

- 입점 기업별 **RFM** 분석(3그룹) 진행

이 데이터에서는 고객에 대한 정보를 알 수 없기 때문에 업체를 기준으로 하였다. 입점 기업 데이터에 해당하는 업체명 컬럼에서 1개의 결측값을 발견하여 해당 데이터를 삭제한 데이터를 이용하여 분석을 하였다. 또한, 결제 방법의 차이로 ‘지니 태블릿’과 ‘지니 태블릿(후불집행)’이 분리된 것을 확인하여 ‘지니 태블릿(후불집행)’은 모두 ‘지니 태블릿’으로 바꿔주었다.

RFM 분석에 앞서 데이터를 대략적으로 파악하기 위해 총 판매 금액이 높은 상위 10개 기업과 그 기업별 매출 점유율을 확인하였다. 그리고 상위 3개 기업에서 수익성이 높을 것으로 기대되는 상품이 무엇인지 살펴 보았다. 다음으로, 판매 빈도수가 많은 상위 10개 기업을 확인하였다. 이때, 판매 빈도수는 해당 기업이 존재하는 행의 개수로 계산하였다. 이 경우, 판매 금액을 함께 고려하여 판매 빈도수는 2위, 판매 금액은 4위인 ‘지니’에서 수익성이 높을 것으로 기대되는 상품이 무엇인지 살펴 보았다.

RFM이란 **Recency(R)**, **Frequency(F)**, **Monetary(M)**의 약자로 이 분석에서는 업체의 최근 판매 일자, 3년간의 판매 횟수, 3년간의 판매 금액을 바탕으로 분석을 수행하였다. 각각의 변수를 고려하여 1부터 4 사이의 점수를 매기고 세 점수를 합산한 것을 기준으로 그룹을 나누어 어느 업체의 상품이 많이 판매되는지 확인해보기로 하였다.

Recency는 데이터상의 마지막 판매일자에 1을 더하여 가상 분석일을 설정해준 뒤, 기업별로 이 날짜에서 마지막 판매일자를 뺀 값으로 지정하였다. **Frequency**는 판매 목록에서 업체명의 개수를 세서 구했고 **Monetary**는 기업별로 판매금액의 총합을 구하였다. 이때, **Recency**의 값은 낮을수록, **Frequency**와 **Monetary**의 값은 클수록 좋다. 그렇기에 해석을 쉽게 할 수 있도록 **Recency**의 값이 낮으면 4를 부여하고 높으면 1을 부여하여 **Recency**의 등급을 매겼다. **Frequency**와 **Monetary**는 각각 값이 높으면 4를, 낮으면 1을 부여하여 **Frequency**와 **Monetary**의 등급을 매겼다. 그리고 각 등급의 합을 **RFM_score**로 할당하여 이 값이 9 이상이면 매출 기여도가 큰 그룹, 5 이상 8 이하이면 매출 기여도가 중간인 그룹, 4 이하이면 매출 기여도가 작은 그룹으로 간주하고 기업을 분류하였다.

- 매출 시각화

결제방법과 할부기간은 고려하지 않았으며, 주문일자를 기반으로 연도와 월을 분할한 뒤 각각 새로운 열에 할당하여 판매금액의 합을 도출하였다.

1. 연도별, 월별 매출

각 연도와 월을 int 인덱스로 채운 '연도', '월' 컬럼을 새로 생성한 뒤 **groupby** 함수를 이용하여 연도별, 월별 매출에 대한 시각화를 진행하였다.

2. 월별 순이익

먼저, 매출 시각화 과정에선 '상품명' 컬럼에서 '판매안함' 항목들을 모두 제거했다. 다음으로, 월별 매출 순이익 계산을 위해 할부기간을 고려하고자 '할부기간'이 적용된 항목들과 미적용된 항목들을 분리하였다. 총 매출계산에 필요한 '연도', '월', '업체명', '할부기간' 등 분석에 필수적인 컬럼들만 모아 데이터를 만들었다. 그 후, 할부기간이 적용된 상품과 미적용된 상품에 대한 월별 매출을 각각 계산한 후 합산하여 월별 순이익 시각화를 진행하였다. 마지막으로, 업체별로 월별 매출액을 계산하여 어떤 업체가 해당 월에 가장 많은 매출을 내었는지 확인하였다.

- 결제 방법에 따른 분석

결제 방법이 없거나 '-이거나 '+ 포'인 값은 모두 결측치로 간주하고 제거하였다. '신 + 포'와 같이 결제 방법이 다양한 경우에는 첫 번째 값을 결제 방법으로 여기고 분석을 진행하였다. 이때, 첫 번째 값은 각 결제 방법의 첫 글자에 해당한다. 예를 들어, '신'은 '신용카드', '적'은 '적립금'을 의미한다. 이를 토대로 데이터를 변환하여 결제 방법에 따른 판매금액의 합계와 평균을 구한 뒤 결제금액과 결제 방법의 연관성을 살펴보았다.

- 시간대별 이용자 수 분석

시간대별 이용자 수를 파악하기 위해 '주문번호'와 '주문일자' 컬럼을 활용하여 **groupby** 함수를 통해 주문 시간대와 주문자 고유번호(중첩되는 번호는 한 사람으로 판단하여 계산 적용) 추출 및 데이터화 작업한 뒤 시각화를 진행하였다.

4. 분석 결과

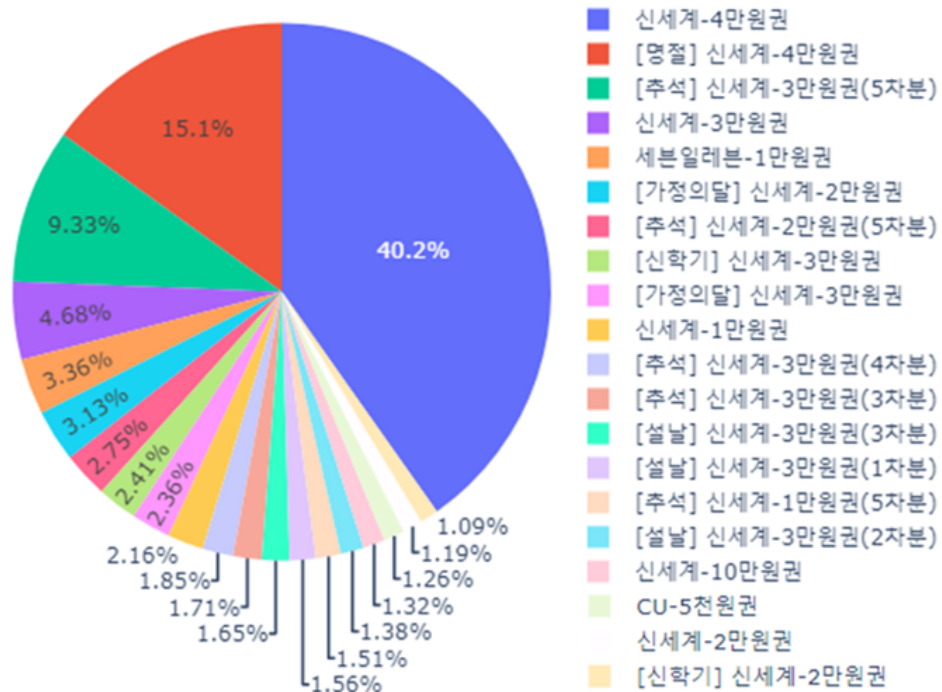
- 입점 기업별 **RFM** 분석 진행

	업체명	판매금액	매출 점유율
1	다우기술	2728817754	33.423
2	지니 태블릿	2155496000	26.401
3	천재태블릿	1393524000	17.068
4	지니	865960870	10.606
5	에이스디포	223196200	2.734
6	오피스웍	176806100	2.166
7	라운웍스	104317040	1.278
8	쥬크박스	74799690	0.916
9	블루스톤익스체인지	54878470	0.672
10	알파	30844914	0.378

판매 금액이 높은 상위 10개 기업

판매금액이 높은 상위 10개 기업을 확인한 결과, 1위는 '다우기술'로 총 판매금액이 약 2,728,817,754원(약 27억 원)이고 매출 점유율은 약 33%로 나타나 가장 높았다. 2위와 3위는 각각 '지니 태블릿'과 '천재 태블릿'으로 모두 학습 태블릿을 판매하고 있는 기업이었고 판매 금액은 각각 2,155,496,000원(약 21억 원)과 1,393,524,000원(약 13억 원)이었다. 매출 점유율은 각각 약 26%와 17%였다.

다우기술 상품 중 매출금액이 큰 상위 20개 상품



위에서 살펴본 10개의 기업 중 상위 3개 기업에서 수익성이 높은 상품을 파악해보았다. 먼저, ‘다우기술’에서 판매 금액이 높은 상위 20개 상품을 확인해보았다. 18개는 신세계 상품권이었고 나머지 둘은 각각 세븐일레븐 모바일 금액권(1만원)과 CU 모바일 금액권(1만원)이었다. 신세계 상품권 중에서도 13개의 상품은 신학기, 가정의 달, 추석 등의 시기에 특가로 판매된 것으로 나타나 상품권을 한정 판매하는 이벤트가 대체로 성공적이었음을 알 수 있다. 또한, 소비자들이 상품 교환권보다는 금액권을 더 선호함을 확인할 수 있다.

‘지니 태블릿’과 ‘천재 태블릿’의 경우, 판매하는 상품의 종류가 많지 않기 때문에 판매 금액이 높은 상위 5개 상품만을 확인해보았다. 모두 학습 태블릿이고 결제 방식에서만 차이가 있었다.

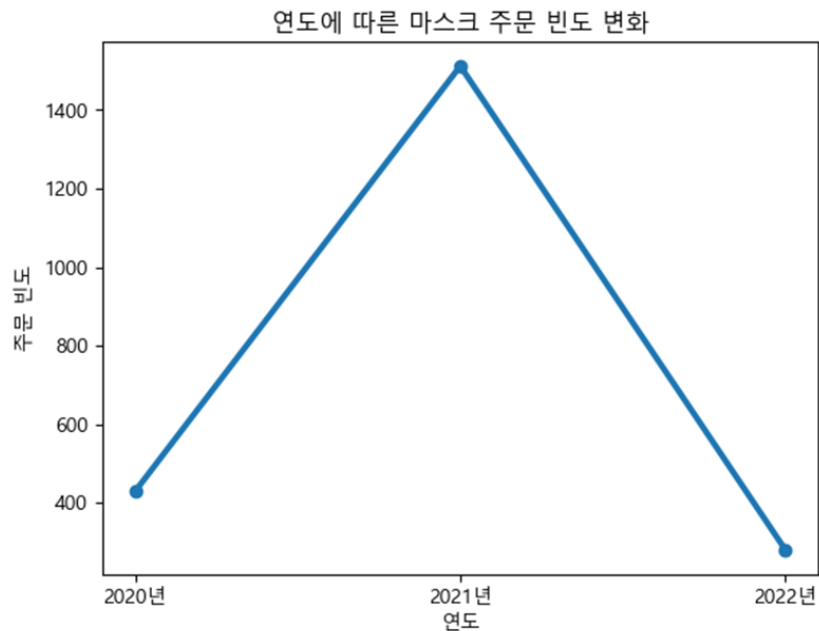
업체명 판매빈도		
1	다우기술	65981
2	지니	44562
3	쥬크박스	12369
4	에이스디포	5951
5	오피스퀵	5718
6	거영아이앤디	2901
7	알파	2887
8	유엔알코리아	2428
9	지니 태블릿	2364
10	라운웍스	2284

판매 빈도가 높은 상위 10개 기업

판매 빈도가 높은 상위 10개 기업을 확인해보면, ‘다우기술’(65981건)이 1위로 나타났으며 2위는 ‘지니’(44567건), 3위는 ‘쥬크박스’(12359건)로 나타났다. 판매 금액 기준 2위였던 ‘지니 태블릿’은 주문 빈도 9위로 나타났으며 2364건을 차지했다. 수명이 길고 결제 금액이 큰 태블릿보다 사무용품 및 문구류가 주문 빈도가 높은 것을 확인해 볼 수 있다. 그럼에도 판매빈도가 높은 상위 10개 기업에 지니 태블릿이 속한다는 점은 주목할 만하다.

판매금액	
상품명	
[예약판매특가] HB 황사방역 마스크 KF-94(흰색)/대형 1Set 100개입/2월15일발송	25740000
[거울방학 판촉물] 스마트 HBM-거울방학 판촉물 흔들이 손난로-1Box 120개입	22375100
[특가/무배] 국산/개별포장-보건용 마스크 KF94-대형 1갑 50개입	19836600
[스마트 HBE] 지사용-방문 영업 패키지	19000000
[스마트 HBM] 지사용-방문 영업 패키지	19000000
[21년 신학기 홍보물 특가] HB 수학 황사방역 마스크 KF-94(흰색)/대형 1Set 100개입	18960000
[스마트 HBM] 거울방학 판촉물-핫팩 1Box 120개입	13780000
[거울방학 판촉물] 스마트 HBE-거울방학 판촉물 흔들이 손난로-1Box 120개입	13588600
[해법독서논술] 글꾸러미 바인더	12060800
HBS 에코백 - 1세트 12개입	12060000

‘지니’의 상품 중 매출 금액이 큰 상위 10개 상품



판매 빈도 순위는 2위이고 판매 금액 순위는 4위인 '지니'에서 주목할 만한 상품을 파악해보았다. '지니'에서 판매 금액이 큰 상위 10개 상품을 확인한 결과, 이 기업에서는 대체로 마스크, 손난로 등의 생활용품과 바인더 같은 교구를 판매하는 것으로 나타났다. 마스크의 경우, 코로나라는 특수한 상황으로 인해 판매 빈도가 높은 것으로 예상된다.

	Recency	Frequency	MonetaryValue	R	F	M	RFM_score	RFM_class
업체명								
거영아이앤디	12	2901	24970280	4	4	4	12	High
다우기술	1	65981	2728817754	4	4	4	12	High
블루스톤익스체인지	11	1217	54878470	4	4	4	12	High
알파	7	2887	30844914	4	4	4	12	High
에이스디포	6	5951	223196200	4	4	4	12	High
오피스퀵	7	5718	176806100	4	4	4	12	High
유엔알코리아	8	2428	16111200	4	4	4	12	High
쥬크박스	7	12369	74799690	4	4	4	12	High
지니	8	44562	865960870	4	4	4	12	High
천재태블릿	7	1772	1393524000	4	4	4	12	High

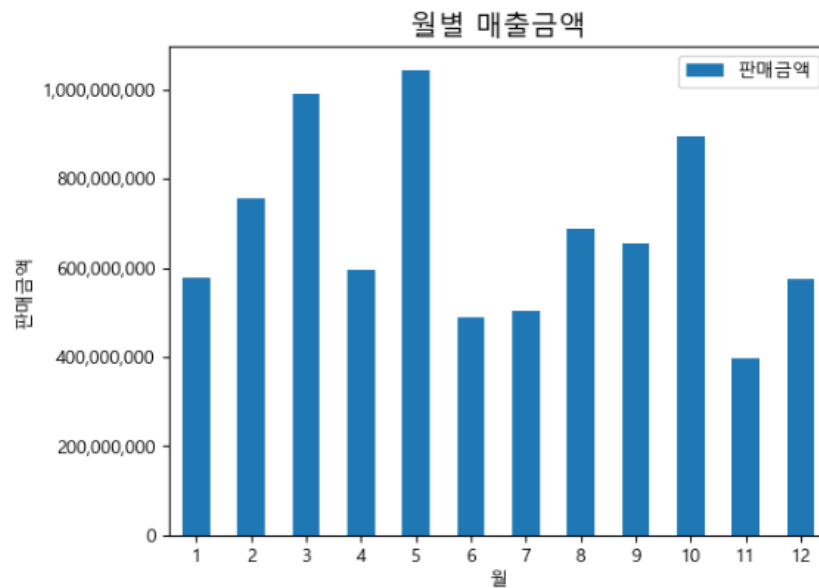
High 그룹 중 RFM_score가 12점인 기업

RFM분석으로 상위그룹으로 분류된 High그룹 중 RFM_score가 12점인 기업을 확인해 보았다. 판매 금액이 가장 크고 판매 빈도가 가장 높았던 '다우기술'은

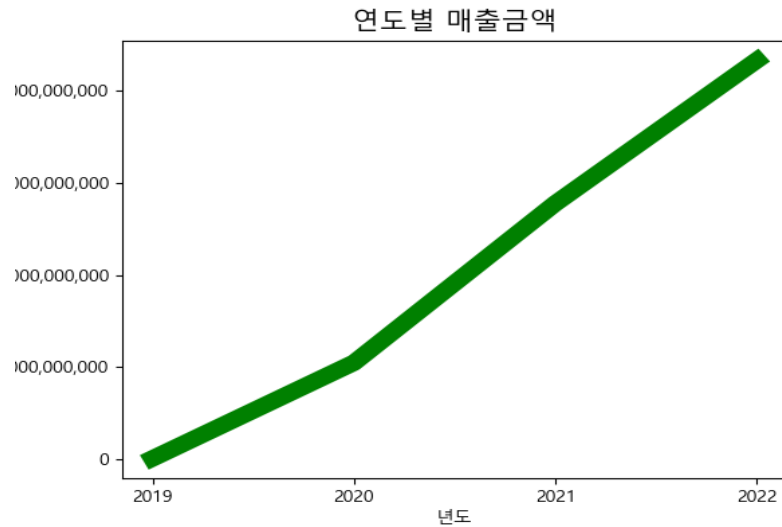
RFM분석 결과에서도 상위 그룹에 속해있는 것을 확인 할 수 있었다. 앞서 살펴본 결과와 RFM 분석 결과를 종합적으로 확인해보면 ‘다우기술’, ‘지니’, ‘천재 태블릿’ 등에서 판매하는 상품을 중심으로 이벤트를 진행하면 좋을 것으로 판단된다.

- 매출 시각화

- 1. 연도별, 월별 시각화

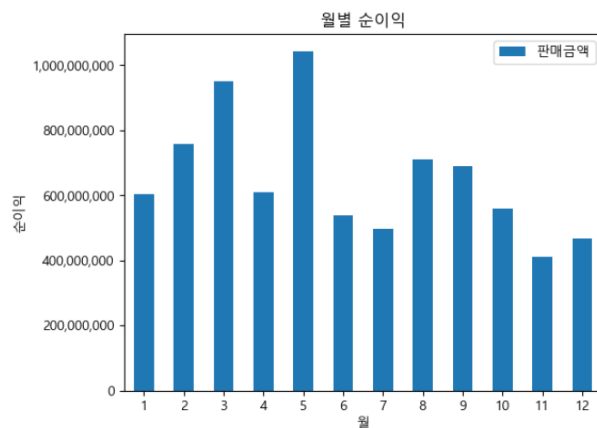


월별 판매금액을 분석한 결과, 신학기 시작인 3월에는 989,788,885원으로 2위를 차지하고 있으며, 5월이 1,044,829,970원으로 1위를 차지하고 있다. 또한, 중간고사와 추석이 겹친 10월은 895,017,703원으로 3위를 차지하고 있다.



연도별 판매금액을 살펴보면 **2019년 12월**부터 데이터가 시작되었으며, 이로 인해 초기 판매금액이 적은 것을 확인할 수 있다. 그러나 매년 판매금액이 약 **2배씩** 상승하고 있어 우상향하는 그래프를 볼 수 있다. **2019년**에 데이터가 적은 것을 감안하더라도 매년 쇼핑몰 매출이 성장하고 있음을 알 수 있다.

b. 월별 순이익

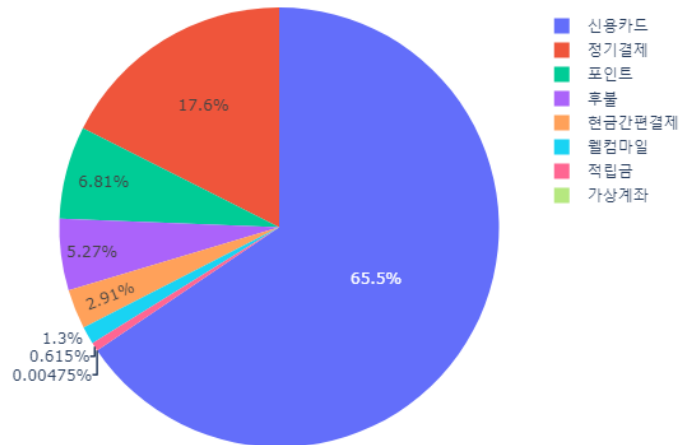


업체명	년도	월	할부기간	판매금액
16676 지니 태블릿	2020	10	12	768000
16680 지니 태블릿	2020	10	12	1920000
16681 지니 태블릿	2020	10	12	768000
16703 지니 태블릿	2020	10	24	1920000
16704 지니 태블릿	2020	10	24	1920000
...
156132 천재태블릿	2022	10	18	1980000
156804 천재태블릿	2022	10	18	3168000
156918 천재태블릿	2022	10	12	396000
157563 천재태블릿	2022	10	12	396000
157566 천재태블릿	2022	10	12	396000

월별 순이익도 위와 같은 월별 매출금액 그래프와 비슷한 추세로 **5월**과 **3월**이 각각 **1, 2위**를 차지하여 가장 높은 매출을 기록하고 있다. 반면 매출금액 그래프에서 매출 **3위**였던 **10월**이 순이익 그래프에서는 평균치 정도로 떨어진 것으로 보아 **10월**에는 할부로 결제한 상품들이 집중되어 판매된 것으로 예상되었다. 할부기간이 적용된 항목들만 추출하여 확인한 결과, 해당 상품들이 모두 태블릿 종류 상품이고 **10월**에 집중되어 있는 것을 알 수 있었다.

- 결제 방법에 따른 분석

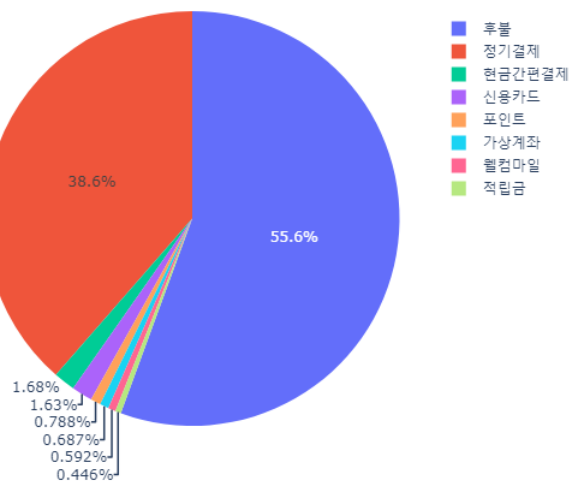
결제방법별 판매금액 합계



판매금액	
결제방법	
가상계좌	339,990원
신용카드	4,687,854,783원
헬컴마일	92,798,695원
적립금	43,975,975원
정기결제	1,256,518,000원
포인트	486,920,060원
현금간편결제	208,048,535원
후불	377,088,000원

결제 방법에 따른 판매금액의 합계를 기준으로 보면, 신용카드를 통한 판매금액이 압도적으로 컸고 정기결제와 포인트, 후불결제의 순으로 크게 나타났다. 대체로 분할 결제가 가능하거나 결제일을 미룰 수 있는 수단을 이용했을 때 판매 금액이 크게 나타났다.

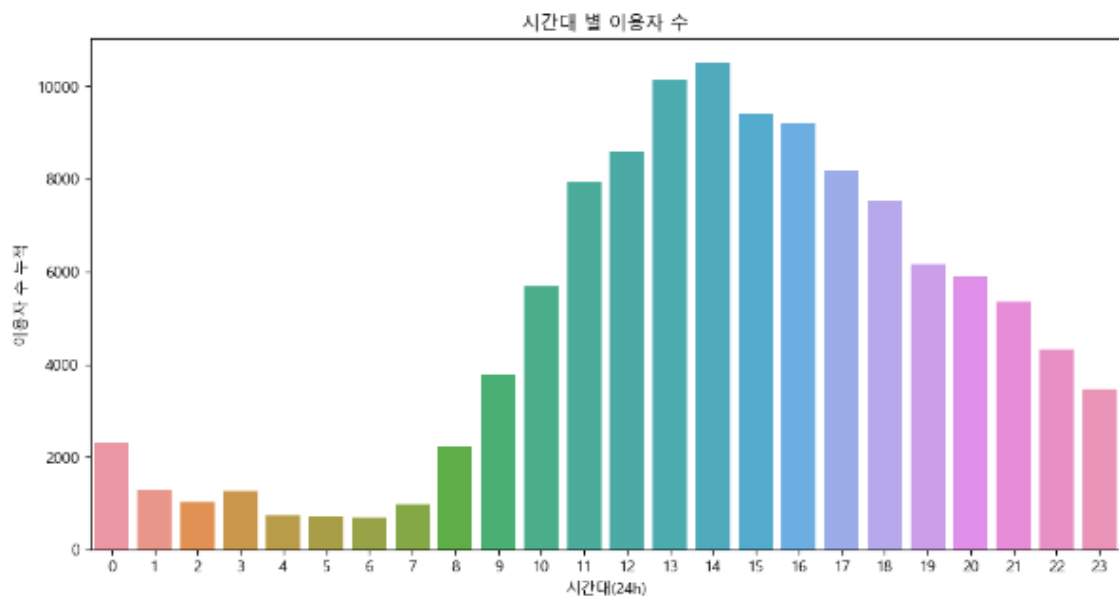
결제방법별 판매금액 평균



판매금액	
결제방법	
가상계좌	17,000원
신용카드	40,346원
헬컴마일	14,646원
적립금	11,044원
정기결제	954,801원
포인트	19,515원
현금간편결제	41,585원
후불	1,376,234원

다음으로, 결제 방법별 판매금액의 평균을 살펴보면 후불결제-정기결제 순으로 크게 나타났고 나머지 결제 방법의 판매금액 간 차이는 크게 나타나지 않았다. 합계를 분석했을 때와 비슷하게, 큰 금액을 결제해야 할 때는 대체로 후불결제나 정기결제와 같이 분할 결제가 가능하거나 결제일을 미룰 수 있는 수단이 자주 활용된 것을 확인할 수 있었다. 평균 판매금액은 적지만 판매금액의 합계는 가장 큰 신용카드의 경우, 사람들이 가장 많이 사용하는 수단이어서 이러한 결과가 나타난 것임을 유추할 수 있다.

- 시간대별 이용자 수 분석



그래프 분석 결과 주문량이 제일 많은 시간은 오후 2시이며 대체로 오전 11시에서 오후 6시 사이가 주문량이 많은 시간대라고 볼 수 있다. 반대로 새벽시간대에는 주문량이 현저히 떨어지는 것을 확인 할 수 있다.

5. 결론 및 활용방안

1. 판매 금액과 판매 빈도가 모두 높으면서 RFM 점수가 12인 ‘다우기술’과 ‘지니’, ‘천재 태블릿’ 등의 업체 상품들을 중심으로 이벤트를 기획하면 좋을 것이다. ‘다우기술’의 모바일 상품권(금액권), ‘천재 태블릿’의 태블릿 등이 가장 많은 판매금액을 기록하였다. ‘지니’ 기업의 상품의 경우, 일시적으로 판매 빈도가 높은 것으로 보이는 마스크와 상품의 특성이

모호한 방문 영업 패키지를 제외한 나머지 상품을 고려하는 것이 타당할 것으로 보인다.

2. 월별 판매금액을 분석한 결과, 3월, 5월, 10월에 매출이 높게 나타난 것을 확인할 수 있었다. 이러한 달에는 한정 판매나 할인 이벤트를 진행하면 더 많은 수익을 창출할 수 있을 것으로 보인다. 특히 10월에는 가격대가 있는 태블릿 종류의 상품들이 할부결제를 통해 판매되는 것이 확인된 만큼 그에 맞는 할인이나 사은품 증정 같은 이벤트를 기획하는 것이 매출 성과에 도움을 줄 것이라 판단된다.

3. 시간대별 이용자 수를 분석한 결과, 주로 오전 11시에서 오후 6시 사이의 시간대에 이용자 수가 많게 나타났고 그중 최고 주문량 시간대는 오후 2시대였다. 이벤트에 대한 광고를 효과적으로 할 수 있도록 이 시간대를 집중적으로 공략하여 이벤트를 노출하면 이벤트 참여자 수를 늘릴 수 있을 것으로 예상된다.

4. 구매자들의 결제 방법 중 신용카드 결제와 정기 결제가 가장 많은 비중을 차지했고 프로모션을 진행하는 시기에 판매율을 늘리기 위해 신용카드 결제시 추가 포인트 적립이나 정기 결제 구독 시 추가 혜택 제공 (예: 할인권) 등 해당 결제 수단을 이용한 프로모션을 진행한다.

6. 제한점

1. 본 분석에서 활용한 일반 RFM 모형에서는 R, F, M 모두에 가중치를 1:1:1로 부여하여 변수의 중요성을 판단하기에 어려움이 있다. 따라서, 기업에서 판매하는 상품별 특성을 파악하여 가중치를 부여할 수 있는 모델을 만든다면 수익 창출에 도움이 되는 기업과 상품을 파악하기에 더 효과적일 것으로 예상된다.

2. 특가로 판매하는 이벤트의 경우에는 ‘오피스퀵’의 상품도 고려하면 좋을 것이다. ‘특가’로 판매된 상품을 판매 금액이 높은 순으로 상위 20개를 추출했을 때, 15개는 ‘다우기술’의 상품권, 3개는 ‘지니’의 마스크, 2개는 ‘오피스퀵’의 A4 용지였다. 또한, ‘오피스퀵’은 판매 빈도와 판매 금액을 기준으로 나열한 상위 10개 기업에 모두 속하였고 RFM_score의 값이 12였기에 유의미한 결과를 도출할 수 있을 것으로 예상된다. 특가 이벤트 외에도 A4 용지를 구매한 고객들이 어떤 상품을 함께 샀는지 살펴봄으로써 묶음 상품을 구성해보면 좋을 것이다. 혹은 이들을 대상으로 해당 상품 중 하나를 선택하면 기존 가격보다 저렴하게 구입할 수 있는 방식을 적용할 수 있을 것으로 보인다.