고객구매여정을 통한 nasmedia 기여 고객군 선별

구매전환율과 이탈률을 중심으로

Team 나이스 크루 김서윤 김채원 백서희 이동근 이은지

Index

팀목표 03

상황분석 04

문제정의 및 목표수립 07

분석 스토리보드 08

용어 정의 09

그룹별 차이 15

모델링 18

대시보드 28

인사이트 도출 36

프로젝트 의의 46

팀 나이스크루의

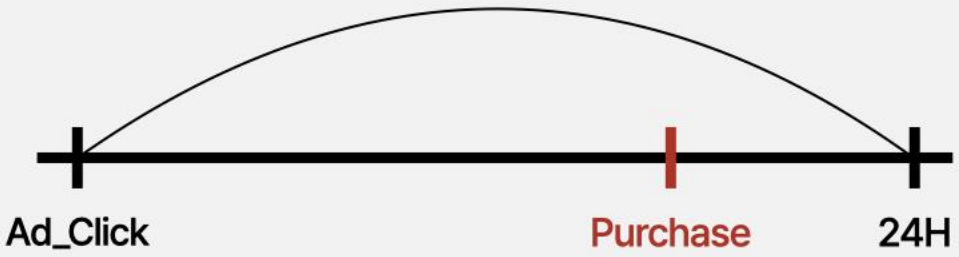
주요 목표

: 나스미디어의 수익 증진

저희 팀은 나스미디어 수익 증진을 목표로
고객구매여정을 통한 **나스미디어 기여 고객군을 선별**하고자 합니다.

상황분석

나스미디어의 **수익 충족 조건**



나스미디어 광고클릭 후, 24시간 내에 구매로 연결

⋮

고객이 제품을 처음 인지하게 되는 시점부터 24시간 내에
구매로 이어지기 위해선 **고객구매여정** 고려가 필요

상황분석

고객구매여정 분석의 중요성

01
개인 / 그룹 별 행동 패턴 확인

개인 및 그룹 별 구매여정 패턴을 이해하여,
나스미디어 **기여 고객군 특성 파악**

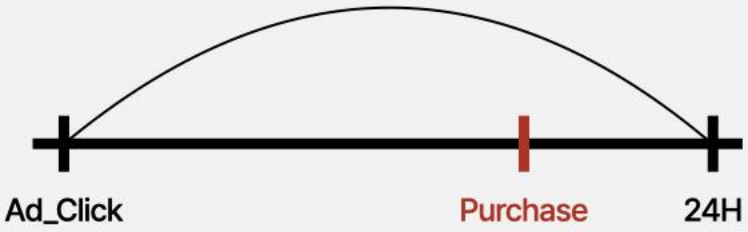
02
구매전환율과 이탈률 파악

구매전환율과 이탈률을 예측하여,
나스미디어 **기여 고객군 선별**

고객구매여정을 분석해 구매전환율 및 이탈률을 예측하고,
나스미디어에 기여할 수 있는 고객군을 선별하고자 합니다.

문제정의

주요 분석 요소



광고 클릭 후 **24시간 이내에 구매로 이어졌을 때 수익이 발생**하므로

고객구매여정을 바탕으로 **구매 전환율** 및 **이탈률** 예측이 필요

⋮

구매전환율 & 이탈률 예측을 위해선

이벤트 사이의 Interval에 대한 분석이 필요.

목표수립

주요 목표

구매전환율 & 이탈률
예측 AI 모델 개발

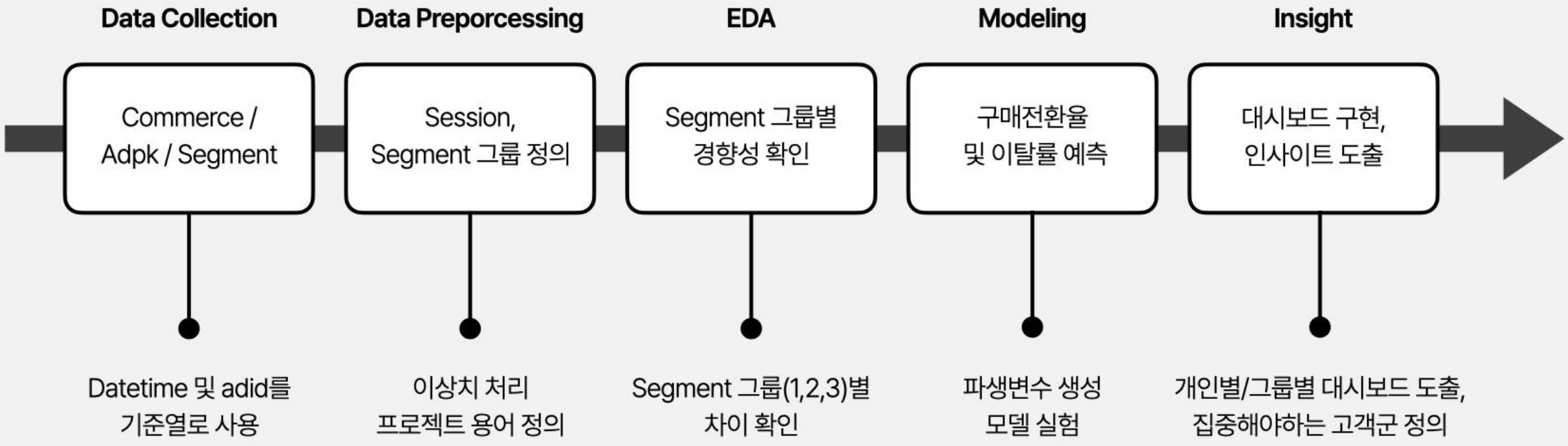
대시보드를 통한
인사이트 도출

❶ **그룹**을 기준으로 **일반화**된 대시보드 도출

❷ **개별 고객**을 기준으로 **개인화**된 대시보드 도출

분석 계획 수립

분석 스토리보드



프로젝트 **용어 정의**

01
구매전환율과 이탈률

예측하고자 하는 Target 값

02
Session

고객구매여정 단위

03
Segment 그룹

유사한 패턴을 가진 고객 그룹

용어 정의

구매전환율 & 이탈률 정의

예측하고자 하는 Target 값

구매전환율

광고를 클릭하여 사이트에 들어온 방문자가
구매, 장바구니 등 광고주가 원하는 특정 행위로 전환 하는 비율

⌄

광고 클릭으로 사이트에 들어온 방문자가 구매할 확률

이탈률

특정 기간 동안 앱 사용을 중단한 사용자의 비율

⌄

30분동안 아무런 이벤트 발생이 없어서 구매여정이 끊길 확률
(나스미디어 ad_conversion 기준이 24시간이므로, 이탈 역시 동일 기준 적용)

용어 정의

Session 정의

고객구매여정 단위

	date_time	event_nm	
7364	2024-01-03 14:58:12	① ad_clicked	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758c
7369	2024-01-03 15:02:49	search_item	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758c
7370	2024-01-03 15:02:57	search_item	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758c
7371	2024-01-03 15:04:10	items_clicked	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758c
7372	2024-01-03 15:05:00	items_clicked	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758c
7374	2024-01-03 15:05:23	purchase	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758c
7375	2024-01-03 15:05:23	ad_conversion	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758c
7841	2024-01-03 18:07:04	② ad_clicked	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758d
8364	2024-01-03 20:52:00	items_clicked	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758d
8365	2024-01-03 20:52:03	add_to_cart	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758d
8367	2024-01-03 20:52:08	items_clicked	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758d
8369	2024-01-03 20:52:33	ad_conversion	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758d
8370	2024-01-03 20:52:33	purchase	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758d

동일한 adid에 대해 **여러 번의 광고 클릭이 발생**하므로
각 클릭에서 비롯된 **고객구매여정을 개별적으로 고려**해야함.

용어 정의

Session 정의

고객구매여정 단위

동일 adid Ad_click, search, item_click, Ad_click, cart, item_click

Session 1 동일 adid Ad_click, search, item_click

Session 2 동일 adid Ad_click, cart, item_click

Session은 **Ad_click**으로 끊어지는 새로운 구매여정

용어 정의

Session 정의

고객구매여정 단위

		date_time	event_nm	session	session_idx
1	7364	2024-01-03 14:58:12	ad_clicked	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758da574d_1	1
	7369	2024-01-03 15:02:49	search_item	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758da574d_1	1
	7370	2024-01-03 15:02:57	search_item	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758da574d_1	1
	7371	2024-01-03 15:04:10	items_clicked	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758da574d_1	1
	7372	2024-01-03 15:05:00	items_clicked	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758da574d_1	1
	7374	2024-01-03 15:05:23	purchase	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758da574d_1	1
	7375	2024-01-03 15:05:23	ad_conversion	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758da574d_1	1
2	7841	2024-01-03 18:07:04	ad_clicked	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758da574d_2	2
	8364	2024-01-03 20:52:00	items_clicked	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758da574d_2	2
	8365	2024-01-03 20:52:03	add_to_cart	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758da574d_2	2
	8367	2024-01-03 20:52:08	items_clicked	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758da574d_2	2
	8369	2024-01-03 20:52:33	ad_conversion	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758da574d_2	2
	8370	2024-01-03 20:52:33	purchase	0f729c7a8a292712358c659361491a1f2dfc4df6474636a5aab083d758da574d_2	2

용어 정의

Segment 그룹 정의

나스미디어 수익 창출 조건 = **24시간 이내 구매**
이를 기준으로 **Segment 그룹 정의**

Group 1

광고 Click O
24시간 이내 구매 O

Group 2

광고 Click O
24시간 이후에 구매 O

Group 3

광고 Click O
한 세션 동안 구매 X

그룹 정의

각 그룹은 확실한 **차이점**을 보일까?

각 그룹별 Event Interval 차이, adid별 Session 수 차이를
분석해 **그룹별 차이점**을 도출했습니다.

각 그룹의 차이점

그룹별 Event Interval 차이

Group 1

평균
0일 0시간 28분 28초

중앙값
0일 0시간 10분 59초

Group 2

평균
0일 1시간 06분 42초

중앙값
0일 0시간 40분 18초

Group 3

평균
0일 1시간 45분 45초

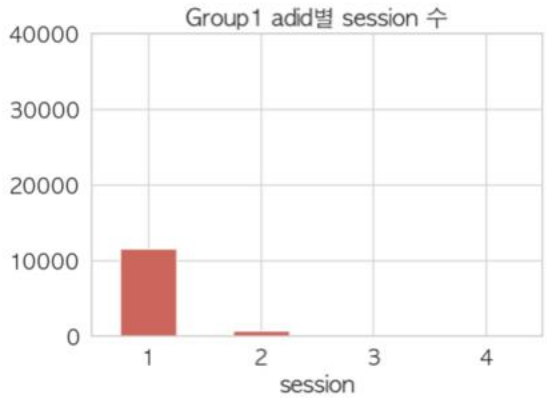
중앙값
0일 0시간 50분 36초

Event Interval Group1 << Group2 << Group3

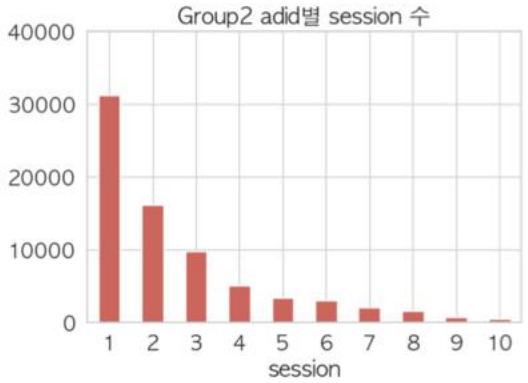
각 그룹의 차이점

adid별 Session 수 차이

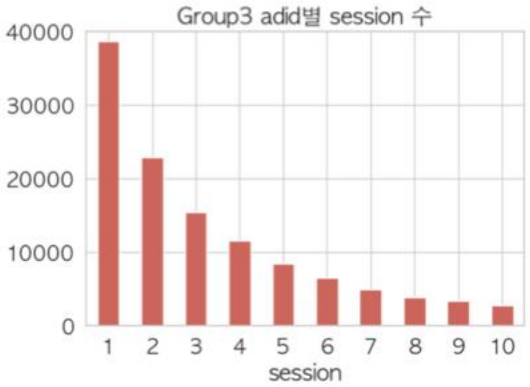
Group 1



Group 2



Group 3



Session 수 Group1 << Group2 << Group3

Session 비중 Group1 << Group2 << Group3

모델링

파생변수 생성 &
구매전환율 및 **이탈률** 예측 모델링

Session, Interval, Event와 관련된 **파생변수**를 생성하고,
각 Session의 Sequence에 초점을 맞춰 **예측 모델링**을 진행했습니다.

파생변수 생성

Session 관련

❶ **['session']**
adid + 각 Session 번호로 이루어진 고유한 변수

Interval 관련

❶ **['conversion_time']**
각 Event 사이의 Interval 에 대한 변수

❷ **['elapse_time']**
첫 Event부터 현재 Event까지의 누적 경과시간을 반영한 변수

Event 관련

❶ **['누적_검색횟수']**
❷ **[' 누적_장바구니횟수']**
❸ **['누적_상품클릭횟수']**
한 Session 안의 고객 행동 횟수를 누적해서 더한 변수

❺ **['관여도 점수']**
고객 행동(상품 클릭, 장바구니, 검색)의 횟수를 점수화한 변수로, 사용자의 정보탐색 정도를 종합하여 반영함.

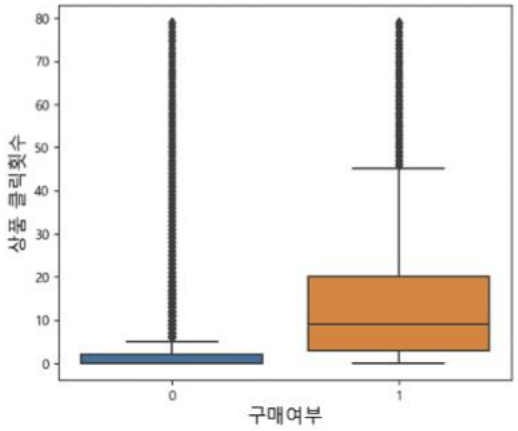
모델링

파생변수 생성

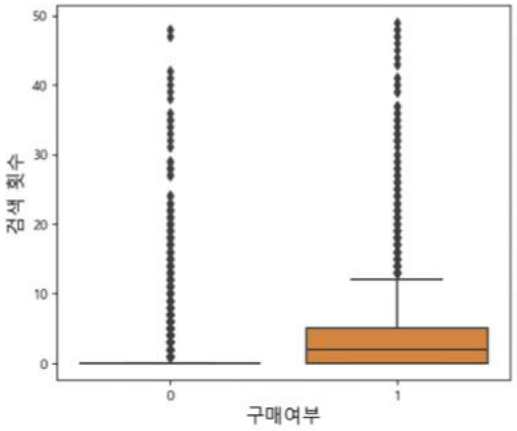
Event 관련

누적 횟수 : 한 Session 안의 고객 행동 횟수를 누적해서 더한 변수

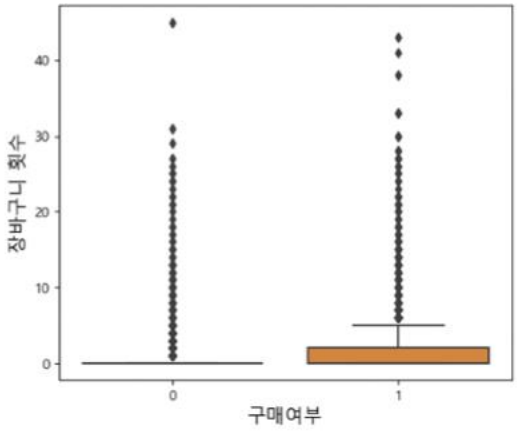
❶ ['누적_상품클릭횟수']



❷ ['누적_검색횟수']



❸ [' 누적_장바구니횟수']



파생변수 생성

Event 관련

관여도 점수 : 고객 행동(상품 클릭, 장바구니, 검색)의 횟수를 점수화한 변수

Session별 전체 행동(상품클릭, 검색, 장바구니) 횟수를 합산해 100점 기준으로 환산

Session	상품 클릭횟수	검색횟수	장바구니 횟수	관여도 점수
343ea4200bf5c.....	9	5	25	11.54

Session	상품 클릭횟수	검색횟수	장바구니 횟수	관여도 점수
343ea4200bf5c.....	279	23	15	100

* (Max 행동개수 = 319개 → 100점)

Sampling을 통한 데이터 불균형 해결

Under Sampling

데이터 Sampling 전

Lable	Target2(구매여부)	Target3(이탈여부)
0	84%	3%
1	16%	97%

데이터 Sampling 후

Lable	Target2(구매여부)	Target3(이탈여부)
0	50%	30%
1	50%	70%

모델링

주요 아이디어

1. Session 단위로 학습 진행

adid 01

Session 1
ad_click ··· purchase

각 session의 고객구매여정 sequence를 학습

Session 2
ad_click ··· cart

각 session의 고객구매여정 sequence를 학습

⋮

Session k
ad_click ··· search

각 session의 고객구매여정 sequence를 학습

구매 및 이탈과 직결되는 '고객구매여정'을 찾기 위하여 **각 Session의 Sequence를 학습**

모델링

주요 아이디어

2. 이전 Session의 영향력을 반영

adid 01

Session 1

ad_click ··· purchase

Session 2

ad_click ··· cart

⋮

Session k

ad_click ··· search

단, 이전 Seesion이 **24시간 이내**에 발생한 경우,
그 영향력을 고려하기 위해서 **이전 Session의 Sequence**를 **현재 Session**에 **반영**

* 이때, **이전 Session의 Sequence는 다운그레이드**

Session 단위로 학습을 진행하면서, **24시간 이내에 발생한 이전 session의 영향력**을 반영

모델링

사용 모델

기준 1. 모델 Size

< 구매전환율 예측 >

	GRU	LSTM	Transformer	CNN + Transformer
Loss	**0.3004**	0.3509	0.5061	0.5225
Accuracy	**0.8664**	0.8372	0.7527	0.7700
F1 Score	**0.8663**	0.8372	0.7476	0.7670

< 이탈률 예측 >

	GRU	LSTM	Transformer	CNN + Transformer
Loss	**0.1519**	0.2067	0.3547	0.3215
Accuracy	**0.9345**	0.9155	0.8356	0.8471
F1 Score	**0.8997**	0.8756	0.8024	0.8237

* A5000 GPU 1장에 대해 20 Epochs(Early Stopping 적용) 실험

모델링

사용 모델

기준 2. Attention Mechanism의 위치

< 구매전환율 예측 >

	Multi-head Attention	Local Attention	Global Attention	Hierarchical Attention	Local + Global Attention
Loss	0.2426	0.2709	**0.2506**	0.3075	**0.2212**
Accuracy	0.8893	0.8865	**0.8879**	0.8703	**0.8998**
F1 Score	0.8882	0.8865	**0.8876**	0.8703	**0.8996**

< 이탈률 예측 >

	Multi-head Attention	Local Attention	Global Attention	Hierarchical Attention	Local + Global Attention
Loss	0.2011	0.2068	0.2170	**0.1457**	**0.1324**
Accuracy	0.9219	0.9061	0.9116	**0.9463**	**0.9519**
F1 Score	0.8946	0.8807	0.8868	**0.9225**	**0.9300**

* **Hierarchical Attention**의 경우, 로그 이벤트를 초/분/시간 단위로 구분하여 알고리즘 구성

모델링

사용 모델

최종 모델 : Local Attention + Global Attention in Hierarchical GRU

Session 1

ad_click ··· purchase

초(S) **Feature Vector**

분(M) **Feature Vector**

시간(H) **Feature Vector**

GRU

다양한 시간 스케일에서의 패턴을 학습하기 위하여 **기존 GRU 구조에 Hierarchical 구조 반영**

대시보드

예측 구매전환율 및 이탈률을 활용한 **실시간 모니터링**을 위해서는?

모델에서 예측한 구매전환율과 이탈률을 바탕으로
고객 모니터링 대시보드를 구성했습니다.

Customer Monitoring Dashboard | Nasmdia Customer Journey Map

Group (전체) Gender (전체) Age (전체) Rfm Grade (전체)

KPI ▶

Interval

Total : 27s
Today : 29s

전날대비 ▲7.41%

25% : 11s
75% : 130s

예측 구매전환율

Total : 53.32%
Today : 38.18%

전날대비 ▼20.68%

25% : 0.21%
75% : 99.96%

예측 이탈률

Total : 91.23%
Today : 92.37%

전날대비 ▲0.02%

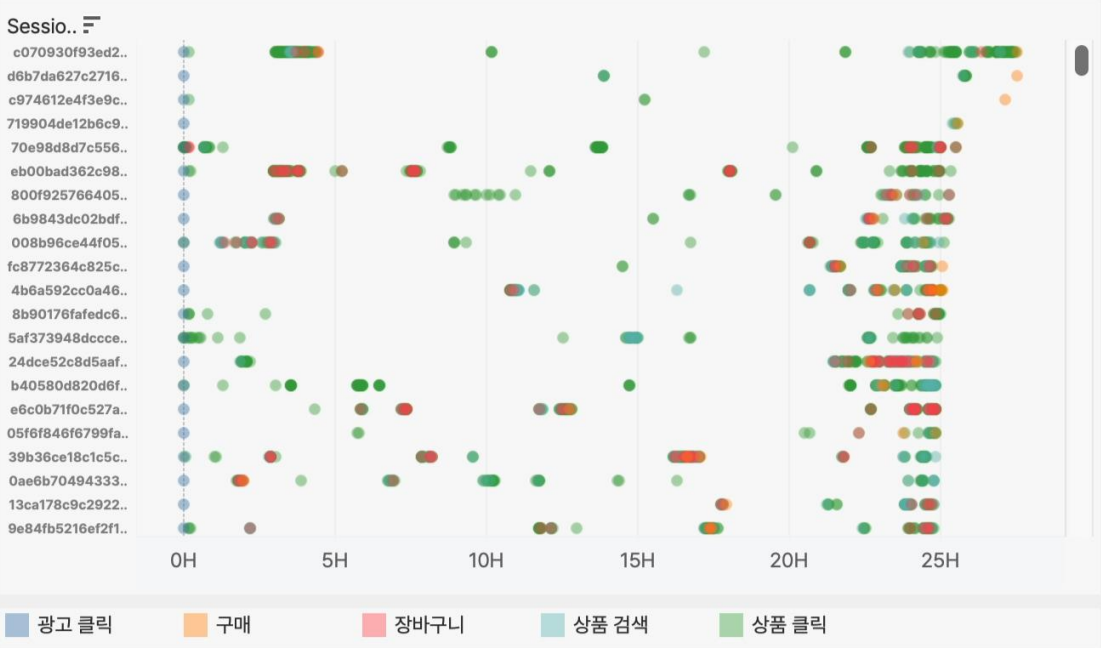
25% : 99.55%
75% : 100.00%

Involvement ▶

Customer List ▶

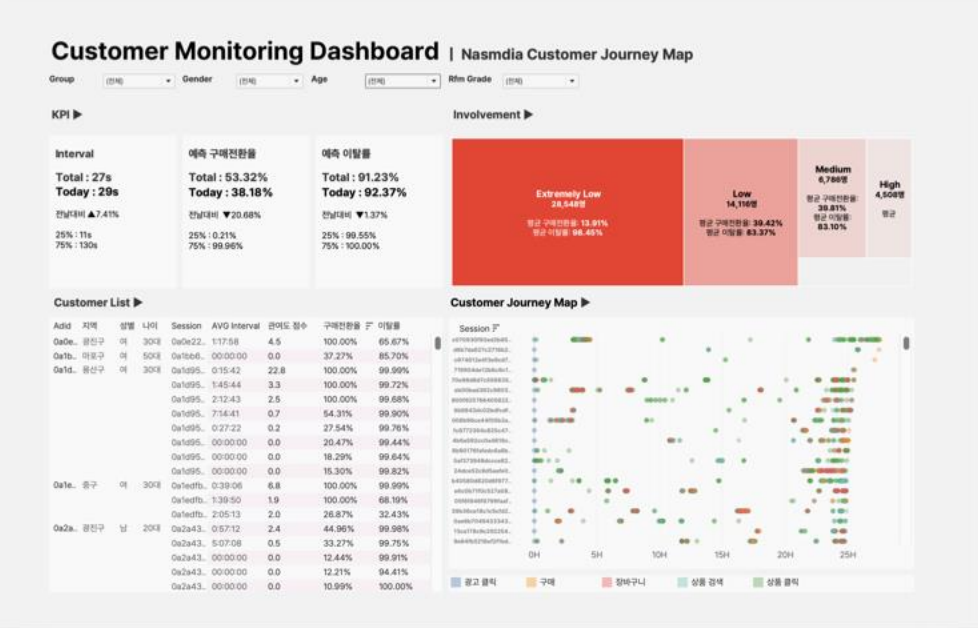
Adid	지역	성별	나이	Session	AVG Interval	관여도 점수	구매전환율	이탈률
54d..	동작구	남	50대	54d88e..	3:56:35	0.7	26.99%	99.75%
				54d88e..	2:25:07	0.5	23.66%	99.51%
54d..	강서구	여	50대	54d361..	1:37:51	1.0	79.99%	45.02%
				54d361..	00:00:00	0.0	25.62%	99.99%
				54d361..	00:00:00	0.0	24.24%	100.00%
54d..	양천구	남	50대	54dd2a..	0:06:02	12.7	100.00%	87.60%
				54dd2a..	0:35:58	10.3	100.00%	99.94%
				54dd2a..	0:05:58	0.0	100.00%	95.90%
				54dd2a..	0:14:54	5.0	100.00%	100.00%
				54dd2a..	1:19:59	1.2	100.00%	98.51%
				54dd2a..	0:59:21	1.0	45.29%	100.00%
				54dd2a..	0:24:44	0.7	28.42%	98.56%
				54dd2a..	0:02:03	0.7	21.98%	99.91%
				54dd2a..	2:41:15	1.0	11.62%	89.65%
				54dd2a..	0:03:48	0.7	11.18%	100.00%
54e6..	광진구	여	50대	54e6f0..	1:07:22	5.7	100.00%	99.83%
				54e6f0..	00:00:00	0.0	26.68%	97.92%
54e5	동작구	여	40대	54e552..	1:00:55	6.4	100.00%	99.33%

Customer Journey Map ▶



대시보드 설명

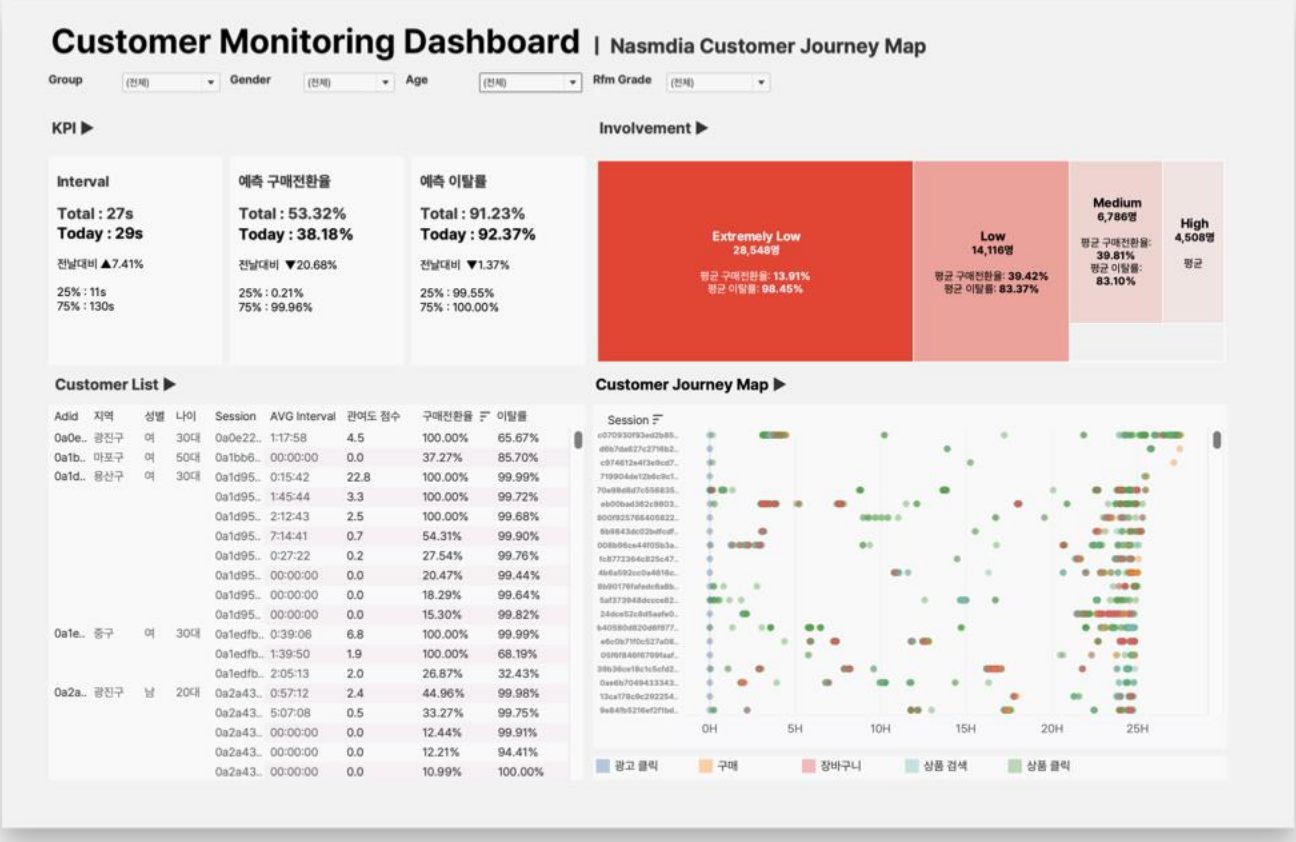
💡 **Tableau**를 사용하여 분석형 대시보드로 제작



대시보드 상단의 **고객 특성 필터**와 **관여도 점수 필터**를 조정하여
원하는 조건의 군집에 대한 KPI 지표, 고객 리스트, 고객여정지도를 확인할 수 있습니다.

대시보드

대시보드 구성요소



대시보드

대시보드 구성요소

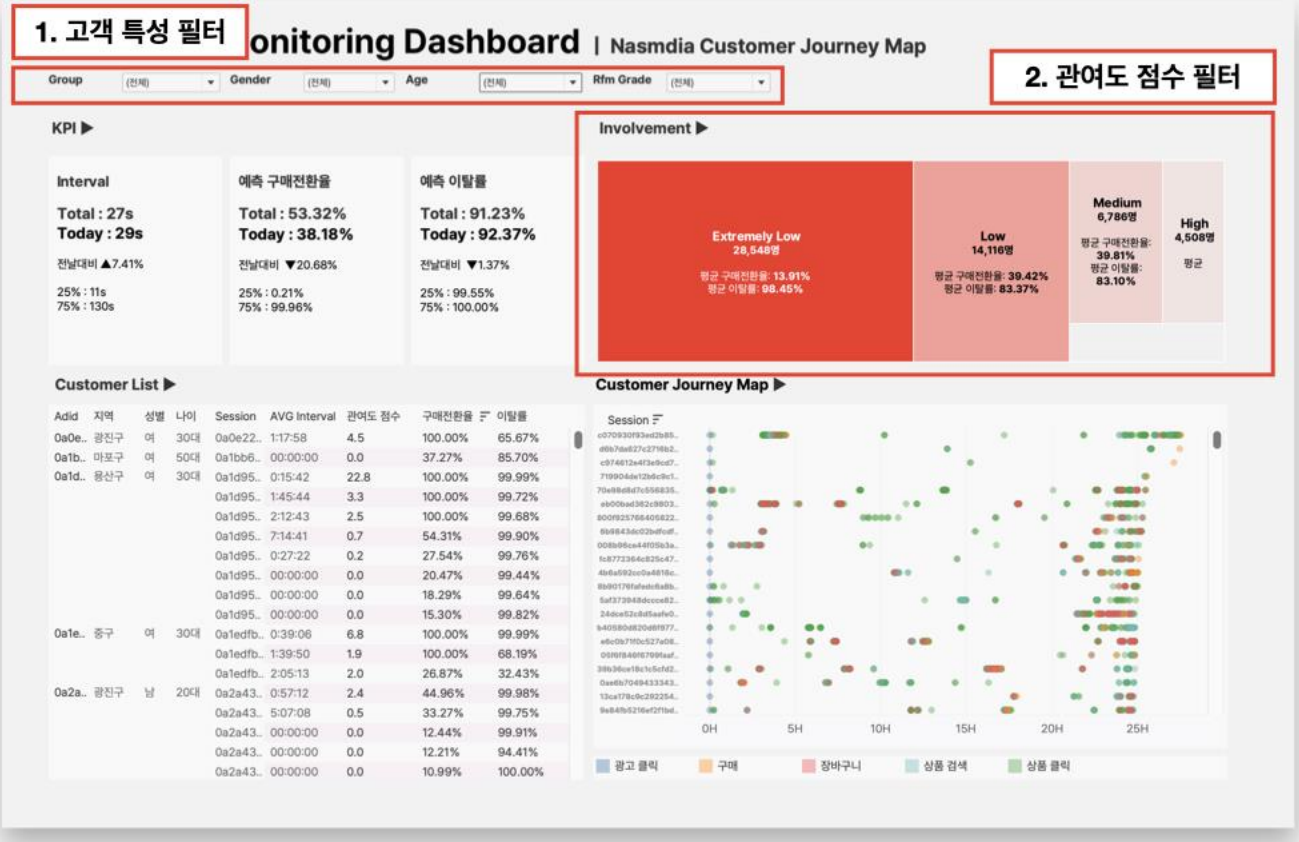
01 고객 특성 필터

: 그룹, 나이, 성별, RFM 등급으로 다양한 조건에 따라 데이터 필터링

02 관여도 점수 필터

: 관여도 점수를 5단계로 나눠 분류,
각 단계별 고객 수에 따라 맵 비율이 변화

→ **특정 조건에 맞는 고객 집중 분석 가능**



* 4분위수를 기준으로 Extremely High / High / Medium / Low /Extremely Low(0)로 분류

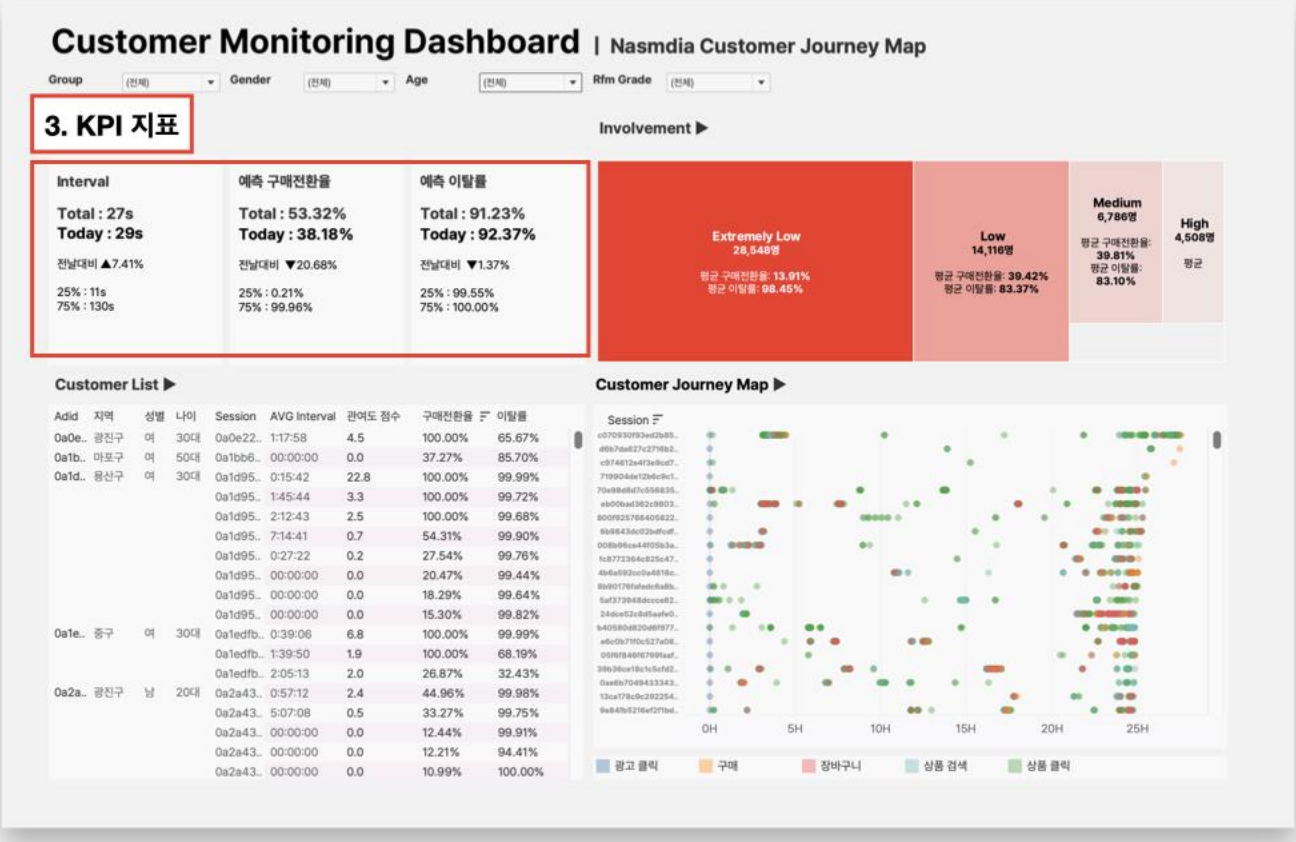
대시보드 구성요소

03 KPI 지표

: 위에서 필터링한 조건에 따른 KPI 지표 제공

- 전체 기간 Interval, 예측 구매전환율(평균), 예측 이탈률(평균)
- 당일 Interval, 예측 구매전환율(평균), 예측 이탈률(평균)
- 전날 대비 상승 및 하락 정도와 25%, 75% 구간도 함께 파악 가능

→ **시간에 따른 성과 변화를 모니터링**

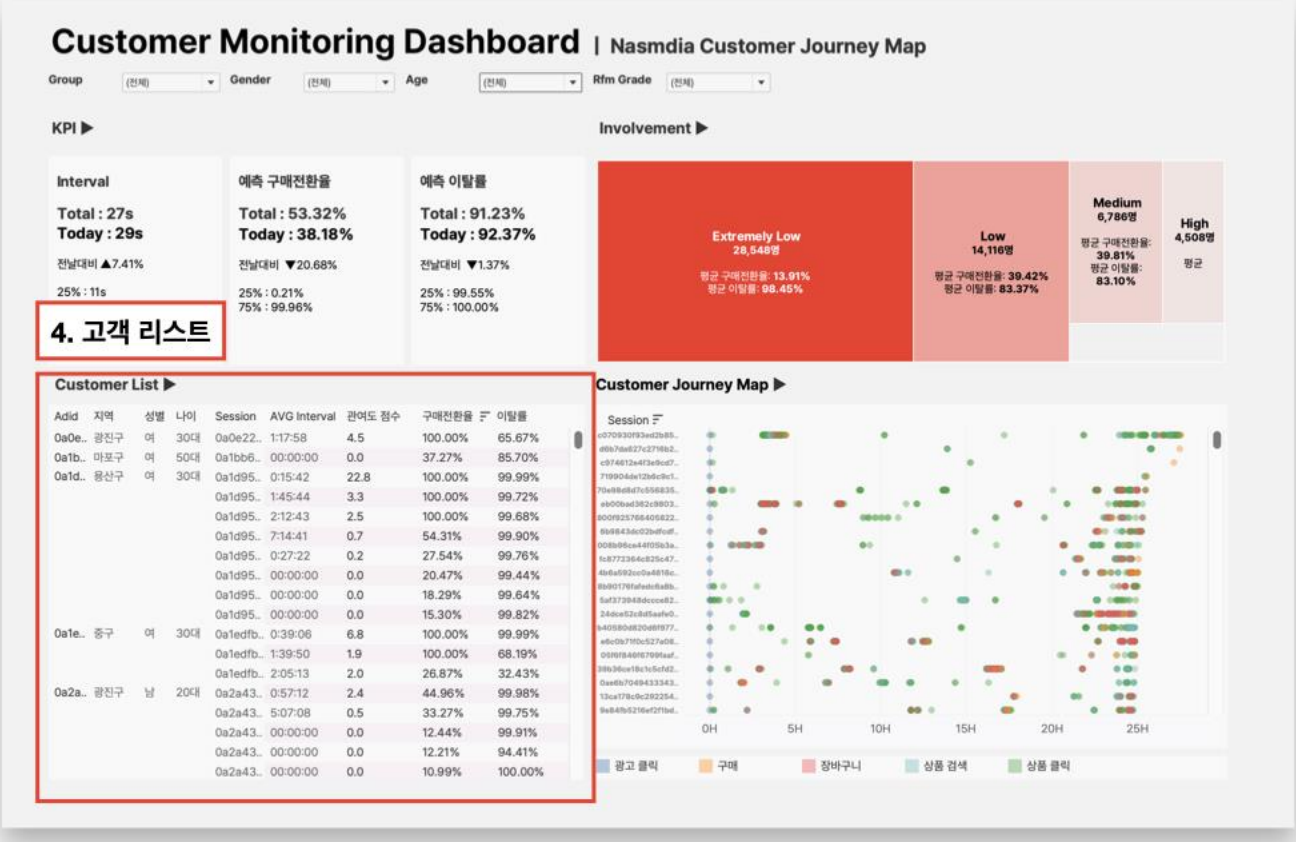


대시보드 구성요소

04 고객 리스트

- 위에서 필터링한 조건에 따라 고객리스트 출력
- adid, 지역, 성별, session, 평균 Interval, 관여도 점수, 구매전환율, 이탈률

→ **각 고객의 상세 정보 확인**

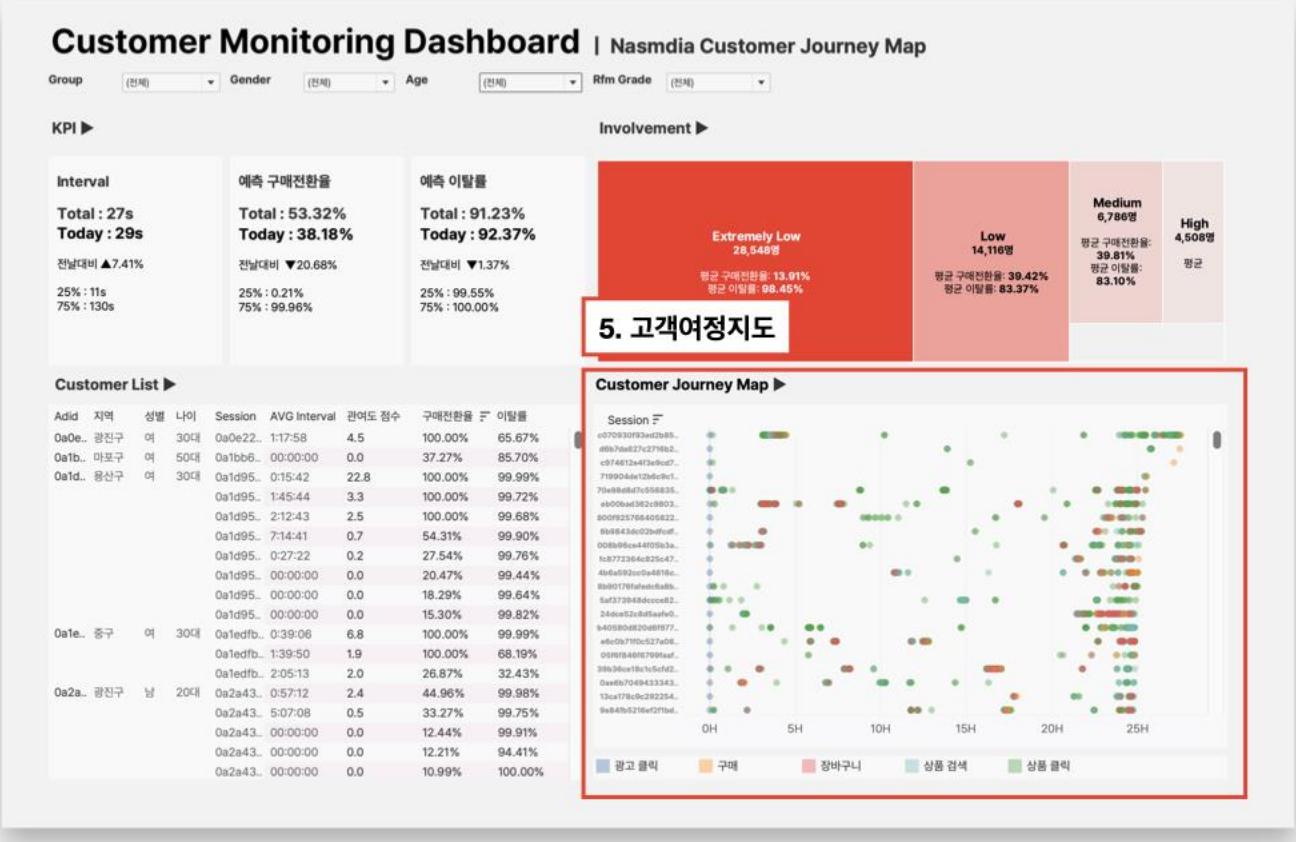


대시보드 구성요소

05 고객여정지도

- 각 session별 ad_click(00:00:00)으로부터
  가장 마지막 Event 시간까지 고객행동을 나타낸 그래프

→ <u>각 고객의 행동 패턴과 고객 여정의 주요 지점 이해</u>



인사이트 도출

대시보드로 알아보는 **그룹별 고객구매여정**은?

모델에서 예측한 구매전환율과 이탈률을 바탕으로
고객 모니터링 대시보드를 구성했습니다.

인사이트 도출

Group1 고객구매여정

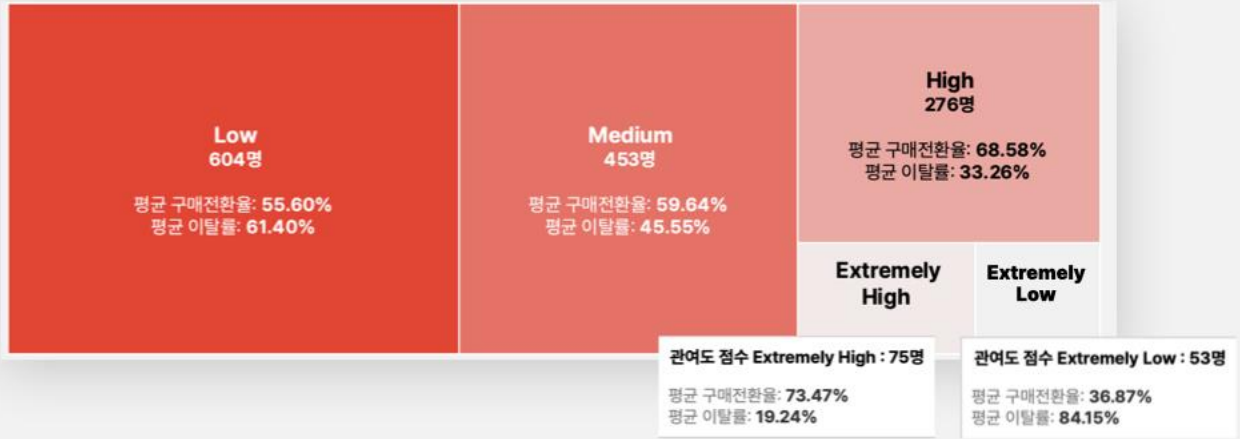
KPI

Interval	예측 구매전환율	예측 이탈률
Total : 21s	**Total : 73.27%**	**Total : 36.33%**
Today : 22s	**Today : 61.91%**	**Today : 52.32%**
전날대비 ▲0.00%	전날대비 ▼14.31%	전날대비 ▲40.72%
25% : 9s 75% : 58s	25% : 22.93% 75% : 99.96%	25% : 12.37% 75% : 96.07%

- **Interval** 중앙값 : **21초**
  75% 지점 : **58초**
- **평균 구매전환율** 65.88%
- **평균 이탈률** 36.33%

→ Interval이 짧은 고객군
→ 구매까지 걸리는 시간이 짧은(24시간 이내)고객군

Involvement



- **관여도 점수**(상위 3단계) **Low** >> Medium >> High

→ 상품 탐색에 소극적인 고객군 다수

Group1 : 상품 구매 의도가 분명한 고객군 (or 충동구매 고객군)

인사이트 도출

Group2 고객구매여정

KPI

Interval	예측 구매전환율	예측 이탈률
Total : 26s **Today : 29s**	**Total : 98.19%** **Today : 98.17%**	**Total : 97.10%** **Today : 99.41%**
전날대비 ▲11.54%	전날대비 ▼0.88%	전날대비 ▲0.30%
25% : 11s 75% : 85s	25% : 100.00% 75% : 100.00%	25% : 99.88% 75% : 100.00%

💡 **Interval** 중앙값 : **26초**
75% 지점 : **85초**

💡 **평균 구매전환율** 98.19%

💡 **평균 이탈률** 97.10%

→ Interval이 긴 고객군 (Group1 : 21초)
→ 구매까지 걸리는 시간이 긴(24시간 이후)고객군

Involvement

💡 **관여도 점수**(상위 3단계) **High** >>Extremely High >>Low

→ 상품 탐색에 적극적인 고객군 다수

Group2 : 상품 탐색 과정이 복잡해서 24시간 이후에 구매하는 고객군

인사이트 도출

Group3 고객구매여정

KPI

Interval	예측 구매전환율	예측 이탈률
Total : 31s **Today : 32s**	**Total : 12.37%** **Today : 11.25%**	**Total : 91.98%** **Today : 97.37%**
전날대비 ▲6.67%	전날대비 ▲9.78%	전날대비 ▲0.23%
25% : 12s 75% : 298s	25% : 0.02% 75% : 16.74%	25% : 99.67% 75% : 100.00%

- **Interval** 중앙값 : **31초**
  75% 지점 : **298초**
- **평균 구매전환율** 12.37%
- **평균 이탈률** 91.98%

→ Interval이 더욱 긴 고객군(Group2 : 26초)
→ 24시간 이후에도 구매를 하지 않는 고객군

Involvement

- **관여도 점수**(상위 3단계) **Extremely Low** >> Low >> Medium

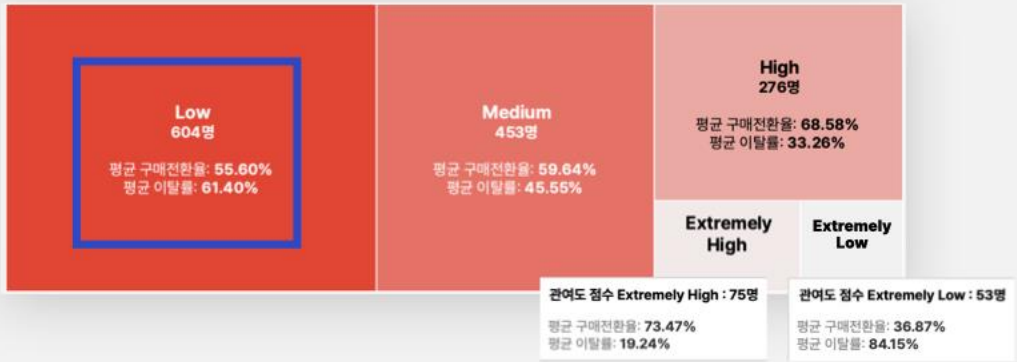
→ 고객 행동 수 = 0인 Session 다수
→ 상품 탐색에 매우 소극적인 고객군 다수

Group3 : 상품 탐색 과정이 늘어져 이탈하는 고객군 (or 잘못 클릭)

인사이트 도출

Group별 관여도 비교

Group1



Group2

Group3

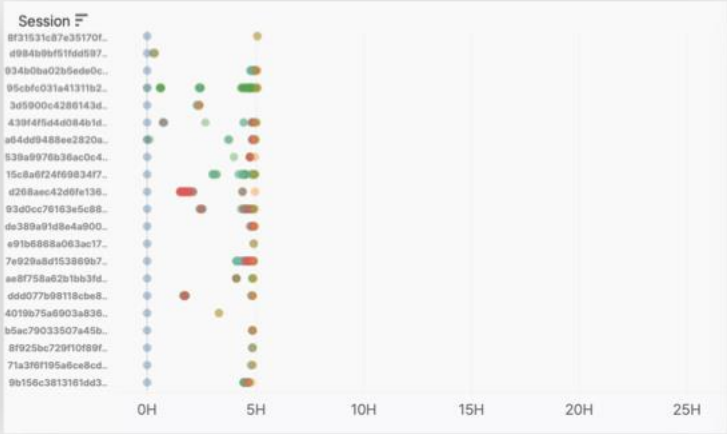
💡 **Group2**가 **Group1과 Group3**에 비해
관여도 점수가 매우 높다 (High ↔ Low, Extremely Low)

= 이벤트 수가 많다
= 상품 탐색 과정이 복잡하다

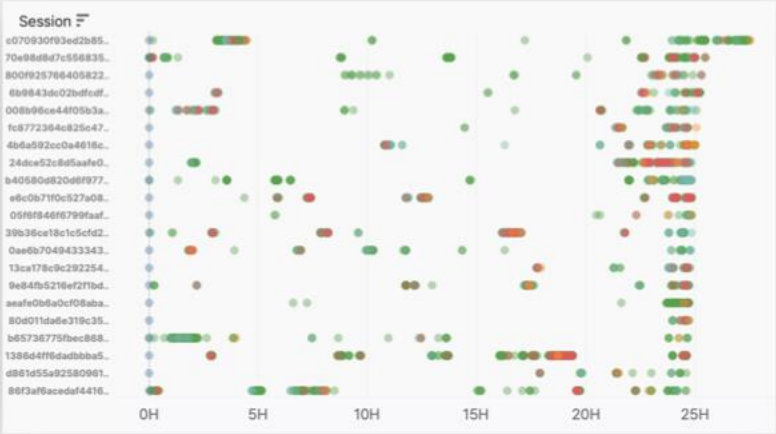
인사이트 도출

Group별 여정지도 비교

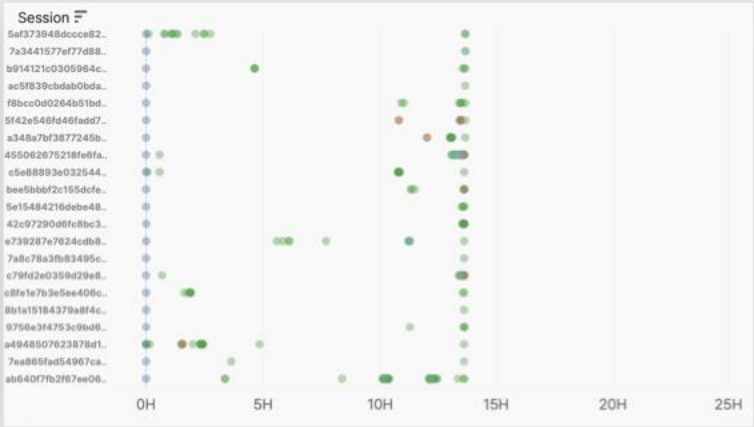
Group 1



Group 2



Group 3



광고 클릭 | 상품 클릭 | 상품 검색 | 장바구니 | 구매

💡 **Group2**가 Event 전환이 가장 활발하며, 구매까지 걸리는 시간이 길다.

인사이트 도출

나스미디어가
집중해야 하는 고객

Group2

: **관여도 점수**가 높고, **예상 이탈률**도 높은 고객군

→ Group1, Group3에 비해 상품 구매 시 **고민이 많은 Group** (긴 시간 동안 많은 행동을 수행)

💡 **고민시간을 단축시켜주면 구매로 이어질 확률이 가장 높은 Group**

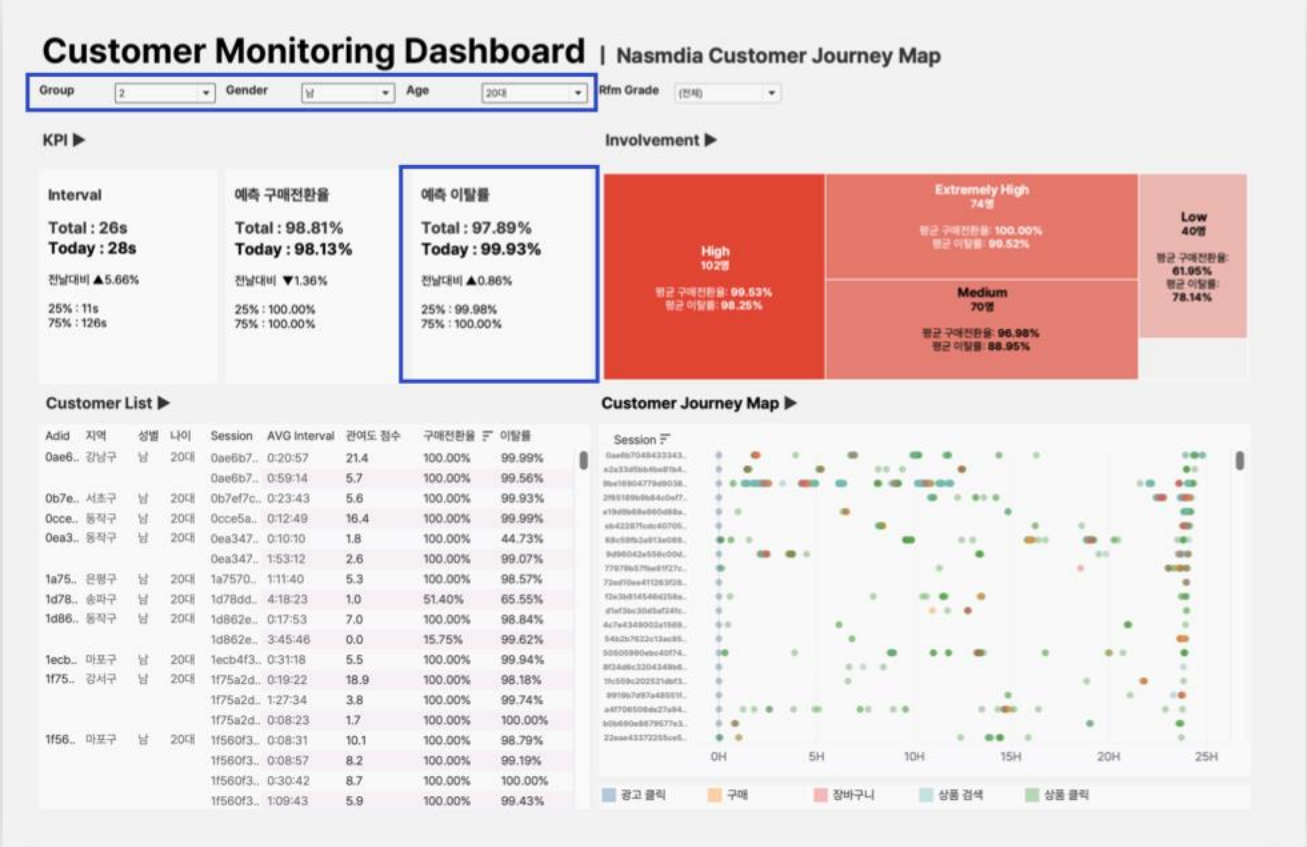
인사이트 도출

나스미디어가 집중해야 하는 고객

Group2

: 20대 남성

→ 예측 이탈률이 가장 높은 나이대와 성별



인사이트 도출

나스미디어가 집중해야 하는 고객

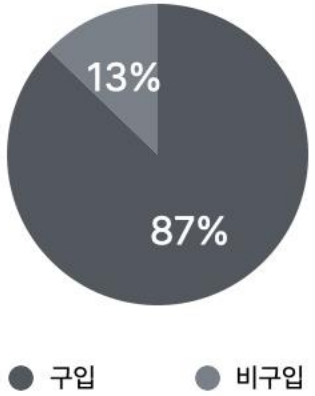
Group2

: 20대 남성 고객들의 소비 특징

정보 탐색 과정이 단순하다. (주체적인 정보 탐색 X)
→ 구매여정 절차를 **단순화**

신규 브랜드/신제품(저관여) 수용도 및 주요 구매 상황

수용도



구매 상황

주변 친구/지인이 추천할 때	40.2%
매장 점원이 추천할 때	30.7%
매장에서 테스트를 해보고	27.6%

[Base 스킨케어 화장품 직접 구매자, N-884명]
오픈 서베이, '남성 뷰티 카테고리 리포트 2020'

인사이트 도출

나스미디어가 집중해야 하는 고객

Group2

: 20대 남성 고객들의 소비 특징

정보 탐색 과정이 단순하다. (주체적인 정보 탐색 X)
→ 구매여정 절차를 **단순화**

불편을 줄이는 전략

- **복잡한 정보 제공하지 않기**
  : 빠르고 쉽게 받아들일 수 있도록 필요한 정보만 제공하자.

- **힉스의 법칙**
  : 선택지가 너무 다양하면 빠른 구매가 이루어지지 않을 수 있다.

절차의 단순화를 통해 **24시간 이내에 구매로 이어질 수 있게 해야 한다.**

프로젝트 의의

고객 행동 패턴 및 특성 파악 가능

개인 및 그룹별 구매여정을 패턴을 바탕으로 고객 행동 패턴 및 특성 파악 가능

인사이트 제공

예측 모델을 기반으로 한 실시간 모니터링 대시보드를 통해 인사이트 제공 가능

이를 바탕으로 나스미디어 기여 고객과 집중해야하는 고객군을 선별하고, 그 특성을 파악함으로써 수익 증진에 기여

감사합니다

Team 나이스 크루 김서윤 김채원 백서희 이동근 이은지