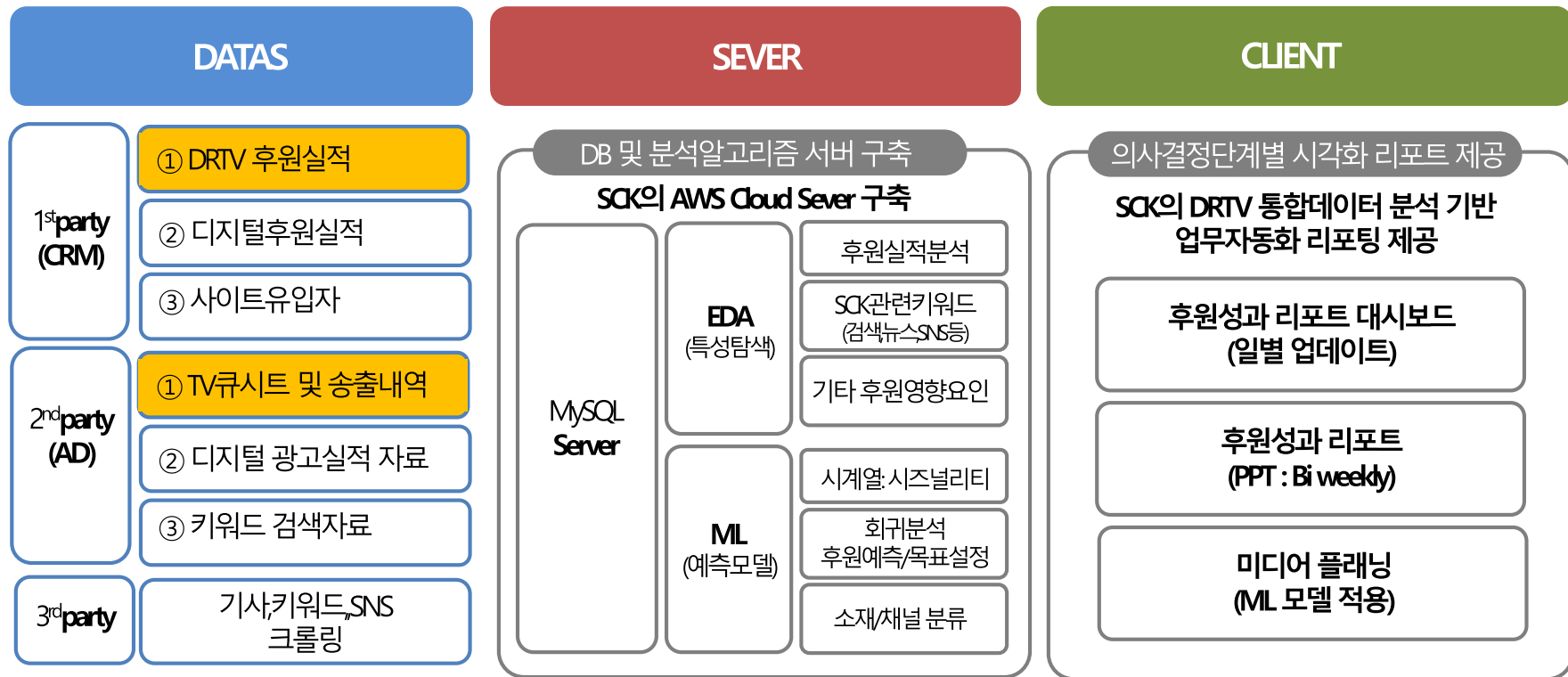


세이드더칠드런 DRTV 데이터분석 솔루션

세이브더칠드런 DRTV 데이터분석 시스템

각 파트에서 발생하는 데이터를 SCK DRTV 데이터 DB서버 구축하여 의사결정 단계에서 자동화 리포팅 제공



후원실적데이터와 DRTV 송출내역자료 분석에 용이한 형태로 정제 순차적으로 크롤링수집 데이터 및 디지털실적, CRM 데이터 추가

[후원실적데이터]

```
donour = pd.read_excel("./datas/Donour.xlsx", '@후원내역')
donour.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 34840 entries, 0 to 34839
Data columns (total 11 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Date        34840 non-null  datetime64[ns]
1   Time        34840 non-null  object
2   Daypart     34840 non-null  object
3   Dnis        34840 non-null  object
4   Station     34840 non-null  object
5   I/B콜수     34840 non-null  int64
6   응대호      34840 non-null  int64
7   정기후원    34840 non-null  int64
8   후원금액    34840 non-null  int64
9   Film        32258 non-null  object
10  연월        34840 non-null  object
dtypes: datetime64[ns](1), int64(4), object(6)
memory usage: 2.9+ MB
```

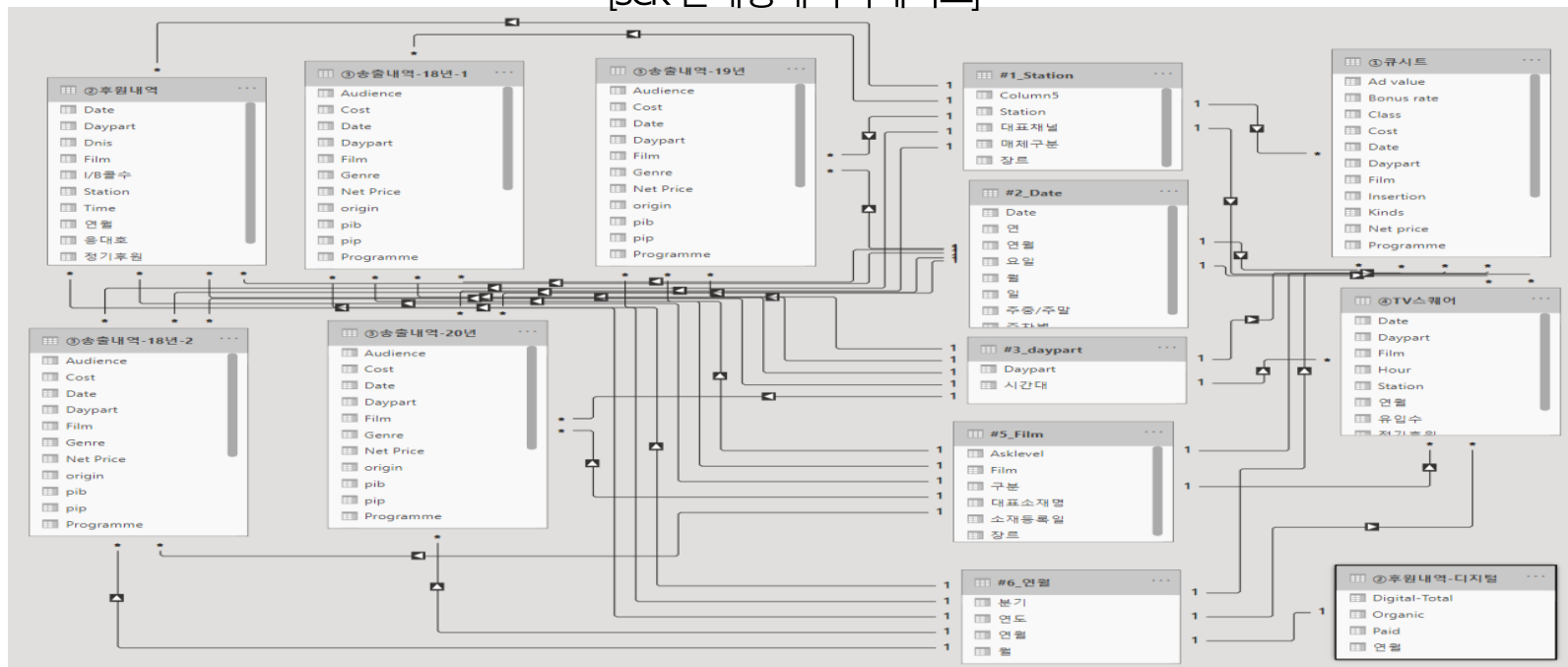
[DRTV송출내역]

```
rating = pd.concat([rating_18_1, rating_18_2, rating_19, rating_20], axis = 0).reset_
rating.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 48432 entries, 0 to 48431
Data columns (total 20 columns):
#   Column      Non-Null Count  Dtype
---  -
0   index       48432 non-null  int64
1   Station     48432 non-null  object
2   Date        48432 non-null  datetime64[ns]
3   Time        48432 non-null  object
4   Secs        48432 non-null  int64
5   Film        48432 non-null  object
6   Programme   48432 non-null  object
7   Genre       48337 non-null  object
8   Net Price   48432 non-null  float64
9   Cost        48432 non-null  float64
10  Audience    48432 non-null  float64
11  Salehouse   0 non-null      float64
12  promocode   0 non-null      float64
13  pib         48432 non-null  object
14  pip         47119 non-null  object
15  origin      0 non-null      float64
16  telno       0 non-null      float64
17  SPOT        48432 non-null  int64
18  Daypart     48432 non-null  object
19  연월        48432 non-null  object
dtypes: datetime64[ns](1), float64(7), int64(3), object(9)
memory usage: 7.4+ MB
```

소스가 다른 후원실적과 송출내역 데이터를 관계형데이터베이스로 설계
AWS클라우드에 아래 형태의 DB서버 구축하여 마케팅팀차원에서 데이터통합관리

[SCK 관계형데이터베이스]

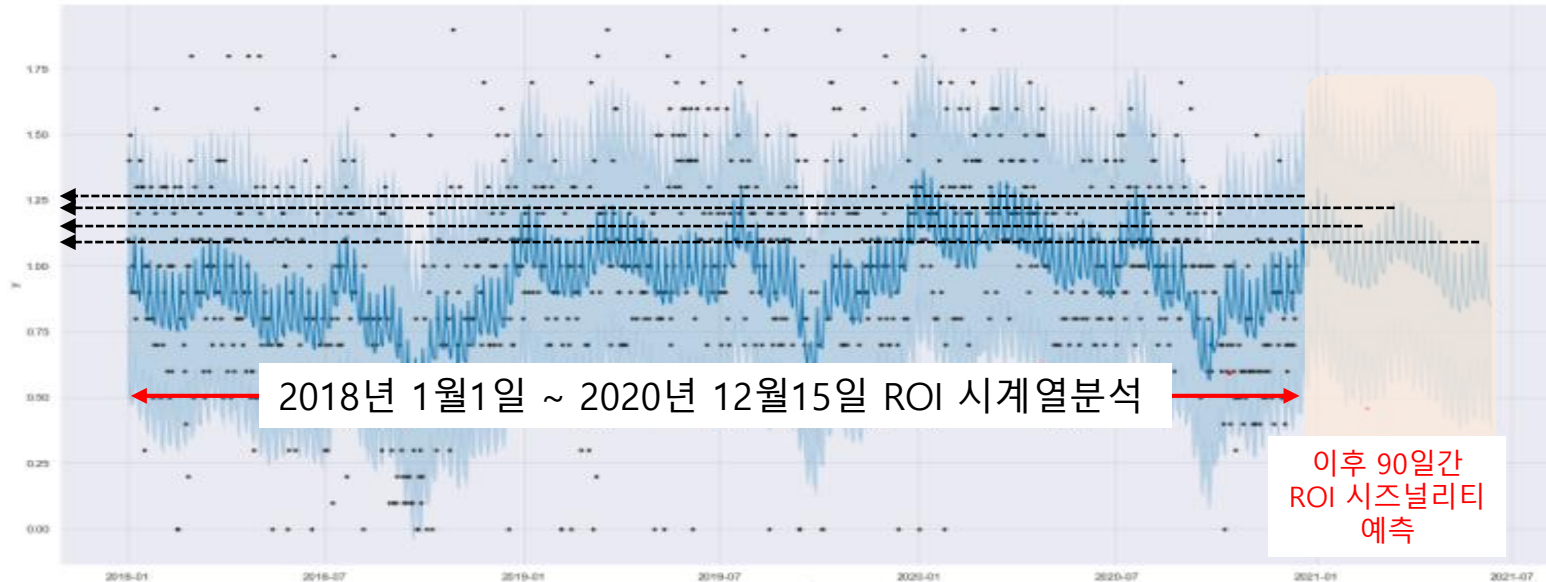


후원예측 ML 예시

① 시계열예측

SCK의 과거 후원 및 송출내역 자료에서 ROI를 추출하여
이후 특정기간(90일) 동안의 ROI 추이를 예측하여 예산 설정 등에 활용
1월, 3월, 6월에 ROI 상승하고 2월, 4월, 5월은 상대적으로 ROI 하락되는 시즌

[SCK 시계열 ROI 추이]



※ 실제 세이브더칠드런 데이터 분석결과입니다.

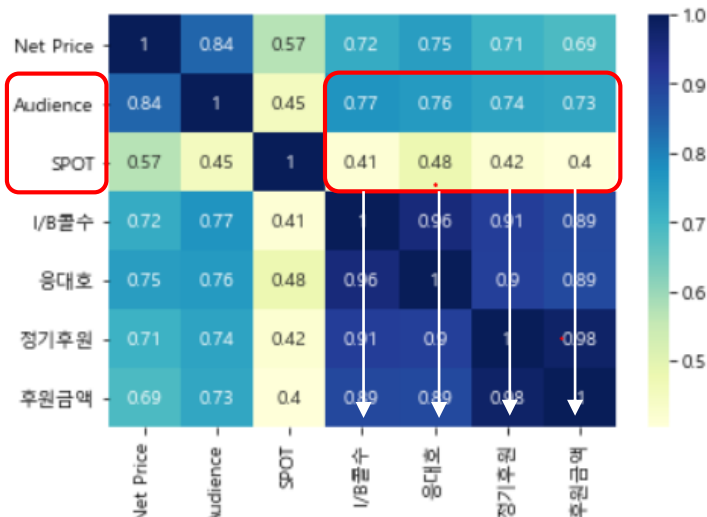
후원예측 ML 예시

② 회귀분석: 후원목표설정

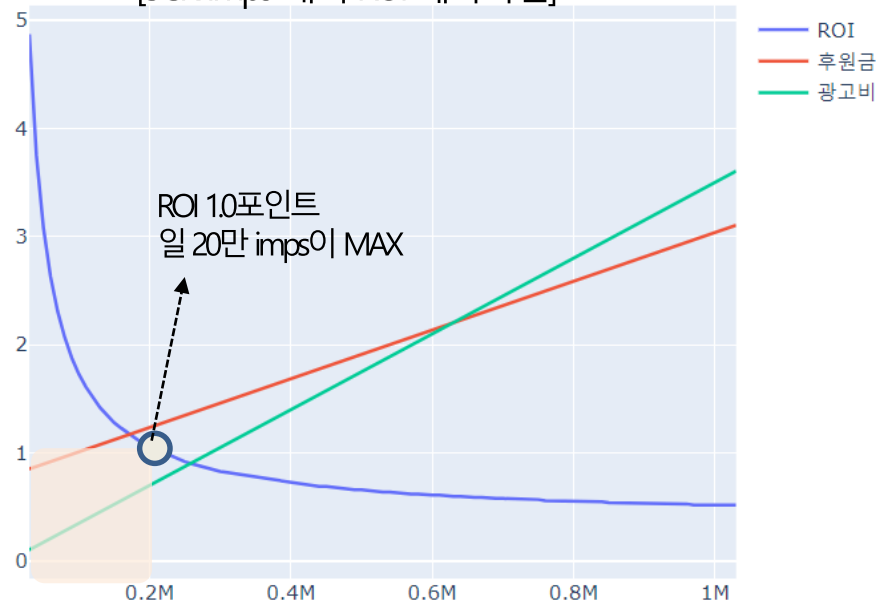
DRTV광고의 어떤 특성들이 IB콜, 후원개발에 영향을 주는 지 분석
 이 분석을 기반으로 ROI 목표 설정 및 적정 광고노출량 등 목표설정
 3년간의 데이터로 봤을 때 Audience-IB콜간의 관계 상관관계수 0.77으로 SPOT 수 보다 상관관계 높음

[SCK 각요소별 상관관계 분석]

```
sns.heatmap(datas_1.corr(), cmap="YlGnBu", annot=True)
plt.show()
```



[SCK Imps 대비 ROI 예측곡선]



※ 실제 세이브더칠드런 데이터 분석결과입니다.

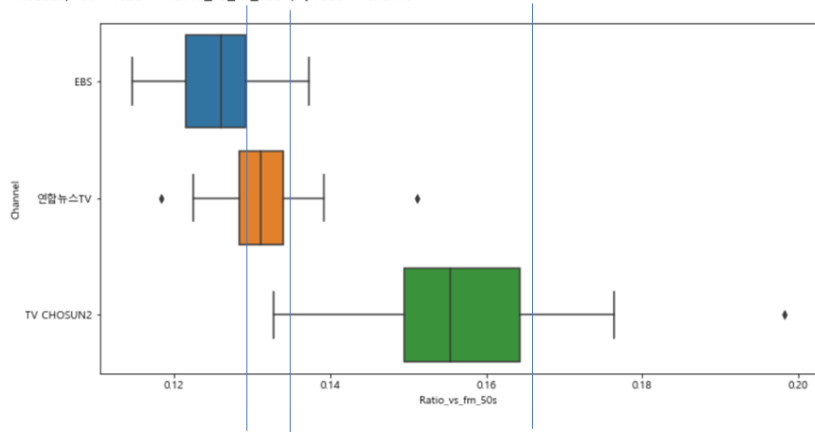
③ 분류모델 : 적정소재 및 유사채널등 파악

ROI효과채널인 EBS, 연합뉴스, TV Chosun2와 유사 채널 분류하기 위한 채널 성분분석
3개채널을 구분하는 유의미한 성분은 <평균시청률>, <50대여성시청비율>, <Reach5+>

[3개채널 분류 기준 ① 50대여성시청비율]

```
# 채널 시청자 중 50대 여성 비중
plt.figure(figsize=(12,6))
sns.boxplot(x='Ratio_vs_fm_50s', y='Channel', data=ch_component, orient='h')
```

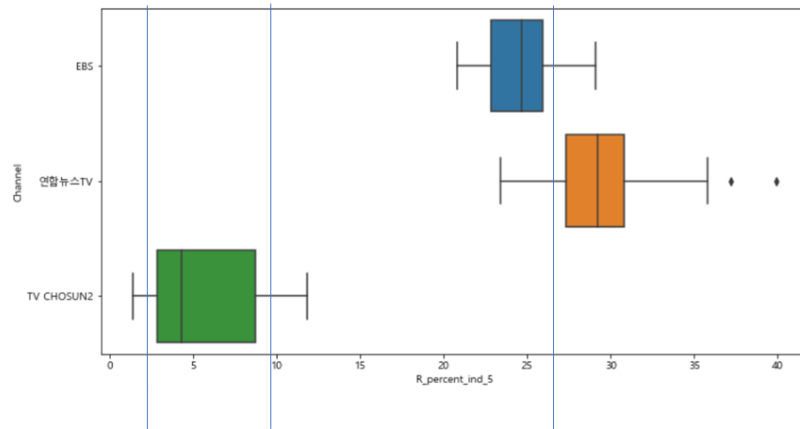
<AxesSubplot: xlabel='Ratio_vs_fm_50s', ylabel='Channel'>



[3개채널 분류 기준 ② Reach 5+ 비중]

```
# 채널의 Reach5+
plt.figure(figsize=(12,6))
sns.boxplot(x='R_percent_ind_5', y='Channel', data=ch_component, orient='h')
```

<AxesSubplot: xlabel='R_percent_ind_5', ylabel='Channel'>



※ 실제 세이브더칠드런 데이터 분석결과입니다.

③ 분류모델 : 적정소재 및 유사채널등 파악

성분분석하는 모델을 112개의 채널의 최근 1개월간의 성분으로 적용하여 분류
이 분류를 통해서 신규채널의 Test 및 적정소재를 선택하는 데 활용 가능

[112개 채널 중 3개 채널과 유사한 채널별로 분류]

X_ch[X_ch["ch_label"] == 1].head(10) # EBS

	Rating_ind	Ratio_vs_fm_50s	R_percent_ind_5	ch_label	Channel
62	0.062	0.148023	18.5	1	JTBC2
65	0.079	0.127799	22.0	1	KBS JOY
70	0.070	0.150309	19.5	1	KBS드라마
77	0.073	0.134871	22.1	1	MBC드라마넷
82	0.076	0.135352	24.1	1	OCN
83	0.045	0.135194	16.5	1	OCN Movies
86	0.061	0.145137	18.7	1	Olive
93	0.053	0.123300	16.5	1	SBS funE
96	0.092	0.137060	25.8	1	SBS Plus
99	0.052	0.145343	16.4	1	SKY

X_ch[X_ch["ch_label"] == 2].head(10) # TV Chosun2

	Rating_ind	Ratio_vs_fm_50s	R_percent_ind_5	ch_label	Channel
7	0.002	0.000000	0.1	2	대교베이비TV
8	0.021	0.157295	10.3	2	드라마H
9	0.009	0.125827	1.8	2	라이프타임
10	0.004	0.076202	0.3	2	리얼TV
11	0.016	0.091032	1.4	2	바둑TV
12	0.008	0.118263	1.3	2	발리어즈TV
13	0.044	0.047118	1.7	2	뿌요TV
14	0.008	0.034270	2.2	2	애니원
15	0.003	0.041508	0.3	2	애니플러스
16	0.005	0.014160	1.7	2	어린이TV

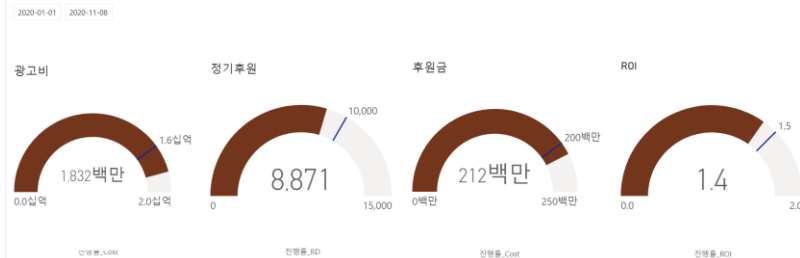
X_ch[X_ch["ch_label"] == 3].head(10) # 연합뉴스

	Rating_ind	Ratio_vs_fm_50s	R_percent_ind_5	ch_label	Channel
0	1.046	0.118338	57.2	3	KBS2
1	0.783	0.110310	62.7	3	MBC
2	1.005	0.112662	66.3	3	SBS
3	0.356	0.121634	46.7	3	JTBC
4	0.200	0.130613	34.9	3	채널A
5	0.234	0.136265	34.9	3	MBN
6	0.530	0.138861	37.3	3	TV CHOSUN
18	0.237	0.131531	27.1	3	연합뉴스TV
73	0.103	0.121697	26.8	3	MBC every1
111	0.407	0.115999	48.2	3	tvN

※ 실제 세이버더칠드런 데이터 분석결과입니다.

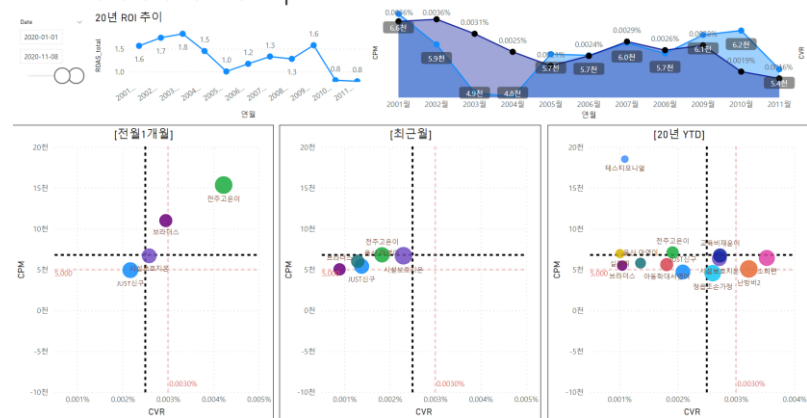
위클리리포트와 Prebuy Report 등 PPT 자동화로 업무 간소화
현재 사용 중인 세이브더칠드런 파워BI 대시보드에 추가 EDA 및 ML 등 내용 업데이트

세이브더칠드런 2020년 집행실적 개요



※ 실제 세이브더칠드런 파워BI 대시보드

SCK 성과최적화 매트릭스 | 소재



SCK DRTV 데이터분석 시스템 구축 스케줄

총 3개월 간 서버구축 - ML 초기 분석 모델 적용한 시스템 - 시각화

