

Sangyoon.s Portfolio

KEYWORDS

데이터 마이닝 (Airbnb new user bookings: 데이터분석)

브랜딩 (Plan A 창업 : 여행일정추천 앱서비스)

오퍼레이션 (D.CAMP : 사업개발팀 인턴)

PM (한국공간데이터: 프로젝트매니저 & 서비스기획)

KEYWORDS

● **데이터 마이닝** (Airbnb new user bookings: 데이터분석)

브랜딩 (Plan A 창업 : 여행일정추천 앱서비스)

오퍼레이션 (D.CAMP : 사업개발팀 인턴)

PM (한국공간데이터: 프로젝트매니저 & 서비스기획)

● 데이터 마이닝 (Airbnb new user bookings: 데이터분석)

분석 설계

```
'data.frame': 213451 obs. of 16 variables:
 $ id
47281 129610 2144 59779 40805 ...
 $ date_account_created : Factor w/ 1634 levels "2010-01-01","2010-01-02"...: 171 502 263 696 249 1 2 3 4 4 ...
 $ timestamp_first_active : num 2.01e+13 2.01e+13 2.01e+13 2.01e+13 2.01e+13 ...
 $ date_first_booking : Factor w/ 1977 levels "","2010-01-02"...: 1 1 194 960 35 2 4 10 190 3 ...
 $ gender : Factor w/ 4 levels "unknown","FEMALE"...: 1 2 2 1 1 2 2 1 ...
 $ age : num NA 38 55 42 41 NA 46 47 50 46 ...
 $ signup_method : Factor w/ 3 levels "basic","facebook"...: 2 2 1 2 1 1 1 1 1 ...
 $ signup_flow : int 0 0 3 0 0 0 0 0 0 ...
 $ language : Factor w/ 25 levels "ca","ca","da"...: 6 6 6 6 6 6 6 6 6 ...
 $ affiliate_channel : Factor w/ 8 levels "api","content"...: 3 8 3 3 3 4 4 3 4 ...
 $ affiliate_provider : Factor w/ 18 levels "adobe","bing"...: 6 9 5 5 5 13 3 5 3 ...
 $ first_affiliate_tracked : Factor w/ 8 levels "","linked","social opt"...: 8 8 8 8 8 8 8 5 ...
 $ signup_app : Factor w/ 4 levels "Android","IOS"...: 4 4 4 4 4 4 4 4 ...
 $ first_device_type : Factor w/ 9 levels "Android Phone"...: 6 6 9 6 6 6 6 6 ...
 $ first_browser : Factor w/ 52 levels "unknown","Android Browser"...: 1 9 9 24 16 9 9 42 42 16 ...
 $ country_destination : Factor w/ 12 levels "AU","CA","DE"...: 8 8 12 10 12 12 12 12 12 ...

1 gxn9p8hntm 2010-06-28 2.008032e+13 unknown NA facebook
2 820tgsjqx7 2011-05-25 2.009052e+13 MALE 38 facebook
3 4rt3pwnxta 2010-09-28 2.002061e+13 2010-08-02 FEMALE 56 basic
4 bjl56pjhuk 2011-12-05 2.009103e+13 2012-09-08 FEMALE 42 facebook
5 6f7eehulop4 2010-09-14 2.009121e+13 2010-02-16 unknown 41 basic
6 osr2jw1jor 2010-01-01 2.010010e+13 2010-01-02 unknown NA basic

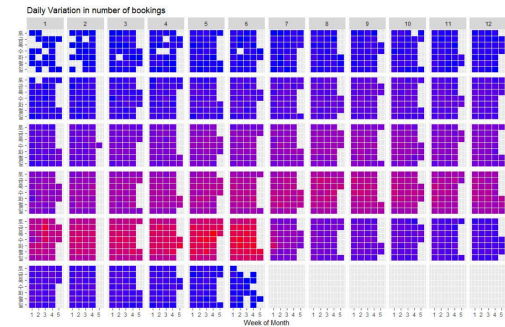
 $ signup_flow language affiliate_channel affiliate_provider first_affiliate_tracked signup_app
1 0 en direct google untracked Web
2 0 en direct google untracked Web
3 3 en direct untracked Web
4 0 en direct untracked Web
5 0 en direct untracked Web
6 0 en other org Web

 $ first_device_type first_browser country_destination
1 Mac Desktop Chrome NCF
2 Mac Desktop Chrome NCF
3 Windows Desktop IE US
4 Mac Desktop Firefox other
5 Mac Desktop Chrome US
6 Mac Desktop Chrome US
```

1차 데이터 탐색
: 유저들의 가입경로 및 예약정보 파악

1차 데이터 시각화
: 데이터의 상관관계를 통한 유의미한 패턴 파악

Heat map



<데이터 전처리>

<신규유저 히트맵 분석>

● 데이터 마이닝 (Airbnb new user bookings: 데이터분석)

Categorical analysis

```
> summary(model)

Call:
glm(formula = booking ~ gender + language_class + affiliate_channel +
  affiliate_provider_class + signup_app + signup_method + first_device_type,
  family = binomial(link = "logit"), data = user.train)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.7270 -0.9680 -0.7473  1.2102  2.5920

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 203115  on 149578  degrees of freedom
Residual deviance: 189371  on 149552  degrees of freedom
AIC: 189425

Number of Fisher Scoring iterations: 4
```

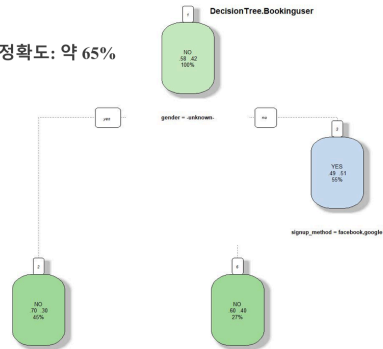
<신규유저 유입경로 모델링>

Tree 모형 생성

Confusion Matrix and Statistics

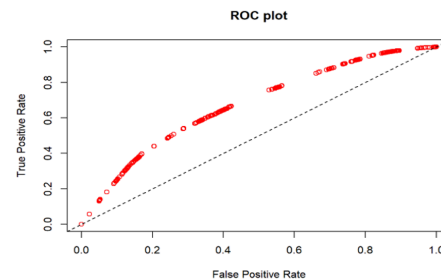
	Reference	
Prediction	NO YES	
NO	30029 15741	
YES	6751 10956	
Accuracy	0.649	
95% CI	(0.6455, 0.6527)	
No Information Rate	0.5854	
P-Value [Acc > NIR]	< 2.2e-16	
Kappa	0.2414	
McNemar's Test P-Value	< 2.2e-16	
Sensitivity	0.8194	
Specificity	0.4104	
Pos Pred Value	0.6905	
Neg Pred Value	0.6187	
Prevalence	0.5834	
Detection Rate	0.4780	
Detection Prevalence	0.7237	
Balanced Accuracy	0.6149	
'Positive' Class	NO	

정확도: 약 65%



<Decision Tree분석>

ROC plot



<ROC 분석>

Random Forest

Reference

	Reference	
Prediction	NO YES	
NO	30403 15792	
YES	6752 10960	
Accuracy	0.6472	
95% CI	(0.6435, 0.6509)	
No Information Rate	0.5814	
P-Value [Acc > NIR]	< 2.2e-16	
Kappa	0.2393	
McNemar's Test P-Value	< 2.2e-16	
Sensitivity	0.8183	
Specificity	0.4007	
Pos Pred Value	0.6581	
Neg Pred Value	0.6188	
Prevalence	0.5814	
Detection Rate	0.4757	
Detection Prevalence	0.7225	
Balanced Accuracy	0.6140	
'Positive' Class	NO	

```
> booking_rf
Model Details:
=====
H2OBinomialModel: drf
Model ID: DRF_model_A_154540885929_939
Model Summary:
  number_of_trees number_of_internal_trees model_size_in_bytes min_depth max_depth mean_depth min_leaves
1             50                50              21281             5             5             5
2             50                50              21281             5             5             5
3             50                50              21281             5             5             5
4             50                50              21281             5             5             5
5             50                50              21281             5             5             5
6             50                50              21281             5             5             5
7             50                50              21281             5             5             5
8             50                50              21281             5             5             5
9             50                50              21281             5             5             5
10            50                50              21281             5             5             5

H2OBinomialMetrics: drf
** Reported on training data. **
** Metrics reported on Out-Of-Bag training samples **

MSE: 0.221903
RMSE: 0.4710599
LogLoss: 0.6208233
Mean Per-Class Error: 0.4221559
AUC: 0.6788105
Gini: 0.5476211

Confusion Matrix (vertical: actual; across: predicted) for F1-optimal threshold:
      NO YES Error Rate
NO 21500 85968 0.755230 ~65996/87388
YES 6537 56610 0.580482 ~5537/62156
Totals 26027 122617 0.476354 ~71535/149544

Maximum Metrics: Maximum metrics at their respective thresholds
  metric threshold value use
1      max f1 0.329135 0.612849 267
2      max f2 0.327468 0.784547 325
3      max f0point0 0.444256 0.563953 134
4      max accuracy 0.489392 0.648928 19
5      max precision 0.590272 0.682785 19
6      max recall 0.076277 1.000000 388
7      max specificity 0.628225 0.999966 0
8      max absolute_auc 0.467807 0.252245 150
9      max min_per_class_accuracy 0.417519 0.622564 176
10     max max_per_class_accuracy 0.455008 0.604840 143

Gains/Lift Table: Extract with 'h2o.gainLift(<model>, <data>)' or 'h2o.gainLift(<model>, valid=<T/F>, xval=<T/F>)'
```

정확도: 약 65.02%

<Random Forest 분석>

KEYWORDS

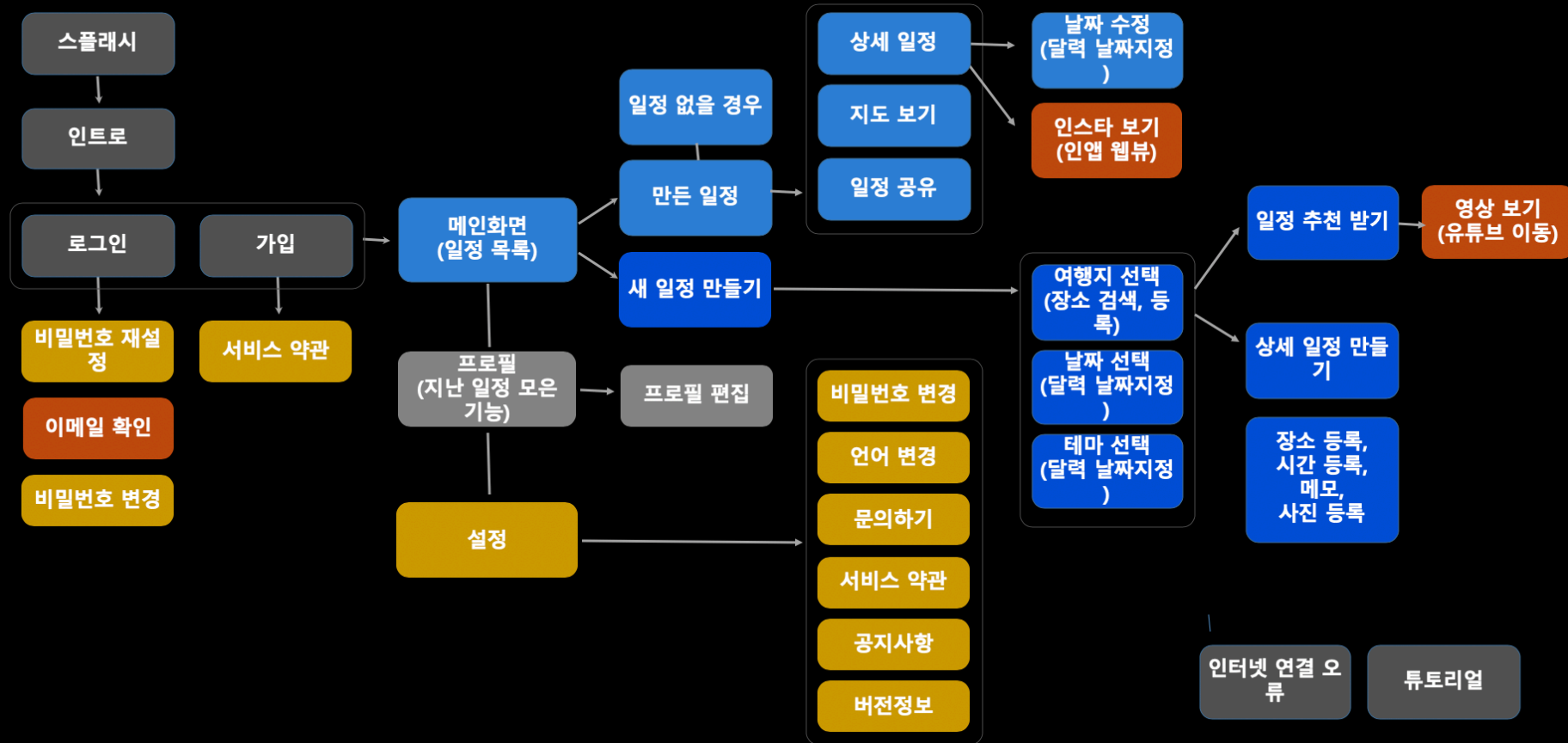
데이터 마이닝 (Airbnb new user bookings: 데이터분석)

● 브랜딩 (Plan A 창업 : 여행일정추천 앱서비스)

오퍼레이션 (D.CAMP : 사업개발팀 인턴)

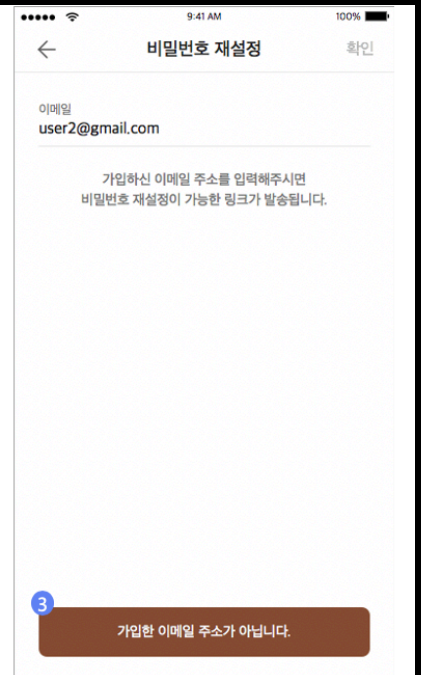
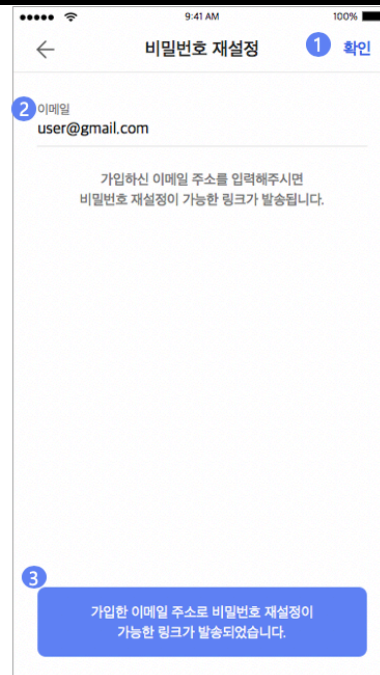
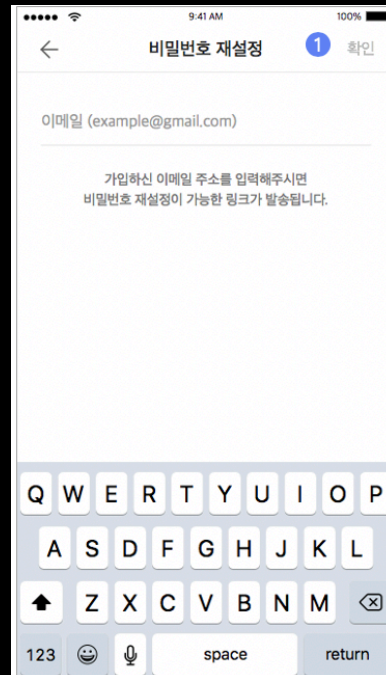
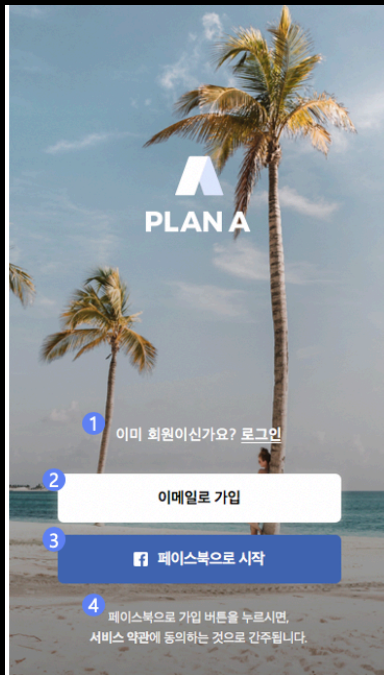
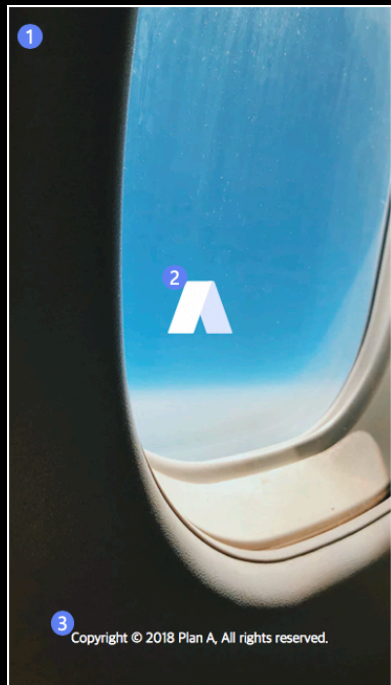
PM (한국공간데이터: 프로젝트매니저 & 서비스기획)

● 브랜딩 (Plan A 창업: 여행일정추천 앱서비스)



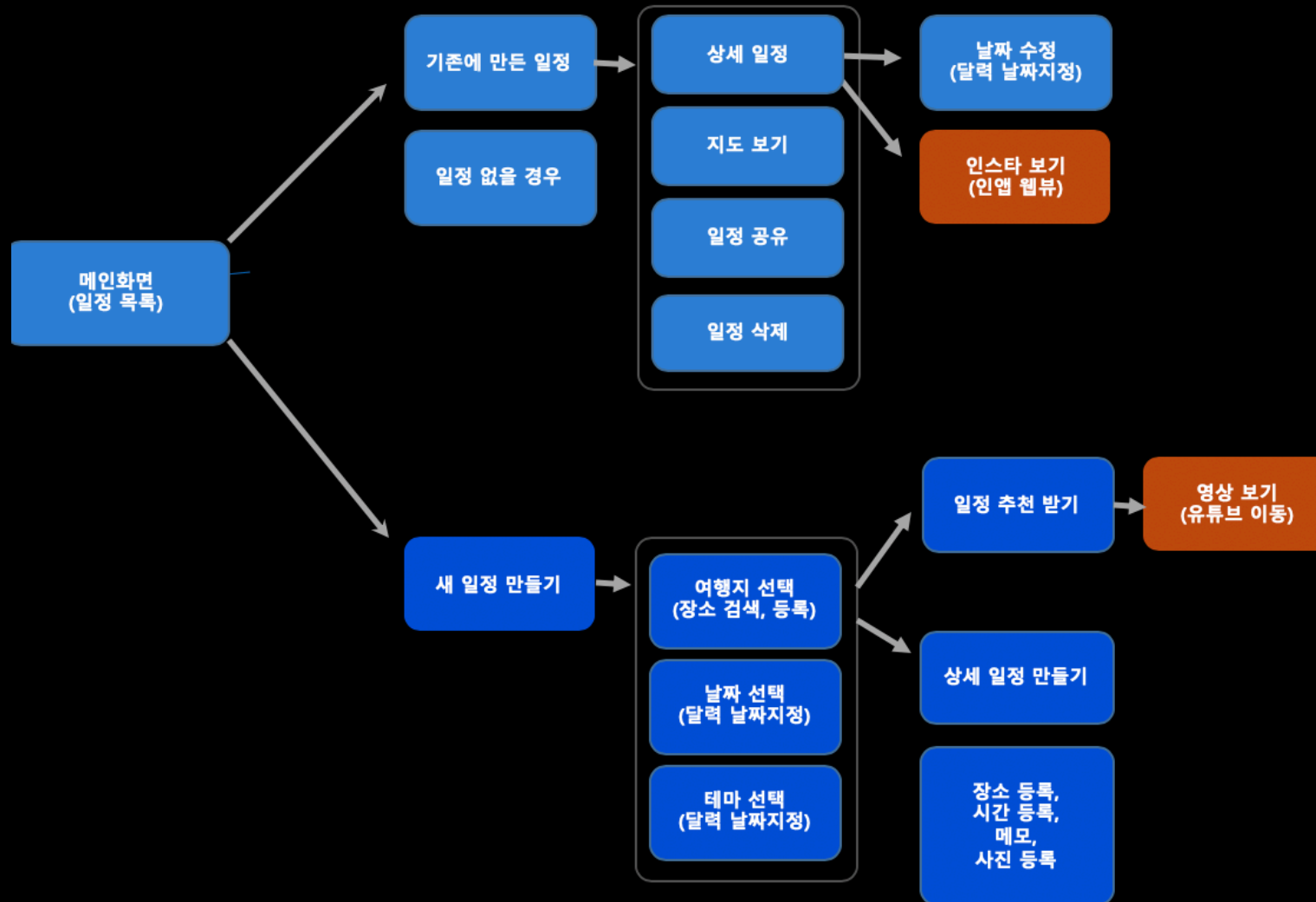
<Plan A | flowchart>

● 브랜딩 (Plan A 창업: 여행일정추천 앱서비스)

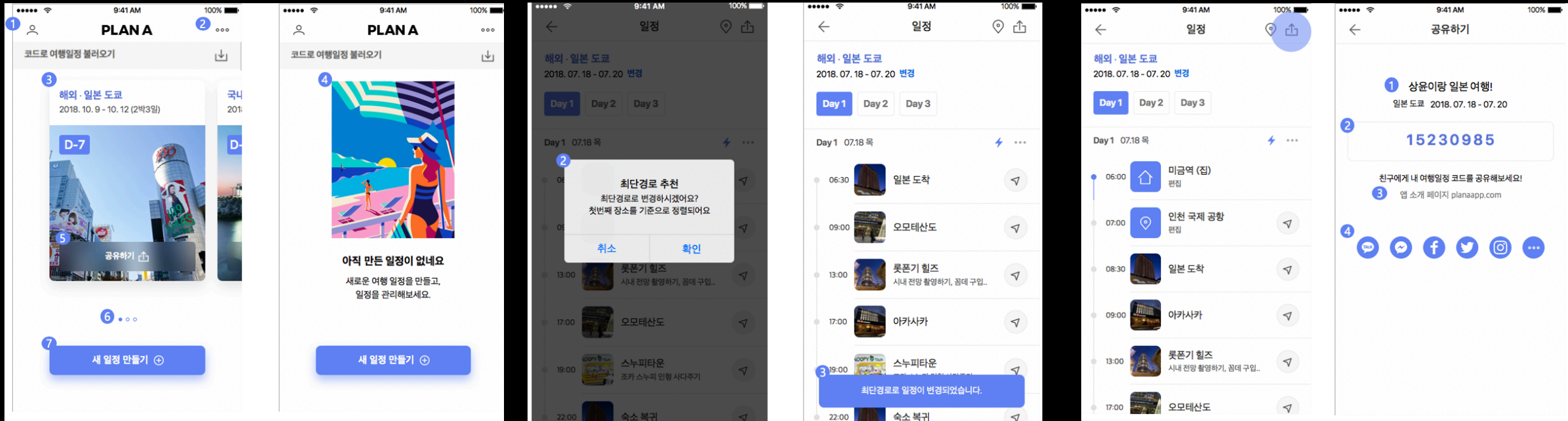


<스플래시 및 인트로 | 로그인>

● 브랜딩 (Plan A 창업: 여행일정추천 앱서비스)

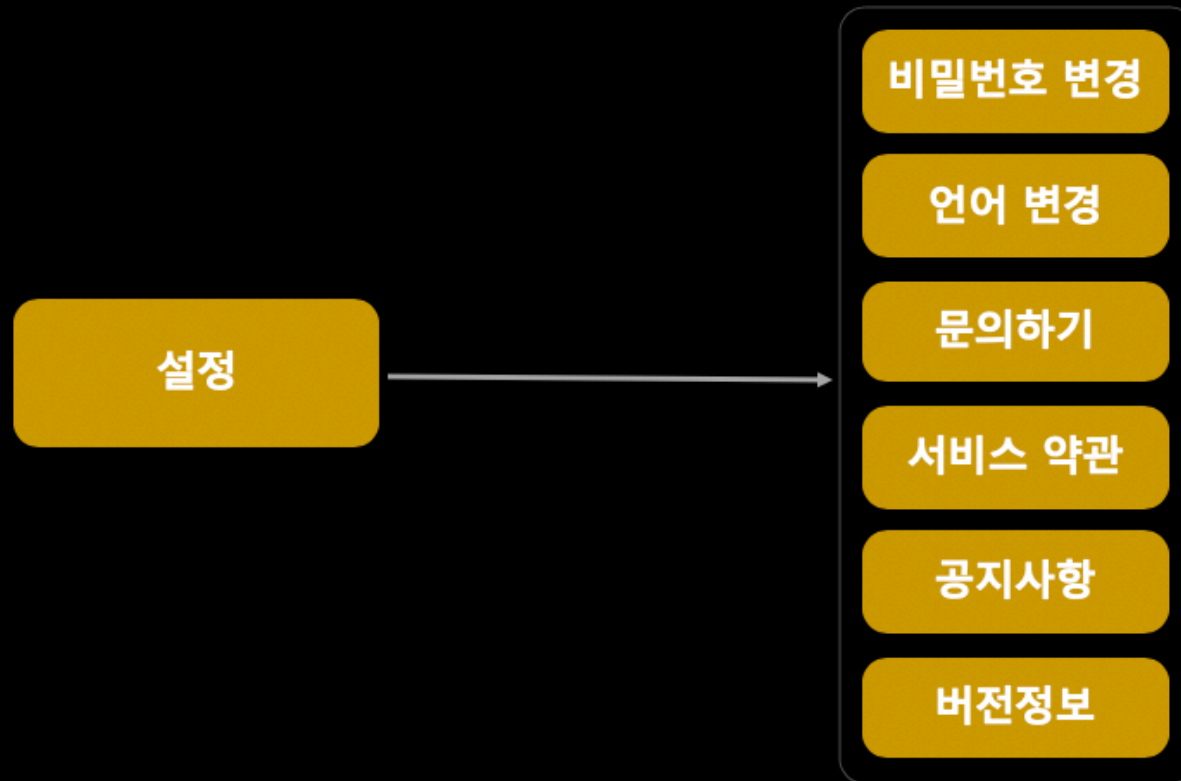


● 브랜딩 (Plan A 창업: 여행일정추천 앱서비스)

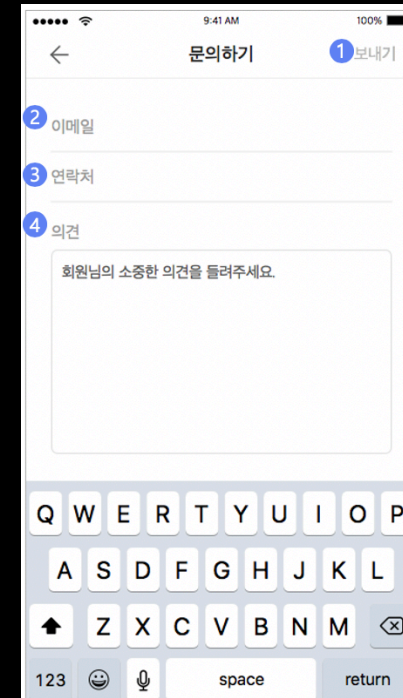
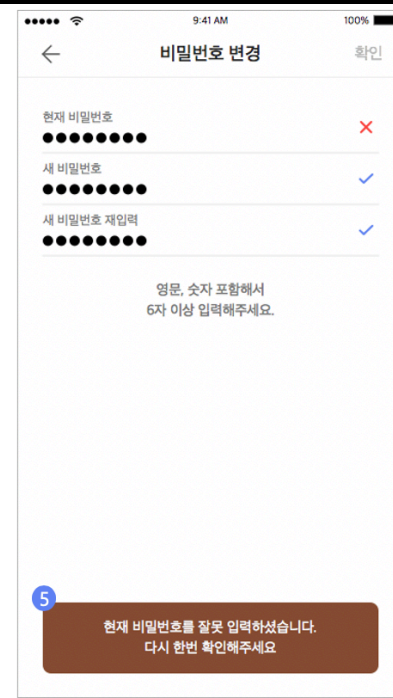
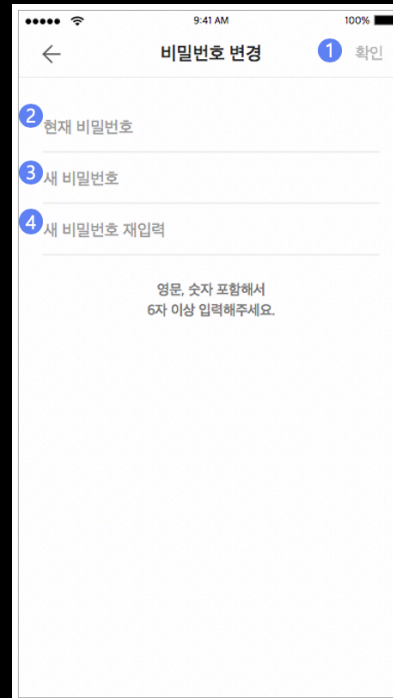
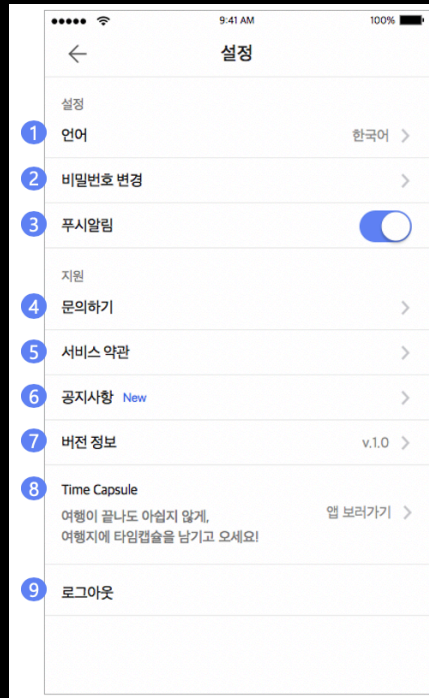


<메인화면 - 새 일정 만들기 | 일정>

● 브랜딩 (Plan A 창업: 여행일정추천 앱서비스)



● 브랜딩 (Plan A 창업: 여행일정추천 앱서비스)



<설정 및 기타 | 설정>

KEYWORDS

데이터 마이닝 (Airbnb new user bookings: 데이터분석)

브랜딩 (Plan A 창업 : 여행일정추천 앱서비스)

● **오퍼레이션** (D.CAMP : 사업개발팀 인턴)

PM (한국공간데이터: 프로젝트매니저 & 서비스기획)

● 오퍼레이션 (D.CAMP : 사업개발팀 인턴)



<디캠프 (은행권청년창업재단), 사업개발팀>

데모데이 오퍼레이션 & 대외협력 사업기획

● 오퍼레이션 (D.CAMP : 사업개발팀 인턴)



<글로벌 방문 프로그램>

Western sydney Univ. 방문프로그램 기획



<IF 2019 패스티벌>

글로벌 IR 프로그램 & 핀테크 프로그램 담당

KEYWORDS

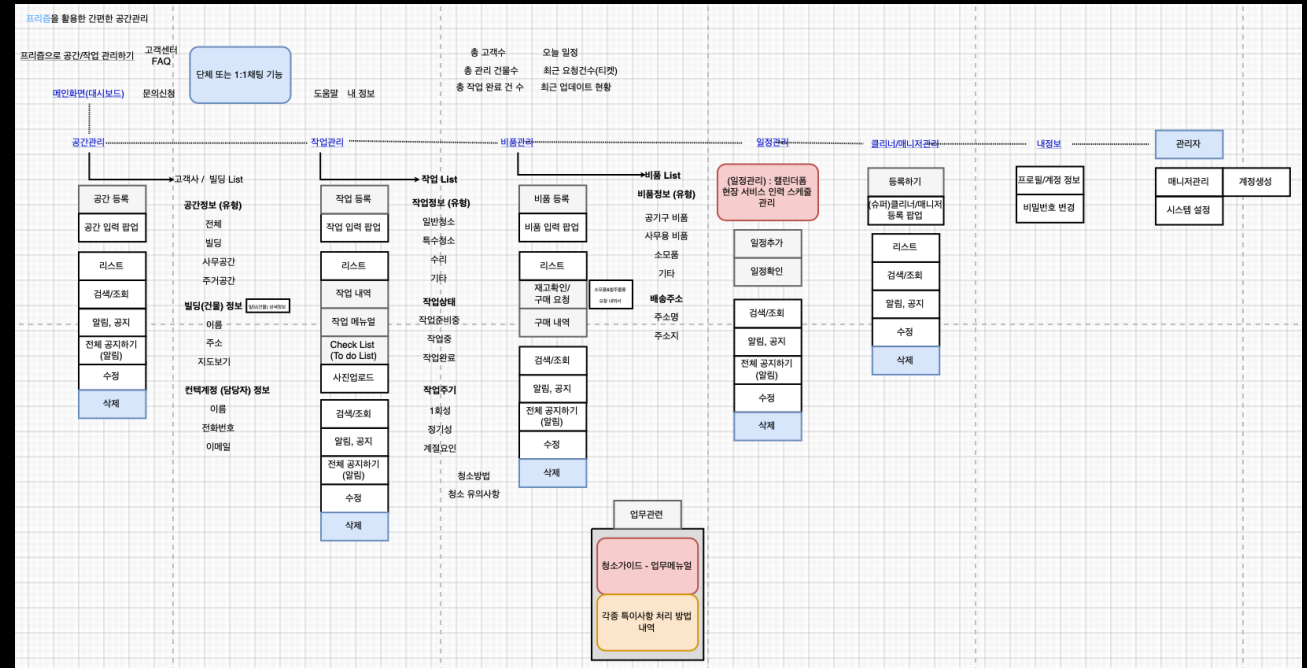
데이터 마이닝 (Airbnb new user bookings: 데이터분석)

브랜딩 (Plan A 창업 : 여행일정추천 앱서비스)

오퍼레이션 (D.CAMP : 사업개발팀 인턴)

● **PM** (한국공간데이터: 프로젝트매니저 & 서비스기획)

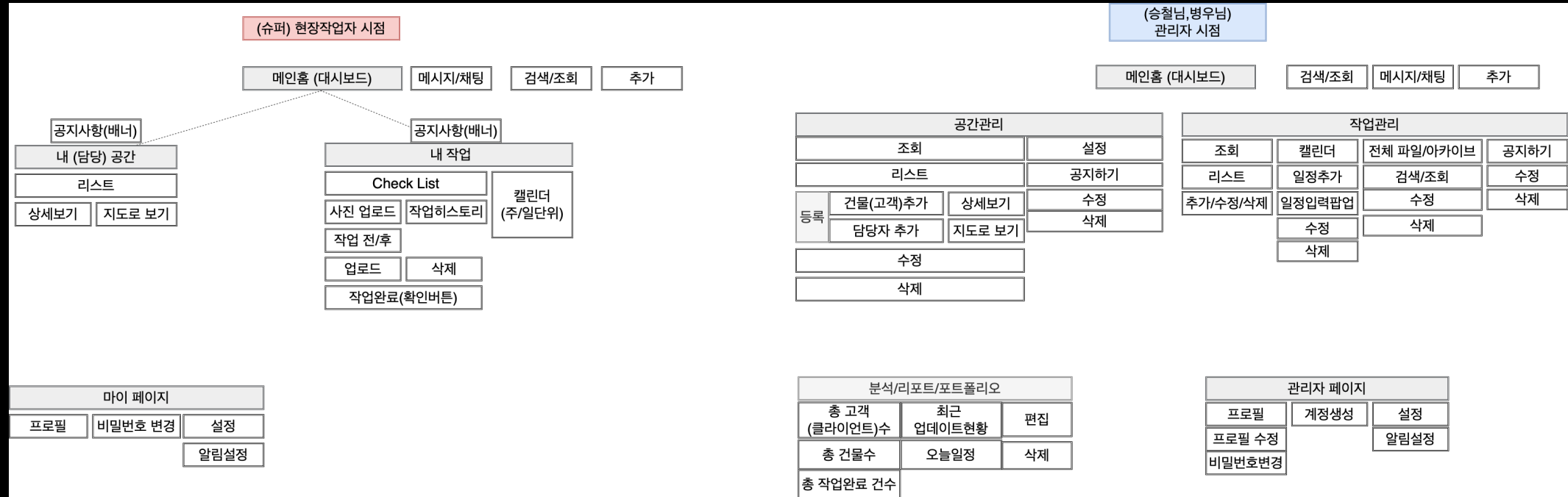
● PM (한국공간데이터: 프로젝트매니저 & 서비스기획)

[illegible][illegible]

<현장&관리자 IA구조>

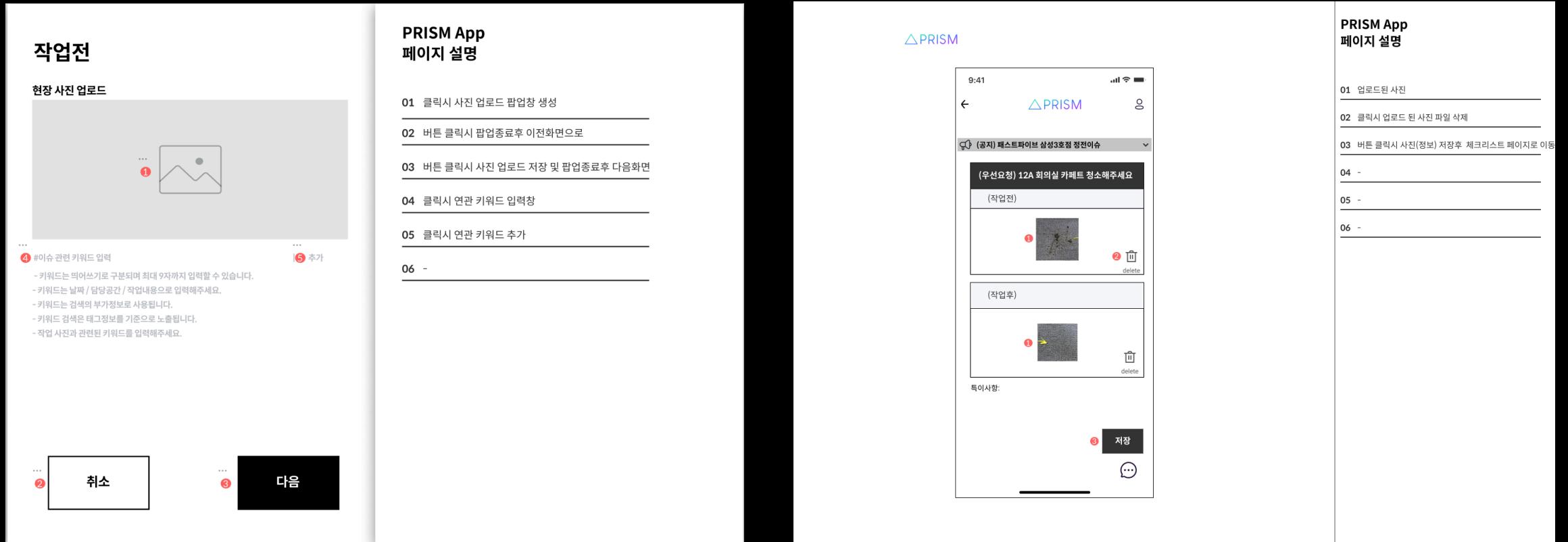
<어드민 페이지 플로우차트>

● PM (한국공간데이터: 프로젝트매니저 & 서비스기획)



<프리즘 메뉴구성>

● PM (한국공간데이터: 프로젝트매니저 & 서비스기획)



<현장 작업자 전용 앱 페이지 스토리보드>

● PM (한국공간데이터: 프로젝트매니저 & 서비스기획)

Prism

클리너 어드민님 안녕하세요!

공간 리스트 (전체보기) 공지사항

1F 정렬 필터

건물명	건물주소	건물정보	(고객)회사명	지점구분	컨택정보	캘린더	아카이브(파일)
☐	하이브로빌딩	서울시 강남구 삼성 1동 157-33	상세보기	페스트파이버	실상 3호점	0	-
☐							
☐							
☐							
☐							
☐							
☐							
☐							
☐							
☐							
☐							
☐							

추가하기

추가하기 버튼 클릭시 새 공간등록 팝업창성

Prism

클리너 어드민님 안녕하세요!

대시보드 (전체보기) 공지사항

작업 수행률(완료) Show: This week -

2021년 02월 26일 (현재)

Sun 13 Mon 14 Tue 15 Wed 16 Thu 17 Fri 18 Sat 19

페스트파이버 12층 배관점검 Reminder
Due date: 2021.02.26
담당자: 서민재 Completed

팁스타운 정기청소 Call
Due date: 2021.02.26
담당자: 세라 Ended

세탁기 누수점검 Event
Due date: 2021.02.26
담당자: 레베카 Completed

Show more

작업요청 건 수 Show: Monthly -

Closed deals

1 Dec 8 Dec 16 Dec 31 Dec

146

작업 수행률 Show: This month -

60%

Active Completed Ended

<관리자용 어드민 웹 페이지>

E.O.D