



BANK TELEMARKETING

권기남, 박수영, 유민형, 조수범

데이터 탐색

모델 적합

고객 특성 파악

결과 종합

CONTENTS

01 데이터 탐색

- I. 변수 개요
- II. 숫자형 변수 시각화
- III. 범주형 변수 시각화
- IV. 분석 계획

age	job	marital	education	default	housing	loan	contact	month	day_of_week
30	blue-collar	married	basic.9y	no	yes	no	cellular	may	fri
39	services	single	high.school	no	no	no	telephone	may	fri
25	services	married	high.school	no	yes	no	telephone	jun	wed
38	services	married	basic.9y	no	unknown	unknown	telephone	jun	fri
47	admin.	married	university.degree	no	yes	no	cellular	nov	mon

Bank-additional.csv

```
> str(a)
'data.frame': 4119 obs. of 21 variables:
 $ age      : int  30 39 25 38 47 32 32 41 31 35 ...
 $ job      : Factor w/ 12 levels "admin.", "blue-collar",...: 2 8 8 8 1 8 1 3 8 2 ...
 $ marital  : Factor w/ 4 levels "divorced", "married",...: 2 3 2 2 2 3 3 2 1 2 ...
 $ education : Factor w/ 8 levels "basic.4y", "basic.6y",...: 3 4 4 3 7 7 7 7 6 3 ...
 $ default  : Factor w/ 3 levels "no", "unknown",...: 1 1 1 1 1 1 1 2 1 2 ...
 $ housing  : Factor w/ 3 levels "no", "unknown",...: 3 1 3 2 3 1 3 3 1 1 ...
 $ loan     : Factor w/ 3 levels "no", "unknown",...: 1 1 1 2 1 1 1 1 1 1 ...
 $ contact  : Factor w/ 2 levels "cellular", "telephone": 1 2 2 2 1 1 1 1 1 2 ...
 $ month    : Factor w/ 10 levels "apr", "aug", "dec",...: 7 7 5 5 8 10 10 8 8 7 ...
 $ day_of_week : Factor w/ 5 levels "fri", "mon", "thu",...: 1 1 5 1 2 3 2 2 4 3 ...
 $ duration  : int  487 346 227 17 58 128 290 44 68 170 ...
 $ campaign  : int  2 4 1 3 1 3 4 2 1 1 ...
 $ pdays    : int  999 999 999 999 999 999 999 999 999 999 ...
 $ previous  : int  0 0 0 0 0 2 0 0 1 0 ...
 $ poutcome  : Factor w/ 3 levels "failure", "nonexistent",...: 2 2 2 2 2 1 2 2 1 2 ...
 $ emp.var.rate : num  -1.8 1.1 1.4 1.4 -0.1 -1.1 -1.1 -0.1 -0.1 1.1 ...
 $ cons.price.idx: num  92.9 94 94.5 94.5 93.2 ...
 $ cons.conf.idx : num  -46.2 -36.4 -41.8 -41.8 -42 -37.5 -37.5 -42 -42 -36.4 ...
 $ euribor3m    : num   1.31 4.86 4.96 4.96 4.19 ...
 $ nr.employed  : num  5099 5191 5228 5228 5196 ...
 $ y            : Factor w/ 2 levels "no", "yes": 1 1 1 1 1 1 1 1 1 1 ...
```

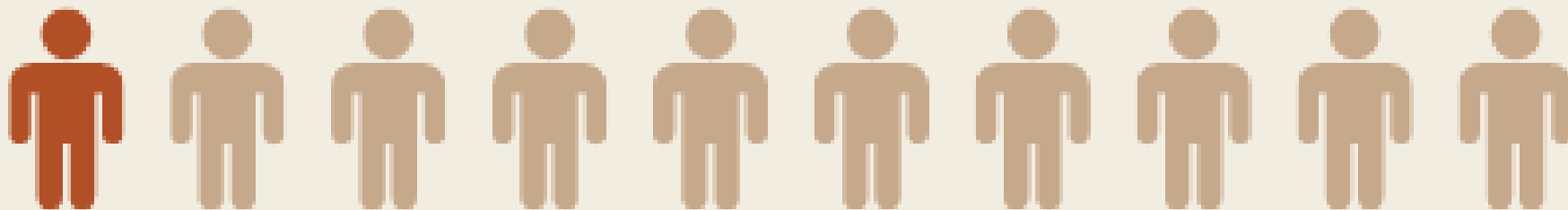
< Input 변수 >

- 은행 클라이언트 관련 변수
- 현 캠페인의 (마지막) 컨택 관련 변수
- 이전 캠페인 관련 변수
- 사회적/경제적 지표

< Output 변수(desired target) >

- 클라이언트의 정기 예금 상품 가입 여부

Yes!



● yes ● no



은행 client
변수

1. Age: 나이(numeric)
2. Job: job 타입 (12 categories)
3. Marital: 기혼 여부(4 categories)
4. Education: (8 cate- ; Portugal System)
5. Default: 파산 여부(3 cate- ; y, n, unknown)
6. Housing: 주택 대출 여부(3 cate- ; y, n, u)
7. Loan: 개인 대출 여부(3 cate- ; y, n, u)

8. Contact: 컨택 방법(2 categories ; cellular, telephone)
9. Month: 마지막으로 컨택한 월(10 cate-; 3-12月)
10. Day of week: 마지막 컨택한 요일(5 cate-; 월-금)
11. Duration: 마지막 컨택 시 통화한 시간(초)(numeric)
12. Campaign: 이번 캠페인 동안 이 클라이언트에게 컨택한 총 횟수(numeric)

현 캠페인의
마지막 컨택
변수이전 캠페인
변수

13. Pdays: 이전 캠페인 이래로 마지막으로 컨택한 후 지난 날짜 수(numeric)
14. Previous: 이전 캠페인에서 컨택한 횟수(numeric)
15. Poutcome: 이전 마케팅 캠페인의 결과(3 categories ; failure, success, nonexistent)

16. Emp.var.rate : Employment Variation Rate(고용변화율) (num; 분기 지표)
17. Cons.price.idx : 소비자 물가 지수(CPI) (num ; 월별 지표)
18. Cons.conf.idx : 소비자 신뢰 지수(CCI) (num; 월별 지표)
19. Euribor3m : Euro Interbank Offered Rate(Euribor) 3 months(num; 일별)
20. Nr.employed : 고용자 수(num; 분기별 지표)

사회/경제적
변수

은행 client
변수

1. Age: 나이(numeric)
2. Job: job 타입 (12 categories)
3. Marital: 기혼 여부(4 categories)

타겟 변수에 매우 영향력이 있다고 한다. 컨택되기 전에는 이 변수 자체를 알 수가 없지만 컨택의 끝에는 타겟변수를 분명히 알 수 있다. 따라서, 이 설명변수는 오직 벤치마켓의 목적으로만 포함되고 그렇지 않으면(현실 예측 모델 등) 제거되어야 한다.

현재의 지역경제 상황, 고용상태, 6개월 후의 지역경제, 고용 및 가계 수입에 대한 전망을 조사한 지표

8. Contact: 컨택 방법(2 categories ; cellular, telephone)
9. Month: 마지막으로 컨택한 월(10 categories ; 3-12월)
10. Day of week: 마지막 컨택한 요일
11. Duration: 마지막 컨택 시 통화한 시간
12. Campaign: 이번 캠페인 동안 이 클라이언트가 캠페인에 참여했는지 여부

Employment Variation : 경제적 상황의 변화로 인해 얼마나 많은 사람들이 고용되거나 해고되는지의 변동성

이전 캠페인
변수

13. Pdays: 이전 캠페인 이래로 마지막으로 컨택한 후 지난 날짜 수(numeric)
14. Previous: 이전 캠페인에서 컨택한 횟수(numeric)
15. Poutcome: 이전 마케팅 캠페인의 결과(3 categories ; failure, success, nonexistent)

16. Emp.var.rate : Employment Variation Rate
17. Cons.price.idx : 소비자 물가 지수(Consumer Price Index)
18. Cons.conf.idx : 소비자 신뢰 지수(Consumer Confidence Index)
19. Euribor3m : Euro Interbank Offered Rate 3 months
20. Nr.employed : 고용자 수(numeric; 분기별)

EURO InterBank Offered Rate(Euribor) 3 months ; daily Euribor rates는 Large 유럽 은행들 간 자금을 빌려주는 평균 이자율로, 유럽 경제 시장을 잘 보여줌(e.g. interest rate swaps, interest rate futures, saving accounts and mortgages).

현재 캠페인의
마지막 컨택
변수

```
> table(a$pdays)
```

```

 0    1    2    3    4    5    6    7    9
 2    3    4   52   14    4   42   10    3
10   11   12   13   14   15   16   17   18
 8    1    5    2    1    2    2    1    2
19   21  999
 1    1 3959

```

999 means
client was not
previously
contacted.

이전 캠페인 변수

- 1. Age: 나이(numeric)
- 2. Job: job 타입 (12 categories)
- 3. Marital: 기혼 여부(4 categories)

- 4. Education: (8 cate- ; Portugal System)
- 5. **Default**: 파산 여부(3 cate- ; y, n, unknown)
- 6. Housing: 주택 대출 여부(3 cate- ; y, n, u)
- 7. Loan: 개인 대출 여부(3 cate- ; y, n, u)

```
> table(a$default)
```

```

      no unknown      yes
    3315      803         1

```

- 8. Contact: 컨택 방법(2 categories ; cellular, telephone)
- 9. Month: 마지막으로 컨택한 월(10 cate-; 3-12月)
- 10. Day of week: 마지막 컨택한 요일(5 cate-; 월-금)
- 11. Duration: 마지막 컨택 시 통화한 시간(초)(numeric)
- 12. Campaign: 이번 캠페인 동안 이 클라이언트에게 컨택한 총 횟수(numeric)
- 13. Pdays: 이전 캠페인 이래로 마지막으로 컨택한 후 지난 날짜 수(numeric)
- 14. Previous: 이전 캠페인에서 컨택한 횟수(numeric)
- 15. Poutcome: 이전 마케팅 캠페인의 결과(3 categories ; failure, success, nonexistent)
- 16. Emp.var.rate : Employment Variation Rate(고용변화율) (num; 분기별 지표)
- 17. Cons.price.idx : 소비자 물가 지수(CPI) (num ; 월별 지표)
- 18. Cons.conf.idx : 소비자 신뢰 지수(CCI) (num; 월별 지표)
- 19. Euribor3m : Euro Interbank Offered Rate(Euribor) 3 months(num; 일별)
- 20. Nr.employed : 고용자 수(num; 분기)

사회/경제적 변수

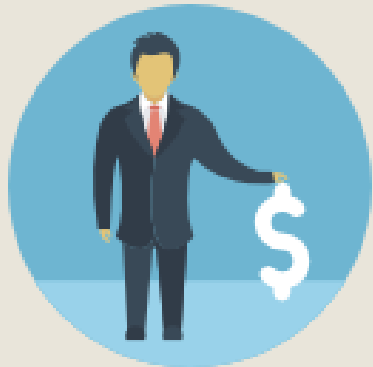


BANCO DE PORTUGAL
EUROSISTEMA



INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

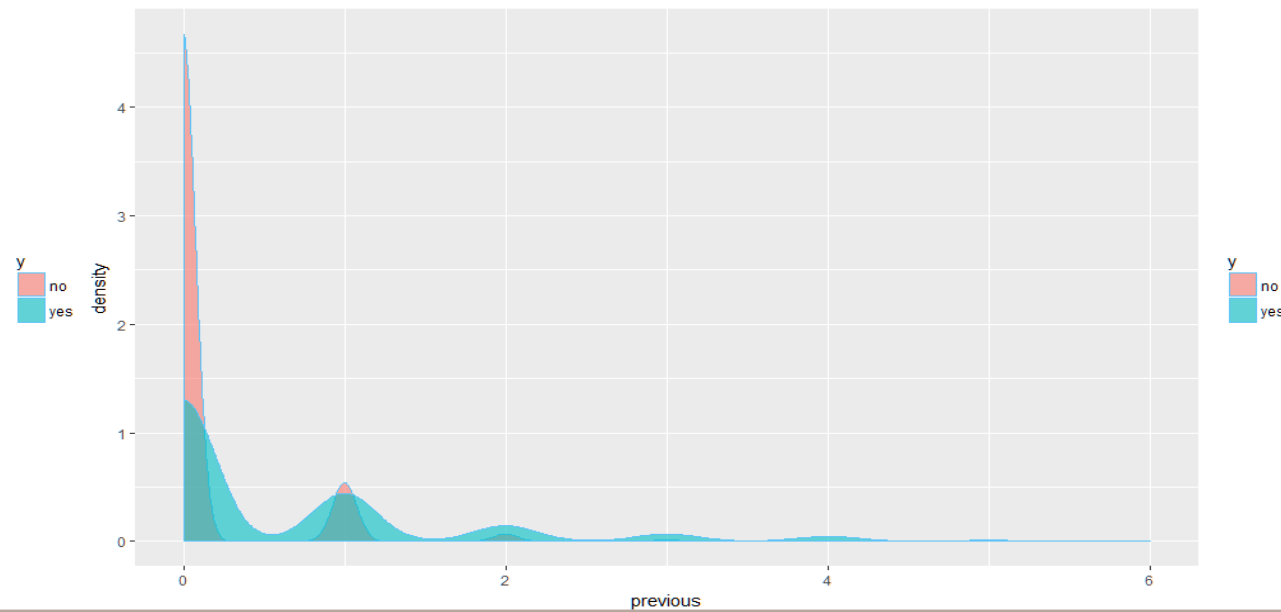
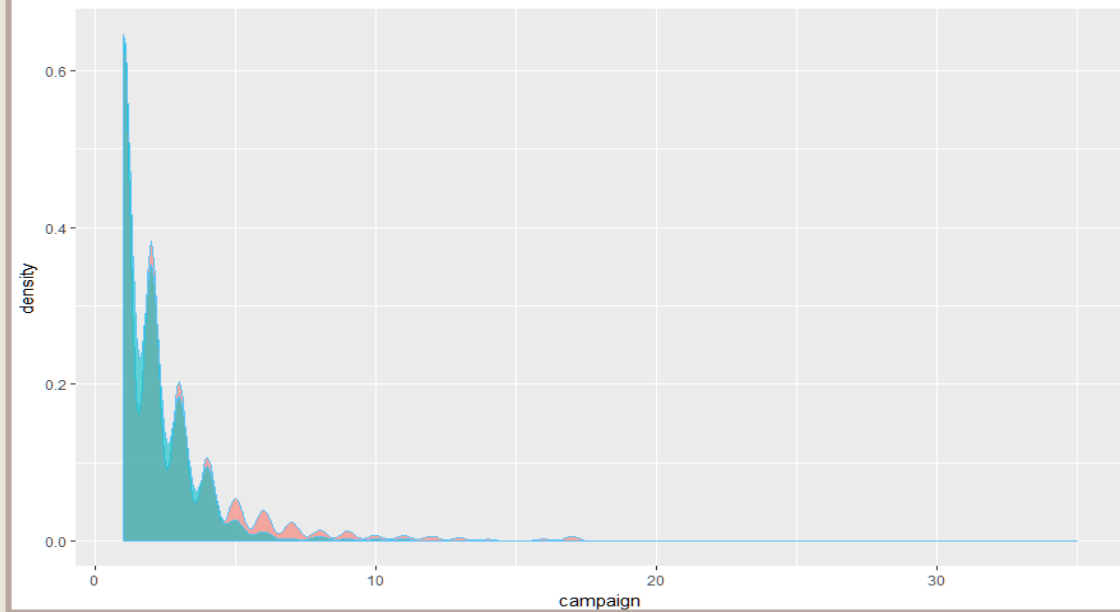
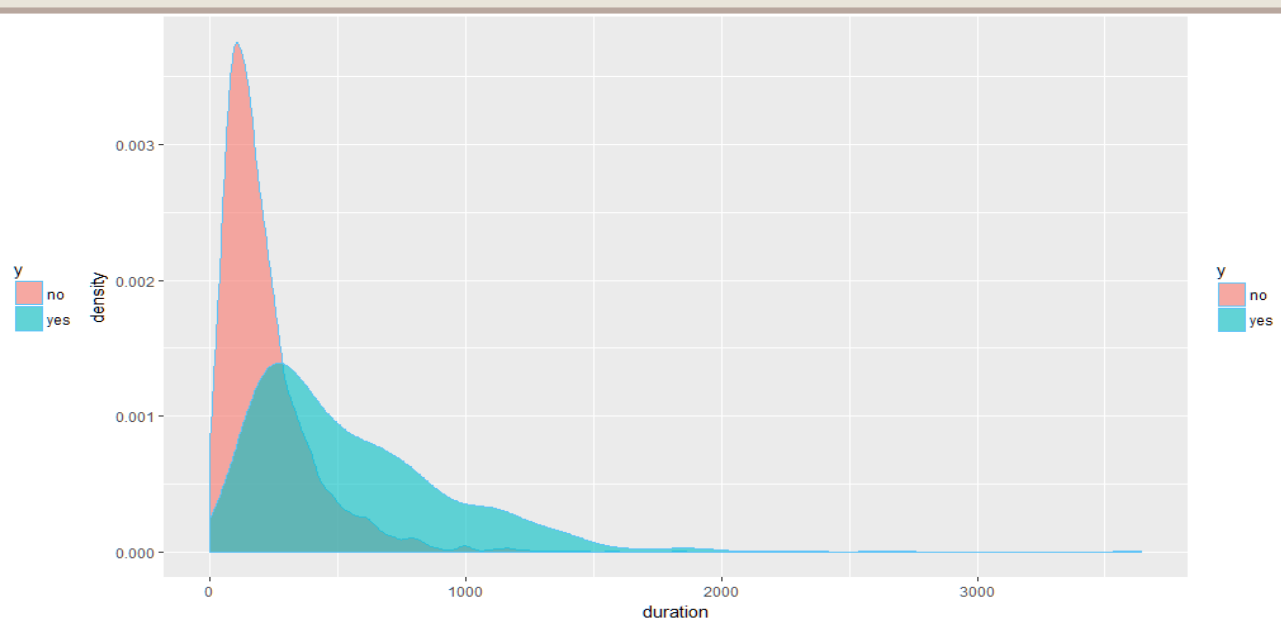
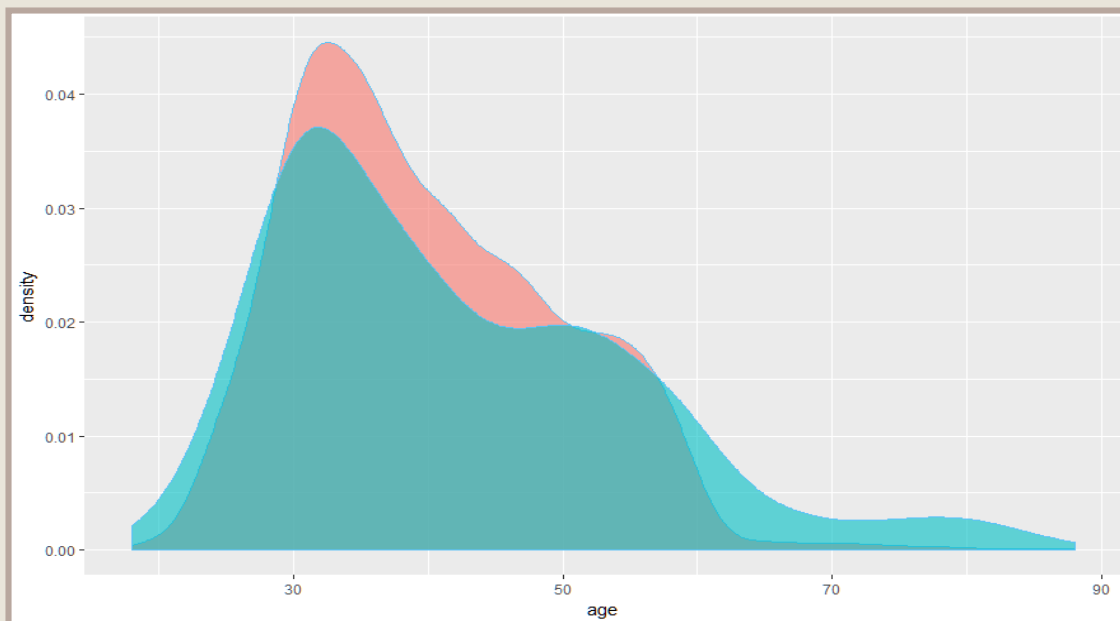
Employment	Statistical data	SDMX data	Source: INE E-mail contact: Mrs. Sónia Torres
Unemployment	Statistical data	SDMX data	Source: INE E-mail contact: Mrs. Sónia Torres
Wages/Earnings	Statistical data	SDMX data	Source: INE E-mail contact: Mrs. Sónia Torres
Consumer Price Index	Statistical data	SDMX data	Source: INE E-mail contact: Ms. Cristina Fernandes

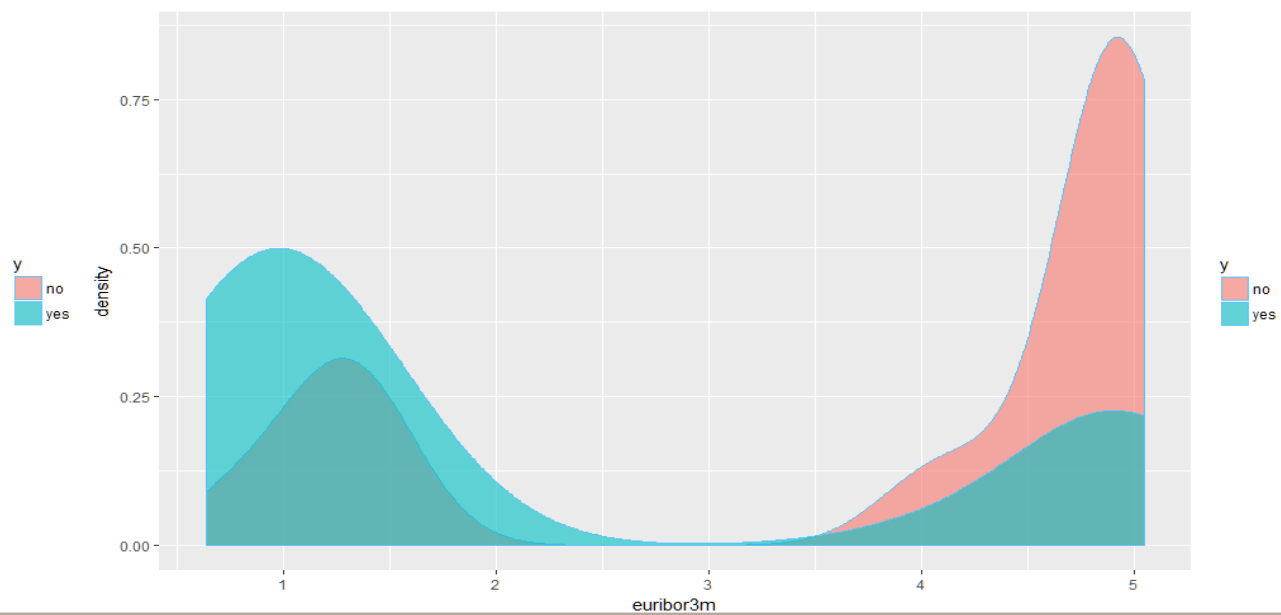
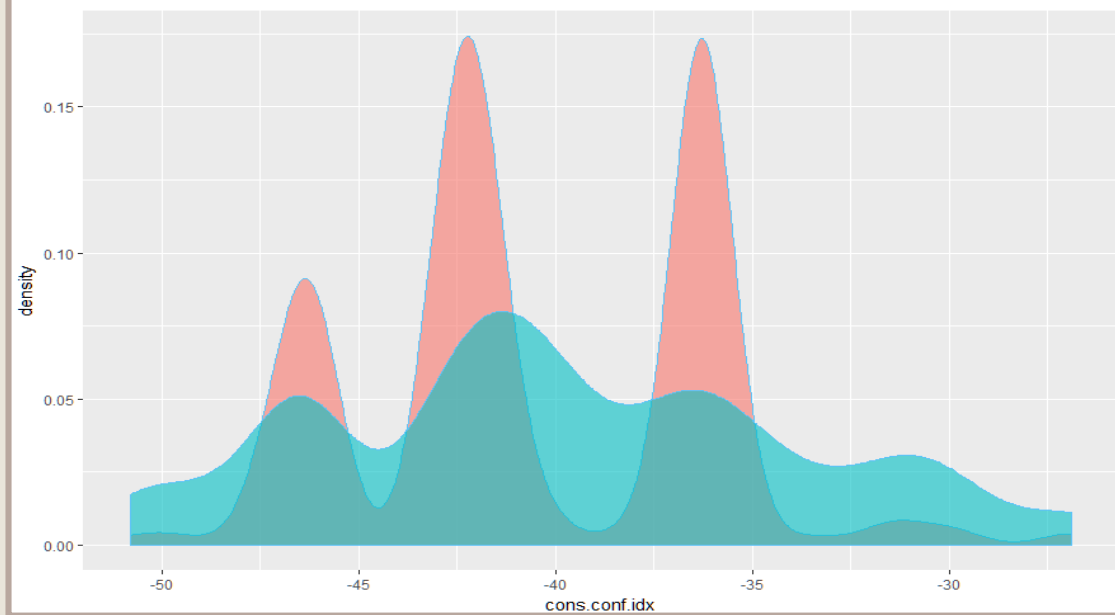
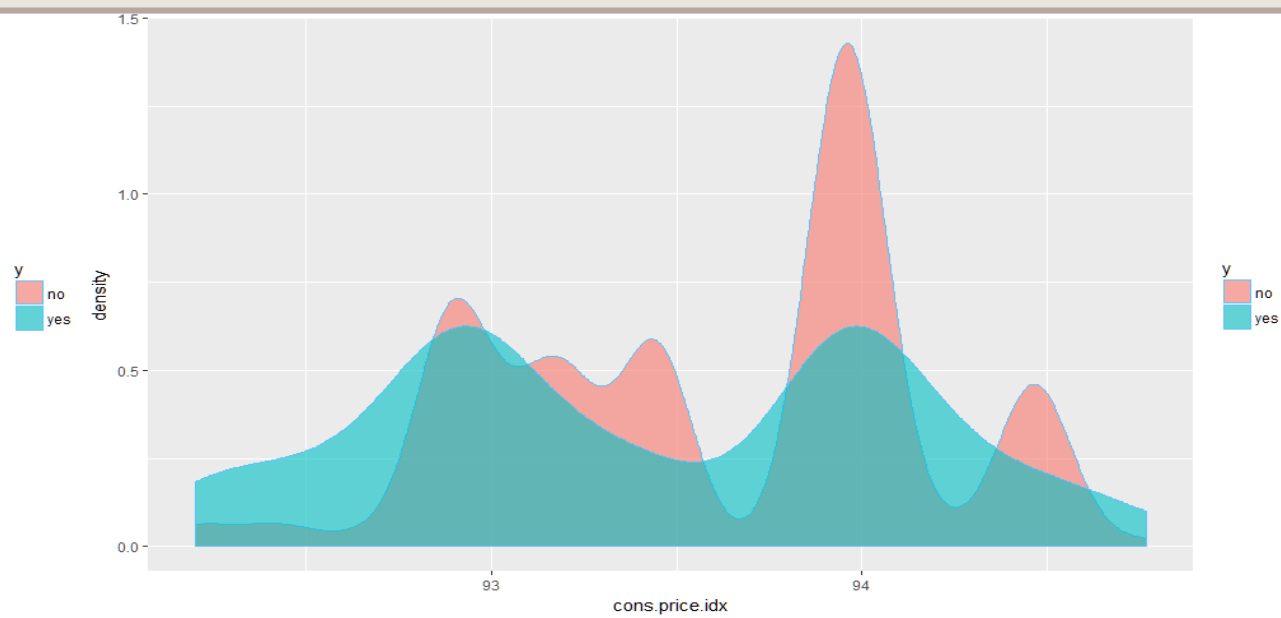
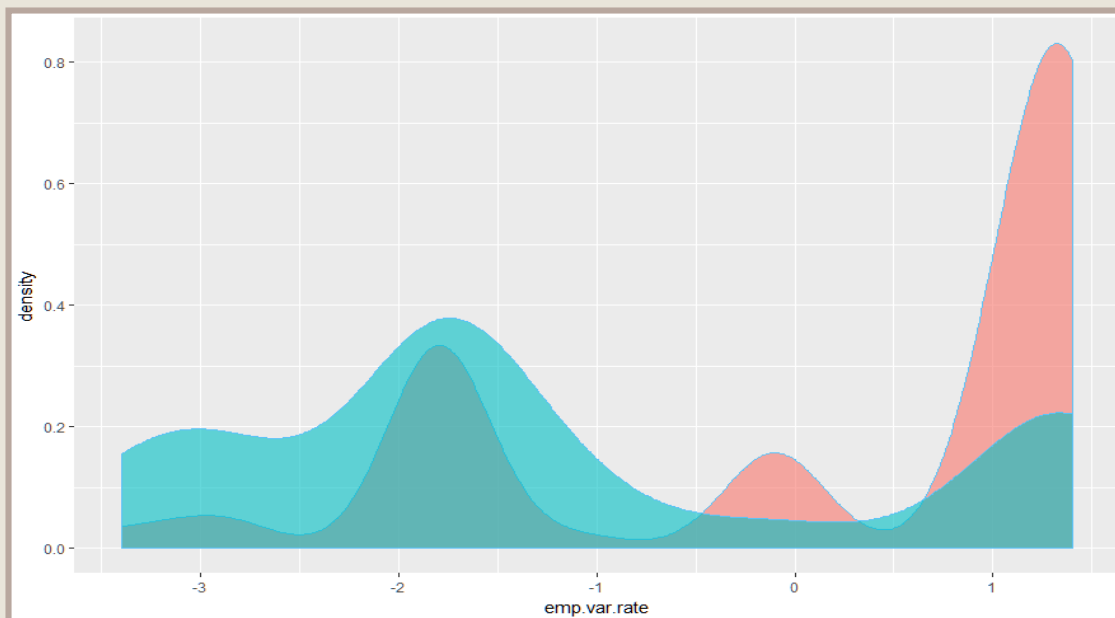


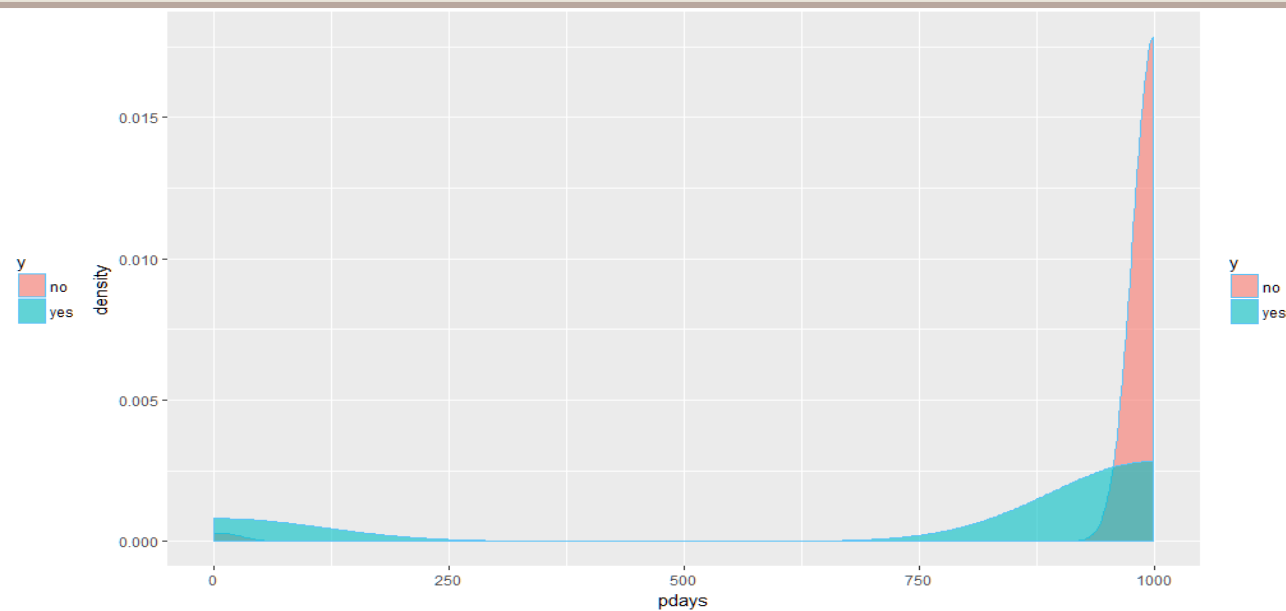
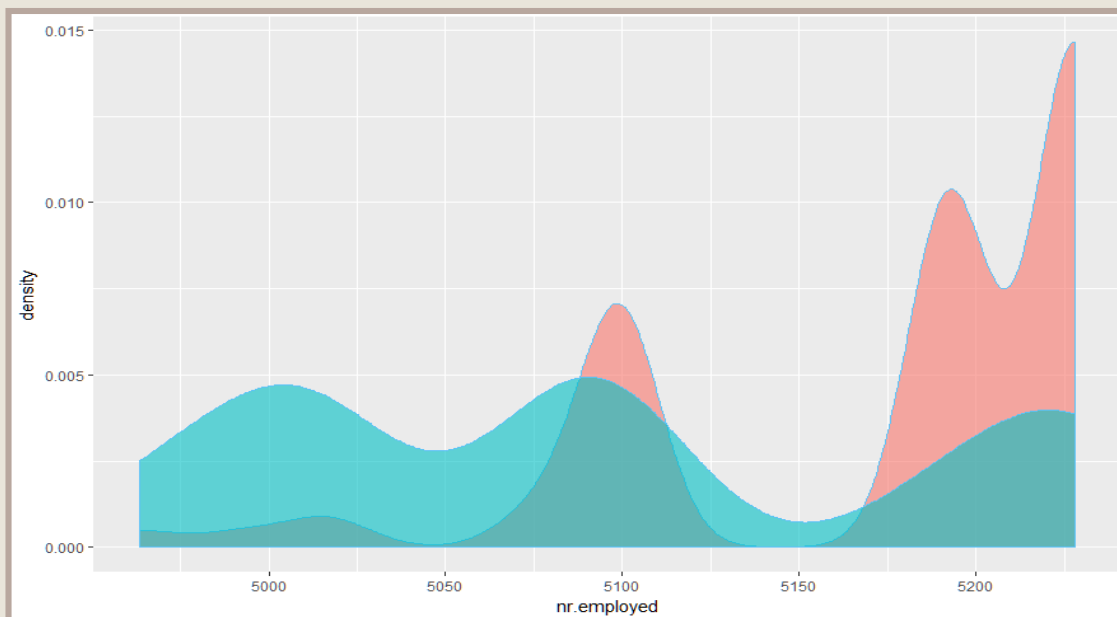
Consumer price index (Year-on-year growth rate - Base 2012 - %) by Geographic localization and Special aggregates; Monthly - Statistics Portugal, Consumer price index

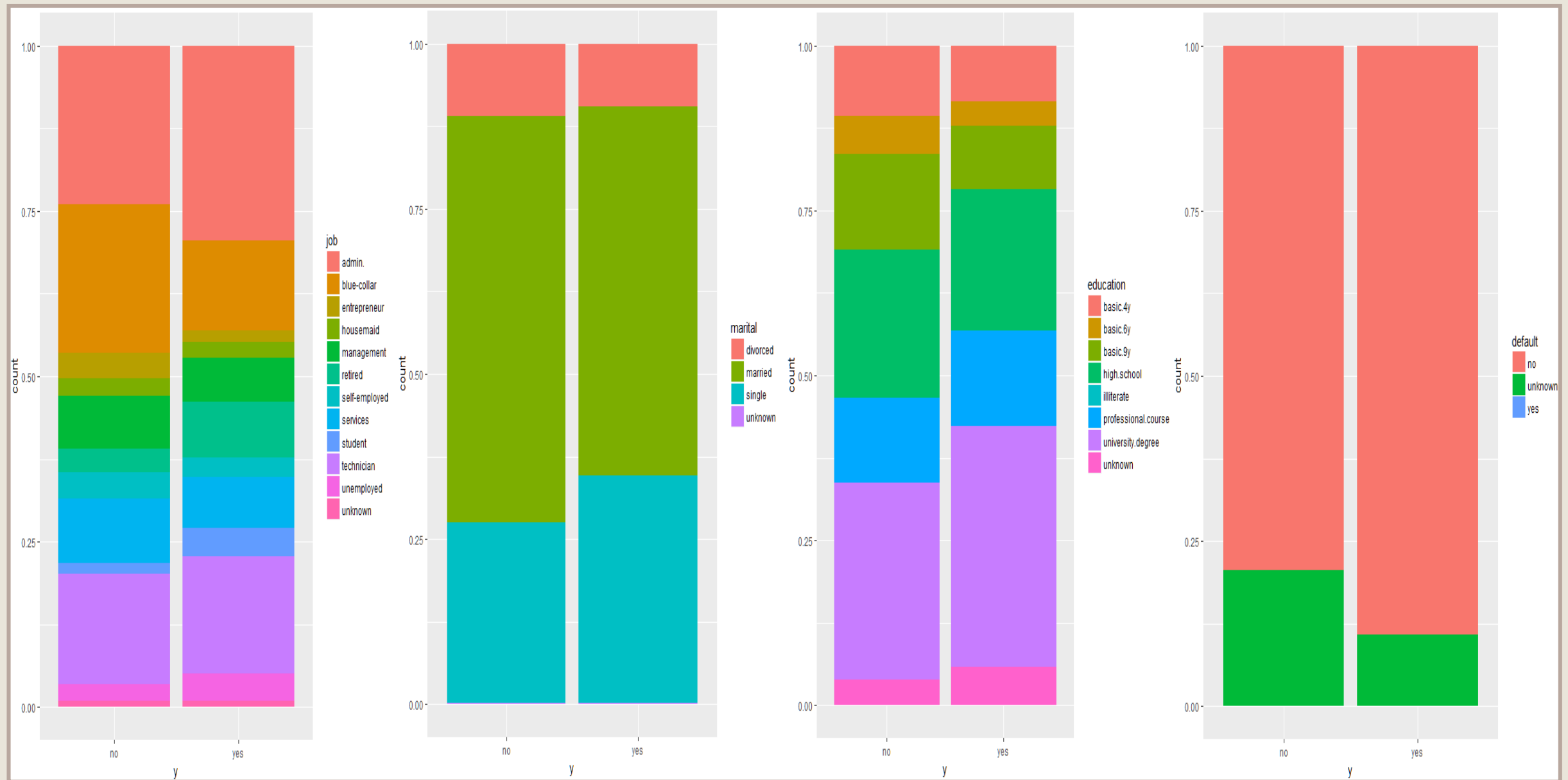
Employed population (Series 2011 - No.) by Place of residence (NUTS - 2013), Sex and Occupation; Quarterly - Statistics Portugal, Labour force survey
Note(s):

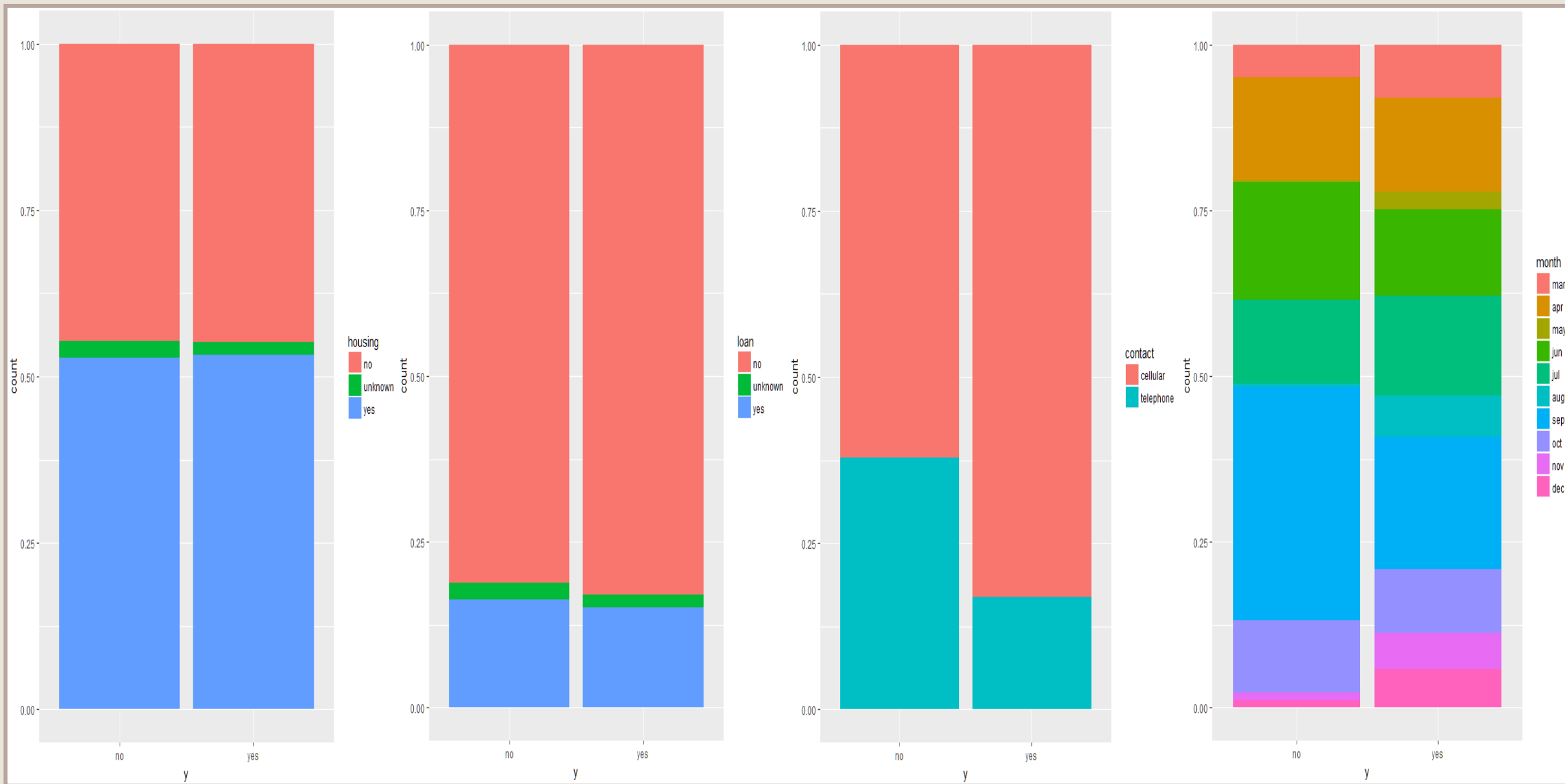
(1) The values were calibrated using the population estimates calculated from the final results of Census 2011.

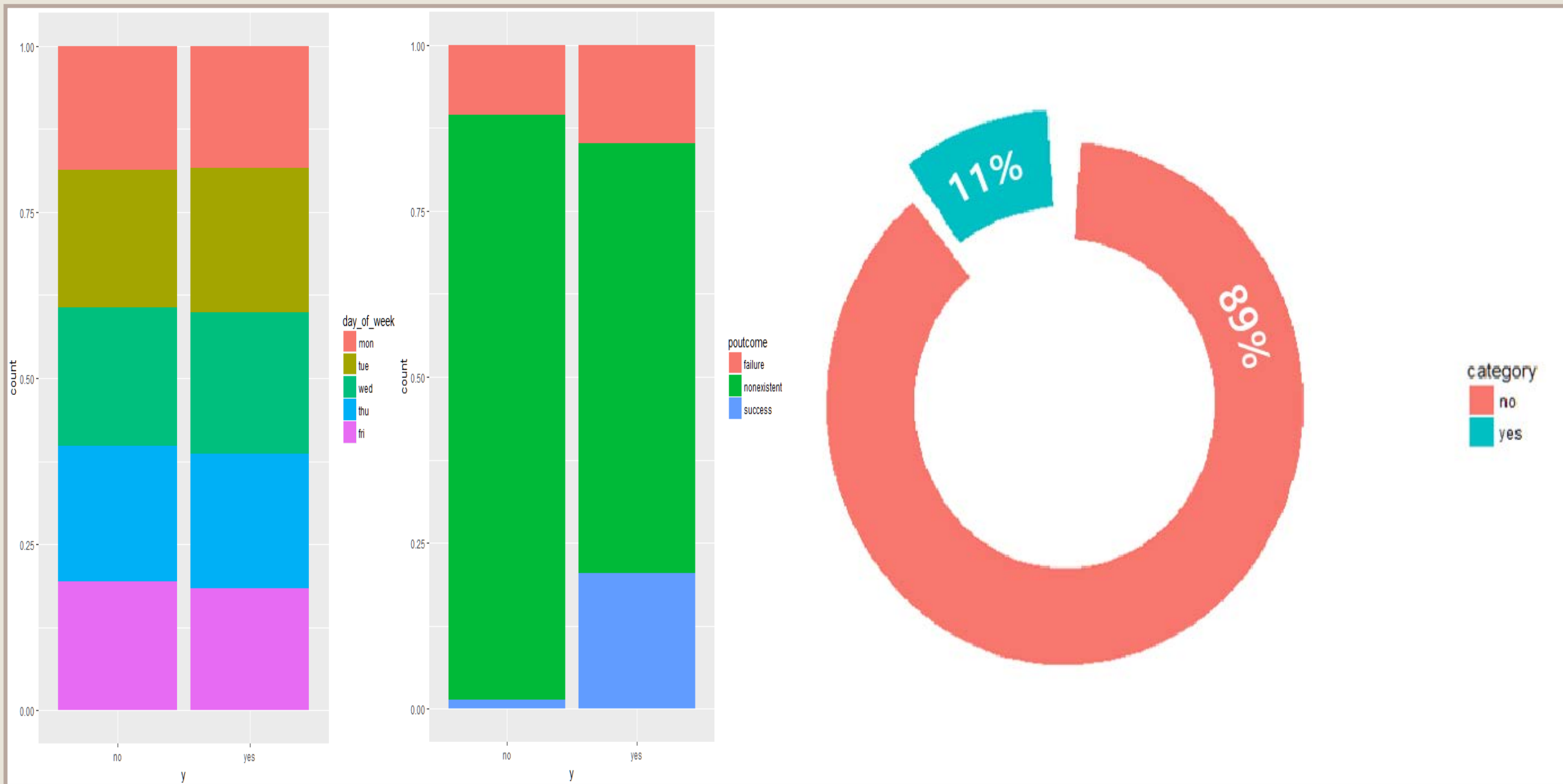














The classification goal is

to predict if the client will subscribe (yes/no)

+ **Subscribe Yes** 고객의 특징?



~~신경망~~~~SVM~~

~~블랙박스 모형 제외~~

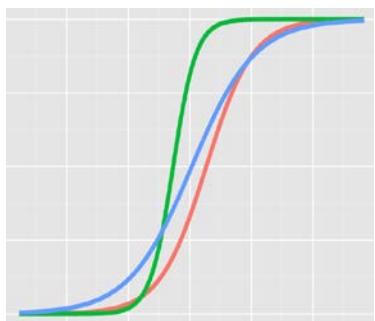


- 분류 규칙 학습

rules를 통해 중요한 변수를
한눈에 알 수 있음

- 로지스틱 회귀분석

회귀 계수를 통해
변수의 영향력 파악



- 연관규칙 분석

알아낸 규칙을 통해
고객의 특성 파악

02 모델 적합

- I. 분류 규칙 학습
- II. 로지스틱 회귀 분석
- III. 연관 규칙 분석

```
> model_tree
```

```
c5.0
```

```
4119 samples
 20 predictor
 2 classes: 'no', 'yes'
```

```
No pre-processing
```

```
Resampling: Cross-Validated (10 fold)
```

```
Summary of sample sizes: 3707, 3707, 3707
```

```
Resampling results:
```

```
Accuracy   Kappa
0.9113817  0.4399253
```

```
Confusion Matrix and Statistics
```

```
      Reference
Prediction no  yes
no      3612  260
yes      56   191
```

```
Accuracy : 0.9233
95% CI : (0.9147, 0.9312)
```

```
No Information Rate : 0.8905
P-Value [Acc > NIR] : 9.42e-13
```

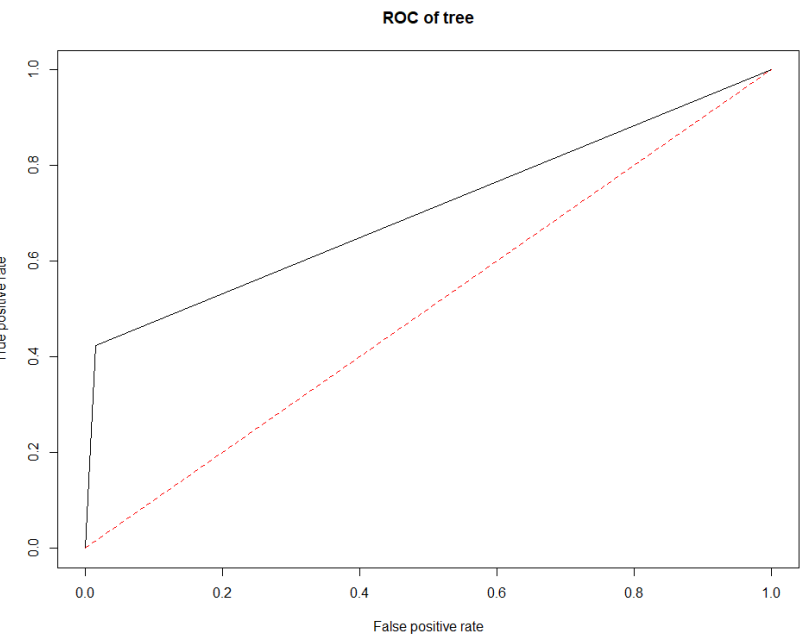
```
Kappa : 0.5092
```

```
McNemar's Test P-Value : < 2.2e-16
```

```
Sensitivity : 0.42350
Specificity : 0.98473
Pos Pred Value : 0.77328
Neg Pred Value : 0.93285
Prevalence : 0.10949
Detection Rate : 0.04637
Detection Prevalence : 0.05997
Balanced Accuracy : 0.70412
```

```
'Positive' Class : yes
```

Actual	Predicted		Row Total
	no	yes	
no	3612 0.985 0.933	56 0.015 0.227	3668 0.891
yes	260 0.576 0.067	191 0.424 0.773	451 0.109
Column Total	3872 0.940	247 0.060	4119



Accuracy
0.9233

AUC
0.7041

Kappa
0.5092

Fscore
0.5472

```
> model_glm
```

```
Generalized Linear Model
```

```
4119 samples
```

```
20 predictor
```

```
2 classes: 'no', 'yes'
```

```
No pre-processing
```

```
Resampling: Cross-Validated (10 fold)
```

```
Summary of sample sizes: 3707, 3708, 3706
```

```
Resampling results:
```

```
Accuracy Kappa
```

```
0.9121104 0.4607297
```

```
Confusion Matrix and Statistics
```

```

      Reference
Prediction no yes
no      3577 252
yes      91 199

```

```
Accuracy : 0.9167
95% CI : (0.9079, 0.925)
```

```
No Information Rate : 0.8905
```

```
P-Value [Acc > NIR] : 1.299e-08
```

```
Kappa : 0.4937
```

```
Mcnemar's Test P-Value : < 2.2e-16
```

```
Sensitivity : 0.44124
```

```
Specificity : 0.97519
```

```
Pos Pred Value : 0.68621
```

```
Neg Pred Value : 0.93419
```

```
Prevalence : 0.10949
```

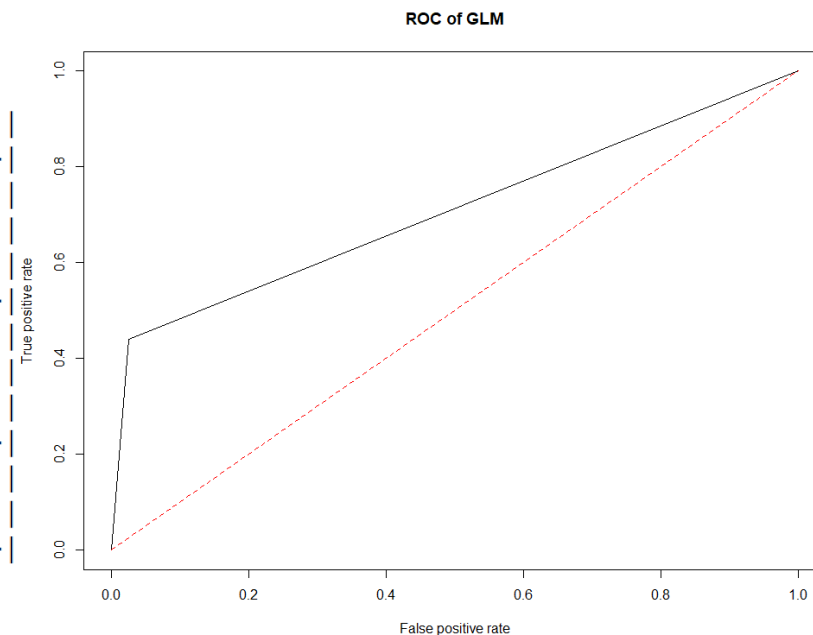
```
Detection Rate : 0.04831
```

```
Detection Prevalence : 0.07041
```

```
Balanced Accuracy : 0.70822
```

```
'Positive' Class : yes
```

Actual	Predicted		Row Total
	no	yes	
no	3577 0.975 0.934	91 0.025 0.314	3668 0.891
yes	252 0.559 0.066	199 0.441 0.686	451 0.109
Column Total	3829 0.930	290 0.070	4119



Accuracy
0.9167

AUC
0.7082

Kappa
0.4937

Fscore
0.5371

Data preprocessing - dummy coding

```
> str(a)
'data.frame': 4119 obs. of 21 variables:
 $ age      : int  30 39 25 38 47 32 32 41 31 35 ...
 $ job      : Factor w/ 12 levels "admin.", "blue-collar",...: 2 8 8 8 1 8 1 3 8 2 ...
 $ marital  : Factor w/ 4 levels "divorced", "married",...: 2 3 2 2 2 3 3 2 1 2 ...
 $ education : Factor w/ 8 levels "basic.4y", "basic.6y",...: 3 4 4 3 7 7 7 7 6 3 ...
 $ default  : Factor w/ 3 levels "no", "unknown",...: 1 1 1 1 1 1 1 2 1 2 ...
 $ housing  : Factor w/ 3 levels "no", "unknown",...: 3 1 3 2 3 1 3 3 1 1 ...
 $ loan     : Factor w/ 3 levels "no", "unknown",...: 1 1 1 2 1 1 1 1 1 1 ...
 $ contact  : Factor w/ 2 levels "cellular", "telephone": 1 2 2 2 1 1 1 1 1 2 ...
 $ month    : Factor w/ 10 levels "apr", "aug", "dec",...: 7 7 5 5 8 10 10 8 8 7 ...
 $ day_of_week : Factor w/ 5 levels "fri", "mon", "thu",...: 1 1 5 1 2 3 2 2 4 3 ...
 $ duration : int  487 346 227 17 58 128 290 44 68 170
 $ campaign : int  2 4 1 3 1 3 4 2
 $ pdays   : int  999 999 999 999
 $ previous : int  0 0 0 0 0 2 0 0
 $ poutcome : Factor w/ 3 levels "f
 $ emp.var.rate : num -1.8 1.1 1.4 1.4
 $ cons.price.idx: num  92.9 94 94.5 94.
 $ cons.conf.idx : num -46.2 -36.4 -41.
 $ euribor3m     : num  1.31 4.86 4.96 4
 $ nr.employed   : num  5099 5191 5228 5
 $ y             : Factor w/ 2 levels "n
```

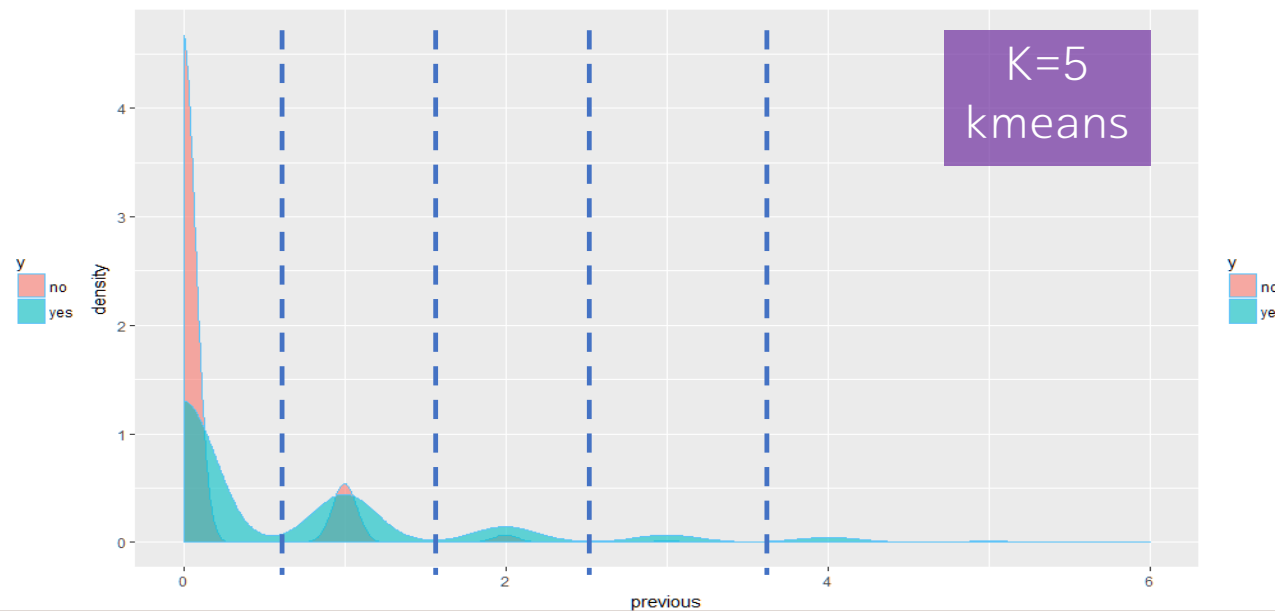
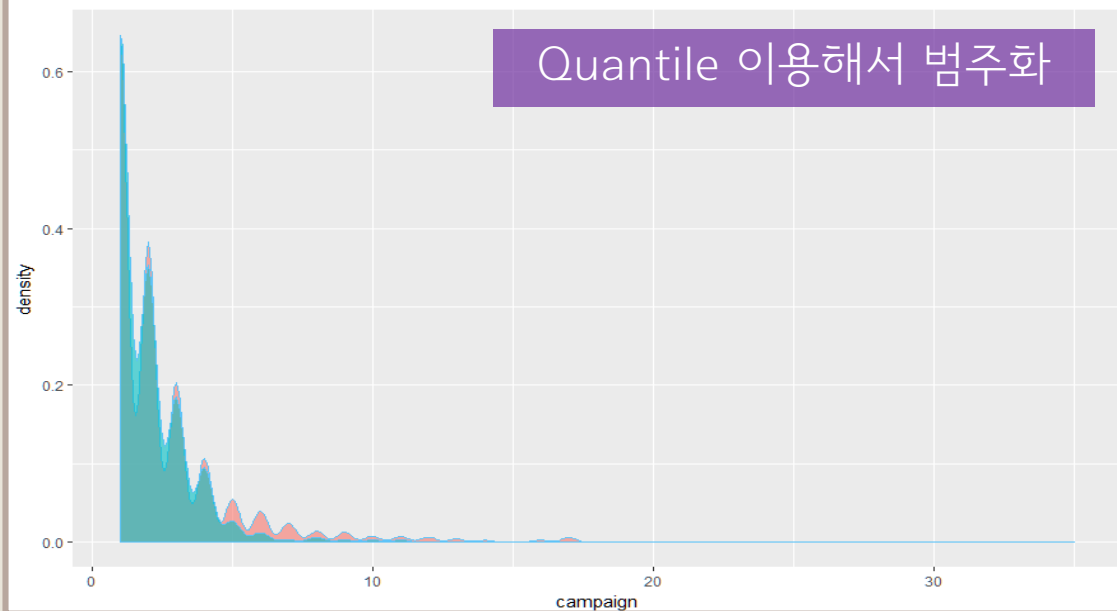
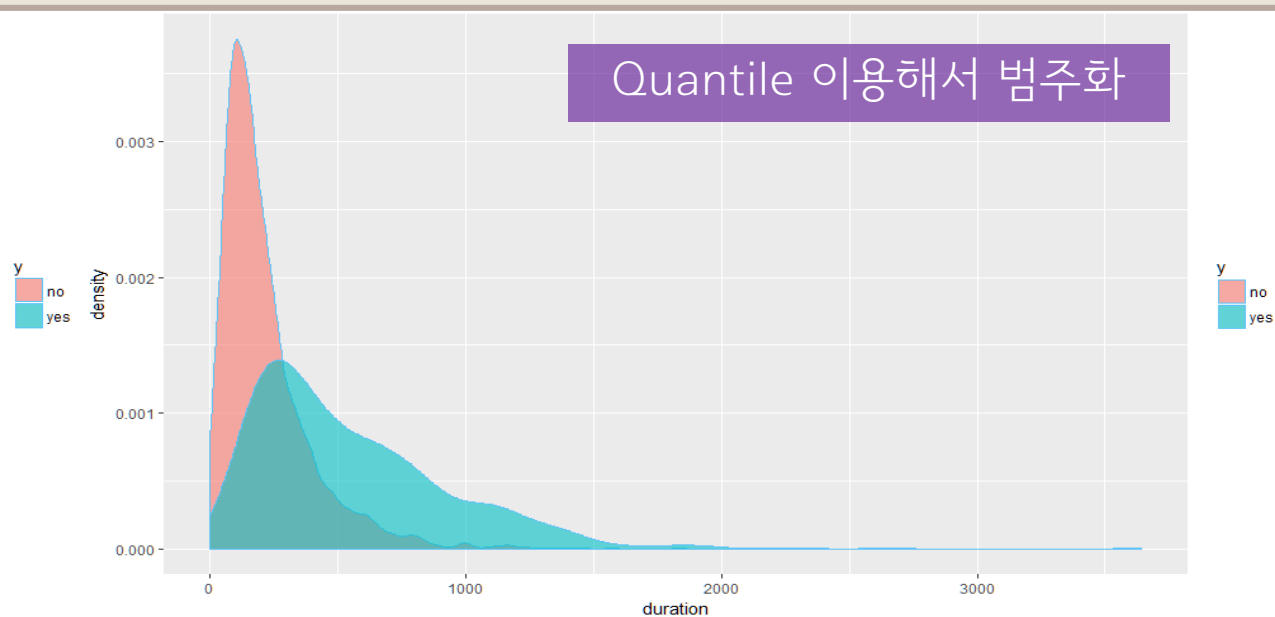
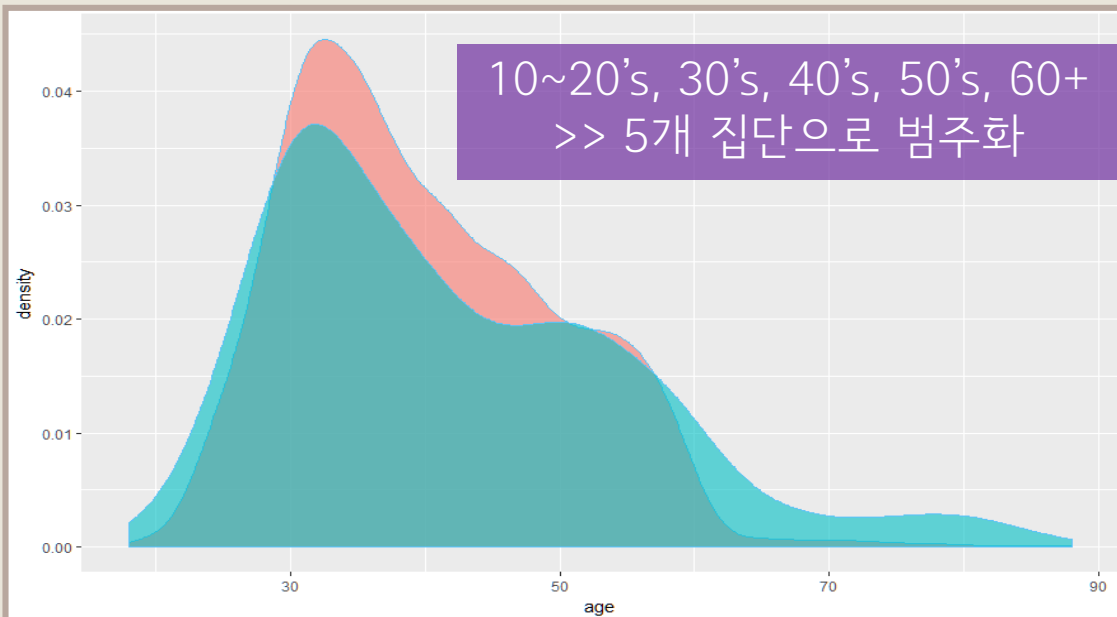
숫자형 변수를 범주화

```
> summary(bank[,c(1,11,12:14,16:20)])
```

age	duration	campaign	pdays	previous
Min. :18.00	Min. : 0.0	Min. : 1.000	Min. : 0.0	Min. :0.0000
1st Qu.:32.00	1st Qu.: 103.0	1st Qu.: 1.000	1st Qu.:999.0	1st Qu.:0.0000
Median :38.00	Median : 181.0	Median : 2.000	Median :999.0	Median :0.0000
Mean :40.11	Mean : 256.8	Mean : 2.537	Mean :960.4	Mean :0.1903
3rd Qu.:47.00	3rd Qu.: 317.0	3rd Qu.: 3.000	3rd Qu.:999.0	3rd Qu.:0.0000
Max. :88.00	Max. :3643.0	Max. :35.000	Max. :999.0	Max. :6.0000

emp.var.rate	cons.price.idx	cons.conf.idx	euribor3m	nr.employed
Min. : -3.40000	Min. :92.20	Min. : -50.8	Min. :0.635	Min. :4964
1st Qu.: -1.80000	1st Qu.:93.08	1st Qu.: -42.7	1st Qu.:1.334	1st Qu.:5099
Median : 1.10000	Median :93.75	Median : -41.8	Median :4.857	Median :5191
Mean : 0.08497	Mean :93.58	Mean : -40.5	Mean :3.621	Mean :5166
3rd Qu.: 1.40000	3rd Qu.:93.99	3rd Qu.: -36.4	3rd Qu.:4.961	3rd Qu.:5228
Max. : 1.40000	Max. :94.77	Max. : -26.9	Max. :5.045	Max. :5228

어떤 기준으로 범주화 할것인가?

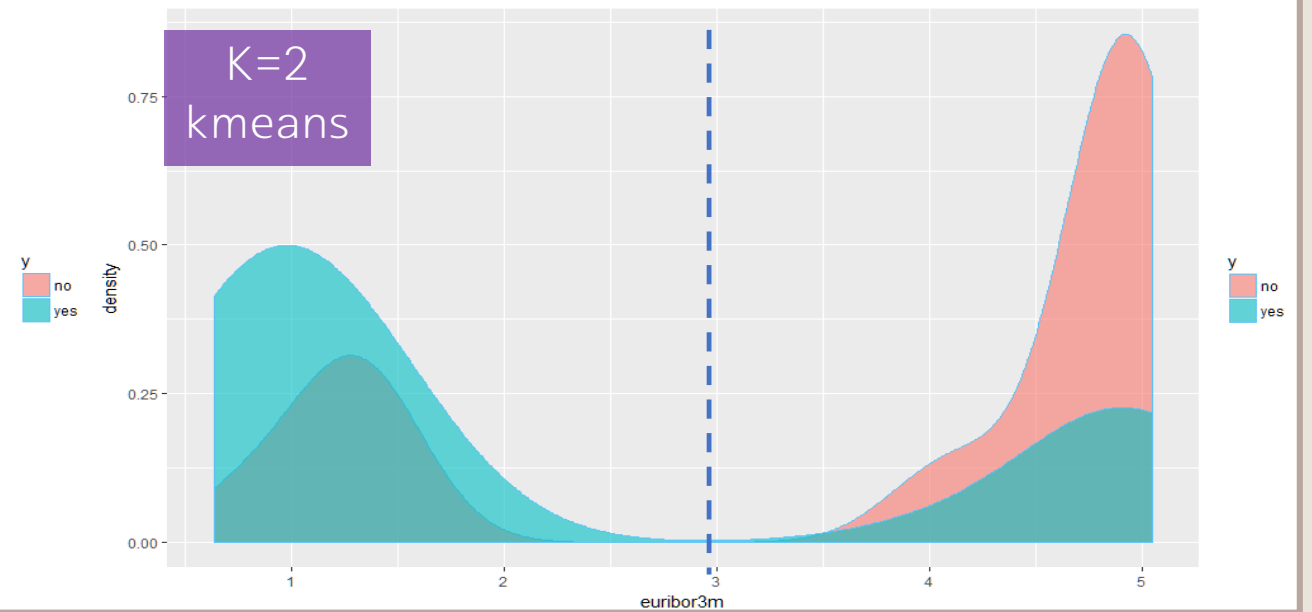
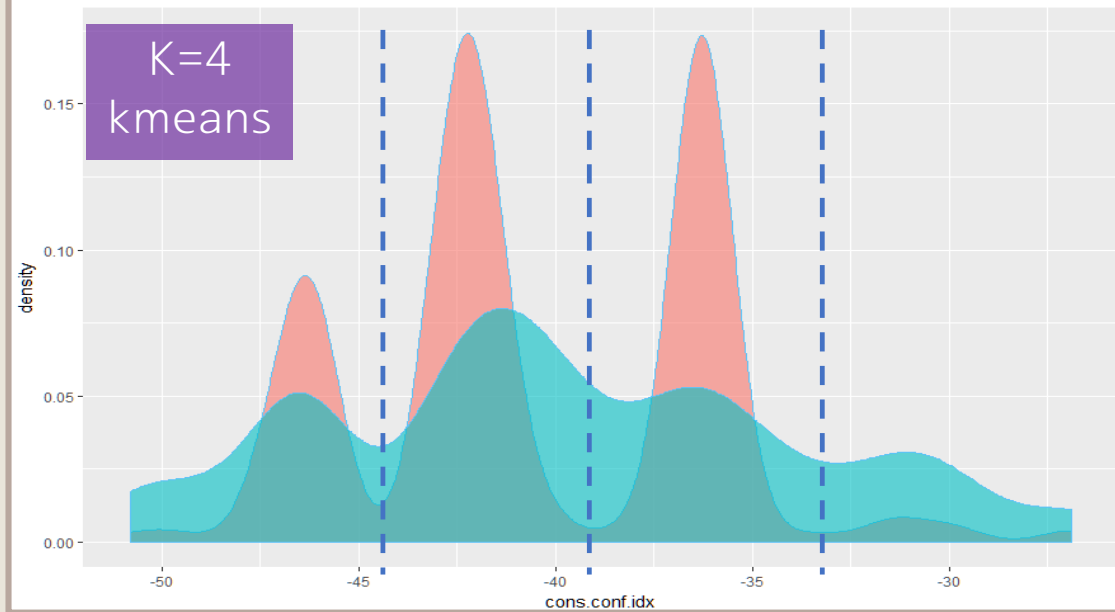
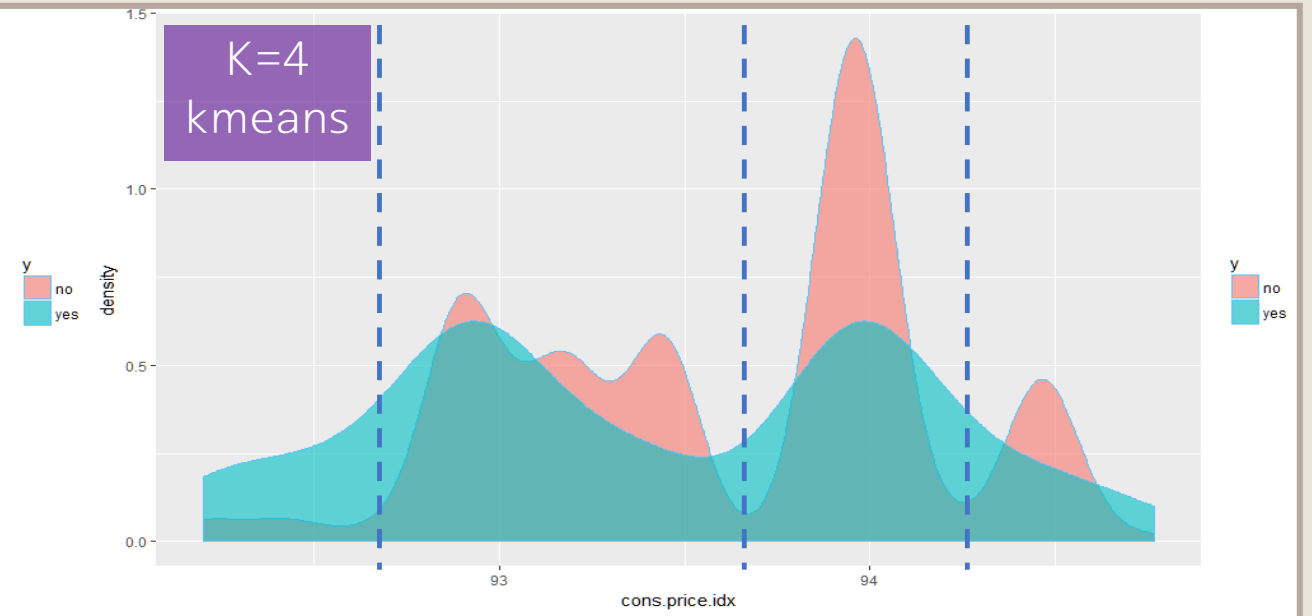
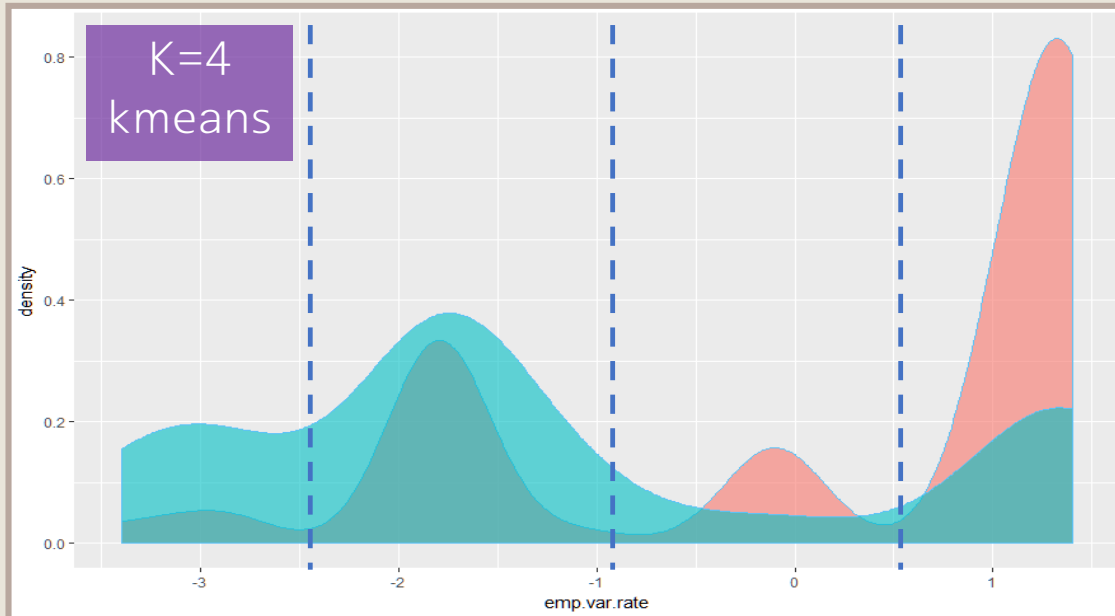


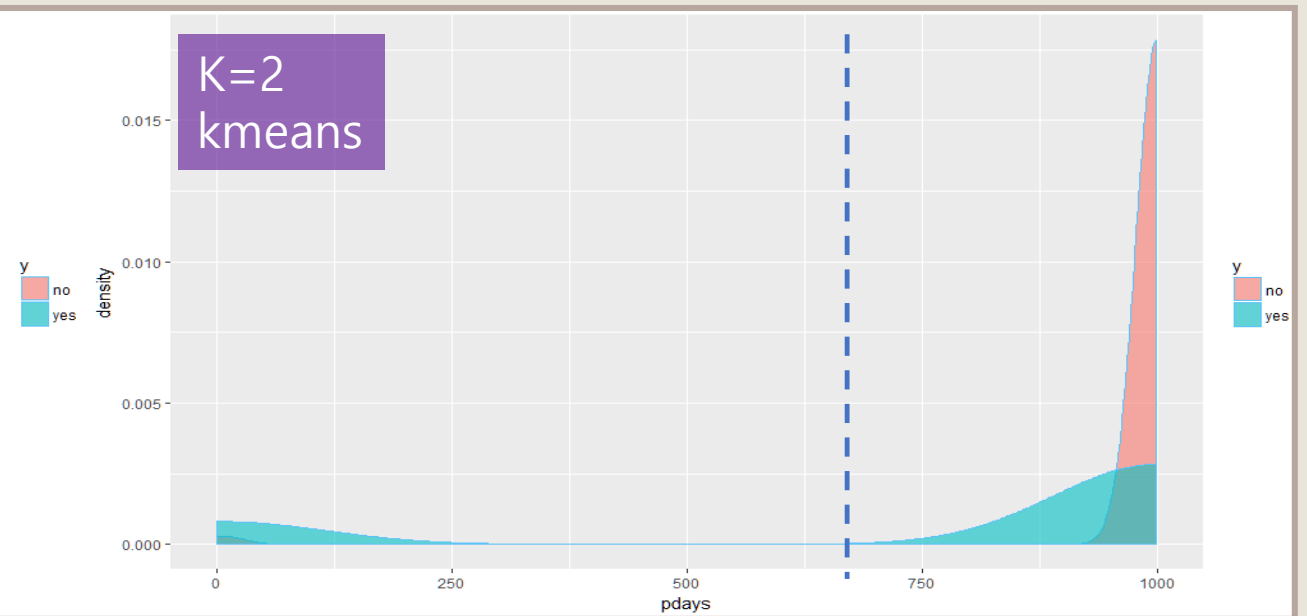
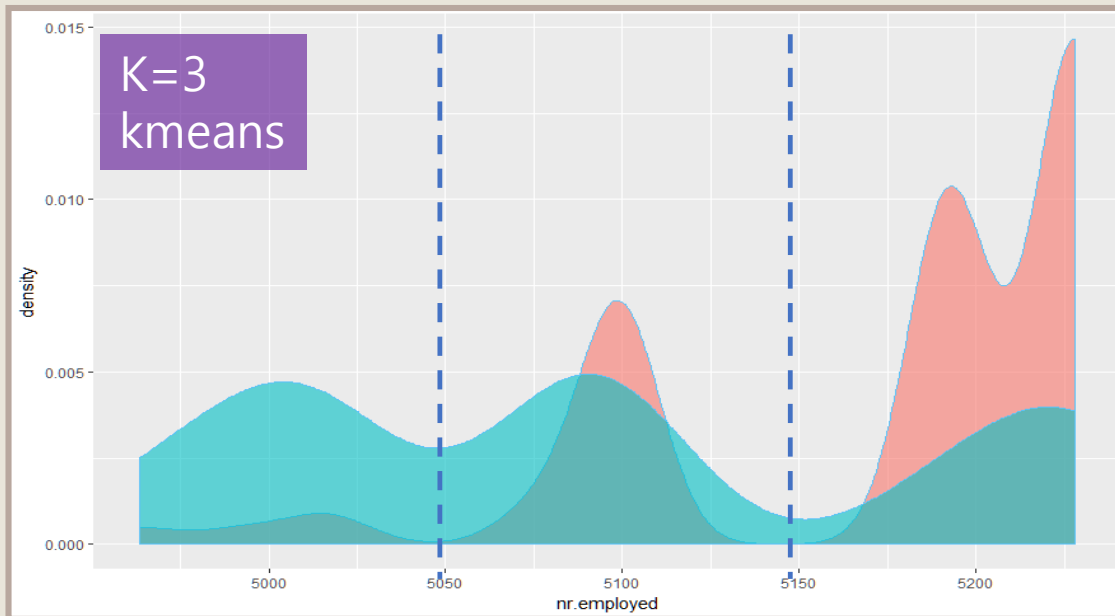
02 모델 적합

I | 분류 규칙 학습

II | 로지스틱 회귀분석

III | 연관규칙 분석





전처리 결과

```
> summary(bank[,c(1,11,12:14,16:20)])
  age      duration  campaign    pdays      previous emp.var.rate
1020s: 551    Q1:1035    Q1:1764   Group1: 160   Group1:3998   Group1: 289
30s  :1688    Q2:1038    Q2:1039   Group2:3959   Group2: 103   Group2:1053
40s  :1048    Q3:1017    Q3: 549           Group3: 14   Group3: 393
50s  : 722    Q4:1029    Q4: 767           Group4:  2   Group4:2384
60+  : 110                        Group5:  2
cons.price.idx cons.conf.idx euribor3m  nr.employed
Group1:2050     Group1: 868   Group1:1342   Group1: 355
Group2: 710     Group2:1484   Group2:2777   Group2: 987
Group3: 884     Group3: 191           Group3:2777
Group4: 475     Group4:1576
```

변수	Age	Duration	Campaign	Pdays	Previous	Emp. var.rate	Cons. price.idx	Cons. conf.idx	Euribor 3m	Nr. employed
Group 1	10, 20대	0 - 103	1	0 - 21	0 - 1	-3.4 ~ -2.9	92.201 ~ 93.444	-50.8 ~ -45.9	0.635 ~ 1.811	4963.6 ~ 5023.5
Group 2	30대	103 - 181	2	999	2 - 3	-1.8 ~ -1.1	93.749 ~ 93.918	-42.7 ~ -41.8	3.329 ~ 5.045	5076.2 ~ 5099.1
Group 3	40대	181 - 317	3		4	-0.2 ~ -0.1	93.994 ~ 94.215	-40.8 ~ -38.3		5176.3 ~ 5228.1
Group 4	50대	317 - 3643	4 - 35		5	1.1 ~ 1.4	94.465 ~ 94.767	-37.5 ~ -26.9		
Group 5	60대 이상				6					

```
> b = as(bank, "transactions")
> summary(b)
transactions as itemMatrix in sparse format with
 4119 rows (elements/itemsets/transactions) and
 92 columns (items) and a density of 0.2282609
```

most frequent items:

previous=Group1	pdays=Group2
3998	3959
loan=no	(Other)
3349	68002

element (itemset/transaction) length distribution:

sizes
21
4119

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
21	21	21	21	21	21

includes extended item information - examples:

	labels	variables	levels
1	age=1020s	age	1020s
2	age=30s	age	30s
3	age=40s	age	40s

includes extended transaction information - examples:

	transactionID
1	1
2	2
3	3

트랜잭션 데이터로 변환

```
> inspect(b[1])
```

```
items
[1] {age=30s,
    job=blue-collar,
    marital=married,
    education=basic.9y,
    default=no,
    housing=yes,
    loan=no,
    contact=cellular,
    month=may,
    day_of_week=fri,
    duration=Q4,
    campaign=Q2,
    pdays=Group2,
    previous=Group1,
    poutcome=nonexistent,
    emp.var.rate=Group2,
    cons.price.idx=Group1,
    cons.conf.idx=Group1,
    euribor3m=Group1,
    nr.employed=Group2,
    y=no}
```

트랜잭션 데이터 예시
transactionID

03 고객 특성 파악

- I. 분류 규칙 학습
- II. 로지스틱 회귀 분석
- III. 연관 규칙 분석

Trial	Rules	
	No	Errors
0	7	351(8.5%)
1	5	526(12.8%)
2	7	501(12.2%)
3	4	399(9.7%)
4	11	350(8.5%)
5	6	714(17.3%)
6	11	506(12.3%)
7	9	482(11.7%)
8	5	398(9.7%)
9	4	350(8.5%)
boost		316(7.7%) <<

----- Trial 9: -----

Rules:

Rule 9/1: (1834.2/46.4, lift 1.4)
duration <= 807
nr.employed > 5076.2
-> class no [0.974]

통화 시간 807초 이하 &
고용자 수 5076 이상

Rule 9/2: (1364.4/92, lift 1.4)
duration <= 237
-> class no [0.932]

통화 시간 237초 이하

Rule 9/3: (775.8/276.4, lift 2.6)
duration > 237
nr.employed <= 5076.2
-> class yes [0.643]

통화 시간 237초 이상 &
고용자 수 5076 이하

Rule 9/4: (674.3/247.3, lift 2.5)
duration > 807
-> class yes [0.633]

통화 시간 807초 이상

3,6,8,12월, 지난번 결과 좋음, 실직, 싱글

변수의 값이 커지면 로짓이 커진다

coefficients	
monthmar	2.474173e+00
cons.price.idx	1.409037e+00
poutcomesuccess	1.327271e+00
monthdec	8.877394e-01
poutcomenonexistent	5.792310e-01
monthjun	5.517150e-01
monthaug	4.943325e-01
jobunemployed	3.362143e-01
maritalsingle	3.276619e-01
educationbasic.6y	3.219139e-01
educationuniversity.degree	3.188565e-01
day_of_weekwed	3.125670e-01
monthoct	2.816571e-01
maritalmarried	2.781896e-01
maritalunknown	2.769494e-01
jobhousemaid	2.476944e-01
educationunknown	2.426285e-01
jobtechnician	2.208664e-01
educationbasic.9y	2.011987e-01
monthsep	1.540329e-01
educationhigh.school	1.515309e-01
defaultunknown	1.452750e-01
day_of_weekmon	1.424766e-01
jobservices	1.250967e-01
previous	1.212589e-01
educationprofessional.course	1.168553e-01

회귀 계수를 통해 변수의 영향력 파악

day_of_weekthu	1.023339e-01
monthjul	9.392480e-02
cons.conf.idx	6.441508e-02
day_of_weektue	4.485134e-02
age	8.667927e-03
duration	5.260092e-03
nr.employed	2.155554e-03
pdays	-4.959130e-04
housingyes	-6.730706e-02
jobstudent	-7.367403e-02
campaign	-9.987546e-02
loanyes	-1.144361e-01
euribor3m	-1.609976e-01
jobretired	-2.068715e-01
`jobblue-collar`	-2.268691e-01
jobmanagement	-3.136148e-01
monthmay	-3.228269e-01
monthnov	-3.241112e-01
jobunknown	-4.623714e-01
housingunknown	-5.595670e-01
`jobself-employed`	-7.304994e-01
jobentrepreneur	-7.739526e-01
emp.var.rate	-8.653427e-01
contacttelephone	-9.556266e-01
defaultyes	-8.818083e+00
educationilliterate	-1.144379e+01

변수의 값이 커지면 로짓이 작아진다

문맹, 파산한 상태, 집전화, 기업가, 자영업자

```
> inspect(rule.sorted)
```

	lhs	rhs	support	confidence	lift	count
[1]	{contact=cellular,nr.employed=Group1}	=> {y=yes}	0.03738772	0.5133333	4.688293	154
[2]	{contact=cellular,euribor3m=Group1,nr.employed=Group1}	=> {y=yes}	0.03738772	0.5133333	4.688293	154
[3]	{default=no,contact=cellular,nr.employed=Group1}	=> {y=yes}	0.03617383	0.5120275	4.676366	149
[4]	{default=no,contact=cellular,euribor3m=Group1,nr.employed=Group1}	=> {y=yes}	0.03617383	0.5120275	4.676366	149
[5]	{default=no,loan=no,contact=cellular,nr.employed=Group1}	=> {y=yes}	0.03034717	0.5102041	4.659713	125
[6]	{default=no,loan=no,contact=cellular,euribor3m=Group1,nr.employed=Group1}	=> {y=yes}	0.03034717	0.5102041	4.659713	125
[7]	{loan=no,contact=cellular,nr.employed=Group1}	=> {y=yes}	0.03107550	0.5079365	4.639003	128
[8]	{loan=no,contact=cellular,euribor3m=Group1,nr.employed=Group1}	=> {y=yes}	0.03107550	0.5079365	4.639003	128

```
> inspect(rule.sorted.no[1:8])
```

	lhs	rhs	support	confidence	lift	count
[1]	{poutcome=nonexistent,euribor3m=Group2}	=> {y=no}	0.6229667	0.9521336	1.069203	2566
[2]	{poutcome=nonexistent,nr.employed=Group3}	=> {y=no}	0.6229667	0.9521336	1.069203	2566
[3]	{poutcome=nonexistent,euribor3m=Group2,nr.employed=Group3}	=> {y=no}	0.6229667	0.9521336	1.069203	2566
[4]	{pdays=Group2,poutcome=nonexistent,euribor3m=Group2}	=> {y=no}	0.6229667	0.9521336	1.069203	2566
[5]	{previous=Group1,poutcome=nonexistent,euribor3m=Group2}	=> {y=no}	0.6229667	0.9521336	1.069203	2566
[6]	{pdays=Group2,poutcome=nonexistent,nr.employed=Group3}	=> {y=no}	0.6229667	0.9521336	1.069203	2566
[7]	{previous=Group1,poutcome=nonexistent,nr.employed=Group3}	=> {y=no}	0.6229667	0.9521336	1.069203	2566
[8]	{pdays=Group2,poutcome=nonexistent,euribor3m=Group2,nr.employed=Group3}	=> {y=no}	0.6229667	0.9521336	1.069203	2566

휴대폰 컨택, 고용자 수가 적음, Euribor 작음,
파산 하지 않음, 개인 대출 하지 않음

지난번 결과 기록 없음, Euribor 큼, 고용자 수 많
음, 이전에 컨택 적게 함, 최근에 연락

04 결과 종합

- I. 종합 하기
- II. 경영 전략 제언

방법	분류 규칙 학습		로지스틱 회귀분석		연관 규칙 분석	
분류	Yes	No	Yes	No	Yes	No
분류에 영향을 주는 특성	<ul style="list-style-type: none"> 통화 시간이 237초 이상이고 고용자 수가 5076명 이하 통화 시간이 807초 이상 	<ul style="list-style-type: none"> 통화 시간이 807초 이하이고 고용자 수가 5076 이상 통화 시간이 237초 이하 	<ul style="list-style-type: none"> 3, 6, 8, 12월 소비자 물가 지수가 높을수록 지난번 결과가 성공적일수록 실업자 싱글 6년, 대출 	<ul style="list-style-type: none"> 집전화 고용 변화율이 높을수록 기업가, 자영업자 집 대출 여부 알 수 없음 	<ul style="list-style-type: none"> 휴대폰 컨택 고용자 수 적음 Euribor 작음 파산 하지 않음 개인 대출하지 않음 	<ul style="list-style-type: none"> 이전 결과 기록 없음 Euribor 클수록 고용자 수 많을수록 이전에 컨택하지 않음 최근 연락

1. 고용지표가 클수록 가입하지 않을 가능성이 크다.
2. Euribor가 높을수록 예금 상품을 가입하지 않는다.
3. 고객과의 마케팅 통화 시간을 길게 유도한다.
4. 기존 고객의 마케팅에 집중한다.
5. 분기 전환 시점을 공략한다.
6. 집전화 대신 휴대전화로 공략한다.
7. 집 대출이 있거나 정보를 알 수 없으면 거절하는 경향이 있다.
8. 파산, 문맹 같은 극단적인 경우는 제외한다.





Q & A