



## Desafio Data Science Big Data & Analytics

---

Relatório Técnico

10 de julho de 2019

Candidata: Andreza Jardelino da Silva

---

### 1 Estudo de caso

Para a execução deste teste foi utilizado uma pesquisa relacionada à campanha de marketing realizada por uma instituição bancária Portuguesa. Para a amostragem, foram feitas chamadas telefônicas, que na sua grande maioria houve a necessidade de mais um contato por cliente. O sucesso dessa campanha seria a adesão por parte do cliente do depósito bancário à prazo, a qual foi obtida resposta do tipo binária.

Dessa forma, por se ter uma classificação na variável de interesse, sendo  $Y=0$  (não aderiu) e  $Y=1$  (aderiu), pode-se realizar análises baseadas na distribuição Bernoulli e árvore de classificação/regressão. Segue-se na Tabela 1 uma descrição das variáveis a serem avaliadas.

**Tabela 1:** Tabela resumo das descrições das variáveis em estudo

Variáveis	Descrições
y	Se o cliente aderiu ou não a proposta da campanha
age	idade do cliente
job	tipo de profissão realizada pelo cliente
marital	estado civil do cliente
education	nível de escolaridade do cliente
default	se o cliente tem dívida
balance	saldo médio anual do cliente (em Euro)
housing	se o cliente tem empréstimo imobiliário
loan	se o cliente tem empréstimo pessoal
contact	tipo de contato estabelecido com o cliente
day	último dia de contato do mês com o cliente
month	último mês de contato com o cliente
duration	tempo de duração estimado do último contato com o cliente
camping	número de contatos realizados por cliente
pdays	número de dias após o último contato com o cliente da campanha anterior
previous	número de contatos realizados por cliente na campanha anterior
poutcome	resultado da campanha de marketing anterior

## 2 Análise Estatística

```
#####  
## Carregando o conjunto de dados  
#####  
setwd("C:/Users/adaja/Desktop/Nova pasta/analise")  
dados <- read.csv("bank_full.csv", header = T, sep = ";")  
dados <- na.omit(dados)
```

```
#####  
## Dimensao e visualizacao das variaveis observadas  
#####  
dim(dados)
```

```
[1] 45211    17
```

```
head(dados, n=5)
```

	age	job	marital	education	default	balance	housing	loan	contact
1	58	management	married	tertiary	no	2143	yes	no	unknown
2	44	technician	single	secondary	no	29	yes	no	unknown
3	33	entrepreneur	married	secondary	no	2	yes	yes	unknown
4	47	blue-collar	married	unknown	no	1506	yes	no	unknown
5	33	unknown	single	unknown	no	1	no	no	unknown

	day	month	duration	campaign	pdays	previous	poutcome	y
1	5	may	261	1	-1	0	unknown	no
2	5	may	151	1	-1	0	unknown	no
3	5	may	76	1	-1	0	unknown	no
4	5	may	92	1	-1	0	unknown	no
5	5	may	198	1	-1	0	unknown	no

```
#####  
# Resumo dos dados  
#####  
summary(dados)
```

age	job	marital	education
Min. :18.00	blue-collar:9732	divorced: 5207	primary : 6851
1st Qu.:33.00	management :9458	married :27214	secondary:23202
Median :39.00	technician :7597	single :12790	tertiary :13301
Mean :40.94	admin. :5171		unknown : 1857
3rd Qu.:48.00	services :4154		
Max. :95.00	retired :2264		
	(Other) :6835		

default	balance	housing	loan	contact
no :44396	Min. : -8019	no :20081	no :37967	cellular :29285
yes: 815	1st Qu.: 72	yes:25130	yes: 7244	telephone: 2906
	Median : 448			unknown :13020
	Mean : 1362			
	3rd Qu.: 1428			
	Max. :102127			

day	month	duration	campaign
Min. : 1.00	may : 13766	Min. : 0.0	Min. : 1.000
1st Qu.: 8.00	jul : 6895	1st Qu.: 103.0	1st Qu.: 1.000
Median : 16.00	aug : 6247	Median : 180.0	Median : 2.000
Mean : 15.81	jun : 5341	Mean : 258.2	Mean : 2.764
3rd Qu.: 21.00	nov : 3970	3rd Qu.: 319.0	3rd Qu.: 3.000
Max. : 31.00	apr : 2932	Max. : 4918.0	Max. : 63.000
	(Other): 6060		

pdays	previous	poutcome	y
Min. : -1.0	Min. : 0.0000	failure: 4901	no : 39922
1st Qu.: -1.0	1st Qu.: 0.0000	other : 1840	yes: 5289
Median : -1.0	Median : 0.0000	success: 1511	
Mean : 40.2	Mean : 0.5803	unknown: 36959	
3rd Qu.: -1.0	3rd Qu.: 0.0000		
Max. : 871.0	Max. : 275.0000		

#####

## 2.1 Qual profissão tem mais tendência a fazer um empréstimo? De qual tipo?

Inicialmente, construiu-se um conjunto de dados com as variáveis tipos de empréstimos (housing e loan), de forma que ficassem organizadas em um único vetor.

```
## Construindo e organizando as observacoes no conjunto novo
dad1 <-
  dados[, ] %>%
  group_by(job, housing) %>%
  summarise(total_housing = length(housing))

dad2 <-
  dados[, ] %>%
  group_by(job, loan) %>%
  summarise(total_loan = length(loan))

resumo1 <- data.frame(dad1)
resumo2 <- data.frame(dad2)

resumo1$housing <- ifelse(resumo1$housing == "yes", "housing.yes", "housing.no")
colnames(resumo1) <- c("job", "type", "total")

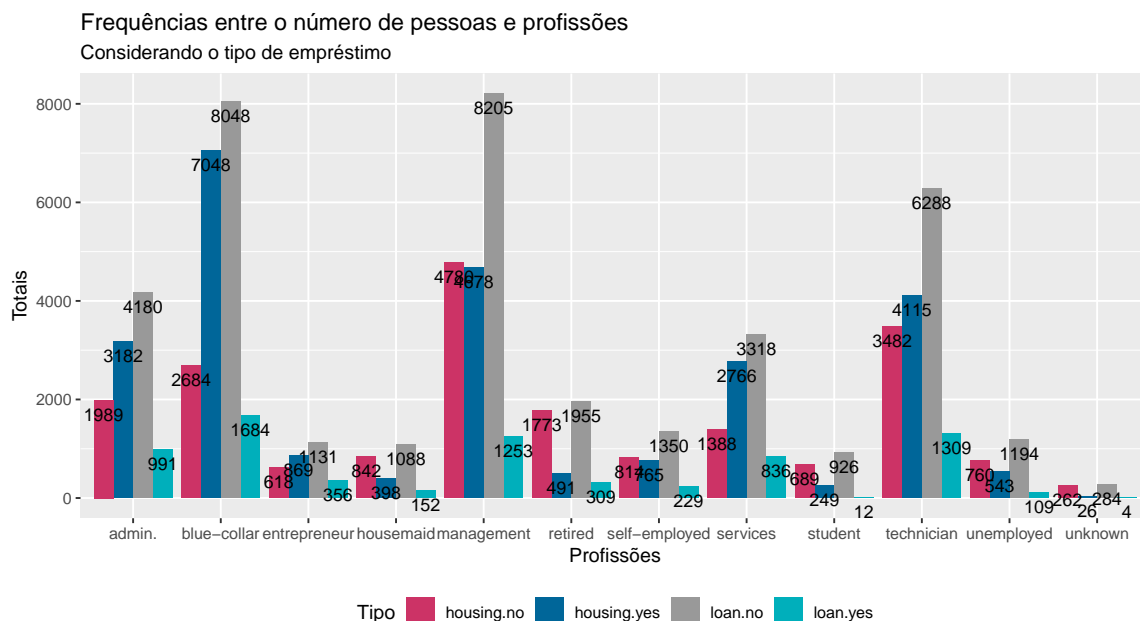
resumo2$loan <- ifelse(resumo2$loan == "yes", "loan.yes", "loan.no")
colnames(resumo2) <- c("job", "type", "total")

# Novo conjunto de dados com os totais para cada classe
resumo <- rbind.data.frame(resumo1, resumo2)
```

Pode-se dizer com base na figura 1, que o profissional da categoria operário (blue-collar) apresentou uma maior tendência à obtenção de algum tipo de empréstimo. Além disso, o empréstimo mais propenso a adesão da classe operária

será referente à imóvel, em que 7048 pessoas tinham empréstimo imobiliário no momento da pesquisa de um total de 9732 clientes dessa categoria, resultando em 72% dos trabalhadores com esse tipo de empréstimo. Por outro lado, 1684 afirmaram ter o empréstimo pessoal de um total de 9732 clientes, resultando em 17% dos trabalhadores que obtiveram esse empréstimo.

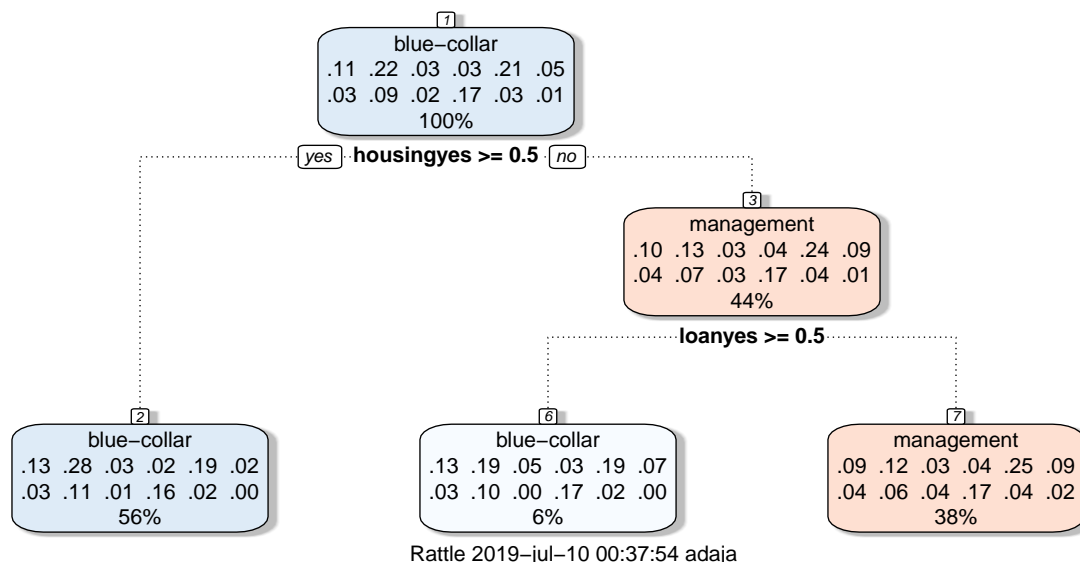
```
## Fazendo o grafico
ggplot(data = resumo, aes(x = job, y = total, fill = type)) +
  geom_bar(stat = "identity", position = position_dodge()) +
  geom_text(aes(label = total), vjust = 1.6, color = "black",
            position = position_dodge(0.9), size = 3.5) +
  scale_fill_manual(name = "Tipo",
                    values = c("#CC3366", "#006699", "#999999", "#00AFBB")) +
  theme_grey() +
  theme(legend.position="bottom") +
  labs(title = "Frequências entre o número de pessoas e profissões",
        subtitle="Considerando o tipo de empréstimo",
        y = "Totais",
        x = "Profissões")
```



**Figura 1:** Gráfico de frequência entre o número de pessoas que aderiram a campanha de marketing e profissões declaradas pelos clientes, por tipo de empréstimo

A fim de verificar as conclusões mencionadas anteriormente foi construída a árvore de classificação (Figura 2). Observou-se que a sugestão acerca da categoria operária foi corroborada, visto que o número de clientes com empréstimo imobiliário da classe operária foi 28% contra 24% referente a profissão de gestor.

```
## Confirmando os resultados pela arvore de classificacao
Mod1<- train(job ~ housing + loan, data = dados, method = "rpart")
fancyRpartPlot(Mod1$finalModel)
```



**Figura 2:** Árvore de classificação para as variáveis avaliadas

## 2.2 Fazendo uma relação entre número de contatos e sucesso da campanha. Quais são os pontos relevantes a serem observados?

Primeiramente, se faz necessário a seleção de clientes que mostraram resposta positiva para a proposta dessa campanha de marketing.

```

# Criando um vetor numerico
dados$Y <- as.numeric(ifelse(dados$y=="no", 0, 1))

## Selecionando as observacoes que contem apenas resposta considerada
# sucesso, ou seja, Y = 1 (aderiu a proposta)
dad <-
  dados[dados$Y==1,] %>%
  group_by(campaign) %>%
  summarise(resp=sum(Y))

# Novo conjunto de dados
resumo2 <- data.frame(dad)

```

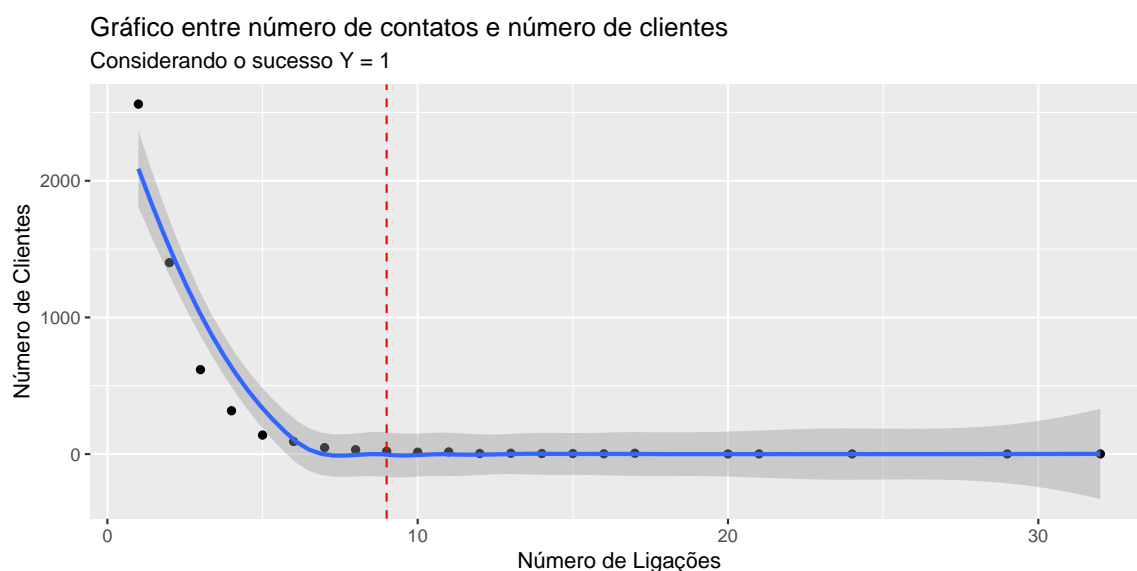
De acordo com a Figura 3, teve-se uma relação exponencial decrescente entre o número de cliente que aderiu à proposta e a quantidade de ligações recebidas. Notou-se pelo gráfico da curva que quanto mais ligações forem realizadas, menor será as chances de aceite da proposta pela campanha. Além disso, verificou-se que a partir de nove ligações houve uma similaridade quanto ao número de aceitação, e que esta tende a um, ou seja, não será necessário quantidades excessivas de ligações para conseguir a adesão do cliente.

```

## Fazendo o grafico
ggplot(resumo2, aes(x = campaign, y = resp, group = 1)) +
  geom_point() +

```

```
geom_smooth(method = loess, se = T) +
geom_vline(xintercept = 9, linetype = "dashed", col = "red") +
theme(legend.position = "right", legend.direction = "vertical") +
labs(title="Gráfico entre número de contatos e número de clientes",
      subtitle = "Considerando o sucesso Y = 1",
      y = "Número de Clientes",
      x = "Número de Ligações") +
theme_grey()
```



**Figura 3:** Relação entre o número de contatos realizados por cliente e o número de clientes que aderiram à proposta da campanha

Adicionalmente, foi ajustado o modelo de regressão no qual foi inserido no preditor linear um efeito aleatório a nível de indivíduo, visto que há uma variabilidade devido ao perfil de cada cliente. Assim, baseado-se na análise de variância e adequação do ajuste por meio do gráfico semi-normal de probabilidade com envelope de simulação (Figura 4), se concluiu com 5% de significância, que houve indícios de influência da quantidade de contatos no sucesso da campanha. Em outras palavras, é necessário realizar ligações para se obter à adesão do cliente.

```
## Ajustando o modelo considerando o efeito aleatorio a nivel de individuo.
ind <- seq(1:length(resumo2$resp))

# Modelo proposto assumindo a distribuicao Poisson, uma vez que os dados apresentam
# a caracteristica de contagem.
Mod2 <- glmer(resp ~ campaign + (1|ind),
              data = resumo2, family = poisson)

# Analise de variancia
summary(Mod2)

Generalized linear mixed model fit by maximum likelihood (Laplace
Approximation) [glmerMod]
```

```

Family: poisson ( log )
Formula: resp ~ campaign + (1 | ind)
Data: resumo2

      AIC      BIC   logLik deviance df.resid
188.7   192.0   -91.4   182.7     19

Scaled residuals:
    Min      1Q  Median      3Q      Max
-0.7816 -0.2627 -0.0455  0.1414  7.5829

Random effects:
Groups Name      Variance Std.Dev.
ind      (Intercept) 0.4467  0.6684
Number of obs: 22, groups: ind, 22

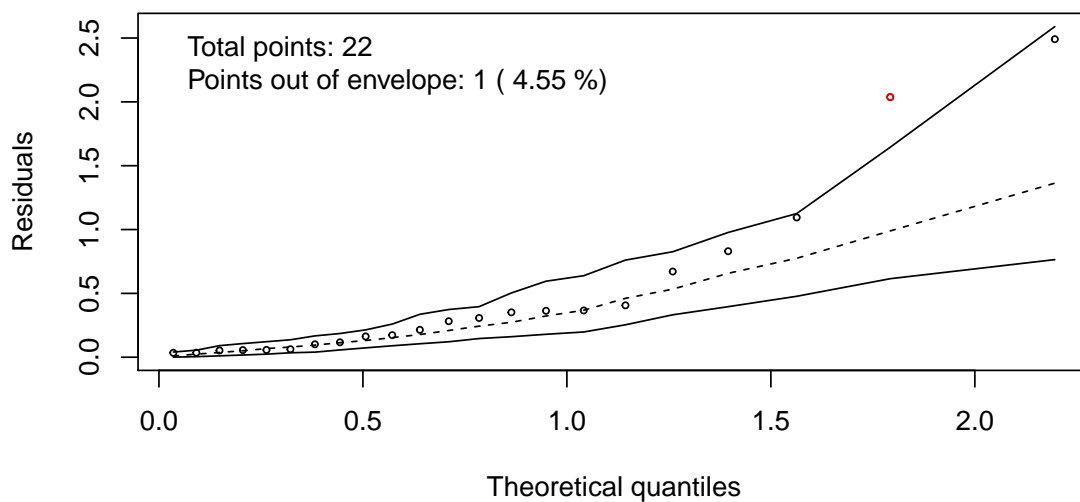
Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  6.91114    0.37657  18.353  <2e-16 ***
campaign     -0.35738    0.03877  -9.217  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
      (Intr)
campaign -0.890

# Qualidade do ajuste para o modelo proposto
hnp(Mod2, verb.sim = F, paint.out = T, print.on = T)

Poisson-normal model

```



**Figura 4:** Gráfico semi-normal de probabilidade com envelope de simulação

### 2.3 Baseando-se nos resultados de adesão desta campanha qual o número médio e o máximo de ligações que você indica para otimizar a adesão?

Segundo a análise descritiva dos dados apresentados a seguir, a média foi 2,14. No entanto, a fim de otimizar o sucesso da campanha pode-se considerar o arredondamento desse valor médio para 3 ligações, pois se conseguiu um acréscimo de, aproximadamente, 500 adesões (Figura 3). Nesse sentido, o mais indicado seria realizar, em média, 3 ligações por cliente. A sugestão ainda pode ser justificada pelos valores dos quantis, em que observou-se uma confiabilidade de sucesso em torno de 75% a 86% para essa quantidade de contatos.

```
# Separando apenas os que aderiram a campanha
```

```
sucesso <- dados[dados$Y == "1",]
```

```
# Resumo dos dados
```

```
summary(sucesso$campaign)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.000	1.000	2.000	2.141	3.000	32.000

Para a quantidade máxima de ligações, o mais indicado seria o total de 5 ligações por cliente, visto que esse número consegue retornar com 5% de significância uma boa aceitação dessa campanha de marketing.

```
# Quantis das probabilidades correspondentes para cada quantidade de ligação
```

```
quantile(sucesso$campaign, probs = seq(0,1, 0.01))
```

0%	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%	11%	12%	13%	14%
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15%	16%	17%	18%	19%	20%	21%	22%	23%	24%	25%	26%	27%	28%	29%
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
30%	31%	32%	33%	34%	35%	36%	37%	38%	39%	40%	41%	42%	43%	44%
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
45%	46%	47%	48%	49%	50%	51%	52%	53%	54%	55%	56%	57%	58%	59%
1	1	1	1	2	2	2	2	2	2	2	2	2	2	2
60%	61%	62%	63%	64%	65%	66%	67%	68%	69%	70%	71%	72%	73%	74%
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
75%	76%	77%	78%	79%	80%	81%	82%	83%	84%	85%	86%	87%	88%	89%
3	3	3	3	3	3	3	3	3	3	3	3	4	4	4
90%	91%	92%	93%	94%	95%	96%	97%	98%	99%	100%				
4	4	4	5	5	5	6	7	8	10	32				

```
#Construindo o grafico referente a quantidade ideal para o numero de
```

```
# ligacoes por cliente
```

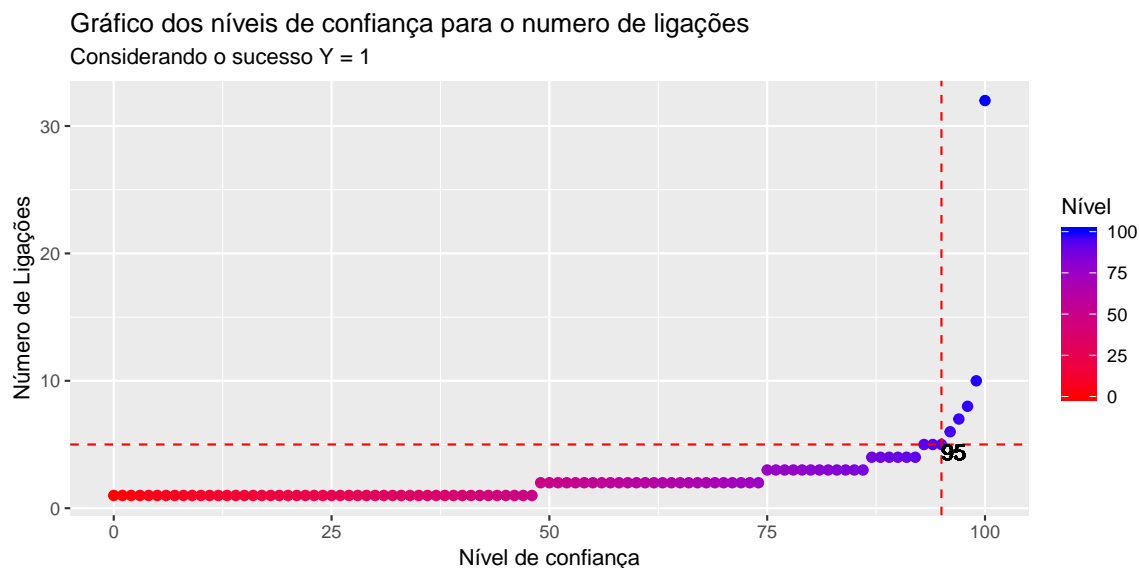
```
quant <- data.frame(quantile(sucesso$campaign, probs = seq(0,1, 0.01)))
```

```
p <- data.frame(NL = quant[,], prop = seq(0, 100, by = 1))
```

```
ggplot(p, aes(x = prop, y = NL, color = prop)) +  
  geom_point(size = 2) +  
  theme(legend.position = "bottom") +  
  scale_color_gradient(name = "Nível", low="red", high="blue") +  
  geom_vline(xintercept = 95, linetype = "dashed", col = "red") +  
  geom_hline(yintercept = 5, linetype = "dashed", col = "red") +
```



```
labs(title="Gráfico dos níveis de confiança para o numero de ligações ",
      subtitle = "Considerando o sucesso Y = 1",
      y = "Número de Ligações",
      x = "Nível de confiança") +
geom_text(x = 95, y = 5, label = "95", col = "black", vjust = 1,
          hjust = 0, parse = T) +
theme_gray()
```



**Figura 5:** Gráfico referente ao nível de confiança para o número máximo de ligações mais indicada por cliente

## 2.4 O resultado da campanha anterior tem relevância na campanha atual?

Para a obtenção da resposta foi avaliada se há, ou não, uma possível relação de dependências entre campanha anterior e atual. Nesse sentido, realizando um teste de associação segue-se o estudo.

```
# Construir uma tabela com os dados da campanha atual (CA) e campanha
# anterior (CP)
dad_camp <- data.frame(CP = dados$poutcome, CA = dados$y)

# Filtrar o data set com as observações inerentes as campanhas estudadas
filtro <- ifelse(dad_camp$CP == "other" | dad_camp$CP == "unknown",
                 F, T)

# Selecionando apenas as campanhas do total
camp <- dad_camp[filtro,]
camp <- data.frame(CP = factor(camp$CP), CA = factor(camp$CA))

# Visualizando as ocorrencias na tabela bidimensional
table(camp)
```

	CA	
CP	no	yes
failure	4283	618
success	533	978

Logo, obtido o p-valor  $< 0.05$ , pode-se dizer que rejeita-se a hipótese de nulidade ao nível de 5% de significância. Em outras palavras, há uma relação de dependência entre os resultados das campanhas atual e anterior, estatisticamente. Assim, procedimentos, ou condições oferecidas anteriormente, induzirão no resultado da campanha atual.

```
# Realizando o teste do qui-quadrado, o qual tem as seguintes
# hipoteses:
#
# H_0: Os dados são independentes
# H_a: Os dados não são independentes
```

```
(X_2<-chisq.test(table(camp)))
```

Pearson's Chi-squared test with Yates' continuity correction

```
data: table(camp)
```

```
X-squared = 1675.1, df = 1, p-value < 2.2e-16
```

## 2.5 Qual o fator determinante para que o banco exija um seguro de crédito?

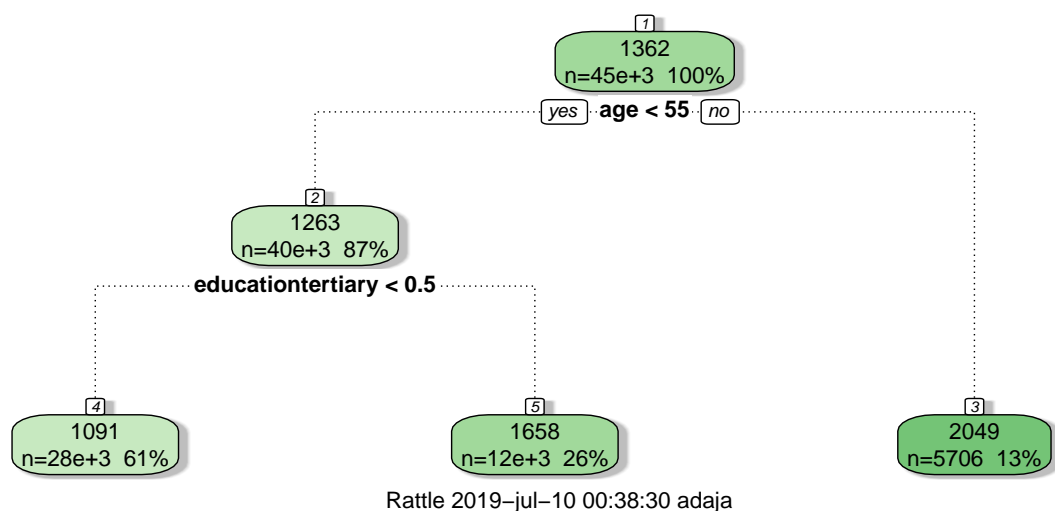
Para a análise de crédito julga-se necessário apenas as variáveis que estão condicionadas ao perfil do cliente. Foi considerado como variável de interesse para avaliar se o cliente será bom pagador, ou não, o saldo médio anual (balance), visto que esse consegue explicar a movimentação financeira do cliente. Dessa forma, observou-se por meio da Figura 6, que para a exigência de um seguro de crédito bancário é necessário que se leve em consideração a idade desse.

```
# Fazendo pela arvore de regressao
```

```
dadbank <- dados[, c(1:8)]
```

```
Mod5.1<- train(balance ~ ., data = dadbank, method = "rpart")
```

```
fancyRpartPlot(Mod5.1$finalModel)
```



**Figura 6:** Árvore de regressão para as variáveis avaliadas

## 2.6 Quais são as características mais proeminentes de um cliente que possua empréstimo imobiliário?

Considerou-se para esta análise a variável resposta binária crédito imobiliário (housing), na qual obteve-se  $Y = 0$  (não tem empréstimo) e  $Y = 1$  (tem empréstimo). Inicialmente foi realizado uma análise por meio do modelo linear generalizado, com função de ligação logística.

```
# Transformando a variavel resposta em numerica para a modelagem
dados$h <- ifelse(dados$housing == "no", 0, 1)
dados$h <- as.numeric(dados$h)

# Separando os dados em treino (60%) e teste (40%)
Train <- createDataPartition(dados$h, p=0.6, list=FALSE)
training <- dados[ Train, ]
testing <- dados[ -Train, ]

# Ajustando o modelo para verificar o que influencia um cliente ter
# emprestimo imobiliario
train_control = trainControl(method="repeatedcv", number=5, repeats=5)
mod_fit <- train(h ~ age + balance + job + marital + education + default,
                 data=training, method="glm",
                 family="binomial", trControl=train_control)
```

Os efeitos estimados referente ao modelo logístico mostrou que todas as covariáveis influenciam no saldo médio anual. Apenas alguns níveis não apresentaram diferenças em relação ao fator de referência, tais como atividades profissionais referente ao empreendedorismo e o nível de escolaridade secundário.

```
summary(mod_fit)
```

Call:

NULL

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.9982	-1.1699	0.7077	1.0144	2.4256

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	2.284e+00	9.968e-02	22.911	< 2e-16	***
age	-3.981e-02	1.589e-03	-25.060	< 2e-16	***
balance	-2.194e-05	4.585e-06	-4.787	1.70e-06	***
`jobblue-collar`	5.349e-01	5.090e-02	10.509	< 2e-16	***
jobentrepreneur	4.716e-02	8.145e-02	0.579	0.562584	
jobhousemaid	-9.175e-01	9.041e-02	-10.147	< 2e-16	***
jobmanagement	-2.386e-01	5.403e-02	-4.415	1.01e-05	***
jobretired	-1.017e+00	8.233e-02	-12.351	< 2e-16	***
`jobself-employed`	-4.529e-01	7.812e-02	-5.798	6.71e-09	***
jobservices	1.337e-01	5.720e-02	2.337	0.019460	*
jobstudent	-1.738e+00	1.055e-01	-16.470	< 2e-16	***
jobtechnician	-2.796e-01	4.886e-02	-5.723	1.05e-08	***
jobunemployed	-6.959e-01	8.305e-02	-8.380	< 2e-16	***
jobunknown	-2.278e+00	2.549e-01	-8.938	< 2e-16	***

```

maritalmarried    -1.464e-01  4.219e-02  -3.471 0.000519 ***
maritalsingle     -4.936e-01  4.866e-02 -10.146 < 2e-16 ***
educationsecondary 9.975e-02  4.331e-02   2.303 0.021277 *
educationtertiary  -2.079e-01  5.291e-02  -3.930 8.49e-05 ***
educationunknown  -2.222e-01  7.605e-02  -2.922 0.003479 **
defaultyes        -2.250e-01  9.570e-02  -2.351 0.018711 *

```

---

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

(Dispersion parameter for binomial family taken to be 1)

```

Null deviance: 37253 on 27126 degrees of freedom
Residual deviance: 34133 on 27107 degrees of freedom
AIC: 34173

```

Number of Fisher Scoring iterations: 4

Para o ajuste do modelo de regressão logística observou-se que esse apresentou uma acurácia em torno de 63% e uma predição de, aproximadamente, 68%.

```

# Validacao do modelo de regressao proposto
pred <- as.factor(ifelse(predict(mod_fit, newdata = testing) < 0.5, 0, 1))
ref <- as.factor(testing$h)
confusionMatrix(pred,ref)

```

Confusion Matrix and Statistics

```

      Reference
Prediction  0    1
0  3477 2213
1  4587 7807

      Accuracy : 0.624
      95% CI   : (0.6169, 0.631)
No Information Rate : 0.5541
P-Value [Acc > NIR] : < 2.2e-16

```

Kappa : 0.2165

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

```

      Sensitivity : 0.4312
      Specificity : 0.7791
Pos Pred Value   : 0.6111
Neg Pred Value   : 0.6299
Prevalence       : 0.4459
Detection Rate   : 0.1923
Detection Prevalence : 0.3146
Balanced Accuracy : 0.6052

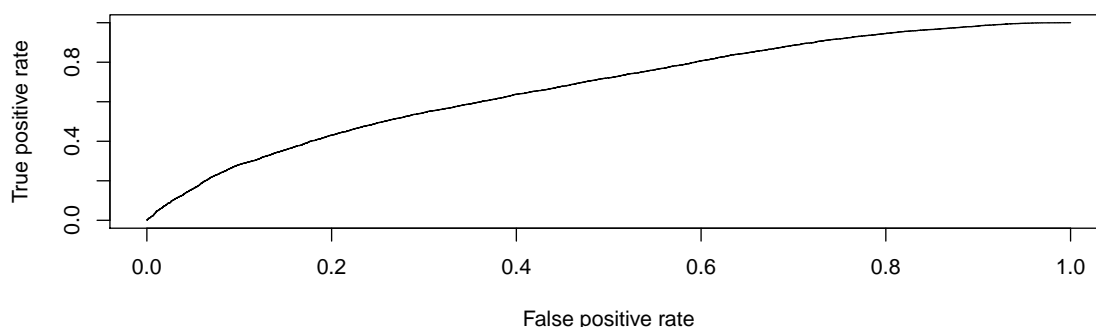
```

'Positive' Class : 0

```
# Curva ROC para o modelo de regressao proposto
prob <- predict(mod_fit, newdata = testing)
pred <- prediction(prob, testing$h)
perf <- performance(pred, measure = "tpr", x.measure = "fpr")
(auc <- performance(pred, measure = "auc") @ y.values[[1]])

[1] 0.6760498

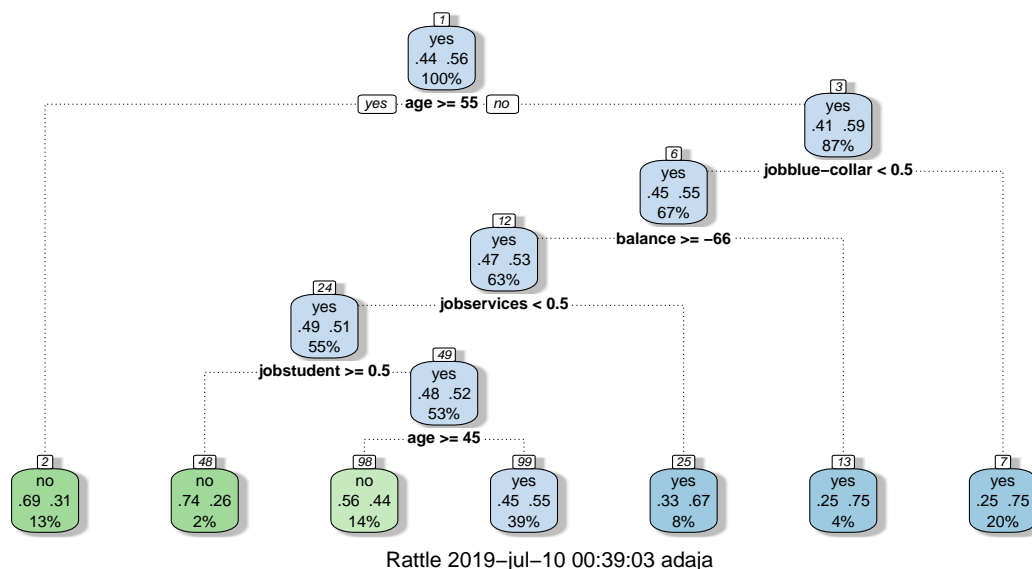
plot(perf)
```



**Figura 7:** Curva ROC para o modelo de regressão logística

Além disso, ao considerar o modelo avaliado pela árvore de classificação, tem-se que a acurácia foi 63%. Nesse sentido, tanto para o modelo logístico quanto para o modelo por árvore de regressão houve uma consistência quanto à capacidade preditiva.

```
## Realizando a analise por meio da arvore de classificacao
Mod6<- train(housing ~ ., data = dadbank, method = "rpart")
fancyRpartPlot(Mod6$finalModel)
```



**Figura 8:** Árvore de classificação para as variáveis avaliadas

```
# validacao da arvore
pred2 <- predict(Mod6, dadbank)
obs2 <- dadbank$housing
confusionMatrix(pred2,obs2)
```

Confusion Matrix and Statistics

	Reference	
Prediction	no	yes
no	8100	4772
yes	11981	20358

Accuracy : 0.6294  
 95% CI : (0.625, 0.6339)  
 No Information Rate : 0.5558  
 P-Value [Acc > NIR] : < 2.2e-16  
  
 Kappa : 0.2215  
  
 McNemar's Test P-Value : < 2.2e-16  
  
 Sensitivity : 0.4034  
 Specificity : 0.8101  
 Pos Pred Value : 0.6293  
 Neg Pred Value : 0.6295  
 Prevalence : 0.4442  
 Detection Rate : 0.1792  
 Detection Prevalence : 0.2847  
 Balanced Accuracy : 0.6067  
  
 'Positive' Class : no

No entanto, a especificidade do modelo da árvore de decisão foi maior, evidenciando uma maior capacidade do ajuste de identificar os chamados falso-positivos.

Portanto, baseados na árvore de decisão, pode-se concluir com 5% de significância, que o perfil de uma pessoa com empréstimo imobiliário é um cliente que possui uma faixa etária inferior a 55 anos, é da categoria operária e tem um saldo médio anual negativo de 66 euros.