INTRODUÇÃO

A Análise de dados embora ainda não tão explorada dentro das organizações é sem dúvida uma ferramenta indispensável ferramenta indispensável dentro dessas empresas. Nesse cenário de grades incertezas e instabilidade tona-se indispensável o uso da Estatísticas e suas técnicas para a compreensão de uma grande quantidade de dados que é gerado por essas organizações.

O desenvolvimento tecnológico, oriundo das descobertas científicas, tem apoiado o próprio desenvolvimento científico, ampliando, em várias ordens de grandeza, a capacidade de obter informações de acontecimentos e fenômenos que estão sendo analisados. Uma grande massa de informação deve ser processada antes de ser transformada em conhecimento. Portanto, cada vez mais necessita-se de ferramentas estatísticas que apresentem uma visão mais global do fenômeno, que aquela possível numa abordagem univariada. A denominação "Análise Multivariada" corresponde a um grande número de métodos e técnicas que utilizam, simultaneamente, todas as variáveis na interpretação teórica do conjunto de dados obtidos. (NETO, 2004).

Os métodos estatísticos, para analisar variáveis, estão dispostos em dois grupos: um que trata da estatística, que olha as variáveis de maneira isolada – a estatística univariada, e outro que olha as variáveis de forma conjunta – a estatística multivariada.

A Estatística Multivariada inclui os métodos de análise das relações de múltiplas variáveis dependentes e/ou múltiplas variáveis independentes, quer se estabeleçam ou não relações de causa/efeito entre estes dois grupos. São também incluídos na Estatística Multivariada os métodos de análise das relações entre indivíduos caracterizados por duas ou mais variáveis.

Só os métodos de Estatística Multivariada permitem que se explore a performance conjunta das variáveis e se determine a influência ou importância de cada uma, estando as restantes presentes.

OBJETIVOS

O Objetivo desse presente trabalho é mostrar a aplicação de técnicas estatísticas multivariadas no processamento do banco de dados (churn_missing) .O banco de dados contém 10.000 observações e 15 variáveis que são :

CustomerId: ID do cliente ,

Surname: Sobrenome ,

CreditScore: Pontução de crédito real ,

Geography: Local ,

Gender: Género ,

Age: Idade ,

Tenure: Número de meses em que o cliente permaneceu com o scompany ,

Balance: Saldo devedor do cartão de crédito ,

NumOfProdutos: Número de produtos consumido ,

HasCrCard: Tem crédito ,

IsActiveMember: Se o cliente estiver ativado no banco ,

EstimatedSalary: Salário estimado ,

Exited: se o cliente deixou de assinar o serviço ou não (1 ou 0 respectivamente).

METODOLOGIA

O Software utilizado na realização desse presente trabalho foi o R Studio um software livre de ambiente de desenvolvimento integrado para R, uma linguagem de programação para gráficos e cálculos estatísticos.

As técnicas utilizadas nesse trabalho para uma análise multivariada foram, **PCA (análise de componentes principais)** e **CLUSTER HIERARQUICO.**

A analise de componentes principais ACP ou PCA é a técnica de análise exploratória que usamos quando temos um conjunto de dados qualitativos (numéricos) e ainda não encontramos a variável dependente e as independentes. Se as variáveis do nosso conjunto forem categóricas é melhor usar o MCA ou Análise de correspondência múltipla.

Em machine learning, o PCA é uma técnica que pertence ao conjunto de "não supervisionada". Através da técnica do PCA e possível reduzir o número de dados sem perda significativa de informações e por consequência Facilitar a interpretação dos dados.

A análise de cluster pretende agrupar os dados em grupos por forma a constituir grupos em que os seus elementos sejam o mais parecidos entre si. os grupos sejam o mais diferente entre si, permite criar um centróide de cada grupo que a caracterização do elemento médio de cada grupo. Isto permite caracterizar o elemento típico de um grupo e as diferenças típicas entre grupos

PCA

LIMPANDO A MEMÓRIA E FIXANDO O DIRETÓRIO

```
rm(list = ls(all = TRUE)) setwd("~/R")
```

INSTALANDO OS PACOTES NECESSÁRIOS

```
library(FactoMineR)
library(factoextra) library(ggplot2)
library(factoextra)
```

LENDO CONJUNTO DE DADOS

```
churn_missing <- read.csv("churn_missing.csv")
View(churn_missing)
```

REMOVENDO DADOS NULOS

```
churn_missing<- na.omit(churn_missing)
```

CRIANDO SUBCONJUNTO

```
churn_missing.active<- churn_missing[1:10000,1:15]
head(churn_missing.active[1:15],4)
```

X	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard
1	1	15634602	Hargrave	619	France	Female	42	2	0.0	1	1
3	3	15619304	Onio	502	France	Female	42	8	159660.8	3	1
4	4	15701354	Boni	699	France	Female	39	1	0.0	2	0
5	5	15737888	Mitchell	850	Spain	Female	43	2	125510.8	1	1

IsActiveMember	EstimatedSalary	Exited
1	101348.88	1
0	113931.57	1
0	93826.63	0
1	79084.10	0

CARREGANDO PACOTES

```
library(missMDA)
```

SUBISTITUINDO DADOS FALTANTES

O banco de dados apresentou alguns dados faltantes que foram substituídos com o uso do comando abaixo.

```
imputePCA(churn_missing)
```

	X	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts
1	1	1	15634602	Hargrave	619	France	Female	42	2	0.00	1
3	3	3	15619304	Onio	502	France	Female	42	8	159660.80	3
4	4	4	15701354	Boni	699	France	Female	39	1	0.00	2
5	5	5	15737888	Mitchell	850	Spain	Female	43	2	125510.82	1
8	8	8	15656148	Obinna	376	Germany	Female	29	4	115046.74	4
12	12	12	15737173	Andrews	497	Spain	Male	24	3	0.00	2
13	13	13	15632264	Kay	476	France	Female	34	10	0.00	2
14	14	14	15691483	Chin	549	France	Female	25	5	0.00	2
15	15	15	15600882	Scott	635	Spain	Female	35	7	0.00	2
16	16	16	15643966	Goforth	616	Germany	Male	45	3	143129.41	2
18	18	18	15788218	Henderson	549	Spain	Female	24	9	0.00	2
20	20	20	15568982	Hao	726	France	Female	24	6	0.00	2
21	21	21	15577657	McDonald	732	France	Male	41	8	0.00	2
25	25	25	15625047	Yen	846	France	Female	38	5	0.00	1
26	26	26	15738191	Maclean	577	France	Male	25	3	0.00	2
27	27	27	15736816	Young	756	Germany	Male	36	2	136815.64	1
28	28	28	15700772	Nebechi	571	France	Male	44	9	0.00	2
30	30	30	15656300	Lucciano	411	France	Male	29	0	59697.17	2
31	31	31	15589475	Azikiwe	591	Spain	Female	39	3	0.00	3
32	32	32	15706552	Odinakachukwu	533	France	Male	36	7	85311.70	1
34	34	34	15659428	Maggard	520	Spain	Female	42	6	0.00	2
35	35	35	15732963	Clements	722	Spain	Female	29	9	0.00	2
36	36	36	15794171	Lombardo	475	France	Female	45	0	134264.04	1
38	38	38	15729599	Lorenzo	804	Spain	Male	33	7	76548.60	1
44	44	44	15755196	Lavine	834	France	Female	49	2	131394.56	1
46	46	46	15754849	Tyler	776	Germany	Female	32	4	109421.13	2
47	47	47	15602280	Martin	829	Germany	Female	27	9	112045.67	1
48	48	48	15771573	Okagbue	637	Germany	Female	39	9	137843.80	1
49	49	49	15766205	Yin	550	Germany	Male	38	2	103391.38	1
50	50	50	15771873	Buccho	776	Germany	Female	37	2	103769.22	2
51	51	51	15616550	Chidiebele	698	Germany	Male	44	10	116363.37	2
52	52	52	15768193	Trevisani	585	Germany	Male	36	5	146050.97	2
54	54	54	15702298	Parkhill	655	Germany	Male	41	8	125561.97	1
56	56	56	15760861	Phillipps	619	France	Male	43	1	125211.92	1
57	57	57	15630053	Tsao	656	France	Male	45	5	127864.40	1
58	58	58	15647091	Endrizzi	725	Germany	Male	19	0	75888.20	1
59	59	59	15623944	T'ien	511	Spain	Female	66	4	0.00	1
60	60	60	15804771	Velazquez	614	France	Male	51	4	40685.92	1
61	61	61	15651280	Hunter	742	Germany	Male	35	5	136857.00	1
63	63	63	15702014	Jeffrey	555	Spain	Male	33	1	56084.69	2
64	64	64	15751208	Pirozzi	684	Spain	Male	56	8	78707.16	1
65	65	65	15592461	Jackson	603	Germany	Male	26	4	109166.37	1
66	66	66	15789484	Hammond	751	Germany	Female	36	6	169831.46	2
68	68	68	15641582	Chibugo	735	Germany	Male	43	10	123180.01	2
69	69	69	15638424	Glauert	661	Germany	Female	35	5	150725.53	2
70	70	70	15755648	Pisano	675	France	Female	21	8	98373.26	1
72	72	72	15620344	McKee	813	France	Male	29	6	0.00	1
73	73	73	15812518	Palermo	657	Spain	Female	37	0	163607.18	1
76	76	76	15780961	Cavenagh	735	France	Female	21	1	178718.19	2
77	77	77	15614049	Hu	664	France	Male	55	8	0.00	2
79	79	79	15575185	Bushell	757	Spain	Male	33	5	77253.22	1
80	80	80	15803136	Postle	416	Germany	Female	41	10	122189.66	2
82	82	82	15663706	Leonard	777	France	Female	32	2	0.00	1
83	83	83	15641732	Mills	543	France	Female	36	3	0.00	2
86	86	86	15805254	Ndukaku	652	Spain	Female	75	10	0.00	2
89	89	89	15622897	Sharpe	646	France	Female	46	4	0.00	3
94	94	94	15640635	Capon	769	France	Male	29	8	0.00	2
97	97	97	15738721	Graham	773	Spain	Male	41	9	102827.44	1
98	98	98	15693683	Yuille	814	Germany	Male	29	8	97086.40	2
100	100	100	15633059	Fanucci	413	France	Male	34	9	0.00	2
101	101	101	15808582	Fu	665	France	Female	40	6	0.00	1
104	104	104	15776605	Bradley	528	Spain	Male	36	7	0.00	2
105	105	105	15804919	Dunbabin	670	Spain	Female	65	1	0.00	1
108	108	108	15812878	Parsons	785	Germany	Female	36	2	99806.85	1
109	109	109	15602312	Walkom	605	Spain	Male	33	5	150092.80	1
110	110	110	15744689	T'ang	479	Germany	Male	35	9	92833.89	1

	HasCrCard	IsActiveMember	EstimatedSalary	Exited
1	1	1	101348.88	1
3	1	0	113931.57	1
4	0	0	93826.63	0
5	1	1	79084.10	0
8	1	0	119346.88	1
12	1	0	76390.01	0
13	1	0	26260.98	0
14	0	0	190857.79	0
15	1	1	65951.65	0
16	0	1	64327.26	0
18	1	1	14406.41	0
20	1	1	54724.03	0
21	1	1	170886.17	0
25	1	1	187616.16	0
26	0	1	124508.29	0
27	1	1	170041.95	0
28	0	0	38433.35	0
30	1	1	53483.21	0
31	1	0	140469.38	1
32	0	1	156731.91	0
34	1	1	34410.55	0
35	1	1	142033.07	0
36	1	0	27822.99	1
38	0	1	98453.45	0
44	0	0	194365.76	1
46	1	1	126517.46	0
47	1	1	119708.21	1
48	1	1	117622.80	1
49	0	1	90878.13	0
50	1	0	194099.12	0
51	1	0	198059.16	0
52	0	0	86424.57	0
54	0	0	164040.94	1
56	1	1	113410.49	0
57	1	0	87107.57	0
58	0	0	45613.75	0
59	1	0	1643.11	1
60	1	1	46775.28	0
61	0	0	84509.57	0
63	0	0	178798.13	0
64	1	1	99398.36	0
65	1	1	92840.67	0
66	1	1	27758.36	0
68	1	1	196673.28	0
69	0	1	113656.85	0
70	1	0	18203.00	0
72	1	0	33953.87	0
73	0	1	44203.55	0
76	1	0	22388.00	0
77	1	1	139161.64	0
79	0	1	194239.63	0
80	1	0	98301.61	0
82	1	0	136458.19	1
83	0	0	26019.59	0
86	1	1	114675.75	0
89	1	0	93251.42	1
94	1	1	172290.61	0
97	0	1	64595.25	0
98	1	1	197276.13	0
100	0	0	6534.18	0
101	1	1	161848.03	0
104	1	0	60536.56	0
105	1	1	177655.68	1
108	0	1	36976.52	0
109	0	0	71862.79	0
110	1	0	99449.86	1

EXCLUIR COLUNAS

Foram retiradas 11 colunas do banco de dados pois elas não são significativas para esse estudo em questão.

```
churn_missing<-churn_missing[,-c(1,2,3,4,5,6,7,8,10,14,15)]
head(churn_missing)
```

As 4 variáveis que vão ser trabalhada nessa análise serão o Número de pro dutos que o cliente tem no banco, se ele é um cliente ativo ou não, se tem c artão de crédito e o Número de meses que o cliente permaneceu no banco.

```
Tenure NumOfProducts HasCrCard IsActiveMember
1      2             1         1              1
2      1             1         0              1
3      8             3         1              0
4      1             2         0              0
5      2             1         1              1
6      8             2         1              0
```

PADRONIZAÇÃO

```
 churn_missing.active <-scale(churn_missing)
```

GERANDO PCA

```
res.pca<-PCA(churn_missing, graph = F)
View(res.pca)
```

EXTRAIR AUTOVALORES

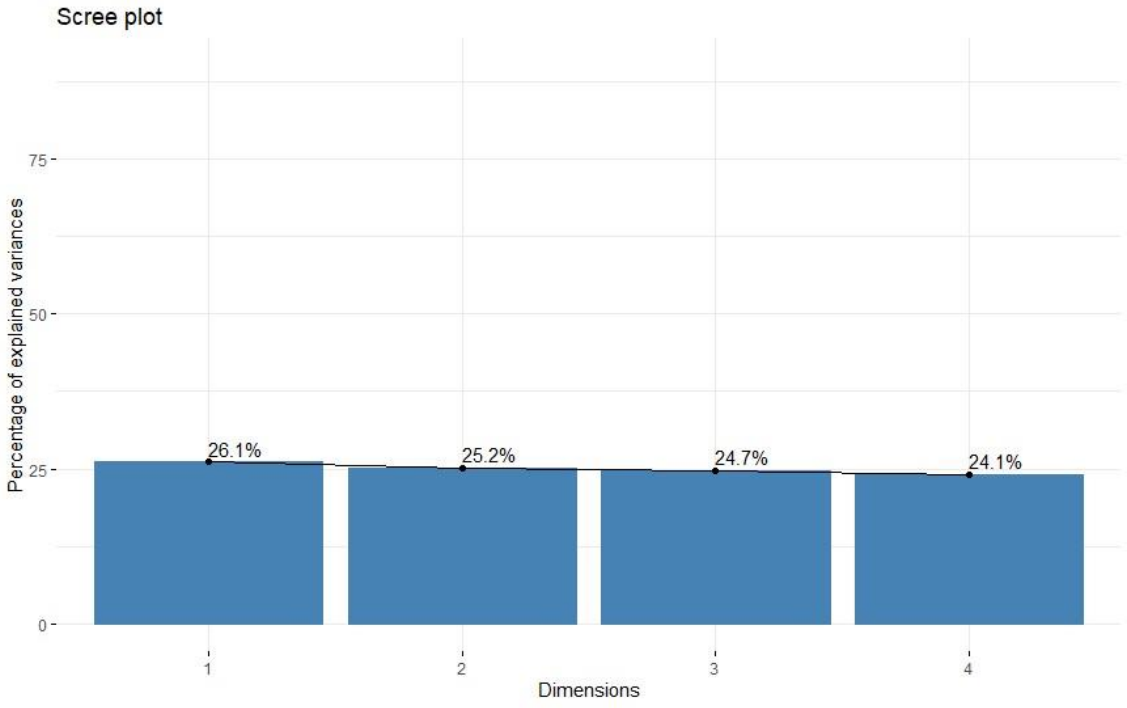
```
eig.val<- get_eigenvalue(res.pca) eig.val
 res.pca <- prcomp(churn_missing,  scale = TRUE)
```

	eigenvalue	variance.percent	cumulative.variance.percent
Dim.1	1.0431586	26.07896	26.07896
Dim.2	1.0074298	25.18574	51.26471
Dim.3	0.9865259	24.66315	75.92786
Dim.4	0.9628857	24.07214	100.00000

Podemos usar o tamanho dos autovalores para determinar o número de componentes principais. usando o critério Kaiser, usaremos somente os componentes principais com os autovalores que são maiores que 1.

PLOTAR O GRÁFICO

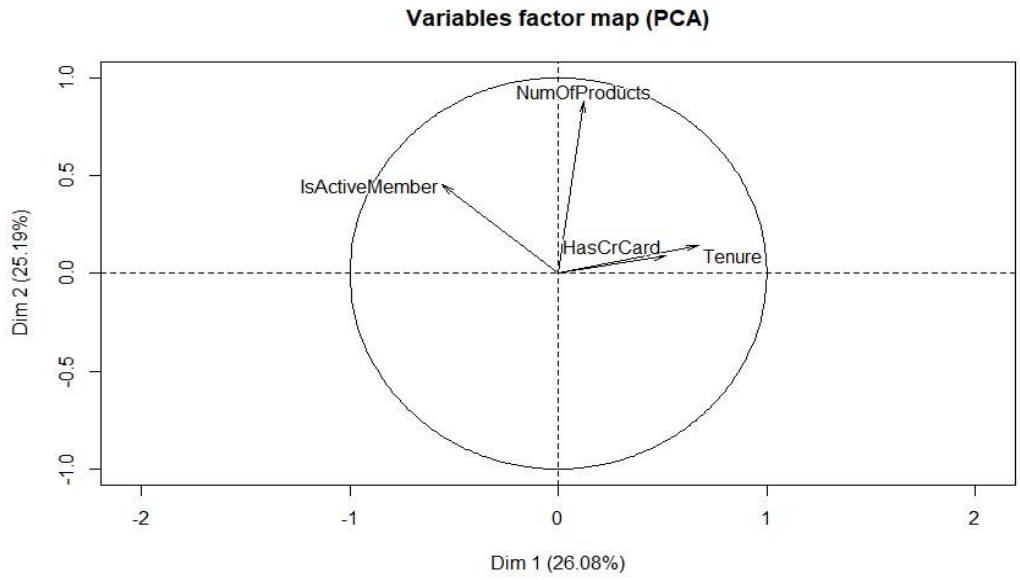
```
fviz_eig(res.pca, addlabels = TRUE, ylim = c(0, 90))
```



Esse gráfico mostra a porcentagem das variâncias estimadas. Podemos notar que quase 100% das informações (variações) contidas nos dados são retidas pelos quatro primeiros componentes principais.

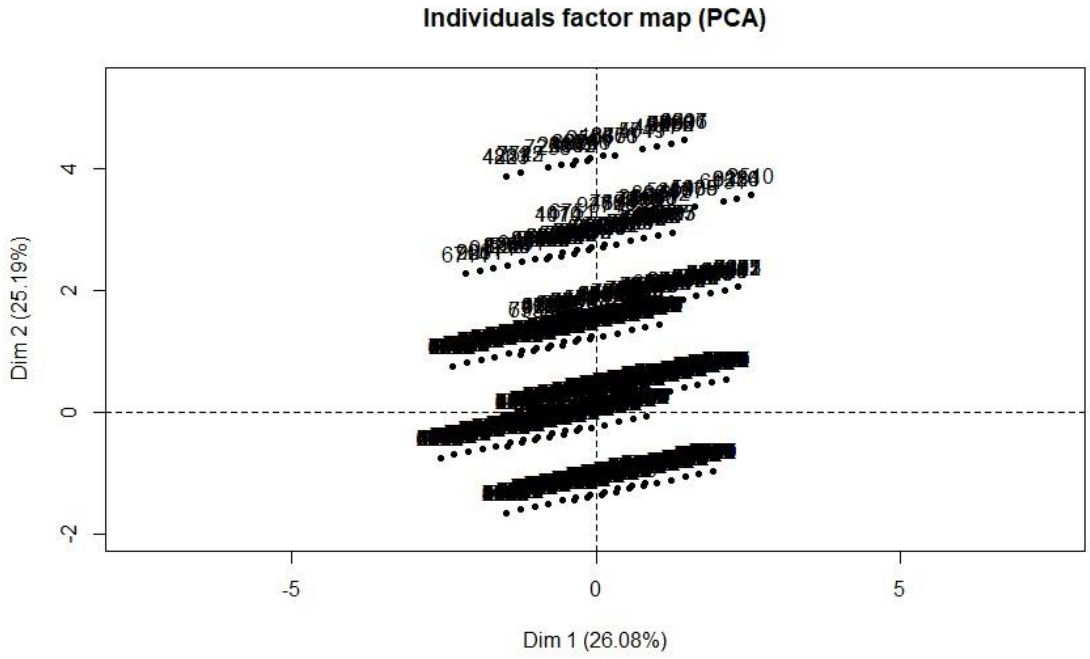
Outros gráficos que mostram a posição das variáveis e a concentração dos dados.

```
plot(PCA(churn_missing))
```



Nesse outro gráfico de correlação das variáveis mostra a relação de todas as variáveis onde as variáveis positivamente correlacionadas são agrupadas enquanto as variáveis negativamente correlacionadas são posicionadas em lados opostos da origem dos gráficos (quadrantes opostos).

No gráfico acima temos 3 variáveis positivamente correlacionadas no lado superior direito que são o número de produtos que o cliente possui no banco, se tem ou não cartão de crédito e o número de meses que o cliente permaneceu no banco. Já a variável ISACTIVEMENBER que diz se o cliente está ativo ou não tem uma correlação negativa com as demais variáveis, ou seja , ser um cliente ativo ou não, não implica no cliente possuir ou não algum produto financeiro



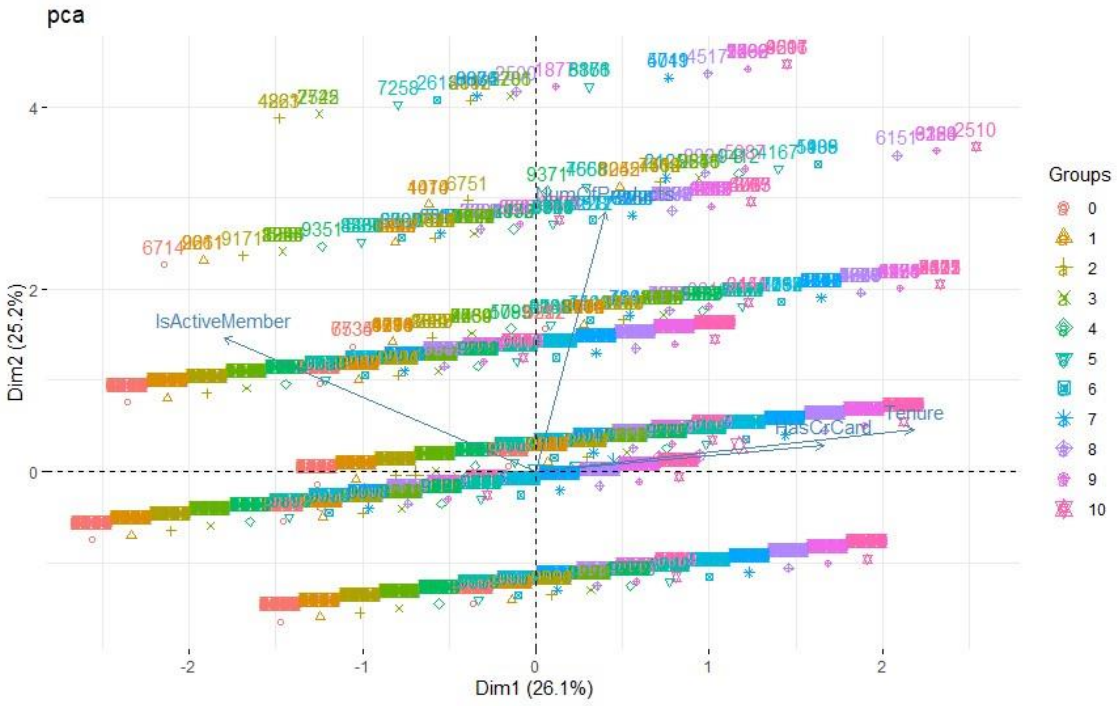
CLUSTER

CRIAR GRUPO CLUSTER

```
grupo<- as.factor(churn_missing [,1])
```

PLOTAR GRÁFICO BIPLOT

```
fviz_pca_biplot(res.pca , habillage = grupo , title = " GRUPO")
```

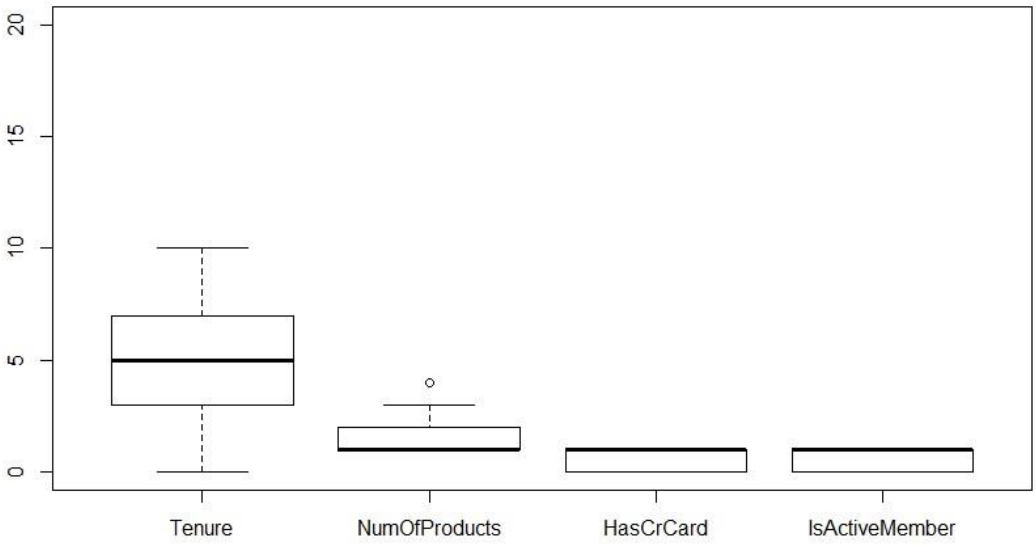


Nesse gráfico mostra os grupos de clusters.

PLOTANDO O GRÁFICO BOXPLOT

O gráfico Boxplot avalia diversos pontos como simetria dos dados, valores discrepantes, concentração de valores.

```
  boxplot(anadados, ylim= c
(0,20))
```



Podemos notar no gráfico acima que a variável TENURE apresenta os seus dados concentrados entre o primeiro e o segundo quartil, já a variável NUMOFPRODUCTS apresentou dados discrepantes (dados que parecem não fazer parte do conjunto de dados). Observa-se que a maioria dos clientes possuem um único produto.

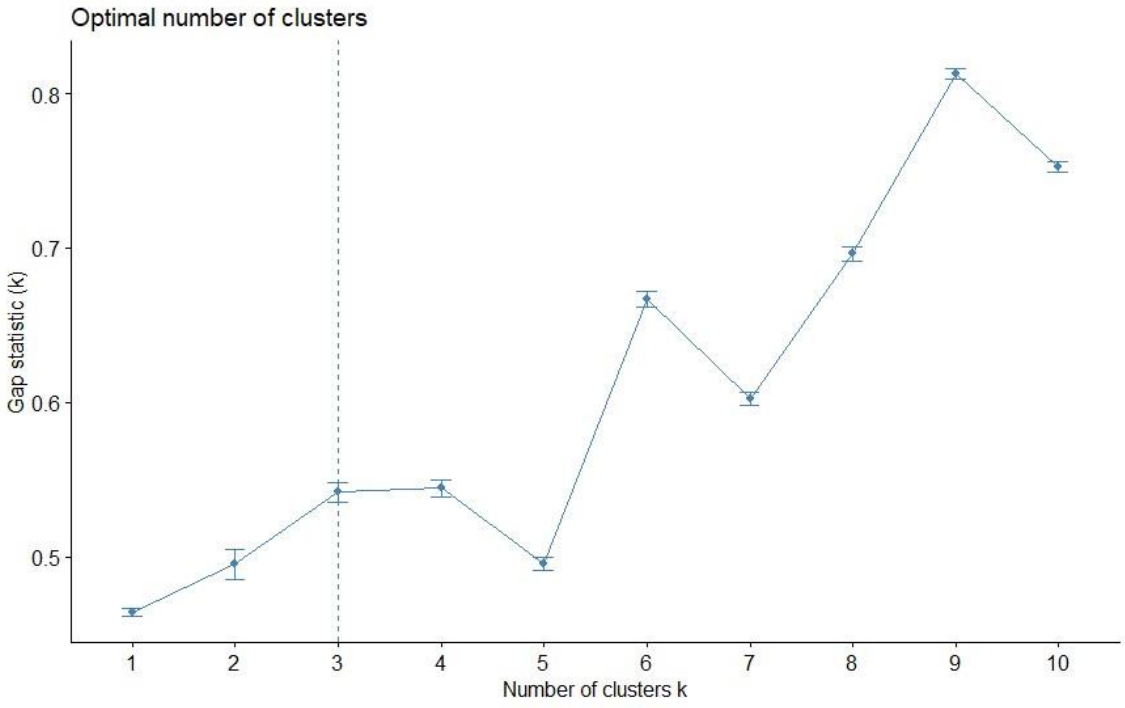
TRANSFORMAR EM ESCALA

```
dados<- scale(anadados)
```

DEFINIR QUANTIDADDE DE CLUSTERS

```
fviz_nbclust(dados, kmeans, method =  "gap_stat", print.summary = TRUE
)
```

```
Clustering k = 1,2,..., K.max (= 10): .. done
Bootstrapping, b = 1,2,..., B (= 100)  [one "." per sample]:
.................................................. 50
.................................................. 100
```



Conforme resultado do gráfico acima mostra que a quantidade de clusters ideal para se trabalhar nesse conjunto de dados é 3. Já que a variável ISACTIVEMERBER já mostrou um correlação negativa ela foi tirada no nosso modelo pois não representa o problema em questão.

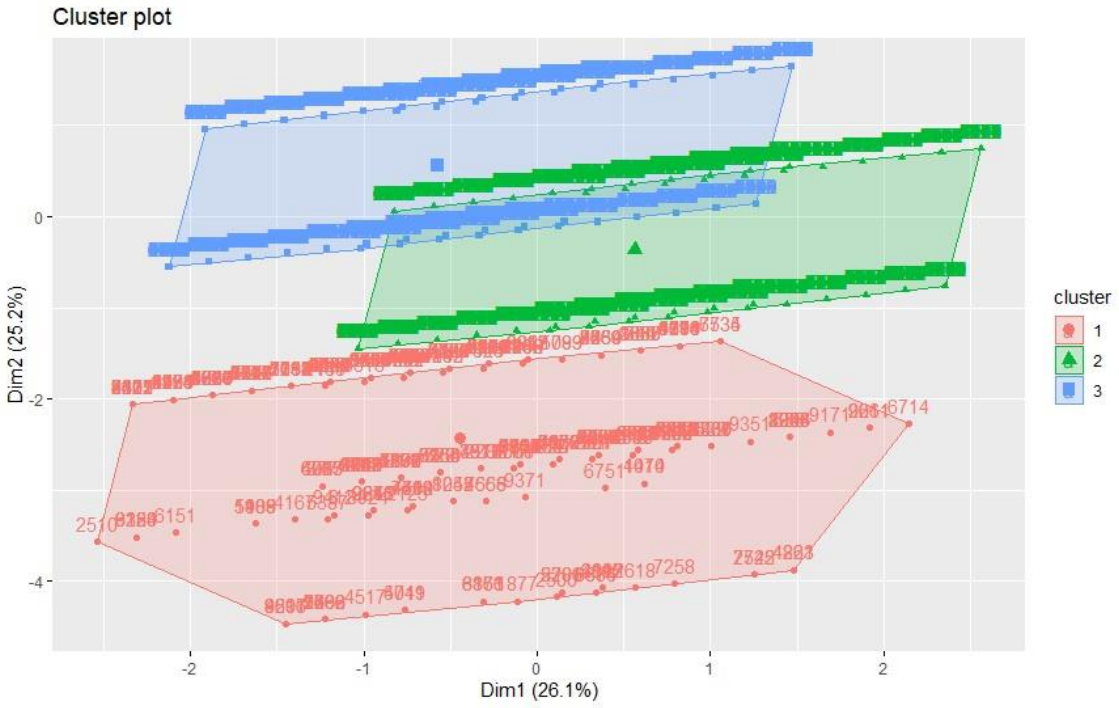
Através dos 3 clusters apontados pelo modelo podemos verificar a correlação dessas variáveis e através delas poder verificar se tem ou não algum problema e suas possíveis soluções.

GERAR KMENS

```
dados_kmeans<- kmeans(dados, 3)
```

VISUALIZAR O GRÁFICO

```
 fviz_cluster(dados_kmeans, data =
dados)
```



Nesse gráfico podemos observar os 3 principais grupos de clusters de forma separada. Em rosa temos o grupo de clientes e o tempo de permanência no banco, em verde temos o grupo de clientes que possuem algum produto financeiro e em azul temos os grupo de clientes que possuem cartão de crédito.

```
dados_kmeans$cluster
```

```
  [1] 2 2 1 3 2 3 2 1 2 2 3 3 3 2 2 3 2 3 2 2 3 3 2 2 2 2 3 2 2 1 2 3 2 2 3 2 2 2 2 3 3 2 3 2 2 2 2 2 3 3 3 3 3 2 3
 [58] 3 3 2 3 3 3 2 2 2 3 2 2 3 1 3 2 2 2 3 2 2 2 3 3 3 3 2 3 2 3 3 1 2 1 3 3 2 2 2 2 2 3 3 2 3 3 3 2 2 2 2 3 3 2 3 3 2
115] 2 2 3 3 3 3 3 3 3 3 2 3 3 2 3 3 3 2 2 2 2 2 2 2 2 3 2 3 2 2 3 3 2 3 3 3 2 3 3 3 3 2 3 3 3 3 2 3 3 2 2 3 2 3 3 2 3
172] 3 3 3 2 2 2 3 2 2 3 2 2 2 3 3 2 2 3 3 3 3 3 3 2 2 2 2 3 2 2 2 2 3 2 3 2 2 3 2 3 3 3 2 2 2 2 3 3 2 3 2 1
229] 3 3 2 3 2 2 2 3 2 1 1 2 3 3 2 2 3 3 3 2 2 3 3 2 2 2 3 2 2 3 2 2 3 2 2 2 3 3 2 2 3 2 3 2 1 2 3 2 2 2 3 2 3 2 2 2 2 2
286] 3 3 3 2 3 2 3 3 3 2 3 2 1 2 2 3 3 3 2 3 2 2 3 3 3 2 3 3 3 2 3 3 2 3 3 2 2 3 3 3 3 3 3 2 2 2 3 2 3 3 3 2 3 3 2 3
343] 2 3 3 3 3 2 2 2 2 2 3 2 2 3 2 3 2 2 2 2 2 3 2 2 3 2 2 2 2 2 2 2 1 3 3 1 2 3 2 3 3 2 3 2 2 3 3 2 2 3 3 2 2 3 2 2
400] 3 3 2 2 2 2 3 3 3 2 3 3 2 2 1 3 3 3 2 3 3 3 3 3 3 3 3 3 2 3 3 2 2 3 3 3 2 3 3 3 2 3 3 2 2 3 2 3 2 3 3 3 2 3 3 2 3
457] 2 3 2 2 2 3 3 2 2 3 2 1 3 2 2 3 3 3 3 3 2 2 2 2 3 1 3 3 2 3 2 2 2 2 2 2 3 3 2 3 2 2 2 2 2 3 2 3 3 3 3 2 2 2 3 3
514] 2 2 3 1 3 2 2 3 2 2 2 3 2 2 3 3 2 2 2 2 3 3 3 3 3 1 3 2 2 2 3 1 3 3 3 2 1 3 2 3 3 3 1 3 3 2 2 2 2 2 3 2 2 3 2 1
571] 2 3 3 2 2 2 3 2 3 2 3 2 2 3 3 3 3 2 3 3 2 3 3 2 3 2 3 2 3 2 3 3 2 3 2 1 2 2 3 2 2 2 3 2 3 3 3 2 2 2 3 3 3 2 2 3 2
628] 3 2 3 3 3 3 2 3 3 3 2 2 3 3 3 2 3 3 3 2 2 2 3 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 3 2 3 2 2 2 3 2 2 3 2 2 2 3 2 3
685] 3 3 2 3 3 3 3 3 3 3 2 2 2 2 2 2 2 3 3 2 2 2 1 2 3 2 3 3 3 3 3 3 1 3 3 3 3 2 2 2 3 3 2 2 2 1 3 3 2 3 3 3 2 2 2 2 3 3 2
742] 2 2 1 2 2 3 3 2 2 2 3 2 3 3 3 3 2 3 2 2 3 3 3 3 3 1 3 2 3 3 2 3 3 2 3 3 3 2 3 3 2 3 3 2 3 2 2 3 2 2 2 2 2 3 2 2 2 3 2 3 3 2 3
799] 2 2 2 2 2 2 2 3 3 2 2 3 3 2 2 1 3 3 3 3 2 3 3 3 3 3 2 3 2 2 2 2 3 3 2 2 2 2 3 3 3 2 3 3 3 3 3 3 2 2 3 3 2 2 2
856] 3 3 2 2 3 3 3 2 2 3 2 2 3 3 2 2 1 2 3 2 3 2 2 3 2 3 1 3 3 2 2 3 2 2 2 3 2 3 1 3 2 3 3 3 3 3 2 2 3 3 2 2 3 2 3 2 2
913] 3 3 2 3 2 2 3 3 2 3 1 2 2 2 2 2 2 3 3 2 3 2 1 2 2 2 2 3 2 3 3 3 3 2 3 2 2 3 2 3 2 2 2 3 2 2 3 2 1 2 2 2 3 3
970] 2 2 2 3 2 3 2 3 3 3 3 3 3 2 3 2 2 3 3 2 3 3 2 2 3 3 1 2 2 2 3
 reached getOption("max.print") -- omitted 9000 entries ]
```

CONCLUSÃO

A Análise multivariada foi de extrema importância na análise desse banco de dados que apresentava grande quantidade de variáveis para ser analisadas simultaneamente. Pois somente através da análise multivariada podemos definir os grupos de variáveis, quais as variáveis que tinham maior significância para ser analisadas.

O PCA mostrou alguns dados faltantes no banco de dados após corrigimos esses dados, tiramos também as variáveis qualitativas que não podem ser analisadas por esse tipo de método estatístico e depois eliminamos algumas variáveis que não foram relevante para a análise , O PCA também mostrou a direção dessas variáveis a suas dimensões , se essas variáveis apresentavam uma significância positiva ou negativa, quanto aos outliers também nos mostrou qual variável apresentava algum dados discrepante.

A Análise de cluster nos possibilitou ver essas variáveis de uma forma mais clara demostrando no gráfico os grupos de clusters e seu posicionamento no gráfico, assim como também poder separar quais os números de clusters que seriam de maior importância para o estudo. O número de clusters ideal foram 3 assim como mostrado no último gráfico podemos ver cada grupo ordenado e separado devidamente e suas proporções.

No último gráfico segundo resultado do número ideal de clusters a ser estudado relacionamos o tempo que o cliente permaneceu no banco com o número de produtos que eles tinham e se dentre esses produtos algum era cartão de crédito.

Segundo o gráfico de clusters podemos perceber que a variável tempo de permanência do cliente no banco (TENURE) não tem relação se o cliente possui ou não um produto financeiro e se esse mesmo é um cartão de crédito. Já as outras duas variáveis que é o grupo de pessoas que tem algum produto financeiro e as pessoas que tem cartão de crédito se relacionam mesmo que de uma forma pequena .

A Estatística Multivariada serve para cruzar variáveis e através deste cruzamento analisar algum tipo de problema dentro de um grupo , nesse caso serviu para identificar o grupo de pessoas que ainda não tem cartão de crédito pois o grupo apareceu com um

grande porcentagem de pessoas sem esse serviço , o que poderia da uma direção a empresa a para analisar o porque que essas pessoas ainda não tem esse produto financeiro e intensificar a oferta desse produto dentro da empresa.

ANA PAULA DE SOUZA VANDERLEY

SALVADOR - 2019