

UNIVERSIDADE DE PERNAMBUCO – UPE
ESCOLA POLITÉCNICA DE PERNAMBUCO - POLI
ESPECIALIZAÇÃO EM CIÊNCIA DOS DADOS E ANALYTICS

PROJETO PRÁTICO DE ESTATÍSTICA COMPUTACIONAL

RUBEM MORAIS NOVELLINO FERRAZ

RECIFE

2020

Sumário

1. Escolha da base de dados	3
2. Problema de regressão e suas variáveis	3
3. Estatísticas descritivas	5
4. Gráficos.....	6
5. Correlação e gráficos de dispersão	10
6. Boxplot.....	13
7. Regressão	14
8. Monte Carlo	15

1. Escolha da base de dados

A base de dados escolhida é o microdados do Censo da Educação Superior de 2019. Estes são publicados anualmente e trazem informações sobre diversos temas relacionados aos alunos, os cursos, os docentes, o local da oferta e as IES em si. Esta última servirá como base para o estudo.

Os arquivos utilizados são públicos e podem ser consultados no site do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) por meio do seguinte link: http://download.inep.gov.br/microdados/microdados_educacao_superior_2019.zip

2. Problema de regressão e suas variáveis

O problema a ser estudado é relativo à influência das despesas em áreas diversas de uma IES na sua respectiva receita.

Serão utilizados os sete tipos de despesas disponíveis, ou seja:

- Com pessoal docente;
- Com pessoal técnico;
- Com encargos de pessoal;
- Com custeio;
- Com investimentos;
- Com pesquisa e
- Outras.

Por mais que não seja o ponto fundamental de várias IES, conforme sua categoria administrativa, utilizaremos as receitas para este estudo. Como elas estão divididas em ‘própria’, ‘transferência’ e ‘outros’, optamos por utilizar a própria pois deve estar mais ligada ao seu corpo técnico.

Por fim, a categoria administrativa será usada como qualitativa nominal. Ela, bem como as demais variáveis citadas, está descrita no ‘dicionário de variáveis’. Como indicado no

mesmo, as demais variáveis são quantitativas contínuas. O nome, descrição e valores possíveis das variáveis que serão utilizadas seguem:

NOME DA VARIÁVEL	DESCRIÇÃO DA VARIÁVEL	TIPO	TAM	DESCRIÇÃO DAS CATEGORIAS
TP_CATEGORIA_ADMINISTRATIVA	Tipo da Categoria Administrativa	Num	1	1. Pública Federal 2. Pública Estadual 3. Pública Municipal 4. Privada com fins lucrativos 5. Privada sem fins lucrativos 6. Privada - Particular em sentido estrito 7. Especial 8. Privada comunitária 9. Privada confessional
VL_DESPESA_PESSOAL_DOCENTE	Valor das despesas com remuneração de pessoal - docentes ativos da IES ou da mantenedora	Num	14	
VL_DESPESA_PESSOAL_TECNICO	Valor das despesas com remuneração de pessoal - docentes ativos da IES ou da mantenedora	Num	14	
VL_DESPESA_PESSOAL_ENCARGO	Valor das despesas com benefícios e encargos sociais de todo o pessoal da IES ou da mantenedora	Num	14	
VL_DESPESA_CUSTEIO	Valor das despesas para a manutenção e custeio da IES ou da mantenedora (Não inclui despesas com pessoal.)	Num	14	
VL_DESPESA_INVESTIMENTO	Valor das despesas de investimentos (despesas de capital) realizadas nas IES ou na mantenedora	Num	14	
VL_DESPESA_PESQUISA	Valor das despesas com Pesquisa e Desenvolvimento da IES ou da mantenedora	Num	14	
VL_DESPESA_OUTRA	Valor de dispêndios não contemplados nos demais campos relativos à despesa	Num	14	
VL_RECEITA_PROPRIA	Valor das receitas próprias auferidas pela Mantenedora ou pela IES no ano de referência	Num	14	

3. Estatísticas descritivas

Após leitura dos dados usando a linguagem R, obtemos as seguintes estatísticas descritivas para as variáveis quantitativas:

VL_DESPESA_PESSOAL_DOCENTE (x 10⁶)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Sd.	Var
0.000	0.580	3.373	65.204	34.629	6780.000	191.0717	36508.4

VL_DESPESA_PESSOAL_TECNICO (x 10⁶)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Sd.	Var
0.0000	0.2475	0.9435	15.9453	5.6434	1317.8635	61.76276	3814.638

VL_DESPESA_PESSOAL_ENCARGO (x 10⁶)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Sd.	Var
0.0000	0.2959	1.7258	27.0545	13.3537	1803.2958	70.62087	4987.307

VL_DESPESA_CUSTEIO (x 10⁶)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Sd.	Var
0.0000	0.1582	1.9339	60.6847	33.5490	1590.8835	181.6245	32987.45

VL_DESPESA_INVESTIMENTO (x 10⁶)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Sd.	Var
0.0000	0.0000	0.1305	16.5218	3.4346	535.5536	56.50109	3192.373

VL_DESPESA_PESQUISA (x 10⁶)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Sd.	Var
0.0000	0.0000	0.0000	1.0263	0.1543	181.4141	5.805779	33.70707

VL_DESPESA_OUTRA (x 10⁶)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Sd.	Var
0.0000	0.0014	0.4035	50.4734	8.2585	1115.2670	145.2738	21104.47

VL_RECEITA_PROPRIA (x 10⁶)

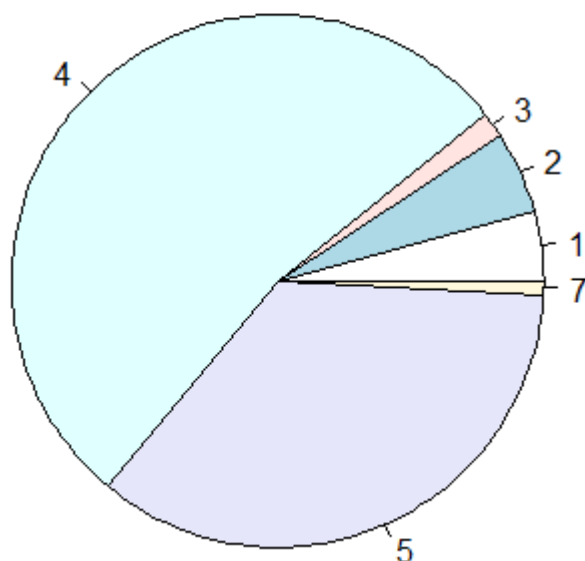
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Sd.	Var
0.000	1.092	7.170	182.545	61.781	3609.180	500.851	250851.6

4. Gráficos

A fim de melhor visualizar as estatísticas calculadas, vamos nos utilizar de gráficos adequados para cada caso.

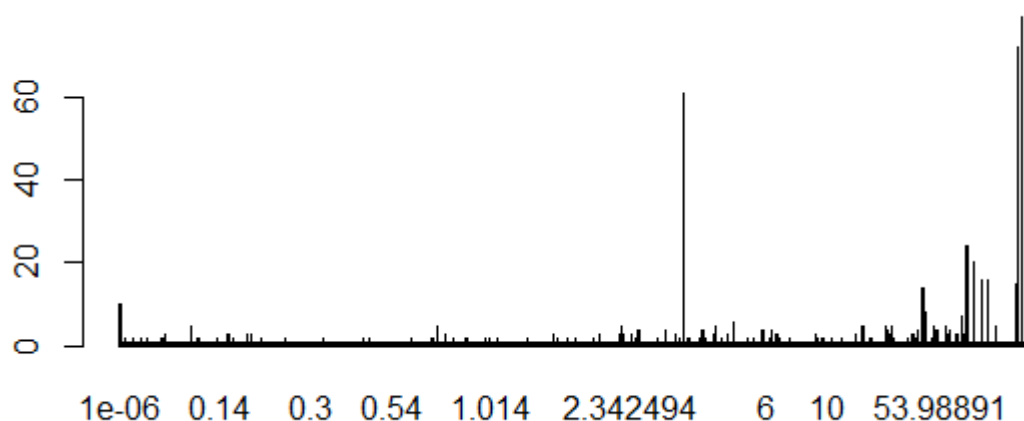
TP_CATEGORIA_ADMINISTRATIVA

Categoria Administrativa

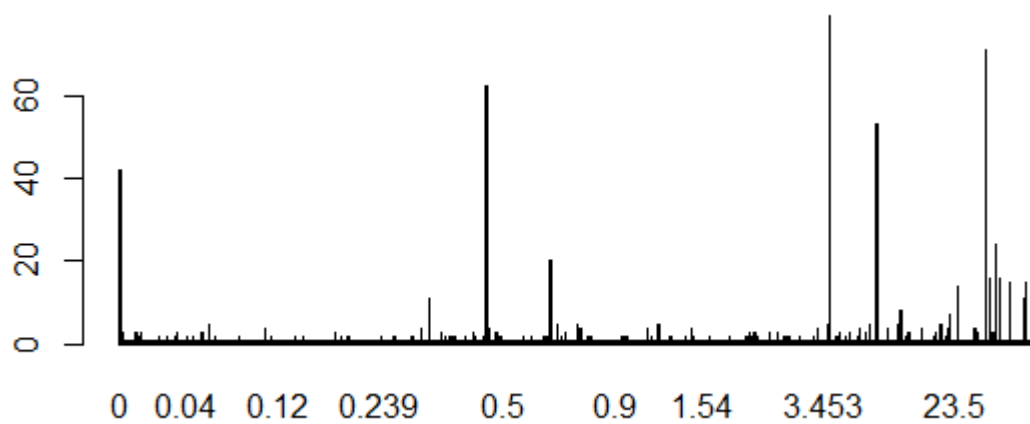


A partir do gráfico de setores, podemos perceber que mais da metade das IESs presentes na nossa base são da categoria 4, ou seja, 'privada com fins lucrativos'. Seguido por 'privada sem fins lucrativos' de categoria 5, públicas em diferentes esferas (números 1, 2 e 3) e uma pequena parcela se enquadrando em 'especiais', representadas pelo número 7. Estas correspondências podem ser verificadas no tópico 2.

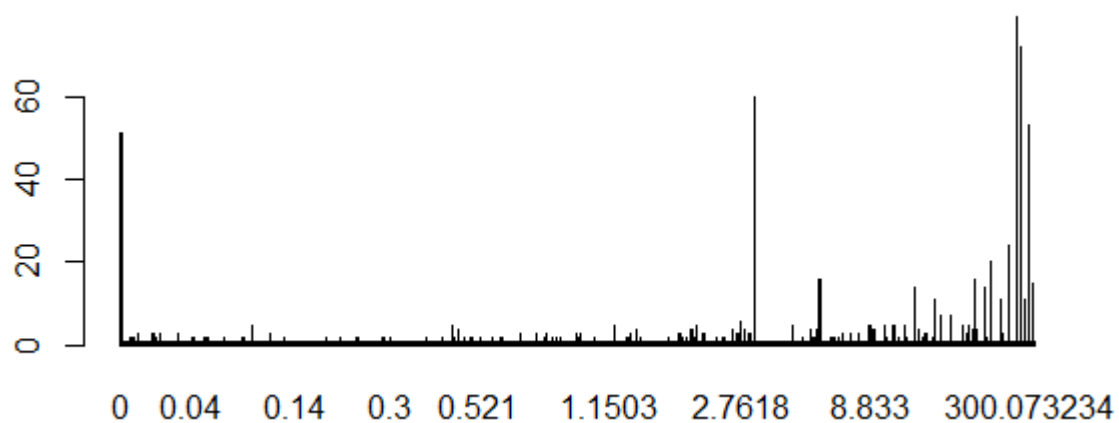
VL_DESPESA_PESSOAL_DOCENTE



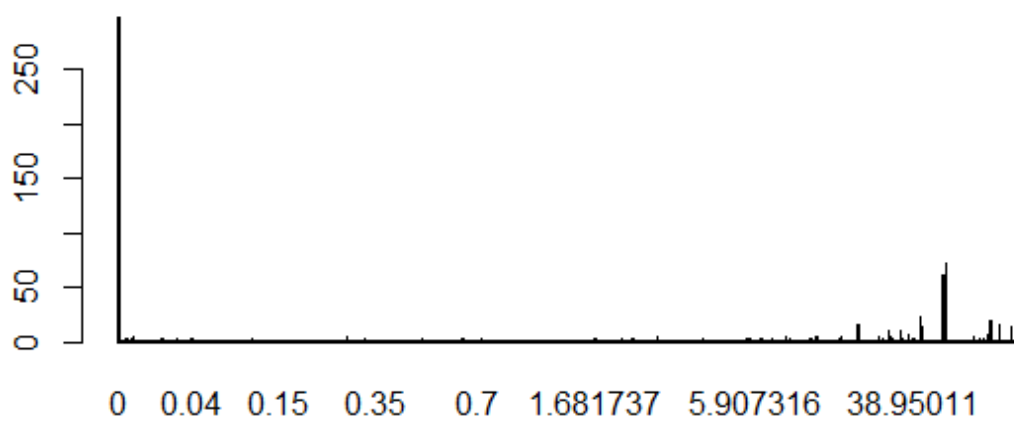
VL_DESPESA_PESSOAL_TECNICO



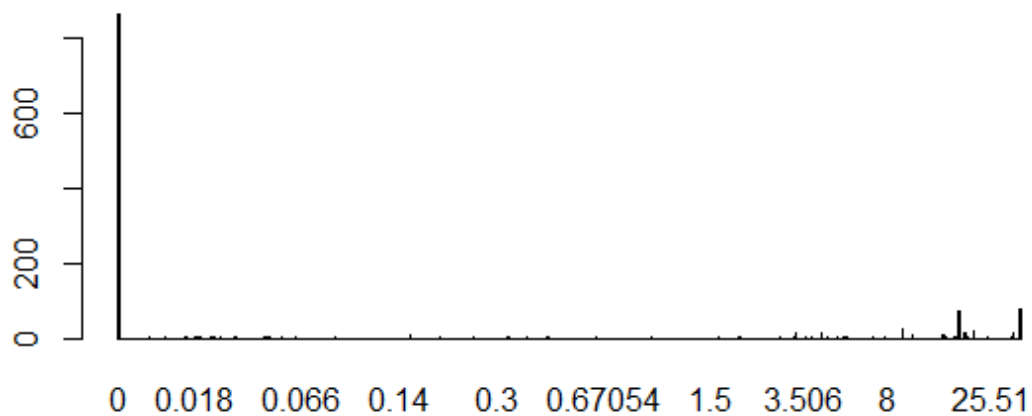
VL_DESPESA_PESSOAL_ENCARGO



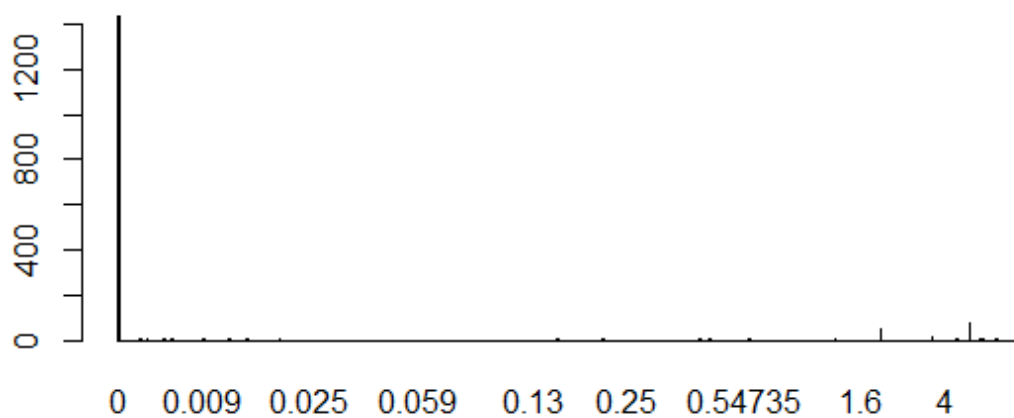
VL_DESPESA_CUSTEIO



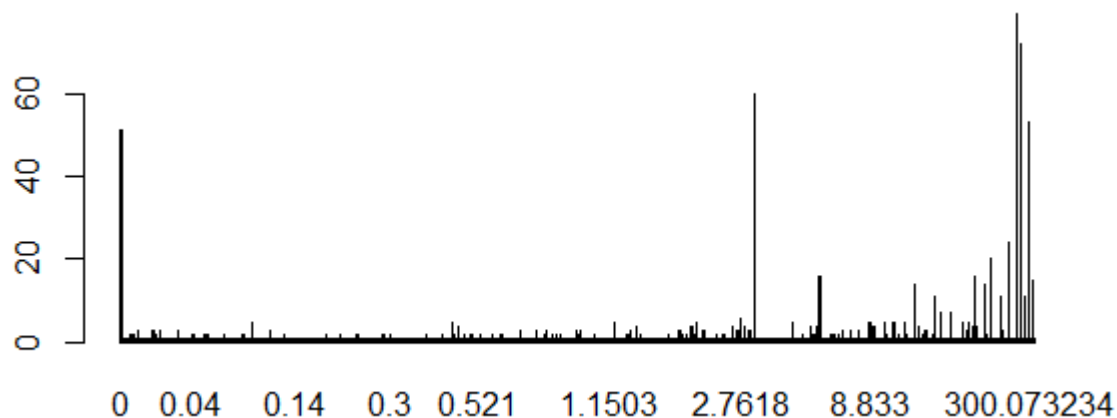
VL_DESPESA_INVESTIMENTO



VL_DESPESA_PESQUISA

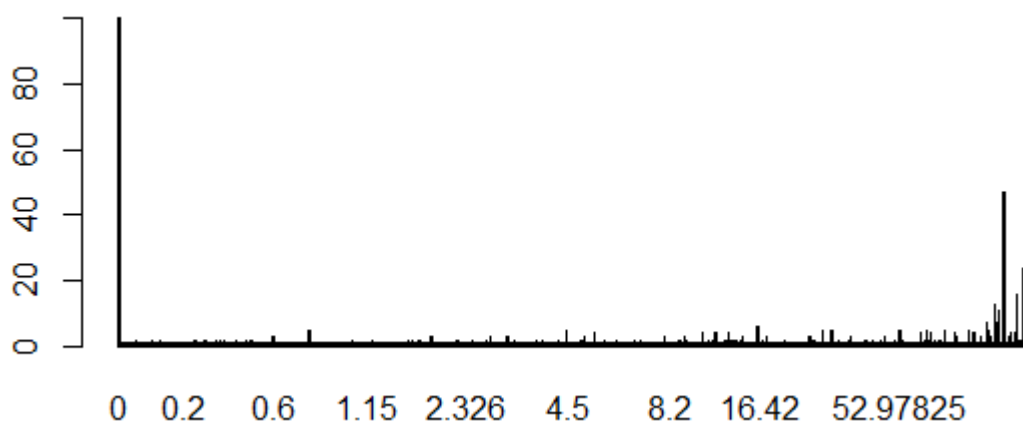


VL_DESPESA_OUTRA



Pelos gráficos de barras, podemos perceber que as despesas com pessoal e aquelas enquadradas em outros gastos costumam ter valores muito parecidos para várias IES, que pode ser visto por algumas barras que se destacam. Enquanto os demais, custeio, investimento e pesquisa costumam variar bastante de acordo com a IES

VL_RECEITA_PROPRIA ($\times 10^6$)



Pelo gráfico podemos perceber que a quantidade de IES com receita até, aproximadamente 53 milhões é bem distribuída, porém, é muito mais provável que a receita própria seja maior que isso, visto o tamanho das barras depois desse valor.

5. Correlação e gráficos de dispersão

Testando as correlações com a receita própria, na mesma ordem da utilizada anteriormente temos:

Docente: $r = 0.4595305$,

Técnico: $r = 0.2543674$,

Encargos: $r = 0.6025158$,

Custeio: $r = 0.8651195$,

Investimentos: $r = 0.7535071$,

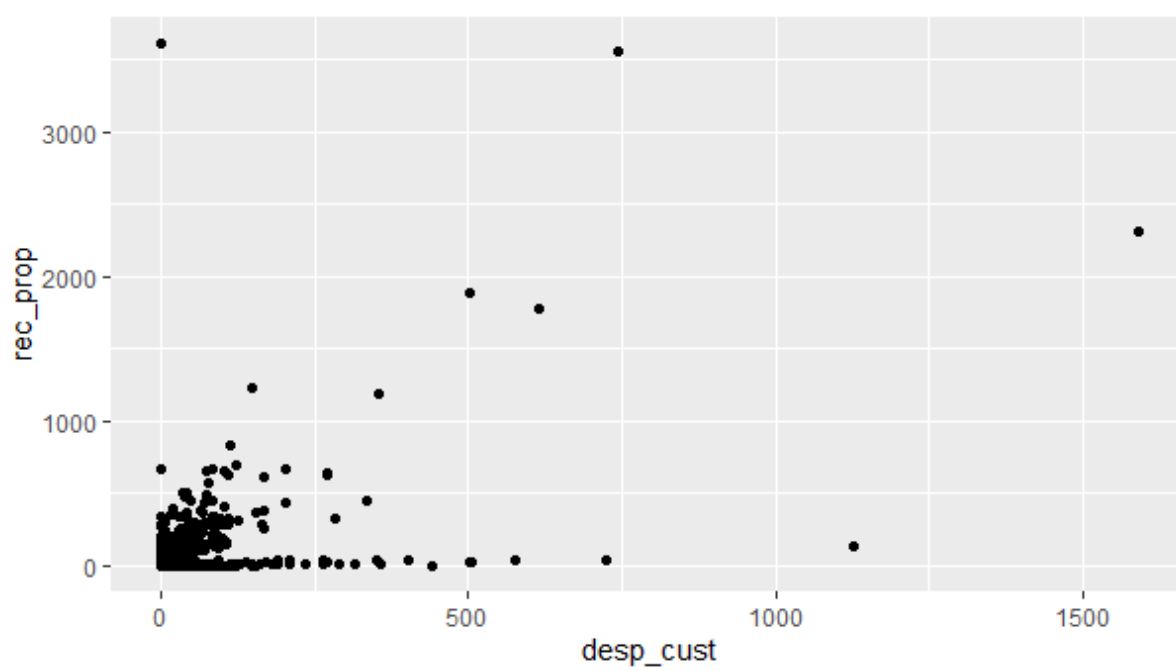
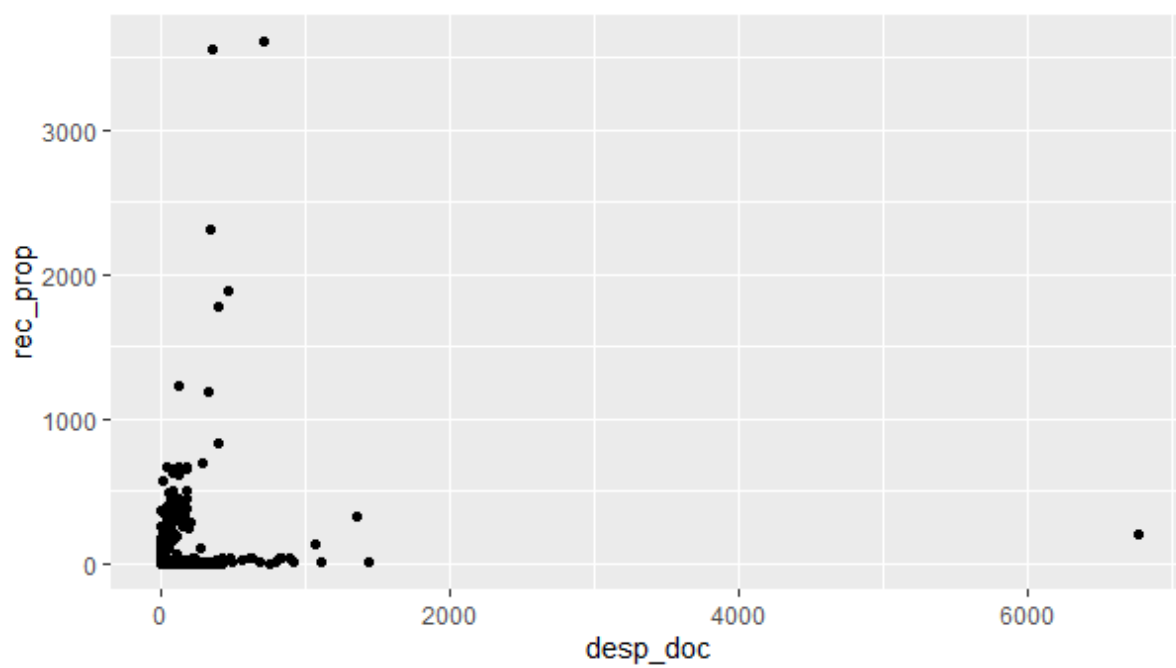
Pesquisa: $r = 0.1353863$,

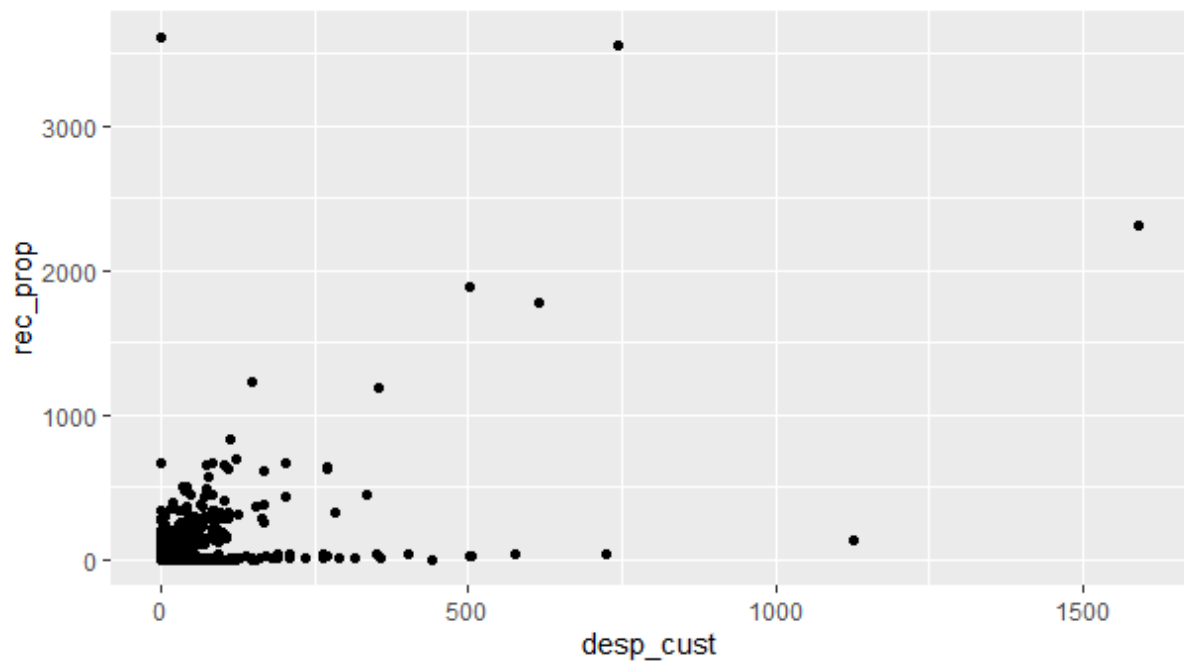
Outros: $r = 0.7870562$.

O que indica que apenas as despesas com docentes, encargos, custeio, investimentos e outros possuem uma correlação de moderada a forte com custeio.

Testando estas cinco restantes entre si, percebemos que os binômios encargos-custeio ($r = 0.731182$) bem como investimento-outros ($r = 0.8809097$) possuem correlação forte. Assim, escolhe-se não utilizar, também, encargos por sua menor correlação com a receita bem como ‘outros’ devido à sua natureza ampla.

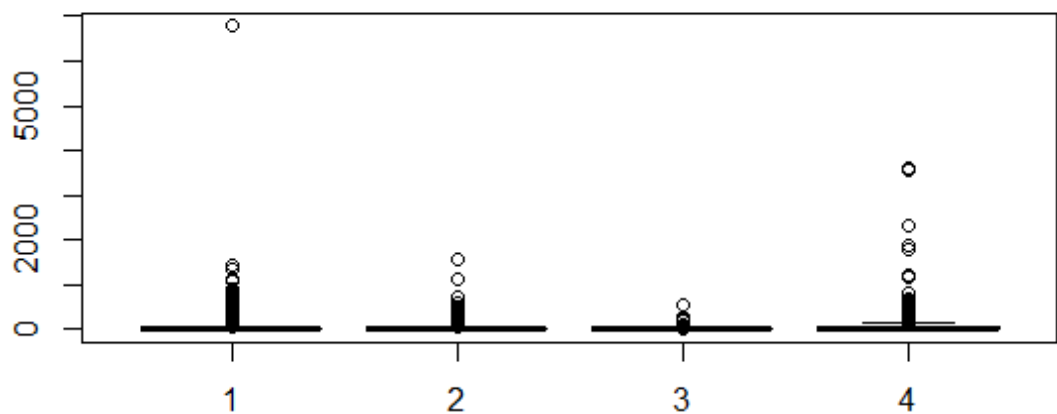
Assim, temos os seguintes gráficos de dispersão para docentes, custeio e investimento x receita, respectivamente:





6. Boxplot

Fazendo o gráfico boxplot para as quatro variáveis quantitativas escolhidas temos:



O que indica muitos outliers e que os demais resultados estão muito próximos uns dos outros.

7. Regressão

Fazendo a regressão com múltipla obtemos o seguinte resultado:

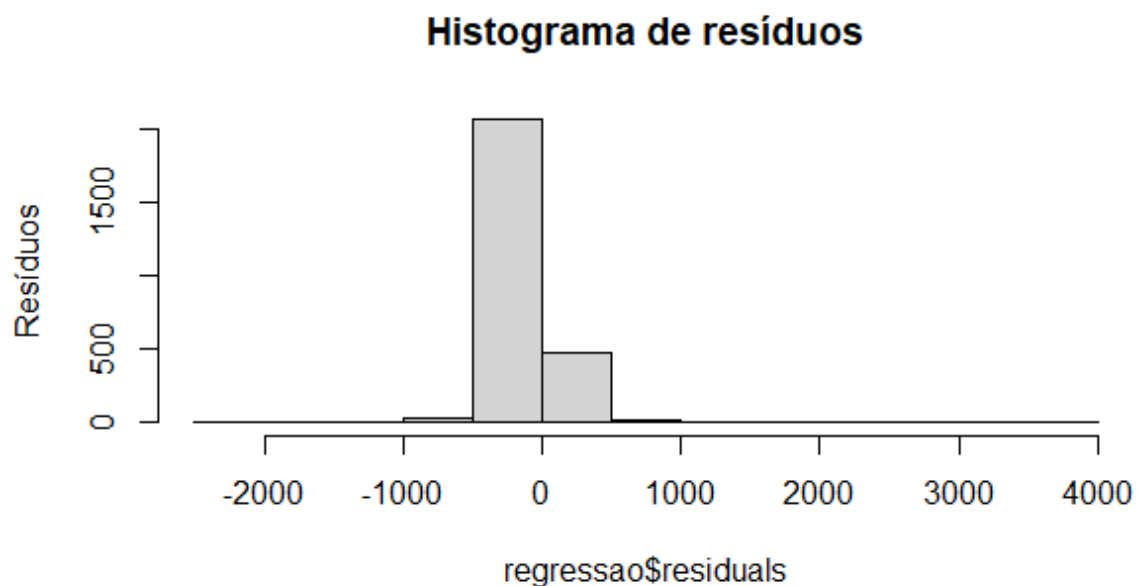
$$R = 27,9725 - 0.0111 * D + 1.8121 * C + 2.7438 * I$$

Onde:

- R = Receita própria
- D = Despesa com docentes
- C = Despesa com custeio
- I = Despesa com investimentos

Observando o resultado obtido constata-se que, talvez D não tenha sido uma boa escolha de variável, já que apesar de ter uma correlação positiva moderada, o valor do seu coeficiente indica uma correlação negativa.

Além disso, os coeficientes indicam que além da receita base de R\$27,97, para cada real despendido com docentes a receita cai R\$0,01. Bem como para cada R\$1,00 destinado a custeio e investimento a receita aumenta em R\$1,81 e R\$2,74, respectivamente. Todos os valores aqui estão divididos por 10^6 com base em seu valor na base de dados original.



8. Monte Carlo

Usando o método de partição Hold Out com a divisão típica de 2/3 para treino e 1/3 para teste temos que usando o modelo 1, 2 e 3 com, respectivamente, todas variáveis, as 3 selecionadas anteriormente e apenas a com maior correlação (custeio):

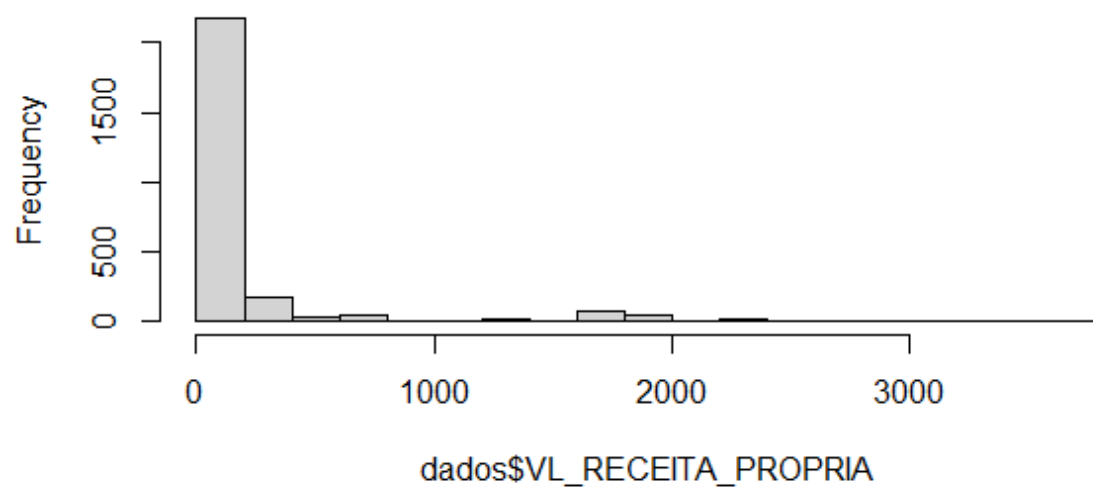
- Modelo 1:
 - $R^2 = 0.8181097$
 - Erro absoluto = 64.98439
 - Erro médio quadrático = 212.5828
- Modelo 2:
 - $R^2 = 0.7987718$
 - Erro absoluto = 87.45439
 - Erro médio quadrático = 223.4335
- Modelo 3:
 - $R^2 = 0.7569394$
 - Erro absoluto = 107.9539
 - Erro médio quadrático = 250.4941

O que nos mostra que quanto menos variáveis usamos, pior fica o modelo como podemos ver pelos 3 indicadores, R^2 mais distante de 1, erro absoluto e erro médio quadrático maior também.

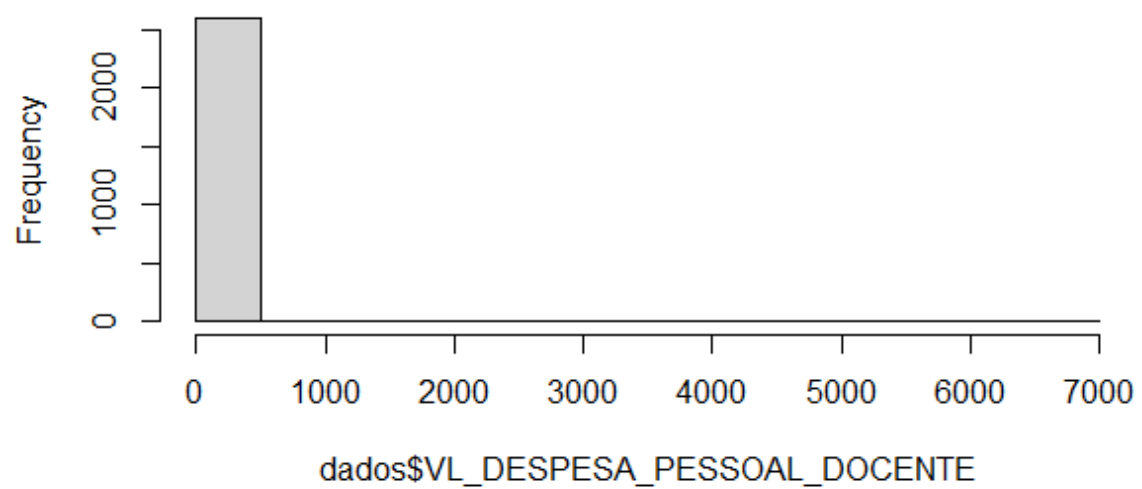
9. Teste de normalidade

Realizando os testes de Shapiro-Wilk e Kolmogorov-Smirnov, obtemos valores de p da ordem de 10^{-16} o que indica que a distribuição não é normal ($< 0,05$). Esse resultado pode ser confirmando pelos histogramas das variáveis.

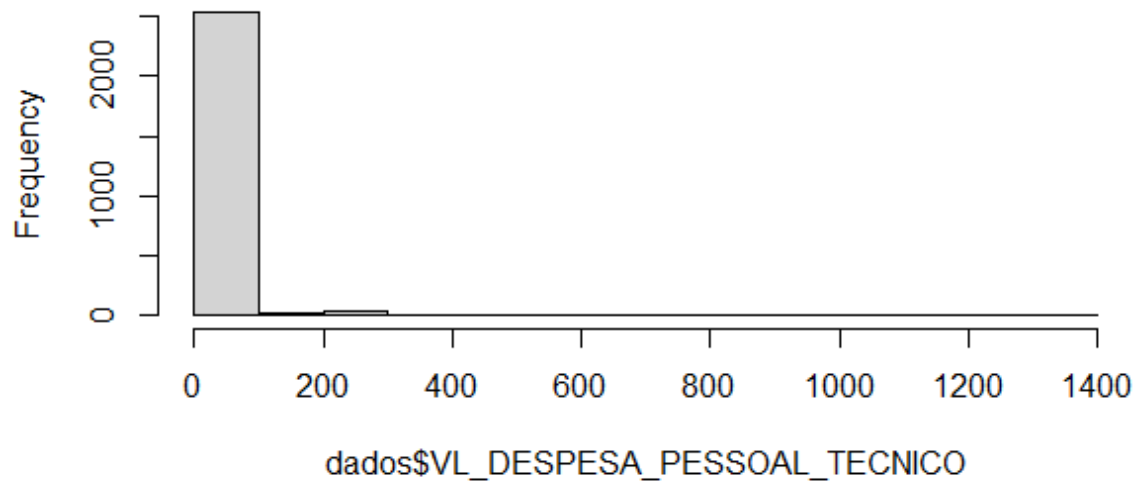
Histogram of dados\$VL_RECEITA_PROPRIA



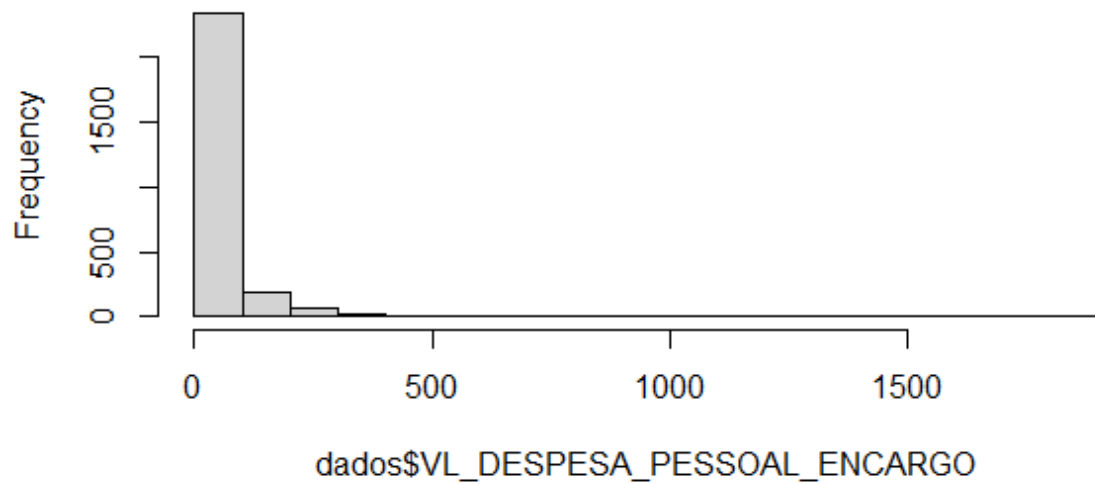
Histogram of dados\$VL_DESPESA_PESSOAL_DOCENTE



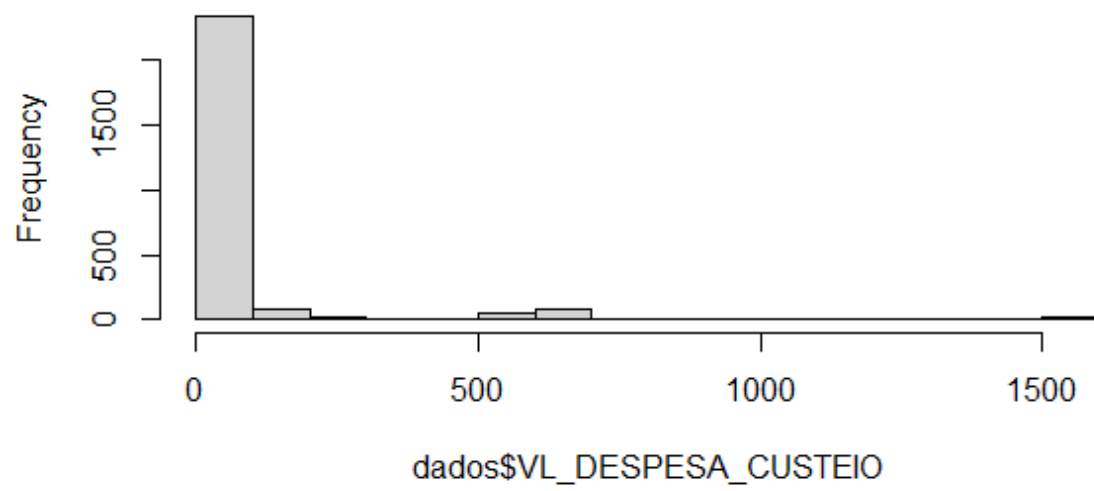
Histogram of dados\$VL_DESPESA_PESSOAL_TECNICO



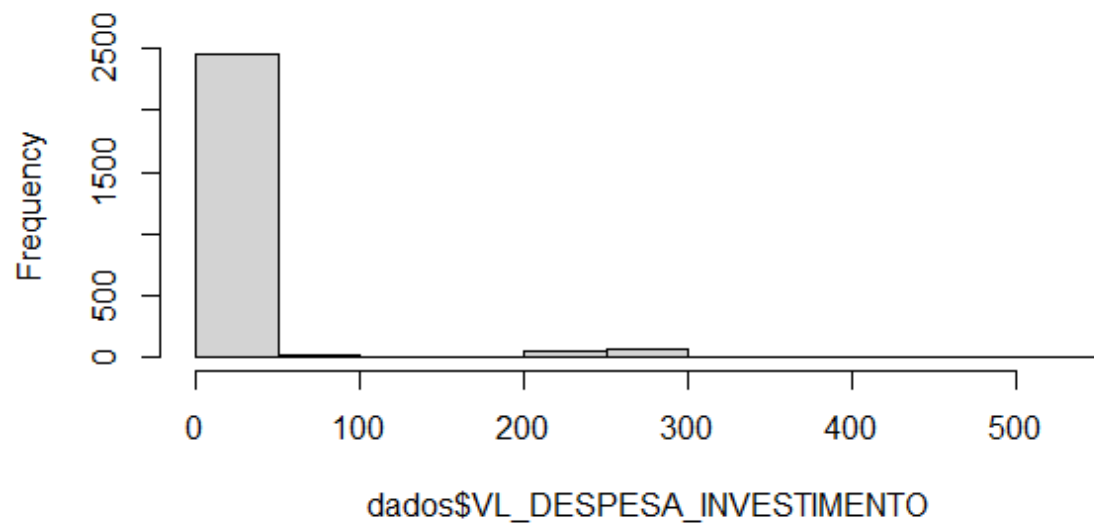
Histogram of dados\$VL_DESPESA_PESSOAL_ENCARGO



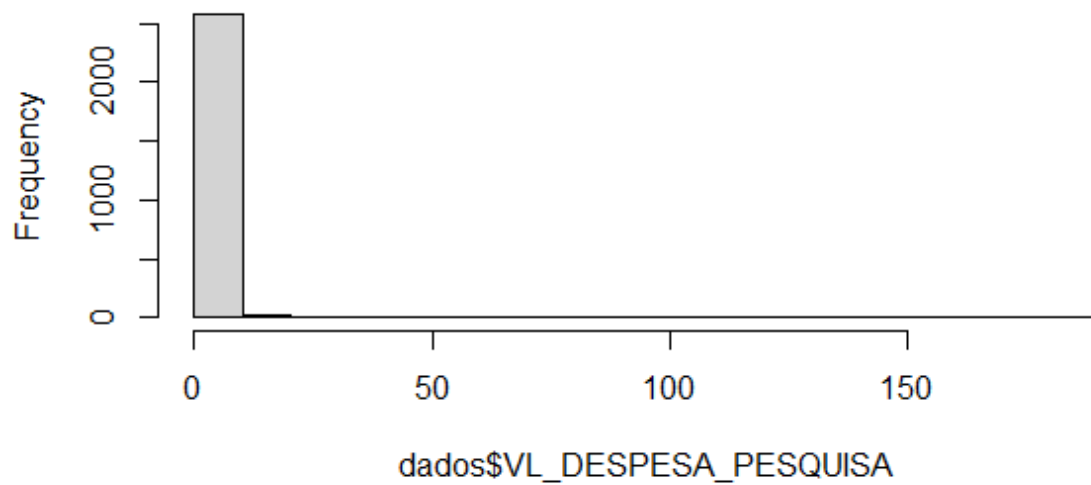
Histogram of dados\$VL_DESPESA_CUSTEIO



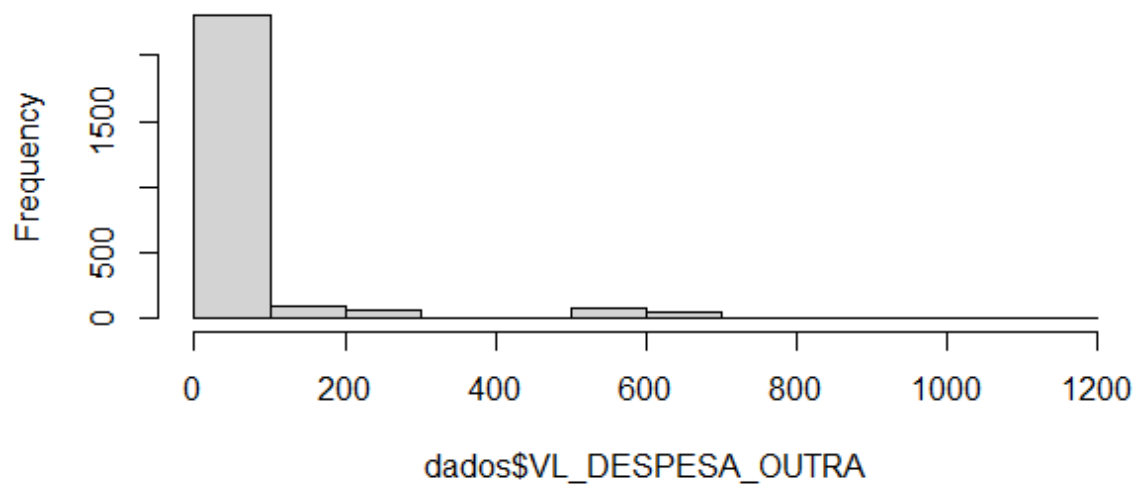
Histogram of dados\$VL_DESPESA_INVESTIMENTO



Histogram of dados\$VL_DESPESA_PESQUISA



Histogram of dados\$VL_DESPESA_OUTRA



10. Teste de hipótese e Anova

Como a distribuição das variáveis não é normal, não faz sentido fazer o teste de hipótese, entretanto, podemos realizar o Anova.

Testando separadamente cada uma contra a categoria administrativa, todos os custos obtiveram o código ‘***’ indicando que existe diferença dos custos entre categorias administrativas com nível de significância de 99,99%.

Para a receita própria, o teste não indica diferença entre as categorias administrativas.

11.Código

As linhas de código utilizadas para chegar a tais resultados estão transcritos aqui e foram utilizados na linguagem R.

```
getwd()
```

```
dados = read.csv2('SUP_IES_2019.CSV', header = TRUE, sep = '|')
```

```
cat_admin = dados$TP_CATEGORIA_ADMINISTRATIVA
```

```
dados$VL_DESPESA_PESSOAL_DOCENTE =  
as.numeric(dados$VL_DESPESA_PESSOAL_DOCENTE)/1e6
```

```
dados$VL_DESPESA_PESSOAL_TECNICO =  
as.numeric(dados$VL_DESPESA_PESSOAL_TECNICO)/1e6
```

```
dados$VL_DESPESA_PESSOAL_ENCARGO =  
as.numeric(dados$VL_DESPESA_PESSOAL_ENCARGO)/1e6
```

```
dados$VL_DESPESA_CUSTEIO = as.numeric(dados$VL_DESPESA_CUSTEIO)/1e6
```

```
dados$VL_DESPESA_INVESTIMENTO =  
as.numeric(dados$VL_DESPESA_INVESTIMENTO)/1e6
```

```
dados$VL_DESPESA_PESQUISA = as.numeric(dados$VL_DESPESA_PESQUISA)/1e6
```

```
dados$VL_DESPESA_OUTRA = as.numeric(dados$VL_DESPESA_OUTRA)/1e6
```

```
rec_prop = as.numeric(dados$VL_RECEITA_PROPRIA)/1e6
```

```
desp_doc = dados$VL_DESPESA_PESSOAL_DOCENTE
```

```
desp_tec = dados$VL_DESPESA_PESSOAL_TECNICO
```

```
desp_enc = dados$VL_DESPESA_PESSOAL_ENCARGO
```

```
desp_cust = dados$VL_DESPESA_CUSTEIO
```

```
desp_invest = dados$VL_DESPESA_INVESTIMENTO
```

```
desp_pesq = dados$VL_DESPESA_PESQUISA
```

```
desp_out = dados$VL_DESPESA_OUTRA
```

```
rec_prop = dados$VL_RECEITA_PROPRIA
```

```
summary(desp_doc)
```

```
sd(desp_doc)
```

```
var(desp_doc)
```

```
summary(desp_tec)
```

```
sd(desp_tec)
```

```
var(desp_tec)
```

```
summary(desp_enc)
```

```
sd(desp_enc)
```

```
var(desp_enc)
```

```
summary(desp_cust)
```

```
sd(desp_cust)
```

```
var(desp_cust)
```

```
summary(desp_invest)
```

```
sd(desp_invest)
```

```
var(desp_invest)
```

```
summary(desp_pesq)
```

```
sd(desp_pesq)
```

```
var(desp_pesq)
```

```
summary(desp_out)
```

```
sd(desp_out)
```

```
var(desp_out)
```

```
summary(rec_prop)
```

```
sd(rec_prop)
```

```
var(rec_prop)
```

```
Values = table(dados$TP_CATEGORIA_ADMINISTRATIVA)
```

```
labels =paste(names(Values))
```

```
pie(Values, labels = labels, main = 'Categoria Administrativa')
```

```
barplot(table(desp_doc))
```

```
barplot(table(desp_tec))
```

```
barplot(table(desp_enc))
```

```
barplot(table(desp_cust))
```

```
barplot(table(desp_invest))
```

```
barplot(table(desp_pesq))
```

```
barplot(table(desp_out))
```

```
barplot(table(rec_prop))
```

```
cor(desp_doc,rec_prop)
```

```
cor(desp_tec,rec_prop)
```

```
cor(desp_enc,rec_prop)
```

```
cor(desp_cust,rec_prop)
```

```
cor(desp_invest,rec_prop)
```

```
cor(desp_pesq,rec_prop)
```

```
cor(desp_out,rec_prop)
```

```
cor(desp_doc,desp_enc)
```

```
cor(desp_doc,desp_cust)
```

```
cor(desp_doc,desp_invest)
```

```
cor(desp_doc,desp_out)
```

```
cor(desp_enc,desp_cust)
```

```
cor(desp_enc,desp_invest)
```

```
cor(desp_enc,desp_out)
```

```
cor(desp_cust,desp_invest)
```

```
cor(desp_cust, desp_out)
```

```
cor(desp_invest, desp_out)
```

```
ggplot(dados, aes(x = as.numeric(VL_DESPESA_PESSOAL_DOCENTE)/1e6, y =  
as.numeric(VL_RECEITA_PROPRIA)/1e6)) +
```

```
  geom_point()
```

```
ggplot(dados, aes(x = desp_cust, y = rec_prop)) +
```

```
  geom_point()
```

```
ggplot(dados, aes(x = desp_invest, y = rec_prop)) +
```

```
  geom_point()
```

```
boxplot(desp_doc, desp_cust, desp_invest, rec_prop)
```

```
boxplot(desp_doc[desp_doc != 0], desp_cust[desp_cust != 0], desp_invest[desp_invest !=  
0], rec_prop[rec_prop != 0])
```

```
regressao = lm(rec_prop~desp_doc+desp_cust+desp_invest)
```

```
as.numeric(dados$VL_RECEITA_PROPRIA) - regressao$fitted.values * 1e6
```

```
hist(regressao$residuals, main = 'Histograma de resíduos', ylab = 'Resíduos')
```

```
boxplot(regressao$residuals)
```

```
#MC
```

```
ModeloAjustado2 = 0
```

```
ErroAbsoluto2 = 0
```


ErroMedioQuadratico2 = 0

#inicio do LOOP de MONTE CARLO

for (i in 1:50){

ind = createDataPartition(dados\$VL_RECEITA_PROPRIA,p = 2/3, list = FALSE)

train.data = dados[ind,]

test.data = dados[-ind,]

modelo2 =lm(VL_RECEITA_PROPRIA ~ VL_DESPESA_OUTRA +
VL_DESPESA_PESQUISA + VL_DESPESA_PESSOAL_ENCARGO +
VL_DESPESA_PESSOAL_TECNICO + VL_DESPESA_PESSOAL_DOCENTE +
VL_DESPESA_CUSTEIO + VL_DESPESA_INVESTIMENTO, data = train.data)

ValoresPreditos2 = predict(modelo2,newdata=data.frame(test.data))

ModeloAjustado2 [i] = R2(ValoresPreditos2, test.data\$VL_RECEITA_PROPRIA)

ErroAbsoluto2 [i] = MAE(ValoresPreditos2, test.data\$VL_RECEITA_PROPRIA)

ErroMedioQuadratico2 [i] = RMSE(ValoresPreditos2,
test.data\$VL_RECEITA_PROPRIA)

}

mean(ModeloAjustado2)

mean(ErroAbsoluto2)

mean(ErroMedioQuadratico2)

ModeloAjustado3 = 0

ErroAbsoluto3 = 0

ErroMedioQuadratico3 = 0

#inicio do LOOP de MONTE CARLO

```

for (i in 1:50){

  ind = createDataPartition(dados$VL_RECEITA_PROPRIA,p = 2/3, list = FALSE)

  train.data = dados[ind, ]

  test.data = dados[-ind, ]

  modelo3 =lm(VL_RECEITA_PROPRIA ~ VL_DESPESA_CUSTEIO , data = train.data)

  ValoresPreditos3 = predict(modelo3,newdata=data.frame(test.data))

  ModeloAjustado3 [i] = R2(ValoresPreditos3, test.data$VL_RECEITA_PROPRIA)

  ErroAbsoluto3 [i] = MAE(ValoresPreditos3, test.data$VL_RECEITA_PROPRIA)

  ErroMedioQuadratico3          [i]          =          RMSE(ValoresPreditos3,
test.data$VL_RECEITA_PROPRIA)

}

mean(ModeloAjustado3)

mean(ErroAbsoluto3)

mean(ErroMedioQuadratico3)


shapiro.test(dados$VL_RECEITA_PROPRIA)

ks.test(dados$VL_RECEITA_PROPRIA,'pnorm',mean(dados$VL_RECEITA_PROPRIA
),sd(dados$VL_RECEITA_PROPRIA))


shapiro.test(dados$VL_DESPESA_PESSOAL_DOCENTE)

ks.test(dados$VL_DESPESA_PESSOAL_DOCENTE,'pnorm',mean(dados$VL_DESPESA_PESSOAL_DOCENTE),sd(dados$VL_DESPESA_PESSOAL_DOCENTE))


shapiro.test(dados$VL_DESPESA_PESSOAL_TECNICO)

ks.test(dados$VL_DESPESA_PESSOAL_TECNICO,'pnorm',mean(dados$VL_DESPESA_PESSOAL_TECNICO),sd(dados$VL_DESPESA_PESSOAL_TECNICO))

```

```
shapiro.test(dados$VL_DESPESA_PESSOAL_ENCARGO)
```

```
ks.test(dados$VL_DESPESA_PESSOAL_ENCARGO,'pnorm',mean(dados$VL_DESPESA_PESSOAL_ENCARGO),sd(dados$VL_DESPESA_PESSOAL_ENCARGO))
```

```
shapiro.test(dados$VL_DESPESA_CUSTEIO)
```

```
ks.test(dados$VL_DESPESA_CUSTEIO,'pnorm',mean(dados$VL_DESPESA_CUSTEIO),sd(dados$VL_DESPESA_CUSTEIO))
```

```
shapiro.test(dados$VL_DESPESA_INVESTIMENTO)
```

```
ks.test(dados$VL_DESPESA_INVESTIMENTO,'pnorm',mean(dados$VL_DESPESA_INVESTIMENTO),sd(dados$VL_DESPESA_INVESTIMENTO))
```

```
shapiro.test(dados$VL_DESPESA_PESQUISA)
```

```
ks.test(dados$VL_DESPESA_PESQUISA,'pnorm',mean(dados$VL_DESPESA_PESQUISA),sd(dados$VL_DESPESA_PESQUISA))
```

```
shapiro.test(dados$VL_DESPESA_OUTRA)
```

```
ks.test(dados$VL_DESPESA_OUTRA,'pnorm',mean(dados$VL_DESPESA_OUTRA),sd(dados$VL_DESPESA_OUTRA))
```

```
hist(dados$VL_RECEITA_PROPRIA)
```

```
hist(dados$VL_DESPESA_PESSOAL_DOCENTE)
```

```
hist(dados$VL_DESPESA_PESSOAL_TECNICO)
```

```
hist(dados$VL_DESPESA_PESSOAL_ENCARGO)
```

```
hist(dados$VL_DESPESA_CUSTEIO)
```

```
hist(dados$VL_DESPESA_INVESTIMENTO)
```

```
hist(dados$VL_DESPESA_PESQUISA)
```

```
hist(dados$VL_DESPESA_OUTRA)
```

```
t.test(dados$VL_DESPESA_PESSOAL_DOCENTE,dados$VL_DESPESA_PESSOAL_
TECNICO, alternative = c("two.sided"))
```

```
boxplot(dados$VL_RECEITA_PROPRIA ~
dados$TP_CATEGORIA_ADMINISTRATIVA)
```

```
boxplot(dados$VL_DESPESA_PESSOAL_DOCENTE ~
dados$TP_CATEGORIA_ADMINISTRATIVA)
```

```
boxplot(dados$VL_DESPESA_PESSOAL_TECNICO ~
dados$TP_CATEGORIA_ADMINISTRATIVA)
```

```
boxplot(dados$VL_DESPESA_PESSOAL_ENCARGO ~
dados$TP_CATEGORIA_ADMINISTRATIVA)
```

```
boxplot(dados$VL_DESPESA_CUSTEIO ~
dados$TP_CATEGORIA_ADMINISTRATIVA)
```

```
boxplot(dados$VL_DESPESA_INVESTIMENTO ~
dados$TP_CATEGORIA_ADMINISTRATIVA)
```

```
boxplot(dados$VL_DESPESA_PESQUISA ~
dados$TP_CATEGORIA_ADMINISTRATIVA)
```

```
boxplot(dados$VL_DESPESA_OUTRA ~
dados$TP_CATEGORIA_ADMINISTRATIVA)
```

```
summary(aov(dados$VL_RECEITA_PROPRIA ~
dados$TP_CATEGORIA_ADMINISTRATIVA))
```

summary(aov(dados\$VL_DESPESA_PESSOAL_DOCENTE dados\$TP_CATEGORIA_ADMINISTRATIVA))	~
summary(aov(dados\$VL_DESPESA_PESSOAL_TECNICO dados\$TP_CATEGORIA_ADMINISTRATIVA))	~
summary(aov(dados\$VL_DESPESA_PESSOAL_ENCARGO dados\$TP_CATEGORIA_ADMINISTRATIVA))	~
summary(aov(dados\$VL_DESPESA_CUSTEIO dados\$TP_CATEGORIA_ADMINISTRATIVA))	~
summary(aov(dados\$VL_DESPESA_INVESTIMENTO dados\$TP_CATEGORIA_ADMINISTRATIVA))	~
summary(aov(dados\$VL_DESPESA_PESQUISA dados\$TP_CATEGORIA_ADMINISTRATIVA))	~
summary(aov(dados\$VL_DESPESA_OUTRA dados\$TP_CATEGORIA_ADMINISTRATIVA))	~