

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Trabalho de Estatística

Celso Junio Simões de Oliveira Santos

Felipe Matheus Guimarães dos Santos

Leandro Marques Venceslau de Souza

Lucas Marinho Moraes dos Reis

Matheus Ferreira Coelho

Willian Douglas Brito Braga

• INTRODUÇÃO

Um banco de dados foi formulado a partir de um questionário aplicado na turma TN do curso de Estatística e Probabilidades, ministrado pela professora Lourdes Coral Contreras Montenegro em 2019/02.

Este trabalho tem como objetivo estabelecer e analisar relações entre as diferentes variáveis presentes nessa database, de forma a extrair informações relevantes, tudo isso pautado na teoria estatística aprendida em sala de aula.

• METODOLOGIA

Os dados foram estudados com o auxílio do software - “R: The Project for Statistical Computing”. O código utilizado para criar os gráficos, tabelas assim como extrair as informações do dataset estão disponíveis no apêndice.

Quanto às análises : exibimos um recurso gráfico informacional cuja forma de exposição depende da natureza da variável observada – variáveis quantitativas contínuas, por exemplo, têm como preferência a utilização de histogramas.

Na hora de estabelecer inter-relações entre as diferentes variáveis poderão ser utilizadas diferentes medidas quando aplicável (ambas as variáveis são de natureza quantitativa). São elas :

- Medidas de Tendência Central (Média, Mediana, Moda).
- Medidas de Variabilidade (Amplitude entre terceiro e segundo quartil, Desvio Padrão).

Caso realizadas comparações quanto aos níveis de dispersão de variáveis com medidas e/ou tipos diferentes (i.e altura e sexo) será utilizado o Coeficiente de Variação.

- RESULTADOS

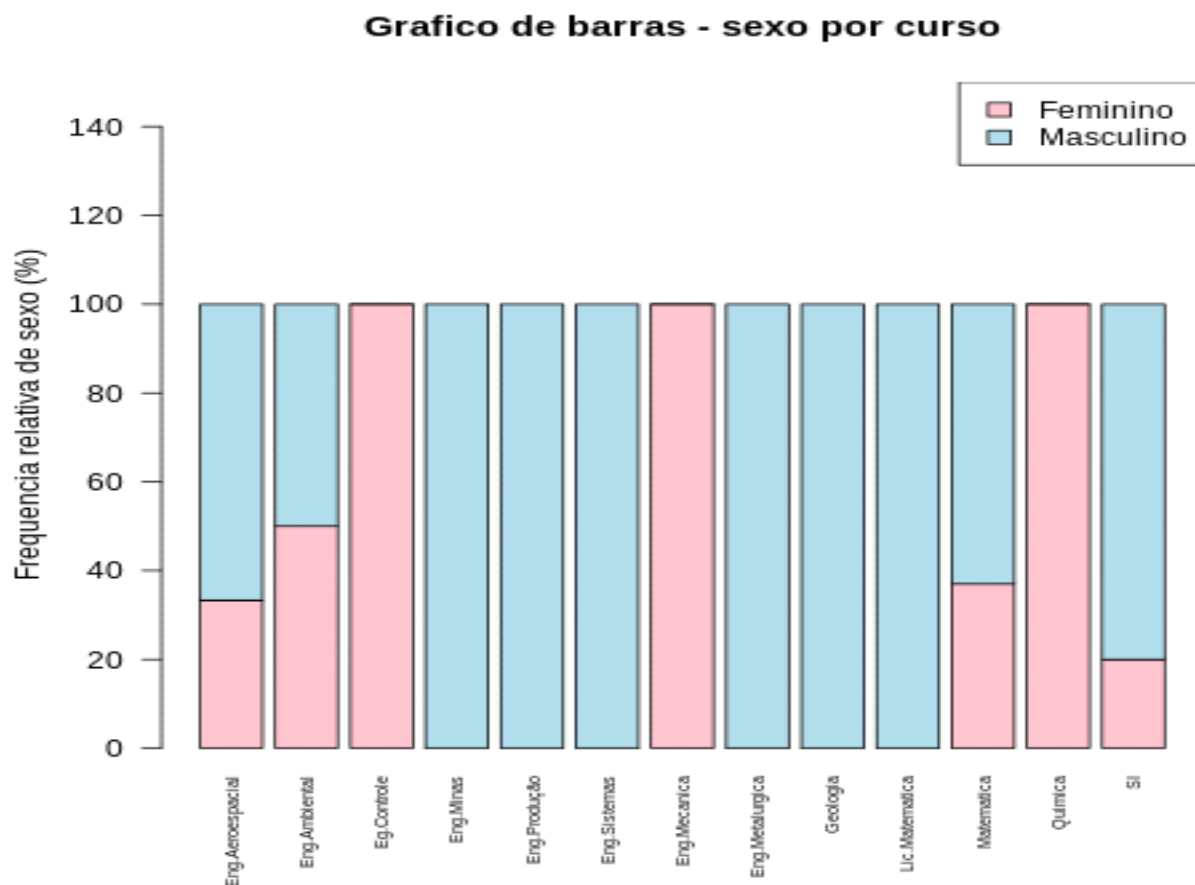


Figura 1: Gráfico de Barras – Sexo por Curso

O gráfico acima mostra que a frequência relativa de homens é maior que a de mulheres para a maioria dos casos observados (9 em 13), com exceção para os cursos – Engenharias Ambiental, Controle de Automação (EC) e Mecânica, assim como Química. Outro aspecto interessante é que existe um conjunto de 8 cursos frequentados apenas por um sexo, sendo 6 deles somente por homens e dois deles (EC, Química) apenas por mulheres. Por fim, há um único curso (Engenharia Ambiental) com uma distribuição quase simétrica entre homens e mulheres, logo percebe-se, dentro dessa base de dados que representa uma amostra dos graduandos da UFMG, uma clara tendência a extremos quanto à distribuição sexo/curso aliado a uma prevalência percentual de indivíduos do sexo masculino.

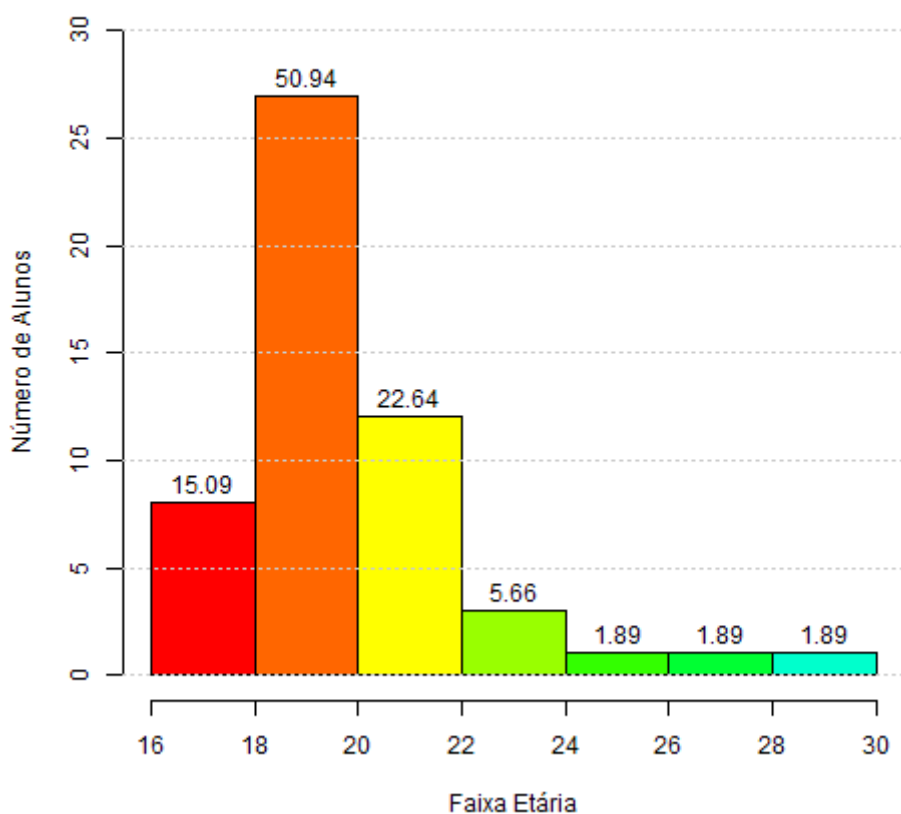


Figura 2: Histograma – Número de Alunos por Faixa Etária

Média Aritmética: 20,2 Anos Mediana: 20 Anos Moda: 19 Anos

A partir da análise dos dados da Figura 2: Histograma - Número de Alunos X Faixa Etária, nota-se que grande parte (88,67%) dos alunos da turma TN do curso de Estatística e Probabilidades da UFMG tem entre 16 e 22 anos, sendo que a maioria (50.94%) dos alunos tem entre 18 e 20 anos (Vale-se observar que essa distribuição é bem próxima da Mediana (20 Anos) e da Moda (19 Anos) da Idade). Além disso, as faixas dos 24 aos 26, 26 aos 28, e 28 aos 30 anos tem o menor número de alunos e ocorrem em mesma porcentagem de alunos (1.89%) cada uma. Afinal, dado que, segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), a idade que define um jovem brasileiro é de 15 aos 24 anos, pode-se afirmar que a turma TN citada é composta por uma população majoritariamente de jovens.

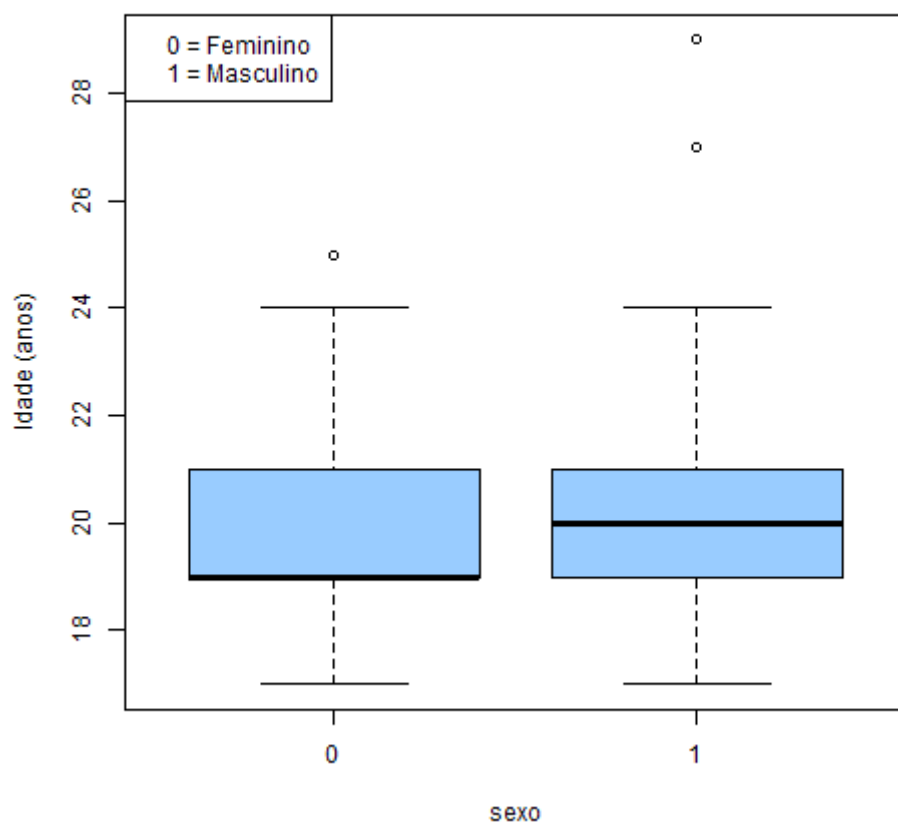


Figura 3: Boxplot – Idade por Sexo

Ao analisar esse Boxplot observa-se que a idade por sexo dos alunos da turma TN do curso de Estatística e Probabilidades da UFMG tem aproximadamente a mesma variância, visto que as duas “caixas” dos boxplots têm praticamente a mesma altura. Apesar disso, as mulheres possuem mediana de idade igual a 19 e os homens possuem mediana de idade igual a 20. É importante denotar que a mediana para as mulheres está próxima ao segundo quartil, o que leva os dados à apresentarem assimetria positiva. No que tange aos homens, no entanto, temos dados com característica simétrica devido à posição central da mediana entre o terceiro e segundo quartis. A moda de idade de homens e mulheres é 19. Nota-se que existe um outlier no boxplot das mulheres, que é uma aluna com 25 anos de idade, e dois outliers no boxplot dos homens, que são um aluno com 27 anos e outro com 28. O gráfico nos auxilia a perceber que a turma TN não tem muita diferença de idade entre os alunos.

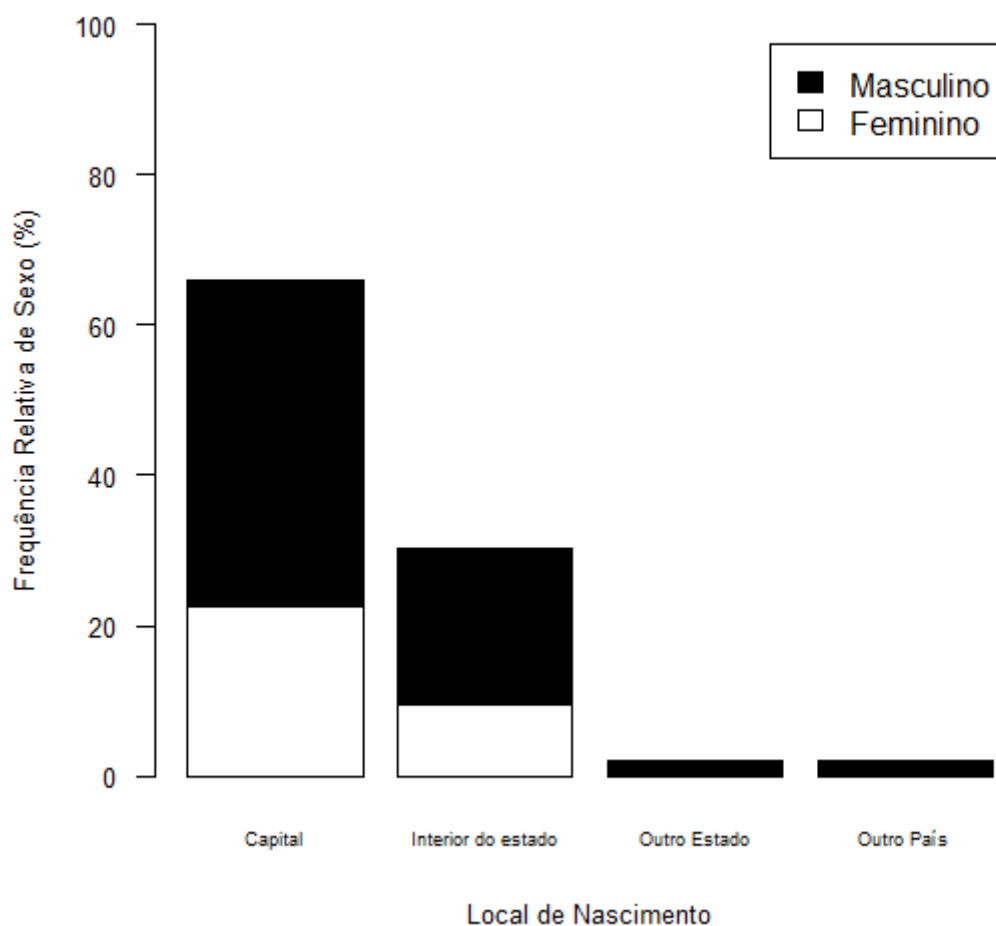


Figura 4: Gráfico de Barras – Sexo por Local de Nascimento

A partir da análise dos dados representados pelo gráfico acima observa-se que mais de 60% dos alunos entrevistados nasceram na capital - Belo Horizonte, sendo que dessa porcentagem cerca de 22% são mulheres e 43% são homens. Os alunos nascidos no interior do estado representam mais 30% da turma, sendo 10% mulheres 20% homens. Como esperado, os alunos que nasceram em outro estado assim como aqueles que nasceram em outro país representam uma porcentagem muito baixa do total, além de que todos são homens. O gráfico nos auxilia a perceber o seguinte - tanto nos alunos da turma nascidos na capital quanto nos nascidos no interior do estado os homens são a maioria, e não possuímos nenhuma mulher que nasceu em outro estado ou outro país no banco de dados estudado.

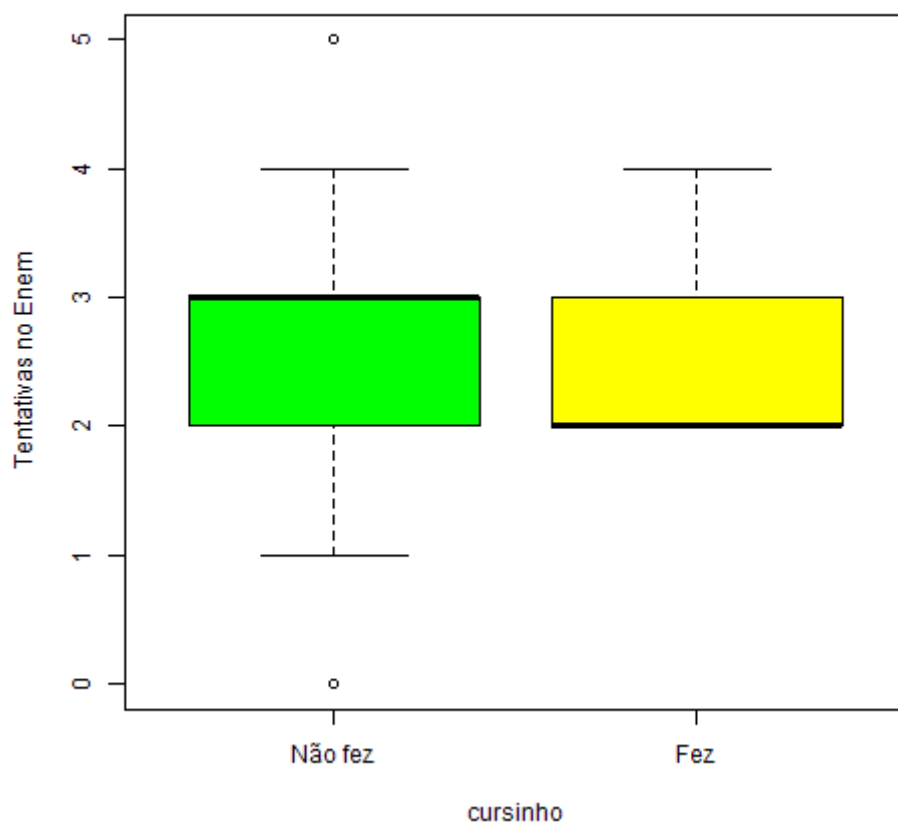


Figura 5: Boxplot - Tentativas ENEM

Analisando o Boxplot acima é possível perceber que a posição e a altura das caixas é semelhante e, assim, ambas possuem uma variabilidade próxima. Ademais, ambas caixas são assimétricas, entretanto, a concentração de tentativas daqueles que não fizeram Cursinho se encontra em 3, enquanto a concentração daqueles que fizeram se encontra em 2. Logo, observa-se que aqueles alunos que fizeram cursinho conseguiram alcançar a nota desejada com menos tentativas em relação aos que não fizeram cursinho; mesmo que essa diferença seja mínima, isso nos mostra que o cursinho tem um impacto positivo no processo seletivo do Enem. Há ainda, a presença de dois Outliers dentre os alunos que não fizeram cursinho: um no extremo superior com 5 tentativas e outro, peculiar, no extremo inferior com 0 tentativas. Essa peculiaridade está no fato de que tal aluno não fez nenhuma tentativa no Enem para entrar no curso, logo é clara a possibilidade de termos um aluno transferido ou até mesmo de intercâmbio em nossa turma, o que é algo incomum.

Principais contribuições na escolha do curso por Sexo				
<i>Contribui</i>	<i>Marcações Femininas</i>	<i>Frequência Relativa(F)</i>	<i>Marcações Masculinas</i>	<i>Frequência Relativa(M)</i>
Salário	0	0,00%	1	2,78%
Mercado de Trabalho	0	0,00%	4	11,11%
Vocação	14	82,35%	19	52,78%
Estabilidade	0	0,00%	0	0,00%
Interesse pela área de pesquisa	0	0,00%	0	0,00%
Salário, Mercado de Trabalho	1	5,88%	5	13,89%
Salário, Vocação	1	5,88%	0	0,00%
Mercado de Trabalho, Vocação	0	0,00%	2	5,55%
Vocação, Interesse pela área de pesquisa	0	0,00%	1	2,78%
Salário, Mercado de Trabalho, Vocação	1	5,88%	3	8,33%
Salário, Vocação, Estabilidade	0	0,00%	1	2,78%
Total:	17	100%	36	100%

Figura 6: Tabela - Principais Contribuições na Escolha do Curso por Sexo

Na Tabela de Principais contribuições na escolha do curso por sexo podemos observar que a maioria dos alunos, tanto do sexo feminino quanto do masculino, tiveram como principal fator contribuinte a Vocação na hora de escolher o curso. Além disso, somente 3 alunas escolheram outros fatores conjuntos que foram: Salário, Mercado de trabalho; Salário, Vocação; e Salário, Mercado de trabalho, Vocação. Ou seja, Todas as Alunas, por exceção de uma, assinalaram "Vocação" como fator contribuinte. Já os alunos do sexo masculino, apesar de possuírem também a vocação como um fator determinante, apresentaram um quadro de opções um pouco mais distribuído quando comparado com as mulheres, marcando mais fatores conjuntos e tendo o salário, assim como o mercado de trabalho, como fatores significativos para sua escolha. Por fim, reiteramos que a Vocação é o fator mais significativo, o que pode indicar que os alunos da turma TN possuem um personalidade e opinião definida quanto ao que querem em sua maioria, tendo poucos contribuintes externos em sua decisão de carreira profissional.

Tabela de Frequências para Influência na escolha do curso		
Influência	Frequência Absoluta	Frequência Relativa
Parentes	3	5.66%
Professor	12	22.54%
Pesquisa na Mídia	3	5.66%
Amigos	0	0,00%
Orientação Vocacional	5	9.43%
Psicólogo	1	1.89%
Namorado	0	0,00%
Mercado aquecido	0	0,00%
Gosto Pessoal	1	1.89%
Parentes, Amigos	1	1.89%
Parentes, Orientação Vocacional	1	1.89%
Parentes, Mercado aquecido	1	1.89%
Parentes, Gosto Pessoal	1	1.89%
Professor, Pesquisa na Mídia	1	1.89%
Professor, Amigos	5	9.43%
Professor, Orientação Vocacional	3	5.66%
Pesquisa na Mídia, Amigos	0	0,00%
Pesquisa na Mídia, Orientação Vocacional	2	3.77%
Amigos, Orientação Vocacional	1	1.89%
Parentes, Professor, Amigos	2	3.77%
Parentes, Professor, Orientação Vocacional	2	3.77%
Parentes, Pesquisa na Mídia, Amigos	1	1.89%
Parentes, Pesquisa na Mídia, Orientação Vocacional	1	1.89%
Professor, Pesquisa na Mídia, Amigos	1	1.89%
Professor, Amigos, Orientação Vocacional	3	5.66%
Pesquisa na mídia, Orientação Vocacional, Psicólogo	1	1.89%
Professor, Amigos, Orientação Vocacional, Psicólogo, Namorado	1	1.89%
Total:	53	100,00%

Figura 7: Tabela – Influências por Curso

Na Tabela de Frequências para Influência na escolha do Curso podemos observar que o professor é o fator que mais exerce influência na escolha, com uma frequência relativa de 22,54%, baseando-se nos dados da turma TN. Nota-se também que há uma influência significativa de Parentes, Pesquisa na Mídia e Orientação vocacional na escolha, visto que possuem frequências entre 5,66% e 9,43%.

Analisando também as opções individuais, juntamente com as marcadas em conjunto, podemos observar que um total de 28 alunos apontaram o professor como influência, o que representa uma frequência relativa de 52,83%, ou seja, a maioria dos alunos, consideram o professor como um fator de influência determinante na escolha do curso. Isso nos mostra o quanto o poder que a influência desses professores têm sobre as escolhas futuras de seus discípulos. Por fim, visando um quadro geral, as opções foram marcadas de forma bem distribuída, o que pode indicar uma variabilidade razoável nas influências.

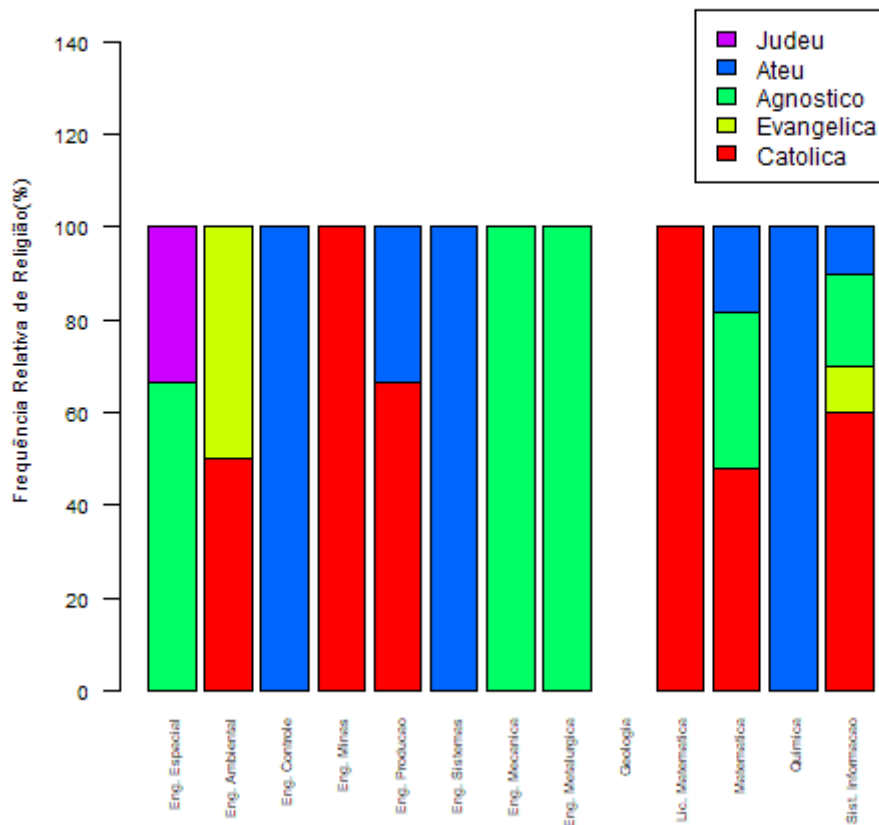


Figura 8: Gráfico de Barras – Religião por Curso

Quanto às frequências relativas, as únicas situações em que há frequência de 100% para uma categoria de religião se aplica aos cursos que possuem apenas um aluno na turma, como visto nos cursos de Engenharia de Controle e Automação, Engenharia de Sistemas, Química e Licenciatura em Matemática. A partir do segundo curso com menos alunos na turma - Engenharia Ambiental - com 2 alunos, já há a presença de duas categorias. Outro detalhe é a ausência de Espíritas em todos os cursos. Em um contexto geral é visto nessa turma de graduandos uma frequência por curso mais notável das categorias Católica, Agnóstica e Ateu.

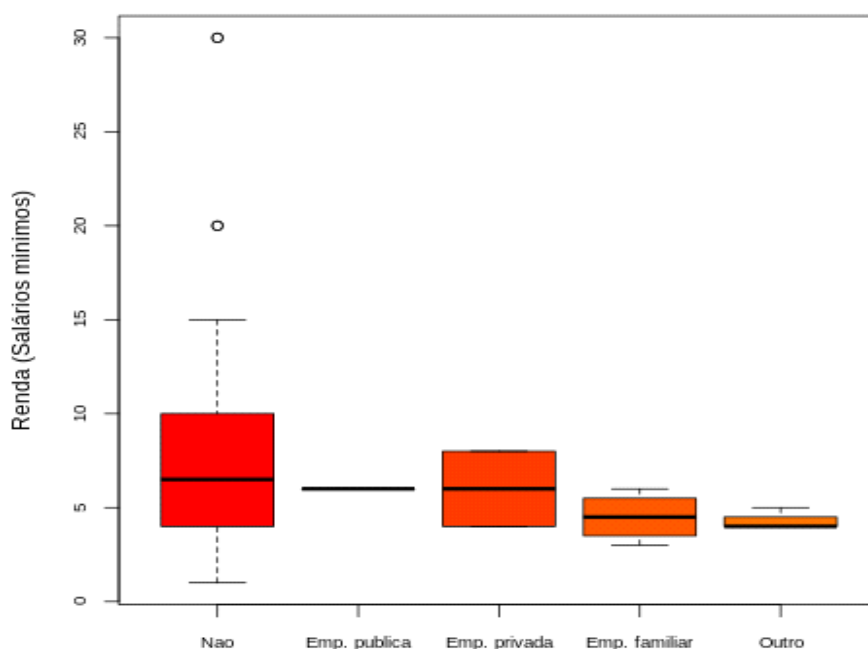


Figura 9: Boxplot – Renda por Trabalho

Na comparação das variáveis renda por trabalho temos que a caixa representante dos alunos que não trabalham é superior em amplitude contra as outras categorias dessa variável, demonstrando que o grupo dos estudantes que não possuem emprego (maioria da turma) apresentam tanto a menor renda quanto a maior renda, dando destaque para os dois outliers presentes nessa categoria, de 20 e 30 salários mínimos. Isso pode se dar devido ao fato de que muitos dos estudantes dessa turma serem jovens graduandos. Outro detalhe é o fato da categoria de trabalho em empresa pública não possuir variabilidade, sendo visível apenas o traço da mediana, cujo valor é 6. O motivo disso se deve ao fato de apenas um estudante se encaixar nessa categoria, logo delimitando a média, mediana, moda, e os quartis. Outra informação interessante é como as caixas das categorias empresa privada e familiar são simétricas, demonstrando como os dados coletados, quando ordenados, estão bem distribuídos sobre essa relação entre as variáveis renda e trabalho. Por fim, observamos a categoria “outro” com a segunda menor amplitude e menor valor máximo, além de ser assimétrica com sua mediana tocando o seu mínimo. Por fim, concluímos que, devido à grande variabilidade quanto às rendas familiares, é difícil estabelecer uma relação clara entre a atuação em determinado tipo de empresa e a renda.

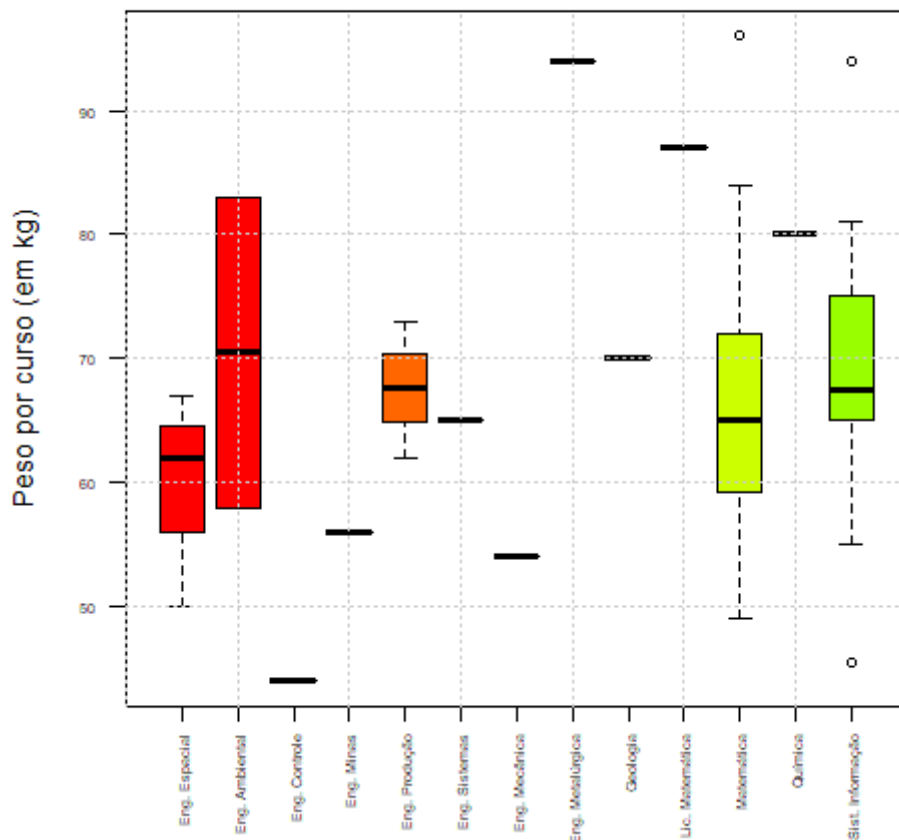


Figura 10: Boxplot – Peso (kg) por Curso

A partir desse gráfico podemos concluir que o curso com a maior mediana é Engenharia Mecânica, totalizando 95 kg. Percebe-se a simetria do curso de Engenharia Produção, cuja mediana é de 68 kg e sua variabilidade é a menor dentre os cursos com mais de um membro. Já o curso de Engenharia Ambiental também apresenta simetria dos dados, com mediana de 70 kg, porém sua variabilidade é a mais alta. Ademais, o curso de Sistemas de Informação conta com dois outliers um acima do terceiro quartil e outro abaixo do primeiro quartil. Para o curso de Engenharia Espacial temos assimetria à direita, mediana de 62 kg e uma variabilidade relativamente baixa. Por último, o curso de Matemática possui um outlier com mais de 95 kg.

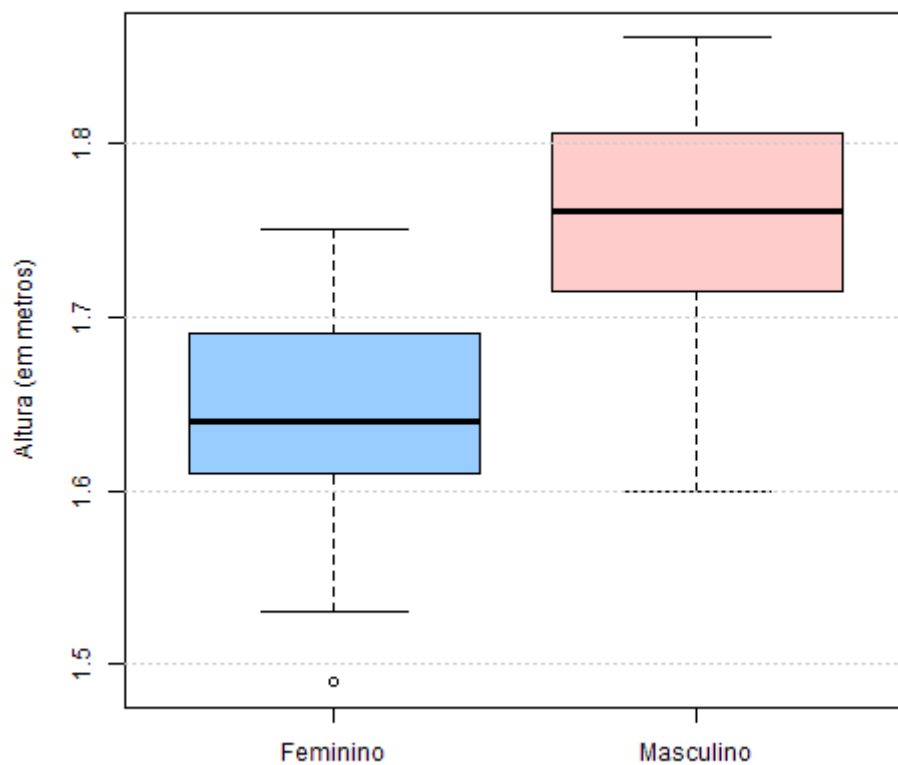


Figura 11: Boxplot – Altura (m) por Sexo

Indivíduos do sexo masculino (mediana 1,76) entrevistados são, em geral, mais altos que os do sexo feminino (mediana 1,64). Além disso, a amplitude total quanto à altura também é maior para os homens. Os dados para as mulheres apresenta leve assimetria positiva, uma vez que a mediana se encontra mais próxima ao segundo quartil, enquanto os dados para os homens é simétrico. Há também a presença de um outlier, uma mulher com altura inferior à 1,5 m.

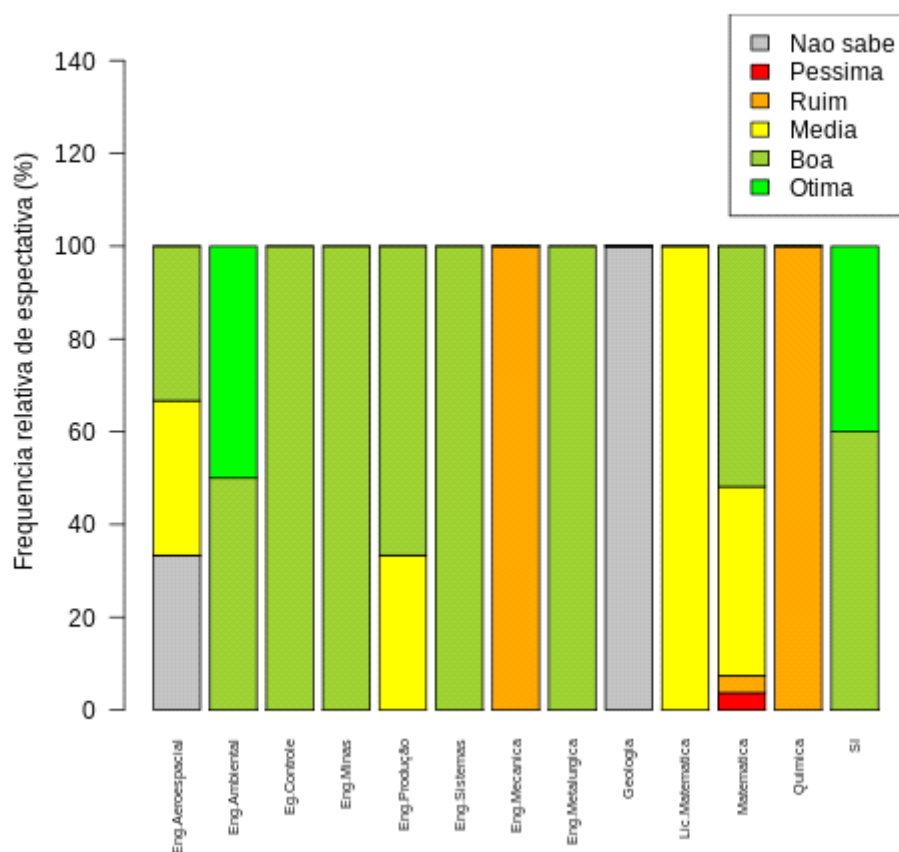


Figura 12: Gráfico de Barras – Expectativa por Curso

Tendo como referência o gráfico de barras acima e a suas frequências relativas apresentadas pode-se apontar o seguinte: Apesar de termos dois cursos com 100% de frequência da categoria “Ruim”, os cursos presentes apresentam expectativas bem positivas em sua maioria, dominando frequência em 7 de 13 cursos, ou 8 de 13, considerando a categoria “Ótima” como acumulativa no caso do curso de Engenharia Ambiental. Além disso, a categoria “Péssima” é quase inexistente estando presente em apenas em 1 curso, o de Matemática, representando menos de 10% de todas as opiniões coletadas nesse curso. Com esse gráfico de barras temos clara dominância de boas expectativas.

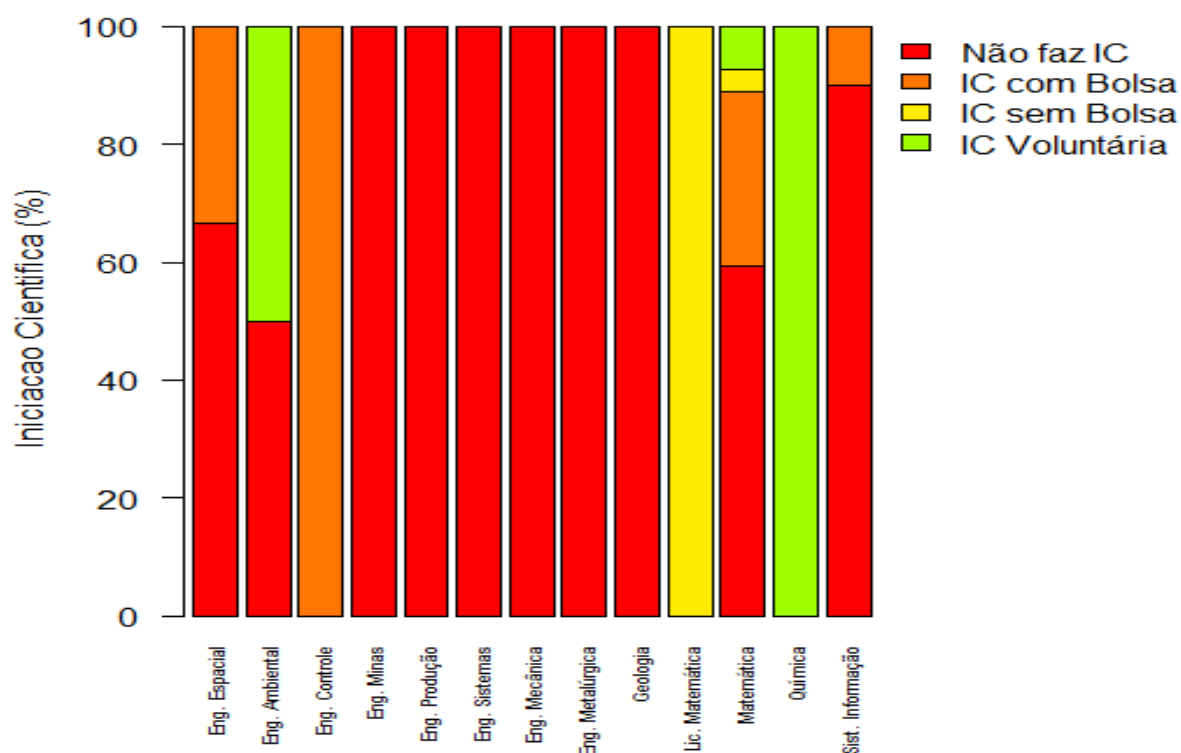


Figura 13: Gráfico de Barras – Iniciação Científica por Curso

Em seis dos treze cursos não há nenhum aluno que realiza Iniciação Científica (IC), sendo que, desses seis cursos, cinco são cursos da Escola de Engenharia. Contudo, nem todos os cursos na área de Engenharia tiveram nenhuma participação, a exemplo das Engenharias: Espacial, Ambiental e de Controle. Ademais, observa-se que para os cursos de Licenciatura em Matemática, Química e Engenharia, todos os alunos estão matriculados em IC, entretanto, deve-se lembrar que a distribuição de alunos por curso é desigual, e que nos cursos citados existe apenas um representante em turma, o que explica essa participação em IC relativamente superior. Para mais, em três cursos: Química, Eng. Ambiental e Matemática, temos alunos que realizam Iniciação Científica Voluntária.

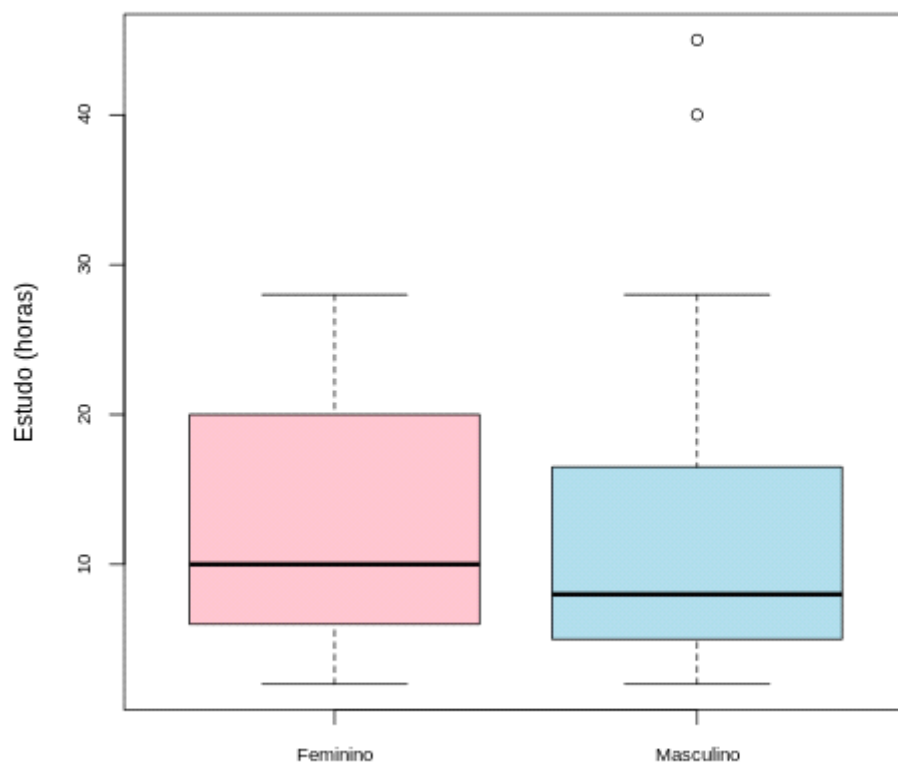


Figura 14: Boxplot – Horas de Estudo por Sexo

Pode-se fazer as seguintes afirmações sobre a distribuição das horas de estudo entre os dois sexos nessa turma - primeiramente é visto como a amplitude total, desconsiderando os outliers, é praticamente idêntica, ou seja, há homens e mulheres que estudam no mínimo e máximo, em limite, o mesmo tempo. Além disso, tem-se os outliers expressivos na caixa da categoria masculina, chegando a ultrapassar 40 horas de estudos semanais. Em questão de mediana é visto como as duas caixas possuem uma distribuição de dados assimétrica positiva. Por mais, a categoria feminina possui uma maior dispersão, por sua caixa possuir maior altura (maior diferença entre o terceiro e o segundo quartil), além de estudar mais que os homens, já que sua mediana (10 horas) é maior que a da categoria masculina (aproximadamente 8 horas).

• CONCLUSÃO

Em suma, ao extrair informações dessa base de dados vemos que uma parcela significativa delas vai de encontro com o observado no censo da UFMG. A maioria católica, a disparidade entre a representação feminina e masculina no meio acadêmico, a predominância de indivíduos pertencentes à classe média e classe média baixa. Tudo isso corrobora nossa constatação de que a turma TN é um mini-retrato representativo do todo formado pela UFMG.

• APÊNDICE.

Tabelas foram feitas com o auxílio da ferramenta - Google Docs

“R: The Project for Statistical Computing” : <https://www.r-project.org/>

"Quick-R by DataCamp": <https://www.statmethods.net/index.html>

“Censo Ufmg” : <https://www.ufmg.br/censo/>

Segue o código em R utilizado para gerar os recursos gráficos:

```
dados<- read.csv("code.csv", sep=";", dec=".", header=T); attach(dados);
```

Figura 1: Gráfico de Barras – Sexo por Curso

```
curso_labels <- c("Eng.Aeroespacial", "Eng.Ambiental", "Eng.Controle", "Eng.Minas",  
                  "Eng.Produção", "Eng.Sistemas", "Eng.Mecanica", "Eng.Metalurgica",  
                  "Geologia", "Lic.Matematica", "Matematica", "Quimica",  
                  "SI")  
  
sexo_labels <- c("Feminino", "Masculino")  
  
sex_cur_freq_a <- table(sexo, curso)  
  
sex_cur_freq_r <- round(prop.table(sex_cur_freq_a, margin = 2) * 100, 2)  
  
png(file = "barplot_sexo_curso.png")  
  
barplot (  
  sex_cur_freq_r,  
  ylab = "Frequencia relativa de sexo (%)",  
  names.arg = curso_labels,  
  col = c("pink", "lightblue"),  
  las = 2,  
  ylim = c(0,150),  
  cex.names = 0.6,  
)  
  
legend("topright", sexo_labels, fill = c("pink", "lightblue"))  
  
dev.off()
```

#Figura 2: Histograma - Número de Alunos por Faixa Etária

```
tab_idade<- table(idade)

tab_idade_rel<- round(prop.table(t5)*100, 2)

png("faixaEtariaXqtdeAlunos.png")

hist_idade<- hist(idade, freq=T, xlab= "Faixa Etária", ylab= "Número de Alunos",
                  ylim= c(0, 30), col=cores, main=" ");

grid(nx=NA, ny=NULL);

text(h$mids, h$counts, labels=round(tab_idade_rel, 2), adj=c(0.5, -0.5));

dev.off()
```

#Figura 3: Boxplot - Idade por Sexo

```
png("idadeXsexo.png")

boxplot(idade~sexo, ylab= "Idade (anos)", main= " ", col = "lightblue")

legend("topleft", legend=c("0 = Feminino", "1 = Masculino"))

dev.off()
```

#Figura 4: Gráfico de Barras - Sexo por Local de Nascimento

```
t8<- table(local, sexo)

tLocal<- round(prop.table(t8)*100,2)

nomes.linha<- c("Capital", "Interior do estado", "Outro Estado", "Outro País")

nomes.coluna<- c("Feminino", "Masculino")

dimnames(tLocal)<- list(nomes.linha, nomes.coluna)

barplot(t(tLocal),

        ylab="Frequência Relativa de Sexo (%)",

        xlab= "Local de Nascimento",

        names.arg=nomes.linha,

        cex.lab=0.8,

        cex.axis=0.8,

        cex.names=0.6,

        col= c("white", "black"),

        legend.text=T,

        ylim=c(0,100),las= 1)
```

#Figura 5: Boxplot – Tentativas Enem

```

png("boxplot-altura-sexo.png")

boxplot(enem~cursinho,main="",ylab="Tentativas no Enem",names=c("Nao fez","Fez"),col=c("green","yellow"))

grid(nx=NA, ny=NULL)

dev.off()

```

#Figura 6 e 7: Tabelas

Foram feitas manualmente usando planilhas, conforme sugerido pela professora.

#Figura 8: Gráfico de Barras - Religião por Curso

```

tab_curso_relig<- table(curso, religiao)

tab_curso_relig_relat<- round(prop.table(tab_curso_relig)*100, 2)

nomes.coluna<- c("Catolica", "Evangelica", "Agnostico", "Ateu", "Judeu")

nomes.linha<- c("Eng. Espacial", "Eng. Ambiental", "Eng. Controle", "Eng. Minas", "Eng. Producao",

               "Eng. Sistemas", "Eng. Mecanica", "Eng. Metalurgica", "Geologia", "Lic. Matematica", "Matematica",

               "Quimica", "Sist. Informacao")

dimnames(tab_curso_relig_relat)<- list(nomes.linha, nomes.coluna)

tab_curso_relig_relat<- round(prop.table(tab_curso_relig, margin=1)*100,2)

nomes.coluna<- c("Catolica", "Evangelica", "Agnostico", "Ateu", "Judeu")

nomes.linha<- c("Eng. Espacial", "Eng. Ambiental", "Eng. Controle", "Eng. Minas",

               "Eng. Producao", "Eng. Sistemas", "Eng. Mecanica", "Eng. Metalurgica",

               "Geologia", "Lic. Matematica", "Matematica", "Quimica", "Sist. Informacao")

dimnames(tab_curso_relig_relat)<- list(nomes.linha, nomes.coluna)

png("religiaoXcurso.png")

barplot(t(tab_curso_relig_relat),

       ylab="Frequência Relativa de Religião(%)",

       names.arg=nomes.linha,

       cex.lab=0.8,

       cex.axis=0.8,

       cex.names=0.6,

       col= rainbow(length(nomes.coluna)),

       legend.text=T,

```

```

ylim=c(0,150),

las= 2)

dev.off()

```

Figura 9: Boxplot – Renda por Trabalho

```

trabalha_labels <- c("Nao", "Emp. publica", "Emp. privada", "Emp. familiar",
                    "Outro")

renda = as.numeric(as.character(renda))
png(file = "boxplot_renda_trabalha.png")
boxplot (
  renda ~ trabalha,
  ylab   = "Renda (Salários mínimos)",
  col    = rainbow(length(trabalha)),
  names  = trabalha_labels,
  cex.axis = 0.8
)
dev.off()

```

#Figura 10: Boxplot – Peso (kg) por Curso

```

peso= as.numeric(as.character(peso))

png("boxplot_peso_curso.png")

boxplot(peso~curso, xlab= "", ylab="Peso por curso (em kg)", col= rainbow(length(curso)),

        names= legCurso, cex.axis=0.6, cex.lab=1.2, las=2 )

grid(nx= NULL, ny= NULL)

dev.off()

```

#Figura 11: Boxplot - Altura (m) por Sexo

```

altura= as.numeric(as.character(altura))

png("boxplot_altura_sexo.png")

boxplot(altura~sexo, xlab= "", ylab="Altura (em metros)", main="",

        col= c("Lightblue", "Pink"), names= c("Feminino", "Masculino"))

grid(nx= NA, ny= NULL)

dev.off()

```

Figura 12: Gráfico de Barras – Expectativa por Curso

```

expec_labels <- c("Nao sabe", "Pessima", "Ruim",
                 "Media" , "Boa" , "Otima")

```

```

curso_labels <- c("Eng.Aeroespacial", "Eng.Ambiental", "Eg.Controle", "Eng.Minas",
                  "Eng.Produção", "Eng.Sistemas", "Eng.Mecanica", "Eng.Metalurgica",
                  "Geologia", "Lic.Matematica", "Matematica", "Quimica",
                  "SI")

exp_cur_freq_a <- table(expect, curso)
exp_cur_freq_r <- round(prop.table(exp_cur_freq_a, margin = 2) * 100, 2)
png(file = "barplot_expec_curso.png")
barplot (
  exp_cur_freq_r,
  ylab = "Frequencia relativa de expectativa (%)",
  names.arg = curso_labels,
  col = c("grey", "red", "orange", "yellow", "yellowgreen", "green"),
  las = 2,
  ylim = c(0,150),
  cex.names = 0.6,
)
legend("topright", expec_labels, fill = c("grey", "red", "orange", "yellow",
      "yellowgreen", "green"))
dev.off()

```

#Figura 13: Gráfico de Barras – Iniciação Científica por Curso

```

t1 <- table(ic,curso)

t1 <- round(prop.table(t1,margin=2)*100,2)

legendaCurso = c("Eng. Espacial", "Eng. Ambiental", "Eng. Controle", "Eng. Minas", "Eng. Produção",
                  "Eng. Sistemas", "Eng. Mecânica", "Eng. Metalúrgica", "Geologia", "Lic. Matemática",
                  "Matemática", "Química", "Sist. Informação")

cor = rainbow(length(legendaCurso))

barCursoIc <- barplot(t1,names.arg = legendaCurso,ylab = "Iniciacao Cientifica (%)", xlab = "",
                      ylim = c(0,101),
                      xlim = c(0,23),
                      cex.names=0.6,
                      col = cor,
                      xaxs="i",
                      las=2,
)

```

```

legend(
  "topright",
  legend = c("Não faz IC", "IC com Bolsa", "IC sem Bolsa", "IC Voluntária"),
  fill = cor,
  bty = "n"
)

```

#Figura 14: Boxplot - Horas de Estudo por Sexo

```

png("horasXsexo.png")

boxplot(h_estudo~sexo, ylab= "Horas de estudo", main= " ", col = "green")

legend("topleft", legend=c("0 = Feminino", "1 = Masculino"))

dev.off()

```