

**UNIVERSIDADE REGIONAL INTEGRADA DO ALTO URUGUAI E DAS MISSÕES**  
**URI - ERECHIM**  
**CURSO DE CIÊNCIAS DA COMPUTAÇÃO**

**GUSTAVO MARMENTINI**  
**MURILO FERANTI**  
**VINICIUS DUFLOTH**

**TRABALHO DE ESTATÍSTICA**

Trabalho da disciplina de estatística, visando análise de dados de uma amostra sorteada pelo grupo a partir de uma população fornecida pelo professor.

Prof. Claodomir Antônio Martinazzo

**ERECHIM**  
**2019**

## Introdução

Esse trabalho é baseado no conjunto de dados captados pelo Prof. Claodomir Antônio Martinazzo de estudantes da URI. Nos foi disponibilizado em planilha para que obtivéssemos aleatoriamente 40 registros que comporiam a amostra da pesquisa realizada.

A partir destes dados, realizamos diversos gráficos, fórmulas e conclusões, demonstrando o aprendizado estatístico obtido durante todo o semestre.

## Desenvolvimento

Primeiramente, classificamos as variáveis disponibilizadas. No Quadro 1 estão demonstradas as variáveis que foram obtidas na pesquisa, e suas respectivas classificações:

Quadro 1: Classificação das variáveis

Curso: Qualitativa Nominal	Cidade: Qualitativa Nominal
Sexo: Qualitativa Nominal	Altura: Quantitativa Contínua
Fumo: Qualitativa Nominal	Internet: Qualitativa Nominal
Idade: Quantitativa Contínua	Peso: Quantitativa Contínua
Ano de nascimento do aluno: Quantitativa Discreta	Posse do computador: Qualitativa Nominal
Grau de instrução do pai: Qualitativa Ordinal	IMC: Quantitativa Contínua

Fonte: Dados da pesquisa

Em sequência, encontramos a média, mediana, mínimo, máximo, primeiro e segundo quartis, variância e desvio padrão das variáveis ano de nascimento, peso e altura da população e amostra.

Tabela 1: Medidas descritivas da população

Variáveis	N	Média	Mediana	Mínimo	Máximo	Quartil 1	Quartil 2	Variância	Desvio Padrão
Ano de Nasc.	215	1991.3442	1992	1978	1996	1990	1992	10.8716	3.2972
Peso	215	63.2028	60	40	110	54	60	148.775	12.1973
Altura	215	1.678	1.67	1.45	1.94	1.62	1.67	0.0081	0.09

Fonte: Dados da pesquisa.

Tabela 2: Medidas descritivas da amostra

Variáveis	N	Média	Mediana	Mínimo	Máximo	Quartil 1	Quartil 2	Variância	Desvio Padrão
Ano de Nasc.	40	1992.225	1993	1985	1996	1991.75	1993	6.6917	2.5868
Peso	40	63.98	60.5	46	93.5	53.75	60.5	157.9763	12.5689
Altura	40	1.69	1.67	1.54	1.9	1.6275	1.675	0.0089	0.0943

Fonte: Dados da pesquisa.

Foi possível notar uma diferença nas medidas descritivas entre amostra e população.

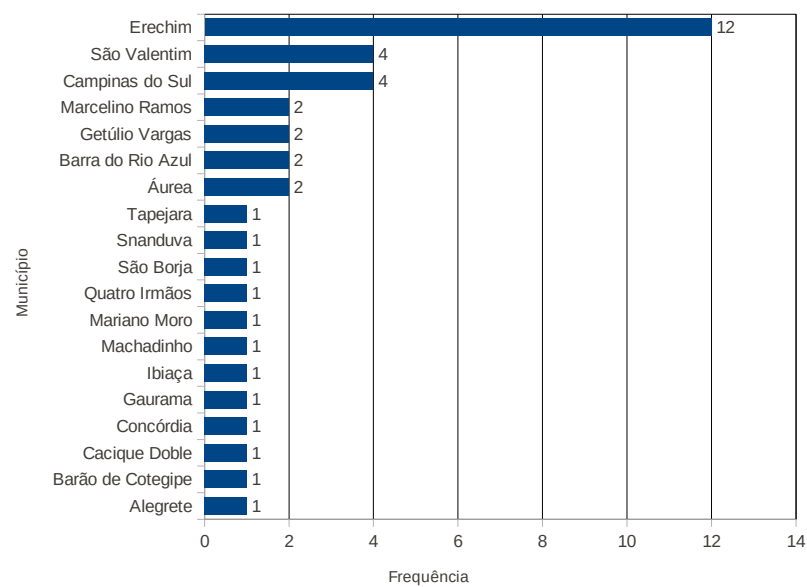
Partimos então, para desenhar uma tabela geográfica com a variável Municípios assim como sua representação em gráfico de barras.

Tabela 3: Municípios

Cidade	Quantidade
Alegrete	1
Barão de Cotegipe	1
Cacique Doble	1
Concórdia	1
Gaurama	1
Ibiaça	1
Machadinho	1
Mariano Moro	1
Quatro Irmãos	1
São Borja	1
Snanduva	1
Tapejara	1
Áurea	2
Barra do Rio Azul	2
Getúlio Vargas	2
Marcelino Ramos	2
Campinas do Sul	4
São Valentim	4
Erechim	12
<b>Total</b>	<b>40</b>

Fonte: Dados da pesquisa

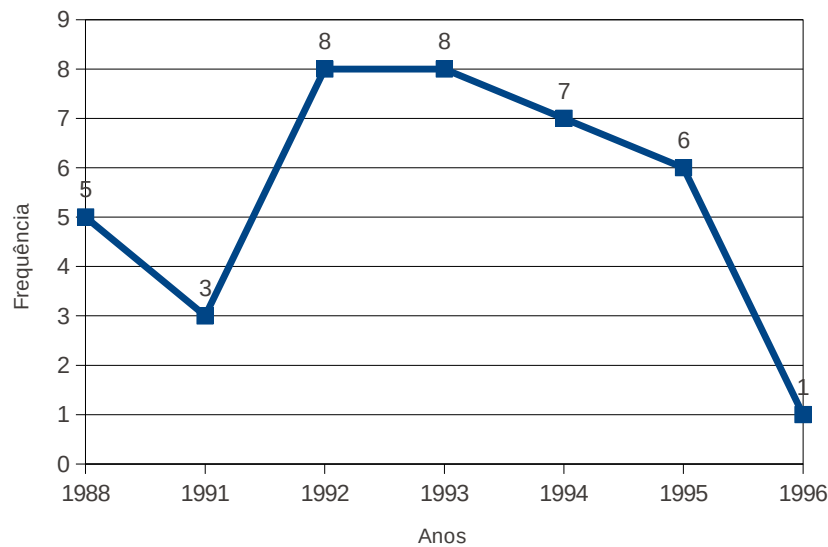
Gráfico 1: Geográfico



Fonte: Dados da pesquisa

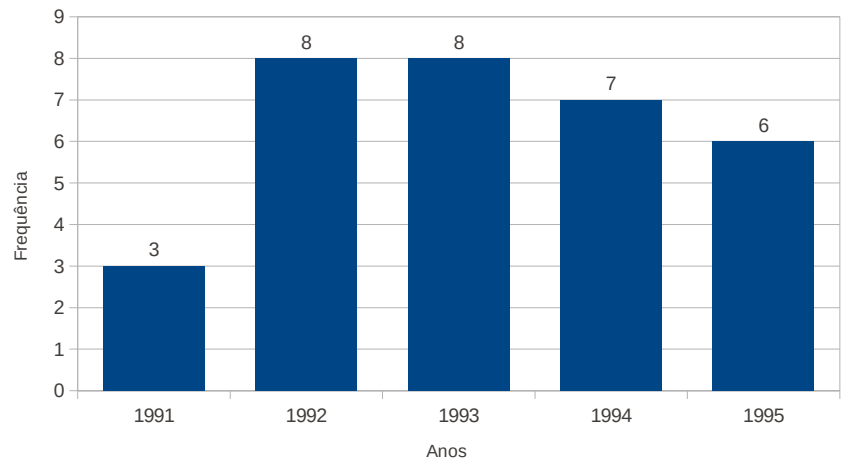
Também foi elaborada uma tabela histórica baseada nos anos de nascimento e representada em gráfico de linhas. Foram utilizados 7 anos e em sequência, 5 anos dentre estes 7.

Gráfico 2: Histórico, 7 anos



Fonte: Dados da pesquisa

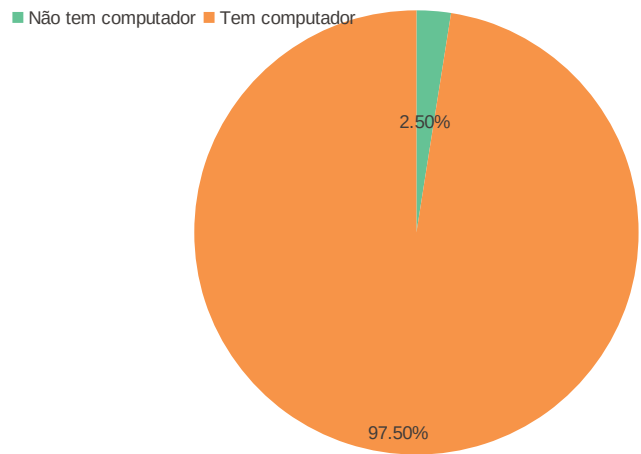
Gráfico 3: Histórico, 5 anos



Fonte: Dados da pesquisa

Fazendo uma análise da variável Computador(posse) foi possível perceber uma clara dominância de estudantes que possuem computador, como representado no Gráfico 4.

Gráfico 4: Posse de computador



Fonte: Dados da pesquisa

Para análise da frequência de alunos que nasceram em certos anos, foi criada uma tabela de distribuição de frequências, representada na Tabela 4.

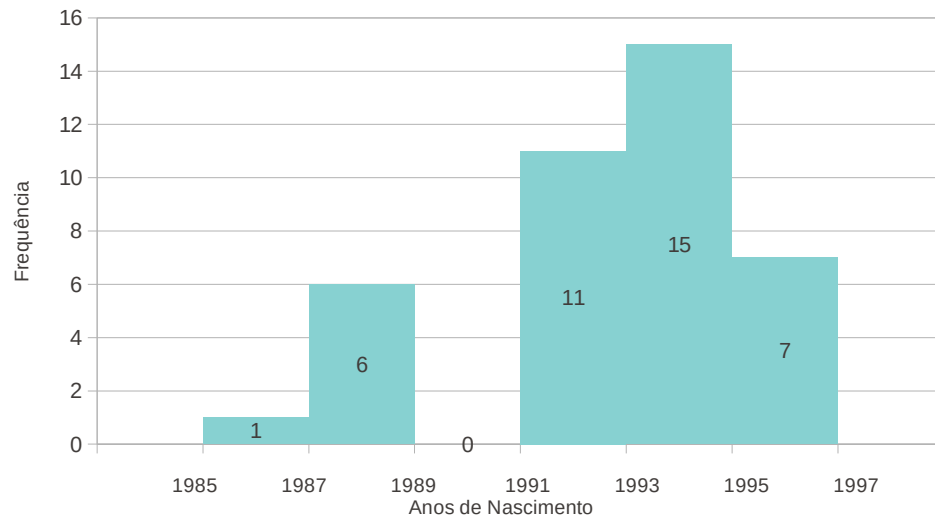
Tabela 4: Frequência de anos de nascimento

Classe		Fi	xi	Fac	fi(%)	FacR(%)
1985	1987	1	1986	1	2.5	2.5
1987	1989	6	1988	7	15	17.5
1989	1991	0	1990	7	0	17.5
1991	1993	11	1992	18	27.5	45
1993	1995	15	1994	33	37.5	82.5
1995	1997	7	1996	40	17.5	100
Total		40	-	-	100	-

Fonte: Dados da pesquisa

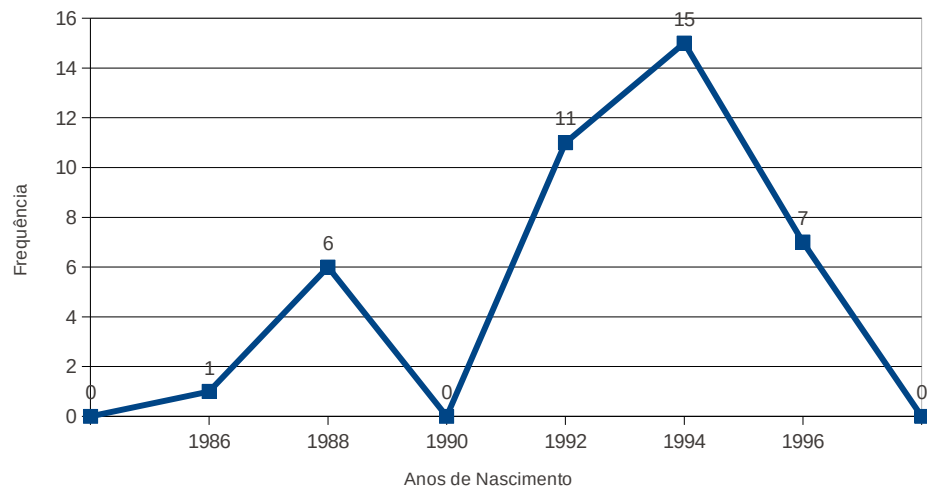
A partir dos dados da Tabela 4, foram desenhados gráficos em histograma, polígono e ogiva.

Gráfico 5: Frequência Idades - histograma



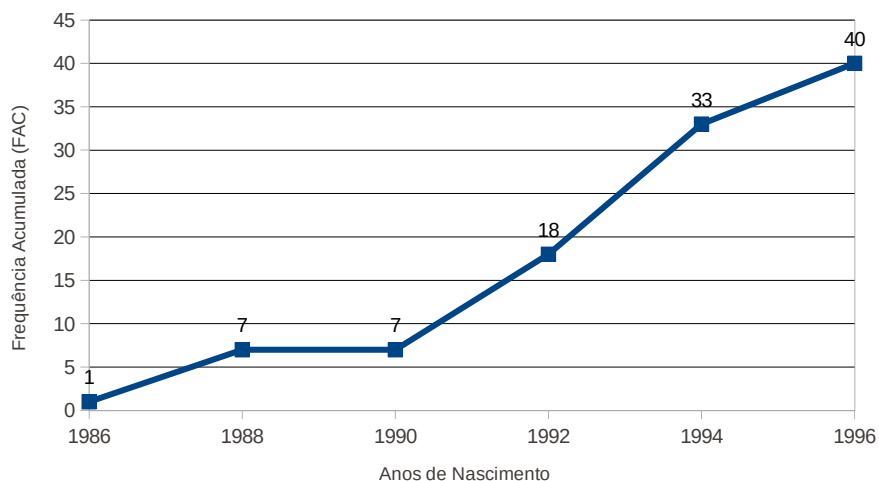
Fonte: Dados da pesquisa

Gráfico 6: Frequência Idades – polígono



Fonte: Dados da pesquisa

Gráfico 7: Frequência Idades – ogiva

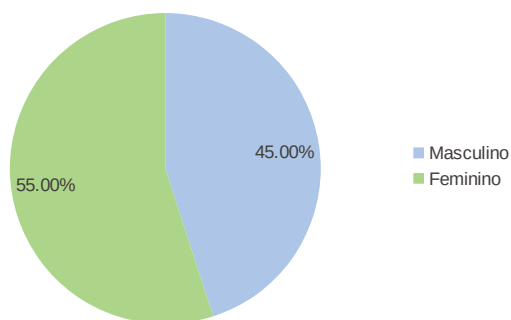


Fonte: Dados da pesquisa

Um fato curioso, e que fica explícito nos Gráficos 5 e 6 é que em nossa amostra, não tivemos nenhum estudante que nasceu entre os anos 1989 e 1991.

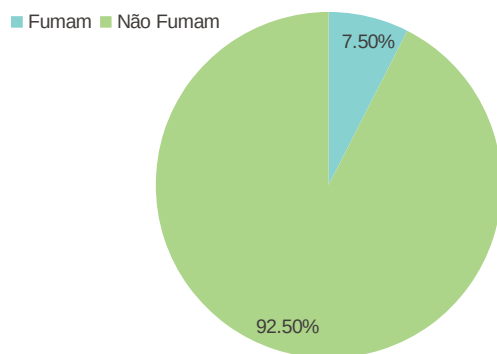
Para estudo das variáveis sexo, fumo, atividade física e internet, foram desenhados gráficos no formato pizza.

Gráfico 8: Sexo



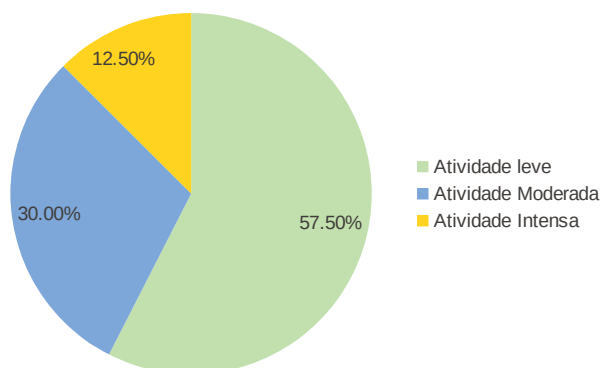
Fonte: Dados da pesquisa

Gráfico 9: Fumo



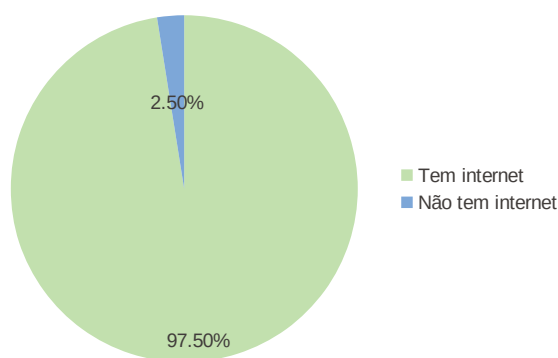
Fonte: Dados da pesquisa

Gráfico 10: Atividade Física



Fonte: Dados da pesquisa

Gráfico 11: Internet



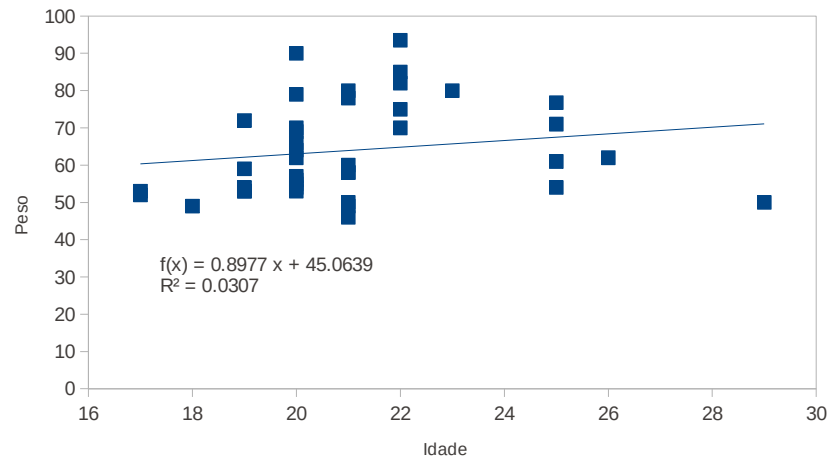
Fonte: Dados da pesquisa

Estudando a correlação entre posse de computador e internet, em nossa amostra obtivemos uma correlação perfeita. Ou seja, todos aqueles que possuíam computador também possuíam internet. E quem não possuía computador, não possuía internet.

Para estudo da relação entre as variáveis idade e peso, foi desenhado um gráfico de dispersão.



Gráfico 12: idade e peso



Fonte: Dados da pesquisa

Observamos que o grau de correlação entre as variáveis Peso e Altura é de apenas 0.18, indicando uma fraca correlação.

O coeficiente de determinação portanto, mostrou-se de 3%, indicando também uma fraca capacidade de determinação.

A equação de regressão, para a idade a partir do peso é  **$f(x) = 0.8977 x + 45.0639$**

Para finalizar, realizamos um cálculo do erro amostral médio da população e da amostra para as variáveis de peso e altura. O erro médio da população foi de 0.03, enquanto que o da amostra foi de -0.05.

Portanto, pode-se notar que a confiança no resultado da altura estimada foi melhor com o maior número de dados provido pela população.

## Conclusão

Podemos botar em prática neste trabalho todos os conhecimentos obtidos no semestre e visualizar na prática a análise de diferentes variáveis e comparações entre populações e amostras.

Percebemos o trabalho como muito proveitoso e acreditamos ter consolidado o aprendizado esperado.