



**Programa de Mestrado Profissional
em Matemática em Rede Nacional
Coordenação do PROFMAT**

**JOSÉ FELIPE ESTOLANO BATISTA DOS
SANTOS**

***MATEMÁTICA, ESTATÍSTICA E
LINGUÍSTICA: UM RELATO DE
EXPERIÊNCIA
INTERDISCIPLINAR NO
CONTEXTO DA ESCOLA BÁSICA***

**Orientador:
Humberto José Bortolossi**

**UNIVERSIDADE
FEDERAL
FLUMINENSE**

**NITERÓI
FEVEREIRO/2015**

José Felipe Estolano Batista dos Santos

***Matemática, Estatística e Linguística: Um Relato de
Experiência Interdisciplinar no Contexto da Escola
Básica***

Niterói – RJ

Fevereiro / 2015

Ficha catalográfica elaborada pela Biblioteca de Pós-graduação em Matemática da UFF

S939 Santos, José Felipe Estolano Batista dos
Matemática, Estatística e Linguística: Uma Relato de Ex-
periência Interdisciplinar no Contexto da Escola Básica / José Fe-
lipse Estolano Batista dos Santos. – Niterói: [s.n.], 2015.
71 f.

Orientador: Humberto José Bortolossi
Dissertação (Mestrado Profissional em Matemática em Rede
Nacional PROFMAT) – Universidade Federal Fluminense, 2015.

1. Interdisciplinaridade. 2. Estatística. 3. Matemática. 4.
Linguística. I. Título.

CDD: 510.7

Todos os direitos reservados.

É proibida a reprodução total ou parcial sem autorização da universidade, do autor e do orientador.

José Felipe Estolano Batista dos Santos

***Matemática, Estatística e Linguística: Um Relato de
Experiência Interdisciplinar no Contexto da Escola
Básica***

Dissertação apresentada à Coordenação do
Mestrado Profissional em Matemática em Rede
Nacional – PROFMAT da Universidade Federal
Fluminense para a obtenção do título de Mestre
em Matemática

Orientador:

Humberto José Bortolossi

MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL – PROFMAT
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
UNIVERSIDADE FEDERAL FLUMINENSE

Niterói – RJ

Fevereiro / 2015

Dissertação de Mestrado Profissional sob o título “*Matemática, Estatística e Linguística: Um Relato de Experiência Interdisciplinar no Contexto da Escola Básica*”, defendida por José Felipe Estolano Batista dos Santos e aprovada em 6 de fevereiro de 2015, em Niterói, Rio de Janeiro, pela banca examinadora constituída pelos professores:

Humberto José Bortolossi
Doutor em Matemática pela PUC-Rio
Orientador

Alexandre Sousa da Silva
Doutor em Estatística pela UFRJ

Cecília de Souza Fernandez
Doutora em Matemática pela Kent State University

José Teixeira Cal Neto
Doutor em Matemática pela PUC-Rio

Vanda Maria Cardozo de Menezes
Doutora em Letras pela UFRJ

*Dedico esse trabalho à minha esposa Claudia, meus filhos Talles e Theo, ao meu irmão
Leandro, ao meu Pai José e especialmente à minha mãe Verônica, que sempre vislumbrou e
acreditou nesse momento.*

Agradecimentos

Agradeço a Deus por me dar saúde, dedicação e paciência.

Agradeço ao meu professor orientador Humberto (sem suas orientações não conseguiria escrever metade da tese).

Agradeço à minha mãe por sempre me incentivar, à minha esposa por me apoiar e aos meus filhos que me motivam e inspiram.

Agradeço a oportunidade que me foi dada pelo projeto PROFMAT e aos meus colegas de mestrado que passaram pelos mesmos “apertos”.

Não posso me esquecer de agradecer também aos alunos colaboradores e geradores de material para a pesquisa. Muito obrigado a todos.

A Matemática é o alfabeto com o qual DEUS escreveu o universo... (Galileu Galilei)

Não há ramo da Matemática, por abstrato que seja, que não possa um dia vir a ser aplicado aos fenômenos do mundo real. (Nicolai Lobachevsky)

Resumo

O *software* multiplataforma gratuito “Estatística das Letras, Palavras e Períodos (LPP)” (<<http://www.uff.br/cdme/desktop/lpp/lpp-br.html>>) permite realizar análises estatísticas de sequências de caracteres. Mais precisamente, o aplicativo conta o número de letras, dígitos, acentos, sinais de pontuação, palavras e períodos de um texto. Ele também calcula o número de letras por palavra e o número de palavras por período (apresentando medidas de posição e dispersão destas variáveis quantitativas), os períodos mais longos, os períodos mais curtos, as palavras mais longas e as palavras mais curtas. Existem três módulos especiais: o primeiro módulo permite usar a distribuição de frequência das letras para decodificar mensagens criptografadas com a Cifra de César, o segundo módulo permite investigar leis de potência na distribuição de frequência das palavras num texto (Lei de Zipf) e o terceiro permite analisar o crescimento do vocabulário ao longo do texto. O objetivo principal do *software* é o de oferecer um ambiente interativo no qual aluno e professor possam experimentar, explorar e apreciar o uso da Estatística em Linguística e, a partir desta articulação, aprender conceitos estatísticos no contexto linguístico, bem como também aprender conceitos linguísticos usando Estatística. Neste trabalho relatamos nossa experiência em usar este *software* com duas turmas do Ensino Médio, uma turma de um curso técnico em Informática e uma turma de Especialização para Professores de Matemática da Escola Básica.

Palavras-chave: ensino e aprendizagem de Matemática e Estatística; interdisciplinaridade; Linguística.

Abstract

The free multiplatform software “Statistics of Letters, Words and Periods” (<<http://www.uff.br/cdme/desktop/lpp/lpp-en.html>>) performs statistical analysis of sequences of characters. More precisely, the application counts the number of letters, digits, accents, punctuation marks, words and periods of a text. It also calculates the number of letters per word and words per period (showing position and dispersion measures of these quantitative variables), the longest periods, the shortest periods, the longest words and the shortest words. There are three special modules: the first module allows the use the frequency distribution of the letters to decode messages encrypted with the Caesar cipher, the second module allows the investigation of power laws in frequency distributions of words in a text (Zipf’s law), and the third one allows the study of vocabulary richness. The main objective of our software is to provide an interactive environment in which students and teachers can experiment, explore and enjoy the use of statistics in Linguistics, and through this articulation, learn statistical concepts in the linguistic context and, also, learn linguistics concepts using Statistics. In this work, we report our experience in using this software with two High School classes, a class of a technical course in Informatics and a class for in-service teachers of Mathematics.

Keywords: teaching and learning of Mathematics and Statistics; interdisciplinarity; Linguistics.

Sumário

1	Introdução	p. 10
1.1	Ensino e Aprendizagem da Estatística: PCN	p. 10
1.2	Ensino e Aprendizagem da Estatística: GAISE	p. 12
1.3	Proposta de Nosso Trabalho	p. 14
2	O Contexto do Trabalho	p. 16
2.1	Instituições de Ensino	p. 16
2.2	Descrição dos Recursos Computacionais Usados	p. 21
2.3	Descrição das Atividades	p. 25
2.4	Cronograma Simplificado	p. 29
2.5	Laboratórios de Informática	p. 30
3	Relato de Experiência	p. 31
3.1	Grupo 1: Estatística e Criptografia	p. 31
3.2	Grupo 2: Atividades Mais Abrangentes	p. 37
3.3	Grupo 3: Análise Estatística por Autor	p. 48
4	Considerações Finais	p. 53
	Apêndice A – Questionário de Sondagem	p. 57
	Apêndice B – Formulário do Grupo 1	p. 58
	Apêndice C – Formulário do Grupo 2	p. 61

Apêndice D – Formulário do Grupo 3

p. 69

Referências Bibliográficas

p. 70

1 *Introdução*

Neste capítulo apresentamos orientações e metodologias recomendadas por alguns documentos oficiais no Brasil e nos Estados Unidos para o ensino e a aprendizagem da Estatística e, em seguida, neste contexto, apresentamos nossa proposta de trabalho.

1.1 **Ensino e Aprendizagem da Estatística: PCN**

Criado em 1997 pelo Governo Federal, os Parâmetros Curriculares Nacionais – PCN – constituem um referencial para a Educação Básica em todo o País. Com o objetivo de orientar o planejamento escolar, sua proposta não é a criação de um modelo curricular homogêneo e impositivo, mas, sim, a busca de referenciais nos quais o sistema educacional brasileiro possa se organizar. Para o Ensino Médio, os PCN destacam três competências principais a serem desenvolvidas para a área das Ciências da Natureza, Matemática e suas Tecnologias:

- *representação e comunicação, que envolvem a leitura, a interpretação e a produção de textos nas diversas linguagens e formas textuais características dessa área do conhecimento;*
- *investigação e compreensão, competência marcada pela capacidade de enfrentamento e resolução de situações-problema, utilização dos conceitos e procedimentos peculiares do fazer e pensar das ciências;*
- *contextualização das ciências no âmbito sócio-cultural, na forma de análise crítica das ideias e dos recursos da área e das questões do mundo que podem ser respondidas ou transformadas por meio do pensar e do conhecimento científico.*

(Brasil, 2002, p. 113, grifos nossos)

Para a competência “contextualização das ciências no âmbito sociocultural”, os PCN estabelecem uma forte conexão com tecnologia:

*Ciência e Tecnologia na Atualidade. Na área: **Reconhecer e avaliar o desenvolvimento tecnológico contemporâneo, suas relações com as ciências, seu papel na vida humana, sua presença no mundo cotidiano e seus impactos na vida social.** Na Matemática: Acompanhar criticamente o desenvolvimento tecnológico contemporâneo, tomando contato com os avanços das*

novas tecnologias nas diferentes áreas do conhecimento para se posicionar frente às questões de nossa atualidade. Utilizar o conhecimento matemático como apoio para compreender e julgar as aplicações tecnológicas dos diferentes campos científicos.

(Brasil, 2002, p. 118, grifos nossos)

Os PCN dividem o conhecimento matemático em quatro grandes blocos: (1) *Números e Operações*; (2) *Funções*; (3) *Geometria* e (4) *Análise de Dados e Probabilidade*. Com relação à unidade temática de *Análise de Dados e Probabilidade*, os PCN reforçam o seu caráter interdisciplinar e indicam os conteúdos e habilidades que devem ser trabalhados de maneira específica:

A Estatística e a Probabilidade devem ser vistas, então, como um conjunto de ideias e procedimentos que permitem aplicar a Matemática em questões do mundo real, mais especialmente aquelas provenientes de outras áreas. Devem ser vistas também como formas de a Matemática quantificar e interpretar conjuntos de dados ou informações que não podem ser quantificados direta ou exatamente.

(Brasil, 2002, p. 126, grifo nosso)

Estatística: descrição de dados; representações gráficas; análise de dados: médias, moda e mediana, variância e desvio padrão.

- *Identificar formas adequadas para descrever e representar dados numéricos e informações de natureza social, econômica, política, científico-tecnológica ou abstrata.*
- *Ler e interpretar dados e informações de caráter estatístico apresentados em diferentes linguagens e representações, na mídia ou em outros textos e meios de comunicação.*
- *Obter médias e avaliar desvios de conjuntos de dados ou informações de diferentes naturezas.*
- *Compreender e emitir juízos sobre informações estatísticas de natureza social, econômica, política ou científica apresentadas em textos, notícias, propagandas, censos, pesquisas e outros meios.*

(Brasil, 2002, p. 127, grifos nossos)

Também gostaríamos de registrar aqui o que os PCN recomendam com relação ao processo de avaliação em projetos como ação metodológica:

Ao professor são oferecidas incessantemente muitas oportunidades de observação e avaliação no desenrolar de seu trabalho com os alunos. Muitas vezes, usamos as informações, mas não mantemos nenhum registro delas, outras vezes recolhemos informações que já possuíamos de que não necessitamos ou das quais nunca faremos uso. Pontuar, registrar e relatar são procedimentos comuns numa avaliação que se integra ao ensino.

Uma vez estabelecidos acordos entre professor e alunos sobre os objetivos gerais e específicos do trabalho, é possível definir aspectos a serem avaliados

conjuntamente a cada etapa do ensino. Um exemplo disso seria a forma de avaliar os alunos ao final do projeto proposto anteriormente sobre o perfil dos alunos da classe. Tendo em mãos todas as anotações sobre atividades e as produções dos alunos, o que o professor pode registrar sobre esse percurso de cada um deles? O que ele tem a dizer sobre a posição em que cada aluno se encontra em relação aos objetivos estabelecidos para esse projeto:

- *identificar os dados relevantes entre as informações obtidas;*
- *identificar diferentes formas de quantificar dados numéricos;*
- *selecionar diferentes formas para representar um dado ou conjunto de dados e informações, reconhecendo as vantagens e limites de cada uma delas;*
- *traduzir uma situação dada em determinada linguagem em outra;*
- *ler e compreender diferentes tipos de textos com informações em linguagem matemática;*
- *ler e interpretar dados ou informações apresentados em tabelas, gráficos, esquemas, diagramas, árvores de possibilidades, fórmulas, equações ou representações geométricas;*
- *identificar as relações entre os dados obtidos e as suas regularidades;*
- *extrair e sistematizar as principais conclusões e identificar problemas a serem enfrentados;*
- *elaborar possíveis estratégias para enfrentar os problemas levantados, buscando, se necessário, novas informações e conhecimentos.*

Além disso, em relação a esse projeto, é possível avaliar como o aluno se encontra em relação às competências de natureza geral como:

- *expressar-se com clareza, utilizando a linguagem matemática;*
- *produzir textos analíticos, fazendo uso da linguagem matemática;*
- *compreender e emitir juízos próprios de forma analítica e crítica, posicionando-se com argumentação clara e consistente;*
- *compreender a Matemática como parte integrante da cultura contemporânea;*
- *compreender formas pelas quais a Matemática influencia nossa interpretação do mundo atual.*

(Brasil, 2002, p. 128, grifos nossos)

1.2 Ensino e Aprendizagem da Estatística: GAISE

Por uma iniciativa de Joan Garfield, a *American Statistical Association* publicou dois relatórios intitulados GAISE (*Guidelines for Assessment and Instruction in Statistics Education*) com orientações para a instrução e a avaliação em Educação Estatística: (Garfield et al, 2005) para o Ensino Superior e (Franklin et al, 2007) para o Ensino Básico. O propósito principal destes documentos foi o de complementar e dar mais profundidade e amplitude às orientações dos *Principles and Standards for School Mathematics* (2000) do *National Council of Teachers of Mathematics* (NCTM). Registramos a seguir algumas recomendações principais dadas por estes documentos:

Recomendação 1: Enfatizar o letramento estatístico e desenvolver o pensamento estatístico. Definimos o letramento estatístico como o entendimento da linguagem básica da Estatística (e.g., conhecer o que os termos e símbolos estatísticos significam e saber ler gráficos estatísticos) e as ideias fundamentais da Estatística. [...]

O pensamento estatístico tem sido definido como o tipo de pensamento que estatísticos usam ao abordar ou resolver problemas estatísticos. O pensamento estatístico tem sido descrito como o entendimento da necessidade de dados, a importância na aquisição de dados, a onipresença da variabilidade e a quantificação da variabilidade.

(Garfield et al, 2005, p. 14, tradução nossa)

Recomendação 2: Usar dados reais.

É importante usar dados reais ao se ensinar Estatística para ser autêntico ao se considerar questões relacionadas em como e por que os dados foram produzidos ou coletados, e relacionar a análise ao contexto do problema. Usar conjuntos de dados reais de interesse dos estudantes também é uma boa maneira de engajá-los na reflexão sobre os dados e os conceitos estatísticos relevantes. [...] Um aspecto importante ao se trabalhar com dados reais é que isto ajuda os estudantes a formularem boas perguntas e a usarem dados para respondê-las de forma apropriada com base em como os dados foram produzidos.

(Garfield et al, 2005, p. 16, tradução nossa)

Recomendação 3: Evidenciar o entendimento conceitual no lugar de um mero conhecimento de procedimentos.

Muitos cursos introdutórios contêm muito material, e os estudantes terminam com uma coleção de ideias que são entendidas apenas superficialmente, que não estão bem integradas, e que são facilmente esquecidas. Se os estudantes não entendem os conceitos importantes, existe pouco valor em conhecer um conjunto de procedimentos. Se eles entendem bem os conceitos, então procedimentos particulares serão facilmente aprendidos. [...]

(Garfield et al, 2005, p. 17, tradução nossa)

Recomendação 4: Promover o aprendizado ativo em sala de aula.

Usar métodos de aprendizado ativo em classe é uma forma valiosa de promover o aprendizado colaborativo, permitindo que os estudantes aprendam uns com os outros. O aprendizado ativo permite que os estudantes descubram, construam, e entendam ideias estatísticas importantes modelando assim o pensamento estatístico. [...] Outros benefícios do aprendizado ativo são a prática que os estudantes adquirem comunicando-se na linguagem estatística e o aprendizado de trabalho em grupo. [...]

(Garfield et al, 2005, p. 18, tradução nossa)

Recomendação 5: Usar tecnologia para desenvolver o entendimento conceitual e analisar dados.

Tecnologia mudou a maneira como estatísticos trabalham e deveria mudar o que e como ensinamos [Estatística]. [...] Ahamos que tecnologia

deveria ser usada para analisar dados, permitindo assim que os estudantes se concentrem na interpretação dos resultados e no teste de hipóteses, ao invés de cálculos mecânicos. Ferramentas tecnológicas também podem ser usadas para ajudar os estudantes a visualizarem conceitos e desenvolverem um entendimento de ideias abstratas via simulações. [...] Independentemente das ferramentas usadas, é importante ver o uso da tecnologia não apenas como uma maneira de calcular números, mas como uma maneira de explorar ideias conceituais e melhorar o aprendizado do aluno também. Advertimos contra o uso da tecnologia apenas pela tecnologia [e.g., entrar com 100 números em uma calculadora gráfica para depois calcular resumos estatísticos] ou por uma falsa precisão (cálculo dos resultados com várias casas decimais).

(Garfield et al, 2005, p. 19, tradução nossa)

Recomendação 6: Usar avaliações que melhorem e monitorem o aprendizado do estudante. Os estudantes irão valorizar aquilo que você vai avaliar. Então, as avaliações precisam estar alinhadas com os objetivos de aprendizagem. As avaliações precisam se concentrar no entendimento de ideias-chaves, e não apenas em habilidades, procedimentos e respostas calculadas.

(Garfield et al, 2005, p. 21, tradução nossa)

1.3 Proposta de Nosso Trabalho

Como implementar efetivamente as várias orientações recomendadas para o ensino da Estatística? Como articular competências, conteúdos, habilidades, tecnologia e interdisciplinaridade? Como usar dados reais promovendo o aprendizado ativo em sala de aula? As comunidades de Educação Matemática e de Educação Estatística vêm realizando várias pesquisas no intuito de obter respostas para essas perguntas (os próprios documentos GAISE dão alguns exemplos de atividades alinhadas com suas recomendações para uso em sala de aula).

Neste contexto, Bortolossi (2013) sugere uma série de atividades com o uso de recursos tecnológicos na análise estatística (distribuições de frequência das letras, dígitos, acentos, sinais de pontuação e palavras; distribuições de frequência do número de letras por palavra e o número de palavras por período de um texto e suas medidas de posição e dispersão) de textos em vários idiomas. Segundo Bortolossi (2013), o objetivo principal é o de oferecer um ambiente interativo no qual aluno e professor possam experimentar, explorar e apreciar o uso da Estatística em Linguística e, a partir dessa articulação, aprender conceitos estatísticos no contexto linguístico, bem como também aprender conceitos linguísticos usando Estatística: ao realizar as atividades propostas, os alunos poderão, através de experimentações, perceber como a Estatística se constitui em uma ferramenta útil e poderosa para se identificar padrões e variações em textos, isto é,

como as distribuições de frequência das letras, dígitos, sinais de pontuação, palavras e períodos e suas medidas de posição e dispersão, representadas numericamente e graficamente, podem ajudar a identificar propriedades que caracterizam um determinado idioma, um autor, um estilo literário, uma língua artificial, uma ideologia, etc. (o que constitui um prelúdio às áreas modernas de linguística computacional e mineração de texto). O autor destaca outra característica importante de sua proposta: é muito fácil conseguir na Internet dados reais para serem analisados (livros, poemas, discursos, letras de músicas, etc.).

Neste trabalho apresentamos um relato de nossa experiência em adaptar e aplicar as atividades sugeridas por Bortolossi (2013) para duas turmas do Ensino Médio, uma turma do Ensino Técnico e uma turma de especialização em Matemática.

No Capítulo 2 apresentamos o perfil dos alunos participantes do trabalho (obtido por meio de um questionário de sondagem), descrevemos os aplicativos educacionais indicados por Bortolossi (2013) e apresentamos o planejamento inicial das atividades. No Capítulo 3, descrevemos quais foram as escolhas que tivemos que fazer ao longo do processo e as respostas dos alunos durante a realização das atividades. Para terminar, no Capítulo 4, apresentamos nossas considerações finais.

2 *O Contexto do Trabalho*

2.1 Instituições de Ensino

Este trabalho foi desenvolvido no segundo semestre de 2013 com a participação de alunos de duas instituições de ensino no Estado do Rio de Janeiro, a saber,

1. o Colégio Estadual Professor Renato Azevedo em Cabo Frio, com duas turmas de Resolução de Problemas Matemáticos do 2º Ano do Ensino Médio regular e uma turma do 2º Módulo (período) do Curso de Educação Profissional de Técnico em Informática;
2. a Universidade Federal Fluminense em Niterói, com uma turma de Análise de Dados do curso de especialização em Matemática para professores.

2.1.1 Colégio Estadual Professor Renato Azevedo

TURMAS DE RESOLUÇÃO DE PROBLEMAS MATEMÁTICOS

Como parte do Projeto de Currículo Mínimo iniciado no ano de 2013 pelo governo do Estado do Rio de Janeiro (SEE-RJ, 2013), a disciplina RPM (Resolução de Problemas Matemáticos) é oferecida aos alunos do 2º ano do Ensino Médio, normalmente dada no mesmo turno, com dois tempos seguidos por semana (cada tempo com 50 minutos). Se trata de uma disciplina obrigatória, mas que não reprova o aluno.

Duas turmas de RPM participaram do trabalho nos 3º e 4º bimestres de 2013: as de número 2001 (com 32 alunos) e 2002 (com 27 alunos). Apesar do tópico “Análise de Dados” estar indicado para ser desenvolvido nesta disciplina nos 3º e 4º bimestres do 2º ano do Ensino Médio (SEE-RJ, 2013, p. 17), constatamos que a maioria dos alunos das duas turmas tinham pouco domínio sobre certos conceitos básicos abordados em Estatística: verificamos que eles sabiam ler e interpretar diagramas estatísticos e reconheciam a média aritmética, mas desconheciam as definições de frequência, moda, mediana e desvio padrão. Na rede estadual do Rio de Janeiro, a disciplina Estatística é abordada como Tratamento da Informação ao longo dos quatro anos do Ensino Fundamental e como Introdução à Estatística no 2º bimestre do 3º ano do Ensino Médio

(o que mostra, para o caso do Ensino Médio, um descompasso entre o que sugere o Projeto de Currículo Mínimo e a distribuição curricular).

TURMA DO CURSO TÉCNICO DE INFORMÁTICA

O Curso de Educação Profissional de Técnico em Informática, com duração de três semestres, é oferecido pela Secretaria de Educação do Estado do Rio de Janeiro a alunos que já concluíram o Ensino Médio. No Colégio Estadual Professor Renato Azevedo em Cabo Frio, este curso é noturno e ele foi autorizado em 17 de agosto de 2004 nos termos das Resoluções SEE nº 2349 de 11.12.2000 e nº 2355 de 18.01.2001. A expedição de diplomas está sob a égide da Lei de Diretrizes e Bases da Educação Nacional nº 9394/96.

Pela disciplina semestral Estatística (obrigatória e passível de reprovação), os oito alunos da turma INFO-2 do 2º módulo (período) deste curso técnico participaram do trabalho nos 3º e 4º bimestres de 2013. Observamos que, destes oito alunos, metade havia concluído o Ensino Médio há mais de 20 anos e que, como no caso das turmas de Resolução de Problemas Matemáticos, os alunos tinham pouco domínio de conceitos básicos de Estatística como frequência, moda e mediana.

Nas Figuras 2.1 a 2.7, a seguir, resumimos as respostas a algumas das perguntas elaboradas em um questionário de sondagem inicial (Apêndice A). Observação: um dos oito alunos da turma INFO-2 não entregou o seu questionário.

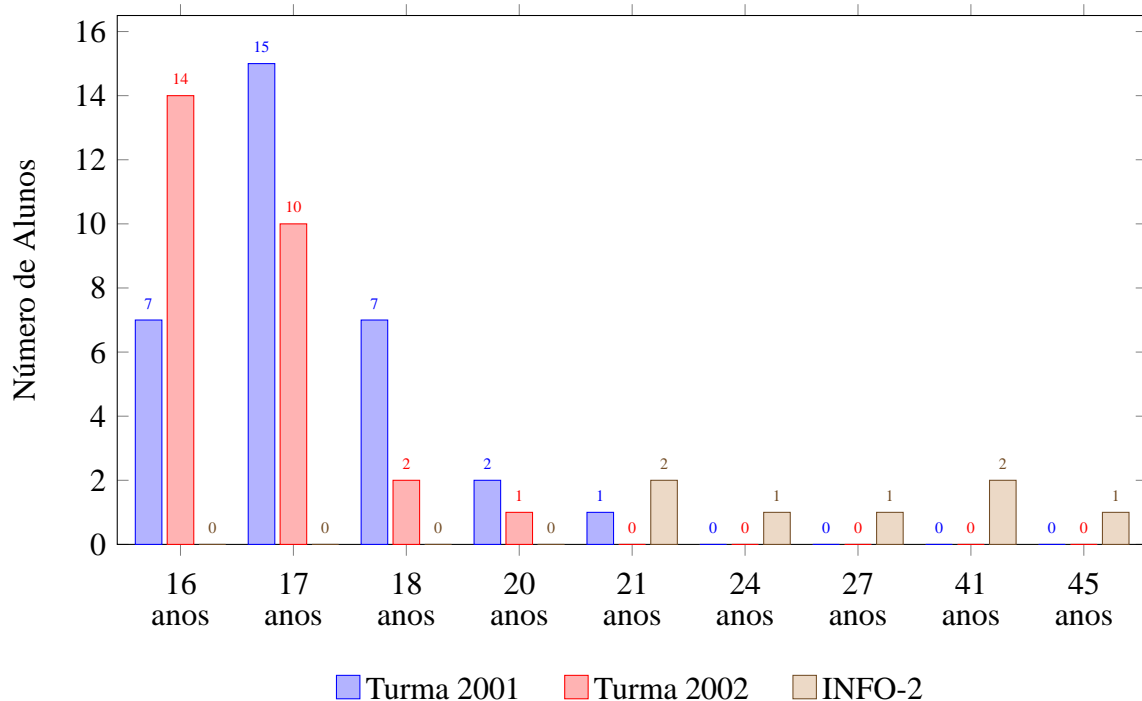


Figura 2.1: Idade.

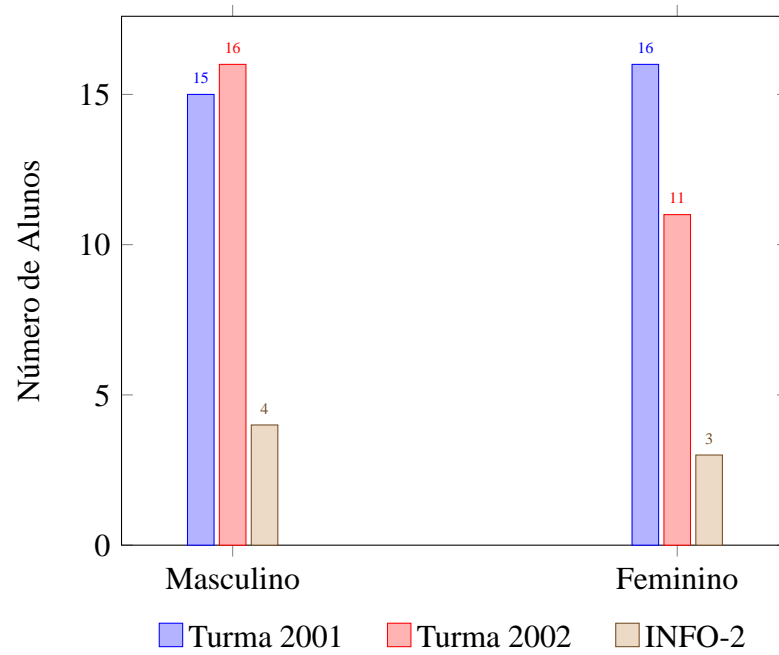


Figura 2.2: Sexo.

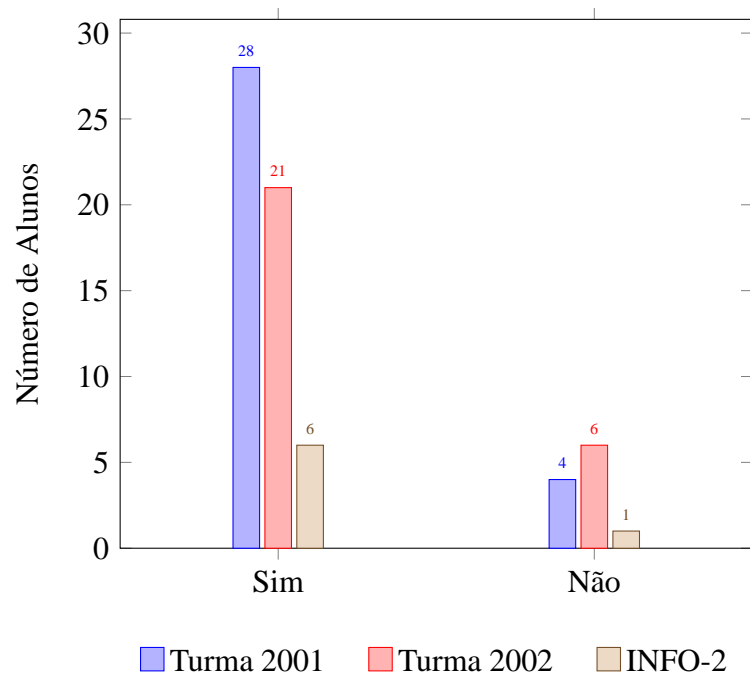


Figura 2.3: Possui computador em casa?

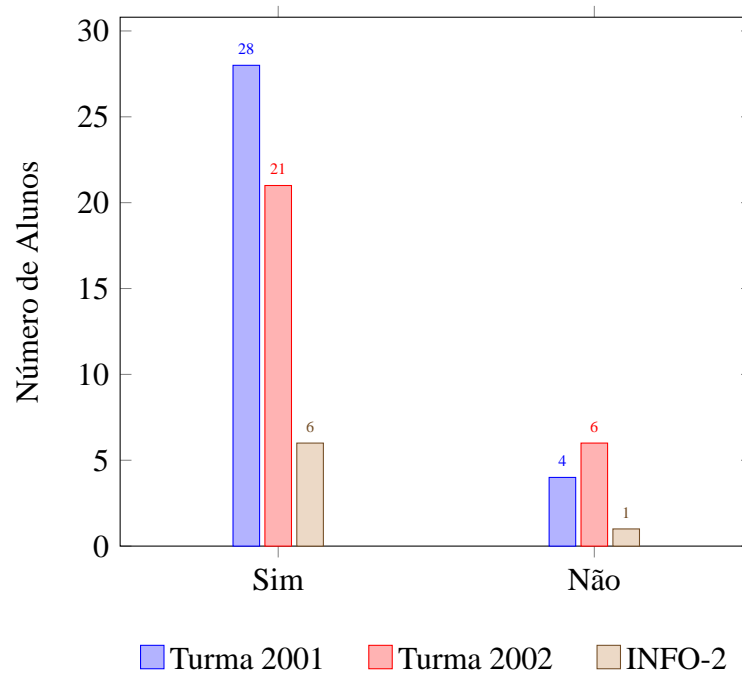


Figura 2.4: Possui acesso à Internet em casa?

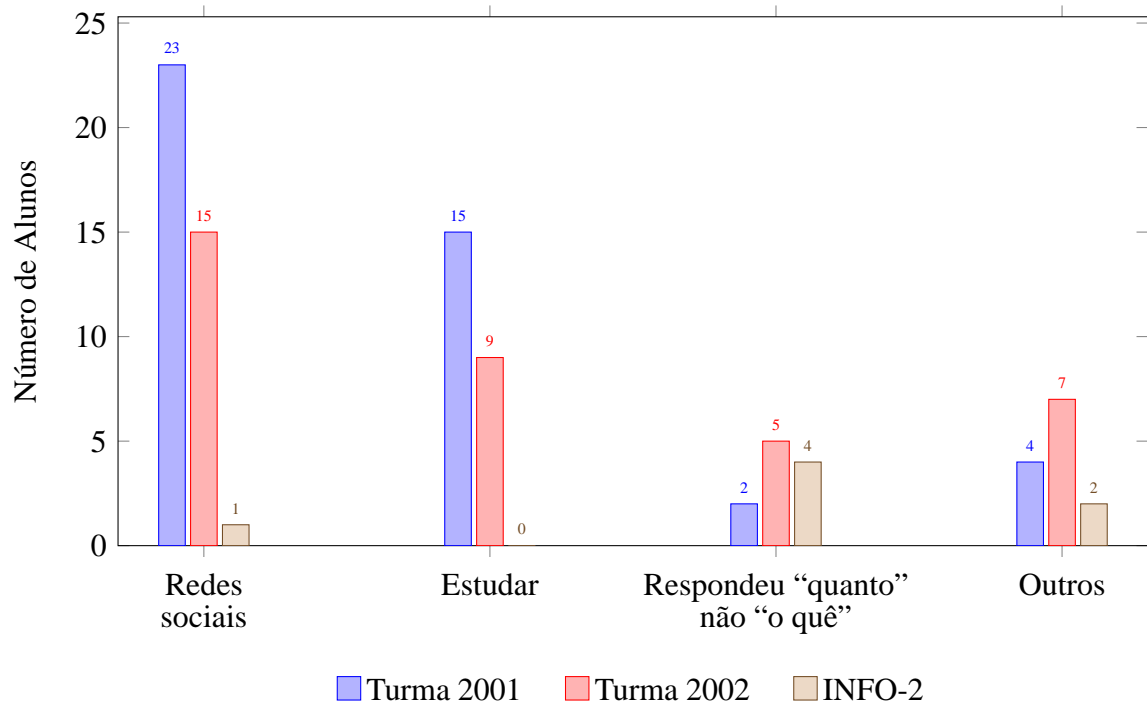


Figura 2.5: Qual é o uso mais frequente que faz do computador?

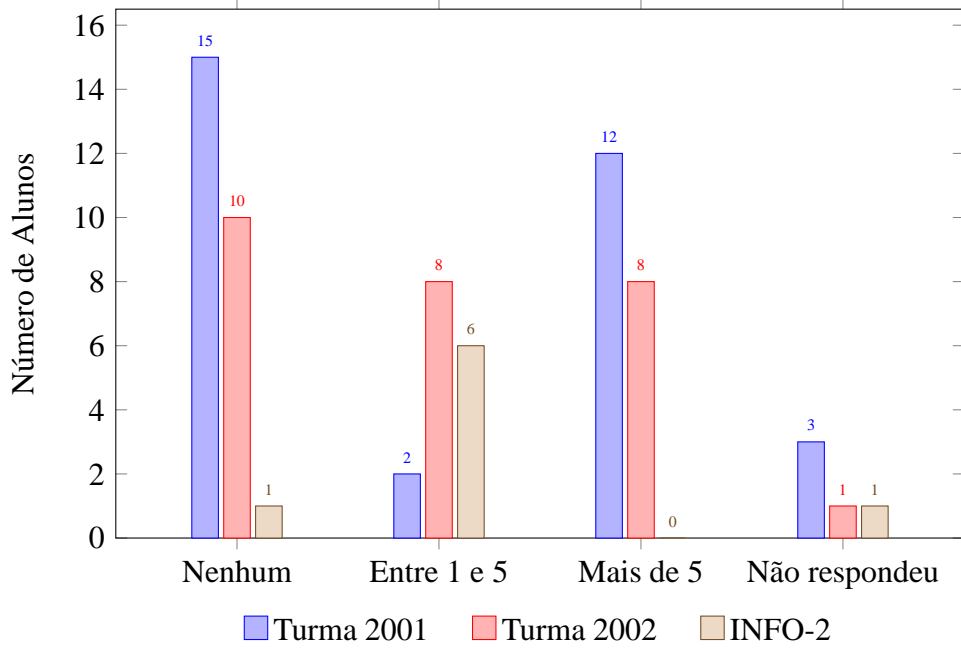


Figura 2.6: Quantos livros você costuma ler por ano?

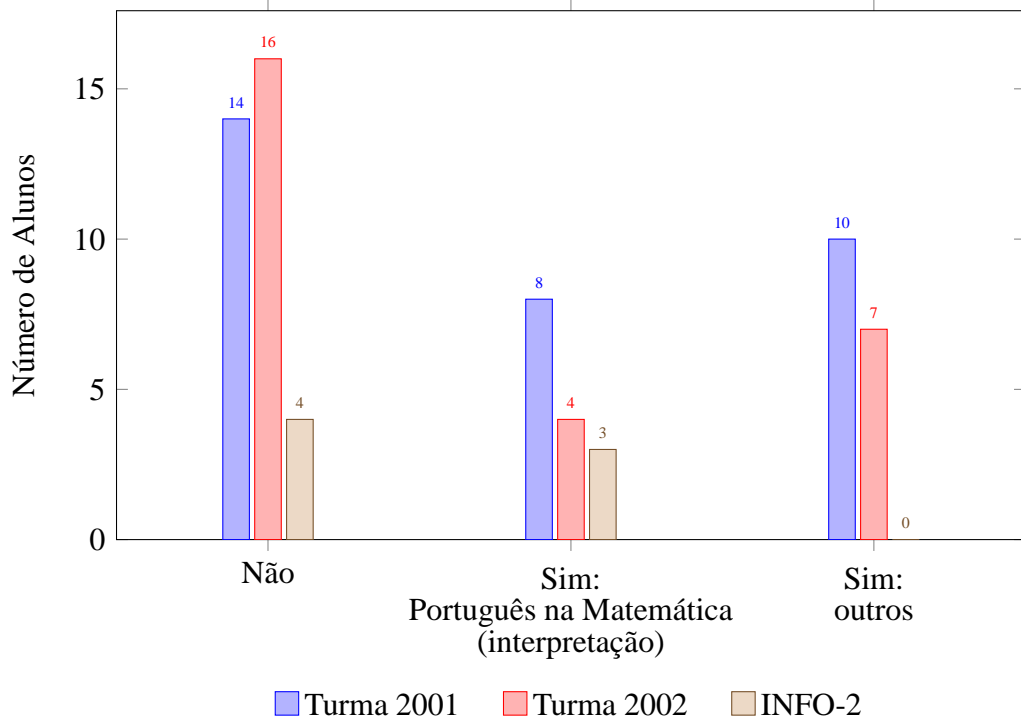


Figura 2.7: Você acha que existe alguma relação entre Matemática e Português? Qual?

Note, na Figura 2.5 que 11 alunos (cerca de 16% do total de alunos que responderam ao questionário) interpretaram de maneira equivocada a Pergunta 6: ao invés de qualificarem, eles quantificaram o uso do computador. Com relação à Figura 2.7, na categoria “Sim: outros”, destacamos algumas respostas dadas pelos alunos: “Sim, todas as duas são chatas.”, “As duas são matérias.”, “Sim, as duas utilizam letras.”, “Sim. São essenciais.”.

2.1.2 Universidade Federal Fluminense

Participaram também do desenvolvimento do trabalho 12 alunos da disciplina Análise de Dados do curso de especialização em Matemática para professores da Universidade Federal Fluminense. Esta é uma disciplina obrigatória do segundo semestre da grade curricular do curso. Observamos que os alunos deste curso são todos professores que atuam no Ensino Básico e que as atividades foram realizadas sob a supervisão do orientador Humberto Bortolossi, então professor da disciplina.

Uma versão simplificada e adaptada do questionário de sondagem do Apêndice A foi aplicada. Quando indagados sobre a existência de alguma relação entre Matemática e Português, todos os 12 alunos responderam que sim. Entre as justificativas dadas, destacamos: “Sim, pois para ler os textos matemáticos e resolver as questões matemáticas é preciso compreendê-los e isto se dá, primeiramente, pela interpretação de texto.”, “Sim. Matemática e Português são formas de expressão da comunicação e do pensamento.”, “Sim, pois na língua existem muitas palavras que denotam contagem (um, bis, muitos, algum, nenhum) e outras que denotam operações ou relações.”. Incluímos no questionário uma pergunta mais específica: “Você acha que existem aplicações da Matemática em Português? Em caso afirmativo, quais?”. Dos 12 alunos, 8 responderam que sim, 1 respondeu que não e 3 responderam que não sabiam. Entre as justificativas para a resposta sim, destacamos: “Sim. Na métrica das poesias.”, “Tanto em Português quanto em Matemática existem parábola, elipse e hipérbole. Já vi uma discussão que afirmava que essas palavras surgiram primeiro na Matemática e depois suas interpretações foram utilizadas no Português.”, “Sim, na Criptografia, um código matemático em Português.”.

2.2 Descrição dos Recursos Computacionais Usados

As atividades sugeridas por Bortolossi (2013) têm como eixo norteador o *software* gratuito LPP (Estatística das Letras, Palavras e Períodos) do *Projeto de Conteúdos Digitais para O Ensino e A Aprendizagem de Matemática e Estatística* (CDME) (<http://www.uff.br/cdme/>) do

Instituto de Matemática e Estatística da Universidade Federal Fluminense¹ (Bortolossi, 2009). Escrito na linguagem Java, para Windows, Linux e Mac OS, este aplicativo pode ser armazenado em um servidor para uso *on-line* pela Internet ou, alternativamente, ele pode ser baixado ou disponibilizado em CDs e *pen drives* para uso *off-line* (isto é, em um navegador, mas sem a necessidade de uma conexão com a Internet)². Seus cinco módulos principais compartilham uma interface gráfica muito simples: basicamente, existe uma área onde o usuário pode digitar um texto diretamente, abrir um arquivo de texto de seu computador ou usar a técnica de “copiar e colar” (ctrl+c/ctrl+v). Ao pressionar o botão “Processar!”, o texto fornecido será então processado e os resultados da análise disponibilizados numericamente e graficamente. Descreveremos agora, com mais detalhes, cada módulo que compõe o *software*:

- (1) O primeiro módulo do *software* LPP permite codificar/decodificar um texto usando a técnica de permutação das letras (isto é, as letras são trocadas pelo programa) (Figura 2.8). Para o processo de decodificação, está disponível um recurso que permite contar o número de letras do texto. A partir da distribuição de frequência resultante, o usuário pode estabelecer estratégias para descobrir a chave de decodificação.

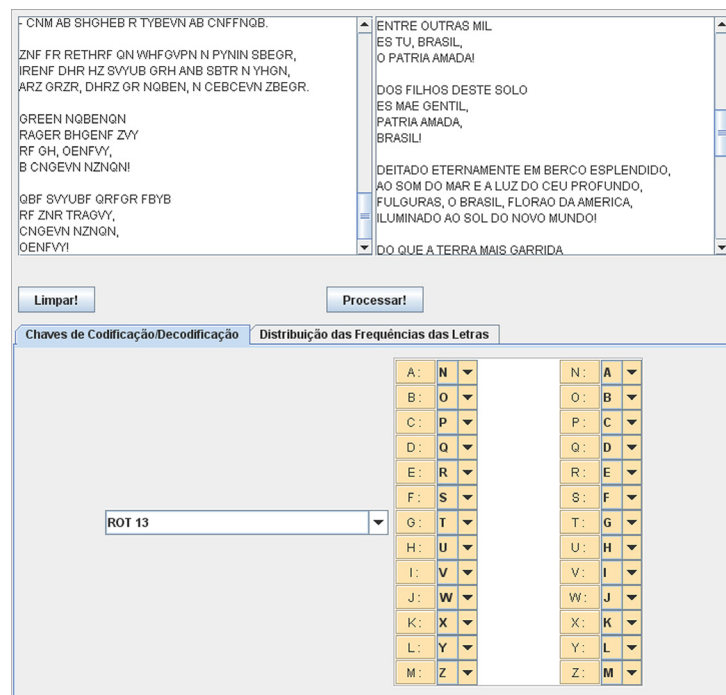


Figura 2.8: Módulo 1 (Estatística das Letras e Criptografia).

¹Este projeto contou com financiamento do Ministério da Educação e do Ministério da Ciência e Tecnologia pelo Edital de Conteúdos Digitais de 2007. A proposta do edital era o de apoiar a produção de recursos digitais gratuitos (*softwares*, vídeos, áudio e experimentos educacionais) para as áreas de Língua Portuguesa, Biologia, Química, Física e Matemática do Ensino Médio

²Com a atualização feita pela Oracle em janeiro de 2014, tanto o uso *on-line* quanto o *off-line* de *applets* Java ficou bem menos amigável, obrigando o usuário a fazer configurações especiais de acesso e a confirmar o uso cada vez que o aplicativo é executado.

- (2) O segundo módulo conta o número de letras, dígitos, acentos, sinais de pontuação, palavras e períodos de um texto (que pode estar em Português, Inglês, Espanhol, Francês ou Alemão). Ele também calcula o número de letras por palavra e o número de palavras por período (apresentando média, mediana, moda, variância e desvio padrão destas variáveis quantitativas), os períodos mais longos, os períodos mais curtos, as palavras mais longas e as palavras mais curtas. As informações das várias tabelas podem ser reordenadas clicando-se várias vezes no cabeçalho da coluna correspondente (Figura 2.9).

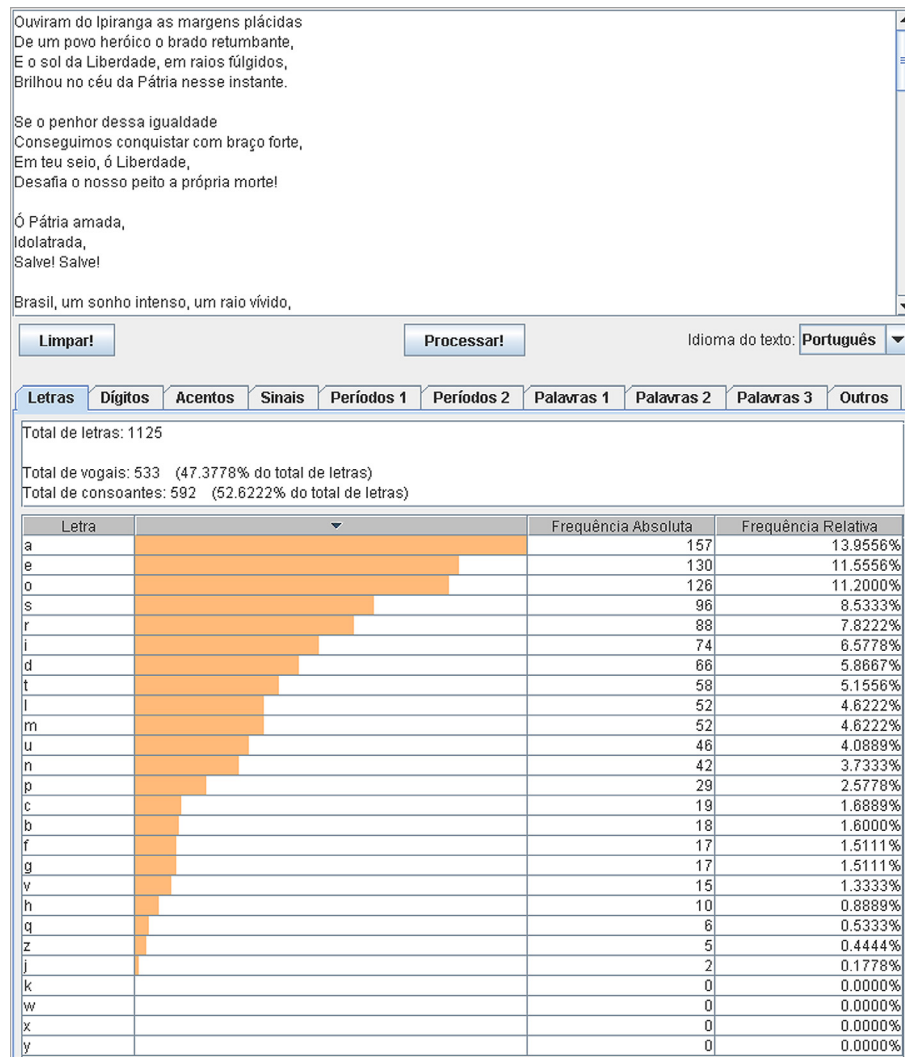


Figura 2.9: Módulo 2 (Estatística das Letras, Palavras e Períodos) com a distribuição de frequência das letras do Hino Nacional Brasileiro.

- (3) O terceiro módulo calcula a frequência das palavras de um texto oferecendo um ambiente conveniente para o estudo da Lei de Zipf. Esta lei, proposta pelo professor de linguística George Kingsley Zipf (1902-1950) da Universidade de Harvard, propõe que, em um texto com um número grande de palavras, a frequência f de ocorrência de uma palavra como

uma função da posição r deste evento em uma lista ordenada pela frequência de ocorrência é dada por uma lei de potências:

$$f \sim \frac{C}{r^a},$$

com C e a constantes, $a \sim 1$. A Figura 2.10 ilustra a Lei de Zipf para o romance “Memórias Póstumas de Brás Cubas” de Machado de Assis (com $C = 3689,775$ e $a = 0,925$). Nota: os eixos horizontal e vertical do gráfico apresentado na figura estão em escala logarítmica.

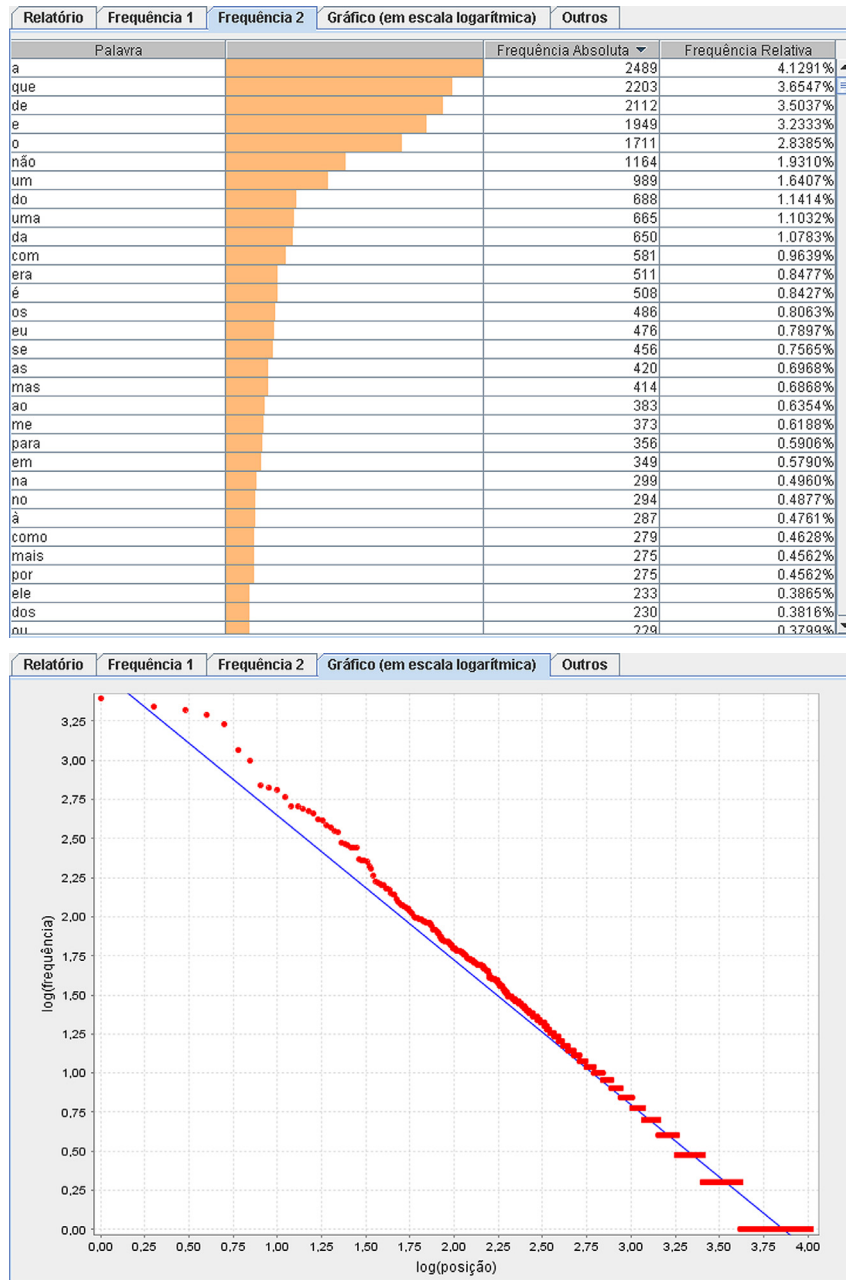


Figura 2.10: Módulo 3 (A Lei de Zipf).

- (4) O quarto módulo analisa o crescimento do vocabulário em um texto, isto é, ele exibe, por

meio de um gráfico, qual é o número de palavras diferentes em função do número de palavras lidas do início até o final do texto. A Figura 2.11 exibe o gráfico do número de palavras diferentes em função do número de palavras para o romance “Memórias Póstumas de Brás Cubas” de Machado de Assis: este texto tem 60279 palavras no total, com 10449 palavras diferentes (17,3344% do número total de palavras).

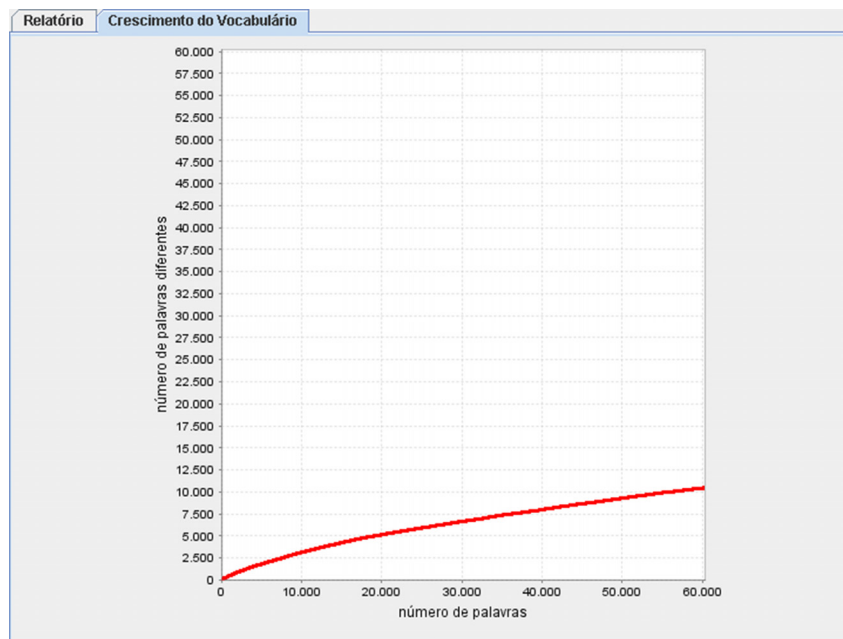


Figura 2.11: Módulo 4 (O Crescimento do Vocabulário em um Texto).

- (5) O quinto e último módulo do *software* LPP apresenta um recurso gráfico (usado principalmente na Internet) para descrever os termos mais frequentes de um determinado texto. O tamanho da fonte em que a palavra é apresentada é uma função da frequência da palavra no texto: palavras mais frequentes são desenhadas em fontes de tamanho maior, palavras menos frequentes são desenhadas em fontes de tamanho menor.

2.3 Descrição das Atividades

O desenvolvimento de todo o trabalho teve como fio condutor um conjunto de atividades adaptadas de Bortolossi (2009), divididas em três grupos, que descreveremos a seguir.

2.3.1 Grupo 1: Estatística e Criptografia

O objetivo principal deste grupo de atividades é mostrar como a distribuição de frequência das letras em textos de um determinado idioma pode ser usada para decodificar uma mensagem

que foi codificada por meio de uma cifra de César. Nesta técnica criptográfica, as letras do alfabeto são permutadas. Por exemplo, a letra “a” é trocada pela letra “b”, a letra “b” pela letra “c”, a letra “c” pela letra “d”, . . . , a letra “y” pela “z” e, por fim, a letra “z” pela letra “a”, de modo que o texto

A NOITE, VOVO KOWALSKY VE O IMA CAIR NO PE DO PINGUIM QUEIXOSO E
VOVO POE ACUCAR NO CHA DE TAMARAS DO JABUTI FELIZ

fica codificado, por esta permutação, como

B OPJUF, WPWP LPXBMTLZ WF P JNB DBJS OP QF EP QJOHVJN RVFJYPTP F
WPWP QPF BDVDBS OP DIB EF UBNBSBT EP KBCVUJ GFMJA.

Para auxiliar na identificação da permutação usada e, assim, decodificar a mensagem, determina-se a distribuição de frequência das letras do texto codificado³. Em Português, a letra mais frequente do texto codificado provavelmente é aquela associada à letra “a” pela permutação, uma vez que, em geral, é a letra “a” a letra mais frequente em textos longos em Português (é claro, isto pode não acontecer dependendo do tamanho e do estilo do texto). Do mesmo modo, a segunda letra mais frequente do texto codificado é aquela associada à letra “e”. Mais geralmente, segundo Quaresma (2008), as 16 letras mais frequentes são, em ordem decrescente de frequência, “a”, “e”, “o”, “s”, “r”, “i”, “d”, “n”, “m”, “u”, “t”, “c”, “l”, “p”, “v” e “h”.

Com o objetivo de motivar as atividades computacionais, planejamos uma aula expositiva inicial (apresentada com o uso de um projetor multimídia) com o seguinte conteúdo: o significado de criptografia e criptologia, aspectos históricos, aplicações e duas técnicas antigas de criptografia (a cítala espartana e a cifra de César). Para compor os *slides* da apresentação, usamos como referência principal o material produzido por Bezerra, Malagutti & Rodrigues (2010).

Todas as atividades computacionais deste grupo são desenvolvidas tendo como base o Módulo 1 do *software* LPP em um estudo dirigido cujo roteiro é apresentado no Apêndice B.

O objetivo da Atividade [01] é fazer com que o aluno perceba que a distribuição de frequência é uma característica do idioma e que esta distribuição pode mudar dependendo do tamanho e da qualidade da amostra. Na Atividade [02], o aluno é desafiado a decodificar um texto em Português criptografado com uma cifra de César. O texto em questão é a letra da música

³Este método de codificação é explicado e faz parte da trama do conto “O Escaravelho de Ouro” do escritor americano Edgar Allan Poe (1809–1849).

“Garota de Ipanema” composta por Vinícius de Moraes e Antônio Carlos Jobim. A permutação que decodifica o texto é a seguinte:

A	↔	D
B	↔	C
C	↔	B
D	↔	A
E	↔	H
F	↔	G
G	↔	F
H	↔	E
I	↔	L
J	↔	M
K	↔	
L	↔	I
M	↔	V
N	↔	Q
O	↔	P
P	↔	O
Q	↔	N
R	↔	U
S	↔	T
T	↔	S
U	↔	R
V	↔	
W	↔	
X	↔	
Y	↔	J
Z	↔	Z

As Atividades [04] a [06] procuram conduzir o aluno a compreenderem um pouco mais a Matemática envolvida na cifra de César. Mais especificamente: espera-se, com a Atividade [04], que o aluno, ao interagir com o *software*, perceba que as permutações do tipo ROT⁴ são permutações circulares; a Atividade [05] indaga o aluno sobre quantas chaves de codificação/decodificação diferentes podem ser construídas com uma cifra de César (o que equivale a perguntar quantas permutações de 26 letras diferentes existem); o objetivo da Atividade [06] é investigar quais são as cifras de César cuja chave de decodificação é igual a própria chave de codificação usada (o que equivale a perguntar quais permutações de 26 letras diferentes são involuções).

A Atividade [07] pretende verificar se o aluno percebeu que em textos mais longos a distribuição de frequência tem mais chance de se aproximar daquela descrita por Quaresma (2008) e que, conseqüentemente, tais textos seriam mais fáceis de serem decodificados usando-se a distribuição de frequência das letras.

Finalizamos com a Atividade [08], que pede para que os alunos comentem as atividades realizadas neste primeiro grupo.

⁴Se $L(0) = A, L(1) = B, \dots, L(25) = Z$, então para cada inteiro $k \geq 0$, permutação ROT k é definida por ROT $k(L(i)) = L((i+k) \bmod 26)$, com $0 \leq i \leq 25$. A permutação ROT₁₃ foi usada no início da Internet principalmente em fóruns e grupos de discussão.

2.3.2 Grupo 2: Atividades Mais Abrangentes

O objetivo principal deste grupo de atividades é estender o escopo da análise estatística para palavras e períodos, incluindo medidas de posição e dispersão das variáveis quantitativas, bem como estilos literários relacionados com a frequência e disposição das letras em textos.

Todas as atividades computacionais deste grupo são desenvolvidas tendo como base o Módulo 2 do *software* LPP em um estudo dirigido cujo roteiro é apresentado no Apêndice C. Logo na Atividade [01] foram fornecidos seis textos para que os alunos determinassem estatísticas das letras, acentos e sinais de pontuação, o número total de palavras e o número total de palavras diferentes. Estes textos foram selecionados intencionalmente de séries escolares diferentes com o propósito de levar os alunos a perceberem, por meio de uma análise estatística, como certas características do texto mudam de acordo com a série (Atividade [02]). Este tipo de exercício é sugerido pelo documento GAISE (Franklin, 2007, p. 16).

Na Atividade [03] foi pedido para que os alunos, com o auxílio do *software*, encontrassem, caso existisse, uma palavra em Português com o número de consoantes superior a 70% e outra superior a 90% do número total de letras da palavra. Não poderiam ser utilizadas siglas, algarismos romanos, onomatopeias e palavras incorporadas de outros idiomas. Essa atividade tem por objetivo fazer com que o aluno reflita sobre como as palavras do nosso idioma se compõem com relação à distribuição de vogais e consoantes.

Em Computação, *stop words* são palavras sem significado (como artigos, pronomes e proposições) que são filtradas antes ou depois de certos tipos de processamento de um texto em linguagem natural (como, por exemplo, o de identificar palavras-chaves de um texto). Na Atividade [04] foi solicitado aos alunos que descobrissem *stop words* mais frequentes do Português analisando textos longos diversos.

As Atividades [05], [06] e [09] foram planejadas com o intuito de colocar a média aritmética e o desvio padrão em um contexto linguístico: a média do número de letras por palavra da frase na Atividade [05] é igual a 4 e o desvio padrão é igual a 0. Na Atividade [06], a média também é igual a 4, mas o desvio padrão é igual a 3,08220. Na Atividade [09] os alunos são desafiados a comporem um texto longo cuja média do número de letras por palavra seja exatamente igual a 4 (com qualquer desvio padrão).

Um *lipograma* é uma composição literária que se caracteriza pela omissão deliberada de determinadas letras do alfabeto em seu texto (isto é, estas letras têm frequência zero no texto). O desafio aqui é criar textos onde as letras mais frequentes de um determinado idioma não são usadas. É fácil escrever um texto longo em Português sem as letras “k”, “w” e “y”, mas é

difícil escrever um texto longo em Português sem a letra “a”⁵. Um *pangrama*, por sua vez, é um período onde são usadas todas as letras do alfabeto de uma determinada língua. Pangramas são usados para testar fontes de computador e letreiros luminosos. Aqui está um em Português: “À noite, vovô Kowalsky vê o ímã cair no pé do pinguim queixoso e vovó põe açúcar no chá de tâmaras do jabuti feliz”. As Atividades [07] e [08] pedem para que os alunos tentem construir, respectivamente, um lipograma e um pangrama.

As Atividades [10], [11] e [12] tratam da disposição das vogais em palavras e orações. Mais especificamente, a Atividade [10] pede para que os alunos escrevam uma palavra onde as cinco vogais estejam todas presentes. Na Atividade [11], inspirados pela palavra

Honorificabilitudinitatibus

(uma palavra em latim usada por William Shakespeare (1564-1616)) que alterna consoantes e vogais, os alunos são convidados a pensar em palavras do Português com a mesma propriedade. A Atividade [12] pergunta para os alunos o que a seguinte oração “Dama do Lago, apenas você sabe como atar veterana medusa!” tem de especial. Note que em todas as palavras desta oração as consoantes e vogais aparecem alternadamente.

Finalizamos com a Atividade [13], que pede para que os alunos comentem as atividades realizadas neste segundo grupo.

2.3.3 Grupo 3: Análise Estatística por Autor

Neste último grupo, propusemos aos alunos que se organizassem em equipes e que cada equipe usasse o Módulo 2 do *software* LPP para analisar estatisticamente obras diferentes de um mesmo autor brasileiro. O objetivo é investigar se autores diferentes possuem estatísticas diferentes ou não. Para esta última etapa, os alunos usaram a ficha disponível no Apêndice D.

2.4 Cronograma Simplificado

As atividades para as Turmas de RPM e INFO-02 foram planejadas para serem realizadas durante o segundo semestre de 2013, com a divisão descrita na tabela a seguir.

⁵Um dos expoentes na arte de lipogramas foi o espanhol Alonso de Alcalá y Herrera (1599-1682), que escreveu cinco novelas curtas em forma de lipograma, cada uma sem uma das cinco vogais: *Los Dos Soles de Toledo*, *La Carrozza con las Damas*, *La Perla de Portugal*, *La Peregrina Ermitaña* e *La Serrana de Sintra*. Na França temos Georges Perec (1936-1982) com seu lipograma *La Disparition* sem a letra “e” (a letra mais frequente do Francês). Gilbert Adair (1944-2011), em *A Void*, traduziu *La Disparition* para o Inglês e, nesta tradução, também omitiu a letra “e” (a letra mais frequente do Inglês).

Grupo	Número de Semanas
1	2
2	4
3	3

Para a Turma da Especialização da UFF, as atividades foram aplicadas nas três últimas semanas do semestre.

2.5 Laboratórios de Informática

Para as Turmas 2001 e 2002, o laboratório de informática (Figura 2.12) era constituído por 13 CPUs e 26 monitores (cada CPU com 2 monitores), um computador exclusivo do professor e um projetor multimídia. As CPUs rodavam o sistema operacional Linux Educativo com acesso à Internet (com alguns *sites* bloqueados).

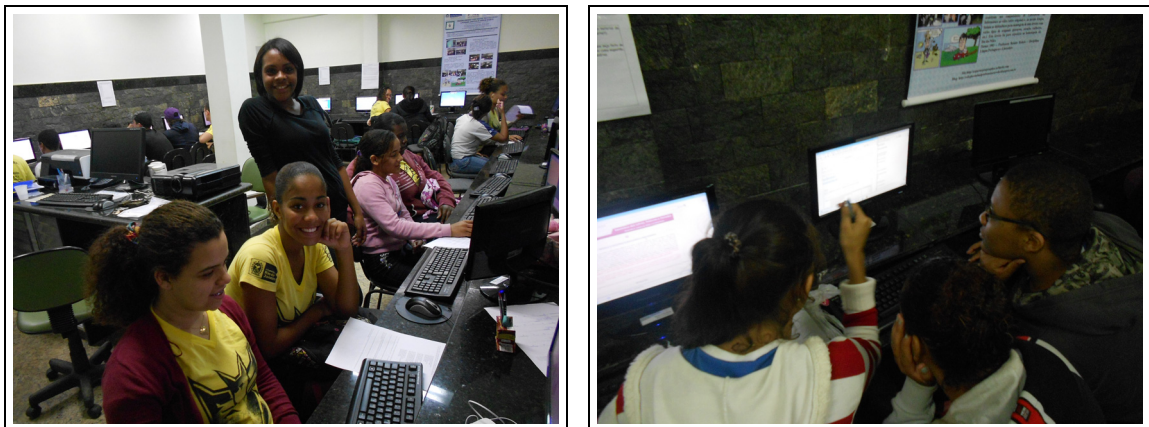


Figura 2.12: Alunos trabalhando no laboratório de informática da escola.

Para a Turma INFO-2, o laboratório de uso exclusivo tinha 18 computadores (cada um com seu monitor). Os computadores rodavam o sistema operacional Windows XP com acesso à Internet.

Para a Turma da Especialização, o laboratório era constituído por 24 computadores (cada um com seu monitor) e um projetor multimídia. Os computadores rodavam o sistema operacional Linux (Kubuntu) com acesso à Internet.

3 *Relato de Experiência*

3.1 Grupo 1: Estatística e Criptografia

As atividades nas Turmas 2001 e 2002 foram realizadas em grupos, mais precisamente, foram doze grupos (cinco duplas, seis trios e um quarteto) na Turma 2001 e nove grupos (um individual, quatro duplas, dois trios e dois quartetos) na Turma 2002. Já nas turmas INFO-2 e Especialização as listas de atividades foram respondidas individualmente.

Os alunos das Turmas 2001, 2002 e INFO-2 tiveram, uma semana antes da aplicação das atividades do Grupo 1, um tempo de aula (50 minutos) de ambientação no uso do *software*. Durante esta etapa de preparação e experimentação livre, aproveitamos também para alertar os alunos para o fato de que a distribuição da frequência de letras do texto é sensível ao seu tamanho: por exemplo, quanto maior for o texto escolhido, maior será a chance de que a letra “A” seja a mais frequente.

Atividade [01] (distribuição de frequência de textos em Português)

Por ser uma atividade bem simples e sistemática (processar os textos com o *software* e anotar alguns dados em uma tabela), a maioria dos alunos conseguiu fazer o que foi proposto. Do total de 164 tabelas preenchidas nessa atividade, 43 continham algum erro em seu preenchimento. Muito dos erros cometidos foram equívocos na hora de preencher a tabela, como: falta de atenção ao selecionar o texto corretamente (alguns alunos selecionaram a página *web* inteira onde se encontrava o texto e não apenas o texto), preenchimento errado da distribuição das frequências das letras (por não entenderem como teriam que preencher essa parte da tabela ou por terem selecionado o texto de maneira errada) e omissão de alguns dados do texto escolhido (por não colocarem o nome do autor ou o título do texto). Nessa atividade, ocorreram também pequenas variações no número total de letras do texto e a substituição do nome do compositor de uma letra de música pelo nome de uma banda que a interpreta. Encontramos quinze (de um total de quarenta e oito) tabelas com equívocos dessa natureza na Turma 2001, seis (de um total de 36) tabelas na Turma 2002 e 19 (de um total de 32) tabelas na Turma INFO-2. Um aluno da

Turma 2002 e um aluno da Turma INFO-2 escolheram textos com o idioma errado (não eram textos em Português). Na Turma da Especialização, tivemos apenas três erros: duas tabelas com erros no nome do autor do texto escolhido (em um deles o nome foi omitido) e em outra esqueceram de informar o total do número de letras. Essa turma produziu, no total, 52 tabelas.

Com relação às escolhas dos textos, podemos destacar o seguinte: apenas os alunos da Turma da Especialização se preocuparam em analisar textos mais longos como livros e notícias publicadas na *web*. É desta turma, por exemplo, o texto mais longo analisado: o livro “Os Lusíadas” de Luís de Camões com 246998 letras (Figura 3.1).

TEXTO 1	
Título do texto:	Os Lusíadas
Autor do texto:	Luís de Camões
Idioma do texto:	Português
Número total de letras do texto:	261876 246998
← Mais frequente	
Cinco letras mais frequentes:	A E O S R N I T D M U C L P V Q G F H B J Z X Y W K
Menos frequente →	

Figura 3.1: Letras mais frequentes do livro “Os Lusíadas” de Luís de Camões.

Os alunos do Ensino Médio (Turmas 2001 e 2002) e da Turma INFO-2 optaram por letras de músicas (a grande maioria) e poesias resultando, portanto, em textos mais curtos. O texto mais longo coletado pela Turma 2001 foi “Vagalumes” de Ivo Mozart, Adriel de Menezes e Luiz Tomim (música tocada pela banda musical Pollo) com 1411 letras, na Turma 2002 foi “Carta de Amor” de Maria Bethânia com 2503 letras e, na Turma INFO-2, foi “Faroeste Caboclo” de Renato Russo com 5140 letras.

TEXTO 2	
Título do texto:	LEGIÃO URBANA
Autor do texto:	RENATO RUSSO
Idioma do texto:	Português
Número total de letras do texto:	5140
← Mais frequente	
Cinco letras mais frequentes:	A E O R I
Menos frequente →	

Figura 3.2: Letras mais frequentes da música “Faroeste Caboclo” do Renato Russo.

A Tabela 3.1 exibe qual foi a letra mais frequente nos diversos textos analisados pelas quatro turmas. Observamos que, como esperado, os textos mais longos (como aqueles analisados pela Turma da Especialização) tiveram a letra “a” como sendo a mais frequente. Uma particularidade que ocorreu e que gostaríamos de mencionar foi a análise feita por um grupo da Turma 2002 da

letra da música “Pela Luz dos Olhos Teus” de Vinícius de Moraes. Este texto possui 388 letras, sendo a letra “o” a mais frequente com 54 ocorrências, enquanto que a letra “a” fica apenas na quinta colocação entre as mais frequentes com 33 aparições (fato este que pode ser justificado pelo uso frequente das palavras “olhos” e “dos”).

Letra mais frequente	Turma 2001	Turma 2002	INFO-2	Especialização
a	em 13 textos	em 9 textos	em 8 textos	em 26 textos
e	em 18 textos	em 10 textos	em 19 textos	em 19 textos
o	em 3 textos	em 4 textos	–	em 4 textos

Tabela 3.1: Letras mais frequentes nos textos da Atividade 1.

Atividade [02] (decodificação de um texto criptografado com uma Cifra de César)

Com exceção da Turma da Especialização, essa foi uma atividade que requereu bastante paciência e atenção dos alunos. Muitos deles, quando se depararam com o texto criptografado, começaram a tentar achar a chave de decodificação sem estratégia alguma, por tentativa e erro. Alguns alunos tiveram a ideia de tentar adivinhar as letras de palavras pequenas como artigos, pronomes e preposições. Nas Turmas 2001, 2002 e INFO-2 foram poucos alunos (ou grupos) que atentaram para a ideia de recorrer à tabela de distribuição de frequência das letras da Língua Portuguesa (dada na semana anterior) para iniciar a decodificação do texto. Frente a esta dificuldade, foi necessário fazer uma intervenção e, assim, sugerimos explicitamente para os alunos que utilizassem a tabela de frequências para tentar decodificar a mensagem criptografada. Mesmo assim, três alunos da Turma 2001 e 4 alunos da turma INFO-2 não responderam a essa questão. O fato de muitos grupos das Turmas 2001 e 2002 e alunos da Turma INFO-2 terem escolhido textos relativamente pequenos na Atividade 1 (o que resultou em um grande número de textos com predominância da letra “e”) pode ter atrapalhado na escolha adequada de uma chave de decodificação.

Os alunos da Turma da Especialização conseguiram decodificar o texto com facilidade. Eles utilizaram inicialmente a distribuição das letras, mas, depois, eles começaram a decodificar as palavras menores com uma ou duas letras (estratégia semelhante a feita por alguns grupos e alunos das Turmas 2001, 2002 e INFO-2).

Uma vez que o texto criptografado é bem conhecido (a letra da música “Garota de Ipanema” de Vinícius de Moares), observamos que alguns alunos conseguiram identificá-lo logo após descodificarem apenas algumas poucas letras.

Atividade [03] (envio de mensagens criptografadas por e-mail)

Essa atividade não foi obrigatória. A intenção aqui era apenas para divertimento dos alunos.

Na Turma da Especialização, apenas um par de alunos realizou a atividade. Nas demais turmas, não temos registro de que algum aluno a tenha realizado.

Atividade [04] (permutações do tipo ROT)

Os alunos das Turmas 2001, 2002 e INFO-2 mostraram uma certa dificuldade em descrever de forma clara e objetiva como atuam as permutações do tipo ROT, apesar de tais rotações terem sido trabalhadas na aula anterior de ambientação no uso do *software* (Figuras 3.3 e 3.4). Mesmo alguns alunos da Turma da Especialização tiveram este tipo de dificuldade (Figura 3.5). Não obstante, ainda de maneira informal, alguns alunos conseguiram dar uma resposta adequada, em nossa opinião (Figura 3.6).

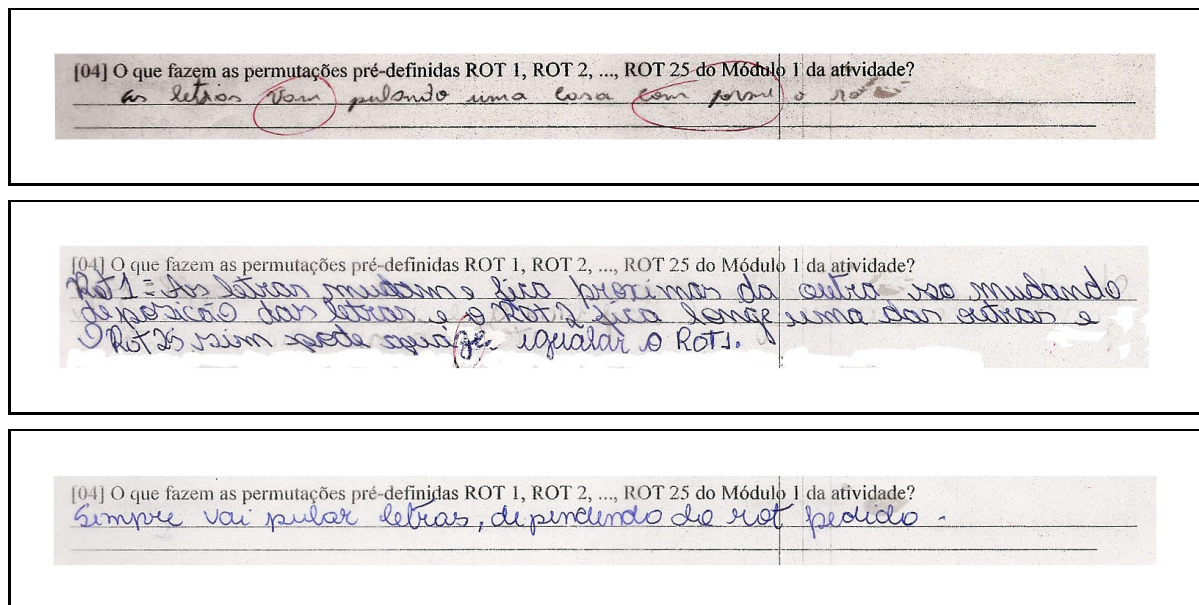


Figura 3.3: Algumas respostas dadas para a Atividade [04] das Turmas 2001 e 2002.

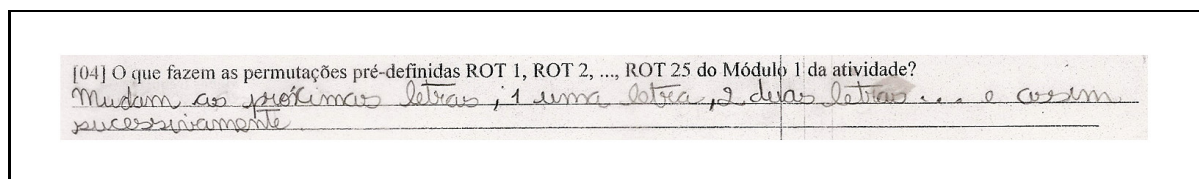


Figura 3.4: Resposta dada para a Atividade [04] da Turma INFO-2.

Atividade [05] (número de permutações das 26 letras)

Essa foi a atividade com o maior número de erros. Dos doze grupos da Turma 2001, oito deram como resposta " $25 \times 24 \times \dots \times 3 \times 2 \times 1$ " (isto é, $25!$ ao invés de $26!$), um grupo deu 27 como resposta, dois deram " $A = 25, B = 24, C = 23, \dots, Y = 1, Z = 0$ " e um grupo deixou a atividade em branco (assim, nenhum grupo acertou a atividade). Dos nove grupos da Turma 2002, quatro grupos acertaram a questão, um grupo deu como resposta " $26 \times 26 \times 24 \times 23 \times$

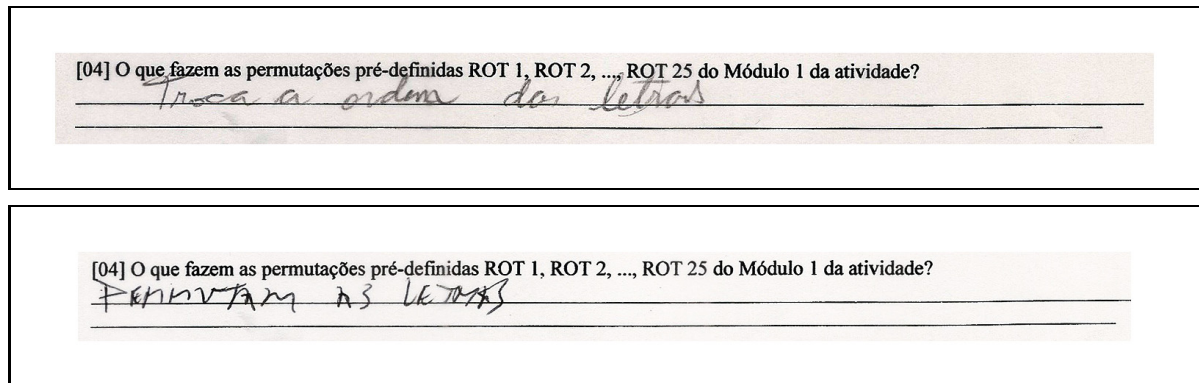


Figura 3.5: Algumas respostas dadas para a Atividade [04] da Turma da Especialização.

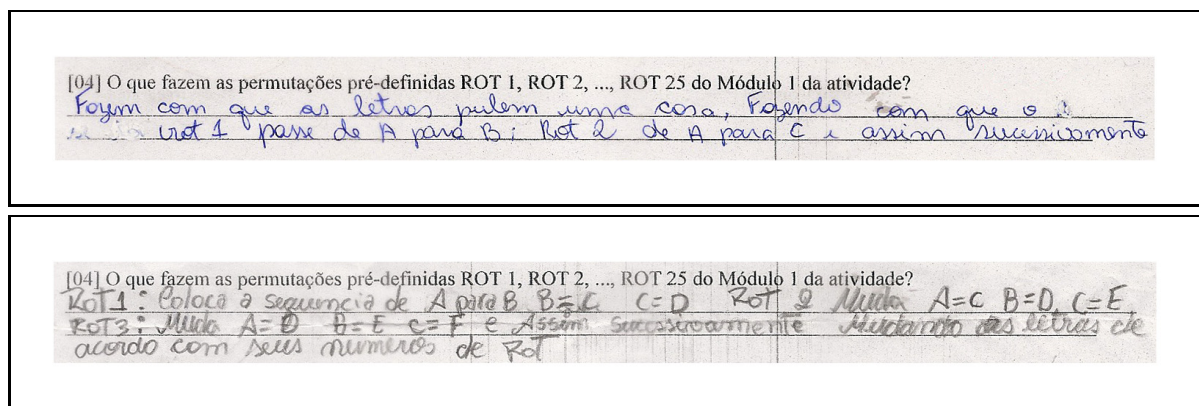


Figura 3.6: Algumas respostas mais elaboradas para a Atividade [04] da Turma 2001 (primeira imagem) e da Turma INFO-2 (segunda imagem).

$\dots \times 3 \times 2 \times 1$ ”, dois escreveram “26, 25, 24, 23, 22, ..., 3, 2, 1”, outro escreveu “ $26 > 25 > 24 > \dots > 3 > 2 > 1$ ” e outro escreveu “ $A = 25, B = 24, C = 23, \dots, Y = 1, Z$ ”.

Na Turma INFO-2 (formada por alunos que já concluíram o Ensino Médio), todos os oito alunos deram respostas erradas: “25, 24, 23, ..., 3, 2, 1”, “A tem 25 combinações, B tem 24 combinações, ...”, “24, 25, 26, 27, ...”, “24/25/26/27”, “26 a mesma quantidade de letras”, “24 – 25 – 26 – ...”.

Na Turma da Especialização (formada por alunos que já concluíram o Ensino Superior), um aluno deu como resposta “26!”, oito deram “ $26! - 1$ ” (pois não consideraram ROT0, a transformação identidade, como uma permutação, uma vez que, neste caso, não há troca de letras), um escreveu “troca a ordem das letras”, outro respondeu “ $26(25 + 24 + 23 + \dots + 2 + 1)$ ”, outro escreveu “cada letra pode assumir $23 \cdot 23^{23}$ possibilidades”, outro escreveu “28”.

Por que ocorreram tantos erros em uma questão de combinatória tão básica e padrão? Para o caso dos alunos do 2º Ano do Ensino Médio das Turmas 2001 e 2002, uma justificativa talvez seja a de que, pelo Currículo Mínimo 2012 do Estado do Rio de Janeiro (SEEDUC-RJ, 2012),

o tópico Análise Combinatória é abordado somente no 3º Ano do Ensino Médio. Na Turma INFO-2, por sua vez, havia alunos que terminaram o Ensino Médio há mais de 15 anos. Frente a essa dificuldade, foi necessária, durante a execução dessa atividade, outra intervenção com o objetivo de auxiliar os alunos a construírem suas respostas corretamente (foi feita uma breve menção sobre o princípio fundamental da contagem e alguns exemplos de permutações de letras com palavras menores).

Atividade [06] (permutações que são involuções)

Este exercício pede para os alunos listarem algumas permutações cujas inversas são as próprias permutações, isto é, permutações que são *involuções*. Apesar de existirem

$$26! \cdot \sum_{k=0}^{13} \frac{1}{(n-2 \cdot k)! \cdot 2^k \cdot k!} = 532\,985\,208\,200\,576$$

involuções de 26 letras (Pemmaraju & Skiena, 2003, p. 105), a ideia aqui é que os alunos listassem as involuções pré-definidas (ROT13 e DUPLA).

Os alunos das Turmas 2001, 2002 e INFO-2 tiveram dificuldades em interpretar o que era solicitado (e, conseqüentemente, em executar a tarefa). A única ajuda dada aqui foi a própria sugestão do enunciado da atividade: a exemplo da permutação REVERSA, procurar por uma permutação cuja chave de decodificação é a própria permutação. Pensávamos que tal dica seria suficiente para uma compreensão do que se queria.

Os resultados obtidos para esta atividade foram os seguintes: na Turma 2001 tivemos oito respostas corretas (todos responderam “ROT13”) e quatro erradas (“Não, se eu por o ROT5 e tenho que diminuir a quantidade de letra (sic) do alfabeto por 5, por o resultado e reverter.”, “Não”, “A25, B24, C23, D22, etc.”); na Turma 2002 oito grupos erraram (“Não”, “Não, somente o REVERSO”, “Só o REVERSO”, “so o reverso fas isto” (sic), “não, pois a cada ROT pula exatamente um letra e assim a diante” (sic), “somente o REVERSO que volta para o seu texto original”) e um grupo deixou a atividade em branco. Na Turma INFO-2 um aluno respondeu “Se você processar com ROT1, copiar e colar, e processar com o ROT25, você terá o texto original.” (não é a resposta correta, mas mostra que o aluno percebeu uma propriedade da composição de permutações), outro aluno simplesmente descreveu o que faz a permutação REVERSO e seis responderam “ROT0” (que está correta).

Na Turma da Especialização foram dez respostas (ROT0, ROT13, DUPLA, ROT26), dois alunos responderam “ROT0 e ROT25” e um aluno respondeu “Sim, basta aplicar a rotação ROT1 25 vezes”.

Atividade [07] (decodificação de textos pequenos e grandes: o que é mais fácil?)

Na Turma 2001 tivemos uma preferência de oito a quatro para os textos pequenos; na Turma 2002, foi de cinco a quatro para os textos pequenos; na Turma INFO-2 tivemos quatro alunos preferindo os textos pequenos, três os grandes e um aluno respondeu que a dificuldade é a mesma. Por fim, na Turma da Especialização, o resultado foi de onze a dois para os textos considerados grandes. Como uma justificativa não foi exigida nessa atividade, muitos alunos deram apenas suas opiniões sem mais nenhum comentário adicional. Isso aconteceu em sete dos doze grupos da Turma 2001, em oito dos nove grupos da Turma 2002, em cinco dos oito alunos da Turma INFO-2 e em três dos treze alunos da Turma da Especialização.

Entre os que justificaram suas escolhas pelos textos pequenos destacamos: "... é mais fácil de identificar." (Turma 2002) e "... porque tem poucas letras." (Turma 2001). Entre as justificativas para a escolha por textos grandes, destacamos: "... pois teremos mais precisão a descobrir as letras que aparecem mais." (Turma 2001), "Porque fica bem mais fácil o entendimento." (Turma INFO-2), "... pois fica mais fácil perceber as regularidades." (Especialização), "... porque as palavras se repetem mais." (Especialização) e "... pois possuem uma frequência de letras mais fiel ao idioma." (Especialização).

Atividade [08] (comentários dos alunos)

As Figuras 3.7, 3.8, 3.9 e 3.10 ilustram, respectivamente, algumas das respostas dadas para esta atividade pela Turma 2001, Turma 2002, Turma INFO-2 e Turma da Especialização.

3.2 Grupo 2: Atividades Mais Abrangentes

Antes de dar início às atividades do Grupo 2, foi feito um trabalho nas Turmas 2001, 2002 e INFO-2 envolvendo os conceitos de média, variância e desvio padrão. Em dois tempos de cinquenta minutos foram apresentadas as respectivas definições e fórmulas, assim como exemplos e exercícios. Julgamos essencial essa aula para as Turmas 2001 e 2002, pois esta era a primeira vez que os alunos dessas turmas estavam aprendendo esses conceitos (na Turma INFO-2 esse conteúdo já estava previsto na grade curricular). Também foi realizado, na semana seguinte, em dois tempos de aula, um primeiro contato dos alunos dessas três turmas com as funcionalidades do Módulo 2 do *software*.

As Turmas 2001 e 2002 foram divididas em sete equipes cada (em geral, não foram mantidas as equipes que realizaram as atividades do Grupo 1). Na Turma INFO-2 tivemos três duplas, um trabalho individual e uma dupla de alunos não entregou o trabalho depois de feito.

[08] Comente a atividade. O que você achou? Você aprendeu algo com ela? O quê? Se precisar de mais espaço para escrever a sua resposta, use o verso desta folha.

eu achei uma atividade legal e diferente, através desses sistemas de codificação pode se criar e desenvolver códigos.

[08] Comente a atividade. O que você achou? Você aprendeu algo com ela? O quê? Se precisar de mais espaço para escrever a sua resposta, use o verso desta folha.

achei interessante pelo conhecimento que adquiri mesmo que eu não utilize em nada, o fato de saber se é gratificante, se conhecia de cripto grafia, se não via conexão a arte de fazer.

[08] Comente a atividade. O que você achou? Você aprendeu algo com ela? O quê? Se precisar de mais espaço para escrever a sua resposta, use o verso desta folha.

achamos este trabalho muito interessante pois (nossa) anteriormente não conhecíamos nem mesmo este programa. Estamos até pensando em utilizá-lo no nosso dia a dia. Aprendemos que se um texto precisarmos enviar uma mensagem subliminar a gente consegue.

[08] Comente a atividade. O que você achou? Você aprendeu algo com ela? O quê? Se precisar de mais espaço para escrever a sua resposta, use o verso desta folha.

show, sim, aprendo que tem como descobrir a letra que aparece com mais frequência no texto e que a frase muda conforme eu aumento o número de vogais.

Figura 3.7: Algumas respostas dadas para a Atividade [08] da Turma 2001.

[08] Comente a atividade. O que você achou? Você aprendeu algo com ela? O quê? Se precisar de mais espaço para escrever a sua resposta, use o verso desta folha.

Essa atividade é bem legal, eu nunca tinha ouvido falar sobre a criptografia, mais agora eu aprendi e gostei, é bem interessante.

[08] Comente a atividade. O que você achou? Você aprendeu algo com ela? O quê? Se precisar de mais espaço para escrever a sua resposta, use o verso desta folha.

Achamos interessante. Aprendimos muito com ela. A criatividade de usar humano em decodificar uma mensagem para que ela fique legível apenas para quem a interessa realmente, para usar códigos, estratégias de guerra. Muito interessante e criativo.

[08] Comente a atividade. O que você achou? Você aprendeu algo com ela? O quê? Se precisar de mais espaço para escrever a sua resposta, use o verso desta folha.

Acho muito interessante, e produtivo, pois foi uma atividade diferente. Aprendemos a prestar mais atenção em cada letra de música, a conhecer mais certos letras.

Figura 3.8: Algumas respostas dadas para a Atividade [08] da Turma 2002.

[08] Comente a atividade. O que você achou? Você aprendeu algo com ela? O quê? Se precisar de mais espaço para escrever a sua resposta, use o verso desta folha.

É bem legal fazer essa atividade, vi que a matemática está dentro do português.

[08] Comente a atividade. O que você achou? Você aprendeu algo com ela? O quê? Se precisar de mais espaço para escrever a sua resposta, use o verso desta folha.

*Foi de ter importância, a melhora na
trouxe a possibilidade de nos inspirando em com
novos conhecimentos na matemática.*

Figura 3.9: Algumas respostas dadas para a Atividade [08] da Turma INFO-2.

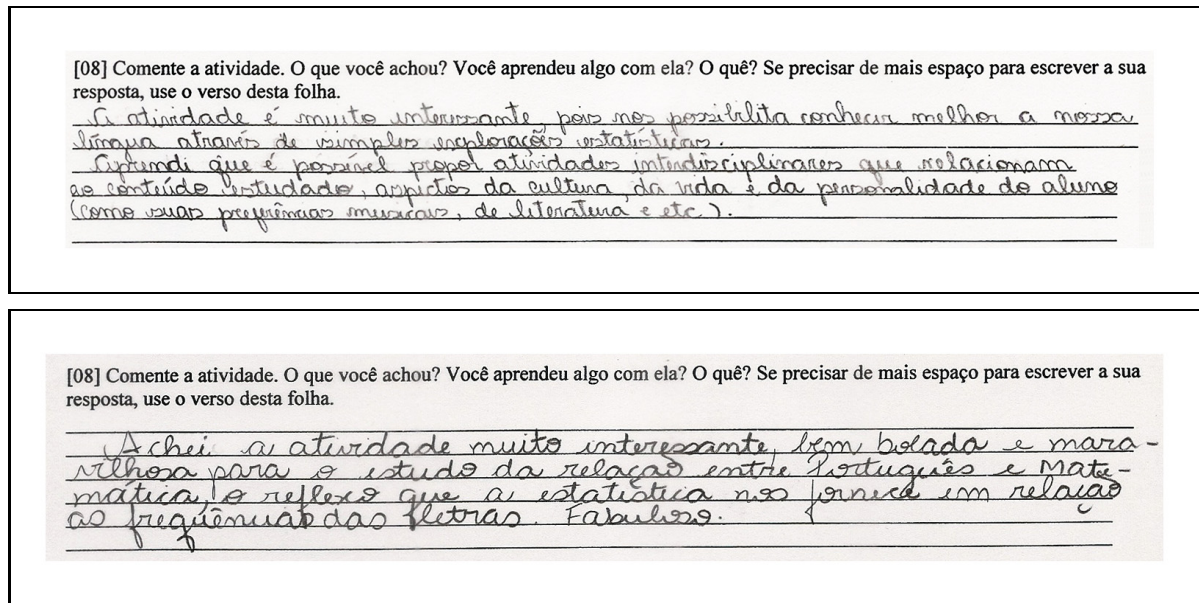


Figura 3.10: Algumas respostas dadas para a Atividade [08] da Turma da Especialização.

A Turma da Especialização seguiu o formulário original proposto por Bortolossi (2013) (<<http://www.uff.br/cdme/desktop/lpp/lpp-br.html>>) no lugar daquele usado pelas Turmas 2001, 2002 e INFO-2 (ver Apêndice C). Deste modo, no que se segue, para a Turma da Especialização, comentaremos apenas algumas das atividades comuns aos dois formulários.

Atividade [01] (análise estatística de textos pré-selecionados)

A primeira Atividade desse Grupo foi bastante semelhante à primeira atividade do Grupo 1. Foi deixado em cada computador do laboratório um arquivo contendo os textos que seriam analisados nessa tarefa. Em relação aos seis textos trabalhados, tivemos os seguintes resultados:

- **Texto 1:** na Turma 2001, tivemos seis tabelas perfeitamente preenchidas e uma tabela com os dados estatísticos errados. Não conseguimos identificar o motivo dos erros, pois os valores preenchidos não se encaixavam em nenhum outro texto trabalhado nessa atividade. Na Turma 2002, os sete grupos preencheram corretamente a tabela, porém, tivemos três grupos que erraram o número de letras do texto (eles puseram, no lugar, a média de letras por palavra do texto). Na Turma INFO-2, os quatro trabalhos entregues estavam totalmente corretos.
- **Texto 2:** na Turma 2001, tivemos cinco grupos com o preenchimento correto e dois grupos com erro. Novamente, não conseguimos encontrar um motivo para tal erro, pois os valores preenchidos foram diferentes dos possíveis valores para todos os outros textos dessa atividade (como os dois grupos cometeram o mesmo erro e acertaram todos os outros textos da atividade, muito possivelmente eles fizeram o trabalho em parceria).

Na Turma 2002, os sete grupos realizaram a tarefa corretamente e o mesmo ocorreu na Turma INFO-2.

- **Texto 3:** todos os sete grupos das turmas 2001 e 2002 fizeram o preenchimento correto das tabelas e, na Turma INFO-2, tivemos um grupo que preencheu a tabela com dados errados sobre o texto (a dupla selecionou a última frase do texto anterior da folha com os textos da atividade).
- **Texto 4:** todos os grupos das três turmas preencheram as tabelas de maneira correta. Na Turma 2002, tivemos três grupos que inverteram as respostas do Texto 3 com as do Texto 4.
- **Texto 5:** todos os grupos das três turmas preencheram as tabelas de maneira correta.
- **Texto 6:** nesse texto ocorreram bastante erros na sua coleta de dados. Muitos dos alunos não selecionaram o texto inteiro (o texto estava em três páginas). Tivemos, na Turma 2001, quatro grupos que coletaram de maneira correta os dados do texto e três grupos que erraram durante esse processo (eles colheram os dados somente da parte do texto que estava na primeira folha). Na Turma 2002, foram apenas três grupos que conseguiram realizar a tarefa de maneira correta, enquanto quatro grupos não conseguiram (um grupo selecionou as palavras “TEXTO 6” e “Fotojornalismo” e os outros três grupos selecionaram o Texto 5 junto com o Texto 6). Pela Turma INFO-2, foram dois trabalhos corretos e dois com preenchimento errado das tabelas (copiaram somente a parte do texto que estava na primeira folha).

A tabela a seguir resume o desempenho das três turmas no processamento estatístico dos seis textos desta atividade.

	Texto 1			Texto 2			Texto 3			Texto 4			Texto 5			Texto 6		
	C	P	E	C	P	E	C	P	E	C	P	E	C	P	E	C	P	E
Turma 2001	6	1	0	4	3	0	7	0	0	7	0	0	7	0	0	4	1	2
Turma 2002	3	4	0	6	1	0	6	1	0	6	1	0	7	0	0	3	4	0
INFO-2	3	1	0	4	0	0	3	1	0	4	0	0	3	1	0	2	0	2

Legenda: C = correta, P = parcialmente correta, E = errada ou em branco

Atividade [02] (classificação dos textos quanto ao grau de escolaridade)

A Atividade 2 gerou uma certa polêmica entre os integrantes dos grupos das Turmas 2001 e 2002 no momento de escolher um critério que adequasse cada texto a uma determinada faixa

etária escolar (a saber, os cinco primeiros anos do Ensino Básico e um texto do Ensino Médio). Como, na Turma INFO-2, o trabalho foi feito em no máximo duplas de alunos, um critério de comum acordo foi mais facilmente alcançado. Nesta turma, apenas uma dupla não justificou seu critério de faixa etária. Mais do que acertar a faixa etária de cada texto, nosso objetivo principal era o de estimular um debate entre os alunos para que estes estabelecessem e justificassem critérios de classificação com o uso da Estatística. Não foi feita nenhuma interferência e sugestão na escolha dos critérios escolhidos pelos alunos. Cabe ressaltar também que o preenchimento incorreto das tabelas da Atividade 1 pode ter tido influência nas respostas finais dadas pelos grupos e duplas. Seguem alguns critérios indicados pelas turmas:

- Turma 2001: “O critério foi a dificuldade de interpretação, tamanho do texto, porém havia alguns pequenos, porém difíceis.”; “Para o 1º ano tem que ser O Gato e Peixe porque tem poucas palavras e poucos acentos.”.
- Turma 2002: “Pela quantidade de letras, e pela facilidade de leitura. Porque é um texto mais maduro que exige entendimento.”.
- Turma INFO-2: “O nosso critério foi palavras e textos pequenos e de fácil entendimento para cada ano. Porque o último texto possui palavras, pontuação e a construção textual mais complexa e completa.”

Atividade [03] (palavras com 70% e 90% de consoantes)

Essa atividade foi considerada como difícil para grande parte dos alunos que participaram dela. De início, foram feitas afirmações sobre a impossibilidade de se achar uma solução para a atividade. Por esse motivo, todos os alunos que acertaram buscaram como recurso a utilização de nomes próprios como solução (uma vez que a manipulação das letras é mais fácil). Tivemos, na Turma 2001, três acertos (com os nomes próprios “Hallyne”, “Kevilyn” e “Camylla”), dois grupos disseram que não existia solução e dois grupos não responderam. Na Turma 2002, foram quatro acertos (o nome “Ygor” apareceu três vezes e o nome “Laryssa” uma vez), dois grupos afirmaram não conhecer uma resposta e um grupo disse não existir solução para a atividade. Na Turma INFO-2, foram dois acertos (“Brás” e “Kíssila”) e duas respostas erradas (“cansaço” e “intrísecos”). Na Turma da Especialização, as palavras que apareceram com o número de consoantes igual ou superior a 70% foram: “trás”, “traz”, “problemas”, “trechos”, “driblar”, “trem”, “cruz” e “programas”. Um aluno não fez a atividade.

Observação: não conhecemos, em Português, uma palavra (que não seja sigla, número romano, onomatopeia ou incorporada de outros idiomas) cujo número de consoantes seja maior do que ou igual a 90% do número total de letras da palavra. Neste contexto, as palavras com a maior proporção de consoantes que encontramos foram “tchãs” e “trens” (80% de consoantes).

Atividade [04] (stop words)

Para essa atividade foi necessário dar uma explicação sobre o significado do conceito de *stop words*, pois os alunos das três turmas não estavam conseguindo compreender a definição somente pelo enunciado da atividade. Os resultados foram os seguintes: na Turma 2001 tivemos cinco grupos que conseguiram sugerir algumas *stop words* frequentes da nossa língua (“que” e “de” apareceram em todos os trabalhos) e dois grupos deixaram a atividade em branco. Na Turma 2002, seis grupos fizeram suas sugestões (as palavras “que”, “de” e “a” que apareceram em todos os trabalhos) e um grupo deixou a atividade em branco. Na Turma INFO-2, todos conseguiram sugerir *stop words* (a palavra “que”, que apareceu em todos os trabalhos, e as palavras “e” e “da” presentes em dois trabalhos). A Turma da Especialização citou como *stop words*: “a”, “que” e “de”. Dois alunos não fizeram a atividade.

Atividade [05] (período com todas as palavras com 4 letras)

Essa atividade foi bastante simples e não houve nenhum tipo de questionamento ou dúvidas. Porém, ocorreram erros, como a omissão do valor do desvio padrão e valores errados para a média ou o desvio padrão. Os resultados foram: seis acertos na Turma 2001 e um erro no valor do desvio padrão (0,3499271 no lugar de 0). Na Turma 2002, cinco grupos acertaram a atividade e dois não colocaram o valor do desvio padrão (um deles alegou que não precisaria por, pois o valor era igual a média). Na Turma INFO-2, tivemos apenas um erro na média (7 no lugar de 4), todos os demais grupos acertaram a atividade. Todos os alunos da Turma da Especialização acertaram esta atividade, com exceção de um que não a fez.

Atividade [06] (período cuja média do número de letras por palavra é igual a 4)

Os resultados desta atividade foram os seguintes: na Turma 2001, foram cinco acertos (apenas um grupo fez a comparação) e dois erros; na Turma 2002 tivemos cinco acertos (apenas um deles com uma justificativa, porém sem nenhum fundamento) e duas respostas em branco; na Turma INFO-2, foram três acertos e uma dupla errou o valor do desvio padrão. Com exceção de um que não a fez, todos os alunos da Turma da Especialização acertaram esta atividade (mas ninguém fez a comparação solicitada).

Seguem algumas respostas: “4,0 e 3,082207. A única diferença é que os desvios são diferentes.” da Turma 2001 (o curioso é que esse grupo errou o desvio padrão da Atividade [05], colocando 0,3499271 no lugar de 0); “A média de letras 4 é igual e o desvio padrão é diferente.” da Turma 2001; “4,0 e 3,082207. O número é menor na média e o número 6 é maior.” da Turma 2002 (essa foi a resposta, com a justificativa sem fundamento que mencionamos anteriormente); “4 e 0”. Não tem desvio padrão em nenhum dos dois.” da Turma INFO-2 (o aluno que deu esta

resposta errou a média e acertou o desvio padrão da Atividade [05]).

Atividade [07] (lipogramas)

Essa foi uma atividade difícil, porém, como havia uma premiação para as melhores respostas, os grupos se dedicaram mais para a sua realização. O conceito de *lipograma* foi facilmente entendido e, por esse motivo, os alunos logo perceberam de que não se tratava de uma atividade trivial. Infelizmente, não saíram textos originais muito extensos. Os textos mais longos foram copiados da Internet. Não houve respostas erradas (isto é, textos com a letra “a”) ou em branco nessa atividade.

Na Turma 2001, a equipe que ganhou o prêmio escreveu o seguinte lipograma (adaptado de um lipograma da *Internet*): “Um minuto do viver que Deus nos premiou é bem indescritível, formoso, feliz, gostoso e único. Ouvindo o vento, que corre sempre e, nele incluído o sol quente, bonito, com luz forte, que nos permite sentir, depois, o escuro.”. O vencedor da Turma 2002 foi o texto “Pense em um tempo onde você pode ser você mesmo sem terem preconceito com seu modo de escolher viver. Onde você goste de sorrir e ser feliz em todo momento.”. Na Turma INFO-2, o lipograma que ganhou o prêmio foi “Eu fui ao boliche, joguei muito e perdi tudo! Fui no jogo do bicho joguei perdi tudo, mesmo perdendo continuei o jogo.”.

Um aluno da Turma da Especialização conseguiu escrever um lipograma com 255 palavras (os demais alunos não fizeram esta atividade):

Longe de ti

Oh! O tempo, nosso inimigo. Lhe vi recentemente no último domingo, porém é como se um século inteiro escorresse pelos minutos sem fim. E eu que sempre fui um ser meio bobo e um pouco louco, me vejo neste delírio sem conseguir decidir o melhor rumo e por onde devo seguir. Penso que só quero ser feliz com você e por isso, te peço: me impressione, me emocione, me impulsione e me leve neste sonho louco de vivermos juntos este momento único. E eu deliro e vibro. Sim! É você! É você que eu quero. E quero muito. Os meus sentimentos me consomem me consomem como se meu corpo fosse somente um estorjo vivo do sentimento que tenho por ti. E cuido de ti com todo zelo e respeito possíveis. E hoje vendo-o de novo, sorrindo feito um bebê que reencontrou seu brinquedo preferido. Te olho e o brilho dos seus olhos é como o sol que se reflete em mim. Te beijo e o gosto do seu beijo é doce como mel. E todo o meu corpo responde o pulso no mesmo ritmo que voa. . . De novo o tempo. Despedimo-nos, porém o gosto do seu beijo, o cheiro do seu perfume, o toque de seus dedos, insistem em me perseguir. Tenho certo que deixei muito de mim em você e muito de você ficou em mim. E mesmo sem perceber, você foi o melhor sonho que vivi. Porém, hoje tenho que entender que o melhor é viver longe de ti.

Atividade [08] (pangramas)

Para essa atividade, também colocada sob a forma de um desafio valendo um prêmio, foi

necessária uma explicação com exemplos sobre a definição de *pangrama*, uma vez que a maioria dos alunos das Turmas 2001, 2002 e INFO-2 não compreenderam o conceito apenas com o enunciado da atividade. Por valer um prêmio, existiu aqui também uma maior dedicação por parte dos alunos, mas, diferentemente da atividade sobre lipogramas, alguns erros ocorreram nesta atividade. Os resultados foram os seguintes: na Turma 2001, três grupos construíram corretamente um pangrama. O vencedor foi: “Hoje eu acordei muito feliz, quando de repente avistei de longe uma mulher que deixava seu batom cair no chão toda hora, e percebi que era minha amiga Kawany.” Dois grupos erraram (um escreveu mais de uma frase e, no outro, faltaram as letras “w” e “y”) e dois grupos deixaram a atividade em branco. Na Turma 2002, tivemos apenas dois grupos que conseguiram escrever um pangrama. O vencedor foi: “Quando anoiteceu Kátia pegou o violão e fez uma bela serenata chamada sexta feira para João e Yan e ela fez um show.”. Os outros cinco grupos erraram por omissão de letras (as mais comuns foram “k”, “y” e “w”). Na Turma INFO-2, tivemos duas duplas que escreveram de maneira correta um pangrama. A resposta vencedora foi: “Washington, Yasmin e Kelly foram assistir na floresta encantada, lindos gafanhotos azuis que celebram a pequena terra das jovens bruxas.”. O aluno que fez sozinho escreveu mais de uma frase e a outra dupla omitiu a letra “h”. Na Turma da Especialização, os cinco alunos que fizeram a atividade acertaram. Aqui está um pangrama elaborado por um dos alunos: “Dona Clotides faxinava a cozinha de Wesley quando encontrou uma caixa com kiwi, jabuticabas e pêssegos.”.

Atividade [09] (texto longo cuja média do número de letras por palavra é igual a 4)

As Turmas 2001 e 2002 tiveram bastante dificuldade para realizar essa atividade, o terceiro e último desafio valendo um prêmio. Muitos grupos não estavam conseguindo estabelecer uma estratégia para escrever uma frase com tal especificidade (isto é, com a média do número de letras por palavra igual a 4). Por este motivo, sugerimos aos alunos que revissem as Atividades [05] e [06] e que fossem criando frases em que cada duas palavras o total de letras fossem oito.

A Turma INFO-2 conseguiu realizar a tarefa com mais facilidade (a mesma dica também foi dada a eles) com os seguintes períodos: “O sapato é vermelho e o vestido também é vermelho.”, “O gato roeu a roupa do rei de Roma, a rainha Ruinha resolveu remendá-lo.”, “João Luiz come maçã.” e “Amor, faça tudo como você quer!”. A vencedora foi a segunda resposta.

Na Turma 2001, tivemos apenas dois períodos corretos: “Caio pediu para sair de casa amanhã de noite.”, “Jack caiu da escada ontem de noite.” (essa resposta também apareceu na Turma 2002). O período vencedor foi o primeiro. Tivemos ainda duas respostas erradas e três em branco.

Na Turma 2002, foram cinco respostas corretas: “Jogo Dota toda hora.”, “Papai é feioso.”, “Todo dia vamos pescar os peixes no lago.” (essa última resposta apareceu duas vezes nessa turma) e, novamente, “Jack caiu da escada ontem de noite.” (muito provavelmente a Turma 2002 deve ter criado este período primeiro, uma vez que as aulas da Turma 2002 eram às quintas e da Turma 2001 às sextas). Por falta de originalidade nessa turma, a vencedora foi a primeira resposta. Tivemos nessa turma duas respostas erradas.

Dois alunos da Turma da Especialização conseguiram escrever textos mais longos nos quais a média do número de letras por palavra é igual 4:

Jana

Jana é uma menina que acabou de entrar na adolescência, ela tem quatorze anos, seus olhos verdes e sua cor morena jambo são iguais aos da sua querida mãe, contudo seus gestos e seu modo de falar em público lembra muito seu pai. As propagandas da televisão já começam a mostrar sinais do Natal, isso denuncia o que estamos no final do ano, e nesse período sua família vai doar algumas roupas para um grande orfanato. Por isso, sua mãe pediu para ela arrumar o armário. A jovem menina olha para seu sapato que é azul, ela ama essa cor e todos os seus tons. Diz que lembra os seus pais, pois eles se casaram em uma bonita tarde a céu aberto, com nuvens brancas e o resto todo era azul. E a igreja lá atrás bem bonita toda branca, cheia de flores. Assim, ela vê nas fotos, já que na época do casório ela era muito miúda e como não se lembra desse dia tão feliz para os pais dela, Mauro e Mara.

A flor amarela

A bela flor amarela estava no jardim da casa azul da Mirela. Durante o dia inteiro a Mirela cuida da flor, mas à noite ela admira pela janela, a linda flor amarela que enfeita o jardim dela. Mas, neste dia, algo estranho aconteceu, a flor amarela não estava mais ali. O que houve? A flor morreu... É o ciclo da vida, meu amor. E logo no seu jardim nascerá outra linda flor. E a Mirela muito triste, continuou a cuidar do seu jardim para nele floresça a alegria de um novo dia.

Atividade [10] (palavras que usam as cinco vogais e palavras só com consoantes)

Todos os trabalhos entregues acertaram a primeira parte desta atividade, isto é, escrever uma palavra onde as cinco vogais aparecem. Entre as respostas estão “Aurélió”, “curiosidades”, “produtividade”, “consensuais”, “costureira”, “preguiçosa”, “neutralizou”, “inconstitucionalissimamente”, “bagunceiro”, “comunidade”, “quixotesca”, entre outras. Tivemos também palavras repetidas: “cabeçudinho”, “opulência” e “anoiteceu”. Com exceção de um aluno da Turma 2001 e de um aluno da Turma da Especialização que escreveram “DVD”, todos os demais alunos escreveram desconhecendo uma palavra em Português formada apenas por consoantes.

Atividade [11] (palavras que alternam consoantes e vogais)

Esta foi uma atividade cujo grau de dificuldade foi considerado moderado pelos alunos. Alguns grupos não quiseram ter o trabalho de procurar por palavras diferentes daquelas que inicialmente apareceram como solução por outros grupos e, como consequência disto, vários grupos utilizaram as mesmas palavras. Os resultados foram: na Turma 2001, cinco grupos acertaram a atividade (sendo três com a palavra “paralelepípedo”, os dois outros com as palavras “patati” e “retiro”, um grupo afirmou não existir resposta e um grupo deixou a questão em branco; na Turma 2002, todos os sete grupos acertaram a atividade, porém tivemos a palavra “cataratas” se repetindo quatro vezes e “paralelepípedo” se repetindo três vezes. Na Turma INFO-2, tivemos três acertos (“aparecer”, “natureza” e o aluno que fez sozinho deu como resposta “número”, “sapato”, “apenas”, “veterana” e “medusa”) e uma dupla deixou a atividade em branco. Com exceção de um aluno, que não fez esta atividade, todos os demais alunos da Turma da Especialização acertaram a questão (“arara”, “humano”, “atividade”, “paralelepípedo”, “parede”, “boneca”, “comodidade”, entre outras palavras).

Atividade [12] (períodos que alternam consoantes e vogais)

Esta é outra atividade relatada como complicada pelos alunos, uma vez que requeria uma maior atenção na estrutura das palavras que constavam no período dado. Foi dada a sugestão para que analisassem o período pelo *software*, mas não foi feita nenhuma menção para que eles buscassem alguma relação com a atividade anterior. Os alunos que analisaram o texto pelo *software* tiveram como resposta “número de vogais igual ao número de consoantes”, mas não perceberam a alternância entre vogais e consoantes. Os resultados foram os seguintes: na Turma 2001, tivemos seis respostas com “número de vogais igual ao número de consoantes” e um grupo deixou a atividade em branco; na Turma 2002, foram cinco respostas desse tipo e dois grupos disseram que a média do número de letras por palavra era igual a 4 (o que está errado); a Turma INFO-2 teve o pior desempenho nessa atividade: apenas uma resposta indicou que as vogais e as consoantes estavam intercaladas; duas respostas estavam erradas: “Ele é especial porque é um pangrama pois possui todas as letras do nosso alfabeto.” e “Que só tem um período.” (este último período, curiosamente, foi dado pelo aluno que respondeu na atividade anterior as palavras “apenas”, “veterana” e “medusa”); por fim, uma resposta estava em branco. Na Turma da Especialização, três alunos disseram que o número de vogais da frase é igual ao número de consoantes; seis alunos disseram que as vogais e consoantes aparecem de forma intercalada e um aluno não fez a atividade.

Atividade [13] (comentários dos alunos)

A Figura 3.11 mostra algumas respostas dadas para a Atividade [13] das Turmas 2001, 2002

e INFO-2, respectivamente.

The figure consists of three vertically stacked rectangular boxes, each containing a question and a handwritten response. The question in each box is: "[13]Faça um comentário sobre a atividade. Você aprendeu alguma coisa com esse módulo 2?Comente .".

- Top box (Turma 2001):** The response is written in blue ink and reads: "Sim, Acharmos interessante os desafios aprendemos o que e- Lipeograma, Pangrama e etc.. #showdebola".
- Middle box (Turma 2002):** The response is written in blue ink and reads: "Esse trabalho confundiu minha cabeça quase explodiu meu cérebro mais foi legal,".
- Bottom box (Turma INFO-2):** The response is written in blue ink and reads: "Aprendi curiosidades, como quantas palavras tem nos textos, que é muito difícil escrever um português textos sem a letra 'A'".

Figura 3.11: Algumas respostas dadas para a Atividade [13] das Turmas 2001, 2002 e INFO-2, respectivamente.

3.3 Grupo 3: Análise Estatística por Autor

Neste terceiro e último grupo de atividades, os alunos das Turmas 2001 e 2002 formaram grupos de tamanhos variados e o combinado foi que cada integrante analisasse duas obras de um autor comum escolhido pelo grupo. Na Turma INFO-2, as pesquisas foram feitas individualmente, cabendo a cada aluno colher informações sobre cinco obras de um mesmo autor. A Turma da Especialização não realizou as atividades deste último grupo.

A primeira parte da atividade foi a escolha do autor a ser pesquisado por cada grupo (além dos dados coletados sobre as suas obras, foi pedido também uma biografia do autor). Aqui cabe ressaltar que os maiores grupos escolheram aqueles autores com um maior acervo disponível na *Internet*. O site exclusivo de consulta utilizado por todos os grupos foi o <<http://www.dominiopublico.gov.br>>. A segunda parte foi o preenchimento do formulário do Apêndice D. Nessa etapa, com os alunos já familiarizados com o *software*, a única novidade foi a inclusão da chamada “taxa de vocabulário”: o número de palavras diferentes dividido pelo número total de palavras no texto (na realização da atividade, alguns alunos se equivocaram fazendo a divisão inversa). O uso da taxa de vocabulário foi justificado da seguinte maneira: “quanto mais

próximo de um ela for, mais diversificado (em termos de palavras diferentes) é o texto”. Porém, não foi relacionado no trabalho (apenas comentado com os alunos) que a taxa de vocabulário varia com o tamanho do texto (é de se esperar que, à medida em que os textos ficam mais longos, a taxa de vocabulário deve diminuir). A terceira e última etapa da atividade consistia em analisar os dados coletados sobre as diferentes obras literárias e tentar encontrar algo peculiar, repetitivo, semelhante ou até mesmo uma arritmia, algo que não se encaixasse em um padrão. Os autores analisados foram os seguintes:

- Aluísio de Azevedo (1857-1913): um grupo da Turma 2002 e um aluno da Turma INFO-2 analisaram onze obras desse autor, a saber, “O Japão”, “Filomena Borges”, “O Esqueleto”, “Mattos, Malta ou Matta?”, “Aos Vinte Anos”, “O Homem”, “Uma Lágrima de Uma Mulher”, “Demônios”, “A Condessa Vésper”, “Flores de Laranjeira” e “O Cortiço”.
- Casimiro de Abreu (1839-1860): um grupo da Turma 2001 analisou seis obras desse autor, a saber, “Saudade”, “Quando Tu Choras”, “Deus”, “O Que é Simpatia”, “Poesia de Amor” e “Desejo”.
- Castro Alves (1847-1871): um aluno da Turma INFO-2 analisou quatro obras desse autor, a saber, “Adormecida”, “Adeus”, “A Canoa Fantástica” e “A Cachoeira”.
- Eça de Queirós (1845-1900): um grupo da Turma 2002 e um aluno da Turma INFO-2 analisaram doze obras desse autor, a saber, “Alves e Cia”, “A Relíquia”, “Contos”, “A Ilustre Casa de Ramires”, “A Cidade e a Serra”, “Cartas D’Amor”, “O Mistério da Entrada de Sintra”, “O Mandarim”, “O Primo Basílio”, “A Capital”, “Carta de Inglaterra” e “Civilização”.
- Euclides da Cunha (1866-1909): um grupo da Turma 2001 e um aluno da Turma INFO-2 analisaram quatro obras desse autor, a saber, “À Margem da História”, “Terra Sem História (Amazônia)”, “Ondas e Outros Poemas Esparsos” e “Os Sertões”.
- Joaquim Manoel Macedo (1820-1882): um aluno da Turma INFO-2 analisou cinco obras desse autor, a saber, “Luxo e Vaidade”, “Luneta Mágica”, “Romance de Uma Velha”, “Amor e Pátria” e “A Moreninha”.
- José de Alencar (1829-1877): um grupo da Turma 2001 e outra da Turma 2002 analisaram doze obras desse autor, a saber, “A Pata da Gazela”, “Senhora”, “Cinco Minutos”, “O Que é Casamento?”, “Como e Porque sou Romancista”, “A Viuvinha”, “A Alma do Lázaro”, “Lucíola”, “Diva”, “Encarnação”, “Asas de Um Anjo” e “Mãe”.
- Lima Barreto (1881-1922): um grupo da Turma 2001 e um aluno da Turma 2002 analisaram duas obras desse autor, a saber, “A Mulher do Anacleto” e “A nova Califórnia”.
- Machado de Assis (1839-1908): um grupo da Turma 2001, outro da Turma 2002 e um aluno da Turma INFO-2 analisaram vinte e duas obras desse autor, a saber, “Dom Casmurro”, “A Igreja do Diabo”, “Adão e Eva”, “A Carteira”, “Memórias Póstumas de Brás Cubas”,

“A Semana”, “A Mão e a Luva”, “Jogo do Bicho”, “Mariana”, “Casa Velha”, “A Cartomante”, “Ayres e Vergueiro”, “A Desejada das Gentes”, “A Herança”, “A Ela”, “Iaiá Garcia”, “A Chave”, “Diana”, “Helena”, “As Bodas de Luís Duarte”, “Um Homem” e “Bons Dias”.

Nem todos os grupos e alunos conseguiram manter o acordo inicial proposto para a atividade, isto é, para as Turmas 2001 e 2002, cada pessoa deveria analisar duas obras e, para os alunos da Turma INFO-2, cada pessoa deveria analisar cinco obras. De fato, a conclusão desse grupo de atividades ficou bastante prejudicada por ter sido realizada no final do ano letivo e, infelizmente, também não tivemos tempo de fazer uma conclusão final dos dados coletados com os alunos. Sobre a terceira etapa da Atividade (detectar padrões nas obras analisadas), somente três alunos da Turma INFO-2 conseguiram escrever algo sobre o tema. A título de ilustração, colocamos na Figura 3.12 um exemplo de formulário do Grupo 3 preenchido e, na Figura 3.13,

um exemplo de análise feita.

Análise de Livros	
TEXTO	
Título do texto(fonte):	À MARGEM DA HISTÓRIA
Autor do texto:	EUCLIDES DA CUNHA
Idioma do texto:	Português
Número total de letras do texto:	147354
Cinco letras mais frequentes:	A - E - O - S - R
Cinco Stop Words mais frequentes:	FIM - LHAS - PELAS - ALGUNS - DESDE
Total de palavras(tp):	28080
Total de palavras diferentes(tpd):	8407
Índice de vocabulário(tpd/tp):	0,29
Cinco palavras mais frequentes:	TERRA - HISTORIA - VIDA - AMAZONAS - DI
Média de número de letras por palavras:	5,2694801
Desvio padrão:	3,3562877
Mediana:	5
Palavra mais longa:	21 LETRAS
Média de palavras por período:	26,0724234
Período mais longo:	198
Período mais curto:	1
Os três acentos mais utilizados:	´ ^ ~
Os três sinais mais utilizados:	, . -

Figura 3.12: Análise estatística da obra “À Margem da História” de Euclides da Cunha feita por um aluno da Turma INFO-02.

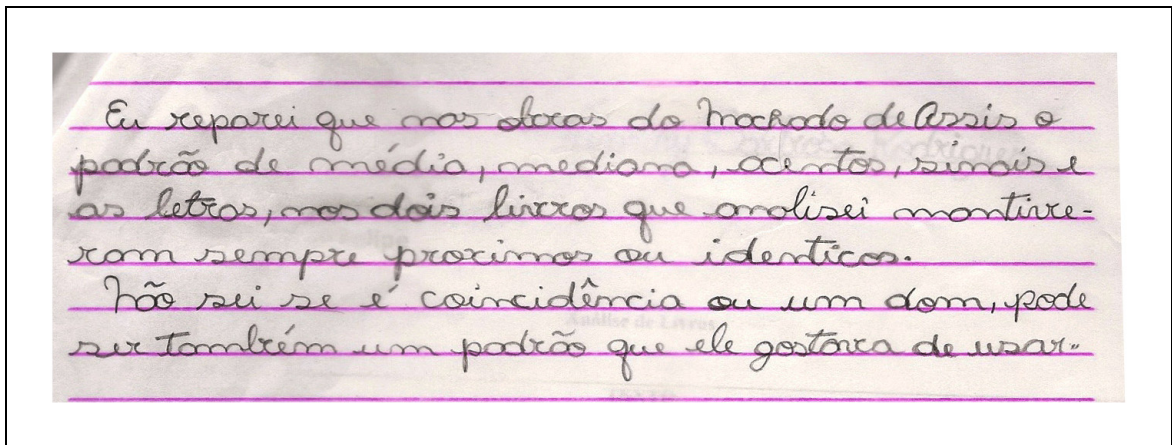


Figura 3.13: Conclusão dada por um aluno da Turma INFO-2 que analisou obras de Machado de Assis.

4 *Considerações Finais*

Nossa experiência neste trabalho mostrou que alguns ajustes finos podem ser feitos nas atividades propostas originalmente por Bortolossi (2009) e Bortolossi (2013):

- No enunciado da Atividade [01] do formulário do Grupo 1, indicar uma quantidade mínima e máxima de letras para cada um dos cinco textos em Português a serem analisados (para evitar que um aluno escolha, por exemplo, apenas textos pequenos, como aconteceu). Alguns alunos não conseguiram interpretar o que se queria com as cinco letras mais frequentes. Sugerimos que o quadro desta parte da atividade seja mais explícito, por exemplo:

	1ª letra mais frequente	2ª letra mais frequente	3ª letra mais frequente	4ª letra mais frequente	5ª letra mais frequente
Cinco letras mais frequentes:					

- No enunciado da Atividade [02] do formulário do Grupo 1, pedir para que o aluno indique explicitamente a permutação usada na decodificação do texto preenchendo uma tabela do tipo:

A	↔
B	↔
C	↔
D	↔
E	↔
F	↔
G	↔
H	↔
I	↔
J	↔
K	↔
L	↔
M	↔
N	↔
O	↔
P	↔
O	↔
R	↔
S	↔
T	↔
U	↔
V	↔
W	↔
X	↔
Y	↔
Z	↔

- Antes da Atividade [05] do Grupo 1, incluir atividades preparatórias para o princípio fundamental da contagem, considerando-se primeiro alfabetos com poucas letras. Por exemplo, suponha que o alfabeto tivesse apenas três letras, “A”, “B” e “C”, quantas chaves de codificação diferentes poderiam ser definidas? E se o alfabeto tivesse quatro letras? Cinco? Vinte e seis?
- Para enfatizar o que se quer na Atividade [06] do Grupo 1, pedir explicitamente no enunciado para que o aluno codifique um texto com a chave REVERSA e, então, decodifique-o com a mesma chave.
- No enunciado da Atividade [07] do formulário do Grupo 1 (sobre o que é mais fácil decodificar: textos grandes ou textos pequenos), pedir para que o aluno justifique sua resposta.
- Deixar as Atividades [01] e [02] do Grupo 2 para o final do formulário, pois acreditamos que depois de experimentar com os outros vários exercícios, talvez os alunos estejam mais confortáveis em usar o número de letras por palavras e o número de palavras por período como critérios adicionais para se identificar a faixa escolar de um texto. Ainda para estas atividades, tomar o cuidado de, na próxima vez, deixar textos longos em uma única página ou, então, alertar os alunos para o fato de que o texto ocupa mais de uma página.
- Incluir no formulário do Grupo 2 mais exercícios do tipo [05] e [06], pois eles são importantes para exercitar os conceitos de média e desvio padrão no contexto linguístico.
- Nos enunciados das Atividades [07] e [08] do Grupo 2, incluir explicitamente a dica para o aluno usar o programa para verificar se os textos criados são, de fato, lipogramas e pangramas.
- No enunciado da Atividade [12] do Grupo 2, pedir explicitamente para que o aluno faça uma comparação com a Atividade [11].

De um ponto de vista mais geral, temos as seguintes sugestões:

- Incluir na programação dos *softwares* um esquema que registre e envie para o professor as várias ações feitas pelos alunos na resolução dos exercícios propostos (o que eles tentaram e em qual sequência). Este registro certamente seria um instrumento valioso para uma análise mais aprofundada das estratégias e dos pensamentos dos alunos.
- No lugar de fazer com que os alunos preencham formulários impressos, que os dados sejam digitados diretamente em um Formulário Google (<<https://www.google.com/intl/pt-BR/forms/about/>>). Com isto, o processo de coleta e análise das respostas fica centralizado e mais fácil.
- Talvez, antes de usar o computador, selecionar alguns exercícios dos Grupos 1 e 2 para serem feitos com papel, lápis e calculadora. Isto pode ajudar na fixação dos conceitos e, também, pode levar o aluno a perceber melhor o papel da tecnologia e de como ela nos livra de tarefas e cálculos rotineiros.

Para registro, destacamos algumas das dificuldades encontradas durante a realização do trabalho:

- A falta de recursos do laboratório utilizado pelas Turmas 2001, 2002 e INFO-2 criou ao longo das aulas um certo “desgaste” geral (por parte dos alunos e do professor). Tivemos diversos problemas: poucos computadores, computadores inoperantes, “lentos” e sem *Internet*, sem a versão desejada do Java, além de uma certa dificuldade de se trabalhar em Linux (no caso das Turmas 2001 e 2002).
- Não obtivemos sucesso em agregar uma participação mais efetiva de professores de Português neste experimento interdisciplinar. Consultados, o que conseguimos foram, no início do trabalho, apenas algumas ideias e sugestões de temas (como, por exemplo, trabalhar com textos de Vinícius de Moares, por conta da comemoração de seu centenário de nascimento).
- O fato das Turmas 2001 e 2002 não reprovarem na disciplina “Resolução de Problemas Matemáticos” pode ter afetado o desempenho destas turmas nas tarefas propostas de um modo geral mas, principalmente, no final do ano. Neste contexto, em nossa opinião, as atividades que valeram prêmios foram estratégias eficazes no sentido de atrair e motivar os alunos destas turmas para a realização dos trabalhos.
- Outro item bastante importante é estabelecer um domínio de busca na *Internet*, para que os alunos somente recorram àqueles *sites* pré-estabelecidos. Isso, infelizmente, não aconteceu. Essa medida serve para ajudar na correção das atividades de textos analisados, fazendo do texto em questão único, sem diferentes versões, refrões a mais, etc.”.

Apesar destas dificuldades enfrentadas, acreditamos que a experiência, como um todo, foi positiva. Pelas falas dos alunos, por exemplo, podemos constatar que, de fato, muitos deles perceberam a conexão interdisciplinar entre Matemática e Português: “Sim, achamos interessante os desafios, aprendemos o que é lipograma, pangrama e etc.” (Turma 2001); “Aprendi curiosidades, como quantas palavras tem no texto, que é muito difícil escrever em Português textos sem a letra A” (Turma INFO-2); “A atividade é muito interessante, pois nos possibilita conhecer melhor a nossa língua através de simples explorações estatísticas. Aprendi que é possível propor atividades interdisciplinares que relacionam ao conteúdo estudado, aspectos da cultura, da vida e da personalidade do aluno (como suas preferências musicais, de literatura e etc).” (Turma da Especialização). Também é de nossa opinião que as atividades propostas de fato promovem as várias competências, conteúdos, habilidades e ações metodológicas recomendadas pelos PCN e pelo documento GAISE.

Pretendemos, neste ano de 2015, aplicar as atividades novamente, procurando, desta vez, enfatizar mais os conceitos estatísticos de média e desvio padrão no contexto linguístico e, com

sorte, incluir uma parceria efetiva com os professores de Português. Outro tópico que pretendemos incluir é Lei de Zipf (o que permitirá fazer conexões com funções potência, logaritmos e regressão linear).

APÊNDICE A – Questionário de Sondagem



Universidade Federal Fluminense

UNIVERSIDADE FEDERAL FLUMINENSE
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

- 1) Idade: _____
- 2) Sexo: M () F ()
- 3) Profissão: _____
- 4)
- a) Possui computador em casa? Sim () Não ()
- b) Possui acesso à internet em casa? Sim () Não ()
- c) Caso não tenha internet em casa, onde costuma acessar a internet?
- _____
- _____
- 5) Possui: Tablet (Ipad ou Android) ? Sim () Não ()
Smartphone (Iphone)? Sim () Não ()
- 6) Qual é o uso mais freqüente que faz do computador?
- _____
- _____
- 7) Que tipo de livro gosta de ler?
- _____
- 8) Quantos livros você costuma ler por ano?
- a) menos que 1. () b) entre 1 e 5. () c) mais que 5. ()
- 9) Qual tipo de música você prefere?
- _____
- 10) Você acha que existe alguma relação entre matemática e português? Qual?
- _____
- _____
- 11) Já participou de alguma pesquisa antes? Qual tipo?
- _____
- 12) Qual disciplina você mais gosta de estudar na escola?
- _____
- 13) Quando terminar o 3° ano do ensino médio você pretende:
- a) fazer faculdade sim () não ()
- b) fazer (ou já faz) curso técnico sim () não ()
- c) fazer outro curso profissionalizante qualquer sim () não ()
- d) apenas trabalhar sim () não ()

TEXTO 3	
Título do texto:	
Autor do texto:	
Idioma do texto:	Português
Número total de letras:	

	← Mais frequente					Menos frequente →				
Cinco letras mais frequentes:										

TEXTO 4	
Título do texto:	
Autor do texto:	
Idioma do texto:	Português
Número total de letras:	

	← Mais frequente					Menos frequente →				
Cinco letras mais frequentes:										

TEXTO 5	
Título do texto:	
Autor do texto:	
Idioma do texto:	Português
Número total de letras:	

	← Mais frequente					Menos frequente →				
Cinco letras mais frequentes:										

[02] Usando o Módulo 1 da atividade, tente decifrar a mensagem dada abaixo (copie e cole o texto para o aplicativo). Lembre-se que para ordenar uma coluna, basta clicar várias vezes no cabeçalho da coluna correspondente.

PIED NRH BPLTD JDLT ILQAD
 JDLT BEHLD AH FUDBD
 H HID JHQLQD
 NRH MHJ H NRH ODTTD
 QRJ APBH CDIDQBP

D BDJLQEP AP JDU

JPBD AP BPUOP APRUDAP

AP TPI AH IODQHJD

P THR CDIDQBDAP H JDLT NRH RJ OPHJD

H D BPLTD JDLT ILQAD NRH HR YD ML ODTTDDU

DE, OPU NRH HTSPR SDP TPZLQEP?

DE, OPU NRH SRAP H SDP SULTSH?

DE, D CHIHZD NRH HXLTSB

D CHIHZD NRH QDP H TP JLQED

NRH SDJCHJ ODTTDD TPZLQED

DE, TH HID TPRCHTTH

NRH NRDQAP HID ODTTDD

P JRQAP LQSHLULQEP TH HQBEH AH FUDBD

H GLBD JDLT ILQAP

OPU BDRTD AP DJPU

- [03] Usando o Módulo 1, criptografe uma mensagem e envie (por exemplo, por e-mail) para um colega seu decodificar.
- [04] O que fazem as permutações pré-definidas ROT 1, ROT 2, . . . , ROT 25 do Módulo 1 da atividade?
- [05] Quantas chaves de codificação/decodificação diferentes é possível definir com o programa do Módulo 1?
- [06] Se você aplicar a permutação pré-definida REVERSA a um texto já codificado com esta mesma permutação, então você obterá o texto original. Dê outro exemplo de permutação com esta propriedade. Alguma permutação pré-definida do tipo ROT possui esta propriedade?
- [07] Em sua opinião, o que é mais fácil decodificar, textos grandes ou textos pequenos?
- [08] Comente a atividade. O que você achou? Você aprendeu algo com ela? O quê? Se precisar de mais espaço para escrever a sua resposta, use o verso desta folha.

APÊNDICE C – Formulário do Grupo 2

Neste apêndice apresentamos as questões que orientaram o trabalho sobre a estatística das letras, palavras e períodos de textos em Português.

[01] Usando o Módulo 2 da atividade, analise seis textos (na folha anexa) escritos em Português referente aos anos do Ensino Fundamental I e um do ensino médio. Preencha as tabelas abaixo, indicando as cinco letras, os cinco acentos e os cinco sinais mais frequentes com os percentuais das frequências relativas correspondentes. Lembre-se que para ordenar uma coluna, basta clicar várias vezes no cabeçalho da coluna correspondente.

TEXTO 1	
Título do texto:	O Gato e A Raposa
Idioma do texto:	Português
Total de palavras do texto:	
Total de palavras diferentes:	
Número total de letras do texto:	
Número total de acentos do texto:	
Número total de sinais do texto:	

	← Mais frequente					Menos frequente →				
Cinco letras mais frequentes:										
Cinco acentos mais frequentes:										
Cinco sinais mais frequentes:										

TEXTO 2	
Título do texto:	Carta para O Papai
Idioma do texto:	Português
Total de palavras do texto:	
Total de palavras diferentes:	
Número total de letras do texto:	
Número total de acentos do texto:	
Número total de sinais do texto:	

	← Mais frequente					Menos frequente →				
Cinco letras mais frequentes:										
Cinco acentos mais frequentes:										
Cinco sinais mais frequentes:										

TEXTO 3	
Título do texto:	Matam ou Engordam?
Idioma do texto:	Português
Total de palavras do texto:	
Total de palavras diferentes:	
Número total de letras do texto:	
Número total de acentos do texto:	
Número total de sinais do texto:	

	← Mais frequente					Menos frequente →				
Cinco letras mais frequentes:										
Cinco acentos mais frequentes:										
Cinco sinais mais frequentes:										

TEXTO 4	
Título do texto:	O Leão e O Ratinho
Idioma do texto:	Português
Total de palavras do texto:	
Total de palavras diferentes:	
Número total de letras do texto:	
Número total de acentos do texto:	
Número total de sinais do texto:	

	← Mais frequente					Menos frequente →				
Cinco letras mais frequentes:										
Cinco acentos mais frequentes:										
Cinco sinais mais frequentes:										

TEXTO 5	
Título do texto:	Gato e Peixe
Idioma do texto:	Português
Total de palavras do texto:	
Total de palavras diferentes:	
Número total de letras do texto:	
Número total de acentos do texto:	
Número total de sinais do texto:	

	← Mais frequente					Menos frequente →				
Cinco letras mais frequentes:										
Cinco acentos mais frequentes:										
Cinco sinais mais frequentes:										

TEXTO 6	
Título do texto:	Fotojornalismo
Idioma do texto:	Português
Total de palavras do texto:	
Total de palavras diferentes:	
Número total de letras do texto:	
Número total de acentos do texto:	
Número total de sinais do texto:	

	← Mais frequente			Menos frequente →	
Cinco letras mais frequentes:					
Cinco acentos mais frequentes:					
Cinco sinais mais frequentes:					

- [02] Você conseguiria dizer qual dos textos analisados é referente ao 1^o ano do Ensino Fundamental I? E qual texto é o mais adequado ao 2^o ano do Ensino Fundamental II? Tente identificar os textos para o 3^o, 4^o e 5^o anos. Explique qual foi o seu critério de escolha. Obviamente que o último texto é típico do Ensino Médio. Porquê?
- [03] Você conhece alguma palavra em Português com 3 ou mais letras cujo número de consoantes seja igual ou superior a 70%? E superior a 90%? Não vale usar siglas, números romanos, onomatopeias ou palavras do inglês que foram incorporadas ao nosso idioma.
- [04] *Stop words* são palavras sem significado (como artigos, pronomes e proposições) que são filtradas antes ou depois de certos tipos de processamento de um texto em linguagem natural (como, por exemplo, o de identificar palavras-chaves de um texto). Usando o Módulo 2 com um texto longo (tente analisar um livro inteiro), sugira *stop words* (mais frequentes) para o Português.
- [05] Qual é média do número de letras por palavra no período abaixo? E o seu desvio padrão?
Luiz, você sabe como isso caiu aqui?
- [06] Qual é média do número de letras por palavra no período abaixo? E o seu desvio padrão?
O sapato é vermelho.
Compare com a média e o desvio padrão do exercício anterior.
- [07] **DESAFIO (vale um prêmio):** escreva um LIPOGRAMA em Português sem utilizar a letra “a” (isto é, escreva um texto em Português sem que letra “a” apareça). Ganhará

o prêmio o grupo que conseguir escrever um lipograma com o maior número de palavras. Importante: o texto tem que ter sentido! Não vale usar frases desconexas!

[08] **DESAFIO (vale um prêmio):** escreva um PANGRAMA no nosso idioma (isto é, escreva uma frase onde aparecem todas as letras do nosso alfabeto). Ganhará o prêmio a equipe que conseguir escrever o pangrama mais imaginativo!

[09] **DESAFIO (vale um prêmio):** escreva um texto cuja média do número de letras das palavras seja exatamente igual a 4. Ganhará o prêmio o grupo que conseguir escrever um texto com o maior número de palavras. Importante: o texto tem que ter sentido! Não vale usar frases desconexas!

[10] Escreva, se é que existe, uma palavra em Português que use todas as cinco vogais. E uma palavra em Português que use apenas consoantes?

[11] “Honorificabilitudinitatibus” é a palavra com o maior número de letras (27 no total) usada pelo poeta e dramaturgo inglês William Shakespeare (1564-1616). Ela aparece na Cena I do Ato V de sua obra “Love’s Labour Lost” (Ato V, Cena I) e seu significado é o seguinte: “estado de ser capaz de atingir honras”. Note que a palavra alterna consoantes e vogais. Escreva, se é que existe, uma palavra em Português com 6 ou mais letras, onde as consoantes e vogais aparecem de forma alternada.

[12] O que o período abaixo tem de especial?

Dama do Lago, apenas você sabe como atar veterana medusa!

[13] Faça um comentário sobre a atividade. Você aprendeu alguma coisa com esse Módulo 2? Comente.

Aqui estão os seis textos indicados na Questão [01]:

O GATO E A RAPOSA

Fonte: Simulado da Escola Estadual Luiz Francisco Ribeiro, MG.

O gato e a raposa andavam sempre juntos pelo mundo. Eram muito amigos, apesar de a raposa estar sempre desvalorizando o colega.

— Amigo gato, por que não aprende mais truques para fugir dos cachorros que nos perseguem? Sempre ouvi dizer que você é tão inteligente. Será verdade?

— Sei subir rapidamente em árvores. É o que me basta. Os cachorros não vão me pegar.

— Você só sabe isso? Eu sei 99 truques diferentes! Conheço mil manhas, cada uma melhor que a outra. Finjo-me de morta, me escondo nas folhas secas, nas moitas, corro em zigue-zague, disfarço minhas pegadas...

Enquanto a raposa falava distraidamente das suas habilidades se aproximavam dali dois cachorros. O gato muito esperto subiu rapidamente na árvore. Quando a raposa percebeu a presença dos cachorros, teve que sair em disparada para fugir deles.

— Pobre comadre raposa. É sempre preferível saber bem uma só coisa, a saber, mal noventa e nove coisas diversas.

CARTA PARA O PAPAÍ

Pai,

Posso brincar na casa do Felipe depois que eu terminar a lição?

Diz que sim, sim, sim!!!

Beijos

Júlia.

MATAM OU ENGORDAM

Adaptado de Ciência Hoje das Crianças. Rio de Janeiro: SBPC, ano 6. n. 30, p. 20-23.

Tem uma coisa que os adultos dizem que eu tenho certeza de que aborrece as crianças: “Vá lavar as mãos antes de comer! Ela está cheia de micróbios. Não coma esse troço que caiu no chão! Lave logo o machucado, senão os micróbios tomam conta!” Daí a criança vai logo pensando: “Coisa chata essa de micróbio!” E eles vão ficando com essa fama de monstros, sempre prontos a atacar em caso de desleixo.

Mas sem micróbios e bactérias também não dá para viver, porque há um montão deles que são essenciais para manter vida em nosso planeta. Quando a gente vai lavar as mãos antes de comer fica até meio desapontado, pois não vê micróbio nenhum. E acha aquilo um exagero. É que os micróbios são microscópicos.

Os micróbios – não há como negar – são responsáveis por uma série de aborrecimentos: gripe, sarampo, tifo, malária, febre amarela, paralisia infantil e um bocado de coisas mais. Mas também há inúmeros micróbios benéficos, que decompõem o corpo morto das plantas e animais, transformando suas moléculas complexas em moléculas pequenas, aproveitáveis na nutrição das plantas.

O vilão de nossa história, portanto, não é totalmente malvado. Se ele desaparecesse, nós também acabaríamos junto com ele.

O LEÃO E O RATINHO

Fonte: Fábulas de Esopo – Companhia das Letrinhas

Um leão, cansado de tanto caçar, dormia espichado à sombra de uma boa árvore. Vieram uns ratinhos passear em cima dele e ele acordou. Todos conseguiram fugir, menos um, que o leão prendeu embaixo da pata. Tanto o ratinho pediu e implorou que o leão desistiu de esmagá-lo e deixou que fosse embora.

Algum tempo depois, o leão ficou preso na rede de uns caçadores. Não conseguia se soltar, e fazia a floresta inteira tremer com seus urros de raiva. Nisso, apareceu o ratinho. Com seus dentes afiados, roeu as cordas e soltou o leão.

GATO E PEIXE

Rosa Clement

MEU GATO SE CHAMA PEIXE, MEU PEIXE SE CHAMA GATO. GATO NÃO GOSTA DE PEIXE, MAS PEIXE GOSTA DE GATO. PEIXE DORME NUM SAPATO, NÃO CABE MESMO NUM PRATO.

FOTOJORNALISMO

Olavo Bilac

Vem perto o dia em que soará para os escritores a hora do irreparável desastre e da derradeira desgraça. Nós, os rabiscadores de artigos e notícias, já sentimos que nos falta o solo debaixo dos pés... Um exército rival vem solapando os alicerces em que até agora assentava a nossa supremacia: é o exército dos desenhistas, dos caricaturistas e dos ilustradores. O lápis destronará a pena: *ceci tuera cela*.

O público tem pressa. A vida de hoje, vertiginosa e febril, não admite leituras demoradas, nem reflexões profundas. A onda humana galopa, numa espumarada bravia, sem descanso. Quem não se apressar com ela será arrebatado, esmagado, exterminado. O século não tem tempo a perder. A eletricidade já suprimiu as distâncias: daqui a pouco quando um europeu espirrar, ouvirá incontinenti o “Deus te ajude” de um americano. E ainda a ciência humana há de achar o meio de simplificar e apressar a vida por forma tal que os homens já nascerão com dezoito anos, aptos e armados para todas as batalhas da existência.

Já ninguém mais lê artigos. Todos os jornais abrem espaço às ilustrações copiosas, que pelos olhos da gente com uma insistência assombrosa. As legendas são curtas e incisivas: toda a explicação vem da gravura, que conta conflitos e mortes, casos alegres e casos tristes.

É provável que o jornal-modelo do século 20 seja um imenso animatógrafo, por cuja tela vasta passem reproduzidos, instantaneamente, todos os incidentes da vida cotidiana. Direis que as ilustrações, sem palavras que as expliquem, não poderão doutrinar as massas nem fazer uma propaganda eficaz desta ou daquela idéia política. Puro engano. Haverá ilustradores para a sátira, ilustradores para a piedade.

(...) Demais, nada impede que seja anexado ao animatógrafo um gramofone de voz tonitruosa, encarregado de berrar ao céu e à terra o comentário, grave ou picante, das fotografias.

E convenhamos que, no dia em que nós, cronistas e noticiaristas, houvermos desaparecido da cena – nem por isso se subverterá a ordem social. As palavras são traidoras, e a fotografia é fiel. A pena nem sempre é ajudada pela inteligência; ao passo que a máquina fotográfica funciona sempre sob a égide da soberana Verdade, a coberto das inumeráveis ciladas da Mentira, do Equívoco, e da Miopia intelectual. Vereis que não hão de ser tão freqüentes as controvérsias. . .

(...)

Não insistamos sobre os benefícios da grande revolução que a fotogravura vem fazer no jornalismo. Frisemos apenas este ponto: o jornal-animatógrafo terá a utilidade de evitar que nossas opiniões fiquem, como atualmente ficam, fixadas e conservadas eternamente, para gáudio dos inimigos. . . Qual de vós, irmãos, não escreve todos os dias quatro ou cinco tolices que desejariam ver apagadas ou extintas? Mas, ai! de todos nós! Não há morte para as nossas tolices! Nas bibliotecas e nos escritórios dos jornais, elas ficam catalogadas (...)

No jornalismo do Rio de Janeiro, já se iniciou a revolução, que vai ser a nossa morte e a opulência dos que sabem desenhar. Preparemo-nos para morrer, irmãos, sem lamentações ridículas, aceitando resignadamente a fatalidade das coisas, e consolando-nos uns aos outros com a cortesia de que, ao menos, não mais seremos obrigados a escrever barbaridades. . .

Saudemos a nova era da imprensa! A revolução tira-nos o pão da boca, mas deixa-nos aliviada a consciência.

APÊNDICE D – Formulário do Grupo 3

TEXTO	
Título do texto (fonte):	
Autor do texto:	
Idioma do texto:	Português
Autor do texto:	
Número total de letras do texto:	
Cinco letras mais frequentes:	
Cinco <i>stop words</i> mais frequentes:	
Total de palavras (tp):	
Total de palavras diferentes (tpd):	
Taxa de vocabulário (tpd/tp):	
Cinco palavras mais frequentes:	
Média do número de letras por palavra:	
Desvio padrão:	
Média do número de letras por palavra:	
Desvio padrão:	
Mediana:	
Palavra mais longa:	
Média de palavras por período:	
Período mais longo:	
Período mais curto:	
Os três acentos mais utilizados:	
Os três sinais mais utilizados:	

Referências Bibliográficas

- Bezerra, D. J.; Malagutti, P. L.; Rodrigues, V. C. S. *Aprendendo Criptologia de Forma Diver-tida*. Departamento de Matemática, Universidade Federal de São Carlos, 2010.
- Bortolossi, H. J. *Estatística das Letras, Palavras e Períodos*. Software educacional, Instituto de Matemática e Estatística, Universidade Federal Fluminense, 2009. Disponível em: <<http://www.uff.br/cdme/lpp/>>. Acessado em: 20 de junho de 2013.
- Bortolossi, H. J. *Estatística das Letras, Palavras e Períodos*. Revista do Professor de Matemática, Sociedade Brasileira de Matemática, p. 26-32, 2013.
- Brasil. MEC. *Parâmetros Curriculares Nacionais: Ensino Médio*. Secretaria de Educação Média e Tecnológica, Ministério da Educação, Brasília, 1999.
- Brasil. MEC. *PCN+ Ensino Médio: Orientações Educacionais Complementares Aos Parâmetros Curriculares Nacionais. Ciências da Natureza, Matemática e Suas Tecnologias*. Secretaria de Educação Média e Tecnológica, Ministério da Educação, Brasília, 2002.
- Brasil. MEC. *Orientações Curriculares para O Ensino Médio: Ciências da Natureza, Matemática e suas Tecnologias*. Secretaria de Educação Básica, Ministério da Educação, Brasília, 2006.
- Franklin, C. et al. *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework*. American Statistical Association, 2007. Disponível em: <<http://www.amstat.org/education/gaise/>>. Acessado em: 5 de março de 2014.
- Garfield, J. et al. *Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report*. American Statistical Association, 2005. Disponível em: <<http://www.amstat.org/education/gaise/>>. Acessado em: 5 de março de 2014.
- MacKay, R. J.; Oldford, W. *Stat 231 Course Notes Fall 1994*. University of Waterloo, 1994.
- Pemmaraju, S.; Skiena, S. *Computational Discrete Mathematics: Combinatorics and Graph Theory with Mathematica[®]*. Cambridge Univeristy Press, 2003.
- Pfannkuch, M.; Wild, C. J. *Towards An Understanding of Statistical Thinking*. Em: Ben-Zvi, D.; Garfield, J. (Eds.) *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking*. Kluwer Academic Publishers, 2004.
- Quaresma, P. *Frequency Analysis of The Portuguese Language*. Technical Report, Department of Mathematics, University of Coimbra, Portugal, 2008. Disponível em: <<http://www.mat.uc.pt/~pedro/cientificos/Cripto/CISUC-TR200803.pdf>>. Acessado em: 21 de junho de 2014.
- Secretaria de Estado de Educação do Rio de Janeiro (SEEDUC-RJ), *Currículo Mínimo – Conteúdo Pedagógico*, 2012. Disponível em: <http://www.conexaoprofessor.rj.gov.br/curriculo_aberto.asp>. Acessado em: 13 de outubro de 2010.

Shaughnessy, J. M. *Research On Statistics Learning and Reasoning*. Em: Lester, F. K. Jr. (Editor) *Second Handbook of Research on Mathematics Teaching and Learning (Volume 2)*. Information Age Publishing, 2007.

Wild, C.; Pfannkuch, M. *Statistical Thinking in Empirical Enquiry*. *International Statistical Review*, v. 67, n. 3, p. 223-265, 1999.