UNIVERSIDADE DO GRANDE RIO "Prof. José de Souza Herdy"

UNIGRANRIO

PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA (PROPEP)

MESTRADO PROFISSIONAL EM ENSINO DAS CIÊNCIAS NA EDUCAÇÃO BÁSICA

JOSÉ CARLOS COELHO SARAIVA

SISTAT: FERRAMENTA COMPUTACIONAL COMO PROPOSTA PARA O ENSINO DA ESTATÍSTICA

Duque de Caxias

2016

JOSÉ CARLOS COELHO SARAIVA

SISTAT: FERRAMENTA COMPUTACIONAL COMO PROPOSTA PARA O ENSINO DA ESTATÍSTICA

Dissertação apresentada à Universidade do Grande Rio "Prof. José de Souza Herdy", como parte dos requisitos para a obtenção de grau de Mestre em Ensino das Ciências na Educação Básica.

Orientadora: Prof. Dra. Eline das Flores Victer

Co Orientador: Prof. Dr. Angelo Santos Siqueira

S243s Saraiva, José Carlos Coelho.
SISTAT: ferramenta computacional como proposta para o ensino da estatística / José Carlos Coelho Saraiva. – 2016.
102 f.: il.; 30 cm.

Dissertação (mestrado em Ensino das ciências na Educação Básica) – Universidade do Grande Rio "Prof. José de Souza Herdy", Escola de Educação, Ciências, Letras, Artes e Humanidades, 2016.
"Orientadora: Eline das Flores Victer".
"Co-Orientador: Angelo Santos Siqueira".
Bibliografia: f. 31-36.

1. Educação. 2. Estatística. 3. Ferramentas – Ensino auxiliador por computador. 4. Aprendizagem - Ensino. 5. Inovações tecnológicas. I. Victer, Eline das Flores. II. Siqueira, Angelo Santosi. III. Universidade do Grande Rio "Prof. José de Souza Herdy. IV. Título.

CDD – 370

JOSÉ CARLOS COELHO SARAIVA

SISTAT: FERRAMENTA COMPUTACIONAL COMO PROPOSTA PARA O ENSINO DA ESTATÍSTICA

Dissertação apresentada como requisito parcial para obtenção de grau de Mestre em Ensino das Ciências na Educação Básica no curso de Pós-Graduação em Ensino das Ciências na Educação Básica da Universidade do Grande Rio "Prof. José de Souza Herdy".

Dissertação aprovada em 28 de MARÇO de 2016.

Prof. Drª. Eline das Flores Victer - Orientadora
UNIGRANRIO – Universidade do Grande Rio

Prof. Dr. Angelo Santos Siqueira - Co orientador
UNIGRANRIO– Universidade do Grande Rio

Prof. Dr. Abel Rodolfo Garcia Lozano
UNIGRANRIO – Universidade do Grande Rio

Prof. Dr. Valdir Agustinho de Melo
UEZO – Centro Universitário Estadual da Zona Oeste

Prof. Dr. Raimundo José Macário Costa
UFRRJ – Universidade Federal Rural do Rio de Janeiro

RESUMO

A estatística faz parte da linguagem matemática, que muitas vezes é vista como de difícil entendimento, possuindo, ainda, simbologia própria que necessita de apresentação prévia para sua adequada compreensão. Desta forma, a leitura e o entendimento das informações disponibilizadas nos mais variados meios de comunicação ficam comprometidos quando não há uma abordagem e/ou um trabalho adequado junto aos alunos da educação básica. Em nossa prática dentro e fora de sala de aula, temos percebido a escassez de recursos tecnológicos e/ou metodologias de ensino e aprendizagem alternativo-complementares direcionadas ao ensino da estatística que auxiliem o professor no desenvolvimento e ampliação de conteúdos específicos do tema. Por isso, consideramos que a presente pesquisa e a construção de nosso produto educacional se justificam como possibilidade de contribuição para a melhoria do processo de ensino e aprendizagem, ao proporcionar ao docente e seu aluno novas formas de lidar com o conhecimento estatístico envolvendo novas tecnologias, pois a microinformática se encaixa perfeitamente nesta lacuna, por meio da possibilidade de se desenvolver softwares com objetivos especificamente educacionais. Nosso produto educacional, denominado Sistat, cujo desenvolvimento é o objetivo principal de nossa pesquisa, é uma ferramenta computacional cujo escopo é favorecer o processo de ensino aprendizagem dos conteúdos da estatística descritiva, de forma que permita ao aluno ter um foco maior nos conceitos e técnicas estatísticos e na análise das informações, compreendendo como os cálculos foram realizados, sem que haja o gasto excessivo de tempo com complexos e repetitivos cálculos.

Palavras-chave: Estatística. Ferramenta Computacional. Ensino-aprendizagem.

ABSTRACT

The statistical is part of the mathematical language, which is often seen as difficult to understand, having also own symbology which requires prior presentation to be fully understood. This way, the reading and the information understanding provided by many ways of communication are impaired when there is no approach and or appropriate work with basic education students. In our practice inside or outside of the classroom, we have noticed lack of technological resources and teaching methodologies and alternative learning directed to teaching Statistics to assist the teacher in the development and expansion of specific theme content. Then we guess that our survey and development of our own educational material is justified as a possible contribution to the improvement of teaching and learning providing the teacher and students new ways to dealing with statistical knowledge involving technologies as microcomputers which fits very well in this gap, through possibility to software developing with specific educational goals.

Our Educational material called Systat which its development is the main goal of our survey it's a computational tool with scope to support teaching process of descriptive statistics contents which students will be able to have a better understanding of concepts, statistical technics and information analysis. Students will understand how calculation were carried out without excessive spending time with a complex and repetitive calculations.

Keywords: Statistics. Computational Tool.Teaching-learning.

AGRADECIMENTOS

À **Deus**, que durante todo esse e trabalho e por toda a minha vida tem me acompanhado e iluminado, sendo o maior dos amigos e o melhor dos terapeutas. Obrigado pelo alívio e socorro nos momentos difíceis (que não foram poucos) e principalmente por sempre ter segurado minha mão, mesmo (eu) não sendo digno.

Aos meus **pais**, Lidia e José (*in memorian*), e minha **esposa**, Valéria, meus exemplos maiores de amor, apoio, companheirismo e amizade.

Aos meus **filhos**, Arthur e Pedro, que são minha mola propulsora e me dão coragem de enfrentar todos os desafios, principalmente pelo amor que sentem por mim, o mais puro e sincero.

Aos meus **avós**, Anibal (*in memorian*), Florinda (*in memorian*), José e Maria (*in memorian*) pelos conselhos, pela doçura e pelas experiências de vida compartilhadas.

Aos meus **irmãos**, Alexandre e Aline, que apesar dos seus compromissos diários sempre se preocupam comigo.

Aos meus **cunhados** e **sobrinhos**, Adilson, André, Anne, Áurea, Camilla, Caroline, Clara, Davi, Janaína, Miguel, Nadine e Ruy (juntamente com meus irmãos), pelos nossos momentos de lazer, momentos estes que renovam minha energia.

A todos os meus **amigos**, que mesmo distantes, torcem por mim, em especial Ed Carlos, Bruno Resende, Everton Lucas e Marcelo Conceição.

Aos meus **amigos**, Erundino Gonzalez, Gilson Rodrigues, Luiz Lima, Luiz Vieira (*in memorian*), Ricardo Garcia e Rubens Costa, que me proporcionaram oportunidades profissionais e de aprendizado sem preço, além de serem grandes conselheiros e amigos.

Aos meus **orientadores**, Dr. Angelo Siqueira e Dra. Eline Flores, que acreditaram em mim, estando sempre disponíveis e dispostos a ajudar. Mostraram-me que há mais que pesquisadores e resultados atrás de uma dissertação, existem vidas humanas.

LISTA DE FIGURAS

Figura 1 – Distribuição Normal..... **27**
Figura 2 – Assimetria..... **27**
Figura 3 – Correlação Linear..... **29**
Figura 4 – Modelo de desenvolvimento em Cascata..... **36**
Figura 5 – Fluxograma..... **38**
Figura 6 – Tela Principal..... **40**
Figura 7 – Estatísticas da Pesquisa..... **41**
Figura 8 – Memória de Cálculo..... **41**
Figura 9 – Organização e tratamento da informação..... **42**
Figura 10 – Representação gráfica..... **42**
Figura 11 – Correlação Linear / Regressão linear..... **43**
Figura 12 – Salvamento e exportação de dados..... **43**

SUMÁRIO

1	INTRODUÇÃO…….	9
2	REVISÃO BIBLIOGRÁFICA E REFERENCIAL TEÓRICO …..……..	13
2.1	Informática Educativa no Brasil……	15
2.2	Ensino da Matemática/Estatística e Tecnologia……...	16
3	A ESTATÍSTICA NO SYSTAT…….	19
3.1	Definições Básicas da Estatística……	19
4	METODOLOGIA…….	32
4.1	Desenvolvimento da Ferramenta Computacional……	33
4.1.1	A Linguagem de Programação Embarcadero Delphi……..	33
4.1.2	Engenharia de Software……..	35
4.1.3	Fluxograma……....	37
5	PRODUTO EDUCACIONAL…….	39
5.1	Elementos do Produto Educacional……	40
5.2	Utilização em sala de aula (Teste inicial do produto educacional).	44
6	RESULTADOS DE APLICAÇÃO DO PRODUTO……	46
6.1	Metodologia e aplicação……	46
6.2	Conclusão da aplicação do produto……	60
7	CONSIDERAÇÕES FINAIS……..	61
REFERÊNCIAS……		62
APÊNDICES……..		64

1. INTRODUÇÃO

A Estatística está cada vez mais presente no dia-a-dia das pessoas, inclusive profissionalmente. Somos diariamente “bombardeados” por informações das mais diversas áreas, através de todos os meios de comunicação e isso acontece independente de tomarmos conhecimento da presença da Estatística. Desta forma, acabamos por fazer uso da Estatística involuntariamente, com maior ou menor intensidade.

Segundo Lopes (1998), apesar de sua constatada importância nas mais variadas áreas do conhecimento, permitindo uma análise, inclusive, de questões sociais e econômicas, o ensino da estatística não pode se limitar a aplicação de fórmulas e cálculos, resolvendo exercícios mecanicamente. É necessário o desenvolvimento do Pensamento Estatístico que consiste, principalmente, em saber utilizar conceitos estatísticos para resolver problemas.

De acordo com pesquisas realizadas, entre 2008 e 2010, pelo grupo denominado Processo de Ensino e Aprendizagem em Matemática – PEA-MAT da PUC-SP, através do projeto PEA-ESTAT (Processo de Ensino e Aprendizagem Envolvendo Raciocínio Estatístico e Probabilístico) constatou-se o desconforto dos professores ao abordarem os conteúdos estatísticos em sala de aula, entrando em sintonia com outras pesquisas realizadas pelo próprio grupo. Observou-se também, nestas pesquisas, que a maioria dos livros didáticos e/ou apostilas sugeridas/fornecidas pelas redes de ensino direcionam o foco para atividades com predominância de cálculos matemáticos e leitura de gráficos, limitando-se normalmente à leitura dos eixos.

Gal (2002) relata que no quadro teórico para a pesquisa em educação estatística encontramos vários trabalhos realizados tanto no Brasil quanto no exterior, todos com o objetivo de intensificar o desenvolvimento de habilidades para a análise exploratória de dados. Particularmente, pela importância de seus conceitos e reflexões desenvolvidos em suas pesquisas, destacamos o letramento estatístico.

Cazorla (2006), afirma que enfrentamos desafios na inserção dos conteúdos estatísticos na educação básica, pois os professores durante sua formação inicial abordaram esses conteúdos no contexto do Ensino Superior, sem observar seus

aspectos didáticos, principalmente para apresentação a crianças e adolescentes, além da pouca oferta de recursos didáticos e/ou computacionais.

Carzola (2006) observa ainda, como estratégia didática, a utilização de poucos dados com o objetivo de facilitar a aprendizagem dos cálculos (que, mesmo assim, consomem muito tempo de aula), reduz a possibilidade de os alunos perceberem o potencial do poder inferencial das medidas estatísticas. Desta forma, torna-se necessário aumentar a base de dados e reduzir o tempo gasto com cálculos complexos já abordados anteriormente, implicando necessariamente na utilização de recursos computacionais.

Além disso, o rápido e constante desenvolvimento na área da computação, aliado a disseminação do uso do computador nas instituições de ensino, tornando-o uma ferramenta complementar ao processo de ensino-aprendizagem, gerou um terreno fértil para o desenvolvimento de ferramentas computacionais focadas na área educacional. Ferramentas essas que podem auxiliar, construir gráficos, simular e agilizar cálculos, analisar informações que, por muitas vezes, envolve manipulação de massa de dados, o que acaba sendo um processo demorado e cansativo, quando realizado manualmente.

Godino (1995), considera que se deve apresentar formas de processar dados em softwares estatísticos no ensino da estatística, levando em consideração que o computador não deve ser visto somente como um recurso de cálculo, mas também como um recurso didático.

Ainda segundo Godino (1995), de uma forma geral, a utilização do computador como ferramenta no processo de ensino-aprendizagem se dá em duas áreas: como objeto de estudo no ensino da microinformática e como recurso didático de apoio no ensino de outras áreas. Seguindo por essa segunda área de atuação da microinformática, em concordância com as afirmações de nossos referenciais teóricos, que apontam uma escassez de recursos computacionais voltados para o ensino da estatística, sentimos motivação a desenvolver uma ferramenta computacional que tem por objetivo auxiliar o ensino-aprendizagem da Estatística Descritiva, agilizando e demonstrando cálculos, construindo distribuições de frequências, plotando gráficos e correlacionando variáveis.

É importante frisar que o objetivo de nossa ferramenta computacional é permitir que o aluno foque nos conceitos e técnicas estatísticos, como, por exemplo, na análise das informações, e, inclusive, compreender como os cálculos foram realizados.

Norteamo-nos pelos conceitos de Godino (1995), que menciona os objetivos didáticos no uso dos computadores na educação estatística, dos quais destacamos: proporcionar ao aluno a capacidade de: coletar, organizar, depurar, reunir, representar e analisar sistemas de dados; utilizar um software que analise um sistema de dados é um objetivo importante na atualidade, no entanto, há uma defasagem na compreensão dos conceitos estatísticos e na aplicação dos cálculos (softwares de uso intuitivo podem auxiliar nestes cálculos); esclarecer aos alunos de que a praticidade de aplicação dos procedimentos estatísticos promovida pela informática pode aproximar o perigo do uso inadequado da estatística; compreender que existem diversas formas de resolução de um problema estatístico.

Desta forma, o objetivo geral do presente trabalho é desenvolver uma ferramenta computacional de uso livre e offline (sem a necessidade de conexão com a internet), que auxilie o ensino dos conteúdos da estatística descritiva, tais como, organização de dados, distribuições de frequências, representações gráficas, medidas de posição dispersão, assimetria e curtose, correlação e regressão, abordados tanto na educação básica, como também em algumas disciplinas do ensino superior.

Nesta pesquisa, além de aliar as indicações de nossos referenciais teóricos, à crescente verificação da necessidade de adequarmos nossas escolas ao mundo cada vez mais tecnológico, acreditamos que podemos estimular o desenvolvimento de mais ferramentas computacionais voltadas ao ensino, de forma a contribuir positivamente na qualidade do processo de ensino desses conteúdos.

Essa ferramenta deve ser de fácil manuseio e o mais simples possível, pois observando outras ferramentas computacionais utilizadas na educação matemática e/ou estatística, percebemos que, em algumas delas, como o Microsoft Excel, é necessário a utilização de algum tipo de linguagem de programação ou sequência de comandos para realização das operações, que desvia o foco do que realmente é importante, o processo de ensino-aprendizagem dos conceitos de estatística descritiva.

Sentimo-nos motivados a contribuir com as questões relacionadas à exploração dos conteúdos de estatística, chegando ao seguinte questionamento: **É possível desenvolver uma ferramenta computacional para o auxílio do ensino da estatística?**

Desta forma, este trabalho apresenta o produto educacional SISTAT, desenvolvido durante nossa pesquisa, suas especificações, desenvolvimento e avaliação. Assim, além da introdução fornecida neste capítulo, elaborou-se mais 4 capítulos que se seguem.

2. REVISÃO BIBLIOGRÁFICA, REFERENCIAL TEÓRICO E METODOLOGIA DO TRABALHO

De acordo com os Parâmetros Curriculares Nacionais (PCN) (BRASIL, 2001), que foram criados pelo Ministério da Educação como sugestão de roteiro mínimo para o professor seguir em sala de aula, para a disciplina Matemática no Ensino Fundamental, temos o bloco "TRATAMENTO DA INFORMAÇÃO" que inclui: Contagem, Probabilidade e Estatística. Dentro dessa perspectiva, espera-se que o aluno entre em contato com as noções da Estatística Descritiva, conhecendo os seguintes tópicos: Coleta e tratamento de dados, tabelas, gráficos e quantificações.

> A análise de dados tem sido essencial em problemas sociais e econômicos, como nas estatísticas relacionadas a saúde, populações, transportes, orçamentos e questões de mercado. Propõe-se que constitua o terceiro eixo ou tema estruturador do ensino, e tem como objetos de estudo os conjuntos finitos de dados, que podem ser numéricos ou informações qualitativas, o que dá origem a procedimentos bem distintos daqueles dos demais temas, pela maneira como são feitas as quantificações, usando-se processos de contagem combinatórios, frequências e medidas estatísticas e probabilidades. Este tema pode ser organizado em três unidades temáticas: Estatística, Contagem e Probabilidade. (BRASIL, 1999, p.126)

Ainda segundo os PCN (BRASIL, 2001, p.134), para a disciplina de Matemática no Ensino Médio, houve uma divisão em três blocos, dos quais um deles trata da "ANÁLISE DE DADOS" que também inclui: Contagem, Probabilidade e Estatística. No entanto, de uma forma mais aprofundada, permitirá que o aluno, além de dominar seus conteúdos, seja capaz de analisá-los, interpretá-los e tomar decisões. O estudo dos conteúdos relacionados a esse bloco, segundo os PCN, favorece "o desenvolvimento de certas atitudes, como posicionar-se criticamente, fazer previsões e tomar decisões frente às informações veiculadas pela mídia, livros e outras fontes".

Cazorla (2006), em sua pesquisa, apresenta resultados que apontam para o crescente interesse de pesquisadores e professores em melhor trabalhar os conceitos de Estatística na educação básica, principalmente no ensino médio. Aponta também para a necessidade de uma comunidade docente consciente da importância dos resultados que venham a ser obtidos a partir de investigações, pois

desenvolvendo o ensino da Estatística, teremos um valioso instrumento para o exercício da cidadania.

Gal (2006) entende que uma pessoa inserida em uma sociedade industrializada é considerada letrada em estatística quando consegue ler criticamente informações estatísticas, expressando opiniões sobre as consequências dessas informações e das conclusões fornecidas. Dentro deste contexto, estão incluídos os conceitos e procedimentos básicos de estatística, como apresentações gráficas e tabulares, medidas de resumo, medidas de tendência central e de variabilidade.

Lopes (1998) afirma que atender a necessidade básica de formação do aluno, considerando uma sociedade informatizada, requer levá-lo ao desenvolvimento do pensamento estatístico e probabilístico. Afirma, ainda, que a Estatística tem um papel essencial na formação do cidadão, uma vez que possibilita lidar com a aleatoriedade e o acaso.

Penteado & Borba (2005), apresenta argumentos acerca do uso das Tecnologias da Informação e Comunicação (TIC) no processo de ensino-aprendizagem, discutindo resultados de investigações realizadas com alunos e professores utilizando softwares no ensino-aprendizagem de matemática.

A necessidade da criação de um software educacional para o ensino de estatística vem ao encontro das ideias de Godino (1995) em seu trabalho intitulado ¿Qué aportan los ordenadores a La enseñanza y aprendizaje de la estadística?, que trata da análise da influência de computadores no desenvolvimento da metodologia de ensino da estatística, apresentando objetivos educacionais que podem ser alcançados a partir da disponibilidade de softwares voltados para o ensino. Afirmando que a criação de um ambiente operacional que permita criar, observar, refletir e interagir sobre fenômenos que são difíceis de explorar, sem uma ferramenta de software.

Como suporte para a área de programação e desenvolvimento computacional baseamo-nos em Cantu (2004), que realizou seus estudos em desenvolvimento de softwares na linguagem de programação Borland Delphi.

Complementando o referencial teórico utilizado nesse trabalho temos HOEL (1979), como suporte para os conceitos estatísticos.

2.1 INFORMÁTICA EDUCATIVA NO BRASIL

Segundo MORAES (2003), a informática educativa no Brasil teve seu início na década de 1970, quando se discutiu o uso de computadores no ensino da Física, em um seminário promovido pela Universidade de São Carlos, assessorado por um especialista da Universidade de Dathmouth/EUA.

Neste mesmo período, iniciava-se no Brasil a ideia de que uma informatização da sociedade poderia proporcionar um desenvolvimento social, político, tecnológico e econômico.

Pouco tempo depois, o Ministério da Educação identificou que a relação entre a informática e a educação era necessária para o sucesso de informatização da sociedade. Em 1982, assumiu o compromisso de criar instrumentos e mecanismos necessários que permitissem o desenvolvimento das primeiras investigações na área. Já nesta época, o Plano Nacional de Desenvolvimento indicava que o uso das tecnologias educacionais seria um catalisador de melhorias na qualidade na educação.

Ainda segundo MORAES (2003), os precursores responsáveis pelas primeiras investigações acerca do uso de computadores na educação em nosso país encontravam-se na Universidade Federal do Rio de Janeiro (UFRJ), Universidade Estadual de Campinas (UNICAMP) e Universidade Federal do Rio Grande do Sul (UFRGS).

No final da década de 1960, a UFRJ foi a primeira instituição a se envolver no uso do computador na educação. Nessa época o computador era objeto de estudo e pesquisa, ensejando na criação de uma disciplina voltada para o ensino de informática. Em 1973, o Núcleo de Tecnologia Educacional para a Saúde, iniciou a aplicação da informática como tecnologia educacional no contexto da avaliação formativa, além da utilização do recurso em simulação.

Ainda em 1973, na UFRGS, surgiram as primeiras iniciativas de estudo com recurso instrumental de terminais de teletipo e display, em um experimento de Física para alunos do 3º grau. O Centro de Processamento de Dados (CPD-UFRGS) desenvolveu o software SISCAI, utilizado em experiência de avaliação de alunos de pós-graduação em educação. Estas e outras experiências foram realizadas até 1980 em computadores de grande porte, quando o computador era visto como recurso

auxiliar do professor no ensino e na avaliação, analisando a atitude e a ansiedade dos alunos em processos interativos com o computador.

Em 1975, na UNICAMP, um grupo de pesquisadores, sob a coordenação do professor Ubiratan D'Ambrósio, desenvolveu o documento *Introdução a computadores* para ser utilizado com alunos do 2º grau (atualmente, ensino médio), financiado pelo MEC, através de convênio com o Programa de Reformulação e Melhoria do Ensino (PREMEN), atualmente extinto. Em 1976, após visita ao Instituto de Tecnologia de Massachusetts (MIT/USA), um grupo de professores iniciou um projeto que investigaria a utilização de computadores em educação, com a formação de um grupo interdisciplinar das áreas de computação, linguística e psicologia educacional. No ano seguinte, passou a envolver crianças sob a coordenação de dois mestrandos em computação.

Segundo MORAES (1997), a implantação de um programa de informática na educação no Brasil foi iniciada através do Seminário Nacional de Informática em Educação, em 1981, na Universidade de Brasília. Nesse contexto originou-se o projeto EDUCOM, que possuía uma filosofia diferente no uso do computador na educação, nas áreas de Matemática, Física, Química, Letras e Biologia, passando a ter um papel de ferramenta para aprendizagem, ao invés de uma máquina de ensinar. Neste formato, o aluno seria o construtor de sua própria aprendizagem por exploração e descoberta. Em 1986, foram criados programas com a finalidade de capacitar professores de 1º e 2º graus (atualmente ensino fundamental e médio) em informática e iniciou-se a implantação de infraestruturas de suporte nas secretarias estaduais de educação, escolas técnicas e universidades.

De acordo com Quartiero (2007), em 1997, foi lançado pelo MEC o Programa Nacional de Informática na Educação, com o propósito de inserir tecnologia de informática nas escolas da rede pública de ensino, com o objetivo de implantar uma política pública de informatização educativa e criar centros de pesquisa e capacitação na área.

2.2 ENSINO DA MATEMÁTICA/ESTATÍSTICA E TECNOLOGIA

Podemos constatar a importância da utilização de tecnologias educacionais, conforme palavras de D'Ambrósio (1990):

> Creio que um dos maiores males que a escola pratica é tomar a atitude de que computadores, calculadoras e coisas do gêneros não são para as escolas dos pobres. Ao contrário: uma escola de classe pobre necessita expor seus alunos a esses equipamentos que estarão presentes em todo o mercado de futuro imediato. Se uma criança de classe pobre não vê na escola um computador, como jamais terá oportunidade de manejá-lo em sua casa, estará condenada a aceitar os piores empregos que se lhe ofereçam. Nem mesmo estará capacitada para trabalhar como um caixa num grande magazine ou num banco. É inacreditável que a educação matemática ignore isso. Ignorar a presença de computadores e calculadoras é condenar os estudantes a uma subordinação total a subempregos. (p.17).

D'Ambrósio (1990), incentiva o uso de tecnologias na educação matemática, e afirma que é mais que um elemento motivador em sala de aula, propicia aos alunos novas formas de geração e disseminação de conhecimento. São possibilidades de se vivenciar o saber matemático em um contexto tecnológico, com qualidade, que para o jovem é muito importante em sua qualificação profissional em qualquer área. As possibilidades tecnológicas podem e devem ser exploradas por uma política educacional comprometida com a manutenção de uma educação que se integre com uma sociedade que se transforma a cada dia.

Ideias semelhantes são encontradas nos PCN (BRASIL,1997) e apresentadas às organizações escolares, onde a tecnologia é apontada como um elemento que agrega possibilidades ao processo de ensino e aprendizagem de Matemática, devido a sua presença cada vez maior na sociedade moderna. A utilização de recursos tecnológicos, como aplicativos abertos, permitem ao aluno a criação de situações-problema dentro de sua realidade cultural, com a possibilidade de explorá-las ativamente.

Segundo Jolliffe (2007), em se tratando do ensino da Estatística, as maiores alterações são resultado da revolução tecnológica. Destaca ainda, o fato de que a tecnologia reduz o tempo de cálculo, o que permite que se trabalhe em sala de aula com aplicações reais, realizando projetos e investigações estatísticas. Hoje em dia é possível pedir que os alunos resolvam situações-problema baseados em dados reais e que apresentem os resultados de uma maneira que não era realizável no passado.

Batanero (2003) diz que entre outras potencialidades é possível a criação de micromundos virtuais, onde o aluno pode explorar os conceitos de probabilidades, substituindo as demonstrações formais por raciocínios mais intuitivos e inferindo de uma forma mais espontânea. O conceito de simulação no ensino de probabilidades pode ser largamente explorado, tendo em vista a dificuldade de ir além dos tradicionais jogos de sorte-azar e de problemas simples, que normalmente são resolvidos através de métodos analíticos e combinatórios. A simulação (especialmente com computador) permite que experimente fenômenos aleatórios num curto espaço de tempo, variando as condições e observando os resultados.

3. A ESTATÍSTICA NO SISTAT

O que modernamente se conhece como Ciências Estatísticas, ou simplesmente Estatística, é um conjunto de técnicas e métodos de pesquisa que, entre outros tópicos envolve o planejamento do experimento a ser realizado, a coleta qualificada dos dados, a inferência, o processamento, a análise e a disseminação das informações. A Estatística tem por objetivo fornecer métodos e técnicas para lidarmos, racionalmente, com situações sujeitas a incertezas.

A palavra Estatística derivou-se de “Status” (Estado) e significava um conjunto de descrições e dados relativos ao Estado, representando assim, uma poderosa ferramenta para os administradores.

A História nos informa que cerca de 3000 anos AC, os babilônios, chineses e egípcios já realizavam processos de contagem com as populações (censos). A Bíblia exemplifica levantamentos populacionais em Israel objetivando descobrir pessoas aptas para a guerra. Um dos levantamentos estatísticos mais completos que se tem notícia (considerando a época) foi realizado no século XI na Inglaterra por determinação de Guilherme “o conquistador”, no qual deveriam estar incluídas informações sobre terras, empregados, animais existentes, forma de uso da terra, entre outras, tudo para servir de base para o cálculo de impostos.

Na Idade Média colhiam-se informações, geralmente com finalidades tributárias ou bélicas. A partir do século XVI começaram a surgir as primeiras análises sistemáticas de fatos sociais, como batizados, casamentos e funerais. Por volta do século XVIII, as tabelas tornaram-se mais completas, surgiram as primeiras representações gráficas e o cálculo das probabilidades, e a Estatística deixou de ser uma simples catalogação de dados numéricos para se tornar uma nova ciência.

3.1 – DEFINIÕES BÁSICAS DA ESTATÍSTICA

A utilização da Estatística é cada vez mais acentuada em qualquer atividade profissional da vida moderna. Nos seus mais diversificados ramos de atuação, as pessoas estão frequentemente expostas à Estatística, utilizando-a com maior ou menor intensidade, Toledo (1985).

População e amostra

Na Estatística, dois conceitos são muitos utilizados, chamados de População ou Universo Estatístico e Amostra. **População ou universo estatístico** é o conjunto da totalidade dos indivíduos sobre o qual se faz uma inferência recebe o nome de população ou universo, isto é, população é o conjunto constituído por **todas** as unidades experimentais (ou observacionais) que apresentem pelo menos uma característica comum.

Amostra

A amostra pode ser definida como um subconjunto, isto é, uma parte da população, através da qual se faz um estudo estatístico, cujos resultados podem ser replicados para a população.

Variável

Variável é, convencionalmente, o conjunto de resultados possíveis de um fenômeno, isto é, uma característica qualquer de interesse em um estudo estatístico associada a uma população ou amostra. É assim chamada por apresentar variação entre os elementos que compões uma população ou amostra. É possível distingui-las como Qualitativas e Quantitativas.

Variável qualitativa recebe essa classificação quando suas observações são expressas por atributos, propriedades ou condições particulares. Podem ser nominais ou ordinais. **Variável Qualitativa Nominal** não existe uma ordenação entre os possíveis resultados.**Variável Qualitativa Ordinal** existe uma ordenação entre os possíveis resultados.

Variável quantitativa recebe essa classificação quando suas observações são expressas por valores numéricos. Em geral resulta de medição, enumeração, contagem ou cálculo. Podem ser Contínuas ou Discretas. A **Variável Quantitativa Contínua** seus possíveis resultados estão associados aos conjuntos dos Reais ($\mathbb{R}$), costuma receber informações de medições ou relações, enquanto a **Variável Quantitativa Discreta** seus possíveis resultados serão sempre valores inteiros.

Apresentação dos dados em tabelas

Tabela é um quadro que resume informações. Tendo em vista que uma tabela deverá ser uma forma objetiva de se apresentar o comportamento de variáveis, o

que se deve buscar são representações simples que possibilitem ao observador a compreensão do fenômeno em estudo. Uma tabela deve apresentar a seguinte estrutura: **Título,** deve conter o suficiente para que sejam respondidas as seguintes questões: O que está representando?, Onde ocorreu? e Quando ocorreu?; **Corpo,** é representado por linhas e colunas dentro dos quais serão registrados as informações e os dados numéricos; e **Rodapé,** é reservado para observações pertinentes a tabela, bem como para o registro e a identificação da fonte dos dados.

Séries estatísticas

Denominamos série estatística a qualquer coleção de dados referentes a uma mesma ordem de classificação. As séries estatísticas são geralmente divididas em: séries temporais, séries geográficas e séries específicas. **Séries Temporais,** caracterizam-se por apresentar o tempo variável, mantendo-se fixos o local e a espécie; **Séries Geográficas,** são aquelas em que o fator geográfico (local) varia, permanecendo fixos o tempo e a espécie e **Séries Específicas,** que tem por características variar o fator especificativo (fenômeno), enquanto permanecem fixos tempo e o local.

Distribuição de frequências

Ao estudarmos conjuntos de dados numéricos com uma grande quantidade de elementos, é conveniente organizá-los e resumi-los em tabelas chamadas de **distribuição de frequências**. **Tabela primitiva** é um agrupamento de dados não ordenados numericamente e **Rol** é grupamento de dados após sua ordenação numérica.

Para a construção de distribuição de frequências é necessário dispor pelo menos de dados em uma tabela primitiva ou rol. A **Distribuição de Frequências** é assim chamada porque relaciona variáveis quantitativas com contagens (ou frequências) do número de valores que se enquadram em grupo de dados ou uma categoria. Uma distribuição de frequências pode ser: sem intervalos de classes ou com intervalos de classes.

O primeiro passo na construção de uma distribuição de frequências com intervalos de classes, é a determinação do **número de classes** (k) e, consequentemente, a determinação da **amplitude do intervalo de classe** (h). Para determinarmos o número de classes de uma distribuição, vamos utilizar a **Regra de**

Sturges*, que nos dá o número de classes (k) em função do número de valores da variável (n): $k = 1 + 3{,}22.\log n$. Para calcularmos a amplitude das classes (h), basta dividirmos a amplitude total (AT) pelo número de classes (k), isto é, $h = \frac{AT}{k}$, onde, AT é a amplitude total, que é calculada da seguinte forma: AT = Max_{rol}– Min_{rol}.

Elementos de uma distribuição de frequências

Existem alguns elementos que fazem parte de uma distribuição de frequências. Citaremos apenas os principais: **Limites de classe –** são os extremos de cada classe. Representamos por L_i o limite superior da classe i, e por ℓ_i o limite inferior da classe i; **Ponto médio da classe (x_i) –** é o ponto que divide o intervalo de classe em duas partes iguais, isto é, $x_i = \frac{L_i + l_i}{2}$; **Frequência acumulada (F_i) –** é a soma das freqüências simples dos valores inferiores ou iguais ao valor dado na linha k, ou seja, $F_k = f_1 + f_2 + \cdots + f_k$; **Frequência relativa simples em percentual (fr_i%) –** são os quocientes entre suas respectivas freqüências simples e o total. O resultado multiplica-se por 100, isto é, $fr_i\% = \frac{f_i}{N}.100$; **Frequência relativa acumulada em percentual (Fr_i%) –** é a soma das freqüências relativas simples dos valores inferiores ou iguais ao valor dado na linha k, ou seja, $Fr_k\% = fr_1\% + fr_2\% + \ldots + fr_k\%$.

Representação gráfica

Um gráfico estatístico é uma forma de apresentação dos dados estatísticos, cujo objetivo inicial é o de produzir em geral, uma interpretação mais rápida do fenômeno em estudo. A representação gráfica de um fenômeno deve obedecer a certos requisitos fundamentais: **Simplicidade,** o gráfico deve ser destituído de detalhes de importância secundária, que possam levar o observador a uma análise errônea; **Clareza –** o gráfico deve possibilitar uma correta interpretação dos valores representativos do fenômeno em estudo; e **Veracidade –** o gráfico deve sempre expressar a verdade sobre o fenômeno em estudo. Os gráficos utilizados para representar distribuições de frequências simples e acumuladas são gráficos tipicamente de análise.

* Essa fórmula apresenta pequenas variações de autor para autor. Optamos em trabalhar segundo Fonseca (1996).

Os principais gráficos utilizados para representar distribuições de frequências com dados não agrupados em classes são: gráfico em linhas, gráfico em barras verticais ou horizontais, gráfico em setores e ramos e folhas. Enquanto os principais gráficos utilizados para representar distribuições de frequências com dados agrupados em classes são: Histograma, Polígono de Frequência e Ogiva. Temos também, o gráfico Boxplot que pode representar distribuições de frequências com dados agrupados ou não agrupados em classes. Cabe informar que os gráficos acima citados podem ser construídos com o auxílio do SISTAT.

Não há uma maneira única de representar graficamente uma série estatística. A escolha do gráfico mais apropriado ficará a critério do analista. Contudo, os elementos simplicidade, clareza e veracidade devem ser considerados durante a elaboração de um gráfico.

Medidas de posição

O estudo sobre distribuições de frequências, permite-nos descrever, o comportamento global de uma variável. Ocorre, todavia que trabalhar com uma distribuição de frequências completa, muitas vezes, é difícil, razão pela qual se costuma lançar mão de determinadas medidas. Essas medidas sumarizam certas características importantes da distribuição de frequências. Há diversas medidas que possibilitam condensar as informações dentro da fase analítica da Estatística Descritiva. Dentre as medidas de posição mais importantes estão as **medidas de tendência central**. Essas medidas recebem tal denominação pelo fato dos dados observados tenderem, em geral, a se agrupar em torno dos valores centrais. Dentre as medidas de tendência central, destacamos: a média aritmética; a moda e a mediana.

A **média aritmética** é, de modo geral, a mais comum de todas as mensurações numéricas descritivas. Outras medidas de posição, também importantes, são as **separatrizes**, que têm como objetivo dividir uma distribuição em partes iguais. Dentre as principais separatrizes, destacamos: quartis, quintis, decis e percentis. As medidas de tendência central e as separatrizes podem ser calculadas, tanto para dados não agrupados em classes, quanto para dados agrupados em classes.

As medidas de tendência central sintetizam os dados de uma distribuição, passando-nos o comportamento geral das observações estudadas. Existem vários tipos de medidas utilizadas como medidas de tendência central.

A **média aritmética** é o valor obtido quando somamos todos elementos de uma série e o dividimos pela quantidade de dados. Para calcularmos a média aritmética de uma série de dados não agrupados em classes temos $\rightarrow \bar{x} = \frac{\Sigma x_i}{n}$, onde x_i são os valores da série de dados e n é o total de dados. Já para os dados agrupados em classes temos $\rightarrow \bar{x} = \frac{\Sigma x_i f_i}{n}$, onde x_i são os pontos médios das classes, f_i são as frequências simples e n é o total.

A **moda** é a ocorrência que aparece com maior frequência dentre uma série de valores, isto é, é o valor que mais se repete. Para calcularmos a moda de uma série de dados não agrupados em classes, basta-nos fazer essa rápida contagem. Porém, quando os dados encontram-se agrupados em uma distribuição de frequências com intervalos de classes, vamos precisar usar uma pequena fórmula.

Nos **dados não agrupados em classes** $\rightarrow$ mo = valor que mais se repete e para **dados agrupados em classes** $\rightarrow mo = l_i + h.\frac{f_{mo} - f_{ant}}{2.f_{mo} - (f_{ant} + f_{post})}$, onde l_i é o limite inferior da classe modal, h é a amplitude da classe modal, f_{mo} é a frequência simples da classe modal, f_{ant} é a frequência anterior à classe modal e f_{post} é a frequência posterior à classe modal. A moda, diferentemente das outras medidas de tendência central, pode ser obtida mesmo que a variável seja qualitativa.

A **Mediana** é o número que divide uma série ordenada de dados em duas outras séries com a mesma quantidade de elementos. Assim como na moda, quando os dados aparecem agrupados em classes precisaremos de uma fórmula. Dados não agrupados em classes $\rightarrow$ md = valor central da série de valores ordenada, e dados agrupados em classes $\rightarrow md = l_i + h.\frac{E_{md} - F_{ant}}{f_{md}}$, onde l_i é o limite inferior da classe mediana, h é o intervalo da classe mediana, E_{md} é o elemento mediano, onde $E_{md} = \frac{n}{2}$, F_{ant} é a frequência acumulada anterior à classe mediana e f_{md} é a frequência simples da classe mediana.

As **separatrizes** dividem uma distribuição ordenada em partes iguais. Dentre as principais separatrizes destacamos os quartis, os quintis, os decis e os percentis.

Os **Percentis** são medidas que dividem uma distribuição ordenada em cem partes iguais. Variam de P_1 até P_{99}. Para calcularmos qualquer percentil, devemos proceder, da seguinte maneira: **Dados não agrupados em classes** →Ordenar os dados. Calcular a posição do j-ésimo percentil, $Ep_j = \frac{j.n}{100}$, onde n é número total de dados. Em seguida, localize o j-ésimo percentil, proporcionalmente, de acordo com a posição Ep_j encontrada. E para d**ados agrupados em classes** → $p_j = l_i + h.\frac{Ep_j - F_{ant}}{fp_j}$

onde l_i é o limite inferior da classe do percentil j, h é a amplitude da classe do percentil j, Ep_j é a posição do percentil j e vale $Ep_j = \frac{j.n}{100}$, F_{ant} é a frequência acumulada anterior à classe do percentil j e fp_j é a frequência simples da classe do percentil j.

Medidas de dispersão e assimetria

Um conjunto de valores pode ser convenientemente sintetizado, por meio de procedimentos matemáticos, isto é, através do cálculo da média, moda ou mediana. No entanto, quando se trata de interpretar dados estatísticos, mesmo aqueles já convenientemente simplificados, é necessário ter-se uma ideia retrospectiva de como se apresentavam na distribuição de frequências. Assim, não é o bastante dar uma das medidas de posição para caracterizar perfeitamente um conjunto de valores. Vemos, então que a média, ainda que considerada como um número que tem a faculdade de representar uma série de valores, não pode, por si mesma, detectar o grau de homogeneidade ou heterogeneidade entre os valores que compõem um conjunto.

Trabalhamos as seguintes medidas de dispersão:Desvio padrão (s) ou (σ); Variância (s^2) ou (σ^2) e Coeficiente de variação (cv). Estas medidas podem ser calculadas, tanto para dados não agrupados em classes, quanto para dados agrupados em classes.

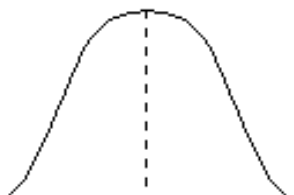
O **Desvio Padrão** de um conjunto de dados é uma medida que nos fornece a variação dos valores em relação à media aritmética. Assim, para calcularmos o desvio padrão de uma série de dados, usaremos para **Dados não agrupados em classes** → $s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}}$, onde x_i são os valores da série de dados, $\bar{x}$ é a média aritmética dos dados e n é o total de dados. E para d**ados agrupados em classes**

$\rightarrow s = \sqrt{\frac{1}{n-1}\left(\Sigma x_i^2 . f_i - \frac{(\Sigma x_i . f_i)^2}{n}\right)}$, onde x_i são os pontos médios das classes, f_i são as frequências simples e n é o total de dados. A variância de um conjunto de dados amostrais é igual a s^2 (desvio padrão ao quadrado).

O desvio padrão por si só não nos diz muita coisa. Assim, um desvio padrão de duas unidades pode ser considerado pequeno para uma série de valores cujo valor médio é 100; no entanto, se a média for 10, o mesmo não pode ser dito. Além disso, o fato de o desvio padrão ser expresso na mesma unidade dos dados limita o seu emprego quando desejamos comparar duas ou mais série de valores. Para contornarmos essas dificuldades e limitações, utilizaremos o **coeficiente de variação**. O coeficiente de variação é o quociente entre o desvio padrão e a média aritmética: $cv = \frac{s}{\bar{x}}.100$. Geralmente multiplica-se o coeficiente de variação por 100, para darmos o resultado em porcentagem. Para classificarmos uma distribuição em relação à sua variabilidade, tomando como base o coeficiente de variação, usaremos o seguinte critério: Baixa dispersão: cv≤ 15%; Média dispersão: 15% <cv< 30%; Alta dispersão: cv≥ 30%

Muitas medidas extraídas de pesquisas na área biomédica, medidas de produtos fabricados em série e também os erros de algumas medidas dão origem a gráficos, em forma de sino, simétricos em torno da média (Figura 1). Todas essas medidas têm distribuições que se aproximam da Distribuição Normal. A Distribuição Normal é uma distribuição conhecida e muito estudada, tendo inclusive todas as suas probabilidades calculadas. Assim, é importante sabermos se uma distribuição está próxima da Normal, pois, neste caso recairíamos em uma situação muito conhecida, onde poderíamos estudar a distribuição inicial utilizando a distribuição normal.

Figura 1 – Distribuição Normal

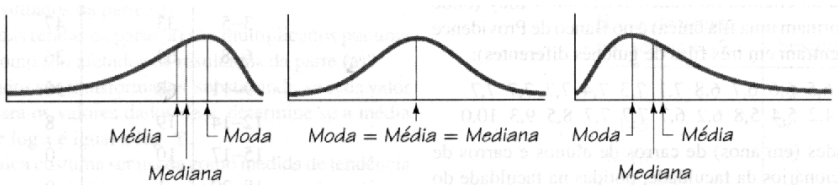


Média = Moda = Mediana

Fonte: Toledo (1985)

Dessa forma, um conceito muito utilizado em Estatística para detectar se uma determinada distribuição se aproxima da distribuição é a **assimetria**. Denominamos assimetria o grau de deslocamento lateral de uma distribuição em relação a uma distribuição simétrica, denominada distribuição normal. Este grau de deslocamento pode ser calculado através de uma fórmula conhecida como *coeficiente de correlação de Pearson*, descrito pela seguinte expressão: $As = \frac{3(\bar{x}-md)}{\sigma}$,onde $\bar{x}$ é a média aritmética, md é a mediana e s é o desvio padrão.Podemos ainda classificar uma distribuição quanto à sua assimetria. A Figura 2 ilustra tais situações:

Figura 2 – Assimetria



Distribuição Assimétrica para a Esquerda ou Distribuição Assimétrica Negativa: A média, e a mediana estão à esquerda da moda, isto é, As < 0.

Distribuição simétrica (ou assimetria nula): A média, a moda e a mediana coincidem, isto é, As = 0

Distribuição Assimétrica para a Direita ou Distribuição Assimétrica Positiva: A média, e a mediana estão à direita da moda, isto é, As > 0.

Fonte: Triola (2008)

Para classificarmos uma distribuição quanto a intensidade de sua assimetria, usamos o seguinte critério:Assimetria muito fraca: $0 < |As| \leq 0,5$; Assimetria fraca: $0,5 < |As| \leq 1,5$ e Assimetria forte: $|As| > 1,5$. Se uma distribuição de frequência for assimétrica muito forte, a medida de posição mais adequada para ser utilizada é a mediana.

Correlação e regressão

Existem situações nas quais interessa estudar o comportamento conjunto de duas variáveis. Por exemplo, dados peso e estatura de pessoas, pode haver interesse em estabelecer em que medida aumenta o peso, quando a estatura aumenta. O comportamento conjunto de duas variáveis quantitativas pode ser mensurado através de técnicas estatísticas conhecidas pelo nome de **correlação** *e* **regressão**.

O termo **correlação** significa "relação nos dois sentidos: descreve a associação entre duas variáveis, não fazendo julgamento sobre se uma é causa ou consequência da outra".A correlação é usada para estudar quão consistentemente duas variáveis mudam em conjunto. Quando isso ocorre, dizemos que há uma correlação ou uma covariação, cuja direção e magnitude podem ser quantificadas.

A correlação indica o grau de associação entre duas variáveis, ao passo que a **regressão** "diz respeito à capacidade de prever um valor baseado no conhecimento do outro". É importante ao fazer uma análise de regressão, determinar qual a variável dependente, ou seja, aquela que flutua em função da variável independente.

Neste capítulo identificaremos e exibiremos quando existe alguma **correlação linear** entre duas variáveis, ou seja, quando uma delas estiver, matematicamente, relacionada com a outra por meio de uma expressão linear. Tal expressão é bastante utilizada para realizar aproximações ou predições de eventos que estão, de uma forma ou de outra, relacionados às variáveis em estudo.

Após a obtenção dos dados experimentais, aos quais identificaremos por duas variáveis, digamos x e y, podemos dispô-los em forma tabelar para uma melhor visualização, ou apresentá-los em forma gráfica, indicando numa linha horizontal os valores de x e numa linha vertical, os valores de y. Tal apresentação gráfica recebe o nome de **diagrama de dispersão**.

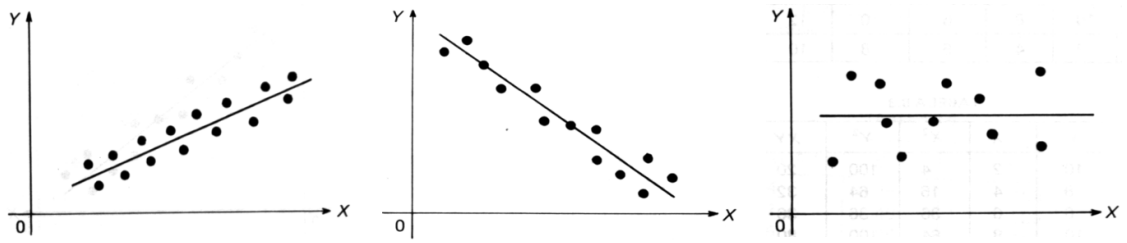
Num diagrama de dispersão, é muito comum ser possível visualizar uma curva regular que se aproxima dos pontos marcados (dados), denominada **curva de ajustamento**. Quando esta curva lembra uma reta e os dados parecem estar bem próximos dela, diz-se que há uma **correlação linear** entre as variáveis.

O **Coeficiente de Correlação Linear de Pearson**[1] (r) mede o grau de relacionamento linear entre os valores de duas variáveis em uma amostra de tamanho *n*. No caso das variáveis *x* e *y*, citadas acima, usaremos para o seu cálculo, a expressão $r = \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\sqrt{\left(\Sigma x^2 - \frac{(\Sigma x)^2}{n}\right)\left(\Sigma y^2 - \frac{(\Sigma y)^2}{n}\right)}}$. Onde o valor de r poderá variar de -1 a 1, isto é, $-1 \leq r \leq 1$.Se $r > 0$, diremos que a correlação linear é positiva; se $r < 0$, diremos que a correlação linear é negativa e ser = 0, diremos que não existe correlação linear entre as variáveis.

Dadas duas variáveis x e y, denominamos **diagrama de dispersão** ao gráfico formado pelos pares ordenados (x, y) representados em um sistema de coordenadas cartesianas.

As Figuras 3 abaixo ilustram alguns exemplos do que pode acontecer com as variáveis *x* e *y* em um diagrama de dispersão:

Figura 3 – Correlação Linear



Correlação Linear Positiva
$0 \leq r \leq 1$

Correlação Linear Negativa
$-1 \leq r \leq 0$

Correlação Nula
r = 0

Fonte: Toledo (1985)

Note que, o fato da correlação linear ser nula, não implica em afirmar a inexistência de correlação entre as variáveis x e y. A correlação pode ser, por exemplo: quadrática, exponencial, etc.

Devemos interpretar o valor calculado de r. Como já foi afirmado acima, o valor de r poderá variar de –1 a 1 (inclusive). O sinal indica somente se as duas variáveis variam no mesmo sentido (+) ou em sentidos opostos (–). Se o valor de r

[1] Karl Pearson (1857-1936), matemático britânico, foi um principais contribuintes da Estatística Moderna.

está próximo de 0, as duas variáveis não variam em conjunto. Mas se r está próximo de –1 ou 1, concluímos que existe "*correlação linear significativa"* entre x e y.

Como a expressão "*próximo de*" é vaga, os coeficientes de correlação, assim como muitas outras medidas estatísticas, são mais informativos quando expressos como *intervalos de confiança.* Assim, pode-se informar que um dado valor de r está dentro de um determinado "*Intervalo de Confiança*". O seguinte critério nos orienta no sentido de entender sobre as questões seguintes: "*correlação linear significativa"* e "*próximo de*".

Até aqui analisamos dados descritos por duas variáveis com o objetivo de determinar se havia correlação linear significativa entre ambas. Supondo então que esta correlação é linear, precisamos determinar a equação da reta que a representa. Essa reta é chamada "*reta de regressão"*, e sua equação é a "*equação de regressão*". Dada uma coleção de dados amostrais de dupla entrada, chama-se equação de regressão à expressão: $\hat{y}(x) = ax + b$

O gráfico da equação de regressão é chamado reta de regressão (ou reta de melhor ajuste, ou reta de mínimos quadrados). Essa definição expressa uma relação entre x, a variável independente, e $\hat{y}$, a variável dependente. É importante lembrar que estamos investigando apenas relações lineares. Os valores dos parâmetros a e b, citados acima, são obtidos com o auxílio das seguintes expressões: $a = \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\left(\Sigma x^2 - \frac{(\Sigma x)^2}{n}\right)}$ e $b = \bar{y} - a\bar{x}$, onde n é o número de observações; $\bar{x}$ é a média aritmética dos valores da variável x, isto é, $\left(\bar{x} = \frac{\Sigma x}{n}\right)$ e $\bar{y}$ é a média aritmética dos valores da variável y, isto é, $\left(\bar{y} = \frac{\Sigma y}{n}\right)$.

As equações de regressão podem ser úteis quando usadas para predizer (estimar) o valor de uma variável, quando conhecemos o valor da outra variável. Se a reta de regressão se ajusta bem aos dados, então tem sentido utilizar sua equação para fazer predições, desde que não ultrapassemos os limites dos valores disponíveis (exceto quando considerações teóricas ou experimentais demonstrem tal possibilidade). Entretanto, só devemos utilizar a equação de regressão se o coeficiente de correlação de Pearson (r) exceder o valor apresentado na Tabela 1, indicando a existência de uma correlação linear significativa.

Tabela 1 – Valores Críticos do Coeficiente de Correlação Linear de Pearson

n	α = .05	α = .01
4	,950	,999
5	,878	,959
6	,811	,917
7	,754	,875
8	,707	,834
9	,666	,798
10	,632	,765
11	,602	,735
12	,576,	,708
13	,553	,684
14	,532	,661
15	,514,	,641
16	,497	,623
17	,482	,606
18	,468	,590
19	,456	,575
20	,444	,561
25	,396	,505
30	,361	,463
35	,335	,430
40	,312	,402
45	,294	,378
50	,279	,361
60	,254	,330
70	,236	,305
80	,220	,286
90	,207	,269
100	,196	,256

4. METODOLOGIA

Após busca por recursos tecnológicos de livre reprodução disponíveis e voltados para o ensino da estatística, encontramos apenas o software StatD (que foi descontinuado). Constatando a escassez deste tipo de ferramenta, considerando a proposta do PCN (2002, p.118), com base nas indicações de nossos referenciais teóricos e com o objetivo de contribuir com as questões relacionadas à exploração dos conteúdos de estatística, chegamos ao seguinte questionamento: **É possível desenvolver uma ferramenta computacional para o auxílio do ensino da estatística?**

> Nos dias de hoje, deve-se reconhecer e acompanhar o desenvolvimento tecnológico, 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)

Segundo Borba e Penteado (2005), utilizar um novo recurso educacional em sala de aula não significa que os demais serão exterminados, no entanto, é necessário repensar a relevância de todos eles.

Partindo da análise de uma atividade educacional qualquer, envolvendo a utilização de um software em sala de aula, pode-se determinar os objetivos e limitações que o mesmo alcançará. Além dos aspectos humanos, como motivação, por exemplo, que raramente são considerados nas técnicas usuais da Engenharia de Software, levamos em consideração as possibilidades de aplicação em uma atividade educacional, respeitando os princípios teóricos que a orientam e dão embasamento.

Prosseguindo para a determinação de nossos objetivos gerais e específicos, surgiram os seguintes questionamentos: a) Qual o motivo da atividade educacional? b) Que necessidade a motiva? c) Quais conceitos serão abordados? d) Quem serão os sujeitos da atividade? Uma das ideias deste trabalho é propor um recurso computacional com uma interface simples que dispense necessidade do aprendizado de novas linguagens de programação ou sequência de comandos, além dos conteúdos abordados em sala de aula.

A ferramenta computacional produzida será de livre reprodução e utilização, a qual será incluída no repositório dos Recursos Educacionais Abertos (REA)[2].

Nossa pesquisa é qualitativa e nosso trabalho não é baseado em observações empíricas ou na análise de fenômenos humanos e/ou sociais ligados à educação. Nossa problemática se baseia em como implementar os conteúdos da estatística descritiva em uma ferramenta computacional de forma objetiva e de fácil manuseio, e que não dependa de programas auxiliares como os aplicativos desenvolvidos para rodar em JVM – Java Virtual Machine (Máquina Virtual Java) ou Microsoft .NET.

A identificação e delimitação dos pontos a serem abordados partiram das indicações de nossos referenciais teóricos.

Como linguagem de programação, optamos pelo Embarcadero Delphi, que é segundo Cantú (2003) uma solução para o desenvolvimento de aplicações desktop, aplicações multicamadas e cliente/servidor, verdadeiramente nativas para Microsoft Windows. Além de, futuramente, podermos utilizar o mesmo código para gerar softwares que rodem tanto no IOS (Sistema operacional móvel da Apple Inc) como no Android (Sistema operacional móvel de propriedade pelo Google), proporcionando a portabilidade.

4.1.Desenvolvimento da Ferramenta Computacional

4.1.1. A linguagem de programação Embarcadero Delphi

A linguagem de programação hoje conhecida como Embarcadero Delphi teve sua origem em 1970, com o lançamento da linguagem Procedure Pascal pela Borland Software Corporation, ganhando maior popularidade após 1980, com o nome comercial Turbo Pascal.

Na década de 1990 passou a utilizar o conceito Object Pascal, que dava suporte a orientação a objetos, conhecida como Programação Orientada a Objetos

[2]O termo Recursos Educacionais Abertos (REA) foi cunhado no Fórum de 2002 da UNESCO sobre Softwares Didáticos Abertos e designa "os materiais de ensino, aprendizagem e investigação em quaisquer suportes, digitais ou outros, que se situem no domínio público ou que tenham sido divulgados sob licença aberta que permite acesso, uso, adaptação e redistribuição gratuitos por terceiros. O licenciamento aberto é construído no âmbito da estrutura existente dos direitos de propriedade intelectual, tais como se encontram definidos por convenções internacionais pertinentes, e respeita a autoria da obra"

(POO). Segundo Bezerra (2015), hoje em dia, a orientação a objetos é um dos padrões mais difundidos, principalmente por ter apresentado uma grande evolução em questões ligadas à segurança e ao reaproveitamento de código, o que é muito importante no desenvolvimento de softwares.

Em 1995, a Borland Software Corporation lançou um ambiente visual, para o desenvolvimento de aplicações Microsoft Windows que utilizava a linguagem de programação Object Pascal, batizado de Delphi. Esta versão contava com recursos como conexão com banco de dados, programação orientada a objetos e ambiente de desenvolvimento rápido (RAD – Rapid Application Development).

Nos anos seguintes, com o lançamento de novas versões, mais recursos foram disponibilizados aos desenvolvedores, como a criação de aplicações multicamadas, manipulação de arquivos XML, suporte a Web Services e plataformas .NET. Houve também, o lançamento de uma versão chamada Kylix, que foi a primeira ferramenta RAD para desenvolvimento de aplicações Linux. O desenvolvimento do Kylix foi posteriormente descontinuado.

Em 2006, a Borland libera para uso público a linha Turbo, que foi o resultado de uma divisão da Suíte Borland Developer Studio (BDS 2006) em duas versões: A Explorer, com download e utilização gratuitos, direcionada a estudantes, iniciantes em programação e desenvolvedores por hobby e a Professional, versão paga, destinada a uso Profissional. Neste mesmo ano, criou uma subsidiária chamada CodeGear, que passou a cuidar do ambiente de desenvolvimento Delphi.

Por fim, em 2011 a Borland vendeu sua divisão responsável pelo desenvolvimento do Delphi, a CodeGear, para a Embarcadero Technologies, empresa norte-americana fundada em 1983, produtora de várias ferramentas para desenvolvimento de softwares para Microsoft Windows e outros sistemas operacionais.

Atualmente, o Delphi encontra-se na versão 23.0, intitulada Delphi 10 Seattle, lançada em Agosto de 2015.

4.1.2. Engenharia de Software

A ferramenta computacional desenvolvida neste trabalho foi batizada de SISTAT (Sistema Estatístico). O processo de construção de uma ferramenta computacional é uma atividade de Engenharia de Software, termo criado na década de 1960 e oficializado em 1968 na NATO Conference on Software Engineering (Conferência sobre Engenharia de Software da OTAN).

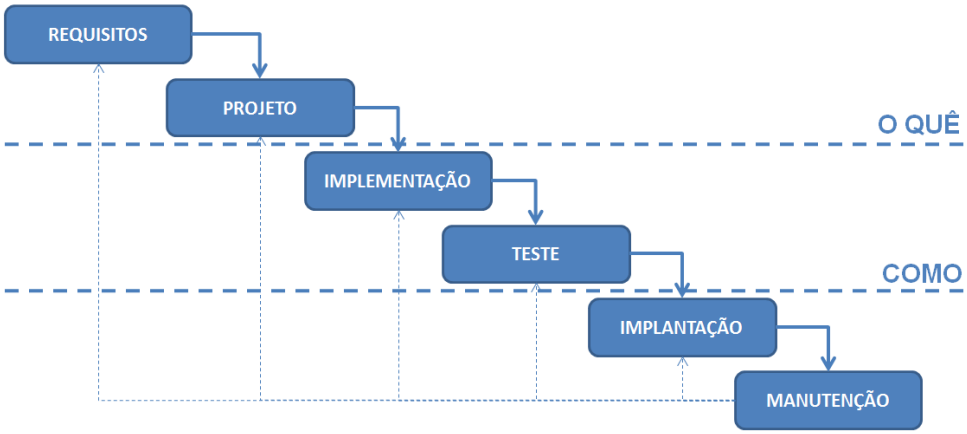
A engenharia de software segue um conjunto de métodos e técnicas e possui princípios científicos que envolvem o uso de modelos abstratos e precisos, permitindo especificar, projetar, implementar e manter um sistema computacional, concentrando-se nos aspectos práticos do processo produtivo, Pressman (2002).

Segundo Friedrich Ludwig Bauer, Engenharia de Software é a criação e a utilização de sólidos princípios de engenharia a fim de obter software de maneira econômica, que seja confiável e que trabalhe eficientemente em máquinas reais.

Ainda segundo Pressman (2002), todas as atividades e tarefas envolvidas no desenvolvimento, operação e manutenção de um software, isto é, toda a vida do sistema, desde a definição de seus requisitos até o encerramento de seu uso, são definidas pelo modelo de ciclo de vida, que normalmente é a primeira escolha a ser feita. O modelo de ciclo de vida determina a estrutura hierárquica de todos os processos envolvidos.

No desenvolvimento do software, optamos pelo ciclo de vida clássico, também conhecido como cascata, nesse processo de desenvolvimento, a modelação ocorre em função do ciclo da engenharia convencional, seguindo uma abordagem sistemática e sequencial. As fases são estanques e mesmo com as setas indicando retorno a fases anteriores, cada fase é vista isoladamente (Figura 4). O resultado de uma fase é entrada para outra.

Figura 4: Modelo de desenvolvimento em Cascata



Fonte: Royce (1970)

Nesse contexto, temos:

- **Requisitos:** Através da observação do processo de ensino-aprendizagem e o levantamento de suas carências, mapeamos as metas a serem atingidas, como o sistema deve se comportar, tanto em funcionalidade como em performance, especificando os requisitos necessários, inclusive na escolha da linguagem de programação.
- **Projeto:**Com base nos requisitos, o sistema foi decomposto em partes para que fosse possível criar as funções do sistema com segurança e representá-las de forma que pudessem ser transformadas em programas executáveis posteriormente.
- **Implementação:** Nesta fase, focamos nos algoritmos e na linguagem de programação (Delphi), criando as sintaxes. Nesse momento nasceu e se desenvolveu o código-fonte do SISTAT. Esbarramos em algumas limitações que provocaram algumas adaptações em relação ao projeto inicial.
- **Teste:** Testamos o software em ambiente controlado, confrontando com as especificações definidas nos requisitos. Verificamos as funcionalidades, os cálculos e gráficos gerados, classificando de acordo com os seguinte termos:

- Defeito (Fail): Instrução ou definição incorreta;
- Falha (Failure): Resultados incorretos;
- Erro (Mistake): Falha resultante de ação humana;

Nesta fase, definimos o padrão de instalação.

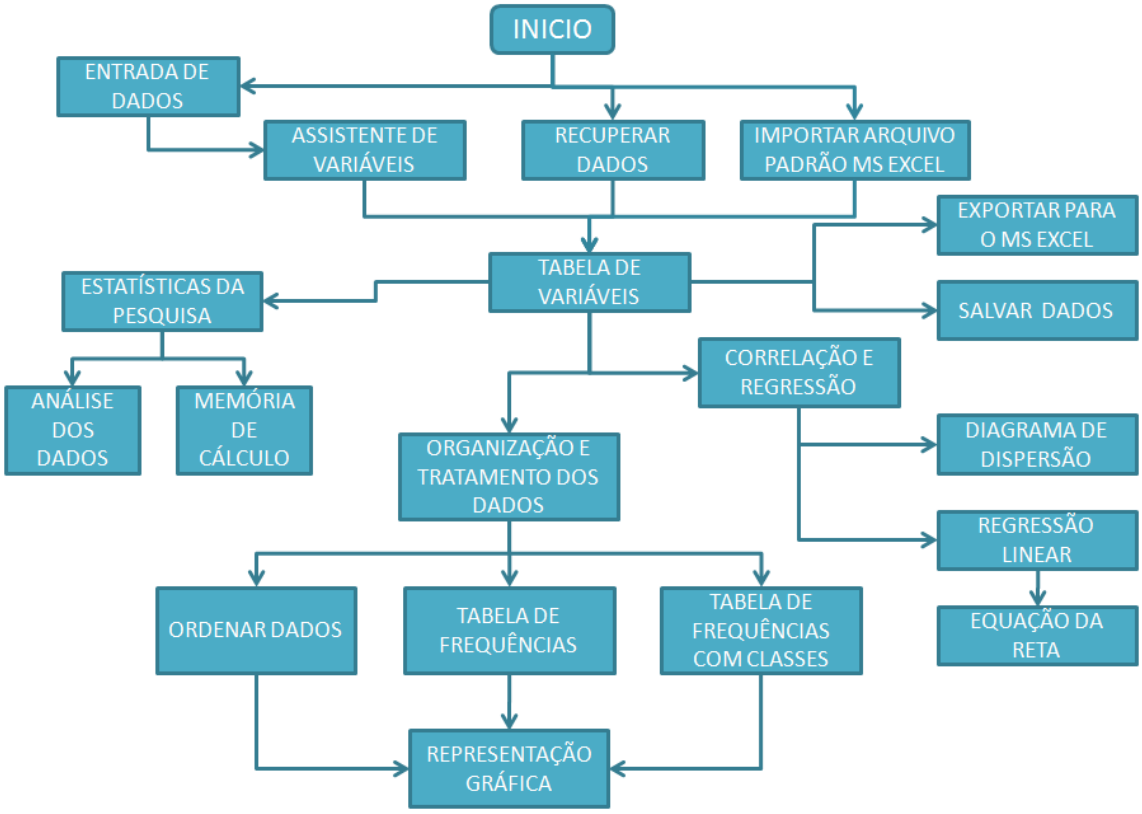
- **Implantação:** Rodamos o software em laboratórios educacionais, solicitando que usuários sem conhecimentos estatísticos fizessem uso do mesmo através de simulações sugeridas. As anomalias identificadas nessa fase foram reparadas e submetidas a novos testes. Verificou-se uma alternância entre as fases Implementação/Teste/Implantação até a liberação final.
- **Manutenção:** Será a fase mais longa do ciclo, visto que nela acontecerá a correção de erros remanescentes, adaptação a novas situações e necessidades.

4.1.3. Fluxograma

Segundo Debastiani (2015), Fluxograma é um diagrama que representa um processo ou um algoritmo, utiliza gráficos que ilustram uma sequência operacional. Possui aplicação em processos de fabricação ou comercialização de produtos, estratégias de marketing, aplicação de questionários, dentre outros. Nos projetos de desenvolvimento de software, utilização de fluxogramas proporciona uma racionalização dos processos e permite uma melhor compreensão e posterior otimização do código fonte.

Na Figura 5 segue o fluxograma de nossa ferramenta, embora não esteja demonstrado, há atalhos entre as áreas do sistema, com o objetivo de facilitar as comparações e análises.

Figura 5: Fluxograma



Fonte: Autor

5. PRODUTO EDUCACIONAL

Nossa ferramenta computacional foi desenvolvida como uma proposta de recurso tecnológico a auxiliar o processo de ensino-aprendizagem dos conteúdos da estatística descritiva.

Uma das principais preocupações foi com a facilidade de uso, inclusive a partir do processo de instalação, que acontece de forma bem simples, esclarecendo quais modificações serão realizadas no computador, e em poucas janelas já se inicia o processo de instalação. Após instalado, é criado um ícone na Área de Trabalho que será o atalho para acesso a nossa ferramenta computacional. Da mesma forma, caso necessário, a desinstalação será realizada de forma simples.

Foi desenvolvida sob a concepção de não consumir muitos recursos do sistema operacional, ser leve e rápido. Como requisitos mínimos temos o Quadro 1.

Quadro 1: Requisitos mínimos para a execução do SISTAT

Sistema Operacional Microsoft Windows XP ou posterior;
Processador de 300 megahertz (MHz) ou mais de velocidade;
128 megabytes (MB) ou superior;
1,5 gigabytes (GB) ou mais de espaço disponível em disco rígido;
Adaptador de vídeo e monitor super VGA (800X600) ou superior;
Teclado e mouse (ou dispositivo apontador compatível);

Fonte: Dados da Pesquisa

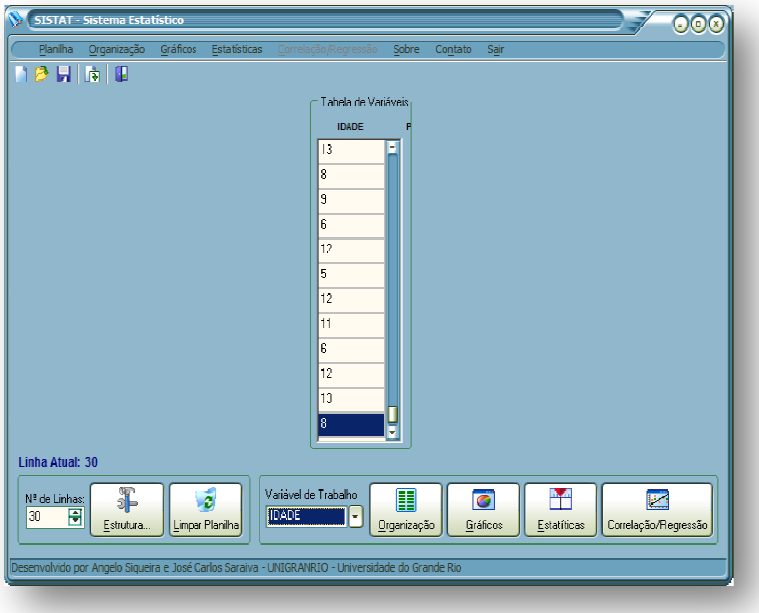
Uma massa de dados foi disponibilizada juntamente com o software para padronização dos estudos no primeiro contato feito pelos alunos, como sugestão. Essa massa de dados é composta por 52 ocorrências, dividas em 3 variáveis. Nomeadas como “SEXO” (Variável Qualitativa Nominal), “IDADE” (Variável Quantitativa Contínua) e “Nº DE FILHOS” (Variável Quantitativa Discreta).

5.1 ELEMENTOS DO PRODUTO EDUCACIONAL

A nossa ferramenta computacional pode receber até 10 variáveis simultâneas, os dados são apresentados na tela principal do software em forma tabulada. Como métodos de entrada de dados temos:

- Digitação diretamente através de tabela na tela principal;
- Recuperação de dados anteriormente salvos a partir do Systat;
- Importação a partir da área de transferência do Windows; e
- Importação a partir de arquivo no formato do Microsoft Excel.

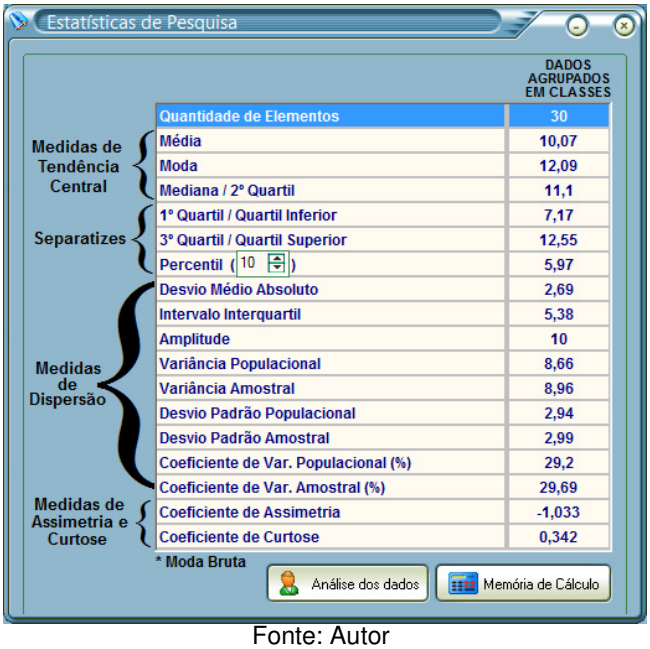
Figura 6: Tela Principal



Fonte: Autor

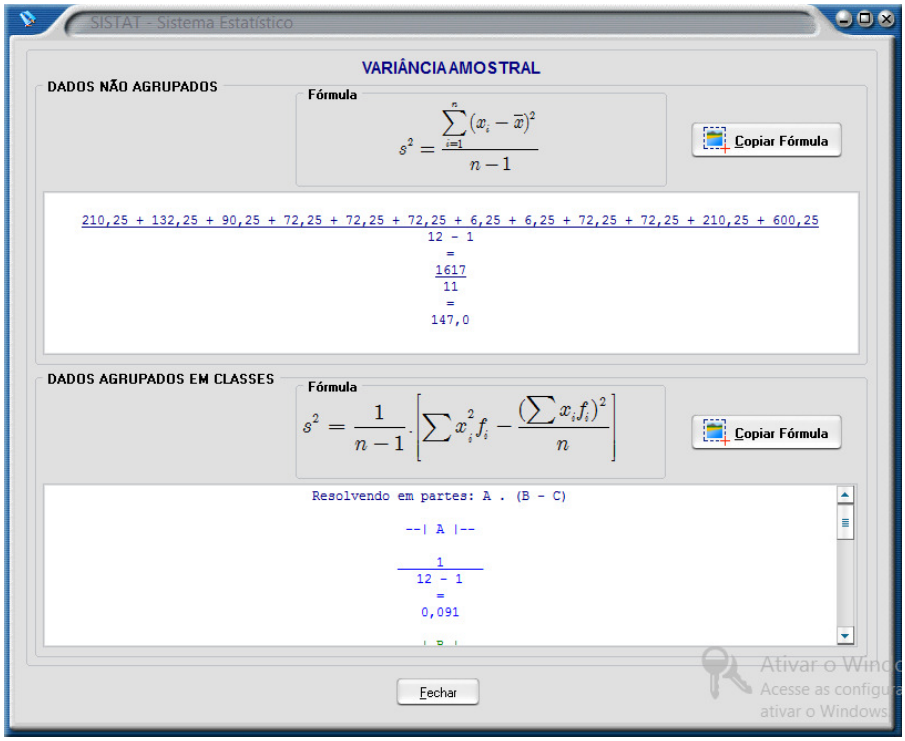
Com a entrada de dados concluída (Figura 6) pode-se verificar as estatísticas da pesquisa (Figura 7), quando teremos contato com as medidas estatísticas, bem como os métodos utilizados para seu cálculo (Figura 8), nesta mesma área podemos optar em iniciar a análise dos dados.

Figura 7: Estatísticas da Pesquisa



Fonte: Autor

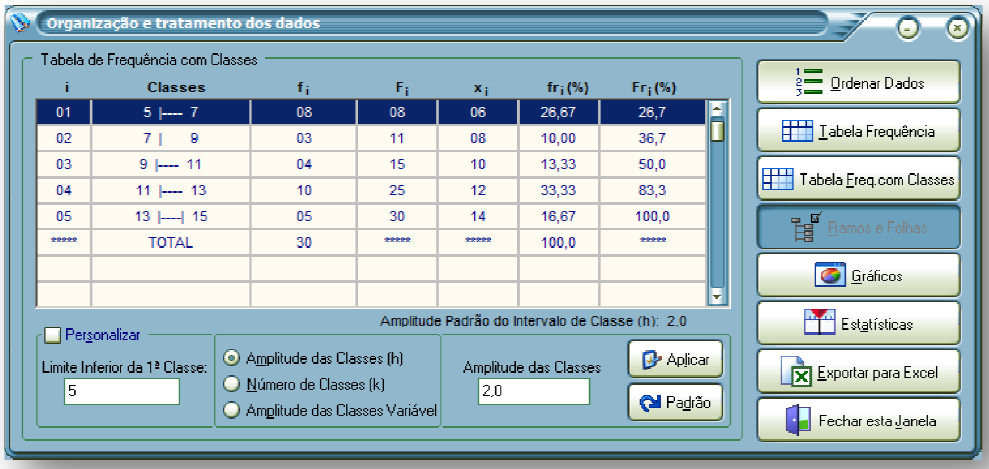
Figura 8: Memória de Cálculo



Fonte: Autor

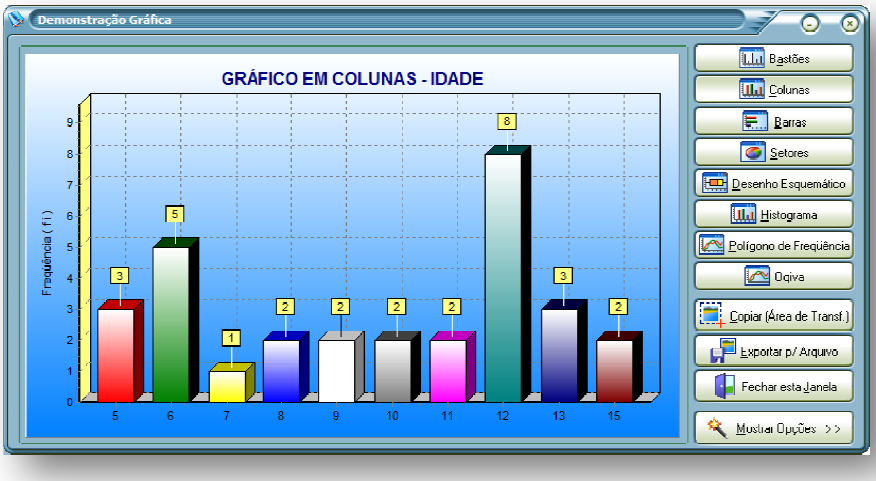
Ainda com os dados carregados no software, podemos ter acesso aos recursos de organização (Figura 9), tratamento da informação e representação gráfica (Figura 10).

Figura 9: Organização e tratamento da informação



Fonte: Autor

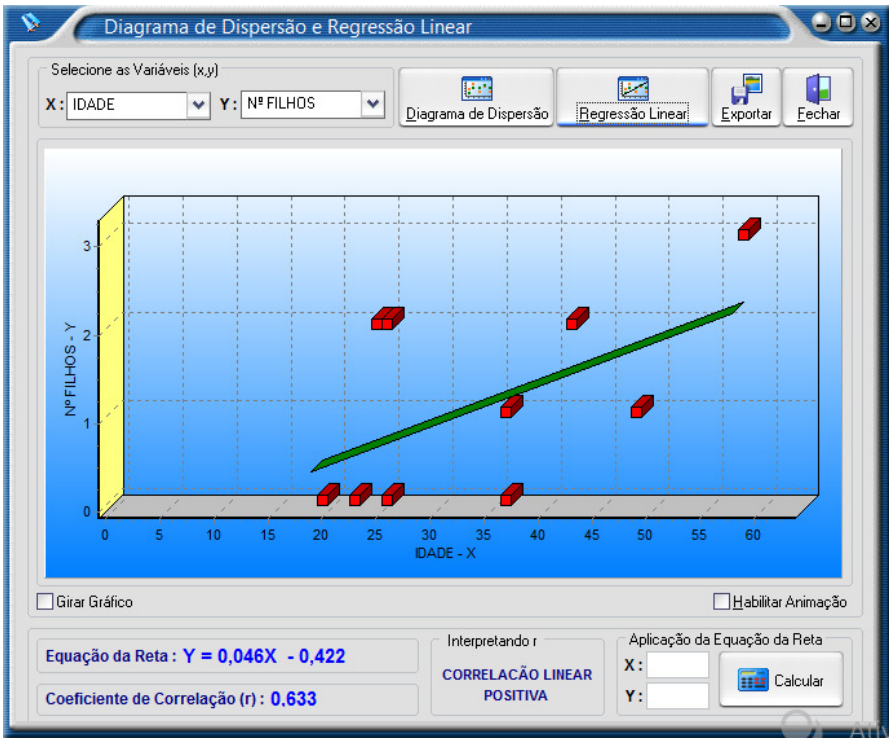
Figura 10: Representação gráfica



Fonte: Autor

Na hipótese de verificar se há correlação linear entre duas variáveis, basta selecionar as variáveis de interesse na respectiva área do software. Bem como, simular a regressão linear desta mesma correlação, se houver (Figura 11).

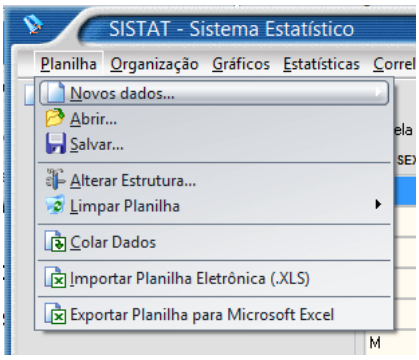
Figura 11: Correlação Linear / Regressão linear



Fonte: Autor

Todos os dados utilizados podem ser exportados para uma planilha do Microsoft Excel ou salvos em arquivo para posterior recuperação (Figura 12).

Figura 12: Salvamento e exportação de dados



Fonte: Autor

O layout de arquivos (txt) para exportação/importação no software “sistat” pode ser encontrado na integra no Apêndice 3.

5.2. UTILIZAÇÃO EM SALA DE AULA (Teste inicial do Produto Educacional)

Com o objetivo de se verificar a funcionalidade do nosso produto educacional, se é de fácil utilização, se possuía alguma inconsistência em seu código fonte, realizamos testes de utilização.

Participaram da pesquisa 26 alunos da disciplina de Estatística Básica, do curso de graduação em Ciências Contábeis da Universidade do Grande Rio. Os encontros foram fotografados e ocorreram em 3 quartas-feiras das 19:00hs às 21:10hs em um dos laboratórios do Campus Carioca Shopping, com um intervalo de 10 minutos para o coffe-break, entre 20:20hs e 20:30hs.

Os participantes foram acomodados em computadores individualmente. Nesta oportunidade, foi solicitado que explorassem o software por conta própria sem a utilização de material para consulta, intervenção do professor ou de outros colegas, tentando identificar empiricamente o seu funcionamento, por um período de 40 minutos. Após, foi-lhes fornecida uma tabela primitiva, contendo 80 ocorrências de valores referentes ao faturamento médio mensal de empresas tributadas sob o regime de apuração do Lucro Presumido:

Quadro 2: Tabela primitiva sugerida (Faturamento médio mensal)

45.985,00	25.698,00	25.111,00	27.563,00	14.563,00	52.639,00	98.136,00	86.325,00
169.302,00	56.985,00	26.589,00	21.410,00	15.478,00	26.596,00	27.129,00	9.267,00
25.326,00	44.875,00	58.745,00	60.253,00	65.293,00	41.025,00	368.842,00	44.268,00
36.202,00	25.448,00	368.521,00	85.365,00	45.698,00	76.325,00	25.685,00	17.362,00
95.298,00	14.885,00	55.636,00	74.298,00	33.698,00	54.823,00	63.024,00	135.262,00
17.452,00	36.298,00	43.527,00	13.985,00	26.521,00	168.302,00	18.992,00	86.238,00
55.025,00	25.655,00	19.429,00	22.681,00	362.178,00	54.226,00	25.496,00	74.156,00
26.514,00	30.256,00	58.965,00	146.025,00	39.267,00	32.105,00	30.924,00	43.620,00
36.851,00	118.306,00	61.485,00	84.526,00	29.653,00	29.126,00	46.208,00	31.897,00
26.587,00	98.574,00	84.215,00	30.806,00	70.268,00	91.839,00	29.993,00	256.385,00

Fonte: Dados fictícios

Os alunos receberam o manual do SISTAT (Apêndice A) e uma apresentação de como utilizá-lo, fazendo uma relação dos recursos do software com os conteúdos estatísticos.

Nos 80 minutos seguintes, os alunos teriam que verificar as medidas realizadas pelo software, bem como os gráficos e tabelas. Foram colhidas as informações a respeito do que foi produzido e as opiniões dos alunos (Apêndice B), e esclarecidas as dúvidas que surgiram durante esse contato. Solicitamos, ainda, que os alunos redigissem um pequeno relato (Apêndice B) sobre esse primeiro contato com o SISTAT.

Os alunos participantes possuíam conhecimento prévio dos seguintes tópicos da estatística descritiva: População e amostra; classificação de variáveis; construção de tabelas; séries estatísticas; distribuição de frequências com e sem intervalos de classes; representação gráfica (gráficos em linha, barras verticais/horizontais, setores, histograma, polígono de frequências e ogiva); medidas de posição (média, moda, mediana e separatrizes).

Os alunos demonstraram entusiasmo na utilização do nosso software educacional, principalmente quando foram informados que ele auxiliaria nos cálculos. Eles sinalizaram que os cálculos eram muito trabalhosos e pelo fato do software agilizar essa fase já os interessaria muito. A medida que foram conhecendo os demais recursos do software, o entusiasmo foi aumentando, principalmente ao observarem os cálculos das medidas estatística, a geração dos gráficos, etc. No Apêndice B encontram-se mais detalhes do teste inicial em sala de aula.

6. RESULTADOS DE APLICAÇÃO DO PRODUTO

Nossa pesquisa tem como objetivo principal desenvolver uma ferramenta computacional, como proposta para o processo de ensino-aprendizagem dos conteúdos de estatística descritiva, devido à escassez de recursos tecnológicos gratuitos.

Concluídas as fases de desenvolvimento e simulação, sentimo-nos preparados para testá-lo em sala de aula.

As atividades que serão descritas neste relatório foram realizadas no laboratório do campus Carioca Shopping da Universidade do Grande Rio, com alunos da disciplina de Estatística Básica do curso de Ciências Contábeis da Universidade do Grande Rio.

O objetivo deste relatório é descrever a aplicação de nosso produto em sala de aula (laboratório) e sua avaliação por parte dos alunos.

6.1. METODOLOGIA E APLICAÇÃO

Nossos testes foram realizados em 2 dias, com o mesmo grupo de alunos.

Os alunos inicialmente receberam aulas regulares em sala de aula, através da utilização de quadro negro e datashow, quando foram abordados os seguintes conteúdos da estatística descritiva: população e amostra; classificação de variáveis; construção de tabelas; séries estatísticas; distribuição de frequências com e sem intervalos de classes; representação gráfica (Gráficos em linha, barras verticais/horizontais, setores, Histograma, Polígono de Frequências e Ogiva);

No dia 11 de março de 2015, agendamos o laboratório do Campus Vicente de Carvalho, das 19:00hs às 21:10hs.. Participaram 26 alunos da disciplina de Estatística Básica do Curso de Ciências Contábeis.

Nos primeiros 40 minutos, os alunos foram orientados a abrir o SISTAT e explorá-lo à vontade, podendo navegar entre os menus, botões e telas livremente. Nesse primeiro momento, foi solicitado que reconhecessem no software os conceitos vistos em sala de aula, podendo a qualquer momento tirar dúvidas com o professor instrutor.

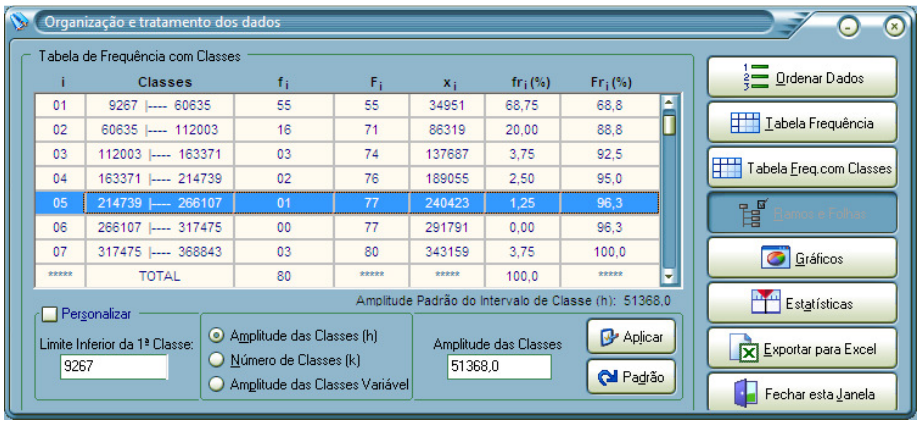
Nos 90 minutos seguintes, os alunos receberam uma tabela primitiva com 80 ocorrências, referentes ao faturamento médio mensal de 80 empresas. A partir desta base de dados, os alunos deveriam colocar os dados em Rol, construir Distribuições

de Frequências com e sem intervalos de classe, verificar se a Regra de Sturges foi utilizada, Calcular a Média, a moda e a Mediana, gerar 2 gráficos de livre escolha.

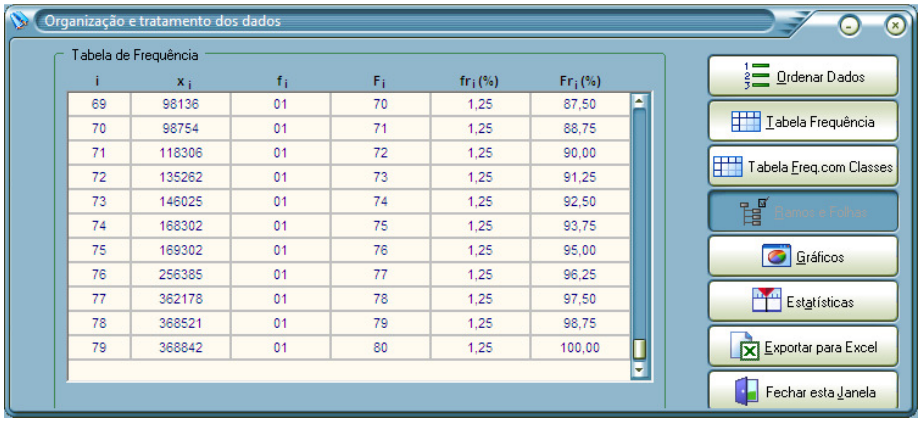
Embora o software possua recurso de importação de dados a partir de uma base gerada anteriormente, os próprios alunos parametrizaram as variáveis a serem utilizadas e digitaram manualmente os dados da tabela primitiva fornecida. Muitas dúvidas surgiram, no entanto, percebeu-se que alguns alunos instintivamente ou observando o aluno ao lado prosseguiu utilizando o software sem apresentar dúvidas.

Nas páginas seguintes seguem os resultados apresentados por 3 alunos escolhidos aleatoriamente, para preservar a identidades desses alunos, chamamos a seguir de ALUNO A, ALUNO B e ALUNO C.

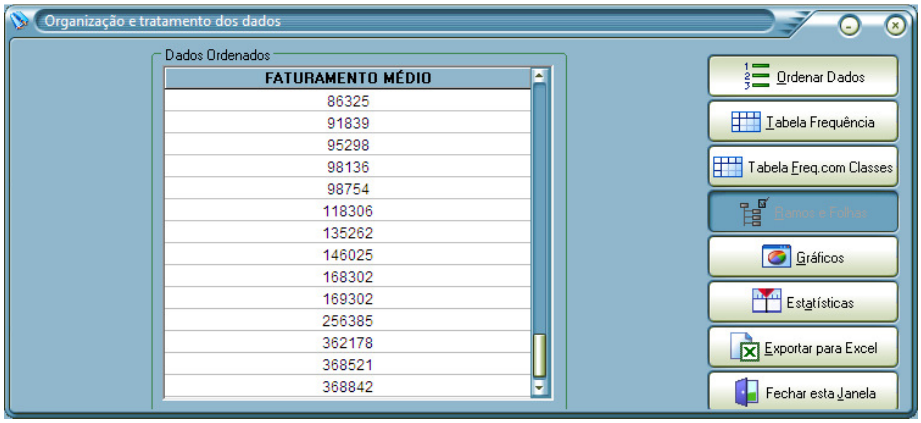
ALUNO A

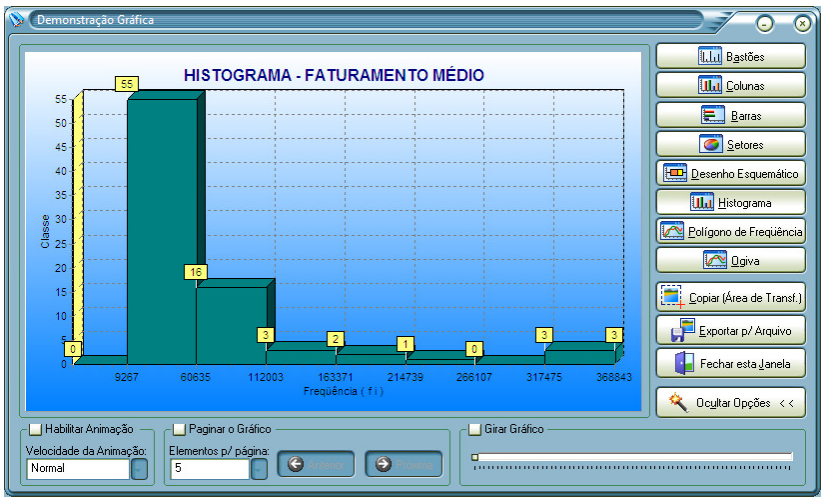


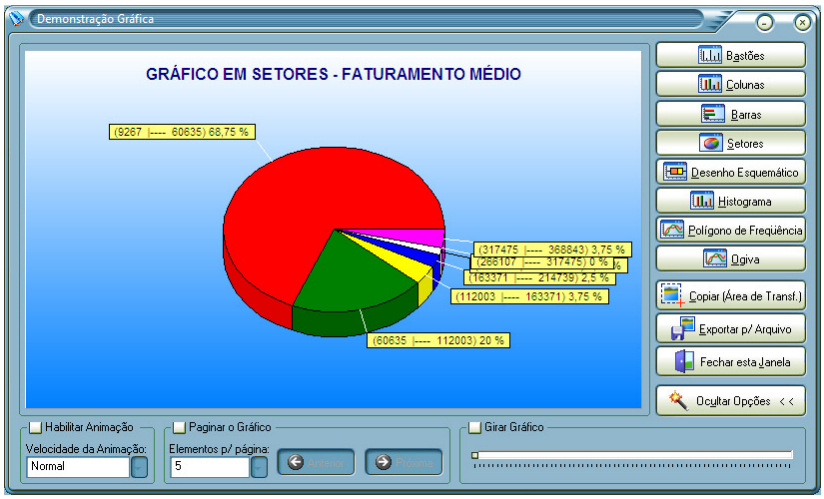
i	Classes	f_i	F_i	x_i	fr_i (%)	Fr_i (%)
01	9267 |---- 60635	55	55	34951	68,75	68,8
02	60635 |---- 112003	16	71	86319	20,00	88,8
03	112003 |---- 163371	03	74	137687	3,75	92,5
04	163371 |---- 214739	02	76	189055	2,50	95,0
05	214739 |---- 266107	01	77	240423	1,25	96,3
06	266107 |---- 317475	00	77	291791	0,00	96,3
07	317475 |---- 368843	03	80	343159	3,75	100,0
*****	TOTAL	80	*****	*****	100,0	*****

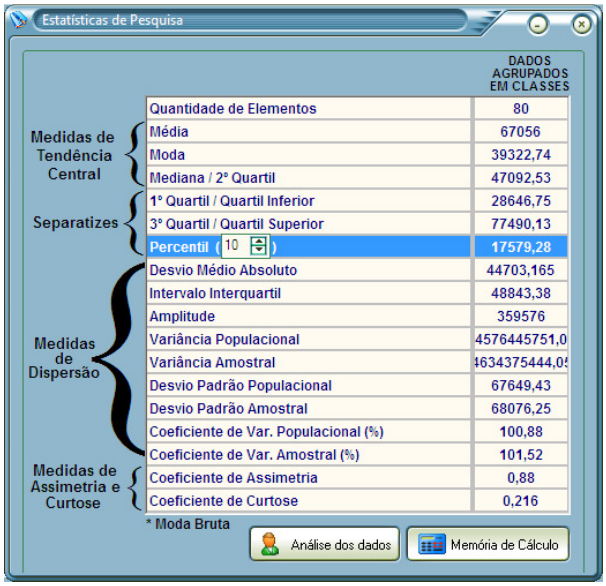


i	x_i	f_i	F_i	fr_i (%)	Fr_i (%)
69	98136	01	70	1,25	87,50
70	98754	01	71	1,25	88,75
71	118306	01	72	1,25	90,00
72	135262	01	73	1,25	91,25
73	146025	01	74	1,25	92,50
74	168302	01	75	1,25	93,75
75	169302	01	76	1,25	95,00
76	256385	01	77	1,25	96,25
77	362178	01	78	1,25	97,50
78	368521	01	79	1,25	98,75
79	368842	01	80	1,25	100,00


Organização e tratamento dos dados
Dados Ordenados
FATURAMENTO MÉDIO
86325
91839
95298
98136
98754
118306
135262
146025
168302
169302
256385
362178
368521
368842
Ordenar Dados
Tabela Frequência
Tabela Freq.com Classes
Ramos e Folhas
Gráficos
Estatísticas
Exportar para Excel
Fechar esta Janela


Demonstração Gráfica
HISTOGRAMA - FATURAMENTO MÉDIO
Classe
Freqüência (f i)
9267
60635
112003
163371
214739
266107
317475
368843
Bastões
Colunas
Barras
Setores
Desenho Esquemático
Histograma
Polígono de Freqüência
Ogiva
Copiar (Área de Transf.)
Exportar p/ Arquivo
Fechar esta Janela
Ocultar Opções
Habilitar Animação
Velocidade da Animação:
Normal
Paginar o Gráfico
Elementos p/ página:
5
Girar Gráfico


Demonstração Gráfica
GRÁFICO EM SETORES - FATURAMENTO MÉDIO
(9267 |---- 60635) 68,75 %
(317475 |---- 368843) 3,75 %
(266107 |---- 317475) 0 %
(163371 |---- 214739) 2,5 %
(112003 |---- 163371) 3,75 %
(60635 |---- 112003) 20 %
Bastões
Colunas
Barras
Setores
Desenho Esquemático
Histograma
Polígono de Freqüência
Ogiva
Copiar (Área de Transf.)
Exportar p/ Arquivo
Fechar esta Janela
Ocultar Opções
Habilitar Animação
Velocidade da Animação:
Normal
Paginar o Gráfico
Elementos p/ página:
5
Girar Gráfico



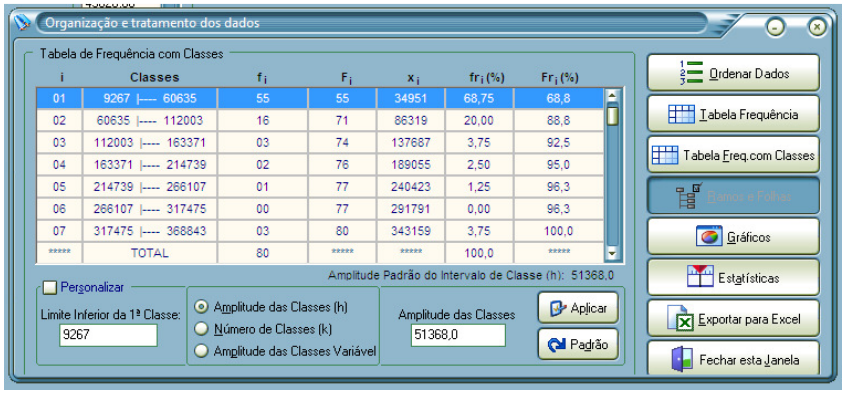
Estatísticas de Pesquisa

		DADOS AGRUPADOS EM CLASSES
	Quantidade de Elementos	80
Medidas de Tendência Central	Média	67056
	Moda	39322,74
	Mediana / 2º Quartil	47092,53
Separatizes	1º Quartil / Quartil Inferior	28646,75
	3º Quartil / Quartil Superior	77490,13
	Percentil (10)	17579,28
Medidas de Dispersão	Desvio Médio Absoluto	44703,165
	Intervalo Interquartil	48843,38
	Amplitude	359576
	Variância Populacional	4576445751,0
	Variância Amostral	4634375444,0
	Desvio Padrão Populacional	67649,43
	Desvio Padrão Amostral	68076,25
	Coeficiente de Var. Populacional (%)	100,88
	Coeficiente de Var. Amostral (%)	101,52
Medidas de Assimetria e Curtose	Coeficiente de Assimetria	0,88
	Coeficiente de Curtose	0,216

* Moda Bruta

Análise dos dados | Memória de Cálculo

ALUNO B



Organização e tratamento dos dados

Tabela de Frequência com Classes

i	Classes	f_i	F_i	x_i	fr_i (%)	Fr_i (%)
01	9267 |---- 60635	55	55	34951	68,75	68,8
02	60635 |---- 112003	16	71	86319	20,00	88,8
03	112003 |---- 163371	03	74	137687	3,75	92,5
04	163371 |---- 214739	02	76	189055	2,50	95,0
05	214739 |---- 266107	01	77	240423	1,25	96,3
06	266107 |---- 317475	00	77	291791	0,00	96,3
07	317475 |---- 368843	03	80	343159	3,75	100,0
*****	TOTAL	80	*****	*****	100,0	*****

Amplitude Padrão do Intervalo de Classe (h): 51368,0

Personalizar

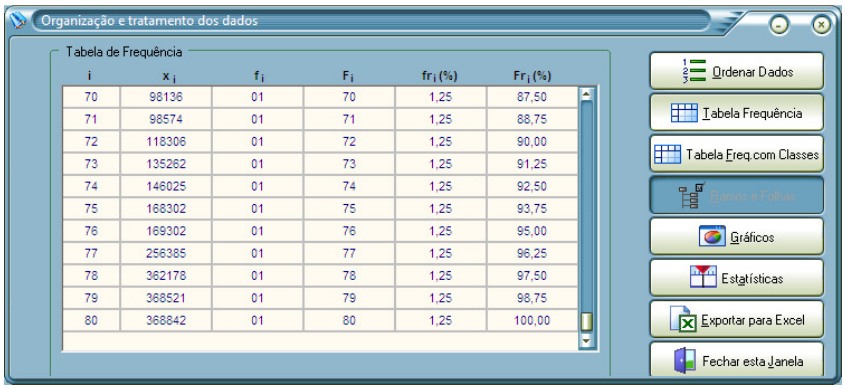
Limite Inferior da 1ª Classe: 9267

Amplitude das Classes (h) | Número de Classes (k) | Amplitude das Classes Variável

Amplitude das Classes 51368,0

Aplicar | Padrão

Ordenar Dados | Tabela Frequência | Tabela Freq.com Classes | Ramos e Folhas | Gráficos | Estatísticas | Exportar para Excel | Fechar esta Janela

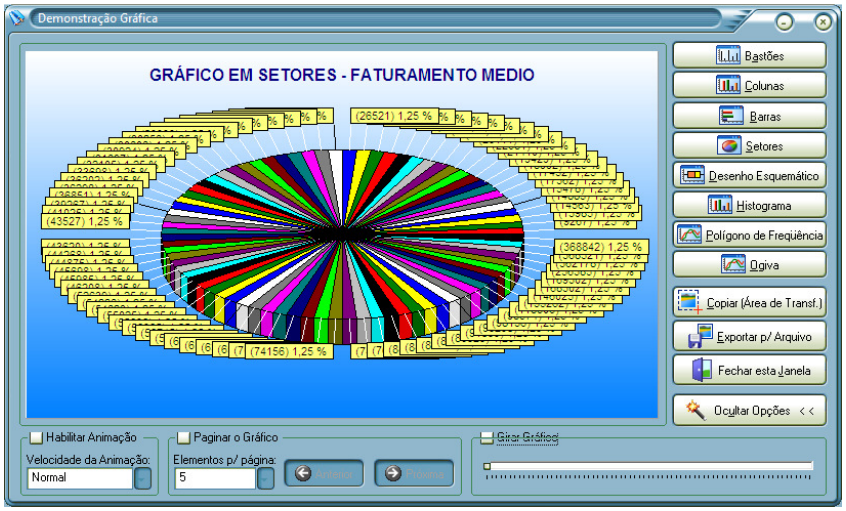


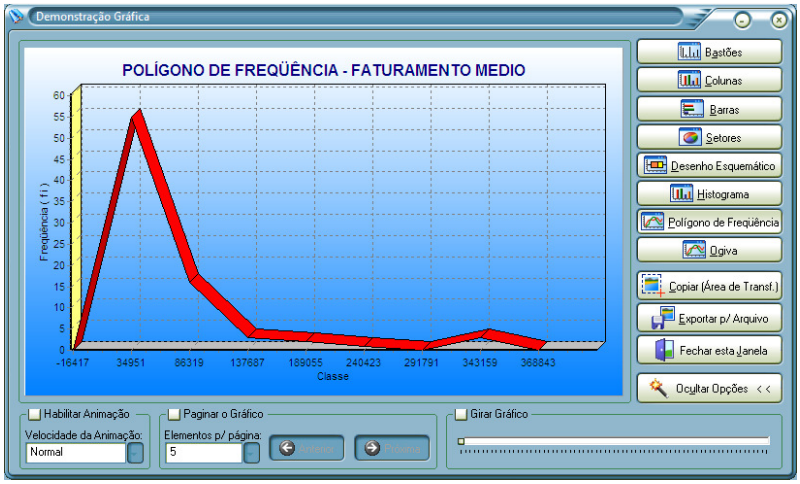
Organização e tratamento dos dados

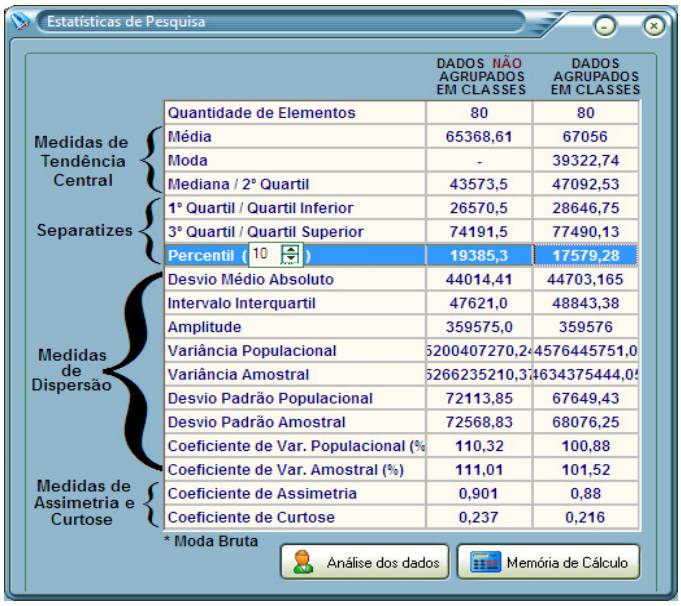
Tabela de Frequência

i	x_i	f_i	F_i	fr_i (%)	Fr_i (%)
70	98136	01	70	1,25	87,50
71	98574	01	71	1,25	88,75
72	118306	01	72	1,25	90,00
73	135262	01	73	1,25	91,25
74	146025	01	74	1,25	92,50
75	168302	01	75	1,25	93,75
76	169302	01	76	1,25	95,00
77	256385	01	77	1,25	96,25
78	362178	01	78	1,25	97,50
79	368521	01	79	1,25	98,75
80	368842	01	80	1,25	100,00

Ordenar Dados | Tabela Frequência | Tabela Freq.com Classes | Ramos e Folhas | Gráficos | Estatísticas | Exportar para Excel | Fechar esta Janela



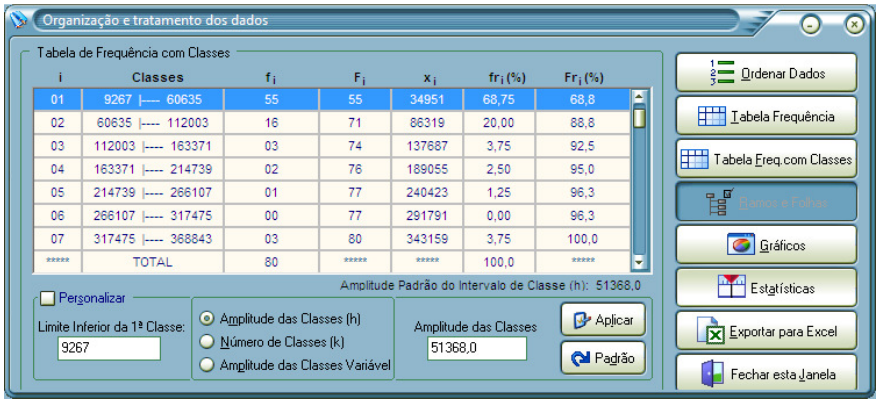




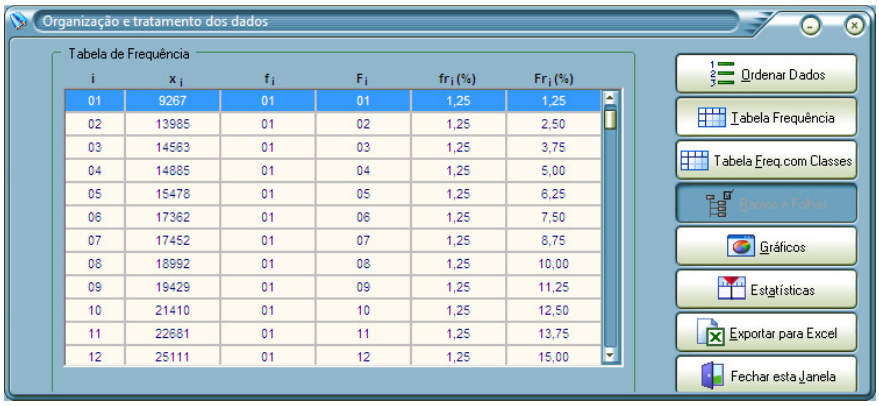
		DADOS NÃO AGRUPADOS EM CLASSES	DADOS AGRUPADOS EM CLASSES
	Quantidade de Elementos	80	80
Medidas de Tendência Central	Média	65368,61	67056
	Moda	-	39322,74
	Mediana / 2º Quartil	43573,5	47092,53
Separatizes	1º Quartil / Quartil Inferior	26570,5	28646,75
	3º Quartil / Quartil Superior	74191,5	77490,13
	Percentil (10)	19385,3	17579,28
Medidas de Dispersão	Desvio Médio Absoluto	44014,41	44703,165
	Intervalo Interquartil	47621,0	48843,38
	Amplitude	359575,0	359576
	Variância Populacional	5200407270,2	4576445751,0
	Variância Amostral	5266235210,3	4634375444,05
	Desvio Padrão Populacional	72113,85	67649,43
	Desvio Padrão Amostral	72568,83	68076,25
	Coeficiente de Var. Populacional (%	110,32	100,88
	Coeficiente de Var. Amostral (%)	111,01	101,52
Medidas de Assimetria e Curtose	Coeficiente de Assimetria	0,901	0,88
	Coeficiente de Curtose	0,237	0,216

* Moda Bruta

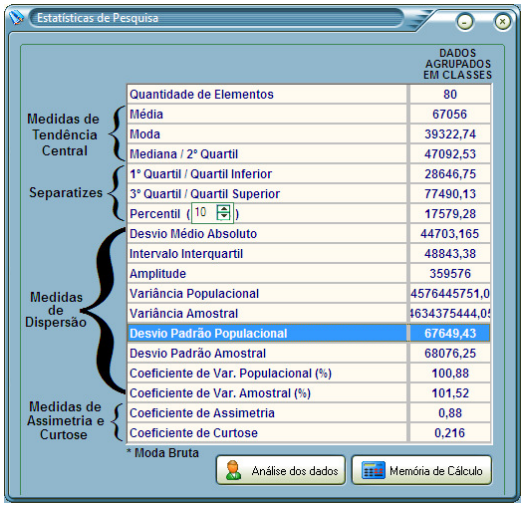
ALUNO C



i	Classes	f_i	F_i	x_i	fr_i (%)	Fr_i (%)
01	9267 |---- 60635	55	55	34951	68,75	68,8
02	60635 |---- 112003	16	71	86319	20,00	88,8
03	112003 |---- 163371	03	74	137687	3,75	92,5
04	163371 |---- 214739	02	76	189055	2,50	95,0
05	214739 |---- 266107	01	77	240423	1,25	96,3
06	266107 |---- 317475	00	77	291791	0,00	96,3
07	317475 |---- 368843	03	80	343159	3,75	100,0
*****	TOTAL	80	*****	*****	100,0	*****

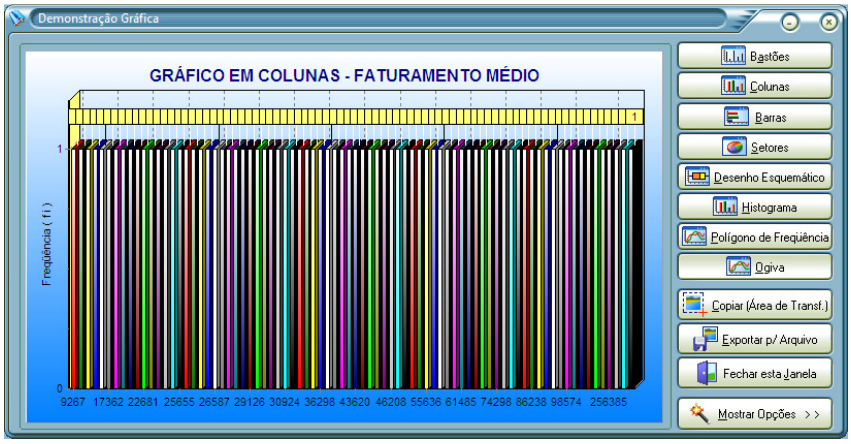


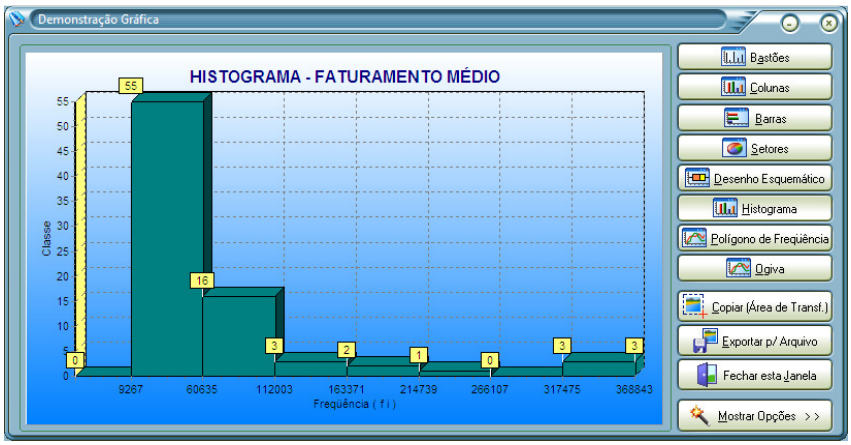
i	x_i	f_i	F_i	fr_i (%)	Fr_i (%)
01	9267	01	01	1,25	1,25
02	13985	01	02	1,25	2,50
03	14563	01	03	1,25	3,75
04	14885	01	04	1,25	5,00
05	15478	01	05	1,25	6,25
06	17362	01	06	1,25	7,50
07	17452	01	07	1,25	8,75
08	18992	01	08	1,25	10,00
09	19429	01	09	1,25	11,25
10	21410	01	10	1,25	12,50
11	22681	01	11	1,25	13,75
12	25111	01	12	1,25	15,00



		DADOS AGRUPADOS EM CLASSES
	Quantidade de Elementos	80
Medidas de Tendência Central	Média	67056
	Moda	39322,74
	Mediana / 2° Quartil	47092,53
Separatizes	1° Quartil / Quartil Inferior	28646,75
	3° Quartil / Quartil Superior	77490,13
	Percentil (10)	17579,28
Medidas de Dispersão	Desvio Médio Absoluto	44703,165
	Intervalo Interquartil	48843,38
	Amplitude	359576
	Variância Populacional	4576445751,0
	Variância Amostral	1634375444,0
	Desvio Padrão Populacional	67649,43
	Desvio Padrão Amostral	68076,25
	Coeficiente de Var. Populacional (%)	100,88
	Coeficiente de Var. Amostral (%)	101,52
Medidas de Assimetria e Curtose	Coeficiente de Assimetria	0,88
	Coeficiente de Curtose	0,216

* Moda Bruta





Ao final, foi solicitado que os alunos formulassem um depoimento a respeito da utilização do software, como por exemplo, pontos positivos, negativos e sugestões. Transcrevemos abaixo 3 depoimentos escolhidos aleatoriamente:

ALUNO A:

Não acho que existam pontos negativos, pode ser que não haja computadores para todos se a turma fosse grande, o que não é o caso.

Existem mais pontos positivos, porque as tabelas sendo vistas são de mais fácil assimilação e também o fato de manuseá-las e não ficar só no papel. É possível em sala de aula sim, seria como uma aula online, sendo com a explicação imediata do professor, quando houver as dúvidas.

No caso seria em laboratório de informática.

Aquelas tabelas da aula anterior foram muito boas, como já saiam todos os resultados; foi bem prática. Como as explicações e os exercícios já tinham sido dados anteriormente foi bem mais fácil usar as ferramentas.

ALUNO B:

Sistat, gostei muito do sistema, não tenho como apontar um PONTO NEGATIVO, pois achei o sistema muito interessante, fácil e prático de mexer. Como tenho que apontar um ponto negativo, na verdade é mais uma opinião: o programa podia dar uma melhorada na sua barra de ferramenta, para que o aluno visualize melhor, e fazer alguns ajustes na parte gráfica. PONTOS POSITIVOS: Como dito acima, lidar com o programa é simplesmente fácil e rápido, sabendo que para isso é preciso conhecer sobre o assunto e ter praticado em sala de aula conceitos básicos de estatística. Também é interessante a forma que o programa apresenta as fórmulas estatística utilizadas para chegar ao resultado. USO EM SALA DE AULA, Sim é possível! Mas não descartando que, antes disso, o aluno pratique bastante e estude as fórmulas, faça bastante exercícios para treinar a parte teórica e dominar o assunto, sendo assim, será muito mais fácil usá-lo em sala de aula. Fazer o uso do programa em sala de aula seria bastante dinâmico, pelo fato de a tecnologia estar tão avançada. Pois somente a parte teórica não ajudaria tanto ao trabalhar com dados estatísticos em um mercado de trabalho!

ALUNO C:

Pontos Negativos: Visualização dos gráficos, na teoria foi muito melhor a visualização, já no sistema complicou um pouco.

Pontos Positivos: Facilita muito quanto aos resultados. Não esquecendo da parte teórica pra poder entender os resultados.

Uso em sala de aula: Seria bom e dinâmico

É possível? Sim

Como? Aulas no laboratório, assim como foi apresentado a turma.

Nas aulas em sala de aula posteriores ao 1º encontro, os alunos receberam aulas regulares abordando os seguintes conteúdos:

- Medidas de Posição (Média, Moda, Mediana e Separatrizes);

- Medidas de Dispersão e Assimetria (Desvio Padrão, Variância, Coeficiente de Variação e Assimetria);
- Correlação e Regressão.

No segundo encontro, realizado no dia 03 de junho de 2015, os alunos já estavam familiarizados com os conteúdos da ementa da disciplina de estatística básica. Reservamos o laboratório do Campus Carioca Shopping da Universidade do Grande Rio pelo mesmo intervalo de tempo do encontro anterior, das 19:00hs às 21:10hs. Os alunos receberam uma tabela contendo informações fictícias de uma eventual pesquisa realizada, como segue na página seguinte:

Pesquisa realizada com 30 alunos de uma Instituição de ensino superior

SEXO*	CURSO	PERÍODO	DESEMPENHO	SATISFAÇÃO
1	ADMINISTRAÇÃO	1	7,50	89
0	CONTABILIDADE	3	6,00	57
1	CONTABILIDADE	5	9,00	95
1	ECONOMIA	4	9,80	98
0	ECONOMIA	2	8,30	77
1	CONTABILIDADE	1	7,60	79
0	ECONOMIA	2	8,20	85
1	ESTATÍSTICA	5	7,40	80
1	CONTABILIDADE	4	8,60	89
0	ADMINISTRAÇÃO	6	9,40	90
1	MEDICINA	2	8,60	90
0	MEDICINA	7	7,50	85
0	ADMINISTRAÇÃO	5	8,10	78
0	CONTABILIDADE	4	9,00	80
1	ECONOMIA	1	8,60	95
0	ESTATÍSTICA	2	7,50	80
1	CONTABILIDADE	8	7,10	60
0	ESTATÍSTICA	6	8,60	90
1	ADMINISTRAÇÃO	3	4,50	60
0	ESTATÍSTICA	2	7,60	88
1	CONTABILIDADE	5	8,50	95
0	ADMINISTRAÇÃO	8	8,40	93
0	DIREITO	7	9,60	95
0	DIREITO	4	8,50	90
1	ESTATÍSTICA	5	7,20	80
0	ADMINISTRAÇÃO	2	9,60	90

1	ESTATÍSTICA	1	8,50	80
0	ECONOMIA	5	7,40	88
0	ADMINISTRAÇÃO	6	8,20	70
1	CONTABILIDADE	3	8,90	85

Dados fictícios

Inicialmente, foi solicitado que classificassem as variáveis, configurassem e alimentassem o software com esses dados. Posteriormente, os alunos deveriam verificar as medidas encontradas e responder o questionário abaixo utilizando o software como ferramenta de apoio:

1 - É possível afirmar que existe relação entre o grau de satisfação do aluno e seu desempenho (nota)? Por que?

2 - Em qual conceito estatístico você se basearia para responder a pergunta anterior?

3 - Informe a participação percentual de cada curso em relação ao total da amostra.

4 - Determine o valor da assimetria da variável desempenho. Em sua opinião, por que a assimetria está tão próxima de zero?

Nas páginas seguintes verificamos os resultados apresentados por 3 alunos selecionados aleatoriamente:

ALUNO A:

QUESTÕES (BASEADAS NA TABELA DO ANVERSO DESTA FOLHA)

1 - É possível afirmar que existe relação entre o grau de satisfação do aluno e seu desempenho (nota)? Por que?

Sim. Pois através do diagramas de dispersão podemos observar que no geral, quanto maior a nota dos alunos, maior seu grau de satisfação.

2 - Em qual conceito estatístico você se basearia para responder a pergunta anterior?

Através do conceito de correlação

3 - Informe a participação percentual de cada curso em relação ao total da amostra.

Direito - 6,7%
Contabilidade - 26,67%
Administração - 23,33%
Economia - 16,67%
Estatística - 20%
Medicina - 6,67%

4 - Determine a assimetria. Em sua opinião, por que a assimetria está tão próxima de zero?

O coeficiente de assimetria é 0,315 (moda = 8,5; mediana = 8,43; média = 8,4) Isso se dá pois os dados coletados que foram apresentados possuem baixa dispersão. São valores muito próximos.

ALUNO B:

QUESTÕES (BASEADAS NA TABELA DO ANVERSO DESTA FOLHA)

1 - É possível afirmar que existe relação entre o grau de satisfação do aluno e seu desempenho (nota)? Por que?

Sim, de acordo com o gráfico quanto maior o desempenho, maior é a satisfação do aluno

2 - Em qual conceito estatístico você se basearia para responder a pergunta anterior?

Coeficiente de correlação linear

3 - Informe a participação percentual de cada curso em relação ao total da amostra.

Administração - 23,33%; Contabilidade - 26,67%; Direito - 6,67%
Economia - 16,67%; Estatística - 20%; Medicina - 6,67%

4 - Determine a assimetria. Em sua opinião, por que a assimetria está tão próxima de zero?

Assimetria para a direita, Está próximo de zero porque a média, moda e mediana tem valores muito próximos

ALUNO C:

QUESTÕES (BASEADAS NA TABELA DO ANVERSO DESTA FOLHA)

1 - É possível afirmar que existe relação entre o grau de satisfação do aluno e seu desempenho (nota)? Por que?

Sim, pois o valor 0,7506 do coeficiente de correlação está mais próximo de 1. Pois quanto maior o desempenho (maior a satisfação).

2 - Em qual conceito estatístico você se basearia para responder a pergunta anterior?

Regressão linear e Correlação

3 - Informe a participação percentual de cada curso em relação ao total da amostra.

(Administração) 23,33 / Contabilidade 26,67 / Direito 6,67 / Economia 16,67 / Estatística 20,00 Medicina 6,67

4 - Determine a assimetria. Em sua opinião, por que a assimetria está tão próxima de zero?

AS = -0,315 porque a moda, a média e a mediana estão próximas entre si.

FOTOS REGISTRADAS DURANTE DA APLICAÇÃO DO SOFTWARE

6.2. CONCLUSÃO DA APLICAÇÃO DO PRODUTO

A instalação do software nos computadores foi realizada pela equipe de TI do laboratório, que nos reportou não ter encontrado dificuldades.

Encontramos alguns problemas relacionados a restrições de segurança em alguns computadores do laboratório, mas não interferiram na utilização do software. Já estão sendo tratadas.

Observamos uma grande aceitação por parte dos alunos, inclusive alguns demonstraram entusiasmo, no sentido de colaborar com a nossa pesquisa, como por exemplo, buscando falha nos cálculos.

Por ser uma atividade bem diferente do cotidiano em sala de aula, os alunos foram mais participativos e colaborativos. Notamos que os alunos se mostraram focados no uso e na testagem da ferramenta.

Registramos uma aluna com um pouco de dificuldade, devido a pouca experiência no uso de computadores.

De modo geral, entendemos como positiva essa experiência realizada. Constatamos que se trata de um projeto promissor, no qual melhorias serão implementadas.

7. CONSIDERAÇÕES FINAIS

A presente pesquisa nos proporcionou a oportunidade de conhecer, através de nossos referenciais teóricos, como as tecnologias de informação podem contribuir com professores e alunos na construção do conhecimento. Apontando ainda que as ferramentas computacionais aplicadas ao processo de ensino-aprendizagem podem contribuir enormemente, servindo como suporte, não substituindo, contudo, a figura do professor.

Segundo Moraes (1997), precisamos ainda partir do princípio que o professor sinta-se motivado a inovar e criar situações inovadoras e desafiadoras, utilizando os diversos recursos didáticos disponíveis, inclusive as *ferramentas computacionais*, proporcionando um ambiente cheio de inovações e novas experiências.

Constatamos que há uma carência de recursos tecnológicos e/ou metodologias de ensino-aprendizagem alternativo-complementares para a disciplina de Estatística.

E, de acordo com nossas pesquisas e referenciais teóricos, dando continuidade às pesquisas já existentes, isto é, objetivando uma linha de pesquisa progressiva, pretendemos acrescentar novas funcionalidades ao software, norteados por resultados obtidos através de aplicações do produto em sala de aula.

Entendemos que os objetivos propostos foram atingidos, assim como nosso questionamento apresentado, pois foi possível desenvolver uma ferramenta computacional de uso livre direcionada para o ensino da estatística descritiva.

Sugerimos que o professor, ao utilizar o *Sistat*, conheça bem suas funcionalidades, definindo os objetivos de suas aulas de forma a provocar situações que despertem o interesse dos alunos, pois vão perceber que o software dará respostas imediatas, ou seja, o computador fará os cálculos, plotagens e análises. Portanto, somente o uso do software, sem acompanhamento e/ou estratégias bem definidas, não garante a aprendizagem.

REFERÊNCIAS

BATANERO, C. **La simulación como instrumento de modernización em probabilidad**. Educación y Pedagogia, Págs 35, 37-64.

BEZERRA, Eduardo. **Princípios de Análise e Projeto de Sistemas com UML**. Rio de Janeiro: Elsevier, 2015.

BORBA, M. C.,PENTEADO, M. G. **Informática e Educação Matemática**. Belo Horizonte: Autêntica, 2005.

BRASIL. Ministério da Educação. **Parâmetros Curriculares Nacionais**, Brasília: 1999.

BRASIL (1997), Secretaria de Educação Fundamental. **Parâmetros Curriculares Nacionais: Matemática – Ensino1ª a 4ª série e 5ª a 8ª série**. Brasília: MEC/SEF

CAMPOS, Rui J. A. **Cálculo Numérico Básico**. São Paulo: Atlas, 1978.

CANTU, M. **Dominando Delphi 7: a bíblia**. São Paulo: Makron Books, 2003.

CAZORLA, I. M. (2006).**Teaching statistics in Brazil**. In Rossman, A.; Chance, B. (Eds.), Proc. Seventh Intern. Conf. Teaching Statistics. Intern. Ass. for Stat. Education, Salvador (Brazil). CD ROM., 2006.

D'AMBRÓSIO, U. **Etnomatemática: Arte ou Técnica de Explicar ou Conhecer**. São Paulo, Editora Ática, 1990.

DEBASTIANI, Carlos Alberto. **Definindo Escopo em Projetos de Software**. São Paulo: Novatec, 2015.

FONSECA, J. S, MARTINS, G. Curso de Estatística. São Paulo, Atlas. 1996.

GAL, Iddo. **Adult's Statistical Literacy: meanings, components, responsibilities**. International Statistical Review, 70(1), 1-25, 2002.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. 4. São Paulo: Atlas, 2008.

GODINO, J. D., **¿Qué aportan los ordenadores al aprendizaje y La enseñanza de la estadística?** UNO: Revista didáctica de las matemáticas, 5, 1995 p. 45-56

HOEL, Paul G. **Estatística elementar**. Tradução: Carlos Roberto Vieira Araújo. São Paulo: Atlas, 1979.

JOLLIFFE, F. **The changing brave new world of statitics assessment**. The Netherlands, Voorburg. International Stastitical Institute, CD-ROM, 2007.

LOPES, Celi Aparecida Espasadin. **A probabilidade e a Estatística no Ensino Fundamental: uma análise curricular.** 1998. 125.p. Dissertação (Mestrado em Educação) Faculdade de Educação da UNICAMP: 1998. Disponível em <http://www.ime.unicamp.br/~lem/publica/ce_lopes/est_prop.pdf>. Acesso em: 14/09/2013.

MORAES, M. C. **Informática Educativa no Brasil: Um pouco de história.** Em Aberto. Brasília, 1993.

MORAES, M. C. **Informática Educativa no Brasil: Uma história vivida, algumas lições aprendidas**. Revista Brasileira de Informática na Educação. (UFSC), Setembro/1997.

NATO Science Commitee, Garmish, Germany. Friedrich Ludwig BAUER. 7-11 Oct. 1968, Brussels, Scientific Affair Division, NATO (1969) p,231.

PEA-MAT – PUC-SP. <Disponível em: http://www.pucsp.br/pensamentomatematico/index.html>

PRESSMAN, ROGER S. **Engenharia de Software**. 5ª Edição. Rio de Janeiro: McGraw-Hill, 2002.

QUARTIERO, E. M. **Da máquina de ensinar à máquina de aprender: pesquisas em tecnologia educacional**. 2007. Disponível em <http://intranet.ufsj.edu.br/rep_sysweb/File/vertentes/Vertentes_29/elisa_quartiero.pdf>Acesso emNov/2015.

ROYCE, Winston W. **Managing the developement of large software systems.**1970.

TOLEDO,G. L. &Ovalle, I.Z. **Estatística Básica**. 2ª Edição. São Paulo, 1985.

TRIOLA, M.F. **Introdução à Estatística**. LTC. 2008.

APÊNDICE A

MANUAL DE UTILIZAÇÃO DO SISTAT

ENTRADA DE DADOS

A entrada de dados pode ser feita de várias formas, podemos:

- Digitar os dados coletados a partir da tabela primitiva[1];
- Copiar e colar grupos de dados a partir planilhas eletrônicas (Microsoft Excel, Libreoffice, etc...);
- Digitar informações diretamente em um tabela de Distribuição de Frequências;
- Digitar informações diretamente em um tabela de Distribuição de Frequências com Intervalos de Classe;
- Importar informações digitadas anteriormente e gravadas em arquivo digital.

INFORMANDO A TABELA PRIMITIVA

Para informar a tabela primitiva, no menu superior clique em PLANILHA → NOVOS DADOS → DADOS BRUTOS.

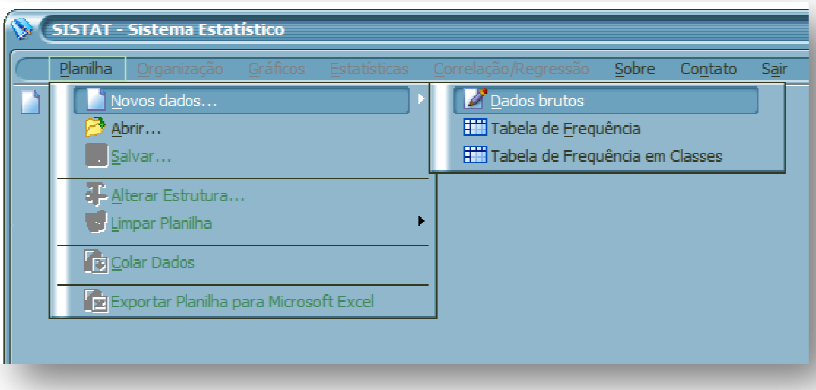


Figura 1 - Entrada de dados - Tabela Primitiva

Será apresentada uma nova janela onde definiremos o número de variáveis da pesquisa e os tipos de dados que receberão.

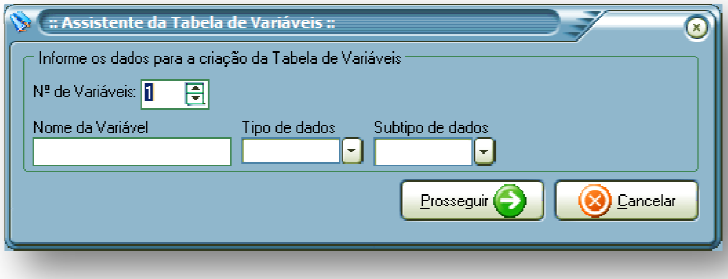


Figura 2 - Entrada de dados - Tabela Primitiva

ANALISANDO APENAS 1 VARIÁVEL

No campo "Nº de Variáveis" selecione "1", para entrar com dados de apenas 1 variável. Informe o nome da variável (exemplo: Idade, Peso, Time de futebol) e o respectivo Tipo de dados[2].

O tipo de dados que a variável irá receber será informado de acordo com sua classificação: Qualitativa[2.1] ou Quantitativa[2.2]. Bem como seu subtipo, Sendo Qualitativa (Nominal ou Ordinal) ou Quantitativa (Discreta ou Contínua).

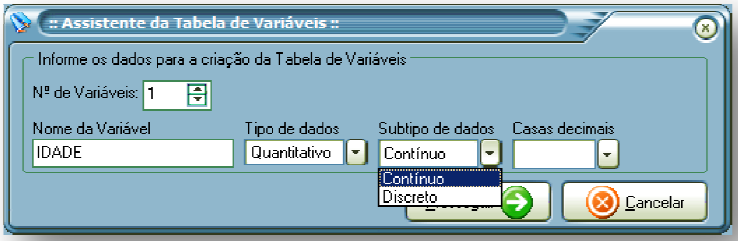


Figura 3 - Entrada de dados - Tabela Primitiva

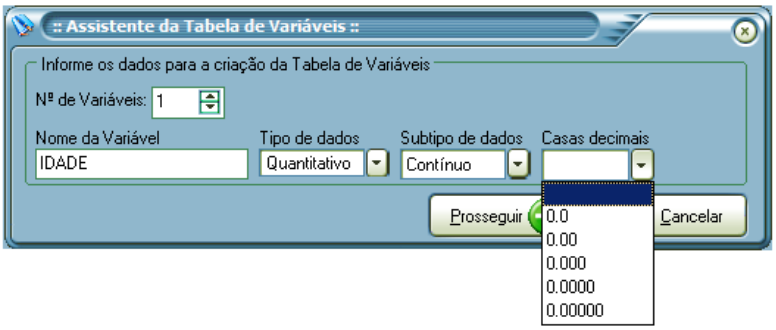


Figura 4 - Entrada de dados - Tabela Primitiva

Na hipótese de trabalharmos com uma variável Quantitativa Contínua, podemos informar o número de casas decimais para que o sistema padronize a entrada de dados.

Clique em PROSSEGUIR, na tela principal será apresentada a tabela para edição dos dados da pesquisa.

DEFININDO O NÚMERO DE OCORRÊNCIAS (n)

No canto inferior esquerdo da Dasboard[3] definimos o nº de ocorrências (Σx_i) que nossa pesquisa possui.

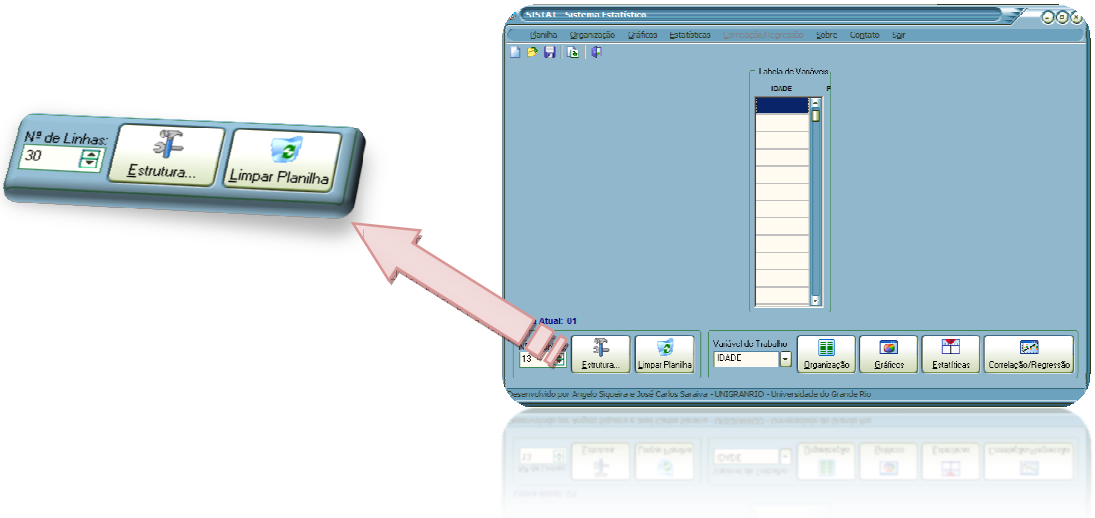


Figura 5 - Entrada de dados - Tabela Primitiva

ENTRADA DE DADOS - DIGITANDO A TABELA PRIMITIVA

A entrada de dados é semelhante às planilhas eletrônicas que já conhecemos (Microsoft Excel, LibreOffice, etc...), onde cada ocorrência será lançada em uma célula, em qualquer ordem que os dados se encontrem. Para mudar o foco para a próxima célula, utilize a tecla ENTER ou SETA PARA BAIXO (↓).

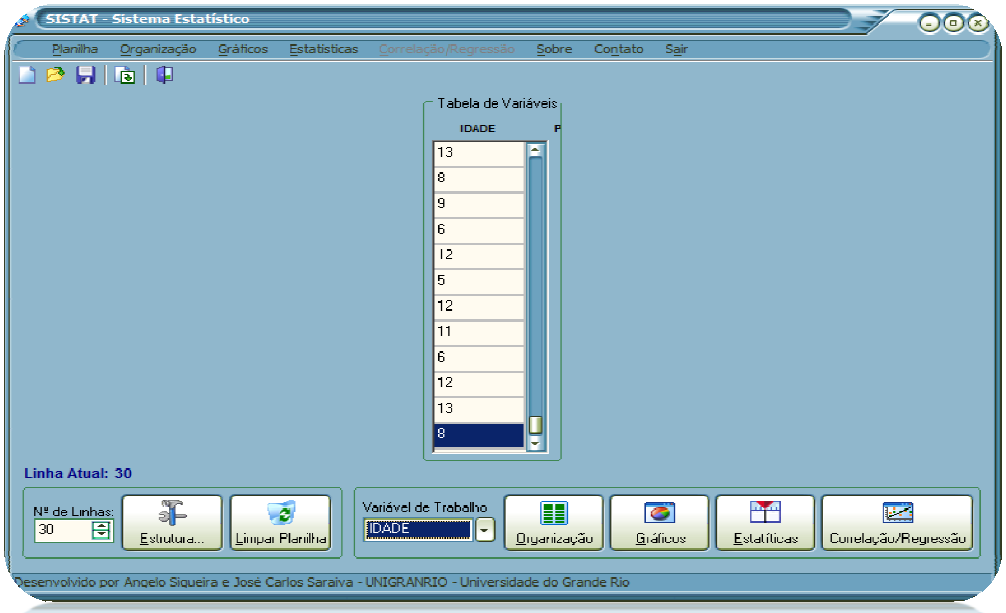


Figura 6 - Entrada de dados - Tabela Primitiva

COLANDO DADOS EXTERNOS

Caso os dados já estejam digitados em uma planilha eletrônica (Microsoft Excel, LibreOffice, etc...), podemos utilizar a área de transferência do Windows para transportar essas informações para a tabela primitiva, poupando-nos do retrabalho.

Basta informar o número de ocorrências que será informada (Tópico 1.1.2), selecionar as células que serão copiadas na planilha eletrônica, clicar no botão copiar ou (CTRL + C), dentro do Sistat selecionar a primeira célula que receberá e no menu superior clicar em PLANILHA → COLAR DADOS.

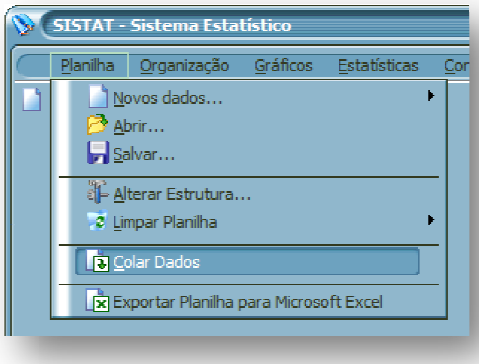


Figura 7 - Entrada de dados - Tabela Primitiva

INFORMANDO A TABELA DE DISTRIBUIÇÃO DE FREQUÊNCIAS

Caso já tenhamos a tabela de Distribuição de frequências sem intervalos de classe[4], podemos informá-la diretamente. Para isso, no menu superior clique em PLANILHA → NOVOS DADOS → TABELA DE FREQUÊNCIA.

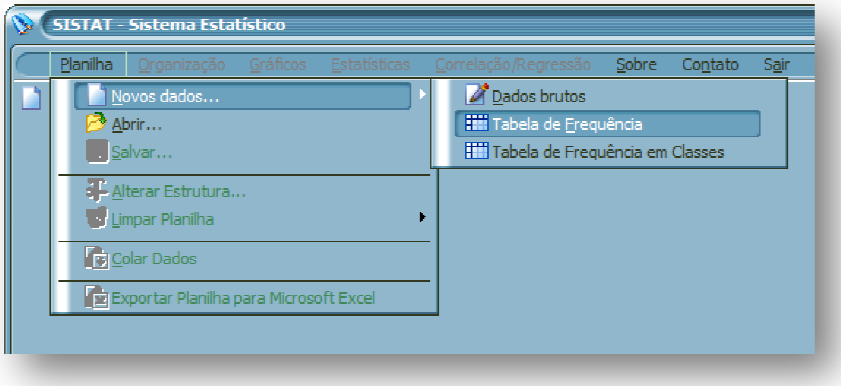


Figura 8 - Entrada de dados - Tabela de Frequência

Na tela principal do software será criada uma tabela de Distribuição de frequências, onde poderemos lançar os dados x_i (valor da variável[5]) e f_i (frequência simples[6]) desta distribuição. As

colunas Fi (Frequência Acumulada[7]), fr_i(%) (Frequência relativa simples em percentual[8]) e Fr_i(%) (Frequência relativa acumulada em percentual[9]) serão calculados automaticamente. Para relembrar o cálculo, consulte a seção 6.6.3.1 (pág. 25).

Informe o tipo de dados que a tabela receberá e o nome da Variável:

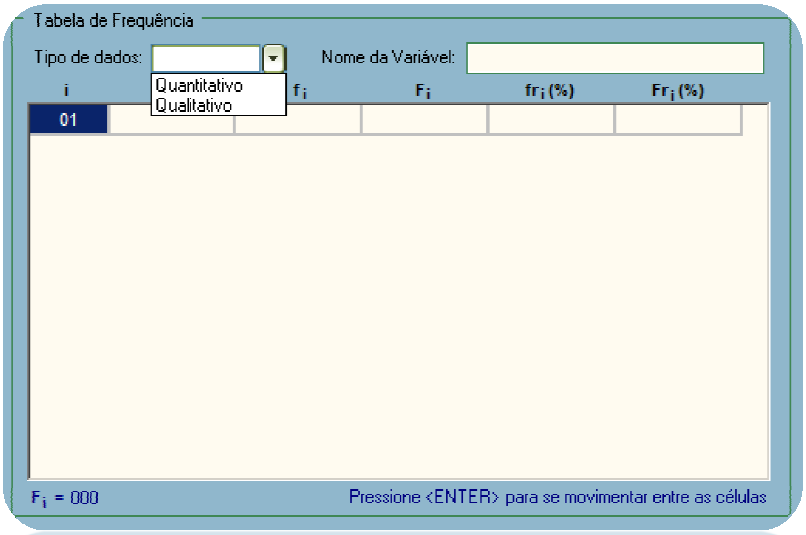
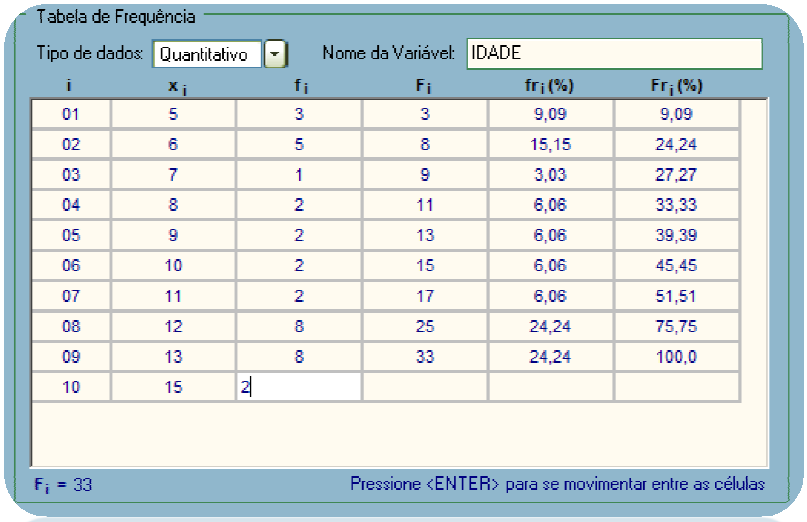


Figura 9 - Entrada de dados - Tabela de Frequência

Como se observa na figura 11, temos uma distribuição de frequências sem intervalos de classe em processo de lançamento. Os dados foram lançados de forma ordenada (x_i), no entanto, poderiam ser lançados fora de ordem.



i	x_i	f_i	F_i	fr_i (%)	Fr_i (%)
01	5	3	3	9,09	9,09
02	6	5	8	15,15	24,24
03	7	1	9	3,03	27,27
04	8	2	11	6,06	33,33
05	9	2	13	6,06	39,39
06	10	2	15	6,06	45,45
07	11	2	17	6,06	51,51
08	12	8	25	24,24	75,75
09	13	8	33	24,24	100,0
10	15	2			

Figura 10 - Entrada de dados - Tabela de Frequência

INFORMANDO A TABELA DE DISTRIBUIÇÃO DE FREQUÊNCIAS COM INTERVALOS DE CLASSE

Caso já tenhamos a tabela de Distribuição de frequências com intervalos de classe[10], podemos informá-la diretamente. Para isso, no menu superior clique em PLANILHA → NOVOS DADOS → TABELA DE FREQUÊNCIA EM CLASSE.

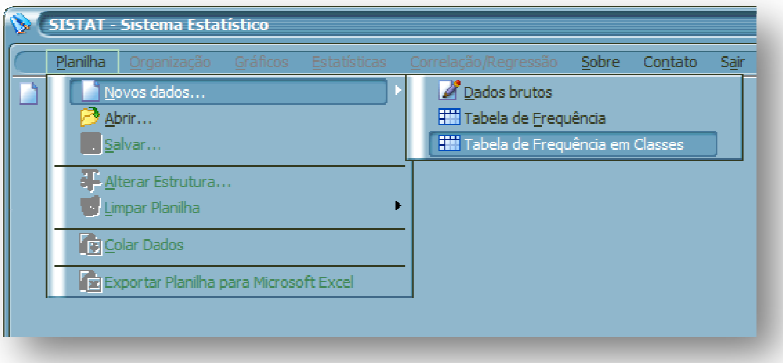


Figura 11 - Entrada de dados - Tabela de Frequência em Classes

Na Tela Principal do software será criada uma tabela de Distribuição de Frequências com Intervalos de Classe. Podemos trabalhar com Intervalos de Classe Fixos[11] ou Intervalos de Classe Variáveis[12].

INTERVALOS DE CLASSE FIXOS

Para Distribuição de Frequências com Intervalos de Classe Fixos, selecione na caixa "Tipo de Intervalo" a opção FIXOS, informemos o "Número de Classes (k)[13]", a "Amplitude das Classes (h)[14]" e o "Limite Inferior da 1ª Classe (L_i)[15]".

O sistema gerará a tabela e liberará a coluna f_i (frequência simples[6]) para inserção de dados. As colunas Fi (Frequência Acumulada[7]), x_i (Ponto Médio[16]), fr_i(%)
frequência relativa simples em percentual[8]) e Fr_i(%) (Frequência relativa acumulada em percentual[9]) serão calculados automaticamente.

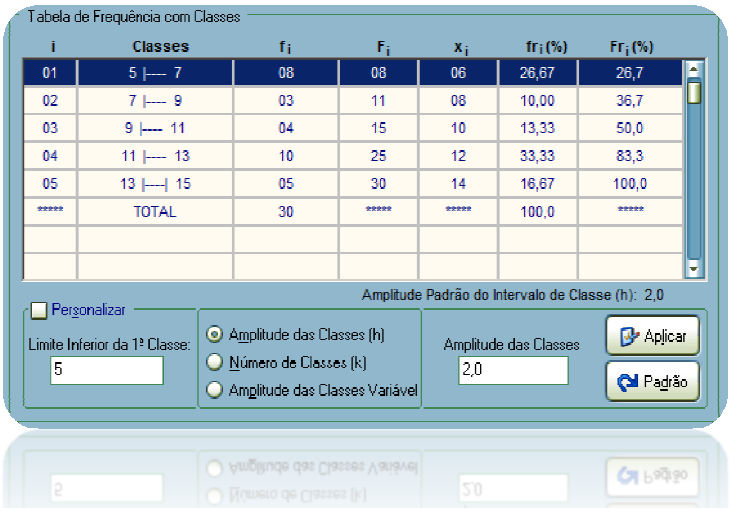


Figura 12 - Entrada de dados - Tabela de Frequência em Classes

INTERVALOS DE CLASSE VARIÁVEIS

Para Distribuição de Frequências com Intervalos de Classe Variáveis, selecione na caixa "Tipo de Intervalo" a opção VARIÁVEIS,será apresentada uma janela para construção dos limites e números de classe.

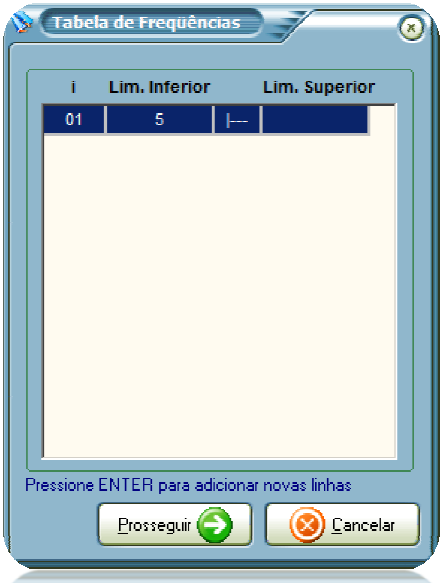


Figura 13 - Tabela de Frequência em Classes

Informe os Limites Inferior e Superior da 1ª Classe e os superiores das demais classes, utilize a tecla ENTER para que o cursor se direcione para a próxima célula.

Após concluir o preenchimentos dos limites superior e inferior de todas as classes, clique na tecla PROSSEGUIR.

O sistema gerará a tabela e liberará a coluna f_i (frequência simples[6]) para inserção de dados. As colunas Fi (Frequência Acumulada[7]), x_i (Ponto Médio[16]), fr_i(%) (Frequência relativa simples em percentual[8]) e Fr_i(%) (Frequência relativa acumulada em percentual[9]) serão calculados automaticamente. Para relembrar o cálculo, consulte a seção 6.6.3.1 (pág. 25).

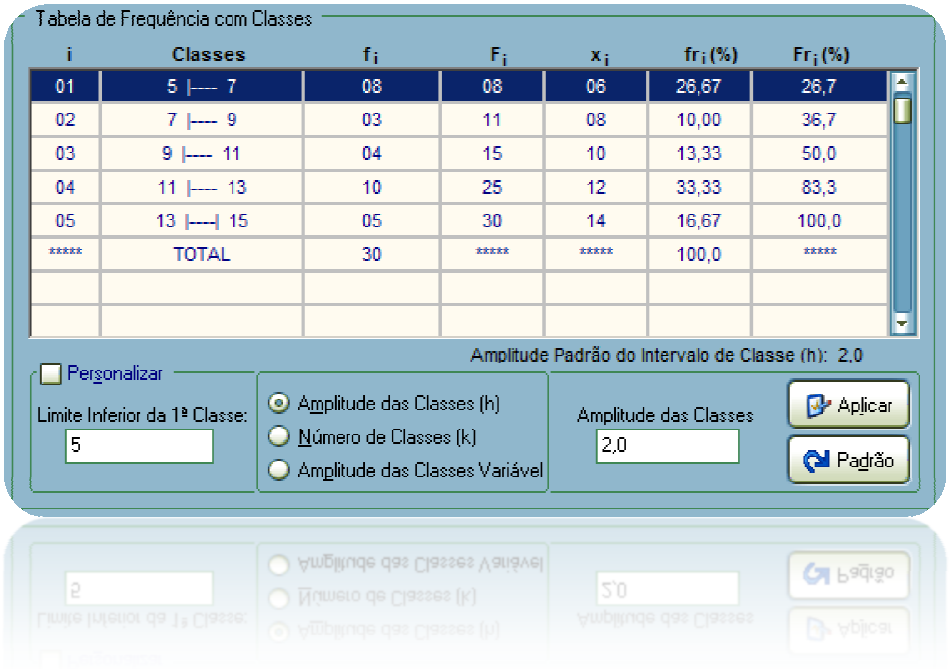


Figura 14 - Entrada de dados - Tabela de Frequência em Classes

RECUPERANDO DADOS DE PESQUISA SALVOS ANTERIORMENTE

É possível recuperar dados lançados anteriormente, podendo, desta forma, retomar um trabalho de pesquisa iniciado em outro momento. Para recuperar uma massa de dados salva, no menusuperior clique em PLANILHA → ABRIR.

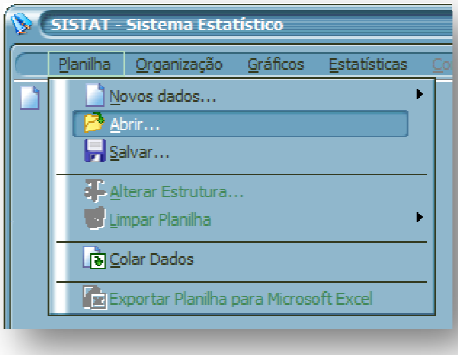


Figura 15 - Recuperar dados salvos

O Sistat salva os dados em arquivos gerados a partir da codificação ASCII, com extensão do arquivo .sta. Sua estrutura está detalhada no anexo 1.

Será exibida uma Caixa de Diálogo que exibirá apenas os arquivos que tenham sido gerados pelo Sistat, selecione o arquivo desejado e clique em "ABRIR".

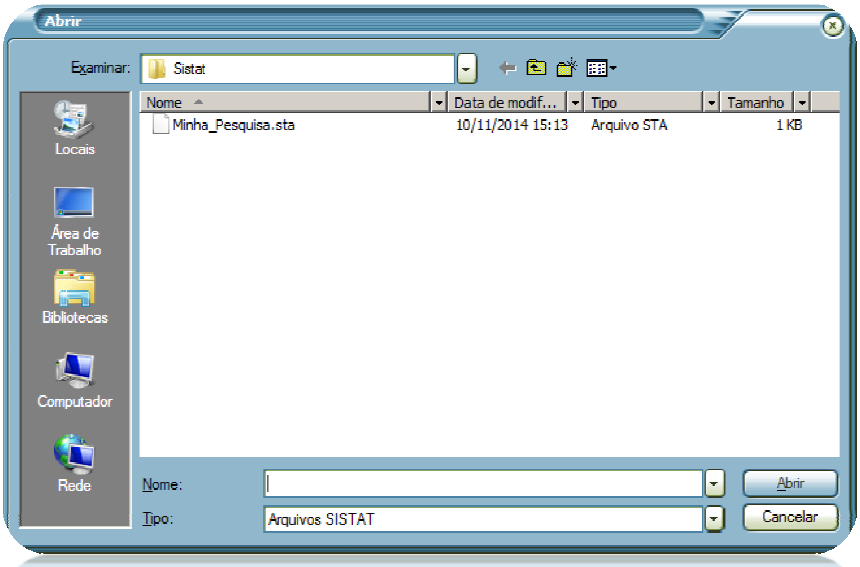


Figura 16 - Tabela de Frequência em Classes

Após a abertura do arquivo, o Sistat importará as configurações da(s) variável(is) e os dados brutos lançados serão recuperados.

7.2 - SALVANDO DADOS DA PESQUISA

Podemos salvar as informações digitadas no Sistat, para que no futuro possamos recuperá-las e continuar nosso estudo estatístico. Serão salvas as informações a respeito da quantidade de variáveis, seus tipos e os dados lançados.

Para salvar, no menu superior clique em PLANILHA → SALVAR.

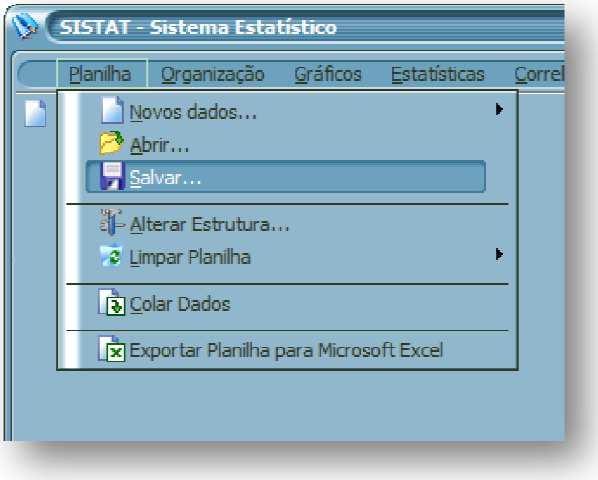


Figura 17 - Salvar dados lançados

Será apresentada uma Caixa de Diálogo, informe o local de salvamento, o nome do arquivo e clique em SALVAR.

O Sistat gera arquivos a partir da codificação ASCII, com extensão do arquivo .sta.

<<<< DETALHAR ESTRUTURA DO ARQUIVO >>>>>>

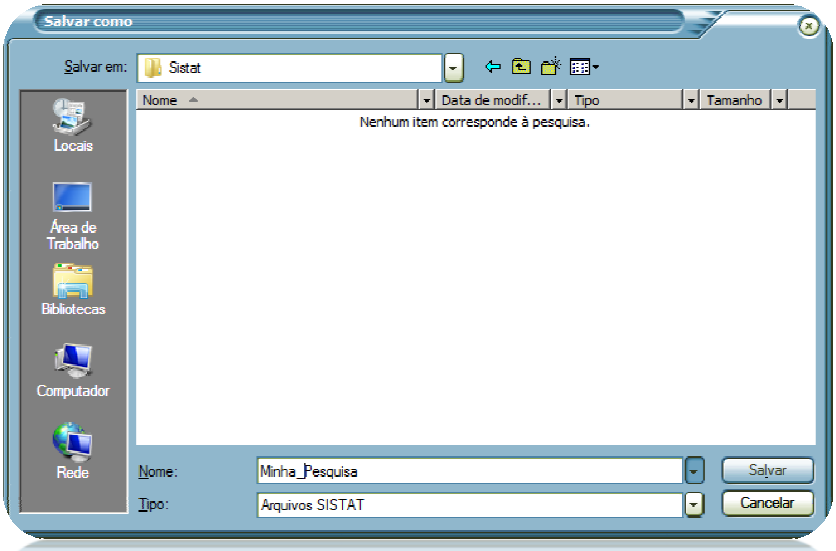


Figura 18 - Salvar dados lançados

7.3 - ALTERANDO A ESTRUTURA DA TABELA PRIMITIVA

Mesmo após termos definido quais variáveis iremos trabalhar e o lançamento de suas respectivas informações, ainda assim podemos alterar a estrutura de nossa origem de dados, no sentido de incluir ou excluir variável, alterar o tipo de dados que receberão, etc.

Para isso, no menu superior clique em PLANILHA → ALTERAR ESTRUTURA.

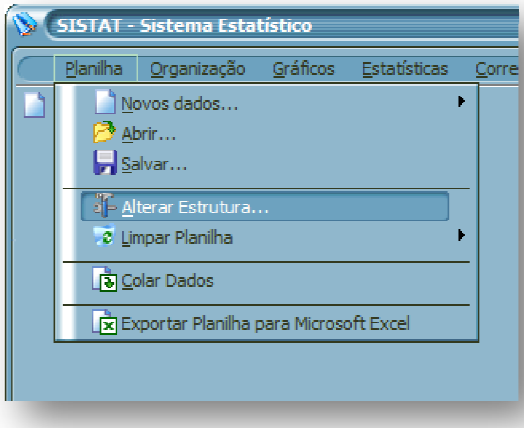


Figura 19 - Alterar Estrutura da Tabela de Variáveis

Será apresentada a janela do Assistente da Tabela de Variáveis. No campo "Nº de variáveis" podemos aumentar ou diminuir a quantidade de variáveis. Informe o(s) nome(s) da(s) variável(is) (exemplo: Idade, Peso, Time de futebol) e o(s) respectivo(s) Tipo(s) de dados[2].

O tipo de dados que cada variável irá receber, deverá ser informado de acordo com sua classificação: Qualitativa[2.1] ou Quantitativa[2.2]. Bem como seu subtipo, Sendo Qualitativa (Nominal ou Ordinal) ou Quantitativa (Discreta ou Contínua).

Na hipótese de trabalharmos com uma variável Quantitativa Contínua, podemos informar o número de casas decimais para que o sistema padronize a entrada de dados.

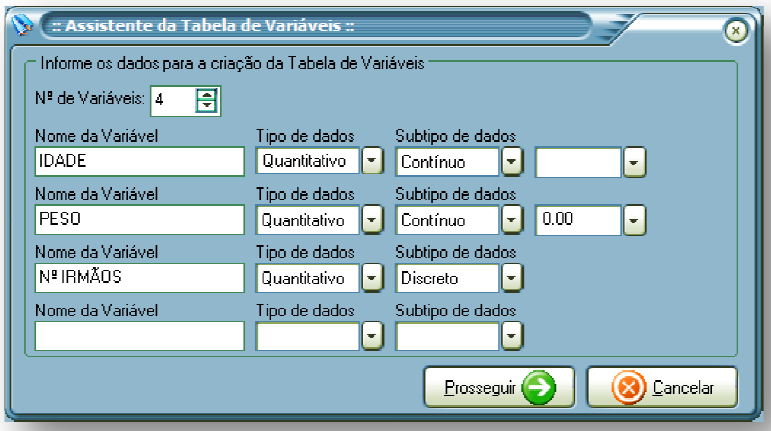


Figura 20 - Alterar Estrutura da Tabela de Variáveis

Clique em PROSSEGUIR, na Tela Principal será(ão) apresentada(s) a(s) tabela(s) para edição dos dados da pesquisa.

EXPORTANDO TABELA PRIMITIVA PARA O MICROSOFT EXCEL

É possível exportar os dados digitados na tabela primitiva para a planinha eletrônica Microsoft Excel. No menu superior clique em PLANILHA → EXPORTAR PLANILHA PARA MICROSOFT EXCEL.

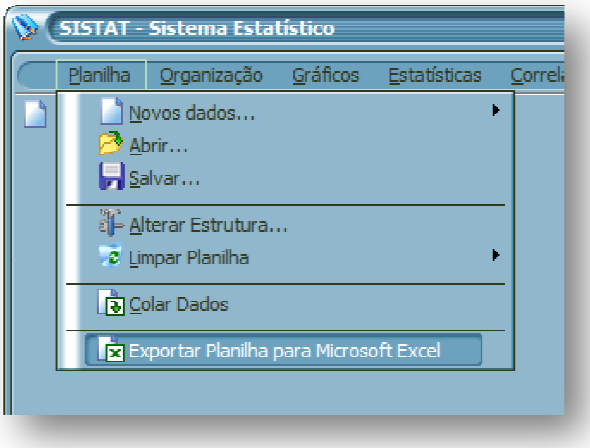


Figura 21 - Exportação de dados

Automaticamente será aberto o Microsoft Excel (deverá ser instalado previamente) e criada uma nova planilha contendo as dados informados na tabela primitiva, inclusive tendo como cabeçalho o(s) nome(s) da(s) variável(s).

APAGANDO DADOS DA TABELA PRIMITIVA

Caso seja necessário apagar dados de uma tabela primitiva, é possível realizar essa operação de forma parcial ou total.

APAGANDO DADOS POR SELEÇÃO

Para apagar os dados de forma parcial ou por seleção de uma tabela primitiva, com o auxílio do mouse ou do teclado (combinando as teclas CTRL+↓ ou CTRL+↑) selecione a área que deseja apagar. Em seguida, no menu superior clique em PLANILHA → LIMPAR PLANILHA → APENAS SELEÇÃO.

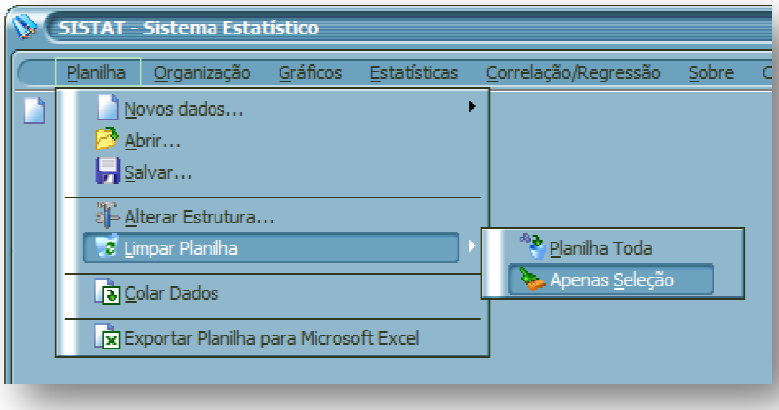


Figura 22 - Apagar dados da tabela primitiva por seleção

APAGANDO TODOS OS DADOS

Para apagar todos os dados de uma tabela primitiva, no menu superior clique em PLANILHA → LIMPAR PLANILHA → PLANILHA TODA.

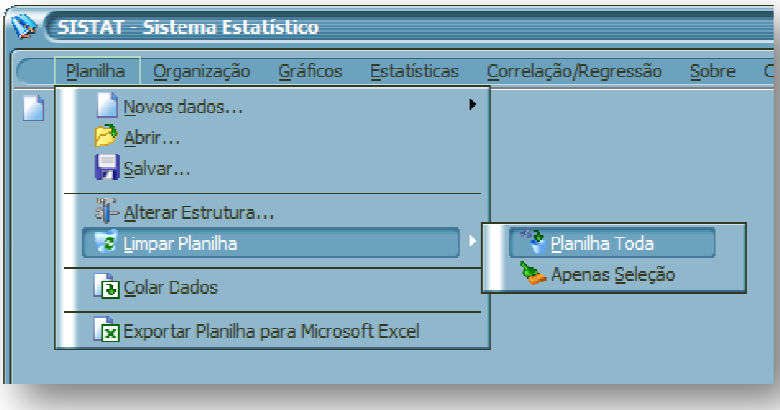


Figura 23 - Apagar tabela primitiva

ORGANIZAÇÃO E TRATAMENTO DA INFORMAÇÃO

Antes de iniciar a organização e tratamento da informação, devemos selecionar a variável de trabalho. Para isso, na região centro-inferior da Tela Principal clique na Combobox abaixo de "Variável de Trabalho" e selecione a variável de trabalho desejada.

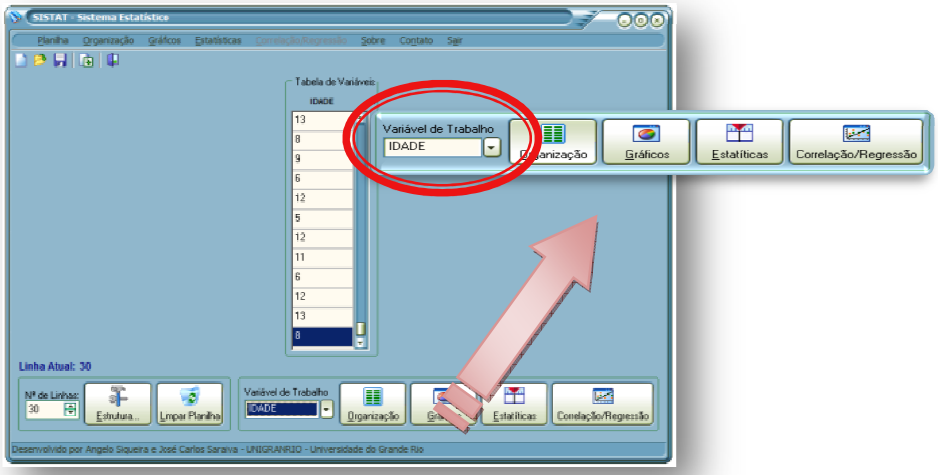


Figura 24 –Organização da Informação

ROL

Para organizar uma massa de dados, isto é, colocá-la em Rol[16], no menu superior clique em ORGANIZAÇÃO → CLASSIFICAR DADOS.

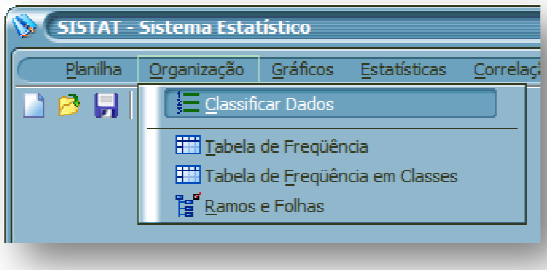


Figura 25 – Organização da Informação

Será apresentada uma nova janela com uma lista vertical das ocorrências para a variável selecionada, organizada de forma crescente.

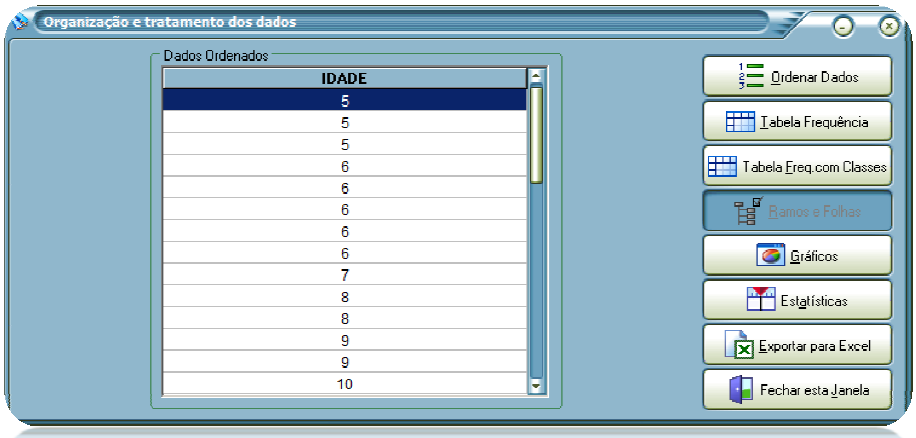


Figura 26 – Organização da Informação

Tomando como exemplo a base de dados disponibilizada e selecionando a variável IDADE como variável de trabalho, teremos o seguinte Rol[16]

5 - 5 - 5 - 6 - 6 - 6 - 6 - 6 - 7 - 8 - 8 - 9 - 9 - 10 - 10 - 11 - 11 - 12 - 12 - 12 - 12 - 12 - 12 - 12 - 12 - 13 - 13 - 13 - 15 - 15

TABELA DE DISTRIBUIÇÃO DE FREQUÊNCIAS

Como visto na seção 6.6.3 – Página 22, o tratamento dos dados resulta na construção da tabela de distribuição de freqüências.

Para termos acesso a tabela de distribuição de freqüências, no menu superior clique em ORGANIZAÇÃO → TABELA DE FREQUÊNCIAS.

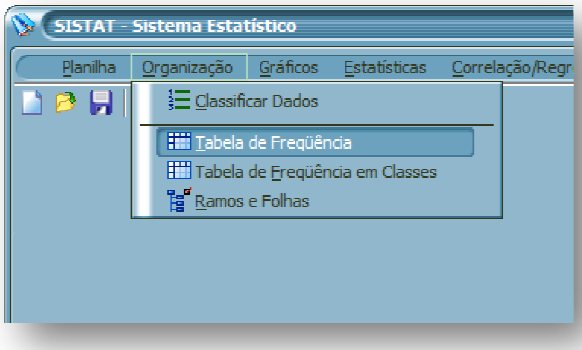
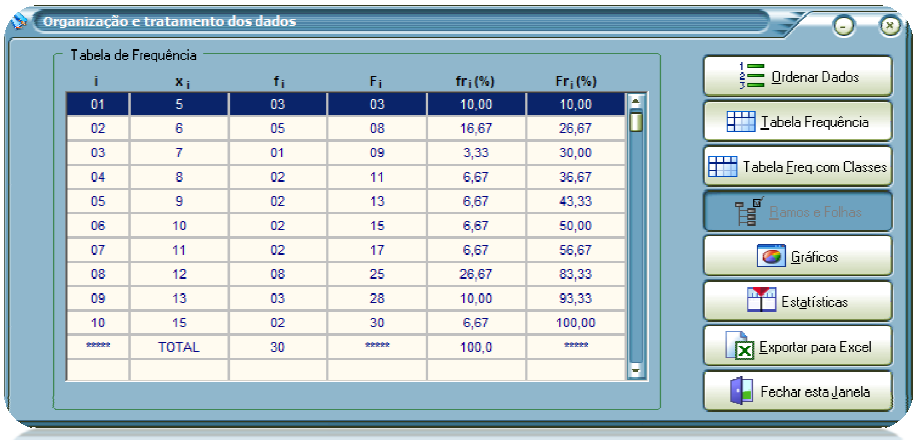


Figura 27 – Distribuição de Frequências

Será apresentada uma janela com a tabela de Distribuição de frequências, gerada a partir da tabela primitiva que contém os dados da variável IDADE. Esta tabela contém as seguinte colunas: i (índice), x_i (valor da variável[5]) e f_i (frequência simples[6]), F_i (Frequência Acumulada[7]), fr_i(%) (Frequência relativa simples em percentual[8]) e Fr_i(%) (Frequência relativa acumulada em percentual[9])



i	x_i	f_i	F_i	fr_i (%)	Fr_i (%)
01	5	03	03	10,00	10,00
02	6	05	08	16,67	26,67
03	7	01	09	3,33	30,00
04	8	02	11	6,67	36,67
05	9	02	13	6,67	43,33
06	10	02	15	6,67	50,00
07	11	02	17	6,67	56,67
08	12	08	25	26,67	83,33
09	13	03	28	10,00	93,33
10	15	02	30	6,67	100,00
*****	TOTAL	30	*****	100,0	*****

Figura 28 – Distribuição de Frequências

7.5.3 – TABELA DE DISTRIBUIÇÃO DE FREQUÊNCIAS COM INTERVALOS DE CLASSE

Como visto na seção 6.6.3 – Página 24, o tratamento dos dados resulta na construção da tabela de Distribuição de Frequências com Intervalos de Classe.

Para termos acesso a tabela de distribuição de Frequências com Intervalos de Classe, no menu superior clique em ORGANIZAÇÃO → TABELA DE FREQUÊNCIAS EM CLASSES.

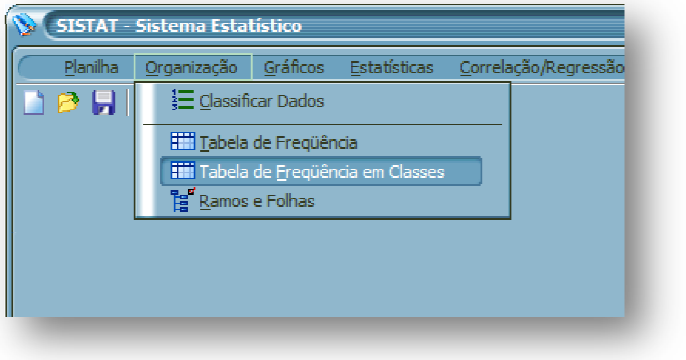


Figura 29 – Distribuição de Frequências *com Intervalos de Classe*

Será apresentada uma janela com a tabela de Distribuição de frequências, gerada a partir da tabela primitiva que contém os dados da variável IDADE. Esta tabela contém as seguinte colunas: i (índice), Classes[18], f_i (frequência simples[6]), F_i (Frequência Acumulada[7]), x_i (ponto médio[17]), fr_i(%) (Frequência relativa simples em percentual[8]) e Fr_i(%) (Frequência relativa acumulada em percentual[9])

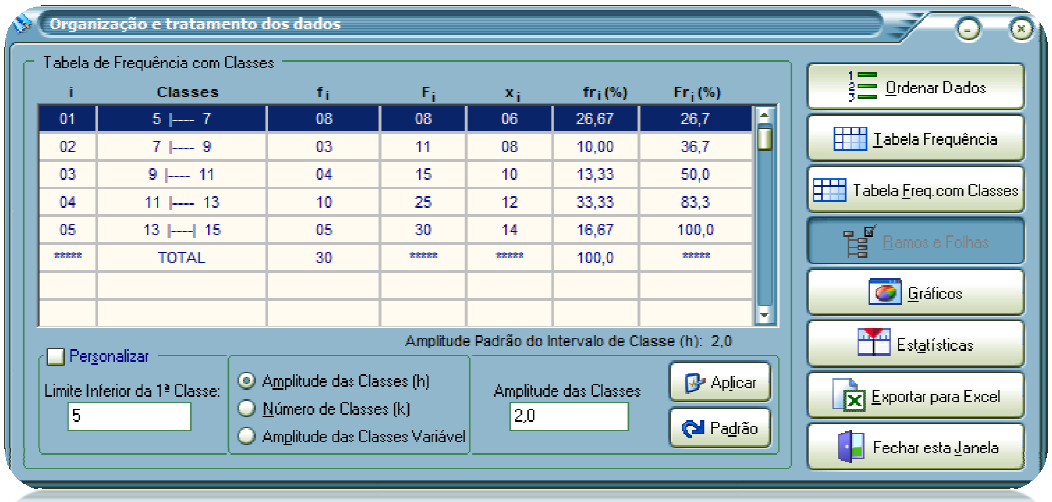


Figura 30 – Distribuição de Frequências com Intervalos de Classe

7.5.3 – DIAGRAMA DE RAMOS E FOLHAS

Embora o Diagrama de Ramos e Folhas seja um recurso semi-gráfico, ele estabelece uma relação com a forma de organização e apresentação dos dados semelhante com a Distribuição de Frequências.

Para visualizar o Diagrama de Ramos e Folhas[19], no menu superior clique em ORGANIZAÇÃO → RAMOS E FOLHAS. Importante frisar que essa opção só estará disponível na hipótese de a Variável de Trabalho selecionada ser Quantitativa[2.2] Discreta.

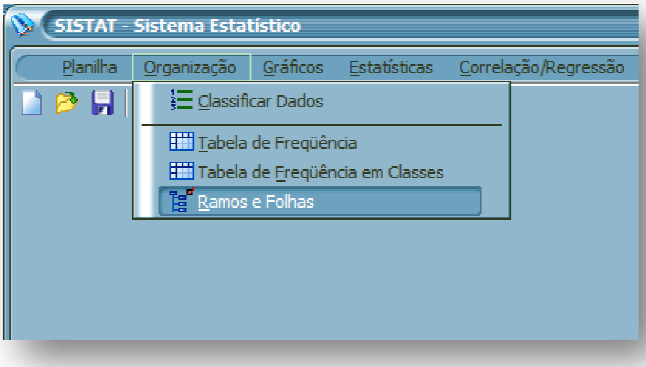


Figura 31 – Diagrama de Ramos e Folhas

Será apresentada uma nova janela com o Diagrama de Ramos e Folhas, como a seguir, na figura 32:

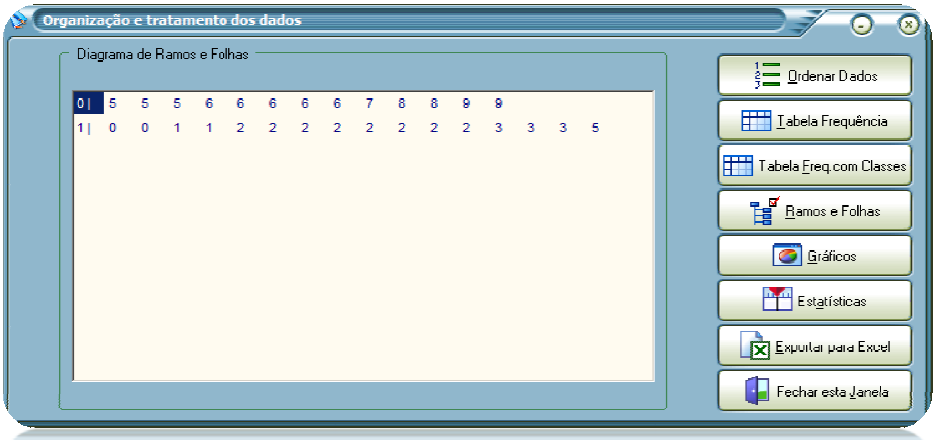


Figura 32 – Diagrama de Ramos e Folhas

OBSERVAÇÃO:

Todas as opções acessadas através do menu superior, opção ORGANIZAÇÃO, podem ser acessadas através do botão localizado na parte centro-inferior da Tela Principal (Figura 33).

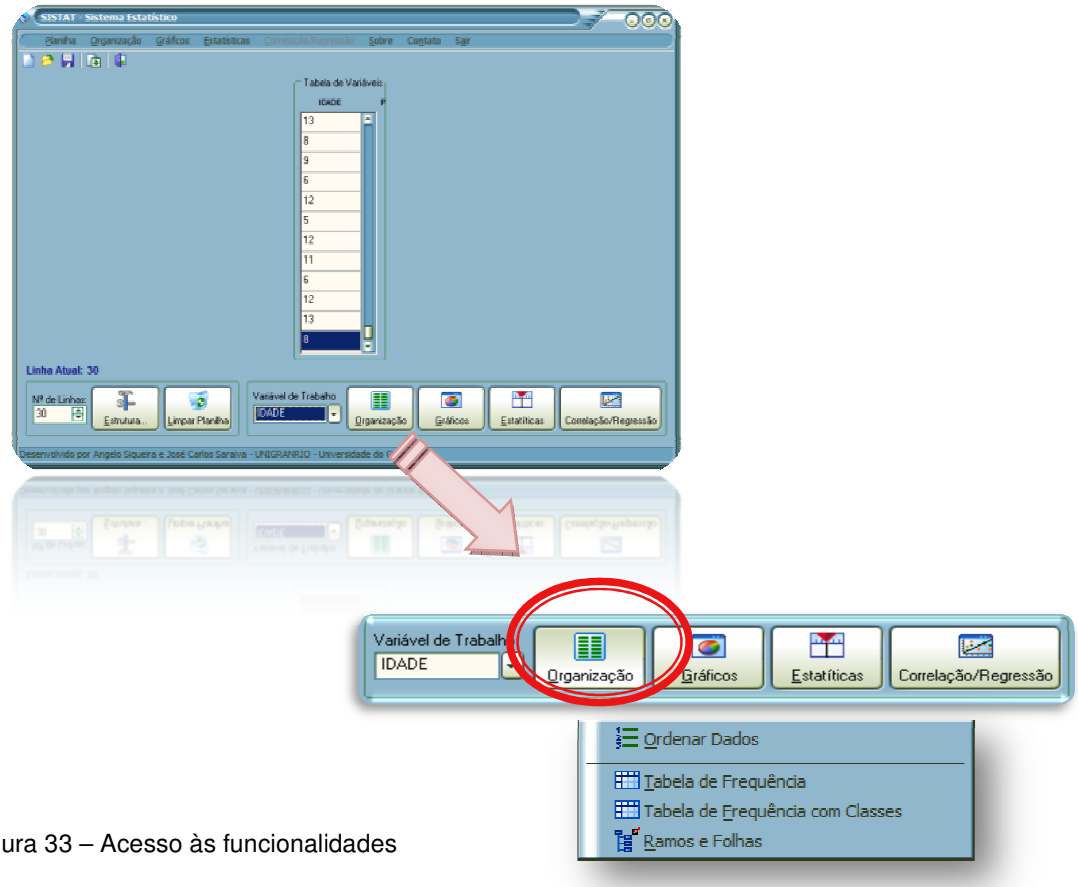


Figura 33 – Acesso às funcionalidades

7.6. REPRESENTAÇÃO GRÁFICA

Para acessar a área dos gráficos, localiza na região centro-inferior da Tela Principal, clique na Caixa de Listagem abaixo de "Variável de Trabalho" e selecione a variável de trabalho desejada. Para fins de padronização, nos exemplos a seguir utilizaremos a variável IDADE, importada a partir da base de dados disponibilizada.

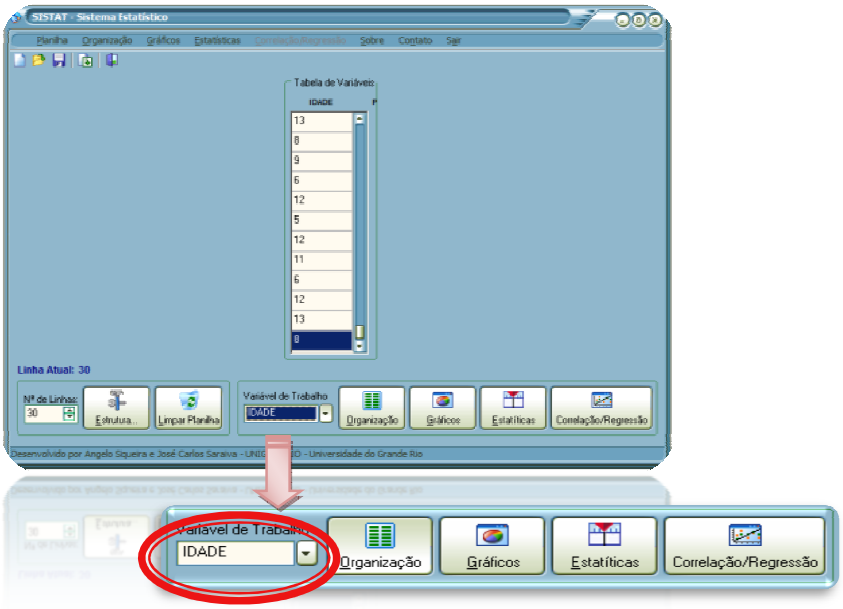


Figura 34–Selecionando a Variável de Trabalho

7.6.1 - REPRESENTAÇÃO GRÁFICA (DISTRIBUIÇÃO DE FREQUÊNCIAS SIMPLES)

Para distribuições de frequências simples, podemos gerar com o Sistat as seguintes representações gráficas: Gráficos em Bastões[19], Gráficos em Colunas[20] (Barras Verticais), Gráficos em Barras[21] (Barras Horizontais), Gráficos em Setores[22] (Pizza).

Após selecionar a variável de trabalho, escolha o gráfico desejado através das opções listadas no menu superior, opção GRÁFICOS.

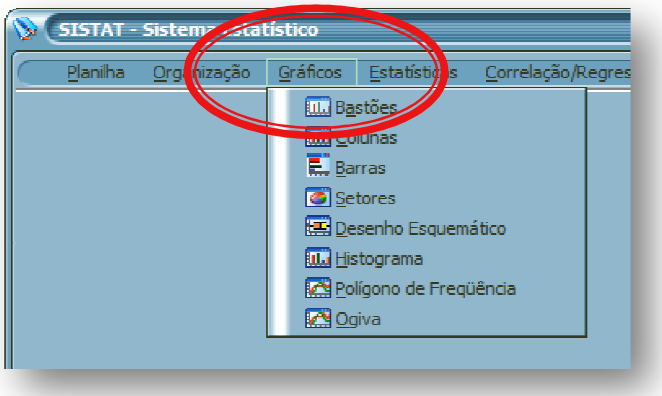


Figura 35–Representação Gráfica

Ao selecionar a representação gráfica desejada, será apresentada uma nova janela, de acordo com as figuras a seguir:

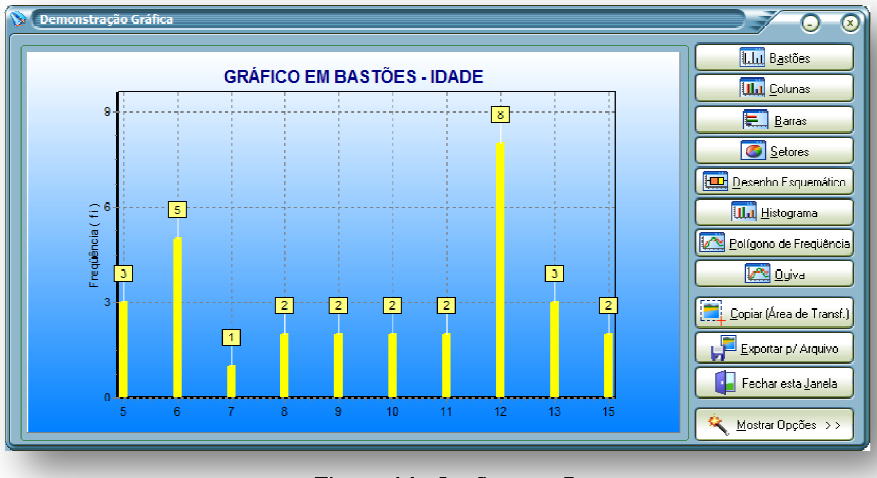


Figura 36–Gráfico em Bastões

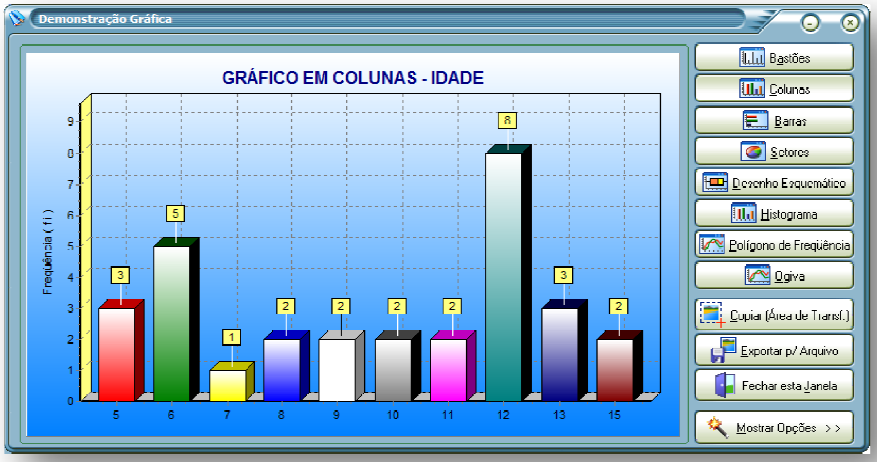


Figura 37–Gráfico em Colunas

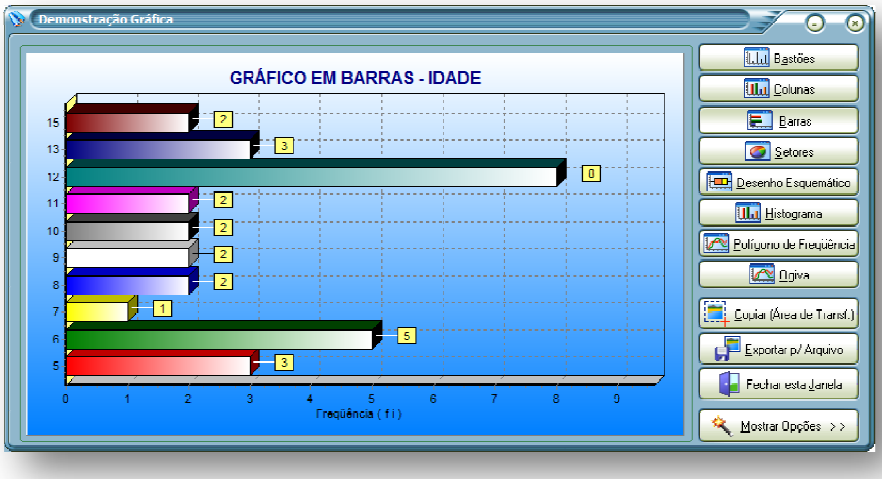


Figura 38 –Gráfico em Barras

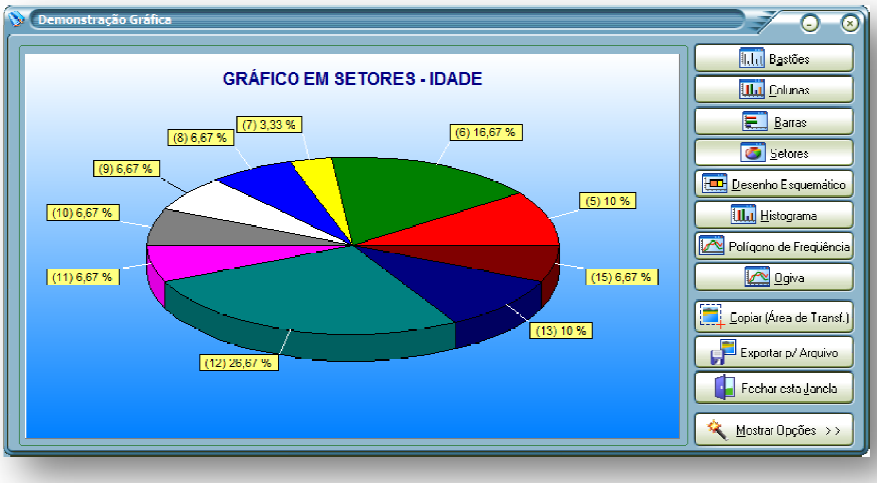


Figura 39 –Gráfico em Setores

7.6.2 - REPRESENTAÇÃO GRÁFICA (DISTRIBUIÇÃO DE FREQUÊNCIAS COM INTERVALOS DE CLASSE)

Para distribuições de Frequências com Intervalos de Classe, podemos gerar com o Sistat as seguintes representações gráficas: Histograma[23], Polígono de Frequências[24]eOgiva[25].

Após selecionar a variável de trabalho, escolha o gráfico desejado através das opções listadas no menu superior, opção GRÁFICOS.

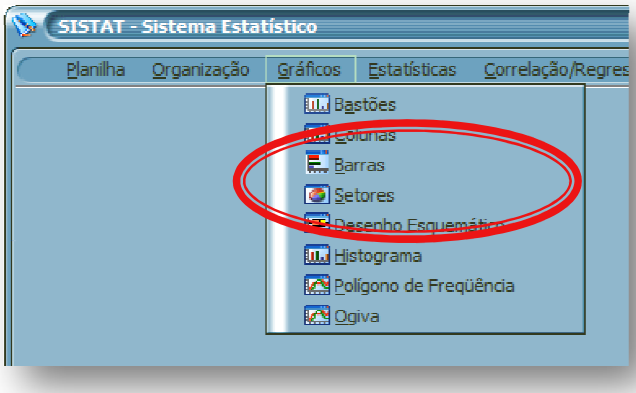


Figura 40 –Representação Gráfica

Ao selecionar a representação gráfica desejada, será apresentada uma nova janela, de acordo com as figuras a seguir:

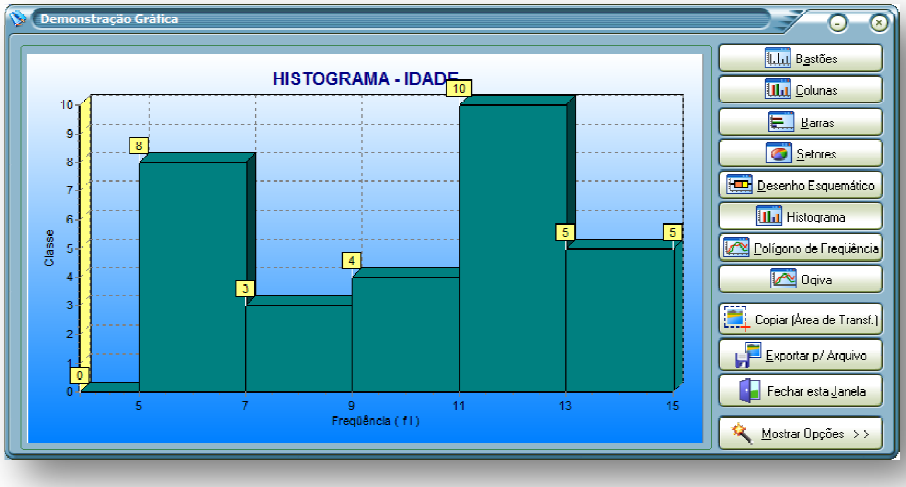


Figura 41 – Histograma

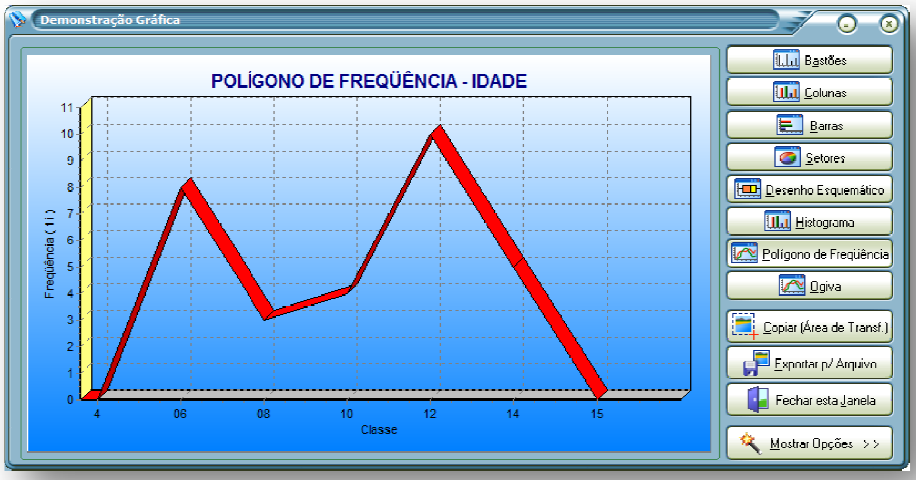


Figura 42 – Polígono de Frequência

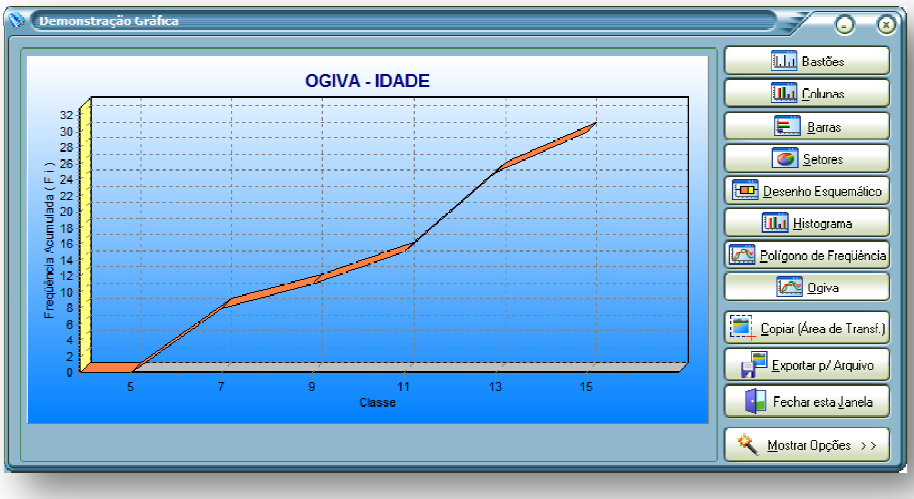


Figura 43 – Ogiva

7.6.3 - REPRESENTAÇÃO GRÁFICA (DESENHO ESQUEMÁTICO/BOXPLOT)

Para distribuições de frequências com ou sem Intervalos de Classe, podemos gerar com a representação gráfica chamada Desenho Esquemática (Boxplot).

Após selecionar a variável de trabalho, escolha o gráfico desejado através das opções listadas no menu superior, opção GRÁFICOS.

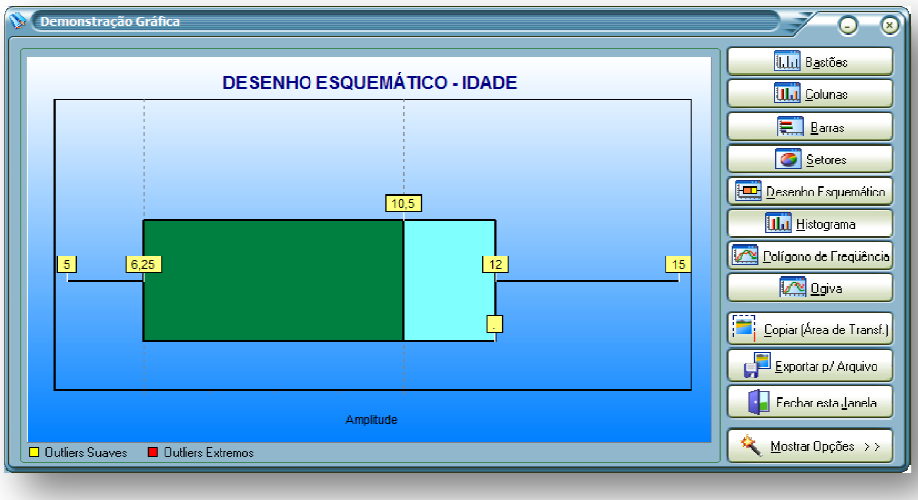


Figura 44 – Desenho Esquemático / BoxPlot

7.6.4 – EXPORTAR OU SALVAR GRÁFICOS

Os gráficos gerados podem ser utilizados em outros aplicativos ou exportados para arquivos no formato jpeg. Na janela de visualização dos gráficos encontramos o botão “EXPORTAR P/ ARQUIVO”, que salva o gráfico visualizado em um arquivo no formato Jpeg e o botão “COPIAR (ÁREA DE TRANSF.)” que envia o gráfico visualizado para a área de transferência do Windows, permitindo que possa ser colado (CTRL+V) em outro aplicativo (como por exemplo, um editor de textos).

Figura 45 – Salvar/Exportar

7.7 - CÁLCULO DAS MEDIDAS ESTATÍSTICAS

Como visto na seção 6.8 – Página 29, as medidas estatísticas têm como objetivo resumir as informações coletadas e podemos visualizá-las separadamente para Distribuições de Frequências Simples e Distribuições de Frequências com Intervalos de Classe, com o SISTAT temos acesso às principais medidas estatísticas: Medidas de Tendência Central[26], Separatrizes[27], Medidas de Dispersão[28] e Medidas de Assimetria[29] e Curtose[30].

7.7.1 - MEDIDAS ESTATÍSTICAS (DADOS NÃO AGRUPADOS EM CLASSES)

Voltando à seção 5.2 (Distribuição de Frequências Simples), onde vimos a tabela de Distribuições de Frequências, ao clicarmos no botão Estatísticas nos é apresentada a janela com estas medidas.

As estatísticas a seguir são referentes a variável IDADE, importada a partir da base de dados disponibilizada.

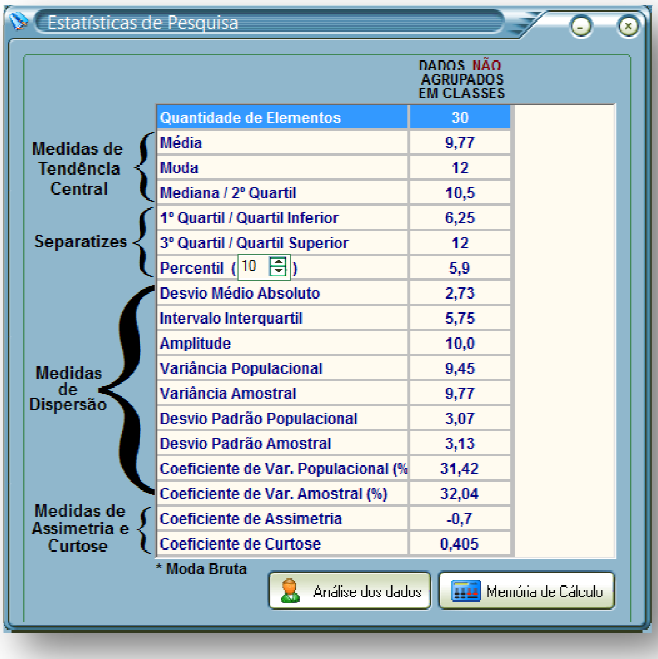


Figura 46 – Estatísticas da Pesquisa (Dados não agrupados em Classes)

7.7.2 - MEDIDAS ESTATÍSTICAS (DADOS AGRUPADOS EM CLASSES)

Voltando à seção 5.3 (Distribuição de Frequências Simples com Intervalos de Classe), onde vimos a tabela de Distribuições de Frequências com Intervalos de Classe, ao clicarmos no botão Estatísticas a janela com estas medidas.

As estatísticas a seguir são referentes a variável IDADE, importada a partir da base de dados disponibilizada.

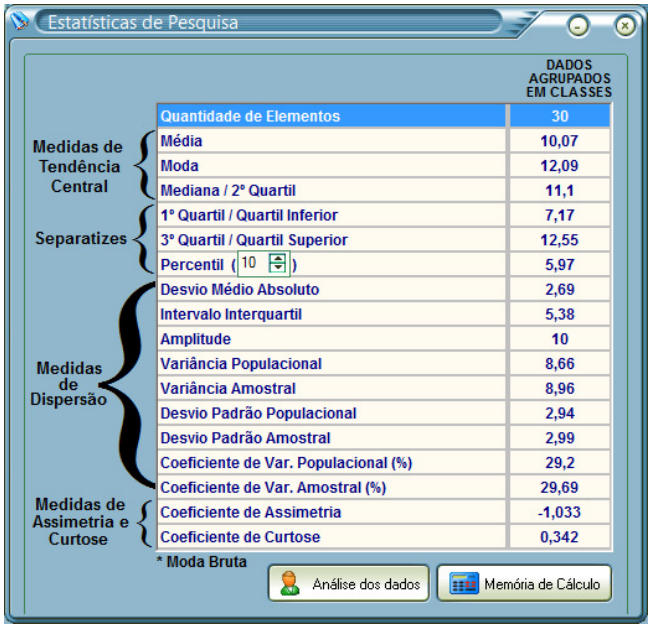


Figura 47 – Estatísticas da Pesquisa
(Dados agrupados em Classes)

7.7.3 - MEDIDAS ESTATÍSTICAS (COMPARATIVO)

É possível, ainda, compararmos as estatísticas da nossa pesquisa simulando distribuições de frequências com dados agrupados em classes e não agrupados em classes.

Para acessarmos o comparativo, no menu superior clique em ESTATÍSTICAS → ESTATÍSTICAS DESTA PESQUISA.

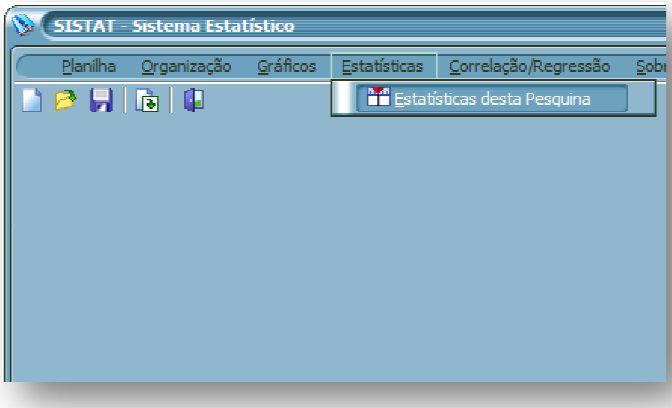


Figura 48 – Estatísticas da Pesquisa (Comparativo)

Na janela apresentada, nos será disponibilizada as medidas para distribuições de freqüências agrupadas ou não em classes.

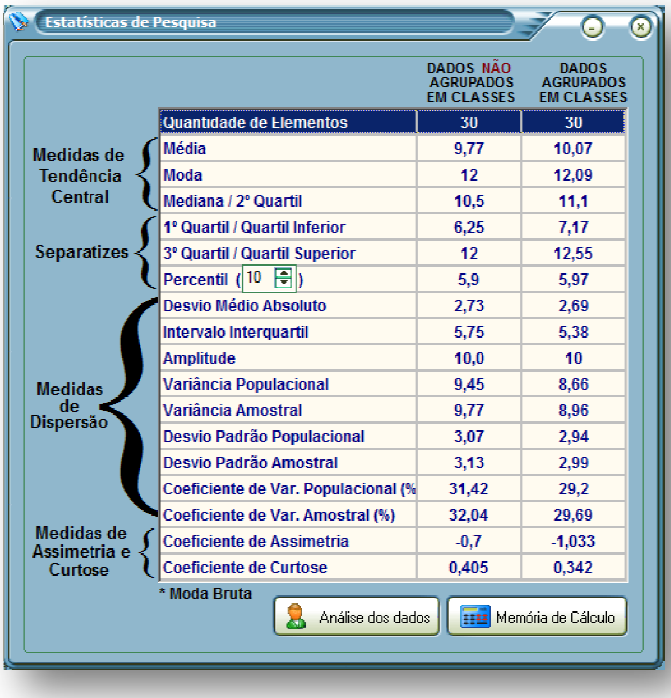


Figura 49 – Estatísticas da Pesquisa (Comparativo)

7.8 – CORRELAÇÃO E REGRESSÃO

Antes de iniciarmos os estudos da Correlação e da Regressão, temos que observar se possuímos informações de, no mínimo, 2 variáveis. Caso tenhamos apenas dados de 1 variável, devemos consultar a *seção 3 - Alterando a Estrutura da tabela primitiva* para adicionarmos uma ou mais variáveis à tabela de variáveis e seus respectivos dados.

Para acessar a área da Correlação e Regressão, na região inferior direita da Dashboard[3] clique no botão CORRELAÇÃO/REGRESSÃO.

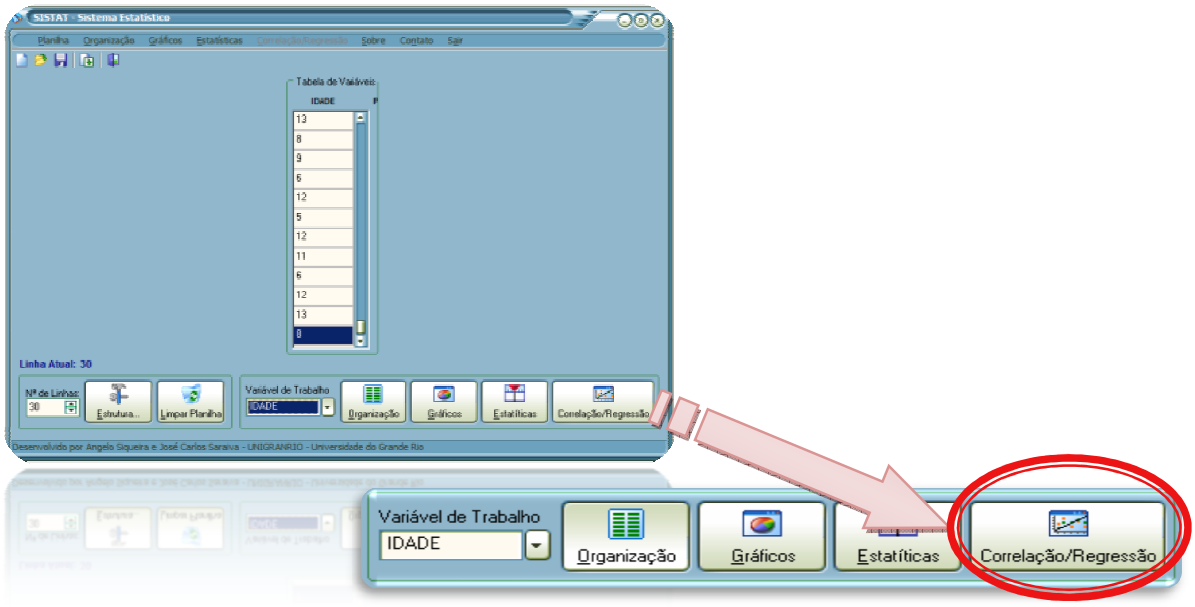
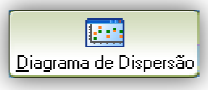


Figura 50 – Correlação/Regressão

Será apresentada a janela a seguir, onde devemos selecionar as variáveis que trataremos como X e Y.

Após selecioná-las poderemos visualizar o Diagrama de Dispersão[31], clicando no botão



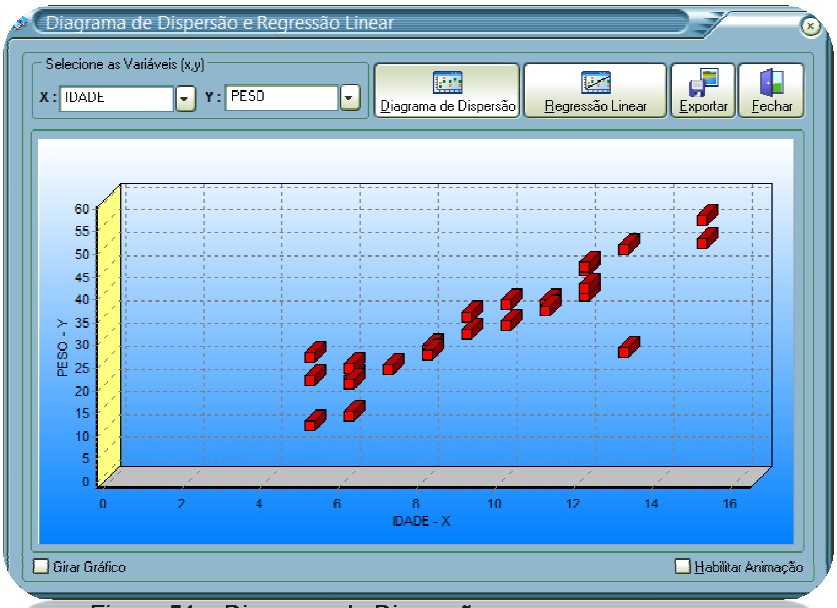
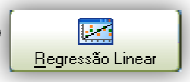


Figura 51 – Diagrama de Dispersão

Para termos acesso ao Coeficiente de Correlação Linear de Pearson (r)32, a Equação da Reta de Regressão33 clique no botão 

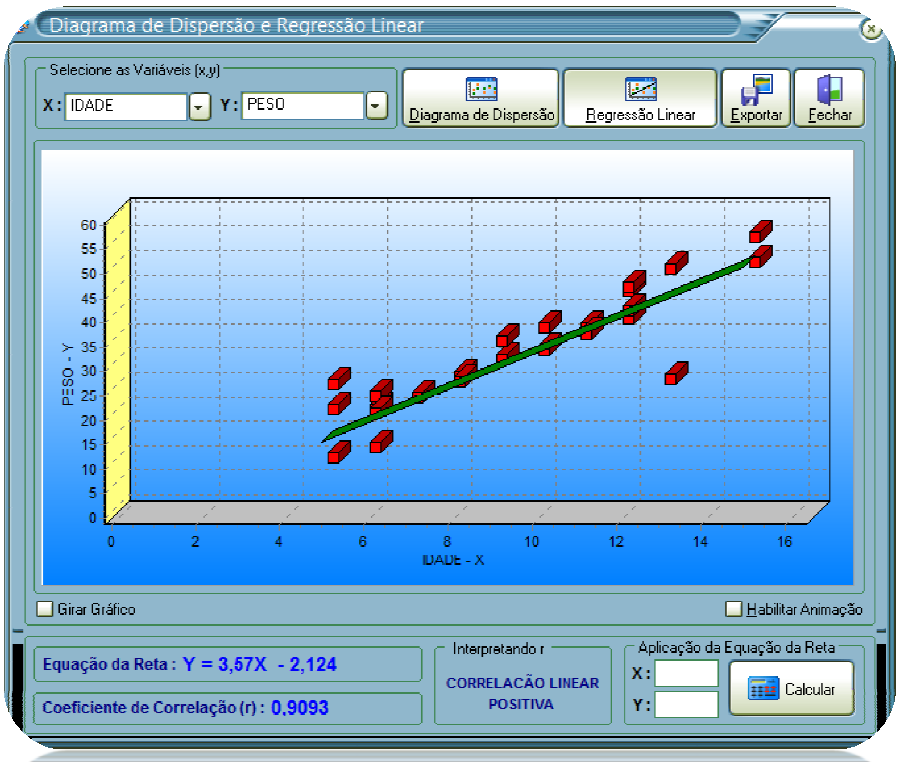


Figura 52 – Diagrama de Dispersão

APÊNDICE B

LAYOUT DE ARQUIVOS TXTPARA EXPORTAÇÃO/IMPORTAÇÃO NO SOFTWARE "SISTAT"

1. ESTRUTURA DOS DADOS

Os arquivos deverão ser gerados no modo texto, no padrão ASCII MS-DOS, com a seguinte estrutura e convenção:

1.1. ESTRUTURA

Os arquivos deverão estar estruturados da seguinte forma:

***HEADER:** Será formado pelas 1ª e 2ª linhas de cada arquivo e conterão a quantidade de variáveis e o número de ocorrências (n).

Registro do tipo HEADER (1ª linha) para todos os arquivos

Descrição do Campo	Tipo	Tamanho (em bytes)	Observações	Campo Obrigatório
Quantidade de Variáveis	Inteiro	02	-	Sim

Registro do tipo HEADER (2ª linha) para todos os arquivos

Descrição do Campo	Tipo	Tamanho (em bytes)	Observações	Campo Obrigatório
Número de Ocorrências (n)	Inteiro	08	-	Sim

Exemplo (Distribuição contendo 3 variáveis e com até 5 ocorrências cada):
3
5

***DADOS:** Deverão ser preenchidos de acordo com a estrutura de cada arquivo descrito neste documento.

Ocorrerá a partir da 3ª linha e receberá todas as ocorrências da tabela primitiva, na tela principal. Cada linha receberá apenas uma ocorrência.

Método: Serão carregadas todas as ocorrências da 1ª variável, após carrega-se todas as ocorrências da 2ª variável e assim sucessivamente:

Registro do tipo DADOS para todos os arquivos

Descrição do Campo	Tipo	Tamanho (em bytes)	Observações	Campo Obrigatório
Ocorrências	Texto	25	-	Sim

Exemplo (Distribuição contendo 3 variáveis e com até 5 ocorrências cada):
32,90
32,00
34,00
4,00
4,00
VERDE
AZUL
PRETO
VERMELHO
PRETO
ALTA

BAIXA
BAIXA
MEDIA
BAIXA

***CONFIGURAÇÃO:** Deverão ser preenchidos de acordo com a estrutura de cada arquivo descrito neste documento.
Ocorrerá a partir da linha imediatamente posterior a última linha de dados.
Método: Serão gerados 10 grupos de configuração de variável, contendo 4 linhas cada. Cada configuração de variável deve ser gerada na mesma ordem dos registros do tipo DADOS.
Registro do tipo CONFIGURAÇÃO para todos os arquivos

Descrição do Campo	**Tipo**	**Tamanho (em bytes)**	**Observações**	**Campo Obrigatório**
Nome da Variável	Texto	25	Permite Vazio	Sim
Tipo dos dados	Inteiro	1	-1 = Variável não informada 0 = Variável Quantitativa 1 = Variável Qualitativa	Sim
Subtipo dos dados	Inteiro	1	-1 = Variável não informada Se Tipo = 0 0 = Variável Quantitativa Contínua 1 = Variável Quantitativa Discreta Se Tipo = 1 0 = Variável Qualitativa Nominal 1 = Variável Qualitativa Ordinal	Sim
Casas decimais	Inteiro	1	Se Tipo = 0 e Subtipo = 0: 0 = “” 1 = “0.0” 2 = “0.00” 3 = “0.000” 4 = “0.0000” 5 = “0.00000”	Sim

Exemplo (Distribuição contendo 3 variáveis e com até 5 ocorrências cada):
PESO
0
0
0
COR
1
0
-1
CLASSE SOCIAL
1
0
-1

-1
-1
-1

-1
-1
-1

-1

-1
-1

-1
-1
-1

-1
-1
-1

-1
-1
-1

-1
-1
-1

***FACILITADOR:** Deverão ser preenchidos de acordo com a estrutura de cada arquivo descrito neste documento.
Ocorrerá a partir da linha imediatamente posterior a última linha do registro CONFIGURAÇÃO. Esse registro recebe os possíveis resultados que uma variável do tipo Qualitativa pode receber, a fim de facilitar o processo de entrada de dados.
Método: Serão gerados 10 grupos de registro do tipo FACILITADOR, na mesma sequência das variáveis no registro do tipo CONFIGURAÇÃO. Na mesma ordem das variáveis informadas no registro do tipo CONFIGURAÇÃO. A 1ª linha receberá a quantidade de registros facilitadores de lançamento e as subseqüentes os valores destes registros.
Registro do tipo FACILITADOR para todos os arquivos

Descrição do Campo	**Tipo**	**Tamanho (em bytes)**	**Observações**	**Campo Obrigatório**
Número de Possíveis Ocorrências	Inteiro	1		Sim
Ocorrência	Texto	25	Se Número de Possíveis Ocorrências = 0, este registro deve ser suprimido	Sim

Exemplo (Distribuição contendo 3 variáveis e com até 5 ocorrências cada):
0
4
AZUL
PRETO
VERDE
VERMELHO
3
ALTA
BAIXA
MEDIA

0
0
0
0
0
0
0

1.2. CONVENÇÃO

Deverá ser utilizada a seguinte convenção para a coluna Tipo que consta das tabelas de descrição dos layouts do arquivo:

* Tipo Texto: Deverá ser alinhado à esquerda, podendo ser preenchido com espaçamentos em branco (Com exceção quando se tratar de variáveis quantitativas. Neste caso, somente receberá valores numéricos, com o separador de decimais [,]);

* Tipo Inteiro: Numérico sem casas decimais, não podendo ter caracteres especiais, tais como / . , ( ) $ “” + - etc.;

No caso de valores positivos:

Exemplo: R$ 120,00 deverá ser informado com o formato 120,00
R$ 1.120,00 deverá ser informado com o formato 1120,00

No caso de valores negativos: O sinal de menos deve vir imediatamente antes do conjunto de caracteres numéricos, ou seja, não deve(m) existir espaço(s) entre o sinal e o conjunto de caracteres.

Exemplo: R$ -120,00, deverá ser informado com o formato: -120,00

No caso de não haver dados para um determinado campo, este poderá ser deixado sem informação.

ANEXO 1 - Parecer do CEP-Unigranrio ao Projeto de Pesquisa

UNIVERSIDADE DO GRANDE RIO PROFESSOR JOSÉ DE SOUZA HERDY - UNIGRANRIO

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: ENSINO DA ESTATÍSTICA: FERRAMENTA COMPUTACIONAL COMO PROPOSTA PARA O PROCESSO PEDAGÓGICO
Pesquisador: JOSE CARLOS COELHO SARAIVA
Área Temática:
Versão: 1
CAAE: 38722114.5.0000.5283
Instituição Proponente: UNIVERSIDADE UNIGRANRIO
Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 907.629
Data da Relatoria: 09/12/2014

Apresentação do Projeto:
Pesquisa tem como foco principal encontrar uma forma de favorecer, por intermédio de recursos tecnológicos, o processo de ensino aprendizagem de Estatística na educação básica, verificando, inclusive, a possibilidade de uma extensão e/ou aprofundamento dos conteúdos aplicados. Através de estudos estatísticos junto a profissionais de ensino, pretende-se mapear o nível de introdução dos conteúdos em sala de aula, bem como as principais dificuldades e expectativas dos alunos em relação a estes conteúdos. Pretendemos desenvolver uma ferramenta computacional com base os resultados dessas análises. Para a validação de nosso produto educacional faremos a aplicação desta ferramenta em uma turma do ensino básico como forma de testagem, avaliação e reestruturação de nosso produto. Dentre os referenciais teóricos que destacaremos para nosso processo de construção e análise da pesquisa, temos como norte os Parâmetros Curriculares Nacionais

Objetivo da Pesquisa:
Objetivo Primário:

Desenvolver uma ferramenta (software) computacional voltada para o ensino da estatística, com formato e metodologia definidos no decorrer da pesquisa.

Endereço: Rua Prof. José de Souza Herdy, 1160
Bairro: 25 de Agosto **CEP:** 25.071-202
UF: RJ **Município:** DUQUE DE CAXIAS
Telefone: (21)2672-7733 **Fax:** (21)2672-7733 **E-mail:** cep@unigranrio.com.br

Objetivo Secundário:

Mapear o nível de introdução dos conteúdos em sala de aula, bem como as principais dificuldades e expectativas. Com base nos resultados desses resultados, determinar-se-ão o formato, os conteúdos e as metodologias a serem aplicadas.

Avaliação dos Riscos e Benefícios:

Riscos:

A pesquisa não apresenta riscos, se houver uma orientação metodológica e acompanhamento criterioso das etapas do estudo, principalmente na fase de coleta de dados.

Benefícios:

Contribuir com o processo de ensino-aprendizagem dos conteúdos de estatística, não limitando sua aplicação apenas a educação básica.

Comentários e Considerações sobre a Pesquisa:

Identificar se é possível contribuir com o processo de ensino-aprendizagem dos conteúdos da estatística descritiva. Para tal, temos a proposição de desenvolver uma ferramenta computacional (software) que será testada em sala da aula, durante a aplicação dos conteúdos da disciplina descrita em tela. Haverá orientação e intervenção do pesquisador a fim de identificar os pontos positivos e, porventura, negativos nesta implementação. O grupo testado será formado por alunos do curso de Ciências Contábeis da Universidade do Grande Rio. A pesquisa apresenta cronograma e orçamento detalhado.

Considerações sobre os Termos de apresentação obrigatória:

Apresenta Termo de Consentimento Livre Esclarecido e o Termo de Risco Confidencialidade para uso de arquivos, registro e similares.

Recomendações:

Recomendo que o projeto apresentado ao CEP deverá ter seus resultados divulgados em meios públicos,através de artigos e textos.

Conclusões ou Pendências e Lista de Inadequações:

A pesquisa é importante e deve ser desenvolvida. Considero o projeto aprovado sem pendências.

Situação do Parecer:

Aprovado

UNIVERSIDADE DO GRANDE RIO PROFESSOR JOSÉ DE SOUZA HERDY - UNIGRANRIO

Continuação do Parecer: 907.629

Necessita Apreciação da CONEP:
Não

Considerações Finais a critério do CEP:
Prezado Pesquisador,

O Comitê de ética em Pesquisa da Unigranrio atendendo o previsto na Resolução nº 466/12 do CNS/MS aprovou o referido projeto na reunião ocorrida em 10 de dezembro de 2014. Caso o (a) pesquisador (a) altere a pesquisa é necessário que o projeto retorne ao Sistema Plataforma Brasil para uma futura avaliação e emissão de novo parecer. Lembramos que o (a) pesquisador (a) deverá encaminhar o relatório da pesquisa após a sua conclusão, como um compromisso junto a esta Instituição e o sistema da Plataforma Brasil.

Cordialmente,
CEP/Unigranrio.

DUQUE DE CAXIAS, 10 de Dezembro de 2014

Assinado por:
Renato Cerqueira Zambrotti
(Coordenador)

Endereço: Rua Prof. José de Souza Herdy, 1160
Bairro: 25 de Agosto **CEP:** 25.071-202
UF: RJ **Município:** DUQUE DE CAXIAS
Telefone: (21)2672-7733 **Fax:** (21)2672-7733 **E-mail:** cep@unigranrio.com.br

ANEXO 2

PORTARIA NORMATIVA - MESTRADO PROFISSIONAL

Portaria Normativa Nº 17, de 28 de dezembro de 2009 - dispõe sobre o Mestrado Profissional no âmbito da Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES.

O **Ministro de Estado da Educação**, no uso de suas atribuições,
Considerando a necessidade de estimular a formação de mestres profissionais habilitados para desenvolver atividades e trabalhos técnico-científicos em temas de interesse público;
Considerando a necessidade de identificar potencialidades para atuação local, regional, nacional e internacional por órgãos públicos e privados, empresas, cooperativas e organizações não-governamentais, individual ou coletivamente organizadas;
Considerando a necessidade de atender, particularmente nas áreas mais diretamente vinculadas ao mundo do trabalho e ao sistema produtivo, a demanda de profissionais altamente qualificados;
Considerando as possibilidades a serem exploradas em áreas de demanda latente por formação de recursos humanos em cursos de pós-graduação stricto sensu com vistas ao desenvolvimento sócioeconômico e cultural do País;
Considerando a necessidade de capacitação e treinamento de pesquisadores e profissionais destinados a aumentar o potencial interno de geração, difusão e utilização de conhecimentos científicos no processo produtivo de bens e serviços em consonância com a política industrial brasileira;
Considerando a natureza e especificidade do conhecimento científico e tecnológico a ser produzido e reproduzido;
Considerando a relevância social, científica e tecnológica dos processos de formação profissional avançada, bem como o necessário estreitamento das relações entre as universidades e o setor produtivo; e, finalmente,
Considerando o disposto na Lei nº 9.394, de 20 de dezembro de 1996, e as deliberações do Conselho Técnico-Científico da Educação Superior - CTC-ES e as deliberações do Conselho Superior da CAPES,
Resolve:

Art. 1º A Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES regulará a oferta de programas de mestrado profissional mediante chamadas públicas e avaliará os cursos oferecidos, na forma desta Portaria e de sua regulamentação própria.

Art. 2º O título de mestre obtido nos cursos de mestrado profissional reconhecidos e avaliados pela CAPES e credenciados pelo Conselho Nacional de Educação - CNE e validados pelo Ministro de Estado da Educação tem validade nacional.

Art. 3º O mestrado profissional é definido como modalidade de formação pós-graduada stricto sensu que possibilita:

I. a capacitação de pessoal para a prática profissional avançada e transformadora de procedimentos e processos aplicados, por meio da incorporação do método científico, habilitando o profissional para atuar em atividades técnico-científicas e de inovação;
II. a formação de profissionais qualificados pela apropriação e aplicação do conhecimento embasado no rigor metodológico e nos fundamentos científicos;
III. a incorporação e atualização permanentes dos avanços da ciência e das tecnologias, bem como a capacitação para aplicar os mesmos, tendo como foco a gestão, a produção técnico-científica na pesquisa aplicada e a proposição de inovações e aperfeiçoamentos tecnológicos para a solução de problemas específicos.

Art. 4º São objetivos do mestrado profissional:

I. capacitar profissionais qualificados para o exercício da prática profissional avançada e transformadora de procedimentos, visando atender demandas sociais, organizacionais ou profissionais e do mercado de trabalho;
II. transferir conhecimento para a sociedade, atendendo demandas específicas e de arranjos produtivos com vistas ao desenvolvimento nacional, regional ou local;
III. promover a articulação integrada da formação profissional com entidades demandantes de naturezas diversas, visando melhorar a eficácia e a eficiência das

organizações públicas e privadas por meio da solução de problemas e geração e aplicação de processos de inovação apropriados;

IV. contribuir para agregar competitividade e aumentar a produtividade em empresas, organizações públicas e privadas.

Parágrafo único. No caso da área da saúde, qualificam-se para o oferecimento do mestrado profissional os programas de residência médica ou multiprofissional devidamente credenciados e que atendam aos requisitos estabelecidos em edital específico.

Art. 5º Os cursos de mestrado profissional a serem submetidos à CAPES poderão ser propostos por universidades, instituições de ensino e centros de pesquisa, públicos e privados, inclusive em forma de consórcio, atendendo necessária e obrigatoriamente aos requisitos de qualidade fixados pela CAPES e, em particular, demonstrando experiência na prática do ensino e da pesquisa aplicada.

Parágrafo único. A oferta de cursos com vistas à formação no Mestrado Profissional terá como ênfase os princípios de aplicabilidade técnica, flexibilidade operacional e organicidade do conhecimento técnico-científico, visando o treinamento de pessoal pela exposição dos alunos aos processos da utilização aplicada dos conhecimentos e o exercício da inovação, visando a valorização da experiência profissional.

Art. 6º As propostas de cursos de mestrado profissional serão apresentadas à CAPES mediante preenchimento por meio eletrônico via internet do Aplicativo para Cursos Novos - Mestrado Profissional (APCN-MP), em resposta a editais de chamadas públicas ou por iniciativa própria das instituições, dentro de cronograma estabelecido periodicamente pela agência.

Art. 7º A proposta de Mestrado Profissional deverá, necessária e obrigatoriamente:

I. apresentar estrutura curricular objetiva, coerente com as finalidades do curso e consistentemente vinculada à sua especificidade, enfatizando a articulação entre conhecimento atualizado, domínio da metodologia pertinente e aplicação orientada para o campo de atuação profissional

II. possibilitar a inclusão, quando justificável, de atividades curriculares estruturadas das áreas das ciências sociais aplicadas correlatas com o curso, tais como legislação, comunicação, administração e gestão, ciência política e ética;

III. conciliar a proposta ao perfil peculiar dos candidatos ao curso;

IV. apresentar, de forma equilibrada, corpo docente integrado por doutores, profissionais e técnicos com experiência em pesquisa aplicada ao desenvolvimento e à inovação;

V. apresentar normas bem definidas de seleção dos docentes que serão responsáveis pela orientação dos alunos;

VI. comprovar carga horária docente e condições de trabalho compatíveis com as necessidades do curso, admitido o regime de dedicação parcial;

VII. prever a defesa apropriada na etapa de conclusão do curso, possibilitando ao aluno demonstrar domínio do objeto de estudo com plena capacidade de expressar-se sobre o tema;

VIII. prever a exigência de apresentação de trabalho de conclusão final do curso.

§ 1º O corpo docente do curso deve ser altamente qualificado, conforme demonstrado pela produção intelectual constituída por publicações específicas, produção artística ou produção técnico científica, ou ainda por reconhecida experiência profissional, conforme o caso.

§ 2º A qualificação docente deve ser compatível com a área e a proposta do curso, de modo a oferecer adequadas oportunidades de treinamento para os estudantes e proporcionar temas relevantes para o seu trabalho de mestrado.

§ 3º O trabalho de conclusão final do curso poderá ser apresentado em diferentes formatos, tais como dissertação, revisão sistemática e aprofundada da literatura, artigo, patente, registros de propriedade intelectual, projetos técnicos, publicações tecnológicas; desenvolvimento de aplicativos, de materiais didáticos e instrucionais e de produtos, processos e técnicas; produção de programas de mídia, editoria, composições, concertos, relatórios finais de pesquisa, softwares, estudos de caso, relatório técnico com regras de sigilo, manual de operação técnica, protocolo experimental ou de aplicação em serviços, proposta de intervenção em procedimentos clínicos ou de serviço pertinente, projeto de aplicação ou adequação

tecnológica, protótipos para desenvolvimento ou produção de instrumentos, equipamentos e kits, projetos de inovação tecnológica, produção artística; sem prejuízo de outros formatos, de acordo com a natureza da área e a finalidade do curso, desde que previamente propostos e aprovados pela CAPES.

§ 4º Para atender situações relevantes, específicas e esporádicas, serão admitidas proposições de cursos com duração temporária determinada.

Art. 8º O desempenho dos cursos de mestrado profissional será acompanhado anualmente e terá avaliação com atribuição de conceito a cada três anos pela CAPES.

§ 1º O credenciamento dos cursos de mestrado profissional pelo CNE terá validade de três anos, sendo renovado a cada avaliação trienal positiva pela CAPES.

§ 2º Quando da avaliação de proposta de curso novo, ou de sua avaliação trienal, o Mestrado Profissional receberá da CAPES graus de qualificação variando dos conceitos 1 a 5, sendo o conceito 3 o mínimo para aprovação.

§ 3º A proposta de curso avaliada seguirá para o CNE para aprovação e credenciamento e posterior autorização do MEC para o funcionamento do curso.

Art. 9º A análise de propostas de cursos, bem como o acompanhamento periódico e a avaliação trienal dos cursos de mestrado profissional, serão feitas pela CAPES utilizando fichas de avaliação próprias e diferenciadas.

Parágrafo único. A avaliação será feita por comissões específicas, compostas com participação equilibrada de docentes-doutores, profissionais e técnicos dos setores específicos, reconhecidamente qualificados para o adequado exercício de tais tarefas.

Art. 10 Em complemento ao disposto no art. 7º, constituem parâmetros para o acompanhamento e a avaliação trienal dos cursos os seguintes indicadores, relativos à produção do corpo docente e, em especial, do conjunto docentes-orientadores-alunos:

I. produção intelectual e técnica pertinente à área, regular nos últimos três anos e bem distribuída entre os docentes, contemplando:
   a. artigos originais, artigos de revisão da literatura e publicações tecnológicas;
   b. patentes e registros de propriedade intelectual e de softwares, inclusive depósito de software livre em repositório reconhecido ou obtenção de licenças alternativas ou flexíveis para produção intelectual, desde que demonstrado o uso pela comunidade acadêmica ou pelo setor produtivo;
   c. desenvolvimento de aplicativos e materiais didáticos e instrucionais e de produtos, processos e técnicas;
   d. produção de programas de mídia;
   e. editoria;
   f. composições e concertos;
   g. relatórios conclusivos de pesquisa aplicada;
   h. manuais de operação técnica, protocolo experimental ou de aplicação ou adequação tecnológica;
   i. protótipos para desenvolvimento de equipamentos e produtos específicos;
   j. projetos de inovação tecnológica;
   k. produção artística;
   l. outros formatos, de acordo com a natureza da área e a finalidade do curso, a critério da CAPES;

II. informações sobre o destino dos egressos do curso, empregabilidade e trajetória profissional;

III. informações, recomendações e observações que constem de relatórios e pareceres das comissões examinadoras de avaliação dos trabalhos de conclusão do mestrado dos estudantes;

IV. dimensão e eficácia dos processos de interação com organizações, empresas e instituições da área de especialização e atuação do curso;

V. informações de outra natureza, além daquelas constantes nos relatórios anuais, sobre a produção técnico-científica, produção intelectual e a atividade acadêmica do curso, quando for o caso.

Art. 11 Salvo em áreas excepcionalmente priorizadas, o mestrado profissional não pressupõe, a qualquer título, a concessão de bolsas de estudos pela CAPES.

Art. 12 Os cursos de mestrado profissional já existentes devem providenciar, ao longo do triênio, as mudanças e atualizações que se mostrarem necessárias para a devida adequação ao disposto nesta Portaria.

Art. 13 Fica revogada a Portaria nº 7 de 22 de junho de 2009.

Art. 14 Esta Portaria entra em vigor na data da sua publicação.

Fernando Haddad.
Publicado no Diário Oficial Nº 248,
em 29 de dezembro de 2009.