



Universidade do Estado do Rio de Janeiro

Centro de Tecnologia e Ciências

Instituto de Matemática e Estatística

Rafael Teixeira Silva

**Interpretando dados do cotidiano: o ensino de Estatística
na educação básica**

Rio de Janeiro

2015

Rafael Teixeira Silva

**Interpretando dados do cotidiano: o ensino de Estatística
na educação básica**



Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-graduação de Mestrado Profissional em Matemática em Rede Nacional (PROFMAT), da Universidade do Estado do Rio de Janeiro.

Orientador: Prof. Dr. Francisco Mattos

Rio de Janeiro

2015

CATALOGAÇÃO NA FONTE
UERJ / REDE SIRIUS / BIBLIOTECA CTC-A

T266 Silva, Rafael Teixeira.
Interpretando dados do cotidiano: o ensino de Estatística na educação básica
/ Rafael Teixeira Silva. – 2015.
89 f. : il.

Orientador: Francisco Mattos.
Dissertação (Mestrado Profissional em Matemática em Rede Nacional /
PROFMAT) - Universidade do Estado do Rio de Janeiro, Instituto de
Matemática e Estatística.

1. Estatística – Estudo e ensino. 2. Educação básica. 3. Práticas de ensino. I.
Mattos, Francisco. II. Universidade do Estado do Rio de Janeiro. Instituto de
Matemática e Estatística. III. Título.

CDU 519.2:37

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta
dissertação, desde que citada a fonte.

Assinatura

Data

Rafael Teixeira Silva

**Interpretando dados do cotidiano: o ensino de Estatística
na educação básica**

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-graduação de Mestrado Profissional em Matemática em Rede Nacional (PROFMAT), da Universidade do Estado do Rio de Janeiro.

Aprovada em 06 de janeiro de 2015.

Banca Examinadora:

Prof. Dr. Francisco Roberto Pinto Mattos

Universidade do Estado do Rio de Janeiro

Prof. Dr. Augusto César de Castro Barbosa

Universidade do Estado do Rio de Janeiro

Prof. Dr. Rafael Garcia Barbastefano

Centro Federal de Educação Tecnológica Celso Suckow da Fonseca

Rio de Janeiro

2015

DEDICATÓRIA

À minha filha Clara, um presente recebido durante o desenvolvimento do mestrado.

AGRADECIMENTOS

Ao Prof. Dr. Francisco Mattos pelo convívio, apoio, incentivo, compreensão que tornaram possível a conclusão desta dissertação.

A todos os professores do PROFMAT pelas contribuições dadas para a minha formação acadêmica.

Aos meus companheiros do curso pelo convívio no compartilhar de ideias, dúvidas, sugestões na construção de novos conhecimentos.

A todos os meus amigos pelo apoio e incentivo.

À Marcella que me deu um lindo presente, minha filha, durante o desenvolvimento do mestrado.

Aos meus pais Sergio e Regina e minha irmã Carolina pelo incentivo durante todo o desenvolvimento deste meu mestrado.

“(…) resta, então, a tarefa mais difícil: colocar a cultura científica em estado de mobilização permanente, substituir o saber fechado e estático por um conhecimento aberto e dinâmico, dialetizar todas as variáveis experimentais, oferecer enfim à razão motivos para evoluir”.

Bachelard

RESUMO

SILVA, Rafael Teixeira. *Interpretando dados do cotidiano: o ensino de Estatística na educação básica*. 2015. 89 f. Dissertação (Mestrado Profissional em Matemática em Rede Nacional / PROFMAT) - Universidade do Estado do Rio de Janeiro, Instituto de Matemática e Estatística, Rio de Janeiro, 2015.

A estatística é uma ciência com seus conceitos e métodos de coleta, organização e análise de informações que fazem parte dos currículos escolares da educação básica, na prática dos professores de matemática, para que os alunos compreendam, analisem e formem opinião crítica em relação às questões econômicas e sociais. A presente pesquisa buscou refletir sobre as práticas pedagógicas do professor de matemática no ensino de estatística no ensino médio, tendo como base as orientações para o ensino de estatísticas nas propostas dos Parâmetros Curriculares do Ensino Médio, as contribuições da aprendizagem significativa no ensino de estatística, com o uso das tecnologias na educação, através da proposta de planos de trabalho que abordem os conteúdos do ensino de estatística e a utilização do *software livre Calc*. Em relação aos caminhos metodológicos foi realizada uma pesquisa bibliográfica, utilizando o método de abordagem dedutivo, através de documentação indireta tendo como fonte de pesquisa os trabalhos científicos com foco no ensino e na aprendizagem da Estatística e da Probabilidade na Educação Básica. O desenvolvimento desta pesquisa possibilitou evidenciar caminhos metodológicos a serem desenvolvidos por professores de matemática na educação básica que contribuam na interpretação de dados do cotidiano a partir de análise de tabelas, análise de gráficos, medidas de posição, medidas de dispersão e linhas de tendência, utilizando como ferramentas as Tecnologias da Informação e Comunicação tendo como fundamentação teórica as contribuições de David Ausubel o conceito de aprendizagem significativa.

Palavras-chave: Práticas pedagógicas. Ensino. Estatística. Professor de matemática.

ABSTRACT

SILVA, Rafael Teixeira. Interpreting data from the everyday: the teaching of Statistics in basic education. 2015. 89 f. Dissertação (Mestrado Profissional em Matemática em Rede Nacional / PROFMAT), Universidade do Estado do Rio de Janeiro, Instituto de Matemática e Estatística, Rio de Janeiro, 2015.

Statistics is a science with its concepts and methods of data collection, organization and analysis of information that are part of the curricula of basic education in the practice of mathematics teachers, so that students understand, analyze, critique and form opinion on the issues economic and social. This research sought to reflect on teaching practices of mathematics teachers in the teaching of statistics in high school, based on the guidelines for the teaching of statistics in the proposals of the High School Curriculum Standards, the contributions of significant learning in teaching statistics with the use of technology in education, through the proposed work plans that address the content of teaching and the use of statistical software *livreCalc*. Regarding methodological approaches a literature search was conducted using the method of deductive approach, using indirect documentation as source of scientific research with a focus on teaching and learning of Statistics and Probability in Elementary Education jobs. The development of this research highlight possible methodological approaches to be developed by mathematics teachers in basic education which will help in the interpretation of the data daily from analysis of charts, graphical analysis of the position measurements, measures of dispersion and trend lines using as the tools of Information and Communication Technologies as having theoretical foundation contributions of David Ausubel's concept of meaningful learning.

Keywords: Pedagogic practices. Education. Statistics. Mathematics teacher.

LISTA DE FIGURAS

Figura 1	– Apresentação do Software LibreOffice.....	39
Figura 2	– Exemplo de Regressão Linear Simples.....	40
Figura 3	– Interface do Software.....	41
Figura 4	– Construindo a Planilha.....	42
Figura 5	– Fórmula MÉDIA.....	43
Figura 6	– Cálculo da Mediana.....	44
Figura 7	– Cálculo da Moda.....	45
Figura 8	– Medidas de Dispersão.....	46
Figura 9	– Variância Populacional.....	47
Figura 10	– Cálculo da Variância Amostral.....	48
Figura 11	– Cálculo do Coeficiente de Variação Amostral.....	49
Figura 12	– Observando o Desvio Padrão Amostral.....	50
Figura 13	– Cálculo do Coeficiente de Variação Amostral.....	51
Figura 14	– Cálculo do Coeficiente de Variação Populacional.....	52
Figura 15	– Inserir funções no LibreOffice Calc.....	53
Figura 16	– Categorias Estatísticas.....	54
Figura 17	– Cálculo do Desvio Padrão com base em uma Amostra.....	55
Figura 18	– Função dando Erro.....	56
Figura 19	– Algoritmo para a Geração da Planilha.....	57
Figura 20	– Efetuação de Gráfico e a Linha de Tendência.....	59
Figura 21	– Tipo de Gráfico.....	60
Figura 22	– Intervalo de Dados 1.....	60
Figura 23	– Intervalo de Dados 2.....	61
Figura 24	– Intervalo de Dados 3.....	61
Figura 25	– Série de Dados.....	62
Figura 26	– Elementos do Gráfico.....	62
Figura 27	– Pontos Gráficos.....	63
Figura 28	– Linha de Tendência para a Série de Dados.....	64
Figura 29	– Cálculo das Funções 1.....	65
Figura 30	– Cálculo das Funções 2.....	66

Figura 31 –	Dados do IBGE quanto o Grau de Instrução e Rendimento das Pessoas....	68
Figura 32 –	Nível de Instrução e Renda Nominal 1.....	74
Figura 33 –	Nível de Instrução e Renda Nominal 2.....	75

LISTA DE TABELAS

Tabela 1 –	População do Brasil – 2007.....	29
Tabela 2 –	Alunos e Idades (Prestaram do Vestibular).....	30
Tabela 3 –	Medidas de Dispersão.....	31
Tabela 4 –	Consumo Mensal no Cartão de Crédito.....	37
Tabela 5 –	Estimativas do Modelo.....	38
Tabela 6 –	Nível de Instrução e Quantidade de Pessoas.....	69
Tabela 7 –	Rendimento Mensal Domiciliar Per Capita.....	70
Tabela 8 –	Média da Classe de Rendimentos Nominais.....	71
Tabela 9 –	Nível de Instrução e Valor esperado Salário Mínimo.....	72
Tabela 10 –	Nível de Instrução e Valor esperado Salário Mínimo.....	81
Tabela 11 –	Nível de Instrução e Valor esperado Salário Mínimo.....	81
Tabela 12 –	Nível de Instrução e Valor esperado Salário Mínimo.....	82

LISTA DE GRÁFICOS

Gráfico 1 –	Nível de Instrução.....	69
Gráfico 2 –	Renda Nominal mensal Domiciliar Per Capita.....	71
Gráfico 3 –	Dispersão.....	73
Gráfico 4 –	Linha de Tendência.....	75

SUMÁRIO

	INTRODUÇÃO.....	13
1	PARÂMETROS CURRICULARES NACIONAIS ENSINO MÉDIO.....	15
2	O USO DA TECNOLOGIA E A APRENDIZAGEM SIGNIFICATIVA.	20
2.1	O Uso das Tecnologias da Informação e Comunicação na Educação.....	20
2.2	Contribuições do Computador e da Rede para a	23
	Aprendizagem Aprendizagem Significativa.....	24
3	CONTEUDOS MATEMÁTICOS EXPLORADOS.....	27
3.1	Medidas de posição.....	28
3.1.1	<u>Média aritmética.....</u>	28
3.1.2	<u>Mediana.....</u>	28
3.1.3	<u>Moda.....</u>	29
3.1.4	<u>Exemplos de medidas de posição.....</u>	29
3.2	Medidas de dispersão.....	31
3.2.1	<u>Amplitude.....</u>	32
3.2.2	<u>Variância.....</u>	32
3.2.3	<u>Desvio padrão.....</u>	33
3.2.4	<u>Coeficiente de variação.....</u>	33
3.3	Método dos Mínimos Quadrados Ordinários.....	34
3.3.1	<u>Coeficiente de Determinação.....</u>	36
3.3.2	<u>Exemplo de mínimos quadrados.....</u>	37
4	CONSTRUÇÃO DE PLANILHAS.....	39
5	PLANOS DE TRABALHO COM USO DE SOFTWARE NO ENSINO	
	DE ESTATÍSTICA.....	77
5.1	Introdução.....	77
5.2	Desenvolvimento.....	77
5.2.1	<u>Atividade 1.....</u>	78
5.2.2	<u>Atividade 2.....</u>	78
5.2.3	<u>Atividade 3.....</u>	78
5.2.4	<u>Atividade 4.....</u>	79
5.3	Primeiro plano de aula, utilizando as três primeiras atividades.....	79
5.4	Segundo plano de aula, utilizando as quatro atividades.....	81
	CONCLUSÕES.....	83

REFERÊNCIAS	85
--------------------------	-----------

INTRODUÇÃO

A Estatística é uma das ciências que está cada vez mais presente no nosso dia a dia através das tecnologias da informação e comunicação temos acesso a informações estatísticas na área da educação, saúde, segurança, etc., pois na maioria das áreas esta ciência está presente.

De acordo com Lopes (2010, p. 47):

A presença constante da Estatística no mundo atual tornou-a uma realidade na vida dos cidadãos, levando-o à necessidade de ensinar estatística a um número de pessoas cada vez maior. Consequentemente, nas últimas décadas a maioria dos países introduziu, nos seus currículos de matemática, conteúdos de estatística, probabilidade e combinatória desde o início da escolaridade.

Neste cenário, o cidadão brasileiro precisa interagir, interpretar e inserir estas informações estatísticas, sendo capaz de analisar estes dados criticamente para que possa atuar no mundo em que vive de forma mais participativa, consciente e coerente.

Na educação, o processo de aprendizagem de conhecimentos estatísticos cada vez mais passa a fazer parte dos currículos escolares, na prática pedagógica dos professores de matemática para que a população cada vez mais tenha condições de refletir, analisar e formar opinião crítica em relação às questões econômicas e sociais.

Segundo Lopes (2010, p. 48):

Em um levantamento realizado em 1986 pelo International Statistical Institute (ISI), constatasse, nos relatórios enviados pelos diversos países, uma insatisfação pelo ensino da Estatística, em particular nas escolas dos anos elementares, donde o seu ensino tem sido ignorado.

Atualmente, aproximadamente duas décadas após este levantamento, percebe-se que, embora a inserção da Estatística e da Probabilidade seja recomendada nas propostas curriculares de Matemática, na maioria dos países do mundo, ainda não tem sido prioridade na escola, nem nos programas de formação inicial e contínua de professores que ensinam Matemática

A presente pesquisa procurou refletir sobre a questão de: “Como o ensino de estatística pode contribuir para que os alunos do ensino médio compreendam mais suas realidades, para uma ação crítica e reflexiva que leve a construção da cidadania?”

De acordo com Lopes (2010, p. 49-50):

Os documentos curriculares e, também, alguns pesquisadores têm justificado a pertinência da Estatística no currículo de Matemática da Educação Básica, ao se considerar que a complexidade da sociedade obrigou a quantificar muito da informação gerada. A Estatística, com seus conceitos e métodos para coletar, organizar e analisar informações diversas, tem se revelado um poderoso aliado nesse desafio que é transformar a informação bruta em dados que permitem ler e compreender uma realidade. Talvez por isso, ela, a Estatística, tenha-se tornado uma presença constante no cotidiano das pessoas, gerando um amplo consenso em torno

da idéia segundo a qual a literacia estatística deva ser uma prioridade da sociedade moderna, ou seja, de uma cidadania com responsabilidade social. Assim, é preciso que a Estatística tenha destaque nas aulas de Matemática.

Tendo como objetivo:

- Identificar nos Parâmetros Curriculares do Ensino Médio as orientações para o ensino de estatística;

- Demonstrar o passo a passo para construção de uma planilha que calcule automaticamente estatísticas descritivas, regressões e façam gráficos.

- Propor planos de trabalho que abordem os conteúdos do ensino de estatística propostos pelos Parâmetros Curriculares do Ensino Médio e utilizem o *software* livre *Calc*.

Para o alcance destes objetivos foi realizada uma pesquisa bibliográfica, através de documentação indireta tendo como fonte de pesquisas trabalhos científicos com foco no ensino e na aprendizagem da Estatística e da Probabilidade na Educação Básica.

A estrutura desta pesquisa foi desenvolvida em seis capítulos.

No primeiro capítulo se apresentou as orientações dos Parâmetros Curriculares Nacionais para o Ensino Médio, para a área de Ciências da Natureza, Matemática e suas Tecnologias, onde estão contidos os conteúdos para o ensino de estatística.

O segundo capítulo abordou as contribuições do uso das Tecnologias da Informação e Comunicação para a educação e as contribuições da aprendizagem significativa no ensino de estatística.

O capítulo três demonstra os conteúdos matemáticos explorados no trabalho, com medidas de posição, dispersão e o Método dos Mínimos Quadrados Ordinários.

No quarto capítulo se abordou a construção de planilhas, no intuito de demonstrar o passo a passo para sua construção, a fim de que qualquer aluno ou professor possa construir.

O capítulo cinco apresenta abordagens metodológicas em forma de planos de trabalho com o uso do software no ensino de estatística, para o desenvolvimento de conceitos que levem o aluno a compreensão das características da sociedade em que vive e contribua na tomada de decisões em seu cotidiano onde estão presentes a variabilidade e a incerteza.

Por fim, o capítulo seis contém as conclusões do trabalho.

O desenvolvimento desta pesquisa possibilitou evidenciar caminhos metodológicos a serem desenvolvidos por professores de matemática na educação básica que contribuam na interpretação de dados do cotidiano a partir de análise de tabelas, análise de gráficos, medidas de posição, medidas de dispersão e linhas de tendência, utilizando como ferramentas as

Tecnologias da Informação e Comunicação tendo como fundamentação teórica as contribuições de David Ausubel o conceito de aprendizagem significativa

1 PARÂMETROS CURRICULARES NACIONAIS ENSINO MÉDIO

Este capítulo apresentará as orientações em relação ao ensino de estatística, presente nos Parâmetros Curriculares Nacionais para o Ensino Médio, documento este que tem como base os princípios presentes na Lei de Diretrizes e Bases da Educação Nacional nº 9.394/96, elaborado por especialistas e educadores de todo o Brasil para estimular e apoiar a reflexão sobre o currículo escolar, o planejamento das aulas e a prática pedagógica, fornecendo orientações para os professores na construção de novas abordagens e metodologias.

Os Parâmetros Curriculares foram organizados em áreas do conhecimento apresentadas em documento específico, a saber:

- Linguagens, Códigos e suas Tecnologias;
- Ciências da Natureza, Matemática e suas Tecnologias;
- Ciências Humanas e suas Tecnologias

A opção por área de conhecimento deve-se ao fato de procurar oferecer uma educação de base científica e tecnológica onde são incluídas as competências que os alunos deverão alcançar em cada área.

O ensino de estatística é abordado na área de Ciências da Natureza, Matemática e suas Tecnologias que deve, de acordo com o PCN (2000, p. 20):

[...] contemplar formas de apropriação e construção de sistemas de pensamento mais abstratos e ressignificados, que as trate como processo cumulativo de saber e de ruptura de consensos e pressupostos metodológicos. A aprendizagem de concepções científicas atualizadas do mundo físico e natural e o desenvolvimento de estratégias de trabalho centradas na solução de problemas é finalidade da área, de forma a aproximar o educando do trabalho de investigação científica e tecnológica, como atividades institucionalizadas de produção de conhecimentos, bens e serviços.

Os conteúdos de estatística inserido nesta área das Ciências da Natureza, Matemática e suas Tecnologias ao contemplar estas formas de apropriação do conhecimento científico contribuem para a compreensão e tratamento das informações divulgadas pelos meios de comunicação que utilizam dados estatísticos através de gráficos, tabelas, e representações possibilitando ao aluno coletar, organizar, interagir, interpretar e inferir com os dados.

Em relação aos conhecimentos matemáticos os Parâmetros Curriculares Nacionais para o Ensino Médio, apresentam o papel formativo e instrumental da Matemática.

Segundo o PCN (2000, p. 40):

Em seu papel formativo, a Matemática contribui para o desenvolvimento de processos de pensamento e a aquisição de atitudes, cuja utilidade e alcance transcendem o âmbito da própria Matemática, podendo formar no aluno a capacidade de resolver problemas genuínos, gerando hábitos de investigação, proporcionando confiança e desprendimento para analisar e enfrentar situações novas, propiciando a formação de uma visão ampla e científica da realidade, a percepção da beleza e da harmonia, o desenvolvimento da criatividade e de outras capacidades pessoais.

No que diz respeito ao caráter instrumental da Matemática no Ensino Médio, ela deve ser vista pelo aluno como um conjunto de técnicas e estratégias para serem aplicadas a outras áreas do conhecimento, assim como para a atividade profissional. Não se trata de os alunos possuírem muitas e sofisticadas estratégias, mas sim de desenvolverem a iniciativa e a segurança para adaptá-las a diferentes contextos, usando-as adequadamente no momento oportuno.

Nesse sentido, é preciso que o aluno perceba a Matemática como um sistema de códigos e regras que a tornam uma linguagem de comunicação de idéias e permite modelar a realidade e interpretá-la.

Assim, os números e a álgebra como sistemas de códigos, a geometria na leitura e interpretação do espaço, a estatística e a probabilidade na compreensão de fenômenos em universos finitos são subáreas da Matemática especialmente ligadas às aplicações.

O ensino de conteúdos de estatística faz parte de uma subárea da disciplina de Matemática. Segundo Cazorla (2002), antes dos Parâmetros Curriculares Nacionais, a estatística era um dos últimos conteúdos dos livros de matemática e com isso, na maioria das vezes, não era ensinado. Mas, esta situação vem sendo mudada com iniciativas conforme nos relata NOGUEIRA (2012) em seu artigo sobre: “A Inclusão da Estatística como Componente Curricular do Ensino Médio, nas Escolas Estaduais Baianas”, onde apresenta uma experiência vivenciada em uma escola estadual da Bahia sobre a proposta curricular do ensino de estatística ofertada na parte diversificada do currículo para alunos de 6 turmas de 2ª série do ensino médio com objetivo de discutir sobre a necessidade de um ensino mais significativo para o aluno, tendo como desafio tratá-lo na proposta curricular do ensino médio, pois de acordo com Nogueira (2012, p. 2):

Nossa justificativa está apoiada no fato de que a Estatística quando ensinada como um tópico da Matemática faz com que quase sempre seja priorizado os cálculos, as fórmulas, em detrimento do seu tratamento como um conteúdo mais abrangente, de cunho social e que deve possuir outras intencionalidades em seu tratamento.

O ensino de estatística foi desenvolvido através de uma disciplina, com carga horária de 40 h/anuais, inserida em um projeto denominado “A Estatística e os fenômenos sociais”, refletindo sobre a presença da estatística nas situações cotidianas, trabalhando a construção de gráficos e seus tipos (barras, setores e linhas).

Os alunos desta escola estadual da Bahia desenvolveram este projeto em sete etapas:

- 1ª etapa: Escolhas de temas pelos grupos;
- 2ª etapa: Pesquisa dos temas;
- 3ª etapa: Elaboração de um questionário de pesquisa;

- 4ª etapa: Aplicação dos questionários nas diversas turmas da escola;
- 5ª etapa: Análise dos dados coletados nos questionários;
- 6ª etapa: Construção dos gráficos;
- 7ª etapa: Culminância - Exposição dos trabalhos no pátio da Escola.

O desenvolvimento dessa experiência reflete segundo Nogueira (2012, p. 6):

Ao mesmo tempo mostra que a possibilidade da inclusão da Estatística como um componente curricular para o Ensino Médio só foi possível com a proposta da criação do Ensino Médio Inovador e sua implementação nas escolas do Estado da Bahia. As modificações na escola advindas dessa proposta contemplou exclusivamente a reorganização curricular do Ensino Fundamental e Médio.

Nesse contexto, a Escola aqui referida mostrou sua autonomia e optou pela escolha da Estatística como um componente da parte diversificada da área de Matemática e suas tecnologias, como forma de estar oportunizando sua clientela a se aprofundarem no estudo da Estatística, já que até então a disciplina só era vista como um conteúdo dentro da disciplina da Matemática, na 3ª série do Ensino Médio.

O exemplo sobre esta forma de trabalhar a estatística no ensino médio incentiva a construção de novos caminhos que possibilitam a aprendizagem do tratamento dessas informações e o desenvolvimento de pesquisas que propõe na reflexão sobre as práticas pedagógicas do professor de matemática no ensino de estatística ao ensino médio voltada para a interpretação dos dados do cotidiano.

Os Parâmetros Curriculares do Ensino Médio complementam e ampliam os conteúdos voltados para o Ensino Fundamental que iniciaram a aproximação do aluno desenvolvendo no 1º ciclo.

De acordo com o PCN (1997, p. 52):

- Leitura e interpretação de informações contidas em imagens.
- Coleta e organização de informações.
- Criação de registros pessoais para comunicação das informações coletadas.
- Exploração da função do número como código na organização de informações (linhas de ônibus, telefones, placas de carros, registros de identidade, bibliotecas, roupas, calçados).
- Interpretação e elaboração de listas, tabelas simples, de dupla entrada e gráficos de barra para comunicar a informação obtida.
- Produção de textos escritos a partir da interpretação de gráficos e tabelas.

Observa-se que nestas séries iniciais o foco do ensino de estatística está na coleta, na organização e na interpretação dos dados pelos alunos a partir de situações cotidianas vivenciadas por eles.

Deve-se inserir nestas séries, também, o conceito de probabilidade que deve ser trabalhado.

Segundo Lopes (2010, p. 56),

[...] em todos os níveis de ensino, devendo possibilitar aos alunos adquirir competência para tirar conclusões de experiências simples, na forma de jogos de

dados, roletas, moedas, em que a possibilidade de ganhar seja, ou não, a mesma para os diversos participantes, e a discussão gerada com base nos resultados obtidos permita aos alunos confrontarem-se com os termos de “certo, possível, impossível, provável”.

No 2º ciclo do Ensino Fundamental há orientação para os conteúdos do tratamento das informações onde está inserido o ensino de estatística:

De acordo com o PCN (1997, p. 62-63):

- Coleta, organização e descrição de dados.
- Leitura e interpretação de dados apresentados de maneira organizada (por meio de listas, tabelas, diagramas e gráficos) e construção dessas representações.
- Interpretação de dados apresentados por meio de tabelas e gráficos, para identificação de características previsíveis ou aleatórias de acontecimentos.
- Produção de textos escritos, a partir da interpretação de gráficos e tabelas, construção de gráficos e tabelas com base em informações contidas em textos jornalísticos, científicos ou outros.
- Obtenção e interpretação de média aritmética.
- Exploração da idéia de probabilidade em situações-problema simples, identificando sucessos possíveis, sucessos seguros e as situações de “sorte”.
- Utilização de informações dadas para avaliar probabilidades.
- Identificação das possíveis maneiras de combinar elementos de uma coleção e de contabilizá-las usando estratégias pessoais.

Verifica-se nestas séries que os alunos coletam, organizam e interpretam a informação, formulando conjecturas, conclusões e fundamentando-as.

Lopes colabora com os objetivos principais do estudo estatístico no ensino fundamental refletindo que.

Segundo Lopes (2010, p. 57):

À construção de tabelas de frequência e aos gráficos de barras, trabalhados nos ciclos anteriores, acrescentam-se os gráficos de setores, os polígonos de frequência e os pictogramas. A mediana surge como a medida de tendência central que aparece neste ciclo e, conjuntamente com a média e a moda, deve constituir os instrumentos para sintetizar e analisar a informação. Encontram-se, ainda, referências, nos programas curriculares, para a importância de os alunos compararem duas distribuições, como forma de discutir as várias medidas. O conceito de probabilidade é retomado como mencionado anteriormente, ou seja, com base em situações de jogo, os alunos devem familiarizar-se com resultados possíveis em uma situação aleatória.

Os Parâmetros Curriculares Nacionais para o Ensino Fundamental enfatizam além dos conteúdos, orientações para os professores abordarem estes conhecimentos levando em conta os interesses dos alunos, e fornecem sugestões para o desenvolvimento destes conteúdos:

De acordo com o PCN (1997, p. 85):

Assim, por exemplo, trabalhando com datas de aniversário pode-se propor a organização de uma lista com as informações sobre o assunto. Um critério para organizar essa lista de nomes precisa ser definido: ordem alfabética, meninos e meninas, etc. Quando a lista estiver pronta, as crianças a analisam e avaliam se as informações podem ser encontradas facilmente. O professor pode então propor a elaboração de uma outra forma de comunicar os aniversariantes de cada mês, orientando-as, por exemplo, a construir um gráfico de barras.

Na construção de gráficos é importante verificar se os alunos conseguem ler as informações neles representadas. Para tanto, deve-se solicitar que dêem sua interpretação sobre gráficos e propor que pensem em perguntas que possam ser respondidas a partir deles.

Outros dados referentes aos alunos, como peso, altura, nacionalidade dos avós, times de futebol de sua preferência, podem ser trabalhados e apresentados graficamente.

A construção de tabelas e gráficos que mostram o comportamento do tempo durante um período (dias ensolarados, chuvosos, nublados) e o acompanhamento das previsões do tempo pelos meios de comunicação indicam a possibilidade de se fazer algumas previsões, pela observação de acontecimentos.

No Ensino Médio são trabalhados novos conhecimentos possibilitando que este aluno continue aprendendo para o desenvolvimento de sua autonomia e sua capacidade de pesquisa, de competências para a produção e transmissão de conhecimentos.

De acordo com o PCN (2000, p. 42):

- compreender os conceitos, procedimentos e estratégias matemáticas que permitam a ele desenvolver estudos posteriores e adquirir uma formação científica geral;
- aplicar seus conhecimentos matemáticos a situações diversas, utilizando-os na interpretação da ciência, na atividade tecnológica e nas atividades cotidianas;
- analisar e valorizar informações provenientes de diferentes fontes, utilizando ferramentas matemáticas para formar uma opinião própria que lhe permita expressar-se criticamente sobre problemas da Matemática, das outras áreas do conhecimento e da atualidade;
- desenvolver as capacidades de raciocínio e resolução de problemas, de comunicação, bem como o espírito crítico e criativo;
- utilizar com confiança procedimentos de resolução de problemas para desenvolver a compreensão dos conceitos matemáticos;
- expressar-se oral, escrita e graficamente em situações matemáticas e valorizar a precisão da linguagem e as demonstrações em Matemática;
- estabelecer conexões entre diferentes temas matemáticos e entre esses temas e o conhecimento de outras áreas do currículo;
- reconhecer representações equivalentes de um mesmo conceito, relacionando procedimentos associados às diferentes representações;
- promover a realização pessoal mediante o sentimento de segurança em relação às suas capacidades matemáticas, o desenvolvimento de atitudes de autonomia e cooperação.

Verifica-se que não há um objetivo voltado para o ensino de estatística, mas enfatiza-se uma atenção ao desenvolvimento de valores, habilidades e atitudes em relação ao conhecimento e nas relações entre alunos e professores. Mas, em suas orientações finais para a área de Ciências da Natureza, Matemática e suas Tecnologias.

De acordo com o PCN (2000, p. 55):

Quando a escola promove uma condição de aprendizado em que há entusiasmo nos fazeres, paixão nos desafios, cooperação entre os participantes, ética nos procedimentos, esta construindo a cidadania em sua prática, dando as condições para a formação dos valores humanos fundamentais, que são centrais entre os objetivos da educação.

Aproveitando estas orientações em relação ao ensino de estatística conclui-se que de acordo com os Parâmetros Curriculares para o Ensino Médio é preciso desenvolver práticas

pedagógicas que partam do interesse dos alunos, motivando-os, desafiando-os, incentivando-os a serem cidadãos comprometidos com os valores humanos fundamentais.

2 O USO DA TECNOLOGIA E A APRENDIZAGEM SIGNIFICATIVA

O presente capítulo refletirá sobre o uso das Tecnologias da Informação e Comunicação, o uso de software, o conceito, os princípios e as estratégias de aprendizagem significativa para as aulas de estatística na educação básica.

2.1 O Uso das Tecnologias da Informação e Comunicação na Educação

A sociedade vem sofrendo mudanças em sua organização, na produção e consumo dos bens, na forma de diversão, de ensinar e aprender com o desenvolvimento das tecnologias. Segundo Moran “todos estamos reaprendendo a conhecer, a comunicar-nos, a ensinar; reaprendendo a integrar o humano e o tecnológico; a integrar o individual, o grupal e o social.” (2000, p.61).

Nas escolas as tecnologias da informação e comunicação estão cada vez mais presentes e, tem impactado tanto a prática do professor como os processos de aprendizagem.

Segundo Spritzer (2009, p. 123):

Os computadores constituem uma síntese de conhecimentos científicos e técnicos, sendo produtos do estudo sistemático de dispositivos físicos e a aplicação de uma série de inovações tecnológicas. Nesse sentido, os computadores, entendidos como a interação entre *hardware* e *software*, não se reduzem ao significado instrumental que frequentemente se atribui ao termo “tecnologia”.

A presença das tecnologias na escola tem impactado diversas maneiras de ensinar.

Segundo Moran (2000, p. 11):

Perdemos tempo demais, aprendemos muito pouco, desmotivamo-nos continuamente. Tanto professores como alunos temos a clara sensação de que muitas aulas convencionais estão ultrapassadas. Mas para onde mudar? Como ensinar e aprender em uma sociedade mais interconectada?

De acordo com Masetto (2000, p.133),

[...] não se valorizou adequadamente o uso da tecnologia visando a tornar um processo de ensino-aprendizagem mais eficiente e mais eficaz. Se nos perguntarmos o porquê deste fato, encontraremos em algumas situações, por exemplo, a convicção de que o papel da escola em todos os níveis é o de “educar” seus alunos – entendendo por “educação” transmitir um conjunto organizado e sistematizado de conhecimentos de diversas áreas, desde a alfabetização, passando por matemática, língua portuguesa, ciências, história, geografia, física, biologia e outras, até aqueles conhecimentos próprios de uma formação profissional nos cursos de graduação de uma faculdade – e exigir deles memorização das informações que lhes são passadas e sua reprodução nas provas e avaliações. Além disso, espera-se que a escola possa transmitir valores e padrões de comportamentos sociais próprios da sociedade em que se vive.

As Tecnologias da Informação e Comunicação tem sido utilizadas, mas permeadas por concepções ultrapassadas de aprendizagem, que insistem em velhos modelos de transmissão, memorização dos conhecimentos, não integrando o processo de aprendizagem à vida, com alunos passivos que não são sujeitos do seu processo de aprendizagem.

Dessa forma, há uma preocupação em definir quais são os caminhos para que esta mudança ocorra na busca de uma educação e um ensino de qualidade que contribua na integração de todas as dimensões da vida.

Segundo Moran (2000, p. 13):

Educar é colaborar para que professores e alunos – nas escolas e organizações – transformem suas vidas em processos permanentes de aprendizagem. É ajudar os alunos na construção da sua identidade, do seu caminho pessoal e profissional – do seu projeto de vida, no desenvolvimento de habilidades de compreensão, emoção e comunicação que lhes permitam encontrar seus espaços pessoais, sociais e profissionais e tornar-se cidadãos realizados e produtivos.

A partir desta concepção de educar verifica-se um grande desafio na construção de caminhos que levem professores e alunos nas instituições escolares estarem motivados, comprometidos e com condições intelectuais, emocionais, comunicacionais e éticas para este processo de aprendizagem contínuo.

As resistências em relação a esta concepção de educar estão presentes nas organizações, nos professores e nos alunos que acomodados na concepção tradicional de aprendizagem não querem alterar seus papéis.

As organizações com suas estruturas fechadas, autoritárias, sem infraestrutura adequada precisam de mudanças que as tornem inovadoras, dinâmicas, participativas, com projeto pedagógico coerente com a prática pedagógica.

Os professores desmotivados, cansados, trabalhando em vários lugares necessitam de boas condições de trabalho, formação contínua, remuneração digna, preparados intelectualmente e emocionalmente.

Os alunos acomodados em serem meros receptores, fazem o mínimo necessário para serem aprovados, pois valorizam mais o diploma do que a aprendizagem, desinteressados, indisciplinados quando deveriam ser motivados a desenvolver suas capacidades intelectuais e emocionais.

As Tecnologias da Informação e Comunicação contribuem nas mudanças necessárias para a melhoria da educação, porém não resolvem questões de fundo.

Segundo Moran (2000, p.15):

Nosso desafio maior é caminhar para um ensino e uma educação de qualidade, que integre todas as dimensões do ser humano. Para isso precisamos de pessoas que façam essa integração em si mesmas no que concerne aos aspectos sensorial, intelectual, emocional, ético e tecnológico, que transitem de forma fácil entre o pessoal e o social, que expressem nas suas palavras e ações que estão sempre evoluindo, mudando, avançando.

Diante deste desafio, é acrescentado que grandes dificuldades estão presentes.

De acordo com Moran (2000, p.16):

As mudanças demorarão mais do que alguns pensam, porque nos encontramos em processos desiguais de aprendizagem e evolução pessoal e social, Não temos muitas instituições e pessoas que desenvolvam formas avançadas de compreensão e integração que possam servir como referência. Predomina a média – a ênfase no intelectual, a separação entre teoria e prática.

Temos grandes dificuldades no gerenciamento emocional, tanto no pessoal como no organizacional, o que dificulta o aprendizado rápido. São poucos os modelos vivos de aprendizagem integradora, que junta teoria e prática, que aproxima o pensar do viver.

Em 2013 constata-se que este processo de mudanças, no âmbito educacional, continua lento em seus avanços, com experiências positivas pontuais, desde 2000, já tinha afirmação de que um dos impeditivos está no autoritarismo presente nas relações humanas.

Segundo Moran (2000, p. 16):

O autoritarismo da maior parte das relações humanas interpessoais, grupais e organizacionais espelha o estágio atrasado em que nos encontramos individual e coletivamente em termos de desenvolvimento humano, de equilíbrio pessoal, de amadurecimento social. E somente podemos educar para a autonomia, para a liberdade com processos fundamentalmente participativos, interativos, libertadores, que respeitem as diferenças, que incentivem, que apoiem, orientados por pessoas e organizações livres.

Relações mais participativas, interativas estão sendo incentivadas e construídas em alguns ambientes educacionais revertendo este quadro, pois não se pode desconsiderar que na construção do conhecimento:

De acordo com Moran (2000, p. 18):

O conhecimento não é fragmentado mas interdependente, interligado, intersensorial. Conhecer significa compreender todas as dimensões da realidade, captar e expressar essa totalidade de forma cada vez mais ampla e integral. Conhecemos mais e melhor conectando, juntando, relacionando, acessando o nosso objeto de todos os pontos de vista, por todos os caminhos, integrando-os da forma mais rica possível.

Na sociedade da informação se convive com várias formas de processamento da informação (sequencial, hipertextual ou multimídico) e a prática educativa não deve usar uma ou outra forma de lidar com as informações e, sim utilizar todas em diversos momentos do processo de ensino e aprendizagem.

Segundo Moran (2000, p.29-30):

Ensinar e aprender exigem hoje muito mais flexibilidade espaço-temporal, pessoal e de grupos, menos conteúdos fixos e processos mais abertos de pesquisa e de comunicação. Uma das dificuldades atuais é conciliar a extensão da informação, a variedade de fontes de acesso, com o aprofundamento da sua compreensão, em espaços menos rígidos, menos engessados. Temos informações demais e dificuldade em escolher quais são significativas para nós e em conseguir integrá-las dentro da nossa mente e da nossa vida.

A aquisição da informação, dos dados, dependerá cada vez menos do professor. As tecnologias podem trazer, hoje, dados, imagens, resumos de forma rápida e atraente. O papel do professor – o papel principal – é ajudar o aluno a interpretar esses dados, a relacioná-los, a contextualizá-los.

O papel do professor é o de mediador do processo de aprendizagem orientando intelectualmente e emocionalmente seus alunos, motivando-os, incentivando-os, estimulando-os a compreenderem, avaliarem as informações e a reelaborarem e a adaptarem nos seus contextos pessoais, colaborando para que eles ampliem o grau de compreensão da realidade em que vivem e a integrá-los a novas sínteses provisórias.

2.2 Contribuições do Computador e da Rede para a Aprendizagem

De acordo com Behrens (2000, p. 74):

O reconhecimento da era digital como nova forma de categorizar o conhecimento não implica descartar todo o caminho trilhado pela linguagem oral e escrita, nem mistificar o uso indiscriminado de computadores no ensino, mas enfrentar com critério os recursos eletrônicos como ferramentas para construir processos metodológicos mais significativos para aprender.

O uso do computador e da rede como recursos didáticos não garante em si o desenvolvimento de estratégias para uma ação educativa de qualidade, pois mais que as características técnicas destes equipamentos, se faz necessário que professores realizem a exploração didática destas tecnologias.

A melhora da aprendizagem com o uso desta tecnologia deve ser analisada com critérios pedagógicos.

Segundo Spritzer (2009, p.132-133):

1. O aproveitamento que se faz das características próprias da ferramenta informática: a capacidade de interação aluno/informação; a possibilidade de individualização, isto é, que os programas levem em conta as características individuais dos alunos; a capacidade de animação de figuras e sons que enriqueçam didaticamente o programa; a capacidade de simulação e a capacidade de retroalimentar a aprendizagem dos alunos.

2. A contribuição para aprendizagem desde uma perspectiva inovadora, isto é, que favoreça a participação solidária entre os alunos; que possibilite a pesquisa, a aprendizagem por descoberta e a recriação dos conhecimentos; que apresente uma visão integradora em sua concepção, e que propicie o tratamento interdisciplinar dos temas do currículo.

3. As modalidades de trabalho em aula: o impacto da utilização do computador sobre a aprendizagem varia em relação direta com o tamanho do grupo e compartilha seu uso. Em geral, os alunos que trabalham em duplas obtêm maiores resultados, ajudam-se mutuamente na interpretação e resolução do conteúdo da lição.

Esses critérios pedagógicos devem ser levados em conta na escolha do *software* que será utilizado na sala de aula, pois o uso da tecnologia não depende apenas de suas características, mas das atividades, objetivos, do ambiente de trabalho, do papel do professor, da concepção de aprendizagem.

Nesta dissertação a proposta de aulas com o uso de *software* para o ensino de estatística terá como fundamentação teórica a concepção de aprendizagem significativa.

2.3 Aprendizagem Significativa

David Ausubel (1918-2008) é um representante do cognitivismo ou psicologia da cognição que tem como foco o processo de compreensão, transformação e armazenamento de informações, apresentando em 1963 sua concepção teórica de ensino e aprendizagem oposta à concepção behaviorista que estava presente nas escolas naquela época.

Segundo Masini e Moreira (2001, 17):

O conceito mais importante na teoria de Ausubel é o de aprendizagem significativa. Para Ausubel, aprendizagem significativa é um processo pelo qual uma nova informação se relaciona com um aspecto relevante da estrutura de conhecimento do indivíduo. Ou seja, neste processo uma nova informação interage com uma estrutura de conhecimento específica, a qual Ausubel define como conceito *subsunçor* ou, simplesmente, *subsunçor* (*subsumer*), existentes na estrutura cognitiva do indivíduo. A aprendizagem significativa ocorre quando a nova informação ancora-se em subsunçores relevantes preexistentes na estrutura cognitiva de quem aprende.

Na concepção behaviorista não consideravam o que o aluno sabia e compreendiam que ele só aprenderia se fosse ensinado por outra pessoa, pois consideravam a influência do meio sobre o aluno. Mas, segundo Ausubel a aprendizagem significativa acontece quando se amplia e reconfigura os conhecimentos existentes na estrutura mental e, dessa forma, possibilitando ao aluno relacionar e acessar novos conhecimentos. Os professores tem o papel de propor situações que favoreçam o processo de aprendizagem onde novos conceitos serão aprendidos de forma significativa, em interação com conceitos já armazenados.

Para que a aprendizagem significativa ocorra se pressupõe as seguintes condições:

De acordo com Masini e Moreira (2001, p. 23-24):

- o material a ser aprendido seja potencialmente significativo para o aprendiz, ou seja, relacionável a sua estrutura de conhecimento de forma não-arbitrária e não-litera (substantiva);

- o aprendiz manifeste uma disposição de relacionar o novo material de maneira substantiva e não arbitrária a sua estrutura cognitiva.

A primeira dessas condições depende, obviamente, de pelo menos dois fatores principais, quais sejam, a natureza do material a ser aprendido e a natureza da estrutura cognitiva do aprendiz.

[...] A outra condição traz implícito que, independentemente de quão potencialmente significativo seja o material a ser aprendido, se a intenção do aprendiz é, simplesmente, a de memorizá-lo arbitrariamente e literalmente, tanto o processo de aprendizagem como seu produto serão mecânicos ou sem significado. (Reciprocamente, independente de quão predisposto para aprender estiver o indivíduo, nem o processo nem o produto serão significativos se o material não for potencialmente significativo).

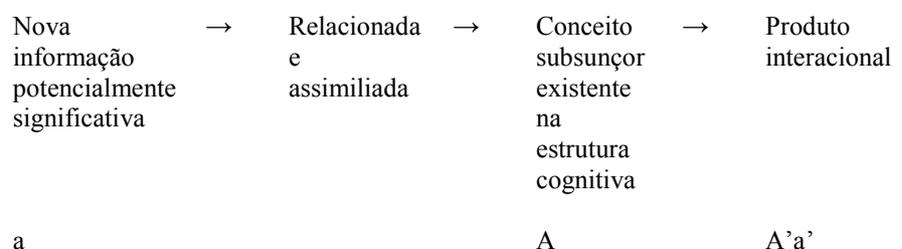
A utilização de situações-problema novas e não familiares ao aluno, a proposição de atividades que sequencialmente dependam uma das outras, a solicitação de que os alunos diferenciem ideias relacionadas a uma determinada proposição e a identificação pelo aluno dos elementos de um conceito. Ou ainda, a proposição em uma lista que contem outros elementos de conceitos e proposições similares são caminhos para verificar se essas condições foram atendidas na qual o aluno evidencia que a aprendizagem significativa realmente ocorreu.

Segundo Moreira (1999) Ausubel ao apresentar o princípio da assimilação ou teoria da assimilação deixa mais claro e preciso o processo de aquisição e organização de significados na estrutura cognitiva.

Segundo Moreira (1991, p. 24),

[...] o resultado da interação que ocorre na aprendizagem significativa entre o novo material a ser aprendido e a estrutura cognitiva existente é uma assimilação de antigos e novos significados a qual contribui para a diferenciação dessa estrutura. No processo de assimilação, mesmo após o aparecimento dos novos significados, a relação entre as ideias-âncoras e as assimiladas permanece na estrutura cognitiva.

O princípio da assimilação que, de acordo com Ausubel, possui valor explanatório tanto para a aprendizagem como para a retenção, pode ser representado esquematicamente da seguinte maneira:



No ensino de conceitos de matemáticos a assimilação é o processo que acontece quando um conceito a, por exemplo, regressão linear, deve ser aprendido sob um conceito subsunçor A - função de 1º e 2º grau já adquirido na estrutura cognitiva. Assim, como demonstra a

representação esquemática do princípio da assimilação, de Ausubel, apresentado por Moreira, a nova informação A e o conceito subordinador A são modificados nesta interação.

Os produtos interacionais $A' a'$ sofrerão outras modificações ao longo do tempo, envolvendo novas aprendizagens e perda de capacidade de reprodução de ideias subordinadas.

De acordo com Masini e Moreira (2001, p. 27),

[...] imediatamente após a aprendizagem significativa, cujo resultado é um produto interacional do tipo $A'a'$, começa um segundo estágio de assimilação: a assimilação obliteradora. As novas informações tornam-se, espontânea e progressivamente, menos dissociáveis de suas idéias-âncoras (subsunçores) até que não mais estejam disponíveis, isto é, não mais reproduzíveis como entidades individuais. Atinge-se, assim, um grau de dissociabilidade nulo, e $A'a'$ reduz-se simplesmente a A .

A aprendizagem significativa ocorre além da elaboração dos conceitos subordinadores, pois é possível, também, fazer interações entre esses conceitos. Dessa forma, para uma facilitação dessa aprendizagem se faz necessário que os recursos contribuam para a aquisição de uma estrutura cognitiva organizada.

Masini e Moreira (2001, p.32) afirmam que:

Ausubel sustenta o ponto de vista de que cada disciplina acadêmica tem uma estrutura articulada e hierarquicamente organizada de conceitos que constitui o sistema de informações dessa disciplina. Acredita que esses conceitos estruturais podem ser identificados e ensinados a um aluno, constituindo para ele um sistema de processamento de informações, um verdadeiro mapa intelectual que pode ser usado para analisar o domínio particular da disciplina e nela resolver problemas.

Segundo Masini e Moreira, na visão de Ausubel, o professor de cada disciplina deve primeiramente identificar os conceitos básicos dos conteúdos e como eles estão estruturados, para que esses conceitos estruturais facilitem o processo de ensino e aprendizagem do aluno.

O processo de ensino-aprendizagem tendo como referencial teórico a aprendizagem significativa deve ser planejado com práticas, exercícios, réplicas reflexivas que proporcionem aos alunos e professores experiências afetivas positivas.

3 CONTEUDOS MATEMÁTICOS EXPLORADOS

Antes de começar a ver estatísticas, é preciso conhecer os dados, saber que tipo de variável está sendo estudada. Há dois grandes grupos, as variáveis quantitativas e as qualitativas.

As variáveis quantitativas (ou numéricas) são divididas em dois subconjuntos, as quantitativas discretas e as quantitativas contínuas, onde os discretos resultam em contagem de eventos como, por exemplo, número de filhos, número de batimentos cardíacos por minuto, etc. Já as contínuas são obtidas por algum tipo de medição, exemplos, altura, peso, pressão arterial, etc.

As variáveis qualitativas (ou categóricas), também são divididas em dois subconjuntos, as qualitativas nominais e as qualitativas ordinais. Onde a nominal, não possui ordem, não tem hierarquia, como por exemplo, tipo sanguíneo, sexo do indivíduo, etc. Já as ordinais, possuem hierarquia e/ou ordem, exemplo, nível escolar, estágio de uma doença, etc.

Quando planejamos uma pesquisa, utilizando técnicas estatísticas, precisamos saber se usaremos informações de toda a população, ou só de parte dela (amostra). Com isso devemos prever as vantagens e desvantagens do uso de toda a população ou não. A população é o conjunto finito ou infinito de elementos, e a amostra (parte da população) é um subconjunto da população.

A vantagem de se utilizar a população, é que se está utilizando todos os dados, com isso temos muito mais confiabilidade nos resultados. A desvantagem é que em geral, não é possível se ter acesso às informações de toda a população, por motivos de custo alto, ou até mesmo por total impossibilidade.

Já a amostra, possui como maior desvantagem, a não utilização de todas as informações, o que faz, com que os resultados das estatísticas tenham uma margem de erro. Em contrapartida, as vantagens são o baixo custo, o fácil acesso e quando se busca uma amostra representativa da população, se tem resultados muito próximos aos da população. Por exemplo, quando se faz um senso para saber a expectativa de voto em um presidente da república, só se pergunta para uma amostra das pessoas, ou seja, só para alguma delas, por ser muito difícil perguntar a toda a população, por isso a amostra tem uma margem de erro, pois não se tem a certeza da intenção de voto de todos, mas se tem a intenção de votos de uma amostra representativa, onde a partir dela, se pode estimar o possível candidato eleito.

Para uma análise e interpretação de dados, além de observar seus gráficos, é preciso analisar algumas estatísticas descritivas deles, como as medidas de posição (média, moda, mediana, etc.) e medidas de dispersão (amplitude, variância, desvio padrão, etc.). Nesse capítulo será explorado um pouco mais dessas estatísticas, detalhando-as com fórmulas e exemplificando.

3.1 Medidas de posição

As medidas de posição, são utilizadas em geral para saber a localização dos dados, e uma idéia sobre o comportamento deles, utilizando alguns parâmetros. Uma aplicação interessante é usar as propriedades e relações entre as medidas de posição, aqui serão trabalhados três delas: a média aritmética, a mediana e a moda.

3.1.1 Média aritmética

É calculada somando-se os valores de todas as observações e dividindo-se essa soma pelo número de observações, assim, só pode ser calculado, se os dados forem quantitativos. A fórmula para esse cálculo segue abaixo.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad , \quad (1)$$

onde X_i são as observações e n é a quantidade de observações.

3.1.2 Mediana

É o valor que divide o conjunto de dados em duas partes tais que abaixo e acima da mediana encontram-se 50% das observações. O cálculo da mediana requer que os dados estejam ordenados. Se o número de observações for ímpar, a mediana é o valor central; se o número de observações for par, a mediana é a média dos dois valores centrais. A mediana em comparação a média é interessante na interpretação dos dados, pois ela desfavorece os valores atípicos (outliers), ao contrário da média. Essa estatística pode ser feita em dados quantitativos, ou qualitativos ordinais. As fórmulas seguem abaixo:

$$M_e = \begin{cases} X_{\frac{(n+1)}{2}} , & \text{se "n" é ímpar} \\ \frac{X_{\frac{n}{2}} + X_{\frac{(n+2)}{2}}}{2} , & \text{se "n" é par} \end{cases} . \quad (2)$$

3.1.3 Moda

É o valor mais freqüente dos dados, a moda pode ser calculada tanto nos dados quantitativos, como nos qualitativos, ela é sempre a observação com maior freqüência. Podemos classifica-la como unimodal, quando só possui uma observação com maior freqüência, bimodal, quando se tem duas observações que são mais freqüentes e com a mesma freqüência, multimodal, quando três ou mais observações são as mais freqüentes, ou amodal, quando não se tem nenhuma observação mais freqüente que as outras.

3.1.4 Exemplos de medidas de posição

Abaixo segue uma tabela com a população do Brasil no ano de 2007, por grande região.

Tabela 1 - População do Brasil - 2007

NÚMERO DE HABITANTES EM 2007	
Brasil	183.987.291
Região Norte	14.623.316
Região Nordeste	51.534.406
Região Centro-Oeste	13.222.854
Região Sudeste	77.873.120
Região Sul	26.733.595

Fonte: <http://noticias.uol.com.br/ultnot/2007/12/21/ult23u860.jhtm>

Inicialmente o que observamos nesses dados é a quantidade de habitantes por região, a variável região é qualitativa nominal, ou seja, não é possível se calcular a média nem a mediana das regiões, porém pode ser observada a moda. Outro fato importante é que os dados

representam toda a população do Brasil e então são populacionais. A moda nesse caso é a região Sudeste, por possuir uma maior frequência (mais habitantes), 77.873.120 pessoas.

Outro exemplo é a idade de dezesseis pessoas que prestaram o vestibular.

Tabela 2 – Alunos e Idades (Prestaram do Vestibular)

Alunos	Idade em anos
Amarildo	18
Bernardo	17
Carolina	17
Diego	18
Elton	19
Fernanda	16
Gustavo	17
Helio	25
Ivana	17
João	18
Karina	35
Leonardo	20
Mônica	18
Natália	19
Otávio	18
Pedro	19

Fonte: O Rafael Teixeira Silva, 2014.

A primeira coisa a se reparar, é que a variável idade em anos é quantitativa discreta, então podemos ter a média, a mediana e a moda. Outra informação importante é que se trata de uma amostra, pois não possui a idade de todas as pessoas que prestaram o vestibular.

A média é igual ao somatório das idades, dividido pelo número de pessoas, então: $(18 + 17 + 17 + 18 + 19 + 16 + 17 + 25 + 17 + 18 + 35 + 20 + 18 + 19 + 18 + 19) / 16$

Média = $311 / 16 = 19,4$ anos.

Para se calcular a mediana, primeiro é preciso ordenar as observações. Temos uma quantidade par de observações, e assim é preciso fazer a média aritmética das observações, entre as posições $X(n/2)$ e $X(n/2 + 1)$, então a mediana é $(18 + 18)/2 = 18$ anos.

A moda é o valor mais frequente, no caso 18 anos, com cinco observações, trata-se de distribuição unimodal.

Interpretando esses dados, é interessante, notar que a moda e a mediana foram iguais a 18 anos, o que revela qual usualmente costuma ser a idade das pessoas que prestam o vestibular. Porém o cálculo da média nos dá um valor maior, 19.4 anos, isso ocorre porque a média considera os elementos fora do padrão, que é o caso da Karina com 35 anos, uma idade bem acima das outras pessoas.

3.2 Medidas de dispersão

Ao se fazer as medidas de posição, temos boas estatísticas para entender a posição dos dados, porém ela sozinha ainda não é suficiente para se interpretar os dados. Ao observar o exemplo abaixo sobre a idade de algumas pessoas, isso deve ficar mais claro.

Tabela 3 – Medidas de Dispersão

Nomes	Idade
João	18
Fernanda	18
Gustavo	18
Rafael	18

(a)

Nomes	Idade
Daniel	10
Bruno	26
Juliana	18
Pedro	18

(b)

Fonte: O Rafael Teixeira Silva, 2014..

Visivelmente os dois grupos possuem comportamentos diferentes, porém se olharmos só as medidas de posição, não é possível reparar essa distinção, pois a média, a mediana e a moda dos dois grupos é igual a 18 anos. Para se reparar essa diferença, calcula-se algumas medidas de dispersão, também conhecidas como medidas de variação.

As medidas de dispersão (ou variação), são bem interessantes, pois é com elas que se vê a variabilidade das observações, onde quanto maior essa medida, maior será sua variabilidade, essas medidas de posição só podem ser calculadas com dados quantitativos. Nesse estudo vamos trabalhar com três principais medidas, a amplitude, a variância e o desvio padrão.

3.2.1 Amplitude

A amplitude é uma medida de variação, usada para se ver qual a diferença entre o maior valor da observação e o menor valor da observação. Ou seja, no caso dos dois grupos acima, o primeiro tem amplitude igual a 0 ano e o segundo tem amplitude igual 16 anos, portanto, já se pode ver que há diferença entre os dois grupos.

3.2.2 Variância

Essa medida de dispersão, possui dos casos distintos, uma fórmula para dados populacionais e outra para dados amostrais. Como veremos a seguir.

A variância populacional é usada para se perceber qual a variação em relação a média populacional dos dados, dividida pelo número de observações, onde esta é de unidades ao quadrado, sua fórmula é:

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2}{N} \quad (3)$$

A variância amostral é muito parecida com a populacional, porém existem algumas diferenças. Por se tratar de uma amostra, temos a dispersão dos dados em relação a média amostral. Para seu cálculo devemos efetuar a divisão pelo tamanho da amostra menos um. É necessário esse ajuste na fórmula, com a intenção de criar uma medida de dispersão não tendenciosa. A fórmula da variância amostral fica assim:

$$S^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} = S^2 = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2 \quad (4)$$

No exemplo dos dois grupos de idades, por se tratar de algumas pessoas, ou seja uma amostra, então usa-se a segunda fórmula. No primeiro grupo, não há dispersão nenhuma, pois as quatro pessoas possuem a mesma idade, já o segundo grupo não, então precisa-se calcular:

$$\text{Var. Amo.} = [(10 - 18)^2 + (26 - 18)^2 + (18 - 18)^2 + (18 - 18)^2] / (4 - 1)$$

$$\text{Var. Amo.} = [64 + 64 + 0 + 0] / 3$$

Var. Amo. = 42,67

3.2.3 Desvio padrão

Como a variância é uma unidade ao quadrado, pois tinha que ser, para se efetuar a soma de diferenças, existe uma outra medida de dispersão, conhecida como desvio padrão, que utiliza as mesmas unidades dos dados. Assim, ela é calculada como a raiz quadrada da variância sendo assim, existem também dois tipos de desvio padrão, o amostral e o populacional, como na variância. Em caso de amostra, raiz quadrada da variância amostral e se for população, raiz quadrada da variância populacional. A fórmula do desvio padrão populacional e amostral seguem abaixo:

$$\begin{aligned} S^2 &= \frac{\sum(x-\bar{x})^2}{n-1} \\ \sigma^2 &= \frac{\sum(x-\mu)^2}{N} \end{aligned} \quad (5)$$

onde S^2 é a variância amostral e o σ^2 é a variância populacional.

No caso do exemplo dos grupos de idades, vai ser usado o desvio padrão amostral, como o primeiro grupo não possui nenhuma dispersão, o seu desvio padrão vai ser 0 ano, já o segundo grupo, o desvio padrão amostral, vai ser a raiz quadrada da variância amostral:

$$\text{Desv. Pad. Amo.} = \sqrt{\text{Var. Amo.}}$$

$$\text{Desv. Pad. Amo.} = \sqrt{42,67}$$

$$\text{Desv. Pad. Amo.} = 6,53 \text{ anos.}$$

3.2.4 Coeficiente de variação

O coeficiente de variação revela qual é a proporção do desvio padrão em relação a média, por ser adimensional, ele é utilizado para comparar a variabilidade de duas variáveis distintas. Por exemplo, comparar a variabilidade entre o peso dos alunos, em quilos, com a altura deles, em metros. Nesse caso, esperamos que a variação (desvio padrão por exemplo) será maior para o peso em quilos do que da altura em metros, porém usando o coeficiente de variação, é possível comparar a variabilidade dessas duas variáveis (peso e altura). A fórmula é dada por:

$$CV = \frac{\sigma}{\mu} \text{ ou } \frac{S}{x} \quad (6)$$

No caso dos grupos de idade, o primeiro grupo não possui variabilidade e o coeficiente de variação amostral do segundo grupo, vai ser igual a:

$$CV = \text{Des. Pad. Amo.} / \text{Média Amostral}$$

$$CV = 6,53 / 18$$

$$CV = 0,36 = 36 \%$$

3.3 Método dos Mínimos Quadrados Ordinários (MQO)

Após observar os gráficos e suas medidas de posição e de dispersão, para interpretar e entender melhor os dados, é possível fazer regressões lineares a fim de estimar o que vai acontecer com certa variável no futuro, ou até mesmo explicar o comportamento dessa variável, observando a sua causa e efeito.

Regressão linear é um método estatístico, usado para estimar uma variável dependente, usando uma variável independente. Na regressão os valores estimados são os valores esperados, em média, por isso há alguns erros, pois os dados são de variáveis aleatórias.

A técnica utilizada para se fazer essa regressão linear, é o método dos mínimos quadrados, que por ser uma otimização, ele quer achar o melhor ajuste para os dados, visando sempre minimizar a soma dos quadrados dos resíduos, onde esses resíduos são as diferenças entre o valor estimado e o observado, maximizando o grau de ajuste do modelo dos dados. Mas essa técnica possui alguns axiomas para poder ser utilizada de forma correta. O erro deve ser distribuído aleatoriamente, com média zero e sua distribuição normal e independente, e ainda as variáveis envolvidas devem possuir uma relação linear entre si. Com isso, fica garantido que o estimador de mínimos quadrados é não-enviesado (estimador centrado, ou seja, quando a média das estimativas obtidas para todas as amostras possíveis são iguais ao parâmetro a estimar) e de mínima variância linear na variável resposta.

O modelo estimado é do formato: $Y_i = a + b \cdot X_i + e_i$ onde Y_i é a variável dependente, o regressando, X_i é a variável independente, o regressor, $a + b \cdot X_i$ é a reta de regressão dos dados, onde a é o intercepto da reta de regressão, b é a declividade da reta e e_i é o termo de erro.

A reta de regressão fornece o valor médio (valor esperado) de Y como função condicional ao valor de X :

$$E(Y/X) = a + b \cdot X$$

O coeficiente b é a variação esperada em Y associada a uma variação de uma unidade de X e o a é o nível ou altura da reta de regressão.

O método dos Mínimos Quadrados Ordinários consiste em posicionar a reta de regressão “no meio da nuvem de pontos”, de tal forma que torne o somatório dos resíduos iguais a zero e minimize o somatório dos resíduos quadráticos, ou seja, minimizando a função abaixo:

$$\sum_{i=1}^n (e_i)^2 \quad , \quad (7)$$

que é o mesmo que minimizar

$$S(a, b) = \sum_{i=1}^n (y_i - a - bx_i)^2 \quad . \quad (8)$$

Para se minimizar a função acima, é preciso usar uma técnica da disciplina de Cálculo, que consiste em derivar a função $S(a,b)$ em relação a a e b e igualar a zero.

$$\frac{\partial S}{\partial a} = -2 \sum_{i=1}^n (y_i - a - bx_i) = 0$$

$$\frac{\partial S}{\partial b} = -2 \sum_{i=1}^n x_i (y_i - a - bx_i) = 0 \quad (9)$$

onde efetuando os cálculos e dividindo a derivação de S em relação a a por $2 \cdot n$,

$$\begin{aligned} \frac{-2 \sum_{i=1}^n y_i}{2n} + \frac{2 \sum_{i=1}^n a}{2n} + \frac{2 \sum_{i=1}^n bx_i}{2n} &= \frac{0}{2n} \\ \frac{-\sum_{i=1}^n y_i}{n} + \frac{\sum_{i=1}^n a}{n} + \frac{b \sum_{i=1}^n x_i}{n} &= 0 \\ -\bar{y} + a + b\bar{x} &= 0 \\ a &= \bar{y} - b\bar{x} \end{aligned} \quad (10)$$

Nesse caso, \bar{y} é a média da variável Y e \bar{x} é a média de X , substituindo esses resultados nas fórmulas abaixo e efetuando os cálculos.

$$-2 \sum_{i=1}^n x_i (y_i - \bar{y} + b\bar{x} - bx_i) = 0$$

$$\begin{aligned}
\sum_{i=1}^n [x_i (y_i - \bar{y}) + x_i b (\bar{x} - x_i)] &= 0 \\
\sum_{i=1}^n x_i (y_i - \bar{y}) + b \sum_{i=1}^n x_i (\bar{x} - x_i) &= 0 \\
b &= \frac{\sum_{i=1}^n x_i (y_i - \bar{y})}{\sum_{i=1}^n x_i (\bar{x} - x_i)} \tag{11}
\end{aligned}$$

Para utilizar o MQO de forma correta, é preciso que as seguintes hipóteses sejam válidas:

1º. A distribuição condicional do erro dado X tem média zero, ou seja, $E(e/X) = 0$.

Por exemplo, se e e X forem correlacionados, então $E(e/X) \neq 0$.

Essa hipótese implica que b é não viesado (estimador centrado).

2º. (X_i, Y_i) , $i=1, \dots, n$ são i.i.d (independentes e identicamente distribuídas).

Ou seja, são seleções independentes e identicamente distribuídas considerando a distribuição conjunta. Isso ocorre se, por exemplo, (X, Y) forem obtidos via amostragem aleatória simples, permitindo construir a distribuição amostral de a e b.

3º. Valores muito aberrantes (outliers) em X e/ou Y são raros

Tecnicamente, equivale a X e Y terem os quartos momentos finitos e diferentes de zero. Os outliers podem resultar em valores sem sentido para b.

3.3.1 Coeficiente de determinação

Um indicador simples e conhecido para se ver a apuração do modelo é o coeficiente de determinação, também chamado de R^2 , esse varia entre 0 e 1, quanto mais próximo de um, mais o modelo pode ser explicado pela variável independente. Porém, o coeficiente de determinação tem algumas falhas, pois ele é apenas o quadrado da correlação das variáveis X e Y, mas é considerado um indicador. Sua fórmula é dada por:

$$R^2 = 1 - \frac{SQ_{res}}{SQ_{tot}} \tag{12}$$

onde SQ_{res} é a soma do quadrado dos resíduos e SQ_{tot} é a soma dos quadrados do total.

3.3.2 Exemplo de mínimos quadrados

Vamos imaginar que foi pego uma amostra aleatória simples, de três pessoas do cadastro de um banco, onde exibia a informação de quanto o cliente consumia no cartão de crédito e de quanto era sua renda. A intenção aqui é fazer um modelo de regressão linear simples, onde se quer estimar a renda mensal do cliente, usando como variável independente seu consumo mensal no cartão de crédito.

Tabela 4 – Consumo Mensal no Cartão de Crédito

Clientes	Consumo mensal no cartão de crédito em reais	Renda mensal em reais
João	1200,00	2000,00
Daniel	860,00	1300,00
Marcos	1330,00	2500,00

Fonte: O Rafael Teixeira Silva, 2014.

Média de Y = $(2000 + 1300 + 2500) / 3 = 1933,33$ reais.

Média de X = $(1200 + 860 + 1330) / 3 = 1130,00$ reais.

Utilizando a fórmula do b, temos:

$$b = [1200*(2000 - 1933,33) + 860*(1300 - 1933,33) + 1330*(2500 - 1933,33)] / [1200*(1200 - 1130) + 860*(860 - 1130) + 1330*(1330 - 1130)] = b = 2,4533$$

$$a = 1933,33 - 2,4533*1130,00 = a = -838,91$$

Então o modelo seria:

$$Y = -838,91 + 2,4533*X$$

Calculando as estimativas do modelo, para ver se ele está assertivo.

Tabela 5 – Estimativas do Modelo

Consumidores	Consumo mensal no cartão de crédito em reais	Renda mensal em reais	Estimativa da renda
João	1200,00	2000,00	2105,07
Daniel	860,00	1300,00	1270,94
Marcos	1330,00	2500,00	2424,00

Fonte: O Rafael Teixeira Silva, 2014.

Como forma de ver se o modelo é bom, vamos usar o indicador do coeficiente de determinação (R^2).

$$R^2 = 1 - [(2105,07 - 2000)^2 + (1270,94 - 1300)^2 + (2424,00 - 2500)^2] / [(2000^2 + 1300^2 + 2500^2) - 3 \cdot (1933,33)^2] = R^2 = 0,9757 = 97,57 \%$$

Portanto o modelo parece estar bem estimado, pois o coeficiente de determinação está perto de um.

4 CONSTRUÇÃO DAS PLANILHAS

Nos exercícios feitos até agora, foram usados amostras pequenas, de apenas três ou quatro elementos, como quase todas as fórmulas de estatística possuem o somatório de muitos termos, imagina se fossem dez observações ou cem, ou pior mil, etc. Seriam muitas contas, e consumiria muito tempo para chegar a bons resultados de interpretação de dados. A solução é o uso de softwares que possuem essas fórmulas estatísticas já prontas, onde pode-se utilizar apenas um menu e o software calcula todas essas estatísticas para você exibindo também os gráficos para visualização dos dados. Nesse estudo utilizamos o software LibreOffice Calc. A Figura 1 mostra uma típica planilha de cálculo desse software.

Figura 1 - Apresentação do Software LibreOffice

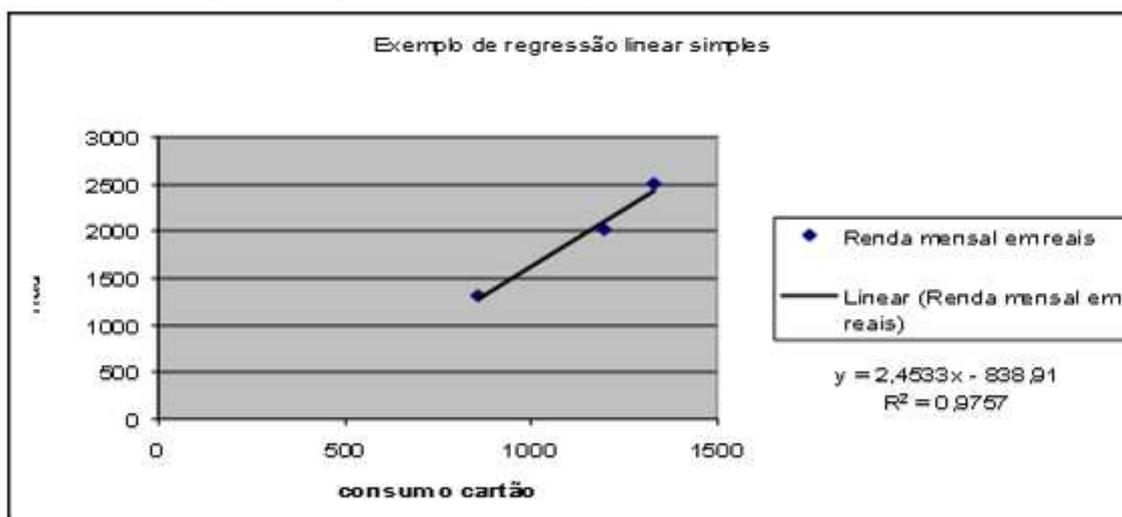
Observação	Variável	Estatísticas	Cálculo
		Média	
		Mediana	
		Moda	
		Amplitude	0,00
		Variância Populacional	
		Variância Amostral	
		Desvio Padrão Populacional	
		Desvio Padrão Amostral	
		Coefficiente Variação amostral	
		Coefficiente Variação populacional	

Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

O *LibreOffice* foi escolhido, por ser um software de fácil acesso e gratuito, sendo possível baixá-lo pela internet de forma rápida e segura, através do link www.libreoffice.org/download/. Outra vantagem dele é por funcionar em planilha, de fácil manuseio e com muitas informações por menu, onde o usuário não necessita ter grandes formações lógicas e programação. Além disso, o Calc possui as principais funções estatísticas e formatos de gráficos. Um fator importante para a escolha desse software, foi o motivo de ser bastante usado e com isso ter muitos usuários, podendo assim, ocorrer uma troca de informação, caso precise utilizar algum tipo de dado específico.

Para exemplificar a vantagem do software, foi feito o exercício anterior de regressão linear simples, representado na Figura 2, utilizando simples cliques no menu do Calc.

Figura 2 - Exemplo de Regressão Linear Simples



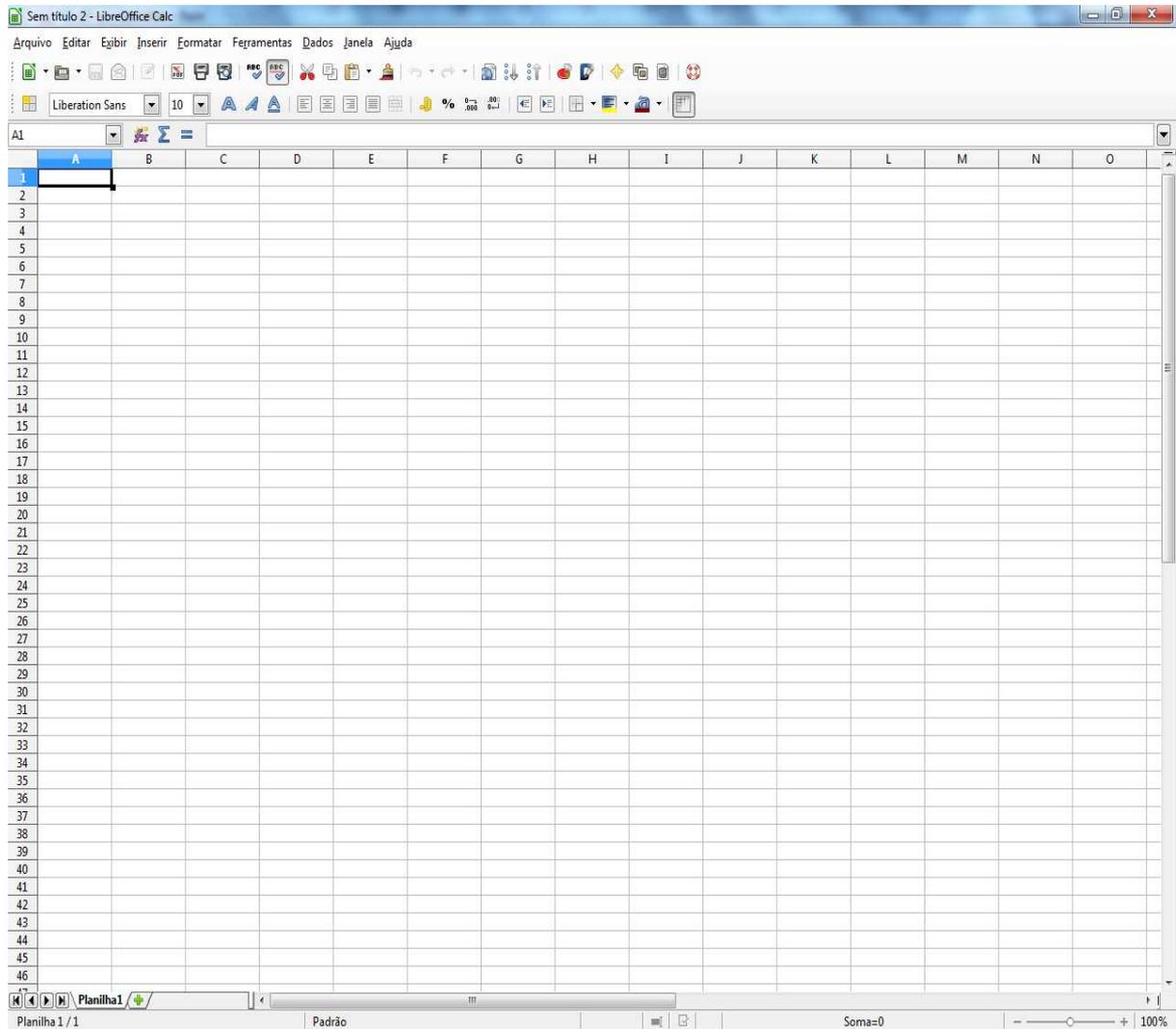
Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Assim, observamos melhor a visualização dos dados, da reta de regressão, e há ainda a vantagem de chegar a resultados estatísticos de forma muito mais rápida e eficiente, sem correr o risco de errar em cálculos por exemplo.

A fim de contribuir na educação do ensino básico, foi feita uma planilha com as principais estatísticas e respectivo gráfico, para qualquer usuário leigo no entendimento do software poder trabalhar. Nesse capítulo, vou mostrá-la funcionando e o passo a passo para que qualquer professor ou aluno possa montar e praticar a interpretação de dados.

A seguir a Figura 3, que mostra a interface do *software*.

Figura 3 - Interface do Software



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Como foi mencionado anteriormente, este software funciona por planilha, onde cada retângulo é uma célula para ser preenchida como quiser, com fórmula, número, gráfico, etc. No caso acima a célula selecionada foi a da coluna A com a linha 1, portanto a A1.

Vamos ao passo a passo de construção da planilha:

Figura 4 - Construindo a Planilha

1	A	B	C	D	E	F	G	H
1	Observação	Variável	Estatísticas	Cálculo				
2			Média					
3			Mediana					
4			Moda					
5			Amplitude	0,00				
6			Variância Populacional					
7			Variância Amostral					
8			Desvio Padrão Populacional					
9			Desvio Padrão Amostral					
10			Coefficiente Variação amostral					
11			Coefficiente Variação populacional					
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								
35								
36								
37								
38								
39								
40								
41								
42								
43								
44								
45								
46								

Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Observando a Figura 4, podemos perceber que na primeira coluna (A) da planilha vão ficar as observações como, por exemplo, o nome das pessoas, as datas, os itens, os casos, etc. Já na segunda coluna (B) vão ficar os valores dessas observações, para alguma variável, seja ela temperatura, altura, peso, etc. Como o software não só faz cálculos para variáveis quantitativas, todas as variáveis utilizadas daqui para frente, vão ser numéricas. A terceira coluna (C) possui os nomes das estatísticas que o software vai calcular e por ao lado delas, na quarta coluna (D).

Agora vamos observar o cálculo de cada uma dessas estatísticas, começando pela média, utilizando a fórmula MÉDIA, como representado na Figura 5.

Figura 5 – Fórmula MÉDIA

The screenshot shows the LibreOffice Calc interface. The formula bar at the top displays the formula: `=SEERRO(MÉDIA(B1:B10000);'')`. The spreadsheet below has the following structure:

	A	B	C	D	E	F	G	H
1	Observação	Variável	Estatísticas	Cálculo				
2			Média					
3			Mediana					
4			Moda					
5			Amplitude	0,00				
6			Variância Populacional					
7			Variância Amostral					
8			Desvio Padrão Populacional					
9			Desvio Padrão Amostral					
10			Coefficiente Variação amostral					
11			Coefficiente Variação populacional					
12								
13								
14								
15								
16								
17								

Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Foi utilizada primeiramente a função SEERRO, pois como a planilha possui células vazias, utilizando essa função, é possível calcular as estatísticas, considerando apenas o número de observações que se tem, podendo esse número aumentar ou diminuir, contanto que fique entre o intervalo de uma e dez mil observações. Essa função SEERRO possui dois argumentos: o primeiro é o valor da célula caso não tenha erro, ou seja, considerando apenas as células preenchidas e de forma numérica, o segundo é o valor alternativo.

Então, pode-se perceber que foi utilizada a função MÉDIA no primeiro argumento da função SEERRO, onde calcula-se a média da variável e o resultado fica na célula D2. A função MÉDIA possui apenas um argumento, que é o intervalo de dados, que está na coluna B.

Após calcular a média dos dados, será feito também o cálculo da mediana, como representado na Figura 6.

Figura 6 – Cálculo da Mediana

The screenshot shows the LibreOffice Calc interface. The formula bar at the top displays the formula `=SEERRO(MED(B1:B10000);'')`. The spreadsheet below has the following structure:

A	B	C	D	E	F	G	H
1	Observação	Variável	Estatísticas	Cálculo			
2			Média				
3			Mediana				
4			Moda				
5			Amplitude	0,00			
6			Variância Populacional				
7			Variância Amostral				
8			Desvio Padrão Populacional				
9			Desvio Padrão Amostral				
10			Coefficiente Variação amostral				
11			Coefficiente Variação populacional				
12							
13							
14							
15							
16							
17							
18							

Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Para se calcular a mediana, utilizou-se como o primeiro argumento da função SEERRO a função MED, essa última possui apenas um argumento, sendo este o intervalo de dados, a coluna B (variável). O resultado dessa função será o valor da célula D3, esta função ordena os valores e seleciona o valor central, ou seja, a mediana dos dados.

Outra estatística descritiva interessante é a moda, que pode ser calculada, utilizando a função, como mostra a Figura 7.

Figura 7 – Cálculo da Moda

1	A	B	C	D	E	F	G	H
1	Observação	Variável	Estatísticas	Cálculo				
2			Média					
3			Mediana					
4			Moda					
5			Amplitude	0,00				
6			Variância Populacional					
7			Variância Amostral					
8			Desvio Padrão Populacional					
9			Desvio Padrão Amostral					
10			Coefficiente Variação amostral					
11			Coefficiente Variação populacional					
12								
13								
14								
15								
16								
17								
18								
19								

Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

O valor da célula D4 vai ser a moda da variável, ou seja, o valor mais freqüente, onde foi utilizada a função MODO no primeiro argumento da função SEERRO. A função MODO possui um único argumento, o intervalo de dados, a coluna B, onde é visto qual o valor possui a maior freqüência.

A fim de ver as medidas de dispersão, é calculada a amplitude dos dados, representada na Figura 8.

Figura 8 – Medidas de Dispersão

1	A	B	C	D	E	F	G	H
1	Observação	Variável	Estatísticas	Cálculo				
2			Média					
3			Mediana					
4			Moda					
5			Amplitude	0.00				
6			Variância Populacional					
7			Variância Amostral					
8			Desvio Padrão Populacional					
9			Desvio Padrão Amostral					
10			Coefficiente Variação amostral					
11			Coefficiente Variação populacional					
12								
13								
14								
15								
16								
17								
18								

Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Para se calcular a amplitude da variável, é utilizada a diferença de duas funções, a MÁXIMO e a MÍNIMO, onde é utilizada primeiramente a função SEERRO, no intuito de só se observar as células preenchidas de forma quantitativa, então a primeira função SEERRO leva como primeiro argumento a função MÁXIMO, onde esta última, possui apenas um argumento, o intervalo de dados, a coluna B da planilha; o segundo argumento da função SEERRO é "", para indicar vazio, não considerado no cálculo. Após calcular o valor máximo dos dados, se faz a diferença do valor mínimo deles, utilizando a função SEERRO como primeiro argumento a função MÍNIMO, sendo esta última, com apenas um argumento, o intervalo de dados, a variável. O segundo argumento da função SEERRO é o "". Com isso a célula D5 vai receber o valor da amplitude.

Caso os dados sejam populacionais, deve-se calcular a variância populacional, utilizando o passo a passo que se segue.

Figura 9 – Variância Populacional

Observação	Variável	Estatísticas	Cálculo
		Média	
		Mediana	
		Moda	
		Amplitude	0,00
		Variância Populacional	
		Variância Amostral	
		Desvio Padrão Populacional	
		Desvio Padrão Amostral	
		Coefficiente Variação amostral	
		Coefficiente Variação populacional	

Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Primeiramente, utiliza-se a função SEERRO, onde o seu primeiro argumento vai ser a função VARP, sendo esta, com um argumento, o intervalo de dados, por isso foi selecionada a coluna B, \$B\$1:\$B\$10000, indicando a seleção da célula B1 até a célula B10000, como mostra a Figura 9, lembrando que mesmo selecionada de um a dez mil células, a função SEERRO só vai permitir fazer o cálculo da variância populacional, em células preenchidas de forma numérica, caso não seja, o segundo argumento da função SEERRO é "", vazio. A célula D6 vai receber o valor da variância populacional.

Caso os dados não sejam populacionais e sim amostrais, como na maioria dos casos, é preciso calcular a variância amostral, como o passo a seguir, representado na Figura 10.

Figura 10 – Cálculo da Variância Amostral

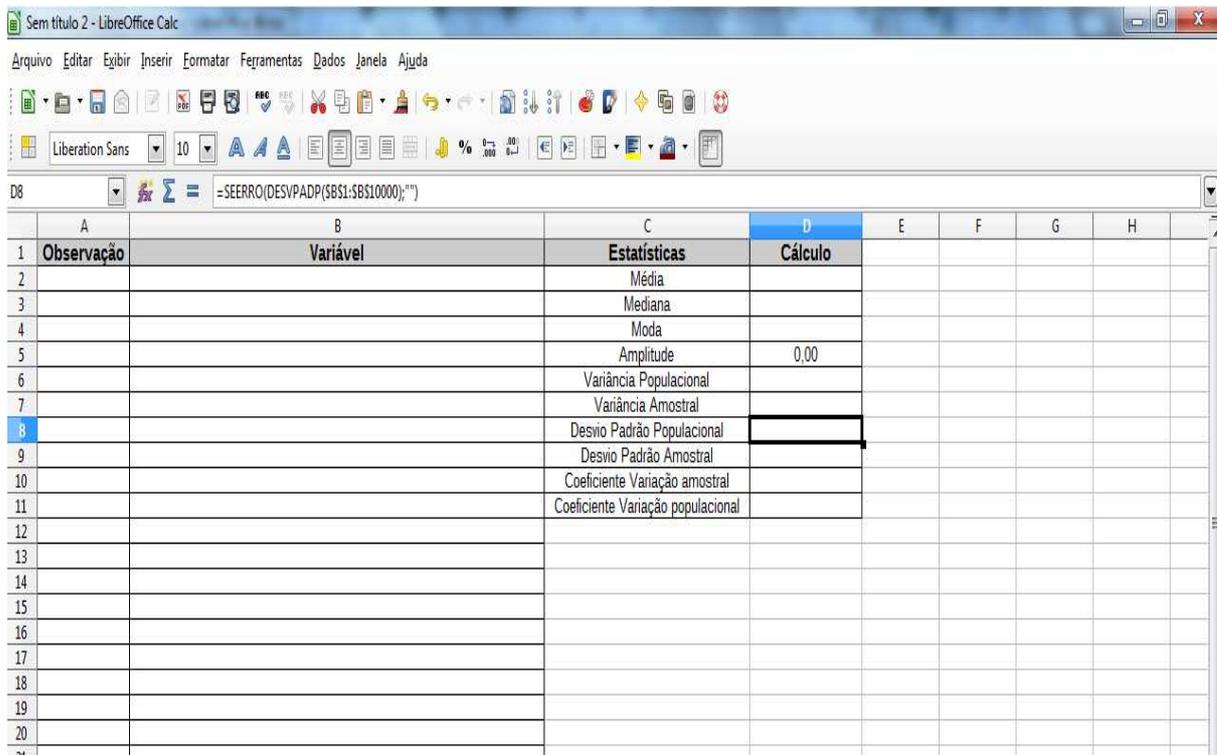
1	A	B	C	D	E	F	G	H
1	Observação	Variável	Estatísticas	Cálculo				
2			Média					
3			Mediana					
4			Moda					
5			Amplitude	0,00				
6			Variância Populacional					
7			Variância Amostral					
8			Desvio Padrão Populacional					
9			Desvio Padrão Amostral					
10			Coefficiente Variação amostral					
11			Coefficiente Variação populacional					
12								
13								
14								
15								
16								
17								
18								
19								
20								

Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

O primeiro argumento da função SEERRO é a função DESVPAD e possui somente um argumento, o intervalo de dados, a coluna B, a variável. O segundo argumento de SEERRO é “”, vazio. A célula D9 vai assumir o valor do desvio padrão amostral.

Para poder comparar a variabilidade de variáveis diferentes, calcula-se o coeficiente de variação, onde este também pode ser populacional ou amostral, dependendo se estão todos os dados, ou só parte deles. Para calcular o coeficiente de variação amostral é preciso o cálculo da média amostral e do desvio padrão amostral, como se segue.

Figura 11 - Cálculo do Desvio Padrão Populacional



	A	B	C	D	E	F	G	H
1	Observação	Variável	Estatísticas	Cálculo				
2			Média					
3			Mediana					
4			Moda					
5			Amplitude	0,00				
6			Variância Populacional					
7			Variância Amostral					
8			Desvio Padrão Populacional					
9			Desvio Padrão Amostral					
10			Coefficiente Variação amostral					
11			Coefficiente Variação populacional					
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								

Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

A função DESVPAD possui um argumento, o intervalo de dados, coluna B da planilha, calculando o desvio padrão populacional da variável, colocando o seu valor na célula D8, para só realizar cálculos em células preenchidas numericamente, foi usada a função DESVPAD no primeiro argumento da função SEERRO, onde o segundo argumento desta é “”, vazio, como mostra a Figura 11.

Mas caso os dados sejam amostrais, é preciso olhar o desvio padrão amostral, calculado na célula D9 da planilha, representado na Figura 12.

Figura 12 – Observando o Desvio Padrão Amostral

Observação	Variável	Estatísticas	Cálculo
		Média	
		Mediana	
		Moda	
		Amplitude	0,00
		Variância Populacional	
		Variância Amostral	
		Desvio Padrão Populacional	
		Desvio Padrão Amostral	
		Coefficiente Variação amostral	
		Coefficiente Variação populacional	

Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

O primeiro argumento da função SEERRO é a função DESVPAD, esta só possui um argumento, o intervalo de dados, a coluna B, a variável. O segundo argumento de SEERRO é "", vazio. A célula D9 vai assumir o valor do desvio padrão amostral.

Para poder comparar a variabilidade de variáveis diferentes, calcula-se o coeficiente de variação, onde este também pode ser populacional ou amostral, dependendo se estão todos os dados, ou só parte deles. Para calcular o coeficiente de variação amostral é preciso o cálculo da média amostral e do desvio padrão amostral, como mostra a Figura 13.

Figura 13 – Cálculo do Coeficiente de Variação Amostral

1	A	B	C	D	E	F	G	H
	Observação	Variável	Estatísticas	Cálculo				
2			Média					
3			Mediana					
4			Moda					
5			Amplitude	0,00				
6			Variança Populacional					
7			Variança Amostral					
8			Desvio Padrão Populacional					
9			Desvio Padrão Amostral					
10			Coefficiente Variação amostral					
11			Coefficiente Variação populacional					
12								
13								
14								
15								
16								
17								
18								
19								

Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Agora usamos a função SEERRO, onde o primeiro argumento é a divisão entre o desvio padrão amostral (célula D9) e a média amostral (célula D2). Lembrando que o segundo argumento da função SEERRO é “”, vazio. A célula D10 vai possuir o valor do coeficiente de variação amostral.

Caso os dados sejam populacionais, a média calculada na célula D2 é uma média populacional; então usa-se ela e o desvio padrão populacional para se calcular o coeficiente de variação populacional, representado na Figura 14.

Figura 14 – Cálculo do Coeficiente de Variação Populacional

The screenshot shows the LibreOffice Calc interface. The formula bar at the top displays the formula $=SEERRO(\$D\$8/\$D\$2;'')$. The spreadsheet has the following structure:

1	A	B	C	D	E	F	G	H
1	Observação	Variável	Estatísticas	Cálculo				
2			Média					
3			Mediana					
4			Moda					
5			Amplitude	0,00				
6			Variancia Populacional					
7			Variancia Amostral					
8			Desvio Padrão Populacional					
9			Desvio Padrão Amostral					
10			Coeficiente Variação amostral					
11			Coeficiente Variação populacional					
12								
13								
14								
15								
16								
17								

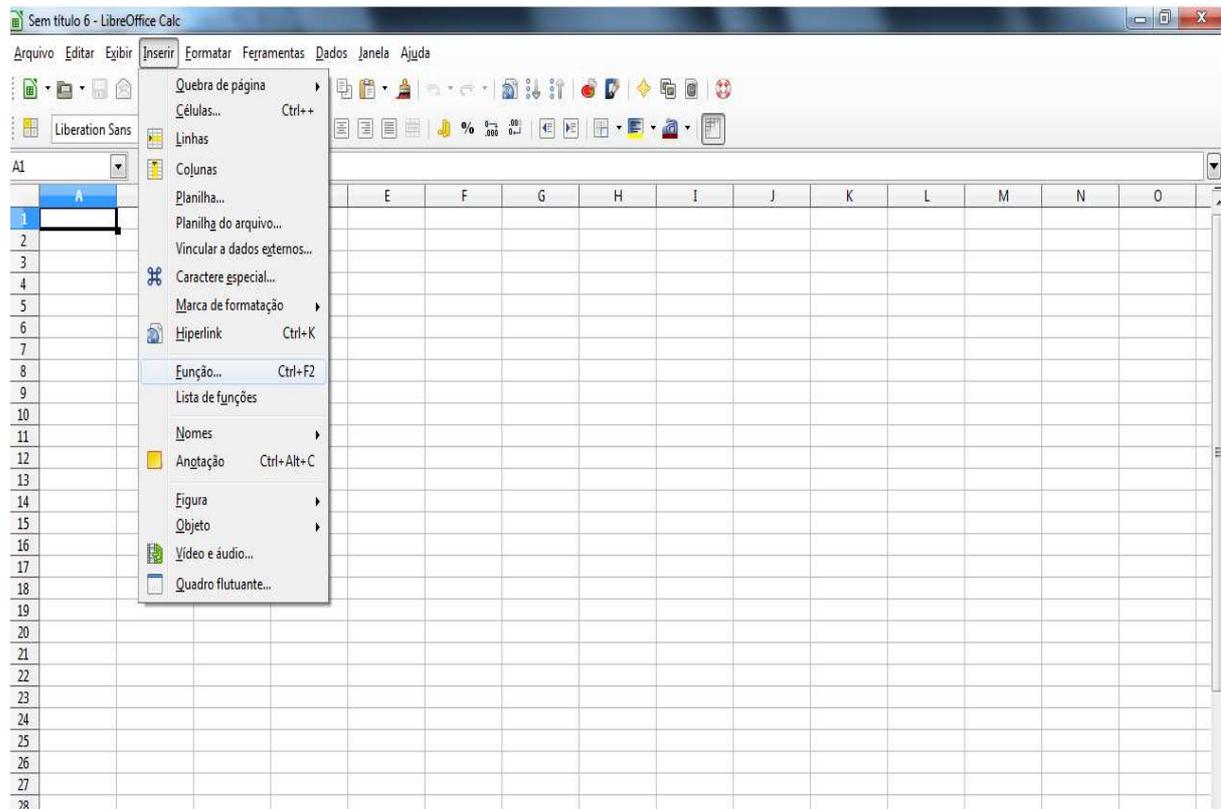
Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

O primeiro argumento da função SEERRO é a divisão do desvio padrão populacional (célula D8) pela média populacional (célula D2), o segundo argumento é “”, vazio. O valor do coeficiente de variação populacional, será inserido na célula D11.

Para calcular essas e outras funções estatísticas, segue o passo a passo.

Primeiro clique em Inserir na parte superior da planilha, onde vai abrir as opções abaixo, depois é preciso clicar na opção função (Ctrl+F2), como mostra a Figura 15, onde abrirá uma nova janela.

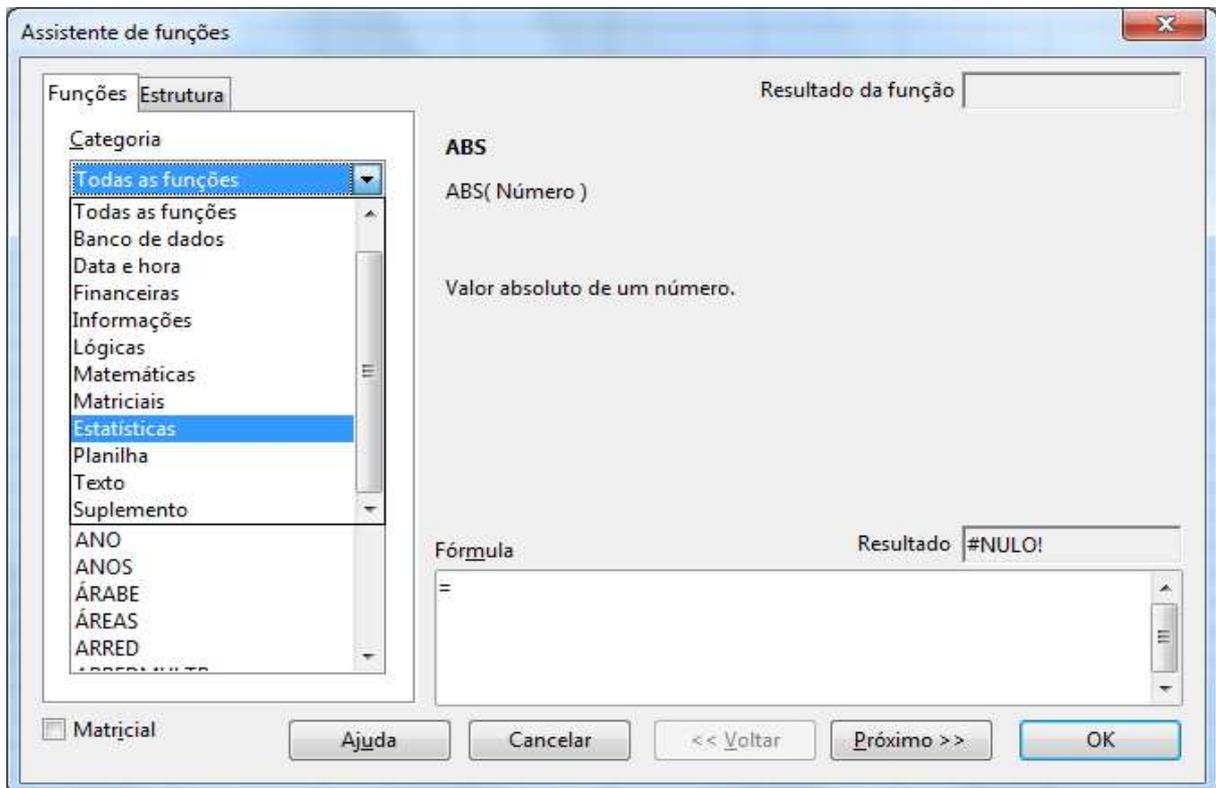
Figura 15 – Inserir funções no LibreOffice Calc.



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Essa nova janela está representada na Figura 16. No intuito de facilitar a visualização das funções estatísticas, selecione a Categoria Estatísticas.

Figura 16 – Categorias Estatísticas

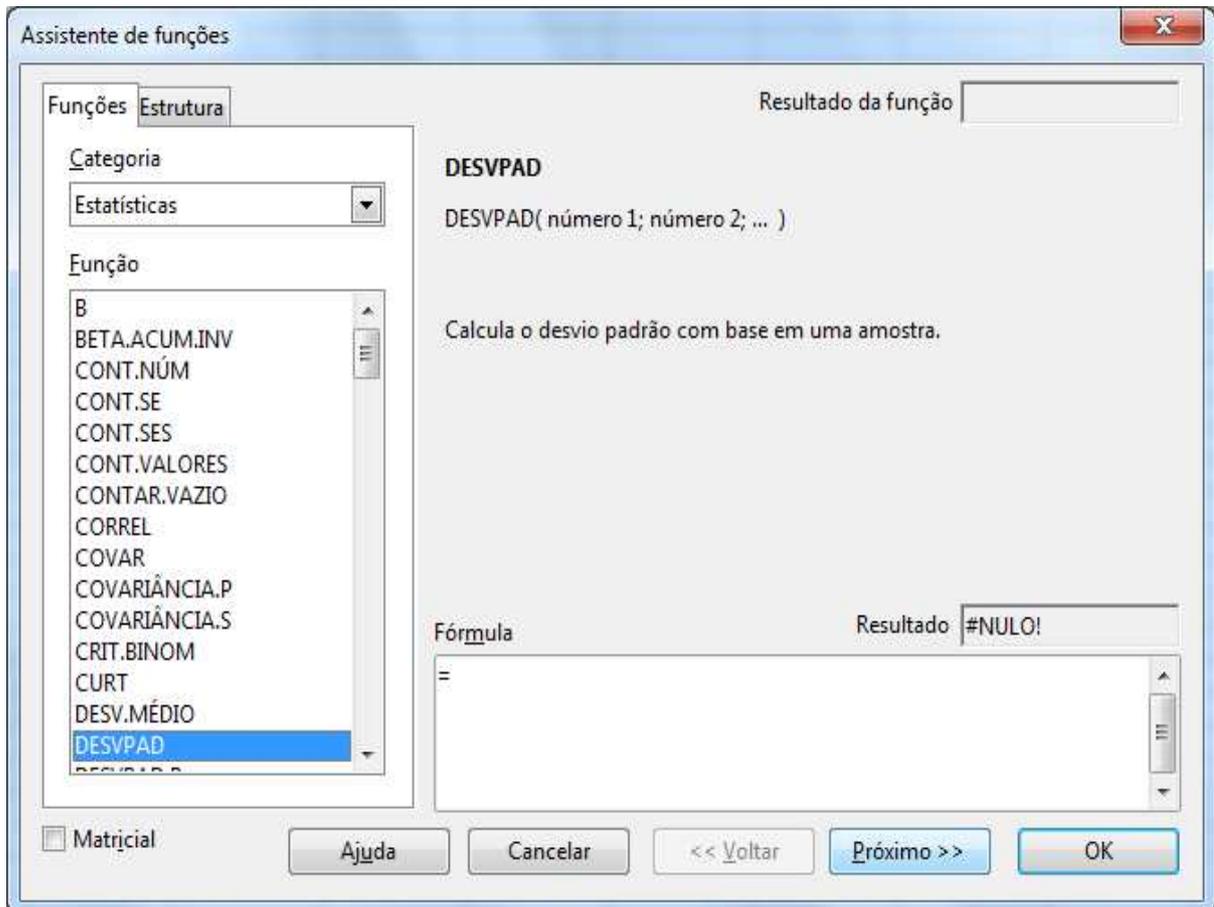


Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Há um assistente de funções que contém mais de cem funções estatísticas e o usuário só precisa selecionar a função que deseja na tela representada na Figura 17, onde clicando na função, esta possui a explicação do que ela representa e dos seus argumentos.

Como pode-se perceber, na Figura 17, foi escolhida a função DESVPAD, que calcula o desvio padrão com base em uma amostra e onde o seu argumento é o intervalo de dados.

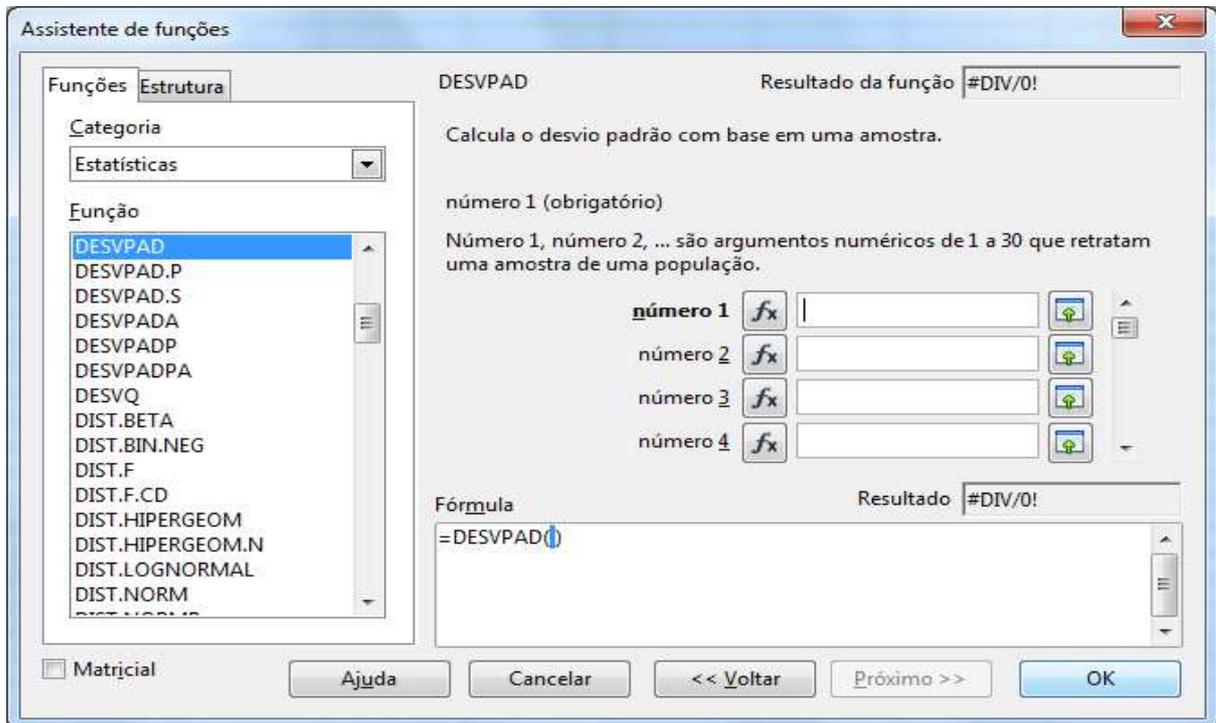
Figura 17 – Cálculo do Desvio Padrão com base em uma Amostra



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Ao clicar em Próximo>>, é preciso preencher os argumentos da função, no caso o intervalo de dados, que pode ser feito selecionando os dados, ou digitando o intervalo na caixa de diálogo no lado direito do número 1 fx. O valor que a função retorna vai aparecer em Resultado da função na parte superior à direita. Ao clicar em OK; esse resultado vai ser o valor da célula selecionada antes de clicar em Inserir. No caso, o resultado da função está dando erro (#Div/0!), pois não possuem dados preenchidos ainda, como mostra a Figura 18.

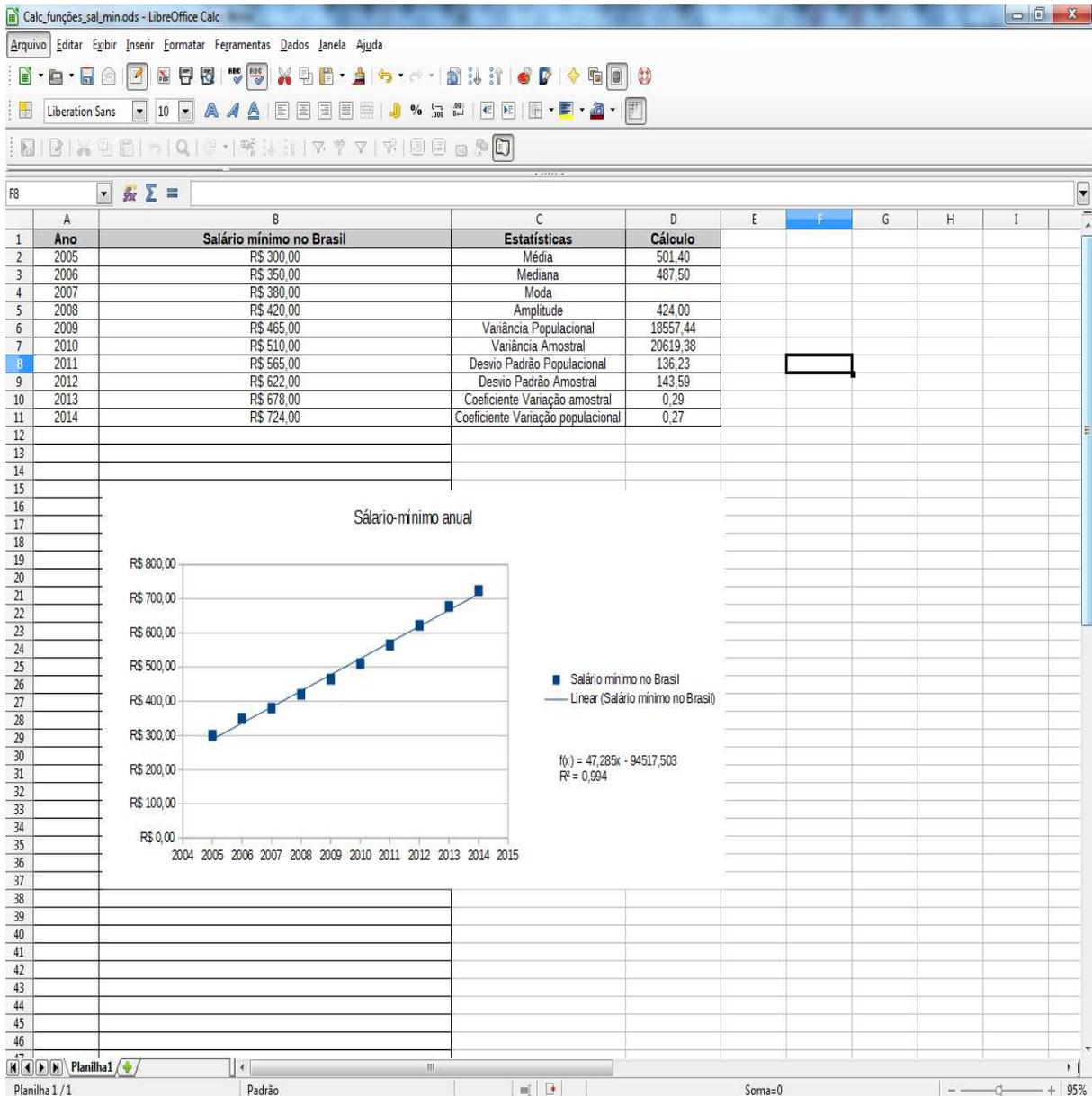
Figura 18 – Função dando Erro



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Entendendo melhor o algoritmo para a geração da planilha, vamos observar o funcionamento dela, com o exemplo do valor histórico do salário-mínimo, desde 2005 até 2014.

Figura 19 – Algoritmo para a Geração da Planilha



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Como pode-se observar na Figura 19, temos quatro colunas na planilha, a primeira, a coluna A, indica o ano vigente, para observarmos uma série temporal, já a coluna B possui o valor do salário-mínimo no Brasil, de acordo com o respectivo ano. A coluna C possui os nomes das estatísticas a serem calculadas na última coluna, a D. Com isso, temos alguns resultados descritivos dos dados, como média do salário-mínimo dos últimos 10 anos igual a R\$ 501,40, mediana igual a R\$ 487,50 e não possui moda, pois por se tratar de salário-mínimo e haver inflação, a ser corrigida anualmente, o valor do salário-mínimo só cresceu nos últimos dez anos, com isso não se teve repetição do valor em anos distintos, portanto essa amostra é amodal. Já as medidas de dispersão apresentam outros dados, como a amplitude igual a R\$ 424,00. Por se tratar apenas dos dados dos últimos dez anos, foi visualizado então a variância amostral igual a

20619,38 e desvio padrão amostral de R\$ 143,59. Com isso se tem um coeficiente de variação amostral, relativamente alto, de 0,29. Portanto, os dados tiveram uma grande variabilidade, o que já era de se esperar, porque o salário-mínimo só aumentou com o tempo.

A fim de observar o comportamento desses dados e fazer previsões do valor do salário-mínimo para o próximo ano, foi feito o gráfico de dispersão, Salário-mínimo anual, onde se tem os valores deles anualmente. Além do gráfico, foi feita uma reta de regressão linear simples, com o crescimento do salário em função do ano, onde se pode fazer estimativas do valor do salário nos próximos anos. A equação da reta está na parte direita do gráfico, $f(x) = 47,285 * X - 94517,503$, onde o X é a variável independente, que indica o ano e $f(x)$ é a estimativa do salário-mínimo para aquele respectivo ano; o coeficiente de determinação, o R^2 foi igual a 0,994, o que indica que o modelo estimado é bom para estimar o valor do salário-mínimo nos próximos anos. Por exemplo, o valor estimado do salário-mínimo, de acordo com esse modelo de regressão linear simples é de:

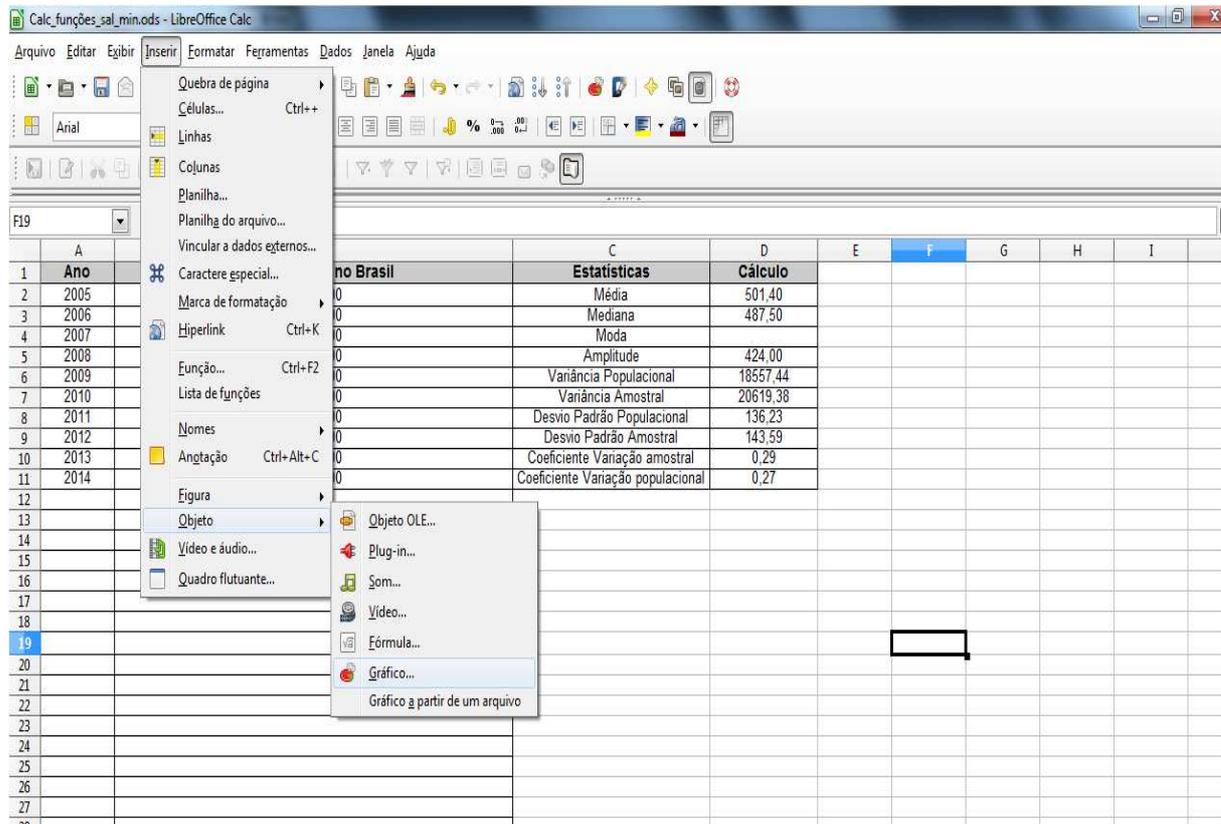
$$\text{Salário estimado de 2015} = 47,285 * (2015) - 94517,503 = 761,772 = \mathbf{R\$ 761,77}.$$

Porém, o mundo não é sempre igual, outras variáveis interferem no crescimento do valor do salário-mínimo, como ter alguns anos com a inflação com oscilações enormes, a mudança de governo, entre outras.

Para efetuar o gráfico e a linha de tendência, com sua equação, segue o passo a passo.

Primeiro, clique em Inserir, depois vá para opção Objeto e, em seguida, clique em Gráfico, como mostra a Figura 20.

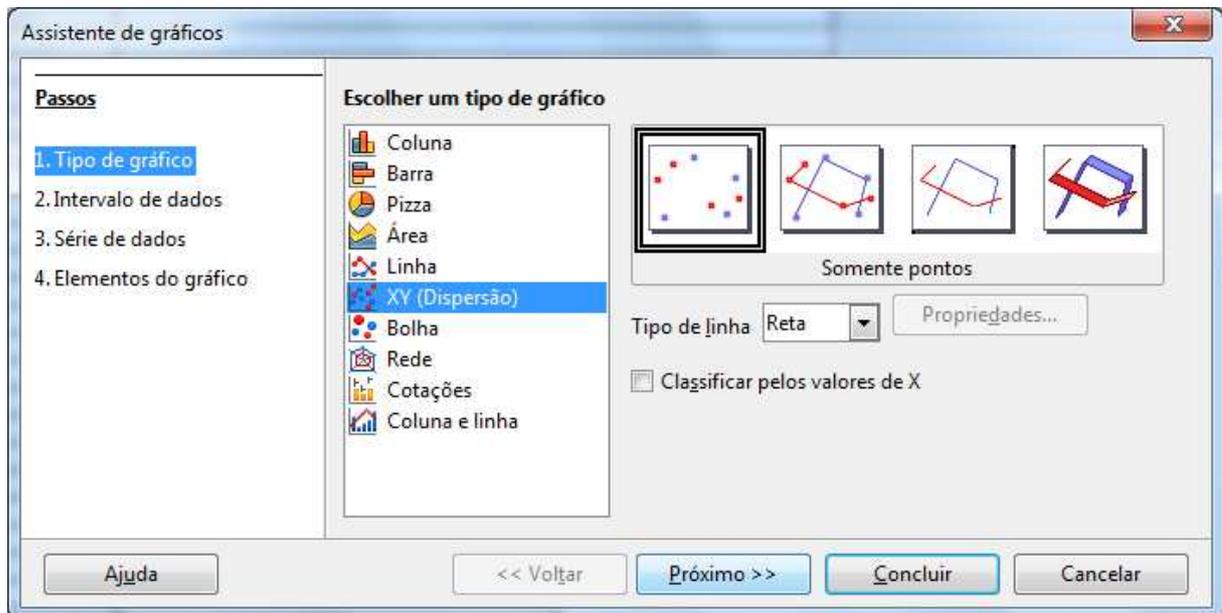
Figura 20 – Efetuação de Gráfico e a Linha de Tendência



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Após clicar em Gráfico, abrirá a janela representada na Figura 21, onde é preciso escolher o tipo de gráfico como XY (Dispersão), somente pontos e clicar em Próximo.

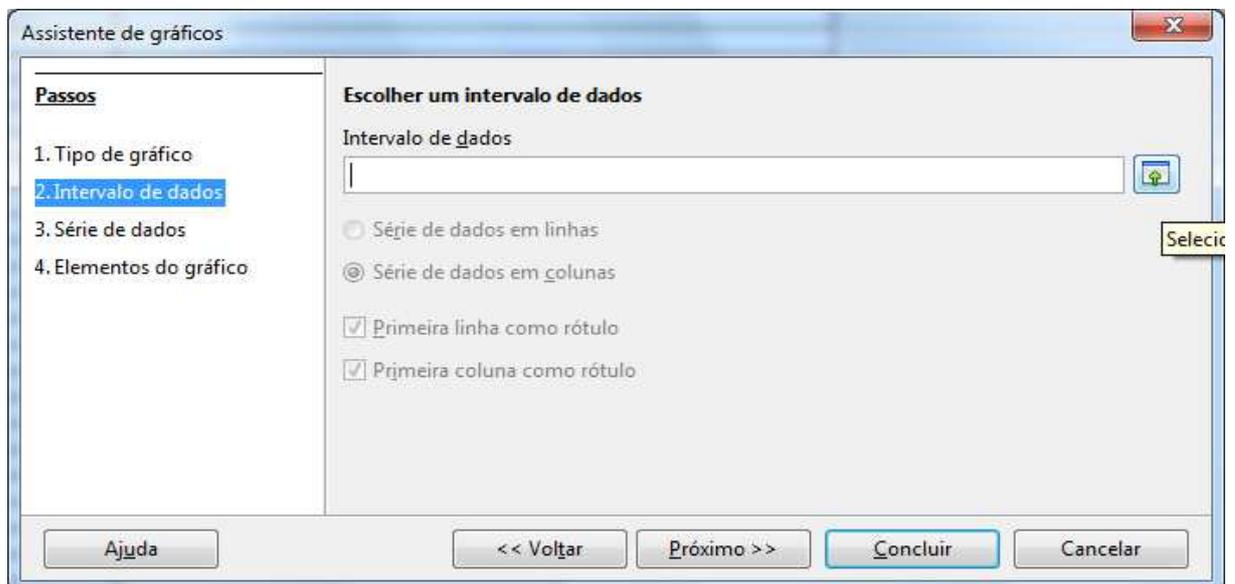
Figura 21 – Tipo de Gráfico



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Quando clicar em Próximo, abrirá uma nova janela, representada na Figura 22, nessa é preciso clicar à direita, para poder selecionar o intervalo de dados.

Figura 22 – Intervalo de Dados 1



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Surgirá então uma janela pequena, como mostra a Figura 23, somente para seleção dos dados, na planilha.

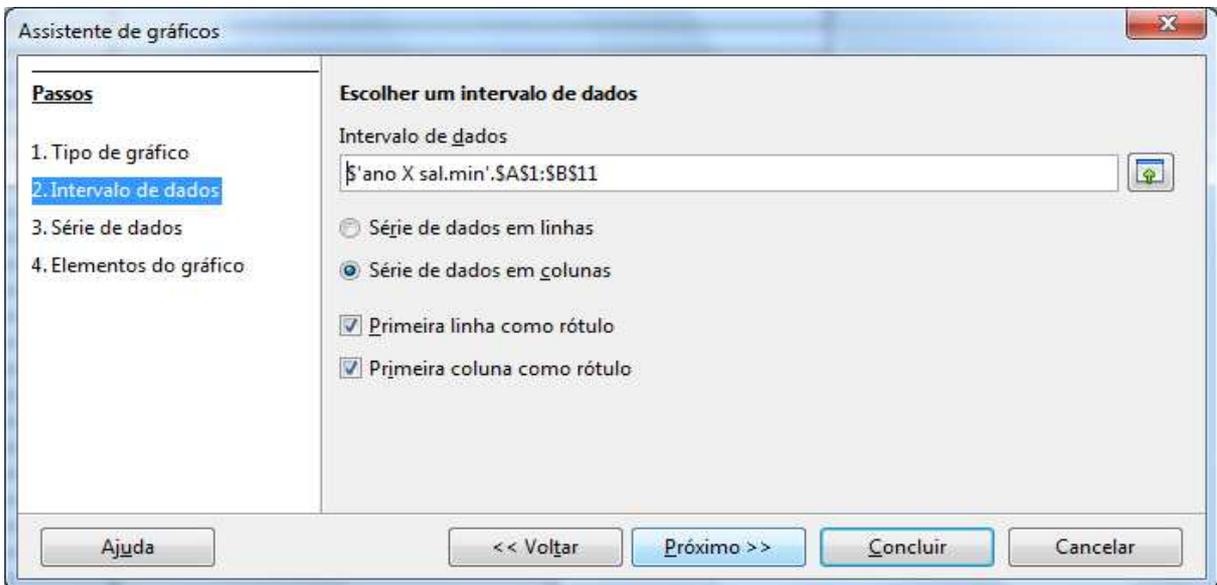
Figura 23 – Intervalo de Dados 2



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Logo, após o passo anterior ter sido efetuado, automaticamente surgirá a janela representada na Figura 24, onde dessa vez essa estará preenchida e então é só clicar em Próximo.

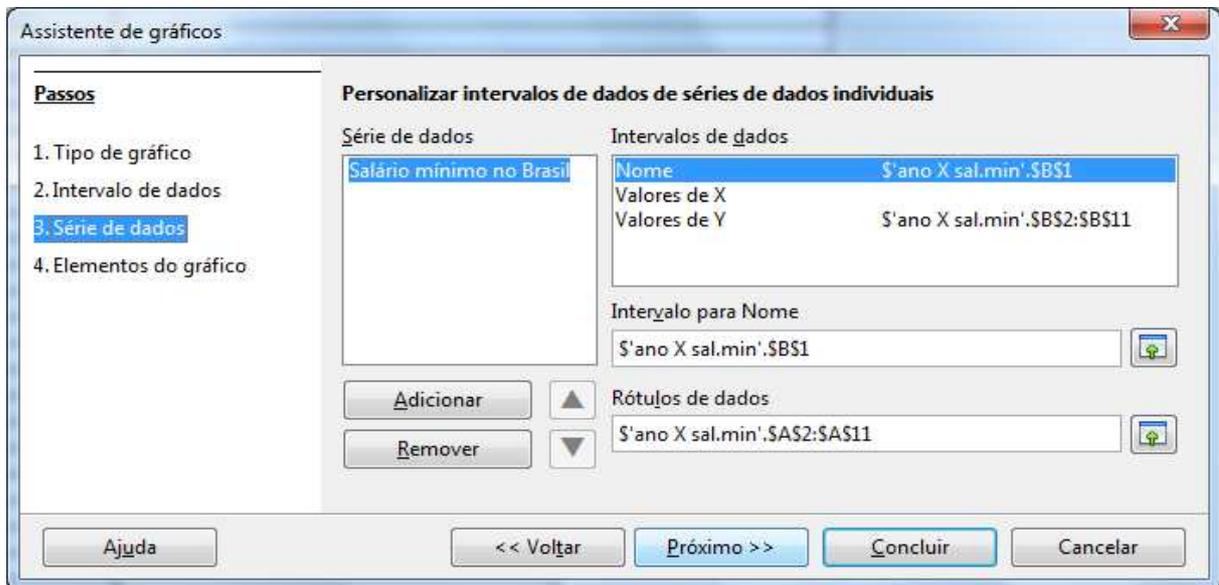
Figura 24 – Intervalo de Dados 3



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

No próximo passo, aparecerá uma nova janela, como representada na Figura 25, onde é possível preencher com o nome para o intervalo e o rótulo dos dados.

Figura 25 – Série de Dados



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Esse novo passo é para pôr os nomes dos eixos e dos títulos do gráfico. Após esse passo é só clicar em Concluir, como mostra a Figura 26 e o gráfico será gerado na planilha.

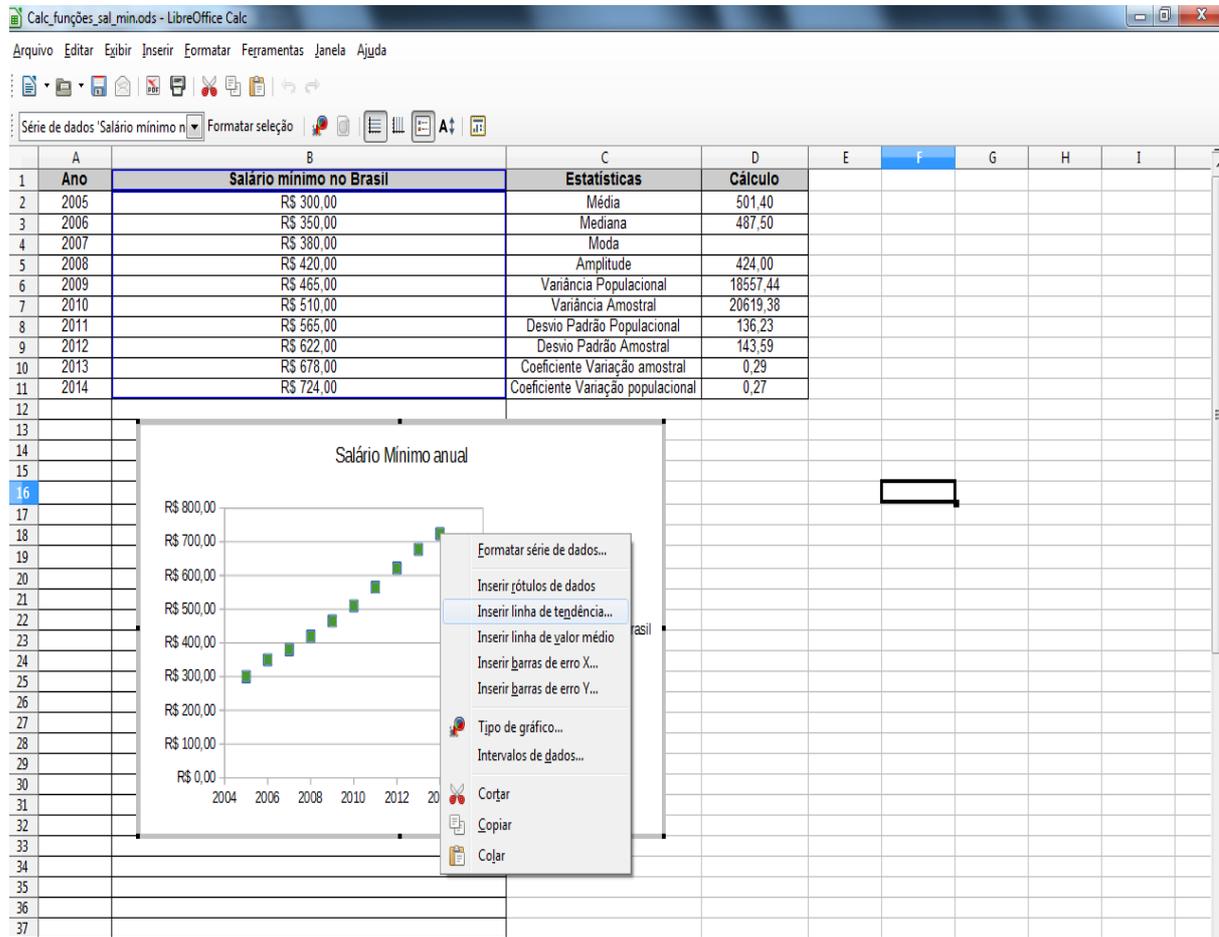
Figura 26 – Elementos do Gráfico



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Agora é preciso selecionar os pontos do gráfico. Clicando em um deles, logo após ter feito essa seleção, clique com o botão direito do mouse nesse ponto, escolhendo a opção Inserir linha de tendência, como mostra a Figura 27.

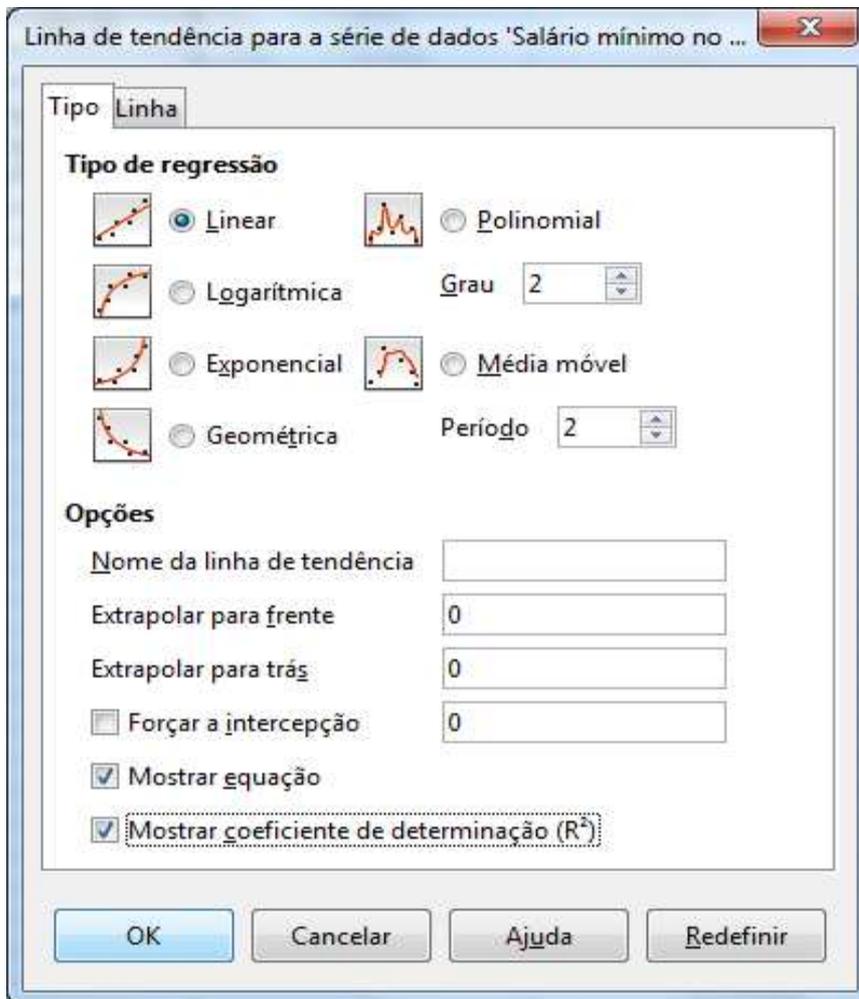
Figura 27 – Pontos Gráficos



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Após clicar, abrirá uma nova janela, representada na Figura 28, nela é preciso selecionar as opções Linear, Mostrar equação e Mostrar coeficiente de determinação (R^2), depois é só clicar em OK.

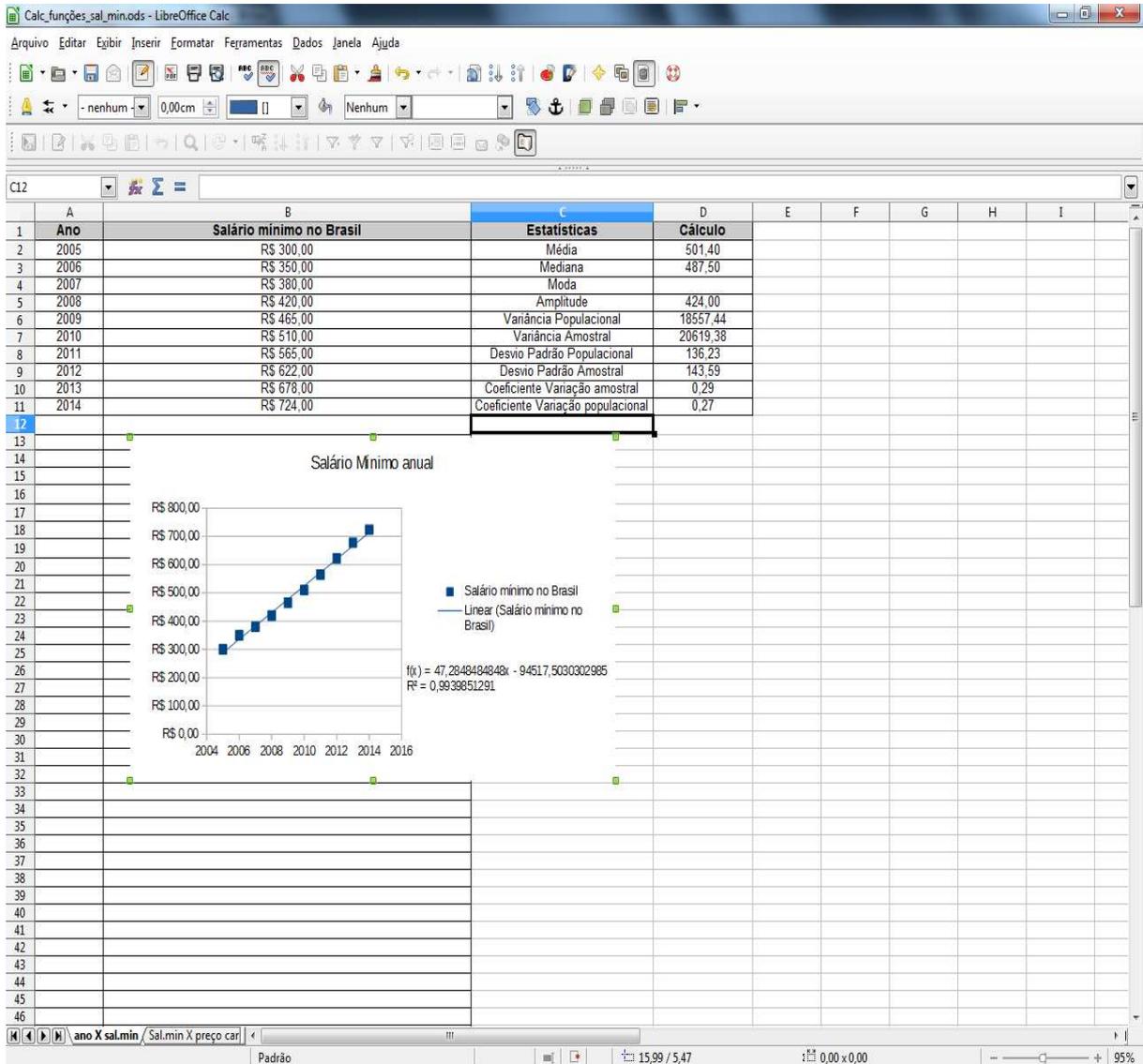
Figura 28 – Linha de Tendência para a Série de Dados



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

E o gráfico estará pronto, já com a reta de regressão linear, com a equação dessa reta e com o coeficiente de determinação dela. Como mostra a Figura 29.

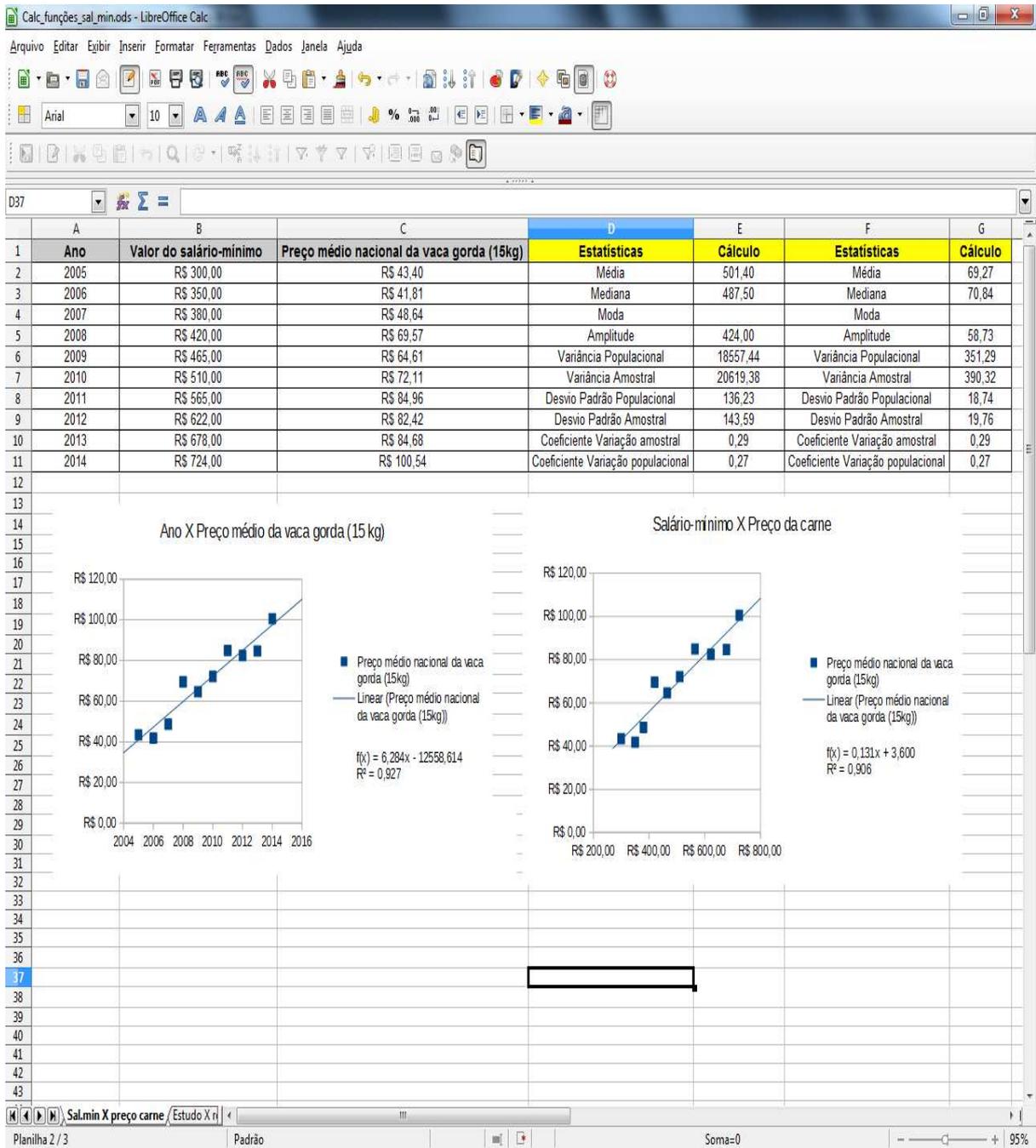
Figura 29 – Regressão Linear



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Outro exemplo interessante é usar o valor do salário-mínimo para estimar o valor da carne de vaca gorda (15 kg), como mostra a Figura 30.

Figura 30 – Estimar preço da carne



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Observando a Figura 30, pode-se perceber que agora estamos trabalhando com três variáveis distintas, onde a primeira é o ano vigente, a segunda é o valor do salário-mínimo em seu respectivo ano vigente e a terceira é o preço médio nacional da vaca gorda (15 kg). Como já vimos anteriormente, as estatísticas descritivas do salário-mínimo nos últimos dez anos, vamos reparar agora nos resultados do preço da carne (vaca gorda 15 kg), sendo assim: média R\$ 69,27; mediana R\$ 70,84; não possui moda, pois não possuem frequências iguais dois anos

distintos, o que é de se imaginar por se tratar de dados quantitativos contínuos, portanto, a amostra é amodal. Olhando agora as medidas de dispersão, temos a amplitude igual a R\$ 58,73, a variância amostral, por se tratar de uma amostra, é de 390,32, com isso o desvio padrão amostral é R\$ 19,76. O coeficiente de variação amostral foi de 0,29. Como há uma semelhança grande nos coeficientes de variação das variáveis salário-mínimo e preço médio nacional da vaca gorda, pode-se concluir que a variabilidade das duas em relação a média é parecida.

Foram feitos também dois gráficos e as respectivas linhas de tendência, onde o primeiro é representado no eixo horizontal o ano vigente e no eixo vertical o preço médio nacional da vaca gorda (15 kg). Pode-se perceber que, quando os anos foram passando, o preço médio da carne foi aumentando, gerando o seguinte modelo de regressão linear simples.

Estimativa do preço médio nacional anual da vaca gorda = $6,284 * (\text{ANO}) - 12558,614$, com um coeficiente de determinação alto, de 0,927, indicando uma alta correlação entre as duas variáveis. Isto indica que o modelo é bom para estimar o preço médio nacional da vaca gorda (15 kg).

Para o ano de 2015, a estimativa do preço médio nacional anual da vaca gorda (15 kg) é de:

$$E(\text{preço da vaca gorda} / X = 2015) = 6,284 * (2015) - 12558,614 = \text{R\$ } 103,65.$$

Como foi visto, o salário-mínimo tinha uma relação bem semelhante ao preço da vaca gorda, quando tínhamos como base o ano vigente, pois quando se passam os anos, o salário-mínimo cresce e o preço médio da carne sobe também. No intuito de reparar essa relação, foi feito o segundo gráfico de dispersão, onde o eixo horizontal é o salário-mínimo e no eixo vertical o preço médio nacional anual da vaca gorda (15 kg), gerando a equação: Estimativa do preço médio nacional da vaca gorda = $0,131 * (\text{Sal. Mín.}) + 3,600$.

Com um coeficiente de determinação de 0,906, portanto, o salário-mínimo e o preço da vaca gorda possuem uma correlação grande, onde como mostra no gráfico, quando um cresce, o outro também aumenta.

Podem ser feitas estimativas do valor médio nacional anual da vaca gorda (15 kg), caso o salário-mínimo seja R\$ 1000,00.

$$E(\text{preço da vaca gorda} / \text{Sal. Mín.} = \text{R\$ } 1000,00) = 0,131 * (1000,00) + 3,6 = \text{R\$ } 134,60.$$

Um exemplo interessante é ver o quanto o grau de instrução interfere na renda das pessoas. Segue a tabela representada na Figura 31, retirada do site do IBGE, com dados de 2010.

Figura 31 – Dados do IBGE quanto o Grau de Instrução e Rendimento das Pessoas

Tabela 3542 - Pessoas de 15 anos ou mais de idade, residentes em domicílios particulares, exceto as cuja condição no domicílio era pensionista, empregado(a) doméstico(a) ou parente do(a) empregado(a) doméstico(a), por nível de instrução, segundo os grupos de idade e as classes de rendimento nominal mensal domiciliar per capita											
Brasil											
Variável = Pessoas de 15 anos ou mais de idade, residentes em domicílios particulares, exceto as pessoas cuja condição na família era pensionista, empregado(a) doméstico(a) ou parente do(a) empregado(a) doméstico(a) (Pessoas)											
Grupos de idade = Total											
Ano = 2010											
Nível de instrução	Classes de rendimento nominal mensal domiciliar per capita										
	Total	Até 1/8 de salário mínimo	Mais de 1/8 a 1/4 de salário mínimo	Mais de 1/4 a 1/2 salário mínimo	Mais de 1/2 a 1 salário mínimo	Mais de 1 a 2 salários mínimos	Mais de 2 a 3 salários mínimos	Mais de 3 a 5 salários mínimos	Mais de 5 a 10 salários mínimos	Mais de 10 salários mínimos	Sem rendimento
Total	143.043.983	5.438.308	8.208.612	23.161.501	39.791.812	33.850.227	11.224.640	8.222.193	5.381.696	2.563.147	5.201.849
Sem instrução e fundamental incompleto	64.492.115	4.145.362	5.727.057	14.136.044	20.674.946	12.330.175	2.630.566	1.258.712	486.775	169.362	2.933.117
Fundamental completo e médio incompleto	27.328.460	868.058	1.529.518	4.893.387	8.433.517	6.974.004	1.905.424	1.090.928	505.181	169.532	958.911
Médio completo e superior incompleto	37.795.869	409.126	909.041	3.909.526	9.727.168	11.924.830	4.512.284	3.114.961	1.661.830	574.242	1.052.862
Superior completo	13.427.539	15.762	42.996	222.544	956.181	2.621.218	2.176.366	2.757.592	2.727.910	1.650.011	256.959

Nota:

- 1 - Os dados são da Amostra.
- 2 - Salário mínimo utilizado: R\$ 510,00.
- 3 - A categoria **Sem rendimento** inclui as pessoas que receberam somente em benefícios.

Fonte: IBGE - Censo Demográfico

Fonte: IBGE – Censo Demográfico.

Repare que a variável nível de instrução não é quantitativa, por isso não é possível calcular média nem medidas de dispersão, mas em compensação ela é qualitativa ordinal, ou seja, mesmo sendo categórica, ainda é possível ordená-la, com isso pode-se calcular além da sua moda, a sua mediana.

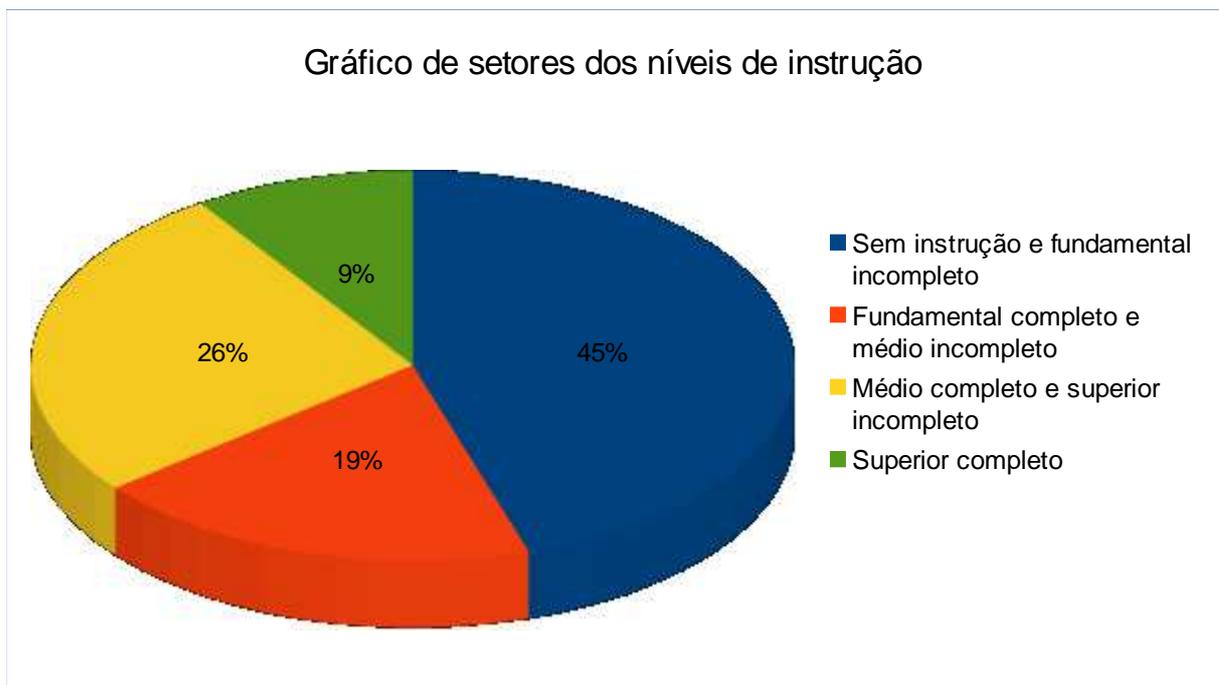
Tabela 6 – Nível de Instrução e Quantidade de Pessoas

Nível de instrução	Pessoas
Total	143.043.983
Sem instrução e fundamental incompleto	64.492.115
Fundamental completo e médio incompleto	27.328.460
Médio completo e superior incompleto	37.795.869
Superior completo	13.427.539

Fonte: Ibge, Censo Demográfico.

Observando a Tabela 6, a moda é sem instrução e fundamental incompleto, com 64.492.115 pessoas, o que corresponde a 45,09 % da população e a mediana é fundamental completo e médio incompleto, pois depois de ordenado, essa classe varia de 45,09% até 64,19%. Somente 9,39% dessa população possui ensino superior completo. O Gráfico 1 mostra a proporção de cada nível de instrução.

Gráfico 1 – Nível de Instrução



Fonte: O Rafael Teixeira Silva, 2014.

Tabela 7 – Rendimento Mensal Domiciliar Per Capita.

Classes de rendimento nominal mensal domiciliar per capita	Pessoas
Total	143.043.983
Sem rendimento	5.201.849
Até 1/8 de salário mínimo	5.438.308
Mais de 1/8 a 1/4 de salário mínimo	8.208.612
Mais de 1/4 a 1/2 de salário mínimo	23.161.501
Mais de 1/2 a 1 de salário mínimo	39.791.812
Mais de 1 a 2 de salários mínimos	33.850.227
Mais de 2 a 3 de salários mínimos	11.224.640
Mais de 3 a 5 de salários mínimos	8.222.193
Mais de 5 a 10 de salários mínimos	5.381.696
Mais de 10 salários mínimos	2.563.147

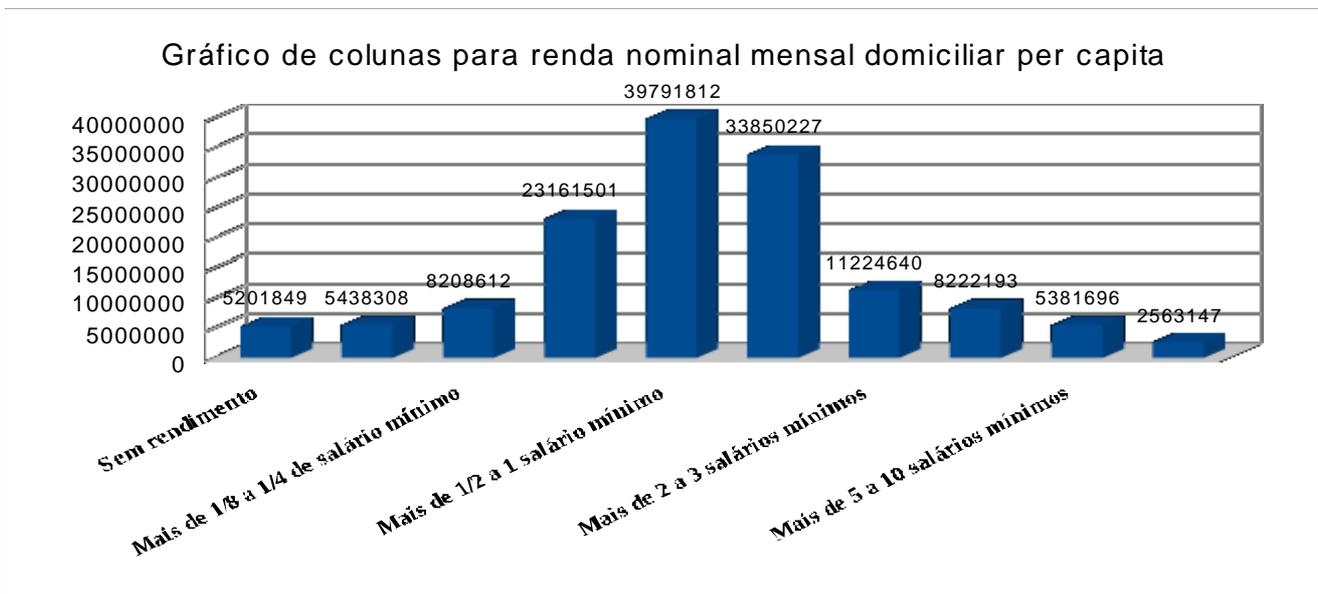
Fonte: Ibge.

Analisando também separadamente a variável classes de rendimento nominal mensal domiciliar per capita, representada na Tabela 7, percebemos que é uma variável qualitativa ordinal, pois os dados estão em classes com intervalos e essas são umas maiores que as outras, podendo assim pôr em uma hierarquia e ser ordenada. Com isso, não é possível se ter a média, nem as medidas de dispersão, mas pode-se calcular sua moda e sua mediana.

A classe modal é mais de meio até um salário mínimo, com 39.791.812 pessoas, o que corresponde a 27,82% dessa população. A mediana foi também a classe de meio a um salário mínimo, pois depois de ordenados os dados, essa classe contém de 29,37% até 57,19%. Apenas 1,79% dessa população tem mais de 10 salários mínimos como rendimento nominal mensal domiciliar per capita.

O Gráfico 2, mostra como estão distribuídas as classes de rendimento nominal, onde pode-se perceber que a maioria dessa população varia entre 1/4 a 2 salários mínimos.

Gráfico 2 – Renda Nominal Mensal Domiciliar Per Capita



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Na variável classe de rendimento, pode ser feita uma transformação na variável, a fim de transformá-la em quantitativa. Para esse caso específico, foi criada uma nova variável, que ao invés de ter os intervalos nas classes, vai se ter a média desses intervalos. Por exemplo, a classe 1/8 a 1/4 de salário mínimo, vai se transformar em $(1/8 + 1/4)/2 = 0,1875$ e assim por diante, representado na Tabela 8.

Tabela 8 – Média da Classe de Rendimentos Nominais

Classes de rendimentos nominais	Média da classe de rendimento nominais
Sem rendimento	0
Até 1/8 de salário mínimo	0,0625
Mais de 1/8 a 1/4 de salário mínimo	0,1875
Mais de 1/4 a 1/2 de salário mínimo	0,375
Mais de 1/2 a 1 de salário mínimo	0,75
Mais de 1 a 2 de salários mínimos	1,5
Mais de 2 a 3 de salários mínimos	2,5
Mais de 3 a 5 de salários mínimos	4
Mais de 5 a 10 de salários mínimos	7,5
Mais de 10 salários mínimos	12,5

Fonte: O Rafael Teixeira Silva, 2014.

Outra variável interessante, que é possível se calcular, é o valor esperado da renda nominal mensal domiciliar per capita, por nível de instrução, utilizando a razão entre, a soma dos produtos do número de pessoas com a média da classe de rendimentos nominais, e o total de pessoas em seus respectivos níveis de instrução. Consideremos o exemplo que se segue.

O valor esperado do rendimento nominal para o total dos níveis de instrução é igual a:

$$(0,0625*5438308 + 0,1875*8208612 + 0,375*23161501 + 0,75*39791812 + 1,5*33850227 + 2,5*11224640 + 4*8222193 + 7,5*5381696 + 12,5*2563147 + 0*5201849) / 143043983 = \mathbf{1,56970042 = Valor esperado de sál. mín. do total.}$$

Assim foi feita a Tabela 9:

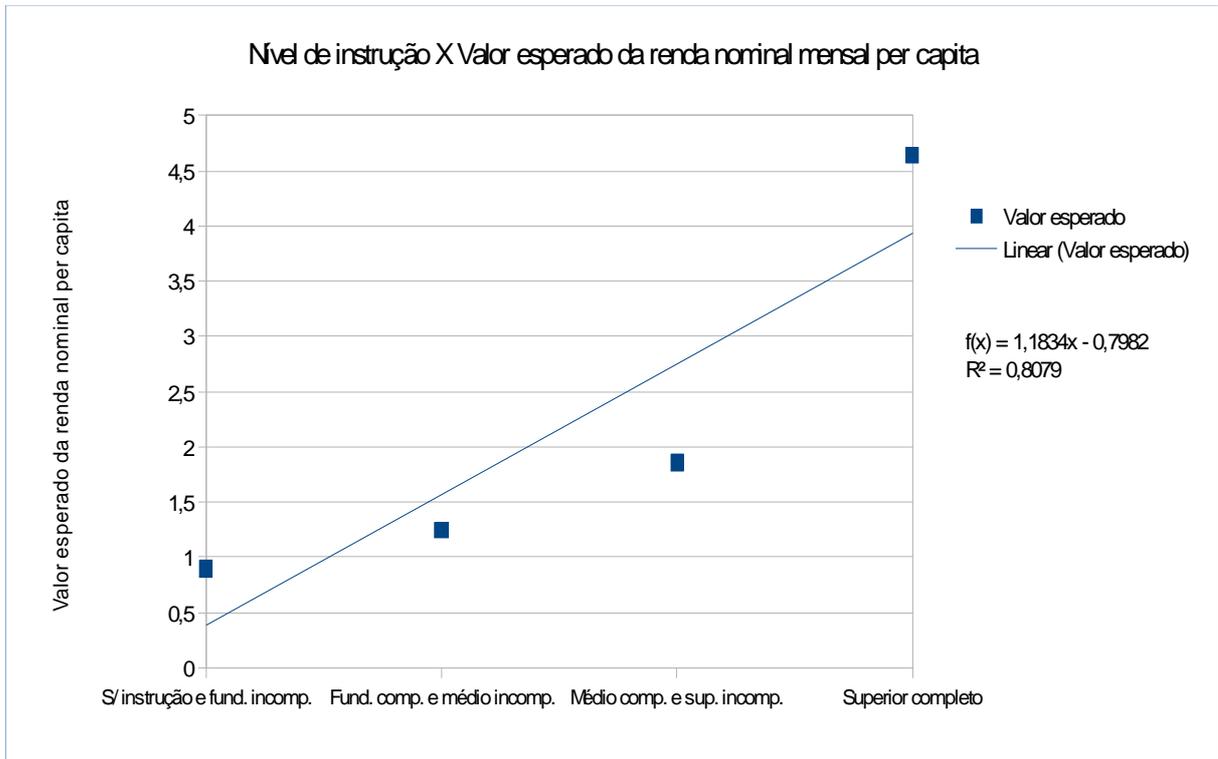
Tabela 9 – Nível de Instrução e Valor esperado Salário Mínimo

Nível de instrução	Valor esperado de sál. Mín.
S/ instrução e fund. incomp.	0,8995595401
Fund. comp. e médio incomp.	1,2440321052
Médio comp. e sup. incomp.	1,8580605413
Superior completo	4,6395105704
Total	1,56970042

Fonte: O Rafael Teixeira Silva, 2014.

Com isso, é possível se fazer um gráfico de dispersão, com o eixo horizontal sendo o nível de instrução e o eixo vertical sendo o valor esperado de salários mínimos domiciliar per capita. Depois inserir a linha de tendência, para se ter uma equação a fim de estimar o comportamento da renda em média, conforme a pessoas vai evoluindo em seus níveis de instrução, gerando assim o Gráfico 3.

Gráfico 3 – Dispersão

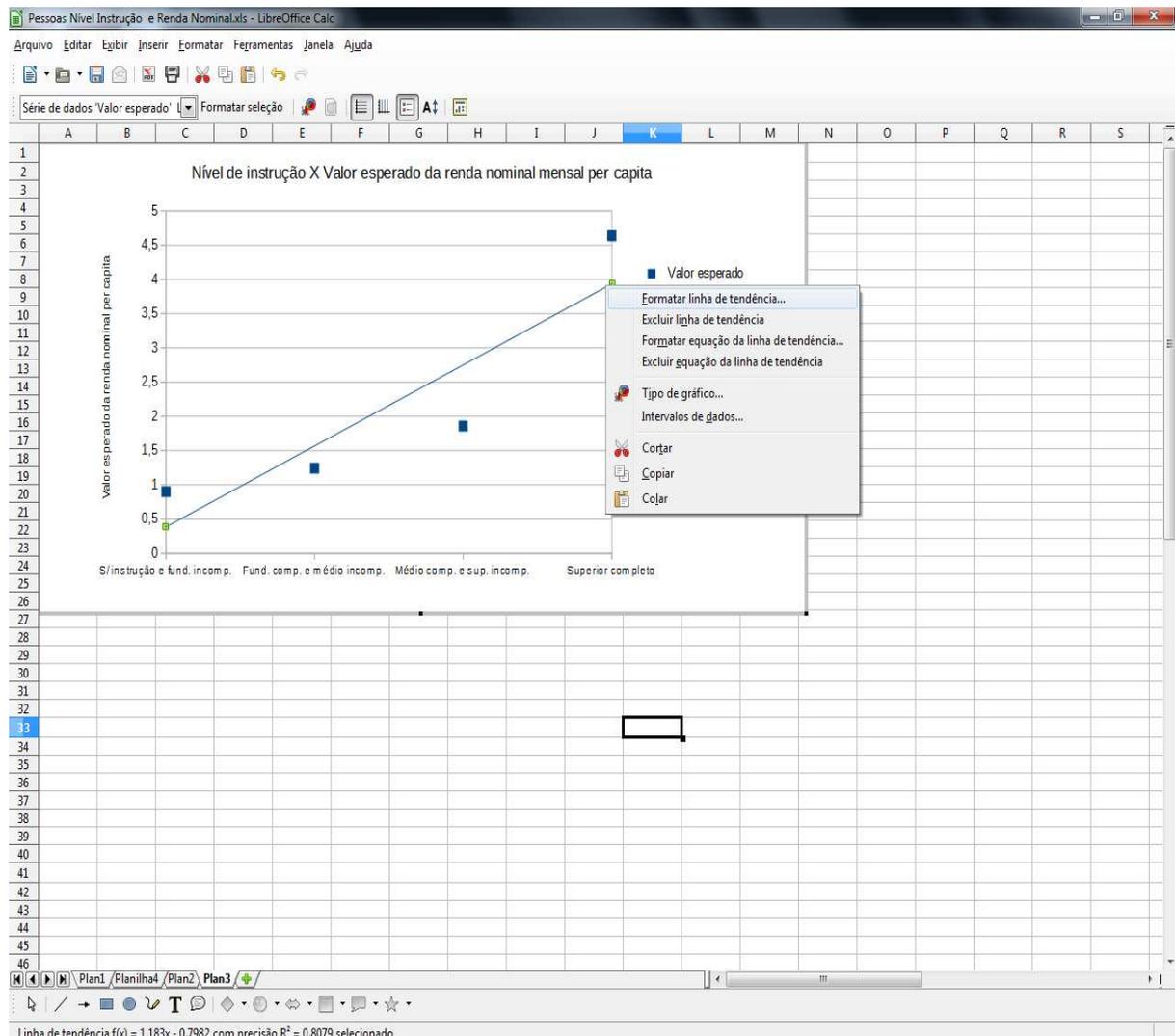


Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Percebe-se que a regressão linear simples gerou um coeficiente de determinação de 0,8079, representando uma boa explicação, porém, ao se olhar os pontos nos dados, parece ter uma distribuição parecida com outro tipo de equação, já conhecida no ensino médio, a equação de segundo grau, gerando assim uma regressão polinomial. Para gerá-la, siga os passos que seguem.

Depois de gerado o Gráfico 3, selecione a linha de tendência, clicando em cima dela. Após esse passo, clique com o botão direito do mouse, e escolha a opção formatar linha de tendência como mostra a Figura 32.

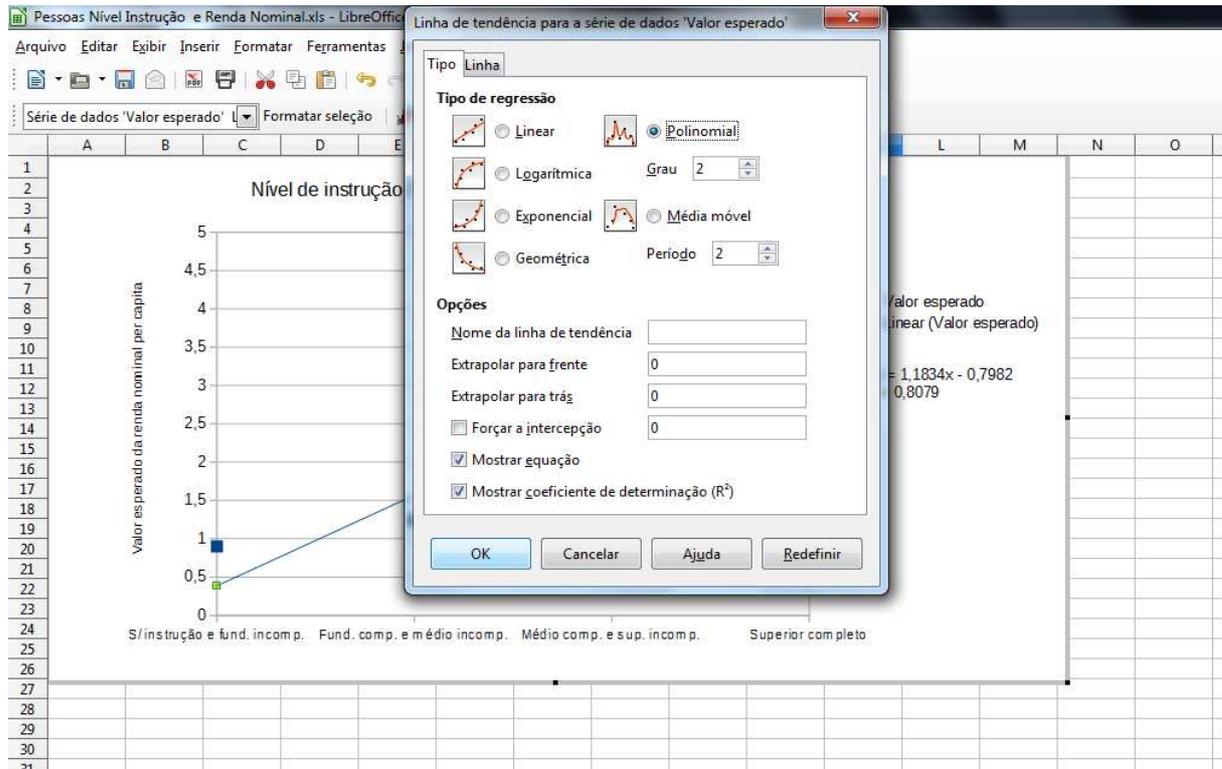
Figura 32 – Pessoas Nível de Instrução e Renda Nominal 1



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Após o clique, surgirá a janela, representada na Figura 33, então é preciso selecionar o tipo de regressão como polinomial, com grau 2, depois clicar em OK.

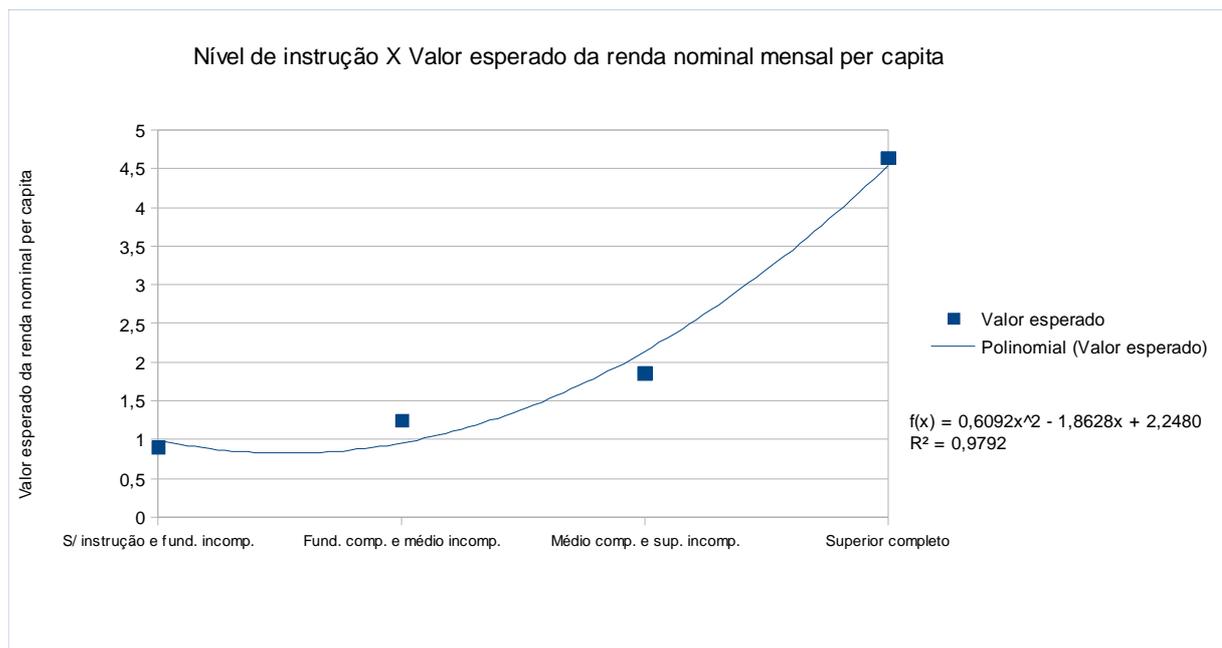
Figura 33 – Pessoas Nível de Instrução e Renda Nominal 2



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

E assim vai ser criada a nova linha de tendência, do tipo de regressão polinomial, como mostra o Gráfico 4.

Gráfico 4 – Linha de Tendência



Fonte: Simulação feita por Rafael Teixeira Silva no LibreOffice Calc, 2014.

Repare que a nova linha de tendência, gerada pela regressão polinomial, se aproxima bem mais do real valor esperado do que a linha linear, onde a linha com equação de segundo grau teve como coeficiente de determinação de 0,9792. Portanto, essa última é melhor para se fazer estimativas. Um cuidado que é preciso se tomar, quando se usa uma regressão polinomial, é que conforme vai se aumentando a variável independente, no caso o nível de instrução, a variável dependente tende a explodir, ou seja, tende a assumir valores muito aberrantes, o que faz com que seja possível a regressão linear ficar melhor do que a polinomial, mas no caso presente a polinomial se mostrou superior.

5 PLANOS DE TRABALHO COM USO DE *SOFTWARE* NO ENSINO DE ESTATÍSTICA

Neste capítulo serão apresentados dois planos de trabalho propondo o ensino de estatística, para alunos do ensino médio com o uso software livre Calc.

5.1 Introdução

O estudo de técnicas estatísticas contribui para interpretar e modelar diversas situações-problema no dia a dia, como questões matemáticas, fenômenos físicos, econômicos, biológicos, entre outros.

O presente plano de trabalho pretende refletir e propor que o processo de aprendizagem das técnicas estatísticas podem ser mais bem assimiladas com a utilização das Tecnologias da Informação e Comunicação através do uso de computadores e do software LibreOffice Calc, complementados com o uso de outros recursos como o da Calculadora em sala de aula.

O plano de trabalho das técnicas de interpretações de dados será desenvolvido em seis tempos de cinquenta minutos para o desenvolvimento dos conteúdos e avaliação da aprendizagem, sendo que dois deles serão realizados no laboratório de informática e quatro em sala de aula.

Ao final das aulas de estatística básica, o aluno deverá alcançar as seguintes competências:

- Entender a diferença entre população e amostra, sabendo quando utilizar cada uma delas e reconhecer de que tipo é cada informação (dados).
- Interpretar de forma correta, gráficos de barras, setores, colunas, dispersão, linhas e histogramas.
- Calcular estatísticas descritivas dos dados.
- Aplicar as técnicas de estatística, como as de regressão linear em situações-problemas do cotidiano.

5.2 Desenvolvimento

As técnicas estatísticas para interpretação de dados, serão desenvolvidas em quatro atividades.

5.2.1 Atividade 1

- Competência: O aluno entender o que é estatística e suas aplicabilidades, conhecer tipos de dados e diferença entre amostra e população.
- Pré-requisitos: Conjunto dos números reais e seus subconjuntos.
- Tempo de duração: 50 minutos.
- Recursos utilizados: Livro didático.
- Organização da turma: Individual
- Objetivos: Apresentar o que é estatística e para que serve, indicando diferentes tipos de dados e diferenças entre usar a população ou a amostra.
- Metodologia: Apresentação de variados tipos de dados, como quantitativos (discreto ou contínuo), qualitativos (nominal ou ordinal), população e amostra.

5.2.2 Atividade 2

- Competência: Interpretar os gráficos e conhecer tipos diferentes desses.
- Pré-requisitos: Conhecimentos de plano cartesiano, pontos mínimos e máximos.
- Tempo de duração: 50 minutos.
- Recursos utilizados: Livro didático, jornal, revista, software LibreOffice Calc.
- Organização da turma: Individual
- Objetivos: Apresentar diferentes tipos de gráficos de diferentes tipos de dados.
- Metodologia: Apresentação em sala de gráficos de colunas, barras, setores, linhas e histogramas presentes em jornais, livros, revistas, etc.

5.2.3 Atividade 3

- Competência: Calcular estatísticas descritivas, como medidas de posição e de variação.
- Pré-requisitos: Aritmética.
- Tempo de duração: 100 minutos.
- Recursos utilizados: LibreOffice Calc e calculadora.
- Organização da turma: Turma dividida em grupos de 2 alunos, tornando favorável o trabalho colaborativo.
- Objetivos: Calcular as medidas de posição, variação e saber analisar.

5.2.4 Atividade 4

- Competência: Auxiliar o aluno a perceber o formato do gráfico de dispersão de duas variáveis.
- Pré-requisitos: Plano Cartesiano; Gráficos de Funções
- Tempo de duração: 100 minutos.
- Recursos utilizados: LibreOffice Calc ; Laboratório de Informática.
- Organização da turma: Turma dividida em grupos de 2 ou 3 alunos tornando favorável o trabalho colaborativo.
- Objetivos: : Entender o gráfico de dispersão das duas variáveis, observando tendências e reparando o comportamento de uma variável em função da outra, utilizando técnicas estatísticas para estimar valores.
- Metodologia: A atividade será realizada no laboratório de informática da escola após instalado em todos os computadores o LibreOffice, utilizar funções do software para a realização de regressões lineares e obtenções de gráficos.

5.3 Primeiro plano de aula, utilizando as três primeiras atividades

1) Escolha aleatoriamente apenas 10 pessoas da turma e pergunte para que time elas torcem.

Em seguida responda:

- a) Classifique os dados quanto os seus tipos.
- b) É uma amostra ou uma população da turma e qual o seu tamanho?
- c) Qual o valor da moda? E de que tipo é ela?

- d) Qual o valor da mediana?
- e) Qual o valor da mediana?
- f) Qual o valor de suas medidas de dispersão?
- f) Esboce um gráfico de setores que represente esses dados.

2) Escolha aleatoriamente apenas 10 alunos da sala de aula e pergunte sua idade, com esse subconjunto da turma, responda:

- a) Classifique os dados quanto os seus tipos.
- b) É uma amostra ou a população e qual o seu tamanho?
- c) Qual valor da moda, sua unidade de medida e de que tipo é?
- d) Qual o valor da média aritmética e sua unidade de medida?
- e) Qual o valor da mediana e sua unidade de medida?
- f) Qual o valor da amplitude e sua unidade de medida?
- g) Qual o valor da variância e sua unidade de medida?
- h) Qual o valor do desvio padrão e sua unidade de medida?
- i) Qual o valor do coeficiente de variação e sua unidade de medida?
- j) Faça um gráfico de colunas que represente os dados.

3) Repare que foi feita muita conta no exercício anterior. A fim de efetuar esses cálculos de forma muito mais rápida, vamos utilizar o software LibreOffice Calc no laboratório e ver os resultados desse segundo exercício. Utilize a planilha feita no capítulo 4, colocando na coluna A , a partir da célula A2, os nomes dos 10 alunos e na célula A1 digite “Nomes”. A coluna B vai ser preenchida com a idades em anos do respectivo aluno da coluna A, ou seja, digite na célula B1 “Idades” e a partir da B2 preencha com a idade do respectivo aluno. A planilha automaticamente vai efetuar os cálculos das estatísticas descritivas. Observe a velocidade com a qual foi feita esses cálculos e compare o resultado com os feitos a mão.

4) Ainda utilizando a planilha, adicione nos seus dados o professor, digitando na célula A12 o nome dele e na B12 sua idade.

- a) O que acontece com média e por que?
- b) O que acontece com a moda e por que?
- c) O que acontece com a mediana e por que?
- d) O que acontece com as medidas de dispersão e por que?
- 5) Agora adicione o nome e a idade de todos os alunos da turma.

- a) As medidas de posição ficaram parecidas com a amostra da turma?
- b) As medidas de dispersão ficaram parecidas com a amostra da turma?
- c) Faça um gráfico de colunas que represente os dados.
- d) Compare os gráficos com 10 alunos e com toda a turma.

6) Utilize dados qualitativos e quantitativos de jornais, revistas ou da internet e preencha a planilha, analisando as estatísticas descritivas e seus gráficos, tentando entender esses dados apenas olhando esses resultados estatísticos.

5.4 Segundo plano de aula, utilizando as quatro atividades

- 1) A Tabela 10 mostra a frequência absoluta das pessoas segundo sua idade.

Tabela 10 – Exemplo da idade de um grupo de pessoas

Idade em anos	Quantidade de pessoas (Freq. Absoluta)
15	100
16	100
17	100
18	100
19	100

Fonte: O Rafael Teixeira Silva, 2014.

- a) Qual o tipo da variável idade?
- b) Qual a é a média aritmética da variável?
- c) Qual é a mediana?
- d) Qual é a moda?
- e) Quais são os valores das medidas de dispersão?
- f) Faça um gráfico de colunas para a variável idade.

- 2) A Tabela 11 contém a frequência absoluta das pessoas segundo sua classe de altura.

Tabela 11 – Exemplo da altura de um grupo de pessoas

Altura em metros	Quantidade de pessoas
Entre 1,51 e 1,60	120
Entre 1,61 e 1,70	175
Entre 1,71 e 1,80	120
Entre 1,81 e 1,90	85

Fonte: O Rafael Teixeira Silva , 2014.

- a) Qual o tipo de variável de classes de altura?
- b) Qual a é a média aritmética da variável?

- c) Qual é a mediana?
- d) Qual é a moda?
- e) Quais são os valores das medidas de dispersão?
- f) Faça um gráfico de setores para a variável classes de altura.
- g) É possível transformar essa variável em quantitativa? Se sim, dê uma solução.

1) Dada a Tabela 12, com a frequência absoluta das pessoas por idade, segundo a classe da sua altura, responda as questões abaixo.

Tabela 12 – Exemplo de intervalo da altura das pessoas segundo a idade

Idade em anos	Altura em metros				Total
	Entre 1,51 e 1,6	Entre 1,61 e 1,7	Entre 1,71 e 1,8	Entre 1,81 a 1,9	
Total	120	175	120	85	500
15	40	30	20	10	100
16	30	40	15	15	100
17	25	40	20	15	100
18	15	35	30	20	100
19	10	30	35	25	100

Fonte: O Rafael Teixeira Silva , 2014.

- a) Transforme a variável altura em quantitativa. Qual será o valor de cada classe?
 - b) Qual o valor esperado da altura, para cada idade?
 - c) Faça um gráfico de linha, só com pontos, do valor esperado da altura, segundo a idade.
 - d) Faça a linha de tendência, utilizando o método de regressão linear simples, comentando a equação e seu coeficiente de determinação.
 - e) Qual a estimativa para o valor esperado da altura de uma pessoa de 20 anos?
 - f) Qual a estimativa para o valor esperado da altura de uma pessoa de 50 anos? É uma boa estimativa? Por que?
 - h) Agora selecione a linha de tendência do gráfico, clique com o botão direito e escolha opção formatar linha de tendência. Ao abrir a nova janela, selecione o tipo de regressão logarítmica, repare a diferença e compare as duas regressões, mas qual vai estimar melhor os dados para idades grandes, por exemplo, 50 anos? Por que?
- 1) Procure em revistas, jornais ou na internet exemplos de duas variáveis em que uma seja dependente da outra, fazendo assim gráficos e estatísticas descritivas desses dados, para, por fim, fazer uma regressão linear simples utilizando o método dos mínimos quadrados.

CONCLUSÕES

A estatística está cada vez mais presente no cotidiano das pessoas, seja para ler um jornal, assistir um programa de TV, na escola, na internet, em redes sociais, ou até mesmo para fazer compras no supermercado. A nossa sociedade possui cada vez mais demandas de dados, de diversas formas e classes, e é cada vez mais desejável interpretar e entender esses dados do cotidiano, a fim de otimizar decisões. Essa análise pode interferir diretamente no voto eleitoral, que produto vai comprar no mercado, qual escola vai matricular seu filho, para a simples análise se concordamos ou não com a demissão de um técnico de futebol ou da venda de um jogador de futebol. Por estarmos rodeados de dados e estatísticas, cada vez mais, se faz necessário a inclusão de estudos de técnicas para interpretar e analisar esses dados no ensino básico. Na atualidade, os jovens já começam a ter acesso a informações, gráficos, tabelas, quadros e dados, em geral, havendo assim a necessidade de saber interpretá-los.

A percepção do aumento da inclusão da estatística na sociedade e na visão da população determina, dentre outros, à maioria dos cursos de ensino superior possuírem grades de estatística, mesmo que básica ou voltadas especificamente para a especialidade do curso. Para facilitar e aprimorar o aprendizado dos futuros graduandos, mais uma vez, se vê imprescindível a visão estatística no ensino básico, não só com fórmulas e contas, mas principalmente a utilização de dados do dia a dia das pessoas, para os estudantes conseguirem perceber aplicações da matemática estatística e assim conseguirem adquirir uma afeição maior pelo estudo das ciências exatas e não ficando com medo do aprendizado matemático.

Para ajudar prover essa demanda de análise de informações, foi feito nesse estudo a apresentação de algumas técnicas estatísticas, tipos de dados, gráficos, medidas de posição, de dispersão e modelos de regressão linear simples, onde este último, mesmo não estando na ementa do ensino básico, pode ser interpretado por conhecimentos dos alunos em função do primeiro grau e sua aprendizagem significativa.

Como para fazer estatísticas, em geral, é preciso fazer muitos cálculos, o que pode ficar cansativo aos alunos e utilizar muito tempo de aula, uma ajuda para a fluidez da aula é a utilização do laboratório de informática das escolas, no intuito de utilizar o software LibreOffice Calc, para agilizar os cálculos, fazerem simulações e o aluno analisar com maior perícia os dados. As vantagens de utilização desse software são: por ser gratuito, pode ser baixado na internet; por trabalhar em planilhas, ter uma utilização grande pelo menu e obter bastantes fórmulas e funções estatísticas já prontas nele. Tudo isso facilita o acesso e a utilização do LibreOffice Calc por qualquer pessoa leiga.

O presente trabalho contém um produto, também feito no LibreOffice Calc, para utilização de qualquer pessoa, com qualquer tipo de variável quantitativa, onde está explícito o algoritmo para a fabricação da planilha e utilização dela. São vistos alguns exemplos de dados,

em que a planilha faz gráfico, calcula estatísticas descritivas dos dados e gera linhas de tendência dos dados.

Por fim, foi apresentado nesse estudo, dois planos de aula com atividades e exercícios para serem feitos no laboratório de informática, com a utilização do LibreOffice Calc. Assim, é preciso fazer uma avaliação, que está integrada às práticas pedagógicas, voltada para a autonomia e a consciência crítica dos alunos. O professor tem o papel de fazer aprender e não simplesmente ensinar, mas através de uma relação democrática, dialogada, de troca, de reciprocidade, de negociações na relação professor-aluno, desenvolver capacidades de modo que o aluno possa se auto-avaliar, aprender a aprender, analisar e superar seus erros.

O processo avaliativo não deve estar delimitado em etapas: início, meio e fim, mas por momentos contínuos e simultâneos de mobilização de experiências educativa e expressão do conhecimento por educadores e educandos, momentos provisórios e complementares que só podem ser analisados em seu conjunto.

O aluno será avaliado de modo integral, o que inclui conhecimentos (saberes), habilidades (saber-fazer), valores e atitudes (saber-ser), pressupondo a utilização de diferentes procedimentos propostos nas atividades neste plano de trabalho.

Outras pesquisas voltadas para o processo de aprendizagem da estatística na educação básica precisam ser desenvolvidas para que o professor de matemática se sinta mais seguro em trabalhar com seus alunos os conteúdos de estatística, cumprindo dessa forma as orientações que constam nos Parâmetros Curriculares.

REFERÊNCIAS

- BEHRENS, Marilda Aparecida. Projetos de aprendizagem colaborativa num paradigma. In: BEHRENS, M.P., MASETTO, M.T., MORAN, J.M. *Novas tecnologias e mediação tecnológica*. Campinas, SP: Papyrus, 2000.
- BRASIL. Ministério da Educação e Desporto. Secretaria do Ensino Fundamental. *Parâmetros curriculares nacionais: matemática*. Brasília: MEC / Secretaria do Ensino Fundamental, 1997.
- BRASIL. Ministério da Educação e Desporto. Secretaria do Ensino Fundamental. *Parâmetros curriculares nacionais (5ª a 8ª séries): Matemática*. Brasília: MEC / Secretaria de Ensino Fundamental, 1998.
- BRASIL. Ministério da Educação e Desporto. Secretaria de Educação Média e Tecnológica. *Parâmetros curriculares nacionais ensino médio: orientações educacionais complementares aos parâmetros curriculares nacionais – Ciências da Natureza, Matemática e suas Tecnologias*. Brasília: MEC / Secretaria de Educação Média e Tecnológica, 2002.
- HOFFMANN, J. *Avaliação: mito & desafio*. Uma perspectiva construcionista. Porto Alegre, RS: Mediação, 2003.
- LOPES, Celi Espasandin. Os desafios para a educação estatística no currículo de Matemática. In: ALMOULOU, S.A., COUTINHO, C.Q.S., LOPES, C.E. (org.). *Estudos e reflexões em educação estatística*. Campinas, SP: Mercado de Letras, 2010.
- MASETTO, Marcos T. Mediação pedagógica e o uso da tecnologia. In: BEHRENS, M.P., MASETTO, M.T., MORAN, J.M. *Novas tecnologias e mediação tecnológica*. Campinas – SP. Papyrus, 2000.
- MASINI, E.F.S e MOREIRA, M.A. *Aprendizagem significativa: a teoria de David Ausubel*. São Paulo. Centauro, 2001.
- MORAN, José Manuel. Ensino e aprendizagem inovadores com tecnologias audiovisuais e telemáticas. In: BEHRENS, M.P., MASETTO, M.T., MORAN, J.M. *Novas tecnologias e mediação tecnológica*. Campinas – SP. Papyrus, 2000.
- MOREIRA, M.A. *Aprendizagem significativa*. Brasília: Editora Universidade de Brasília, 1999.
- NOGUEIRA, Lemerton Matos. *A inclusão da Estatística como componente curricular no ensino médio, nas escolas estaduais baianas*. VII Congresso Norte Nordeste de Pesquisa e Inovação – CNNEPI. Palmas – Tocantis – 2012. Disponível em: <<http://propi.ifto.edu.br/ocs/index.php/connepi/vii/paper/viewFile/5612/908>>. Acesso em: 23 de fevereiro de 2014.
- PAIVA, M. *Matemática Paiva*. 1. ed. São Paulo: Moderna, 2009, 1 ano.

ROTEIROS DE AÇÃO – *Função polinomial do 2 grau* – Curso de aperfeiçoamento oferecido por CECIERJ referente ao 1º ano do Ensino Medio – 3º bimestre/2013 Disponível em: <<http://projetoeduc.cecierj.edu.br/>>. Acesso em: 25 de agosto de 2013.

SPRITZER, Ilda Maria de Paiva Almeida. A tecnologia e a educação. In: BITTENCOURT, P.C. e SPRITZER, I.M.P.A. *Educação tecnológica: tecnologias da informação e comunicação: Módulo 2*. Rio de Janeiro. CEFET/RJ, 2009.

MÉTODO dos mínimos quadrados. Disponível em: <http://pt.wikipedia.org/wiki/M%C3%A9todo_dos_m%C3%ADnimos_quadrados>. Acesso em: 20 de setembro de 2014.

AGRO link Cotações. Disponível em: <<http://www.agrolink.com.br/cotacoes/Historico.aspx?e=9839&p=1772&l=13142>>. Acesso em: 25 de setembro de 2014.

INSTITUTO Brasileiro de Geografia e Estatística. Disponível em: <www.ibge.gov.br>. Acesso em: 26 de setembro de 2014.

BAKER, John; SUGDEN, Stephen J. Spreadsheets in Education –The First 25 Years. *Spreadsheets in Education*, v. 1, n. 1, 2007.