



**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**  
**CAMPUS PONTA GROSSA**  
**DEPARTAMENTO DE PÓS-GRADUAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO**  
**PPGEP**

**CELSO BILYNKIEVYCZ DOS SANTOS**

**ANÁLISE DOS RESULTADOS DO *WHOQOL-100***  
**UTILIZANDO *DATA MINING***

**PONTA GROSSA**  
**DEZEMBRO - 2007**

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

**CELSO BILYNKIEVYCZ DOS SANTOS**

**ANÁLISE DOS RESULTADOS DO *WHOQOL-100*  
UTILIZANDO *DATA MINING***

Dissertação apresentada como requisito parcial à obtenção do título de Mestre em Engenharia de Produção, do Programa de Pós-Graduação em Engenharia de Produção, Área de Concentração: Gestão Industrial, do Departamento de Pesquisa e Pós-Graduação, do Campus Ponta Grossa, da UTFPR.

**Orientador:** Prof. Luciano Scandelari, Dr.

**PONTA GROSSA**

**DEZEMBRO – 2007**

S237 Santos, Celso Bilynkiewicz dos  
Análise dos resultados do WHOLQOL-100 utilizando Data Mining. / Celso  
Bilynkiewicz dos Santos. -- Ponta Grossa: [s.n.], 2007.  
108 f.: il. ; 30 cm.

Orientador: Prof. Dr. Luciano Scandelari

Dissertação (Mestrado em Engenharia da Produção) - Universidade Tecnológica  
Federal do Paraná, Campus Ponta Grossa. Programa de Pós-Graduação em Engenharia  
de Produção. Ponta Grossa, 2007.

1. Mineração de dados. 2. Qualidade de vida. 3. Gestão do conhecimento. I.  
Scandelari, Luciano. II. Universidade Tecnológica Federal do Paraná, Campus Ponta  
Grossa. III. Título.

CDD 658.5



Universidade Tecnológica Federal do Paraná –  
Campus de Ponta Grossa  
Gerência de Pesquisa e Pós-Graduação  
**PROGRAMA DE PÓS-GRADUAÇÃO EM  
ENGENHARIA DE PRODUÇÃO**



## TERMO DE APROVAÇÃO

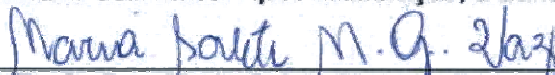
Título de Dissertação Nº 56

**ANÁLISE DOS RESULTADOS DO WHOQOL-100. UTILIZANDO DATA MINING**

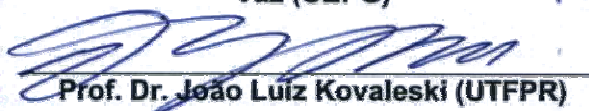
por

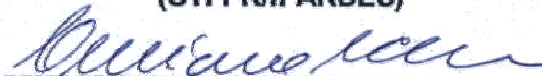
**Celso Bilynkievycz dos Santos**

Esta dissertação foi apresentada às **16 horas** da manhã do dia **07 de dezembro de 2007** como requisito parcial para a obtenção do título de **MESTRE EM ENGENHARIA DE PRODUÇÃO**, com área de concentração em **Gestão Industrial**, linha de pesquisa em **Gestão do Conhecimento e Inovação**, Programa de Pós-Graduação em Engenharia de Produção. O candidato foi argüido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado.

  
\_\_\_\_\_  
**Profa. Dra. Maria Salete Marcon Gomes  
Vaz (UEPG)**

  
\_\_\_\_\_  
**Profa. Dra. Deborah Ribeiro Carvalho  
(UTFPR/IPARDES)**

  
\_\_\_\_\_  
**Prof. Dr. João Luiz Kovaleski (UTFPR)**

  
\_\_\_\_\_  
**Prof. Dr. Luciano Scandelari (UTFPR) -  
Orientador**

Visto do Coordenador:

\_\_\_\_\_  
**Kazuo Hatakeyama (UTFPR)  
Coordenador do PPGE**

*Dedico este trabalho em memória a minha querida irmã caçula,  
Shirlei Alves dos Santos, falecida em 2006 e a minha querida avó,  
Maria Chicoski Bilynkievycz, falecida em 2005.*

## AGRADECIMENTOS

*Antes de tudo a Deus por ter me acompanhado nesta jornada.*

*Agradeço ao meu orientador, Prof. Luciano Scandelari, por sua dedicação e pela forma motivadora que conduziu toda a orientação deste trabalho.*

*Aos meus familiares, pelo incentivo e pela compreensão nos momentos de ausência.*

*À minha esposa Morgana e filhas, Bianca e Camila, principais motivos da minha existência.*

*Ao meu amigo e colega de trabalho, Cristiano Delabernarda, pelo apoio e por ter assumido muitas das minhas responsabilidades de trabalho.*

*À Universidade Estadual de Ponta Grossa (UEPG), pela liberação concedida para cursar o mestrado.*

*Ao Conselho de Administração de 2006 da UEPG, principalmente aos professores Elias Zahi Fadel, Paulo Roberto Godoy, Carlos Roberto Berger, Graciete Tozetto Góes, Atair Justino, Eduardo Caíres e João Alfredo Madalozzo, pelos debates favoráveis durante o processo de liberação.*

*Ao Centro de Processamento de Dados da UEPG, principalmente aos servidores, Carlos Alberto Volpi e Christiani Borsato de Ramos, pelo apoio.*

*Ao Magnífico Reitor da UEPG, Prof. João Carlos Gomes, pelo apoio.*

*A todos os servidores da UEPG, que colaboraram com a pesquisa.*

*Ao servidor, Luiz César dos Santos Lima da UTFPR pelo apoio nos trâmites burocráticos.*

*Ao Prof. Luiz Alberto Pilatti, pelo incentivo.*

*Aos demais professores do Mestrado, Antonio Carlos de Francisco, Dácio Roberto dos Reis, Hélio Gomes de Carvalho, João Luiz Kovaleski, Kazuo Hatakeyama e Magda Lauri Gomes Leite, pela dedicação e pelo conhecimento transmitido durante os dois anos de formação.*

*A Prof<sup>ª</sup> Regina Maria Beninca Schwingel, pela colaboração.*

*Aos meus colegas e amigos de pós-graduação, principalmente ao Julio César Ferreira, Vicente Toniolo Zander, Marcos Aurélio Zoldan e Paula Michelle Purcidonio, pelas palavras de incentivo, nos momentos de dificuldades.*

*À Universidade Tecnológica Federal do Paraná, por oportunizar esta titulação.*

*.... nunca ande pelo caminho traçado,  
pois ele conduz somente até onde os  
outros já foram ...*

*Alexandre Graham Bell*



## RESUMO

A Qualidade de Vida é um estado de bem-estar, que se intercorrelaciona com diversos aspectos do cotidiano e que pode sofrer interferências de fatores externos, muitas vezes difíceis de serem identificados, principalmente quando são estudados grandes grupos de trabalhadores com diferentes características. A identificação destes fatores, potencialmente implícitos, poderá auxiliar gestores de empresas em suas tomadas de decisões para melhoria do estado de qualidade de vida de seus colaboradores e, conseqüentemente, da sua produtividade. Este trabalho propõe uma análise complementar utilizando técnicas de *Data Mining* de maneira inovadora. Foi realizado um estudo de caso envolvendo agentes universitários de uma instituição pública e, como instrumento de coleta dos dados, o *WHOQOL-100*. Para aquisição de conhecimento implícito, foram percorridas as etapas do Processo de Descoberta de Conhecimento em Base de Dados, utilizando-se na etapa de Mineração de Dados, técnicas de Regras de Associação. Conclui-se que as técnicas de *Data Mining*, mais especificamente, as de Regras de Associação, podem ser utilizadas para complementar a análise dos resultados de um instrumento de pesquisa estruturada para *surveys*, principalmente quando se dispõe de muitos dados adicionais e se deseja descobrir conhecimento implícito. O planejamento e a utilização combinada de medidas de interesse subjetivas e objetivas permitiu diminuir os custos operacionais das tarefas de descrição do Processo de Aquisição de Conhecimento em Base de Dados.

### Palavras-chave

Mineração de Dados, *WHOQOL-100*, Qualidade de Vida, Processo de Descoberta de Conhecimento em Base de Dados.

## ABSTRACT

Quality of Life is a well-being state which links several different aspects of everyday life and which also may be suffer interferences from outside factors, being those at times difficult to be realized, mainly when are studied carries large groups of workers with different characteristics. The identification of such factors, potentially implicit, might serve managers as for their decisions when it comes to improve the quality of life of their collaborators and consequently the productivity. The following study throws light over a complementary analysis using Data Mining techniques in an innovative mode. A Study of case was carried out approaching worker of college from a public institution, using the WHOQOL-100 as a data collector. In order to obtain implicit knowledge, the Procedure of Knowledge Discovery in Databases phases was taken, using it on the Data Mining, Rules of Association techniques. It is concluded that the Data Mining techniques, more specifically, the Rules of Association ones, can be used as an addition when it comes to analysis results of an instrument of study based upon surveys, mainly when a huge number of additional data is held and the implicit knowledge is the main focus to be figured. The planning as well as the use of combined measures of subjective and objective interest allowed the decrease of operational tasks costs as for the very description of the Procedure of Knowledge Discovery in Databases, Rules of Association.

### **Keywords**

Data Mining, WHOQOL-100, Quality of Life, Knowledge Discovery in Databases, Rules of Association.

## LISTA DE FIGURAS

Figura 1 – Frequência de publicações envolvendo o assunto <i>Data Mining</i> nas Edições dos ENEGEPs dos anos de 2003, 2004 e 2005, distribuídas por áreas temáticas.....	24
Figura 2 – Frequência de Dissertações e Teses Publicadas no Portal CAPES, envolvendo os Temas: <i>WHOQOL</i> , <i>Data Mining</i> .....	26
Figura 3 – Frequência de Facetas por Domínio .....	31
Figura 4 – Distribuição das Questões por Domínio .....	31
Figura 5 - Distribuição dos Tipos de Escalas Respostas do <i>WHOQOL-100</i> .....	33
Figura 6 - Processo de KDD x <i>Data Mining</i> .....	40
Figura 7 – Etapas do Processo KDD.....	41
Figura 8 – Etapas do Processo de KDD Desenvolvidas na Pesquisa.....	60
Figura 9 – Relação de Dependência entre as Questões, Facetas e Domínios de QV .....	62
Figura 10 – Equação para Inversão de Escala.....	64
Figura 11 - Escala de Classificação dos Resultados.....	65
Figura 12 - Médias: Domínio, QGQV e Média de QV.....	67
Figura 13 - Aspectos de Qualidade de Vida.....	68
Figura 14 – Atributos inter-relacionados com o Domínio Meio Ambiente.....	73
Figura 15 - Atributos inter-correlacionados com o Atributo Sanção Disciplinar .....	74
Figura 16 - Atributos inter-correlacionados com o Atributo Sanção Disciplinar .....	74
Figura 17 – Regra de Produção .....	85

## LISTA DE TABELAS

Tabela 1 – Núcleo de diferentes culturas que contribuíram na elaboração do <i>WHOQOL-100</i> e seus respectivos países .....	29
Tabela 2 - Aspectos e contemplados em cada domínio do <i>WHOQOL-100</i> .....	30
Tabela 3 – Questões Gerais sobre Qualidade de Vida .....	32
Tabela 4 - Relação entre a Escala de Likert e as Escalas: %, Avaliação, Capacidade e Frequência. ....	32
Tabela 5 – Questões Negativas do <i>WHOQOL-100</i> .....	34
Tabela 6 – Grupos de Questões que contemplam diferentes Facetas.....	36
Tabela 7 - Estágios no desenvolvimento do <i>WHOQOL-100</i> (Grupo <i>WHOQOL</i> , 1995) .....	37
Tabela 8 - Problemas e respectivas Tarefas e principais Técnicas de <i>Data Mining</i> utilizadas para sua solução.....	46
Tabela 9 – Itens dos Registros em uma Base de Dados e sua Distribuição .....	53
Tabela 10 - Classificação da Nível de Qualidade de Vida.....	65
Tabela 11 – Relação de Estudos .....	66
Tabela 12 - Índice de Consistência Interna dos Indicadores de QV Antes e Após a Categorização dos Dados.....	67
Tabela 13 – Percentual dos Atributos Metas atendidos pelos Atributos Previsores..	72
Tabela 14 – Total de RA com os Resultantes .....	78
Tabela 15 – Regras de Associação com Conseqüente: Média de Qualidade de Vida (MQV) .....	79
Tabela 16 - Regras de Associação com Conseqüente: Questões Gerais de Qualidade de Vida (QGQV) .....	80
Tabela 17 - Regras de Associação com Conseqüente: Domínio Meio Ambiente (D5) .....	81

Tabela 18 - Regras de Associação com Conseqüente: Aspecto Dor e Desconforto (F1) .....	82
Tabela 19 - Regras de Associação com Conseqüente: Aspecto Segurança Física e Proteção (F16) .....	83
Tabela 20 - Regras de Associação com Conseqüente a Informação de Cadastro: Sanção Disciplinar .....	84
Tabela 21 - Regras de Associação com Conseqüente a Informação de Cadastro: Disfunção Definitiva .....	84
Tabela 22 – Resultados Estatísticos dos Domínios de Qualidade de Vida e a Média de QV .....	99
Tabela 23 – Resultados Estatísticos dos Aspectos de Qualidade de Vida .....	100
Tabela 24 - Relação de Dados Cadastrais da Amostra .....	101
Tabela 25 – Intercorrelação entre os Indicadores de Qualidade de Vida .....	103

## LISTA DE EQUAÇÕES

Equação 1 - Suporte .....	50
Equação 2 - Confiança.....	51
Equação 3 - <i>Lift</i> .....	51
Equação 4 – Rules Interset ou <i>Leverage</i> .....	52
Equação 5 - Convicção .....	53
Equação 6 - Inversão de Escala.....	64
Equação 7 – Conversão da Escala de Likert para Centesimal.....	64

## LISTA DE ABREVIATURAS E SIGLAS

BDI	- <i>Beck Depression Inventory</i>
BHS	- <i>Beck Hopelessness Scale</i>
C	- Índice de Confiança
Cv	- Índice de Convicção
D	- Domínio de Qualidade de Vida
DM	- <i>Data Mining</i>
F	- Faceta ou Aspecto de Qualidade de Vida
EQVF	- Escala de Qualidade de Vida de Flanagan
IQLIE	- <i>Quality of life in Epilepsy Inventory</i>
L	- Índice de <i>Lift</i>
LM	- <i>Machine learning</i>
Lv	- Índice de <i>Leverage</i>
KDD	- <i>Knowledge in Data Discovery</i>
MQV	- Média de Qualidade de Vida
ICT	- Índice de Capacidade para o Trabalho
ODBC	- <i>Open Data Base Connectivity</i>
PCCS	- Plano de Carreira, Cargos e Salários
QV	- Qualidade de Vida
QGVQ	- Questões Gerais de Qualidade de Vida
RA	- Regras de Associação
RI	- <i>Rules Interest</i>
S	- Índice de Suporte
SQL	- <i>Structured Query Language</i>
WHOQOL	- <i>World Health Organization quality of life assessment instrument</i>

## LISTA DE SÍMBOLOS

- U** - União
- - Sentido de implicação
- ¬** - Negação lógica



## SUMÁRIO

AGRADECIMENTOS	iv
RESUMO	vi
ABSTRACT	vii
LISTA DE FIGURAS	viii
LISTA DE TABELAS	ix
LISTA DE EQUAÇÕES	xi
LISTA DE ABREVIATURAS E SIGLAS	xii
LISTA DE SÍMBOLOS	xiii
SUMÁRIO	xiv
<b>1 INTRODUÇÃO</b>	<b>18</b>
1.1 DEFINIÇÃO DO OBJETO	19
1.2 PROBLEMA DE PESQUISA	19
1.3 JUSTIFICATIVA	19
1.4 OBJETIVOS	21
1.4.1 Objetivo Geral	21
1.4.2 Objetivos Específicos	21
<b>2 REVISÃO DE LITERATURA</b>	<b>23</b>
<b>3 INSTRUMENTO DE AVALIAÇÃO DE QUALIDADE DE VIDA - WHOQOL-100</b>	<b>29</b>
3.1.1 Composição do <i>WHOQOL-100</i>	30
3.1.2 Agrupamento das Perguntas	35
3.1.3 Construção do Instrumento	37
3.1.4 Validação do instrumento	37
3.1.5 Versão final	38
3.1.6 Outras Versões do <i>WHOQOL</i>	38
<b>4 DATA MINING x KDD</b>	<b>40</b>
4.1 O PROCESSO de KDD	40
4.1.1 Etapas do Processo KDD	41
4.2 <i>DATA MINING</i>	43
4.2.1 Classificação do <i>Data Mining</i>	44
4.2.2 Objetivos de <i>Data Mining</i>	44
4.2.3 Tarefas e Técnicas de <i>Data Mining</i>	45
4.2.4 Softwares de <i>Data Mining</i>	47
<b>5 MEDIDAS DE INTERESSE</b>	<b>49</b>
5.1 Medidas de Interesse Subjetivas	49
5.2 Medidas de Interesse Objetivas	49
5.2.1 Suporte	50
5.2.2 Confiança	50
5.2.3 <i>Lift</i>	51

5.2.4	<i>Rules Interest</i>	51
5.2.5	Convicção	52
6	<b>METODOLOGIA</b>	<b>56</b>
6.1	DESENHO DA PESQUISA	56
6.2	INSTRUMENTO, FONTE DE PESQUISA E INTERVENÇÃO EM AMBIENTES SOCIAIS E INSTITUCIONAIS	57
6.3	DADOS DA PESQUISA	57
6.4	APLICAÇÃO DO INSTRUMENTO	58
6.5	ANÁLISE DOS DADOS	59
7	<b>ANÁLISE DOS RESULTADOS DO WHOQOL-100 UTILIZANDO DATA MINING</b>	<b>60</b>
7.1	Pré-processamento para <i>Data Mining</i>	60
7.1.1	Implementação da Base de Dados	61
7.1.2	Conhecimento da Base de Dados	67
7.1.3	Enriquecimento da Base de Dados	69
7.1.4	Seleção e Transformação dos Dados	70
7.2	<i>Data Mining</i>	75
7.3	Pós-processamento de <i>Data Mining</i>	75
7.3.1	Transformação	75
7.3.2	Seleção das Regras de Associação	75
7.3.3	Interpretação dos Resultados	85
7.3.4	Aquisição de Conhecimento	85
8	<b>CONCLUSÕES E RECOMENDAÇÕES</b>	<b>88</b>
8.1	SUGESTÕES PARA TRABALHOS FUTUROS	89
	<b>REFERÊNCIAS</b>	<b>91</b>
	ANEXO A – ESTATÍSTICA DESCRITIVA DOS INDICADORES DE QUALIDADE DE VIDA – DOMÍNIOS, QGQV E MQV	99
	ANEXO B – ESTATÍSTICA DESCRITIVA DOS INDICADORES DE QUALIDADE DE VIDA - ASPECTOS	100
	ANEXO C – DADOS CADASTRAIS UTILIZADOS NA MINERAÇÃO DE DADOS	101
	ANEXO D – INTERCORRELAÇÃO ENTRE OS INDICADORES DE QUALIDADE DE VIDA DA AMOSTRA	103
	ANEXO E – MODELOS DE RELATÓRIOS OFERECIDOS AOS PARTICIPANTES DA PESQUISA	104
	APÊNDICE A – TRABALHOS PUBLICADOS DURANTE O MESTRADO	105

## 1 INTRODUÇÃO

A Qualidade de Vida (QV), definida pela Organização Mundial de Saúde (OMS), como "um estado de completo bem-estar físico, mental e social, que não consiste apenas na ausência de doença ou de enfermidade", foi um tema de estudo muito discutido ao longo dos anos, em diferentes momentos e perspectivas, principalmente pós-revolução industrial.

Atualmente essa temática, também é foco de estudo da Engenharia de Produção, principalmente quando se aborda a Qualidade de Vida no Trabalho (QVT), pois as empresas estão em constante busca por qualidade, modificando suas atitudes, relacionadas à vida de seus consumidores internos (colaboradores) e externos, e cada vez mais, preocupadas com suas responsabilidades sociais.

Alguns indicadores podem ser freqüentemente monitorados pelos seus gestores, para auxiliar no planejamento de ações e tomadas de decisões, que garantam níveis de Qualidade de Vida satisfatórios.

Para realização deste monitoramento existem diversos instrumentos que avaliam a Qualidade de Vida, seja ela observada apenas no ambiente de trabalho ou de forma global. Alguns possuem protocolos específicos para análise básica dos dados, outros dependem da criatividade de análise do pesquisador.

Neste trabalho foi realizado um estudo de caso utilizando técnicas de *Data Mining* para complementar a análise dos resultados de uma ferramenta de coleta e análise de dados.

Escolheu-se para estudo de caso a Universidade Estadual de Ponta Grossa (UEPG), principalmente pela acessibilidade aos dados e por ser considerada uma das instituições que mais emprega na região dos Campos Gerais, com grande diversificação de mão-de-obra.

## 1.1 DEFINIÇÃO DO OBJETO

Este estudo propõe inovar a análise dos resultados *WHOQOL-100*<sup>1</sup>, com a utilização de técnicas de *Data Mining*, definindo-se o objeto como: “Análise dos Resultados do *WHOQOL-100* utilizando técnicas de *Data Mining*”.

## 1.2 PROBLEMA DE PESQUISA

A Qualidade de Vida é um estado de bem-estar influenciado por um conjunto de fatores.

A identificação dos fatores que prejudicam este estado pode auxiliar na elaboração de medidas que contribuam para a sua melhoria. Porém, na maioria das vezes, eles são difíceis de serem identificados, devido a sua complexidade e interdependência, principalmente quando se deseja monitorar grupos de trabalhadores com diferentes características.

Primeiramente surge o problema de como colher o maior número de dados padronizados e depois de como identificar os fatores que poderiam estar prejudicando a Qualidade de Vida e que poderiam passar despercebidos aos olhos dos gestores de qualidade da empresa.

Então, neste contexto, definiu-se a seguinte pergunta a nortear a pesquisa:

Como extrair conhecimento implícito de Qualidade de Vida?

## 1.3 JUSTIFICATIVA

Partindo das premissas de que a QV deve ser observada de forma global (FOX-RUSHBY & PARKER, 1995) e, de que o processo de tomada de decisão pode ser auxiliado por ferramentas de apoio à decisão (FAYYAD et al., 1996), busca-se neste trabalho, a utilização de uma ferramenta de avaliação de qualidade de vida global e de uma ferramenta de apoio à tomada de decisões.

Nesta perspectiva, procurou-se um instrumento que melhor avaliasse a qualidade de vida de forma global.

---

<sup>1</sup> *WHOQOL-100* – Instrumento de Avaliação de Qualidade de Vida da Organização Mundial de Saúde (OMS).

Adotou-se como objeto de estudo, o Instrumento de Avaliação de Qualidade de Vida - *WHOQOL-100*, considerado por Fleck (1995), um instrumento complexo, por contemplar a qualidade de vida nos seus diferentes aspectos.

Este instrumento possui uma sintaxe para análise estatística dos dados, que permite apenas a obtenção das médias dos domínios e dos aspectos de qualidade de vida. Outros tipos de análises estatísticas mais específicas devem ser implementadas pelo pesquisador.

Considerando-se o grande número de questões do instrumento e de resultados gerados pela sintaxe, mais a possibilidade de acréscimo de dados cadastrais e, considerando-se também, as limitações da holística humana, recorreu-se à Inteligência Artificial (IA) para complementar a análise.

Segundo Fayyad (1998), o ser humano não está preparado para interpretar grandes volumes de dados e/ou espaços multidimensionais. Nesta conjuntura, a Tecnologia da Informação surge como a ferramenta necessária para exercer esta função, mais especificamente a IA, com suas técnicas de *Data Mining*.

A função do *Data Mining* é garimpar informações ocultas em grandes bases de dados, ampliando ao máximo as comparações entre os dados e tornando o conhecimento implícito em explícito.

As técnicas tradicionais exigem que, inicialmente, hipóteses sejam estabelecidas, para que então, sejam construídas consultas a fim de comprovar ou não estas suposições, enquanto as técnicas de *Data Mining* têm maior independência do estabelecimento de hipóteses iniciais.

O instrumento *WHOQOL-100* foi escolhido, por ser um instrumento que avalia qualidade de vida de forma global, desenvolvido e validado pela Organização Mundial da Saúde e, que contempla os diversos aspectos inerentes à qualidade de vida de um indivíduo.

Foi escolhido o *Data Mining*, como objeto de estudo, por se tratar de uma área da Inteligência Artificial que permite descobrir conhecimento implícito em bases de dados e, que pode contribuir nas tomadas de decisões dos gestores de qualidade de vida.

Outro fator motivador é a inexistência de trabalhos utilizando técnicas de *Data Mining* em dados coletados a partir do *WHOQOL-100*, pois os trabalhos existentes, geralmente se limitam à análise proposta pelo Grupo *WHOQOL*, para o instrumento, o que caracteriza o presente trabalho como inovador.

## 1.4 OBJETIVOS

### 1.4.1 Objetivo Geral

Adquirir conhecimento implícito dos resultados do *WHOQOL-100*.

### 1.4.2 Objetivos Específicos

Neste trabalho buscou-se atender aos seguintes objetivos específicos:

1. Computar os indicadores de Qualidade de Vida da amostra através de Linguagem de Consulta Estruturada (SQL);
2. Definir um modelo de dependência entre questões, facetas e domínios de Qualidade de Vida do *WHOQOL-100*;
3. Definir uma categorização para os indicadores de Qualidade de Vida;
4. Classificar a Qualidade de Vida da amostra em seus diferentes domínios e aspectos;
5. Definir os piores indicadores de Qualidade de Vida da amostra;
6. Encontrar padrões de comportamento interessantes na Média de Qualidade de Vida da amostra;
7. Encontrar padrões de comportamento interessantes na Média das Questões Gerais de Qualidade de Vida da amostra;
8. Encontrar os itens cadastrais previsores dos indicadores de Qualidade de Vida da amostra;
9. Encontrar padrões de comportamento interessantes nos piores indicadores de Qualidade de Vida da amostra;
10. Encontrar os indicadores de Qualidade de Vida, previsores de dados cadastrais da amostra;

11. Encontrar padrões de comportamento interessantes nos indicadores de Qualidade de Vida, previsores de dados cadastrais da amostra;
12. Avaliar o uso de medidas de interesse.

## 2 REVISÃO DE LITERATURA

Neste capítulo é apresentado o “estado da arte”, evidenciando os principais pesquisadores, trabalhos e aspectos, publicados sobre os assuntos do tema, demonstrando a existência de lacunas entre eles e a inexistência de trabalhos semelhantes.

Esta revisão de literatura, segundo Luna (1997), classifica-se como empírica, pois procura explicar como o problema vem sendo pesquisado do ponto de vista metodológico.

Os assuntos do tema desta pesquisa, *Data Mining* e *WHOQOL-100*, são objetos de estudo da Engenharia de Produção, porém, pertencem a áreas distintas, respectivamente, às áreas de Tecnologias da Informação e da Saúde. Apesar deste diferencial, assemelham-se pelas contemporaneidades, pois tiveram elevada conotação científica na última década. O primeiro tema teve seu apogeu em 1995, quando a comunidade científica reunida na Cidade Montreal – Canadá, depois de vários Workshops (FAYYAD & UTHURUSAMY, 1994) realizados com a Temática Prospecção de Dados (FAYYAD & UTHURUSAMY, 1995), definiu o Termo “Knowledge Discovery in Databases – KDD”, ao processo de Aquisição de Conhecimento em Base de Dados, onde o *Data Mining* está inserido, como uma das etapas deste processo. Enquanto o segundo recebeu destaque em 1994, quando *WHOQOL GROUP*, publicou o Instrumento de Avaliação de Qualidade de Vida da Organização Mundial da Saúde – *WHOQOL-100* (*WHOQOL GROUP*, 1994).

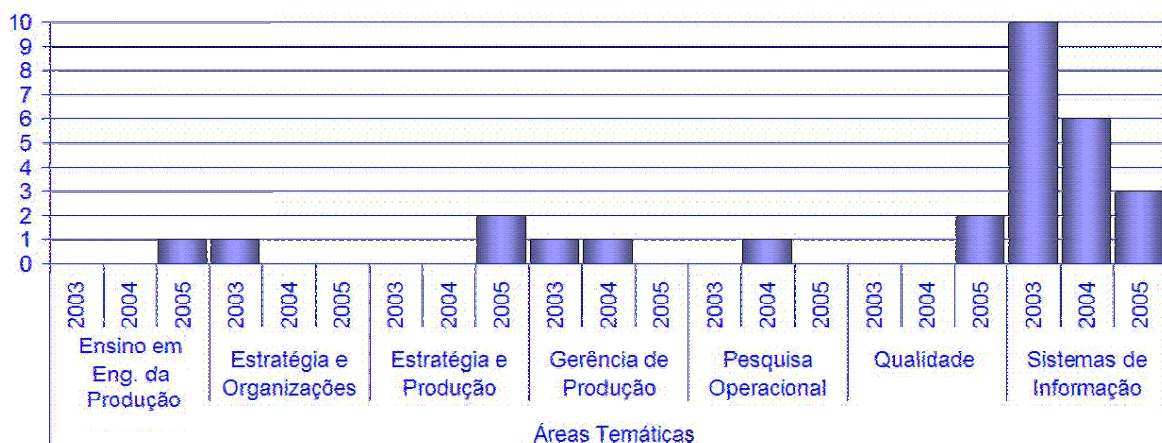
Buscaram-se literaturas nas três últimas edições do Encontro Nacional de Engenharia de Produção, que se assemelhassem ao tema proposto e envolvessem os dois assuntos, no entanto, não foram encontradas. Provavelmente, pela diferença de áreas e pela inovação sugerida, não existem relatos de trabalhos relacionando os dois assuntos.

Procuraram-se também, trabalhos com os assuntos abordados de formas separadas, o que resultaram em 29 publicações. Destas, apenas uma abordava o assunto *WHOQOL* e as demais tratavam do assunto *Data Mining* em diferentes áreas da Engenharia de Produção.



Na única pesquisa que tratava do assunto *WHOQOL*, os pesquisadores Rugiski, Pilatti e Scandelari (2005) verificaram o nível de utilização do instrumento de avaliação da Qualidade de Vida da Organização Mundial da Saúde - *WHOQOL-100*, através da *internet* em *sites* na língua portuguesa. Neste trabalho, os *sites* foram categorizados e quantificados, chegando-se à seguinte leitura: 30 *sites* em língua portuguesa; 76,67% dos quais apresentam trabalhos científicos, na grande maioria artigos; 86,96% dos temas dos artigos relacionados às Ciências da Saúde e Ciências Humanas; e que a produção acadêmica está concentrada nas instituições: USP e a UFRGS.

Os demais trabalhos mencionados anteriormente abordavam pesquisas com o assunto *Data Mining* nas áreas de: Ensino em Engenharia de Produção; Estratégia e Organizações; Estratégia e Produção; Gerência de Produção; Pesquisa Operacional; Qualidade e Sistemas de Informação, conforme Figura 1.



**Figura 1 – Frequência de publicações envolvendo o assunto *Data Mining* nas Edições dos ENEGEPs dos anos de 2003, 2004 e 2005, distribuídas por áreas temáticas**

A Figura 1 apresenta a freqüência das publicações envolvendo o assunto *Data Mining* nas três últimas edições dos ENEGEPs, distribuídas por áreas temáticas. Observa-se, que no ano de 2003 foram publicados 12 artigos envolvendo o assunto, nas seguintes áreas: 01 em Gerência de Produção; 01 em Estratégia e Organizações e 10 em Sistemas de Informação. No ano de 2004, também foram publicados 10 trabalhos, distribuídos nas seguintes áreas de temática: 01 em Gerência de Produção; 01 em Pesquisa Operacional; 06 em Sistemas de Informação. No ano de 2005 foram publicados 08 trabalhos,

distribuídos nas seguintes áreas de temática: 01 em Ensino em Engenharia de Produção; 02 em Estratégia e Produção; 02 em Qualidade e 03 em Sistemas de Informação.

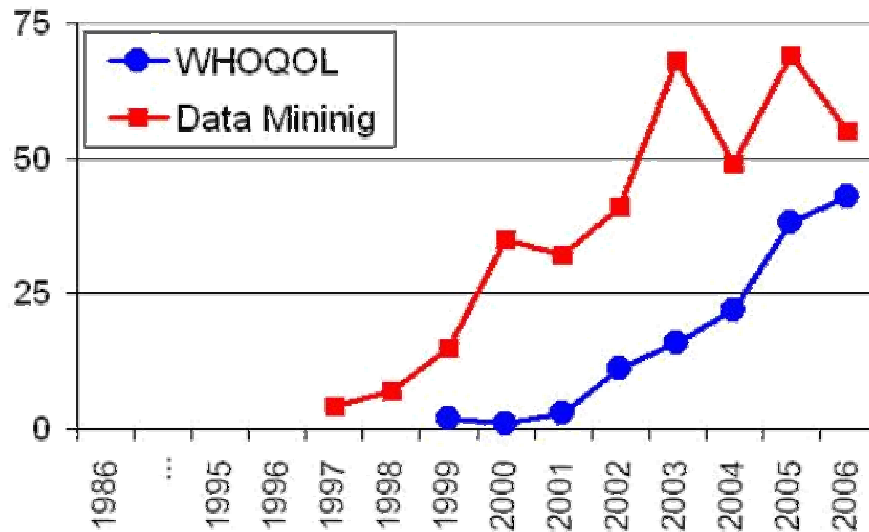
A partir destas informações, observa-se que os assuntos são objetos de estudo da área de Engenharia de Produção.

Além do Evento Nacional da área, também foram realizadas buscas para encontrar trabalhos relacionados nas fontes de pesquisa de trabalhos científicos da base de dados da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES e no banco de teses de instituições brasileiras detentoras de cursos de Engenharia de Produção, destacando as seguintes informações sobre os assuntos objetos de estudo.

A partir do Banco de Monografias, Dissertações e Teses da Universidade de São Paulo (USP) e da Universidade Federal de Santa Catarina (UFSC), procurou-se trabalhos envolvendo o assunto *WHOQOL* publicados nos Programas de Engenharia de Produção e observou-se que existem poucos estudos. Entre eles podemos citar o trabalho de Martins (2005). No entanto, envolvendo o assunto *Data Mining*, encontram-se inúmeros trabalhos, em sua maioria, avaliando e comparando a utilização de algoritmos mineradores, como por exemplo, os trabalhos de Kulkarni (2007).

A pesquisadora Martins (2005), em seu estudo, analisou profissionais de enfermagem, categorizados por turno de trabalho e possuidores ou não de outro emprego, verificando o aspecto da qualidade de vida (QV) e capacidade para o trabalho (CT), utilizando o *WHOQOL - Breaif* e o índice de capacidade para o trabalho (ICT). Seus resultados indicaram que o QV e ICT estão diretamente relacionados.

Analisando o Banco de Teses e Dissertações do Portal da CAPES, observa-se a existência de vários trabalhos envolvendo os assuntos do tema desta pesquisa em diferentes campos, além da área de Engenharia de Produção (Figura 2).



**Figura 2 – Frequência de Dissertações e Teses Publicadas no Portal CAPES, envolvendo os Temas: *WHOQOL*, *Data Mining***

Observa-se, a partir da Figura 2, um notável crescimento nos últimos anos de Dissertações ou Teses envolvendo esses assuntos, acumulando ao longo de 11 anos, 136 trabalhos sobre *WHOQOL* e 375 sobre *Data Mining*.

O primeiro trabalho publicado, por Polyana (1997), estudava técnicas que auxiliassem no desenvolvimento de ferramentas de KDD, da revisão e adaptava técnicas da programação denominada "Generalização Menos Geral Relativa" (ILP) para as técnicas de *Data Mining*. Seu trabalho envolveu basicamente a integração de duas linhas de pesquisa, Banco de dados relacional e técnicas de programação ILP.

Moreno (1999) realiza revisão sistemática sobre trabalhos que relacionam qualidade de vida e laringectomia, concluindo que a maioria dos estudos que propunham mensurar constructo de QV em laringectomizados não o faziam com rigor metodológico e, essas relações não seguiam orientação básica e a definição de QV estabelecida pelo *WHOQOL*.

Em relação a trabalhos publicados em periódicos sobre os assuntos, observa-se que as primeiras pesquisas sobre o assunto *WHOQOL*, tinham como propósito a sua criação do instrumento Internacional (HERRMAN et. al.,1993).

Em 1995, depois da sua publicação como instrumento oficial da OMS, passou a ser validado em diversos países.

O pesquisador Fleck, componente do GRUPO *WHOQOL*, foi o precursor dos trabalhos envolvendo *WHOQOL* no Brasil. Sua equipe traduziu para a língua portuguesa e realizou os primeiros testes de validação do *WHOQOL-100* (Fleck et. al., 1999), *WHOQOL-Bref* (FLECK, et. al. 2000), *WHOQOL-OLD* (FLECK, et. al. 2006) e *WHOQOL-HIV* (ZIMPEL & FLECK, 2007).

Alguns estudos utilizam simultaneamente o *WHOQOL* com outros instrumentos de avaliação, com propósitos de validação concorrente ou para análise aglomerada. Na primeira situação, podemos citar os trabalhos de Fleck (1999), utilizando o Inventário de Depressão de Beck (BDI) e Escala de desesperança de Beck (BHS). Na segunda situação, os trabalhos de Moraes e Souza (2005) e, Martins (2002), que utilizaram o *WHOQOL*, respectivamente, com a Escala de Qualidade de Vida de Flanagan (EQVF) e com o Índice de Capacidade para o Trabalho (ICT).

Masthoff et. al (2005) validou o uso do *WHOQOL-100*, em uma população holandesa de pacientes psiquiátricos adultos.

Meldolesi et. al (2006) utilizou em seu estudo, análise múltipla de regressão linear, para verificar a relação entre *WHOQOL-100* e outros instrumentos específicos para avaliar a QV na epilepsia, entre eles o QOLIE – 31. Em seu trabalho sugerem que a QV em pacientes epiléticos pode ser substancialmente afetada pela presença e severidade dos sintomas depressivos e, em menor grau, os sintomas da ansiedade.

Recentemente, Meldolesi et. al (2007), divulga os resultados de outro estudo semelhante, concluindo que a melhoria na qualidade de vida global e principais domínios de qualidade de vida estão associados à diminuição da depressão, ansiedade e irritação.

O *WHOQOL*, ainda se encontra em processo de validação e adaptação para alguns países. Neste ano foi publicada a validação do *WHOQOL-Tawain*, adaptado para a cultura taiwanense (YAO, 2007).

Sobre o assunto *Data Mining*, observa-se trabalhos, como o de Mcgarry (2005), que avaliam o uso das algumas unidades de medidas de interesse.

McGarry (2005), apresenta revisão de algumas Medidas de Interesse disponíveis na literatura para avaliar e classificar o produto da mineração de dados. E afirma que a grande lacuna existente na área de *Data Mining* está em saber como a investigação poderá combinar Medidas de Interesse Objetivas e Subjetivas em um único índice.

Sobre o assunto *Data Mining*, atualmente se observa que a preocupação da comunidade científica concentra-se, principalmente, na busca de medidas de interesse que possam unir medidas objetivas e subjetivas de forma mais autônoma (MCGARRY, 2005; LENCA et. al., 2008).

A maioria dos trabalhos envolvendo o assunto *WHOQOL* está relacionada à área da saúde, enquanto os que envolvem *Data Mining* aparecem em diferentes segmentos da área de Engenharia de Produção.

Todos os trabalhos mencionados assemelham-se com o tema proposto, embora de formas independentes, pois os assuntos envolvidos estão inseridos em diferentes áreas, incluindo as de Engenharia de Produção.

Segundo Silva e Menezes (2001), a revisão de literatura é essencial para fornecer informações para impedir a duplicação de pesquisas sobre o mesmo aspecto do tema e favorecer definições adjacentes mais precisas do problema a ser estudado. Portanto, esta revisão de literatura foi finalizada, afirmando-se que, o principal diferencial entre o tema e os demais trabalhos está na forma de análise dos dados, pois neles são utilizadas apenas técnicas estatísticas para este fim, enquanto neste, sugere-se uma inovação complementando a análise dos dados com a aplicação de técnicas de *Data Mining*.

### 3 INSTRUMENTO DE AVALIAÇÃO DE QUALIDADE DE VIDA - WHOQOL-100

Na década de noventa, ocorreu uma proliferação de instrumentos para avaliação da QV, cuja maioria foi desenvolvida nos Estados Unidos. Cresceu também, nesse período, o interesse pela tradução destes instrumentos para aplicação em outros países (BULLINGER, ANDERSON E CELLA, 1993).

O aproveitamento transcultural de qualquer instrumento de avaliação traduzido, sempre foi um tema polêmico e contestado (FOX-RUSHBY & PARKER, 1995).

A procura de um instrumento de avaliação de qualidade de vida dentro de uma expectativa legitimamente universal fez com que a Organização Mundial da Saúde desenvolvesse um projeto colaborativo internacional. Este projeto resultou na elaboração do *WHOQOL-100*, um instrumento de QV composto por 100 itens (FOX-RUSHBY & PARKER, 1995).

O projeto foi elaborado com a cooperação de 15 núcleos (Tabela 1) de diferentes culturas que trabalharam simultaneamente.

**Tabela 1 – Núcleo de diferentes culturas que contribuíram na elaboração do *WHOQOL-100* e seus respectivos países**

Núcleo	País	Núcleo	País	Núcleo	País
Bangkok	Tailândia	Madras	Índia	São Petersburgo	Rússia
Barcelona	Espanha	Melbourne	Austrália	Seattle	EUA
Bath	Reino Unido	Nova Delhi	Índia	Tilburg	Holanda
Beer-Sheeva	Israel	Panamá	Panamá	Tóquio	Japão
Harare	Zimbabwe	Paris	França	Zagreb	Croácia

A versão do *WHOQOL-100* na língua portuguesa foi desenvolvida no Brasil em 1999, sob a coordenação do pesquisador Dr. Marcelo Pio de Almeida Fleck no Departamento de Psiquiatria e Medicina Legal da Universidade Federal do Rio Grande do Sul – UFRGS.

### 3.1.1 Composição do WHOQOL-100

O WHOQOL-100 é composto por 100 questões que analisa seis domínios: Físico, Psicológico, Nível de Independência, Relações Sociais, Meio Ambiente e Espiritualidade / Crenças Pessoais.

O instrumento proporciona uma avaliação minuciosa de 25 facetas, das quais uma corresponde às questões de QV geral, e as demais correspondem a 24 aspectos que estão distribuídas entre os 6 domínios (Tabela 2).

**Tabela 2 - Aspectos e contemplados em cada domínio do WHOQOL-100**

Domínios		Aspectos Contemplados	
Nº	Nome	Nº	Descrição
I	Físico	01	Dor e desconforto
		02	Energia e fadiga
		03	Sono e repouso
II	Psicológico	04	Sentimentos Positivos
		05	Pensar, aprender, memória e concentração
		06	Auto-estima
		07	Imagem corporal e aparência
		08	Sentimentos negativos
III	Nível de Independência	09	Mobilidade
		10	Atividades da vida cotidiana
		11	Dependência de medicação ou de tratamentos
		12	Capacidade de trabalho
IV	Relações Sociais	13	Relações pessoais
		14	Suporte (apoio) social
		15	Atividade sexual
V	Meio Ambiente	16	Segurança física e proteção
		17	Ambiente no lar
		18	Recursos financeiros
		19	Cuidados de saúde e sociais: disponibilidade e qualidade
		20	Oportunidades de adquirir novas informações e habilidades
		21	Participação em, e oportunidades de recreação/ lazer
		22	Ambiente físico: poluição/ ruído/ trânsito/ clima
		23	Transporte
VI	Aspectos Espirituais/ Religião/ Crenças pessoais	24	Espiritualidade / religião / crenças pessoais

O Domínio Meio Ambiente é o mais complexo, composto por 8 facetas (Figura 3) que totalizam 32 % do questionário (Figura 4).

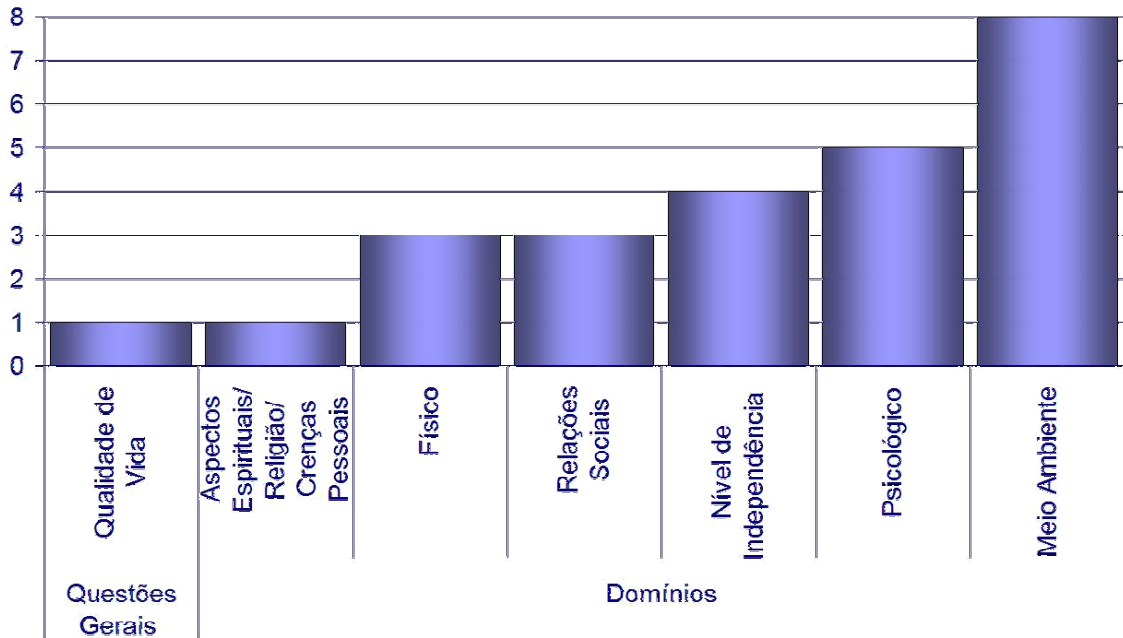


Figura 3 – Frequência de Facetas por Domínio

A Figura 3 apresenta a quantidade de facetas distribuídas por domínios do WHOQOL-100.

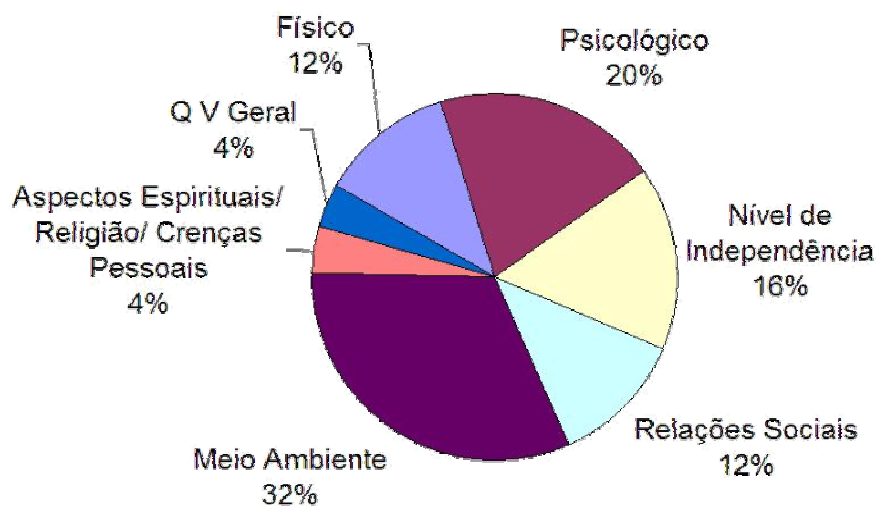


Figura 4 – Distribuição das Questões por Domínio

A partir da Figura 4, observa-se que 96% do questionário corresponde aos domínios, e 4% às Questões Gerais de Qualidade de Vida (QGQV). Estas questões



gerais (Tabela 3) examinam as maneiras em que uma pessoa avalia a sua qualidade de vida global, saúde e bem-estar.

Tabela 3 – Questões Gerais sobre Qualidade de Vida

Questão		
N	Código	Pergunta
81	G1	Como você avaliaria sua qualidade de vida?
53	G2	Quão satisfeito(a) você está com a qualidade de sua vida?
54	G3	Em geral, quão satisfeito(a) você está com a sua vida?
55	G4	Quão satisfeito(a) você está com a sua saúde?

O WHOQOL apresenta sete grupos de respostas possíveis, correspondentes à escala de Likert<sup>2</sup>, que qualificam as questões dentro de uma grandeza de avaliação, capacidade e frequência (Tabela 4).

Tabela 4 - Relação entre a Escala de Likert e as Escalas: %, Avaliação, Capacidade e Frequência.

Likert		1	2	3	4	5
(%)		0	25	50	75	100
Avaliação	Satisfação	muito insatisfeito	insatisfeito	nem satisfeito nem insatisfeito	satisfeito	muito satisfeito
	Qualidade	muito ruim	ruim	nem ruim nem bom	bom	muito bom
				nem ruim nem boa	boa	muito boa
Felicidade	muito infeliz	infeliz	nem feliz nem infeliz	feliz	muito feliz	
Capacidade	nada	muito pouco	mais ou menos	bastante	extremamente	
			médio	muito	completamente	
Frequência	nunca	raramente	às vezes	repetidamente	sempre	

As questões predominantemente tem respostas de grandeza de capacidade, como se pode observar a partir da Figura 5.

<sup>2</sup> escala de Likert – escala com grandezas gradativas

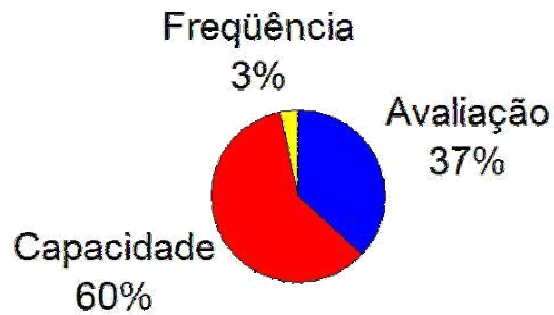


Figura 5 - Distribuição dos Tipos de Escalas Respostas do WHOQOL-100

Suas questões podem ser classificadas, quanto à natureza da pergunta, em positivas e negativas. Como exemplos de questionamento positivo, tem-se: *O quanto você aproveita a vida?* e de questionamentos negativos: *Quão facilmente você fica cansado (a)?*. Estes dois grupos de questões correspondem respectivamente a 70% e 30% do questionário.

A Tabela 5 apresenta questões negativas, suas facetas e seus domínios correspondentes que merecem destaque por receberem um tratamento pré-estatístico diferenciado das demais.

Tabela 5 – Questões Negativas do WHOQOL-100

Código	Pergunta	Faceta	Domínio
F1.2	Você se preocupa com sua dor ou desconforto (físicos)?	Dor e Desconforto	Físico
F1.4	Em que medida você acha que sua dor (física) impede você de fazer o que você precisa?		
F1.1	Com que frequência você sente dor (física)?		
F1.3	Quão difícil é para você lidar com alguma dor ou desconforto?		
F2.2	Quão facilmente você fica cansado(a)?	Energia e fadiga	Físico
F2.4	O quanto você se sente incomodado(a) pelo cansaço?		
F3.2	Você tem alguma dificuldade para dormir (com o sono)?	Sono e repouso	
F3.4	O quanto algum problema como o sono lhe preocupa?		
F6.2	Quanta confiança você tem em si mesmo?	Auto-estima	Psicológico
F8.2	Quão preocupado(a) você se sente?	Sentimentos negativos	
F8.3	Quanto algum sentimento de tristeza ou depressão interfere no seu dia-a-dia?		
F8.4	O quanto algum sentimento de depressão lhe incomoda?		
F8.1	Com que frequência você tem sentimentos negativos, tais como mau humor-desespero-ansiedade-depressão?		
F7.3	Há alguma coisa em sua aparência que faz você não se sentir bem?	Imagem corporal e aparência	Nível de Independência
F9.3	O quanto alguma dificuldade de locomoção lhe incomoda?	Mobilidade	
F9.4	Em que medida alguma dificuldade em mover-se afeta a sua vida no dia-a-dia?		
F10.4	Quanto você se sente incomodado por alguma dificuldade em exercer as atividades do dia-a-dia?	Atividades da vida cotidiana	
F10.2	Em que medida você tem dificuldade em exercer suas atividades do dia-a-dia?	Dependência de medicação ou de tratamentos	Nível de Independência
F11.2	Quanto você precisa de medicação para levar a sua vida no dia-a-dia?		
F11.3	Quanto você precisa de algum tratamento médico para levar sua vida diária?		
F11.4	Em que medida a sua qualidade de vida depende do uso de medicamentos ou de ajuda médica?		
F11.1	Quão dependente você é de medicação?	Relações pessoais	Relações Sociais
F13.1	Quão sozinho você se sente em sua vida?		
F15.4	Você se sente incomodado(a) por alguma dificuldade na sua vida sexual?		
F16.3	O quanto você se preocupa com sua segurança?	Segurança física e proteção	Meio Ambiente
F18.2	Você tem dificuldades financeiras?	Recursos financeiros	
F18.4	O quanto você se preocupa com dinheiro?		
F22.2	Quão preocupado(a) você está com o barulho na área que você vive?	Ambiente físico: (poluição/ ruído/ trânsito/ clima)	
F23.2	Em que medida você tem problemas com transporte?	Transporte	Meio Ambiente
F23.1	Em que medida você tem meios de transporte adequados?		

### 3.1.2 Agrupamento das Perguntas

As questões do WHOQOL-100 estão agrupadas em sete segmentos (Tabela 6). A maioria destes segmentos contempla diferentes facetas dentro de um único contexto. Cada segmento possui instruções e exemplos para melhor compreensão das perguntas.

O primeiro grupo contém perguntas sobre o quanto o avaliado tem sentido algumas situações, como por exemplo, sentimentos positivos, tais como, felicidade ou satisfação.

No segundo questiona-se o quanto o entrevistado tem se sentido ou é capaz de fazer certas coisas, como por exemplo, atividades diárias, tais como: lavar-se, vestir-se e alimentar-se.

O terceiro grupo contém perguntas sobre o quão satisfeito, feliz ou sentiu-se bem o entrevistado a respeito de vários aspectos de sua vida, como por exemplo, na sua vida familiar ou a respeito da sua disposição.

O quarto grupo questiona com que frequência o avaliado sentiu ou experimentou certos acontecimentos, como por exemplo, o apoio de sua família ou amigos ou se teve experiências negativas, tais como, um sentimento de insegurança.

No quinto grupo pergunta-se sobre as atividades que tomam a maior parte do tempo e disposição do entrevistado.

O sexto grupo, pergunta sobre quanto é capaz de se locomover. Está relacionado à habilidade física de mover-se, permitindo a realização de atividades de seu agrado ou por necessidade.

O último grupo refere-se às crenças pessoais, e o quanto elas afetam a qualidade de vida do entrevistado. As questões dizem respeito à religião, à espiritualidade e outras crenças.

O protocolo do questionário não permite a alteração da ordem das perguntas e do *layout* original para padronizar as pesquisas a respeito do assunto e permitir comparações (WHOQOL GROUP, 1995).

Tabela 6 – Grupos de Questões que contemplam diferentes Facetas

Grupo	As questões seguintes são sobre:	Exemplos	Explicação	Pág.	Número de Questões
1	O quanto você tem sentido algumas coisas.	Sentimentos positivos tais como: felicidade ou satisfação.	Se você sentiu estas coisas "extremamente", coloque um círculo no número abaixo de "extremamente". Se você não sentiu nenhuma destas coisas, coloque um círculo no número abaixo de "nada". Se você desejar indicar que sua resposta se encontra entre "nada" e "extremamente", você deve colocar um círculo em um dos números entre estes dois extremos.	2	39
2	Quanto completamente você tem sentido ou é capaz de fazer certas coisas.	Atividades diárias tais como: lavar-se, vestir-se e comer.	Se você foi capaz de fazer estas atividades completamente, coloque um círculo no número abaixo de "completamente". Se você não foi capaz de fazer nenhuma destas coisas, coloque um círculo no número abaixo de "nada". Se você desejar indicar que sua resposta se encontra entre "nada" e "completamente", você deve colocar um círculo em um dos números entre estes dois extremos.	6	13
3	O quanto <i>satisfeito (a), feliz ou bem</i> você se sentiu a respeito de vários aspectos de sua vida.	Na sua vida familiar ou a respeito da energia (disposição) que você tem.	Indique quão satisfeito (a) ou não satisfeito (a) você está em relação a cada aspecto de sua vida e coloque um círculo no número que melhor represente como você se sente sobre isto.	8	33
4	Com que frequência você sentiu ou experimentou certas coisas.	O apoio de sua família ou amigos ou você teve experiências negativas, tais como, um sentimento de insegurança.	Se, nas duas últimas semanas, você não teve estas experiências de nenhuma forma, circule o número abaixo da resposta "nunca". Se você sentiu estas coisas, determine com que frequência você as experimentou e faça um círculo no número apropriado. Então, por exemplo, se você sentiu dor o tempo todo nas últimas duas semanas, circule o número abaixo de "sempre".	12	3
5	Qualquer "trabalho" que você faça.	<i>Trabalho</i> aqui significa qualquer atividade principal que você faça. Pode incluir trabalho voluntário, estudo em tempo integral, cuidar da casa, cuidar das crianças, trabalho pago ou não.	<i>Trabalho</i> , na forma que está sendo usada aqui, quer dizer as atividades que você acha que tomam a maior parte do seu tempo e energia.	13	4
6	Quanto você é capaz de locomover-se bem.	Relação à sua habilidade física de mover o seu corpo,	Habilidade que permite você realizar as coisas que gostaria de fazer, bem como, as coisas que necessite fazer.	13	4
7	Suas <i>crenças pessoais</i> e o quanto elas afetam a sua qualidade de vida.	As questões dizem respeito à religião, à espiritualidade e outras crenças que você possa ter.	Uma vez mais, elas referem-se às duas últimas semanas.	14	4

### 3.1.3 Construção do Instrumento

A confecção de um questionário, à primeira vista parece ser simples, mas exige um treinamento prévio e avaliação cuidadosa das suas características temáticas, da cultura da população entrevistada e da linguagem utilizada (VASCONCELOS, 2002).

Considerando-se essas necessidades, o *WHOQOL-100*, foi desenvolvido inicialmente, com a participação de 15 centros, distribuídos pelo mundo em países de culturas e línguas diferentes (FLECK, 2002).

O instrumento foi construído em quatro estágios apresentados na Tabela 7:

**Tabela 7 - Estágios no desenvolvimento do *WHOQOL-100* (Grupo *WHOQOL*, 1995)**

Estágio	Objetivos	Método	Produto
Clarificação do conceito	Estabelecimento de um consenso para uma definição de qualidade de vida e para uma abordagem internacional da avaliação de qualidade de vida	Revisão por especialistas internacionais	Definição de qualidade de vida Definição de um protocolo para o estudo
Estudo piloto qualitativo	Exploração do conceito de qualidade de vida através das culturas e geração de questões	Revisão por especialistas dos Grupos focais Painel escrito por especialistas e leigos	Definição de domínio e subdomínios Elaboração de um conjunto de questões
Desenvolvimento de um Piloto	Refinamento da estrutura do <i>WHOQOL</i> . Redução do conjunto de questões.	Administração do <i>WHOQOL</i> piloto em 15 centros para 250 pacientes e 50 normais	Padronização de um questionário de 300 questões
Teste de campo	Estabelecimento de propriedades psicométricas do <i>WHOQOL</i>	Aplicação em grupo homogêneo de pacientes	Estrutura comum de domínios Conjunto de 100 questões Escala de respostas equivalentes nas diferentes línguas

O protocolo de aplicação do *WHOQOL-100* sugere a manutenção do mesmo layout do instrumento original, portanto não sofrerá nenhuma alteração.

### 3.1.4 Validação do instrumento

O *WHOQOL-100* encontra-se validado pela OMS, desde 1995 e, atualmente o instrumento está disponível em mais de vinte idiomas.

A sua validação ocorreu durante o estágio de aplicação do seu teste piloto, administrada em 15 centros para 250 pacientes (enfermos) e 50 indivíduos normais. Esta etapa tinha como objetivo o refinamento da estrutura do instrumento, reduzindo o conjunto de questões, o que resultou na padronização de um questionário de 300 questões (*WHOQOL GROUP*, 1995).

### 3.1.5 Versão final

O estabelecimento de propriedades psicométricas do *WHOQOL-100*, foi realizado durante a aplicação do teste de campo, no último estágio de desenvolvimento do instrumento, resultando em uma estrutura comum de domínios, composta de 100 questões, com escala de respostas equivalentes nas diferentes línguas (*WHOQOL GROUP*, 1995).

A partir desta versão final, procura-se a validação em diferentes países.

Fleck (2002), publicou os resultados da validação do *WHOQOL-100* no Brasil, destacando que o instrumento:

- apresentou boas consistências internas;
- suas características psicométricas, consistência interna, validade discriminante, validade concorrente e validade de conteúdo preencheram os critérios básicos em uma avaliação preliminar.
- a capacidade discriminante foi observada em todos os domínios exceto no domínio 6 (espiritualidade/ religião/ crenças pessoais);
- na avaliação de fidedignidade teste-reteste não houve diferença significativa entre as médias.

Segundo Fleck (2002) o *WHOQOL-100* não é apenas um instrumento para avaliar Qualidade de Vida relacionada à saúde e, pode ser útil para analisar Qualidade de Vida em outras situações que não a de saúde-doença.

### 3.1.6 Outras Versões do *WHOQOL*

A aplicação do *WHOQOL* em determinados estudos, como por exemplo, em pesquisas epidemiológicas, onde a avaliação de QV é apenas uma das variáveis em estudo, tornou-se muito demorada. Para diminuir o tempo de preenchimento do

instrumento, o Grupo de Qualidade de Vida da OMS, resolveu diminuir o número de questões de forma que preservasse as características psicométricas satisfatórias, desenvolvendo assim, uma versão abreviada do WHOQOL-100, denominada WHOQOL-bref (WHOQOL GROUP, 1998).

O WHOQOL-bref é composto de 26 questões, das quais, 02 são sobre a QV do ponto de vista do avaliado, e as demais representam cada uma das 24 facetas do WHOQOL-100. As questões foram selecionadas por critérios psicométrico e conceitual, a partir dos dados dos testes de campo de 20 centros distribuídos em 18 países (WHOQOL GROUP, 1998).

Na seleção por critérios conceituais considerou-se que o novo instrumento deveria manter a abrangência do instrumento original, permitindo a avaliação de 24 facetas. E na escolha por critério psicométrico selecionou-se apenas uma questão por faceta, ou seja, a que mais se correlacionasse com a média final de todas facetas do WHOQOL-100 (WHOQOL GROUP, 1998).

Em seguida, um grupo de especialistas, após verificar a correspondência conceitual dos domínios, substituiu seis questões, três por outras que melhor representavam seus respectivos domínios e outras três, pertencentes ao Domínio Meio Ambiente. Essas questões foram substituídas porque se assemelhavam muito com as do Domínio Psicológico (WHOQOL GROUP, 1998).

Para finalizar, realizou-se uma análise fatorial confirmatória para redução do instrumento em apenas quatro domínios. Os domínios resultantes desta análise foram: Físico, Psicológico, Relações Sociais e Meio Ambiente (WHOQOL GROUP, 1998).

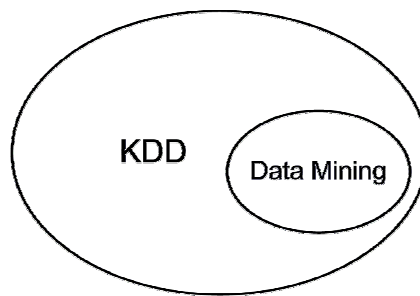
Além do WHOQOL-100 e *Bref*, os grupos focais auxiliaram no desenvolvimento de outras versões, são eles: WHOQOL-7, WHOQOL-OLD, WHOQOL-SRBP, WHOQOL-HIV, e recentemente, o WHOQOL-Tawain (YAO, 2007), que não serão abordados neste trabalho.



## 4 DATA MINING X KDD

O *Data Mining* compreende apenas uma das fases do processo KDD (*Knowledge Discovery in Databases*), referente à aplicação de algoritmos para extração de padrões em um processo de prospecção de conhecimento em bases de dados, envolvendo a automação da identificação e do reconhecimento de padrões (CARVALHO, 1999). Esse processo é um conjunto de procedimentos contínuos que compartilham o conhecimento descoberto a partir de bases de dados (CARVALHO, 1999).

A Figura 6 representa a relação entre o processo de KDD e *Data Mining*, onde KDD se refere a todo processo de descoberta do conhecimento e *Data Mining* compreende a etapa referente à extração de padrões.(CARVALHO, 1999).



Fonte: ADRIAANS, 1996

**Figura 6 - Processo de KDD x *Data Mining***

### 4.1 O PROCESSO de KDD

Historicamente, noções de descoberta de conhecimento, receberam várias nomenclaturas, incluindo mineração de dados, extração de conhecimento, descoberta de informação, arqueologia de dados, colheita de informação (AMARAL, 2001; VASCONCELOS, 2002).

Esse processo surgiu em 1989, com o propósito de encontrar o conhecimento existente em uma base de dados e enfatizar o alto nível das aplicações dos métodos de prospecção de dados (AMARAL, 2001).

Em 1995, depois da realização de uma série de *workshops* (FAYYAD & UTHURUSAMY, 1994) sobre o assunto, foi realizada na cidade de Montreal no

Canadá, a primeira Conferência Internacional de Prospecção de Dados (FAYYAD & UTHURUSAMY, 1995), quando se padronizou o termo “Descoberta de Conhecimento em Base de Dados” como KDD - *Knowledge Discovery in Databases*.

Esse processo tem por objetivo a extração do conhecimento implícito previamente desconhecido e a busca de informações potencialmente útil sobre os dados (FAYYAD, 1998).

#### 4.1.1 Etapas do Processo KDD

Segundo Fayyad (1998) são cinco os procedimentos básicos para a implantação deste processo: seleção dos dados, pré-processamento (limpeza e exploração), transformação, *Data Mining*, interpretação dos resultados (Figura 7).

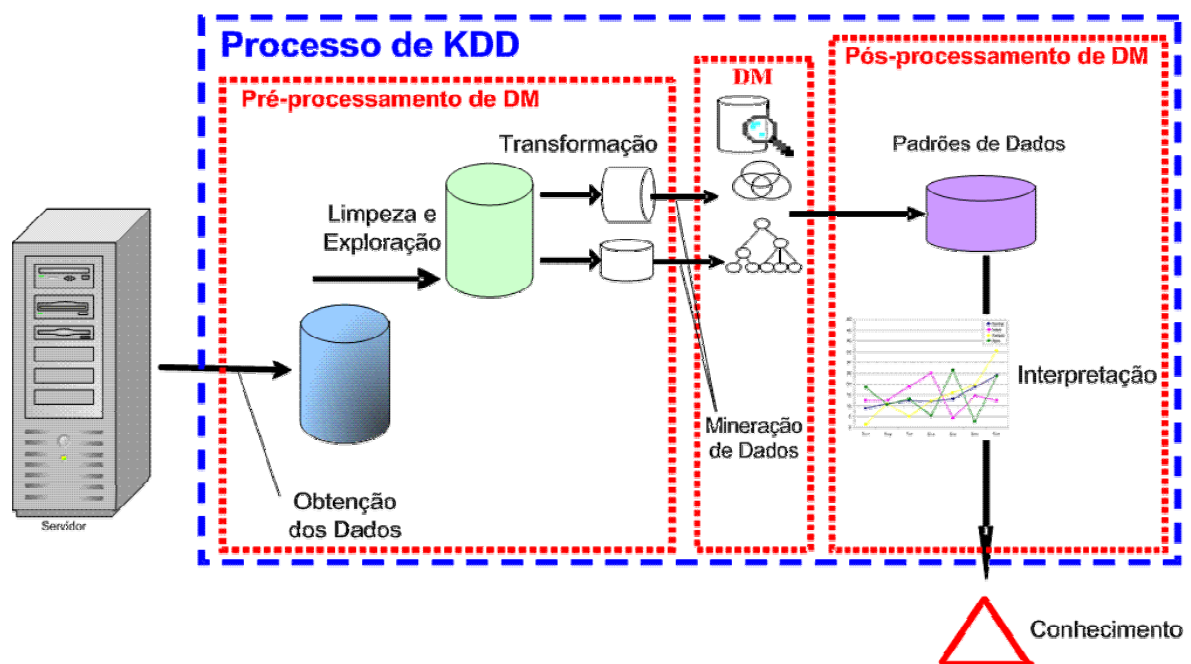


Figura 7 – Etapas do Processo KDD

#### a) Seleção dos dados

Etapa de compreensão do domínio e dos objetivos da tarefa de *Data Mining*, onde é realizada a seleção do conjunto de origem, de um subconjunto das variáveis ou, ainda, de uma amostra (CARVALHO,1999). Os dados normalmente são

extraídos de um banco de dados, ou de um *data warehouse*<sup>3</sup>, criado para servir às diversas necessidades de análise.

### **b) Limpeza e Exploração**

Nesta etapa, são decididas as estratégias e são realizadas a identificação e limpeza dos dados a fim de remover ruídos<sup>4</sup>, dados incompletos, repetições de registros, problemas de tipagem e tratar a ausência de dados de alguns atributos, visando adequá-los aos algoritmos.

### **c) Transformação**

Os dados pré-processados devem ainda passar por uma transformação que os armazena adequadamente, visando facilitar o uso das técnicas de *Data Mining*, pois existem diversos tipos de algoritmos e cada um necessita de uma entrada específica, além das conversões de dados, criação de novas variáveis e categorização de variáveis contínuas (FIGUEIRA, 1998).

Normalmente, os algoritmos de mineração não podem acessar os dados em seu formato nativo, seja em razão da forma como são armazenados, seja pela normalização adotada na modelagem do banco, então é necessária a conversão desses para um formato apropriado. Pode-se, ainda, sumarizar os dados a fim de reduzir o número de variáveis sob consideração (CARVALHO, 1999).

### **d) Data Mining**

Esta etapa caracteriza-se pela aplicação do algoritmo que diante da tarefa adequada será capaz de extrair conhecimento implícito e útil de um banco de dados (LOPES, 1999).

---

<sup>3</sup> Warehouse - armazém de dados.

<sup>4</sup> ruídos – dados estranhos e/ou inconsistentes

Consiste na efetiva aplicação do algoritmo escolhido sobre os dados a serem analisados, de acordo com a tarefa a ser cumprida, e com o objetivo de localizar os padrões desejados (CARVALHO, 1999).

#### **e) Interpretação dos resultados**

As informações resultantes das etapas anteriores são interpretadas e avaliadas de forma que se selecione o conhecimento resultante de todo o processo (CARVALHO, 1999), com possível retorno aos passos anteriores (FAYYAD et al., 1996).

Essas sub-fases não necessitam ser executadas nessa ordem, algumas dessas operações podem ser parcialmente executadas durante a construção de um Banco de Dados ou *Data Warehouse* (LOPES, 1999).

Em alto nível de abstração pode-se dizer que as etapas de KDD fazem parte de três grandes grupos: pré-processamento, aplicação de um algoritmo de *Data Mining* e pós-processamento (MICHALSKI & KAUFMAN, 1998).

## **4.2 DATA MINING**

*Data Mining* é um método de extrair informação válida, previamente desconhecida e de máxima abrangência geralmente de grandes bases de dados, usando-as para tomada de decisões, que permitem aos usuários explorar e inferir informação útil a partir dos dados, descobrindo relacionamentos escondidos no banco de dados (ADRIAANS & ZABTINGE, 1996).

O *Data Mining* é um área de estudo da Inteligência Artificial, que descende de três conceitos: da estatística, da inteligência artificial (IA) e da *Machine Learning* (LM). Do primeiro herdou seus conceitos clássicos de: distribuição normal, variância, desvio simples, análise de regressão, análise de discriminante, análise de conjuntos e intervalos de confiança; do segundo, recebeu conceitos heurísticos, de como o homem pensa para a solução de seus problemas, e do último, ganhou seus

algoritmos, desenvolvidos com a união dos conceitos de IA e LM (FAYYAD et al., 1996).

Segundo Lemos (2003), o uso de *Data Mining* para construção de modelos traz as seguintes vantagens:

- fácil compreensão;
- possibilita a análise de grandes bases de dados;
- descobre informações surpreendentes;
- as variáveis não necessitam de recodificação;
- os modelos são precisos.

#### 4.2.1 Classificação do *Data Mining*

Rodrigues (2000) desenvolveu uma proposta para classificação da temática da *Data Mining*, a partir da holística do usuário, que pode auxiliar no processo de implementação de um projeto de descoberta do conhecimento.

Segundo Rodrigues (2000), a partir da visão do usuário, o *Data Mining* pode ser classificado por três dimensões:

- Tipos de Problemas: através dos problemas gerais de *Data Mining*, independente do tipo de técnica ou ferramenta utilizada.
- Métodos de Abordagem: a partir das principais metodologias ou ferramentas para solução de problemas de *Data Mining*.
- Áreas de Aplicação: através das áreas onde os tipos de problemas e métodos de abordagem interagem gerando soluções.

#### 4.2.2 Objetivos de *Data Mining*

As diversas tarefas desenvolvidas em *Data Mining* têm como objetivo elementar, a predição e a descrição (FAYYAD et al, 1996).

A predição utiliza atributos para prognosticar o desconhecido ou os valores

futuros de outras variáveis de interesse, enquanto a descrição contempla o que foi descoberto nos dados sob o ponto de vista da interpretação humana (Carvalho, 1999).

Em outras palavras, os objetivos de *Data Mining* são de natureza (YIN, 2001):

- Explanatória: explicar algum evento ou medida observada. Questionamentos como, por exemplo: Porque a venda de sorvetes caiu em Ponta Grossa?
- Confirmatória: confirmar uma hipótese. Uma companhia de seguros, por exemplo, pode querer examinar os registros de seus clientes para determinar se famílias de duas rendas têm mais probabilidade de adquirir um plano de saúde do que famílias de apenas uma renda.

#### 4.2.3 Tarefas e Técnicas de *Data Mining*

As tarefas de KDD são dependentes do domínio da aplicação e do interesse do usuário. Cada tarefa de KDD extrai um tipo diferente de conhecimento do banco de dados, podendo requerer um algoritmo diferente para a extração de conhecimento (LOPES, 1999).

Segundo Viana (2004) as principais classes de tarefas de KDD, são: associação, classificação, *clustering*.

Técnicas de *Data Mining* são procedimentos diferentes de exploração dos dados, utilizados para retirar diferentes formas de conhecimento.

Alguns autores (RODRIGUES, 2000), apresentam um grande número de técnicas de *Data Mining*. No entanto, cinco delas englobam as demais (CARVALHO, 1999). São elas: técnicas de redes neurais, regras de indução, árvores de decisão, análise de séries temporais e visualização (CARVALHO, 1999; CARVALHO, 2001).

A Tabela 8 apresenta as técnicas mais utilizadas na mineração de dados de acordo com as principais tarefas de KDD:

**Tabela 8 - Problemas e respectivas Tarefas e principais Técnicas de *Data Mining* utilizadas para sua solução**

Problema	Tarefa	Técnica	Exemplos de Algoritmos mais Utilizados
Predição	Regressão	Regressão	Cart, M5
	Classificação	Árvore de Decisão	J48
Descrição	Associação	Regras de Associação	Apriori, Dhp, Abs, Sampling
	Agrupamento	Particionamento	K-Means
		Hierarquia	Hac

Fonte: Criada a partir das definições de Viana (2004).

As técnicas de *Data Mining* são de caráter universal e podem ser praticadas por meio de algoritmos diferentes.

A escolha da técnica mais apropriada depende dos objetivos da pesquisa, do tipo de extração desejada e das características dos dados.

#### 4.2.3.1 Tarefa de Associação

A tarefa de Associação (AGRAWAL & SRIKANT, 1994), tem como objetivo descobrir regras de associação (RA), que são expressões:  $X \rightarrow Y$ , onde  $X$  e  $Y$  são conjuntos de itens,  $X \cap Y = \emptyset$ . Cada regra é um relacionamento SE ( $X$ ) ENTÃO ( $Y$ ), onde os conjuntos de itens  $X$  e  $Y$  freqüentemente ocorrem juntos em uma mesma transação (registro), que indicam que na base de dados contém instâncias nas quais existe o conjunto  $X$  e também contem a presença de  $Y$ .

Um exemplo de uma regra do tipo  $X \rightarrow Y$  poderia ser: 90% dos consumidores que compram carvão e carne (elemento  $X$ ), também compram cerveja (elemento  $Y$ ). O valor 90% representa o valor de confiança da regra, ou seja, representa o número de consumidores que compram cerveja, dividido pelo número de consumidores que compraram carvão e carne.

Além do valor de confiança, uma RA pode ser avaliada através de outras medidas de interesse que serão apresentadas no capítulo 5 trabalho.

Existem vários algoritmos de tarefas de associação, entre eles, o algoritmo *Apriori*, muito conhecido e eficiente para essa finalidade (REFFAT et al, 2004) e que será utilizando neste trabalho.

O Apriori foi desenvolvido por Agrawal e Srikant (1994), e está disponibilizado gratuitamente na Internet.

O algoritmo permite trabalhos com grande número de atributos, gerando várias alternativas combinatórias entre eles. Realiza buscas contínuas em toda a base de dados, mantendo um ótimo desempenho em termos de tempo de processamento (AGRAWAL & SRIKANT, 1994).

Os algoritmos de tarefa de associação têm a desvantagem da exigência de maior tempo de pós-processamento, principalmente quando são envolvidos muitos itens, tornando trabalhosa a identificação de regras interessantes.

Para amenizar este problema, os pesquisadores, entre eles, Brin et. al. (1997), desenvolveram medidas de interesses, que serão descritas no capítulo 5 deste trabalho.

#### 4.2.4 Softwares de *Data Mining*

Atualmente, existem inúmeros programas de *Data Mining* vendidos comercialmente.

Segundo Rodrigues (2000) os softwares de *Data Mining* podem ser classificados em sete grupos de acordo com suas principais características.

- **Suítes:** são pacotes com acessórios para todo o KDD; desde interfaces para acesso a banco de dados, manipulação de dados, desenvolvimento do modelo e auxílio na análise dos resultados.
- **Softwares de Classificação:** possuem a especialidade de gerar modelos de classificação a partir de um banco de dados. Podem utilizar diferentes abordagens como classificação múltipla, árvores de decisão, regras de associação, redes neurais, e estatística bayesiana.
- **Softwares de *Clustering*:** são utilizados para a determinação de *clusters* e subgrupos do conjunto de dados.



- **Softwares de Análises de Relacionamentos e Associações:** utilizados para a determinação de regras de associações e construção de redes de relacionamentos.
- **Softwares de Padrões Seqüenciais:** aplicações utilizadas para a descoberta de padrões seqüenciais no banco de dados.
- **Softwares de Transformação de Dados e Limpeza:** utilizado nas etapas de pré-processamento de *Data Mining*.

Além dos softwares proprietários existem os softwares livres, como por exemplo, o WEKA (*Waikato Environment for Knowledge Analysis*), utilizado no desenvolvimento desta pesquisa.

Este software possui um ambiente gráfico desenvolvido em linguagem Java pelo Departamento de Ciência da Computação da Universidade de Waikato na Nova Zelândia (WITTEN & FRANK, 2005).

O WEKA compreende um conjunto de implementações de algoritmos de pré-processamento de dados e de diversas técnicas de *Data Mining*.

Neste ambiente, após a etapa de transformação dos dados para formato *arff* (*attribute relation file format*), um único arquivo de entrada, pode ser utilizado por diferentes algoritmos nas demais etapas do processo de KDD.

Durante a etapa de conhecimento do domínio, é possível gerar informações estatísticas básicas (média, máximo, mínimo e desvio padrão) de cada atributo e, também, gerar gráficos da freqüência dos dados correlacionando-os com um atributo de referência.

Ferramentas de *Data Mining* em ambientes de softwares livres como, por exemplo, o WEKA, permite às empresas de pequeno porte, o acesso às tecnologias de ponta, que podem auxiliá-las em suas tomadas de decisões e na economia de gastos com licenças de softwares.

## 5 MEDIDAS DE INTERESSE

As Medidas de Interesse são indicadores estatísticos utilizadas para avaliar a qualidade de uma forma de representação padrões e podem ser classificadas em Subjetivas e Objetivas (MCGARRY, 2005).

Nas Medidas de Interesse Subjetivas, o grau de interesse da regra é avaliado pelo usuário, levando em consideração as suas expectativas e conhecimento prévio dos dados (SILBERSCHATZ & TUZHILIN, 1995). Enquanto, nas medidas de interesse objetivas, as regras são mensuradas a partir de unidades estatísticas.

A adoção da técnica de avaliação mais adequada depende do grau de integração do usuário no processo de descoberta de conhecimento.

Neste estudo as medidas de interesse foram utilizadas para filtrar regras de associação.

### 5.1 Medidas de Interesse Subjetivas

As técnicas subjetivas de avaliação de uma regra, geralmente operam comparando as crenças de um usuário contra os padrões descobertos através de *Data Mining* (MCGARRY, 2005).

Através dessas medidas normalmente se determina, se um padrão é “ útil ” e/ou “inesperado”. No entanto, o padrão encontrado, poderá receber diferentes julgamentos, dependendo do usuário e dos seus objetivos (SILBERSCHATZ & TUZHILIN, 1995).

### 5.2 Medidas de Interesse Objetivas

As medidas de interesse objetivas mais conhecidas para avaliar uma RA são Suporte e a Confiança. A primeira corresponde ao índice de ocorrências de registros com os mesmos antecedentes da RA na base de dados, e o segundo corresponde ao índice de registros com os mesmos antecedentes que atendem a RA.

A utilização apenas destes parâmetros de medidas, proporciona a geração de muitas regras, entre elas, regras óbvias, de baixo interesse, redundante ou contraditório, tornando a análise muitas vezes impraticável.

Para amenizar estes problemas, pesquisadores desenvolveram outras medidas estatísticas para avaliar as regras, entre eles, Tan (2002) apresenta em seu estudo aproximadamente vinte medidas de interesse.

Neste trabalho foram utilizadas cinco medidas, disponíveis no software WEKA, que serão descritas a seguir, são elas: Suporte, Confiança, *Lift*, *Rule Interest* (RI), Convicção.

### 5.2.1 Suporte

O suporte corresponde ao índice de frequência de um determinado item ou grupo de itens na base de dados é calculado através da Equação 1.

$$S_{(x)} = \frac{R_{(x)}}{Tr} \quad (1)$$

onde:

$S_{(x)}$  é o índice de suporte;

$R_{(A)}$  é o número de registros com o objeto x; e

Tr é o número total de registros na base de dados.

### 5.2.2 Confiança

Uma regra de associação é composta de dois elementos, um antecedente (condição) e um conseqüente (resposta). O antecedente deve implicar ( $\rightarrow$ ) na ocorrência do conseqüente.

Dada a seguinte RA:  $A \rightarrow B$ , o seu valor de confiança é calculado a partir da Equação 2.

Esta medida representa, dentre as transações que contém o item “A”, o índice de vezes que as transações que também contém o item “B”

$$C_{(A \rightarrow B)} = \frac{S_{(A \cup B)}}{S_{(A)}} \quad (2)$$

onde:

$C_{(A \rightarrow B)}$  é o índice de confiabilidades da RA;

$S_{(A \cup B)}$  é suporte do antecedente e conseqüente, juntos; e

$S_{(A)}$  é o suporte do antecedente.

Os valores de confiança variam de 0 a 1.

### 5.2.3 Lift

O valor de *Lift*, utilizado para verificação do índice de dependência do conseqüente em relação ao seu antecedente, é calculado a partir da Equação 3.

$$L_{(A \rightarrow B)} = \frac{C_{(A \rightarrow B)}}{S_{(B)}} \quad (3)$$

onde:

$L_{(A \rightarrow B)}$  é o índice de *Lift* da RA;

$C_{(A \rightarrow B)}$  é a confiabilidade da RA; e

$S_{(B)}$  é o suporte do conseqüente na base de dados.

O valores de *Lift* variam de 0 a  $\infty$ .

O valor do índice igual a 1 indica independência ( $Lift = 1$ ) entre os elementos da RA. Enquanto para os demais valores indicam, dependência positiva ( $Lift > 1$ ) ou negativa ( $Lift < 1$ ).

### 5.2.4 Rules Interest

O valor de RI, também conhecido por *Leverage*, nome do seu criador, também é utilizado para verificação do índice de dependência do conseqüente em relação ao seu antecedente, é calculado a partir da Equação 4:

$$RI_{(A \rightarrow B)} = S_{(real)} - S_{(Esperado)} \quad (4)$$

onde:

$RI_{(A \rightarrow B)}$  é o índice de *Lift* da RA;

$S_{(real)}$  é o suporte rel, equivalente a  $S_{(A \cup B)}$  ; e

$S_{(esperado)}$  é o suporte esperado, equivalente a  $S_{(A)} * S_{(B)}$  .

Os valores dos índices de RI são bem menores que os de *Lift*, e variam de  $-\infty$  a 0.

O valor do índice igual a zero indica independência ( $RI = 1$ ) entre os elementos da RA. Enquanto para os demais valores indicam dependência positiva ( $RI > 1$ ) ou negativa ( $RI < 1$ ).

A diferença entre os dois índices: *Lift* e RI, está relacionada ao valor de suporte, o primeiro é mais preciso para identificar o nível de dependência entre elementos de RA de maior suporte, enquanto o segundo, para os de menor suporte.

Os dois índices apenas apresentam o grau de dependência entre os antecedentes e conseqüente, porém, não identificam o elemento implicado.

### 5.2.5 Convicção

A medida de *Convicção* foi introduzida por Brin et. al. (1997) e é utilizada para indicar o grau de implicação do antecedente no conseqüente, que é calculado a partir da Equação 5.

Segundo Brin et. al. (1997), as regras potencialmente implícitas (consideradas interessantes), possuem o índice de convicção entre 1.1 a 5.0.

$$CV_{(A \rightarrow B)} = \frac{S_{(A)} * (1 - S_{(B)})}{S_{(A)} - S_{(A \cup B)}} \tag{5}$$

onde:

$CV_{(A \rightarrow B)}$  é o índice de convicção da ra;

$S_{(A)}$  é o suporte do antecedente;

$S_{(\neg B)}$  é o suporte sem o conseqüente, equivale a  $1 - S_{(B)}$ ; e

$S_{(A \cup B)}$  é suporte do antecedente sem o conseqüente, equivale a  $S_{(A)} - S_{(A \cap B)}$

O valor do índice igual a zero indica independência ( $CV = 0$ ) entre os elementos da RA. Enquanto para os demais valores, indicam dependência positiva ( $CV > 0$ ) ou negativa ( $CV < 0$ ).

Para exemplificar, temos na Tabela 9, à esquerda, uma base de dados com 10 registros compostos de três itens cada e, à direita, a distribuição de cada item na base de dados:

**Tabela 9 – Itens dos Registros em uma Base de Dados e sua Distribuição**

N Registro	Itens			Distribuição dos itens na base de dados							
				♣	♠	♥	♦	●	■	○	
1	♣	●	♦	X			X	X			
2	♣	●	♠	X	X			X			
3	♣	●	♠	X	X			X			
4	♣	●	♠	X	X			X			
5	♣	♦	■	X			X			X	
6	♥	♦	♠		X	X	X				
7	♥	■	♠		X	X				X	
8	♥	■	♠		X	X				X	
9	♥	○	♦			X	X				X
10	♥	○	■			X				X	X
Total				5	6	5	4	4	4	4	2
Suporte				0.5	0.6	0.5	0.4	0.4	0.4	0.4	0.2

Adotou-se uma representação simbólica dos itens, para não se gerar pré-concepções em relação à dependência dos dados, pois elas poderiam interferir subjetivamente na identificação de RA ditas interessantes.

O suporte de cada item do banco de dados (objeto) foi calculado através da Equação 1.

Para exemplificar a aplicação dos índices, no processo de seleção objetiva das RA, temos as seguintes regras como suas respectivas unidades de interesse:

$$(1) \clubsuit \rightarrow \bullet \quad S_{(\clubsuit)} = 0.4; S_{(\bullet)} = 0.5; S_{(\clubsuit \cup \bullet)} = 0.4; C = 1; L = 2; RI = 0.2; Cv = 0$$

$$(2) \bullet \rightarrow \clubsuit \quad S_{(\bullet)} = 0.5; S_{(\clubsuit)} = 0.4; S_{(\bullet \cup \clubsuit)} = 0.4; C = 0.8; L = 2; RI = 0.2; Cv = 3$$

$$(3) \blacklozenge \rightarrow \spadesuit \quad S_{(\blacklozenge)} = 0.4; S_{(\spadesuit)} = 0.6; S_{(\blacklozenge \cup \spadesuit)} = 0.1; C = 0.25; L = 0.42; RI = -0.14; Cv = 0.53$$

$$(4) \spadesuit \rightarrow \blacklozenge \quad S_{(\spadesuit)} = 0.4; S_{(\blacklozenge)} = 0.4; S_{(\spadesuit \cup \blacklozenge)} = 0.1; C = 0.17; L = 0.42; RI = -0.14; Cv = 0.72$$

$$(5) \blacklozenge \rightarrow \heartsuit \quad S_{(\blacklozenge)} = 0.4; S_{(\heartsuit)} = 0.5; S_{(\blacklozenge \cup \heartsuit)} = 0.2; C = 0.5; L = 1; RI = 0; Cv = 1$$

$$(6) \heartsuit \rightarrow \blacklozenge \quad S_{(\heartsuit)} = 0.5; S_{(\blacklozenge)} = 0.4; S_{(\heartsuit \cup \blacklozenge)} = 0.2; C = 0.4; L = 1; RI = 0; Cv = 1$$

Observando-se as RA 1 e 2, a partir das unidades de interesse de *Lift* e *RI*, conclui-se que existe dependência positiva entre o antecedente e conseqüente. E a partir do índice de Convicção, sabe-se que o item “●” depende do item “♣”, e o inverso não ocorre.

As RA 2 e 4 indicam que existe dependência negativa entre os termos “◆” e “♠”; que a medida que um deles ocorre na base de dados, diminui a possibilidade da ocorrência do outro e; que existe a possibilidade do elemento “◆” não ocorrer em função da ocorrência de “♠” de 0.53 vezes (RA 2), enquanto, a possibilidade do elemento “♠” não ocorrer em função da ocorrência de “◆” de 0.72 vezes (RA 4).

Como exemplos de RA que indicam independência entre os itens, temos as RA 5 e 6. Estas indicam que os elementos “♦” e “♥”, apesar de aparecerem juntas em 20% das transações, não dependem um do outro para ocorrerem.



## 6 METODOLOGIA

Neste capítulo estão descritos: o desenho da pesquisa; o instrumento e fonte de pesquisa e intervenção em ambientes sociais e institucionais; dados da pesquisa, aplicação do instrumento e, a análise dos dados na pesquisa.

### 6.1 DESENHO DA PESQUISA

Segundo Vasconcelos (2002) a pesquisa pode ser classificada sob 5 aspectos principais, quanto: ao enquadramento do objeto; o tipo do objeto e as fontes a serem investigadas; a natureza dos dados e da análise; os objetivos, aplicações e tipo de conhecimento implicado.

Quanto ao objeto essa pesquisa é classificada como Temática ou Focal simples, pois eleger como objeto uma temática específica, com foco preciso para produção de um novo conhecimento com contribuição para os debates teóricos, científicos e técnicos com subsídio para tomada de decisões de caráter gerencial.

Quanto ao tipo do objeto e as fontes a serem investigadas classifica-se como Meta-análise, pois constitui um tipo de análise secundária, ainda não utilizada, usada para comparar os dados internos dos estudos quantitativos.

Quanto à natureza dos dados e da análise é classificada como Mista, pois a natureza dos dados é qualitativa e quantitativa.

Quanto aos objetivos, aplicações e tipo de conhecimento implicado é considerada Aplicada, pois articula dentro de diferentes ciências (Engenharia de Produção, Saúde, Ciências da Computação).

Considerando a definição da lógica que determina as possibilidades de generalização dos resultados encontrados e a combinação de métodos quantitativos e qualitativos, a pesquisa também pode ser classificada com empírica, com lógica de replicação, pois a solidificação desse conhecimento se faz por meio das avaliações posteriores da comunidade científica e de novos testes através de replicações integrais ou parciais dos mesmos procedimentos de investigação (VASCONCELOS, 2002).

Em relação aos níveis de análise e do número de casos, aplicou-se o *WHOQOL-100* através de amostragem probabilística por conglomerado, atingindo grupos de indivíduos de diversos níveis e segmentos, totalizando 267 agentes universitários (31,84% da população), equivalentes a 45,95% dos funcionários lotados no Campus da UEPG, em Uvaranas.

A relação entre o tempo em que se realiza a pesquisa e o tempo do fenômeno observado é classificado como “durante”, pois o processo principal de investigação é construído concomitantemente à ocorrência do fenômeno (VASCONCELOS, 2002).

Considerando os tipos de corte no tempo do fenômeno e o significado teórico e intrínseco das marcas temporais feitas no tempo do fenômeno, será utilizado na pesquisa um corte transversal para focar o objeto de investigação, pois se busca a análise do fenômeno em um momento específico (VASCONCELOS, 2002).

## **6.2 INSTRUMENTO, FONTE DE PESQUISA E INTERVENÇÃO EM AMBIENTES SOCIAIS E INSTITUCIONAIS**

Segundo Vasconcelos (2002), os recursos e fontes desta pesquisa se classificam como primários e as fontes bibliográficas como secundárias. A primeira classificação se dá pela falta de tratamento analítico em relação às informações e dados. A segunda classificação é atribuída ao fato da fonte de pesquisa contemplar em tese todos os tipos de investigação.

Quanto ao tipo de entrevista, o *WHOQOL* é classificado como, estruturada para *surveys*, pois possui redação fixa e perguntas cuja ordem, formulação e estratégia de entrevista são invariáveis para todos os informantes.

Outra característica dos modelos *surveys*, segundo Vasconcelos (2002) é a possibilidade de tratamento estatístico das respostas, também presente no instrumento *WHOQOL-100*.

## **6.3 DADOS DA PESQUISA**

Uma variável consiste em qualquer atributo e um fenômeno classificado em duas ou mais categorias, pois os acontecimentos sociais e humanos sempre

apresentam predicados que descrevem suas características (VASCONCELOS, 2002).

As variáveis podem ser classificadas em simples, quantitativas, descontínuas ou discretas, contínuas e multidimensionais complexas.

Em uma concepção de causalidade mecanicista, as variáveis, ainda podem ser classificadas em: variáveis independentes, dependentes e “outras” (VASCONCELOS, 2002). A segunda sofre, interferência da primeira e as duas juntas, sofrem influências da terceira.

Neste estudo foram utilizados os diversos tipos de variáveis, apresentados por Vasconcelos (2002), porém, todas passaram por um processo final de categorização, antes da aplicação do *Data Mining*.

#### 6.4 APLICAÇÃO DO INSTRUMENTO

O instrumento possui um protocolo de aplicação da entrevista, disponível no *site* do grupo de pesquisa focal do Brasil<sup>5</sup>.

As orientações para sua aplicação serão descritas nos próximos parágrafos, com algumas modificações.

Inicialmente, o questionário (WHOQOL-100) foi distribuído para aproximadamente 300 agentes universitários, no período de julho a setembro de 2007, que foram orientados individualmente, quanto ao preenchimento do instrumento, de forma auto-aplicável.

Para motivar o preenchimento do instrumento e para maior fidedignidade das respostas, prometeu-se a cada entrevistado, um relatório individual (Anexo D), diagnosticando a sua qualidade de vida, por meio de dois modelos, um gráfico e outro categórico, que informavam o seu nível de qualidade de vida.

Este procedimento, de caráter inovador, garantiu um retorno de 88.67% dos questionários distribuídos na pesquisa.

---

<sup>5</sup> <http://www.ufrgs.br/psiq/WHOQOL-manual.html>

Os entrevistados foram informados sobre o objetivo da aplicação da ferramenta, o modo de aplicação e o destino dos dados obtidos. E que deveriam ficar à vontade para esclarecimentos de dúvidas ao longo da aplicação.

Recomendou-se ao entrevistado que, ao concordar em responder o questionário, seria essencial a obtenção do seu consentimento, assinado pelas duas partes.

Aconselhou-se ao avaliado de que não deveria responder o instrumento acompanhado de familiar, cônjuge ou amigos.

Explicou-se, que as questões deveriam ser respondidas em um único momento.

Não foi realizado o preenchimento da folha de dados demográficos, como sugere o protocolo de pesquisa, pois cada questionário, foi identificado por um código, através do qual foram recuperados os dados cadastrais do entrevistado, sem a sua identificação nominal.

## **6.5 ANÁLISE DOS DADOS**

A análise de dados foi realizada em dois momentos. No primeiro foi utilizado o protocolo tradicional (*WHOQOL GROUP*, 1995), porém, sem a utilização do *software* SPSS. E no segundo momento, utilizou-se o Processo de KDD, como um protocolo complementar, destinado à aquisição de conhecimento implícito.

## 7 ANÁLISE DOS RESULTADOS DO WHOQOL-100 UTILIZANDO DATA MINING

Nesta proposta de pesquisa, as técnicas de *Data Mining* foram utilizadas para encontrar conhecimento implícito no relacionamento entre os resultados do WHOQOL-100 e informações cadastrais da amostra.

Para utilização das técnicas de *Data Mining* na análise dos dados WHOQOL-100, foram utilizadas as etapas do processo de KDD (Figura 8), que serão descritas a seguir.

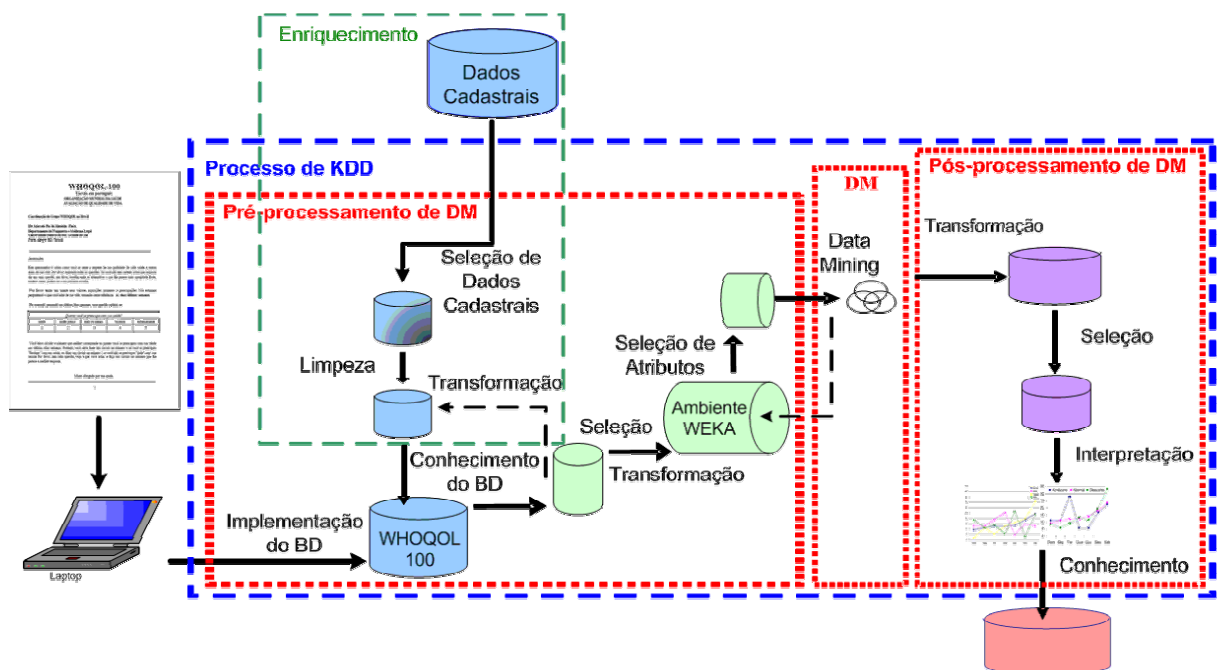


Figura 8 – Etapas do Processo de KDD Desenvolvidas na Pesquisa

O processo de KDD, segundo Michalski & Kaufman (1998), é composto basicamente de três macro-etapas: Pré-processamento de *Data Mining*, *Data Mining* e Pós-processamento de *Data Mining*.

### 7.1 Pré-processamento para *Data Mining*

Esta etapa foi realizada através de cinco sub-etapas. Iniciou-se com a Implementação da Base de Dados e foi finalizada com a transformação dos dados.

### 7.1.1 Implementação da Base de Dados

A base de dados foi implementada a partir do entendimento da sintaxe desenvolvida para o Software SPSS (GRUPO *WHOQOL*, 1995).

O banco de dados foi composto, basicamente de um formulário de entrada, que alimentava uma tabela na escala de *Likert*. Esta tabela serviu de referência para a atualização de quatro tabelas, assim denominadas: Dados Categóricos; Resultados das Questões em escala centesimal; Resultados das Facetas e Domínios na escala de *Likert*; Resultados das Facetas e Domínios em escala centesimal; e Resultados Categóricos.

A Tabela "Dados Categóricos" foi atualizada com as respostas categóricas de cada questão do *WHOQOL-100*.

Para atualização da Tabela "Resultados das Facetas e Domínios para a escala de *Likert*", foi elaborada uma SQL baseada na relação de dependência dos dados (Figura 9).

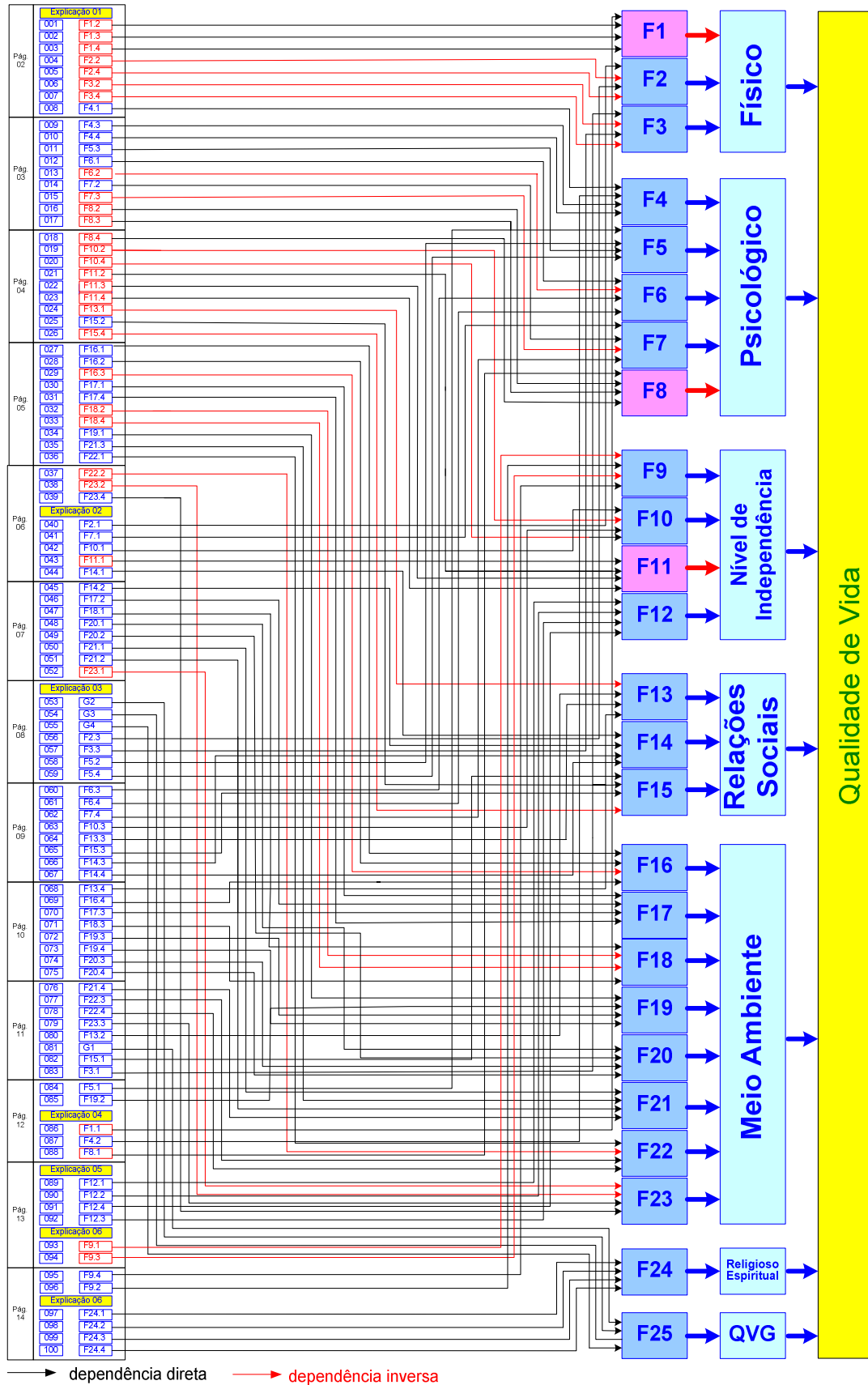


Figura 9 – Relação de Dependência entre as Questões, Facetas e Domínios de QV

Segundo Sferra e Corrêa (2003), um modelo descreve as dependências significativas entre variáveis. Elas podem ocorrer em dois níveis: estruturado e quantitativo. No nível estruturado o grau de dependência é especificado geralmente em forma de gráfico, apresentando quais variáveis estão localmente dependentes. O nível quantitativo especifica o grau de dependência, usando alguma escala numérica.

A análise de dependência visa o estudo da dependência de uma ou mais variáveis em relação a outras e, pode ser classificada, segundo Castro (1995):

- quanto à dependência: funcional ou aleatória.
- quanto à tendência: linear ou não, direta ou inversa;
- quanto à relação: dependência causal, de interdependência, dependência indireta, concordância e correlação espúria<sup>6</sup>.

Em função do conhecimento prévio das relações dos dados e da sua origem podemos classificar os atributos, segundo Crespo (2002), em dependência funcional linear direta e dependência funcional linear inversa (Figura 9).

Através do entendimento dos comandos e artifícios matemáticos utilizados na Sintaxe do SPSS, e também do entendimento das relações de dependência entre dados (respostas na escala *Likert*) e resultados (facetos e domínios), foram desenvolvidas SQL para atualização de tabelas com resultados.

As facetos compostas por questões negativas e positivas e os domínios compostos por facetos com esta característica, são tratados na sintaxe através de recursos matemáticos que possibilitam a inversão dos escores originais. Daí a importância em identificar as relações de dependência de proporções diretas ou inversas entre os atributos para evitar erros na análise dos dados.

Devido a essa necessidade de inversão de escalas de respostas, foi desenvolvida uma equação matemática e incorporada às SQL de atualização (Figura 10) para a análise de dos dados.

---

<sup>6</sup> correlação espúria - é um equívoco de interpretação de relações de mera causalidade, associações temporais entre dois fenômenos. Pois o fato de dois fenômenos ocorrerem ao mesmo momento não permite a dedução de que um seja ocasionado pelo outro, ou seja, a causa dele (BATANERO,1999).



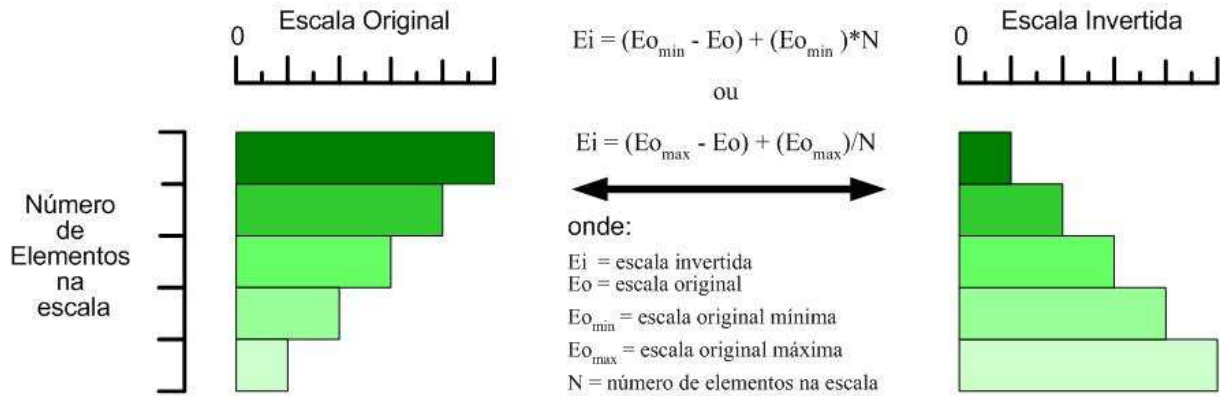


Figura 10 – Equação para Inversão de Escala

A sintaxe original do SPSS para *WHOQOL-100* (WHOQOL GROUP, 1995), não transforma os escores finais da escala de Likert para uma escala centesimal, como ocorre na sintaxe do *WHOQOL-BREF* (WHOQOL GROUP, 1998).

Para melhor compreensão e comparação dos indicadores de QV, em relação aos valores máximos e mínimos possíveis, os dados foram transformados na Tabela “Resultados em escala centesimal”, através da seguinte equação, incluída em uma SQL:

$$E_{(Centesimal)} = \frac{(E_{(Likert)} - 4) * 25}{4} \tag{7}$$

onde:

$E_{(Centesimal)}$  é escala centesimal; e

$E_{(Likert)}$  é escala likert.

Como as técnicas de DM de RA utilizadas na pesquisa requerem dados categorizados, buscou-se na literatura uma classificação para os resultados do *WHOQOL-100*. Observou-se, no entanto, que os trabalhos existentes, geralmente apresentam comparações entre grupos ou entre estados temporais das amostras, como podem ser observados no trabalho de Kluthcovsky (2005), onde a pesquisadora compara seus resultados com outros trabalhos, sem possuir um parâmetro de classificação.

O grupo WHOQOL (1995) menciona que não existe uma normalização padrão para os resultados, e que havendo essa necessidade, sugerem que seja realizada através do seu respectivo grupo focal.

Tendo em vista a necessidade de classificação dos resultados e, conhecendo-se os valores máximo e mínimo da escala, adotou-se essa escala centesimal, composta de cinco medianas: 10, 30, 50, 70 e 90, todas com um desvio padrão de 9.9 unidades centesimais.

Através de SQL, os resultados da Tabela “Resultados em escala centesimal” formam classificados em variáveis ordinárias de avaliação de QV na Tabela “Resultados Categóricos”, considerando as características dos Aspectos ou Domínio e os intervalos da Tabela 10.

Tabela 10 - Classificação da Nível de Qualidade de Vida

Classe	Aspectos ou Domínios Positivos		Aspectos Negativos	
	$\geq$	$<$	$\geq$	$<$
Péssimo	0	20	80	=100
Ruim	20	50	60	80
Médio	40	60	40	60
Bom	60	80	20	40
Excelente	80	=100	0	20

Esta classificação pode ser melhor representada através da Figura 11 em uma escala de *Likert*.

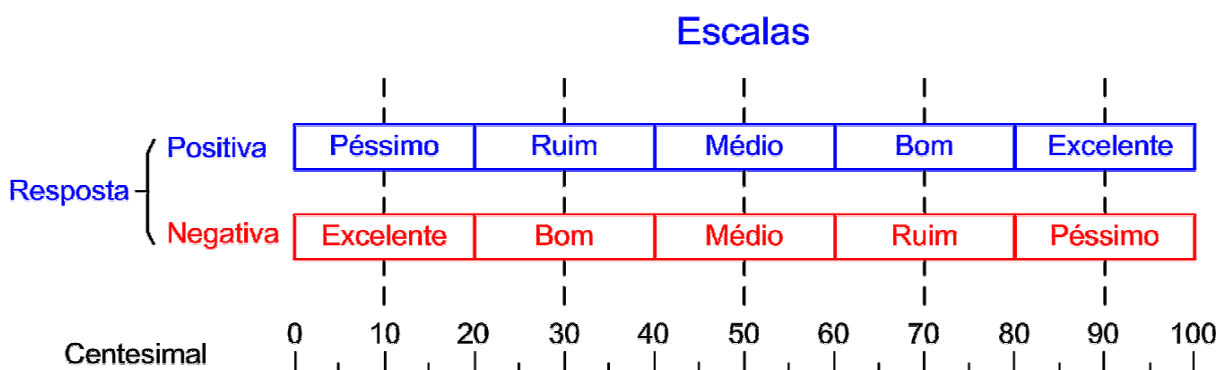


Figura 11 - Escala de Classificação dos Resultados

Para avaliar a consistência interna entre questões, facetas e domínios, após a categorização dos resultados, utilizou-se um filtro para seleção de atributos, na expectativa de que, cada domínio, quando selecionado como atributo meta<sup>7</sup>, apresentasse dependência por suas respectivas facetas e, essa conseqüentemente, pelas questões do questionário.

Para validar esta proposta de classificação, foram utilizados 793 registros do WHOQOL-100, colhidos de pesquisas realizadas em outros momentos, em distintas áreas, distribuídos na Tabela 11.

**Tabela 11 – Relação de Estudos**

LOCAL / PROFISSÃO	PERÍODO	N
Receita Federal / TTN	2007	54
Universidade / Agentes Universitários	2007	266
Universidade / Docentes	2007	19
Faculdade / Acadêmicos	2007	130
Indústria Cervejarias / Diversas	2006	114
Indústria Cervejarias / Diversas	2006	110
Indústria de Grãos / Diversas	2006	60
Indústria Metalúrgica / Diversas	2005	20
Agência De Correios / Carteiros	2004	40
Total		793

Todos os atributos e metas atenderam à expectativa, apresentando entre seus atributos previsores, seus respectivos atributos dependentes.

Outra análise foi realizada utilizando-se o coeficiente de *Conbrach*, usado para verificar o índice de consistência interna dos dados, chegando-se aos resultados da Tabela 12.

<sup>7</sup> atributo meta – termo adotado neste trabalho para identificar o atributo de referência utilizado na seleção dos atributos relevantes de um grupo de atributos. Geralmente este termo é utilizado em tarefas de *Data Mining* classificação.

**Tabela 12 - Índice de Consistência Interna dos Indicadores de QV Antes e Após a Categorização dos Dados**

Itens	Cronbach's Alpha		Número de casos		Número de Itens
	A	B	A	B	
Questões	0.897	0.897	762	762	100
Facetas	0,840	0.890	765	769	25
Domínios	0.804	0.771	766	769	6
Questões + Facetas	0.914	0.922	762	762	125
Domínios + MQV	0.853	0.828	766	769	7
Facetas + Domínios	0.894	0.920	765	769	31
Facetas + Domínios + MQV	0.902	0.924	765	769	32
Questões + Facetas + Domínios	0.930	0.928	762	762	131
Questões + Facetas + Domínios + MQV	0.932	0.929	762	762	132

A - dados numéricos

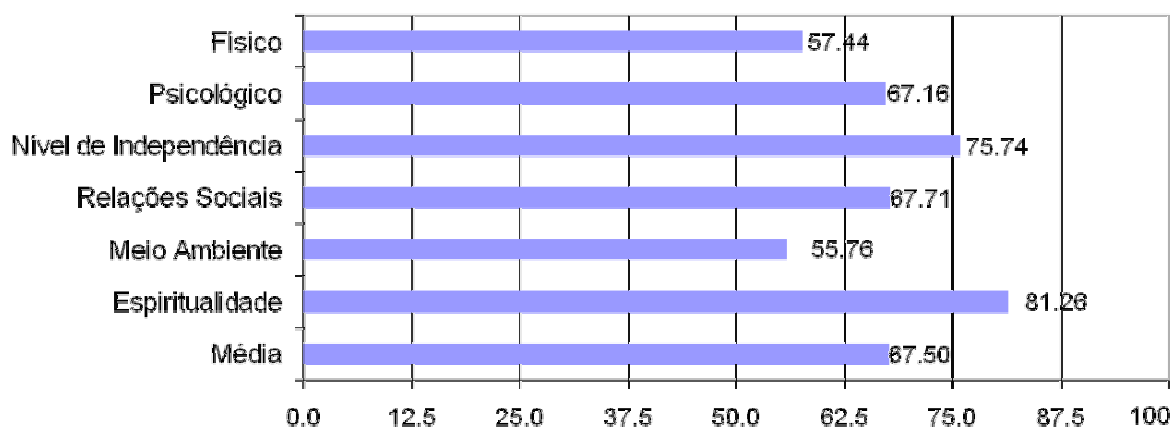
B - dados categóricos

Através dos coeficientes de *Cronbach's*, observa-se que após a categorização dos dados, mantêm-se satisfatórias as suas intercorrelações.

### 7.1.2 Conhecimento da Base de Dados

Para o conhecimento da base de dados, algumas SQLs foram implementadas, dando origem a um relatório no MS Access, a gráficos e tabelas no MS Excel, que serão apresentadas a seguir.

Após o desenvolvimento da sintaxe do *WHOQOL-100*, através dos softwares MS Access e MS Excel, chegou-se aos seguintes resultados dos indicadores de qualidade de vida.



**Figura 12 - Médias: Domínio, QGQV e Média de QV**

O Domínio Meio Ambiente apresentou os piores indicadores de QV, enquanto o Domínio Espiritualidade apresentou os melhores escores de QV.

Outras informações estatísticas são apresentadas no Anexo A e B.

Foi convencionada, uma média de QV, a partir da média aritmética entre os domínios e QGQV, que resultou em 67,5% da possibilidade máxima da escala de QV.

A Figura 13 apresenta a média dos aspectos de QV.

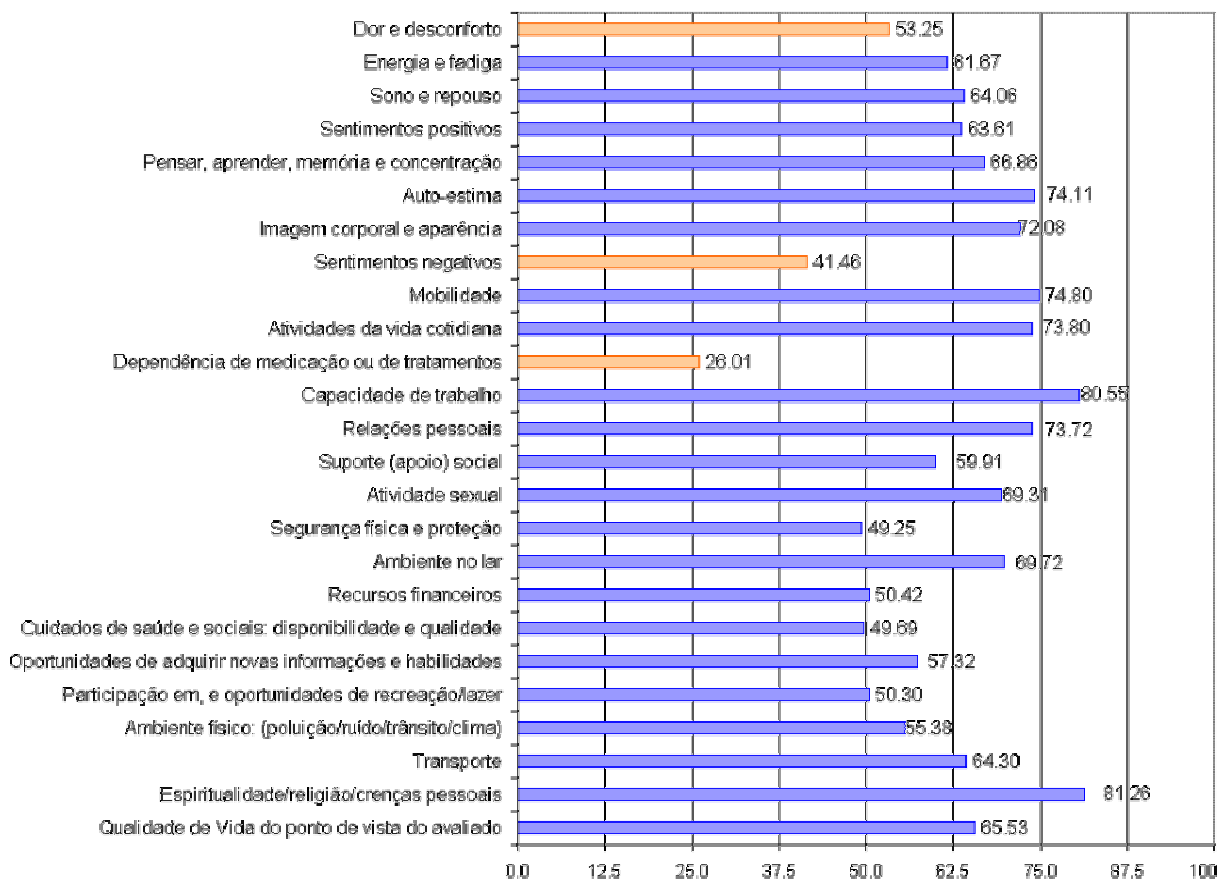


Figura 13 - Aspectos de Qualidade de Vida

Os aspectos Dor e Desconforto, Sentimento Negativos e Dependência por Medicação ou de Tratamento, são apresentados, na Figura 13, em cores diferenciadas, por possuírem características diferentes dos demais indicadores, pois são classificados como aspectos negativos e, portanto, quanto menor seus valores melhor o aspecto de QV.

O conhecimento da Base de Dados permitiu identificar os piores domínios e aspectos de qualidade de vida, que foram definidos com atributos metas para investigação através de RA, são eles: Domínio Meio Ambiente; Aspecto Dor e Desconforto e Cuidados de Saúde e Sociais (pior aspecto do pior domínio).

### **7.1.3 Enriquecimento da Base de Dados**

O enriquecimento da base de dados foi realizado com a inclusão de 32 atributos cadastrais dos agentes universitários, disponibilizados pela instituição objeto de pesquisa.

Além dos dados cadastrais, foram incluídos outros dados obtidos através de informações disponíveis em três atos públicos: Portaria R 161/06 da UEPG; Resolução Conjunta N 001/06 SETI/SEAP e Lei 15.050/06 do Governo do Estado do Paraná. Do primeiro, que descreve a transição dos servidores antes, durante e após a implantação de um Plano de Carreira, Cargos e Salários (PCCS), retirou-se os seguintes atributos: Classe, Série, Nível e Função. Do segundo, que descreve o perfil profissiográfico dos agentes universitários, retiraram-se as competências de cada cargo, com o aproveitamento de informações de trabalhos realizados anteriormente (SANTOS et. al, 2007). E finalizando, do último foram aproveitados apenas os valores de salários (Anexo III da Lei), que após cruzamento com as informações do primeiro foram categorizadas em níveis salariais.

#### **7.1.3.1 Limpeza dos Dados**

Esta etapa, apesar de apresentada nesta seqüência, iniciou-se durante o conhecimento da base de dados enriquecida, onde foram realizadas SQL específicas que resultaram na identificação de ruídos e registros duplicados, que em seguida foram eliminados ou substituídos por dados apropriados.

#### **7.1.3.2 Conhecimento Base de Dados Enriquecida**

Esta etapa foi realizada através da ferramenta de Pré-processamento do WEKA, que resultou nas informações da Tabela 24 (Anexo C).

### 7.1.3.3 Transformação dos Dados Enriquecidos

Como já foi mencionado anteriormente, para utilização da técnica de Regras de Associação, o algoritmo minerador Apriori, exige que os dados de entrada estejam categorizados. Então, todos os atributos do tipo numérico foram substituídos por duas novas tipagens, lógica e categórica, totalizando 128 atributos cadastrais.

### 7.1.4 Seleção e Transformação dos Dados

A seleção e transformação dos dados foram realizadas no ambiente WEKA.

Inicialmente configurou-se um *driver* ODBC para conectar o software WEKA ao MS Access. Através desta integração foi possível executar SQL no ambiente WEKA, que fez a transformação dos dados para o formato arff.

Foram selecionados através de SQL 160 atributos: 32 indicadores de QV e 128 dados cadastrais.

#### 7.1.4.1 Seleção de Atributos

Nesta sub-etapa foram selecionados os atributos relevantes, para evitar a geração na etapa subsequente de regras óbvias ou casuais (resultante de correlações espúrias).

Para a seleção desses atributos foi utilizado o filtro *AttributeSelectio*, composto pelo avaliador de atributos *CfsSubsetEval* (*Correlation-based Feature Subset Selection*), através do método de busca *Best First*.

Durante a análise dos dados, através das ferramentas de pré-processamento do WEKA, observou-se que muitos atributos do tipo lógico, como por exemplo, os atributos correspondentes às competências ou o atributo “Deficiência Física”, evidenciavam baixos ou nenhum ganho de informação, devido à baixa ou inexistente frequência de algumas classes. Então, os atributos lógicos, com as classes de menor frequência, inferiores ao valor mínimo do suporte adotado foram eliminados, pois naturalmente, estas classes, não apareceriam nas RA, enquanto as classes contrárias estariam presentes em grande parte das regras. Este procedimento resultou na diminuição de 17 atributos, o que diminuiria consideravelmente o número de RA com baixo grau de interesse (com conhecimento óbvio).

Os demais atributos com baixo ganho de informação evidente foram mantidos, ficando em função dos algoritmos de seleção a sua eliminação. Observou-se que a supressão de alguns desses atributos interferia na seleção de determinados grupos de atributos, devido ao fato de que, o algoritmo avaliador utilizado na seleção, busca preferencialmente, intercorrelações envolvendo classes de baixa frequência.

Esse algoritmo avalia o valor de um subconjunto de atributos considerando a capacidade preditiva individual de cada atributo, juntamente com o grau de redundância entre eles (HALL, 1998).

A seleção dos dados foi realizada de forma acíclica, buscando-se relacionamentos entre cada um dos indicadores de QV (Facetas, Domínios e MQV) e os dados cadastrais, gerando 32 grupos de dados.

Cada grupo de dados foi selecionado de forma supervisionada. Separando-se, inicialmente, apenas um indicador de QV e todos dados cadastrais (111 atributos) em um único arquivo. Cada indicador de qualidade de vida foi configurado como atributo meta (atributo de referência), durante a aplicação do filtro para seleção dos atributos relevantes em cada grupo de dados.

Depois de realizados os procedimentos acima mencionados, foram encontrados 14 atributos classificados como previsores dos aspectos e domínios de QV da amostra, apresentados na Tabela 13.



**Tabela 13 – Percentual dos Atributos Metas atendidos pelos Atributos Previsores**

Item	Atributo Previsor	% Atributos Metas	
		Domínios	Facetas
1	Lotação	100	100
2	Bairro	83.33	96
3	Últimas Férias	50	48
4	CSN	33.33	40
5	Alfabetizado	16.67	0
6	Competência: Agilidade	16.67	0
7	Competência: Atenção	16.67	0
8	Competência: Condicionamento Físico	16.67	0
9	Competência: Trabalho em Equipe	16.67	8
10	Função Anterior	16.67	48
11	Sanção Disciplinar	16.67	0
12	Competência: Asseio Pessoal	0	4
13	Disfunção Temporária	0	4
14	Partições em Atividades Administrativas	0	4

O atributo lotação foi selecionado em todos lotes de dados como previsor dos diferentes aspectos e domínios de QV.

O atributo bairro foi classificado previsor de 83.33% dos Domínios e 96% dos aspectos de qualidade de vida. Não é considerado atributo previsor apenas do Domínio Físico e do Aspecto: Participação em oportunidades de recreação/ lazer.

O salário representado pelo atributo “CSN” foi considerado atributo previsor de dois domínios: Psicológico e Meio Ambiente e, 40% dos aspectos de Qualidade de Vida.

O período do gozo das últimas férias foi classificado como atributo previsor dos domínios: “Psicológico”, “Relações Sociais” e “Meio Ambiente” e, de 48% dos aspectos de qualidade de vida.

A função ocupada antes do PCCS foi considerada um atributo previsor do domínio Meio Ambiente e 48% dos aspetos de qualidade de vida.

O domínio “Meio Ambiente” apresentou o maior número de atributos previsores (Figura 14).

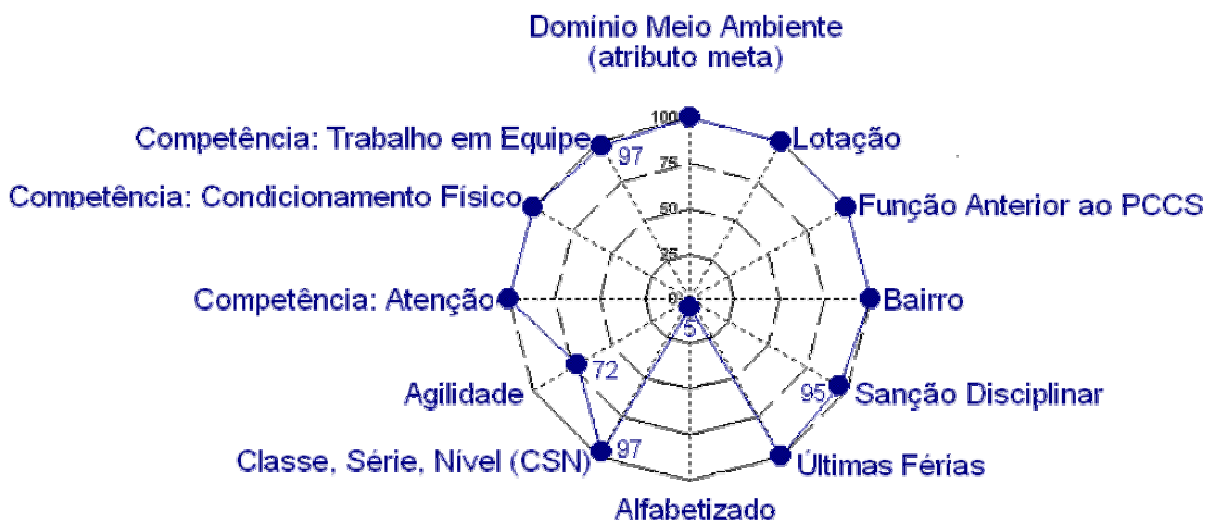


Figura 14 – Atributos inter-relacionados com o Domínio Meio Ambiente

A partir da Figura 14, observa-se que o Atributo “Alfabetizado”, apresenta a menor intercorrelação com o grupo (5%). Apesar da baixa intercorrelação, este atributo foi mantido, porque a sua classe menos freqüência possuía suporte maior que o suporte mínimo adotado.

A competência “Trabalho em equipe” foi classificada como um atributo predictor dos aspectos “Cuidados de saúde e sociais” e “Ambiente Físico”, ambos pertencentes ao domínio “Meio Ambiente”.

A competência “Asseio Pessoal” foi selecionada como um atributo predictor do aspecto “Atividades da Vida Cotidiana”.

A participação em atividades administrativas, como por exemplo, comissões de sindicância, comissões eleitorais, etc, foram classificados como atributos previsores do aspecto “Ambiente Físico”.

A Disfunção Temporária foi selecionada como um atributo predictor da Capacidade de Trabalho.

O processo de KDD é um processo contínuo e suas etapas podem ocorrer de forma acíclica.

A identificação de alguns atributos como previsores de determinados comportamentos, podem abrir caminho para outras pesquisas de investigação.

Neste trabalho, após a constatação de que o atributo “Sanção Disciplinar” e “Disfunção Temporária”, são atributos previsores de qualidade de Vida, mais especificamente do Domínio Meio Ambiente (Figura 15), decidiu-se verificar se existia dependência positiva destes atributos, com outros indicadores de qualidade de vida. Esta investigação resultou na seleção dos indicadores de qualidade de vida que poderiam estar contribuindo na ocorrência de sanções disciplinares ou disfunção temporárias.

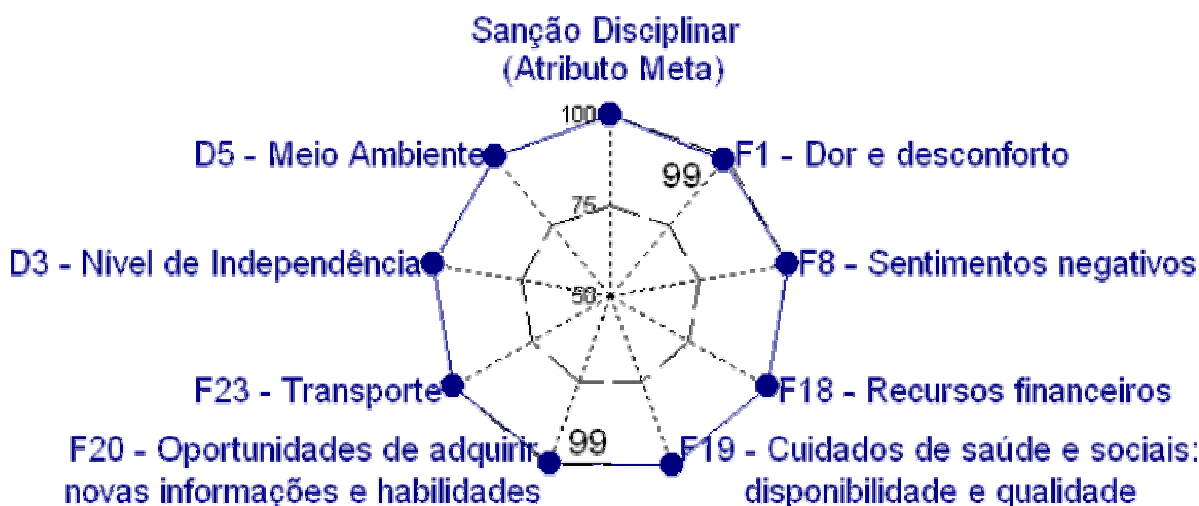


Figura 15 - Atributos inter-correlacionados com o Atributo Sanção Disciplinar

A Figura 15 apresenta oito atributos previsores do atributo “Sanção Disciplinar”.

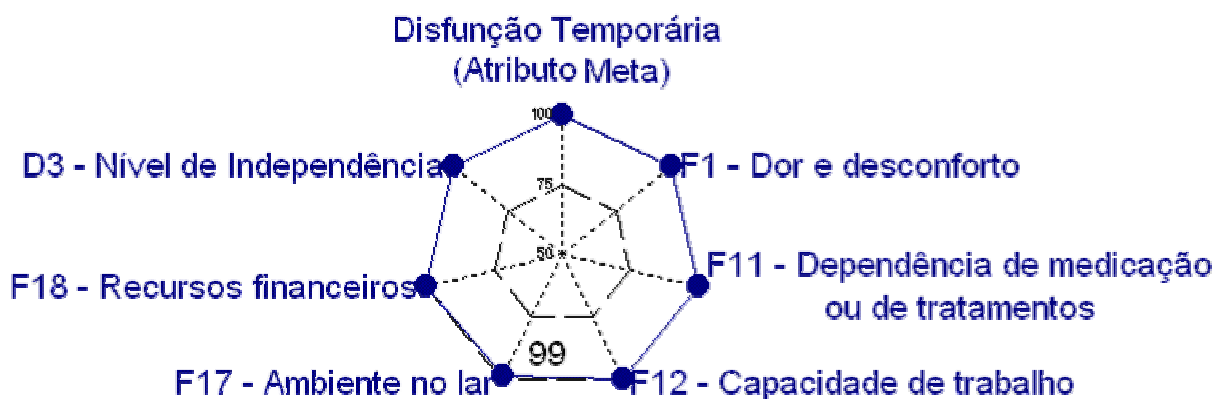


Figura 16 - Atributos inter-correlacionados com o Atributo Sanção Disciplinar

A Figura 10 apresenta seis indicadores de qualidade de vida que poderiam estar interferindo na ocorrência de disfunções temporárias.

Esta etapa de seleção de dados permitiu identificar os dados relevantes para a realização da mineração de dados.

Uma vez identificada a correlação, buscaram-se modelos que possam prever as variáveis dependentes em função de variáveis previsoras (BATANERO, 1999).

## **7.2 Data Mining**

Esta etapa também foi realizada de forma acíclica. Após cada seleção de atributos (etapa anterior), foi aplicado o algoritmo Apriori.

A configuração do algoritmo no Ambiente WEKA, permitiu definir apenas duas unidades de medidas de interesse, o suporte e mais uma (Confiabilidade ou *Lift* ou *Leverage* ou Convicção). Então, definiram-se as seguintes configurações mínimas: Suporte = 0.01 (equivalente a três registros) e Convicção = 1.1 (dependência positiva). Mantendo-se as demais configurações padrões.

Estes procedimentos resultaram na geração 327157 RA, distribuídas em 34 arquivos de saída.

## **7.3 Pós-processamento de Data Mining**

Esta etapa iniciou-se com a transformação dos resultados da mineração de dados e foi finalizada na sub-etapa de conhecimento adquirido.

### **7.3.1 Transformação**

Após a geração das Regras de Associação, os resultados foram salvos em uma base de dados com os resultados identificados por atributo meta.

### **7.3.2 Seleção das Regras de Associação**

No processo de seleção das RA utilizaram-se medidas de interesse subjetivas e objetivas.

Definiu-se como medida de interesse a subjetividade, a preferência por regras que possuíam no seu conseqüente, os piores indicadores de qualidade de vida (identificados na etapa de conhecimento da base de dados); os dados

cadastrais (identificados na etapa de seleção de dados) que possivelmente sofreriam interferência dos indicadores de QV; a média de Qualidade de Vida e; as Questões Gerais de Qualidade de Vida.

Segundo Silberschatz & Tuzhilin (1995), o grau de interesse de uma RA é avaliado pelo usuário, levando em consideração as suas expectativas e conhecimento prévio dos dados.

Como expectativa do usuário, procurou-se identificar entre os indicadores de QV, os atributos mais preditivos entre eles, pois a inferência direta sobre o atributo previsor poderia contribuir na alteração dos atributos dependentes. Então, buscou-se identificar os atributos previsores do maior número de atributos dependentes, considerando-se preferencialmente, aqueles que não possuíam dependência funcional. Esta identificação foi possível através do uso do filtro *AttributeSelectio*, que resultou na contabilidade apresentada na Tabela 25, onde foi possível identificar como atributos mais promissores para sofrer algum tipo de inferência, as QGQV e o Aspecto Segurança física e proteção (F16).

O atributo QGQV foi considerado atributo previsor de 11 indicadores de QV, e não possui dependência funcional de nenhum deles. Enquanto o Aspecto Segurança Física e Proteção foi considerado atributo previsor de 11 indicadores de qualidade de vida, dois quais, apenas dois possuíam dependência funcional (D5 e MGQV).

As medidas de interesse objetivas adotadas foram:

- dependência positiva do atributo meta;
- dependência positiva potencialmente implícita (*Convicção* < 5.0);
- confiança  $\geq 0.60$  (equivalente ao mínimo de 3/5 dos antecedentes);
- separadas por classe do atributo meta; e
- ordenadas pelos maiores valores de: Suporte, Confiança, Convicção, *Lift* e *Leverage*; e
- maior número de itens no antecedente da RA.

Adotou-se o valor mínimo de confiança de 60%, equivalente a 3/5 do antecedente, pois este valor melhor contemplaria RA de menor e maior valor de suporte, e porque, em regras com menor apoio, o menor valor de confiança encontrado seria de 67%, equivalente ao menor valor absoluto para essas configurações.

Após delimitada a estratégia de seleção, as RA foram selecionadas através de SQL específicas ou de processos de buscas simples (em arquivos menores) e contabilizadas na, onde se pode verificar a significativa contribuição das unidades de medidas na diminuição dos custos com o processo de pós-processamento.

A partir dos dados da Tabela 14, observou-se uma diminuição no número de RA avaliadas de 4,20% de forma subjetiva, e de 99% de forma objetiva.

As RA dos indicadores de qualidade de vida e atributos cadastrais escolhidos para investigação, foram organizados em tabelas e, ordenados por classes e por maior suporte (Tabela 15 à Tabela 21), para facilitar as suas interpretações na etapa a seguir.

Tabela 14 – Total de RA com os Resultantes

Atributos metas da Intercorrelação:  Indicadores de Qualidade de Vida e Dados Cadastrais		Número de Atributos Intercorrelacionados	Total de RA							Interessantes / Geradas
			Definidas no Algoritmo S >= 0.01  Cv >= 1.1	Processo de Busca C >= 0.60 L >= 1.01 ou Lv >= 0.01 Cv <= 5.0						
				Classes do Atributo Meta						
				Péssimo	Ruim	Médio	Bom	Excelente	Total	
MQV*		2	36	0	1	0	0	8	9	25%
Domínios de QV	D1	1	10							
	D2	4	338							
	D3	2	46							
	D4	3	163							
	D5*	11	198000	0	0	1336	222	0	1558	0.79%
	D6	2	26							
Aspectos de QV	F1*	3	136	0	14	11	2	0	27	19,85%
	F2	3	253							
	F3	5	1096							
	F4	3	269							
	F5	4	371							
	F6	3	161							
	F7	4	708							
	F8	3	214							
	F9	2	42							
	F10	4	896							
	F11	2	29							
	F12	6	4366							
	F13	4	787							
	F14	4	252							
	F15	2	36							
	F16*	4	755	0	18	49	6	0	73	9,67 %
	F17	3	110							
	F18	3	104							
	F19	5	2358							
	F20	3	281							
	F21	1	12							
	F22	6	481							
	F23	4	300							
	F24	2	26							
QGQV*		4	730	0	0	25	43	0	68	9.32 %
Sanção Disciplinar*		8	101166	Sim = 35					35	0.01 %
Disfunção Temporária*		6	12599	Sim = 11					11	0.09 %
Total		14	327157						1781	0.5 %

\* Atributos selecionados através de medidas de Interesse Subjetivas.

Tabela 15 – Regras de Associação com Conseqüente: Média de Qualidade de Vida (MQV)

MQV	Antecedente → MQV (Suporte (S); Confiança (C); Lift (L); Leverage (Lv); Convicção (Cv))
Excelente S= 0.62	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas
Bom S= 0.10	<p>1. Bairro = Oficinas → MQV = Bom (S= 0.056; C= 0.93; L= 1.5; Lv= 0.02; Cv = 2.85)</p> <p>2. Bairro = Cará-Cará → MQV = Bom (S= 0.023; C= 1; 1.61; Lv= 0.01; Cv= 2.28)</p> <p>3. Lotação = Colégio Agrícola Estadual "Augusto Ribas" E Bairro = Uvaranas → MQV = Bom (S= 0.023; C= 1; L= 1.61; Lv= 0.01; Cv= 2.28)</p> <p>5. Lotação = Seção de Paisagismo E Bairro = Uvaranas → MQV = Bom (S = 0.019;C =1; L = 1.61; Lv = 0.01; Cv = 1.9)</p> <p>7. Lotação = Colégio Agrícola Estadual "Augusto Ribas" → MQV = Bom (S = 0.053; C =0.85; L = 1.36; Lv = 0.01; Cv = 1.65)</p> <p>8. Lotação = Laboratório Universitário de Análises Clínicas - Central → MQV = Bom (S = 0.015;C =1; L = 1.61; Lv = 0.01; Cv = 1.52)</p> <p>9. Bairro = Núcleo Pimentel → MQV = Bom (S = 0.015;C =1; L = 1.61; Lv = 0.01; Cv = 1.52)</p> <p>14. Lotação = Departamento de Odontologia → MQV = Bom (S = 0.053;C =0.79; L = 1.27; Lv = 0.01;Cv = 1.33)</p>
Médio S= 0.21	17. Lotação= Seção de Obras → MQV = Médio (S = 0.019; C =0.6; L = 2.75; Lv = 0.01; Cv = 1.3)
Ruim S= 0.005	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas
Péssimo S= 0.004	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas
SR S= 0.053	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas



**Tabela 16 - Regras de Associação com Conseqüente: Questões Gerais de Qualidade de Vida (QGQV)**

QGQV	Antecedente → QGQV (Suporte (S); Confiança (C); Lift (L); Leverage (Lv); Convicção (Cv))
Excelente S = 0.154	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas
Bom S = 0.508	50. Função Anterior = Jardineiro Últimas Férias = 7/2006 → F25 = Bom (S = 0.030; C = 1; L = 1.97; Lv = 0.01; Cv = 3.94) 72. Lotação = Seção de Paisagismo Últimas Férias = 7/2006 → F25 = Bom (S = 0.026; C = 1; L = 1.97; Lv = 0.01; Cv = 3.45) 73. Lotação = Seção de Paisagismo Função Anterior = Jardineiro Últimas Férias = 7/2006 → F25 = Bom (S = 0.026; C = 1; L = 1.97; Lv = 0.01; Cv = 3.45) 154. Lotação = Seção de Paisagismo Bairro = Uvaranas → F25 = Bom 5 (S = 0.019; C = 1; L = 1.97; Lv = 0.01; Cv = 2.46) 155. Lotação = Seção de Paisagismo Função Anterior = Jardineiro Bairro = Uvaranas → F25 = Bom (S = 0.019; C = 1; L = 1.97; Lv = 0.01; Cv = 2.46) 156. Lotação = Seção de Paisagismo Bairro = Uvaranas Últimas Férias = 7/2006 → F25 = Bom (S = 0.019; C = 1; L = 1.97; Lv = 0.01; Cv = 2.46) 157. Função Anterior = Jardineiro Bairro = Uvaranas Últimas Férias = 7/2006 → F25 = Bom (S = 0.019; C = 1; L = 1.97; Lv = 0.01; Cv = 2.46) ... 510. Lotação = Colégio Agrícola Estadual "Augusto Ribas" → F25 = Bom (S = 0.049; C = 0.69; L = 1.36; Lv = 0.01; Cv = 1.28) 654. Lotação = Departamento de Odontologia → F25 = Bom (S = 0.053; C = 0.64; L = 1.27; Lv = 0.01; Cv = 1.15)
Médio S = 0.248	255. Função Anterior = Ajudante Geral Bairro = Uvaranas → F25 = Médio (S = 0.019; C = 0.8; L = 3.22; Lv = 0.01; Cv = 1.88) 256. Função Anterior = Ajudante Geral Bairro = Uvaranas Últimas Férias = 12/2006 → F25 = Médio (S = 0.019; C = 0.8; L = 3.22; Lv = 0.01; Cv = 1.88) 357. Lotação = Seção Elétrica → F25 = Médio (S = 0.015; C = 0.75; L = 3.02; Lv = 0.01; Cv = 1.5) 358. Bairro = Palmeirinha → F25 = Médio (S = 0.015; C = 0.75; L = 3.02; Lv = 0.01; Cv = 1.5). 358. Bairro = Palmeirinha → F25 = Médio (S = 0.015; C = 0.75; L = 3.02; Lv = 0.01; Cv = 1.5). 359. Lotação = Setor de Alimentação – CAAR E Função Anterior = Auxiliar de Cozinha → F25 = Médio (S = 0.015; C = 0.75; L = 3.02; Lv = 0.01; Cv = 1.5) 360. Lotação = Seção Elétrica E Últimas Férias = 12/2006 → F25 = Médio (S = 0.015; C = 0.75; L = 3.02; Lv = 0.01; Cv = 1.5) 361. Função Anterior = Ajudante Geral E Últimas Férias = 1/2007 → F25 = Médio (S = 0.015; C = 0.75; L = 3.02; Lv = 0.01; Cv = 1.5)
Ruim S = 0.079	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas
Péssimo S = 0.008	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas
SR S = 0.004	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas

Tabela 17 - Regras de Associação com Conseqüente: Domínio Meio Ambiente (D5)

D5	Antecedente → D5 (Suporte (S); Confiança (C); Lift (L); Leverage (Lv); Convicção (Cv))
Excelente S= 0.011	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas
Bom S= 0.538	<p>175006. Sanção Disciplinar = Não E Últimas Férias = 12/2006 E Alfabetizado = Não E Atenção = Não E Condicionamento Físico = Não → D5 = Bom (S = 0.117; C = 0.74; Lift = 2.1; Lv = 0.05; Cv = 2.23)</p> <p>174998. Sanção Disciplinar = Não E Agilidade = Não E Atenção = Não → D5 = Bom (S = 0.117; C = 0.74; L = 2.1; Lv = 0.05; Cv = 2.23)</p> <p>177829. Sanção Disciplinar = Não E Atenção = N → D5 = Bom (S = 0.128; C = 0.74; L = 2.08; Lv = 0.05; Cv = 2.2)</p> <p>187288. Sanção Disciplinar = Não E Ultimas Férias = 12/2006 E Atenção = Não 33 → D5 = Bom (S = 0.124; C = 0.73; L = 2.06; Lv = 0.05; Cv = 2.13)</p> <p>...</p>
Médio S= 0.353	<p>44495. Sanção Disciplinar = Sim e Ultimas Férias = 12/2006 E Agilidade = Sim E Atenção = Sim E Condicionamento Físico = Sim E Trabalho em Equipe = Sim → D5 = Médio (S = 0.038; C = 1; L = 1.86; Lv = 0.02; Cv = 4.62)</p> <p>...</p> <p>163629. Função Anterior = Ajudante Geral e Sanção Disciplinar = Sim E Ultimas Férias = 12/2006 E Atenção = Sim E Condicionamento Físico = Sim E Trabalho em Equipe = Sim → D5 = Médio (S = 0.019; C = 1; L = 1.86; Lv = 0.01; Cv = 2.31)</p> <p>...</p> <p>193932. Sanção Disciplinar = Sim E Condicionamento Físico = Sim → D5 = Médio (S = 0.068; C = 0.83; L = 1.85; Lv = 0.02; Cv = 2.08)</p> <p>...</p>
Ruim S= 0.086	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas
SR S= 0.011	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas

**Tabela 18 - Regras de Associação com Conseqüente: Aspecto Dor e Desconforto (F1)**

F1	Antecedente → F1 (Suporte (S); Confiança (C); Lift (L); Leverage (Lv); Convicção (Cv))
Excelente S = 0.053	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas
Bom S = 0.162	32. CSN = IA4 → F1= Bom (S = 0.015; C = 0.75; L = 4.64; Lv = 0.01; Cv = 1.68) 33. Lotação = CAIC E CSN = IB2 → F1= Bom (S = 0.015; C = 0.75; L = 4.64; Lv = 0.01; Cv = 1.68)
Médio S = 0.350	5. CSN = IB5 → F1= Médio (S = 0.034; C = 0.89; L = 2.54; Lv = 0.02; Cv = 2.93) 11. Lotação = CDR → F1= Médio (S = 0.023; C = 0.83; L = 2.38; Lv = 0.01; Cv = 1.95) 12. CSN = IA3 → F1= Médio (S = 0.011; C = 1; L = 2.86; Lv = 0.01; Cv = 1.95) 13. Lotação = CDR E CSN = IB5 → F1= Médio (S = 0.011; C = 1; L = 2.86; Lv = 0.01; Cv = 1.95) 14. Bairro = Centro E CSN = IIA4 → F1= Médio (S = 0.011; C = 1; L = 2.86; Lv = 0.01; Cv = 1.95) 19. Bairro = Centro → F1= Médio (S = 0.086; C = 0.7; L = 1.99; Lv = 0.03; Cv = 1.87) 34. Bairro = Ronda → F1= Médio (S = 0.019; C = 0.8; L = 2.29; Lv = 0.01; Cv = 1.63) 53. CSN = IIA4 → F1= Médio (S = 0.041 ; C = 0.64; L = 1.82; Lv = 0.01; Cv = 1.43) 57. Bairro = Oficinas → F1= Médio (S = 0.056; C = 0.6; L = 1.72; Lv = 0.01; Cv = 1.39) 73. Bairro = Vila Liane → F1= Médio (S = 0.015; C = 0.75; L = 2.15; Lv = 0.01; Cv = 1.3)
Ruim S = 0.387	1. Bairro = Uvaranas E CSN = IIIA2 → F1= Ruim (S = 0.030; C = 1; L = 2.58; Lv = 0.02; Cv = 4.9) 2. Bairro = Uvaranas E CSN = IIB5 → F1= Ruim (S = 0.023; C = 1; L = 2.58; Lv = 0.01; Cv = 3.68) 4. Lotação = Setor de Alimentação - CAAR E CSN = IIB5 → F1= Ruim (S = 0.019; C = 1; L = 2.58; Lv = 0.01; Cv = 3.06) 7. Lotação = Setor de Alimentação - CAAR → F1= Ruim (S = 0.045; C = 0.83; L = 2.15; Lv = 0.02; Cv = 2.45) 8. Lotação = Setor de Alimentação - CAAR Bairro = Uvaranas → F1= Ruim (S = 0.015; C = 1; L = 2.58; Lv = 0.01; Cv = 2.45) 20. CSN = IIIA2 → F1= Ruim (S = 0.079; C = 0.7; L = 1.81; Lv = 0.04; Cv = 1.84) 21. Bairro = Cara-Cará → F1= Ruim (S = 0.023; C = 0.83; L = 2.15; Lv = 0.01; Cv = 1.84) 22. CSN = IIIC11 → F1= Ruim (S = 0.023; C = 0.83; L = 2.15; Lv = 0.01; Cv = 1.84) 23. CSN = IIIC5 → F1= Ruim (S = 0.011; C = 1; L = 2.58; Lv = 0.01; Cv = 1.84) 24. Lotação = CAIC E CSN = IIIC11 → F1= Ruim (S = 0.011; C = 1; L = 2.58; Lv = 0.01; Cv = 1.84) 25. Lotação = Seção de Vigilância Patrimonial Bairro = Uvaranas CSN = IIIA2 → F1= Ruim (S = 0.011; C = 1; L = 2.58; Lv = 0.01; Cv = 1.84) 26. Lotação = Setor de Alimentação - CAAR Bairro = Uvaranas CSN = IIB5 → F1= Ruim (S = 0.011; C = 1; L = 2.58; Lv = 0.01; Cv = 1.84) 37. CSN = IIB5 → F1= Ruim (S = 0.049; C = 0.69; L = 1.79; Lv = 0.01; Cv = 1.59) 40. Lotação = Seção de Vigilância Patrimonial E CSN = IIIA2 → F1= Ruim (S = 0.019; C = 0.8; L = 2.07; Lv = 0.01; Cv = 1.53)
Péssimo S = 0.041	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas
SR S = 0.008	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas

**Tabela 19 - Regras de Associação com Conseqüente: Aspecto Segurança Física e Proteção (F16)**

F16	Antecedente → F16 (Suporte (S); Confiança (C); Lift (L); Leverage (Lv); Convicção (Cv))
Excelente S= 0.038	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas
Bom S= 0.553	325. Lotação = CDR E Função Anterior = Instrutor de Prática Desportiva → F16 = Bom (S= 0.015; C= 0.75; L = 4.87; Lv = 0.01; Cv = 1.69) 326. Lotação = CDR E Função Anterior = Instrutor de Prática Desportiva Ultimas Férias =12/2006 4 → F16 = Bom 3 (S= 0.015; C= 0.75; L= 4.87; Lv = 0.01; Cv = 1.69) 467. Função Anterior = Instrutor de Prática Desportiva → F16 = Bom (S = 0.019; C =0.6; L = 3.89; Lv = 0.01; Cv = 1.41) 468. Função Anterior = Instrutor de Prática Desportiva e Ultima Ferias = 12/2006 → F16 = Bom (S = 0.019; C =0.6; L = 3.89; Lv = 0.01; Cv = 1.41) 469. Função Anterior = Auxiliar de Laboratório e Bairro = Uvaranas → F16 = Bom (S = 0.019; C =0.6; L = 3.89; Lv = 0.01; Cv = 1.41) 469. Função Anterior = Auxiliar de Laboratório e Bairro = Uvaranas → F16 = Bom (S = 0.019; C =0.6; L = 3.89; Lv = 0.01; Cv = 1.41) 470. Função Anterior = Auxiliar de Laboratório e Bairro = Uvaranas E Últimas Férias = 12/2006 → F16 = Bom (S = 0.019; C =0.6; L = 3.89; Lv = 0.01; Cv = 1.41)
Médio S= 0.154	69. Função Anterior = Jardineiro E Ultimas Férias = 7/2006 → F16 = Médio (S = 0.030; C =1; L = 1.81; Lv = 0.01; Cv = 3.58) 94. Função Anterior = Marceneiro → F16 = Médio (S = 0.026; C =1; L = 1.81; Lv = 0.01; Cv = 3.13) 95. Lotação = Seção de Paisagismo E Ultimas Férias =7/2006 → F16 = Médio (S = 0.026; C =1; L = 1.81; Lv = 0.01; Cv = 3.13) 96. Lotação = Seção de Paisagismo E Função Anterior = Jardineiro Ultimas Férias =7/2006 7 → F16 = Médio 7 (S = 0.026; C =1; L = 1.81; Lv = 0.01; Cv = 3.13) 153. Lotação = Seção de Produtos de Madeira E Função Anterior = Marceneiro → F16 = Médio (S = 0.023; C =1; L = 1.81; Lv = 0.01; Cv = 2.68) 154. Função Anterior = Marceneiro E Ultimas Férias =12/2006 → F16 = Médio (S = 0.023; C =1; L = 1.81; Lv = 0.01; Cv = 2.68) ....  735. Bairro = Santa Paula Últimas Férias = 12/2006 → F16= Médio (S = 0.188;C =0.8; L = 1.45; Lv = 0; Cv = 1.12) 736. Lotação = Seção de Transportes Função Anterior = Motorista Últimas Férias =12/2006 5 → F16 = Médio 4 (S = x; C =0.8; L = 1.45; Lv = 0; Cv = 1.12)
Ruim S= 0.214	92. Bairro = Palmeirinha → F16 = Ruim (S= 0.015; C =1; L = 4.67; Lv = 0.01; Cv = 3.14) 152. Lotação = Setor de Alimentação - CAAR E Função Anterior = Cozinheiro 7 → F16 = Ruim 6 (S= 0.026; C =0.86; L = 4; Lv = 0.02; Cv = 2.75) 199. Função Anterior = Cozinheiro → F16 = Ruim (S = 0.034; C =0.78; L = 3.63; Lv = 0.02; Cv = 2.36) 200. Lotação = Setor de Alimentação - CAAR E Função Anterior = Cozinheiro E Últimas Férias = 12/2006 → F16 = Ruim (S = 0.023; C = 0.83; L= 3.89; Lv = 0.01; Cv = 2.36) 202. Bairro = Palmeirinha E Últimas Férias = 12/2006 → F16 = Ruim (S = 0.011; C =1; L = 4.67; Lv = 0.01; Cv = 2.36) 203. Lotação = Seção de Obras E Função Anterior = Ajudante Geral E Últimas Férias = 12/2006 → F16 = Ruim (S = 0.011; C = 1; L = 4.67; Lv = 0.01; Cv = 2.36) 229. Função Anterior = Cozinheiro E Últimas Férias = 12/2006 → F16 = Ruim (S = 0.030; C =0.75; L = 3.5; Lv = 0.02; Cv = 2.1) 244. Lotação = Seção de Obras → F16 = Ruim (S = 0.019; C =0.8; L = 3.73; Lv = 0.01; Cv = 1.96) 246. Lotação = Seção de Obras E Últimas Férias = 12/2006 → F16 = Ruim 4 (S = 0.019; C =0.8; L = 3.73; Lv = 0.01; Cv = 1.96) 247. Função Anterior =Cozinheiro E Bairro = Uvaranas → F16 = Ruim 4 (S = 0.019; C =0.8; L = 3.73; Lv = 0.01; Cv = 1.96) 248. Função Anterior =Cozinheiro E Bairro = Uvaranas Últimas Férias = 12/2006 → F16 = Ruim (S = 0.019; C =0.8; L = 3.73; Lv = 0.01; Cv = 1.96) 376. Lotação =Setor de Alimentação - CAAR E Bairro = Uvaranas → F16 = Ruim (S= 0.015; C = 0.75; L = 3.5; Lv = 0.01; Cv = 1.57) 377. Função Anterior = Ajudante Geral - Últimas Férias = 1/2007 → F16 = Ruim (S= 0.015; C = 0.75; L = 3.5; Lv = 0.01; Cv = 1.57) 378. Lotação =Setor Alimentação - CAAR E Função Anterior = Cozinheiro E Bairro = Uvaranas → F16 = Ruim (S= 0.015; C = 0.75; L = 3.5; Lv = 0.01; Cv= 1.57) 379. Lotação =Setor Alimentação - CAAR E Bairro = Uvaranas E Últimas Férias = 12/2006 → F16 = Ruim (S= 0.015; C = 0.75; L = 3.5; Lv = 0.01; Cv= 1.57) 380. Lotação =Setor Alimentação - CAAR E Função Anterior = Cozinheiro E Bairro = Uvaranas E Últimas Férias =12/2006 → F16=Ruim (S=0.015;C=0.75;L=3.5;Lv=0.01;Cv=1.57)
Péssimo S= 0.34	Não foram encontrados padrões com os valores mínimos das unidades das Medidas de Interesse Objetivas adotadas

**Tabela 20 - Regras de Associação com Conseqüente a Informação de Cadastro: Sanção Disciplinar**

<b>Sanção Disciplinar</b>	<b>Antecedente → Sanção Disciplinar (Suporte (S); Confiança (C); Lift (L); Leverage (Lv); Convicção (Cv))</b>
Sim S= 0.083	2085. F18 = Ruim E F19 = Péssimo → Sanção Disciplinar = Sim (S = 0.023; C = 0.83; L= 10.08; Lv = 0.02; C = 2.75) 2951. F8 = Péssimo E F20 = Médio → Sanção Disciplinar = Sim (S = 0.011; C = 1; L = 12.09; Lv = 0.01; Cv = 2.75) 2952. F8 = Péssimo E D3 = Médio → Sanção Disciplinar = Sim (S = 0.011; C = 1; L = 12.09; Lv = 0.01; Cv = 2.75) 2953. F8 = Péssimo E D3 = Médio E D5 = Médio → Sanção Disciplinar = Sim (S = 0.011; C = 1; L = 12.09; Lv = 0.01; Cv = 2.75)
Não S= 0.917	428. F20 = Bom E D5 = Bom → Sanção Disciplinar = Não (S = 0.207; C = 1; L = 1.09; Lv = 0.02; Cv = 4.55) 737. D5 = Bom → Sanção Disciplinar = Não (S = 0.0372; C = 0.99; L = 1.08; Lv = 0.03; Cv = 3.89)

**Tabela 21 - Regras de Associação com Conseqüente a Informação de Cadastro: Disfunção Definitiva**

<b>Disfunção Temporária</b>	<b>Antecedente → Disfunção Definitiva (Suporte (S); Confiança (C); Lift (L); Lev (Lv); Convicção (Cv))</b>
Sim S= 0.098	237. F11 = Médio E F12 = Bom E F18 = Ruim E D3 = Bom → Disfunção Temporária = Sim (S = 0.015;C =1;L =10.23; Lv = 0.01;Cv = 3.61) 669. F1 = Péssimo E F12 = Ruim → Disfunção Temporária = Sim (S = 0.011;C =1; L = 10.23; Lv = 0.01; Cv = 2.71) 669. F1 = Péssimo E F12 = Ruim → Disfunção Temporária = Sim (S = 0.011;C =1; L = 10.23; Lv = 0.01; Cv = 2.71) 670. F1 = Péssimo E F17 = Excelente → Disfunção Temporária = Sim (S = 0.011;C =1; L = 10.23; Lv = 0.01; Cv = 2.71) 671. F1 = Péssimo E D3 = Ruim → Disfunção Temporária = Sim (S = 0.011;C =1; L = 10.23; Lv = 0.01; Cv = 2.71) 672. F12 = Ruim E D3 = Ruim → Disfunção Temporária = Sim (S = 0.011;C =1; L = 10.23; Lv = 0.01; Cv = 2.71) 673. F1 = Péssimo E F12 = Ruim E D3 = Ruim → Disfunção Temporária = Sim (S = 0.011;C =1; L = 10.23; Lv = 0.01; Cv = 2.71) 1981. D3 = Ruim → Disfunção Temporária = Sim (S = 0.015;C =0.75; L = 7.67; Lv = 0.01; Cv = 1.8) 1980. F1 = Péssimo E F18 = Ruim → Disfunção Temporária = Sim (S = 0.023;C =0.67; L = 6.82; Lv = 0.01; Cv = 1.8)
Não S= 0.902	116. F11 = Excelente E F17 = Excelente → Disfunção Temporária = Não (S = 0.169;C =1; L = 1.11; Lv = 0.02; Cv = 4.4) 199. F17 = Excelente E D3 = Excelente → Disfunção Temporária = Não (S = 0.154;C =1; L = 1.11; Lv = 0.02; Cv = 4.01) 281. F12 = Excelente E F17 = Excelente E D3 = Excelente → Disfunção Temporária = Não (S =0.135;C =1; L = 1.11; Lv = 0.01;Cv = 3.52) 316. F11 = Excelente E F12 = Excelente E F17 = Excelente → Disfunção Temporária = Não (S = 0.131;C =1; L = 1.11; Lv = 0.01;Cv = 3.42)

### 7.3.3 Interpretação dos Resultados

Para auxiliar na interpretação das RA, o ambiente WEKA permite configurar o algoritmo para apresentar a frequência, na base de dados, de todos os elementos envolvidos nas RA (conseqüentes ou antecedentes). Algumas dessas informações, como o suporte de cada classe do elemento conseqüente da RA, foram incluídas nas tabelas de apresentadas anteriormente.

Nesta etapa as principais Regras de Associação foram transformadas em Regras de Produção, para facilitar a sua interpretação, conforme o modelo a seguir.

<p><b>SE</b></p> <p>Lotação = Setor de Alimentação CAAR</p> <p><b>ENTÃO</b></p> <p>D5 = Médio</p> <p>Medidas de Interesse Objetivas (S = 0.045; C = 0.92; Cv = 2.77)</p>
--

Figura 17 – Regra de Produção

Esta regra de associação indica que 4.5% dos servidores do campus de Uvaranas trabalham no Setor de Alimentação do Colégio Agrícola Augusto Ribas e que 92% destes servidores tem o domínio de Qualidade de Vida: Meio Ambiente classificado como Médio. E que a possibilidade de um servidor ter a sua qualidade de vida classificada como média é 2.77 vezes maior, caso esteja lotado no Setor de Alimentação do Colégio Augusto Ribas.

### 7.3.4 Aquisição de Conhecimento

Nesta etapa ocorre a transformação do conhecimento implícito em explícito.

Esta transformação ocorreu através da análise dos resultados dos indicadores de qualidade de vida e da interpretação das RA de associação.

A média da QV da amostra pode ser classificada, segundo a escala adotada, como boa, mas percebe-se que pode ser melhorada em até 32,5%. Essa melhoria pode ser alcançada inferindo inicialmente, nos piores indicadores de QV, entre eles, os mais preditivos. E, posteriormente, nos fatores passíveis de inferências,

apresentados nas RA como elementos implicadores das piores classes de dos indicadores de QV.

O Aspecto Segurança Física e Proteção apresentou o prior indicador de QV do Domínio Meio Ambiente.

Segundo o grupo *WHOQOL* (1998), esta faceta examina o indivíduo no sentimento de segurança e proteção de danos físicos. As ameaças para a sua segurança podem surgir a partir de qualquer fonte, como outras pessoas, ou a opressão política. Esta faceta é susceptível de atuar diretamente sobre a pessoa na sensação de liberdade. Deste modo, permite-se identificar, pessoas que têm oportunidade de viver sem constrangimentos e, pessoas que vivem em um estado ou proximidade de opressão ou insegurança.

Entre as facetas de qualidade de vida, observa-se que o Aspecto Dor e Desconforto apresentou o pior escore. Esta faceta explora as desagradáveis sensações físicas sentidas por uma pessoa e como estas sensações perturbam e interferem na sua vida (*WHOQOL*, 1998).

A partir da interpretação das RA selecionadas, sabe-se que existe um padrão de comportamento entre os indivíduos com os piores indicadores, que serão descritos a seguir.

Os agentes universitários que estão lotados na Seção de Obras merecem atenção especial, pois obtiveram a menor classificação da MGQV com padrão de comportamento.

Em relação as QGQV, que examinam as maneiras pelas quais uma pessoa avalia a sua qualidade de vida global, saúde e bem-estar, a menor classificação (QGQV = Média) atribuída com padrão de comportamento, foi dada aos servidores: ocupantes da função de Ajudante Geral, que residem no Bairro de Uvaranas; lotados na Seção Elétrica; lotados no Setor de Alimentação do CAAR que ocupavam a Função de auxiliar de Cozinha e os residentes no Bairro da Palmeirinha.

Os servidores que sofreram sanção disciplinar e/ou ocupavam a função de Ajudante Geral e/ou que a função exige determinadas competências (Agilidade e/ou Atenção e/ou Condicionamento Físico e/ou Trabalho), possui a menor classificação do Domínio Meio Ambiente, com um padrão de comportamento.

Apresentaram o Aspecto Dor e Desconforto classificado como ruim, os servidores: lotados no CAAR; que residem no Bairro Cara-Cará; que recebem salário máximo de 3 salários mínimos e/ou estão lotados na Seção de Alimentação do CAAR ou Seção de Vigilância ou no CAIC e/ou residem no Bairro de Uvaranas.

Em relação ao Aspecto Segurança Física e Proteção, obtiveram a pior classificação (F16 = ruim), os servidores: lotados na Seção Elétrica ou Seção de Obras ou Seção de Alimentação do CAAR e/ou que residem nos Bairros: Cará-Cará, Palmeirinha e Uvaranas.

Observou-se, que alguns indicadores de QV poderiam estar contribuindo para a ocorrência de sanção disciplinar ou de disfunção temporária de servidores.

O Aspecto Recursos Financeiros e Cuidados de Saúde e Sociais, classificados respectivamente com ruim e péssimos; ou o aspecto Sentimentos Negativos e o Domínio Nível de Independência e/ou Domínio Meio Ambiente, classificados, respectivamente como péssimo e médios, poderiam estar contribuindo na ocorrência de uma sanção disciplinar.

Observou-se que a combinação de pelo menos dois dos indicadores de QV a seguir, poderiam estar contribuindo com a ocorrência de um desvio de função temporária do servidor:

- F11 - Dependência de medicação ou de tratamentos – médio;
- D3 - Nível de independência – ruim;
- F1 - Dor e desconforto – ruim;
- F18 - Recursos financeiros – ruim e;
- F12 - Capacidade de trabalho – péssimo;

Finalizando, observa-se que merecem atenção especial dos responsáveis pela gestão de QV da instituição, principalmente, os servidores ocupantes das funções de Ajudante Geral ou Cozinheiro; lotados na Seção Elétrica ou Seção de Obras ou Seção de Alimentação do CAAR; que residem nos Bairros Cará-Cará ou Palmeirinha.



## 8 CONCLUSÕES E RECOMENDAÇÕES

As técnicas de *Data Mining*, mais especificamente, as técnicas de Regras de Associação, podem ser utilizadas para complementar a análise dos resultados de um instrumento de pesquisa estruturada para *surveys*, principalmente quando se dispõe de muitos dados adicionais e se deseja descobrir conhecimento implícito.

O planejamento e a utilização combinada de medidas de interesse subjetivas e objetivas são de extrema importância para a economia de custos operacionais das tarefas de KDD de descrição.

O uso de medidas de interesse, neste trabalho, reduziu o número de RA a serem investigadas em 4,20% e 99.5%, através da respectiva utilização de medidas de interesse subjetivas e objetivas. Isto evidencia a importância do processo de filtragem, principalmente, através do uso das medidas de interesse objetivas; e do processo de planejamento, através das medidas de interesse subjetivas.

Os seguintes padrões de comportamento, potencialmente implícitos, dos agentes universitários da instituição objeto da avaliação de qualidade de vida foram evidenciados:

- a lotação do trabalhador é determinante na sua qualidade de vida, em seus diferentes aspectos domínios;
- o bairro residencial interfere na maioria dos domínios e aspectos de QV, exceto no Domínio Físico e Aspecto Participação em oportunidades de recreação/ lazer.
- o salário e o período de férias interferem em alguns aspectos de qualidade de vida;
- a exigência ou não de algumas competências da função, também pode contribuir na classificação da sua QV;
- as funções mais específicas, ocupadas na instituição, também possuem padrões de comportamento em relação à QV, principalmente no Domínio e Aspectos de QV do Meio Ambiente.

- a ocorrência de sanção disciplinar contribui para a diminuição na qualidade de vida no Domínio Meio Ambiente; e tem maiores chances de ocorrer, quando alguns aspectos de Qualidade de Vida encontram-se classificados, como péssimo ou ruim; e
- uma disfunção temporária tem maiores possibilidades de ocorrer, quando existe a combinação de baixos indicadores de qualidade de vida.

Através desses resultados comprova-se a importância do uso de técnicas de Mineração de Dados, no monitoramento de fatores que possam estar interferindo nos indicadores de Qualidade de Vida, pois com a utilização apenas dos métodos estatísticos propostos no instrumento *WHOQOL*, seria possível apenas a identificação dos índices de QV da amostra ou de determinados segmentos, enquanto através do uso das técnicas de *Data Mining*, foi permitido identificar entre os dados disponíveis, fatores que contribuem na predição dos indicadores de qualidade de vida.

O produto de um processo de KDD, quando aplicado em um instrumento de avaliação de qualidade de vida, poderá auxiliar os gestores de qualidade de vida da instituição em suas tomadas de decisões. Assim como, realização de treinamentos, acompanhamento psicológico, assistência social e outras tarefas, com focos mais precisos podem ser aplicadas.

A utilização das técnicas de *Data Mining* poderá contribuir na elaboração de instrumentos mais específicos para determinados grupos ou culturas. E, também, podem ser utilizadas para definir questões de re-testes para avaliar a eficiência de inferências pontuais.

## 8.1 SUGESTÕES PARA TRABALHOS FUTUROS

Sugere-se a realização de estudos para identificação das relações de independência, dependência positiva e dependência negativa entre indicadores de qualidade de vida.

Aplicação de taxonomia para generalizar regras por domínios e aspectos de qualidade de vida.

Desenvolvimento de instrumentos mais específicos para determinados grupos ou culturas, utilizando técnicas de *Data Mining*

Desenvolvimento de modelos de classificação de Qualidade de Vida, baseados em árvore de decisão ou redes neurais, para facilitar a identificação dos fatores que mais interferem em determinados grupos de trabalhadores.

## REFERÊNCIAS

ADRIAANS, P; ZABTINGE, D. **Data Mining**. Harlow: Addison-Wesley, 1996. 176p.

AGRAWAL, R.; EVFIMIEVSKI; A.; KIERNAN, J.; VELU, R. Auditing disclosure by relevance ranking. In: ACM SIGMOD INTERNATIONAL CONFERENCE ON MANAGEMENT OF DATA, 2007, Beijing, China. **Proceedings ...** Ney York: ACM, 2007. p.79-90.

AGRAWAL, R.; SRIKANT, R. Fast algorithms for mining association rules. In: CONFERENCE ON VERY LARGE DATABASES, 20., 1994, Santiago de Chile, Chile. **Proceedings ...** San Francisco, CA: Morgan Kaufmann, 1994. p.487-499.

AMARAL, F. C. N. **Data Mining: Técnicas e Aplicações para o Marketing Direto**. 1. ed. São Paulo: Berkeley, 2001, 110p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6022 Informação e Documentação** - Artigo em publicação periódica científica impressa. Rio de Janeiro: ABNT, 2003.

BATANERO, C. Análise Exploratória de dados nos cursos de segundo grau. In: CONFERÊNCIA INTERNACIONAL "EXPERIÊNCIAS E EXPECTATIVAS DO ENSINO DE ESTATÍSTICA - DESAFIOS PARA O SÉCULO XXI", 1999, Florianópolis. **Atas ....** Florianópolis: UFSC, 1999.

BRIN, S.; MOTWANI, R.; ULMAN, J. D.; TSUR, S. Dynamic Itemset counting and implication rules for market basket data. In: ACM SIGMOD INTERNATIONAL CONFERENCE ON MANAGEMENT OF DATA, 1997, Tucson, Arizona, USA. **Proceedings ...** New York: AMC, 1997, p.255-264.

BULLINGER, M.; ANDERSON, R.; CELLA, D. Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. **Quality of Life Research**. v.2, n.6, p.451-459, dec. 1993.

CARVALHO, D. R. **Data Mining Através de Indução de Regras e Algoritmos Genéticos**, 1999. 142p. Dissertação (Mestrado em Informática Aplicada) - Pontifícia Universidade Católica do Paraná, Curitiba, 1999.

CARVALHO, D. R.; BASSI, P. Data Mining para avaliação do aprendizado dos cursos de computação. In: INTERNATIONAL CONFERENCE ON ENGINEERING AND COMPUTER EDUCATION, 3., 2003, São Vicente/Santos. **Proceedings ...** São Paulo, 2003. p.16-19.

CARVALHO, D. R.; FREITAS A. A.; EBECKEN, N. F. F. A critical review of rule surprisingness measures. In: INTERNATIONAL CONFERENCE OF DATA MINING, 5., 2003, Rio de Janeiro. **Proceedings ....** Ashurst: WIT Press, 2003. p.545-556.

CARVALHO, L. A. V. **Data Mining: a Mineração de Dados no Marketing, Medicina, Economia, Engenharia e Administração**. São Paulo: Erica, 2001. 237p.

CASTRO, A. E. **Las tablas de contingencia y su enseñanza. Qué podemos aprender de las investigaciones realizadas?** Uno: Revista de Didáctica de las Matemáticas, Barcelona, v.3, p.89-100, 1995.

CRESPO, A. A. **Estatística fácil**. 18.ed. São Paulo: Saraiva, 2002.

FAYYAD, U. M. Diving into databases: SQL is helpless in the face of massive, accumulating data stores. **Database programming and design**. San Mateo, v.11, n.3, p.24-31, mar. 1998.

FAYYAD, U. M.; UTHURUSAMY, R. INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 1., 1995. Montreal, Canadá: AAAI Press, 1995, 348p.

FAYYAD, U. M.; UTHURUSAMY, R. WORKSHOP, TECHNICAL REPORT KNOWLEDGE DISCOVERY IN DATABASES, 1994. Seattle, WA: AAAI Press, 1994, 473p.

FAYYAD, U. et al. **Advances in Knowledge Discovery and Data Mining**. Menlo Park, CA: AAAI Press / The MIT Press. 1996, 625p.

FIGUEIRA, R. **Mineração de dados e bancos de dados orientados a objetos**. 1998, 96f. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1998.

FLECK, M. P. A. A Application of the Portuguese version of the instrument for the assessment of the quality of life of the World Health Organization (WHOQOL-100), **Revista Saúde Pública**. São Paulo, v.33, n.2, p.198-205, abr. 1999.

FLECK, M. P. A. et al. Aplicação da versão em português do instrumento WHOQOL-bref. **Revista Saúde Pública**. São Paulo, v.34, n.2, p.178-83, abr. 2000.

FLECK, M. P. A.; CHACHAMOVICH, E., TRENTINI, C. Development and validation of the Portuguese version of the WHOQOL-OLD module. **Revista Saúde Pública**, São Paulo, v.40, n.5, p.785-791, oct. 2006.

FOX-RUSHBY, J.; PARKER, M. Culture and the measurement of health-related quality of life, **Journal of European Applied Psychology**, Paris, v.45, n.4, p.257-263, 1995.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 5. ed. São Paulo: Atlas, 1999. 208p.

GONÇALVES, E. C. Regras de associação e suas medidas de interesse objetivas e subjetivas. **Journal of Computer Science**. Lavras: UFLA, v.4, n.1, p.27-36, mar. 2005.

HALL, M. A. **Correlation-based feature subset selection for machine learning**. Department of Computer Science, University of Waikato, Hamilton, New Zealand, 1998.

HSU, W. et al. Exploration mining in diabetic patients databases: findings and conclusions. In: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, Conference on Knowledge Discovery in Data, 6., 2000, Boston, MA, USA. **Proceedings ...** New York: ACM, 2000. p.430-436.

HUSSAIN, F.; LIU, H.; LU, H. Exception rule mining with a relative interestingness measure. In: PACIFIC-ASIA CONFERENCE, 4, 2000. Kyoto. **Proceedings ...** Kyoto: Springer, 2000. p.86-96.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Centro de Documentação e Disseminação de Informação. **Normas para apresentação tabular**. 3.ed. Rio de Janeiro: IBGE, 1993. 62p.

KLUTHCOVSKY, A. C. G. C. **Qualidade de vida dos agentes comunitários de saúde de um município do interior do Paraná**, 2005. 118f. Dissertação (Mestrado em Enfermagem) - Universidade de São Paulo, Ribeirão Preto, 2005.

KULKARNI, G.; FATHI, Y. A. Very large scale neighborhood search algorithm for the q-mode problem. London, **IIE Transactions**, v. 39, n.10; p.971, oct. 2007.

LAKATOS, E. M.; MARCONI, M. A. **Fundamentos de metodologia científica**. 4. ed. São Paulo: Atlas, 2001, 288 p.

LEMOS, P. E. **Análise de crédito bancário com o uso de Data Mining: redes neurais e árvores de decisão**. 2003, 147p. Dissertação (Mestrado em Ciências) - Universidade Federal do Paraná, Curitiba, 2003.

LENCA, P. et al. On selecting interestingness measures for association rules: user oriented description and multiple criteria decision aid. **European Journal of Operational Research**. Amsterdam, v.184, n.2; p.610-626, jan. 2008.

LOPES, C. H. P. **Classificação de registros em banco de dados por evolução de regras de associação: utilizando algoritmos genéticos**, 1999, 136p. Dissertação (Mestre em Engenharia Elétrica e Sistemas de Computação) -. Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 1999.

LUNA, S. V. **Planejamento de pesquisa: uma introdução**. São Paulo: EDUC, 1997.

MASTHOFF, E. D. et. al. Validation of the WHO Quality of Life assessment instrument (WHOQOL-100) in a population of Dutch adult psychiatric outpatients. **European Psychiatry**. Issy les Moulineaux, France, v.20, n7, p.465-473, nov. 2005.

MCGARRY, K. A survey of interestingness measures for knowledge discovery. **The Knowledge Engineering Review**, Cambridge University, New York, USA, v.20, n.1, p.39–61, mar. 2005.

MELDOLESI, G. N. et al. Factors associated with generic and disease-specific quality of life in temporal lobe epilepsy. **Epilepsy Research**. Amsterdam, v.69, n.2, p.135-146, may. 2006.

MELDOLESI, G. N. et al. Changes in depression, anxiety, anger, and personality after resective surgery for drug-resistant temporal lobe epilepsy: a 2-year follow-up study. **Epilepsy Research**. Amsterdam, v.77, n.1, p.22-30, oct. 2007.

MICHALSKI, R.; KAUFMAN, K. **Data Mining and Knowledge Discovery**: a review of Issues and multistrategy approach. In: MICHALSKI, R. S.; BRATKO, I. KUBAT, M. Machine Learning and Data Mining Methods and Applications, England: John Wiley, 1998.

MORAES, J. F. D.; SOUZA, V. B. A. Fatores associados ao envelhecimento bem-sucedido de idosos socialmente ativos da região metropolitana de Porto Alegre. **Revista Brasileira de Psiquiatria**. São Paulo, v.27, n.4, p.302-308, dez. 2005.

MORENO, A. B. **Avaliação da qualidade de vida em pacientes laringectomizados**: Uma Revisão Sistemática, 1999, 111p. Dissertação (Mestrado em Saúde Coletiva) - Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 1999.

POLYANA, A. **Descoberta de conhecimento via provador de teoremas**, 1997. 87p. Dissertação (Mestrado em Sistemas de Computação) - Instituto Militar de Engenharia, Rio de Janeiro, 1997.

REFFAT, R.; GERO, J.; PENG, W. Improving the management of building life cycle: A Data Mining approach. In: CRC CONSTRUCTION INNOVATION CONFERENCE, 2004, Brisbane, Austrália. **Proceedings ...** Brisbane: CRC, 2004.



RODRIGUES, A. M. **Técnicas de Data Mining Classificadas do Ponto de Vista do Usuário**. Rio de Janeiro, 2000. 116f. Dissertação (Mestrado em Engenharia de Produção) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2000.

RUGISKI, M.; PILATTI, L. A.; SCANDELARI, L. WHOQOL-100 e sua utilização: uma pesquisa na Internet. In: PILATTI, A. L.; KOVALESKI, J. L.; OLIVERIA, S. L. **Temas em Engenharia de Produção I**. Jundiaí: Fontoura, 2005, cap.3, p.21-30.

SILBERSCHATZ, A.; TUZHILIN, A. On subjective measures of interestingness in knowledge discovery. In: INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 1., 1995, Montreal, Canadá. **Proceedings ...** Chicago, USA: AAAI Press, 1995, p.275-281.

SILVA, E. L. MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. 3. ed. Florianópolis: Laboratório de Ensino à Distância da UFSC, 2001.

SUZUKI, E. Autonomous discovery of reliable exception rules. In: INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 3., 1997, Newport Beach, California, USA. **Proceedings ...** Chicago, USA: AAAI Press, 1997, p.259-263.

UNIVERSIDADE ESTADUAL DE PONTA GROSSA. **Manual de normalização bibliográfica para trabalhos científicos**. Ponta Grossa: UEPG, 2005. 131p.

UNIVERSIDADE ESTADUAL DE PONTA GROSSA. Portaria R. nº 161 - UEPG de 1º de agosto de 2006. Procede a correlação de funções e o enquadramento em três etapas, na carreira técnica universitária para os servidores ocupantes do cargo de Agente Universitário da UEPG, nos termos do anexo que passa a integrar este ato legal. Relator: Paulo Roberto Godoy. **Atos Oficiais da Universidade Estadual de Ponta Grossa. Ponta Grossa**, 1º ago. 2006. p.1-30.

VASCONCELOS, B. S. **Mineração de regras de classificação com sistemas de banco de dados objeto - relacional**. Estudo de caso: regras de classificação de litofácies de poços de petróleo, 2002, 128f. Dissertação (Mestrado em Informática) - Universidade Federal de Campina Grande, Campina Grande, 2002.

VASCONCELOS, E. M. **Complexidade e pesquisa interdisciplinar**: epistemologia e metodologia operativa. Petrópolis: Vozes, 2002.

VIANA, R. Mineração de dados: teoria e prática. **Revista SQL Magazine**, Rio de Janeiro: Neofício, n.10, 2004.

WEKA. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka>>. Acesso em: 10 fev. 2007.

WITTEN, I. H.; FRANK, E. **Data Mining**: Practical machine learning tools and techniques, 2.ed., San Francisco, CA: Morgan Kaufmann, 2005. 525p.

WHOQOL GROUP. Study protocol for the World Health Organization project to develop a quality of life assessment instrument (WHOQOL). **Quality of Life Research**. v.2, n.2, p.153-159, apr. 1993.

WHOQOL GROUP. The development of the World Health Organization quality of life assessment instrument (the WHOQOL). In: Orley, J.; Kuyken, W. **Quality of life assessment: international perspectives**. Heidelberg: Springer Verlag, 1994, p.41-60.

WHOQOL GROUP. Development of the WHOQOL: Rationale and current status. **International Journal of Mental Health**, v.23, n.3, p.24-56, 1994.

WHOQOL GROUP. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. **Social Science and Medicine**. Amsterdam, v.41, n.10, p.1403-1409, nov. 1995.

WHOQOL GROUP. The World Health Organization Quality of Life assessment: development and general psychometric properties. **Social Science and Medicine**. Amsterdam, v.46, n.12, p.1569-1585, jun. 1998.

YIN, R. K. **Estudo de caso**: planejamento e métodos. 2.ed. Porto Alegre: Bookman, 2001, 205 p.

YAO, G.; WANG, J.D.; CHUNG, C. W. Cultural adaptation of the WHOQOL questionnaire for Taiwan. **Journal of the Formosan Medical Association**. Amsterdam, v.106, n.7, p.592-597, jul. 2007.

ZIMPEL, R. R.; FLECK, M. P. Quality of life in HIV-positive Brazilians: application and validation of the WHOQOL-HIV Brazilian version, **Aids Care**. London, v.19, n.7, p.923-930, aug. 2007.

## ANEXO A – ESTATÍSTICA DESCRITIVA DOS INDICADORES DE QUALIDADE DE VIDA – DOMÍNIOS, QGQV E MQV

**Tabela 22 – Resultados Estatísticos dos Domínios de Qualidade de Vida e a Média de QV**

Estatística	Físico	Psicológico	Nível de Independência	Relações Sociais	Meio Ambiente	Espiritualidade Crenças pessoais	Média de Qualidade de Vida
Média	57.44	67.16	75.74	67.71	55.76	81.26	67.5
Erro padrão	0.968	0.771	0.907	0.853	0.731	0.925	0.642
Mediana	58.31	67.5	76.56	68.34	55.43	81.25	67.37
Modo	64.56	67.5	85.93	68.75	52.31	75	75.18
Desv. Pad.	15.73	12.6	14.79	13.7	11.87	15.06	10.22
Var. Amos.	247.4	158.7	218.6	187.6	141	226.7	104.4
Curtose	-0.49	0.018	0.27	0.579	0.008	0.466	-0.04
Assimetria	0.051	-0.15	-0.59	-0.24	0.165	-0.64	0.034
Intervalo	77.06	71.25	75	87.5	75	81.25	62.25
Mínimo	14.56	28.75	25	12.5	25	18.75	34.93
Máximo	91.62	100	100	100	100	100	97.18
Contagem	264	266	266	258	264	265	253

## ANEXO B – ESTATÍSTICA DESCRITIVA DOS INDICADORES DE QUALIDADE DE VIDA - ASPECTOS

Tabela 23 – Resultados Estatísticos dos Aspectos de Qualidade de Vida

Estadística	Dor e Desconforto	Energia e fadiga	Sono e repouso	Sentimentos Positivos	Pensar, aprender, memória e concentração	Auto-estima	Imagem corporal e aparência	Sentimentos negativos	Mobilidade	Atividades da vida cotidiana	Dependência de medicação ou de tratamentos	Capacidade de trabalho	Relações pessoais	Suporte (apoio) social	Atividade sexual	Segurança física e proteção	Ambiente no lar	Recursos financeiros	Cuidados de saúde e sociais: disponibilidade e qualidade	Oportunidades de adquirir novas informações e habilidades	Participação em, e oportunidades de recreação/ lazer	Ambiente físico: poluição/ ruído/ trânsito/ clima	Transporte	Espiritualidade/ religião/ crenças pessoais	Questões Gerais de Qualidade de Vida
Média	53.25	61.67	64.06	63.61	66.86	74.11	72.08	41.46	74.8	73.8	26.01	80.55	73.72	59.91	69.31	49.25	69.72	50.42	49.69	57.32	50.3	55.38	64.3	81.26	65.53
Erro padrão	1.08	1.142	1.474	1.097	0.913	0.865	1.053	1.329	1.235	1.053	1.537	0.949	0.901	1.093	1.195	0.902	0.996	1.17	1.092	1.097	1.163	0.889	1.444	0.925	0.983
Mediana	56.25	62.5	68.75	62.5	68.75	75	75	43.75	75	75	25	75	75	62.5	75	50	75	50	50	56.25	50	56.25	68.75	81.25	68.75
Modo	62.5	56.25	75	62.5	75	75	75	50	100	81.25	0	75	75	62.5	81.25	50	75	50	56.25	50	43.75	62.5	50	75	75
Desv. Pad.	17.59	18.55	23.9	17.92	14.86	14.11	17.17	21.67	20.03	17.15	25.12	15.48	14.72	17.82	19.2	14.68	16.22	19.08	17.81	17.82	19.01	14.48	23.6	15.06	16.03
Var. Amos.	309.3	344.1	571.3	321.2	220.7	199.1	294.7	469.8	401.3	294.1	630.9	239.5	216.8	317.6	368.5	215.5	263.1	364.2	317.2	317.6	361.4	209.6	556.8	226.7	257
Curtose	0.294	-0.59	-0.96	0.31	0.399	0.237	-0.3	-0.71	-0.43	-0.49	-0.33	0.14	0.698	-0.23	0.891	0.671	0.523	-0.09	0.149	-0.04	-0.52	0.425	-0.61	0.466	-0.01
Assimetria	-0.51	-0.02	-0.28	-0.45	-0.51	-0.36	-0.38	0	-0.42	-0.36	0.802	-0.53	-0.64	0.053	-0.8	0.107	-0.57	0.201	-0.37	-0.24	0.088	-0.26	-0.42	-0.64	-0.4
Intervalo	100	81.25	93.75	100	81.25	68.75	87.5	100	93.75	75	100	75	81.25	87.5	100	93.75	93.75	100	100	100	93.75	100	100	81.25	81.25
Mínimo	0	18.75	6.25	0	18.75	31.25	12.5	0	6.25	25	0	25	18.75	12.5	0	6.25	6.25	0	0	0	6.25	0	0	18.75	18.75
Máximo	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Contagem	265	264	263	266	265	266	266	266	263	265	266	266	266	266	258	265	265	266	266	264	266	265	266	265	266

## ANEXO C – DADOS CADASTRAIS UTILIZADOS NA MINERAÇÃO DE DADOS

Tabela 24 - Relação de Dados Cadastrais da Amostra

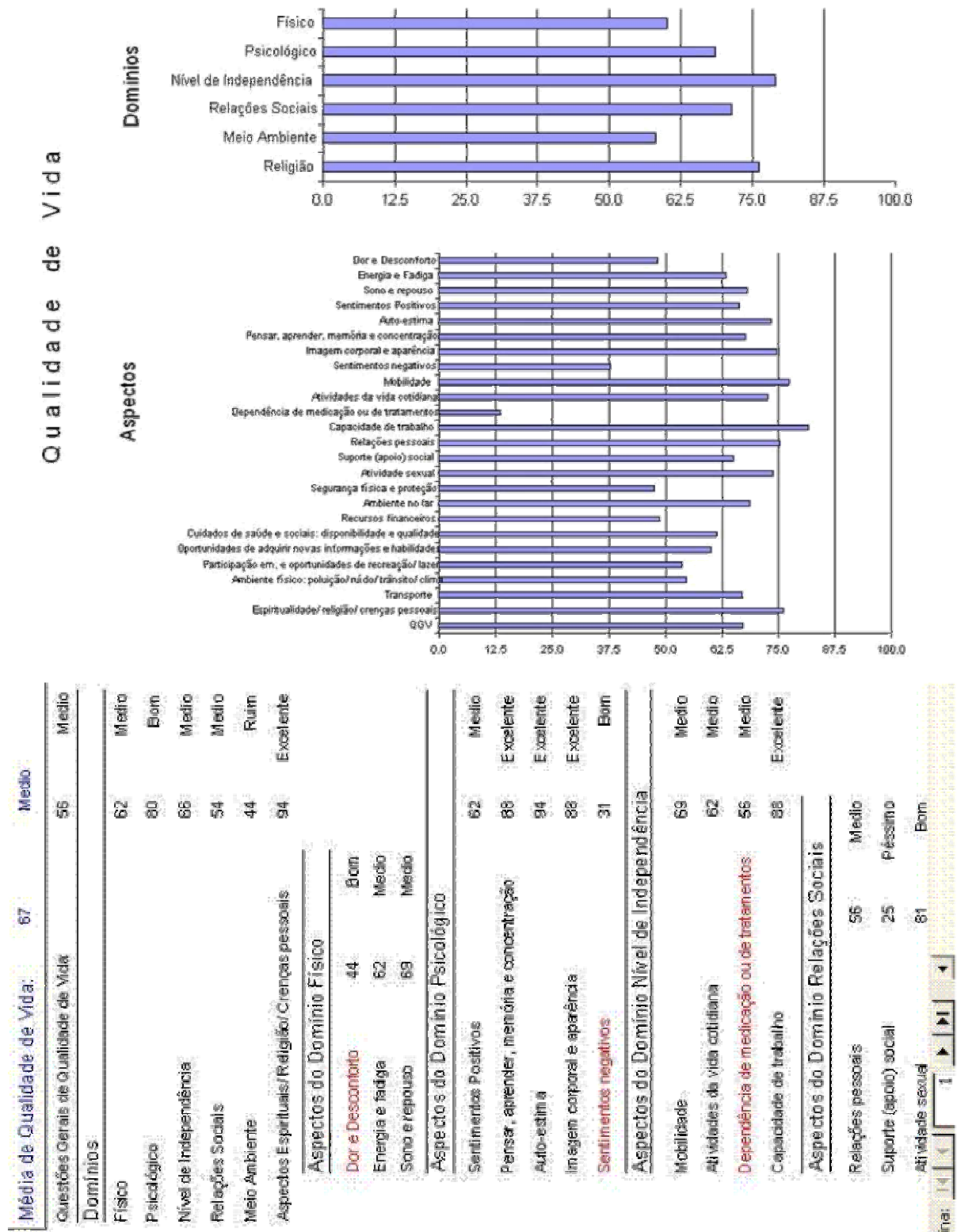
Item	Atributo	Tipo	Valores		Descrição	Fonte
			Distintos	Único		
1	Lotação	categórico	22	21 %	Menor lotação hierárquica	Cadastro do Recursos Humanos
2	Função atual	categórico	41	16 %	Função ocupada após a implantação do PCCS	
3	Sexo	categórico	2	0 %		
4	Idade	numérico	39	2 %		
5	Escolaridade	categórico	8	0 %		
6	Estado Civil	categórico	6	0 %		
7	Cor	categórico	3	0 %		
8	Deficiência	lógico	1	0 %		
9	RTSE	numérico	4	0 %	Registros de cargos ocupados no estado	
10	Dependentes	numérico	8	0 %	Número de dependentes	
11	NPE	numérico	3	0 %	Número de progressos por escolaridade	
12	Bairro	categórico	59	8 %		
13	FG ou CC	lógico	2	0 %	Função Gratificada ou Cargo de Comissão	
14	Participações em Atividades	numérico	10	1 %	Participações em atividades administrativa na qualidade de membro de comissões.	
15	Periculosidade ou Insalubridade	lógico	2	0 %	Trabalha em local perigoso ou insalubre	
16	Afastamento	numérico	3	0 %	Afastamento para pós-graduação/ licença sem vencimento / <u>à disposição</u> de outra instituição	
17	Licença Médica	numérico	18	1 %		

18	NPA	numérico	19	3 %	Número de participações em atividades	
19	Sanção Disciplinar	numérico	8	1 %		
20	Transferência de Lotação	numérico	10	3 %		
21	Últimas Férias	categórico	16	1 %	Mês e ano das últimas férias	
22	Disfunção Definitiva	lógico	2	0 %		
23	Disfunção Temporária	numérico	5	1 %		
24	Alfabetizado	lógico	2	0 %	Escolaridade adquirida após o ingresso na instituição	
25	Primário	lógico	2	0 %	Escolaridade adquirida após o ingresso na instituição	
26	1º. Grau	lógico	2	0 %	Escolaridade adquirida após o ingresso na instituição	
27	2º. Grau	lógico	2	0 %	Escolaridade adquirida após o ingresso na instituição	
28	Especialização	lógico	2	0 %	Escolaridade adquirida após o ingresso na instituição	
29	Graduação	lógico	2	0 %	Escolaridade adquirida após o ingresso na instituição	
30	Mestrado	lógico	2	0 %	Escolaridade adquirida após o ingresso na instituição	
31	Doutorado	lógico	2	0 %	Escolaridade adquirida após o ingresso na instituição	
32	TS	numérico	29	1 %	Tempo de serviço na instituição	
33 - 101	Competências (99)	lógico	2	0 – 1 %	Competência de cada cargo	Perfil Profissiográfico
102	Função Anterior	categórico	58	9 %	Função anterior ao PCCS	
103	CSN	categórico	60	7 %	Classe, Série e Nível atual.	Portaria R 161/06
104	Série 2 Etapa do PCCS	categórico	4	0%	Série na 2 Etapa de Transição do PCCS, que indica progressão ou não por escolaridade.	
105	Nível 3 Etapa do PCCS	categórico			Nível na 2 Etapa de Transição do PCCS, que indica progressão por tempo de serviço.	
106	Salário Base	numérico	60	7 %		Lei Estadual 15.050/06





## ANEXO E – MODELOS DE RELATÓRIOS OFERECIDOS AOS PARTICIPANTES DA PESQUISA



## APÊNDICE A – TRABALHOS PUBLICADOS DURANTE O MESTRADO

HEIL, O. G.; SANTOS, C. B.; HAUSER, M. W. Biomechanic analysis of the march in gestantes through kinematic variable and antropométricas in the city of Ponta Grossa – Paraná. **The FIEP Bulletin**. v.76, p.188-190, 2006.

GRALHA, S. R.; BAGGIO, R.; SANTOS, C. B.; LAUFER-NETO, J.; GOMES, J. C.; GOMES, O. M. M. Estudo comparativo sobre a influência do fotoiniciador das resinas compostas no processo de polimerização. **Brazilian Ora Research**, São Paulo, v.20, p.76, 2006.

BAGGIO, R.; GRALHA, S. R.; LAUFER-NETO, J.; BAGGIO, R.; SANTOS, C. B.; GOMES, J. C.; GOMES, O. M. M. Influência da distância da ponta do fotopolimerizador nas propriedades da resina composta. **Brazilian Ora Research**, São Paulo, v.20, p.141, 2006.

SANTOS, C. B.; ZANDER, V. T.; SCANDELARI, L.; ROMANO, C. A.; CARVALHO, D. R. Data Mining Aplicado na Base de Dados das Categorias Mais Vendidas de um Supermercado. In: KOVALESKI, J. L.; PILATTI, A. L.; GUARNIERI, P. **Temas em Engenharia da Produção II**. Jundiaí: Fontoura, 2006, cap.8, p.87-100.

GRALHA, S. R.; BAGGIO, R.; SANTOS, C. B.; LAUFER-NETO, J.; GOMES, J. C.; GOMES, O. M. M. Estudo comparativo sobre a influência do fotoiniciador das resinas compostas no processo de polimerização. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE PESQUISA ODONTOLÓGICA, 23., 2006, Atibaia, SP. **Resumos...** São Paulo: Santos, 2006.

BAGGIO, R.; GRALHA, S. R.; LAUFER-NETO, J.; BAGGIO, R.; SANTOS, C. B.; GOMES, J. C.; GOMES, O. M. M. Influência da distância da ponta do fotopolimerizador nas propriedades da resina composta. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE PESQUISA ODONTOLÓGICA, 23., 2006, Atibaia, SP. **Resumos...** São Paulo: Santos, 2006.

MENDES, J. T.; LISBOA, A. H.; BORSATO, L. A.; HAUSER, M. W.; SANTOS, C. B.; VIEIRA, J. C. Análise Periodontal em 111 Pacientes Correlacionando com seus Fatores de Risco. In: SALÃO DE INICIAÇÃO CIENTÍFICA DO CESCAGE, 2006. **Anais...** Ponta Grossa: CESCAGE, 2006.

HEIL, O. G.; SANTOS, C. B.; HAUSER, M. W. Biomechanic analysis of the march in gestantes through kinematic variable and antropométricas in the city of Ponta Grossa – Paraná. In: CONGRESSO INTERNACIONAL DE EDUCAÇÃO FÍSICA E CONGRESSO DE FISIOTERAPIA DO MERCOSUL, 21, 2006, Foz do Iguaçu, PR. **Proccedings...** Foz do Iguaçu: New World Gráfica, 2006. p.188-190.

SANTOS, C. B; ZANDER, V. T.; SCANDELARI, L.; ROMANO, C. A.; CARVALHO, D. R. Data Mining Aplicado na Base de Dados das Categorias Mais Vendidas de um Supermercado. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 26., 2006, Vila Galé, Fortaleza, CE. **Anais...** Rio de Janeiro: ABEPRO, 2006.

QUILICI, R. F. M.; SANTOS, C. B; XAVIER, A. A. P.; SCANDELARI, L. A. Utilização do WHOQOL-100 para Verificar o Nível de Satisfação da Qualidade de Vida no Trabalho em Relação a Dor e Desconforto: Estudo de Caso. In: SIMPÓSIO DE ENGENHARIA DE PRODUÇÃO, 13., 2006, Bauru, SP. **Anais...** Bauru: UNESP, 2006.

VAZ, M. S. M. G.; ZAMBIANCO JR, R. A.; ANTONELLI, R. A.; SANTOS, C. B.; BENACCHIO, A. Metamapa: Armazenamento e Indexação de Mapas Digitais Via Metadados. In: CONGRESSO DE TECNOLOGIAS PARA GESTÃO DE DADOS E METADADOS DO CONE SUL, 5., 2007, Guarapuava, PR. **Anais...** Guarapuava: UNICENTRO, 2007.

PEDROSO, B.; PILATTI, L. A.; CANTORANI, J. R. H. ; SANTOS, C. B. ; MARTINS, M. E. A. . Proposta de Um Instrumento de Avaliação da Qualidade de Vida com Base no Lazer. In: JORNADA DA PRODUÇÃO CIENTÍFICA DA EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA DA REGIÃO SUL, 1., 2007, Florianópolis, SC. **Anais...** Florianópolis: CEFET/SC, 2007. v. I. p. 1-5.

PEDROSO, B.; PILATTI, L. A. ; KOVALESKI, J. L. ; CANTORANI, J. R. H. ; SANTOS, C. B. ; MARTINS, M. E. A. . Construção da Faceta Lazer Adaptada ao WHOQOL-100. In: CONGRESSO INTERNACIONAL DE ADMINISTRAÇÃO, 2007, Ponta Grossa, PR, Brasil. **Anais...** Ponta Grossa: UEPG, 2007.

SANTOS, C. B.; SCANDELARI, L. ; CARVALHO, D. R. ; VAZ, M. S. M. G. ; GOMES, J. C. Data Mining para Análise das Necessidades de Treinamento das Competências do Perfil Profissiográfico de Servidores de Uma Instituição Pública. In: SIMPÓSIO DE ENGENHARIA DA PRODUÇÃO: GESTÃO DE DESEMPENHO EM SISTEMAS PRODUTIVOS, 14, 2007, Bauru, SP. **Anais...** Bauru: USP, 2007.

SANTOS, C. B.; SCANDELARI, L. ; CARVALHO, D. R. ; VAZ, M. S. M. G. ; GOMES, J. C. Data Mining para Classificação das Funções de Uma Instituição Pública a Partir das Semelhanças Entre Suas Competências. In: SIMPÓSIO DE ENGENHARIA DA PRODUÇÃO: GESTÃO DE DESEMPENHO EM SISTEMAS PRODUTIVOS, 14, 2007, Bauru, SP. **Anais...** Bauru: USP, 2007.

SANTOS, C. B.; SCANDELARI, L. ; CARVALHO, D. R. ; VAZ, M. S. M. G. ; GOMES, J. C. Data Mining para Definição de Pacotes de Treinamentos de Competências de Uma Instituição Pública. In: SIMPÓSIO DE ENGENHARIA DA PRODUÇÃO: GESTÃO DE DESEMPENHO EM SISTEMAS PRODUTIVOS, 14, 2007, Bauru, SP. **Anais...** Bauru: USP, 2007.

PEDROSO, B.; PILATTI, L. A. ; CANTORANI, J. R. H. ; SANTOS, C. B. Trabalho e Lazer: Um Instrumento de Avaliação da Qualidade de Vida. In: SEMINÁRIO DE INICIAÇÃO CIENTÍFICA E TECNOLÓGICA DA UTFPR, 7, 2007, Curitiba, PR. **Anais...** Curitiba: UTFPR, 2007.

PEDROSO, B.; PILATTI, L. A. ; CANTORANI, J. R. H. ; SANTOS, C. B. ; MARTINS, M. E. A. Desenvolvimento do L-QOL-70: Um Instrumento de Avaliação da Qualidade de Vida com Base no Lazer. In: JORNADA NACIONAL DA PRODUÇÃO CIENTÍFICA EM EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA, 2., 2007, São Luís, MA. **Anais...** São Luís: CEFET/MA, 2007.

TEODORO, A.; SANTOS, C. B.; SCANDELARI, L. KALINOWSKI, F. G. Comparação da Qualidade de Vida de Idosos Hipertensos Fisicamente Ativos e Sedentários. In: SIMPÓSIO DE EDUCAÇÃO FÍSICA DO SUL DO BRASIL, 19, 2007, Ponta Grossa, PR. **Anais...** Ponta Grossa: UEPG, 2007.

SANTOS, C. B.; SCANDELARI, L. ; HAUSER, M. W. ; PEDROSO, B. ; PILATTI, L. A. Quality of Life Indicators of College Students. **The FIEP Bulletin**, Foz do Iguaçu. v.78, n.2, p.411-413, 2008.

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)