

## 5

### A Usabilidade das Estatísticas Públicas

O capítulo anterior descreveu as facilidades de acesso às informações estatísticas e este capítulo descreve, exemplifica e mostra a usabilidade destas informações utilizando a tecnologia OLAP. No item 5.1 é realizada uma comparação entre o processo de análise exploratória das informações da PNAD em formato (texto) e formato OLAP. No item 5.2, novas visões são criadas montando cenários para responder às perguntas de uma análise exploratória da PNAD realizada através do protótipo. Finalizando no item 5.3, com a análise das características e os métodos de análise do cubo que implementam as aferições de uniformidade, intuitividade, amigabilidade e compreensibilidade .

#### 5.1.

##### Comparação do processo atual com o proposto

As Figuras 23 e 24 mostram tanto o processo de análise exploratória atual quanto o proposto. Vê-se que no processo atual, o especialista realiza várias atividades para dar visibilidade às informações para o analista e no processo proposto ele não existe, ou seja, o especialista não precisa realizar tais atividades e o analista acessa e usa os dados diretamente.

No processo atual temos as seguintes atividades:

- **Analista:** Planeja a pesquisa identificando o público alvo, selecionando as variáveis de análise coletadas na PNAD e, para a exploração de dados, monta um esquema de tabelas para responder às perguntas de acordo com a necessidade de análise.
- **Especialista:** Com base no esquema de tabela definido pelo Analista, o especialista acessa os dados disponíveis na Web, transforma e carrega para o seu próprio ambiente de análise. Após realizar a tradução dos dados em informação, ele constrói o esquema de tabelas definido pelo analista. Assim, cada tabela populada pelo consultor é enviada para o analista para validação e dependendo da análise o ciclo é repetido.

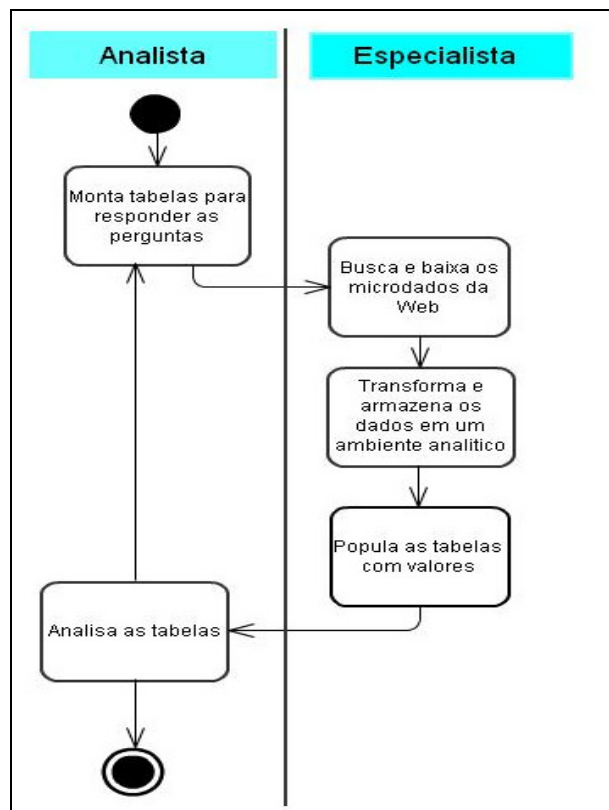


Figura 23 - Diagrama de atividades em cenários sem a aplicação da sistemática proposta

Dificuldades do método:

- O consultor não tem autonomia em suas pesquisas, sempre precisando de um especialista para decodificação dos microdados.
- O analista não tem transparência, pois os dados são transformados em informações por terceiros.
- Falta um mecanismo de análise que responda às perguntas de forma rápida, o que torna o processo lento e dispendioso.

No processo proposto temos as seguintes tarefas:

Após a aplicação desta sistemática, pelos institutos de pesquisa, os dados coletados nas pesquisas tornam-se prontos para serem explorados a partir de softwares existentes no mercado, como por exemplo, o Excel.

**Analista:** Acessa o cubo via Web, através de um software específico, e explora as informações contidas na PNAD.

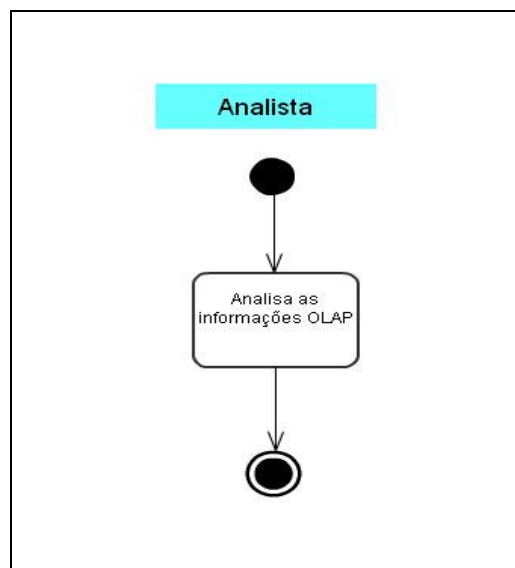


Figura 24 - Diagrama de atividades em cenários, após a aplicação da sistemática proposta

## 5.2.

### Usabilidade da Informação estatística em formato cúbico

Para disponibilizar as informações estatísticas em formato cúbico foram utilizados os conceitos, técnicas e ferramentas OLAP. O uso do XMLA garante a interoperabilidade entre as aplicações OLAP (cliente e servidor), oferecendo flexibilidade na escolha de hardware, ferramenta e plataforma para acessar os dados em formato cúbico. Assim, a avaliação da usabilidade das informações estatísticas deve ser realizada com base nos métodos que o XMLA oferece para a tecnologia OLAP e não com base nas aplicações que implementam estes métodos.

Segundo (Cappelli, 2009), “a capacidade de usabilidade da informação é identificada através da aferição de práticas que implementam características de uniformidade, intuitividade, simplicidade, amigabilidade e compreensibilidade”. Dessa forma, mecanismos OLAP que implementem as características citadas dão indícios do grau de usabilidade que esse trabalho propõe para as Estatísticas Públicas. As características são descritas como:

**Uniformidade:** Capacidade de manter uma única forma

O cubo, por ser um modelo de dados multidimensional único, permite que a navegação e a pesquisa ocorram no mesmo modelo. Dessa forma, geralmente, as ferramentas OLAP apresentam as funcionalidades necessárias em uma única tela.

**Intuitividade:** Capacidade de ser utilizado sem aprendizado prévio

O usuário que estiver familiarizado com o formato tabular (linhas e colunas) para o entendimento de dados quantitativos, pode acessar o demo construído neste trabalho e aprender a utilizar todas as funcionalidades à disposição através da tecnologia OLAP.

**Simplicidade:** Capacidade de não apresentar dificuldade ou obstáculos

Os objetos de manipulação são apenas as dimensões e as medidas, assim com poucas ações gera-se variadas combinações de dados.

**Amigabilidade:** Capacidade de uso sem esforço.

O uso do XMLA permite o acesso aos metadados do cubo explicando os conceitos utilizados na pesquisa estatística na interface, tornando-a mais amigável. Normalmente o cubo é apresentado com suas dimensões em forma hierárquica.

**Compreensibilidade:** Capacidade de ser entendido

A organização lógica da área de trabalho é uma matriz, em que o usuário pode colocar as dimensões de análise nas linhas ou colunas da matriz e as medidas nas células correspondentes possibilitando assim uma melhor organização das informações permitindo fácil entendimento.

### 5.3.

#### **Análise Exploratória da PNAD após a sistemática**

Esta seção tem o objetivo de descrever o processo de exploração da PNAD para os anos 2007 e 2008, elencando as perguntas da pesquisa com a interpretação dos resultados sob os conceitos utilizados para a elaboração da PNAD e o processo de execução de construção da resposta. Essa breve pesquisa foi utilizada para mostrar a usabilidade na análise exploratória utilizando no protótipo construído e disponibilizado neste trabalho.

Embora haja inúmeras aplicações que suportem XMLA, optou-se para usar o Excel 2007, para essa demonstração, por ter o uso amplamente disseminado. O Excel 2007 oferece acesso aos cubos via Web e suporta XMLA a contento, podendo ser substituído por qualquer outra aplicação, web ou não, para o uso do protótipo construído neste trabalho.

A pesquisa iniciou com a definição da população a ser estudada com seus aspectos e níveis de análise. Ao longo do processo de exploração das informações foram realizadas referências ao modelo conceitual de produção da PNAD e outras fontes.

### **Definição da população estudada e os níveis de análise**

Estudo da população brasileira. A população estudada foi definida segundo as características socioeconômicas. Os níveis geográficos escolhidos para análise foram o Brasil como um todo e o conjunto das regiões metropolitanas (RM's<sup>15</sup>).

Inicia-se querendo saber o percentual da população jovem brasileira em relação a toda a população. A população estudada foi definida segundo o critério etário para a classificação dos jovens da OMS<sup>16</sup>, segundo a qual jovens são aqueles indivíduos entre 15 e 24 anos.

1- Qual o tamanho da juventude no nível de análise Brasil?

### **Construção da resposta**

Para criar esse cenário, a dimensão “V0101- Ano referência” foi arrastada para as colunas e a dimensão “V8005-Faixa-Idade do morador” para as linhas e em seguida com base nos dados da tabela 6 foi gerado o gráfico, conforme a mostra na figura 25, para enriquecer a visualização.

---

<sup>15</sup> Uma região metropolitana é definida por um conjunto de municípios integrados e contíguos a uma cidade principal (metrópole) e que tem serviços públicos e infraestrutura comuns ou partilhados. Dessa forma, é possível construir uma base de informações que não diz respeito apenas à cidade principal, pois o entorno muitas vezes é fundamental para mapear a dinâmica das grandes cidades. No Brasil existem, atualmente, 26 RMs nos estados de São Paulo, Minas Gerais e Santa Catarina. No Paraná existe mais de uma RM.

<sup>16</sup> Organização Mundial de Saúde

Tabela 6 - Distribuição da população segundo faixa etária.

População Estimada	2007	2008
V8005-Faixa-Idade do morador		
00 a 14 anos	48.307.829	46.954.185
15 a 24 anos	34.107.951	33.531.719
25 a 34 anos	30.803.983	30.840.899
35 a 44 anos	26.905.673	26.979.535
45 ou mais	49.694.894	51.646.457
<b>Total Geral</b>	<b>189.820.330</b>	<b>189.952.795</b>

### Interpretação dos resultados

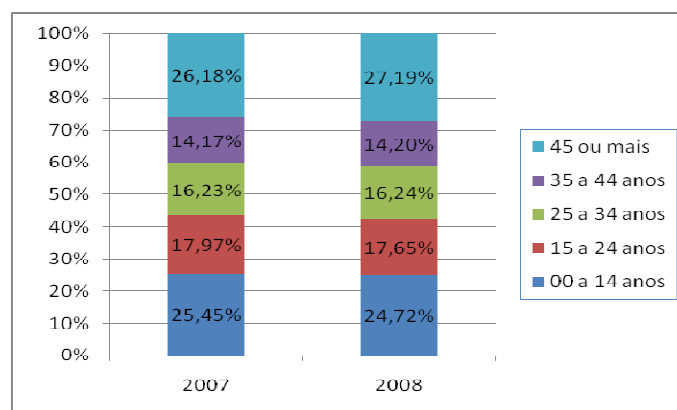


Figura 25 - População total do Brasil segundo faixa etária, 2007 e 2008

O Brasil é considerado um país de jovens, mas esse retrato, como pode ser visto no gráfico acima, vem se modificando, dada à transição demográfica marcada pela redução da fecundidade e mortalidade, junto com o aumento da expectativa de vida. Ao longo dos anos analisados, essa dinâmica fica visível, pois, observa-se no Brasil a redução do número de crianças. A juventude, aqui definida pela faixa de 15 a 24 anos de idade, mantém percentuais muito próximos e é nas faixas a partir de 45 anos de idade que observamos aumentos.

No contexto socioeconômico. O Brasil é um dos países mais desiguais do mundo e, nas últimas décadas, passou por algumas mudanças significativas com impactos no mercado de trabalho, rendimentos e, conseqüentemente, na desigualdade social.

2 - Qual a distribuição da população brasileira por faixa de renda per capita?

### Construção da resposta

Para criar esse cenário, a dimensão “V0101- Ano referência” foi arrastada para as colunas e a dimensão “V4750-Faixa-Rendimento Mensal Familiar per capita” para as linhas, aplicado o filtro excluindo os membros “Não aplicável”, “Sem declaração” e em seguida, com base nos dados da tabela 7 foi gerado o gráfico, conforme mostra a Figura 26.

Tabela 7 – Faixa de renda per capita familiar – população total

População Estimada V4750-Faixa-Rendimento	V0101-Ano	
	2007	2008
Sem rendimento	4.012.562	3.158.825
Até ¼ salário mínimo	21.228.828	15.871.239
Mais de ¼ até ½ salário	35.520.834	30.535.455
Mais de ½ até 1 salário	50.973.830	49.307.925
Mais de 1 até 2 salários	41.696.833	48.730.242
Mais de 2 até 3 salários	13.443.177	16.054.056
Mais de 3 até 5 salários	9.879.862	11.268.919
Mais de 5 salários	7.692.065	9.059.856
<b>Total geral</b>	<b>184.447.991</b>	<b>183.986.517</b>

### Interpretação dos resultados

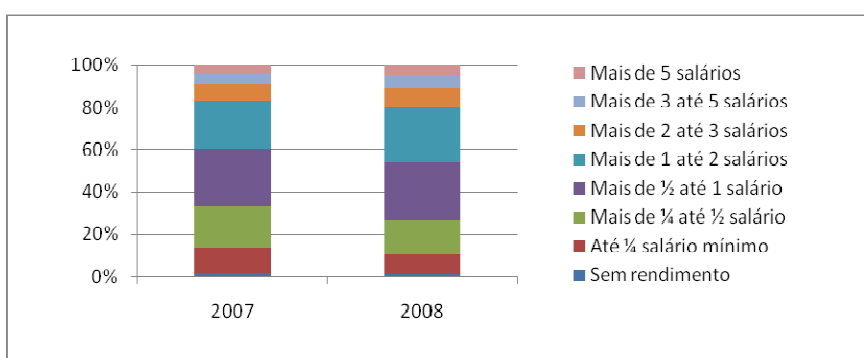


Figura 26 – Faixa de Renda per capita familiar do Brasil, 2007 e 2008

Nesta análise experimenta-se um ciclo ainda pouco definido em relação às faixas de renda mais altas, cristalizando o seu núcleo de pobreza. Os dados assim desagregados mostram a pequena, porém presente, o aumento da faixa

de extrema pobreza. Este estudo, no entanto, define como sendo de baixa renda aqueles indivíduos em famílias com até meio salário mínimo, uma decisão que levou em consideração as definições mais comuns de pobreza.

No contexto de elevada desigualdade de renda e condições de vida, os grupos sociais definidos por cor não vivenciam as possibilidades e os limites de recursos da mesma forma. As desigualdades acumuladas pela população negra do país fazem dessa dimensão um elemento essencial para a análise.

3 - Qual a distribuição da população por faixa de renda, segundo a cor?

### Construção da resposta

Para criar esse cenário, a dimensão “V0101- Ano referência” foi arrastada para o filtro do relatório e utilizou-se o filtro para o ano 2007. Seguidamente as dimensões “V4750-Faixa-Rendimento Mensal Familiar per capita” e “V0404-Cor ou Raça” para as linhas e colunas respectivamente. Foi aplicado os filtro excluindo os membros “Não aplicável” e “Sem declaração” das linhas e selecionando apenas os membros “Branca” e “Negra”(preta + parda) nas colunas e finalizando com a geração do gráfico com base nos dados da tabela 8, conforme mostra a Figura 27.

Tabela 8 – Faixa de renda segundo a cor, Brasil – 2007

V0101-Ano de referência 2007		
Estimativa População	2007	
	Branca	Negra
Sem rendimento	1.552.691	2.395.585
Até ¼ salário mínimo	5.733.717	15.345.154
Mais de ¼ até ½ salário	12.440.546	22.958.638
Mais de ½ até 1 salário	23.339.670	27.278.436
Mais de 1 até 2 salários	24.616.818	16.737.245
Mais de 2 até 3 salários	9.262.207	4.045.377
Mais de 3 até 5 salários	7.467.440	2.278.759
Mais de 5 salários	6.296.367	1.128.354
<b>Total geral</b>	<b>90.709.456</b>	<b>92.266.548</b>



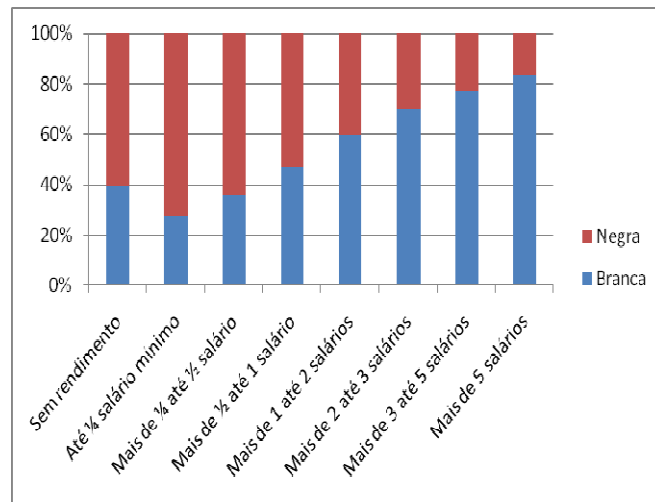


Figura 27- – Faixa de renda segundo a cor, Brasil – 2007

### Interpretação dos resultados

As disparidades entre brancos e negros são existentes, indicando que são persistentes. As disparidades só reduzem um pouco quando consideramos a baixa renda (até meio salário mínimo), mas invertendo a análise, podemos ver que os negros são quase todos pobres.