

ANAIS DO  
VI SIMPÓSIO NACIONAL DOS PROFESSORES  
UNIVERSITÁRIOS DE HISTÓRIA

Organizado pelo Prof. *Eurípedes Simões de Paula*.

**TRABALHO LIVRE E TRABALHO  
ESCRAVO.**

VOLUME II

XLIV

Coleção da *Revista de História* sob a direção  
do Prof. Eurípedes Simões de Paula.



SÃO PAULO — BRASIL  
1973.

A APLICAÇÃO DE COMPUTADORES NA ANÁLISE HISTÓRICA QUANTITATIVA NO BRASIL.  
(Apresentação de um exemplo no setor da Demografia Histórica) (\*).

---

*MARIA LUIZA MARCÍLIO*

do Departamento de História da Faculdade de Filosofia, Ciências e Letras de Assis (SP).

INTRODUÇÃO.

A intensiva e cada vez mais ampla utilização dos recursos oferecidos pela Informática nos diversos campos das ciências naturais, físicas e matemáticas permitiu o aperfeiçoamento e sofisticação progressivos dos processos, programas e análises de Computação eletrônica, inclusive da melhoria da própria máquina. A aplicação dos computadores acabou por penetrar nos setores das Ciências Humanas promovendo uma primeira revolução nas formas de aproveitamento de dados, de experimentação de métodos, da aparição de comportamentos e mentalidades novas de abordagem, tratamento e formas de trabalho, propiciando a cooperação efetiva interdisciplinar de equipes em laboratórios centralizados.

As primeiras tentativas bem sucedidas em Ciências Humanas do uso de computadores foram realizadas naqueles setores onde tradicionalmente se utilizavam com frequência procedimentos de análise quantitativa, matemático-estatística. Somente a partir de uma fase relativamente recente é que se assiste à extensão cada vez mais ampla do uso das máquinas eletrônicas em pesquisas nas Ciências Humanas em seus vários setores. Certas áreas foram, é óbvio, mais precoces na sua aceitação, outras relutaram mais. A Economia e a Demografia são sem dúvida as disciplinas que primeiro deram lugar às pesquisas quantitativas

---

(\*) . — Comunicação apresentada na 3ª sessão de estudos, Equipe D, no dia 8 de setembro de 1971 (*Nota da Redação*).

através da análise em computadores. Elas foram seguidas pela Psicologia e Lingüística. A Sociologia, e a Etnologia estão entre as disciplinas que aceitaram com mais reticências a utilização de métodos matemáticos e de computação eletrônica. Mas também não é menos verdade que o setor mais resistente às inovações continua sendo o da História, quer quanto à quantificação, quer — e porisso mesmo — quanto ao uso dos recursos eletrônicos de análise.

“Tôdas as disciplinas chegaram hoje a um ponto em que elas fazem uso importante da instrumentação matemática. O que era verdadeiro somente para a economia antes da Segunda Guerra Mundial é agora verdadeiro para a Psicologia, a Sociologia, a Antropologia, a Lingüística, sem falar na Demografia” (1).

A timidez manifesta no setor da História deve-se seguramente ao desenvolvimento tardio da quantificação e da busca ainda continuada de “modelos”, definições, e metodologias específicas de operação históricas, sem mencionar dificuldades de caráter externo e material, de infra-estrutura, encontradas geralmente para a implantação de uma mentalidade quantitativa e do desenvolvimento de pesquisas no setor (2). Por outro lado, os projetos de pesquisas, que vêm sendo realizados em centros científicos que se multiplicam, têm mostrado as imensas possibilidades abertas para a História pela quantificação e pela Informática.

“O emprêgo das máquinas eletrônicas pode levar a recolocar em questão os métodos e inclusive a concepção da História”

conforme admite Schneider (3). Este mesmo Autor pondera ainda que

---

(1). — BOUDON (Raymond), *Modèles et méthodes mathématiques*. in “Tendances principales de la recherche dans les sciences sociales et humaines”. Partie 1, sciences sociales. Paris, UNESCO, 1970, p. 629.

(2). — O artigo de PRICE (Jacob), *Principales tendencias de la investigación cuantitativa reciente en el campo de la historia*. In “Perspectivas de la História Economica cuantitativa en America Latina”. México, CLACSO (1), 1970, p. 9-33 indica bem a natureza das dificuldades existentes para o pleno desenvolvimento da quantificação em História em nosso continente.

(3). — SCHNEIDER (Jean), *La machine et l'Histoire: de l'emploi des moyens mécaniques et électroniques dans la recherche historique*. Moscou, Editions Naouke, 1970 (trabalho apresentado ao XII Congrès International des Sciences Historiques. Moscou, 16-23 aout, 1970) p. 70.

“a utilização dos computadores pode conduzir a ciência histórica a definir mais claramente um objeto e seus métodos sobre o duplo plano dos estudos sincrônicos e diacrônicos” (4).

Em Demografia corrente como em Demografia Histórica a própria natureza desta Ciência propiciou a utilização de máquinas de computação eletrônica de forma ampla, quase rotineira. No caso específico da Demografia Histórica numerosas experiências já existem realizadas especialmente em França, União Soviética, Estados Unidos, Canadá, Inglaterra (5).

A aplicação dos novos recursos de investigação — os calculadores eletrônicos — permitiram então à Demografia o que se chama hoje as “simulações”. Nos computadores tôdas as possíveis complicações e construção de “modelos” de população podem ser realizadas. Pode-se, pelos modelos, observar a evolução de valores médios mas também sua variância. Modificando as leis de probabilidades obtêm-se simulações diferentes e o demógrafo realiza então as condições de uma pseudo-experimentação.

“Os calculadores eletrônicos permitiram recentemente às pesquisas baseadas na *aproximação de documentos* abandonar o domínio do artesanato onde a insuficiência de meios os havia acantonado até então. O princípio destas pesquisas reside na aproximação de documentos estabelecidos em ocasiões diferentes. A aproximação no tempo não é a única possível. Pode-se também aproximar em uma mesma data documentos estabelecidos por diferentes administradores” (6).

O historiador das populações está em condições hoje de manipular centenas de documentos, milhares de fichas e dados sobre micro-populações do passado ou vastas coletividades, através da análise eletrônica, o que outrora era impraticável e mesmo inconcebível de ser realizado. Mas a multiplicação do uso dos novos recursos tecnológicos tem sido apanágio de países mais desenvolvidos.

---

(4). — *Idem, Ibidem*, p. 10.

(5). — Informações maiores sobre centros de pesquisa e experiências realizadas em Demografia Histórica em Institutos de pesquisas multidisciplinares nos países citados Vide a obra coletiva da UNESCO citada e também GUILLAUME (P.) e POUSSOU (J. P.), *Démographie Historiques*. Paris, Armand Colin, 1970.

(6). — BOURGEOIS-PICHAT (Jean), *La Démographie*. In “Tendances principales de la recherche dans les sciences sociales et humaines”, *ob. cit.*, p. 488.

Para os países não industrializados que manipulam mal ainda as modernas técnicas da Informática, as dificuldades práticas do uso dos novos meios de investigação por vêzes desanimam o pesquisador antes mesmo dêle tentar os primeiros passos. Examinemos alguns dêstes problemas.

### DIFICULDADES QUE ATRASAM O USO DE COMPUTADORES EM CIÊNCIAS HUMANAS NO BRASIL.

A utilização dos recursos da ciência de Computação está no nascedouro no Brasil. Com relativa maior penetração nos setores privados — bancos, emprêsas, etc., — sua introdução normal nas Universidades e Centros de Pesquisas é bastante recente e não se pode dizer ainda que a Computação eletrônica esteja integrada à vida acadêmica. Dentro das nossas Universidades são — como seria natural que o fosse — as físicas, e as matemáticas que a introduziram primeiro e que maior uso fazem dêstes recursos. No entretanto, algumas pesquisas pioneiras nas áreas humanas começam a ser realizadas, particularmente em economia, e em sociologia quantitativa, além da demografia, no Brasil. Sendo a própria informática um setor novo entre nós e quase integralmente dependente do exterior em experiência científica, dificuldades de caráter prático são extremamente grandes.

As observações que fazemos a seguir partem de uma experiência isolada e porisso mesmo limitativas. Mas acreditamos que a discussão dêstes tipos de problemas pode facilitar e economizar esforços dos que buscam a aplicação de recursos aperfeiçoados no setor da História quantitativa do Brasil.

1. — Os trabalhos de pesquisa, processamento, análise e interpretação em História quantitativa apresentam como característica básica a de exigirem antes de mais nada a existência de uma equipe de profissionais de variadas disciplinas reunidos em um Centro de investigações e com objetivos comuns. A inexistência de tais centros ou laboratórios de trabalhos apresenta-se como a mais séria dificuldade para o desenvolvimento dêstes tipos de pesquisas. A variedade de tarefas que começam com a busca, seleção e coleta de dados básicos, padronisadamente, em arquivos e bibliotecas, a discussão dos métodos de trabalhos, escôlha de técnicas de tratamento desde as estatístico-matemáticas às específicas conforme o tipo de pesquisa (demográfica, econômica, sociológica, psicológica, lingüística, etc. em História) impõem a presença de uma equipe que se proponha a metas comuns. Vários países já compreenderam há algum tempo esta necessidade e poderíamos exemplificar com as experiências bem sucedidas que vêm sendo realizadas na Universidade de Caen, com Chaunu, no *Institut National*

*d'Etudes Démographiques* com Louis Henry à frente, com a equipe de demógrafos e economistas de História Quantitativa de Cambridge, com o grupo do Centro Latino Americano de Demografia (Santiago) e outros. No Brasil, os estudos quantitativos históricos realizados em equipes pluridisciplinares e com especialistas de formação variada são raros, diríamos inexistentes. O que há são tentativas de pesquisadores isolados.

2. — Qualquer pesquisa em História quantitativa com uso de computadores exige grande dispêndio em tempo e em custos operacionais o que, mesmo com o apóio financeiro de fundações de amparo à pesquisa, exorbita o quadro de pesquisas individuais. A constituição de séries cronológicas de preços, de estatísticas vitais (Batismos ou Nascimentos, Casamentos e Óbitos), educacionais, de rendimentos agrícolas (por sementes ou por produtividade dos solos ou do trabalhador), o levantamento de palavras e conceitos em textos, e outros exemplos de pesquisas quantitativas históricas demanda tempo bastante longo na busca, coleta e seleção dos dados. Tempo gasto nas deslocações do pesquisador pelos nossos vários arquivos descentralizados e esparsos por um vasto território (sem contar os arquivos portugueses ou de outros países estrangeiros, quando as despesas de locomoção do pesquisador se avolumam ainda mais). Apesar das técnicas modernas de microfilmagem e de cópias em papel (sistemas Xerox, Copi-flo, ou fotocópia) acelerarem e facilitarem o trabalho do pesquisador, ainda assim a falta de centralização arquivística e de orgânica interna dos arquivos nacionais existentes obrigam a presença e a busca artesanal nesta fase da pesquisa. Mas para que os dados colhidos possam ser utilizados pela máquina é necessário que sejam coletados de forma o mais possível padronizada, para que possam ser depois facilmente transformados em códigos numéricos para a entrada em Computadores. A preparação dos dados para a máquina acaba também por consumir tempo ainda grande no decorrer da pesquisa. E em História quantitativa, é obvio, não são documentos isolados, não representativos que entram em análise, mas séries ou conjuntos de informações as mais completas possíveis que permitam reproduzir uma imagem real ou o mais representativa possível da realidade ou universo em estudo.

3. — A falta de pessoal habilitado em programação eletrônica ligado às características das disciplinas humanas talvez seja o aspecto mais sério, entre nós, de utilização dos recursos oferecidos pela computação. Um verdadeiro abismo de incomunicabilidade existe aqui, entre, de um lado, os nossos escassos programadores de computação, preparados todos eles para operarem em setores estranhos às problemáticas específicas das pesquisas humanas, e, de outro lado, os cientistas humanos e sociais quase inteiramente desinformados sobre a me-

cânica mais elementar e as possibilidades oferecidas pela máquina eletrônica. Daí gerar-se um quase impossível diálogo entre os dois setores, que na prática já estão fisicamente dissociados dentro de nossas Universidades, resultando em ensaios e erros de ambos os lados e no conseqüente encarecimento inútil, em tempo e em recursos da pesquisa. Em computação o custo conta-se por segundos de utilização das máquinas, sem pensarmos na parte não menos custosa da preparação dos dados para a linguagem da máquina. Os pesquisadores em ciências humanas, por falta de uma formação sistemática sobre computação, acabam por acreditar que em matéria de computação basta apertar botões, para se obter, em segundos, resultados que demandariam anos de cálculos e análises. Ainda se busca entre nós uma linguagem comum entre os dois campos, o que já foi possível em outros países, através de cursos básicos elementares de programação eletrônica ministrados exclusivamente para cientistas de ciências humanas.

4. — Na utilização racional dos meios eletrônicos na pesquisa histórica busca-se a escolha prévia de variáveis a serem analisadas e possivelmente a constituição na máquina de modelos de trabalho. As numerosas experiências efetuadas em centros de pesquisas existentes tem demonstrado as possibilidades e limitações geradas em História Quantitativa pelo uso de computadores. Na falta de equipes de pesquisas pluridisciplinares no Brasil no campo da História Quantitativa, os pesquisadores isolados nacionais têm normalmente se servido de modelos e métodos aplicados em outros países, esforçando-se por adaptá-los às características da realidade brasileira, o que pode criar distorções.

5. — A inexistência de cursos estáveis e de formação sobre as técnicas e métodos em História quantitativa impedem a preparação de quadros novos de especialistas levando ao auto-didatismo desestimulante, anti-econômico e retardatário, de elementos isolados interessados em trabalhar no setor.

Estas e outras dificuldades que não são exclusivas do Brasil tem dificultado a implantação de estudos quantitativos históricos entre nós.

A título de exemplo de uma pesquisa que vem sendo realizada por nós e onde foi utilizado o recurso da computação, comunicamos a seguir os estágios técnicos seguidos para a concretização do projeto que visa à reconstrução da estrutura e da dinâmica da população da então Capitania (mais tarde Província) de São Paulo e suas linhas básicas de tendência. O objetivo é mais o de demonstrar a viabilidade e possibilidade dentro das nossas precárias condições da pesquisa histórica, quer da História Quantitativa, quer ainda e dentro dela, da aplicação das técnicas de computação.

## ETAPAS DO PROCESSAMENTO ELETRÔNICO DOS DADOS CENSITÁRIOS ANTIGOS DA POPULAÇÃO PAULISTA.

A pesquisa que vimos realizando desde 1968 visa portanto à reconstituição da população paulista a partir da segunda metade do século XVIII até nossos dias.

A base empírica do projeto é constituída pelos *censos* demográficos levantados para tôda a região desde o final do período colonial, durante a época imperial e a republicana, censos variados que atendiam a finalidade diversas e que se apresentam também sob distintas formas (7). Teòricamente êstes dados são operados dentro da variada metodologia oferecida pela Demografia visando com ela reconstituir a população a partir de um plano amostral probabilístico, aplicado ao levantamento dos dados dos primeiros censos, do tratamento estatístico dos dados assim levantados, da correção dos resultados demográficos obtidos através da aplicação de modelos de Tábuas de Vida e de populações estáveis (Lotka), de elaborar as linhas evolutivas gerais da população e de variáveis dela, para finalmente poder interpretar a realidade histórico-social em que está incluída.

A parte mais trabalhosa e longa da pesquisa é sem dúvida a que se refere ao levantamento e tratamento dos primitivos censos da população — *as listas nominativas de habitantes* — que constituem os elementos cronológicos iniciais do trabalho. Êste material censitário conservado para tôdas as vilas que compunham o quadro geográfico da Capitania paulista é estabelecido a partir de 1765 e sua prática anual mantém-se mais ou menos contínua e homogênea até por volta de 1830 quando perde ràpidamente a sistemática, a universalidade, a integralidade para se extingüir de vez depois da segunda metade do século XIX. Êsse material encontra-se conservado no Arquivo Público do Estado de São Paulo. Sòmente estas primitivas listas nominativas sofreram tratamento amostral na fase de levantamento de suas informações sócio-econômico-demográficas, e também sòmente seus dados foram tratados para serem processados em computador. O Computador utilizado foi B/360 do Departamento de Física da Universidade de São Paulo. Os resultados dos demais censos mais recentes usados no estudo foram aproveitados das publicações censitárias existentes desde o trabalho de

---

(7). — Para a descrição das fontes censitárias paulistas, formas de levantamentos, agentes censores e finalidades dos primitivos censos nominativos paulistas vide MARCÍLIO (M. L.), *La ville de São Paulo: peuplement e population. 1750-1850*, Rouen, Université de Rouen. 1968; *Idem* e LISANTI (Luís), *Problèmes des sources en Histoire quantitative du Brésil*. Trabalho a ser apresentado no *Colloque International d'Histoire quantitative du Brésil*. Paris, 1971.

Pedro Müller (1836) até os atuais censos demográficos pelo IBGE, passando pelos levantamentos populacionais publicados nos Relatórios de Presidentes de Províncias e dos Ministérios do Império. Todos êstes últimos censos publicados não entram pois dentro das considerações que fazemos nesta comunicação.

Cinco datas foram escolhidas dentro de critérios fixados, para a primeira fase em que se fundamenta a pesquisa: 1765, 1798, 1808, 1818 e 1828 (8).

Nossa intenção é a de apresentar a seguir os três estágios da preparação dêsses dados amostrais levantados para a linguagem da máquina e os resultados obtidos. São êles:

1. — organização e padronização da forma de levantamento dos dados contidos nas listas nominativas de habitantes paulistas;
2. — operações de codificação e de perfuração dos cartões IBM;
3. — programação e processamento eletrônico pròpriamente dito (9).

#### 1. — *Organização e padronização dos dados.*

Antes de se poder manipular os dados de população contidos nas listas de habitantes da antiga região paulista (São Paulo e Paraná), impunha-se uma tarefa de conhecimento prévio, ano por ano, vila por vila, da existência, estado de conservação e integridade dos documentos.

Estabelecidas as datas censais para o levantamento, necessária se fazia a organização do material. Apesar das listas de população estarem conservadas em latas no Arquivo Público do Estado de São Paulo, classificadas por vilas e em ordem cronológica, o conteúdo destas latas nem sempre correspondia à sua classificação externa e pelo menos

---

(8). — Para não sobrecarregar o texto, e, como foge à sua finalidade, não apresentaremos aqui as técnicas da amostragem utilizadas e seus critérios. Para maiores informações a respeito remetemos os interessados ao nosso trabalho: *Algunos aspectos de la fuerza de trabajo, en la Capitanía (Provincia) de São Paulo* apresentado nas III Jornadas Internacionales de História Economica y Social. Buenos Aires, Instituto Torcuato di Tella. 1970.

(9). — Uma vez terminadas estas operações em Computador encontramos um projeto de pesquisa bem similar realizado a partir dos "padrones" de Caracas por BRADY (T. M.) e LOMBARDI (J. V.), *The application of computers to the analysis of Census data: the bishopric of Caracas 1780-1820*, in "Population and Economics. Proceedings of section V (Historical Demography) of the Fourth Congrès of the International Economic History Association". Winnipeg. University of Manitoba Press. 1970, pp. 271-278.

apresentava-se em grande mistura, cuja ordem precisava ser restabelecida para poderem ser utilizadas.

Forçosamente algumas das vilas da região não dispunham mais, ou não tiveram realizados, seus levantamentos nominativos nas datas da pesquisa. Para não se perder a integralidade territorial da população a solução que adotamos foi a da escolha da data imediatamente anterior e na falta desta a da data imediatamente posterior para completar tôdas as vilas naquelas datas. Solução não ideal, é obvio, mas cujo mérito está em não apresentar claros de dados para alguns poucos municípios.

Para a reorganização do material era necessária prèviamente uma familiaridade bastante grande com o processo de povoamento, a nomenclatura e as mudanças eventuais nos territórios e nos nomes das vilas, enfim com a própria geografia de tôda a região paulista da época.

As técnicas de levantamento do material constituiu a etapa seguinte. Etapa fundamental em que entram em jôgo o rigor na aplicação do plano amostral, o cuidado na correta, integral e verdadeira coleta e na transcrição das informações e a exatidão na leitura paleográfica do documento. Desta etapa, do planejamento das técnicas a serem nela empregadas, dependeria em grande parte o sucesso ou insucesso posterior do uso dos procedimentos da Computação Eletrônica. Assim é que os objetivos finais do pesquisador precisam ser claramente definidos nesta fase: quais as variáveis que pretende analisar, o tipo de tratamento que objetiva (demográfico, social, etc.), os resultados que pretende atingir. No caso específico da pesquisa em questão a formação teórica demográfica encaminha e informa o estabelecimento do plano de *padronização* das informações contidas no material para seu equacionamento com: 1). — a linguagem da máquina; 2). — o tratamento demográfico posterior dos dados.

## 2. — *Codificação e Perfuração dos cartões IBM.*

Antes de iniciar as operações necessárias para o uso dos Computadores, os dados poderiam ser levantados diretamente do documento para as fôlhas de codificação. No entanto, a codificação direta do documento, porquanto economizasse tempo precioso, na fase de coleta, incorreria em sério e irremediável risco de tornar as obrigatórias etapas de correção da codificação e da coerência interna dos dados, mais longas do que o desdobramento das tarefas em duas partes: 1. — transcrição direta das informações do documento em formulários especialmente concebidos de forma padronizada; 2. — passagem dêstes dados para fôlhas de codificação especiais. Por cautela e para permitir a manipulação constante e à mão dos dados do arquivo, preferi-

mos depender maior tempo na coleta padronizada prévia dos dados amostrais. Realizada esta longa operação poderemos a qualquer momento voltar aos dados originais transcritos, controlar os erros eventuais de amostragem, codificação e perfuração, acompanhar rapidamente as operações de correção da coerência interna dos dados realizada pelo computador.

Recopiados em formulários os dados primários passou-se à fase de codificação em fôlhas especialmente concebidas. A operação de codificação corresponde em estabelecer uma relação direta entre cada variável que se pretende analisar e um código numérico e que este cubra a totalidade das informações disponíveis classificadas de acordo com os interesses da pesquisa.

Para tanto é preciso definir as variáveis e classificá-las em categorias para se estabelecer o plano da codificação.

Em nossa pesquisa foram definidas as seguintes variáveis tomadas das duas unidades elementares, pontos de partidas: o "fogo" (domicílio) e o indivíduo:

1. — Data.
2. — Vila.
3. — Critério da Amostra.
4. — Sexo.
5. — Idade.
6. — Estado Civil.
7. — Côr.
8. — Naturalidade.
9. — Situação Familiar 1 (situação de cada indivíduo dentro do fogo em relação ao *chefe de fogo*, incluindo livres e escravos).
10. — Situação Familiar 2 (situação de cada indivíduo dentro do fogo, agora em relação a cada um dos *chefes de família* eventualmente presentes no mesmo domicílio, incluindo igualmente livres e escravos).
11. — Profissão do Chefe de fogo 1: setor de atividades primárias.
12. — Profissão do Chefe de fogo 2: setor de atividades secundárias.
13. — Profissão do Chefe de fogo 3: terciário.
14. — Profissão do Chefe de fogo 4: Outras atividades.

Em cada uma destas variáveis foram definidas categorias ou indicadores que especificam melhor cada uma daquelas. Para exempli-

ficar a variável Profissão do Chefe de fogo 1: setor primário compreende as categorias:

- a. — senhor de engenho.
- b. — lavrador (agricultor, roceiro, etc.).
- c. — fazendeiro de criar.
- d. — criador e lavrador.
- e. — extração natural.
- f. — pescador.
- g. — pescador e lavrador.
- h. — mineiro ou fiscador.
- i. — senhor de engenhoca.

A operação de codificação consistiu então na transferência das informações escritas para seus códigos numéricos respectivos, realizada em fôlha especial de codificação.

A seguir iniciaram-se as operações de perfuração e verificação dos códigos em cartões IBM, cujo total final totalizou aproximadamente 30.000 cartões.

### 3. — *Programação e Análise em Computador.*

Com as operações de codificação e perfuração completadas, o estágio analítico do projeto inicia-se.

A primeira série de questões a colocar na máquina, através de programas especiais concebidos, e com a linguagem FORTRAN, refere-se à coerência interna e integralidade dos dados contidos nos cartões. Através da listagem simples obtida pôde-se conseguir detectar os erros de coerência e procurar corrigir os cartões incorretamente perfurados ou codificados indevidamente.

A seguir o técnico foi solicitado a conceber um programa que seria aplicado separadamente a cada um dos cinco censos — para se chegar às seguintes listagens:

1. — das freqüências simples de cada uma das variáveis e suas respectivas categorias separadamente e com as devidas porcentagens sobre os totais gerais de cada censo. Ter-se-á então, por exemplo, a distribuição simples da população por sexo, ou por grupos de idades, ou por côr, etc. e em cada data.

2. — das tabelas com números absolutos e vários tipos de porcentagens com cruzamento de duas entradas. Tôdas as combinações podem ser realizadas mesmo que seu interêsse seja restrito ou nulo. A seleção das tabelas é feita “a posterior”.

3. — das tabelas mais complicadas com cruzamentos de três e até cinco variáveis com tôdas as porcentagens.

Com a informação derivada destas primeiras séries de cálculos que abrangem, no total geral, mais de 130.000 indivíduos distribuídos nos seus quase 18.000 fogos da amostra e nas cinco datas, pode-se começar a colocar na máquina questões mais elaboradas. Por exemplo, questões específicas sôbre as peculiaridades entre fogos com grande número de escravos e sem escravos; sôbre o dimensionamento e suas tendências dos domicílios e das famílias nucleares; sôbre as migrações internas e externas; sôbre os processos de povoamento de sub-regiões da Capitânia, sôbre a fôrça de trabalho, etc. .

A distribuição por sexo e idade em censos com intervalos regulares permite a simulação em computadores de programas buscando a aplicação de modêlos demográficos de populações estáveis, para se chegar às taxas de fertilidade, natalidade e mortalidade através de modêlos de Tábuas de Vida e outros, etc. .

Os historiadores econômicos podem se servir das listas nominativas para estabelecer a análise em Computadores dos primórdios do desenvolvimento econômico paulista. E êsse material e as primeiras tabelas saídas da máquina são indispensáveis para a análise, como dissemos, dos fenômenos do povoamento, migração, mestiçagem, tendências numéricas da escravidão, da fôrça de trabalho, da urbanização, e do dimensionamento da família e dos domicílios.

Sòmente depois de concluídas as fases descritas e com a obtenção das tabelas com resultados quantitativos simples, simulados ou construídos a partir de modêlos, chega-se à etapa de interpretação quantitativa e qualitativa dos resultados, inseridos no contexto histórico.

\*

\*        \*

## INTERVENÇÕES.

Da Profa. *Odah Regina Guimarães Costa* (ICH/UFP. Curitiba. Paraná).

Pergunta:

Até que ponto pode-se confiar nos registros paroquiais no que se refere aos dados relativos à história demográfica e obtidos nos registros de batizados, *causa mortis* e de óbitos?

Qual a margem de êrro?

\*

Da Profa. *Antonietta de Aguiar Nunes* (FFCL/São Bento. PUC. São Paulo).

Pede os seguintes esclarecimentos:

Onde foi feita a computação de dados?

Quem fez a programação na exposição como “analistas”?

Qual a sua programação na exposição como “analistas”?

Qual a sua formação e como se deu a correção do programa para dar os dados esperados?

Há outros especialistas em Demografia trabalhando também com computação?

\*

Da Profa. *Geraldina Pôrto Witter* (IP/USP. São Paulo).

Solicita os seguintes esclarecimentos:

Qual o processo de amostragem usado?

Que porcentagem da população de fogos foi amostrado para cada centro urbano? E para a província como um todo?

Que técnica(s) usaram para a codificação?

Qual o critério para a avaliação do grau de alfabetização do dono de cabeça da casa?

Usaram alguma das escalas convencionais para esta avaliação? Usaram pelo menos dois juízes? Foi feita correlação das avaliações feitas pelos mesmos, pelo menos para 25% da amostra?

Estão planejando fazer a significância estatística da diferença entre as porcentagens e médias?

\*

\* \*

## RESPOSTAS DA PROFESSORA MARIA LUIZA MARCÍLIO.

À Profa. *Odah Regina Guimarães Costa*.

Diz que a partir da elaboração do método de “reconstituição de famílias” por *Henry-Fleury* a série de registros paroquiais da época moderna, que até então era raramente explorada, cientificamente, pelos historiadores, passou a constituir a fonte principal e mais importante do mais atual setor da História Quantitativa, isto é, da *Demografia Histórica*.

Mas, evidentemente, como qualquer espécie de série de documentos utilizados em História Quantitativa, a dos Registros Paroquiais têm suas limitações. Estas são impostas, fundamentalmente, pelos próprios documentos, pelo estado de conservação em que se encontram. Partindo do pressuposto da existência, para uma determinada paróquia, ou conjunto de paróquias que se pretende estudar, da existência das três séries de batismos, casamentos e óbitos, para que sua análise tenha validade científica é imperiosa sua crítica. Para tanto, os próprios autores citados elaboraram, igualmente, uma metodologia científica específica.

Assim é que o pesquisador necessita examinar rigorosamente suas séries, para constatar se são completas, universais e representativas da população, ou de parte da população que quer examinar (por exemplo população de livres). Precisa ainda verificar se suas séries são contínuas durante o período que o ocupa.

Claro está que, especialmente para a série de óbitos, e em nosso caso brasileiro dos séculos XVIII e XIX, não se pode esperar uma integralidade perfeita na cobertura dos registros. Mas, não é este também o caso dos censos e das estatísticas vitais das atuais populações das áreas pouco desenvolvidas?

A Demografia corrente e a Demografia Histórica desenvolveram técnicas elaboradas para a correção de dados básicos falhos ou omissos, através da aplicação de modelos matemáticos de populações teóricas. Mas estes sofisticados recursos nem sempre são necessários para a análise demográfica dos Registros Paroquiais.

\*

À Profa. *Antonietta de Aguiar Nune*.

Afirma que os dados foram processados no Computador B/3.600 do Instituto de Física da Universidade de São Paulo. Esta parte do seu programa de pesquisas contou também com o concurso financeiro da Fundação de Amparo à Pesquisa do Estado de São Paulo, a quem agradece. Na própria Física foram elaborados os programas.

É evidente que outros especialistas em Demografia, dos grandes centros de pesquisas, já trabalham, há algum tempo, com os recursos atuais da informática. Para a América Latina, no entanto, e até onde vão os seus conhecimentos, este programa de pesquisas é pioneiro. Pela primeira vez tenta-se os procedimentos eletrônicos na análise de dados quantitativos históricos, com fins demográficos, na região.

\*

À Profa. *Geraldina Pôrto Witter*.

Declara que o plano de amostragem para os seus dados referentes às listas nominativas de habitantes de tôdas as vilas que compunham a então Capitania, depois Província de São Paulo, nas datas escolhidas (1765, 1798, 1808, 1818 e 1828), foi estabelecido em várias etapas. O elemento de base tomado foi o “fogo”, isto é, o *domicílio* segundo o seu conceito demográfico. Cada vila constituiu um *estrato*, na amostra, e dentro de cada vila foram estabelecidos três critérios de natureza sócio-econômica, para os “fogos”. Assim classificou em fogos: 1) . — sem escravos; 2) . — com até 39 escravos e 3) . — fogos grandes, ou com 40 escravos ou mais. Para cada *estrato amostral* considerado foi sorteado o primeiro fogo, dentro da técnica do *começo casual simples*. Tomou-se para cada critério uma proporção de 1/8 dos fogos.

A codificação dos dados amostrais probabilísticos levantados obedeceu o esquema de acôrdo com as informações contidas nos antigos censos e de agrupamento das variáveis seguindo critérios teóricos: demográficos e econômicos.

Para ficar em um só exemplo, as várias categorias de profissões dos chefes dos “fogos” foram codificadas dentro dos três setores básicos de atividades econômicas definidos por Colin Clark e para cada um deles foram estabelecidos sub-grupos de atividades. Assim, para o setor secundário os sub-grupos (10) foram, entre outros, atividades artesanais com metais, com couros, com madeiras, etc. .

O contrôle da amostra foi realizado após o levantamento.

Apenas agora os trabalhos de levantamento, codificação, processamento e correção dos dados foram terminados. Começou então a análise dos mesmos, que começa pelos testes estatísticos de verificação da significância da amostra e dos dados, para terminar na interpretação dos resultados demográficos analisados dentro de uma perspectiva especialmente social.