

PROF. DOUTOR LUÍS FARINHA
DOCENTE DA UNIVERSIDADE NOVA DE LISBOA
INVESTIGADOR INSTITUTO HISTÓRIA
CONTEMPORÂNEA



INFORMÁTICA E CIÊNCIAS SOCIAIS E HUMANAS

A Informática é um instrumento de trabalho que permite um processo sofisticado de construção do conhecimento e de apresentação desse conhecimento a um público generalizado, quer através de suportes locais, quer, e especialmente, através de meios de comunicação à distância. Pelo rigor obtido na construção do conhecimento e pela eficácia conseguida na sua comunicação, a Informática exige uma cuidadosa programação de cada um dos trabalhos pretendidos e executados.

Para começar, deve contar-se com o apoio dos mais entendidos, quer na adequação do(s) programa(s) escolhidos para a execução do trabalho e sua conveniente utilização, quer nas formas mais bem conseguidas de apresentação do trabalho ao público. Dependendo da natureza da informação a tratar e dos meios de análise indispensáveis à sua interpretação, assim a nossa escolha deve recair sobre um ou outro dos programas disponíveis e indispensáveis: normalmente um programa de tratamento de texto e um outro de registo e tratamento de dados, quando não mesmo o auxílio de um programa de cálculo para operações de contagem. São muitos os programas actualmente disponíveis no mercado: do mundialmente conhecido e utilizado Word para tratamento de texto, aos já mais diversificados programas de recolha e trata-

mento de dados (Acess ou SPSS), ou ainda à também muito utilizada tabela de Excel.

Não menos importante é a forma como programamos a apresentação das conclusões, depois de resumidos e interpretados os dados. Esta preocupação é tanto maior quanto o espaço/tempo de que dispomos para comunicar as nossas conclusões se torna escasso e diverso num mundo global: deverá adquirir formas diferenciadas de acordo com os suportes/meios (livro, conferência, Internet), com o público a que se destina (especialistas, meio académico, comunicação social, público em geral) e com os suportes linguísticos utilizados (língua nacional ou internacional). Não raro é habitual assistirmos a um “despejar” de informação numa conferência com base num suporte informático que foi pensado para comunicar numa reunião de pares. O público perde o essencial, o comunicador



fica frustrado e o objecto da comunicação não chega ao seu destinatário. O mesmo se poderia considerar com a informação que disponibilizamos para um qualquer “sítio” da Internet ou aquela que é seleccionada para ser apresentada em livro.

Cada um destes trabalhos (de recolha, tratamento, análise e interpretação de dados e apresentação de conclusões), pequenos ou grandes, é um projecto – com fases de planeamento e execução, com avaliação contínua e contínua correcção, e com avaliação final. E, mais interessante ainda, com possibilidades infinitas de manipulação posterior, através de correcções, cortes ou acréscimos da informação existente. Se é verdade que um trabalho em ciência nunca está verdadeiramente acabado, então a Informática aí está para nos facilitar esse trabalho incessante de correcção e reavaliação constante.

Desta vez, dirigirei especialmente a minha preocupação para a Informática como instrumento auxiliar de construção do conhecimento em ciências sociais e humanas, muito particularmente em História. Do mesmo modo, não irei valorizar toda a (importante e necessária) experiência prática de aquisição básica do trabalho com processadores de texto ou programas de

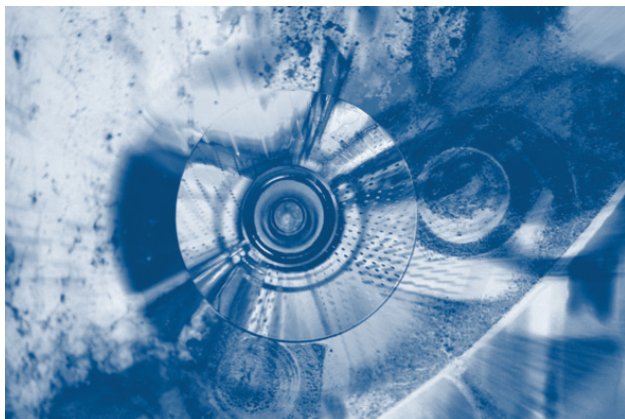
recolha e tratamento de dados. Sendo conhecimentos indispensáveis, exigem apenas aprendizagem inicial e muita prática. Não foi esse o caminho dos calígrafos que todos nós já fomos? Difícil, dirão... Fica-nos a vantagem de pensar que a história que “começou na Suméria” está agora a recomeçar connosco. Não é um privilégio?

Fora das minhas preocupações estão também todos os elementos de análise estatística indispensáveis à realização do trabalho informático.

1. PONTO DE PARTIDA

Alguém alguma vez terá afirmado que a estatística é a “grande ciência” do mundo contemporâneo. Sabemos bem como com ela se mede, ajusta, ajuíza e conclui de forma categórica, com base nas conclusões matemáticas de que parte. Também sabemos como ela pode ser incorrectamente manipulada por economistas, por financeiros, por políticos ou pelos meios sofisticados da informação que utilizamos.

O que é certo é que não saberemos nunca pensar e viver sem o recurso a um “pensamento estatístico”: apreciamos, medimos, calculamos, ajuizamos e concluímos com base na apreciação de muitos (ou poucos) casos semelhantes, mesmo se os nossos instrumentos de medição são tão rudimentares como a memória, o cálculo mental ou a visão. E, normalmente, não nos enganamos. Apenas pensamos com “menos rigor”. Se, por acaso, num dado mês, excedemos o orçamento individual, não é por falta de noção do que fazemos, mas antes por vontade de compensar o excesso com o benefício



que daí recebemos. Se os nossos rendimentos não variarem durante algum tempo (coisa habitual e vulgar), estaremos seguramente condenados a conhecer ao pormenor o nosso “cabaz mensal” durante um tempo muito longo. Ou seja, o nosso “pensamento estatístico” (por cálculo, por memória...) chega-nos bem para saber até onde pode ir a bolsa. E, para isso, não precisamos de grandes cálculos matemáticos.

Mas apesar de tudo, os produtos vão variando nas montras, os preços (normalmente) subindo, assim como a nossa avidez por um maior consumo. Nesse caso, o melhor é socorreremo-nos do computador. E as/os donas/os de casa fazem-no, nos tempos que correm. Elas e eles sabem que introduzem rigor na forma de gerir o orçamento individual ou familiar.

Ora, é esse rigor que introduzimos com a utilização da Informática na construção do conhecimento em ciências sociais e humanas. O meu ponto de partida é a História, aquela habitual ciência dos “factos únicos e irrepetíveis”, como costuma ser pensada e referida.

Nada mais verdadeiro e nada mais falso igualmente. Se reduzirmos a História a “um acontecimento”, porque esse foi o “acontecimento que fez história” (do dia, do ano, do milénio), então, sim, ficaremos satisfeitos com o seu estudo qualitativo – e, mesmo assim, teremos tendência para o arrumarmos sincrónica e diacronicamente com o maior número de acontecimentos (ou factos) da mesma categoria dos tempos e espaços mais próximos. Só assim o tornamos para nós inteligível. Chamamos a isto a integração no contexto. Mesmo uma revolução – não falamos nós de “revoluções em cadeia” para nos referirmos às revoluções liberais e independentistas que deram origem ao mundo contemporâneo na Europa e na América? E, no entanto,



resolveremos nós o nosso problema dizendo que a revolução portuguesa de 1820 e a revolução brasileira do mesmo período são revoluções liberais da primeira fase do liberalismo? Sim, naquilo que as une. Não naquilo que as separa. Sobre um fundo comum (o liberalismo do período romântico) teremos necessidade de descobrir as particularidades de cada um dos processos revolucionários. O que é mais importante? Saber que houve cerca de meia centena de revoluções liberais de 1770 a 1830, ou saber que a revolução liberal portuguesa se resolveu por uma monarquia constitucional e a brasileira por um processo diferente, com Império e República? As duas coisas são importantes. A primeira para concluirmos sobre a existência de um “movimento avassalador” de mudanças políticas “quase inevitáveis” (aqui está algo que a análise histórica não permite, mas que a análise estatística permite pensar como provável). A segunda, para percebermos que, apesar das similitudes e das ideias partilhadas, cada



processo individual se distingue pela sua singularidade. E essa é a grande “imposição” exigida pela mudança - necessária! Sem singularidade, não há mudança - em História...como em Biologia. São fenómenos semelhantes no seu processo evolutivo.

Mas, e a análise estatística só nos será útil para o mundo contemporâneo, onde a estatística parece reduzir tudo a números e séries numéricas? Não, claro que não. É tão útil contarmos os milhares de processos judiciais montados pelos tribunais políticos dos regimes autoritários (através disso saberemos muito sobre a natureza do regime), como contar as “marcas” religiosas e sociais deixadas pelos peregrinos medievais nas suas deambulações – por elas seremos capazes de traçar os itinerários das suas viagens.

Daqui decorre a necessidade de pensar que a quantidade e a qualidade são dimensões indispensáveis do estudo do objecto social e histórico, ambas explicativas *per se*, mas também implicadas de forma dialéctica. O significado estatístico de um facto repetitivo ganha sentido pela quantidade de vezes que ocorre, mas só e na medida em que lhe atribuímos qualidades que o aproximam a todos os factos da mesma série. E sabemos nós, os cientistas sociais, como por vezes nos “custa” arrumar numa mesma categoria factos (ou objectos) complexos e diversos na sua singularidade. Para o fazer, normalmente, simplificamos e reduzimos esse facto ou acontecimento ao essencial. Objectivamos o que é objectivável (=repetitivo), mas perdemos muita da singularidade.

Daqui decorre uma das grandes dificuldades da construção de uma base de dados: a decisão sobre as variáveis que nos ajudam a construir o “facto estatístico”. Se forem em excesso, aproximamo-nos da diversidade infinita e corremos o risco de perder a “tendência essencial”, caindo na dispersão. Se, pelo contrário, for em défice, garantimos a obtenção de uma “tendência essencial”, mas empobrecemos a diversidade. É um problema, mas não irresolúvel, porque a Informática permite correcções posteriores ao preenchimento da base de dados: resumos ou exclusões, no caso de variáveis em excesso; acréscimos, no caso de variáveis em falta – este um problema mais difícil de resolver quando a decisão foi a de não recolher os dados na fase da consulta da informação. É por isso que, na recolha de dados, será sempre mais



vantajosa uma recolha mais vasta (mesmo excessiva), quer de informação “objectivável” em variáveis, quer de “observações” complementares de carácter mais qualitativo. Perde-se mais tempo na consulta e registo, ganha-se tempo e qualidade na análise e conclusão. Pense-se só – no caso da História – num daqueles arquivos de consulta difícil e

pouco acessível e na necessidade de o consultar mais do que uma vez...

Um assento desta natureza vai permitir-nos quantificar melhor e, ao mesmo tempo, fazer uma análise qualitativa da informação recolhida.

Desta vez vou dirigir-me, especialmente, para o campo das metodologias quantitativas.

2. DADOS

2.1. Recolha

Normalmente começamos pela construção de uma nova base de dados porque as que existem já feitas se não adequam às nossas necessidades. Nada de extraordinário! O nosso trabalho está muito simplificado, porque as bases que nos chegam já foram muito trabalhadas por especialistas em sofisticados programas informáticos. No fundo, é como se nos sentássemos calmamente a uma mesa de trabalho para desenhar as nossas velhas matrizes de duas entradas a que vamos acrescentando as variáveis que nos parecem quantificáveis a partir da nossa informação, e de acordo com os objectivos e metas que pretendemos atingir numa determinada fase do estudo.

E este, sim, é o verdadeiro trabalho do investigador: adequar o quadro da investigação aos objectivos pretendidos e às condições da informação disponível. Dir-se-á que, em História, os arquivos que consultamos nos reservam sempre “deliciosas” e imprevisíveis surpresas que apetece não perder. É verdade. Mas o investigador não é um administrativo que preenche tabelas de preços numa qualquer repartição pública! Em cada momento da sua

investigação, deverá estar em condições de julgar a forma como a base está a responder às suas necessidades ou não. E principalmente, no início, não poderá nunca começar a elaborar a sua base sem ter uma “ideia” clara sobre a natureza genérica do material que vai encontrar. Por isso, uma pesquisa prévia sobre a informação a consultar é sempre indispensável.

Durante a fase de consulta e registo, os dois domínios – a informação e o registo com fins de tratamento estatístico – deverão continuar a ser sempre as principais preocupações. Como historiadores, a nossa tendência é, normalmente, para nos embrenharmos nos documentos (que nos aguçam a curiosidade) e em esquecer a “qualidade” dos registos. Se isso acontecer muitas vezes, as correcções posteriores serão sempre mais trabalhosas. Podem chegar ao ponto da segunda consulta... todos temos experiência disso. Mas será vantajoso? Ao invés, se vestirmos a pele do “informático”, teremos tendência para registar o “essencial”, de modo a chegar o mais rapidamente ao fim. Especialmente se a nossa base de dados é muito volumosa (em registos e variáveis). Aqui perdemos a hipótese de irmos a fazer uma boa análise qualitativa dos dados. Já vimos como esta componente (sendo esclarecida pela quantidade) é fundamental para compreender a diversidade que existe por



dentro de cada série numérica. É assim como se, semicerrando os olhos, só vissemos uma enorme floresta de árvores todas iguais e monótonas. Basta afinar o olhar para ver que não são iguais essas árvores...

Estas “perturbações”, no trabalho de levantamento e registo de dados, podem ser sempre corrigidas a cada momento e, especialmente, em momentos de análise do trabalho realizado até aí. Da “qualidade” dos dados obtidos depende o rigor com que elaboramos as nossas análises e conclusões. Sem rigor, o resultado do estudo sobre uma base de dados pode ser duplamente negativo e um autêntico embuste – maior que as tradicionais análises qualitativas, feitas com o auxílio de instrumentos rudimentares. De facto, até podemos dar-lhe clareza nas formas de apresentação, ou fundamentar os dados em rigorosa análise matemática, que as conclusões não deixarão de ser falsas. A suposta “cientificidade” aprofunda o embuste. É por isso que, de tempos a tempos, se torna necessário “observar amostras” do



material recolhido. Elas nos darão a justa medida da correcção do nosso trabalho nessa fase.

2.2. Tratamento e análise interpretativa

Quem constrói uma base de dados, espera ansiosamente descobrir a “tendência central”, que parece reduzir tudo ao mais simples, à palavra genética que traduz “toda a verdade”. Este é um anseio justo, por vezes depois de um trabalho de recolha longo e difícil. Até porque essa

PARADOXO QUÂNTICO É OBSERVADO DIRECTAMENTE PELA PRIMEIRA VEZ

Na mecânica quântica, um campo na vanguarda da física onde a ciência muitas vezes se funde com a filosofia, grande parte do nosso conhecimento é baseado em probabilidades.

Um grupo de Investigadores no Japão acaba de levar um dos paradoxos fundamentais da mecânica quântica -

uma questão até agora meramente teórica - para o laboratório, para o âmbito da experimentação, e observar directamente algumas das mais “assombrosas questões da mecânica quântica.”

PARADOXO DE HARDY

O paradoxo de Hardy, o axioma segundo o qual não podemos fazer inferências sobre os acontecimentos

“palavra-chave” que tudo traduz, sendo a mais simples na forma, nos parece também a “grande verdade”. No entanto, deveremos precaver-nos para a ingenuidade de tais anseios. Basta pensar na exiguidade de algumas fontes arquivísticas, truncadas ou semi-destruídas para vermos até que ponto essa “verdade” não deixará de ser sempre contingente.

Para o historiador, sobrarão sempre a necessidade de complementar a informação recolhida com uma boa análise qualitativa. Especialmente quando os dados recolhidos são diminutos. Embora a complexidade não seja menor com as grandes séries numéricas – elas contêm, como vimos, tanto de unidade como de diversidade. Esta última apenas foi alienada, de forma consciente, para conseguir ver a floresta. Mas a realidade é a floresta e as árvores, cada uma delas...

No entanto, descobrir uma “tendência essencial” é de uma enorme importância, qualquer que seja a realidade que estudamos: política, económica, cultural, social ou material. Não explica tudo, apenas ajuda a explicar. Mas é fundamental para percebermos as continuidades, as repetições, os padrões comportamentais, etc.

Chega, portanto, a fase em que, muitas vezes a contragosto, temos que nos desfazer

do “acessório”. Em que chegamos à conclusão que escolhemos variáveis a mais, ou em que algumas delas só fazem sentido se associadas entre si. Por si só, pouco nos dizem.

A manipulação dos dados é, pois, essencial nesta fase. Precisamos de limitar variáveis, de resumir dados, por vezes, através do auxílio de programas de cálculo. As transferências desses dados, em ambiente *Windows* não são difíceis e permitem-nos simplificar dois elementos essenciais: o cruzamento das variáveis, com vista a uma análise mais rica e aprofundada (coisa que só a Informática nos permite fazer a níveis maiores de tamanho e complexidade) e a apresentação desses dados em formas de tabelas, gráficos, histogramas, etc. Portanto, legíveis de forma simples por nós e pelos outros, mesmo sem o auxílio de complicados textos.

Podemos ficar pela análise quantitativa dos elementos colhidos. Só por si ela é uma componente fundamental da explicação em História, especialmente em História serial, com aplicações práticas na História Económica e Social, mas também em todos os outros domínios da História. Porém, raramente nos restringimos a essa análise quantitativa,

passados que não tenham sido observados directamente, reconhecendo que o próprio acto da observação afecta a realidade que procuramos desvendar, coloca um dilema que os físicos quânticos têm procurado superar há décadas.

Como se pode observar na mecânica quântica, sistemas atómicos e subatómicos que ocorrem em escala tão pequena que não podem ser descritos em termos clássicos, se o próprio acto de olhar para eles pode mudá-los permanentemente?

Num artigo publicado na revista *New Journal of Physics*, Investigadores da Universidade Osaka, no Japão, explicam como utilizaram uma medida técnica que tem um impacto quase imperceptível sobre a experimentação que lhes permitiu compilar objectivamente resultados prováveis em escala subatómica.

EXPERIÊNCIA DE LUCIEN HARDY

O teste, baseado na experiência do pensamento de Lucien Hardy, que utilizou interferómetros - instrumentos que



até pela necessidade de contar com todos os dados recolhidos, mesmo os mais singulares, para proceder à interpretação das “tendências essenciais”: uma coisa são os dados estatísticos, as dominâncias e as médias que nos permitem traçar as linhas gerais de tendência; outra, a sua interpretação, a qual só poderá ser conseguida com o concurso de uma detalhada análise qualitativa. Daí a importância de todos os dados qualitativos obtidos durante a fase de recolha dos dados.

2.3. Apresentação da Informação

A apresentação da informação – em forma de texto, de gráficos, de tabelas ou de histogramas, entre outros formatos - permitirá sempre a sua utilização com diferentes objectivos, de acordo com a finalidade a que se destina. Porém, será vantajoso que a apresentação inicial tenha potencialidades para ser desdobrada em formas diferentes de comunicação – da comunicação para validação entre pares (em forma de artigo científico ou de conferência), até às formas de divulgação mais simplificadas, como acontece nos “sítios” ou Páginas da Internet ou em artigos de revistas de divulgação.

A regra é: uma boa apresentação – tanto no rigor como na clareza da forma - dos dados estudados poderá ter sempre diferentes adaptações, de acordo com os fins e utilidades.

3. CONCLUSÃO

A Informática é um imprescindível auxiliar de tratamento da informação, nesta “Idade” da informação abundante (quase excessiva), com importantes aplicações a períodos e temáticas sociais e históricas de períodos muito anteriores,

mesmo da “Idade” pré-estatística. Veja-se, a título de exemplo, o estudo estatístico dos artefactos com que se vai construindo a pré-história.

Com o seu adequado concurso, nunca o trabalho em ciências sociais e humanas esteve tão simplificado como hoje. Pela sua versatilidade de utilização, ficam simplificadas três tarefas fundamentais da investigação: a recolha e tratamento de informação; a análise, interpretação e apresentação dos estudos empreendidos; e, por fim, mas não menos relevante, o arquivamento e manipulação da informação com possibilidades contínuas de correcção, corte e acrescentamento dos dados recolhidos anteriormente. Por fim, mas também muito importante, através da Internet, a possibilidade de transporte e comunicação instantânea e a longas distâncias de todas as produções, com infinitas possibilidades de trabalho cooperativo entre parceiros de diferentes locais do Mundo.

Dito tudo isto, não convém deificar o instrumento: ele não pode prescindir de outros instrumentos como os procedimentos estatísticos ou o processamento correcto do texto final – não será o computador que vai ensinar o autor a escrever... E, principal e definitivamente, da capacidade única do cientista social (sociólogo, historiador, etc.) de imaginar problemas, hipóteses e explicações para os problemas que pretende resolver. ▀

BIBLIOGRAFIA (MUITO SUCINTA)

BRYMAN, ALAN E DUNCAN CRAMER, ANÁLISE DE DADOS EM CIÊNCIAS SOCIAIS, OEIRAS, CELTA, 2003

OLIVEIRA, MANUEL ARMANDO, INTRODUÇÃO À INFORMÁTICA PARA AS CIÊNCIAS SOCIAIS, UNIVERSIDADE ABERTA, LISBOA, 2001

PESTANA, MARIA E JOÃO CAGEIRO, ANÁLISE DE DADOS PARA AS CIÊNCIAS SOCIAIS: A COMPLEMENTARIDADE DO SPSS, LISBOA, SÍLABO, 2000