

George Welling hield op de NWD van 2006 een prachtig en gloedvol betoog over ‘computationele geschiedenis’. De aanwezigen waren het er meteen over eens: hier liggen de mogelijkheden om zeer authentiek, origineel en vernieuwend onderwijsmateriaal te maken voor het oprapen. Wie neemt deze uitdaging aan?

Geschiedenis en wiskunde: wat hebben die met elkaar te maken?

Inleiding

Toen ik in 1971 mijn eindexamen gymnasium alfa met een vette voldoende voor wiskunde afsloot, was ik ervan overtuigd dat ik me nooit meer in zou laten met iets dat ook maar in de verste verten leek op wiskunde. Meneer Klant was een goede leraar, maar de liefde voor het vak heeft hij mij nooit kunnen bijbrengen en meestal stond hij garant voor één van de onvoldoendes op mijn rapporten. Eigenlijk hoopte ik stiekem dat hij aanwezig zou zijn op de Nationale Wiskunde Dagen van 2006 om mij daar een praatje te zien houden. Ik heb hem ontmoet op een reünie van mijn oude school en weet dat hij nog steeds alles geïnteresseerd volgt. Na dat examen ben ik geschiedenis gaan studeren en dacht zo de garantie te hebben nooit meer met formules of statistiek geconfronteerd te worden. Ergens heeft er een wissel verkeerd gestaan, want het heeft allemaal anders uitgepakt.

De narratieven

Op het eerste gezicht lijken geschiedbeoefening en wiskunde niet al te veel met elkaar te maken te hebben, maar de oorzaak daarvan ligt hoofdzakelijk in het soort geschiedenis dat momenteel populair is. In de geschiedschrijving voeren momenteel de ‘narratieven’ de boventoon. Zij willen vooral het verleden weer tot leven laten komen in een mooi verhaal, waarbij subjectivisme geen enkel bezwaar is. Zo liggen de boekhandels vol met stapels biografieën en historische monografieën die zich vooral met ideeëngeschiedenis bezighouden. De historicus brengt ons zijn volkomen eigen visie op het verleden, natuurlijk wel onderbouwd met veel literatuurverwijzingen, maar geen tweede historicus zal ooit over hetzelfde onderwerp hetzelfde schrijven. Het is een eindeloze discussie en wie er gelijk heeft, hangt af van de bronnen en visies die je gebruikt.

Als je nu niet geïnteresseerd bent in de geschiedenis van grote mannen of vrouwen, maar meer in de ontwikkeling van de hele samenleving en vooral in de meer materiële aspecten daarvan, is er momenteel niet zoveel van je gading te vinden. Want als je je wilt verdiepen in groepen

mensen in het verleden, dan is het vrijwel niet te doen om de mensen die zo’n groep vormen allemaal individueel te volgen. Het zal nodig zijn om gegevens op een geaggregeerd niveau te bekijken. Dat wordt dus tellen, gemiddelden berekenen, kijken of er niet te veel gevallen zijn die het gemiddelde verpesten, met andere woorden eenvoudige statistiek. En daar moet je bij het merendeel van de huidige generatie historici niet mee aankomen.

De cliometristen

Dat is wel anders geweest. In de zestiger en zeventiger jaren van de vorige eeuw was er zelfs een groep historici, die zich Cliometristen noemden, in hun naam verwijzend naar de muze van de geschiedenis en het doen van metingen. Deze kwantitatieve historici zijn een beetje naar de periferie van het historisch métier verdreven, maar ze bestaan nog steeds en ik schaam me zelfs niet om mijzelf ook een kwantitatief historicus te noemen. Sterker nog: de tak van sport die ik beoefen, noem ik meestal ‘computationele geschiedenis’. Helaas is deze term, die ik in mijn proefschrift gelanceerd heb, nog niet zo erg aangeslagen. Wat ik ermee wil aanduiden, is die tak van de geschiedbeoefening die niet zonder computationele hulpmiddelen is te bedrijven. Daarbij kun je aan verschillende dingen denken. Met behulp van computers kunnen we tegenwoordig immense hoeveelheden gegevens analyseren. Dergelijke analyses waren vroeger of niet mogelijk, of niet binnen één mensenleven uit te voeren. Daarnaast kunnen we modellen maken met zoveel variabelen, dat mensen al snel het overzicht verliezen. Maar onze geduldige slaaf, de computer, is bereid alles voor ons door te rekenen, zolang hij maar zijn portie stroom krijgt. In de meeste wetenschappen is de computer wat vroeger doorgedrongen, maar zelfs in de geschiedwetenschap is er een kleine groep enthousiastelingen die zich internationaal hebben verenigd in de *Association for History and Computing*, die in Vlaanderen en Nederland een tak heeft, de Vereniging voor Geschiedenis en Informatica (VGI). Zowel de Association als de Vereniging hebben een groot probleem: ze preken hoofdzakelijk voor eigen parochie. Ze bestaan al enige decennia, maar het lukt niet echt om

de ‘mainstream’ historici te bereiken. Natuurlijk is er tegenwoordig geen weldenkende historicus meer die zonder een computer kan, maar de meeste gebruiken hem slechts als tekstverwerker en om informatie op het web mee te zoeken. Toen ik uitgenodigd werd voor de Wiskundedagen schoot het door me heen dat dit misschien een mogelijkheid was om de zaak andersom aan te pakken. Misschien kunnen we wiskundigen geïnteresseerd krijgen in historische gegevens. En daarom heb ik een verhaal gehouden over mijn eigen onderzoek en de gegevens waarop dat gebaseerd is. Waarom zou dat voor wiskundigen leuk of leerzaam kunnen zijn? Wel, in de discussie bleek al snel dat vele aanwezigen uit de onderwijswereld de mogelijkheden zagen van de dataset om allerhande statistische bewerkingen mee uit te leggen. Er is schijnbaar een schrikbarend gebrek aan interessante datasets om bewerkingen op los te laten. Mijn datasets zijn allemaal vrij toegankelijk en mogen voor ieder denkbaar doel gebruikt worden.

't Paalgeld

Observantie		van den		Ontfang van 't Paal-geld	
A		Gl.	St.	P.	
A.L. Schoonfe Aal	't laft van 12 ton	0	2	0	
Aluyn 't laft tegen	4000 pondt	0	5	0	
Aluyn 't vat tegen	1600 pondt	0	2	0	
Almagro of roo aarde de ton		0	0	8	
Amandelen de baal		0	1	0	
Amydon of Snyzel liet laft tegen	4000 pondt	0	2	0	
Anchovis de 100 tonnetjes		0	2	0	
Anijs de baal		0	1	0	
Appelen voor de tonne		0	0	8	
Arpays 't laft tegen 4000 pond		0	3	0	
as boras 't laft 4000 pondt		0	0	0	
afch Weed-aich 't laft van 12 tonnen		0	0	8	
afijn 't vat van 4 Orhooffen		0	3	0	

Alle droge vaten		Gl.	St.	P.
Duale het vat		0	2	8
E		0	2	8
Eweten 't laft		0	0	8
Ebbenhout 't laft tegen 4000 p.		0	5	0
F		0	0	8
Fijgen van de 100 korven		0	5	0
Fijgen 10 tonnetjes		0	1	0
Fijgen dodie Vygen 10 coppel		0	2	0
Fijcet hout 't laft tegen 4000 pond		0	2	8
Fufnyen 't vat		0	2	8
G		0	10	0
Gaaren het laft tegen 4000 pond		0	10	0
Gaaren 't vat		0	2	8
Ganckneppels 't groot hendert		0	6	0
Galnoten de baal		0	1	0
Garft 't laft		0	0	8
Geers 't laft van 12 tonnen		0	2	8
Gember de baal		0	1	0
Glas van de kift van Ooftern		0	0	4
Glas de choof of korff van 'Normandien		0	1	0
Glas van de waarde van 100 guld.		0	4	0
Gomme de baal of kas		0	1	0
Loffe gom het laft		0	10	0

fig. 1 bladzij uit 'Observantie van den Ontfang van 't Paalgeld

Wat heb ik dan precies onderzocht en wat stellen die datasets dan voor? Ik heb mij bezig gehouden met de achttiende-eeuwse handel van Amsterdam. Daarvoor heb ik een register van een belasting op inkomende scheepvaart van over zee in Amsterdam voor een beperkt aantal jaren bewerkt. Deze belasting heette het Paalgeld en de belastingregisters noemt men de *Havenboeken van de Heffing van het Paalgeld* te Amsterdam. Niet alle delen van deze registers zijn bewaard gebleven, maar we beschikken toch over een vrijwel volledige reeks vanaf 1771 tot 1832. Er is ook nog een register van het jaar 1742 overgebleven. De originelen van deze boeken zijn ondertussen goed opgeborgen, nadat er microfilms van gemaakt zijn die te raadplegen zijn in het Gemeente Archief Amsterdam en in het Streekarchief West-Friesland. Oorspronkelijk komen de boeken uit het stadsarchief van Enkhuizen, omdat deze stad het recht had verworven om deze belasting te heffen in alle steden aan de Zuiderzee. Als wederdienst had Enkhuizen de verplichting te zorgen

voor de bebakening en betonning van de Zuiderzee. Tot 1832 heeft Enkhuizen deze taak inderdaad uitgevoerd. Daarna heeft de centrale overheid dit overgenomen. Voor mijn proefschrift heb ik deze gegevens gedeeltelijk in een computerbestand ingevoerd. De boeken zelf zijn per jaar in twee delen onderverdeeld, waarvan het eerste deel de Europese handel betreft en het tweede deel de handel met het gebied dat onder het monopolie van de West-Indische Compagnie viel, grofweg de transatlantische handel. Van dit register heb ik de Europese handel ingevoerd voor de jaren 1742 en 1771 – 1786, voor de transatlantische handel ben ik doorgegaan tot 1817. Dat alles is ondergebracht in een grote database. In de tabel betreffende de Europese handel vind je 52094 schepen die in die periode in Amsterdam binnenvoeren en hun 104041 ladingen zijn ondergebracht in een daaraan gekoppelde tabel. De tabel betreffende de transatlantische handel bevat voor de hele periode 5737 scheepsbewegingen. Deze laatste tabel is verreweg het eenvoudigst. In sommige jaren hield de klerk precies bij wanneer de schepen in de haven arriveerden, maar in andere gevallen weten we alleen het jaar en de volgorde van binnenkomen. Van ieder schip is de naam van het schip genoteerd, de familienaam van de schipper, soms zijn voornaam, maar meestal zijn initialen, de haven van afvaart, en de hoeveelheid belasting die er betaald was, uitgedrukt in guldens en stuivers. Dat lijkt allemaal heel eenvoudig, maar de problemen die zich voordoen, zijn legio. Er was in die tijd geen sprake van standaardisering van namen, niet voor persoonsnamen, niet voor scheepsnamen, maar ook niet voor plaatsnamen. Ik heb geprobeerd om bij het invoeren van de gegevens zo dicht mogelijk bij de oorspronkelijke vorm te blijven. Met behulp van standaardisatietabellen heb ik later orde in de chaos trachten te scheppen.

In de paalgeld-database vinden we ook de gekoppelde tabellen van schepen en hun ladingen uit de Europese handel van 1742 en 1771 – 1786. Volgens de normalisatieregels voor relationele databases dien je herhalende groepen in een afzonderlijke tabel onder te brengen en aangezien één schip één of meer ladingen kan vervoeren, is er hier sprake van zo'n herhalende groep. Dat maakt voor het werken met de gegevens eigenlijk weinig uit, als de relatie tussen de tabellen maar goed is vastgelegd. In de database vinden we ook een aantal tabellen die nodig zijn om van allerhande vormen van namen te komen tot een voorkeursvorm, alsmede hulptabellen om tot categorieën te komen. Hierbij heb ik gebruik gemaakt van een coderingssysteem dat professor Johansson heeft ontworpen voor de digitalisering van de Deens Sonttolregisters en dat later ook gebruikt is door Jake Knopper in zijn studie over de Galjoetsgeldregisters. Daarnaast heb ik van alle voorkomende havens met behulp van een uitgave van de Amerikaanse overheid de afstand in zeemijlen naar Amsterdam berekend en aan de plaatsentabel toegevoegd¹.

Vervolgens kon de analyse beginnen en eigenlijk is die analyse niets anders dan een exercitie in beschrijvende

statistiek. Natuurlijk is het niet zo ingewikkeld om met behulp van SQL-queries allerlei totalen te berekenen, maar dat zegt allemaal nog niet zo bar veel. Het wordt veel leuker als je het relatieve aandeel in de handel van verschillende havens kunt bekijken. Hoe ligt het voor hout en hoe ligt het voor wijn of andere goederen? Daarbij doet zich dan het vervelende probleem voor dat de vermelding van de hoeveelheid goederen in vele vaak niet meer in gebruik zijnde maten is. Gelukkig is de heffingsvoet van het Paalgeld beschreven in de *Observantie van den Ontfang van 't Paalgeld*, zodat ik kon besluiten me niet druk te maken over al die verschillende maten. Als er in de *Observantie* staat dat je voor één zak anijs één stuiver belasting moest betalen, dan ging ik er vanuit dat als een schipper een gulden belasting voor anijs moest betalen, hij derhalve $\frac{100}{5} = 20$ balen anijs aan boord moest hebben, ook al staat er in de bron misschien wel zakken of kisten, of wat dan ook. Door van de belasting uit te gaan en dan terug te rekenen naar de maat zoals beschreven in de *Observantie*, voorkom je eindeloos omrekenen van de ene maat naar de andere. Waarschijnlijk zal er wel een afwijking in zitten, maar die heb ik voor lief genomen. Ook prettig van de *Observantie* is dat erin beschreven staat dat er voor iedere vijfentwintig gulden waarde van niet in de lijst opgenomen goederen één stuiver betaald moest worden, met andere woorden een 0,2% belasting *ad valorem*. En dan kan je grove schattingen maken van de waarde van de handel: vermenigvuldig de geheven belasting met 500 en je hebt de waarde van de goederen. Dat is natuurlijk niet precies, maar het is wel een goede basis voor een schatting. Het blijkt trouwens in de bron dat veel schippers hun lading naar waarde opgaven en niet naar gewicht, hetgeen een argument te meer is om uitgaande van de heffingen een schatting van de totale waarde te maken.

Het is hier niet de plaats om allerlei historische conclusies aan dit materiaal te verbinden, maar om erop te wijzen dat al mijn berekeningen te verifiëren zijn en misschien leuke statistische exercities voor het onderwijs zijn. Uit de gesprekken op de Nationale Wiskunde Dagen bleek mij dat er veel behoefte is aan dit soort materiaal, om te laten zien dat je met hoeveelheden gegevens die niet meer handmatig te verwerken zijn, statistische bewerkingen kunt doen die meer inzicht in de massa gegevens toestaan. Dat je bijvoorbeeld historische vragen zou kunnen bedenken en dan als opdracht geven hoe je dat met deze gegevens zou kunnen aantonen. Eén van de belangrijke zaken die je hiermee zou kunnen uitleggen, is het belang van het gebruik van de standaarddeviatie naast het gebruik van gemiddelden. Je kunt voor iedere havenplaats de gemiddelde waarde van de lading berekenen die daarvandaan kwam en dat lijkt een heel mooi resultaat. Maar als je dan de standaarddeviatie ernaast legt, dan blijkt het gemiddelde wel erg beïnvloed te zijn door de extreme waarden. Binnenkort zal ik ook een andere bron beschikbaar stellen, namelijk de Lastgelden van maart 1743 tot en met december 1748. Het betreft hier een ruime 13000 schepen

die in die periode in Amsterdam binnenvoeren. Het is een hele eenvoudige bron: in tegenstelling tot de *Registers van de Heffing van het Paalgeld* ontbreekt er hier een beschrijving van de lading, omdat dit lastgeld geheven werd naar rato van het bruto laadvermogen van het schip. De lading is dus niet vermeld, maar wel de omvang van het schip uitgedrukt in lasten, zodat we met behulp van deze bron ook iets kunnen zeggen over de omvang van de schepen. En ook dan weer komen er leuke dingen naar voren. Neem bijvoorbeeld de gemiddelde grootte van schepen afkomstig uit wat in de bron aangeduid wordt als 'Noorwegen'. De gemiddelde grootte over deze bijna vijf jaar is 96,8 last, maar in werkelijkheid is er geen enkel schip binnengekomen van 96 of 97 last.

Verrassend...

Als je naar de grafiek kijkt, zie je al snel dat er een heel groot aantal kleine schepen naar Amsterdam komt uit Noorwegen, maar ook vrij forse aantallen erg grote schepen van meer dan 100 last.

Als je met dergelijke bronnen analyses gaat maken, blijkt al snel dat het simpelweg tellen van de schepen die in Amsterdam zijn binnengekomen eigenlijk dwaasheid is, omdat je dan de kleinste scheepjes, die bijvoorbeeld uit de buurt van Hamburg en Bremen kwamen en niet meer dan acht tot twaalf last maten, gelijk stelt aan de schepen van meer dan 200 last die ook arriveerden. En dan hebben we het alleen nog maar over de afmetingen. Als we de waarde erbij gaan betrekken, die we met behulp van de Paalgeldregisters kunnen schatten, verandert het hele beeld weer, want het zijn niet altijd de grootste schepen die de kostbaarste lading vervoerden. Grote schepen waren bedoeld voor bulkgoederen, en die zijn meestal niet zo kostbaar.

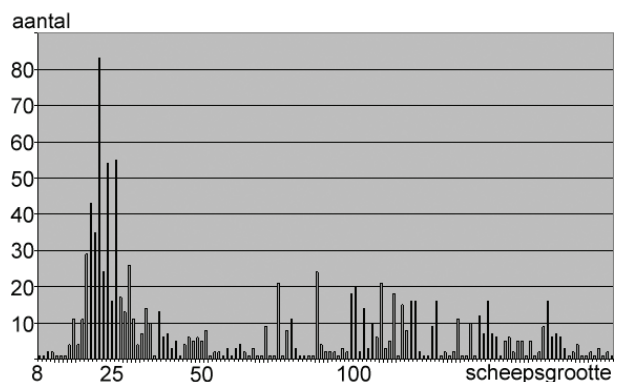


fig. 2 lastgeld maart 1743 – december 1748: schepen afkomstig uit Noorwegen (scheepsgrootte in lasten)

Als we de gegevens uit de database van de Paalgeldregisters gaan koppelen aan een Geografisch Informatie Systeem kunnen we heel aardig laten zien wat er aan de hand is. Een Geografisch Informatie Systeem biedt de mogelijkheid om resultaten van queries op de database te koppelen aan geografische eenheden, zodat geografische relaties direct inzichtelijk worden.

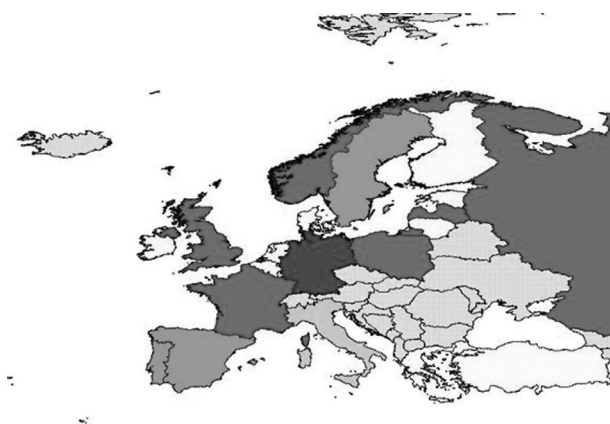


fig. 3 representatie van de aantallen schepen in 1773

Op het eerste plaatje zien we de aantallen schepen die in 1773 in Amsterdam binnenvoeren. De absolute aantallen zijn niet zo interessant: naarmate de kleur donkerder is, kwamen er meer schepen vandaan. (Helaas is dit afgebeeld op een moderne kaart, maar ik had geen andere voorhanden). Als we nu daarna kijken naar een afbeelding voor hetzelfde jaar waarin niet de aantallen schepen, maar de waarde van de lading de mate van donkerheid van de kleur bepaalt, dan zien we een verrassende verschuiving.

Frankrijk, Spanje, Portugal en Engeland blijken qua waarde veel hoger te liggen dan de bekende handelspartners uit de Oostzee, terwijl de Oostzeehandel toch altijd beschouwd werd als de ‘moedernegotie’. De handel met Portugal en Spanje, die qua aantal schepen niet al te veel voorstelt, blijkt in de waarde aan de top te zitten. Heel opvallend is dat Rusland – en in die tijd spreken we dan eigenlijk alleen over het gebied rond St. Petersburg en rond Archangel – zowel qua aantallen als qua waarde aan de bovenkant van het spectrum zit.

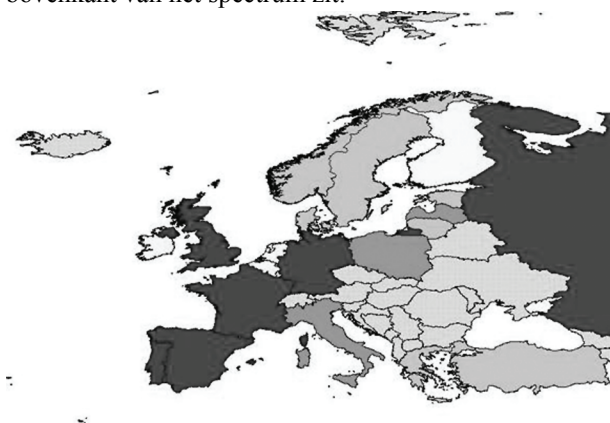


fig. 4 representatie van de waarde van de lading

Dit zijn maar een paar voorbeelden van het gebruik van deze dataset. Ik heb zelf seizoenspatronen voor de verschillende takken van handel berekend, de correlatie tussen de gevaren afstand en de waarde van de lading, de correlatie tussen de gevaren afstand en de afmetingen van de schepen, verdelingstestjes over de opbouw van de po-

pulatie, noem maar op. Het maken van een database met zoveel gegevens is een ontzettend karwei en het zou eigenlijk zonde zijn als het gebruik ervan beperkt zou blijven tot mijn statistische exercities om een aantal historische problemen in een ander daglicht te plaatsen. Met deze gegevens kun je laten zien dat alleen via de statistiek een bruikbaar inzicht in zulke grote hoeveelheden gegevens te krijgen is. Op de Wiskundedagen was er een aantal docenten aanwezig die direct veel vragen wisten te bedenken die hun leerlingen hiermee zouden moeten kunnen beantwoorden. Zodra ik het Lastgeld helemaal vrij geef, ben ik ook geïnteresseerd in de oplossing van het volgende probleem. De bron van het lastgeld begint pas in april van 1743. Wel heb ik beschikking over de gegevens uit de Paalgeldregisters van 1742 en die van het Lastegeld van 1744 – 1748. Wat zou nu de meest verantwoorde manier zijn om schattingen te maken voor die eerste drie maanden van 1743? Wie het weet, mag het zeggen.

Tot slot

Wiskunde en geschiedenis blijken soms meer met elkaar te maken te hebben dan voor de hand liggend lijkt. Vrijwel alle zaken in het leven hebben een getalsmatig aspect, zowel in de tegenwoordige als in de verleden tijd. In bepaalde takken van de geschiedwetenschap worden grote datasets ontwikkeld, die niet alleen voor historisch onderzoek bruikbaar zijn, maar die ook zeer goed gebruikt kunnen worden om bepaalde onderdelen van de statistiek mee te onderwijzen. Op die manier heeft ook de enorme tijdsinvestering die nodig is om zulke datasets tot stand te brengen een rendement dat maatschappelijk een stuk aanvaardbaarder is.

George Welling

Informatiekunde, Rijksuniversiteit Groningen

Noot

[1] Distances between Ports, Pub. 151, Prepared and published by the National Imagery and Mapping Agency, Bethesda, Maryland USA, Eleventh edition 2001.

Literatuur

George Welling (1998). *The Prize of Neutrality. Trade relations between Amsterdam and North America 1771-1817*. A study in computational history. Hilversum.

Alle hier genoemde bestanden zijn te vinden op: <http://www.let.rug.nl/~welling/paalgeld/appendix.htm>

Oproep redactie:

De redactie van de *Nieuwe Wiskrant* wil in een vervolgartikel graag de lesideeën, die ongetwijfeld met de datasets van George Welling gemaakt gaan worden, alsmede de oplossing van het probleem publiceren!