

HET MAKEN VAN THEMATISCHE KAARTEN MET KAART⁺ EN VU-STAT

Jelke Bethlehem *)

1. Inleiding

KAART⁺ is een programma voor het maken van thematische kaarten. Statistische gegevens kunnen op geografische kaarten worden weergegeven door middel van het kleuren of arceren van gebieden. Maar ook biedt het programma de mogelijkheid om de statistische informatie in de vorm van staafdiagrammen in de gebieden te plaatsen.

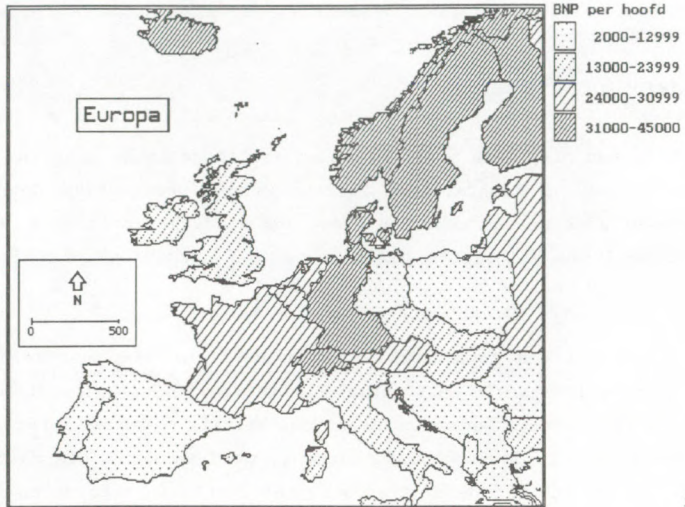
Voor het maken van thematische kaarten heeft u twee soorten informatie nodig: geografische informatie en statistische informatie. De geografische informatie is nodig om de kaart te kunnen tekenen. Het betreft met name informatie over grenzen en kusten. De statistische informatie is de informatie die moet worden geanalyseerd. Bij Kaart⁺ zit de statistische informatie in records en ieder record moet corresponderen met een gebied. De door KAART⁺ te verwerken statistische informatie moet zijn aangemaakt met het bijbehorende programma VU-STAT.

KAART⁺ wordt samen met VU-STAT in één pakket geleverd door uitgeverij Anaximander, Postbus 7432, 10007 JK Amsterdam. Als voorbeeld wordt de kaart van Europa met enige statistische informatie bijgeleverd. In figuur 1 staat een voorbeeld van een kaart die hiermee kan worden gemaakt.

*) Centraal Bureau voor de Statistiek, Hoofdafdeling Automatisering,
Postbus 959, 2270 AZ Voorburg

Het pakket draait onder MS-DOS en ondersteunt de grafische kaarten CGA, Hercules, EGA en VGA. Alleen op computers met een EGA- of VGA-kaart kunnen kaarten in kleur worden gemaakt. Het pakket kost fl 245,-, inclusief handleiding (KAART[†]: 41 pagina's, VU-STAT: 77 pagina's). De beoordeling van het pakket is gebaseerd op versie 1.01 van beide programma's.

Figuur 1. Het Bruto Nationaal Product in Europa



1

Het pakket beoogt een eenvoudige, gebruiksvriendelijk en goedkoop alternatief te bieden voor grotere en duurdere pakketten als ATLAS/GRAPHICS. Als doelgroep wordt met name het onderwijs gezien, maar het pakket is volgens de makers ook goed te gebruiken door onderzoeksbureaus. Een aantrekkelijke eigenschap van KAART[†] is dat u zelf op eenvoudige wijze kaarten kunt tekenen, of u kunt ze importeren vanuit het tekenprogramma PC Paintbrush of vanuit ATLAS/GRAPHICS.

VU-STAT is een simpel statistisch pakket. De beschikbare statistische technieken kunt u ook in vrijwel elk ander statistisch pakket treffen. Het aanbod is beperkt tot aan aantal eenvoudige, gangbare technieken. Het zal met name zijn gebruikers in het onderwijs moeten zoeken want voor het 'grote werk' zijn de mogelijkheden te beperkt. VU-STAT kunt u echter niet missen, want het programma is nodig om be-

standen met statistische gegevens in een voor KAART⁺ begrijpelijk formaat te kunnen zetten.

2. Installatie

Het pakket wordt, ingeval van 5.25 inch diskettes, geleverd op drie diskettes: een diskette met KAART⁺, een diskette met VU-STAT, en een diskette met de voorbeeldbestanden. Voor het draaien van het programma is een harde schijf niet noodzakelijk, al is dat natuurlijk wel sneller en eenvoudiger. Installatie van het pakket is zoals installatie behoort te zijn: simpel. Het is slechts een kwestie van kopiëren van bestanden. Het programma ontdekt zelf wel wat voor soort grafische kaart de computer heeft.

Voor het afdrukken van kaarten heeft het pakket de beschikking over een printer-aansturing voor Epson, NEC, Star, IBM en hiermee compatibele printers. Met het bijgeleverde programma *INSTALP.EXE* kunt u een printer-instelling kiezen. Helaas is er geen aansturing aanwezig is voor een laserprinter.

3. De handleiding

De handleiding is lekker klein, en daarom duurt het niet lang voordat u met het pakket aan het werk kunt. Aan KAART⁺ worden slechts 34 pagina's besteed en dat is zeker voldoende om al binnen korte tijd mooie kaarten te kunnen produceren. Dat dit gedeelte van de handleiding zo kort is, heeft zeker te maken met het feit dat de mogelijkheden van KAART⁺ beperkt zijn. Nadat de naam van het geografisch bestand en het statistisch bestand zijn opgegeven, worden de statistische variabelen gespecificeerd die moeten worden verwerkt tot kleuren, arceringen of staven, en klaar is kees!

Het is jammer is dat de handleiding weinig methodologische achtergronden bevat. Hoewel de auteurs ervoor waarschuwen dat 'het mogelijk is om met het programma volstrekt verantwoorde kaarten te maken', gaan ze hier niet verder op in maar verwijzen naar de literatuur, die lang niet iedereen binnen handbereik zal hebben. Dat is een gemiste kans. Je ziet al vaak genoeg kaarten waarin statistische informatie op een verkeerde manier wordt weergegeven. Met een paar eenvoudige regels zouden beginnende gebruikers al op weg kunnen worden geholpen. Een paar

voorbeelden. Voor een nominale variabele moet je geen arceringen met verschillende intensiteiten gebruiken, want dat wekt de suggestie van een ordening. Voor dit soort variabelen is het gebruik van verschillende kleuren aantrekkelijk. Omgekeerd moet je bij variabelen met een ordinaal of hoger meetniveau wel met intensiteiten werken en juist niet met kleuren. Evenzo zouden de auteurs aan kunnen geven welk type variabele zicht leent voor het weergeven in de vorm van een staafdiagram en welke voor het kleuren of arceren van gebieden.

Ook de methodologische achtergrond van VU-STAT is zeer summier. Elke beschikbare statistische techniek wordt in hooguit een paar zinnen beschreven. Formules worden alleen in een appendix gegeven en de berekende grootheden worden niet of nauwelijks uitgelegd. De auteurs merken terecht op dat VU-STAT eigenlijk meer een rekenprogramma is, maar geen cursus statistiek. Ze gaan er kennelijk vanuit dat de gebruiker zijn statistische kennis elders heeft opgedaan.

Afgezien van de hierboven gegeven kritische kanttekeningen, ziet de handleiding er verzorgd uit en is goed leesbaar.

4. KAART⁺

Na het opstarten van KAART⁺ vraagt het programma u eerst om de naam van het kaartbestand, het bestand met de geografische informatie. Om daar een beetje bij te helpen worden de namen van de beschikbare bestanden op het scherm getoond. Het programma is niet zo vriendelijk dat u simpel de cursor op de naam van je keuze kunt zetten; u moet de naam volledig intypen. Na het kaartbestand geselecteerd te hebben, moet het bestand met statistische informatie worden geselecteerd. Het programma suggereert een bestand met dezelfde naam als het kaartbestand. Is dat het goede bestand, dan hoeft u dit slechts te bevestigen met een druk op <Enter>. Moet aan ander bestand worden gebruikt, dan moet u weer een naam overtypen uit de getoonde lijst.

Na de specificatie van de bestanden, gaat het programma naar het hoofdmenu, dat slechts 5 opties kent:

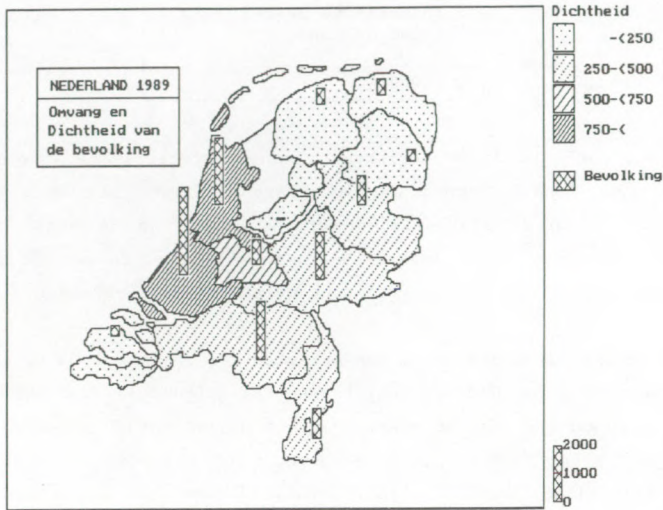
- | |
|----------------------------|
| 1 basiskaart tekenen |
| 2 gebiedskenmerken opgeven |
| 3 thematische kaart |
| 4 ander bestand |
| 0 stoppen |

Met de eerste optie wordt de basiskaart (zonder statistische informatie) op het scherm getekend. Zonodig kunt u de basiskaart wijzigen. De tweede optie is voor de specificatie van de statistische variabelen die op de kaart moeten worden weergegeven. De derde optie tekent de kaart volledig. En met de laatste optie kunt u overgaan naar andere (geografische of statische) bestanden.

Er zijn verschillende manieren om aan basiskaarten te komen. Om te beginnen wordt één basiskaart, die van Europa, meegeleverd (zie figuur 1). Ook kunt u een van de basiskaarten kopen die door de makers van het pakket worden geleverd. Het is echter ook mogelijk zelf basiskaarten te maken. Kaarten die met een tekenpakket zijn getekend, door ATLAS/GRAPHICS zijn gemaakt, of door een scanprogramma zijn gescand, kunnen worden ingelezen in KAART⁺ mits ze in het PCX-formaat zijn opgeslagen zoals dat wordt ondersteund door PC Paintbrush en een aantal andere programma's.

Een heel leuke optie is wat je zou kunnen noemen de 'poor mans digitizer': Neem een kaart over op doorzichtig folie, plak dit folie op de beeldbuis van de monitor en de neem de kaartgegevens over door in de teken-modus met de cursor de lijnen te volgen. Helaas ondersteunt het programma in deze teken-modus alleen cursor-besturing met de pijltjes-toetsen, terwijl het in deze situatie juist zo aardig zou kunnen zijn met een muis te kunnen werken. Maar iemand die de beschikking heeft over een muis en de bijbehorende software, kan de muisbesturing heel simpel buiten het programma om maken. En dan wordt het heel leuk om kaarten in te voeren. Binnen een uur had ik zo de kaart van Nederland, inclusief provinciegrenzen, in de computer. En kleine slorigheidjes bij het tekenen zijn geen probleem, want je kunt inzoomen en op eenvoudige wijze pixel voor pixel corrigeren. Vervolgens ken je namen toe aan de gebieden, en geef je (met de cursor) de plaatsen aan waar de staafdiagrammen moeten komen. Het resultaat van dit werk staat in figuur 2.

Figuur 2. Bevolkingsomvang en -dichtheid per provincie



1

Er dient wel te worden opgemerkt dat een kaart die op een computer met een bepaalde grafische kaart is aangemaakt, niet draait op een computer met een andere grafische kaart. Een kaart die u op een computer met Hercules-kaart heeft gemaakt, kunt u dus niet gebruiken op een machine met een EGA-kaart. Er is echter een hulpprogramma aanwezig dat voor u de conversie naar een ander type grafische kaart kan uitvoeren. Deze conversie wordt niet altijd helemaal vlekkeloos uitgevoerd. Soms wil er wel eens onderbreking voorkomen van de grenzen op de nieuwe kaart, maar dat kan snel worden opgelost door inzoomen en handmatige correctie.

Voor het weergeven van statistische informatie kunt u kiezen uit arcen/kleuren of staafdiagrammen. Alleen op computers met een EGA- of VGA-kaart kan gebruikt worden gemaakt van kleuren. Zoals al eerder is opgemerkt moet de statistische informatie zijn aangemaakt met het programma VU-STAT. Op het moment dat u in KAART^T de variabelen moet opgeven, kunt u door intypen van een vraagteken om een lijstje met beschikbare variabelen vragen. Het is niet erg handig dat de namen van variabelen maar 5 tekens lang mogen zijn. Dat geeft niet veel mogelijkheden om zinvolle namen met een mnemonische waarde te kiezen.

Voor arceren zijn er 10 verschillende arceringen beschikbaar, maar pas op: de volgorde waarin ze worden gepresenteerd is er niet een met een oplopende intensiteit. Arceringen 1 t/m 7 lopen redelijk op, maar dan komt 8 met de hoogste intensiteit, en 9 en 10 hebben weer lagere intensiteiten. Als je dus meer dan 7 arceringen gebruikt, loop je kans een kaart te produceren die verkeerd wordt geïnterpreteerd. De legenda-tekst die mag worden opgegeven bij de verschillende arceringen mag hooguit 12 tekens lang zijn, en bij veel toepassingen is dat veel te weinig. Continue variabelen zult u eerst in klassen moeten verdelen. U kunt dat zelf doen met de hand, maar ook kunt u de computer een suggestie laten doen. Als u in het programma VU-STAT al een klasseindeling hebt aangebracht, dan zal KAART⁺ die als suggestie overnemen.

Er kunnen maximaal zes kenmerken per gebied in de vorm van een staafdiagram worden weergegeven. Hoogte en breedte, kleur en arcering kunnen worden ingesteld. In principe zijn deze mogelijkheden voldoende om statistische informatie op een correcte wijze in een kaart weer te geven. U moet natuurlijk wel weten welk soort variabele u op welke wijze het beste kunt weergeven, maar dat vertelt de handleiding u niet.

Met de derde optie van het hoofdmenu kunt u het resultaat van uw activiteiten op het scherm bekijken. Bij een kaartbestand kunt u maximaal negen kaarten tegelijk maken. Indien uw computer dat toelaat, kunt u de kleuren op het scherm nog wat aanpassen. Uiteraard zult u een geproduceerde kaart ook willen afdrukken. De eenvoudigste mogelijkheid is een screen-dump. Dat lukte aardig toen het programma in de CGA-modus werkte, maar het ging fout in de EGA-modus. Een andere mogelijkheid is het commando <Alt P>, waarmee de kaart wordt afgedrukt met gebruikmaking van de 'eigen snelle printerdriver'. De hoogte/breedte-verhouding waarmee deze kaartjes worden afgedrukt, blijkt af te hangen van de grafische kaart van de computer. En de proporties zijn altijd anders dan die van de kaart op het scherm.

De op de matrixprinter afgedrukte kaartjes zien er redelijk uit, maar zijn niet van overweldigende kwaliteit. Wat node wordt gemist, is de mogelijkheid om de kaartjes af te drukken op een laserprinter. Alhoewel deze mogelijkheid niet is ingebouwd, is hij wel via een omweg te realiseren. Daarvoor heeft u een tekenprogramma of ander programma nodig waarmee u de op de monitor afgebeelde kaart in een bestand kunt opslaan. Dat kan bijvoorbeeld met het bij PC Paintbrush geleverde hulpprogramma FRIEZE, met het bij DR. HALO geleverde programma GRAB, of met het programma INSET (Inset Systems Inc). De landkaarten in deze bespreking zijn

met INSET verwerkt. Dit programma biedt de bijkomende voordelen dat u de hoogte-/breedte-verhouding van een plaatje kunt wijzigen, en dat u de plaatjes op simpele wijze in uw tekstverwerker kunt opnemen.

5. VU-STAT

VU-STAT is een eenvoudig statistisch programma, dat zijn belangrijkste verdienste ontleend aan het feit dat KAART⁺ alleen werkt met bestanden die door dit programma zijn aangemaakt. Na het opstarten van VU-STAT (dat op schijf ineens STAT blijkt te heten) verschijnt het hoofdmenu:

1. Opzet bestand aanmaken
2. Gegevens invoeren
3. Statistische verwerking
4. Gegevens bekijken
5. Extra mogelijkheden

Met de eerste optie kan een nieuw bestand worden aangemaakt. Bij het opzetten van een bestand moet voor elke variabele het type worden opgegeven. Type 1 is voor discrete variabelen waarbij de waarde uit één cijfer mag bestaan. Type 2 is voor numerieke variabelen. Daarbij moet het aantal cijfers achter de decimale punt worden opgegeven. Type 3 variabelen zijn bedoeld voor de situatie waarbij meer dan één antwoord is toegestaan. Per geval mogen maximaal 8 antwoorden opgegeven, waarbij elk antwoord een cijfer is. Namen van variabelen mogen maar 5 posities lang zijn, en dat zal vaak te kort zijn.

Aan de antwoordcodes van type 1 en type 3 variabelen kunnen labels worden toegekend. En bij variabelen van type 2 kan een verdeling in intervallen worden opgegeven, waarbij aan de intervallen ook weer labels kunnen worden gehecht. Labels mogen 10 posities lang zijn. De verdeling in intervallen is niet alleen noodzakelijk voor een aantal statistische technieken in VU-STAT, maar ook kan deze informatie worden overgenomen door KAART⁺.

Gegevens kunnen in record-vorm of in tabel-vorm worden ingevoerd. Het invoeren in record-vorm gaat wat moeizaam, want een <Enter> aan het einde van een record brengt u niet meteen naar het volgende record. Daarvoor is nog een extra toetsaanslag vereist. Ook moet u goed in de gaten houden dat het recordnummer als een

variabele wordt behandeld door het programma. Daar moet u dus steeds overheen springen. Het invoeren in tabel-vorm gaat efficiënter en zal daarom in de meeste gevallen de voorkeur verdienen. Het programma kent een aantal mogelijkheden om de ingevoerde gegevens te manipuleren. Records kunnen worden verwijderd, gegevens kunnen een vaste waarde krijgen, en nieuwe variabelen kunnen worden gemaakt uit bestaande met behulp van expressies. Het is ook mogelijk bestanden te importeren.

De optie statistische technieken geeft in eerste instantie de keuze uit de volgende mogelijkheden:

1. Frekwentietabel
2. Histogram
3. Cirkeldiagram
4. Gemiddelde, spreiding, mediaan, boxplot
5. Kruistabel
6. Lineaire regressie
7. Tijdreeks
8. Geavanceerde technieken

De frekwentietabel geeft aantallen en percentages. Doordat de onderstreping wordt gebruikt voor de lijnen in de tabel, is de vormgeving niet erg mooi. Bij het histogram worden kwalitatieve en kwantitatieve variabelen over één kam gescheerd. Kwalitatieve variabelen worden in een histogram weergegeven in plaats van een staafdiagram (met ruimte tussen de balken, en bij voorkeur met horizontale balken). Het cirkeldiagram is niet de meest aan te bevelen optie uit dit menu. Behalve dat je vraagtekens kunt zetten bij het gebruik van dit type grafiek, is de uitvoering ervan ook niet bijzonder fraai. Het is niet mogelijk de segmenten te kleuren of te arceren. In de boxplot wordt de mediaan op een wat onconventionele wijze weergegeven als klein vierkantje. De handleiding legt uit dat de boxplot zo handig is omdat hij ongevoelig is voor uitschieters. De formulering had beter kunnen luiden dat het juist zo'n goed instrument is om uitbijters te detecteren. In de kruistabel kunnen totalen of percentages worden weergegeven. Ook kan de chi-kwadraat-toetsingsgrootte worden bepaald. Het programma waarschuwt als het verwachte aantal in de cellen kleiner is dan vijf. De layout van de kruistabel is lelijk. Waarom nu ineens voor de lijnen in de tabel stippeltjes gebruiken? De optie 'Lineaire regressie' kan een simpele regressie met één onafhankelijke variabele uitvoeren. Het programma kan een spreidingsdiagram van de twee variabelen, inclusief regressielijn, produceren, en ook een residuenplot. Met de optie

'Tijdreeks' kan worden gezocht naar een additief of multiplicatief seizoenspatroon.

De optie 'Geavanceerde technieken' geeft ook nog toegang tot een menu met 'geavanceerde technieken'. Hiertoe behoren correlatiematrix, multipele regressie-analyse, discriminant-analyse, één- en tweestEEKproeventoetsen (Wilcoxon en Student), Kendall's Tau, Spearman's Rho en Lorentzcurve. Deze technieken worden op zeer summiere wijze afgehandeld. Vaak worden alleen toetsingsgrootheden uitgerekend, maar geen significanties, zodat je met de resultaten eigenlijk nog niets aan kunt.

6. Conclusies

KAART⁺ in combinatie met VU-STAT is een aardig pakket om snel en eenvoudig redelijke kaarten te maken. De user-interface is nog wat primitief, maar daar valt wel mee te leven. Alhoewel deze versie het pakket (1.01) zeker al goed bruikbaar is, zouden een aantal verbeteringen en aanvullingen het nog aantrekkelijker kunnen maken. Genoemd kunnen worden:

- Het selecteren van bestanden en variabelen via cursorbesturing, dus zonder de namen te hoeven intypen.
- De print-routine de kaarten in juiste proporties laten afdrukken.
- Een muis-aansturing voor de teken-routine.
- Meer methodologische achtergrond over kartografie in de handleiding.
- Langere teksten toestaan voor namen van variabelen en legenda's.