SOUTH AFRICAN POSTCODE GEOGRAPHY

Marisa Lombaard

Professional: GIS (Social Statistics)
Statistical Support and Informatics
Statistics South Africa
Email: marisal@statssa.gov.za

Abstract

In South Africa a postal code is a four-digit number that represents a certain area in the country. Each area is serviced by specific post offices, postal agencies, postal depots or mail collection points.

Postal codes might not receive a lot of attention in daily life, although we use them without even thinking. They are important as spatial units for storing and mapping certain kinds of thematic data such as geodemographic and lifestyle data for business purposes. Postal codes can be used as basic spatial data units in geographic information systems. They are part of the group of spatial units – along with administrative boundaries, place name boundaries, police districts, and enumeration areas - that form 'area footprints', making them receptacles for information. Without these, thematic data cannot be mapped. Therefore, spatial postcode creation and preserving previous versions is crucial.

INTRODUCTION

Postcode geography in Africa is characterised by a lack and unavailability of authorized spatial data. In South Africa postcode geography has never been used as an output form for demographic or census data. Postal codes (street and box) are readily available in text format, but no spatial set has been released by the South African Postal Office (SAPO).

The paper deals with the status and availability of South African spatial postal codes as points and areas, as well as the implications linking demographic information to the spatial sets. It will explain where postal codes fit into the current existing Statistics South Africa geographical frame, and the role it can play in a national address register.

BACKGROUND

In South Africa, a postal code is a four-digit code that represents an area. As in other counties, postal codes facilitate the national postal agency, in this case the South African Post Office (SAPO), to automate sorting and delivery of mail.

The four digit numeric number range starts from 0001 and stretches up to 9999. The ranges have been divided into postbox and street codes. Both these codes are linked to a place name, sorting centre (HUB) and province (Table 1). The provincial divisions do not correspond fully to the existing South African provincial boundaries but rather to post office distribution areas and regions (Fig 1).

Table 1 Post Office Sorting Lines (SAPO, 2004)

Province	HUB	Range	
		Start	End
Gauteng	Pretoria 1	0001	0204
Mpumalanga	Pretoria 2	0205	0698
Limpopo	Polokwane (Pietersburg)	0699	0999
Mpumalanga	Pretoria 3	1000	1199
	Nelspruit	1200	1399
Gauteng	Germiston	1400	1699
	Heidelberg	1438	1444
	Krugersdorp	1700	1799
	KDP/Soweto	1800	1870
	Vanderbijlpark	1871	1990
	Witspos (Johannesburg)	2000	2199
Mpumalanga	Pretoria 4	2200	2494
North West	Krugersdorp	2495	2519
	Potchefstroom	2520	2709
	Mafikeng	2710	2899
KwaZulu Natal	Ladysmith	2900	3199
	Pietermaritzburg	3200	3309
	Ladysmith	3310	3599

|			
	Durmail 2	3600	3799
	Richards Bay	3800	3990
	Durmail 1	3991	4179
	Port Shepstone	4180	4299
	Durmail 2	4300	4641
	Port Shepstone	4642	4730
Eastern Cape	Umtata	4735	4739
KwaZulu Natal	Port Shepstone	4740	4799
Eastern Cape	Umtata	4800	4899
	East London	4920	5049
	Umtata	5050	5199
	East London	5200	5750
	Port Elizabeth	5751	6499
Western Cape	George	6500	6699
	Worcester	6700	6899
	Beaufort West	6900	7099
	Cape Mail	7100	8179
Northern Cape	Upington	8180	8299
	Kimberley	8300	8799
	Upington	8800	8999
Free State	Bloemfontein 1	9300	9409
	Welkom	9410	9699
	Bloemfontein 2	9700	9999
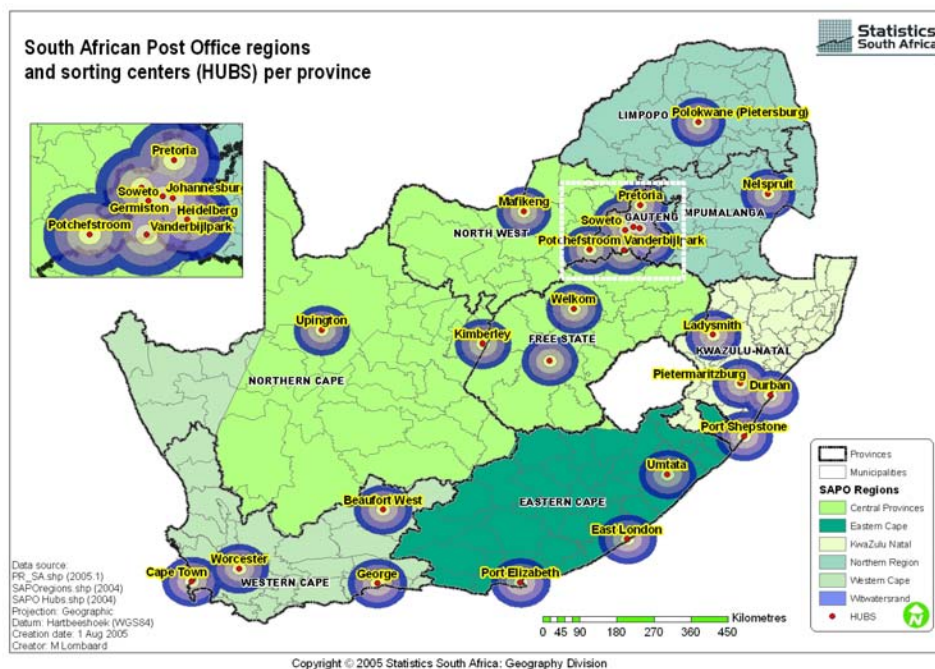
Figure 1. South African Post Office regions and sorting centers per province

POSTAL CODE HISTORY

A postal code is a series of letters and/or digits appended to a postal address for the purpose of sorting mail.

Germany was the world's first country with a postal code system in the early 1960s. The United States followed a couple of years later. The vast majority of the world's national postal services have postal code systems. A few do not: Ireland (previously only Dublin) and Hong Kong, for example, do not have postal codes, while New Zealand's postcode system is only used for the presorting of mail in bulk.

Although postal codes are usually assigned to geographical areas, sometimes this is not the case: special codes may be assigned to institutions with large volumes of post, such as government agencies and large commercial companies (Wikipedia, 2005).

Germany

Postal codes in Germany, known as *Postleitzahl* (pl. Postleitzahlen, acronym PLZ), consist of five digits, which indicate the wider area (first 2 digits), and the postal district (last 3 digits). The present system was introduced in 1993. Before reunification, both the Federal Republic of Germany (FRG) and the German

Democratic Republic (GDR) used 4 digit postal codes (e.g., the pre-1993 Frankfurt postal code was 6000 and afterwards 60000 – 60624) (Wikipedia, 2005).

United States of America
A ZIP Code is the postal code used by the United States Postal Service. *ZIP* is an acronym for 'Zoning Improvement Plan' and the basic code format consists of five numerical digits. The postal service implemented postal zones for large cities in 1943. In 1983, the US Postal Service began using an expanded ZIP Code system called "ZIP+4", which are often called "plus-four codes" or "add-on codes." A ZIP+4 code uses the basic 5-digit ZIP plus an additional 4-digits to identify a geographic segment within the 5-digit delivery area, such as a city block, a group of apartments, an individual high-volume receiver of mail, or any other unit that could use an extra identifier to aid in efficient mail sorting and delivery. For Post Office boxes, the general (but not invariable) rule is that each box has its own ZIP+4 code.

   ZIP Codes are numbered with the first digit representing a certain group of U.S. states, the second and third digits together representing a region in that group (or perhaps a large city), and the fourth and fifth digits representing more specific areas, such as small towns or regions of that city. The main town in a region (if applicable) often gets the first ZIP Codes for that region; afterwards, the numerical order often follows the alphabetical order of town names. Like area codes, ZIP Codes are sometimes divided and changed, especially when a rural area becomes suburban. Typically, the new ZIP codes become effective once announced, and a grace period (e.g., one year) is provided in which the new and old ZIP codes are used concurrently, so that postal patrons in the affected area can notify correspondents, order new stationery, etc. ZIP codes also change when postal boundaries are realigned (Wikipedia, 2005).

United Kingdom
UK Postal codes are alphanumeric. These codes were introduced by the Royal Mail over a fifteen year period from 1959 to 1974. They have been widely adopted not just for their original purpose of automating the sorting of mail but for many other purposes as well.

   The format of UK Postal codes is generally: LD DLL or LLD DLL or LDD DLL or LLDD DLL or LLDL DLL or LDL DLL, where L signifies a letter and D a digit. It is a hierarchical system, working from left to right - the first letter or pair of letters represents the area, the following digit or digits represent the district within that area, and so on. Each postcode generally represents a street, part of a street, or a single premises.

   The part of the code before the space is the *outward code* used to direct mail from one sorting office to the destination sorting office (the alphabetic part identifying one of 124 postal districts), while the part after the space is the *inward code* used to sort the mail into individual postmen's delivery rounds, each separate code usually identifying the address to within 80 properties, although large businesses may have a unique code. The letters in the outward code can only be one of a set of known combinations, which usually gives some clue to its geographical location.

   In the London area Postal codes are slightly different, being based on the old system of London postal districts which predated by many years the introduction of Postal codes in the 1960s. In central London, the letters WC and EC (West Central and East Central) and in the rest of London, N, NW, SW, SE, W and E are used.

   Until the 1960s, cities such as Belfast, Birmingham, Edinburgh, Glasgow, Leeds, Liverpool, Manchester, Newcastle upon Tyne and Sheffield were divided into different postal districts, each with a number (e.g., Toxteth in Liverpool was *Liverpool 8*). When the national postcode system was introduced, these were incorporated into it, so that Postal codes in Toxteth would start with *L8*, followed by the rest of the postcode. Northern Ireland was the last area of the UK to be postcoded, between 1970 and 1974. While Belfast was already divided into postal districts, rural areas, known as townlands, posed an additional problem, as many roads were not named, and houses were, similarly, not numbered. Consequently, many people living in such areas shared the same postal address, which still occurs in the Republic of Ireland (Wikipedia, 2005).

   Following the 1987 Chorley report, the 1991 British Census was to be based around postal geography. However, due to spiralling development costs, it was decided that the census area units be reliably linked to the census units used in 1981. The idea of basing census geography on Postal codes was shelved until the planning of the 2001 Census (Martin 1992). The growth of Geographical Information Systems (GIS) through the 1990's led to a dramatic increase in the use of digital geographic data. The popularity of the postcode as a spatial unit was reinforced by the Chorley report of 1987 and 'Postal codes the new geography' by Raper et al. (1992). This led to the reengineering of census output geography in the planning of the 2001 Census count.

   The 2001 Census of England and Wales was different to all previous undertakings of the population count; for the first time the collection and output of the Census were based on separate geographies. Data collected during the enumeration of the 2001 Census of England and Wales were for the first time stored at individual level rather than being accumulated into EDs (Enumeration Districts) and then stored, as had previously been the case.

   When the data are stored at the individual level they can be aggregated into many different spatial units, including - for the first time in the United Kingdom - postal geographies. Postal codes had no spatial

boundaries; therefore to enable the production of census data based on postal geography, postcode boundaries were to be created. They were formed by Thiessen polygons that were created around the centre of each address point using the 'Ordnance Survey Address Point™' (accurate to 1m), and afterwards the polygons of each address were merged within a postcode (Martin 1997 & 2000).

Australia

Australian Postal codes are numeric, consisting of four digits. They were introduced in the early 1970s by the Postmaster-General's Department (PMG), the predecessor of Australia Post. Postal codes are allocated to geographic areas to facilitate the efficient processing and delivery of mail to customers. The current four-digit numeric postcode system was introduced in 1967 in association with the first mechanised mail-processing centre in Australia.

The government land administration agencies in each state are responsible for gazetting locality names and boundary positions. Postal codes are only allocated to localities officially gazetted by state land agencies. In the majority of circumstances, a postcode covers an area comprising of more than one locality.

The decision as to whether a new postcode or an existing postcode is to be allocated to a locality is based on operational efficiency. Because the adoption of new or changed Postal codes by customers is slow, changes are only made where significant reasons for change are established. A postcode change will only be considered if such a change leads to either enhanced service to customers or operational efficiency to the organisation. Any such change will involve consultation with the local council/shire and residents (Australian Post, 2005).

In 1991, Australia Post teamed up with the Australian Surveying and Land Information Group (AUSLIG) with the goal to map and digitise the postcode boundaries of Australia and make them available for use by industry, commerce and government. To cover this vast continent with the high degree of accuracy required, over 2,000 base maps ranging in scale from 1:1,000,000 to 1:10,000 were chosen and distributed to Australia Post's post offices. The postal managers plotted the postcode boundaries onto the maps, which were then checked and handed back to AUSLIG to be digitised. By February 1992 AUSLIG had completed the process of digitising the postcode boundaries from the maps (Borrett, 1993).

People's Republic of China (PRC)

Postal codes in the People's Republic of China are six-digit numbers. The number is built from provincial capitals, municipalities and sub-provincial cities. For example, 100000–109999 is Beijing; 102800 is Hebei and 200000–209999 is Shanghai; 202400 is Zhejiang. Hong Kong and Macau do not follow this postal system.

India

Postal codes in India are also 6 digits, and they follow the countries administrative boundaries. These codes differentiate between 28 states and 6 union territories. For example, Gujarat (state), Ahmedabad (locality) with postal code 380001 compared to Arunachal Pradesh (state), Ramkrishna Mission (locality) with postal code 791113 (Wikipedia, 2005).

Africa

Out of a list (Pritchard, 2003) of 62 African countries postal codes could only be identified in 22 (35.5%) countries at the time this paper was written. None of these country's postal codes could be found to be available in an official spatial format.

Table 2. Postal codes in Africa (Pritchard, 2003)

Country	Postal codes	Country	Postal codes
Algeria	17791	Mozambique	407
Ascension Island	1	Niger	65
Cape Verde	989	Nigeria	44673
Ethiopia	38	Reunion	88
Guinea-Bissau	91	South Africa	9693
Kenya	1099	St Helena	13
Lesotho	2104	Sudan	173
Liberia	1	Swaziland	84
Madagascar	888	Tunisia	1426
Mayotte	22	Western Sahara	2
Morocco	1245	Zambia	9

Morroco
Morocco's postal codes also follow the administrative divisions and subdivisions of the country. The country is divided in 16 wilaya'at (states) that are also called economic regions. These states are subdivided into prefectures and provinces. The prefectures tend to be the more urbanized areas and the provinces are the more rural areas around smaller towns and localities.

Table 3 Example of Morrocan administrative subdivisions including Postal codes

Division	HASC	Postcode	Type	Capital
Fès	MA.FB.FS	14.231	perfectures	Fès
Boulemane	MA.FB.BO	14.131	province	Boulemane
Marrakech	MA.MK.MR	7.351	perfectures	Marrakech
Essaouira	MA.MK.ES	7.211	province	Essaouira

*HASC Hierarchical administrative subdivision codes

Madagascar
The Malagasy postal codes are a numeric 3 digit range based on 6 regions, (1) Antananarivo (2) Antsiranana (3) Fianarantsoa (4) Mahajanga (5) Toamasina (6) Toliara and the different localities forms these regions. For example, Antananarivo (region) Soavinandriana (locality) has the postal code 118 and Mahajanga (region) Tsaratanàna (locality) has the postal code 421.

Kenya
In Kenya no post gets delivered to a street address but rather to a post office box. The Kenyan postal code is a numeric 5 digit range based on post office (point). For example, Ngara Rd (Post office) has the postal code 0 0600 and North Kinangop (Post office) has the postal code 2 0154.

SOUTH AFRICAN POSTAL CODES

Spatial availability
Postal codes were introduced in South Africa in the mid 1970's, with introduction of the automated sorting. Currently postcode (street and box) information is only available in a booklet and text format from the South African Postal Office (SAPO). To date, no official spatial postcode data have been released by SAPO. It is clear that this information is not adequate for a spatial data environment. The result of this is that private data vendors create and sell these boundaries with added information, and they achieve this by using existing place names information, the National Address Dictionary (NAD) and sectional property schemes.

Postal code creation based on place name
The only way to create an existing postcode spatially is to capture a postman's walk/area either from paper maps or from GPS coordinates. It can also be created from existing spatial place name databases or other data sources, like NAD and recently captured traditional addresses.
   When creating it purely from place names the SAPO text file (Table 4) is needed. This file can be downloaded from the website http://www.sapo.co.za. It includes street and box Postal codes linked to town and place names. It is useful in getting the data in a workable format but the spatially friendliness of the set can be improved. The file consists of around 15 000 records.
   The disadvantage of using this file in creating a spatial link to place names is that the street code (polygon) and box code (point) need to be separated. Secondly, the place names and town names that link to these codes are names given by SAPO for their own purposes. A lot of duplication is evident within the SAPO text file, and not all the place and town names are recognized by other spatial place name databases.

Table 4 Example of SAPO postal code text file

Place name	Box code	Street code	Town name
ABBOTSFORD		2192	JOHANNESBURG
ABBOTSFORD		5241	EAST LONDON
KRUISFONTEIN	0200		ROSSLYN
KRUISFONTEIN	6306	6300	HUMANSDORP

When a test was done creating Postal codes from place names (polygons), and specific sub placenames (SP_SA.shp), the following was the process and outcomes (Fig 2).

Table 5 Spatial linking of street Postal codes – process and outcome

Process	Outcome
Data preparation Input SAPO text file Split box and street codes Link to SAPO province and HUBS Eliminate duplicate records (placename and province key)	Box codes without duplications Street codes without duplications

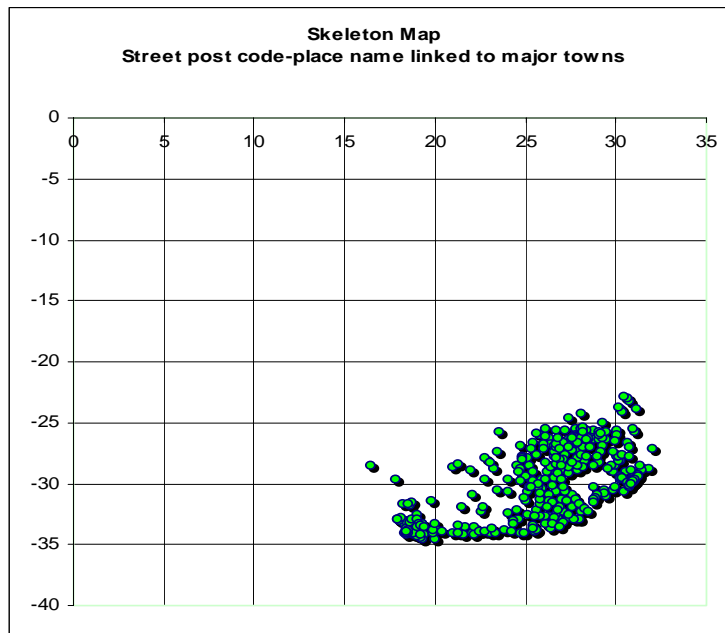| Spatial creation of street codes (polygon)	
Join street code file place names to existing spatial set polygons Dissolve spatial street code selection on code	14 000 records Unique 2424
Spatial creation of street codes (points)	400 unique
Join street code file place names to existing town points set	

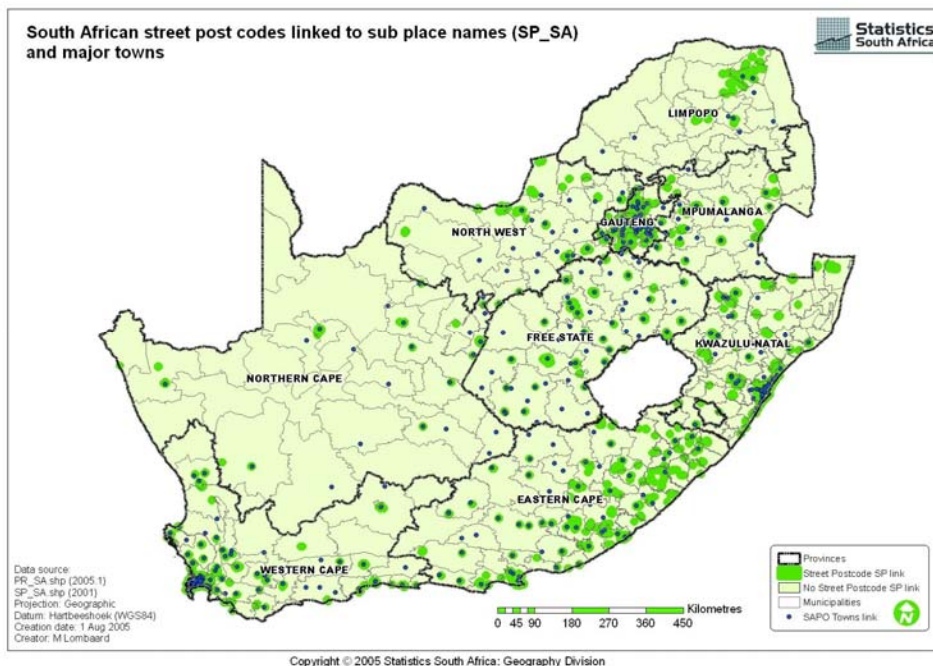Figure 2 Skeleton map of street post code /place name (SP) linked to major towns

Figure 3 Street Postal codes link to sub place names (SP_SA) and major towns

In the street code case a splintered layer was created (Fig 3). The reason for this is that not all farm, tribal and small holding areas get mail delivered to their door and therefore no street Postal codes exist in these areas, unlike in urban areas. In these cases the mail is delivered to postboxes with in nearby towns.

In some cases data vendors supply spatial coverage for Postal codes for the entire South Africa. Different spatial data sources were used for added quality. Postal codes were derived according to nearest town or place name in cases where Postal codes did not exist spatially, as in the case of farm or tribal areas.

Postal codes and the Statistics South Africa geographical frame

The current geographical frame is built from enumeration areas (EAs), place names (sub and main places), municipalities, district councils, and the nine provinces. Statistics South Africa uses all these geographical levels to aggregate data for different operational purposes. Boundary changes are being kept to the minimum but they do occur. If a postcode layer can be established it could become an area footprint over time and it will be slotted in between the main and sub place name level of the geographic frame.

Currently Postal codes can be linked to demographic data either using sub place names or the small area layer (SAL), through the use of the centroid or the percentage to area method. It must be kept in mind that the outcome of such methods will not be as accurate as data collected with the postcode as an output boundary.

The most difficult part of data aggregation currently for Postal codes are that they do not follow administrative boundaries, e.g. provincial boundaries. With an increase in data demand, specific demographic information at lower geographic levels, Postal codes could become the key to integrating diverse data sets in all spheres of South Africa.

Postal codes and a national address register

When a South African national address register concept is compiled together with Postal codes, the South African geography is lifted into a higher dimension; one which requires the South African Postal Office (SAPO) to work hand in hand with Statistics South Africa and data users.

During future censuses it will become necessary to distinguish between collection and output formats. A digital address base feeds into the unit data of a country. It can be used both in creating collection areas (EA) and output areas (postcode). None of the scales at which the census data are published are representative of any real world features; for example, the size of an EA represents the amount of ground a person can cover in the census timeframe of 21days. The census data are not published at an EA level. This is to ensure anonymity of individual census respondents. No information on individuals or their homes must be released singly or in a form where it can be interpreted from information for a specific area (Duke-Williams & Rees 1998).

Table 6 Future options for South African Postal codes

Options	Advantage	Disadvantages
Keep existing postcode format	No changes in SAPO work streams	Not spatial – needs digitizing from existing maps or from GPS co-ordinates of certain areas. Limited use by data users – lost opportunity Demographic data will always be derived from EA's and/or place names, influencing accuracy
Create new and/or transitional Postal codes **Prerequisite** *An official spatial postcode set has to be released and maintained*	Improved mail sorting and delivery Solve 4 digit postcode number shortages Forms a direct link between Postal codes and addresses Postcode as data output area more valuable and versatile	Postcode confusion – run dual system for a period

USES OF POSTAL CODES

Postal codes have obvious advantages over other output areas because they are widely used and understood. It can become the key to spatial strategic thinking and the ideal for data collection and distribution. A postcode area is large enough not to influence personal record confidentiality, and as a geographical unit it generally stable over time.

It is and can be effective in:
- Demographic operations (Statistics South Africa)
  - being an output and comparison unit during a Census and surveys
  - coding Business Register Records
  - quality measure on collecting survey questionnaires
- Geo Marketing - where data on population, economy and business can be linked and analysed in a three-dimensional way
- Property management

- Health
  - rare diseases research
  - hospital market share comparisons
- Optimum locality determination
- Emergency planning systems
- Community profiles
- Service delivery comparisons

The concept of using Postal codes as the key to unlock increased profits is applicable to most organisations, but few have explored this option. Most organisations don't think spatially and they don't really consider where their customers are located and the impact that this has on their organizational structure and customer service. Even fewer organisations consider how the location of existing customers can be used to predict where potential customers can be found.

A greater problem is that, until recently, even the organisations that think spatially come across many problems (Borrette,1993). In the case of South Africa and Africa the main problem is a lack of spatial data overall, especially postcode data.

SUMMARY

From the beginning Postal codes were created to assist in automated sorting and delivery of mail. South Africa's case is no different, and that is why Postal codes are based on the nearest major city rather than on an administrative area, causing postal boundaries to cross provincial and administrative boarders.

It would be incorrect to suggest that Postal codes themselves don't cause problems when used as a geographic base. In South Africa an official postcode set still have to be released although some sets do exist.

The use of Postal codes as part of the Stats SA geographical frame will result in a re-thinking of, or incorporation into, existing spatial formats. Postal codes will have to be incorporated as part of collection and output of data on a national level. Spatial address points will assist in forming a unit data source in developing collection and output geography formats.

Using Postal codes to link information will provide another spatial unit for storing and mapping different data such as geodemographic and lifestyle data for business purposes, as well as providing a more stable unit for time series comparisons.

Postal codes are widely recognized and used as a reference system and for that reason forms a useful geographic building block.

Postal codes can become the key to spatial strategic thinking and the ideal for data collection and distribution.

REFERENCES

1) Duke-Williams, O & Rees, P. (1998), Can Census Offices publish statistics for more than one small area geography? An analysis of the differencing problem in statistical disclosure, in The International Journal of Geographical Information Science, Vol. 12 No. 6 pp 579-605.
2) Martin, D. (1992), Postal codes and the 1991 Census of Population issues, problems and prospects, in Transactions of the Institute of British Geographers Vol. 17 350 -7
3) Martin, D. (1997) From enumeration districts to output areas: experiments in the automated creation of a census output geography, working paper No. 38
4) Martin, D. (2000), Towards the geographies of the 2001 UK Census of population, in the Transactions of the Institute of British Geographers Vol. 25 pp 321 - 332.
5) Martin, D. Nolan, A & Tranmer, M. (2001), The application of zone-design methodology in the 2001 UK Census, in Environment and Planning A, Vol. 33, pp 1949-1962, Pion.
6) Raper, J. Rhind, D. & Shepherd, J. (1992), Postal codes: The New Geography, Longman, Harlow.
7) SAPO (2004), 2004 Postal codes Alphabetical List
8) Vickers, D. (2003), The difficulty of linking two differently aggregated spatial datasets: Using a look-up table to link postal sectors and 1991 Census enumeration districts

Electronic References
9) Australian Post (2005), Assignment of Postal codes
10) Borrette, L. (1993), Profiting from Postal codes
11) Pritchard (2003), The Global Gazetteer
12) Wikipedia (2005), Wikipedia, the free encyclopedia