

Minimum geocoding match rates: an international study of the impact of data and areal unit sizes - Andreson

Il se base sur des données criminales, une variété d'unités spatiales d'analyse, provenant de divers pays et un test de configuration de points spatiaux à base locale pour trouver les niveaux de taux de correspondance nécessaires pour conserver les modèles spatiaux des données d'origine lorsque des adresses manquent.

Etudie l'impact de ne pas coder un morceau de données spatialement référencées.

Concentré que sur l'Australie, New South Wales.

Positional error in automated geocoding of residential addresses : Cayo and Talbot => précision de la position était meilleure dans les zones urbaines que dans les zones suburbaines ou rurales

Shah et al. 2014 ou Edwards et al(2014) => sur des études de criminalité, ils retrouvent plus d'information sur des unités de surface plus grandes => meilleur taux de correspondance et niveaux plus élevés de précision de positionnement.

Etude internationale !!

Dans cette étude, il semble que les résultats suggèrent que les niveaux nécessaires de réussite du géocodage sont inférieurs à ceux trouvés dans les recherches précédentes.

Géocodage minimum acceptable : 85% => résultat pertainent pour ceux qui font des analyses du crime sur des micro-lieux pour leur unité d'analyse (point ou segments de rues). Mais de façon générale, nécessite un minimum de 80%.

Ratcliffe, 2004 => 85% sur des données de crime.

Cependant, le minimum n'a pas besoin d'être aussi stricte : une fois que le ratio event/aire atteint 10, (1000 événements pour 100 unités d'aire par exemple), une correspondance de 50% est suffisante pour atteindre un S-Index \geq à 90%

Cette étude peut même atteindre des taux très bas qui peuvent être choquant, et donc à éviter.

An effective and efficient approach for manually improving geocoded data – Daniel Goldberg

Il n'y a pas juste la qualité du géocodage qui compte mais aussi l'exactitude => cette exactitude correspond à la qualité de la position du point dans l'espace.

Avantage d'employer une approche manuelle pour améliorer la qualité du géocodage ?

Possibilité d'utiliser un regéocodage provenant du web pour avoir un meilleur géocodage => utilisation principalement de la plateforme de géocodage USC.

Amélioration de la qualité du géocodage avec google maps => précision au niveau du centroïde de la rue et des intersections de rues.

Dans l'étude, passage d'une amélioration de géocodage de 83 à 86%.

Conclusion de l'étude : correction manuelle du géocodage est une solution intéressante mais se traduit par une augmentation des taux de correspondance + fonctionne bien sur petit et grands ensembles de donnée.

The Impact of Incomplete Geocoding on Small Area Population Estimates

Possibilité d'un passage d'un taux de confiance de 75 à 90% sur les zones urbaines à un taux de 40% voire moins sur les zones rurales.

Des modèles peuvent remédier à ce problème : modèle d'allocation de sous-zones, modèle d'automates cellulaires stochastiques, modèle basé sur la régression, etc.

Autre possibilité : manuel avec des images photo aériennes

“En outre, ils doivent savoir si les informations sont incomplètes le géocodage est associé à des facteurs démographiques et socio-économiques importants, conduisant peut-être à des estimations biaisées pour des groupes démographiques spécifiques. Maintenant, il n'existe aucune recherche publiée sur ces relations pour guider les démographes appliqués dans l'utilisation de données géoréférencées.”

L'étude suggère qu'un géocodage incomplet introduit potentiellement des erreurs dans les estimations de populations sur petites régions. Cependant, incertitude vis-à-vis de la qualité de ces estimations, puisqu'elles peuvent varier de 20 à 60%. Lunn et col => 10% en se basant sur une procédure de régression.