

SAMENVATTINGEN - RÉSUMÉS - SUMMARIES

De macht van het getal

SVEN VRIELINCK

Een statistische kritiek van de bevolkingsstatistieken in België (1801-1976)

Eén van de belangrijkste problemen waarmee het kritische onderzoek van overheidsstatistieken kampt, is de vraag naar de betrouwbaarheid van deze bronnen. Diverse studies wekken de indruk dat de kwaliteit van het cijfermateriaal in de loop der tijd is toegenomen, maar zekerheid en kennis over het verloop en de reikwijdte van dit proces ontbreken. Bij gebrek aan adequate technieken waarmee de cijfermassa kan worden geëvalueerd, blijft een zweem van twijfel hangen. In dit artikel onderzoeken we hoe de betrouwbaarheid van de meest elementaire tabellen uit de volkstellingen en de officiële statistieken van de bevolkingsbeweging doorheen de 19^{de} en het grootste deel van de 20^{ste} eeuw is geëvolueerd. Na een algemene beschrijving van de bronnen en van de kritische literatuur ter zake (delen I en II) vergelijken we bevolkingscijfers van gemeenten met de voorspelde verdeling volgens de Wet van Benford (deel III). Deze techniek blijkt doeltreffend te zijn om onregelmatigheden in het cijfermateriaal op te sporen voor zover de anomalieën een patroon van systematische afwijking vertonen. Uit de analyses blijkt dat de gedetecteerde bias in de datasets grotendeels mag toegeschreven worden aan fouten en gebreken bij de verwerking van individuele telresultaten. Deze verklaring is aantoonbaar op grond van een comparatieve studie van een aantal casussen. In grote lijnen toont deze studie een gestage verbetering van de datakwaliteit in de bestudeerde periode aan. De periode van het midden tot het voorlaatste decennium van de 19^{de} eeuw wordt gekenmerkt door een terugval in dit

proces. De neergang werd gestuit in de jaren 1880, toen de centrale overheid meer toezicht op de statistische productie in de gemeenten ging uitoefenen.

La force du nombre

SVEN VRIELINCK

Une analyse critique des statistiques de population en Belgique (1801-1976)

La question de la fiabilité des sources est l'un des principaux problèmes auxquels est confrontée la recherche critique relative aux statistiques des pouvoirs publics. Différentes études tendent à indiquer que la qualité du matériel chiffré s'est améliorée au fil du temps mais on manque de certitudes et de connaissances quant au déroulement et à la portée de ce processus.

En l'absence de techniques appropriées permettant l'évaluation de la masse chiffrée, il subsiste une pointe de doute. Dans cet article, nous aborderons la question de l'évolution de la fiabilité des tableaux les plus élémentaires issus des recensements et des statistiques officiels des mouvements de population au fil du 19^e siècle et durant la majeure partie du 20^e siècle. Après une présentation générale des sources et de la littérature critique en la matière (parties I et II), nous comparerons les chiffres de population des communes avec la répartition prévue sur base de la loi de Benford (partie III). Cette technique conserve son efficacité dès qu'il s'agit de détecter des irrégularités dans le matériel chiffré dans la mesure où les anomalies présentent un caractère d'aberration systématique.

Sur base des analyses, il apparaît que les écarts détectés dans les bases de données

peuvent en grande partie être imputés à des erreurs et à des manquements dans le traitement des comptages individuels. Cette explication peut être démontrée sur base d'une étude comparative d'une série de cas. De manière générale, cette étude démontre une amélioration progressive de la qualité des données au fil de la période envisagée.

La période qui va du milieu jusqu'à l'avant dernière décennie au 19^e siècle se caractérise par un recul du processus. Le déclin s'estompe dans les années 1880 lorsque les autorités centrales se sont mises à exercer une plus grande surveillance sur la production statistique des communes.

Force of Numbers

SVEN VRIELINCK

A Statistical Criticism of Population Statistics in Belgium (1801-1976)

One of the most important problems in critical research involving official statistics is the question of the reliability of the source. Studies suggest that the reliability of data has improved over time, but we lack absolute certainty and understanding of how this process occurs and develops.

In the absence of appropriate techniques for evaluating a large quantity of data, there must always be a certain amount of doubt about the reliability of the material. This article investigates how the reliability of the most basic tables from censuses and official statistics on demographic development evolved throughout the 19th and much of the 20th centuries. After a general presentation of the sources and relevant critical literature (parts I and II),

we will compare the population figures of municipalities with the projected distribution according to Benford's Law (part III). This technique proves to be effective in tracking data irregularities, as long as the anomalies exhibit a pattern of systematic deviation.

Our analysis shows that the detected bias in sets of data can be largely attributed to errors in processing of results of individual countings. This is demonstrated by comparing several examples. In general terms, this study shows a slow but gradual improvement of the quality of data gathered during this period, except for a relapse from the middle to penultimate decades of the 19th century. The decline was halted when the central government began to more closely scrutinise statistical collection in its municipalities.