

Assessing of Social Comfort of the Russian Population with Big Data

Authors : Marina Shakleina, Konstantin Shaklein, Stanislav Yakiro

Abstract : The digitalization of modern human life over the last decade has facilitated the acquisition, storage, and processing of data, which are used to detect changes in consumer preferences and to improve the internal efficiency of the production process. This emerging trend has attracted academic interest in the use of big data in research. The study focuses on modeling the social comfort of the Russian population for the period 2010-2021 using big data. Big data provides enormous opportunities for understanding human interactions at the scale of society with plenty of space and time dynamics. One of the most popular big data sources is Google Trends. The methodology for assessing social comfort using big data involves several steps: 1. 574 words were selected based on the Harvard IV-4 Dictionary adjusted to fit the reality of everyday Russian life. The set of keywords was further cleansed by excluding queries consisting of verbs and words with several lexical meanings. 2. Search queries were processed to ensure comparability of results: the transformation of data to a 10-point scale, elimination of popularity peaks, detrending, and deseasoning. The proposed methodology for keyword search and Google Trends processing was implemented in the form of a script in the Python programming language. 3. Block and summary integral indicators of social comfort were constructed using the first modified principal component resulting in weighting coefficients values of block components. According to the study, social comfort is described by 12 blocks: 'health', 'education', 'social support', 'financial situation', 'employment', 'housing', 'ethical norms', 'security', 'political stability', 'leisure', 'environment', 'infrastructure'. According to the model, the summary integral indicator increased by 54% and was 4.631 points; the average annual rate was 3.6%, which is higher than the rate of economic growth by 2.7 p.p. The value of the indicator describing social comfort in Russia is determined by 26% by 'social support', 24% by 'education', 12% by 'infrastructure', 10% by 'leisure', and the remaining 28% by others. Among 25% of the most popular searches, 85% are of negative nature and are mainly related to the blocks 'security', 'political stability', 'health', for example, 'crime rate', 'vulnerability'. Among the 25% most unpopular queries, 99% of the queries were positive and mostly related to the blocks 'ethical norms', 'education', 'employment', for example, 'social package', 'recycling'. In conclusion, the introduction of the latent category 'social comfort' into the scientific vocabulary deepens the theory of the quality of life of the population in terms of the study of the involvement of an individual in the society and expanding the subjective aspect of the measurements of various indicators. Integral assessment of social comfort demonstrates the overall picture of the development of the phenomenon over time and space and quantitatively evaluates ongoing socio-economic policy. The application of big data in the assessment of latent categories gives stable results, which opens up possibilities for their practical implementation.

Keywords : big data, Google trends, integral indicator, social comfort

Conference Title : ICQLP 2021 : International Conference on Quality of Life and Psychology

Conference Location : Istanbul, Türkiye

Conference Dates : September 27-28, 2021