



TRANSFORMASI SAMPLING DI ERA DIGITAL: VALIDITAS PENGGUNAAN BIG DATA SEBAGAI POPULASI DALAM PENELITIAN SOSIAL

Muammar Khaddafi¹ Naira Azzahra² Ardiana Rahmadania³ Devi Aulia Sari⁴
Azizah Muhshana Siregar⁵

¹⁻⁵ Jurusan Akuntansi, Fakultas Ekonomi dan Bisnis, Universitas Malikussaleh

*Penulis Korespondensi: khaddafi@unimal.ac.id¹ naira.230420045@mhs.unimal.ac.id²
ardiana.230420056@mhs.unimal.ac.id³ devi.230420047@mhs.unimal.ac.id⁴
azizah.230420135@mhs.unimal.ac.id⁵

Abstract. *The rapid advancement of digital technology has transformed data collection methods in social research. The emergence of Big Data enables researchers to access large-scale, diverse, and real-time data, leading to a shift from conventional sampling techniques toward the use of digital data as a representation of research populations. This study aims to analyze the transformation of sampling methods in the digital era and examine the validity of using Big Data as a population in social research. The study employs a descriptive qualitative approach through a literature review by analyzing recent scientific publications on Big Data, sampling methods, data validity, and social research methodology. The findings indicate that Big Data offers significant advantages in terms of data volume, variety, and timeliness, allowing researchers to capture social phenomena more comprehensively. However, challenges remain regarding data representativeness, selection bias, data quality, and privacy protection. Therefore, the use of Big Data as a research population requires appropriate methodological considerations to ensure the validity and reliability of research findings.*

Keywords: *Big Data, sampling transformation, population, social research, validity*

Abstrak. Perkembangan teknologi digital telah mengubah metode pengumpulan data dalam penelitian sosial. Kehadiran Big Data memungkinkan peneliti memperoleh data dalam jumlah yang sangat besar, beragam, dan tersedia secara real-time sehingga mendorong transformasi dari teknik sampling konvensional menuju pemanfaatan data digital sebagai representasi populasi penelitian. Penelitian ini bertujuan untuk menganalisis transformasi teknik sampling di era digital serta mengkaji validitas penggunaan Big Data sebagai populasi dalam penelitian sosial. Penelitian ini menggunakan pendekatan kualitatif deskriptif melalui studi literatur dengan menganalisis berbagai publikasi ilmiah terkini mengenai Big Data, teknik sampling, validitas data, dan metodologi penelitian sosial. Hasil penelitian menunjukkan bahwa Big Data memiliki keunggulan dalam menyediakan data yang berukuran besar, beragam, dan diperoleh secara cepat sehingga mampu menggambarkan fenomena sosial secara lebih komprehensif. Namun, penggunaan Big Data masih menghadapi berbagai tantangan, seperti representativitas data, bias seleksi, kualitas data, serta perlindungan privasi. Oleh karena itu, pemanfaatan Big Data sebagai populasi penelitian memerlukan pertimbangan metodologis yang tepat agar validitas dan reliabilitas hasil penelitian tetap terjaga.

Kata kunci: Big Data, transformasi sampling, populasi, penelitian sosial, validitas.

1. LATAR BELAKANG

Perkembangan teknologi informasi dan komunikasi telah membawa perubahan yang signifikan terhadap berbagai aspek kehidupan, termasuk dalam bidang penelitian sosial. Kemajuan teknologi digital telah menghasilkan data dalam jumlah yang sangat besar (Big Data) yang berasal dari berbagai sumber, seperti media sosial, mesin pencari, transaksi elektronik, aplikasi digital, serta berbagai platform daring lainnya. Karakteristik Big Data yang mencakup volume data yang besar, kecepatan pembaruan data, dan keragaman jenis data membuka peluang baru bagi peneliti untuk memahami berbagai fenomena sosial secara lebih komprehensif dan berbasis data.

Dalam metodologi penelitian sosial, teknik sampling merupakan salah satu tahapan penting yang bertujuan memperoleh sampel yang mampu merepresentasikan karakteristik populasi. Selama ini, pengambilan sampel umumnya dilakukan melalui teknik probability sampling maupun non-probability sampling, dengan mempertimbangkan prinsip representativitas agar hasil penelitian dapat digeneralisasikan terhadap populasi yang diteliti. Pendekatan tersebut telah menjadi dasar dalam berbagai penelitian sosial karena dinilai mampu menghasilkan data yang valid dan reliabel apabila diterapkan secara tepat.

Seiring dengan berkembangnya teknologi digital, paradigma pengumpulan data mengalami transformasi yang cukup mendasar. Kehadiran Big Data memungkinkan peneliti memperoleh data dalam jumlah yang sangat besar tanpa selalu bergantung pada proses pengambilan sampel sebagaimana diterapkan dalam penelitian konvensional. Kondisi ini memunculkan pergeseran paradigma dalam penelitian sosial, yaitu dari pendekatan berbasis sampel menuju pemanfaatan data digital berskala besar sebagai sumber informasi penelitian. Meskipun demikian, penggunaan Big Data sebagai representasi populasi masih menimbulkan berbagai perdebatan, terutama terkait representativitas data, potensi bias, validitas, serta reliabilitas hasil penelitian yang dihasilkan.

Berbagai penelitian terdahulu menunjukkan bahwa Big Data memiliki potensi yang besar dalam mendukung penelitian sosial karena mampu menyediakan informasi secara real-time dan menggambarkan perilaku masyarakat dalam skala yang luas. Namun

demikian, tidak seluruh data digital dapat dianggap mewakili populasi secara keseluruhan, mengingat adanya kemungkinan bias algoritma, bias seleksi, serta ketimpangan akses teknologi di berbagai kelompok masyarakat. Oleh karena itu, diperlukan kajian yang lebih mendalam mengenai validitas penggunaan Big Data sebagai populasi dalam penelitian sosial agar pemanfaatannya tetap memenuhi prinsip-prinsip metodologi penelitian yang ilmiah.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menganalisis transformasi teknik sampling di era digital serta mengkaji validitas penggunaan Big Data sebagai populasi dalam penelitian sosial. Hasil penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan metodologi penelitian sosial, khususnya dalam memahami peluang, tantangan, dan implikasi penggunaan Big Data sebagai sumber data penelitian.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif deskriptif dengan metode studi literatur (library research). Pendekatan ini dipilih untuk memperoleh pemahaman yang komprehensif mengenai transformasi teknik sampling di era digital serta mengkaji validitas penggunaan Big Data sebagai populasi dalam penelitian sosial berdasarkan berbagai kajian ilmiah yang telah dipublikasikan.

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari berbagai artikel ilmiah nasional dan internasional, buku referensi, serta publikasi ilmiah lainnya yang membahas Big Data, teknik sampling, validitas penelitian, representativitas data, dan metodologi penelitian sosial. Sumber-sumber literatur dipilih berdasarkan tingkat relevansi dengan topik penelitian, kredibilitas penerbit, serta tahun publikasi yang relatif mutakhir untuk memastikan informasi yang digunakan sesuai dengan perkembangan ilmu pengetahuan.

Teknik pengumpulan data dilakukan melalui proses penelusuran literatur menggunakan basis data ilmiah, seperti Google Scholar, Scopus, SpringerLink, ScienceDirect, dan Wiley Online Library. Literatur yang telah diperoleh kemudian diseleksi berdasarkan kesesuaian tema penelitian dan dianalisis secara sistematis.

Analisis data dilakukan dengan menggunakan teknik analisis isi (content analysis) melalui tahapan identifikasi, klasifikasi, evaluasi, dan sintesis terhadap berbagai hasil penelitian terdahulu. Tahapan tersebut bertujuan untuk mengidentifikasi perubahan paradigma teknik sampling di era digital, mengevaluasi kelebihan dan keterbatasan penggunaan Big Data sebagai populasi penelitian, serta menyusun kesimpulan mengenai validitas pemanfaatan Big Data dalam penelitian sosial.

3. HASIL DAN PEMBAHASAN

1. Transformasi Teknik Sampling di Era Digital

Perkembangan teknologi digital telah membawa perubahan yang signifikan terhadap teknik sampling dalam penelitian sosial. Jika pada penelitian konvensional peneliti harus menentukan sampel melalui teknik probability sampling maupun non-probability sampling, maka di era digital proses pengumpulan data dapat dilakukan melalui berbagai platform digital yang menghasilkan data dalam jumlah besar (Big Data). Perubahan ini memungkinkan peneliti memperoleh informasi secara lebih cepat, luas, dan berkelanjutan tanpa harus melakukan pengambilan sampel secara manual.

Transformasi tersebut menunjukkan bahwa teknik sampling tidak lagi hanya berfokus pada pemilihan sebagian anggota populasi, tetapi juga memanfaatkan data digital yang tersedia secara real-time. Meskipun demikian, teknik sampling konvensional tetap memiliki peran penting dalam memastikan representativitas data, sehingga penggunaan Big Data sebaiknya dipandang sebagai pelengkap, bukan pengganti metode sampling yang telah berkembang dalam metodologi penelitian sosial.

2. Big Data sebagai Sumber Populasi Penelitian

Big Data merupakan kumpulan data berukuran sangat besar yang dihasilkan dari berbagai aktivitas digital, seperti media sosial, transaksi elektronik, mesin pencari, sensor digital, dan aplikasi berbasis internet. Big Data memiliki lima karakteristik utama, yaitu Volume (jumlah data yang besar), Velocity (kecepatan data dihasilkan), Variety (keragaman jenis data), Veracity (keakuratan data), dan Value

(nilai atau manfaat data). Karakteristik tersebut menjadikan Big Data sebagai sumber informasi yang potensial dalam penelitian sosial.

Pemanfaatan Big Data memungkinkan peneliti mengamati perilaku masyarakat dalam skala yang lebih luas dibandingkan dengan metode survei konvensional. Namun, meskipun jumlah datanya sangat besar, tidak semua data digital dapat dianggap mewakili seluruh populasi karena terdapat perbedaan karakteristik pengguna teknologi di setiap kelompok masyarakat.

3. Validitas Penggunaan Big Data

Validitas merupakan aspek penting dalam penelitian karena berkaitan dengan kemampuan data dalam menggambarkan kondisi yang sebenarnya. Penggunaan Big Data dapat meningkatkan validitas penelitian apabila data yang digunakan relevan, akurat, dan sesuai dengan tujuan penelitian. Selain itu, jumlah data yang besar memungkinkan peneliti memperoleh gambaran fenomena sosial secara lebih komprehensif.

Meskipun demikian, ukuran data yang besar tidak selalu menjamin validitas hasil penelitian. Big Data masih memiliki keterbatasan, seperti data yang tidak lengkap, informasi yang tidak terverifikasi, dan adanya pengaruh algoritma platform digital. Oleh karena itu, peneliti perlu melakukan proses validasi data sebelum menarik kesimpulan penelitian.

4. Representativitas dan Bias Data

Salah satu tantangan utama dalam penggunaan Big Data adalah representativitas data. Tidak semua anggota masyarakat menggunakan media digital secara merata sehingga data yang diperoleh cenderung hanya mewakili kelompok tertentu. Kondisi ini dapat menimbulkan selection bias, yaitu bias akibat tidak semua individu memiliki peluang yang sama untuk menjadi bagian dari data penelitian.

Selain itu, algoritma pada platform digital juga dapat memengaruhi jenis informasi yang muncul sehingga berpotensi menghasilkan algorithmic bias. Oleh karena itu, peneliti perlu mempertimbangkan karakteristik sumber data dan melakukan evaluasi terhadap kemungkinan bias agar hasil penelitian tetap objektif.

5. Peluang Penggunaan Big Data

Pemanfaatan Big Data memberikan berbagai peluang bagi penelitian sosial, terutama dalam memahami perilaku masyarakat secara cepat dan berkelanjutan. Data digital memungkinkan peneliti menganalisis opini publik, pola konsumsi, mobilitas masyarakat, hingga tren sosial yang berkembang melalui media digital.

Selain itu, Big Data juga dapat dimanfaatkan sebagai dasar penyusunan kebijakan publik, evaluasi program pemerintah, dan pengambilan keputusan berbasis data. Dengan demikian, penggunaan Big Data mampu meningkatkan efisiensi penelitian sekaligus menghasilkan informasi yang lebih aktual dibandingkan metode konvensional.

6. Tantangan Metodologis

Di samping berbagai keunggulannya, penggunaan Big Data juga menghadapi tantangan metodologis yang cukup kompleks. Permasalahan utama meliputi perlindungan privasi, keamanan data, etika penelitian, serta kualitas data yang belum tentu sesuai dengan kebutuhan penelitian. Peneliti juga harus memastikan bahwa data yang digunakan diperoleh secara legal dan tetap menghormati hak privasi individu.

Selain itu, penggunaan Big Data memerlukan kemampuan analisis data yang memadai agar informasi yang diperoleh dapat diinterpretasikan secara tepat. Oleh karena itu, integrasi antara teknik sampling konvensional dan Big Data menjadi salah satu pendekatan yang dapat meningkatkan kualitas penelitian sosial.

7. Analisis Hasil Penelitian

Berdasarkan hasil kajian literatur, transformasi teknik sampling di era digital menunjukkan bahwa Big Data memberikan peluang baru dalam pengumpulan data penelitian sosial. Keunggulan Big Data terletak pada kemampuannya menyediakan data dalam jumlah besar, beragam, dan diperbarui secara real-time. Namun, pemanfaatannya masih memerlukan perhatian terhadap validitas, representativitas, serta potensi bias data.

Dengan demikian, Big Data tidak dapat sepenuhnya menggantikan teknik sampling konvensional. Sebaliknya, kombinasi antara metode sampling yang tepat dan pemanfaatan Big Data akan menghasilkan data penelitian yang lebih valid, reliabel, dan mampu memberikan gambaran yang lebih komprehensif mengenai fenomena sosial.

4. KESIMPULAN DAN SARAN

Berdasarkan hasil kajian yang telah dilakukan, transformasi teknik sampling di era digital menunjukkan adanya perubahan yang signifikan dalam metodologi penelitian sosial. Kehadiran Big Data memberikan peluang baru dalam proses pengumpulan data karena mampu menyediakan informasi dalam jumlah besar, beragam, dan diperbarui secara real-time. Pemanfaatan Big Data memungkinkan peneliti memperoleh gambaran fenomena sosial secara lebih luas serta mendukung proses analisis yang lebih cepat dibandingkan dengan pendekatan sampling konvensional.

Di sisi lain, penggunaan Big Data sebagai representasi populasi penelitian masih memerlukan perhatian terhadap aspek representativitas, potensi bias, kualitas data, serta validitas hasil penelitian. Atas dasar itu, pemanfaatan Big Data perlu didukung oleh penerapan prosedur metodologis yang tepat sehingga data yang diperoleh tetap valid, reliabel, dan dapat dipertanggungjawabkan secara ilmiah. Integrasi antara teknik sampling konvensional dan Big Data diharapkan mampu menghasilkan penelitian sosial yang lebih akurat, komprehensif, serta memberikan kontribusi bagi pengembangan ilmu pengetahuan dan penyusunan kebijakan berbasis bukti.

DAFTAR REFERENSI

- Kim, J. K., & Wang, Z. (2019). Sampling Techniques for Big Data Analysis. *International Statistical Review*, 87(S1), S177–S191. <https://doi.org/10.1111/insr.12290>
- Daikeler, J., Fröhling, L., Sen, I., Birkenmaier, L., Gummer, T., Schwalbach, J., Silber, H., Weiß, B., Weller, K., & Lechner, C. (2025). Assessing Data Quality in the Age of Digital Social Research: A Systematic Review. *Social Science Computer Review*, 43(5). <https://doi.org/10.1177/08944393241245395>
- Freese, J., & Jin, O. (2025). Online Nonprobability Samples. *Annual Review of Sociology*, 51, 109–128. <https://doi.org/10.1146/annurev-soc-090524-043117>

- Vicente, P. (2023). Sampling Twitter Users for Social Science Research: Evidence from a Systematic Review of the Literature. Quality & Quantity. <https://doi.org/10.1007/s11135-023-01615-w>
- Klinke, D., Jacobsen, J., Dieres, M., Schwander, H., von Scheve, C., & Specht, J. (2026). Social Media Sampling Is an Effective Way to Access Hard-to-Survey Populations and Low-Prevalence Groups. *International Journal of Social Research Methodology*, 29(3), 463–482. <https://doi.org/10.1080/13645579.2025.2564866>
- Pollard, M. S., Robbins, M. W., & Griswold, M. (2026). A Demonstration of Propensity-Score Weighting to Adjust a Social Media Nonprobability Sample Survey of Political Attitudes. *Public Opinion Quarterly*, 90(2), 536–565. <https://doi.org/10.1093/poq/nfaf071>
- Bergquist, P., et al. (2025). Survey Sampling in the Global South Using Facebook Advertisements. *Political Science Research and Methods*, 13(4), 781–797. <https://doi.org/10.1017/psrm.2025.18>