

**KAJIAN STATISTICAL DAN COST EFFICIENCY
DALAM PENENTUAN GUGUS SAMPEL BLOK SENSUS TERBAIK
(Studi Kasus: *Sampling Design* Susenas-2015 di Kabupaten Natuna)**

Wiwik Andriyani Lestari Ningsih¹, I Made Arcana²

¹Badan Pusat Statistik; ²Politeknik Statistika STIS

e-mail: wiwik.andriyani@bps.go.id, arcana@stis.ac.id

DOI: 10.14710/medstat.11.2.93-105

Article Info:

Received: 26 April 2018

Accepted: 12 February 2019

Available Online: 4 March 2019

Keywords: *sampling design, all possible samples, statistical efficiency, cost efficiency*

Abstract: Two aspects of efficiency that should be considered in applying sampling design of a survey are statistical efficiency and cost efficiency. Efficiency in statistical aspect improves precision of estimators obtained by the survey data, whereas efficiency in cost aspect provides an economic survey. The purpose of this research is to evaluate the both efficiencies in all possible census blocks (CBs) sample set and to identify the best CBs sample set in the 2015 National Socio-Economic Survey (Susenas). Therefore, a computer program for calculating statistical, and cost efficiency aspects was developed in this research to determine the best sample set of CBs among all possible sample set of CBs based on sampling design of the 2015 Susenas implemented in Natuna District, Kepulauan Riau Province. The best possible sample set of CBs is determined by considering statistical efficiency aspect, cost efficiency aspect, as well as combination of those two aspects. The result showed that the best sample set of CBs on statistical efficiency aspect provided the CBs sample set having minimum value of RSE index; evaluation on cost efficiency aspect provided the best CBs sample set having minimum value of total cost estimated using the total score of accessibility index; and evaluation on both efficiency aspects provided the best CBs sample set having minimum value of RSE index and minimum value of total score of accessibility index.

1. PENDAHULUAN

Sampling design merupakan tahapan penting dari metodologi survei yang mencakup keseluruhan aspek yang meliputi proses penentuan dan pemilihan sampel dalam suatu survei. Survei Sosial Ekonomi Nasional (Susenas) yang merupakan salah satu survey rumah tangga yang secara rutin dilakukan oleh Badan Pusat Statistik (BPS) menerapkan *two stages stratified sampling* sejak pelaksanaan Susenas 2015, dimana

stratifikasinya didasarkan pada kombinasi dari status wilayah (perkotaan dan perdesaan) dan tingkat kesejahteraan rumah tangga menggunakan pendekatan indeks kesejahteraan/*wealth index* (kesejahteraan tinggi, menengah, dan rendah).

Lehtonen dan Pahkinen (2004) mengungkapkan bahwa dalam suatu *sampling design* terdapat dua aspek efisiensi yang harus dipertimbangkan, yaitu *statistical efficiency* (efisien pada penduga yang dihasilkan) dan *cost efficiency* (efisien pada biaya pelaksanaan). Aspek *statistical efficiency* berhubungan dengan nilai presisi dari suatu penduga yang ditandai dengan nilai *sampling variance* yang minimum, sementara aspek *cost efficiency* berkaitan dengan biaya minimum ketika proses pengumpulan data di lapangan. Pada kenyataannya, kedua aspek tersebut bertolak belakang, artinya untuk memperoleh penduga dengan presisi yang baik (*statistical efficiency* tinggi) diperlukan jumlah sampel yang besar yang pada gilirannya berakibat pada meningkatnya biaya survei (*cost efficiency* rendah).

Pada tahapan pemilihan sampel blok sensus (BS) dalam *sampling design* Susenas, perlu dipertimbangkan aspek presisi estimasi indikator pokok dari setiap gugus sampel BS yang kemungkinan terpilih. Gugus sampel BS yang diharapkan adalah gugus sampel yang menunjukkan presisi tinggi, ditandai dengan nilai estimasi *relative standard error* (RSE) yang relatif kecil. Namun, adanya BS di daerah sulit yang termasuk dalam gugus sampel, berdampak pada tingkat efisiensi biaya operasional pengumpulan data di lapangan yang berpeluang munculnya permintaan pergantian sampel BS akibat beban pendataan yang lebih berat, waktu pendataan yang lebih lama, dan biaya yang lebih besar. Untuk itu gugus sampel BS yang terpilih diharapkan memuat relatif sedikit BS yang berlokasi di daerah yang aksesnya sulit, sehingga dapat mengefisienkan biaya operasional pengumpulan data di lapangan. Oleh karena itu, penelitian ini bertujuan untuk mengkaji *statistical efficiency* dan *cost efficiency* dalam penentuan gugus sampel BS yang paling efisien dengan melakukan evaluasi terhadap keseluruhan kandidat gugus sampel BS yang dapat dihasilkan dari *sampling design* Susenas 2015 dengan wilayah studi kasus di Kabupaten Natuna Provinsi Kepulauan Riau.

2. TINJAUAN PUSTAKA

2.1 *Statistical Efficiency* dalam Pemilihan Gugus Sampel

Department of Economic and Social Affairs Statistics Division (2005) menyatakan bahwa salah satu ukuran untuk mengukur tingkat presisi dalam suatu survei adalah *sampling variance* yang merupakan ukuran keragaman dari distribusi sampel suatu penduga yang dihasilkan. Evaluasi terhadap suatu penduga sering kali menggunakan ukuran *standard error*, yaitu akar kuadrat dari *sampling variance*. Nilai *standard error* digunakan untuk mengestimasi besarnya *sampling error* yang ditampilkan dalam tiga bentuk ukuran yang berbeda, yaitu nilai absolut *standard error*, *relative standard error* (RSE), dan *confidence intervals* (selang kepercayaan).

Salah satu metode penghitungan nilai *standard error* pada data hasil survei kompleks seperti Susenas menggunakan pendekatan metode linearisasi dengan mengaplikasikan metode perluasan *Taylor series*. Pendekatan ini secara luas digunakan karena dapat diterapkan pada desain sampel dan statistik secara umum yang merupakan fungsi linear dari statistik yang dihasilkan, seperti rata-rata atau total, dengan koefisien yang diperoleh dari proses diferensiasi pada komponen yang diperlukan.

2.2 *Cost Efficiency* dalam Pemilihan Gugus Sampel

Cochran (1977) mengungkapkan bahwa salah satu tujuan dari penerapan *sampling techniques* dalam suatu survei adalah agar tercapai efisiensi pada proses penarikan sampel, penduga yang dihasilkan maupun biaya yang dibutuhkan. Aspek *cost efficiency* berkaitan dengan kebutuhan biaya pelaksanaan survei yang besarnya dapat ditekan seminimal mungkin. Dalam penghitungan estimasi *cost efficiency* digunakan pendekatan indikator kategori akses, karena berdasarkan Litman (2015) bahwa *accessibility* berhubungan dengan kemudahan dalam meraih barang, jasa, aktivitas, dan tujuan, yang secara keseluruhan disebut kesempatan. Dengan kata lain, akses mengacu pada keseluruhan kemampuan seseorang untuk meraih layanan dan aktivitas, dan oleh karenanya waktu dan biaya merupakan komponen yang tidak terpisahkan.

Beberapa faktor umum yang mempengaruhi akses menurut Litman (2015) digunakan dalam penentuan dan pemilihan beberapa variabel yang terdapat dalam data Podes-2014, yaitu:

1. Kondisi perjalanan menggunakan kendaraan bermotor (*motor vehicle travel conditions*), meliputi kecepatan kendaraan, kemampuan, dan keselamatan dalam berkendara.
2. Kualitas mode pendukung lainnya dalam perjalanan (*quality of other modes*), meliputi berjalan, bersepeda, peristirahatan publik, kecepatan jasa antar, kesempatan, kenyamanan, kemampuan, dan keselamatan.
3. Koneksi jaringan transportasi (*transport network connectivity*), meliputi kepadatan jalan, koneksi jalan raya, jarak antar perjalanan, kualitas koneksi antar mode, seperti kenyamanan menuju stasiun transportasi publik.
4. Penggunaan lahan (*land use proximity*), meliputi pengembangan jalan dan penggabungan jalan.

Berdasarkan variabel-variabel yang telah ditentukan dengan referensi tersebut, selanjutnya dilakukan perhitungan indeks untuk menentukan tingkat kategori akses. Penghitungan indeks kategori akses dilakukan dengan menerapkan *Principal Component Analysis (PCA) Polychoric*, yang merupakan perluasan dari PCA yang dikembangkan sejak awal abad ke-20 oleh Pearson Hotelling pada tahun 1933 untuk data kategorik (Kolenikov dan Angeles, 2004).

3. METODE PENELITIAN

3.1. Sumber dan Metode Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang dihasilkan oleh Badan Pusat Statistik (BPS) dalam bentuk *master frame* Susenas 2015 di Kabupaten Natuna, Provinsi Kepulauan Riau sebagai dasar dalam pembentukan *all possible samples* BS untuk penghitungan penimbang; Sensus Penduduk (SP) 2010 untuk penghitungan RSE dan *difference*; serta Potensi desa (Podes) 2014 untuk dasar penghitungan indeks kategori akses. Struktur data yang digunakan secara lebih rinci dapat dilihat pada Lampiran 1.

Sampling design yang diterapkan pada Susenas 2015 untuk tingkat estimasi kabupaten/kota adalah *two stages stratified sampling* (Subdit Pengembangan Desain Sensus dan Survei BPS, 2015), dengan tahapan sebagai berikut:

Tahap 1:

- a. Memilih 25% BS dari populasi yang ada secara *probability proportional to size* (PPS) dengan ukuran jumlah rumah tangga hasil SP2010 di setiap strata BS, yang selanjutnya digunakan sebagai *sampling frame* dalam proses pemilihan BS.
- b. Memilih BS sejumlah n sesuai alokasi jumlah BS secara *systematic* di setiap strata. Pada proses pemilihan BS dibentuk semua kemungkinan gugus sampel (*all possible samples*) BS. Untuk Kabupaten Natuna, jumlah sampel BS yang telah ditentukan oleh BPS adalah sebanyak 44 BS.

Tahap 2: Memilih 10 rumah tangga berdasarkan hasil pemutakhiran pada BS terpilih secara *systematic sampling* dengan menerapkan *implicit stratification* berdasarkan pendidikan kepala rumah tangga.

3.2. Metode Analisis

Secara umum, terdapat tiga tahapan yang dilakukan dalam penelitian ini yaitu penyiapan informasi kategori akses, pembentukan *all possible samples* BS, dan penentuan *possible samples* BS terbaik. Penyiapan informasi kategori akses merupakan tahapan awal untuk melengkapi informasi pada setiap BS dalam *master frame* Susenas 2015 sehingga bisa dilakukan penentuan gugus sampel BS terbaik dari aspek *cost efficiency*. Informasi mengenai kategori akses tersebut dibentuk menggunakan beberapa variabel data Podes 2014. Pembentukan kategori tersebut dilakukan dengan tahapan sebagai berikut:

1. Melakukan penghitungan indeks kategori akses desa dengan pendekatan analisis *Polychoric Principal Component Analysis (Polychoric PCA)* melibatkan enam variabel kategori akses berdasarkan data Podes 2014 untuk menghitung nilai indeks kategori akses (*accessibility index*) untuk setiap desa dengan menerapkan rumus (1).

$$I_g = \alpha_{g1k} + \alpha_{g2k} + \dots + \alpha_{g6k} \quad (1)$$

Keterangan:

- I_g : indeks akses desa ke-g
 - α_{g1k} : skor untuk desa ke-g variabel pertama kategori ke-k
 - α_{g2k} : skor untuk desa ke-g variabel kedua kategori ke-k
 - α_{g6k} : skor untuk desa ke-g variabel keenam kategori ke-k
2. Membagi seluruh desa menjadi tiga kelompok kategori berdasarkan *cutting point* nilai kuartil yang dihitung berdasarkan data empiris nilai indeks kategori akses, yaitu
 - Kategori 1 : tingkat akses mudah (nilai indeks < Q1)
 - Kategori 2 : tingkat akses menengah (nilai indeks antara Q1 dan Q3)
 - Kategori 3 : tingkat akses sulit (nilai indeks > Q3)
 3. Menetapkan setiap kategori wilayah tersebut (1, 2, dan 3) masing-masing dengan skor/bobot 1, 2 dan 3, yang digunakan dalam melakukan estimasi biaya kategori akses pada setiap BS.

Metode sampling yang digunakan untuk membentuk seluruh kemungkinan gugus sampel BS tersebut sesuai dengan metode *sampling* Susenas 2015. Membentuk seluruh kemungkinan gugus sampel BS dalam penelitian ini membutuhkan data *sampling frame* BS yang diperoleh dari data *master frame* BS. Proses pemilihan sampel untuk menghasilkan *sampling frame* BS tersebut dilakukan BPS menggunakan metode sampling *PPS* dengan ukuran jumlah rumah tangga. Selanjutnya, berdasarkan *sampling frame* BS tersebut dilakukan proses pemilihan sampel dengan *systematic sampling* sehingga menghasilkan *all possible sample* BS per strata, yang kemudian dilakukan proses kombinasi yang melibatkan seluruh strata untuk memperoleh *all possible samples* BS.

Prosedur untuk menentukan gugus sampel BS terbaik dari aspek *statistical efficiency* didasarkan pada indikator yang tercakup dalam data SP 2010 (data individu dan rumah tangga) yang digabung dengan *master frame* BS, sehingga bisa dilakukan pemilihan beberapa variabel untuk indikator pendidikan (anggota rumah tangga) dan perumahan (rumah tangga). Data *master frame* BS yang sudah dilengkapi informasi strata serta beberapa variabel yang dibutuhkan, selanjutnya dikompilasi dengan data *all possible samples* yang telah terbentuk. Data *master frame* BS selanjutnya digunakan sebagai dasar untuk penghitungan persentase *difference* dan *RSE* pada setiap variabel yang telah ditentukan dengan menerapkan formula yang tercantum pada persamaan (2), (3), (4) dan (5). Persamaan tersebut merupakan hasil modifikasi dari rumus umum yang sudah ada (Sharon, 2010).

$$\text{persentase } difference = \frac{r-R}{R} \times 100\% \quad (2)$$

$$r = \frac{y}{x} = \frac{\sum_h \sum_i \sum_j w_{hi} y_{hij}}{\sum_h \sum_i \sum_j w_{hi} x_{hij}}, \quad w_{hi} = \frac{z_h}{z_{hi} n_h} \quad (3)$$

Keterangan:

- r : rasio untuk karakteristik yang diteliti berdasarkan data sampel
- R : rasio untuk karakteristik yang diteliti berdasarkan data populasi
- y : estimasi total untuk variabel Y
- x : estimasi total untuk variabel X
- y_{hij} : nilai karakteristik variabel Y pada rumah tangga ke- j , BS ke- i dan strata ke- h
- x_{hij} : nilai karakteristik variabel X pada rumah tangga ke- j , BS ke- i dan strata ke- h
- w_{hi} : penimbang sampel (*sampling weight*) strata ke- h pada BS ke- i
- z_{hi} : total rumah tangga pada BS ke- i , strata ke- h
- z_h : total rumah tangga pada strata ke- h
- n_h : jumlah sampel BS strata ke- h

$$\text{persentase } RSE = \frac{\sqrt{v(r)}}{r} \times 100\% \quad (4)$$

$$v(r) = \sum_h \left((1 - f_h) \frac{n_h}{n_h - 1} \sum_i \left(z_{hi} - \frac{z_h}{n_h} \right)^2 \right) \quad (5)$$

$$f_h = \frac{n_h}{N_h}, \quad z_{hi} = \frac{1}{x} (y_{hi} - r \cdot x_{hi}), \quad z_h = \sum_i z_{hi}$$

Keterangan:

- RSE : *relative standard error* untuk setiap variabel
- $v(r)$: *sampling variance* untuk rasio setiap variabel
- f_h : *sampling fraction* BS pada strata ke- h
- N_h : jumlah populasi BS pada strata ke- h
- y_{hi} : estimasi total variabel Y pada blok sensus ke- i strata ke- h
- x_{hi} : estimasi total variabel X pada blok sensus ke- i strata ke- h

Dengan menggunakan persentase nilai *RSE* yang telah dihitung, dilakukan penghitungan nilai indeks untuk setiap kombinasi indikator yang dipilih/ditentukan menggunakan rumus pada persamaan (6).

$$I_{rse}(i) = \frac{1}{p} \sum_{j=1}^p \frac{(RSE_{ij} - RSE_{j.min})}{(RSE_{j.max} - RSE_{j.min})} \quad (6)$$

Keterangan:

- $I_{rse(i)}$: nilai indeks *RSE* untuk gugus sampel ke-*i*
- p : banyaknya indikator yang dilibatkan
- RSE_{ij} : persentase nilai *RSE* gugus sampel ke-*i*, indikator ke-*j*
- $RSE_{j(min)}$: nilai minimum *RSE* pada indikator ke-*j*
- $RSE_{j(max)}$: nilai maksimum *RSE* pada indikator ke-*j*

Oleh karena dalam penelitian ini melibatkan empat indikator, maka akan terdapat 15 nilai indeks yang berbeda yang merupakan kombinasi dari empat indikator yang terlibat. Penghitungan indeks tersebut dimaksudkan untuk membuat agregasi dari beberapa indikator yang terlibat ke dalam satu nilai sehingga lebih mudah untuk diperbandingkan.

Prosedur untuk menentukan gugus sampel BS terbaik dari aspek *cost efficiency* membutuhkan data *master frame* BS yang disatukandengan data Podes 2014 yang sudah memuat informasi kategori akses sehingga terbentuk *master frame* BS yang dilengkapi informasi strata dan kategori akses. *Master frame* BS tersebut kemudian digabungkan dengan *all possible samples* yang telah dibentuk, yang siap untuk dihitung nilai estimasi biaya (*C*) untuk setiap *possible samples* BS yang terbentuk.

$$C = \frac{3}{2} \sum_{i=1}^3 (n_i \cdot c_i); \quad i = 1, 2, 3 \quad (7)$$

Keterangan:

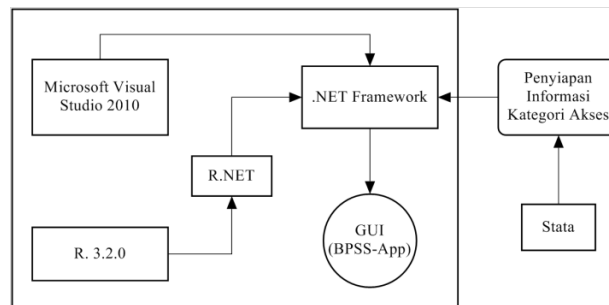
- C : estimasi total biaya menggunakan skor kategori akses
- n_i : banyaknya sampel BS dengan kategori akses ke-*i*
- c_i : estimasi biaya untuk kategori akses ke-*i*

Rumus (7) merupakan modifikasi dari bentuk umum dalam menghitung total biaya, yaitu $C = nC_1 + nmC_2$ (n : banyaknya sampel blok sensus; m : banyaknya sampel rumah tangga; C_1 : biaya antar blok sensus; C_2 : biaya dalam blok sensus). Aliaga dan Ren (2006) mengungkapkan bahwa secara umum perbandingan C_1/C_2 bernilai 20. Komponen C_1 dalam penelitian ini merupakan pendekatan untuk biaya antar BS (skor kategori akses).

Penentuan gugus sampel BS terbaik dari aspek *statistical efficiency* didasarkan pada indikator yang telah ditetapkan yang memiliki nilai indeks *RSE* terkecil. Sedangkan, penentuan gugus sampel BS terbaik dari aspek *cost efficiency* didasarkan pada nilai estimasi terkecil untuk total biaya dengan mempertimbangkan skor kategori akses (*C*) gugus sampel BS yang terbentuk. Dengan demikian, gugus sampel BS yang memuat BS dengan kategori akses mudah paling banyak akan ditetapkan sebagai gugus sampel BS terbaik. Lebih jauh lagi, penentuan gugus sampel BS terbaik untuk gabungan aspek *statistical* dan *cost efficiency* menggunakan nilai indeks *RSE* yang terkecil untuk gugus sampel BS tersebut, sekaligus juga menghasilkan estimasi total biaya yang relatif kecil pada skor kategori akses.

4. HASIL DAN PEMBAHASAN

Rancangan arsitektur aplikasi komputer yang dihasilkan untuk melakukan proses pemilihan gugus sampel BS terbaik ditampilkan pada Gambar 1 yang terdiri dari: a) *Microsoft Visual Studio* 2010 (C#) dengan *Framework.NET* yang digunakan untuk merancang *Graphical User Interface* (GUI) dari aplikasi yang dibuat, b) R 3.2.0 yang digunakan untuk membuat fungsi-fungsi penghitungan statistik dari aplikasi, c) *Connector R.NET* yang digunakan sebagai penghubung antara R dengan C#; dan d) *Stata* yang digunakan sebagai alat bantu dalam penyediaan informasi kategori akses.



Gambar 1 Rancangan Arsitektur Aplikasi

Informasi kategori akses atau variabel akses dalam penelitian ini diperoleh dengan menggunakan bantuan Stata yang *syntax*-nya dibuat sendiri menggunakan pendekatan *Polychoric PCA*. Hasil yang diperoleh bisa dilihat pada Gambar 1. Dari gambar tersebut dapat dilihat *polychoric correlation* antar variabel, *eigen value*, proporsi varians yang dijelaskan tiap *PCA* yang terbentuk menggunakan enam variabel yang digunakan (terdapat di Lampiran 1, serta *score* dari tiap kategori dari variabel yang digunakan. Dalam pembentukan indeks akses dibutuhkan nilai penimbang yang diperoleh dari nilai *score principal component* yang pertama dengan alasan bahwa *principal component* yang pertama mempunyai kontribusi terbesar dalam menjelaskan variabilitas data. Dari output tersebut, dapat ditunjukkan bahwa besarnya variabilitas yang bisa dijelaskan oleh *principal component* yang pertama relatif besar yaitu 52,44% dengan *eigen value* sebesar 3,15.

```

    . polychoricpca v1 v2 v3 v4 v5 v6, score(indeksAkses) nscore(1)
  
```

Polychoric correlation matrix						
	v1	v2	v3	v4	v5	v6
v1	1					
v2	.91359476	1				
v3	.43386417	.47676482	1			
v4	.42746235	.51367545	-.14414049	1		
v5	.26559704	.34889535	-.22098572	.97388098	1	
v6	.08189597	.07849141	-.21866716	.80480452	.78937408	1

Principal component analysis			
k	Eigenvalues	Proportion explained	Cum. explained
1	3.146473	0.524412	0.524412
2	2.060042	0.343340	0.867752
3	0.538487	0.089748	0.957500
4	0.199070	0.033178	0.990679
5	0.059622	0.009937	1.000616
6	-0.003694	-0.000616	1.000000

Scoring coefficients				
Variable	Coeff. 1	Coeff. 2	Coeff. 3	
v1	1	-0.278194	-0.376150	0.256746
	4	0.263989	0.356943	-0.243637
	6	0.842177	1.138718	-0.777249
v2	1	-0.239051	-0.295617	0.157159
	2	0.377278	0.466553	-0.248033
v3	1	-0.025594	-0.537528	-0.758253
	2	0.014056	0.295210	0.416432
	3	0.055205	1.159425	1.635519
v4	1	-0.560628	0.168555	0.016974
	2	-0.036505	0.010975	0.001105
	3	0.508974	-0.153025	-0.015410
v5	1	-0.481833	0.248027	-0.034857
	2	-0.008388	0.004318	-0.000607
	3	0.470228	-0.242053	0.034018
v6	1	-0.722179	0.628839	-0.772308
	2	-0.184026	0.160241	-0.196800
	3	0.387041	-0.337017	0.413907

Gambar 2 Output Saat Pembentukan Tambahan Data Podes-2014 Menggunakan Stata (Informasi Kategori Akses/Variabel Akses)

4.1. Gugus Sampel BS Terbaik dari Aspek Statistical Efficiency

Output yang dihasilkan dari aplikasi dalam menentukan gugus sampel BS terbaik dari aspek *statistical efficiency* adalah berupa daftar ID dari BS yang tercakup dalam lima gugus sampel BS dengan nilai indeks RSE terkecil untuk setiap kombinasi dari empat indikator yang terlibat dilengkapi dengan grafik distribusi dari *all possible samples* yang

dibentuk. Nilai indeks RSE pada lima gugus sampel BS terbaik dari aspek *statistical efficiency* untuk setiap kombinasi empat indikator yang terlibat, ditampilkan pada Tabel 1.

Tabel 1 Nilai Indeks RSE Lima Gugus Sampel BS Terbaik dari Aspek *Statistical Efficiency* Berdasarkan Indikator yang Ditentukan

No	Indikator	PS-1	PS-2	PS-3	PS-4	PS-5
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	R1	0,0000	0,0096	0,0129	0,0224	0,0233
2	R2	0,0000	0,0036	0,0139	0,0164	0,0211
3	R3	0,0000	0,0044	0,0272	0,0274	0,0290
4	R4	0,0000	0,0025	0,0088	0,0095	0,0121
5	R1, R2	0,1910	0,0254	0,0262	0,0317	0,0578
6	R1, R3	0,0267	0,0441	0,0510	0,0566	0,0684
7	R1, R4	0,2847	0,2895	0,2902	0,2930	0,2997
8	R2, R3	0,0356	0,0471	0,0471	0,0583	0,0614
9	R2, R4	0,0209	0,0286	0,0600	0,0619	0,0715
10	R3, R4	0,3810	0,3891	0,3938	0,3991	0,4071
11	R1, R2, R3	0,0400	0,0548	0,0587	0,0750	0,0830
12	R1, R2, R4	0,2005	0,2039	0,2106	0,2222	0,2260
13	R2, R3, R4	0,2606	0,2660	0,2762	0,2833	0,2878
14	R1, R3, R4	0,2703	0,2710	0,2835	0,2859	0,2870
15	R1, R2, R3, R4	0,2041	0,2126	0,2194	0,2222	0,2300

Perbandingan nilai RSE berdasarkan indikator aspek *statistical efficiency* dengan nilai RSE gugus sampel terpilih pada Susenas 2015 disertai informasi *ranking* dalam *all possible samples* ditampilkan pada Tabel 2. Terlihat bahwa nilai RSE untuk indikator terpilih yang dihasilkan pada Susenas 2015 masih berada dalam jangkauan nilai RSE, walaupun nilai RSE yang diberikan tidak minimum. Hal ini menunjukkan bahwa kualitas data hasil Susenas melalui pengendalian *sampling error* masih memungkinkan untuk ditingkatkan dengan memilih gugus sampel BS yang memiliki *statistical efficiency* yang tinggi.

Tabel 2 Perbandingan Nilai RSE Susenas Maret-2015 dengan *All Possible Samples* Aspek *Statistical Efficiency* Berdasarkan Empat Indikator

No	Indikator	Nilai RSE		
		<i>All Possible Samples</i>		Susenas Maret 2015
(1)	(2)	<i>Min</i>	<i>Max</i>	(Urutan)
(1)	(2)	(3)	(4)	(5)
1	R1	6,4274	9,3399	6,9017 (48)
2	R2	3,6426	5,0053	4,7183 (97)
3	R3	6,8616	10,1896	7,6529 (55)
4	R4	5,4101	8,3465	7,5362 (106)

4.2. Gugus Sampel BS Terbaik dari Aspek Cost Efficiency

Output aplikasi berupa daftar ID dari BS yang dilengkapi nilai total untuk kategori akses yang minimum sebagai estimasi biaya total pada lima gugus sampel BS terbaik dari aspek *cost efficiency* dapat dilihat pada Tabel 3. Dari tabel tersebut juga bisa dilihat berapa jumlah gugus sampel untuk tiap lima gugus sampel blok sensus terbaik yang memiliki nilai estimasi total cost sama (tabel 3 baris 2), yang memberikan peluang lebih besar untuk memilih gugus sampel BS yang digunakan dalam survei.

Tabel 3 Nilai Total Kategori Akses untuk Lima Gugus Sampel BS Terbaik dari Aspek *Cost Efficiency*

No	Keterangan	PS-1	PS-2	PS-3	PS-4	PS-5
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	Total Kategori Akses (Nilai estimasi total <i>cost</i>)	115,5	117	118,5	120	121,5
2	Jumlah <i>possible samples</i>	1	3	5	7	10

Keterangan: PS = *Possible Samples*

Perbandingan lima gugus sampel BS terbaik yang dihasilkan aplikasi terhadap gugus sampel BS yang terpilih pada Susenas 2015 ditampilkan pada Tabel 4. Terlihat bahwa gugus sampel BS Susenas 2015 tidak termasuk dalam lima gugus sampel BS terbaik dari aspek *cost efficiency* yang dihasilkan oleh aplikasi. Kondisi ini memperlihatkan bahwa *cost efficiency* gugus sampel BS pada Susenas dapat dioptimalkan dengan memilih gugus sampel BS melalui aplikasi yang dikembangkan dalam penelitian ini.

Tabel 4 Perbandingan Nilai Estimasi Total *Cost* Kategori Akses Susenas Maret-2015 dengan *All Possible Samples* Aspek *Cost Efficiency*

No	Keterangan	<i>All Possible Samples</i>		Susenas Maret-2015 (Urutan)
		<i>Min</i>	<i>Max</i>	
(1)	(2)	(3)	(4)	(5)
1	Nilai estimasi total <i>cost</i>	115,5	136,5	127,5 (9)
2	Jumlah <i>possible samples</i>	1	1	18

4.3. Gugus Sampel BS Terbaik untuk Kedua Aspek (*Statistical* dan *Cost Efficiency*)

Output berupa daftar ID dari BS dilengkapi nilai indeks RSE dan nilai estimasi biaya total (*cost*) untuk lima gugus sampel BS terbaik dari aspek kombinasi antara *statistical* dan *cost efficiency* ditampilkan pada Lampiran 2. Sementara, perbandingan dengan gugus sampel BS yang terpilih pada Susenas 2015 dapat dilihat pada Tabel 5. Gugus sampel BS yang dihasilkan dengan bantuan aplikasi menunjukkan tingkat efisiensi yang lebih tinggi dibandingkan gugus sampel BS yang terpilih pada Susenas 2015 untuk setiap indikator yang digunakan.

Tabel 5 Perbandingan Nilai *RSE* dan Estimasi *Cost* Susenas Maret-2015 dengan *All Possible Samples* Aspek Kombinasi antara *Statistical Efficiency* dan *Cost Efficiency*

No	Indikator	Nilai <i>RSE</i> (Estimasi <i>Cost</i>)						
		<i>All Possible Samples</i>				Susenas Maret-2015		
(1)	(2)	<i>Min</i>		<i>Max</i>		[Urutan]		
		(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	R1	6,4274	(123)	9,3399	(130,5)	6,9017	(127,5)	[48]
2	R2	3,6426	(135)	5,0053	(123)	4,7183	(127,5)	[97]
3	R3	6,8616	(123)	10,1896	(126)	7,6529	(127,5)	[55]
4	R4	5,4101	(130,5)	8,3465	(123)	7,5362	(127,5)	[106]

5. KESIMPULAN

Berdasarkan hasil yang telah diuraikan sebelumnya, kesimpulan yang diperoleh dari penelitian ini adalah sebagai berikut:

1. Aplikasi yang dikembangkan memberikan gugus sampel BS alternatif yang dapat meningkatkan efisiensi pelaksanaan survei, baik pada aspek *statistical efficiency* yang ditunjukkan dengan nilai indeks *RSE* yang minimum berdasarkan indikator yang dapat ditentukan oleh pengguna aplikasi, serta pada aspek *cost efficiency* yang ditunjukkan dengan nilai estimasi *cost* yang minimum berdasarkan kategori akses.
2. Evaluasi terhadap gugus sampel BS dalam Susenas 2015 memperlihatkan bahwa efisiensi hasil survei dapat ditingkatkan dengan pemilihan gugus sampel BS yang dapat mengoptimalkan *statistical* dan *cost efficiency*, yang pada gilirannya akan meningkat kualitas data hasil survei yang dilakukan.

DAFTAR PUSTAKA

- Aliaga, A. dan Ren, R. 2006. *Cluster Optimal Size for Demographic and Health Surveys.*: ORC Macro, Inc., United States.
- Badan Pusat Statistik (BPS). *Pendataan Potensi Desa (Podes) 2014 dan Sensus Penduduk 2010*. URL: <http://sirusa.bps.go.id>
- Subdit Pengembangan Desain Sensus dan Survei BPS. 2015. *Kumpulan Metodologi Sensus dan Survei Tahun 2015*. Badan Pusat Statistik, Jakarta.
- Cochran, W. G. 1977. *Sampling Techniques*. John Wiley & Sons, Ltd., England.
- Department of Economic and Social Affairs Statistics Division. 2005. *Designing Household Survey Samples: Practical Guidelines*. United Nations (UN), New York.
- Kolenikov, S. dan Angeles, G. 2014. *The Use of Discrete Data in Principal Component Analysis for Socio-Economic Status Evaluation*. URL: <http://tinyurl.com/nf9qkqq>
- Lehtonen, R. dan Pahkinen, E. 2004. *Practical Methods for Design and Analysis of Complex Surveys (2nd ed)*. John Wiley & Sons, Ltd., England.
- Litman, T. 2015. *Evaluating Accessibility for Transportation Planning*. URL: <http://www.vtpi.org/access.pdf>

United Nations Development Programme (UNDP). 2006. *ICT and Human Development: Towards Building a Composite Index for Asia*. Elsevier. New Delhi.

Sharon, L. L. 2010. *Sampling: Design and Analysis (2nd ed)*. Nelsen Education, Ltd., Canada.

Supranto, J. 2007. *Teknik Sampling: Untuk Survey dan Eksperimen*. Rineka Cipta, Jakarta.

Verma, V. 2013. *Sampling Elusive Populations: Applications to Studies of Child Labour*. International Labour Organization (ILO), Geneva.

Lampiran 1 Jenis dan Manfaat Data dalam Pembentukan *All Possible Samples* Blok Sensus serta Variabel yang Digunakan

No	Jenis Data	Manfaat dalam Pembentukan <i>All Possible Samples</i> Blok Sensus	Variabel	Keterangan Variabel
(1)	(2)	(3)	(4)	(5)
1	<i>Master Frame</i> Susenas-2015 (blok sensus) [type: .dbf]	Penghitungan penimbang (<i>sampling weight</i>)	BSC1_RT STRATA_SSN FRM180RB	Kode blok sensus Strata Susenas Blok sensus yang terpilih menjadi <i>sampling frame</i>
2	SP-2010 (<i>art</i> dan <i>ruta</i>) [type: .dbf]	Penghitungan <i>RSE</i> dan <i>difference</i> untuk evaluasi dari aspek <i>statistical efficiency</i> .	R213 (kode 3) R214A (kode 2) R404 (kode 5) R406 (kode 4)	Ijazah/STTB tertinggi: Tamat SLTP/MTs/ sederajat [R2] Tidak dapat membaca dan menulis huruf latin [R1] Bahan bakar utama untuk memasak sehari-hari: kayu [R4] Fasilitas tempat buang air besar: tidak ada [R3]
3	Podes-2014 (desa) [type: .dta]	Pembentukan kategori akses sebagai dasar penghitungan <i>cost</i> untuk evaluasi dari aspek <i>cost efficiency</i> .	R306A & R1002B_K2 R1002B_K7 R1001A R1001C1 R1001C2 R1001C3	Keberadaan kantor kepala desa/lurah & Jarak tempuhnya ke kantor bupati/ walikota. Biaya transportasi ke kantor bupati/ walikota. Lalu lintas dari dan ke desa/kelurahan. Keberadaan angkutan umum. Operasional angkutan umum yang utama. Jam operasi angkutan umum yang utama.

Sumber: Badan Pusat Statistik (BPS)

Lampiran 2. Nilai Indeks RSE dan Estimasi Total *Cost* Lima Gugus Sampel Blok Sensus Terbaik dari Aspek Kombinasi antara *Statistical* dan *Cost Efficiency*

No	Indikator	PS-1		PS-2		PS-3		PS-4		PS-5	
		<i>Stat</i>	<i>Cost</i>	<i>Stat</i>	<i>Cost</i>	<i>Stat</i>	<i>Cost</i>	<i>Stat</i>	<i>Cost</i>	<i>Stat</i>	<i>Cost</i>
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
1	R1	0,0000	123,0	0,0096	124,5	0,0129	129,0	0,0224	120,0	0,0233	130,5
2	R2	0,0000	135,0	0,0036	129,0	0,0139	133,5	0,0164	127,5	0,0211	133,5
3	R3	0,0000	123,0	0,0044	126,0	0,0272	121,5	0,0274	129,0	0,0290	124,5
4	R4	0,0000	130,5	0,0025	132,0	0,0088	126,0	0,0095	132,0	0,0121	133,5
5	R1, R2	0,1910	129,0	0,0254	127,5	0,0262	135,0	0,0317	133,5	0,0578	126,0
6	R1, R3	0,0267	126,0	0,0441	121,5	0,0510	132,0	0,0566	127,5	0,0684	127,5
7	R1, R4	0,2847	129,0	0,2895	123,0	0,2902	133,5	0,2930	127,5	0,2997	132,0
8	R2, R3	0,0356	126,0	0,0471	123,0	0,0471	132,0	0,0583	129,0	0,0614	124,5
9	R2, R4	0,0209	132,0	0,0286	130,5	0,0600	126,0	0,0619	133,5	0,0715	124,5
10	R3, R4	0,3810	126,0	0,3891	129,0	0,3938	132,0	0,3991	135,0	0,4071	127,5
11	R1, R2, R3	0,0400	126,0	0,0548	132,0	0,0587	129,0	0,0750	135,0	0,0830	124,5
12	R1, R2, R4	0,2005	133,5	0,2039	127,5	0,2106	132,0	0,2222	126,0	0,2260	135,0
13	R2, R3, R4	0,2606	129,0	0,2660	135,0	0,2762	126,0	0,2833	132,0	0,2878	127,5
14	R1, R3, R4	0,2703	126,0	0,2710	129,0	0,2835	135,0	0,2859	132,0	0,2870	124,5
15	R1, R2, R3, R4	0,2041	129,0	0,2126	135,0	0,2194	126,0	0,2222	127,5	0,2300	132,0

Keterangan: PS = *Possible Samples*; *Stat* = Nilai Indeks *RSE*; *Cost* = Nilai Estimasi Total *Cost*