

サンプリング方法の例

サンプリングには様々な方法がありますが、対象者の属性が特定の地域や年齢、性別、世帯主など、偏らないことが重要です。偏りが無い形で抽出する方法のひとつは、地域や年齢で第一優先順位、第二優先順位をつけて対象者を並べ替え、一定の間隔で抽出していく方法です。

B	C	D	E	F	G	H	I	J
被保険者番号	生年月日	性別	郵便番号	住所	小学校区	氏名	抽出用番号	
19640	19460103	1	4600001	〇〇町1丁目5	青山	〇〇××	1	✓
1587861	19440402	1	4600002	〇〇町2丁目8	青山	〇〇××	2	
15893	19440401	1	4600002	〇〇町2丁目3	青山	〇〇××	3	
175603	19420601	1	4600001	〇〇町1丁目44	青山	〇〇××	4	
159875	19401103	2	4600002	〇〇町2丁目25	青山	〇〇××	1	✓
1265789	19401101	2	4600002	〇〇町2丁目1	青山	〇〇××	2	
534981	19390301	2	4600001	〇〇町1丁目32	青山	〇〇××	3	
123578	19380202	1	4600001	〇〇町1丁目31	青山	〇〇××	4	
56875	19371002	2	4600002	〇〇町2丁目4	青山	〇〇××	1	✓
658789	19371001	2	4600002	〇〇町2丁目32	青山	〇〇××	2	
45646	19340802	2	4600002	〇〇町2丁目17	青山	〇〇××	3	
389156	19340801	2	4600001	〇〇町1丁目26	青山	〇〇××	4	
4568789	19400201	2	4600003	△町2丁目6	赤川	〇〇××	1	✓
789156	19390202	1	4600003	△町2丁目12	赤川	〇〇××	2	
7895132	19331002	2	4600003	△町2丁目10	赤川	〇〇××	3	
1357892	19320401	1	4600003	△町1丁目5	赤川	〇〇××	4	
55568	19300101	2	4600003	△町1丁目2	赤川	〇〇××	1	✓
578084	19280201	2	4600003	△町1丁目33	赤川	〇〇××	2	
873078	19481002	2	4600004	××町3丁目17	白石	〇〇××	3	
456578	19401101	1	4600004	××町3丁目9	白石	〇〇××	4	
554897	19390402	1	4600004	××町3丁目9	白石	〇〇××	1	✓
387896	19360401	2	4600004	××町2丁目2-203	白石	〇〇××	2	
789156	19350302	1	4600004	××町3丁目51	白石	〇〇××	3	

1. 第一順位＝小学校区、第二順位＝年齢で並べ替えた状態にする。これにより、小学校区ごとに若い順から高齢の順にデータが並べ替えられる。
2. 抽出用番号という列を作り、母数÷サンプル数の値の小数点を切り上げた数字を、繰り返し循環するように入力する。
(例：高齢者が 3350 人いる中から 1000 人を抽出する場合、 $3350 \div 1000 = 3.35$ なので、1～4 を繰り返し循環させる)
3. フィルタ機能を使う等して、抽出用番号で「1」がついた人だけ、抽出人数（上記の例でいうと 1000 人）になるまで上から順にフラグを立てる。
4. 3の作業で抽出された人数が、最終的に抽出する人数に足りない分は、抽出用番号に「2」がついた人にリストの上から順にフラグを立てることで補う。