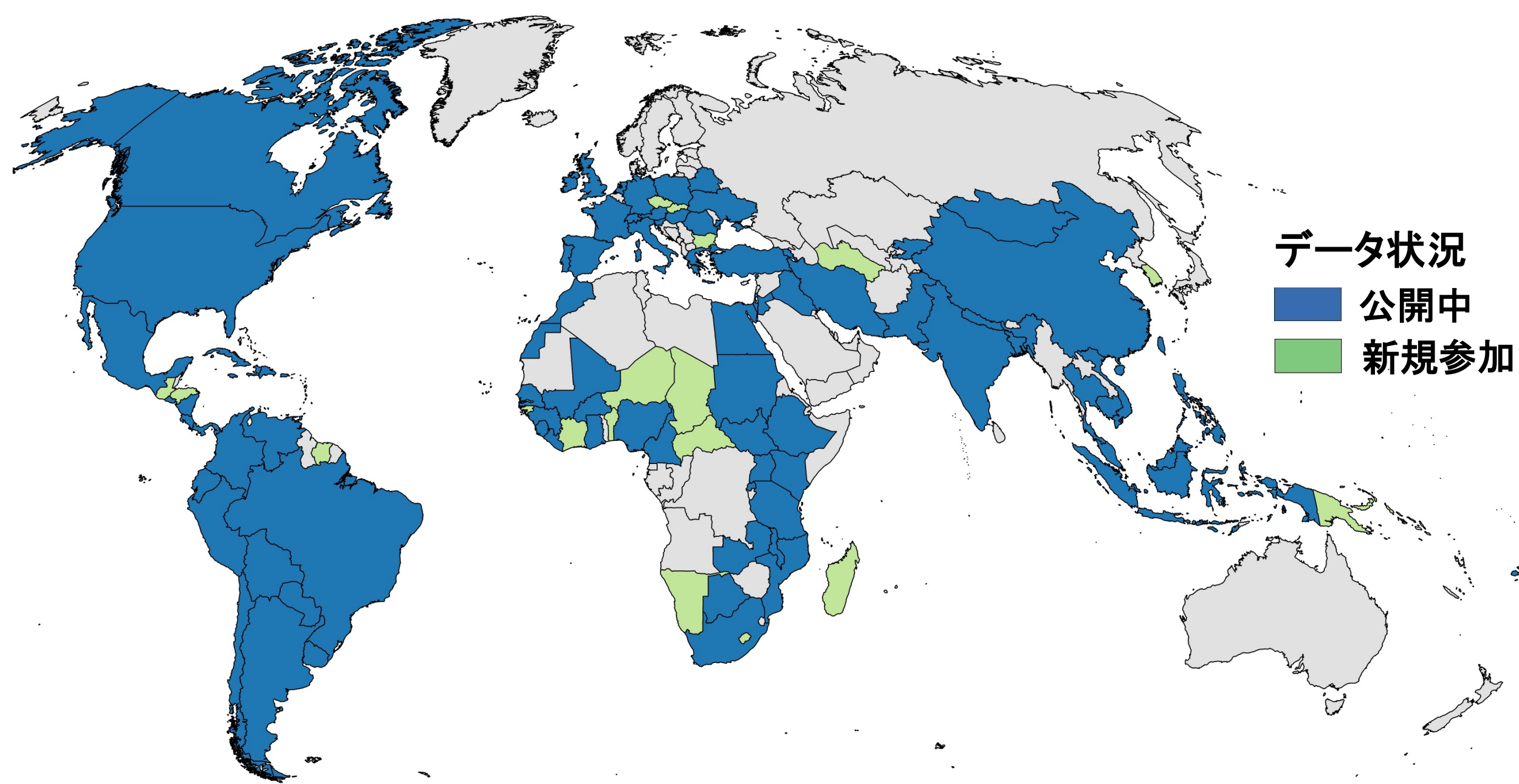


概要

IPUMS(International Integrated Public Use Microdata Series)
ミネソタ人口センターと世界各国の統計局による共同プロジェクトである。
IPUMS が統合した人口マイクロデータは、現在85カ国におよぶ301のセンサス(国勢調査)から、6億7,200万人の個人に関する世界最大規模の情報を保有している。プロジェクトでは、資格要件を満たしたすべての研究者にこれらのデータを無料で提供している。研究者は自分のPCにマイクロデータをダウンロードし、任意の統計パッケージで分析することができる。
このプロジェクトには100カ国以上が参加し、1万4,000人以上の研究者がデータ利用の登録をしている。IPUMS は毎年およそ20件の標本(サンプルデータ)を追加している。それらは新たに参加する国のデータを加える場合と、すでにIPUMSに参加している国の最新のセンサスを加えてデータシリーズを拡張する場合とがある。

<https://international.ipums.org>

IPUMS-Internationalの対象地域



現在、IPUMSは85カ国のデータを公開している。大半の国が複数の標本を提供しているため、時間を追った分析が可能である。他にも多くの国がIPUMSにデータを委託、またはデータ提供を申し出ている。

目標

IPUMS の4つの主目標：

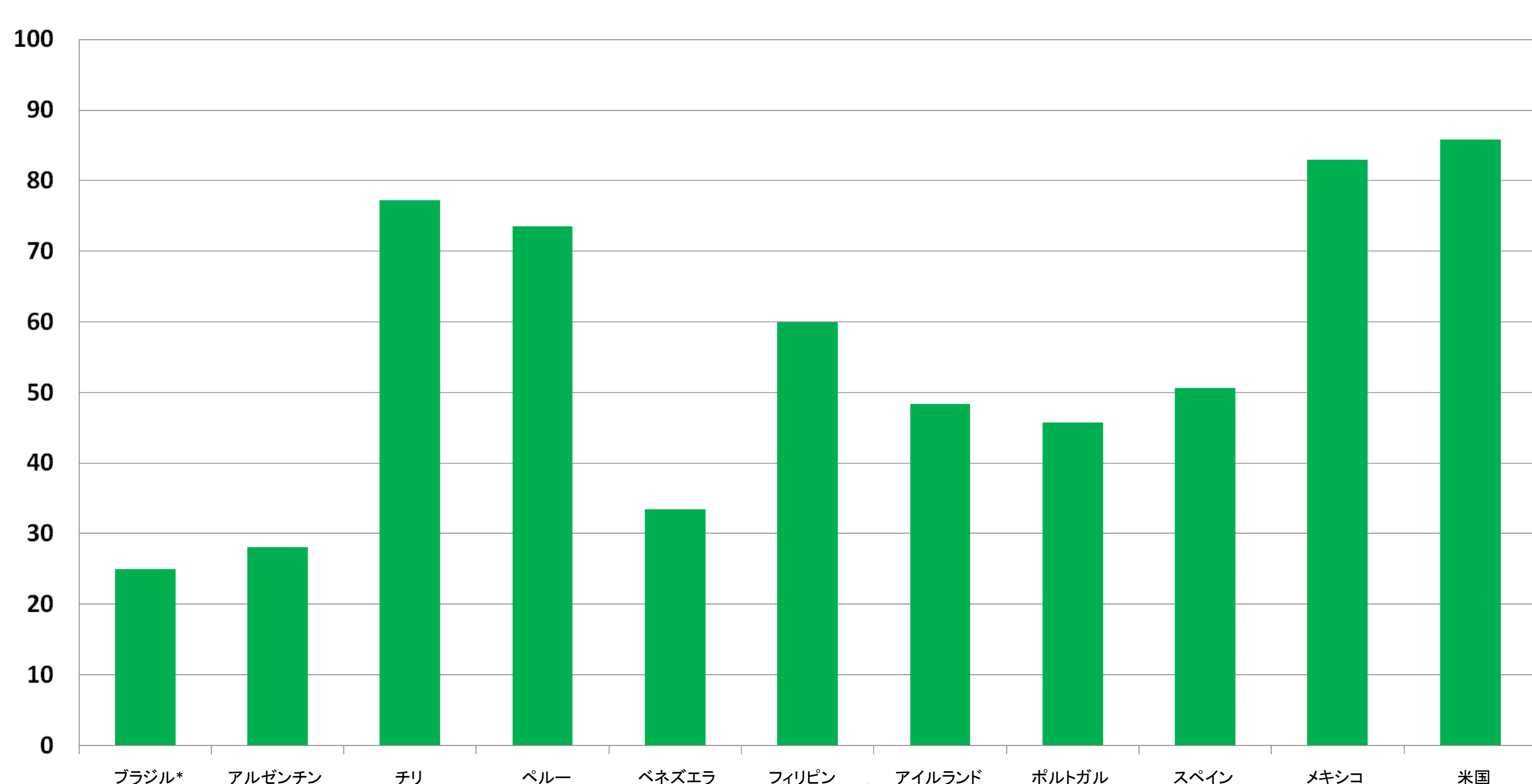
- **保存** - 世界中のセンサスデータは、メディアの劣化とその取り込みに関する知識が失われ、データ損失の危機にある。IPUMSは、このプロジェクトでの提供が許可されていないデータの場合でも、すべてのセンサスデータのアーカイブに取り組んでいる。このプロジェクトは、古いデータを回復するために、多数のデータ復旧プロジェクトに資金を提供している。
- **秘密保護** IPUMSは、開示リスクを最小限に抑えるため、各国の統計局と協力している。この対策では、地域間のデータのスワッピング、小さなカテゴリまたは重要なグループを大きなカテゴリに結合すること、および連続変数のトップコーディングが含まれている。一般的に、人口が25,000人未満の地域は識別されない。
- **統一化** - IPUMS は標本間において、変数コードとラベルを統一化することで、時間と国を越えた比較分析を容易にする。
- **無料公開** - IPUMS はデータへのアクセスを平等化することを目指す。IPUMS のすべてのデータはどこでも無償で研究者に提供される。すべての研究者は IPUMS に登録しなければならず、優先的にアクセスできる者はいない。

統一化

IPUMS は単にデータをアーカイブに保存するのではなく、データベース内のすべての標本に同じコードとラベルが適用できるよう、国を越えて変数を統一している。したがって、IPUMS全体で婚姻区分を示す状態変数は1種類であり、あらゆる時と場所に対して同じコードは同じ内容を意味する。統一化後に生じる整合性の問題を喚起するため、各資料にはその内容が記載されている。研究者は、必要に応じて統一化前のオリジナルデータにアクセスすることもできる。

統一化の意図は、次に例を示すような比較分析を促すことにある。この例は、11のセンサスからブラジル人の成人を統合したものである。

国外に住むブラジル人：中等教育を受けた人の割合(2000年前後)



ウェブによる公開システム

IPUMS のデータ提供は、単なるデータセットのダウンロードではない。
このプロジェクトが構築した高度なウェブベースの公開システムでは、ユーザーが求めるセンサスのあらゆる組み合わせを閲覧し、そこから変数を選択することができる。このシステムの閲覧中、ユーザーは関連文献、コード、一覧テキストを表示することができる。これらはユーザーの関心がある標本情報のみを示すために 絞り込まれている。
研究者は、世界各国の統一化された変数か、またはオリジナルデータを示す統一化前の標本固有の変数のいずれかを選択することができる。
抽出システムでは、研究者が要求するあらゆる標本の組み合わせから、選択した変数を含む統合的なデータセットを生成する。したがって、データ分析では、時間と場所を変数として扱うことができるため、比較が容易となる。データベースに収められた6億7,200万人全員の変数を含めて1つのデータセットを生成することも可能である。

公開中の標本

アルゼンチン	1970, 1980, 1991, 2001, 2010	マラウイ	1987, 1998, 2008
アルメニア	2001, 2011	マレーシア	1970, 1980, 1991, 2000
オーストリア	1971, 1981, 1991, 2001, 2011	マリ	1987, 1998, 2009
バングラデシュ	1991, 2001, 2011	メキシコ	1960, 1970, 1990, 1995, 2000
ベラルーシ	1999, 2009		2005, 2010, 2015
ボリビア	1976, 1992, 2001	モンゴル	1989, 2000
ボツワナ	1981, 1991, 2001, 2011	モロッコ	1982, 1994, 2004
ブラジル	1960, 1970, 1980, 1991, 2000	モザンビーク	1997, 2007
	2010	ネパール	2001
ブルキナファソ	1985, 1996, 2006	オランダ	1960, 1971, 2001
カンボジア	1998, 2008	ニカラグア	1971, 1995, 2005
カメルーン	1976, 1987, 2005	ナイジェリア	2006, 2007, 2008, 2009, 2010
カナダ	1971, 1981, 1991, 2001, 2011	パキスタン	1973, 1981, 1998
チリ	1960, 1970, 1982, 1992, 2002	パレスチナ	1997, 2007
中国	1982, 1990, 2000	パナマ	1960, 1970, 1980, 1990, 2000
コロンビア	1964, 1973, 1985, 1993, 2005		2010
コスタリカ	1963, 1973, 1984, 2000, 2011	パラグアイ	1962, 1972, 1982, 1992, 2002
キューバ	2002	ペルー	1993, 2007
ドミニカ共和国	1960, 1970, 1981, 2002, 2010	フィリピン	1990, 1995, 2000
エクアドル	1962, 1974, 1982, 1990, 2001	ポーランド	1978, 1988, 2002, 2011
	2010	ポルトガル	1981, 1991, 2001, 2011
エジプト	1986, 1996, 2006	プエルトリコ	1970, 1980, 1990, 2000, 2005
エルサルバドル	1992, 2007		2010
エチオピア	1984, 1994, 2007	ルーマニア	1977, 1992, 2002, 2011
フィジー	1966, 1976, 1986, 1996, 2007	ルワンダ	1991, 2002
フランス	1962, 1968, 1975, 1982, 1990	セントルシア	1980, 1991
	1999, 2006, 2011	セネガル	1988, 2002
ドイツ	1970, 1971, 1981, 1987	シエラレオネ	2004
ガーナ	1984, 2000, 2010	スロベニア	2002
ギリシャ	1971, 1981, 1991, 2001, 2011	南アフリカ	1996, 2001, 2007, 2011
ギニア	1983, 1996	南スーダン	2008
ハイチ	1971, 1982, 2003	スペイン	1981, 1991, 2001, 2011
ハンガリー	1970, 1980, 1990, 2001, 2011	スーダン	2008
インド	1983, 1987, 1993, 1999, 2004	スイス	1970, 1980, 1990, 2000
	2009	タンザニア	1988, 2002, 2012
インドネシア	1971, 1976, 1980, 1985, 1990	タイ	1970, 1980, 1990, 2000
		トリニダード・ト	
	1995, 2000, 2005, 2010	バゴ	1970, 1980, 1990, 2000, 2011
イラン	2006, 2011	トルコ	1985, 1990, 2000
イラク	1997	ウガンダ	1991, 2002
アイルランド	1971, 1979, 1981, 1986, 1991	ウクライナ	2001
	1996, 2002, 2006, 2011	英国	1991, 2001
イスラエル	1972, 1983, 1995	米国	1960, 1970, 1980, 1990, 2000
イタリア	2001		2005, 2010
ジャマイカ	1982, 1991, 2001	ウルグアイ	1963, 1975, 1985, 1996, 2006
ヨルダン	2004		2011
ケニア	1969, 1979, 1989, 1999, 2009	ベネズエラ	1971, 1981, 1990, 2001
キルギスタン	1999, 2009	ベトナム	1989, 1999, 2009
リベリア	1974, 2008	ザンビア	1990, 2000, 2010

大部分のIPUMS 標本が各国の人口の5%または10%を代表している。標本サイズの中央値は828,000 人分の記録である。データベースには、全体で1960年から2015年までの672,173,839人の情報が含まれる。サンプル調査(サーベイ)の標本は少ない。

付加価値

IPUMS は、統一化とウェブによる公開システムの他にも、提供データには以下のような付加価値と効用を与える。

- 「ポイント」変数 - 各個人の母、父、配偶者の行番号を示す。配偶者の年齢や母親の雇用形態など、新たな変数の生成を容易にする。
- GIS 境界ファイル - 空間分析とマッピングを容易にする。境界ファイルはほぼすべての国で、第1および第2の行政レベルにおいて利用ができる。
- オンラインによる図表の作成 - データを統計パッケージに読み込まなくとも、素早い分析を可能にする。

謝辞

IPUMSプロジェクトは、U.S. National Science FoundationおよびNational Institute of Child Health and Human Developmentからご支援頂いた。さらに、参加各国の統計局からデータのご提供を頂いた。