EPOP
Understanding Entrepreneurship in the US

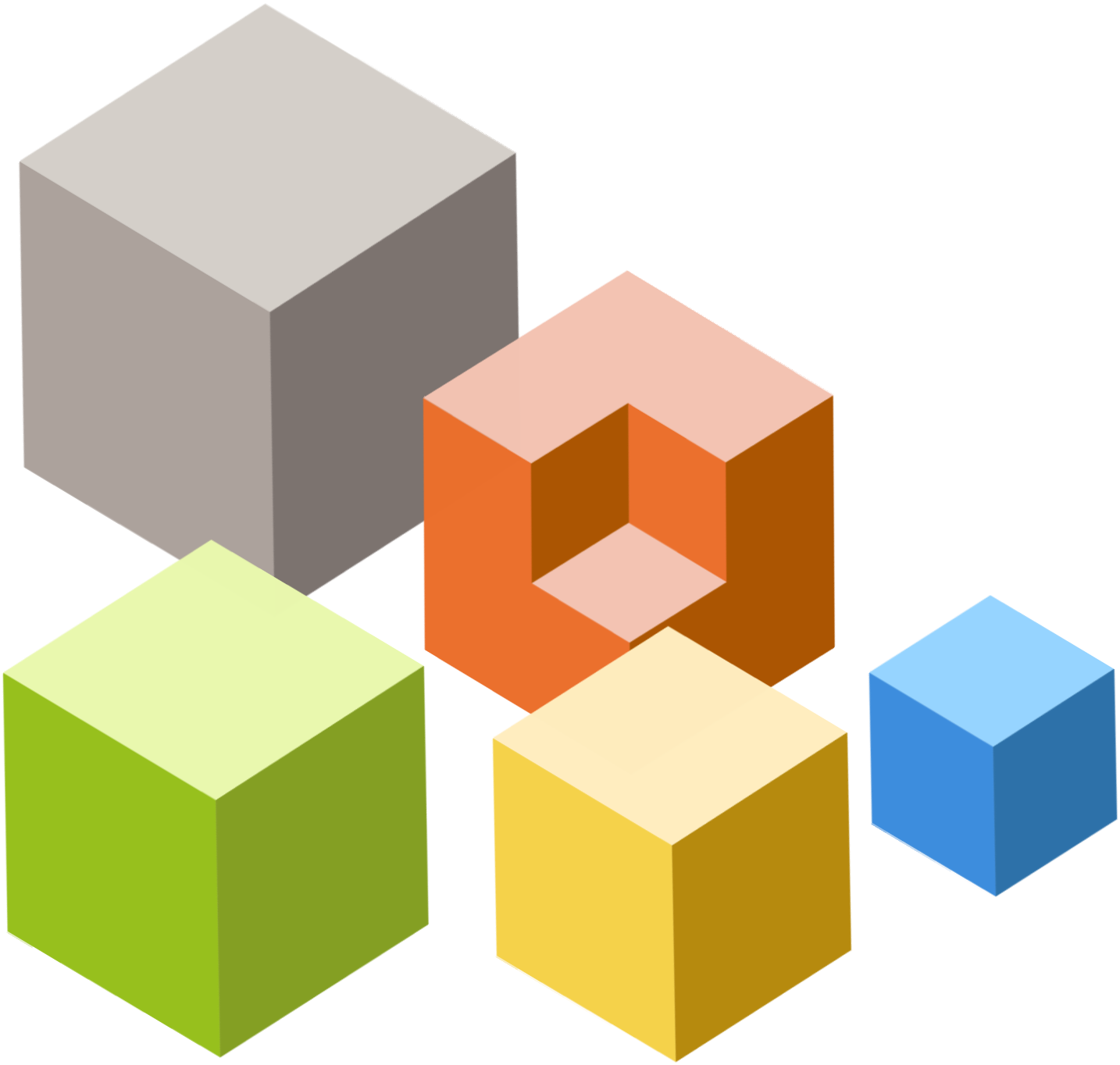
ISSUED: December 30, 2024
VERSION: 2

Entrepreneurship in the Population Survey

EPOP:2024 Data User Guide

Created by
NORC at the University of Chicago
55 East Monroe Street, 30th Floor
Chicago, IL 60603
(312) 759-4000 Main
(312) 759-4004 Fax

Point of Contact
The NORC EPOP Research Team
EPOPresearch@norc.org



The Entrepreneurship in the Population Survey Project is being conducted by researchers at NORC at the University of Chicago with funding from a grant from the Ewing Marion Kauffman Foundation. Questions about this research project should be directed to EPOPresearch@norc.org.

The full title of the survey is "The Entrepreneurship in the Population Survey" and the abbreviation is EPOP Survey. In referencing the project or document, follow these standards:

Full Project Title:	The Entrepreneurship in the Population Survey Project: 2024
Project Abbreviation:	EPOP Survey
Survey Cycle Abbreviation:	EPOP:2024
Full Document Title:	Entrepreneurship in the Population Survey Data User Guide: 2024
Document Abbreviation:	EPOP:2024 Data User Guide
Citation:	Entrepreneurship in the Population Survey Data User Guide: 2024." NORC at the University of Chicago. October 15, 2024. https://EPOP.norc.org.

TABLE OF CONTENTS

1. Overview 1
About the EPOP Survey 1
Sponsor and Partner 1
About this Document 2
2. EPOP Survey Design and Response Rates 3
Target Population 3
Sample Design 3
Sample Sources 3
Data Collection and Response Rates 4
3. Survey Content 5
Significant Changes from EPOP:2023 to EPOP:2024 5
Screener 5
Job Type and Gig Work Measurement 6
Entrepreneurial Activities 6
Pathways and Prioritization 7
1 Current Business Owners 7
2 Current Freelancers 8
3 Nascent Entrepreneur 8
4 Former Business Owners 8
5 Former Freelancer 8
6 Withdrawn Entrepreneur 8
7 Non-Entrepreneur 8
Ancillary Questions by Entrepreneurship Categories 9
Pursuing Entrepreneurship Section 9
Business Ownership Operations Section 10
Non-Entrepreneur Section 10
Demographics 11
Primary 11

Secondary 11
Difference in Demographic Survey Administration by Sample Type 12
4. Working with EPOP Data Files 13
Administrative Variables 13
Location Variables 14
Weighting Variables 14
Data File Conventions 15
Variables Name 15
Reserve Codes 17
Data Protection 17
Weights 19
Development of Weights 19
How to Use Weights 21
Variance Estimation 22
Using the .csv data file 24
Using the STATA data file 24
Using the SAS data file 24
5. Reporting, Dissemination and Future Files 26
Abbreviations and Citations 26
EPOP Website: News and Publications 26
Where to find other Publications 27
Publish your analysis on the EPOP Website 27
Anticipated Data Release Schedule 27
References 28
Appendix 29
Appendix A: Entrepreneurship Pathways and Topical Areas in EPOP:2024 ..... A-1
Appendix B: Survey Questionnaire Changes from EPOP:2023 to EPOP:2024 ..... B-1
Appendix C: Cross Round Variable Changes to Reference when comparing or merging data from EPOP:2023 to EPOP:2024 ..... C-1

1. Overview

About the EPOP Survey

The Entrepreneurship in the Population Survey, or EPOP Survey, has been conducted in 2022, 2023 and 2024; two more annual collections are planned in 2025 to 2026. The survey is designed to understand the scope of entrepreneurial activities from adults 18 years and older in United States and result in a variety of measures of entrepreneurial behavior including current and former business ownership, whether individuals are currently taking or have in the past taken steps towards starting a business, the extent to which individuals engage in freelance work, and engagement with the "gig economy." In addition to capturing the characteristic profile of the individuals involved in these various entrepreneurial activities across the U.S., the survey collects information on the behaviors, challenges, and resources available to individuals during the entrepreneurial process.

Information about the EPOP Survey methods, data availability, publications, and access to data user support may be found on the project's website: EPOP.norc.org.

EPOP Survey Manager and Conservator

NORC at the University of Chicago (NORC) is developing and conducting the EPOP Survey Project with grant funding from the Ewing Marion Kauffman Foundation. NORC is responsible for collecting, maintaining, disseminating, and safeguarding the resulting EPOP Survey data. For the project, NORC is both the manager of the enterprise and conservator of the resulting data.

NORC is an independent research institution that delivers reliable data and rigorous analysis to guide critical programmatic, business, and policy decisions. We conduct objective, non-partisan research to help inform people in government, nonprofits, and businesses making decisions on key issues of the day. Our research addresses important issues like employment, education, and health care. Since 1941, NORC has conducted groundbreaking studies, created and applied innovative methods and tools, and advanced principles of scientific integrity and collaboration. Today, government, corporate, and nonprofit clients around the world partner with NORC to transform increasingly complex information into useful knowledge. For more information, visit norc.org and connect with us via Twitter (twitter.com/norcnews) or Facebook (facebook.com/NORCatUofC.)

Sponsor and Partner

The Ewing Marion Kauffman Foundation is a private, nonpartisan foundation based in Kansas City, MO., which seeks to build inclusive prosperity through a prepared workforce and entrepreneur-focused economic development. The Foundation uses its $3 billion in assets to

change conditions, address root causes, and break down systemic barriers so that all people – regardless of race, gender, or geography – have the opportunity to achieve economic stability, mobility, and prosperity.

For more information, visit their website at Kauffman.org or connect with Kauffman via Twitter (twitter.com/kauffmanfdn) or Facebook (facebook.com/kauffmanfdn).

ABOUT THIS DOCUMENT

This document is designed to help data users analyze and better understand the EPOP:2024 Public Use File data. If there is methodology that is not covered in this document, a full methodology report for the 2024 survey may be found on EPOP.norc.org.

2. EPOP Survey Design and Response Rates

Target Population

The target population of the EPOP Survey includes noninstitutionalized adults 18 years or older in the United States.

The sample design supports the following estimation objectives:

- National estimates of entrepreneurial activity by demographics such as race/ethnicity, gender, age, and education, but not necessarily by the cross of these demographic variables,
- State-level estimates of entrepreneurial activity by race/ethnicity and gender, but not necessarily by the cross of these variables, and
- Metropolitan statistical area (MSA) level estimates of entrepreneurial activity for the top 50 MSAs by population[1] by race/ethnicity and gender, but not necessarily by the cross of these variables.

Sample Design

A stratified sampling design is used to achieve these objectives. Each state that does not contain a top 50 MSA constitutes a primary sampling stratum or a geography. For states that contain one or more of the top 50 MSAs, each MSA and the rest of state outside MSAs make a primary sampling stratum. For example, seven (7) strata or geographies are defined for the state of California, including the six (6) MSAs within the state plus the rest of the state. In addition, MSAs that are made up of counties from multiple states are divided into multiple primary sampling strata, one for each state. For example, Minneapolis-St. Paul-Bloomington, MN-WI, contains counties from both Minnesota and Wisconsin. One important objective of the study is to support estimation and analysis of entrepreneurship characteristics of underrepresented minorities, particularly Black and Hispanic individuals, within states and MSAs. Therefore, each primary stratum is further divided into three secondary sampling strata: Hispanic, non-Hispanic Black, and non-Hispanic Other.

Sample Sources

The study sample is selected from three frame sources:

1. NORC's AmeriSpeak® Panel,

[1] Top 50 MSAs are defined according to total population size in the 2020 decennial census. In the remainder of this document, "MSAs" refer to these top 50 MSAs.

2. An address-based sample (ABS) frame built from the United States Postal Service (USPS) Delivery Sequence File (DSF), and
3. Opt-in online survey panels.

Samples selected from the AmeriSpeak Panel and the ABS frame are probability samples with explicit stratification and known sample selection probabilities while the sample obtained from the opt-in online survey panels is a nonprobability sample with unknown frame coverage and unknown selection probabilities. Subsequent to data collection, the completed surveys from the three (3) samples are combined through a tree model to generate a set of combined sample weights for estimation.

DATA COLLECTION AND RESPONSE RATES

EPOP:2024 survey data collection began on February 28, 2024, for the AmeriSpeak® sample, March 8, 2024 for the ABS sample, and May 30, 2024 for the opt-in online survey panel samples. Differential data collection protocols were followed for each of the sample types. After sending survey requests by USPS letter, email, and prompting calls, data collection ended on July 8, 2024 for both ABS and AmeriSpeak Panel samples. Data were primarily collected via an online survey; computer-assisted telephone interviewing was a secondary mode and available upon request. The survey was available in both English and Spanish. All participants were compensated for their participation.

The response rate varied by sample type. For the AmeriSpeak Panel sample, the response rate was 35.4%, and for the ABS sample, the response rate was 9.3%. For the opt-in survey panels, the response rate is not reported.

3. Survey Content

Significant Changes from EPOP:2023 to EPOP:2024

Based on data review of the EPOP:2023 data, some updates were made to the EPOP:2024 questionnaire. Appendix B lists all survey item updates, including the specific change made, variable name, questionnaire section, and update type and category. The three main types of updates are:

1. New items,
2. Removed items, and
3. Changes to existing items.

Changes to existing items includes logic updates and response option changes. In the EPOP:2024 questionnaire, a few response option changes were made to the PE_CAPITAL and BO_ADDFINANCE series of items to make options clearer and remove options that might have caused confusion for respondents.

A few items were dropped for EPOP: 2024 including some business operations questions (e.g., BO_EMPWAGES) and non-entrepreneurial items (some items from the GP_CONSIDER series). A handful of items were also added to the questionnaire that measured how many owners are/were involved in the business and some additional demographic questions related disability and incarceration history. Please see Appendix B for details.

In addition to the outlined changes to the EPOP survey itself, NORC added several questions that were funded by the Association for Enterprise Opportunities (AEO). Details of these additional questions can be found in the Ancillary Questions by Entrepreneurship Category section related to microbusiness owners of this document.

Screener

The screener section of the survey determines a respondent's working status (e.g., currently employed, retired, student, etc.) and, if working, their working arrangements and any potential entrepreneurial activities they might be engaged in. Through a multiple step process, the screener section identifies various possible entrepreneurial activities capturing current and former business ownership, current and former freelance/consultant/independent contracting work, and any current new business planning as well as situations where respondents were planning to start a business of some type at one point but withdrew from the planning process. Additionally, the screener was designed to capture flexible work arrangements provided by the gig economy.

Employment Status Measurement

The EPOP Survey first establishes the respondent's work status by asking, "*In the last week, did you work for pay at a job or business?*" Following the results of Abraham and Amaya (2019), the questionnaire also asks, "*In the last week, did you do ANY work for pay, even for as little as one hour?*" By asking this follow-up question of respondents who report they are not working, the survey ensures more informal work activities are captured and asked about which is important for determining an accurate measurement of gig work and the full suite of entrepreneurial activities.

Job Type and Gig Work Measurement

After establishing employment status, the EPOP survey collects key job information from those employed to construct a typology for entrepreneurs. This includes three main job types: (1) self-employed/business owner (respondents who select they either own their own business or are freelancers), (2) working for a for-profit or non-for-profit company, or (3) working for the government. This information is collected for both primary and secondary jobs.

Additionally, respondents are asked if the primary or secondary job is gig work. Given the potential for lack of clarity in what counts as gig work, the survey includes extensive examples of gig work activities and includes a definition of gig work in the main text of the question:

> "*Some people earn money through short, paid tasks or jobs online or in-person that are conducted through companies that coordinate payment for the service. This is sometimes referred to as 'gig work.'*"

A final question is asked to determine if respondents are engaged in gig work regardless of the prior responses about the primary and secondary jobs to ensure all gig work activities are reported even if that gig work is not the primary or secondary job.

Entrepreneurial Activities

Once job information has been captured, the EPOP Survey asks questions to gauge entrepreneurial activities directly. These survey items include a series of questions designed to determine if a respondent currently owns a business but does not work at it; has owned any sort of business enterprise in the past which is now closed; is planning a new business enterprise; or considered starting a business in the past, but ultimately withdrew from the enterprise. In EPOP:2023, the survey broadened the set of respondents asked former ownership and freelancer questions and nascent business owner questions to make each entrepreneurial activity fully independent of one another. More information on these screener changes are detailed in the EPOP 2023 User Guide, Subsection 3, "Significant Changes from EPOP:2022 to EPOP:2023".

.

1. **Former business ownership and freelancer.** Respondents are asked if they have ever owned a business or freelanced and if so when this activity stopped. Importantly, some respondents at this step report they currently own a business even though it is not reported as a current job. This likely reflects individuals who are currently passive business owners, and the business ownership is not considered a job. The results presented below combine these business owners with those who report business ownership as their primary or secondary job.

2. **Nascent entrepreneur.** To measure whether respondents are currently taking steps towards owning a business venture of any type, respondents are asked, "*Are you, alone or with others, currently trying to start a new business, including any form of self-employment, freelancing, consulting, or independent contracting, or selling any goods or services to others?*"

3. **Withdrawn entrepreneur.** Respondents are asked if they have ever considered starting a business, but withdrew from planning the enterprise: "*Have you, alone or with others, ever considered starting a new business, including any form of self-employment, freelancing, consulting, or independent contracting, or selling any goods or services to others but decided to wait or change your mind?*"

PATHWAYS AND PRIORITIZATION

The EPOP Survey screener is intentionally designed to capture the full range of entrepreneurial activities in which an individual might be engaged. As a result, some respondents qualify for multiple categories. To limit the burden on survey participants, each respondent is assigned to just one entrepreneurship category for follow-up survey questions using a priority order schema. That priority schema and description of each entrepreneurship category are presented here and shown in Table 1.

1 Current Business Owners

Respondents who report they currently own a business. Importantly, some respondents report that they currently own a business even though it was not reported as a current job. This likely reflects individuals who are currently passive business owners, but for whom the business ownership is not considered a job. The results presented below combine these business owners with those who report business ownership as one of their two primary jobs. Therefore, "current business owners" includes individuals who report that they still own a business even if it is not one of their two primary jobs.

2 Current Freelancers

Respondents who report they are currently freelancers, consultants, or independent contractors. Like the current business owner category, this category includes individuals who report that they are freelancers, consultants, or contractors even if they do not report their freelance work as one of their two primary jobs.

3 Nascent Entrepreneur

To measure whether respondents are currently taking steps towards owning a business, respondents are asked "*Are you, alone or with others, currently trying to start a new business, including any form of self-employment, freelancing, consulting, or independent contracting, or selling any goods or services to others?*" For the purposes of survey categorization, this classification does not condition on specific steps being taken towards entrepreneurship (such as Bennet and Chatterji (2019)), but this information is available in the EPOP Survey's follow questions. In this way, individual researchers can create measures suited to different definitions of nascent business development.

4 Former Business Owners

Respondents who answer they used to own a business but are no longer current business owners.

5 Former Freelancer

Respondents who report they were at one time a freelancer, consultant, or independent contractor but are no longer engaged in freelance work.

6 Withdrawn Entrepreneur

Respondents who answer yes to the following question regarding whether they have considered starting a business, "*Have you, alone or with others, ever considered starting a new business, including any form of self-employment, freelancing, consulting, or independent contracting, or selling any goods or services to others but decided to wait or change your mind?*"

7 Non-Entrepreneur

Respondents who are not engaged in any of the previous six entrepreneurial activities. These respondents receive "general population" questions.

Table 1. Assigned Entrepreneurship Categories by Reported Entrepreneurship Activities

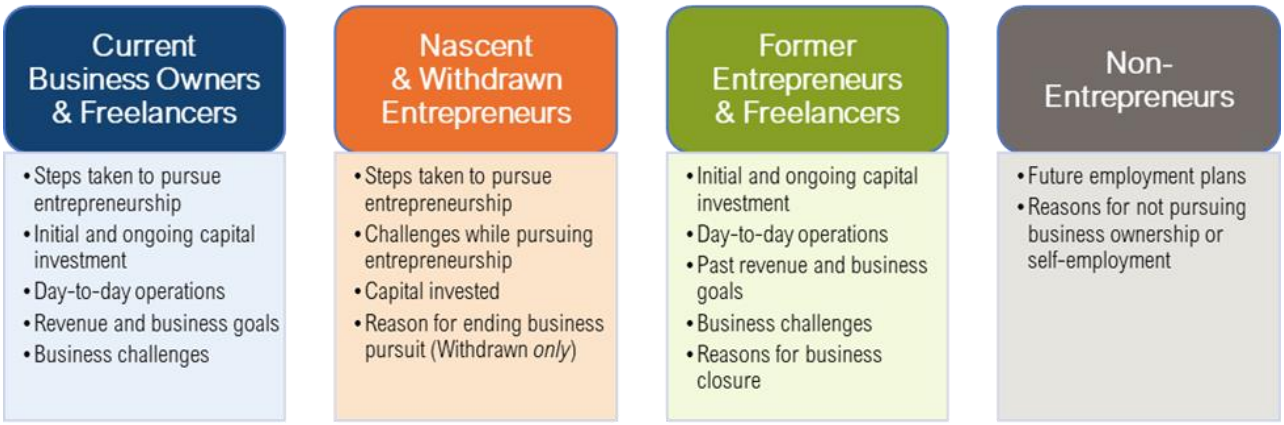
Survey Pathway Priority	Assigned Entrepreneurship Category	Total Surveys	Reported Entrepreneurship Activity					
			Current Business Ownership	Current Freelancing	Entrepreneur-ship Planning	Former Business Ownership	Former Freelancing	Withdrawn Entrepreneur-ship Planning
1	Current Business Owner	4,653	4,653	2,670	2,313	891	1,203	2,644
2	Current Freelancer	4,397	0	4,397	1,407	900	994	2,080
3	Nascent Entrepreneur	1,667	0	0	1,667	332	585	1,276
4	Former Business Owner	2,667	0	0	0	2,667	1,356	783
5	Former Freelancer	2,682	0	0	0	0	2,682	872
6	Withdrawn Entrepreneur	3,249	0	0	0	0	0	3,249
7	Non-Entrepreneur	14,199	0	0	0	0	0	0
	Total	33,514	4,653	7,067	5,387	4,780	6,820	10,904

Source: NORC, Entrepreneurship in the Population Survey: 2024.

ANCILLARY QUESTIONS BY ENTREPRENEURSHIP CATEGORIES

The focus of the EPOP Survey ancillary or follow-up questions asked of each assigned entrepreneurship category is shown in Figure 1 and are briefly described below. Survey questions for current and former business owners and freelancers focus on concepts such as the operations and goals of the activities, whereas nascent and withdrawn entrepreneurs are asked more about concepts such as challenges starting a business. Non-entrepreneurs are asked more general questions about future employment plans as well as reasons for not pursing business ownership or self-employment.

Figure 1. EPOP Survey Content Focus by Assigned Entrepreneurship Categories



Pursuing Entrepreneurship Section

The questions in this section focus on the steps respondents took to pursue starting a business or working for themselves. The topics covered within this section are asked of current business owners and Freelancers as well as nascent and withdrawn entrepreneurs.

Business Ownership Operations Section

The questions in this section focus on the day-to-day operations of business owners and freelancers/consultants/independent contractors when their businesses were in full operation. For former business owners and freelancers, the questions pertained to the last year when their business was in operation. Specifically, the topics covered in this section include questions on when they started the business or self-employment and general descriptions of the type of business, legal status of the business, and how they came up with the idea for the business or self-employment.

For current and former business owners and freelancers, this section also asked for the types and amounts of additional financing they requested and/or received to continue the business or self-employment, the number and types of employees they used in their business, how much time they spent managing or working in their business, their revenue and profit/loss margin, their goals for the next five years, their biggest challenges facing their business (or former business), and their post-entrepreneurship plans and exit strategy.

Finally, all entrepreneurship categories were asked to indicate the industry that best classifies their current, former, or idea for a business.

Non-Entrepreneur Section

The questions in this section focus on respondents who did not have any prior experience with business ownership and were not currently taking steps to own a business or be self-employed. These respondents are asked to provide reasons why they have not considered entrepreneurship and to describe their current work arrangements including how long they've been in their current job, how many coworkers they have, how much they've worked in the past year, what type of benefits they receive with their work, and their likelihood of starting a new job or changing jobs in the future.

See Appendix A from a detailed graphic showing the question topics asked across all the assigned entrepreneurship groups.

Microbusiness Owner Section

In addition to the outlined changes to the EPOP survey itself, NORC added several questions that were funded by the Association for Enterprise Opportunities (AEO). These 18 questions focused on microbusiness owners defined as businesses with fewer than 10 full-time or part-time W-2 employees to explore critical areas such as access to business support services, financing needs beyond startup, and the adoption of cutting-edge technologies like generative AI. These questions were answered by 2,316 survey completers. These additions aimed to uncover the unique barriers and opportunities that entrepreneurs face and identify resources that could help drive success. Questions were integrated into the EPOP:2024 questionnaire and responses are

included in the EPOP:2024 Public Use and Restricted Use data files. Table 2 below shows the variable names of each of the AEO questions.

Table 2. Microbusiness Owner Variable Names

Variable Names	
BO_IMPACT_1	TECH_4
BO_IMPACT_2	TECH_5
BO_TAX_FILING_1	TECH_6
BO_SERVICES_1	TECH_7
BO_SERVICES_2	TECH_8
BO_SERVICES_3	TECH_9
TECH_1	TECH_10
TECH_2	TECH_11
TECH_3	TECH_12

DEMOGRAPHICS

The EPOP Survey asks a comprehensive set of demographic questions. The demographic questions asked at the start of the survey are critical for weighting and are considered primary. The remaining demographic survey items are asked last and considered secondary.

Primary

The primary demographic variables that were asked of the ABS and nonprobability samples included age, gender, ethnicity, race, household income, education, and number of adults in the household. For the AmeriSpeak sample, panel member data on age, gender, education, ethnicity, and race were already known and were not asked again as part of the EPOP Survey questionnaire.

Secondary

Secondary demographic variables included student status, health insurance and benefits, debt amount, marital status, number of household members, number of children in the household, military status, and citizenship status. These secondary demographic questions were asked of all respondents from all sample types.

Difference in Demographic Survey Administration by Sample Type

Data users should be aware that demographic variables were collected somewhat differently for the different sample types. The AmeriSpeak Panel members collected all primary variables and most secondary variables when the individual was originally recruited for the panel and the EPOP Survey did not re-ask those questions of those sample members. The ABS and opt-in online survey sample respondents were asked all demographic questions during EPOP survey administration. To mitigate inconsistency and potential bias across samples, the EPOP Survey aligned the demographic survey items to the AmeriSpeak Panel's version.

4. Working with EPOP Data Files

EPOP Survey data is available via both a restricted use data file (RUF) and a public use data file (PUF) for each annual survey. Both data file types include variables for all survey questions in a format suitable for analysis without disclosure risk. Many variables in the RUF and PUF represent all response choices in the original survey questions. However, based on the need to protect respondent identity, question responses have been aggregated as necessary. In some cases, as with open-ended questions, variables were omitted entirely as the answers are unique to the respondent. The EPOP:2024 Questionnaire shows how the survey item was asked and the data file codebooks (described below) will show if any response options were aggregated for the restricted and public use analysis files. Together, these documents allow data users to clearly see where and how survey item aggregation was implemented.

The RUF provides more finely grained response options aligned more closely to the original response choices in the survey. To obtain the RUF, data file data users must undergo training on disclosure and publishing considerations and sign an agreement. (Please see the "Data Protection" section of this document for more information on the disclosure review process and disclosure considerations). The PUF presents some survey data items in broader categories. The PUF is available on the project website and is available to all interested data users without the need for a formal data use agreement.

Administrative Variables

Both the RUF and PUF contain a set of administrative variables. Administrative variables include information relevant to how the survey was administered, how data was edited, or sample information. Table 3 shows the administrative variables included in the files. Note that only the RUF contains the EPOP Survey sample type variable (SAMP_TYPE). More information on these variables is provided in the data codebooks.

Table 3. EPOP Survey Data File Administrative Variables

Variable Name	Variable Label	RUF or PUF
R_SUID	RESPONDENT ID	Both
SAMP_TYPE	EPOP SAMPLE TYPE	Both
PARTIAL_CASE	INDICTOR FOR PARTIAL COMPLETE CASES	Both
Q_LANGUAGE	SURVEY LANGUAGE	Both
Q_MODE	SURVEY MODE	Both
DEM_AGE_IMPUTED	INDICATOR FOR IMPUTED AGE VALUES	Both

Location Variables

Both the RUF and the PUF include the respondents' residence location data. Each file includes a nine-category Census division variable and four-category Census region derived from the respondent's county or ZIP code. The RUF also includes a county and state variable. Some respondents provided their ZIP code if they could not identify their county. ZIP code was then linked to county. Note that many ZIP codes cross multiple counties. For these cases where county was identified from ZIP codes, the case was coded to the county with the highest population based on 2021 Census estimates. The derived variable ZIP_TO_M_COUNTY is set to '1' for these cases. Table 4 shows the location variables.

Table 4. EPOP Survey Location Variables

Variable Name	Variable Label	RUF or PUF
CENSUS_DIV_DRV	CENSUS DIVISION DERIVED FROM ZIP CODE	Both
REGION_DRV	CENSUS REGION DERIVED FROM ZIP CODE	Both
COUNTY_DRV	COUNTY FIPS CODE DERIVED FROM ZIP CODE	Only RUF
DEM_STATE	RESPONDENT STATE	Only RUF

Weighting Variables

In addition to the survey weight (WTSURVY), design variables (PSU, STRATA) are included to allow for calculating accurate standard errors (more information on the weighting process is in the "Weights" section of this document). The RUF also includes a weight constructed for only the probability sample (WTPROB). Table 5 shows the available survey weight variables.

Table 5. EPOP Survey Weight Variables

Variable Name	Variable Label	Included in:
WTSURVY	SURVEY WEIGHT: APPLIES TO ALL CASES	Both
WTPROB	PROBABILITY SAMPLE WEIGHT: FOR WORK WITH ONLY ABS AND AMERISPEAK SAMPLES	Only RUF
STRATA	SAMPLING STRATA	Both
PSU	PRIMARY SAMPLING UNIT (PSU)	Both

Data File Conventions

Variables Name

In most instances, variable names within the RUF and PUF match each other and the variable names in the EPOP Survey Questionnaire. In instances when survey responses were aggregated to protect respondent confidentiality, variable names have been modified. For example, the original variable name in the questionnaire for the highest level of education is "DEM_EDU." Both the RUF and PUF require a different level of aggregation based on disclosure considerations. The variable recoded for the RUF is appended with "_RUF." The variable recoded for the PUF is appended with "_PUF." When a variable is recoded using the same level of aggregation for both the RUF and the PUF, the variable name is appended with "_DRV." For instance, the nine-category census division the same grouping for both the PUF and the RUF. This variable is named "CENSUS_DIV_DRV" in both files. Table 5 shows the variable name convention used to indicate which variables are modified for the RUF or PUF.

Beginning with the 2023 data collection round, additional naming conventions were developed to track differences between EPOP survey rounds. Efforts were made to use the same variable across rounds but in a small number of circumstances, this was not possible. In some instances, improvements to the survey content or flow required the use of new variable names and at times different aggregations were necessary to preserve respondent confidentiality during the disclosure review process. When this occurred, we added a 'Y2' to the variable to indicate that the content or the code-frame of this variable changed from the Year 1 round.

If a variable changed in the Year 2 but underwent no further changes in the Year 3, the 'Y2' naming convention was retained. For example, the response options to the survey question S_FORMFREE_STAT_1 (Are you still working for yourself as a consultant, freelancer, or independent contractor either full-time or part-time?) changed from the Year 1 to the Year 2. No further changes were made in the Year 3 so the Year 2 variable name was retained. Therefore, the delivery variable name changed from 'S_FORMFREE_STAT_1' in Year 1 to 'S_FORMFREE_STAT_1_Y2' in Year 2 and remained 'S_FORMFREE_STAT_1_Y2' in Year

3. There were no further changes to the variable in the Year 3, so the Year 2 name was retained. Therefore, variables that contain 'Y2' in the variable name, indicate a previously existing question had a variable format change from the Year 1 to the Year 2 but no further changes in the Year 3. Variables with a 'Y3' in the variable name indicate there was a change in a previously existing variable's format from the Year 2 to the Year 3. Tables 6 and 7 provide examples of how the 'Y2' naming convention is used.

Table 6. EPOP Survey Data Variable Name Conventions indicating Aggregation

Variable Source	Name Convention	Example
Original Questionnaire Variable	No change	DEM_EDU
Aggregated for RUF	_RUF	DEM_EDU_RUF
Aggregated for PUF	_PUF	DEM_EDU_PUF
RUF Aggregations updated in Year 2, retained in Year 3	_Y2_RUF	BO_REVENUE_1_Y2_RUF
RUF Aggregations updated in Year 3	_Y3_RUF	BO_STARTBIZ_1_Y3_RUF
PUF Aggregations updated in Year 2, retained in Year 3	_Y2_PUF	BO_REVENUE_1_Y2_PUF
PUF Aggregations updated in Year 3	_Y3_PUF	BO_NUMEMPLOY_1_Y3_PUF
Aggregated for RUF and PUF	_DRV	CENSUS_DIV_DRV

Table 7. EPOP Survey Data Variable Name Conventions indicating Across-Round Changes in Questionnaire Content

Variable Source	Name Convention	Example
EPOP Year 1, 2, and 3 variables	No change	PE_GIGREASON_1_1
EPOP Year 2 and 3	_Y2	PE_CHALLENGE_1_8_Y2
EPOP Year 3 variable	_Y3	PE_CAPITAL_5_1_Y3

When merging files from multiple rounds together, these naming conventions will prevent data users from accidentally combining variables that have a different meaning. In some instances, variables that have answer choices specific to Year 3 can be recoded and combined with Year 1 or Year 2 variables. Tables in Appendix C show which variables can be combined and is provided with the EPOP:2024 RUF and PUF downloads. Data users should reference these tables for guidance on how to combine these variables.

Reserve Codes

When respondents skipped or refused questions or indicated they did not know the response to a question, the response is coded with a reserve code value. Similarly, data points that present a disclosure risk either in isolation or in combination with other data points are masked with a reserve code value. Table 8 shows the list of reserve codes used in the EPOP Survey data files.

Table 8. EPOP Survey Reserve Code Values

Reserve Code	Label
-3	Missing
-5	Don't' Know
-7	Suppressed

DATA PROTECTION

To protect the EPOP Survey data from allowing the potential re-identification of respondents, these four Statistical Disclosure Limitation (SDL) techniques have been used:

1. **Recoding.** Recoding can be used for both categorical and continuous variables. For categorical variables, it involves combining smaller categories into larger categories. It can also be used for continuous variables to code numbers into categories.
2. **Local Suppression.** Local suppression creates missing values to replace some the values.
3. **Rounding.** Rounding is applied to continuous variables like numbers to make the data harder to re-identify.
4. **Micro-aggregation.** Micro-aggregation clusters records into small groups and then the average is released as the value for some of the sensitive units.

The variables that have undergone recoding, rounding, or micro-aggregation are indicated in the data files with variable names with the suffixes _PUF, _RUF, or _DRV (see Table 6). Variables that have undergone local suppression are in Table 9. All suppressed values in the PUF and RUF are indicated by -7 (see Table 8). For a full description of the various recoding schemas used in the PUF and RUF and the SDL methods used in the EPOP Survey data files, see the methodology report on EPOP.norc.org.

Table 9. EPOP Survey Restricted and Public Data File Suppression Count by Variable

Variable Name	Suppression Case Counts
RUCC_DRV	135
DEM_AGE	160
RACE	24
DEM_GENDER	20
DEM_EDU	765
DEM_MARITAL	505
DEM_HHINC	995
DEM_MILITARY_1	445
DEM_CITIZEN	507
DEM_HOUSEHOLD_2_Y2_DRV	8
DEM_HOUSEHOLD_3_Y2_DRV	4
DEM_STUDENT	290
BO_INDUSTRY1	233
BO_STARTBIZ_1	71
BO_EMPLOYEES_1_1	86
BO_EMPLOYEES_1_2	40
BO_EMPLOYEES_1_3	30
BO_EMPLOYEES_1_4	6
BO_EMPLOYEES_1_5	9
BO_EMPLOYEES_1_6	14
BO_EMPLOYEES_1_7	16
BO_NUMEMPLOY_1_Y3_PUF	175
BO_STARTBIZ_1_Y2_PUF	71
BO_WORKHOME_1	64
BO_PLMARGIN_1	130
BO_REVENUE_1	29
DEM_DISABILITY_1A_PUF	112
DEM_DISABILITY_1B_PUF	222
DEM_DISABILITY_1C_PUF	160
DEM_DISABILITY_1D_PUF	147
DEM_DISABILITY_1E_PUF	254

Variable Name	Suppression Case Counts
DEM_DISABILITY_1F_PUF	243
DEM_DISABILITY_1G_PUF	225

Suppressed values are noted with a value of -7 in the data except in cases where only one category is suppressed. When only one category of the variable is suppressed, suppressed cases are coded as -3 so they are intermixed with other missing values.

WEIGHTS

Development of Weights

The EPOP Survey contains two sets of weights: probability sample weights for probability samples (AmeriSpeak and ABS samples), and combined sample weights for the combined probability and nonprobability samples. The combined sample weights are available in both the RUF and PUF, whereas the probability sample weights are available only in the RUF.

Creation of the probability sample weights follows these steps:

1. **AmeriSpeak sample base weights.** Computed as the AmeriSpeak Panel weights divided by the probabilities of selection from the AmeriSpeak Panel to the study sample.

2. **ABS sample base weights.** Computed as the inverse of the selection probabilities that account for both the first and second phase of the ABS sample selection.

3. **Adjustment for unknown eligibility.** This adjustment is applied to the ABS sample because the eligibility status of some sample addresses is not determined at the end of the survey. Through this adjustment, the weights assigned to cases with known eligibility are inflated to account for the eligible cases among the cases with unknown eligibility. This adjustment is not applied to the AmeriSpeak sample because all AmeriSpeak samples are assumed to be eligible.

4. **Adjustment for interview nonresponse.** The interview nonresponse adjustments inflated the weights w_{2i} assigned to eligible complete cases so that they represented the incomplete cases among the eligible sample members. Nonresponse adjustments required information about both respondents and nonrespondents. A lot of information was available for the AmeriSpeak panel sample nonrespondents. On the other hand, only limited information was available for the ABS sample nonrespondents based on commercial data appendage. Therefore, the nonresponse adjustments for the AmeriSpeak and ABS samples were carried out separately using different adjustment cells.

For the AmeriSpeak sample, adjustment cells were constructed by cross-classifying:

a. *Geography* (or primary sampling strata),
b. *Race/Ethnicity* (Hispanic/Non-Hispanic Black, and All Other),
c. *Age* (18-34, 35-64, 65 and older),
d. *Education* (Some college or less and bachelor's degree or above) and,
e. *Gender* (Male and Female).

For the ABS sample, there was no sample frame information on education or age, so adjustment cells were defined by cross-classifying:

a. *Geography* (or primary sampling strata),
b. *Race/Ethnicity* (Hispanic and Non-Hispanic Black, and Other), and
c. *Gender* (Male and Female).

5. **Combined interview nonresponse adjusted weights for probability samples.** The interview nonresponse adjusted weights computed for AmeriSpeak and ABS sample completes are then combined by geography and race/ethnicity group where the composition factor is proportional to the number of completed interviews from each sample source. This combination is carried out such that the combined sample represents the target population for each geography and race/ethnicity group.

6. **Raking to derive probability sample final weights.** Raking benchmarks are developed using the 2022 American Community Survey (ACS) 1-year estimates. Raking adjustments are conducted within each state and MSA using these dimensions:

a. State by Race and Ethnicity (Non-Hispanic White, Non-Hispanic Black, Hispanic, Non-Hispanic Other)
b. State by Gender (Male, Female)
c. State by Age (18-29, 30-39, 40-49, 50-64, 65+)
d. State by Education (High School/GED or Less, Some College, and BA and Above)
e. State by Number of Adults in Household (1 adult in Household, 2 adults in Household, 3 or more adults in Household)
f. *MSA by Race and Ethnicity* (Non-Hispanic White, Non-Hispanic Black, Hispanic, Non-Hispanic Other)
g. *MSA by Gender* (Male, Female)
h. *MSA by Age* (18-29, 30-39, 40-49, 50-64, 65+)

i. *MSA by Education* (High School/GED or Less, Some College, and BA and Above)
j. *MSA by Number of Adults in Household* (1 adult in Household, 2 adults in Household, 3 or more adults in Household)

Weights for the combined probability and nonprobability samples are then developed through a regression tree model.

7. **Decision tree.** Implement the Classification and Regression Tree (CART) supervised machine learning algorithm to generate a decision tree. The decision tree splits the combined sample into a set of leaves.

8. **Estimate inclusion probabilities for the nonprobability sample.** The inclusion probability of the nonprobability sample is computed as the weighted empirical proportion of nonprobability sample cases in the leaf.

9. **Compute the combined sample weights.** The combined sample weights are computed from both known probability sample selection probabilities and estimated nonprobability sample inclusion probabilities. The combined sample weights are calibrated through a ratio adjustment such that the sum of the combined sample weights equal to the sum of the original probability weights within the leaf.

10. **Raking adjustment.** The combined weights are raked to population benchmarks by geography and by race/ethnicity within each geography.

How to Use Weights

The final EPOP:2024 analysis data contains 33,514 respondents, including 12,512 respondents from the probability sample and 21,002 respondents from the nonprobability sample. Two provided weights can be used to generate approximately unbiased estimates of the population, the combined weight (WTSURVY) and the probability sample weight (WTPROB). WTSURVY is used for generating estimates using the full sample which combines the probability and nonprobability samples. WTPROB is used for generating estimates based on only the probability sample.

Probability sample weights **(**WTPROB**)** are developed for AmeriSpeak and ABS samples to correct for potential bias due to unequal sample selection probabilities, nonresponse, and coverage errors. These weights can be used to produce unbiased national estimates, state and MSA level estimates, and estimates for other domains defined by the user. Any software package that can handle sample weights should produce correct weighted points estimates. The probability sample weights are available only in the restricted use file.

The combined sample weights (WTSURVY) are available in both the restricted use file and public use file and make analysis and reporting for many smaller domains possible due to its larger sample size. The combined sample weight can be used to produce approximately unbiased national estimates, state and MSA level estimates, and estimates for other domains defined by the user.

Variance Estimation

The EPOP Survey uses a complex sample design that needs to be accounted for in variance estimation. Otherwise, statistical software will likely underestimate standard errors of estimates. To facilitate variance estimation, we provide sample weights as well as two design variables: PSU and STRATA.

For samples selected from the AmeriSpeak Panel, these variables are pseudo-PSUs and pseudo STRATA that are defined to represent the first stage PSUs and STRATA associated with the NORC National Frame (Master Sample) that was used as the sampling frame for AmeriSpeak Panel recruitment sampling. The ABS and opt-in samples are not clustered, so each PSU is a single sample unit, and each STRATA is an MSA or the rest of a state outside MSAs. Using PSU and STRATA with either probability sample-only weights or combined weights will provide approximately unbiased variance estimates.

Standard variance estimation method can be used to approximate the variance of estimates based on the combined probability and nonprobability sample. NORC can provide additional information to support proper variance estimation.

The sample code provided in Figure 2 shows examples of how variable AVAR can be analyzed using corrections for weighting and sample design in R and Stata. This example uses the combined probability and nonprobability samples with the corresponding weight variable, WTSURVY. SAS users can use PROC SURVEYFREQ and PROC SURVEYMEANS to calculate the design-corrected standard errors.

Figure 2. EPOP Survey Sample Stata and R Code

STATA

Load data	`use EPOP_YR3_PUF.dta,clear`
Set survey design	`svyset [pweight=WTSURVY], ///` `  strata(STRATA) psu(PSU) singleunit(scaled)`
Weighted mean	`svy: mean AVAR`
Weighted percentage	`svy: proportion AVAR`
Weighted total	`svy: total AVAR`
Weighted one-way table	`svy: tabulate AVAR`
Subset mean	`svy, subpop (if SUBGROUP==1): mean AVAR`
Specifying subgroups	`svy: mean AVAR, over(GROUPVAR)`

R

Install & load required packages	`install.packages(c("tidyverse", "survey"))`
	`library(haven)`
	`library(survey)`

Load data	`mydata <- read_dta("EPOP_YR3_PUF.dta")`
Set survey design	`mydesign <- svydesign(id = ~PSU, weights = ~WTSURVY, strata = ~STRATA, data = mydata, nest = TRUE)`
Singleton PSU correction	`options(survey.lonely.psu = "adjust")`
Weighted mean	`svymean(~AVAR, mydesign, na.rm = TRUE)`
Weighted total	`svytotal(~AVAR, mydesign, na.rm = TRUE)`
Weighted one-way table	`svytable(~AVAR, mydesign)`
Subset mean	`svymean(~AVAR, subset(mydesign, SUBGROUP == 1), na.rm = TRUE)`
Specifying subgroups	`svyby(~AVAR, by = ~GROUPVAR, mydesign, svymean)`

Statistical software may return errors when conducting variance estimation on subsamples and/or variables with a large number of observations with missing values. STRATA and PSU were created so that there was a minimum number of respondents within a STRATA/PSU cell. However, if all respondents within a cell are missing on a variable, it will be impossible to calculate the standard error. This is sometimes referred to as a "lonely PSU" or "singleton PSU." If the dataset is subset (to current entrepreneurs, for example), this error becomes more likely to happen. In these situations, you may receive an error such as this:

```
STATA error handling: "missing standard error because of stratum with
single sampling unit"
```

The best workaround to avoid this type of error is to manually combine the single-PSU stratum with a similar stratum. Alternatively, the sample code provided addresses the lonely PSU issue using automatic adjustments. In Stata, the correction is made with the svyset option "singleunit(scaled)" and in R, with the command "options (survey.lonely.psu = "adjust")." These methods of adjustment involve taking variance averages from stratum with multiple sampling units. Users should refer to their software documentation for more information on automatic adjustment methods before implementing them in their own research.

DATA FILE FORMATS

Both the RUF and PUF are available in three file formats: .csv, SAS, and STATA. Each file provides a different set of meta-data requiring different accompanying programs. As an example, the PUF package will include the data files and formatting programs shown in Table 10.

Table 10. EPOP Survey PUF Package File and Formatting Program Contents

File	Name
CSV data file	EPOP_YR3_PUF.csv
STATA data file with formatting applied	EPOP_YR3_PUF.dta
SAS data file	EPOP_YR3_PUF.sas7bdat
SAS program containing labels	EPOP_YR3_PUF_LABELS.sas
SAS program containing formats	EPOP_YR3_PUF_FORMATS.sas
SAS program to apply labels and formats	EPOP_YR3_PUF_APPLY_FORMATS_LABELS.sas

Using the .csv data file

Users of the .csv can refer to the variable names in the header column and review the data codebooks to retrieve variable format information.

Using the STATA data file

The STATA file provided is a formatted file. The file contains all labels and format information. STATA users can simply import that file and run frequencies to review variable formats and variable labels.

Using the SAS data file

Users of the SAS file may apply label and variable format information by applying the provided label and format definitions. To do this, open the provided SAS program 'EPOP_YR3_PUF_APPLY_FORMATS_LABELS.sas' and update the folder reference to the location where the data user has saved the EPOP data files and programs. Then run the program to apply the formatting information.

CODEBOOKS

A separate codebook is provided for the RUF and PUF. Each codebook includes an index of all variables included in the file. For each variable, a table is presented containing the variable name, variable label, original question text, and any survey skip logic. For most variables, a frequency table was the appropriate format to report answer choices. Each frequency table includes unweighted and weighted counts and percentages for each response choice and reserve code (i.e., -3-missing, -5-don't know, -7-suppressed). Continuous variables were reported as a table containing descriptive statistics: valid n, mean, median, min, max. Variables that were not continuous but contained many categories (e.g., the case identifier [R_SUID]) were reported in a frequency table where rows were grouped by valid and reserve code categories.

DATA USER SUPPORT

If you are having issues accessing the link for the PUF, specific files that were sent, or you have other questions about the EPOP Survey data or methods, please contact the EPOP research team at EPOPresearch@norc.org.

5. Reporting, Dissemination and Future Files

Abbreviations and Citations

The full title of the survey is "The Entrepreneurship in the Population Survey" and the abbreviation is EPOP Survey. In referencing a specific year, follow these standards:

Full Project Title: The Entrepreneurship in the Population Survey Project: 2024

Project Abbreviation: EPOP Survey

Survey Cycle Abbreviation: EPOP:2024

User Guide Citation: "Entrepreneurship in the Population Survey User Guide: 2024." NORC at the University of Chicago. October 15, 2024. EPOP.norc.org.

Data File Citations: "Entrepreneurship in the Population (EPOP) Survey Project Restricted Use Data File: <year>." NORC at the University of Chicago. <month> <day>, <year>. EPOP.norc.org.

"Entrepreneurship in the Population (EPOP) Survey Project Public Use Data File: <year>." NORC at the University of Chicago. <month> <day>, <year>. EPOP.norc.org.

EPOP Website: News and Publications

The EPOP Survey project website at EPOP.norc.org posts up-to-date news and information on research using the EPOP Survey data. As researchers use the data to write journal articles, research briefs, book chapters, presentations, and other products, the EPOP research team will post links to their publications on the website at their request.

Other individuals and organizations—in government, non-profit, and for-profit sectors—will also find EPOP data compelling. Ideally, they will use the data for a variety of purposes including policy action, advocacy, media releases, and proposals. Should changes to policies or programs be made based on the EPOP Survey data, the EPOP research team would appreciate being notified and will create a post about it, if permissible.

Where to find other Publications

As the EPOP Survey data gets analyzed by NORC and other researchers and mentioned by news media, a repository of EPOP related research, publications and media mentions will be available on the EPOP website: EPOP.norc.org.

Publish your analysis on the EPOP Website

The EPOP research team welcomes information on research using EPOP Survey data. Please contact EPOPresearch@norc.org if you have analyzed EPOP data and would like your research displayed on the website.

ANTICIPATED DATA RELEASE SCHEDULE

The overall timeline for implementing this project is July 1, 2021, to June 30, 2026. Table 11 shows the anticipated release of future year EPOP Survey data releases for Public Use Files.

Table 11. EPOP Survey Future Data Release Schedule

Data Release	Anticipated Release Period
2022	October 2022
2023	October 2023
2024	October 2024
2025	August – September 2025
2026	August – September 2026

Any changes to these release periods will be posted on the EPOP website: EPOP.norc.org.

REFERENCES

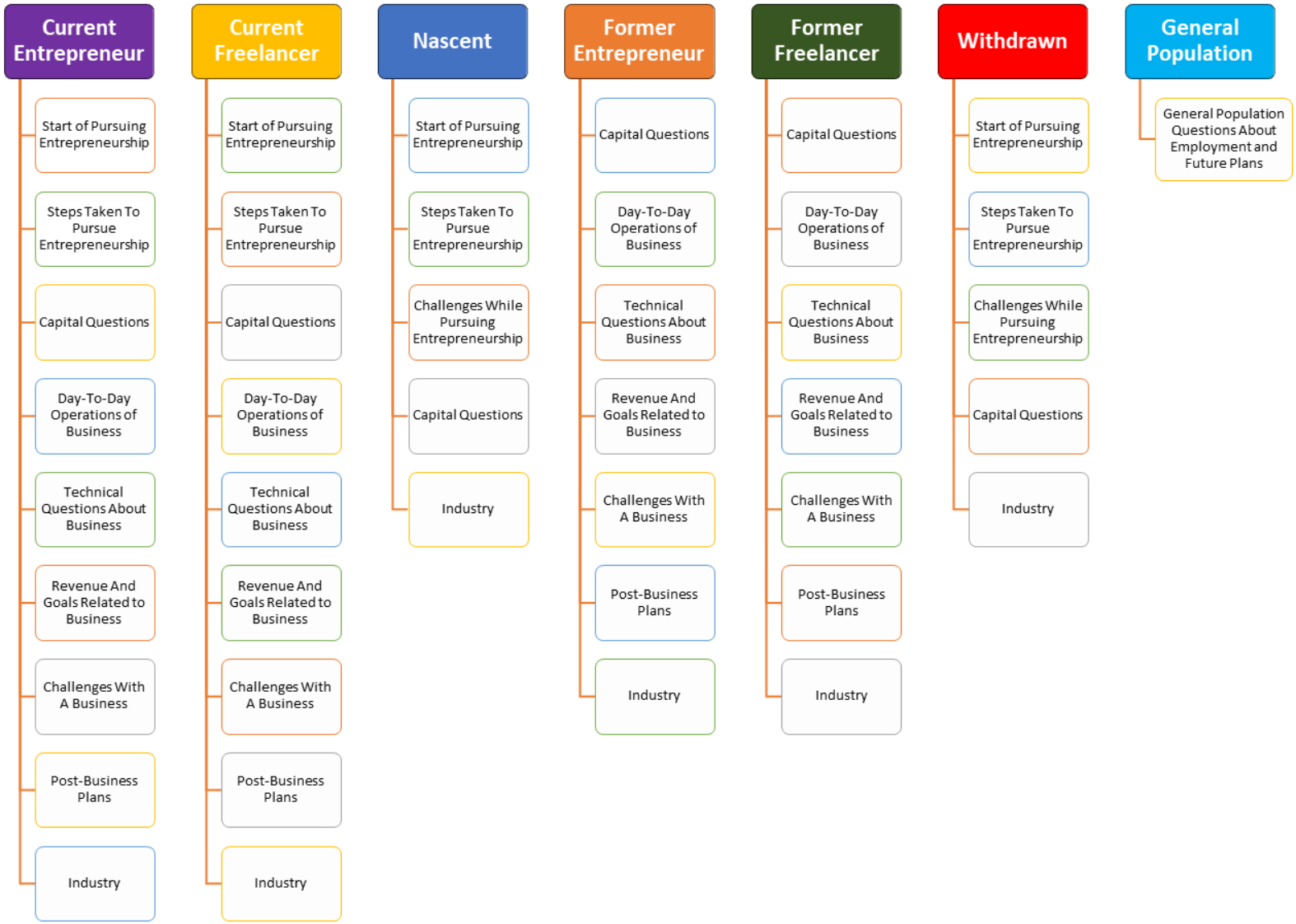
Abraham, K.G. and Amaya, A., (2019). Probing for Informal Work Activity. *Journal of Official Statistics*, 35(3), pp.487-508. https://doi.org/10.2478/jos-2019-0021.

"Entrepreneurship in the Population (EPOP) Survey Project Questionnaire: 2022." NORC at the University of Chicago. October 12, 2022. EPOP.norc.org.

"Entrepreneurship in the Population (EPOP) Survey Project Questionnaire: 2023." NORC at the University of Chicago. October 12, 2023. EPOP.norc.org.

APPENDIX

APPENDIX A: ENTREPRENEURSHIP PATHWAYS AND TOPICAL AREAS IN EPOP:2024



APPENDIX B: SURVEY QUESTIONNAIRE CHANGES FROM EPOP:2023 TO EPOP:2024

#	Change	Variable	Section	Type	Category
1	Added new item that asks about the number of owners for the business (for current and former business owners).	BO_OWNERSHIP_1	Business Operations	New Item	New Item
2	Added new item that follows up BO_OWNERSHIP_1 to ask about the number of owners if respondent selects "Self and other."	BO_OWNERSHIP_2	Business Operations	New Item	New Item
3	Added new item that asks about motivation for starting own business/freelancing.	PE_MOTIVE_1	Pursuing Entrepreneurship	New Item	New Item
4	Added new disability item to demographics that asks about difficulties with seeing, hearing, walking, etc.	DEM_DISABILITY_1	End Demographics	New Item	New Item
5	Added new disability item to demographics that asks if respondent receives disability income.	DEM_DISABILITY_2	End Demographics	New Item	New Item
6	Added new ADDFINANCE item asking about respondent's reason for seeking additional financing after start-up.	BO_ADDFINANCE_REASON	Business Operations	New Item	New Item
7	Added new item to end demos asking if respondent was previously incarcerated.	DEM_INCARCERATED	End Demographics	New Item	New Item
8	Updated response options to add "I did not need grant funding" option, removed "It would be difficult to adhere to the grant accounting and reporting requirements", and changed "The grant/s would only cover a portion of the funding needed" to "Available grant funding amount was too small."	PE_CAPITAL_5	Pursuing Entrepreneurship	Changed Item	Response Option Change
9	Updated response options to add "I did not need funding from a bank" option, removed "Did not want to take on debt", and combined "Expected unfavorable interest rate" and "Expected unfavorable loan repayment terms" into one item.	PE_CAPITAL_6	Pursuing Entrepreneurship	Changed Item	Response Option Change
10	Updated response options to add "I did not need funding from these sources" option, removed "Did not want to take on debt", and combined "Expected unfavorable interest rate" and "Expected unfavorable loan repayment terms" into one item.	PE_CAPITAL_7	Pursuing Entrepreneurship	Changed Item	Response Option Change
11	Updated response options to add "I did not need grant funding" option, removed "It would be difficult to adhere to the grant accounting and reporting requirements", and changed "The grant/s would only cover a portion of the funding needed" to "Available grant funding amount was too small."	BO_ADDFINANCE_6	Business Operations	Changed Item	Response Option Change

#	Change	Variable	Section	Type	Category
12	Updated response options to add "I did not need funding from a bank" option, removed "Did not want to take on debt", and combined "Expected unfavorable interest rate" and "Expected unfavorable loan repayment terms" into one item.	BO_ADDFINANCE_7	Business Operations	Changed Item	Response Option Change
13	Updated response options to add "I did not need funding from these sources" option, removed "Did not want to take on debt", and combined "Expected unfavorable interest rate" and "Expected unfavorable loan repayment terms" into one item.	BO_ADDFINANCE_8	Business Operations	Changed Item	Response Option Change
14	Removed BO_EMPWAGES item.	BO_EMPWAGES	Business Operations	Removed Item	Removed Item
15	Removed BUSINESS_COSTS_1 item.	BUSINESS_COSTS_1	Business Operations	Removed Item	Removed Item
16	Removed BUSINESS_COSTS_2 item.	BUSINESS_COSTS_2	Business Operations	Removed Item	Removed Item
17	Removed GP_CONSIDER_4 item.	GP_CONSIDER_4	Non-Entrepreneurial	Removed Item	Removed Item
18	Removed GP_CONSIDER_5 item.	GP_CONSIDER_5	Non-Entrepreneurial	Removed Item	Removed Item
19	Removed GP_CONSIDER_6 item.	GP_CONSIDER_6	Non-Entrepreneurial	Removed Item	Removed Item

APPENDIX C: CROSS ROUND VARIABLE CHANGES TO REFERENCE WHEN COMPARING OR MERGING DATA FROM EPOP:2023 TO EPOP:2024

Table C-1. EPOP Round Specific Variable Names

EPOP Survey Round(s)	Prior Round Variable Name	EPOP:2024 Variable Name
2022 & 2023	PE_CAPITAL_5_1	PE_CAPITAL_5_1_Y3
2022 & 2023	PE_CAPITAL_5_2	PE_CAPITAL_5_2_Y3
2022 & 2023	PE_CAPITAL_5_3	PE_CAPITAL_5_3_Y3
2022 & 2023	PE_CAPITAL_5_4	PE_CAPITAL_5_4_Y3
2022 & 2023	PE_CAPITAL_5_5	*n/a: Item not included*
Not available	*n/a: Item not included*	PE_CAPITAL_5_5_Y3
2022 & 2023	PE_CAPITAL_5_6	*n/a: Item not included*
2022 & 2023	PE_CAPITAL_5_7	PE_CAPITAL_5_6_Y3
Not available	*n/a: Item not included*	PE_CAPITAL_5_7_Y3
2022 & 2023	PE_CAPITAL_6_1	PE_CAPITAL_6_1_Y3
2022 & 2023	PE_CAPITAL_6_2	PE_CAPITAL_6_2_Y3
2022 & 2023	PE_CAPITAL_6_3	*n/a: Item not included*
Not available	*n/a: Item not included*	PE_CAPITAL_6_3_Y3
2022 & 2023	PE_CAPITAL_6_4	*n/a: Item not included*
2022 & 2023	PE_CAPITAL_6_5	*n/a: Item not included*
2022 & 2023	PE_CAPITAL_6_6	PE_CAPITAL_6_4_Y3
2022 & 2023	PE_CAPITAL_6_7	PE_CAPITAL_6_5_Y3
Not available	*n/a: Item not included*	PE_CAPITAL_6_6_Y3
2022 & 2023	PE_CAPITAL_7_1	PE_CAPITAL_7_1_Y3
2022 & 2023	PE_CAPITAL_7_2	PE_CAPITAL_7_2_Y3
2022 & 2023	PE_CAPITAL_7_3	n/a: Item not included
Not available	n/a: Item not included	PE_CAPITAL_7_3_Y3
2022 & 2023	PE_CAPITAL_7_4	n/a: Item not included
2022 & 2023	PE_CAPITAL_7_5	n/a: Item not included

EPOP Survey Round(s)	Prior Round Variable Name	EPOP:2024 Variable Name
2022 & 2023	PE_CAPITAL_7_6	PE_CAPITAL_7_4_Y3
2022 & 2023	PE_CAPITAL_7_7	PE_CAPITAL_7_5_Y3
Not available	n/a: Item not included	PE_CAPITAL_7_6_Y3
2022 & 2023	BO_ADDFINANCE_6_1	BO_ADDFINANCE_6_1_Y3
2022 & 2023	BO_ADDFINANCE_6_2	BO_ADDFINANCE_6_2_Y3
2022 & 2023	BO_ADDFINANCE_6_3	BO_ADDFINANCE_6_3_Y3
2022 & 2023	BO_ADDFINANCE_6_4	BO_ADDFINANCE_6_4_Y3
2022 & 2023	BO_ADDFINANCE_6_5	n/a: Item not included
Not available	n/a: Item not included	BO_ADDFINANCE_6_5_Y3
2022 & 2023	BO_ADDFINANCE_6_6	n/a: Item not included
2022 & 2023	BO_ADDFINANCE_6_7	BO_ADDFINANCE_6_6_Y3
2022 & 2023	BO_ADDFINANCE_7_1	BO_ADDFINANCE_7_1_Y3
2022 & 2023	BO_ADDFINANCE_7_2	BO_ADDFINANCE_7_2_Y3
Not available	n/a: Item not included	BO_ADDFINANCE_7_3_Y3
2022 & 2023	BO_ADDFINANCE_7_3	n/a: Item not included
2022 & 2023	BO_ADDFINANCE_7_4	n/a: Item not included
2022 & 2023	BO_ADDFINANCE_7_5	n/a: Item not included
2022 & 2023	BO_ADDFINANCE_7_6	BO_ADDFINANCE_7_4_Y3
2022 & 2023	BO_ADDFINANCE_7_7	BO_ADDFINANCE_7_5_Y3
2022 & 2023	BO_ADDFINANCE_8_1	BO_ADDFINANCE_8_1_Y3
2022 & 2023	BO_ADDFINANCE_8_2	BO_ADDFINANCE_8_2_Y3
Not available	n/a: Item not included	BO_ADDFINANCE_8_3_Y3
2022 & 2023	BO_ADDFINANCE_8_3	n/a: Item not included
2022 & 2023	BO_ADDFINANCE_8_4	n/a: Item not included
2022 & 2023	BO_ADDFINANCE_8_5	n/a: Item not included
2022 & 2023	BO_ADDFINANCE_8_6	BO_ADDFINANCE_8_4_Y3
2021	BO_CHALLENGE_2_1	BO_CHALLENGE_2_1_Y2

EPOP Survey Round(s)	Prior Round Variable Name	EPOP:2024 Variable Name
2021	BO_CHALLENGE_2_2	BO_CHALLENGE_2_2_Y2
2021	BO_CHALLENGE_2_3	BO_CHALLENGE_2_3_Y2
2021	BO_CHALLENGE_2_4	n/a: Item not included in Year 2
2021	BO_CHALLENGE_2_5	BO_CHALLENGE_2_4_Y2
2021	BO_CHALLENGE_3_1	BO_CHALLENGE_3_1_Y2
2021	BO_CHALLENGE_3_2	BO_CHALLENGE_3_2_Y2
2021	BO_CHALLENGE_3_3	BO_CHALLENGE_3_3_Y2
not available	n/a: Item not included in Year 1	BO_CHALLENGE_3_4_Y2
2021	BO_CHALLENGE_3_4	BO_CHALLENGE_3_5_Y2
2021	BO_CHALLENGE_4_1	BO_CHALLENGE_4_1_Y2
2021	BO_CHALLENGE_4_2	BO_CHALLENGE_4_2_Y2
2021	BO_CHALLENGE_4_3	BO_CHALLENGE_4_3_Y2
2021	BO_CHALLENGE_4_4	BO_CHALLENGE_4_4_Y2
2021	BO_CHALLENGE_4_5	BO_CHALLENGE_4_5_Y2
2021	BO_CHALLENGE_4_6	BO_CHALLENGE_4_6_Y2
2021	BO_CHALLENGE_4_7	BO_CHALLENGE_4_7_Y2
2021	BO_CHALLENGE_4_8	BO_CHALLENGE_4_8_Y2
not available	n/a: Item not included in Year 1	BO_CHALLENGE_4_9_Y2
not available	n/a: Item not included in Year 1	BO_CHALLENGE_4_10_Y2
2021	BO_CHALLENGE_4_9	BO_CHALLENGE_4_11_Y2
2021	BO_CHALLENGE_5_1	BO_CHALLENGE_5_1_Y2
2021	BO_CHALLENGE_5_2	BO_CHALLENGE_5_2_Y2
2021	BO_CHALLENGE_5_3	BO_CHALLENGE_5_3_Y2
2021	BO_CHALLENGE_5_4	BO_CHALLENGE_5_4_Y2
2021	BO_CHALLENGE_5_5	BO_CHALLENGE_5_5_Y2
not available	n/a: Item not included in Year 1	BO_CHALLENGE_5_6_Y2
2021	BO_CHALLENGE_5_6	BO_CHALLENGE_5_7_Y2

EPOP Survey Round(s)	Prior Round Variable Name	EPOP:2024 Variable Name
2021	BO_CHALLENGE_ENDA	BO_CHALLENGE_ENDA_Y2
2021	BO_CHALLENGE_ENDB	BO_CHALLENGE_ENDB_Y2
2021	BO_CHALLENGE_ENDC	BO_CHALLENGE_ENDC_Y2
2021	BO_CHALLENGE_ENDD	BO_CHALLENGE_ENDD_Y2
2021	BO_CHALLENGE_ENDE	BO_CHALLENGE_ENDE_Y2
2021	BO_CHALLENGE_ENDF	BO_CHALLENGE_ENDF_Y2
2021	BO_CHALLENGE_ENDG	BO_CHALLENGE_ENDG_Y2
2021	BO_CHALLENGE_ENDH	BO_CHALLENGE_ENDH_Y2
2021	BO_CHALLENGE_ENDI	BO_CHALLENGE_ENDI_Y2
2021	BO_CHALLENGE_ENDJ	BO_CHALLENGE_ENDJ_Y2
2021	BO_CHALLENGE_ENDK	BO_CHALLENGE_ENDK_Y2
2021	BO_CHALLENGE_ENDL	BO_CHALLENGE_ENDL_Y2
2021	BO_CHALLENGE_ENDM	BO_CHALLENGE_ENDM_Y2
not available	n/a: Item not included in Year 1	BO_CHALLENGE_ENDN_Y2
2021	BO_CHALLENGE_ENDN	BO_CHALLENGE_ENDO_Y2
2021	BO_CHALLENGE_ENDO	BO_CHALLENGE_ENDP_Y2
2021	BO_CHALLENGE_ENDP	BO_CHALLENGE_ENDQ_Y2
2021	BO_CHALLENGE_ENDQ	BO_CHALLENGE_ENDR_Y2
2021	BO_CHALLENGE_ENDR	BO_CHALLENGE_ENDS_Y2
2021	BO_CHALLENGE_ENDS	BO_CHALLENGE_ENDT_Y2
2021	BO_CHALLENGE_ENDT	BO_CHALLENGE_ENDU_Y2
2021	BO_CHALLENGE_ENDU	BO_CHALLENGE_ENDV_Y2
not available	n/a: Item not included in Year 1	BO_CHALLENGE_ENDW_Y2
not available	n/a: Item not included in Year 1	BO_CHALLENGE_ENDX_Y2
2021	BO_CHALLENGE_ENDV	BO_CHALLENGE_ENDY_Y2
2021	BO_CHALLENGE_ENDW	BO_CHALLENGE_ENDZ_Y2
2021	BO_CHALLENGE_ENDX	BO_CHALLENGE_ENDAA_Y2

EPOP Survey Round(s)	Prior Round Variable Name	EPOP:2024 Variable Name
2021	BO_CHALLENGE_ENDY	BO_CHALLENGE_ENDBB_Y2
2021	BO_CHALLENGE_ENDZ	BO_CHALLENGE_ENDCC_Y2
not available	n/a: Item not included in Year 1	BO_CHALLENGE_ENDDD_Y2
2021	DEM_HOUSEHOLD_1_RUF	n/a: Item not included in Year 2; refer to DEM_HOUSENUM_DRV
2021	DEM_HOUSEHOLD_2_RUF	DEM_HOUSEHOLD_1_Y2_DRV
2021	DEM_HOUSEHOLD_3_RUF	DEM_HOUSEHOLD_2_Y2_DRV
2021	DEM_HOUSEHOLD_4_RUF	DEM_HOUSEHOLD_3_Y2_DRV
2021	DEM_HOUSEHOLD_5_RUF	DEM_HOUSEHOLD_4_Y2_DRV
2021	DEM_HOUSENUM_RUF	DEM_HOUSENUM_DRV
2021	GP_CONSIDER_1_1	GP_CONSIDER_1_1_Y2
2021	GP_CONSIDER_1_2	GP_CONSIDER_1_2_Y2
2021	GP_CONSIDER_1_3	GP_CONSIDER_1_3_Y2
2021	GP_CONSIDER_1_4	GP_CONSIDER_1_4_Y2
2021	GP_CONSIDER_1_5	GP_CONSIDER_1_5_Y2
2021	GP_CONSIDER_1_6	GP_CONSIDER_1_6_Y2
2021	GP_CONSIDER_1_7	GP_CONSIDER_1_7_Y2
2021	GP_CONSIDER_1_8	GP_CONSIDER_1_8_Y2
2021	GP_CONSIDER_1_9	GP_CONSIDER_1_9_Y2
2021	GP_CONSIDER_1_10	GP_CONSIDER_1_10_Y2
2021	GP_CONSIDER_1_11	GP_CONSIDER_1_11_Y2
2021	GP_CONSIDER_1_12	GP_CONSIDER_1_12_Y2
2021	GP_CONSIDER_1_13	GP_CONSIDER_1_13_Y2
2021	GP_CONSIDER_1_14	GP_CONSIDER_1_14_Y2
2021	GP_CONSIDER_1_15	GP_CONSIDER_1_15_Y2
2021	GP_CONSIDER_1_16	GP_CONSIDER_1_16_Y2
not available	n/a: Item not included in Year 1	GP_CONSIDER_1_17_Y2
2021	GP_CONSIDER_1_17	GP_CONSIDER_1_18_Y2

EPOP Survey Round(s)	Prior Round Variable Name	EPOP:2024 Variable Name
2021	PE_CHALLENGE_1_1	PE_CHALLENGE_1_1_Y2
2021	PE_CHALLENGE_1_2	PE_CHALLENGE_1_2_Y2
2021	PE_CHALLENGE_1_3	PE_CHALLENGE_1_3_Y2
2021	PE_CHALLENGE_1_4	PE_CHALLENGE_1_4_Y2
2021	PE_CHALLENGE_1_5	PE_CHALLENGE_1_5_Y2
not available	n/a: Item not included in Year 1	PE_CHALLENGE_1_6_Y2
not available	n/a: Item not included in Year 1	PE_CHALLENGE_1_7_Y2
2021	PE_CHALLENGE_1_6	PE_CHALLENGE_1_8_Y2

Table C-2. EPOP Round Specific Variable Code Frames Across EPOP 2022, 2023, and 2024

EPOP:2022 Variable Name	EPOP:2023 Variable Name	EPOP:2024 Variable Name	EPOP:2022 Response Values	EPOP:2023 Response Values	EPOP:2024 Response Values
BO_STARTBIZ_1_PUF	BO_STARTBIZ_1_Y2_PUF	BO_STARTBIZ_1_Y3_PUF	1: Before 1980	1: Before 1990	1: Before 1990
			2: 1980-1989	1: Before 1990	
			3: 1990-1999	2: 1990-1999	2: 1990-1999
			4: 2000-2009	3: 2000-2009	3: 2000-2009
			5: 2010-2014	4: 2010-2014	4: 2010-2014
			6: 2015-2016	5. 2015-2016	5. 2015-2016
			7: 2017-2018	6: 2017-2018	6: 2017-2018
			8: 2019	7: 2019	7: 2019
			9: 2020	8: 2020	8: 2020
			10:2021	9: 2021	9: 2021
			n/a: value not included	10: 2022+	10:2022-2024
BO_STARTBIZ_1_RUF	BO_STARTBIZ_1_Y2_RUF	BO_STARTBIZ_1_Y3_RUF	1: Before 1970	1: Before 1980	1: Before 1980
			2: 1970-1979	1: Before 1980	
			3: 1980-1989	2: 1980-1989	2: 1980-1989
			4: 1990-1999	3: 1990-1999	3: 1990-1999
			5: 2000-2009	4: 2000-2009	4: 2000-2009
			6: 2010	5: 2010	5: 2010
			7: 2011	6: 2011	6: 2011
			8: 2012	7: 2012	7: 2012
			9: 2013	8: 2013	8: 2013
			10: 2014	9: 2014	9: 2014
			11: 2015	10: 2015	10: 2015
			12: 2016	11: 2016	11: 2016
			13: 2017	12: 2017	12: 2017
			14: 2018	13: 2018	13: 2018
			15: 2019	14: 2019	14: 2019
			16: 2020	15: 2020	15: 2020
			17: 2021	16: 2021	16: 2021
			n/a: value not included	17: 2022	17: 2022
			n/a: value not included	18: 2023	18: 2023
					19: 2024
BO_NUMEMPLOY_1_PUF	BO_NUMEMPLOY_1_Y2_PUF	BO_NUMEMPLOY_1_Y3_PUF	0: 0	0: 0	0: 0
			1: 1-4	1: 1-4	1: 1-9
			2: 5-9	2: 5-9	2: 10-49
			3: 10-19	3: 10-19	3: 50-99

EPOP:2022 Variable Name	EPOP:2023 Variable Name	EPOP:2024 Variable Name	EPOP:2022 Response Values	EPOP:2023 Response Values	EPOP:2024 Response Values
			4: 20-49	4: 20-49	4: 100-199
			5: 50-74	5: 50-99	5: 200+
			6: 75-99	5: 50-99	
			7: 100+	6: 100-199	
			7: 100+	7: 200+	

Table C-3. EPOP Round Specific Variable Code Frames Across EPOP:2022 and EPOP:2023 – EPOP:2024

EPOP:2022 Variable Name	EPOP:2023 & EPOP:2024 Variable Name	EPOP:2022 Response Values	EPOP:2023 & EPOP:2024 Response Values
BO_EXITSTRAT_1	BO_EXITSTRAT_1_Y2	1. Sold your business at a loss	1. Sold your business at a loss
		2. Sold your business at more or less break even	2. Sold your business at more or less break even
		3. Sold your business at a profit	3. Sold your business at a profit
		4. Bankruptcy or liquidation	4. Bankruptcy or liquidation
		5. Transferred business to a family member	5. Transferred business to a family member
		6. Did not complete any forms/paperwork, just stopped working or taking work	6. Did not complete any forms/paperwork, just stopped working or taking work
		n/a: value not included in Year 1	7. Transferred business to a non-family member
		n/a: value not included in Year 1	8. Converted the business to an employee ownership model
		7. Other, specify: [TEXTBOX]	9. Other, specify: [TEXTBOX]
S_FORMFREE_STAT_1	S_FORMFREE_STAT_1_Y2	1. Yes, I am still working for myself as a freelancer, consultant, or independent contractor	1. Yes
		2. No, I stopped working as a freelancer, consultant, or independent contractor within the last 5 years	2. No
		3. No, I stopped working as a freelancer, consultant, or independent contractor more than 5 years ago	2. No
S_GIGPLATFORM_DRV	S_GIGPLATFORM_Y2_DRV	1: Confirmed Gig Platform: Services	1: Confirmed Gig Platform: Services
		2: Confirmed Gig Platform: Selling/Renting of Goods	2: Confirmed Gig Platform: Selling/Renting of Goods

EPOP:2022 Variable Name	EPOP:2023 & EPOP:2024 Variable Name	EPOP:2022 Response Values	EPOP:2023 & EPOP:2024 Response Values
		3: Confirmed Gig Platform: Online Surveys	3: Confirmed Gig Platform: Online Surveys
		4: Payment Provider	4: Payment Provider
		5: Unconfirmed Gig Work	5: Unconfirmed/unlikely Gig Work
		6: Unlikely Gig Work	5: Unconfirmed/unlikely Gig Work
DEM_MARITAL_PUF	DEM_MARITAL_Y2_PUF	1: Married	1: Married/Cohabitating
		2: Widowed	2: Widowed/Divorced/Separated
		3: Divorced/Separated	2: Widowed/Divorced/Separated
		4: Single	3: Single
		5: Cohabitating	1: Married/Cohabitating
BO_NUMEMPLOY_1_RUF	BO_NUMEMPLOY_1_Y2_RUF	0: 0	0: 0
		1: 1	1: 1
		2: 2	2: 2
		3: 3	3: 3
		4: 4	4: 4
		5: 5	5: 5
		6: 6	6: 6
		7: 7	7: 7
		8: 8	8: 8
		9: 9	9: 9
		10: 10-14	10: 10-14
		11: 15-19	11: 15-19
		12: 20-29	12: 20-29
		13: 30-49	13: 30-49
		14: 50-74	14: 50-74
		15: 75-99	15: 75-99

EPOP:2022 Variable Name	EPOP:2023 & EPOP:2024 Variable Name	EPOP:2022 Response Values	EPOP:2023 & EPOP:2024 Response Values
		16: 100-199	16: 100-199
		17: 200+	17: 200-499
		17: 200+	18: 500-999
		17: 200+	19: 1000+
BO_REVENUE_1_PUF	BO_REVENUE_1_Y2_PUF	1: 0-99	0: 0
		1: 0-99	1: 1-99
		2: 100-499	2: 100-499
		3: 500-999	3: 500-999
		4: 1,000-4,999	4: 1,000-4,999
		5: 5,000-9,999	5: 5,000-9,999
		6: 10,000-24,999	6: 10,000-24,999
		7: 25,000-49,999	7: 25,000-49,999
		8: 50,000-74,999	8: 50,000-74,999
		9: 75,000-99,999	9: 75,000-99,999
		10: 100,000-249,999	10: 100,000-249,999
		11: 250,000-499,999	11: 250,000-499,999
		12: 500,000-999,999	12: 500,000+
		13: 1,000,000+	12: 500,000+
BO_REVENUE_2_PUF	BO_REVENUE_2_Y2_PUF	1: 0-99	0: 0
		1: 0-99	1: 1-99
		2: 100-499	2: 100-499
		3: 500-999	3: 500-999
		4: 1,000-4,999	4: 1,000-4,999
		5: 5,000-9,999	5: 5,000-9,999
		6: 10,000-24,999	6: 10,000-24,999
		7: 25,000-49,999	7: 25,000-49,999

EPOP:2022 Variable Name	EPOP:2023 & EPOP:2024 Variable Name	EPOP:2022 Response Values	EPOP:2023 & EPOP:2024 Response Values
		8: 50,000-74,999	8: 50,000-74,999
		9: 75,000-99,999	9: 75,000-99,999
		10: 100,000-249,999	10: 100,000-249,999
		11: 250,000-499,999	11: 250,000+
		12: 500,000+	11: 250,000+

NORC at the University of Chicago

© Copyright 2024. NORC at the University of Chicago