

## Accessing CMIE Consumer Pyramids Data [Restricted Access for Duke University Users]

With funding support from DCID, DUPRI, SSRI and generous individual contributions by faculty members from Fuqua (*Prof. Sharique Hasan*); Sanford (*Profs. Marc Jeuland, Manoj Mohanan, Marcos Rangel*) and Political Science (*Prof. Erik Wibbels*), we now have subscription access to the CMIE Consumer Pyramids Data. This document provides a brief summary about the data, documentation, and access.

This dataset, collected by Centre for Monitoring Indian Economy (CMIE) is the world's largest household panel survey; it interviews over 174,000 households (and all members) three times a year since the panel was started in 2014. As of May 2021, the 23<sup>rd</sup> wave is currently being collected. Duke users will have access to all waves of data collected (starting in 2014) until the end of our subscription period (May 2022).

### 1. Overview of data:

Main landing page <https://consumptionpyramidsdx.cmie.com>

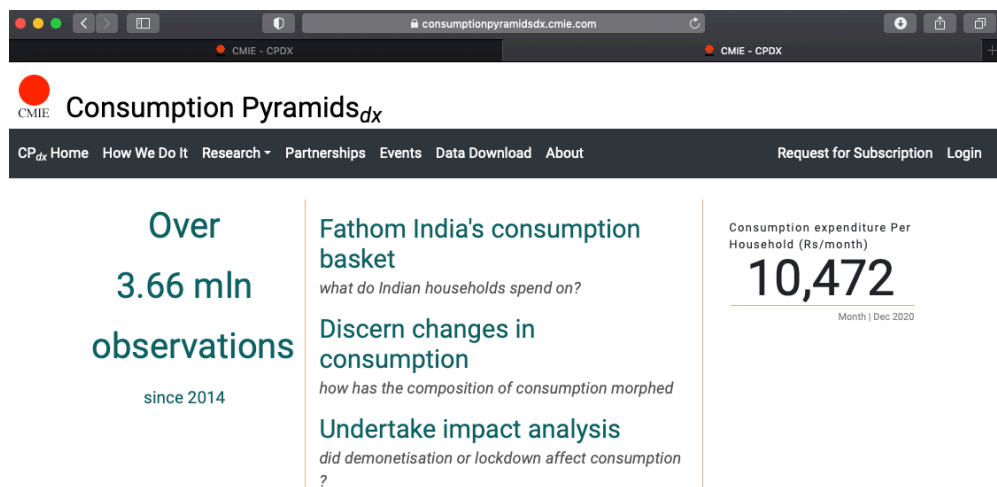
CMIE has produced a series of videos that will help understand the data, its panel structure, as well as answer questions about its implementation, data quality etc.:

- Survey Design: <https://youtu.be/MbCo2KxzvBM>
- Sample Survival / Response: <https://youtu.be/Lwkl8LJQP4k>
- Weights: <https://youtu.be/UQN20Ipz3rl>

Details on the data documentation are available on the website (click on "How we do it").

### 2. Accessing the data

**Please note:** CMIE data is a subscription-based service. Our subscription enables users at Duke University to access the data. If your research team includes co-authors from institutions outside Duke, they will need to get their own subscription in order to access the data.



CP<sub>dx</sub> Home How We Do It Research Partnerships Events Data Download About Request for Subscription Login

Over  
3.66 mln  
observations  
since 2014

Fathom India's consumption basket  
*what do Indian households spend on?*

Discern changes in consumption  
*how has the composition of consumption morphed*

Undertake impact analysis  
*did demonetisation or lockdown affect consumption?*

Consumption expenditure Per Household (Rs/month)  
**10,472**  
Month | Dec 2020

In order to access the data, you will need to first sign up / register – the link for registration is on the top banner "Request for Subscription".

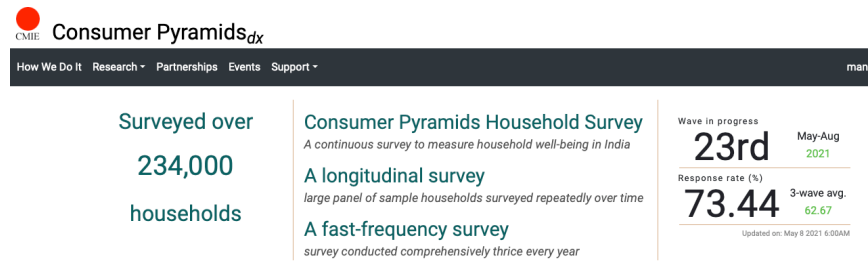
Once you register, you will get an email from [p.intimation@cmie.com](mailto:p.intimation@cmie.com) confirming your account creation, username and password.

## Institutional Data Access:

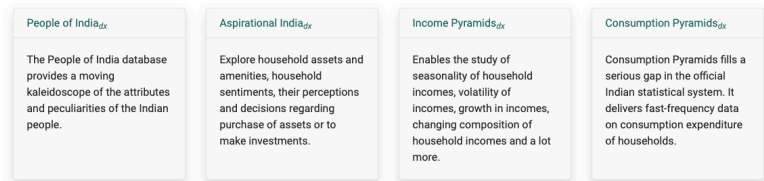
We have set up the institutional subscription so that our **users can access the data from all IP**

**addresses that are part of Duke University.** What this means is:

When you login using your credentials from anywhere on campus (university, not health system) you will be able to see a tab that allows you to download data. Once you log in from Duke servers:



### A fast-frequency portrayal of living standards of Indian households



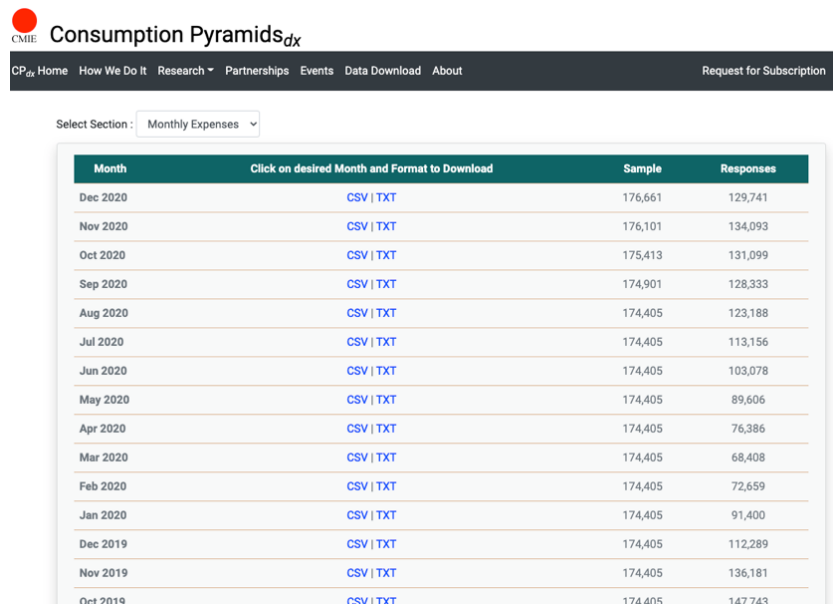
The **four boxes** link directly to 4 components / modules of this database. When you access the module for the first time, you will be asked to accept the user agreement.

- [People Of Indiadx](#) – demographic / employment, time use, financial assets
- [Aspirational Indiadx](#) – assets, perceptions, expectations,
- [Income Pyramidsdx](#) - Household / Member Income
- [Consumption Pyramidsdx](#) - Monthly/Weekly expenses

Full list of variables: <https://consumerpyramidsdx.cmie.com/kommon/bin/sr.php?kall=wlist>

After clicking on the database, you will see a tab for “data download”. Data from all waves of that module will be available as either CSV or TXT files.. for example:

Download, merge, and have fun researching!!



Month	Click on desired Month and Format to Download	Sample	Responses
Dec 2020	<a href="#">CSV</a>   <a href="#">TXT</a>	176,661	129,741
Nov 2020	<a href="#">CSV</a>   <a href="#">TXT</a>	176,101	134,093
Oct 2020	<a href="#">CSV</a>   <a href="#">TXT</a>	175,413	131,099
Sep 2020	<a href="#">CSV</a>   <a href="#">TXT</a>	174,901	128,333
Aug 2020	<a href="#">CSV</a>   <a href="#">TXT</a>	174,405	123,188
Jul 2020	<a href="#">CSV</a>   <a href="#">TXT</a>	174,405	113,156
Jun 2020	<a href="#">CSV</a>   <a href="#">TXT</a>	174,405	103,078
May 2020	<a href="#">CSV</a>   <a href="#">TXT</a>	174,405	89,606
Apr 2020	<a href="#">CSV</a>   <a href="#">TXT</a>	174,405	76,386
Mar 2020	<a href="#">CSV</a>   <a href="#">TXT</a>	174,405	68,408
Feb 2020	<a href="#">CSV</a>   <a href="#">TXT</a>	174,405	72,659
Jan 2020	<a href="#">CSV</a>   <a href="#">TXT</a>	174,405	91,400
Dec 2019	<a href="#">CSV</a>   <a href="#">TXT</a>	174,405	112,289
Nov 2019	<a href="#">CSV</a>   <a href="#">TXT</a>	174,405	136,181
Oct 2019	<a href="#">CSV</a>   <a href="#">TXT</a>	174,405	147,743

## Coordination Suggestion:

Once you have become a bit more familiar with the data structure, merging, and esp the weights, I would suggest we form a group at Duke to exchange notes on data management, coding, cleaning etc that will reduce inefficiencies and also help our respective research teams move faster. Also, perhaps we can form sub-groups of people interested in specific questions so we do not end up duplicating effort.

**Trouble shooting:**

- If you are not able to access data (**after registration and you are on a Duke server**), please scroll down to the bottom of the page and click on “diagnosis”.
- We do not think access via VPN works. The best way to access data remains a quick visit to campus 😊