

Details of problem statement: MPCE

Introduction:

The Household Consumption Expenditure Survey (HCES) is designed to collect information on consumption of goods and services by the households. The survey also collects some auxiliary information on household characteristics and demographic particulars of the members of the households. Information collected in HCES is useful for understanding the consumption pattern, standard of living and hence, well-being of the households. Besides, the data of the survey provides budget shares of different commodity groups that is used for preparation of the weighting diagram for compilation of official Consumer Price Indices (CPIs). The data collected in HCES is also utilized, inter-alia, for deriving various other macroeconomic indicators. The survey covered the whole of the Indian Union except a few inaccessible villages in the Andaman and Nicobar Islands. Conducted in 2022-23, HCES collected information from **8,723 villages** and **6,115 urban blocks** spread over the entire country covering **2,61,746 households (1,55,014 in rural areas and 1,06,732 in urban areas)**. The survey also collected some auxiliary information on household characteristics and demographic particulars of the members of the households in Questionnaire - HCQ: Household Characteristics.

Variables and Metadata:

Train and test data set are created by randomly splitting whole datasets into 80:20 ratio (209396 and 52350 Households respectively). Both train and test datasets have two-two files (HH_Level_train_data, Person_Level_train_data, HH_Level_test_data, Person_Level_test_data) each . One file represents household-level data, and the other file represents person-level data, from which participants can create household-level variables.

Variables in Household Level Datasets:

Variables	Type	Description	Remarks
HH_ID	Character	Household ID	
Sector	Category	Sector (i.e. Rural/Urban)	1: rural, 2: urban
State	Category	State	
NSS-Region	Category	NSS- Region	
District	Category	District	
Household Type	Category	Type of Household	
Religion of the head of the household	Category	Religion	Hinduism-1 Islam-2 Christianity –3 Sikhism-4 Jainism-5 Buddhism-6 Zoroastrianism-7

Variables	Type	Description	Remarks
			Others-9 Not reported-0
Social Group of the head of the household	Category	Social Group	scheduled tribe (ST)-1 scheduled caste (SC)-2 other backward class (OBC)-3 others-9 Not reported-0
HH Size (For FDQ)	Numeric	Size of Household	
NCO_3D	Category	3 digit National Classification of Occupation code 2015	
NIC_5D	Category	5 digit National Industrial Classification-2008	
Is_online_Clothing_Purchased_Last365	Category	Which online item purchased/Paid during last 365 days - Clothing	Yes: 1
Is_online_Footwear_Purchased_Last365	Category	Which online item purchased/Paid during last 365 days- Footwear	Yes: 1
Is_online_Furniture_fixturesPurchased_Last365	Category	Which online item purchased/Paid during last 365 days - Furniture & fixtures	Yes: 1
Is_online_Mobile_Handset_Purchased_Last365	Category	Which online item purchased/Paid during last 365 days - Mobile handset	Yes: 1
Is_online_Personal_Goods_Purchased_Last365	Category	Which online item purchased/Paid during last 365 days- Personal goods (laptop/PC, tablet, clock, watch, spectacles, contact lenses, etc.)	Yes: 1
Is_online_Recreation_Goods_Purchased_Last365	Category	Which online item purchased/Paid during last 365 days- Goods for recreation (TV, camera, pen-drive, musical instruments etc.)	
Is_online_Household_Appliances_Purchased_Last365	Category	Which online item purchased/Paid during last 365 days -	Yes: 1

Variables	Type	Description	Remarks
		Cooking & other household appliances	
Is_online_Crockery_Utensils_Purchased_Last365	Category	Which online item purchased/Paid during last 365 days - Crockery & utensils	Yes: 1
Is_online_Sports_Goods_Purchased_Last365	Category	Which online item purchased/Paid during last 365 days - Sports goods	Yes: 1
Is_online_Medical_Equipment_Purchased_Last365	Category	Which online item purchased/Paid during last 365 days- Medical equipment	Yes: 1
Is_online_Bedding_Purchased_Last365	Category	Which online item purchased/Paid during last 365 days- Bedding	Yes: 1
Is_HH_Have_Television	Category	Whether household possessed one or more item as on the date of the survey- Television	Yes: 1
Is_HH_Have_Radio	Category	Whether household possessed one or more item as on the date of the survey- Radio	Yes: 1
Is_HH_Have_Laptop_PC	Category	Whether household possessed one or more item as on the date of the survey - Laptop/PC	Yes: 1
Is_HH_Have_Mobile_handset	Category	Whether household possessed one or more item as on the date of the survey- Mobile handset	Yes: 1
Is_HH_Have_Bicycle	Category	Whether household possessed one or more item as on the date of the survey- Bicycle	Yes: 1
Is_HH_Have_Motorcycle_scooter	Category	Whether household possessed one or more item as on the date of the survey- Motorcycle, scooter	Yes: 1
Is_HH_Have_Motorcar_jeep_van	Category	Whether household possessed one or more item as on the	Yes: 1

Variables	Type	Description	Remarks
		date of the survey- Motor car/jeep/van	
Is_HH_Have_Trucks	Category	Whether household possessed one or more item as on the date of the survey- Trucks	Yes: 1
Is_HH_Have_Animal_cart	Category	Whether household possessed one or more item as on the date of the survey - Animal cart	Yes: 1
Is_HH_Have_Refrigerator	Category	Whether household possessed one or more item as on the date of the survey- Refrigerator	Yes: 1
Is_HH_Have_Washing_machine	Category	Whether household possessed one or more item as on the date of the survey- Washing machine	Yes: 1
Is_HH_Have_Airconditioner_aircooler	Category	Whether household possessed one or more item as on the date of the survey- Air conditioner/air cooler	Yes: 1
TotalExpense	Numeric	Total Expenditure of Household	

Variables in Person Level Data:

Variables	Type	Description	Remarks
HH_ID	Character	Household ID	
Person Srl No.	Character	Serial Number of Person	
Relation to head (code)	Categorical		relation to head: self-1, spouse of head-2, married child-3, spouse of married child-4, unmarried child-5, grandchild-6, father/mother/father-in-law/mother-in-law-7, brother/sister/brother-in-law/sister-inlaw/other relatives-8, servants/employees/other non-relatives-9
Gender	Numeric		
Age(in years)	Numeric		
Marital Status (code)	Categorical		marital status: never married – 1, currently married – 2, widowed – 3, divorced/separated – 4
Highest educational level attained (code)	Categorical		highest educational level attained: not literate (i.e. not able to read or write a simple message with understanding in any language) -01 literate with non-formal education (like NFEC, AEC, TLC, literate without any schooling, etc.)-02 literate with formal education o below primary-03 primary -04 upper primary/middle -05 secondary -06 higher secondary -07 diploma /certificate course (up to secondary)-08 diploma/certificate course (higher secondary)-10

Variables	Type	Description	Remarks
			diploma/certificate course (graduation & above)-11 graduate -12 post graduate and above -13
Total year of education completed	Numeric		
Whether used internet from any location during last 30 days	Categorical		yes - 1, no - 2
No. of days stayed away from home during last 30 days	Numeric		
No. of meals usually taken in a day	Numeric		
No. of meals taken during last 30 days from school, balwadi etc.	Numeric		
No. of meals taken during last 30 days from employer as perquisites or part of wage	Numeric		
No. of meals taken during last 30 days others	Numeric		
No. of meals taken during last 30 days on payment	Numeric		
No. of meals taken during last 30 days at home	Numeric		

References:

1. Imputing Poverty Indicators without Consumption Data: An Exploratory Analysis: **Hai-Anh H. Dang, Talip Kilic, Kseniya Abanokova, Calogero Carletto**
2. Evaluating Household Consumption Patterns: Comparative Analysis Using Ordinary Least Squares and Random Forest Regression Models: **En Lee, Thian Song Ong, Yvonne Lee**
3. An Analysis of the Determinants of Household Consumption Expenditure and Poverty in India: **Almas Heshmati, Esfandiar Maasoumi and Guanghua Wan**