MSM Assignment-1 REPORT

Based on PLFS 2020-21 Data, NSO

Dhruv Choudhary	**2021B3A73142H**
Jyotirmoy Singh	**2021B3A72513H**
Sthitaprajna	**2021B3A71082H**
Aditya Kumar Jain	**2021B3A82573H**
Mugdha Gupta	**2021B3A72724H**
C Varun	**2021B3AA3031H**
Ananya Agrawal	**2021B3AA1935H**
Rachit Vaghani	**2021B3AA1457H**

20-11-22

Report accompanied with Data Extraction and Preliminary Data Analysis

Data for Development

Ministry of Statistics and Programme Implementation
Government of India

Q3 a:- Calculate some summary statistics of the variables & comment on the same.

INTRODUCTION

Variables , we are calculating summary statistics for are as per variables concerned with PLFS:

1. Age
2. Years of formal education & Vocational Training
3. Earnings

Age (V020)

Mean :	31.22	Kurtosis :	2.43	Skewness :	0.45

- From the sample population, we can infer that the working population is relatively young since the median population is bit less than the mean population.
- The median age of the sample was found to be 29 which is consistent, since it is periodic labour force survey.
- Since the median and mean are very close we can infer the data is uniform and not very skewed, which is proved, since skewness is 0.45 and not much deviated by outliers.
- Since the kurtosis is less than 3, the dataset has lighter tails, than an ordinary normal distribtuion.

Years of Formal Education (V024) & Vocational Training(V026)

Mean :	6.88	Kurtosis :	1.92	Skewness :	0.09

- Only about 25 % percent of the population has passed at least the metric examinations.
- Since kurtosis is less than 3, there are only a few people who are either uneducated or study for a very long time (Lighter Tails).
- As 50% of the population has completed formal education till at least 8 std and has completed vocational training 6 years, we can infer that they are independent in terms of small jobs like plumbing, electrician etc.

Salaried Earnings (V134)

Mean :	1862.31	Kurtosis :	7819.71	Skewness :	38.70

- We can conclude that 90% of the people earn very less/negligible.
- Kurtosis is very high, which indicates they have heavy tails and many outliers.
- This shows very high uneven distribution of wealth.

Q3 b :- Make a plot of any set of relevant variables. Discuss the results

INTRODUCTION:

The plost describes the relationship between the type of work a person does and the Household expenditure of that person. The mean of a household's monthly expenditure is compared against the type of work a person does in two areas - rural and urban.

We also compared the mean expenditure between households in rural and urban areas.

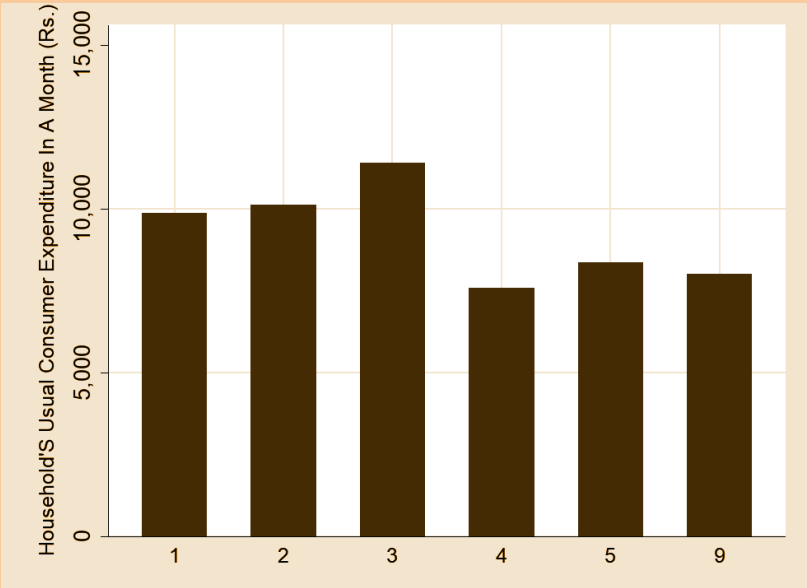
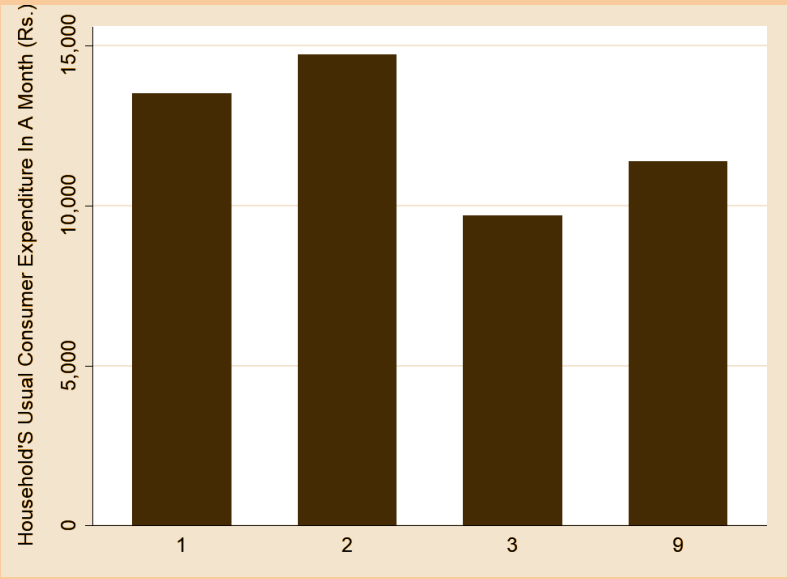


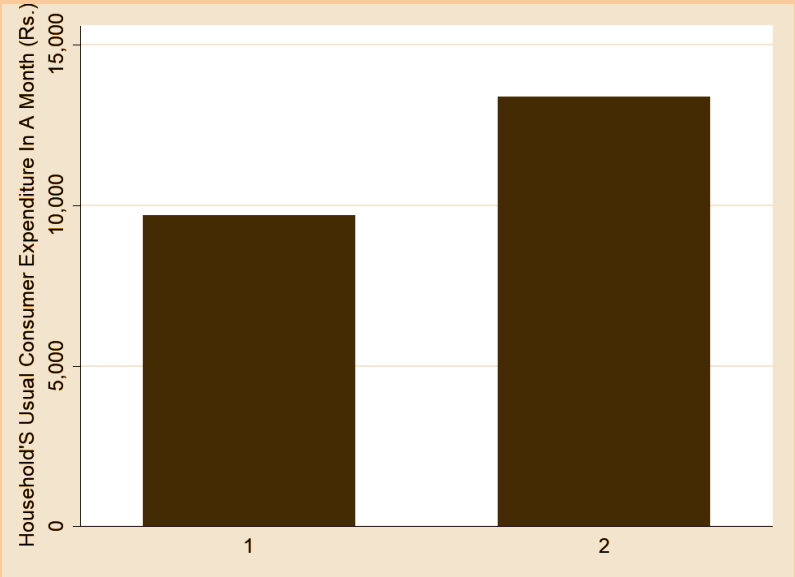
Fig 1.1

For rural areas:
1- self-employed in: agriculture,
2- non-agriculture ,
3-regular wage/salary earning,
4- casual labour in: agriculture,
5- non-agriculture ,
9-others



For urban areas:
1-self-employed,
2-regular wage/salary earning,
3-casual labour ,
9-others .

Fig 1.2



1-Rural , 2-Urban

Fig 1.3

RESULTS/ INTERPRETATION

Usual consumer expenditure (in Rs.) in rural and urban areas would give us fair idea about the income distribution and purchasing power of people in both the areas.

- With reference to Fig. 1.1, we can infer people who work as casual labourers in rural areas typically spend more each month on home essentials. The least expensive household expenses are those involved in agriculture. However, those who are doing non agriculture (i.e., own land) business and others have almost the same household expenses.

- With reference to Fig. 1.2, salary earners in urban regions typically have the greatest monthly household expenses, followed by self-employed individuals. Casual workers had the lowest monthly household expenditures, much like rural areas. The poor pay of casual workers in urban and rural areas may be one of the causes of this. An increase in their income could make it easier for them to meet their daily regular needs.

- With reference to Fig. 1.3, we discover that there is a sizable difference in household expenditure when we compare rural and urban areas. This disparity may be brought on by low household income in rural areas or by price variations in both areas for products and services.The cause is an income disparity, rural areas need to have more employment possibilities to help close the gap.

Q3 c:- State-level variation of a variable selected from the dataset and discuss the findings.

INTRODUCTION

Discussed underneath is the state-level variation of "**Spell of Unemployment".**

Here, we have compared spell of unemployment with State/UTs, which will help states to reassess their employment programs or enhance them in order to decrease unemployment rate and get rid of **frictional unemployment**.

Observations of for a few states (Kerala ,Nagaland, Daman and Diu) are discussed below:

Kerala

Most number of people who went through the spell of unemployment were from Kerala than from any other states and union territories. This was caused because of high population growth in the earlier period and low economic growth created a heavy backlog of unemployment in the State of Kerala.

Nagaland

Most people in Nagaland were unemployed for more than 3 years. Unemployment in Nagaland among the educated could be mostly attributed to lack of sufficient jobs. This could be due to very less industrialization and less jobs availabilty.

Daman and Diu

Daman & Diu had the least number of people undergoing the spell of unemployment in India. One of the prominent reasons behind it could be Daman and Diu being one of the best tourist attractions across India, generating thousands of jobs. Also, less population helped everyone get employed.

PS :- The code list for States and UT are attached at last.

Q3d: Choose any two or more variables in the NSS dataset and describe their characteristics in any manner you deem to be fit [example, state-level distribution, rural-urban split, across social groups, etc.]. The selected variables should be from different levels and interpret the results

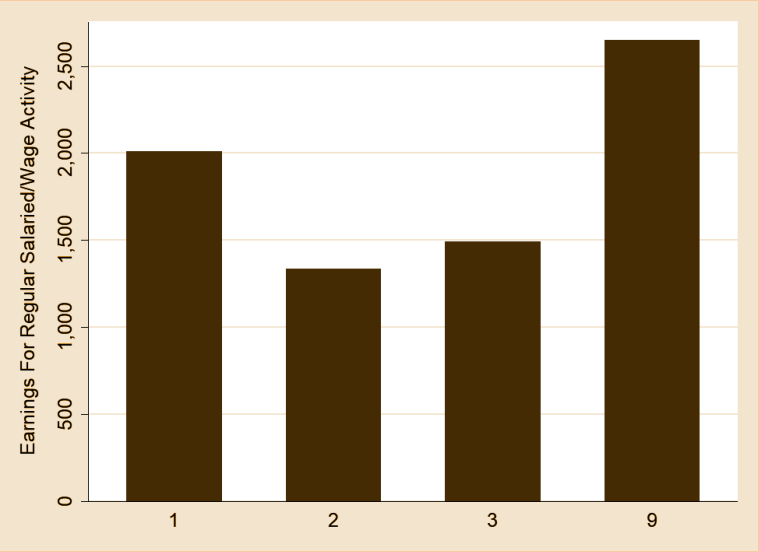
INTRODUCTION

Here we have used different variables such as V005 (Sector : Urban/Rural), V057(Spell of Unemployment) V151 (Reason for leaving the last place of employment), V134 (Earning of regular/salaried employment).

INFERENCE

- Most number of people are unemployed for a particular duration in urban as compared to rural area
- More number of people in urban area are unemployed for more than 5 years than people of rural area
- Less number of people in urban areas are unemployed for less than equal to 6 months as compared to that of rural areas. This means short term unemployment is a big problem in rural areas. Whereas long term unemployment is a threatening problem in rural areas.
- Most people have migrated in search of employment as well as loss of job/closure of units/lack of employment opportunities from rural area to urban area. Many people from rural areas have migrated for studies in that urban area.
- Very high number of people in rural areas have been displaced from rural areas mainly due to loss of job/lack of unemployment. This means that there is a lack of employment opportunities in rural areas, which shall be more focused upon.
- People from urban areas have left it mainly to pursue higher education. Even employment and universities in both areas could reduce migration from both rural and urban areas.
- Scheduled tribe peoples' mean salaried income is high as compared to OBCs, SC and STs. General category people earn way higher than backward class people. Scheduled caste people are earning much lesser than other backward classes.
- Thus, more attention should be paid on educating backward classes and giving them enough opportunities to explore.

Fig 1.4



Social Groups
1-scheduled tribe (ST),
2-scheduled caste (SC),
3-other backward class,
9-others

Q3e:Create one new variable based upon the variables given in the dataset and find out the association of this new variable with one of the existing variables.

INTRODUCTION

The new variable we define is “Earnings Per Hour” : Which is the summation of all earnings from casual labour activity ,salaried and self employed activities/ Total number of hours worked.

We use this value to compare the hourly wages and their stark difference between rural and urban areas.

This is an apt description and the wage gap between urban and rural India, and why rural India needs to catch up with urban India : Thus the association.

STATE CODE :

01	Jammu & kashmir	19	West Bengal
02	Himachal Pradesh	20	Jharkhand
03	Punjab	21	Odisha
04	Chandigarh	22	Chhattisgarh
05	Uttrakhand	23	Madhya Pradesh
06	Haryana	24	Gujarat
07	Delhi	25	Daman and Diu
08	Rajasthan	26	D & N Haveli
09	Uttar Pradesh	27	Maharashtra
10	Bihar	28	Andhra Pradesh
11	Sikkim	29	Karnataka
12	Arunachal Pradesh	30	Goa
13	Nagaland	31	Lakshadweep
14	Manipur	32	Kerala
15	Mizoram	33	Tamil Nadu
16	Tripura	34	Puducherry
17	Meghalaya	35	Andaman and Nicobar Island
18	Assam	36	Telangana