

PLFS

Pawan

2025-11-02

```
library(data.table)
library(readr)
library(tidyverse)
library(readxl)
library(gt)
options(digits=3)

## Read Data
fhhh1819 <- readRDS("/home/pawan/plfsdata/PLFS-18-19/data/hhfv201819.rds")
fvpp1819 <- readRDS("/home/pawan/plfsdata/PLFS-18-19/data/perfv201819.rds")
fhhh1920 <- readRDS("/home/pawan/plfsdata/PLFS-19-20/HHV1_2019_20.rds")
fvpp1920 <- readRDS("/home/pawan/plfsdata/PLFS-19-20/PERV1_2019_20.rds")

fhhh2021 <- readRDS("/home/pawan/plfsdata/PLFS-20-21/HHV1_2020_21.rds")
fvpp2021 <- readRDS("/home/pawan/plfsdata/PLFS-20-21/PERV1_2020_21.rds")

fhhh2122 <- readRDS("/home/pawan/plfsdata/PLFS-21-22/HHV1_2021_22.rds")
fvpp2122 <- readRDS("/home/pawan/plfsdata/PLFS-21-22/PERV1_2021_22.rds")

as.character(fvpp1819$visit) > fvpp1819$visit
as.character(fvpp1819$quarter) > fvpp1819$quarter
as.integer(fvpp1819$hhd_no) > fvpp1819$hhd_no

names(fvpp1819)[names(fvpp1819) == "42_status_code"] <- "status_code_sub"
names(fvpp1819)[names(fvpp1819) == "42_nic_code"] <- "nic_code_sub"
names(fvpp1819)[names(fvpp1819) == "42_nco_code"] <- "nco_code_sub"

as.numeric(fvpp1819$status_code) > fvpp1819$status_code
as.numeric(fvpp1819$status_code_sub) > fvpp1819$status_code_sub
```

```

ifelse(fvpp1819$status_code==11,11,
ifelse(fvpp1819$status_code==12,12,
ifelse(fvpp1819$status_code==21,21,
ifelse(fvpp1819$status_code==31,31,
ifelse(fvpp1819$status_code==41,41,
ifelse(fvpp1819$status_code==51,51,NA))))))->fvpp1819$upa_code

fvpp1819[!is.na(upsa_code),upsa_code:=upsa_code]
fvpp1819[is.na(upsa_code)&status_code_sub<=51,upsa_code:=status_code_sub]
fvpp1819[is.na(upsa_code),upsa_code:=9999]

ifelse(fvpp1819$status_code==11,fvpp1819$nic_code,
ifelse(fvpp1819$status_code==12,fvpp1819$nic_code,
ifelse(fvpp1819$status_code==21,fvpp1819$nic_code,
ifelse(fvpp1819$status_code==31,fvpp1819$nic_code,
ifelse(fvpp1819$status_code==41,fvpp1819$nic_code,
ifelse(fvpp1819$status_code==51,fvpp1819$nic_code,
ifelse(fvpp1819$status_code_sub==11,fvpp1819$nic_code_sub,
ifelse(fvpp1819$status_code_sub==12,fvpp1819$nic_code_sub,
ifelse(fvpp1819$status_code_sub==21,fvpp1819$nic_code_sub,
ifelse(fvpp1819$status_code_sub==31,fvpp1819$nic_code_sub,
ifelse(fvpp1819$status_code_sub==41,fvpp1819$nic_code_sub,
ifelse(fvpp1819$status_code_sub==51,fvpp1819$nic_code_sub,9999)))))))->fvpp1819$upsa_nic

substr(x=fvpp1819$upsa_nic,start=1,stop=2)->fvpp1819$upsa_nic
as.numeric(fvpp1819$upsa_nic)->fvpp1819$upsa_nic

ifelse(fvpp1819$upsa_nic<=3,"Agriculture",
ifelse(fvpp1819$upsa_nic>=10&fvpp1819$upsa_nic<=33,"Manufacturing",
ifelse(fvpp1819$upsa_nic>=41&fvpp1819$upsa_nic<=43,"Construction",
ifelse(fvpp1819$upsa_nic>=45&fvpp1819$upsa_nic<=99,"Services","Others"))))->fvpp1819$upsa_nic

factor(fvpp1819$upsa_nic_main,levels=c("Agriculture","Construction",
"Manufacturing","Services","Others"))->fvpp1819$upsa_nic

fvpp1819$upsa_nic_detail<-"Others"

ifelse(fvpp1819$upsa_nic==1,
ifelse(fvpp1819$upa==11&fvpp1819$upa_nic==1,"Self-employed",
ifelse(fvpp1819$upa==12&fvpp1819$upa_nic==1,"Self-employed",
ifelse(fvpp1819$upa==21&fvpp1819$upa_nic==1,"Self-employed",
ifelse(fvpp1819$upa==31&fvpp1819$upa_nic==1,"Long-term worker",

```

```

ifelse(fvpp1819$upa==41&fvpp1819$upa_nic==1,"Casual worker",
ifelse(fvpp1819$upa==51&fvpp1819$upa_nic==1,"Casual worker",
ifelse(fvpp1819$usa==11&fvpp1819$usa_nic==1,"Self-employed",
ifelse(fvpp1819$usa==12&fvpp1819$usa_nic==1,"Self-employed",
ifelse(fvpp1819$usa==21&fvpp1819$usa_nic==1,"Self-employed",
ifelse(fvpp1819$usa==31&fvpp1819$usa_nic==1,"Long-term worker",
ifelse(fvpp1819$usa==41&fvpp1819$usa_nic==1,"Casual worker",
ifelse(fvpp1819$usa==51&fvpp1819$usa_nic==1,"Casual worker",
fvpp1819$upsa_nic_detail))))))))),fvpp1819$upsa_nic_detail)->fvpp1819$upsa_nic_detail

ifelse(fvpp1819$upsa_nic==10,"Manufacture of food products & beverages",
fvpp1819$upsa_nic_detail)->fvpp1819$upsa_nic_detail

ifelse(fvpp1819$upsa_nic==11,"Manufacture of food products & beverages",
fvpp1819$upsa_nic_detail)->fvpp1819$upsa_nic_detail

ifelse(fvpp1819$upsa_nic==12,"Manufacture of tobacco products",
fvpp1819$upsa_nic_detail)->fvpp1819$upsa_nic_detail

ifelse(fvpp1819$upsa_nic==13,"Manufacture of textile & apparel",
fvpp1819$upsa_nic_detail)->fvpp1819$upsa_nic_detail

ifelse(fvpp1819$upsa_nic==14,"Manufacture of textile & apparel",
fvpp1819$upsa_nic_detail)->fvpp1819$upsa_nic_detail

ifelse(fvpp1819$upsa_nic==47,"Retail trade except motor vehicle",
fvpp1819$upsa_nic_detail)->fvpp1819$upsa_nic_detail

ifelse(fvpp1819$upsa_nic==85,"Education",
fvpp1819$upsa_nic_detail)->fvpp1819$upsa_nic_detail
ifelse(fvpp1819$upsa_nic==86,"Health care",
fvpp1819$upsa_nic_detail)->fvpp1819$upsa_nic_detail
ifelse(fvpp1819$upsa_nic==41,"Construction",
fvpp1819$upsa_nic_detail)->fvpp1819$upsa_nic_detail
ifelse(fvpp1819$upsa_nic==42,"Construction",
fvpp1819$upsa_nic_detail)->fvpp1819$upsa_nic_detail
ifelse(fvpp1819$upsa_nic==43,"Construction",
fvpp1819$upsa_nic_detail)->fvpp1819$upsa_nic_detail

factor(fvpp1819$upsa_nic_detail,levels=c("Self-employed",
"Long-term worker",

```

```

    "Casual worker",
    "Construction",
    "Manufacture of food products & beverages",
    "Manufacture of tobacco products",
    "Manufacture of textile & apparel",
    "Retail trade except motor vehicle",
    "Education",
    "Health care"))->fvpp1819$upsa_nic_detail

fvpp1819[age>=15,.(upsa_code,sector,sex,weight,upsa_nic_main,upsa_nic,upsa_nic_detail)]->t
ifelse(t$upsa_code!=9999,"Worker","Non-worker")->t$Worker
reshape2::melt(t,id=c("Worker","sector","sex"),m="weight")->t
reshape2::dcast(t,sex+sector~Worker,sum,margins=c("sector","sex","Worker"))->t5
round(t5$Worker*100/t5[,5],2)->t5$Worker

t5[c(1:6),c(2,1,4)]->t5
names(t5)[3] <- "2018-19"

as.character(fvpp1920$visit)->fvpp1920$visit
as.character(fvpp1920$quarter)->fvpp1920$quarter
as.integer(fvpp1920$hhd_no)->fvpp1920$hhd_no

names(fvpp1920)[names(fvpp1920) == "X42_status_code"] <- "status_code_sub"
names(fvpp1920)[names(fvpp1920) == "X42_nic_code"] <- "nic_code_sub"
names(fvpp1920)[names(fvpp1920) == "X42_nco_code"] <- "nco_code_sub"

as.numeric(fvpp1920$status_code)->fvpp1920$status_code
as.numeric(fvpp1920$status_code_sub)->fvpp1920$status_code_sub

ifelse(fvpp1920$status_code==11,11,
ifelse(fvpp1920$status_code==12,12,
ifelse(fvpp1920$status_code==21,21,
ifelse(fvpp1920$status_code==31,31,
ifelse(fvpp1920$status_code==41,41,
ifelse(fvpp1920$status_code==51,51,NA))))))->fvpp1920$upa_code

fvpp1920[!is.na(upa_code),upsa_code:=upa_code]
fvpp1920[is.na(upsa_code)&status_code_sub<=51,upsa_code:=status_code_sub]
fvpp1920[is.na(upsa_code),upsa_code:=9999]

```

```

ifelse(fvpp1920$status_code==11,fvpp1920$nic_code,
ifelse(fvpp1920$status_code==12,fvpp1920$nic_code,
ifelse(fvpp1920$status_code==21,fvpp1920$nic_code,
ifelse(fvpp1920$status_code==31,fvpp1920$nic_code,
ifelse(fvpp1920$status_code==41,fvpp1920$nic_code,
ifelse(fvpp1920$status_code==51,fvpp1920$nic_code,
ifelse(fvpp1920$status_code_sub==11,fvpp1920$nic_code_sub,
ifelse(fvpp1920$status_code_sub==12,fvpp1920$nic_code_sub,
ifelse(fvpp1920$status_code_sub==21,fvpp1920$nic_code_sub,
ifelse(fvpp1920$status_code_sub==31,fvpp1920$nic_code_sub,
ifelse(fvpp1920$status_code_sub==41,fvpp1920$nic_code_sub,
ifelse(fvpp1920$status_code_sub==51,fvpp1920$nic_code_sub,
ifelse(fvpp1920$status_code_sub==51,fvpp1920$nic_code_sub,9999)))))))))))>fvpp1920$upsa_nic

substr(x=fvpp1920$upsa_nic,start=1,stop=2)>fvpp1920$upsa_nic
as.numeric(fvpp1920$upsa_nic)>fvpp1920$upsa_nic

ifelse(fvpp1920$upsa_nic<=3,"Agriculture",
ifelse(fvpp1920$upsa_nic>=10&fvpp1920$upsa_nic<=33,"Manufacturing",
ifelse(fvpp1920$upsa_nic>=41&fvpp1920$upsa_nic<=43,"Construction",
ifelse(fvpp1920$upsa_nic>=45&fvpp1920$upsa_nic<=99,"Services","Others")))>fvpp1920$upsa_nic

factor(fvpp1920$upsa_nic_main,levels=c("Agriculture","Construction",
"Manufacturing","Services","Others"))>fvpp1920$upsa_nic

fvpp1920$upsa_nic_detail<-"Others"

ifelse(fvpp1920$upsa_nic==1,
ifelse(fvpp1920$upa==11&fvpp1920$upa_nic==1,"Self-employed",
ifelse(fvpp1920$upa==12&fvpp1920$upa_nic==1,"Self-employed",
ifelse(fvpp1920$upa==21&fvpp1920$upa_nic==1,"Self-employed",
ifelse(fvpp1920$upa==31&fvpp1920$upa_nic==1,"Long-term worker",
ifelse(fvpp1920$upa==41&fvpp1920$upa_nic==1,"Casual worker",
ifelse(fvpp1920$upa==51&fvpp1920$upa_nic==1,"Casual worker",
ifelse(fvpp1920$usa==11&fvpp1920$usa_nic==1,"Self-employed",
ifelse(fvpp1920$usa==12&fvpp1920$usa_nic==1,"Self-employed",
ifelse(fvpp1920$usa==21&fvpp1920$usa_nic==1,"Self-employed",
ifelse(fvpp1920$usa==31&fvpp1920$usa_nic==1,"Long-term worker",
ifelse(fvpp1920$usa==41&fvpp1920$usa_nic==1,"Casual worker",
ifelse(fvpp1920$usa==51&fvpp1920$usa_nic==1,"Casual worker",
fvpp1920$upsa_nic_detail))))))),fvpp1920$upsa_nic_detail)>fvpp1920$upsa_nic_detail

```

```

ifelse(fvpp1920$upsa_nic==10,"Manufacture of food products & beverages",
      fvpp1920$upsa_nic_detail)->fvpp1920$upsa_nic_detail

ifelse(fvpp1920$upsa_nic==11,"Manufacture of food products & beverages",
      fvpp1920$upsa_nic_detail)->fvpp1920$upsa_nic_detail

ifelse(fvpp1920$upsa_nic==12,"Manufacture of tobacco products",
      fvpp1920$upsa_nic_detail)->fvpp1920$upsa_nic_detail

ifelse(fvpp1920$upsa_nic==13,"Manufacture of textile & apparel",
      fvpp1920$upsa_nic_detail)->fvpp1920$upsa_nic_detail

ifelse(fvpp1920$upsa_nic==14,"Manufacture of textile & apparel",
      fvpp1920$upsa_nic_detail)->fvpp1920$upsa_nic_detail

ifelse(fvpp1920$upsa_nic==47,"Retail trade except motor vehicle",
      fvpp1920$upsa_nic_detail)->fvpp1920$upsa_nic_detail

ifelse(fvpp1920$upsa_nic==85,"Education",
      fvpp1920$upsa_nic_detail)->fvpp1920$upsa_nic_detail
ifelse(fvpp1920$upsa_nic==86,"Health care",
      fvpp1920$upsa_nic_detail)->fvpp1920$upsa_nic_detail
ifelse(fvpp1920$upsa_nic==41,"Construction",
      fvpp1920$upsa_nic_detail)->fvpp1920$upsa_nic_detail
ifelse(fvpp1920$upsa_nic==42,"Construction",
      fvpp1920$upsa_nic_detail)->fvpp1920$upsa_nic_detail
ifelse(fvpp1920$upsa_nic==43,"Construction",
      fvpp1920$upsa_nic_detail)->fvpp1920$upsa_nic_detail

factor(fvpp1920$upsa_nic_detail,levels=c("Self-employed",
                                         "Long-term worker",
                                         "Casual worker",
                                         "Construction",
                                         "Manufacture of food products & beverages",
                                         "Manufacture of tobacco products",
                                         "Manufacture of textile & apparel",
                                         "Retail trade except motor vehicle",
                                         "Education",
                                         "Health care"))->fvpp1920$upsa_nic_detail
fvpp1920[age>=15,.(upsa_code,sector,sex,weight,upsa_nic_main,upsa_nic,upsa_nic_detail)]->t
ifelse(t$upsa_code!=9999,"Worker","Non-worker")->t$Worker
reshape2::melt(t,id=c("Worker","sector","sex"),m="weight")->t

```

sex	sector	2018-19	2019-20
1	Total	71.0	73.0
1	Rural	72.2	74.4
1	Urban	68.5	69.9
2	Total	23.3	28.7
2	Rural	25.5	32.2
2	Urban	18.4	21.3

```

reshape2::dcast(t,sex+sector~Worker,sum,margins=c("sector","sex","Worker"))->t6
round(t6$Worker*100/t6[,5],2)->t6$Worker

t6[c(1:6),c(2,1,4)]->t6
names(t6)[3] <- "2019-20"

merge(t5,t6,by=c("sex","sector"))->t
levels(t$sector)<-c("Rural","Urban","Total")
##levels(t$sex)<-c("Men", "Women")
gt(t)

## fvhh1819[,UID:=paste(schedule,quarter,
##                         visit,sector,state,district,
##                         nss_region,stratum,sub_stratum,sub_sample,
##                         fod_sub_region,fsu,sample_sg_no,
##                         second_stage_stratum, hhd_no)]
```

```

## Calculate LFPR, WFPR, and UR for 2018-19
names(fvpp1819)[names(fvpp1819) == "42_status_code"] <- "status_code_sub"
names(fvpp1819)[names(fvpp1819) == "42_nic_code"] <- "nic_code_sub"
names(fvpp1819)[names(fvpp1819) == "42_nco_code"] <- "nco_code_sub"

fvpp1819[(age >= 15) & (status_code %in% c(11,12,21,31,41,42,51,81) |
           status_code_sub %in% c(11,12,21,31,41,42,51,81))] -> lf1819

fvpp1819[(age >= 15) & (status_code %in% c(11,12,21,31,41,42,51) |
           status_code_sub %in% c(11,12,21,31,41,42,51))] -> wf1819

lf1819[,.LFPR = round((sum(weight)/fvpp1819[(age >= 15),sum(weight)])*100,1))]->l1

wf1819[,.WFPR = round((sum(weight)/fvpp1819[(age >= 15),sum(weight)])*100,1))]->w1
```

```

lf1819$status_code %in% c(81) & is.na(status_code_sub)
,.(Unemployment_Rate=round((sum(weight)/lf1819[,sum(weight)])*100,1))]->ur1

cbind(l1, w1, ur1)->a1
a1$Year<-"2018-19"

## Calculate LFPR, WFPR, and UR for 2019-20
names(fvpp1920)[names(fvpp1920) == "X42_status_code"] <- "status_code_sub"
names(fvpp1920)[names(fvpp1920) == "X42_nic_code"] <- "nic_code_sub"
names(fvpp1920)[names(fvpp1920) == "X42_nco_code"] <- "nco_code_sub"

fvpp1920[(age >= 15) & (status_code %in% c(11,12,21,31,41,42,51,81) |
           status_code_sub %in% c(11,12,21,31,41,42,51,81))] -> lf1920
fvpp1920[(age >= 15) & (status_code %in% c(11,12,21,31,41,42,51) |
           status_code_sub %in% c(11,12,21,31,41,42,51))] -> wf1920

lf1920[,.LFPR = round((sum(weight)/fvpp1920[(age >= 15),sum(weight)])*100,1))]->l2

wf1920[,.WFPR = round((sum(weight)/fvpp1920[(age >= 15),sum(weight)])*100,1))]->w2

lf1920$status_code %in% c(81) & is.na(status_code_sub)
,.(Unemployment_Rate=round((sum(weight)/lf1920[,sum(weight)])*100,1))]->ur2

cbind(l2, w2, ur2)->a2
a2$Year<-"2019-20"

## Calculate LFPR, WFPR, and UR for 2020-21
names(fvpp2021)[names(fvpp2021) == "Status.Code...32"] <- "Status_Code"
names(fvpp2021)[names(fvpp2021) == "Industry.Code..NIC....33"] <- "NIC_Code"
names(fvpp2021)[names(fvpp2021) == "Occupation.Code..NCO....34"] <- "NCO_Code"

names(fvpp2021)[names(fvpp2021) == "Status.Code...43"] <- "Status_Code_Sub"
names(fvpp2021)[names(fvpp2021) == "Industry.Code..NIC....44"] <- "NIC_Code_Sub"
names(fvpp2021)[names(fvpp2021) == "Occupation.Code..NCO....45"] <- "NCO_Code_Sub"

fvpp2021[(Age >= 15) & (Status_Code %in% c(11,12,21,31,41,42,51,81) |
           Status_Code_Sub %in% c(11,12,21,31,41,42,51,81))] -> lf2021
fvpp2021[(Age >= 15) & (Status_Code %in% c(11,12,21,31,41,42,51) |
           Status_Code_Sub %in% c(11,12,21,31,41,42,51))] -> wf2021

```

```

lf2021[,.LFPR = round((sum(weight)/fvpp2021[(Age >= 15),sum(weight)])*100,1)]->l3
wf2021[,.WFPR = round((sum(weight)/fvpp2021[(Age >= 15),sum(weight)])*100,1)]->w3

lf2021[Status_Code %in% c(81) & is.na(Status_Code_Sub)
,.(Unemployment_Rate=round((sum(weight)/lf2021[,sum(weight)])*100,1))]->ur3

cbind(l3, w3, ur3)->a3
a3$Year<-"2020-21"

## Calculate LFPR, WFPR, and UR for 2021-22
names(fvpp2122)[names(fvpp2122) == "Status.Code...32"] <- "Status_Code"
names(fvpp2122)[names(fvpp2122) == "Industry.Code..NIC....33"] <- "NIC_Code"
names(fvpp2122)[names(fvpp2122) == "Occupation.Code..NCO....34"] <- "NCO_Code"

names(fvpp2122)[names(fvpp2122) == "Status.Code...45"] <- "Status_Code_Sub"
names(fvpp2122)[names(fvpp2122) == "Industry.Code..NIC....46"] <- "NIC_Code_Sub"
names(fvpp2122)[names(fvpp2122) == "Occupation.Code..NCO....47"] <- "NCO_Code_Sub"

fvpp2122[(Age >= 15) & (Status_Code %in% c(11,12,21,31,41,42,51,81) |
           Status_Code_Sub %in% c(11,12,21,31,41,42,51,81))] -> lf2122
fvpp2122[(Age >= 15) & (Status_Code %in% c(11,12,21,31,41,42,51) |
           Status_Code_Sub %in% c(11,12,21,31,41,42,51))] -> wf2122

lf2122[,.LFPR = round((sum(weight)/fvpp2122[(Age >= 15),sum(weight)])*100,1)]->l4
wf2122[,.WFPR = round((sum(weight)/fvpp2122[(Age >= 15),sum(weight)])*100,1)]->w4

lf2122[Status_Code %in% c(81) & is.na(Status_Code_Sub)
,.(Unemployment_Rate=round((sum(weight)/lf2122[,sum(weight)])*100,1))]->ur4

cbind(l4, w4, ur4)->a4
a4$Year<-"2021-22"
rbind(a1, a2, a3, a4)->e1
gt(e1)

```

LFPR	WFPR	Unemployment_Rate	Year
50.2	47.3	5.8	2018-19
53.5	50.9	4.8	2019-20
54.5	52.2	4.2	2020-21
55.2	52.9	4.1	2021-22