The Invisible Workforce: Insights from MOSPI's PLFS & TUS on Women's Domestic Contributions

Software and Codes

Software: R

Version: 2023.03.0

Codes:

Libraries:

```
library(foreign)
library(tidyverse)
library(ggalluvial)
```

Figure 1: Alluvial Plot depicting Participation and Average Time spent in a day

```
#Reading the consolidated data containing Time spent and Participation Rates from TUS Report-2019
dt_tus=read.csv("C:/MOSPI Hackathon/TUS_ConsolidatedTable.csv",header=T)
dt_tus$Category <- ifelse(dt_tus$Category  == "Unpaid Domestic Service Activities", "Unpaid\nDomestic\nServices",
                          "Unpaid\nCaregiving")
dt_tus$SubActivity=factor(dt_tus$SubActivity, levels=unique(dt_tus$SubActivity),labels = c("Cooking","Cleaning","DIY decor","Textile & \nfootwear care","Household \nmanage","Pet care","Shopping","Travelling","Others")) #Short names for sub-activities

# Alluvial Plot
ggplot(dt_tus, aes(axis1 = Gender, axis3 = SubActivity, y = TimeSpent)) +
  geom_alluvium(aes(fill = ParticipationRate), width = 1/12, knot.pos = 0.3) +
  geom_stratum(fill = "white", color = "black") +
  scale_fill_gradient(low = "lightblue", high = "darkblue", name = "Participation \nRate (%)") +
# Color gradient
  theme_bw() + theme( panel.grid.major = element_blank(),
                      panel.grid.minor = element_blank(), axis.line = element_line(colour = "black")) +
  labs(title = "Gender wise Breakdown of Unpaid Domestic Service Activities",
       caption = "*Flow width represents time spent, color shade represents participation rate",
       x = " ", y = "Average Time (in minutes) spent in a day")+
  geom_text(stat = "stratum", aes(label = after_stat(stratum)),
            size = 4, vjust = 0.5) +
```

```
  facet_wrap(~ Region)+theme(strip.text = element_text(size = 14, face =
"bold",color="darkgreen"),
                        plot.title = element_text(size = 18, face = "bold", hjust = 0.5,color="white"),
                        plot.caption=element_text(hjust=0,color="white",size=12),
                        axis.title.y = element_text( face = "bold", color = "white"))+
  theme( axis.text.x=element_blank(),axis.ticks.x=element_blank(),axis.text.y =
element_text(color="white"))+
  theme(plot.background = element_rect(fill = "#69922d"))
```

Figure 2: Region & Age-Group wise Estimated Wage per Hour for NIC Division-97 from PLFS 2019-20

```
#Reading the data

plfs_person_data1 <- read.dta("C:/MOSPI Hackathon/Data/PLFS_Data_2019-
20/Data_Tables/PERFV_2019-20.dta")#Person wise record for visit-1

#Compute Weekly Wage for Household Activity using PLFS data

wage=function(nic_code){

 N=nrow(plfs_person_data1)

 week_wage=rep(0,N)

 week_hours=rep(0,N)

 week_wage2=rep(0,N)

 week_hours2=rep(0,N)

 for(i in 1:N){


Activity_codes1=c(plfs_person_data1$b6q5_3pt1_Act1[i],plfs_person_data1$b6q5_3pt2_Act1[i],plfs
_person_data1$b6q5_3pt3_Act1[i],plfs_person_data1$b6q5_3pt4_Act1[i],plfs_person_data1$b6q5_
3pt5_Act1[i],plfs_person_data1$b6q5_3pt6_Act1[i],plfs_person_data1$b6q5_3pt7_Act1[i])

   index_Act1=which(Activity_codes1==nic_code)

   k=length(index_Act1)

   if(k>0){

    for(j in 1:k){

      r=index_Act1[j]

      if(r!=6 & r!=7){

      name_hour=paste("b6q6_3pt",r,"_Act1_per_fv",sep="")

      name_wage=paste("b6q9_3pt",r,"_Act1_per_fv",sep="")

      index_hour=which(colnames(plfs_person_data1)==name_hour)
```

```
      index_wage=which(colnames(plfs_person_data1)==name_wage)
      week_wage[i]= week_wage[i]+as.numeric(plfs_person_data1[i,index_wage])
      week_hours[i]= week_hours[i]+as.numeric(plfs_person_data1[i,index_hour])
      }else if(r==6){
        name_hour=paste("b6q6_3pt",r,"_Act1",sep="")
        name_wage=paste("b6q9_3pt",r,"_Act1",sep="")
        index_hour=which(colnames(plfs_person_data1)==name_hour)
        index_wage=which(colnames(plfs_person_data1)==name_wage)
        week_wage[i]= week_wage[i]+as.numeric(plfs_person_data1[i,index_wage])
        week_hours[i]= week_hours[i]+as.numeric(plfs_person_data1[i,index_hour])
      }else{
        name_hour=paste("b6q6_3pt",r,"_Act1",sep="")
        name_wage="b6q9_Act1_3pt7"
        index_hour=which(colnames(plfs_person_data1)==name_hour)
        index_wage=which(colnames(plfs_person_data1)==name_wage)
        week_wage[i]= week_wage[i]+as.numeric(plfs_person_data1[i,index_wage])
        week_hours[i]= week_hours[i]+as.numeric(plfs_person_data1[i,index_hour])
      }
     }
    }
```

```
Activity_codes2=c(plfs_person_data1$b6q5_3pt1_Act2[i],plfs_person_data1$b6q5_3pt2_Act2[i],plfs_person_data1$b6q5_3pt3_Act2[i],plfs_person_data1$b6q5_3pt4_Act2[i],plfs_person_data1$b6q5_3pt5_Act2[i],plfs_person_data1$b6q5_3pt6_Act2[i],plfs_person_data1$b6q5_3pt7_Act2[i])
   index_Act2=which(Activity_codes2==nic_code)
   k2=length(index_Act2)
   if(k2>0){
     for(j in 1:k2){
       r=index_Act2[j]
       if(r!=5 & r!=6 & r!=7){
         name_hour=paste("b6q6_3pt",r,"_Act2_per_fv",sep="")
```

```
      name_wage=paste("b6q9_3pt",r,"_Act2_per_fv",sep="")
      index_hour=which(colnames(plfs_person_data1)==name_hour)
      index_wage=which(colnames(plfs_person_data1)==name_wage)
      week_wage2[i]= week_wage2[i]+as.numeric(plfs_person_data1[i,index_wage])
      week_hours2[i]= week_hours2[i]+as.numeric(plfs_person_data1[i,index_hour])
    }else if(r==5){
      week_wage2[i]= week_wage2[i]+as.numeric(plfs_person_data1[i,98])
      week_hours2[i]= week_hours2[i]+as.numeric(plfs_person_data1[i,97])
    }else if(r==6){
      week_wage2[i]= week_wage2[i]+as.numeric(plfs_person_data1[i,108])
      week_hours2[i]= week_hours2[i]+as.numeric(plfs_person_data1[i,107])
    }else{
      week_wage2[i]= week_wage2[i]+as.numeric(plfs_person_data1[i,118])
      week_hours2[i]= week_hours2[i]+as.numeric(plfs_person_data1[i,117])
    }
   }
  }
 }

data2=data.frame("Sex"=plfs_person_data1$b4q5_per_fv,"Age"=plfs_person_data1$b4q6_per_fv,"Hours_Worked"=week_hours2,"Wages_Earned"=week_wage2,"Location"=plfs_person_data1[,36],"Education"=plfs_person_data1[,22])

data1=data.frame("Sex"=plfs_person_data1$b4q5_per_fv,"Age"=plfs_person_data1$b4q6_per_fv,"Hours_Worked"=week_hours,"Wages_Earned"=week_wage,"Location"=plfs_person_data1[,36],"Education"=plfs_person_data1[,22])

data=rbind(data1,data2)

# Categorizing into Rural and Urban

data <- data %>%
  mutate(AreaType = case_when(
    Location <= 19  ~ "Rural",
    Location > 19 & Location <= 29 ~ "Urban",
    TRUE ~ "No Fixed Workplace"  # For any unexpected values
```

```
 ))
 return(data)
 }
dt_hh_work=wage(nic_code = 97)
```

Figure 2 (a): Region wise Estimated Wage per Hour

```
##Ploting the estimated wage rate
index=which(dt_hh_work$Hours_Worked!=0 & dt_hh_work$Wages_Earned!=0)
index_female=which(dt_hh_work$Sex==2)
index_male=which(dt_hh_work$Sex==1)
index_rural=which(dt_hh_work$AreaType=="Rural")
index_urban=which(dt_hh_work$AreaType=="Urban")
index_male_wage_rural=intersect(intersect(index,index_male),index_rural)
index_male_wage_urban=intersect(intersect(index,index_male),index_urban)
index_female_wage_rural=intersect(intersect(index,index_female),index_rural)
index_female_wage_urban=intersect(intersect(index,index_female),index_urban)
index_female_wage_total=intersect(index,index_female)
index_male_wage_total=intersect(index,index_male)
average_wage=rep(0,6)
average_wage[1]=mean(dt_hh_work$Wages_Earned[index_female_wage_rural]/(dt_hh_work$Hours_Worked[index_female_wage_rural]))
average_wage[2]=dt_hh_work$Wages_Earned[index_female_wage_urban]/(dt_hh_work$Hours_Worked[index_female_wage_urban])
average_wage[3]=dt_hh_work$Wages_Earned[index_female_wage_total]/(dt_hh_work$Hours_Worked[index_female_wage_total])
average_wage[4]=dt_hh_work$Wages_Earned[index_male_wage_rural]/(dt_hh_work$Hours_Worked[index_male_wage_rural])
average_wage[5]=dt_hh_work$Wages_Earned[index_male_wage_urban]/(dt_hh_work$Hours_Worked[index_male_wage_urban])
average_wage[6]=dt_hh_work$Wages_Earned[index_male_wage_total]/(dt_hh_work$Hours_Worked[index_male_wage_total])

dt_wage=data.frame("Gender"=c(rep("Female",3),rep("Male",3)),
```

```
                "Region"=rep(c("Rural","Urban","Rural+Urban"),2),
"Average_Wage"=average_wage)
dt_wage$Region=factor(dt_wage$Region,levels=c("Rural","Urban","Rural+Urban"))
ggplot(dt_wage, aes(x = Average_Wage, y = Region)) +
 geom_line() +geom_point(aes(color = Gender), size = 3) +
 scale_color_manual(values = c("Male" = "blue", "Female" = "red"))+
 theme_bw() + labs(y=" ",x="Wages",title = "Estimated Average Wage per Hour")+ xlim(40,69)+
 geom_text(aes(label =
round(average_wage,2)),angle=45,size=3,vjust=0.5,hjust=1.2,color="darkblue")+
 theme(panel.grid.major = element_blank(),
   panel.grid.minor = element_blank(),
   axis.line = element_line(colour = "black"),
   plot.title = element_text(size = 16, face = "bold",hjust=0.5),  # Title formatting
   axis.text.x = element_text( size = 12, face = "bold") , axis.text.y = element_text( size = 12, face =
"bold") , axis.title.y = element_text( face = "bold", color = "darkolivegreen") , axis.title.x =
element_text(size=12))+theme(plot.background = element_rect(fill = "yellow"))+coord_flip()
```

Figure 2 (b): Age-Group wise Estimated Wage per Hour

```
wage_gender_age=function(dt_hh_work,gender,region){
 if(region!= "Rural" & region!="Urban"){
  dt_hh_work=dt_hh_work
 }else{
  dt_hh_work=dt_hh_work[dt_hh_work$AreaType==region,]
 }
  dt=dt_hh_work[dt_hh_work$Sex==gender,]
  dt$Age=as.numeric(dt$Age)
  dt21=dt[which(dt$Age>=15 & dt$Age<=29),]
  dt22=dt[which(dt$Age>=15 & dt$Age<=59),]
  dt23=dt[which(dt$Age>=60),]
  index1=which(dt21$Hours_Worked!=0 & dt21$Wages_Earned!=0)
  wages_per_hour1=(dt21$Wages_Earned[index1])/(dt21$Hours_Worked[index1])
  index2=which(dt22$Hours_Worked!=0 & dt22$Wages_Earned!=0)
  wages_per_hour2=(dt22$Wages_Earned[index2])/(dt22$Hours_Worked[index2])
  index3=which(dt23$Hours_Worked!=0 & dt23$Wages_Earned!=0)
  wages_per_hour3=(dt23$Wages_Earned[index3])/(dt23$Hours_Worked[index3])

average_wage=c(mean(wages_per_hour1),mean(wages_per_hour2),mean(wages_per_hour
3))
 return(average_wage)
```

```
 }

wage_male_rural=wage_gender_age(dt_hh_work,1,"Rural")#1-male, 2-female
wage_male_urban=wage_gender_age(dt_hh_work,1,"Urban")
wage_male_total=wage_gender_age(dt_hh_work,1,"Total")
wage_female_rural=wage_gender_age(dt_hh_work,2,"Rural")#1-male, 2-female
wage_female_urban=wage_gender_age(dt_hh_work,2,"Urban")
wage_female_total=wage_gender_age(dt_hh_work,2,"Total")
#Plot for Age wise Estimated Wage Rate
dt_wage_agewise=data.frame(Region=rep(c(rep("Rural",3),rep("Urban",3),rep("Rural+Urban
",3)),2),
                    Gender=c(rep("Male",9),rep("Female",9)),
                    Age_Group=rep(c("15-29","15-59","60+"),6),
                    Average_Wage=c(wage_male_rural,wage_male_urban,wage_male_total,
                            wage_female_rural,wage_female_urban,wage_female_total))
dt_wage_agewise$Region=factor(dt_wage_agewise$Region,levels=c("Rural","Urban","Rural
+Urban"))
dt_wage_agewise$Gender=factor(dt_wage_agewise$Gender,levels=c("Female","Male"))
ggplot(dt_wage_agewise, aes(x = Average_Wage, y = Age_Group)) +
 geom_line() +geom_point(aes(color = Gender), size = 3) +
 scale_color_manual(values = c("Male" = "blue", "Female" = "red")) + geom_text(aes(label =
round(dt_wage_agewise$Average_Wage,2)),angle=45,size=3,vjust=0.5,hjust=1.2,color="dark
blue")+
 facet_wrap(~Region) +  # Facet by Region
 labs(title = "Estimated Average Wage per Hour",
     y = "Age Groups",
     x = "Wage",
     fill = "Gender") +
 theme_bw() +  theme(
   panel.grid.major = element_blank(),
   panel.grid.minor = element_blank(),
   axis.line = element_line(colour = "black"),
   strip.text = element_text(size = 14, face = "bold", color = "darkgreen"),  # Larger facet
labels
   plot.title = element_text(size = 16, face = "bold",hjust=0.5),  # Title formatting
   axis.text.x = element_text( size = 12, face = "bold") ,
   axis.title.y = element_text( face = "bold", color = "darkolivegreen"),
   axis.title.x = element_text(face="bold",size=14))+
 theme(plot.background = element_rect(fill = "yellow"))+coord_flip()
```

Figure 3: Estimated Monthly Expenditure saved on Unpaid Domestic Activities

```
time_male_rural=c(94, 96, 115)#Time- Average time in minutes per day

time_male_urban=c(94, 93, 105)
```

```
time_male_total=c(94 ,95, 112)
time_female_rural=c(301 ,318, 247)
time_female_urban=c(264 ,306 ,238)
time_female_total=c(291 ,315 ,245)
#Participation Percentage
part_female_rural=c(88.7,93.6,79.1)/100
part_male_rural=c(25.3,31.2, 35.6)/100
part_female_urban=c(79.0 ,89.4 ,76.6)/100
part_male_urban=c(21.9,24.1, 32.6)/100
part_female_total=c(85.8,92.3, 78.3)/100
part_male_total=c(24.2 ,28.9 ,34.7)/100
#Average Income per day -- Time/60* Wage_per_Hour will be income for 1 day
mi_male_rural=wage_male_rural*time_male_rural*(30/60)*part_male_rural
mi_male_urban=wage_male_urban*time_male_urban*(30/60)*part_male_urban
mi_male_total=wage_male_total*time_male_total*(30/60)*part_male_total
mi_female_rural=wage_female_rural*time_female_rural*(30/60)*part_female_rural
mi_female_urban=wage_female_urban*time_female_urban*(30/60)*part_female_urban
mi_female_total=wage_female_total*time_female_total*(30/60)*part_female_total
Region=c(rep("Rural",6),rep("Urban",6),rep("Rural+Urban",6))
Gender=c(rep(c("Female","Male"),9))
Age_Group=rep(c(rep("15-29",2),rep("15-59",2),rep("60+",2)),3)
Monthly_income=c(mi_female_rural[1],mi_male_rural[1],
                 mi_female_rural[2],mi_male_rural[2],
                 mi_female_rural[3],mi_male_rural[3],
                 mi_female_urban[1],mi_male_urban[1],
                 mi_female_urban[2],mi_male_urban[2],
                 mi_female_urban[3],mi_male_urban[3],
                 mi_female_total[1],mi_male_total[1],
                 mi_female_total[2],mi_male_total[2],
                 mi_female_total[3],mi_male_total[3])
dt_income_plfs_tus=data.frame(Region,Gender,Age_Group,Monthly_income)
```

```
dt_income_plfs_tus$Region=factor(dt_income_plfs_tus$Region,levels=c("Rural","Urban","Rural+Urb
an"))

#Plot

ggplot(dt_income_plfs_tus, aes(x = Age_Group, y = Monthly_income, fill = Gender)) +

 geom_bar(stat = "identity", position = position_dodge(), color = "black",width=0.4) +

 scale_fill_manual(values = c("Male" = "blue", "Female" = "red")) + ylim(0,7600)+

 geom_text(aes(label =
round(dt_income_plfs_tus$Monthly_income,2)),angle=40,size=3,vjust=0,hjust=0.0,color="black")+

 facet_wrap(~Region) +  # Facet by Region

 labs(title = "Estimated Monthly Expenditure Saved for Unpaid Domestic Activities",

     x = "Age Groups",

     y = "Monthly Expenditure",

     fill = "Gender") +

 theme_bw() +  theme(

   panel.grid.major = element_blank(),

   panel.grid.minor = element_blank(),

   axis.line = element_line(colour = "black"),

   strip.text = element_text(size = 14, face = "bold", color = "black"),  # Larger facet labels

   plot.title = element_text(size = 16, face = "bold",hjust=0.5),  # Title formatting

   axis.text.x = element_text( size = 12, face = "bold") ,

   axis.title.y = element_text( face = "bold"),

   axis.title.x = element_text(face="bold",size=14))+

 theme(plot.background = element_rect(fill = "yellow"))
```

Figure 4: Comparison over different levels of Education

Figure 4 (a): Percentage of persons participating - TUS 2019

```
#Percentage of persons participating in different activities for different levels of education

dt_percent_participation=data.frame("Region"=c(rep("Rural",10),rep("Urban",10)),

"Education"=rep(c(rep("Not Literate",2),rep("Below
Primary",2),rep("Primary",2),rep("Middle",2),rep("Secondary and above",2)),2),

"Gender"=rep(c("Female","Male"),10),

"Participation_percent"=c(89.8,33,54.9,19,76.8,26.4,84.3,28.6,89.1,28.6,84.9,25,45.3,12.2,71.0,20.1,
79.8,22.4,86.3,24.7))
```

```
dt_percent_participation$Education=factor(dt_percent_participation$Education,levels=c("Not Literate","Below Primary","Primary","Middle","Secondary and above"))

ggp_percent_part=ggplot(dt_percent_participation, aes(x = Education, y = Participation_percent, fill = Gender)) +

 geom_bar(stat = "identity", position = position_dodge(), color = "black",width=0.4) +

 scale_fill_manual(values = c("Male" = "blue", "Female" = "red")) +  # Male = blue, Female = red

 geom_text(aes(label = round(Participation_percent,2)),angle=40,size=3,vjust=0,hjust=-0.15,color="black")+

 facet_wrap(~Region) +  # Facet by Region

 labs(title = "Percentage of persons participating in Unpaid Household Activity for different levels of education",

     x = "Education Level",

     y = "Percentage of persons participating",

     fill = "Gender") +

 theme_bw() +  theme(

   panel.grid.major = element_blank(),

   panel.grid.minor = element_blank(),

   axis.line = element_line(colour = "black"),

   strip.text = element_text(size = 14, face = "bold", color = "black"),  # Larger facet labels

   plot.title = element_text(size = 14, face = "bold",hjust=0.5,color="white"),  # Title formatting

   axis.text.x = element_text(angle = 45, hjust = 1,size=10,face="bold",color="white"),

   axis.title.x = element_text(face="bold",size=14,color="white"),

   axis.text.y = element_text(color="white"),

   axis.title.y = element_text(color="white")

 )+theme(plot.background = element_rect(fill = "#69922d"))

ggp_percent_part
```

Figure 4 (b): Average time spent in a day per participant - TUS 2019

```
time=c(299,305, 306, 307 ,298,284 ,285 ,296, 310, 291,

    109 ,93 ,97 ,94 ,95,102 ,93 ,89, 95, 94)

dt_time=data.frame("Gender"=c(rep("Female",10),rep("Male",10)),

          "Education"=rep(c("Not Literate","Below Primary","Primary","Middle","Secondary and above"),4),

          "Region"=rep(c(rep("Rural",5),rep("Urban",5)),2),
```

```
                "Average_Time"=time)
dt_time$Education=factor(dt_time$Education,levels=c("Not Literate","Below
Primary","Primary","Middle","Secondary and above"))
ggp_time=ggplot(dt_time, aes(x = Education, y =Average_Time, fill = Gender)) +
  geom_bar(stat = "identity", position = position_dodge(), color = "black",width=0.4) +
  scale_fill_manual(values = c("Male" = "blue", "Female" = "red")) +  # Male = blue, Female = red
  geom_text(aes(label = round(Average_Time,2)),angle=40,size=3,vjust=0,hjust=-0.15,color="black")+
  facet_wrap(~Region) +  # Facet by Region
  labs(title = "Average time (in minutes) spent in a day in Unpaid Household Activity for different
levels of education",
       x = "Education Level",
       y = "Average Time",
       fill = "Gender") +
  theme_bw() +  theme(
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    axis.line = element_line(colour = "black"),
    strip.text = element_text(size = 14, face = "bold", color = "black"),  # Larger facet labels
    plot.title = element_text(size = 14, face = "bold",hjust=0.5,color="white"),  # Title formatting
    axis.text.x = element_text(angle = 45, hjust = 1,size=10,face="bold",color="white"),
    axis.title.x = element_text(face="bold",size=14,color="white"),
    axis.text.y = element_text(color="white"),
    axis.title.y = element_text(color="white")
  )+theme(plot.background = element_rect(fill = "#69922d"))
ggp_time
```

Figure 4 (c): Estimated wage per hour for participating individuals - PLFS (2019-20) & TUS

```
#Boxplots for Estimated Wage by Rural/Urban, Male/Female, Education/Level
Boxplots_wage_edu=function(dt,wages_per_hour){

dt$Education=factor(dt$Education,levels=c("01","03","04","05","06","07","08","10","11","12","13"),
                    labels=c("Not Literate","03","Others","Below
Primary","Primary","Middle",rep("Secondary and above",5)))
```

```
index1=which(dt$Education=="03" | dt$Education=="Others")
index2=which(dt$AreaType!="Rural" & dt$AreaType!="Urban")
dt=dt[-c(index1,index2),]
wages_per_hour1=wages_per_hour[-c(index1,index2)]
dt_plot=data.frame("Gender"=dt$Sex,"Level_of_Education"=dt$Education,"Wage_per_Hour"=wages
_per_hour1,"Region"=dt$AreaType)
dt_plot <- dt_plot %>%
 mutate(Gender = case_when(
  Gender == 1  ~ "Male",
  TRUE ~ "Female"
 ))
ggp_box=ggplot(dt_plot, aes(x = Level_of_Education, y = Wage_per_Hour, fill = Gender)) +
 geom_boxplot(outlier.shape = NA, width = 0.6) +  # Boxplot with black borders
 scale_fill_manual(values = c("Male" = "blue", "Female" = "red")) +  # Custom gender colors
 facet_wrap(~Region) +  # Facet by Region
 labs(title = "Estimated Wage per Hour Distribution by Education Level",
     x = "Education Level",
     y = "Estimated Wage per Hour",
     fill = "Gender") +ylim(0,100)+
 theme_bw() +
 theme(panel.grid.major = element_blank(),
  panel.grid.minor = element_blank(),
  strip.text = element_text(size = 14, face = "bold", color = "black"),  # Larger facet labels
  plot.title = element_text(size = 14, face = "bold",hjust=0.5,color="white"),  # Title formatting
  axis.text.x = element_text(angle = 45, hjust = 1,size=10,face="bold",color="white"),
  axis.title.x = element_text(face="bold",size=14,color="white"),
  axis.text.y = element_text(color="white"),
  axis.title.y = element_text(color="white"),
  legend.position = "bottom"
 )+theme(plot.background = element_rect(fill = "#69922d"))
```

```
print(ggp_box)
}
ggp_box_wage=Boxplots_wage_edu(dt=dt_hh_work[index,],wages_per_hour)
```

Figure 4 (d): Estimated Average Wage per hour - PLFS (2019-20) & TUS

```
#Heatmaps for Mean Wage Rate by Rural/Urban, Male/Female, Education/Level
Heatmaps_wage_edu=function(dt,wages_per_hour){

dt$Education=factor(dt$Education,levels=c("01","03","04","05","06","07","08","10","11","12","13"),
                    labels=c("Not Literate","03","Others","Below
Primary","Primary","Middle",rep("Secondary and above",5)))
 index1=which(dt$Education=="03" | dt$Education=="Others")
 index2=which(dt$AreaType!="Rural" & dt$AreaType!="Urban")
 dt=dt[-c(index1,index2),]
 wages_per_hour1=wages_per_hour[-c(index1,index2)]

dt_plot=data.frame("Gender"=dt$Sex,"Level_of_Education"=dt$Education,"Wage_per_Hour"=wages
_per_hour1,"Region"=dt$AreaType)
 dt_plot <- dt_plot %>%
  mutate(Gender = case_when(
   Gender == 1  ~ "Male",
   TRUE ~ "Female"
  ))
 female_index=which(dt_plot$Gender=="Female")
 male_index=which(dt_plot$Gender=="Male")
 urban_index=which(dt_plot$Region=="Urban")
 rural_index=which(dt_plot$Region=="Rural")
 not_lit=which(dt_plot$Level_of_Education=="Not Literate")
 below_primary=which(dt_plot$Level_of_Education=="Below Primary")
 primary=which(dt_plot$Level_of_Education=="Primary")
 middle=which(dt_plot$Level_of_Education=="Middle")
 secondary=which(dt_plot$Level_of_Education=="Secondary and above")
 average_wage=rep(0,20)
```

```
 average_wage[1]=mean(wages_per_hour[c(female_index,rural_index,not_lit)])
 average_wage[2]=mean(wages_per_hour[c(female_index,rural_index,below_primary)])
 average_wage[3]=mean(wages_per_hour[c(female_index,rural_index,primary)])
 average_wage[4]=mean(wages_per_hour[c(female_index,rural_index,middle)])
 average_wage[5]=mean(wages_per_hour[c(female_index,rural_index,secondary)])
 average_wage[6]=mean(wages_per_hour[c(female_index,urban_index,not_lit)])
 average_wage[7]=mean(wages_per_hour[c(female_index,urban_index,below_primary)])
 average_wage[8]=mean(wages_per_hour[c(female_index,urban_index,primary)])
 average_wage[9]=mean(wages_per_hour[c(female_index,urban_index,middle)])
 average_wage[10]=mean(wages_per_hour[c(female_index,urban_index,secondary)])
 average_wage[11]=mean(wages_per_hour[c(male_index,rural_index,not_lit)])
 average_wage[12]=mean(wages_per_hour[c(male_index,rural_index,below_primary)])
 average_wage[13]=mean(wages_per_hour[c(male_index,rural_index,primary)])
 average_wage[14]=mean(wages_per_hour[c(male_index,rural_index,middle)])
 average_wage[15]=mean(wages_per_hour[c(male_index,rural_index,secondary)])
 average_wage[16]=mean(wages_per_hour[c(male_index,urban_index,not_lit)])
 average_wage[17]=mean(wages_per_hour[c(male_index,urban_index,below_primary)])
 average_wage[18]=mean(wages_per_hour[c(male_index,urban_index,primary)])
 average_wage[19]=mean(wages_per_hour[c(male_index,urban_index,middle)])
 average_wage[20]=mean(wages_per_hour[c(male_index,urban_index,secondary)])
 dt_box=data.frame("Gender"=c(rep("Female",10),rep("Male",10)),
                   "Region"=rep(c(rep("Rural",5),rep("Urban",5)),2),
                   "Education"=rep(c("Not Literate","Below Primary","Primary","Middle","Secondary and
above"),4),
                   "Average_Wage"=average_wage)
dt_box$Education=factor(dt_box$Education,levels=c("Not Literate","Below
Primary","Primary","Middle","Secondary and above"))
 ggp_box=ggplot(dt_box, aes(x = Education, y = Gender, fill = Average_Wage)) +
   geom_tile(color = "white") +  # Heatmap blocks with white borders
   scale_fill_gradient(low = "lightblue", high = "darkblue", name = "Wage per Hour") +
   geom_text(aes(label = round(Average_Wage,2)),color="yellow",face="bold")+
   facet_wrap(~ Region) +  # Facet by Region
```

```
  labs(title = "Estimated Average Wage per Hour by Education Level",
       x = "Education Level",
       y = " ",
       fill = "Gender") +
  theme_bw() +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        strip.text = element_text(size = 14, face = "bold", color = "black"),  # Larger facet labels
        plot.title = element_text(size = 14, face = "bold",hjust=0.5,color="white"),  # Title formatting
        axis.text.x = element_text(angle = 45, hjust = 1,size=10,face="bold",color="white"),
        axis.title.x = element_text(face="bold",size=14,color="white"),
        axis.text.y = element_text(color="white"),
        axis.title.y = element_text(color="white")
  )+theme(plot.background = element_rect(fill = "#69922d"))
 print(ggp_box)
}
ggp_heatmap_mean_wage=Heatmaps_wage_edu(dt=dt_hh_work[index,],wages_per_hour)
```

Figure 5: Comparison of usual PLFS Monthly Expenditure and Total monthly household work valuation

```
#Impact on Household Expenditure
#Count the number of females and males in each household
NH=nrow(plfs_hh_data1)
NP=nrow(plfs_person_data1)
#location=rep(0,NH)
number_females=rep(0,NH)
number_males=rep(0,NH)
number_females_age1=rep(0,NH)#15-29
number_females_age2=rep(0,NH)#15-59
number_females_age3=rep(0,NH)#60+
number_males_age1=rep(0,NH)#15-29
```

```
number_males_age2=rep(0,NH)#15-59
number_males_age3=rep(0,NH)#60+
plfs_hh_data1[,21]=as.numeric(plfs_hh_data1[,21])
plfs_person_data1[,20]=as.numeric(plfs_person_data1[,20])
k=1
for(i in 1:NH){
 hh_size=plfs_hh_data1[i,21]
 for(j in 1:hh_size){
  ages=plfs_person_data1[k:(k+hh_size),20]
  gender=plfs_person_data1[k:(k+hh_size),19]
 }
 gender_freq=as.data.frame(table(gender))
 number_females[i]=gender_freq$Freq[2]
 number_males[i]=gender_freq$Freq[1]
 number_females_age1[i]=length(intersect(which(ages>=15 & ages<=29),which(gender==2)))
 number_females_age2[i]=length(intersect(which(ages>=15 & ages<=59),which(gender==2)))
 number_females_age3[i]=length(intersect(which(ages>=60),which(gender==2)))
 number_males_age1[i]=length(intersect(which(ages>=15 & ages<=29),which(gender==1)))
 number_males_age2[i]=length(intersect(which(ages>=15 & ages<=59),which(gender==1)))
 number_males_age3[i]=length(intersect(which(ages>=60),which(gender==1)))
 k=k+hh_size+1
}
estimation_table=data.frame(number_females,number_males,number_females_age1,number_fema
les_age2,

number_females_age3,number_males_age1,number_males_age2,number_males_age3,
                    "PLFS_Consumption_Expenditure"=as.numeric(plfs_hh_data1[,25]))
estimation_table[is.na(estimation_table)] <- 0
engaged_females=(estimation_table$number_females)*(81.2)/100
engaged_males=(estimation_table$number_males)*(26.1)/100
engaged_females_age1=(estimation_table$number_females_age1)*(85.8)/100
engaged_females_age2=(estimation_table$number_females_age2)*(92.3)/100
```

```
engaged_females_age3=(estimation_table$number_females_age3)*(78.3)/100
engaged_males_age1=(estimation_table$number_males_age1)*(24.2)/100
engaged_males_age2=(estimation_table$number_males_age2)*(28.9)/100
engaged_males_age3=(estimation_table$number_males_age3)*(34.7)/100
#Estimate Total Hours Spent on Unpaid Work
time_engaged_females_age1=engaged_females_age1*291
time_engaged_females_age2=engaged_females_age2*315
time_engaged_females_age3=engaged_females_age3*245
time_engaged_males_age1=engaged_males_age1*94
time_engaged_males_age2=engaged_males_age2*95
time_engaged_males_age3=engaged_males_age3*112
#Compute Unpaid Work Value
UPWV_engaged_females_age1=time_engaged_females_age1*wage_female_total[1]*(30/60)
UPWV_engaged_females_age2=time_engaged_females_age2*wage_female_total[2]*(30/60)
UPWV_engaged_females_age3=time_engaged_females_age3*wage_female_total[3]*(30/60)
UPWV_engaged_males_age1=time_engaged_males_age1*wage_male_total[1]*(30/60)
UPWV_engaged_males_age2=time_engaged_males_age2*wage_male_total[2]*(30/60)
UPWV_engaged_males_age3=time_engaged_males_age3*wage_male_total[3]*(30/60)
#Expense on Household Activities Saved
Saved_Expense=UPWV_engaged_females_age2+UPWV_engaged_females_age3+
  UPWV_engaged_males_age2+UPWV_engaged_males_age3
#Total Imputed Expense
Total_Expense=estimation_table$PLFS_Consumption_Expenditure+Saved_Expense
```

Figure 5 (a): Boxplots showing distribution of Monthly Household Expenditure

```
dt_expense=data.frame("PLFS_Expenditue"=estimation_table$PLFS_Consumption_Expenditure,
                      "Total_Imputed_Expense"=Total_Expense)
df_long <- pivot_longer(dt_expense, cols = everything(), names_to = "Expenditure", values_to =
"Value")
df_long$Expenditure=factor(df_long$Expenditure,levels=c("PLFS_Expenditue" ,
"Total_Imputed_Expense"),label=c("PLFS Consumption \nExpenditue", "Estimated Expenditure
\n(Paid + Unpaid)"))
# Boxplot
```

```
ggp_expenditure_boxplot=ggplot(df_long, aes(x = Expenditure, y = Value, fill = Expenditure)) +
 geom_boxplot(width=0.3) +ylim(0,40000)+ labs(title = "Distribution of Monthly Consumption Expenditure",
                                              x = " ",y = "Monthly Expenditure") +
 scale_fill_manual(values = c("#f5b041","#1e8449"))+
 theme_bw() +
 theme(panel.grid.major = element_blank(),
       panel.grid.minor = element_blank(),
       plot.title = element_text(size = 14, face = "bold",hjust=0.5,color="white"),  # Title formatting
       axis.text.x = element_text(angle = 45, hjust = 1,size=12,face="bold",color="white"),  # Rotate x-axis labels
      axis.text.y = element_text(color="white"),
       legend.position = "bottom",axis.title.x = element_text(face="bold",size=14),
       axis.title.y = element_text( face = "bold", color = "white") )+
 theme(plot.background = element_rect(fill = "#661a6c"), legend.position = "none")
ggp_expenditure_boxplot
```

Figure 5 (b): Dot Plot for Monthly Average Household Expenditure

```
avg_expenditure=data.frame(Expenditure=c("PLFS_Expenditue","Estimated Expenditure"),Value=c(mean(estimation_table$PLFS_Consumption_Expenditure),mean(Total_Expense)))
avg_expenditure$Expenditure=factor(avg_expenditure$Expenditure,levels=c("PLFS_Expenditue","Estimated Expenditure"),
                                   label=c("PLFS Consumption \nExpenditue", "Estimated Expenditure \n(Paid + Unpaid)"))
ggp_expenditure_line=ggplot(avg_expenditure, aes(x = Expenditure, y = Value,group = 1)) +
 geom_point(aes(color = Expenditure), size = 5) +
 scale_color_manual(values = c("PLFS Consumption \nExpenditue"="#f5b041","Estimated Expenditure \n(Paid + Unpaid)"="#1e8449"))+
 geom_line(size = 1)  +  geom_text(aes(label = round(Value,2)), vjust = 1.7,hjust=0.5, size = 3.5,color="darkblue")+
 labs(title = "Average Monthly Consumption Expenditure",
      x = " ",y = "") +ylim(0,24000)+
  theme_bw() +
 theme(panel.grid.major = element_blank(),
```

```
    panel.grid.minor = element_blank(),
    plot.title = element_text(size = 14, face = "bold",hjust=0.5,color="white"),  # Title formatting
    axis.text.x = element_text(angle = 45, hjust = 1,size=12,face="bold",color="white"),  # Rotate x-axis labels
    axis.text.y = element_text(color="white"),
    legend.position = "bottom",
    axis.title.y = element_text( face = "bold", color = "white") )+
  theme(plot.background = element_rect(fill = "#661a6c"), legend.position = "none")
ggp_expenditure_line
```