Country: Peru

Exercise 7

7.1)

The Demographic and Health Surveys (DHS) provides information about inter-partner violence (IPV) suffered by women from their male partners, asking them if they have experienced any severe violence. The DHS dataset includes data from Peru, from 2012. The initial sample for Peru comprised 23,888 individuals. After subsetting according to the variables of interest to this analysis and removing missing values, the number of observations dropped to 13,477. In this context, the share of women who reported suffering from IPV was 16.52%.

7.2)

Peru is strongly diverse in terms of ethnic composition. The DHS dataset comprises five ethnic groups, labeled according to their mother tongue: "Spanish speaker", "Quechua", "Aymara other", "Indigenous", and "foreign language". In such a diverse scenario, the causes are related to factors such as origin, gender, or status (Iguales & Oxfam, 2015). Therefore, in this analysis ethnicity and education were used to differentiate groups and identify the most marginalized.

The ethnic groups (variable "ethnic") were grouped and relabeled as "Spanish", "Quechua" and "Aymara and Indigenous", and "foreign language" speakers were not considered since they comprised only 6 observations, with no IPV being reported. The variable "class" was defined according to the education level of the respondent, encompassing "low" (0 and 1), "medium" (2), and "top" (3).

a)

The graphs below illustrate the distribution of IPV cases among ethnic (**Figure 1**) and class (**Figure 2**) groups.

In terms of ethnic groups, IPV cases were more frequently observed among Quechua women, in which 26.6% of cases are concentrated. In terms of class, IPV cases were more recurrent among women with low education levels, in which 22.1% of the cases are concentrated.

Figure 1

Proportion of IPV cases among the main ethnic groups in Peru

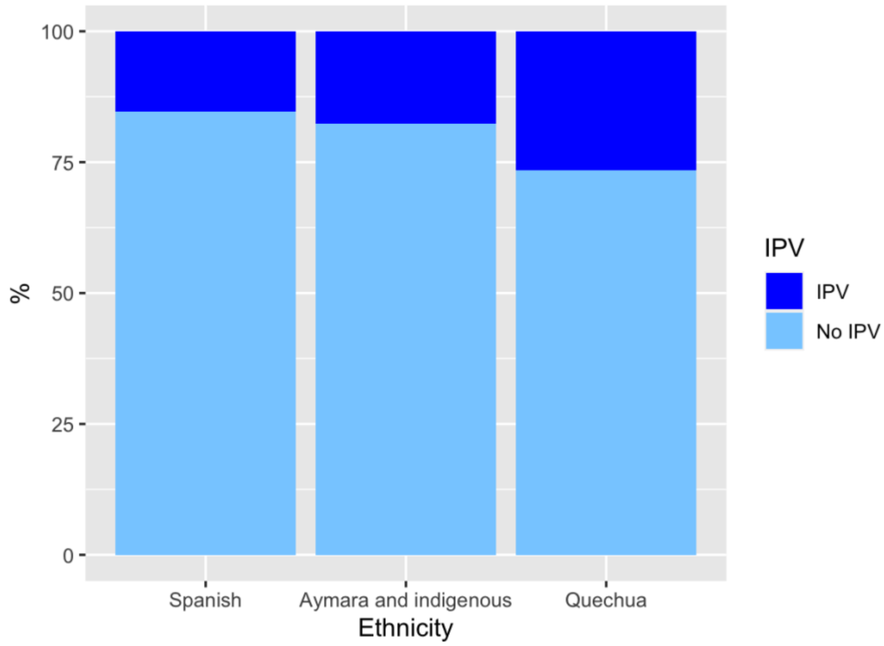
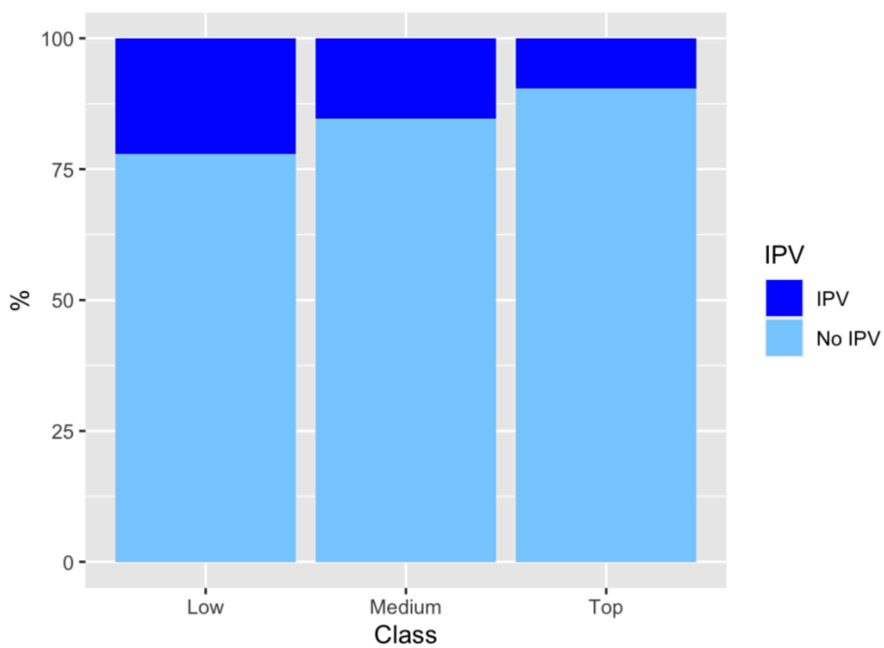


Figure 2

Proportion of IPV cases among the classes in Peru

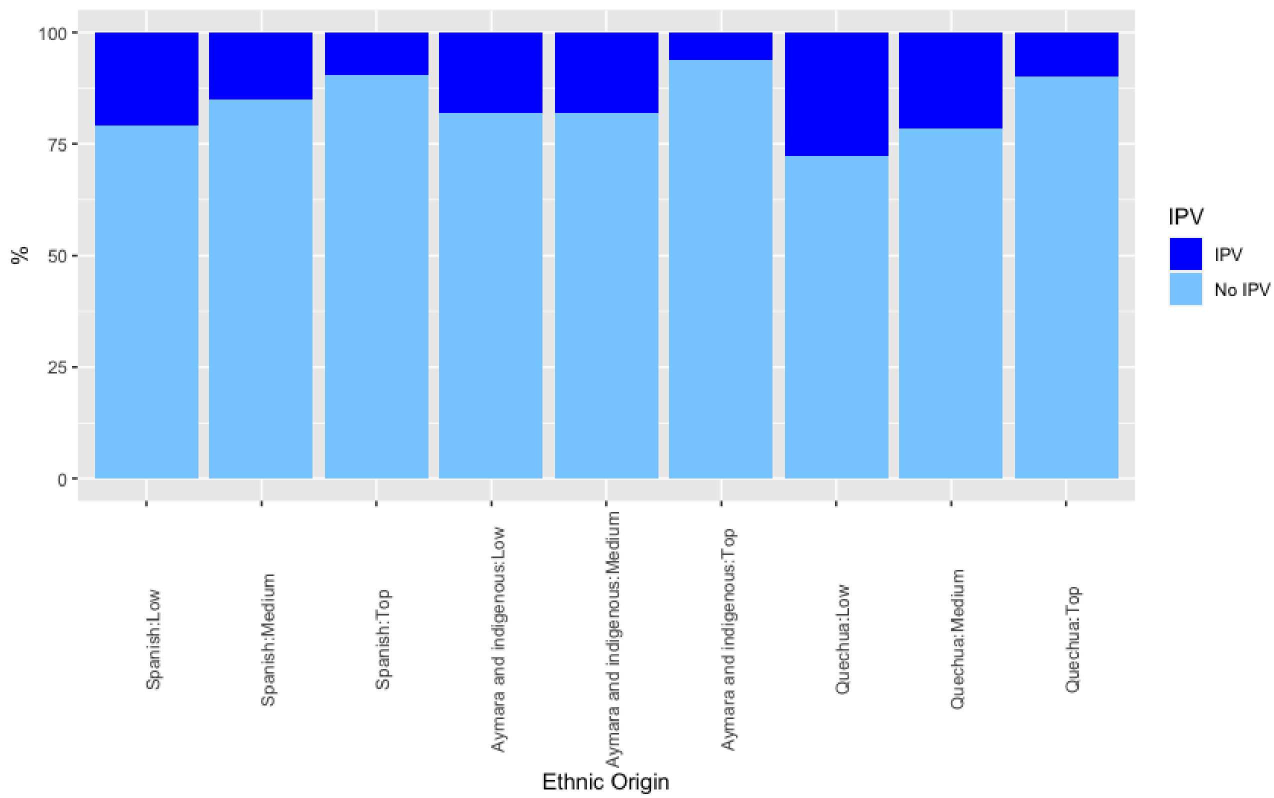


b)

Figure 3 depicts the proportion of IPV cases among ethnic and class groups in Peru. The interaction between the two variables reinforces patterns observed in Figures 1 and 2. IPV is more prevalent among women of low class, regardless of their ethnic group. However, among Quechua women IPV is more prevalent.

Figure 3

Occurrence of IPV cases among ethnic groups and classes in Peru



7.3)

The results from the intersectional approach (Figure 3) demonstrated a similar scenario to the one depicted from the separate approaches to class and ethnicity. However, the relevance of such an approach relies on the support to better comprehend the structural factors that shape disparities (Short & Zacher, 2022). In this case, an intersectional analysis illustrates how the distribution of cases varies among classes of each group, reinforcing the idea that low-income women are more vulnerable to IPV and, therefore, solutions should encompass a multitude of approaches (e.g. economic, health, psychological). As proposed by Al-Faham, Davis, & Ernst (2019), the between-category relationships reveal the most vulnerable individuals inside a social division when cross-cutting issues intensify their disadvantages.

A setback of an intersectional approach is the complexity it brings to policymakers in identifying subgroups and proposing effective solutions that take their specificities into account, bringing complexity to the agenda, especially in the case of a delicate topic such as IPV. Equally, the specificities of an intersectional approach might deviate the attention from the main root of the problem to obstacles that do not necessarily derive from the unequal power dynamics in a society.

Nevertheless, since this is a recurring issue, the agenda to discuss ethnic group inequalities is already displayed in Peru. In that case, understanding the intergroup variation might redirect solutions to the most vulnerable individuals, but not to the expanse of other subgroups (in this case, women of high and medium social classes and from more advantaged ethnic groups) that also demand directed policies.

7.4)

a)

The logistic regression results (see **Table 1**) show education level (CLASS), ethnicity (ETHNIC), and age (AGE) associated with Experienced Severe Violence (DEP) in Peru. Here's a breakdown of the findings in terms of Odds Ratios (OR) and their interpretation:

Table 1

Logistic Regression Model Predicting DEP in Peru

Characteristic	OR[1]	95% CI[1]	p-value
CLASS			
Low	—	—	
Medium	0.80	0.71, 0.88	<0.001
Top	0.42	0.36, 0.48	<0.001
ETHNIC			
Spanish	—	—	
Aymara and indigenous	1.03	0.78, 1.36	0.8
Quechua	1.46	1.27, 1.69	<0.001
AGE	1.04	1.03, 1.04	<0.001

[1] OR = Odds Ratio, CI = Confidence Interval

The level of education plays a role in DEP risk. Compared to the reference group with low education, individuals with a medium level of education have a 20% decrease in the odds of experiencing DEP. This effect is even stronger for those with top education, who have a 58% decrease in odds compared to the low education group. Ethnicity is another factor associated with DEP risk. Individuals of Quechua ethnicity have 46% increased odds of experiencing DEP compared to those with a Spanish ethnic background. However, the model doesn't find a significant association for the Aymara and Indigenous group. The analysis shows a positive association between age and DEP risk. With each year of age increase, the odds of experiencing DEP go up by 4%.

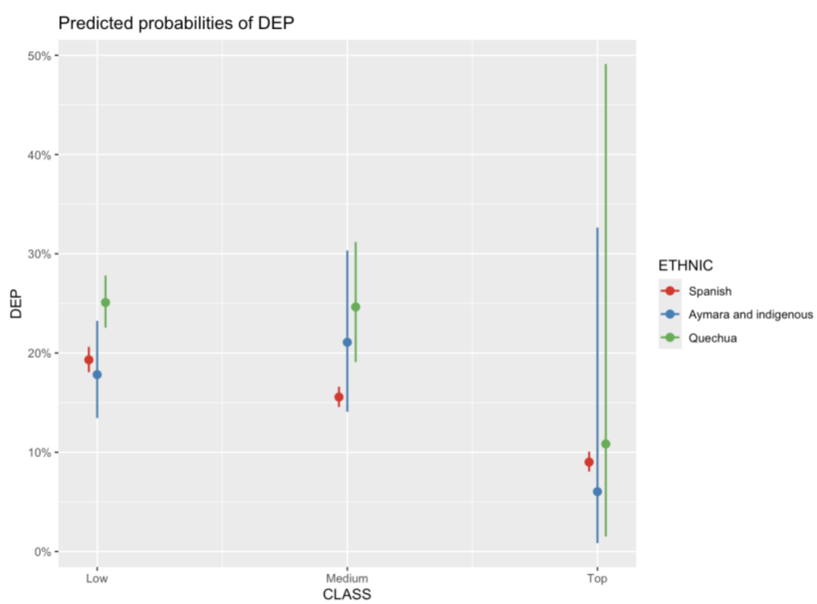
It's important to clarify that these are just observational associations. The model doesn't necessarily imply that education level directly causes a decrease in DEP risk. There could be other factors at play.

b)

This logit model estimates the association between Intimate Partner Violence (DEP) and several factors. It controls for the interaction between education level (CLASS) and ethnicity (ETHNIC), along with age (AGE). The results are shown in Figure 4.

Figure 4

Plot for Logistic Regression Model Predicting DEP in Peru, Controlling Interaction between CLASS and ETHNIC



This analysis shows the substantial influence of education level on the likelihood of experiencing DEP. Notably, individuals with higher education attainment, categorized into Medium and Top Class, consistently exhibit a reduced predicted probability of DEP compared to those with lower education levels, irrespective of ethnicity. This observation suggests that education serves as a protective factor against DEP, highlighting its pivotal role in mitigating such adverse experiences.

Furthermore, variations in the predicted probability of DEP are observed across different ethnicities. Spanish individuals generally demonstrate the lowest predicted probability across all education levels, contrasting with Quechua individuals who exhibit a comparatively higher predicted probability of DEP. However, it's essential to acknowledge that the model does not reveal a statistically significant interaction between education and ethnicity. Specifically, the presence of large standard errors (SE) and wide confidence intervals (CI) in the predicted probability for the CLASS interacting with the ETHNIC group (see **Table 2**) suggests that the impact of education on DEP risk may not vary significantly across ethnicities. There could be no interaction or more complex relationships among those factors.

Table 2

Logistic Regression Model Predicting DEP in Peru, Controlling Interaction between CLASS and ETHNIC

Characteristic	OR[1]	95% CI[1]	p-value
AGE	1.04	1.03, 1.04	<0.001
CLASS * ETHNIC			
Low * Spanish	1.97	0.37, 36.5	0.5
Medium * Spanish	1.52	0.28, 28.1	0.7
Top * Spanish	0.81	0.15, 15.1	0.8
Low * Aymara and indigenous	1.78	0.32, 33.5	0.6
Medium * Aymara and indigenous	2.20	0.38, 41.9	0.5
Top * Aymara and indigenous	0.53	0.02, 14.5	0.7
Low * Quechua	2.76	0.51, 51.1	0.3
Medium * Quechua	2.69	0.48, 50.4	0.4
Top * Quechua			

[1] OR = Odds Ratio, CI = Confidence Interval

References

Al-Faham, H., Davis, M. A., & Ernst, R. (2019). Intersectionality. From theory to practice. *Annual Review of Law and Social Science, 15*, 247-265.

Iguales & Oxfam. (2015). *Inequality in Peru: Reality and Risks.* Retrieved from: https://cng-cdn.oxfam.org/peru.oxfam.org/s3fs-public/file_attachments/Inequality%20in%20Peru.%20Reality%20and%20Risks_4.pdf

Rutstein, Shea O., and Ann Way. 2014. The Peru Continuous DHS Experience. DHS Occasional Papers No. 8. Rockville, Maryland, USA: ICF International.

Short, S. E., & Zacher, M. (2022). Women's Health: Population Patterns and Social Determinants. *Annu. Rev. Sociol., 48*, 277–98.

R packages:

Enzmann JAIRscadwbD, Schwartz M, Jain N, Kraft S (2023). _descr: Descriptive Statistics_. R package version 1.1.8, <https://CRAN.R-project.org/package=descr>.

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Lüdecke D (2023). _sjPlot: Data Visualization for Statistics in Social Science_. R package version 2.8.15, <https://CRAN.R-project.org/package=sjPlot>.

Rich B (2023). _table1: Tables of Descriptive Statistics in HTML_. R package version 1.4.3, <https://CRAN.R-project.org/package=table1>.

Sjoberg DD, Whiting K, Curry M, Lavery JA, Larmarange J. Reproducible summary tables with the gtsummary package. The R Journal 2021;13:570–80. https://doi.org/10.32614/RJ-2021-053.

Wickham H, Miller E, Smith D (2023). _haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files_. R package version 2.5.4, <https://CRAN.R-project.org/package=haven>.

Appendix

```
# Session 07
# Packages
library(haven)
library(descr)
library(ggplot2)
library(table1)
library(gtsummary)
library(sjPlot)

#Read in data
DATA00<-read_dta("Peru.dta")  # 23888 obs.
DATA01<-subset(DATA00, !(is.na(DATA00$d107)))
DATA01<-subset(DATA01, !(is.na(DATA01$v131)))
DATA01<-subset(DATA01, !(is.na(DATA01$v106)))

#New Variable: IPV
DATA01$IPV<-"NA"
DATA01$IPV[DATA01$d107==0]<-"No IPV"
DATA01$IPV[DATA01$d107==1]<-"IPV"
DATA01$IPV<- as.factor(DATA01$IPV)

#New Variable: DEP
DATA01$DEP<-"NA"
DATA01$DEP[DATA01$d107==0]<-0
DATA01$DEP[DATA01$d107==1]<-1
DATA01$DEP<- as.numeric(DATA01$DEP)

#New Variable: ETHNIC
# 1=Spanish, 2=Quechua, 3=Aymara and indigenous, 4=Other, 5=Foreign language
attributes(DATA01$v131)
```

```
# drop the ethnic group "foreign language" (6 obs.)
DATA01 <- subset(DATA01, v131 < 5)  # 13477 obs.

DATA01$ETHNIC<-"NA"
DATA01$ETHNIC[DATA01$v131==1]<-"Spanish"
DATA01$ETHNIC[DATA01$v131==2]<-"Quechua"
DATA01$ETHNIC[DATA01$v131==3]<-"Aymara and indigenous"
DATA01$ETHNIC[DATA01$v131==4]<-"Aymara and indigenous"
DATA01$ETHNIC<- as.factor(DATA01$ETHNIC)
DATA01$ETHNIC<- relevel(DATA01$ETHNIC, ref = "Spanish")

#New Variable: EDUCATION
DATA01$CLASS<-"NA"
DATA01$CLASS[DATA01$v106==0]<-"Low"
DATA01$CLASS[DATA01$v106==1]<-"Low"
DATA01$CLASS[DATA01$v106==2]<-"Medium"
DATA01$CLASS[DATA01$v106==3]<-"Top"
DATA01$CLASS <- as.factor(DATA01$CLASS)

#New Variable: AGE
DATA01$AGE<-DATA01$v007-DATA01$v010

#Frequencies
freq(DATA01$IPV)

#Column
crosstab(DATA01$IPV,DATA01$ETHNIC, prop.c=TRUE, xlab = "Ethnic", ylab = "IPV")
crosstab(DATA01$IPV,DATA01$CLASS, prop.c=TRUE, xlab = "Class", ylab = "IPV")

#Bar Chart: 2 Dimensional
# Ethnic
TABLE1 <- table(DATA01$IPV,DATA01$ETHNIC)
```

```
TABLE2 <- prop.table(TABLE1,2)
TABLE3 <- as.data.frame(TABLE2)
TABLE3$Percent <- TABLE3$Freq*100
ggplot(TABLE3, aes(fill=Var1, x=Var2, y=Percent))+
  geom_bar(stat="identity",position="stack")+
  ylab("%")+
  xlab("Ethnicity")+
  scale_fill_manual(name = "IPV",
                    labels = c("IPV", "No IPV"),
                    values = c("blue1", "skyblue1"))

# Class
TABLE4 <- table(DATA01$IPV,DATA01$CLASS)
TABLE5 <- prop.table(TABLE4,2)
TABLE6 <- as.data.frame(TABLE5)
TABLE6$Percent <- TABLE6$Freq*100
ggplot(TABLE6, aes(fill=Var1, x=Var2, y=Percent))+
  geom_bar(stat="identity",position="stack")+
  ylab("%")+
  xlab("Class")+
  scale_fill_manual(name = "IPV",
                    labels = c("IPV", "No IPV"),
                    values = c("blue1", "skyblue1"))

# Intersectionality
DATA01$INT=DATA01$ETHNIC:DATA01$CLASS
crosstab(DATA01$IPV, DATA01$INT, prop.c=TRUE, xlab = "Intersectionality: Ethnic and Class", ylab =
"IPV")

#Bar Chart: 3 Dimensional (nicer)
TABLE1 <- table(DATA01$IPV,DATA01$INT)
TABLE2 <- prop.table(TABLE1,2)
```

```
TABLE3 <- as.data.frame(TABLE2)
TABLE3$Percent <- TABLE3$Freq*100
ggplot(TABLE3, aes(fill=Var1, x=Var2, y=Percent))+
  geom_bar(stat="identity",position="stack")+
  ylab("%")+
  xlab("Ethnic Origin")+
  scale_fill_manual(name = "IPV",
                    labels = c("IPV", "No IPV"),
                    values = c("blue1", "skyblue1")) +
  theme(axis.text.x = element_text(angle = 90))


# Regression
table1::label(DATA01$AGE) <- "AGE"
table1(~DEP+ETHNIC+CLASS+AGE | IPV, data = DATA01)

MODEL1 <- glm(DEP ~ CLASS+ETHNIC+AGE, data=DATA01, family=binomial())
tbl_regression(MODEL1, exponentiate = TRUE)

MODEL2 <- glm(DEP ~ CLASS:ETHNIC+AGE, data=DATA01, family=binomial())
tbl_regression(MODEL2, exponentiate = TRUE)
plot_model(MODEL2, type = "pred", terms = c("CLASS","ETHNIC"))
```