Scatter_plots

Rajneesh

2024-01-08

Overall Structure of the Document

I tried to explore interesting relationships mainly found from the correlation plot and tried to understand further

Case 1: Education Descripency Between Males and Females

Case 2: Checking if educated families had more pashmina goats

Case 3: If wealtheir Families had more Pashmina Goats

Case 4: Explaining Kiang threat perception using kiang estimates

Note: I am assuming 10% as level of significance for p-values

Loading Required Libraries

```
# Library
library(ggplot2)
library(hrbrthemes)
library(ggpmisc)
library(ggpubr)
library(readxl)
library(tidyverse)
```

Loading data

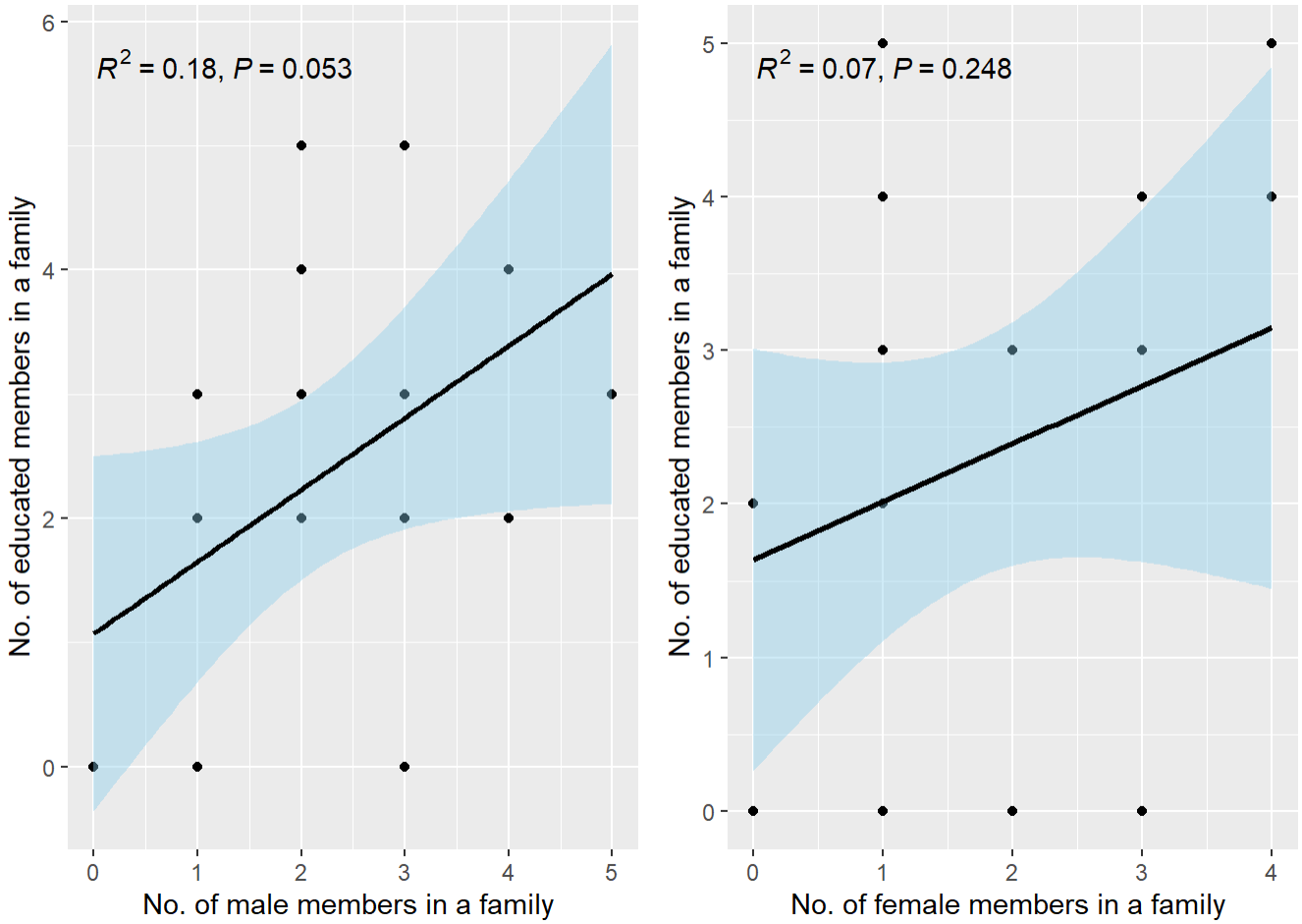
```
data <- read_excel("E:/GIS/Amit's Social Analytics/Cleaned_dataset.xlsx")%>%
  mutate(Percapita_pashmina=Pashmina_goats/HH_TFM)%>%
  mutate(ED_percapita=ED/HH_TFM)
```

Case 1: Education Descripency Between Males and Females

```
# Education VS Males
ED_vs_M= ggplot(data, aes(x=M, y=ED)) +
geom_point() +
geom_smooth(method=lm , color="black", fill="skyblue", se=TRUE) +
stat_poly_eq(use_label(c("R2", "p")))+
xlab("No. of male members in a family ") +
ylab("No. of educated members in a family")

# Education VS Females
ED_vs_F=ggplot(data, aes(x=F, y=ED)) +
geom_point() +
geom_smooth(method=lm , color="black", fill="skyblue", se=TRUE) +
stat_poly_eq(use_label(c("R2", "p")))+
xlab("No. of female members in a family ") +
ylab("No. of educated members in a family")

ggarrange(ED_vs_M,ED_vs_F, ncol = 2, nrow = 1)
```



Seems like the increase in number of males in family meant more educated member in family (weak relationship) but that was not the case with increase in number of females (non-significant p-value)

Case 2: Checking if educated families had more pashmina goats

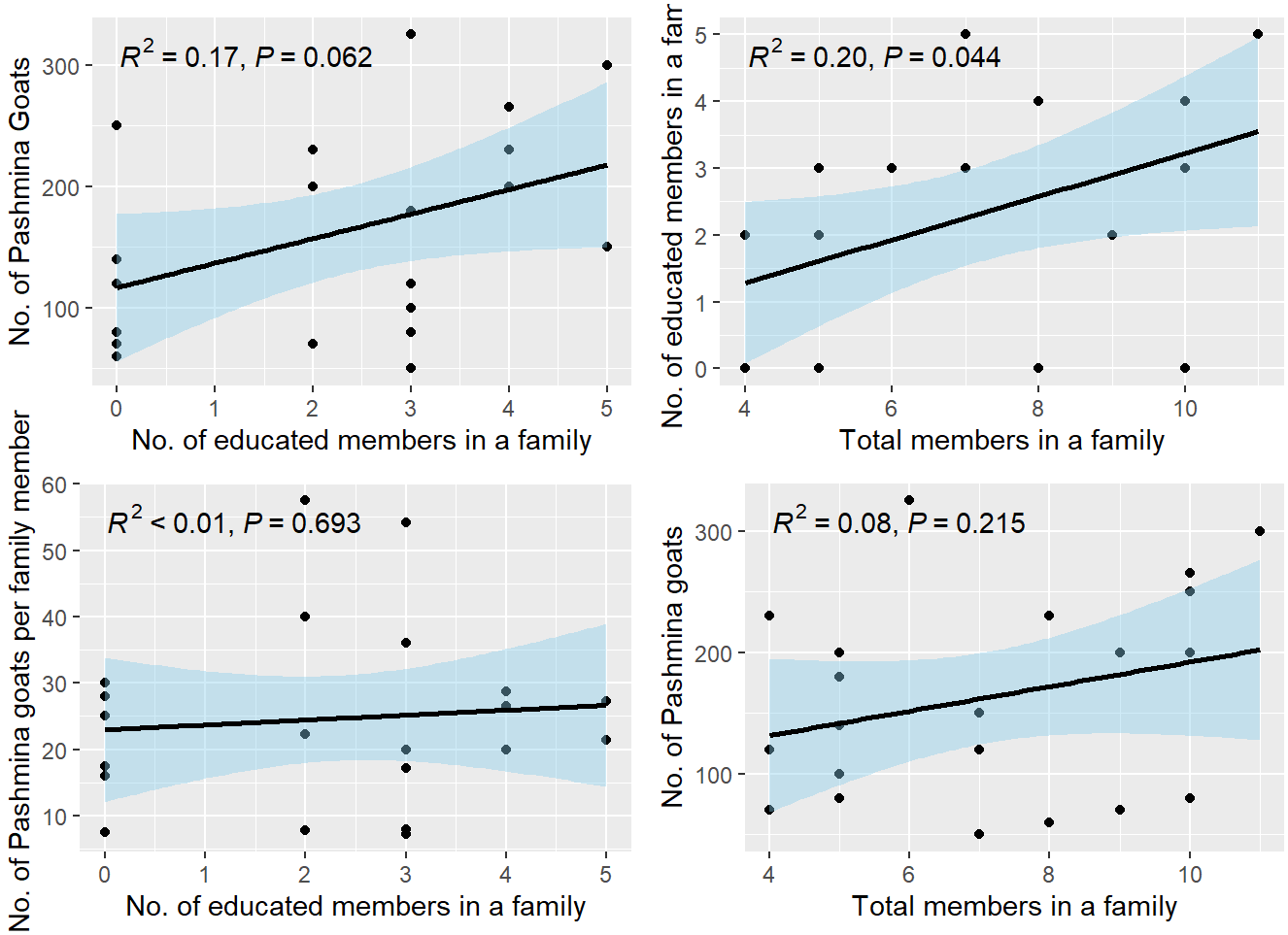
```
# Education VS Pashmina Goats
ED_vs_Pash= ggplot(data, aes(x=ED, y=Pashmina_goats)) +
geom_point() +
geom_smooth(method=lm , color="black", fill="skyblue", se=TRUE) +
stat_poly_eq(use_label(c("R2", "p")))+
ylab("No. of Pashmina Goats") +
xlab("No. of educated members in a family")

# Education VS Total Family Numbers
ED_vs_TFM= ggplot(data, aes(x=HH_TFM, y=ED)) +
geom_point() +
geom_smooth(method=lm , color="black", fill="skyblue", se=TRUE) +
stat_poly_eq(use_label(c("R2", "p")))+
xlab("Total members in a family") +
ylab("No. of educated members in a family")

#Education vs Percapita Pashmina goats
ED_vs_percapita_pashmina= ggplot(data, aes(x=ED ,y=Percapita_pashmina)) +
geom_point() +
geom_smooth(method=lm , color="black", fill="skyblue", se=TRUE) +
stat_poly_eq(use_label(c("R2", "p")))+
ylab("No. of Pashmina goats per family member ") +
xlab("No. of educated members in a family")

#Total Family members vs Pashmina goats
TFM_vs_pashmina= ggplot(data, aes(x=HH_TFM ,y=Pashmina_goats)) +
geom_point() +
geom_smooth(method=lm , color="black", fill="skyblue", se=TRUE) +
stat_poly_eq(use_label(c("R2", "p")))+
ylab("No. of Pashmina goats") +
xlab("Total members in a family")

ggarrange(ED_vs_Pash,ED_vs_TFM,ED_vs_percapita_pashmina, TFM_vs_pashmina,ncol = 2, nrow = 2)
```



Step 1: Upon first look seems like as no. of educated family members increased, the number of pashmina goats increased (top left panel).

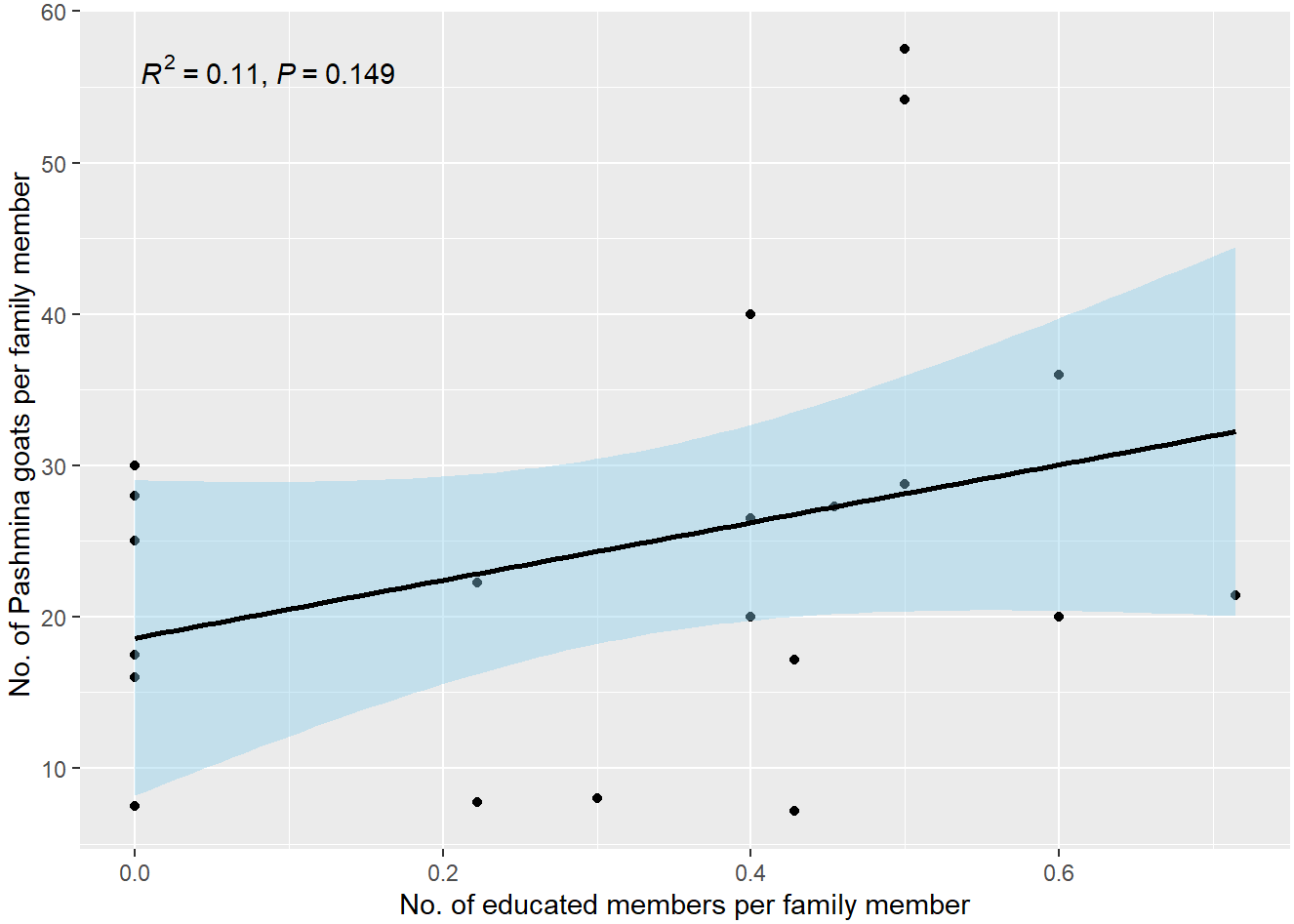
Step 2: However that could have been happening because of larger families and thus more educated members. Top right panels shows that larger families had more educated members.

Step 3: Bottom left panel also show that there wasn't increase in per capita pashmina goats (per each member of family) with increase in educated members of family.

Step 4: But the number of Pashmina goats was also not dependent upon total number of family member (bottom right panel) or full time workers (figure not mentioned here, but checked separately)

There seems like a weak effect of educated members on total pashmina goats.

```
#per capita Education vs Percapita Pashmina goats
ggplot(data, aes(x=data$ED_percapita ,y=Percapita_pashmina)) +
geom_point() +
geom_smooth(method=lm , color="black", fill="skyblue", se=TRUE) +
stat_poly_eq(use_label(c("R2", "p")))+
ylab("No. of Pashmina goats per family member ") +
xlab("No. of educated members per family member")
```



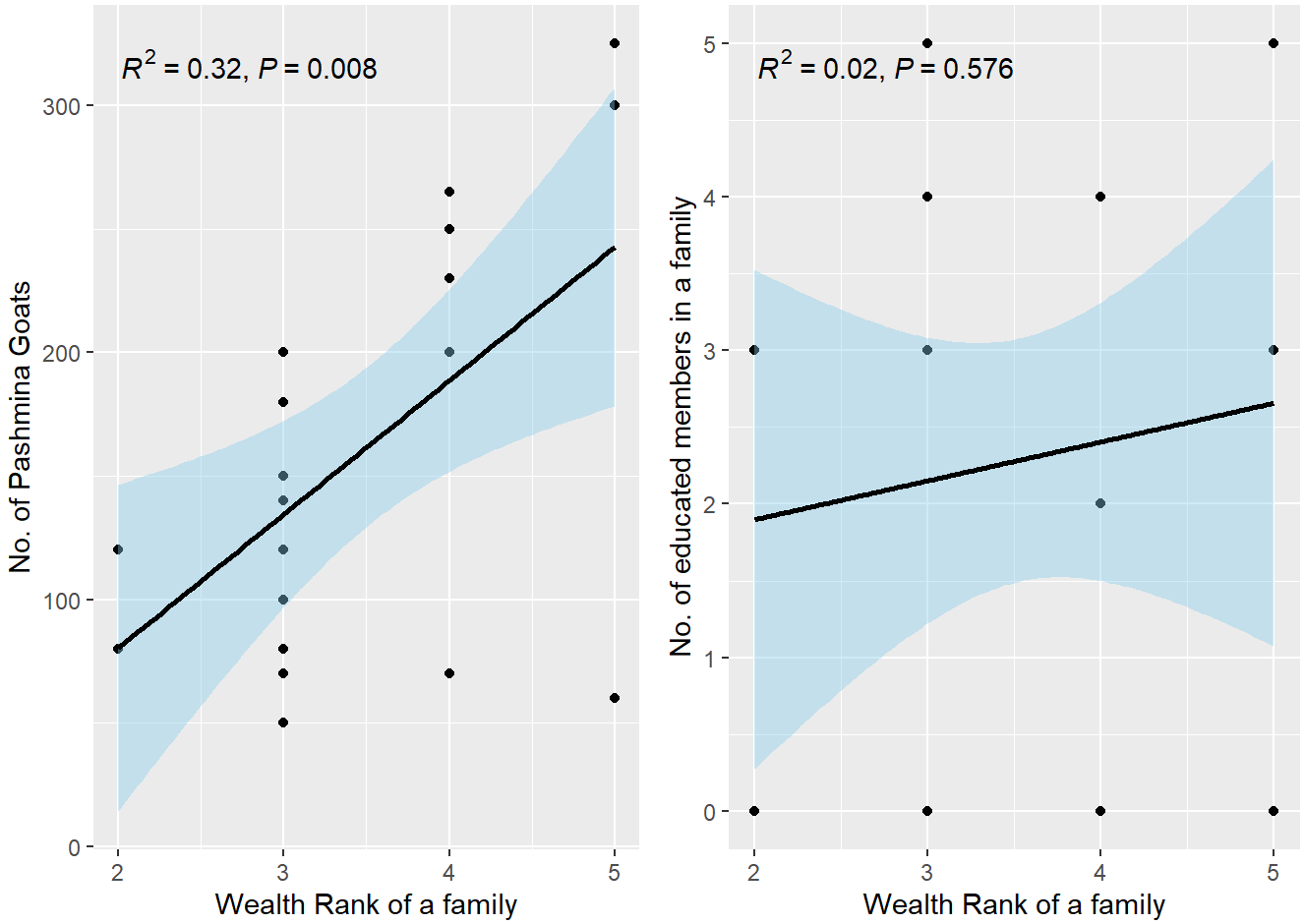
The above figure checks if percapita pashmina goats increased with per capita increase in eductaed members however the p-value is not signicant

Case 3: If wealtheir Families had more Pashmina Goats

```
# Wealth VS Pashmina Goats
Wealth_vs_Pash= ggplot(data, aes(x=wealth_rank, y=Pashmina_goats)) +
geom_point() +
geom_smooth(method=lm , color="black", fill="skyblue", se=TRUE) +
stat_poly_eq(use_label(c("R2", "p")))+
ylab("No. of Pashmina Goats") +
xlab("Wealth Rank of a family")

# Wealth VS ED
Wealth_vs_ED= ggplot(data, aes(x=wealth_rank, y=ED)) +
geom_point() +
geom_smooth(method=lm , color="black", fill="skyblue", se=TRUE) +
stat_poly_eq(use_label(c("R2", "p")))+
ylab("No. of educated members in a family") +
xlab("Wealth Rank of a family")

ggarrange(Wealth_vs_Pash,Wealth_vs_ED,ncol = 2, nrow = 1)
```



Rather than educated members (Case 2) in a family wealth seems like a better correlated with number of Pashmina goats (left panel).

I also checked if wealthy families were more eductaed (right panel). But that was not the case.

Case 4: Explaining Kiang threat perception using kiang estimates

```
# Kiang VS wealth rank
Kiang_vs_wealthrank= ggplot(data, aes(x=wealth_rank, y=data$kiang_exp)) +
geom_point() +
geom_smooth(method=lm , color="black", fill="skyblue", se=TRUE) +
stat_poly_eq(use_label(c("R2", "p")))+
ylab("Estimated Kiang numbers") +
xlab("Wealth rank of a family")

# Kiang VS Total livestock
Kiang_vs_livestock= ggplot(data, aes(x=data$total_livestock, y=data$kiang_exp)) +
geom_point() +
geom_smooth(method=lm , color="black", fill="skyblue", se=TRUE) +
stat_poly_eq(use_label(c("R2", "p")))+
ylab("Estimated Kiang numbers") +
xlab("Total no. of livestock")
# Kiang VS TFM
Kiang_vs_TFM= ggplot(data, aes(x=HH_TFM, y=data$kiang_exp)) +
geom_point() +
geom_smooth(method=lm , color="black", fill="skyblue", se=TRUE) +
stat_poly_eq(use_label(c("R2", "p")))+
ylab("Estimated Kiang numbers") +
xlab("Total members in a family")

# Kiang VS FT
Kiang_vs_FT= ggplot(data, aes(x=FT, y=data$kiang_exp)) +
geom_point() +
geom_smooth(method=lm , color="black", fill="skyblue", se=TRUE) +
stat_poly_eq(use_label(c("R2", "p")))+
ylab("Estimated Kiang numbers") +
xlab("Full time workers in a family")

ggarrange(Kiang_vs_wealthrank, Kiang_vs_livestock ,Kiang_vs_TFM, Kiang_vs_FT,ncol = 2, nrow =
2)
```

Contrary to our earlier expectation that Kiang threat perception could be relalted to wealth (top left panel) or livestock number (top right panel), the kiang threat perception was related to family size.

I think this could have been because of as more family member have higher chances of encountering kiang (bottom left panel)

Further I checked if full time working members could explan kiang threat perception (botton right panel), however that doesn't seems to be the case as p-value is non-significant.