

ANALISIS PREDICTIVO

# ANALISIS EXPLORATORIO DE DATOS

*Profesores*

*Ezequiel Martín Eliano Sombory*

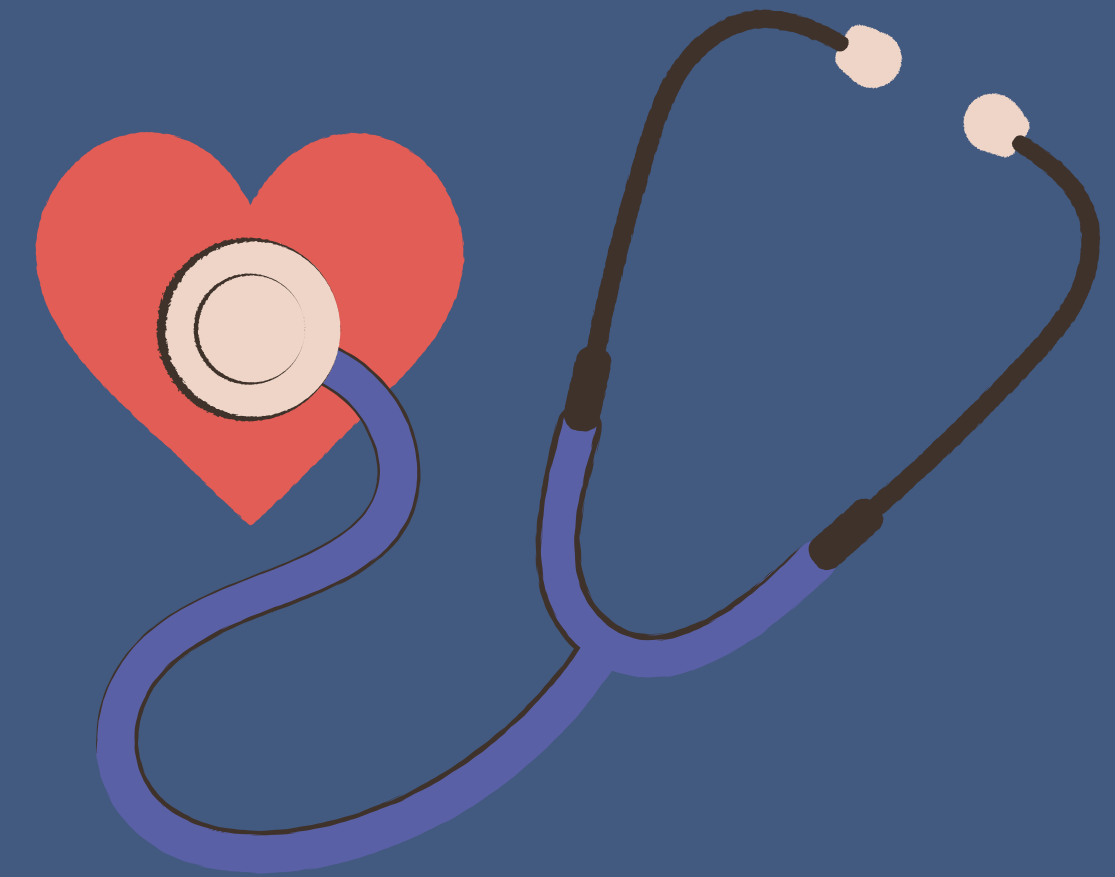
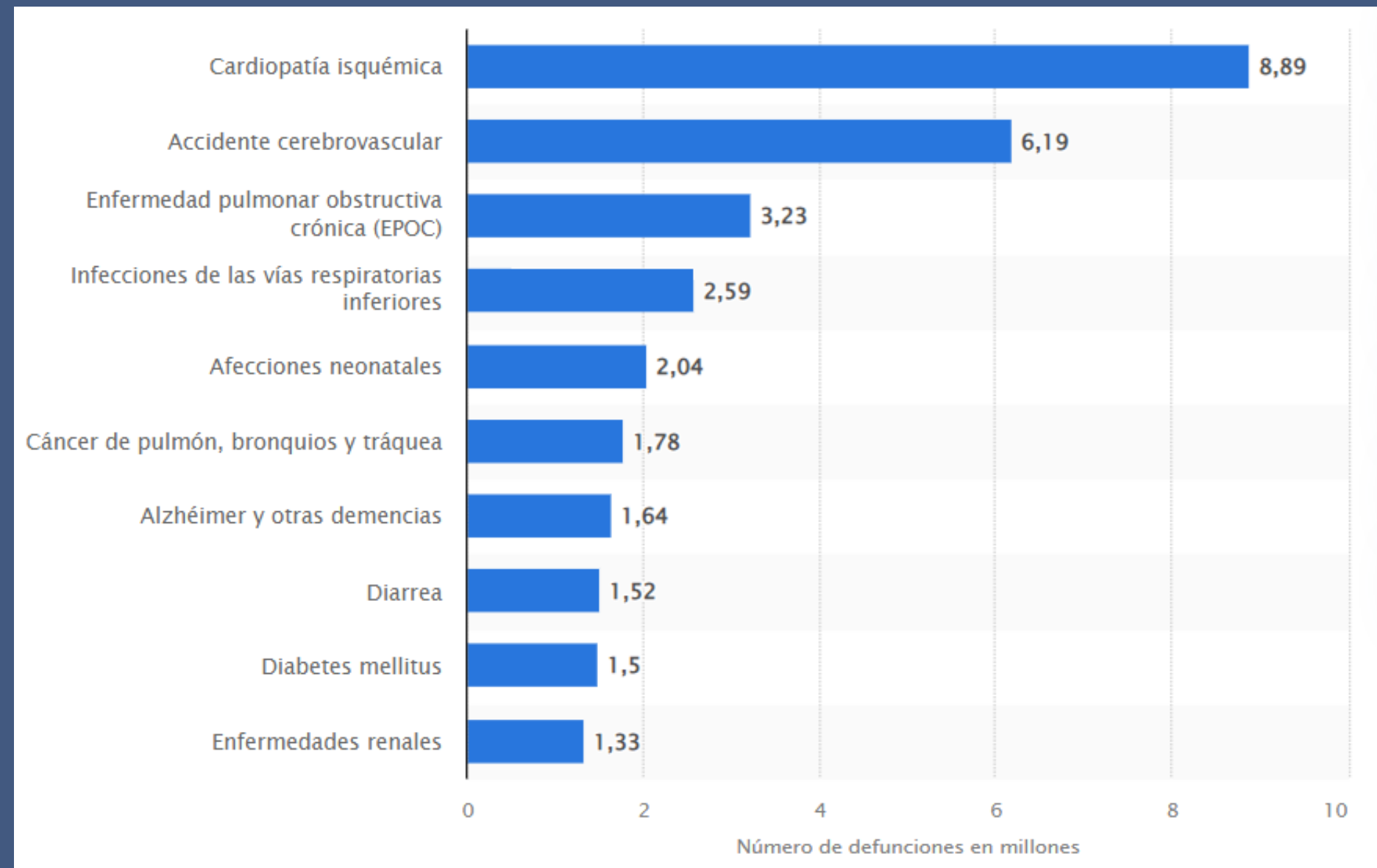
*Leonardo Andrés Caravaggio*

*Francisco Valentin*



JUAN INGACIO SCORZA

# PRESENTACION CASO DE NEGOCIO



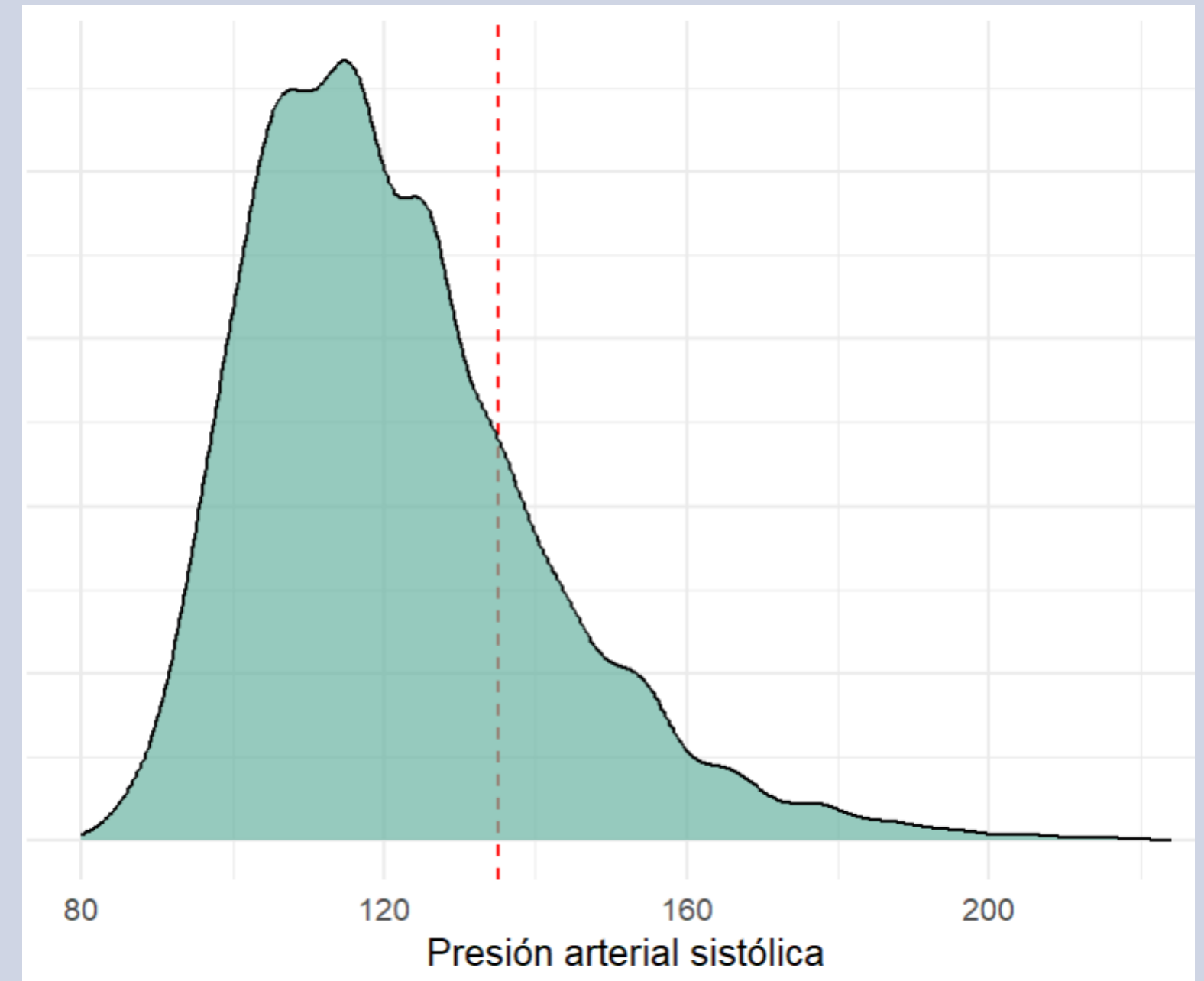
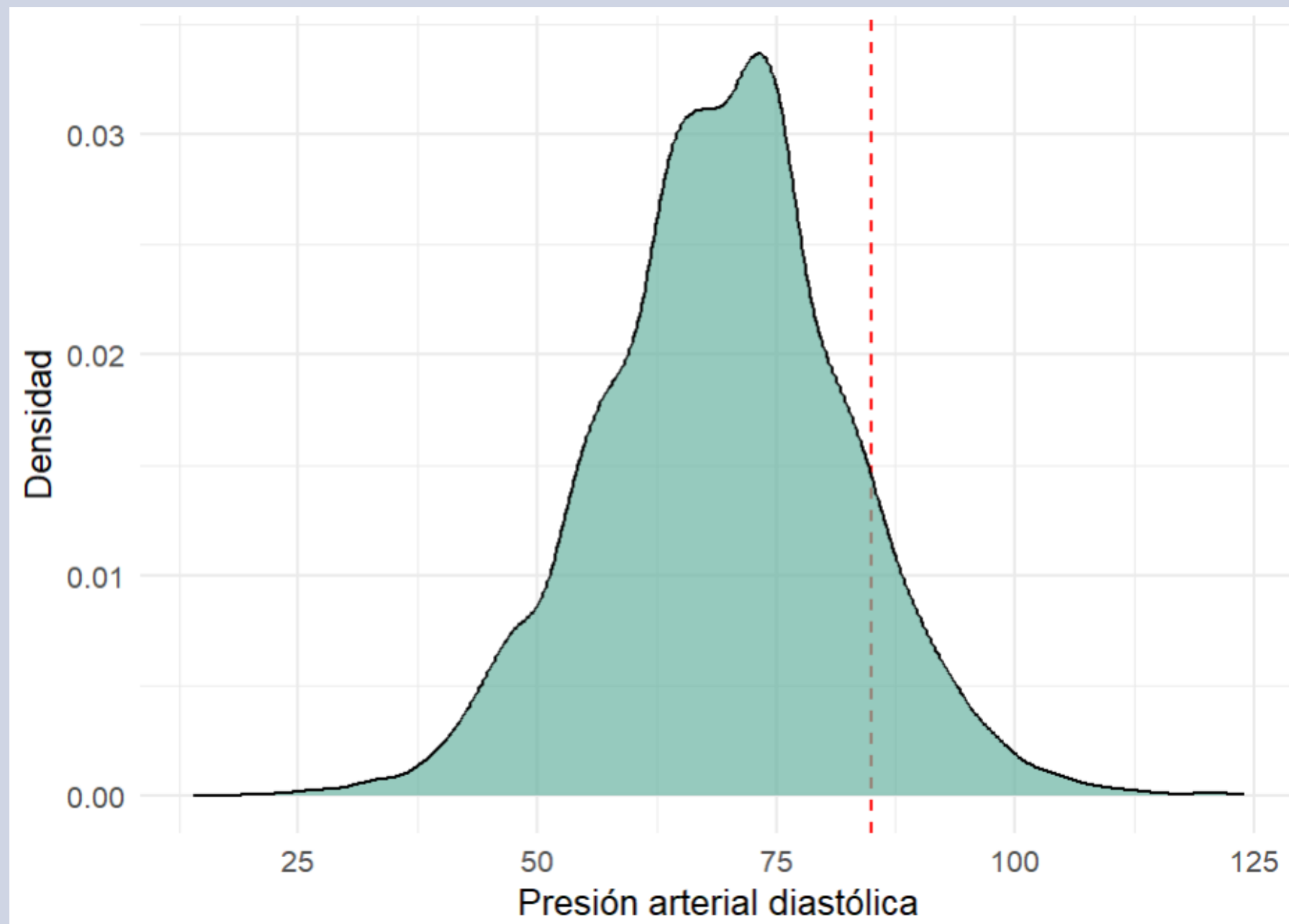
# BASE DE DATOS

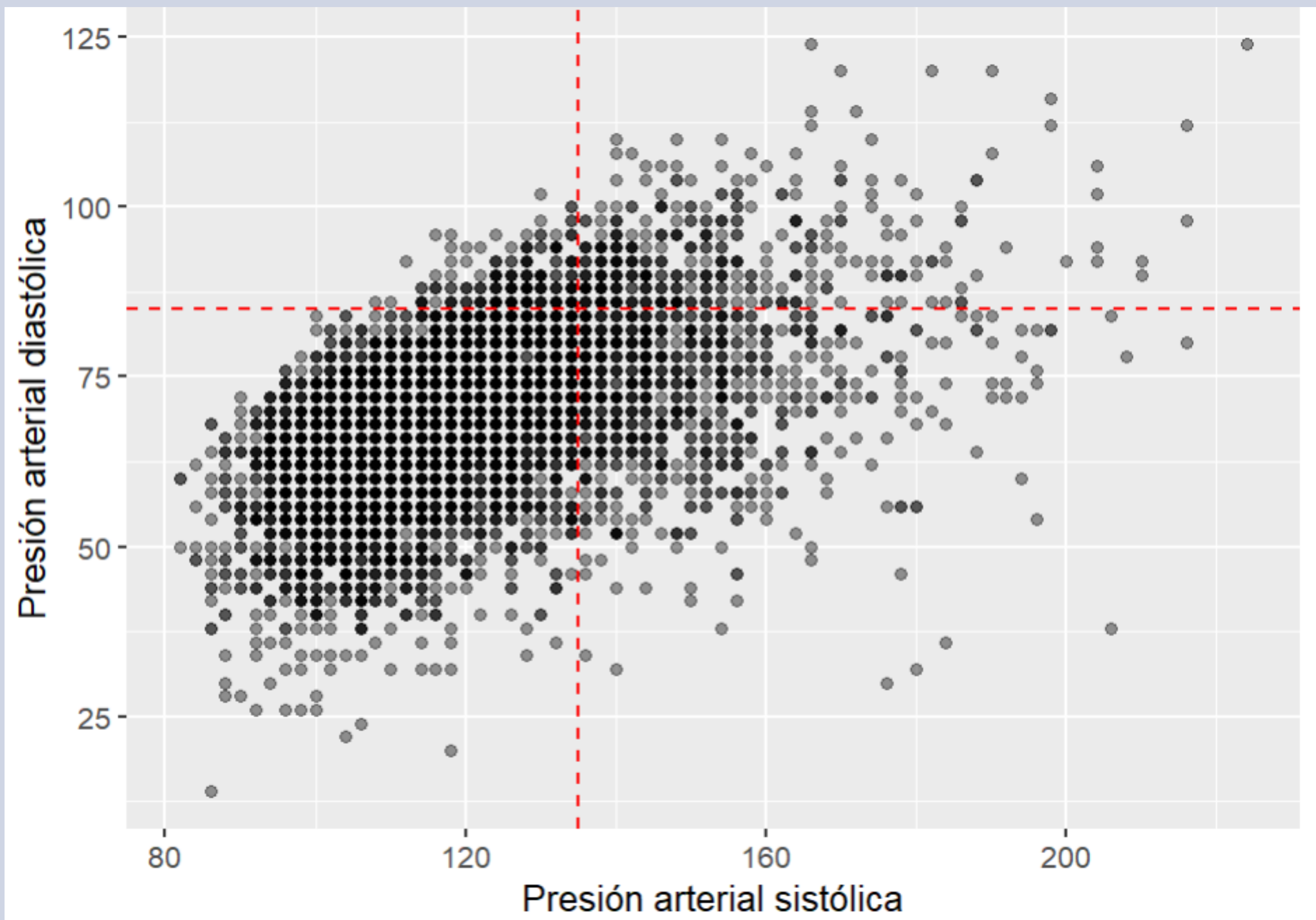
- ① *Fuente*
- ② *Alcance y Limitaciones*
- ③ *Procesamiento*



seqn	age	race_ethnicity	gender	education_child	education_adult	household_income	country_birth	energy_intake
94964	48	Mexican American	Male	NA	Less than 9th grade	\$35,000 to \$44,999	Others	175
95002	26	Mexican American	Male	NA	9-11th grade (Includes 12th grade with no diploma)	\$35,000 to \$44,999	Others	148
95044	67	Non-Hispanic White	Female	NA	College graduate or above	\$75,000 to \$99,999	US	213
95074	37	Non-Hispanic Black	Female	NA	College graduate or above	\$100,000 and Over	US	184
95105	67	Non-Hispanic White	Male	NA	Some college or AA degree	\$35,000 to \$44,999	US	331
95140	42	Mexican American	Female	NA	Some college or AA degree	\$100,000 and Over	US	238
95184	21	Non-Hispanic White	Female	NA	Some college or AA degree	\$20,000 and Over	US	207
95190	28	Mexican American	Female	NA	Some college or AA degree	\$65,000 to \$74,999	US	364
95202	23	Non-Hispanic White	Female	NA	College graduate or above	\$100,000 and Over	US	287
95210	44	Non-Hispanic White	Female	NA	Some college or AA degree	\$100,000 and Over	Others	145
95215	52	Non-Hispanic White	Female	NA	Some college or AA degree	\$ 0 to \$ 4,999	US	277
95227	22	Non-Hispanic Asian	Male	NA	College graduate or above	\$75,000 to \$99,999	Others	185
95266	58	Non-Hispanic White	Female	NA	High school graduate/GED or equivalent	\$100,000 and Over	US	234
95326	72	Non-Hispanic White	Male	NA	College graduate or above	\$75,000 to \$99,999	US	249
95393	28	Non-Hispanic Asian	Female	NA	College graduate or above	\$100,000 and Over	Others	232
95394	47	Mexican American	Male	NA	9-11th grade (Includes 12th grade with no diploma)	\$35,000 to \$44,999	Others	247

# ANALISIS DE LAS VARIABLES OBJETIVO



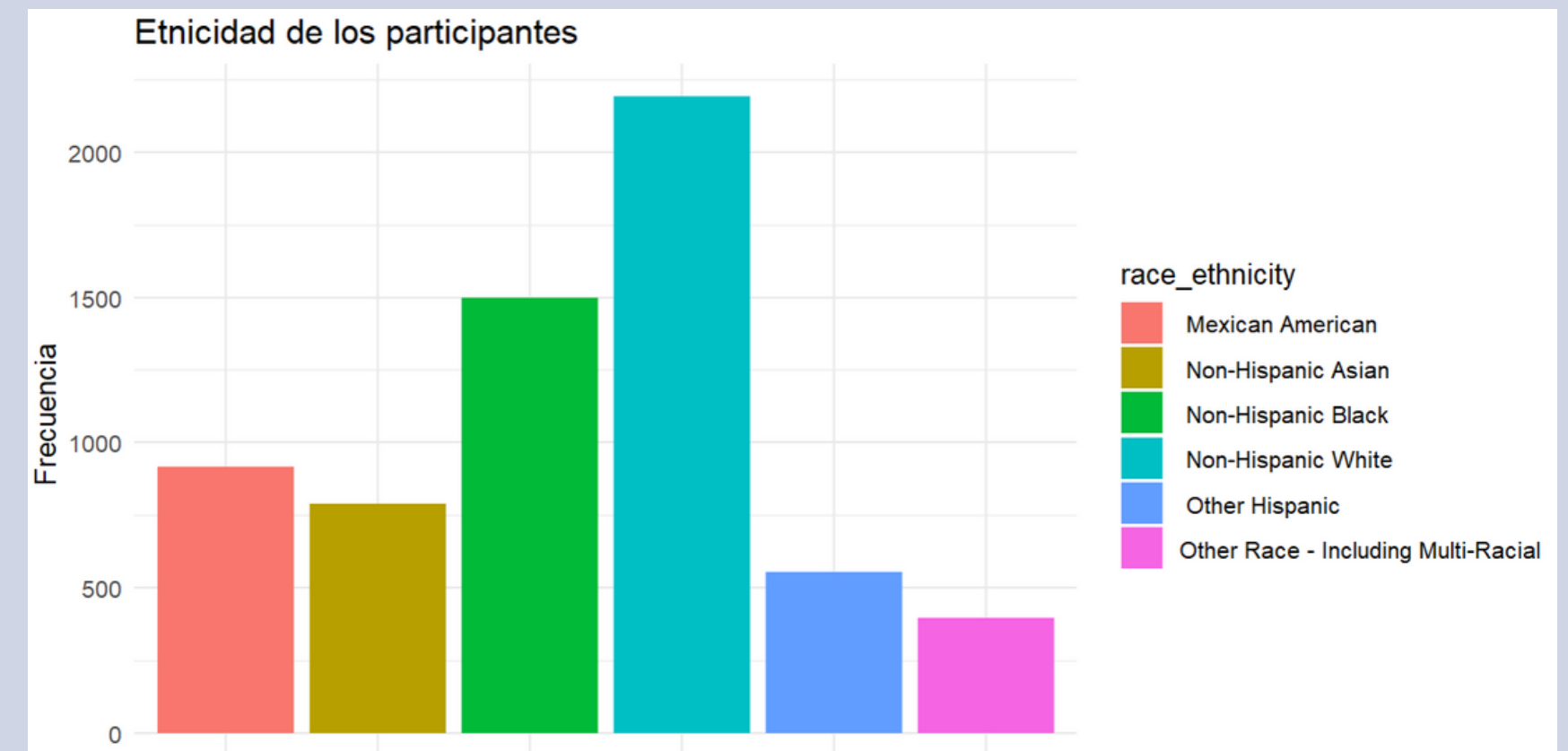
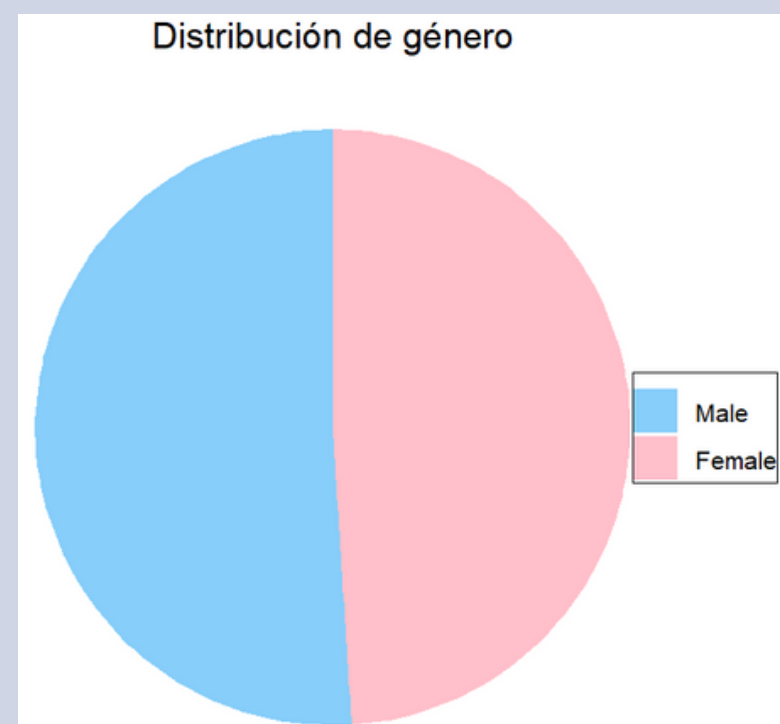
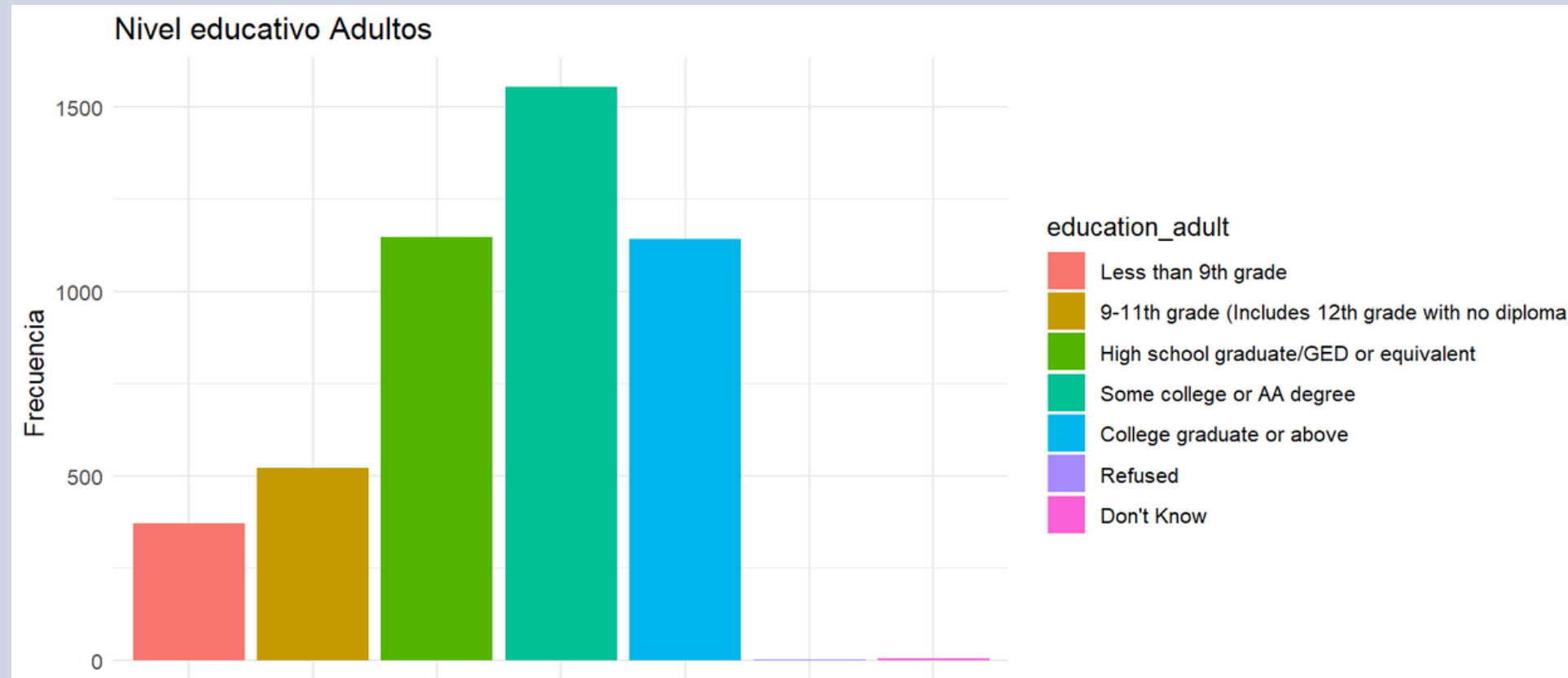


# ANALISIS DE DATOS FALTANTES

education_child	74.6693955
total_percent_fat	46.8828715
alcohol_past_12mo	32.4464736
education_adult	25.3463476
alcohol_ever	24.0396725
energy_intake_dr2	13.2399244
sugar_intake	13.2399244
alcohol_intake_dr2	13.2399244
water_intake_dr2	13.2399244
sodium_table_dr2	12.9722922
systolic_bp	9.6032746
diastolic_bp	9.6032746
monthly_income	5.5100756
household_income	4.1404282
weight	0.9130982

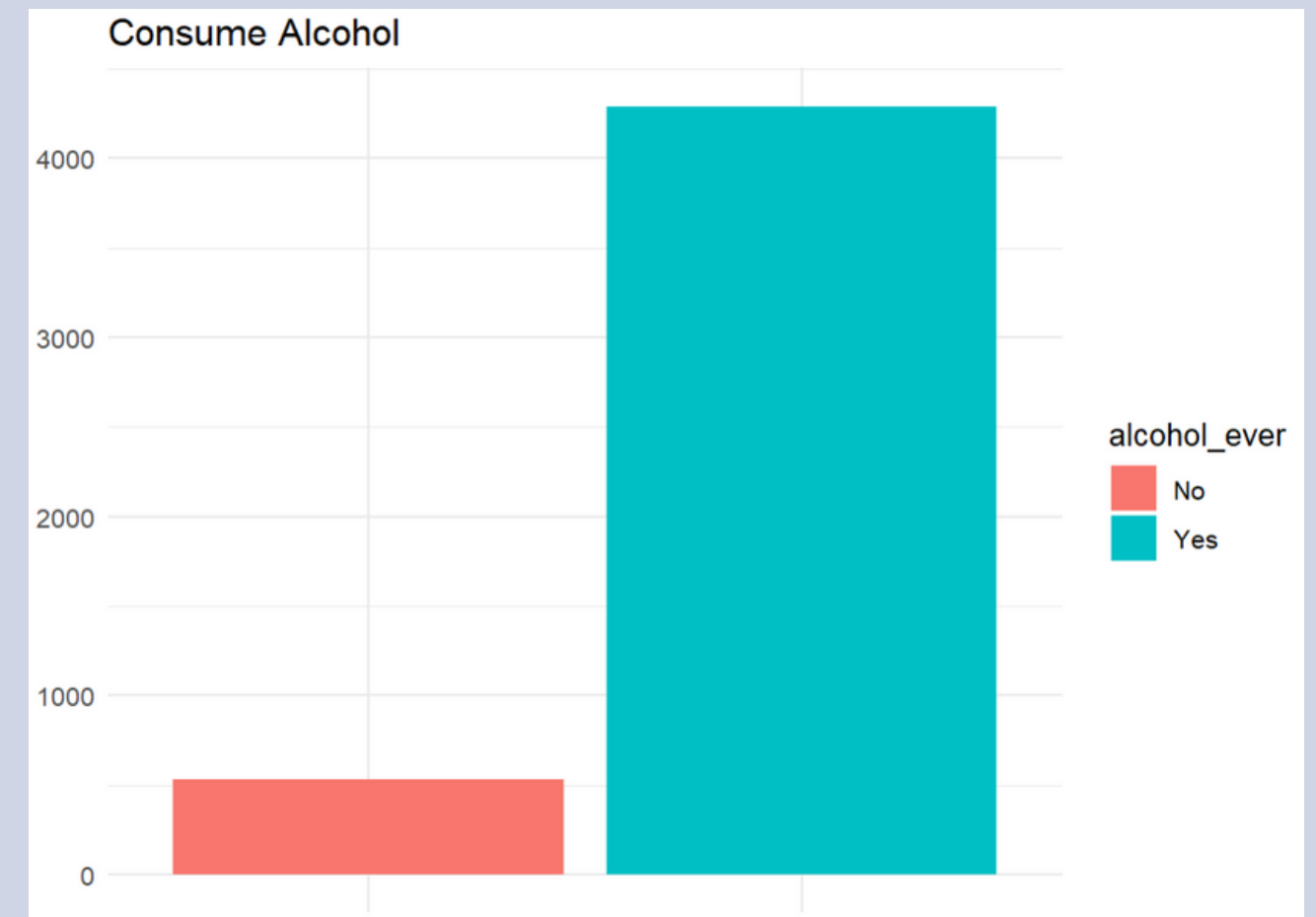
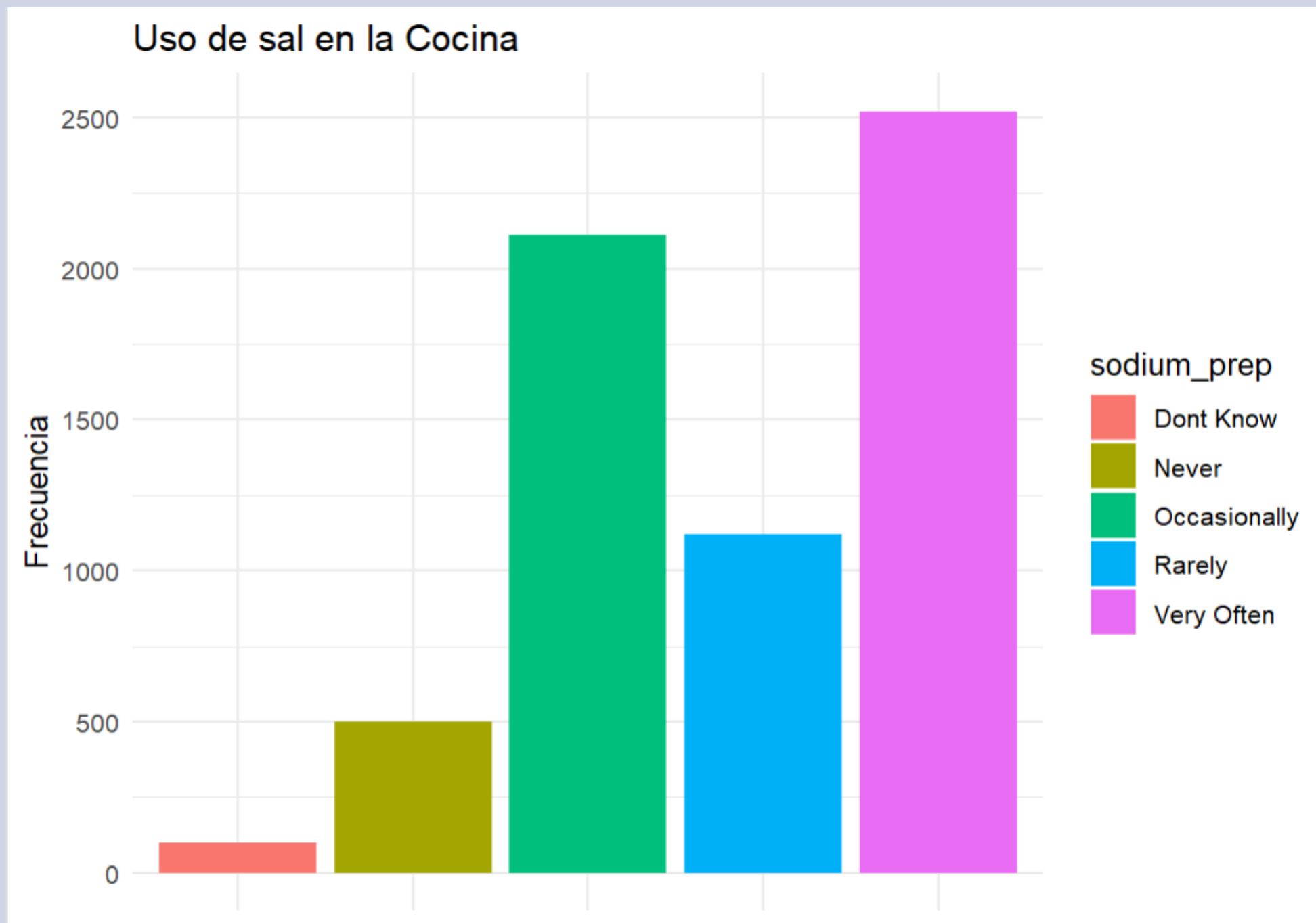
seqn	0.0000
age	0.0000
race_ethnicity	0.0000
gender	0.0000
country_birth	0.0000
energy_intake	0.0000
DR1DRSTZ	0.0000
sodium_prep	0.0000
sodium_table	0.0000
water_intake	0.0000
alcohol_intake	0.0000
health_insurance	0.0000

# ANALISIS DE VARIABLES CATEGORICAS





# ANALISIS DE VARIABLES CATEGORICAS

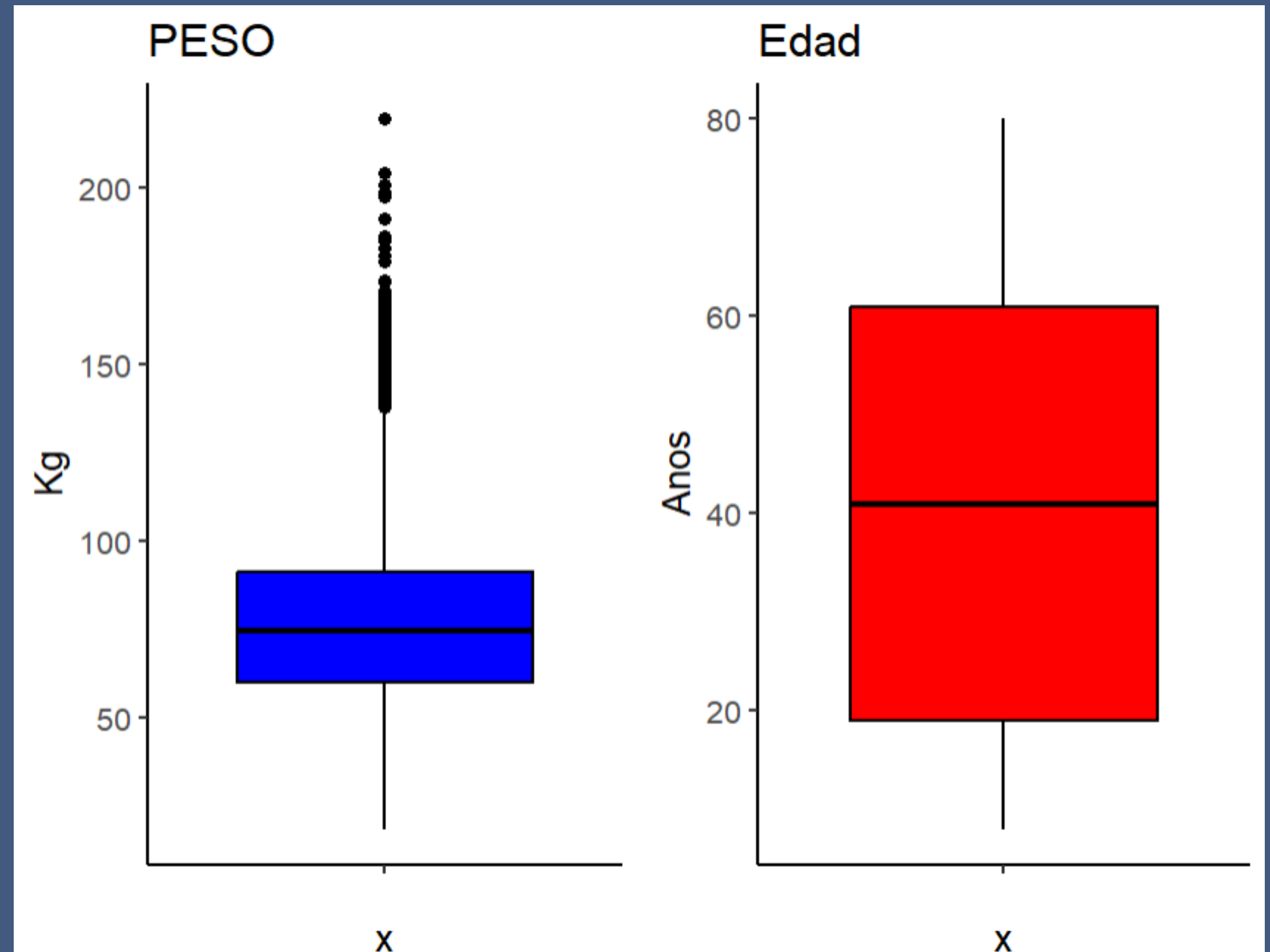


# ANALISIS VARIABLES NUMERICAS

water_intake	
Min.	: 0
1st Qu.:	300
Median	: 795
Mean	: 1110
3rd Qu.:	1521
Max.	: 15360

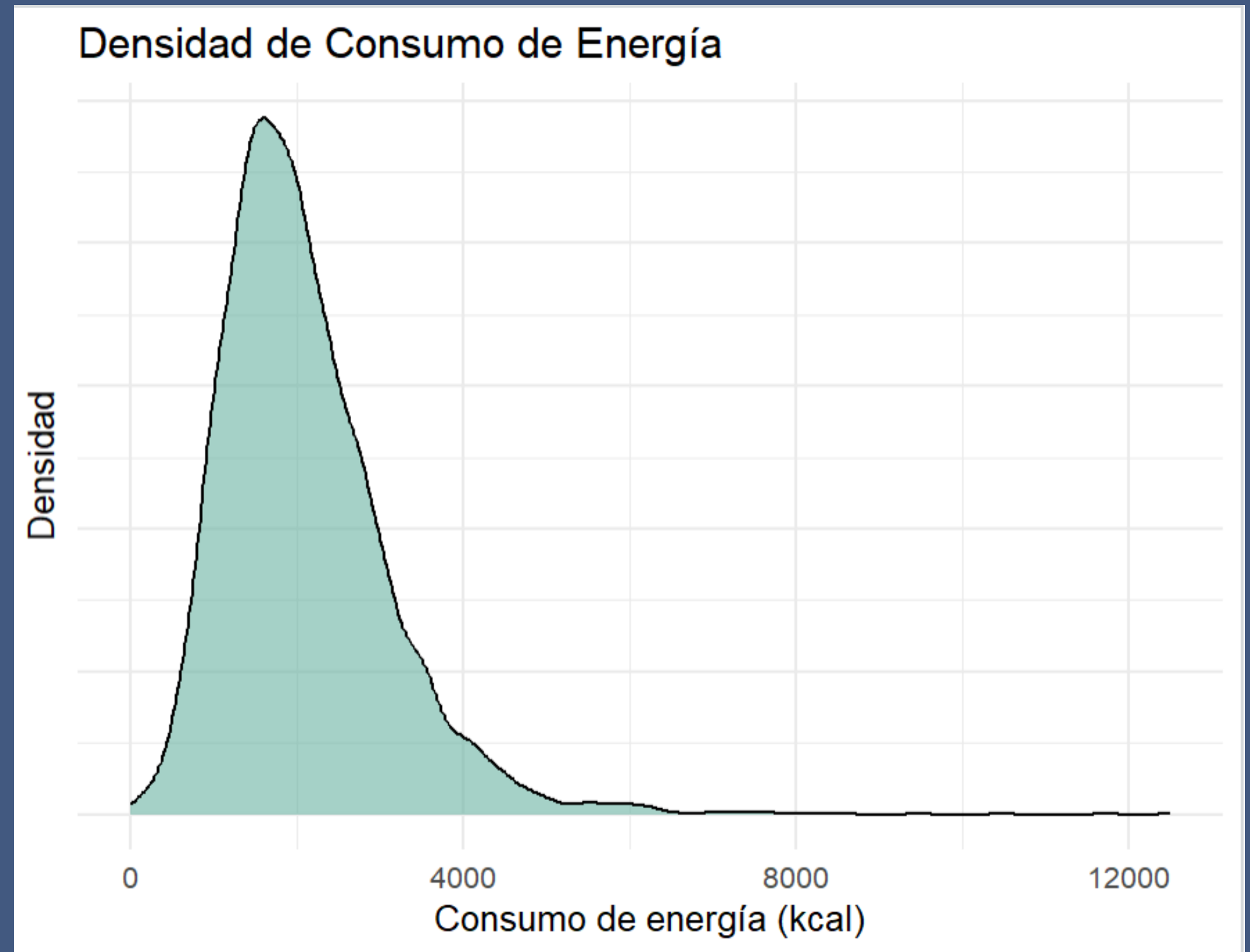
---

alcohol_intake	
Min.	: 0.000
1st Qu.:	0.000
Median	: 0.000
Mean	: 6.891
3rd Qu.:	0.000
Max.	: 406.800

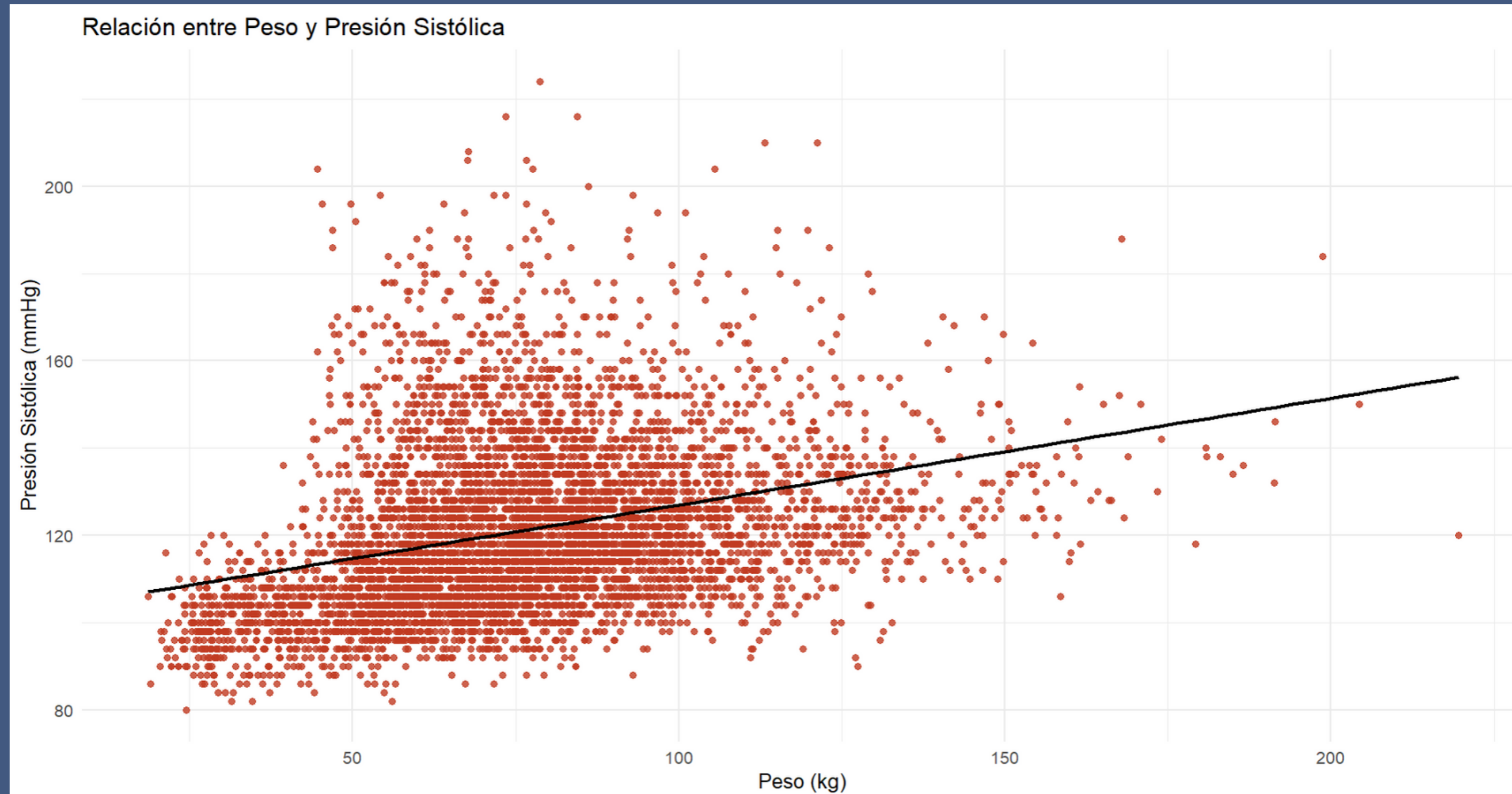


# ANALISIS VARIABLES NUMERICAS

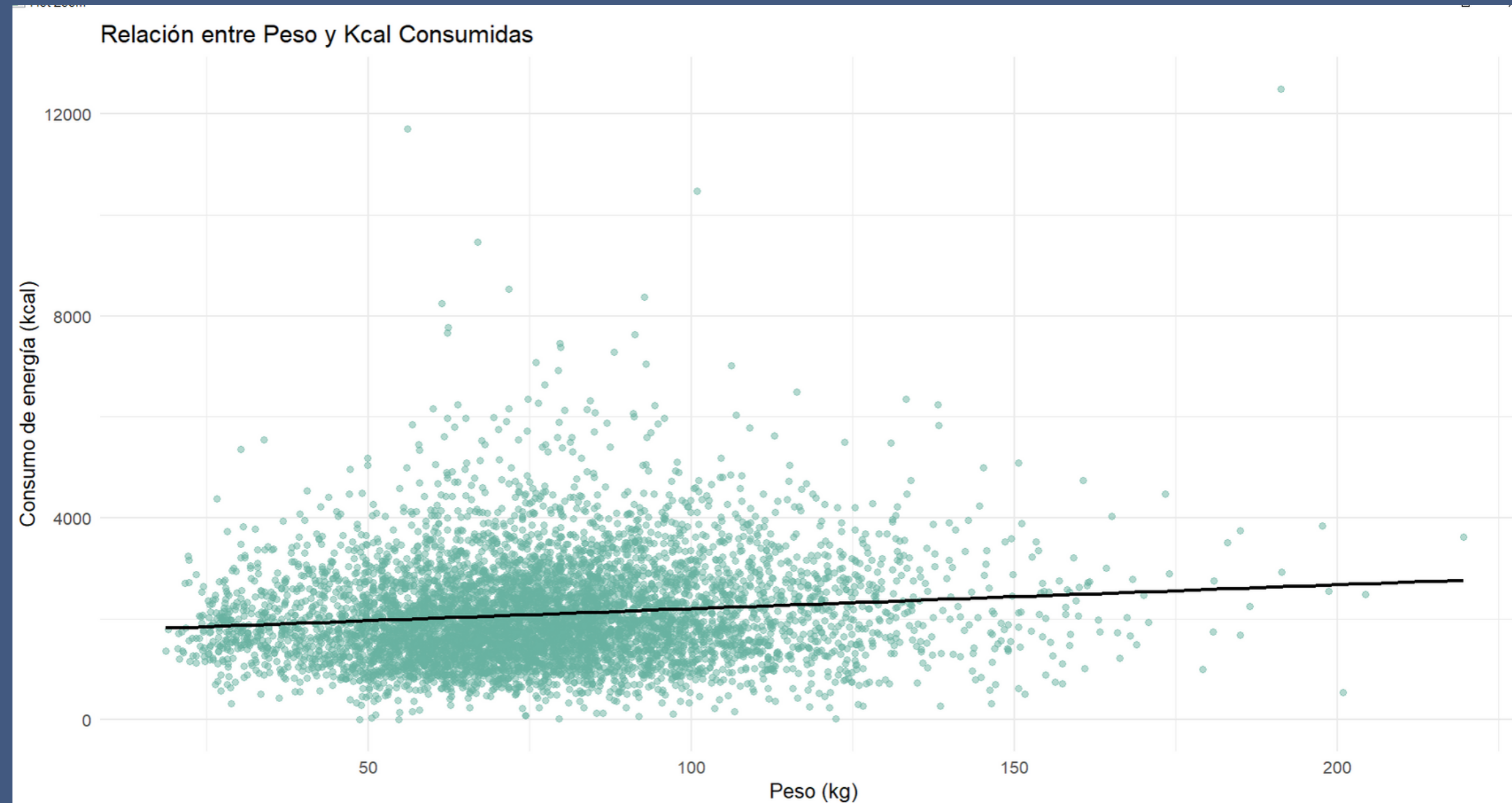
```
total_percent_fat
Min.      :12.10
1st Qu.   :25.90
Median    :32.10
Mean      :32.32
3rd Qu.   :39.20
Max.      :56.10
NA's      :2978
```



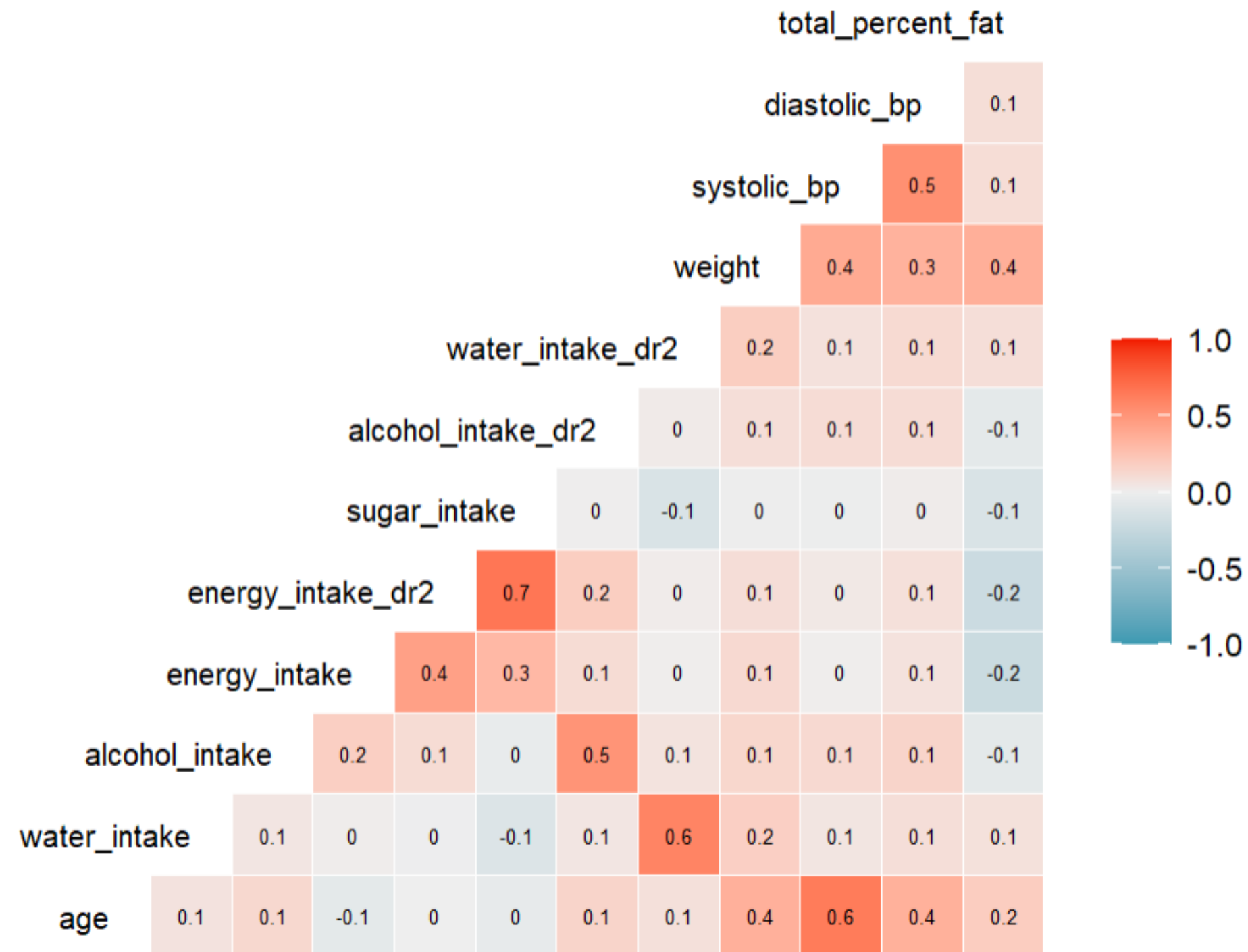
# ANALISIS CORRELACIONES



# ANALISIS CORRELACIONES



# ANALISIS CORRELACIONES





# ANALISIS DE CATEGORICAS Y OBJETIVO

## PRUEBAS ANOVA

### *Education*

	Df	Sum Sq	Mean Sq	F value
education_adult	6	25543	4257	11.52
Residuals	4274	1579722	370	
Pr(>F)				
education_adult	7.97e-13	***		

### *Race\_ethnicity*

	Df	Sum Sq	Mean Sq	F value
race_ethnicity	5	27988	5598	15.17
Residuals	4275	1577276	369	
Pr(>F)				
race_ethnicity	8.37e-15	***		

### *Gender*

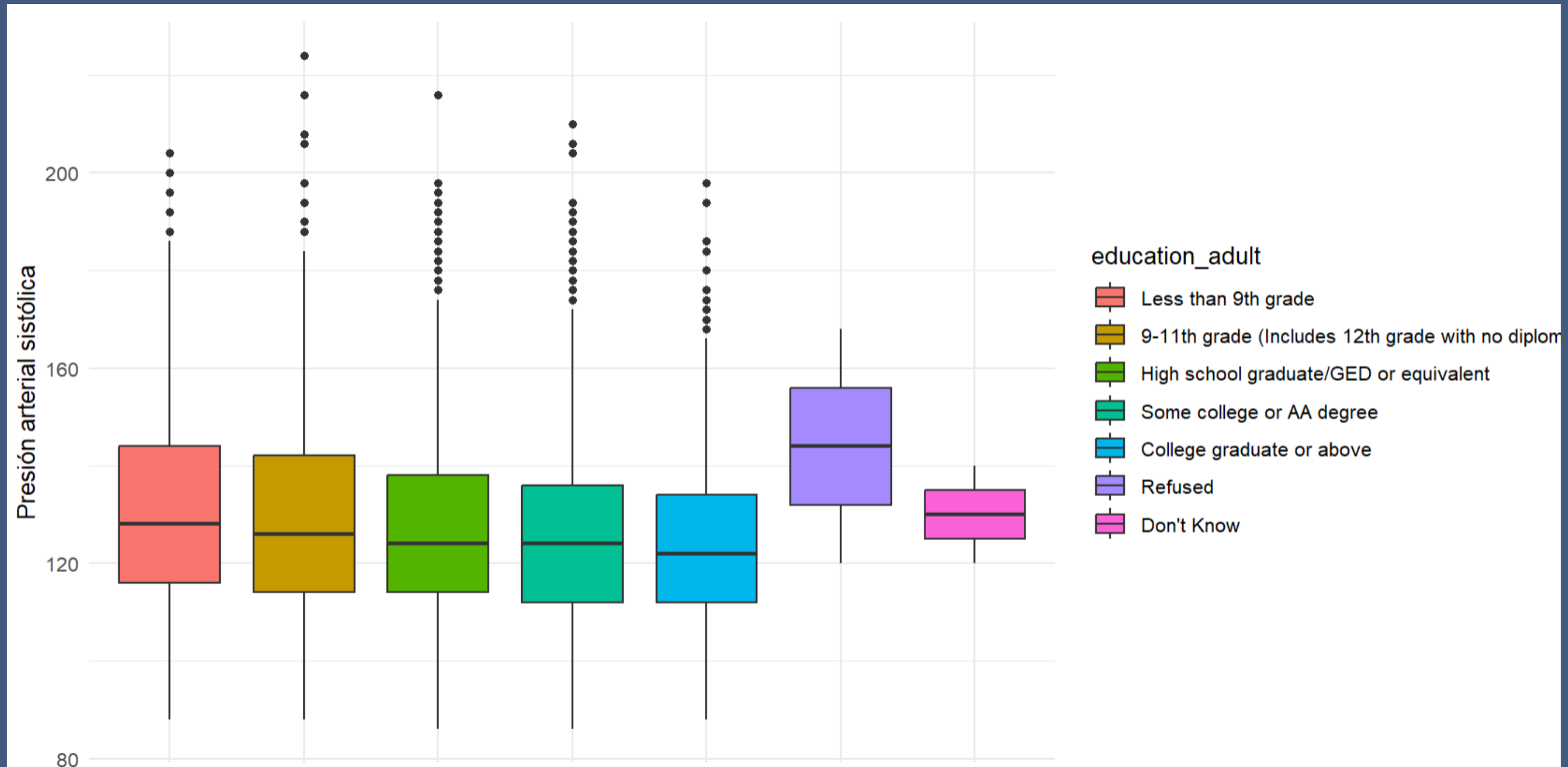
	Df	Sum Sq	Mean Sq	F value
gender	1	6267	6267	16.77
Residuals	4279	1598998	374	
Pr(>F)				
gender	4.3e-05	***		

### *Sodium\_prep*

	Df	Sum Sq	Mean Sq	F value
sodium_prep	4	20046	5012	13.52
Residuals	4276	1585218	371	
Pr(>F)				
sodium_prep	5.95e-11	***		

# ANALISIS DE CATEGORICAS Y OBJETIVO

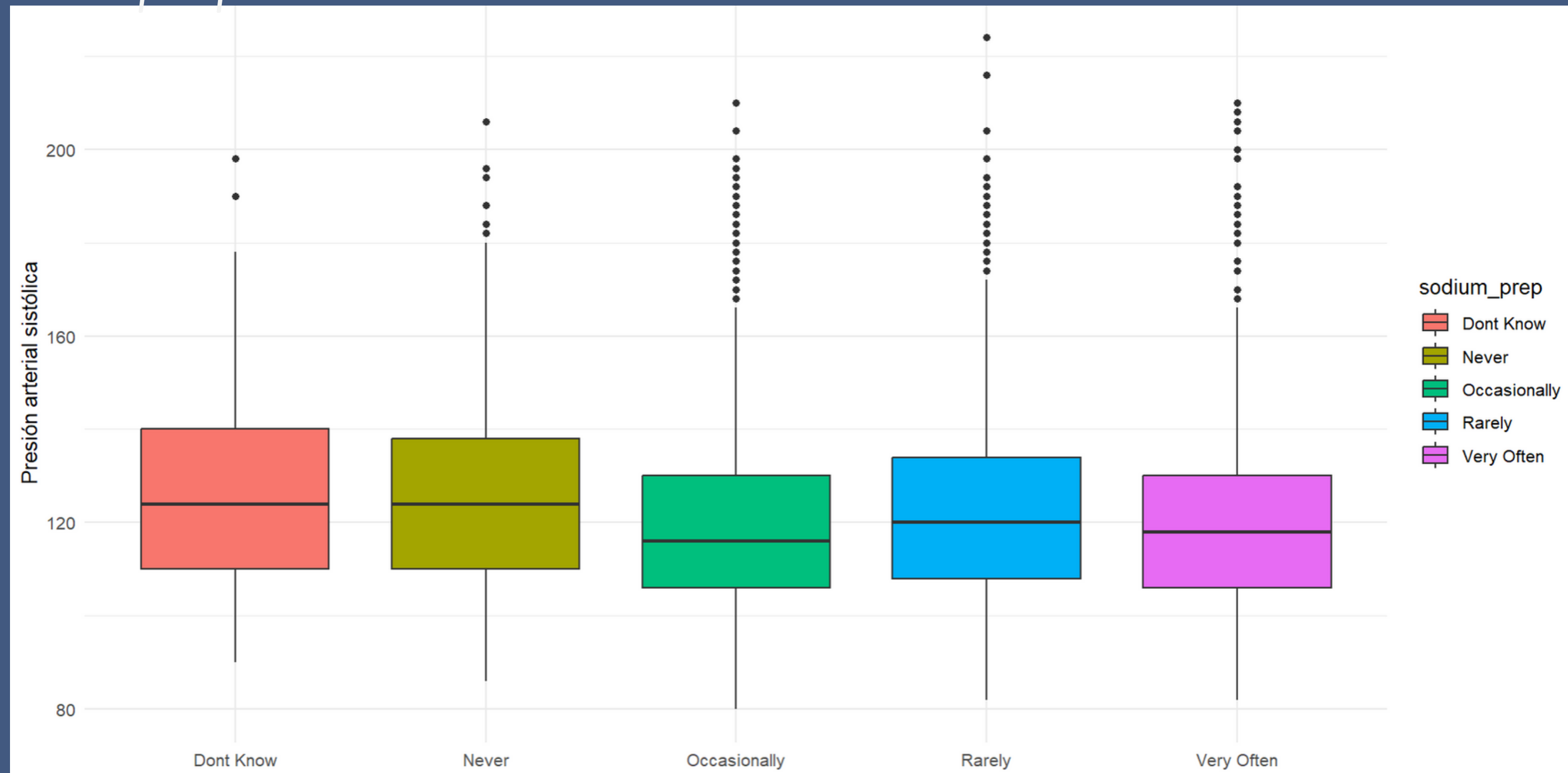
## *Education*





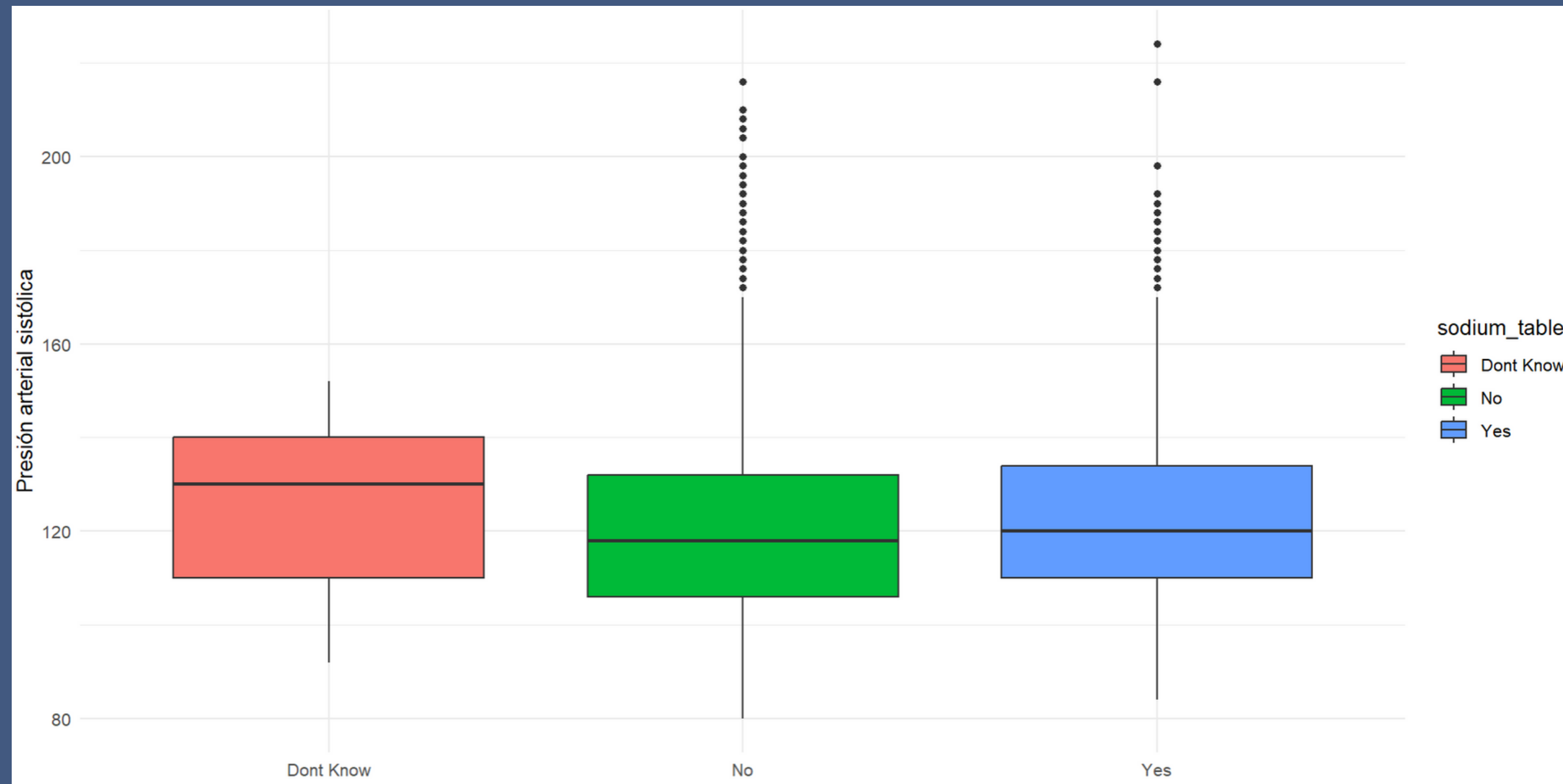
# ANALISIS DE CATEGORICAS Y OBJETIVO

*Sodium\_prep*



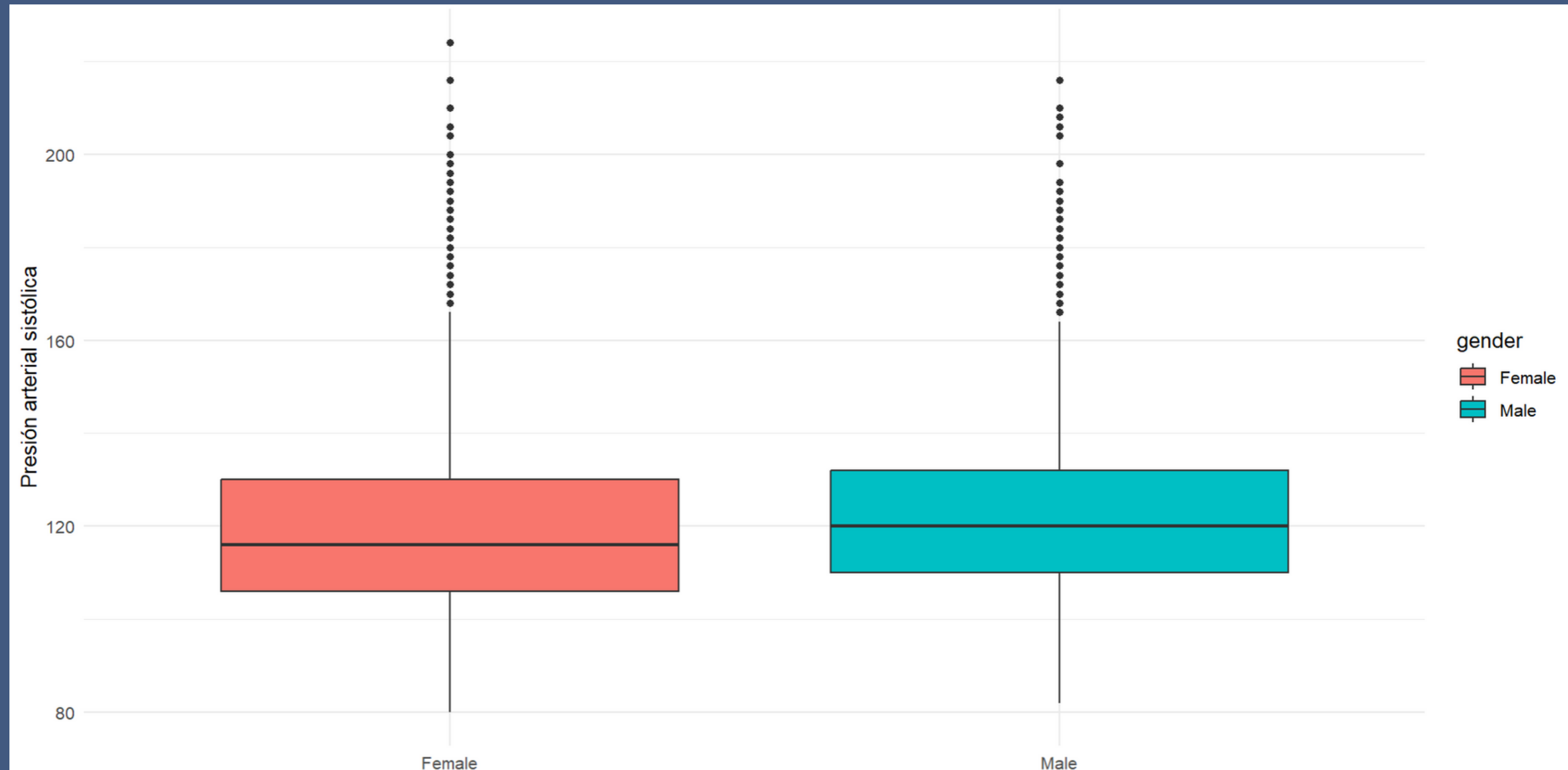
# ANALISIS DE CATEGORICAS Y OBJETIVO

# Sodium\_table



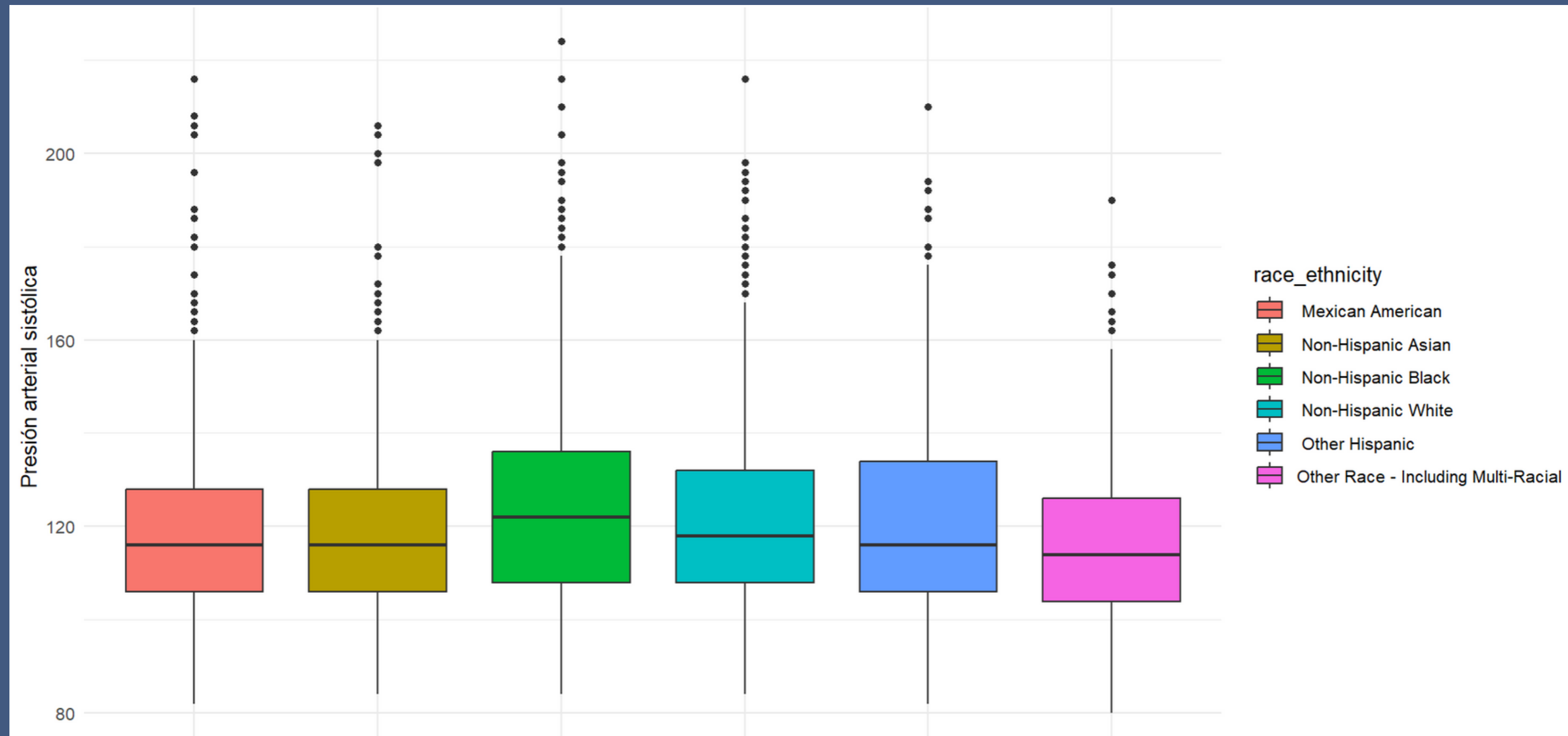
# ANALISIS DE CATEGORICAS Y OBJETIVO

*Gender*



# ANALISIS DE CATEGORICAS Y OBJETIVO

*Race\_ethnicity*



# CONCLUSIONES

