Assignment 2

A2-1) HINCOME9 and POVFAM9 (summary + variance)

HINCOME9
Min 23,247 · 1st Qu. 28,877 · Median 31,546 · Mean 32,419 · 3rd Qu. 35,112 · Max 49,584
Variance: 23,659,025

POVFAM9
Min 4.93 · 1st Qu. 8.65 · Median 10.56 · Mean 11.82 · 3rd Qu. 14.23 · Max 36.52
Variance: 21.44704

Commands used: summary(), var()

A2-2)

	sd	cv
HINCOME9	4864.0544	0.1500
POVFAM9	4.6311	0.3918

Commands used: sd(), mean(), CV = sd/mean

A2-3)

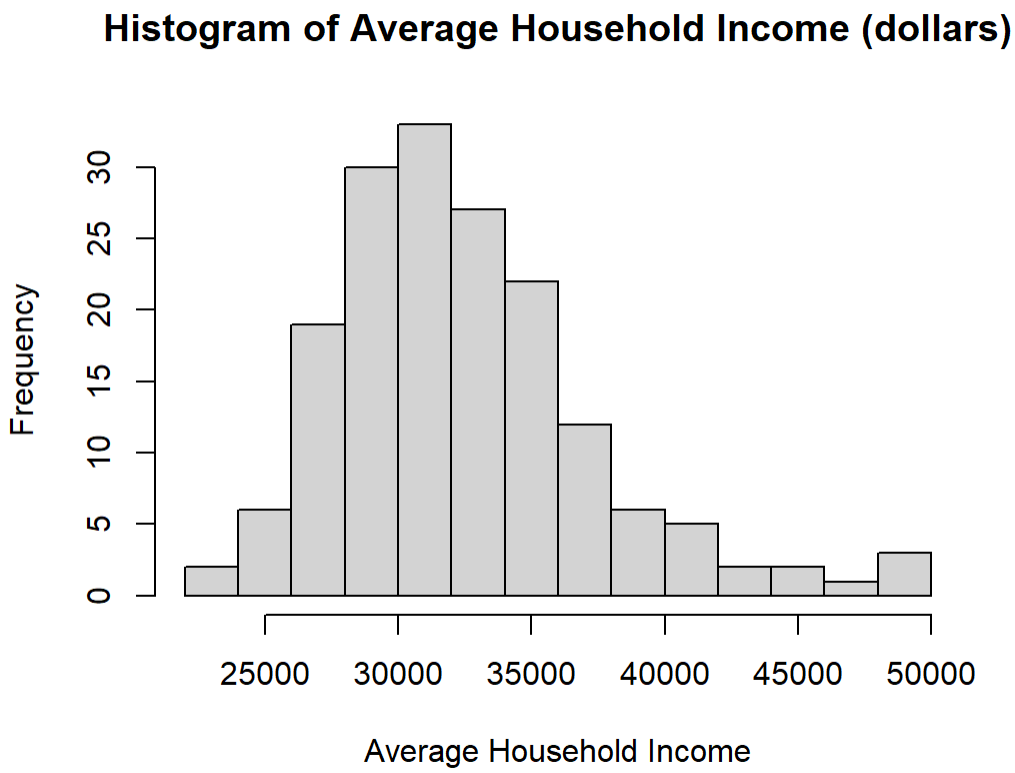
SD suggests Hincome9 is more variable (4864.05 vs 4.63). However, the CV whic h accounts for units shows PovFam9 is more variable (0.3918 vs 0.1500). I use CV for the comparison.

A2-4)

	Region	Frequency	Percent
1	MIDWEST	49	28.82
2	NORTHEAST	14	8.24
3	SOUTH	76	44.71
4	WEST	31	18.24

Commands used: table (), as.data.frame(), names(), round()

A3-1)



Commands used: hist() with main, xlab, breaks=12

A3-2)

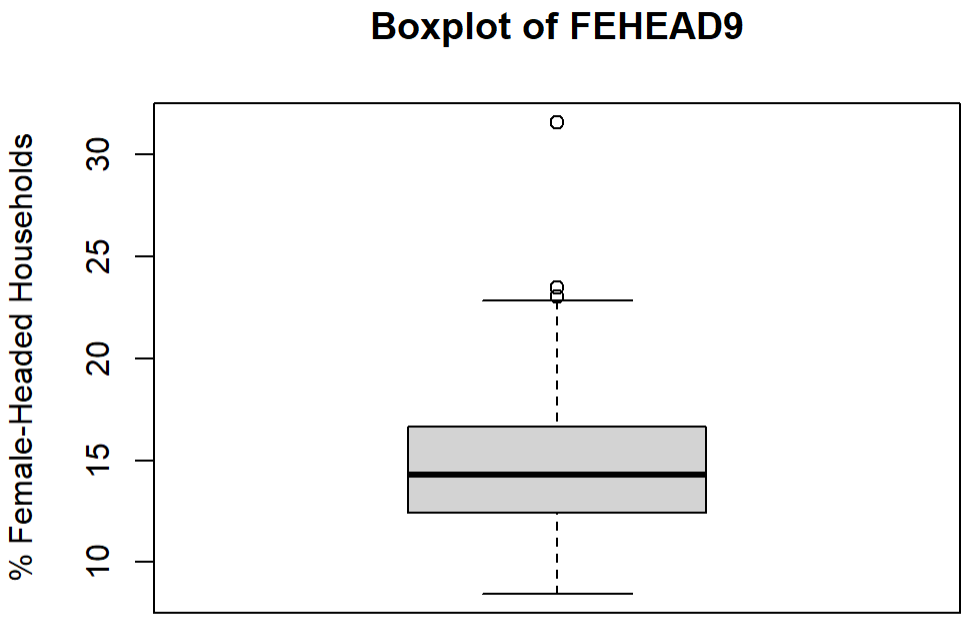
50%	75%
31546.45	35111.89

Commands used: quantile()

A3-3)

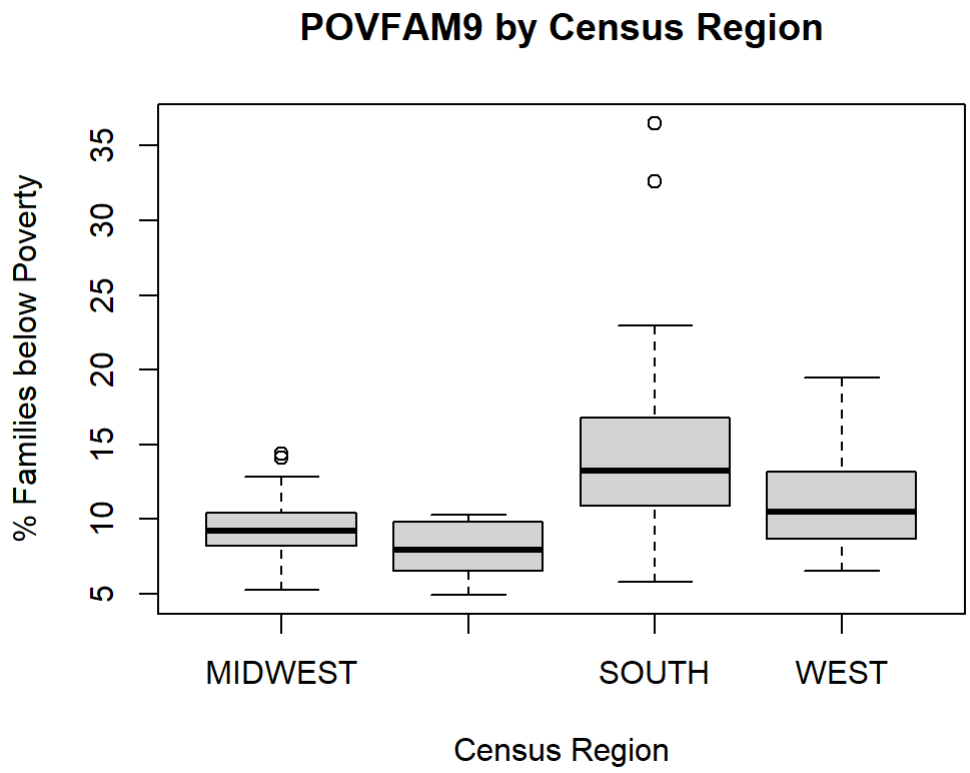
Mostly concentrated around 28–35k, with a right tail out to 50k, so I'd call it positively skewed

A4-1)



Median ≈ 11%, moderate IQR, and a few high outliers.

A4-2)



SOUTH shows the highest median POVFAM9 and widest spread (IQR) with some high outliers. MIDWEST/NORTHEAST are generally lower and tighter; WEST is in between.