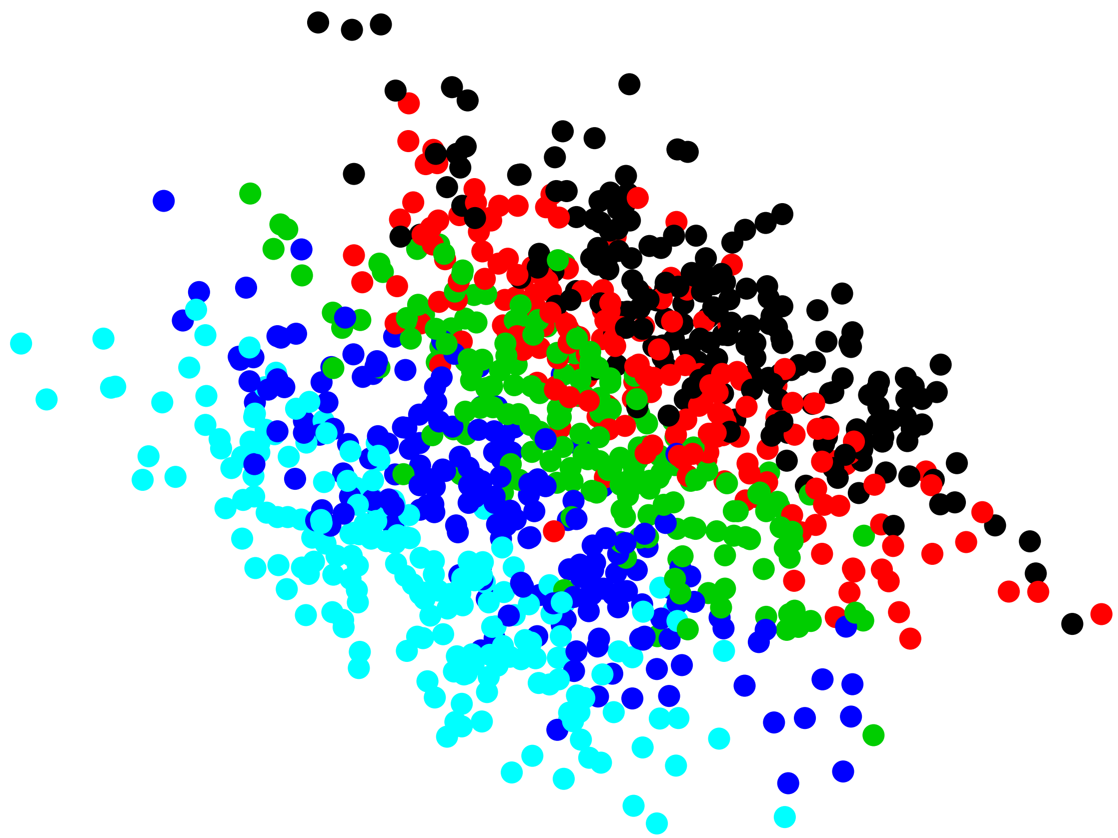
MULTILEVEL MODELS



WHAT ARE MULTILEVEL MODELS?

Multilevel models are an extension of linear regression models used to analyze nested or clustered data.

They are also known as mixed models and hierarchical linear models because why have one name when you can have three?

WHAT ARE MULTILEVEL MODELS?

Nested/clustered data can take many forms:

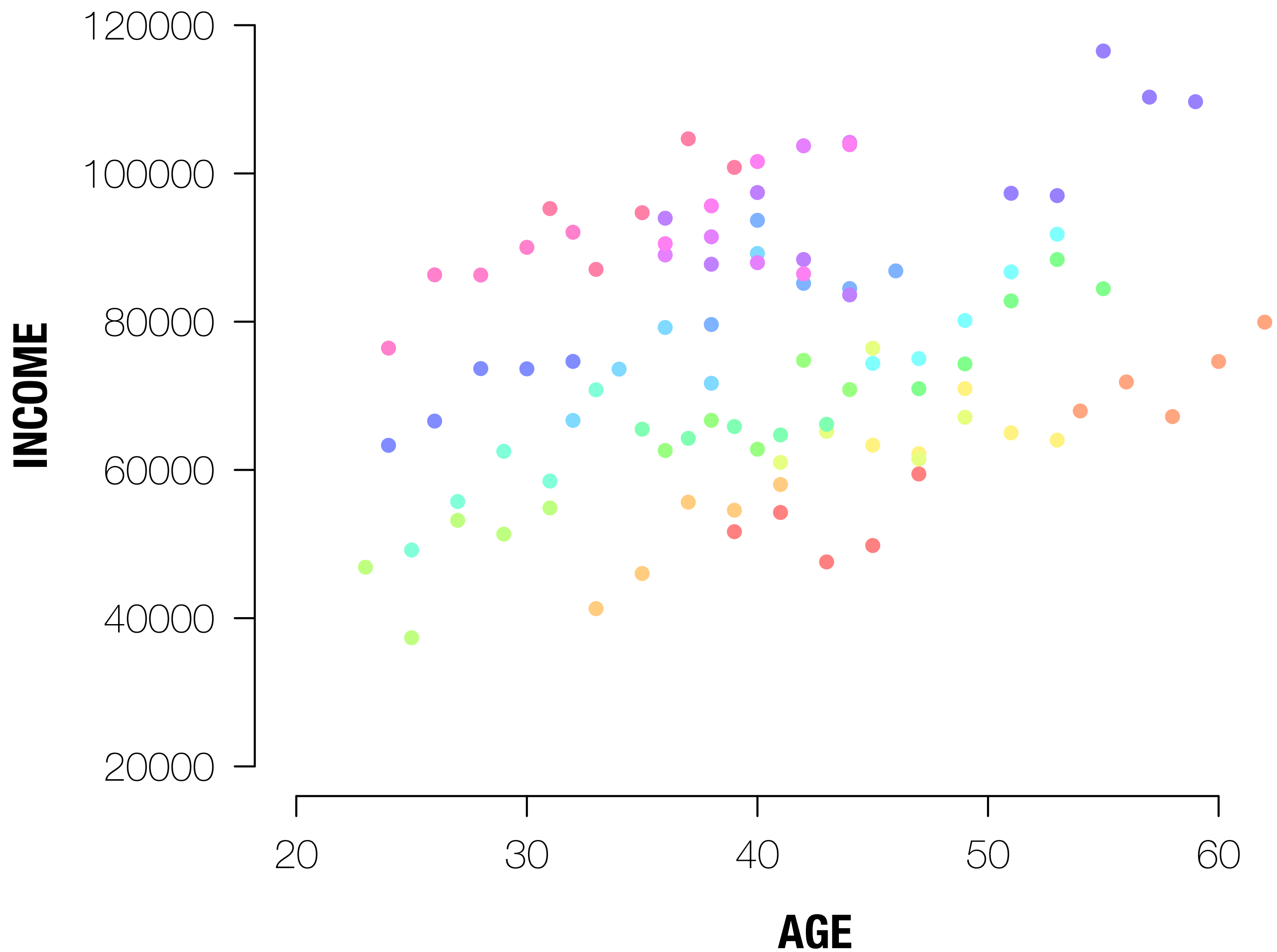
- Individuals within regions
- Longitudinal or within-subjects designs (observations within individuals)
- Partners within a couple
- Children in classrooms (in schools)

WHY USE MULTILEVEL MODELS?

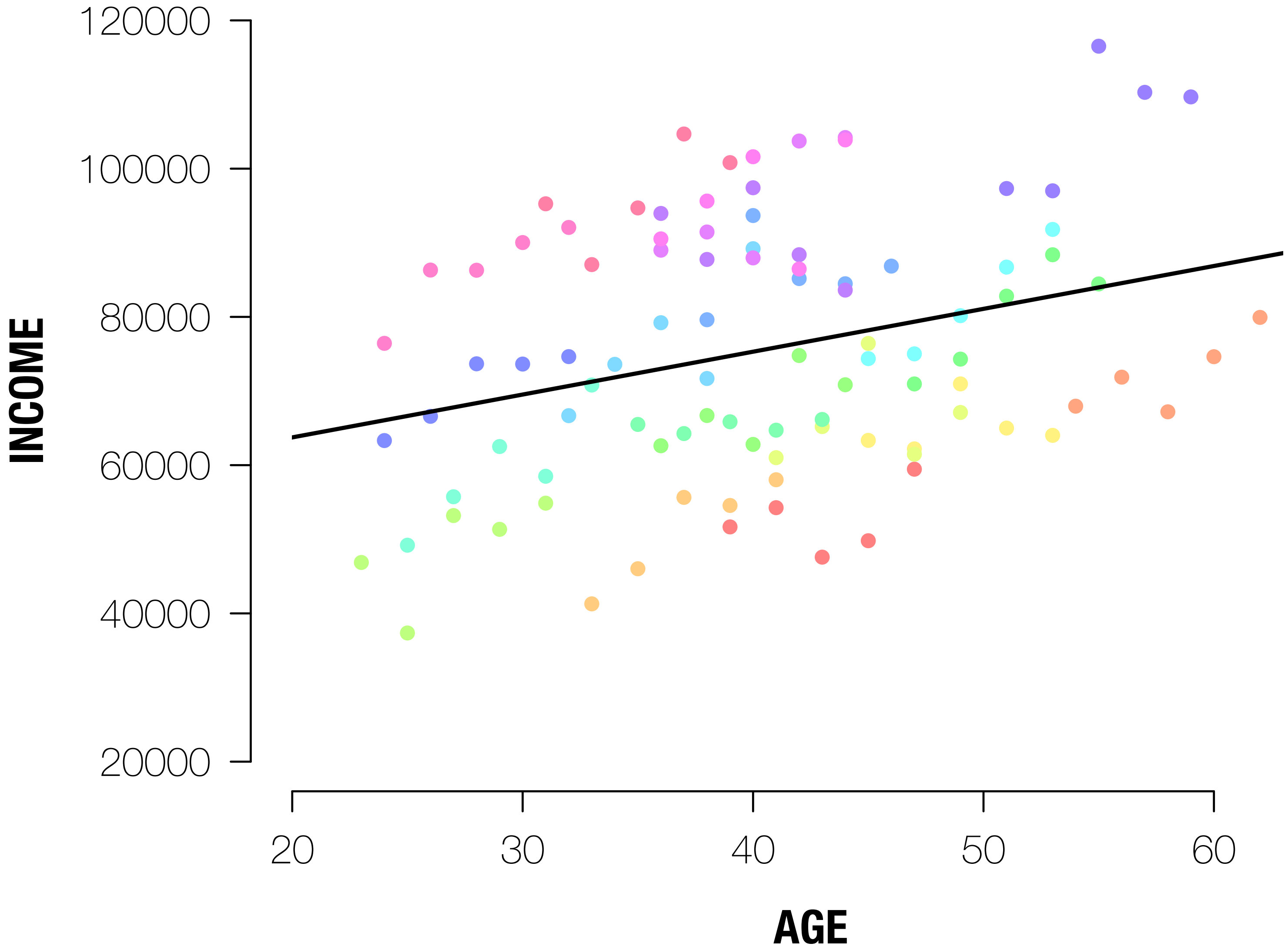
Nested/clustered data (may) violate the assumption made in normal linear modelling of independent errors.

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

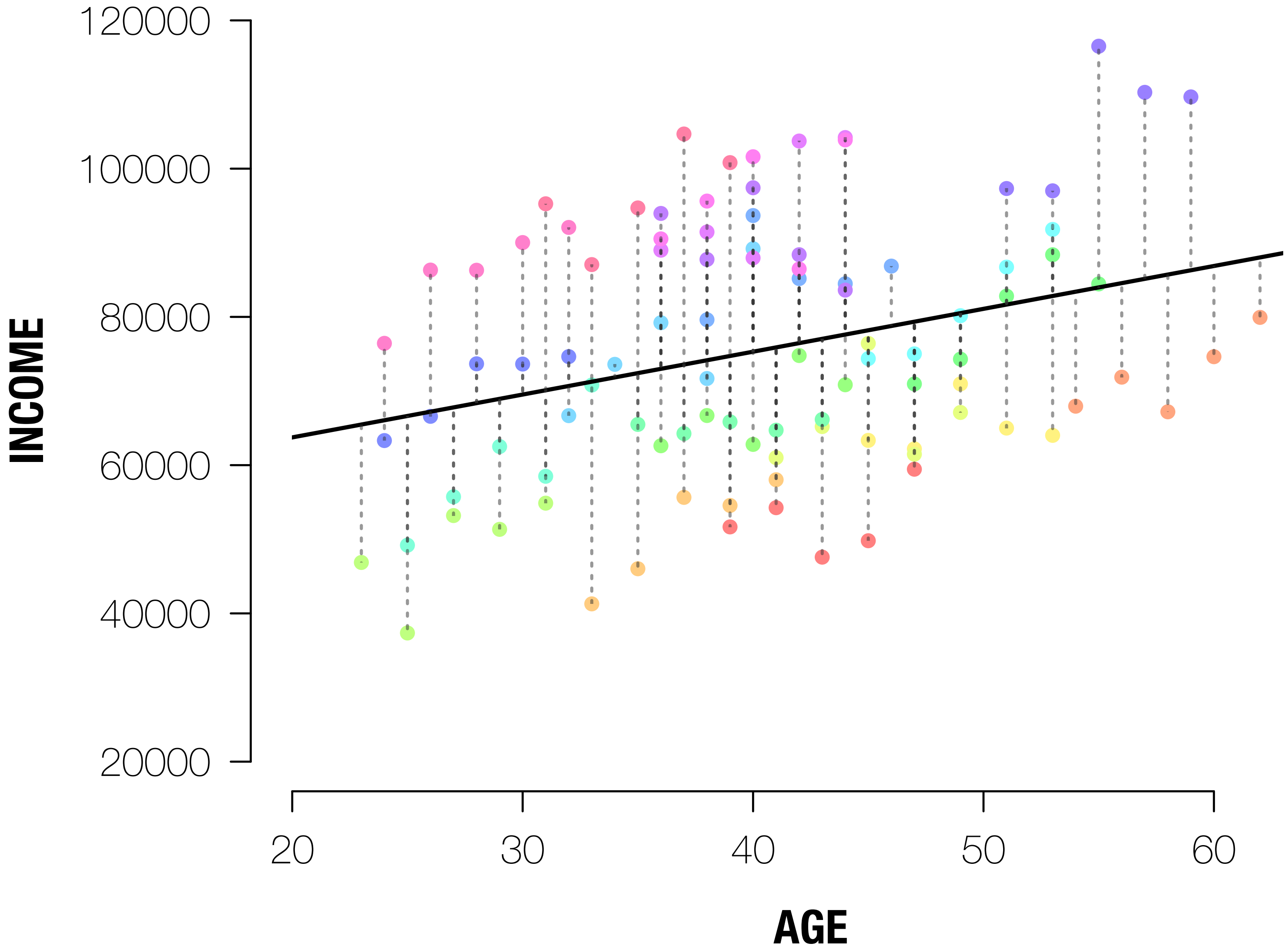
WHY USE MULTILEVEL MODELS?



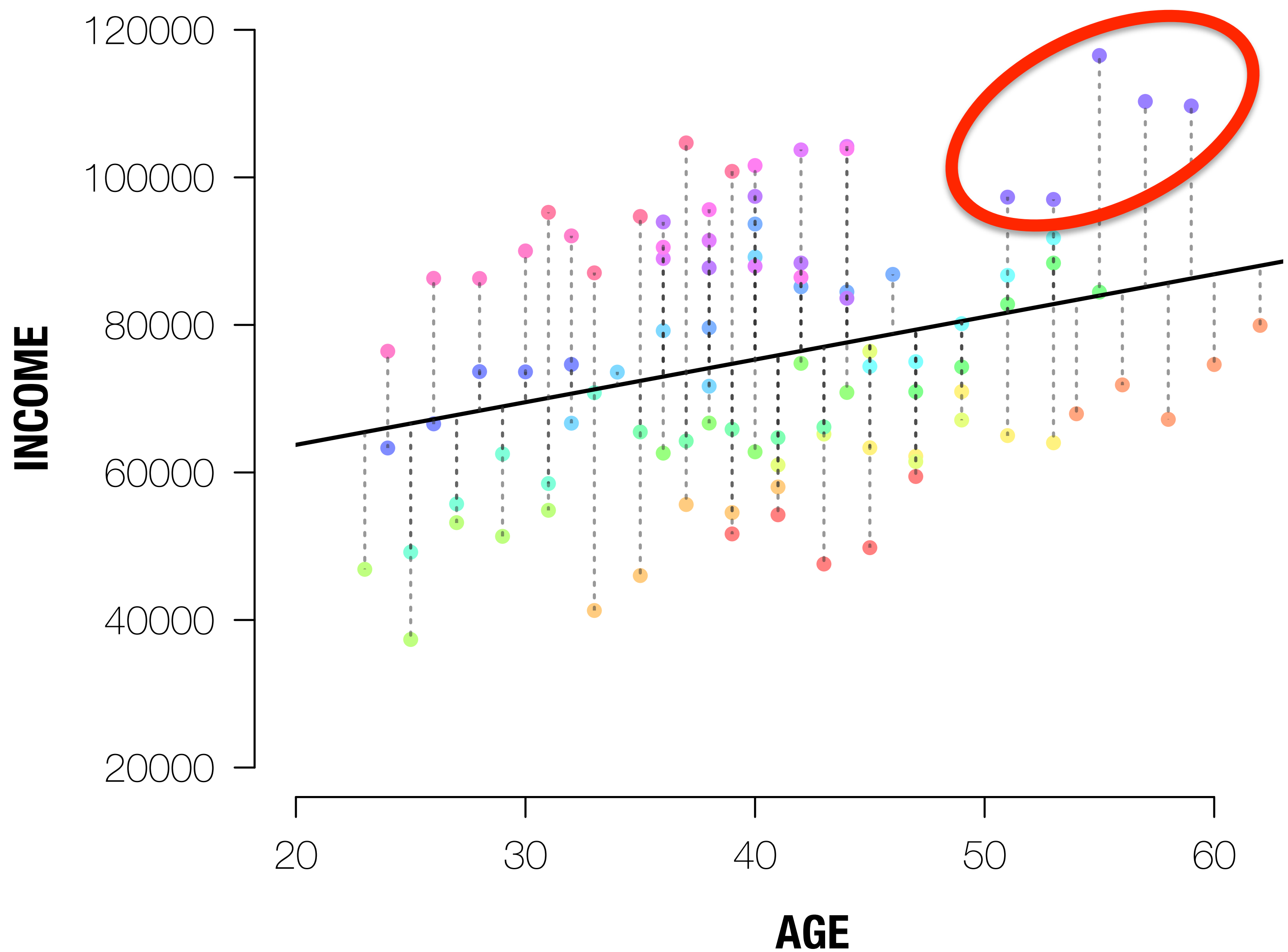
WHY USE MULTILEVEL MODELS?



WHY USE MULTILEVEL MODELS?



WHY USE MULTILEVEL MODELS?



WHY USE MULTILEVEL MODELS?

Relationships can differ across levels:

- E.g. what is the between-country relationship between meat consumption and physical health?
- What is the within-country relationship?

WHAT DO MULTILEVEL MODELS DO?

Multilevel models allow us to account for nested/clustered data structures and estimate relationships at different levels.

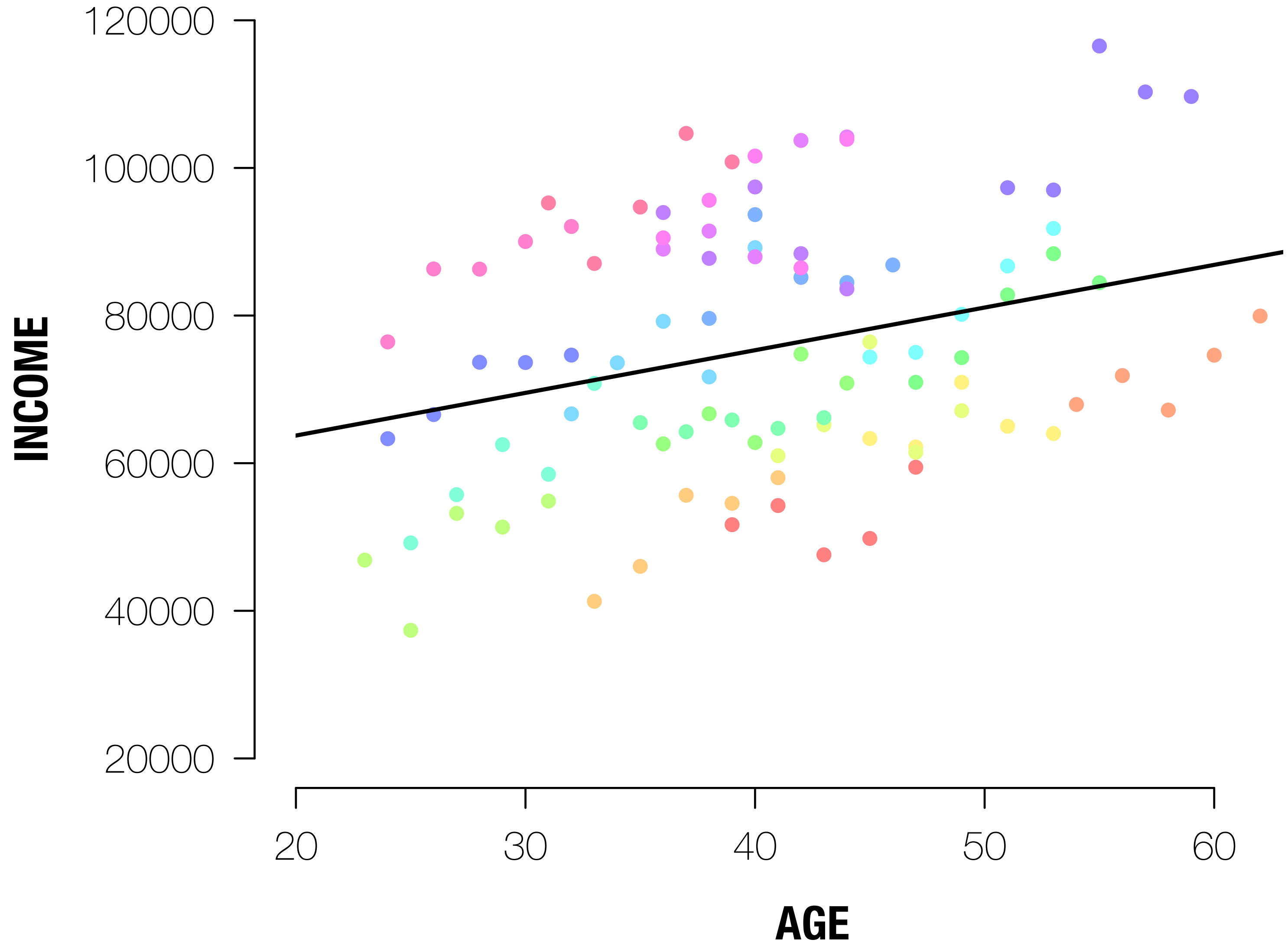
RANDOM INTERCEPTS

The simplest and most common multilevel model extends on linear regression by including a random intercept.

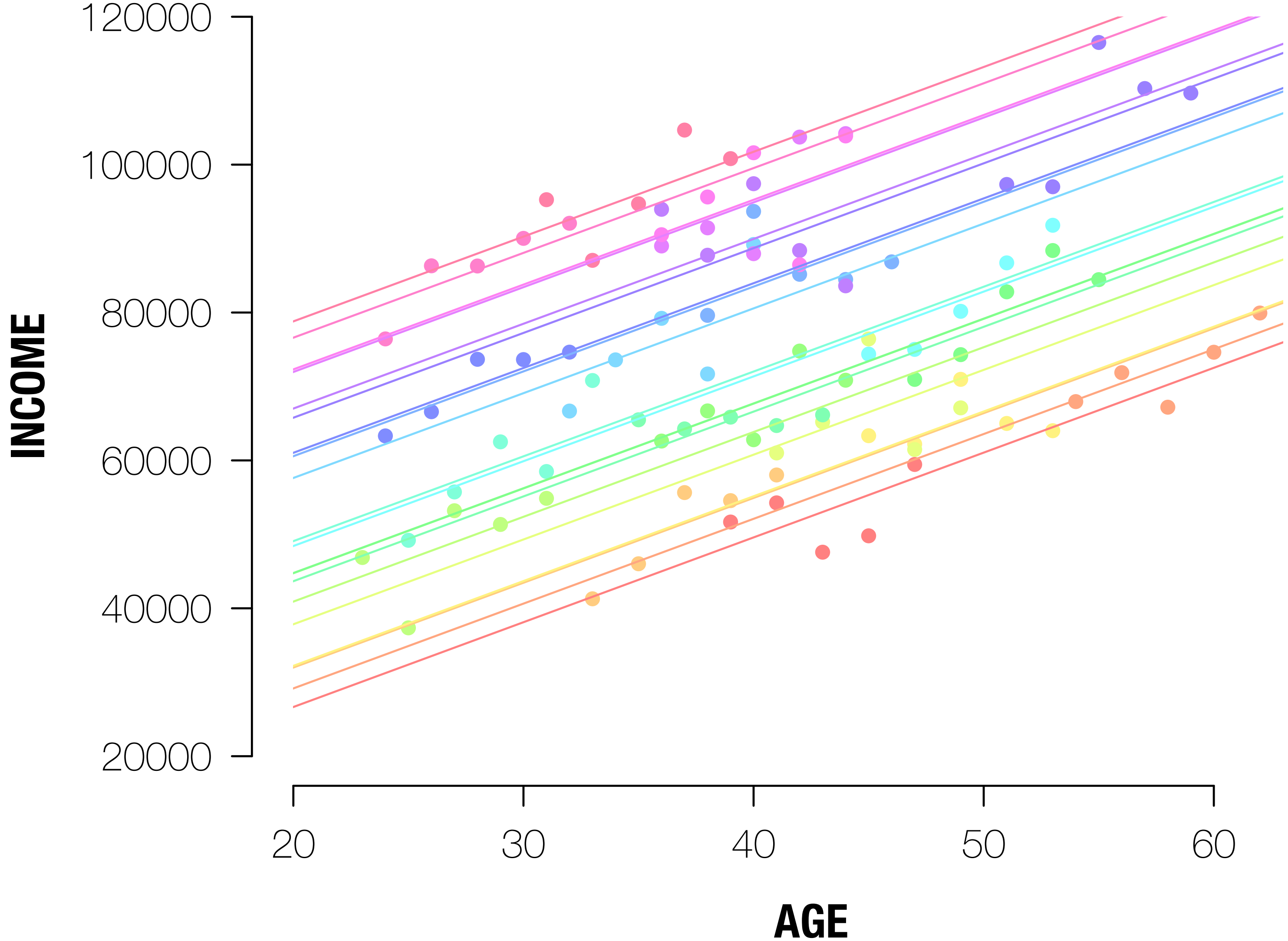
This divides a random error term ε_i into cluster-level error ζ_j (we usually index level 1 variables with i and level 2 with j) and error due to an individual variation within groups ε_{ij}. We therefore fit a model:

$$Y_{ij} = \alpha + \beta X_{ij} + \zeta_j + \varepsilon_{ij}$$

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$



$$Y_{ij} = \alpha + \beta X_{ij} + \zeta_j + \varepsilon_{ij}$$



$$Y_{ij} = \alpha + \beta X_{ij} + \zeta_j + \varepsilon_{ij}$$

This is called a random intercept model because we can think of it as allowing the intercept α to vary between clusters:

$$= \alpha_j + \beta X_{ij} + \varepsilon_{ij}$$

where:

$$\alpha_j = \alpha + \zeta_j$$

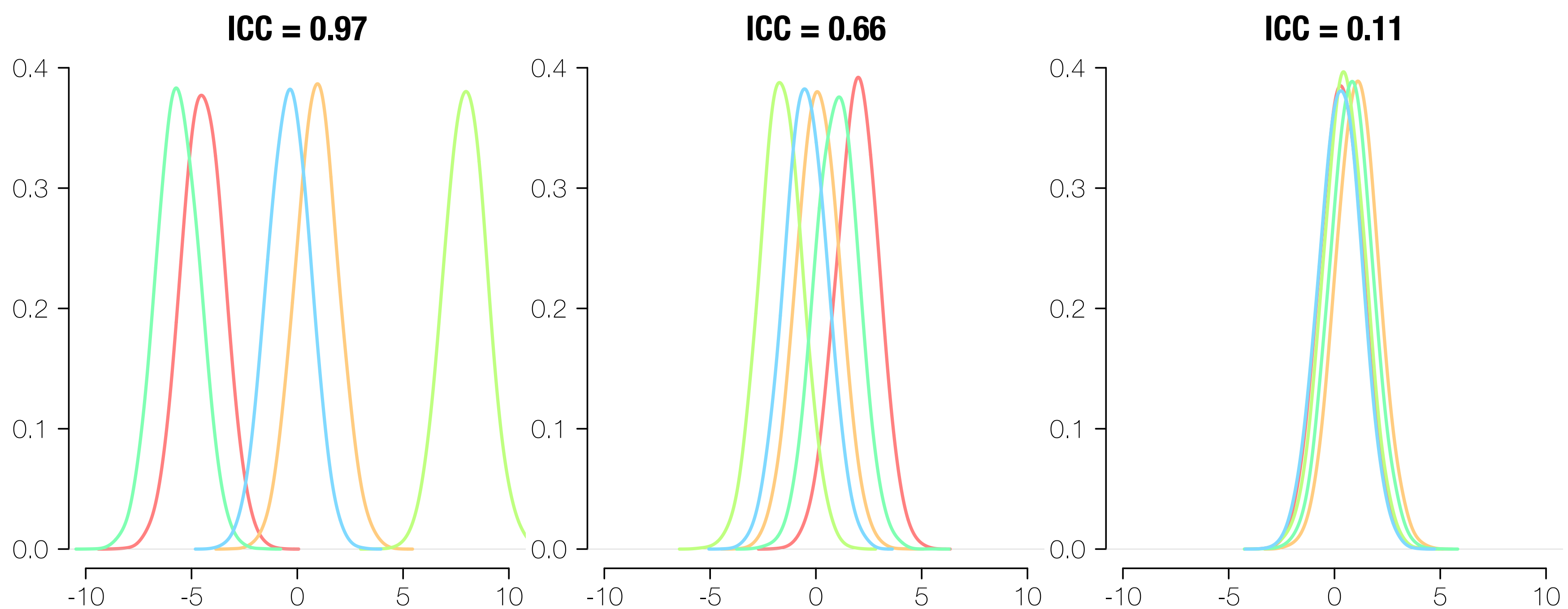
RANDOM INTERCEPT MODELS

Random intercept models allow us to:

- Estimate an intraclass correlation and test whether a significant amount of variance is occurring at the cluster level.
- Estimate relationships/effects at different levels of analysis
- Estimate cross-level interaction effects.

Intraclass correlation is an estimate of the amount of error variance that can be attributed to the clustering variable:

$$\text{ICC} = \frac{\text{Variance of } \zeta}{\text{Variance of } \zeta + \text{Variance of } \varepsilon}$$



Significance tests of random effects ask if the variance of random effects is greater than would be expected by chance.

If so, this suggests that your outcome may be affected or related to variables at the clustering/higher level.

If not, this suggests the opposite: you are likely primarily looking at a within group/level 1 phenomenon (though cross-level interactions remain possible).

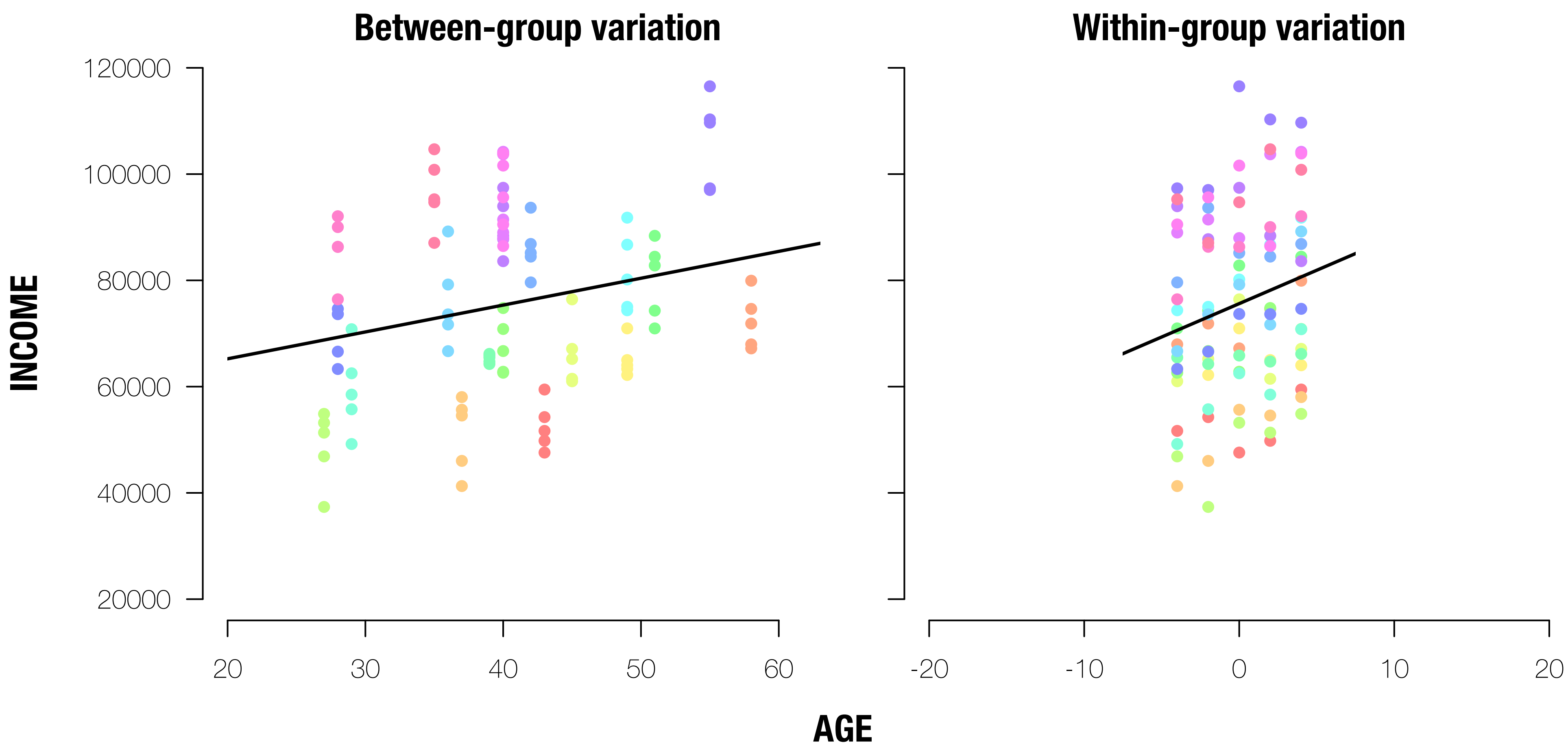
Including random intercepts allows us to examine effects at different levels. For example, X_{ij} may have both between-cluster variance and within-cluster variance (e.g., beef eating):

$\bar{X}_j$ = group-level means of X_{ij} (between)

$(X_{ij} - \bar{X}_j)$ = variation from group means (within)

$$Y_{ij} = \alpha + \beta_1(X_{ij} - \bar{X}_j) + \beta_2\bar{X}_j + \zeta_j + \varepsilon_{ij}$$

X_{ij} may have both between-cluster variance and within-cluster variance:



Other variables may have only between-cluster variance or only within-cluster variance:

X_{ij} = a level 1 predictor with no between-cluster variation

Z_j = a level 2 predictor with no within-cluster variation

$$Y_{ij} = \alpha + \beta_1 X_{ij} + \beta_2 Z_j + \zeta_j + \varepsilon_{ij}$$

Random intercept models also allow us to estimate cross-level interactions. These occur when effects of level 1 and level 2 variables depend on each other (e.g., daily variation in positive experiences (a within-person variable) has a stronger effect on individuals with higher neuroticism (a between-person variable).

X_{ij} = daily positive events
Z_j = trait-level neuroticism
Y_{ij} = daily reported well-being

$$Y_{ij} = \alpha + \beta_1 X_{ij} + \beta_2 Z_j + \beta_3 X_{ij} * Z_j + \zeta_j + \varepsilon_{ij}$$

Multiple random intercepts can be added to model data with multiple levels of clustering (e.g., observations within students within classrooms).

$$Y_{ijk} = \alpha + \beta X_{ijk} + \eta_k + \zeta_{jk} + \varepsilon_{ijk}$$

Random effects don't have to be fully nested. Cross-classified designs have multiple levels of clustering where cluster memberships can overlap (e.g. random effects for country and year of measurement).

$$Y_{ijk} = \alpha + \beta X_{ijk} + \eta_k + \zeta_j + \varepsilon_{ijk}$$

RANDOM SLOPES

Another extension of multilevel modelling includes random slopes.

This allows fitted slopes on level 1 variables (β) to vary according to cluster. We now divide our random error term ε_i into cluster-level error ζ_j, error due to group-level differences in the effect of X_{ij} γ_j, and individual variation not accounted for by ζ_j or γ_j ε_{ij}.

$$Y_{ij} = \alpha + \beta X_{ij} + \zeta_j + \gamma_j X_{ij} + \varepsilon_{ij}$$

$$Y_{ij} = \alpha + \beta X_{ij} + \zeta_j + \gamma_j X_{ij} + \varepsilon_{ij}$$

This is called a random slope model because we can think of it as allowing the intercept α and slope β to vary between clusters:

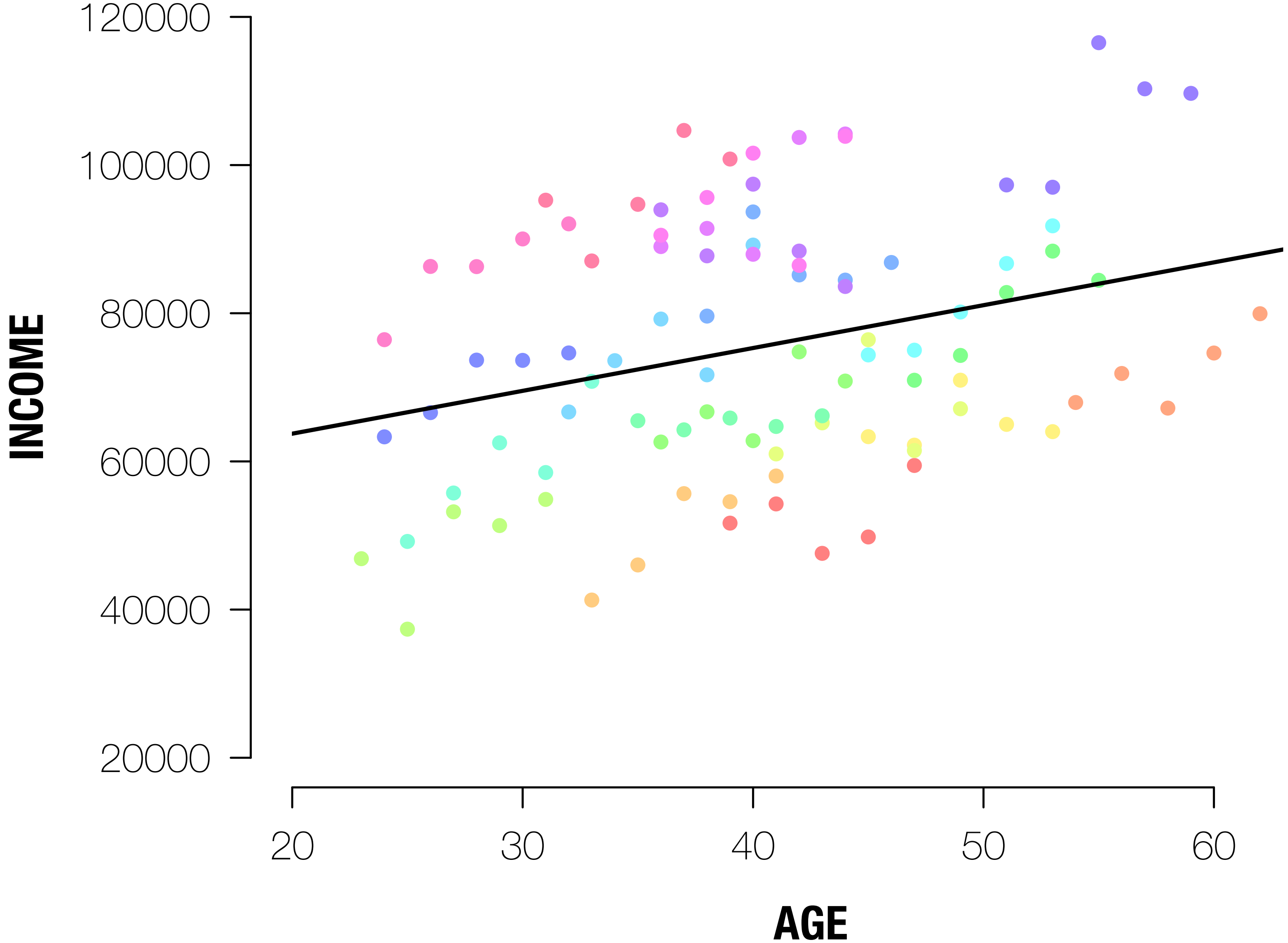
$$= \alpha_j + \beta_j X_{ij} + \varepsilon_{ij}$$

where:

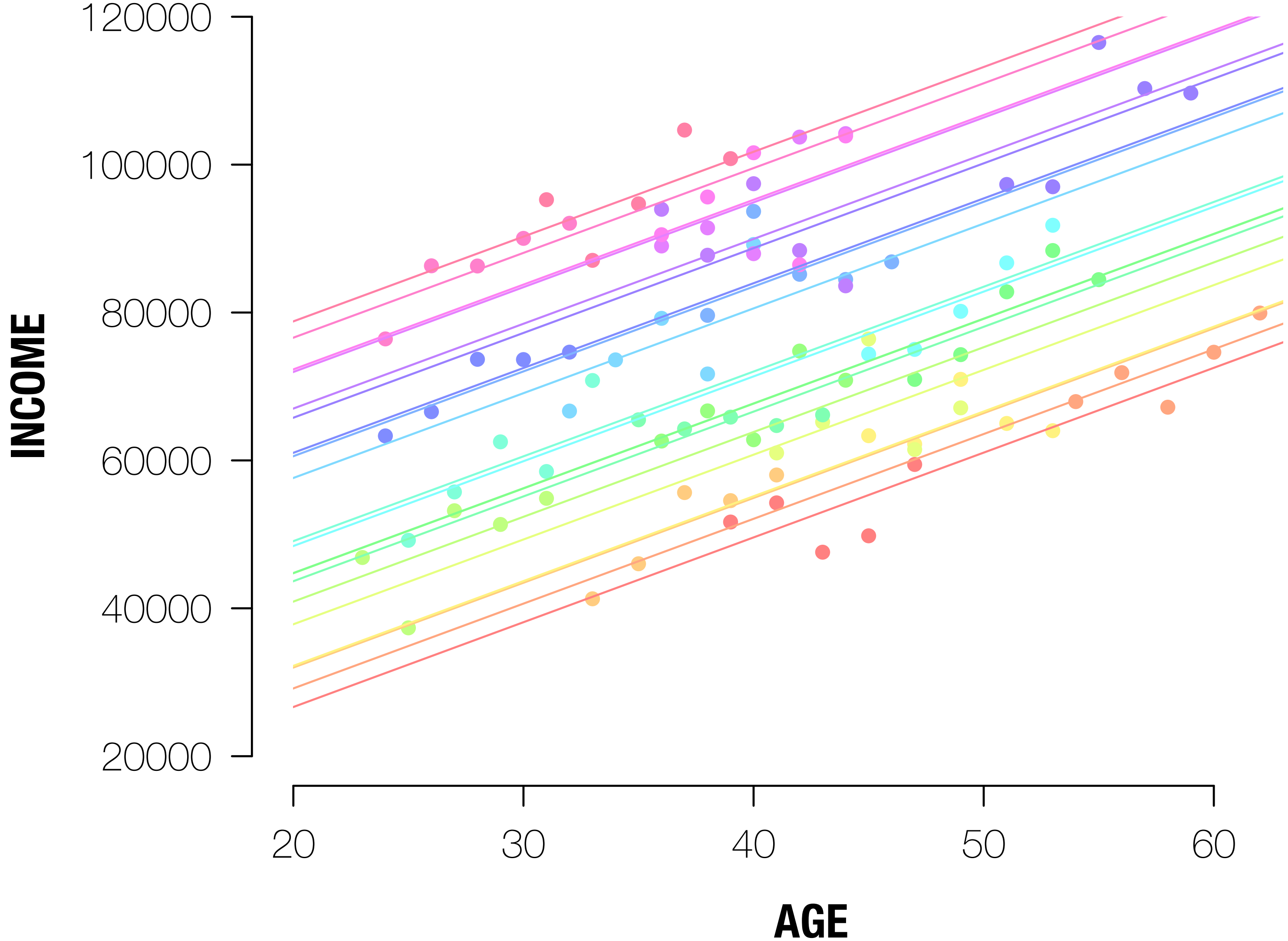
$$\alpha_j = \alpha + \zeta_j$$

$$\beta_j = \beta + \gamma_j$$

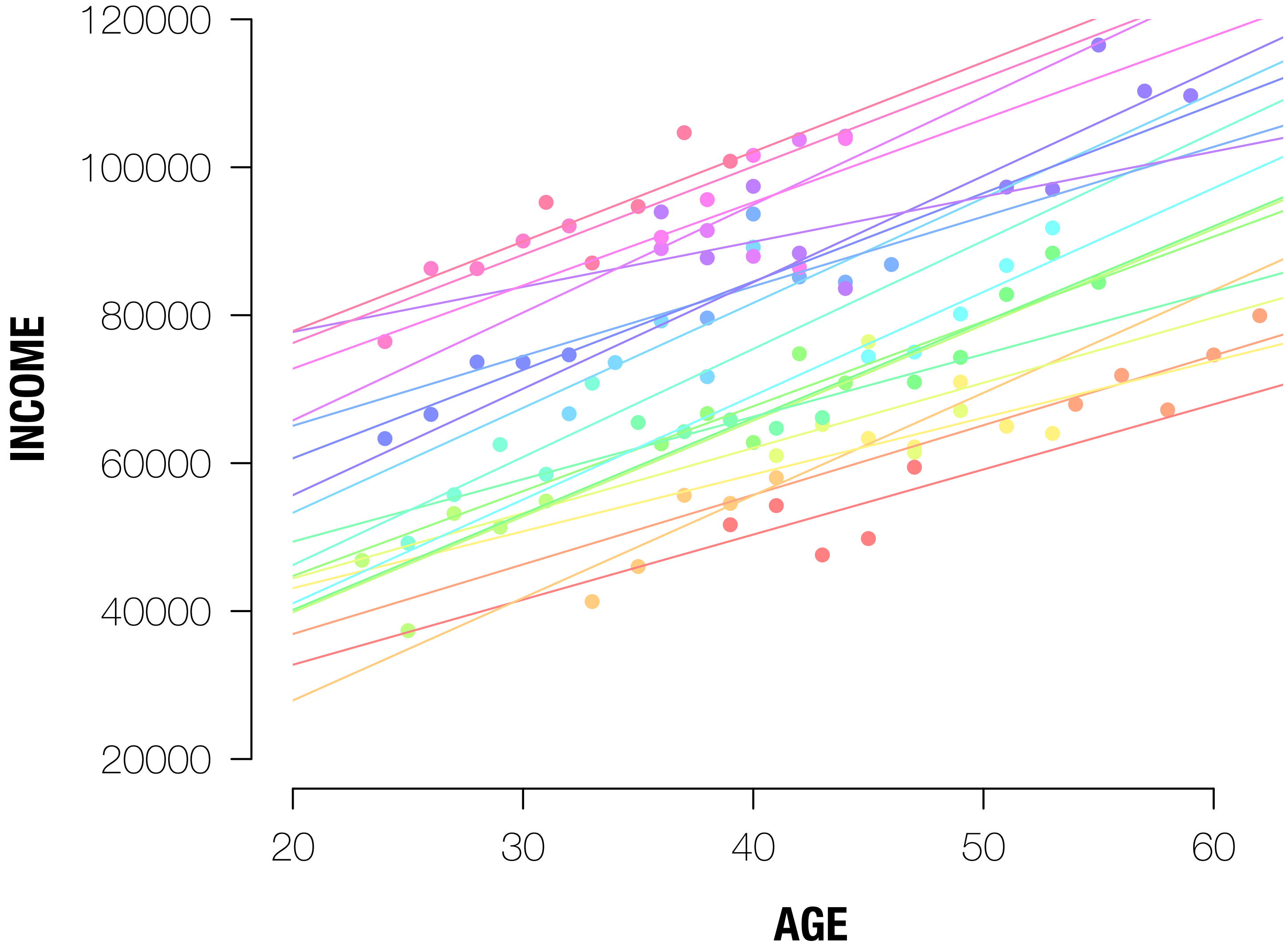
$$Y_i = \alpha + \beta X_i + \varepsilon_i$$



$$Y_{ij} = \alpha + \beta X_{ij} + \zeta_j + \varepsilon_{ij}$$



$$Y_{ij} = \alpha + \beta X_{ij} + \zeta_j + \gamma_j X_{ij} + \varepsilon_{ij}$$



RANDOM SLOPE MODELS

Random slopes don't make much sense without:

- Random intercepts: if slopes differ by group, unlikely that intercepts would not also differ
- Fixed effects: without a fixed slope, we are assuming average slope is 0.

RANDOM SLOPE MODELS

Random slopes allow us to test whether a significant amount of variance is accounted for by allowing slopes to differ between groups

A significant amount of variance in random slopes suggests that there may be a cross-level interaction occurring. I.e., there is group-level variance in slopes on some X_{ij} that a group level predictor Z_j may be able to explain.

RANDOM SLOPE MODELS

Multiple random slopes can be added to models, and can be added at multiple levels (not your highest level), but I mean, come on.

$$Y_{ijk} = \alpha + \beta_1 X_{ijk} + \beta_2 Z_{jk} + \gamma_{jk} X_{ijk} + \lambda_k Z_{jk} + \eta_j + \zeta_{jk} + \varepsilon_{ijk}$$

ASSUMPTIONS

Assumptions of multilevel models include:

- Model is correctly specified, no omitted confounds (rarely justified with observational data)
- Random effects are i.i.d, and normally distributed around mean of 0 (also rarely justified – why would New Jersey's random slope be independent from New York's?). Robust SEs can help with violations of normality and homoskedasticity.