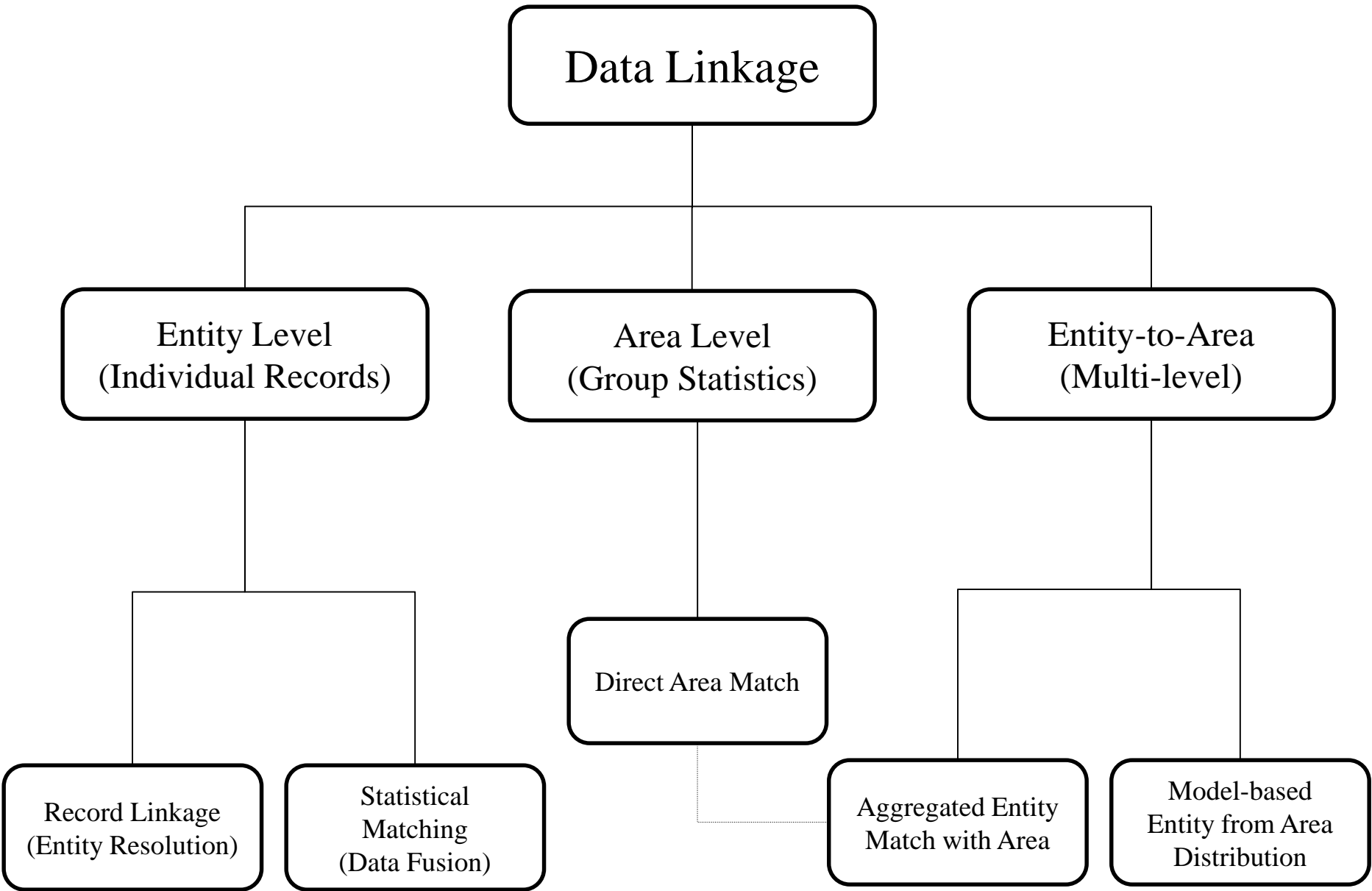


Combining Data by Statistical Matching, Imputation and Modeling

Purpose for combining data

- Improve coverage
 - Survey data from different frames (e.g. landline and cell phone)
- Increase sample size
 - Meta analysis
 - Combining probability sample with nonprobability sample (improves coverage as well)
- Bring together variables from different files
 - Neighborhood Air quality measurements



Statistical Matching

- Record's measurements are at the same level
- Little-to-no overlap of records across samples

	Y_1	Y_2	...	Y_q	X_1	X_2	...	X_p	Z_1	Z_2	...	Z_r
Sample 1	y_{111}	y_{112}	...	y_{11q}	x_{111}	x_{121}	...	x_{11p}				
	y_{121}	y_{122}	...	y_{12q}	x_{121}	x_{122}	...	x_{12p}				
	\vdots				\vdots							
	y_{1n_11}	y_{1n_12}	...	y_{1n_1q}	x_{1n_11}	x_{1n_12}	...	x_{1n_1p}				
Sample 2					x_{211}	x_{212}	...	x_{21p}	z_{211}	z_{212}	...	z_{21r}
					x_{221}	x_{222}	...	x_{22p}	z_{221}	z_{222}	...	z_{22r}
					\vdots				\vdots			
					x_{2n_21}	x_{2n_22}	...	x_{2n_2p}	z_{2n_21}	z_{2n_22}	...	z_{2n_2r}