Input Components

ML Model h
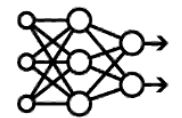
Fairness Query $F(\Omega)$

Biased Data D_Δ

Data Collection Diagram $\mathcal{G}$
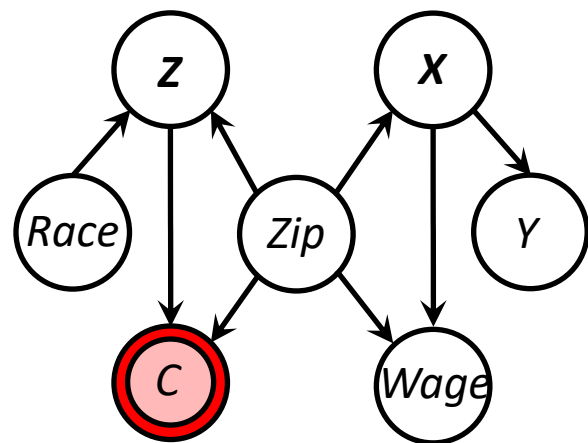
Crime Type	Age	Zip Code	Race	Y	Prediction
Assault	26	90210	White	Low Risk	High Risk
...	...	...	...		

Range of Accessibility to **External Data Sources** $\mathcal{A}_\Omega$

No Access: No external statistics available

Partial: Access to statistics about sensitive attributes and **some** variables that affect data selection

Sufficient: Access to statistics about sensitive attributes and **all** variables that affect data selection

Possible Repairs of D_Δ

Crime Type	Age	Zip Code	Race	Y	Prediction
Assault	26	90210	White	Low Risk	High Risk
Robbery	31	60610	Black	High Risk	High Risk
...	...	...	...	...	...

$\rightarrow 0.2$

Crime Type	Age	Zip Code	Race	Y	Prediction
Assault	26	90210	White	Low Risk	High Risk
Burglary	33	90210	Black	High Risk	High Risk
...	...	...	...	...	...

$\rightarrow 0.1$

$\cdots$

Crime Type	Age	Zip Code	Race	Y	Prediction
Assault	26	90210	White	Low Risk	High Risk
Theft	32	60610	Black	Low Risk	High Risk
...	...	...	...	...	...

$\rightarrow 0.5$

Possible Fairness Query Answers

Consistent Range of the fairness query answer $F(\Omega)$

$[CLB, F(\Omega), CUB]$

Consistent **L**ower **B**ound

Consistent **U**pper **B**ound