

$$d = 2r \arcsin \left(\sqrt{\sin^2 \left(\frac{\phi_2 - \phi_1}{2} \right) + \cos(\phi_1) \cos(\phi_2) \sin^2 \left(\frac{\lambda_2 - \lambda_1}{2} \right)} \right)$$

where ϕ_1 and ϕ_2 correspond to latitude 1 and latitude 2,
 λ_1 and λ_2 correspond to longitude 1 and longitude 2,
and r is the radius of the Earth.

The latitude and longitude coordinates for each house and each metro station are substituted into the haversine function for (ϕ_1, λ_1) & (ϕ_2, λ_2) respectively. The distance from a single house to all 91 metro stations are calculated, then the minimum distance is selected and added to the dataframe.