

Green Cab Taxi Data Quality Plan

Variable Names	Data Quality Issue	Handling Strategy
All variables	Duplicate rows	Drop rows
Dropoff time/Pickup time	Dropoff time before pickup time	Drop rows
Pickup time	Pickup time before 2021	Drop rows
Dropoff time	Dropoff time after 31 st Mar 2024	Drop rows
Passenger count	Missing values	Drop rows
Trip duration	Duration > 5hrs	Drop rows
Passenger Count	Count = 0	Drop rows
Passenger Count	Count > 6	Drop rows