Use of Multiple regression in prediction of schedule

The schedule can be predicted using the multiple regression where the response variable can be taken as the time difference between two stops for a given shift . This can be retrieved for a given month along with other parameters as weather , passengers (dwell time ) and traffic parameters . This will lead to the formulation of a following model .

Yi = α + β1Xi\_1 + · · · + βp Xi\_p +epsilon

Where X\_i’s are independent variables weather (will be a discrete variable-Rainy , Sunny , Cloudy , Adverse ) , traffic will be a discrete variable(Busy , Light) and passenger dwell time is a continuous variable . Here because the data is large therefore we make an assumption that errors are normally distributed . The final result which we will get is

E(Time difference between two stops in a given shift on a day/X\_i)= α + β1 *weather* + β1 *passenger\_dwell \_time*+ β3 *traffic\_conditions*

The result can be interpreted as that how does incremental increase in one of the variables affects the time difference between two stops in a given shift on a day. The above model will predict the time difference between the two stops given the value of the any of the three independent variables .