

## Supplementary Materials A: List of variables in the merged data set and their types

VariableName	Type	interested	Recode as 9999 (NOT REPORTED / UNKNOWN)	Not useful				
SPEEDREL	BINARY	1	9999	8,9				
DRUGS	BINARY	1	9999	9,8				
VPROFILE	NOMINAL	1	9999	9,8				
VSURCOND	NOMINAL	1	9999	98,99				
ROLLOVER	NOMINAL	1			9			
HAZ_CNO	NOMINAL	1			88			
CF1	NOMINAL	1			99			
CF2	NOMINAL	1			99			
CF3	NOMINAL	1			99			
P_SF1	NOMINAL	1			99			
P_SF2	NOMINAL	1			99			
P_SF3	NOMINAL	1			99			
TRAV_SP	INTERVAL	1			998,999	***		
LAND_USE	ORDINAL	1		9,8	***			
NUMOCCS	INTERVAL	1		97,99				
VSPD_LIM	INTERVAL	1		99,98,0				
AGE_IM	Interval	1						
ALCHL_IM	nominal	1						
BDYTYP_IM	NOMINAL	1						
FIRE_EXP	BINARY	1						
HAZ_INV	BINARY	1						
HOUR_IM	INTERVAL	1						
IMPACT1_IM	NOMINAL	1						
J_KNIFE	NOMINAL	1						
LGTCOIN_IM	NOMINAL	1						
MAXSEV_IM	TARGET	1						
MDLYR_IM	NOMINAL	1						
PERALCH_IM	BINARY	1						
REGION	NOMINAL	1						
SEX_IM	BINARY	1						
SAFETY_RATING	INTERVAL	1						
VE_TOTAL	INTERVAL	1						
WKDY_IM	NOMINAL	1						
WRK_ZONE	NOMINAL	1						
DRUGRES1	NOMINAL	0	9999		999,997,095			
DRUGRES2	NOMINAL	0	9999		999,997,095			
DRUGRES3	NOMINAL	0	9999		999,997,095			
VNUM_LAN	ORDINAL	0	9999	8,9		RECODE: 2,3,4,5 AS 1		
VTCONT_F	NOMINAL	0	9999	8,9				
DRUGTST1	NOMINAL	0	9999	9,6				
DRUGTST2	NOMINAL	0	9999	9,6				
DRUGTST3	NOMINAL	0	9999	9,6				
ALC_STATUS	NOMINAL	0	9999	9,8				
DEFORMED	NOMINAL	0	9999	9,8				
DSTATUS	NOMINAL	0	9999	9,8				
EMER_USE	NOMINAL	0	9999	9,8				
HOSPITAL	NOMINAL	0	9999	9,8				
VALIGN	NOMINAL	0	9999	9,8				
VTRAFWAY	NOMINAL	0	9999	9,8				
V_CONFIG	NOMINAL	0	9999	98,99				
LOCATION	NOMINAL	0	9999	99,28,98				

## Supplementary Materials A: List of variables in the merged data set and their types

REST_USE	NOMINAL	0	9999	99,97,98				
AIR_BAG	NOMINAL	0	9999	99,98				
BUS_USE	NOMINAL	0	9999	99,98				
CARGO_BT	NOMINAL	0	9999	99,98				
REL_ROAD	NOMINAL	0	9999	99,98				
SPEC_USE	NOMINAL	0	9999	99,98				
TYP_INT	NOMINAL	0	9999	99,98				
ATST_TYP	NOMINAL	0	9999	99,98,95				
HAZ_PLAC	NOMINAL	0			8			
HAZ_REL	NOMINAL	0			8			
SCH_BUS	BINARY	0			8			
DR_PRES	BINARY	0			9			
INT_HWY	BINARY	0			9			
PCRASH4	NOMINAL	0			9			
ROLINLOC	NOMINAL	0			9			
TOW_VEH	NOMINAL	0			9			
TOWED	NOMINAL	0			9			
WEATHER1	NOMINAL	0			9			
WEATHER2	NOMINAL	0			9			
MAKE	NOMINAL	0			99			
VEH_SC1	NOMINAL	0			99			
VEH_SC2	NOMINAL	0			99			
GVWR	ORDINAL	0	0,8,9					
ACC_TYPE	nominal	0	4,5,9,10,15,16,32,33,42,43,48,49,52,53,62,63,66,67,74,75,84,84,90,91,98,99					
VTRAFCON	NOMINAL	0	99,98,97					
ALC_RES	INTERVAL	0	99,98,97,96,95					
STR_VEH	ID	0	999,0					
AGE	DROP	0						
ALCOHOL	DROP	0						
BODY_TYP	DROP	0						
CASENUM	ID	0				RECODE: 997 AS 152		
DAY_WEEK	DROP	0				DISCUSS		
DR_SF1	NOMINAL	0						
DR_SF2	NOMINAL	0				TO MAKE IT ORDINAL RECODE AS MISSING?		
DR_SF3	NOMINAL	0						
DR_SF4	NOMINAL	0						
DR_ZIP	DROP	0						
DRINKING	DROP	0						
EJECT_IM	NOMINAL	0						
EJECTION	DROP	0						
EVENT1_IM	NOMINAL	0						
HARM_EV	DROP	0						
HAZ_ID	DROP	0						
HIT_RUN	DROP	0						
HITRUN_IM	BINARY	0						
HOUR	DROP	0						
ID1	ID	0						
IMPACT1	DROP	0						
INJ_SEV	DROP	0						
INJSEV_IM	DROP	0						
LGT_COND	DROP	0						
M_HARM	DROP	0						

## Supplementary Materials A: List of variables in the merged data set and their types

MAN_COLL	DROP	0					
MANCOL_IM	NOMINAL	0					
MAX_SEV	DROP	0					
MAX_VSEV	DROP	0					
MCARR_I1	ID	0					
MCARR_I2	ID	0					
MCARR_ID	ID	0					
MINUTE	DROP	0					
MINUTE_IM	INTERVAL	0					
MOD_YEAR	DROP	0					
MODEL	NOMINAL	0					
MONTH	NOMINAL	0					
MXVSEV_IM	DROP	0					
NO_INJ_IM	INTERVAL	0					
NUM_INJ	DROP	0				HAVE STATE INFO.	
NUM_INJV	DROP	0					
NUMINJ_IM	DROP	0					
P_CRASH1	DROP	0					
P_CRASH2	NOMINAL	0					
P_CRASH3	NOMINAL	0					
PCRASH1_IM	NOMINAL	0					
PCRASH5	NOMINAL	0					
PEDS	INTERVAL	0					
PER_NO	ID	0					
PER_TYP	NOMINAL	0					
PERMVIT	INTERVAL	0					
PERNOTMVIT	INTERVAL	0					
PJ	ID	0					
PSU	DROP	0					
PSUSTRAT	DROP	0					
PVH_INVL	INTERVAL	0					
RELJCT1	DROP	0					
RELJCT1_IM	BINARY	0					
RELJCT2	DROP	0					
RELJCT2_IM	NOMINAL	0					
REST_MIS	NOMINAL	0					
SEAT_IM	NOMINAL	0					
SEAT_POS	DROP	0					
SEX	DROP	0					
STRATUM	DROP	0					
UNITTYPE	DROP	0					
V_ALCH_IM	BINARY	0					
VE_FORMS	INTERVAL	0					
VEH_ALCH	DROP	0					
VEH_NO	ID	0					
VEVENT_IM	NOMINAL	0					
VIN	ID	0					
WEATHER	DROP	0					
WEATHR_IM	NOMINAL	0					
YEAR	NOMINAL	0					
WEIGHT	weight					****CASE WEIGHT	