

DATA SET: d4 LABEL: DATE CREATED: 01NOV16:12:14:37

Number of Observations: 2152

Number of Variables: 35

Organization of file: One Record per Problem List item*patient (ID)

Variable Name	Variable Label (VAR)	VAR Type	VAR Format	VAR Length	Mean	Range of Values	Frequency Category	Frequency	Percent
ANNOTATED	Annotation added	Num		8	1	1.00-1.00		2152	100.00
CERTAINTY	Certainty	Num		8	8.241171	2.00-10.00		2152	100.00
CHANGE_DUE_TO_IRR_	Change due to IRR	Char	\$	1		.-.			
CHANGE_MADE_DUE_TO_HUMAN_ERROR_	Change made due to human error?	Char	\$	1		.-.			
COMPLETE	Whether complete	Char	\$	3		.-.			
DIAGNOSIS_NAME		Char	\$DX	5		ADA	Alcohol/Drug Abuse	163	7.57
						ADHD	Attention Def/Hyperact. Disorder	46	2.14
						asthm	Asthma	359	16.68
						Breas	Breast Cancer	38	1.77
						cad	Coronary Art. Disease	151	7.02
						chf		103	4.79
						CHF	CHF	12	0.56
						ckd	Chronic Kidney Disease	158	7.34
						CKD	Chronic Kidney Disease	158	7.34
						COPD		116	5.39
						CVA	Ischemic Stroke	72	3.35
						Dep	Depression	441	20.49
						DMnoC	Diabetes wo comp	260	12.08

DATA SET: d4 LABEL: DATE CREATED: 01NOV16:12:14:37

Number of Observations: 2152

Number of Variables: 35

Organization of file: One Record per Problem List item*patient (ID)

Variable Name	Variable Label (VAR)	VAR Type	VAR Format	VAR Length	Mean	Range of Values	Frequency Category	Frequency	Percent
						DMwC	DM with comp	101	4.69
						hiv	HIV	132	6.13
						HIV	HIV	132	6.13
DTE_LASTCLAIM	Date of latest claim	Num	MMDDYY	8	01/07/12	03/20/10-03/01/13		2152	100.00
E	Encounter EHR signal	Char		1		.-.			
ENC	Enc GS review	Char	\$	8		.-.			
ENCOUNTERDATE		Num	MMDDYY	8	03/06/10	03/25/08-01/16/12		2122	98.61
ERROR	Error found in GS	Char	\$	4		.-.			
FALSE_POSITIVES	False + / decoy	Char	\$	1		.-.			
FIRSTDATE_AMBPROBLIST		Num	MMDDYY	8	04/18/08	11/11/67-01/12/12		730	33.92
FIRSTDATE_ENCDX		Num	MMDDYY	8	05/19/10	04/21/08-01/26/12		931	43.26
FIRSTDATE_HOSPADMDX		Num	MMDDYY	8	07/14/10	05/23/08-04/13/12		223	10.36
HAS	Has	Char	\$	8		.-.			
HOSPADMIT	Hosp Admit dx GS review	Char	\$	8		.-.			
HOSPDATE	deid Hospital Date	Num	DATE	8	18491.861	17741.00-19096.00		223	10.36
HOSPDIC	Hosp Discharge dx GS review	Char	\$	8		.-.			

DATA SET: d4 LABEL: DATE CREATED: 01NOV16:12:14:37

Number of Observations: 2152

Number of Variables: 35

Organization of file: One Record per Problem List item*patient (ID)

Variable Name	Variable Label (VAR)	VAR Type	VAR Format	VAR Length	Mean	Range of Values	Frequency Category	Frequency	Percent
IRR	Inter-rater reliability score	Char	\$	1		.-.			
M	MedHistory EHR signal	Char		1		.-.			
MEDHX	Med History GS review	Char	\$	8		.-.			
NOTFOUNDREASON	HAS = NO reason	Char	\$	72		.-.			
NUM_DIAGNOSES	# of times diagnosis found	Num		8	3.94847775	1.00-71.00		1708	79.37
NUM_ENCDX	# of times dx in encounters	Num		8	5.1544477	0.00-58.00		1023	47.54
NUM_HOSPDY	# of times hosp dx found	Num		8	0.45063539	0.00-13.00		1023	47.54
OTHERDX	Diagnosis found elsewhere	Char	\$	40		.-.			
PATIENT_ID	Deid patient ID	Num		8	89143.1701	60385.00-189710.00		2152	100.00
PL	Problem List EHR Signal	Char		1		.-.			
PROBLIST	Problem List GS review	Char	\$	8		.-.			
PTDX_ID	DEID Primary key	Num		8	1404.43913	2.00-2800.00		2152	100.00

DATA SET: d4 LABEL: DATE CREATED: 01NOV16:12:14:37

Number of Observations: 2152

Number of Variables: 35

Organization of file: One Record per Problem List item*patient (ID)

Variable Name	Variable Label (VAR)	VAR Type	VAR Format	VAR Length	Mean	Range of Values	Frequency Category	Frequency	Percent
R	Rule EHR signal	Char		1		.-.			
RULE	actual rule fired	Char		306		.-.			
RULECHK	Rule GS review	Char	\$	8		.-.			
TRUE_NEGATIVE		Num		8	0.95771375	0.00-1.00		2152	100.00

Obs	warning	Variable Name	VAR Format	Frequency Category
1	All Non-Missing values are the same for:	ANNOTATED		