

number of symptoms	Samples full Dataset	Samples project Dataset
1	443	17
2	6448	282
3	25451	1481
4	47283	3391
5	57989	5189
6	51116	6004
7	34037	5110
8	16425	3130
9	5980	1507
10	1534	509
11	231	82
12	8	3
Total	246945	26705

Project Diseases distribution	
Disease	samples
cystitis	1219
nose disorder	1218
vulvodynia	1218
complex regional pain syndrome	1217
spondylosis	1216
conjunctivitis due to allergy	1215
esophagitis	1215
hypoglycemia	1215
peripheral nerve disorder	1215
vaginal cyst	1215
diverticulitis	1214
gastrointestinal hemorrhage	1214
acute bronchitis	1213
fungal infection of the hair	1212
infectious gastroenteritis	1212
pneumonia	1212
spontaneous abortion	1212
sprain or strain	1212
gout	1211
arthritis of the hip	1210
marijuana abuse	1210
strep throat	1210