Overview:

The document will outline the format of the raw dataset, uses of each attribute, and steps to process the dataset.

candidates.csv Format:

	seriesuid	coordX	coordY	coordZ	class
0	1.3.6.1.4.1.14519.5.2.1.6279.6001.100225287222...	-56.08	-67.85	-311.92	0
1	1.3.6.1.4.1.14519.5.2.1.6279.6001.100225287222...	53.21	-244.41	-245.17	0
2	1.3.6.1.4.1.14519.5.2.1.6279.6001.100225287222...	103.66	-121.80	-286.62	0
3	1.3.6.1.4.1.14519.5.2.1.6279.6001.100225287222...	-33.66	-72.75	-308.41	0
4	1.3.6.1.4.1.14519.5.2.1.6279.6001.100225287222...	-32.25	-85.36	-362.51	0

1. Dataset Cancer Identification determined by Class in candidates file
    a. Will be used to determine if parental class is cancerous

2. coordN determines coordinate on X/Y/Z plane
    a. Will be used to create an image from voxel coords

3. seriesuid identifies each group of pixels
    a. Will be used to identify pixels and parental class

Abstract Design Philosophy:

1. Create images through coordX and coordY pixel locations of same series
2. Label image as cancer class
3. Run image segmentation to identify seperate entities in image
4. Vectorize image for model input