cancer_test_results

November 29, 2017

1 Cancer Test Results

```
In [1]: # load dataset
        import numpy as np
        import pandas as pd

In [2]: df = pd.read_csv('cancer_test_data.csv')
        df.head()
```

Out[2]:

	patient_id	test_result	has_cancer
0	79452	Negative	False
1	81667	Positive	True
2	76297	Negative	False
3	36593	Negative	False
4	53717	Negative	False

```
In [9]: # number of patients
        num_patients = df.shape[0]
        num_patients
```

Out[9]: 2914

```
In [13]: # number of patients with cancer
         df_pos_cancer = df[df['has_cancer'] == True]
         cancer_pos_count = df_pos_cancer.shape[0]
         cancer_pos_count
```

Out[13]: 306

```
In [14]: # number of patients without cancer
         df_neg_cancer = df[df['has_cancer'] == False]
         cancer_neg_count = df_neg_cancer.shape[0]
         cancer_neg_count
```

Out[14]: 2608

```
In [10]: # proportion of patients with cancer
         cancer_pos_count / num_patients
```

```
Out[10]: 0.10501029512697323

In [11]: # proportion of patients without cancer
         cancer_neg_count / num_patients

Out[11]: 0.8949897048730268

In [18]: # proportion of patients with cancer who test positive
         df_pos_cancer_pos_test = df_pos_cancer[df_pos_cancer['test_result'] == 'Positive']
         df_pos_cancer_pos_test.shape[0]/cancer_pos_count

Out[18]: 0.9052287581699346

In [19]: # proportion of patients with cancer who test negative
         df_pos_cancer_neg_test = df_pos_cancer[df_pos_cancer['test_result'] == 'Negative']
         df_pos_cancer_neg_test.shape[0]/cancer_pos_count

Out[19]: 0.09477124183006536

In [20]: # proportion of patients without cancer who test positive
         df_neg_cancer_pos_test = df_neg_cancer[df_neg_cancer['test_result'] == 'Positive']
         df_neg_cancer_pos_test.shape[0]/cancer_neg_count

Out[20]: 0.2036042944785276

In [21]: # proportion of patients without cancer who test negative
         df_neg_cancer_neg_test = df_neg_cancer[df_neg_cancer['test_result'] == 'Negative']
         df_neg_cancer_neg_test.shape[0]/cancer_neg_count

Out[21]: 0.7963957055214724
```