

- test_statistics.py result

```
shaanjot@shaanjot-Aspire-E5-575:~/Desktop/homework1$ nosetests tests/test_statistics.py --nologcapture
...
-----
Ran 3 tests in 0.074s

OK
shaanjot@shaanjot-Aspire-E5-575:~/Desktop/homework1$
```

- Event Statistics

```
shaanjot@shaanjot-Aspire-E5-575:~/Desktop/homework1/src$ python event_statistics.py
Time to compute event count metrics: 0.09053277969360352s
(1, 8635, 982.014, 1, 12627, 498.118)
Time to compute encounter count metrics: 1.5098578929901123s
(1, 203, 23.038, 1, 391, 15.452)
Time to compute record length metrics: 0.2446153163909912s
(0, 1972, 127.532, 0, 2914, 159.2)
shaanjot@shaanjot-Aspire-E5-575:~/Desktop/homework1/src$
```

Metric	Deceased Patient	Alive Patient	Function to complete
Event Count			
1. Average Event Count	982.014	498.118	
2. Max Event Count	8635	12627	
3. Min Event Count	1	1	
Encounter Count			
1. Average Encounter Count	23.038	15.452	
2. Max Encounter Count	203	391	
3. Min Encounter Count	1	1	
Record Length			
1. Average Record Length	127.532	159.2	
2. Max Record Length	1972	2914	
3. Min Record Length	0	0	

- Model Performance on training data

Model	Accuracy	AUC	Precision	Recall	F-Score
Logistic Regression	0.9545	0.9454	0.9869	0.8988	0.9408
SVM	0.9940	0.9945	0.9882	0.9970	0.99259

Decision Tree	0.7763	0.7475	0.7921	0.6011	0.6835
---------------	--------	--------	--------	--------	--------

- Model performace on test data

Model	Accuracy	AUC	Precision	Recall	F-Score
Logistic Regression	0.7380	0.7375	0.6804	0.7333	0.7058
SVM	0.7380	0.7388	0.6767	0.7444	0.70899
Decision Tree	0.6714	0.6569	0.6329	0.5555	0.5917

Possible strategies to improve test performance will be to towards a more suitable model family so that may be there is no overfitting

- Cross-Validation

CV Strategy	Accuracy	AUC
KFold	0.7213	0.7075
Randomized	0.7678	0.7343

- For my model, I decided to implement using RandomForestClassifier after the traning data is preprocessed using MinMaxScalar, in order to improve test performance I implemented a GridSearch validation to fit data.