Data Intake Report Insurance

Name: Insurance

Report date: 25 January 2022 Internship Batch: LISUM05

Version: 1.0

Data intake by: Selaelo Kholofelo Ramokgopa

Data intake reviewer:

Data storage location: computer

Tabular data details:

Total number of observations	1338
Total number of files	1
Total number of features	7
Base format of the file	.csv
Size of the data	73.3 KB

Proposed Approach: Data pre-processing

- Information on the dataset was obtained
- Dataset was checked for null values and duplicates
- Column headers were viewed
- Pair wise relationships in the data were plotted
- Flask deployment

0

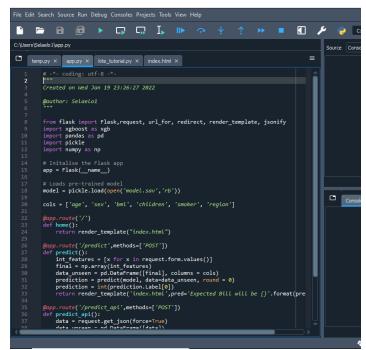
Xgboost model was selected and fit to data

```
1250 18648.42170
1257 11305.93455
139 2166.73200
919 5245.22690
....
298 38746.35510
954 20009.63365
950 11534.87265
399 1631.66830
1217 4058.11610
Name: charges, Length: 402, dtype: float64

In [21]: model.fit(x, y)
[05:52:46] WARNING: ..\src\learner.cc:1115: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'm ulti:softprob' was changed from 'merror' to 'mlogloss'. Explicitly set eval_metric if you'd like to restore the old behavior.

Out[21]: XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1, colsample_bylevel=1, colsample_bylevel=0, colsample_bylevel=1, colsample_bylevel=0, and colsample=0, and colsample=0,
```

App.py file was created



Html file was created

```
### Spyder (Python 3.7)

File Edit Search Source Run Debug Consoles Projects Tools View Help

**C:\Users\Selecto\1\spyder.py3\putosave\ndex.html

| C:\Users\Selecto\1\spyder.py3\putosave\ndex.html
| C:\Users\Se
```

0

0

0



Requirement.txt was created

0

0

0

0

```
Requirement.txt was creater than the property of the property
```

Model was deployed in flask

