```
In [ ]:  1
```

executed in 11.6s, finished 16:21:01 2020-10-07

```
In [1]:  1  from GlucoCheck.glucoCheck import glucoCheckOps
         2  import pandas as pd
         3  import random
         4  import numpy as np
         5  from tqdm.auto import tqdm
         6
         7  from scipy import stats
         8
         9  import random
        10  import re
        11  from dateutil.parser import parse
        12
        13  import warnings
        14  warnings.filterwarnings('ignore')
        15
        16  import os
        17
```

executed in 12.8s, finished 12:22:58 2020-10-28

```
Using TensorFlow backend.
```

```
In [2]:  1 ▼  def createGap(df,start,end):
         2        """
         3        Creating a Gap
         4 ▼      input:
         5            start: seed
         6            end: seed + gap
         7 ▼      output:
         8            df: dataframe with index => DisplayTime value => GlucoseValue
         9        """
        10
        11        #df = readData()
        12        l = len(df.index)
        13 ▼      if end>l:
        14            end = l
        15
        16 ▼      for i in range(start,end):
        17            df['GlucoseValue'][i]=float("NaN")
        18
        19        return df
```

executed in 43ms, finished 12:22:58 2020-10-28

In [3]:
```python
#Extract Data
data = pd.read_csv("~/Desktop/NCSA_genomics/Python - notebooks/Data/O
data = data[data['subjectId']=='OD552']
data = data.reset_index(drop=True)
data
```
executed in 524ms, finished 12:22:58 2020-10-28

Out[3]:

	subjectId	Display Time	GlucoseValue
0	OD552	4/16/25 11:17	95
1	OD552	4/16/25 11:22	86
2	OD552	4/16/25 11:27	81
3	OD552	4/16/25 11:32	81
4	OD552	4/16/25 11:37	82
...	...	...	...
11439	OD552	6/7/25 16:49	238
11440	OD552	6/7/25 16:54	233
11441	OD552	6/7/25 16:59	229
11442	OD552	6/7/25 17:04	224
11443	OD552	6/7/25 17:09	215

11444 rows × 3 columns

In [ ]:
```python

```
executed in 50ms, finished 15:27:50 2020-10-16

In [67]:
```python
#1 week after : 1890, 1974, 2003, 2196, 2378, 2581, 2751, 3190, 3223,
#2 weeks after: 3600, 3797, 3828, 3939, 4210, 4353, 4567, 4890, 5102,
#3 weeks after: 5500, 5681, 5727, 5893, 5919, 6060, 6143, 6250, 6492,
#4 weeks after: 7000, 7296, 7384, 7557, 7572, 7698, 7839, 7934, 8190,

seed_points = [7000, 7296, 7384, 7557, 7572, 7698, 7839, 7934, 8190,

#
```
executed in 4ms, finished 13:56:37 2020-10-28

In [68]:
```python
# obj = glucoCheckOps()
```
executed in 7ms, finished 13:56:37 2020-10-28

In [ ]:
```python

```
executed in 5ms, finished 00:11:30 2020-10-23

In [69]:

```python
#for gap size 50
ioa_gap50 = list()


for seed in tqdm(seed_points):
    start = seed
    end = seed+49

    dataWithMissing = data.copy()
    dataWithMissing = createGap(dataWithMissing,start,end)

    dataBeforeGap = dataWithMissing[:seed]

    obj = glucoCheckOps()
#     obj.train(dataBeforeGap);
    imputed_data = obj.impute(dataWithMissing)

    ioa = obj.index_agreement(np.asarray(imputed_data['GlucoseValue']

    del obj

    ioa_gap50.append(ioa)

ioa_gap50
```

executed in 17m 49s, finished 14:14:27 2020-10-28

100%　　　　　　　　　　　　　　　　　　　10/10 [26:23<00:00, 158.36s/it]

```
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
```

```
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
```

Out[69]: [0.41809306223243115,
 0.6161928056908199,
 0.6355850387056549,
 0.5002538589001625,
 0.8056102951583577,
 0.7706399802062551,
 0.20408098205041691,
 0.38504288993596725,
 0.40434005193970435,
 0.44914782931882713]

In [70]:
```python
#for gap size 30
ioa_gap30 = list()

for seed in tqdm(seed_points):
    start = seed
    end = start+29

    dataWithMissing = data.copy()
    dataWithMissing = createGap(dataWithMissing,start,end)

    dataBeforeGap = dataWithMissing[:seed]

    obj = glucoCheckOps()
#   obj.train(dataBeforeGap);
    imputed_data = obj.impute(dataWithMissing)

    ioa = obj.index_agreement(np.asarray(imputed_data['GlucoseValue'][start:

    del obj

    ioa_gap30.append(ioa)

ioa_gap30
```

executed in 17m 54s, finished 14:32:21 2020-10-28

100%   10/10 [17:54<00:00, 107.41s/it]

```
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
```

```
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
Gap < 50; We use LSTM imputations
Training Model...


Model trained successfully!
```

Out[70]: [0.8072498260484534,
 0.5949163312857446,
 0.10690645622339023,
 0.051481973831164485,
 0.1481376490739157,
 0.007468192866499468,
 0.4510406771910811,
 0.6292705344332856,
 0.1999853165986084,
 0.7923369152243919]

In [71]:
```python
1  gap_size 12
2  gap15 = list()
3
4
5  for seed in tqdm(seed_points):
6      start = seed
7      end = start+12
8
9      dataWithMissing = data.copy()
10     dataWithMissing = createGap(dataWithMissing,start,end)
11
12     dataBeforeGap = dataWithMissing[:seed]
13
14     obj = glucoCheckOps()
15     obj.train(dataBeforeGap);
16     imputed_data = obj.impute(dataWithMissing)
17
18     ioa = obj.index_agreement(np.asarray(imputed_data['GlucoseValue'][start:end-
19
20     del obj
21
22     ioa_gap15.append(ioa)
23
24  gap15
```

executed in 37.6s, finished 14:32:59 2020-10-28

100%   10/10 [00:37<00:00, 3.75s/it]

```
Gap < 15; We use the spline imputations
Gap < 15; We use the spline imputations
Gap < 15; We use the spline imputations
Gap < 15; We use the spline imputations
Gap < 15; We use the spline imputations
Gap < 15; We use the spline imputations
Gap < 15; We use the spline imputations
Gap < 15; We use the spline imputations
Gap < 15; We use the spline imputations
Gap < 15; We use the spline imputations
```

Out[71]:
```
[0.40122356741668785,
 0.14898086043236058,
 0.1631342624361869,
 0.2760294232888253,
 0.7066065243790824,
 0.25118516375890454,
 0.1371920355703775,
 0.10330506036017661,
 0.9980211170698682,
 0.6469001586477505]
```

```
In [72]:  1  size = 100
          2  ioa = list()
          3
          4  for seed in tqdm(seed_points):
          5      start = seed
          6      end = seed+99
          7
          8      dataWithMissing = data.copy()
          9      dataWithMissing = createGap(dataWithMissing,start,end)
         10
         11      dataBeforeGap = dataWithMissing[:seed]
         12
         13      obj = GlucoCheckOps()
         14      obj.train(dataBeforeGap);
         15      imputed_data = obj.impute(dataWithMissing)
         16
         17      if isinstance(imputed_data, pd.DataFrame):
         18          ioa = obj.index_agreement(np.asarray(imputed_data['GlucoseValue'][start:end-
         19          ioa_gap100.append(ioa)
         20      ▼
         21      ioa_gap100.append(0)
         22
         23  obj
         24
         25
         26
         27  ioa
```

executed in 406ms, finished 14:32:59 2020-10-28

100%   10/10 [00:00<00:00, 19.16it/s]

```
We cannot impute this data
We cannot impute this data
We cannot impute this data
We cannot impute this data
We cannot impute this data
We cannot impute this data
We cannot impute this data
We cannot impute this data
We cannot impute this data
We cannot impute this data
```

Out[72]: [0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

In [73]:

```python
#for gap size 5
ioa_gap5 = list()
# fb_gap5 = list()
# mad_gap5 = list()
# rmse_gap5 = list()
# mape_gap5 = list()

for seed in tqdm(seed_points):
    start = seed
    end = start+4

    dataWithMissing = data.copy()
    dataWithMissing = createGap(dataWithMissing,start,end)

    dataBeforeGap = dataWithMissing[:seed]

    obj = glucoCheckOps()
#     obj.train(dataBeforeGap);
    imputed_data = obj.impute(dataWithMissing)

    ioa = obj.index_agreement(np.asarray(imputed_data['GlucoseValue']

    del obj

    ioa_gap5.append(ioa)



ioa_gap5
```

executed in 447ms, finished 14:32:59 2020-10-28

100%   10/10 [00:00<00:00, 24.40it/s]

```
Gap < 5; We use the linear imputations
Gap < 5; We use the linear imputations
Gap < 5; We use the linear imputations
Gap < 5; We use the linear imputations
Gap < 5; We use the linear imputations
Gap < 5; We use the linear imputations
Gap < 5; We use the linear imputations
Gap < 5; We use the linear imputations
Gap < 5; We use the linear imputations
Gap < 5; We use the linear imputations
```

Out[73]:
```
[0.7096774193548361,
 0.4415845091209638,
 0.2720751916893387,
 0.9540034071550259,
```

```
        0.8945686900958462,
        0.7318348144275977,
        0.41943638236120384,
        0.8985228563936718,
        0.7749703238935044,
        0.923076923076923]
```

In [ ]:
```
1
```

In [74]:
```
OA1= pd.DataFrame({'Seeds':seed_points, 'Gap:5':ioa_gap5, 'Gap:12':ioa_gap15
OA2
3
```

executed in 104ms, finished 14:33:00 2020-10-28

Out[74]:

	Seeds	Gap:5	Gap:12	Gap:30	Gap:50	Gap:100
0	7000	0.709677	0.401224	0.807250	0.418093	0
1	7296	0.441585	0.148981	0.594916	0.616193	0
2	7384	0.272075	0.163134	0.106906	0.635585	0
3	7557	0.954003	0.276029	0.051482	0.500254	0
4	7572	0.894569	0.706607	0.148138	0.805610	0
5	7698	0.731835	0.251185	0.007468	0.770640	0
6	7839	0.419436	0.137192	0.451041	0.204081	0
7	7934	0.898523	0.103305	0.629271	0.385043	0
8	8190	0.774970	0.998021	0.199985	0.404340	0
9	8250	0.923077	0.646900	0.792337	0.449148	0

In [ ]:
```
1
```

In [75]:
```
1    IOA.to_csv("~/Desktop/4week.csv")
```

executed in 12ms, finished 14:33:00 2020-10-28

In [ ]:
```
1
```

In [ ]:
```
1
```

In [76]:
```
1    import matplotlib.pyplot as plt
2    import matplotlib.ticker as ticker
3    import seaborn as sns
```

executed in 15ms, finished 14:33:00 2020-10-28

In [77]:
```
1 ▾  # IOA
```

executed in 9ms, finished 14:33:00 2020-10-28

In [78]:
```python
gaps = [5,15,30,50,100]
ioa = []
ioa.append(IOA['Gap:5'].mean())
ioa.append(IOA['Gap:12'].mean())
ioa.append(IOA['Gap:30'].mean())
ioa.append(IOA['Gap:50'].mean())
ioa.append(IOA['Gap:100'].mean())
ioa
```

executed in 25ms, finished 14:33:00 2020-10-28

Out[78]:
```
[0.7019750517568913,
 0.38325781733602204,
 0.3788793872776535,
 0.5188986794138597,
 0.0]
```

In [79]:
```python
plt.figure(figsize=(14,8))
plt.title("Avg. IOA vs Gap size")
sns.set(style="white")
fig = sns.lineplot(x = gaps, y = ioa, palette="tab10", linewidth=1.25

fig.set_xlabel('Gaps')
fig.set_ylabel('IOA value')
```

executed in 952ms, finished 14:33:01 2020-10-28
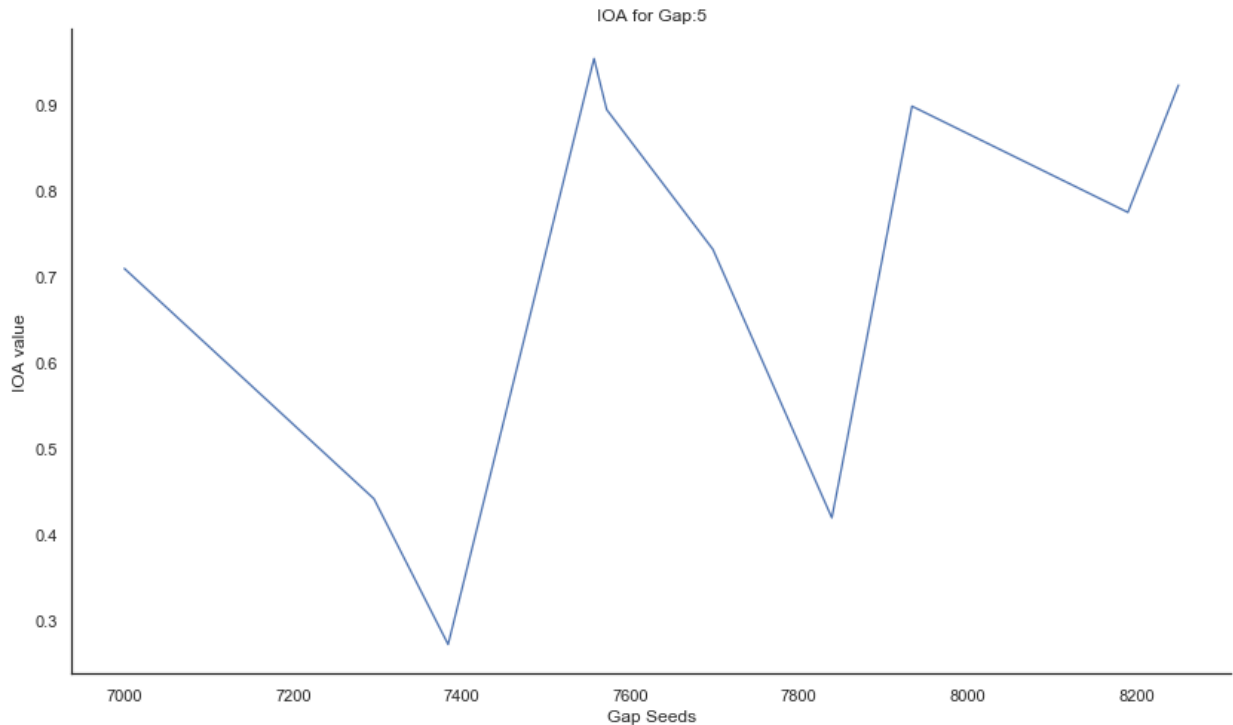
Out[79]: Text(0, 0.5, 'IOA value')
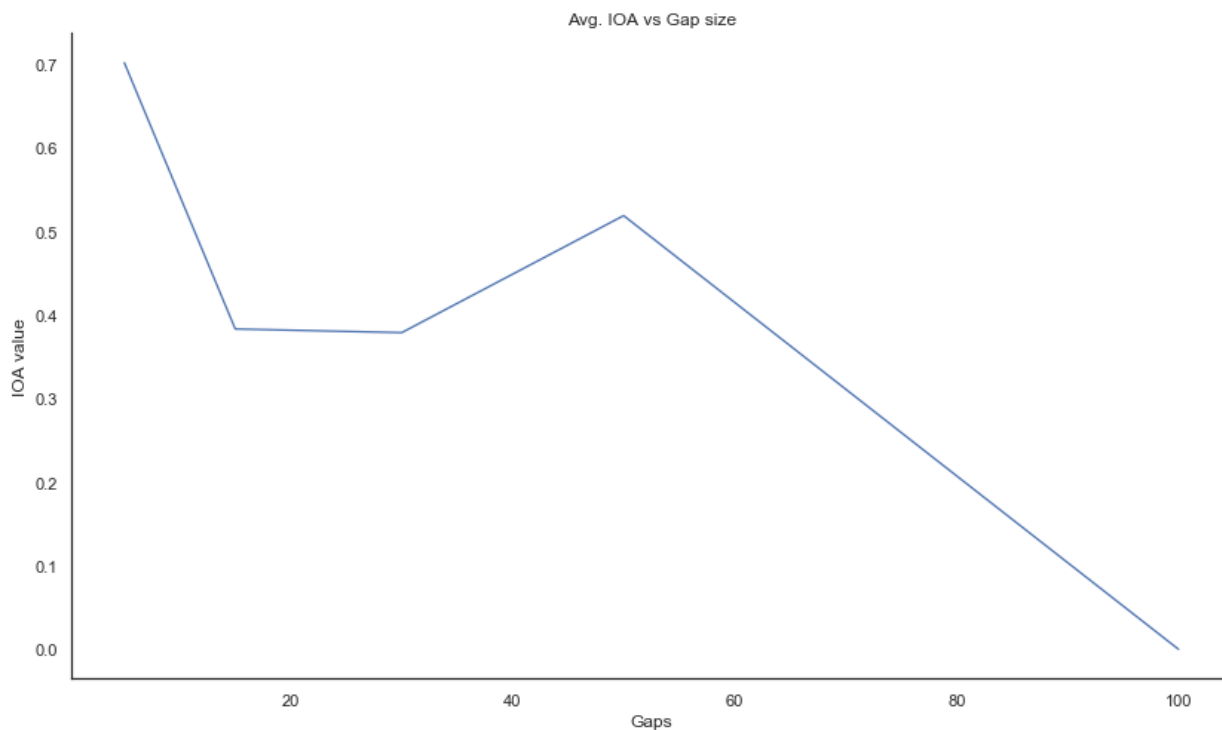
In [ ]:
```
1
```

In [ ]:
```
1
```

executed in 8ms, finished 14:05:45 2020-10-14

In [80]:
```python
plt.figure(figsize=(14,8))
plt.title("IOA for Gap:5")
sns.set(style="white")
fig = sns.lineplot(x = seed_points, y = IOA['Gap:5'], data = IOA, pal
sns.despine()

fig.set_xlabel('Gap Seeds')
fig.set_ylabel('IOA value')
```
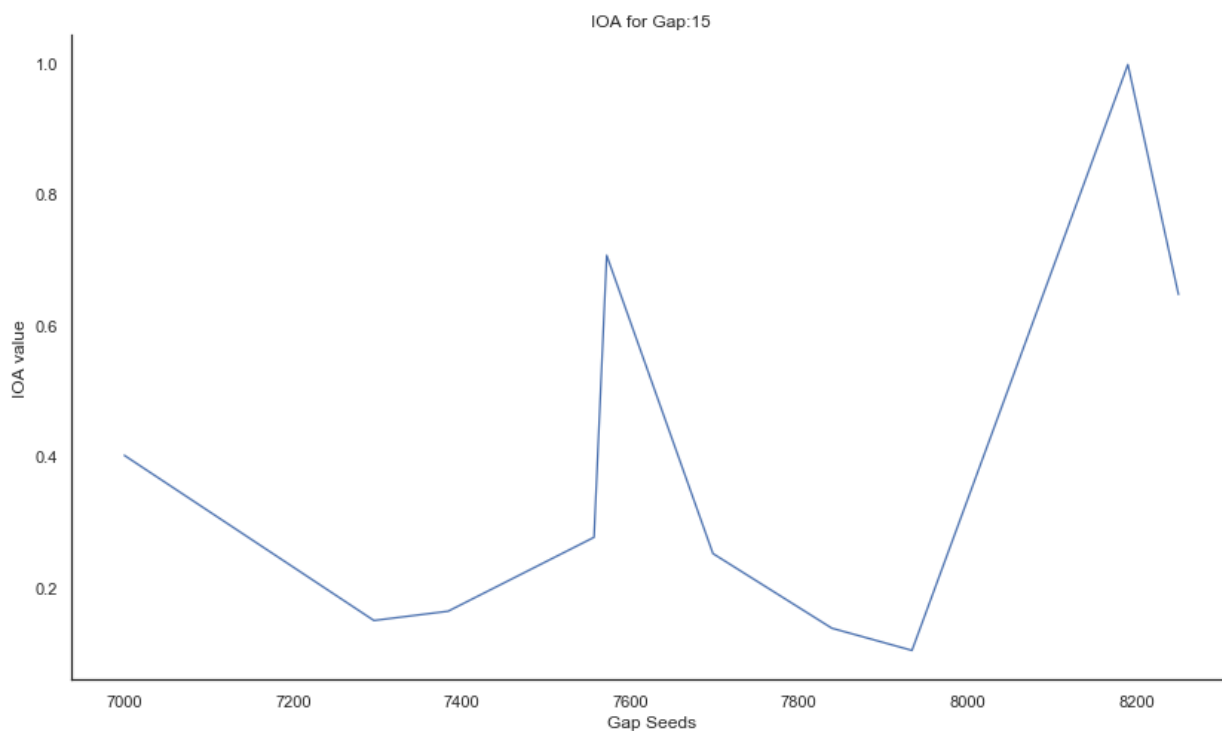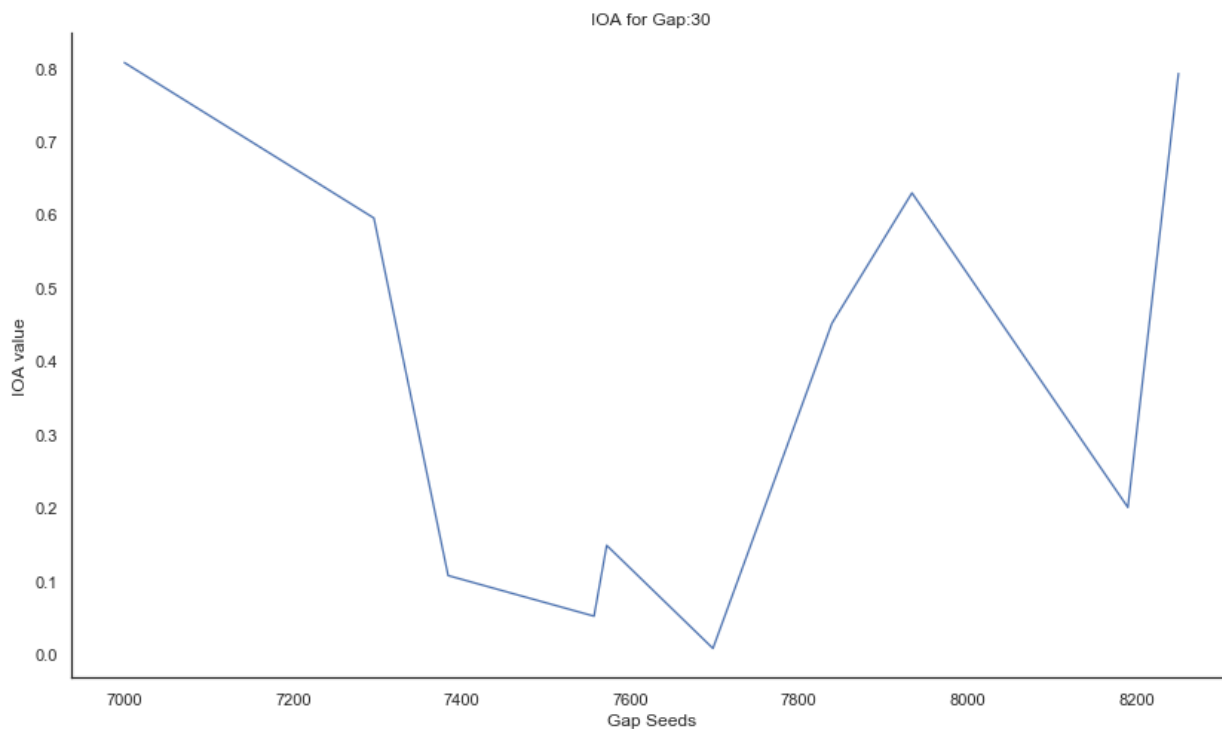executed in 553ms, finished 14:33:01 2020-10-28

Out[80]: Text(0, 0.5, 'IOA value')

In [81]:

```python
plt.figure(figsize=(14,8))
plt.title("IOA for Gap:15")
sns.set(style="white")
fig = sns.lineplot(x = seed_points, y = IOA['Gap:12'], data = IOA, pa
sns.despine()

fig.set_xlabel('Gap Seeds')
fig.set_ylabel('IOA value')
```

executed in 543ms, finished 14:33:02 2020-10-28

Out[81]: Text(0, 0.5, 'IOA value')

In [82]:

```
1
2   plt.figure(figsize=(14,8))
3   plt.title("IOA for Gap:30")
4   sns.set(style="white")
5   fig = sns.lineplot(x = seed_points, y = IOA['Gap:30'], data = IOA, pa
6   sns.despine()
7
8   fig.set_xlabel('Gap Seeds')
9   fig.set_ylabel('IOA value')
```
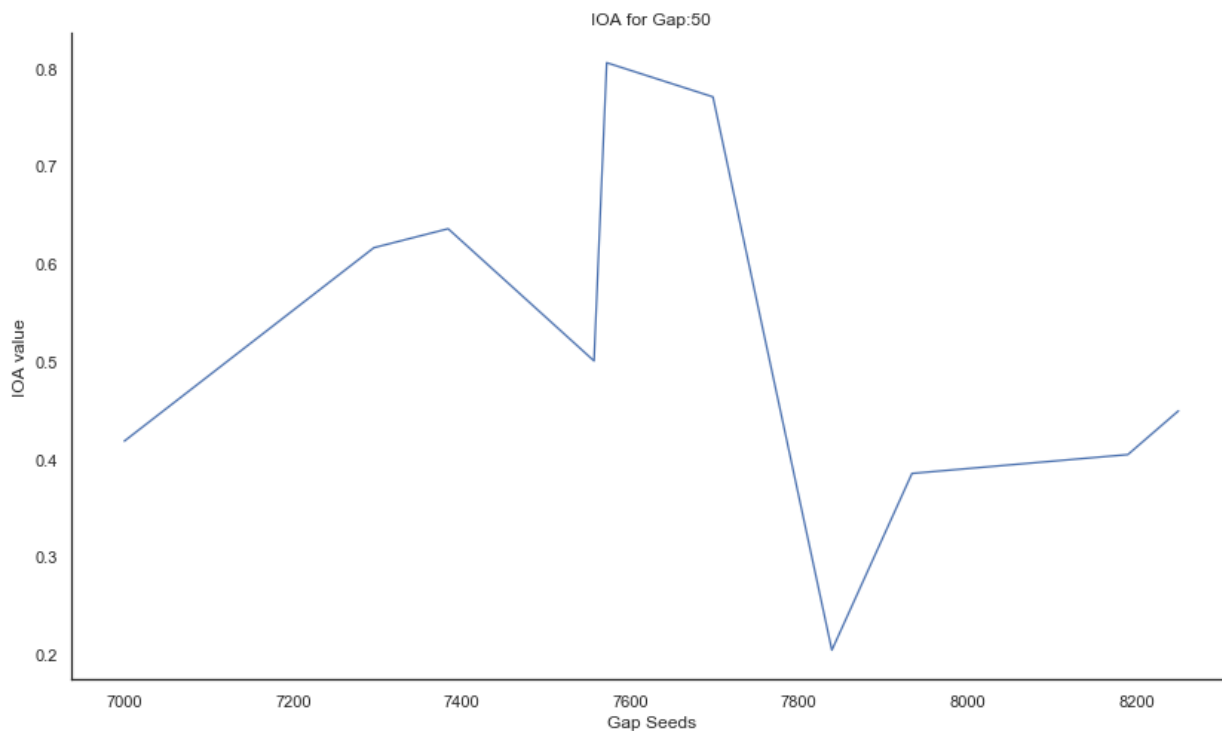
executed in 557ms, finished 14:33:03 2020-10-28

Out[82]: Text(0, 0.5, 'IOA value')

In [83]:
```python
plt.figure(figsize=(14,8))
plt.title("IOA for Gap:50")
sns.set(style="white")
fig = sns.lineplot(x = seed_points, y = IOA['Gap:50'], data = IOA, pa
sns.despine()

fig.set_xlabel('Gap Seeds')
fig.set_ylabel('IOA value')
```
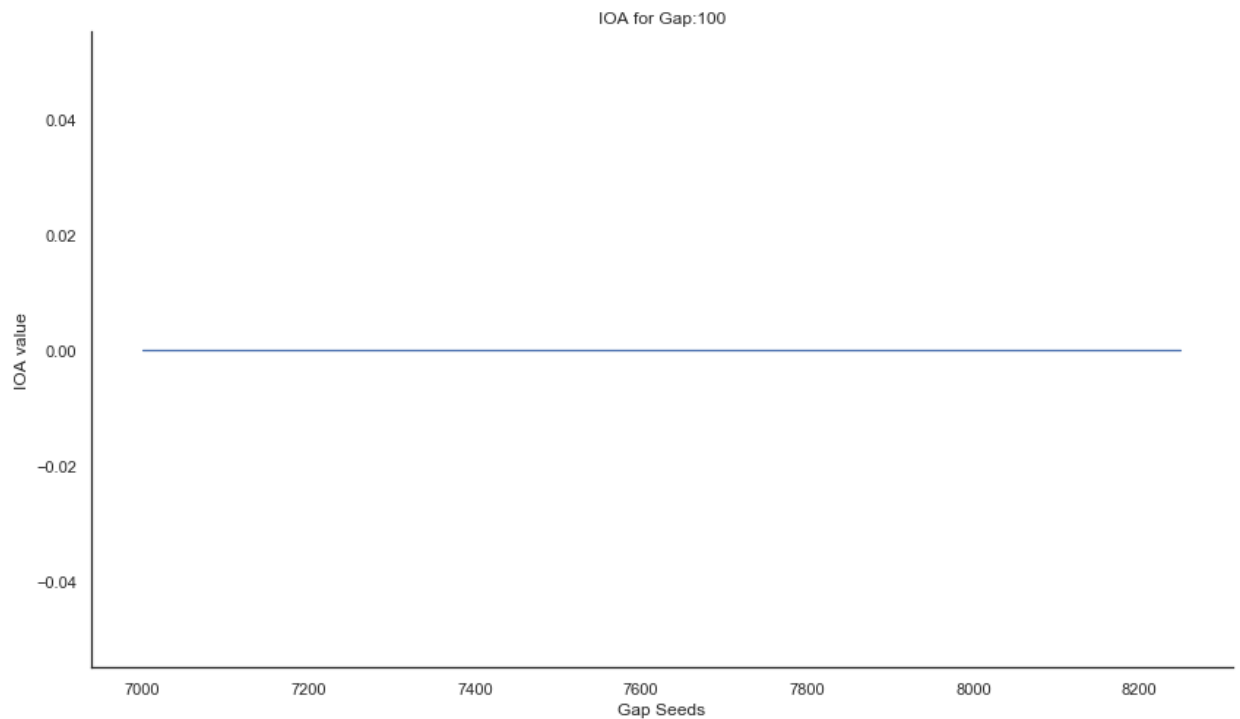executed in 539ms, finished 14:33:03 2020-10-28

Out[83]: Text(0, 0.5, 'IOA value')

In [84]:
```python
plt.figure(figsize=(14,8))
plt.title("IOA for Gap:100")
sns.set(style="white")
fig = sns.lineplot(x = seed_points, y = IOA['Gap:100'], data = IOA, p
sns.despine()
fig.set_xlabel('Gap Seeds')
fig.set_ylabel('IOA value')
```
executed in 528ms, finished 14:33:04 2020-10-28

Out[84]: Text(0, 0.5, 'IOA value')


```
###
```

In [ ]:

In [22]:
```
1 ▾ # MAD = pd.DataFrame({'Gap:5':mad_gap5, 'Gap:15':mad_gap15, 'Gap:30':m
2   # MAD
```
executed in 7ms, finished 12:36:57 2020-10-28

In [23]:
```
1 ▾ # FB = pd.DataFrame({'Gap:5':fb_gap5, 'Gap:15':fb_gap15, 'Gap:30':fb_
2   # FB
```
executed in 7ms, finished 12:36:57 2020-10-28

In [24]:
```
1 ▾ # RMSE = pd.DataFrame({'Gap:5':rmse_gap5, 'Gap:15':rmse_gap15, 'Gap:3
2   # RMSE
```
executed in 8ms, finished 12:36:57 2020-10-28

In [25]:
```
1 ▾ # MAPE = pd.DataFrame({'Gap:5':mape_gap5, 'Gap:15':mape_gap15, 'Gap:3
2   # MAPE
```
executed in 55ms, finished 12:36:58 2020-10-28

In [26]:
```
1 ▾ # IOA.to_csv("~/Desktop/NCSA_genomics/Python - notebooks/GlucoCheck/M
2   # FB.to_csv("~/Desktop/NCSA_genomics/Python - notebooks/GlucoCheck/Me
3   # RMSE.to_csv("~/Desktop/NCSA_genomics/Python - notebooks/GlucoCheck/.
4   # MAPE.to_csv("~/Desktop/NCSA_genomics/Python - notebooks/GlucoCheck/.
5   # MAD.to_csv("~/Desktop/NCSA_genomics/Python - notebooks/GlucoCheck/M
```
executed in 11ms, finished 12:36:58 2020-10-28

In [ ]:
```
1
```