

Important: Read this!

This recipe does not work with the standard conda environment.

If you are in the standard environment, do this:

1. Stop Jupyter
2. Activate QIIME2 environment on conda
3. Do `jupyter serverextension enable --py qiime2 --sys-prefix`
4. Start Jupyter inside QIIME2 environment

Note that other recipes will not work inside this environment.

Check this out!

This is based on on [QIIME2 Fecal Microbiota Transplant example \(https://docs.qiime2.org/2018.8/tutorials/fmt/\)](https://docs.qiime2.org/2018.8/tutorials/fmt/) (for the command line). You are strongly advised to read it before proceeding.

There is an [amazing example \(http://nbviewer.jupyter.org/gist/tkoscio/29de5198a4be81559a075756c2490fde\)](http://nbviewer.jupyter.org/gist/tkoscio/29de5198a4be81559a075756c2490fde) of using the Artifact API using the "Moving Pictures" tutorial of QIIME 2 produced by Tomasz Kościółek. I use a more convoluted approach than Tomasz's in order to go a little deeper in terms of understanding of the Python internals. That is more of a learning experience on the internals than a practical recommendation. **My recommendation is to use Tomasz's dialect, not mine.**

Getting the data

```
In [ ]: !wget https://data.qiime2.org/2018.8/tutorials/fmt/sample_metadata.tsv
!wget https://data.qiime2.org/2018.8/tutorials/fmt/fmt-tutorial-demux-1-10p.qza
!wget https://data.qiime2.org/2018.8/tutorials/fmt/fmt-tutorial-demux-2-10n.qza
```

The recipe

```
In [1]: import pandas as pd

from qiime2.metadata.metadata import Metadata
from qiime2.metadata.metadata import CategoricalMetadataColumn
from qiime2.sdk import Artifact
from qiime2.sdk import PluginManager
from qiime2.sdk import Result
```

```
In [2]: pm = PluginManager()
demux_plugin = pm.plugins['demux']
#demux_emp_single = demux_plugin.actions['emp_single']
demux_summarize = demux_plugin.actions['summarize']
nm_plugins
```

```
Out[2]: {'alignment': <qiime2.plugin.plugin.Plugin at 0x7f843047e908>,
'composition': <qiime2.plugin.plugin.Plugin at 0x7f8430481be0>,
'cutadapt': <qiime2.plugin.plugin.Plugin at 0x7f8430493ba8>,
'dada2': <qiime2.plugin.plugin.Plugin at 0x7f84301e8dd8>,
'deblur': <qiime2.plugin.plugin.Plugin at 0x7f842fcea668>,
'demux': <qiime2.plugin.plugin.Plugin at 0x7f842fd41f98>,
'diversity': <qiime2.plugin.plugin.Plugin at 0x7f8430466128>,
'emperor': <qiime2.plugin.plugin.Plugin at 0x7f842fc2d8d0>,
'feature-classifier': <qiime2.plugin.plugin.Plugin at 0x7f842fd41668>,
'feature-table': <qiime2.plugin.plugin.Plugin at 0x7f842fd50898>,
'gneiss': <qiime2.plugin.plugin.Plugin at 0x7f843020dac8>,
'longitudinal': <qiime2.plugin.plugin.Plugin at 0x7f84304dc4a8>,
'metadata': <qiime2.plugin.plugin.Plugin at 0x7f843052b978>,
'phylogeny': <qiime2.plugin.plugin.Plugin at 0x7f84304931d0>,
'quality-control': <qiime2.plugin.plugin.Plugin at 0x7f84304e6f60>,
'quality-filter': <qiime2.plugin.plugin.Plugin at 0x7f8430525b38>,
'sample-classifier': <qiime2.plugin.plugin.Plugin at 0x7f8430560a20>,
'taxa': <qiime2.plugin.plugin.Plugin at 0x7f84361b56d8>,
'types': <qiime2.plugin.plugin.Plugin at 0x7f8436254898>,
'vsearch': <qiime2.plugin.plugin.Plugin at 0x7f84361a7f98>}
```

```
In [3]: print(demux_summarize.description)
demux_summarize_signature = demux_summarize.signature
print(demux_summarize_signature.inputs)
print(demux_summarize_signature.parameters)
print(demux_summarize_signature.outputs)
```

```
Summarize counts per sample for all samples, and generate interactive positional
quality plots based on `n` randomly selected sequences.
OrderedDict([('data', ParameterSpec(qiime_type=SampleData[JoinedSequencesWithQuali
ty | PairedEndSequencesWithQuality | SequencesWithQuality], view_type=<class '
q2_demux._summarize._visualizer._PlotQualView'>, default=NOVALUE, description='T
he demultiplexed sequences to be summarized.'))])
OrderedDict([('n', ParameterSpec(qiime_type=Int, view_type=<class 'int'>, defaul
t=10000, description='The number of sequences that should be selected at random
for quality score plots. The quality plots will present the average positional q
ualities across all of the sequences selected. If input sequences are paired end
, plots will be generated for both forward and reverse reads for the same `n` se
quences.'))])
OrderedDict([('visualization', ParameterSpec(qiime_type=Visualization, view_type
=None, default=NOVALUE, description=NOVALUE))])
```

```
In [4]: seqs1 = Result.load('fmt-tutorial-demux-1-10p.qza')
sum_data1 = demux_summarize(seqs1)

print(dir(sum_data1))
print(type(sum_data1.visualization))
print(dir(sum_data1.visualization))
sum_data1.visualization
[ '__add__', '__class__', '__contains__', '__delattr__', '__dict__', '__dir__', '__doc__', '__eq__', '__format__', '__ge__', '__getattr__', '__getitem__', '__getnewargs__', '__gt__', '__hash__', '__init__', '__iter__', '__le__', '__len__', '__lt__', '__module__', '__mul__', '__ne__', '__new__', '__reduce__', '__reduce_ex__', '__repr__', '__rmul__', '__setattr__', '__sizeof__', '__str__', '__subclasshook__', '__fields__', 'count', 'index', 'visualization']
<class 'qiime2.sdk.result.Visualization'>
[ '__class__', '__delattr__', '__dict__', '__dir__', '__doc__', '__eq__', '__format__', '__ge__', '__getattr__', '__gt__', '__hash__', '__init__', '__le__', '__lt__', '__module__', '__ne__', '__new__', '__reduce__', '__reduce_ex__', '__repr__', '__setattr__', '__sizeof__', '__str__', '__subclasshook__', '__weakref__', '__alias__', '__archiver__', '__destructor__', '__from_data_dir__', '__is_valid_type__', '__repr_html__', '__citations__', '__export_data__', '__extension__', '__extract__', '__format__', '__get_index_paths__', 'load', 'peek', 'save', 'type', 'uuid']
```

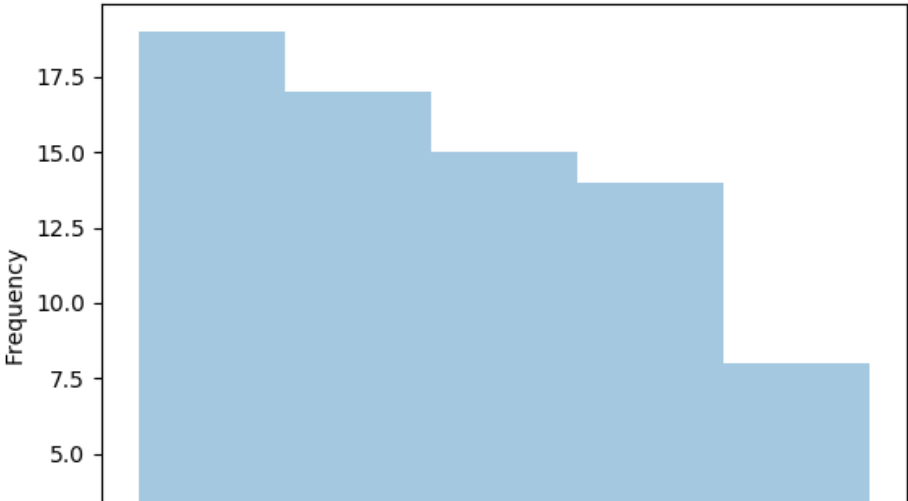
Out[4]:

Overview

Interactive Quality Plot

Demultiplexed sequence counts summary

Minimum:	1208
Median:	4324.0
Mean:	4578.50684932
Maximum:	8918
Total:	334231



```
In [5]: seqs2 = Result.load('fmt-tutorial-demux-2-10p.qza')
sum_data2 = demux_summarize(seqs2)

print(dir(sum_data2))
print(type(sum_data2.visualization))
print(dir(sum_data2.visualization))
sum_data2.visualization
[ '__add__', '__class__', '__contains__', '__delattr__', '__dict__', '__dir__', '__doc__', '__eq__', '__format__', '__ge__', '__getattr__', '__getitem__', '__getnewargs__', '__gt__', '__hash__', '__init__', '__iter__', '__le__', '__len__', '__lt__', '__module__', '__mul__', '__ne__', '__new__', '__reduce__', '__reduce_ex__', '__repr__', '__rmul__', '__setattr__', '__sizeof__', '__str__', '__subclasshook__', '__fields__', 'count', 'index', 'visualization']
<class 'qiime2.sdk.result.Visualization'>
[ '__class__', '__delattr__', '__dict__', '__dir__', '__doc__', '__eq__', '__format__', '__ge__', '__getattr__', '__gt__', '__hash__', '__init__', '__le__', '__lt__', '__module__', '__ne__', '__new__', '__reduce__', '__reduce_ex__', '__repr__', '__setattr__', '__sizeof__', '__str__', '__subclasshook__', '__weakref__', '__alias__', '__archiver__', '__destructor__', '__from_data_dir__', '__is_valid_type__', '__repr_html__', '__citations__', '__export_data__', '__extension__', '__extract__', '__format__', '__get_index_paths__', 'load', 'peek', 'save', 'type', 'uuid']
```

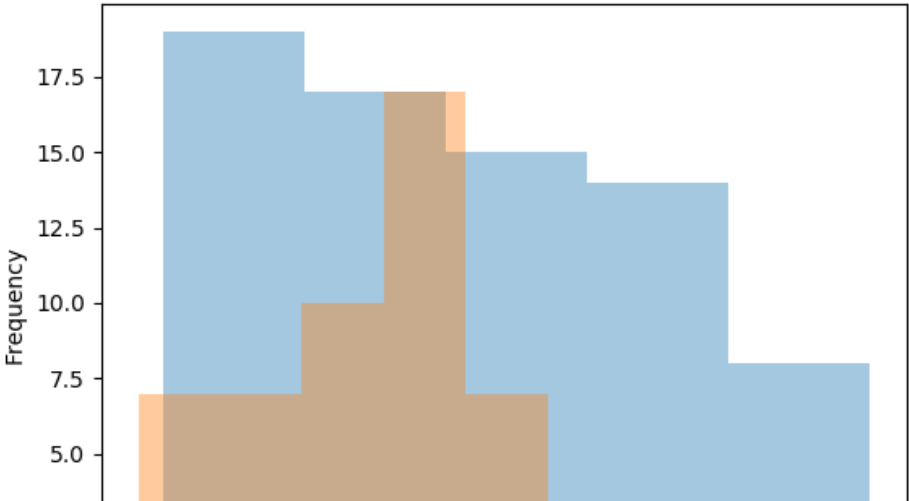
Out[5]:

Overview

Interactive Quality Plot

Demultiplexed sequence counts summary

Minimum:	930
Median:	3522.0
Mean:	3315.89583333
Maximum:	5400
Total:	159163



In [6]: *#Quality control*

```
dada2_plugin = pm.plugins['dada2']  
dada2_denoise_single = dada2_plugin.actions['denoise_single']  
qual_control1 = dada2_denoise_single(demultiplexed_seqs=seqs1,  
                                     trunc_len=150 trim_left=13)
```

Running external command line application(s). This may print messages to stdout and/or stderr.

The command(s) being run are below. These commands cannot be manually re-run as they will depend on temporary files that no longer exist.

Command: run_dada_single.R /tmp/qiime2-archive-uidiox5/5db90b62-c6e4-4c09-8c79-c0cdfbe2cea0/data /tmp/tmpeu6ezzg7/output.tsv.biom /tmp/tmpeu6ezzg7/track.tsv /tmp/tmpeu6ezzg7 150 13 2.0 2 Inf consensus 1.0 1 1000000 NULL 16

In [7]: *qual_control2 = dada2_denoise_single(demultiplexed_seqs=seqs2,*
trunc_len=150 trim_left=13)

Running external command line application(s). This may print messages to stdout and/or stderr.

The command(s) being run are below. These commands cannot be manually re-run as they will depend on temporary files that no longer exist.

Command: run_dada_single.R /tmp/qiime2-archive-de622u8k/5f459ea8-c6a8-438e-b42f-35e4f86166d7/data /tmp/tmpi0zda_vn/output.tsv.biom /tmp/tmpi0zda_vn/track.tsv /tmp/tmpi0zda_vn 150 13 2.0 2 Inf consensus 1.0 1 1000000 NULL 16

In [8]:

metadata_plugin = pm.plugins['metadata']
metadata_tabulate = metadata_plugin.actions['tabulate']
stats_meta1 = metadata_tabulate(input=qual_control1.denoising_stats.view(Metadata))
stats_meta1.visualization

Out[8]:

Download metadata TSV file (metadata.tsv)

This file won't necessarily reflect dynamic sorting or filtering options based on the interactive table below.

Search:

sample-id	input	filtered	denoise
#q2:types	numeric	numeric	numeric
0bb5e24a.aa34.48e7.b1f9.e761ac2dc6b4	6621	6488	6488
101c02ce.4f8d.4394.83a3.0c406831d934	6634	6545	6545
104e5902.1c3e.417f.bb22.124bfca61a61	7952	7795	7795
11c4be47.fdce.4135.8acf.0894f2da5ede	3049	3017	3017
11c70d46.e0a4.4797.ad7b.3677fe93d73b	5204	5143	5143
11f3f37a.3698.4031.8183.75f6f2c6c04c	2664	2642	2642
124dc1db.9029.41c6.afff.b73a692e19b1	3554	3517	3517
15129220.6eda.44ad.9511.be0f6e027f5d	6633	6516	6516
17f96e8a.dc68.4aef.98db.8aaaab702e33	4359	4313	4313
183645ce.cacd.4744.8440.de4eb377dfea	5908	5853	5853
1fe57dd4.58b3.4fde.b8a1.f73db5614909	3986	3894	3894
20aa4aaf.8242.48d9.a225.e9f6b3e8938c	8667	8469	8469
212f0ce6.9123.4e38.abf3.db944de7f1c1	2687	2657	2657
21a4ec1a.bcae.4e10.be82.66328d2dad4f	7084	6947	6947

Open in a: [new window \(http://localhost:8888/qiime2/view/3028e5f9-5aff-4ae6-a163-41942d4a1ccc/\)](http://localhost:8888/qiime2/view/3028e5f9-5aff-4ae6-a163-41942d4a1ccc/)

In [9]: stats_meta2 = metadata_tabulate(input=qual_control2.denoising_stats.view(Metadata))
stats_meta2_visualization

Out[9]: Download metadata TSV file (metadata.tsv)

This file won't necessarily reflect dynamic sorting or filtering options based on the interactive table below.

Search:

sample-id #q2:types	input numeric	filtered numeric	denoise numeric
00954e18.1dcf.40fc.a52e.255f5cda6f9c	2767	2752	2752
0476f5aa.a4f4.4835.ada1.4e72fb601e0a	4696	4650	4650
04ea008a.d444.4833.988c.a2a3f5d4ad23	1361	1355	1355
08ea0c6b.7a30.4074.bbf6.557480fc5285	3901	3859	3859
0eb12bd1.4548.4004.8dcd.1a71de8b1df5	5018	4937	4937
1037164d.a836.4ebe.bf98.92294a0b60ab	3234	3203	3203
1230c3fd.efa4.4ad9.8f01.499141965b11	5400	5328	5328
14c94c5e.aaf2.414f.a72f.74ccc134b294	4610	4559	4559
1654ccec.557a.40f3.b4d9.b263f9e2ad4e	4217	4193	4193
18f853f2.0e63.400e.baac.cad966c6c0a0	2949	2923	2923
1c4ca381.9cc3.4313.a11e.df350bb73ec9	2104	2087	2087
1fe19636.3708.4a87.8e5a.cf10d8530d4e	930	923	923
230994e3.5682.432f.b4e4.454ca1901e0f	3737	3702	3702
28e84479.bd9f.4d55.bb03.1411b71bdc9b	3317	3297	3297

Open in a: [new window \(http://localhost:8888/qiime2/view/8fb65ec5-1576-4cdb-9191-e387c04bf9f7/\)](http://localhost:8888/qiime2/view/8fb65ec5-1576-4cdb-9191-e387c04bf9f7/)

In [10]: ft_plugin = pm.plugins['feature-table']
ft_merge = ft_plugin.actions['merge']
ft_merge_seqs = ft_plugin.actions['merge_seqs']
ft_summarize = ft_plugin.actions['summarize']
ft_tab_seqs = ft_plugin.actions['tabulate_seqs']

table_merge = ft_merge(tables=[qual_control1.table, qual_control2.table])
seqs_merge = ft_merge_seqs(data=[qual_control1.representative_sequences, qual_control2.representative_sequences])

In [13]:

ft_sum = ft_summarize(table=table_merge.merged_table)
ft_sum_visualization

Out[13]:

Overview

Interactive Sample Detail

Feature Detail

Table summary

Metric	Sample
Number of samples	121
Number of features	787
Total frequency	475,397

Frequency per sample

	Frequency
Minimum frequency	923.0
1st quartile	2,594.0
Median frequency	3,764.0
3rd quartile	4,906.0
Maximum frequency	8,412.0
Mean frequency	3,928.900826446281

Frequency per sample detail (csv (sample-frequency-detail.csv) | html (sample-frequency-detail.html))

Open in a: [new window \(http://localhost:8888/qiime2/view/2d2f3759-6cc4-447a-b62f-372e248c8f07/\)](http://localhost:8888/qiime2/view/2d2f3759-6cc4-447a-b62f-372e248c8f07/)


```
In [12]: tab_seqs = ft_tab_seqs(data=seqs_merge.merged_data)
tab_seqs_visualization
```

Out[12]: To BLAST a sequence against the NCBI nt database, click the sequence and then click the *View report* button on the resulting page.

To download a raw FASTA file of your sequences, click here (sequences.fasta).

Click on a Column header to sort the table.

Feature ID	Sequence
002e78333d6cf2b11aa7a5ba03dd2c68	AGCGTTATCCGGATTTATTGGGTTTAAAGGGCGCGTAGGCCGGC (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi?ALIGNMENT_VIEW&QUERY=AGCGTTATCCGGATTTATTGGGTTTAAAGGGCGCGTAGGCCGGC)
0046913ae6f9e12dbd889671ed26c09d	AGCGTTGTCCGGAATTATTGGGCGTAAAGGGCGCGCAGGCCGGC (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi?ALIGNMENT_VIEW&QUERY=AGCGTTGTCCGGAATTATTGGGCGTAAAGGGCGCGCAGGCCGGC)
009c3d1fd56cf5682c875f959d9fee33	AGCGTTATCCGGATTTATTGGGTGTAAAGGGTGCGTAGACGGG (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi?ALIGNMENT_VIEW&QUERY=AGCGTTATCCGGATTTATTGGGTGTAAAGGGTGCGTAGACGGG)
00c7ff8350c1f1874014b473b1890bf3	GGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCCGGT (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi?ALIGNMENT_VIEW&QUERY=GGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCCGGT)
014ed0286a2651609cc4a145b7ea6788	GGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCCGGC (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi?ALIGNMENT_VIEW&QUERY=GGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCCGGC)
02812d27f4bf6ec2c62d954495d6ede5	AGCGTTGTCCGGAATTATTGGGCGTAAAGGGTACGTAGGCCGGT (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi?ALIGNMENT_VIEW&QUERY=AGCGTTGTCCGGAATTATTGGGCGTAAAGGGTACGTAGGCCGGT)
028eb2e964dcc235351e1380676e4d77	AGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGGTAGCCGGC (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi?ALIGNMENT_VIEW&QUERY=AGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGGTAGCCGGC)

Open in a: [new window \(http://localhost:8888/qiime2/view/3889631a-8101-4234-b9a5-8f8aa1fee838/\)](http://localhost:8888/qiime2/view/3889631a-8101-4234-b9a5-8f8aa1fee838/)

In []: