edgar_xbrl

September 29, 2021

[3]:
```
from io import BytesIO
from zipfile import ZipFile, BadZipFile
import requests
from datetime import date
from pathlib import Path
import pandas_datareader.data as web
import pandas as pd
import json
from pprint import pprint
import matplotlib.pyplot as plt
import matplotlib.ticker as mticker
```

[4]:
```
%matplotlib inline
plt.style.use('fivethirtyeight')
data_path = Path('data') # perhaps set to external harddrive to accomodate␣
 ↪large amount of data
```

0.1 Download FS & Notes

The following code downloads and extracts all historical filings contained in the Financial Statement and Notes (FSN) datasets for the given range of quarters:

Downloads over 40GB of data!

[3]:
```
SEC_URL = 'https://www.sec.gov/files/dera/data/
 ↪financial-statement-and-notes-data-sets/'

today = pd.Timestamp(date.today())
this_year = today.year
this_quarter = today.quarter

past_years = range(2014, this_year)
filing_periods = [(y, q) for y in past_years for q in range(1, 5)]
filing_periods.extend([(this_year, q) for q in range(1, this_quarter + 1)])
for i, (yr, qtr) in enumerate(filing_periods, 1):
    print(yr, qtr, end=' ', flush=True)
    filing = f'{yr}q{qtr}_notes.zip'
    path = data_path / f'{yr}_{qtr}' / 'source'
```

```
    if not path.exists():
        path.mkdir(exist_ok=True, parents=True)

    response = requests.get(SEC_URL + filing).content
    try:
        with ZipFile(BytesIO(response)) as zip_file:
            for file in zip_file.namelist():
                local_file = path / file
                if local_file.exists():
                    continue
                with local_file.open('wb') as output:
                    for line in zip_file.open(file).readlines():
                        output.write(line)
    except BadZipFile:
        continue
```

```
2014 1 2014 2 2014 3 2014 4 2015 1 2015 2 2015 3 2015 4 2016 1 2016 2 2016 3
2016 4 2017 1 2017 2 2017 3 2017 4 2018 1 2018 2 2018 3 2018 4 2019 1
```

0.2 Save to parquet

The data is fairly large and to enable faster access than the original text files permit, it is better to convert the text files to binary, columnar parquet format (see Section 'Efficient data storage with pandas' in chapter 2 for a performance comparison of various data-storage options compatible with pandas DataFrames):

[4]:
```
for f in data_path.glob('**/*.tsv'):
    file_name = f.stem  + '.parquet'
    path = Path(f.parents[1]) / 'parquet'
    if (path / file_name).exists():
        continue
    if not path.exists():
        path.mkdir(exist_ok=True)
    try:
        df = pd.read_csv(f, sep='\t', encoding='latin1', low_memory=False)
    except:
        print(f)
    df.to_parquet(path / file_name)
```

0.3 Metadata json

[5]:
```
file = data_path / '2018_3' / 'source' / '2018q3_notes-metadata.json'
with file.open() as f:
    data = json.load(f)

pprint(data)
```

```
{'@context': 'http://www.w3.org/ns/csvw',
 'dialect': {'delimiter': '\t', 'header': True, 'headerRowCount': 1},
 'tables': [{'tableSchema': {'aboutUrl': 'readme.htm',
                             'columns': [{'datatype': {'base': 'string',
                                                       'maxLength': 20,
                                                       'minLength': 20},
                                          'dc:description': 'Accession Number. '
                                                            'The 20-character '
                                                            'string formed '
                                                            'from the 18-digit '
                                                            'number assigned '
                                                            'by the Commission '
                                                            'to each EDGAR '
                                                            'submission.',
                                          'name': 'adsh',
                                          'required': 'true',
                                          'titles': ['Accession Number']},
                                         {'datatype': {'base': 'decimal',
                                                       'maxLength': 10,
                                                       'minInclusive': 0},
                                          'dc:description': 'Central Index Key '
                                                            '(CIK). Ten digit '
                                                            'number assigned '
                                                            'by the Commission '
                                                            'to each '
                                                            'registrant that '
                                                            'submits filings.',
                                          'name': 'cik',
                                          'titles': ['Central Index Key']},
                                         {'datatype': {'base': 'string',
                                                       'maxLength': 150},
                                          'dc:description': 'Name of '
                                                            'registrant. This '
                                                            'corresponds to '
                                                            'the name of the '
                                                            'legal entity as '
                                                            'recorded in EDGAR '
                                                            'as of the filing '
                                                            'date.',
                                          'name': 'name',
                                          'titles': ['Registrant']},
                                         {'datatype': {'base': 'string',
                                                       'maxLength': 4},
                                          'dc:description': 'Standard '
                                                            'Industrial '
                                                            'Classification '
                                                            '(SIC). Four digit '
                                                            'code assigned by '
```

```
                    'the Commission as '
                    'of the filing '
                    'date, indicating '
                    "the registrant's "
                    'type of business.',
 'name': 'sic',
 'titles': ['Standard Industrial '
            'Classification Code']},
{'datatype': {'base': 'string',
              'maxLength': 2,
              'minLength': 2},
 'dc:description': 'The ISO 3166-1 '
                   'country of the '
                   "registrant's "
                   'business address.',
 'name': 'countryba',
 'titles': ['Business Address Country',
            'Country (B)']},
{'datatype': {'base': 'string',
              'maxLength': 2,
              'minLength': 2},
 'dc:description': 'The state or '
                   'province of the '
                   "registrant's "
                   'business address, '
                   'if field '
                   'countryba is US '
                   'or CA.',
 'name': 'stprba',
 'titles': ['Business Address State '
            'or Province',
            'State (B)']},
{'datatype': {'base': 'string',
              'maxLength': 30},
 'dc:description': 'The city of the '
                   "registrant's "
                   'business address.',
 'name': 'cityba',
 'titles': ['Business Address City',
            'City (B)']},
{'datatype': {'base': 'string',
              'maxLength': 10},
 'dc:description': 'The zip code of '
                   "the registrant's "
                   'business address.',
 'name': 'zipba',
 'titles': ['Business Address Zip or '
            'Postal Code',
```

```
            'Zip (B)']},
{'datatype': {'base': 'string',
              'maxLength': 40},
 'dc:description': 'The first line of '
                   'the street of the '
                   "registrant's "
                   'business address.',
 'name': 'bas1',
 'titles': ['Business Address Street '
            '1',
            'Street1 (B)']},
{'datatype': {'base': 'string',
              'maxLength': 40},
 'dc:description': 'The second line '
                   'of the street of '
                   "the registrant's "
                   'business address.',
 'name': 'bas2',
 'titles': ['Business Address Street '
            '2',
            'Street2 (B)']},
{'datatype': {'base': 'string',
              'maxLength': 12},
 'dc:description': 'The phone number '
                   'of the '
                   "registrant's "
                   'business address.',
 'name': 'baph',
 'titles': ['Business Address Phone',
            'Phone (B)']},
{'datatype': {'base': 'string',
              'maxLength': 2,
              'minLength': 2},
 'dc:description': 'The ISO 3166-1 '
                   'country of the '
                   "registrant's "
                   'mailing address.',
 'name': 'countryma',
 'titles': ['Mailing Address Country',
            'Country (M)']},
{'datatype': {'base': 'string',
              'maxLength': 2,
              'minLength': 2},
 'dc:description': 'The state or '
                   'province of the '
                   "registrant's "
                   'mailing address, '
                   'if field '
```

```
                    'countryma is US '
                    'or CA.',
 'name': 'stprma',
 'titles': ['Mailing Address State or '
            'Province',
            'State (M)']},
{'datatype': {'base': 'string',
              'maxLength': 30},
 'dc:description': 'The city of the '
                   "registrant's "
                   'mailing address.',
 'name': 'cityma',
 'titles': ['Mailing Address City',
            'City (M)']},
{'datatype': {'base': 'string',
              'maxLength': 12},
 'dc:description': 'The zip code of '
                   "the registrant's "
                   'mailing address.',
 'name': 'zipma',
 'titles': ['Mailing Address Zip or '
            'Postal Code',
            'Zip (M)']},
{'datatype': {'base': 'string',
              'maxLength': 40},
 'dc:description': 'The first line of '
                   'the street of the '
                   "registrant's "
                   'mailing address.',
 'name': 'mas1',
 'titles': ['Mailing Address Street1',
            'Street1 (M)']},
{'datatype': {'base': 'string',
              'maxLength': 40},
 'dc:description': 'The second line '
                   'of the street of '
                   "the registrant's "
                   'mailing address.',
 'name': 'mas2',
 'titles': ['Mailing Address Street2',
            'Street1 (M)']},
{'datatype': {'base': 'string',
              'maxLength': 2,
              'minLength': 2},
 'dc:description': 'The country of '
                   'incorporation for '
                   'the registrant.',
 'name': 'countryinc',
```

```
 'titles': ['Country of Incorporation',
            'Incorporation Country']},
{'datatype': {'base': 'string',
              'maxLength': 2,
              'minLength': 2},
 'dc:description': 'The state or '
                   'province of '
                   'incorporation for '
                   'the registrant, '
                   'if countryinc is '
                   'US or CA, '
                   'otherwise NULL.',
 'name': 'stprinc',
 'titles': ['State or Province of '
            'Incorporation',
            'Incorporation State']},
{'datatype': {'base': 'string',
              'maxLength': 9},
 'dc:description': 'Employee '
                   'Identification '
                   'Number, 9 digit '
                   'identification '
                   'number assigned '
                   'by the Internal '
                   'Revenue Service '
                   'to business '
                   'entities '
                   'operating in the '
                   'United States.',
 'name': 'ein',
 'titles': ['EIN',
            'Employee Identification '
            'Number']},
{'datatype': {'base': 'string',
              'maxLength': 150},
 'dc:description': 'Most recent '
                   'former name of '
                   'the registrant, '
                   'if any.',
 'name': 'former',
 'titles': ['Former Name']},
{'datatype': {'base': 'string',
              'maxLength': 8,
              'minLength': 8},
 'dc:description': 'Date of change '
                   'from the former '
                   'name, if any.',
 'name': 'changed',
```

```
 'titles': ['Date of Name Change']},
{'datatype': {'base': 'string',
              'maxLength': 5},
 'dc:description': 'Filer status with '
                   'the Commission at '
                   'the time of '
                   'submission: '
                   '1-LAF=Large '
                   'Accelerated, '
                   '2-ACC=Accelerated,
'
                   '3-SRA=Smaller '
                   'Reporting '
                   'Accelerated, '
                   '4-NON=Non-
Accelerated, '
                   '5-SML=Smaller '
                   'Reporting Filer, '
                   'NULL=not '
                   'assigned.',
 'name': 'afs',
 'titles': ['Status',
            'Accelerated Filer '
            'Status']},
{'datatype': {'base': 'decimal',
              'maxInclusive': 1,
              'minInclusive': 0},
 'dc:description': 'Well Known '
                   'Seasoned Issuer '
                   '(WKSI). An issuer '
                   'that meets '
                   'specific '
                   'Commission '
                   'requirements at '
                   'some point during '
                   'a 60-day period '
                   'preceding the '
                   'date the issuer '
                   'satisfies its '
                   'obligation to '
                   'update its shelf '
                   'registration '
                   'statement.',
 'name': 'wksi',
 'titles': ['Well-known Seasoned '
            'Issuer']},
{'datatype': {'base': 'string',
              'maxLength': 4},
```

```
 'dc:description': 'Fiscal Year End '
                   'Date.',
 'name': 'fye',
 'titles': ['FY End Date']},
{'datatype': {'base': 'string',
              'maxLength': 20},
 'dc:description': 'The submission '
                   'type of the '
                   "registrant's "
                   'filing.',
 'name': 'form',
 'titles': ['Submission Type',
            'Filing Type',
            'EDGAR Form Type']},
{'datatype': {'base': 'string',
              'maxLength': 8,
              'minLength': 8},
 'dc:description': 'Balance Sheet '
                   'Date.',
 'name': 'period',
 'titles': ['Report Period',
            'Date of Balance Sheet']},
{'datatype': {'base': 'string',
              'maxLength': 4,
              'minLength': 4},
 'dc:description': 'Fiscal Year Focus '
                   '(as defined in '
                   'EFM Ch. 6).',
 'name': 'fy',
 'titles': ['Fiscal Year']},
{'datatype': {'base': 'string',
              'maxLength': 2,
              'minLength': 2},
 'dc:description': 'Fiscal Period '
                   'Focus (as defined '
                   'in EFM Ch. 6) '
                   'within Fiscal '
                   'Year. The 10-Q '
                   'for the 1st, 2nd '
                   'and 3rd quarters '
                   'would have a '
                   'fiscal period '
                   'focus of Q1, Q2 '
                   '(or H1), and Q3 '
                   '(or M9) '
                   'respectively, and '
                   'a 10-K would have '
                   'a fiscal period '
```

```
                   'focus of FY.',
  'name': 'fp',
  'titles': ['Fiscal Period']},
 {'datatype': {'base': 'string',
               'maxLength': 8},
  'dc:description': 'The date of the '
                    "registrant's "
                    'filing with the '
                    'Commission.',
  'name': 'filed',
  'titles': ['Date Filed']},
 {'datatype': {'base': 'date',
               'format': 'YYYYMMDD '
                         'HH:MM:SS.S'},
  'dc:description': 'The acceptance '
                    'date and time of '
                    "the registrant's "
                    'filing with the '
                    'Commission. '
                    'Filings accepted '
                    'after 5:30pm EST '
                    'are considered '
                    'filed on the '
                    'following '
                    'business day.',
  'name': 'accepted',
  'titles': ['Acceptance Datetime']},
 {'datatype': {'base': 'decimal',
               'maxInclusive': 255,
               'minInclusive': 0},
  'dc:description': 'Previous Report.  '
                    'TRUE indicates '
                    'that the '
                    'submission '
                    'information was '
                    'subsequently '
                    'amended prior to '
                    'the end cutoff '
                    'date of the data '
                    'set.',
  'name': 'prevrpt',
  'required': 'true',
  'titles': ['Previous Report Flag',
             'Subsequently Amended '
             'Flag']},
 {'datatype': {'base': 'decimal',
               'maxInclusive': 255,
               'minInclusive': 0},
```

```
 'dc:description': 'TRUE indicates '
                   'that the XBRL '
                   'submission '
                   'contains '
                   'quantitative '
                   'disclosures '
                   'within the '
                   'footnotes and '
                   'schedules at the '
                   'required detail '
                   'level (e.g., each '
                   'amount).',
 'name': 'detail',
 'required': 'true',
 'titles': ['Detail Tagged']},
{'datatype': {'base': 'string',
              'maxLength': 32},
 'dc:description': 'The name of the '
                   'submitted XBRL '
                   'Instance Document '
                   '(EX-101.INS) type '
                   'data file. The '
                   'name often begins '
                   'with the company '
                   'ticker symbol.',
 'name': 'instance',
 'titles': ['Instance Filename']},
{'datatype': {'base': 'decimal',
              'maxInclusive': 32767,
              'minInclusive': 0},
 'dc:description': 'Number of Central '
                   'Index Keys (CIK) '
                   'of registrants '
                   '(i.e., business '
                   'units) included '
                   'in the '
                   'consolidating '
                   "entity's "
                   'submitted filing.',
 'name': 'nciks',
 'required': 'true',
 'titles': ['Number of '
            'Coregistrants']},
{'datatype': {'base': 'string',
              'maxLength': 120},
 'dc:description': 'Additional CIKs '
                   'of co-registrants '
                   'included in a '
```

```
                    'consolidating '
                    "entity's EDGAR "
                    'submission, '
                    'separated by '
                    'spaces. If there '
                    'are no other '
                    'co-registrants '
                    '(i.e., nciks = '
                    '1), the value of '
                    'aciks is NULL. '
                    'For a very small '
                    'number of filers, '
                    'the list of '
                    'co-registrants is '
                    'too long to fit '
                    'in the field. '
                    'Where this is the '
                    'case, PARTIAL '
                    'will appear at '
                    'the end of the '
                    'list indicating '
                    'that not all '
                    "co-registrants' "
                    'CIKs are included '
                    'in the field; '
                    'users should '
                    'refer to the '
                    'complete '
                    'submission file '
                    'for all CIK '
                    'information.',
  'name': 'aciks',
  'titles': ['Additional Coregistrant '
             'CIKs']},
 {'datatype': {'base': 'decimal'},
  'dc:description': 'Public float, in '
                    'USD, if provided '
                    'in this '
                    'submission.',
  'name': 'pubfloatusd',
  'titles': ['Public Float']},
 {'datatype': {'base': 'string',
               'maxLength': 8},
  'dc:description': 'Date on which the '
                    'public float was '
                    'measured by the '
                    'filer.',
  'name': 'floatdate',
```

```
                                 'titles': ['Public Float Measurement '
                                            'Date']},
                                {'datatype': {'base': 'string',
                                              'maxLength': 255},
                                 'dc:description': 'If the public '
                                                   'float value was '
                                                   'computed by '
                                                   'summing across '
                                                   'several tagged '
                                                   'values, this '
                                                   'indicates the '
                                                   'nature of the '
                                                   'summation.',
                                 'name': 'floataxis',
                                 'titles': ['Public Float Axis']},
                                {'datatype': {'base': 'decimal',
                                              'maxInclusive': 255,
                                              'minInclusive': 0},
                                 'dc:description': 'If the public '
                                                   'float was '
                                                   'computed, the '
                                                   'number of terms '
                                                   'in the summation.',
                                 'name': 'floatmems',
                                 'titles': ['Public Float Members']}],
                  'primaryKey': 'adsh'},
  'url': 'sub.tsv'},
 {'tableSchema': {'aboutUrl': 'readme.htm',
                  'columns': [{'datatype': {'base': 'string',
                                            'maxLength': 256},
                               'dc:description': 'The unique '
                                                 'identifier (name) '
                                                 'for a tag in a '
                                                 'specific taxonomy '
                                                 'release.',
                               'name': 'tag',
                               'required': 'true',
                               'titles': ['Localname']},
                              {'datatype': {'base': 'string',
                                            'maxLength': 20},
                               'dc:description': 'For a standard '
                                                 'tag, an '
                                                 'identifier for '
                                                 'the taxonomy; '
                                                 'otherwise the '
                                                 'accession number '
                                                 'where the tag was '
                                                 'defined.',
```

```
 'name': 'version',
 'required': 'true',
 'titles': ['Namespace', 'Taxonomy']},
{'datatype': {'base': 'decimal',
              'maxInclusive': 1,
              'minInclusive': 0},
 'dc:description': '1 if tag is '
                   'custom '
                   '(version=adsh), 0 '
                   'if it is '
                   'standard. Note: '
                   'This flag is '
                   'technically '
                   'redundant with '
                   'the version and '
                   'adsh fields.',
 'name': 'custom',
 'required': 'true',
 'titles': []},
{'datatype': {'base': 'decimal',
              'maxInclusive': 1,
              'minInclusive': 0},
 'dc:description': '1 if the tag is '
                   'not used to '
                   'represent a '
                   'numeric fact.',
 'name': 'abstract',
 'required': 'true',
 'titles': []},
{'datatype': {'base': 'string',
              'maxLength': 20},
 'dc:description': 'If abstract=1, '
                   'then NULL, '
                   'otherwise the '
                   'data type (e.g., '
                   'monetary) for the '
                   'tag.',
 'name': 'datatype',
 'titles': []},
{'datatype': {'base': 'string',
              'maxLength': 1},
 'dc:description': 'If abstract=1, '
                   'then NULL; '
                   'otherwise, I if '
                   'the value is a '
                   'point in time, or '
                   'D if the value is '
                   'a duration.',
```

```
 'name': 'iord',
 'titles': ['Instant or Duration']},
{'datatype': {'base': 'string',
              'maxLength': 1},
 'dc:description': 'If datatype = '
                   'monetary, then '
                   "the tag's natural "
                   'accounting '
                   'balance from the '
                   'perspective of '
                   'the balance sheet '
                   'or income '
                   'statement (debit '
                   'or credit); if '
                   'not defined, then '
                   'NULL.',
 'name': 'crdr',
 'titles': ['Credit or Debit']},
{'datatype': {'base': 'string',
              'maxLength': 512},
 'dc:description': 'If a standard '
                   'tag, then the '
                   'label text '
                   'provided by the '
                   'taxonomy, '
                   'otherwise the '
                   'text provided by '
                   'the filer. A tag '
                   'which had neither '
                   'would have a NULL '
                   'value here.',
 'name': 'tlabel',
 'titles': ['Label']},
{'datatype': {'base': 'string',
              'maxLength': 2048},
 'dc:description': 'The detailed '
                   'definition for '
                   'the tag, '
                   'truncated to 2048 '
                   'characters. If a '
                   'standard tag, '
                   'then the text '
                   'provided by the '
                   'taxonomy, '
                   'otherwise the '
                   'text assigned by '
                   'the filer. Some '
                   'tags have '
```

```
                                                'neither, in which '
                                                'case this field '
                                                'is NULL.',
                                    'name': 'doc',
                                    'titles': ['Documentation']}],
                        'primaryKey': ['tag', 'version']},
  'url': 'tag.tsv'},
 {'tableSchema': {'aboutUrl': 'readme.htm',
                        'columns': [{'datatype': {'base': 'string',
                                                  'maxLength': 34},
                                     'dc:description': 'MD5 hash of the '
                                                'segments field '
                                                'text. Although '
                                                'MD5 is unsuitable '
                                                'for cryptographic '
                                                'use, it is used '
                                                'here merely to '
                                                'limit the size of '
                                                'the primary key.',
                                     'name': 'dimh',
                                     'required': 'true',
                                     'titles': ['Dimension Hash']},
                                    {'datatype': {'base': 'string',
                                                  'maxLength': 1024},
                                     'dc:description': 'Concatenation of '
                                                'tag names '
                                                'representing the '
                                                'axis and members '
                                                'appearing in the '
                                                'XBRL segments. '
                                                'Tag names have '
                                                'their first '
                                                'characters '
                                                '"Statement", last '
                                                '4 characters '
                                                '"Axis", and last '
                                                '6 characters '
                                                '"Member" or '
                                                '"Domain" '
                                                'truncated where '
                                                'they appear. '
                                                'Namespaces and '
                                                'prefixes are '
                                                'ignored because '
                                                'EDGAR validation '
                                                'guarantees that '
                                                'the local-names '
                                                'are unique with a '
```

```
                                                        'submission. Each '
                                                        'dimension is '
                                                        'represented as '
                                                        'the pair '
                                                        '"{axis}={member};"
'
                                                        'and the axes '
                                                        'concatenated in '
                                                        'lexical order. '
                                                        'Example: '
'"LegalEntity=Xyz;Scenario=Restated;" '
                                                        'represents the '
                                                        'XBRL segment with '
                                                        'dimension '
                                                        'LegalEntityAxis '
                                                        'and member '
                                                        'XyzMember, '
                                                        'dimension '
'StatementScenarioAxis '
                                                        'and member '
                                                        'RestatedMember.',
                                         'name': 'segments',
                                         'titles': []},
                                        {'datatype': {'base': 'decimal',
                                                      'maxInclusive': 1,
                                                      'minInclusive': 0},
                                         'dc:description': 'TRUE if the '
                                                           'segments field '
                                                           'would have been '
                                                           'longer than 1024 '
                                                           'characters had it '
                                                           'not been '
                                                           'truncated, else '
                                                           'FALSE.',
                                         'name': 'segt',
                                         'required': 'true',
                                         'titles': ['Segments Truncated']}],
                            'primaryKey': 'dimh'},
             'url': 'dim.tsv'},
            {'tableSchema': {'aboutUrl': 'readme.htm',
                             'columns': [{'datatype': {'base': 'string',
                                                       'maxLength': 20,
                                                       'minLength': 20},
                                          'dc:description': 'Accession Number. '
                                                            'The 20-character '
                                                            'string formed '
                                                            'from the 18-digit '
                                                            'number assigned '
```

```
                    'by the Commission '
                    'to each EDGAR '
                    'submission.',
  'name': 'adsh',
  'required': 'true',
  'titles': ['Accession Number']},
 {'datatype': {'base': 'string',
               'maxLength': 255},
  'dc:description': 'The unique '
                    'identifier (name) '
                    'for a tag in a '
                    'specific taxonomy '
                    'release.',
  'name': 'tag',
  'required': 'true',
  'titles': ['Localname']},
 {'datatype': {'base': 'string',
               'maxLength': 20},
  'dc:description': 'For a standard '
                    'tag, an '
                    'identifier for '
                    'the taxonomy; '
                    'otherwise the '
                    'accession number '
                    'where the tag was '
                    'defined.',
  'name': 'version',
  'required': 'true',
  'titles': ['Namespace']},
 {'datatype': {'base': 'string',
               'maxLength': 8,
               'minLength': 8},
  'dc:description': 'The end date for '
                    'the data value, '
                    'rounded to the '
                    'nearest month '
                    'end.',
  'name': 'ddate',
  'required': 'true',
  'titles': ['Data Date']},
 {'datatype': {'base': 'decimal',
               'minInclusive': 0},
  'dc:description': 'The count of the '
                    'number of '
                    'quarters '
                    'represented by '
                    'the data value, '
                    'rounded to the '
```

```
                    'nearest whole '
                    'number. "0" '
                    'indicates it is a '
                    'point-in-time '
                    'value.',
 'name': 'qtrs',
 'required': 'true',
 'titles': ['Quarters']},
{'datatype': {'base': 'string',
              'maxLength': 50},
 'dc:description': 'The unit of '
                   'measure for the '
                   'value.',
 'name': 'uom',
 'required': 'true',
 'titles': ['Unit of Measure']},
{'datatype': {'base': 'string',
              'maxLength': 34},
 'dc:description': 'The 32-byte '
                   'hexadecimal key '
                   'for the '
                   'dimensional '
                   'information in '
                   'the DIM data set.',
 'name': 'dimh',
 'titles': ['Dimension Hash']},
{'datatype': {'base': 'decimal',
              'maxInclusive': 32767,
              'minInclusive': 0},
 'dc:description': 'A positive '
                   'integer to '
                   'distinguish '
                   'different '
                   'reported facts '
                   'that otherwise '
                   'would have the '
                   'same primary key. '
                   'For most '
                   'purposes, data '
                   'with iprx greater '
                   'than 1 are not '
                   'needed. The '
                   'priority for the '
                   'fact based on '
                   'higher precision, '
                   'closeness of the '
                   'end date to a '
                   'month end, and '
```

```
                   'closeness of the '
                   'duration to a '
                   'multiple of three '
                   'months. See '
                   'fields dcml, durp '
                   'and datp below.',
  'name': 'iprx',
  'titles': ['Fact Preference']},
 {'datatype': {'base': 'decimal'},
  'dc:description': 'The value. This '
                   'is not scaled, it '
                   'is as found in '
                   'the Interactive '
                   'Data file, but is '
                   'rounded to four '
                   'digits to the '
                   'right of the '
                   'decimal point.',
  'name': 'value',
  'titles': []},
 {'datatype': {'base': 'string',
               'maxLength': 512},
  'dc:description': 'The plain text of '
                   'any superscripted '
                   'footnotes on the '
                   'value, if any, as '
                   'shown on the '
                   'statement page, '
                   'truncated to 512 '
                   'characters.',
  'name': 'footnote',
  'titles': ['Footnote Text']},
 {'datatype': {'base': 'decimal',
               'minInclusive': 0},
  'dc:description': 'Number of bytes '
                   'in the plain text '
                   'of the footnote '
                   'prior to '
                   'truncation; zero '
                   'if no footnote.',
  'name': 'footlen',
  'required': 'true',
  'titles': ['Footnote Length']},
 {'datatype': {'base': 'decimal',
               'minInclusive': 0},
  'dc:description': 'Small integer '
                   'representing the '
                   'number of '
```

```
                    'dimensions. Note '
                    'that this value '
                    'is a function of '
                    'the dimension '
                    'segments.',
  'name': 'dimn',
  'required': 'true',
  'titles': ['Number of Dimensions']},
 {'datatype': {'base': 'string',
               'maxLength': 256},
  'dc:description': 'If specified, '
                    'indicates a '
                    'specific '
                    'co-registrant, '
                    'the parent '
                    'company, or other '
                    'entity (e.g., '
                    'guarantor). NULL '
                    'indicates the '
                    'consolidated '
                    'entity. Note that '
                    'this value is a '
                    'function of the '
                    'dimension '
                    'segments.',
  'name': 'coreg',
  'titles': ['Coregistrant']},
 {'datatype': {'base': 'decimal'},
  'dc:description': 'The difference '
                    'between the '
                    'reported fact '
                    'duration and the '
                    'quarter duration '
                    '(qtrs), expressed '
                    'as a fraction of '
                    '1. For example, a '
                    'fact with '
                    'duration of 120 '
                    'days rounded to a '
                    '91-day quarter '
                    'has a durp value '
                    'of 29/91 = '
                    '+0.3187.',
  'name': 'durp',
  'titles': ['Duration Preference']},
 {'datatype': {'base': 'decimal'},
  'dc:description': 'The difference '
                    'between the '
```

```
                                                        'reported fact '
                                                        'date and the '
                                                        'month-end rounded '
                                                        'date (ddate), '
                                                        'expressed as a '
                                                        'fraction of 1. '
                                                        'For example, a '
                                                        'fact reported for '
                                                        '29/Dec, with '
                                                        'ddate rounded to '
                                                        '31/Dec, has a '
                                                        'datp value of '
                                                        'minus 2/31 = '
                                                        '-0.0645.',
                                         'name': 'datp',
                                         'titles': ['Date Preference']},
                                        {'datatype': {'base': 'decimal',
                                                      'maxInclusive': 32767,
                                                      'minInclusive': -32768},
                                         'dc:description': 'The value of the '
                                                           'fact "decimals" '
                                                           'attribute, with '
                                                           'INF represented '
                                                           'by 32767.',
                                         'name': 'dcml',
                                         'titles': ['Decimals']}],
                              'foreignKeys': [{'columnReference': 'adsh',
                                               'reference': {'columnReference':
'adsh',
                                                             'resource':
'sub.tsv'}},
                                              {'columnReference': 'dimh',
                                               'reference': {'columnReference':
'dimh',
                                                             'resource':
'https://wwww.sec.gov/files2018q3.zip#path=dim.tsv'}},
                                              {'columnReference': ['tag',
                                                                   'version'],
                                               'reference': {'columnReference':
['tag',
'version'],
                                                             'resource':
'https://wwww.sec.gov/files2018q3.zip#path=tag.tsv'}}],
                              'primaryKey': ['adsh',
                                             'tag',
                                             'version',
                                             'ddate',
                                             'qtrs',
```

```
                                  'uom',
                                  'dimh',
                                  'iprx']},
  'url': 'num.tsv'},
 {'tableSchema': {'aboutUrl': 'readme.htm',
                  'columns': [{'datatype': {'base': 'string',
                                            'maxLength': 20,
                                            'minLength': 20},
                               'dc:description': 'Accession Number. '
                                                 'The 20-character '
                                                 'string formed '
                                                 'from the 18-digit '
                                                 'number assigned '
                                                 'by the Commission '
                                                 'to each EDGAR '
                                                 'submission.',
                               'name': 'adsh',
                               'required': 'true',
                               'titles': ['Accession number']},
                              {'datatype': {'base': 'string',
                                            'maxLength': 255},
                               'dc:description': 'The unique '
                                                 'identifier (name) '
                                                 'for a tag in a '
                                                 'specific taxonomy '
                                                 'release.',
                               'name': 'tag',
                               'required': 'true',
                               'titles': ['Localname']},
                              {'datatype': {'base': 'string',
                                            'maxLength': 20},
                               'dc:description': 'For a standard '
                                                 'tag, an '
                                                 'identifier for '
                                                 'the taxonomy; '
                                                 'otherwise the '
                                                 'accession number '
                                                 'where the tag was '
                                                 'defined. For '
                                                 'example, '
                                                 '"invest/2013" '
                                                 'indicates that '
                                                 'the tag is '
                                                 'defined in the '
                                                 '2013 INVEST '
                                                 'taxonomy.',
                               'name': 'version',
                               'required': 'true',
```

```
 'titles': ['Namespace', 'Taxonomy']},
{'datatype': {'base': 'string',
              'maxLength': 8,
              'minLength': 8},
 'dc:description': 'The end date for '
                   'the data value, '
                   'rounded to the '
                   'nearest month '
                   'end.',
 'name': 'ddate',
 'required': 'true',
 'titles': ['Data Date']},
{'datatype': {'base': 'decimal',
              'minInclusive': 0},
 'dc:description': 'The count of the '
                   'number of '
                   'quarters '
                   'represented by '
                   'the data value, '
                   'rounded to the '
                   'nearest whole '
                   'number. A point '
                   'in time value is '
                   'represented by 0.',
 'name': 'qtrs',
 'required': 'true',
 'titles': ['Quarters']},
{'datatype': {'base': 'decimal',
              'maxInclusive': 32767,
              'minInclusive': -32768},
 'dc:description': 'A positive '
                   'integer to '
                   'distinguish '
                   'different '
                   'reported facts '
                   'that otherwise '
                   'would have the '
                   'same primary key. '
                   'For most '
                   'purposes, data '
                   'with iprx greater '
                   'than 1 are not '
                   'needed. The '
                   'priority for the '
                   'fact based on '
                   'higher precision, '
                   'closeness of the '
                   'end date to a '
```

```
                    'month end, and '
                    'closeness of the '
                    'duration to a '
                    'multiple of three '
                    'months. See '
                    'fields dcml, durp '
                    'and datp below.',
 'name': 'iprx',
 'titles': ['Fact Preference',
            'Preferred Fact Sort '
            'Key']},
{'datatype': {'base': 'string',
              'maxLength': 5},
 'dc:description': 'The ISO language '
                   'code of the fact '
                   'content.',
 'name': 'lang',
 'titles': ['Language']},
{'datatype': {'base': 'decimal',
              'maxInclusive': 32767,
              'minInclusive': -32768},
 'dc:description': 'The value of the '
                   'fact "xml:lang" '
                   'attribute, en-US '
                   'represented by '
                   '32767, other "en" '
                   'dialects having '
                   'lower values, and '
                   'other languages '
                   'lower still.',
 'name': 'dcml',
 'titles': ['Language Preference',
            'Language Sort Key']},
{'datatype': {'base': 'decimal'},
 'dc:description': 'The difference '
                   'between the '
                   'reported fact '
                   'duration and the '
                   'quarter duration '
                   '(qtrs), expressed '
                   'as a fraction of '
                   '1. For example, a '
                   'fact with '
                   'duration of 120 '
                   'days rounded to a '
                   '91-day quarter '
                   'has a durp value '
                   'of 29/91 = '
```

```
                  '+0.3187.',
 'name': 'durp',
 'titles': ['Duration Preference']},
{'datatype': {'base': 'decimal'},
 'dc:description': 'The difference '
                   'between the '
                   'reported fact '
                   'date and the '
                   'month-end rounded '
                   'date (ddate), '
                   'expressed as a '
                   'fraction of 1. '
                   'For example, a '
                   'fact reported for '
                   '29/Dec, with '
                   'ddate rounded to '
                   '31/Dec, has a '
                   'datp value of '
                   'minus 2/31 = '
                   '-0.0645.',
 'name': 'datp',
 'titles': ['Date Preference']},
{'datatype': {'base': 'string',
              'maxLength': 34},
 'dc:description': 'The 32-byte '
                   'hexadecimal key '
                   'for the '
                   'dimensional '
                   'information in '
                   'the DIM data set.',
 'name': 'dimh',
 'titles': ['Dimension Hash']},
{'datatype': {'base': 'decimal',
              'minInclusive': 0},
 'dc:description': 'Small integer '
                   'representing the '
                   'number of '
                   'dimensions, '
                   'useful for '
                   'sorting. Note '
                   'that this value '
                   'is function of '
                   'the dimension '
                   'segments.',
 'name': 'dimn',
 'required': 'true',
 'titles': ['Number of Dimensions']},
{'datatype': {'base': 'string',
```

```
                'maxLength': 256},
 'dc:description': 'If specified, '
                   'indicates a '
                   'specific '
                   'co-registrant, '
                   'the parent '
                   'company, or other '
                   'entity (e.g., '
                   'guarantor). NULL '
                   'indicates the '
                   'consolidated '
                   'entity. Note that '
                   'this value is a '
                   'function of the '
                   'dimension '
                   'segments.',
 'name': 'coreg',
 'titles': ['Coregistrant']},
{'datatype': {'base': 'decimal',
              'maxInclusive': 1,
              'minInclusive': 0},
 'dc:description': 'Flag indicating '
                   'whether the value '
                   'has had tags '
                   'removed.',
 'name': 'escaped',
 'required': 'true',
 'titles': []},
{'datatype': {'base': 'decimal',
              'minInclusive': 0},
 'dc:description': 'Number of bytes '
                   'in the original, '
                   'unprocessed '
                   'value. Zero '
                   'indicates a NULL '
                   'value.',
 'name': 'srclen',
 'required': 'true',
 'titles': ['Source Length']},
{'datatype': {'base': 'decimal',
              'minInclusive': 0},
 'dc:description': 'The original '
                   'length of the '
                   'whitespace '
                   'normalized value, '
                   'which may have '
                   'been greater than '
                   '8192.',
```

```
 'name': 'txtlen',
 'required': 'true',
 'titles': ['Text Length']},
{'datatype': {'base': 'string',
              'maxLength': 512},
 'dc:description': 'The plain text of '
                   'any superscripted '
                   'footnotes on the '
                   'value, as shown '
                   'on the page, '
                   'truncated to 512 '
                   'characters, or if '
                   'there is no '
                   'footnote, then '
                   'this field will '
                   'be blank.',
 'name': 'footnote',
 'titles': ['Footnote Text']},
{'datatype': {'base': 'decimal',
              'minInclusive': 0},
 'dc:description': 'Number of bytes '
                   'in the plain text '
                   'of the footnote '
                   'prior to '
                   'truncation.',
 'name': 'footlen',
 'required': 'true',
 'titles': ['Footnote Length']},
{'datatype': {'base': 'string',
              'maxLength': 255},
 'dc:description': 'The value of the '
                   'contextRef '
                   'attribute in the '
                   'source XBRL '
                   'document, which '
                   'can be used to '
                   'recover the '
                   'original HTML '
                   'tagging if '
                   'desired.',
 'name': 'context',
 'titles': ['Context Ref']},
{'datatype': {'base': 'string',
              'maxLength': 2048},
 'dc:description': 'The value, with '
                   'all whitespace '
                   'normalized, that '
                   'is, all sequences '
```

```
                                                          'of line feeds, '
                                                          'carriage returns, '
                                                          'tabs, '
                                                          'non-breaking '
                                                          'spaces, and '
                                                          'spaces having '
                                                          'been collapsed to '
                                                          'a single space, '
                                                          'and no leading or '
                                                          'trailing spaces. '
                                                          'Escaped XML that '
                                                          'appears in EDGAR '
                                                          '"Text Block" tags '
                                                          'is processed to '
                                                          'remove all '
                                                          'mark-up '
                                                          '(comments, '
                                                          'processing '
                                                          'instructions, '
                                                          'elements, '
                                                          'attributes). The '
                                                          'value is '
                                                          'truncated to a '
                                                          'maximum number of '
                                                          'bytes. The '
                                                          'resulting text is '
                                                          'not intended for '
                                                          'end user display '
                                                          'but only for text '
                                                          'analysis '
                                                          'applications.',
                                               'name': 'value',
                                               'titles': []}],
                                 'foreignKeys': [{'columnReference': 'adsh',
                                                  'reference': {'columnReference':
'adsh',
                                                                'resource':
'sub.tsv'}},
                                                 {'columnReference': 'dimh',
                                                  'reference': {'columnReference':
'dimh',
                                                                'resource':
'https://wwww.sec.gov/files2018q3.zip#path=dim.tsv'}},
                                                 {'columnReference': ['tag',
                                                                      'version'],
                                                  'reference': {'columnReference':
['tag',
'version'],
```

```
                                                    'resource':
'https://wwww.sec.gov/files2018q3.zip#path=tag.tsv'}}],
                            'primaryKey': ['adsh',
                                           'tag',
                                           'version',
                                           'ddate',
                                           'qtrs',
                                           'dimh',
                                           'iprx']},
             'url': 'txt.tsv'},
            {'tableSchema': {'aboutUrl': 'readme.htm',
                             'columns': [{'datatype': {'base': 'string',
                                                       'maxLength': 20,
                                                       'minLength': 20},
                                          'dc:description': 'Accession Number. '
                                                            'The 20-character '
                                                            'string formed '
                                                            'from the 18-digit '
                                                            'number assigned '
                                                            'by the Commission '
                                                            'to each EDGAR '
                                                            'submission.',
                                          'name': 'adsh',
                                          'required': 'true',
                                          'titles': ['Accession Number']},
                                         {'datatype': {'base': 'decimal',
                                                       'minInclusive': 0},
                                          'dc:description': 'Represents the '
                                                            'report grouping. '
                                                            'The numeric value '
                                                            'refers to the "R '
                                                            'file" as computed '
                                                            'by the renderer '
                                                            'and posted on the '
                                                            'EDGAR website. '
                                                            'Note that in some '
                                                            'situations the '
                                                            'numbers skip.',
                                          'name': 'report',
                                          'required': 'true',
                                          'titles': ['Report Number']},
                                         {'datatype': {'base': 'string',
                                                       'maxLength': 1},
                                          'dc:description': 'The type of '
                                                            'interactive data '
                                                            'file rendered on '
                                                            'the EDGAR '
                                                            'website, H = .htm '
```

```
                   'file, X = .xml '
                   'file.',
  'name': 'rfile',
  'required': 'true',
  'titles': ['Report File Type']},
 {'datatype': {'base': 'string',
               'maxLength': 2},
  'dc:description': 'If available, one '
                   'of the menu '
                   'categories as '
                   'computed by the '
                   'renderer: '
                   'C=Cover, '
                   'S=Statements, '
                   'N=Notes, '
                   'P=Policies, '
                   'T=Tables, '
                   'D=Details, '
                   'O=Other, and '
                   'U=Uncategorized.',
  'name': 'menucat',
  'titles': ['Menu Category']},
 {'datatype': {'base': 'string',
               'maxLength': 512},
  'dc:description': 'The portion of '
                   'the long name '
                   'used in the '
                   'renderer menu.',
  'name': 'shortname',
  'titles': ['Short Name']},
 {'datatype': {'base': 'string',
               'maxLength': 512},
  'dc:description': 'The '
                   'space-normalized '
                   'text of the XBRL '
                   'link "definition" '
                   'element content.',
  'name': 'longname',
  'titles': ['Long Name']},
 {'datatype': {'base': 'string',
               'maxLength': 255},
  'dc:description': 'The XBRL '
                   '"roleuri" of the '
                   'role.',
  'name': 'roleuri',
  'titles': ['Role URI']},
 {'datatype': {'base': 'string',
               'maxLength': 255},
```

```
                                        'dc:description': 'The XBRL roleuri '
                                                          'of a role for '
                                                          'which this role '
                                                          'has a matching '
                                                          'shortname prefix '
                                                          'and a higher '
                                                          'level menu '
                                                          'category, as '
                                                          'computed by the '
                                                          'renderer.',
                                        'name': 'parentroleuri',
                                        'titles': ['Parent Role URI']},
                                       {'datatype': {'base': 'decimal',
                                                     'minInclusive': 0},
                                        'dc:description': 'The value of the '
                                                          'report field for '
                                                          'the role where '
                                                          'roleuri equals '
                                                          'this '
                                                          'parentroleuri.',
                                        'name': 'parentreport',
                                        'titles': ['Parent Report']},
                                       {'datatype': {'base': 'decimal',
                                                     'maxInclusive': 32767,
                                                     'minInclusive': 0},
                                        'dc:description': 'The highest '
                                                          'ancestor report '
                                                          'reachable by '
                                                          'following '
                                                          'parentreport '
                                                          'relationships. A '
                                                          'note (menucat = '
                                                          'N) is its own '
                                                          'ultimate parent.',
                                        'name': 'ultparentrpt',
                                        'titles': ['Ultimate Parent']}],
                            'foreignKeys': [{'columnReference': 'adsh',
                                             'reference': {'columnReference':
'adsh',
                                                           'resource':
'sub.tsv'}}],
                            'primaryKey': ['adsh', 'report']},
              'url': 'ren.tsv'},
             {'tableSchema': {'aboutUrl': 'readme.htm',
                              'columns': [{'datatype': {'base': 'string',
                                                        'maxLength': 20,
                                                        'minLength': 20},
                                           'dc:description': 'Accession Number. '
```

```
                    'The 20-character '
                    'string formed '
                    'from the 18-digit '
                    'number assigned '
                    'by the Commission '
                    'to each EDGAR '
                    'submission.',
  'name': 'adsh',
  'required': 'true',
  'titles': ['Accession Number']},
 {'datatype': {'base': 'decimal',
               'minInclusive': 0},
  'dc:description': 'Represents the '
                    'report grouping. '
                    'The numeric value '
                    'refers to the "R '
                    'file" as computed '
                    'by the renderer '
                    'and posted on the '
                    'EDGAR website. '
                    'Note that in some '
                    'situations the '
                    'numbers skip.',
  'name': 'report',
  'required': 'true',
  'titles': []},
 {'datatype': {'base': 'decimal',
               'minInclusive': 0},
  'dc:description': 'Represents the '
                    "tag's "
                    'presentation line '
                    'order for a given '
                    'report. Together '
                    'with the '
                    'statement and '
                    'report field, '
                    'presentation '
                    'location, order '
                    'and grouping can '
                    'be derived.',
  'name': 'line',
  'required': 'true',
  'titles': []},
 {'datatype': {'base': 'string',
               'maxLength': 2},
  'dc:description': 'The financial '
                    'statement '
                    'location to which '
```

```
                    'the value of the '
                    '"report" field '
                    'pertains.',
  'name': 'stmt',
  'titles': ['Statement']},
 {'datatype': {'base': 'decimal',
               'maxInclusive': 1,
               'minInclusive': 0},
  'dc:description': '1 indicates that '
                    'the value was '
                    'presented '
                    '"parenthetically" '
                    'instead of in '
                    'fields within the '
                    'financial '
                    'statements. For '
                    'example: '
                    'Receivables (net '
                    'of allowance for '
                    'bad debts of USD '
                    '200 in 2012) USD '
                    '700',
  'name': 'inpth',
  'required': 'true',
  'titles': ['Parenthentical']},
 {'datatype': {'base': 'string',
               'maxLength': 256},
  'dc:description': 'The tag chosen by '
                    'the filer for '
                    'this line item.',
  'name': 'tag',
  'required': 'true',
  'titles': ['Localname']},
 {'datatype': {'base': 'string',
               'maxLength': 20},
  'dc:description': 'The taxonomy '
                    'identifier if the '
                    'tag is a standard '
                    'tag, otherwise '
                    'adsh.',
  'name': 'version',
  'required': 'true',
  'titles': ['Namespace', 'Taxonomy']},
 {'datatype': {'base': 'string',
               'maxLength': 50},
  'dc:description': 'The XBRL link '
                    '"role" of the '
                    'preferred label, '
```

```
                                                  'using only the '
                                                  'portion of the '
                                                  'role URI after '
                                                  'the last "/".',
                                   'name': 'prole',
                                   'titles': ['Preferred Role']},
                                  {'datatype': {'base': 'string',
                                                'maxLength': 512},
                                   'dc:description': 'The text '
                                                  'presented on the '
                                                  'line item, also '
                                                  'known as a '
                                                  '"preferred" '
                                                  'label.',
                                   'name': 'plabel',
                                   'titles': ['Label']},
                                  {'datatype': {'base': 'decimal',
                                                'maxInclusive': 1,
                                                'minInclusive': 0},
                                   'dc:description': 'Flag to indicate '
                                                  'whether the prole '
                                                  'is treated as '
                                                  'negating by the '
                                                  'renderer.',
                                   'name': 'negating',
                                   'required': 'true',
                                   'titles': []}],
                     'foreignKeys': [{'columnReference': ['adsh',
                                                          'report'],
                                      'reference': {'columnReference':
['adsh',
'report'],
                                                    'resource':
'ren.tsv'}},
                                     {'columnReference': ['tag',
                                                          'version'],
                                      'reference': {'columnReference':
['tag',
'version'],
                                                    'resource':
'tag.tsv'}}],
                     'primaryKey': ['adsh', 'report', 'line']},
          'url': 'pre.tsv'},
         {'tableSchema': {'aboutUrl': 'readme.htm',
                     'columns': [{'datatype': {'base': 'string',
                                               'maxLength': 20,
                                               'minLength': 20},
                                  'dc:description': 'Accession Number. '
```

```
                    'The 20-character '
                    'string formed '
                    'from the 18-digit '
                    'number assigned '
                    'by the Commission '
                    'to each EDGAR '
                    'submission.',
  'name': 'adsh',
  'required': 'true',
  'titles': ['Accession Number']},
 {'datatype': {'base': 'decimal',
               'maxInclusive': 255,
               'minInclusive': 0},
  'dc:description': 'Sequential number '
                    'for grouping arcs '
                    'in a submission.',
  'name': 'grp',
  'required': 'true',
  'titles': ['Group']},
 {'datatype': {'base': 'decimal',
               'minInclusive': 255},
  'dc:description': 'Sequential number '
                    'for arcs within a '
                    'group in a '
                    'submission.',
  'name': 'arc',
  'required': 'true',
  'titles': []},
 {'datatype': {'base': 'decimal',
               'maxInclusive': 1,
               'minInclusive': 0},
  'dc:description': 'Indicates a '
                    'weight of -1 '
                    '(TRUE if the arc '
                    'is negative), but '
                    'typically +1 '
                    '(FALSE).',
  'name': 'negative',
  'required': 'true',
  'titles': ['Negative Weight']},
 {'datatype': {'base': 'string',
               'maxLength': 256},
  'dc:description': 'The tag for the '
                    'parent of the arc',
  'name': 'ptag',
  'required': 'true',
  'titles': ['Parent Tag']},
 {'datatype': {'base': 'string',
```

```
                                                  'maxLength': 20},
                                   'dc:description': 'The version of '
                                                     'the tag for the '
                                                     'parent of the arc',
                                   'name': 'pversion',
                                   'required': 'true',
                                   'titles': ['Parent Namespace']},
                                  {'datatype': {'base': 'string',
                                                  'maxLength': 255},
                                   'dc:description': 'The tag for the '
                                                     'child of the arc',
                                   'name': 'ctag',
                                   'required': 'true',
                                   'titles': ['Child Tag']},
                                  {'datatype': {'base': 'string',
                                                  'maxLength': 20},
                                   'dc:description': 'The version of '
                                                     'the tag for the '
                                                     'child of the arc',
                                   'name': 'cversion',
                                   'required': 'true',
                                   'titles': ['Child Namespace']}],
                     'foreignKeys': [{'columnReference': 'adsh',
                                      'reference': {'columnReference':
'adsh',
                                                    'resource':
'sub.tsv'}},
                                     {'columnReference': ['ptag',
                                                          'pversion'],
                                      'reference': {'columnReference':
['tag',
'version'],
                                                    'resource':
'tag.tsv'}},
                                     {'columnReference': ['ctag',
                                                          'cversion'],
                                      'reference': {'columnReference':
['tag',
'version'],
                                                    'resource':
'tag.tsv'}}],
                     'primaryKey': ['adsh', 'grp', 'arc']},
          'url': 'cal.tsv'}]}
```

0.4 Data Organization

For each quarter, the FSN data is organized into eight file sets that contain information about submissions, numbers, taxonomy tags, presentation, and more. Each dataset consists of rows and

fields and is provided as a tab-delimited text file:

File	Dataset	Description
SUB	Submission	Identifies each XBRL submission by company, form, date, etc
TAG	Tag	Defines and explains each taxonomy tag
DIM	Dimension	Adds detail to numeric and plain text data
NUM	Numeric	One row for each distinct data point in filing
TXT	Plain Text	Contains all non-numeric XBRL fields
REN	Rendering	Information for rendering on SEC website
PRE	Presentation	Detail on tag and number presentation in primary statements
CAL	Calculation	Shows arithmetic relationships among tags

0.5 Submission Data

The latest submission file contains around 6,500 entries.

[9]:
```
sub = pd.read_parquet(data_path / '2018_3' / 'parquet' / 'sub.parquet')
sub.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 6492 entries, 0 to 6491
Data columns (total 40 columns):
adsh            6492 non-null object
cik             6492 non-null int64
name            6492 non-null object
sic             6490 non-null float64
countryba       6481 non-null object
stprba          5899 non-null object
cityba          6481 non-null object
zipba           6477 non-null object
bas1            6481 non-null object
bas2            2804 non-null object
baph            6481 non-null object
countryma       6447 non-null object
stprma          5905 non-null object
cityma          6447 non-null object
zipma           6446 non-null object
mas1            6447 non-null object
mas2            2761 non-null object
countryinc      5935 non-null object
stprinc         5631 non-null object
ein             6491 non-null float64
former          3618 non-null object
changed         3618 non-null float64
afs             6415 non-null object
wksi            6492 non-null int64
fye             6489 non-null float64
```

```
form           6492 non-null object
period         6492 non-null int64
fy             6492 non-null int64
fp             6492 non-null object
filed          6492 non-null int64
accepted       6492 non-null object
prevrpt        6492 non-null int64
detail         6492 non-null int64
instance       6492 non-null object
nciks          6492 non-null int64
aciks          130 non-null object
pubfloatusd    639 non-null float64
floatdate      640 non-null float64
floataxis      3 non-null object
floatmems      4 non-null float64
dtypes: float64(7), int64(8), object(25)
memory usage: 2.0+ MB
```

0.5.1 Get AAPL submission

The submission dataset contains the unique identifiers required to retrieve the filings: the Central Index Key (CIK) and the Accession Number (adsh). The following shows some of the information about Apple's 2018Q1 10-Q filing:

[10]:
```
name = 'APPLE INC'
apple = sub[sub.name == name].T.dropna().squeeze()
key_cols = ['name', 'adsh', 'cik', 'name', 'sic', 'countryba', 'stprba',
            'cityba', 'zipba', 'bas1', 'form', 'period', 'fy', 'fp', 'filed']
apple.loc[key_cols]
```

[10]:
```
name                    APPLE INC
adsh          0000320193-18-000100
cik                         320193
name                    APPLE INC
sic                          3571
countryba                      US
stprba                         CA
cityba                  CUPERTINO
zipba                       95014
bas1           ONE APPLE PARK WAY
form                         10-Q
period                   20180630
fy                           2018
fp                             Q3
filed                    20180801
Name: 386, dtype: object
```

0.6 Build AAPL fundamentals dataset

Using the central index key, we can identify all historical quarterly filings available for Apple, and combine this information to obtain 26 Forms 10-Q and nine annual Forms 10-K.

0.6.1 Get filings

[9]:
```
aapl_subs = pd.DataFrame()
for sub in data_path.glob('**/sub.parquet'):
    sub = pd.read_parquet(sub)
    aapl_sub = sub[(sub.cik.astype(int) == apple.cik) & (sub.form.isin(['10-Q',␣
 ↪'10-K']))]
    aapl_subs = pd.concat([aapl_subs, aapl_sub])
```

We find 15 quarterly 10-Q and 4 annual 10-K reports:

[10]:
```
aapl_subs.form.value_counts()
```

[10]:
```
10-Q    15
10-K     4
Name: form, dtype: int64
```

0.6.2 Get numerical filing data

With the Accession Number for each filing, we can now rely on the taxonomies to select the appropriate XBRL tags (listed in the TAG file) from the NUM and TXT files to obtain the numerical or textual/footnote data points of interest.

First, let's extract all numerical data available from the 19 Apple filings:

[11]:
```
aapl_nums = pd.DataFrame()
for num in data_path.glob('**/num.parquet'):
    num = pd.read_parquet(num).drop('dimh', axis=1)
    aapl_num = num[num.adsh.isin(aapl_subs.adsh)]
    print(len(aapl_num))
    aapl_nums = pd.concat([aapl_nums, aapl_num])
aapl_nums.ddate = pd.to_datetime(aapl_nums.ddate, format='%Y%m%d')
aapl_nums.to_parquet(data_path / 'aapl_nums.parquet')
```

```
738
1345
707
961
1001
905
951
1277
937
751
```

```
923
793
1364
1271
682
805
942
919
952
```

In total, the nine years of filing history provide us with over 18,000 numerical values for AAPL.

[12]:
```
aapl_nums.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 18224 entries, 84837 to 5467444
Data columns (total 15 columns):
adsh        18224 non-null object
tag         18224 non-null object
version     18224 non-null object
ddate       18224 non-null datetime64[ns]
qtrs        18224 non-null int64
uom         18224 non-null object
iprx        18224 non-null float64
value       18176 non-null float64
footnote    68 non-null object
footlen     18224 non-null int64
dimn        18224 non-null int64
coreg       0 non-null object
durp        18224 non-null float64
datp        18224 non-null float64
dcml        18224 non-null float64
dtypes: datetime64[ns](1), float64(5), int64(3), object(6)
memory usage: 2.2+ MB
```

0.7 Create P/E Ratio from EPS and stock price data

We can select a useful field, such as Earnings per Diluted Share (EPS), that we can combine with market data to calculate the popular Price/Earnings (P/E) valuation ratio.

[15]:
```
stock_split = 7
split_date = pd.to_datetime('20140604')
split_date
```

[15]:
```
Timestamp('2014-06-04 00:00:00')
```

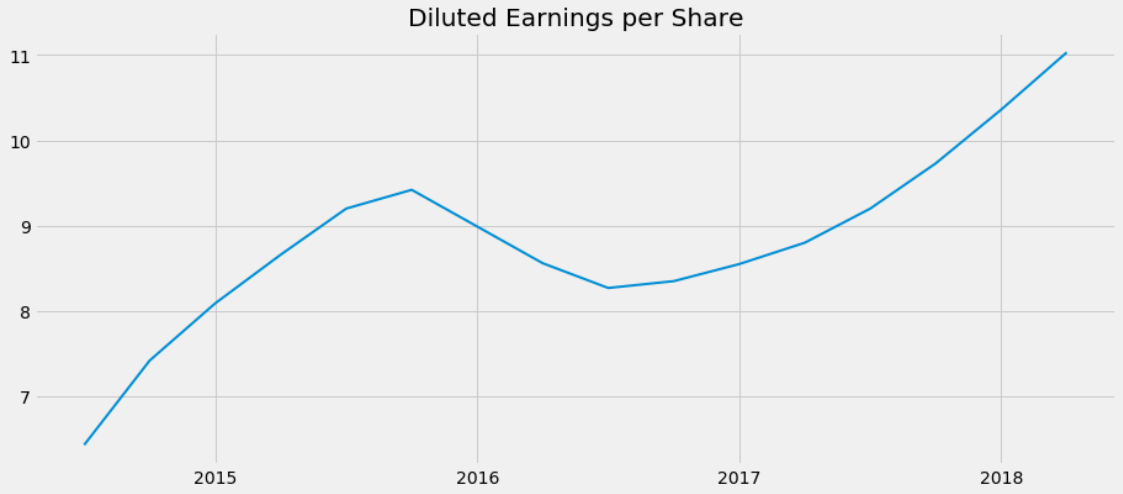
We do need to take into account, however, that Apple split its stock 7:1 on June 4, 2014, and Adjusted Earnings per Share before the split to make earnings comparable, as illustrated in the following code block:

[16]:
```
# Filter by tag; keep only values measuring 1 quarter
eps = aapl_nums[(aapl_nums.tag == 'EarningsPerShareDiluted')
                & (aapl_nums.qtrs == 1)].drop('tag', axis=1)

# Keep only most recent data point from each filing
eps = eps.groupby('adsh').apply(lambda x: x.nlargest(n=1, columns=['ddate']))

# Adjust earnings prior to stock split downward
eps.loc[eps.ddate < split_date,'value'] = eps.loc[eps.ddate < split_date,␣
 ↪'value'].div(7)
eps = eps[['ddate', 'value']].set_index('ddate').squeeze().sort_index()
eps = eps.rolling(4,min_periods=4).sum().dropna()
```

[17]:
```
eps.plot(lw=2, figsize=(14, 6), title='Diluted Earnings per Share')
plt.xlabel('')
plt.savefig('diluted eps', dps=300);
```



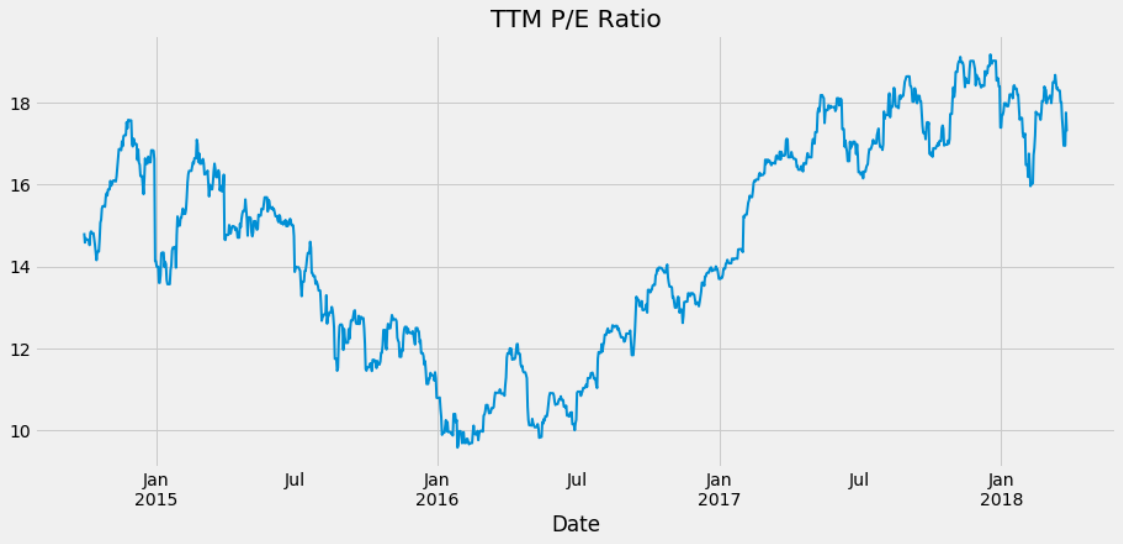
[18]:
```
symbol = 'AAPL.US'

aapl_stock = (web.
              DataReader(symbol, 'quandl', start=eps.index.min())
              .resample('D')
              .last()
             .loc['2014':eps.index.max()])
aapl_stock.info()
```

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 1275 entries, 2014-09-30 to 2018-03-27
Freq: D
Data columns (total 12 columns):
```

```
Open           877 non-null float64
High           877 non-null float64
Low            877 non-null float64
Close          877 non-null float64
Volume         877 non-null float64
ExDividend     877 non-null float64
SplitRatio     877 non-null float64
AdjOpen        877 non-null float64
AdjHigh        877 non-null float64
AdjLow         877 non-null float64
AdjClose       877 non-null float64
AdjVolume      877 non-null float64
dtypes: float64(12)
memory usage: 129.5 KB
```

[19]:
```
pe = aapl_stock.AdjClose.to_frame('price').join(eps.to_frame('eps'))
pe = pe.fillna(method='ffill').dropna()
pe['P/E Ratio'] = pe.price.div(pe.eps)
pe['P/E Ratio'].plot(lw=2, figsize=(14, 6), title='TTM P/E Ratio');
```

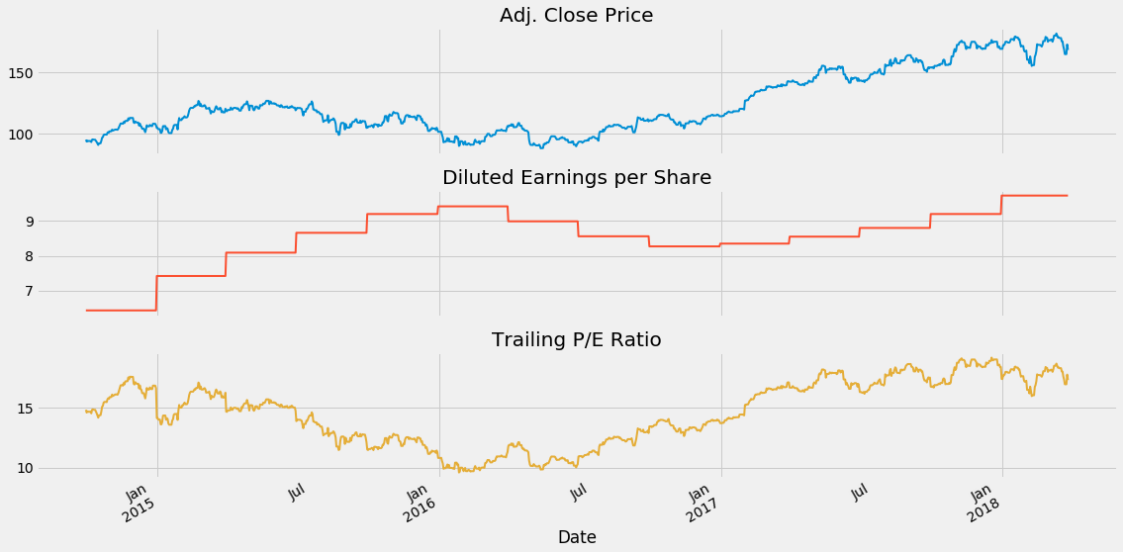


[20]:
```
pe.info()
```

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 1275 entries, 2014-09-30 to 2018-03-27
Freq: D
Data columns (total 3 columns):
price          1275 non-null float64
eps            1275 non-null float64
P/E Ratio      1275 non-null float64
dtypes: float64(3)
```

```
memory usage: 39.8 KB
```

[21]:
```
axes = pe.plot(subplots=True, figsize=(16,8), legend=False, lw=2)
axes[0].set_title('Adj. Close Price')
axes[1].set_title('Diluted Earnings per Share')
axes[2].set_title('Trailing P/E Ratio')
plt.tight_layout();
```



0.8 Explore Additional Fields

The field `tag` references values defined in the taxonomy:

[22]:
```
aapl_nums.tag.value_counts()
```

[22]:
```
DebtInstrumentInterestRateEffectivePercentage
573
CashAndCashEquivalentsAtCarryingValue
570
SalesRevenueNet
544
AvailableForSaleSecuritiesNoncurrent
532
AvailableForSaleSecurities
532
AvailableForSaleSecuritiesCurrent
532
AvailableForSaleSecuritiesAmortizedCost
532
AvailableForSaleSecuritiesAccumulatedGrossUnrealizedLossBeforeTax
```

```
476
AvailableForSaleSecuritiesAccumulatedGrossUnrealizedGainBeforeTax
476
OperatingIncomeLoss
447
SeniorNotes
374
DerivativeInstrumentsGainLossRecognizedInOtherComprehensiveIncomeEffectivePortio
nNet                                306
DebtInstrumentCarryingAmount
295
DebtInstrumentInterestRateStatedPercentage
287
StockRepurchasedAndRetiredDuringPeriodShares
255
AllocatedShareBasedCompensationExpense
231
DerivativeFairValueOfDerivativeAsset
204
DerivativeInstrumentsGainLossReclassifiedFromAccumulatedOCIIntoIncomeEffectivePo
rtionNet                            201
DerivativeFairValueOfDerivativeLiability
180
StockRepurchasedAndRetiredDuringPeriodValue
175
ConcentrationRiskPercentage1
172
StockholdersEquity
168
CommonStockDividendsPerShareDeclared
159
PropertyPlantAndEquipmentGross
152
DerivativeNotionalAmount
142
NonoperatingIncomeExpense
134
NetIncomeLoss
130
PaymentsOfDividends
125
OtherComprehensiveIncomeLossReclassificationAdjustmentFromAOCIOnDerivativesBefor
eTax                                120
IncomeLossFromContinuingOperationsBeforeIncomeTaxesExtraordinaryItemsNoncontroll
ingInterest                         118
                                    …
UnrecognizedTaxBenefitsPeriodIncreaseDecrease
```

```
2
RepaymentsOfAssumedDebt
2
ShareBasedCompensationArrangementByShareBasedPaymentAwardOptionsExercisableWeigh
tedAverageExercisePrice                1
SharebasedCompensationArrangementBySharebasedPaymentAwardOptionsExercisableIntri
nsicValue1                             1
ProceedsFromRepaymentsOfShortTermDebt
1
StockIssuedDuringPeriodSharesStockOptionsExercised
1
ShareBasedCompensationArrangementByShareBasedPaymentAwardOptionsExpectedToVestIn
trinsicValueAtPeriodEnd                1
LossContingencySubsidiariesImpactedNumber
1
UnrecordedUnconditionalPurchaseObligationBalanceOnThirdAnniversary
1
UnrecordedUnconditionalPurchaseObligationBalanceOnFirstAnniversary
1
ShareBasedCompensationArrangementsByShareBasedPaymentAwardOptionsExercisesInPeri
odWeightedAverageExercisePrice         1
IncomeTaxReconciliationTaxSettlementsDomestic
1
ShareBasedCompensationArrangementsByShareBasedPaymentAwardOptionsGrantsInPeriodW
eightedAverageExercisePrice            1
RestrictedInvestmentsIncreaseDecrease
1
UnrecordedUnconditionalPurchaseObligationBalanceOnFourthAnniversary
1
ShareBasedCompensationArrangementByShareBasedPaymentAwardOptionsExpectedToVestWe
ightedAverageExercisePrice             1
SharebasedCompensationArrangementbySharebasedPaymentAwardOptionsNumberofSharesof
CommonSharesAwardedUponSettlement      1
ShareBasedCompensationArrangementByShareBasedPaymentAwardOptionsGrantsInPeriod
1
TaxCutsAndJobsActOf2017MeasurementPeriodAdjustmentIncomeTaxExpenseBenefit
1
PreferredStockSharesAuthorized
1
SalesRevenueServicesGross
1
ShareBasedCompensationArrangementByShareBasedPaymentAwardOptionsVestedAndExpecte
dToVestOutstandingNumber               1
ResultOfLegalProceedingsAwardUpHeld
1
UnrecordedUnconditionalPurchaseObligationBalanceOnFifthAnniversary
1
```

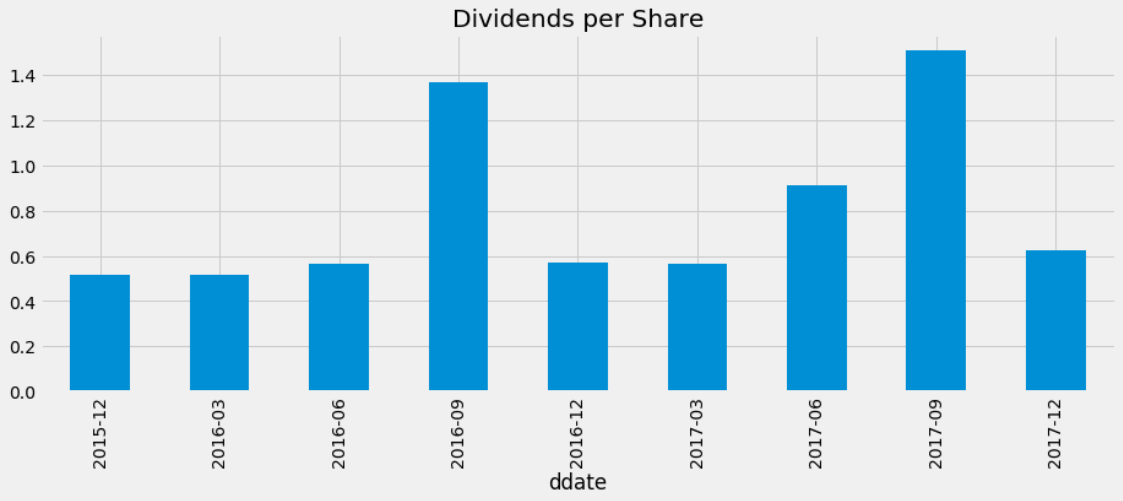
```
UnrecordedUnconditionalPurchaseObligationDueAfterFiveYears
1
UnrecordedUnconditionalPurchaseObligationBalanceOnSecondAnniversary
1
ShareBasedCompensationArrangementByShareBasedPaymentAwardOptionsExercisableNumbe
r                                                 1
ResultOfLegalProceedingsAdditionalAmountAwarded
1
DeferredTaxLiabilityRelatedToAmountsThatMayBeRepatriated
1
LossContingencyRangeOfPossibleLossMaximum
1
Name: tag, Length: 427, dtype: int64
```

We can select values of interest and track their value or use them as inputs to compute fundamental metrics like the Dividend/Share ratio.

0.8.1 Dividends per Share

[23]:
```
fields = ['EarningsPerShareDiluted',
          'PaymentsOfDividendsCommonStock',
          'WeightedAverageNumberOfDilutedSharesOutstanding',
          'OperatingIncomeLoss',
          'NetIncomeLoss',
          'GrossProfit']
```

[24]:
```
dividends = (aapl_nums
             .loc[aapl_nums.tag == 'PaymentsOfDividendsCommonStock', ['ddate',␣
 ↪'value']]
             .groupby('ddate')
             .mean())
shares = (aapl_nums
          .loc[aapl_nums.tag ==␣
 ↪'WeightedAverageNumberOfDilutedSharesOutstanding', ['ddate', 'value']]
          .drop_duplicates()
          .groupby('ddate')
          .mean())
df = dividends.div(shares).dropna()
ax = df.plot.bar(figsize=(14, 5), title='Dividends per Share', legend=False)
ax.xaxis.set_major_formatter(mticker.FixedFormatter(df.index.strftime('%Y-%m')))
```



0.9 Bonus: Textual Information

[15]:
```
txt = pd.read_parquet(data_path / '2016_2' / 'parquet' /  'txt.parquet')
```

AAPL's adsh is not avaialble in the txt file but you can obtain notes from the financial statesments here:

[17]:
```
txt.head()
```

[17]:

	adsh	tag	version	ddate	qtrs	iprx	lang	dcml	durp	datp	dimh	dimn	coreg	escaped	srclen	txtlen	footnote	footlen	context
0	0000014693-16-000160	AdvertisingCostsPolicyTextBlock	us-gaap/2015	20160430	4	0	en-US	32767	0.0	0.0	0x00000000	0	None	1	425	112	None	0	FD2016Q4YTD
1	0000014693-16-000160	AmendmentFlag	dei/2014	20160430	4	0	en-US	32767	0.0	0.0	0x00000000	0	None	0	5	5	None	0	FD2016Q4YTD
2	0000014693-16-000160	ComprehensiveIncomeNoteTextBlock	us-gaap/2015	20160430	4	0	en-US	32767	0.0	0.0	0x00000000	0	None	1	82857	2106	None	0	FD2016Q4YTD
3	0000014693-16-000160	EntityFilerCategory	dei/2014	20160430	4	0	en-US	32767	0.0	0.0	0x00000000	0	None	0	23	23	None	0	FD2016Q4YTD
4	0000014693-16-000160	ScheduleOfComprehensiveIncomeLossTableTextBlock	us-gaap/2015	20160430	4	0	en-US	32767	0.0	0.0	0x00000000	0	None	1	67007	1686	None	0	FD2016Q4YTD

```
                                                  value
0  Advertising costs. We expense the costs of adv…
1                                                  false
2  ACCUMULATED OTHER COMPREHENSIVE INCOME The fol…
3                            Large Accelerated Filer
4  The following table presents the components of…
```