Outline

The previous qparquet example showed pulling a full table in to memory from a parquet file.

This example expands on that to allow control of which columns are extracted.

Imports

```
//parquet library prints many warnings - ignore for this example
p)import warnings
p)warnings.filterwarnings("ignore")
```

```
//Import pandas, numpy, and pyarrow
p)import pandas as pd
p)import numpy as np
p)import pyarrow as pa
p)import pyarrow.parquet as pq
```

Find columns

Start by creating a function which extracts the list of columns in a parquet file.

```
p)def getColumnNames(file):
    return (pq.read_schema(file)).names
```

This can be called from python

```
p)print(getColumnNames('example.parquet'))
```

```
['one', 'two', 'three', '__index_level_0__']
```

Next we pull the function in to q

```
getColumnNames:.p.get`getColumnNames
getColumnNames["example.parquet"]`
```

```
"one"
"two"
```

```
    "three"
    "__index_level_0__"
```

Extract columns

The read_table function has an optional argument to only extract a sublist of columns from the parquet file.

We create a function which will return only the columns we choose.

The resulting data is converted to pandas and then to a dictionary for preperation to export to q

```
p)def getColumns(file, cols):
    table=pq.read_table('example.parquet', columns=cols)
    return (table.to_pandas()).to_dict('list')
```

Now you can control which columns return

```
p)print(getColumns('example.parquet',['one', 'two']))
```

```
{'one': [-1.0, 0.0, 2.5], 'two': ['foo', 'bar', 'baz']}
```

Next we pull the function in to q. The data is returned as a dictionary. To convert to a table use flip

```
getColumns:.p.get`getColumns
getColumns["example.parquet";("one";"two")]`
```

```
one| -1    0     2.5
two| "foo" "bar" "baz"
```

Extract only one column

A small wrapper around previsouly created function exposes this functionality

```
getColumn:{[file;column] first value getColumns[file;enlist column]`}
```

We can choose the exact needed column.

```
getColumn["example.parquet";"one"]
```

```
-1 0 2.5
```