Factor Research and FactSet’s Integration with MATLAB

Christopher Thomas
Content & Technology Solutions Sales Specialist

Joe Importico, CFA, FRM
Vice President, Quantitative Product Specialist
Agenda

- Introduction
  - FactSet Overview
  - FactSet Content
  - MATLAB Integration

- Creating a Factor Allocation Model
**Client Base**

- 2,525 clients
- 51,014 users
- 95% Client Retention

**About FactSet**

- Founded in 1978
- Public since 1996
- Dual listed on NYSE & NASDAQ (FDS)
- 6,399 employees worldwide
- FY 2013 was FactSet’s 17th consecutive year of positive earnings growth
Trusted by the Industry’s Best

+ Top media players trust FactSet as a reliable source for financial information

+ Every day, FactSet data is used by thousands of professionals – and cited in leading publications
FactSet DataFeed Solutions

- Fundamentals
- Estimates
- Ownership
- Benchmarks
- Economics
- Events & Transcripts
- People
- Private Equity
- M&A
- Pricing
- Debt Capital Structure
- Corporate Governance
- Corporate Activism
- Industry & Sector Classifications
- Entity Hierarchy
- Supply Chain
- Geographical Revenue
FactSet and MATLAB Integration

+ Access full scope and coverage of FactSet content
+ Eliminate the need to download/upload files manually
  + Outsource data management
+ Retrieve FactSet data in a MATLAB-friendly format
+ Take advantage of quick and easy setup
  + [www.factset.com/download/statlink](http://www.factset.com/download/statlink)
+ Seamless integration with other FactSet applications
Extract Data in MATLAB

`^=FF_ENTRPR_VAL_EBITDA_OPER(QTR,-20AY,0)`
MATLAB Request Output

```
>> segdata = FF_ExtractVectorFormula(factset, 'IBM', 'GE', 'FF_SEGMENT_RPT_DATA', 0, 0, 'SALES', 'BUS', 'SEG'), FF_SEGMENT_RPT_LABELS (ANN, 0, 0, BUS, 'SEG'))

Summary
---
Entities: 2 Items: 3 Dates: 0 Layout: IC
Orientation: None Rows: 15
- Item (1): Type: Entity Desc: Id
- Item (2): Type: Double Desc: ff_segment_rpt_data (Segment Report Data - A)
- Item (3): Type: String Desc: ff_segment_rpt_labels (Segment Report Labels - A)

>> segdata

segdata =

Id: ('IBM' 'IBM' 'IBM' 'IBM' 'IBM' 'GE' 'GE' 'GE' 'GE' 'GE' 'GE' 'GE' 'GE' 'GE')

ff_segment_rpt_data: [20131231 30551 25892 18396 14371 2022 20131231 42917 23777 21411 18186 16615 8313 6721 5873]

ff_segment_rpt_labels: (1x15 cell)

```
Help Menu and Documentation

**FF_EXTRACTALPHA_TESTING_SNAPSHOT**

The ExtractAlphaTestingSnapshot function provides access in MATLAB to data from AlphaTesting model results. Alpha Testing is a tool available in the FactSet workstation used to assess the relationship between one or more variables and subsequent returns over time. A subscription to Alpha Testing in FactSet is necessary to extract this data in MATLAB.

1. **Alpha Testing**

   The Alpha Testing application in FactSet is used to build models specifying the factors to test, the historical context, and customizing fractile assignments. After building and running a model, the data can be viewed in the FactSet workstation in overview charts, an overall report or in detailed reports for any specific fractile or time period. For a more comprehensive overview of Alpha Testing refer to FactSet Online Assistant page 13550.

2. **FF_EXTRACTALPHA_TESTING_SNAPSHOT in MATLAB**

   The following syntax is for extracting data using the ExtractAlphaTesting function in MATLAB:

   ```matlab
   data = FF_ExtractAlphaTestingSnapshot('factset', 'useStat', 'headers', 'model', 'report', 'items', 'resultType', 'date', 'security', 'reportSetting', 'reportSettingName', 'sortOrder', 'sortCol', {'param value'});
   ```
FactSet Plugin Integration Requirements

+ MATLAB Datafeed Toolbox®
+ MATLAB version 2007a or higher
+ FactSet subscription
Factor Modeling & Evaluating Results
Quantitative Workflow

**RESEARCH**

- **Setup**
  - Build and test universes and factors for backtesting.

- **Backtest**
  - Test whether your factors identify companies with superior performance. Use your factors to create alphas.

- **Simulate**
  - Run rules-based or optimizer-based simulations to see how potential strategies would have performed historically.

**CONSTRUCT**

- **Optimize**
  - Use your alphas and desired risk/return level to see what trades you should make to improve performance.

**EVALUATE**

- **Analyze**
  - Analyze performance and manage the risk of your portfolios.
Benefits of Backtesting on FactSet

+ Historical screening and data management
  + Historical CUSIP/SEDOL stitching across databases

+ Avoid look-head and survivorship bias
  + Point-in-Time databases
  + Lagging Fundamental Content

+ Integrate commercial and proprietary data

+ Seamless integration with other FactSet applications
  + Optimization/Simulation
  + PA
  + SPAR
  + Intraday Views
Factor Construction

+ Choose Factors
  + FCF Yield
  + Relative Strength Index
  + Accruals Ratio
  + FY1 EPS Growth

+ Import Factor Data into MATLAB
+ Standardize Data
  + Z-Score

+ Develop Factor Allocations
  + Run a series of cross-sectional regressions
  + Extract Factor Betas
Import Factor Data into MATLAB

Command Window

```
>> data.Identifier(1:10)
ans =
'17290810' '80898M10' '91301710' '65546510' '00130810' '02003910' '20911510' '55983010' '67622010' '82617010'
>> data.Company_Name(1:10)
ans =
Columns 1 through 8
'Cintas Corporation' 'SGMA Corporation' 'United Technology...' 'National City Corp.' 'The AES Corporation' 'Alltel Corp.' 'Consolidated Edi...' 'Lockheed Martin ...'
Columns 9 through 10
'Office Depot, Inc.' 'Siebel Systems Inc.'
>> data.PCF_Yield(1:10)
ans =
0.0033 0.0482 0.0481 -0.0065 -0.0354 0.0152 0.0461 0.0284 0.0312 0.0047
```
Standardized Data – Free Cash Flow Yield

```
>> zscore1 = (data.FCF_Yield-nanmean(data.FCF_Yield))/nanstd(data.FCF_Yield);
```

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>85</td>
<td>-0.0591</td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>-0.0336</td>
<td></td>
</tr>
<tr>
<td>87</td>
<td>-0.0603</td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>0.0454</td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>-0.0163</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>0.0253</td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>0.0027</td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>-0.0268</td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>-0.1177</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>-0.0217</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>-0.0453</td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>NaN</td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>-0.0305</td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>-0.0238</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>-0.0644</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>-0.0243</td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>-0.0633</td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>-0.0148</td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>0.0093</td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>-0.0298</td>
<td></td>
</tr>
</tbody>
</table>
Factor Betas

Command Window

```matlab
>> reg1.Coefficients.Estimate(2)
ans =
   -1.2454e-05
>> h=rc;
>> reg1 = fitlm(data.Universe_Returns,zscore1)
reg1 =

Linear regression model:
  y ~ 1 + x1

Estimated Coefficients:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>tStat</th>
<th>pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>2.6062e-06</td>
<td>0.0056648</td>
<td>0.0046007</td>
<td>0.99963</td>
</tr>
<tr>
<td>x1</td>
<td>-1.2454e-05</td>
<td>0.00027438</td>
<td>-0.045389</td>
<td>0.9638</td>
</tr>
</tbody>
</table>

Number of observations: 31047, Error degrees of freedom: 31045
Root Mean Squared Error: 1
R-squared: 6.47e-08, Adjusted R-Squared -3.13e-05
F-statistic vs. constant model: 0.00206, p-value = 0.964

>> reg1.Coefficients.Estimate(2)
ans =
   -1.2454e-05
```
Factor Construction Continued

+ Upload factor Betas to FactSet
+ Assign Betas (weights) to factors within Alpha Testing
Analyzing Performance

+ Can dynamically allocating amongst factors improve performance?
Risk and Return Profile

+ Can time-varying weights improve the risk and return profile?

![Factor Risk vs. Return Chart]

- Universe Return - 1 (Positive)
- Universe Return - 1 (Negative)
- FCF Yield
- Alpha Timing
- CF Accruals Ratio
- Alpha
- 14 Day RSI
- FY1 EPS Growth
FactSet Integration with MATLAB Workflow

FactSet -> Upload analyzed data -> MATLAB

MATLAB -> Upload analyzed data -> FactSet
Questions?

+ Christopher Thomas
  + CThomas@FactSet.com

+ Joe Importico, CFA, FRM
  + Jimportico@FactSet.com

Thank you.