MICROSOFT" EXCEL" FOR STOCK AND OPTION TRADERS

BUILD YOUR OWN ANALYTICAL TOOLS FOR HIGHER RETURNS



JEFFFAUGE IN OPTIONS TRADING

Vice President, Publisher: Tim Moore Associate Publisher and Director of Marketing: Amy Neidlinger Executive Editor: Jim Boyd Editorial Assistant: Pamela Boland Operations Manager: Gina Kanouse Senior Marketing Manager: Julie Phifer Publicity Manager: Laura Czaja Assistant Marketing Manager: Megan Colvin Cover Designer: Chuti Prasertsith Managing Editor: Kristy Hart Project Editor: Betsy Harris Copy Editor: Cheri Clark Proofreader: Kathy Ruiz Indexer: Erika Millen Senior Compositor: Gloria Schurick Manufacturing Buyer: Dan Uhrig © 2011 by Pearson Education, Inc.

Publishing as FT Press Upper Saddle River, New Jersey 07458

This book is sold with the understanding that neither the author nor the publisher is engaged in rendering legal, accounting, or other professional services or advice by publishing this book. Each individual situation is unique. Thus, if legal or financial advice or other expert assistance is required in a specific situation, the services of a competent professional should be sought to ensure that the situation has been evaluated carefully and appropriately. The author and the publisher disclaim any liability, loss, or risk resulting directly or indirectly, from the use or application of any of the contents of this book.

FT Press offers excellent discounts on this book when ordered in quantity for bulk purchases or special sales. For more information, please contact U.S. Corporate and Government Sales, 1-800-382-3419, corpsales@pearsontechgroup.com. For sales outside the U.S., please contact International Sales at international@pearson.com.

Company and product names mentioned herein are the trademarks or registered trademarks of their respective owners.

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

Printed in the United States of America

First Printing April 2011

ISBN-10: 0-13-713182-8 ISBN-13: 978-0-13-713182-2

Pearson Education LTD. Pearson Education Australia PTY, Limited. Pearson Education Singapore, Pte. Ltd. Pearson Education North Asia, Ltd. Pearson Education Canada, Ltd. Pearson Educación de Mexico, S.A. de C.V. Pearson Education—Japan

Pearson Education Malaysia, Pte. Ltd.

Library of Congress Cataloging-in-Publication Data

Augen, Jeffrey.

Microsoft Excel for stock and option traders : build your own analytical tools for higher returns / Jeffrey Augen.

p. cm.

ISBN 978-0-13-713182-2 (hbk. : alk. paper)

1. Investment analysis—Computer programs. 2. Investment analysis—Mathematical models. 3. Microsoft Excel (Computer file) I. Title.

HG4515.5.A94 2011

332.640285'554—dc22

To Lisa, who changed everything when she said: "Why don't you just calculate the integral between those two points and chart the value as it changes over time?" This page intentionally left blank

Contents

Preface	1
Chapter 1 Introduction—The Value of Information	7
The Struggle for a Statistical Edge	7
Fingerprinting the Market	12
Graphical Approaches to Discovering Price-Change Relationships	20
Focusing on a Statistical Anomaly	25
Capitalizing on Rare Events	53
Predicting Corrections	54
Brief Time Frames	
Summary	
Further Reading	
Endnotes	60
Chapter 2 The Basics	63
Spreadsheet Versus Database	63
Managing Date Formats	65
Aligning Records by Date	69
Decimal Date Conversion	91
Volatility Calculations	94
Descriptive Ratios	
Creating Summary Tables	116
Discovering Statistical Correlations	
Creating Trendlines	147

Summary	149
Further Reading	151
Endnotes	152
Chapter 3 Advanced Topics	153
Introduction	153
Time Frames	155
Building and Testing a Model	158
Sample Results	178
Index	

Acknowledgments

would like to thank the team that helped pull the book together. First must be Jim Boyd, who encour-Laged me to continue the project and always seems willing to explore new areas and concepts. This book would never have made it to print without advice and direction from Jim. Once again it was my pleasure to work with Betsy Harris, who always does a terrific job turning a rough manuscript into a polished, productionquality book. In that regard, I must also thank Cheri Clark, who carefully read every word and made corrections that put the finishing touch on the work. Finally, I'd like to acknowledge the important contributions of a friend-Robert Birnbaum. Over the past several months, Robert has helped shape my thinking about the statistical relevance of trends-ideas which surfaced in some of the key examples and continue to weigh heavily in my own investing.

This page intentionally left blank

About the Author

eff Augen, currently a private investor and writer, has spent more than a decade building a unique intellectual property portfolio of databases, algorithms, and associated software for technical analysis of derivatives prices. His work, which includes more than a million lines of computer code, is particularly focused on the identification of subtle anomalies and price distortions.

Augen has a 25-year history in information technology. As cofounding executive of IBM's Life Sciences Computing business, he defined a growth strategy that resulted in \$1.2 billion of new revenue and managed a large portfolio of venture capital investments. From 2002 to 2005, Augen was President and CEO of TurboWorx Inc., a technical computing software company founded by the chairman of the Department of Computer Science at Yale University. His books include Trading Realities, Day Trading Options, Trading Options at Expiration, The Option Trader's Workbook, and The Volatility Edge in Options Trading. He currently teaches option trading classes at the New York Institute of Finance and writes a weekly column for Stocks, Futures and Options magazine.

This page intentionally left blank

Preface

n August 2010, Cisco stock (ticker: CSCO) hovered just a few cents below \$25. Several analysts identi-fied the stock as a strong buy. They pointed to the rising demand for network infrastructure that, among other things, was being driven by explosive growth in online video gaming and Internet television. Cisco, they believed, would continue to dominate the consumer market while benefiting from a weak dollar and low manufacturing costs. They must have been wrong because the stock fell 15% when earnings were released on August 11. The price continued to decline until August 31, when it bottomed out at \$19-24% below its previous high. About the time that everyone had given up and turned bearish, the stock began to rally. On November 10 the price was, once again, back up to \$24.50. Then came another earnings report and another sharp decline. The price immediately fell 16% and continued plunging until, on December 3, it once again bottomed out at \$19. These bizarre dynamics played out a third time, with the stock rallying steadily to \$22 on February 9, 2011, before falling back to \$18.92 the very next day after earnings were released—another 14% decline. Figure P.1 displays Cisco closing prices from June 1, 2010, to February 11, 2011.



FIGURE P.1 CSCO closing prices June 1, 2010 to February 11, 2011.

Wild unpredictability doesn't seem to discourage speculators because the trading volume for Cisco remains shockingly high. Moreover, the number of investors who bet on the direction of the stock seems to peak just before and after earnings the most dangerous times of all. For example, the trading volume climbed above 125 million shares on February 9, 2011 (before earnings), and skyrocketed to 560 million shares the next day after the numbers were released. Each of the previously outlined events was accompanied by a similar pattern of extremely high volume the day before earnings were announced and even higher volume the day after. Who would trade this completely unpredictable stock? Why does the volume soar at the most dangerous times when anything seems possible? More importantly, why do analysts believe that they know enough to make predictions? The answers are simple. Analysts can make all the predictions they want because it's not their money that ends up being lost, and speculators always believe they can find a bargain. As a group, investors tend to be arrogant. They typically believe that they have unique insights and that these insights give them an advantage over the market. One of the most common mistakes is relying on traditional off-the-shelf technical indicators that often prove to be even less reliable than fundamental analysis. The Cisco story represents one of the best examples of the problem.

Various technical indicators signaled that the stock would continue to rally just before each of the sharp declines displayed in Figure P.1. They were clearly wrong. Moreover, technical indicators cannot be valid if the underlying trend being analyzed is statistically insignificant. Yet technical analysts routinely talk about moving-average crosses, momentum, or relative strength, without any reference to the statistical strength of the underlying trend being studied. We can compile the relevant statistics for any stock in just a few seconds by loading the information into a spreadsheet and applying Excel's r-squared function. Not surprisingly, the test reveals that most trends appearing on stock charts have very low statistical significance. For Cisco, a relatively weak r-squared value of 0.7 is achieved less than 30% of the time using a 10day sliding window. Highly significant trends with r-squared values above 0.9 occur with a frequency less than 5%. Table P.1 displays r-squared data for 2 years of Cisco stock.

	rsq>.9	rsq>.8	rsq>.7	rsq>.6	rsq>.5
Days	24	96	146	189	238
Percent	4.8%	19.4%	29.5%	38.2%	48.1%

TABLE P.1Compiled r-squared values for Cisco stock February 2009 toFebruary 2011. Calculations span a 10-day sliding window.

The table is divided into columns that reveal the number and percentage of days appearing in trends with minimum r-squared values listed in the column headings. In some sense the data represents a dose of reality. It is common, for example, to hear a technical analyst turn bullish because the 50-day moving average has crossed above the 200-day moving average. However, it is unlikely that you will ever hear the same analyst report the r-squared value of the current trendline. Fortunately, however, most good trading platforms have an r-squared function that can display on a chart, and the data can be exported to a spreadsheet where more detailed analysis can be used to study different length windows and combinations of indicators. This kind of analysis can be used to validate, invalidate, or tune combinations of indicators and give investors an edge against the market. In today's complex computer-driven markets, this kind of analysis can make the difference between winning and losing.

Modern trading platforms always include sophisticated tools for back-testing indicators and strategies. But before a strategy can be tested, it must first be developed, and that development is best accomplished on a foundation of statistical analysis. Spreadsheets and databases are the perfect platform for that kind of technical work. In most cases the process involves a sequence of basic questions designed to reveal the statistical behavior of a stock following a specific set of conditions. There is virtually no limit to the size, number, or complexity of the experiments that can be performed to search for unique correlations that are not generally known to the market.

This book is designed to help technically minded private investors learn to run just a little faster than the market. A few years ago the discussion would have been too complex to be generally useful because it would have been focused on data mining strategies in large databases. That has all changed. Most of the complex statistical analysis and model building that a few years ago could only be accomplished at the institutional level is now within the reach of any investor with a trading platform and a copy of Microsoft Excel. This book is built on that theme. It is designed to help investors learn to translate complex questions into simple spreadsheet models. The discussions span a range from simple conditionals and logical expressions to relatively complex VBA programs that generate statistical summary tables. My goal was to include content that can add value to the efforts of a wide range of investors and to challenge everyone to improve their analytical capabilities.

This page intentionally left blank

INDEX

A

AAPL (Apple Computer), 128-129 ABS function, 109 Access, 16 actual volatility of S&P 500 index, 53 - 54aligning records by date AlignMultipleTickers() function, 85-89 AlignRecords() functions, 76-77 AlignRecordsComplete() function, 80-82 correctly aligned records for two tickers, 74-76 flagged date alignments, 71-72 handling ascending and descending dates, 80-83 record consistency verification with date checksum, 69-70 RemoveSpaces() function, 78-79 removing unmatched records after date alignment, 78 AlignMultipleTickers() function, 85-89 AlignRecords() function, 76-77 AlignRecordsComplete() function, 80-82

Amazon.com price change model column descriptions for sample experiment, 163-166 conditional and logical expressions to evaluate successive price changes, 158-163 downward trend-reversal results for two years of daily price changes, 178-182 experimental trend-following results for two years of daily price changes, 173-174 minute-level time frame analysis, 156 processing trend statistics, 174-177 r-squared calculation and associated data columns, 166-168 r-squared (RSQ) function, 163 sequential price-change tests, 169-171 summary data for failed reversal days following downward spikes larger than 1.1%, 185-186

summary data for reversal days following downward spikes larger than 1.1%, 183-184 tabulation of final results, 172 TrendStat() function, 174-177 Anadarko Petroleum, 145 AND function, 110, 171 Apple Computer (AAPL), 128-129, 147-148 implied volatility, 57 price spikes, 46-49

B

back-testing strategies, 154 Black-Scholes values, 68 Boeing, 141 brief time frames, 57-58 building models. *See* Amazon.com price change model

C

calendar effects at broad market level, 17-20 explained, 9 at individual security level, 12-15 Oil Services HOLDRs exchange traded fund (OIH) example, 14 TOM (turn-of-month) effect, 10 candlestick patterns consecutive pairs of spinning top candles, 112-113 descriptive ratios, 109-115 explained, 108-109 capacity of spreadsheets, 64-65 CBOE Futures Exchange (CFE), 53 CBOE S&P 500 Three-Month Variance Futures, 53 CBOE Volatility Index (VIX), 35,44 VIX/true ratio, 54-56

CFE (CBOE Futures Exchange), 53 Cisco (CSCO) case study, 1-4 CME Group, 145 ColumnLetter function, 88 ColumnNumber function, 88 consecutive pairs of spinning top candles, 112-113 consistency of records, verifying, 69-70 context discovery, 51 converting dates to decimal dates. 91-94 corrections predicting, 54-56 sharp market corrections, 53 correlations, discovering, 128 MAX function, 132 MIN function, 132 PEARSON function, 130-131 Pearson product-moment correlation coefficient, 131-136 compared to r-squared, 163 formula, 129-130 Pearson correlation matrix for first-pass hierarchy, 136 Pearson correlation matrix for 10 tickers, 135-136 Pearson correlation matrix for 6 tickers, 131 PEARSON function, 130 Pearson r calculation for 6 tickers spanning 10 days, 131 price-change responses of 16 stocks/ETFs to two sets of market conditions during 2008 banking collapse, 143-146

price-change responses of 17 stocks/ETFs to two sets of market conditions, 139-142 scatterplot for two closely related stocks, 128-129 Countrywide Financial Corp., 137 CSCO (Cisco) case study, 1-4

D

data mining experiments, discovering price-change relationships with, 21-24 databases versus spreadsheets, 63-65 DataRow variable, 126-127 DATE function, 66 dates aligning records by AlignMultipleTickers() function, 85-89 AlignRecords() function, 76-77 AlignRecordsComplete() function, 80-82 correctly aligned records for two tickers, 74-76 date alignments flagged through indirection, 72 date alignments flagged with IF, 71-72 handling ascending and descending dates, 80-83 record consistency verification with date checksum, 69-70 RemoveSpaces() function, 78-79 removing unmatched records after date alignment, 78

converting to decimal dates, 91-94 date alignments, flagging through indirection, 72 with IF function, 71-72 date formats, 65-69 converting text to, 66-67 default starting dates, 67 recognition of, 65 two- versus four-digit formats, 68 functions DATE, 66 DATEVALUE, 66 **YEARFRAC**, 91, 94 DATEVALUE function, 66 decimal date conversion, 91-94 default starting dates, 67 descriptive ratios, 108-115 direction-neutral volatility distortions, 12-15 dollar sign (\$), 161 DuPont Co., 137

E

efficiency, 11 Efficient Market Hypothesis (EMH), 11 EIA (Energy Information Agency) Petroleum Status Report, 14 Eli Lilly, 140 EMH (Efficient Market Hypothesis), 11 Energy Information Agency (EIA) Petroleum Status Report, 14 ESRX (Express Scripts), 27 evaluating successive price changes. See Amazon.com price change model

event extraction, 50 event-based clustering, 146 Excel spreadsheets. *See* spreadsheets expiration-day behavior, 25-33 Express Scripts (ESRX), 26 extracting events, 50

F

Fama, Eugene, 11 flagging date alignments through indirection, 72 with IF function, 71-72 functions. See also methods ABS, 109 AlignMultipleTickers, 85-89 AlignRecords, 76-77 AlignRecordsComplete, 80-82 AND, 110, 171 ColumnLetter, 88 ColumnNumber, 88 DATE, 66 DATEVALUE, 66 IF, 71-72, 171 IFERROR, 168 INDIRECT, 72-74 LEFT, 66 MAX, 132 MID, 66 MIN, 132 PEARSON, 130-131 RemoveSpaces, 78-79 RIGHT, 66 RSQ, 163 spikes, 119-121 TrendStat, 174-177 YEARFRAC, 91, 94

G

General format, 65 German Central Bank, 9 GLD (SPDR Gold Trust) price-change relationships, 21-24 Google, 140

H

high-low prices changes changes greater than 5% for S&P 500 index (1987). 39-41 changes greater than 5% for S&P 500 index (January 1990-November 2007), 35-37 changes greater than 8% for S&P 500 index (January 1990-December 2010), 37-39 S&P 500 historical volatility (1929), 43-45S&P 500 historical volatility (1987), 41-42historical volatility, calculating, 96-98

Ι

IBM, purchase of Lotus Development Corporation, 8 IF function, 71-72, 171 IFERROR function, 168 implied volatility of S&P options, 53-54 Import Wizard, 65 INDIRECT function, 72-74 indirection, pointing to records with, 72 insider trading, 8-9 intraday volatility, calculating, 100-102 IterationIndex variable, 121

J-K-L

LEFT function, 66 Lotus Development Corporation, purchase by IBM, 8

Μ

managing date formats, 65-69 converting text to dates, 66-67 default starting dates, 67 two- versus four-digit formats, 68 Marathon Oil, 140 market fingerprint, tracking, 13 market inefficiencies, discovering database/spreadsheet approach, 12 - 20graphical approaches, 20-24 MAX function, 132 methods. See also functions Range.Delete, 79 Range.Insert, 84 SpecialCells, 79 Microsoft Access, 16 MID function, 66 MIN function, 132 minute-level time frame analysis, 156-157 models. See Amazon.com price change model multiple tickers calculating volatility across, 98-100 record alignment program for, 85-89

N-O numbers. See dates

objects, WorksheetFunction, 117 Oil Services HOLDRs exchange traded fund (OIH) price changes, 12-15 option prices, calculating, 67

P

patterns, candlestick consecutive pairs of spinning top candles, 112-113 descriptive ratios, 109-115 explained, 108-109 PEARSON function, 130-131 Pearson product-moment correlation coefficient, 131-136 compared to r-squared, 163 formula, 129-130 Pearson correlation matrix for 10 tickers, 135-136 Pearson correlation matrix for 6 tickers, 131 Pearson correlation matrix for first-pass hierarchy, 136 PEARSON function, 130 Pearson r calculation for 6 tickers spanning 10 days, 131 Petroleum Status Report, 14 Philadelphia Gold/Silver Index (XAU) price-change relationships, 21-24 pinning effect, 25-33 pointers, SummaryRow, 123 polynomial trendlines, 149 predicting corrections, 54-56

price Amazon.com price change model, 158 column descriptions for sample experiment, 163-166 conditional and logical expressions to evaluate successive price changes, 158-161, 163 downward trend-reversal results for two years of daily price changes, 178-182 experimental trend-following results for two years of daily price changes, 173-174 processing trend statistics, 174-177 r-squared (RSQ) function, 163 r-squared calculation and associated data columns, 166-168 sequential price-change tests, 169-171 summary data for failed reversal days following downward spikes larger than 1.1%, 185-186 summary data for reversal days following downward spikes larger than 1.1%, 183-184 tabulation of final results, 172 TrendStat() function, 174-177

Cisco (CSCO) case study, 1-4 expiration-day behavior, 25-33 high-low prices changes changes greater than 5% for S&P 500 index (1987), 39-41 changes greater than 5% for S&P 500 index (January 1990-November 2007). 35-37 changes greater than 8% for S&P 500 index (January 1990-December 2010), 37-39 S&P 500 historical volatility (1929), 46-49S&P 500 historical volatility (1987), 41-42option prices, calculating, 67 price distortions, 17 price spikes, 48-49 Excel VBA price-spikesummary program, 119-123 price spike calculations, 46-49, 102-108, 123 - 128summary table format, 117 price-change relationships, discovering database/spreadsheet approach, 12-20 graphical approaches, 20-24 responses to market conditions price-change responses of 16 stocks/ETFs to two sets of market conditions during 2008 banking collapse, 143-146

price-change responses of 17 stocks/ETFs to two sets of market conditions, 139-142 time series correlation, 51-52

Q-R queries (SQL), 32

r-squared (RSQ) function, 3-4, 163 Range.Delete method, 79 Range.Insert method, 84 RangeString variable, 83 ratios, descriptive ratios, 108-115 records, aligning by date AlignMultipleTickers() function, 85-89 AlignRecords() function, 76-77 AlignRecordsComplete() function, 80-82 correctly aligned records for two tickers, 74-76 flagged date alignments, 71-72 handling ascending and descending dates, 80-83 record consistency verification with date checksum, 69-70 RemoveSpaces() function, 78-79 removing unmatched records after date alignment, 78 RemoveSpaces() function, 78-79 removing unmatched records after date alignment, 78 Research in Motion (RIMM), 128-129, 144 **RIGHT** function, 66 RIMM (Research in Motion), 128-129, 144 RSQ function, 163

S

scatterplot for two closely related stocks, 128-129 semi-strong efficiency, 11 sequential price-change tests, 169-171 sharp market corrections, 53 SPDR Gold Trust (GLD) price-change relationships, 21-24 SpecialCells method, 79 spikes() function, 119-121 spreadsheets Amazon.com price change model column descriptions for sample experiment, 163-166 conditional and logical expressions to evaluate successive price changes, 158-163 downward trend-reversal results for two years of daily price changes, 178-182 experimental trend-following results for two years of daily price changes, 173-174 processing trend statistics, 174-177 r-squared (RSQ) function, 163 r-squared calculation and associated data columns, 166-168 sequential price-change tests, 169-171

summary data for failed reversal days following downward spikes larger than 1.1%, 185-186 summary data for reversal days following downward spikes larger than 1.1%, 183-184 tabulation of final results, 172 TrendStat() function. 174-177 capacity of, 64-65 compared to databases, 63-65 dates aligning records by, 69-91 converting to decimal dates, 91-94 date formats, 65-69 descriptive ratios, 108-115 power of, 153-155 statistical correlations, discovering, 128 MAX function, 132 MIN function, 132 PEARSON function. 130-131 Pearson product-moment correlation coefficient, 129-136, 163 price-change responses of 17 stocks/ETFs to two sets of market conditions, 139-142 scatterplot for two closely related stocks, 128-129 summary tables, creating, 116 Excel VBA price-spikesummary program, 119-123

price spike calculations, 123-128 summary table format: price spikes organized by ticker, 117 time frame analysis, 155-157 trendlines, creating, 147-149 volatility calculations, 94-96 across multiple tickers. 98-100 historical volatility, 96-98 intraday volatility, 100-102 price spike calculations, 102-108 SQL queries, 32 standard deviations, comparing price changes in, 141-146 statistical correlations. discovering, 128 MAX function, 132 MIN function, 132 PEARSON function, 130-131 Pearson product-moment correlation coefficient. 131-136 compared to r-squared, 163 formula, 129-130 Pearson correlation matrix for 10 tickers, 135-136 Pearson correlation matrix for 6 tickers, 131 PEARSON function, 130 price-change responses of 17 stocks/ETFs to two sets of market conditions, 139-142 scatterplot for two closely related stocks, 128-129 StdDevRange variable, 125 strong-form efficiency, 11 successive price changes, evaluating, 160-163

successive price changes, evaluating. *See* Amazon.com price change model summary tables, creating, 116, 123-128 Excel VBA price-spike-summary program, 119-123 summary table format: price spikes organized by ticker, 117 SummaryRow pointer, 123

T

tables (summary), creating, 116 Excel VBA price-spike-summary program, 119-123 price spike calculations, 123-128 summary table format: price spikes organized by ticker, 117 text, converting to dates, 66-67 time brief time frames, 57-58 time frame analysis, 155-157 time series correlation, 51-52 TOM (turn-of-month) effect, 10 tracking market fingerprint, 13 trendlines, creating, 147-149 TrendStat() function, 174-177 turn-of-month (TOM) effect, 10

U-V

United Health Group, 141 United Parcel Service, 140 unmatched records, removing after date alignment, 78

variables DataRow, 126-127 IterationIndex, 121

RangeString, 83 StdDevRange, 125 WindowLength, 125 verifying record consistency, 69-70 VIX (CBOE Volatility Index), 35.44 VIX/true ratio, 54-56 volatility, 94-96 actual volatility of S&P 500 index, 53-54 calculating across multiple tickers, 98-100 CBOE Volatility Index (VIX), 35,44 historical volatility calculating, 96-98 S&P 500 historical volatility (1929), 43-45S&P 500 historical volatility (1987), 41-42implied volatility of S&P options, 53-54 intraday volatility, 100, 102 price spike calculations, 102-108 VIX/true ratio, 54-56

W-X-Y-Z

Wal-Mart, 140, 144 weak-form efficiency, 11 Welteke, Ernst, 9 WindowLength variable, 125 wizards, Import Wizard, 65 WorksheetFunction object, 117

XAU (Philadelphia Gold/Silver Index) price-change relationships, 21-24

YEARFRAC function, 91, 94