Marketing Data Science

Modeling Techniques in Predictive Analytics with R and Python

THOMAS W. MILLER

Publisher: Paul Boger
Editor-in-Chief: Amy Neidlinger
Executive Editor: Jeanne Glasser Levine
Operations Specialist: Jodi Kemper
Cover Designer: Alan Clements
Managing Editor: Kristy Hart
Manufacturing Buyer: Dan Uhrig

Published by Pearson Education, Inc.
Old Tappan New Jersey 07675

Printed in the United States of America

First Printing May 2015

ISBN-10: 0-13-388655-7
ISBN-13: 978-0-13-388655-9

Pearson Education LTD.
Pearson Education Australia PTY, Limited.
Pearson Education Singapore, Pte. Ltd.
Pearson Education 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 Control Number: 2015937911

Contents

C.13 Two Month's Salary

I never understood why giving a diamond was the social norm when proposing marriage. As I began searching for an engagement ring, two thoughts kept racing through my mind: "How will I be able to find the right diamond?" and "What is this thing going to cost me?" It goes without saying that my fiancée-to-be is worth the expense, but very seldom in our lives do we spend two month's salary on a product we know so little about.

Most guys are like me. They do not want to spend a lot of time talking to jewelers, doing extensive research, and comparing prices. So for the sake of my male cohort, I took my statistical education to the streets to find out what goes into diamond pricing and value.

I visited ten brick-and-mortar jewelers where I talked with salespeople, tracked data, and viewed more than one hundred diamonds. Then I visited seven online jewelers, gathering information on more than three hundred additional diamonds from two active stores. All observations in my data set represented round-cut diamonds. Although prices of alternative shapes or cuts might be comparable, I only looked at round-cut stones because that shape was the most common, held the most value, and was the only one my girlfriend wanted.

Shortly after beginning my research, I realized why a diamond is the perfect gift to represent an engagement. A diamond symbolizes your choice in a mate because a perfect one is very rare and all of them are unique, complete with imperfections and positive aspects that make them sparkle.

Uniqueness in diamonds is measured using four characteristics called the four Cs: color, clarity, carat, and cut. These traits combine to give a diamond its brilliance and fire. A low level of any one of these attributes can significantly decrease a diamond's value. Here is what I learned about the four Cs.

Carat. Carat is the standard unit of weight used for gemstones (one carat equals 0.200 grams or 200 milligrams). Diamonds are rounded to the nearest hundredth of a carat or point. A 1.27-carat diamond is said to be "one

---

Written by Brian A. Pope.

hundred and twenty-seven points." Typical diamond sizes vary from one-quarter to three carats. Diamonds are sized in one-quarter-carat increments, and jewelers typically carry stock of diamonds at each one-quarter-carat increment. According to jewelry store personnel, not only does price increase with the weight of a stone, but, as a diamond passes each one-quarter-carat threshold, its price jumps correspondingly.

Color. Because diamonds are formed through heat and pressure, the presence of various gases can cause them to take on various tints. Some diamonds are clear. Others have a yellow or brown tint. The Gemological Institute of America (GIA) has established a standard color scale for grading diamonds from D to Z based on tint or color. This scale was used by all twelve of the jewelers I visited. It breaks color grades down into categories like "colorless" and "near colorless." Jewelers indicate that the price of a diamond decreases as you move away from a D grade, which is considered perfectly colorless. In most cases, however, differences in color grade can only be seen when diamonds are compared with one another.

Clarity. The clarity of a diamond measures the purity of the stone. There are often carbon pockets that form imperfections in diamonds called inclusions. Clarity summarizes the number and size of inclusions. The GIA has created a scale that rates inclusions by their visibility to the naked eye. From a flawless (FL) diamond to one that has slight inclusions (SI1 and SI2), salespeople will tell you that the price and value of a diamond decreases as the number of noticeable inclusions increases. But when you shop, you will rarely see a perfectly flawless diamond, and most often you cannot visually detect inclusions at the VVS or VS levels.

Cut. As you go from one jeweler to the next, carat, color, and clarity are defined and measured in a generally universal way. A grade D diamond is perfectly colorless. A diamond with I2 clarity will have plainly visible flaws. And a 1.03-carat diamond has the same weight anywhere you shop. That leaves the type and quality of a stone's cut to differentiate diamond products. The type of cut determines the shape of the diamond, but I limited my study to round-cut diamonds. Determining the quality of cut was more problematic.

I often felt like I was being deceived when salespeople explained why their cut scale was the only appropriate way to measure the quality of cut. A few

jewelers used three criteria that the GIA says make an ideally cut stone: depth, symmetry, and polish. Variations in depth and symmetry can cause a diamond to lose its brilliance. In addition to these two qualities, the overall finish or polish of the stone can have a substantial effect on how well it shines. In the end, I simplified my definition of the cut variable based on my shopping experiences. Regardless of what was said about cut, most jewelers would show two levels of cut. One of the levels would be described as ideal and the other non-ideal. The difference between ideally and non-ideally cut diamonds is not likely to be noticeable to the naked eye, but a diamond will undoubtedly cost more if a jeweler describes it as ideally cut. In addition to the four Cs, I wanted to see if price varied across sales channels. I gathered data from three separate types of jewelers.

Independent Jewelers. These businesses are usually not in an enclosed mall. They are limited to a single community rather than chain stores. Many of the independent jewelers I visited operated at only one location. At independent jewelers I would be given a selection of seven to ten round-cut stones, and store personnel took a non-pressured approach to the sales process.

Mall Jewelers. Located within enclosed malls, many of these jewelers were local branches of national chains. I found the selection of stones to be higher in number but lower in quality. The main factor that I did not like here was the pushy nature of the sales force. I often felt like I was buying a used car.

Internet Jewelers. I looked for online jewelers to complete my analysis. I found two stores with a vast selection of stones. I took a sample of more than three hundred stones from the over four thousand round-cut diamonds available at these two stores. Although online jewelers provided pictures of about half their stones, I would find it difficult to buy a diamond I could not see in person.

Now that the data have been gathered and coded according to the rules summarized in table C.15, I need to figure out which diamond to buy my girlfriend. Furthermore, some of the jewelers are asking questions about why I am collecting this information. One of the independent jewelers is interested in my study. He thinks he might be able to use the results to guide his own diamond buying.

Table C.15. *Diamonds Data: Variable Names and Coding Rules*

Variable	*Description and Coding*	
carat	Weight of the diamond in carats (1 carat = 200 milligrams)	
color	D = 1	I = 6
	E = 2	J = 7
	F = 3	K = 8
	G = 4	L = 9
	H = 5	M = 10
clarity	FL = 1	SI1 = 7
	IF = 2	SI2 = 8
	VVS1 = 3	I1 = 9
	VVS2 = 4	I2 =10
	VS1 = 5	I3 = 11
	VS2 = 6	
cut	Not Ideal = 0	
	Ideal = 1	
channel	Mall = 0	
	Independent = 1	
	Internet = 2	
store	Goodman's = 1	Kay = 7
	Chalmer's = 2	Zales = 8
	Fred Meyer = 3	Danford = 9
	R. Holland = 4	Blue Nile = 10
	Ausman's = 5	Ashford = 11
	University = 6	Riddle's = 12
price	Price in U.S. dollars (April 2001)	