

A collage of business-related items including a calculator, a compass, a pen, and various charts and graphs. The background is a dark, textured surface with a grid pattern. The calculator is in the top left, the compass is in the top right, and the pen is in the bottom left. The charts and graphs are scattered across the background, with a large pie chart in the center. The text "Multiple Regression Analysis" is overlaid in the center in a large, white, sans-serif font. Below it, the subtitle "How Does Product Type Impact Sales Across the Enterprise" is also in white, but in a smaller font. The overall aesthetic is professional and analytical.

Multiple Regression Analysis

How Does Product Type Impact Sales Across the Enterprise



Agenda

- The background of this project
- The project goals
- Description of the data
- Prediction Performance of Different Models
- Prediction Results of Random Forest Model for Test Data
- The Predicted Sales Volume of Four Product Types
- Answer the Business Questions

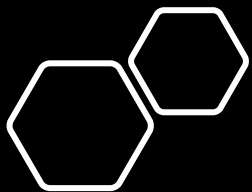
The background of this project

- The sales team have been tracking the sales performance of specific product types.
- The sales team would like us to include the “product type” attribute in our previous predictions to better understand how specific product types perform against each other.
- The sales team have asked our team to analyze historical sales data and then make sales volume predictions for a list of new product types.

A blurred background image of a business meeting. Several people in professional attire are gathered around a table. One person is holding a smartphone, and another is holding a white coffee cup. In the foreground, a laptop screen displays a circular chart. The overall scene suggests a collaborative work environment.

The Project Goals

- Predicting sales of four different product types: PC, Laptops, Netbooks and Smartphones
- Assessing the impact services reviews and customer reviews have on sales of different product types



Description of the data

- 80 observations
- 18 variables

ProductType	ProductNum	Price	x5StarReviews	x4StarReviews	x3StarReviews	x2StarReviews	x1StarReviews	PositiveService Review
PC	101	949	3	3	2	0	0	2
PC	102	2249.99	2	1	0	0	0	1
PC	103	399	3	0	0	0	0	1
Laptop	104	409.99	49	19	8	3	9	7
Laptop	105	1079.99	58	31	11	7	36	7

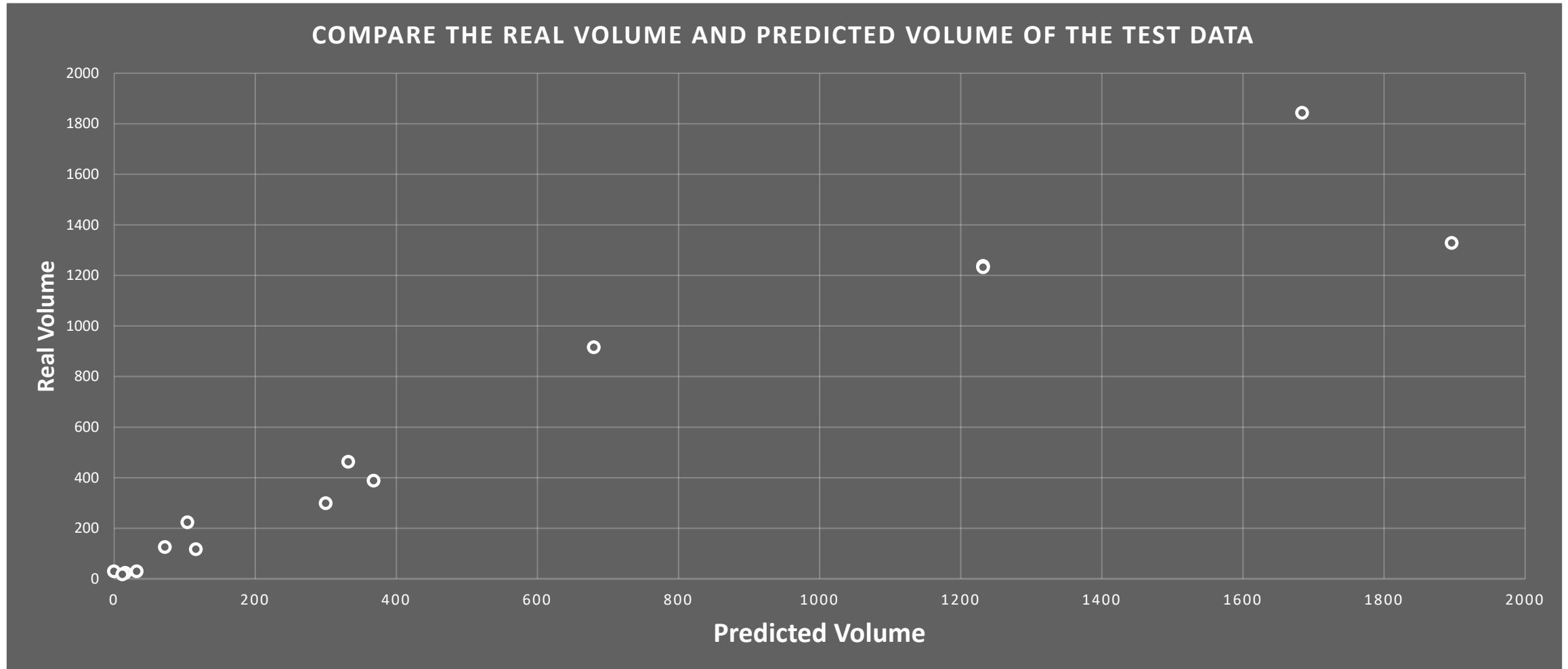
NegativeService Review	Recommend product	BestSellersRank	ShippingWeight	ProductDepth	ProductWidth	ProductHeight	ProfitMargin	Volume
0	0.9	1967	25.8	23.94	6.62	16.89	0.15	12
0	0.9	4806	50	35	31.75	19	0.25	8
0	0.9	12076	17.4	10.5	8.3	10.2	0.08	12
8	0.8	109	5.7	15	9.9	1.3	0.08	196
20	0.7	268	7	12.9	0.3	8.9	0.09	232

	RMSE	Rsquared
SVM	365.5822410	0.9013156
Random Forest	170.9811713	0.9287063
Gradient Boosting	208.3635127	0.8954733

Prediction Performance of Different Models

- Lower the RMSE better is model performance.
- Higher the Rsquared better is model performance.
- The Random Forest Model is the best Model for Prediction the Volume in the project.

Prediction Results of Random Forest Model for Test Data



Variable Importance Analysis

The most importance variables for predicting Sales Volume are Customer reviews (5 Star, 4 Star, 3 Star, 2 Star, 1 Star) and Services reviews (Positive)

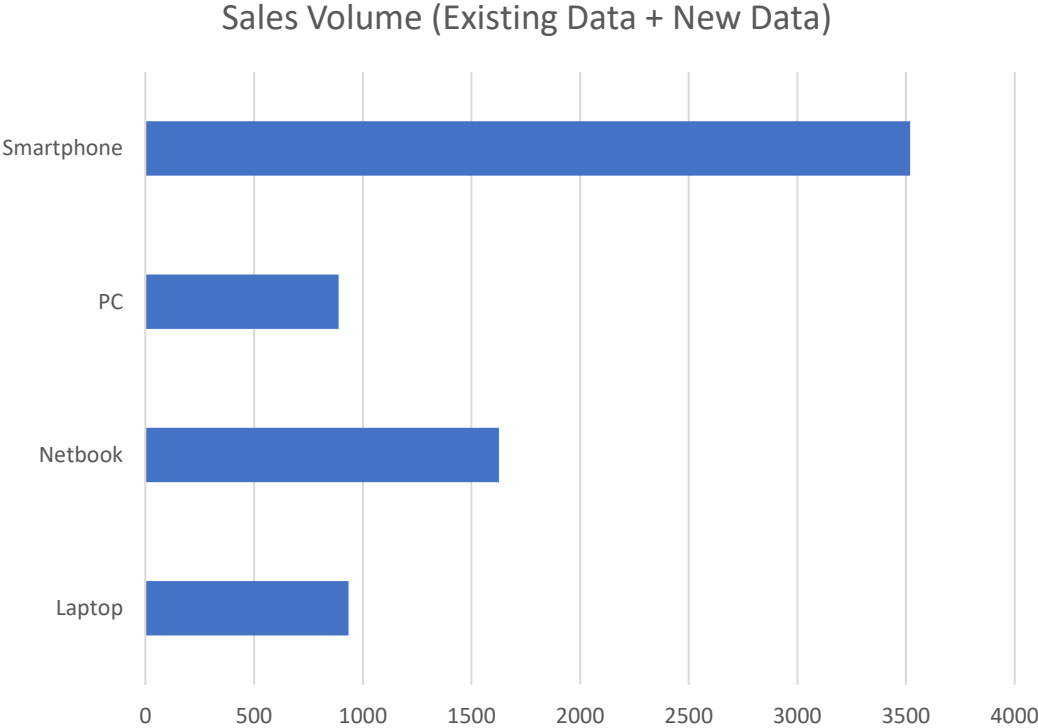
Variables	Importance
5Star.	100.00
Positive	88.64
4Star.	65.00
1Star.	57.67
2Star.	47.90
3Star.	40.35
ProductTypeExtendedWarranty	33.86
ProductTypeLaptop	22.37
Price	22.10
ProductTypeAccessories	21.08
ProductTypeDisplay	21.08
ProductNum	20.78
Depth	20.60
ProductTypeSmartphone	20.27
ProductTypeNetbook	19.89
ProductTypePrinter	19.54
ProductTypeSoftware	19.28
Negative	18.82
Width	17.44
ProfitMargin	16.81

The Predicted Sales Volume of Four Product Types

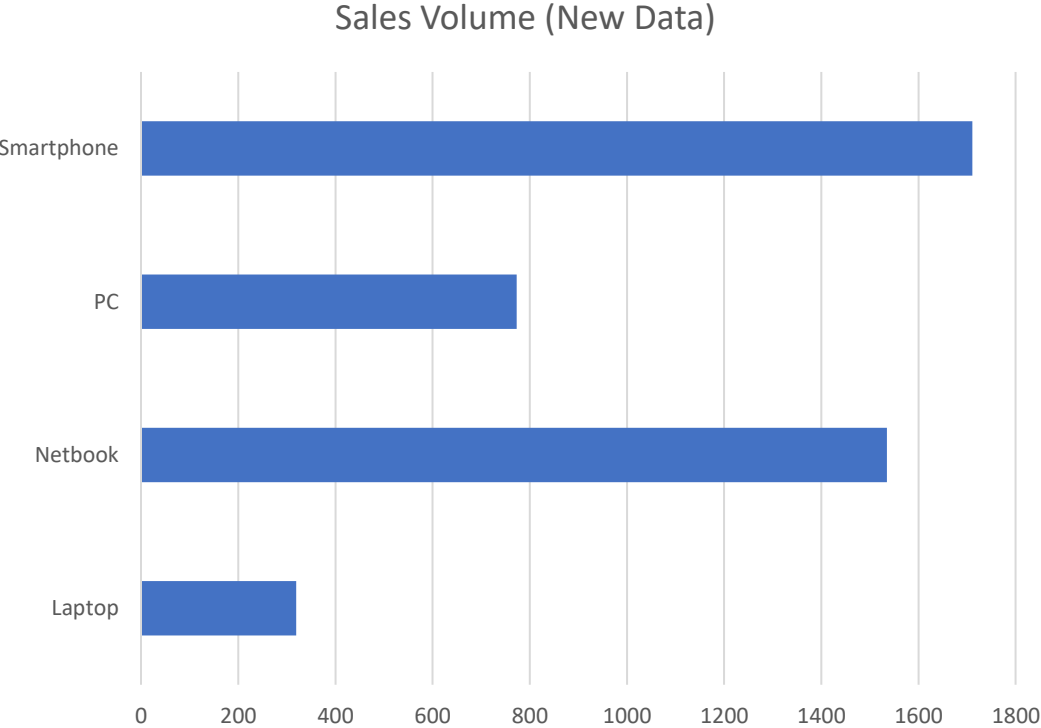
ProductType	ProductNum	Price	5Star	4Star	3Star	2Star	1Star	Positive	Negative	Recommend	BestSeller	Profit	Volume
											Rank	Margin	
Laptop	173	1199	74	10	3	3	11	11	5	0.8	111	0.1	264.0391
Laptop	175	1199	7	2	1	1	1	2	1	0.6	4446	0.15	29.5448
Laptop	176	1999	1	1	1	3	0	0	1	0.3	2820	0.23	25.03027
Netbook	178	399.99	19	8	4	1	10	2	4	0.6	4140	0.08	88.808
Netbook	180	329	312	112	28	31	47	28	16	0.7	2699	0.09	1140.474
Netbook	181	439	23	18	7	22	18	5	16	0.4	1704	0.11	291.1225
Netbook	183	330	3	4	0	1	0	1	0	0.7	5128	0.09	14.1824
PC	171	699	96	26	14	14	25	12	3	0.7	2498	0.25	478.4224
PC	172	860	51	11	10	10	21	7	5	0.6	490	0.2	294.5576
Smartphone	193	199	99	26	12	16	35	8	6	0.4	1277	0.11	451.3475
Smartphone	194	49	100	26	37	33	48	14	6	0.6	16966	0.12	582.4656
Smartphone	195	149	42	8	4	4	9	4	1	0.7	6316	0.15	243.9923
Smartphone	196	300	50	19	13	20	22	5	7	0.6	44465	0.11	433.2585

The Predicted Sales Volume of Four Product Types

Predicted Sales Volume of Total Data



Predicted Sale Volume of New Data



Answer the Business Questions

The regression models can help the to predict the sales volume in different product types and can decide which are the most important factors for prediction, which are very useful to develop the sales strategy.

According to our prediction results, the best sales are netbook and smartphone. I will suggest the sales team put more effort to sale these two product types.



Thank you!

Presented by Wanyun Ho
howanyun@gmail.com