Amazon Prices and Sales Benjamin Lee

benylee2004@gmail.com

University of Washington, Tacoma

Abstract

Amazon's ratings system plays a crucial role in guiding consumer purchases and shaping product visibility in the marketplace. It operates primarily through user-generated reviews and star ratings, reflecting customers' experiences and satisfaction levels. High ratings generally lead to increased sales and better search rankings, while negative ratings can significantly hinder a product's success. The system encourages transparency but is also subject to manipulation, as some sellers may attempt to game the ratings through fake reviews. Overall, Amazon's ratings influence consumer behavior, drive competition among sellers, and impact on the overall shopping experience on the platform. Understanding the dynamics of this system is essential for both buyers and sellers navigating the e-commerce landscape.

Data Dictionary:

Product ID: The identification of the product.

Product Name: The name of the product.

Category: The type of product (electronic, computer)

Discounted Price: The price after a discount

Actual Price: The price of the appliance or product

Discount Price Percentage: The percentage of the price being discounted

Rating: The rating of the product on a 1-5 scale

Rating Count: How many people or users have rated the product

About Product: A description of the product

User ID: The identification of the user

User Name: The name of the user

Review ID: The identification of the reviewer

Review Title: A summary of the review

Review Content: The description of the review

Img Link: The image of the product

Product Link: The link to purchase or view the product

Data Understanding

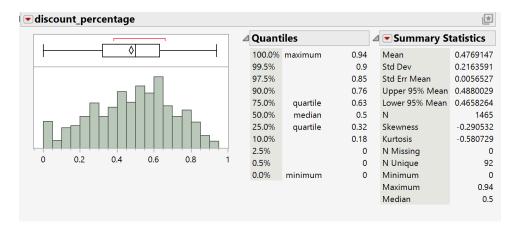
Univariate Properties:

Feature	Var Type	Data Type	Count	Missin g	%	Uniqu e	Min	Q1	Med	Q3	Max	Mean	SD	Skew	Kurt
Product Id	Feature	Qual	1465	0	0	0									
Product Name	Feature	Qual	1465	0	0	1,337									
Category	Feature	Qual	1465	0	0	211									
Discounted Price	Feature	Qual	1465	0	0	0									
Actual Price	Feature	Qual	1465	0	0	0									
Discount Percentage	Feature	Quant	1465	0	0	698.7	0	0.32	0.5	0.63	0.94	0.479	0.2164	-0.29	-3.59
Rating	Label	Quant	1465	0	0	25	2	4	4.1	4.3	5	4.097	0.292	-1.24	7.34
Rating Count	Feature	Quant	1465	0	00	1,108	2	1,091. 50	4,75 6	15,189	98,250	11,945.42	17,617	2.6	10.59
About Product	Feature	Qual	1465	0	0	1,293									
UserID	Feature	Quant	1465	0	0	0									
UserName	Feature	Qual	1465	0	0	1,194									
ReviewID	Feature	Quant	1465	0	0	0									
Review Title	Feature	Qual	1465	0	0	1,194									
Review Content	Feature	Qual	1465	0	0	1,212									
ImgLink	Feature	Qual	1465	0	0	1,412									
ProductLink	Feature	Qual	1465	0	0	1,465									

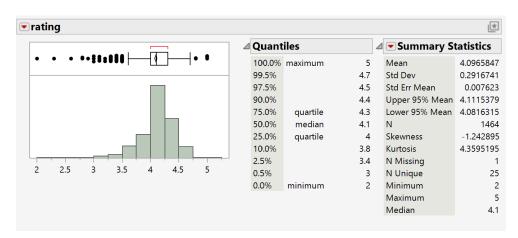
Univariate Analysis

Histogram

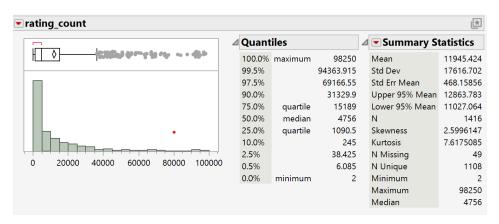
Discount Percentage



Rating



Rating Count

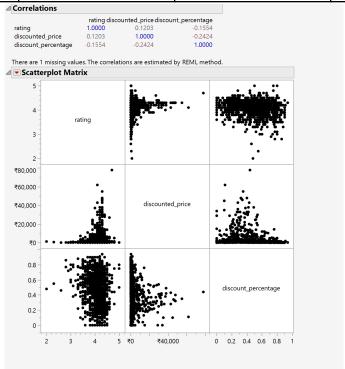


Bivariate Analysis Summary Table:

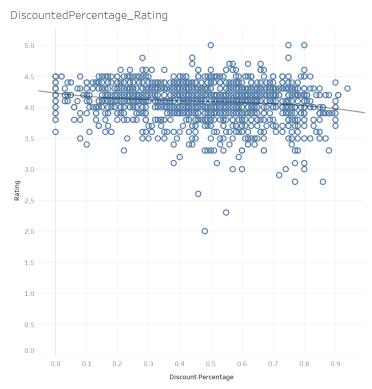
Feature	Analysis	Effect Size	P-Value
Discount Percentage	R-Squared	0.0254672	p<0.0001
Rating Count	R-Squared	0.0634647	p<0.0001
Product ID	T-test	0.092263	0
Product Name	F-stat	0.093047	p<0.0001
Category	F-stat	0.590718	p<0.0001
Discounted Price	T-test	1.80236	p<0.0001
Actual Price	T-test	0.080225	p>0.8473
About Product	F-stat	0.096231	p<0.0001
User ID	T-test	0.096231	p<0.0001
UserName	F-stat	0.104415	0
Review ID	T-test	0.104415	0
Review Title	F-stat	0.104415	0
Review Content	F-stat	0.102862	0
Img Link	F-stat	0.088257	p<0.0001
Product Link	F-stat	0.085074	0

Correlation Matrix

	Rating	Discounted Price	Discount Percentage
Rating	1	0.1203	-0.1554
Discounted Price	0.1203	1	-0.2424
Discount Percentage	-0.1554	-0.2424	1



Discounted Percentage



Discount Percentage vs. Rating.

R²: 0.0254672

P-value: < 0.0001

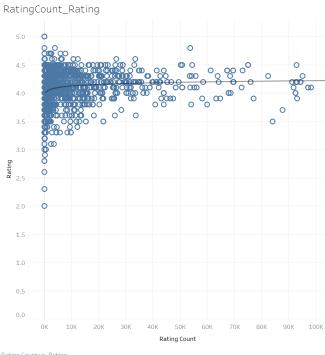
Equation: Rating = -0.775401*Discount Percentage $^3 + 1.06369*$ Discount Percentage $^2 + -0.610195*$ Discount Percentage +4.22967

Coefficients

Term	Value	StdErr	t-value	p-value
Discount Percentage^3	-0.775401	0.561563	-1.38079	0.167555
Discount Percentage^2	1.06369	0.754866	1.40911	0.159015
Discount Percentage	-0.610195	0.294566	-2.0715	0.0384873
intercept	4.22967	0.0341021	124.03	< 0.0001

Summary: There is an effect of *Discounted Percentage* on *Rating*. ($R^2 = 0.0254672$ and p < 0.0001).

Rating Count



Rating Count vs. Rating

R2: 0.0634647

P-value: < 0.0001

Equation: $\ln (Rating) = 0.00987962*\ln (Rating Count) + 1.32612$

Coefficients

Term	Value	StdErr	t-value	p-value
In (Rating Count)	0.0098796	0.0010096	9.78532	< 0.0001
intercept	1.32612	0.0084805	156.372	< 0.0001

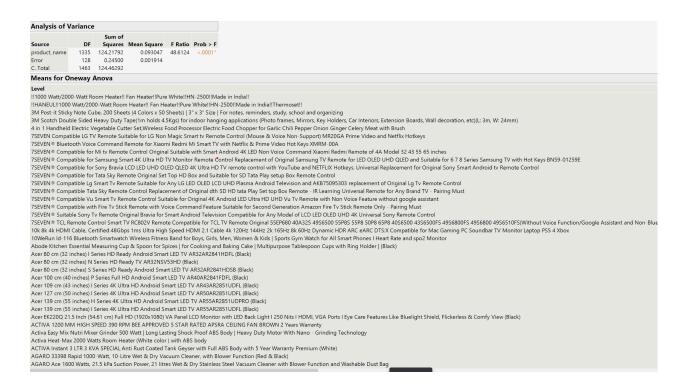
Summary: There is an effect of *Rating Count* on *Rating*. ($R^2 = 0.0634647$ and p < 0.0001).

ProductID

Analysis o			m of								
Source	DF		ares	Mos	n Square		F Ratio	Drob	, E		
product id	1349	124.4		IVIC	0.092263		1 Katio	FIUD	- 1		
Error	114		0000		0.000000				•		
C. Total	1463	124.4			0.000000						
Means fo											
Level		mber	Me	an	Std Erro		Lower	05%	Ha	per 95%	
BOB1F6GQP		1		0000	Stu Elloi	0		.8000	Op	3.8000	
B0B1MDZV		1		0000		0		.9000		3.9000	
BOB1NX6JTI		1		0000		0		.3000		4.3000	
B0B1YVCJ2Y		1		0000		0		.3000		4.3000	ı
BOB1YY6JJL		1		0000		0		.3000		4.3000	
B0B1YZ9CB	8	1		0000		0		1.3000		4.3000	
B0B1YZX72I		1		0000	•	0		.3000		4.3000	
B0B2C5MJN		1		0000		0		.1000		4.1000	
B0B2CPVXH		1		0000		0		.1000		4.1000	
B0B2CWRDI		1		0000		0		.2000		4.2000	
B0B2CZTCL		1		0000		0		.8000		3.8000	
B0B2DD8B0	28	1	4.10	0000		0		.1000		4.1000	
B0B2DD66G		1		0000		0		.5000		4.5000	
B0B2DJ5RV	Q	1	4.30	0000		0	4	.3000		4.3000	
B0B2DJDCP	X	2	3.90	0000		0	3	.9000		3.9000	
B0B2DZ5S6	R	1	4.00	0000		0	4	.0000		4.0000	
B0B2PQL5N	3	1	4.20	0000		0	4	.2000		4.2000	
B0B2RBP83I	>	1	4.00	0000		0	4	.0000		4.0000	
B0B2X35B1I	<	1	4.10	0000		0	4	.1000		4.1000	
B0B3CPQ5P	F	1	4.30	0000		0	4	.3000		4.3000	
B0B3CQBRB	14	1	4.30	0000		0	4	.3000		4.3000	
B0B3D39RK	V	1	4.30	0000		0	4	.3000		4.3000	
B0B3DV7S9		1		0000		0		.6000		3.6000	
B0B3G5XZN		1		0000		0		.5000		4.5000	
B0B3JSWG8		1		0000		0		.1000		3.1000	
B0B3MMYH		1		0000		0		.2000		4.2000	
B0B3MQXN		1		0000		0		.0000		4.0000	
B0B3MWYC	-	2		0000		0		.2000		4.2000	
B0B3N7LR6		2		0000		0		.3000		4.3000	
B0B3N8VG2		1		0000		0		.0000		4.0000	
B0B3NDPCS	-	1		0000		0		.3000		4.3000	
B0B3RHX6B		1		0000		0		.9000		3.9000	
B0B3RRWSF		2		0000		0		.3000		4.3000	
B0B3RS9DN		1		0000		0		.3000		4.3000	
B0B3RSDSZ	-	1		0000		0		.3000		4.3000	
BOB3TBY2Y	Κ.	1	4.30	0000		0	4	.3000		4.3000	

Summary: There is an effect of *Product ID* on *Rating*. ($R^2 = 0.092263$ and p=0).

Product Name



Number	Mean	Std Error	Lower 95%	Upper 95%
1	4.50000	0.04375	4.4134	4.5866
1	3.40000	0.04375	3.3134	3.4866
1	4.40000	0.04375	4.3134	4.4866
1	3.90000	0.04375	3.8134	3.9866
1	3.20000	0.04375	3.1134	3.2866
1	3.00000	0.04375	2.9134	3.0866
1	3.30000	0.04375	3.2134	3.3866
1	3.60000	0.04375	3.5134	3.6866
1	3.60000	0.04375	3.5134	3.6866
1	3.90000	0.04375	3.8134	3.9866
1	4.30000	0.04375	4.2134	4.3866
1	4.20000	0.04375	4.1134	4.2866
1	3.50000	0.04375	3.4134	3.5866
1	3.70000	0.04375	3.6134	3.7866
1	3.30000	0.04375	3.2134	3.3866
1	3.90000	0.04375	3.8134	3.9866
1	3.40000	0.04375	3.3134	3.4866
1	4.60000	0.04375	4.5134	4.6866
1	4.10000	0.04375	4.0134	4.1866
1	4.10000	0.04375	4.0134	4.1866
1	4.30000	0.04375	4.2134	4.3866
1	4.30000	0.04375	4.2134	4.3866
1	4.30000	0.04375	4.2134	4.3866
1	4.30000	0.04375	4.2134	4.3866
1	4.30000	0.04375	4.2134	4.3866
1	4.30000	0.04375	4.2134	4.3866
1	4.30000	0.04375	4.2134	4.3866
1	4.30000	0.04375	4.2134	4.3866
1	4.20000	0.04375	4.1134	4.2866
1	3.90000	0.04375	3.8134	3.9866
1	3.80000	0.04375	3.7134	3.8866
1	4.10000	0.04375	4.0134	4.1866
1	3.80000	0.04375	3.7134	3.8866
1	4.00000	0.04375	3.9134	4.0866
1	4.10000	0.04375	4.0134	4.1866

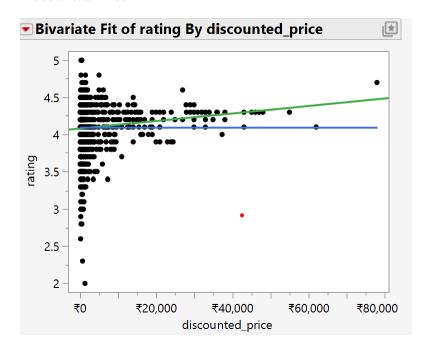
Summary: There is an effect of *Product Name* on *Rating*. ($R^2 = 0.093047$ and p < 0.0001).

Category

Analysis	Analysis of Variance								
		Sum of							
Source	DF	Squares	M	ean Squar	e F Ratio	Prob > F			
category 1	8	4.72574		0.59071	8 7.1782	<.0001*			
Error	1455	119.73718		0.08229	4				
C. Total	1463	124.46292							
Means for Oneway Anova									
Level		Numbe	r	Mean	Std Error	Lower 95	%	Upper 95%	
Car&Motor	bike		1	3.80000	0.28687	3.23	373	4.3627	
Computers	&Accessor	ies 4	53	4.15497	0.01348	4.12	285	4.1814	
Electronics		5	26	4.08175	0.01251	4.05	572	4.1063	
Health&Per	rsonalCare		1	4.00000	0.28687	3.43	373	4.5627	
Home&Kito	chen	4	47	4.04072	0.01357	4.01	141	4.0673	
Homelmpro	ovement		2	4.25000	0.20285	3.85	521	4.6479	
MusicalInst	ruments		2	3.90000	0.20285	3.50)21	4.2979	
OfficeProdu	ucts		31	4.30968	0.05152	4.20	086	4.4107	
Toys&Gam	es		1	4.30000	0.28687	3.73	373	4.8627	
Std Error us	es a poole	d estimate of	err	or variance					

Summary: There is an effect of *Category* on *Rating*. (R²= 0.590718 and p<0.0001).

Discounted Price



Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Ratio			
Model	1	1.80236	1.80236	21.4824			
Error	1462	122.66056	0.08390	Prob > F			
C. Total	1463	124.46292		<.0001*			

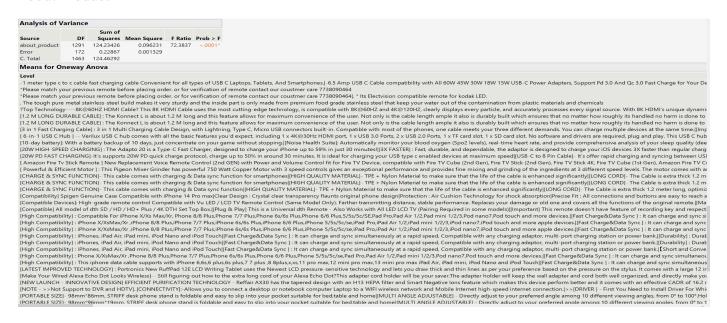
Summary: There is an effect of *Discounted Price* on *Rating*. (R²= 1.80236 and p<0.0001).

Actual Price

Analysis	of Varian	ce			
Source	DF	Sum of Squares		are F Ratio	Prob > F
actual_price		35.94090			
Error	1015	88.52202			
C. Total	1463	124.46292			
Means fo	or Oneway				
Level	Number	Mean S	Std Error	Lower 95%	Upper 95%
₹1,000	10	4.13000	0.09339	3.9467	4.3133
₹1,010	1	4.10000	0.29532	3.5205	4.6795
₹1,020	2	4.10000	0.20882	3.6902	4.5098
₹1,052	1	4.30000	0.29532	3.7205	4.8795
₹1,075	1	4.10000	0.29532	3.5205	4.6795
₹1,090	2	4.30000	0.20882	3.8902	4.7098
₹1,099	19	4.11579	0.06775	3.9828	4.2487
₹1,39,900	1	4.70000	0.29532	4.1205	5.2795
₹1,100	3	4.00000	0.17050	3.6654	4.3346
₹1,109	1	4.40000	0.29532	3.8205	4.9795
₹1,111	1	4.20000	0.29532	3.6205	4.7795
₹1,129	1	4.00000	0.29532	3.4205	4.5795
₹1,130	1	4.20000	0.29532	3.6205	4.7795
₹1,150	1	4.10000	0.29532	3.5205	4.6795
₹1,190	3	4.13333	0.17050	3.7988	4.4679
₹1,199	10	4.03000	0.09339	3.8467	4.2133
₹1,200	3	3.96667	0.17050	3.6321	4.3012
₹1,208	2	4.10000	0.20882	3.6902	4.5098
₹1,230	1	4.10000	0.29532	3.5205	4.6795
₹1,245	2	3.90000	0.20882	3.4902	4.3098
₹1,249	2	3.85000	0.20882	3.4402	4.2598
₹1,282	1	4.20000	0.29532	3.6205	4.7795
₹1,290	3	4.16667	0.17050	3.8321	4.5012
₹1,295	5	4.30000	0.13207	4.0408	4.5592
₹1,299	21	4.16667	0.06444	4.0402	4.2931
₹1,300	3	4.06667	0.17050	3.7321	4.4012
₹1,339	1	4.20000	0.29532	3.6205	4.7795
₹1,345	2	3.85000	0.20882	3.4402	4.2598
₹1,350	1	4.30000	0.29532	3.7205	4.8795
₹1,390	2	4.20000	0.20882	3.7902	4.6098
₹1,395	1	4.20000	0.29532	3.6205	4.7795
₹1,399	4	4.00000	0.14766	3.7102	4.2898
₹1,400	3	4.30000	0.17050	3.9654	4.6346
₹1,440	1	4.20000	0.29532	3.6205	4.7795
₹1,445	1	3.90000	0.29532	3.3205	4.4795

Summary: There is *not an* effect of *Actual Price* on *Rating*. ($R^2 = 0.080225$ and p > 0.8473).

About Product



Number	Mean	Std Error	Lower 95%	Upper 95%	
Number 2	4 50000	0.02578	4 4491	4.5509	
1	3.60000	0.02576	3.5280	3.6720	
1	4.00000	0.03646	3.9280	4.0720	
1	3.40000	0.03646	3.9280	3.4720	
1	4.40000	0.03646	4.3280	4.4720	
2	4.40000	0.03646	4.5260	4.4720	
1	4.10000	0.02376	4.0280	4.1720	
1	4.00000	0.03646	3,9280	4.1720	
1	4.10000	0.03646	4.0280	4.1720	
2	4.00000	0.03040	3 9491	4.1720	
1	4.10000	0.02578	4.0280	4.1720	
3	4.20000	0.03040	4.1584	4.1720	
1	3.30000	0.02105	3.2280	3.3720	
1	3.40000	0.03646	3.3280	3.4720	
2	4.30000	0.02578	4.2491	4.3509	
2	4.20000	0.02578	4.1491	4.2509	
6	4.15000	0.01489	4.1206	4.1794	
1	4.30000	0.03646	4.2280	4.3720	
1	3.70000	0.03646	3.6280	3.7720	
1	4.30000	0.03646	4.2280	4.3720	
1	4 20000	0.03646	4.1280	4.2720	
2	4.20000	0.02578	4.1491	4.2509	
1	4.20000	0.03646	4.1280	4.2720	
1	3.70000	0.03646	3.6280	3.7720	
1	4.20000	0.03646	4.1280	4.2720	
1	4.20000	0.03646	4.1280	4.2720	
1	4.20000	0.03646	4.1280	4.2720	
2	3.90000	0.02578	3.8491	3.9509	
1	4.10000	0.03646	4.0280	4.1720	
1	4.30000	0.03646	4.2280	4.3720	
1	3.80000	0.03646	3.7280	3.8720	
1	4.20000	0.03646	4.1280	4.2720	
1	4.10000	0.03646	4.0280	4.1720	
2	4.20000	0.02578	4.1491	4.2509	

Summary: There is an effect of *AboutProduct* on *Rating*. ($R^2 = 0.096231$ and p<0.0001).

Analysis of Variance
 Squares
 Mean Square
 F Ratio
 Prob > F

 124.46292
 0.104415
 ...

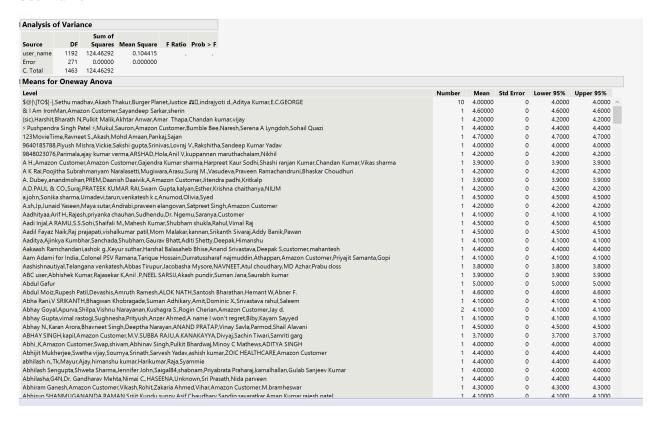
 0.00000
 0.000000
 ...
 124.46292 C. Total

AE2JTMRKTUOIVIZWS2WDGTMNTU4Q,AF4QXCB32VC2DVF703DGFNQVFFNQ,AGAFYHMPFGVPR3MOS4QAZLAWPW3A,AGNNWLEF6V57TKIFIM7SWHNFAIQQ,AFVIPOPKMOCVCX3CMXUJHMWDIMGA,AH6MFUU725GG4KA3XTALSTU2ILHA,AGQYTSKE2UBYAR2 AESDRGJUSPRSINPYZTZAUGEWBA.AEOSTNYHDXOWBNGSZYEGBYMRTZA.AHJRAOGQG4TA8VMMSZOFFSXDOFA.AEDWANALRIDBB4CQK/PMSARCRM7QAEJUXXLGBAZWERWSDADPBGVHTQ.AHJSDSSHOYMITCEBG8W4CT7OCQA.GAGNOMSSWDMHSADDHAGAGGATA8WWASALDBB4CQAK/PMSARCRM7QAEJUXXLGBAZWERWSDADPBGVHTQ.AHJSDSSHOYMITCEBG8W4CT7OCQA.GAGNOTSCTICLEFTONGS.
AESJIMEZHC22EA7YZAUQFTVOUIFQ.AEHZSZRWOOR6UFKX4YRXJMQMJ3MQ.AFZAJPITUPDCOSNY6ASVRJOECMQ.AEPYNLLYHRUBHDIAFPM4OGPALCSQ.AF3IH3T2BPU4MB3U34MNSKQRHJ4Q.AF5IMXZUTWVPW4XMKAEJCJNGVPQ.AHHTWGSVW6ENNVUTEPA
AEBPTERVUSYMSYFVAISICDAXS6FA_AGF5SDMXESGXZRPJKSKBAXMVGGHGQ.AESVMTV93HENV5JBNTJB4LCDHOAA_AHBWFF4SDSJBHNRBRAZWAJRXVAZAAGSOXJBMILSSAMV7OYCRRBWJGCAGPCHJSMCCHUGVOUSGAUCNCQ.AEEEIMNIJHHC3CN3H
AESJGKSUSKSPSDSHYHCKRAOLDOSA,AFUGHF4SZAFXRDSZHMAGBGRSKPRICQAFKANDLOSSWOLDUDISSWOLDJUJNOLDJUJNOLDSSWOLDJUJNOLDJUJNOLDSSWOLDJUJNOLDSSWOLDJUJNOLDJUJNOLDSSWOLDJUJNOLDJUJNOLDSSWOLDJUJNOLDSSWOLDJUJNOLDJUJNOLDSSWOLDJUJNOLDJUJNOLDSSWOLDJUJNOLDSSWOLDJUJNOLDSSWOLDJUJNOLDJUJNOLDJUJNOLDSSWOLDJUJ AE4L3MBEACOHT7Y7GGWO72DUJ6SA AELHNY42N2WPTZN5R42GOW2RO5SA AG5ST5LJL5DUJYA5AH5V4EFXO4JO AG5V75LPUVIVS3DFZNN4UEK2FBJA AEOSYGYF7LVSIVOEEXPC6VCMLBJO AH40BRDNZVTVDLWTAV2V3B422AVO AHTBKWSU5IG46LZWJ3OOZI AESTYLENSINGUS KINDSCASALESEASINIMBEZZOUV YRWEZWIFL CHANGHERWINKINGHINGHINGH ZOOR ZEDNELZER KINDSCASSE ASSESSE ASSESS ASSESSE ASSESSE

Number	Mean	Std Error	Lower 95%	Upper 95%	
2	4.40000	0	4.4000	4.4000	\wedge
2	4.00000	0	4.0000	4.0000	
1	3.50000	0	3.5000	3.5000	
1	4.20000	0	4.2000	4.2000	
1	3.50000	0	3.5000	3.5000	
1	4.10000	0	4.1000	4.1000	
1	4.20000	0	4.2000	4.2000	
1	4.50000	0	4.5000	4.5000	
1	4.30000	0	4.3000	4.3000	
1	4.30000	0	4.3000	4.3000	
3	4.20000	0	4.2000	4.2000	
1	4.00000	0	4.0000	4.0000	
1	4.30000	0	4.3000	4.3000	
1	4.10000	0	4.1000	4.1000	
2	4.00000	0	4.0000	4.0000	
1	4.40000	0	4.4000	4.4000	
1	4.30000	0	4.3000	4.3000	
1	3.80000	0	3.8000	3.8000	
1	4.00000	0	4.0000	4.0000	
1	4.20000	0	4.2000	4.2000	
1	3.80000	0	3.8000	3.8000	
1	4.30000	0	4.3000	4.3000	
1	3.70000	0	3.7000	3.7000	
1	4.20000	0	4.2000	4.2000	
1	3.60000	0	3.6000	3.6000	
1	4.30000	0	4.3000	4.3000	
1	4.30000	0	4.3000	4.3000	
1	4.00000	0	4.0000	4.0000	
1	3.80000	0	3.8000	3.8000	
1	4.10000	0	4.1000	4.1000	
2	4.10000	0	4.1000	4.1000	
1	4.30000	0	4.3000	4.3000	
1	4.10000	0	4.1000	4.1000	
1	4 10000	0	4 1000	4 1000	

Summary: There is an effect of *UserID* on *Rating*. ($R^2 = 0.104415$ and p = 0).

UserName



Summary: There is an effect of *UserName* on *Rating*. ($R^2 = 0.104415$ and p = 0).

Review ID

Analysis	or varia	ince																					
		Sum of																					
ource	DF		Mean Square	F Ratio Prob	> F																		
eview_id	1192	124.46292	0.104415																				
rror	271	0.00000	0.000000																				
C. Total	1463	124.46292																					
Means fo	r Onew	ay Anova																					
_evel																Num	ber	Mean	Std Error	Lowe	95%	Upper 95%	
R1AKJKNRE	IBCV4.R2	ZG9F0E80XA	WO.R39LC2YR7L	3N4E.R2ADNFHJ	2J8A7L.R	R3VV8VK7	7НООҮ	YOS.RX	XGEG3E	BUDZOW	VO						1	4.10000			4.1000		^
				M5MD2S,RY4W)								3NGWIR	B2VT.R2V	WOO592F	U73V7		- 1	4.30000			4.3000		
				5T4MUX,R11RM													1	4.30000		0	4.3000		
				4J.R3JPZAMX1OI													1	3.80000		0	3.8000	3.8000	
				IFB.R4FCBHSKL9													1	3.50000		0	3,5000		
				E8OJA,RA9R916JI											BIOM		- 1	4.40000		-	4.4000		
				N6FTO4,R1LUISQ8													- 1	4.30000		-	4.3000		
				MGJX5,RILQMI1I													- 1	4.20000		0	4.2000		
				C4Q.R2XA4OT30													- 1	4.10000			4.1000		
				QCD.RWBH0HJW													- 1	4.10000			4.1000		
				BO,R3JH7SHSXDT													3	4,30000		-	4,3000		
		KQY8Y6U3Y4		,	,			-,		,			_,				- 1	4.40000			4,4000		
				WG94P7,R2J2KA	10UGEH	31 R2O717	IZD5DO	ORYLN	N R1B31	1TOG8VE	WWH R7	7K5A II5Y	VIMCI R1I	IMH92PE	PV73Y		- 1	4.20000		0	4.2000		
				JX5O5.R1N5UJPJ													4	4.00000		n	4.0000		
				NG6WLD0.RTH4II				_									- 1	3,90000			3,9000		
				MAHXJ.R2CXOL													- 1	4.20000		-	4.2000		
				66T4P,R3KK4GT9										,	XIII LLIC		3	4.20000		-	4.2000		
				W67Z,R371Z2WN											161.0		1	4.40000		0	4.4000		
				MMXW,R1SO4U													- 1	4.40000			4,4000		
				OH22B.R2EG8VX													- 1	3.90000		•	3,9000		
				N7,R1QAZXMA5											OWA		- 1	3.80000		-	3.8000		
				HO1Z,R3GQL7YKA													3	4.50000			4.5000		
				HEGNR,RXG29ZH											LOS26		1	3.80000			3.8000		
				9D,R1P01XZPNV											10320		- 1	3.90000			3.9000		
				GYLM25.R2U6XBE											JEE80		- 1	3.80000		-	3,8000		
				C5VJ,R1LJCG64H													- 1	3.90000			3.9000		
				UNXW,R1VQUW											SAORR		- 1	4.20000		-	4.2000		
				3DWY5,R3QBLT1													2	4.40000		0	4.4000		
				543JO84.R2FKWV													1	4.50000			4.5000		
				DY5RZ02.R1KMS2													- 1	4.30000		•	4.3000		
				DNRBK,R3JJ8CIA		,											- 1	4.20000			4.2000		
				LG4KR,R2TWSF8L													- 1	4.00000			4.0000		
				JOKR.R2ICWHHE.													- 1	3.50000			3.5000		
			-	(YB14D.R3375SV						-	-						- 1	4.20000		-	4.2000		
				WLQ7,R38X9EM0													- 1	4.50000			4.5000		
				8S6G9O3.R3H5V													1	4.00000		0	4.0000		
				* * * * * * * * * * * * * * * * * * * *														4.00000			4.0000	4.0000	

Summary: There is an effect of *ReviewID* on *Rating*. ($R^2 = 0.104415$ and p=0).

Review Title

Allalysis U	of Variar	ice			
		Sum of			
ource	DF		Mean Square		Prob > F
eview_title		124.46292	0.104415		
Tor	271	0.00000	0.000000		
. Total		124.46292			
	r Onewa	y Anova			
evel					
			money TV, exc		
			r a USB 2.0 Devi Sound clarity.,Go		
			lent 🛘 ,Really Go		
			ok,Better,Nice,W		
			use,Good choic		
V2A Route	erUPS,Dec	ent Product,F	roduct quality,E	ffective and	d function
			k temp control, l		
			iter2900watt ned		
			get. The volume		
			as missing from but Wow facto		
			e C Device,Good		
			, Brightness is g		
			purchase,Best r		
			or Money,Am no		
			priced Item,A bi		
			hey Work Well a		
			e for money!,No		
			e specially in No		
			n.,Easy to use,Su nyc hairdryee,Go		
			ormance,Great S		
			res craving for a		
			h scope for imp		
			wire,Good produ		
			ty Music But Cor		t thanks us
A well-pricer	d product.	Lenthy cord.		mes with co	
			Product is work	ing as expe	ompromise ected.,Leng
Above avera			e Money 🛮,Okie	ting as expe e,Phone is e	ompromise ected.,Leng excellent,Pu
Above avera Absolutely lo	oving it!,Ex	cellent produ	e Money 🛮 ,Okie uct,Very nice,Co	ing as expe e,Phone is e impact way	ompromise ected.,Leng excellent,Pu to make co
Above avera Absolutely lo Absolutely w	oving it!,Ex orth the p	cellent produ rice,Best pro	e Money II,Okie uct,Very nice,Co duct on low valu	ting as expe e,Phone is e impact way ue, definitel	ompromise ected.,Leng excellent,Pu to make co ely gofor thi
Above avera Absolutely lo Absolutely w Adhesion,Go	oving it!,Ex orth the pood	cellent produ rice,Best pro ct,Does not s	e Money 🛮 ,Okie uct,Very nice,Co	ting as expe e,Phone is e empact way ue, definitel note,good,N	ompromise ected.,Leng excellent,Pu to make co ly gofor thi Not up to t

Number	Mean	Std Error	Lower 95%	Upper 95%	
1	3.90000	0	3.9000	3.9000	^
1	4.30000	0	4.3000	4.3000	
1	3.80000	0	3.8000	3.8000	
1	4.40000	0	4.4000	4.4000	
1	4.00000	0	4.0000	4.0000	
4	4.30000	0	4.3000	4.3000	
1	4.30000	0	4.3000	4.3000	
1	4.20000	0	4.2000	4.2000	
1	4.20000	0	4.2000	4.2000	
1	4.10000	0	4.1000	4.1000	
1	3.90000	0	3.9000	3.9000	
1	4.30000	0	4.3000	4.3000	
7	4.00000	0	4.0000	4.0000	
1	4.00000	0	4.0000	4.0000	
1	4.10000	0	4.1000	4.1000	
1	4.30000	0	4.3000	4.3000	
1	4.00000	0	4.0000	4.0000	
1	3.90000	0	3.9000	3.9000	
1	3.70000	0	3.7000	3.7000	
1	4.00000	0	4.0000	4.0000	
1	4.40000	0	4.4000	4.4000	
1	4.20000	0	4.2000	4.2000	
1	4.40000	0	4.4000	4.4000	
1	4.10000	0	4.1000	4.1000	
1	3.80000	0	3.8000	3.8000	
1	3.90000	0	3.9000	3.9000	
1	3.60000	0	3.6000	3.6000	
3	4.00000	0	4.0000	4.0000	
6	4.10000	0	4.1000	4.1000	
1	4.50000	0	4.5000	4.5000	
1	4.50000	0	4.5000	4.5000	
1	4.40000	0	4.4000	4.4000	
1	4.10000	0	4.1000	4.1000	
1	4.30000	0	4.3000	4.3000	

Summary: There is an effect of *ReviewTitle* on *Rating*. ($R^2 = 0.104415$ and p=0).

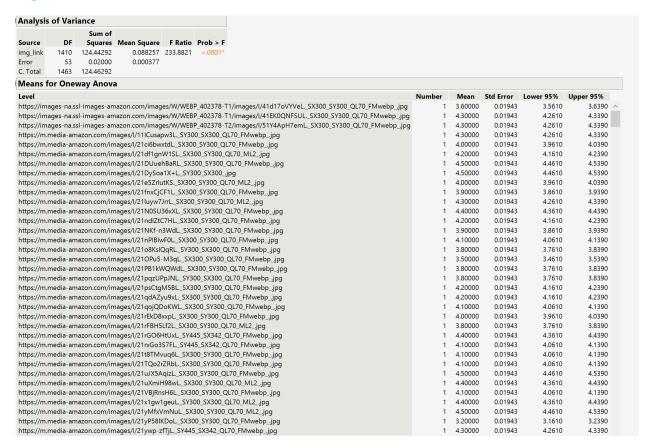
Review Content



Number	Mean	Std Error	Lower 95%	Upper 95%	
1	4.60000	0	4.6000	4.6000	\wedge
1	4.10000	0	4.1000	4.1000	
1	3.90000	0	3.9000	3.9000	
1	4.30000	0	4.3000	4.3000	
1	3.50000	0	3.5000	3.5000	
1	3.80000	0	3.8000	3.8000	
1	4.10000	0	4.1000	4.1000	
1	3.60000	0	3.6000	3.6000	
1	4.30000	0	4.3000	4.3000	
1	4.30000	0	4.3000	4.3000	
1	3.90000	0	3.9000	3.9000	
1	3.20000	0	3.2000	3.2000	
1	3.60000	0	3.6000	3.6000	
1	3.80000	0	3.8000	3.8000	
1	3.80000	0	3.8000	3.8000	
1	4.30000	0	4.3000	4.3000	
1	3.80000	0	3.8000	3.8000	
1	3.80000	0	3.8000	3.8000	
1	4.00000	0	4.0000	4.0000	
2	4.30000	0	4.3000	4.3000	
1	3.70000	0	3.7000	3.7000	
1	4.30000	0	4.3000	4.3000	
1	3.90000	0	3.9000	3.9000	
2	4.20000	0	4.2000	4.2000	
1	4.20000	0	4.2000	4.2000	
1	4.30000	0	4.3000	4.3000	
5	3.80000	0	3.8000	3.8000	
1	3.80000	0	3.8000	3.8000	
1	4.10000	0	4.1000	4.1000	
1	4.30000	0	4.3000	4.3000	
1	3.30000	0	3.3000	3.3000	
1	4.10000	0	4.1000	4.1000	
1	4.20000	0	4.2000	4.2000	
1	4.10000	0	4.1000	4.1000	
)

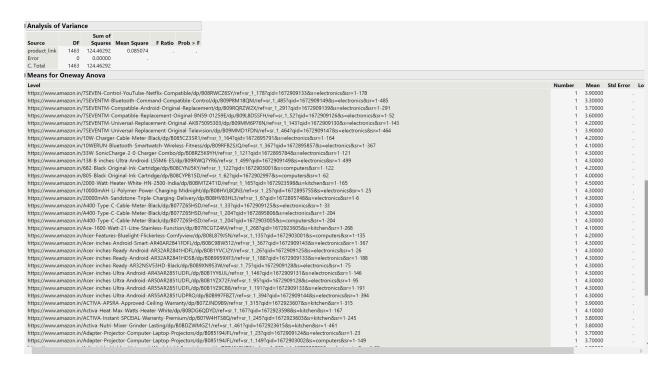
Summary: There is an effect of *ReviewContent* on *Rating*. ($R^2 = 0.104415$ and p = 0).

Img Link



Summary: There is an effect of *ImgLink* on *Rating*. (R²= 0.088257 and p<0.0001).

Product Link



Summary: There is an effect of *ProductLink* on *Rating*. ($R^2 = 0.085074$ and p < 0.0001).

Data Preparation:

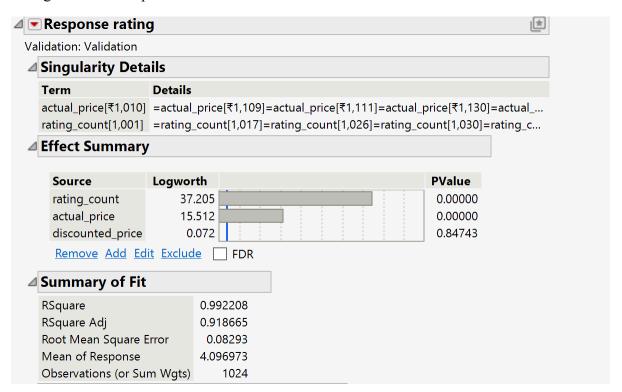
- No missing data points or outliers were found in the data
- No skewness correction was needed
- Category Unique Ordered
- I tried grouping the other variables, but they were not able to be grouped properly
- Changed Review ID to Geographical (Country/Region)
- Changed Actual price to number (whole)
- Changed Rating to Continuous

Model Testing:

Model Iteration	Features	Adj.R ²	RMSE
1	Category, Rating	0.236491	0.25661
	Count, Discounted		
	Price, Actual Price		
2	Rating, ProductID,	1	3.973e-8
	UserID, ReviewID		

3	DiscountedPrice,	0.918665	0.08293
	ActualPrice,		
	RatingCount		
4	RatingCount,	0.917859	0.08334
	ActualPrice,		
	DiscountPercentage		
5	RatingCount,	1	0
	ActualPrice,		
	ReviewContent		

The best performing model was Model 3, which included Discounted Price, Actual Price, and Rating Count as our predictors.



This model had the highest Adjusted R^2 at 0.918665 and the best RMSE at 0.08293, showing that this is the best fit while at the same time, minimizing prediction error.

Findings:

- 1. Rating Count, Actual Price, and Discounted Price are the strongest predictors of Rating.
 - a. The Rating Count significantly affects the Rating, with higher Rating Counts leading to higher Ratings.
 - b. The Actual Price plays another significant role affecting the Rating, both going hand in hand on affecting the Rating.
- 2. Category, Rating Count, Discounted Price, Actual Price are the lowest predictors of Rating.
 - a. The Category significantly affected the Rating, causing the predictors to be lowered.
- 3. RatingCount, ActualPrice, ReviewContent, ProductID, UserID, ReviewID all had little to no impact on the Rating's predictors.
 - a. This suggest that Ratings were primarily based off of the Rating Count and the Actual Price.
 - b. Also shows that the ID's had little to no significance.

These findings provide valuable information on Ratings and how they work. Rating Counts (Along with the Product ID, User ID, Review ID) and Price (Actual Price, Discounted Price) affect the likelihood of purchasing the item, while the type (Category) affects the impact on what the object is being bought, and how often it is purchased. The Ratings can be used to determine what objects are being bought, how frequent they are purchased, and what causes people to purchase said object.

Comments:

Tableau Dashboard URL:

https://public.tableau.com/views/FinalProject_17420642913920/DiscountedPercentage_Rating?: language=en-US&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link

Why I Selected These Measures:

Rating, Discount Percentage, Rating: These variables help Amazon understand different results on what or how discounts affect the Ratings, and how often a person would purchase the item.

The variables I selected all significantly impact ratings based on Amazon reviews. This table is designed to enable users to compare ratings among various options, illustrating how different configurations can influence user satisfaction. It's important to observe rating fluctuations that occur with variations in features, as well as to analyze how different brands rank in customer feedback for similar offerings. It is also important to note that the more a Rating Count there is,

the more likely a person is to look through the Rating's in a closer manure to ensure that they are buying something that they believe is worth it. Additionally, examining rating changes for specific configurations from a single brand provides valuable insight into brand-specific performance and user experience, allowing for a deeper understanding of the overall market as well as individual brand distinctions. Finally, the choice of Amazon was a conscious choice, as the use of Amazon is something that most people in Washington America have used. Amazon is one of the leading companies throughout the entire world, so to understand how they use Ratings, as well as understanding the customer was very interesting.