

Linear Regression

Prediction Face Creams Price

By: Rana Alzahrani

Introduction

- The face cream industry is currently experiencing huge growth.
 - The Market looks set to skyrocket as people look for quick and easy solutions.
 - Price manipulation
- By doing web scrubbing from IHERRB.com

Workflow

01

Web Scraping

02

Cleaning the
data

03

Linear
Regression Model

04

feature
engineering

05

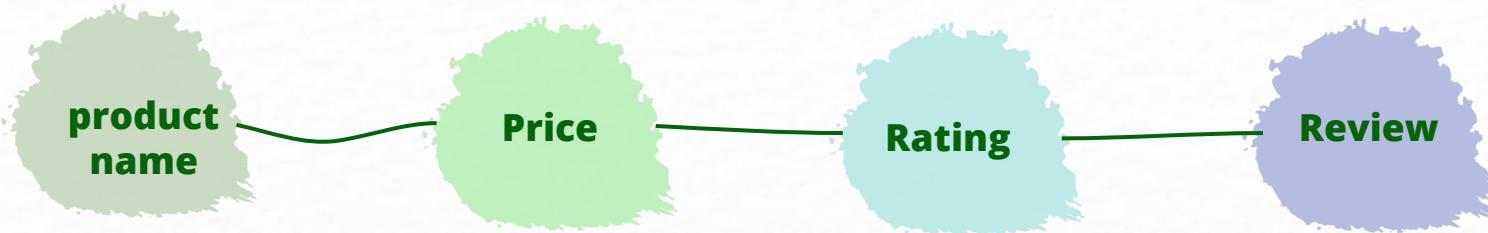
Best model

01

Web scraping

Web scraping

- scraped 1000+ product by using BeautifulSoup



02

Data cleaning

Data cleaning

	product_title	Price	Review	Rating
0	PrescriptSkin, Retinol Night Cream, 1.55 oz (44 g)	AU\$22.95	965.0	4.5/5 - 965 Reviews
1	PrescriptSkin, Vitamin C Moisturizer, Enhanced... ... 1.75 oz (50 g)	AU\$14.35	1562.0	4.5/5 - 1562 Reviews
2	PrescriptSkin, Stem Cell Cream, 2.25 oz (64 g)	AU\$22.95	1167.0	4.5/5 - 1167 Reviews
3	Sympli Beautiful, All Day Hydration Cream, 2 oz (56 g)	AU\$19.13	1356.0	4.4/5 - 1356 Reviews
4	Eucerin, Q10 Anti-Wrinkle Face Cream, 1.7 oz (48 g)	AU\$16.45	20523.0	4.5/5 - 20523 Reviews

03

Linear Regression Model

Correlation

Price	1	-0.17	0.059
Review	-0.17	1	-0.02
Rating	0.059	-0.02	1

04

feature engineering

Add New column (brand_label)

	Name	Price	Review	Rating	brand	brand_label
0	PrescriptSkin, Retinol Night Cream, 1.55 oz (44 g)	22.95	965.0	4.5	PrescriptSkin	131
1	PrescriptSkin, Vitamin C Moisturizer, Enhanced... ...ntioxidant Serum, 1.7 fl oz (50 ml)	14.35	1562.0	4.5	PrescriptSkin	131
2	PrescriptSkin, Stem Cell Cream, 2.25 oz (64 g)	22.95	1167.0	4.5	PrescriptSkin	131
3	Sympli Beautiful, All Day Hydration Cream, 2 oz (56 g)	19.13	1356.0	4.4	Sympli Beautiful	174
4	Eucerin, Q10 Anti-Wrinkle Face Cream, 1.7 oz (48 g)	16.45	20523.0	4.5	Eucerin	62

New Correlation

	Price	Review	Rating	brand_label
Price	1	-0.17	0.059	0.18
Review	-0.17	1	-0.02	-0.02
Rating	0.059	-0.02	1	0.0036
brand_label	0.18	-0.02	0.0036	1

Feature Engineering

- First score with out dummy variables for testing : 0.0701
- score with dummy variables testing : 0.0904

Result :

Model	Train	Test	Val
Linear Regression	0.025	0.054	0.049
Polynomial	0.038	0.090	0.074

05

Best model

The chosen model

As we can see here, for the polynomial regression which will give the best test value than from Linear regression and fits the data.

Recommendations

The Polynomial regression best model I prefer to :

- Add more features.
- Collect more production the data



THANKS