

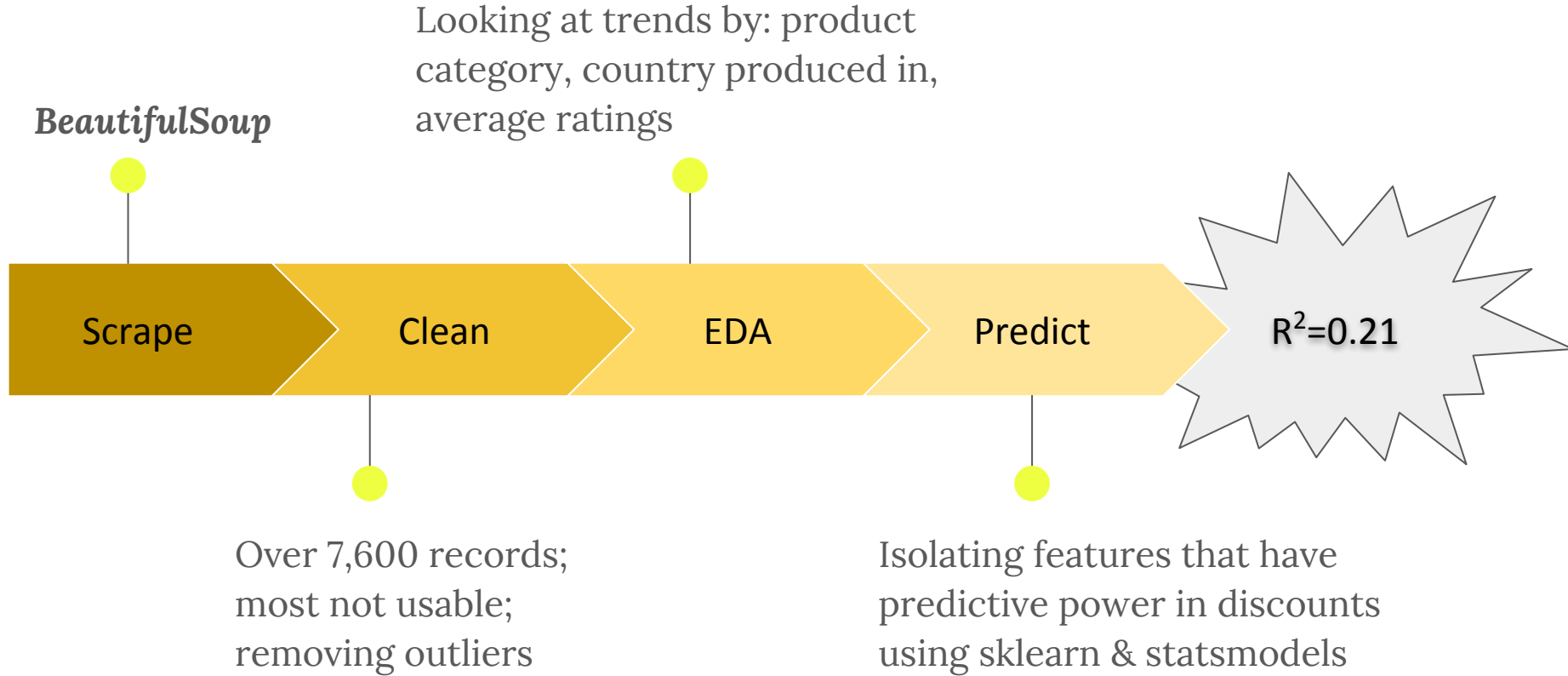


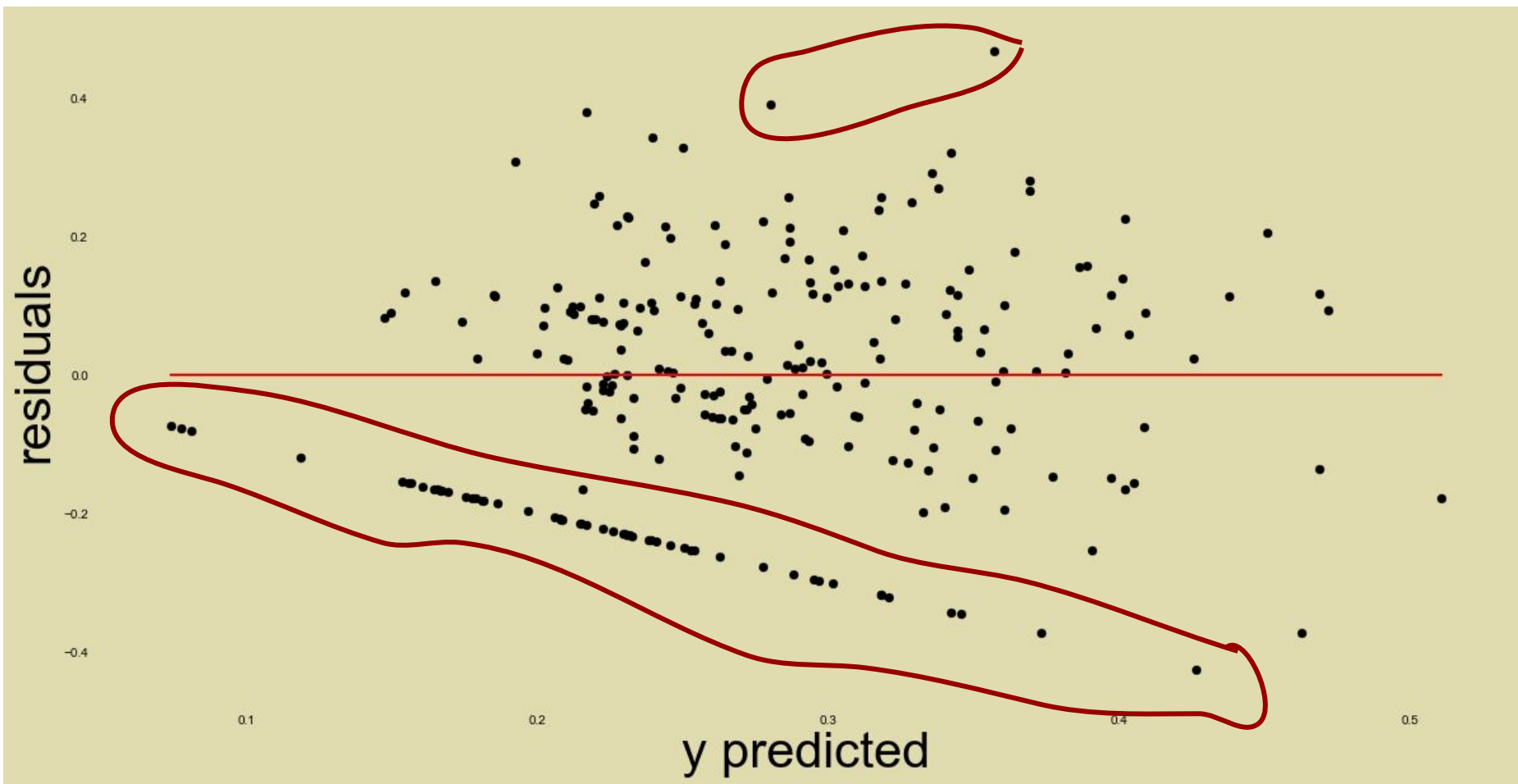
Predicting Discounts on Sur la table

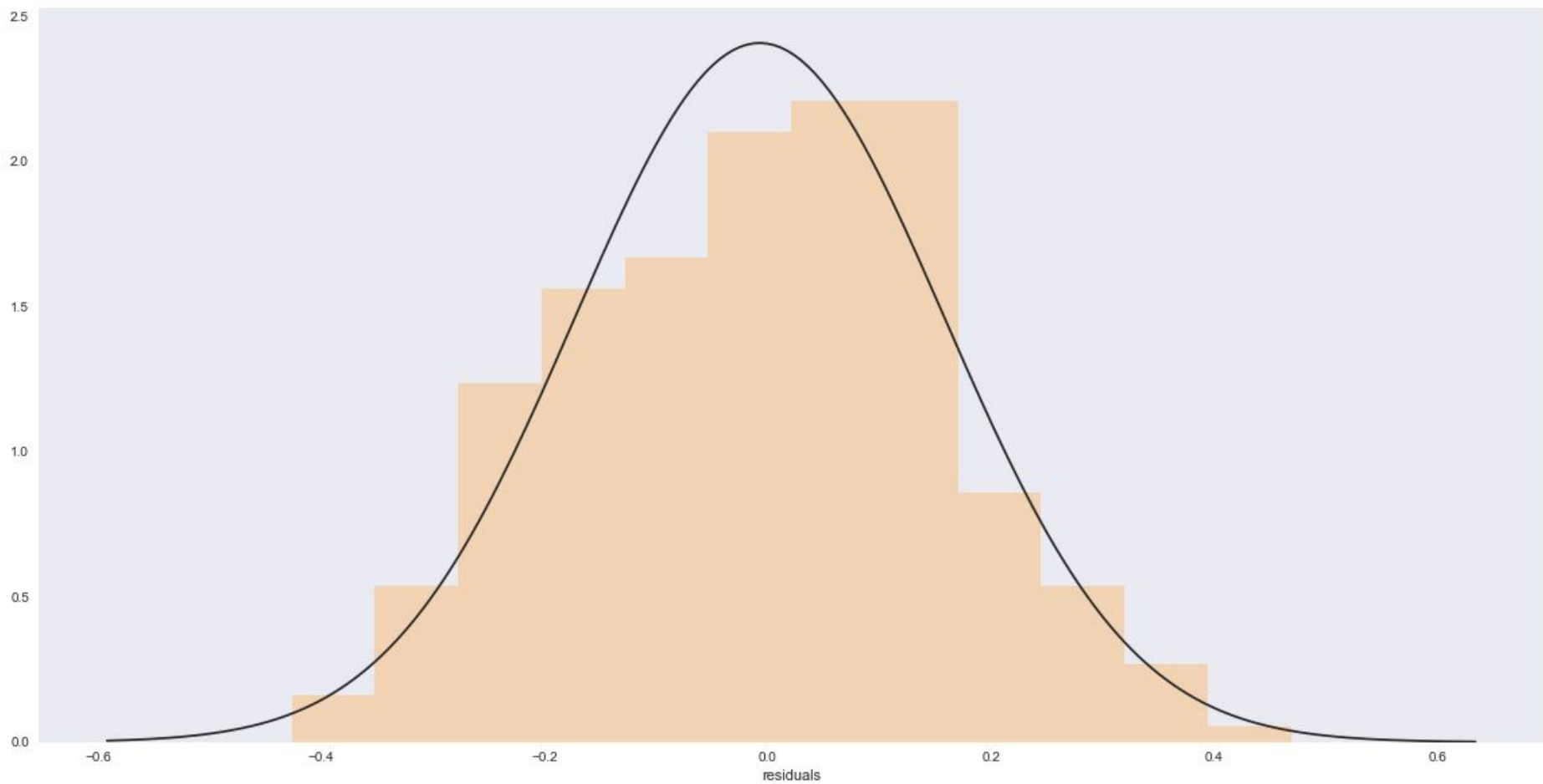
Lisa Sabovic

What is Sur La Table?

- It means “On the Table” in French
- The second-largest specialty cookware brand in the US
- Not super affordable...
- But has lots of discounted products!
- **Objective: let's try to predict a discount on items of at least \$100!**







residuals

...and what does $R^2=0.21$ tell us?

Features in this model can only explain 21.3% of variance in how much a product was discounted.

Dep. Variable:	percentSaved	R-squared:	0.213
Model:	OLS	Adj. R-squared:	0.207
Method:	Least Squares	F-statistic:	31.20
Date:	Thu, 24 Jan 2019	Prob (F-statistic):	3.94e-28
Time:	18:32:55	Log-Likelihood:	191.40
No. Observations:	581	AIC:	-370.8
Df Residuals:	575	BIC:	-344.6
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	0.1801	0.015	11.756	0.000	0.150	0.210
suggPrice	0.0002	3.67e-05	4.095	0.000	7.83e-05	0.000
answers	0.0016	0.000	9.765	0.000	0.001	0.002
France	-0.0560	0.020	-2.869	0.004	-0.094	-0.018
Germany	0.0596	0.026	2.275	0.023	0.008	0.111
USA	-0.1355	0.025	-5.338	0.000	-0.185	-0.086

Omnibus:	17.854	Durbin-Watson:	2.087
Prob(Omnibus):	0.000	Jarque-Bera (JB):	10.848
Skew:	0.177	Prob(JB):	0.00441
Kurtosis:	2.432	Cond. No.	1.37e+03

...and which features does this include?

1. As MSRP increases, so does the discount rate
2. As the amount of answers increases, so does the discount rate
3. If an item is made in either France or USA, the discount rate *decreases relative to if its made in China*
4. If an item is made in Germany, the discount rate *increases relative to if its made in China*

Dep. Variable:	percentSaved	R-squared:	0.213			
Model:	OLS	Adj. R-squared:	0.207			
Method:	Least Squares	F-statistic:	31.20			
Date:	Thu, 24 Jan 2019	Prob (F-statistic):	3.94e-28			
Time:	18:32:55	Log-Likelihood:	191.40			
No. Observations:	581	AIC:	-370.8			
Df Residuals:	575	BIC:	-344.6			
Df Model:	5					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	0.1801	0.015	11.756	0.000	0.150	0.210
suggPrice	0.0002	3.6e-05	4.095	0.000	7.83e-05	0.000
answers	0.0016	0.000	9.765	0.000	0.001	0.002
France	-0.0560	0.020	-2.869	0.004	-0.094	-0.018
Germany	0.0596	0.026	2.275	0.023	0.008	0.111
USA	-0.1355	0.025	-5.338	0.000	-0.185	-0.086
Omnibus:	17.854	Durbin-Watson:	2.087			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	10.848			
Skew:	0.177	Prob(JB):	0.00441			
Kurtosis:	2.432	Cond. No.	1.37e+03			

BeautifulSoup



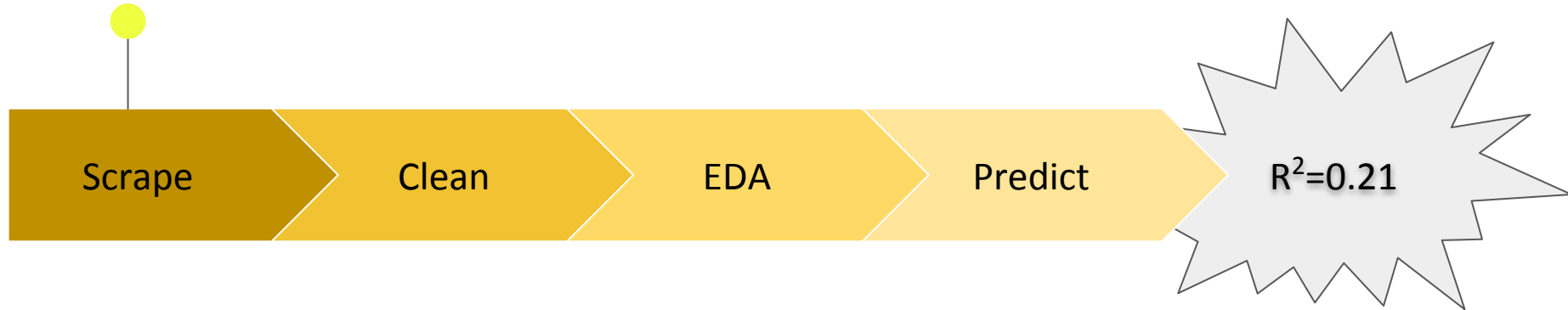
Scrape

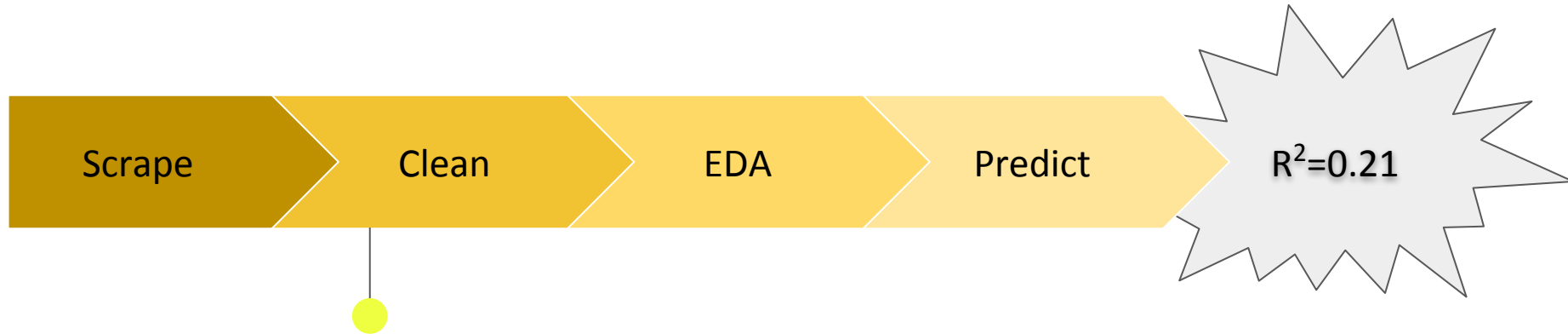
Clean

EDA

Predict

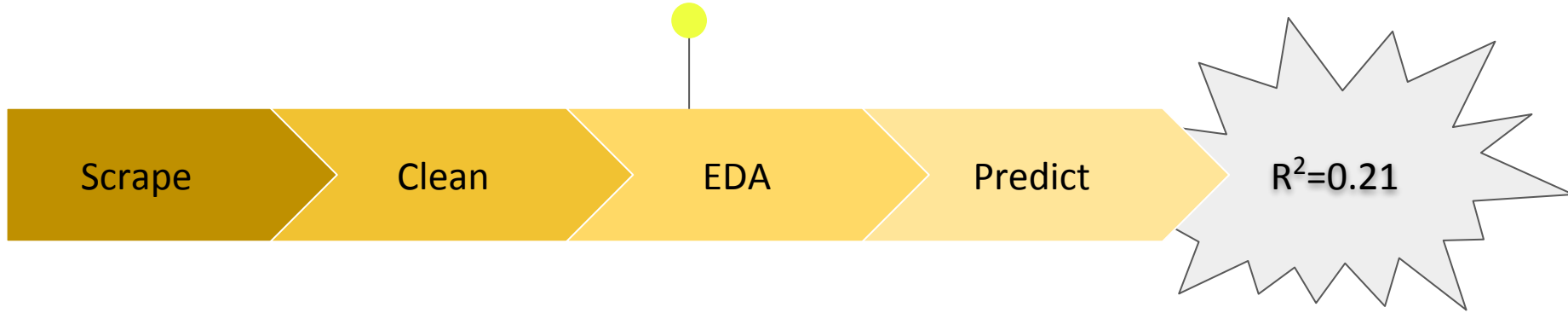
$R^2=0.21$



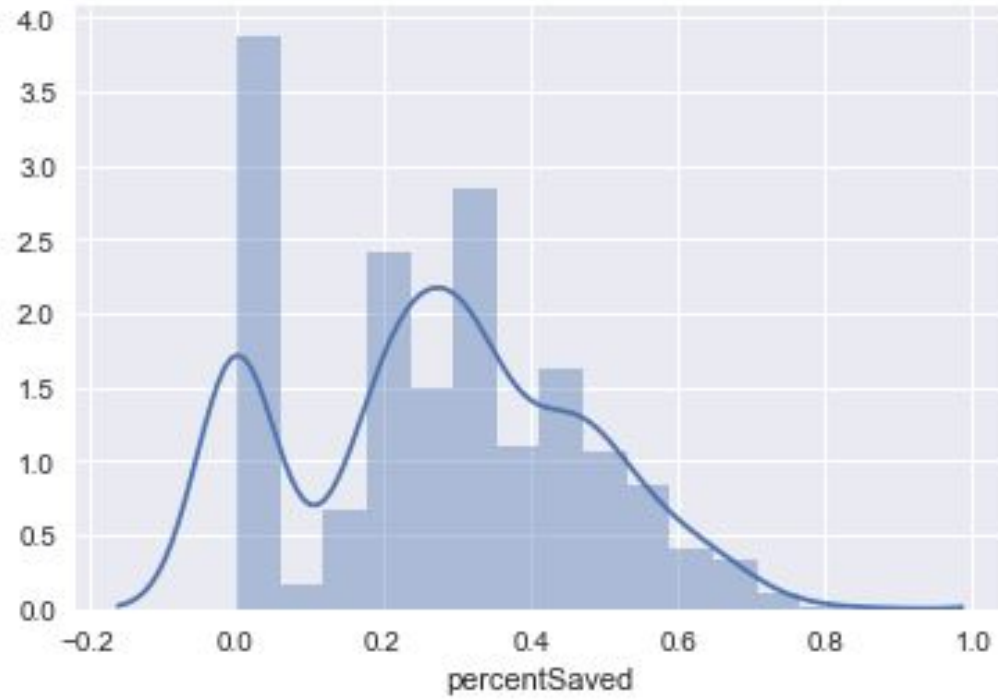


1. Outliers (items valued more than 3 std from mean)
2. Missing data
3. Suggested price lower than sale price (clearly a mistake)

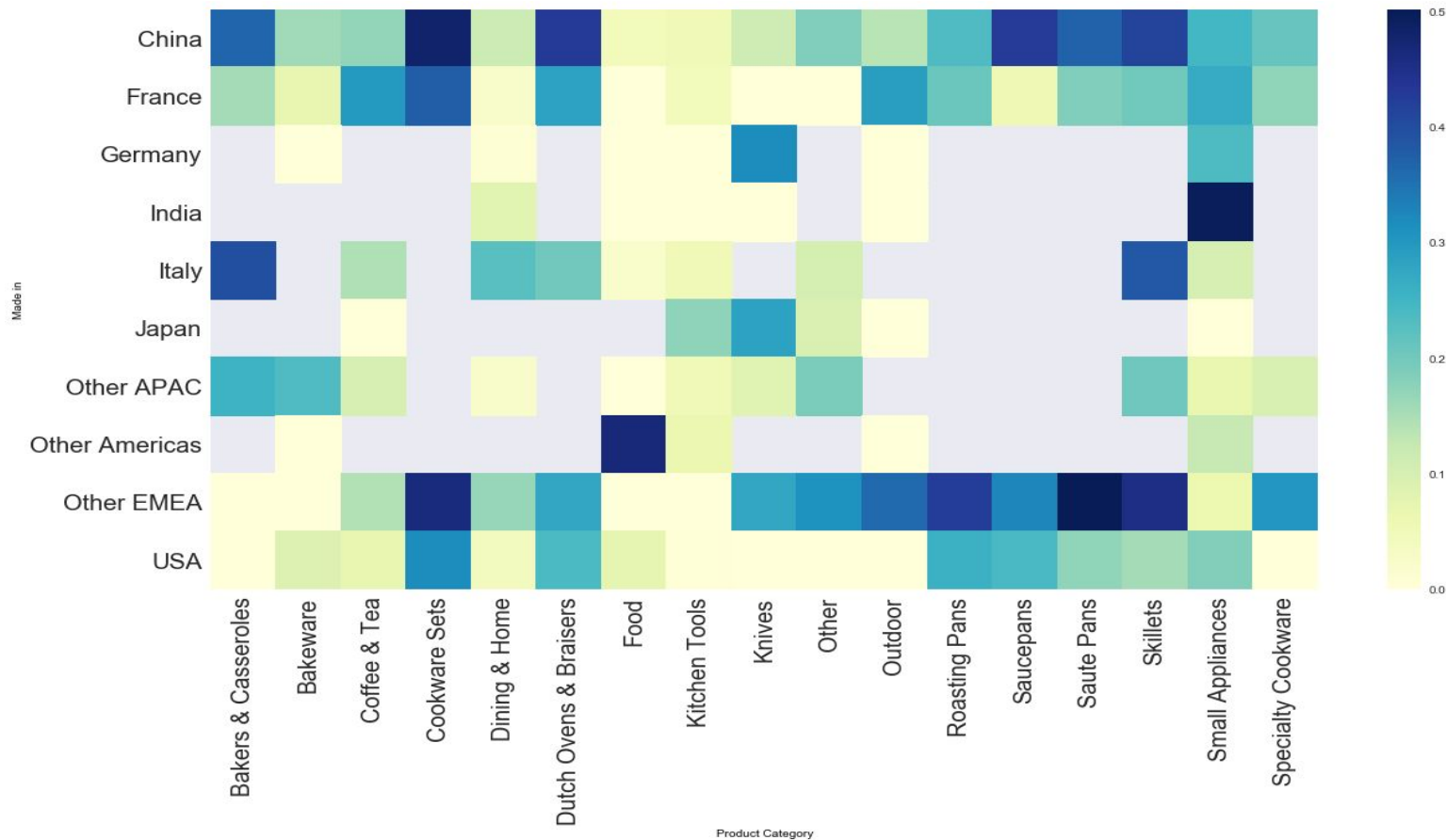
Looking at trends by: product
category, country produced in,
average ratings

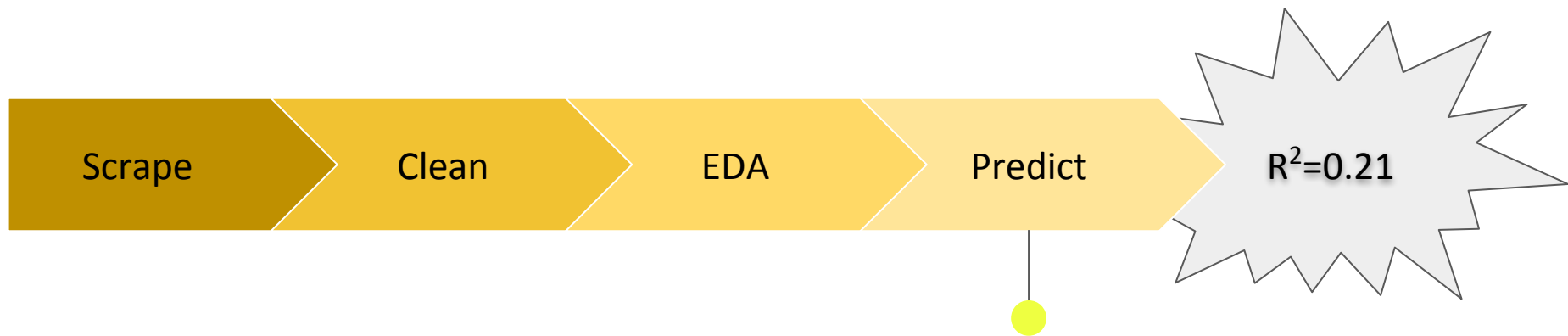


Distribution of Discounts



Discounts by Country and Category





Isolating features that have
predictive power in discounts
using sklearn & statsmodels

Model Methodology

- Plain linear regression resulted in the highest R^2
- Adding polynomial features resulted in extreme overfitting
- Alternated between sklearn's `LinearRegression()` and statsmodels' OLS to narrow down features using p values of coefficients
- Created interaction variables, no positive change
- Created standard deviation variables, no positive change
- Created dummy variables for country, slight positive change!

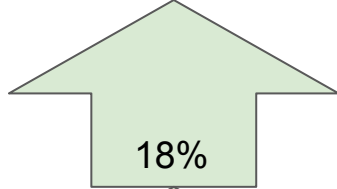


Model Takeaways

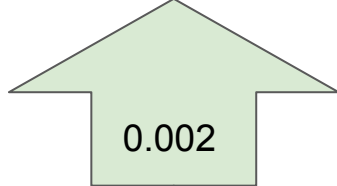
- Predicted a discount when there wasn't one
- Maybe should have introduced other variables into the model (ex: Sur La Table brand vs not)
- Classification might work better (will there be a discount or not)
- Reform the question into: given there IS a discount, how much is it?



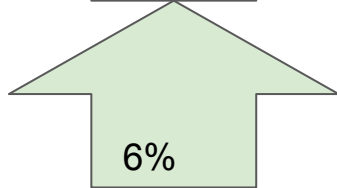
***So if you want a to get a
good discount at Sur La
Table....***



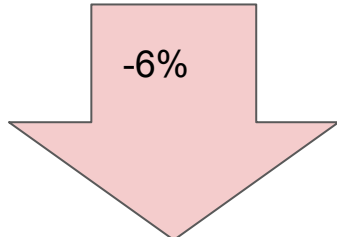
*MSRP



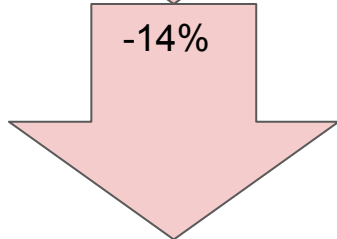
*Answers



*Germany



*France



*USA





Thank you!



Appendix

Scraping

Not pictured:

- Average review
- Product Category
- SKU



Staub Steam Griddle, 12"

★★★★★ 76 Reviews | 24 Questions | 126 Answers

\$149⁹⁶

Sugg. \$357.00

You save 57%

Special Offer:

Receive a free decorative knob by mail (a \$43 value) when you purchase \$200 worth of Staub cookware. Click for details.

COLOR:



QUANTITY:

1

ADD TO CART

FIND IN STORE

ADD TO REGISTRY

ADD TO WISHLIST

Price Guarantee

Hassle-Free Returns

Free Shipping Over \$59 With Code SHIPFREE

Customer Gallery



DESCRIPTION

Inspired by the success of their popular Steam Grill, Staub has developed a brand-new essential for healthy cooking: the cast iron Steam Griddle. The large, flat surface provides plenty of room for browning, searing and warming without much oil—and the enamel interior requires no seasoning, releasing foods easily right out of the box. A large, domed glass lid offers ample space for larger pieces of food like thick steaks and pork chops. Like all of Staub's renowned cast iron cookware, the Steam Griddle offers incredible heat retention and distribution. The tight-fitting glass lid locks in heat and moisture, so you can deglaze and steam-cook foods, sealing in flavorful juices.

With its large size and versatile shape, the Steam Griddle is great for

FEATURES Show

SPECIFICATIONS Hide

- Manufacturer: Staub
- Material: Enameled cast iron, glass
- Dimensions: 12" diameter (15.5" wide with handle) x 1.25" (4.5" tall with lid)
- Warranty: Lifetime warranty
- Made in France

VIDEOS



Staub Steam Griddle Burger