



THE POI NORTHEAST REGIONAL SUMMIT

Achieving Growth Through RGM,
Retail Execution, eCommerce,
and Analytics Excellence



RGM Using Wine Profile Clustering Analysis and Prediction Modeling Using Machine Learning

Sunny Yurasek
Constellation Brands, LLC

- Data is available, **EVERYWHERE**
- Machine learning, AI is not new
 - Airline, hotel industry
 - Volume and scalability
- Too big to fail?
 - Amazon, Forever 21



The Power of Words

“The everyday act of typing a word or phrase into a compact, rectangular white box leaves a small trace of truth that, when multiplied by millions, eventually reveals profound realities.”

– Seth Stephens-Davidowitz **Everybody lies**

Machine learning project via NLP

- What's NLP- Natural Language Processing?
- TF-IDF
- Data sourced from WineMag 130K wine review entries

In [69]: `df_wine.head(20)`

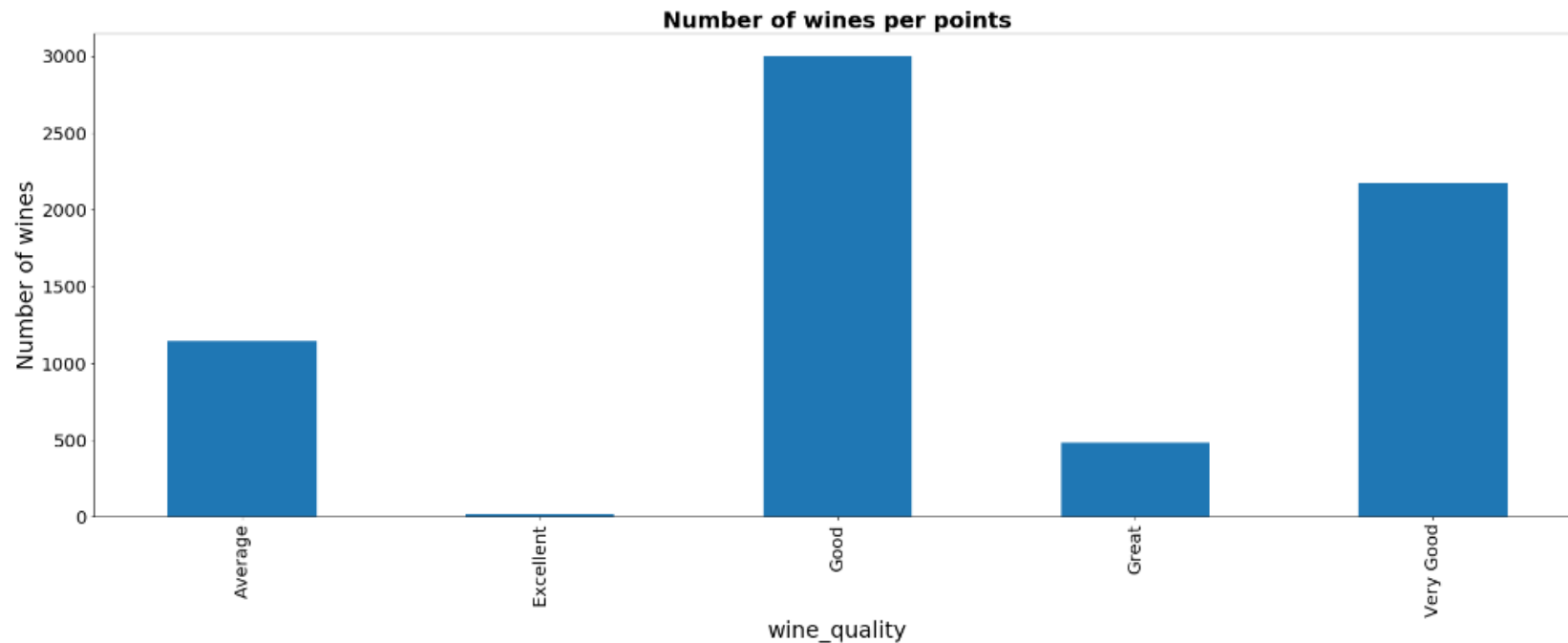
Out[69]:

	country	description	points	price	variety	wine_quality	wine_rank	price_val
39623	France	Made from organically grown grapes and boastin...	89	14.0	Rhône-style Red Blend	Very Good	3	0-20
96742	New Zealand	With only 6.5 g/l of residual sugar, this weig...	90	19.0	Pinot Gris	Very Good	3	0-20
34380	Israel	Fleshy black plum and berry aromas open the bo...	90	40.0	Cabernet Sauvignon	Very Good	3	30-50
47819	Argentina	An intense Malbec-led marauder with rubbery bl...	91	50.0	Red Blend	Very Good	3	30-50
71476	Argentina	Overt oak is the first thing you encounter on ...	88	20.0	Malbec	Good	2	0-20
122812	Italy	Here's a standout Sangiovese-based super Tusca...	92	120.0	Sangiovese	Very Good	3	Above 100
32589	New Zealand	This is an amply endowed, round wine, with no ...	91	40.0	Pinot Noir	Very Good	3	30-50
147428	France	A second label of the famous second growth, Ch...	85	50.0	Bordeaux-style Red Blend	Good	2	30-50
57127	Australia	Heady and superripe, this is a huge, mouthfill...	95	50.0	Shiraz	Great	4	30-50
45056	Argentina	Fiery and clipped on the nose, but then it set...	85	12.0	Cabernet Sauvignon	Good	2	0-20
137840	Italy	Typical of this hot vintage. La Fiammenga is a	87	28.0	Nebbiolo	Good	2	20-30

Machine learning project continued

- Simply across all points into ranking of 5 categories for ease of analytics.

```
Out[70]: <matplotlib.axes._subplots.AxesSubplot at 0x19e4cbe0>
```



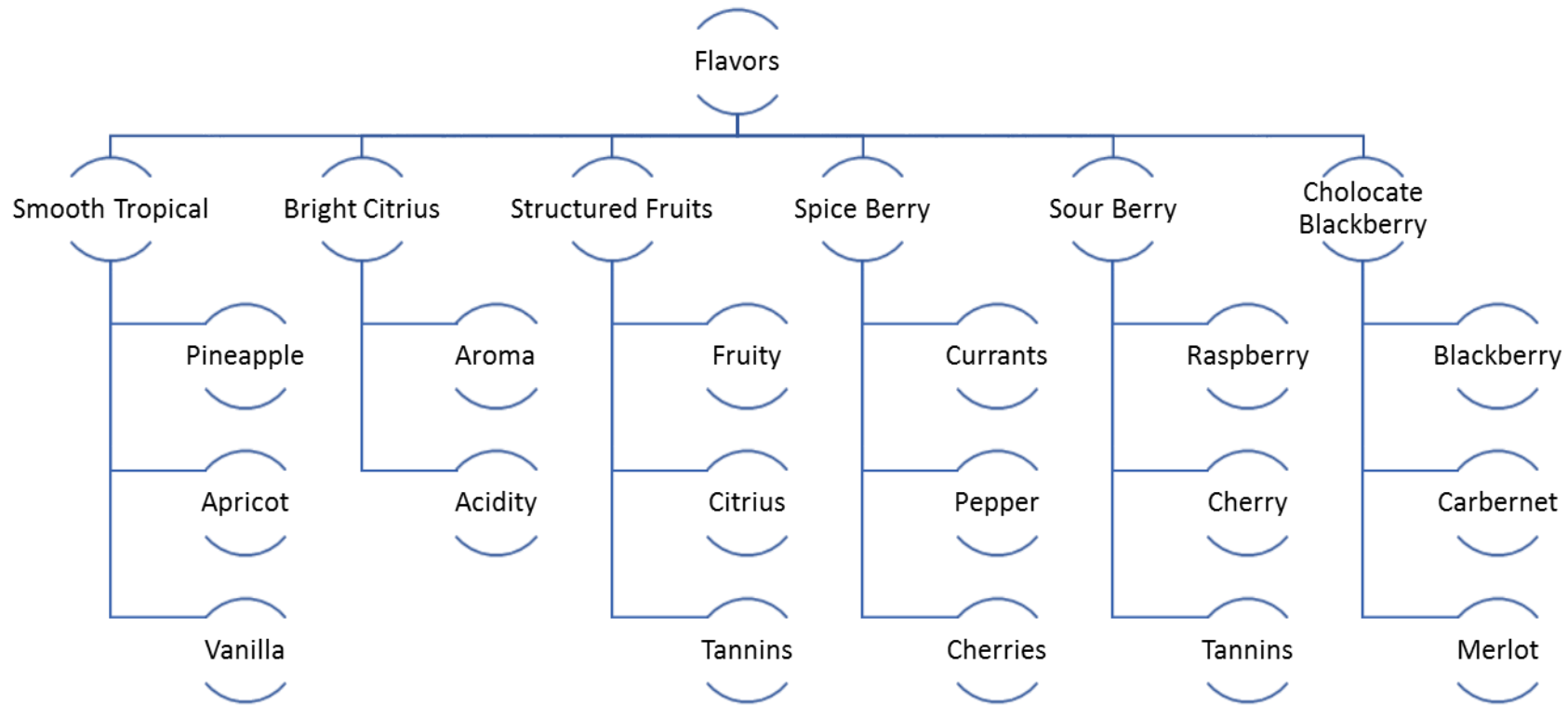
Machine learning project continued...

Data cleaning, stop words limitation, punctuation, and then drop words length >5

Clustering analyses. Iterative process

Acidity Cherry	Tannins Structrued Fruits	Tropical Vanila	Bright Aroma Citrus	Spice Pepper Berry	Chocolate Blackberry
Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
cherry	fruits	pineapple	flavors	cherries	blackberry
raspberry	acidity	chardonnay	aromas	blackberries	cabernet
flavors	tannins	flavors	finish	raspberries	flavors
tannins	character	vanilla	palate	currants	tannins
finish	fruity	acidity	citrus	flavors	cherry
aromas	flavors	apricot	acidity	tannins	merlot
palate	texture	creamy	theres	spices	sauvignon
acidity	structure	tropical	offers	pepper	chocolate
pepper	attractive	buttered	tannins	finish	finish
offers	citrus	finish	bright	acidity	aromas

Machine learning project continued...



Machine learning project continued...

```
In [65]: # Training model
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.1, random_state=101)
rfc = RandomForestClassifier()
rfc.fit(X_train, y_train)

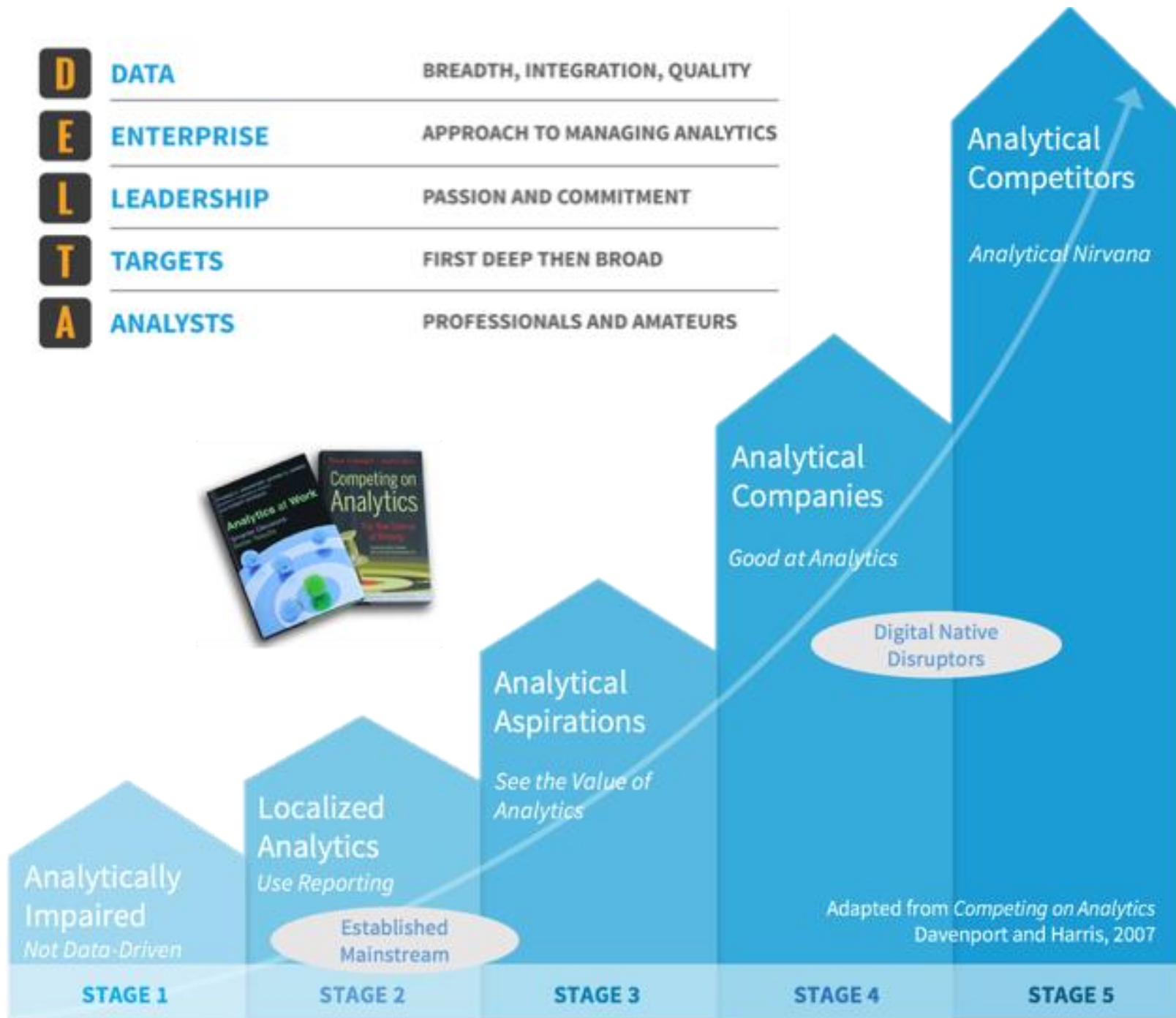
# Testing model
predictions = rfc.predict(X_test)
print(classification_report(y_test, predictions))
```

	precision	recall	f1-score	support
1	0.67	0.45	0.54	459
2	0.55	0.83	0.66	1149
3	0.64	0.46	0.53	908
4	0.76	0.13	0.23	216
5	1.00	0.08	0.14	13
avg / total	0.62	0.59	0.56	2745

REFLECTION:

- Reviews and quality of wine can be further evaluated to improve the accuracy of prediction. Additional models, more data all can be further tested to improve the fitness.
- Understanding and interpreting the results is critical. Able to communicate using storyline will bridge advanced analytics with practical business action

D	DATA	BREADTH, INTEGRATION, QUALITY
E	ENTERPRISE	APPROACH TO MANAGING ANALYTICS
L	LEADERSHIP	PASSION AND COMMITMENT
T	TARGETS	FIRST DEEP THEN BROAD
A	ANALYSTS	PROFESSIONALS AND AMATEURS



Before we run, let's ask ourselves these 4 questions

1. Does XXX see data as a key function as sales, finance, marketing and IT?
2. Is there a sufficient number of data individuals possess modeling and statistical skillset? (aka, data scientists)
3. Are these data scientists rooted with deep understanding of each business functions?
4. Is there a governance structure in place to help stakeholders in making decisions and prioritize opportunities?

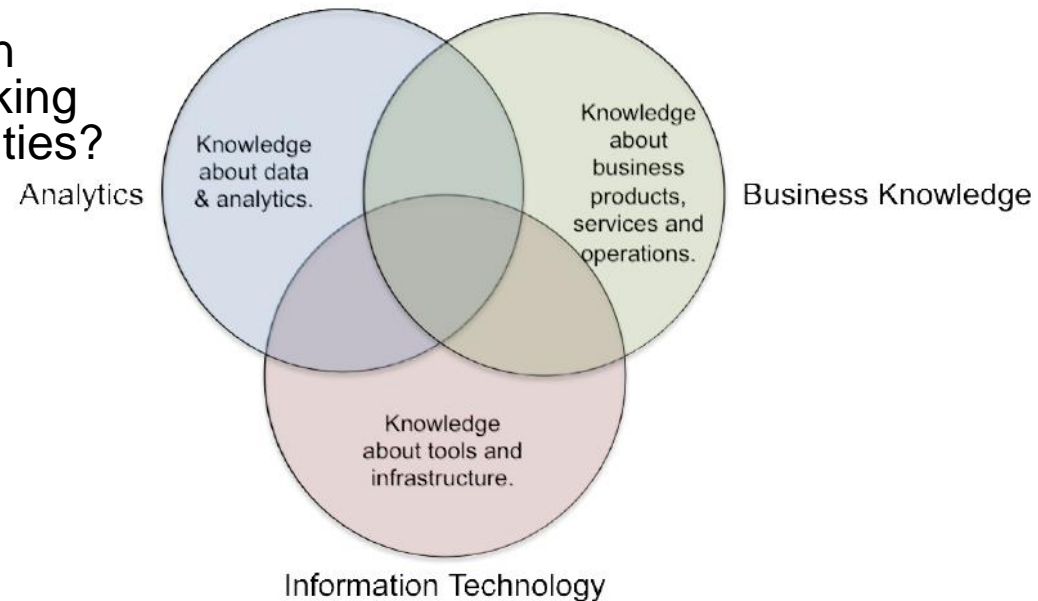


Fig. 1. The knowledge required by data scientists

Sample ML projects

“Your most unhappy customers are your greatest source of learning.”—Bill Gates

Sales/Marketing

1. Consumer Engagement/Personalization TBA
2. Forward trend prediction based on social listening
3. Marketing effectiveness Scenario Planning

HR/Finance

4. Weather temperature effect decomposition
5. HR turnover prediction/At Risk Index

Supply Chain

6. Inventory on hand prediction
7. Production line Preventative maintenance

Winery/Brewery

8. Image recognition for crop health monitoring
9. Targeted water irrigation system
10. Yield prediction

Consumer Engagement/Personalization

- Overview
 - Connecting with our consumer should be 24/7 in the digital age. Research showed consumer who digitally searches product information converts to in-store purchase. CBI is in a unique position to support TBA across beer, wine and spirits portfolio. Platform to bridge is a critical piece to understand. 4 strategies to consider: respond to desire, curated offering, coach behavior, and automatic execution.
- Impact Proposition
 - Digital influence matters, and direct to consumer strategy can improve experiences, boost operational efficiencies and lower the marketing and customer acquisition cost.
 - Instead of push, this can be an interactive process to pull information from consumer and enhance personalization such as **product recommendation** across TBA as well as **flavor personalization** of high-end liquor (high west/tequila/whisky)
 - The rich preference data collected can potentially drive **personalized promotion offering**
- Key Collaborate Stakeholders
 - Strategy, E-Commerce, DTC, Growth
- Reference
 - McGraw-Hill via digital learning , Nike's embedded chips in shoes, Diageo's personalization of Jonnie Walker
 - <https://hbr.org/2019/05/the-age-of-continuous-connection>
 - <https://www.campaignlive.co.uk/article/diageo-digital-chief-using-ai-means-letting-go-creativity-stifling-perfectionism/1465148>

Reality Check

- No leadership buy-in
- Business constraints
- Poor execution of the project
- Lacking good user interface
- Missing clear and measurable KPI's
- Low adoption

What does it all mean to you?

Act

Listen

Learn

Grow!

Q&A