



Analyzing Coffee Reviews

NLP to build recommendations and predictions



Agenda

- Data** 01
Coffee Review Intro
- Natural Language Processing** 02
Cleaning, preparing, and embedding
- Topic Modeling** 03
Non-negative Matrix Factorization
- Visualizations** 04
Seeing insights
- Applet** 05
Recommendations, Modeling, and more

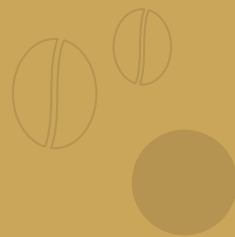




01

Data

www.coffeereview.com





93

Souvenir Coffee

Costa Rica Cloza Estate

Roaster Location: Berkeley, California

Coffee Origin: Alajuela, Costa Rica

Roast Level: Medium-Light

Agtron: 58/80

Est. Price: \$19.00/12 ounces

Review Date: February
2021

Aroma: 9

Acidity
/Structure : 8

Body: 9


Flavor: 9

Aftertaste: 8

Blind Assessment: Crisp, balanced, richly nut-toned. Nutella, red apple, freesia, agave syrup, cedar in aroma and cup. Very sweet in structure with brisk acidity; plush, syrupy-smooth mouthfeel. Sweetly nut-toned finish supported by freesia-like floral tones.

Notes: Produced by Kathia Zamora of Cloza Estate from trees of the Caturra and Catuai varieties of Arabica and processed by the washed method. Souvenir Coffee, based in Berkeley, California, goes by the motto, “People first, then coffee.” Visit www.souvenir-coffee.com for more information.

The Bottom Line : A pleasing, delicately spice-toned, richly nutty Costa Rica cup, classically balanced.





93

Souvenir Coffee

Costa Rica Cloza Estate

Roaster Location: Berkeley, California

Coffee Origin: Alajuela, Costa Rica

Roast Level: Medium-Light

Agtron: 58/80

Est. Price: \$19.00/12 ounces

Review Date: February
2021

Aroma: 9

Acidity 8

/Structure :

Body: 9


Flavor: 9

Aftertaste: 8

Blind Assessment: Crisp, balanced, richly nut-toned. Nutella, red apple, freesia, agave syrup, cedar in aroma and cup. Very sweet in structure with brisk acidity; plush, syrupy-smooth mouthfeel. Sweetly nut-toned finish supported by freesia-like floral tones.

Notes: Produced by Kathia Zamora of Cloza Estate from trees of the Caturra and Catuai varieties of Arabica and processed by the washed method. Souvenir Coffee, based in Berkeley, California, goes by the motto, “People first, then coffee.” Visit www.souvenir-coffee.com for more information.

The Bottom Line : A pleasing, delicately spice-toned, richly nutty Costa Rica cup, classically balanced.





93

Souvenir Coffee

Costa Rica Cloza Estate

Roaster Location: Berkeley, California
Coffee Origin: Alajuela, Costa Rica
Roast Level: Medium-Light
Agtron: 58/80
Est. Price: \$19.00/12 ounces

Review Date:	February 2021
Aroma:	9
Acidity	8
/Structure :	
Body:	9
Flavor:	9
Aftertaste:	8

Blind Assessment: Crisp, balanced, richly nut-toned. Nutella, red apple, freesia, agave syrup, cedar in aroma and cup. Very sweet in structure with brisk acidity; plush, syrupy-smooth mouthfeel. Sweetly nut-toned finish supported by freesia-like floral tones.

Notes: Produced by Kathia Zamora of Cloza Estate from trees of the Caturra and Catuai varieties of Arabica and processed by the washed method. Souvenir Coffee, based in Berkeley, California, goes by the motto, “People first, then coffee.” Visit www.souvenir-coffee.com for more information.

The Bottom Line : A pleasing, delicately spice-toned, richly nutty Costa Rica cup, classically balanced.



93

Souvenir Coffee

Costa Rica Cloza Estate

Roaster Location: Berkeley, California

Coffee Origin: Alajuela, Costa Rica

Roast Level: Medium-Light

Agtron: 58/80

Est. Price: \$19.00/12 ounces

Review Date: February
2021

Aroma: 9

Acidity 8

/Structure :

Body: 9


Flavor: 9

Aftertaste: 8

Blind Assessment: Crisp, balanced, richly nut-toned. Nutella, red apple, freesia, agave syrup, cedar in aroma and cup. Very sweet in structure with brisk acidity; plush, syrupy-smooth mouthfeel. Sweetly nut-toned finish supported by freesia-like floral tones.

Notes: Produced by Kathia Zamora of Cloza Estate from trees of the Caturra and Catuai varieties of Arabica and processed by the washed method. Souvenir Coffee, based in Berkeley, California, goes by the motto, “People first, then coffee.” Visit www.souvenir-coffee.com for more information.

The Bottom Line : A pleasing, delicately spice-toned, richly nutty Costa Rica cup, classically balanced.



Blind Assessment: Crisp, balanced, richly nut-toned. Nutella, red apple, freesia, agave syrup, cedar in aroma and cup. Very sweet in structure with brisk acidity; plush, syrupy-smooth mouthfeel. Sweetly nut-toned finish supported by freesia-like floral tones.

The background is a solid dark teal color. It is decorated with several circles and capsule shapes. In the top left, there are two small white-outlined capsules and one large solid dark teal circle. In the top right, there is one large solid dark teal circle. In the bottom left, there is one large solid dark teal circle. In the bottom right, there is one large solid dark teal circle, one medium white-outlined capsule, and one small white-outlined capsule.

5959

Blind Reviews

Natural Language Processing

02





Goal

Turn the words in a review into a numerical embedding

Capture word choice, frequency with a review, uniqueness across reviews

Avoid common words and errors

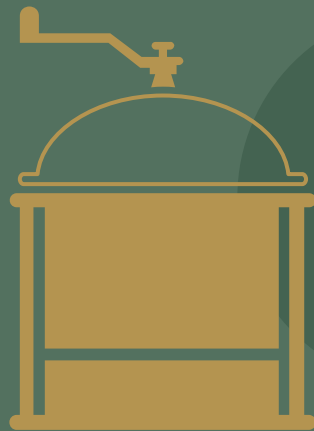
Cleaning



03

Topic Modeling

Non-negative Matrix Factorization



NMF

General Overview

- Create potential clusters of reviews
- Assign each review a score for assignment to clusters
- Find meaning in clusters to understand topics
- By-product: Reduce each review to a smaller feature space

NMF

Process

- Perform NMF fit and transform
 - Across various numbers of topics
 - Across numerous word embeddings
 - Visualize using scree plots
- Inspect most important words in topics to ascertain meaning



04

Visualizations

Origin and Roast Level by Group

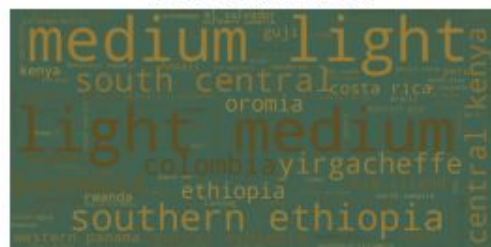
Bright, Floral, Citrus



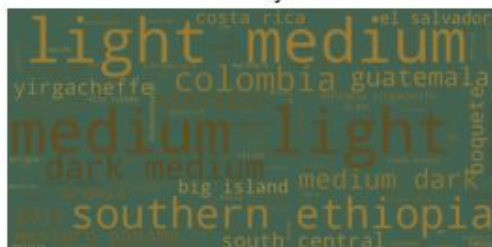
Chocolate, Dark, Woody



Tart, Sweet, Smooth



Cacao, Nutty, Clean



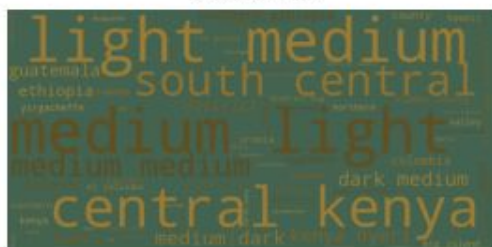
Sweet, Nut, Pine



Juicy, Honey, Cacao



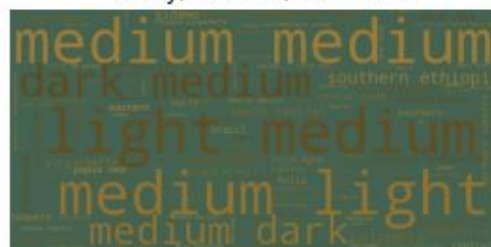
Red Berries



Nutty, Caramel, Woody



Cherry, Vinuous, Chocolate



Reviews by Groups

Bright, Floral, Citrus



Chocolate, Dark, Woody



Tart, Sweet, Smooth



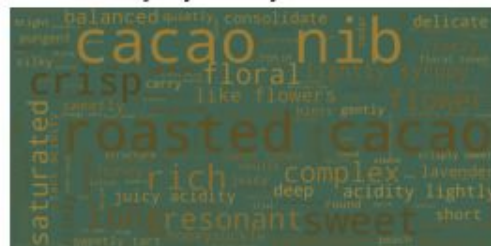
Cacao, Nutty, Clean



Sweet, Nut, Pine



Juicy, Honey, Cacao



Red Berries



Nutty, Caramel, Woody

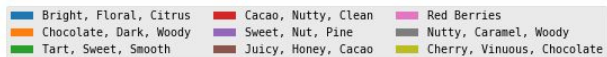


Cherry, Vinuous, Chocolate

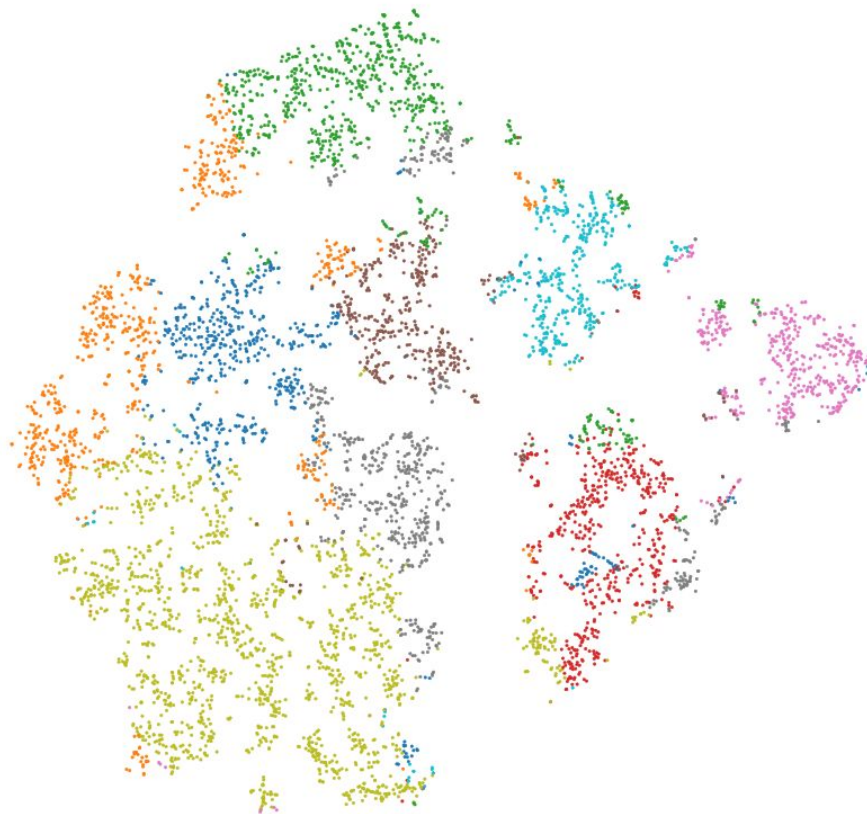


All Coffee Reviews Clustered by Category

Unsupervised topic modeling with NMF based on textual content + 2D-embedding with t-SNE:



Blind Reviews from Coffee-Review.com



05

Applet

Recommendations, Predictions, and More



Coffee Recommendations

Find Similar
Reviews



NMF Vector
Creates a five
dimensional
representation of
each review

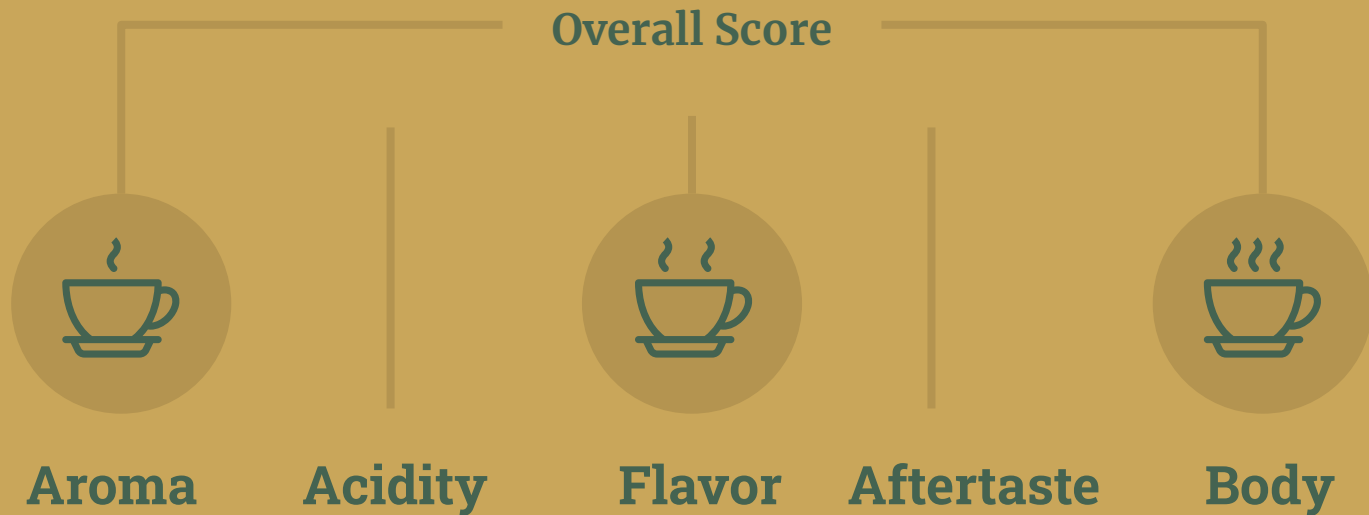


Pairwise Distance
Use cosine similarity to
compare reviews to each
other



Recommendation
Pick the review that is
most like the input

Reviews to Overall Score



Subcategories to Overall Score



Reviews to Subcategories

Aroma
Body
Flavor
Aftertaste
Acidity



NMF Vector
As feature vector



Review Length
Word count and number
of different words



Sentiment Analysis
Vader
Positive, Negative,
Compound



93

Souvenir Coffee

Costa Rica Cloza Estate

Roaster Location: Berkeley, California

Coffee Origin: Alajuela, Costa Rica

Roast Level: Medium-Light

Agtron: 58/80

Est. Price: \$19.00/12 ounces

Review Date: February
2021

Aroma: 9

Acidity
/Structure : 8

Body: 9


Flavor: 9

Aftertaste: 8

Blind Assessment: Crisp, balanced, richly nut-toned. Nutella, red apple, freesia, agave syrup, cedar in aroma and cup. Very sweet in structure with brisk acidity; plush, syrupy-smooth mouthfeel. Sweetly nut-toned finish supported by freesia-like floral tones.

Notes: Produced by Kathia Zamora of Cloza Estate from trees of the Caturra and Catuai varieties of Arabica and processed by the washed method. Souvenir Coffee, based in Berkeley, California, goes by the motto, “People first, then coffee.” Visit www.souvenir-coffee.com for more information.

The Bottom Line : A pleasing, delicately spice-toned, richly nutty Costa Rica cup, classically balanced.



Next up: Test

g1-M

- ☐ Home
- ☐ Recommendations
- ☐ Score from Text
- ☒ Score from Score
- ☐ Generated Reviews

Score Predictor (if you have the details)

Predict an overall score from subcategories

acidity



body



balance



brightness



mouth



With subcategory scores as shown above, I predict your coffee to be **overall 8.0**

The background is a solid dark teal color. It is decorated with several geometric shapes: large solid dark teal circles, smaller solid dark teal circles, and several white-outlined circles with a vertical line through the center, resembling capsules or pills. These shapes are scattered around the central text.

Questions?



93

Souvenir Coffee

Costa Rica Cloza Estate

Roaster Location: Berkeley, California

Coffee Origin: Alajuela, Costa Rica

Roast Level: Medium-Light

Agtron: 58/80

Est. Price: \$19.00/12 ounces

Review Date: February
2021

Aroma: 9

Acidity 8

/Structure :

Body: 9


Flavor: 9

Aftertaste: 8

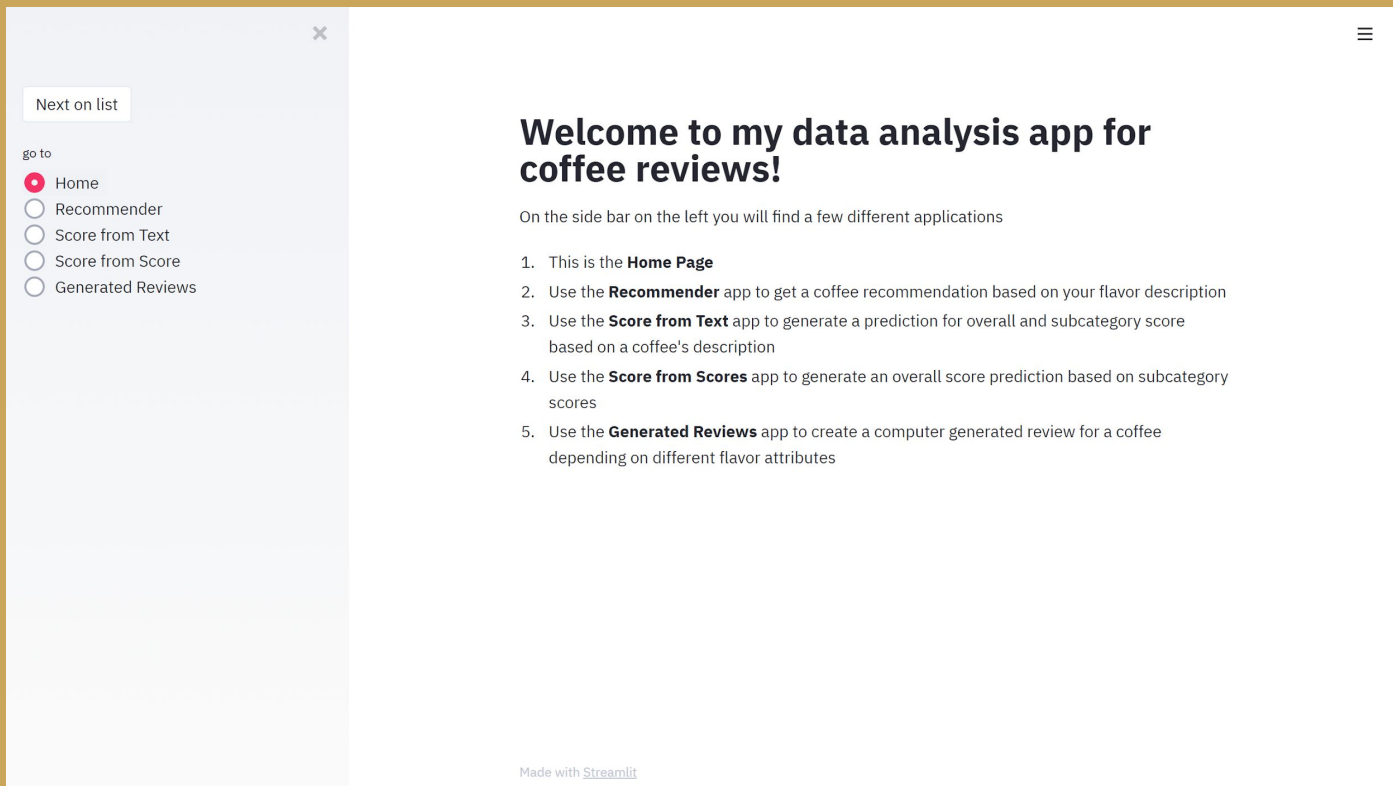
Blind Assessment: Crisp, balanced, richly nut-toned. Nutella, red apple, freesia, agave syrup, cedar in aroma and cup. Very sweet in structure with brisk acidity; plush, syrupy-smooth mouthfeel. Sweetly nut-toned finish supported by freesia-like floral tones.

Notes: Produced by Kathia Zamora of Cloza Estate from trees of the Caturra and Catuai varieties of Arabica and processed by the washed method. Souvenir Coffee, based in Berkeley, California, goes by the motto, “People first, then coffee.” Visit www.souvenir-coffee.com for more information.

The Bottom Line : A pleasing, delicately spice-toned, richly nutty Costa Rica cup, classically balanced.



Streamlit



Applet Recommendation

Coffee Recommender

Get a new coffee recommendation

Give a couple sentences here of how you describe your ideal coffee. Try to include as much as you can about your desired flavor profile.

Excited to recommend a coffee for you!

Applet Recommendation

Coffee Recommender

Get a new coffee recommendation

Give a couple sentences here of how you describe your ideal coffee. Try to include as much as you can about your desired flavor profile.

Crisp, balanced, richly nut-toned. Nutella, red apple, freesia, agave syrup, cedar in aroma and

Based on your input coffee, I recommend you try:

Small Eyes Cafe who roast a bean from Santa Barbara, Honduras .

It could be described as:

Crisply sweet, nut-toned, richly floral. Almond butter, lilac, lemon verbena, red apple, oak in aroma and cup. Sweet structure with soft, round acidity; very full, syrupy-smooth mouthfeel. Lemon verbena supports the resonantly nut-toned finish.

Score Predictions

Score Predictor

Predict coffee scores from reviews

Provide a couple sentence descripton of the flavors, acid level, aroma, aftertaste, and body of your coffee.

|

Excited to predict the score of your coffee!

Score Prediction

Score Predictor

Predict coffee scores from reviews

Provide a couple sentence descripton of the flavors, acid level, aroma, aftertaste, and body of your coffee.

Crisp, balanced, richly nut-toned. Nutella, red apple, freesia, agave syrup, cedar in aroma and

Based on your input coffee, I predict it to receive a score of: 92.32

An aroma score of (out of 10): 8.59

An acidity score of (out of 10): 8.24

An aftertaste score of (out of 10): 8.06

A flavor score of (out of 10): 8.72

A body score of (out of 10): 8.21

92.3

93

Souvenir Coffee

Costa Rica Cloza Estate

Roaster Location: Berkeley, California

Coffee Origin: Alajuela, Costa Rica

Roast Level: Medium-Light

Agtron: 58/80

Est. Price: \$19.00/12 ounces

Review Date: February
2021

Aroma: 9

Acidity 8

/Structure : 9

Body: 9

Flavor: 9

Aftertaste: 8

8.59

8.24

8.06

8.72

8.21

Blind Assessment: Crisp, balanced, richly nut-toned. Nutella, red apple, freesia, agave syrup, cedar in aroma and cup. Very sweet in structure with brisk acidity; plush, syrupy-smooth mouthfeel. Sweetly nut-toned finish supported by freesia-like floral tones.

Notes: Produced by Kathia Zamora of Cloza Estate from trees of the Caturra and Catuai varieties of Arabica and processed by the washed method. Souvenir Coffee, based in Berkeley, California, goes by the motto, “People first, then coffee.” Visit www.souvenir-coffee.com for more information.

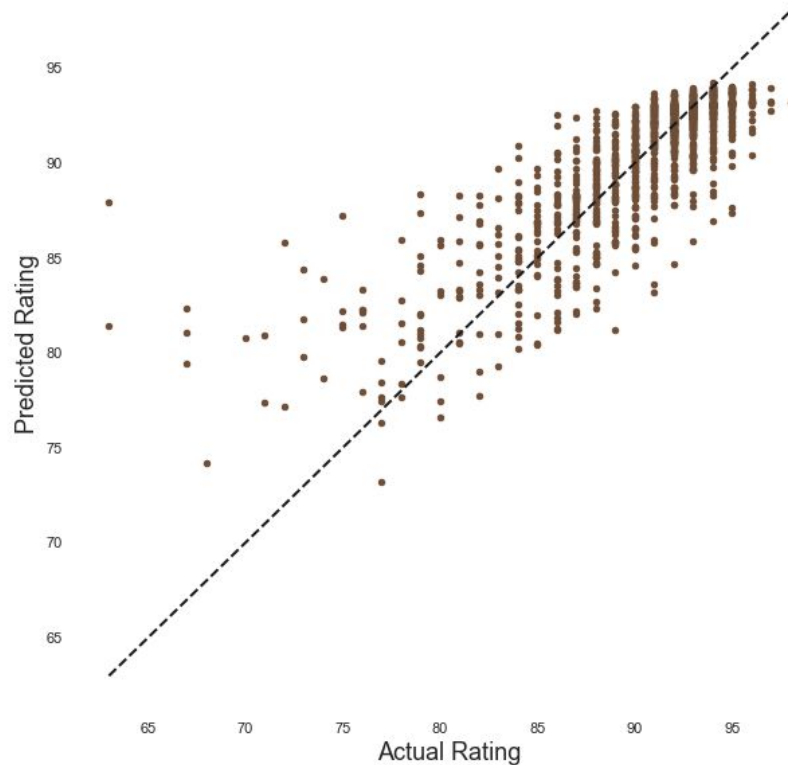
The Bottom Line : A pleasing, delicately spice-toned, richly nutty Costa Rica cup, classically balanced.

RFR Model (from text)

$$r^2 = .688$$

MAE = 1.7 points

Can we predict coffee ratings from reviews using a RandomForest?

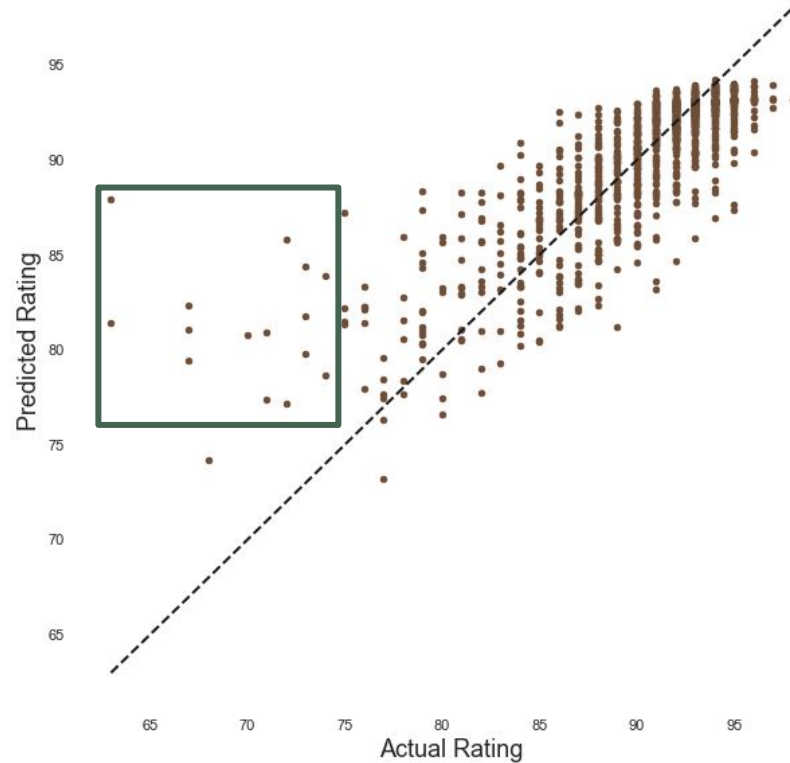


RFR Model (from text)

Errors:

Negative reviews
Older reviews
Instant Coffee

Can we predict coffee ratings from reviews using a RandomForest?

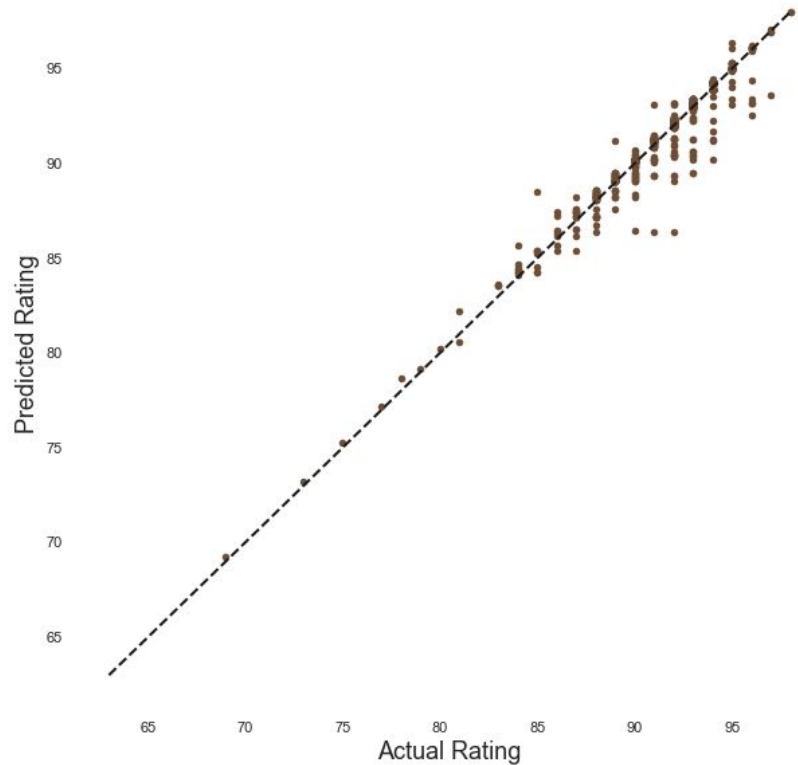


Linear Model (from scores)

$$r^2 = .943$$

$$\text{MAE} = .36 \text{ points}$$

How does the predicted rating of a coffee compare to reality?



Linear Model (from scores)

Coefficients

Aroma = 1.04

Body = 0.75

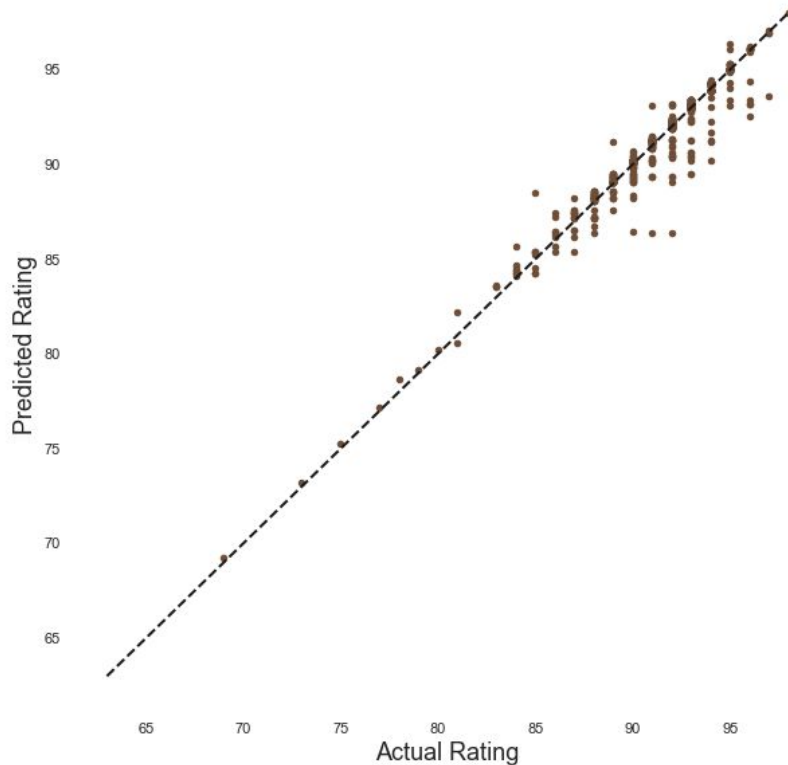
Flavor = 1.00

Aftertaste = 1.14

Acidity = 0.89

Intercept = 51.65

How does the predicted rating of a coffee compare to reality?



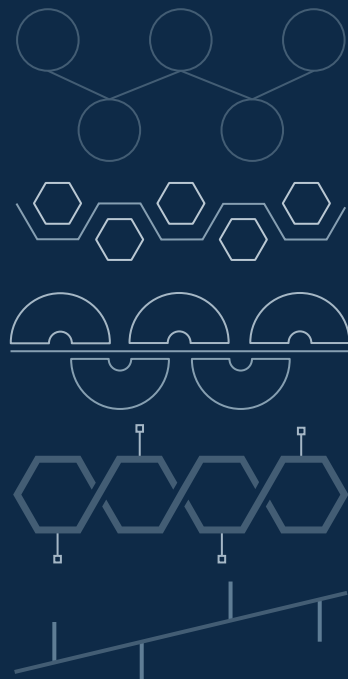
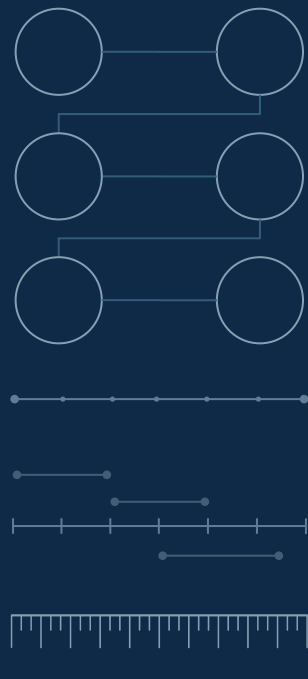
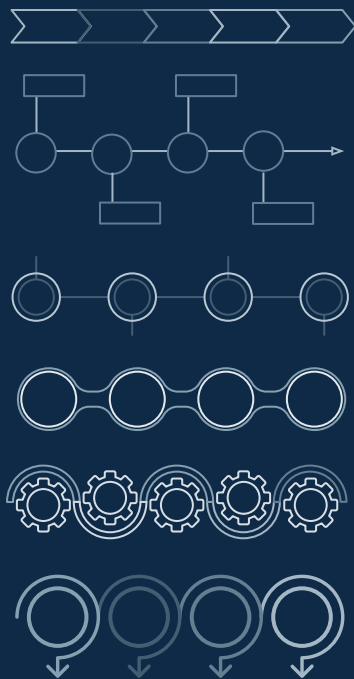
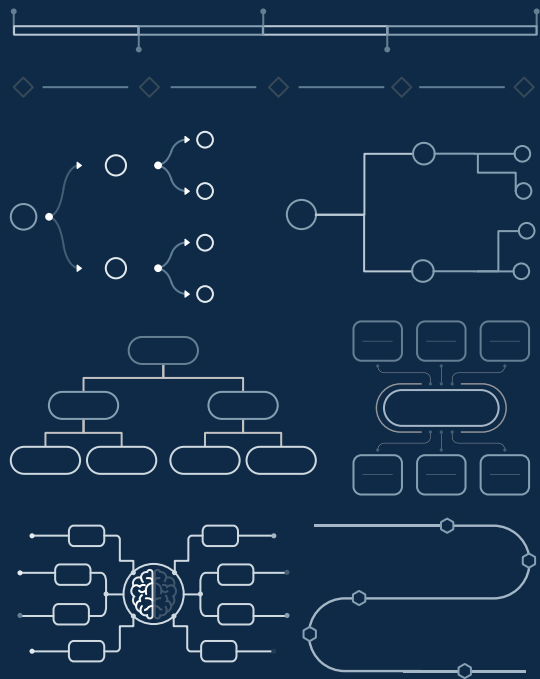
Use our editable graphic resources...

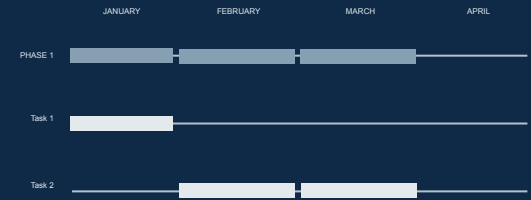
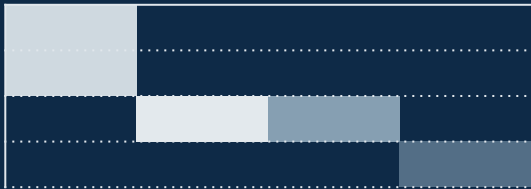
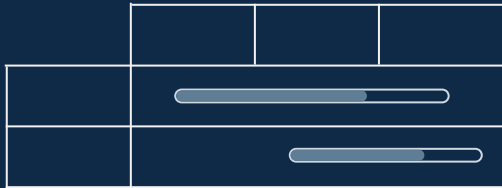
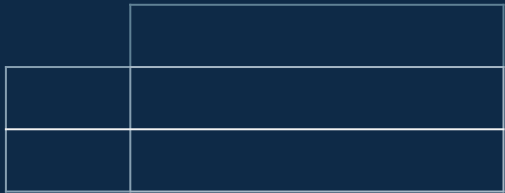
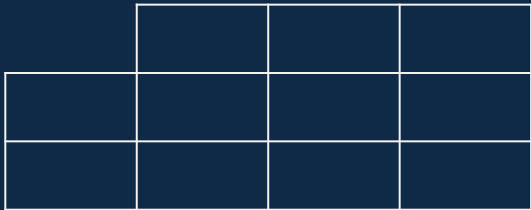
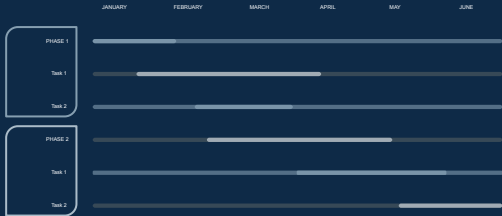
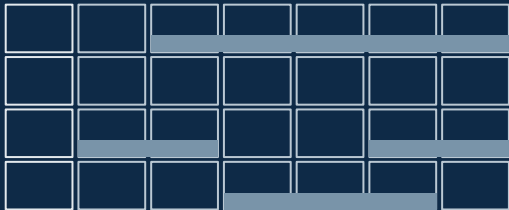
You can easily [resize](#) these resources without losing quality. To [change the color](#), just ungroup the resource and click on the object you want to change. Then, click on the paint bucket and select the color you want.

Group the resource again when you're done. You can also look for more [infographics](#) on [Slidesgo](#).

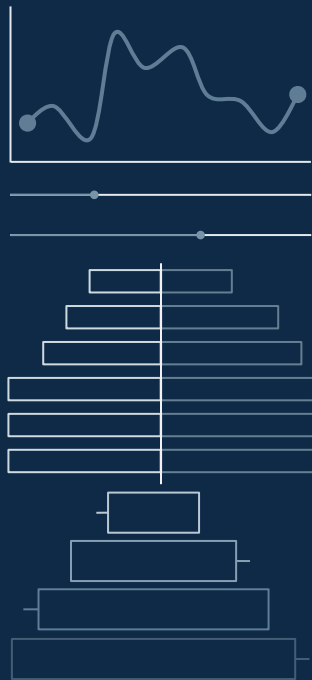
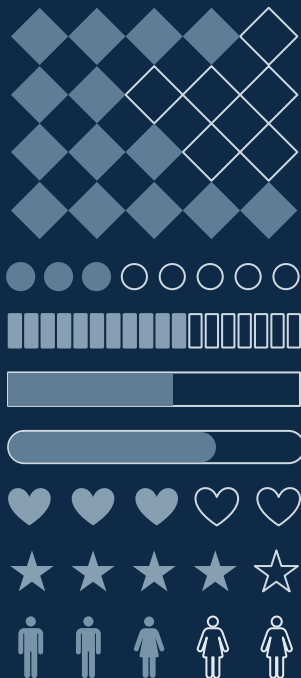
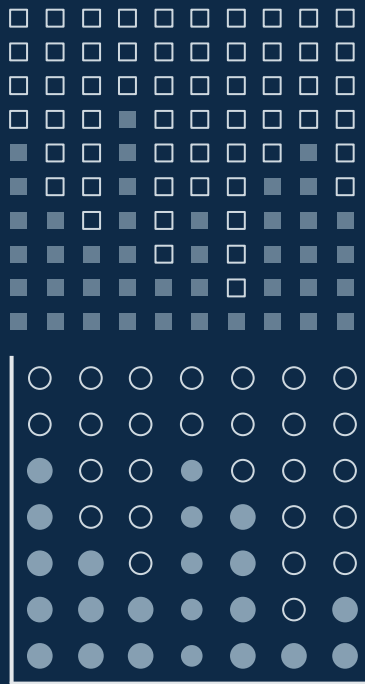












...and our sets of editable icons

You can resize these icons, keeping the quality.

You can change the stroke and fill color; just select the icon and click on the paint bucket/pen.

In Google Slides, you can also use Flaticon's extension, allowing you to customize and add even more icons.



Educational Icons



Medical Icons



Business Icons



Teamwork Icons



Help & Support Icons



Avatar Icons



Creative process Icons



Performing Arts Icons



Nature Icons



SEO & Marketing Icons



