



Lifestyle Grocery Recommender

By: Nicholas Vazquez





Hello!

I'm Nicholas an Economist
turned Data Scientist
currently based in Miami. Feel
free to contact me if you find
yourself enjoying any part of
this presentation.
Github on the last slide.

1

Start With Why?

Let's start with the goal of this project.



Start with why?: Goal

So I set out to build a new way to find your alternative favourite food that fits your lifestyle. A lifestyle grocery recommender with a website that allows visitors to input a paragraph describing the food they wished they had in their current lifestyle choice.



Project process steps

Data process

Create a sample dataset. Decide on lifestyle filters for products.
Create data pipeline for product reviews and then user input text.

Model

Using TF-IDF vectorizer for product reviews and the user input, I will use Cosine Similarity to judge the similarity of the vectors to return the 10 best products a user should try next.

Flask app

Using Flask I will create a website that allows users to input a paragraph and I will return recommendations in real time on a table with product name and a link to the product on amazon.

2

Data Processing

Let's start processing
the data



5,074,160

Product Reviews

297,209

Distinct Products

98,532 users

Over 15 years!

3

Filtering the data

Building a lifestyle grocery recommender requires going through each item and seeing if they are part of the lifestyle group.



Filtering the Data: Lifestyles

Vegetarian

Low sugar

Organic

Paleo

Gluten free

Non gmo

Keto

Pescaterian

Kosher

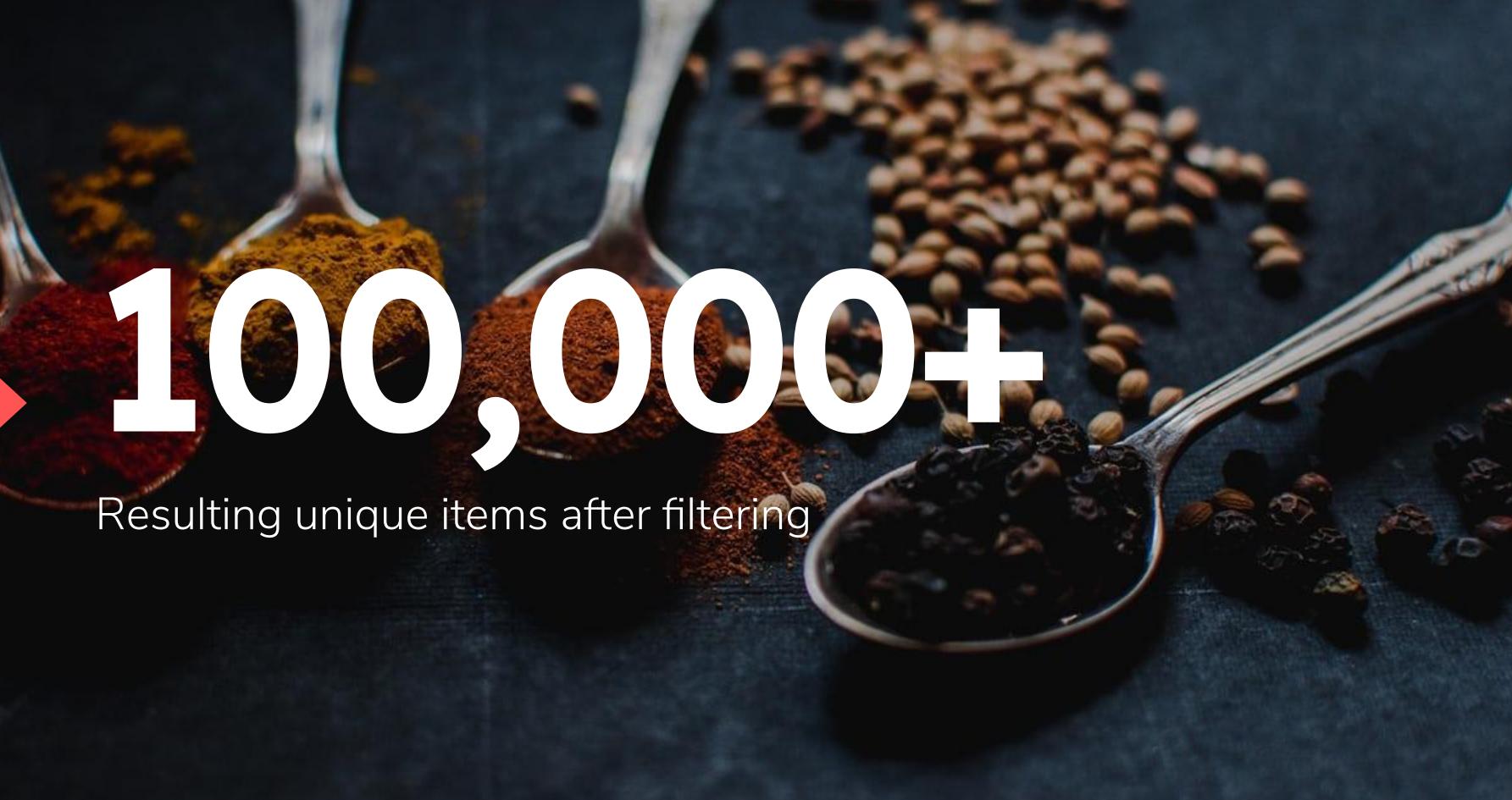
Vegan

Low fat

Halal

Ovo lacto

Fat free



100,000+

Resulting unique items after filtering

4

Text Processing

Now I will do some natural language processing on my filtered tables.



Random sampling

Before I do this I am going to work with only a random sample of 150,000 reviews due to the constraints of my computer ram.



Text Processing

“5while” → “while”

- Custom function to remove numbers in words.

“sunday’s” →

“sunday”

- Drop all english punctuations

“Abandonment” →

“Abandon”

- Then I removed the stem from words

“I love data” →

“I”, “love”, “data”

- Tokenized and dropped stop words



“

'i', 'make', 'claims',
'sports', 'nutrition', 'or',
'nutrition', 'matter',
'expertise', 'i', 'call',
'snack', 'however', 'i',
'say', 'much', 'better',
'average', 'tasting',
'sports', 'bar', 'snack'



Text Processing

Once I had this data clean I ran the TF-IDF. (Term Frequency - Inverse Document Frequency)



Big concept

What does the TF-IDF vectorizer mean for his model?

This means that words show up a lot in our reviews, I will divide it by a large number since it doesn't mean a lot if it's in most of the reviews. On the other hand, if a word pops up very infrequently in these reviews it's probably much more important.



5

Connecting pipelines

Next step was to move my code into a pipeline that would work for the final flask app.



Connecting Pipelines

This includes:

- Getting my pipeline into concise python files I can call from my app.
- Running my Text pipeline once and saving that to a .json so I don't have to re-use computing power for the most intensive part of the app.
- Running the TF-IDF once on AWS and saving it since I want to be able to recommend to users in real-time and training the model takes a couple of minutes.
- Allow a user to input text and have it run through the pipeline and be served recommendations



6

Content-based Recommender

The final step before being able to serve recommendations to the user



Content Based Recommender

Once the pipeline was done for incoming text and reviews was vectorized the final step was to compare the vectors using cosine similarity.

So why cosine similarity?

- Curse of dimensionality. So many features. So euclidean distance is a bad choice cause everything will be far apart. But, cosine similarity would be best because it uses angles so it's a better choice for sparse matrices.



7

Flask app

Flask is an easy and quick way to deploy websites using python



Final app

Lifestyle Grocery-list Recommender

Find new and exciting products that suit your lifestyle

Find Product Recommendations Alternatives to Products You Already Love

First version of website

Lifestyle Grocery-list Recommender Home About Alternative Product Recommender Content Based Recommender

Lifestyle Grocery-list Recommender

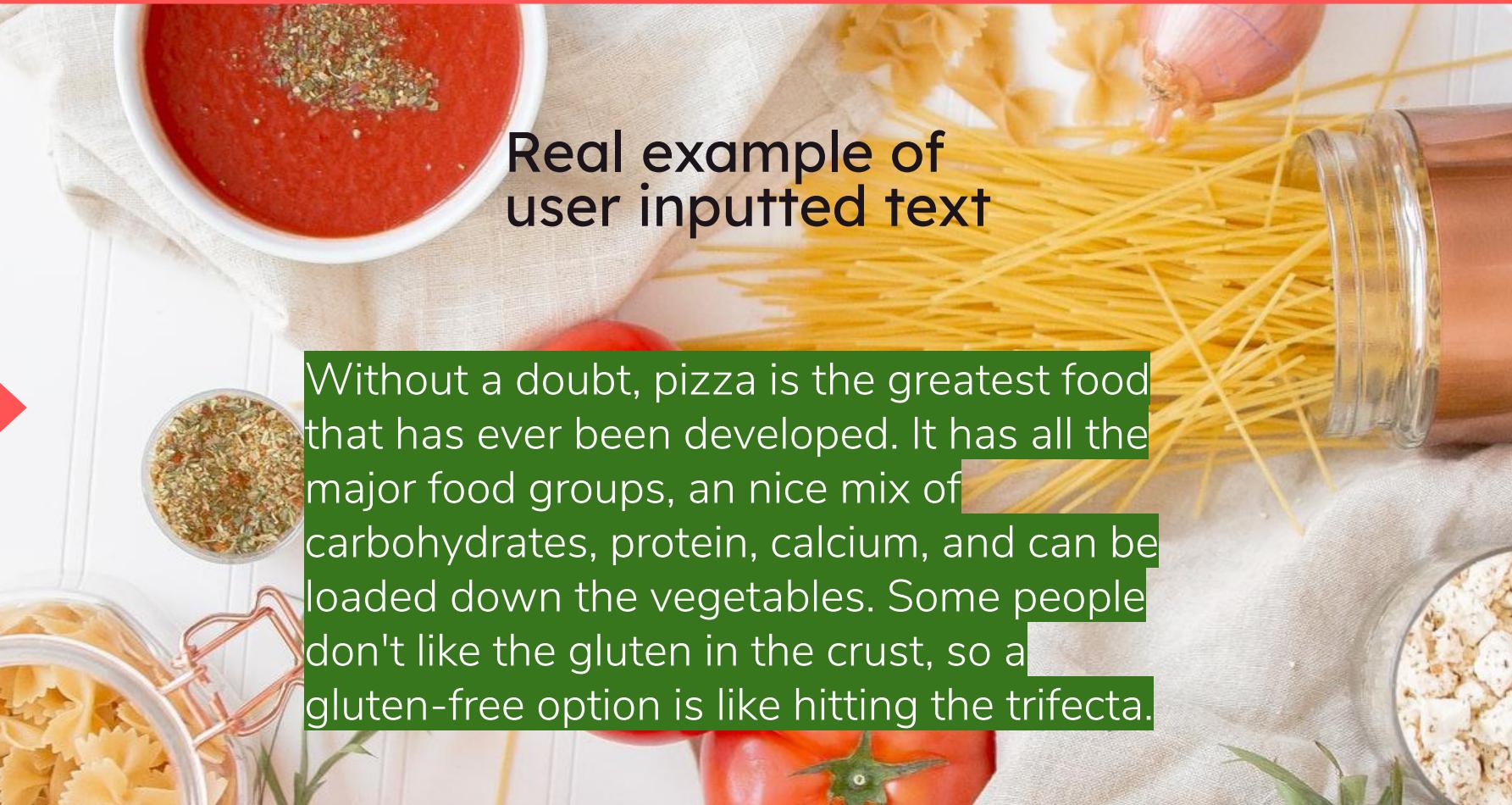
Find new and exciting products
that suit your lifestyle

[Find Product Recommendations](#)

[Alternatives to Products you already love](#)

Home About EDA Content based recommender





Real example of user inputted text

Without a doubt, pizza is the greatest food that has ever been developed. It has all the major food groups, a nice mix of carbohydrates, protein, calcium, and can be loaded down the vegetables. Some people don't like the gluten in the crust, so a gluten-free option is like hitting the trifecta.

Final app

Your Input :

Without a doubt, pizza is the greatest food that has ever been developed. It has all the major food groups, an nice mix of carbohydrates, protein, calcium, and can be loaded down the vegetables. Some people don't like the gluten in the crust, so a gluten-free option is like hitting the trifecta.

Recommendations

> > > > > > >

#	name	Url
1	Betty Crocker Gluten Free Pizza Crust Mix 10.5 oz Box	http://www.amazon.com/dp/B00LM8PH02
2	Betty Crocker Bisquick Baking Mix, Gluten Free Pancake and Baking Mix, 16 Oz Box (Pack of 3)	http://www.amazon.com/dp/B004391DK0
3	Katz Gluten Free Small Pizza Crust 5" Dairy, Nut and Gluten Free Kosher (6 Packs of 4 Crusts, 8.5 Ounce Each)	http://www.amazon.com/dp/B0064NN97S
4	Simple Mills Almond Flour Mix, Pumpkin Muffin & Bread, Naturally Gluten Free, 9 oz	http://www.amazon.com/dp/B00DB8KKDK
5	Bold Organics Meat Lovers Pizza, 12.5 Ounce (Frozen)	http://www.amazon.com/dp/B00D85T4ZQ
6	Chebe Bread Pizza Crust Mix, Gluten Free, 7.5-Ounce Box (Pack of 8)	http://www.amazon.com/dp/B001ACMCNA
7	Udi's Gluten Free Gourmet Pizza Crust, Dairy Free, Nut Free, 2 Pack (Frozen)	http://www.amazon.com/dp/B0036C3TR0
8	Inspiration Mixes Gluten Free Alannah's Cookie Mix, 13 Ounce	http://www.amazon.com/dp/B00AERCMRI
9	Bob's Red Mill Gluten Free All-Purpose Baking Flour, 44-ounce (Pack of 4)	http://www.amazon.com/dp/B000KEPBGS

Next Moves:

- N - grams to expand vocabulary
- Recommender using a hybrid model
- Using different filters for finding alternative foods when you have a food in mind.



Thanks!

Any questions?

You can find me at:

- [Github.com/nicoivazquez](https://github.com/nicoivazquez)
- nicoletovazquez@gmail.com



Final app

Lifestyle Grocery-list Recommender

Find new and exciting products that suit your lifestyle

Find Product Recommendations Alternatives to Products You Already Love

Thanks!

Any questions?

You can find me at:

- [Github.com/nicoivazquez](https://github.com/nicoivazquez)
- nicoletovazquez@gmail.com



Instructions for use

EDIT IN GOOGLE SLIDES

Click on the button under the presentation preview that says "Use as Google Slides Theme".

You will get a copy of this document on your Google Drive and will be able to edit, add or delete slides.

You have to be signed in to your Google account.

More info on how to use this template at slidescarnival.com/help-use-presentation-template

This template is free to use under [Creative Commons Attribution](#) license. You can keep the Credits slide or mention SlidesCarnival and other resources used in a slide footer.

EDIT IN POWERPOINT®

Click on the button under the presentation preview that says "Download as PowerPoint template". You will get a .pptx file that you can edit in PowerPoint.

Remember to download and install the fonts used in this presentation (you'll find the links to the font files needed in the [Presentation design slide](#))



Let's review some concepts



Is the color of gold, butter and ripe lemons. In the spectrum of visible light, yellow is found between green and orange.



Is the colour of the clear sky and the deep sea. It is located between violet and green on the optical spectrum.



Is the color of blood, and because of this it has historically been associated with sacrifice, danger and courage.



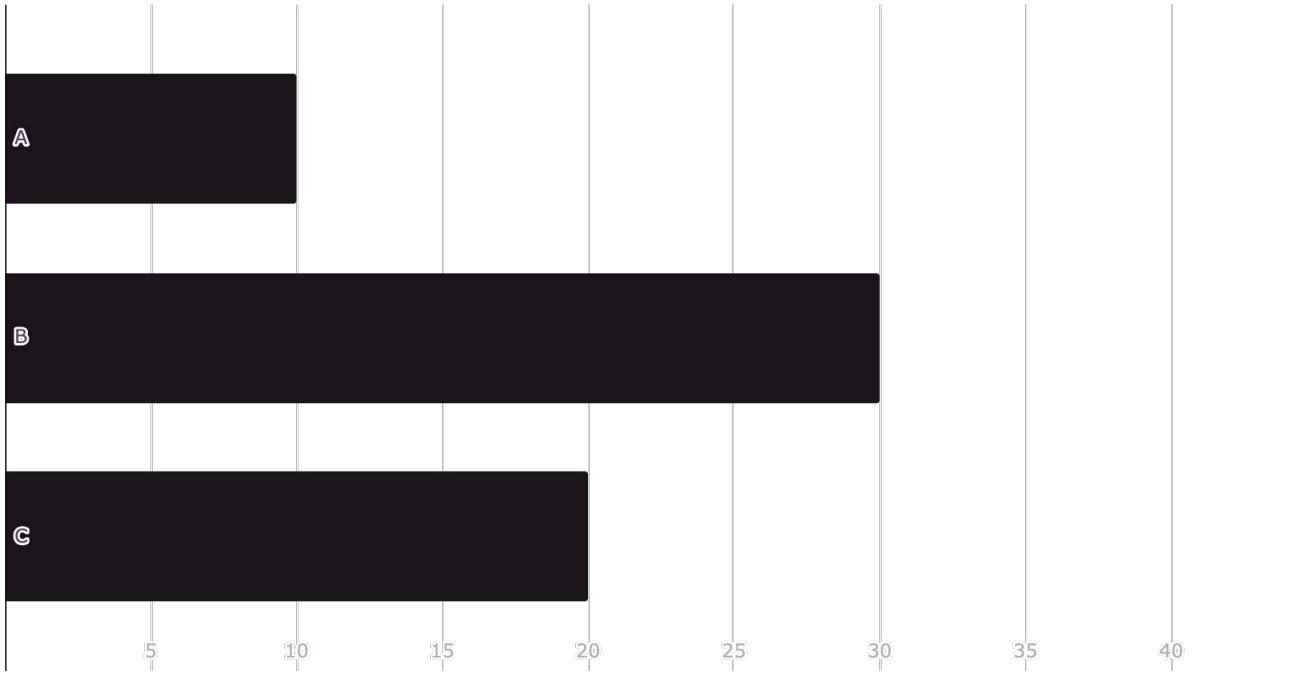
Is the color of gold, butter and ripe lemons. In the spectrum of visible light, yellow is found between green and orange.



Is the colour of the clear sky and the deep sea. It is located between violet and green on the optical spectrum.



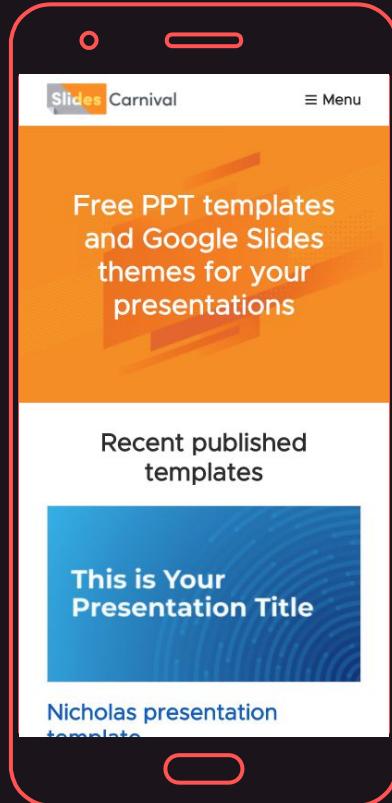
Is the color of blood, and because of this it has historically been associated with sacrifice, danger and courage.



You can insert graphs from Excel or Google Sheets

Mobile project

Show and explain your web, app or software projects using these gadget templates.





Want big impact?
Use big image.

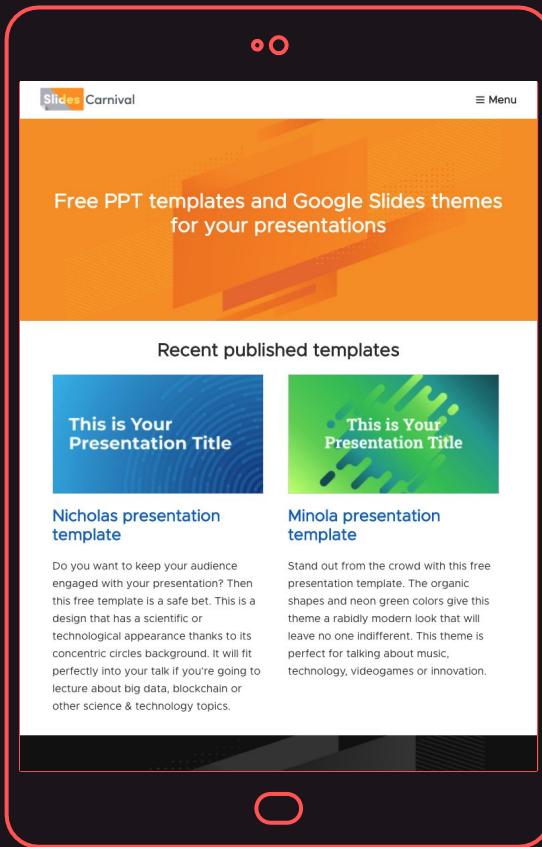
“

If a model is created in
a jupyter notebook
but no one ever uses
it, was it ever created?



Tablet project

Show and explain your web, app or software projects using these gadget templates.



Connecting Pipelines

White

Is the color of milk and fresh snow, the color produced by the combination of all the colors of the visible spectrum.

Black

Is the color of ebony and of outer space. It has been the symbolic color of elegance, solemnity and authority.



In two or three columns

Yellow

Is the color of gold, butter and ripe lemons. In the spectrum of visible light, yellow is found between green and orange.

Blue

Is the colour of the clear sky and the deep sea. It is located between violet and green on the optical spectrum.

Red

Is the color of blood, and because of this it has historically been associated with sacrifice, danger and courage.

Presentation design

This presentation uses the following typographies:

- Titles: Lexend Deca
- Body copy: Nunito Sans

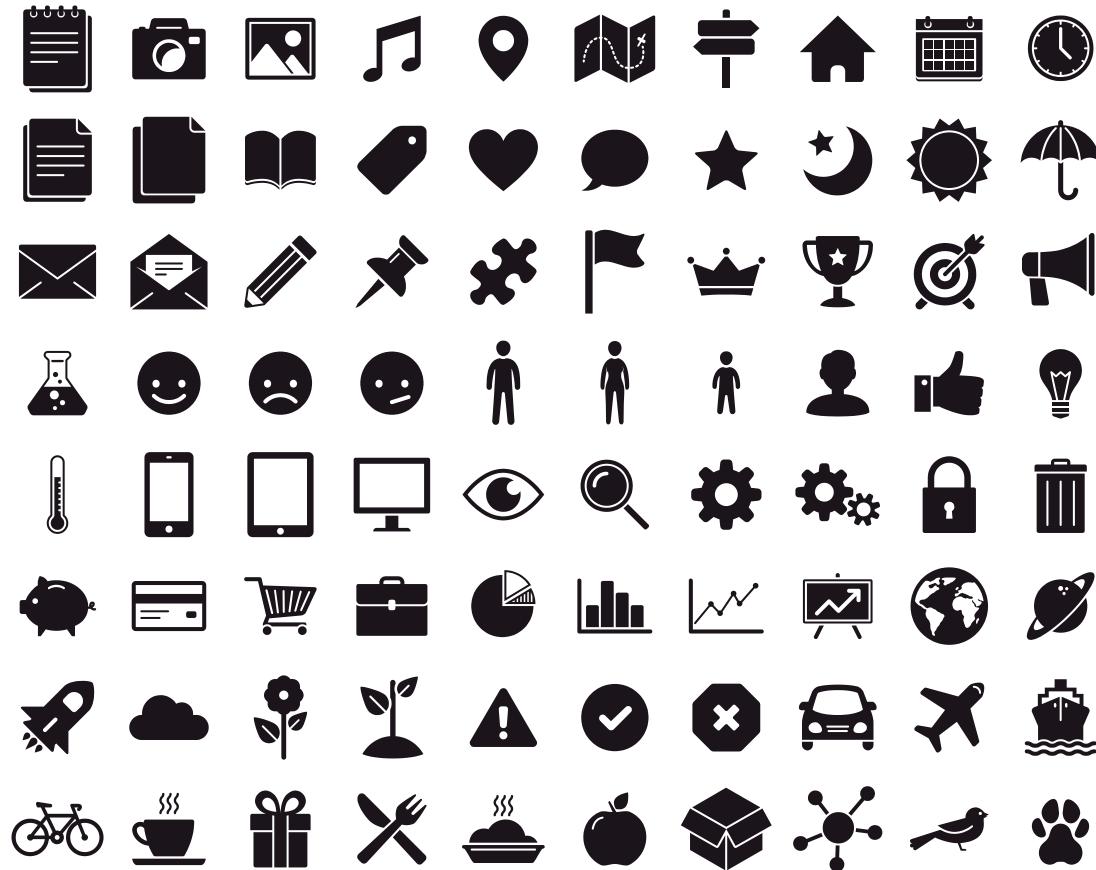
Download for free at:

<https://www.lexend.com/>

<https://www.fontsquirrel.com/fonts/nunito-sans>

You don't need to keep this slide in your presentation. It's only here to serve you as a design guide if you need to create new slides or download the fonts to edit the presentation in PowerPoint®





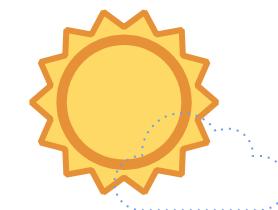
SlidesCarnival icons are editable shapes.

This means that you can:

- Resize them without losing quality.
- Change fill color and opacity.
- Change line color, width and style.

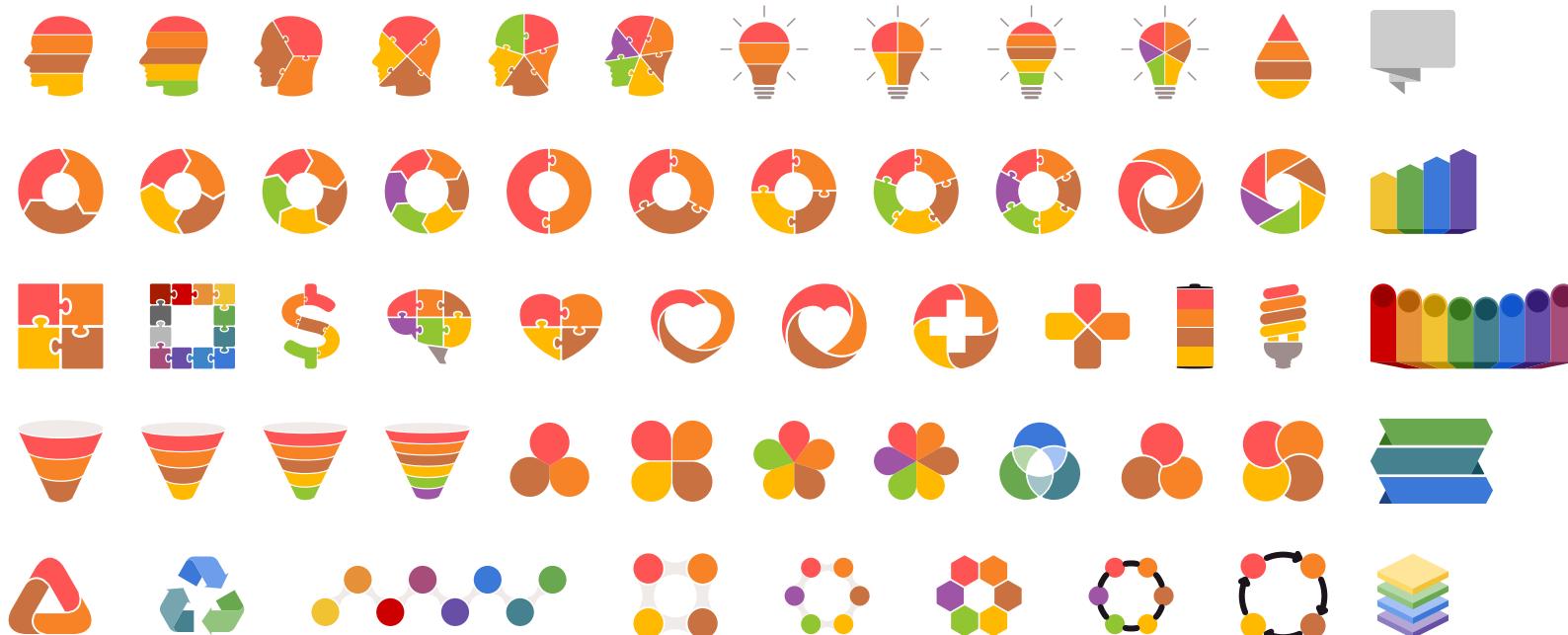
Isn't that nice? :)

Examples:



Find more icons at
slidescarnival.com/extr-free-resources-icons-and-maps

Diagrams and infographics

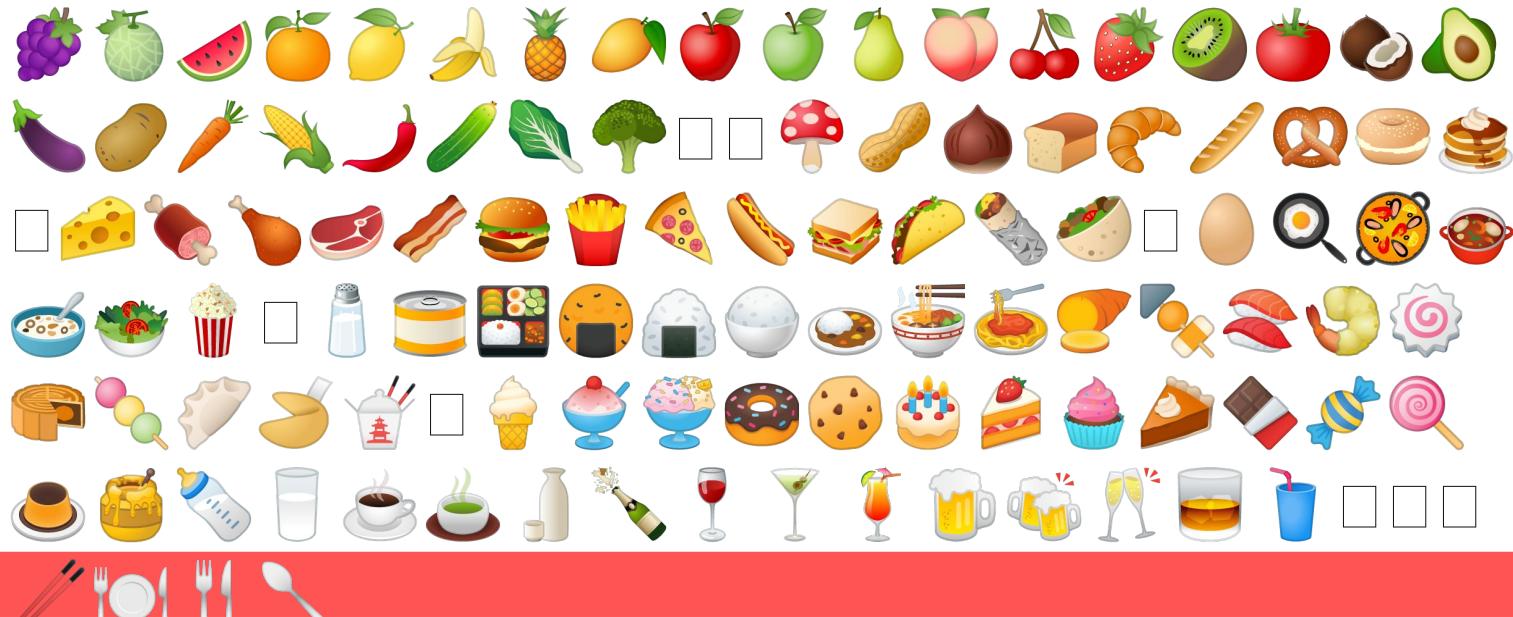


You can also use any emoji as an icon!

And of course it resizes without losing quality.

How? Follow Google instructions

<https://twitter.com/googledocs/status/730087240156643328>





Free templates for all your presentation needs



For PowerPoint and
Google Slides



100% free for personal
or commercial use



Ready to use,
professional and
customizable



Blow your audience
away with attractive
visuals