Calculating GHGE impacts and carbon labels for generic meals

Christian Reynolds, Berill Takacs, Ianko Ignatiev, Dinko Tenev, Victor Penev

LEAP 2021, 10:50 - 11:05, 6 December 2021

Dr Christian Reynolds
Centre for Food Policy, City, University of London
@sartorialfoodie
christian.reynolds@city.ac.uk

Centre for Food Policy
Shaping an effective food system

We need sustainable recipes tools and data

- **Public engagement/communication need**

#1 ask

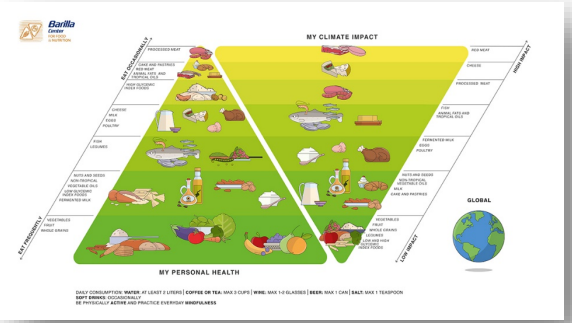
TAKE A BITE OUT OF CLIMATE CHANGE

“how/what can I cook sustainably this at home?”

“what are the impacts of this recipe?”

We need this information to empower citizens!

People do not think in **ingredients**, they think in **recipes**



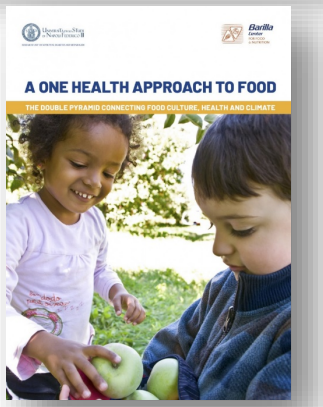
- **Industry need**

Need for communication around sustainable menu development and recipe design.

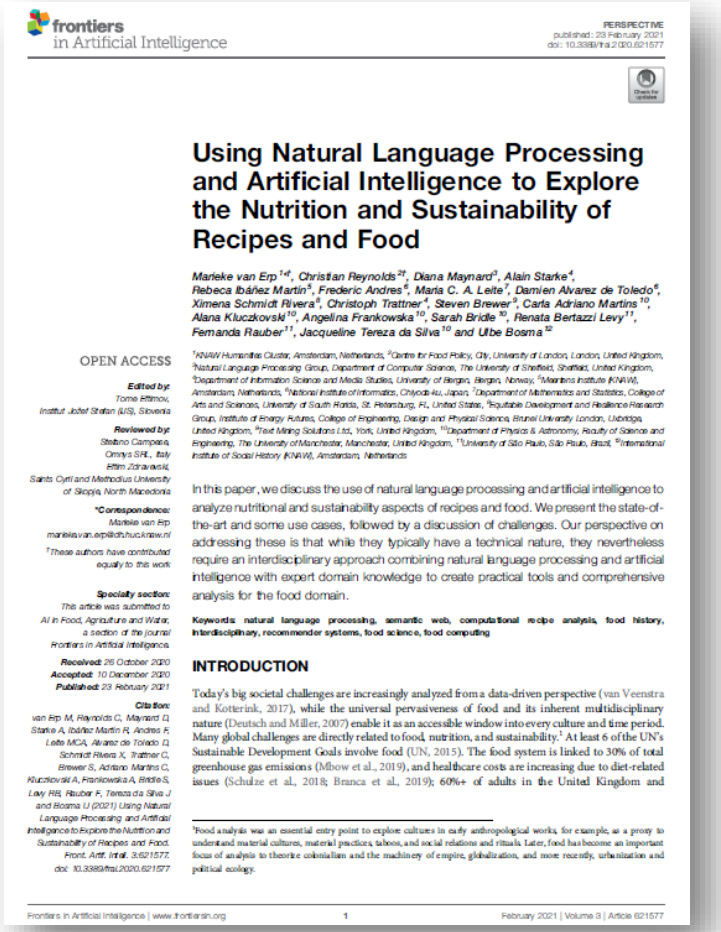
- **Policy need**

Need for data / visualisations of nutrition and food education, pack and portion advice etc.

Are there recipes that meet or are within the **Eat-Lancet** ?



This builds on previous NLP and recipe work



frontiers in Artificial Intelligence

PERSPECTIVE
published: 23 February 2021
doi: 10.3389/frai.2020.621577

Using Natural Language Processing and Artificial Intelligence to Explore the Nutrition and Sustainability of Recipes and Food

Marieke van Erp[1*†], Christian Reynolds[2†], Diana Maynard[3], Alain Starke[4], Rebeca Ibáñez Martín[5], Frederic Andres[6], Maria C. A. Leite[7], Damien Alvarez de Toledo[8], Ximena Schmidt Rivera[9], Christoph Trattner[4], Steven Brewer[3], Carla Adriano Martins[10], Alana Kluczkovski[10], Angelina Frankowska[10], Sarah Bridle[10], Renata Bertazzi Levy[11], Fernanda Rauber[11], Jacqueline Tereza da Silva[10] and Ulbe Bosma[12]

[1]KNAW Humanities Cluster, Amsterdam, Netherlands, [2]Centre for Food Policy, City, University of London, London, United Kingdom, [3]Natural Language Processing Group, Department of Computer Science, The University of Sheffield, Sheffield, United Kingdom, [4]Department of Information Science and Media Studies, University of Bergen, Bergen, Norway, [5]Maastricht Institute (KNAW), Amsterdam, Netherlands, [6]National Institute of Informatics, Chiyoda-ku, Japan, [7]Department of Mathematics and Statistics, College of Arts and Sciences, University of South Florida, St. Petersburg, FL, United States, [8]Equitable Development and Resilience Research Group, Institute of Energy Futures, College of Engineering, Design and Physical Science, Brunel University London, Uxbridge, United Kingdom, [9]Text Mining Solutions Ltd, York, United Kingdom, [10]Department of Physics & Astronomy, Faculty of Science and Engineering, The University of Manchester, Manchester, United Kingdom, [11]University of São Paulo, São Paulo, Brazil, [12]International Institute of Social History (KNAW), Amsterdam, Netherlands

OPEN ACCESS

Edited by: Tome Eftimov, Institut Jožef Stefan (IJS), Slovenia

Reviewed by: Stefano Campese, Omnys SRL, Italy; Riste Stojanov, Saints Cyril and Methodius University of Skopje, North Macedonia

***Correspondence:** Marieke van Erp marieke.van.erp@dh.huc.knaw.nl

[†]These authors have contributed equally to this work

Specialty section: This article was submitted to AI in Food, Agriculture and Water, a section of the journal Frontiers in Artificial Intelligence.

Received: 26 October 2020
Accepted: 10 December 2020
Published: 23 February 2021

Citation: van Erp M, Reynolds C, Maynard D, Starke A, Ibáñez Martín R, Andres F, Leite MCA, Alvarez de Toledo D, Schmidt Rivera X, Trattner C, Brewer S, Adriano Martins C, Kluczkovski A, Frankowska A, Bridle S, Levy RB, Rauber F, Tereza da Silva J and Bosma U (2021) Using Natural Language Processing and Artificial Intelligence to Explore the Nutrition and Sustainability of Recipes and Food. Front. Artif. Intell. 3:621577. doi: 10.3389/frai.2020.621577

In this paper, we discuss the use of natural language processing and artificial intelligence to analyze nutritional and sustainability aspects of recipes and food. We present the state-of-the-art and some use cases, followed by a discussion of challenges. Our perspective on addressing these is that while they typically have a technical nature, they nevertheless require an interdisciplinary approach combining natural language processing and artificial intelligence with expert domain knowledge to create practical tools and comprehensive analysis for the food domain.

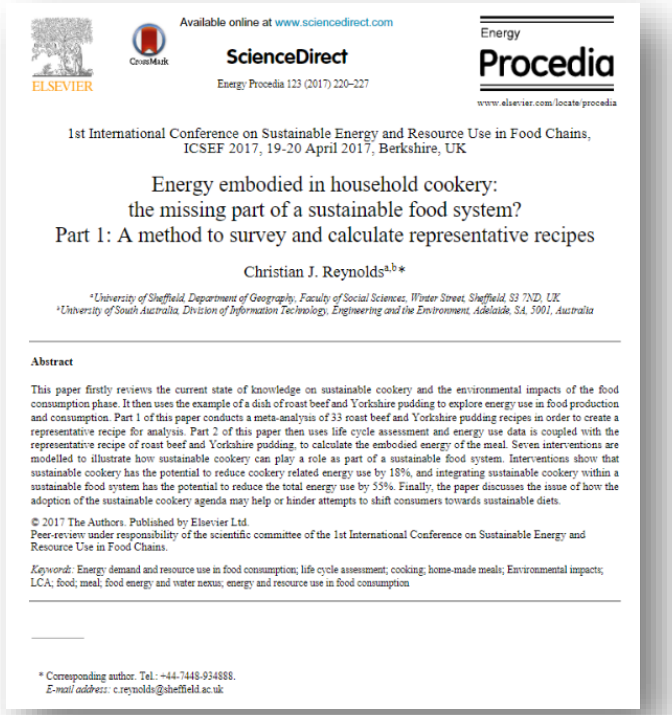
Keywords: natural language processing, semantic web, computational recipe analysis, food history, interdisciplinary, recommender systems, food science, food computing

INTRODUCTION

Today's big societal challenges are increasingly analyzed from a data-driven perspective (van Veenstra and Kotterink, 2017), while the universal pervasiveness of food and its inherent multidisciplinary nature (Deutsch and Miller, 2007) enable it as an accessible window into every culture and time period. Many global challenges are directly related to food, nutrition, and sustainability.[1] At least 6 of the UN's Sustainable Development Goals involve food (UN, 2015). The food system is linked to 30% of total greenhouse gas emissions (Mbow et al., 2019), and healthcare costs are increasing due to diet-related issues (Schulze et al., 2018; Branca et al., 2019); 60%+ of adults in the United Kingdom and

[1]Food analysis was an essential entry point to explore cultures in early anthropological works, for example, as a proxy to understand material cultures, material practices, taboos, and social relations and rituals. Later, food has become an important focus of analysis to theorize colonialism and the machinery of empire, globalization, and more recently, urbanization and political ecology.

https://dx.doi.org/10.3389%2Ffrai.2020.621577



Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 123 (2017) 220–227

Energy Procedia

www.elsevier.com/locate/procedia

1st International Conference on Sustainable Energy and Resource Use in Food Chains, ICSEF 2017, 19-20 April 2017, Berkshire, UK

Energy embodied in household cookery: the missing part of a sustainable food system? Part 1: A method to survey and calculate representative recipes

Christian J. Reynolds[a,b]*

[a]University of Sheffield, Department of Geography, Faculty of Social Sciences, Winter Street, Sheffield, S3 7ND, UK
[b]University of South Australia, Division of Information Technology, Engineering and the Environment, Adelaide, SA, 5001, Australia

Abstract

This paper firstly reviews the current state of knowledge on sustainable cookery and the environmental impacts of the food consumption phase. It then uses the example of a dish of roast beef and Yorkshire pudding to explore energy use in food production and consumption. Part 1 of this paper conducts a meta-analysis of 33 roast beef and Yorkshire pudding recipes in order to create a representative recipe for analysis. Part 2 of this paper then uses life cycle assessment and energy use data is coupled with the representative recipe of roast beef and Yorkshire pudding, to calculate the embodied energy of the meal. Seven interventions are modelled to illustrate how sustainable cookery can play a role as part of a sustainable food system. Interventions show that sustainable cookery has the potential to reduce cookery related energy use by 18%, and integrating sustainable cookery within a sustainable food system has the potential to reduce the total energy use by 55%. Finally, the paper discusses the issue of how the adoption of the sustainable cookery agenda may help or hinder attempts to shift consumers towards sustainable diets.

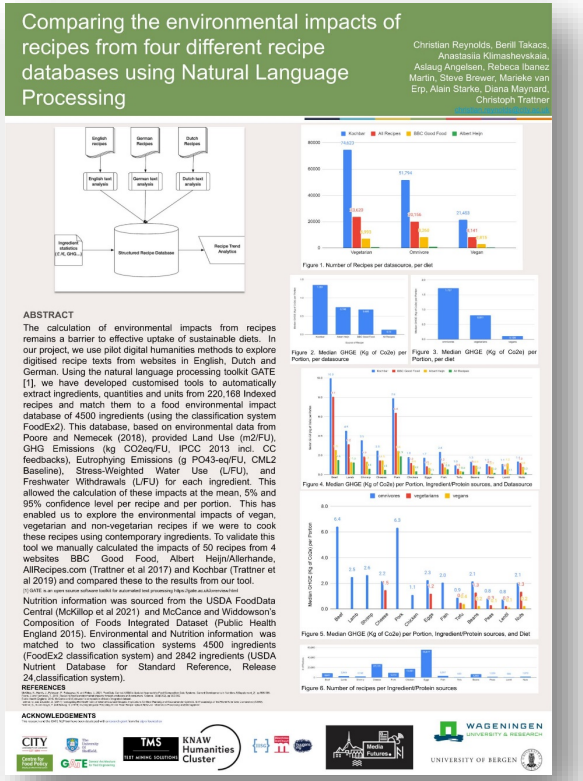
Peer-review under responsibility of the scientific committee of the 1st International Conference on Sustainable Energy and Resource Use in Food Chains.

Keywords: Energy demand and resource use in food consumption; life cycle assessment; cooking; home-made meals; Environmental impacts; LCA; food; meal; food energy and water nexus; energy and resource use in food consumption

* Corresponding author. Tel.: +44-7448-934888.
E-mail address: c.reynolds@sheffield.ac.uk

https://doi.org/10.1016/j.egypro.2017.07.245

LEAP 2021 Poster for the project: Communicating the environmental impact of plant based recipes – funded by the Alpro foundation (2021).



Comparing the environmental impacts of recipes from four different recipe databases using Natural Language Processing

Christian Reynolds, Benill Takacs, Anastasiia Klimashevskaia, Aslaug Angelsen, Rebeca Ibanez Martin, Steve Brewer, Marieke van Erp, Alain Starke, Diana Maynard, Christoph Trattner

ABSTRACT

The calculation of environmental impacts from recipes remains a barrier to effective uptake of sustainable diets. In our project, we use pilot digital humanities methods to explore digitised recipe texts from websites in English, Dutch and German. Using the natural language processing toolkit GATE [1], we have developed customised tools to automatically extract ingredients, quantities and units from 220,168 Indexed recipes and match them to a food environmental impact database of 4500 ingredients (using the classification system FoodEx2). This database, based on environmental data from Poore and Nemecek (2018), provided Land Use (m2/FU), GHG Emissions (kg CO2eq/FU, IPCC 2013 incl. CC feedbacks), Eutrophying Emissions (g PO43-eq/FU, CML2 Baseline), Stress-Weighted Water Use (L/FU), and Freshwater Withdrawals (L/FU) for each ingredient. This allowed the calculation of these impacts at the mean, 5% and 95% confidence level per recipe and per portion. This has enabled us to explore the environmental impacts of vegan, vegetarian and non-vegetarian recipes if we were to cook these recipes using contemporary ingredients. To validate this tool we manually calculated the impacts of 50 recipes from 4 websites BBC Good Food, Albert Heijn/Allerhande, AllRecipes.com (Trattner et al 2017) and Kochbar (Trattner et al 2019) and compared these to the results from our tool.

Nutrition information was sourced from the USDA FoodData Central (McKillop et al 2021) and McCance and Widdowson's Composition of Foods Integrated Dataset (Public Health England 2015). Environmental and Nutrition information was matched to two classification systems 4500 ingredients (FoodEx2 classification system) and 2842 ingredients (USDA Nutrient Database for Standard Reference, Release 24,classification system).

REFERENCES

ACKNOWLEDGEMENTS

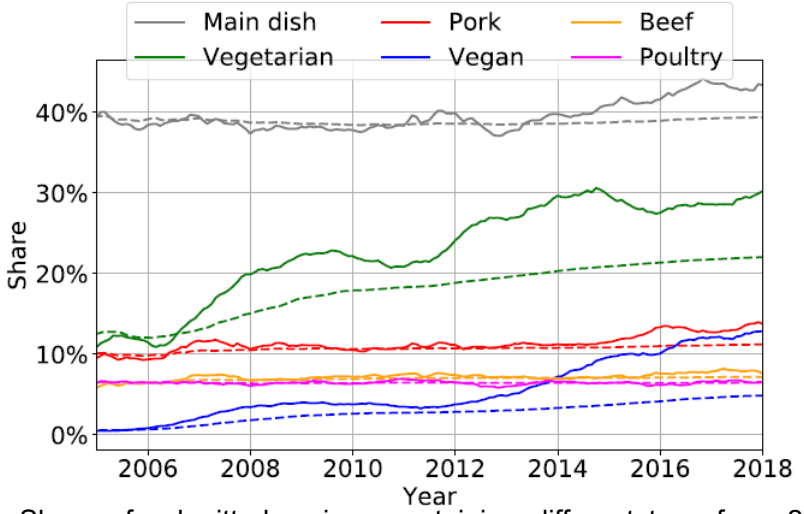
Multiple studies already

- Nutritional and health studies (Reinivuo et al., 2009; Trattner et al., 2017)
- Computational linguistics (Jurafsky, 2015),
- Computational gastronomy (Jain et al.,2015)
- Online shopping recommendations (Aiello et al., 2019)
- Semantic web (Haussmann et al., 2019)

This is still a young field of investigation!

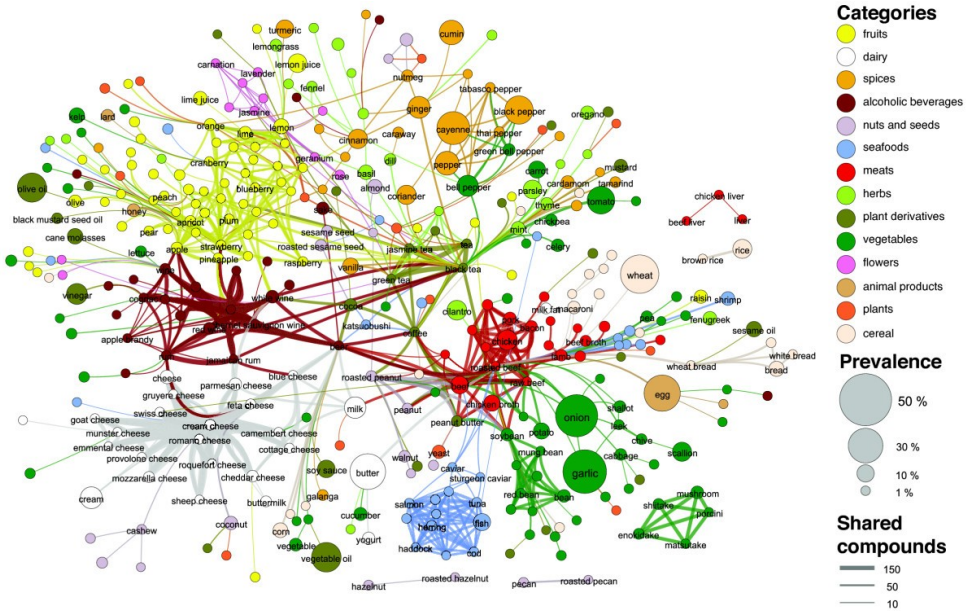
Other Recipe NLP research

Analysis of submitted recipes (Asano and Biermann, 2019)

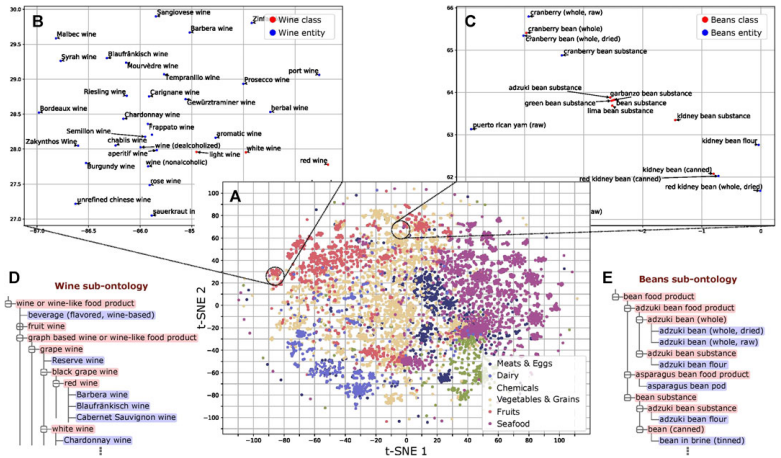


Share of submitted recipes containing different tags from 2005-2018 as a time-series. https://doi.org/10.1038/s41893-019-0316-0

Flavour networks (Ahnert 2013) 10.1186/2044-7248-2-4

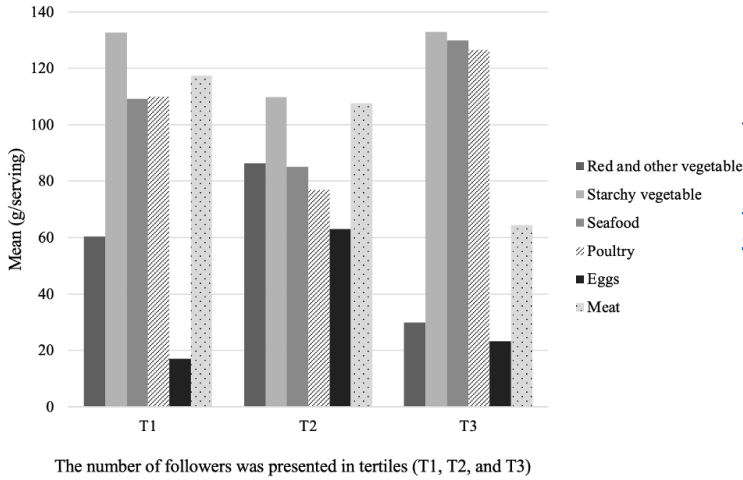


Existing food ontology groupings



Using Word Embeddings to Learn a Better Food Ontology https://doi.org/10.3389/frai.2020.584784

Healthfulness Assessment of Recipes Shared on Pinterest



The number of followers was presented in tertiles (T1, T2, and T3)

Relationship between pinners' popularity and their recipe ingredients. (Cheng et al 2021) https://doi.org/10.2196/25757

Ecolabels are becoming mainstream

There are many (10+) environmental Ecolables now emerging with different food label designs, with combined and multicomponent scores

Sustainable food profiling models to inform the development of food labels that account for nutrition and the environment: a systematic review

Anne Charlotte Bunge, Kremlin Wickramasinghe, Jessica Renzella, Michael Clark, Mike Rayner, Holly Rippin, Afton Halloran, Nia Roberts, João Breda

Sustainable food profiling models (SFPMs) are the scientific basis for the labelling of food products according to their environmental and nutritional impact, allowing consumers to make informed choices. We identified ten SFPMs that score individual foods according to at least two environmental indicators, with the most common being greenhouse gas emissions (n=10) and water use (n=8). Six models additionally assessed the nutritional quality of foods and presented different methods to combine nutritional and environmental indicators. Key advantages of identified models include a wide range in system boundaries, reference units, approaches for defining cutoff values, design proposals for food labelling schemes, and the comprehensive geographical scope of the lifecycle inventory databases used in the development phase of the model. Key disadvantages of identified models include inconsistent methods for food classification and poor replicability due to unclear methods, unavailable code for environmental and nutritional impact calculation, and unclear cutoff values. We found that few SFPMs to date account for at least two environmental impact factors, and even fewer include nutritional values or other dimensions of sustainability. This systematic review highlights the need to use consistent components and to develop national and international reference values for the classification of sustainable food to enable standardised food labelling.

Lancet Planet Health 2021; 5: e818–26

NCD Office, World Health Organization, Moscow, Russia (A C Bunge MSc, K Wickramasinghe DPhil, H Rippin PhD, A Halloran PhD, J Breda PhD); Stockholm Resilience Centre, Stockholm, Sweden (A C Bunge); Nuffield Department of Population Health, University of Oxford, Oxford, UK (J Renzella MPH, M Clark PhD, Prof M Rayner DPhil, N Roberts MSc); Department of Nutrition, Exercise and Sports,

https://doi.org/10.1016/S2542-5196(21)00231-X

		Model name										
		Menu sustainability index[28]	Dietary environmental index[25]	SusDISH-LEH[27]	Avadi[29]	Masset[30]	Leach[31]	CONE-LCA[32]	Eaternity score[22]	Food carbon scope[33]	Beelong Eco-Score[34]	% of SFPMs fulfilling criterion
Factors of environmental impact	Greenhouse gas emission	■	■	■	■	■	■	■	■	■	■	100%
	Land use*	■	■	■	■				■		■	60%
	Biodiversity loss	■		■							■	30%
	Water use†	■	■	■	■		■		■	■	■	80%
	Eutrophication	■				■						20%
	Acidification					■						10%
	Other‡	■		■	■	■	■	■		■	■	80%
Nutritional indicators	Protein		■	■		■			■			40%
	Fibre		■	■	■	■		■	■			50%
	Vitamins§		■	■	■	■			■			50%
	Minerals¶		■	■	■	■		■	■			60%
	Unsaturated fat			■	■	■		■	■			50%
	Saturated fat		■	■		■		■	■			50%
	Sugar		■			■						20%
	Sodium		■	■		■		■	■			50%
	Energy			■	■				■			30%
System boundary	Cradle-to-farm gate		■		■			■	■			40%
	Cradle-to-consumer			■			■				■	30%
	Cradle-to-grave	■				■						20%
	Adjustable									■		10%
Reference value	Mass basis (in g, kg, or tonne)		■	■	■		■		■	■		60%
	Serving size	■						■	■	■	■	40%
	Energy basis (in kcal)					■	■					20%
	Per kg of protein						■					10%

FOUNDATION EARTH

ECO IMPACT A+

Environmental Information

Carbon footprint: **867.0 gCO_2e** per serving (28.0% fair daily food emissions)

Find out more at myemissions.green

Generic Meals and carbon labels

Edamam, a provider of nutrition data and semantic solutions for businesses in the food, health, and wellness sectors (https://developer.edamam.com)

- Integrated a food environmental impact database of 2,842 ingredients (using the classification system of the USDA Nutrient Database for Standard Reference, Release 24). This food environmental impact database was based on environmental data from Poore and Nemecek (2018) and was supplied by City.
- For some items which are not part of USDA food list Edamam used in-house nutrition experts to map them to USDA items.
- Edamam has labeled about **5 million recipes in the English language web** with CO2 labels ranking from A+ (best) to G (worst) and is making those searchable via its Recipe Search API.

Edamam's Generic meals are a database of 180,000+ recipes that encompass more than 90% of what restaurants offer/commonly cooked at home.

- Similar recipes are clustered based on titles after removing certain non essential words from the title. These recipes represent the initial generic meal set.
- Compare recipes based on nutrition and content and remove any outliers. From the rest of the recipes Edamam build a combined recipes for which they also create a distribution of labels and nutrition among the recipe population. CO2e is one of the values which is part of this calculation.
- Edamam matched the CO2e data and carbon labels to the Generic meals database.

EDAMAM
Food. Data. Health

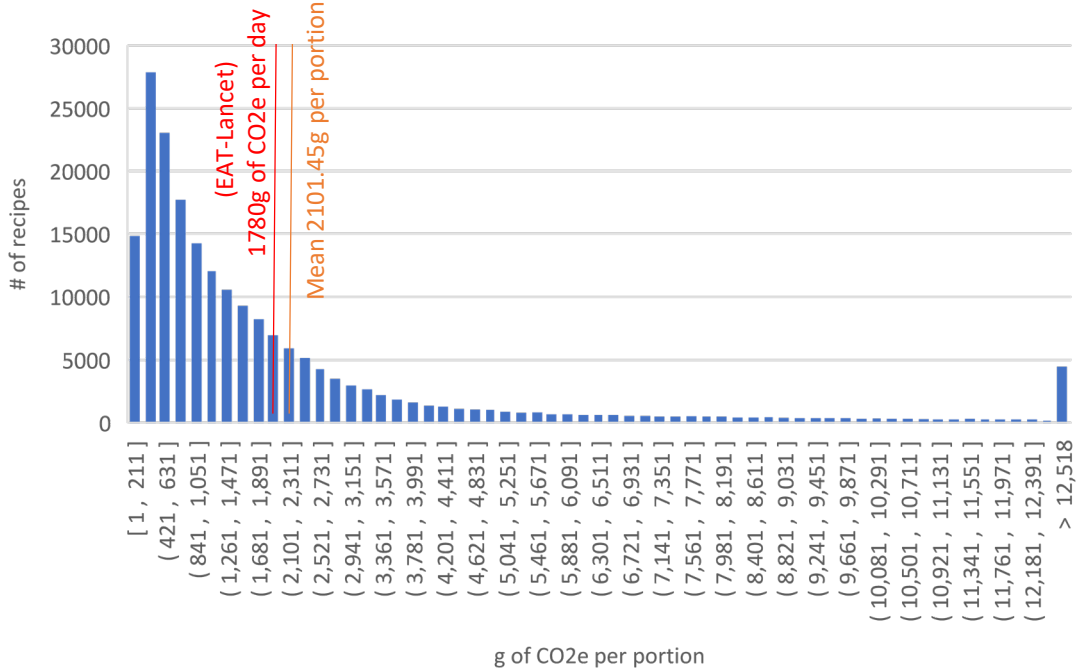
Results: YES! Eat-Lancet compatible recipes!

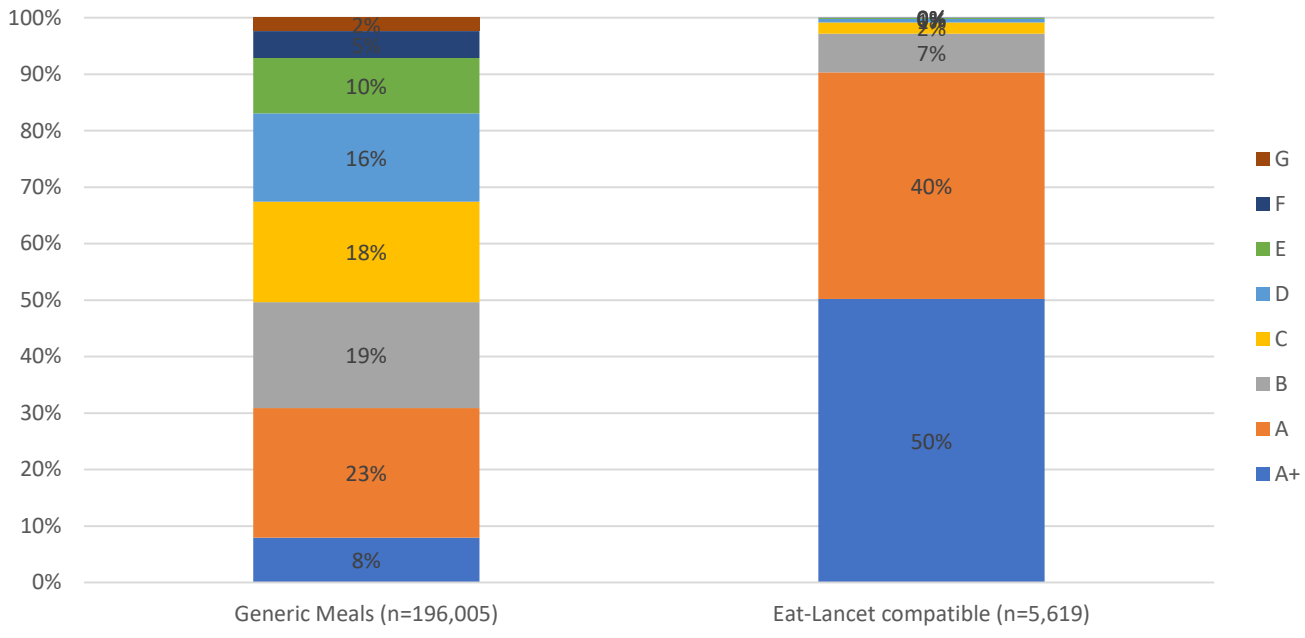
196,005 recipes with 100% ingredients matched to CO2e data. Mean 2101.45g of CO2e per portion, (SD 3472.02g)

Information provided in grams of CO2e per **portion**, per **Kcal**, per g of **protein**

Eat-Lancet recipes: Assume consumption of this recipe is scaled to meet 2500 kcal, and protein 56g, is the scaled recipe below 1780g of CO2e.

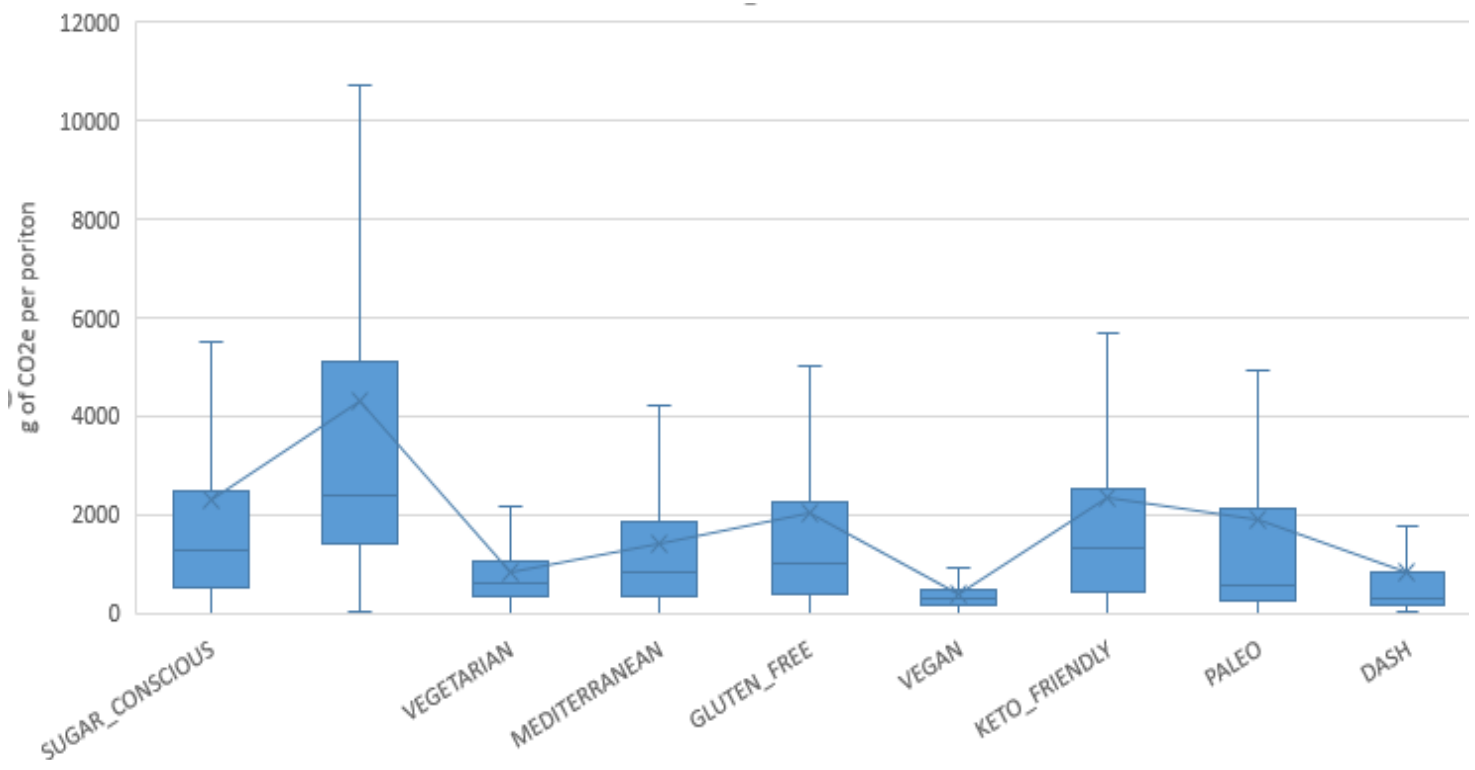
5,619 recipes met this criteria! (2.8%) Mean 180.87g of CO2e per portion, (SD 117.20g)





Different ways to cut the data… Health/Diet

Metadata presented for Meal type, **Health/Diet** type, Cuisine type, Dish type, and Ingredients per recipe

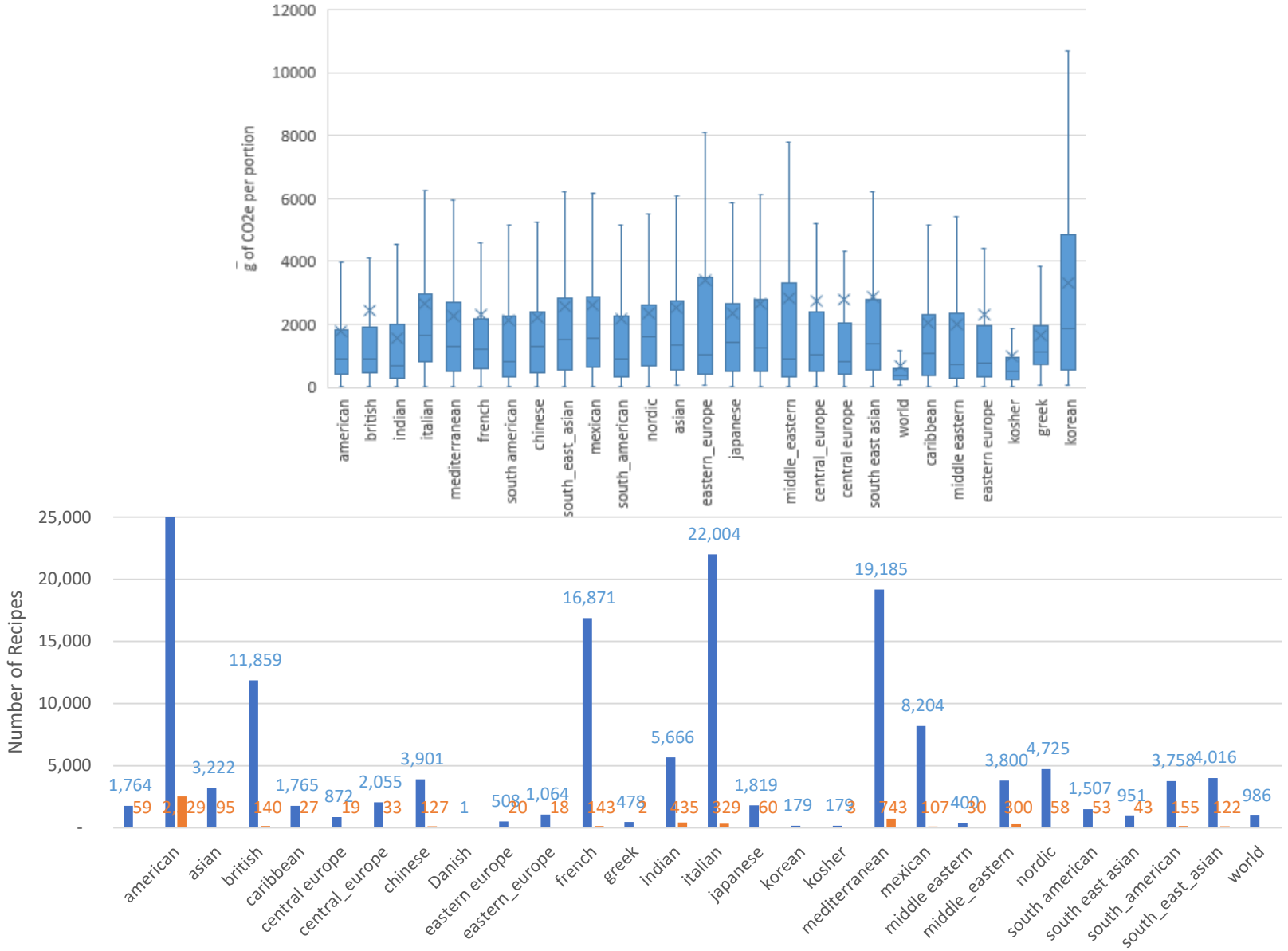


	SUGAR CONSCIOUS	No Classification	VEGETARIAN	MEDITERRANEAN	GLUTEN_FREE	VEGAN	KETO FRIENDLY	PALEO	DASH
Count	49,690	29,031	111,263	37,869	81,000	24,651	22,372	11,270	7,086
Avg. g CO2e per portion	2,313.34	4,320.09	833.55	1,417.64	2,013.42	402.28	2,349.80	1,881.94	816.31

Different carbon impact spreads across Diet choice types, but also the number of recipes matters!

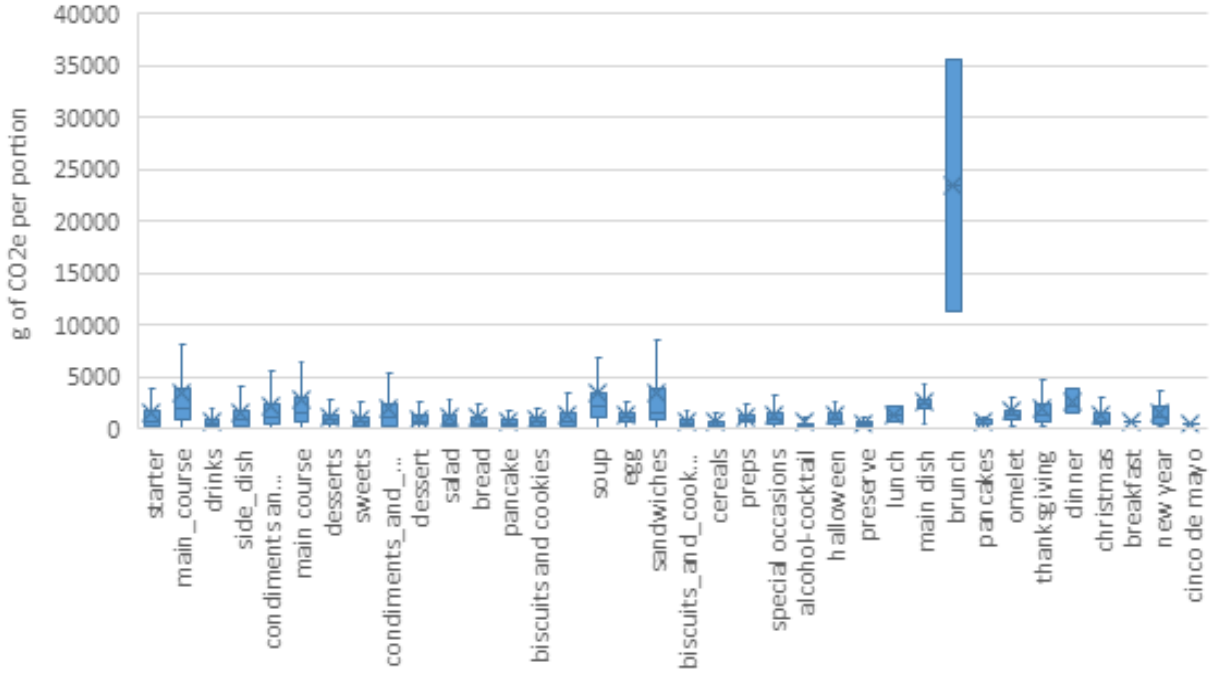
DASH, Vegan, and Vegetarian recipes had the lowest mean, median and IQR of any specific health/diet type.

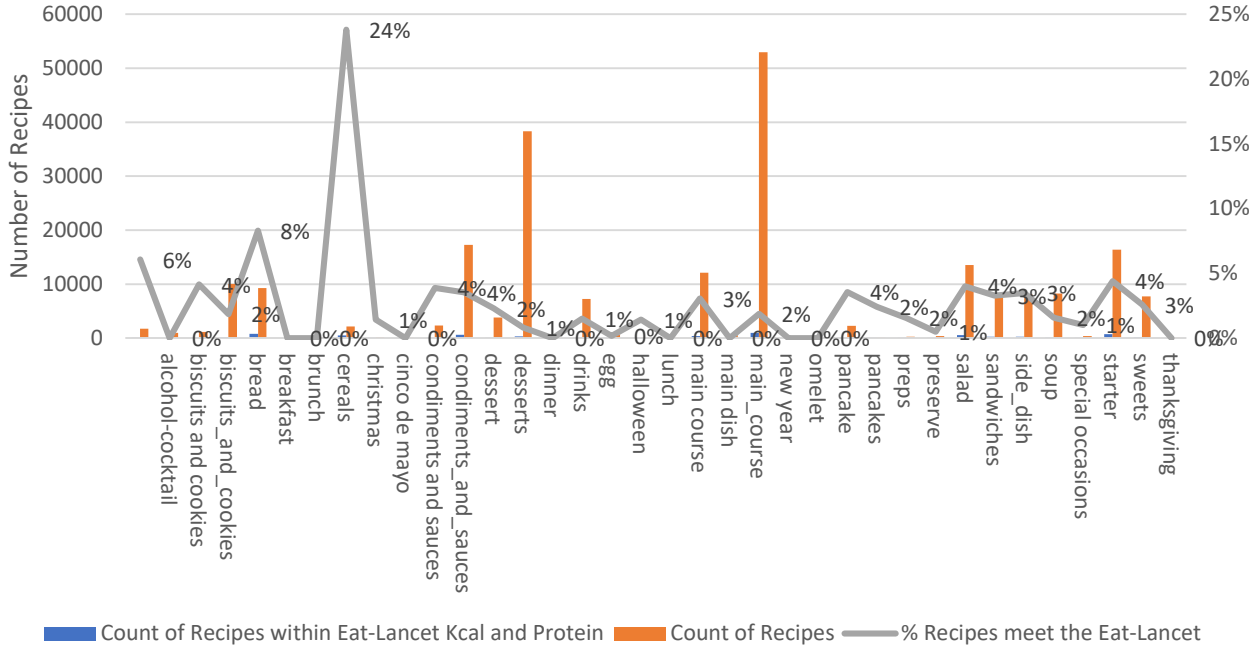
Different ways to cut the data… Cuisine type



There are a % of recipes in most cuisines that meet the Eat-Lancet

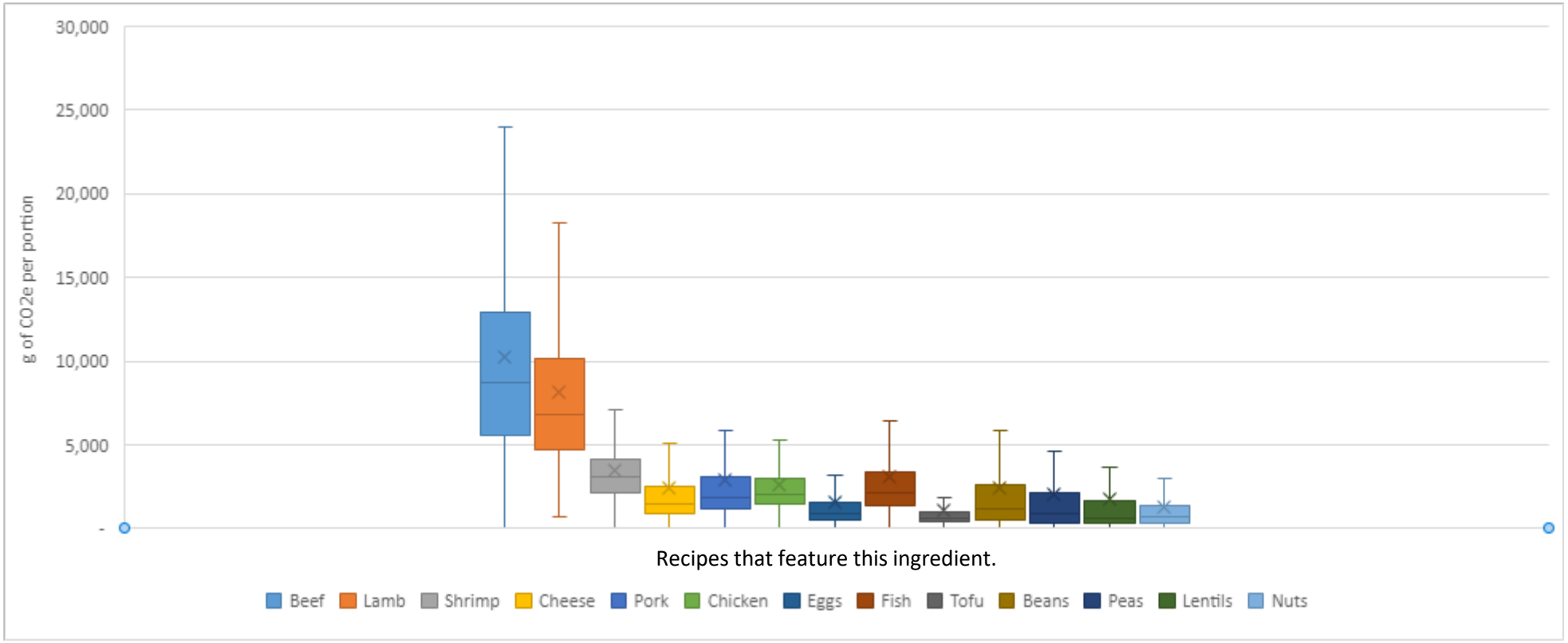
Different ways to cut the data… Dish type





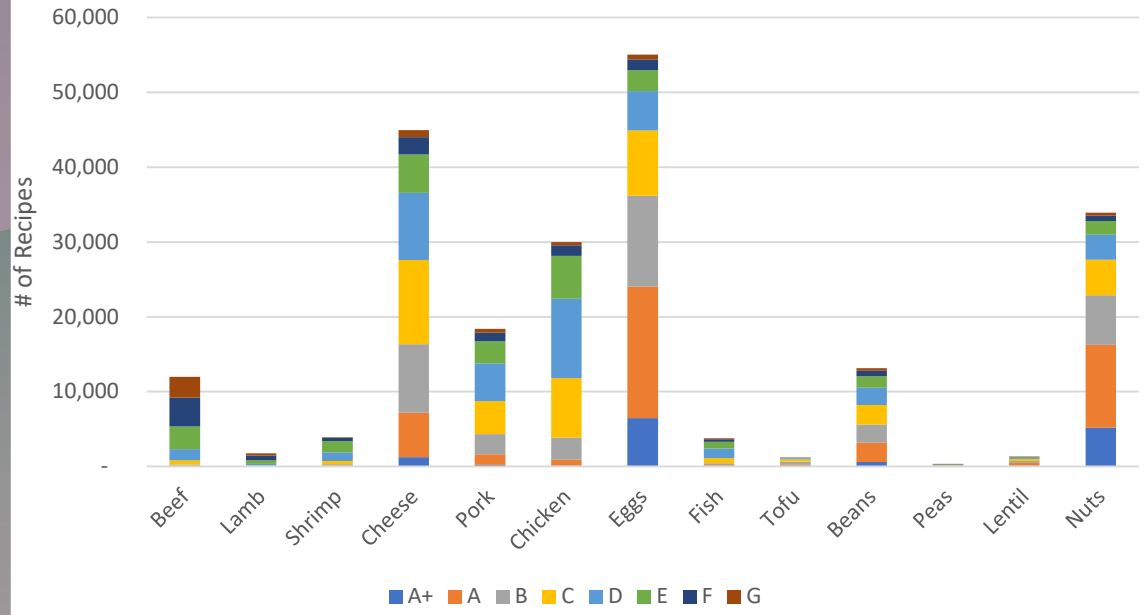
There are a % of recipes that meet the Eat-Lancet – Dish types vary in footprint, but a problem with sample size/tagging

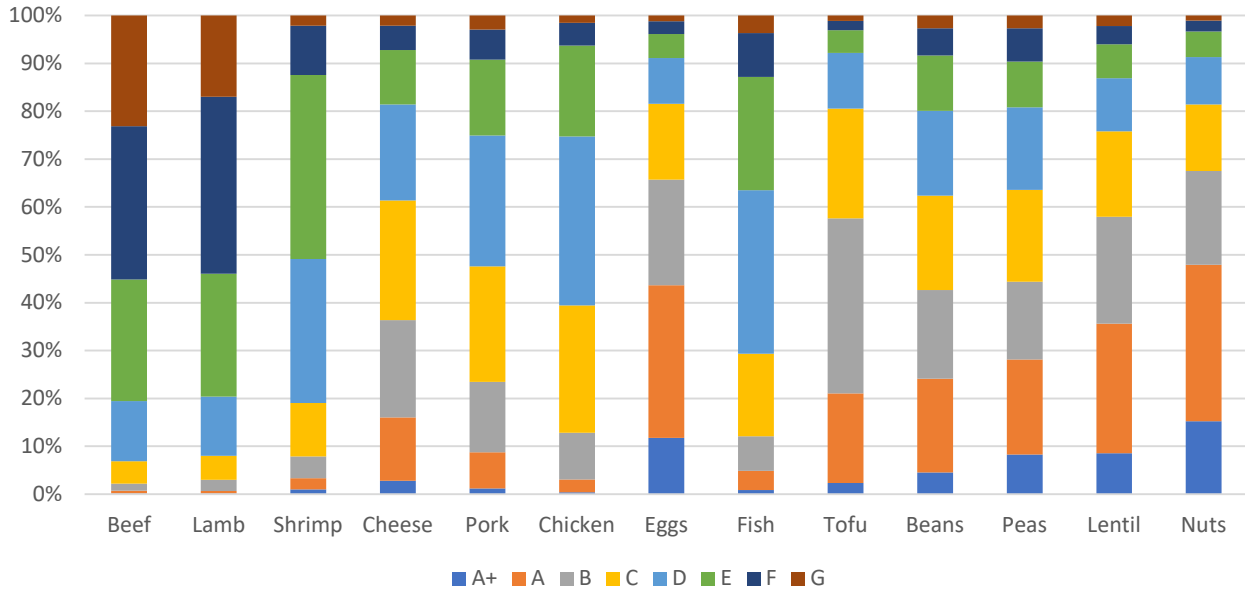
Different ways to cut the data… Ingredients



	Beef	Lamb	Shrimp	Cheese	Pork	Chicken	Eggs	Fish	Tofu	Beans	Peas	Lentil	Nuts
Mean g of CO2e per portion	10,265.96	8,139.05	3,448.71	2,388.032	2,890.13	2890.13	1,552.63	3,086.02	1,054.26	2,473.38	2,057.60	1,742.12	1,289.52
Count	11,984	1,776	3,890	44,959	18,411	18,411	55,074	3,795	1,168	13,157	302	1,312	33,835
# of Eat-Lancet	0	0	4	48	17	14	542	8	12	608	31	206	1802
% Eat-Lancet	0.0%	0.0%	0.1%	0.1%	0.1%	0.1%	1.0%	0.2%	1.0%	4.6%	10.3%	15.7%	5.3%

Different ways to cut the data… Ingredients





Different carbon label spreads across ingredient types, but also the number of recipes matters!

Key take-aways

- We have a database for CO2e of ~200,000 commonly cooked recipes in the English language (web)
  - Information provided in grams of CO2e per **portion**, per **Kcal**, per g of **protein** and **carbon labels**
  - This database, and API can easily be used on menus, cookbooks etc.
- Recipes from different cuisines, dishes, health/diets, and protein sources all can **NOW** be cooked to meet the Kcal and Protein requirements set out by the EAT-Lancet.
- DASH, Vegan, and Vegetarian recipes had the lowest mean, median and IQR of any specific health/diet type.
- We need to think about how carbon/eco labels convey complexity when compared to specific diet requirements (e.g Eat-Lancet).

Please do get in touch

Dr Christian Reynolds
Centre for Food Policy, City, University of London
@sartorialfoodie ***christian.reynolds@city.ac.uk***

The Centre for Food Policy, City, University of London offers the following courses

- **Nutrition and Food Policy BSc (Hons), with Distance Learning option**
  Undergraduate degree
- **Food Policy MSc/PGDip/PGCert, with Distance Learning option**
  Postgraduate taught degree
- **PhD/MPhil Food Policy**
  Postgraduate research degree

https://www.city.ac.uk/prospective-students/courses/postgraduate/food-policy

Thank you again to all my numerous collaborators and Edamam!