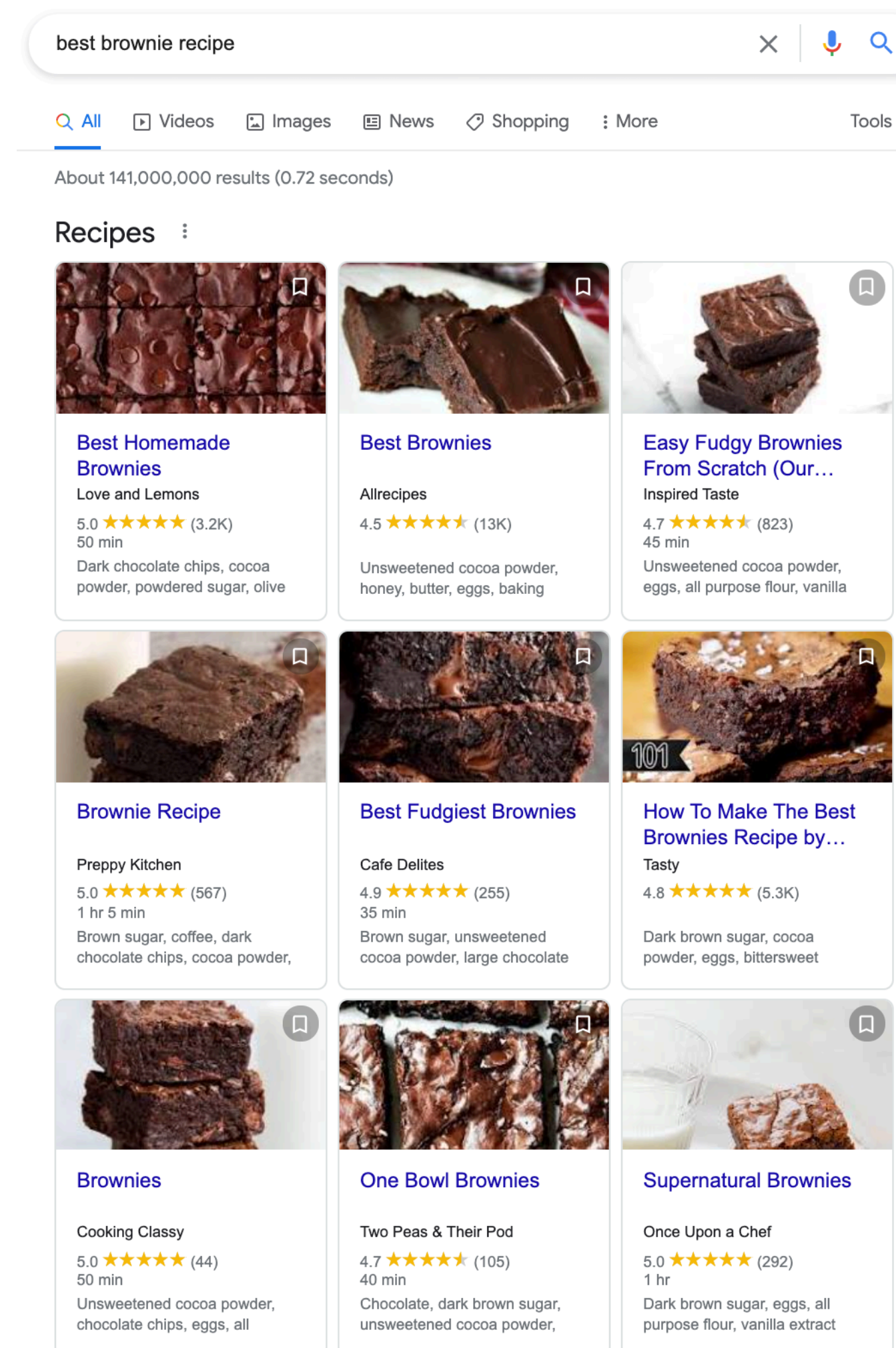# Using data science to improve cooking…?

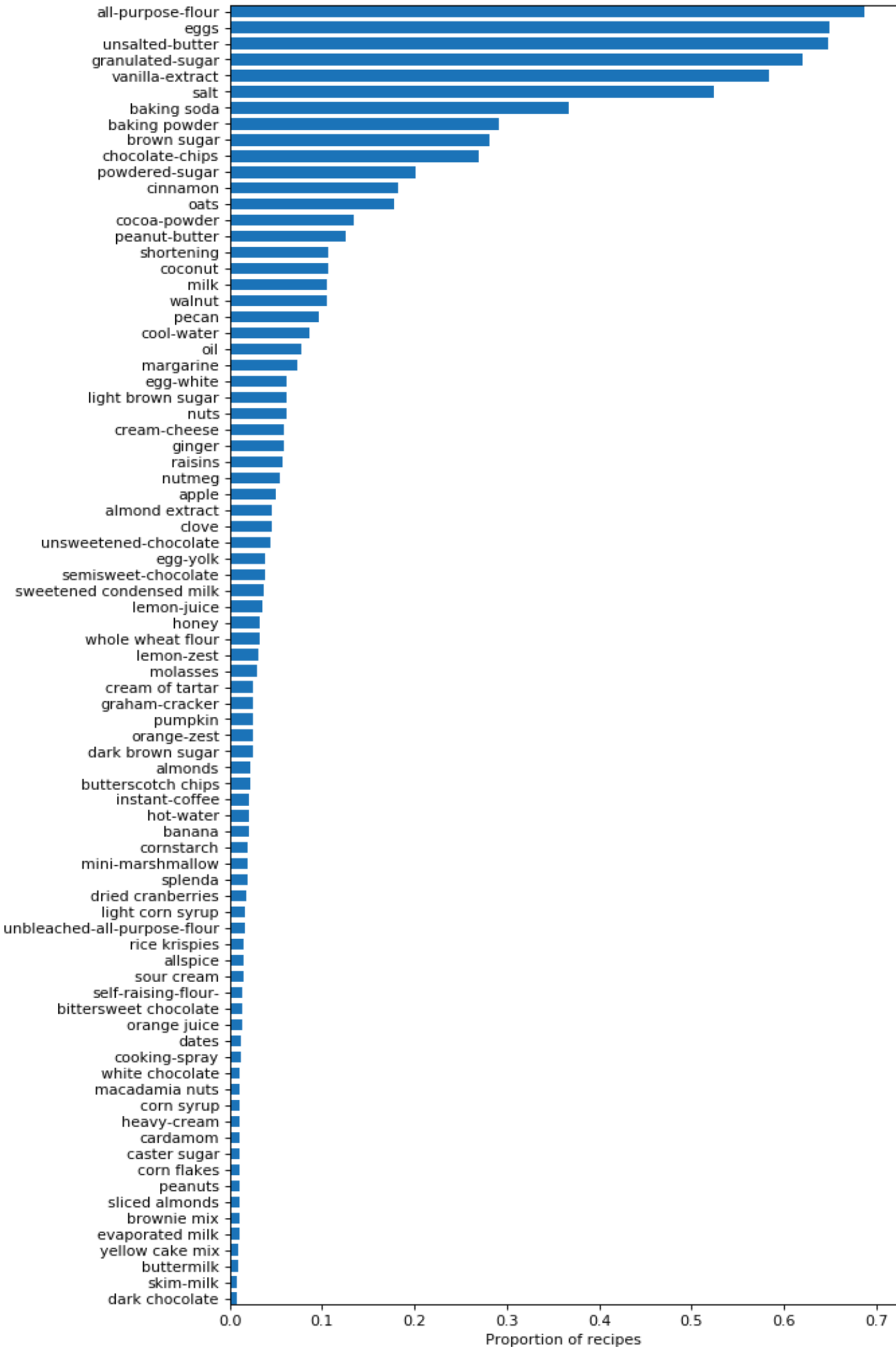- A myriad of data exists online in the space of recipes + ratings

Using data science to improve cooking...?

- A myriad of data exists online in the space of recipes + ratings

- Could data science approaches be used to improve recipe development?

- *Idea*: build a tool that provides suggestions for recipe enhancement using existing corpus of recipes + user ratings

- *Purpose*: use by recipe hosting websites (food.com) to offer users suggestions on **how** to improve their recipe

Using data science to improve cooking…?

- **Mechanics**: Use data on recipes + ratings from food.com on Kaggle

	name	id	minutes	contributor_id	submitted	tags	nutrition	n_steps	steps	description	ingredients	n_ingredients
30	get the sensation brownies	27087	70	37305	2002-05-03	['weeknight', 'time-to-make', 'course', 'main-...	[174.4, 13.0, 67.0, 5.0, 4.0, 26.0, 7.0]	11	['preheat oven to 350 degrees', 'grease 13 x 9...	a yummy thick chocolate mint brownie from york...	['butter', 'sugar', 'vanilla', 'eggs', 'all-pu...	9
72	one pot brownies	32271	40	41809	2002-06-25	['60-minutes-or-less', 'time-to-make', 'course...	[1389.5, 129.0, 410.0, 26.0, 42.0, 211.0, 51.0]	8	['melt chocolate and butter in heavy pot', 'le...	this recipe was submitted by mrs. stuart padno...	['unsweetened chocolate squares', 'butter', 's...	7
94	symphony brownies	44895	45	51224	2002-10-29	['60-minutes-or-less', 'time-to-make', 'course...	[5728.3, 470.0, 1941.0, 87.0, 140.0, 443.0, 22...	10	['preheat oven to 350 degrees , or 325 degrees...	a	['betty crocker fudge brownie mix', 'eggs', 'w...	5
103	turtle squares	58651	30	35635	2003-04-07	['30-minutes-or-less', 'time-to-make', 'course...	[172.4, 16.0, 51.0, 1.0, 2.0, 26.0, 6.0]	15	['preheat oven to 350 degrees f', 'spray a 13 ...	for lovers of pecans and chocolate...	['flour', 'brown sugar', 'butter', 'pecans', '...	5
111	1 brownies in the world best ever	333281	40	985201	2008-10-27	['60-minutes-or-less', 'time-to-make', 'course...	[138.4, 10.0, 50.0, 3.0, 3.0, 19.0, 6.0]	10	['heat the oven to 350f and arrange the rack i...	these are the most; chocolatey, moist, rich, d...	['bittersweet chocolate', 'unsalted butter', '...	9

Using data science to improve cooking…?

- **Mechanics**: Use data on recipes + ratings from food.com on Kaggle

  - Build predictive models to assess which ingredients are associated with recipe ratings

- Case study: *Brownie* recipes (N = 12,440; N = 515 with at least 20 reviews)

  - Data cleaning: combining redundant ingredients (e.g. all-purpose flour, flour, white flour, etc.)

Using data science to improve cooking...?

- **Mechanics**: Use data on recipes + ratings from food.com on Kaggle

  - Build predictive models to assess which ingredients are associated with recipe ratings

- Case study: ***Brownie*** recipes (N = 12,440; N = 515 with at least 20 reviews)

  - Data cleaning: combining redundant ingredients (e.g. all-purpose flour, flour, white flour, etc.)

- Presence of unique ingredients accounts for **~25%** of variance in recipe ratings

Brownie ingredients associated with HIGHER review ratings (N=515 recipes)

Brownie ingredients associated with LOWER review ratings (N=515 recipes)