Get The Best Ořechovka Using Design of Experiments

Data Show 2019

Vojtěch FILIPEC

whoami

☢️ nuclear physicist

🎯 data scientist

🏢 professional bio:

- banks: applied statistics in credit risk
- fintech, pharma: consultancy in analytics

Contents

About my project

1. ořechovka:
   a. what is it
   b. problem: many recipes
2. design of experiments
   a. what is it
   b. how-to in four steps
3. optimal recipe

About DoE

1. the same problem elsewhere
2. what is DoE good for
3. data analysis

goal: a new useful method in your toolbox

What is ořechovka and what is the problem here?

What is ořechovka?

semi-sweet liqueur

simple basic recipe

problem: many variations

Problem: Many variations

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
1	source	# nuts	alcohol	alcohol	alcohol	extraction	extraction	sirup		% alc. after mixing							
2			[l]	[%]	form	[weeks]	method	sugar [g]	water[g]		total sugar	clove [ks]	cinnamon	muscat (nut/blo	lemon peel	other spices	pr
3	https://www.apetitonline	24 ks	1	50	vodka, vínovice,	6	s kusy koreni	500	500	28,2	1500	3	ano				
4	https://www.idnes.cz/h	1 kg	1,3358695	30	1l rum+ 200 ml H	rada	jen palenka nebo vc. sirupu	250	500	20,3	1000	5	spetka	spetka	1 lzice, kandovana		po
5	https://prima-receptar.c	tři hrsti	1	40	vodka-režná-slivo	3-4	jen palenka ("z alud.liker")	250	500	24,5	750	ano	ano	ano	ano	bady an	po
6	https://abecedaz.ahrady	15 ks	1	40	vodka	6	s kusy koreni	500	500	22,6	1500	4	cela		z 1/2 citronu	3 jalovce	po
7	https://www.toprecepty.	24 ks	1	40	konak, slivovice	6	s kusy koreni	500	500	22,6	1500		15 g	4 g			po
8	https://www.recepty.onli	20 ks	1	40	zitna	3	s kusy koreni bez vanilky	200	500	24,9	600	1	kousek			5 jalovce, vanilka	15
9	**https://www.jaktak.cz**	20 ks	1	80	lih	4	s kusy koreni i sirupem	750	1250	30,1	900	5 g	2 ks rozlamat		1/2 balicku pome	rozkrajeny citron	po
10	https://www.jaktak.cz/j	dvoji macerace															
11																	

Problem: Many variations

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
1	source	# nuts	alcohol	alcohol	alcohol	extraction	extraction	sirup		% alc. after mixing							
2			[l]	[%]	form	[weeks]	method	sugar [g]	water[g]		total sugar	clove [ks]	cinnamon	muscat (nut/blo	lemon peel	other spices	pr
3	https://www.apetitonline	24 ks	1	50	vodka, vínovice,	6	s kusy koreni	500	500	28,2	1500	3	ano				
4	https://www.idnes.cz/h	1 kg	1,3358695	30	1l rum+ 200 ml H	rada	jen palenka nebo vc. sirupu	250	500	20,3	1000	5	spetka	spetka	1 lzice, kandovana		po
5	https://prima-receptar.c	tři hrsti	1	40	vodka-režná-slivo	3-4	jen palenka ("z alud.liker")	250	500	24,5	750	ano	ano	ano	ano	bady an	po
6	https://abecedaz ahrady	15 ks	1	40	vodka	6	s kusy koreni	500	500	22,6	1500	4	cela		z 1/2 citronu	3 jalovce	po
7	https://www.toprecepty.	24 ks	1	40	konak, slivovice	6	s kusy koreni	500	500	22,6	1500		15 g	4 g			po
8	https://www.recepty.onli	20 ks	1	40	zitna	3	s kusy koreni bez vanilky	200	500	24,9	600	1	kousek			5 jalovce, vanilka	15
9	https://www.jaktak.cz	20 ks	1	80	lih	4	s kusy koreni i sirupem	750	1250	30,1	900	5 g	2 ks rozlamat		1/2 balicku pome	rozkrajeny citron	po
10	https://www.jaktak.cz/j	dvoji macerace															
11																	

Important ingredients are known

How to get the best combination? (hint: run a few)

Which ones to try?

Problem: Many variations

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
1	source	# nuts	alcohol	alcohol	alcohol	extraction	extraction	sirup		% alc. after mixing							
2			[l]	[%]	form	[weeks]	method	sugar [g]	water[g]		total sugar	clove [ks]	cinnamon	muscat (nut/blo	lemon peel	other spices	pr
3	https://www.apetitonline	24 ks	1	50	vodka, vínovice,	6	s kusy koreni	500	500	28,2	1500	3	ano				
4	https://www.idnes.cz/h	1 kg	1,3358695	30	1l rum+ 200 ml H	rada	jen palenka nebo vc. sirupu	250	500	20,3	1000	5	spetka	spetka	1 lzice, kandovana		po
5	https://prima-receptar.c	tři hrsti	1	40	vodka-režná-slivo	3-4	jen palenka ("z alud.liker")	250	500	24,5	750	ano	ano	ano	ano	bady an	po
6	https://abecedaz.ahrady	15 ks	1	40	vodka	6	s kusy koreni	500	500	22,6	1500	4	cela		z 1/2 citronu	3 jalovce	po
7	https://www.toprecepty.	24 ks	1	40	konak, slivovice	6	s kusy koreni	500	500	22,6	1500		15 g	4 g			po
8	https://www.receptyonli	20 ks	1	40	zitna	3	s kusy koreni bez vanilky	200	500	24,9	600	1	kousek			5 jalovce, vanilka	15
9	https://www.jaktak.cz	20 ks	1	80	lih	4	s kusy koreni i sirupem	750	1250	30,1	900	5 g	2 ks rozlamat		1/2 balicku pome	rozkrajeny citron	po
10	https://www.jaktak.cz/j	dvoji macerace															
11																	

Important ingredients are known

How to get the best combination? (hint: run a few)

Which ones to try?

(crucial)

derive process description

naive approach: thousands of combinations

Design of Experiments

What combinations shall we try?

Design of Experiments (DoE)

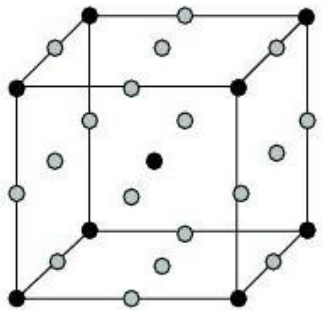
naive approach: too *many* experiments

DoE approach:

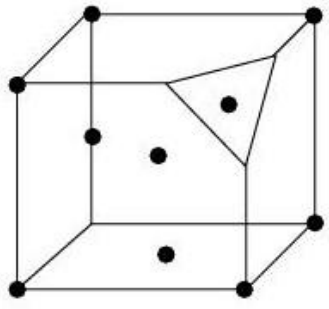
- *a few designed* experiments
- process description
- optimisation

optimal DoE: stronger assumptions for reduced number of runs

cause-effect relation



Full factorial



D-optimal design

credits: https://www.researchgate.net/figure/Classical-design-of-experiments_fig3_308164853

How-to in four steps

1. describe

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 m + \beta_4 m^2 + \beta_5 xm + \beta_6 v + \beta_7 d + \epsilon$$

2. generate designs

3. run

4. analyse

1. Describe + 2. Generate designs

inputs:

- reduction: 12 → 4
- min's + max's
- assumed relations

input	symbol	unit	min	max
walnuts	**x**	g	330	550
sugar	**m**	g	600	1500
initial water	**v**	g	0	851
extraction	**d**	week	3	6

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 m + \beta_4 m^2 + \beta_5 xm + \beta_6 v + \beta_7 d + \epsilon$$

1. Describe + 2. Generate designs

inputs:

- reduction: 12 → 4
- min's + max's
- assumed relations

designs:

- classical, optimal, others
- selection criteria
- software: R + SAS

input	symbol	unit	min	max
walnuts	**x**	g	330	550
sugar	**m**	g	600	1500
initial water	**v**	g	0	851
extraction	**d**	week	3	6

3. Run

4. Analyse

objective: taste *rank* (ordinal)

surprisingly hard task

univariate objective

details: blog

The Best Recipe

Final - and Ultimate - Recipe

ingredients:

- 550 g green walnuts (picked at June 24) **[input x]**
- 1 liter of 80 % alcohol **[input v]**
- 600 g plain sugar **[input m]**
- spices
- water to dilute the alcohol

steps:

- soak nuts (cleansed, sliced) to the alcohol;
- keep the jar lid untight at a sunny place, stir occasionally
- strain the nuts after 3-6 weeks **[input d]**
- add the spices and the sirup: put the water on to boil, dissolve the sugar, let it cool,
- let it ferment; the longer, the better

Thank You

📫 linkedin.com/in/vojtech-filipec/

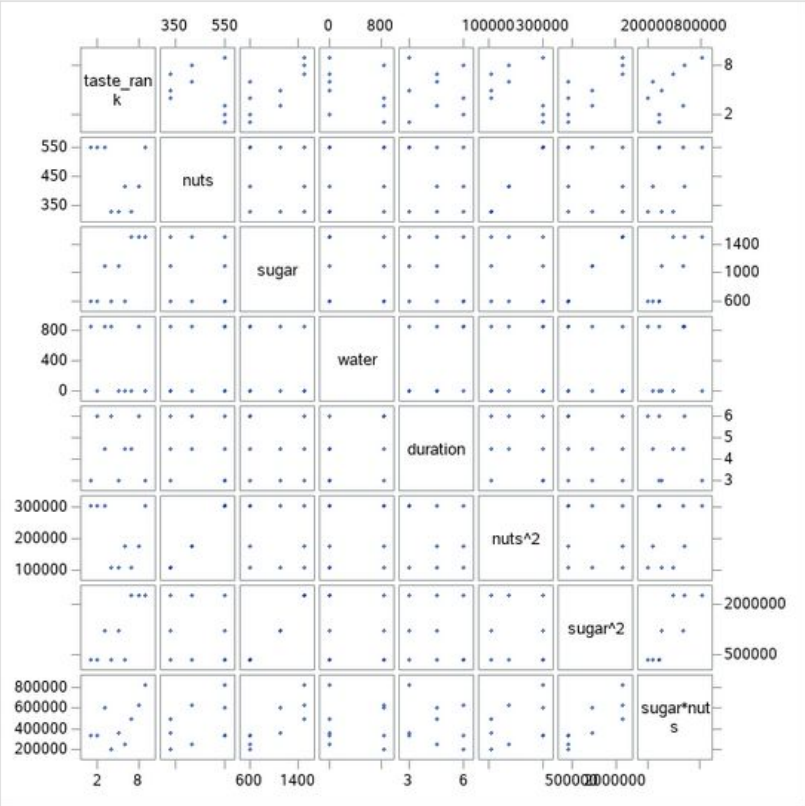
🔎 github.com/vojtech-filipec

credits: https://www.festisite.com/posters/uncle-sam-wants-you/

Additional Sources

- My blog about ořechovka: https://vojtech-filipec.github.io/pages/orechovka1.html
- NIST/SEMATECH e-Handbook of Statistical Methods: https://www.itl.nist.gov/div898/handbook/pmd/section3/pmd3.htm
- monography: Optimal Design of Experiments: A Case Study Approach: https://www.wiley.com/en-us/Optimal+Design+of+Experiments%3A+A+Case+Study+Approach-p-9780470744611

Data analysis



taste *rank* ⇒ordinal regression

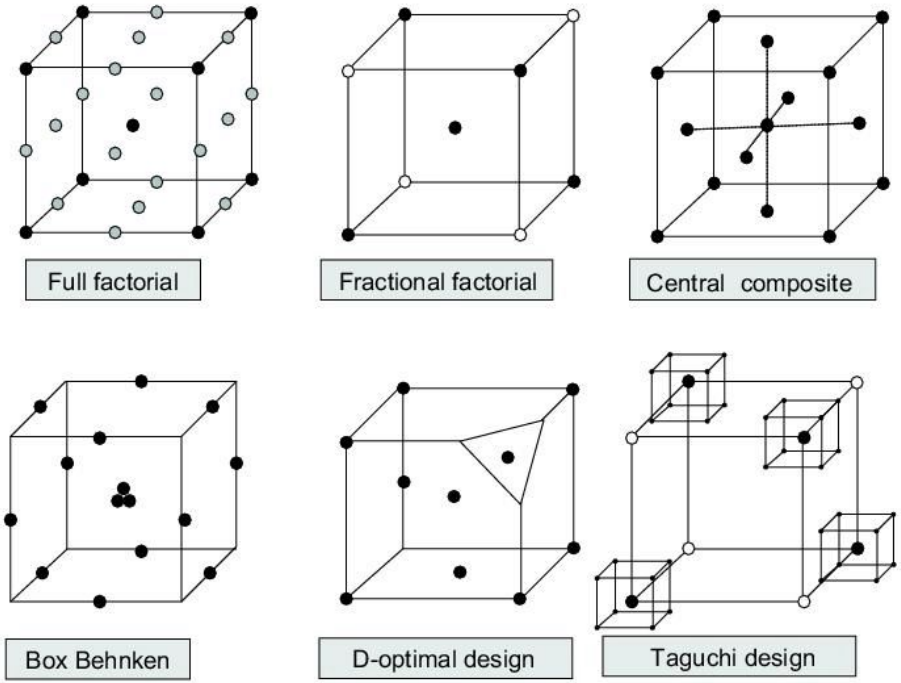
(linear or beta regression)

non-trivial task:

- 9 batches (o.o.w. 3 extra sweet)
- variable significance
- (quasi-)complete separation issue
- ?∃ of max. likelihood estimate

conclusion: diversity of recipes justified

Types of designs



https://www.researchgate.net/profile/Tp_Scholcz/publication/308164853/figure/fig3/AS:406881282609155@1474019465440/Classical-design-of-experiments.png