

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

DEPARTMENT OF
INDUSTRIAL ENGINEERING 

Machine learning

Lesson #8

Academic year 2025-2026

Prof. Pierantonio Facco

CAPE-Lab, Computer-Aided Process Engineering Laboratory

Email: pierantonio.facco@unipd.it

URL: <https://research.dii.unipd.it/capelab/>

Example on food consumption

Food example

- Let's consider the example of food consumption in different European countries:
 - 16 observations
 - European countries
 - 20 variables
 - food consumption
- **AIM: understanding** how food consumption pattern vary between different countries

	observations
1	Germany
2	Italy
3	France
4	Holland
5	Belgium
6	Luxembourg
7	England
8	Portugal
9	Austria
10	Switzerland
11	Sweden
12	Denmark
13	Norway
14	Finland
15	Spain
16	Ireland

	variables
1	Ground Coffee
2	Instant Coffee
3	Tea
4	Sweetner
5	Biscuits
6	Powder Soup
7	Tin Soup
8	Instant Potatoes
9	Frozen Fish
10	Frozen Vegetables
11	Apples
12	Oranges
13	Tin Fruit
14	Jam
15	Garlic
16	Butter
17	Margarine
18	Olive Oil
19	Yoghurt
20	Crisp Bread

Food example: data observation

Country	Ground Coffee	Instant Coffee	Tea	Sweetner	Biscuits	Powder Soup	Tin Soup	Instant Potatoes	Frozen Fish	Frozen Vegetables	Apples	Oranges	Tin Fruit	Jam	Garlic	Butter	Margarine	Olive Oil	Yoghurt	Crisp Bread
Germany	90	49	88	19	57	51	19	21	27	21	81	75	44	71	22	91	85	74	30	26
Italy	82	10	60	2	55	41	3	2	4	2	67	71	9	46	80	66	24	94	5	18
France	88	42	63	4	76	53	11	23	11	5	87	84	40	45	88	94	47	36	57	3
Holland	96	62	98	32	62	67	43	7	14	14	83	89	61	81	15	31	97	13	53	15
Belgium	94	38	48	11	74	37	23	9	13	12	76	76	42	57	29	84	80	83	20	5
Luxembourg	97	61	86	28	79	73	12	7	26	23	85	94	83	20	91	94	94	84	31	24
England	27	86	99	22	91	55	76	17	20	24	76	68	89	91	11	95	94	57	11	28
Portugal	72	26	77	2	22	34	1	5	20	3	22	51	8	16	89	65	78	92	6	9
Austria	55	31	61	15	29	33	1	5	15	11	49	42	14	41	51	51	72	28	13	11
Switzerland	73	72	85	25	31	69	10	17	19	15	79	70	46	61	64	82	48	61	48	30
Sweden	97	13	93	31	██████	43	43	39	54	45	56	78	53	75	9	68	32	48	2	93
Denmark	96	17	92	35	66	32	17	11	51	42	81	72	50	64	11	92	91	30	11	34
Norway	92	17	83	13	62	51	4	17	30	15	61	72	34	51	11	63	94	28	2	62
Finland	98	12	84	20	64	27	10	8	18	12	50	57	22	37	15	96	94	17	██████	64
Spain	70	40	40	██████	62	43	2	14	23	7	59	77	30	38	86	44	51	91	16	13
Ireland	30	52	99	11	80	75	18	2	5	3	57	52	46	89	5	97	25	31	3	9

Food example: PCA model building

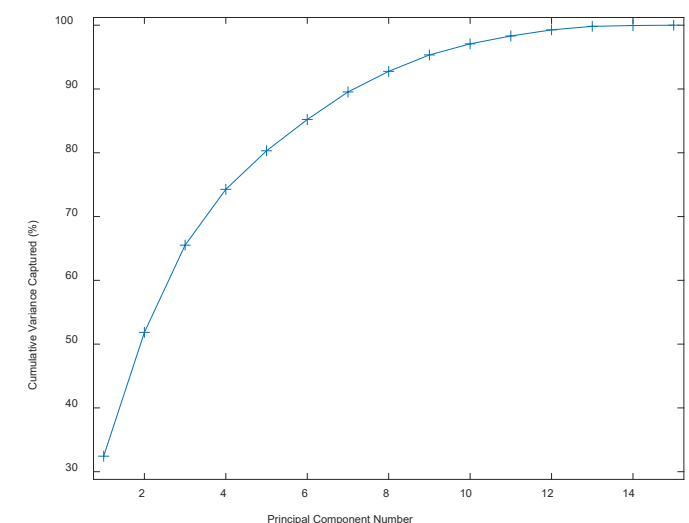
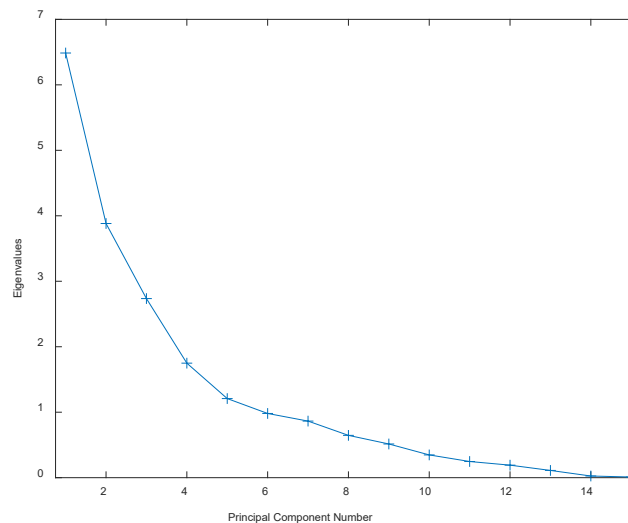
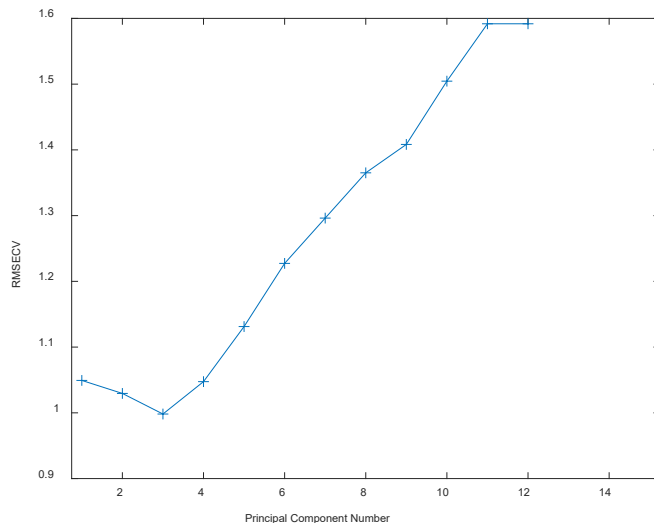
■ Preprocessing:

- data observation
- autoscaling
- missing data imputation (not part of this course)
 - use Matlab® command: `mdcheck`

■ PCA model building:

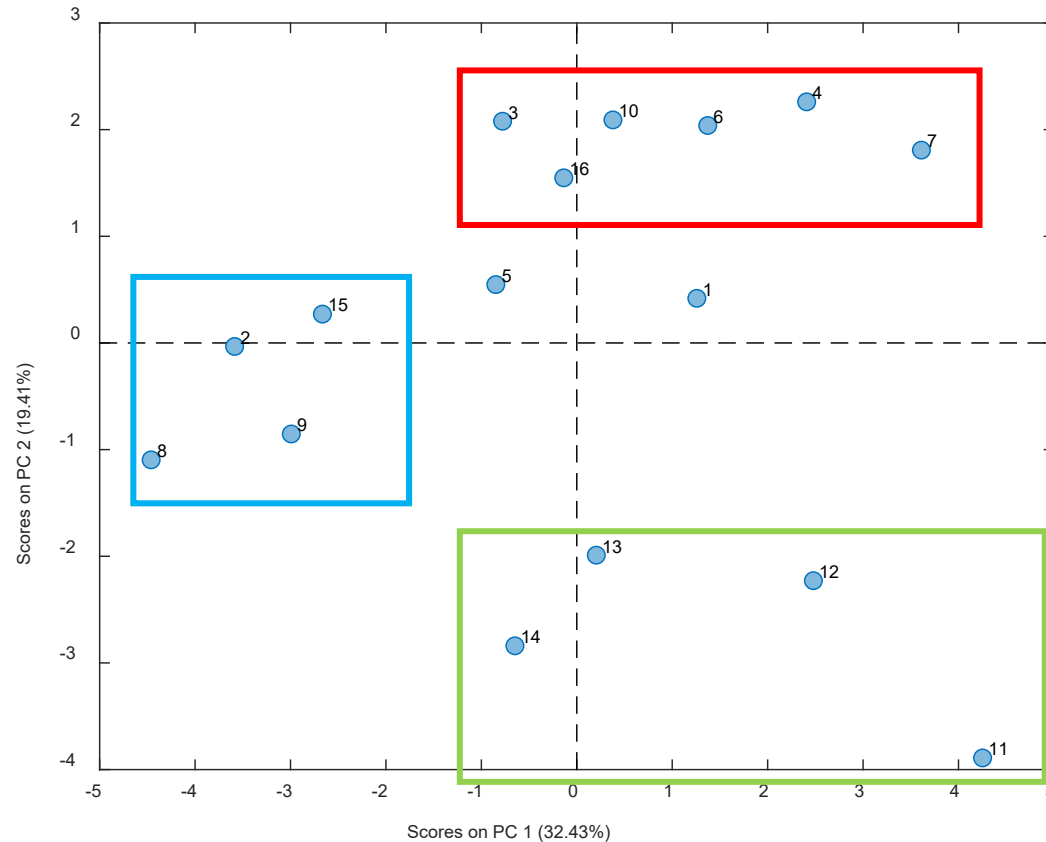
- model structure
 - $RMSECV = \sqrt{\frac{PRESS}{N}}$
 - 7-10 PCs are a good choice to explain a lot of variance
 - but 3 PCs guarantee best fitting performance in CV

PC	Eigenval.	Exp. variance (%)	Cum. explained variance (%)	RMSECV
1	6.49	32.43	32.43	1.05
2	3.88	19.41	51.84	1.03
3	2.74	13.68	65.52	1.00
4	1.75	8.75	74.27	1.05
5	1.21	6.04	80.31	1.13
6	0.98	4.9	85.21	1.23
7	0.86	4.32	89.53	1.30
8	0.65	3.23	92.76	1.37
9	0.52	2.57	95.34	1.41
10	0.35	1.74	97.08	1.50



Food example: relation among observations

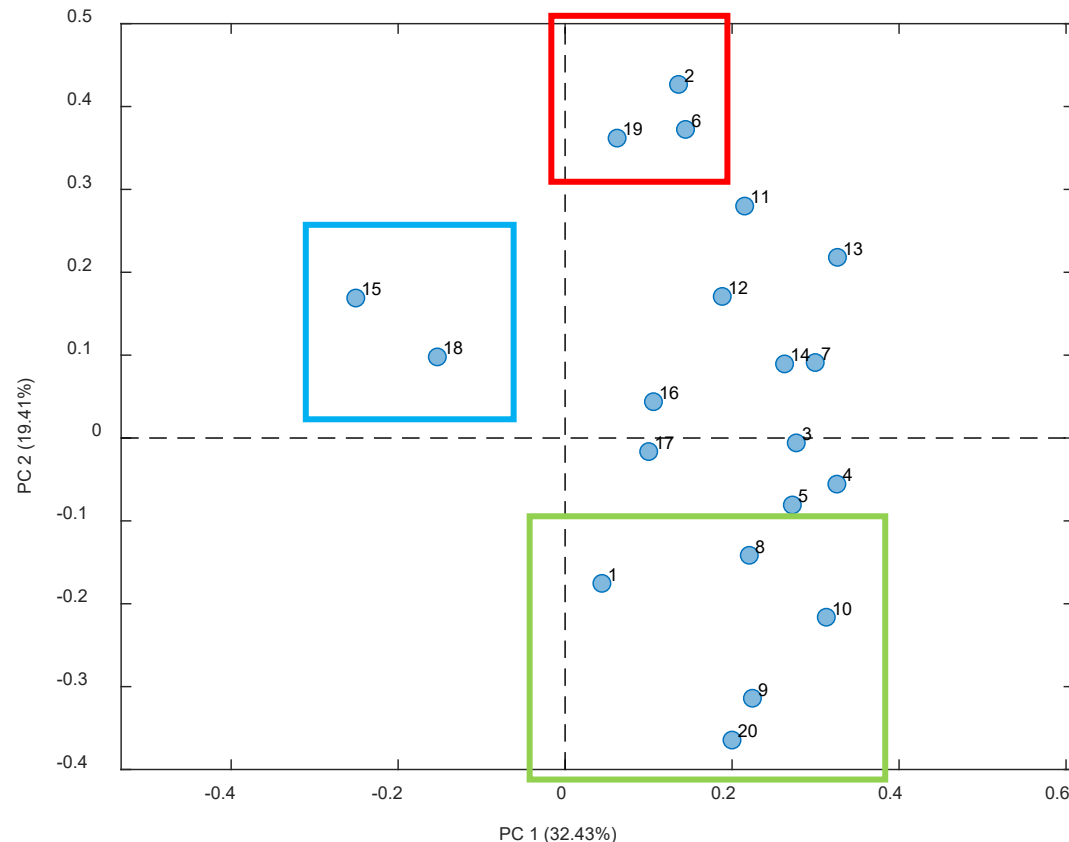
- The score plot identifies the similarities (and differences) among observations:
 - countries with similar food habits are close in the score plot



	observations
1	Germany
2	Italy
3	France
4	Holland
5	Belgium
6	Luxembourg
7	England
8	Portugal
9	Austria
10	Switzerland
11	Sweden
12	Denmark
13	Norway
14	Finland
15	Spain
16	Ireland

Food example: correlation between variables

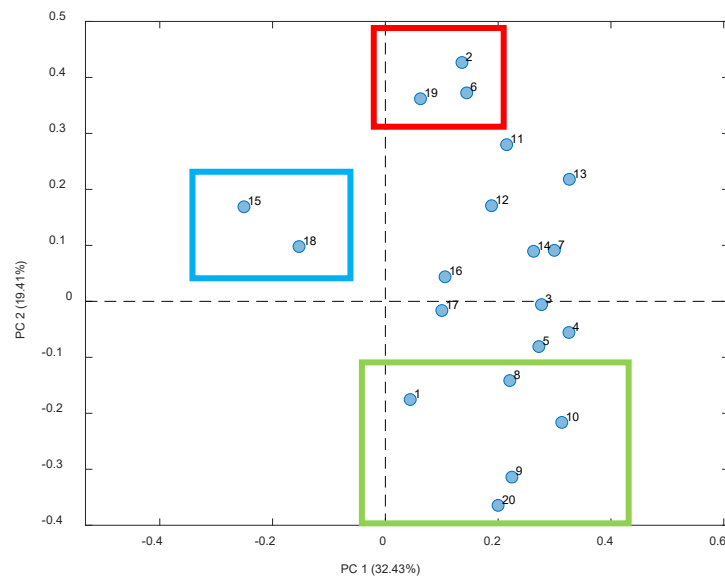
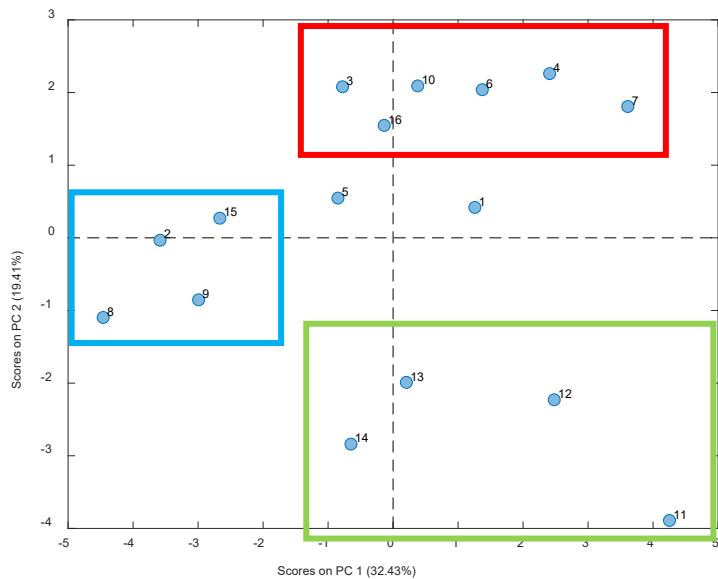
- The loading plot gives a clear idea of variable correlation
 - foods whose consumption is associated to the consumption of other types of foods



	variables
1	Ground Coffee
2	Instant Coffee
3	Tea
4	Sweetner
5	Biscuits
6	Powder Soup
7	Tin Soup
8	Instant Potatoes
9	Frozen Fish
10	Frozen Vegetables
11	Apples
12	Oranges
13	Tin Fruit
14	Jam
15	Garlic
16	Butter
17	Margarine
18	Olive Oil
19	Yoghurt
20	Crisp Bread

Food example: bi-plot

- The joint reading of scores and loadings (**bi-plot**) give important information for the variables that motivate similarities among observations
 - information on what food product motivates the similarities and the differences among countries



	observations
1	Germany
2	Italy
3	France
4	Holland
5	Belgium
6	Luxembourg
7	England
8	Portugal
9	Austria
10	Switzerland
11	Sweden
12	Denmark
13	Norway
14	Finland
15	Spain
16	Ireland

	variables
1	Ground Coffee
2	Instant Coffee
3	Tea
4	Sweetner
5	Biscuits
6	Powder Soup
7	Tin Soup
8	Instant Potatoes
9	Frozen Fish
10	Vegetables
11	Apples
12	Oranges
13	Tin Fruit
14	Jam
15	Garlic
16	Butter
17	Margarine
18	Olive Oil
19	Yoghurt
20	Crisp Bread

Food example: data observation

Country	Ground Coffee	Instant Coffee	Tea	Sweetner	Biscuits	Powder Soup	Tin Soup	Instant Potatoes	Frozen Fish	Frozen Vegetables	Apples	Oranges	Tin Fruit	Jam	Garlic	Butter	Margarine	Olive Oil	Yoghurt	Crisp Bread
Germany	90	49	88	19	57	51	19	21	27	21	81	75	44	71	22	91	85	74	30	26
Italy	82	10	60	2	55	41	3	2	4	2	67	71	9	46	80	66	24	94	5	18
France	88	42	63	4	76	53	11	23	11	5	87	84	40	45	88	94	47	36	57	3
Holland	96	62	98	32	62	67	43	7	14	14	83	89	61	81	15	31	97	13	53	15
Belgium	94	38	48	11	74	37	23	9	13	12	76	76	42	57	29	84	80	83	20	5
Luxembourg	97	61	86	28	79	73	12	7	26	23	85	94	83	20	91	94	94	84	31	24
England	27	86	99	22	91	55	76	17	20	24	76	68	89	91	11	95	94	57	11	28
Portugal	72	26	77	2	22	34	1	5	20	3	22	51	8	16	89	65	78	92	6	9
Austria	55	31	61	15	29	33	1	5	15	11	49	42	14	41	51	51	72	28	13	11
Switzerl	73	72	85	25	31	69	10	17	19	15	79	70	46	61	64	82	48	61	48	30
Sweden	97	13	93	31		43	43	39	54	45	56	78	53	75	9	68	32	48	2	93
Denmark	96	17	92	35	66	32	17	11	51	42	81	72	50	64	11	92	91	30	11	34
Norway	92	17	83	13	62	51	4	17	30	15	61	72	34	51	11	63	94	28	2	62
Finland	98	12	84	20	64	27	10	8	18	12	50	57	22	37	15	96	94	17		64
Spain	70	40	40		62	43	2	14	23	7	59	77	30	38	86	44	51	91	16	13
Ireland	30	52	99	11	80	75	18	2	5	3	57	52	46	89	5	97	25	31	3	9

... and now let's practice



... per sempre a fianco a me!

