

Noah Caulfield

Dr. Edgar Eduardo Ceh Varela

CS 460

3/4/2024

HW2: Data Preprocessing and Visualization

1.

```

1. size color gender price weight bought
2 XL red female 199.0 500.0 yes
3 L green male 89.0 450.0 no
4 M blue NaN 99.0 300.0 yes
5 NaN green female 129.0 99.0 no
6 M red female 79.0 410.0 yes
7 M green male 89.0 99.0 no

```

2.

```

1. size color gender price weight bought
2. 0 XL red female 199.0 500.0 yes
3. 1 L green male 89.0 450.0 no
4. 2 M blue male NaN 300.0 yes
5. 3 M green female 129.0 NaN no
6. 4 M red female 79.0 410.0 yes
7. 5 M green male 89.0 NaN no

```

3.

	weight	weight_cut_light	weight_cut_normal	weight_cut_heavy
0	75.0	1	0	0
1	78.5	0	1	0
2	85.0	0	0	1
3	91.0	0	0	1
4	84.5	0	0	1
5	83.0	0	0	1
6	68.0	1	0	0

4.

	currency	number
0	[PLN, USD]	2
1	[EUR, USD, PLN, CAD]	4
2	[GBP]	1
3	[JPY, CZK, HUF]	3
4	[]	0

5.

	currency	PLN_flag
0	[PLN, USD]	1
1	[EUR, USD, PLN, CAD]	1
2	[GBP]	0
3	[JPY, CZK, HUF]	0
4	[]	0

6.

```
[[-0.4785 -0.2137 -0.4584 -0.5114  0.0636]
 [ 1.3742  0.3895  1.2924  1.4055 -0.616 ]
 [ 0.3777 -0.0178  0.3905  0.2783  0.7424]
 [-0.4986 -0.4185 -0.4427 -0.5531  0.4723]
 [-0.7449 -1.1397 -0.7236 -0.745  0.0771]
 [ 1.8496  2.1747  1.9719  1.8048  1.3537]
 [ 2.2534  0.5208  2.2644  2.4466  0.5363]
 [ 0.9819 -1.0039  0.9382  0.8882 -0.0203]
 [-0.2294 -0.8236 -0.236  -0.3987  0.6429]
 [-0.069  -0.6544 -0.1338 -0.1639 -2.1663]
 [-0.2666  1.2665 -0.3344 -0.3457 -0.7709]
 [ 0.7557 -0.1692  0.7035  0.6847 -0.7183]]
```

7.

```

[[-0.3491 -1.4385 -0.4117 -0.3905 -1.8637]
 [-0.2047  0.3126 -0.1337 -0.2759  1.0781]
 [-0.3293 -0.2151 -0.3174 -0.3644 -1.5799]
 [ 1.0274  2.0898  1.0469  0.9176  0.3163]
 [ 1.829   0.696   1.7637  1.7838 -0.3337]]
[[-0.4681 -0.1417 -0.4447 -0.486   0.2934]
 [ 1.3645  0.4996  1.3064  1.3344 -0.3917]
 [ 0.3788  0.0665  0.4043  0.264   0.9777]
 [-0.4879 -0.3594 -0.429  -0.5256  0.7054]
 [-0.7315 -1.1261 -0.71   -0.7079  0.307  ]]

```

8. Part 1:

	FileName	Content
0	BattleCreekDec19_2019.txt	Thank you. Thank you. Thank you to Vice Presid...
1	BemidjiSep18_2020.txt	There's a lot of people. That's great. Thank y...
2	CharlestonFeb28_2020.txt	Thank you. Thank you. Thank you. All I can say...
3	CharlotteMar2_2020.txt	I want to thank you very much. North Carolina,...
4	CincinnatiAug1_2019.txt	Thank you all. Thank you very much. Thank you ...
5	ColoradorSpringsFeb20_2020.txt	Hello Colorado. We love Colorado, most beautif...
6	DallasOct17_2019.txt	Thank you. Thank you very much. Hello Dallas. ...
7	DesMoinesJan30_2020.txt	I worked so hard for this state. I worked so h...
8	FayettevilleSep19_2020.txt	What a crowd, what a crowd. Get those people o...
9	FayettevilleSep9_2019.txt	Thank you everybody. Thank you and Vice Presi...
10	FreelandSep10_2020.txt	We brought you a lot of car plants, Michigan. ...
11	GreenvilleJul17_2019.txt	Thank you very much. Thank you. Thank you. Tha...
12	HendersonSep13_2020.txt	Thank you, thank you. Wow. Wow, and I'm thrill...
13	HersheyDec10_2019.txt	Well, thank you to Vice President Pence. Than...
14	LasVegasFeb21_2020.txt	Well, thank you very much. And hello Las Vegas...
15	LatrobeSep3_2020.txt	So thank you Pennsylvania, very much. I'm thri...
16	LexingtonNov4_2019.txt	Thank you very much and thank you to the origi...
17	MilwaukeeJan14_2020.txt	Well thank you very much. And I'm thrilled to ...
18	MindenSep12_2020.txt	Well, I thank you very much. So I want to star...
19	MinneapolisOct10_2019.txt	Thank you very much. Thank you, Minnesota. Thi...
20	ModestoSep17_2020.txt	Thank you. Thank you very much. Thank you very...

Part 2:

	Content	City	Date	Day	Month	Year
0	Thank you. Thank you. Thank you to Vice Presid...	BattleCreek	2019-12-19	19	12	2019
1	There's a lot of people. That's great. Thank y...	Bemidji	2020-09-18	18	09	2020
2	Thank you. Thank you. Thank you. All I can say...	Charleston	2020-02-28	28	02	2020
3	I want to thank you very much. North Carolina,...	Charlotte	2020-03-02	2	03	2020
4	Thank you all. Thank you very much. Thank you ...	Cincinnati	2019-08-01	1	08	2019
5	Hello Colorado. We love Colorado, most beautif...	ColoradoSprings	2020-02-20	20	02	2020
6	Thank you. Thank you very much. Hello Dallas. ...	Dallas	2019-10-17	17	10	2019
7	I worked so hard for this state. I worked so h...	DesMoines	2020-01-30	30	01	2020
8	What a crowd, what a crowd. Get those people o...	Fayetteville	2020-09-19	19	09	2020
9	Thank you everybody. Thank you and Vice Presi...	Fayetteville	2019-09-09	9	09	2019
10	We brought you a lot of car plants, Michigan. ...	Freeland	2020-09-10	10	09	2020
11	Thank you very much. Thank you. Thank you. Tha...	Greenville	2019-07-17	17	07	2019
12	Thank you, thank you. Wow. Wow, and I'm thrill...	Henderson	2020-09-13	13	09	2020
13	Well, thank you to Vice President Pence. Than...	Hershey	2019-12-10	10	12	2019
14	Well, thank you very much. And hello Las Vegas...	LasVegas	2020-02-21	21	02	2020
15	So thank you Pennsylvania, very much. I'm thri...	Latrobe	2020-09-03	3	09	2020
16	Thank you very much and thank you to the origi...	Lexington	2019-11-04	4	11	2019
17	Well thank you very much. And I'm thrilled to ...	Milwaukee	2020-01-14	14	01	2020
18	Well, I thank you very much. So I want to star...	Minden	2020-09-12	12	09	2020
19	Thank you very much. Thank you, Minnesota. Thi...	Minneapolis	2019-10-10	10	10	2019
20	Thank you, thank you very much. Thank you very...	Mosinee	2020-09-17	17	09	2020
21	Hello, Manchester, and I am thrilled to be in ...	NewHampshire	2020-02-10	10	02	2020
22	Thank you very much everybody. Thank you. Wow...	NewHampshire	2019-08-15	15	08	2019
23	Hello, everybody. Hello, everybody. Wow. Hello...	NewHampshire	2020-08-28	28	08	2020
24	Wow, thank you. Thank you, New Mexico. Thank ...	NewMexico	2019-09-16	16	09	2019
25	Wow, that's a big crowd. This is a big crowd. ...	Ohio	2020-09-21	21	09	2020
26	Thank you very much, Phoenix. We love to be b...	Phoenix	2020-02-19	19	02	2020
27	Doesn't have the power. Doesn't have the stayi...	Pittsburgh	2020-09-22	22	09	2020
28	Hello, Houston. I am so thrilled to be here in...	Texas	2019-09-23	23	09	2019
29	Well, thank you very much. Vice President Mike...	Toledo	2020-01-09	9	01	2020
30	Thank you. thank you. So we begin. Oklahoma. w...	Tulsa	2020-06-20	20	06	2020

Part 3:

1	There's a lot of people. That's great. Thank y...	Bemidji	2020-09-18	18	09	2020	0.000355	0.000059	0.000237	0.000000
2	Thank you. Thank you. Thank you. All I can say...	Charleston	2020-02-28	28	02	2020	0.000528	0.000106	0.000106	0.000000
3	I want to thank you very much. North Carolina,...	Charlotte	2020-03-02	2	03	2020	0.000300	0.000600	0.000150	0.000450
4	Thank you all. Thank you very much. Thank you ...	Cincinnati	2019-08-01	1	08	2019	0.000734	0.000367	0.000857	0.000245
5	Hello Colorado. We love Colorado, most beautif...	ColoradoSprings	2020-02-20	20	02	2020	0.000429	0.000086	0.000172	0.000086
6	Thank you. Thank you very much. Hello Dallas. ...	Dallas	2019-10-17	17	10	2019	0.000286	0.000381	0.000476	0.000190
7	I worked so hard for this state. I worked so h...	DesMoines	2020-01-30	30	01	2020	0.000513	0.000598	0.000171	0.000171
8	What a crowd, what a crowd. Get those people o...	Fayetteville	2020-09-19	19	09	2020	0.000371	0.000062	0.000247	0.000124
9	Thank you everybody. Thank you and Vice Presi...	Fayetteville	2019-09-09	9	09	2019	0.001196	0.000217	0.000217	0.000435
10	We brought you a lot of car plants, Michigan. ...	Freeland	2020-09-10	10	09	2020	0.000592	0.000099	0.000000	0.000395
11	Thank you very much. Thank you. Thank you. Tha...	Greenville	2019-07-17	17	07	2019	0.000566	0.000189	0.000283	0.000660
12	Thank you, thank you. Wow. Wow, and I'm thrill...	Henderson	2020-09-13	13	09	2020	0.000336	0.000224	0.000000	0.000336
13	Well, thank you to Vice President Pence. Than...	Hershey	2019-12-10	10	12	2019	0.000407	0.000306	0.000102	0.000204
14	Well, thank you very much. And hello Las Vegas...	LasVegas	2020-02-21	21	02	2020	0.000072	0.000145	0.000290	0.000072
15	So thank you Pennsylvania, very much. I'm thri...	Latrobe	2020-09-03	3	09	2020	0.000321	0.000080	0.000000	0.000161
16	Thank you very much and thank you to the origi...	Lexington	2019-11-04	4	11	2019	0.001341	0.000223	0.000559	0.000447
17	Well thank you very much. And I'm thrilled to ...	Milwaukee	2020-01-14	14	01	2020	0.000423	0.000634	0.000106	0.000211
18	Well, I thank you very much. So I want to star...	Minden	2020-09-12	12	09	2020	0.000427	0.000213	0.000142	0.000142
19	Thank you very much. Thank you, Minnesota. Thi...	Minneapolis	2019-10-10	10	10	2019	0.000172	0.000000	0.000773	0.000258
20	Thank you, thank you very much. Thank you very...	Mosinee	2020-09-17	17	09	2020	0.000550	0.000206	0.000000	0.000138
21	Hello, Manchester, and I am thrilled to be in ...	NewHampshire	2020-02-10	10	02	2020	0.000608	0.000152	0.000000	0.000456
22	Thank you very much everybody. Thank you. Wow...	NewHampshire	2019-08-15	15	08	2019	0.000887	0.000197	0.001282	0.000099
23	Hello, everybody. Hello, everybody. Wow. Hello...	NewHampshire	2020-08-28	28	08	2020	0.001663	0.000333	0.000111	0.000443
24	Wow, thank you. Thank you, New Mexico. Thank ...	NewMexico	2019-09-16	16	09	2019	0.000522	0.000348	0.000435	0.000087
25	Wow, thank you very much. This is a big crowd...	Ohio	2020-08-04	04	08	2020	0.000405	0.000070	0.000000	0.000105

Part 4:

	Y_M	r_vote	r_tax	r_campaign	r_economy
0	2019_Nov	0.000539	0.000216	0.000108	0.000108
1	2019_Sep	0.001196	0.000217	0.000217	0.000435
2	2020_Sep	0.000550	0.000206	0.000000	0.000138
3	2020_Feb	0.000528	0.000106	0.000106	0.000000
4	2020_Sep	0.000718	0.000539	0.000090	0.000539
5	2020_Sep	0.000427	0.000213	0.000142	0.000142
6	2019_Sep	0.000522	0.000348	0.000435	0.000087
7	2019_Oct	0.000172	0.000000	0.000773	0.000258
8	2019_Sep	0.000000	0.002010	0.000402	0.000804
9	2020_Jun	0.000445	0.000178	0.000178	0.000089
10	2020_Mar	0.000300	0.000600	0.000150	0.000450
11	2020_Feb	0.000429	0.000086	0.000172	0.000086
12	2020_Aug	0.001663	0.000333	0.000111	0.000443

Part 5:

