

**Федеральное государственное образовательное бюджетное учреждение
высшего образования**

«ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ РФ»

Департамент анализа данных и машинного обучения

Отчет по практике №6

по дисциплине «эконометрика»

Студента группы ПМ23-1

Факультета
информационных
технологий и анализа
больших данных

Тищенко И.С.

Преподаватель
Михайлова С.С.

Москва 2024

1. Варианты работы определяется следующим образом, по последней цифре порядкового номера:

Последняя цифра номера в списке группы	ВАРИАНТ
1, 5	1
2, 6, 9	2
3, 7, 0	3
4, 8	4

Согласно этой таблице, у меня вариант 3. Но этот вариант понадобится лишь в задании 3.2 .

2. Разработать алгоритм компонентного анализа на языке Python.

- a. Для начала, нужно центрировать данные. Для этого из матрицы экзогенной переменной нужно вычесть ее среднее значение.
- b. На полученных центрированных данных составим матрицу ковариаций.
- c. Для полученной матрицы ковариаций вычислим собственные значения и собственные векторы.
- d. Отсортируем собственные значения по убыванию и возьмем по их индексам векторы.
- e. Составим матрицу собственных векторов по отсортированным по убыванию соответствующим собственным значениям.
- f. Умножим матрицу центрированных данных на матрицу векторов.
- g. Соответствующим значением PCA для каждого объекта будет первое значение координат для каждого вектора из результата прошлого пункта.

3.

- a. Перейдем к заданию 3.1

Применим разработанный алгоритм к датасету.

Доля объясненной дисперсии (вклад каждой компоненты):

0.53578359; 0.17582261; 0.13514881; 0.07013181; 0.06386609
;0.01924709

Факторные нагрузки (вклад признаков в компоненты):

	x1	x2	x3	x4	x5	PCA_score
PC1	0.584892	0.317800	0.690811	0.045668	0.148361	0.235771
PC2	0.481138	-0.339428	-0.340384	-0.487029	-0.227244	0.498595
PC3	0.383696	-0.191943	0.124163	-0.220006	-0.353221	-0.792055
PC4	0.184029	-0.634601	0.011681	0.153899	0.724522	-0.121085
PC5	0.480578	0.444735	-0.621468	0.362585	0.175198	-0.151234
PC6	0.119850	-0.382677	0.071713	0.746433	-0.495942	0.175872

Ранжированные страны по удовлетворённости:

№	Страна	PCA_Score
35	Парагвай	29,04924
41	Гондурас	26,99154
1	Израиль	26,98324
34	Боливия	26,55881
33	Таиланд	26,29496
39	Гана	26,20219
36	Филиппины	26,11995
40	Намибия	25,36624
7	Польша	25,22839
27	Гайана	24,69975
8	Чили	24,67121
47	Марокко	24,37005
48	Никарагуа	22,23468
32	Шри-ланка	19,24653
31	Сальвадор	19,08698
6	Португалия	18,68722
2	Греция	18,25616
28	Ямайка	17,91133
45	Таджикистан	17,52262
26	Эквадор	16,62729
67	Судан	14,96176
46	Вьетнам	14,37059
30	Доминиканская Республика	12,63961
49	Индия	12,47689

42	Индонезия	12,07743
18	Казахстан	11,2305
52	Бангладеш	10,51209
50	Камбоджи	9,534377
15	Перу	9,325298
66	Малави	9,141965
55	Йемен	9,126656
43	Киргизия	8,896796
10	Хорватия	8,883375
5	Венгрия	8,116765
3	Словакия	6,82587
12	Тринидад и Тобаго	6,055459
77	Мозамбик	5,896722
22	Иран	5,833436
4	Эстония	3,015837
76	Чад	2,522276
62	Ангола	0,451156
44	Юар	-0,14153
78	Нигер	-0,31725
68	Афганистан	-1,09256
16	Албания	-2,42762
57	Мавритания	-3,48785
20	Босния и Герцеговина	-4,08717
19	Азербайджан	-4,36555
23	Югославия	-4,62726
29	Турция	-5,22078
13	Сербия	-5,68825
58	Нигерия	-6,08764
9	Латвия	-7,46487
17	Россия	-7,7387
72	Центрально-Африканская Республика	-9,01173
54	Камерун	-9,87834
38	Молдова	-10,5669
37	Ботсвана	-10,7323
79	Конго	-12,1205
11	Болгария	-12,2954
14	Белоруссия	-16,3086
69	Гвинея	-17,2788
75	Либерия	-18,4359
59	Уганда	-18,6942
70	Эфиопия	-20,2469
21	Украина	-21,4123
53	Гана	-21,9249
64	Замбия	-22,8266
51	Кения	-22,8276

25	Армения	-23,1714
80	Зимбабве	-24,8926
74	Буркина-Фасо	-28,7353
56	Мадагаскар	-28,7567
60	Сенегал	-31,5106
61	Гаити	-31,5603
65	Руанда	-31,8307
24	Грузия	-33,4472
73	Мали	-34,5584
63	Танзания	-35,0007
71	Сьерра Леон	-43,2314

Исходя из ранжированных данных – Парагвай – лучшая страна.

б. Перейдем к заданию 3.2

Сделаем аналогичные действия для данных из задания 3.2. Мой вариант - Вариант 3, а значит, мои признаки из экзогенной переменной: x_2 ; x_4 ; x_5 ; x_6 ; x_8 ; x_9 ; x_{10} .

Доля объясненной дисперсии (вклад каждой компоненты):

9.26088181e-01; 7.18164959e-02; 1.33235626e-03; 4.08136189e-04;

1.80205529e-04; 1.43967129e-04; 3.06579631e-05;

Факторные нагрузки (вклад признаков в компоненты):

	x_1	x_2	x_3	x_4	x_5	x_6	PCA_score
PC1	0.015968	0.035834	-0.003201	-0.000293	-0.210680	-0.011864	0.976690
PC2	0.016506	-0.007660	-0.008392	-0.016211	0.977118	-0.015468	0.210563
PC3	0.073979	0.158361	-0.137684	-0.285273	-0.015870	-0.931859	-0.022299
PC4	0.592966	0.779150	-0.042090	0.129238	0.007726	0.146844	-0.034930
PC5	0.350346	-0.344860	0.053958	0.819433	0.000025	-0.289712	0.003834
PC6	0.720007	-0.493762	-0.025452	-0.471008	-0.022723	0.121525	0.002694
PC7	0.035186	0.061443	0.987741	-0.091298	0.005150	-0.104849	0.000217

Предоставим ранжированные данные о регионах РФ:

	Регион	PCA_Score
56	Тюменская область	750,0184
17	г.Москва	608,7571
76	Сахалинская область	465,6594
78	Чукотский Автономный округ	447,1291
70	Республика Саха	137,767
19	Республика Коми	135,1055
27	г.Санкт-Петербург	115,5961
24	Мурманская область	91,05058
75	Магаданская область	90,33092
64	Красноярский край	74,29252
69	Томская область	67,23712
71	Камчатский край	61,4935
9	Московская область	59,97951
43	Республика Татарстан	53,26285
21	Вологодская область	51,95824
20	Архангельская область	49,3314
23	Ленинградская область	45,37505
46	Пермский край	34,17005
51	Самарская область	31,92814
8	Липецкая область	25,44222
55	Свердловская область	25,24135
73	Хабаровский край	17,21786
66	Кемеровская область	15,84185
0	Белгородская область	14,03121
49	Оренбургская область	11,43919
22	Калининградская область	10,67449
57	Челябинская область	-2,55971
65	Иркутская область	-4,91882
40	Республика Башкортостан	-6,27086
68	Омская область	-9,31639
67	Новосибирская область	-12,803
18	Республика Карелия	-12,9263
48	Нижегородская область	-18,6083
72	Приморский край	-20,908
74	Амурская область	-22,9811
25	Новгородская область	-24,219
44	Удмуртская Республика	-27,4283
32	Волгоградская область	-30,9118
16	Ярославская область	-30,963
77	Еврейская Автономная область	-36,9424
31	Астраханская область	-40,4987
30	Краснодарский край	-42,8999
5	Калужская область	-44,1823
61	Республика Хакасия	-45,9072
59	Республика Бурятия	-51,6504

7	Курская область	-56,3221
63	Забайкальский край	-57,9455
15	Тульская область	-59,2353
33	Ростовская область	-59,4746
14	Тверская область	-62,0545
45	Чувашская Республика	-63,1192
52	Саратовская область	-64,9179
12	Смоленская область	-67,4591
53	Ульяновская область	-72,0119
3	Воронежская область	-74,101
42	Республика Мордовия	-75,5548
2	Владимирская область	-75,8919
11	Рязанская область	-77,199
6	Костромская область	-77,8488
62	Алтайский край	-79,5506
10	Орловская область	-81,494
47	Кировская область	-81,8604
41	Республика Марий Эл	-83,7133
54	Курганская область	-86,2055
39	Ставропольский край	-89,9705
50	Пензенская область	-91,7233
13	Тамбовская область	-91,9739
26	Псковская область	-96,3864
58	Республика Алтай	-99,8974
1	Брянская область	-100,666
29	Республика Калмыкия	-110,151
37	Карачаево-Черкесская Республика	-111,002
34	Республика Дагестан	-112,01
36	Кабардино-Балкарская Республика	-115,553
60	Республика Тыва	-116,367
28	Республика Адыгея	-117,202
4	Ивановская область	-118,136
38	Республика Северная Осетия - Алания	-118,824
35	Республика Ингушетия	-157,612

Таким образом Тюменская область заняла лидирующую позицию в экономических показателях.