

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ЭКОНОМИКЕ

УДК 519.237.8 : 330

Кластерный анализ качества жизни населения стран мира с применением пакета STATISTICA®

© 2012 г. В.И. Перова, Б. Хаймер

Нижегородский государственный университет им. Н.И. Лобачевского
mmes@mm.unn.ru

Проведена многомерная кластеризация 169 стран мира по показателям, составляющим Индекс развития человеческого потенциала. Применены две процедуры кластеризации – иерархический метод и метод *K*-средних. Приведено сравнение результатов кластеризации.

Ключевые слова: человеческий капитал, Индекс развития человеческого потенциала, страны мира, кластерный анализ, иерархический метод, метод *K*-средних, пакет STATISTICA®.

Обеспечение устойчивого экономического роста является одной из важнейших задач реальной экономической политики государств. Несмотря на то, что мировая экономика в целом показывает стабильные темпы роста, экономический рост варьируется в различных странах мира. В настоящее время в качестве важнейшего внутреннего фактора экономического развития рассматривается человеческий капитал, который оказывает серьезные воздействия на производительность труда и на темпы экономического роста [1; 2]. Он представляет собой совокупность накопленных профессиональных знаний, умений и навыков, получаемых в процессе образования и повышении квалификации, которые в последствии могут приносить доход – в виде заработной платы, процента или прибыли [3]. В связи с этим является актуальным проведение анализа уровня жизни государств с целью определения перспектив их дальнейшего развития и улучшения качества жизни населения на основе данных, отражающих влияние различных форм капитала и демографической ситуации. В данной работе базой для исследования были взяты показатели [4]:

Expected_live – Ожидаемая продолжительность жизни при рождении (лет);

Avg_stady – Средняя продолжительность обучения (лет);

Expected_stady – Ожидаемая продолжительность обучения (лет);

VND – ВНД на душу населения (доллары США).

Эти показатели входят в состав Индекса развития человеческого потенциала (ИРЧП) – метода измерения человеческого развития, который был введен Программой развития ООН в 1990 г. в Докладе о развитии человека.

Двумя эффективными процедурами кластерного анализа [5; 6; 7] – иерархическим методом и методом *K*-средних – была проведена многомерная кластеризация 169 стран мира на основе стандартизованных данных. В работе применялась одна из наиболее признанных в

мировой практике статистическая система обработки и анализа данных STATISTICA® [8]. При проведении кластеризации иерархическим методом использовались инструменты кластерного анализа [5; 6]: мера расстояния между объектами – расстояние городских кварталов (манхэттенское расстояние)

$$M_{im} = \left| X_i^{(1)} - X_m^{(1)} \right| + \left| X_i^{(2)} - X_m^{(2)} \right| + \dots + \left| X_i^{(k)} - X_m^{(k)} \right|$$

и правило связи – метод полной связи

$$\delta_{\max}(G_p, G_q) = \max_{x_p \in G_p, x_q \in G_q} \delta(x_p, x_q).$$

В результате кластеризации иерархическим методом анализируемая совокупность стран была объединена в 3 кластера, имеющих следующий состав:

В кластер № 1 вошли 32 страны: Норвегия, Австралия, Новая Зеландия, США, Ирландия, Лихтенштейн, Нидерланды, Канада, Швеция, Германия, Япония, Швейцария, Франция, Израиль, Финляндия, Исландия, Бельгия, Дания, Испания, Гонконг (Китай), Греция, Италия, Люксембург, Австрия, Соединённое королевство, Сингапур, Андорра, ОАЭ, Бруней Даруссалам, Катар, Республика Корея, Кувейт.

Кластер № 2 составила 81 страна, в том числе Российская Федерация, Чешская Республика, Словения, Словакия, Польша, Венгрия, Португалия, Эстония, Литва, Латвия, Украина, Бразилия, Турция, Алжир, Китай, Казахстан, Беларусь, Армения, Мальта, Кипр, Бахрейн и другие страны.

Кластер № 3 образовали 56 стран, в числе которых Индия, Пакистан, Афганистан, Камбоджа, Никарагуа, Марокко и страны центральной и южной Африки.

В табл. 1 и табл. 2 представлены средние значения показателей в кластерах и интервалы изменения каждого показателя:

Таблица 1
Средние значения показателей в кластерах

Номер кластера	Средние значения			
	Expected_live	Avg_stady	Expected_stady	VND
1	80.3	10.1	15.6	41100.22
2	72.6	8.8	13.2	10938.22
3	57.3	4.0	8.5	2200.96

Таблица 2

Интервалы изменения показателей в кластерах

Номер кластера	Интервал (минимум, максимум)			
	<i>Expected_live</i>	<i>Avg_stady</i>	<i>Expected_stady</i>	<i>VND</i>
1	76 – Катар, 83.2 – Япония	2.5 – Новая Зеландия, 12.6 – Норвегия	11.5 – ОАЭ, 20.5 – Австралия	22917 – Исландия, 81011 – Лихтенштейн
2	52 – ЮАР, 96.3 – Таиланд	4.7 – Мальдивские Острова, 12.3 – Чешская Республика	10.4 – Вьетнам, 16.7 – Словения	2020 – Таджикистан, 26664 – Бахрейн
3	44.6 – Афганистан, 73.8 – Никарагуа	1.2 – Мозамбик, 7.2 – Зимбабве	4.3 – Нигер, 11.2 – Кабо-Верде	176 – Зимбабве, 22218 – Экваториальная Гвинея

Из табл. 1 и табл. 2 следует, что в *кластер № 1* вошли страны с высокими показателями уровней материального благосостояния, продолжительности жизни и обучения населения. *Кластер № 2* составили страны со средними уровнями среднедушевого дохода, продолжительности жизни и обучения населения. Страны, объединившиеся в *кластер № 3*, имеют низкий уровень материального благосостояния населения, продолжительности жизни и обучения.

Рассмотрим, например, показатели, характеризующие ИРЧП в Российской Федерации и Алжире (табл. 3), поскольку один из авторов – Хаймер Б. является гражданином Алжира и в настоящее время обучается в магистратуре ННГУ. Обе страны распределились в кластер № 2. Из табл. 1 и табл. 3 следует, что показатели в России выше показателей для Алжира и выше средних значений по кластеру, кроме ожидаемой продолжительности жизни при рождении (*Expected_live*).

Таблица 3

–	<i>Expected_live</i>	<i>Avg_stady</i>	<i>Expected_stady</i>	<i>VND</i>
Россия	67.2	8.8	14.1	15258
Алжир	72.9	7.2	12.8	8320

При проведении кластеризации методом K -средних было произведено разбиение исходных данных также на 3 кластера. Здесь в качестве меры расстояний использовано евклидово расстояние

$$d_{im} = \left[(X_i^{(1)} - X_m^{(1)})^2 + (X_i^{(2)} - X_m^{(2)})^2 + \dots + (X_i^{(k)} - X_m^{(k)})^2 \right]^{1/2},$$

а в качестве правила связи – центроидный метод [5; 6]. На рис. 1 представлен график средних значений показателей в каждом кластере:

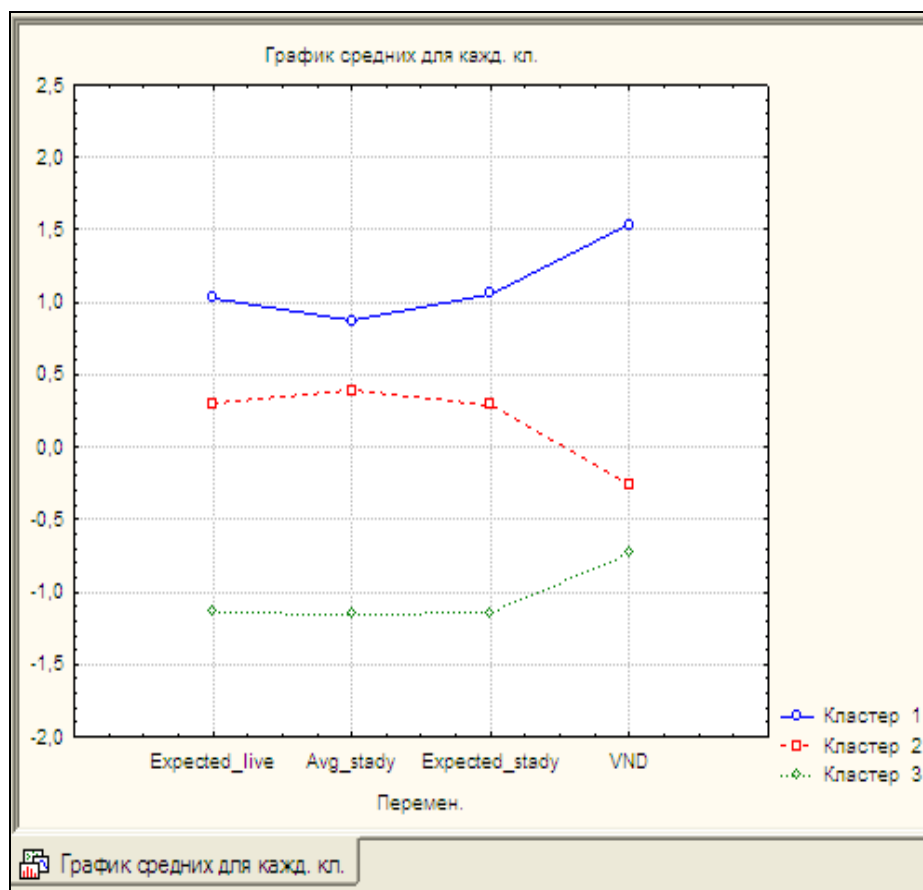


Рис. 1. Графики средних значений показателей в кластерах

В результате было получено, что кластер № 1 образовали 39 стран, кластер № 2 – 75 стран и кластер № 3 – 55 стран. По сравнению с результатами иерархической кластеризации страны: Чешская Республика, Словения, Словакия, Мальта, Кипр, Бахрейн, Португалия перешли из кластера № 2 со средними показателями в кластер № 1, характеризующийся высокими показателями, а Никарагуа – из кластера № 3 с низкими показателями в кластер № 2. Различия в результатах можно объяснить тем, что в иерархическом методе были взяты в качестве правила связи метод полной связи и в качестве метрики – манхэттенское расстояние, а в методе K -средних всегда используются, соответственно, центроидный метод и евклидово расстояние.

Таким образом, проведенные исследования позволяют оценить качество жизни населения стран мира и способствовать нахождению путей его повышения.

Список литературы

1. Кузнецов Ю.А. Оптимальное управление экономическими системами: Учебное пособие. Нижний Новгород: Изд-во Нижегородского госуниверситета, 2008, 449 с.
2. Кузнецов Ю.А., Мичасова О.В. Обобщение модели экономического роста с учетом накопления человеческого капитала. I. // Вестник ННГУ. Серия: Математическое моделирование. Оптимальное управление. – 2010. – №1. – С. 168–175.
3. Кузнецов Ю.А. Роль человеческого капитала в инновационной экономике, основанной на знаниях // Актуальные проблемы социально-экономического развития и пути их решения: Материалы Первой Региональной научно-практической конференции (Дзержинск, 18 марта 2010 г.): Нижний Новгород: Изд-во НИСОЦ, 2010. – С. 21–27.
4. Доклад о развитии человека 2010. Реальное богатство народов: пути к развитию [Электронный ресурс]: Пер. с англ., ПРООН. – М.: Изд-во «Весь Мир», 2010, 244 с. – Режим доступа: http://www.un.org/ru/development/hdr/2010/hdr_2010_complete.pdf, свободный.
5. Балабанов А.С., Стронгина Н.Р. Анализ данных в экономических приложениях: учебное пособие. Н. Новгород: Изд-во ННГУ, 2004, 135 с.
6. Кузнецов Ю.А., Перова В.И. Кластерный анализ в экономических приложениях с применением ППП «STATISTICA»: Учебно-методическое пособие. Нижний Новгород: Нижегородский госуниверситет, 2009, 88 с.
7. Перова В.И., Сиверухина А.С. Кластерный анализ уровня развития государств с применением информационных технологий // Актуальные проблемы социально-экономического развития и пути их решения: Материалы Первой Региональной научно-практической конференции (Дзержинск, 18 марта 2010 г.): Нижний Новгород: Изд-во НИСОЦ, 2010.– С. 228–232.
8. Боровиков В.П. Программа STATISTICA для студентов и инженеров. М.: КомпьютерПресс, 2001, 301 с.