

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра социальной информатики

**ВОЗМОЖНОСТИ КЛАСТЕРНОГО АНАЛИЗА В
ИССЛЕДОВАНИИ СОЦИАЛЬНОЙ СТРАТИФИКАЦИИ**

(автореферат бакалаврской работы)

студента 5 курса 531 группы
направления 09.03.03 - Прикладная информатика
профиль Прикладная информатика в социологии
Социологического факультета
Невейкина Андрея Валерьевича

Научный руководитель
старший преподаватель

_____ Ю.А. Седавкина
подпись, дата

Зав. кафедрой
кандидат социологических наук, доцент

_____ И.Г. Малинский
подпись, дата

Саратов 2022

ВВЕДЕНИЕ

Актуальность исследования. Социальная стратификация в общем виде, представляет собой разделение какой-либо части общества на группы, содержание которых будет схожим, но сами группы будут отличаться друг от друга.

Страты – слои, на которые можно разбить человеческое общество по различным критериям, например, по силе, влиянию, положению и т.д.

Человечество во все времена своего существования, так или иначе делилось на страты. Ярким примером может служить сословие, рабство, касты в Индии и т.д. В настоящее время выделяют такие виды глобальной стратификации как: экономическая, политическая, профессиональная.¹

Говоря о стратификации, речь зачастую идет о больших выборках респондентов, что в свою очередь усложняет процесс исследования. В 30-ые годы XX века, было внедрено такое понятие, как кластерный анализ. Историческим «отцом» данного термина принято считать Беркли Роберта Трайона.

Говоря кратко, кластер – объединение нескольких схожих элементов в одну группу.

Кластерный анализ – математический метод обработки статистических данных, задача которого стоит в том, чтобы разделить какую-либо выборку на группы. Члены одной группы должны обладать схожими проявлениями переменных, а члены разных групп различными.²

Возможности кластерного анализа довольно подробно можно рассмотреть в ПО SPSS Statistics.

¹ Социальная структура общества [Электронный ресурс] URL: <https://foxford.ru/wiki/obschestvoznanie/sotsialnaya-struktura-obschestva-stratifikatsiya-obschestva> (Дата обращения 04.03.2022) Загл. с экрана яз. русс.

² Кластерный анализ [Электронный ресурс] URL: https://science.fandom.com/ru/wiki/Кластерный_анализ (Дата обращения 04.03.2022) Загл. с экрана Яз.рус.

SPSS statistics – компьютерная программа для статистической обработки данных, предназначенная для проведения прикладных исследований в общественных науках.¹

Норман Най, Хедли Халл и Дейл Бент разработали первую версию системы в 1968 году.

Вышеописанная информация позволяет определить цели и задачи исследования.

Объект исследования: Кластерный анализ в SPSS Statistics

Предмет исследования: Возможности кластерного анализа в исследовании социальной стратификации

Цель исследования: проанализировать возможности кластерного анализа в исследовании социальной стратификации (на примере программы IBM SPSS Statistics).

Задачи исследования:

1. Рассмотреть кластерный анализ как способ социальной стратификации.
2. Выявить разницу между видами кластеризации в SPSS.
3. На основе полученных статистических данных, вывести и проинтерпретировать дендрограмму.
4. Выделить социальные классы при помощи одного из способов кластеризации в SPSS.

Эмпирической базой выпускной квалификационной работы являются результаты авторского исследования на тему «Экономическая стратификация жителей города Саратова» (N=100 человек);

Структура бакалаврской работы. Выпускная квалификационная работа состоит из введения, двух разделов, заключения, списка источников и приложений.

¹ SPSS Statistics [Электронный ресурс] URL: https://spravochnick.ru/sociologiya/statisticheskii_analiz_dannyh_v_spss/ (Дата обращения 8.03.2022) Загл.с экрана яз. русс.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Первая глава «Теоретический пласт в исследовании социальной стратификации» содержит в себе теоретический разбор таких понятий как, иерархический кластерный анализ и теории социальной стратификации.

Иерархический кластерный анализ при помощи определенного алгоритма, выявляет относительно схожие группы. Суть алгоритма же состоит в том, чтобы рассмотреть каждое наблюдение как отдельный кластер, после чего последовательно объединить кластеры, до тех пор, пока все кластеры не объединятся в один.

Важно отметить, что Главной особенностью иерархического кластерного анализа в SPSS является то, что исследователю не обязательно знать, число кластеров, на которое необходимо разделить статистику. При помощи различных методов, мер, которые представлены в меню SPSS Statistics, а также дендрограммы или сосульчатой диаграммы процедура иерархического кластерного анализа автоматически предлагает исследователю несколько вариаций разбиений.

Дендрограмма- древовидная диаграмма, которая содержит определенное количество уровней, каждый из которых соответствует одному из шагов последовательного увеличения кластеров. Главная суть дендрограммы, состоит в том, чтобы наглядно показать, на какое возможное число кластеров мы можем разбить наше исследование.

Были рассмотрены различные методы и меры сходства иерархического кластерного анализа.

Для лучшего понимания социальной стратификации, были рассмотрены следующие теории:

- Теория К. Маркса
 - Статусная теория М. Вебера
 - Теория социальной стратификации П. А. Сорокина
 - Функционализм

Анализируя эти теории, было принято решение провести авторское исследование на тему: «Экономическая стратификация жителей города Саратова». Также данные теории позволили выдвинуть больше индикаторов, для выявления социальных классов в Авторском исследовании.

Глава 2 «Методологические основания изучения экономической стратификации жителей города Саратова»

Объект исследования: Жители города Саратова

Жители города Саратова
Пол
Возраст
Образование
Социальный статус
Семейное положение
Ежемесячный доход

Предмет исследования: Материальное положение респондентов в обществе

Материальное положение респондента в обществе
Оценка материального положения
Оценка уверенности в будущем
Оценка самостоятельности
Оценка коммуникабельности
Оценка лидерских качеств
Оценка значимости в обществе в повседневной жизни
Наличие недвижимости
Наличие авто средства
Наличие вкладов/инвестиций

Цель: стратифицировать респондентов по экономическому признаку

Задачи:

1. Определить оценку Материального положения
2. Выявить оценку личностных характеристик респондентов
3. Определить число возможных социальных классов
4. Выявить зависимость оценок личностных характеристик респондента от их материального положения

Для выявления социального портрета респондентов, были выведены частотные таблицы, и получены следующие данные:

В исследовании приняли участие 100 человек 38% женщин, 62% мужчин. Средний возраст респондентов находится в диапазоне от 20 до 65 лет. Большинство респондентов (44%) имеет среднее профессиональное образование. Подавляющая часть опрошенных (79%) имеет место работы. Основная часть респондентов (57%) находятся в зарегистрированных отношениях (женаты /замужем). Средний доход респондентов находится в промежутке от 30 до 60 т.р.

Посчитав личностные оценки респондентов, получаем следующие результаты:

- Средняя оценка материального положения - 2,98 балла.
- Средняя оценка уверенности в будущем - 3,33 балла.
- Средняя оценка самостоятельности - 3,73 балла.
- Средняя оценка коммуникабельности – 3,35 балла.
- Средняя оценка лидерских качеств – 3,16 балла.
- Средняя оценка значимости в обществе – 3,17 балла

По наличию недвижимости респонденты довольно равно распределились на две группы. Первая (49%) не является собственником недвижимого имущества, вторая (46%) имеет в собственности 1 единицу недвижимого имущества.

Так же стало ясно, что большая часть респондентов (63%) имеет в собственности 1 единицу авто средства.

И наконец, основная часть респондентов (46%) не имеет вклады /инвестиции.

В главе III «Экономическая стратификация жителей города Саратова», была проведена кластеризация ранее полученных данных. Был использован Иерархический кластерный анализ. Была выведена дендрограмма по методу Варда, в качестве меры был использован «Квадрат расстояния Евклида».

Для того, чтобы содержание кластеров было «открыто», были использованы таблицы сопряженности. В поле строка вносились переменные 1-15. В поле столбец, была внесена переменная «WardMethod».

После чего для каждого кластера были рассчитаны средние значения.

ЗАКЛЮЧЕНИЕ

В ходе исследования было опрошено 100 респондентов. Из них 62% мужчин и 38 % женщин. При помощи дендрограммы построенной по методу Варда, среди выборки респондентов было определено 3 социальных класса.

В первый класс вошло 37% респондентов. (67,57% мужчин, 32,43% женщин). Средняя оценка материального положения для которых равна 1,95, что можно интерпретировать как «денег хватает на еду, но покупка одежды вызывает трудности».

Какую-либо возрастную группу выделить не приходится, ввиду равного распределения респондентов. Данный класс в основном имеет среднее профессиональное образование. Люди этого класса, как правило работающие, однако стоит отметить, что большая часть безработных людей, входит именно в этот класс. Семейное положение яркой зависимости не имеет, половина респондентов этого звена холоста/ не замужем, другая половина напротив, жената/ замужем. Так же в подтверждение оценки материального положения, данный класс имеет низкий ежемесячный доход (до 20 т.р.)

Для этого класса характерны довольно низкие оценки личностных характеристик.

- Уверенность в будущем - 2,41 балла
- Самостоятельность - 2,78 балла
- Коммуникабельность - 2,81
- Лидерские качества - 2,22

- Значимость в обществе в повседневной жизни - 2,57

Как правило члены данного кластера не являются собственниками недвижимости.

Половина респондентов в классе имеет авто средство, другая имеет в собственности 1 единицу.

Так же членам данного кластера не свойственно иметь вклады/инвестиции.

Во второй класс вошло 45% респондентов.(53,33% мужчин, 46,67% женщин).

Средняя оценка материального положения для этого класса равна 3,22, это можно интерпретировать как «денег хватает на питание и одежду, но покупка вещей длительного пользования вызывает затруднения».

В основном данный класс имеет возрастную группу 20-30 лет, уровень образования. Для этого класса можно выделить как среднее профессиональное, так и высшее (бакалавриат) образование. Подавляющая часть опрошенных этой страты имеет место работы. Этому классу свойственно иметь зарегистрированный брак. Основной ежемесячный доход равен 30-45 т.р.

Для этого класса характерны слегка выше среднего оценки личностных характеристик.

- Уверенность в будущем – 3,67 балла
- Самостоятельность - 4 балла
- Коммуникабельность – 3,47 балла
- Лидерские качества – 3,42 балла
- Значимость в обществе в повседневной жизни – 3,27 балла

Большая часть респондентов этого кластера имеет в собственности 1 ед. недвижимости, и 1 ед. авто средства

Так же респондентам этого кластера свойственно иметь не большие вклады/инвестиции (до 100 т.р.).

В третий кластер вошло 18% респондентов из них 72,22% мужчины, и 27,78 женщины.

Данные респонденты имеют высокую оценку материального положения равную 4,5 что можно интерпретировать как «Мы можем покупать вещи длительного пользования, затруднительна только покупка действительно дорогих вещей» и «Мы можем без труда приобретать вещи длительного пользования, купить машину».

В данный класс входят респонденты от 30-65 лет, для которых свойственно иметь высшее (бакалавриат) образование. 100% респондентов этого звена имеют место работы. Основная часть женаты/ замужем. Ежемесячный доход составляет более 60 т.р.

Для этого кластера характерны высокие оценки личностных характеристик

- Уверенность в будущем – 4,39 балла
- Самостоятельность – 5балла
- Коммуникабельность – 4,17балла
- Лидерские качества – 4,44балла
- Значимость в обществе в повседневной жизни – 4,17балла

Как правило респонденты этого звена имеют в собственности как минимум 1 квартиру.

Так же высокий коэффициент наличия авто средства показывает, что данный класс людей имеет в собственности 1 или более ед. авто средств.

Данный класс имеет вклады/инвестиции более 100 т.р.

Таким образом, данное исследование позволило убедиться, в том, что при помощи кластерного анализа можно более чем успешно стратифицировать респондентов по экономическому признаку. При помощи дендрограммы построенной по методу Варда, было выявлено 3 четких, сильно разнящихся класса (рисунок 5).

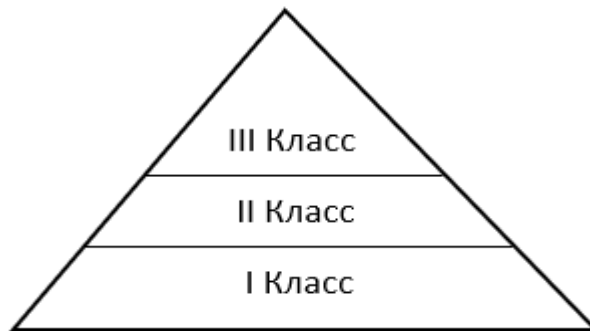


Рисунок 1 - Социальная пирамида опрошенных респондентов

При помощи таблиц сопряженности, содержание выявленных классов было «открыто».

Анализ литературы так же показал, что для большой выборки респондентов, следует использовать метод к-средних.

Была выявлена зависимость оценок личностных характеристик респондентов от их материального положения в обществе. Согласно этой зависимости, классы находящиеся у вершины пирамиды, обладают более высокими оценками личностных характеристик. Классы находящиеся у основания пирамиды имеют напротив, самые низкие оценки личностных качеств.