

Введение в статистический анализ данных

Алексей Бессуднов

НИУ ВШЭ, факультет социологии

О чем этот курс?

По окончании курса вы сможете:

- ▶ Понимать статистические модели, используемые в исследовательской литературе
- ▶ Самостоятельно проводить базовый анализ данных в пакете SPSS с использованием синтаксиса

Зачем вам это нужно?

- ▶ Для того чтобы сдать зачет и экзамен

Зачем вам это нужно?

- ▶ Для того чтобы сдать зачет и экзамен
- ▶ В большинстве компаний, так или иначе занимающихся исследованиями, используются статистические методы (не только в маркетинговых исследованиях, но и в финансах, страховании и т.д.)

Зачем вам это нужно?

- ▶ Для того чтобы сдать зачет и экзамен
- ▶ В большинстве компаний, так или иначе занимающихся исследованиями, используются статистические методы (не только в маркетинговых исследованиях, но и в финансах, страховании и т.д.)
- ▶ Статистические методы необходимы в академических исследованиях

Зачем вам это нужно?

- ▶ Для того чтобы сдать зачет и экзамен
- ▶ В большинстве компаний, так или иначе занимающихся исследованиями, используются статистические методы (не только в маркетинговых исследованиях, но и в финансах, страховании и т.д.)
- ▶ Статистические методы необходимы в академических исследованиях
- ▶ Самое главное: культура обоснованного высказывания

Как строится этот курс

- ▶ Три модуля: два в этом году, один в следующем
- ▶ На одну лекцию будут приходиться три занятия в компьютерном классе
- ▶ Письменная домашняя работа каждые две недели (предполагающая использование SPSS)
- ▶ Работы, сданные после дедлайна, не принимаются; плагиат наказывается
- ▶ Зачет после каждого модуля (в компьютерном классе); экзамен в конце 4-го модуля
- ▶ Оценка: домашние задания (40%), зачеты (два по 20%), экзамен (20%)
- ▶ Все материалы курса будут выкладываться на <http://sites.google.com/site/bessudnov/teaching>

Пример 1. Явка на выборах и партийное голосование: Москва, 2011

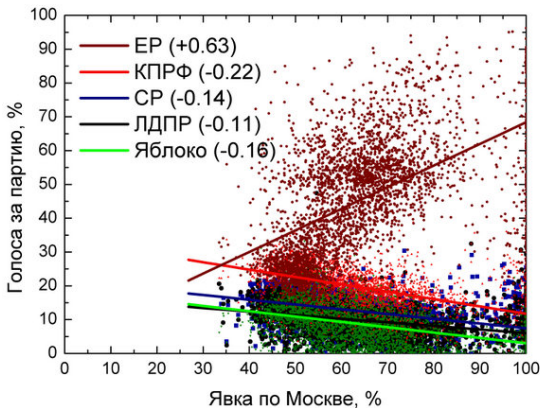


Рис.: Источник: <http://oude-rus.livejournal.com>

Явка на выборах и партийное голосование: Великобритания, 2010, консерваторы

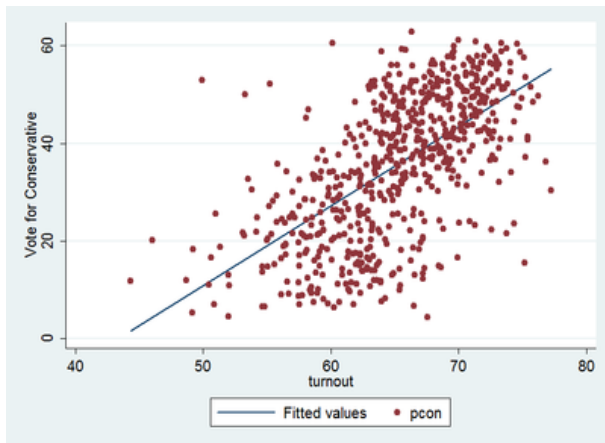


Рис.: Источник: The Electoral Commission

Явка на выборах и партийное голосование: Великобритания, 2010, лейбористы

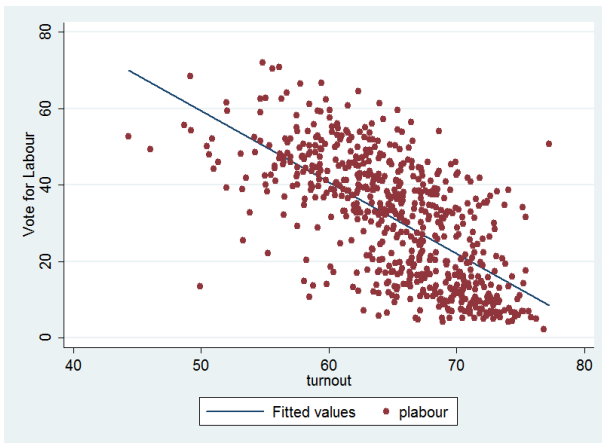


Рис.: Источник: The Electoral Commission

Пример 2. Ханс Рослинг: двести лет, которые изменили мир

Пример 3. Карта ценностей Инглхарта-Велцеля

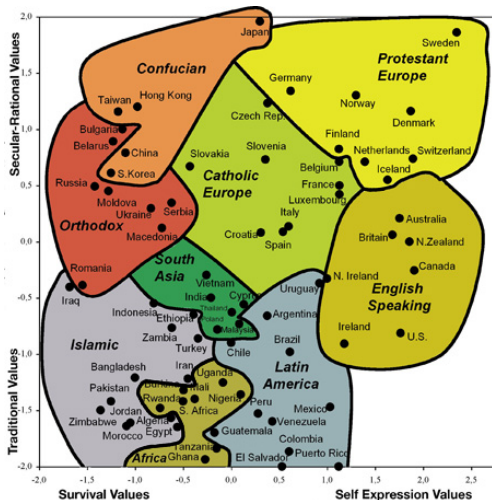


Рис.: Источник: R.Inglehart and Ch.Welzel. "Changing Mass Priorities: The Link Between Modernization and Democracy." Perspectives on Politics 2010 (vol 8, No. 2, p. 554.

Особенности статистического анализа в разных социальных науках

- ▶ Социология: опросы, сравнительно большие выборки

Особенности статистического анализа в разных социальных науках

- ▶ Социология: опросы, сравнительно большие выборки
- ▶ Политология: относительно небольшие выборки стран

Особенности статистического анализа в разных социальных науках

- ▶ Социология: опросы, сравнительно большие выборки
- ▶ Политология: относительно небольшие выборки стран
- ▶ Психология: эксперименты, относительно небольшие выборки

Особенности статистического анализа в разных социальных науках

- ▶ Социология: опросы, сравнительно большие выборки
- ▶ Политология: относительно небольшие выборки стран
- ▶ Психология: эксперименты, относительно небольшие выборки
- ▶ Эконометрика: акцент на идентификации причинно-следственных связей

Источники данных для статистического анализа

- ▶ Опросы населения

Источники данных для статистического анализа

- ▶ Опросы населения
- ▶ Данные опросов, доступные для анализа: ESS, ISSP, RLMS, GSS, WVS, etc. (см. также “Единый архив социологических данных”)

Источники данных для статистического анализа

- ▶ Опросы населения
- ▶ Данные опросов, доступные для анализа: ESS, ISSP, RLMS, GSS, WVS, etc. (см. также “Единый архив социологических данных”)
- ▶ Микроданные переписей (в США, Британии, Швеции, других странах)

Источники данных для статистического анализа

- ▶ Опросы населения
- ▶ Данные опросов, доступные для анализа: ESS, ISSP, RLMS, GSS, WVS, etc. (см. также “Единый архив социологических данных”)
- ▶ Микроданные переписей (в США, Британии, Швеции, других странах)
- ▶ Агрегированная статистика на уровне стран и регионов

Немного о философии количественного анализа данных

- ▶ Исследует среднюю тенденцию, а не индивидуальные факты

Немного о философии количественного анализа данных

- ▶ Исследует среднюю тенденцию, а не индивидуальные факты
- ▶ Описательная и объяснительная логика в анализе данных

Немного о философии количественного анализа данных

- ▶ Исследует среднюю тенденцию, а не индивидуальные факты
- ▶ Описательная и объяснительная логика в анализе данных
- ▶ Наблюдение и эксперимент

Немного о философии количественного анализа данных

- ▶ Исследует среднюю тенденцию, а не индивидуальные факты
- ▶ Описательная и объяснительная логика в анализе данных
- ▶ Наблюдение и эксперимент
- ▶ Выборка и популяция. Оценка (estimation) и инференция (inference)

Программы для статистического анализа данных

- ▶ Excel
- ▶ SPSS
- ▶ Stata
- ▶ R
- ▶ SAS, Statistica, Matlab, eViews, etc.

Обзор содержания курса: Третий модуль

- ▶ Тема 1. Введение в анализ данных. Практические занятия: введение в SPSS. Работа с синтаксисом

Обзор содержания курса: Третий модуль

- ▶ Тема 1. Введение в анализ данных. Практические занятия: введение в SPSS. Работа с синтаксисом
- ▶ Тема 2. Описательная статистика и статистические графики

Обзор содержания курса: Третий модуль

- ▶ Тема 1. Введение в анализ данных. Практические занятия: введение в SPSS. Работа с синтаксисом
- ▶ Тема 2. Описательная статистика и статистические графики
- ▶ Тема 3. Статистические гипотезы и их тестирование. Тест на равенство средних и пропорций в группах. Дисперсионный анализ

Обзор содержания курса: Третий модуль

- ▶ Тема 1. Введение в анализ данных. Практические занятия: введение в SPSS. Работа с синтаксисом
- ▶ Тема 2. Описательная статистика и статистические графики
- ▶ Тема 3. Статистические гипотезы и их тестирование. Тест на равенство средних и пропорций в группах. Дисперсионный анализ
- ▶ Тема 4. Таблицы сопряженности и их анализ. Связь между категориальными переменными

Обзор содержания курса: Третий модуль

- ▶ Тема 1. Введение в анализ данных. Практические занятия: введение в SPSS. Работа с синтаксисом
- ▶ Тема 2. Описательная статистика и статистические графики
- ▶ Тема 3. Статистические гипотезы и их тестирование. Тест на равенство средних и пропорций в группах. Дисперсионный анализ
- ▶ Тема 4. Таблицы сопряженности и их анализ. Связь между категориальными переменными
- ▶ Тема 5. Корреляция и парная регрессия

Обзор содержания курса: Третий модуль

- ▶ Тема 1. Введение в анализ данных. Практические занятия: введение в SPSS. Работа с синтаксисом
- ▶ Тема 2. Описательная статистика и статистические графики
- ▶ Тема 3. Статистические гипотезы и их тестирование. Тест на равенство средних и пропорций в группах. Дисперсионный анализ
- ▶ Тема 4. Таблицы сопряженности и их анализ. Связь между категориальными переменными
- ▶ Тема 5. Корреляция и парная регрессия
- ▶ Зачет

Обзор содержания курса: Четвертый модуль

- ▶ Тема 6. Множественная линейная регрессия (4 недели)

Обзор содержания курса: Четвертый модуль

- ▶ Тема 6. Множественная линейная регрессия (4 недели)
- ▶ Тема 7. Регрессионный модели для бинарных и категориальных зависимых переменных (4 недели)

Обзор содержания курса: Четвертый модуль

- ▶ Тема 6. Множественная линейная регрессия (4 недели)
- ▶ Тема 7. Регрессионный модели для бинарных и категориальных зависимых переменных (4 недели)
- ▶ Тема 8. Регрессионный анализ и идентификация причинно-следственных связей

Обзор содержания курса: Четвертый модуль

- ▶ Тема 6. Множественная линейная регрессия (4 недели)
- ▶ Тема 7. Регрессионный модели для бинарных и категориальных зависимых переменных (4 недели)
- ▶ Тема 8. Регрессионный анализ и идентификация причинно-следственных связей
- ▶ Зачет + экзамен

Обзор содержания курса: Первый модуль в следующем году (А.В.Рыжова)

- ▶ Методы классификации: факторный и кластерный анализ

- ▶ А.О.Крыштановский, “Анализ социологических данных”
- ▶ Ю.Н.Тюрин, А.А.Макаров, “Анализ данных на компьютере”
- ▶ Ю.Н.Толстова, “Анализ социологических данных”
- ▶ К.Доугерти, “Введение в эконометрику”
- ▶ A.Agresti, B.Finley, “Statistical methods for the social sciences”