|  |  |
| --- | --- |
|  | Институт образования **НИУ ВШЭ**основы Анализа данных (в stata) и интерпретация результатов исследований в образовании. курс 1захаров а.б., ab.zakharov@gmail.com, вт.,ср., 11.00-13.00 |

##### цели

Курс является начальным в ознакомлении слушателей с основами анализа данных и ориентирован на формирование общего представления об идеях, лежащихв основебазовыхвидов анализа. Центральная задача – научить слушателей понимать суть изучаемых методов анализа данных, работать с данными и выполнять анализ в STATAи научить трактовать результаты анализа, представленные в табличной, графической или текстовой форме. Особенностью курса является то, что его материал излагается на примерах, связанных с областью образования: слушатели знакомятся с результатами современных исследований в этой области и тренируются в анализе данных исследований и мониторингов качества образования.

|  |  |
| --- | --- |
| ***Цель*** | ***Чем проверяется*** |
| Освоить управление данными | Задания после прохождения темы 1 |
| Освоить базовые понятия статистики | Задания после прохождения темы 2 |
| Приобрести опыт анализа данных | Задания после прохождения тем 3-7, контрольная работа |
| Научиться интерпретировать результаты анализа, понимая особенности используемых данных, исследовательского дизайна, возможности и ограничения методов анализа | Задания после прохождения тем 2-7, контрольная работа |

##### ТЕМЫ ЗАНЯТИЙ

|  |  |  |
| --- | --- | --- |
| ***Дата*** | ***Содержание занятия*** | ***Что делать, сдавать к этой дате*** |
| Февраль | **Тема 1. Основы работы с STATA. Управление данными** |  |
| Февраль | **Тема 2. Базовые понятия статистики применительно к исследованиям в образовании. Описательная статистика** | Выполнить задание, заданное после предыдущей темы |
| Март | **Тема 3. Графическое представление данных** | Выполнить задание, заданное после предыдущей темы |
| Март | **Тема 4. Выборка: основные понятия. Типы выборки. Внешняя валидность** | Выполнить задание, заданное после предыдущей темы |
| Апрель | **Тема 5. Дизайн исследования. Внутренняя валидность. Возможности и ограничения разных типов исследовательского дизайна для оценки каузальных связей в образовании** | Выполнить задание, заданное после предыдущей темы |
| Апрель | **Тема 6. Связь переменных, характеристики связи** | Выполнить задание, заданное после предыдущей темы |
| Май | **Тема 7. Проверка гипотез. Сравнение средних. Дисперсионный анализ. Интерпретация результатов** | Выполнить задание, заданное после предыдущей темы |
| Май | **Контрольная работа** | Определение метода анализа по поставленному исследовательскому вопросу, подготовка данных к анализу, представление описательной статистики и анализ данных. Интерпретация полученных результатов в виде текста – до 0.5 стр. |

##### оценка

|  |  |
| --- | --- |
| ***Требование*** | ***К-во баллов*** |
| Домашнее задание | 20 |
| Контрольная работа | 40 |
| Всего максимально | 160 |

Конечная оценка зависит от числа баллов

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Баллы | 160 | 144 | 128 | 112 | 96 | 80 | 64 | 48 | 32 | 16 |
| Оценка | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |

##### Описание заданий

Самостоятельная работа слушателей включает три типа заданий:

а) освоение теоретической литературы. Источники по статистике и анализу данных представлены в разделе «Список обязательной литературы».

б) ознакомление с примерами анализа выполненного в области образования. Слушатель сдает текст объемом 0,5 - 1 стр.

в) самостоятельная работа с данными, собранными в ходе международных мониторингов качества образования или в результате исследования в области образования и интерпретация полученных результатов. Для оценивания сдаются база данных, do-file, log-file, и при прохождении тем 2, 3 6, 7 – таблицы и графики с результатами анализа, текст с интерпретацией результатов (до 0,5 страницы).

Последние два типа заданий детальнее описаны ниже.

**Тема 1. Основы работы с STATA. Управление данными**

Подготовить исследовательский вопрос, выбрать данные (timss, pisa, pirls, мониторинг экономики образования и т.п.). Выбрать переменные в соответствии с исследовательским вопросом. Открыть данные. Переименовать и перекодировать переменные. Изменить метки переменных. Удалить лишние переменные. Сохранить данные

**Тема 2. Базовые понятия статистики применительно к исследованиям в образовании. Описательная статистика**

Проанализировать один из отчетов о результатах TIMSS 2011 или PISA 2012 (на выбор слушателя). Охарактеризовать результаты России с использованием мер центральной тенденции и разброса – в сопоставлении с другими странами. О чем говорят результаты? Объем до0,5 страницы.

Создать таблицу с описательными статистиками для переменных, отобранных при прохождении темы 1. Охарактеризовать шкалы отобранных переменных.

**Тема 3. Графическое представление данных**

Интерпретировать графики одного из отчетов о результатах TIMSS 2011 или PISA 2012 (на выбор слушателя).

Используя полученную в ходе выполнения темы 1 базу данных,построить графики. Должны быть использованы все пройденные виды графиков (графики распределения, точечные, круговые, ящичковые диаграммы и др.). Интерпретировать полученные графики.

**Тема 4. Выборка: основные понятия. Типы выборки. Внешняя валидность**

Описать выборку одного из исследований (из списка рекомендованных или других статей с результатами «количественного» исследования – на выбор слушателя): что является генеральной совокупностью, единицей наблюдения, какой тип выборки, ее объем. Какие возможности и ограничения несет данная выборка. Объем 0,5 страница

**Тема 5. Дизайн исследования. Внутренняя валидность. Возможности и ограничения разных типов исследовательского дизайна для оценки каузальных связей**

Описать дизайн одного из исследований (из списка рекомендованных или других статей с результатами «количественного» исследования – на выбор слушателя). Указать возможности и ограничения дизайна для каузального анализа. Объем 0,5 страницы

**Тема 6. Связь переменных, характеристики связи**

По одной из прочитанных книг или статей (из списка рекомендованных или других статей – на выбор слушателя), содержащих описание связи переменных в образовании, дать интерпретацию результатов исследования: какие задачи исследования? Какие переменные и шкалы? Каковы характеристики связи переменных? О чем говорит обнаруженная связь между переменными? Объем 0,5 – 1 страница

Проанализировать связь переменных, выбирая метод в зависимости от типа шкал, показать значимость. Интерпретировать результаты.

**Тема 7. Проверка гипотез. Сравнение средних. Дисперсионный анализ. Интерпретация результатов**

Проанализировать одну из прочитанных книг или статей с результатами дисперсионного анализа в образовании. Интерпретировать результаты.

Поставить исследовательский вопрос, подобрать переменные, подготовить их к анализу, выполнить анализ, используя пройденные методы (T-тестслучайной выборки для сравнения с заданным значением, T-тест для двух независимых выборок, Т-тест для парных выборок, ANOVA, MANOVA).

##### Списокобязательнойлитературы

1. Baum, C. 2006. An Introduction to Modern Econometrics Using Stata, Stata Press.
2. Cameron, C., Trivedi, P. 2009. Microeconometrics Using Stata. A Stata Press Publication.
3. Hamilton, L. 2009. Statistics with Stata. Updated for Version 10. Brooks/Cole Cencage Learning.
4. Long, J. S. 2009. The Workflow of Data Analysis: Principles and Practice Stata Press.
5. Longest, K. C. 2012. Using Stata for quantitative analysis. Los Angeles [etc.] SAGE Publications
6. Mitchell, M. N. 2010. Data management using Stata: a practical handbook. Stata press.
7. Mitchell, M. N. 2008. A Visual Guide to Stata Graphics. Stata press.
8. Rabe-Hesketh, S., Everitt, B. 2007. A handbook of statistical analysis using STATA. 4th edition. Taylor&FrancisGroup.

##### Напоминания

1. Студент считается имеющим академическую задолженность, если он получил оценку ниже 4 баллов (по 10-балльной шкале) по итогам промежуточной/окончательной аттестации по Дисциплине, либо не явился на экзамен (не сдал работу в установленный срок) без уважительной причины.
2. Студенты, имеющие единовременно академические задолженности по трем или более разным Дисциплинам, к пересдачам результатов промежуточной аттестации по Дисциплинам не допускаются и подлежат отчислению непосредственно после получения академической задолженности по третьей Дисциплине.
3. Запрещается пересдача результатов промежуточной аттестации по Дисциплине с целью повышения удовлетворительной оценки (от 4 баллов по 10-ти балльной шкале).
4. Пересдача по одной и той же Дисциплине допускается не более двух раз.