ECON 6009
Graduate Seminar
Memorial University of Newfoundland

Lecture 1-*Workflow of data analysis*

INTRODUCTION

The objectives

- To strengthen the knowledge the student gained in previous Economics modules about **how to conduct academic research** and present it to an academic audience
- To develop students' capacity for self-motivated learning and problem solving during the practical process of conducting academic research

The tools

STATA 16 available at Grad Resource Room and check **GradPlan**

The tools

MiKTeX, which is free-source together with an editing interface of your choice (some of which are also free)

The tools

Scott Long's book

The Workflow of Data Analysis Using Stata by J. Scott Long, Stata Press, 2009

Why

- Replication: *being able to do things twice*
- Efficiency through automatization: *Not having to do things twice*
- Quality *(avoiding errors)*
- Collaboration
- Research ethics (replicability, data storage, confidentiality)

Advantages

- By making your work *look* professional, it provides a great signal of quality
- By making your work professional, it makes your work more professional

Advantages

- It saves you tons of time in the long run
- It saves you tons of effort in the long run
- It saves you lots of grief from trying to solve errors, looking for files, looking for references, etc.

Disadvantages

- It takes time at the beginning
- It takes some effort at the beginning
- Some other researchers (older or from other disciplines) might not be using tools that are 100% compatible

Replication

- For your own purposes
- For your colleagues
- For people who are trying to learn from you
- For journal reviewers
- For ethical purposes
- If someone challenges your work results
- If someone wants to replicate your work

Automatization

- Saves time
- Saves you effort
- Avoids mistakes
- Helps you to think more logically and carefully
- Helps you document your steps

Steps in the workflow

- Data collection
- Data cleaning
- Analysis
- Documenting analysis
- Presenting results
- Storing/protecting data
- Storing/protecting results

All these steps require careful

- Planning
- Organization
- Documentation
- Execution

Planning

- Think a bit before doing
- It will help in the end
- Periodically reassess your plans!

Planning

- Planning
- Organization
- Documentation
- Execution

Planning

- Planning
- Organization
- Documentation
- Execution

Planning

- Planning
- Organization
- Documentation
- Execution

Questions?

Any questions?

Any suggestions?

Any complaints?