In [1]:
```
from IPython.display import IFrame
import ipynb_style
from epstata import Stpy
import pandas as pd
from itertools import combinations
from importlib import reload
```

In [2]:
```
reload(ipynb_style)
ipynb_style.clean()
#ipynb_style.presentation()
#ipynb_style.pres2()
```

Out[2]:

Data Workflows in Stata and Python

(http://www.stata.com) (https://www.python.c

Data Workflows in Stata and Python

Dejan Pavlic, Education Policy Research Initiative, University of Ottawa

Stephen Childs (presenter), Office of Institutional Analysis, University of Calgary

(http://ucalgary.ca) (http://uottaw
(http://socialsciences.uottawa.ca/
epri/eng/index.asp)

EPRI Education Policy Research Initiative

Initiative de recherche sur les politiques de l'éducation

Introduction

About this talk

Objectives

- know what Python is and what advantages it has
- know how Python can work with Stata

Please save questions for the end. Or feel free to ask me today or after the conference.

Outline

- Introduction
  - Overall
  - Motivation
  - About Python
- Building Blocks
  - Running Stata from Python
  - Pandas
  - Python language features
- Workflows
  - ETL/Data Cleaning
  - Stata code generation
  - Processing Stata output

About Me

- Started using Stata in grad school (2006).
- Using Python for about 3 years.
- Post-Secondary Education sector

- University of Calgary - Institutional Analysis (https://oia.ucalgary.ca/Contact)
- Education Policy Research Initiative (http://socialsciences.uottawa.ca/irpe-epri/eng/index.asp - University of Ottawa (a Stata shop)

Motivation

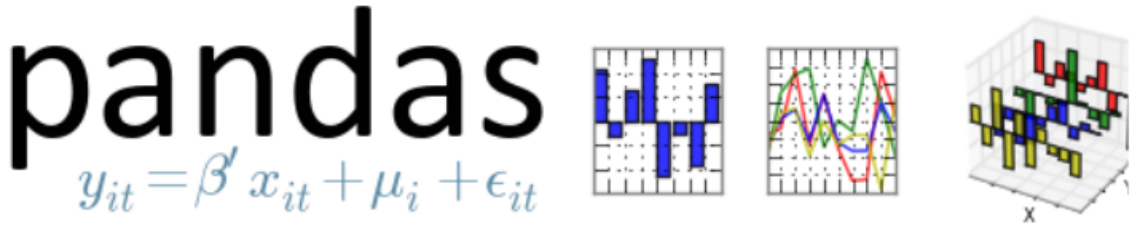
- Python is becoming very popular in the data world.
- Python skills are widely applicable.
- Python is powerful and flexible and will help you get more done, faster.

About Python

The Python Language

- General purpose programming language
- Name comes from Monty Python
- Python 2 vs. 3 - use Python 3
- "batteries included"

Scientific Python



(http://pandas.pydata.org)

(http://matplotlib.org/)

(http://www.numpy.org)