Python for Tabular Data & Visualization

GitHub Repository: http://github.com/emonson/pandas-datamatters

Monday

Session 1	9:30-10:40am	**Data structs**
Break	10:40-10:55am	
Session 2	10:55am-12:05pm	**DataFrame acc**
Lunch	12:05-1:20pm	
Session 3	1:20-2:30pm	**Built-in Plotting**
Break	2:30-2:45pm	
Session 4	2:45-3:55pm	**Tidy data & plot**

Tuesday

Session 1	9:30-10:40am	**Melt**
Break	10:40-10:55am	
Session 2	10:55am-12:05pm	**Merge**
Lunch	12:05-1:20pm	
Session 3	1:20-2:30pm	**Groupby**
Break	2:30-2:45pm	
Session 4	2:45-3:55pm	**& Pivot tables**

Python for Tabular Data & Vis
Groupby & Pivot Tables

Eric Monson, PhD

Visualization specialist
Center for Data and Visualization Sciences

Data Matters 2020

Repository: https://github.com/emonson/pandas-datamatters

Conceptual introduction

How do we look for patterns in a data set of individual events?

Order Date	Sales	Product Category
10/13/10	$261.54	Office Supplies
10/1/12	$10,123.02	Office Supplies
10/1/12	$244.57	Office Supplies
7/10/11	$4,965.76	Technology
8/28/10	$394.27	Office Supplies
8/28/10	$146.69	Furniture
6/17/11	$93.54	Office Supplies
6/17/11	$905.08	Office Supplies
3/24/11	$2,781.82	Office Supplies
2/26/10	$228.41	Office Supplies
11/23/10	$196.85	Office Supplies
11/23/10	$124.56	Office Supplies
6/8/12	$716.84	Office Supplies
6/8/12	$1,474.33	Technology
8/4/12	$80.61	Office Supplies
5/30/11	$1,815.49	Furniture

Conceptual introduction

We add **numbers up (aggregate)** to see summaries

Sum of Sales
$14,915,600.82

Conceptual introduction

We add numbers up (aggregate) to see summaries,
subset by categories to make comparisons

Product Categories ↓▼	Sum of Sales
Technology	$5,984,248.18
Furniture	$5,178,590.54
Office Supplies	$3,752,762.10

Conceptual introduction

We add numbers up (aggregate) to see summaries,
subset by categories to make comparisons,
and **by dates to see trends**

Sum of Sales	2009	2010	2011	2012
Technology	$1,701,825.48	$1,397,142.14	$1,366,806.58	$1,518,473.98
Furniture	$1,472,671.72	$1,252,665.11	$1,269,661.09	$1,183,592.62
Office Supplies	$1,034,642.25	$899,873.56	$800,349.03	$1,017,897.26

Conceptual introduction

Sum of Sales	2009	2010	2011	2012
Technology	$1,701,825.48	$1,397,142.14	$1,366,806.58	$1,518,473.98
Furniture	$1,472,671.72	$1,252,665.11	$1,269,661.09	$1,183,592.62
Office Supplies	$1,034,642.25	$899,873.56	$800,349.03	$1,017,897.26

We visualize to make those comparisons and trends more clear