



Introduction to Matplotlib and Line Plots

Introduction

The aim of these labs is to introduce you to data visualization with Python as concrete and as consistent as possible. Speaking of consistency, because there is no *best* data visualization library available for Python - up to creating these labs - we have to introduce different libraries and show their benefits when we are discussing new visualization concepts. Doing so, we hope to make students well-rounded with visualization libraries and concepts so that they are able to judge and decide on the best visualization technique and tool for a given problem *and* audience.

Please make sure that you have completed the prerequisites for this course, namely ["Python for Data Science"](#) and ["Data Analysis with Python"](#), which are part of this specialization.

Note: The majority of the plots and visualizations will be generated using data stored in *pandas* dataframes. Therefore, in this lab, we provide a brief crash course on *pandas*. However, if you are interested in learning more about the *pandas* library, detailed description and explanation of how to use it and how to clean, munge, and process data stored in a *pandas* dataframe are provided in our course ["Data Analysis with Python"](#), which is also part of this specialization.

Table of Contents

Exploring Datasets with *pandas*

pandas is an essential data analysis toolkit for Python. From their [website](#):

1. [Exploring Datasets with *pandas*](#)
 - 1.1 [The Dataset: Immigration to Canada from 1980 to 2013](#)
 - 1.2 [pandas Basics](#)
 - 1.3 [pandas Intermediate: Indexing and Selection](#)

2. [Visualizing Data using Matplotlib](#)
The course heavily relies on *pandas* for data wrangling, analysis, and visualization. We encourage you to spend some time and familiarize yourself with the *pandas* API Reference: <http://pandas.pydata.org/pandas-docs/stable/10min.html>

