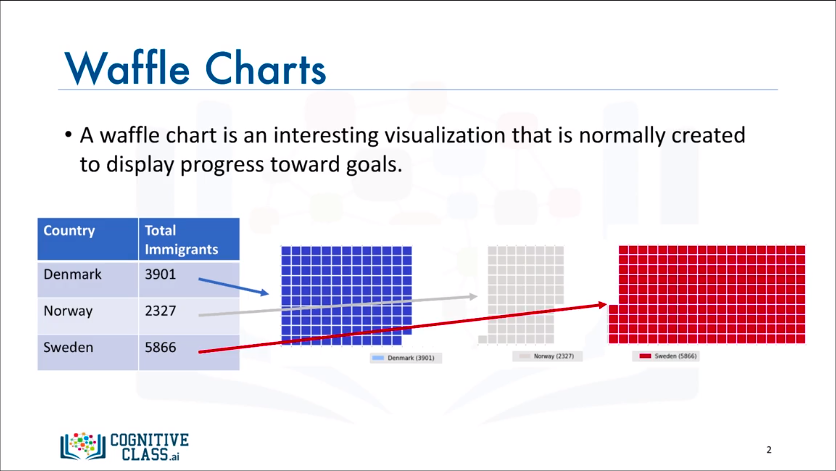
Advanced Visualization Tools

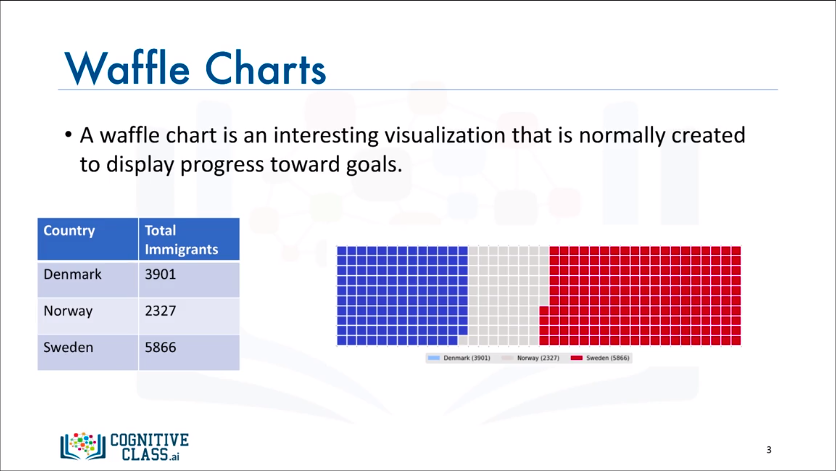
**Waffle Charts**



* A waffle chart is a great way to visualize data in relation to a whole or to highlight progress against a given threshold.

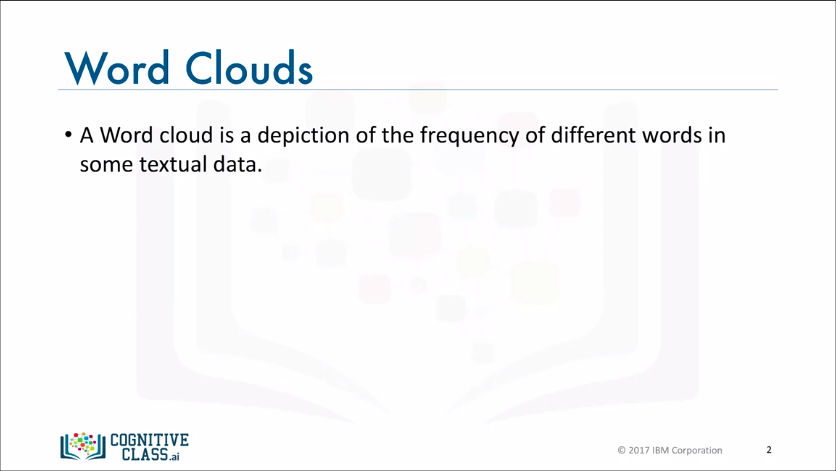


* For example, say immigration from Scandinavia to Canada is comprised only of immigration from Denmark, Norway, and Sweden, and we're interested in visualizing the contribution of each of these countries to the Scandinavian immigration to Canada.
* The main idea here is for a given waffle chart whose desired height and width are defined, the contribution of each country is transformed into a number of tiles that is proportional to the country's contribution to the total, so that more the contribution the more the tiles…

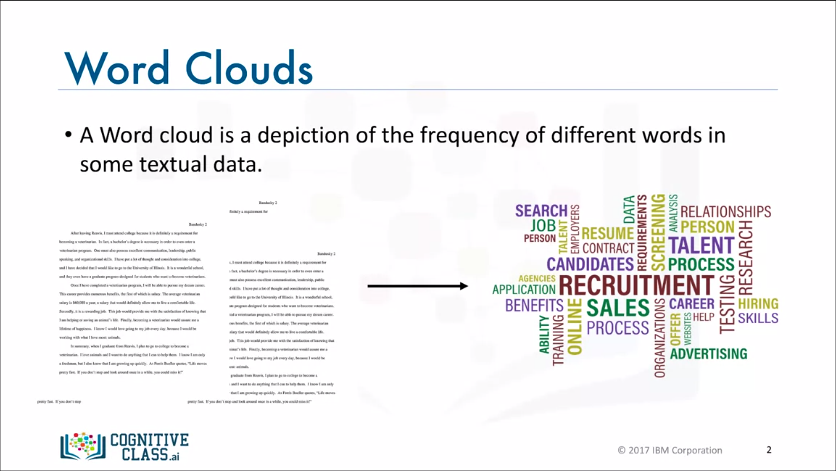


* …resulting in what resembles a waffle when combined. Hence the name waffle chart.
* Unfortunately, Matplotlib does not have a built-in function to create waffle charts. However, there’s a way to use Python function to create a waffle chart.

**Word Clouds**

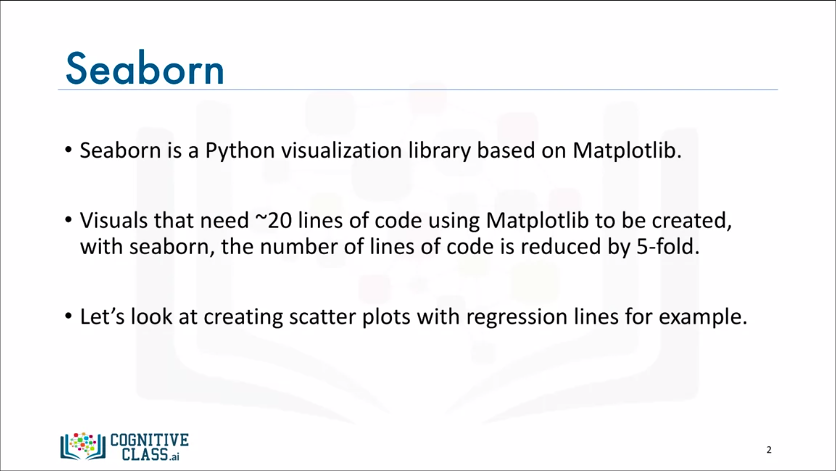


* A word cloud is simply a depiction of the importance of different words in the body of text. A word cloud works in a simple way; the more a specific word appears in a source of textual data the bigger and bolder it appears in the world cloud.

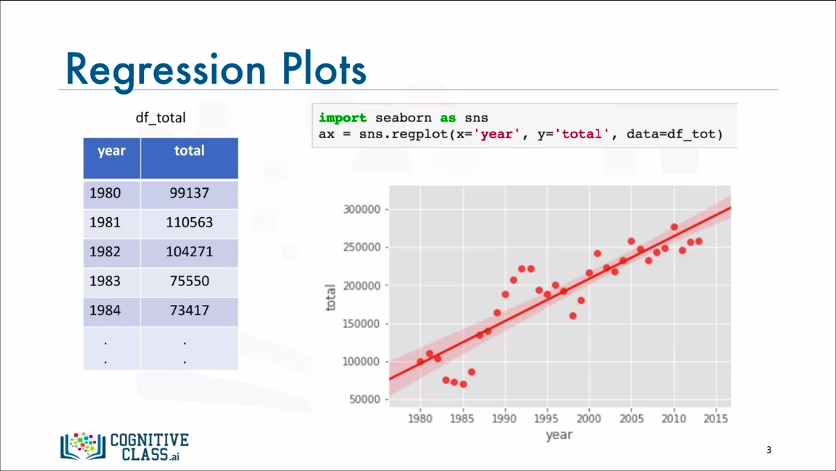


* For example, given some text data on recruitment (on the left), we generate a cloud of words like this (on the right). This cloud is telling us that words such as recruitment, talent, candidates, and so on, are the words that really stand out in these text documents.
* And assuming that we didn't know anything about the content of these documents, a word cloud can be very useful to assign a topic to some unknown textual data.
* Unfortunately, just like waffle charts, Matplotlib does not have a built-in function to generate word clouds. However, luckily a Python library for cloud word generation that was created by Andreas Mueller is publicly available. So, we will use Mueller's word cloud generator, and we will also create interesting word clouds superimposed on different background images.

**Seaborn and Regression Plots**



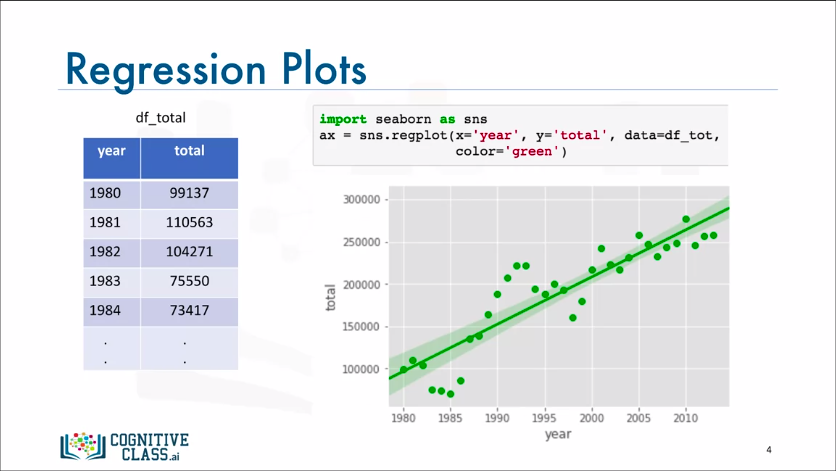
* Although Seaborn is another data visualization library, it is actually based on Matplotlib. It was built primarily to provide a high-level interface for drawing attractive statistical graphics, such as regression plots, box plots, and so on.
* Seaborn makes creating plots very efficient. Therefore, with Seaborn you can generate plots with code that is 5 times less than with Matplotlib. Let's see how we can use Seaborn to create a statistical graphic.



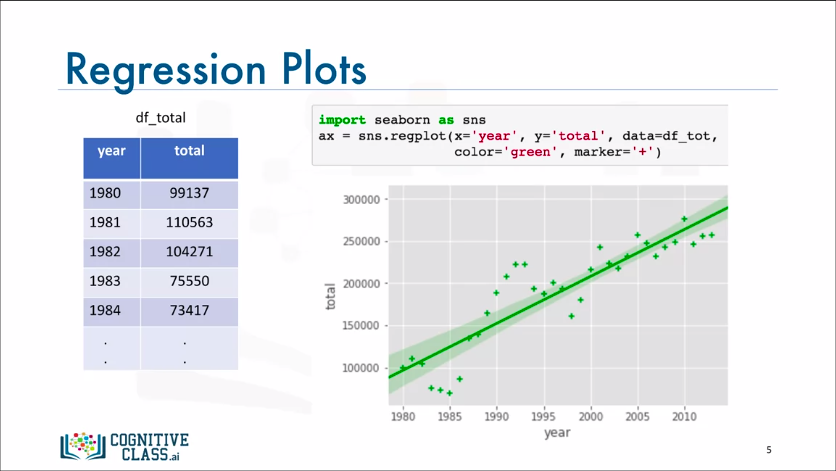
* Let's look into regression plots. Let's say we have a DataFrame called df\_total of total immigration to Canada from 1980 to 2013 with the year in one column and the corresponding total immigration in another column, and say we're interested in creating a scatter plot along with a regression line to highlight any trends in the data.
* With Seaborn, you can do all this with literally one line of code. The way to do this,

1. We first import Seaborn and let's import it as sns
2. Then, we call the Seaborn regplot function
3. We basically tell it to use the dataframe df\_total
4. And to plot the column year on the horizontal axis and the column total on the vertical axis

* And the output of this one line of code is a scatter plot with a regression line and not just that, but also 95% confidence interval. Isn't that really amazing?



* Seaborn's regplot function also accepts additional parameters for any personal customization. So, we can change the color for example using the color parameter. Let's go ahead and change the color to green.



* Also, we can change the marker shape as well using the marker parameter. Let's go ahead and change the shape of our markers to a + marker instead of the default circular marker.