

Pie Charts (4:14) | Pie Chart: x

https://courses.cognitiveclass.ai/courses/course-v1:CognitiveClass+DV0101EN+v1/courseware/bd64ccd56ad4ea1afe870e2...

Data Visualization with Python

Pie Charts

In this video, we will learn about another visualization tool: the pie chart,

0:03 / 4:14

Speed 1.0x

ENG US 20:39 10/02/2021

Pie Charts (4:14) | Pie Chart: x

https://courses.cognitiveclass.ai/courses/course-v1:CognitiveClass+DV0101EN+v1/courseware/bd64ccd56ad4ea1afe870e2...

DV0101EN - Pie Charts 4:14

Pie Chart

Party	Color
Liberal	Red
Conservative	Blue
New Democratic	Orange
Bloc Québécois	Teal
Green	Green

with Matplotlib? Before we go over the code to do that, let's do a quick recap

0:36 / 4:14

Speed 1.0x

ENG US 20:39 10/02/2021

Dataset - Recap

Type	Coverage	OdName	AREA	AreaName	REG	RegName	DEV	DevName	1980	...	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
0	Immigrants	Foreigners	Afghanistan	935	Asia	5501	Southern Asia	902	Developing regions	16	...	2978	3436	3009	2652	2111	1746	1758	2203	2635
1	Immigrants	Foreigners	Albania	908	Europe	925	Southern Europe	901	Developed regions	1	...	1450	1223	856	702	560	716	561	539	620
2	Immigrants	Foreigners	Algeria	903	Africa	912	Northern Africa	902	Developing regions	80	...	3616	3626	4807	3623	4005	5393	4752	4325	3774
3	Immigrants	Foreigners	American Samoa	909	Oceania	957	Polynesia	902	Developing regions	0	...	0	0	1	0	0	0	0	0	0
4	Immigrants	Foreigners	Andorra	908	Europe	925	Southern Europe	901	Developed regions	0	...	0	0	1	1	0	0	0	0	1

with Matplotlib? Before we go over the code to do that, let's do a quick recap

Dataset - Processed

	Continent	Region	DevName	1980	1981	1982	1983	1984	1985	1986	...	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
Country																					
Afghanistan	Asia	Southern Asia	Developing regions	16	39	39	47	71	340	496	...	3436	3009	2652	2111	1746	1758	2203	2635	2004	58639
Albania	Europe	Southern Europe	Developed regions	1	0	0	0	0	0	1	...	1223	856	702	560	716	561	539	620	603	15699
Algeria	Africa	Northern Africa	Developing regions	80	67	71	69	63	44	69	...	3626	4807	3623	4005	5393	4752	4325	3774	4331	69439
American Samoa	Oceania	Polynesia	Developing regions	0	1	0	0	0	0	0	...	0	1	0	0	0	0	0	0	0	6
Andorra	Europe	Southern Europe	Developed regions	0	0	0	0	0	0	2	...	0	1	1	0	0	0	0	1	1	15

df_canada

15,699 and so on. And let's name our dataframe df_canada. So now

Pie Charts (4:14) | Pie Chart

https://courses.cognitiveclass.ai/courses/course-v1:CognitiveClass+DV0101EN+v1/courseware/bd64ccd56ad4ea1afe870e2...

Pie Chart

```
df_continents = df_canada.groupby('Continent', axis = 0).sum()
```

	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	...	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total	
Continent																						
Africa	3951	4363	3819	2671	2639	2650	3782	7494	7552	9894	...	27523	29188	28284	29890	34534	40892	35441	38283	38543	618948	
Asia	31025	34314	30214	24696	27274	23850	28739	43203	47454	60256	...	158253	149054	133459	139894	141434	163845	146894	152218	155075	3317794	
Europe	38780	44802	42720	24638	22287	20844	24370	46698	54726	60893	...	35955	33053	33495	34692	35078	33425	26778	29177	28691	1410947	
Latin America and the Caribbean	13081	15215	16769	15427	13678	15171	21179	28471	21904	25060	...	24747	24676	26011	26547	26867	28818	27856	27173	24950	765148	
Northern America	9378	10030	9074	7100	6661	6543	7074	7705	6469	6790	...	8394	9613	9463	10190	8995	8142	7677	7892	8503	241142	
Oceania	1942	1839	1675	1018	878	920	904	1200	1181	1539	...	1585	1473	1693	1834	1860	1834	1548	1679	1775	55174	

df_continents. The resulting dataframe has six rows, each

2:23 / 4:14

Speed 1.0x

ENG US 20:40 10/02/2021

Pie Charts (4:14) | Pie Chart

https://courses.cognitiveclass.ai/courses/course-v1:CognitiveClass+DV0101EN+v1/courseware/bd64ccd56ad4ea1afe870e2...

DV0101EN - Pie Charts 4:14

Pie Chart

```
import matplotlib as mpl
import matplotlib.pyplot as plt

df_continents['Total'].plot(kind='pie')
plt.title('Immigration to Canada by Continent (1980-2013)')
plt.show()
```

Immigration to Canada by Continent (1980-2013)

proportion of immigration to Canada from 1980 to 2013. In the lab session, we will

youtube

3:18 / 4:14

Speed 1.0x

ENG US 20:41 10/02/2021

Pie Charts (4:14) | Pie Chart | The Data Science Course 2021: | +

<https://courses.cognitiveclass.ai/courses/course-v1:CognitiveClass+DV0101EN+v1/courseware/bd64ccdf56ad4ea1afe870e2...>

Pie Charts

Immigration to Canada by Continent (1980-2013)

Immigration to Canada by Continent (1980-2013)

the use of pie charts under any circumstances. Most argue that pie charts

Box Plots (3:42) | Box Plots (3:42) | The Data Science Course 2021: | +

<https://courses.cognitiveclass.ai/courses/course-v1:CognitiveClass+DV0101EN+v1/courseware/bd64ccdf56ad4ea1afe870e2...>

Data Visualization with Python

Box Plots

In this video, we will learn about another visualization tool: the box plot,

0:01 / 3:42

Speed 1.0x

ENG US 20:43 10/02/2021

Box Plots (3/42) | Box Plots | The Data Science Course 2021: | SWAYAM

https://courses.cognitiveclass.ai/courses/course-v1:CognitiveClass+DV0101EN+v1/courseware/bd64ccdf56ad4ea1afe870e2...

DV0101EN - Box Plots 3:42

Box Plots

Outliers

Maximum

Third Quartile

Median

First Quartile

Minimum

Inter Quartile Range (IQR)

let's see how we can create a box plot with Matplotlib. Before we go over the

youtube

1:01 / 3:42

Speed 1.0x

ENG US 20:44 10/02/2021

Box Plots (3/42) | Box Plots | The Data Science Course 2021: | SWAYAM

https://courses.cognitiveclass.ai/courses/course-v1:CognitiveClass+DV0101EN+v1/courseware/bd64ccdf56ad4ea1afe870e2...

DV0101EN - Box Plots 3:42

Dataset - Recap

Type	Coverage	OdName	AREA	AreaName	REG	RegName	DEV	DevName	1980	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013		
0	Immigrants	Foreigners	Afghanistan	935	Asia	5501	Southern Asia	902	Developing regions	16	...	2978	3436	3009	2652	2111	1746	1758	2203	2635	2004
1	Immigrants	Foreigners	Albania	908	Europe	925	Southern Europe	901	Developed regions	1	...	1450	1223	856	702	560	716	561	539	620	603
2	Immigrants	Foreigners	Algeria	903	Africa	912	Northern Africa	902	Developing regions	80	...	3616	3626	4807	3623	4005	5393	4752	4325	3774	4331
3	Immigrants	Foreigners	American Samoa	909	Oceania	957	Polynesia	902	Developing regions	0	...	0	0	1	0	0	0	0	0	0	0
4	Immigrants	Foreigners	Andorra	908	Europe	925	Southern Europe	901	Developed regions	0	...	0	0	1	1	0	0	0	0	1	1

code to do that, let's do a quick recap of our dataset. Recall that each row

youtube

1:06 / 3:42

Speed 1.0x

ENG US 20:44 10/02/2021

Box Plots (3/42) | Box Plots | The Data Science Course 2021: | Data Science for Engineers - Co... |

https://courses.cognitiveclass.ai/courses/course-v1:CognitiveClass+DV0101EN+v1/courseware/bd64ccd56ad4ea1afe870e2...

DV0101EN - Box Plots 3:42

Dataset - Processed

Country	Continent	Region	DevName	1980	1981	1982	1983	1984	1985	1986	...	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
Afghanistan	Asia	Southern Asia	Developing regions	16	39	39	47	71	340	496	...	3436	3009	2652	2111	1746	1758	2203	2635	2004	58639
Albania	Europe	Southern Europe	Developed regions	1	0	0	0	0	0	1	...	1223	856	702	560	716	561	539	620	603	15699
Algeria	Africa	Northern Africa	Developing regions	80	67	71	69	63	44	69	...	3626	4807	3623	4005	5393	4752	4325	3774	4331	69439
American Samoa	Oceania	Polynesia	Developing regions	0	1	0	0	0	0	0	...	0	1	0	0	0	0	0	0	0	6
Andorra	Europe	Southern Europe	Developed regions	0	0	0	0	0	0	2	...	0	1	1	0	0	0	0	1	1	15

each row. This should make retrieving rows pertaining to specific countries a

COGNITIVE CLASS

youtube

1:39 / 3:42

Speed 1.0x

ENG US 20:44 10/02/2021

Box Plots (3/42) | Box Plots | The Data Science Course 2021: |

https://courses.cognitiveclass.ai/courses/course-v1:CognitiveClass+DV0101EN+v1/courseware/bd64ccd56ad4ea1afe870e2...

Box Plots

```
import matplotlib as mpl
import matplotlib.pyplot as plt

df_japan = df_canada.loc(['Japan'], years).transpose()
df_japan.plot(kind='box')

plt.title('Box plot of Japanese Immigrants from 1980-2013')
plt.ylabel('Number of Immigrants')
plt.show()
```

Box plot of Japanese Immigrants from 1980-2013

Number of Immigrants

Japan

Canada from 1980 to 2013. In the lab session, we explore box plots in more

COGNITIVE CLASS

3:23 / 3:42

Speed 1.0x

ENG US 20:46 10/02/2021

Scatter Plots (4:17) | Scatter Plot: x The Data Science Course 2021: x +

https://courses.cognitiveclass.ai/courses/course-v1:CognitiveClass+DV0101EN+v1/courseware/bd64ccdf56ad4ea1afe870e2...

DV0101EN - Scatter Plots 4:17

Watch later Share

Data Visualization with Python

Scatter Plots

In this video, we will learn about an additional visualization tool: the

CLAS AI

YouTube

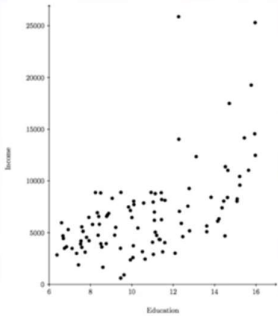
0:01 / 4:17

Speed 1.0x HD ENG US 20:46 10/02/2021

Scatter Plots (4:17) | Scatter: x The Data Science Course 2021: x +

https://courses.cognitiveclass.ai/courses/course-v1:CognitiveClass+DV0101EN+v1/courseware/bd64ccdf56ad4ea1afe870e2...

Scatter Plots



A scatter plot with 'Education' on the x-axis (ranging from 8 to 16) and 'Income' on the y-axis (ranging from 0 to 25000). The plot shows a positive correlation between education and income, with data points clustered at lower education levels and more spread out at higher education levels.

COGNITIVE

higher income than an individual with fewer years of education. So how can we

0:43 / 4:17

Speed 1.0x HD ENG US 20:47 10/02/2021

Dataset - Recap

Type	Coverage	OdName	AREA	AreaName	REG	RegName	DEV	DevName	1980	...	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
0	Immigrants	Foreigners	Afghanistan	935	Asia	5501	Southern Asia	902	Developing regions	16	...	2978	3436	3009	2652	2111	1746	1758	2203	2635
1	Immigrants	Foreigners	Albania	908	Europe	925	Southern Europe	901	Developed regions	1	...	1450	1223	856	702	560	716	561	539	620
2	Immigrants	Foreigners	Algeria	903	Africa	912	Northern Africa	902	Developing regions	80	...	3616	3626	4807	3623	4005	5393	4752	4325	3774
3	Immigrants	Foreigners	American Samoa	909	Oceania	957	Polynesia	902	Developing regions	0	...	0	0	1	0	0	0	0	0	0
4	Immigrants	Foreigners	Andorra	908	Europe	925	Southern Europe	901	Developed regions	0	...	0	0	1	1	0	0	0	0	1

create a scatterplot with Matplotlib? Before we go over the code to do that,

Dataset - Processed

	Continent	Region	DevName	1980	1981	1982	1983	1984	1985	1986	...	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
Country																					
Afghanistan	Asia	Southern Asia	Developing regions	16	39	39	47	71	340	496	...	3436	3009	2652	2111	1746	1758	2203	2635	2004	58639
Albania	Europe	Southern Europe	Developed regions	1	0	0	0	0	0	1	...	1223	856	702	560	716	561	539	620	603	15699
Algeria	Africa	Northern Africa	Developing regions	80	67	71	69	63	44	69	...	3626	4807	3623	4005	5393	4752	4325	3774	4331	69439
American Samoa	Oceania	Polynesia	Developing regions	0	1	0	0	0	0	0	...	0	1	0	0	0	0	0	0	0	6
Andorra	Europe	Southern Europe	Developed regions	0	0	0	0	0	0	2	...	0	1	1	0	0	0	0	1	1	15

lot easier. Also let's add an extra column which represents the cumulative

Scatter Plots

```
import matplotlib as mpl
import matplotlib.pyplot as plt

df_total.plot(
    kind='scatter',
    x='year',
    y='total',
)

plt.title('Total Immigrant population to Canada from 1980 - 2013')
plt.xlabel('Year')
plt.ylabel('Number of Immigrants')
plt.show()
```

df_total	
year	total
1980	99137
1981	110563
1982	104271
1983	75550
1984	73417
.	.
.	.

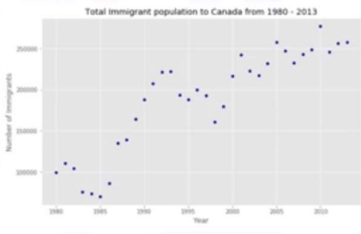
give it a title and we label its axes. Finally, we call the show function to

Scatter Plots

```
import matplotlib as mpl
import matplotlib.pyplot as plt

df_total.plot(
    kind='scatter',
    x='year',
    y='total',
)

plt.title('Total Immigrant population to Canada from 1980 - 2013')
plt.xlabel('Year')
plt.ylabel('Number of Immigrants')
plt.show()
```



display the figure. And there you have it. A scatter plot that shows total