

Introduction to R Markdown

REPORTING WITH R MARKDOWN



Amy Peterson
Head of Core Curriculum at DataCamp

R Markdown file

```
1 ---  
2 title: "Investment Report"  
3 output: html_document  
4 ---
```

YAML header

```
5  
6 ```{r data, include = FALSE}  
7 library(readr)  
8  
9 investment_annual_summary <- read_csv("https://  
assets.datacamp.com/production/repositories/5756/  
datasets/d0251f26117bbcf0ea96ac276555b9003f4f7372/  
investment_annual_summary.csv")  
10  
11
```

```
12 ### Investment Annual Summary
```

```
13  
14 The `investment_annual_summary` dataset provides a  
summary of the dollars in millions provided to each  
region for each fiscal year, from 2012 to 2018.
```

text

```
15  
16 ```{r}  
17 investment_annual_summary  
18 ...
```

code chunk

Formatting text and adding img

```
1 ---  
2 title: "Title"  
3 output: html_document  
4 ---  
5  
6 # Header 1  
7 ## Header 2  
8 ### Header 3  
9  
10 `code`  
11  
12 **bold** or __bold__  
13  
14 *italics* or _italics_  
15  
16 ~~strikethrough~~  
17  
18 ``{r}  
19 # Comment  
20 ``  
21  
22 [World Bank IFC Data](https://www.kaggle.com/theworldbank/  
ifc-investment-summary-and-services-projects)  
23  
24   
25
```

Title

Header 1

Header 2

Header 3

code

bold or **bold**

italics or *italics*

~~strikethrough~~

Comment

World Bank IFC Data



Output types

```
1 ---  
2 title: "Investment Report"  
3 author: "Amy"  
4 output: html_document  
5 ---  
6
```

Investment Report
Amy

```
1 ---  
2 title: "Investment Report"  
3 author: "Amy"  
4 output: pdf_document  
5 ---  
6
```



Formatting the date

```
1 ---  
2 title: "Investment Report"  
3 date: ``r format(Sys.time(), '%d %B %Y')``  
4 output: html_document  
5 ---  
6
```

Investment Report
08 April 2020

Text

%A or %a	Weekday
%B or %b	Month

Numeric

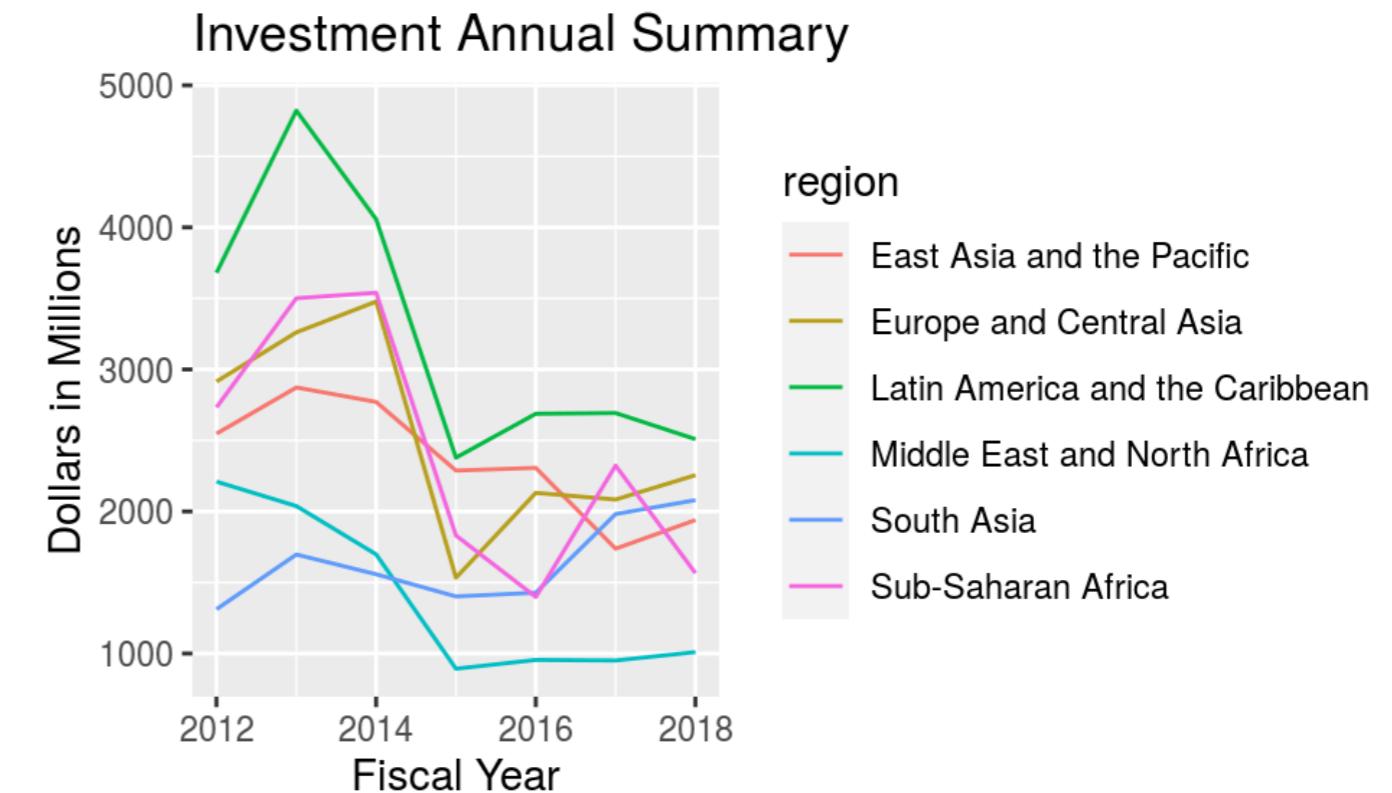
%d	Decimal date
%m	Decimal month
%Y or %y	Year

Figure dimensions

```
26  ```{r investment-annual-summary, fig.width = 5, fig.height = 3}
27  ggplot(investment_annual_summary, aes(x = fiscal_year, y =
28    dollars_in_millions, color = region)) +
29    geom_line() +
30    labs(
31      title = "Investment Annual Summary",
32      x = "Fiscal Year",
33      y = "Dollars in Millions"
34  ```

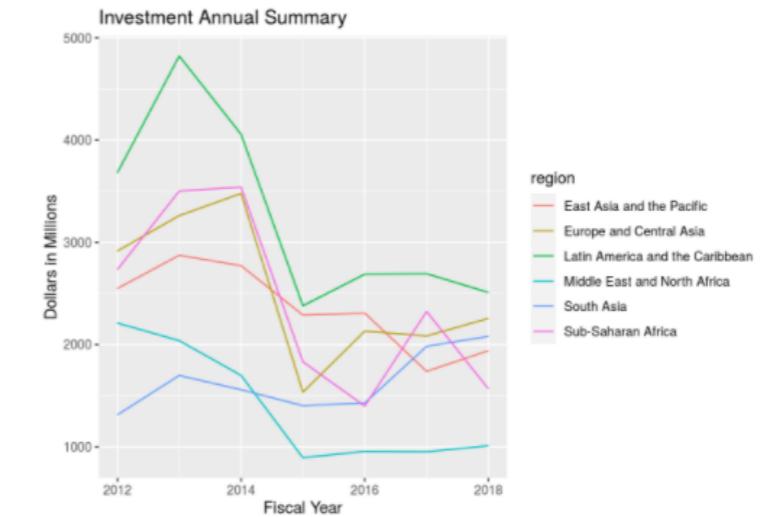
26  ```{r investment-annual-summary, fig.dim = c(5, 3)}
27  ggplot(investment_annual_summary, aes(x = fiscal_year, y =
28    dollars_in_millions, color = region)) +
29    geom_line() +
30    labs(
31      title = "Investment Annual Summary",
32      x = "Fiscal Year",
33      y = "Dollars in Millions"
34  ```


```



Output dimensions

```
26  ```{r investment-annual-summary, out.width = '50%'}
27  ggplot(investment_annual_summary, aes(x = fiscal_year, y =
28    dollars_in_millions, color = region)) +
29    geom_line() +
30    labs(
31      title = "Investment Annual Summary",
32      x = "Fiscal Year",
33      y = "Dollars in Millions"
34    )
35  ````
```



- `out.width`
- `out.height`

Figure alignment

```
26 ````{r investment-annual-summary, out.width = '80%', fig.align = 'left'}
```

```
27 ggplot(investment_annual_summary, aes(x = fiscal_year, y =
```

```
28 dollars_in_millions, color = region)) +
```

```
29   geom_line() +
```

```
30   labs(
```

```
31     title = "Investment Annual Summary",
```

```
32     x = "Fiscal Year",
```

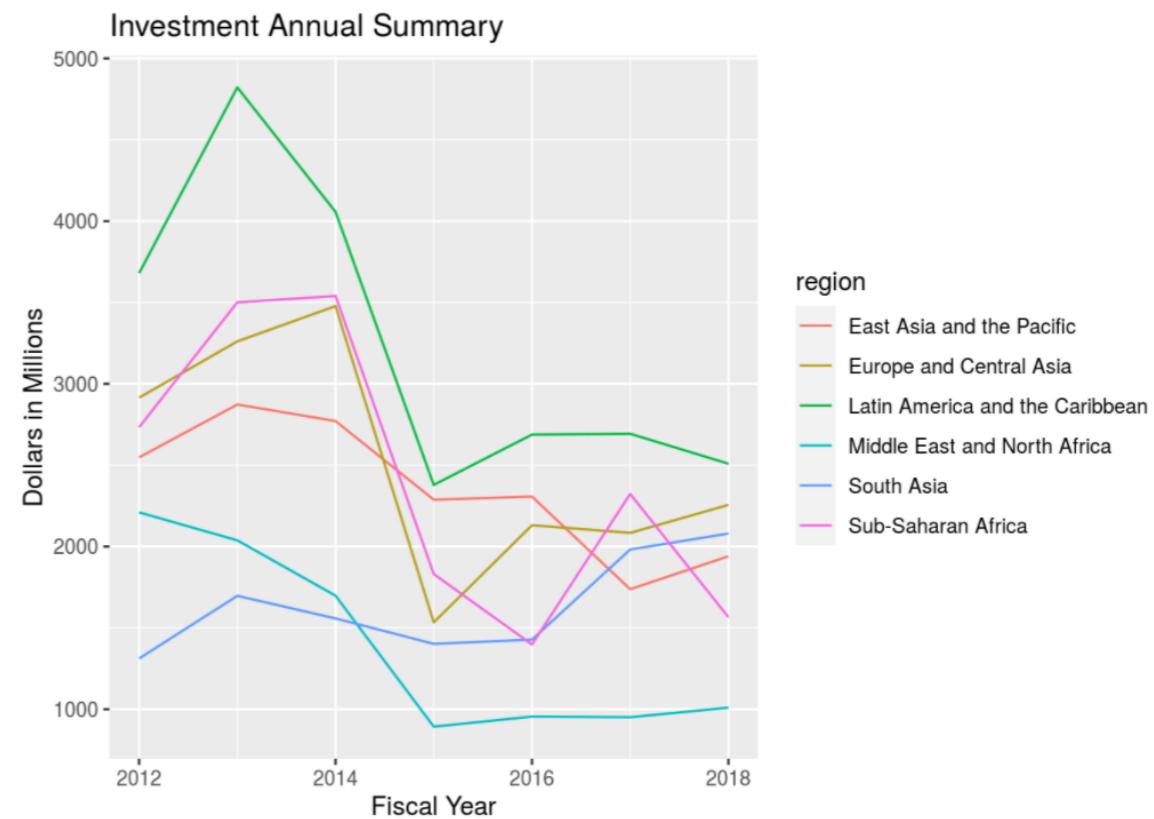
```
33     y = "Dollars in Millions"
```

```
34   )
```

```
35 ````
```

- `fig.align`
 - `'left'`
 - `'right'`
 - `'center'`

```
ggplot(investment_annual_summary, aes(x = fiscal_year, y = dollars_in_millions, color = region)) +  
  geom_line() +  
  labs(  
    title = "Investment Annual Summary",  
    x = "Fiscal Year",  
    y = "Dollars in Millions"  
)
```



Adding captions

```
26  ````{r investment-annual-summary, out.width = '85%', fig.cap = 'Figure 1.1 The  
27  Investment Annual Summary for each region for the 2012 to 2018 fiscal years.'}  
28  ggplot(investment_annual_summary, aes(x = fiscal_year, y = dollars_in_millions,  
29  color = region)) +  
30  geom_line() +  
31  labs(  
32  title = "Investment Annual Summary",  
33  x = "Fiscal Year",  
34  y = "Dollars in Millions")  
35  ````
```

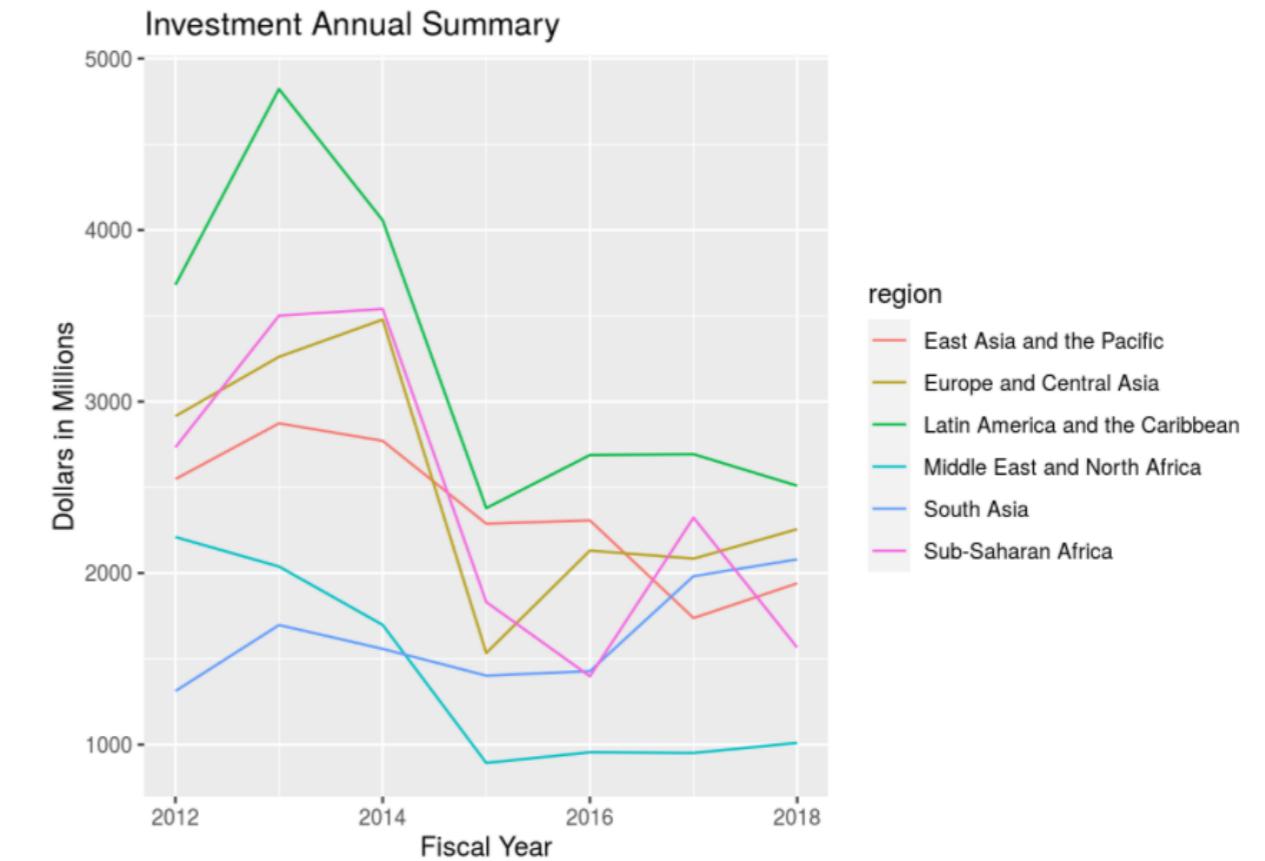


Figure 1.1 The Investment Annual Summary for each region for the 2012 to 2018 fiscal years.

Local vs. global options

```
26  ````{r investment-annual-summary, fig.align = 'center'}
27  ggplot(investment_annual_summary, aes(x = fiscal_year, y =
28    dollars_in_millions, color = region)) +
29    geom_line() +
30    labs(
31      title = "Investment Annual Summary",
32      x = "Fiscal Year",
33      y = "Dollars in Millions"
34    )
35  ```
36
37
38  ### Investment Projects in Indonesia
39
40  The `investment_services_projects` dataset provides information about each
41  investment project from 2012 to 2018. Information listed includes the project
42  name, company name, sector, project status, and investment amounts.
43
44  ````{r indonesia-investment-projects, fig.align = 'center'}
45  indonesia_investment_projects <- investment_services_projects %>%
46    filter(country == "Indonesia")
47
48  ggplot(indonesia_investment_projects, aes(x = date_disclosed, y =
49    total_investment, color = status)) +
50    geom_point() +
51    labs(
52      title = "Investment Services Projects in Indonesia",
53      x = "Date Disclosed",
54      y = "Total IFC Investment in Dollars in Millions"
55    )
56  ````
```

```
7  ````{r setup, include = FALSE}
8  knitr::opts_chunk$set(fig.align = 'center', echo = TRUE)
9  ````
```

Bulleted lists

```
23 ### Investment Annual Summary
24 The `investment_annual_summary` dataset provides a summary of the
25 dollars in millions provided to each of the following regions for
26 each fiscal year, from 2012 to 2018:
27
28 - Region
29   - East Asia and the Pacific
30   - Europe and Central Asia
31   - Latin America and the Caribbean
32   - Middle East and North Africa
33   - South Asia
34   - Sub-Saharan Africa
```

Numbered lists

```
23 ### Investment Annual Summary
24 The `investment_annual_summary` dataset provides a summary of the
25 dollars in millions provided to each of the following regions for
26 each fiscal year, from 2012 to 2018:
27
28 1. East Asia and the Pacific
29 2. Europe and Central Asia
30 3. Latin America and the Caribbean
31 4. Middle East and North Africa
32 5. South Asia
33 6. Sub-Saharan Africa
```

Investment Annual Summary

The `investment_annual_summary` dataset provides a summary of the dollars in millions provided to each region for each fiscal year, from 2012 to 2018.

- Region
 - East Asia and the Pacific
 - Europe and Central Asia
 - Latin America and the Caribbean
 - Middle East and North Africa
 - South Asia
 - Sub-Saharan Africa

Investment Annual Summary

The `investment_annual_summary` dataset provides a summary of the dollars in millions provided to each region for each fiscal year, from 2012 to 2018.

Region

1. East Asia and the Pacific
2. Europe and Central Asia
3. Latin America and the Caribbean
4. Middle East and North Africa
5. South Asia
6. Sub-Saharan Africa

Adding table caption

```
44  ````{r tables}
45  kable(indonesia_investment_projects_2012_summary,
46  |   col.names = c("Project Name", "Status", "Total Investment"),
47  |   align = "ccc",
48  |   caption = "Table 1.1 The total investment summary for each
49  |   project in Indonesia for the 2012 fiscal year.")
50  ````
```

```
kable(indonesia_investment_projects_2012_summary, col.names = c("Project Name",
"Status", "Total Investment"), align = "ccc", caption = "Table 1.1 The total inv
estment summary for each project in Indonesia in the 2012 fiscal year.")
```

Table 1.1 The total investment summary for each project in Indonesia in the 2012 fiscal year.

Project Name	Status	Total Investment
FHP Indonesia I	Active	25
LMS Toll Project	Hold	NA
CIMB Niaga Sr.	Completed	75
BTPN Loan II	Active	250
Medco Power 2011	Completed	25
Wintermar Group	Active	60

Code option summary

	Code is run	Code appears in report	Results appear in report
include = FALSE	Yes	No	No
echo = FALSE	Yes	No	Yes
eval = FALSE	No	Yes	No

```
67  ```{r indonesia-investment-projects-2012, collapse = TRUE}
68  indonesia_investment_projects_2018 <- investment_services_projects %>%
69  | filter(country == "Indonesia",
70  | | | | date_disclosed >= "2011-07-01",
71  | | | | date_disclosed <= "2012-06-30")
72
73 ggplot(indonesia_investment_projects_2012, aes(x = date_disclosed, y =
74   total_investment, color = status)) +
75   geom_point() +
76   labs(
77     title = "Investment Services Projects in Indonesia in 2012",
78     x = "Date Disclosed",
79     y = "Total IFC Investment in Dollars in Millions"
80   )```

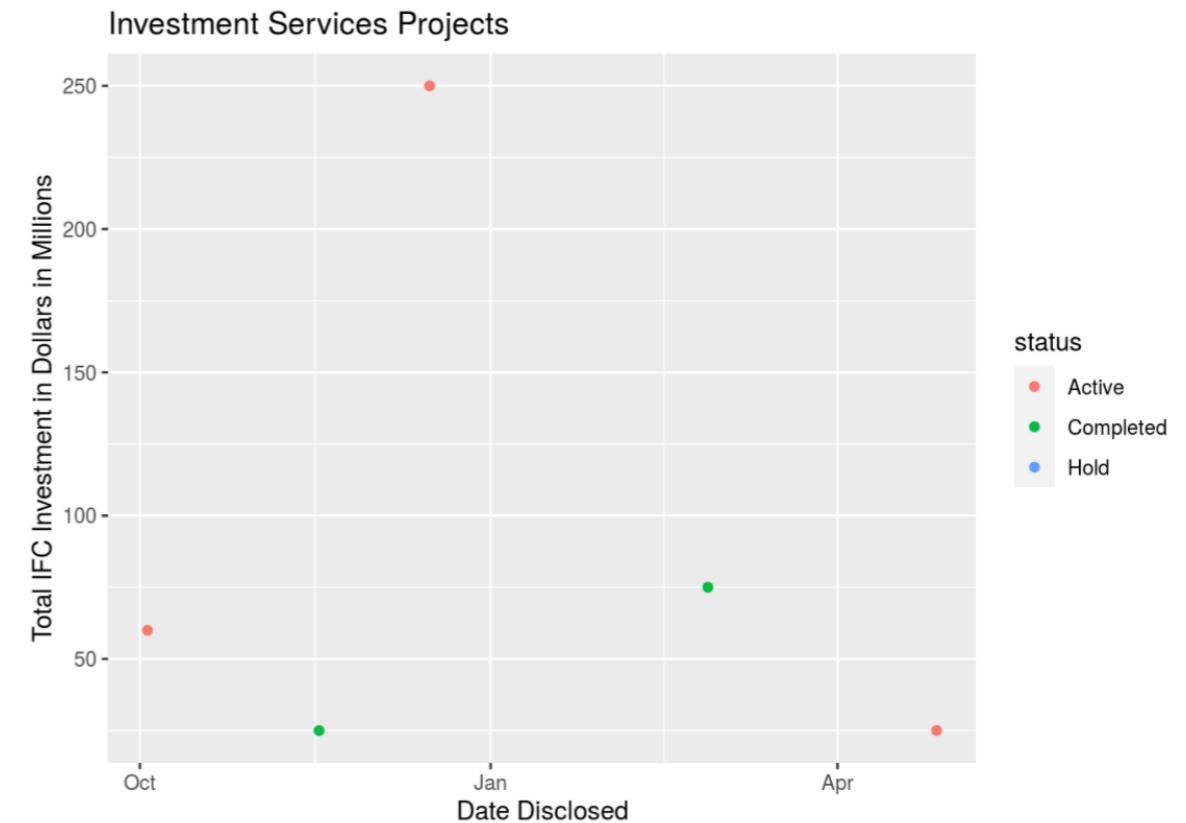
```

```
ggplot(indonesia_investment_projects_2012, aes(x = date_disclosed, y = total_
investment, color = status)) +
  geom_point() +
  labs(
    title = "Investment Services Projects",
    x = "Date Disclosed",
    y = "Total IFC Investment in Dollars in Millions"
  )
## Warning: Removed 1 rows containing missing values (geom_point).
```

Warnings

```
53 ````{r indonesia-investment-projects-2012, warning = FALSE}
54 indonesia_investment_projects_2012 <- investment_services_projects %>%
55   filter(country == "Brazil",
56         date_disclosed >= "2011-07-01",
57         date_disclosed <= "2012-06-30")
58
59 ggplot(indonesia_investment_projects_2012, aes(x = date_disclosed, y =
60   total_investment, color = status)) +
61   geom_point() +
62   labs(
63     title = "Investment Services Projects in Indonesia in 2012",
64     x = "Date Disclosed",
65     y = "Total IFC Investment in Dollars in Millions"
66   )
````
```

```
ggplot(indonesia_investment_projects_2012, aes(x = date_disclosed, y = total_
investment, color = status)) +
 geom_point() +
 labs(
 title = "Investment Services Projects",
 x = "Date Disclosed",
 y = "Total IFC Investment in Dollars in Millions"
)
```



# Messages

```
11 ```{r data, message = FALSE}
12 library(readr)
13 library(dplyr)
14 library(ggplot2)
15
16 investment_annual_summary <- read_csv("https://assets.datacamp.com/
17 production/repositories/5756/datasets/
18 d0251f26117bbcf0ea96ac276555b9003f4f7372/investment_annual_summary.csv")
19
20 investment_services_projects <- read_csv("https://assets.datacamp.com/
21 production/repositories/5756/datasets/
22 bcb2e39ecbe521f4b414a21e35f7b8b5c50aec64/
23 investment_services_projects.csv")
24
25 ````
```

**Message = FALSE**

Previene que salta del código cualquier anuncio de la formula, sea una advertencia o no.

HTML Viewer ↗

# Investment Report

28 April 2020

```
library(readr)
library(dplyr)
library(ggplot2)
library(knitr)

investment_annual_summary <- read_csv("https://assets.datacamp.com/product
ion/repositories/5756/datasets/d0251f26117bbcf0ea96ac276555b9003f4f7372/in
vestment_annual_summary.csv")
investment_region_summary <- read_csv("https://assets.datacamp.com/product
ion/repositories/5756/datasets/52f5414f6504e0503e86eb1043afa9b3d157fab2/in
vestment_region_summary.csv")
investment_services_projects <- read_csv("https://assets.datacamp.com/prod
uction/repositories/5756/datasets/bcb2e39ecbe521f4b414a21e35f7b8b5c50aec6
4/investment_services_projects.csv")
```

# Errors

```
40 ```{r indonesia-investment-projects, error = TRUE}
41 ggplot(indonesia_investment_projects, aes(x = date_disclosed, y =
42 total_investment, color = status)) +
43 geom_point() +
44 labs(
45 title = "Investment Services Projects in Indonesia",
46 x = "Date Disclosed",
47 y = "Total IFC Investment in Dollars in Millions"
48)
49 ```

50 ggplot(indonesia_investment_projects, aes(x = date_disclosed, y = total_investment,
51 color = status)) +
52 geom_point() +
53 labs(
54 title = "Investment Services Projects in Indonesia",
55 x = "Date Disclosed",
56 y = "Total IFC Investment in Dollars in Millions"
57)
58
59 ## Error in ggplot(indonesia_investment_projects, aes(x = date_disclosed, : object
60 'indonesia_investment_projects' not found
```

`error = TRUE`

Permite que documento se genere aunque se encuentre errores y nos permite ver donde estan esos errores.

# Table of contents

```
1 ---
2 title: "Investment Report"
3 output:
4 | html_document:
5 | | toc: true
6 date: "`r format(Sys.time(), '%d %B %Y')`"
7 ---
```

The file won't knit unless the correct indentation is used in the YAML header.

## TOC depth

```
1 ---
2 title: "Investment Report"
3 output:
4 | html_document:
5 | | toc: true
6 | | toc_depth: 2
7 date: "`r format(Sys.time(), '%d %B %Y')`"
8 ---
```

toc\_depth: Limita los titulos a los que tienen  $\leq$  numero de #.

# Investment Report

08 May 2020

- Datasets
  - Investment Annual Summary
  - Investment Projects from the 2012 to 2018 Fiscal Years
  - Investment Projects in 2018

# Investment Report

08 May 2020

- Datasets

## Datasets

### Investment Annual Summary

# Number sections

```
1 ---
2 title: "Investment Report"
3 output:
4 html_document:
5 toc: true
6 toc_depth: 2
7 number_sections: true
8 date: "`r format(Sys.time(), '%d %B %Y')`"
9 ---
10 ````{r setup, include = FALSE}
11 knitr::opts_chunk$set(fig.align = 'center', echo = TRUE)
12 ````
13
14 ````{r data, include = FALSE}
15 library(readr)
16 library(dplyr)
17 library(ggplot2)
18
19 investment_annual_summary <- read_csv("https://assets.datacamp.com/
20 production/repositories/5756/datasets/
21 d0251f26117bbcf0ea96ac276555b9003f4f7372/investment_annual_summary.csv")
22 investment_services_projects <- read_csv("https://assets.datacamp.com/
23 production/repositories/5756/datasets/
24 bcb2e39ecbe521f4b414a21e35f7b8b5c50aec64/investment_services_projects.csv")
25
26
27 ## Datasets
28
29 ### Investment Annual Summary
```

# Investment Report

08 May 2020

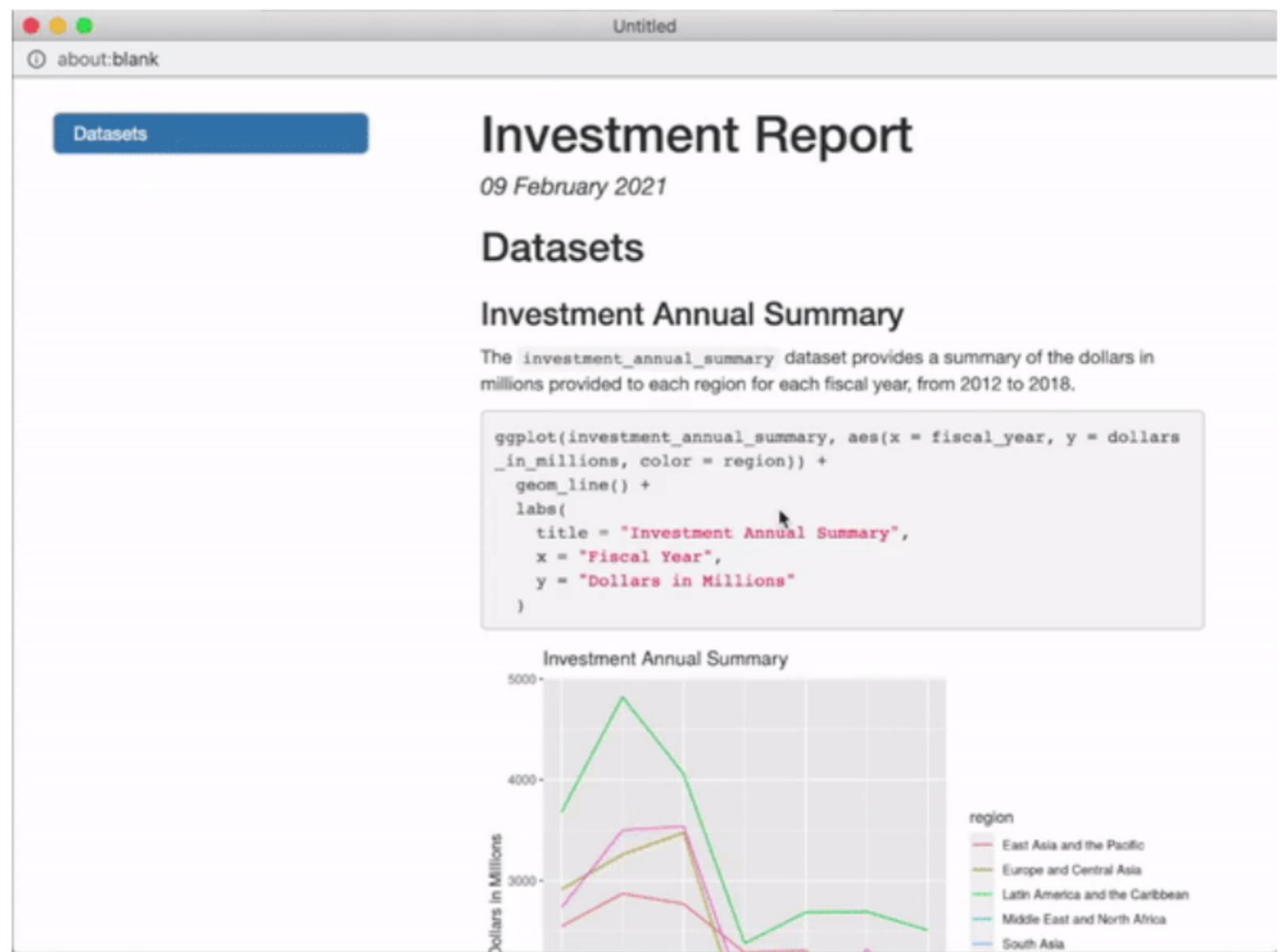
- 0.1 Datasets
  - 0.1.1 Investment Annual Summary
  - 0.1.2 Investment Projects from the 2012 to 2018 Fiscal Years
  - 0.1.3 Investment Projects in 2018

No comienza en 1 porque no tenemos el primer # solo tenemos de ## y ###.

# TOC float

```
1 ---
2 title: "Investment Report"
3 output:
4 html_document:
5 toc: true
6 toc_float: true
7 toc_depth: 3
8 date: "`r format(Sys.time(), '%d %B %Y')`"
9 ---
```

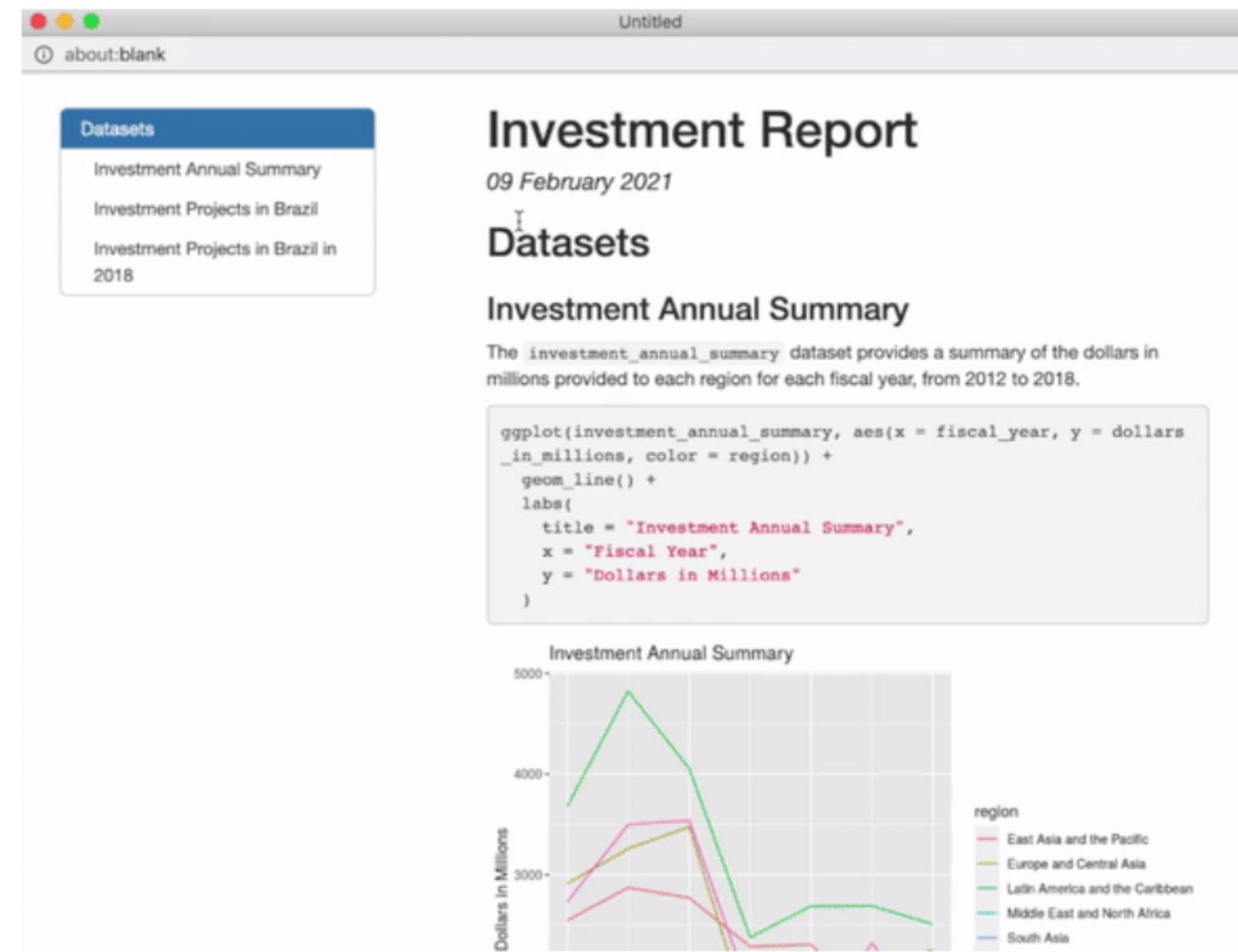
El menú va apareciendo segun la persona baja en el documento.



# TOC float: collapsed

```
1 ---
2 title: "Investment Report"
3 output:
4 html_document:
5 toc: true
6 toc_float:
7 collapsed: false
8 toc_depth: 3
9 date: "`r format(Sys.time(), '%d %B %Y')`"
10 ---
```

El definir el collapse como false ahora el menu se mantiene fijo a la izquierda del documento.



# TOC float: smooth scroll

```
1 ---
2 title: "Investment Report"
3 output:
4 html_document:
5 toc: true
6 toc_float:
7 collapsed: false
8 smooth_scroll: false
9 toc_depth: 3
10 date: "`r format(Sys.time(), '%d %B %Y')`"
11 ---
```

When smooth scroll is set to false in the YAML header, clicking on an item in the table of contents will navigate the reader to that section of the article without animation.

The screenshot shows the RStudio interface. On the left, the YAML header for the 'investment\_report.Rmd' file is displayed. It includes settings for the title, output type (HTML document), and the table of contents (toc). The 'smooth\_scroll' setting is explicitly set to 'false'. The main pane shows the R Markdown code, which includes code chunks for reading CSV files and some R code for data manipulation and plotting. A green button at the bottom right of the editor area is labeled 'Knit HTML'.

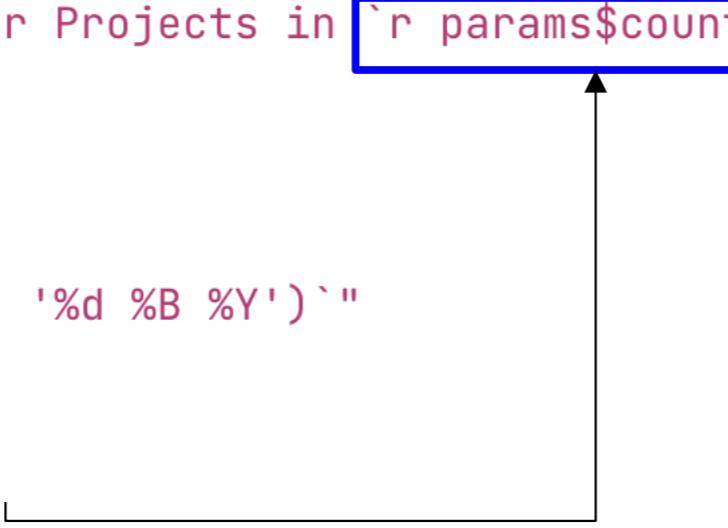
```
investment_report.Rmd
1 ---
2 title: "Investment Report"
3 output:
4 html_document:
5 toc: true
6 toc_float:
7 collapsed: false
8 smooth_scroll: false
9 toc_depth: 3
10 date: "`r format(Sys.time(), '%d %B %Y')`"
11 ---
12 ```{r setup, include = FALSE}
13 knitr::opts_chunk$set(fig.align = 'center', echo = TRUE)
14 ...
15 |
16 ```{r data, include = FALSE}
17 library(readr)
18 library(dplyr)
19 library(ggplot2)
20
21 investment_annual_summary <- read_csv("https://assets.datacamp.com/production/
repositories/5756/datasets/d0251f26117bbcfc0ea96ac276555b9003f4f7372/
investment_annual_summary.csv")
22 investment_services_projects <- read_csv("https://assets.datacamp.com/production/
repositories/5756/datasets/bcb2e39ecbe521f4b414a21e35f7b8b5c50aec64/
investment_services_projects.csv")
23 ...
24
25 ## Datasets
26 ### Investment Annual Summary
```

# Summary

- toc
  - toc\_depth
    - HTML default: 3
    - PDF default: 2
  - number\_sections
- HTML
- toc\_float
    - collapsed
    - smooth\_scroll

# Adding a parameter

```
1 ---
2 title: "Investment Report for Projects in `r params$country`"
3 output:
4 html_document:
5 toc: true
6 toc_float: true
7 date: "`r format(Sys.time(), '%d %B %Y')`"
8 params:
9 country: Indonesia
10 ---
```



# Reviewing the code

```
42 ```{r country-investment-projects}
43 country_investment_projects <- investment_services_projects %>%
44 | filter(country == params$country)
45
46 ggplot(country_investment_projects, aes(x = date_disclosed, y =
47 total_investment, color = status)) +
48 | geom_point() +
49 | labs(
50 | title = "Investment Services Projects in Indonesia",
51 | x = "Date Disclosed",
52 | y = "Total IFC Investment in Dollars in Millions"
53)
54 ````
```

# Reviewing the text

```
39 ### Investment Projects in `r params$country`
40 The `investment_services_projects` dataset provides information
about each investment project in `r params$country` from 2012 to
2018. Information listed includes the project name, company name,
sector, project status, and investment amounts.
```

# Adding parameters to define fiscal year

```
1 ---
2 title: "Investment Report for Projects in `r params$country`"
3 output:
4 html_document:
5 toc: true
6 toc_float: true
7 date: "`r format(Sys.time(), '%d %B %Y')`"
8 params:
9 country: Indonesia
10 year_start: 2011-07-01
11 year_end: 2012-06-30
12 fy: 2012
13 ---
14 `r country-investment-projects-2012}
15 country_investment_projects_2012 <- investment_services_projects %>%
16 filter(country == params$country,
17 date_disclosed >= params$year_start,
18 date_disclosed <= params$year_end)
19
20 ggplot(country_investment_projects_2012, aes(x = date_disclosed, y =
21 total_investment, color = status)) +
22 geom_point() +
23 labs(
24 title = "Investment Services Projects",
25 x = "Date Disclosed",
26 y = "Total IFC Investment in Millions"
27)
28 ...
29
30 59 ### Investment Projects in `r params$country` in `r params$fy`
31 60 The `investment_services_projects` dataset was filtered below to focus on
32 information about each investment project from the `r params$fy` fiscal
33 year, and is referred to as `country_annual_investment_projects`.
```

# Customizing the report

REPORTING WITH R MARKDOWN



Amy Peterson  
Head of Core Curriculum at DataCamp

# Specifying element style

- ```
15 <style>
16
17
18
19 </style>
```
- color
 - background-color
 - font-family
 - font-size

The table of contents

```
15 <style>
16 #TOC {
17   color: #708090;
18   font-family: Calibri;
19   font-size: 16px;
20   border-color: #708090;
21 }
22 body {
23   color: #708090;
24   font-family: Calibri;
25   background-color: #F5F5F5;
26 }
27 pre {
28   color: #708090;
29   background-color: #F8F8FF;
30 }
31 </style>
```

Datasets
Investment Annual Summary
Investment Projects from the 2012 to 2018 Fiscal Years
Investment Projects in Brazil in 2018

Investment Report for Projects in Brazil

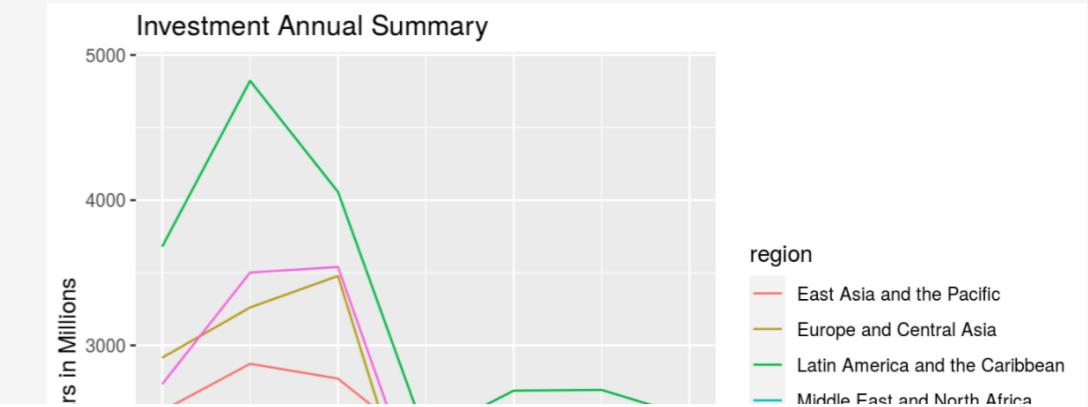
08 May 2020

Datasets

Investment Annual Summary

The `investment_annual_summary` dataset provides a summary of the dollars in millions provided to each region for each fiscal year, from 2012 to 2018.

```
ggplot(investment_annual_summary, aes(x = fiscal_year, y = dollars_in_millions, color = region)) +
  geom_line() +
  labs(
    title = "Investment Annual Summary",
    x = "Fiscal Year",
    y = "Dollars in Millions"
  )
```



The header

```
22 #header {  
23   color: #800000;  
24   background-color: #F5F5F5;  
25   opacity: 0.6;  
26   font-family: Calibri;  
27   font-size: 20px;  
28 }
```

Datasets
Investment Annual Summary
Investment Projects from the 2012 to 2018 Fiscal Years
Investment Projects in Brazil in 2018

Investment Report for Projects in Brazil

08 May 2020

Datasets

Investment Annual Summary

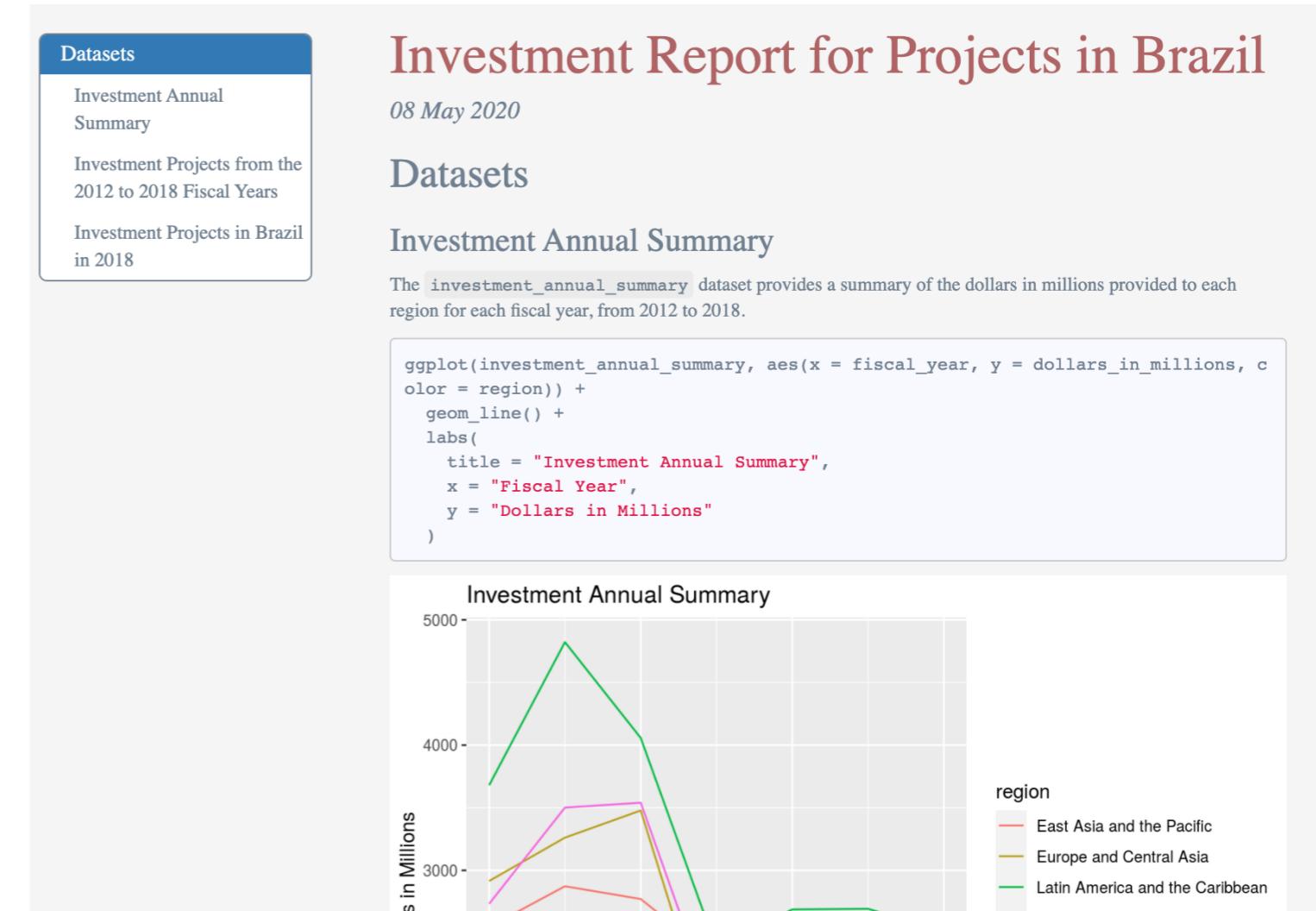
The `investment_annual_summary` dataset provides a summary of the dollars in millions provided to each region for each fiscal year, from 2012 to 2018.

```
ggplot(investment_annual_summary, aes(x = fiscal_year, y = dollars_in_millions, color = region)) +  
  geom_line() +  
  labs(  
    title = "Investment Annual Summary",  
    x = "Fiscal Year",  
    y = "Dollars in Millions"  
)
```



The title, author, and date

```
22 h1.title {  
23   color: #800000;  
24   background-color: #F5F5F5;  
25   opacity: 0.6;  
26   font-family: Calibri;  
27   font-size: 40px;  
28 }  
29 h4.author {  
30   color: #708090;  
31   font-family: Calibri;  
32 }  
33 h4.date {  
34   color: #708090;  
35   font-family: Calibri;  
36 }
```



CSS file

```
1 ---  
2 title: "Investment Report for Projects in `r params$country`"  
3 output:  
4   html_document:  
5     css: styles.css  
6     toc: true  
7     toc_float: true  
8 date: "`r format(Sys.time(), '%d %B %Y')`"  
9 params:  
10    country: Brazil  
11    year_start: 2017-07-01  
12    year_end: 2018-06-30  
13    fy: 2018  
14 ---
```

investment_report.Rmd	styles.css
	1 #TOC { 2 color: #708090; 3 font-family: Calibri; 4 font-size: 16px; 5 border-color: #708090; 6 } 7 h1.title { 8 color: #F08080; 9 background-color: #F5F5F5; 10 opacity: 0.6; 11 font-family: Calibri; 12 font-size: 20px; 13 } 14 h4.author { 15 color: #708090; 16 font-family: Calibri; 17 background-color: #F5F5F5; 18 } 19 h4.date { 20 color: #708090; 21 font-family: Calibri; 22 background-color: #F5F5F5; 23 } 24 body { 25 color: #708090. }