



[Course](#) > [Section...](#) > [4.2 Usi...](#) > [Density...](#)

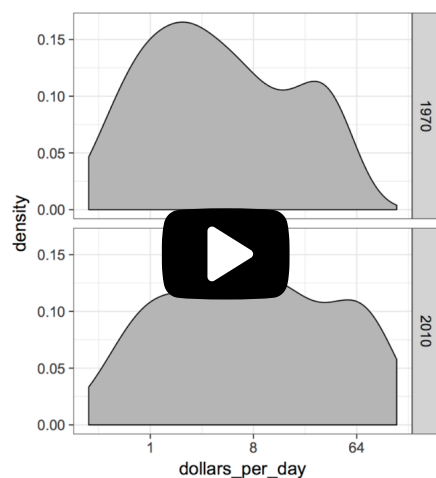
Audit Access Expires Mar 24, 2020

You lose all access to this course, including your progress, on Mar 24, 2020.

Upgrade by Feb 18, 2020 to get unlimited access to the course as long as it exists on the site. [Upgrade now](#)

Density Plots

Density Plots



[Start of transcript. Skip to the end.](#)



RAFAEL IRIZARRY: We have used data exploration

to discover that the income gap between rich and poor countries

has closed considerably during the last forty years.

We use a series of histograms and box plots to see this.

Video



[Download video file](#)

Transcripts

[Download SubRip \(.srt\) file](#)

[Download Text \(.txt\) file](#)

Textbook link

This video corresponds to the following sections:

- The end of the [textbook section on 1970 versus 2010 income distributions](#)
- [Textbook section on accessing computed variables](#)
- [Textbook section on weighted densities](#)

Key points

- Change the y-axis of density plots to variable counts using `..count..` as the y argument.
- The `case_when()` function defines a factor whose levels are defined by a variety of logical operations to group data.
- Plot stacked density plots using `position="stack"`.
- Define a weight aesthetic mapping to change the relative weights of density plots - for example, this allows weighting of plots by population rather than number of countries.

Code: Faceted smooth density plots



```
# see the code below the previous video for variable definitions

# smooth density plots - area under each curve adds to 1
gapminder %>%
  filter(year == past_year & country %in% country_list) %>%
  mutate(group = ifelse(region %in% west, "West", "Developing")) %>%
  summarize(n = n()) %>% knitr::kable()

# smooth density plots - variable counts on y-axis
p <- gapminder %>%
  filter(year == past_year & country %in% country_list) %>%
  mutate(group = ifelse(region %in% west, "West", "Developing")) %>%
  ggplot(aes(dollars_per_day, y = ..count.., fill = group)) +
  scale_x_continuous(trans = "log2")
p + geom_density(alpha = 0.2, bw = 0.75) + facet_grid(year ~ .)
```

Code: Add new region groups with case_when

```
# add group as a factor, grouping regions
gapminder <- gapminder %>%
  mutate(group = case_when(
    .$region %in% west ~ "West",
    .$region %in% c("Eastern Asia", "South-Eastern Asia") ~ "East Asia",
    .$region %in% c("Caribbean", "Central America", "South America") ~ "Latin America",
    .$continent == "Africa" & .$region != "Northern Africa" ~ "Sub-Saharan Africa",
    TRUE ~ "Others"))

# reorder factor levels
gapminder <- gapminder %>%
  mutate(group = factor(group, levels = c("Others", "Latin America",
```

Code: Stacked density plot



```
# note you must redefine p with the new gapminder object first
p <- gapminder %>%
  filter(year %in% c(past_year, present_year) & country %in% country_li
  ggplot(aes(dollars_per_day, fill = group)) +
  scale_x_continuous(trans = "log2")

# stacked density plot
p + geom_density(alpha = 0.2, bw = 0.75, position = "stack") +
  facet_grid(year ~ .)
```

Code: Weighted stacked density plot

```
# weighted stacked density plot
gapminder %>%
  filter(year %in% c(past_year, present_year) & country %in% country_
  group_by(year) %>%
  mutate(weight = population/sum(population*2)) %>%
  ungroup() %>%
  ggplot(aes(dollars_per_day, fill = group, weight = weight)) +
  scale_x_continuous(trans = "log2") +
  geom_density(alpha = 0.2, bw = 0.75, position = "stack") + facet_gr
```

Learn About Verified Certificates

© All Rights Reserved

