This plot nicely shows the distribution of the data. I added a filter to r gapminder %% filter(year == 1952 & gdpPercap > 15000) %% pull(countr which is an exterme outlier. See here: title: "R Notebook" output: html_notebook gapminder %>% filter(year == 1952 & gdpPercap > 15000) library(gapminder) library(dplyr) gapminder %>% *Gapminder* project data To compare global life expectancy between 1952 and 2007 G Chunk 2 ¢ A: Anatomy of a notebook / geom_point(shape = 1) + theme_bw() + filter(gdpPercap < 15000) %>% 50 RMarkdown document 10 colour = continent, 20 = pop/1e6)) + gdpPercap/1000, #divide by 1000 to tidy the x-axis lifeExp, 1952 30 gdpPercap/1000 10 20 30 40 continent O 500 pop/1e+06 O 250 Americas Oceania Europe Asia Africa 1250 1000 750 R Markdown options Headings (h2) 🥕 Aim document type - run chunk - end ``` R Markdown run all chunks - begin ```{r} In-line code / YAML header Code chunk to here B: Rendered preview To compare global life expectancy between 1952 and 2007 Gapminder project data Data R Notebook Plot the data This plot nicely shows the distribution of the data. I added a filter to exclude Kuwait, which is an exterme outlier. See library(dplyr) library(gapminder gapminder %>% gapminder %>% facet_wrap(~year) theme_bw() + $geom_point(shape = 1) +$ ggplot(aes(x filter(gdpPercap < 60000) %>% filter(year == 2007 | year == 1952) %>% filter(year == 1952 & gdpPercap > 15000) size = pop/le6)) + colour = continent, 1952 = lifeExp, = gdpPercap/1000, #divide by 1000 to tidy the x-axis gdpPercap/1000 pop/1e+06 O 1000 continent O 750 O 500 Oceania Asia Americas Africa Hide Hide