Rainfall Maps

Maps of average monthly rainfall in each district for each year from 1901 to 2002. The rainfall data is sourced from <u>India Meteorological Department</u>

(http://www.indiawaterportal.org/articles/meteorological-datasets-download-entire-datasets-various-meteorological-indicators-1901) (IMD), the Indian district boundaries are sourced from the GADM database (http://gadm.org/), and the national boundary of India is extracted from the administrative divisions map of Census of India 2011

(http://www.censusindia.gov.in/2011census/maps/administrative maps/INDIA2011.pdf).

Maps:

How to use: Click on thumbnails to see full-size maps. Hit *Escape* to exit the lightbox.

Legends: These are <u>choropleth maps</u> (https://en.wikipedia.org/wiki/Choropleth). Each district is shaded according to the volume of average monthly rainfall for that district for that year. Colour chart is given with each map. Darker blue indicates greater rainfall. Lighter cyan indicates lesser rainfall.

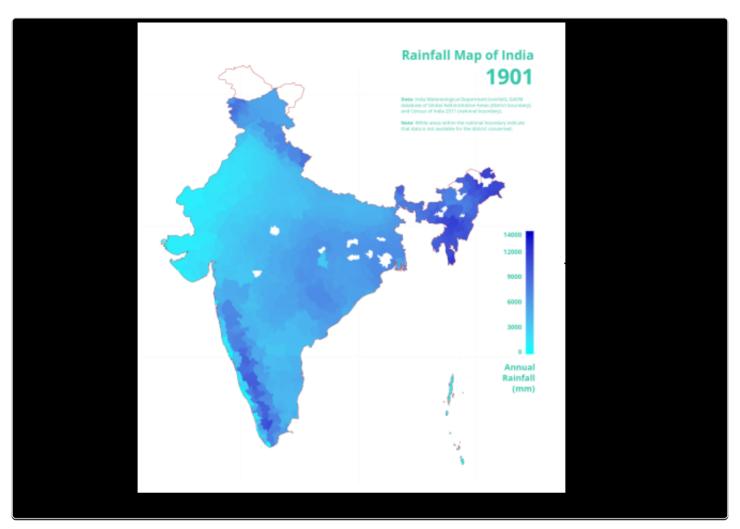
Note on rainfall data: The rainfall data available from IMD does not provide measurements for several districts, such as Ashoknagar, Annupur and Burhanpur in Madhya Pradesh, and Bongaigaon and Marigaon in Assam. In the maps, these districts can be found as white (absent data) areas. Details of the districts with absent data can be found in the file containing the R code for generating these maps (https://github.com/ajantriks/iwp/blob/master/R code/rainfall maps.R).

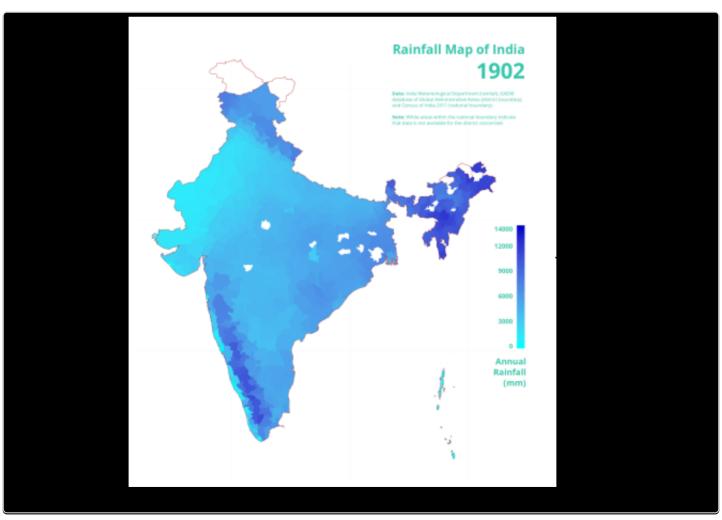
Note on administrative boundaries data: Please note that district and national boundary data available from GADM (http://gadm.org/) and extracted from administrative divisions map of Census of India 2011

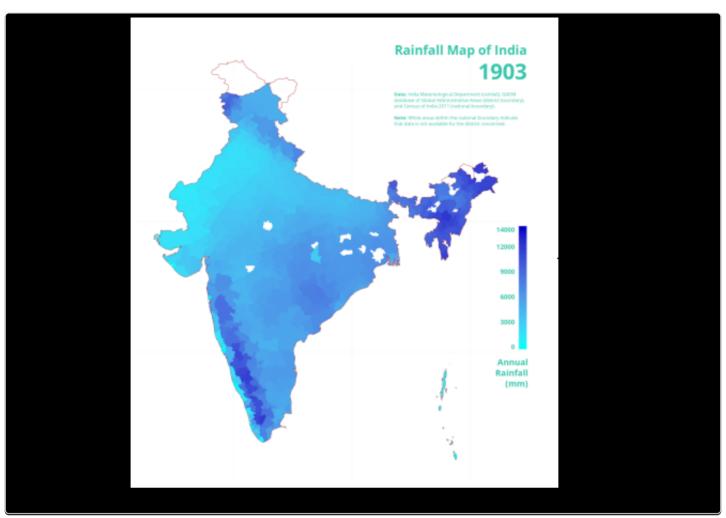
(http://www.censusindia.gov.in/2011census/maps/administrative maps/INDIA2011.pdf) may not accurately represent the official national boundary of the country, and may neither include the most recent list of districts.

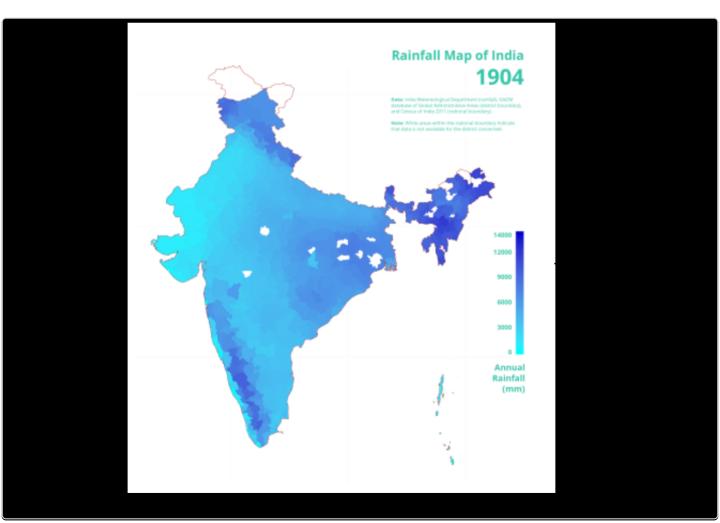
Animated rainfall map: Click here to watch

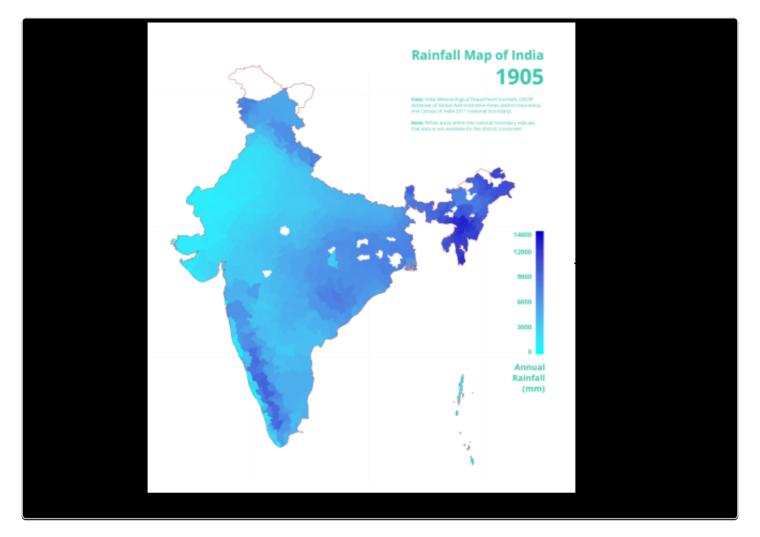
(http://ajantriks.github.io/iwp/images/rainfall maps/imd rainfall 1901-2002 animation.gif) an animated version of the rainfall maps from 1901 to 2002. The original file can also be accessed at GitHub (https://github.com/ajantriks/iwp/blob/master/images/rainfall maps/imd rainfall 1901-2002 animation.gif).

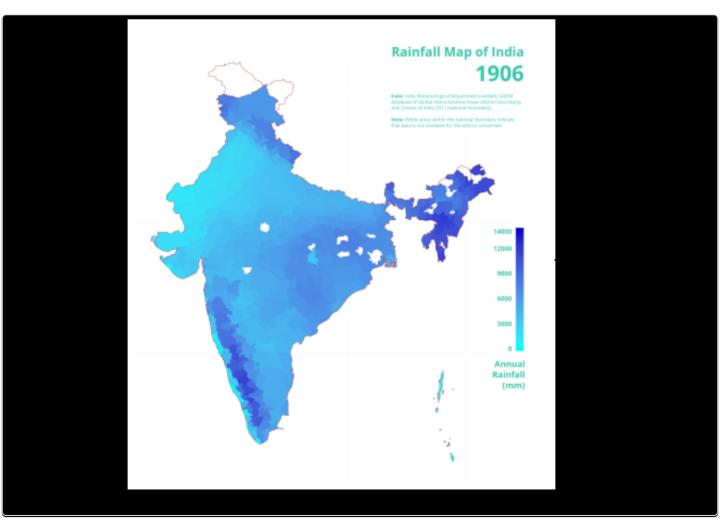


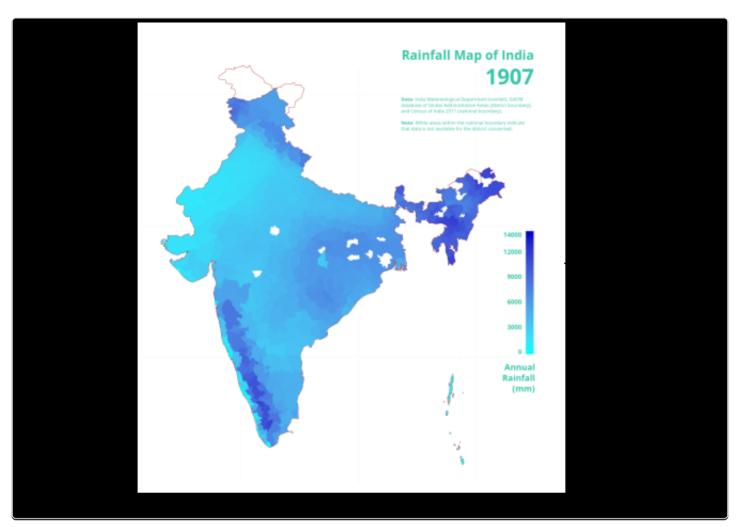


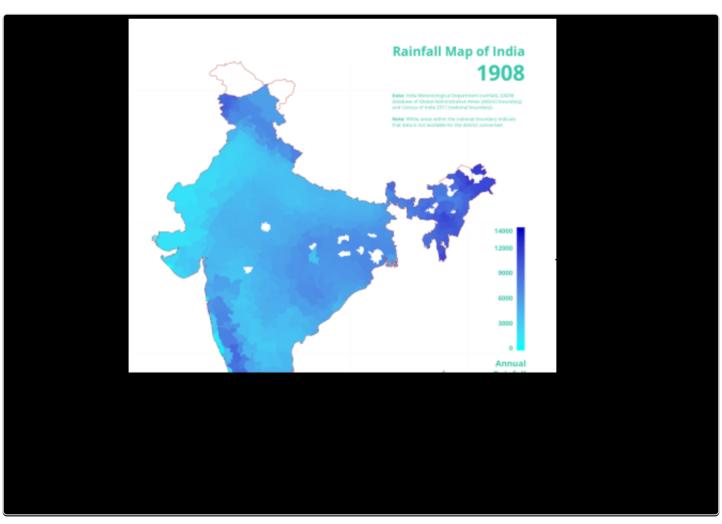




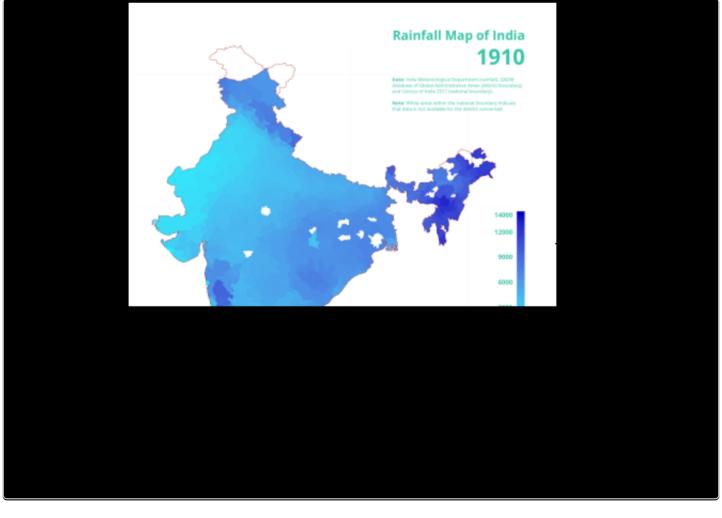




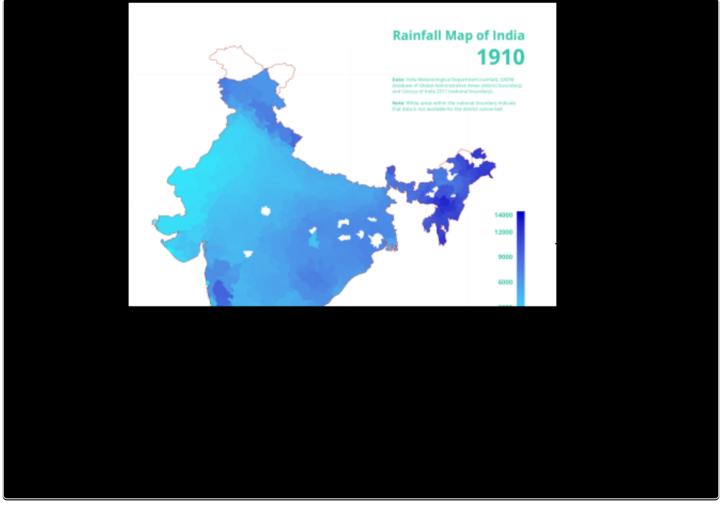




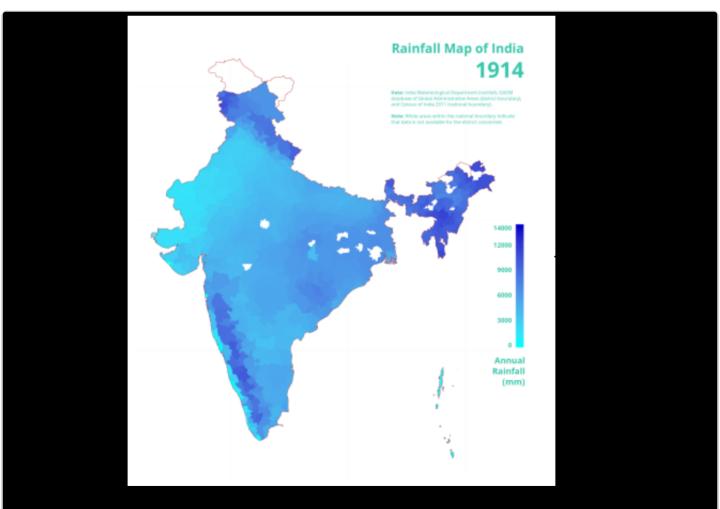


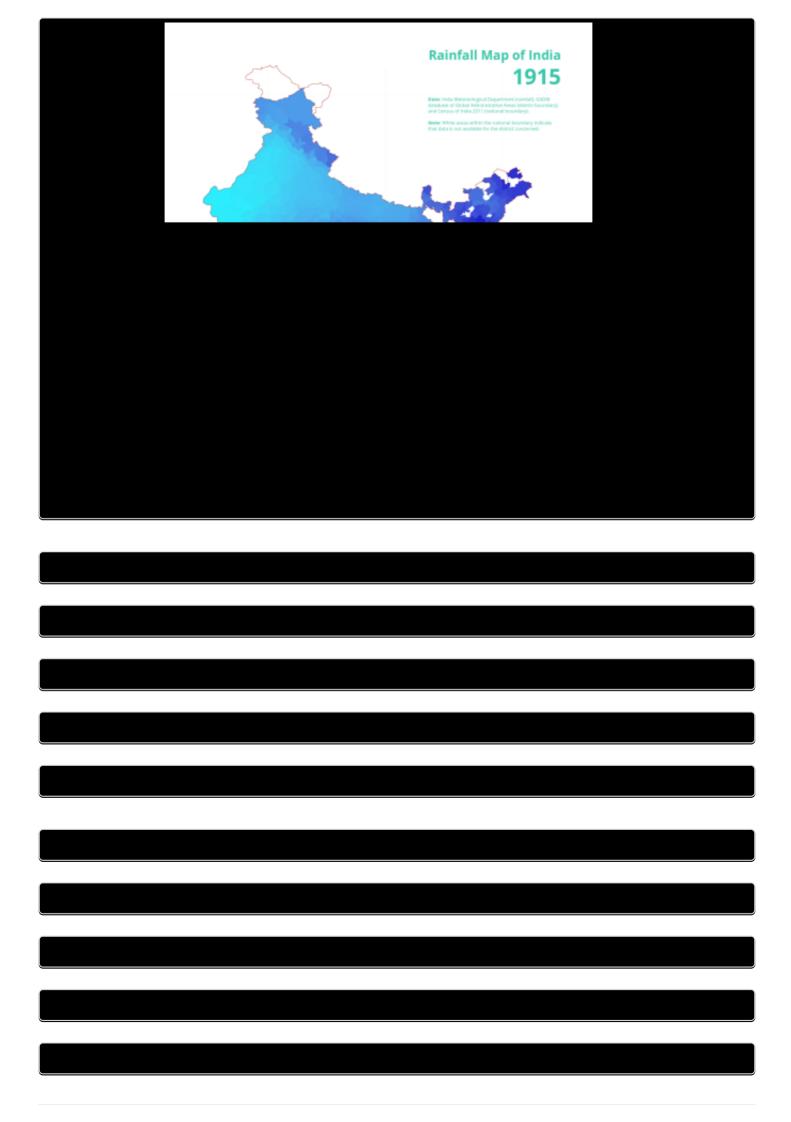


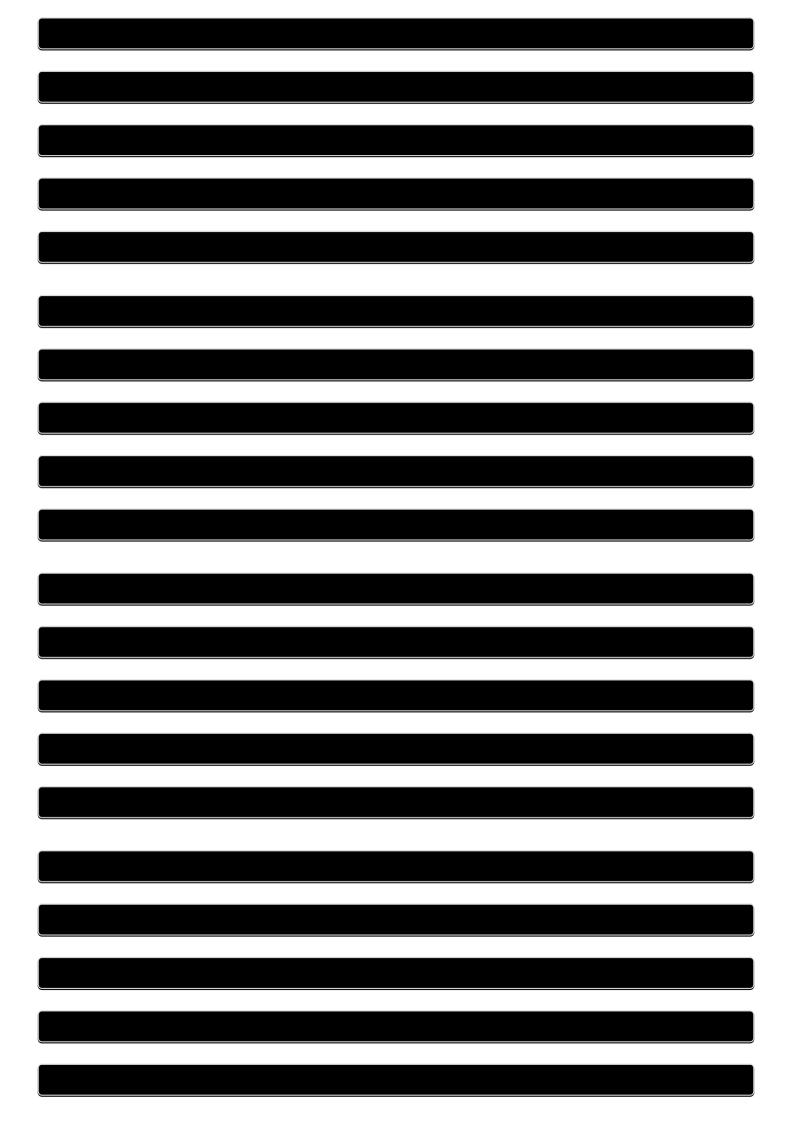


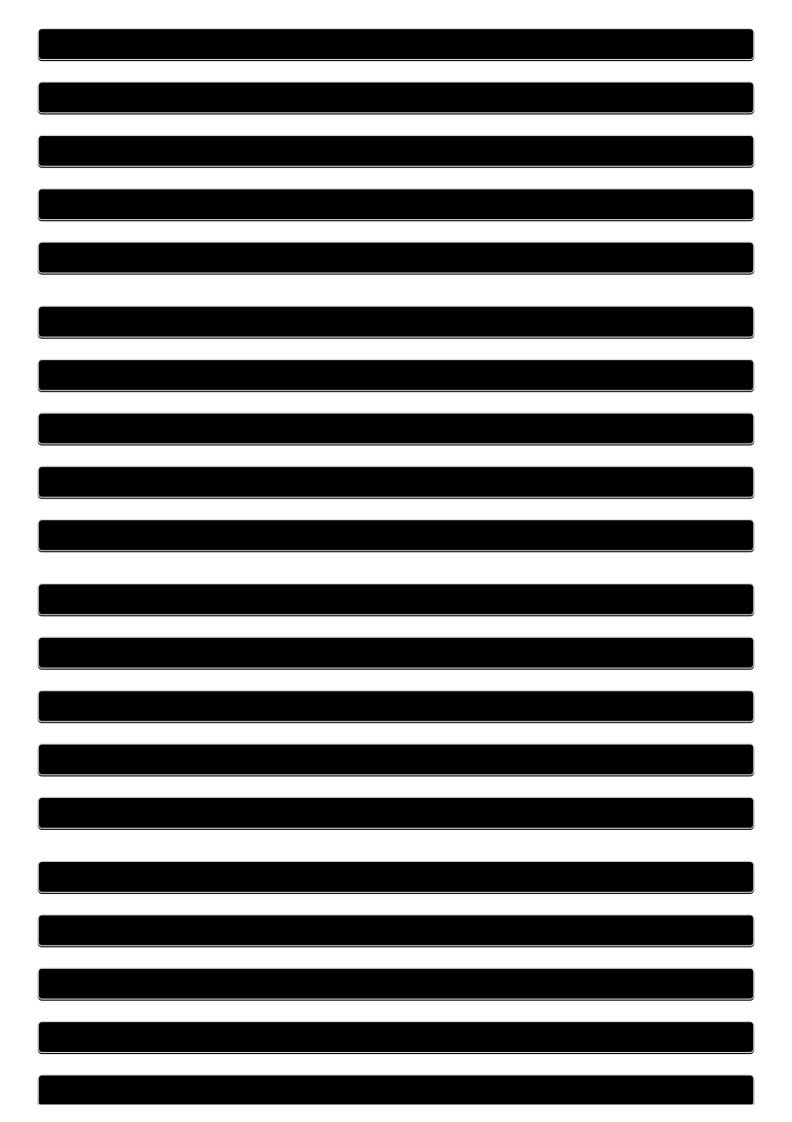


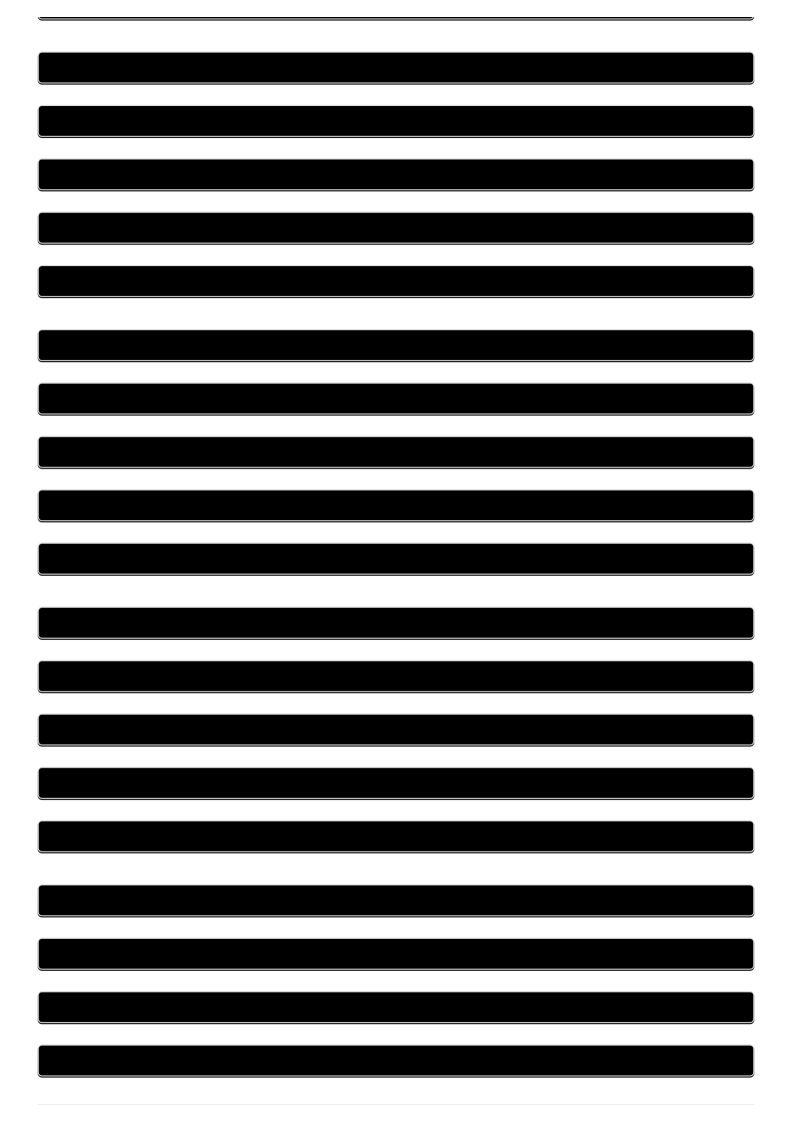


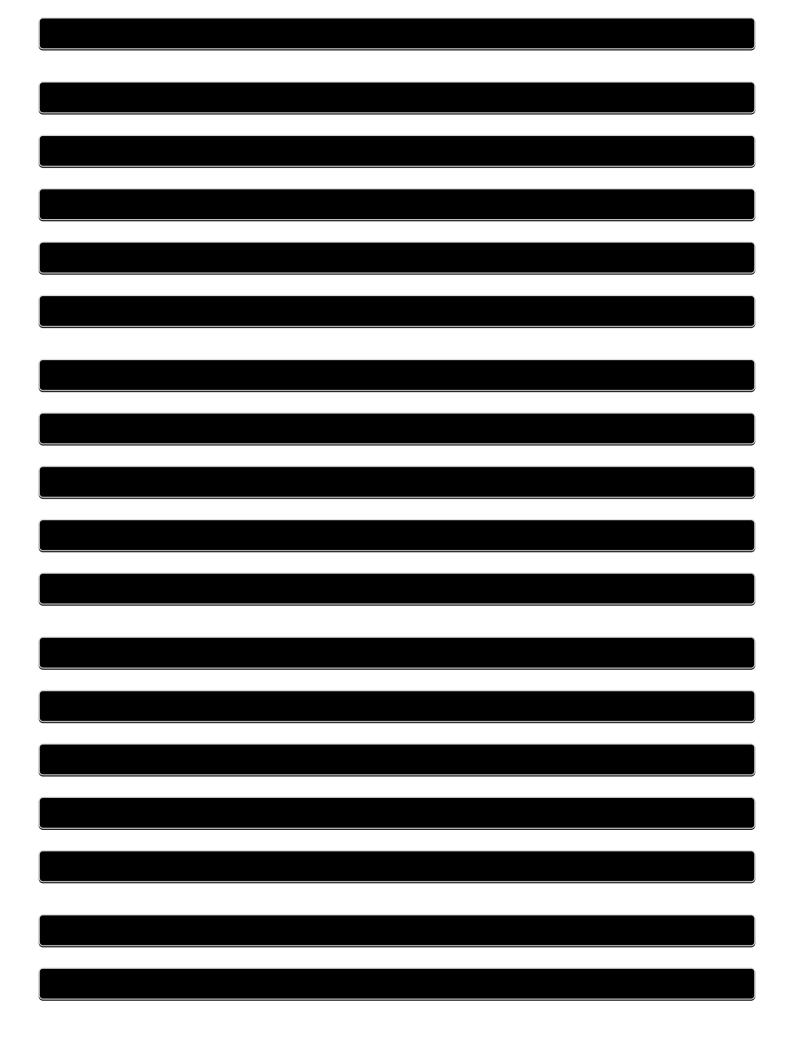












The original IMD rainfall data for 1901-2002 comes in three MS Excel spreadsheets, and each of them contains data for a set of states (for the entire period). The data is organised as follows (with the state and district codes given in a separate sheet of the spreadsheet file):

```
State ID, District ID, Year, January, February, March, April, May, June, July, August, September, October, November, December

1,1,1901,398.24,331.86,563.32,757.26,944.61,802.88,1020.44,432.67,717.18,223.14,44.45,101.16

1,1,1902,255.22,331.97,746.78,962.14,954.47,1116.03,910.09,688.77,542.22,434.55,523.79,159.64

1,1,1903,607.9,589.12,692.71,594.51,830.91,817.7,1094.51,983.38,732.75,165.03,146.73,210.24
```

To use this data along with the Indian district boundaries datset from <u>GADM (http://gadm.org/)</u>, it needs to be re-organised in the following format, where each row gives the annual (aggregate) rainfall (for the year concerned) for a specific district for a specific year:

```
"State.ID", "District.ID", "Year", "Annual.Rainfall"
"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1901, 6337.21
"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1902, 7625.67
"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1903, 7465.49
```

This format is useful because it allows the *District.ID* value of a row to be used to link up the count of annual (aggregate) rainfall for a district (and for a year) to its corresponding district boundary shape (available from the Indian district boundaries data file). To enable this, the name of the districts must match between the rainfall data file and the district boundaries shapefile.

To do this, first IMD rainfall data files are converted from MS Excel spreadsheet format to .csv (Comma Separated Value) format, and read into R using the read.csv command. Then they are combined into a single file ("imd_states_name_1-35.csv") using the rbind command.

```
d1 <- read.csv("imd_states_1-10.csv")
d2 <- read.csv("imd_states_11-25.csv")
d3 <- read.csv("imd_states_26-35.csv")
data <- rbind(d1,d2,d3)
View(data)
write.csv(data, file="imd_states_name_1-35.csv")</pre>
```

Then the state and district codes from the IMD rainfall data file is converted to their respective names, by following the spellings of those state and district names as given in the GADM district boundaries shapefile. The second line below replace the values of *State.ID* by "Jammu and Kashmir" (spelt as in GADM district boundaries shapefile) if the present value is "1". The third line below replace the values of *District.ID* by "Kupwara (Muzaffarabad)" (spelt as in GADM district boundaries shapefile) if the present value is "1" and the corresponding *District.ID* value is "Jammu and Kashmir".

```
d <- read.csv ("imd_states_1-10.csv")

d$State.ID [d$State.ID == "1"] <- "Jammu and Kashmir"

d$District.ID [d$State.ID == "Jammu and Kashmir" & d$District.ID == "1"] <- "Kupwara (Muzaffarabad)"</pre>
```

Once such replacing of the state and district codes by their corresponding names (spelt as in GADM district boundaries shapefile) is done, the data takes the following form:

```
"State.ID", "District.ID", "Year", "January", "February", "March", "April", "May", "June", "July", "August", "September", "October", "November", "December"

"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1901, 398.24, 331.86, 563.32, 757.26, 944.61, 802.88, 1020.44, 432.67, 717.18, 223.14, 44.45, 101.16

"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1902, 255.22, 331.97, 746.78, 962.14, 954.47, 1116.03, 910.09, 688.77, 542.22, 434.55, 523.79, 159.64

"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1903, 607.9, 589.12, 692.71, 594.51, 830.91, 817.7, 1094.51, 983.38, 732.75, 165.03, 146.73, 210.24
```

The three IMD rainfall data spreadsheets are converted to this form and combined into a single .csv file, using the rbind command, and the resultant data file is named "imd_states_name_1-35.csv"

```
d1 <- read.csv("imd_states_1-10.csv")
d2 <- read.csv("imd_states_11-25.csv")
d3 <- read.csv("imd_states_26-35.csv")
data <- rbind(d1,d2,d3)
View(data)
write.csv(data, file="imd_states_name_1-35.csv")</pre>
```

Now the present structure of the table, where each row shows monthly rainfall counts for all months for a single year, needs to be changed to a structure where each rows gives the annual (aggregate) rainfall in a district for an unique combination of year and month. This is done using the melt command from the Rehape2 package of R (http://cran.r-

<u>project.org/web/packages/reshape2/)</u>, by converting the month columns ("January" to "December") into a single column that takes the months as its values.

```
d <- read.csv ("imd_states_name_1-35.csv")
library(reshape2)
dm <- melt(d, id = c('State.ID', 'District.ID', 'Year'))</pre>
```

This converts the data to the following format:

```
"State.ID", "District.ID", "Year", "variable", "value"
"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1901, "January", 398.24
"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1902, "January", 255.22
"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1903, "January", 607.9
```

But this gives the volume of rainfall for each district for each month (for each year), and not the annual (aggregate) rainfall for each district (for each year) that we need to generate the maps concerned.

The following commands generate this annual (aggregate) rainfall for each district (for each year), by taking the average of *value* for all rows with unique combinations of *State.ID*, *District.ID* and *Year*, and multiplying it by "12".

```
dm$Avg.Monthly.Rainfall <- with(dm, ave(value, State.ID, District.ID, Year))
dm$Annual.Rainfall <- with(dm, Avg.Monthly.Rainfall * 12)</pre>
```

This converts the data to the following format:

```
"State.ID", "District.ID", "Year", "variable", "value", "Avg.Monthly.Rainfall", "Annu al.Rainfall"

"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1901, "January", 398.24,528.10083333
3333,6337.21

"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1902, "January", 255.22,635.4725,762
5.67

"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1903, "January", 607.9,622.124166666
667,7465.49
```

The reformatted data table is saved.

```
write.csv(dm, file = "imd_states_name_1-35_annual-rainfall.csv")
```

The resultant data file still has separate rows for all the months of each year. Since, the annual (aggregate) rainfall for each year have already been generated, there is no need for these rows for separate months. The first line below takes a subset of the data with only rows for the month of "January" (for all years); and the next lines remove the variables that are not needed anymore and save the data table again.

```
d2 <- subset(dm, dm$variable == "January")

d2$variable <- NULL

d2$value <- NULL

d2$Avg.Monthly.Rainfall <- NULL

write.csv(d2, file = "imd_states_name_1-35_annual-rainfall-edited.csv")</pre>
```

Now we have the data in the following format, and it is ready for mapping:

```
"State.ID", "District.ID", "Year", "Annual.Rainfall"
"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1901, 6337.21
"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1902, 7625.67
"Jammu and Kashmir", "Kupwara (Muzaffarabad)", 1903, 7465.49
```

The Indian administrative boundaries (shapefile) data file is downloaded from the <u>GADM website</u> (http://gadm.org/country), and opened using the readogR command of the rgdal library of R (http://cran.r-project.org/web/packages/rgdal/index.html). Further, the shapefile data is converted to a R data frame, which can be used by the ggplot2 library (http://cran.r-project.org/web/packages/ggplot2/index.html) for mapping, by using the fortify command.

```
library(sp)
library(rgdal)
i <- readOGR (dsn = "IND_adm/", layer = "IND_adm2")

library(maptools)
gpclibPermit()
library(ggplot2)
library(scales)
m <- fortify(i, region = "NAME_2")</pre>
```

Please note that the Indian administrative boundaries data file from GADM contains state, district and sub-district boundaries. For the purpose of these maps, the district level boundaries are used, by choosing the "IND_adm2" layer, and by fortifying the map data with the "NAME_2" variable that contains the district names.

To create map of annual (aggregate) rainfall across districts for a specific year, the final reorganised rainfall data file is opened and the corresponding subset is taken (the year "1901" is used in the example below).

Please note that if the opened data file has a column titled *X*, remove it using this command:

```
d$X <- NULL
```

```
d <- read.csv ("imd_states_name_1-35_annual-rainfall.csv")
d1901 <- subset (d, d$Year == "1901")</pre>
```

Finally, we draw the map using the <u>ggplot2 library of R (http://cran.r-project.org/web/packages/ggplot2/index.html)</u>.

```
ggplot(d1901, aes(map_id = District.ID)) + geom_map(aes(fill = Annual.Rainfall)
, map = m) + expand_limits(x = m$long, y = m$lat) + scale_fill_gradient(low = "
#00FFFF", high = "#0000C8", space = "Lab", limits = c(0, 14000), breaks = c(0,
3000, 6000, 9000, 12000, 14000)) + theme(panel.background = element_rect(fill =
    "white"), axis.line = element_line(color = "white"), axis.ticks = element_line
    (color = "white"), axis.text = element_text(color = "white")) + guides(fill = g
    uide_colorbar(title = "Annual Rainfall\n(mm)\n", title.theme = element_text(col
    our = "#4444444", size = "40", angle = 0), barwidth = 2, barheight = 30, label.t
    heme = element_text(colour = "#4444444", size = "38", angle = 0))) + xlab("") +
    ylab("")
```

Code and Data:

The entire R code (as given in the documentation) for this visualisation can be downloaded <u>from</u> here (https://github.com/ajantriks/iwp/blob/master/R code/rainfall maps.R).

Data (reorganised from the original data shared by IMD) used to create the above charts can also be downloaded <u>from here (https://github.com/ajantriks/iwp/blob/master/data/imd_states_name_1-35_annual-rainfall.csv)</u>.

Credits and License:

Created by <u>Sumandro Chattapadhyay (http://www.ajantriks.net/)</u> for <u>India Water Portal (http://indiawaterportal.org/)</u>.

All visualisations are developed using R (http://cran.r-project.org/).

Pages written using <u>Bootstrap</u> (<u>http://twitter.github.com/bootstrap/</u>), <u>Google Web Fonts</u> (<u>http://www.google.com/webfonts/</u>), <u>Highlight.js</u> (<u>http://softwaremaniacs.org/soft/highlight/en/</u>) and <u>Lightbox</u> (<u>http://lokeshdhakar.com/projects/lightbox2/</u>).

Code for all the visualisations is available on GitHub (https://github.com/ajantriks/iwp).

Code (but not data) is shared under Creative Commons <u>Attribution-ShareAlike 3.0 Unported</u> (http://creativecommons.org/licenses/by-sa/3.0/) license.

Reorganised <u>data</u> (https://github.com/ajantriks/iwp/tree/master/data) has been shared only for the purpose of checking and validating the charts and maps.

It is necessary to take required permission from the original creator of data before re-using it.

Note: Please link back to this page while sharing or re-using this work.