

World Bank Data Viz

Qi Xue - qx61

October 10, 2021

Data source

World Bank Geocoded Research Release, Version 1_4_2 data

```
knitr::opts_chunk$set(message = FALSE)
knitr::opts_chunk$set(warning = FALSE)

#install.packages("sf")
#install.packages("treemapify")

library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --

## v ggplot2 3.3.5      v purrr 0.3.4
## v tibble 3.1.6       v dplyr 1.0.7
## v tidyr 1.1.4        v stringr 1.4.0
## v readr 2.1.0        v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

library(ggplot2)
library(treemapify)

project <- read_csv("projects.csv")
locations <- read_csv("locations.csv")
ancillary <- read_csv("projects_ancillary.csv")
```

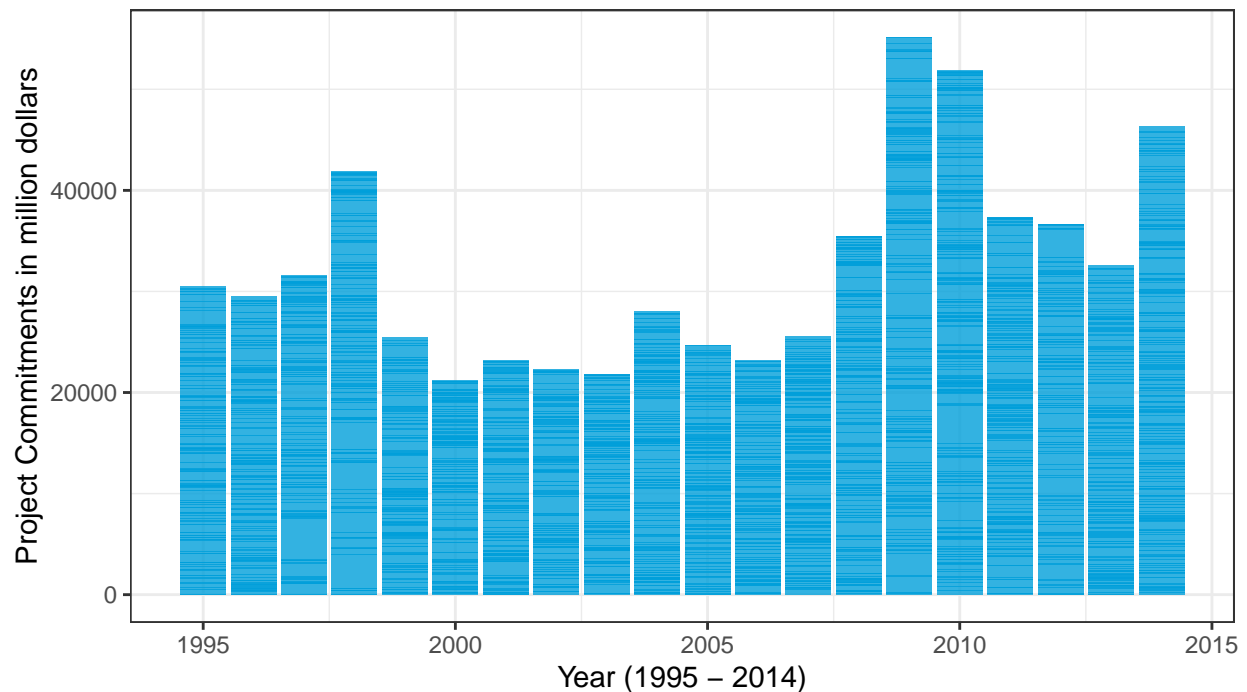
Graph 1 - Project commitments over 1995-2014

```
project %>%
  mutate(year = as.numeric(substring(start_actual_isodate, 1, 4))) %>%
  mutate(m_total_commitments = total_commitments/1000000) %>%
  ggplot(aes(x = year, y = m_total_commitments)) +
```

```
geom_col(fill = "#009FDA", alpha = 0.8) +
theme(plot.margin = unit(c(1, 1, 1, 1), "cm")) +
theme_bw() +
labs(title = paste("Project commitments are on the rise since 2008"),
      subtitle = paste("Total World Bank project commitments from 1995 to 2014"),
      x = "Year (1995 - 2014)",
      y = "Project Commitments in million dollars",
      caption = "\nData Source: AidData | World Bank Geocoded Research Release")
```

Project commitments are on the rise since 2008

Total World Bank project commitments from 1995 to 2014



Data Source: AidData | World Bank Geocoded Research Release

```
ggsave(filename = "project_commitments.png")
```

World Bank project commitments remained relatively stable before 2008. Yet, year 1998 saw a spike in investment may possibly be due to the Asian financial crisis. Project spending also rose after the 2008 financial crisis, and fluctuated more than in previous periods. We see a more active role for the World Bank not only investing in long-term results but crisis response.

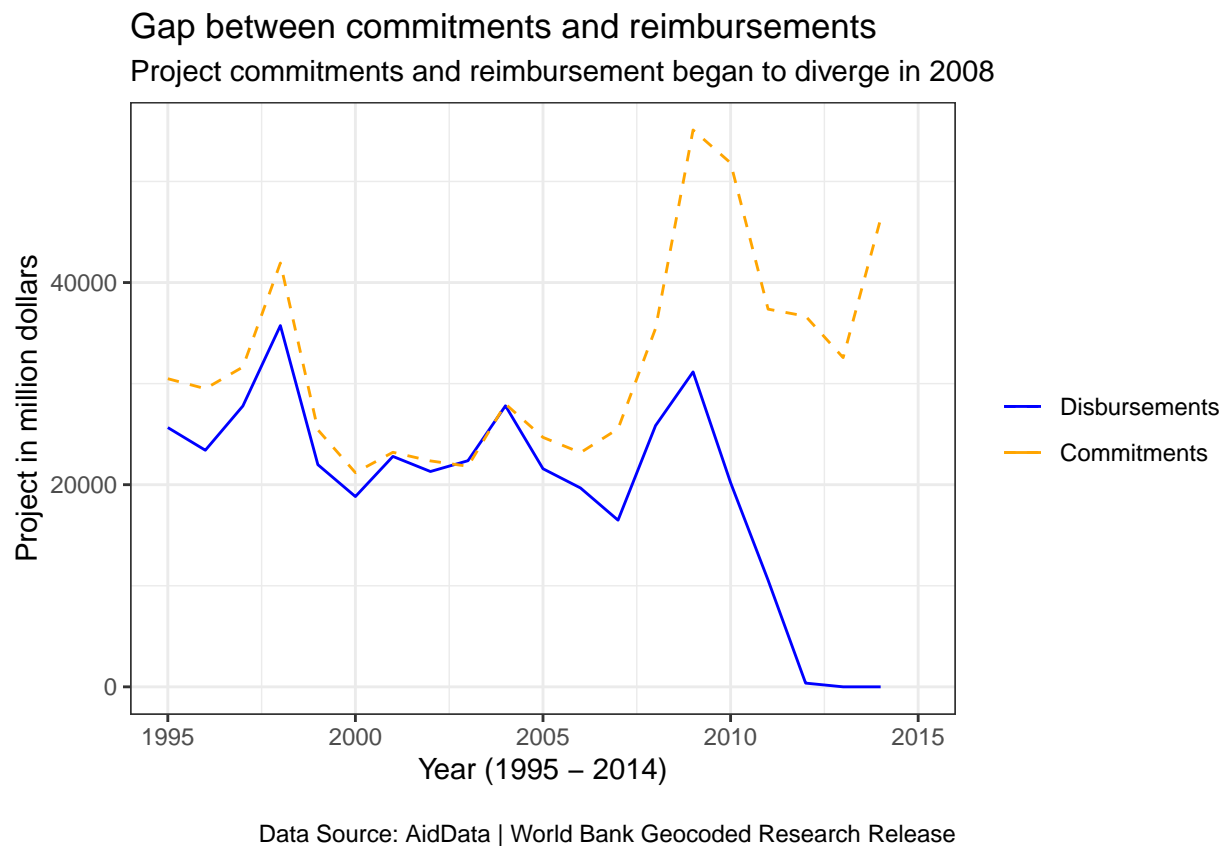
Graph2 - Gap between commitments and reimbursements

```
project %>%
  mutate(m_total_commitments = total_commitments/1000000,
         m_total_disbursements = total_disbursements/1000000,
         year = as.numeric(substring(start_actual_isodate, 1, 4))) %>%
```

```

group_by(year) %>%
summarize(n_m_total_commitments = sum(m_total_commitments, na.rm = TRUE),
          n_m_total_disbursements = sum(m_total_disbursements, na.rm = TRUE)) %>%
ggplot(aes(x = year)) +
geom_line(aes(y = n_m_total_disbursements, color = "Disbursements"),
          group = 1,
          show.legend = TRUE) +
geom_line(aes(y = n_m_total_commitments, color = "Commitments"),
          linetype = "dashed",
          show.legend = TRUE) +
theme_bw() +
scale_color_manual(name = "",
                  values = c("Disbursements" = "blue",
                             "Commitments" = "orange")) +
scale_x_continuous(limits = c(1995, 2015),
                  breaks = c(1995, 2000, 2005, 2010, 2015)) +
scale_y_continuous(limits = c(0, NA)) +
labs(title = paste("Gap between commitments and reimbursements"),
     subtitle = paste("Project commitments and reimbursement began to diverge in 2008"),
     x = "Year (1995 - 2014)",
     y = "Project in million dollars",
     caption = "\nData Source: AidData | World Bank Geocoded Research Release")

```



```

ggsave(filename = "commitments_gap.png")

```

This line graph intends to see the gap between project commitments and disbursements. We would be interested know how the World Bank is balancing of its project budget. Further, we could learn more about when and where projects experience with the largest gaps in funding. Before 2005, the reimbursement could roughly meet the commitments. Yet, project commitments and disbursement began to diverge in 2008, resulting in an enlarging gap.

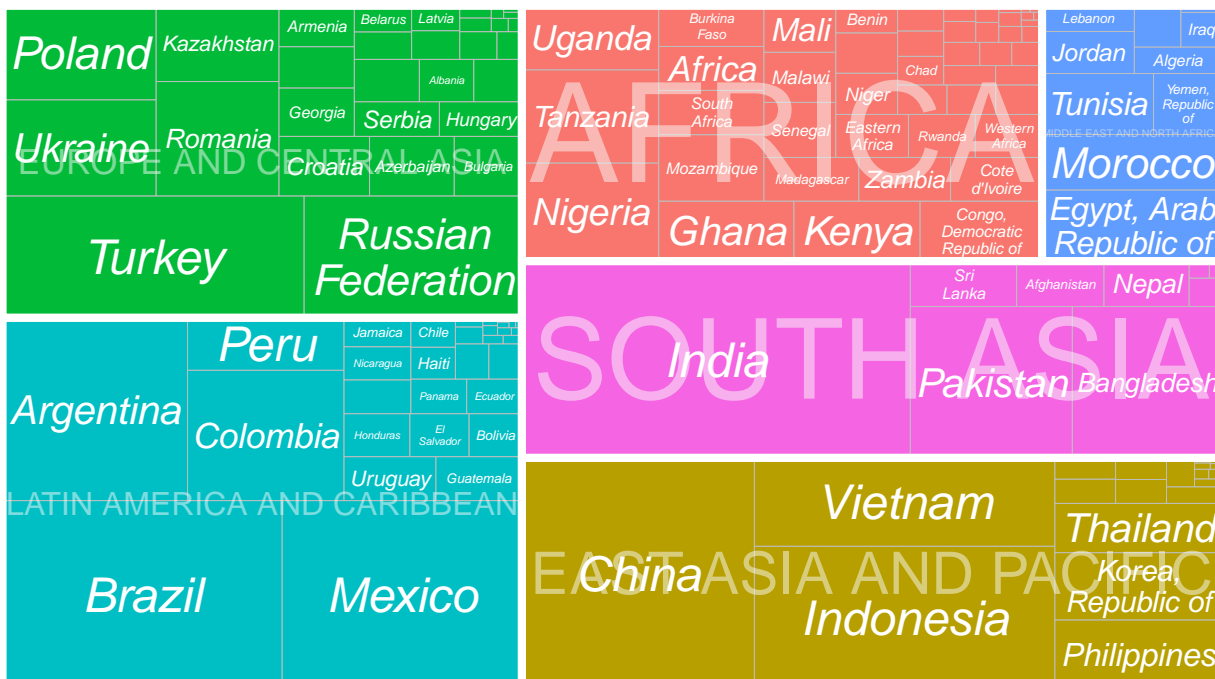
Graph 3 - Top projects destinations

```
ancillary <- ancillary %>%
  rename(project_id = `PROJECT ID`)

project_ancillary <- left_join(project, ancillary, by = "project_id")

project_ancillary %>%
  select(COUNTRY, REGION, total_commitments) %>%
  group_by(COUNTRY, REGION) %>%
  summarize(c_total_commitments = sum(total_commitments)) %>%
  ggplot(aes(area = c_total_commitments,
             label = COUNTRY,
             subgroup = REGION,
             fill = REGION)) +
  geom_treemap() +
  geom_treemap_subgroup_border(color = "white") +
  geom_treemap_subgroup_text(place = "center",
                             grow = TRUE,
                             color = "grey98",
                             alpha = 0.5,
                             min.size = 0) +
  geom_treemap_text(fontface = "italic",
                    color = "white",
                    place = "center",
                    grow = FALSE,
                    reflow = TRUE,
                    na.rm = TRUE) +
  theme(legend.position = 'none') +
  labs(title = "Top three beneficiaries of World Bank projects are India, China and Brazil",
       subtitle = "Between 1995 to 2014, World Bank has heavily involved in projects in these countries",
       caption = "\nData Source: AidData | World Bank Geocoded Research Release")
```

Top three beneficiaries of World Bank projects are India, China and Brazil
Between 1995 to 2014, World Bank has heavily involved in projects in these countries



Data Source: AidData | World Bank Geocoded Research Release

```
ggsave(filename = "destinations.png")
```

This treemap clearly maps out the top project destinations in each region. Between 1995 and 2014, we saw most investments go to India, China, Brazil, Mexico and Turkey. India and China lifted millions out of poverty, contributing to half of the world's poverty reduction. We expect the focus to shift to other developing countries in urgent need of aid and technical assistance.

Graph 4 - Distrbution of World Bank IBRD-IDA projects

```
library(sf)
library(tigris)

options(tigris_use_cache = TRUE)

new_locations <- st_as_sf(locations, coords = c("longitude", "latitude"))

new_locations %>%
  ggplot() +
  geom_sf(size = 0.5, color = "#009FDA", alpha = 0.1) +
  theme_void() +
  labs(title = "Distribution of World Bank IBRD-IDA projects",
       subtitle = "Projects are mostly concentrated in India, Southeast China, Asia-Pacific, and coastli",
       caption = "\nData Source: AidData | World Bank Geocoded Research Release")
```

Distribution of World Bank IBRD–IDA projects

Projects are mostly concentrated in India, Southeast China, Asia–Pacific, and coastlines of Africa and Latin America



Data Source: AidData | World Bank Geocoded Research Release

```
ggsave(filename = "map_locations.png")
```

We obtain this map by plotting the longitude and latitude of each project to get an accurate distribution of World Bank project commitments from 1998 to 2004. Although these projects are commonly shown on the map, most projects are located along the coasts and relatively wealthy countries and regions. Countries and regions that need aid and technical assistance are missing from the map. There is an imperative to penetrate deeper into these remote and harsh areas, instead of picking “low-hanging” fruit. We suggest that the World Bank should use technology and other tools to help these under-invested places thrive.