# exploratory_analysis

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

## Parsed with column specification:
## cols(
##   .default = col_character(),
##   country_id = col_integer(),
##   region_id = col_integer(),
##   implementing_agency_id = col_integer(),
##   implementing_subagency_id = col_integer(),
##   channel_category_id = col_integer(),
##   channel_subcategory_id = col_integer(),
##   channel_id = col_integer(),
##   dac_category_id = col_integer(),
##   dac_sector_code = col_integer(),
##   dac_purpose_code = col_integer(),
##   funding_agency_id = col_integer(),
##   assistance_category_id = col_integer(),
##   aid_type_group_id = col_integer(),
##   activity_id = col_integer(),
##   transaction_type_id = col_integer(),
##   current_amount = col_double(),
##   constant_amount = col_double(),
##   USG_sector_id = col_integer(),
##   submission_id = col_integer()
## )

## See spec(...) for full column specifications.

## Parsed with column specification:
## cols(
##   .default = col_character(),
##   ISO = col_integer(),
##   EVENT_ID_NO_CNTY = col_integer(),
##   YEAR = col_integer(),
##   TIME_PRECISION = col_integer(),
##   INTER1 = col_integer(),
```

```
##   INTER2 = col_integer(),
##   INTERACTION = col_integer(),
##   LATITUDE = col_double(),
##   LONGITUDE = col_double(),
##   GEO_PRECISION = col_integer(),
##   FATALITIES = col_integer(),
##   TIMESTAMP = col_integer()
## )
```

```
## See spec(...) for full column specifications.
```

```
## Parsed with column specification:
## cols(
##   .default = col_character(),
##   ISO = col_integer(),
##   EVENT_ID_NO_CNTY = col_integer(),
##   YEAR = col_integer(),
##   TIME_PRECISION = col_integer(),
##   INTER1 = col_integer(),
##   INTER2 = col_integer(),
##   INTERACTION = col_integer(),
##   LATITUDE = col_double(),
##   LONGITUDE = col_double(),
##   GEO_PRECISION = col_integer(),
##   FATALITIES = col_integer(),
##   TIMESTAMP = col_integer()
## )
```

```
## See spec(...) for full column specifications.
```

```
## Parsed with column specification:
## cols(
##   .default = col_character(),
##   ISO = col_integer(),
##   EVENT_ID_NO_CNTY = col_integer(),
##   YEAR = col_integer(),
##   TIME_PRECISION = col_integer(),
##   INTER1 = col_integer(),
##   INTER2 = col_integer(),
##   INTERACTION = col_integer(),
##   LATITUDE = col_double(),
##   LONGITUDE = col_double(),
##   GEO_PRECISION = col_integer(),
##   FATALITIES = col_integer(),
##   TIMESTAMP = col_integer()
## )
```

```
## See spec(...) for full column specifications.
```

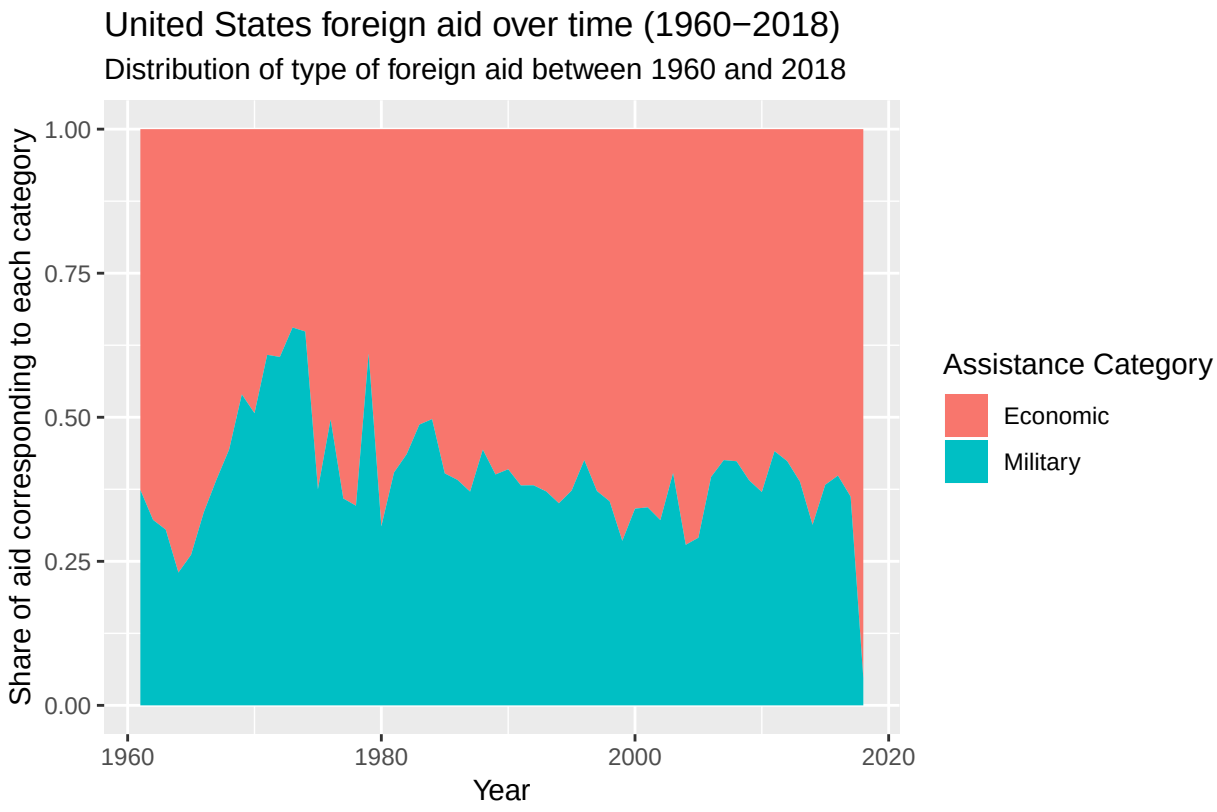
```
foreign_aid  %>%
  filter((region_name != "World") & (fiscal_year > 1960))  %>%
  group_by(fiscal_year, assistance_category_name) %>%
  summarise(total_raw = sum(constant_amount)) %>%
  mutate(share = total_raw / sum(total_raw)) %>% ggplot(aes(x=as.numeric(as.character(fiscal_year)), y =
  geom_area(position = "fill") +
  labs(
    x = "Year",
```

```
    y= "Share of aid corresponding to each category",
    title ="United States foreign aid over time (1960-2018)",
    subtitle = "Distribution of type of foreign aid between 1960 and 2018",
    caption =  "Source: USAID dataset available at https://explorer.usaid.gov/",
    fill = "Assistance Category")
```

```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```

```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```

```
## Warning: Removed 2 rows containing missing values (position_stack).
```

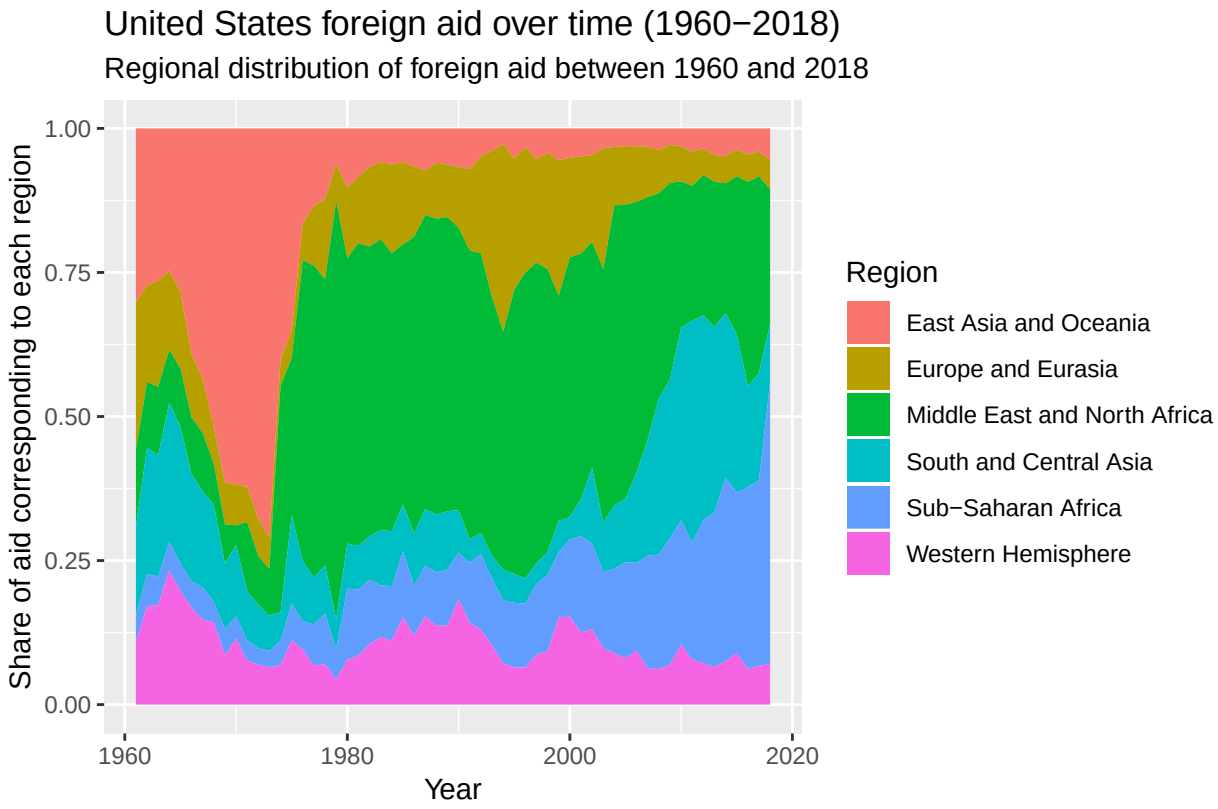



```
foreign_aid %>%
  filter((region_name != "World") & (fiscal_year > 1960))%>%
  group_by(fiscal_year, region_name) %>% summarise(total_raw = sum(constant_amount)) %>%
  mutate(share = total_raw / sum(total_raw)) %>%
  ggplot(aes(x=as.numeric(as.character(fiscal_year)), y = share, fill = region_name)) + geom_area(positi
  labs(
    x = "Year",
    y= "Share of aid corresponding to each region",
    title ="United States foreign aid over time (1960-2018)",
    subtitle = "Regional distribution of foreign aid between 1960 and 2018",
    caption =  "Source: USAID dataset available at https://explorer.usaid.gov/",
    fill = "Region")
```

```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```

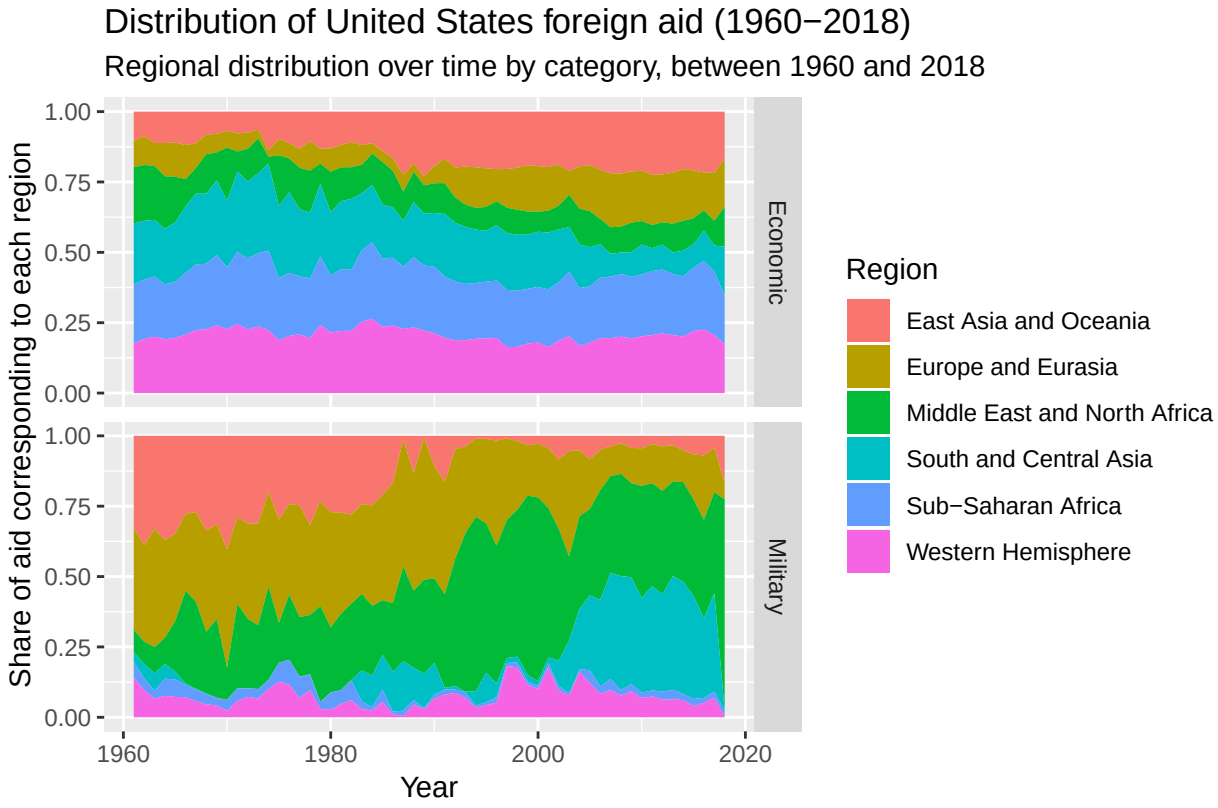
```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```

```
## Warning: Removed 6 rows containing missing values (position_stack).
```




```
foreign_aid %>%
  filter((region_name != "World") & (fiscal_year > 1960))%>%
  group_by(fiscal_year, region_name, assistance_category_name) %>%
  summarise(total_raw = sum(constant_amount)) %>%
  mutate(share = total_raw / sum(total_raw)) %>%
  ggplot(aes(x=as.numeric(as.character(fiscal_year)), y = share, fill = region_name, label = region_name
  geom_area(position = "fill") +
  facet_grid(assistance_category_name ~ .) +
  labs(
  x = "Year",
  y = "Share of aid corresponding to each region",
  title = "Distribution of United States foreign aid (1960-2018)",
  subtitle = "Regional distribution over time by category, between 1960 and 2018",
  caption = "Source: USAID dataset available at https://explorer.usaid.gov/",
   fill = "Region")
```

```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion

## Warning in FUN(X[[i]], ...): NAs introduced by coercion

## Warning: Removed 12 rows containing missing values (position_stack).
```




Source: USAID dataset available at https://explorer.usaid.gov/

```
  ggsave("distribution_aid.pdf", plot=last_plot(), device = "pdf", path="output/")
```

```
## Saving 6.5 x 4.5 in image
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
## Warning: Removed 12 rows containing missing values (position_stack).
```

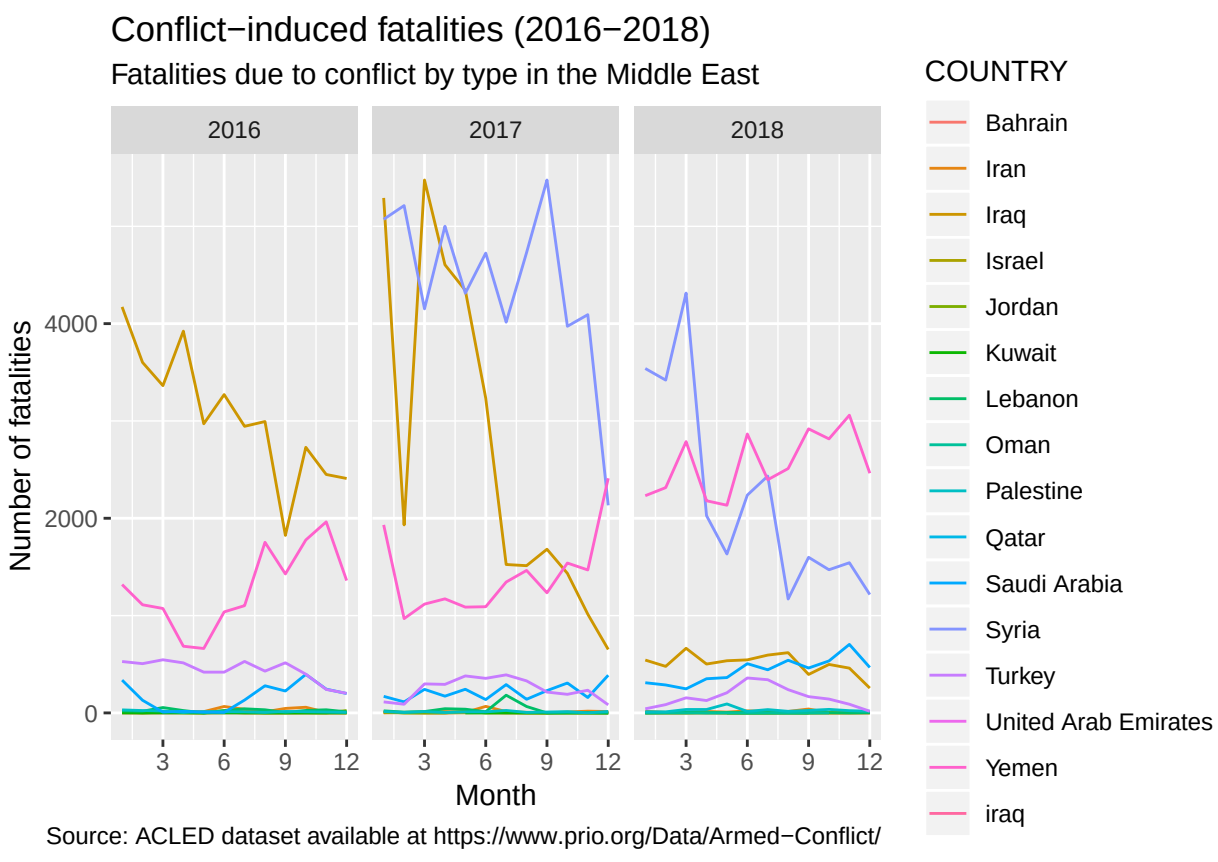
```
# Logarithmic

library(lubridate)
```

```
##
## Attaching package: 'lubridate'
## The following object is masked from 'package:base':
##
##     date
```

```
middle_east %>%
  mutate(DATE_NUM = as.Date(EVENT_DATE, format = "%d-%B-%Y"))  %>%
  mutate(MONTH = month(DATE_NUM)) %>%
  filter(YEAR < 2019)  %>%
  group_by(COUNTRY, MONTH, YEAR) %>% summarize(total_death = sum(FATALITIES)) %>%
  ggplot(mapping = aes(x = MONTH, y = total_death, color = COUNTRY)) +
  geom_line() +
  facet_grid(. ~YEAR  ) + scale_x_continuous(breaks=c(3,6,9,12)) +
  labs(
```

```
  x = "Month",
  y = "Number of fatalities",
  title = "Conflict-induced fatalities (2016-2018)",
  subtitle = "Fatalities due to conflict by type in the Middle East",
  caption = "Source: ACLED dataset available at https://www.prio.org/Data/Armed-Conflict/",
  fill = "Country"
)
```

```
## Warning in strptime(x, format, tz = "GMT"): unknown timezone 'zone/tz/
## 2018i.1.0/zoneinfo/America/Chicago'
```

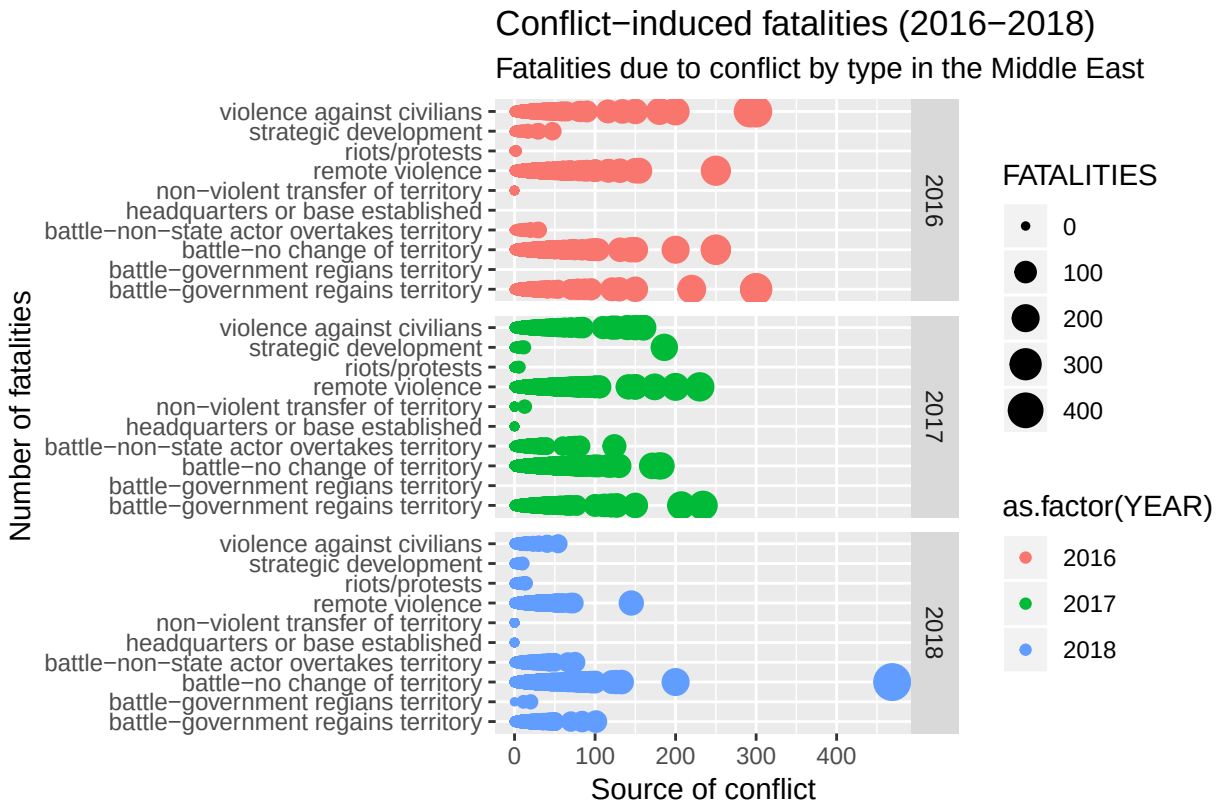



```
  ggsave("fatalities_middleeast_1.pdf", plot=last_plot(), device = "pdf", path="output/")
```

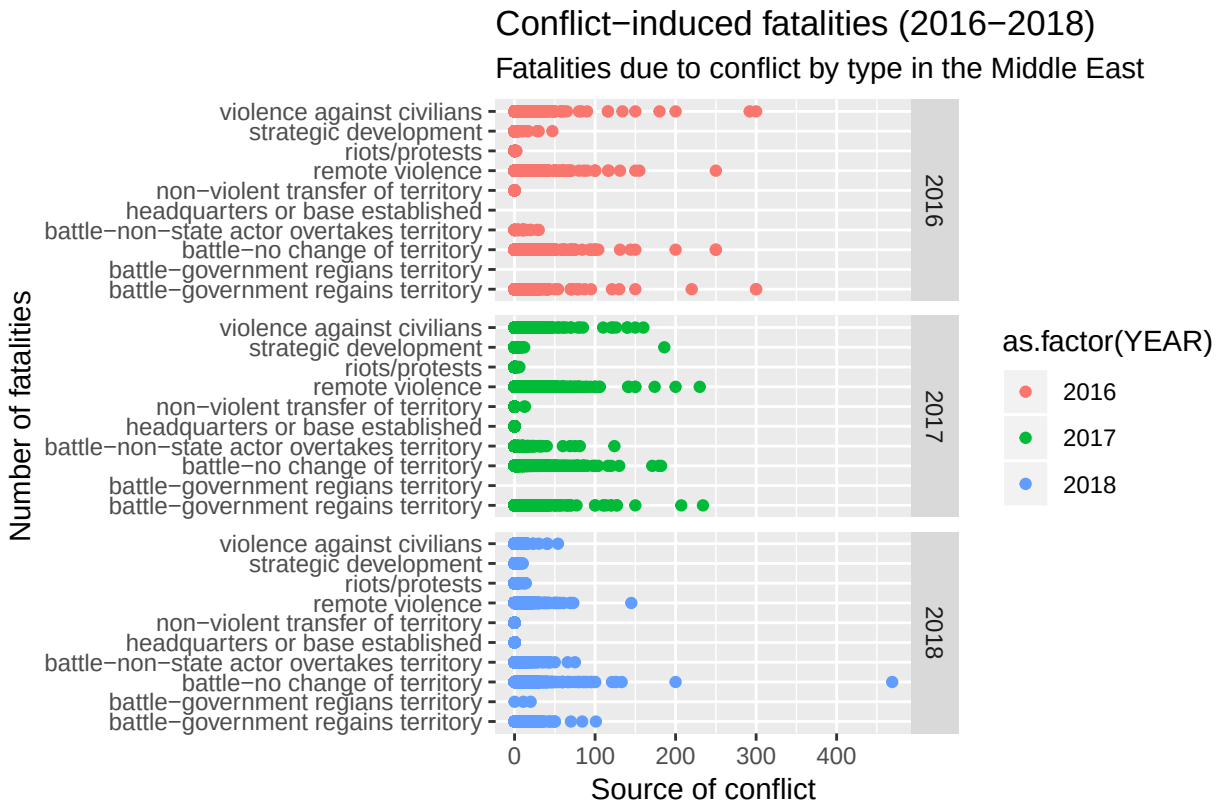
```
## Saving 6.5 x 4.5 in image
```

```
middle_east %>%
  mutate(DATE_NUM = as.Date(EVENT_TYPE, format = "%d-%B-%Y"))  %>%
  mutate(MONTH = month(DATE_NUM)) %>%
  mutate(TYPE = tolower(EVENT_TYPE)) %>%
  filter(YEAR < 2019) %>%
  ggplot(mapping = aes(x = TYPE, y = FATALITIES, color = as.factor(YEAR))) +
  geom_point(aes(size=FATALITIES)) +
  coord_flip()  +
  facet_grid( YEAR ~. ) +
labs(
x = "Number of fatalities",
y = "Source of conflict",
title = "Conflict-induced fatalities (2016-2018)",
```

```
subtitle = "Fatalities due to conflict by type in the Middle East",
caption = "Source: ACLED dataset available at https://www.prio.org/Data/Armed-Conflict/"
)
```

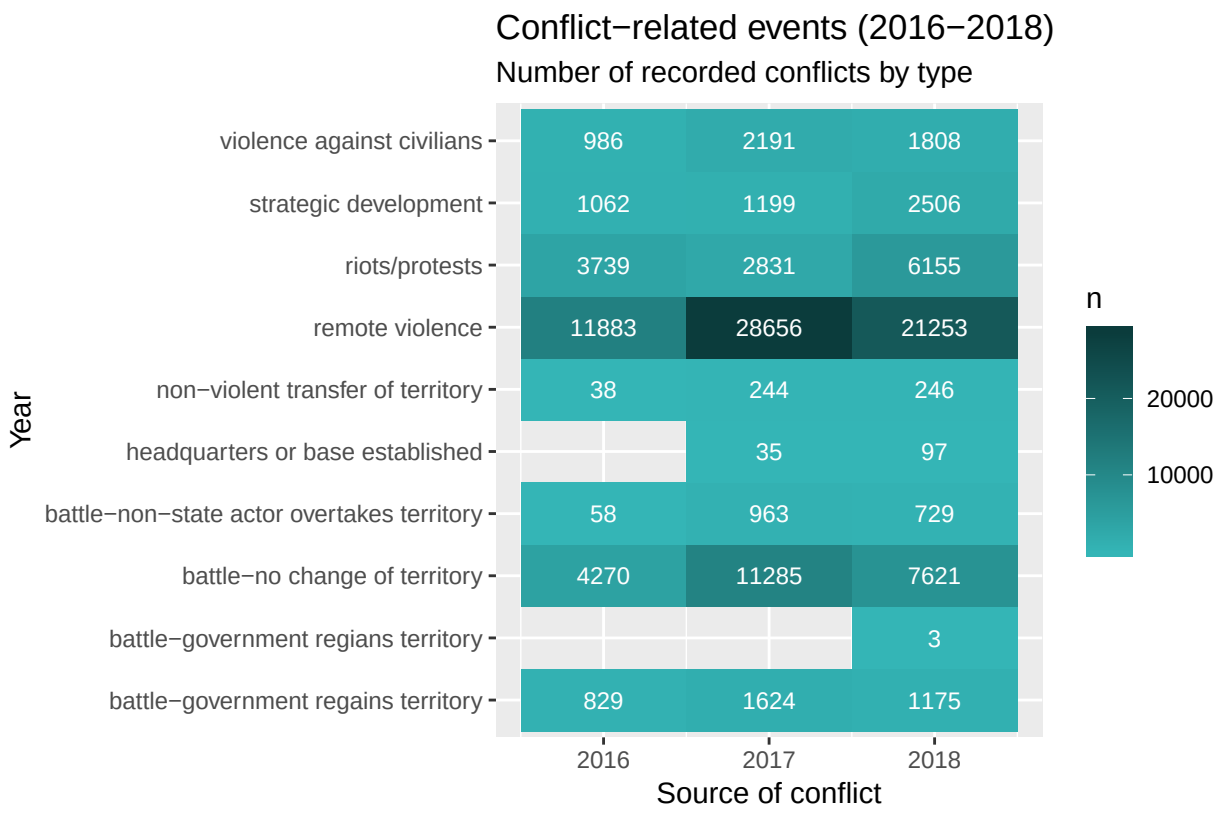



```
middle_east %>%
  mutate(DATE_NUM = as.Date(EVENT_TYPE, format = "%d-%B-%Y"))  %>%
  mutate(MONTH = month(DATE_NUM)) %>%
  mutate(TYPE = tolower(EVENT_TYPE)) %>%
  filter(YEAR < 2019) %>%
  ggplot(mapping = aes(x = TYPE, y = FATALITIES, color = as.factor(YEAR))) +
  geom_point() +
  coord_flip()  +
  facet_grid( YEAR ~. ) +
labs(
x = "Number of fatalities",
y = "Source of conflict",
title = "Conflict-induced fatalities (2016-2018)",
subtitle = "Fatalities due to conflict by type in the Middle East",
caption = "Source: ACLED dataset available at https://www.prio.org/Data/Armed-Conflict/"
)
```



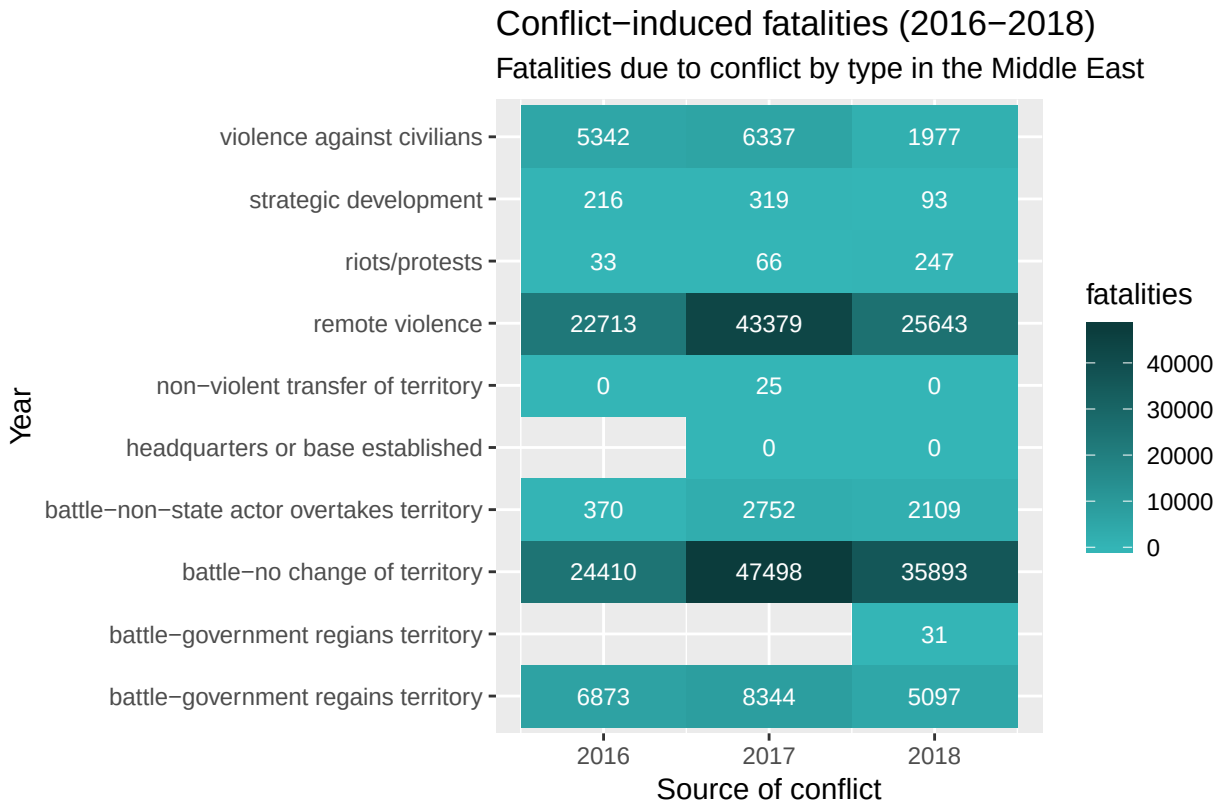

Source: ACLED dataset available at https://www.prio.org/Data/Armed−Conflict/

```
library(RColorBrewer)
middle_east %>%
  mutate(DATE_NUM = as.Date(EVENT_TYPE, format = "%d-%B-%Y"))  %>%
  mutate(MONTH = month(DATE_NUM)) %>%
  mutate(TYPE = tolower(EVENT_TYPE)) %>%
  filter(YEAR < 2019) %>%
  group_by(TYPE, YEAR) %>% summarize(n= n()) %>%
  ggplot(middle_east, mapping = aes(x = TYPE, y = YEAR, fill = n)) +
  geom_tile() +
  coord_flip() +
  scale_fill_gradient(low = "#34B5B6", high = "#0B3C3C") +
  geom_text(aes(label=n), colour = "white", size= 3) +
  labs(
  x ="Year",
  y = "Source of conflict",
  title = "Conflict-related events (2016-2018)",
  subtitle = "Number of recorded conflicts by type",
  caption = "Source: ACLED dataset available at https://www.prio.org/Data/Armed-Conflict/")
```




Source: ACLED dataset available at https://www.prio.org/Data/Armed−Conflict/

```
middle_east %>%
  mutate(DATE_NUM = as.Date(EVENT_TYPE, format = "%d-%B-%Y"))  %>%
  mutate(MONTH = month(DATE_NUM)) %>%
  mutate(TYPE = tolower(EVENT_TYPE)) %>%
  filter(YEAR < 2019) %>%
  group_by(TYPE, YEAR) %>% summarize(fatalities = sum(FATALITIES)) %>%
  ggplot(middle_east, mapping = aes(x = TYPE, y = YEAR, fill = fatalities)) +
  geom_tile() +
  coord_flip() +
  geom_text(aes(label=fatalities), colour = "white", size= 3) +
  scale_fill_gradient(low = "#34B5B6", high = "#0B3C3C") +
  labs(
  x ="Year",
  y = "Source of conflict",
  title = "Conflict-induced fatalities (2016-2018)",
  subtitle = "Fatalities due to conflict by type in the Middle East",
  caption = "Source: ACLED dataset available at https://www.prio.org/Data/Armed-Conflict/")
```




Source: ACLED dataset available at https://www.prio.org/Data/Armed−Conflict/

```
ggsave("fatalities_middleeast_2.pdf", plot=last_plot(), device = "pdf", path="output/")
```

```
## Saving 6.5 x 4.5 in image
```