Lab 2 - Descriptive Statistics and Basic Mapping

GIS III Spring 2020 - Erin Abbott

Contents

Submission by Sunday 4/19

Data Manipulation

- Load Libraries

```
library(sf)
```

```
## Linking to GEOS 3.5.1, GDAL 2.2.2, PROJ 4.9.2
```

```
library(raster)
```

```
## Loading required package: sp
```

```
library(dplyr)
```

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

```
## The following objects are masked from 'package:raster':
##
##     intersect, select, union
```

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

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

```
library(spData)
```

```
## To access larger datasets in this package, install the spDataLarge
## package with: `install.packages('spDataLarge',
## repos='https://nowosad.github.io/drat/', type='source')`
```

```
library(tidyverse)
```

```
## -- Attaching packages ----------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.0     v purrr   0.3.4
## v tibble  3.0.0     v stringr 1.4.0
```

```
## v tidyr   1.0.2     v forcats 0.5.0
## v readr   1.3.1
```

```
## -- Conflicts --------------------------- tidyverse_conflicts() --
## x tidyr::extract() masks raster::extract()
## x dplyr::filter()  masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## x dplyr::select()  masks raster::select()
```

```
library(RColorBrewer)
library(ggplot2)
```

- Data sources:
  - Coffee production dataset: by Roger Bivand, Jakub Nowosad, Robin Lovelace. Available on the GitHub page.
  - World dataset: by Roger Bivand, Jakub Nowosad, Robin Lovelace. Available on the GitHub page.

```
#load coffee production data
cof_data <- coffee_data
class(cof_data)
```

```
## [1] "tbl_df"     "tbl"        "data.frame"
```

```
#load world data
world <- world
class(world)
```

```
## [1] "sf"         "tbl_df"     "tbl"        "data.frame"
```

```
#join coffee and world data
coffee <- left_join(world, cof_data)
```

```
## Joining, by = "name_long"
```

```
head(coffee)
```

```
## Simple feature collection with 6 features and 12 fields
## geometry type:  MULTIPOLYGON
## dimension:      XY
## bbox:           xmin: -180 ymin: -18.28799 xmax: 180 ymax: 83.23324
## CRS:            EPSG:4326
## # A tibble: 6 x 13
##   iso_a2 name_long continent region_un subregion type  area_km2     pop lifeExp
##   <chr>  <chr>     <chr>     <chr>     <chr>     <chr>    <dbl>   <dbl>   <dbl>
## 1 FJ     Fiji      Oceania   Oceania   Melanesia Sove~   1.93e4  8.86e5    70.0
## 2 TZ     Tanzania  Africa    Africa    Eastern ~ Sove~   9.33e5  5.22e7    64.2
## 3 EH     Western ~ Africa    Africa    Northern~ Inde~   9.63e4 NA         NA
## 4 CA     Canada    North Am~ Americas  Northern~ Sove~   1.00e7  3.55e7    82.0
## 5 US     United S~ North Am~ Americas  Northern~ Coun~   9.51e6  3.19e8    78.8
## 6 KZ     Kazakhst~ Asia      Asia      Central ~ Sove~   2.73e6  1.73e7    71.6
## # ... with 4 more variables: gdpPercap <dbl>, geom <MULTIPOLYGON [°]>,
## #   coffee_production_2016 <int>, coffee_production_2017 <int>
```

```
# create coffee production difference variable
coffee1 <- coffee %>% mutate(prod_dif = coffee_production_2017 - coffee_production_2016)
coffee1
```

```
## Simple feature collection with 177 features and 13 fields
## geometry type:  MULTIPOLYGON
## dimension:      XY
```

```
## bbox:           xmin: -180 ymin: -90 xmax: 180 ymax: 83.64513
## CRS:            EPSG:4326
## # A tibble: 177 x 14
##    iso_a2 name_long continent region_un subregion type  area_km2     pop lifeExp
##  * <chr>  <chr>     <chr>     <chr>     <chr>     <chr>    <dbl>   <dbl>   <dbl>
##  1 FJ     Fiji      Oceania   Oceania   Melanesia Sove~   1.93e4  8.86e5    70.0
##  2 TZ     Tanzania  Africa    Africa    Eastern ~ Sove~   9.33e5  5.22e7    64.2
##  3 EH     Western ~ Africa    Africa    Northern~ Inde~   9.63e4 NA         NA
##  4 CA     Canada    North Am~ Americas  Northern~ Sove~   1.00e7  3.55e7    82.0
##  5 US     United S~ North Am~ Americas  Northern~ Coun~   9.51e6  3.19e8    78.8
##  6 KZ     Kazakhst~ Asia      Asia      Central ~ Sove~   2.73e6  1.73e7    71.6
##  7 UZ     Uzbekist~ Asia      Asia      Central ~ Sove~   4.61e5  3.08e7    71.0
##  8 PG     Papua Ne~ Oceania   Oceania   Melanesia Sove~   4.65e5  7.76e6    65.2
##  9 ID     Indonesia Asia      Asia      South-Ea~ Sove~   1.82e6  2.55e8    68.9
## 10 AR     Argentina South Am~ Americas  South Am~ Sove~   2.78e6  4.30e7    76.3
## # ... with 167 more rows, and 5 more variables: gdpPercap <dbl>,
## #   geom <MULTIPOLYGON [°]>, coffee_production_2016 <int>,
## #   coffee_production_2017 <int>, prod_dif <int>
```

- Summary Stats:

```
summary(coffee1$coffee_production_2016)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##     1.0     3.0    38.0   275.1   149.5  3277.0     141
```

```
summary(coffee1$coffee_production_2017)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##    1.00    3.25   43.50  264.15  206.25 2786.00     143
```

```
summary(coffee1$prod_dif)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
## -491.00  -12.00    1.00  -27.82    6.00  113.00     144
```

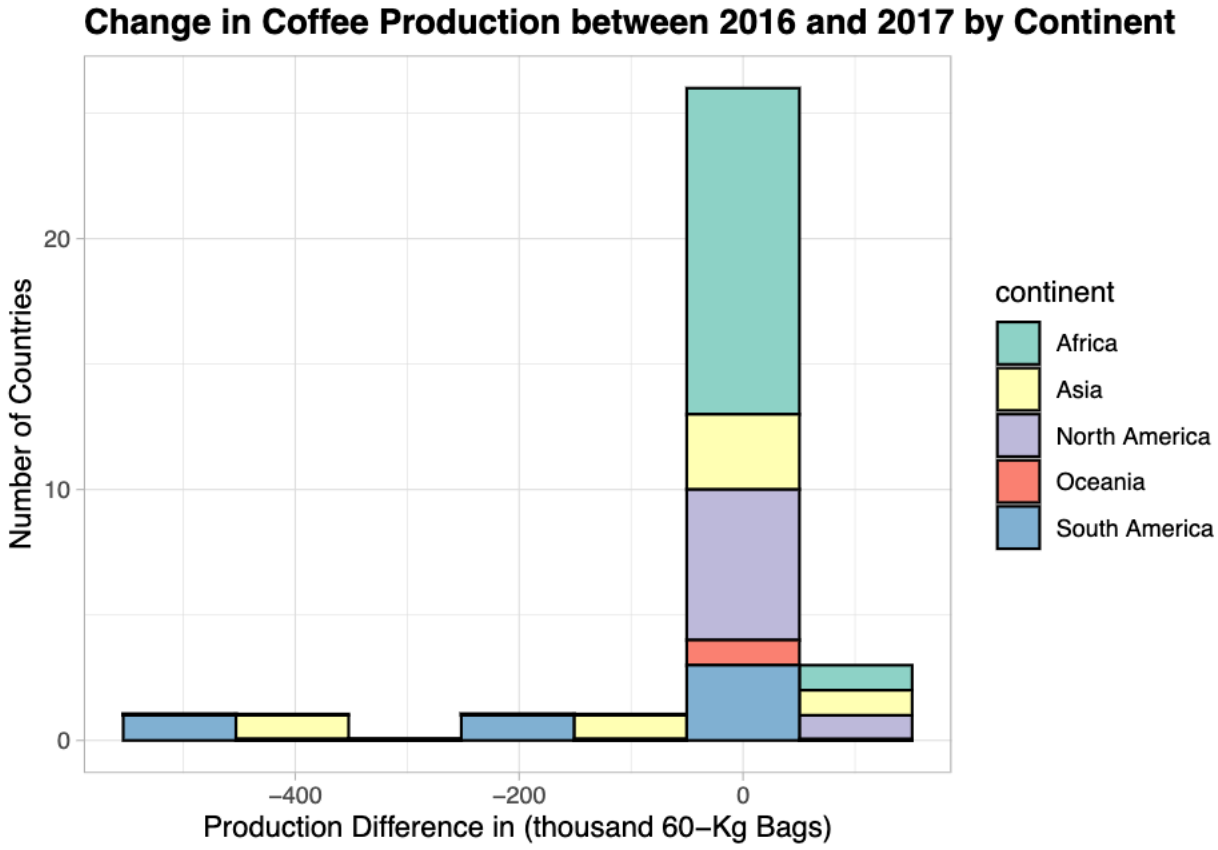
```
summary(coffee$area_km2)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
##     2417    46185   185004   832558   621860 17018507
```

Non-Spatial Data Visualization

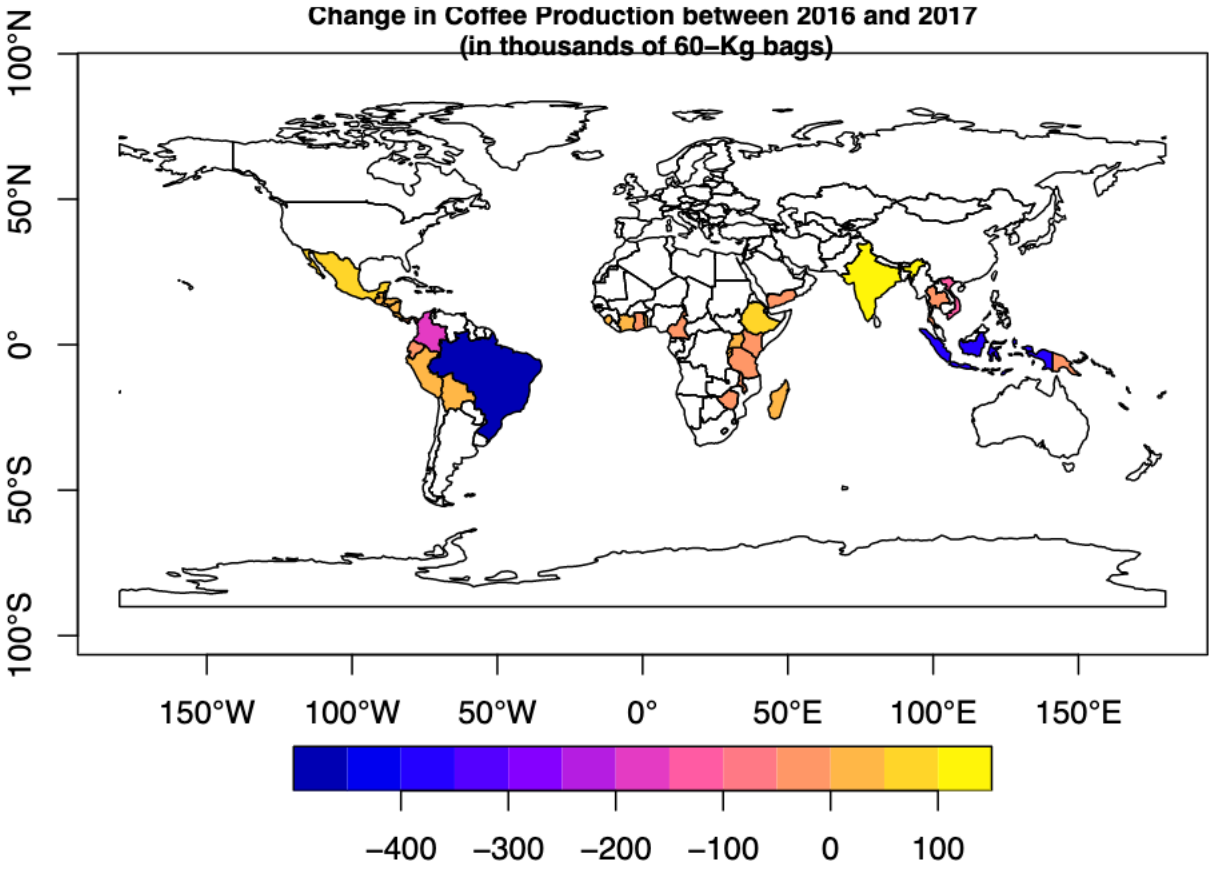
```
ggplot(data=coffee1,aes(prod_dif)) +
  geom_histogram(aes(fill=continent), color="black",bins=7) +
  theme_light() +
  scale_fill_brewer(palette="Set3") +
  ggtitle("Change in Coffee Production between 2016 and 2017 by Continent") +
  xlab("Production Difference in (thousand 60-Kg Bags)") +
  ylab("Number of Countries") +
  theme(plot.title = element_text(face = "bold"))
```

```
## Warning: Removed 144 rows containing non-finite values (stat_bin).
```



Spatial Data Visualization

```
plot(coffee1["prod_dif"], main="Change in Coffee Production between 2016 and 2017\n (in thousands of 60
```



- Rendering in html
- Uploaded to GitHub