Map of Server Logins

Brian High

12/03/2015

Set up our environment

Set document rendering options.

```
# Configure `knitr` options.
library(knitr)
opts_chunk$set(tidy=FALSE, cache=FALSE)
```

Install packages if needed.

```
# Install packages (if necessary)
for (pkg in c("dplyr", "pander", "ggmap")) {
    if (! suppressWarnings(require(pkg, character.only=TRUE)) ) {
        install.packages(pkg, repos="http://cran.fhcrc.org", dependencies=TRUE)
        if (! suppressWarnings(require(pkg, character.only=TRUE)) ) {
            stop(paste0(c("Can't load package: ", pkg, "!"), collapse = ""))
        }
    }
}
```

Read the data

```
cities <- read.csv("cities.csv")
```

Take a look at the data

See which cities in which countries we are trying to map. List the top cities.

```
# Top-ten cities by login count
cities[,1:4] %>% head(10) %>% pandoc.table(style='rmarkdown')
```

city	region	country	logins
Seattle	WA	US	746
Port Orchard	WA	US	45
Cedar Crest	NM	US	30
Duvall	WA	US	28
Kent	WA	US	26
Arcadia	CA	US	25
Portland	OR	US	16
Redmond	WA	US	14
Las Vegas	NV	US	10
Salt Lake City	UT	US	10

Most logins are from Seattle, but what is the percentage of Seattle logins?

```
# Percent of logins from Seattle
percent_seattle <- round(
    sum(cities[cities$city == "Seattle" & cities$region == "WA", "logins"]) /
    sum(cities[, "logins"]) * 100, 1)
percent_seattle
```

```
## [1] 70.3
```

List the top countries.

```
# Top-ten countries by login count
cities %>% group_by(country) %>%
    summarize(logins=sum(logins)) %>% arrange(desc(logins)) -> countries
countries %>% head(10) %>% pandoc.table(style='rmarkdown')
```

country	logins
US	1051
CA	6
HN	3
ZA	1

Find percent of logins from the US.

```
# Percent of logins from US
percent_us <- round(
    sum(cities[cities$country == "US", "logins"]) /
    sum(cities[, "logins"]) * 100, 1)
percent_us
```

```
## [1] 99.1
```

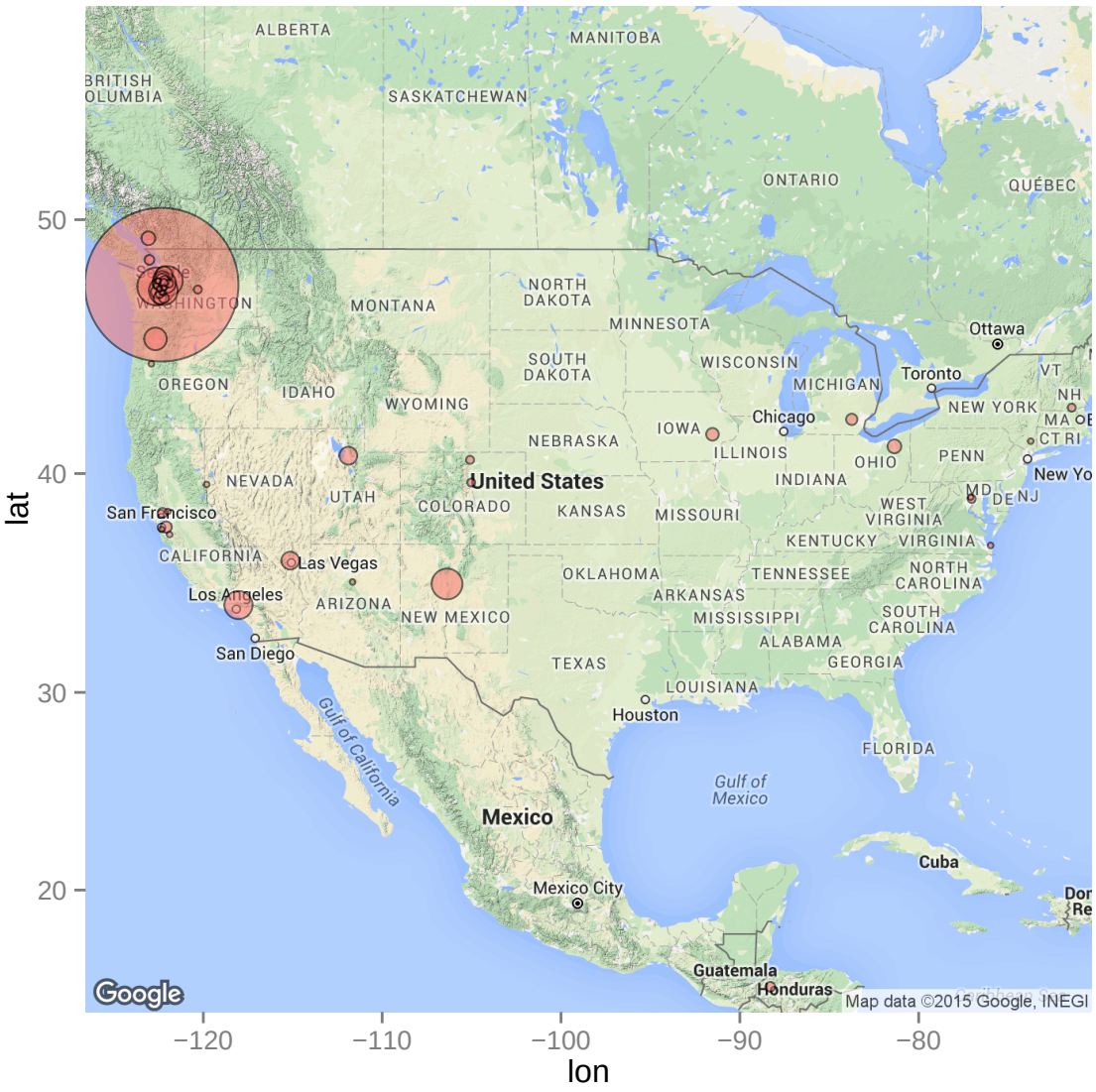
So, 70.3 % of the logins are from Seattle and 99.1 % are from the US.

Make the map

We can center the map on the center of the US (Kansas) and zoom so that we pick up some of the locations in the rest of North America and Central America.

```
# Center the map on Kansas
kansas <- geocode("Kansas")
google_basemap <- get_map(location = c(lon = kansas$lon, lat = kansas$lat),
                          zoom = 4, scale=2)

ggmap(google_basemap) + geom_point(data = cities,
                                   aes(x = longitude, y = latitude,
                                fill = "red", alpha = 0.8),
                                size = (cities$logins)^.5, shape = 21) +
    guides(fill=FALSE, alpha=FALSE, size=FALSE)
```



Map and geocoding sources:

- *ggmap*. version 2.5.2. David Kahle and Hadley Wickham. Web 3 Dec. 2015.
- *Google Maps*. Google, 3 Dec. 2015. Web 3 Dec. 2015.