Exploring Refrigerated Truck Dataset

Import Data:

```
library(ggplot2)
library(tidyverse)
```

```
##   Attaching packages                                           tidyverse 1.3.0
##   tibble  2.1.3      dplyr   0.8.3
##   tidyr   1.0.0      stringr 1.4.0
##   readr   1.3.1      forcats 0.4.0
##   purrr   0.3.3
##   Conflicts                                                    tidyverse_conflicts()
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(rgdal)
```

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

```
## rgdal: version: 1.4-8, (SVN revision 845)
##  Geospatial Data Abstraction Library extensions to R successfully loaded
##  Loaded GDAL runtime: GDAL 2.4.2, released 2019/06/28
##  Path to GDAL shared files: /Library/Frameworks/R.framework/Versions/3.6/Resources/library/rgdal/gdal
##  GDAL binary built with GEOS: FALSE
##  Loaded PROJ.4 runtime: Rel. 5.2.0, September 15th, 2018, [PJ_VERSION: 520]
##  Path to PROJ.4 shared files: /Library/Frameworks/R.framework/Versions/3.6/Resources/library/rgdal/pr
##  Linking to sp version: 1.3-2
```

```
setwd("/Users/JacquelineLincroft/Documents/INSH2102 Bostonography/Assignments/FinalProject/")
truckdf <- read.table(file="Ref_Trucks.csv", header=TRUE, sep=",")

# see which destination cities we could choose from
unique(truckdf$Destination)
```

```
##  [1] ATLANTA       CHICAGO       DALLAS        DENVER        NEW YORK
##  [6] LOS ANGELES   MONTREAL, QUE BOSTON        PHILADELPHIA  MIAMI
## [11] SEATTLE       BALTIMORE
## 12 Levels: ATLANTA BALTIMORE BOSTON CHICAGO DALLAS DENVER LOS ANGELES ... SEATTLE
```

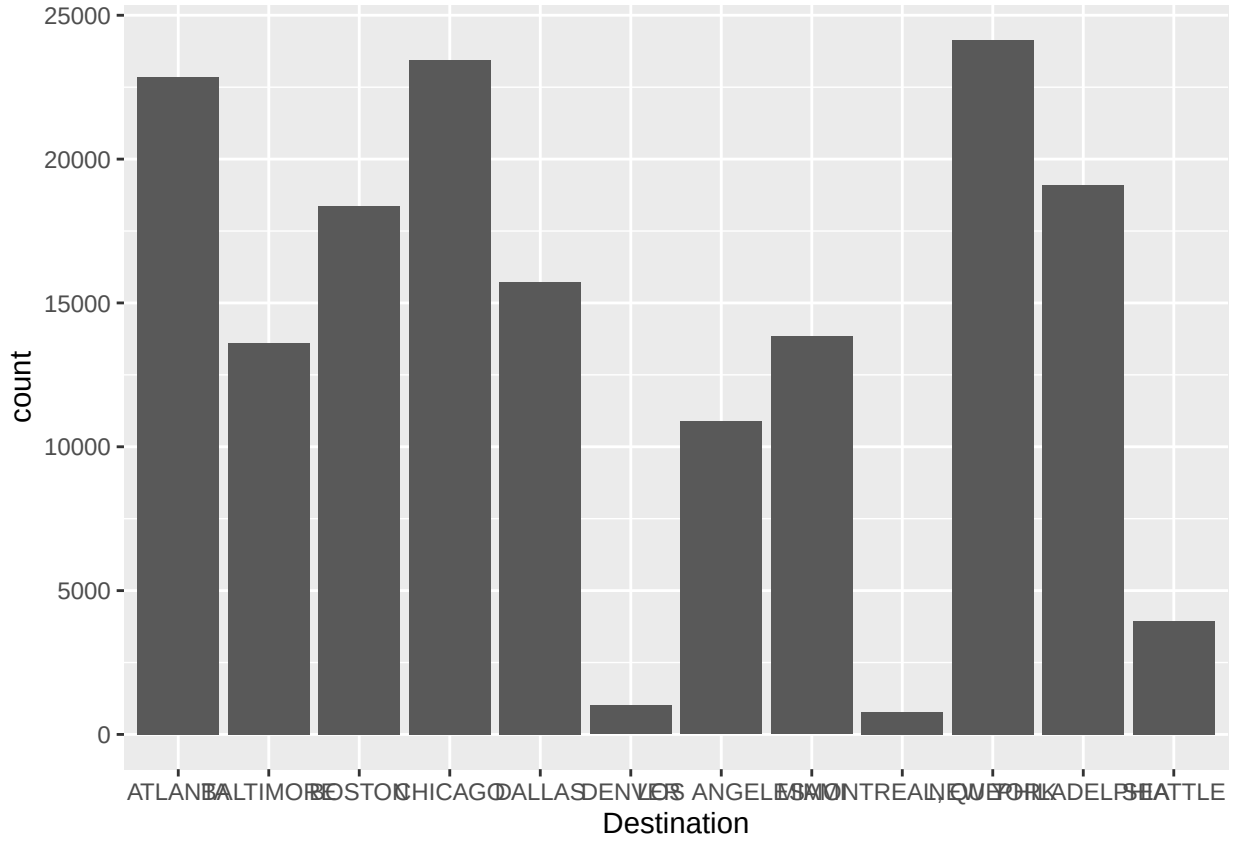
```
# separate out the data for boston
bostondf <- truckdf[truckdf$Destination == "BOSTON", ]
bostondf <- bostondf[order(bostondf$Origin), ]
boston_mean <- mean(bostondf$Distance, na.rm=TRUE)

# separate out the data for LA
ladf <- truckdf[truckdf$Destination == "LOS ANGELES", ]
la_mean <- mean(ladf$Distance, na.rm=TRUE)
```

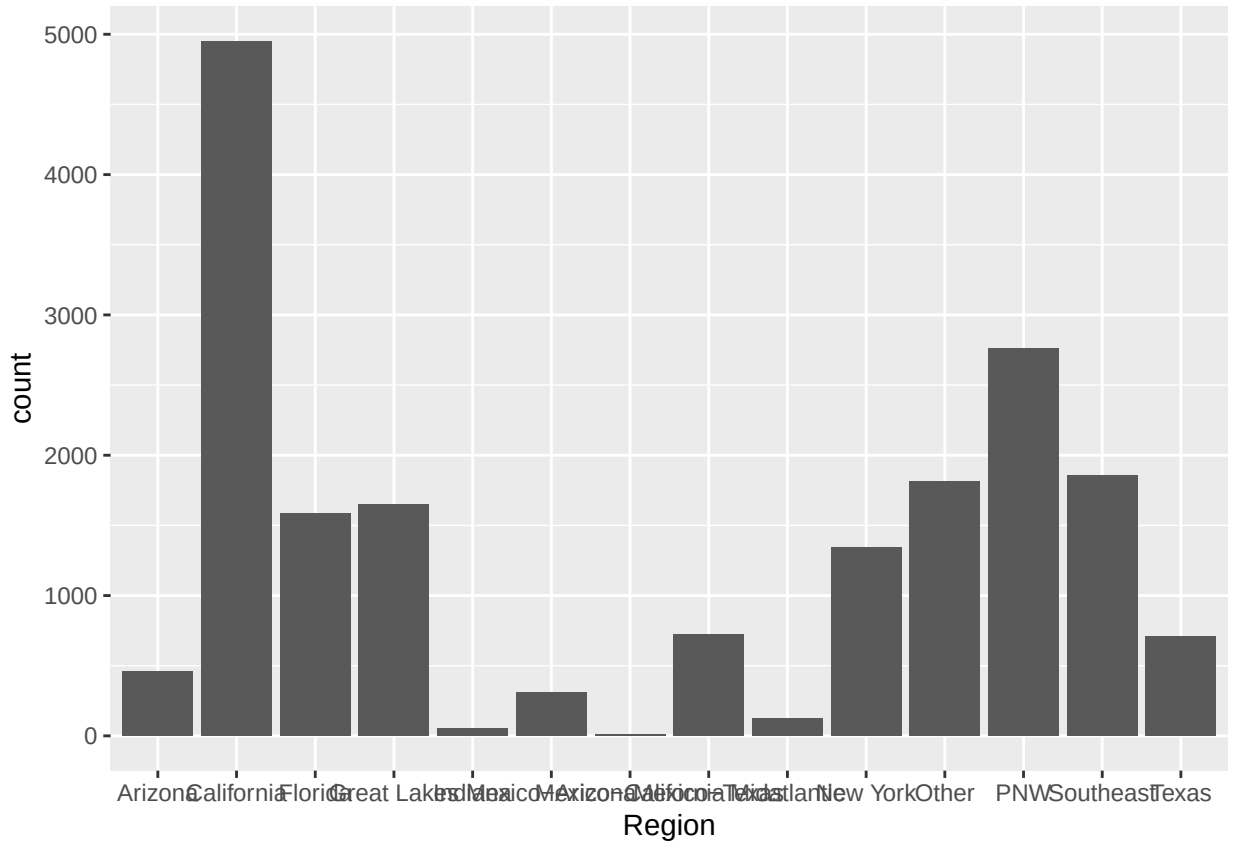
```r
# separate out data for seattle
seattledf <- truckdf[truckdf$Destination == "SEATTLE", ]
seattle_mean <- mean(seattledf$Distance, na.rm=TRUE)
```

Trying Out Some Visualizations:

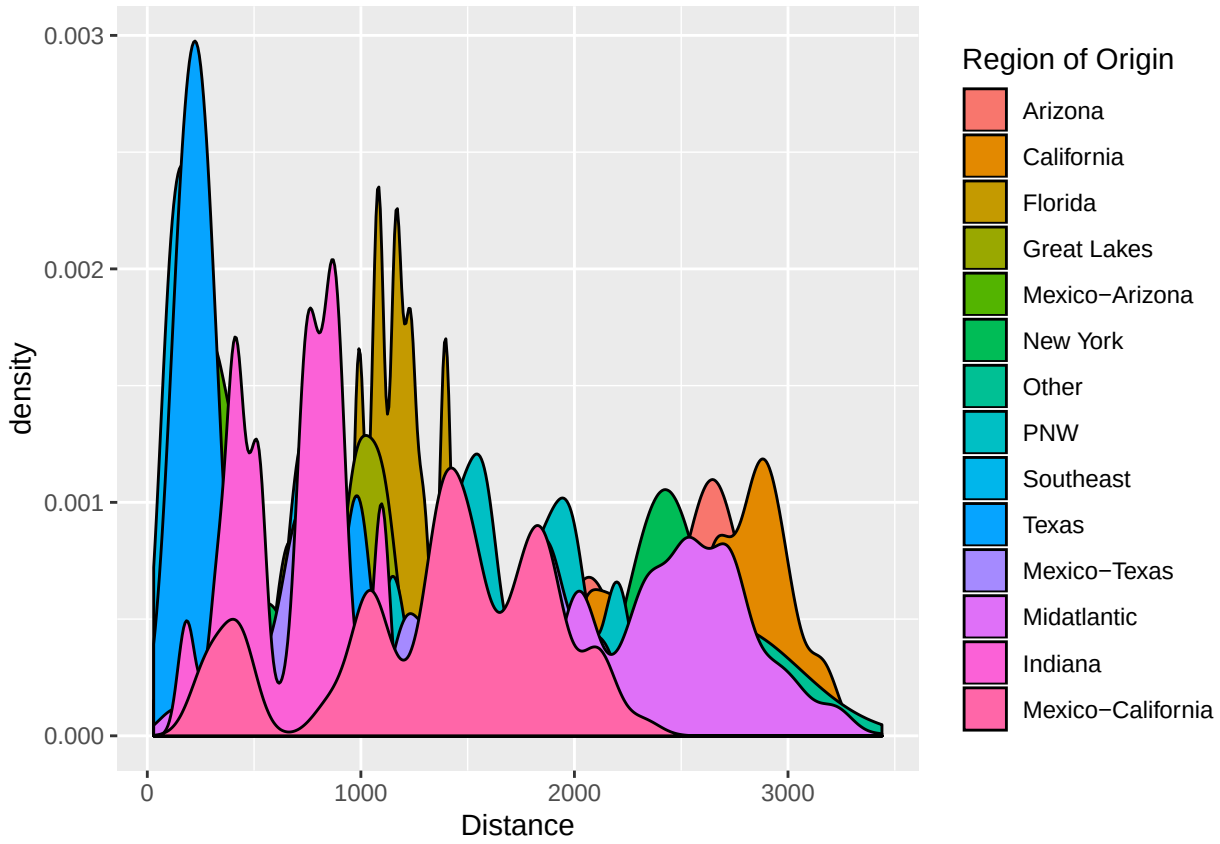
```r
# how much data do we have for each of the destination cities?
hist_destination <- ggplot(data=truckdf, aes(x=Destination)) + geom_histogram(stat="count")
hist_destination
```



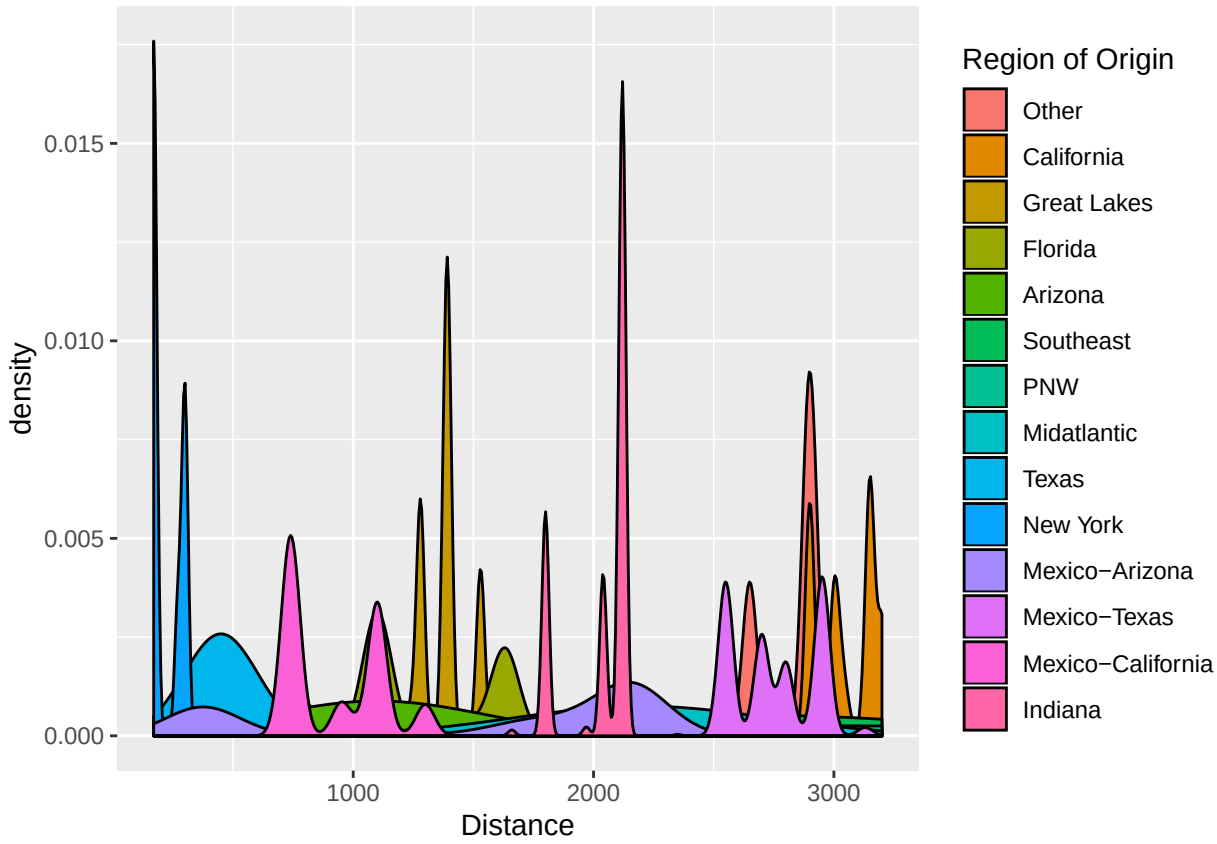

```r
# which regions does boston's food come from
hist_boston_region <- ggplot(data=bostondf, aes(x=Region)) + geom_histogram(stat="count")
hist_boston_region
```




```
# which regions send food the furthest distances?
density_region<-ggplot(data=truckdf, aes(x=Distance, fill=Region)) + scale_fill_hue(name="Region of Orig
density_region
```




```
# which regions send food the furthest distances (to boston)?
density_boston_region<-ggplot(data=bostondf, aes(x=Distance, fill=Region)) + scale_fill_hue(name="Regio
density_boston_region
```



From these visualizations, it looks like food from Texas and New York generally travel the shortest distances, and food from Mexico/Arizona/Southeast have a lot of variability in the distance travelled, although this may be due to the specificity/generality of the way these regions are categorized.

Trying Out Some Geospatial Visualizations:

```
mapdf <- read.csv(file="map.csv", sep=",")
uscitiesdf <- read.csv(file="uscities.csv", sep=",")

# List the regions of origin in the original dataset
unique(truckdf$Region)
```

```
##  [1] Arizona           California        Florida           Great Lakes
##  [5] Mexico-Arizona    New York          Other             PNW
##  [9] Southeast         Texas             Mexico-Texas      Midatlantic
## [13] Indiana           Mexico-California
## 14 Levels: Arizona California Florida Great Lakes Indiana ... Texas
```

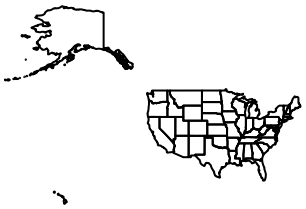
```
# We need to map these 14 regions onto more specific states so that we can use geospatial data
# For now, I'm picking a specific state to represent a general area when necessary (ex. Michigan=Great

mapping = data.frame(Region=c("Arizona","California","Floria", "Great Lakes", "Mexico-Arizona", "New Yo
truckdf <- inner_join(truckdf, mapping, by=c("Region"))

library(geojsonio)
usa_spdf <- geojson_read("us-states.json",  what = "sp")
mexico_spdf <- geojson_read("mexico.json",  what = "sp")
```

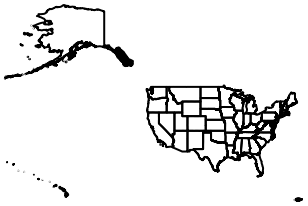
```
usa_github_spdf <- geojson_read("us_states_github.json",  what = "sp")

plot(usa_spdf)
```



```
plot(mexico_spdf)
```

```
plot(usa_github_spdf)
```



```
## this is still in progress, but we're working on getting a map of US+Mexico set up so we can visualiz
setwd("/Users/JacquelineLincroft/Documents/INSH2102 Bostonography/Assignments/FinalProject/PoliticalBou
#pb_spdf <- readOGR(
#  dsn=setwd("/Users/JacquelineLincroft/Documents/INSH2102 Bostonography/Assignments/FinalProject/Polit
#  layer="boundary_l_v2",
#  verbose=FALSE)
```