

Kairavi Chahal
36-315 B Exam 2
Professor Thomas
April 12, 2013

Question 0

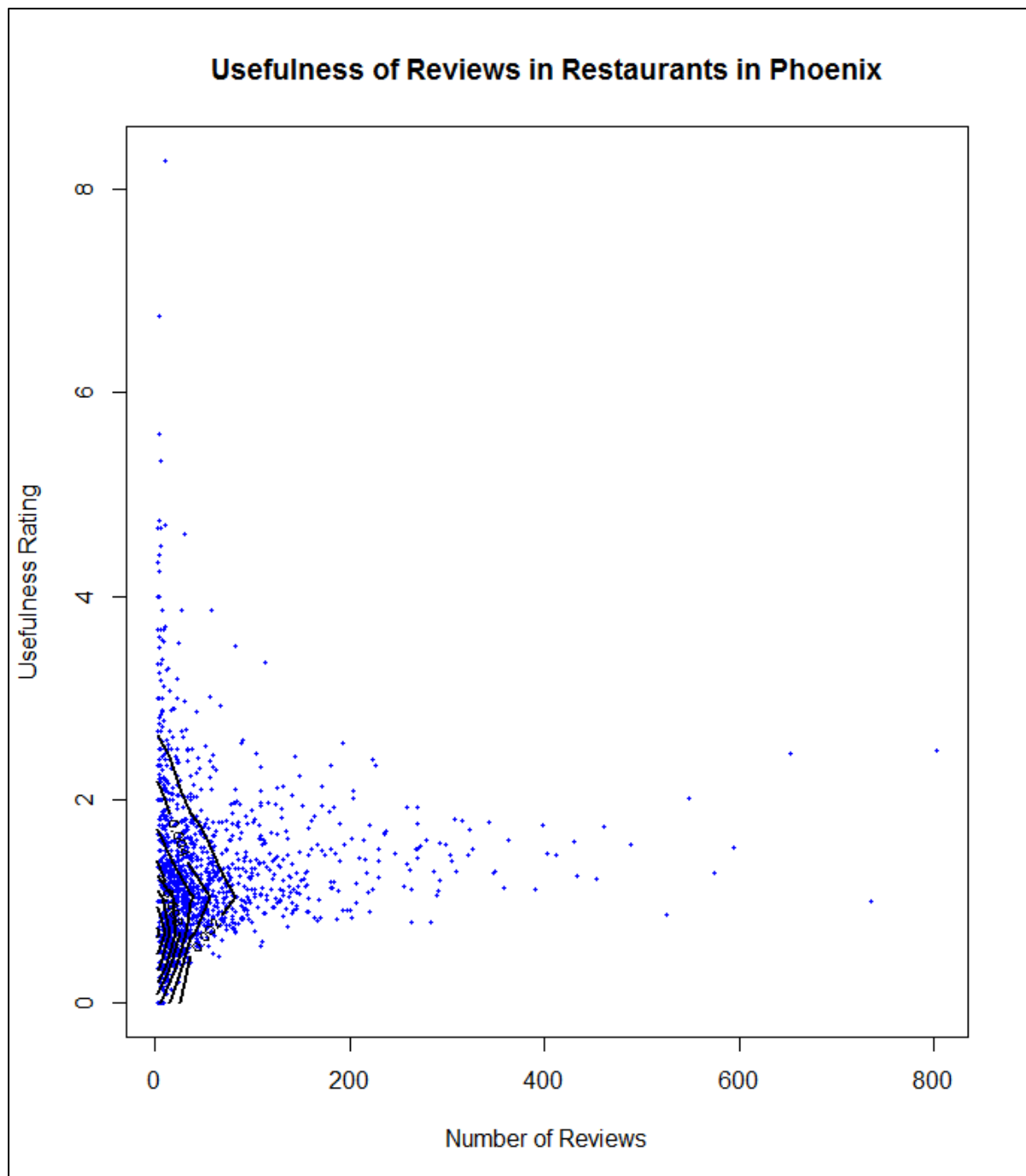
Do you think it will be sunny tomorrow?

You bet your bottom dollar, the sun will come out tomorrow.

Question 2

I used a contour plot with the points overlaid because it is easiest to determine the mode of density from this plot. Alternatively, I could have used a heat map; however I prefer the contour map when plotting points over it, since it is much clearer.

```
reviews.useful = kde2d(review_count, useful)
plot(review_count, useful, pch=16, cex=0.4, col=4,
      main="Usefulness of Reviews in Restaurants in Phoenix",
      xlab="Number of Reviews", ylab="Usefulness Rating")
contour(reviews.useful, add=T)
```

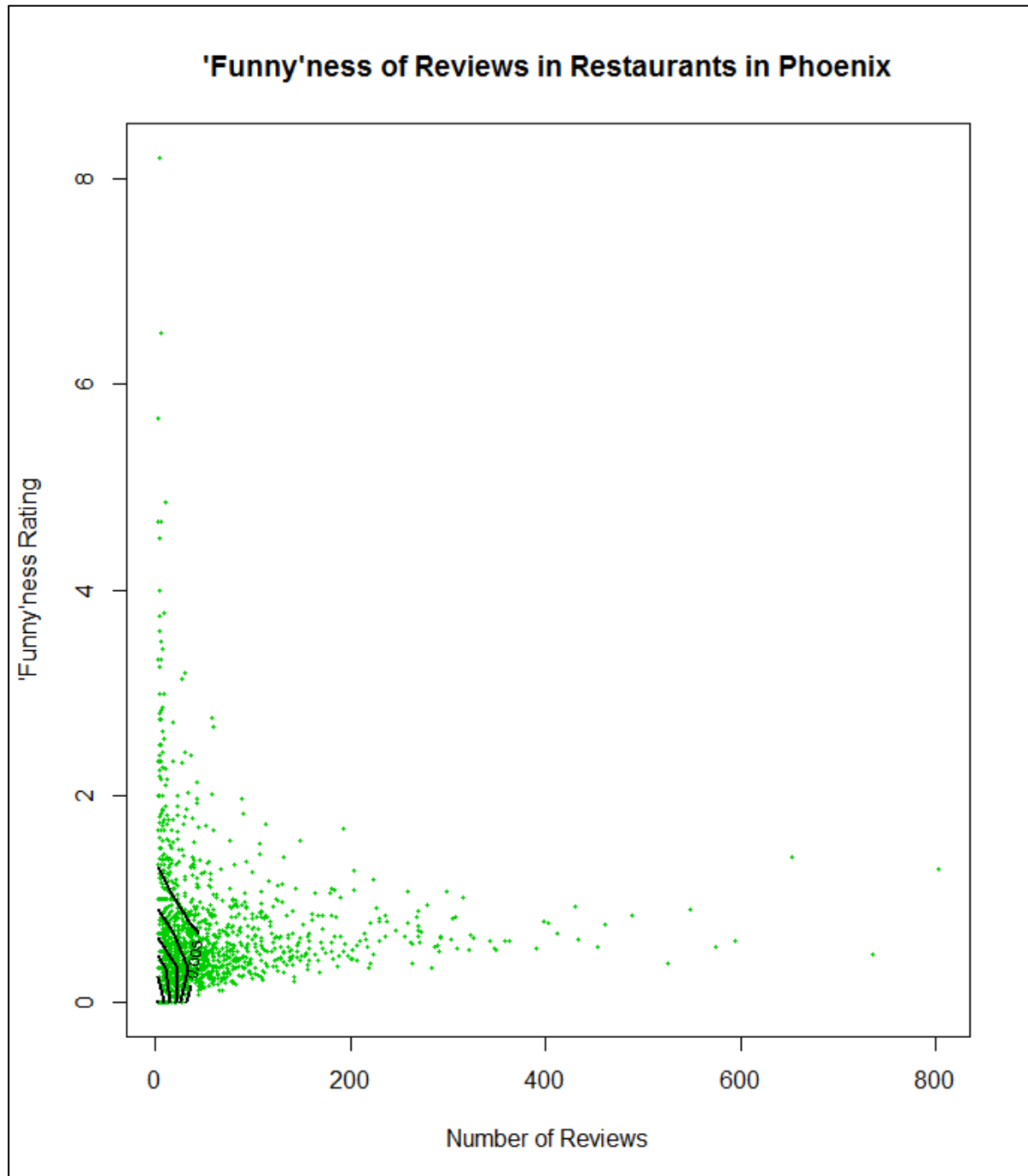


From the above plot, we can see that the highest density is between 0 and 80 reviews and the ratings are between 0 and 1.5.

It seems restaurants that have a greater number of reviews aren't really useful, while those with fewer reviews are more useful. This is to be expected since having less reviews would make each review of a restaurant more valuable.

```
reviews.funny = kde2d(review_count, funny)
plot(review_count, funny, pch=16, cex=0.4, col=3,
      main="'Funny'ness of Reviews in Restaurants in Phoenix",
      xlab="Number of Reviews", ylab="'Funny'ness Rating")
```

```
contour(reviews.funny, add=T)
```



The highest density for 'funny'ness is also between 0 and 80 reviews and the ratings are between 0 and 1.

It seems restaurants that have a greater number of reviews aren't as funny as those with fewer reviews. One guess as to why this might be is because other users rate 'funny'ness relative to the other reviews at a particular restaurant, and hence those with fewer reviews are more likely to be funny.

The distributions for usefulness and 'funny'ness vs. number of reviews are very similar; their highest density areas are concentrated in almost the same region and they both have a negative correlation with number of reviews.

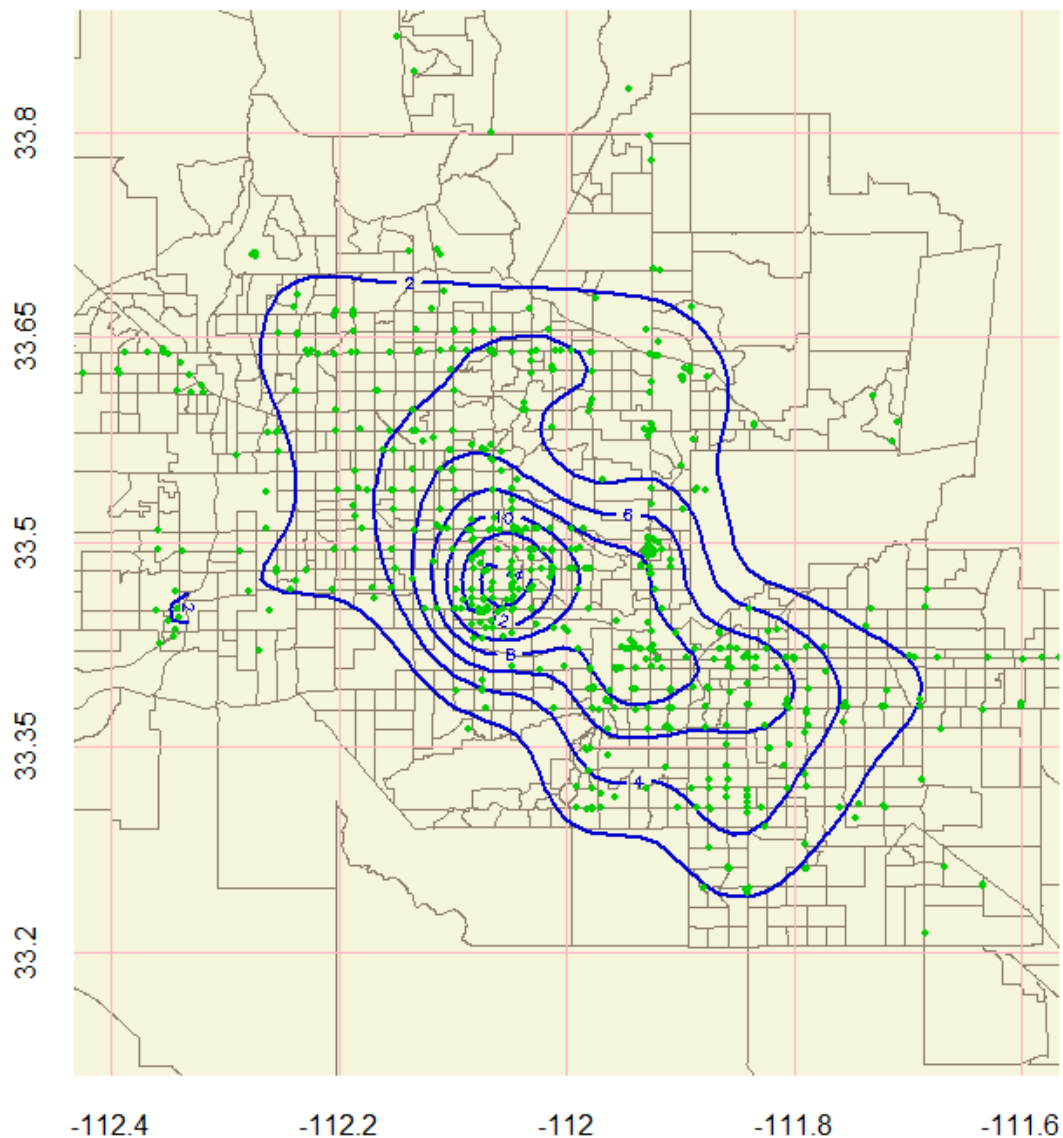
Question 3b

I plotted the kernel density estimate using a contour plot and overlaid it on the map of Phoenix. Again, a heat map could have been used, but it is clearer to use contour plots, when all we want to compare is the local modes. Since we want to compare the location of two cuisine types, it would seem better to plot both contours on the same map. However, I feel that it is not easy to distinguish the patterns when they intersect so much, hence I have plotted two separate maps.

```
mexican = kde2d(longitude[Mexican==1], latitude[Mexican==1], n=100)

plotMap("Density of Mexican Restaurants in Phoenix")
contour(mexican, add=T, col="blue3", lwd=2)
points(longitude[Mexican==1], latitude[Mexican==1], col="green3", cex=0.5,
pch=16)
```

Density of Mexican Restaurants in Phoenix

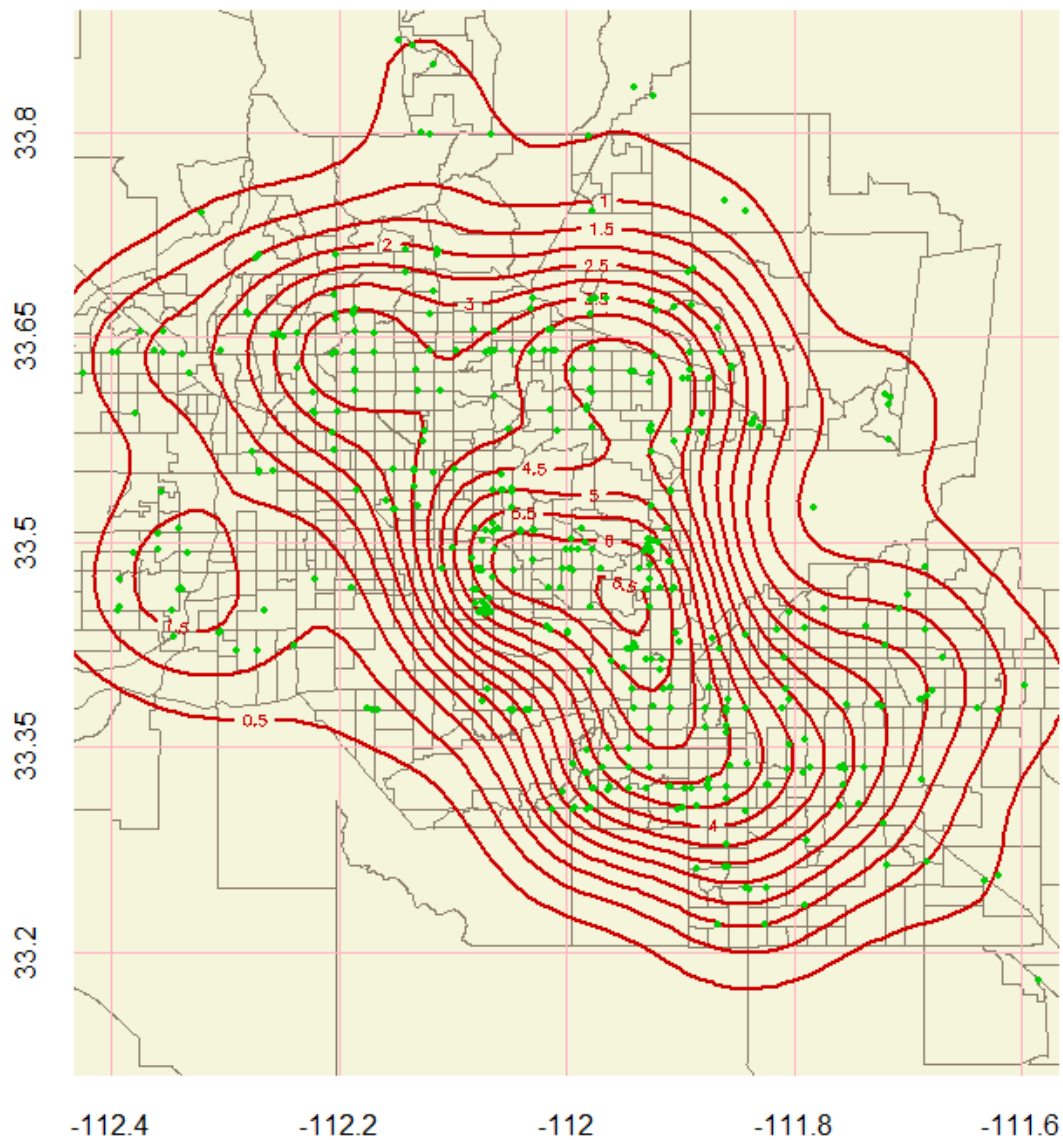


There are two clear modes: one at -112.3, and one at -112.07 longitude. Both are at about the same latitude, that is, 33.4. Another mode could be to the south-east of the mode that is at the center of Phoenix. A smaller mode also exists directly to the east of the central mode. So in total, there are 4 local modes for Mexican restaurants.

```
pizza = kde2d(longitude[Pizza==1], latitude[Pizza==1], n=200)

plotMap("Density of Pizza Restaurants in Phoenix")
contour(pizza, add=T, col="red3", lwd=2)
points(longitude[Pizza==1], latitude[Pizza==1], col="green3", cex=0.5,
pch=16)
```

Density of Pizza Restaurants in Phoenix



There are about four clear modes: one at -112.3, -112.2, and two at -111.95 longitude. Another mode could be to the west of the mode at -111.95 and 33.4. So in total, there are 5 local modes for Pizza restaurants.

The western-most modes for both Mexican and Pizza restaurants coincide at (-112.3, 33.4). While for the other modes, it can be seen that the 3 major Pizza modes surround the one major Mexican restaurant mode in central Phoenix. There is a minor mode at (-111.9, 33.5) that also coincides for both Mexican and Pizza restaurants.

Question 3c

I plotted the kernel density estimate using a contour plot and overlaid it on the map of Phoenix. In this case it would have been hard to use a heat map, since we are plotting two densities on one map.

Best is defined as 4.5 or 5 stars and worst is defined as 0 or 0.5 stars.

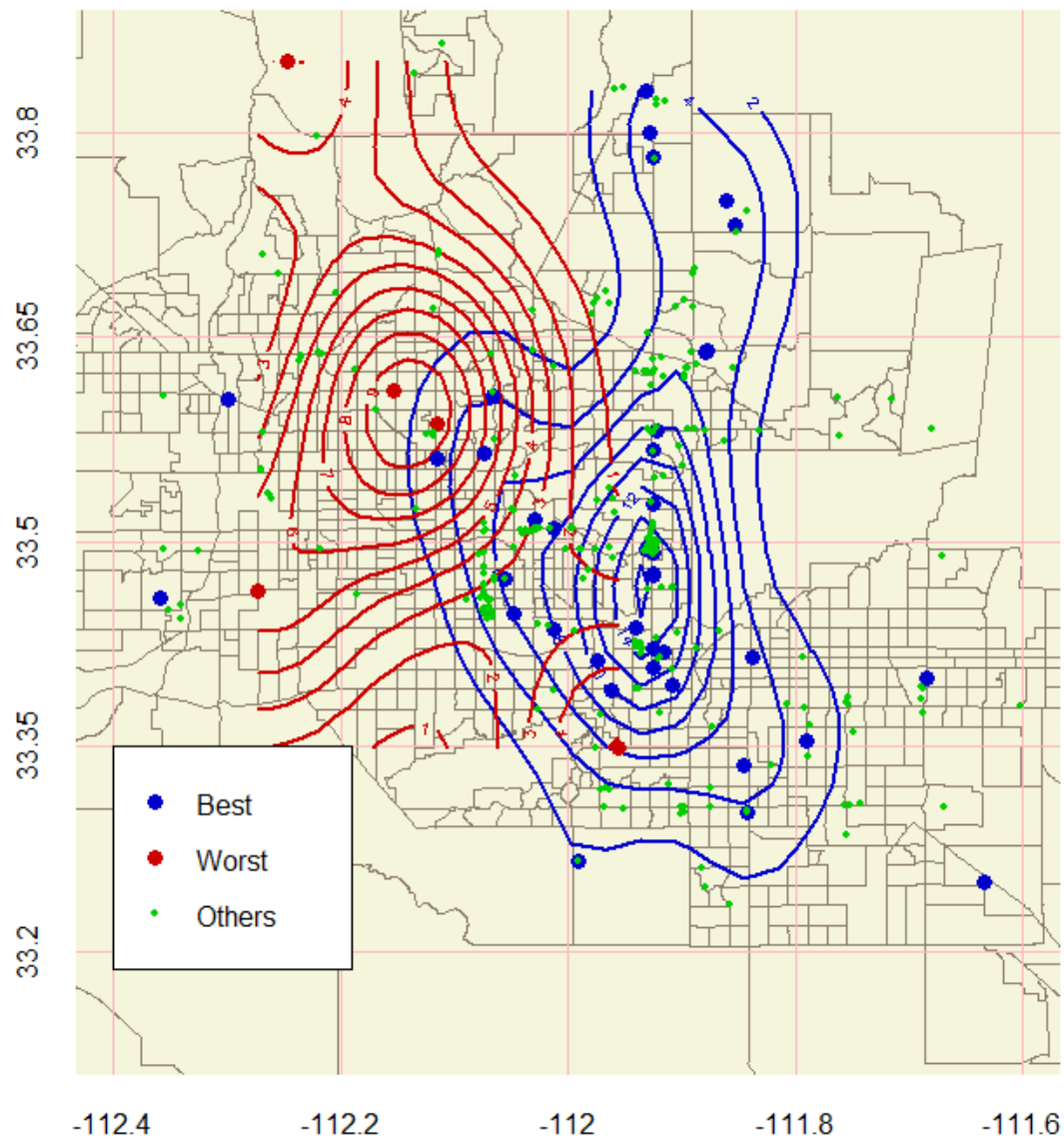
```
new.best = kde2d(longitude[rest[5]==1 & stars>4], latitude[rest[5]==1 &
stars>4])
new.worst = kde2d(longitude[rest[5]==1 & stars<2], latitude[rest[5]==1 &
stars<2])

plotMap("American (New) Restaurants in Phoenix\nby Rating (1 to 5 stars)")
contour(new.best, add=T, col="blue3", lwd=2)
contour(new.worst, add=T, col="red3", lwd=2)

points(longitude[rest[5]==1 & stars>4], latitude[rest[5]==1 & stars>4],
col="blue3", cex=1.2, pch=16)
points(longitude[rest[5]==1 & stars<2], latitude[rest[5]==1 & stars<2],
col="red3", cex=1.2, pch=16)
points(longitude[rest[5]==1 & stars>=2 & stars<=4], latitude[rest[5]==1 &
stars>=2 & stars<=4], col="green3", cex=0.5, pch=16)

legend(-112.4, 33.35, col=c("blue3", "red3", "green3"), c("Best", "Worst",
"Others"), pch=16, pt.cex=c(1.2, 1.2, 0.5))
```

American (New) Restaurants in Phoenix by Rating (1 to 5 stars)



New American restaurants seem to have a pretty clear east/west divide between the worst and best restaurants. The better ones are located mostly in eastern Phoenix, while the worst ones are in western Phoenix. However, there are some 'best' restaurants that are way out in the west, away from the other 'best' restaurants.

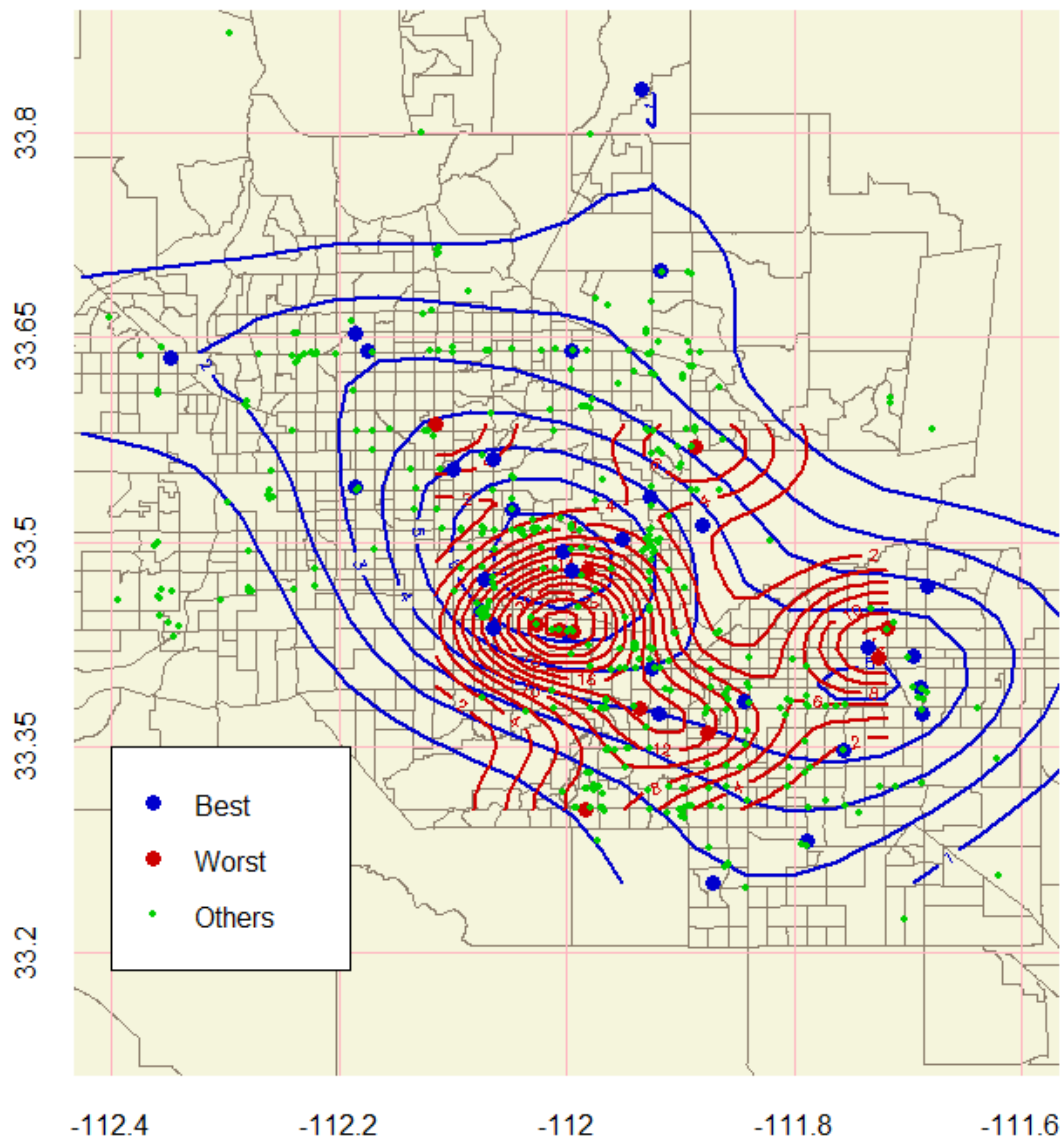

```
trad.best = kde2d(longitude[rest[6]==1 & stars>4], latitude[rest[6]==1 & stars>4])
trad.worst = kde2d(longitude[rest[6]==1 & stars<2], latitude[rest[6]==1 & stars<2])
```

```
plotMap("American (Traditional) Restaurants in Phoenix\nby Rating (1 to 5 stars)")
contour(trad.best, add=T, col="blue3", lwd=2)
contour(trad.worst, add=T, col="red3", lwd=2)
```

```
points(longitude[rest[6]==1 & stars>4], latitude[rest[6]==1 & stars>4], col="blue3", cex=1.2,
pch=16)
points(longitude[rest[6]==1 & stars<2], latitude[rest[6]==1 & stars<2], col="red3", cex=1.2,
pch=16)
points(longitude[rest[6]==1 & stars>=2 & stars<=4], latitude[rest[6]==1 & stars>=2 &
stars<=4], col="green3", cex=0.5, pch=16)
```

```
legend(-112.4, 33.35, col=c("blue3", "red3", "green3"), c("Best", "Worst", "Others"), pch=16,
pt.cex=c(1.2, 1.2, 0.5))
```

American (Traditional) Restaurants in Phoenix by Rating (1 to 5 stars)



Traditional American restaurants don't really seem to differ in location by best/worst. Both have around two modes and both modes seem to coincide. The best restaurants have a larger spread, however and can be found much further outside of Phoenix than the worst restaurants, which are mostly contained in central Phoenix.