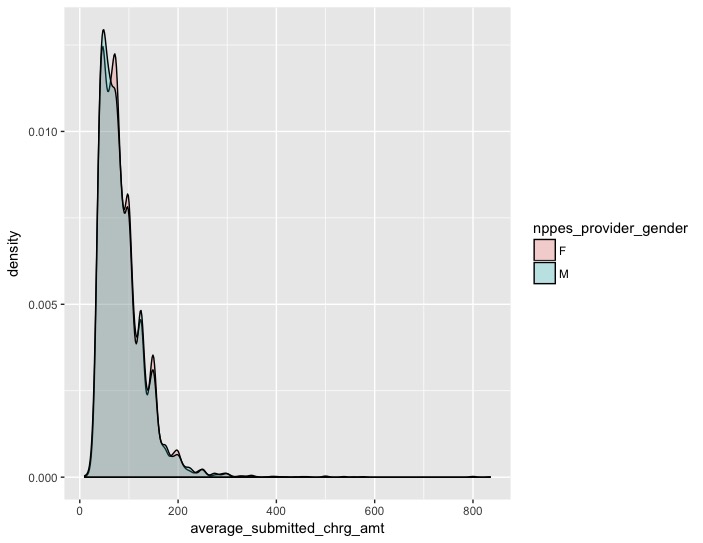
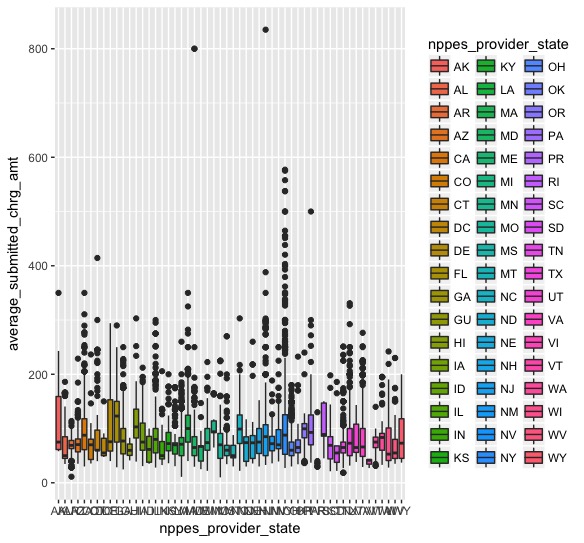
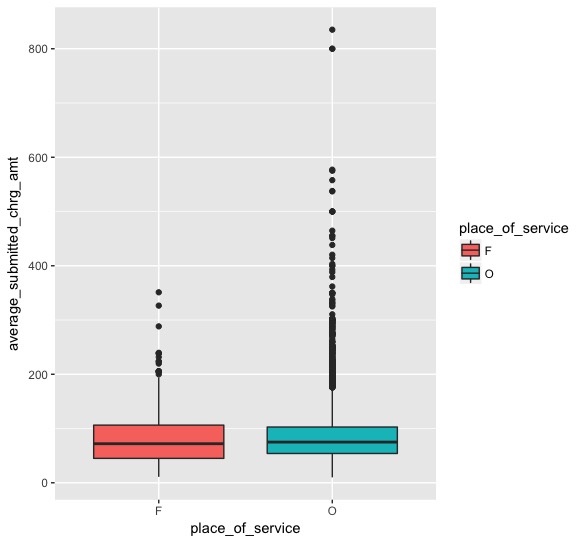
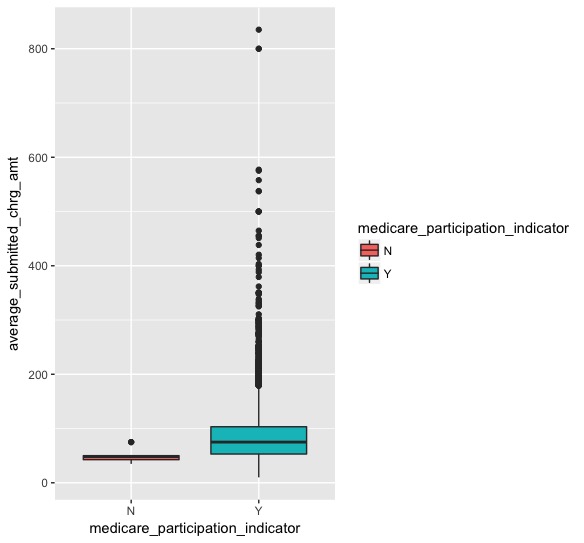
Average submitted charge amount from physicians can vary a lot from individual to individual even for the same procedure. Many factors contribute such differences including state/city expenses level, how experience physician is, if a physician participates in Medicare or not. Knowing the physicians charge patterns can help patients find a better fit based on their economic status and the severity of their diseases. It can also help policy makers understand where to put resource on the place needs the most. Here, we are particularly interested in how procedures in OBGYN (Obstetrics/Gynecology) are charged and if there any distinct patterns of charging. Specifically, we are interested in procedure - G0101 ‘Cervical or vaginal cancer screening; pelvic and clinical breast examination’, which is the top two frequently ordered procedure among OBGYN specialists. For the same procedure, we want to compare what lead to individual physician has different submitted charges. We will first start to explore how some common factors including gender, medicare\_participation\_indicator, place of service, state affects average\_submitted\_chrg\_amt. By histogram, box plot, and density plot as shown below, we can tell that participating in medicare or not and state affect the average submitted charge amount for procedure G0101 significant, while sex and place of service not so much.

|  |
| --- |
|  |



For the second step, we are interested to find if there are specific patterns/groups among physicians which leads to distinct average\_submitted\_chrg\_amts. Our features include Medicare participation indicator, gender, nppes\_provider\_state, line\_srvc\_cnt. We use algorithm ‘partitioning around medoids’ as our clustering method because it can deal with both categorical and continuous data as features. This method uses gower distance to represent distance between categorical vectors. We find that when the number of cluster is 6, the algorithm generates the best outcome.

We then generate the characteristics of those 6 clusters by analyzing the exemplars of each cluster. The result from R is shown as below. The result shows NY, M and physicians in Medicare network have the highest submitted\_charge\_amount. Those physicians have a about 78 line of service count per day. In the meantime, still in New York, female physicians in Medicare network, providing around 31 HCPCS G0101 services charge the submitted the least average charge amount. This, on one hand, reflect the gaps of submitted charges between male and female. It also on the other hand suggesting the huge difference of G0101 charges in New York city. A person who does not have enough money for healthcare can seek physicians with relative low service amount to save the cost.

nppes\_provider\_gender nppes\_provider\_state

703912 F NY

791147 F TX

163284 F PA

8820125 F CA

1995749 M NY

1126226 M TX

medicare\_participation\_indicator line\_srvc\_cnt

703912 Y 31

791147 Y 80

163284 Y 48

8820125 Y 17

1995749 Y 78

1126226 Y 25

average\_submitted\_chrg\_amt

703912 40.0000

791147 91.6875

163284 71.0000

8820125 100.0000

1995749 134.4533

1126226 86.2500