1. What is the question?

   To find the cheapest place to have "Cardio vascular stress test" in Virginia and also in different states of US.

2. Why is it important?

   It is important because it is necessary to get the cheapest place for "Cardio vascular stress test". This can be obtained by visualizing the average cost in the state of Virginia and also by analyzing how average cost varies across the United States in the form of heatmap without considering the practitioner.

3. What have others done to try and solve the question?

   I referred the following site to learn the pig. http://www.rohitmenon.com/index.php/apache-pig-tutorial-part-1/. Here they have used a Movie dataset to extract the required data according to the movie ratings. This really helped me in developing the pig script to extract the data with respect to different practitioners in Virginia and also with respect to different states in US regardless of practitioner.

4. What will I do to solve the problem?
a) Login to atria.cs.odu.edu
b) Secure shell to the Hadoop server.
c) Make sure these files are available undergrad.pig, graduate.pig, makeFile.sh
d) Run the makeFile.sh bash script. It will ask user to extract data for undergrad or gradpart. Further it will ask user to give directory name where output files need to be generated.
e) Once the output file is generated for both undergrad and grad part, I get the output file of grad part to generate heatmap to visualize the average cost across the United States using Tableau.

5. What will I do to prove that I have solved the problem?

   I will find the average cost for "Cardio vascular stress test" in Virginia when it is given by different practitioners. Further, I will find the average cost for "Cardio vascular stress test" across the United States regardless of practitioner and visualize or analyze the data in the form of heatmap using Tableau.

6. What is the conclusion?

   I'm able to find the average cost for "Cardio vascular stress test" in Virginia. Further I was able to generate average cost across the United States and visualized these average cost in the form of heatmap. With this, we were able to know the cheapest place to have "Cardio vascular stress test" get it done in Virginia and also other places in US.