Data Source: https://www.transtats.bts.gov/DL_SelectFields.asp?
Table_ID=236&DB_Short_Name=On-Time

According to a 2010 report made by the US Federal Aviation Administration, the economic price of domestic flight delays entails a yearly cost of 32.9 billion dollars to passengers, airlines and other parts of the economy. More than half of that amount comes from the pockets of passengers who not only lose time waiting for their planes to leave, but they also miss connecting flights, spend money on food and have to sleep on hotel rooms while they're stranded. In this report, I analyzed the 1.936.758 different internal flights in the US for 2008 and examined more for each carriers and airports in detail. Thus I provided a big picture of flights delay condition in U.S. domestic market.

In order to have a geographical overview of the flight condition in U.S., I used Basemap toolkit in Matplotlib. Basemap toolkit is one of several Matplotlib toolkits which lives under the `mpl_toolkits` namespace. I choose this graphical form because Basemap is a useful tool for Python users to have in their virtual toolbelts. The useful thing is that the globe shown here is not a mere image, it is a fully-functioning Matplotlib axes that understands spherical coordinates and which allows us to easily overplot data on the map.

Firstly, I used Basemap to plot the airport location and the number of flights per year in the U.S. Secondly, I explored the delay condition of each airport in detail. Basically, we care more about the departure delay for the following two reasons. First, costumers usually arrive in the departure

airport before the estimated departure time. Thus they have to wait in the airport if the flight is delay in departure. Second, if carriers are allowed higher speed on planes that departed late, the arrival delay could be decrease significantly. Moreover, I only consider flights that delays more than 15 minutes since short delay minutes are acceptable. In other words, I used departure delay time that over 15 minuted to calculate the delay percentage for each carrier and airport. I plotted the airport location and the delay percentage for departed planes in this airport in the U.S. map. Thirdly, I used air chart to demonstrat the percentage of flights per airline and per departure airport.

In terms of channels to encode the data, Basemap provided elegant and meticulous color and scale internally. I draw a high-resolution land-sea mask as an image, with land and ocean colors specified. The land-sea mask is derived from the GSHHS coastline data. This made my map looks vivid and representative. Moreover, I carefully choose the labels and scale for bar charts when comparing flights among airports and airlines. I extracted a subset of columns and redefined the airlines labeling. I also set the legend as "abreviation -> airline name" in order to make clear description for audiences. I redefine the colors for correspondence with the pie charts. In this way, main information was conveyed to audience without reading too much write-ups.

Overall, I think my data visualization satisfied truthfulness, functionality, beauty and insight. It facilitated effective communication. It is truthful since it presents the flights delay condition based on the data I used. It is functional since I have answered questions I planned to explore. For

example, I found that the busiest airports serving over 100 thousands flight per year mainly concentrate in the east and west coasts in the United States. Then I found that Dillingham Airport, King Salmon Airport and Pueblo Memorial Airporta have the highest delay percentage that over 40%.There is some disparity between the carriers. And these graphs present elegant and well designed in terms of the use of typography, color, and composition. As far as the insight, my graphs provide an insightful version for audiences to comprehend information quickly. By using Basemap, audiences could see the geographical overview of the flight condition in U.S. quickly and easily. However, I am not sure if it is enlightening. Even though I have presented the basic information of domestic flights condition, I am not sure if I made readers more informed one owing to space constraints. I think expanding more information could help my analysis more enlightening.