Project 5.3. Orbitz web scraping

Requirement: submit a code to realize the following function:

1. Enter the following information in the code:
   1. Choose Flight Only
   2. City flying from
   3. City flying to
   4. Choose one-way flight on 1/7/2018
   5. Choose 1 adult
2. Return a csv file with the following information:
   1. Cities flying from and flying to
   2. Airports flying from and flying to
   3. Airport codes flying from and flying to
   4. Flight company and flight number
   5. Aircraft type
   6. Taking-off time, arriving time, duration
   7. If connecting flight, all above information of every individual flight, also include connecting airport code
   8. Ticket price
3. If entering a list of [city flying from, city flying to], the code returns multiple csv files.
4. [Bonus 1] Make a comprehensive file of all flights between all cities in the world
5. [Bonus 2] Picture a flight network using Gephi based on the information you obtain from Bonus 1

Codes will be graded based on:

1. The format of the codes
2. The organization of the csv file
3. The speed of the codes
4. The extendibility of the codes
5. The results of Bonus 1
6. The network figure of Bonus 2

Estimated completion time: 60 hours + [Bonus 1] 600 hours + [Bonus 2] 10 hours. Time consumption depends on the codes’ quality.

Hint: for Bonus 1, you may be able to find all flights’ information in other websites, try your best to search.

Suggestive points:

Deadline:

Bonus points:

1. Work completion before the deadline will generate an extra 20% bonus.

Penalty:

1. Late delivery/missing deadline will incur -10% of finally determined points of the whole project per missing day.
2. Serious mistakes that made the submitted results not useful will incur -120% of the whole project.